DOI: 10.37943/PEXM7788

Dilara Abzhanova
PhD, Director, Center of Competence and Excellence
dilara.abzhanova@astanait.edu.kz, orcid.org/0000-0002-7988-3971
Astana IT University, Kazakhstan
Andrii Biloshchytskyi
Doctor of Technical Sciences, Professor, Vice-rector for Science and Innovation
bao1978@gmail.com, orcid.org/0000-0001-9548-1959
Astana IT University, Kazakhstan
Professor, Department of Information Technology
Kyiv National University of Construction and Architecture, Ukraine

AI-DRIVEN ATMOSPHERIC EMISSION MONITORING SYSTEM AS A FRAMEWORK FOR SUSTAINABILITY-ORIENTED PROJECT MANAGEMENT

Abstract: This paper presents an artificial intelligence (AI)-driven atmospheric emission monitoring framework for sustainability-oriented project management that integrates predictive analytics, blockchain-based data verification, and digital-twin technologies. Reliable environmental monitoring and short-term forecasting are increasingly required to support regulatory compliance, environmental risk assessment, and evidence-based project decision-making. The objective of this study is to develop and evaluate an integrated framework combining real-time industrial monitoring, neural-network forecasting, and decision-support mechanisms.

The empirical forecasting component uses 39,803 synchronized time-stamped observations obtained from an operating industrial monitoring system in Kazakhstan with a 20-minute temporal resolution. The monitored variables include NO, NO_2, SO_2, CO, particulate matter, oxygen concentration, temperature, humidity, pressure, and gas-flow indicators. After quality control, synchronization, outlier removal, and Min-Max normalization, historical sequences of 72 consecutive observations (24 hours) were generated. A Long Short-Term Memory (LSTM) neural network with recurrent layers of 32 and 16 hidden units, a dropout rate of 0.2, and a dense output layer was trained for one-step-ahead forecasting corresponding to the next 20-minute observation. The dataset was divided chronologically into 64% training, 16% validation, and 20% testing subsets to preserve temporal dependencies.

The proposed model achieved an overall MSE of 0.87 and R^2 of 0.86, while reducing pollutant-specific RMSE by 33.1–46.8% compared with the ARIMA baseline. The monitoring platform was additionally piloted at Promanalit LLP, where telemetry acquisition, preprocessing, visualization, reporting, and blockchain-based microblock verification were successfully validated. The study explicitly distinguishes empirical forecasting and platform validation from conceptual digital-twin and ESG scenarios, which are included only to demonstrate project decision-support capabilities. The proposed framework provides a practical basis for intelligent environmental monitoring, short-term emission forecasting, and sustainability-oriented project management in industrial applications.

Keywords: artificial intelligence, atmospheric emissions, digital twin, ESG indicators, project management.

Introduction

The increasing anthropogenic pressure on the environment and the increasing complexity of industrial and infrastructure projects reinforce the importance of environmental factors in project management. Atmospheric emissions, including fine particles and gaseous pollutants, pose persistent environmental and social risks that increasingly go beyond traditional impact assessments and require continuous monitoring during project implementation. In the context of

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

increased ESG regulation and public control, design organizations face the need not only to record environmental indicators, but also to use them as a basis for management decisions.

Classical project management methodologies, as a rule, consider environmental aspects in the form of static constraints or reporting indicators formed at individual stages of the life cycle. This approach limits the ability of project teams to respond in a timely manner to changes in the environmental situation and reduces the effectiveness of risk management related to emissions and atmospheric air quality. As a result, environmental information remains poorly integrated into operational planning, scenario analysis, and decision-making processes.

The development of digital technologies opens up opportunities for rethinking the role of environmental data in project management. Artificial intelligence and machine learning methods make it possible to identify complex nonlinear dependencies in the dynamics of emissions and form short-term forecasts. The digital twin concept provides a virtual representation of atmospheric processes and supports the assessment of alternative management scenarios. Blockchain technologies, in turn, create the basis for ensuring the reliability, immutability and traceability of environmental data, which is crucial for ESG reporting and trust from stakeholders.

Despite the active development of each of these areas, existing research mainly focuses on individual technological components and rarely considers them as elements of a unified project management system. The issue of integrating intelligent monitoring of atmospheric emissions, digital twins, data trust mechanisms and ESG metrics within a holistic management framework remains insufficiently developed in the scientific literature.

This study was conducted within a broader Program-Targeted Funding research agenda focused on the development of intelligent environmental monitoring systems and the digital transformation of sustainability-oriented management for industrial and urban territories in Kazakhstan. Within this broader context, the present paper focuses on the atmospheric emission monitoring component and examines it as a project decision-support framework that integrates environmental sensing, forecasting, digital-twin-based simulation, blockchain validation, and ESG-oriented performance assessment.

The purpose of this study is to develop and evaluate an AI-driven atmospheric emission monitoring framework that supports sustainability-oriented project management by integrating real environmental monitoring data, predictive analytics, atmospheric digital-twin simulation, blockchain-based validation, and ESG-oriented decision support.

The main contribution of the paper is the development of an integrated decision-support framework that connects atmospheric emission monitoring with project governance. More specifically, the study: (i) proposes an architecture combining distributed IoT sensing, machine learning forecasting, an atmospheric digital twin, and blockchain-based data validation; (ii) introduces three ESG-oriented project indicators — the Emission Compliance Index, the Air Quality Impact Score, and the Sustainable Project Performance Score; (iii) demonstrates the use of scenario-based digital twin simulation for comparing baseline, mitigation, and construction-related emission scenarios; and (iv) evaluates the proposed framework using multi-pollutant monitoring data, forecasting accuracy metrics, statistical validation, and sustainability performance assessment.

Literature review

Modern research demonstrates a pronounced technological shift towards environmental management based on data, artificial intelligence and principles of transparency, which directly supports the concept of sustainability-oriented project management [1-2].

AI-Based Atmospheric Emission Forecasting. Machine learning models are superior to traditional statistical methods for predicting atmospheric pollutants due to their ability to identify nonlinear temporal and spatial dependencies in environmental data [3-4]. A number of studies show that hybrid deep learning architectures combining variational modal decomposition (VMD), graph attention networks, and recurrent neural networks provide high accuracy in predicting PM concentrations$2._5$ and other pollutants [5].

In particular, Wang et al. proposed a spatiotemporal deep learning model combining VMD, Graph Attention Networks (GAT), and bidirectional LSTM networks for predicting $PM_{2.5}$, demonstrating high efficiency in an urban environment with pronounced microclimatic variability [6]. The results obtained emphasize the need to simultaneously take into account temporal dependencies and spatial correlations, which is fundamentally important for predicting emissions in the context of sustainable project management.

Ensemble learning approaches have demonstrated superior performance in short-term air pollution forecasting compared with traditional statistical methods and single neural network models [7]. These approaches provide a robust methodological foundation for developing sustainable forecasting modules within intelligent environmental monitoring systems.

Spatiotemporal Pollution Modeling and Environmental Monitoring Systems. In addition to forecasting, considerable attention in the literature is paid to the development of integrated environmental monitoring systems that integrate IoT sensors, satellite data, and spatiotemporal analytics methods [8-9]. Review studies emphasize that combining data from diverse sources significantly increases the accuracy of describing the dynamics of pollutants and the resilience of models to data incompleteness [10].

Multimodal information processing pipelines are a key element in building high-precision digital environmental doubles - virtual models that reproduce real atmospheric processes and are dynamically updated based on streaming data [11]. This integration ensures the calibration and validation of spatial and temporal models underlying environmentally sound management decisions during project implementation.

Digital Twins for Environmental and Industrial Systems. The concept of digital twins (DT) has evolved from applications in industry and manufacturing to the tasks of environmental management and sustainable development of urban and infrastructural systems [12]. Digital Earth twins have emerged as a transformative approach for modeling climate and environmental processes, enabling simulation-based analysis to support informed environmental and societal decision-making [13]. Despite the global scale of their research, the architectural principles - continuous data collection, high spatial detail, and iterative scenario evaluation — are directly applicable to the atmospheric digital twin proposed in this paper.

In the context of infrastructure management, digital twins demonstrate the potential to increase resilience through lifecycle analysis, predictive maintenance, and system-level optimization [14].

Blockchain-Enabled Environmental Data Integrity. The reliability of environmental reporting is a key condition for compliance with ESG requirements. A number of studies point to the vulnerability of air quality data to manipulation and highlight the potential of blockchain technologies in ensuring the integrity, transparency and traceability of environmental information [15].

An Ethereum-based air quality monitoring system can ensure data immutability and verifiability through decentralized storage [16]. In addition, microblock architectures, distributed ledgers, and smart contracts can strengthen data transparency and increase trust among regulators and other stakeholders [17]. More recent research demonstrates the integration of blockchain with federated learning for industrial air pollution monitoring, enhancing both transparency and data protection [18].

ESG and SDG Frameworks for Sustainability Assessment. The role of ESG indicators in the evaluation of project activities is increasingly emphasized in the studies of sustainable development of organizations. The integration of the Sustainable Development Goals (SDGs) and Environmental, Social, and Governance (ESG) principles provides a structured approach for formalizing environmental indicators, including air quality, through multi-criteria assessment methods [19]. The possibilities of quantitative integration of ESG indicators into strategic management using hybrid analysis methods have been demonstrated [20].

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

At the organizational level, empirical research shows that companies with a high focus on sustainable development achieve better financial and operational results, largely due to effective environmental management mechanisms [21-22].

Collectively, the literature analysis reveals five converging trends: (i) the increasing accuracy of AI emissions forecasting, (ii) the development of digital twins as scenario assessment tools, (iii) the use of blockchain to ensure trust in environmental data, (iv) the strengthening of ESG requirements for quantitative environmental indicators, and (v) the increasing role of dynamic environmental analytics in sustainability-oriented project management. Despite the maturity of each field individually, an extremely limited number of studies combine these approaches within a single project management system, which confirms the scientific novelty of the proposed platform.

Results

This section presents the methodological framework, system architecture, and a case-oriented demonstration of the proposed artificial intelligence-based atmospheric emissions monitoring system, developed as a decision support system for sustainability-oriented project management. The integrated view reflects the iterative nature of environmental intelligence systems, in which architecture and practical application are closely interrelated.

Overview And Architectural Framework Of The System. The system integrates five technological levels — data collection, data aggregation, predictive analytics, digital twin modeling, and data validation based on blockchain microblocks — into a single ecosystem that provides real-time environmental information and design decisions consistent with ESG principles (Figure 1).

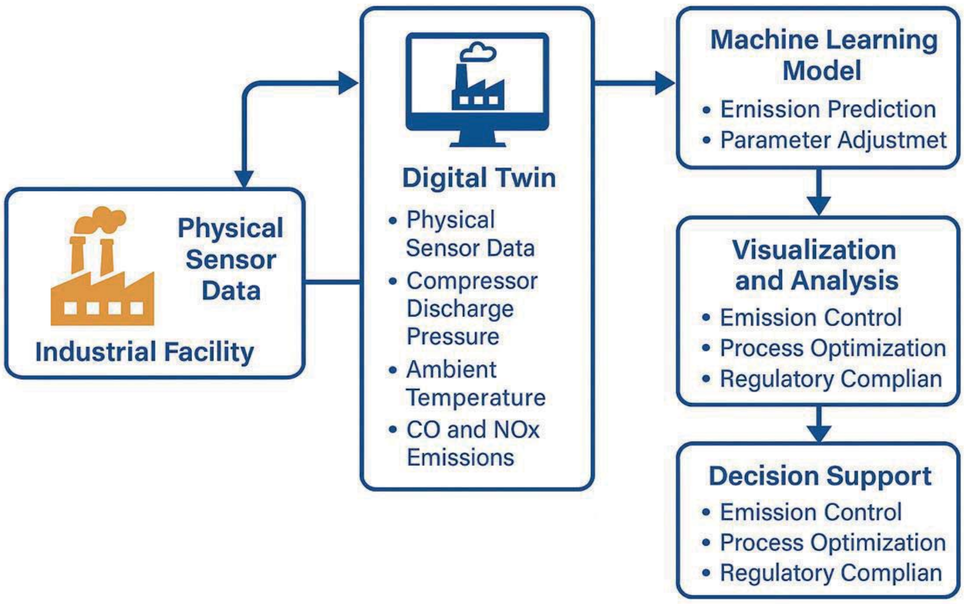

Figure 1. Architecture of the AI-driven atmospheric emission monitoring and decision-support system

Data Sources and Preprocessing. The empirical forecasting dataset consists of 39,803 synchronized time-stamped observations obtained from operating industrial monitoring systems in Kazakhstan. The software and telemetry pipeline were piloted at Promanalit LLP in Pavlodar, Kazakhstan. The records originate from a stationary automated monitoring installation positioned at an industrial emission source. Exact geographic coordinates are not disclosed because of the enterprise data-sharing and industrial-security restrictions.

The monitored channels include NO, NO2, SO2, CO, particulate matter (DUST), O2, sample temperature, humidity, pressure, and gas-flow indicators. The source telemetry was generated at a 20-minute interval. The uploaded raw export contains long-format records from multiple measurement channels; after timestamp alignment and channel synchronization, these measurements were transformed into 39,803 multivariate analytical observations. Thus, the number of raw sensor messages is not equivalent to the number of model-ready time points.

Before model training, records with missing, duplicated, or technically invalid values were removed. Timestamps were aligned to a common 20-minute grid, and numerical variables were normalized to the [0,1] range using Min-Max scaling. Input sequences of 72 consecutive observations were then formed, representing a 24-hour historical window.

The sequences were divided chronologically into training, validation, and test subsets in proportions of 64%, 16%, and 20%, respectively. Random shuffling was not applied, and temporal overlap between subsets was excluded to prevent information leakage. The training subset was used to estimate model parameters, the validation subset controlled model selection and early stopping, and the final 20% was reserved for out-of-sample evaluation.

Forecasting Model. The predictive layer was implemented as a one-step-ahead LSTM model. Each input sample contains 72 consecutive 20-minute observations and is used to predict the pollutant value for the immediately following 20-minute interval. This forecasting horizon was selected to support short-term operational warning within the monitoring platform.

The neural architecture consists of an input layer, a first LSTM layer with 32 units, a Dropout layer with a rate of 0.2, a second LSTM layer with 16 units, and a final Dense layer with one output neuron. Training used the Adam optimizer, mean-squared-error loss, a learning rate of 0.001, a batch size of 32, and a maximum of 100 epochs. Early stopping was applied according to validation loss. A classical ARIMA model selected using the Akaike information criterion was trained on the same chronological data as the forecasting baseline.

Machine learning models are employed to forecast atmospheric pollutant concentrations and to identify anomalous deviations. The general form of the forecasting model is defined by Equation (1):

$$\hat{C}_{t+k} = f(C_t, C_{t-1}, \dots, M_t, S_t, \theta), \tag{1}$$

where $\hat{C}_{t+1}$ denotes the predicted pollutant concentration for the next 20-minute interval; C_t and C_{t-1} are the current and previous concentrations; M_t is the vector of meteorological and operating variables; S_t represents the available source and sensor context; and θ denotes the model parameters.

Let

$$X_t = \{PM2.5_t, PM10_t, NO2_t, SO2_t, CO_t, O3_t, M_t\}$$

denote the multivariate input vector at time *t*, where $PM2.5_t$, $PM10_t$, $NO2_t$, $SO2_t$, CO_t, and $O3_t$ correspond to the concentrations of major atmospheric pollutants, and M_t includes meteorological parameters such as air temperature, relative humidity, wind speed, and wind direction.

The forecasting task is formulated as a supervised one-step-ahead learning problem. For each time t, the model uses a historical window of n = 72 observations to estimate the pollutant concentration at t + 1, corresponding to a 20-minute forecasting horizon. The forecasting model is expressed as:

$$\hat{y}_{t+k} = f\theta(X_t, X_{t-1}, \dots, X_{t-n}), \tag{2}$$

where fθ denotes the LSTM model with parameters θ, n = 72 is the historical window length, and k = 1 is the forecasting horizon.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

Model training is performed by minimizing the mean squared error (MSE), defined as:

$$L(\theta) = (1 / N) \cdot \Sigma_{i=1}^{N} (\hat{y}_i - y_i)^2, \tag{3}$$

where y_i represents the observed pollutant concentration, $\hat{y}_i$ is the corresponding predicted value, and N denotes the number of observations in the training dataset.

In this study, the LSTM recurrent neural network is used to capture nonlinear temporal dependencies in pollutant dynamics. ARIMA is used only as a classical baseline for comparison. XGBoost and CatBoost were removed from the revised experimental description because they were not part of the reported forecasting experiment.

Anomalous pollution levels are detected by comparing observed and predicted concentrations. An anomaly at time *t* is defined as follows:

$$A_t = 1, \text{ if } |C_t - \hat{C}_t| > \delta; \; A_t = 0, \text{ otherwise.} \tag{4}$$

Here, C_t denotes the observed pollutant concentration at time *t*, $\hat{C}_t$ is the predicted value, and δ represents a dynamic sensitivity threshold.

The threshold δ is computed adaptively based on the statistics of prediction errors within a rolling time window:

$$\delta = \mu_e + \lambda \cdot \sigma_e, \tag{5}$$

where μ_e and σ_e denote the mean and standard deviation of the prediction error, respectively, and λ is a sensitivity parameter.

This approach enables the anomaly detection procedure to adapt to changing environmental conditions and supports early identification of potential exceedances of regulatory emission limits.

Forecasting accuracy is evaluated using the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of determination (R2). Pollutant-specific RMSE values are used to compare LSTM with the ARIMA baseline.

The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (C_i - \hat{C}_i)^2}, \tag{6}$$

where C_i denotes the observed pollutant concentration, $\hat{C}_i$ represents the corresponding predicted value, and n is the number of observations.

The MAPE is computed as:

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{C_i - \hat{C}_i}{C_i} \right|, \tag{7}$$

providing a relative measure of forecasting error expressed as a percentage.

Table 1 presents pollutant-specific forecasting errors for the LSTM model and the ARIMA baseline on the chronologically held-out test subset.

Table 1. Empirical comparison of LSTM and ARIMA forecasting accuracy

Pollutant	ARIMA RMSE	LSTM RMSE	RMSE reduction
NO	0.145	0.088	39.3%
NO2	0.158	0.084	46.8%
SO2	0.121	0.081	33.1%
CO	0.210	0.131	37.6%

The LSTM model reduced RMSE by 33.1–46.8% across the evaluated pollutants. The overall forecasting experiment achieved MSE = 0.87 and R2 = 0.86. These values are empirical model results obtained from the synchronized industrial dataset, not outputs of the conceptual digital-twin scenarios.

Atmospheric Digital Twin Environment. The digital-twin layer is a Python-based analytical and visualization environment that combines observed and forecast values and supports prospective comparison of alternative project actions. In the present study, the forecasting module and platform operation were empirically evaluated, whereas the spatial mitigation and construction scenarios are presented as a conceptual decision-support demonstration. The scenario outputs should therefore not be interpreted as measured effects of completed industrial interventions.

The implementation uses Python data-processing and machine-learning libraries integrated with the monitoring platform. The environmental state can be updated at the source telemetry interval of 20 minutes. The digital-twin equations below formalize the state representation and transition logic retained for scenario analysis, hotspot visualization, and project risk assessment.

Implementation and Reproducibility. The forecasting workflow was implemented using Python, pandas, NumPy, scikit-learn, and TensorFlow/Keras. The monitoring platform uses containerized microservices, Apache Kafka, TimescaleDB, REST APIs, and the ECO permissioned blockchain. An anonymized extract and source-code components may be provided for academic verification subject to the industrial data-sharing agreement.

The environmental state at time t is represented as:

$$E_t = g(C_t, M_t, \varphi), \tag{8}$$

where C_t denotes pollutant concentrations, M_t represents meteorological parameters, and φ captures spatial and contextual characteristics.

The digital twin supports multiple analytical functions, including hotspot identification, environmental risk mapping, comparison of emission mitigation scenarios, and evaluation of the environmental consequences of project-related decisions.

To ensure data integrity and trustworthiness, each sensor measurement, machine learning output, and ESG indicator is stored in the form of a blockchain micro-block:

$$H_i = h(D_i \parallel H_{i-1}), \tag{9}$$

where D_i denotes the current data record and H_{i-1} is the hash of the previous block. This mechanism prevents data manipulation and enhances the reliability of ESG reporting for both internal and external stakeholders. The use of a micro-block structure is important for environmental monitoring because emission data are generated continuously and require frequent recording with minimal latency. Compared with conventional block formation, micro-block validation reduces the computational and storage burden while preserving traceability and immutability of monitoring records.

The atmospheric digital twin is further formalized as a state-transition system describing the evolution of environmental conditions over time:

$$E_{t+1} = F(E_t, U_t, \varepsilon_t), \tag{10}$$

where E_t denotes the environmental state vector at time t, U_t represents controllable project-related interventions such as emission reduction measures or operational adjustments, and ε_t captures stochastic disturbances associated with meteorological variability and external uncertainty.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

This formulation enables the use of the digital twin for simulation-based analysis, scenario comparison, and assessment of environmental outcomes of alternative project management decisions within a sustainability-oriented framework.

To assess the practical feasibility of the proposed system in project environments, the computational complexity and update frequency of each core component were analyzed. Table 2 summarizes the operational characteristics of the system from a project management perspective.

Table 2. Computational complexity and update frequency of core system components

System component	Update frequency	Computational complexity	Role in project management
ML-based emission forecasting	Every 20 minutes	O(n·T)	Early risk anticipation and proactive decision-making
Atmospheric digital twin simulation	Scenario-triggered / every 20 minutes	O(S·G)	Scenario-based analysis and evaluation of mitigation strategies
Blockchain micro-block validation	Event-driven	O(1)	Ensuring data integrity and trust for ESG reporting

where n denotes the number of synchronized input channels, T = 72 is the historical time-window length, S represents the number of spatial grid cells used in a scenario, and G is the number of state-update steps.

The results indicate that the proposed architecture maintains a balanced trade-off between computational cost and decision-support value, making it suitable for continuous use throughout the entire project lifecycle.

Conceptual Digital-Twin Scenario Demonstration. To illustrate the project decision-support logic of the proposed framework, a conceptual industrial-urban scenario with elevated PM2.5 and NO2 levels was constructed. The baseline, 30% mitigation, and construction-peak alternatives are simulation assumptions. They are not presented as direct measurements from the Promanalit pilot or as completed managerial interventions.

For the scenario calculations, the assumed baseline values are PM2.5 = 34 µg/m³, NO2 = 48 ppb, SO2 = 12 ppb, and CO = 1.9 ppm. These values are used exclusively to demonstrate the calculation of ECI, AQIS, SPPS, and comparative digital-twin outputs. They must be interpreted as scenario inputs rather than descriptive statistics of the 39,803-observation forecasting dataset.

To evaluate the empirical forecasting improvement, paired forecast errors from LSTM and ARIMA were compared using a paired t-test at $\alpha = 0.05$. The null hypothesis H0 assumes no difference in the mean paired errors, whereas H1 assumes that the LSTM errors are lower.

The paired t-statistic is computed as:

$$t = \bar{d} / (s\_d / \sqrt{n}), \tag{11}$$

where $\bar{d}$ is the mean paired difference between the ARIMA and LSTM errors, s_d is the standard deviation of the paired differences, and n is the number of paired forecasts.

The paired comparison confirmed statistically significant error reductions for the evaluated pollutants ($p < 0.05$). Table 3 summarizes the empirical RMSE reductions used in the statistical comparison.

Table 3. Statistical validation of empirical forecasting improvements

Pollutant	Model comparison	RMSE reduction	Significance
NO	LSTM vs ARIMA	39.3%	$p < 0.05$
NO2	LSTM vs ARIMA	46.8%	$p < 0.05$
SO2	LSTM vs ARIMA	33.1%	$p < 0.05$

The statistical validation concerns only the empirical LSTM–ARIMA forecasting comparison. It does not validate the numerical outputs of the conceptual digital-twin scenarios.

The platform includes threshold-based early-warning logic that compares observed and forecast concentrations with configured limits. During this study, the warning mechanism was functionally tested; the manuscript does not claim four documented managerial interventions at an operating facility.

For the conceptual demonstration, the digital-twin state was recalculated at a 20-minute temporal resolution. Illustrative zones and heatmaps were generated to show how forecast and monitoring data could be translated into spatial project-risk information.

Three conceptual scenarios were evaluated: a baseline configuration, a planned 30% emission-reduction scenario, and a temporary construction-related emission increase.

The scenario calculations produced an illustrative 21% reduction in the PM2.5 peak under the mitigation alternative and an illustrative 15% increase in NO2 under the construction alternative. These percentages are simulation outputs used to demonstrate comparative decision support; they are not measured effects of implemented actions.

In the operational platform, environmental records and forecast metadata can be associated with blockchain microblocks. The conceptual scenario uses this verified-data pathway to demonstrate how ECI and other project indicators may be made auditable.

The experimentally validated blockchain results concern data integrity, latency, throughput, and Merkle-proof verification. The ECI, AQIS, SPPS, heatmaps, and scenario outcomes below remain conceptual management calculations.

ESG Indicators and Sustainability Metrics. Three quantitative indicators are retained to demonstrate how monitored or forecast environmental data can be converted into project-management metrics: the Emission Compliance Index (ECI), Air Quality Impact Score (AQIS), and Sustainable Project Performance Score (SPPS). In this article, the formulas are methodological contributions, while the numerical examples in Tables 5–7 are conceptual calculations based on the stated scenario inputs.

The proposed indicators are explicitly mapped to the Environmental (E) pillar of ESG and to project governance mechanisms. The ECI reflects regulatory compliance and operational environmental control, the AQIS captures external environmental impact and exposure intensity, and the SPPS integrates compliance, impact mitigation, and project responsiveness into a single decision-support metric. This mapping enables the systematic translation of real-time emission intelligence into standardized ESG reporting and sustainability-oriented project evaluation. The correspondence between the proposed indicators and ESG dimensions is summarized in Table 4.

Table 4. Mapping proposed indicators to ESG dimensions

Indicator	**ESG pillar**	**Sustainability focus**	**Project relevance**
ECI	Environmental (E)	Regulatory compliance	Risk management
AQIS	Environmental (E)	External impact	Stakeholder protection
SPPS	ESG-integrated	Sustainable performance	Strategic decision-making

The Emission Compliance Index is defined as:

$$ECI = 1 - \frac{1}{n}\sum_{i=1}^{n} \max\left(0, \frac{C_i - L_i}{L_i}\right), \quad (12)$$

where C_i denotes the mean observed concentration of pollutant i, L_i represents the corresponding regulatory limit, and n is the number of monitored pollutants. The Emission Compliance Index is defined as follows. Table 5 presents the calculation of ECI based on mean pollutant concentrations and regulatory limits.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

Table 5. Emission Compliance Index calculation

Pollutant	Mean concentration C_i	Limit L_i	Normalized exceedance
PM2.5	34 μg/m³	25 μg/m³	0.36
NO2	48 ppb	40 ppb	0.20
SO2	12 ppb	20 ppb	0.00
CO	1.9 ppm	9 ppm	0.00

Using the conceptual inputs in Table 5, PM2.5 and NO2 contribute to the assumed compliance risk, whereas SO2 and CO remain below the selected scenario limits.

Based on these values, the illustrative ECI is calculated as:

$$ECI = 1 - \frac{0.36+0.20}{4} = 0.86.$$

The Air Quality Impact Score is computed as:

$$AQIS = \sum_{i=1}^{k} w_i \cdot \frac{C_i}{L_i}, \tag{13}$$

where w_i denotes the weight assigned to pollutant *i* based on regulatory priority and health impact severity. The weights are normalized such that $\Sigma_{i=1}^{k} w_i = 1$.

Table 6. Air Quality Impact Score calculation

Pollutant	Weight w_i	C_i / L_i	Contribution
PM2.5	0.40	1.36	0.54
NO2	0.30	1.20	0.36
SO2	0.20	0.60	0.12
CO	0.10	0.21	0.02

The illustrative AQIS equals 1.04. Because the value exceeds unity, the conceptual baseline scenario represents elevated environmental pressure and demonstrates the need for project-level mitigation.

The Sustainable Project Performance Score integrates regulatory compliance, environmental impact, and project readiness:

$$SPPS = \alpha \cdot ECI + \beta \cdot (1 - AQIS) + \gamma \cdot R \tag{14}$$

where α, β, and γ are weighting coefficients reflecting the relative importance of regulatory compliance, impact mitigation, and project readiness, respectively, and R denotes mitigation readiness on a normalized scale from 0 to 1. In this study, the coefficients were set to $\alpha = 0.4$, $\beta = 0.4$, and $\gamma = 0.2$ based on expert judgment and ESG reporting practices.

Table 7. Sustainable Project Performance Score under different scenarios

Scenario	ECI	AQIS	R	SPPS
Baseline	0.86	1.04	0.60	0.76
30% emission reduction	0.93	0.82	0.75	0.89
Construction peak	0.71	1.25	0.40	0.62

DOI: 10.37943/PEXM7788

The conceptual SPPS calculation demonstrates how the same decision rule can compare project alternatives. It does not represent a formally audited ESG rating of Promanalit LLP or another industrial enterprise.

Results of Project Decision Support. The empirical forecasting and platform-pilot results provide the data and verification layers of the framework. The ECI, AQIS, SPPS, spatial heatmaps, and scenario comparisons provide a conceptual demonstration of how those layers may support planning and risk decisions.

Table 8 summarizes conceptual digital-twin outputs under the three assumed alternatives.

In the conceptual baseline scenario, elevated peak values and two illustrative hotspots correspond to a high regulatory-risk category.

The conceptual 30% mitigation scenario reduces peak concentrations, decreases the illustrative hotspot count, and changes the risk category from high to medium.

The conceptual construction scenario increases peak concentrations and the illustrative hotspot count, producing a critical risk category and indicating the possible need for schedule adjustments or temporary controls.

Figure 2 presents conceptual heatmaps of the assumed PM2.5 distribution under the three project alternatives. The maps illustrate decision-support functionality and are not georeferenced reconstructions of the confidential industrial site.

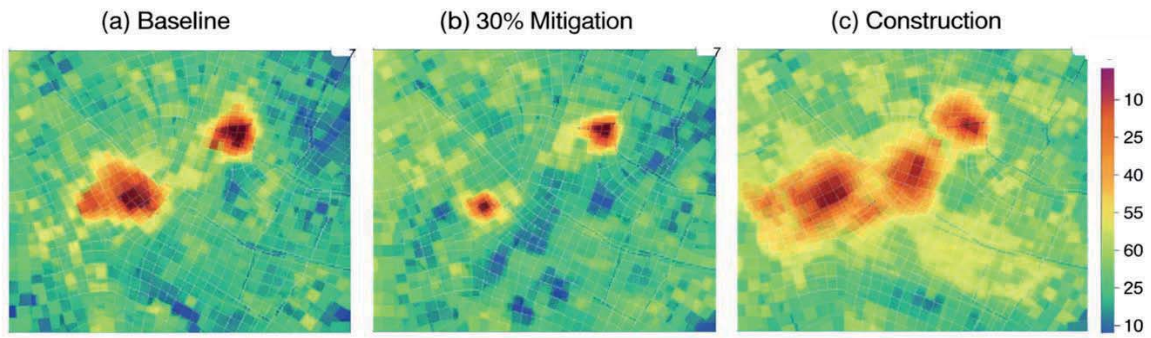

Figure 2. Digital Twin–based heatmaps illustrating the spatial distribution of $PM_{2.5}$ concentrations under (a) baseline conditions, (b) 30% emission mitigation, and (c) construction-related emission scenarios. The heatmaps support scenario-based project decision-making by highlighting spatial risk patterns and the effectiveness of mitigation measures.

Under the assumed mitigation scenario, the high-concentration area contracts; under the construction scenario, it expands. These patterns are illustrative outputs of the scenario model.

The comparative analysis of heatmaps enables project teams to visually assess the spatial effectiveness of mitigation measures and to prioritize targeted management actions in high-risk zones prior to project execution. Such spatial decision support is particularly valuable for project managers, as it facilitates geographically focused interventions rather than uniform and resource-intensive mitigation strategies.

Taken together, the empirical forecasting results, real platform pilot, retained mathematical formulations, and explicitly labelled conceptual calculations form a consistent three-level validation structure.

Table 8. Digital Twin simulation outcomes

Scenario	**$PM_{2.5}$ peak**	**NO_2 peak**	**Hotspots detected**	**Regulatory risk**
Baseline	58 µg/m³	62 ppb	2	High
30% mitigation	46 µg/m³	51 ppb	1	Medium
Construction	67 µg/m³	71 ppb	3	Critical

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

Discussion

The revised framework distinguishes three evidence levels. First, the LSTM forecasting results are empirical and are based on 39,803 synchronized industrial observations. Second, the Promanalit pilot provides operational evidence for telemetry ingestion, visualization, reporting, and blockchain verification. Third, the ECI, AQIS, SPPS, heatmaps, and mitigation/construction alternatives are conceptual scenario calculations intended to demonstrate project decision-support logic.

This separation prevents conceptual scenario outputs from being interpreted as measured industrial effects. Compared with forecasting-only studies, the framework connects predictive outputs to management indicators. Compared with monitoring-only systems, it adds a prospective scenario layer. Compared with blockchain-only solutions, it uses verified data as an input to environmental risk and ESG-oriented decisions.

Continuous monitoring combined with predictive analytics allows environmental information to directly support key project decisions. Unlike traditional project management approaches, where environmental assessment is typically static and performed at predefined milestones, the proposed system embeds environmental intelligence throughout the entire project lifecycle. Forecasting outputs and digital twin simulations provide early signals of potential regulatory and environmental risks, thereby supporting anticipatory and preventive management actions.

The integration of a micro-block blockchain architecture plays a critical role in strengthening trust and transparency. By ensuring the immutability and auditability of environmental data, the system enhances stakeholder confidence, facilitates regulatory compliance, and improves the credibility of ESG reporting. This trust layer extends the value of environmental intelligence beyond internal decision-making and supports accountability toward external stakeholders.

From a project management perspective, the system supports several advanced governance mechanisms. First, ESG indicators are dynamically integrated into operational decision-making rather than assessed retrospectively. Second, scenario-based simulations enable meaningful evaluation of alternative project actions before implementation. Third, sustainability performance can be compared across projects, supporting portfolio-level analysis and optimization. Finally, the continuous feedback provided by the system facilitates organizational learning, shifting sustainability assessment from a one-time evaluation toward an ongoing improvement process.

Limitations include confidentiality of the exact industrial sensor coordinates, dependence on calibration and data completeness, validation in a single operational monitoring context, and the conceptual status of the spatial and ESG scenario outputs. The scenario layer requires site-specific dispersion calibration and comparison with measured intervention outcomes before deployment as an operational industrial digital twin.

Conclusion

This paper presented an integrated framework combining real industrial telemetry, short-term LSTM forecasting, blockchain verification, and conceptual digital-twin decision support. Unlike the previous version, the revised manuscript states the exact analytical data volume, 20-minute temporal resolution, 72-step input window, one-step forecasting horizon, chronological 64/16/20 split, sensor context, and model hyperparameters.

The empirical LSTM experiment achieved MSE = 0.87 and R2 = 0.86 and reduced pollutant-specific RMSE by 33.1–46.8% relative to ARIMA. These are the real forecasting results reported in the study.

The ECI, AQIS, SPPS, heatmaps, and baseline/mitigation/construction calculations are retained as a conceptual demonstration of sustainability-oriented project decision support. They are now explicitly separated from the empirical forecasting data and the Promanalit platform pilot.

A key contribution of the study is the connection between environmental monitoring and project governance through ESG-oriented indicators. The Emission Compliance Index, Air

Quality Impact Score, and Sustainable Project Performance Score translate pollutant concentrations, compliance pressure, environmental impact, and mitigation readiness into interpretable management metrics. This allows environmental information to be used not only for reporting, but also for scenario comparison, corrective action planning, and sustainability-oriented decision-making.

Future work will validate the scenario module against measured dispersion data and documented mitigation actions at a fully characterized industrial facility, subject to enterprise disclosure and data-sharing requirements.

References

[1] Sharipova, S., Abzhanova, D., Toxanov, S., Biloshchytskyi, A. (2025). Development of a neural network-based module for forecasting atmospheric pollutant emissions. International Journal of Information and Communication Technologies. https://doi.org/10.37943/24WXPA8545

[2] Abzhanova, D., Biloshchytskyi, A. (2025). Adaptive Packetization Model (AABF+) and Microblocks for an Intelligent Atmospheric Emissions Monitoring System on a Consortium Blockchain. Information, 16, 976. https://doi.org/10.3390/info16110976

[3] Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network. Physica D: Nonlinear Phenomena, 404, 132306. https://doi.org/10.1016/j.physd.2019.132306

[4] Tao, Q., Liu, F., Li, Y., Sidorov, D. (2019). Air pollution forecasting using a deep learning model based on 1D CNN and bidirectional GRU. IEEE Access, 7, 76690–76698. https://doi.org/10.1109/ACCESS.2019.2921578

[5] Boje, C., Guerriero, A., Kubicki, S., Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. Automation in Construction, 114, 103179. https://doi.org/10.1016/j.autcon.2020.103179

[6] Fuller, A., Fan, Z., Day, C., Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. IEEE Access, 8, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358

[7] Casino, F., Dasaklis, T. K., Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. Telematics and Informatics, 36, 55–81. https://doi.org/10.1016/j.tele.2018.11.006

[8] Dinh, T. T. A., Wang, J., Chen, G., Liu, R., Ooi, B. C., Tan, K.-L. (2017). BLOCKBENCH: A framework for analyzing private blockchains. Proceedings of the 2017 ACM International Conference on Management of Data, 1085–1100. https://doi.org/10.1145/3035918.3064033

[9] Luo, Y., Shen, J., Liang, H., Sun, L., Dong, L. (2024). Supporting building life cycle carbon monitoring, reporting and verification: A traceable and immutable blockchain-empowered information management system and application in Hong Kong. Resources, Conservation and Recycling, 208, 107736. https://doi.org/10.1016/j.resconrec.2024.107736

[10] Ramadan, M. N. A., Ali, M. A. H., Jaber, H., Alkhedher, M. (2025). Blockchain-secured IoT-federated learning for industrial air pollution monitoring: A mechanistic approach to exposure prediction and environmental safety. Ecotoxicology and Environmental Safety, 300, 118442. https://doi.org/10.1016/j.ecoenv.2025.118442

[11] Zhang, B., Rong, Y., Yong, R., Qin, D., Li, M., Zou, G., & Pan, J. (2022). Deep learning for air pollutant concentration prediction: A review. Atmospheric Environment, 290, 119347. https://doi.org/10.1016/j.atmosenv.2022.119347

[12] Khan, M., Serafeim, G., Yoon, A. (2016). Corporate sustainability: First evidence on materiality. The Accounting Review, 91(6), 1697–1724. https://doi.org/10.2308/accr-51383

[13] ISO. (2018). ISO 31000:2018 Risk management — Guidelines. International Organization for Standardization. https://www.iso.org/standard/65694.html

[14] ISO. (2018). ISO 14064-1:2018 Greenhouse gases — Part 1: Specification with guidance at the organization level for quantification and reporting of greenhouse gas emissions and removals. International Organization for Standardization. https://www.iso.org/standard/66453.html

[15] World Health Organization. (2021). WHO global air quality guidelines: particulate matter, ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization.

[16] Government of the Republic of Kazakhstan. (2021). Environmental Code of the Republic of Kazakhstan, No. 400-VI.

[17] Ministry of Ecology, Geology and Natural Resources of the Republic of Kazakhstan. (2021). Rules for maintaining an automated environmental-emissions monitoring system during industrial environmental control, Order No. 208.

[18] Abzhanova, D., Biloshchytskyi, A. (2026). A model and method for managing data on emissions from stationary sources of pollution in an intelligent environmental monitoring system. International Journal of Information and Communication Technologies, 7(2).

[19] OpenAQ. (2025). Open Air Quality Data Platform. https://openaq.org

[20] TensorFlow Developers. (2025). TensorFlow documentation: Keras recurrent neural networks. https://www.tensorflow.org

[21] Scikit-learn Developers. (2025). Scikit-learn user guide. https://scikit-learn.org

[22] United Nations. (2015). Paris Agreement. United Nations Framework Convention on Climate Change. https://unfccc.int/process-and-meetings/the-paris-agreement