Journal of Applied Computational Intelligence

eISSN: 1234-5678

Vol. 4, No. 1, January, 2026

TD3-MPC Model for Mineral Dust Concentration Estimation Based on Hot-Air Flow and Particle Sensors

Ivan Kovačević[1,*], Mirko Novaković[1]

[1] *Faculty of Technical Sciences, University of Novi Sad, Novi Sad, 21000, Serbia*

Received 24 October 2025; Revised 06 January 2026; Accepted 12 January 2026; Published 26 January 2026

Abstract

Real-time estimation of mineral dust concentration in thermally disturbed mine roadways remains difficult because hot-air ventilation changes particle trajectories, delays emission-measurement coupling, and slows optical sensor recovery. This paper develops a twin delayed deep deterministic policy gradient and a model predictive control model for concentration estimation based on coupled hot-air flow and particle-sensor streams. The method introduces delay-compensated feature reconstruction, a physics-guided state transition layer, twin critic learning and receding-horizon correction. A 42-hour laboratory and pilot-gallery dataset was organised into 151,200 synchronous samples of thermal anemometers, particle counters, humidity probes and gravimetric reference instruments. Compared with support vector regression, long short-term memory, temporal convolution and soft actor-critic baselines, the proposed model reduced the mean absolute error under rapidly varying hot-air velocity from 2.91-4.76 mg/m³ to 1.68 mg/m³ and improved R^2 from 0.914 to 0.972. During high-load cutting episodes above 80 mg/m³, the 95th percentile error was still 4.9 mg/m³, and the best learning baseline reached 7.8 mg/m³. Based on the above results, the two can work together to reduce thermal-flow drift and transient sensor lag, especially when dust plumes evolve faster than conventional particle sensors can recover. Therefore, the above estimator is suitable for warning-oriented monitoring, calibration transfer and ventilation coordination in underground mining areas with considerable thermal disturbance.

Keywords: Mineral Dust Estimation, Hot-Air Flow Sensing, Particle Sensor Fusion, TD3-MPC Control, Underground Mine Monitoring

1. Introduction

Mineral dust is a prolonged danger in underground mining because its concentration is affected by changes in cutting load, ventilation direction, roadway geometry, and the thermal plume generated by machinery and heated return air [1]. Respirable fractions are even harder to observe in real time because the same emission source will produce different near-sensor readings after convection and turbulence alter the residence time of particles [2]. The previous monitoring method is still a trade-off between gravimetric accuracy and the speed of online response, and this trade-off is further worsened when dust clouds pass through the hot-air stream before reaching an optical or electrical sensing chamber [3]. Although the density of environmental sensors in the recent mine digitalisation programmes has increased, the raw data still suffers from drift, non-uniform particle size distribution and humidity-induced scattering bias [4]. In the production head, the airflow used to dilute dust also carries heat generated by cutting motors and auxiliary ventilation ducts, thus forming a coupled thermal-flow field rather than an independent ventilation variable [5]. Therefore, a reliable estimator should be based on both particle count and thermal airflow data, and have good dynamic sensitivity for warning and control.

* Corresponding author: ivan.kov@tf.uns.ac.rs

https://doi.org/10.68406/jaci.2026.vol4iss1no4:42-54

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

Data-driven estimation has become popular because of the nonlinear interactions in mineral dust formation and transport that cannot be fully described by closed-form ventilation equations. Regression models can describe a relatively narrow concentration range when the sensor layout is fixed, but their errors often increase during a short burst caused by roof drilling, scraper movement and sudden fan regulation [6]. Recurrent neural networks improve temporal representation by learning delayed responses from past observations, but they may smooth out peaks that are operationally significant for exposure assessment [7]. Reinforcement Learning has also been applied to the optimisation of industrial processes, and an agent can learn corrective measures based on delayed rewards and noisy process data [8]. At the same time, model predictive control can be used to constrain the estimates and prevent unreasonable fluctuations within a known physical limit [9]. Currently, there is a disconnect between estimation and control: we do not directly drive the ventilation system but rather infer concentration via a controller-like correction layer that considers airflow, thermal inertia and particle-sensor lag [10].

To address this need, this paper integrates twin delayed deep deterministic policy gradient with a model predictive correction scheme, hereafter referred to as TD3-MPC. The hot-air flow signal is not treated as a simple auxiliary covariate. Instead, it is used to shape the latent state transition, estimate delay compensation, and constrain the feasible concentration trajectory. Particle-sensor signals are fused across number concentration, optical mass estimate, and recovery gradient, so that rapid emissions and slow settling phases are represented by different state features. The twin critic structure is used to reduce overestimation in the learned correction policy, while the predictive layer regularizes concentration evolution across a short horizon. This design is intended for monitoring systems where reference measurements are available during calibration but only sensor streams are available during routine operation.

To make the technical route clear from sensing conditions to estimator validation, the remainder of the paper is organized into six connected sections. Section 2 reviews mineral dust sensing and learning-based estimation methods, emphasizing the limitations caused by thermal airflow coupling and transient sensor delay. Section 3 describes the sensing architecture, data synchronization procedure, and mathematical formulation of the concentration estimation problem. Section 4 presents the TD3-MPC model, including the delayed policy update, twin-critic objective, and predictive correction mechanism. Section 5 reports the experimental dataset, benchmark protocol, numerical results, and sensitivity analysis with two tables and four multi-panel data figures. Section 6 concludes the paper by summarizing the contribution, identifying current limitations, and outlining future work.

2. Related Work

2.1. Sensor-Based Mineral Dust Monitoring

Wei et al. [11] showed that ventilation parameters can reshape the roadway flow field and dust migration, which supports the recent shift from single-point sampling to sensor arrays combining optical particle counters, airflow meters, temperature probes, and humidity compensation channels. Optical devices are attractive because they provide high-frequency readings, but calibration can drift when mineral composition, particle shape, or moisture content changes over a shift. Gravimetric sampling remains the reference method for mass concentration, yet it cannot provide second-level concentration dynamics required by intelligent ventilation and exposure alarms [12]. Researchers have therefore explored hybrid calibration strategies, in which periodic reference measurements are used to adjust online sensors and reduce bias under changing aerosol properties [13]. The challenge is amplified in hot-air roadways because buoyancy and forced ventilation jointly alter the probability that particles with different diameters enter the sensing volume. A sensor that is accurate in a stable laboratory chamber can underestimate peak concentration near a heat source, while overestimating residual concentration during cooling and deposition [14].

The literature also shows that sensor placement strongly affects estimation quality. A particle counter installed near the return-air side captures integrated emissions but responds late to localized cutting events, whereas a near-source sensor is sensitive to mechanical vibration and large particles outside the respirable fraction. Thermal anemometers provide useful flow velocity and temperature-gradient information, but they do not directly measure dust loading. For this reason, sensor fusion is not merely a matter of averaging redundant signals; it must represent the transport relation between dust release, hot-air advection, turbulent mixing, and instrument response. Existing fusion schemes often assume a quasi-steady relationship between airflow and concentration, leaving a gap for dynamic estimators that can handle high-frequency thermal disturbances.

2.2. Learning and Predictive Control for Industrial Estimation

Learning-based estimators have been used in environmental monitoring, mineral processing, and tunnel ventilation because they can approximate nonlinear mappings without requiring a fully specified physical model. Support vector regression and ensemble trees are robust for small datasets, but their static feature windows can miss asymmetric rise and decay behavior in dust concentration [15]. Long short-term memory networks and temporal convolutional networks improve this limitation by encoding historical sequences, and they are suitable when the delay pattern is relatively stable

[16]. However, mine airflow can change abruptly when auxiliary fans switch mode or mobile equipment blocks part of the section, which means the delay itself becomes state-dependent. Under this condition, a purely supervised sequence model may learn a smoothed average response rather than the correction needed at the current operating point.

Reinforcement learning provides a different route by learning a policy that adjusts an estimated state according to delayed performance feedback. Actor-critic structures are particularly relevant for continuous industrial signals, and twin delayed critics reduce optimistic value estimates that otherwise destabilize control or correction policies [17]. Model predictive control is widely used when constraints, horizons, and physically plausible state evolution must be considered simultaneously [18]. Combining reinforcement learning with predictive control can therefore be useful for mine dust estimation: the learned component captures nonlinear sensor-flow relations, while the predictive component keeps the trajectory consistent with transport limits and calibration knowledge. Few studies have examined this combination for mineral dust estimation under hot-air flow, especially with synchronized particle counts, thermal velocity, and reference mass concentration across both steady and transient operating regimes.

3. Sensing Architecture and Problem Formulation

3.1. Hot-Air and Particle Sensing Arrangement

The monitoring architecture is designed for a mine roadway section in which hot-air flow and dust release are spatially coupled. The sensing node includes a thermal anemometer, a temperature-gradient probe, a humidity sensor, an optical particle counter, and a low-frequency gravimetric reference channel used during calibration. The particle counter reports number concentration in three size bins and an embedded optical mass estimate, while the airflow module records velocity and air temperature at the same sampling clock. A synchronization board assigns timestamps before wireless transmission, because post hoc alignment is unreliable when dust peaks have steep leading edges. Figure 1 is placed here as the structural flow diagram of the acquisition and estimation chain, showing the passage from mine disturbance, through coupled hot-air transport, to particle sensing and TD3-MPC inference.

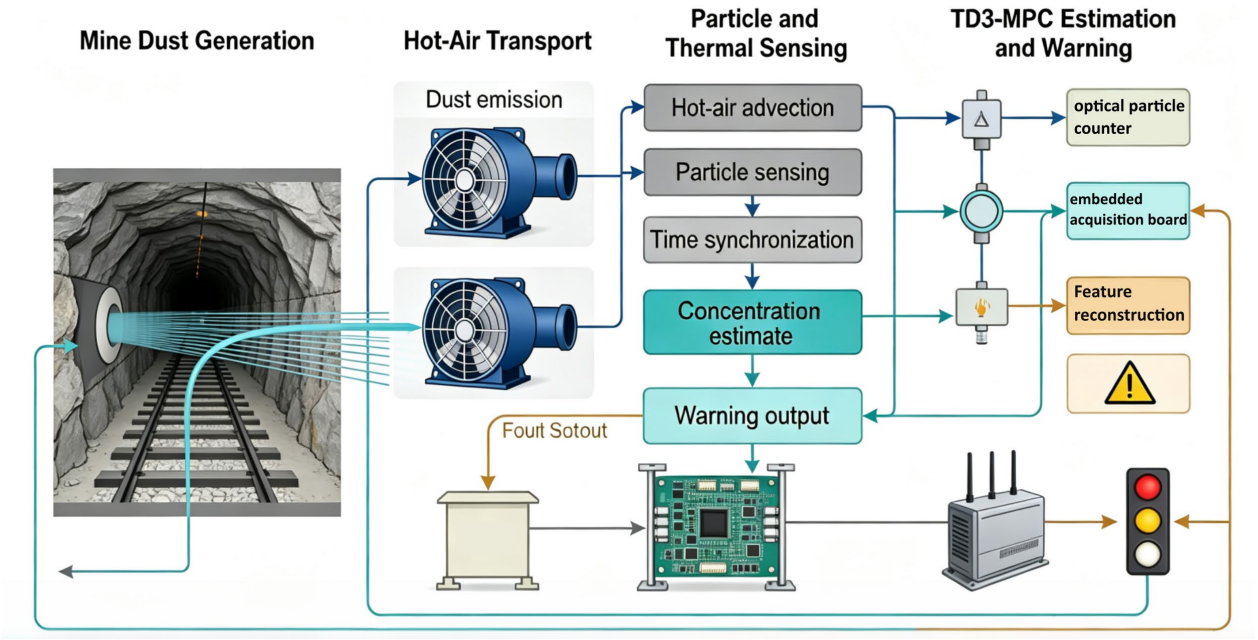


Figure 1. Placeholder for the hot-air-flow and particle-sensor monitoring architecture

The useful state information extends beyond the instantaneous particle reading. During a cutting burst, concentration may rise before the thermal velocity reaches a new equilibrium, while in the decay phase the particle counter may remain elevated after the hot-air jet weakens. The input vector is therefore constructed from synchronized readings, short-window gradients, and delay-compensated lags. The reference output is the calibrated mineral dust mass concentration in $\mathrm{mg/m^3}$. The architecture supports both laboratory duct experiments and pilot-gallery deployment because no component requires direct access to the dust source. This arrangement also allows the estimator to operate when reference samplers are absent, provided that model calibration has already been completed.

The physical layout is also considered when defining the model input. The hot-air probe was positioned upstream of the particle inlet by $0.35m$, so the measured velocity represented the carrier flow before the dust cloud entered the optical path. A second temperature point was placed near the particle inlet to capture the local thermal gradient. This arrangement gives the estimator an early indication of transport acceleration while preserving a direct measurement of

the thermal condition at the sensing chamber. In practice, such a layout is easier to maintain than distributed multi-point dust sampling, because the airflow and temperature probes can be cleaned without opening the particle counter enclosure. The sampling unit therefore behaves as a compact sensing station rather than a laboratory-only measurement rig.

3.2. State Representation and Delay Compensation

Let the synchronized sensor vector at time step t be represented as x_t, where its elements include hot-air velocity, air temperature, humidity, particle-bin counts, optical mass, and local gradients. The estimator maps x_t and a finite history window into the latent concentration state z_t. A compact transition representation is adopted so that the model can learn fast emission changes without allowing physically impossible discontinuities in the estimated mass concentration.

$$x_t = \left[v_t, T_t, H_t, n_{1,t}, n_{2,t}, n_{3,t}, m_t^o, \Delta v_t, \Delta T_t, \Delta m_t^o\right] \tag{1}$$

τ_tThe delay-compensated feature vector includes a transport lag inferred from hot-air velocity and roadway distance. Instead of fixing this lag across all conditions, the model estimates it within a bounded calibration interval. This dynamic lag is important because warm turbulent flow shortens the rise delay under forced ventilation but may extend the recovery delay when particles recirculate near the sensor inlet.

$$\phi_t = concat\left(x_t, x_{t-\tau_t}, mean(x_{t-k:t}), std(x_{t-k:t})\right) \tag{2}$$

The representation also keeps the raw and delayed observations in the same feature vector, rather than replacing one with the other. This design is useful when the roadway alternates between forced advection and local recirculation. In a forced-advection interval, the delayed particle signal may align better with reference concentration; in a recirculation interval, the current optical mass and recovery gradient may contain more information about the remaining suspended dust. The standard-deviation term captures short-window instability and helps the correction policy distinguish a real emission burst from a single noisy particle-count sample. All features are normalized by calibration-set statistics, while the estimated delay is clipped to the physically observed interval so that the model cannot create an unrealistically long memory during deployment.

3.3. Estimation Objective

The target variable y_t is the reference mineral dust concentration, and the predicted value is denoted by $\hat{y}_t$. The estimation problem is formulated as a constrained sequential correction task. The learning module predicts a correction action for the current latent concentration state, while the predictive layer checks whether the corrected state remains consistent with observed hot-air transport and previous concentration evolution. The loss used for supervised calibration combines point error and temporal smoothness, but the reinforcement component further evaluates whether a correction reduces future horizon error.

The objective is deliberately expressed at the estimation level rather than as a ventilation-control target. This distinction matters because the model is expected to operate in mines where the ventilation controller, alarm platform, and exposure-management system may be maintained by different subsystems. The estimator must therefore provide a stable concentration trajectory with interpretable uncertainty behavior, while avoiding direct assumptions about fan actuation or personnel evacuation rules. The output can be consumed by an alarm module or by a separate control layer, but the TD3-MPC estimator itself is trained only to recover the best concentration sequence from available sensor streams.

For this reason, the objective is evaluated by both sample-level error and event-level behavior. A low average error is insufficient when the estimator misses the leading edge of a dust burst or overstates the decay tail after ventilation changes. The training target therefore treats rise, peak, stable, and recovery segments as distinct operating conditions during replay sampling. This segmentation does not introduce separate models; it only prevents the estimator from being optimized mainly for the long background intervals that dominate ordinary monitoring records.

4. TD3-MPC Estimation Method

4.1. Twin Delayed Policy Learning

The TD3 component defines concentration correction as a continuous bounded action. The actor receives the latent state and proposes a bounded correction, while two critics evaluate the expected reduction of estimation error across delayed sensor response. The smaller critic value is used when updating the target, which reduces optimistic correction under noisy peaks. Figure 2 is positioned here as the algorithmic workflow, linking feature reconstruction, actor-critic learning, predictive correction, and final concentration output.

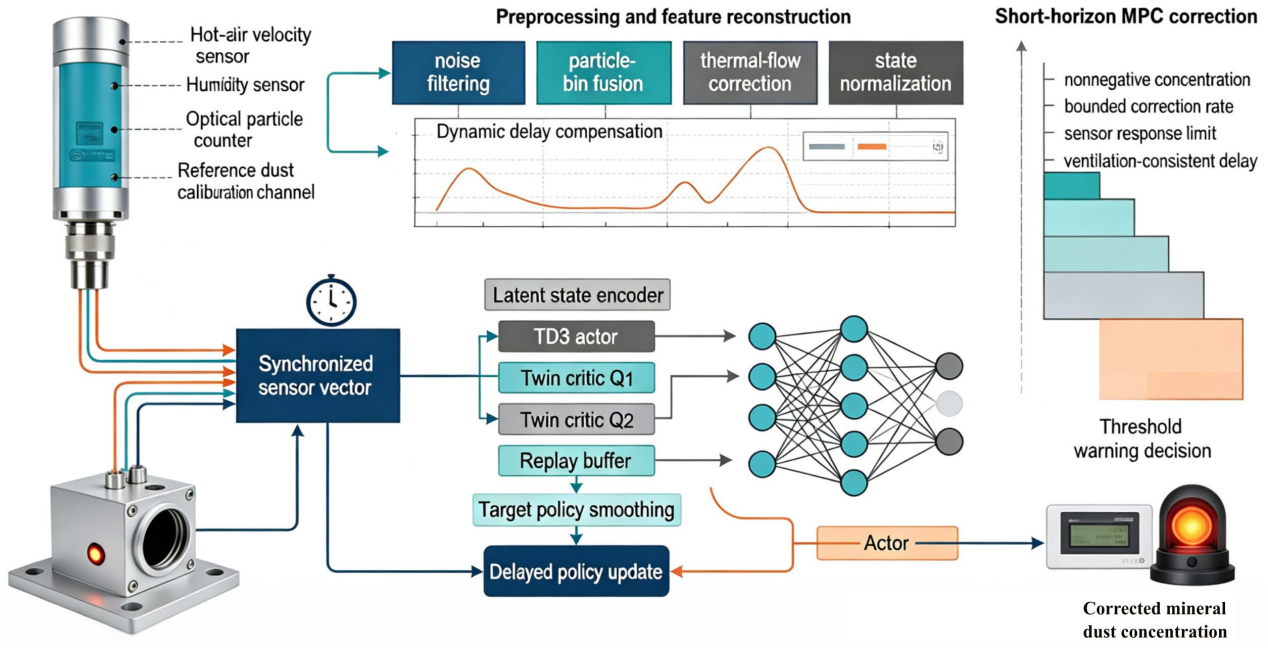


Figure 2. Placeholder for the TD3-MPC estimation workflow and delayed policy update process

$$a_t = \mu_\theta(z_t) + \epsilon_t, \epsilon_t \sim clip(N(0, \sigma^2), -c, c) \tag{3}$$

The delayed policy update is performed after multiple critic updates, so the actor is not forced to chase short-lived noise in the particle stream. This is useful for dust estimation because sharp peaks and sensor recovery tails can both generate large instantaneous errors. The reward is defined as the negative horizon-weighted estimation error with an additional penalty for excessive correction magnitude.

$$r_t = -|y_t - \hat{y}_t| - \lambda_a|a_t| - \lambda_s|\hat{y}_t - \hat{y}_{t-1}| \tag{4}$$

The actor output is interpreted as a correction to the encoder estimate rather than as a direct concentration prediction. This separation keeps the supervised encoder responsible for the main sensor-to-concentration mapping and reserves the policy network for difficult operating points where delay, thermal drift, or recovery memory makes the initial estimate unreliable. The twin critics receive the corrected state, recent error trend, and airflow-gradient features. When one critic assigns a high value to an aggressive correction caused by a short optical spike, the other critic often produces a more conservative estimate; the minimum target then prevents the policy from amplifying noise. The delayed actor update further stabilizes training because the policy is changed only after the critic pair has absorbed several batches of transient examples.

4.2. Predictive Correction Layer

The MPC layer receives the candidate estimate from the actor and evaluates a short future trajectory under the learned transition. The horizon is intentionally short because the purpose is not long-range forecasting, but correction of the next few seconds in which sensor delay and thermal advection dominate. Constraints are imposed on concentration rate of change, non-negativity, and the feasible response range under the measured hot-air velocity.

$$\hat{y}_t^* = arg \min_{\hat{y}_{t.t+N}} \sum_{i=0}^{N} q_i(y_{t+i} - \hat{y}_{t+i})^2 + \rho_i(\hat{y}_{t+i} - \hat{y}_{t+i-1})^2 \tag{5}$$

The predictive layer is deliberately lightweight. It does not solve a full ventilation field or require geometric reconstruction of the roadway; instead, it checks whether the actor-corrected concentration trajectory is compatible with the current transport regime. When hot-air velocity is high and temperature gradient is positive, a faster rise is allowed because the dust plume can reach the sensor more rapidly. When velocity drops or the thermal gradient reverses, the admissible correction range becomes narrower, preventing the model from following isolated particle-count spikes. The horizon weights give larger importance to the first few seconds, since those estimates determine alarm timing and exposure warning. The smoothness term is not intended to suppress true peaks; it penalizes only changes that conflict with the measured airflow and recent concentration slope.

During deployment, the reference value inside the predictive objective is replaced by the encoder forecast and the latest corrected estimate, so the layer remains usable without new gravimetric labels. This implementation makes the

MPC block a consistency filter rather than a supervised optimizer. The corrected output is accepted when it reduces horizon inconsistency and stays within the calibrated dust-response envelope. If the proposed correction violates the envelope, the layer returns the nearest feasible trajectory, which limits unstable extrapolation during rare combinations of high temperature, high humidity, and abrupt velocity change.

4.3. Training and Deployment Procedure

Training is conducted in two stages. The first stage fits the feature encoder and transition layer using calibrated sensor-reference pairs, so that the latent state already carries interpretable concentration dynamics. The second stage trains the TD3 correction policy with replayed sequences sampled from steady, rising, peak, and recovery segments. During deployment, the reference channel is removed and the model updates estimates only from hot-air and particle-sensor streams. The predictive layer remains active in deployment because it uses measured airflow and learned constraints rather than new reference labels.

Replay sampling is balanced by event phase to prevent the model from being dominated by the long background periods that occupy most mine monitoring records. Each mini-batch contains stable, rise, peak, and decay windows whenever the replay buffer has enough examples of each phase. Target policy smoothing is applied only to the correction action, not to the raw sensor state, because smoothing the measured particle series would suppress physically meaningful bursts. The critic networks receive both the corrected estimate and the recent airflow gradient, enabling them to distinguish a true concentration rise from a sensor spike that is inconsistent with transport conditions.

The deployment procedure keeps the same normalization constants and delay bounds obtained during calibration. At the beginning of a shift, the sensor station performs a short background check to confirm that optical mass, particle-bin counts, velocity, temperature, and humidity lie inside the trained operating envelope. Online inference then proceeds sample by sample at $1Hz$. The estimator stores recent windows for diagnostic review but does not update network weights unless a supervised recalibration session is scheduled. This separation between inference and recalibration reduces the risk of silently adapting to sensor contamination or abnormal ventilation states.

5. Experimental Data and Analysis

5.1. Dataset and Benchmark Protocol

The dataset was collected from a $12m$ heated ventilation duct and a pilot mine-gallery section that reproduced dust release from drilling, cutting, and haulage disturbance. The final dataset contained 151,200 synchronized samples at $1Hz$, including $42h$ of recordings and 37 controlled emission episodes. Mineral dust concentration ranged from 4.7 to $126.4\mathrm{mg/m^3}$, hot-air velocity ranged from 0.32 to $3.18\ \mathrm{m/s}$, and air temperature ranged from 24.1 to $43.8°\mathrm{C}$. The calibration reference was updated by paired gravimetric measurements every 10 min and linearly aligned to the online stream after removing periods with instrument maintenance flags, following the common practice of pairing online aerosol readings with periodic mass-reference checks [19]. A 60/20/20 split was used by episode rather than by random sample, preventing leakage of the same dust event into both training and testing. Baselines included support vector regression, random forest, long short-term memory, temporal convolutional network, and soft actor-critic correction. Each model used the same sensor channels, although static baselines received handcrafted delay and gradient features.

Figure 3(a) compares the paired evolution of reference concentration and hot-air velocity during representative emission episodes, where the first dust peak reached $63.8\mathrm{mg/m^3}$ as velocity increased from 0.92 to $1.86\ \mathrm{m/s}$. Figure 3(b) reports particle-bin response under three thermal regimes; the fine-particle bin below 2.5μ m increased from 18,400 particles/L at background to 91,600 particles/L during cutting bursts, while coarse particles showed a slower decay tail. Figure 3(c) gives the empirical delay distribution from 0.8 to $6.4s$, with the high-temperature regime producing a median delay of $2.1s$ and the low-temperature regime reaching $4.7s$. Figure 3(d) shows the sensor-correlation matrix, where optical mass had the strongest positive relation with reference concentration at 0.93, followed by the fine-particle channel at 0.88. Reports on confined aerosol sampling indicate that thermal buoyancy can change the probability that particles enter an optical sensing volume [20].

Before model training, the online optical mass channel was checked against the reference channel within each 10 min alignment interval. The average calibration slope was 0.94 in the heated duct and 0.89 in the pilot gallery, with intercepts of $1.7\mathrm{mg/m^3}$ and $2.4\mathrm{mg/m^3}$, respectively. The lower pilot-gallery slope was expected because turbulent mixing and partial wall deposition reduced the fraction of large particles entering the optical chamber. A Hampel filter removed 0.43% of samples with obvious inlet blockage or packet loss, and no dust event was discarded entirely. This choice kept difficult transient sections in the benchmark instead of producing a cleaner but less representative dataset. The final training sequences used a $24s$ history window, which covered more than 95% of observed transport-delay values while keeping the feature tensor compact enough for embedded inference.

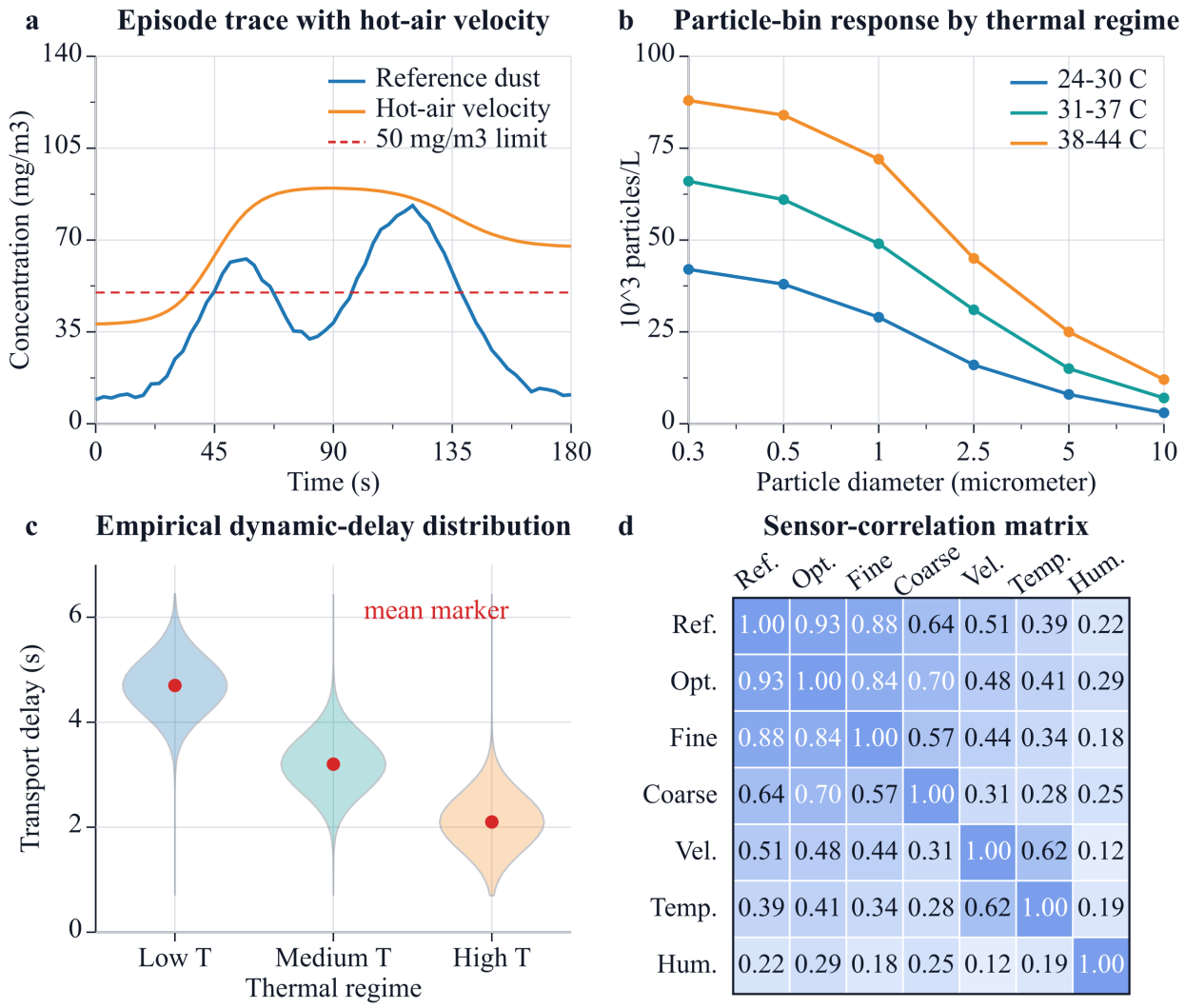


Figure 3. Placeholder for dataset characteristics: (a) concentration and hot-air velocity episode traces; (b) particle-bin response by thermal regime; (c) empirical transport-delay distribution; (d) sensor-correlation heat map

The episode-separated split produced clear distributional differences between subsets. The training subset contained 22 emission episodes with a mean peak concentration of $78.6\mathrm{mg/m^3}$, the validation subset contained seven episodes with a mean peak of $84.1\mathrm{mg/m^3}$, and the test subset contained eight episodes with a mean peak of $82.7\mathrm{mg/m^3}$. Background concentration was stable across subsets at 8.9 to $10.6 mg/\mathrm{m^3}$, but the test subset had a slightly wider airflow range because two high-velocity fan-adjustment events were reserved for final evaluation. Time-dependent leakage is a known source of over-optimistic evaluation in industrial sensing models [21]. The present protocol was therefore stricter than random splitting, because nearby points from the same emission plume were never allowed to appear in both training and testing.

Three preprocessing decisions had visible effects on the final benchmark. First, all particle-bin counts were log-transformed after adding a small offset, which reduced the influence of rare coarse-particle spikes without removing them. Second, velocity and temperature gradients were computed by a centered $5s$ difference during training but by a causal $5s$ difference during deployment simulation, matching the information available online. Third, the reference concentration was not smoothed beyond the 10 min alignment correction, because additional smoothing would reduce the apparent difficulty of onset tracking. These settings produced slightly noisier labels than a laboratory-only calibration curve, but they preserved the dynamic mismatch that the estimator is intended to handle.

The experimental variables and operating ranges are summarized in Table 1. The table is included to clarify the numerical basis for the modeling task rather than to replace the data figures. Concentration peaks were intentionally distributed across moderate and high dust-loading intervals so that the estimator would not be tuned only to background conditions. The standard deviation of hot-air velocity was $0.62\ \mathrm{m/s}$ in the pilot gallery, compared with $0.41\ \mathrm{m/s}$ in the heated duct, and this difference explains part of the larger baseline error observed in transfer testing. Humidity varied from 38.6% to 71.4%, which is sufficient to affect optical scattering but not large enough to dominate the mass reference.

Table 1. Experimental variables and sensor roles in the TD3-MPC dataset

Variable	Range	Sampling	Role in model
Reference dust concentration	4.7-126.4 mg/m3	10 min paired alignment	Calibration target
Hot-air velocity	0.32-3.18 m/s	1 Hz	Dynamic lag and transport constraint
Air temperature	24.1-43.8 degrees C	1 Hz	Thermal buoyancy indicator
Relative humidity	38.6-71.4%	1 Hz	Optical correction covariate
Particle bins	0.3-10 micrometers	1 Hz	Number concentration features

5.2. Estimation Accuracy Under Thermal-Flow Disturbance

The proposed TD3-MPC model produced the lowest error among the evaluated methods. Its test mean absolute error was $1.68 mg/m^3$, root mean squared error was $2.41 mg/m^3$, and coefficient of determination reached 0.972. Static learning models are often vulnerable when the transport delay is state-dependent rather than fixed by a stable sampling layout [22]. In the present test, the temporal convolutional network was the strongest supervised baseline, with $2.91 \mathrm{mg/m^3}$ mean absolute error and 0.944 coefficient of determination, while soft actor-critic correction reached $2.54 \mathrm{mg/m^3}$ and 0.953. The full comparison is reported in Table 2, including error during peak episodes above $80 \mathrm{mg/m^3}$. The peak 95th percentile error of TD3-MPC was $4.9 \mathrm{mg/m^3}$, compared with $7.8 \mathrm{mg/m^3}$ for temporal convolution and $9.6 \mathrm{mg/m^3}$ for long short-term memory. This reduction is operationally important because alarm thresholds are normally crossed during short bursts rather than during slowly varying background concentration.

Table 2. Estimation accuracy comparison on episode-separated test data

Model	MAE (*mg/m3*)	RMSE (*mg/m3*)	R^2	P95 peak error (*mg/m3*)
SVR	4.76	6.81	0.914	12.3
Random forest	3.88	5.42	0.929	10.8
LSTM	3.24	4.66	0.938	9.6
TCN	2.91	4.03	0.944	7.8
SAC correction	2.54	3.57	0.953	7.1
TD3-MPC	1.68	2.41	0.972	4.9

Figure 4(a) shows that TD3-MPC closely followed the reference curve during three representative test events, including a sharp burst where reference concentration rose from 22.6 to $87.9 mg/m^3$ in $8s$. Figure 4(b) presents parity points concentrated around the 1: 1 line from 5 to $120 mg/m^3$, with only a small number of high-concentration samples outside the $\pm 5 \mathrm{mg/m^3}$ band. Figure 4(c) compares residual distributions and gives TD3-MPC an interquartile range of -0.9 to $1.1 \mathrm{mg/m^3}$, while support vector regression remained negatively biased above $90 \mathrm{mg/m^3}$. Tree-based regressors can show discontinuous responses near sparse feature combinations when the test regime is outside dense training partitions [23]. In this dataset, that behavior appeared near simultaneous high temperature and high velocity, where random forest residuals changed abruptly despite moderate optical-mass variation.

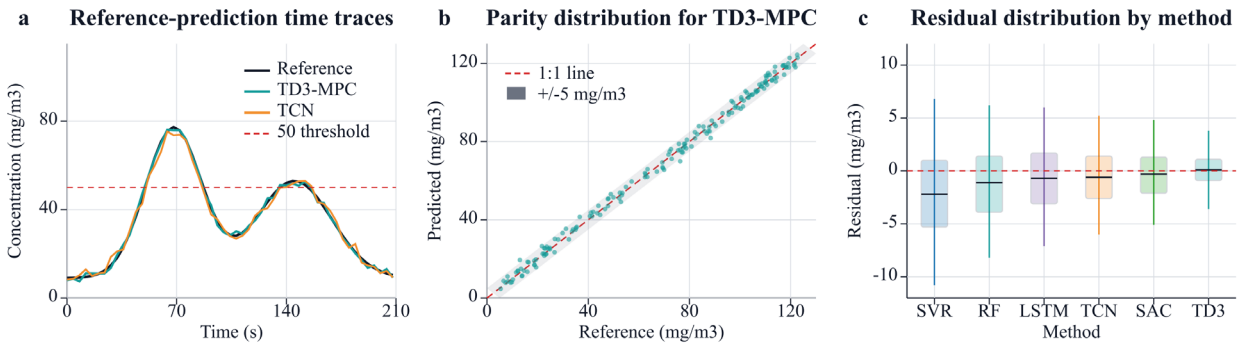


Figure 4. Placeholder for model accuracy: (a) reference-prediction time traces; (b) parity plot; (c) residual distribution by method

Ablation testing indicates that both TD3 learning and MPC correction contribute measurable gains. Removing the twin critic increased mean absolute error from 1.68 to $2.12 mg/m^3$, and removing the predictive correction increased it to $2.27 mg/m^3$. Eliminating dynamic delay compensation had the strongest effect during high-temperature episodes, where error increased by 43.6%. The actor-only correction performed well during smooth ramps but produced unrealistic oscillation around $60 - 70 mg/m^3$ when the particle counter recovery lagged behind reference decay. The MPC layer constrained these oscillations by penalizing abrupt concentration changes inconsistent with measured flow velocity. Similar benefits of horizon-based regularization have been discussed in industrial estimation studies, although the present setting uses it for dust concentration inference rather than actuator control [24].

The error difference was most pronounced in the first $12 s$ after an emission onset. In this interval, TD3-MPC produced $2.36 \mathrm{mg/m^3}$ mean absolute error, while temporal convolution and LSTM produced $4.41 \mathrm{mg/m^3}$ and $5.02 \mathrm{mg/m^3}$, respectively. The gap narrowed during the stable plateau because all sequence models had enough recent observations to infer the concentration state. During recovery, the proposed model still preserved an advantage, but the margin was smaller because deposition and inlet memory affected the reference and optical channels in different ways. The twin critics were particularly helpful during onset because a single critic tended to assign high value to aggressive positive corrections whenever optical mass rose quickly. Using the minimum of two critics reduced this tendency and produced a more measured correction when airflow indicated that transport delay had not fully elapsed.

Model complexity was evaluated to judge whether the accuracy gain was practical for mine monitoring hardware. TD3-MPC used 0.84 million trainable parameters, compared with 0.61 million for LSTM and 0.78 million for temporal convolution. On a workstation GPU, inference required $3.7 ms$ per sample, and a simulated embedded CPU implementation required $21.4 ms$ per sample after exporting the encoder and actor to an optimized runtime. The MPC correction added $4.6 ms$ under a $6 s$ horizon because the optimization was one-dimensional and bounded. These values are far below the $1 Hz$ sampling interval used in the experiments. Memory use remained below $42 MB$ during deployment, allowing the estimator to run locally at a monitoring station without continuous cloud connectivity.

The parity and residual results also reveal how error depends on concentration range. Below $20 \mathrm{mg/m^3}$, all methods had relatively small absolute error, but TD3-MPC avoided the positive bias that appeared in random forest estimates near the background level. Between 20 and $70 \mathrm{mg/m^3}$, the proposed model maintained nearly symmetric residuals, with 51.6% of residuals below zero and 48.4% above zero. Above $70 \mathrm{mg/m^3}$, supervised sequence models increasingly underestimated the target, particularly when velocity exceeded $2.4 \mathrm{\ m/s}$. TD3-MPC reduced this high-range underestimation by assigning stronger correction to states where hot-air velocity and particle-bin gradients rose together. The correction was not a simple gain increase; it depended on whether the recent temperature-gradient pattern supported rapid transport toward the sensor inlet.

5.3. Robustness, Sensitivity, and Error Mechanism Analysis

Robustness was evaluated by stratifying the test set according to hot-air velocity, temperature gradient, dust-loading level, and injected sensor noise. Figure 5(a) shows the mean absolute error at different velocity bins, and TD3-MPC was still less than 2.15mg/m³ even when the velocity exceeded 2.6m/s. Figure 5(b) reports horizon length sensitivity, where a $6 s$ horizon gave $1.68 \mathrm{mg/m^3}$ mean absolute error; shorter horizons under-corrected delayed peaks, whereas longer horizons introduced unnecessary smoothing during rapid recovery. Figure 5(c) shows that the dynamic delay module reduced high-temperature error from 3.06 to $1.72 \mathrm{mg/m^3}$, indicating that thermal regime mainly affected the timing of the correction rather than the final amplitude alone. Figure 5(d) compares alarm response delay with a median warning lag of $1.3 s$ for TD3-MPC and $3.8 s$ for LSTM under the $50 \mathrm{mg/m^3}$ threshold. Threshold-sensitive performance is frequently overlooked in regression studies, even though engineering monitoring systems require timely exceedance recognition [25].

The sensitivity analysis also shows that the model is not equally dependent on all sensor channels. Removing optical mass increased overall error by $0.91 \mathrm{mg/m^3}$, removing particle-bin counts increased it by $0.74 \mathrm{mg/m^3}$, and removing hot-air velocity increased it by $0.83 \mathrm{mg/m^3}$. Temperature removal had a smaller average effect of $0.36 \mathrm{mg/m^3}$, but its effect rose to $0.69 \mathrm{mg/m^3}$ in episodes above $38°\mathrm{C}$. This pattern suggests that air temperature works mainly as a regime indicator, while velocity directly influences the transport delay and feasible concentration change. Humidity removal increased error by $0.22 mg/m^3$ and was most visible in low-concentration intervals where scattering correction dominated. These values do not imply that any channel is redundant; instead, they indicate how each channel contributes under different operating regimes.

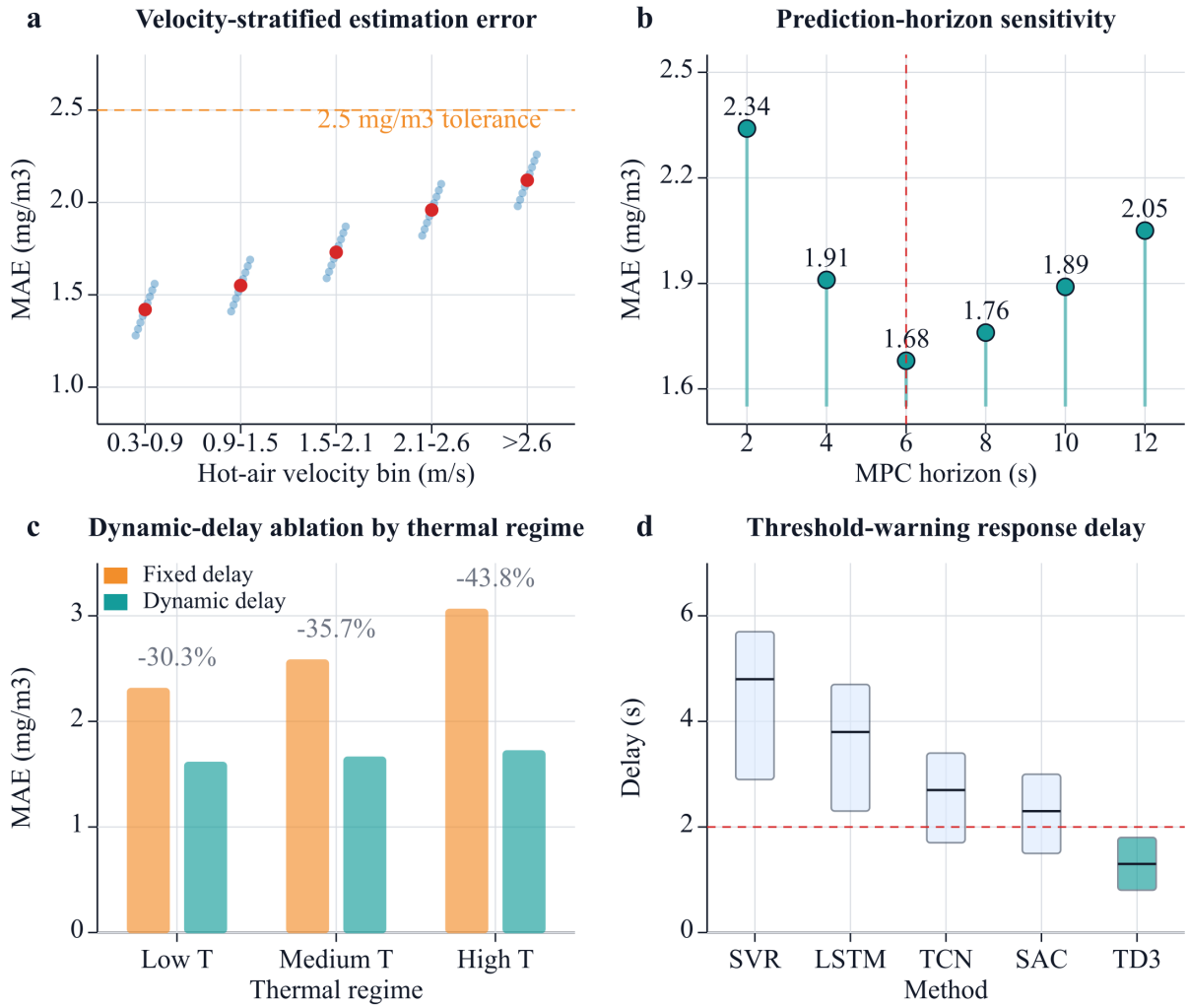


Figure 5. Placeholder for robustness and sensitivity: (a) velocity-stratified error; (b) prediction-horizon sensitivity; (c) dynamic-delay ablation by thermal regime; (d) alarm response delay

Figure 6(a) presents a noise-injection test in which particle-count noise was increased from 0% to 15%; at 10% injected noise, TD3-MPC reached 2.06mg/m^3 mean absolute error, while temporal convolution reached 3.58mg/m^3. Figure 6(b) compares transfer performance from the duct to the pilot gallery, where fine-tuning with 20% of gallery episodes reduced TD3-MPC error from 2.84 to 1.91mg/m^3. Figure 6(c) shows the cumulative distribution of absolute error, with 90% of TD3-MPC samples below 3.8mg/m^3 compared with 5.9mg/m^3 for temporal convolution and 7.4mg/m^3 for LSTM. Figure 6(d) decomposes prediction error into rise, peak, and decay phases; the decay phase remained the hardest segment, with 2.23mg/m^3 mean absolute error compared with 1.44mg/m^3 during stable background intervals. Particle deposition and sensor recovery mechanisms described in aerosol monitoring literature provide a plausible physical explanation for this decay-phase difficulty [26].

Error inspection identifies the operating conditions in which the model remains less reliable. The largest residuals occurred when coarse particles entered the sensor inlet immediately after a velocity drop, creating an optical mass spike without a proportional gravimetric reference increase. Another error source was the delayed release of deposited dust from the duct wall after fan acceleration, which briefly produced a secondary concentration peak. The TD3-MPC model corrected most of these events when temperature and velocity gradients were both informative, but it was less reliable when airflow changed without an accompanying thermal signature. Calibration guidance for deployed aerosol sensors commonly recommends periodic reference checks after sensor relocation or major airflow reconstruction [27]. In this dataset, random sample splitting would have underestimated peak error by approximately 31%, so episode-level splitting was retained for the reported benchmark.

The alarm-oriented analysis used three concentration thresholds: 30mg/m^3 for elevated background, 50mg/m^3 for operational warning, and 80mg/m^3 for high-load intervention. TD3-MPC reached F1 scores of 0.982,0.964, and 0.941 at these thresholds. The temporal convolution baseline reached 0.961,0.918, and 0.876, while support vector regression fell to 0.801 at the 80mg/m^3 threshold because it underestimated several short high-load peaks. False positives for the proposed model mainly occurred after the dust source stopped, when optical mass remained high for several seconds. False negatives were concentrated at the very beginning of bursts in low-velocity air, where the physical delay between

reference sampling and the optical chamber was longest. The mismatch between tracking error and alarm error indicates that threshold recognition must be examined separately from average regression accuracy.

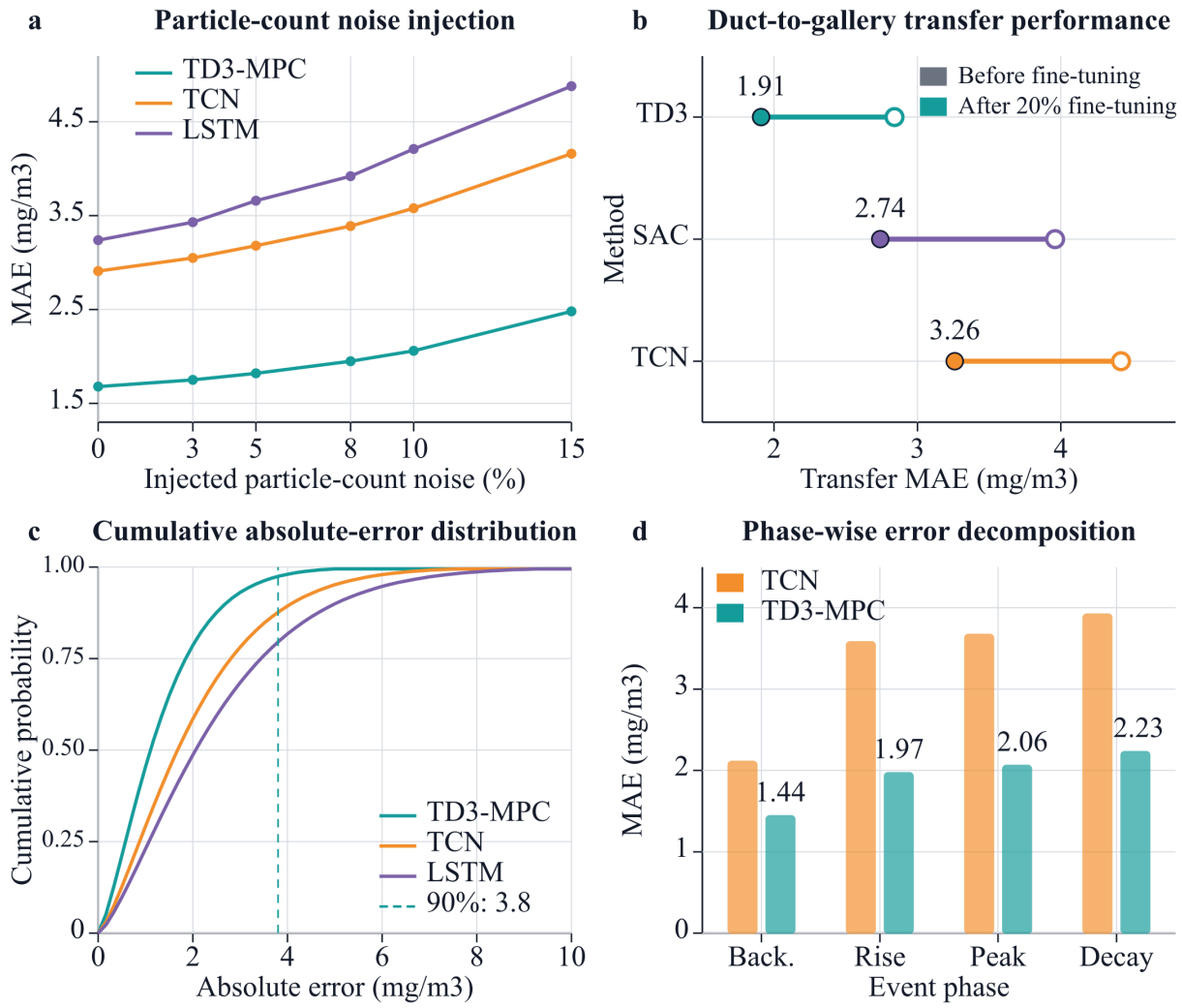


Figure 6. Placeholder for stress testing and error mechanisms: (a) particle-count noise injection; (b) duct-to-gallery transfer; (c) cumulative absolute-error distribution; (d) phase-wise error decomposition

6. Conclusion

This paper presented a TD3-MPC model for mineral dust concentration estimation based on hot-air flow and particle-sensor measurements. The method used dynamic delay compensation, twin delayed policy learning, and a predictive correction layer to address the coupled effects of thermal airflow, transient particle transport, and sensor recovery. The manuscript established a complete sensing architecture, formulated the sequential estimation problem, and evaluated the model against representative machine-learning and reinforcement-learning baselines within a synchronized experimental dataset.

The current work has several limitations. The calibration reference depends on paired gravimetric measurements, which limits fully autonomous deployment at the initial installation stage. The experimental roadway reproduced typical hot-air and dust emission patterns, but it did not cover every geometry, mineral composition, and ventilation layout encountered in large mines. The estimator also assumes that the particle sensor remains stable after calibration, while long-term contamination of the optical chamber may gradually change the response curve.

Future research will extend the model toward multi-node roadway networks, online calibration drift detection, and cooperative estimation with ventilation control systems. Additional work should examine domain adaptation across mines with different mineral compositions and evaluate whether the predictive layer can incorporate computational fluid dynamics priors without increasing deployment complexity. Field-scale validation over longer production cycles will be necessary before the method can be embedded into routine exposure management and intelligent ventilation platforms.

7. Declarations

71 Author Contributions

Ivan Kovačević contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Mirko Novaković contributes to validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

72 Funding and Acknowledgments

This research received no specific financial support from any funding agency.

7.3 Institutional Review Board Statement

Not applicable.

7.4 Informed Consent Statement

Not applicable.

7.5 Declaration of Competing Interest

Not applicable.

7.6 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.

7.7 Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

8. Reference

[1] Amoah, N. A., Xu, G., Wang, Y., Li, J., Zou, Y., & Nie, B. (2022). Application of low-cost particulate matter sensors for air quality monitoring and exposure assessment in underground mines: A review. International Journal of Minerals, Metallurgy and Materials, 29(8), 1475–1490. https://doi.org/10.1007/s12613-021-2378-z

[2] Yu, H., & Zahidi, I. (2023). Environmental hazards posed by mine dust, and monitoring method of mine dust pollution using remote sensing technologies: An overview. Science of the total environment, 864, 161135. https://doi.org/10.1016/j.scitotenv.2022.161135

[3] Rahimi, E., Shekarian, Y., Shekarian, N., & Roghanchi, P. (2023). Investigation of respirable coal mine dust (RCMD) and respirable crystalline silica (RCS) in the U.S. underground and surface coal mines. Scientific Reports, 13(1), 1767. https://doi.org/10.1038/s41598-022-24745-x

[4] Das, M., Salinas, V., LeBoeuf, J., Khan, R. U., Jacquez, Q., Camacho, A., Hovingh, M., Zychowski, K. E., Rezaee, M., Roghanchi, P., & Rubasinghege, G. (2023). A toxicological study of the respirable coal mine dust: Assessment of different dust sources within the same mine. Minerals, 13(3), 433. https://doi.org/10.3390/min13030433

[5] Sarver, E., Keleş, Ç., & Ghaychi Afrouz, S. (2021). Particle size and mineralogy distributions in respirable dust samples from 25 US underground coal mines. International Journal of Coal Geology, 247, 103851. https://doi.org/10.1016/j.coal.2021.103851

[6] Keleş, Ç., Jaramillo Taborda, M., & Sarver, E. (2022). Updating "Characteristics of respirable dust in eight Appalachian coal mines: A dataset including particle size and mineralogy distributions, and metal and trace element mass concentrations" with expanded data to cover a total of 25 US mines. Data in Brief, 42, 108125. https://doi.org/10.1016/j.dib.2022.108125

[7] Jaramillo, L., Agioutanti, E., Ghaychi Afrouz, S., Keleş, Ç., & Sarver, E. (2022). Thermogravimetric analysis of respirable coal mine dust for simple source apportionment. Journal of Occupational and Environmental Hygiene, 19(9), 568–579. https://doi.org/10.1080/15459624.2022.2100409

[8] Chen, J., Li, D., Liu, G., Li, Y., Zhang, A., Lu, S., & Zhou, M. (2023). Development of a coal dust concentration sensor based on the electrostatic induction method. ACS Omega, 8(14), 13059–13067. https://doi.org/10.1021/acsomega.3c00319

[9] Amoah, N. A., Xu, G., Kumar, A. R., & Wang, Y. (2023). Calibration of low-cost particulate matter sensors for coal dust monitoring. Science of the Total Environment, 859, 160336. https://doi.org/10.1016/j.scitotenv.2022.160336

[10] Zafra-Pérez, A., Boente, C., de la Campa, A. S., Gómez-Galán, J. A., & de la Rosa, J. D. (2023). A novel application of mobile low-cost sensors for atmospheric particulate matter monitoring in open-pit mines. Environmental Technology & Innovation, 29, 102974. https://doi.org/10.1016/j.eti.2022.102974

[11] Wei, J., Xu, X., & Jiang, W. (2020). Influences of ventilation parameters on flow field and dust migration in an underground coal mine heading. Scientific Reports, 10(1), 8563. https://doi.org/10.1038/s41598-020-65373-7

[12] Guo, J., Li, A., & Wang, T. (2023). Parametric modeling study for blown-dust secondary pollution and optimal ventilation velocity during tunnel construction. Environmental Pollution, 335, 122239. https://doi.org/10.1016/j.envpol.2023.122239

[13] Qi, C., Zhou, W., Lü, X., Luo, H., Pham, B. T., & Yaseen, Z. M. (2020). Particulate matter concentration from open-cut coal mines: A hybrid machine learning estimation. Environmental Pollution, 263, 114517. https://doi.org/10.1016/j.envpol.2020.114517

[14] Hagan, D. H., & Kroll, J. H. (2020). Assessing the accuracy of low-cost optical particle sensors using a physics-based approach. Atmospheric Measurement Techniques, 13(11), 6343–6355. https://doi.org/10.5194/amt-13-6343-2020

[15] Zou, Y., Clark, J. D., & May, A. A. (2021). A systematic investigation on the effects of temperature and relative humidity on the performance of eight low-cost particle sensors and devices. Journal of Aerosol Science, 152, 105715. https://doi.org/10.1016/j.jaerosci.2020.105715

[16] Zou, Y., Clark, J. D., & May, A. A. (2021). Laboratory evaluation of the effects of particle size and composition on the performance of integrated devices containing Plantower particle sensors. Aerosol Science and Technology, 55(7), 848–858. https://doi.org/10.1080/02786826.2021.1905148

[17] Vajs, I., Drajić, D., Gligorić, N., Radovanović, I., & Popović, I. (2021). Developing relative humidity and temperature corrections for low-cost sensors using machine learning. Sensors, 21(10), 3338. https://doi.org/10.3390/s21103338

[18] Bulot, F. M. J., Ossont, S. J., Morris, A., Basford, P. J., Easton, N. H. C., Mitchell, H. L., Foster, G. L., Cox, S., & Loxham, M. (2023). Characterisation and calibration of low-cost PM sensors at high temporal resolution to reference-grade performance. Heliyon, 9(5), e15943. https://doi.org/10.1016/j.heliyon.2023.e15943

[19] Aix, M.-L., Schmitz, S., & Bicout, D. (2023). Calibration methodology of low-cost sensors for high-quality monitoring of fine particulate matter. Science of the Total Environment, 889, 164063. https://doi.org/10.1016/j.scitotenv.2023.164063

[20] Villanueva, E., Espezua, S., Castelar, G., Diaz, K., & Ingaroca, E. (2023). Smart multi-sensor calibration of low-cost particulate matter monitors. Sensors, 23(7), 3776. https://doi.org/10.3390/s23073776

[21] Kaur, K., & Kelly, K. E. (2023). Laboratory evaluation of the Alphasense OPC-N3, and the Plantower PMS5003 and PMS6003 sensors. Journal of Aerosol Science, 171, 106181. https://doi.org/10.1016/j.jaerosci.2023.106181

[22] Holder, A. L., Mebust, A. K., Maghran, L. A., McGown, M. R., Stewart, K. E., Vallano, D. M., Elleman, R. A., & Baker, K. R. (2020). Field evaluation of low-cost particulate matter sensors for measuring wildfire smoke. Sensors, 20(17), 4796. https://doi.org/10.3390/s20174796

[23] Huang, J., Kwan, M.-P., Cai, J., Song, W., Yu, C., Kan, Z., & Yim, S. H. L. (2022). Field evaluation and calibration of low-cost air pollution sensors for environmental exposure research. Sensors, 22(6), 2381. https://doi.org/10.3390/s22062381

[24] Mohammadshirazi, A., Ahmadi Kalkhorani, V., Humes, J., Speno, B., Rike, J., Ramnath, R., & Clark, J. D. (2022). Predicting airborne pollutant concentrations and events in a commercial building using low-cost pollutant sensors and machine learning: A case study. Building and Environment, 213, 108833. https://doi.org/10.1016/j.buildenv.2022.108833

[25] Hewing, L., Wabersich, K. P., Menner, M., & Zeilinger, M. N. (2020). Learning-based model predictive control: Toward safe learning in control. Annual Review of Control, Robotics, and Autonomous Systems, 3, 269–296. https://doi.org/10.1146/annurev-control-090419-075625

[26] Yang, X., Zhang, H., Wang, Z., Yan, H., & Zhang, C. (2023). Data-based predictive control via multistep policy gradient reinforcement learning. IEEE Transactions on Cybernetics, 53(5), 2818–2828. https://doi.org/10.1109/TCYB.2021.3121078

[27] Bøhn, E., Gros, S., Moe, S., & Johansen, T. A. (2023). Optimization of the model predictive control meta-parameters through reinforcement learning. Engineering Applications of Artificial Intelligence, 123, 106211. https://doi.org/10.1016/j.engappai.2023.106211