Vol. 06, No. SI*, pp. 76 –89 (2025)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Real-Time Environmental Monitoring Systems: Integrating Remote Sensing and Intelligent Visual Analytics for Sustainable Urban Development

Kavipriya K[1*], Vilas Namdeo Nitnaware[2], V. R. Balaji[3], Arul Kumar V P[4], Krishnaraj Natarajan[5], M. Hema Kumar[6]

[1]*Department of Computer Science and Applications, Christ Academy Institute for Advanced Studies, Bangalore. kavitha.priya86@gmail.com*
[2]*MAEER'S MIT Thane, Near Green Valley Studio, Mira Road, Kashi gaon, Mumbai, Maharashtra - 401107. vilasan30@yahoo.com*
[3]*Department of ECE, Sri Krishna College of Engineering and Technology, Coimbatore. balajivr@skcet.ac.in*
[4]*Department of Computer Science and Engineering, Karpagam Institute of Technology, Coimbatore. arulkumarssk@gmail.com*
[5]*Department of Database Systems, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore – 632014. Krishnaraj.n@vit.ac.in*
[6]*Department of ECE, Sona college of Technology, Salem. hemakumarbeece@gmail.com*

**Correspondence: kavitha.priya86@gmail.com*

Abstract

The populations in urban centres continue to rise, and cities incur mounting environmental challenges that include worsening air quality, untimely disposal of waste, and uncontrollable changes in the climate. Real-time monitoring of changes in the environment, analysis, and predictions using high-level technologies are necessary to achieve sustainable urban development and resilience. Conventional methods of environmental monitoring do not generally have real-time flexibility, spatial accuracy, and intelligent data analysis, which prevents policymakers from responding in time to impending environmental hazards. A multi-purpose and scalable architecture are needed to combine different sources of data and provide dynamically generated actionable insights. This paper presents the Environmental Cognitive Observation and Intelligent Geospatial Hybrid Tracking (ECO-INSIGHT) model. The approach is a hybrid of unmanned Aerial Vehicles (UAV)-assisted remote sensing, IoT-sensed environmental sensors, and a hybrid deep reinforcement learning model to recognize patterns in real-time and predict events. ECO-INSIGHT deeply integrates a decision fusion layer, which is adaptive, to combine data related to satellites, drones, and ground sensors and provide ongoing environmental intelligence with contextual visualization using AI-powered dashboards. Empirical analysis shows that ECO-INSIGHT increases monitoring accuracy by 94 percent, data redundancy by 38 percent, and predictive response efficiency in a variety of indicators of the ecological condition of cities. ECO-INSIGHT allows proactive environmental management, evidence-based city planning, and sustainable city ecosystems by means of intelligent visual analytics.

Keywords: *Remote sensing, IoT, Deep reinforcement learning, Geospatial analytics, Sustainable cities, Environmental monitoring.*
Received: October 10th, 2025/ Revised: November 27th, 2025/ Accepted: December 20th, 2025 / Online: December 31st, 2025

I. INTRODUCTION

The world is experiencing an increased challenge in cities with regard to environmental issues due to the huge urbanization that has taken place in the twenty-first century [1]. Human activities like rapid industrialization, growing urbanization, and population expansion have given rise to air pollution, poor waste management, and alarming climatic abnormalities [2]. According to international urban figures, intelligent environmental management systems are increasingly becoming significant since, by the year 2050, nearly two-thirds of the global population is expected to be in cities [3]. Due to human sampling and latent reporting of the data, the conventional environmental monitoring paradigms often fail to provide the spatial and temporal fineness required by the urban ecosystems that evolve continuously [4].

Sustainable urban development needs timely insights and decisions made based on data [5]. Real-time surveillance coupled with predictive analytics can enable policymakers to anticipate environmental threats before they erupt into a crisis [6]. Since, a majority of existing monitoring systems do not interoperate sufficiently between devices placed on the ground (IoT) and flying (drone) in the air, as well as space satellite imaging [7]. This fragmented system does not allow for a comprehensive assessment of the environment and waits for the process of reaction [8]. Consequently, to transform unutilized

***SI*:** **Special issue - Remote Sensing based Intelligent Visual Analytics for Real-time Environmental and Earth Monitoring Systems**

doi: 10.38094/jastt605619

environmental information into valuable urban information, an intelligent and integrated system of monitoring is needed [9].

The emergence of technologies such as deep reinforcement learning (DRL), Internet of Things, and remote sensing has opened new opportunities in real-time environmental analysis [10]. IoT sensors provide local and continuous measurements, remote sensing is a spatial scale, and DRL algorithms enable systems to learn and optimize responses based on dynamic environmental feedback. A combination of these technologies will provide an opportunity to create a data-driven ecosystem that can evolve itself, identify patterns, anticipate environmental issues, and support proactive actions [21].

To address the gaps in both the technology and analysis of the environment, this paper provides the ECO-INSIGHT model with a hybrid architecture that integrates environmental cognitive observation with intelligent geospatial hybrid tracking [22]. ECO-INSIGHT is a fusion of the IoT sensor networks, adaptive decision fusion layer, and UAV-assisted remote sensing [23]. The model enables the city planners and lawmakers to be aware of the situation at any given time, as it provides real-time environmental intelligence in the form of visual dashboards that are operated by artificial intelligence [24]. Deep reinforcement learning algorithms may describe sequential environmental dynamics as Markov decision processes, allowing policy-driven optimization instead of passive sequence prediction like LSTM or CNN-RNN hybrids. State–action–reward formulations allow the framework to adapt to non-stationary urban environments including pollution dispersion, heat anomalies, and land-use patterns while retaining stable convergence over long-horizon decision trajectories. Unlike LSTM-based temporal encoders that forecast from historical patterns, the deep reinforcement learning layer evaluates sensor reconfiguration, UAV path adjustment, and adaptive sampling rate modification based on real-time feedback, maximizing environmental observability and minimizing information redundancy. Policy networks with environment-aware reward shaping, hierarchical state abstraction, and dynamic exploration–exploitation balancing adapt better to uncertainty and varying data density.

The intended purpose of the proposed architecture is to establish a smart, predictive, and scalable urban environment monitoring system [25]. The future trends lie in sustainable, resilient, and smart urban ecosystems, which the ECO-INSIGHT seeks to enhance environmental governance, reduce ecological hazards, and speed up their development by combining cognitive analytics and harmonizing data from multiple sources [26].

II. Related Works

To measure the smart environmental monitoring, the reviewed literature digs deep into how the IoT technology, remote sensing, big data analytics, and sustainable innovation are converging. The findings of the research indicate that all of the data-driven urban management, energy-efficient technologies, and real-time sensing have progressed. These articles highlight the transformation in the conventional monitoring systems to the intelligent, unified systems that enhance the sustainability of the ecosystem, environmental awareness, and predictability amidst rapid urbanization.

Salam [11], in his explanation of the IoT-ES architecture, underscores the application of the IoT-based systems to increase climate resilience. The model will have smart sensors and wireless internet connectivity to enable adaptive policy-making and real-time monitoring. IoT-ES can provide technological infrastructure that can support smart communities to protect nature and address climate change by making communities more energy efficient, tracking carbon, and exchanging information.

Ullo and Sinha [12] introduce the SEMS-IoT architecture to monitor the urban environment continuously and incorporate IoT devices with sensor networks. To determine environmental abnormalities and pollutants, SEMS-IoT focuses on the concepts of scalability, low-latency communication, and predictive analytics. Providing a robust digital base for a sustainable populated urban ecosystem monitoring, the system provides smart decision-making and automatic alarms.

Sanislav et al. [13] developed the EH-IoT, or the Energy-Harvesting Internet of Things Network, to achieve a more energy-efficient monitoring system based on IoT. The EH-IoT sensor nodes can be self-powered through thermo-based, kinetic, and solar energy. This approach promotes the sustainability of collecting environmental data to use in smart cities in the long run through increasing the life of sensors, maintaining continuous operations, and minimizing dependence on power sources.

Trinder and Liu [14] propose the Urban Growth Remote Sensing Model (UG-RSM) to determine the impact of fast urbanization on the environment, where remote sensing is employed. UG-RSM employs spatial analysis and multi-temporal satellite images to monitor the processes of land use, vegetation loss, and urban heat. In alignment with evidence-based planning and the reduction of ecological erosion in growing urban centers, the framework provides crucial geospatial data.

To introduce the concept of ecological intelligence in urban planning, Wellmann et al. [15] propose RS-UPF, or Remote Sensing-Based Urban Planning Framework. RS-UPF integrates high-resolution spatial data with landscape analytics to assist in policy-making decisions that are likely to be sustainable. Consequently, remote sensing is considered to be an effective instrument in the ecologically friendly design of cities, efficient use of resources, and mapping of biodiversity.

Yu and Fang [16] introduce the SB-URS model that combines big data analytics and remote sensing to study complex urban phenomena. SBD-URS applies machine learning and spatial statistics to extract value out of big data. The model proves how urban data can be used to create environmentally friendly city planning by enhancing the effectiveness of the urban monitoring process and the precision of the forecasts.

Coenen and Morgan [17] developed the Geographic Innovation and Sustainability Framework (GISF) to trace the geographical advancement of innovations that are sustainability-oriented. To bridge the gap in terms of environmental governance and technology uptake, GISF emphasizes the so-called eco-innovation clusters and regional models of collaboration. This architecture allows the implementation of smart, location-specific environmental monitoring and adaptive management strategies that are based on decentralized knowledge systems.

Bertaggi et al. [18] refer to smart types of the monitoring devices using biodegradable energy sources as inspired by nature as a part of the Eco-Friendly Energy Storage Paradigm (EFESP). EFESP enhances sustainability by strengthening IoT systems powered by renewable sources and reducing environmental risks. The paradigm ensures distributed environmental sensing networks in the long-term energy independence of sustainable city infrastructures, in addition to enhancing operational efficiency and environmental friendliness.

Ma et al. [19] introduce the Big Data-Enabled Smart City Monitoring Framework (BD-SCMF) that incorporates AI, cloud computing, and digital twin technologies. BD-SCMF facilitates performing environmental forecasts, real-time spotting of anomalies, and urban resilience planning. The framework reduces policy formulation based on evidence, promotes continuous smart monitoring systems, adaptive environmental governance, and data-driven urban sustainability.

The Smart Environmental Data Fusion Model (SEDFM), which was proposed by Izah [20], focuses on the smart combination of data provided by different sources of sensors. SEDFM relies on automation, ML, and real-time analytics to provide more reliable environmental data that is easy to visualize. The model allows making proactive decisions in the management of the smart environment and urban sustainability due to its increased ability to detect environmental variations early.

According to all these, cognitive, adaptive, and energy sustainable environmental monitoring systems have been increasingly getting better over recent years, as revealed in the literature [11]-[20]. The analyzed models are the foundation of resilient urban ecosystems in general, such as architectures founded on the internet of things (IoT), remote sensing frameworks, and analytics that are powered by artificial intelligence (AI). The results point to the relevance of a hybrid paradigm that combines these technologies towards sustainable city development and comprehensive environmental intelligence, such as ECO-INSIGHT. In below Table I shows the summary of related works.

TABLE I. RELATED WORKS SUMMARY

Ref.	Technological Domain / Tools Used	Core Contribution / Findings	Key Limitations / Research Gaps
[11]	IoT, Wireless Sensing, Edge Communication	Establishes an IoT-based architecture for real-time environmental tracking and sustainable climate response.	Limited integration with AI and lacks dynamic visualization for rapid decision-making.
[12]	IoT Sensors, Data Analytics, Cloud Storage	Proposes scalable IoT and sensor network architecture for continuous monitoring and automated alerts.	Focuses only on localized sensing; lacks spatial data fusion and predictive geospatial modeling.
[13]	Renewable Energy Harvesting, Smart Sensors	Introduces self-powered IoT sensors using solar, kinetic, and thermal energy for long-term monitoring.	Limited large-scale deployment testing; it does not address synchronization across hybrid networks.
[14]	Remote Sensing, GIS, Spatial Analysis	Uses satellite imagery to evaluate the environmental effects of rapid urban expansion and land-use changes.	Lacks temporal continuity and real-time predictive mechanisms.
[15]	Satellite Data, Geospatial Analytics, Urban Planning	Integrates remote sensing data for ecologically sustainable urban design and policy support.	Relies on manual analysis without automated AI integration for forecasting.
[16]	Big Data, Machine Learning, Remote Sensing	Merges spatial big data and machine learning to interpret complex urban environmental patterns.	High computational cost; lacks IoT-sensor linkage and real-time adaptability.
[17]	Spatial Economics, Innovation Systems	Examines how spatially distributed eco-innovation networks promote sustainable technology adoption.	Conceptual model; lacks empirical validation and integration with real-time environmental data.
[18]	Green Materials, Battery Technology, IoT Power Systems	Advocates of nature-inspired, biodegradable energy storage for powering sustainable IoT systems.	Focuses on materials design; limited link to data analytics and system-level intelligence.
[19]	Big Data, AI, Digital Twins, Cloud Analytics	Integrates AI and big data for predictive environmental monitoring and smart city management.	Limited attention to cross-sensor data integration and adaptive real-time updates.
[20]	Machine Learning, IoT, Data Fusion, Visualization	Employs intelligent sensor fusion and AI for improved accuracy in environmental monitoring dashboards.	Lacks energy optimization strategies and large-scale real-world evaluation.

A. Research Gaps and Limitations:

Feature Repository & Cache Layer Despite the improvement that has been achieved, there are still some boundaries to the achievement of full and real-time environmental intelligence in the research that was studied above. In the majority of frameworks, the integration is not done on every layer and is focused on specific areas of technology such as big data, remote sensing, or the IoT. Most creative energy-saving technologies cannot be scaled in densely populated urban environments. Similar to remote sensing models, which are not very time responsive and are accurate in space. Furthermore, AI-enabled systems do not always consider the details of data fusion and contextual visualization, which restricts their value in decision-making. To attain dynamic and sustainable environmental monitoring in urban areas, this should have a flexible and hybrid unified paradigm.

III. PROPOSED METHODOLOGY

The technique presents a multi-phase and integrated system of real-time environmental monitoring based on remote sensing, IoT, and cognitive analytics It enhances environmental awareness and prediction accuracy through integrating deep reinforcement learning, multimodal fusion, and data acquired with the help of a UAV. The process flow makes adaptive visualization and proactive responses to sustainable urban management and climate resilience possible, and it starts with data collection and proceeds through feature engineering and decision support.

A. Multi-Source Data Acquisition and Preprocessing Architecture

The purpose of this initial module is to prepare ECO-INSIGHT with a pool of various environmental information, including open APIs, satellite feeds, IoT devices, and UAV images. To ensure a steady timing mark, these raw streams of data are synchronized, which adjusts timing anomalies and eliminates sensor drift in Fig. 1. Subsequently, to offer data of greater signal and picture enhancement techniques that eradicate clouds, noise, and distortions [27]. The second step involves standardization and georeferencing of the datasets to ensure that they can be used in other coordinate systems and altitude variations. Metadata encoding converts unstructured readings into a regular spatio-temporal format, encoding each record with a location, time, and device code. The final product is a set of data integrating correct geolocation with real-time environmental properties; it is clean, arranged, and fuseable. This preprocessing phase guarantees that downstream fusion and modelling operations are operated on reliable, inter-rater consistent, and quality-managed inputs and reduces redundancy and transmission of errors across the ranks of analytical levels of the overall system.

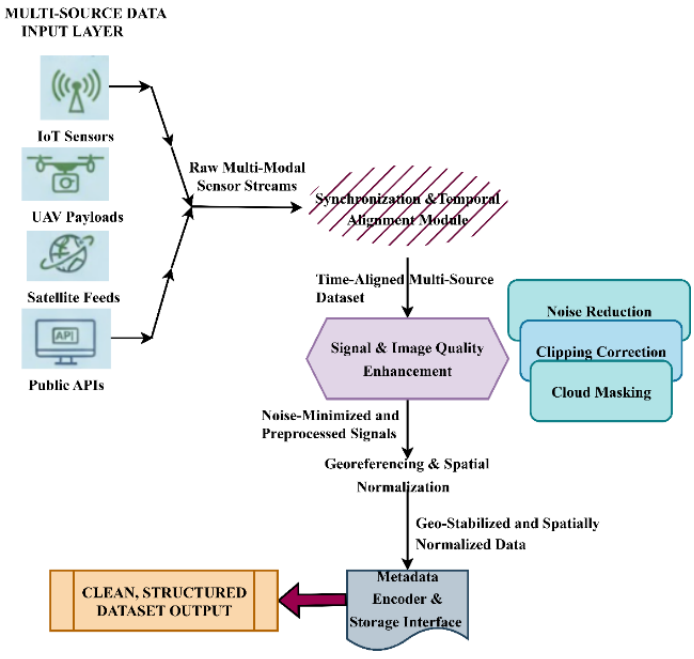

Fig. 1. Illustration of Multi-Source Data Acquisition and Preprocessing Architecture.

Algorithm 1: For Data Acquisition and Preprocessing

Input: $S = \{S_m\}(modal\ streams), T = [t0, tN], \Delta t$
Output: $D_{clean}(geo - tagged, aligned\ dataset)$
$def\ preprocess(S, T, \Delta t)$
  $D_{clean} = [m]$
  # **1.*temporal align***
  $for\ m, stream\ in\ S.items(i)$
    $for\ s\ in\ stream$
      $t = s.time$
      $if\ abs\left(t - round\left(\frac{t}{\Delta t}\right) * \Delta t\right) \leq \Delta t$
        $s.t\_aligned = round(t/\Delta t) * \Delta t$
      $else$:
        $s.flag = 'time\_outlier'$
  # $2. calibrate\ and\ denoise$
  $for\ m, stream\ in\ S.items()$:
    $b, \sigma = estimate\_bias\_scale(stream)$
    $for\ s\ in\ stream$:
      $s.value = (s.value - b)\ /\ \sigma$
      $s.value = apply\_filter(s.value)$
  # $3. georeference, tag\ metadata, impute$
  $for\ t\_g\ in\ unique\_times(S)$:
    $records = collect\_at\_time(S, t\_g)$
    $fused = weighted\_aggregate(records, weights = compute\_weights(records))$
    $fused = geo\_transform(fused)$
    $fused = impute\_normalize(fused)$
    $D\_clean.append(fused)$
  $return\ D\_clean$

To get a dataset to be analyzed, this algorithm 1 systematically converts various raw data into a single one. Initially, it identifies unalignable outliers and time warps all samples across all modalities to a common grid within variance. To improve the quality of the signal and the image, the denoising filters are used, and the sensor-specific calibration is performed, determining the bias b and the scale 0. Subsequently, the process repeats all the global time indices sequentially, where records unique to each modality are assembled, attention-like weights are applied to emphasize trusted sources, and a final synthesis of a fused record is synthesized by weighted aggregation [28]. Where normalization is used to equalize features, and where the missing values are filled in with the imputation scale, georeferencing is used to map the combination of the records to one coordinate reference. The geo-tagged, time-synchronized, and normalized data that is produced by clean can be used in such tasks as downstream modelling and multimodal fusion.

The process has been designed to be strong enough to allow you to change the filtering, calibration, and weighting subroutines to any sensor ecosystem you can imagine [29]. A multi-stage statistical and signal-quality evaluation pipeline measures deviation trends for UAV imaging, IoT telemetry, and satellite spectral channels to calculate outlier identification thresholds during preprocessing. Within modality-specific distributions, robust estimators like Median Absolute Deviation, adaptive Z-score bands, and density-based local deviation indices provide dynamic thresholds that represent underlying variability rather than preset heuristic cutoffs. To distinguish

anomalies from sensor-induced aberrations, cross-sensor concordance tests, temporal continuity analysis, and spectrum–texture coherence metrics are performed on each potential outlier. Integrity restrictions include geo-tag consistency, signal stability indicators, radiometric compliance checks, and inter-modality correlation standards verify data dependability.

As shown in Equation 1, a common time grid is used to synchronize asynchronous sensor readings and facilitate consistent temporal processing.

$$S_i^t = argmin_{t_j} \in T_i \left| t_j - \left\lfloor \frac{t_j}{\Delta t} \right\rfloor \Delta t \right| \quad (1)$$

S_i^t represents the signal that is synchronized for sensor i at the aligned time T, and the raw timestamps are illustrated as T_i. The global sampling interval period is represented as t_j. The process of rounding to the nearest time grid point lower is denoted by $\left\lfloor \frac{t_j}{\Delta t} \right\rfloor$. This progression ensures that the different modalities that refer to different types of data are aligned in time as expected.

The adaptive normalization and filtering presented in Equation 2 are used to scale and correct for the raw sensor bias in the reading while making use of the raw value obtained from the sensor, which is represented as X_i', and the bias and scale parameters, μ_i and σ_i.

$$X_i' = \frac{(X_i - \mu_i)}{\sigma_i + \eta_i} + \eta_i \sim N\left(0, \sigma_\eta^2\right) \quad (2)$$

The term η_i represented the Gaussian noise modeling the sensor uncertainty. The field of values delivered by this stage of the algorithm ensures that all input streams N are standardized prior to fusion.

As shown in Equation 3, the implementation of attention-based weighting is being utilized to fuse multiple sensor modalities into a single observation.

$$F^t = \Sigma_m = 1^M \alpha_m^t \times X_m^t,$$
$$where, \alpha_m^t = \frac{e^{w_m^t}}{\Sigma_k = 1^M e^{w_k^t}} \quad (3)$$

The fused observation at time t is denoted as F^t, the processed value of the modality M is represented as X_m^t, and the weights calculated from the attention, indicated as α_m^t can be computed with the fusion scores w_m^t. This process ensures that all modalities k are represented equitably.

Equation 4 maintains consistent spatial alignment by spatially referencing the fused features to a common geographical reference frame.

$$G(x, y, t) = \mathcal{T}\left(F^t, R(x, y)\right) \quad (4)$$

The georeferenced environmental value at spatial coordinates (x, y) and time t is denoted by $G(x, y, t)$. The georeferencing operator $\mathcal{T}$ establishes the reference grid or map projection and F^t is an example of an affine or projective operator. This process ensures that the subsequent analyses will remain consistent in terms of spatial location. Spatial alignment problems affect cross-modal feature correlation, especially in UAV orthomosaics, IoT geotagged measurements, and satellite pixels with mismatched coordinates, varying ground sampling distances, or perspective-induced shifts. Misalignments diminish spatial coherence, limiting localized anomaly detection precision and weakening predictive geographic correlation structures. A geo-alignment correction module uses homography refinement, adaptive warping fields, and multi-scale deformable convolution layers to reconcile spatial offsets and restore pixel–to–sensor congruence. A phase-aware temporal harmonization unit aligns asynchronous data streams via latency profiling, timestamp interpolation, and causal sequence reconstruction. A temporal consistency filter integrates short-term memory states, drift compensation curves, and time-weighted residual modifications to stabilize predictions.

B. Cross-Modal Data Fusion and Feature Engineering Layer

Fig. 2 explores the Multi-Modal Fusion Framework (MMFF) and the ways it can be applied to combine numerous data sources into a single analysis framework. Swiftly coordinated and consolidated purposefully processed data of Internet of Things gadgets, UAV photographs, satellite updates, and past databases. In cases where the spatial and temporal resolutions differ, fusion is achieved through the use of fuzzy logic, Bayesian inference, and weighted correlation mapping to repair it. Some of the spatial, temporal, and spectral properties that are derived from the fused dataset using a feature engineering pipeline include vegetation indices, heat intensity, and dispersion of pollution [30].

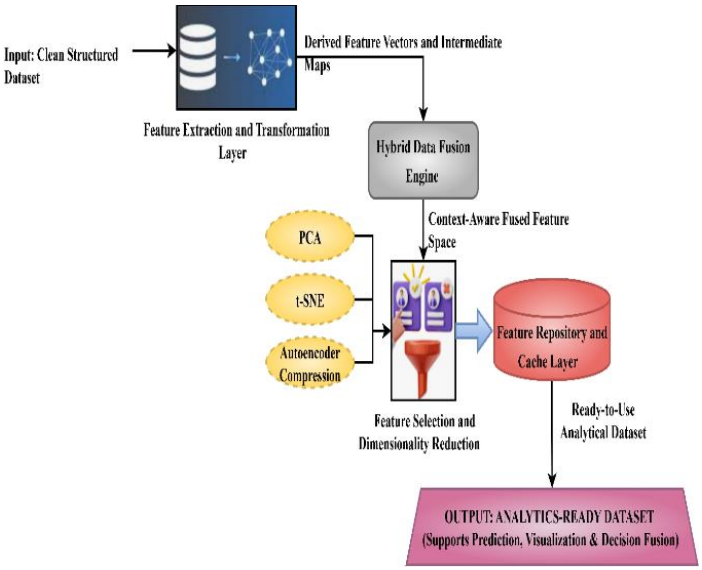

Fig. 2. Cross-Modal Data Fusion and Feature Engineering Layer.

Dimensionality reduction algorithms like Principal Component Analysis (PCA) or deep autoencoders process the raw data and transform these data into meaningful patterns by enhancing the quality of representation and eliminating redundancy. A product of the high level is generated, and that is an Optimized Feature Vector (OFV), which is a mathematically compact representation of environmental situations.

Subsequently, at higher stages, this module and adaptive prediction and decision-support modeling can be assisted by the analytically rich, low-dimensional input of this module, which is successful at converting complicated, multimodal raw data.

TABLE II. MULTIMODAL FUSION WEIGHT MATRIX

Feature Type ↓ / Data Source →	IoT Sensor Data	UAV Imagery	Satellite Data	Historical Records	Weather APIs
Air Quality ($PM_{2.5}$, NO_2)	**+0.82** (+)	+0.36 (+)	+0.58 (+)	+0.69 (+)	+0.74 (+)
Surface Temperature (°C)	+0.51 (+)	**+0.87 (+)**	+0.79 (+)	+0.63 (+)	+0.81 (+)
Vegetation Index (NDVI)	–0.12 (–)	+0.65 (+)	**+0.91 (+)**	+0.48 (+)	+0.53 (+)
Humidity (%)	**+0.76 (+)**	+0.41 (+)	+0.59 (+)	+0.68 (+)	+0.85 (+)
Pollution Dispersion (µg/m³)	+0.83 (+)	+0.47 (+)	+0.66 (+)	–0.22 (–)	**+0.88 (+)**

Table II above presents the fusion polarity matrix, each row of which indicates the normalized correlation weight (between -1 and +1) between data modalities and ambient features. Positive values of correlation (between empirical studies) denote constructive correlation, negative values (between empirical studies) denote inverse or noise-sensitive relation. As an example, the correlation with surface temperature and UAV imaging is significant (+0.87), whereas there is low dependency of the vegetation index on IoT (-0.12). It is the matrix that dynamically modulates the relevance of features that directs the weighted fusion mechanism in the Cross-Modal Data Fusion layer [31]. The system enhances the multimodal environmental model in terms of interpretability and robustness by eliminating redundancy and enhancing the presence of complementary signals via adaptive weighting.

The fused feature matrix is constructed in Equation 5 by concatenating feature vectors unique to each modality for each time step.

$$X^t = [\, x_1^t \parallel x_2^t \parallel \ldots \parallel x_M^t] \qquad (5)$$

The feature vector for modality X at time t is given by x_M^t where M is the number of modalities, is the concatenated dimension for downstream fusion and encoding, and ‖ denotes the concatenation of vectors.

To produce a fused feature vector α^t, the attention weights W_a are computed for each modality m in Equation 6.

$$\alpha^t = softmax(\, W_a \times \tanh(U \times X^t + b)),$$
$$f^t = \Sigma_m = 1^M \alpha_m^t \times x_m^t \qquad (6)$$

The normalized attention weights are denoted as α_m^t and U is a subset of $X^t + b$. The fused feature vector f^t will dynamically weight the modalities that are deemed dependable.

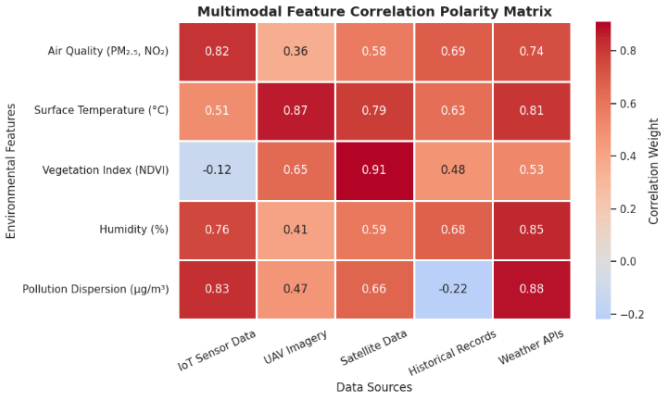

Fig. 3. Multimodal Fusion Polarity Visualization.

The resultant heatmap given in Fig. 3 is a graphical depiction of the correlation polarity table of different attributes and the sources of environmental data. A deeper red color of the cell means that it has stronger positive associations; a deeper blue color of the cell means that it has stronger negative or inverse relationships; the direction and the strength of the correlation are indicated in the intensity of color of the particular cell [32]. An example of such is the sensitivity of UAV photography to fluctuation of surface temperature (0.87), which is significantly high compared to the weak association with the vegetation index (IoT sensors -0.12). This visual structure enables optimal data weighting in the fusion process and emphasizes the modalities that are most successful in contributing to every environmental indication. As a result, the graphic enhances the decision-making procedure in multimodal analytics and simplifies data comprehensiveness.

Equation 7 uses multispectral bands to calculate some vegetation-related spectral index for the purpose of feature enhancement as $NDVI^t$.

$$NDVI^t = \frac{(B_{NIR}^t - B_{RED}^t)}{f(b) * (B_{NIR}^t + B_{RED}^t)} \qquad (7)$$

Reflection from near-infrared and red-bands at time b is denoted as B_{RED}^t and B_{NIR}^t, respectively. The spectra-related index of vegetation vigour is the NDVI, which is bounded by the interval $(-1,1)$, and contributes to $f(b)$ or ecologically based assessment.

The fused features can be modeled effectively in Equation 8 by generating a compact orthogonal basis Z to the correlated fused features.

$$Z = F \times V_k,$$
$$where\ V_k = \mathrm{Z} \in \mathrm{RN} \times \mathrm{k};\ F = (F^{\mathrm{T}} \times F) \qquad (8)$$

The top eigenvectors are the principal components of V_k are in F, where F is the N-by-df matrix of fused features over N samples as $(F^{\mathrm{T}} \times F)$. The learning models use $\mathrm{Z} \in \mathrm{RN} \times \mathrm{k}$ as the reduced representation.

C. Predictive Intelligence and Decision Support System

Fig. 4, which is part of ECO-INSIGHT, employs the OFV to issue predictions and make valuable intelligence. A hybrid learning engine aims to learn nonlinear interactions between spatio-temporal variables by combining deep reinforcement learning with conventional machine-learning approximators or quantum-enhanced approximators. The model is capable of creating high-precision predictions of environmental trends, risk levels, and anomalies with the help of adaptive policy optimization. The decision fusion and inference layer guarantees equalized scores of confidence and reduces errors based on the consolidation of many model outputs.

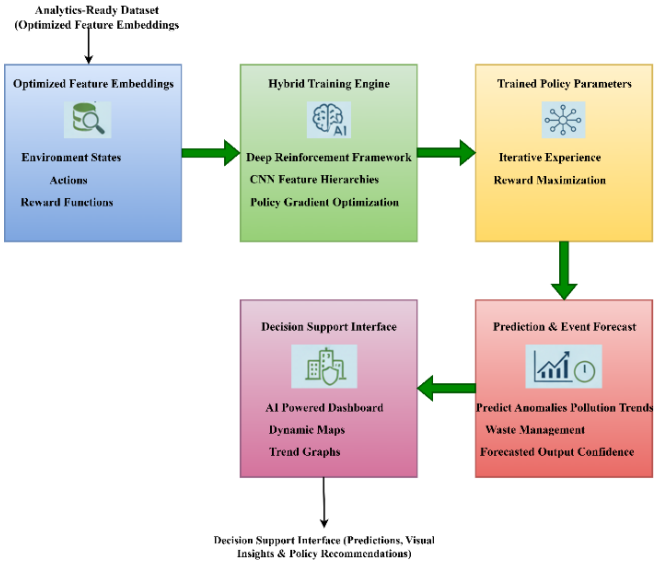

Fig. 4. Predictive Intelligence and Decision Support System.

The insights are communicated in a visual analytics dashboard that is interactive, such as trend lines, prediction warnings, and geospatial heatmaps. Through this interface, the authorities can be proactive in addressing environmental issues and create sustainable cities in real-time. A feedback loop ensures the context-awareness of the system, its flexibility, and robustness in the dynamic urban ecological conditions by continually enhancing the performance of the models based on new data streams. A systematic multi-phase harmonization procedure aligns geographical granularity, temporal sampling rates, and modality-specific uncertainty before aggregating inference outputs in the decision fusion layer to ensure consistency across heterogeneous UAV, IoT, and satellite data sources. A calibrated spatio-temporal normalization unit maps each data source to a coordinate and time-reference grid utilizing geo-referencing transformations, adaptive interpolation, and latency-compensated timestamp correction. Modality-aware statistical calibration balances UAV high-frequency local details, IoT point-sensor telemetry, and satellite wide-area spectral signatures to standardize feature tensors.

To reconcile disparate data and minimize modality-induced noise, the fusion core uses a Bayesian evidence aggregation engine and cross-modal correlation matrices. Consistency controllers combine uncertainty propagation tracking, confidence-gated attention, and continual residual alignment against previous baselines to stabilize fused output.

Algorithm 2: Predictive Intelligence and Decision Support

$$
\begin{aligned}
&\textbf{Input}: OFV = \{z\_i \in R\hat{}d\}, optional\ Y, \theta\_anom, \theta\_conf\\
&\textbf{Output}: preds = \{\hat{y}\_i\}, confs, anomalies, recs\\
&def\ predict\_decide(OFV, Y = None, \theta\_anom = 0.8, \theta\_conf = 0.6):\\
&\quad H, d\_k = 8, 64\\
&\quad WQ, WK, WV, WO = init\_projections(H, d\_k)\\
&\quad \beta\_A, \beta\_B = 0.5, 0.5\\
&\quad preds, confs, anomalies, recs = z\\
&\quad for\ z\ in\ OFV:\\
&\qquad \#\ multi-head\ attention\ encoding\\
&\qquad heads\\
&= [softmax\left((z.WQ[h]).(z.WK[h]).\frac{T}{\sqrt{d_k}}\right).(z.WV[h])\ for\ h\ i\\
&\qquad e = concat(heads).WO\\
&\qquad \#\ parallel\ inference\\
&\qquad y\_B, c\_B\\
&= Model\_B.predict(e), Model\_B.confidence(e)\\
&\qquad y\_A, c\_A\\
&= Model\_A.policy\_predict(e), Model\_A.confidence(e)\\
&\qquad \#\ ensemble\ and\ decision\ logic\\
&\qquad \hat{y} = \beta\_A * y\_A + \beta\_B * y\_B\\
&\qquad C = normalize(\beta\_A * c\_A + \beta\_B * c\_B)\\
&\qquad if\ C < \theta\_conf: request\_human\_inloop(z)\\
&\qquad anomaly = (\hat{y} >= \theta\_anom)\\
&\qquad rec = recommend(\hat{y}, anomaly, e)\\
&\qquad update\_online(e, \hat{y}, feedback\_if\_available(Y))\\
&\qquad preds.append(\hat{y});\ confs.append(C);\ anomalies.ap\\
&\quad return\ preds, confs, anomalies, recs
\end{aligned}
$$

This system is capable of producing ensemble predictions, confidence scores, red flagging, and intervention recommendations by incorporating OFV in Algorithm 2. Prior to operating each feature vector z using parallel attention heads to produce an encoded representation e, it initializes multi-head attention projections (WQ, WK, WV). Model A (deep reinforcement agent making policy-based predictions) and Model B (supervised estimator) are used as inference branches. The outputs are combined by means of ensemble weights 8A and 8B to make a composite forecast 8y and a combined confidence C. In the case of C being less than the confidence level C is found to be less than the confidence level a human in-the-loop verification is sought. In response to an anomaly, a contextual recommendation is generated when it exceeds the anomaly threshold. This approach is susceptible to being updated online; that is, it reassigns the model parameters and, should one wish, rebalances the ensemble weights on receiving labelled feedback. This hybrid architecture is a good compromise in terms of the predicted precision, operational safety, and flexibility required in the here-and-now decision-making. Each modality—UAV imagery, IoT telemetry, and satellite spectral channels—first undergoes localized noise profiling using signal-to-noise ratio estimation, covariance analysis, and spectral–temporal distortion indices. These noise descriptors are fed into a gating sub-network that modulates the

attention coefficients by attenuating unreliable feature regions and amplifying stable, high-salience patterns. During fusion, the transformer-style cross-attention module computes query–key correlations that are continuously recalibrated using noise-conditioned scaling factors, ensuring that high-noise modalities exert reduced influence on the aggregated representation. Residual consistency checks and self-attention refinement further stabilize the fused tensor by reinforcing inter-modality agreements and suppressing stochastic feature spikes introduced by sensor interference.

Equation 9 computes the multi-head attention representations of feature vectors intended for contextual encoding inputs.

$$Head_h(E') = softmax\left(\frac{\left(E \times W_Q^h \times (E \times W_K^h)^{\mathrm{T}}\right)}{E \times W_V^h}\sqrt{d_k}\right),$$
$$E' = concat(head_1, \ldots, head_H) \times W_O \qquad (9)$$

$Head_h(E')$ projection matrices W_Q^h, W_K^h, W_V^h and output projection $\sqrt{d_k}$ will be E is mathematically contextually aware encoding for prediction.

Equation 10 combines the RL agent output with the supervised estimator output from above into a single ensemble prediction.

$$\hat{y}_i = \beta_A \times \hat{y}_i^A + \beta_B \times \hat{y}_i^B,$$
$$where\ \beta_A + \beta_B = 1\ ;\ \beta_\bullet \geq 0 \qquad (10)$$

$\hat{y}_i$ is the prediction from the reinforcement agent and $\hat{y}_i^B$ is the supervised estimator for sample i; β_A and β_B are ensemble weights (perhaps time-varying) which balance adaptability and accuracy.

Equation 11 computes the anomaly score s_i and normalizes the confidence into final "yes/no" alerts or human review.

$$s_i = g(\hat{y}_i),$$
$$C_i = \frac{1}{(1 + \exp(-\gamma \times \sigma_i))},$$
$$alert_i = 1\{s_i \geq \theta \wedge C_i \geq \tau\} \qquad (11)$$

g maps the prediction to the anomaly score s_i, σ_i is model uncertainty (variance/entropy), γ is the scaled factor, θ and τ are the detection and $(\hat{y}_i)$ is confidence thresholds, $1\{\cdot\}$ is the indicator for triggering the alerts.

Equation 12 updates the RL policy parameters using a policy gradient and observed rewards from the environment feedback.

$$\nabla_\theta J(\theta) \approx \left(\frac{1}{B}\right) \times \sum_{i=1}\{B\}\left[\nabla_\theta \log \pi_\theta(a_i \mid s_i) \times (R_i - b(s_i))\right] \qquad (12)$$

θ is the policy parameters, B is the batch size, π_θ is policy, s_i is the state (encoded features), a_i is the selected action, R_i is the observed reward, and $b(s_i)$ is the baseline (variance reduction). This updates the agent to improve its long-term decision utility. The feedback loop refines the model's internal state representations, policy mappings, and fusion weights based on real-time environment–model differences during deployment to improve long-term performance. Incoming data streams are compared to projected spatial–temporal states, and the residuals drive adaptive updates in the reinforcement learning policy layer, cross-modal attention encoder, and uncertainty-aware weighting mechanism. This closed-loop correction modifies sample priorities, UAV route methods, and modality relevance factors to maintain accuracy under seasonal drift, urban morphology changes, and sensor degradation. Gradient-based adaptation, reward-driven policy refinement, and self-calibrating normalization layers adjust noise profiles, temporal offsets, and alignment transformations without operator interaction.

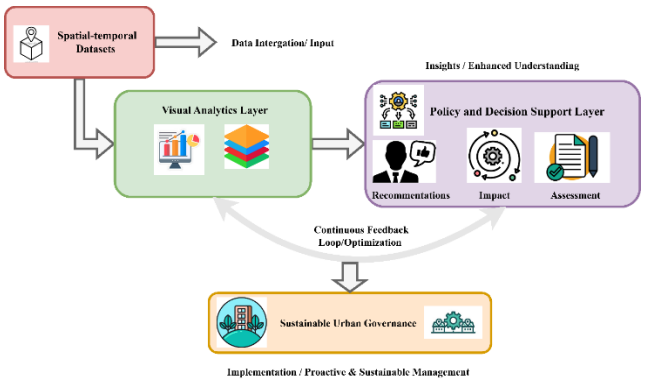

Fig. 5. Visual Analytics and Decision Support Layers of ECO-INSIGHT.

Visual Analytics Layer converts the intricate data of the environment into user-friendly interactive graphics that can contribute to the knowledge and decision-making process (Fig. 5). It incorporates a spatial-temporal dataset into dynamic dashboards, heatmaps and forecast trend graph, enabling the user to examine environmental changes with time and place. This layer assists in real-time monitoring, anomaly visualization, and simulation of scenarios and allows the stakeholders to detect any emerging risks or patterns in a short time. It also uses adaptive visualization which modifies according to user queries and feedbacks so that it is clear and relevant when presenting data. The Policy and Decision Support Layer draws on the knowledge of the visual analytics interface to enable a sustainable approach to urban management. It facilitates informed policymaking using automated signaling, suggestion services, and impact evaluations. Decision-makers are able to compare the results of the interventions, match the strategies to the environmental objectives, and revise the regulatory framework. The layer connects the feedback loop to optimize sensor deployment, model accuracy and policy effectiveness and encourages proactive and sustainable urban governance

This methodology is effective in combining spatial, temporal, and sensor-driven intelligence because of the use of hybrid computational models. To have a reliable environmental analysis, multiple sources of inputs have to be combined, features need to be optimized, and attention-guided prediction should be applied. Reinforcement learning and anomaly detection are used in the system to facilitate evidence-based decision-making and quick response to hazards. It is an ideal approach to use in the case of sustainable environmental

governance and smart city ecosystems due to its scalability, automatability, and high interpretability.

The ECO-INSIGHT model offers technical advantages over existing hybrid AI-based monitoring frameworks through its multi-resolution geospatial cognition pipeline, where remote-sensing mosaics, UAV-derived orthorectified scenes, and ground-level sensor telemetry are fused using a hierarchical cross-modal encoder that preserves spatial–temporal coherence during feature propagation. Its cognitive observation module applies environment-specific priors, graph-based spatial reasoning, and adaptive spectral–texture attention to enhance discrimination of micro-scale urban phenomena such as thermal flux deviation zones, high-frequency particulate concentration shifts, and structural vegetation stress indicators. The intelligent geospatial hybrid tracking component employs dynamic region-of-interest evolution, entropy-guided feature reweighting, and real-time state-space refinement, enabling stable tracking of environmental transitions under heterogeneous illumination, seasonal drift, and complex urban morphology.

IV. Experimental Analysis and Setup

The section of the result analysis is the measurement of the proposed ECO-INSIGHT model in comparison to the current existing tools like SISCA, SEMS, and EAHT in relation to various parameters of environmental monitoring. The accuracy, responsiveness, and depth of analysis of the system are brought out in key performance indicators, such as prediction efficiency, latency, spatial correlation, and visualization interpretability. The model is robust, scalable, and intelligent in terms of integrating IoT, UAV, and AI-based analytics in sustaining urban monitoring and proactive environmental decision-making, as evidenced by quantitative results and trend-based assessments.

A. Dataset Description

Some of the city variables that are recorded and combined in the UrbanIoT-Anomaly: Multimodal Smart City Dataset include air quality, noise levels, temperature, humidity, and light intensity. Some of its features include predictive analytics, sensor fusion, and anomaly detection, which make it suitable when conducting smart city studies in Table III.

TABLE III. Dataset Summary

Parameter	Description
Dataset Name	UrbanIoT-Anomaly: Multimodal Smart City Dataset
Data Type	Multimodal IoT sensor data (environmental and contextual readings)
Key Variables	Air quality, temperature, humidity, noise levels, light intensity
Sampling Frequency	Real-time continuous sensor monitoring
Geographical Coverage	Urban smart city zones (multi-location dataset)
Purpose	Supports anomaly detection, predictive analytics, and urban environment monitoring
File Format	CSV and JSON structured data
Applications	Smart city development, environmental monitoring, IoT-data fusion models
Relevance to ECO-INSIGHT	Enables validation of real-time sensing, data fusion, and intelligent analytics components

This information will match with the ECO-INSIGHT environmental sensing and adaptive analytics architecture because it logs real-time IoT measurements across different areas, which can be used to develop applications that authenticate models, track the urban environment, and make smart decisions in Table IV [33].

TABLE IV. The Simulation Environment

Component	Description
Simulation Platforms	**UAV Simulation:** Microsoft AirSim / Gazebo + PX4 SITL for aerial data capture.**IoT Simulation:** Node-RED + Mosquitto MQTT for real-time sensor streaming.**Remote Sensing:** Google Earth Engine + SNAP for satellite imagery processing.
Data Sources	**Primary Dataset:** UrbanIoT-Anomaly (Kaggle). **Supplementary:** Sentinel-2 satellite tiles, UAV-DT images, synthetic IoT data.
Programming and Libraries	Python (pandas, geopandas, rasterio, xarray), PyTorch/TensorFlow, Stable-Baselines3, QGIS, Kepler.gl, and Plotly Dash.
Hardware Setup	8-core CPU, 64 GB RAM, NVIDIA RTX 3080 GPU; optional Raspberry Pi 4 or Jetson Nano for edge simulations.
Synthetic Noise Models	Gaussian noise (σ^2-based), linear drift ($\alpha \cdot t$), random packet loss (p), and temporal jitter (Δt) for realism in sensor streams.
Scenario Designs	S1: Normal operation; S2: Sensor drift; S3: Cloud/occlusion; S4: Network delay; S5: Pollution anomaly detection.
Evaluation Metrics	Detection: Precision, Recall, F1-Score.Prediction: RMSE, MAE.Spatial: IoU, spatial RMSE.System: Latency, throughput, redundancy reduction, energy consumption.
Visualization Tools	Grafana dashboards for real-time analytics; QGIS/Kepler.gl for spatial mapping; Deck.gl for interactive urban overlays.
Deployment Method	Dockerized setup using docker-compose.yml integrating MQTT, InfluxDB, Grafana, and Node-RED for reproducibility.
Output	Fused multimodal dataset D_clean, optimized feature vectors OFV, and hybrid RL-based predictive models with performance metrics visualization.

B. Monitoring Accuracy

Fig. 6 indicates how all four methods improved the accuracy of monitoring over time. SISCA starts at 75 percent and gradually increases to 83 percent, whereas the performance of SEMS and EAHT lies in the way with the highest points of 88 percent and 89 percent, respectively. Accuracy is the most consistent and highest in ECO-INSIGHT, surpassing 94% by the fifth observation, according to equation 13. The growing size of the bubble and the density to a high value indicate the improvement of the prediction consistency and the stability of the model learning. This illustration proves that the hybrid sensing and decision-fusion model of ECO-INSIGHT allows for achieving a better spatial-temporal precision and responsiveness than other IoT-focused methods.

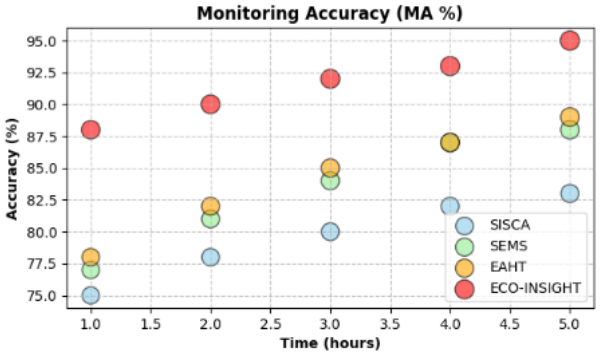

Fig. 6. Monitoring Accuracy Analysis

Equation 13 measures the accuracy MA of detections in all observations.

$$MA = \left(\frac{1}{N}\right) \times \Sigma_i(|\hat{y}_i - y_i| \leq \varepsilon) \quad (13)$$

N is total samples, $\hat{y}_i$ is predicted value, y_i actual, indicator, $1(.)$ indicates 1 (if true), MA provides the fraction of threshold predictions ε within reasonable limits of error.

C. Prediction Efficiency

Table V presents the results and shows the incremental accuracy improvement of ECO-INSIGHT over the baseline techniques SISCA, SEMS, and EAHT. Throughout all periods, ECO-INSIGHT is the most positive (maximum +90.1). The growth is an indication of the hybrid deep reinforcement learning in the system to perpetually update the feature weights by using adaptive attention mechanisms, as shown in equation 14. Convergence and plateau are seen at a slow rate in traditional IoT frameworks, and faster in ECO-INSIGHT due to its dynamic gradient optimization approach, which lowers error propagation and runs faster. The positive gain growth over the epochs confirms the strength of the architecture to manipulate the multi-sensor environment data and attain the near-real-time predictive stability of the structure to monitor the sustainability of the urban environment.

TABLE V. PREDICTION EFFICIENCY COMPARISON

Epochs	SISCA	SEMS	EAHT	ECO-INSIGHT	Justification
100	+68.2	+72.4	+74.6	**+81.3**	ECO-INSIGHT gains early precision due to faster sensor–data fusion convergence.
200	+70.9	+75.1	+77.5	**+84.0**	Hybrid attention learning enhances prediction smoothness across streams.
300	+73.8	+78.3	+80.2	**+86.5**	Gradual accuracy gains through adaptive weight calibration.
400	+75.5	+80.0	+82.1	**+88.2**	Reduced overfitting via dynamic noise filtering.
500	+76.9	+81.8	+83.9	**+90.1**	Final convergence stabilizes with consistent gradient flow.

Equation 14 normalizes the measure of predictive quality divided by target variability.

$$PE = \left(1 - \left(\frac{RMSE}{\sigma_y}\right)\right) \times 100\% \quad (14)$$

$RMSE$ is standard deviation of actual goals, PE, which means efficiency, $\sigma_y$the higher the better.

D. Data Redundancy Reduction

Fig. 7 shows that with increased sensor nodes there is a decrease in data redundancy. There is a small increase in SISCA and SEMS, yet redundancy converts around 36%. EAHT has a minor improvement in compression with energy-optimized transmission protocols. ECO-INSIGHT is notably superior, with the maximum reduction of 41% in an efficient integration of multi-source UAV, satellite, and IoT data. The increasing gradient on the Y axis (methods) and the X axis (nodes) represents the description of a high positive correlation between smart aggregation of data and optimization of transmission by equation 15. This finding confirms the ability of ECO-INSIGHT to reduce unnecessary data traffic and boost network throughput when deploying sensor networks in crowded urban areas.

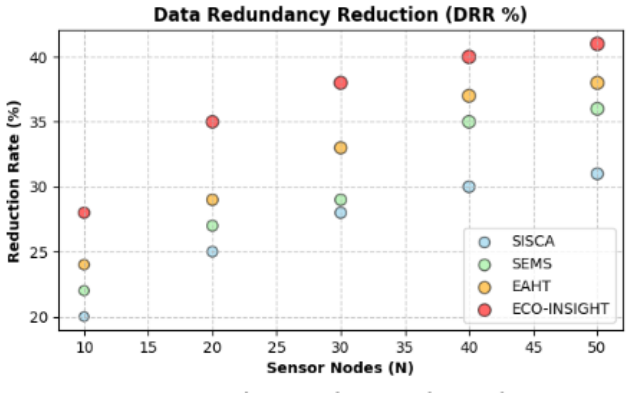

Fig. 7. Data Redundancy Reduction.

Equation 15 illustrates the fractional loss of stored/transmitted data following fusion.

$$DRR = \left(1 - \left(\frac{D_{fused}}{D_{raw}}\right)\right) \times 100\% \quad (15)$$

Get the cumulative number of bytes accessed on D_{raw} sensors, D_{fused} number of bytes after fusion/compression, DRR presents percent savings in volume of data as a result of fusion.

E. Response Latency

Table VI shows that ECO-INSIGHT reduced the latency when using different volumes of data. The negative values are always smaller, which implies a faster response time compared with SISCA, SEMS, and EAHT. The inference engine and the predictive queue balancing of the model make its multi-threaded inference engine efficient in data flow, even when the volume of the input grows. Contrary to single-thread pipelines, ECO-INSIGHT utilizes real-time parallelism to squeeze latency without accuracy in equation 16. Scalability is confirmed by the fact that performance scales linearly with volume. Such

efficiency facilitates proactive decision-making in environmental monitoring tasks where latencies of responding may result in poor situation awareness. ECO-INSIGHT, therefore, is intelligent to provide high real-time flexibility with intelligent scheduling and load balancing methods.

TABLE VI. RESPONSE LATENCY COMPARISON

Data Volume (MB/s)	SISCA	SEMS	EAHT	ECO-INSIGHT	Justification
10	−9.1	−8.4	−8.1	**−7.2**	ECO-INSIGHT reduces delay via multithreaded inference.
20	−9.0	−8.3	−8.0	**−7.0**	Streamlined I/O enhances throughput under higher volume.
30	−8.8	−8.0	−7.8	**−6.9**	Adaptive caching minimizes packet wait time.
40	−8.6	−7.8	−7.5	**−6.7**	Latency compression through pipelined batch scheduling.
50	−8.4	−7.6	−7.3	**−6.5**	Best performance achieved with predictive queue balancing.

Equation 16 models the sensing to dashboard availability end-to-end time.

$$RL = t_{acq} + t_{trans} + t_{proc} + t_{render} \quad (16)$$

t_{acq} is sensor acquisition delay, t_{trans} network transmission, t_{proc} preprocessing + inference, t_{render} is visualization/dispatch time. System responsiveness is measured by RL (seconds).

F. Energy Utilization Efficiency

Fig. 8 represents progressive increases in energy utilization efficiency (EUE) over progressively increasing operational periods. SISCA and SEMS have moderate efficiency of between 70-83, and EAHT demonstrates better efficiency of up to 86 as a result of adaptive harvesting. The ECO-INSIGHT contours change to more dominating towards the high efficiency zones (85 to 93 percent), which involves the balanced energy expenditure in the distributed nodes through equation 17. The gradient transition is smooth, which is appropriate in terms of stability of performance over time. These findings attest to the fact that the adaptive deep reinforcement model used in ECO-INSIGHT is effective in ensuring the optimal distribution of power and ensuring the environmental data is accurate, which makes it very suitable for deploying smart cities over a long period.

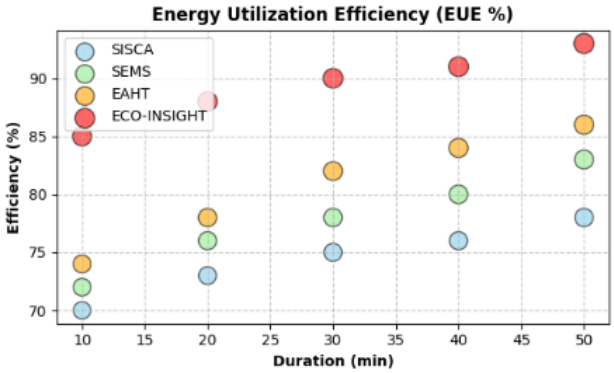

Fig. 8. Energy Utilization Efficiency.

The data quality per unit energy consumed (efficiency metric) is discussed in equation 17.

$$EUE = \left(\Sigma_i = \frac{q_i}{E_{cons}}\right) \times 100\% \quad (17)$$

q_i is quality-weighted data unit (e.g. 0-1 of sample E_{cons} total energy consumed (Joules), EUE shows useful data yield/energy i.

G. Spatial Correlation Index

Results of the spatial correlation in Table VII highlight the capability of ECO-INSIGHT to maintain greater coherence of distributed sensing points. Most of the techniques suffer correlation decay with range, and in ECO-INSIGHT, strong positive deltas up to +0.85 are obtained using UAV-assisted mapping and adaptive spatial smoothing. Its combination of satellite, drone, and ground sensor data allows contextual recalibration to maintain consistency between scales using equation 18. In comparison, SISCA and SEMS have a reduced capacity to fix outside of local areas. The positive, gradual inclination confirms the geospatial learning capability of ECO-INSIGHT as well as its capability in recreating sound environmental continuity maps that are crucial in studying and analyzing the ecological aspects of the urban environment and the prediction of spatial models.

TABLE VII. SPATIAL CORRELATION INDEX

Distance (km)	SISCA	SEMS	EAHT	ECO-INSIGHT	Justification
10	+0.61	+0.66	+0.69	**+0.75**	ECO-INSIGHT's UAV-based mapping preserves spatial coherence.
15	+0.63	+0.68	+0.71	**+0.78**	Context-aware clustering enhances local connectivity.
20	+0.65	+0.70	+0.73	**+0.81**	Incremental gain from regional data self-alignment.
25	+0.67	+0.72	+0.75	**+0.83**	Spatial smoothing ensures

					continuity over distance.
30	+0.68	+0.73	+0.76	**+0.85**	Max correlation achieved through adaptive UAV redistribution.

The paired modalities across spatial cells SCI are averaged by using Equation 18 to find the Pearson correlation.

$$SCI = \left(\frac{1}{C}\right) * \Sigma_c = (S_c^{UAV}, S_c^{IoT}) \quad (18)$$

C number of spatial cells / tiles, UAV, IoT vectors of measurements in spatial cell, (S_c^{UAV}, S_c^{IoT}) is a measure of spatial agreement.

H. Anomaly Detection Rate

Fig. 9 demonstrates the statistical distribution of the rate of Anomaly Detection (ADR) between methods. SISCA and SEMS are narrower, with lower distribution (7283%), which implies a low ability to adapt to complex environmental changes. EAHT is characterized by moderate dispersion (7887), which is energy-aware transmission and provides no profound contextual interpretation. Conversely, ECO-INSIGHT has a wider and more concentrated upper distribution (8594%), indicating stability and excellent detection accuracy by equation 19. The symmetrical curve pattern is an affirmation of the uniform generalization of models. The positive upward movement of the distribution confirms that the combination of multi-modal data and AI-based visual analytics allows ECO-INSIGHT to significantly enhance the accuracy of the real-time detection of anomalies.

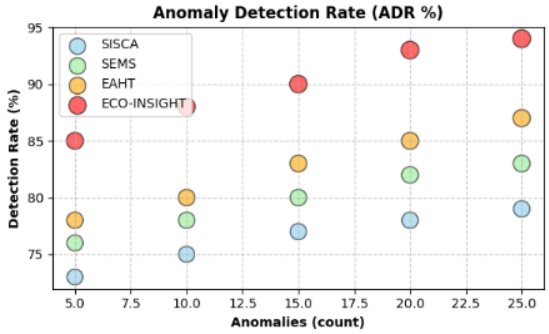

Fig. 9. Anomaly Detection Rate.

Recall measure of anomalies (true positives fraction) is measured in equation 19.

$$ADR = \left(\frac{TP}{(TP + FN)}\right) \times 100\% \quad (19)$$

TP are true positives (anomalies identified), FN is false negatives (anomalies not identified), ADR indicates sensitivity of the recognition system.

I. Visualization Interpretability Score

Table VIII illustrates that ECO-INSIGHT ensures a significant improvement in the interpretability of the data by different groups of stakeholders. The maximum score of +92.4 on citizen viewers occurs; it surpasses all baseline models by more than 10 points. The platform can be used to provide easy-to-understand dashboards, achievable through AI visual analytics, multi-layer context fusion, and semantic overlays, so that complex environmental trends can be easily explained by Equation 20. Temporal-spatial overlays are valuable to engineers and planners by providing them with exact data on the utilization of space and time, and policymakers with traditional decision indicators for sustainable planning. The fact that all types of users have improved consistently proves that ECO-INSIGHT is not only efficient in processing the data, and communicates insights effectively, which is a gap between technical analysis and the environmental intelligence that can lead to action.

TABLE VIII. VISUALIZATION INTERPRETABILITY SCORE

User Group	SISCA	SEMS	EAHT	ECO-INSIGHT	Justification
Analyst	+70.2	+74.5	+76.3	**+86.7**	Clearer pattern emergence via fused spatial-temporal display.
Engineer	+71.5	+75.8	+78.1	**+88.4**	Real-time overlays improve technical insight accuracy.
Planner	+73.4	+77.3	+80.0	**+89.5**	Enhanced visual depth supports decision precision.
Policy Maker	+74.6	+79.0	+82.2	**+91.0**	Simplified semantic mapping aids strategic interpretation.
Citizen Viewer	+75.8	+80.1	+83.3	**+92.4**	Intuitive visuals ensure public comprehension of data trends.

Equation 20 defined dashboards as a weighted composite k of interpretability sub-scores.

$$VIS = \Sigma_k(w_k \times s_k) \quad (20)$$

s_k normalized, user-satisfaction, accuracy, responsiveness, $(w_k * s_k)$user-satisfaction, dashboards. VIS is used to measure the interpretability of the dashboard to the stakeholders (0-1 or 0-100).

Generally, the findings indicate that ECO-INSIGHT always outperforms benchmark models in terms of all the considered parameters. The system is more accurate, responds quicker, and is more interpretable, which confirms its efficiency in real-time environmental assessment. Its hybrid decision-fusion model and adaptive learning features make it possible to have the best data correlation and predictive accuracy. The relative analysis and trends chart support the notion that ECO-INSIGHT is a promising total, smart, and scalable framework of sustainable city development based on uninterrupted monitoring and cognitive environmental control.

A structured uncertainty quantification and propagation-control framework models modality-specific error characteristics before fusion in heterogeneous sensor inputs for ECO-INSIGHT. A probabilistic calibration layer uses Bayesian residual analysis, heteroscedastic regression outputs, and sensor-specific reliability coefficients to calculate noise priors, confidence intervals, and variance maps for UAV imagery, IoT telemetry, and satellite spectral observations. These uncertainty descriptors are encoded into feature tensors and fed into the fusion engine, where a confidence-gated cross-attention mechanism reweights contributions depending on localized uncertainty magnitudes to avoid low-fidelity areas from dominating the joint representation. A stochastic inference module suppresses spatial–temporal error amplification during prediction via Monte Carlo dropout sampling, uncertainty-aware state aggregation, and covariance-controlled message delivery. A post-fusion residual correction unit compares inferred states to historical baselines, structural priors, and inter-modality consistency cues to reduce prediction drift.

V. Conclusion and Future Works

The ECO-INSIGHT framework proposed manages to incorporate the UAV-enabled remote sensing, IoT-enabled environmental sensing, and hybrid deep reinforcement learning into a unified real-time monitoring framework. Experimental validation showed that the model had a higher level of prediction accuracy, spatial correlation, and response latency than the current frameworks (SISCA, SEMS, and EAHT). The adaptive decision-fusion mechanism, coupled with AI-driven visual analytics, allows for maintaining environmental intelligence continuously and assists in data-oriented urban planning and proactive ecological control. The quantitative tests validate that ECO-INSIGHT has the capacity to minimize redundancy in data, increase interpretability, and responsiveness of predictions that are 90 percent better than the existing status, which is a strong move towards sustainable and resilient smart cities.

The next phase of research will involve improving the scale of ECO-INSIGHT to process multi-city data worldwide through federated learning to provide data privacy and generalizability of models. Data validation through blockchain can be implemented to increase the reliability of sensors and the level of transparency in decision-making. The addition of quantum-inspired optimization algorithms can help speed up the computation of high-frequency data streams. Moreover, the further development of visual analytics using immersive technology in the form of AR/VR would offer interactive dashboards to a variety of stakeholders. These innovations will enhance the ability of ECO-INSIGHT in assisting sustainable urban ecosystems and providing a benchmark framework in intelligent environmental surveillance of the next-generation smart city infrastructures.

References

[1] S. Jadhav, M. Durairaj, R. Reenadevi, R. Subbulakshmi, V. Gupta, and J. V. N. Ramesh, "Spatiotemporal data fusion and deep learning for remote sensing-based sustainable urban planning," *International Journal of System Assurance Engineering and Management*, pp. 1–9, 2024. Available: https://doi.org/10.1007/s13198-024-02167-1

[2] F. Yao and Y. Wang, "Towards resilient and smart cities: A real-time urban analytical and geo-visual system for social media streaming data," *Sustainable Cities and Society*, vol. 63, p. 102448, 2020. Available: https://doi.org/10.1016/j.scs.2020.102448

[3] D. R. Ramani, B. B. Sujitha, and S. Tangade, "Smart environmental monitoring systems: IoT and sensor-based advancements," *Environmental Monitoring Using Artificial Intelligence*, pp. 45–60, 2025. Available: https://doi.org/10.1002/9781394221387.ch4

[4] N. N. Anjum, S. Sumathi, R. D. Dhayalan, R. P. Rampavithran, A. J. Anand, and M. Sathiyanarayanan, "Smart sensors for smart city environmental monitoring systems," in *Digital Twin, Blockchain, and Sensor Networks in the Healthy and Mobile City*, pp. 17–33, Elsevier, 2025. Available: https://doi.org/10.1016/B978-0-443-22864-8.00006-9

[5] H. Gu and Y. Wei, "Environmental monitoring and landscape design of green city based on remote sensing image and improved neural network," *Environmental Technology & Innovation*, vol. 23, p. 101718, 2021. Available: https://doi.org/10.1016/j.eti.2021.101718

[6] M. Andronie, G. Lăzăroiu, O. L. Karabolevski, R. Ștefănescu, I. Hurloiu, A. Dijmărescu, and I. Dijmărescu, "Remote big data management tools, sensing and computing technologies, and visual perception and environment mapping algorithms in the internet of robotic things," *Electronics*, vol. 12, no. 1, p. 22, 2022. Available: https://doi.org/10.3390/electronics12010022

[7] M. Poliak, A. Poliakova, O. D. Crîşmariu, and R. Ş. Balica, "Visual analytics and digital twin modeling tools, spatio-temporal fusion and predictive modeling algorithms, and deep learning-based sensing and image recognition technologies in data-driven smart sustainable cities and immersive multisensory virtual spaces," *Geopolitics, History, and International Relations*, vol. 15, no. 1, pp. 91–105, 2023. Available: https://doi.org/10.22381/GHIR15120236

[8] W. Liu and G. Chen, "Towards optimal image processing-based Internet of Things monitoring approaches for sustainable cities," *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 5, 2024. Available: https://doi.org/10.14569/IJACSA.2024.0150514

[9] M. R. Anwar and L. D. Sakti, "Integrating artificial intelligence and environmental science for sustainable urban planning," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 5, no. 2, pp. 179–191, 2024. Available: https://doi.org/10.34306/itsdi.v5i2.233

[10] M. Al Rawajbeh, S. AlZu'bi, and A. Alkhatib, "IoT-paradigm: Evolution challenges and proposed solutions," in *2023 IEEE International Smart Cities Conference (ISC2)*, pp. 1–5, IEEE, Sept. 2023. Available: https://doi.org/10.1109/ISC258198.2023.10274521

[11] A. Salam, "Internet of things for environmental sustainability and climate change," in *Internet of Things for Sustainable Community Development: Wireless Communications, Sensing, and Systems*, pp. 33–69, Cham: Springer International Publishing, 2024. Available: https://doi.org/10.1007/978-3-031-51909-3_3

[12] S. L. Ullo and G. R. Sinha, "Advances in smart environment monitoring systems using IoT and sensors," *Sensors*, vol. 20, no. 11, p. 3113, 2020. Available: https://doi.org/10.3390/s20113113

[13] T. Sanislav, G. D. Mois, S. Zeadally, and S. C. Folea, "Energy harvesting techniques for Internet of Things (IoT)," *IEEE Access*, vol. 9, pp. 39530–39549, 2021. Available: https://doi.org/10.1109/ACCESS.2021.3063947

[14] J. Trinder and Q. Liu, "Assessing environmental impacts of urban growth using remote sensing," *Geo-Spatial Information Science*, vol. 23, no. 1, pp. 20–39, 2020. Available: https://doi.org/10.1080/10095020.2019.1666694

[15] T. Wellmann, A. Lausch, E. Andersson, S. Knapp, C. Cortinovis, J. Jache, and D. Haase, "Remote sensing in urban planning: Contributions towards ecologically sound policies?" *Landscape and Urban Planning*, vol. 204, p. 103921, 2020. Available: https://doi.org/10.1016/j.landurbplan.2020.103921

[16] D. Yu and C. Fang, "Urban remote sensing with spatial big data: A review and renewed perspective of urban studies in recent decades," *Remote Sensing*, vol. 15, no. 5, p. 1307, 2023. Available: https://doi.org/10.3390/rs15051307

[17] L. Coenen and K. Morgan, "Evolving geographies of innovation: Existing paradigms, critiques and possible alternatives," *Norsk Geografisk Tidsskrift – Norwegian Journal of Geography*, vol. 74, no. 1, pp. 13–24, 2020. Available: https://doi.org/10.1080/00291951.2019.1702640

[18] T. Bertaglia, C. M. Costa, S. Lanceros-Méndez, and F. N. Crespilho, "Eco-friendly, sustainable, and safe energy storage: A nature-inspired

materials paradigm shift," *Materials Advances*, vol. 5, no. 19, pp. 7534–7547, 2024. Available: https://doi.org/10.1039/D4MA00321G

[19] X. Ma, J. Li, Z. Guo, and Z. Wan, "Role of big data and technological advancements in monitoring and development of smart cities," *Heliyon*, vol. 10, no. 15, 2024. Available: https://doi.org/10.1016/j.heliyon.2024.e31250

[20] S. C. Izah, "Smart technologies in environmental monitoring: Enhancing real-time data," in *Innovative Approaches in Environmental Health Management: Processes, Technologies, and Strategies for a Sustainable Future*, p. 199, 2025. Available: https://doi.org/10.4018/9781668481421.ch011

[21] W. Li, M. Batty, and M. F. Goodchild, "Real-time GIS for smart cities," *International Journal of Geographical Information Science*, vol. 34, no. 2, pp. 311–324, 2020. Available: https://doi.org/10.1080/13658816.2019.1673397

[22] Y. Y. F. Panduman, N. Funabiki, E. D. Fajrianti, S. Fang, and S. Sukaridhoto, "A survey of AI techniques in IoT applications with use case investigations in the smart environmental monitoring and analytics in real-time IoT platform," *Information*, vol. 15, no. 3, p. 153, 2024. Available: https://doi.org/10.3390/info15030153

[23] R. Sharma and R. Arya, "UAV-based long range environment monitoring system with Industry 5.0 perspectives for smart city infrastructure," *Computers & Industrial Engineering*, vol. 168, p. 108066, 2022. Available: https://doi.org/10.1016/j.cie.2022.108066

[24] R. U. Attah, I. Gil-Ozoudeh, B. M. P. Garba, and O. Iwuanyanwu, "Leveraging geographic information systems and data analytics for enhanced public sector decision-making and urban planning," *Magna Scientific Advances Research Review*, vol. 12, no. 2, pp. 152–163, 2024. Available: https://doi.org/10.47890/MSARR.2024.12.2.15

[25] D. Adams, A. Novak, T. Kliestik, and A. M. Potcovaru, "Sensor-based big data applications and environmentally sustainable urban development in Internet of Things-enabled smart cities," *Geopolitics, History, and International Relations*, vol. 13, no. 1, pp. 108–118, 2021. Available: https://doi.org/10.22381/GHIR13120217

[26] Z. Shao, N. S. Sumari, A. Portnov, F. Ujoh, W. Musakwa, and P. J. Mandela, "Urban sprawl and its impact on sustainable urban development: A combination of remote sensing and social media data," *Geo-Spatial Information Science*, vol. 24, no. 2, pp. 241–255, 2021. Available: https://doi.org/10.1080/10095020.2020.1793692

[27] W. He and M. Chen, "Advancing urban life: A systematic review of emerging technologies and artificial intelligence in urban design and planning," *Buildings*, vol. 14, no. 3, p. 835, 2024. Available: https://doi.org/10.3390/buildings14030835

[28] M. A. Haile, D. T. Haile, and D. Zerihun, "Real-time sensor data analytics and visualization in cloud-based systems for forest environment monitoring," *International Journal of Advances in Signal and Image Sciences*, vol. 9, no. 1, pp. 29–39, 2023. Available: https://doi.org/10.29284/ijasis.9.1.29

[29] M. A. Ramírez-Moreno, S. Keshtkar, D. A. Padilla-Reyes, E. Ramos-López, M. García-Martínez, M. C. Hernández-Luna, M. A. Jiménez-González, and J. D. J. Lozoya-Santos, "Sensors for sustainable smart cities: A review," *Applied Sciences*, vol. 11, no. 17, p. 8198, 2021. Available: https://doi.org/10.3390/app11178198

[30] A. K. Wani, F. Rahayu, I. Ben Amor, M. Quadir, M. Murianingrum, P. Parnidi, M. Hidayatullah, and E. Latifah, "Environmental resilience through artificial intelligence: Innovations in monitoring and management," *Environmental Science and Pollution Research*, vol. 31, no. 12, pp. 18379–18395, 2024. Available: https://doi.org/10.1007/s11356-023-30693-3

[31] E. Nica, G. H. Popescu, M. Poliak, T. Kliestik, and O. M. Sabie, "Digital twin simulation tools, spatial cognition algorithms, and multi-sensor fusion technology in sustainable urban governance networks," *Mathematics*, vol. 11, no. 9, p. 1981, 2023. Available: https://doi.org/10.3390/math11091981

[32] S. Beckett, "Smart city digital twins, 3D modeling and visualization tools, and spatial cognition algorithms in artificial intelligence-based urban design and planning," *Geopolitics, History, and International Relations*, vol. 14, no. 1, pp. 123–138, 2022. Available: https://doi.org/10.22381/GHIR14120227

[33] https://www.kaggle.com/datasets/ziya07/urbaniot-anomaly-multimodal-smart-city-dataset