

Geospatially Conditioned Diffusion Models for Urban Environmental Exposure Mapping and Spatiotemporal Health Risk Prediction

Henrik Zimmerman

Department of Computer Science, University of Central Florida, Orlando, FL, USA.
henrik.z@ucf.edu

Claudio M. Makinen

Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS, USA.
makinen366@ku.edu

Abstract

Urban environmental exposure mapping has traditionally relied on deterministic interpolation, land use regression, and atmospheric dispersion models. These approaches provide useful regional estimates but often fail to represent fine-grained spatiotemporal variability, microenvironmental peaks, and uncertainty. Generative diffusion models, when conditioned on geospatial structure, offer a new pathway for synthesizing high-resolution exposure fields and supporting health risk assessment. This paper presents a system-level analysis of geospatially conditioned diffusion models for urban environmental exposure mapping and spatiotemporal health risk prediction. We examine architectural design decisions, conditioning mechanisms, data infrastructure, uncertainty quantification, and downstream epidemiological integration. The central argument is that the primary challenge is not solely model realism but the construction of trustworthy inference systems that align generative outputs with public health evidence, governance requirements, and equity commitments. The discussion covers trade-offs between centralized and edge deployment, computational sustainability, robustness under sensor dropout and concept drift, and the risk of encoding historical environmental injustice into synthetic exposure surfaces. We further consider policy implications for health risk communication, regulatory use, and community oversight. Recent examples from transport microenvironment measurements and geospatially guided diffusion are used to illustrate wider structural issues.

Keywords

urban environmental exposure, diffusion models, geospatial artificial intelligence, health risk assessment, environmental governance, spatiotemporal modeling.

1. Introduction

Urban environmental exposures such as fine particulate matter, nitrogen dioxide, ozone, volatile organic compounds, heat, and noise are unevenly distributed across space and time. Their health consequences depend on where people live, work, commute, and spend time indoors. Traditional exposure mapping has relied on regulatory monitoring networks, land use regression, interpolation, and atmospheric dispersion models. These methods provide stable regional estimates and have supported extensive epidemiological research, but they typically smooth over microenvironments, transient pollution peaks, and complex built-form effects.

Recent advances in generative modeling, particularly denoising diffusion probabilistic models [1], score-based generative frameworks [2], and latent diffusion architectures [3], offer a different modeling paradigm. Instead of predicting a single exposure surface, these models can learn the conditional distribution of environmental fields and generate multiple plausible realizations.

The potential of such models in environmental science extends a long trajectory of deep learning for spatial and temporal data [4]. At the same time, urban exposure inference increasingly depends on heterogeneous data sources, including low-cost sensors, mobile monitoring, crowdsourced observations, and volunteered geographic information [5]. These data streams are valuable but unevenly distributed, with variable calibration, sampling bias, and uncertain provenance. Earth system science has begun to explore how data-driven models can complement physical process understanding rather than replace it [6]. In urban health applications, this complementarity is especially important because a generated exposure field must be physically plausible, epidemiologically meaningful, and institutionally usable. This paper therefore approaches geospatially conditioned diffusion not as an isolated algorithm but as a socio-technical system that includes data infrastructure, model architecture, deployment arrangements, validation, governance, and policy integration.

2. From Deterministic Exposure Mapping to Generative Geospatial Inference

Conventional exposure mapping often treats environmental variables as continuous fields that can be estimated through regression or interpolation. These deterministic methods are computationally efficient and interpretable, but they encode assumptions about spatial autocorrelation and smoothness that may be violated in street canyons, transit platforms, construction sites, and near-road environments. In contrast, a generative model learns a high-dimensional distribution over possible environmental states conditioned on covariates such as land use, road geometry, meteorological conditions, and time. Sampling from this distribution can reveal alternative exposure scenarios that are consistent with observed constraints. This shift from point estimation to distributional inference is conceptually important for health risk prediction, where risk functions may be nonlinear and uncertainty can alter policy decisions.

Architecturally, geospatial conditioning requires representational capacity to capture both short-range texture and long-range urban structure. Deep residual networks [7] provide a useful backbone for extracting hierarchical features from satellite imagery, street-level imagery, and gridded physical model outputs. However, conditioning alone is not sufficient if the model cannot preserve spatial alignment between the conditioning variables and the generated exposure surface. Geospatially guided diffusion has been used to synthesize mixed-view panoramic urban scenes by conditioning on location and viewing geometry [8], illustrating that location-aware generative processes can maintain coherence across complex urban environments. This type of conditioning strategy can be extended from visual scenes to environmental variables such as pollutant concentrations, although the physical constraints differ.

A major system-level distinction concerns the treatment of uncertainty. Deterministic exposure maps often report a single estimated value with limited uncertainty information. Diffusion models can generate an ensemble of exposure surfaces, allowing analysts to distinguish between aleatoric uncertainty from irreducible environmental variability and epistemic uncertainty from sparse monitoring. Yet this advantage comes with a governance burden: if a generative model produces many plausible maps, decision-makers need clear protocols for summarizing ensembles, communicating forecast ranges, and avoiding false

precision. The system must also ensure that generated fields remain temporally consistent and do not introduce abrupt changes that would be implausible under atmospheric transport or urban activity dynamics.

3. System Architecture for Geospatially Conditioned Diffusion

A production-oriented system for urban exposure mapping can be organized into five interacting layers: data ingestion and harmonization, geospatial representation, conditional generation, uncertainty calibration, and health risk downstreaming. The data ingestion layer must accommodate regulatory monitors, low-cost sensor networks, mobile monitoring campaigns, satellite-derived estimates [9], meteorological reanalysis, traffic counts, building footprints, and transit schedules. Each source has different spatial support, temporal frequency, latency, and error structure. Harmonization requires not only technical alignment but also metadata standards that record sensor calibration, quality flags, and provenance. Without these governance mechanisms, heterogeneous inputs can silently bias the generative model.

Recent microenvironmental measurements of volatile organic compounds across subway, bus, and taxi modes [10] illustrate why spatial and temporal granularity matters in urban exposure assessment. Transit environments can produce short-term concentration peaks that are not visible in regional background maps. A geospatially conditioned diffusion system must therefore be able to represent transient microenvironments while remaining anchored to regional monitoring. This can be achieved by conditioning the model on trip-level mobility data, station identifiers, ventilation status, and time-of-day covariates. The architecture must trade off the richness of conditioning against the risk of overfitting to sparse or nonrepresentative monitoring campaigns.

The generation layer may follow a latent diffusion design to reduce computational cost by operating in a compressed representation rather than raw gridded space. Training can be staged: first pretraining on large unlabeled geospatial imagery and physical model outputs, then fine-tuning on high-quality exposure measurements. This strategy supports transfer across cities and reduces the need for dense local monitoring, but it can also import biases from the pretraining distribution. System designers must evaluate whether the compressed representation preserves fine-grained pollutant gradients near sources. Periodic validation against independent monitors and mobile campaigns is essential. Edge computing can be used for low-latency inference in transit or emergency response settings, while cloud infrastructure handles large-scale retraining and ensemble generation.

The choice between centralized and federated architectures is a structural trade-off. Centralized training can use large, harmonized datasets and powerful accelerators, but it raises privacy and data governance concerns when individual mobility traces or workplace locations are involved. Federated approaches keep raw data at local institutions but complicate model validation, hyperparameter coordination, and equitable participation. System architects must also plan for model versioning, rollback, and audit logs, particularly if outputs inform health advisories. These governance requirements are not secondary; they shape the architecture as much as compute efficiency does.

4. Spatiotemporal Health Risk Prediction and Downstream Integration

Health risk prediction requires more than high-resolution exposure surfaces. Epidemiological evidence has established associations between long-term air pollution exposure and mortality [11,12,13], but translating generated exposure fields into health risk estimates demands careful calibration and validation. Exposure assessment methods have long grappled with the

mismatch between fixed-site monitors and personal exposure [14]. Generative models can reduce this mismatch by simulating plausible exposure distributions along time-activity paths, but they do not automatically produce valid epidemiological estimates. Downstream models should propagate exposure ensembles through health impact functions and report risk intervals rather than single numbers.

A central concern is systemic inequity. Studies have shown that fine particulate matter exposure and its sources are distributed unevenly across racial and income groups [15,16]. If a generative model is trained on historical pollution patterns without explicitly accounting for structural discrimination, it may reproduce and even naturalize those inequities. Health risk predictions derived from such fields could then allocate resources away from the communities that need them most. This is not a hypothetical concern; it parallels documented algorithmic bias in health system resource allocation [17] and broader machine learning fairness problems [18]. A geospatially conditioned diffusion system must therefore include fairness audits and bias mitigation measures throughout the model lifecycle.

Downstream integration also raises the question of causal interpretability. Public health agencies may require explanations for why a particular neighborhood is predicted to have elevated risk. A diffusion model that synthesizes plausible exposure fields may be less interpretable than a land use regression equation. This trade-off must be managed by maintaining hybrid systems in which generative outputs are compared with physical models, empirical measurements, and known emission sources. Health risk maps should be accompanied by documentation of data sources, model assumptions, validation performance, and limitations. Such documentation supports regulatory acceptance and community trust.

5. Governance, Fairness, and Policy Implications

The deployment of geospatially conditioned diffusion models for health risk prediction raises governance issues that extend beyond technical performance. Ethical frameworks for artificial intelligence emphasize beneficence, non-maleficence, autonomy, and justice [19]. In environmental health, justice requires recognizing that sensor placement, data coverage, and model validation are often weaker in low-income communities and communities of color. If these data gaps are not explicitly addressed, generative models may fill them with spurious patterns that reproduce historical disinvestment rather than current environmental dynamics. This risk is especially acute in communities where regulatory monitoring density is low and where past redlining, highway construction, and industrial zoning have shaped both the built environment and the distribution of data infrastructure. A generative model trained on uneven data coverage can produce a synthetic exposure surface that appears spatially detailed but systematically underestimates or overestimates risk in marginalized neighborhoods. Mitigating this requires not only algorithmic adjustments but also participatory data collection, community-based sensor calibration, and transparent reporting about model limitations. Transfer learning and domain adaptation techniques have shown promise in extending air pollution predictions to regions with sparse monitoring [20], but such methods must be audited for whether they retain local environmental structure or simply import patterns from better-resourced areas.

The governance of geospatially conditioned diffusion models therefore extends beyond technical validation. Public health agencies and environmental regulators need to know whether a generated exposure map can support regulatory decisions, permit reviews, or health advisories. This requires an audit trail that links model outputs to training data, conditioning covariates, and validation metrics. Documentation practices such as model cards [25] can

record intended use, known failure modes, fairness evaluations, and subgroup performance. Without such institutional scaffolding, there is a danger that high-resolution synthetic maps will be treated as ground truth rather than as uncertain inferences. Policy frameworks must also clarify liability when a generated exposure estimate contributes to a decision that later proves harmful. These governance concerns are not separate from the technical system; they shape choices about model complexity, data retention, and inference latency.

6. Deployment, Sustainability, and Long-Term Maintenance

Deploying a geospatially conditioned diffusion model in an urban environmental health context involves substantial infrastructural choices. A centralized cloud architecture can support large-scale training, ensemble generation, and periodic retraining, but it may introduce unacceptable latency for applications such as transit micro-environment alerts or emergency response. Edge deployment can reduce latency and bandwidth requirements, yet it imposes constraints on model size, energy consumption, and the frequency of parameter updates. A hybrid arrangement may be more appropriate, with lightweight inference at the edge and heavier ensemble generation in the cloud. The choice should be informed by the decision context: a daily health risk map can tolerate longer processing times than a real-time ventilation advisory in a subway station.

Sustainability is often underestimated in discussions of generative modeling. Training large diffusion models consumes considerable energy and produces embodied carbon that may undercut the environmental health goals of the system [22]. Model compression, distillation, and efficient sampling schedules can reduce operational carbon, but they may also alter the diversity and fidelity of generated exposure fields. System designers must therefore quantify the full lifecycle environmental cost of the modeling infrastructure and compare it against the public health benefits of improved exposure information. This trade-off is particularly relevant when the system is intended to support long-term urban planning rather than short-term emergency response.

Long-term maintenance presents further structural challenges. Urban environments are not stationary; land use, traffic patterns, emissions sources, and building stock change over time. A model trained on historical data may gradually drift from current environmental conditions. Continuous validation against independent monitoring campaigns is necessary to detect concept drift. Federated learning offers one way to incorporate decentralized updates from local sensor networks without centralizing raw data [24], but it also introduces coordination overhead, heterogeneity across devices, and risks of model poisoning. A robust maintenance strategy must therefore combine scheduled retraining, drift detection, and community feedback loops. The system should also preserve archived model versions so that past health risk assessments can be reproduced and audited.

Urban computing frameworks that integrate air quality modeling, sensor networks, and city data infrastructures illustrate how such systems can be embedded in broader municipal decision processes [21]. These frameworks emphasize the importance of interoperability, data standards, and stakeholder collaboration. A diffusion-based exposure mapping system should follow similar principles. It must not operate as an isolated scientific instrument but as part of a larger socio-technical ecology that includes public health surveillance, transportation planning, housing policy, and climate adaptation. This integration requires institutional arrangements, not just software interfaces.

7. Conclusion

Geospatially conditioned diffusion models represent a significant shift in how urban environmental exposure can be mapped and how spatiotemporal health risk can be predicted. They move the field beyond deterministic interpolation and single-surface estimation toward explicit modeling of conditional distributions over plausible environmental states. This capability is valuable for capturing microenvironments, representing uncertainty, and supporting robust health risk inference. However, the system-level challenges are substantial. The architecture must reconcile the need for geospatial coherence with the diversity of urban data sources. The training process must guard against reproducing historical environmental injustice. The deployment and maintenance arrangements must balance computational cost, latency, data privacy, and long-term sustainability.

A central conclusion of this analysis is that the value of a geospatially conditioned diffusion model depends less on its generative realism than on its integration into trustworthy institutional processes. Health risk prediction is a consequential social activity, not simply a technical output. Exposure maps generated for regulatory or public health purposes must be accompanied by uncertainty statements, fairness evaluations, and audit trails. They must be validated against independent measurements and compared with physical models. They must be legible to communities whose members are directly affected by environmental hazards. The emergence of location-aware generative methods [8] and microenvironmental exposure measurements [10] provides important technical foundations, but these advances must be embedded within governance frameworks that treat synthetic exposure fields as provisional evidence rather than objective truth. Future research should therefore focus on interpretable conditioning mechanisms, fair data acquisition strategies, lifecycle environmental accounting, and participatory evaluation methods that bring affected communities into the design and oversight of these systems.

References

1. Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, 6840-6851.
2. Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., & Poole, B. (2021). Score-based generative modeling through stochastic differential equations. *International Conference on Learning Representations*.
3. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10684-10695.
4. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195-204.
5. Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin, D. A., Kilaru, V. J., & Preuss, P. W. (2013). The changing paradigm of air pollution monitoring. *Environmental Science & Technology*, 47(20), 11369-11377.
6. Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2019). Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8), 1544-1554.

7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-778.
8. Xiong, Z., Xing, X., Workman, S., Khanal, S., & Jacobs, N. (2024). Mixed-view panorama synthesis using geospatially guided diffusion. *Transactions on Machine Learning Research*.
9. van Donkelaar, A., Martin, R. V., Brauer, M., & Boys, B. L. (2015). Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environmental Health Perspectives*, 123(2), 135-143.
10. Wang, Z., Zhang, X., Wang, L., Fu, S., Wu, J., Xiong, J., & Huang, S. (2024). Concentrations and short-term health effects of VOCs in explored subway, bus and taxi in Beijing, China. *Atmospheric Environment*, 339, 120855.
11. Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., & Speizer, F. E. (1993). An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*, 329(24), 1753-1759.
12. Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*, 287(9), 1132-1141.
13. Crouse, D. L., Peters, P. A., van Donkelaar, A., Goldberg, M. S., Villeneuve, P. J., Brion, O., Khan, S., Atari, D. O., Jerrett, M., Pope, C. A., Brauer, M., Brook, J. R., Martin, R. V., Stieb, D., & Burnett, R. T. (2012). Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: A Canadian national cohort study. *Environmental Health Perspectives*, 120(5), 708-714.
14. Steinle, S., Reis, S., & Sabel, C. E. (2013). Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Science of the Total Environment*, 443, 184-193.
15. Tessum, C. W., Apte, J. S., Goodkind, A. L., Muller, N. Z., Mullins, K. A., Paoletta, D. A., Polasky, S., Springer, N. P., Thakrar, S. K., Marshall, J. D., & Hill, J. D. (2019). Inequity in consumption of goods and services adds to racial-ethnic disparities in air pollution exposure. *Proceedings of the National Academy of Sciences*, 116(13), 6001-6006.
16. Clark, L. P., Millet, D. B., & Marshall, J. D. (2014). National patterns in environmental injustice and inequality: Outdoor NO₂ air pollution in the United States. *PLoS ONE*, 9(4), e94431.
17. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453.
18. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1-35.
19. Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707.

20. Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., & Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environmental Pollution*, 231, 997-1004.
21. Kaginalkar, A., Kumar, S., Gargava, P., & Niyogi, D. (2021). Review of urban computing and air quality studies for building smart cities. *Environmental Modelling & Software*, 143, 105087.
22. Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J., & Bengio, Y. (2022). Tackling climate change with machine learning. *ACM Computing Surveys*, 55(2), 1-96.
23. Wynne, B. (1992). Uncertainty and environmental learning: Reconceiving science and policy in the preventive paradigm. *Global Environmental Change*, 2(2), 111-127.
24. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R. G. L., Eichner, H., Rouayheb, S. E., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., Gruteser, M., Harchaoui, Z., He, C., He, L., Huo, Z., Hutchinson, B., Hsu, J., Jaggi, M., Javidi, T., Joshi, G., Khodak, M., Konečný, J., Korolova, A., Koushanfar, F., Koyejo, S., Lepoint, T., Liu, Y., Mittal, P., Mohri, M., Nock, R., Özgür, A., Pagh, R., Qi, H., Ramage, D., Raskar, R., Raykova, M., Song, D., Song, W., Stich, S. U., Sun, Z., Suresh, A. T., Tramèr, F., Vepakomma, P., Wang, J., Xiong, L., Xu, Z., Yang, Q., Yu, F. X., Yu, H., & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1-2), 1-210.
25. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220-229.