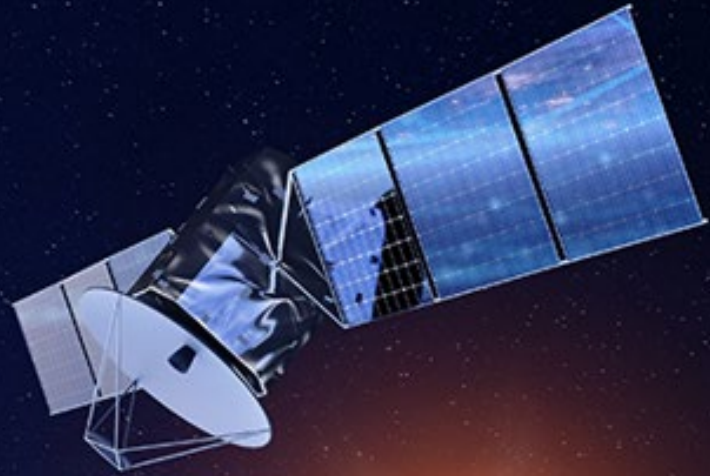


UDENE



URBAN DEVELOPMENT EXPLORATIONS USING NATURAL EXPERIMENTS

D3.1: Use Case Descriptions



The project has received funding from the European Commission's Horizon Europe Research and Innovation Programme under grant agreement No 101131190 with the EUSPA

Deliverable Information			
Work Package	WP3		
Lead partner	Politecnico di Milano (POLIMI)		
Author(s)	Nikola Obrenović (BioSense) Mohamed Rajhi (TUNSA) Onur Lenk (NiK Systems) Vasil Yordanov (POLIMI) Daniele Oxoli (POLIMI) Maria Antonia Brovelli (POLIMI)		
Due date	15.09.2024		
Version number	2.0	Status	Draft
Dissemination level	Public (PU)		

Document Information	
Project Number	101131190
Project Title	Urban Development Explorations using Natural Experiments
Acronym	UDENE
Start Date	1 January 2024
Duration	24 months
Call identifier	HORIZON-EUSPA-2022-SPACE
Topic	HORIZON-EUSPA-2022-SPACE-02-56 Strategic autonomy in developing, deploying and using global space-based infrastructures, services, applications and data 2022
Instrument	HORIZON-RIA
Project URL	www.udene.eu
EU Project Officer	Valeria CATALANO
Disclaimer	This document has been prepared within the scope of UDENE Project which received funding from the European Commission's Horizon Europe Research and Innovation Programme under grant agreement No 101131190 with the EUSPA. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. All the information in this document reflects only the Authors' view. Neither EUSPA nor the EC is responsible for any use that may be made of the information contained therein.



POLITECNICO
MILANO 1863



Revision History			
Revision	Date	Contributor(s)	Description
1.0	20.06.2024	Vasil Yordanov, Daniele Oxoli, Maria Antonia Brovelli (POLIMI)	Initial version
1.0.1	21.06.2024	Murat Ozbayoglu (TOBB ETU)	Made some cosmetic changes
2.0	15.09.2024	Nikola Obrenović (BioSense), Mohamed Rajhi (TUNSA), Onur Lenk (NiK Systems), Vasil Yordanov, Daniele Oxoli, Maria Antonia Brovelli (POLIMI)	Document title corrected; minor proofreading; added page numbers; added subsections in the table of contents; added Subsection 12 "Use case integration with UDENE's data cube and exploration tools" in the UC description form; Revisions according to the reviewers' comments

Quality Control			
Role	Date	Contributor(s)	Approved/ Comment
Internal Review	21.06.2024	Murat Ozbayoglu (TOBB ETU)	Approved
Final Quality Review			

Table of Contents

Table of Contents	3
Table of Figures	4
Purpose and Scope	5
Serbia Use Case	6
1. Use Case Title:	7
2. Urban Development Idea:	7
3. Objective:	8
4. Inverse Urban Development Idea:	8
5. Study area:	8
6. Study area characteristic variables.....	9
7. Temporal scope:	10
8. Input predictor variables	10
9. Outcome variables	12
10. Input dataset/s:	12
11. Methodology:	13
12. Use case integration with UDENE's data cube and exploration tools	16
13. Challenges/risks and mitigation strategies:	17
14. Planning:.....	18
Tunis Use Case	19
1. Use Case Title:	20
2. Urban Development Idea:	20
3. Objective:	20
4. Inverse Urban Development Idea:	21
5. Study area:	21
6. Study area characteristic variables.....	23
7. Temporal scope:	23
8. Input predictor variables	23
9. Outcome variables	24
10. Input dataset/s:	24
11. Methodology:	25
12. Use case integration with UDENE's data cube and exploration tools	28
13. Challenges/risks and mitigation strategies:	28

14. Planning:.....	29
Türkiye Use Case.....	30
1. Use Case Title:	31
2. Urban Development Idea:	31
3. Objective:	31
4. Inverse Urban Development Idea:	31
5. Study area:	32
6. Study area characteristic variables.....	33
7. Temporal scope:	34
8. Input predictor variables	34
9. Outcome variables	36
10. Input dataset/s:	36
11. Methodology:	39
12. Use case integration with UDENE's data cube and exploration tools.....	42
13. Challenges/risks and mitigation strategies:	42
14. Planning:.....	42

Table of Figures

Figure 1. Location of the bridges to be built as part of the outer transportation ring.....	7
Figure 2. The new pedestrian area to be created by closing a street, marked by the curvy blue line.	8
Figure 3. The location of Novi Sad within Serbia (Map data ©2024 Google).....	9
Figure 4. Utilization of the use case models by the Exploration Tool.....	17
Figure 5. Tunis study area.....	22
Figure 6. The General View of the Scenario Based Earthquake Effected area	32
Figure 7. Tectonically active area in wider European and Mediterranean region	33
Figure 8. General scheme of implementation has been given in the below flow chart.	40

Purpose and Scope

The primary purpose of this deliverable is to offer a comprehensive technical description of the three use cases central to the UDENE project. This document aims to provide a detailed account of the methodologies, data integration processes, analytical tools, and anticipated outcomes for each use case. Serving as a critical resource for project stakeholders—including urban planners, policymakers, researchers, and technical teams—the deliverable will clarify the project's objectives, technical strategies, and the application of Earth Observation (EO) data in addressing specific urban development challenges.

The document will present the technical descriptions of the three use cases: Türkiye, Tunisia, and Serbia. Each use case represents a unique urban challenge. For Türkiye, the focus will be on assessing the impact of earthquakes on high-rise buildings. This case study will cover the specific urban challenges related to seismic activity, the technical methodologies employed—including data acquisition and integration of EO data—and the analytical models used to estimate earthquake-induced building damage.

In Tunisia, the case study will be centered on addressing urban heat islands. It will describe the technical approaches used to map and analyze heat distribution. The use of EO data will be detailed, along with strategies for mitigating heat islands and improving urban microclimates. This section will highlight the methodologies and expected results, demonstrating the potential for improved urban planning and public health.

The Serbia use case will focus on mitigating air pollution and traffic congestion. It will describe the detailed methodologies for data collection, integrating EO data with in-situ measurements, and the analytical techniques used to evaluate pollution levels and traffic patterns. Proposed solutions will be discussed, along with their expected outcomes and impacts on urban mobility and air quality.

Use Case Description Form
***Assessment of the urban development plans
from environmental perspectives***

Project Coordinator: WEglobal
Work Package Leader: TOBB-ETÜ
Deliverable Leader: Politecnico di Milano
Use Case Responsible: BioSense

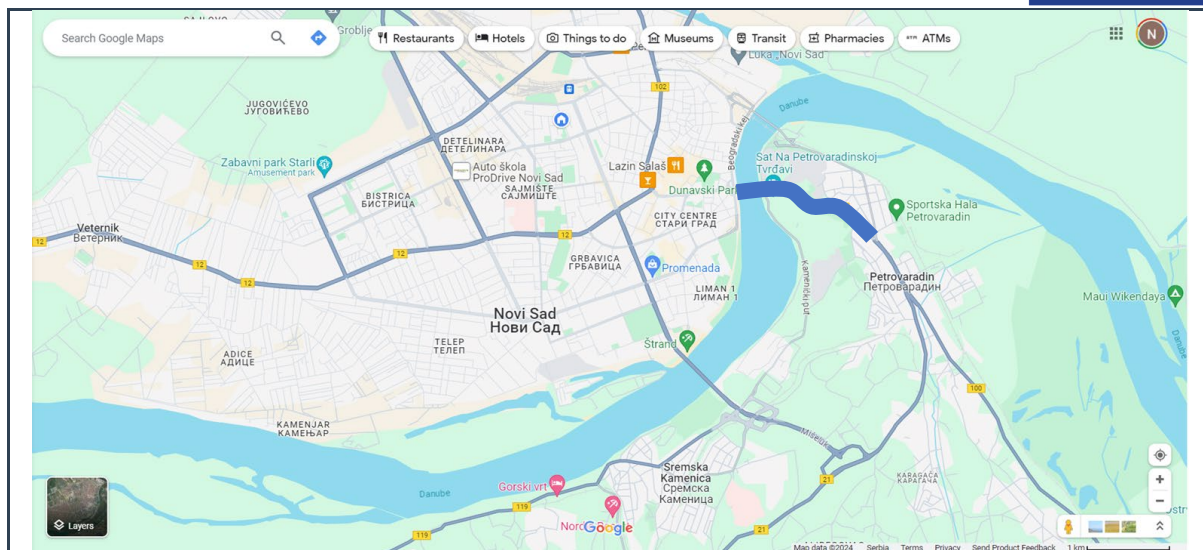


Figure 2. The new pedestrian area to be created by closing a street, marked by the curvy blue line.

3. Objective:

The objective of the use case is to analyze the effects of the planned transportation network changes to the emission of the pollutants, via changes in the city traffic patterns.

The broader research goal is to develop and provide general methodologies which will be initially evaluated on the Novi Sad case study but later extensible and generally applicable for other geographical areas of the similar characteristics.

4. Inverse Urban Development Idea:

- Concept: Determination of feasible transportation network characteristics by defining and solving an optimization problem.
- Objective: To minimize the emission of pollutants.

5. Study area:

The study area will be the city of Novi Sad, in northern Serbia. The city geocoordinates are 45°15'15"N 19°50'33"E, while the city location within Serbia is illustrated in Figure 3.



6. Study area characteristic variables

The city of Novi Sad is located in the south of the Pannonian plain, which makes it exposed to strong winds and large amount of dust. The characteristic attributes of the city are:

- Geocoordinates: 45°15'15"N 19°50'33"E
- Population: 368,967
- Surface: 702.7 km²
- Climate: temperate climate (Köppen climate classification: Cfb)
- Elevation: between 80 and 83 meters a.s.l.
- Characteristic wind: Košava blows from the Carpathians, making it an east-southeasterly wind, and usually brings clear and dry weather. The wind mainly occurs in Fall and Winter, with usual intervals of 2 to 3 days, with the average speed of 26 to 41 km/h. However, the highest recorded wind speed was 130 km/h.
- Climate data is given in the following table:

Climate data for meteo station Rimski Šančevi, Novi Sad					
Month	Winter	Spring	Summer	Fall	Year
Mean daily maximum °C	7.97	22.7	27.3	11.63	17.4
Daily mean °C	3.33	16.87	20.7	6.87	11.9

Mean daily minimum °C	1.9	10.97	14.67	5.25	7
Average snowy days	4.97	0.13	0	2.3	1.13
Average relative humidity (%)	78.83	67.33	69.37	81.97	74.5
Mean monthly sunshine hours	110.87	249.3	274.07	105	221.77
Source: Republic Hydrometeorological Service of Serbia					

7. Temporal scope:

We shall analyse the average daily traffic behaviour and consequent pollutants emissions on characteristic day (e.g., working day, Saturday, Sunday/Holiday) and seasons (e.g., Spring, Summer, Fall, Winter).

8. Input predictor variables

The following variables, given in Table 8.1, will be analysed and tested for the purpose of modelling the stated use case. Along with each variable, we note the originating dataset, from the list of used datasets given in Table 10. A more detailed description of the variables is given below Table 8.1.

Table 8.1. Independent variables for potential inclusion in the use case model(s)

Independent variable	Originating datasets*
Average integration (or closeness) of the transportation network	DS_MCT
Average choice (or betweenness) of the transportation network	DS_MCT
Normalized choice of the transportation network	DS_MCT
Average hourly traffic lane occupancy	DS_OTC
Average hourly vehicle speed	DS_OTC
Hourly traffic volume	DS_OTC
Average hourly number of bicycles	DS_HTC
Average hourly number of motorcycles	DS_HTC
Average hourly number of cars	DS_HTC
Average hourly number of buses	DS_HTC
Average hourly number of light cargo vehicles	DS_HTC
Average hourly number of medium cargo vehicles	DS_HTC
Average hourly number of heavy cargo vehicles	DS_HTC
Average hourly number of car transporters	DS_HTC
Average hourly number of tractors	DS_HTC
Wind speed	DS_ERA5
Dew point	DS_ERA5
Humidity	DS_ERA5
Distribution of the utilized vehicle types on the city level	DS_MCNS
Education levels per household on the city level	DS_MCNS
Occupation distribution on the city level	DS_MCNS
Historical city population	DS_CRS

*DS_ = Dataset

We further provide descriptions of and additional explanations for the used variables:

- One of the fundamental space syntactic measures is integration, or mathematical closeness. This is the calculation of how close or how accessible each spatial element is to all others under each definition of distance, such as the least angular distance. Integration can be used to assess how much potential the space has as a destination for movement, called the to-movement potential, by generating an integration pattern. The average integration is calculated as the arithmetical mean of the integration for all spatial elements.
- Another widely used syntactic measure is choice, or mathematical betweenness. This measures the degree to which each spatial element lies on the shortest paths, under each definition of distance, between any pair of spatial elements. Choice assesses the potential of the movements passing through each space, called the through-movement potential, in contrast to the to-movement potential measured by integration. The average choice is calculated as the arithmetical mean of the choice for all observed spatial elements.
- Normalised choice aims to solve the paradox that segregated designs add more total (and average) choice to the system than integrated ones. It divides total choice by total depth for each segment in the system. This adjusts choice values according to the depth of each segment in the system, since the more segregated is, the more its choice value will be reduced by being divided by a higher total depth number. This would seem to have the effect of measuring choice in a cost-benefit way.
- Average hourly traffic lane occupancy, vehicle speed, and traffic volume are gathered online by the automatic meters. The average lane occupancy and traffic represent the average percentage of utilized lanes and the number of counted vehicles during the observed time period, in our case one hour.
- Average hourly numbers of different vehicle types are manually counted vehicles, for several days in the past, on the selected segments and junctions in Novi Sad.
- Wind speed, dew point, and humidity are the standard meteorological measurements. Since these meteorological occurrences can influence the pollutant emissions, we include them in the model development.
- Distribution of the vehicle types represents the percentage of the utilized vehicle types by the Novi Sad citizens when performing daily routines, i.e., going to work, school, shopping, or for leisure activities.
- The educational levels tell the average number of persons per household that have 1) no education, 2) primary school only, 3) secondary school, 4) bachelor of applied sciences, 5) bachelor of sciences, or 6) higher levels of education.
- The occupational distribution gives percentages of Novi Sad inhabitants that work in the particular sectors, such as industry, agriculture, or households.
- The historical Novi Sad population numbers are used for the year which we have other data for. Therefore, to calculate the absolute numbers for the needed parameters, we rely on the percentages and population counts for the years of interest.

For modelling the inverse use case, we aim to define an optimization problem that will output the desired values of the transportation network characteristics, in order to minimize the pollutant

emissions. The variables listed in Table 8.2 will represent parts of the objective function of the optimization problem.

Table 8.2. Variables for potential inclusion in the inverse use case model(s)

Independent variable	Originating datasets
Nitrogen Dioxide (NO ₂) emissions	DS_S5P, DS_SEPA
Carbon Monoxide (CO) emissions	DS_S5P, DS_SEPA
Sulfur Dioxide (SO ₂) emissions	DS_S5P, DS_SEPA

9. Outcome variables

- Use Case – Primary Variable: Nitrogen Dioxide (NO₂) emission
- Use Case – Secondary Variable (*optional*): Particle matter (PM_{2.5} and PM₁₀) emissions, Carbon Monoxide (CO) emissions, Sulfur Dioxide (SO₂) emissions
- Inverse Use Case – Decision Variables: Average integration (or closeness) of the transportation network, Average choice (or betweenness) of the transportation network, Normalized choice of the transportation network.

10. Input dataset/s:

The following rows describe datasets that will be used in the research and modelling of the pollutant emissions models. For easier referencing, each dataset is assigned a code, which is given before the dataset's description.

a. Remote sensing datasets

Appropriate remote sensing data will be used for the Area of Interest (AoI):

DS_S5P: Pollutant emissions data from ESA Sentinel-5P – 10MB

- *Sentinel-5P data is currently utilized through Google Earth Engine (GEE) platform for experimental purposes, because of GEE's user-friendly interface and efficiency in conserving data storage. As we need to extract multiple satellite images for time series data, using the GEE platform allows us to avoid downloading all the images, thereby saving the storage space required for them. In that case the values to be stored are only data extracted from satellite images that correspond to the region of Novi Sad. During the course of the project, we intent to assess the usage of Copernicus Data Space Ecosystem JupyterLab service, for the same purpose, and switch to this option if technically possible.*

b. In-situ datasets

The following in-situ datasets will be used:

1. DS_SEPA: Historical and current measurements of emissions (NO_x, CO, SO₂, PM_{2.5}, PM₁₀), provided by the Serbian Environmental Protection Agency (SEPA) -1 GB
2. DS_HTC: Historical traffic counts from the Novi Sad City Department of Property and Investments - 1GB

3. DS_MCNS: Citizens transport behaviour and microcensus data, performed in 2008 on 2.5% of the city population for Novi Sad, from the Novi Sad City Department of Property and Investments - 0.5 GB
4. DS_OTC: Online traffic counts from the Novi Sad City Department of Transportation – 1GB

DS_CRS: Census data, performed in 2011 and 2022, for the whole of Serbia on the city and regional levels, from the Statistical Office of the Republic of Serbia – 200MB

c. Datasets from other models

- DS_MCT: Space syntax model results (traffic potential data) from Prof. Marina Carević Tomić. (200 MB)
- DS_ERA5: ERA5 global atmospheric dataset, with emphasis on wind speed, precipitation and humidity.
 - *ERA5 data from the Copernicus Atmosphere Monitoring Service (CAMS) will be retrieved through the API for analysis and model integration. CAMS service enables direct connection and downloading data via API where the data would be stored locally in the form of csv files.*

11. Methodology:

a. Input data assessment

All data has undergone the validation by the data providers.

b. Data Integration

Data will be integrated via geo locations and identifiers of the metering stations.

c. Model implementation

At the macroscopic level, we will develop a hierarchical multivariate model (HMM), to model dependency between traffic and transportation network properties, on one side, and pollutant emissions, on the other. The macroscopic pollutant emissions model will be developed by employing different regression techniques. Firstly, linear and random forest regression models will be deployed while the sensitivity analysis will be used to assess their quality and stability. In the later stages of the research, other hierarchical regression models will be potentially applied.

One of the key approaches for traffic simulation is agent-based modeling (ABM), which requires as input detailed data on population, trips (origin-destination), transport modes, a network of the transportation system, etc. Therefore, our first task is to generate agents that will reflect the real structure of the population with specific demographic characteristics and travel behaviours. However, due to people's privacy protection and lack of data, we need to apply several techniques for generating reliable data for the agents. These techniques include:

- Generation of population data, by applying the Iterative Proportional Fitting (IPF) algorithm is used, which combines microcensus and total census data for determining the contingency tables about citizens' travel habits.
- Generation of households' data, based on specialized Markov Chain Monte Carlo (MCMC) simulations. Th MCMC simulations are hierarchically organized so that the generation of

individuals is performed first, following the matching procedure of these individuals to households.

- Activity assignment: To simulate activity schedules, the Optimisation-based Activity Scheduling with Integrated Simultaneous, OASIS approach (Pougala, Hillel, Bierlaire 2023) is chosen. The OASIS framework integrates multiple choice dimensions (activity participation, scheduling, travel mode, and location choice) simultaneously, offering a comprehensive method to simulate individuals' daily schedules by capturing the trade-offs between these choices. This approach provides a realistic and flexible model of daily scheduling decisions, capturing the complex interactions between various factors that influence individual behavior. The key components of the approach are:
 - Preparation: This step involves data collection and formatting, including gathering detailed information on individual activities, modes of transport, socio-demographic characteristics - the data we can model from the synthetic population, and travel times - including start times and durations, which depend on activity type. Utility functions are defined to capture the preferences of individuals for different schedules based on the collected input data.
 - Choice set generation: Using the Metropolis-Hastings algorithm, we can sample a set of feasible schedules for each individual. This algorithm can sample schedules that reflect realistic constraints and preferences.
 - Discrete choice estimation: Applying Maximum Likelihood Estimation (MLE) for estimation of the utility functions parameters based on the generated choice sets.
 - Simulation of scheduling choices: OASIS simulates daily schedules by drawing from the estimated parameters and solving for the optimal schedule that maximizes total utility subject to various constraints. This simulation captures interactions and dependencies between different decisions, such as the choice of activity, start time, duration, and mode of travel. The constraints include time budgets, mode, and resource availability.

The entire process of agent-based modeling applied in this use case can be summarized with the following pipeline:



Furthermore, at the microscopic level, we will continue development of agent-based models (ABM) with advanced decision-making methodologies, using GAMMA programming language and technology. The output of ABMs will be used to estimate traffic emissions by the EU-standardized COPERT model.

d. Outcome validation

The output of HMM will be validated against aggregated emission estimations obtained from ESA Sentinel-5P data (DS_S5P), using standard statistical error techniques, in particular mean absolute and mean squared errors, as well as R-squared and adjusted R-squared.

The output of ABM will be validated with the average observed traffic counts (DS_HTC and DS_OTC), throughout the city of Novi Sad, and the average pollutant emissions value, obtained from the SEPA metering units in Novi Sad (DS_SEPA), using the same statistical errors.

12. Use case integration with UDENE's data cube and exploration tools

(Briefly introduce specific workflows or examples of data integration.)

As part of the Serbia's use case, several regression and simulation models are being developed, in python and Gamma programming languages. The models will be utilized by the user through the Exploration Tool, where the user will select the desired model and provide arguments for the model inference. Then, the Exploration Tool will execute the selected model and visualize outputs to the user. This workflow is presented as a sequence diagram in Figure 4.

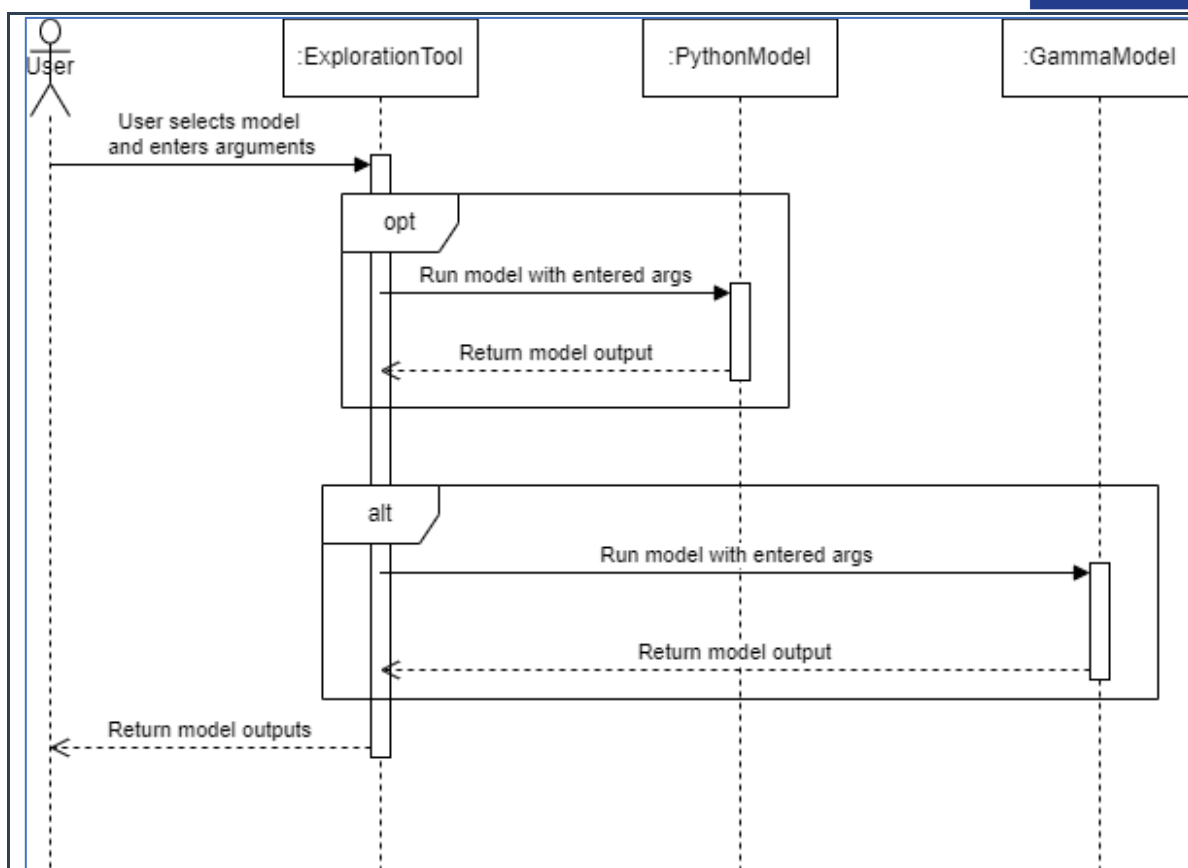


Figure 4. Utilization of the use case models by the Exploration Tool

Given that the exploration tool is developed in python, too, the python models will be integrated using the standard application programming interface (API), i.e., the models' functions will be directly called from the exploration tool's code base. On the other hand, the models developed in Gamma will expose a platform independent interface, e.g., in the form of REST API, which will be initiated from the exploration tool. Since all the models will be pre-trained and fine-tuned, the direct integration with the data cube is not envisioned for the time being. On the other hand, the data residing in the data cube will be utilized for the model development and available for the development of the future 3rd party models.

13. Challenges/risks and mitigation strategies:

(Potential challenges or risks associated with the use case, such as data limitations, technical issues, or external factors)

Challenge/risk	Mitigation strategy
Most of the datasets will not be available for publishing.	We will involve higher research instances, firstly the directorate of BioSense Institute, and University Open Data Officer. Furthermore, we will include our consortium partner, InoSens, due to their expertise, to guide us in the approaches to the public

	<i>authorities whom we intend to gather data from. .</i>
--	--

14. Planning:			
<i>(Internal Gantt per use case with defined subtasks, start and end month)</i>			
<i>Subtask (e.g. EO data collection)</i>	<i>Responsible POC*</i>	<i>Start month</i>	<i>End month</i>
<i>T.3.1.1. EO data collection</i>	<i>BioSense</i>	<i>2</i>	<i>12</i>
<i>T.3.1.2. In-situ and other data collection</i>	<i>BioSense</i>	<i>1</i>	<i>12</i>
<i>T.3.1.2. Model development and validation</i>	<i>BioSense</i>	<i>3</i>	<i>18</i>

**POC – Point of Contact*

Use Case Description Form
Effect of a linked park system on heat load

Project Coordinator: WEglobal
Work Package Leader: TOBB-ETÜ
Deliverable Leader: Politecnico di Milano
Use Case Responsible: TUNSA

1. Use Case Title:

Effect of a linked park system on heat load

2. Urban Development Idea:

The urban development idea revolves around the creation and enhancement of a linked park system within the city to mitigate urban heat islands (UHI). This idea would involve the strategic design and expansion of green spaces, parks, and urban forests to counteract the negative impacts of increased temperatures and create a more comfortable and sustainable urban environment. The idea is to create not just a series of green spaces, but a cohesive system that benefits the environment, the climate, and the community.

3. Objective:

Objectives:

- **Assess Heat Load Reduction within Local Climate Zones:**

The primary objective of the use case is to quantify and analyse the impact of a linked park system, considering Local Climate Zones (LCZs), on reducing the heat load within the urban environment. This includes measuring changes in temperature, heat absorption, and heat distribution in areas where the park system is implemented and assessing how different LCZs influence the effectiveness of park-based cooling strategies. Additionally, exploring the correlation between park cooling indices and LCZ characteristics to provide practical insights for urban planning.

- **Evaluate Park System Efficiency:**

The use case aims to evaluate the efficiency of the linked park system in mitigating urban heat islands by examining its cooling effect during various seasons and under different weather conditions, considering the variations in LCZs across the city.

- **Determine Spatial and Temporal Variations in LCZ-Based Heat Load Reduction:**

This use case seeks to identify spatial and temporal variations in heat load reduction within the city, considering factors such as park location, size, vegetation type, and proximity to residential and commercial areas, while also considering the LCZ classification of different areas.

Goals:

- **Quantify Heat Reduction and LCZ-Based Cooling:**

One of the main goals of the use case is to quantitatively measure the reduction in heat load achieved through the implementation of a linked park system within different LCZs. This will provide concrete data on temperature reductions and heat dissipation, specifically tailored to various LCZ categories.

- *Enhance Urban Resilience:*

The use case aims to contribute to enhancing the city's resilience to climate change by promoting sustainable urban development practices that improve the urban microclimate, thus reducing the health risks associated with extreme heat events across different parts of the city.

- *Inform Decision-Makers:*

The use case intends to provide decision-makers, urban planners, and policymakers with data-driven insights and evidence specific to LCZs to support informed choices regarding green infrastructure investments, urban development strategies, and policies tailored to the diverse urban fabric of the city.

4. Inverse Urban Development Idea:

- Concept:

The core idea is to understand and mitigate the Urban Heat Island (UHI) effect, which refers to the phenomenon where urban and suburban areas experience higher temperatures than their rural counterparts. This approach revolves around the integration of nature-based solutions, technological innovations and urban planning strategies to create cooler urban environments. It focuses on enhancing vegetation cover and improving urban design to facilitate natural cooling processes.

- Objective:

- *Reduce surface temperatures:*

By enhancing the Albedo (reflectivity) of urban surfaces, less solar radiation is absorbed, resulting in cooler surface and air temperatures. Through using cool roofs (materials that reflect more sunlight and absorb less heat) and cool pavements can reduce heat absorption.

- *Improve green infrastructure and water bodies:*

Expanding green spaces such as parks, gardens and greenways, also Incorporating water features like ponds, fountains, and artificial lakes, can reduce surface and air temperatures through shading and evapotranspiration.

- *Improving urban ventilation:*

Designing urban layouts that facilitate airflow. This includes designing cities in a way that maximizes natural ventilation and reduces heat accumulation.

5. Study area:

Greater Tunis is a metropolitan region situated in northern Tunisia along the Mediterranean Sea coastline. It encompasses the capital city of Tunis and its surrounding urban areas. Here are the geographical extent and characteristics of the Greater Tunis study area:

- **Geographical Extent and Characteristics:**

Location: Greater Tunis is positioned in northern Tunisia, bordering the Mediterranean Sea.

Study area coordinates:

Southwest (SW) Corner: 36°49'43.15"N latitude, 10° 7'55.65"E longitude

Southeast (SE) Corner: 36°49'56.69"N latitude, 10°22'35.8"E longitude

Northwest (NW) Corner: 36°54'39.6"N latitude, 10°00'35.7"E longitude

Northeast (NE) Corner: 36°54'58.5"N latitude, 10°14'55.86"E longitude

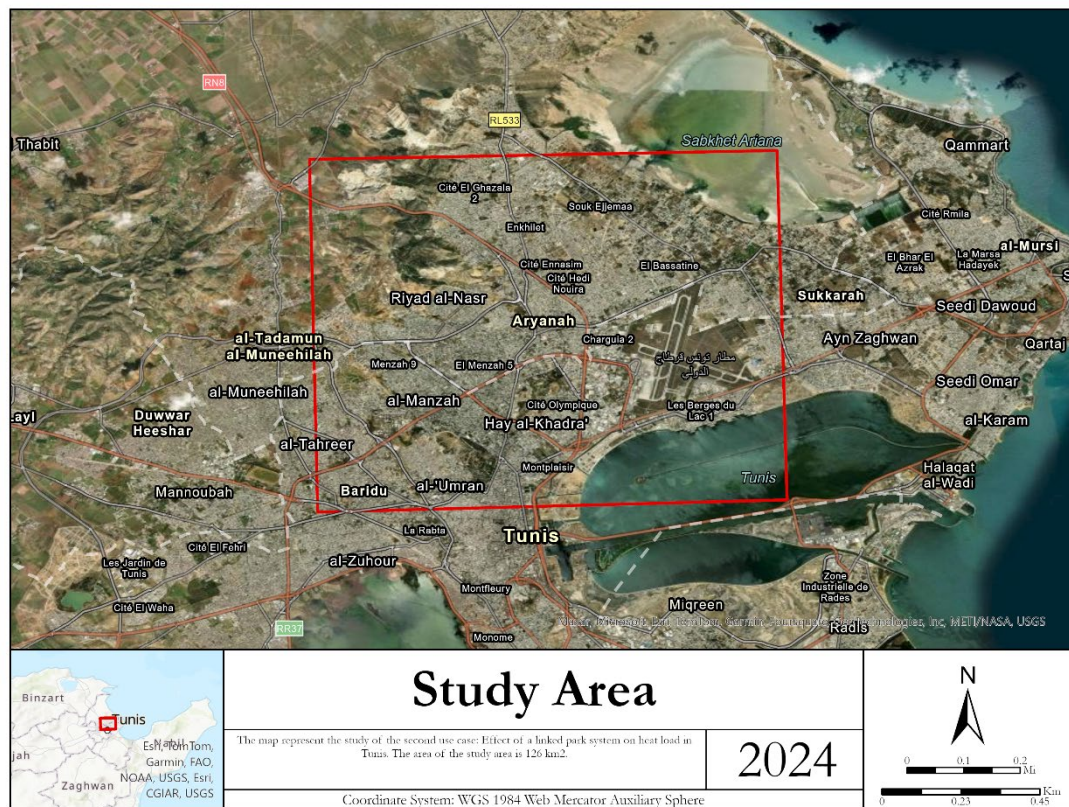


Figure 5. Tunis study area

Region Name: The study area is commonly known as "Greater Tunis" and includes key components such as:

- Tunis: Serving as the capital city of Tunisia, Tunis is located at the heart of Greater Tunis
- Ariana: Situated to the northeast of Tunis, Ariana is a prominent suburb and municipality.
-

6. Study area characteristic variables

- **Surface Area:** The study area covers a surface area of approximately 126 square kilometers.
- **Population:** Greater Tunis is home to a population of about 2.6 million inhabitants.
- **Urbanization:** It is highly urbanized, characterized by a dense concentration of residential, commercial, industrial, and institutional areas.
- **Geographical Features:** The region presents a diverse and complex environment, including the following features:
 - Basins: Greater Tunis includes various basins, which may affect drainage patterns and land use.
 - Plains: It comprises plains that are significant for agriculture, urban development, and infrastructure.
 - Lakes: The region may contain lakes, which could have ecological and recreational importance.
 - Lagoons: Lagoons are part of the coastal ecosystem, providing habitat for various species.

7. Temporal scope:

The temporal scope for the use case spans the previous 10 years, with a specific focus on 2023. This time frame will be the primary focus of the study, enabling an in-depth analysis of data, trends, and events that have occurred over the past decade.

8. Input predictor variables

List of the input predictor variables	
	Local Climate Zones (LCZ)
	Vegetation spectral indices (NDVI, EVI)
	Urban spectral indices (NDBI, BUI, ASI, ENDISI, BSI)
	Land Use Land Cover (LULC)
	Human settlement - Built-up
	Human settlement – population
	Urban Green Spaces (UGS)
	Land surface temperature (LST)
	Water surface spectral indices (MNDWI, NDMI))
	Wind direction and speed
	Humidity
	Topography

9. Outcome variables

- Use Case – Primary Variable: Heat load
- Inverse Use Case – Primary Variable: Surface Urban Heat Island (SUHI)

10. Input dataset/s:

a. Remote sensing datasets

Satellite data						
Data	Spatial resolution	Spectral bands	Applications	Format	Data Source	Link
Landsat Surface Reflectance	30m	7	Vegetation, built-up and water spectral indices	COG	ODC	https://github.com/opendatacube/datacube-dataset-config/blob/main/products/landsat_8_c2_l2.odc-product.yaml
Landsat Surface Temperature	30 m	11	LST	COG	ODC	https://github.com/digitalearthafrika/config/blob/master/products/ls8_st.odc-product.yaml
Sentinel 2	10/20/60 m	13	UGS	COG	ODC	https://github.com/opendatacube/datacube-dataset-config/blob/main/products/s2_l2a.odc-product.yaml

b. In-situ datasets

Dataset	Original format	Georeferencing	Source
Municipal Development plan	Raster	Georeferenced	Municipality

	LCZ	Raster	Georeferenced	WUDAPT	
c. Datasets from other models					
Generated products					
Provider	Product	Spatial resolution	Source	Applications	
Copernicus DEM	GLO-30	30m	https://github.com/opendatacube/datacube-dataset-config/blob/main/products/dem_cop_30.odc-product.yaml (available in the ODC)	Topography	
Copernicus GHSL	GHSL-Built-up	1 km	https://ghsl.jrc.ec.europa.eu/data.php (External to the ODC)	Human settlement	Built up
	GHSL-POP	1km		Human settlement	Population

11. Methodology:

a. Input data assessment

When assessing input data contributed by users, it's essential to perform several critical steps to ensure data quality and suitability. Here's our approach to the assessment of user-contributed data:

- Metadata Analysis:

Reviewing the provided metadata to understand the data's source, collection methodology, units of measurement, and any relevant information about the data's context and limitations.

Checking for documentation on data format, coordinate systems, and any transformations applied to the data.

- Data Quality Assessment:

Identifying any missing or erroneous values within the dataset. This may involve checking for data gaps, outliers, or inconsistencies.

Assessing the completeness of the dataset, ensuring that all required variables are present and appropriately recorded.

- Spatial and Temporal Coverage:

Confirming that the data covers the required geographical area for the specific analysis.

Checking that the data's temporal coverage matches the time range necessary for the analysis.

- *Data Format:*

We will examine the format of the data to determine if it is compatible with our tools.

- *Data Normalization:*

Normalization involves converting data to a common scale or format to facilitate meaningful comparisons and analysis.

We will normalize data units if necessary, ensuring that all variables are using consistent units of measurement.

- *Data Verification:*

Verifying that the data aligns with the user's stated objectives and research questions.

- *Data Security and Privacy:*

We will ensure that any sensitive or personally identifiable information (PII) is appropriately anonymized or masked to protect privacy and comply with data protection regulations.

By following these steps, we can thoroughly assess and prepare user-contributed data to ensure its quality, compatibility, and suitability for the analysis. This process helps minimize potential issues and uncertainties in the data and enhances the reliability of the results derived from it.

b. Data Integration

Integrating Earth Observation (EO) and in-situ data is a critical step in ensuring the accuracy and reliability of the analysis. Here's how the integration will be approached, along with methods for merging these datasets effectively:

- *Data Collection:*

EO data, including air temperature and wind speed, will be collected through remote sensing sources and the in-situ data will be collected from the field such as the ground-based weather stations, and the information on urban features and the land cover.

- *Accuracy Assessment / Verification:*

EO-derived products will be compared with corresponding in-situ measurements to assess accuracy and verify the quality of remote sensing data. Ground-based surveys will provide valuable ground truth data to validate and improve the accuracy of land cover classification obtained from EO data.

- *Validation:*

Ground-based measurements from in-situ sources are often considered as ground truth.

These measurements will serve as the basis for validating EO-derived data. Validation will involve comparing the spatial and temporal agreement between EO and in-situ datasets, assessing discrepancies, and quantifying errors.

Challenges

- *Resolution Mismatch:*

EO data may have different spatial and temporal resolutions compared to in-situ measurements, requiring interpolation or aggregation methods for effective integration.

Methods for Effective Merging

- *Spatial Analysis:* Conducting spatial analysis to align datasets properly.
- *Downscaling:* Applying of the downscaling methods to match the finer resolution of in-situ data with coarser EO data. Techniques like spatial interpolation or disaggregation can be used to bridge the resolution gap.
- *Validation Metrics:* Utilizing of the validation metrics (e.g., RMSE, MAE, correlation coefficients) to quantify the agreement and discrepancies between EO and in-situ measurements. This quantification provides insights into the accuracy of the integrated data.

c. Model implementation

The implementation of the model for assessing Surface Urban Heat Island (SUHI) susceptibility begins with the collection of Local Climate Zone (LCZ) data. This data is then combined with Land Surface Temperature (LST) statistics for each LCZ class. LST data, sourced from Landsat 8 and 9 satellites, is supplemented with additional indices such as the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Built-up Index (NDBI) obtained from Sentinel 2. To enhance the dataset, ancillary information including building type and population density is integrated.

The dataset is prepared by splitting it into training and testing sets. For modelling this data, we will use Random Forest (RF) regressors. The RF regressors are particularly suited for this task due to their robustness in handling large datasets and their ability to model complex relationships between multiple variables. The RF model will analyse the training data and generate predictions on SUHI susceptibility, which are then validated using the testing data. The final output is a set of SUHI susceptibility maps, which pinpoint areas vulnerable to urban heat island effects, providing valuable insights for urban planning and climate adaptation strategies. To minimize model uncertainty and identify the most influential predictors, sensitivity analysis will be conducted. This step is crucial for refining the model to more accurately capture the intricate relationships between predictors and urban heat islands, thereby enhancing its predictive performance and reliability.

d. Outcome validation

Outcome validation is a critical step in any research process to ensure the accuracy and reliability of the results obtained. In our context of assessing the impact of a linked park system on reducing urban heat islands in the Greater Tunis, outcome validation involves confirming that the findings align with expectations, are consistent with real-world observations, and meet the project's objectives. Here's how outcome validation can be conducted:

- **Field Validation:** We will conduct on-site field validation with GNSS technology to compare model predictions with actual conditions on the ground. We will ensure that the observed data collected in the field matches the trends and patterns predicted by the model. Any significant discrepancies should be investigated and addressed.
- **Urban Development Plan:** We will use the Urban Development Plan as a ground truth to do the accuracy assessment of produced predictors related to LULC.
- **Sensitivity Analysis:** Sensitivity analysis involves varying key input parameters to evaluate the stability of the model's outcomes. We will assess how changes in factors like park size, vegetation type, or impervious surface area impact the results.

12. Use case integration with UDENE's data cube and exploration tools

(Briefly introduce specific workflows or examples of data integration.)

The data / datasets available in the Open Data Cube and Digital Earth Africa data cube as COG are going to be called through APIs to be connected to the UDENE data cube, as well as Copernicus products. Other data and datasets will be converted and implemented following the guidelines given in the Data Cube Tools.

13. Challenges/risks and mitigation strategies:

Challenge/risk	Mitigation strategy
EO (Earth Observation) data and in-situ measurements may have different spatial and temporal resolutions, leading to challenges in aligning and merging datasets seamlessly.	Use spatial interpolation or resampling techniques to align datasets at a common spatial resolution. Consider the scale of the analysis and choose an appropriate interpolation method.
EO data may be subject to uncertainties related to atmospheric conditions, sensor characteristics, and retrieval algorithms.	Conduct validation studies comparing EO-derived values with ground truth measurements. Assess the accuracy assessment.

Discrepancies in the scale of measurements, especially in heterogeneous urban environments, can affect the representativeness of integrated datasets.	Utilize advanced EO technologies with higher spatial resolution (open source medium resolution satellite images) to capture fine-scale variations within urban areas
In-Situ data not publicly accessible	Establish direct partnership agreements with local actors for data access such as the municipalities and the Urban Planning Agency of Greater Tunis.

14. Planning:

(Internal Gantt per use case with defined subtasks, start and end month)

Add rows as much as needed

<i>Subtask</i>	<i>Responsible POC*</i>	<i>Start month</i>	<i>End month</i>
<i>T.3.1.1. Input data assessment</i>	TUNSA	02/2024	02/2025
<i>T.3.1.2. Data integration</i>	TUNSA	04/2024	02/2025
<i>T.3.1.2. Model implementation</i>	TUNSA	06/2024	06/2025
<i>T.3.1.2. Outcome validation</i>	TUNSA	04/2025	08/2025

**POC – Point of Contact*

Use Case Description Form
***High-Rise District Effect in the context of
Earthquake Preparedness***

Project Coordinator: WEglobal
Work Package Leader: TOBB-ETÜ
Deliverable Leader: Politecnico di Milano
Use Case Responsible: NiK Systems

1. Use Case Title:

Determination of Having a High-Rise District Effect in the context of Earthquake Preparedness

2. Urban Development Idea:

To ensure earthquake-resistant urbanization, realization of evidence-based land use planning to test the earthquake preparedness of high-rise buildings is an important concern. Sustainable Development Goal 11 (SDG11) titled “Sustainable cities and communities” is constituting specific reference for such an achievement in resilient urbanization. This goal aims to make cities inclusive, safe, resilient, and sustainable, envisaging the indicators:

- a. 11.5.1: Number of deaths, missing persons, and directly affected persons attributed to disasters per 100,000 population.
- b. 11.B: By 2020, substantially increase the number of cities and human settlements adopting and implementing integrated policies and plans towards inclusion, resource efficiency, mitigation and adaptation to climate change, resilience to disasters, and develop and implement, in line with the Sendai Framework for Disaster Risk Reduction 2015-2030, holistic disaster risk management at all levels.

3. Objective:

Implementation of a scenario-based earthquake to determine damages and losses relevant to high rise as part of yielding a casual effect for the urban development.

Goals:

- a. Conducting thorough assessments of the earthquake induced damage & loss to buildings, within high rise district.
- b. Determination of the rate of damage/undamaged buildings.
- c. Assessing the economic impact of the earthquake in the district.
- d. Defining in-situ variables to enable global search results for similar previous earthquake-affected areas.

A devastating earthquake having magnetic moment of $M_w = 7.0$ or greater in tectonically active fault lines is expected to expose hazardous effects wide area considering its source parameters like depth, duration of the earthquake as well as the length of ruptured fault line. The use case implementation naturally addresses physically damaged human-made settlements and their social and economic losses by making use of estimated parameters through ground motion prediction equations in the circumference of the emitted waves from the hypocenter (event location) in such an earthquake region. Additionally, to meet the Project's requirements, a specific area of interest is selected as the worst example which is worth to investigate consequences of irregular and intense construction of high-rise buildings and their mobility infrastructure.

4. Inverse Urban Development Idea:

- Concept: Best urban development characterizing resilient housing in evidence-based land use. To yield inferences for effective building planning scheme if not scenario-based event exposes such those destructions.
- Objective: Development options for minimizing damage/loss by avoiding high-rise (tall) buildings (in case they have been exposed to more than moderate damages by an $M_w \geq 7.0$ earthquake) within an effective building planning scheme.

5. Study area:

The Asian part of Istanbul, located in the Marmara Region, is at risk of a severe earthquake. The study area covering Kadıköy-Ataşehir-Üsküdar District Municipalities of İstanbul Metropolitan Municipality.

- The Geographical Coordinates (Coverage):
 $41^{\circ}00'35''N$ $29^{\circ}03'46''E$
 $40^{\circ}59'44''N$ $29^{\circ}06'58''E$
 $40^{\circ}58'47''N$ $29^{\circ}05'40''E$
 $40^{\circ}59'37''N$ $29^{\circ}02'55''E$
- Population: Approximately 383.596
- Density: Approximately 18.000 people/km²
- Surface Area: The study area covers a surface area of approximately 12 km² grid area.

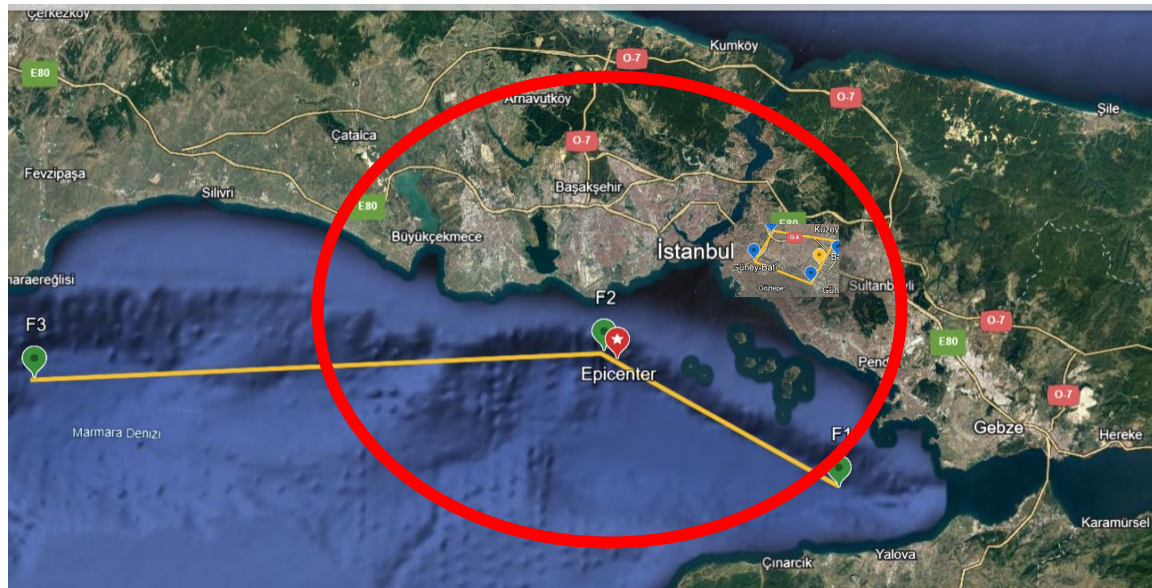


Figure 6. The General View of the Scenario Based Earthquake Effected area (Red circle having epicenter location at the center portraying intensely damaged area), whereas yellow rectangle is the area of interest for the investigation high-rise effect. The bold yellow lines are the active fault line in western segment of North Anatolian Fault Zone.

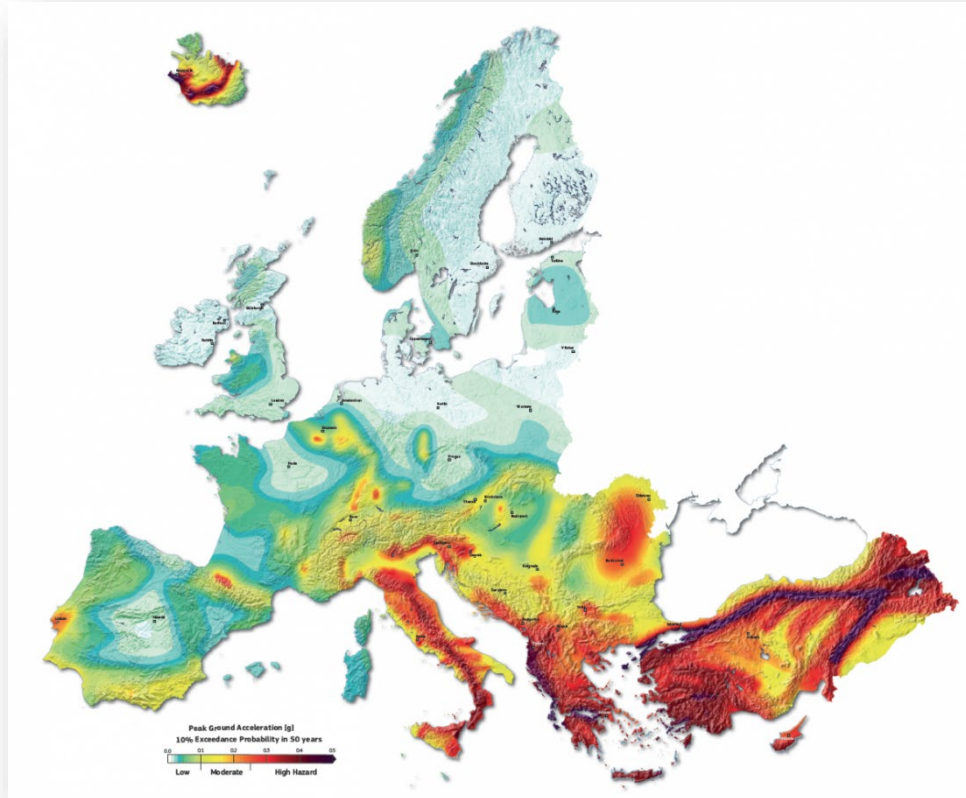


Figure 7. Tectonically active area in wider European and Mediterranean region including event area portrayed with the estimated Peak Ground Accelerations (PGA)*

Peak Ground Accelerations (PGA) from Wiemer S. et al (2016)¹

6. Study area characteristic variables

- Local urban maps and/or satellite images having resolutions adequately identify building information (especially high-rise buildings having equal and more than 8+ (namely 8-16) floors as well as surrounding higher and lower buildings). For building grouping, the European building taxonomy developed within the EU-FP5 RISK-UE project (Lagomarsino and Giovinazzi, 2006)² and model building types of HAZUS-MH (FEMA, 2003)³ will be adopted as respective datasets.
- Active fault line data/data sets (Maps and/or images portraying lines having moderate resolutions of about 1/1 M), since the area of interest is close to active North Anatolian Fault zone segment in Marmara Sea within less than 20 km.

¹ Wiemer S. et al (2016), Seismic Hazard Model 2015 for Switzerland (SUIhaz2015), ETHZ, Switzerland.

² Lagomarsino, S., and Giovinazzi, S., (2006). "Macroscopic and mechanical models for the vulnerability and damage assessment of current buildings." Bulletin of Earthquake Engineering, Vol.4, pp.415-443.

³ FEMA (2003). "HAZUS-MH Technical Manual", Federal Emergency Management Agency, Washington, DC, U.S.A

- Local Disaster (earthquake) maps or geospatial materials showing respective levels of earthquake for the region.
- Earthquake Catalogues giving respective information about previous events. The area is prone to an expected earthquake of about a magnitude 7.0 or greater according to the seismological and historical evidence and records in scientific journals and/or catalogues.
- Replacement Cost (RC) values and Building Damage (DK) State information. These are data sets (non-graphic) used to estimate direct economic loss.
- Structural Geology Micro zonation maps showing soil and ground types of the study area including but not limited the features,
 - Post-Tectonic Deposits
 - Devonian-Carboniferous Rock Units
 - Silurian-Lower Devonian Rock Units
 - Ordovician - Lower Silurian Rock Units

Above geologic conditions are considered with site conditions generally represented by a parameter: the upper 30-m average shear wave-velocity (Vs30). These Vs30 data is compiled from USGS Global Vs30 Maps.
- Human causality estimates for geographical area (from the respect of land use) or for population density (high populated area settled at high-rise buildings) correlating with the size of the earthquake (magnitude or intensity) as direct consequence of damaged buildings are being considered.

7. Temporal scope:

For temporal context, two implementations have been considered:

a. Earthquake event epoch at a specific daytime within scenario-based implementation in Aol (Istanbul Area).

(This comprises recent (updated) building inventory information, satellite images as base layers and updated active fault line (rupture) vector data -if applicable)

b. For global search (exploration) purpose of similar areas in the context of common variables stated in item 6 and 8 throughout world, about 10 years of past window will be considered.

8. Input predictor variables

Earthquake Event (Scenario to be implemented);
(The variables (parameters) to estimate Ground Motion Prediction and Intensity Distribution)

Magnitude

Epicentral coordinates

Focal Depth

Scalar value: Vector data*

Scalar value: Vector data*

Scalar value: Vector data*

Fault Line Rupture V30 Share Wave Velocity <i>(The variables (parameters) used to estimate Building Damages and direct economic losses)</i> Building Types Number of Buildings (population related) and number of Stories Age of Buildings Building Heights Building cost Demographic data	Vector Dataset and/or raster images Site Correction Map (Contour line, gridded data set) * Use case defined data or XML files to be compiled from catalogues https://deprem.afad.gov.tr/event-catalog Building Inventory Tabular and Textual Data* Tabular and Textual Data* Tabular and Textual Data* Tabular and Textual Data* Tabular, Raster Data and Web Service (App) * To be compiled from https://depremzemin.ibb.istanbul/wp-content/uploads/2020/11/Kadikoy.pdf https://ulasav.csb.gov.tr/dataset/34-mahalle-bazli-bina-analiz-verisi and https://land.copernicus.eu/en/products/urban-atlas/building-height-2012
Search/Exploration implementation for determining earthquake vulnerable areas similar to that of Istanbul Aol use case event: Magnitude Epicentral coordinates Global Fault Line Information Global Population Distribution Direct Economic Loss	Scalar value: Vector data* Scalar value: Vector data* Vector Dataset and/or raster images Vector Dataset Land use and human settlements datasets * Use case defined data or XML files to be compiled from national and global catalogues
Inverse Use Case:	

Implementation method is the same as above stated input use case variables, however evaluation will be performed in the context of low- and mid-rise for the building types having less than 8 floors envisaging effective building inventory suggestion.

9. Outcome variables

- Use Case – Primary Variable: EO base data (e.g.: satellite image) portraying Gridded Damaged Buildings with different levels of sophistication and their economic losses due to high-rise effect in the context of land use.
- Inverse Use Case – Primary Variable: EO base data (e.g.: satellite image) portraying Gridded un-damaged buildings having low and mid-rise.

Background of direct and reverse outcomes: The difference lies in the fact of scrutinizing best development options to minimize expected loss from damaged settlements (direct) and undamaged ones meeting certain building conditions (inverse). In the earthquake scenario based on Project requirements, to observe the damaged high-rise buildings (as part of building information like vulnerability classes, construction periods, other attributes etc.) and their economic aspects have been considered as the direct use case outcome, whereas the inverse outcome is to observe undamaged buildings and surrounding infrastructures (their construction characteristics etc.) and their economic gains in the same area subjected the same earthquake moment of magnitude, so to get best development options minimizing expected losses.

10. Input dataset/s:

a. Remote sensing datasets

The following data/datasets will be used in use case implementation.

Dataset	Sensor	Spatial Resolution	Spectral Bands	Temporal Resolution	Applications
Worldview-3 HD	MSI	15cm	4 VNIR	4.5 day	Update recent building inventory
Quickbird	MSI	60 cm	4 VNIR	1-3.5 day	12 years differences to find out building before 2002 after.

b. In-situ datasets

Dataset Name	Data Type	Source	Scale	Source*
Istanbul Province Microzonation Projects	JPEG, PDF	TR	1: 150 000	https://depremezmin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/
Structural Geology Map	Raster	TR	1:150 000	https://depremezmin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/

Geology Map	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Land Suitability	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Structural Geology Distribution Basemap	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Landslide Location Basemap	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Flooding Hazard Map Distribution	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Soil classification by Euro Code Map	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Hydrogeology Distribution Map	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Geology Distribution	Raster	TR	1:150 000	https://depzemzemin.ibb.istanbul/calismalarimiz/tamamlanmis-calismalar/istanbul-ili-mikrobolgeleme-projeleri/	
Distribution Map of the Number of Buildings According to the Number of Floors in Ataşehir	Raster	TR	1:175 000	https://depzemzemin.ibb.istanbul/wp-content/uploads/2020/11/Atasehir.pdf	
Distribution Map of the Number of Buildings According to the Number of Floors in Kadıköy	Raster	TR	1:175 000	https://depzemzemin.ibb.istanbul/wp-content/uploads/2020/11/Kadikoy.pdf	
Distribution Map of the Number of Buildings According to the Number of Floors in Üsküdar	Raster	TR	1:175 000	https://depzemzemin.ibb.istanbul/wp-content/uploads/2020/11/uskudar.pdf	
Neighborhood Based Building Numbers (2017)	CSV	TR	-	https://ulasav.csb.gov.tr/dataset/34-mahalle-bazli-bina-analiz-verisi	
Building Density Map	API	TR	-	https://ulasav.csb.gov.tr/dataset/34-yol-bakim-web-servisi	
Mineral Research and Exploration General Directorate(MTA)	Link	TR	-	http://yerbilimleri.mta.gov.tr/anasayfa.aspx	
Distribution of Building information(damaged) and the after/before 2023	Raster	TR	Dynamic	To be acquired from Province Administrations of Ministry of Environment,Urbanization and Climate Change	

K.MAraş earthquake area satellite images					
*some of the URL's might have location-based access restrictions					
c. Datasets from other models					
Provider	Product	Source	Data Type	Source	
Global Earthquake Model (GEM)	Global Seismic Risk Map - Area - Buildings - Fatalities - Losses	EU	Raster	https://www.globalquakemodel.org/product/global-seismic-hazard-map	
OpenQuake Map Viewer	Number of Buildings	EU	Raster	https://maps.openquake.org/map/global-exposure-map/#8/41.156/29.756	
ESRI/ArcGIS	Global Active Earthquake Faults	TR	Vector	https://www.arcgis.com/apps/mapviewer/index.html?layers=37a384d4c1ef4f56a33a40f291a634e9	
Global Human Settlement Layer	- Built-up Surface - Built-up Height - Built-up Volume - Built-up Characteristic	EU	Raster	https://ghsl.jrc.ec.europa.eu/download.php?ds=land	
World Population Dataset	2000-2010-2015-2020-2022 Population Data	EU+TR	Numeric	https://www.kaggle.com/datasets/iamsouravbanerjee/world-population-dataset	
National Center for Environmental Information (NOAA)	- Year - Region/Country/Area - Earthquake Location - Magnitude - Focal Depth - Hazard Association - Earthquake Effects - Damage - Deaths/Deaths Description - Injuries/Injuries Description - Damage/Damage Description - Houses - Damage/Houses Damage Description	USA	Numeric	https://www.ngdc.noaa.gov/hazel/view/hazards/earthquake/search	

GEM Global Active Faults	• Average Dip • Average Rake • Catalog ID • Catalog Name • Lower Seis Depth • Net Slip Rate • Slip Type • Upper Seis Depth	EU	Vector	https://blogs.openquake.org/hazard/global-active-fault-viewer/
Building Height	• Height • Area	UK	Numeric/3D Visualization	https://buildingheights.emu-analytics.net/main?layers=greatBritain&urbanAreas=&options=&baseMap=darkmatter
OSM Building	Height	TR	Numeric/3D Visualization	https://osmbuildings.org/?lat=40.99896&lon=29.09888&zoom=16.4&tilt=30
Urban Atlas Building Height	Height information in the spatial resolution 10 m for core urban areas in 870 cities across Europe for the 2012 reference year.	EU	Raster	https://land.copernicus.eu/en/products/urban-atlas/building-height-2012
Landscan Population	Remote Sensing-based global data modeling and mapping 30 arc-second (~1km) Density (people/km ²) Count (People).	USA	Raster	https://landscan.ornl.gov/
Digital Elevation Model-GTOPO30	DEM data having 30-arc second resolution (worldwide coverage with 33 tiles)	USGS/USA	Tabular/grid data	https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-global-30-arc-second-elevation-gtopo30
Copernicus GLO-30-Digital Surface Model (DSM)	Global coverage DSM at a resolution of 30 meters. (2023 release)	ESA-Copernicus	Tabular/grid data	https://spacedata.copernicus.eu/collections/copernicus-digital-elevation-model

11. Methodology:

a. Input data assessment

(Only for data contributed by users, if the case)

b. Data Integration

The data integration is nominally realized that the estimated quantities (output data) are visually mapped onto satellite images and/or any geospatial material available in Copernicus Data Space or Global Spaces ensuring the damaged buildings and associated casualty or loss coverages are spatially

mapped. For the methods for merging the datasets and estimated quantities, common GIS tools will be adopted. The challenges may be faced in the phase of creating respective APIs for the integration.

c. Model implementation

(1) Earthquake Event (The general methodology steps within the scenario to be implemented in Istanbul Aol)

(a) For a given earthquake magnitude and epicenter information, estimation of the spatial distribution of selected ground motion parameters through region specific ground motion prediction equations and using shear wave velocity distributions.

(b) Estimation of the building damage and human casualty at different levels of sophistication that commensurate with the availability of inventory of human-built environment.

(c) Estimation of direct economic losses stemming from building damage.



Figure 8. General scheme of implementation has been given in the below flow chart.

The analysis method: The spectral acceleration-displacement-based vulnerability assessment methodology is utilized for building damage estimation. To estimate this, an average number of dwelling units per building type (which is usually a function of the number of floors) is input. The intensity based empirical vulnerability relationships developed by Lagomarsino and Giovinazzi (2006)⁴ are used for the building damage assessment. Using average number of dwellings per building type and the grid-based population data, casualty estimation is computed with Samerdjieva & Badal (2002)

⁴ Lagomarsino, S., and Giovinazzi, S., (2006). "Macro seismic and mechanical models for the vulnerability and damage assessment of current buildings." Bulletin of Earthquake Engineering, Vol.4, pp.415-443

⁵approach. The casualty estimation is done through number of damaged buildings using HAZUS99 (FEMA, 1999) and HAZUS-MH (FEMA, 2003) methodologies. For the rapid estimation of earthquake shaking and losses analyses, Earthquake Loss Estimation Routines (ELER) will be adopted, which was developed under the JRA-3 component of the EU FP-6 NERIES Project.

Economic loss is, essentially, the translation of physical damage into total monetary loss using local estimates of repair and reconstruction costs. The economic loss in the model relies on user-defined loss ratios for the damage states. Damage states are based on the European Macroseismic Scale (EMS-98) assigning seismic intensities in European countries. The loss ratio is defined as the ratio between repair and reconstruction costs of a structure. The loss ratios are used to convert the number of damaged buildings in each grid to cell-based loss ratio values. The monetary value of direct economic losses is obtained by multiplying the loss ratios with the total building cost for each building type of the used inventory. This can be expressed with the following simple equation:

$$Loss (Btype D_k) = LR (D_k) \times RC (Btype)$$

where the loss ratio LR is a function of the building's damage states, namely from D_k ($k=1-5$), and the replacement cost RC is defined for each building type (Btype) in the building database. The overall economic loss is obtained by grid-based aggregation.

The integration and display of estimated quantities with the selected EO base data, we will employ common methodologies provided by GIS tools.

(2)Search/Exploration implementation for determining similar earthquake vulnerable areas similar to that of Istanbul AoI use case event.

Search mechanism in Data Cube will be established and executed according to parameters given in Section 8 above. Once similar earthquake prone areas have been found within acceptance level of uncertainties provided by global EO data, then the estimation of damages and associated parameters will then be computed as given in f (1) above. The validation of the methodology will be tested for those determined areas within the uncertainties associated search parameters and the geophysical constraints.

d. Outcome validation

Damage to the urban environment, and thus economic loss are compared and validated with the results obtained from the evaluation of the same routines in the areas having experienced almost same earthquake event before (E.g.; 6th Feb, 2023 Kahramanmaraş, Pazarcık (Mw=7.8) and Elbistan (Mw=7.5) earthquakes' pre- and post EO images are going to be exploited in the analysis for the comparisons). The buildings having stories 8 + are considered as high-rise settlements. The similar segments of urban areas as that of Istanbul IoA are going to be examined with the images and the building information provided by local administrations along East Anatolian Fault line to evaluate

⁵ Samardjieva, E. and J. Badal, 2002. Estimation of the expected number of casualties caused by strong earthquakes. Bulletin of the Seismological Society of America, 92 (6), pp. 2310-2322, Aug 2002.

general aspects of the algorithms used in the use case. Another validation process is anticipated to use with different earthquake damage routines/applications in the same area of interests.

12. Use case integration with UDENE's data cube and exploration tools

(Briefly introduce specific workflows or examples of data integration.)

The respective data/data sets and APIs are going to be transformed and used into the UDENE Data Cube along with the guidelines given in the Data Cube Tools.

13. Challenges/risks and mitigation strategies:

Challenge/risk	Mitigation strategy
Absence of similarities for the model inputs/parameters in different area of interest (Aoi).	Selection of generalized global inputs/parameters within statistical confidence level of uncertainty expected in the estimation.
High-resolution and representative data sets are required to evaluate earthquake analysis.	Selection of alternative variables is being adopted to get indicative decisions by using other sources in CDSE and global data lakes.

14. Planning:

Subtask	Responsible POC*	Start month	End month
T.3.1.1. Definition of Türkiye use case	NiK System	2	5
T.3.1.2. Compilation of data	NiK System	2	6
T.3.1.3 Experimental Analysis in Aoi	NiK System	4	10
T.3.1.4 Development of API (In connection with GraphQL/REST for DC integration)	NiK System	5	20

*POC – Point of Contact



The project has received funding from the European Commission's Horizon Europe Research and Innovation Programme under grant agreement No 101131190 with the EUSPA.

UDENE

