Harnessing EO and Natural Experiments for Urban Development: The UDENE Approach

Vasil Yordanov[1], Daniele Oxoli[1], Nikola Obrenović[2], Branislav Pejak[2], Onur Lenk[3], *, Nebiye Musaoğlu[4],
Bekir Kapukaya[4], Yücel Erbay[5], Seda Kara[5], Tacettin Yücel İpek[5], Mohamed Rajhi[6], Ahmed El Fadhel[6], Syrine Souissi[6],
Murat Özbayoğlu[7], Hüseyin Pekkan[7], Oğuz Deniz[7], Ali Türker[8], Marina Carević Tomić[9], Maria Antonia Brovelli[1]

[1] Department of Civil and Environmental Engineering (DICA), Politecnico di Milano, 20133 Milano, Italy
(vasil.yordanov, daniele.oxoli, maria.brovelli)@polimi.it
[2] BioSense Institute, University of Novi Sad, 21000 Novi Sad, Serbia - (nikola.obrenovic, branislav.pejak)@biosense.rs
[3] Istanbul University, Istanbul, Türkiye - onur.lenk@istanbul.edu.tr
[4] Istanbul Technical University, Istanbul, Türkiye - (musaoglune,bekirkapulaya)@itu.edu.tr
[5] NiK Insaat Ticaret Ltd. Şti., Istanbul, Turkiye - (yucel, seda,yucel)@nik.com.tr
[6] Tunisian Space Association, Sfax, Tunisia - (ahmedelfadhel.official, m.rajhi.inat)@gmail.com
[7] Department of Artificial Intelligence Engineering, TOBB University of Economics and Technology, Ankara, Türkiye
mozbayoglu@etu.edu.tr
[8] WeGlobal srl, 20152 Milano, Italy - ali.turker@weglobal.org
[9] Department of Architecture and Urbanism, Faculty of Technical Sciences, University of Novi Sad - marinac@uns.ac.rs

Keywords: Earth Observation, Urban Development, Natural Experiments, Digital Twin, Data Cube, Sustainable Cities.

Abstract

Urban Development Explorations using Natural Experiments (UDENE) is a forward-looking initiative under the Horizon Europe program that merges Earth Observation (EO) technologies with urban planning to tackle pressing urban challenges. By utilizing Copernicus satellite imagery and organizing local in-situ data into interoperable data cubes, UDENE provides a comprehensive framework for data-driven decision-making. Further, it is applied the concept of "natural experiments" - real-life changes analyzed with the rigor of controlled studies - to uncover causal relationships in urban development. A primary goal is to incorporate structured urban data into the broader Copernicus data cube federation, enabling consistent analysis of urban impacts across different times and locations. To support this, UDENE develops advanced sensitivity analysis methods for validating and applying multivariate causal models, enhancing predictions on factors such as air pollution, urban heat, mobility, and disaster resilience. To close the gap between high-level EO technologies and real-world planning needs, we are introducing the three core tools: the UDENE's Data Cube, which populates in-situ data EO based analysis-ready data and datasets; the Exploration Tool, which empowers planners and policymakers to simulate, assess, and visualize urban interventions; and a matchmaking tool connecting users with EO-based services. Together, these tools foster informed urban strategies grounded in EO data and causal inference.

1. Introduction

Urbanization presents complex challenges that require innovative, data-driven approaches for sustainable development. The Urban Development Explorations using Natural Experiments (UDENE) project, funded under Horizon Europe, pioneers a framework that integrates Earth Observation (EO) technologies with urban analytics to identify causal links between interventions and outcomes in urban environments. Central to this approach is the concept of natural experiments, where uncontrolled but observable urban changes are treated with scientific rigor, enabling comparative evaluations across space and time.

By structuring satellite and in-situ datasets into interoperable data cubes, UDENE allows stakeholders to model and simulate urban transformations, assess policy impacts, and improve resilience planning. This is achieved through a suite of tools, including the UDENE Data Cube and Exploration Tool, all designed to democratize access to EO-derived insights for planners, scientists, and policy-makers.

The project builds upon a robust body of literature that leverages EO and spatial data science for urban monitoring and planning. Recent studies underscore the growing relevance of EO in analyzing urban thermal environments and informing climate adaptation strategies (Pugliese Viloria et al., 2024). In particular, the use of Local Climate Zones (LCZs) as standardized spatial units has advanced the remote sensing of urban heat islands (UHIs) from basic surface temperature mapping to more integrated assessments (Oxoli et al., 2018; Vavassori et al., 2024). Similarly, advancements in EO-based air pollution modeling have highlighted the effectiveness of combining satellite-derived parameters with in-situ and auxiliary datasets to reconstruct the spatiotemporal dynamics of pollutants (Gianquintieri et al., 2024), particularly in relation to traffic-related interventions, and can be further strengthened by coupling these insights with agent-based traffic simulations and emission estimation models to assess how mobility changes influence urban air quality (Gurram et al., 2019). Furthermore, Earth Observation (EO) data are increasingly used to support scenario-based assessments of earthquake risks, allowing for the estimation of potential building damage and economic losses to guide urban land use decisions and enhance resilience planning.

This paper presents three flagship UDENE use cases: (1) evaluating the environmental effects of transport changes in Novi Sad, Serbia; (2) quantifying urban heat mitigation via green infrastructure in Greater Tunis, Tunisia; and (3) modeling seismic vulnerability in high-rise districts of Istanbul,

---

* Corresponding author

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-4-W18-2025-363-2026 | © Author(s) 2026. CC BY 4.0 License.

Türkiye. Together, these cases showcase the operationalization of UDENE's core tools and the replicability of its framework for urban policy experimentation.

2. Use Case 1: Environmental Impacts of Traffic Interventions in Novi Sad, Serbia

In Novi Sad, UDENE's efforts focus on assessing the environmental impacts resulting from transport infrastructure changes, specifically the construction of new bypass bridges and the pedestrianization of a historical zone. These initiatives aim to mitigate the impact of air pollution and enhance urban mobility, utilizing sophisticated machine learning and agent-based models, as well as Earth Observation (EO) data to provide comprehensive assessments.

2.1 Data Collection and Preprocessing

The study analysis integrates Sentinel-5P satellite data monitoring NO_2 concentrations with comprehensive transport potential datasets, localized traffic records, and air quality measurements from ground-based sensors. Extensive data preparation was performed, including detailed exploratory data analysis, feature engineering, and the aggregation of temporal data from hourly measurements into daily, weekly, and monthly aggregations. Distinct datasets were created by merging street segments with nearby Serbian Environmental Protection Agency (SEPA) measurement devices, with scenarios evaluated both with and without traffic data inclusion.

2.2 Machine Learning and Regression Analysis

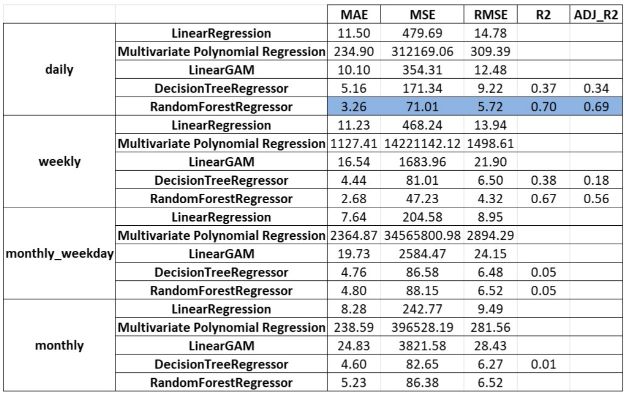
The BioSense Institute developed several multivariate regression models, based on different machine learning algorithms. The models were trained on various datasets utilizing 3-fold cross-validation grouped monthly to prevent data leakage. The methods explored included:

- Linear multi-output regression,
- Multivariate polynomial regression with second-degree polynomial and interaction features,
- Linear Generalized Additive Models (LinearGAM),
- Decision Trees, and
- Random Forest algorithms.

Model performance was rigorously evaluated using metrics such as mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), R^2, and adjusted R^2. The optimal model configuration, in our case based on the Random Forest algorithm, was selected based on adjusted R^2 performance (Table 1) and further trained on the most effective dataset.

2.3 Sensitivity Analysis and Explainability

Sensitivity analyses of the best-performing regression models employed advanced local and global eXplainable Artificial Intelligence (XAI) techniques (Kopanja et al., 2025). Methods such as permutation feature importance, SHapley Additive exPlanations (SHAP), and mean decrease in impurity for tree-based models were applied. Specifically, TreeSHAP was utilized for tree-based methods, while LinearSHAP was deployed for linear regression models to identify influential factors clearly and transparently.



		MAE	MSE	RMSE	R2	ADJ_R2
daily	LinearRegression	11.50	479.69	14.78		
	Multivariate Polynomial Regression	234.90	312169.06	309.39		
	LinearGAM	10.10	354.31	12.48		
	DecisionTreeRegressor	5.16	171.34	9.22	0.37	0.34
	RandomForestRegressor	3.26	71.01	5.72	0.70	0.69
weekly	LinearRegression	11.23	468.24	13.94		
	Multivariate Polynomial Regression	1127.41	14221142.12	1498.61		
	LinearGAM	16.54	1683.96	21.90		
	DecisionTreeRegressor	4.44	81.01	6.50	0.38	0.18
	RandomForestRegressor	2.68	47.23	4.32	0.67	0.56
monthly_weekday	LinearRegression	7.64	204.58	8.95		
	Multivariate Polynomial Regression	2364.87	34565800.98	2894.29		
	LinearGAM	19.73	2584.47	24.15		
	DecisionTreeRegressor	4.76	86.58	6.48	0.05	
	RandomForestRegressor	4.80	88.15	6.52	0.05	
monthly	LinearRegression	8.28	242.77	9.49		
	Multivariate Polynomial Regression	238.59	396528.19	281.56		
	LinearGAM	24.83	3821.58	28.43		
	DecisionTreeRegressor	4.60	82.65	6.27	0.01	
	RandomForestRegressor	5.23	86.38	6.52		

Table 1. Evaluation of the tested regression models and configurations

2.4 Agent-Based Traffic Simulation and Emission Estimation

Utilizing the MATSim simulation framework, the team developed an emissions-calculation module, improving the capacity to estimate urban mobility-related environmental impacts accurately. To overcome the lack of data, characteristic for agent-based modeling, our team developed a state-of-the-art approach for the integrated generation of population and household data (Maglevannaia et al., 2025), inspired by the work of Kukic et al. (2024). Synthetic population data, enriched with detailed household and individual-level characteristics, travel behaviors, and daily activity patterns (Figure 1), underpinned this advanced simulation environment.

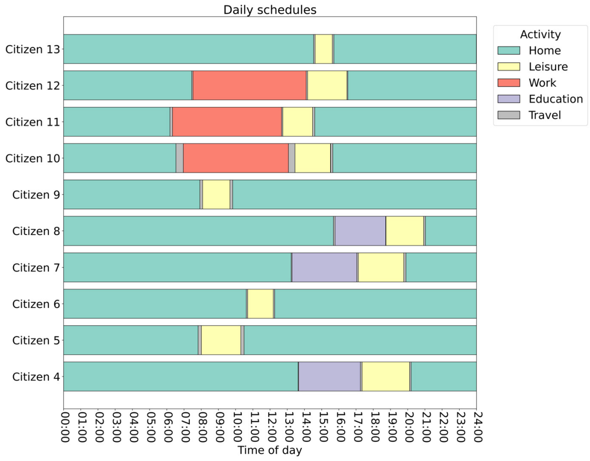


Figure 1. Activity planning

Additionally, the simulation environment, developed using the GAMA platform, integrated customized agent behaviors (including real-time road monitoring), road network preprocessing (addressing road tag conflicts, traffic light logic, intersection detection, and edge segmentation), and comprehensive scenario analyses of proposed urban interventions (new bridges and pedestrian zones). Integration with a Python client application via WebSocket allowed automated, headless simulation runs, facilitating extensive performance benchmarking. A comparative analysis pipeline was also established, enabling the visualization and interpretation of differences in traffic flows between scenarios (Figure 2).

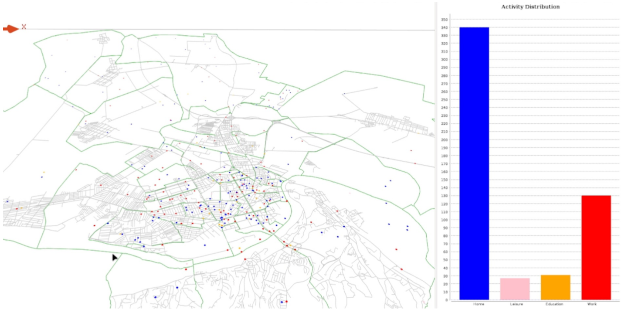


Figure 2. ABM simulation

3. Use Case 2: Urban Heat Island Mitigation in Greater Tunis, Tunisia

This case addresses UHI effects in Greater Tunis and the cooling impact of a linked park system across LCZs. EO data such as Sentinel-2-based vegetation indices and land surface temperature from Landsat - combine with in-situ meteorological data to evaluate thermal mitigation. Urban built-up areas are mapped via Normalized Difference Building Index (NDBI), while Random Forest models explore how green infrastructure influences local heat dynamics.

3.1 Study area

The Greater Tunis metropolitan region, located in northern Tunisia along the Mediterranean coastline, serves as the study area for this research. It encompasses the capital city of Tunis and its surrounding municipalities. The study area covers a surface of about 126 km^2. With a population of approximately 2.6 million, Greater Tunis is a highly urbanized zone, marked by a dense mix of residential, commercial, and industrial infrastructures. Geographically, the region exhibits considerable environmental diversity, featuring plains, coastal lagoons and lakes that contribute to local ecology and microclimates. These natural and anthropogenic characteristics make Greater Tunis an ideal case for investigating urban environmental phenomena such as the Surface Urban Heat Island (SUHI) effect.

3.2 Data Collection and Preprocessing

In this study, we employed a dual-source approach for remote sensing data acquisition, combining manual downloading of Sentinel data via the Copernicus Data Space Ecosystem with a semi-automated retrieval from the Digital Earth Africa Data Cube. These methods allowed us to balance flexibility in spatial-temporal selection with scalable access to preprocessed data.

Manual Data Collection involves downloading Sentinel-2 (for surface reflectance and vegetation indices) and Sentinel-3 (for Land Surface Temperature, LST) imagery from the Copernicus Browser. The user manually selected spatial polygons and timeframes of interest and downloaded Level-1C or Level-2A products. This raw data required significant preprocessing, including atmospheric correction, cloud masking, geometric alignment, and reprojection to match the study area coordinates.

Semi-Automated Data Collection leveraged the Digital Earth Africa Open Data Cube, a cloud-based JupyterLab environment offering access to Analysis Ready Data (ARD). The data were retrieved using spatial-temporal queries and already included corrections for atmosphere and geometry. This approach reduced preprocessing workload and ensured consistency and comparability across space and time. After evaluating data accessibility, consistency, and processing overhead, we selected the semi-automated Digital Earth Africa Data Cube as the primary source, ensuring both data quality and processing efficiency for subsequent modeling.

We used both raster (e.g., GeoTIFF for LST, indices, elevation) and vector data (e.g., shapefiles for study area boundaries), and ensured alignment through coordinate system conversion to EPSG:4326 and spatial resampling to 30 m resolution. For harmonization, GHS and LCZ rasters were interpolated using cKDTree-based inverse distance weighting. Data transformation included Min-Max scaling for bounded indices, RobustScaler for outlier-prone features, and data type optimization to enhance computational performance. One-hot encoding was applied to categorical variables like LCZ class and season. In addition, the Surface Urban Heat Island (SUHI) index was computed (Figure 3) using LST differentials between urban and reference (dense forest) zones at the pixel level and seasonal level.

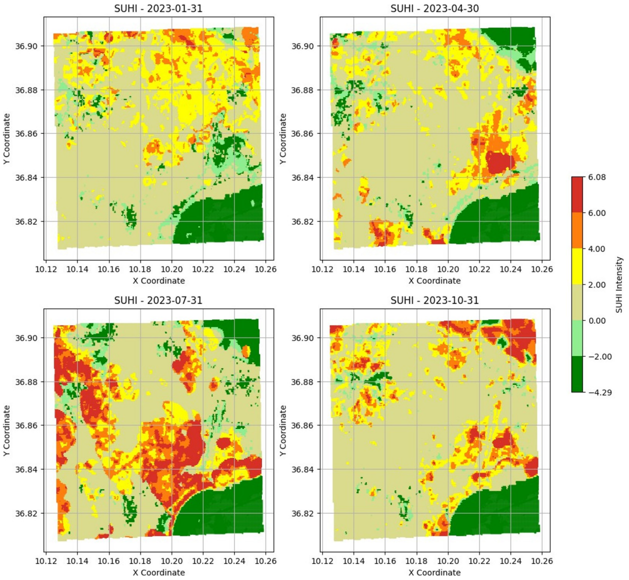


Figure 3. SUHI Intensity Maps

3.2.1 Modeling and Evaluation: The modeling phase translates processed data into actionable insights through predictive algorithms. This study targets the prediction of SUHI intensity using ensemble learning techniques, Random Forest (RF), LightGBM (LGBM), and XGBoost, chosen for their proven performance in environmental and geospatial prediction tasks.

3.2.2 Ensemble Learning Overview. Ensemble learning enhances model robustness by combining multiple ”weak learners”. Two major strategies were applied: Bagging, using RF, which trains multiple independent trees and averages predictions to reduce variance; and Boosting, where models are trained sequentially, each correcting the predecessor’s errors. Boosting algorithms are powerful for capturing complex patterns but require careful tuning to prevent overfitting.

3.2.3 Sensitivity Analysis: Feature importance analysis revealed that normalized indeces for moisture, water, vegetation (NDMI, NDWI, NDVI, EVI) and LCZ classes were the most influential predictors after LST removal. This emphasizes vegetation and moisture as key cooling agents. Simulated scenarios further validated model sensitivity: increasing vegetation indices in LCZ 11/14 slightly raised SUHI due to spatial interactions, while enhancing indices in Urban Green Spaces (UGS) led to 2°C SUHI reduction (Figure 4), confirming the mitigation effect of localized greenery.

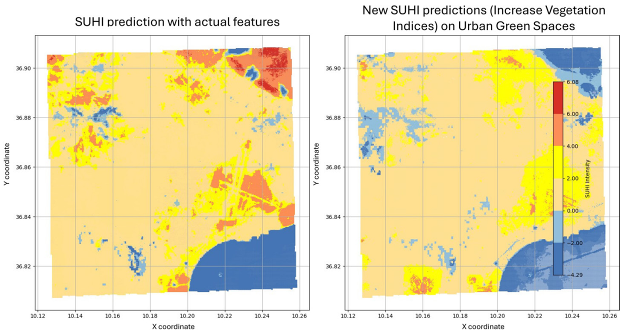


Figure 4. SUHI Intensity Distribution Maps.

4. Use Case 3: Seismic Risk and Urban Resilience in Istanbul, Türkiye

4.1 From Earthquake Hazard to Urban Risk in High-Rise Environments

Tectonically active metropolitan areas, such as Istanbul, face increasing vulnerability due to a combination of natural seismic hazards and human-induced urban development patterns. The use case aims to investigate and visualize the risks associated with high-rise urbanization by estimating potential building damages and the corresponding economic losses resulting from a major earthquake - specifically a scenario event of magnitude Mw = 7.4. Istanbul lies near the North Anatolian Fault Zone (NAFZ), a major right-lateral strike-slip fault separating the Anatolian and Eurasian plates. The fault system poses significant seismic hazard to the Istanbul metropolitan area (Figure 5).

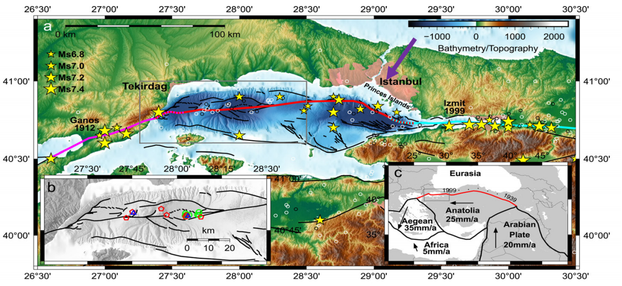


Figure 5. Seismicity and mapped surface fault traces in the Marmara region (Becker et al., 2023). (a) Rupture traces of shown earthquakes and the Main Marmara Fault (MMF) (red line) (b) Repeater sequence centroids (c) Plate tectonic setting and GPS-derived velocities relative to Eurasia.

4.1.1 High-rise Phenomena and Damage Vulnerability:

High-rise buildings (9 or more floors) can be vulnerable to earthquakes with a magnitude of about 7 or greater if they are not well-designed and constructed according to modern seismic standards. Their vulnerability depends on several factors, including their design, construction quality, materials, local seismic hazards, and soil conditions, and the seismic characteristics of the region. In risk assessments, reinforced concrete frame high-rise (9-20 story) buildings are often evaluated carefully due to their potential for significant economic and human losses in the event of failure. To investigate possible risks for high-rise buildings in the area of interest (AoI) Fikirtepe-Kadıköy (purple arrow in Figure 5), our use case implements a scenario-based earthquake shaking in greater Istanbul utilizing Earthquake Loss Estimation Routines (ELER) (Demircioglu et al., 2009) and develops tools to quantify mentioned vulnerabilities. One of the most important data in the estimation process is the composition of building inventory when its damages are considered to ensure a uniform interpretation. The proxy distribution of buildings is aggregated to 150 arcsec grids (≈3.5 km) to form the default building inventory data in greater Istanbul whereas it is 19.6 arcsecs for Fikirtepe AoI.

4.2 Methodology, Analyses and Results

The methodology and respective analyses involve following headings. Ground Motion intensity (Seismic Hazard) estimation:A deterministic earthquake scenario with Mw = 7.4, (Epicenter ;40°53’42” N; 28°43’ 44”E, Depth=13 km) is modeled using the Boore and Atkinson (2008) ground motion prediction equations (GMPEs). The intensity distribution (Figure 6) and relevant parameters like Peak Ground Acceleration (PGA) were estimated.

- Ground Conditions and Site Effects: Vs30 grid data is employed to represent local soil conditions (obtained average Vs30: ≈250–360 m/s).
- Building Inventory and Typology: The analysis focuses on newly constructed high-rise Reinforced Concrete (RC) buildings (≥20 stories) in Fikirtepe, developed under urban renewal programs.

Figure 6. The intensity distribution in Istanbul Damage Modelling and Estimates: The intensity based empirical vulnerability relationships developed by Lagomarsino and Giovinazzi (2006) are used to estimate RC damage grades (D3-D5) (see also Figure 7).

Older buildings (pre-1979) have notably higher damage rates, especially at D3 and D4 levels. This is likely due to inadequate structural design for seismic resistance and poor construction quality, exacerbated by high-rise vulnerability.

For damage estimates in AoI (Figure 8), building vector data for the region was obtained from OSM Buildings, then checked and

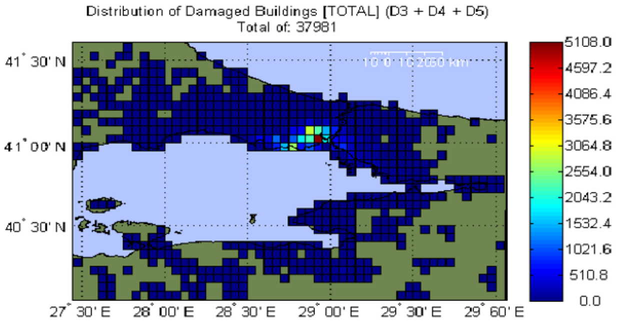


Figure 7. The distribution of damaged RC high-rise buildings for D3-D5 levels. Damage levels: D1-Slight, D2-Moderate, D3-Substantial heavy, D4-Very heavy, D5-Destruction.

Figure 8. Substantial to heavy damage (D3-D5) were computed up to 35–40% of tall RC buildings due to amplified shaking on soft ground and possible structural irregularities (torsion, weak stories, etc.) in AoI.

re-arranged using high-resolution (30 cm HD registered WV3) satellite imagery.

Urban Geometry Analysis: Building layout and street width-to-height ratios are spatially analyzed to evaluate secondary risks, including post-collapse debris blockage of narrow streets. In AoI, the height limit is set at 80 m (approximately 25–27 floors) and there are streets having widths changing to 8-20 m.

Although newer buildings are expected to avoid widespread collapse, estimated buildings in geo-cells shows that partial or total failure of individual towers. Using Bayesian method which directly correlates the street width with the probability of rubble blockage (Byun and D'Ayala, 2022), it has been computed the probability of a complete blockage based on existing street widths (Figure 9).

4.3 Immediate Policy Implications

Urban development planning should adopt minimum width-to-height ratios for streets to prevent complete blockage during collapses. Even for new buildings, traditional code compliance may not be sufficient. Land use and development policies should incorporate seismic risk layers to guide densification towards safe and adequate vertical growth.

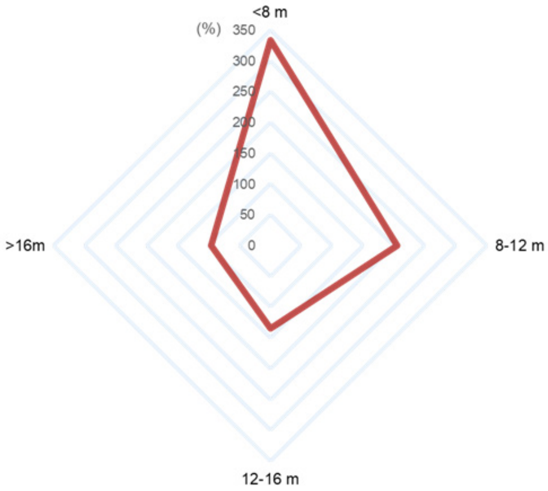


Figure 9. The probability of debris blockage (%) risk by street width.

5. UDENE Data Cube

The UDENE project requires the systematic processing of large-scale, multi-source, and temporally rich datasets to enable the analysis of natural experiments across diverse geographic regions. This type of data management involves more than just storing data; it also encompasses making it analysable, comparable, and reproducible.

The combined use of EO data and in-situ data provided by local institutions (e.g.: use case developers) presents various technical challenges, such as differences in format, resolution, temporal completeness, and coordinate systems. These issues pose significant obstacles to conducting comparative analyses that ensure the scientific validity of natural experiments. Therefore, it is essential to standardize, align, and prepare the data across spatial and temporal dimensions.

To address such a requirement, UDENE has developed a modern data cube infrastructure. The data cube integrates EO satellite imagery and local sensor datasets into a single, analysis-ready structure. This architecture makes it possible to compare intervention scenarios across cities, analyse temporal changes, and evaluate model outputs consistently. Unlike traditional data storage solutions, the data cube offers direct access to data along spatial and temporal dimensions, supports automated queries, enables reproducible scenario-based analyses, and provides scalable analytics for diverse user profiles. Within UDENE, the data cube functions not only as a data repository but also as an intelligent analytical engine.

This infrastructure has been built using modern, flexible, and open-source components. Data processing has been conducted with Python-based libraries (such as xarray, rasterio, and geopandas), while spatial data has been managed using a PostGIS-enabled database. So, it has been aimed to enable seamless integration of EO and in-situ datasets into a unified system with efficient access and high-capacity processing.

One of the most critical steps in the UDENE Data Cube is the harmonization of multi-source data flows. Satellite data is typically available in large-volume raster formats, whereas local measurements (such as air quality, traffic, or building data) often come in vector formats with varying spatial and temporal resolutions. All this data has been standardized using the Cloud Optimized GeoTIFF (COG) format. This conversion enables fast cloud-based data access and facilitates the integration of large datasets into web-based analysis tools. Working directly with COG files reduces processing time and improves reproducibility. Each dataset is accompanied by EO3-compliant metadata files. These files clearly define the spatial and temporal extent, measurement attributes, and data sources, ensuring transparency for users and enabling automated data management systems to function effectively. This metadata structure is a fundamental component of the data cube's sustainability and manageability.

The UDENE Data Cube provides a multi-source infrastructure that integrates EO and in-situ data. EO datasets from open-access sources such as Copernicus and Landsat are used to assess urban environmental indicators including vegetation distribution, air pollution levels, and surface temperature. In-situ datasets, covering traffic intensity, temperature, disaster, and seismic risk, are collected from local and national authorities and are harmonized with EO data for integrated analysis. Additionally, the data cube has been designed with the flexibility to ingest data from external open sources, enabling comparative analyses across cities and scenarios.

The data cube is organized to support both temporal and spatial dimensions. This allows users to investigate changes occurring within specific time intervals and geographic areas. Since all datasets are timestamped, retrospective analyses and direct comparisons between different time periods are possible. For example, data on temperature, air quality, or land use from different dates in each city can be compared to evaluate the effects before and after an intervention.

Data can be grouped into different temporal resolutions—daily, weekly, or monthly-allowing for both short-term impacts and long-term trends to be assessed. Defined Areas of Interest (AOIs) can be used to query the data cube, and temporal changes in those regions can be automatically extracted. This capability facilitates comparative studies between intervention and control areas. The integrated spatial-temporal structure enhances both the scientific accuracy of analyses and the reliability of decision-support processes.

Rather than offering a direct user interface, the UDENE Data Cube operates as a backend service for project tools such as the Exploration Tool and the Matchmaking Tool. These tools query and analyse data within the cube to support user workflows involving discovery, benchmarking, and decision-making. Users can define specific spatial and temporal parameters and conduct analyses accordingly, with results presented as map-based visualizations or comparative data outputs. All these interactions are powered by API services integrated with the data cube infrastructure. Figure 10 shows the workflow between the data cube and other technologies.

The data cube is hosted on an AWS EC2 instance and is configured with a horizontally scalable architecture. Its application components are containerized and orchestrated to ensure high availability and responsiveness under varying user loads. This design enables users to benefit from powerful data querying and analysis capabilities through intuitive and user-friendly interfaces, without requiring direct interaction with the underlying infrastructure.

Ultimately, the UDENE Data Cube is not merely a data storage system - it is a decision-support platform that delivers actionable insights to urban planners, environmental analysts, and disaster risk professionals. Its spatially and temporally aligned, analysis-ready data structure allows users to obtain reliable results without having to process complex and heterogeneous data sources.

One of the most important advantages of the data cube is its support for methodological reproducibility. Because analyses are based on well-defined data structures and timeframes, the same method can be applied repeatedly across different periods and locations. This enhances the credibility of scientific research and supports accountability in public decision-making.

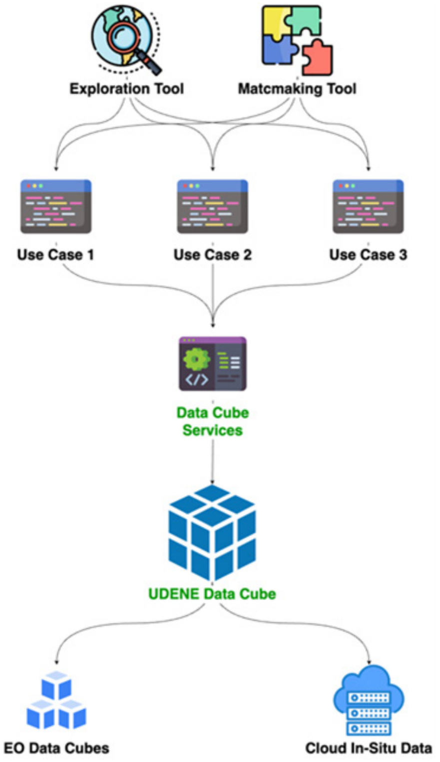


Figure 10. Workflow between UDENE Data Cube and other technologies.

Moreover, the data cube's geographically scalable design makes it adaptable to other cities and countries. Its data model and software architecture allow similar natural experiment approaches to be implemented in new regions. As such, the UDENE Data Cube provides not only a foundation for current case studies but also a versatile basis for future applications in sustainable urban development across Europe and beyond.

6. UDENE Exploration Tool

UDENE's Exploration Tool (ET) offers planners and decision-makers an interactive platform to simulate and assess urban interventions. Built on a Data Cube combining EO and in-situ data from case studies in Serbia, Tunisia, and Türkiye, it enables visualization of impacts on sustainability and resilience. By applying real-world data, the tool supports evidence-based urban planning tailored to environmental and hazard conditions.

The basic working structure of ET is based on finding similar places to the Area of Interest (AOI) around the world using spatial similarities defined by the input choices. Then, the variations in the target variable (such as temperature, economic loss or emission levels as defined in the use cases) are analyzed to capture any significant deviation.

6.1 Tool Architecture and Design Philosophy

The UDENE platform comprises two primary AI-powered tools built on a modern microservices architecture: the Exploration Tool and the Matchmaking Tool. Both tools share a common technical foundation featuring FastAPI backends for high-performance asynchronous operations and React+TypeScript frontends that deliver responsive, intuitive user experiences. This architecture ensures scalability while maintaining the system's ability to process large geospatial datasets within constrained computational resources (4 vCPUs, 16 GiB RAM).

6.2 The Exploration Tool: Interactive Geospatial Analytics

The Exploration Tool serves as the platform's primary interface for interacting with complex geospatial data cubes. Its design prioritizes accessibility without sacrificing analytical depth, offering multiple layers of functionality: Interactive Navigation and Visualization: Users can seamlessly navigate through satellite imagery using standard map controls (zoom, pan, layer toggling) while accessing advanced visualization options including temporal animations, 3D renderings, and dynamic heatmaps. These features transform static satellite data into compelling narratives about urban change.

6.2.1 Advanced Querying Capabilities: The tool implements sophisticated filtering mechanisms that allow users to query data based on spatial boundaries, temporal ranges, and thematic criteria. This functionality enables targeted analysis of specific urban phenomena, from heat island effects to vegetation coverage patterns.

6.2.2 Customizable Dashboards: Recognizing that different stakeholders have varying analytical needs, the tool allows users to create and save personalized dashboard configurations. This feature ensures that urban planners, researchers, and policymakers can each optimize their workflow according to their specific requirements.

6.2.3 Machine Learning Integration: The tool incorporates pre-trained models accessible through API endpoints including /predict and /find_similar, enabling users to perform predictive analytics and identify geographically similar regions based on environmental indices such as NDVI and LST. This capability transforms descriptive data exploration into prescriptive planning insights.

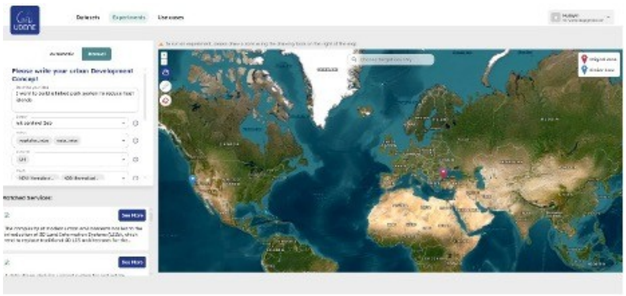


Figure 11. The exploration tool's integrated user interface for running an experiment.

6.3 The Matchmaking Tool: Connecting Users with EO Services

The Matchmaking Tool represents an innovative approach to democratizing access to Earth Observation services. Built as an intelligent recommendation engine, it bridges the gap between user needs and the vast ecosystem of EO products available through eoMALL. Natural Language Processing Pipeline: Users describe their urban development ideas through a simple form interface. The system employs advanced NLP models (Llama + LangChain) to extract key concepts and requirements from these natural language descriptions, eliminating the need for users to understand technical EO terminology. Intelligent Matching Algorithm: Extracted keywords are processed through a sophisticated matching pipeline that leverages both Cosine Similarity metrics and Elasticsearch to rank relevant service providers. The system presents the top matches with detailed explanations of their relevance, ensuring transparency in the recommendation process. Continuous Learning Framework: A built-in feedback mechanism allows users to rate the relevance of matches, creating a continuous improvement loop that refines the algorithm's accuracy over time. This adaptive approach ensures the tool becomes more effective as it accumulates user interaction data.

6.4 Technical Implementation and Deployment

The entire platform is containerized using Docker, ensuring consistent deployment across development and production environments. The services are hosted on AWS infrastructure, leveraging EC2 instances for compute, S3 for storing Cloud-Optimized GeoTIFFs, and RDS PostgreSQL with PostGIS for spatial metadata management. This cloud-native approach ensures both scalability and accessibility while maintaining robust security through carefully configured security groups and network policies.

6.5 Real-World Application and Impact

The platform's effectiveness is demonstrated through multiple use cases across diverse urban contexts. For instance, in Bakırköy, Istanbul, landscape architects have used the tool to model the potential cooling effects of converting Atatürk Airport into an urban park, validating hypotheses about green infrastructure's role in mitigating urban heat islands. Similar applications in Novi Sad (pollution impact assessment), Greater Tunis (green space analysis), and Kadıköy-Ataşehir-Üsküdar (earthquake preparedness) showcase the platform's versatility in addressing varied urban challenges. Through its combination of sophisticated analytics, intuitive interfaces, and intelligent service matching, the UDENE platform represents a significant advancement in making Earth Observation data actionable for urban development stakeholders, fostering data-driven decision-making that can lead to more sustainable and resilient cities.

7. Conclusions

The UDENE project demonstrates a transformative paradigm for urban analytics by integrating natural experiment frameworks with Earth Observation (EO) data and state-of-the-art modeling tools. Across its three diverse case studies, Serbia, Tunisia, and Türkiye, UDENE illustrates how EO-informed insights can move beyond descriptive mapping to support causal inference, predictive modeling, and real-world policy evaluation.

In Novi Sad, the combined use of Sentinel-5P data, in-situ sensors, and agent-based traffic simulations revealed how specific infrastructural changes, like bridge construction and pedestrian zones, can influence local air pollution. The application of multivariate regression and explainable AI (XAI) methods enabled transparent quantification of impact. In Greater Tunis, ensemble learning models were deployed to predict UHI intensity using vegetation indices and land surface temperature data. Results affirmed the cooling efficacy of urban green spaces and local climate zoning strategies, confirming that spatially targeted interventions can yield measurable thermal benefits. The Istanbul case study addressed disaster resilience, simulating a Mw 7.4 earthquake scenario to assess seismic vulnerability in high-rise zones. EO-derived building inventories, combined with ground motion prediction equations and regional vulnerability models, facilitated nuanced risk maps, estimating damage distribution and emergency accessibility.

A key strength of UDENE lies in its modular infrastructure, particularly the UDENE Data Cube, which enables dynamic querying and comparative analysis via the Exploration Tool and intelligent service matching through the Matchmaking Tool. These capabilities reduce the barrier for urban stakeholders to integrate EO data into policy workflows, making advanced analytics accessible even to non-specialist users.

Ultimately, UDENE exemplifies how EO, when coupled with natural experiment methodology and interactive tools, can guide urban interventions toward greater sustainability and resilience. By fostering causal understanding rather than mere correlation, UDENE provides a scientific foundation for evaluating “what works” in urban planning and how those lessons can be transferred across geographies. As urban environments grow increasingly complex and vulnerable, tools like UDENE become indispensable for informed, adaptive governance.

Acknowledgments

This work has received funding from the European Union’s Horizon Europe Research and Innovation Programme under grant agreement No. 101131190 within EUSPA. The authors would like to thank the European Commission for its support.

References

Becker, D., Martínez-Garzón, P., Wollin, C., Kılıç, T., Bohnhoff, M., 2023. Variation of Fault Creep Along the Overdue Istanbul-Marmara Seismic Gap in NW Türkiye. *Geophysical Research Letters*, 50(6), e2022GL101471. https://doi.org/10.1029/2022GL101471. e2022GL101471 2022GL101471.

Boore, D. M., Atkinson, G. M., 2008. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. *Earthquake spectra*, 24(1), 99–138. https://doi.org/10.1193/1.2830434.

Byun, J.-E., D’Ayala, D., 2022. Urban seismic resilience mapping: a transportation network in Istanbul, Turkey. *Scientific reports*, 12(1), 8188. 10.1038/s41598-022-11991-2.

Demircioglu, M., Erdik, M., Hancilar, U., Sesetyan, K., Tuzun, C., Yenidogan, C., Zulfikar, A., 2009. Technical manual-earthquake loss estimation Routine ELERv1. 0. *Bogazici University, Department of Earthquake Engineering, Istanbul*, 133.

Gianquintieri, L., Oxoli, D., Caiani, E. G., Brovelli, M. A., 2024. State-of-art in modelling particulate matter (PM) concentration: a scoping review of aims and methods. *Environment, Development and Sustainability*. https://doi.org/10.1007/s10668-024-04781-5.

Gurram, S., Stuart, A. L., Pinjari, A. R., 2019. Agent-based modeling to estimate exposures to urban air pollution from transportation: Exposure disparities and impacts of high-resolution data. *Computers, Environment and Urban Systems*, 75, 22-34. https://doi.org/10.1016/j.compenvurbsys.2019.01.002.

Kopanja, M., Maglevannaia, P., Obrenović, N., 2025. Integrating ML and XAI with Urban Planning: Air Quality Predictions to Support Traffic Optimization. *Preprint, Zenodo*. https://doi.org/10.5281/zenodo.15582760.

Kukic, M., Li, X., Bierlaire, M., 2024. One-step Gibbs sampling for the generation of synthetic households. *Transportation Research Part C: Emerging Technologies*, 166, 104770. https://doi.org/10.1016/j.trc.2024.104770.

Lagomarsino, S., Giovinazzi, S., 2006. Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings. *Bulletin of Earthquake Engineering*, 4(4), 415–443. https://doi.org/10.1007/s10518-006-9024-z.

Maglevannaia, P., Pejak, B., Brdar, S., Marko, O., Crnojević, V., Obrenović, N., 2025. An integrated framework for population synthesis and activity scheduling for transportation microsimulation. *34th European Conference on Operational Research (EURO 2025)*, Leeds, UK. June 22–25, 2025.

Oxoli, D., Ronchetti, G., Minghini, M., Molinari, M. E., Lotfian, M., Sona, G., Brovelli, M. A., 2018. Measuring Urban Land Cover Influence on Air Temperature through Multiple Geo-Data—The Case of Milan, Italy. *ISPRS International Journal of Geo-Information*, 7(11). https://www.mdpi.com/2220-9964/7/11/421.

Pugliese Viloria, A. d. J., Folini, A., Carrion, D., Brovelli, M. A., 2024. Hazard Susceptibility Mapping with Machine and Deep Learning: A Literature Review. *Remote Sensing*, 16(18). https://www.mdpi.com/2072-4292/16/18/3374.

Vavassori, A., Oxoli, D., Venuti, G., Brovelli, M. A., Siciliani de Cumis, M., Sacco, P., Tapete, D., 2024. A combined Remote Sensing and GIS-based method for Local Climate Zone mapping using PRISMA and Sentinel-2 imagery. *International Journal of Applied Earth Observation and Geoinformation*, 131, 103944. https://doi.org/10.1016/j.jag.2024.103944.