EUREKA CLUSTER ITEA3 Software intensive systems and services

Results of the benchmarking in the defined case Scenarios

Deliverable 3.4

POLicy Data Exploitation & Re-use

POLDER

PROJECT IDENTIFIER	POLDER
PROJECT TITLE	POLicy Data Exploitation & Re-use

DOCUMENT VERSION	1.0
PLANNED DELIVERY DATE	M43 (June 2022)
ACTUAL DELIVERY DATE	M43
DOCUMENT TITLE	Results of the benchmarking in the defined case Scenarios
WORK PACKAGE	WP3
DOCUMENT ACCESS	Public

ABSTRACT	The objective of this document is to carry out an analysis to empirically measure the accuracy and suitability of the proposed algorithms and solutions for each of the proposed use cases.
KEYWORDS	

FUNCTION	DATE	ENTITY
	25/06/2022	ACD
	22/06/2022	FORTEARGE
	22/06/2022	BEIA
	22/06/2022	ACCURO
	23/06/2022	QUOBIS
	27/06/2022	NOMMON
Author	03-01-2022	VEXIZA
	22/05/2022	ACCURO
Reviewer		

PARTNERS CONTRIBUTIONS RECORD

#	Partners	Contributor on Phase 1	Date of Contribution	Contributor on Phase 2	Date of Contribution2
1	ACCURO			X	22/05/2022
2	ACD			X	28/06/2022
3	ARD GROUP				
4	BEIA			X	22/06/2022
5	ForteArGe			X	22/06/2022
6	FCC				
7	Keypro Oy				
8	Mantis Software				
9	NOMMON			X	27/06/2022
10	Quobis			X	23/06/2022
11	SIS				
12	Wealize				
13	Vexiza			X	03/01/2022

CHANGES RECORD

VERSION	DATE	ENTITY	DESCRIPTION OF CHANGES
0.1	03/01/2022	VEXIZA	Base document
0.1.1	22/05/2022	ACCURO	Document structure and guidelines
0.2	15/06/2022	ACD	Added Air quality modelling and monitoring and Energy monitoring information
0.3	22/06/2022	FORTEARGE BEIA ACCURO	Contributed to the regional population estimation, Air quality in Galati industrial area and Image recognition for Smart Tourism, respectively
0.4	23/06/2022	QUOBIS	Added the Voice recognition and sentiment analysis Use case
0.5	27/06/2022	NOMMON	Added the Estimation of tourist flows Use case
0.6	28/06/2022	ACD	Added Air quality modelling and monitoring use case benchmarking
1.0	30/06/2022	ACCURO	Information compiling, final editing and formatting

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1. INTRODUCTION

1.1. DOCUMENT OBJECTIVES AND SCOPE

The objective of this document is to provide an analysis of the algorithms to be used in the defined scenarios and why these are the convenient ones in each case. By going through this document, the reader will be able to understand the theoretical effectiveness of the algorithms for the data available.

1.2. DOCUMENT STRUCTURE

The document is organized in 2 sections. Section 1 presents introductory information and Section 2 is Benchmarking and Testing. The second one is further divided into each of the use cases presented in D3.1. Moreover, each subsection provides information on the data concerning the use cases evaluated, and an explanation why the algorithm is suitable for that situation.

2. BENCHMARKING

A relevant part launching of a new project and project closures is to conduct benchmarking, i.e., to evaluate and analyse the project to be started or the product to be developed with respect to other similar projects or products with which we may have to compete, and take them as a reference point to position ourselves in the market and differentiate from them.

Not only does this allow us to evaluate the possibilities of entering a market, but also to know what currently exists and what aspects can be improved once the project or product development is finalised.

In this section, the consortium has performed a benchmarking of the use cases developed in the project, critically analysing which aspects could be improved in order to be better accepted in the market.

2.1. UC-1: SMART TOURISM (TOURISM)

2.1.1. ESTIMATION OF TOURIST FLOWS

This use case consisted in development of a solution for the monitoring and prediction of tourism flows, integrated by a descriptive and a predictive module. The descriptive module aimed at leveraging from the rich information provided by mobile phone data and its fusion with other data sources for the generation of tourism activity indicators. In particular the generation of three specific indicators was considered by this solution: number of pernoctations, number of daily visits and number of hourly visits per zone in the region of study. For all these indicators the following segmentations were considered: type of visitor (resident, national, no national) age, gender and visit purpose (for residents and national visitors), nationality and length of visit. The predictive solution aimed at the prediction of the mentioned indicators aided by the use of machine learning algorithms and exploiting historical data of the indicators generated by the descriptive module. For the developed of the predictive module a time series analysis based on long sort term memory (LSTM) networks was used.

The case study was tested and demonstrated for the city of Madrid for April 2019. The described indicators were calculated for the whole month of April and predicted only for the last week of April, to test the accuracy of predictions. The first three weeks of April were used to train the predictive model. The results were compared against those estimated for the same week with the descriptive module to assess the accuracy of the predictive solution.

The solution was tested in the city of Madrid. For this test the city was divided in small zones of $1km^2$ and the different indicators were estimated at a zone level. Once the indicators were calculated for each zone, each zone was characterised according to the type and mix of visitors they attract (e.g mainly nationals, mainly international, mixed etc.) and the activities performed in them (main hours of activity and zone's facilities -hotels, restaurants, etc.-)

The developed algorithms for indicators estimation and prediction as well as the zones classifications were integrated in an interactive visual analytic tool for the facilitation of tourism monitoring and scenarios analysis for tourism planning and management.

2.1.1.1. RESULTS

Four main results can be highlighted for this case study: the development of a series of algorithms for the estimation of pernoctation, and presence indicators from mobile phone records, the development of accurate algorithms for the prediction of the same indicators, a methodology to classify different zones of a region accordingly to its tourist activity and the development of the visual analytic tool. From the obtained results we can highlight the following:

Four main results can be highlighted for this case study: the development of a series of algorithms for the estimation of pernoctation, and presence indicators from mobile phone records, the development of accurate algorithms for the prediction of the same indicators, a methodology to classify different zones of a region accordingly to its tourist activity and the development of the visual analytic tool. From the obtained results we can highlight the following:

1. The proposed methodology and the developed algorithms the estimation of tourism activity, form the fusion of mobile phone data with other data sources, has allow us the extraction of relevant tourism indicators not only at a very high level of spatial and temporal disaggregation but also with a rich segmentation (nationality, length of stay, type of activity, etc.). For instance, Figure 1 shows the plot of the evolution of Portuguese visitors during April segmented by the number of overnights staying in Madrid. This enables services providers and public administrations to monitor and assess the punctual effect of policies and events implement in specific areas of the region.
2. The predictive models developed in the context of the project provide an accurate prediction of tourism flows for the near future with a minimum of information required for training the predictive models. Figure 3, shows the results of obtained for the prediction of

number of visitors sleeping in each of the zones in which the city of Madrid was divided. Both the relative and absolute error are small proving the good performance of the model. This kind of models enable services providers to plan for the optimal allocation of resources.

3. The zones characterisation provides a natural division of the region of study as a function of the visitors it attracts and the activities performed in them. As an example, Figure 2 shows the classification of Madrid as a function of the type of visitors present during a given day. This classification allows the identification of those services that attract a given type of visitor, providing tourist planning authorities and services providers to design the services offer according to the type of visitors they expect or desire to attract.
4. The interactive platform enables relevant actors to monitor different areas of the region of study and facilitate the comparison of different scenarios. This supports the benchmarking and analysis of different zones.

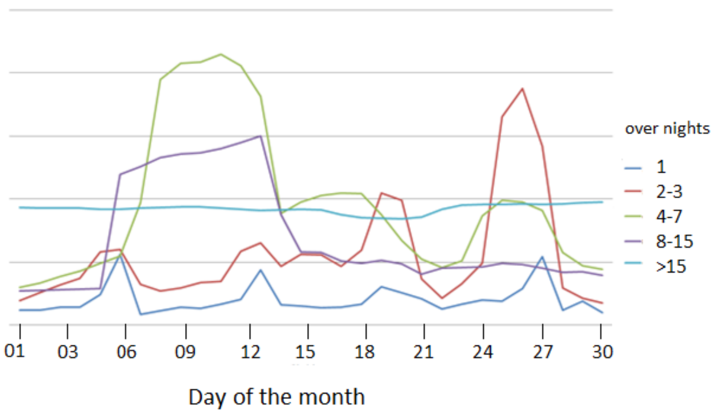


Figure 1. Number of Portuguese visitors sleeping in Madrid during the month of April 2019, segmented by the number of overnights

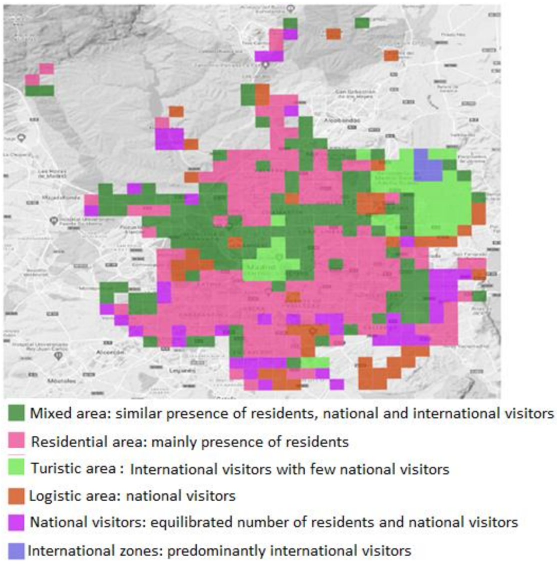


Figure 2. Classification of the different zones of Madrid as a function of the type of visitors it attracts and the activities performed there.

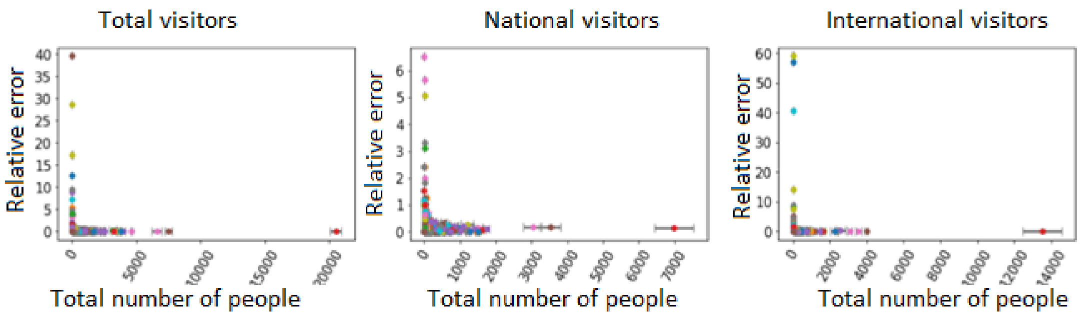


Figure 3. Relative error of the predicted number of visitors sleeping in each zone of the city as a function of the predicted volume. The bars indicate the absolute error. It can be seen that those areas with a higher absolute error correspond to areas with a higher number of visitors. Hence although high in relative terms this error represents a low percentage of the total volume.

2.1.1.2. USE CASE BENCHMARKING

This section provides a benchmarking of the results previously discusses respect to the solutions existing in the market. Although the final product of this use case is the platform which integrate all the different results obtained in this use case we will provided a short discussion of the main advantages and a benchmarking when applicable of the different subproducts that compound it.

Indicators: Monitoring of tourist flows has traditionally been done through the use of surveys. This is an activity that has been mainly performed official statistics offices. These methods have some intrinsic limitations: they are expensive, and as a consequence the studies provide a limited sample size and are not very frequently updated, they are highly aggregated and do not provide longitudinal data. In addition to surveys, tourist administrations (and official statistics) rely on demand information provided by attractions providers (museums, services rental, etc.) and official accommodation's establishments (hotels, hostels, etc.). However, these places do not have the whole picture. For instance, they fail to capture visitors staying at family or frindes, etc.

In the last years some new actors have appear in the sector of tourism to provide relevant indicators extracted from non- conventional data sources (e.g mobile phone data, social networks, etc.). From these actors the main competition to the solutions developed in this use case are:

- **Luca**: Luca is the data unit of the mobile network operator. Lucca has the advantage of having guaranteed mobile access to registration data in those countries in which Telefónica-Movistar has a presence. Through its Smart Steps solution, it offers information on the presence and mobility of tourists. For international tourists only sample's number are provided, i.e the sample is not expanded to the total if visitors and indicators consider only the users connected to Telefonica's/Movistar's network. In contrast with this limitation, the solution developed in POLDER provides total number of visitors (nationals and internationals) together with information about the characteristics of the visitors (e.g age and gender for nationals, nationality, length of visit, etc.)
- **Flux vision**: Flux vision belongs to Orange France. This very similar to that of Luca, with the difference that this one does provide total numbers of visitors foreign. On the other hand, the Orange solution has much more limited when it comes to making a longitudinal analysis of visitors.
- **Positium**: Positium is a company based in Tartu, Estonia. It is one of the first European companies to exploit information from mobile phone data for the study of tourism. The information provided is at the sample level, which, as in the case of Luca, supposes an important limitation.

The describe companies base their solutions on the exploitation of information from mobile data and do not play much attention of the enrichment of them by its fusion with other data sources. On the other hand, with the exception of Positium, which offers certain indicators through an API, the solutions offered by the other companies are through consulting projects designed ad-hoc for each occasion. The main difference between them lies in the ability of the proprietary algorithms developed by each of them to provide an accurate and reliable information. In this sense, it should be noted that Nommon has received a higher technical score than its competitors, such as Luca and Vodafone, in all the public tenders to which it has been presented in Spain in the last years, which indicates that, to date, Nommon's solution is the most advanced in the market.

Prediction: There is a large body of research on the use of machine learning models for the prediction of tourism flows. However, most of it is restricted to academia. Additionally, the predictive models developed so far require a large amount of historical data not easily available and are focused on the prediction of far future. As an example, one can look at the work of Abdelhgany et al. (2010), Andreoni (2006) Wu et al. (2012) and Lin et al. (2013) among others.

Zones classification: None of the commercial products available in the market for tourism business intelligence provide with an analysis and classification of the region. Most of them provide indicators based on pre-stablished zoning systems losing the opportunity of taking advantage of the natural division this classification provides.

Platform: There are different companies and institutions dedicated to the generation of dashboards for tourism monitoring and planning. The demand of these platforms is rising among those regions that aim to become smart destinations. Some of the most examples are:

Specialised tourism business intelligence companies commercialising these platforms, like Mabrian do not generate the raw indicators deployed in the platforms, they buy them from other companies like Nommon. Hence, in this sense rather than competitors are clients of one of the sub-products of the platform developed by Nommon. In addition to the fact that the solution developed in POLDER covers from the indicators generation to the monitoring, analysis and visualisation of them, an advantage of this solution over the platforms currently available in the market is that it not only provides descriptive scenarios of the tourist flows but it also predictive estimations.

2.1.2. VOICE RECOGNITION AND SENTIMENT ANALYSIS

The voice recognition and sentiment analysis tool provides real time information about positive or negative sentiments in a conversation. As first step it identifies the language and then it analyses the voice transcription, providing anonymized information about the sentiments. The conversation or speech can come from an audio collection system or from a video stream.

The outputs (language and real time information about positive or negative sentiments (absolute and subjective values) can be integrated into dashboards or platforms to be taken into account jointly with other information to obtain a variety of services or combined information.

2.1.2.1. RESULTS

The result of this use case is an autonomous real time tool for obtaining information regarding the satisfaction, disgust, conformity, etc. This information is useful for improve the services under study and adapt them to the customers/users requirements. This system can be also integrated with real time communication platforms to add its functionalities.

2.1.2.2. USE CASE BENCHMARKING

There are very accurate solutions on the market for transcription and sentiment analysis, like CloudFactory, Bandwatch, Sprinklr, etc. The main selling point of the system developed in this project is that it can be integrated in a variety of communication systems to better fits the requirements of a complex context. It was not designed to be the best content analyser but it can integrate other analytics tools, even context specific tools, to better fits the sentiment analysis, providing the best algorithms for the most demanding architectures.

2.1.3. IMAGE RECOGNITION SYSTEM FOR SMART TOURISM

For this use case, a new tool capable of analysing images in order to extract features for them has been developed. This tool, based on an Artificial Intelligence module with neural networks, takes as input either static (photos) or dynamic (video) images captured by cameras located in cities to identify in them a variety of objects comprising a wide range of vehicles (cars, motorbikes, trucks, buses, bicycles and scooters), people (either alone or in groups of different sizes), animals (dogs and horses) and objects that can be related to tourism, such as backpacks or suitcases. In addition, it is possible to detect specific situations related to the COVID-19 regulation, implemented to impose a series of safety measures that were mandatory until recently to prevent the spread of the virus, including:

1. Verifying that the safety distance between people is complied with.
2. Detecting the use of facemasks.
3. Generating alerts in case of non-compliance of these security measures.

This use case aims to provide the tourism sector, municipal authorities and service providers with tools for the analysis of population analysis to extract patterns of people's behaviour, especially in tourist areas, and allow users to improve the services offered to the population or to provide more information to their customers, for example, about peaks and off-peak hours of a place or the traffic situation in a given area.

To do this, the system generates a series of indicators based on the data extracted from the images, and displays them on easily customisable control panels, which also includes different types of graphs and maps that allow the user to visualise the information easily and quickly, and to make decisions based on the data consulted.

2.1.3.1. RESULTS

As mentioned above, the system takes as input images, which can be static or dynamic, taken with cameras located at strategic points in the cities. These images are analysed to identify and tag the aforementioned objects of interest. After labelling the objects in the image, the algorithm behind the system counts the objects detected and groups them according to type, as shown in Figure 4.

Figure 4. Object tagging and counting

This information, which is stored in a database, is then used in the control panels and maps to be represented and in the generation of alerts that facilitate decision making.

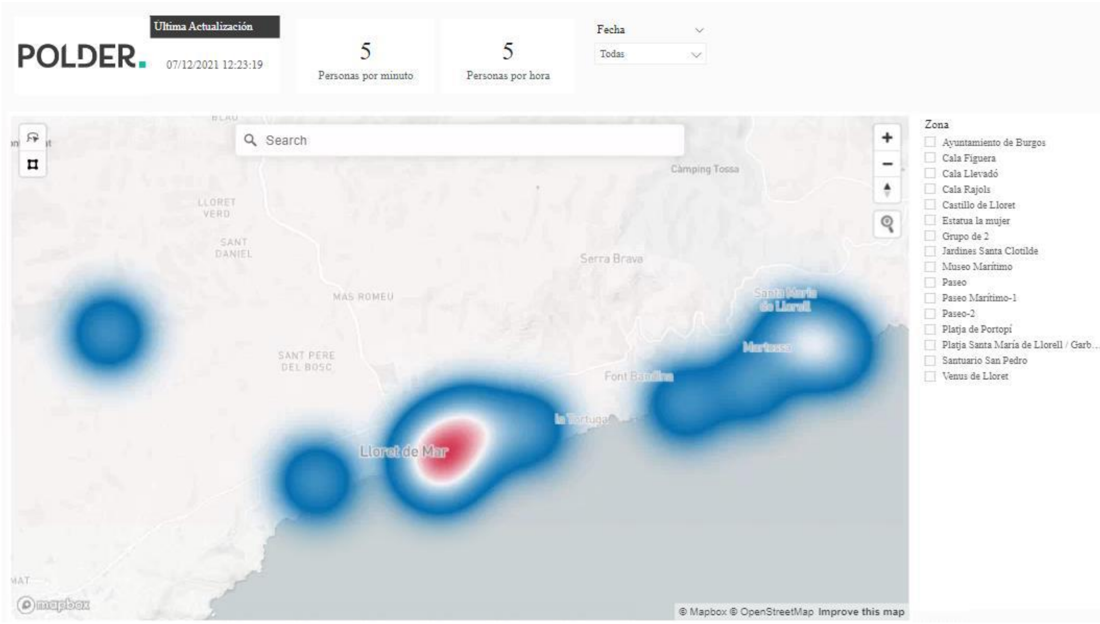


Figure 5. Heat map of the Image recognition system use case

It is important to highlight that although these images are used to extract information, they are not stored anywhere; only the information extracted from them by means of the AI module is stored and historicised.

In this way, the identity of the people and the vehicles appearing in the images analysed is protected, as the only thing of interest for us is the count of these characteristics in general, and not the recognition of individual people or vehicles.

2.1.3.2. USE CASE BENCHMARKING

The use of artificial vision systems in Smart Cities is nothing new: over the last 10 years cities have been implementing numerous systems that use image recognition to collect information about cities and streamline processes mainly related to their management.

Perhaps, one of the first and best-known use cases of AI and image analysis is reading the license plates of vehicles at the entrances of public garages to keep track of the entrances and exits, or for the control of traffic rules violations both in urban areas and on the road. However, there are already many image recognition applications that can help other day-to-day activities of the city, such as detection of occupancy/free spaces in garages, the detection of the use of facemasks and social distance, detection of weapons – unfortunately, of great importance in countries where the use of firearms is allowed and poses a danger to society – and even the prevention of suicide in public spaces.

The possibilities of using computer vision systems in Smart Cities are endless and therefore, there are a large number of very strong competitors such as ATOS, which has a wide range of products for Smart Cities using computer vision and is already working on the integration of 5G technologies in their systems (https://atos.net/en/lp/digital-vision-for-5g/smart-cities).

Focusing more on the computer vision systems for the Smart Tourism use case, the competition is greatly reduced as the demand is not very high yet. The only competitor found regarding Smart Tourism is **Clarifai** (https://www.clarifai.com/). Clarifai is a market leading AI development company since winning the top five places in image classification at the ImageNet 2013 competition. Using image classification technology, not only do they offer AI modules for detecting people and vehicles, but also modules geared towards the hotel sector that allow automated check-in through facial recognition or recognize VIP clients, among others. The company operates mainly in North America.

Although we have only been able to find one competitor, it is an area of knowledge that has recently begun to be explored. An example of this is presented in [1], where it is shown that people from different parts of the world have different visual perception, something that could be used in the tourism sector for improvement purposes depending on the type of visitors to be attracted, or for marketing purposes.

This use case still has a lot of room for improvement in terms of interfaces and development of new functionalities, and still has time to do so before the demand of such solutions increases. This is seen as an advantage for he future, because when this time comes, our solution will be better established than those that are starting to be developed now. On the other hand, the goal of the development realized in this use case was more oriented towards demonstrating the ease of integration with other Smart City tools, something that has been realized during POLDER project.

2.2. UC-2: CITY EXPANSION PLANNING AND URBAN TRANSFORMATION (PLAN)

2.2.1. URBAN POPULATION PREDICTION AND EFFECT ANALYSIS

This Sub-Case deals with the design and development of Population Prediction Engine for city government and policy/decision makers. In this sub-case, the population growth at a given time is predicted.

2.2.1.1. RESULTS

As a result of this subcase, a "Regional Population Estimation Model with High Temporal Resolution" was developed. The main features of the Model are as follows.

- Estimation of populations in time and location with an explainable AI model.
- Animate the estimated populations.
- Provide a base for value-added-services like waste management.

The Model animates the estimated populations with animated maps and heat maps and provides a base for value-added-services like waste management by using Time Use Survey (TUS) Data, GPS Data, Check-in Data from Social media platforms, Land Use Data and Regional Population statistics.

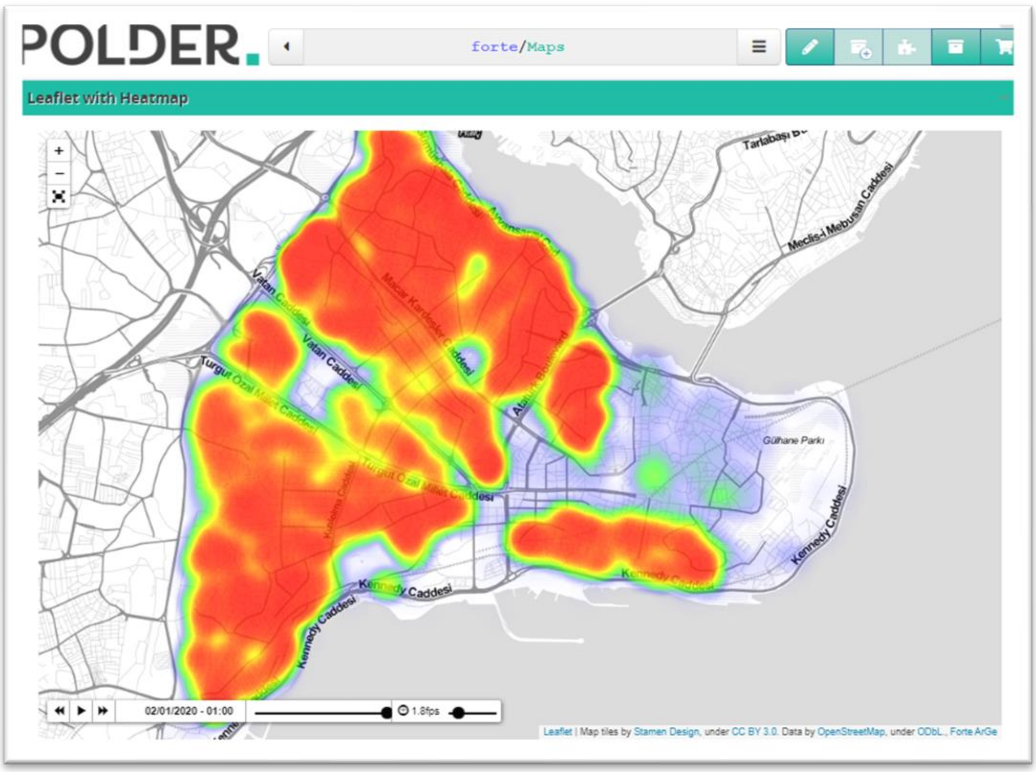


Figure 6. Screenshot of the heat map.

2.2.1.2. USE CASE BENCHMARKING

The main innovation of our solution is that the model employs open-source mobility data to learn mobility behaviors in the city and estimates high resolution populations in time and space. In addition, it can be easily integrated with other estimation models like waste, traffic and provides a base for them.

Although there are smart city estimation models or tools such as crowd estimation (video-based), traffic prediction and modelling, parking availability prediction, air quality/pollution prediction etc. related population changes, we did not find a direct competitor during our research in terms of estimating regional and temporal population and its effects.

2.3. UC-3: CITY MONITORING (MONITOR)

2.3.1. AIR QUALITY MODELLING AND MONITORING AND ENERGY MONITORING

Analysing and detecting abnormal values of the sensors for air quality is a crucial policy for maintaining health protection. Detection of anomalies, sending warning messages and visualizing the sensor values helps to suppliers for management of the air quality.

Air quality is an indicator that measures how air polluted are. On the other words, air quality index describes the level of air pollution that can be identified with the index value between 0 to 500. Higher air quality index value indicates that the higher pollution.

Air quality levels can be divided into six classes that are good, moderate, sensitive group, unhealthy, very unhealthy and hazardous. Each of the class presented with different colours that are illustrated in Table 1.

Table 1. Air Quality Index levels

Air Quality Index		
AQI Category and Color	**Index Value**	**Description of Air Quality**
Good Green	0 to 50	The air quality is good, and pollution offers little or no danger
Moderate Yellow	51 to 100	Some people, particularly those who are unusually sensitive to air pollution, may be at risk
Unhealthy for Sensitive Groups Orange	101 to 150	Members of vulnerable populations may suffer health risks. The broader population, on the other hand, is less likely to be impacted
Unhealthy Red	151 to 200	Sensitive groups can experience serious health effects. On the other hand, general public can also be affected
Very Unhealthy Purple	201 to 300	The risk of the air threatens everyone
Hazardous Maroon	301 and higher	Emergency condition. The air affects more likely everyone

In the scope of this study, air quality modelling system is divided into 2 sub system. Former is based on machine learning approach; latter is notification system that considers the statistical distribution. In this section both approaches will be investigated separately.

Machine learning approaches are the broad field of computer science that aim to teach computing process to machines intelligently as how humans can perform. The models require the same amount of data that should be enough to learn patterns underlaying the data. In this context, the data is acquired through various APIs. The collected data includes the 4 sensor values belongs to 113 countries with 5537 observations. The data is collected carefully in order to provide balanced dataset in order not to biasing. The sensors are $PM_{2.5}$, PM_{10}, NO_2 and O3 which is the most important attributes for determining the air quality index.

2.3.1.1. RESULTS

In modelling process Random Forest is fit to predict air quality. Even though it is a regression problem, the predicted value converted into categories which describes the quality of air in more understandable way. The achieved RMSE for the problem is that 2.61. In order to better understanding how the achieved result is satisfactory, the comparison can be done between actual and predicted values. Figure 7 demonstrates that there is no deviation between actual and predictions. The obtained results are pretty accurate.

	actual	preds
0	88.0	89.08
1	86.0	86.09
2	82.0	83.27
3	84.0	83.46
4	88.0	88.02
5	96.0	95.66
6	106.0	105.34
7	105.0	105.06
8	108.0	107.94
9	112.0	111.81

Figure 7. Comparison between actual and predicted values

Due to the machine learning approaches has complex intrinsic structure, the methods are not purely understandable. However, there are some techniques that can increase the understandability of the models. Feature importance is a method that describes the importance level of each attribute for predicting outcome. In this case, Figure 8 illustrates that PM_{10} is the most salient feature for predicting air quality index.

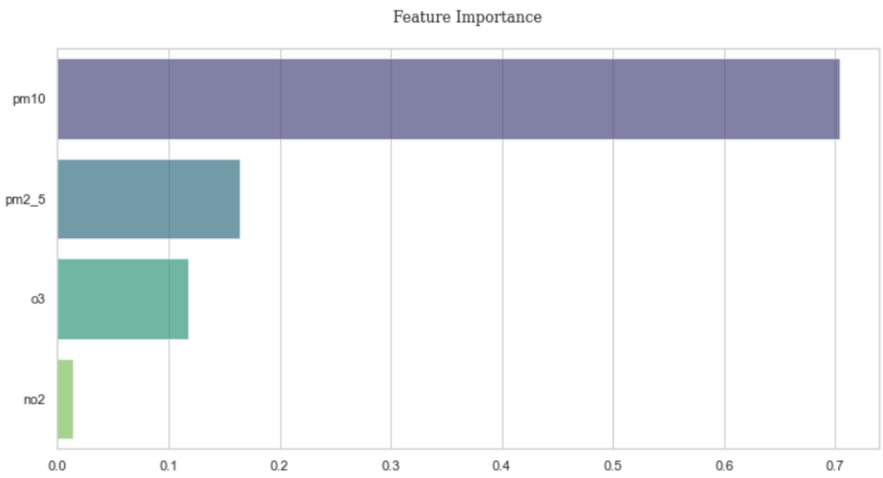


Figure 8. Feature importance of AQI forecasting

There also some ways to interpret prediction mechanism for random forest model. Figure 8 provides a different perspective for the feature importance. It does not only indicates that PM10 is the most important feature but also depicts that increasing value of PM10 causes higher air quality index. This means, the bigger PM10 values causes more air pollution.

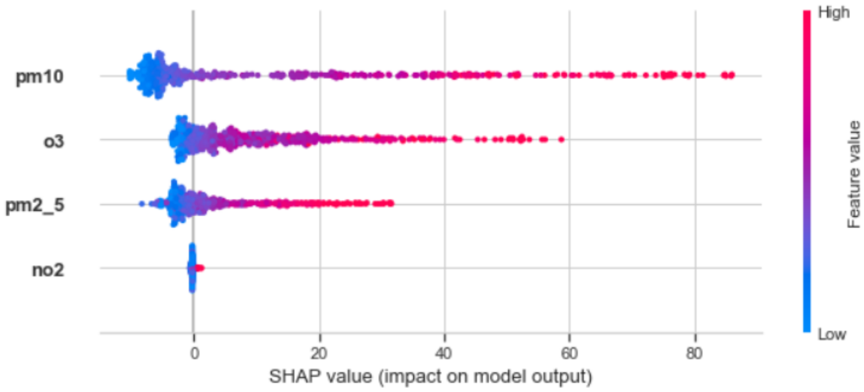


Figure 9. SHAP value of the features

Another way to describe how predictions are occured, Shapley Additive Explanation (SHAP) values can be observed. Figure 9 illustrates that how one observation is handled and prediction is conducted. For the specific example, NO_2, O_3, PM_{10}, and $PM_{2.5}$ are measured 0.66, 79.86, 118.05, 44.71 respectively. The base value of AQI is learnt from dataset is 39.34. The PM_{10} and $PM_{2.5}$ has causes increasing predicted AQI value. This gives an idea why air is predicted as polluted, because PM_{10} and $PM2_{.5}$ has bigger values. Final AQI value is predicted as 124.18 which belongs to “Very High Pollution”.

2.3.1.2. USE CASE BENCHMARKING

Intelligent building system provided by **IoT.nxt** provides monitoring of data such as heating, electricity, temperature control, power demand through a single application for smart buildings. However, there is no artificial intelligence capability to provide a decision support system in this system (https://www.iotnxt.com/smart-buildings/). **Libelium** company proposing solutions for cities parking and waste management in concept of digital ecosystem, company proposes air quality station with enabling analysis on machine learning algorithms. However, company is still lacking of analysing water sensing and gas analysis (https://www.libelium.com/iot-products/air-quality-station/). **BuildingOS** is a cloud-based facilities management system suitable for different types of businesses. BuildingOS offers tools to help users manage building maintenance, assets and environmental data spaces. BuildingOS software provides interactive interfaces (https://www.softwareadvice.com/cafm/buildingos-profile/). Also, **NEXOG CityWatch** is a system-of-systems monitoring platform that provides 360-degree monitoring and control of all city infrastructure provides rule-based, Action-on-Event (AoE) engine. The engine raises an alert when a specific event occurs or a performance threshold is crossed (https://nexog.com/cloud-services/learn-more/citywatch). However, the system is missing AI capabilities.

2.3.2. ENERGY MONITORING

2.3.2.1. RESULTS

Home Energy Management System Monitoring has before optimization and after optimization graphs. These graphs contain; energy consumption data of user sensors, total energy consumption and electricity cost. Time data are given in 5 minutes within a day. The amount of energy consumed by the devices is given in Kw. In the DataSet, the electricity cost values are determined as three tariffs within a day and the unit is Euro.

To see graphics; going to AI/Energy Monitoring from the general menu on the left hand side.

After selected the organization name, domain name and sub-domain name and click the "APPLY" button in Energy Monitoring.

Here, the changes in the before energy optimization and after the optimization of the energy are observed.

Graphic selection should be performed to display the results. There are three graphic types. These; Area Graph, Table (Raw Data) and Chart.

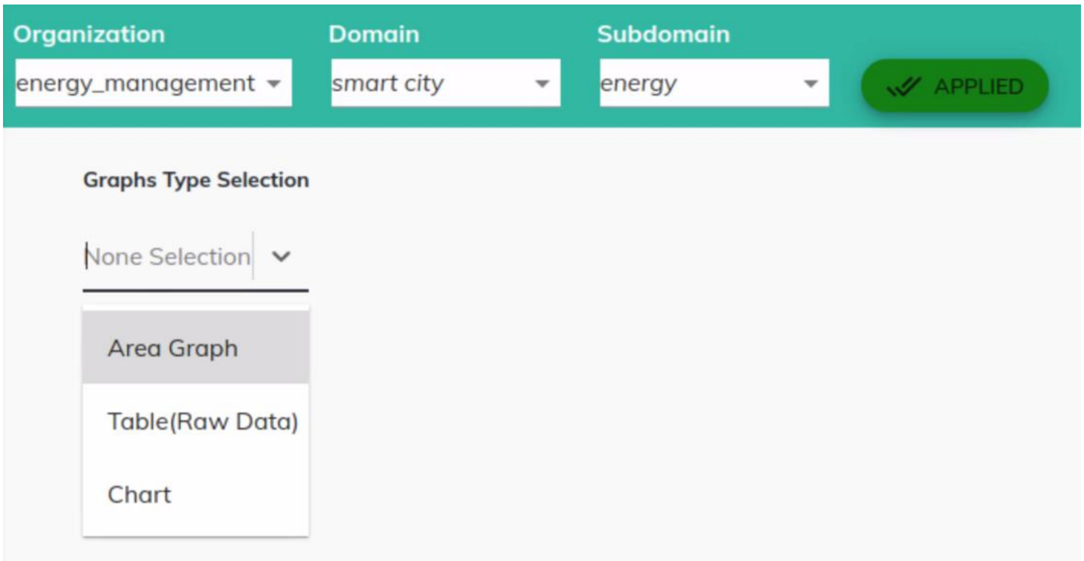


Figure 10. Graphs type selection

In order to see the graphics, sensors selection is needed.

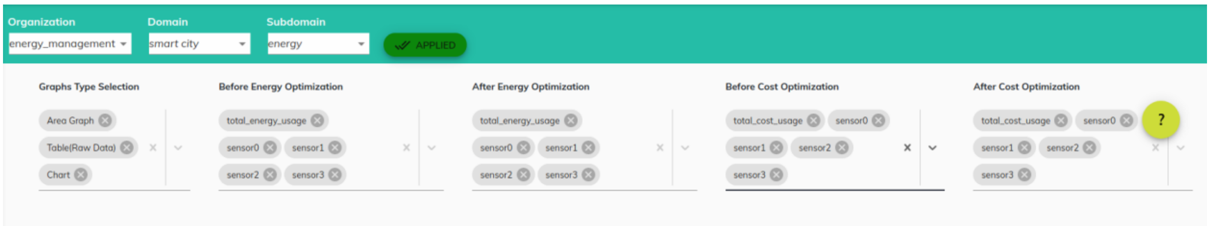
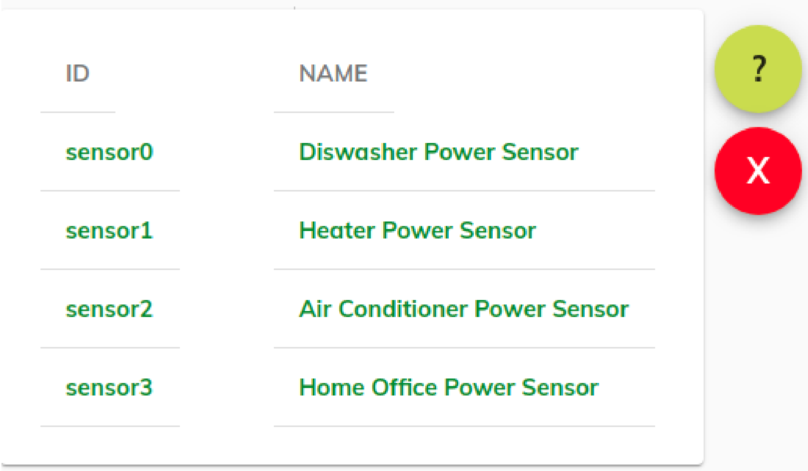


Figure 11. Sensor selection

When the "?" button is clicked, information about the sensors is displayed.

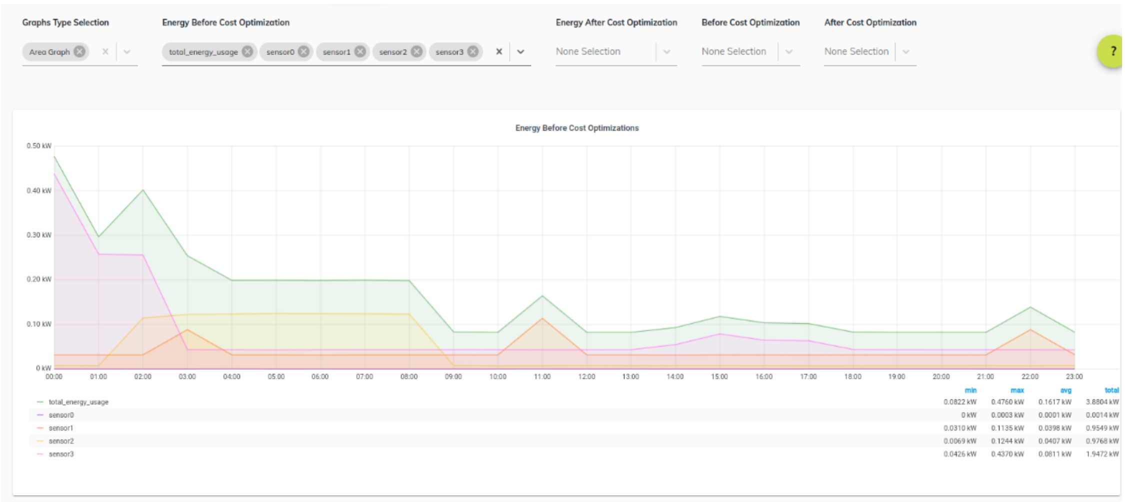


ID	NAME
sensor0	Diswasher Power Sensor
sensor1	Heater Power Sensor
sensor2	Air Conditioner Power Sensor
sensor3	Home Office Power Sensor

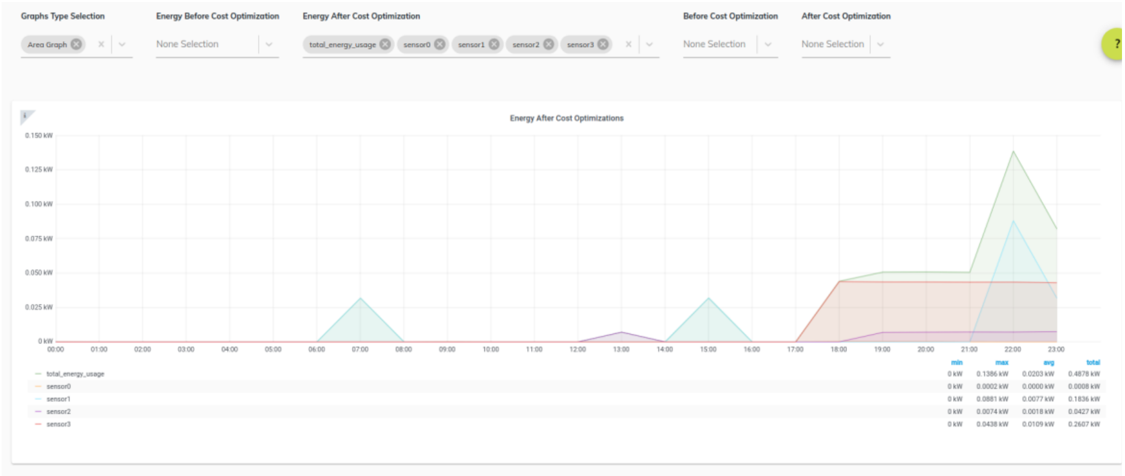
Figure 12. Information about the sensors

These graphs, the total energy consumed by the devices and the energy consumed separately for each device. The time data of this drawing are samples taken at intervals of 5 minutes in a day.

When Area Graph and before and after Energy Optimization are selected:



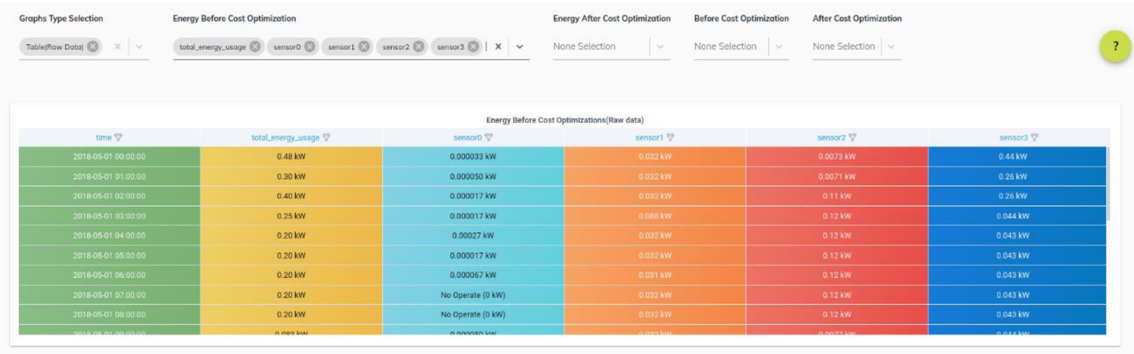
(a)



(b)

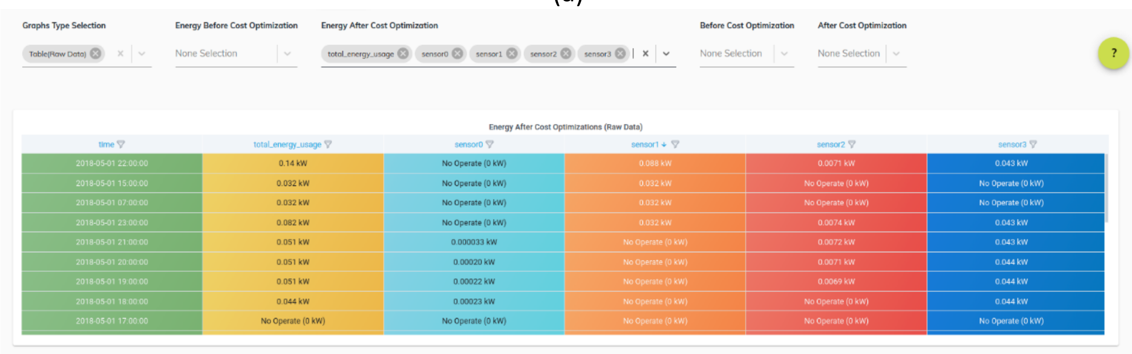
Figure 13. Energy before (a) and after (b) cost optimization – area graph

When table (raw data) and energy before and after cost optimization are selected:



time	total_energy_usage	sensor0	sensor1	sensor2	sensor3
2018-05-01 00:00:00	0.48 kW	0.000033 kW	0.032 kW	0.0073 kW	0.44 kW
2018-05-01 01:00:00	0.30 kW	0.000050 kW	0.032 kW	0.0071 kW	0.26 kW
2018-05-01 02:00:00	0.40 kW	0.000017 kW	0.032 kW	0.11 kW	0.26 kW
2018-05-01 03:00:00	0.25 kW	0.000017 kW	0.088 kW	0.12 kW	0.044 kW
2018-05-01 04:00:00	0.20 kW	0.00027 kW	0.032 kW	0.12 kW	0.043 kW
2018-05-01 05:00:00	0.20 kW	0.000017 kW	0.032 kW	0.12 kW	0.043 kW
2018-05-01 06:00:00	0.20 kW	0.000067 kW	0.031 kW	0.12 kW	0.043 kW
2018-05-01 07:00:00	0.20 kW	No Operate (0 kW)	0.032 kW	0.12 kW	0.043 kW
2018-05-01 08:00:00	0.20 kW	No Operate (0 kW)	0.032 kW	0.12 kW	0.043 kW

(a)

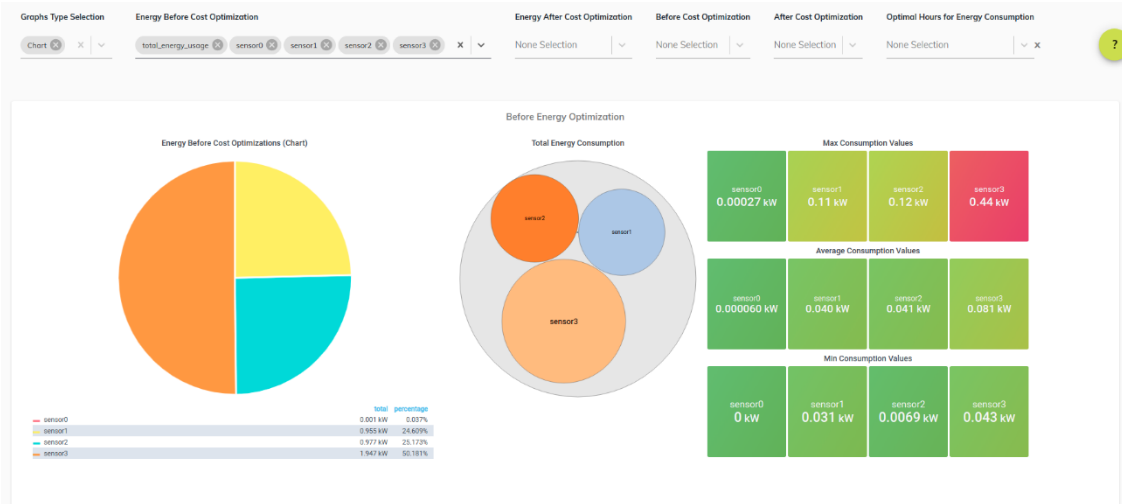


time	total_energy_usage	sensor0	sensor1	sensor2	sensor3
2018-05-01 22:00:00	0.14 kW	No Operate (0 kW)	0.088 kW	0.0071 kW	0.043 kW
2018-05-01 15:00:00	0.032 kW	No Operate (0 kW)	0.032 kW	No Operate (0 kW)	No Operate (0 kW)
2018-05-01 07:00:00	0.032 kW	No Operate (0 kW)	0.032 kW	No Operate (0 kW)	No Operate (0 kW)
2018-05-01 23:00:00	0.082 kW	No Operate (0 kW)	0.032 kW	0.0074 kW	0.043 kW
2018-05-01 21:00:00	0.051 kW	0.000033 kW	No Operate (0 kW)	0.0072 kW	0.043 kW
2018-05-01 20:00:00	0.051 kW	0.00020 kW	No Operate (0 kW)	0.0071 kW	0.044 kW
2018-05-01 19:00:00	0.051 kW	0.00022 kW	No Operate (0 kW)	0.0069 kW	0.044 kW
2018-05-01 18:00:00	0.044 kW	0.00023 kW	No Operate (0 kW)	No Operate (0 kW)	0.044 kW
2018-05-01 17:00:00	No Operate (0 kW)	No Operate (0 kW)	No Operate (0 kW)	No Operate (0 kW)	No Operate (0 kW)

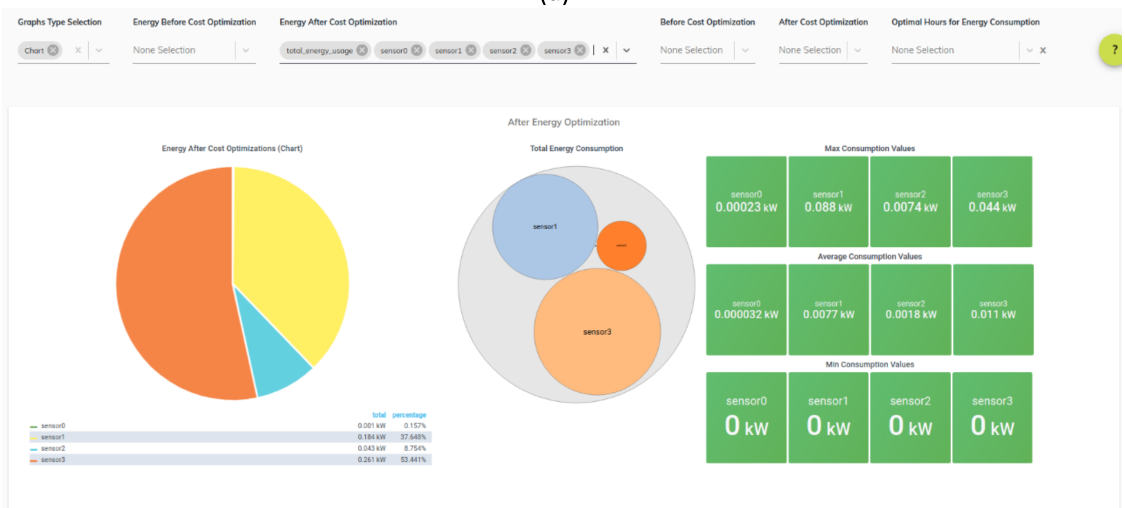
(b)

Figure 14. Energy before (a) and after (b) cost optimization - raw data

When chart and energy before and after cost optimization are selected; this graph type contains pie chart, bubble chart and sensors's maximum, average and minimum consumption values.



(a)



(b)

Figure 15. Energy before (a) and after (b) cost optimization - chart

Interval suggestion system is another part of the proposed method. The DQN time interval suggestion system aims to provide a user-friendly suggestion table that is the interpreted form of the optimization plots. Optimization plots are the visual output of the DQN algorithm and display the insight of the produced results. However, sometimes users may not have enough background to interpret the graphs. In order to solve this problem, we provide a solution that interprets the plots at the background and informs the users what are the optimal ranges for usage of the devices. As an example of the recommended optimal hours of the devices presented in Figure 16.

When Chart and Optimal Hours for Energy Consumption are selected:

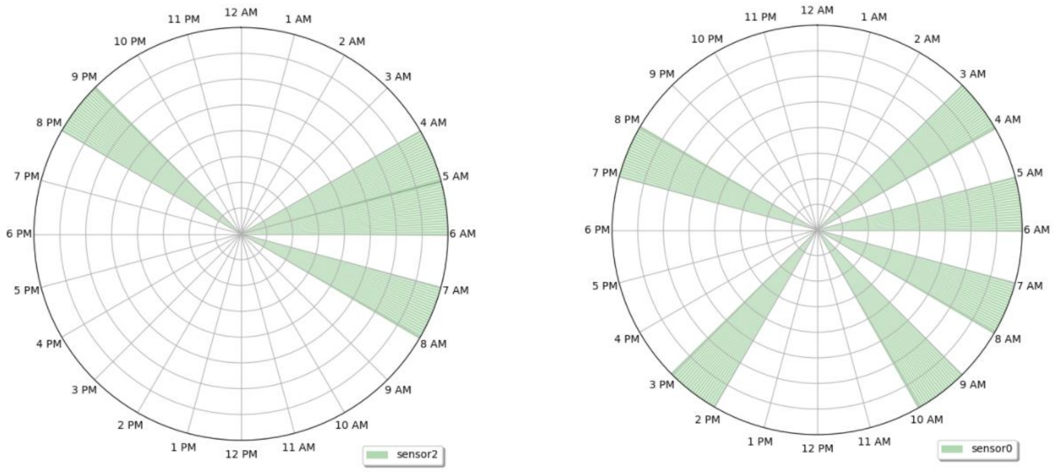


Figure 16. Recommended time intervals

2.3.3. AIR QUALITY MONITORING IN GALATI INDUSTRIAL AREA

The use case was implemented in the industrial area of Galati, where the air quality was monitored using IoT tools, to support local public administration authorities in developing, improving and implementing public policies.

The purpose of the use case was to install stations that monitor the air quality regarding particulate matter concentrations (PM_1, $PM_{2.5}$, PM_{10}), and meteorological parameters from the metallurgical industry area of Galati, which is an important Port city on the Danube, Romania.

2.3.3.1. RESULTS

Our final result of the use case was consisted of polices for industrial air emissions and mobility based on air quality data.

Figure 17presents the air quality monitoring station together with the PMs sensors used, installed in the use case area of Galati.

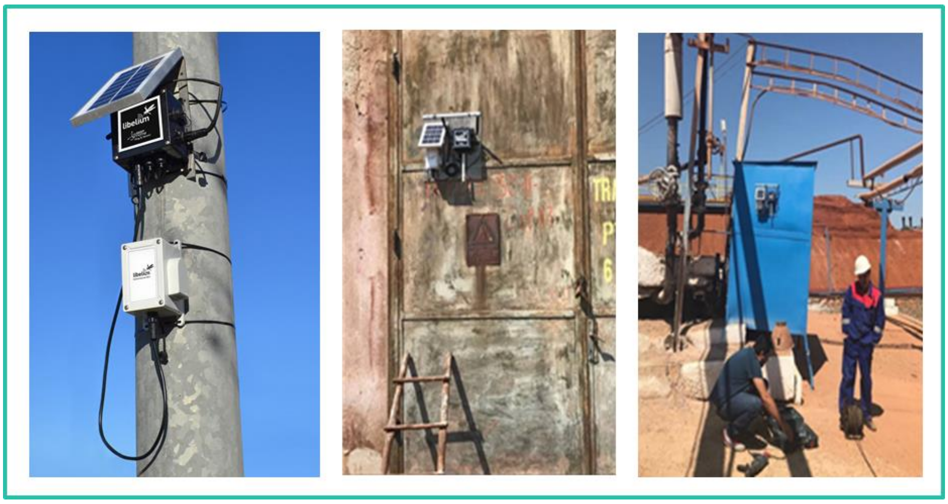


Figure 17. Air quality monitoring station

2.3.3.2. USE CASE BENCHMARKING

The main competitors at national level are:

- URADMonitor, developed in Romania, is an air quality monitoring system providing data regarding various pollutants (NO2, SO2, CO, O3, PM1, PM2.5, PM10) and meteorological data.
- The Airly platform, also present in Romania, is an air quality monitoring system developed in Romania providing real-time data regarding PM1, PM2.5, PM10 and meteorological data, measured with static IoT devices.

Strengths:

- BEIA's strong market position within the smart air quality, management market verticals
- Professional high accuracy sensors for a wide field of applications
- Partnerships with important market players in the smart environmental monitoring market (e.g. Libelium)

Weaknesses:

- Price of the hardware solution, which may be considered high for the Balkan region (professional high-end solution), when compared to sample analysis in laboratory for air quality.

3. CONCLUSION

After this analysis, it can be seen that the use cases developed in POLDER present both advantages and disadvantages compared to its competitors: on the one hand, there are use cases for which the competitor is very scarce due to the novelty of the application of the technologies in these areas, such as the image recognition for smart cities or the urban population prediction; on the other hand, for some use cases there is wide competition, although there are differences between the services they offer, such as the completeness of the information offered (tourist flow estimation use case) or the technologies used (air quality modelling and monitoring).

However, what it is clear is that this project has tried to defend the applicability of technologies that already exist in the market in areas that have not been explored in depth, with and architecture that allows their use together with other types of tools with different functionalities, something that, altogether, makes POLDER results easily adaptable to a wide variety of necessities and situations.

4. REFERENCES

[1] K. Zhang, H. Qiu, J. Wang, C. Li, J. Zhang y D. D. Chen, «Tourist gaze through computer vision: where, what, how and why?,» *Tourism Review,* 2021.