Report on the CDW flow optimisation tool development

IDENER

Version	1.0
WP	2
Date	30 April 2021
Dissemination level	Public
Deliverable lead	IDENER
Authors	Santiago Rodríguez, IDENER Santiago Moreira, IDENER Alex Hall, IDENER
Reviewers	María Lopez, IDENER
Abstract	This report details how IDENER has developed the CDW flow optimisation tool
Keywords	CDW, IT software tool, Seville
License	

Contents

1. Executive Summary

The CDW flow optimisation tool has been developed using a web framework based on Python, in this case, Django. The framework will be supported by specific modules and tools developed in Python for data analysis, machine learning and geographical information systems. Those tools include basic libraries for numerical tools and visualization such as Numpy and Matplotlib and more specific libraries for data analysis and ML such as Scipy, Scikit-Learn or Keras and advanced visualization such as Bokeh or Plotly. The initial deployment of the tool will be done using either Virtual Machines (e.g., with VirtualBox) or Docker containers.

Scrum methodology has been strictly followed for the development and the tasks and subtasks necessary for developing each feature, and sections of the front-end and back-end have been clearly defined, including expected results and goals. The philosophy behind the development has been to approach each feature individually, then create the first version of one of the web apps for citizens and managers (i.e., Lipasam workers) and, then deploy them in a virtual machine.

Thanks to this approach for the development, all the members involved in the development of the CityLoops platform have been able to contribute directly to the programming activities. Those more experienced in mathematical modelling, scientific programming and data analysis easily learned a web framework such as Django and those experience in other web frameworks and/or Python quickly acquire the knowledge for developing the CityLoops platform in Django. In addition, other partners developing software tools for the CityLoops project have already used Django and provided support and tools which can be adapted to the needs of the CityLoops platform in Seville.

2. Introduction

Machine learning and Data-Driven Applications have been gaining popularity in recent years. However, the final deployment requires a combination of different mathematical, programming and web development skills (or creation of graphical applications using frameworks such as Qt5 or GTK+).

Django is a very popular framework for web development with a very smooth learning curve, allowing the full development of web applications from initial concepts to final user products very quickly taking into consideration the most important aspects of security. These advantages overcome some limitations of other frameworks such as verbosity and complexity or steeper learning curves. Besides those limitations related to the development of the software platform, the use of Django as a framework also helps to improve the performance of the applications developed with other popular web frameworks. Dynamic applications did not always perform well, and complex single-page applications may have a significant lag, making them inconvenient to use due to their size.

Even if Django is an easy-to-learn framework for web development, it has been used in developing professional sites and platforms such as Instagram, Spotify, YouTube, BitBucket, DropBox, The Washington Post, Mozilla, Pinterest and even websites of NASA or National Geographic. Another additional advantage consists in using a single web framework other Python module without using communication protocols or separated containers with the back-end operations.

For the back-end part of the application, the data analysis and machine learning have also been powered by Python libraries. The communication between data sources has also been managed with Python when needed.

The initial developments have been provided to the managers for testing in containers or virtual machines to facilitate direct deployment in any computer. Virtual Machines allow preparing a complete operating system with the pre-installed CityLoops platform in a single file which can be directly executed by the user using Virtual Box, KVM, VMware or Hyper-V. Similarly, Docker containers allow running virtual servers to deploy web applications very quickly, even not booting an OS with the main advantage of allowing continuous integration.

A Gitlab repository has been created with a clean Django project to serve as a template to implement the initial features of the CityLoops platform. Once the main functionalities and widgets have been developed and tested, they have been all integrated on a single common application panel (e.g., like a dashboard), and interactions between them have been implemented and tested. This corresponds to Stage 1 of the development. Stage 2 has been the full development of one of the CityLoops platforms (i.e., CDW flow optimisation tool) without taking into consideration aesthetics or final product aspects, only a basic front-end and main functionality. Stage 3, has been reproducing this development in all the remaining applications, obtaining as a result an initial version fully functional of all the CityLoops applications. Finally, Stage 4, will consist of adding the aesthetics and improving UX and GUI, together with the deployment of citizens' web pages for visualization and services.

Now that the technologies for development have been presented. Let us briefly discuss the main goals and functionalities of each platform.

The CityLoops platform for Seville's construction and demolition waste (CDW) will have two main goals: optimization of the fluxes of CDW from citizens and small businesses; and the selection of optimal locations of new "clean points", i.e., intermediate management facilities for construction or other specific waste. The software tool for managers will combine data regarding the economic and geographic growth of the city with data about waste generation and, more specifically, construction and demolition waste, CDW, to support managers on the decision of new locations for future clean points. In the case of the version for citizens, regarding CDW, provide the user with the location of the optimal clean point to fulfil the needs of the citizen to dump construction and demolition waste from their homes and from small business.

> ### CONTEXT
>
> The city of Seville, located in the south of Spain, has 700,000 inhabitants and a surface area of 140.42 km2. It is the most populated city in Andalusia, the fourth in Spain. The construction industry is one of the main economic activities that boosts the economic development of the city.
>
> Construction and Demolition Waste (CDW) represents up to 45% of the total amount of waste generated in the city, i.e., 270,547 tonnes of CDW and 1,309,501 tonnes of excavated soil annually. Currently, only 16.1% of CDW is recycled, mostly for buildings works and road fillings.
>
> The demonstration actions of the CityLoops project represent another step towards advancing Seville's path towards a more circular city, which is aligned with the declaration that the city itself led in 2017, together with more than 200 municipalities in Spain the which underlines the importance of Local Governments to put into practice the commitment, the need to implement the Circular Economy.
>
> Seville's declaration for the Circular Economy is also aligned with its City Model of the Seville 2030 Strategic Plan.
>
> On the other hand, the city of Seville understands its commitment to the 2030 Agenda for sustainable development of the United Nations. It assumes it as the standard of its strategic and sectoral planning. Likewise, the Seville 2030 Strategic Plan is aligned with the Sustainable Development Goals (SDG).
>
> Other local and sectoral initiatives and plans converge in the Seville 2030 Strategic Plan, such as the Local Waste Management Program, currently the draft, which will incorporate, if the results are satisfactory, the actions piloted within the framework the project CityLoops.

The CWD flow optimisation tool is expected to contribute to the following outcomes:

- Advance in the fulfilment of the European, national and regional objectives, in selective collection, recycling and not disposal in a landfill, marked by Directive 851/2018 and 850/2018.
- Strengthen the education, awareness and knowledge of citizens and other socio-economic agents related to the bases of the circular economy in relation to the improvement of CDW management.
- Reduced illegal dumping and abandonment of CDW and increased reused and recycled CDW due to the awareness campaign and improved management and clean points.

IDENER has developed the IT software tool in collaboration with the Seville´s cluster. Seville's cluster of partners for the CityLoops project includes the Municipality of Seville, LIPASAM (Municipal Solid Waste Management company), EMASESA (Municipal Wastewater Treatment Management company) and IDENER (Private Research company). Together these partners are committed to the CityLoops' approach to close the loops of waste material in the city, promoting a circular economy approach to the city's development.

3. Definition of functionalities

3.1 Functionalities of CDW platform for citizens

- Forms for query: The citizen will provide data about its location and the type of waste placed in the "clean point". Using this data and the status of the clean points, the CDW CityLoops platform for citizens will suggest the most suitable "clean point".

- Map showing clean points and highlighting the most suitable one: A widget with a map (using OpenStreet Maps) will be presented to the citizens showing the current location of the "clean points" and after the results of the query, the most suitable one will be highlighted. At the same time, a route from the district origin to the proposed clean point will be suggested to the citizen-based on OpenStreetMap.

- Dashboard showing the current status of the "clean points": Clean points have different containers for different types of waste (e.g., bricks, wood, steel…). A dashboard will show the fill levels of the containers and the timestamp of the presented data.

- Control panel: A set of buttons will be included in the application to clean the forms, do the query, and export information. Logos of the involved partners (Emasesa, Lipasam, Municipality of Seville, Cityloops project) will be included providing additional information.

3.2 Functionalities of CDW platform for managers.

- District representation of Seville Data: The city of Seville will be divided into a discrete representation of districts and overlapped over a real map of the city.

- Data request of each district: Each district will be represented by different social and economic aspects, and the waste generation will be described. The remaining areas must be classified as "urban soil", "urban land scheduled for development", "rural areas" or "natural ecosystem".

- Representation of the growth of the city: The growth of the city will be represented in a time-lapse showing in the last frame the projection of the next 3, 5, 10 years according to the data provided by the manager through the platform.

- Location of potential future "clean points": Using the data provided by the managers and machine learning techniques, the platform will present the areas where "clean points" could be placed in the next few years, according to the previously decided criteria.

4. Stakeholders' identification. Public and private.

The platform for each use case within Seville will have two versions, one for managers and authorities and another for citizens.

The platform version for citizens has only visualization capabilities and widgets using data provided by managers from Lipasam or results from the corresponding CityLoops platform for partners.

Therefore, the applications for managers will consist, at least, in a data collection tool where managers will provide the data to power up the data-driven applications or the visualization widgets in the platforms for citizens. Additionally, the platforms for managers may have additional capabilities for decision support. The applications for citizens show interactive visualizations of the data provided by managers and will use the available widgets to get information about specific questions regarding CDW.

5. Functional Requirements.

This section lists specific functions and object-oriented designs linked to functionalities of the Cityloops platform.

- ORM/SQL database management (included already using Django)
- Machine Learning functionalities: regression, clustering.
- GIS / OpenStreetMap data management.
- Web Apps, Widgets and data visualization.
- Data collection and data analysis.
- PostgreSQL 12.6
- Django 3.2
- OSGeo4W
- Leaflet
- Gdal: (required modules installation: wheel; pipwin; numpy; pandas; shapely; gdal; fiona; pyproj; six; rtree; geopandas)

6. Design Requirements

6.1. List of Components of WebApp for Construction and Demolition Waste (CDW)

Table 1 – Web Form for Query and Data Collection

CDW_APP001	Web Form for Query and Data Collection
Platform version	Basic User or Citizen
Description	This component of the Cityloops platform for CDW is a web form where the citizen or basic user will introduce the parameters of a query to find the closest "clean point" to dump construction waste (debris, metal, wood, or other). Some fields of these data include postal code, district, type of waste, estimated amount. Once all the fields of the form are completed and the button "Search" is pressed, the query is sent to the back-end function, which will provide as answer the "closest suitable clean point or clean points according to the location of the CDW and the type and amount of waste".
Implementation	Django Form object has been implemented and formatted in the front-end of the web-app. A search and decision-making algorithm have been implemented to 1) select the suitable "clean points" according to the types of containers; 2) suggestion of the "clean point" to the user based on how full the existing containers and the distance from the district to the clean point are. The distance is calculated using the Haversine formula.
Input Data and Widgets	Text Edit (Postal Code), Spinbox (Amount of waste in kg), Comboboxes (District, Type of Waste), Date (expected date of disposal), Push Button (Search, create the query)
Output / Results	Once the button is clicked, the closest Clean Point will be highlighted in a map, and the information of the containers and a route will be suggested to the user (citizen) based on OpenStreetMap tools.

Table 2 – OpenStreet Map showing districts highlighted and Clean Points location

CDW_APP002	**OpenStreet Map showing districts highlighted and Clean Points location**
Platform version	Basic User or Citizen / Manager or Admin
Description	Widget imported from OpenStreetMaps showing a map of Seville City, including some additional layers, such as polygons representing the districts of Seville in terms of administrative divisions and waste generation. Locations of the current Sevilla districts and the current "clean points" will be marked in the initial map. Once the query is done and the result of checking the "optimal clean point for your needs" is obtained, this clean point will be highlighted in the map as well as the district of the origin of CDW, and, finally, a route by car will be suggested using OpenStreetMaps.
Implementation	The searching and decision-making function is implemented in Python in the back-end, including the Haversine formula. The results of the most suitable clean point, the distance and the coordinates of district and clean point are passed to the frontend. There, a geodjango application using OpenStreetMaps, calculates and presents the suggested route to the clean point from the corresponding district.
Input Data and Widgets	Map imported from OpenStreetMap library. Additional layers added, including district representation and clean points position. The route is highlighted on the map.
Output / Results	Points, lines and polygons highlighted on the map.

Table 3 – Visualisation of Containers Status in the Clean Points

CDW_APP003	Visualisation of Containers Status in the Clean Points
Platform version	Basic User or Citizen / Manager or Admin
Description	This part of the platform for CDW will present the status and the expected evolution (according to previous data) of the level in each of the containers for the different CDW materials that a citizen can dump in the clean point. Once the clean point is selected in the menu, the corresponding location will be highlighted in the map of CDW_APP002, and the status of each container will be presented in the panel of containers of the selected clean point.
Implementation	Javascript function combined with CSS to create a dashboard for each clean point. This dashboard will show the level of each container.
Input Data and Widgets	ComboBox containing the list of clean points to select. Panel with several plots showing the status of the containers and the projection based on past data.
Output / Results	Bar Plots with the evolution of the containers and line plots with the projected evolution.

CDW_APP004	Updating data of the containers of the clean points
Platform version	Manager or Admin
Description	The responsible for the management of this app will collect the information about the status of the containers in the clean point and will update the corresponding section in the CityLoops CDW webapp. This information will be used in the selection of the most suitable clean point. The information about the status of the containers in the "clean points" corresponding to the different types of construction and demolition waste (e.g., bricks, concrete, wood) will be part of the inputs for the decision-making process when recommending a "clean point" to the citizen.
Implementation	A Django model will be implemented, and a specific interface will be implemented in the front-end to update this information.
Input Data and Widgets	The input of this data may be using a table widget within the CDW platform.
Output / Results	Once the data is loaded (or updated) into the system, the visualisation panel CDW_APP3 of the containers will be refreshed, showing the updated data and a corrected projection for the next 2-3 days.

Table 5 – Loading data for new clean point locations.

CDW_APP005	Loading data of clean points
Platform version	Manager or Admin
Description	The data of the CDW collected in the different clean points will be loaded to the system in order to predict which clean point will receive more CDW in the near future. The CDW, with an increasing trend of CDW reception, may need an extension or a near additional clean point in the future. Select the clean points of Seville receiving the highest amount of CDW.
Input Data and Widgets	Data will be loaded in one or several CSV files to the platform. Widget: Load File Dialog Window for each one of the input files
Implementation	A Django model will be created to store the name and the CSV file of the dataset. In the front-end part, the manager will select the available datasets (previously loaded by the administrator) and will use a form to provide the query as the input to the ML prediction model. Then, the model, implemented in the back-end, will calculate the result (i.e., the prediction), and this information will be presented to the user, together with a visualization of the dataset. In this way, the manager will check if the prediction is consistent with the corresponding data and will increase the trust of the manager on the tool.
Output / Results	According to the results obtained by the back-end calculations, the clean points with the increasing demand of containers for CDW will be highlighted in the OpenStreet Map of the CDW_App002 Current and Expected tons of CDW will be presented on the map over the corresponding polygon of the district.

Table 6 – Manually loading data of CDW generation by district.

CDW_APP006	Loading data of CDW generation by district
Platform version	Manager or Admin
Description	CityLoops Platform for CDW will collect the data provided by managers or admins in CSV files. This data will include social, economic, geographic, and waste-related data of Sevilla districts. This data may include, construction related licenses, number of containers for residues, packages, glass and paper, and other data relevant for the characterisation of the Sevilla district.
Implementation	A Django model will be created to store the name and the CSV file of the dataset. In the front-end part, the manager will select the available datasets (previously loaded by the administrator) and will use a form to provide the query as the input to the ML prediction model. Then, the model, implemented in the back-end, will calculate the result (i.e., the prediction), and this information will be presented to the user, together with a visualization of the dataset. In this way, the manager will check if the prediction is consistent with the corresponding data and will increase the trust of the manager in the tool. In this case, the prediction will be related to the potential generation of CDW based on the expected number of licenses for construction activities directly related to citizens and small enterprises (e.g., new buildings are mostly managed by big construction companies and they will be not considered here).
Input Data and Widgets	Loading CSV or XLSX file or modifying the data directly in the CDW platform (this must be decided)
Output / Results	Updated data about CDW generated in each district of Seville city.

7. ML tools for CDW prediction

The tools for predicting and managing the amount of CDW in the near future using data-driven machine learning applications are divided into two approaches. The two approaches are due to the two different data sets available depending on the source of data. One dataset is provided by the clean points, where the total amount of several types of waste (including CDW such as debris, wood or metal) in each clean point is collected every year.

The other data set contains social-economic information of the Sevilla districts, as well as information of the number of containers for each waste and the number of constructing and renovation works of each district.

7.1. ML App based on clean points dataset.

This dataset will be used to predict the expected amount of CDW (i.e., debris, wood, and metal) using as input the total amount of waste received (in tonnes). A short-term extrapolation will be done using the CDW data of each clean point. It is important to mention that the available time series is quite limited. The data of 2020 and 2021 do not properly represent the reality of CDW generation in Sevilla due to the strong impact of the COVID-19 situation.

In Figure 1, we can see the amount of debris waste compared to the total amount of waste received in the clean points of Sevilla (Los Olivos, Los Naranjos, Las Jacarandas and Los Pinos) for the years 2015, 2016, 2017, and 2018 (increasing size of the point in the figure). We can observe in the figure when plotting all values together, that exists a potential correlation between the total waste received in the clean point and the debris. The same occurs when plotting wood (Figure 2) and metal (Figure 3) against the total waste collected in the clean points.
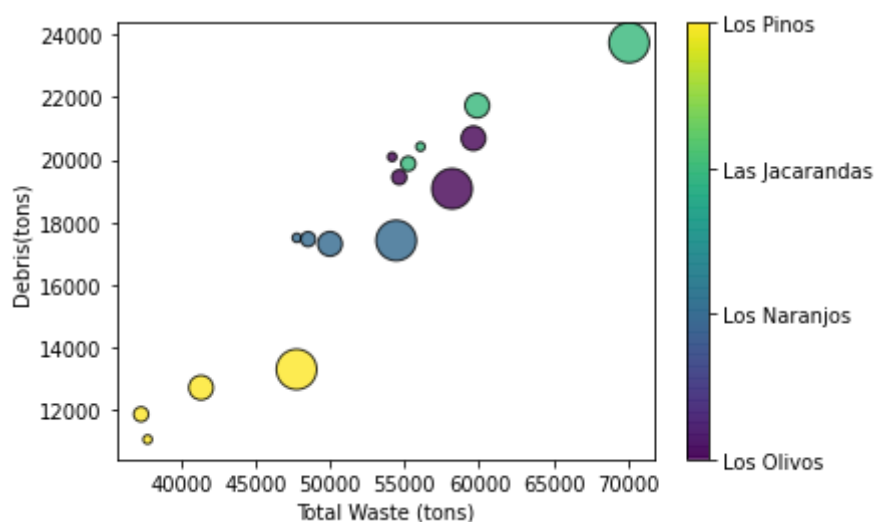
Figure 1. Debris vs Total Waste collected in the clean points of Seville. Each colour represents the data of a specific clean point and the increasing size of the marker corresponds to the years 2015 to 2018.
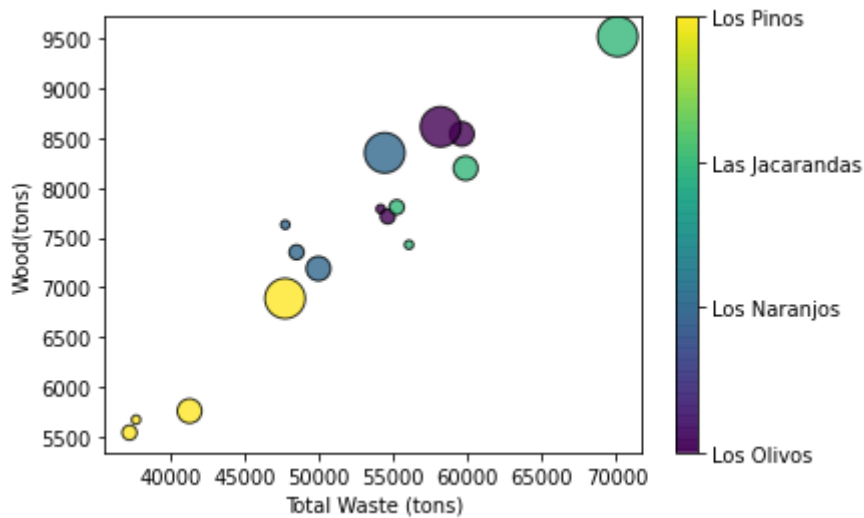
Figure 2. Wood vs Total Waste collected in the clean points of Seville. Each colour represents the data of a specific clean point and the increasing size of the marker corresponds to the years 2015 to 2018.
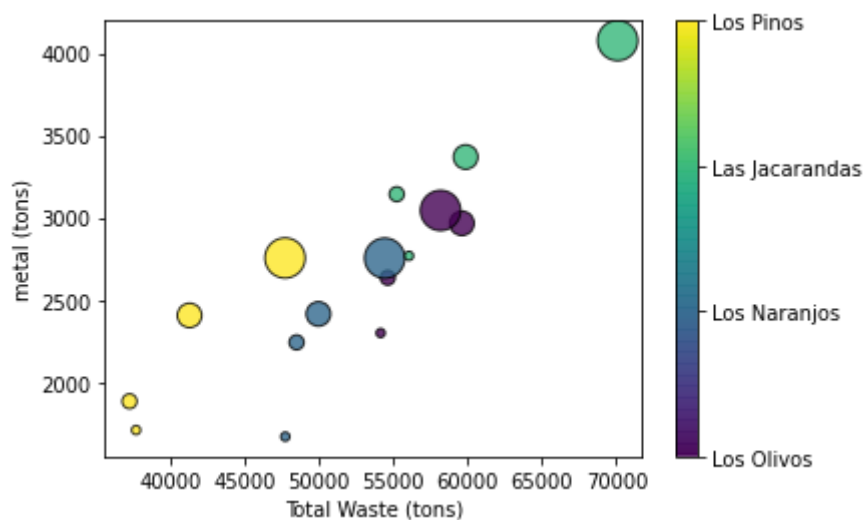
Figure 3. Metal vs Total Waste collected in the clean points of Seville. Each colour represents the data of a specific clean point and the increasing size of the marker corresponds to the years 2015 to 2018.

Something relevant that we can observe in the data of the metal vs total waste. Even if a certain signal can be inferred from data, this is noisier than in the case of wood and debris.

A linear regression model has been applied to the three data series (debris, wood, and metal), splitting first the data into a train set (80%) and test set (20%). Due to the limited amount of data (only 16 points) the model has been applied to predict the whole series of debris, wood, and metal to present the following figures. In the three cases, the applied model has been trained with the 80% of the data but then applied to predict the 100% (not only the 20%) of the data to construct the figures (otherwise, only 3 points would be presented). In the three cases, we can observe how the linear model captures a certain signal of correlation in the data.
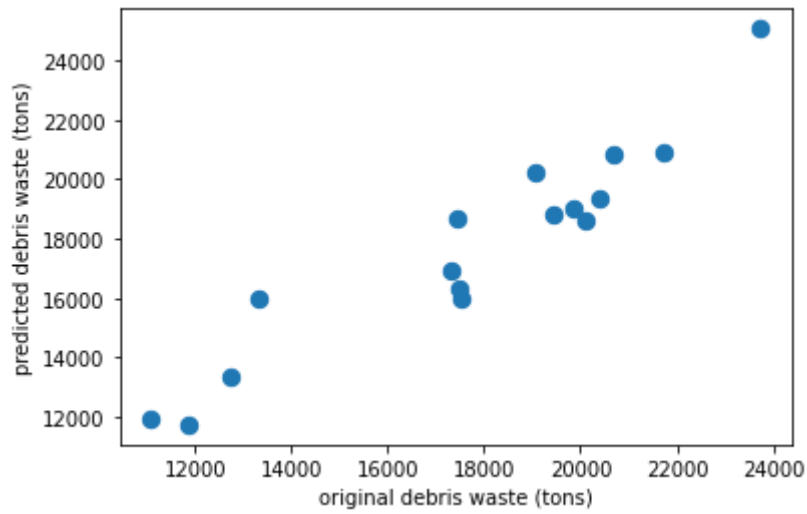
Figure 4. Predicted debris using total waste and linear regression vs original debris data.
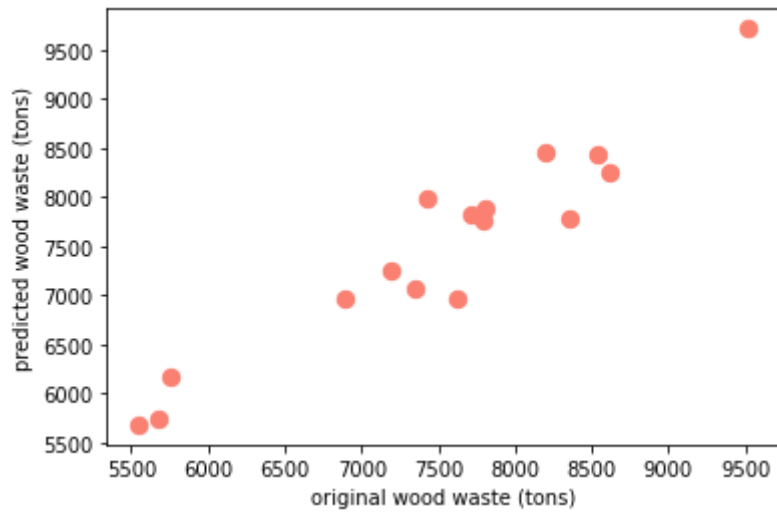
Figure 5. Predicted wood waste using total waste and linear regression vs original wood waste data.
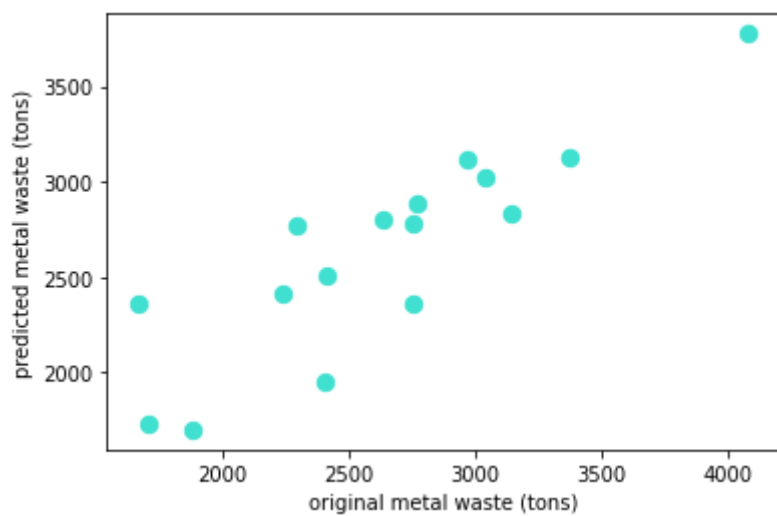
Figure 6. Predicted metal waste using total waste and linear regression vs original metal waste data.

CWD flow optimisation tool   18

The obtained linear regression models will be used to predict the quantity of debris, wood and metal using the expected values of total waste in the clean point. The models will be continuously updated when new data is incorporated. The total amount of waste expected will be estimated using a regression model and then, the total waste predicted will be used to predict specifically the total amount of debris, wood, and metal in each clean point.

As the final result of this ML tool, the manager would obtain information about which clean point is expected to receive a higher quantity of debris, wood and metal (according to the available data and the observed trends). Therefore, it will be able to decide if an extension of one of the clean points is needed or even the creation of a new clean point near to the one with the expected increasing pressure.

7.2. ML app based on Sevilla districts dataset.

The second ML tool which will be used in the prediction and management of CDW in Sevilla is the social-economic and waste data of the Sevilla districts (including population, number of wastes, plastic and glass containers, and construction and renovation licenses). The goal of this tool is to predict the number of restorations, alteration and demolition works based on the information of population and waste of each district, which indirectly represents and an indication of the level of consumption of the habitants of the district and this welfare can potentially be translated into construction and renovation works performed by the habitants themselves or small-medium size construction companies. We have selected those types of construction works because they are more related to citizens and small companies than the new construction licences (new buildings or houses).

The tool will use a Random Forest Regressor to estimate the number of licenses for restoration, alteration, and demolition works for the Sevilla districts based on its population, surface, the total number of waste containers, package containers, paper containers and glass containers. The approach relates the level consumption of Sevilla citizens with the number of containers that are necessary to collect the generated waste and the population and surface of each district with the number of licenses for the previously mentioned construction works.

The data has been divided into a train set (80% of the data) and a test set (20%) for training the Random Forest Regression model. However, the following figure presents the predictions for the whole
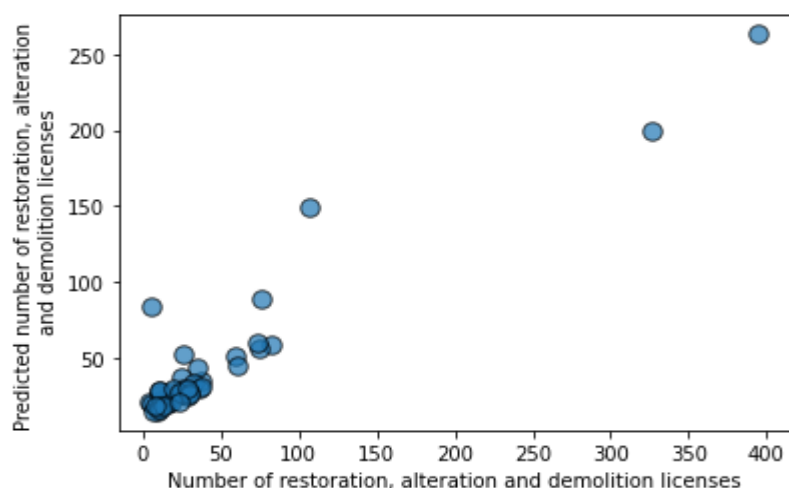

Figure 7. Random Forest Regression of social-economic and waste data to predict the number of licenses of restoration, alteration and demolition works.
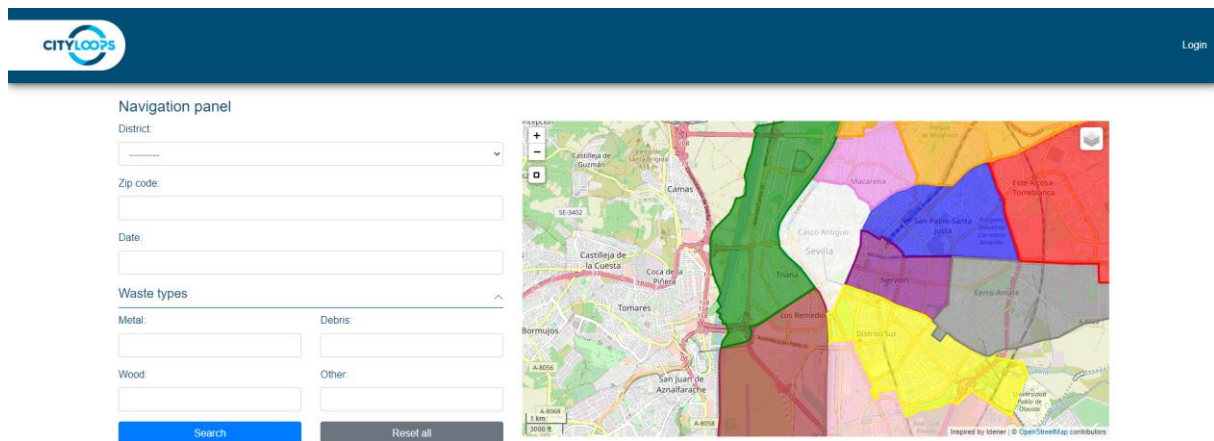
data set and its comparison with the real number of restorations, alteration, and demolition work licences due to the small size of the data set. Before the training, the data has been scaled using a Standard Scaler (X-mean/standard deviation).

8. User Workflow

8.1. Citizen platform

When the user enters the CDW platform, they will encounter the navigation panel with the text inputs for the user's district, zip code and the date the user wants to go to the clean point. Below the navigation panel, the user will see the inputs to enter the amount of waste of each type they will be placed in the clean point. Two buttons are found at the bottom of the form: one that submits the form (the one labelled "Search") and the other one to reset all the text boxes.

At the right of the form, a map will be shown with all the districts of the city highlighted with a different colour for each one of them.

Figure 8. CDW platform main page.

Once the user fills in every input and hits the "Search" button, a column with the results will appear between the original form and the map. In this section, the user can expand each collapsible to show the information of each clean point, including the distance between the selected district and the clean point and the container fill levels. The map will also show the distance between the district that the user selected and the best clean point for their needs.
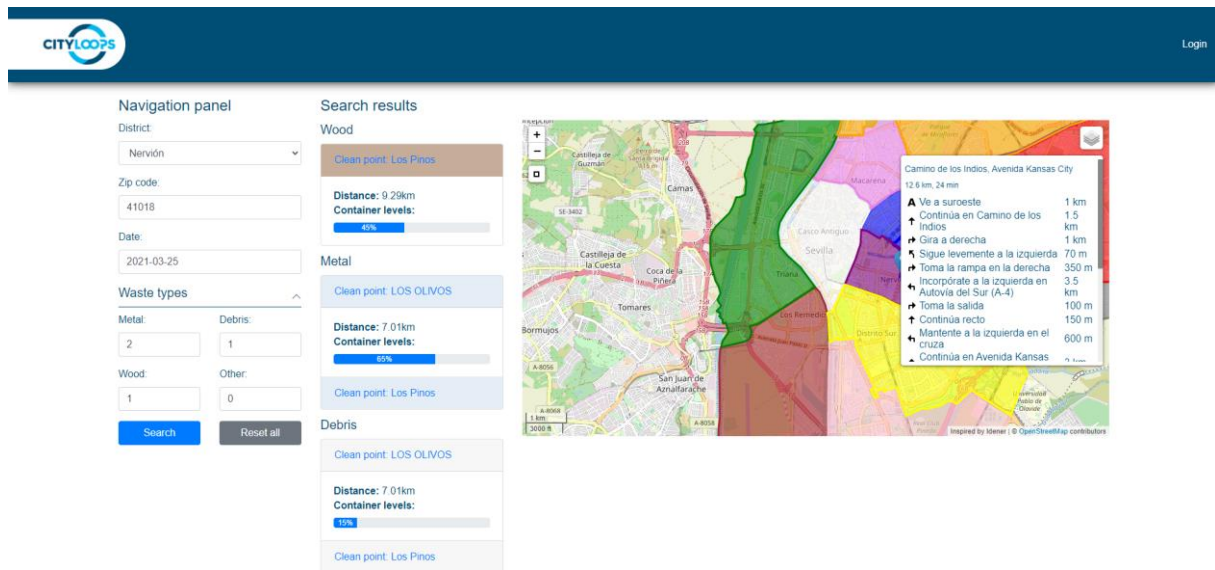
Figure 9. CDW platform main page showing results.

8.2. Manager platform

The manager platform has several pages to perform actions related to the management of the application and its resources.

To access the manager platform, the user must click the "Login" link at the top right corner of the page. This will lead them to the login page, in which they can enter their username and password to access the manager platform.
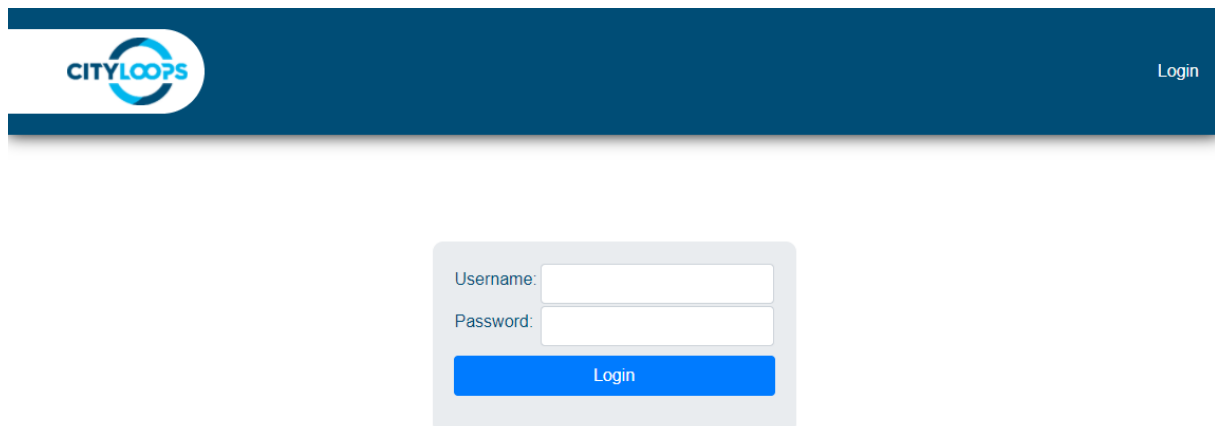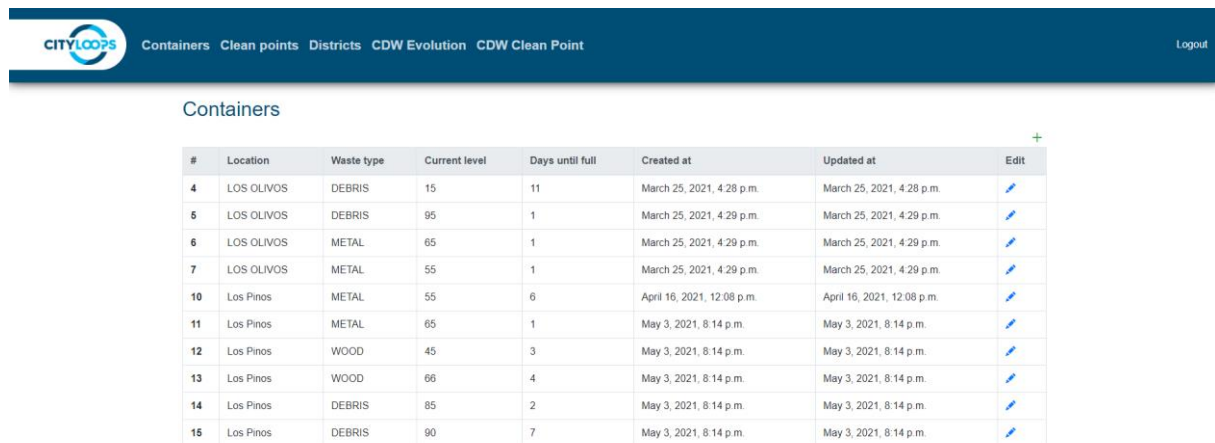

Figure 10. Login screen.

Once the user has logged in to the application, they will be able to access the pages to manage the entities of the application (containers, clean points, and districts). Each of the buttons located at the top bar of the page regarding these entities leads to the list of said entity. Through this list, the user can see the different attributes of each entity created. They can also create a new entity by clicking on the + sign located at the top right of the table, or edit any of the already created ones, by clicking the pencil icon at the end of each row.

Figure 11. Container list page.

When entering the page to create a new entity, the user will be supplied with two ways to create a new entity: to create it manually, entering the data in several inputs or uploading a file that contains the pertaining data. Note that the file uploaded must be in the correct format and contain the needed data to create a new entity.
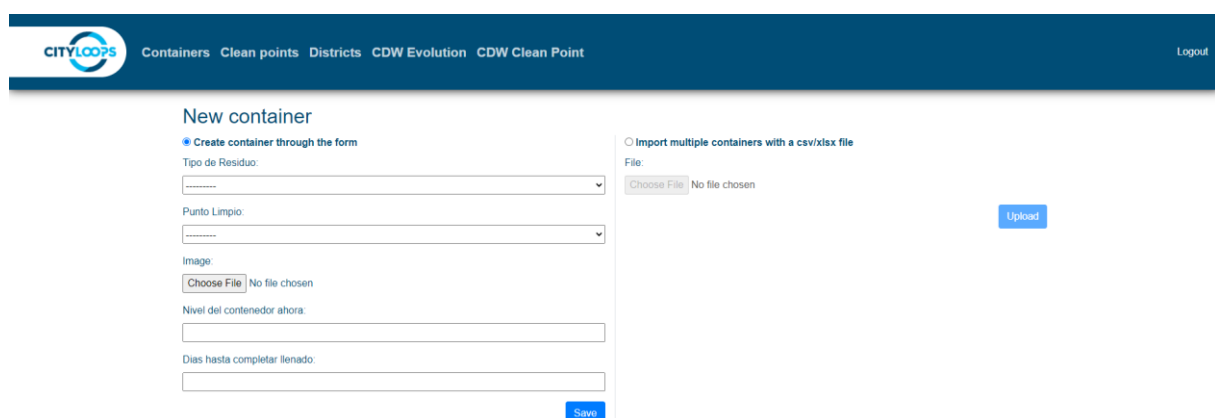

Figure 12. Container creation page. On the left, the form to create the container manually. On the right, the input to upload a file with the containers data to create.

Apart from managing the different entities of the application, the user can access the CDW evolution page and the Clean Point evolution page.
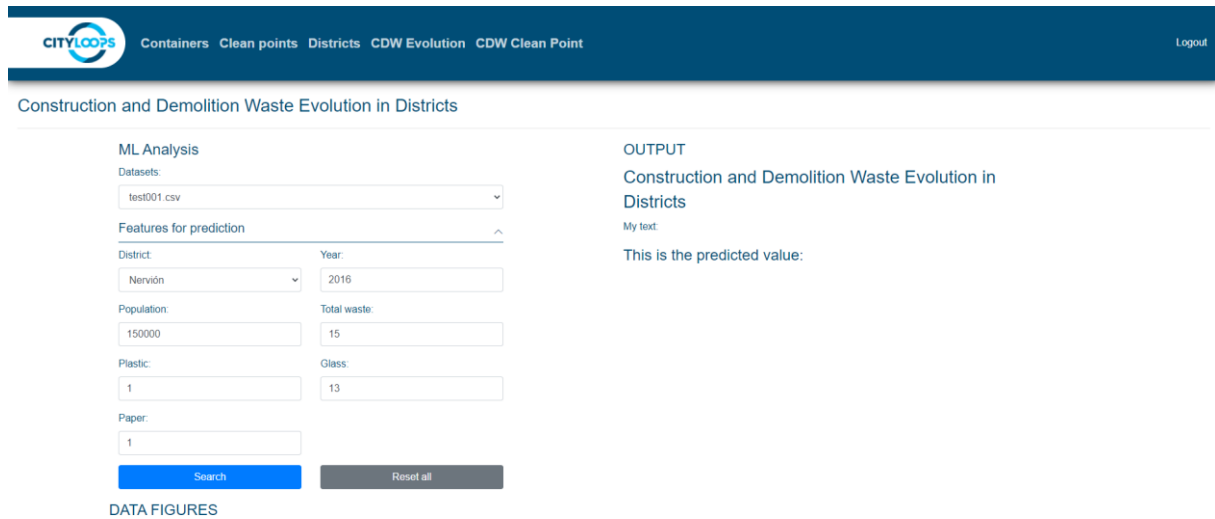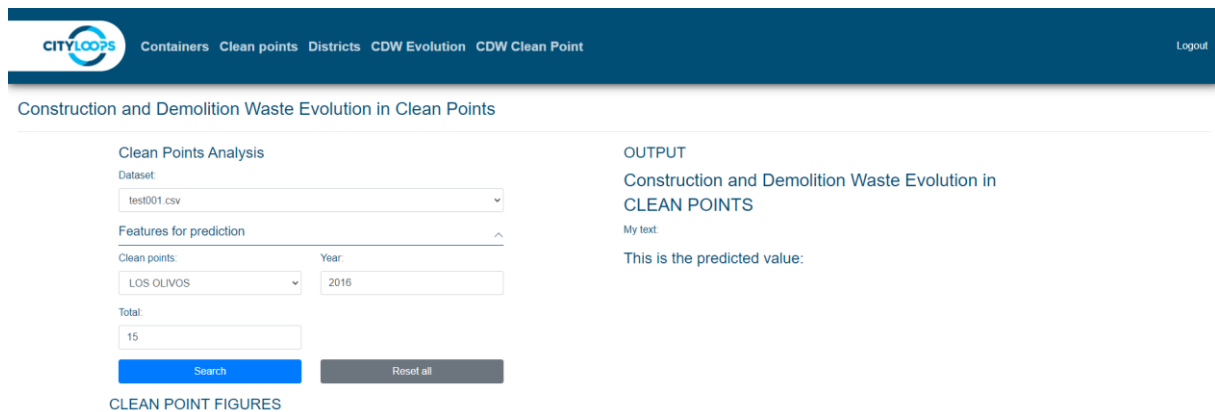

Figure 13. CDW evolution page.

Figure 14. Clean Point evolution page.

9. References

- Seville Declaration: the commitment of cities to the Circular Economy. *Declaración de Sevilla: el compromiso de las ciudades por la Economía Circular.* http://www.femp.es/sites/default/files/multimedia/declaraciondesevilladefmaquetada.pdf
- Seville 2030 Strategic Plan. *Plan Estratégico Sevilla 2030* https://www.sevilla.org/planestrategico2030/documentos/sevilla-2030/plan-estrategico-sevilla-2030/objetivos-y-estrategias-sevilla2030.pdf