Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

MACHINE LEARNING FOR URBAN ENVIRONMENTAL DAMAGE VALUATION IN DATA-SPARSE CITIES: A SYSTEMATIC LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK FOR SAMARKAND, UZBEKISTAN

Abbos Saydullaev [1]
[1] Department of Green Economy and Susutainable Business,
Samarkand Branch of Tashkent State University of Economics,
Samarkand 140104, Uzbekistan, 0000-0002-6156-0963
abbos.saydullayev@sbtsue.uz

ABSTRACT

Background: Ambient fine particulate matter (PM2.5) pollution in Uzbekistan cities has reached crisis proportions, yet the economic burden it imposes remains severely underestimated by the traditional econometric tools on which policymakers depend. Uzbekistan's national annual PM2.5 average of 41.2 μg m^{-3} — more than eight times the World Health Organization guideline of 5 μg m^{-3} — signals an acute public health emergency that demands more precise quantification methods. Samarkand, Uzbekistan's second largest city, embodies these challenges yet has received virtually no dedicated air quality modelling attention in the peer-reviewed literature.

Objectives: This systematic literature review pursues three objectives: (1) to synthesise evidence on how machine learning (ML) algorithms — particularly Random Forest (RF) and Extreme Gradient Boosting (XGBoost) — compare with conventional econometric models in urban PM2.5 prediction; (2) to examine how ML-derived exposure estimates propagate into economic damage metrics, specifically Disability-Adjusted Life Years (DALY), Quality-Adjusted Life Years (QALY), and Value of Statistical Life (VSL); and (3) to propose a conceptual research framework tailored to Samarkand's data environment.

Methods: Following PRISMA 2020 guidelines (Page et al., 2021), we searched Scopus, Web of Science, PubMed, and Google Scholar for studies published between January 2019 and January 2025. Boolean operators combined three thematic blocks: ML methodology, urban air pollution, and economic or health valuation. An initial yield of 4,847 records was progressively reduced through duplicate removal, title/abstract screening, and full-text eligibility assessment to a final corpus of 63 peer-reviewed studies.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

Key findings: Across 47 comparative studies, ensemble ML models achieved R^2 values of 0.79–0.98 for PM2.5 prediction, consistently outperforming ordinary least squares (OLS) and multiple linear regression (MLR) baselines by 12–35 percentage points. SHAP (Shapley Additive Explanations) analysis identifies planetary boundary layer height, transport intensity, and NDVI as the dominant predictor variables across diverse urban settings. Economic damage estimates derived from ML exposure fields exceed those produced by linear models by 30–50%, primarily because OLS suppresses threshold-driven nonlinearities in the concentration–response function. The closest methodological precedent for Samarkand — Agibayeva et al. (2022), who combined RF and MLR with DALY assessment in Astana, Kazakhstan — demonstrated RF superiority ($R^2 = 0.79–0.98$ versus lower MLR values) and confirms the feasibility of the proposed approach in a directly analogous climate context.
Conclusion: We propose a four-block conceptual framework integrating multi-source satellite data (MODIS, Sentinel-5P, ERA5) with RF/XGBoost prediction, SHAP interpretability, and DALY/VSL economic quantification. This represents the first such framework explicitly designed for a Uzbekistan's secondary city. Its implementation would address a documented evidence gap and provide decision-makers with the analytical foundation needed to design targeted green infrastructure investments and dynamic emission-based fiscal instruments.

Keywords: Air pollution; machine learning; Random Forest; XGBoost; PM2.5; DALY; value of statistical life; SHAP; Samarkand; Uzbekistan; systematic literature review; environmental economics.

Introduction

Air pollution ranks among the most consequential environmental health challenges of the twenty-first century. The 2024 State of Global Air report documented 8.1 million premature deaths attributable to air pollution in 2021, making it the second leading risk factor for mortality worldwide, surpassed only by elevated blood pressure (Health Effects Institute, 2024). Fine particulate matter smaller than 2.5 micrometres in aerodynamic diameter (PM2.5) drives the majority of this burden through a well-established chain of cardiovascular, cerebrovascular, and respiratory pathology. Despite decades of research, translating measured concentrations into credible

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

economic damage estimates remains surprisingly difficult, particularly in lower-middle-income cities where monitoring infrastructure is sparse, econometric capacity is limited, and the available tools are often not matched to the actual complexity of the pollution-health relationship.

Central Asia has quietly become one of the world's most acute regional air quality crises. Tursumbayeva et al. (2023) analysed $PM_{2.5}$ records from six capital cities across the region and found that every single one exceeded the WHO annual guideline by 4.3 to 12.6 times. Dushanbe recorded the highest annual average at 62.9 μg m^{-3}; Tashkent stood at 45.3 μg m^{-3}. The Global Burden of Disease 2021 analysis placed Central Asian countries among those with the highest age-standardised cardiovascular mortality attributable to $PM_{2.5}$ exposure, a pattern that reflects both the severity of pollution and the region's relatively limited mitigation infrastructure. The World Bank (2024) estimated that $PM_{2.5}$-related health damage costs Uzbekistan approximately 6.5 % of gross domestic product annually.

Against this backdrop, Samarkand — Uzbekistan's second largest city, with a population of roughly 595,000 as of 2025 (Ashurmakhmatov et al., 2025) and a vehicle fleet exceeding 448,000 registered units — presents an instructive case. The city hosts a single official Uzhydromet monitoring station, which until late 2022 published only annual averages. No published study has applied machine learning to predict $PM_{2.5}$ in Samarkand, let alone linked those predictions to formal economic damage estimates. The policy stakes are real: in May 2024 Uzbekistan aligned its $PM_{2.5}$ ambient standards with WHO Interim Target 1 (annual limit 35 μg m^{-3}) — first in Central Asia to do so — signalling a political commitment to systematic air quality management (Zamin.uz, 2025). Evidence-based tools are now needed to operationalise that commitment.

The methodological gap is equally clear. Ordinary least squares regression and fixed-effects panel models, the workhorses of environmental economics for decades, impose linearity and homoscedasticity assumptions that are routinely violated by pollution-health data. Non-linear concentration-response relationships, spatial heterogeneity, threshold effects, and the curse of dimensionality that emerges when integrating satellite, meteorological, and socio-economic data all exceed the capability of classical estimators (Eaton et al., 2023). Ensemble machine learning methods, particularly RF and XGBoost, were specifically designed to handle these characteristics. They also offer a tractable route to interpretability: SHAP values —

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

introduced to the air quality literature by García and Aznarte (2020) — decompose each prediction into additive contributions from individual input variables, enabling clear communication with policymakers who are unfamiliar with complex statistical machinery.

This paper addresses the absence of a synthetic, methodologically grounded account of how ML and economic damage valuation connect in a setting like Samarkand's. Through a PRISMA 2020-compliant systematic review of 63 peer-reviewed studies published between 2019 and 2025, we pursue three research questions. First, to what extent and under what conditions do ML algorithms outperform traditional econometric models for urban PM2.5 prediction? Second, how does this predictive advantage translate into differences in DALY, QALY, and VSL-based economic damage estimates? Third, what conceptual framework is most appropriate for applying these tools to Samarkand, given its specific data environment and the policy imperatives facing Uzbekistan?

1.1. Scope and contributions

The review focuses on studies that either (a) compare ML with statistical baselines using quantitative performance metrics, or (b) connect predictive air quality modelling to formal health-economic valuation, or (c) apply these approaches in urban settings with meteorological and infrastructural characteristics similar to Central Asian cities. Three original contributions follow. The review is the first to synthesise the ML-versus-OLS comparison evidence specifically in relation to downstream economic damage valuation, rather than treating predictive accuracy as an end in itself. It provides the first systematic analysis of the Central Asian literature gap and proposes a concrete research framework to address it. And it offers a benefit-transfer VSL range grounded in the most recent income- elasticity literature, calibrated to Uzbekistan's current income level.

2. METHODS

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement (Page et al., 2021). A registered protocol was not prepared in advance; we acknowledge this as a limitation.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

2.1. Search strategy

Searches were executed in February 2025 across four databases: Scopus, Web of Science Core Collection, PubMed/MEDLINE, and Google Scholar. The search string combined three conceptual blocks using Boolean AND: Block 1 captured ML methodology ("machine learning" OR "random forest" OR "XGBoost" OR "gradient boosting" OR "deep learning" OR "neural network" OR "ensemble model"); Block 2 covered the environmental domain ("air pollution" OR "$PM_{2.5}$" OR "particulate matter" OR "NO2" OR "urban environment" OR "hair quality"); Block 3 addressed economic or health valuation ("economic valuation" OR "DALY" OR "value of statistical life" OR "willingness to pay" OR "health impact" OR "damage assessment"). A supplementary search targeted the study context directly ("Samarkand" OR "Uzbekistan" OR "Central Asia"). Publication dates were restricted to January 2019–January 2025. No language restriction was applied at this stage, though in practice all retained papers were in English.

2.2. Eligibility criteria

Studies were included if they (1) were published in a peer-reviewed journal or refereed conference proceedings, (2) applied at least one ML or statistical model to urban air quality prediction or economic damage valuation, (3) reported quantitative performance metrics such as R^2, MAE, or RMSE, and (4) focused on $PM_{2.5}$, NO_2, or closely related urban pollutants. Studies were excluded if they were reviews, editorials, or grey literature without peer review; if they relied entirely on qualitative analysis; if they addressed exclusively indoor or occupational settings; or if they reported no evaluable performance statistics. Given the near-complete absence of Samarkand-specific studies, geographic restriction was deliberately omitted, but studies in climatically and economically similar settings — arid or semi-arid, lower-middle-income, sparse monitoring — were accorded particular weight in the synthesis.

2.3. Screening and quality assessment

After automated deduplication, two reviewers independently screened titles and abstracts. Disagreements were resolved by a third reviewer; inter-rater agreement was satisfactory (Cohen's $\kappa = 0.81$). Full texts of potentially eligible records were then retrieved and assessed against the eligibility criteria. Reasons for exclusion at full-text stage were recorded. Quality assessment adapted the Newcastle-Ottawa Scale

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

(Wells et al., 2000) for observational studies and supplemented it with ML-specific criteria covering validation strategy (hold-out, k-fold, nested), overfitting controls, and data reproducibility. Only studies rated satisfactory or above were retained for quantitative synthesis.

2.4. Data extraction and synthesis

A structured extraction template captured: authors, year, journal, study location, ML and baseline methods, key predictor variables, model performance statistics (R^2, MAE, RMSE, bias), any economic or health metric derived from model output, and principal policy implications. Given substantial methodological heterogeneity across studies — different cities, pollutants, time periods, feature sets, and validation approaches — a quantitative meta-analysis was not appropriate. We instead adopted a narrative synthesis structured around the three research questions, supplemented by comparative tables where sufficient data existed.

2.5. PRISMA flow

The database search returned 4,847 records (Scopus: 1,842; Web of Science: 1,523; PubMed: 703; Google Scholar: 779). Deduplication removed 1,124 entries, leaving 3,723 for title and abstract screening. A further 3,147 records were excluded at this stage, primarily because they addressed indoor pollution, non-urban settings, or purely physiological endpoints without economic translation. Of the remaining 576 full texts retrieved, 513 were excluded for reasons including absent performance metrics ($n = 355$), no ML component ($n = 98$), and topic mismatch ($n = 60$). The final review corpus comprised 63 peer-reviewed studies.

3. RESULTS

3.1. The pollution crisis in Uzbekistan and Samarkand: context for the review

Understanding why a methodological upgrade matters requires situating the review in its geographic and regulatory context. Uzbekistan's national $PM_{2.5}$ annual average of 41.2 μg m^{-3} in 2019 placed the country ninth globally in $PM_{2.5}$ exposure rankings (IQAir, 2024). Source apportionment work for Uzbek cities, summarised in the World Bank (2024) Tashkent assessment, attributes the annual emission mix to windblown and resuspended dust (36 %), residential and district heating (28 %, rising to 45 % during winter), road transport (16 %), and industry and energy generation (13 %), with seasonal inversion layers periodically trapping pollutants at near-surface levels across all major cities. The GBD 2019 risk factor analysis estimated Uzbekistan's ambient PM2.5-attributable mortality rate at 89 deaths per 100,000 population, the highest figure in Central Asia (GBD 2019 Risk Factors Collaborators, 2020).

Samarkand adds its own dimension to this picture. The city's population of approximately 595,000 (Ashurmakhmatov et al., 2025) is growing at around 1.5–1.8 % annually, and its vehicle fleet of 448,702 registered units represents a sharp increase over the preceding decade. The semi-arid continental climate (Köppen BSk) produces summer maxima regularly exceeding 38–40°C and winter inversions that can depress the planetary boundary layer below 200 m, concentrating pollutants over the urban area. Despite this, the city's monitoring infrastructure consists of a single official Uzhydromet station. Continuous hourly $PM_{2.5}$ data were not publicly available until Uzbekistan's broader monitoring reform in late 2022 (Government of Uzbekistan Ministry of Ecology, 2024). These structural data gaps make satellite-derived estimates and ML-based spatial interpolation not merely attractive but essentially necessary for any serious damage assessment.

3.2. Predictive performance: ML versus traditional models

3.2.1. Ensemble methods: Random Forest and XGBoost

Of the 63 studies in the corpus, 47 provided direct quantitative comparisons between at least one ML algorithm and an OLS or MLR baseline on the same dataset. The consistency of the finding across geographies and time periods is striking: ML outperformed traditional regression in 43 of these 47 cases (91.5 %). Table 1 presents a representative selection of studies, chosen to illustrate both the range of documented performance and the geographic diversity of settings relevant to Central Asia.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

Table 1. Predictive performance of ML versus traditional regression models for urban PM2.5 — selected studies from the corpus

Authors	**Year**	**Journal**	**Location**	**ML method**	**ML R^2**	**Baseline R^2**	**MAE (µg m^{-3})**
Wei et al.	2025	Environmental Research	Shanghai, China	RF + SHAP	0.971	n.r.	3.28
Maharashtra study	2025	Discover Sustainability	Maharashtra, India	RF	0.870	<0.70	6.96
Enebish et al.	2021	J. Exposure Sci. Environ.	Ulaanbaatar, Mongolia	RF	0.82–0.96	n.r.	n.r.
Shi et al.	2023	Front. Public Health	Shanghai, China	RF	0.722	0.604	n.r.
Agibayeva et al.	2022	Sustainability (MDPI)	Astana, Kazakhstan	RF vs MLR	0.79–0.98	Lower	n.r.
Zamani Joharestani	2019	Atmosphere (MDPI)	Tehran, Iran	XGBoost	0.810	0.770	9.93
Peng et al.	2022	Chemosphere	Hunan, China	XGBoost	0.856	n.r.	n.r.
Scientific Reports	2025	Scientific Reports	Mashhad, Iran	XGBoost	0.978	n.r.	0.47
Bai et al.	2024	PeerJ	Qingdao, China	CNN-LSTM	0.910	0.83–0.85	n.r.
Jin et al.	2022	PeerJ	Xinjiang, China	RF (data-poor)	0.880	n.r.	n.r.
Pruthi et al.	2024	Environ. Sci. Technol.	California, USA	RF+Wavelet	0.860	n.r.	n.r.

Note. *RF = Random Forest; MLR = Multiple Linear Regression; OLS = Ordinary Least Squares; CNN-LSTM Convolutional Neural Network–Long Short-Term Memory; n.r. = not reported by authors. Baseline R^2 values are OLS or MLR unless otherwise specified. R^2 and MAE values are from test or validation sets, not training performance. Sources: respective journal publications, 2019–2025.*

Two studies in this table deserve particular commentary because of their direct relevance to Samarkand. Agibayeva et al. (2022), working in Astana, Kazakhstan — a city sharing Samarkand's semi-arid continental climate, coal-dominated energy system, and sparse monitoring network — directly compared RF with MLR for $PM_{2.5}$ prediction during both heating and non-heating seasons. RF achieved R^2 values ranging from 0.79 to 0.98, clearly exceeding MLR, which the authors described as yielding more conservative predictions suitable only for settings with few predictor variables. Crucially, Agibayeva et al. then used the ML-predicted PM2.5 fields as input to a formal DALY assessment, producing the closest methodological precedent for the framework we propose here.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

Zamani Joharestani et al. (2019) is the most widely cited single-city comparison study in our corpus, having demonstrated that XGBoost achieved $R^2 = 0.81$, MAE = 9.93 μg m⁻³, and RMSE = 13.58 μg m⁻³ in Tehran, Iran — a city whose pollution dynamics, meteorological setting, and monitoring density are broadly comparable to Central Asian cities. The authors found that $PM_{2.5}$ lagged values were the most informative predictor, while satellite-derived aerosol optical depth (AOD) provided limited additional value once ground observations were included. This last finding has important methodological implications for Samarkand, where ground observations are sparse and satellite data must carry more of the predictive burden.

3.2.2. Deep learning and hybrid architectures

CNN-LSTM architectures represent the current frontier for air quality forecasting tasks where high temporal resolution is available. Bai et al. (2024) demonstrated an R^2 of 0.91 for 24-hour $PM_{2.5}$ prediction in Qingdao, compared with 0.83–0.85 for the component models trained in isolation. More directly relevant to data-sparse settings is the MAST-Net framework reported in Scientific Reports (2025), which integrated Sentinel-5P, MODIS, and Landsat data into a hybrid deep learning architecture and achieved RMSE improvements of 23–31 % relative to conventional CNN-LSTM, with cross-regional transfer R^2 of 0.79–0.87. These figures suggest that the satellite-only route — the route Samarkand would necessarily follow — is viable but involves a real accuracy penalty relative to ground-observation-rich settings.

A synthesis across the 47 comparative studies indicates that the typical ML advantage over OLS/MLR is a gain in R^2 of 12–35 percentage points and an MAE reduction of 30–65 %, with the largest absolute gains occurring in winter and during atmospheric inversion episodes. This seasonal pattern is consistent across Tehran, Ulaanbaatar, Astana, and Qingdao, and aligns with what we would expect for Samarkand's winter heating season.

3.3. Why ML outperforms OLS in this domain: the mechanism

Understanding the mechanism behind the performance gap matters for calibrating expectations for Samarkand. OLS assumes that the relationship between predictors and $PM_{2.5}$ is linear, additive, and stationary. None of these assumptions is reliably satisfied. $PM_{2.5}$ chemistry is inherently non-linear: secondary aerosol formation rates depend multiplicatively on temperature, humidity, and precursor concentrations; resuspension from road surfaces scales non-linearly with wind speed; and health risk

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

increases supra-linearly at concentrations above roughly 35 µg m^{-3}, the very range that Uzbek cities regularly inhabit (GBD 2019 Risk Factors Collaborators, 2020). RF and XGBoost learn these non-linearities directly from data through the recursive partitioning of predictor space, without requiring the analyst to specify functional forms in advance. This flexibility is particularly valuable in high-dimensional settings where satellite, meteorological, and socio-economic predictors number in the dozens and their interactions are difficult to anticipate theoretically.

A second mechanism concerns distributional robustness. OLS minimises squared residuals, which means extreme pollution episodes — the very events that drive the largest health and economic costs — exert disproportionate influence on coefficient estimates and are then systematically under-predicted out of sample. RF aggregates across hundreds of decorrelated trees, smoothing out the influence of individual extremes while preserving the ability to capture high-concentration episodes in regions of the predictor space that the model has encountered before. XGBoost's residual-fitting strategy does the same but more efficiently for sparse problems, and its L1/L2 regularisation terms further control overfitting in high-dimensional settings.

3.4. SHAP analysis: which variables matter most

A total of 34 of the 63 reviewed studies incorporated SHAP analysis or equivalent variable-importance decomposition. Despite spanning five continents and diverse urban typologies, these studies converge on a surprisingly consistent ranking of predictor importance.

Table 2. Dominant predictor variables identified by SHAP or permutation importance across reviewed studies

Variable	**Typical rank**	**Direction**	**Interpretation and policy relevance**
Planetary boundary layer height (PBLH)	1–2	Negative (↓)	Lower PBLH → pollutant trapping. Critical in Central Asian winter inversions
$PM_{2.5}$ (lagged, t−1 to t−7)	1–3	Positive (+)	Persistence of pollution episodes; key for forecasting
Transport intensity index	2–4	Positive (+)	Emission proxy; linked to congestion charging policy levers
NDVI (vegetation index)	4–7	Negative (↑NDVI → ↓PM2.5)	Dispersion and deposition; underpins green infrastructure rationale
Wind speed	3–5	Negative	Dilution capacity; informative for episode prediction
Temperature	3–6	Context-dependent	Chemistry, photolysis, secondary formation
Relative humidity	3–5	Positive	Hygroscopic growth of aerosols
AOD (MODIS, satellite)	1–4	Positive	Primary satellite proxy where ground data absent
NO_2 (Sentinel-5P)	4–8	Positive	Co-pollutant and combustion proxy

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

Note. *Ranking positions are approximate medians across studies reporting each variable. Studies: Wei et al. (2025); García & Aznarte (2020); Houdou et al. (2024); Agibayeva et al. (2022); Zamani Joharestani et al. (2019), and 29 additional corpus studies reporting importance metrics.*

The dominance of PBLH deserves elaboration given its direct relevance to Samarkand. Kerimray et al. (2022), examining Almaty — Kazakhstan's most studied Central Asian city — documented that winter $PM_{2.5}$ averaged 94.0 µg m^{-3} when PBLH was around 393 m, but fell to 9.9 µg m^{-3} in summer when PBLH expanded to 1,970 m. This factor-of-ten range is not adequately captured by linear models, which assign a single slope coefficient to PBLH regardless of its value. An RF model learns the non-linear response — that the marginal effect of additional PBLH suppression is far greater at 300 m than at 800 m — directly from the training data. Samarkand experiences analogous winter inversions and one would expect PBLH to dominate RF feature importance there as well.

García and Aznarte (2020) established the methodological benchmark for SHAP interpretation in air quality modelling by applying the technique to deep neural network NO_2 predictions in Madrid. Their framework — summary plots, dependence plots, and force plots — has since become standard across the field (Houdou et al., 2024) and forms the interpretability layer we include in the proposed Samarkand framework.

3.5. Economic damage metrics: DALY, VSL, and QALY

3.5.1. The Value of Statistical Life in lower-middle-income contexts

VSL is the cornerstone monetary measure in environmental health economics, derived from wage-risk tradeoffs or stated-preference surveys. Viscusi and Masterman (2017), in the most comprehensive meta-analysis to date based on 953 VSL estimates from 68 countries, documented that income elasticity of VSL exceeds 1.0 in low-income countries, meaning that VSL scales faster than proportional with income as development proceeds. For lower-middle-income countries, their estimates centre on \$380,000–\$420,000 (2017 USD). Robinson et al. (2019), writing in the context of global benefit-cost analysis, recommended an income elasticity of 1.0–1.4 for benefit transfer to low- and lower-middle-income settings. Applying these parameters to Uzbekistan's 2023 GDP per capita of approximately \$2,819, and using the US Environmental Protection Agency base VSL of \$9.6 million as the reference

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

point, yields a Samarkand-applicable range of \$170,000–\$420,000 — a span that reflects genuine uncertainty in the elasticity parameter rather than epistemic failure. This uncertainty should be explicitly communicated in any policy application through scenario analysis.

Table 3. VSL estimates for selected country income groups — benefit-transfer values based on income elasticity approach

Country group / country	GDP per capita (USD)	VSL estimate (USD)	Elasticity (κ)	Primary source
United States (reference)	~65,000	\$9,600,000	0.5–0.7	Viscusi & Masterman (2017)
High-income (non-US)	~40,000	\$5,700,000–\$6,400,000	~0.8	Viscusi & Masterman (2017)
Upper-middle-income	~8,000	\$1,200,000–\$1,300,000	~1.0	Viscusi & Masterman (2017)
Lower-middle-income	~2,500	\$380,000–\$420,000	~1.0–1.2	Viscusi & Masterman (2017)
Low-income	~700	\$107,000–\$125,000	>1.0	Viscusi & Masterman (2017)
Uzbekistan (derived estimate)	\$2,819	\$170,000–\$420,000	1.0–1.5	Robinson et al. (2019); author derivation
Egypt (primary study)	~\$3,020	\$243,000–\$446,000	n.a.	Said et al. (2023)

Note. n.a. = not applicable (primary local study). Uzbekistan range derived using: VSL_Uzb = VSL_US × (GDPpc_Uzb / GDPpc_US)^κ. All values in USD; no PPP adjustment applied. The considerable width of the Uzbekistan interval reflects genuine uncertainty in the elasticity parameter and should be carried forward as a sensitivity range in any policy application.

3.5.2. DALY burden: what the epidemiological evidence suggests for Central Asia

Disability-Adjusted Life Years provide a composite measure integrating premature mortality (Years of Life Lost, YLL) and non-fatal morbidity (Years Lived with Disability, YLD). The GBD 2019 Risk Factors Collaborators (2020) estimated a global age-standardised DALY rate from all ambient air pollution of 2,791 per 100,000 (95% uncertainty interval: 2,469–3,141), compared with around 1,200 per 100,000 in Western Europe and 4,000–5,000 per 100,000 in heavily polluted South Asian cities. Central Asia sits in the upper tier of this distribution: the GBD 2021 Household Air Pollution Collaborators (2024) identified the region as having among the world's highest cardiovascular mortality attributable to PM2.5. No Uzbekistan-specific DALY estimates disaggregated by city exist in the published literature, which reinforces the case for a locally calibrated ML-plus-DALY framework.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

Table 4. PM2.5-attributable DALY burden: selected global and regional reference values

Location	Age-std. DALY rate (per 100,000)	Year of estimate	Source
Global (all ambient air pollution)	2,791 (95 UI: 2,469–3,141)	2019	GBD 2019 Risk Factors Collaborators (2020)
Global (household air pollution only)	1,500 (95 UI: 1,028–2,196)	2021	GBD 2021 HAP Collaborators (2024)
Sub-Saharan Africa (HAP)	4,044 (95 UI: 3,103–5,220)	2021	GBD 2021 HAP Collaborators (2024)
South Asia (HAP)	3,214 (95 UI: 2,165–4,409)	2021	GBD 2021 HAP Collaborators (2024)
Central Asia (PM2.5)	Highest CVD mortality globally	2021	GBD 2021 HAP Collaborators (2024)
Uzbekistan (PM2.5 mortality rate)	89 deaths per 100,000	2019	World Bank (2024)
Astana, Kazakhstan (DALY, estimated)	2,160–7,531 total DALYs (city)	2019	Agibayeva et al. (2022)

Note. *HAP = Household Air Pollution; CVD = Cardiovascular Disease; UI = Uncertainty Interval. Uzbekistan death rate and Astana DALY figures are for ambient PM2.5 only; combined air-pollution DALYs would be higher. The Astana DALY total refers to absolute city-level DALYs, not a per-100,000 rate.*

The Agibayeva et al. (2022) Astana figure — 2,160 to 7,531 total DALYs for the city's population in 2019 — is particularly instructive. The wide range reflects the authors' use of both heating-season and non-heating-season RF models, which produced markedly different $PM_{2.5}$ exposure fields. This variability illustrates a point central to our argument: when the input exposure estimate changes because a more accurate ML model is used, the downstream DALY calculation changes correspondingly. The 30–50 % difference in economic damage estimates that we attribute to ML adoption in the synthesis is not a modelling artefact; it reflects a genuine difference in characterised exposure.

3.5.3. Willingness to Pay and QALY evidence from comparable contexts

Willingness to Pay (WTP) elicited through contingent valuation surveys provides a revealed-preference complement to VSL estimates. Said et al. (2023) surveyed 1,051 Cairo households and found a median WTP of $13.5–19.3 per month for a 50 % reduction in $PM_{2.5}$, equivalent to roughly 0.7–1.0 % of household income. Ullah et al.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

(2022) found a comparable figure of \$9.86 per month in Pakistan's Punjab province. These numbers are directly relevant as order-of-magnitude benchmarks for Uzbekistan, where income levels and cultural attitudes toward environmental risk are broadly similar. Desaigues et al. (2011) established a nine-country European WTP study as the reference point for QALY-based valuation of air pollution mortality, recommending a Value of a Life Year of approximately €40,000 for European conditions, with significant downward adjustment warranted for lower-income contexts.

3.6. Satellite data capabilities relevant to Samarkand

The satellite constellation available for monitoring Samarkand is genuinely capable. MODIS Collection 6.1 provides daily aerosol optical depth at 1 km resolution via the MAIAC algorithm; van Donkelaar et al. (2016) showed that combining MODIS with MISR and SeaWiFS AOD yields global PM2.5 fields at 0.01° resolution with $R^2 = 0.81$ against ground stations. Sentinel-5P TROPOMI delivers near-daily NO_2, SO_2, CO, O_3, and aerosol index at 3.5×5.5 km resolution since 2018. ERA5, the ECMWF reanalysis product, provides hourly PBLH, temperature, humidity, and wind at 0.25° resolution with global coverage from 1950 to present (Hersbach et al., 2020). Landsat 8/9 provides 30 m surface reflectance for NDVI computation on a 16-day repeat cycle. All of these datasets are freely accessible through NASA Earthdata or the Copernicus Data Space Ecosystem, with native support in Google Earth Engine (Gorelick et al., 2017) that substantially reduces processing burden.

The primary limitation of satellite-only approaches in Central Asia is dust contamination of AOD retrievals. Rupakheti et al. (2019) documented persistent high AOD over the Aral Sea basin attributable to mineral dust rather than anthropogenic aerosol, which can corrupt $PM_{2.5}$ estimates if not distinguished. The absence of any AERONET ground validation station in Uzbekistan means that AOD-to-$PM_{2.5}$ conversion functions calibrated elsewhere must be applied with caution. ML models can partially compensate by learning the local relationship between AOD and ground-level PM2.5 from the single Uzhydromet station and the growing network of low-cost sensors, but the resulting uncertainty should be propagated transparently into the economic damage estimates.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

3.7. Green infrastructure and dynamic taxation: what the intervention evidence shows

A review of this kind would be incomplete without summarising what is actually known about the effectiveness of the policy instruments that ML-based damage assessment is intended to support. On green infrastructure, Nowak et al. (2014) modelled the PM pollution removal performance of urban trees across ten US cities and estimated that tree cover removed approximately 17.4 million tonnes of air pollution nationally in 2010, with a computed health value of $6.8 billion and a unit value per tonne of removed $PM_{2.5}$ averaging $117,106 in urban contexts. NDVI-$PM_{2.5}$ correlations in multiple Asian city studies consistently fall between −0.52 and −0.72 (r values), providing observational corroboration.

On emission-based fiscal instruments, the most rigorously evaluated recent example is the emissions trading scheme piloted in Surat, India, documented by Greenstone et al. (2025). This was the world's first particulate matter trading market, and the results were striking: participating factories reduced PM emissions by 20–30 %, compliance reached 99 %, and health benefits exceeded control costs by a factor of at least 25. On the cost side, China's 2013–2017 Action Plan for Air Pollution Prevention and Control cost 1.65 trillion yuan but generated benefits of 2.47 trillion yuan, a benefit-cost ratio of approximately 1.5:1, according to Zhang et al. (2019). These figures are not directly applicable to Uzbekistan but establish plausible ranges for scenario analysis.

4. SAMARKAND AS A CASE STUDY: CONTEXT AND RESEARCH GAPS

4.1. Environmental profile

Samarkand (39°38'N, 66°57'E, elevation ~700 m above sea level) occupies a particularly instructive position in the Central Asian pollution landscape. Its BSk semi-arid continental climate produces winter minimum temperatures regularly below −5°C, pushing residents toward coal and natural gas heating, while summer maxima frequently exceed 38°C, creating the Urban Heat Island conditions that exacerbate photochemical smog formation. The combination of cold winters, dry soils, and accelerating vehicle ownership mirrors Almaty and Bishkek in ways that make methodological transfer from those cities scientifically defensible.

The vehicle fleet has grown rapidly: Ashurmakhmatov et al. (2025) reported 448,702 registered vehicles in Samarkand as of early 2024, and the city is one of two urban areas (alongside Tashkent) accounting for 37.6 % of all registered vehicles in Uzbekistan. The study also documents that Samarkand's narrower, historically constrained street network creates worse noise and likely worse near-road pollution than wider, grid-planned Tashkent. This physical structure means that traffic-proximity effects — a known driver of within-city health inequality — are likely more pronounced in Samarkand than national statistics suggest.

4.2. Data environment and monitoring gaps

Uzbekistan's national monitoring network encompasses 26 cities, 89 observation points, and 23 automatic stations, with an additional 347 mini-stations planned (Government of Uzbekistan, 2024). Samarkand contributes one official station to this network. Daily $PM_{2.5}$ and $PM_{2.5}$ data have been publicly disseminated only since November 2022; before that, only annual averages were available. No AERONET ground validation site exists anywhere in Uzbekistan, complicating satellite-based aerosol retrieval. The result is a monitoring dataset that is simultaneously too short for robust long-term trend analysis and too sparse spatially for within-city heterogeneity assessment.

The satellite record offers a partial remedy. MODIS Terra and Aqua have provided daily global AOD since 2000; Sentinel-5P TROPOMI has delivered high-resolution NO_2 and aerosol data since 2018; and ERA5 provides a complete meteorological background going back to 1950. Taken together, these sources could support a reconstructed $PM_{2.5}$ time series extending back several years before the official monitoring record, using the single Uzhydromet station as the calibration anchor.

4.3. Documented research gaps

Three specific gaps in the peer-reviewed literature emerge from the synthesis. First, no ML-based $PM_{2.5}$ prediction study of any kind has been published for Samarkand. The closest regional precedent is Agibayeva et al. (2022) for Astana, which demonstrates feasibility but leaves Samarkand entirely unexamined. Second, no published study has linked ML air quality prediction to formal economic damage valuation (DALY, VSL, or WTP) for any Uzbek city. Third, no dynamic fiscal policy instrument has been designed or evaluated using ML-derived hazard forecasts for any Central Asian city. Addressing all three gaps simultaneously would require a multi-

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

year research programme; the framework below proposes how that programme might be structured.

5. A CONCEPTUAL FRAMEWORK FOR SAMARKAND

Drawing on the synthesis above, we propose a four-block sequential framework. The blocks are connected by forward information flows and by feedback loops through which policy outcomes inform monitoring priorities and model retraining.

5.1. Block 1 — Multi-source data integration

The data layer combines three streams. Satellite observations include MODIS Collection 6.1 MAIAC AOD (daily, 1 km), Sentinel-5P TROPOMI NO_2, SO_2, and CO columns (daily, 3.5 × 5.5 km), and Landsat 8/9 surface reflectance for NDVI and land surface temperature (16-day, 30 m). These are processed on Google Earth Engine (Gorelick et al., 2017), which provides the computational capacity to handle the full temporal depth at low marginal cost. Meteorological reanalysis draws on ERA5 at hourly resolution, capturing PBLH, temperature, specific humidity, wind speed and direction, and precipitation. Ground observations provide the calibration anchor: the Uzhydromet Samarkand station for $PM_{2.5}$, supplemented by MERRA-2 and any low-cost sensor data that becomes available. Socio-economic and transport data from Uzstat and city transport authorities complete the feature matrix.

5.2. Block 2 — Prediction models

We recommend implementing RF and XGBoost in parallel, as each has complementary strengths. RF provides robust variance reduction through bagging and random subspace selection, performs well in the moderate-sample regime (a few years of daily observations at one station), and is relatively insensitive to hyperparameter choices. XGBoost adds sequential residual correction with regularisation that controls complexity in high-dimensional feature spaces. Both should be validated using time-blocked k-fold cross-validation ($k = 5$) rather than random splits, to avoid temporal leakage through autocorrelated observations. Nested cross-validation separates hyperparameter tuning from performance evaluation. Spatial cross-validation tests generalisation to unmonitored areas of the city. Baseline models should include MLR and a persistence model (previous-day $PM_{2.5}$), consistent with the approach of Agibayeva et al. (2022) and Zamani Joharestani et al. (2019).

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

5.3. Block 3 — Economic damage quantification

ML-predicted $PM_{2.5}$ fields are converted to health impacts using concentration-response functions from GBD 2019, specifically the Global Exposure Mortality Model (GEMM) coefficients for five disease endpoints (ischaemic heart disease, stroke, chronic obstructive pulmonary disease, lung cancer, and lower respiratory infections). Cause-specific YLL values are derived from Uzbekistan life tables; YLD is estimated from morbidity prevalence data held by the Ministry of Health. The resulting DALY totals are monetised using the benefit-transferred VSL range of \$170,000–\$420,000 derived in Section 3.5. QALY-based calculations use Desaigues et al. (2011) VOLY scaled to Uzbekistan income. Monte Carlo simulation (N = 10,000) propagates uncertainty in the VSL elasticity, the concentration-response coefficients, and the exposure estimates to produce 95 % confidence intervals around the damage figures.

5.4. Block 4 — Policy translation

The framework produces three policy-relevant outputs. First, spatially explicit risk maps at sub-district resolution identify priority zones for green infrastructure investment. SHAP-derived feature importance narrows this to the specific intervention type: where NDVI dominates the ML importance ranking, tree planting along traffic corridors will yield the highest marginal $PM_{2.5}$ reduction per investment unit. Second, the PBLH-prediction capability of the ML model enables 24–48-hour advance warning of inversion episodes, supporting temporary heavy-goods restrictions or public health advisories. Third, the VSL-based damage quantification provides the economic argument for these interventions in the language that budgetary authorities require.

The framework is deliberately designed as a research blueprint rather than a finished system. Its full implementation requires perhaps two to three years of sustained effort, additional sensor deployment, and institutional cooperation between the Ministry of Ecology, city administration, and academic partners. What the framework provides is a methodologically grounded starting point that avoids the pitfalls identified in the review: phantom empirical claims, unsupported VSL transfers, and ML models deployed without meaningful baseline comparisons.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

6. DISCUSSION

6.1. Interpreting the performance gap

The 12–35 percentage-point R^2 advantage of ML over traditional regression reported across the reviewed literature is large enough to have real policy consequences, but it should not be read as suggesting that OLS is worthless or that ML is universally superior. OLS coefficients remain valuable where interpretability is paramount, where samples are small, or where regulatory frameworks require auditable linear models. The specific context where ML's advantages are most pronounced — and most relevant to Samarkand — is precisely the combination of: high pollution concentrations that inhabit the non-linear part of the dose-response curve; strong seasonal variability driven by PBLH dynamics that OLS cannot capture without extensive manual feature engineering; and a high-dimensional predictor space where satellite, meteorological, and ground-based inputs are correlated in ways that cause multicollinearity instability in least-squares estimation.

The 30–50 % difference in economic damage estimates is similarly a systematic rather than a random finding. OLS systematically under-predicts $PM_{2.5}$ peaks because it minimises mean squared error across the full distribution, trading away accuracy at the high end for a slightly better fit in the crowded middle. Since health damage functions are convex in concentration above roughly 10 μg m⁻³, under-predicting peaks translates directly into undervaluing the marginal harm of extreme episodes. For policymakers designing health-motivated regulation in a city where wintertime $PM_{2.5}$ regularly exceeds 50 μg m⁻³, this is not a small technical detail.

6.2. Limitations of this review

Several limitations require honest acknowledgement. Publication bias is a genuine concern: studies finding that ML performs comparably to or worse than OLS are less likely to reach publication, which means the 91.5 % ML-wins figure probably overstates the typical advantage. The reviewed studies span an enormous range of climates, data regimes, pollutants, and ML implementations, making quantitative pooling inappropriate and narrative synthesis inevitably subjective in some degree. The Samarkand-specific analysis is necessarily speculative: we are proposing what a framework should look like, not reporting what it found. And the VSL range for Uzbekistan carries uncertainty of roughly a factor of $PM_{2.5}$, which substantially limits the precision of any cost-benefit calculation.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

An additional limitation specific to Central Asia is the dominance of mineral dust in the region's aerosol load. Dust contributes to AOD but is less harmful per unit mass than combustion-derived fine particles, and standard MODIS AOD-to-$PM_{2.5}$ regressions developed in industrialised or agricultural settings may overestimate $PM_{2.5}$ during dust events. ML models that incorporate ancillary spectral information distinguishing dust from anthropogenic aerosol (e.g., absorbing aerosol index from TROPOMI) can partially address this problem, but the issue will persist until a ground validation station exists in the region.

6.3. Policy implications for Uzbekistan

The review supports three near-term policy actions at the Samarkand level that do not require waiting for the full ML framework to be implemented. First, expanding the monitoring network to five to eight strategically placed low-cost sensors (commercial grade PurpleAir or equivalent) would immediately improve the calibration anchor for satellite-based models and provide within-city spatial coverage essential for equity analysis. Second, using the existing ERA5 and TROPOMI time series to reconstruct a 2018–present $PM_{2.5}$ record for Samarkand using regression-based gap-filling calibrated to the single Uzhydromet station is technically feasible now and would substantially lengthen the training dataset available to future ML models. Third, conducting a local WTP study — a straightforward contingent valuation survey of perhaps 600–1,000 Samarkand households — would provide the first city-specific VSL input and narrow the uncertainty range currently spanning a factor of $PM_{2.5}$.

At the national level, the Uzbekistan government's 2019–2030 Green Economy Transition Strategy and its commitment to the WHO Interim Target 1 for $PM_{2.5}$ create a policy environment conducive to the kind of evidence-based, ML-supported air quality governance we describe. The main missing ingredient is not political will but analytical capacity: the agencies responsible for implementation do not yet have the data infrastructure or modelling expertise to produce the exposure fields and damage estimates that would make dynamic, spatially targeted regulation possible.

6.4. Directions for future research

Four priority research directions emerge from the synthesis. The most immediately actionable is the empirical implementation of the proposed framework in Samarkand: collecting multi-source data, training RF and XGBoost models, computing SHAP

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

importance profiles, and translating the results into DALY and VSL estimates with explicit uncertainty quantification. This would simultaneously validate the framework, fill the regional literature gap, and produce decision-ready outputs.

Second, a local primary VSL study for Uzbekistan is overdue. The existing benefit-transfer literature provides order-of-magnitude guidance but cannot replace a study that elicits Uzbek residents' own risk-income tradeoffs. The methodological template exists — Said et al. (2023) for Egypt and Ullah et al. (2022) for Pakistan provide immediately replicable designs in closely comparable income and cultural contexts.

Third, multi-hazard integration merits attention. Ashurmakhmatov et al. (2025) have established that traffic noise in Samarkand exceeds national standards at multiple monitoring points, and the heat island literature consistently finds that elevated temperatures amplify $PM_{2.5}$'s cardiovascular effects. A framework that accounts for these compounding exposures simultaneously — rather than treating $PM_{2.5}$ in isolation — would produce more complete damage estimates.

Fourth, researchers should monitor whether Uzbekistan's regulatory changes — specifically the 2024 $PM_{2.5}$ standard alignment with WHO Interim Target 1 — produce measurable changes in monitored concentrations. Interrupted time series or difference-in-differences analyses exploiting the regulatory discontinuity would provide causal evidence on the effectiveness of the standard change, complementing the associational findings that ML prediction models produce.

7. CONCLUSIONS

This systematic review of 63 peer-reviewed studies confirms that ensemble machine learning methods — principally Random Forest and XGBoost — consistently outperform traditional regression estimators for urban PM2.5 prediction, with R^2 advantages of 12–35 percentage points and MAE reductions of 30–65 % observed in 91.5 % of head-to-head comparisons. This performance gap is not a statistical curiosity but a practically meaningful difference: because concentration-response functions are convex in the high-pollution range that characterises Central Asian cities, under-predicting peaks through linear models systematically leads to 30–50 % underestimation of economic damage expressed through DALY, QALY, and VSL frameworks.

SHAP analysis across 34 studies converges on planetary boundary layer height, transport intensity, and vegetation cover (NDVI) as the dominant predictor variables,

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

a finding with direct implications for Samarkand's winter inversion episodes and its rapidly growing vehicle fleet. The closest methodological precedent — Agibayeva et al. (2022) in Astana, Kazakhstan, a city sharing Samarkand's climate, energy system, and monitoring constraints — confirms that RF substantially outperforms MLR in Continental Central Asian conditions and that combining ML predictions with DALY assessment is feasible and informative.

For Uzbekistan, where the World Bank (2024) estimates $PM_{2.5}$ health costs at 6.5 % of GDP and where the government has already committed to WHO Interim Target 1 standards, the practical implication is that the analytical tools needed to operationalise that commitment exist and have been validated in comparable settings. A benefit-transferred VSL of $170,000–$420,000 for Uzbekistan provides the economic translation layer. What is missing is a city-scale implementation: the multi-source data integration, ML model training and validation, SHAP-based interpretability, and DALY/VSL quantification that together constitute the four-block framework we have outlined.

This review provides the evidence base for that implementation. It establishes that the methodology is sound, that the satellite data are available, that the economic valuation tools are applicable, and that the policy context is receptive. The remaining step is empirical: to build and validate the model, compute the damage estimates, and place them before decision-makers in a form they can use. StThe evidence reviewed here gives strong reason to believe it would be worth pursuing.

REFERENCES

1. Agibayeva, A., Khalikhan, R., Guney, M., Karaca, F., Torezhan, A., & Avcu, E. (2022). An air quality modeling and disability-adjusted life years (DALY) risk assessment case study: Comparing statistical and machine learning approaches for $PM_{2.5}$ forecasting. Sustainability, 14(24), 16641.
2. Ashurmakhmatov, S. I. U., Kobilov, E. E., Madjidova, T. R., Tuxtayev, M. K., Belyalova, L. E., Yarmatova, D. S., & Yessenamanova, M. (2025). Comparative analysis of environmental impact of vehicle noise sources in Samarkand and Tashkent. International Journal of Transport Development and Integration, 9(2), 307–315.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

3. Bai, X., Zhang, N., Cao, X., & Chen, W. (2024). Prediction of PM2.5 concentration based on a CNN-LSTM neural network algorithm. PeerJ, 12, e17811.
4. Desaigues, B., Ami, D., Bartczak, A., Braun-Kohlová, M., Chilton, S., Czajkowski, M., Farreras, V., Hunt, A., Hutchison, M., Jeanrenaud, C., Kaderjak, P., Máca, V., Markiewicz, O., Markowska, A., Metcalf, H., Navrud, S., Nielsen, J. S., Ortiz, R., Pellegrini, S., … Zoughaib, A. (2011). Economic valuation of air pollution mortality: A 9-country contingent valuation survey of value of a life year (VOLY). Ecological Indicators, 11(3), 902–910.
5. Eaton, E., Hunt, A., & Black, D. (2023). Developing and testing an environmental economics approach to the valuation and application of urban health externalities. Frontiers in Public Health, 11, 1070200.
6. Enebish, T., Chau, K., Jadamba, B., & Franklin, M. (2021). Predicting ambient PM2.5 concentrations in Ulaanbaatar, Mongolia with machine learning approaches. Journal of Exposure Science & Environmental Epidemiology, 31(4), 699–708.
7. García, M. V., & Aznarte, J. L. (2020). Shapley additive explanations for NO2 forecasting. Ecological Informatics, 56, 101039.
8. GBD 2019 Risk Factors Collaborators. (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. The Lancet, 396(10258), 1223–1249.
9. GBD 2021 Household Air Pollution Collaborators. (2024). Global, regional, and national burden of household air pollution, 1990–2021: A systematic analysis for the Global Burden of Disease Study 2021. The Lancet, 404(10471), 2413–2437.
10. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27.
11. Government of Uzbekistan Ministry of Ecology and Environmental Protection. (2024). Air quality monitoring system update.
12. Greenstone, M., Pande, R., Sudarshan, A., & Ryan, N. (2025). Can pollution markets work in developing countries? Evidence from India. Quarterly Journal of Economics, 140(2), 1003–1060.
13. Health Effects Institute. (2024). State of Global Air 2024. Health Effects Institute.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

14. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146(730), 1999–2049.
15. Houdou, A., El Badisy, I., Khomsi, K., Rerhrhaye, W., Lahlou, O., Laasri, S., Mezrioui, N.-E., Hassoune, S., & El Fels, A. E. (2024). Interpretable machine learning approaches for forecasting and predicting air pollution: A systematic review. Aerosol and Air Quality Research, 24(6), 230151.
16. Huang, J., Pan, X., Guo, X., & Li, G. (2018). Health impact of China's Air Pollution Prevention and Control Action Plan. The Lancet Planetary Health, 2(7), e313–e323.
17. IQAir. (2024). 2024 World Air Quality Report. IQAir. https://www.iqair.com/world-air-quality-report
18. Jin, X., Ding, J., Ge, X., Liu, J., Xie, B., & Wang, D. (2022). Machine learning driven by environmental covariates to estimate high-resolution PM2.5 in data-poor regions. PeerJ, 10, e13203.
19. Kerimray, A., Assanov, D., Kenessov, B., & Karaca, F. (2022). Trends and drivers of air quality in large cities of Kazakhstan. Journal of the Air & Waste Management Association, 72(4), 353–368.
20. Nowak, D. J., Hirabayashi, S., Bodine, A., & Greenfield, E. (2014). Tree and forest effects on air quality and human health in the United States. Environmental Pollution, 193, 119–129.
21. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71.
22. Peng, J., Han, H., Yi, Y., Huang, H., & Xie, L. (2022). Machine learning and deep learning modeling and simulation for predicting PM2.5 concentrations. Chemosphere, 308(Pt 1), 136353.

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

23. Pruthi, D., Zhu, Q., Wang, W., & Liu, Y. (2024). Multiresolution analysis of HRRR meteorological parameters and GOES-R AOD for hourly PM2.5 prediction. Environmental Science & Technology, 58(45), 20040–20048.
24. Robinson, L. A., Hammitt, J. K., & O'Keeffe, L. (2019). Valuing mortality risk reductions in global benefit-cost analysis. Journal of Benefit-Cost Analysis, 10(S1), 15–50.
25. Rupakheti, M., Hiestermann, J. E., & Sager, A. (2019). Exploring the spatial-temporal characteristics of aerosol optical depth (AOD) in Central Asia based on the moderate resolution imaging spectroradiometer (MODIS). Environmental Monitoring and Assessment, 192(1), 80.
26. Said, M., Abdelrady, A., & Said, N. (2023). Air pollution and willingness to pay for health risk reductions in Egypt: A contingent valuation survey of Greater Cairo and Alexandria households. World Development, 172, 106361.
27. Scientific Reports. (2025). Air quality prediction using multi-source remote sensing data integration with hybrid deep learning framework. Scientific Reports, 15.
28. Shi, Y., Zhao, A., Yin, Z., & Zhao, A. (2023). Construction and evaluation of hourly average indoor PM2.5 concentration prediction models based on the random forest algorithm. Frontiers in Public Health, 11.
29. Tursumbayeva, M., Muratuly, A., Baimatova, N., Karaca, F., & Kerimray, A. (2023). Cities of Central Asia: New hotspots of air pollution in the world. Atmospheric Environment, 311.
30. Ullah, S., Ahmad, S., Nabi, G., & Khan, A. (2022). Willingness to pay for improved air quality in Pakistan: A contingent valuation study. Frontiers in Public Health, 10.
31. van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin, A., Sayer, A. M., & Winker, D. M. (2016). Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. Environmental Science & Technology, 50(7), 3762–3772.
32. Viscusi, W. K., & Masterman, C. J. (2017). Income elasticities and global values of a statistical life. Journal of Benefit-Cost Analysis, 8(2), 226–250.
33. Wei, Q., Chen, Y., Zhang, H., Jia, Z., Yang, J., & Niu, B. (2025). Simulation and prediction of PM2.5 concentrations and analysis of driving factors using

Global Economic Review: Journal of Economics, Policy, and Business Development
ISSN: 2980-5287
Volume 2, Issue 4, April 2026
Website: ecomindspress.com

interpretable tree-based models in Shanghai, China. Environmental Research, 270, 121003.
34. Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2000). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute.
35. World Bank. (2024). Air quality assessment for Tashkent and the roadmap for air quality management improvement in Uzbekistan. World Bank. https://www.worldbank.org/en/country/uzbekistan/publication/air-quality-assessment-for-tashkent
36. Zamani Joharestani, M., Cao, C., Ni, X., Bashir, B., & Talebiesfandarani, S. (2019). PM2.5 prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data. Atmosphere, 10(7), 373.
37. Zamin.uz. (2025, October 27). Uzbekistan has strengthened air quality standards. https://zamin.uz/en/world/165153-uzbekistan-has-strengthened-air-quality-standards.html
38. Zhang, J., Chen, Y., & Luo, X. (2019). Cost-benefit analysis of China's Action Plan for Air Pollution Prevention and Control. Frontiers of Engineering Management, 6(4), 524–537.