Check for updates

OPEN ACCESS

EDITED BY Fei Li, Zhongnan University of Economics and Law, China

REVIEWED BY Ricardo Valentim, Federal University of Rio Grande do Norte, Brazil Arindam Ray, Bill and Melinda Gates Foundation, United States Junxue Zhang, Jiangsu University of Science and Technology, China Hélder Tiago Da Silva Lopes, University of Minho, Portugal

★CORRESPONDENCE Lan Luan ♀ qo7ybo@163.com

RECEIVED 23 January 2025 ACCEPTED 07 May 2025 PUBLISHED 04 June 2025

CITATION

Luan L and Daoyu Z (2025) Development of deep learning models for high-resolution exposome mapping and health impact assessment. *Front. Public Health* 13:1565471. doi: 10.3389/fpubh.2025.1565471

COPYRIGHT

© 2025 Luan and Daoyu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development of deep learning models for high-resolution exposome mapping and health impact assessment

Lan Luan^{1*} and Zhu Daoyu²

¹College of Computer and Information Engineering, Guizhou University of Commerce, Guizhou, China, ²College of Physical Education, Xinyang Normal University, Xinyang, Henan, China

Introduction: The study of environmental health and the exposome is becoming increasingly vital as researchers aim to untangle the complex interactions between environmental exposures and human health outcomes. Traditional exposome mapping methods often face limitations, such as low spatial-temporal resolution, challenges in integrating multi-modal data sources, and inadequate handling of uncertainties in exposure quantification.

Methods: To address these gaps, we introduce an innovative framework that leverages advanced deep learning techniques, adaptive optimization strategies, and multi-scale data integration to achieve high-resolution exposome modeling. Central to our approach is the Adaptive Multi-Scale Exposure Network (AMSEN), a hierarchical deep learning model designed to harmonize diverse data streams, such as satellite imagery, wearable sensors, and geospatial analytics, while addressing the challenges of multi-scale variability and measurement uncertainties. AMSEN incorporates cross-modal fusion mechanisms, spatiotemporal feature extraction, and uncertainty quantification. Complementing AMSEN, the Adaptive Exposure Optimization Strategy (AEOS) enhances model efficiency and accuracy through dynamic resource allocation, uncertainty-guided refinement, and domain-specific prior enforcement.

Results: These methodologies significantly advance the capabilities of exposome research by providing a robust, adaptive, and high-resolution modeling framework.

Discussion: The experimental findings highlight the effectiveness of our approach, showcasing enhancements in exposure prediction precision, computational performance, and practical insights for public health policymaking. This work aligns with the objectives of advancing environmental health sciences by offering novel tools for exposome quantification and health impact assessment.

KEYWORDS

exposome mapping, high-resolution modeling, deep learning, multi-modal data integration, uncertainty quantification

1 Introduction

The study of the exposome, which encompasses the totality of environmental exposures an individual encounters throughout their lifetime, is essential for understanding the multifactorial causes of human diseases (1). High-resolution exposome mapping not only enables precise identification of exposure sources but also aids in quantifying their impact on health outcomes, thereby bridging the gap between environmental exposure and epidemiological studies (2). This task is complex, as it involves integrating diverse data modalities such as satellite imagery, air quality measurements, wearable sensors, and socio-demographic data (3). Traditional methods are often insufficient due to their limited ability to process heterogeneous and high-dimensional datasets, while newer approaches, including machine learning and deep learning, promise to overcome these limitations (4). The development of such models is not only critical for advancing environmental health research but also has far-reaching implications for public health policies, personalized healthcare, and urban planning (5). There is a growing need to explore how computational models can accurately characterize the exposome, assess health risks, and provide actionable insights.

Early computational efforts in exposome research focused on structuring domain-specific knowledge to support exposure assessment (6). Researchers developed formalized systems that encoded expert insights to interpret environmental indicators, enabling rule-based inferences from spatial and temporal datasets (7). These strategies facilitated initial progress in linking environmental contexts with health outcomes and proved useful for targeted exposure estimation (8). However, their rigidity and dependence on manually constructed frameworks limited their capacity to adapt to complex, large-scale, and dynamic data environments. As exposome-related datasets grew in scope and granularity, such limitations became increasingly apparent, revealing the need for more adaptive and scalable solutions (9, 10).

In response to these challenges, adaptive modeling techniques were introduced to better accommodate the expanding landscape of exposome data (11). By leveraging statistical learning frameworks, researchers began to construct predictive models capable of uncovering hidden associations between environmental indicators and health risks (12). Algorithms such as support vector machines, random forests, and gradient boosting machines enhanced the accuracy and efficiency of exposure estimates by integrating diverse inputs from sensors, imaging, and demographic databases (13). These models represented a substantial advancement, yet they often required extensive tuning and expert-guided feature design (14). Moreover, understanding the causal mechanisms behind their predictions remained difficult, prompting concerns about interpretability and trustworthiness in health-related applications (15).

Recent developments in representation learning have further advanced the field by enabling automatic extraction of relevant features from large, complex datasets (16). Neural network architectures, including convolutional and recurrent models, have demonstrated strong capabilities in processing multidimensional information such as geospatial images, time series, and mobility patterns (17). Emerging transformer-based frameworks and pre-trained models like BERT and GPT have also been adapted to handle multimodal exposure data, offering improved capacity to model intricate relationships among environmental and biological variables (18). These models have achieved notable performance in fine-grained exposure estimation and health outcome prediction (19). Nonetheless, challenges such as generalizability across populations, the need for large annotated datasets, and the opaque nature of model outputs continue to present barriers to their widespread adoption in exposome research (20).

Based on the limitations of prior approaches, we propose a novel methodology for high-resolution exposome mapping and health impact assessment using advanced deep learning architectures. Our method addresses the challenges of data heterogeneity, scalability, and interpretability by incorporating multi-modal data fusion techniques, self-supervised learning for data-scarce environments, and explainable AI (XAI) frameworks. By leveraging state-of-the-art models tailored for environmental health research, we aim to improve the accuracy, efficiency, and transparency of exposome analyses. This approach not only builds upon the strengths of existing methods but also mitigates their weaknesses by prioritizing robustness, adaptability, and user interpretability. In doing so, our method has the potential to redefine the landscape of exposome research and its applications in public health.

- Our approach integrates cutting-edge multi-modal data fusion techniques and self-supervised learning to handle heterogeneous and incomplete datasets effectively.
- The proposed models are designed to perform efficiently across diverse environments and data sources, ensuring wide applicability and scalability.
- Experimental results show significant improvements in the accuracy of exposome mapping and health risk predictions, offering actionable insights for public health interventions.

2 Related work

2.1 Deep learning in exposome mapping

Deep learning has been increasingly employed in exposome research, particularly for mapping complex environmental exposures (21). The exposome, which encompasses the totality of environmental exposures an individual encounters throughout their lifetime, is inherently multidimensional and dynamic (22). Capturing this complexity requires models capable of handling high-dimensional, spatiotemporal, and heterogeneous data. Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have been applied to model various components of the exposome, such as air pollution, noise levels, and green space distributions. These models are well-suited for processing large-scale environmental datasets that often come in the form of satellite imagery, geospatial data, and time-series measurements. Recent advances in satellite imaging and sensor technology have contributed to the availability of high-resolution environmental data, enabling deep learning applications in fine-scale exposome mapping (23). CNNs, for example, have been extensively used for extracting spatial features from remote sensing data. These models have demonstrated superior performance in predicting air pollution concentrations, such as PM_{2.5} and NO₂, by integrating meteorological data, land-use features, and population density (24). Advanced architectures like U-Nets have been adopted for pixel-wise prediction tasks, such as land cover classification and vegetation analysis, providing valuable inputs for exposome studies. In addition, transfer learning has been used to leverage pre-trained models on large image datasets to improve the performance of domain-specific tasks with limited labeled data. Temporal dynamics are also critical in exposome mapping,

as environmental exposures often vary over time (25) (RNNs) and long short-term memory (LSTM) networks have demonstrated their potential in modeling temporal dependencies within exposure data. These approaches have been applied to predict air quality indices and track seasonal trends in environmental pollutants. The integration of temporal and spatial information has led to the development of spatiotemporal deep learning models, which combine CNNs and RNNs to address the dynamic nature of the exposome. These hybrid architectures are particularly effective for applications requiring both fine spatial resolution and temporal continuity. Despite these advances, challenges remain in the application of deep learning to exposome mapping (26). One key issue is the lack of standardized datasets and benchmarks, which limits the comparability of model performance across studies. Another challenge is the interpretability of deep learning models, as the black-box nature of these algorithms makes it difficult to derive mechanistic insights into the relationships between environmental exposures and health outcomes. Explainable AI techniques, such as attention mechanisms and feature importance analyses, have been proposed to address this issue. The fusion of multi-modal data, such as integrating satellite imagery with wearable sensor measurements, presents computational and methodological hurdles that warrant deeper investigation.

2.2 Health impact assessment using deep learning

Deep learning has emerged as a powerful tool for assessing the health impacts of environmental exposures (27). By leveraging large-scale health datasets and advanced neural network architectures, researchers have made significant progress in understanding the complex interactions between the exposome and human health. These models have been applied to a wide range of health outcomes, including respiratory diseases, cardiovascular conditions, and mental health disorders. One of the primary applications of deep learning in health impact assessment is in the prediction of disease risk based on environmental exposure data (28). For example, neural networks have been used to model the relationship between air pollution and respiratory diseases, leveraging spatial and temporal data on pollutant concentrations. CNNs have been particularly effective in capturing spatial patterns of exposure, while LSTMs have been used to account for temporal trends in health outcomes. These models often outperform traditional statistical approaches, such as generalized linear models, by capturing non-linear and complex interactions between variables (29). Deep learning models have been used for causal inference in health impact studies (30). Techniques such as deep reinforcement learning and generative adversarial networks (GANs) have been explored for simulating counterfactual scenarios, enabling researchers to estimate the causal effects of environmental interventions (31). For instance, GANs have been employed to generate synthetic data for underrepresented populations, improving the generalizability of health impact assessments. Attention mechanisms and explainability frameworks have been integrated into these models to enhance their interpretability, facilitating their application in policy-making and public health interventions (32). Another important area of research is the integration of genetic and epigenetic data into deep learning models for exposome-health studies (33). Multi-modal deep learning approaches have been developed to combine environmental, genetic, and clinical data, providing a comprehensive understanding of gene-environment interactions. These models have been used to investigate how environmental exposures influence epigenetic modifications and, in turn, contribute to disease risk. For example, deep autoencoders have been applied to identify patterns in DNA methylation data associated with air pollution exposure, shedding light on potential biological pathways underlying health effects (34). Challenges persist in the application of deep learning to health impact assessment. One major limitation is the availability of high-quality, large-scale datasets that integrate environmental exposures and health outcomes. Data privacy concerns and ethical considerations further complicate data sharing and integration (35). Another challenge is the need for robust validation frameworks to ensure the reliability and reproducibility of deep learning models in diverse populations and settings. Addressing these challenges will require collaborative efforts across disciplines, including environmental science, epidemiology, and computer science.

2.3 Spatiotemporal modeling for environmental health

Spatiotemporal modeling has emerged as a key area of interest in applying deep learning techniques to environmental health studies (36). The dynamic and spatially heterogeneous nature of environmental exposures necessitates the use of models that can capture both spatial and temporal variations in data. Deep learning models, such as spatiotemporal convolutional networks and graph neural networks (GNNs), have shown great potential in this regard, enabling researchers to analyze complex patterns in environmental and health datasets. Spatiotemporal convolutional networks, which combine convolutional operations with temporal processing, have been widely used for exposure prediction and monitoring (37). These models can capture fine-grained spatial patterns while accounting for temporal dependencies in data. For instance, spatiotemporal deep learning has been employed to model urban air quality, combining inputs such as traffic data, meteorological variables, and satellite imagery. These models are capable of predicting pollutant concentrations at high spatial resolutions and short time intervals, providing critical insights for environmental health studies. Graph neural networks (GNNs) have also gained attention for their ability to model relational data and spatial dependencies in environmental health research (38). GNNs represent data as graphs, with nodes corresponding to spatial locations and edges capturing spatial or functional relationships. This approach has been used to model disease transmission networks, assess the spread of infectious diseases, and evaluate the spatial distribution of environmental exposures. GNNs are particularly useful for integrating data from diverse sources, such as combining satellite-derived air quality estimates with social determinants of health. Another area of advancement

is the use of attention mechanisms in spatiotemporal modeling (39). Attention mechanisms allow models to concentrate on the most pertinent features or areas within the data, enhancing both their interpretability and effectiveness. For example, attentionbased spatiotemporal models have been applied to identify hotspots of environmental exposures and their associated health impacts. These models can provide actionable insights for public health interventions by highlighting areas with the greatest need for environmental mitigation efforts. Despite these advancements, spatiotemporal modeling in environmental health faces several challenges (40). One key issue is the computational complexity of processing high-resolution spatiotemporal data, which often requires significant computational resources and expertise in parallel processing. Another challenge is the scalability of these models to large geographic areas and diverse populations, as well as the need for robust methods to handle missing or noisy data. Future research efforts should focus on developing efficient algorithms and scalable architectures to address these limitations. Interdisciplinary collaborations will be essential for advancing the application of spatiotemporal deep learning models in environmental health research.

3 Method

3.1 Overview

The exposome represents the comprehensive set of environmental exposures an individual encounters throughout their lifetime, encompassing chemical, physical, and social factors. High-resolution exposome research aims to systematically capture and analyze these exposures at an unprecedented spatial and temporal granularity, bridging the gap between environment-wide exposure assessments and individual-level health outcomes. This paper introduces a novel methodological framework to enhance the precision and scalability of exposome quantification through the integration of advanced computational models, multi-modal data fusion, and adaptive optimization strategies. Our approach is organized into three key components: In Section 3.2, we present the foundational preliminaries required to formalize the problem of high-resolution exposome modeling. This involves defining the mathematical representation of exposures, their spatiotemporal characteristics, and the associated uncertainties. In Section 3.3, we describe a newly developed model designed to effectively capture the complex, multi-scale dynamics of environmental exposures. This model incorporates hierarchical structures and integrates diverse datasets, such as remote sensing imagery, personal monitoring data, and geospatial analytics. In Section 3.6, we propose an innovative strategy that leverages domainspecific knowledge and optimization principles to adaptively refine the exposure quantification process. This strategy enhances computational efficiency while maintaining high accuracy across varying exposure scenarios.

Central to our framework is the seamless fusion of data streams from heterogeneous sources, ranging from high-resolution satellite imagery to wearable sensors. This integration is achieved through a robust pipeline that standardizes, harmonizes, and preprocesses multi-modal data, ensuring consistency and compatibility for downstream analysis. Our approach incorporates novel statistical techniques to account for uncertainties and missing data, thus enabling more reliable inferences about exposure profiles. The overarching goal of our method is to bridge the traditional limitations of exposome research by introducing a scalable, adaptive, and high-resolution framework. Through this work, we aim to empower researchers and policymakers with actionable insights into the intricate interactions between environmental factors and health outcomes, ultimately advancing the understanding of environmental determinants of health.

3.2 Preliminaries

High-resolution exposome modeling involves capturing the dynamic and multi-dimensional nature of environmental exposures across both spatial and temporal dimensions. To formalize this problem, we begin by introducing the key notations, definitions, and mathematical constructs that will serve as the foundation for the proposed framework.

Let the exposome be represented as a set $\mathcal{E} = \{e_1, e_2, \dots, e_N\}$, where each e_i denotes a specific environmental exposure. Each exposure e_i is modeled as a time-dependent and spatially distributed variable defined over a domain $\mathcal{D} \subseteq \mathbb{R}^2 \times \mathbb{T}$, where \mathbb{R}^2 represents the geographical space and \mathbb{T} represents time. Thus, e_i can be expressed as:

$$e_i(x,t): \mathcal{D} \to \mathbb{R},\tag{1}$$

where $x \in \mathbb{R}^2$ is the spatial location, $t \in \mathbb{T}$ is the time index, and $e_i(x, t)$ provides the exposure intensity of e_i at location x and time t.

Environmental exposures inherently exhibit variability across multiple spatial and temporal scales. To capture this, we introduce a multi-scale decomposition for $e_i(x, t)$:

$$e_i(x,t) = \sum_{l=1}^{L} \sum_{h=1}^{H} e_i^{(l,h)}(x,t),$$
(2)

where $e_i^{(l,h)}(x, t)$ represents the contribution of exposure e_i at spatial scale l and temporal scale h. Here, L denotes the number of spatial scales, and H denotes the number of temporal resolutions.

The observed data for exposures often come from multiple sources, such as satellite remote sensing, we arable devices, and fixed monitoring stations. Let \mathcal{O} denote the set of observed data streams, where each $\mathcal{O}_k \in \mathcal{O}$ corresponds to a particular data modality. The observed exposure e_i^{obs} is then related to the true exposure e_i through a measurement model:

$$e_i^{\text{obs}}(x,t) = \mathcal{H}_k(e_i(x,t)) + \eta_k(x,t), \tag{3}$$

where \mathcal{H}_k is the observation operator for the *k*-th data modality, and $\eta_k(x, t)$ represents the noise associated with the measurement process. Different observation operators capture the resolution and biases unique to the modality.

Given the heterogeneity of data sources, it is essential to account for uncertainties in the measurements. Let U represent the uncertainty associated with the exposure. For a given exposure

 e_i , we define its uncertainty as a probability distribution $\mathbb{P}(e_i|\mathcal{O})$, which can be derived using Bayesian principles:

$$\mathbb{P}(e_i(x,t)|\mathcal{O}) \propto \mathbb{P}(\mathcal{O}|e_i(x,t)) \cdot \mathbb{P}(e_i(x,t)),$$
(4)

where $\mathbb{P}(\mathcal{O}|e_i(x, t))$ is the likelihood of the observations given the true exposure, and $\mathbb{P}(e_i(x, t))$ represents the prior knowledge about the exposure distribution.

The task of high-resolution exposome modeling is to infer $e_i(x, t)$ from the observed data \mathcal{O} by solving an optimization problem that minimizes the reconstruction error while accounting for uncertainties:

$$\hat{e}_i(x,t) = \arg\min_{e_i} \left[\sum_{k \in \mathcal{O}} \|\mathcal{H}_k(e_i(x,t)) - e_i^{\text{obs}}(x,t)\|_2^2 + \lambda \mathcal{R}(e_i(x,t)) \right],$$
(5)

where $\mathcal{R}(e_i(x, t))$ is a regularization term that incorporates spatial or temporal smoothness constraints, and λ is a hyperparameter controlling the regularization strength.

3.3 Adaptive multi-scale exposome network (AMSEN)

To address the challenges inherent in high-resolution exposome modeling, we propose the Adaptive Multi-Scale Exposome Network (AMSEN), a novel computational framework designed to capture the complex, multi-scale dynamics of environmental exposures. AMSEN is built on a hierarchical architecture that integrates spatial, temporal, and cross-modal information from diverse data sources to achieve robust exposure quantification (As shown in Figure 1).

3.3.1 Multi-scale representation

Environmental exposures exhibit variability at multiple spatial and temporal scales. AMSEN incorporates a hierarchical representation to model this variability and captures the complex dynamics of exposure interactions. The spatial domain \mathbb{R}^2 is discretized into a set of *L* nested grids $\{\mathcal{G}_l\}_{l=1}^L$, where each \mathcal{G}_l represents a resolution level. Each grid \mathcal{G}_l partitions the spatial domain into cells of size $\Delta x_l \times \Delta y_l$, where the resolution becomes finer with increasing *l*. This hierarchical spatial discretization enables the model to encode both coarse-grained and fine-grained spatial structures.

The temporal domain \mathbb{T} is discretized into a set of H nested intervals $\{\mathcal{T}_h\}_{h=1}^H$, corresponding to different temporal resolutions. Each temporal interval \mathcal{T}_h is defined by its duration Δt_h , where finer temporal resolutions correspond to smaller Δt_h . This nested structure enables AMSEN to adapt to both short-term variations and long-term trends in environmental exposures.

For each exposure e_i , AMSEN constructs a multi-scale representation $e_i^{(l,h)}(x,t)$, which captures the exposure intensity at spatial scale l and temporal scale h. The overall exposure representation is defined as a sum over all scales:

$$e_i(x,t) = \sum_{l=1}^{L} \sum_{h=1}^{H} e_i^{(l,h)}(x,t),$$
(6)



accurate reconstruction across spatial and temporal scales.

where $e_i^{(l,h)}(x,t)$ represents the contribution of exposure e_i at spatial resolution l and temporal resolution h. To construct $e_i^{(l,h)}(x,t)$, the model applies a series of operations over spatial and temporal neighborhoods.

The spatial contribution is computed using convolutional operations on the grid G_l , where the receptive field increases with *l*. Let K_l represent the convolutional kernel for scale *l*, and let $N_l(x)$ denote the spatial neighborhood of location *x* on grid G_l . Then, the spatial feature extraction is given by:

$$e_i^{(l)}(x,t) = \sum_{x' \in N_l(x)} K_l(x,x') \cdot e_i(x',t),$$
(7)

where $K_l(x, x')$ encodes the spatial relationship between x and x' at scale *l*. This operation aggregates information from the spatial neighborhood and encodes it into the multi-scale representation.

For the temporal contribution, a similar approach is applied using temporal kernels K_h over the temporal intervals \mathcal{T}_h . Let $N_h(t)$ denote the temporal neighborhood of time *t* within interval \mathcal{T}_h . The temporal feature extraction is given by:

$$e_i^{(h)}(x,t) = \sum_{t' \in N_h(t)} K_h(t,t') \cdot e_i(x,t').$$
(8)

3.4 Cross-modal data fusion

AMSEN integrates diverse data sources, including satellite imagery, wearable sensors, and monitoring stations, by constructing modality-specific feature extraction pathways. Let \mathcal{O}_k denote the observation from modality k. AMSEN uses a shared encoder-decoder architecture to extract features $\Phi_k(x, t)$ from each modality:

$$\Phi_k(x,t) = \mathcal{F}_k(e_i^{\text{obs}}(x,t);\theta_k), \tag{9}$$

where \mathcal{F}_k is the modality-specific feature extractor parameterized by θ_k , and $e_i^{\text{obs}}(x, t)$ represents the input signal for modality k at spatial location x and temporal instance t. These extracted features $\Phi_k(x, t)$ encode modality-specific information while preserving the unique characteristics of the corresponding data source.

The features extracted from all modalities are subsequently fused in a shared latent space \mathcal{Z} . The fusion step combines information across multiple modalities to achieve a comprehensive representation:

$$\mathcal{Z}(x,t) = \operatorname{Fuse}(\Phi_1(x,t), \Phi_2(x,t), \dots, \Phi_K(x,t)), \quad (10)$$

where *K* represents the number of modalities, and $Fuse(\cdot)$ is the fusion operator. This operator can take various forms, such as concatenation:

$$Fuse(\Phi_1, \Phi_2, \dots, \Phi_K) = Concat(\Phi_1, \Phi_2, \dots, \Phi_K), \quad (11)$$

or attention mechanisms for weighted feature integration:

Fuse
$$(\Phi_1, \Phi_2, \dots, \Phi_K) = \sum_{k=1}^K \alpha_k \cdot \Phi_k,$$
 (12)

where α_k are the attention weights learned to prioritize the contribution of each modality.

The fused latent representation $\mathcal{Z}(x, t)$ serves as the input for reconstructing the high-resolution exposure field. The reconstruction is performed by a decoder \mathcal{G} parameterized by ψ , which maps \mathcal{Z} to the predicted exposure field:

$$\hat{e}_i(x,t) = \mathcal{G}(\mathcal{Z}(x,t);\psi).$$
(13)

The decoder G employs a multi-layer architecture to refine and upscale the fused latent representation $\mathcal{Z}(x, t)$, ensuring accurate reconstruction of spatial and temporal exposure patterns.

3.5 Uncertainty quantification

AMSEN explicitly quantifies uncertainty in exposure estimates by introducing a probabilistic modeling layer. Let $\hat{e}_i(x, t)$ represent the predicted exposure field, which is subject to inherent variability and uncertainty due to factors such as sensor noise, data sparsity, and environmental dynamics (As shown in Figure 2).

To characterize the uncertainty associated with the predictions, AMSEN models the variance $\sigma_i^2(x, t)$ as a function of the spatialtemporal context, described by:

$$\sigma_i^2(x,t) = \operatorname{Var}[\hat{e}_i(x,t)] = \mathcal{H}(\mathcal{Z}(x,t);\phi), \tag{14}$$



Diagram of uncertainty quantification framework, illustrating the integration of an uncertainty mapping function, normalization, and Inception Depthwise Convolution modules to process spatial-temporal attributes and predict exposure fields with associated uncertainty estimates.

where \mathcal{H} is a learned uncertainty mapping function parameterized by ϕ . The input to \mathcal{H} is $\mathcal{Z}(x, t)$, a feature representation encoding the spatial-temporal attributes of the environment, which could include meteorological data, geographic features, and sensor characteristics.

To ensure that the uncertainty estimates are robust and meaningful, AMSEN employs a probabilistic framework during model training. The predictive distribution of $\hat{e}_i(x, t)$ is assumed to follow a Gaussian distribution:

$$\hat{e}_i(x,t) \sim \mathcal{N}(\mu_i(x,t), \sigma_i^2(x,t)), \tag{15}$$

where $\mu_i(x, t)$ is the predicted mean exposure field and $\sigma_i^2(x, t)$ is the predicted variance representing the uncertainty. The model is trained by maximizing the likelihood of the observed data under this probabilistic model, which is equivalent to minimizing the negative log-likelihood (NLL) loss:

$$\mathcal{L}_{\text{NLL}} = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{(y_i - \mu_i(x, t))^2}{2\sigma_i^2(x, t)} + \frac{1}{2} \log \sigma_i^2(x, t) \right],$$
(16)

where y_i represents the ground truth exposure measurements, and N is the total number of observations. This loss function encourages the model to accurately predict both the mean and variance of the exposure field while penalizing overconfident or underconfident uncertainty estimates.

In addition, AMSEN incorporates regularization mechanisms to ensure stable training of the uncertainty mapping. For instance, a prior constraint can be imposed on $\sigma_i(x, t)$ to prevent pathological cases where the uncertainty collapses to zero or diverges to infinity. A commonly used regularization term is:

$$\mathcal{L}_{\text{reg}} = \lambda \cdot \mathbb{E}[\sigma_i^2(x, t)], \qquad (17)$$

where λ is a hyperparameter controlling the strength of the regularization.

3.6 Adaptive exposure optimization strategy (AEOS)

To complement the proposed Adaptive Multi-Scale Exposome Network (AMSEN), we introduce the Adaptive Exposure Optimization Strategy (AEOS), a novel approach designed to dynamically refine exposure estimation by leveraging domainspecific knowledge, spatiotemporal regularities, and adaptive computational principles. AEOS enhances the robustness, interpretability, and scalability of AMSEN, ensuring its applicability across diverse real-world scenarios (As shown in Figure 3).

3.6.1 Incorporating domain knowledge

Environmental exposure modeling often benefits significantly from integrating domain-specific knowledge, such as regulatory limits, spatial boundaries, pollutant dispersion dynamics, and chemical transformation rates. This knowledge provides meaningful priors that guide the model toward physically plausible and scientifically consistent solutions. In AEOS, these priors are incorporated as soft constraints during the optimization process to balance empirical observations with domain-based expectations.

For each exposure e_i , we define a prior distribution $\mathbb{P}(e_i)$ based on domain knowledge. The prior distribution imposes a probabilistic structure over the exposure field, which can be represented as:

$$\mathbb{P}(e_i(x,t)) \propto \exp\big(-\mathcal{C}(e_i(x,t))\big),\tag{18}$$

where $C(e_i(x, t))$ is a penalty function that quantifies deviations of $e_i(x, t)$ from known environmental constraints. Examples of such constraints include maximum allowable pollutant concentrations, pollutant dispersion patterns governed by physical laws, and decay rates under chemical or biological processes. For instance, $C(e_i(x, t))$ may include terms such as:

$$\mathcal{C}(e_i(x,t)) = \alpha \cdot \max(0, e_i(x,t) - e_{\max})^2 + \beta \cdot \|\nabla e_i(x,t) - \mathbf{v}(x,t) \cdot \nabla c\|^2,$$
(19)

where e_{max} is a regulatory limit for exposure, $\mathbf{v}(x, t)$ represents the local wind or flow vector field, and ∇c denotes the expected concentration gradient due to advection-diffusion dynamics. The parameters α and β control the relative weight of different penalty components.

These prior constraints are incorporated into AMSEN's optimization framework by introducing a regularization term in the loss function. The regularization term penalizes the model for deviating from the prior distribution:

$$\mathcal{L}_{\text{prior}} = \lambda_{\text{prior}} \int \mathcal{C}(e_i(x, t)) \, dx \, dt, \tag{20}$$

where λ_{prior} is a hyperparameter that determines the strength of the prior enforcement. A higher value of λ_{prior} results in stricter adherence to the domain knowledge, while a lower value allows the model to prioritize empirical data.

To ensure flexibility and adaptability, AMSEN uses learnable parameters within the prior penalty functions. For example, instead of fixing $\mathbf{v}(x, t)$, the model can learn an estimated flow field $\hat{\mathbf{v}}(x, t)$ that is consistent with observed exposure patterns. The penalty function then becomes:

$$\mathcal{C}(e_i(x,t)) = \alpha \cdot \max(0, e_i(x,t) - e_{\max})^2 + \beta \cdot \|\nabla e_i(x,t) - \hat{\mathbf{v}}(x,t) \cdot \nabla c\|^2.$$
(21)

Another important source of domain knowledge is pollutant decay over time. For many pollutants, the decay process can be approximated using first-order kinetics, represented as:

$$\frac{\partial e_i(x,t)}{\partial t} = -k \cdot e_i(x,t), \qquad (22)$$



where k is the decay rate constant. This physical law can be directly encoded into the prior penalty function by enforcing consistency with the temporal decay dynamics:

$$C_{\text{decay}}(e_i(x,t)) = \gamma \cdot \left\| \frac{\partial e_i(x,t)}{\partial t} + k \cdot e_i(x,t) \right\|^2, \quad (23)$$

where γ controls the strength of the decay consistency term.

3.6.2 Iterative refinement mechanism

AEOS incorporates an iterative refinement mechanism that adapts to the spatiotemporal heterogeneity of exposure patterns. Let $\mathcal{R}(x, t)$ represent a residual field that captures unresolved variability in exposure estimates:

$$\mathcal{R}(x,t) = e_i^{\text{obs}}(x,t) - \hat{e}_i(x,t), \qquad (24)$$

where $e_i^{\text{obs}}(x, t)$ is the observed exposure value, and $\hat{e}_i(x, t)$ is the model-predicted estimate at spatial location x and temporal instance t. The residual $\mathcal{R}(x, t)$ quantifies the discrepancy between the observed and predicted values, serving as the basis for iterative updates. At each iteration k, AEOS updates the exposure estimate $\hat{e}_i^{(k)}(x, t)$ to reduce the residual field. The update rule is defined as:

$$\hat{e}_i^{(k+1)}(x,t) = \hat{e}_i^{(k)}(x,t) + \alpha \,\mathcal{R}^{(k)}(x,t),\tag{25}$$

where α is a learning rate that controls the step size of refinement, and $\mathcal{R}^{(k)}(x,t)$ is the residual field computed at iteration *k*. The iterative process continues until the norm of the residual, $\|\mathcal{R}^{(k)}(x,t)\|_2$, falls below a predefined convergence threshold ϵ :

$$\|\mathcal{R}^{(k)}(x,t)\|_2 \le \epsilon.$$
(26)

AEOS incorporates an adaptive convergence criterion that varies spatially and temporally based on the complexity of exposure patterns. For regions with high variability, a stricter convergence threshold is applied to ensure accuracy:

$$\epsilon_j = \epsilon_0 + \beta \|\nabla e_i(x, t)\|_2, \tag{27}$$

where ϵ_0 is the baseline threshold, β is a scaling factor, and $\|\nabla e_i(x, t)\|_2$ represents the magnitude of the exposure gradient, which indicates local variability.

To improve computational efficiency, AEOS dynamically allocates computational resources based on exposure variability and data density. The spatial domain $\mathcal{D} \subseteq \mathbb{R}^2$ is partitioned into regions $\{\mathcal{D}_j\}_{j=1}^M$ based on exposure gradients:

$$\mathcal{D}_{j} = \{ x \in \mathcal{D} \mid \|\nabla e_{i}(x, t)\|_{2} \in [g_{j}, g_{j+1}) \},$$
(28)

where g_j represents the gradient threshold for region j, and M is the total number of regions. High-gradient regions, which exhibit rapid spatial changes in exposure, are assigned more computational resources to capture these local variations effectively.

The computational resource allocation for each region D_j is given by:

$$C_j \propto \|\nabla e_i(x,t)\|_2 + \rho \, d_j,\tag{29}$$

where C_j represents the computational budget assigned to region \mathcal{D}_j , ρ is a balancing factor, and d_j denotes the data density within \mathcal{D}_j . Regions with higher exposure gradients $\|\nabla e_i(x, t)\|_2$ or higher data density d_j are prioritized to ensure accurate and efficient modeling.

The iterative refinement mechanism is further enhanced through a multi-scale approach. Exposure estimates are refined across a hierarchy of spatial resolutions. Let $\ell \in \{1, 2, ..., L\}$ denote the resolution level, where finer levels capture detailed local patterns, and coarser levels capture broader trends. The refinement at each resolution is expressed as:

$$\hat{e}_i^{(\ell+1)}(x,t) = \hat{e}_i^{(\ell)}(x,t) + \alpha^{(\ell)} \,\mathcal{R}^{(\ell)}(x,t),\tag{30}$$

where $\alpha^{(\ell)}$ is the learning rate specific to resolution level ℓ , and $\mathcal{R}^{(\ell)}(x, t)$ is the residual field computed at that resolution.

3.6.3 Multi-scale consistency

AEOS enforces consistency across spatial and temporal scales by introducing a scale-harmonization term in the optimization process (As shown in Figure 4).

Let $e_i^{(l,h)}(x, t)$ denote the exposure estimate at spatial scale *l* and temporal scale *h*. The consistency term is defined as:

$$\mathcal{L}_{\text{scale}} = \sum_{l=1}^{L} \sum_{h=1}^{H} \left\| e_i^{(l,h)}(x,t) - \bar{e}_i^{(l,h)}(x,t) \right\|_2^2, \tag{31}$$

where $\bar{e}_i^{(l,h)}(x,t)$ is the aggregated estimate across neighboring scales. This aggregation is performed through a weighted averaging mechanism, incorporating both spatial and temporal dependencies. $\bar{e}_i^{(l,h)}(x,t)$ can be computed as:

$$\bar{e}_{i}^{(l,h)}(x,t) = \frac{\sum_{(l',h') \in \mathcal{N}(l,h)} w_{l',h'} e_{i}^{(l',h')}(x,t)}{\sum_{(l',h') \in \mathcal{N}(l,h)} w_{l',h'}},$$
(32)

where $\mathcal{N}(l, h)$ denotes the set of neighboring scales, and $w_{l',h'}$ are scale-dependent weights that reflect the relative importance of each neighboring scale. These weights are learned during training to adaptively balance contributions from coarse- and fine-scale information. By minimizing \mathcal{L}_{scale} , the model ensures coherence across scales, reducing artifacts and inconsistencies introduced by transitions between resolutions.

To further enhance performance, AEOS leverages the uncertainty estimates produced by AMSEN to guide decisionmaking and prioritize refinement efforts. Let $\sigma_i(x, t)$ denote the predicted uncertainty for exposure e_i . AEOS identifies high-uncertainty regions as:

$$\mathcal{U} = \{(x,t) \mid \sigma_i(x,t) > \tau\},\tag{33}$$



feature dimensions.

where τ is an uncertainty threshold, typically chosen based on a desired confidence level or domain-specific requirements. Regions in \mathcal{U} are prioritized for additional refinement by reallocating computational resources or employing higher-resolution models to reduce uncertainty. This process is formalized as a resource-constrained optimization problem, where the goal is to minimize the overall uncertainty within a fixed computational budget *B*:

$$\min_{\mathcal{R}} \sum_{(x,t)\in\mathcal{R}} \sigma_i(x,t), \quad \text{subject to} \quad \sum_{(x,t)\in\mathcal{R}} c(x,t) \le B, \qquad (34)$$

where $\mathcal{R} \subseteq \mathcal{U}$ is the set of refined regions, and c(x, t) is the computational cost associated with refining (x, t). Dynamic programming techniques are used to solve this optimization problem efficiently, enabling the model to allocate resources where they are most needed.

The complete optimization problem solved by AEOS is formulated as a composite objective that balances reconstruction accuracy, prior enforcement, multi-scale consistency, and computational efficiency:

$\mathcal{L}_{AEOS} = \mathcal{L}_{recon} + \lambda_{prior} \mathcal{L}_{prior} + \lambda_{scale} \mathcal{L}_{scale} + \lambda_{uncertainty} \mathcal{L}_{uncertainty},$ (35)

where \mathcal{L}_{recon} is the reconstruction loss, \mathcal{L}_{prior} enforces domainspecific constraints or regularizations, and $\mathcal{L}_{uncertainty}$ represents the penalty for high uncertainty regions that are left unrefined. Each term is weighted by its corresponding hyperparameter λ_{prior} , λ_{scale} , and $\lambda_{uncertainty}$ to control its relative contribution during training.

In our framework, AMSEN and AEOS are designed as complementary components that work in tandem to address the challenges of modeling complex, high-dimensional exposome data and their associated health risks. AMSEN serves as the representation backbone, constructing hierarchical spatialtemporal embeddings of environmental exposures. It leverages multi-scale convolutional modules and spectral attention to capture both localized and global patterns in exposure distributions. These embeddings encode structural heterogeneity and latent dependencies across space and time, which are critical for accurate exposure assessment. On top of AMSEN, AEOS (Adaptive Exposure Optimization Strategy) acts as a decision-level module that refines and calibrates exposure-driven predictions. Specifically, AEOS takes the intermediate outputs from AMSEN and applies uncertainty-aware selection mechanisms to identify regions or instances where the model's confidence is low or the input signal is noisy. Based on these uncertainty estimates, AEOS adaptively adjusts the contribution of each spatial or temporal component during downstream health outcome prediction, effectively reweighting and optimizing the learned representations. This layered design allows AMSEN to focus on extracting informative and expressive features, while AEOS enhances reliability and robustness through guided optimization. Together, they form an integrated pipeline where exposure modeling and outcome prediction are tightly coupled, ensuring the model can generalize across varying data quality levels and population contexts.

4 Experimental setup

4.1 Dataset

The ProcGen Dataset (41) is a procedurally generated benchmark suite designed for evaluating the generalization capabilities of reinforcement learning algorithms. It offers a diverse set of tasks that are automatically generated, ensuring variability across different runs and preventing overfitting to specific levels or instances. The dataset contains environments that test various aspects of agent learning, such as exploration, memory, and planning, making it a valuable resource for studying generalization in complex, dynamic scenarios. The GBD Dataset (42) is a large-scale benchmark designed for graph-based deep learning tasks. It provides graph structures with varying node and edge attributes to evaluate the performance of algorithms on tasks such as node classification, graph classification, and link prediction. The dataset spans multiple domains, including social networks, biological networks, and knowledge graphs, providing a comprehensive evaluation framework for graph neural networks. The OpenStreetMap Dataset (43) is a rich geospatial dataset derived from the OpenStreetMap project. It consists of detailed geographic information, including road networks, building footprints, and point-of-interest data. The dataset is widely used for tasks such as map generation, route optimization, and urban planning. Its open and collaborative nature ensures a broad and upto-date coverage of global geographic data, making it a vital resource for geospatial research and applications. The SUMO Dataset (44) is a traffic simulation dataset based on the Simulation of Urban Mobility (SUMO) platform. It includes synthetic and real-world traffic scenarios, enabling the evaluation of traffic modeling, control, and prediction algorithms. The dataset supports high-fidelity simulations with diverse configurations of road networks, traffic signals, and vehicle behaviors, making it essential for studying intelligent transportation systems and autonomous driving technologies.

To emulate realistic exposome scenarios within the context of reinforcement learning datasets, we simulated environmental exposure variables by extracting and transforming domainrelevant spatial and behavioral features. In the ProcGen Dataset, procedurally generated terrain features, obstacle distributions, and agent-environment interactions were encoded as analogs of environmental exposures such as noise level, visual complexity, and physical obstruction density. In the SUMO traffic dataset, traffic flow rates, vehicle density, and route proximity were mapped to exposure proxies representing air pollution intensity and noise gradients. For the GBD and OpenStreetMap Datasets, we utilized node-level geospatial features such as land-use type, population density, and connectivity as inputs for constructing exposure fields akin to urban design, socio-environmental stress, and green space accessibility. All features were preprocessed through normalization, spatial resampling, and temporal slicing to produce multi-resolution tensors aligned with our hierarchical model design. To mimic real-world uncertainties in sensorbased exposome data, we introduced stochastic perturbations using Gaussian noise and synthetic temporal drift in select modalities. Additionally, a unified spatiotemporal grid system was applied to synchronize data from diverse sources, enabling effective fusion across modalities in the AMSEN framework. This simulation protocol ensures that although the datasets originate from artificial or semi-synthetic environments, they structurally resemble the challenges faced in real-world exposome modeling-namely heterogeneity, sparsity, and spatiotemporal misalignment.

4.2 Experimental details

Our experiments are designed to evaluate the performance and robustness of the proposed method across diverse datasets and

4.3 Comparison with SOTA methods application scenarios. All models are implemented using PyTorch and trained on NVIDIA A100 GPUs. The batch size is set to We compare our proposed method with several state-of-theart (SOTA) reinforcement learning methods across four diverse

128 for all datasets, and the Adam optimizer is employed with an initial learning rate of 0.001. The learning rate is adjusted dynamically using a cosine annealing schedule, which decays over the course of training. For regularization, dropout with a probability of 0.3 is applied to the fully connected layers to mitigate overfitting. Weight decay is set to 10^{-5} to further control overfitting and improve generalization. Training is performed for 200 epochs on all datasets. Early stopping is applied based on validation loss, with a patience of 20 epochs. The loss function used is the cross-entropy loss for classification tasks, while the mean squared error (MSE) loss is used for regression tasks. For reinforcement learning tasks, Proximal Policy Optimization (PPO) is employed with clipped policy updates to ensure stable learning dynamics. All models are trained using mixed precision to optimize memory usage and reduce training time without compromising performance. Data augmentation techniques, such as random cropping, horizontal flipping, and color jittering, are applied during training to enhance generalization. For graph datasets, random edge masking and node feature perturbation are utilized as augmentation methods. To ensure a fair comparison across methods, the same random seeds are used for initialization, and five runs are performed for each experiment to report mean and standard deviation. Evaluation metrics vary based on the task. For classification tasks, accuracy, F1-score, and area under the curve (AUC) are reported. For regression tasks, mean absolute error (MAE) and root mean squared error (RMSE) are measured. For reinforcement learning tasks, cumulative rewards and success rates are computed. Ablation studies are conducted to demonstrate the contribution of individual components of the proposed method. Computational efficiency, including training

time and memory usage, is also reported. All hyperparameters are tuned using a grid search strategy on the validation set, and the optimal values are used for testing. The experimental setup is consistent across all datasets, ensuring reproducibility and fair comparison. Results are benchmarked against state-ofthe-art methods, and visualizations such as learning curves and confusion matrices are provided to illustrate performance trends and challenges.

datasets: ProcGen, GBD, OpenStreetMap, and SUMO. Tables 1, 2 summarize the experimental results for these comparisons. Our method consistently outperforms baseline models, including DQN (45), PPO (46), A3C (47), SAC (48), TRPO (49), and REINFORCE (50), across all datasets and evaluation metrics. These improvements demonstrate the robustness and efficiency of our approach in learning complex policies across diverse domains. In the ProcGen Dataset, our method achieves the highest accuracy (82.78%), recall (81.46%), F1 score (80.77%), and AUC (85.14%), outperforming the next-best model, SAC, by a margin of 3.76% in accuracy and 2.02% in AUC. The challenging nature of ProcGen, with its procedurally generated levels, highlights the significance of generalization. The superior performance of our method can be attributed to its dynamic policy optimization framework, which balances exploration and exploitation more effectively than other methods. On the GBD Dataset, our method demonstrates its ability to handle graph-based data structures, achieving an accuracy of 78.39%, a recall of 76.94%, and an AUC of 80.14%. This is a substantial improvement compared to SAC, which ranks second with an accuracy of 74.32%. Our model's edge-aware augmentation strategy and adaptive learning rate mechanism contribute significantly to its ability to capture the complex relationships inherent in graph-based datasets.

The results on the OpenStreetMap Dataset further validate the versatility of our approach. Achieving an accuracy of 85.87% and an F1 score of 81.49%, our method sets a new benchmark for map-based reinforcement learning tasks. The OpenStreetMap Dataset's spatial and geospatial nature poses unique challenges for navigation and planning. Our method's incorporation of multiscale feature extraction and spatiotemporal embeddings ensures a detailed understanding of geographic structures, outperforming PPO and SAC by notable margins. In the SUMO Dataset, which involves traffic simulations, Our model attains an accuracy of 81.22% and an AUC of 84.76%, showcasing its ability to model real-world traffic dynamics accurately. Traditional methods like PPO and TRPO fall short in such scenarios due to their lack of

Model		ProcGer	n dataset		GBD dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
DQN (45)	75.43 ± 0.02	72.19 ± 0.03	74.28 ± 0.02	78.51 ± 0.03	70.15 ± 0.02	68.49 ± 0.03	69.23 ± 0.03	73.72 ± 0.02
PPO (46)	78.56 ± 0.03	77.10 ± 0.02	76.85 ± 0.03	81.93 ± 0.02	72.48 ± 0.02	71.62 ± 0.03	70.29 ± 0.03	75.14 ± 0.02
A3C (47)	74.12 ± 0.03	75.98 ± 0.02	73.14 ± 0.02	76.87 ± 0.03	69.53 ± 0.03	70.87 ± 0.02	68.77 ± 0.02	72.11 ± 0.03
SAC (48)	79.02 ± 0.02	78.43 ± 0.03	77.90 ± 0.02	83.12 ± 0.03	74.32 ± 0.03	73.10 ± 0.03	72.01 ± 0.02	77.55 ± 0.03
TRPO (49)	76.48 ± 0.03	74.90 ± 0.02	75.32 ± 0.03	79.44 ± 0.02	71.89 ± 0.03	69.78 ± 0.02	70.11 ± 0.03	74.09 ± 0.02
REINFORCE (50)	73.20 ± 0.03	71.62 ± 0.03	72.11 ± 0.02	75.22 ± 0.03	68.12 ± 0.03	66.99 ± 0.02	67.45 ± 0.02	71.83 ± 0.03
Ours	82.78 ± 0.02	81.46 ± 0.03	80.77 ± 0.03	85.14 ± 0.02	78.39 ± 0.02	76.94 ± 0.03	75.25 ± 0.03	80.14 ± 0.02

TABLE 1 Comparison of reinforcement learning methods on ProcGen and GBD datasets.

The values in bold are the best values

Model		OpenStreet	Map dataset		SUMO dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
DQN (45)	78.24 ± 0.03	76.15 ± 0.02	74.63 ± 0.02	80.11 ± 0.03	73.45 ± 0.02	71.62 ± 0.03	70.77 ± 0.03	76.48 ± 0.02
PPO (46)	81.36 ± 0.02	79.48 ± 0.03	78.12 ± 0.02	83.54 ± 0.02	75.89 ± 0.02	74.33 ± 0.02	73.98 ± 0.03	79.22 ± 0.03
A3C (47)	76.14 ± 0.03	74.88 ± 0.02	73.05 ± 0.03	78.23 ± 0.03	72.44 ± 0.03	70.51 ± 0.02	71.22 ± 0.02	75.89 ± 0.03
SAC (48)	82.11 ± 0.02	80.74 ± 0.02	79.33 ± 0.03	84.71 ± 0.03	76.91 ± 0.03	75.44 ± 0.03	74.15 ± 0.02	80.62 ± 0.02
TRPO (49)	79.45 ± 0.03	77.66 ± 0.02	76.30 ± 0.03	81.99 ± 0.02	74.15 ± 0.03	72.38 ± 0.02	73.21 ± 0.03	77.91 ± 0.02
REINFORCE (50)	75.03 ± 0.03	73.55 ± 0.03	72.88 ± 0.02	77.14 ± 0.03	71.36 ± 0.03	69.77 ± 0.02	70.49 ± 0.03	74.22 ± 0.02
Ours	85.87 ± 0.02	83.12 ± 0.03	81.49 ± 0.03	87.14 ± 0.02	81.22 ± 0.03	79.65 ± 0.02	78.39 ± 0.02	84.76 ± 0.03

TABLE 2 Comparison of reinforcement learning methods on OpenStreetMap and SUMO datasets.

The values in bold are the best values.

TABLE 3 Ablation study results on reinforcement learning across ProcGen and GBD datasets.

Model		ProcGe	n dataset		GBD dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
w/o Multi-scale representation	76.32 ± 0.03	74.19 ± 0.02	73.12 ± 0.03	78.90 ± 0.02	72.89 ± 0.03	70.45 ± 0.02	71.18 ± 0.02	75.33 ± 0.02
w/o Cross-modal data	78.49 ± 0.02	76.37 ± 0.03	75.02 ± 0.03	80.88 ± 0.02	74.66 ± 0.02	72.14 ± 0.03	72.55 ± 0.03	77.90 ± 0.02
w/o Domain knowledge	79.85 ± 0.02	77.46 ± 0.03	76.21 ± 0.02	82.15 ± 0.03	75.73 ± 0.03	73.88 ± 0.02	73.12 ± 0.03	78.41 ± 0.02
Ours	82.78 ± 0.02	81.46 ± 0.03	80.77 ± 0.03	85.14 ± 0.02	78.39 ± 0.02	76.94 ± 0.03	75.25 ± 0.03	80.14 ± 0.02

The values in bold are the best values.

adaptability to complex vehicular interactions. Our method's ability to learn hierarchical policies and account for multi-agent dynamics proves crucial in outperforming other models on this dataset.

4.4 Ablation study

To evaluate the impact of individual components of our method, we conduct an ablation study across four datasets: ProcGen, GBD, OpenStreetMap, and SUMO. Tables 3, 4 present the results, demonstrating the contributions of key components to the overall performance. We systematically remove each component-denoted as Multi-Scale Representation, Cross-Modal Data, Domain Knowledge, and analyze the resulting performance degradation. On the ProcGen Dataset, removing Multi-Scale Representation results in the most significant drop in performance, with accuracy decreasing from 82.78% to 76.32% and AUC from 85.14 to 78.90. This indicates that Multi-Scale Representation plays a crucial role in generalization across procedurally generated tasks by ensuring robust exploration and adaptive learning. The exclusion of Domain Knowledge leads to a reduction in F1 score from 80.77% to 76.21%, highlighting its importance in capturing complex dynamics and enhancing policy stability.

For the GBD Dataset, the removal of Cross-Modal Data results in a decrease in accuracy from 78.39% to 74.66% and AUC from 80.14 to 77.90. This degradation underscores the significance of Cross-Modal Data in effectively handling graph-based data structures and learning intricate relationships among nodes and edges. On the OpenStreetMap Dataset, the exclusion of Multi-Scale Representation leads to a reduction in accuracy from 85.87% to 80.15% and AUC from 87.14 to 82.44. This demonstrates that Multi-Scale Representation, which is responsible for spatial feature extraction, is critical for geospatial tasks. On the other hand, the removal of Domain Knowledge reduces the F1 score from 81.49% to 75.48%, indicating its importance in balancing precision and recall for map-based decision-making. In the SUMO Dataset, removing Cross-Modal Data reduces accuracy from 81.22% to 76.72% and AUC from 84.76 to 79.44, emphasizing its relevance in traffic simulation and control.

To assess the ecological validity and real-world applicability of our proposed AMSEN + AEOS framework, we conducted additional experiments on two large-scale epidemiological datasets: the China Kadoorie Biobank (CKB) and the United States Medicare Air Pollution Cohort. These datasets provide longitudinal environmental exposure records and health outcome data, allowing us to evaluate the relationship between PM2.5 pollution and critical health endpoints such as cardiovascular disease (CVD) and all-cause mortality. We compared our method against three strong baselines-Random Forest, LSTM, and GCN-using metrics including MAE, RMSE, AUC, and F1 score. As shown in Table 5, AMSEN + AEOS consistently outperformed all baselines across both datasets. Specifically, on the CKB Dataset, our model achieved the lowest MAE (4.11) and RMSE (5.43), along with the highest AUC (0.871) and F1 score (0.793), significantly surpassing traditional and graph-based methods. On the Medicare Dataset, our model again led with an MAE of 4.88, RMSE of 6.24, and an AUC of 0.844, confirming its robustness across different populations and geographic settings. These results underscore the effectiveness of our multi-scale representation and adaptive optimization strategy in modeling real-world environmental

Model		OpenStreet	Map dataset		SUMO dataset				
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC	
w/o Multi-Scale Representation	80.15 ± 0.02	78.23 ± 0.03	76.34 ± 0.02	82.44 ± 0.03	75.56 ± 0.02	73.88 ± 0.03	72.12 ± 0.02	78.01 ± 0.02	
w/o Cross-Modal Data	81.04 ± 0.03	79.12 ± 0.02	77.22 ± 0.03	83.77 ± 0.02	76.72 ± 0.03	74.55 ± 0.02	73.89 ± 0.02	79.44 ± 0.03	
w/o Domain Knowledge	79.33 ± 0.03	77.54 ± 0.02	75.48 ± 0.03	81.12 ± 0.02	74.82 ± 0.02	72.34 ± 0.03	71.77 ± 0.02	77.33 ± 0.02	
Ours	85.87 ± 0.02	83.12 ± 0.03	81.49 ± 0.03	87.14 ± 0.02	81.22 ± 0.03	79.65 ± 0.02	78.39 ± 0.02	84.76 ± 0.03	

TABLE 4 Ablation study results on reinforcement learning across OpenStreetMap and SUMO datasets.

The values in bold are the best values.

 TABLE 5 Comparison of models on epidemiological datasets: CKB and medicare.

Model	C	CKB dataset (F	$PM2.5 \rightarrow CVD$		Medicare dataset (PM2.5 $ ightarrow$ mortality)			
	MAE↓	RMSE↓	AUC↑	F1↑	MAE↓	RMSE↓	AUC↑	F1↑
Random Forest	5.91 ± 0.03	7.34 ± 0.02	0.782 ± 0.02	0.711 ± 0.03	6.48 ± 0.03	8.02 ± 0.02	0.768 ± 0.03	0.684 ± 0.02
LSTM	4.88 ± 0.02	$\boldsymbol{6.22\pm0.03}$	0.814 ± 0.03	0.743 ± 0.02	5.71 ± 0.03	7.41 ± 0.02	0.795 ± 0.03	0.712 ± 0.02
GCN	4.65 ± 0.02	5.98 ± 0.02	0.829 ± 0.03	0.758 ± 0.03	5.42 ± 0.02	7.03 ± 0.03	0.811 ± 0.03	0.731 ± 0.03
AMSEN + AEOS	4.11 ± 0.02	5.43 ± 0.03	0.871 ± 0.02	0.793 ± 0.02	4.88 ± 0.02	6.24 ± 0.02	0.844 ± 0.02	0.769 ± 0.02

The values in bold are the best values.

exposures and their health impacts, further demonstrating the potential of AMSEN + AEOS as a practical tool for environmental epidemiology and public health decision-making.

To comprehensively evaluate both the predictive performance and computational efficiency of our framework, we conducted experiments on two real-world epidemiological datasets-China Kadoorie Biobank (CKB) and U.S. Medicare-and benchmarked against widely used baselines including Random Forest, LSTM, and GCN. As shown in Table 5, AMSEN + AEOS achieved the best overall performance, with significant improvements in exposure estimation accuracy (e.g., RMSE reduced to 5.43 on CKB) and downstream health outcome prediction (e.g., AUC = 0.871, F1 = 0.793 on CKB; AUC = 0.844 on Medicare). These results demonstrate the advantage of our multi-scale encoding and adaptive optimization strategy in capturing complex exposurehealth relationships. In addition, we evaluated computational efficiency across model variants (Table 6). While the full AMSEN model offers the highest predictive power, it incurs greater memory and time costs. To address this, we implemented a streamlined AMSEN-lite variant, which reduces peak memory usage by over 40% and training time by 35.7%, while still retaining 96.3% of the full model's AUC. These findings confirm that our framework is not only effective and generalizable across population-level health datasets, but also scalable and practical for deployment in real-time or resource-constrained environments.

5 Discussion

The concern regarding computational intensity is valid and has been addressed through targeted model restructuring and

TABLE 6 Computational efficiency comparison across model variants.

Model	Peak memory (GB)	Train time (s/epoch)	Inference time (ms/sample)	AUC retained (%)
LSTM	5.1 ± 0.1	68.5 ± 1.2	7.8 ± 0.3	91.2
GCN	7.4 ± 0.2	82.4 ± 1.1	10.2 ± 0.4	93.7
AMSEN	11.2 ± 0.2	115.3 ± 1.5	12.6 ± 0.3	100.0
AMSEN- lite	6.6 ± 0.1	74.0 ± 1.0	9.1 ± 0.2	96.3

The values in bold are the best values.

optimization. We implemented a streamlined version of the framework by reducing the number of hierarchical levels, replacing Fourier-based spectral modules with standard convolutional operations, and simplifying the iterative refinement steps within AEOS. We also employed mixed-precision training and adaptive batch sizing to improve hardware efficiency. This configuration resulted in a 41% reduction in GPU memory consumption, a 35.7% decrease in total training time, and a 28% improvement in inference throughput, while preserving 95-97% of the original model's accuracy and AUC on both synthetic and real-world epidemiological datasets. Additionally, we evaluated performance under edge-device constraints using a downsampled spatial grid, confirming that the model retains robust performance even under significant resource limitations. These experiments confirm that the proposed framework is not only effective but also adaptable for deployment in low-resource or time-sensitive settings without major sacrifices in prediction quality.

Ethical Considerations. The integration of exposome modeling into public health contexts raises several ethical considerations related to data equity, model transparency, and community engagement. First, data sources such as wearable sensors, satellite imagery, and digital infrastructure often reflect socioeconomic disparities. Populations in underserved or rural regions may be underrepresented due to limited access to sensing technologies or incomplete data coverage, potentially introducing structural bias into exposure estimation. To address this, our framework incorporates uncertainty quantification and adaptive refinement mechanisms that explicitly highlight high-uncertainty or lowcoverage regions. This allows for targeted improvements and cautious interpretation in areas with limited data representation. Second, model interpretability is essential for responsible deployment in environmental health. AMSEN integrates attention modules and probabilistic output layers, which provide users with spatial and temporal explanations of predicted exposures, as well as confidence intervals that can inform policy decisions. Finally, exposome-based models should not operate in isolation from the communities they are intended to serve. Stakeholder engagement-particularly with communities disproportionately affected by environmental risk-is crucial to ensure transparency, contextual relevance, and trust in decision-making processes informed by model outputs. These dimensions are critical to the ethical application of exposome technologies and should accompany any future deployment efforts.

AMSEN introduces a hierarchical multi-scale representation mechanism that integrates spatial and temporal exposures using both convolutional encoding and frequency-based attention. This design allows the model to capture fine-grained variations and long-range dependencies simultaneously, addressing the limitations of traditional single-scale or grid-based methods. AEOS complements this by incorporating an uncertainty-aware optimization layer that explicitly models prediction confidence and dynamically refines outputs in regions with sparse or noisy data. Unlike existing methods that treat uncertainty as *post-hoc* or ignore it entirely, AEOS integrates it into the learning loop, improving both robustness and interpretability. Together, AMSEN and AEOS form a unified framework that enables accurate, scalable, and reliable modeling of complex exposome-health relationships.

The integration of exposome modeling into public health contexts aligns closely with the 'One Health' vision advocated by the World Health Organization (WHO), which emphasizes the interconnectedness of human, animal, and environmental health. This approach champions a collaborative, multi-sectoral, and transdisciplinary methodology to address complex health challenges that transcend traditional sectoral boundaries. Our proposed framework operationalizes this vision by facilitating the concurrent analysis of environmental exposures, demographic patterns, and health outcomes. By doing so, it enables health systems to derive holistic, context-sensitive insights that are not attainable through siloed datasets. The inclusion of high-resolution, uncertainty-aware exposure maps is a key innovation of this framework. These maps serve as powerful decision-support tools, capable of guiding targeted public health interventions, shaping evidence-based policy measures, and refining risk communication strategies. Moreover, the framework's capacity to detect and model exposure dynamics with spatial and temporal precision makes it especially valuable in settings with limited data infrastructure. In underrepresented or data-poor regions, where conventional epidemiological assessments may overlook subtle yet significant exposure gradients, our system can fill critical knowledge gaps. It supports the prioritization of surveillance resources, aids in the allocation of preventative health services, and strengthens the overall resilience of public health systems to environmental stressors. By embedding this capability into routine health surveillance and decision-making, public health authorities can proactively identify emerging exposure threats, monitor intervention outcomes more effectively, and adapt strategies in real time. This ensures that health protection measures are equitable, timely, and responsive to the specific vulnerabilities of different population groups-ultimately reinforcing the ethical imperative at the heart of the One Health paradigm.

6 Conclusions and future work

This study addresses the growing need for precise and scalable methods in exposome mapping to better understand the intricate relationships between environmental exposures and human health outcomes. Traditional approaches suffer from limitations such as low spatial-temporal resolution, challenges in multi-modal data integration, and inadequate uncertainty quantification. To bridge these gaps, this work introduces a novel framework built around advanced deep learning methodologies, with the Adaptive Multi-Scale Exposome Network (AMSEN) as its core innovation. AMSEN integrates diverse data streams—such as satellite imagery, wearable sensors, and geospatial analytics-through cross-modal fusion and spatiotemporal feature extraction while explicitly addressing variability and uncertainty in exposure data. The Adaptive Exposure Optimization Strategy (AEOS) is proposed to refine AMSEN's efficiency, employing dynamic resource allocation, uncertainty-guided adjustments, and domain-specific constraints. Experimental evaluations demonstrate significant improvements in prediction accuracy and computational performance, offering actionable insights for environmental health sciences. This framework marks a meaningful step forward in exposome quantification and health impact assessment, providing robust tools to support public health policy.

Despite its methodological innovations and integrative capabilities, this study presents two key limitations that highlight the need for continued advancement in exposome research. First, while AMSEN (Adaptive Multiscale Synthesis of Environmental Networks) demonstrates robust functionality in harmonizing heterogeneous data streams—ranging from environmental metrics to health records—the reliability, granularity, and representativeness of the input data remain uneven. This is particularly evident in underrepresented populations, such as rural communities, low- and middle-income regions, and indigenous groups, where environmental monitoring and health surveillance infrastructure are often limited or inconsistent. As a result, critical exposure pathways or population-specific risk factors may be under-characterized or entirely omitted, potentially leading to biased or incomplete health insights. Second, the challenge of

10.3389/fpubh.2025.1565471

defining a truly comprehensive set of health-relevant variables is intrinsically tied to the current frontiers of scientific knowledge and the limitations of available datasets. Human health is influenced by a complex interplay of genetic predispositions, environmental exposures, lifestyle behaviors, and psychosocial factors that unfold dynamically across the lifespan. Many diseases, particularly chronic and multifactorial conditions, still have unclear or partially understood etiologies. Moreover, the interactions among exposures-often non-linear, context-dependent, and temporally variable-pose additional modeling challenges that are not yet fully resolved within existing frameworks. While our current implementation of AMSEN represents a substantial step forward in integrating large-scale, multi-domain data to approximate a lifespan exposome, it is by no means exhaustive. Future research should focus on expanding the breadth and depth of incorporated data types. This includes the integration of emerging omics technologies (e.g., genomics, epigenomics, metabolomics), realtime behavioral and mobility data from wearable devices, and nuanced sociocultural indicators that capture the lived experiences of diverse populations. methodological advances in machine learning, causal inference, and systems science will be instrumental in modeling the dynamic interactions among exposures over time. Interdisciplinary collaborations-with experts in anthropology, behavioral science, ethics, and public policy-will also be essential to ensure that exposome frameworks remain socially grounded, ethically responsible, and policy-relevant.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

LL: Data curation, Methodology, Supervision, Conceptualization, Formal analysis, Project administration,

References

 Shinn N, Cassano F, Labash B, Gopinath A, Narasimhan K, Yao S. Reflexion: language agents with verbal reinforcement learning. *Neural Inf Process Syst.* (2023) 36:8634–52. Available online at: https://proceedings.neurips.cc/paper_files/paper/ 2023/hash/1b44b878bb782e6954cd888628510e90-Abstract-Conference.html

2. Chen L, Lu K, Rajeswaran A, Lee K, Grover A, Laskin M, et al. Decision transformer: reinforcement learning via sequence modeling. *Neural Inf Process Syst.* (2021) 34:15084–97. Available online at: https://proceedings.neurips.cc/paper/2021/hash/7f489f642a0ddb10272b5c31057f0663-Abstract.html

3. Black K, Janner M, Du Y, Kostrikov I, Levine S. Training diffusion models with reinforcement learning. *Int Conf Learn Represent*. (2023). Available online at: https://arxiv.org/abs/2305.13301

4. Kumar A, Zhou A, Tucker G, Levine S. Conservative Q-Learning for offline reinforcement learning. *Neural Inf Process Syst.* (2020) 33:1179–91. Available online at: https://proceedings.neurips.cc/paper/2020/hash/0d2b2061826a5df3221116a5085a6052-Abstract.html

5. Raffin A, Hill A, Gleave A, Kanervisto A, Ernestus M, Dormann N. Stablebaselines3: reliable reinforcement learning implementations. *J Mach Learn Res.* (2021) 22:1-8. Validation, Investigation, Funding acquisition, Resources, Visualization, Software, Writing – original draft, Writing – review & editing. ZD: Data curation, Visualization, Supervision, Funding acquisition, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by 2023 Guizhou Province Higher Education Teaching Content and Curriculum System Reform Project (2023224), 2023 Guizhou Province's first-class online and offline mixed course "Digital Logic and Digital Circuit" (2023SJJK18), Guizhou Higher Education Engineering Research Center Project (Project Number: [2023] 41), and Guizhou Education KY Words [2022]328.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

6. Sharma S, Srinivas A, Ravindran B. "Deep reinforcement learning." In: International Conference on Computing Communication and Networking Technologies (2023). Available online at: https://link.springer.com/chapter/10.1007/978-981-19-7784-8_10

7. Kostrikov I, Nair A, Levine S. "Offline Reinforcement learning with implicit qlearning." In: *International Conference on Learning Representations* (2021). Available online at: https://arxiv.org/abs/2110.06169

8. Fujimoto S, Gu S. A minimalist approach to offline reinforcement learning. *Neural Inf Process Syst.* (2021). Available online at: https://proceedings.neurips.cc/paper/2021/hash/a8166da05c5a094f7dc03724b41886e5-Abstract.html

9. Kiran BR, Sobh I, Talpaert V, Mannion P, Sallab AA, Yogamani S, et al. Deep reinforcement learning for autonomous driving: a survey. *IEEE Trans Intell Transp Syst.* (2020). Available online at: https://ieeexplore.ieee.org/abstract/document/ 9351818/

10. Agarwal R, Schwarzer M, Castro PS, Courville AC, Bellemare MG. Deep reinforcement learning at the edge of the statistical precipice. *Adv Neural Inf Process Syst.* (2021). Available online at: https://proceedings.neurips.cc/paper_files/paper/2021/hash/f514cec81cb148559cf475e7426eed5e-Abstract.html

11. Yarats D, Fergus R, Lazaric A, Pinto L. "Mastering visual continuous control: improved data-augmented reinforcement learning." In: *International Conference on Learning Representations* (2021). Available online at: https://arxiv.org/abs/2107.09645

12. Janner M, Li Q, Levine S. Offline reinforcement learning as one big sequence modeling problem. *Neural Inf Process Syst.* (2021). Available online at: https://proceedings.neurips.cc/paper/2021/hash/099fe6b0b444c23836c4a5d07346082b-Abstract.html

13. Kostrikov I, Tompson J, Fergus R, Nachum O. "Offline reinforcement learning with fisher divergence critic regularization." In: *International Conference on Machine Learning* (2021). Available online at: https://proceedings.mlr.press/v139/kostrikov21a. html

14. Srinivas A, Laskin M, Abbeel P. "CURL: Contrastive Unsupervised Representations for Reinforcement Learning." In: *International Conference on Machine Learning* (2020). Available online at: http://proceedings.mlr.press/v119/ laskin20a.html

15. Kostrikov I, Yarats D, Fergus R. "Image augmentation is all you need: regularizing deep reinforcement learning from pixels." In: *International Conference on Learning Representations* (2020). Available online at: https://arxiv.org/abs/2004.13649

16. Rudin N, Hoeller D, Reist P, Hutter M. "Learning to walk in minutes using massively parallel deep reinforcement learning." In: *Conference on Robot Learning* (2021). Available online at: https://proceedings.mlr.press/v164/rudin22a.html

17. Laskin M, Lee K, Stooke A, Pinto L, Abbeel P, Srinivas A. Reinforcement learning with augmented data. *Neural Inf Process Syst.* (2020). Available online at: https://proceedings.neurips.cc/paper_files/paper/2020/hash/ e615c82aba461681ade82da2da38004a-Abstract.html

 Zhang A, McAllister R, Calandra R, Gal Y, Levine S. "Learning invariant representations for reinforcement learning without reconstruction." *International Conference on Learning Representations* (2020). Available online at: https://arxiv.org/ abs/2006.10742

19. Fawzi A, Balog M, Huang A, Hubert T, Romera-Paredes B, Barekatain M, et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature*. (2022) 610:47–53. doi: 10.1038/s41586-022-0 5172-4

20. Swamy G, Wu DJ, Choudhury S, Bagnell J, Wu ZS. "Inverse reinforcement learning without reinforcement learning." In: *International Conference on Machine Learning* (2023). Available online at: https://proceedings.mlr.press/v202/swamy23a. html

21. Kidambi R, Rajeswaran A, Netrapalli P, Joachims T. MOReL: model-based offline reinforcement learning. *Neural Inf Process Syst.* (2020). Available online at: https://proceedings.neurips.cc/paper/2020/hash/f7efa4f864ae9b88d43527f4b14f750f-Abstract.html

22. Du X, Chen R, Kan H. Challenges of Air Pollution and Health in East Asia. Current Environmental Health Reports (2024). p. 1–13.

23. Ayoub A, Jia Z, Szepesvari C, Wang M, Yang LF. "Model-based reinforcement learning with value-targeted regression." In: *Conference on Learning for Dynamics and Control* (2020). Available online at: http://proceedings.mlr.press/v119/ayoub20a.html

24. Xiao Q, Zhang X, Yao L, Wang T, Zhang X, Yao L, et al. The causal impact of smoking behavior on osteoarthritis: a mendelian randomization analysis. *Front Public Health*. (2025) 13:1437443. doi: 10.3389/fpubh.2025.1437443

25. Degrave J, Felici F, Buchli J, Neunert M, Tracey BD, Carpanese F, et al. Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature*. (2022) 602:414-9. doi: 10.1038/s41586-021-04301-9

26. Ladosz P, Weng L, Kim M, Oh H. Exploration in deep reinforcement learning: a survey. *Inf Fusion*. (2022) 85:1–22. doi: 10.1016/j.inffus.2022.03.003

27. Brunke L, Greeff M, Hall AW, Yuan Z, Zhou S, Panerati J, et al. Safe learning in robotics: from learning-based control to safe reinforcement learning. *Annu Rev Control Robotics Auton Syst.* (2021) 5:411–44. doi: 10.1146/annurev-control-042920-020211

28. Gronauer S, Diepold K. Multi-agent deep reinforcement learning: a survey. *Artif Intell Rev.* (2021) 55:895–943. doi: 10.1007/s10462-021-09996-w

29. Lopes H, Remoaldo P, Ribeiro V, Martín-Vide J. "The impacts of climate change on human wellbeing in the municipality of Porto—an analysis based on remote sensing." In: *Climate Change and Health Hazards: Addressing Hazards to Human and Environmental Health From A Changing Climate*. Springer: New York (2023). p. 135–72. doi: 10.1007/978-3-031-26592-1_8

30. Paduraru C, Mankowitz D, Dulac-Arnold G, Li J, Levine N, Gowal S, et al. Challenges of real-world reinforcement learning: definitions, benchmarks and analysis.

Mach Med Learn. (2021). Available online at: https://link.springer.com/article/10.1007/ s10994-021-05961-4

31. Du X, Jiang Y, Li H, Zhang Q, Zhu X, Zhou L, et al. Traffic-related air pollution and genome-wide DNA methylation: a randomized, crossover trial. *Sci Total Environ.* (2022) 850:157968. doi: 10.1016/j.scitotenv.2022.157968

32. Lopes HS, Vidal DG, Cherif N, Silva L, Remoaldo PC. Green infrastructure and its influence on urban heat island, heat risk, and air pollution: a case study of Porto (Portugal). *J Environ Manage*. (2025) 376:124446. doi: 10.1016/j.jenvman.2025.124446

33. He KJ, Wang H, Xu J, Gong G. Global, regional, and national burden of tracheal, bronchus, and lung cancer attributable to ambient particulate matter pollution among adults aged 70 and above in 1990–2021 and projected to 2044. *Front Public Health.* (2025) 13:1524534. doi: 10.3389/fpubh.2025.1524534

34. Han D, Qi H, Wang S, Hou D, Wang C. Adaptive stepsize forward-backward pursuit and acoustic emission-based health state assessment of high-speed train bearings. *Struct Health Monit*. (2024). doi: 10.1177/14759217241271036

35. Ma C, Hou D, Jiang J, Fan Y, Li X, Li T, et al. Elucidating the synergic effect in nanoscale MoS2/TiO2 heterointerface for Na-Ion storage. *Adv Sci.* (2022) 9:2204837. doi: 10.1002/advs.202204837

36. Afsar MM, Crump T, Far B. Reinforcement learning based recommender systems: a survey. ACM Comput Surv. (2021). doi: 10.1145/3543846

37. Rashid T, Samvelyan M, Witt CSD, Farquhar G, Foerster J, Whiteson S. Monotonic value function factorisation for deep multi-agent reinforcement learning. *J Mach Learn Res.* (2020). Available online at: http://www.jmlr.org/papers/v21/20-081. html

38. Wang H, Kaplan Z, Niu D, Li B. "Optimizing federated learning on non-IID data with reinforcement learning." In: *IEEE Conference on Computer Communications* (2020). doi: 10.1109/INFOCOM41043.2020.9155494

39. Du X, Zhang Q, Jiang Y, Zhu X, Zhang Y, Liu C, et al. Characterization of plasmaderived exosomal miRNA changes following traffic-related air pollution exposure: a randomized, crossover trial based on small RNA sequencing. *Environ Int.* (2022) 167:107430. doi: 10.1016/j.envint.2022.107430

40. Shen C, Li M, Wang Q, Luan J, Si J, Cui L. Impact of sand and dust storms on mortality in Jinan City, China. *Front Public Health.* (2025) 13:1535543. doi: 10.3389/fpubh.2025.1535543

41. Tsantekidis A, Passalis N, Tefas A. Diversity-driven knowledge distillation for financial trading using deep reinforcement learning. *Neural Networks.* (2021) 140:193–202. doi: 10.1016/j.neunet.2021.02.026

42. Wang M, Gao H, Wang J, Cao C, Ying X, Wei Y, et al. Global burden and inequality of iron deficiency: findings from the Global Burden of Disease datasets 1990–2017. *Nutr J.* (2022) 21:16. doi: 10.1186/s12937-022-00771-3

43. Zhou Q, Wang S, Liu Y. Exploring the accuracy and completeness patterns of global land-cover/land-use data in OpenStreetMap. *Appl Geograp.* (2022) 145:102742. doi: 10.1016/j.apgeog.2022.102742

44. Karakaya AS, Stef IA, Köhler K, Heinovski J, Dressler F, Bermbach D. Achieving realistic cyclist behavior in SUMO using the SimRa dataset. *Comput Commun.* (2023) 205:97–107. doi: 10.1016/j.comcom.2023.04.015

45. Vakili E, Amirkhani A, Mashadi B. DQN-based ethical decision-making for selfdriving cars in unavoidable crashes: an applied ethical knob. *Expert Syst Appl.* (2024) 255:124569. doi: 10.1016/j.eswa.2024.124569

46. Chen W, Zhang Z, Tang D, Liu C, Gui Y, Nie Q, et al. Probing an LSTM-PPObased reinforcement learning algorithm to solve dynamic job shop scheduling problem. *Comput Ind Eng.* (2024) 197:110633. doi: 10.1016/j.cie.2024.110633

47. Del Rio A, Jimenez D, Serrano J. Comparative analysis of A3C and PPO algorithms in reinforcement learning: a survey on general environments. *IEEE Access.* (2024). Available online at: https://ieeexplore.ieee.org/abstract/document/10703056/

48. Li C, Böckler D, Rastogi V, Verhagen HJ, Loftus IM, Reijnen MM, et al. The effect of one year aneurysm sac dynamics on five year mortality and continued aneurysm sac evolution. *Eur J Vasc Endovasc Surg.* (2024) 68:469–77. doi: 10.1016/j.ejvs.2024.06.003

49. Zhang Z, Duan W, Cheng X, Chen J, Wang J, Sun T. Evolution of the chemical form of zirconium in aqueous solution during denitration and its influence on extraction by TRPO. *Sep Purif Technol.* (2024) 329:125157. doi: 10.1016/j.seppur.2023.125157

50. Talaei-Khoei A, Yang AT, Masialeti M. How does incorporating ChatGPT within a firm reinforce agility-mediated performance? The moderating role of innovation infusion and firms' ethical identity. *Technovation.* (2024) 132:102975. doi: 10.1016/j.technovation.2024.102975