D4.1. GTool Functional Requirements and Reference Architecture

Lead: Engineering – Ingegneria Informatica S.p.A.

Contributors: Cetaqua Andalucía

Date: January 2021

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

Project deliverable

Project Number	**Project Acronym**	**Project Title**
1922	GOTHAM	Governance Tool for sustainable water resources allocation in the Mediterranean through stakeholder's collaboration. Towards a paradigm shift in groundwater management by end users.

Instrument: Research and Innovation Action	**Thematic Priority:** Mediterranean water co-operation
Title D4.1 GTool Functional Requirements and Reference Architecture	
Contractual Delivery Date 30 November 2020	**Actual Delivery Date** 26 January 2021
Organisation name of lead contractor for this deliverable Engineering – Ingegneria Informatica S.p.A.	**Document version** V1.0
Dissemination level Public X Confidential	**Deliverable Type** Document, Report X Demonstrator
Authors (organisation) ENG, CETAQUA	
Reviewers (organisation) Davide Storelli (ENG) Damián Sánchez (CETAQUA)	
Abstract The overarching objective of the GOTHAM project is to develop and validate a user-driven tool that enables effective groundwater governance to ultimately preserve the quantity and quality of this strategic resource in the Mediterranean basin. The GOTHAM Tool (GTool) uses an integrated methodological approach that targets optimal allocation of water resources from an	

environmental, social and economic perspective, including stakeholder knowledge, priorities and behaviour.

This report aims to define the functional requirements and the reference architecture of the GTool. Functional requirements and use cases are analysed with respect to different type of users that have emerged, covering the three main architectural levels: at the data and analytics layer technically savvy users will be provided with the necessary tools and functionalities to import data from heterogeneous data source and use such data for advanced analytical processes. The results of such analysis will be made available as easy to understand charts and visualizations for the different decision makers involved in the project: water producers, water suppliers, water end-users and water regulators.

The document defines the overall reference architecture of the GTool to be implemented in the rest of the project and identifies some preliminary data sources that will be used to feed the analytical processes.

Keywords

GTool, multicriteria analysis, data visualization, data integration, reference architecture, functional requirements

Disclaimer

Project summary

The overarching objective of the GOTHAM project is to develop and validate a user-driven tool that enables effective groundwater governance to ultimately preserve the quantity and quality of this strategic resource in the Mediterranean basin. The GOTHAM Tool (GTool) uses an integrated methodological approach that targets optimal allocation of water resources from an environmental, social and economic perspective, including stakeholder knowledge, priorities and behaviour.

One of the main strengths of the tool is that it provides a common framework for collaboration and engagement of the different water users (mainly, agricultural communities but also municipal and industrial users), as well as other relevant stakeholders such as water producers/operators and regulator(s). The GTool will enable them to exchange information in order to reach the optimal water governance at each point in time as well as in future scenarios.

The concept of the proposed GTool targets effective groundwater governance for the improvement of the management and preservation of this essential and strategic resource. This effective groundwater management remains an important and complex challenge in the Mediterranean and elsewhere, but is essential to ensure long-term sustainable use of the resource.

In this regard, GOTHAM integrates multicriteria decision methods for stakeholder group decision making and social learning, and use socio- hydrological water balance framework as a theoretical foundation for water allocation to evaluate the dynamic balance between the societal and ecological systems in catchments. GOTHAM project presents a bottom-up decision-making approach inspired in this methodological framework.

GOTHAM project presents a scalable and user-specific tool for decentralising water resources management. The proposed user-based tool leverages six analytical modules:

- The **water balance and water quality dynamics module** uses advanced investigation of the main aquifer formations and real-time monitoring (on site and distant), including preliminary analysis of the background hydrogeological and hydro-meteorological information to create a baseline.
- The **water availability and demand forecasting module** predicts different water scenarios and assess their impact on groundwater quality and quantity status using remote-sensing measurements to model agriculture water demand and assess water availability.
- The **Managed Aquifer Recharge (MAR) and aquifer remediation module** mobilises multicriteria analysis (QGIS environment), including hydrogeological, economic, and chemical (water quality) indicators as well as regulatory restrictions to evaluate the feasibility of MAR schemes.
- The **agro-economic module** simulates the effect of different economic instruments, such as water tariff structures, water markets contexts and incentives for water savings (water demand management) and assessing the economic use values and trade-offs between users in alternative resource allocation scenarios.
- The **user's engagement module** enables to fix water priorities (water boundary conditions) by water users, taking into consideration water resources to meet water demands.

- The **optimised water allocation module** calculates the optimal mix of water sources satisfying their requirements

The GTool uses data visualisation techniques to deliver the results into customisable dashboards tailored for the needs of each stakeholders.

Broad outreach activities will take place in Europe, Lebanon and Jordan, therefore contributing to GOTHAM impact maximization.

The further development and exploitation (beyond the project) of the GTool will be done by CETaqua, both on B2B and B2C approaches.

Table of content

Table of figures

List of tables

1. Introduction

The overarching objective of the present report is to define the functional requirements and the reference architecture of the GTool. Functional requirements and use cases are analysed with respect to different type of users that have emerged. On the one side technically skilled people who aim at collecting data and performing analysis on it with advanced tools and algorithms. On the other side end users, which have been identified already at proposal time as water producers, water suppliers, water end-users and water regulators. They are not interested in advanced analytical tools but in accessing their results and up-to-date data and information. Such information must be easy to understand and must allow evidence-based decision-making processes. From an architectural point of view, already at proposal time, the main layers of the GTool have been defined:

- GTool Data Layer, that will enable a smooth interfacing with the different and heterogeneous data sources of the water domain.
- GTool Analytics Layer, that will provide multicriteria analysis (availability, demand, quality, cost) regarding water allocation scenarios and their impact on the environmental, economic and social sustainability of the water resources.
- GTool Service Layer, that exposes, in different ways, data and results produced by the Analytics Layer and data imported from other sources.

The document defines specific functional requirements for each of the layers and proposes an overall reference architecture to be implemented in the rest of the project.

The rest of the document is structured as follows. In chapter 2 functional requirements and use cases of the GTool have been defined. In chapter 3 the GTool non-functional requirements have been defined. Chapter 4 defines the overall architecture of the GTool and describes its components, together with their main interactions. Chapter 5 is dedicated to the description of the first data sources that have been identified in order to feed the data layer. Chapter 6 concludes this report.

2. Functional Requirements and use cases

2.1. Users

2.1.1. User management

2.1.1.1. Functional Requirements

FUN_REQ_GT_USR_010 – Users of the system

- Data Analyst (DAn): Users with this role have access to the analytical tools in order to perform advanced analytical tasks and save the results for feeding visualizations dedicated to Decision Makers.
- Dashboard Administrator (DA). Users with this role are able to create and manage dataset, charts and dashboards. They enable decision makers to access dashboards.
- Decision maker (DM). Users with this role take decisions on the basis of the dashboards they consult. A decision maker can be one of the following: water producers, water suppliers, water end-users, water regulators. Each subtype of decision maker will access the dashboard dedicated to its role and territory.
- Admin. Users with this role have all possible rights, including granting or revoking the rights of other users.

FUN_REQ_GT_USR_020 - Interfacing with External Identity Providers for authentication

The GTool must be accessible by users already registered on an enabled external IDP (Identity Provider).

FUN_REQ_GT_USR_030 - Assigning roles to a user

Admin users will assign roles to users from external IDPs.

FUN_REQ_GT_USR_040 - Changing the role of a user

It must be possible to change the role of a user.

FUN_REQ_GT_USR_060 - Creating a role

It must be possible for administrative users to create ad-hoc roles based on granular permissions.

FUN_REQ_GT_USR_070 - Modify a Role

It must be possible for administrative users to modify a role by modifying the permissions assigned to it.

FUN_REQ_GT_USR_080 - Deleting a role

It must be possible for administrative users to delete a role.

2.1.1.2. Use Cases

2.1.1.2.1. Use Case: User Authentication

Name	User Authentication
Purpose	It allows users authenticate to the GTool
Actors	All users
Pre-Conditions	The user is not logged into the platform.
Requirements Covered	FUN_REQ_GT_USR_020
Workflow	1. The system displays the authentication interface. 2. The user can access the platform through the Identity Manager. The system delegates the authentication process to the selected Identity Manager (Keyrock, etc.), delegating the user authentication procedure.
Post-Conditions	The system presents the interface provided for that type of user (Administrator or Decision maker).

Table 1 - Use case: User Authentication

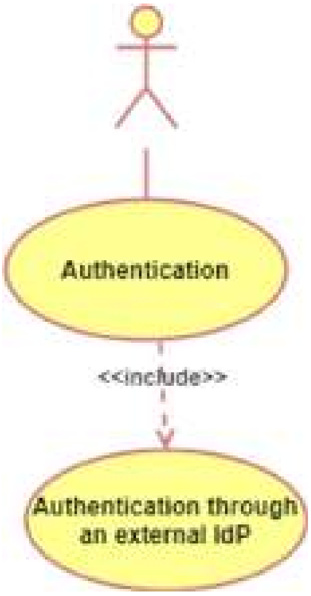

Figure 1 – Use case: User Authentication

2.1.1.2.2. Use Case: Management of user roles

Name	Management of user roles
Purpose	Allows administrative users to manage the roles of registered users
Actors	Admin
Pre-Conditions	User is logged to the platform
Requirements Covered	FUN_REQ_GT_USR_030, FUN_REQ_GT_USR_040
Workflow	1. The system displays the user management interface: 2. The administrator can choose to: a. assign a certain role to a user. b. Change a user's role.
Post-Conditions	The target user have a new role assigned.

Table 2 – Use case: Management of user roles

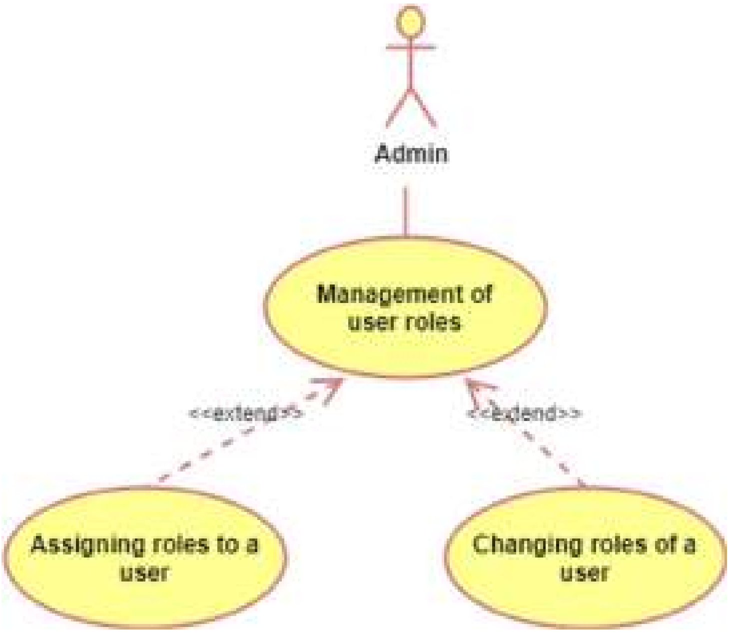

Figure 2 – Use case: Management of user roles

2.1.1.2.3. Use Case: Roles management

Name	Roles management
Purpose	Allows Admin to manage roles
Actors	Admin
Pre-Conditions	User is logged to the platform
Requirements Covered	FUN_REQ_GT_USR_060, FUN_REQ_GT_USR_070, FUN_REQ_GT_USR_080
Workflow	1. The system displays the role management interface. 2. The administrator can choose to: a. Create a new role by assigning specific permissions (for example, access to a single system dataset). b. Modify an existing role by modifying the granular permissions that compose it c. Delete an existing role.
Post-Conditions	A specific role is created, modified or deleted.

Table 3 – Use case: Roles Management

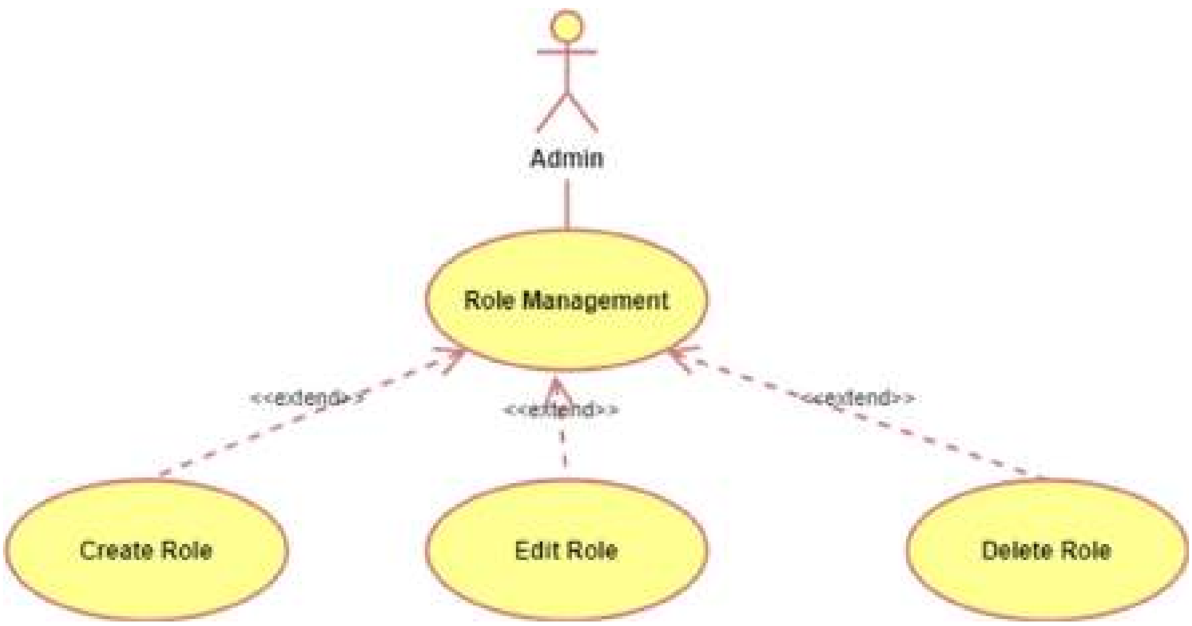

Figure 3 – Use case: Role Management

2.2. Data Layer

2.2.1. Data Source import management

2.2.1.1. Functional Requirements

FUN_REQ_DLY_DS_010 – Import of time-series file from a local filesystem

It must be possible to acquire time series from the local filesystem. These can be in text and csv format.

FUN_REQ_DLY_DS_020 – Import time-series from open data portal

It must be possible to acquire time series from an open data portal. These can be in text and csv format. It must be possible to specify at least the delimiter used and the header line containing the column names (if any). It must be possible to specify synchronization parameters (if required).

FUN_REQ_DLY_DS_030 – Import time-series from a REST service

This feature is useful in case you need to acquire the time series from sources that expose data through a RESTful interface.

FUN_REQ_DLY_DS_040 – Import of GIS file from a local filesystem

It must be possible to acquire GIS files from the local file system.

FUN_REQ_DLY_DS_050 – Import of GIS files from a REST service

This feature is useful in case you need to acquire GIS files from sources for that expose data through RESTful interface.

FUN_REQ_DLY_DS_060 – GIS files format

The supported GIS file formats are: shapefile and ratser files.

FUN_REQ_DLY_DS_070 – Import field measurement

It must be possible to import data from field sensors placed in the area of interest (e.g. piezometric level, water quality, water flows).

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

2.2.1.2. Use Cases

2.2.1.2.1. Use Case: Data Source Management

Name	Data source Management
Purpose	It allows users import data sources in the system.
Actors	DA
Pre-Conditions	The user is logged to the platform.
Requirements Covered	REQ_FUN_DLY_DS_010, REQ_FUN_DLY_DS_020, REQ_FUN_DLY_DS_030, REQ_FUN_DLY_DS_040, REQ_FUN_DLY_DS_050, REQ_FUN_DLY_DS_060, REQ_FUN_DLY_DS_070.
Workflow	1. The user accesses the system and can: a. Import a time series from the local filesystem. b. Import a time series from an open data portal. c. Import a time series using a REST service. d. Import a GIS file from the local filesystem. e. Import a GIS file using a REST service. f. Import of field measurement from sensors.
Post-Conditions	The system, according to the operations carried out by the user: • Imports a time series from the local filesystem storing it in GTool non-spatial RDBMS. • Connects to an open data portal to download the specific time-series file and import it in GTool non-spatial RDBMS. • Connects to the REST service, that returns the time series file (or the url to download it) and then parses the file and uploads the time-series in the non-spatial RDBMS. • Imports a GIS file from the local filesystem storing it in GTool the spatial fileserver of the GTool.

	• Connects to the REST service that returns the GIS file (or the url to dowload it). The file is then stored in spatial fileserver of GTool. • The system imports data retrieved from field sensors. The data are stored in the GTool non-spatial RDBMS.

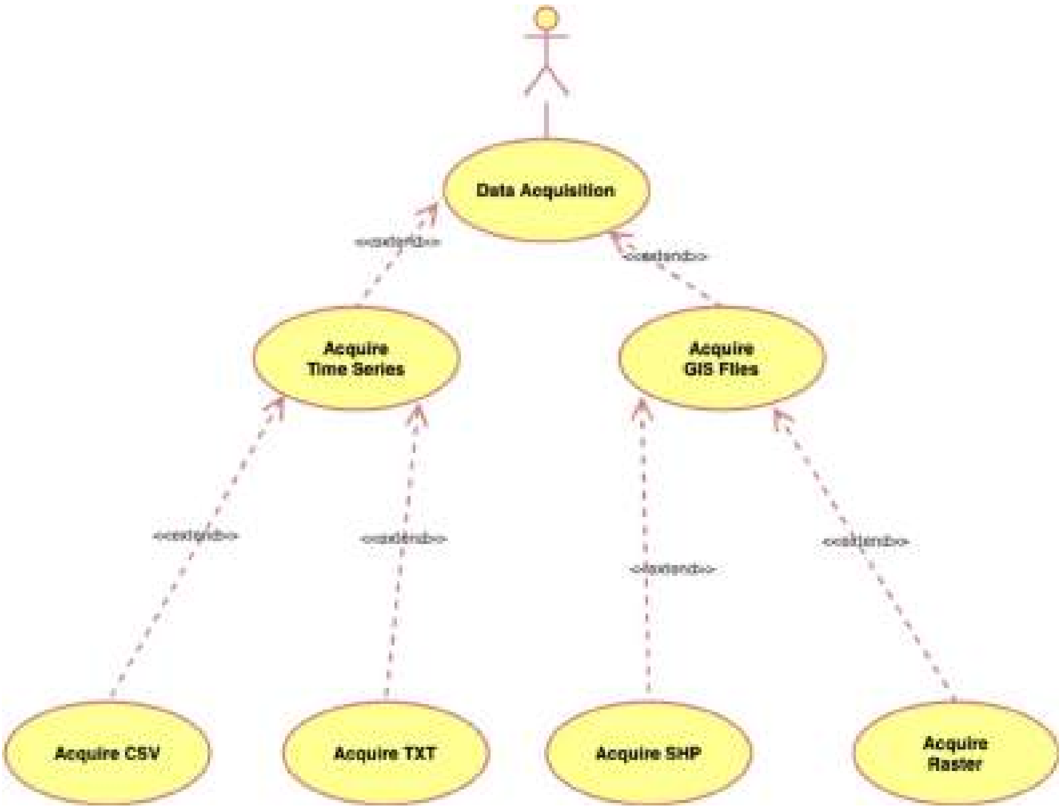

Figure 4 – Use case: data acquisition

2.3. Analytics Layer

2.3.1. Water balance and water quality dynamics

2.3.1.1. Groundwater balance decision support system

FUN_REQ_AL_GWB_010 – Selection of input data

It must be possible to gather data regarding the following time series:

- Rainfall (obtained from online data sources or local filesystem)

- Temperature (minimum, maximum and average) (obtained from online data sources or local filesystem)
- Groundwater level (obtained from online data sources or local filesystem)

The format of the input data will be one of the following:

- .txt or .csv (monthly scale)
- Shapefiles (.shp, .json, .geojson) and raster (.tiff)

FUN_REQ_AL_GWB_020 – Selection of options

It must be possible to select the desired options regarding the following boundary conditions:

- Aquifer material (selection among a pre-defined list of available options)
- Aquifer functioning (selection among a pre-defined list of available options)
- Location
- Surface (km^2)
- Temporal scale of the analysis (it must possible to select among monthly or annual)
- Groundwater budget components to be assessed (selection among a pre-defined list of available options)

The above list may be supplemented by any other variables deemed necessary during the model development phase.

FUN_REQ_AL_GWB_030 – Groundwater budget calculation

Using input data from FUN_REQ_AL_GWB_010 and the boundary conditions and options from FUN_REQ_AL_GWB_020, the model will calculate the aquifer groundwater budget on a monthly or annual scale.

Model output results will refer to groundwater recharge and/or piezometric levels, and will be available in one of the following formats or any other equivalent: .txt, .csv, .xls, .xlsx.

2.3.1.2. Groundwater potential index and mapping

FUN_REQ_AL_GWPI_010 – Selection of input data

It must be possible to gather data regarding the following geospatial information:

- Rainfall maps
- Texture soil maps
- Surface geology
- DEM (digital elevation model) analysis (slope, drainage density, etc.)
- Satellite data regarding land use and/or land cover, such as LANDSAT and Sentinel

- Water points

This data could be obtained from remote sensing services like Copernicus, Sentinel Hub or Google Earth Engine.

FUN_REQ_AL_GWPI_020 – Selection of options

It must be possible to select the desired options regarding the following variables:

- Geographical limits of the study area
- Algorithm to be used (in case several algorithms are available)

FUN_REQ_AL_GWPI_030 – Groundwater potential calculation

Using input data from FUN_REQ_AL_GWPI_010 and the options from FUN_REQ_AL_GWPI_020, the model will calculate the groundwater potential in the selected study area and will present the results as coloured maps, where each colour will represent a different groundwater potential grade.

2.3.1.3. Statistical multivariate analysis

FUN_REQ_AL_SMA_010 – Selection of input data

It must be possible to import data from the local file system. This data will refer to the following physicochemical parameters of groundwater:

- Electrical conductivity
- Temperature
- pH
- Major components concentration in the water (Cl^-, HCO_3^-, SO_4^{2-}, Ca^{2+}, Mg^{2+}, Na^+ and K^+)

The format of the input data will be .txt or .csv (monthly or quarterly scale).

FUN_REQ_AL_SMA_020 – Selection of statistical analysis

It will be possible to select the statistical analysis to be carried out among a close list of available multivariate statistical methods, for instance Principal Components Analysis, Cluster Analysis, Factorial Analysis, Multiple Regression Analysis, and Discriminant Analysis.

FUN_REQ_AL_SMA_030 – Classification of water samples according to their physicochemical characteristics

Model output will correspond to one or several graphs showing homogeneous groups of water samples with a similar physicochemical nature.

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

2.3.2. Water availability and demand forecasting

2.3.2.1. Drought Monitoring and Early Warning System

FUN_REQ_AL_DMEWS_010 – Selection of data

It must be possible to gather data regarding the following time series:

- Streamflow and stored water in reservoirs
- Groundwater level
- Rainfall
- Temperature
- Evapotranspiration
- Water demands (agriculture, urban and industrial)

These time series could come from online data sources or from the local file system.

The format of the input data will be one of the following:

- .txt or .csv (daily or monthly scale)
- Raster files (monthly scale)

FUN_REQ_AL_DMEWS_020 – Selection of forecasting procedures and temporal scale

The user will be able to select the forecasting methodology to be applied (artificial neural networks, support vector machines, autoregressive models…) as well as the desired temporal scale for the analysis to be conducted.

FUN_REQ_AL_DMEWS_030 – Drought status and result of the early warning system

The results of the model regarding current drought status and the early warning system will be shown graphically.

2.3.2.2. Agriculture water demand

FUN_REQ_AL_AWD_010 – Selection of data

It must possible to obtain remote sensing data for the following parameters:

- Actual evapotranspiration
- Crop types and areas
- Remote-sensing data in order to deduce the developed state of crops
- Land use changes

The above data will come from online data services such as Copernicus, Sentinel Hub or Google Earth Engine.

The format of the input data will be one of the following:

- Raster files (monthly scale)
- .txt or .csv (monthly scale)

- ***FUN_REQ_AL_AWD_020 – 'Real time' agriculture water demand***

The output of the remote sensing-based models will be an estimation of the current agriculture water demand in the study area.

2.3.3. Agro-economic module

FUN_REQ_AL_AEM_010 – Input of data

It must be possible to input data of the following parameters:

- Crop value
- Input costs (including water tariff and incentives for water savings)
- Water-yield relationship
- Water value / Demand
- Marginal water value
- Irrigation water demand

All these data will be directly imported from the local file system.

FUN_REQ_AL_AEM_020 – Definition of the resource allocation scenario (boundary conditions set-up)

This should allow the user to define the boundary conditions of the resource allocation scenario selected, in terms of maximum volume of resource available for each water source and each water demand.

FUN_REQ_AL_AEM_030 – Results of the simulation

The output of the model will be the results of the simulation conducted under the criteria previously selected under Functional Requirement AL_AEM_010.

2.3.4. Managed aquifer recharge and aquifer remediation

2.3.4.1. Managed aquifer recharge

FUN_REQ_AL_MAR_010 – Selection of input data

It must be possible to input data on the following parameters:

- Permeability of rocks
- Volume of surplus water resources
- Recharge water quality
- Water quality of the recharged aquifer
- Recharge water price
- Land use
- Water residence time (estimation)

FUN_REQ_AL_MAR_020 – Selection of options

The user will be allowed to choose the following options:

- Geographical limits of the study area
- Regulatory restrictions (if any)

FUN_REQ_AL_MAR_030 – Identification of suitable areas for Managed Aquifer Recharge

The result of the model will be a map of the study area classified according to the suitability for carrying out managed aquifer recharge considering hydrogeological, economic and hydrochemical aspects.

2.3.4.2. Aquifer remediation

FUN_REQ_AL_AR_010 – Selection of options

The user will choose one option for the following parameters:

- Groundwater remediation technique (from a closed list of available remediation methods previously established)
- Aquifer properties (from a closed list of available aquifer properties)
- Pollutant targeted (from a closed list of available pollutants)
- Technical constraints / requirements (if any)

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

FUN_REQ_AL_AR_020 – CAPEX / OPEX calculation

The model will provide the user with an estimation of the capital expenses (CAPEX) and operational expenses (OPEX) necessary to carry out a remediation work on the selected aquifer, considering the remediation method, aquifer type and targeted pollutant chosen in the previous section.

2.3.5. Groundwater bodies' response

FUN_REQ_AL_GWBR_010 – Selection of input data

It must be possible to select the input data regarding the following parameters:

- Initial water balance
- Initial water quality
- Modelling time scale
- Water availability forecasting
- Water demand forecasting
- Climate forecasting

The format of the above input data could be:

- Boundary conditions (fixed information)
- Time series (.txt, .csv)
- Results from previous sections (for instance water availability and water demand forecasting)

FUN_REQ_AL_GWBR_020 – Selection of groundwater governance scenario

The user will be able to select a specific groundwater governance scenario to be modelled, using the input data from the previous section. A detailed description of the fields required to describe the governance scenario will be given in a more advance state of the project.

FUN_REQ_AL_GWBR_030 – Groundwater body status

The result of the modelling will be:

- Groundwater indexes (describing, for instance, level trends, exploitation, seawater intrusion or agriculture pollution)
- Groundwater body status resulting from the selected groundwater governance scenario

2.3.6. User's engagement

FUN_REQ_AL_UE_010 – Definition of boundary conditions of the water system

The user must be able to fill in the following fields necessary to define the boundary conditions of the water system to be modelled:

- Water network and infrastructure
- Water quality thresholds
- Economic constraints
- Volume of water required for each user
- Temporal needs

The format of the data for each of the above options will be defined in a more advanced state of the project.

2.3.7. Use cases

2.3.7.1. Water balance and water quality dynamics

2.3.7.1.1. Groundwater balance decision support system

Name	Groundwater balance decision support system
Purpose	It allows users to calculate the aquifer groundwater budget on a monthly or annual scale.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_GWB_010, FUN_REQ_AL_GWB_020, FUN_REQ_AL_GWB_030.
Workflow	1. The system displays the 'Groundwater balance decision support system' interface. 2. From this interface the user can: i. Select the input data to conduct the analysis ii. Select a set of options iii. Carry out the analysis in order to obtain the aquifer groundwater

	budget
Post-Conditions	The system according to the operations carried out by the user: • Import time series data on rainfall, temperature and groundwater level of the aquifer in .txt, .csv, .shp, .json, .grojson or .tiff format. • Let the user choose several options regarding the following boundary conditions: - Aquifer material - Aquifer functioning - Location - Surface - Temporal scale of the analysis - Groundwater budget components • Calculate the aquifer groundwater budget (on a monthly or annual scale) using the input data and the boundary conditions selected by the user.

2.3.7.1.2. Groundwater potential index and mapping

Name	Groundwater potential index and mapping
Purpose	It allows users to calculate the groundwater potential in the selected study area and present the results as coloured maps.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_GWPI_010, FUN_REQ_AL_GWPI_020, FUN_REQ_AL_GWPI_030.
Workflow	1. The system displays the 'Groundwater potential index and mapping' interface. 2. From this interface the user can: i. Select the input data to conduct the analysis ii. Select a set of options

	iii. Carry out the analysis in order to calculate the groundwater potential
Post-Conditions	The system according to the operations carried out by the user: • Import data about water points in the study area from the local file system. • Let the user choose several options regarding the following parameters: - Geographical limits of the study area - Algorithm to be used (in case several algorithms are available) • Calculate the groundwater potential budget and present the results as coloured maps, where each colour will represent a different groundwater potential grade.

2.3.7.1.3. Statistical multivariate analysis

Name	Statistical multivariate analysis
Purpose	It allows users to classify groundwater samples into homogeneous groups according to their physicochemical characteristics, using multivariate statistical methods.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_SMA_010, FUN_REQ_AL_SMA_020, FUN_REQ_AL_SMA_030.
Workflow	1. The system displays the 'Statistical multivariate analysis' interface. 2. From this interface the user can: i. Select the input data to conduct the analysis ii. Select the statistical method to be applied iii. Carry out the statistical analysis in order to obtain the classification of groundwater samples in homogeneous groups
Post-Conditions	The system according to the operations carried out by the user: • Import data from the local filesystem referring to water electrical conductivity, water temperature, water pH and water major components concentration (Cl^-,

	HCO_3^-, SO_4^{2-}, Ca^{2+}, Mg^{2+}, Na^+ and K^+). The format of the input data will be .txt or .csv (monthly or quarterly scale). • Let the user choose the statistical analysis to be carried out among a close list of available multivariate statistical methods, for instance Principal Components Analysis, Cluster Analysis, Factorial Analysis, Multiple Regression Analysis, and Discriminant Analysis. • Carry out the selected statistical analysis and show one or several graphs with the classification of the groundwater samples according to their physicochemical characteristics.

2.3.7.2. Water availability and demand forecasting

2.3.7.2.1. Drought Monitoring and Early Warning System

Name	Drought Monitoring and Early Warning System
Purpose	It allows users to define the current drought status and display the early warning system graphically.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_DMEWS_010, FUN_REQ_AL_DMEWS_020, FUN_REQ_AL_DMEWS_030.
Workflow	1. The system displays the 'Drought Monitoring and Early Warning System' interface. 2. From this interface the user can: i. Select the input data to conduct the analysis ii. Select the forecasting procedure and temporal scale to be applied iii. Calculate the current drought status and show the early warning system

Post-Conditions	The system according to the operations carried out by the user: • Import time series from the local file system relative to the following parameters: streamflow and stored water in reservoirs, groundwater level, rainfall, air temperature, evapotranspiration and water demands (agriculture, urban and industrial). The format of the input data will be .txt / .csv (daily or monthly scale) or raster files (monthly scale). • Let the user choose the forecasting method to be applied (artificial neural networks, support vector machines, autoregressive models…) as well as the desired temporal scale. • Calculate the current drought status and the early warning system, and show the latter graphically.

2.3.7.2.2. Agriculture water demand

Name	Agriculture water demand
Purpose	It allows users to estimate the current agriculture water demand in the study area.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_AWD_010, FUN_REQ_AL_AWD_020.
Workflow	1. The system displays the 'Agriculture water demand' interface. 2. From this interface the user can: i. Obtain remote sensing data ii. Estimate the current agriculture water demand in the study area.
Post-Conditions	The system according to the operations carried out by the user: • Import data on actual evapotranspiration; crop types and areas; remote-

	sensing data in order to deduce the developed state of crops; and land use changes. The above data will come from online data services such as Copernicus, Sentinel Hub or Google Earth Engine (the specific service will be specified later). • Calculate 'real time' agriculture water demand in the study area.

2.3.7.3. Agro-economic module

Name	Agro-economic module
Purpose	It allows users to simulate several agro-economic parameters based on a resource allocation scenario previously defined by the user.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_AEM_010, FUN_REQ_AL_AEM_020, FUN_REQ_AL_AEM_030.
Workflow	1. The system displays the 'Agro-economic module' interface. 2. From this interface the user can: i. Select the input data to conduct the analysis ii. Define the resource allocation scenario iii. Conduct the simulation
Post-Conditions	The system according to the operations carried out by the user: • Import data of the following parameters: Crop value, Input costs (including water tariff and incentives for water savings), water-yield relationship, water value / demand, marginal water value and irrigation water demand. All these data will be directly imported from the local filesystem. • Let the user define the boundary conditions of the resource allocation scenario selected, in terms of maximum volume of resource available for each water source and each water demand.

	• Carry out the simulation and obtain forecasted results of several agro-economic parameters.

2.3.7.4. Managed aquifer recharge and aquifer remediation

2.3.7.4.1. Managed aquifer recharge

Name	Managed aquifer recharge
Purpose	It allows users to obtain a map of the studied area classified according to the suitability for carrying out managed aquifer recharge, considering hydrogeological, economic and hydrochemical criteria.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_MAR_010, FUN_REQ_AL_MAR_020, FUN_REQ_AL_MAR_030.
Workflow	1. The system displays the 'Managed aquifer recharge' interface. 2. From this interface the user can: i. Select the input data to conduct the analysis ii. Choose among a set of options iii. Conduct the analysis and obtain a map classified according to the higher/lower suitability to implement managed aquifer recharge operations in the study area
Post-Conditions	The system according to the operations carried out by the user: • Let the user load data related to permeability of rocks, volume of surplus water resources, recharge water quality, water quality of the recharged aquifer, recharge water price, land use and water residence time (estimation). • Let the user choose the geographical limits of the study area and to incorporate regulatory restrictions -if there is any-. • Identify the most suitable areas to implement managed aquifer recharge

	operation based on a map of the study area displaying higher / lower suitability according to hydrogeological, economic and hydrochemical aspects.

2.3.7.4.2. Aquifer remediation

Name	Aquifer remediation
Purpose	It provides users with an estimation of the capital expenses (CAPEX) and operational expenses (OPEX) necessary to carry out a remediation work on the selected aquifer, considering the remediation method, aquifer type and targeted pollutant.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_AR_010, FUN_REQ_AL_AR_020.
Workflow	1. The system displays the 'Aquifer remediation' interface. 2. From this interface the user can: i. Select several options ii. Calculate CAPEX / OPEX for the selected set of options
Post-Conditions	The system according to the operations carried out by the user: • Let the user choose one option for the following parameters: – Groundwater remediation technique (from a closed list of available remediation methods previously established) – Aquifer properties (from a closed list of available aquifer properties) – Pollutant targeted (from a closed list of available pollutants) – Technical constraints / requirements (if any) • Calculate an estimation of the capital expenses (CAPEX) and operational expenses (OPEX) necessary to carry out a remediation work on the selected aquifer, considering the remediation method, aquifer type and targeted pollutant chosen in the previous section.

2.3.7.5. Groundwater bodies' response

Name	Groundwater bodies' response
Purpose	It allows users to obtain a set of indexes describing selected groundwater-related parameters, as well as the expected groundwater body status based on a specific groundwater governance scenario selected by the user.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_GWBR_010, FUN_REQ_AL_GWBR_020, FUN_REQ_AL_GWBR_030.
Workflow	1. The system displays the 'Groundwater bodies' response' interface. 2. From this interface the user can: i. Select input data ii. Select a specific groundwater governance scenario iii. Model a set of groundwater indexes as well as the groundwater body status resulting from the selected groundwater governance scenario
Post-Conditions	The system according to the operations carried out by the user: • Let the user select the input data corresponding to the following parameters: – Initial water balance – Initial water quality – Modelling time scale – Water availability forecasting – Water demand forecasting – Climate forecasting The format of the above input data could be: – Boundary conditions (fixed information) – Time series (.txt, .csv) – Results from previous sections (for instance water availability and

	water demand forecasting) • Let the user select a specific groundwater governance scenario to be modelled, using the input data from the previous section. • Calculate: – Groundwater-related indexes describing, for instance, level trends, exploitation, seawater intrusion or agriculture pollution – The expected groundwater body status resulting from the selected groundwater governance scenario

2.3.7.6. User's engagement

Name	User's engagement
Purpose	It allows users to define the boundary conditions of the water system to be modelled.
Actors	DAn
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	FUN_REQ_AL_UE_010.
Workflow	1. The system displays the 'User's engagement' interface. 2. From this interface the user can: i. Complete a close set of fields defining the boundary conditions of the water system to be modelled.
Post-Conditions	The system according to the operations carried out by the user: • Let the user fill in the following fields necessary to define the boundary conditions of the water system: – Water network and infrastructure – Water quality thresholds – Economic constraints – Volume of water required for each user

	– Temporal needs The format of the above input data will be defined in a more advanced state of the project.

2.4. Service Layer

2.4.1. Charts Management

2.4.1.1. Functional Requirements

FUN_REQ_SRV_CRT_010 - Creating a chart

It must be possible to create the charts to represent results from the analytics layer or to display data retrieved from other sources

FUN_REQ_SRV_CRT_020 - Deleting a chart

It must be possible to delete a chart.

FUN_REQ_SRV_CRT_030 - Editing a chart

It must be possible to modify a chart both by type, by data source used and by modifying the specific parameters of the single chart.

FUN_REQ_SRV_CRT_040 - List of created charts

It must be possible to view a list of the charts created and allow the user, for each of them, to access the functions for deleting, modifying and viewing the chart details.

FUN_REQ_SRV_CRT_050 - view details of a chart

It must be possible to view some information on a particular chart in particular, at least the following information must be available:

- Owners. Owners of the chart
- Created By. User who created the chart
- Changed By. Last user who modified the chart
- Created On. Creation date

- Changed On, Modified date
- Name of the chart

FUN_REQ_SRV_CRT_060 - Search for created charts

It must be possible to search the charts by:

- name
- description
- type
- owner
- data source

2.4.1.2. Use cases

2.4.1.2.1. Use Case: Chart Management

Name	Chart Management
Purpose	It allows users to create and manage charts.
Actors	DA
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	REQ_FUN_SRV_CRT_10, REQ_FUN_SRV_CRT_20, REQ_FUN_SVR_CRT_30, REQ_FUN_SVR_CRT_40, REQ_FUN_SVR_CRT_50, REQ_FUN_SVR_CRT_60.
Workflow	1. The system displays the chart management interface with the list of all charts created. 2. The user from this interface can: a. Create a new chart. b. Edit an existing chart. c. Delete an existing chart. d. Search for an existing chart. e. View the details of an existing chart.
Post-Conditions	The system according to the operations carried out by the user: • Creates a new chart • Edits an existing chart • Deletes a chart • Searches among the existing charts and shows the results of the search on the screen. • Displays the details of an existing chart.

Table 4 - Use case: Chart management

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

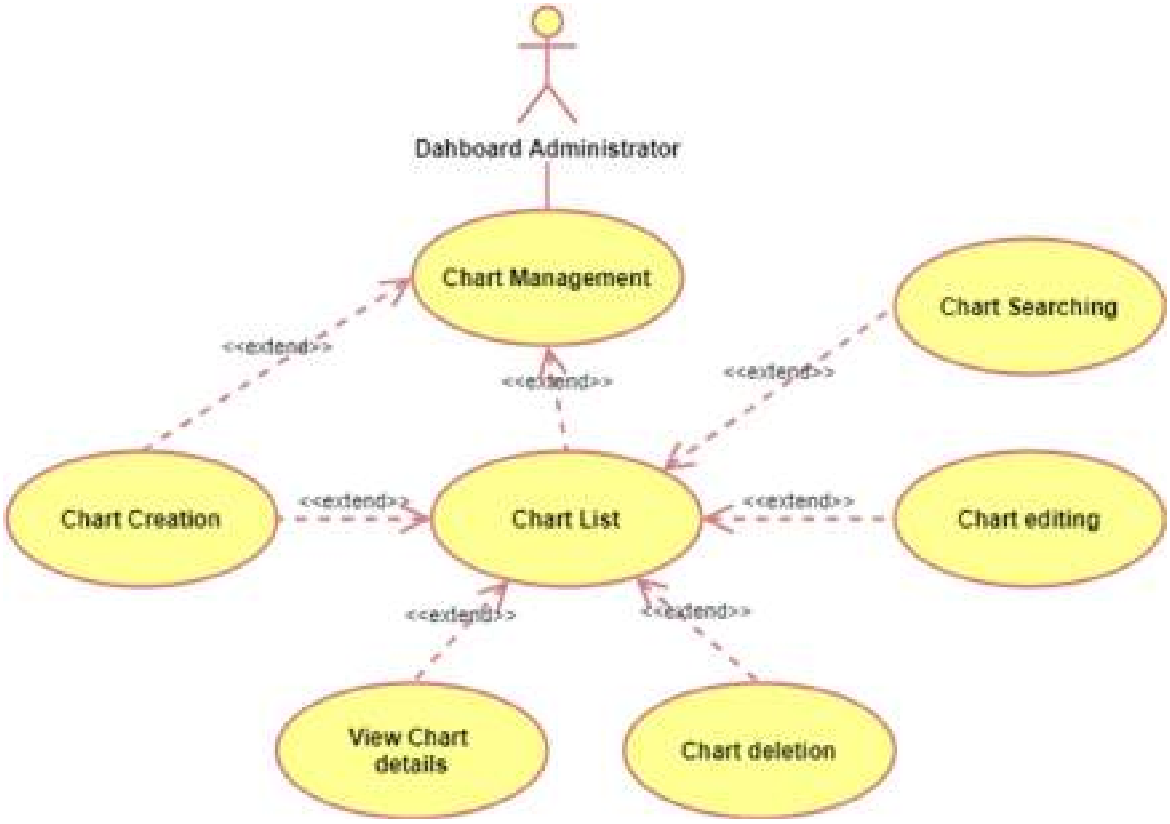

Figure 5 – Use case: chart management.

2.4.2. Dashboard management

2.4.2.1. Functional Requirements

FUN_REQ_SRV_DSH_010 – Creating a Dashboard

It must be possible to create a dashboard. To create a dashboard you must specify at least the name of this.

REQ_FUN_SRV_DSH_020 - Deleting a Dashboard

It must be possible to delete a dashboard.

REQ_FUN_SRV_DSH_030 - Modify Dashboard Data

It must be possible to edit a dashboard by changing, for example, its name.

REQ_FUN_SRV_DSH_040 - Inserting a Chart in a Dashboard

It must be possible to add an existing chart to a dashboard by positioning it appropriately.

REQ_FUN_SRV_DSH_041 - Editing a Chart in a Dashboard

It must be possible to change the position of the chart in the dashboard.

REQ_FUN_SRV_DSH_042 - Deleting a chart in a dashboard

It must be possible to remove a chart from a specific dashboard.

REQ_FUN_SRV_DSH_050 - Inserting a Graphic Element in a Dashboard

It must be possible to add graphic elements to a dashboard such as Header, Tabs, Footer by positioning them appropriately.

REQ_FUN_SRV_DSH_051 - Editing a Graphic Element in a Dashboard

It must be possible to change the position of the graphic elements within a dashboard and, where required, edit the content (e.g. Header and Footer)

REQ_FUN_SRV_DSH_052 - Deleting a Graphic Element in a Dashboard

It must be possible to delete a graphic element from the dashboard

REQ_FUN_SRV_DSH_060 - Adding a Filter to Dashboard

It must be possible to add an element that allows you to filter the charts on the dashboard based on the selection of specific attributes.

REQ_FUN_SRV_DSH_070 - Publishing a Dashboard

It must be possible to publish a dashboard so that it is available for visualization to other users.

REQ_FUN_SRV_DSH_080 – View Dashboard Details

It must be possible to view the details of a dashboard which include at least the following information:

- Name
- Owner
- Whether it is published or not
- CSS
- The charts present in the dashboard

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

REQ_FUN_SRV_DSH_090 – Dashboard List

It must be possible to view a list of the dashboards created and allow, for each of them, to access the functionalities for deleting, modifying and viewing the details of the dashboard.

REQ_FUN_SRV_DSH_100 - Search for a Dashboard

It must be possible to search for dashboards in the system. The search must be based on at least on the following criteria:

- Name
- Owner
- Whether it is published or not

REQ_FUN_SRV_DSH_110 – Enable decision makers to view a specific dashboard

The dashboard manager can enable the view of a dashboard to a group of decision makers on the basis of their role and on the basis of the territory of their competence.

- **Use Cases**

2.4.2.1.1. Use Case: Dashboard Management

Name	Dashboard Management
Purpose	It allows users to create and manage dashboards.
Actors	DA
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	REQ_FUN_SRV_DSH_10, REQ_FUN_ SRV_DSH_20, REQ_FUN_ SRV_DSH_30, REQ_FUN_SRV_ SRV_40, REQ_FUN_ SRV_DSH_41, REQ_FUN_SRV _DSH_42, REQ_FUN_SRV_DSH_50, REQ_FUN_SRV_DSH_60, REQ_FUN_SRV_DSH_70, REQ_FUN_SRV_DSH_80, REQ_FUN_SRV_DSH_90, REQ_FUN_SRV_DSH_100, REQ_FUN_SRV_DSH_110
Workflow	1. The system displays the dashboard management interface with the list of the dashboard managed by the user. 2. The user from that interface can:

	• Create a new dashboard or modify it. Among the operations of creation or modification we have: i. Inserting a chart in the dashboard. ii. Deleting a chart from a dashboard. iii. Edit chart in position dashboard iv. Inserting graphics in the dashboard. v. Editing graphics in dashboard vi. Deleting graphics from the dashboard vii. Insert filter in a dashboard viii. Edit filter in dashboard (location). ix. Delete filter in the dashboard. • Delete an existing dashboard • Search for an existing dashboard • View the details of an existing dashboard • Enable decision makers to view a particular dashboard
Post-Conditions	The system according to the operations performed by the user: 1. Create a new dashboard 2. Edit an existing dashboard 3. Delete a dashboard 4. Search through existing dashboards and show search results on screen 5. Display the details of an existing dashboard. 6. Grant permission to group of decision makers in order to view a particular dashboard

Table 5 – Use case: Dashboard Management

Figure 6 – Use case. Dashboard Management

2.4.3. Dashboard frontend

2.4.3.1. Functional Requirements

REQ_FUN_DSR_DFE_010 – Dashboards List

It must be possible to view a list of the dashboards created and allow, for each of them, to access the functionalities for viewing the details of the dashboard or open it in read only mode.

REQ_FUN_DSR_DFE_020 - Search for a Dashboard

It must be possible to search for dashboards in the system. The search must be based at least on the following criteria:

- Name
- Owner
- Whether it is published or not

REQ_FUN_SRV_DFE_030 – View Dashboard Details

It must be possible to view the details of a dashboard which include at least the following information:

- Name
- Owner
- Whether it is published or not
- CSS
- The charts present in the dashboard

REQ_FUN_SRV_DFE_040 – View/Navigate Dashboard

The user can consult a specific dashboard displaying all the elements (charts, graphics elements) that are present in it and navigating through tabs.

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

2.4.3.2. Use cases

Name	Dashboard Frontend
Purpose	It allows decision makers to consults dashboards that belongs to them
Actors	Decision maker
Pre-Conditions	The user is logged in to the platform.
Requirements Covered	REQ_FUN_SRV_DFE_10, REQ_FUN_SRV_DFE_20, REQ_FUN_SRV_DFE_30, REQ_FUN_SRV_DFE_40.
Workflow	1. The system displays the dashboard frontend interface with the list of dashboards belonging to the user. 2. The user from that interface can: i. View/navigate a specific dashboard ii. Search for a Dashboard specifying several search criteria. iii. View the details of a dashboard (metadata).
Post-Conditions	The system according to the operations performed by the user: 1. Displays the selected dashboard allowing the user to navigate it 2. Displays the result of a research. 3. Display the details (metadata) of a specific dashboard.

Table 6 – Use case: Dashboard Frontend

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

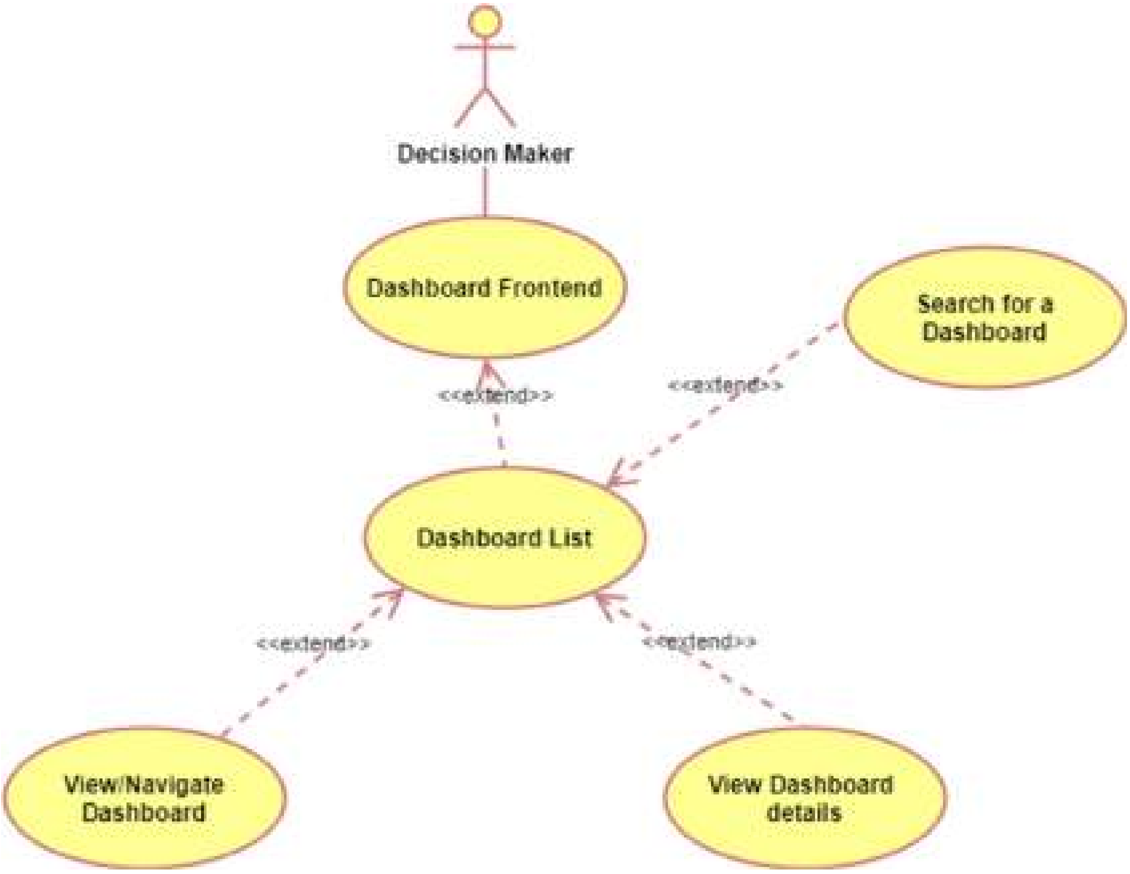

Figure 7 – Use case: Dashboard Frontend

3. Non-Functional Requirements

3.1. Security

In order to guarantee the appropriate security level the following features must be supported:

- access to the databases takes place only through services distributed on machines configured in the DMZ;
- credentials must be blocked in the face of repeated failed authentication attempts;
- logical-physical perimeter protection measures must always be present, such as, for example, firewalls;
- the software systems, the programs used and the antivirus protection must be constantly updated on the servers;

- security measures must be periodically reconsidered and adapted to technical progress and the evolution of risks;
- the user authentication procedure must be protected from the risk of interception of credentials by cryptographic mechanisms of adequate strength;

3.2. Interoperability

It is assured interoperability with external data sources.
It is required that protocols, models and standard or de facto standard data formats are used whenever possible, in order to maximize interoperability with external systems.

3.3. Usability

The usability of software products is a set of attributes that are related to the effort required to use, and to the individual evaluation of such use, by a declared or implied group of users.

The attributes of the usability feature of a software product are:

- *understandability*, which represents the ability to allow the user to understand the software product functionality and how to use it successfully to perform particular tasks under certain conditions of use;
- *learnability*, which represents the ability to allow the user to learn how to use the application;
- *operability*, that is, the ability to allow the user to use and control a software product. This attribute also refers to the user's expectations on the software product controllability, fault tolerance and compliance;
- *attractiveness*, which represents the ability of a software product to be attractive to the user. It relates to the design of the graphic aspect of its interfaces, the use of colours and images, etc.

Among "users" there are operators, end users and other indirect users who are influenced by or dependent on the software. Usability will have to address all environments and product usage scenarios, including preparation for use of the software and evaluation of results.

3.4. Scalability and Performance

To guarantee the appropriate performance level to system, scalability must be ensured.

The system must be able to increase or reduce its performance, resources and functionalities according to user's needs.

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

For the system deployment must be used a Cloud Infrastructure that uses containers because they enable both horizontal and vertical scaling and represents the state of the art of the cloud solution currently available on the market.

4. Architecture

4.1. Overview

The GTool is the integrated environment that will provide all the services, applications and tools developed in the Gotham project. The aim of the GTool is to support effective groundwater governance for the improvement of the management and preservation of this essential and strategic resource. The GTool will allow actual data leveraging in groundwater governance decisions. This section builds upon the requirements and use cases defined in the previous chapters in order to depict and describe the general architecture of the GTool and the specific roles of its components.

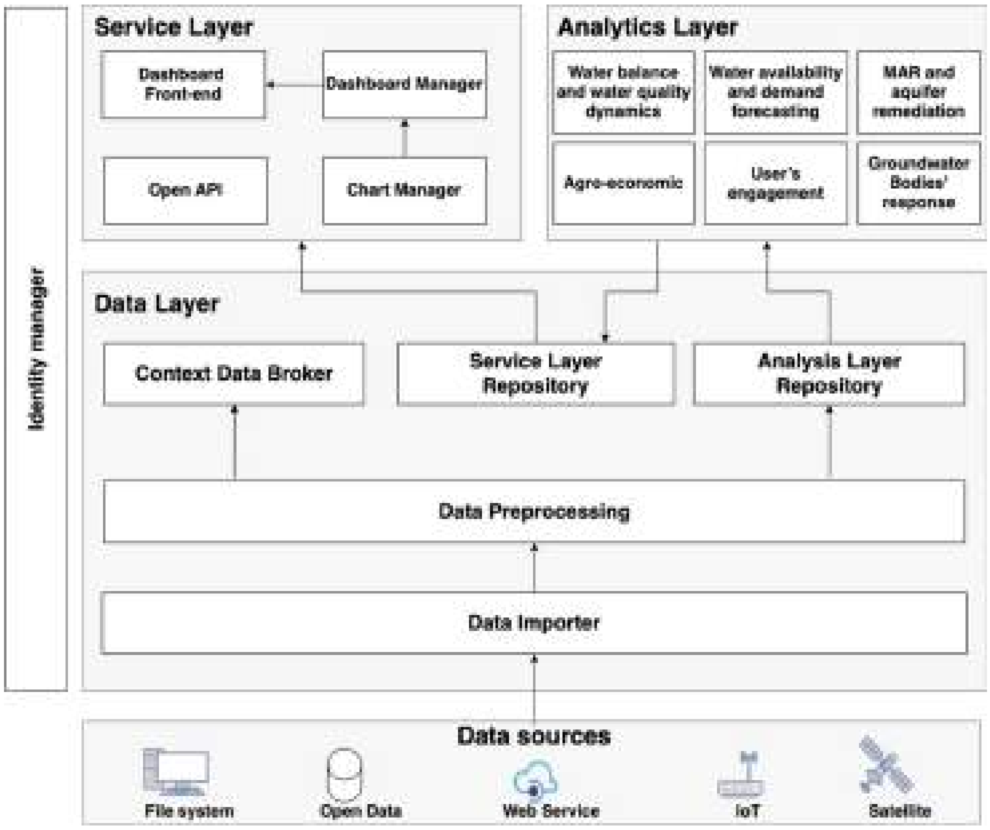

Figure 8 – GTool high level architecture

The image above shows a high level overview of the GTool from the point of view of its main components. It is possible to identify three architectural layers which mainly support the data management process, plus one transversal component dedicated to identity management.

The Data Layer is responsible for collecting, importing and storing the data at any stage of its usage. The very heterogeneous kind of data that are to be used in the GTool (see the data sources depicted in the figure) requires a multi-faceted import component, able to grab different data types and formats and to make them available to the rest of the system. For that purpose the data may need to be pre-processed for cleaning, normalization and harmonization. Data imported from the data sources are stored in the Analysis Layer Repository so that they can be fetched and elaborated by the different analytical modules of the Analytical Layer. Some of the data may be made available through the "Context Data Broker" component, specifically real time data generated by IoT devices. Moreover, some of the imported data may be directly available to the Service Layer for visualization.

The Analytical Layer is the place where actual elaboration of the data is performed in order to generate value added for the decision makers. This layer includes six modules that take as input the data from the Analytical Layer Repository and store their output at the Service Layer Repository. In order to maximize the robustness, interoperability and scalability of the solution, each module has been defined as an independent entity with predefined interfaces for data exchange. Besides, some modules are divided into several entities with clearly differentiated functionalities. This modularization will allow the development of GTool in an environment of continuous integration and improvement, and the application of methodologies such as TDD (Test-driven development) and SCRUM.

The Service Layer is the actual entry point for the GTool end-users (in particular for decision makers). It provides both graphical and programmatic access to the relevant data produced by the Analytical Layer and to data retrieved from other sources The analytical modules store their results in the Service Layer Repository so that charts and dashboards can be created for final users. Such data are both exposed through Open APIs and made available to the Chart Manager component. The latter, in particular, will allow the user to select the appropriate data and generate one of the many charts or map-based visualizations available. Once the charts have been created, they can be put together in a single dashboard through the Dashboard manager component. The dashboard is then made selectively accessible to the final users (i.e. Decision makers, based on their role) through the Dashboard Front-end component.

4.2. Data Layer

4.2.1. Data Importer

4.2.1.1. Description

The Data Importer is the component that allows dashboard admin to import data from external data sources. It can interface with time series or GIS files present in the file system; or third party services, open data portal, or Orion Context Broker. The importer also deals with the import of satellite data (e.g. Copernicus) and the import of data from field sensors. When the data import is successful, the importer publishes the imported data in the Analysis layer repository or make the data available directly to the Service Layer.

4.2.1.2. Interactions

4.2.1.2.1. Import of time series from an open data portal to the Service Layer Repository

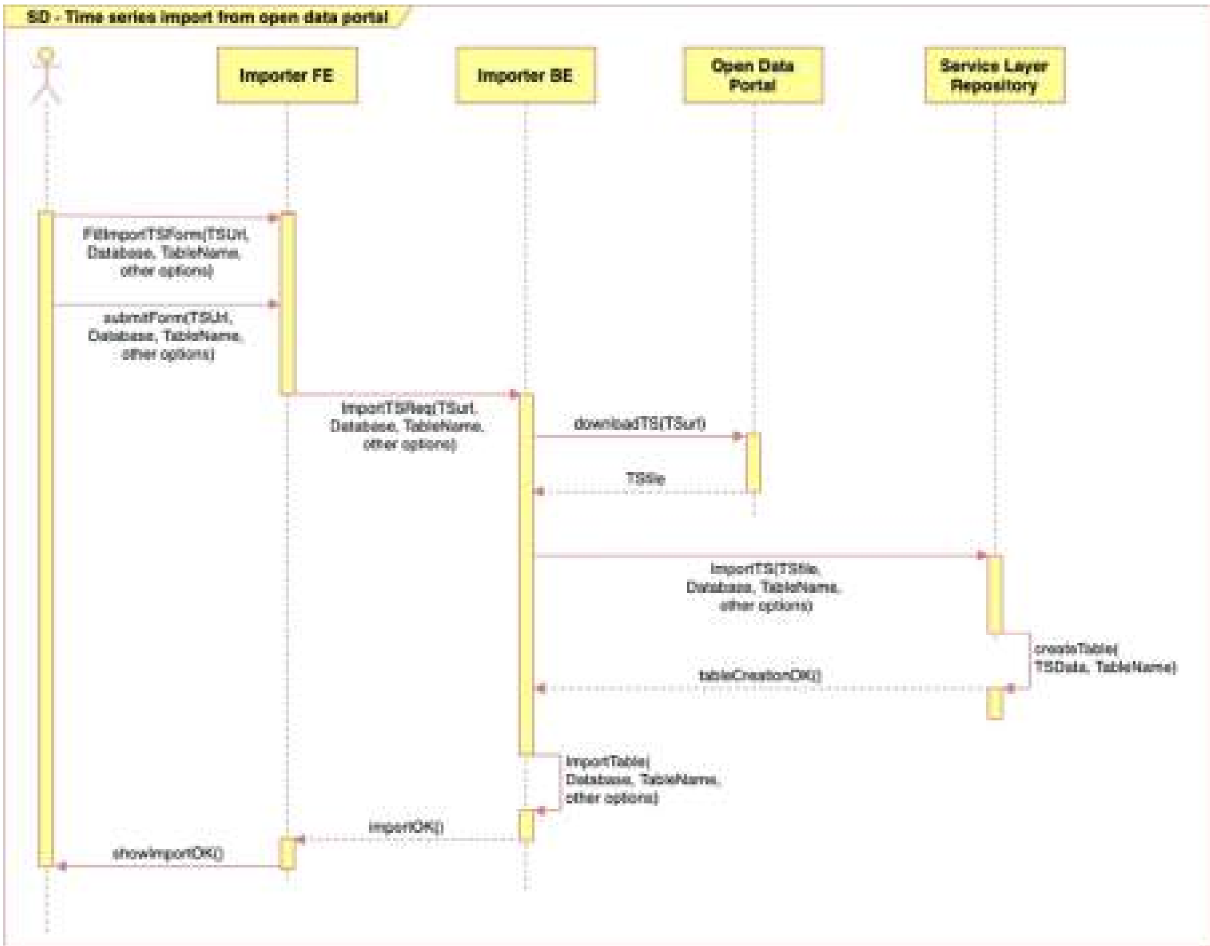

Figure 9 – Import of time series from an open data portal to the Service Layer Repository

4.2.1.2.2. Time series import from local file system to the Service Layer Repository

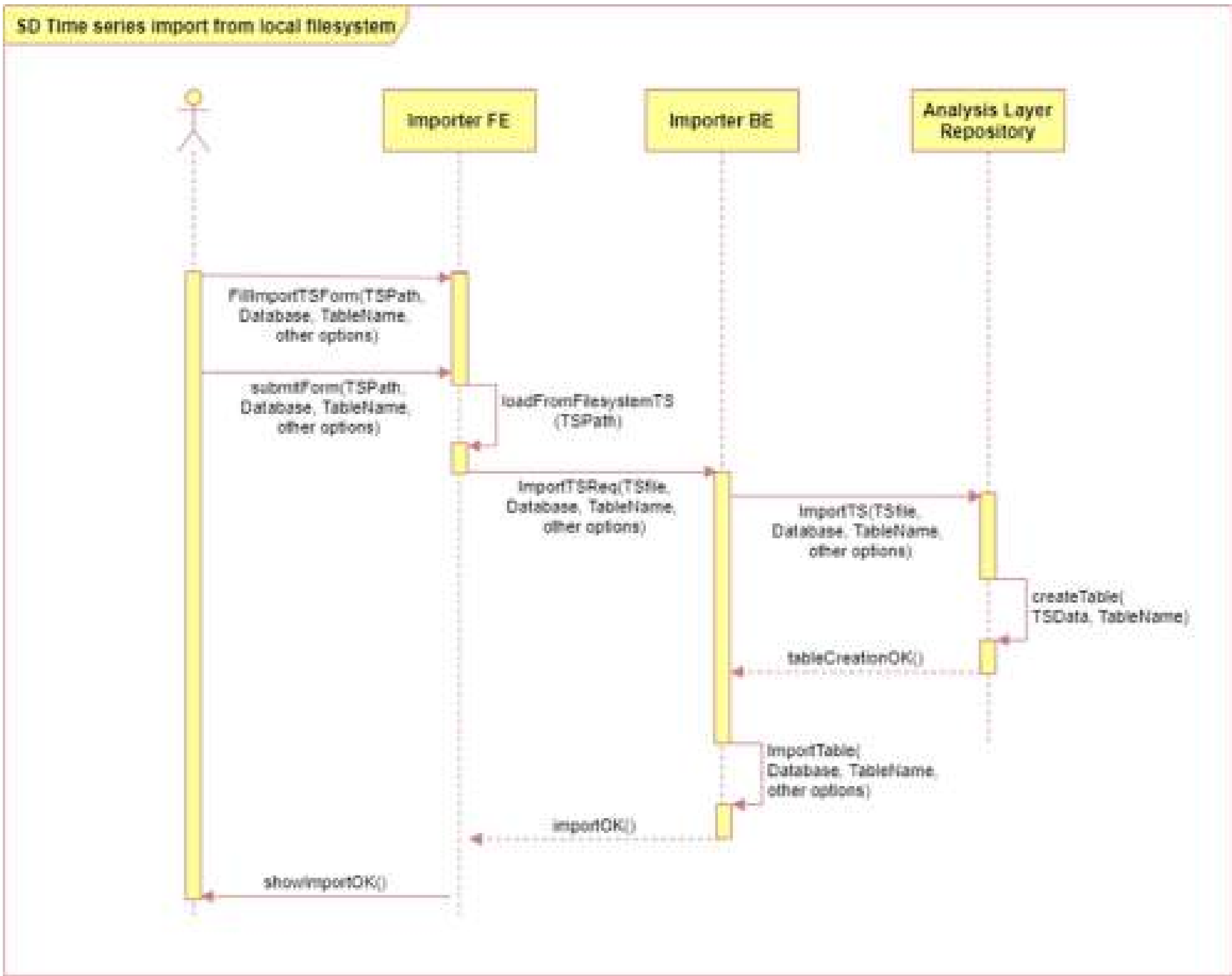

Figure 10 - SD Time series import from local file system to the Service Layer Repository.

4.2.1.2.1. Time series import from local file system to the Analysis Layer Repository

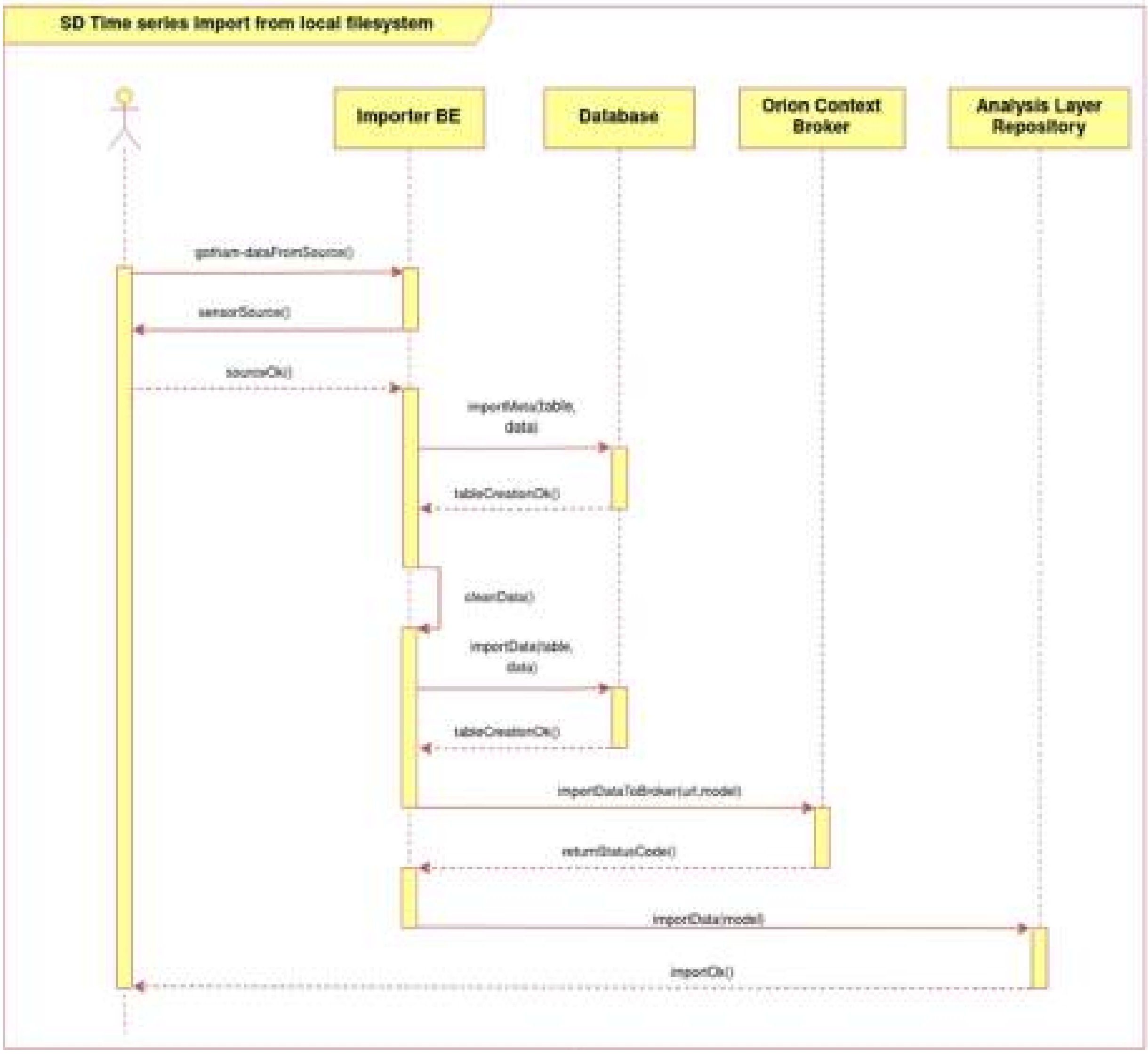

Figure 11 - SD Time series import from local file system to the Analysis Layer Repository

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

4.2.1.2.2. Import of time series from a third-party service to the Service Layer Repository

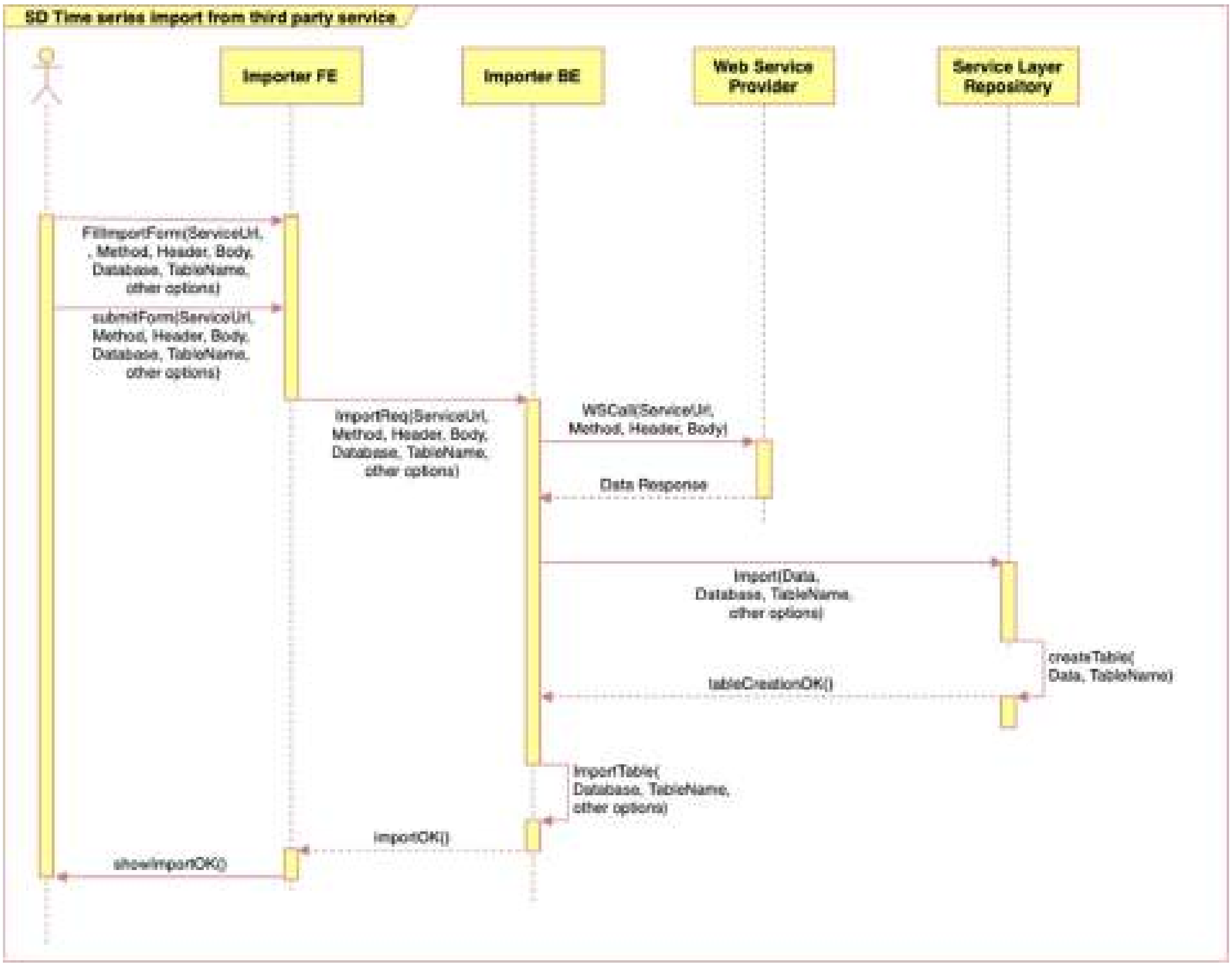

Figure 12 – SD Import of time series from a third-party service to the Service Layer Repository

4.2.1.2.1. Import of time series from a third-party service to the Analysis Layer Repository

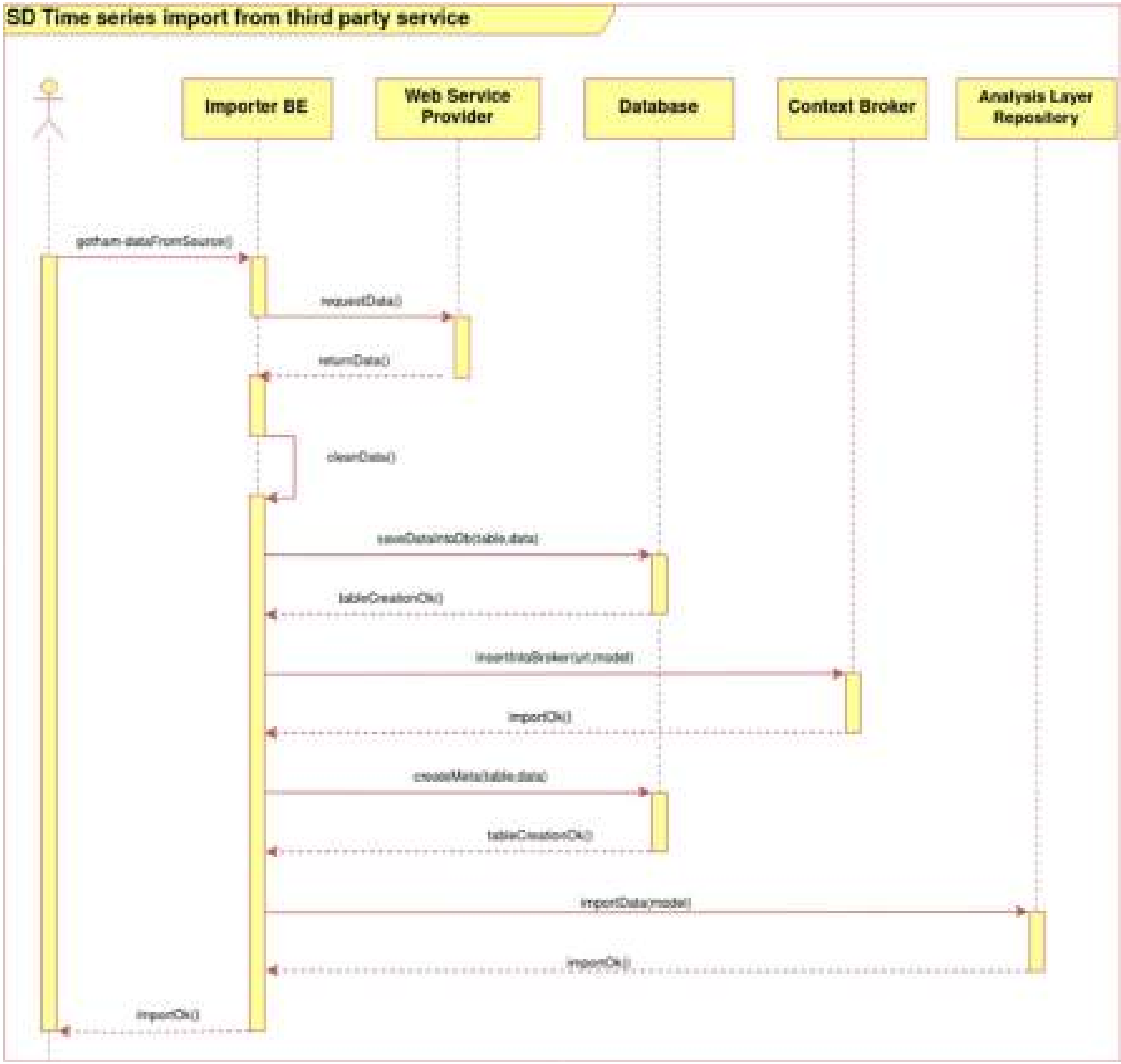

Figure 13 – SD Import of time series from a third-party service to the Analysis Layer Repository

4.2.1.2.2. GIS File import from local file system to the Analysis Layer Repository

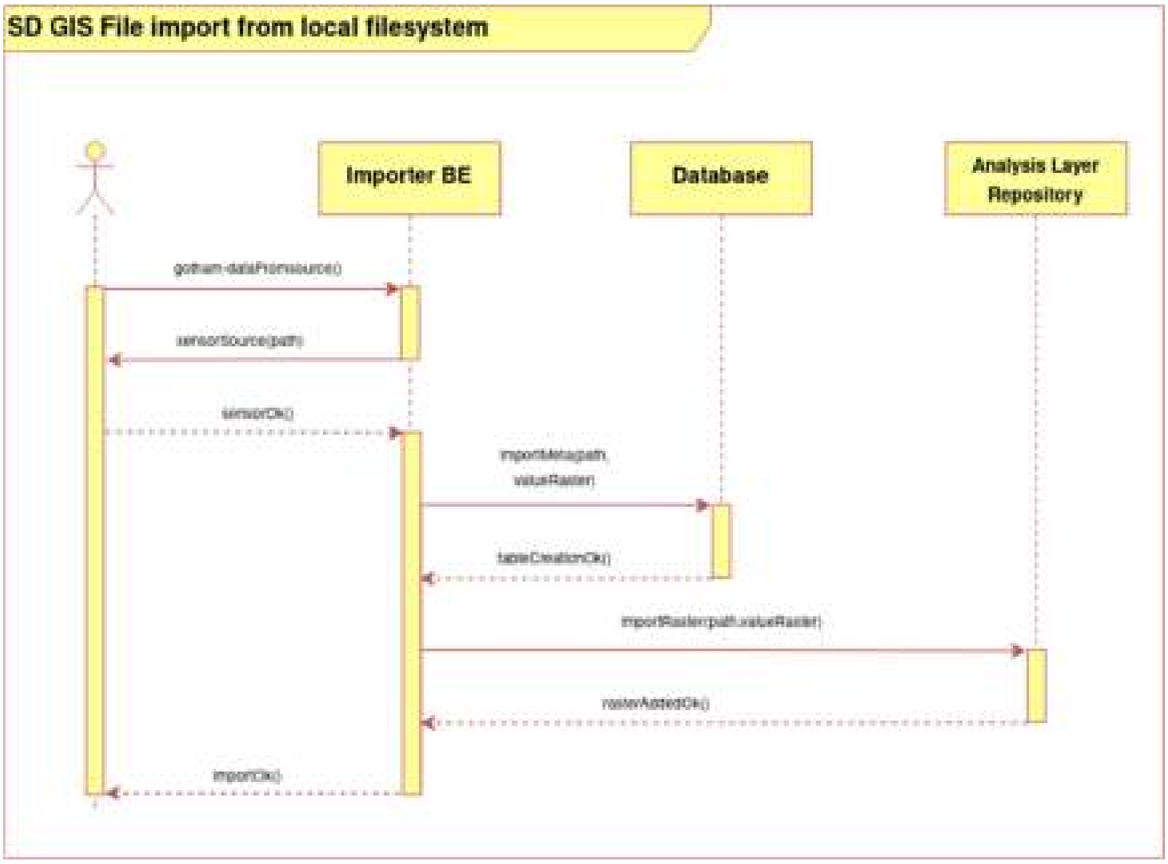

Figure 14 - SD Import of a GIS File from the local file system to the Analysis Layer Repository

4.2.1.2.3. Import of GIS Files from a third party service to the Analysis Layer Repository

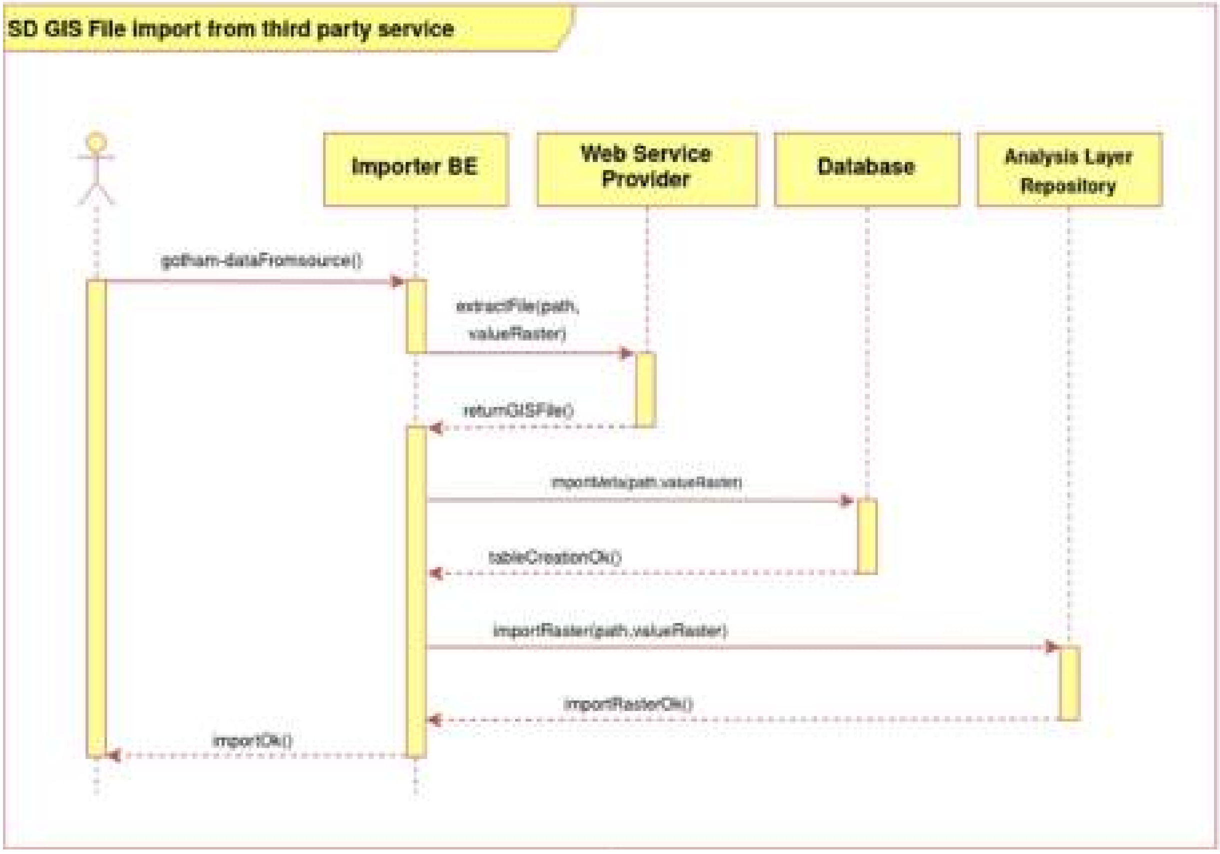

Figure 15 – SD GIF file import from third party services to the Analysis Layer Repository

4.2.2. Analysis Layer Repository

4.2.2.1. Description

The analysis layer repository is made up by the following components:

- *Non-spatial RDBMS*. Used to store time-series.
- *Spatial RDBMS*. Used to store geoJson files.
- *Spatial Fileserver*. Used to store raster files and shape files.
- *Data source manager.* On the top of the previous three components there is the data source manager that is responsible to manage the requests from the Analytics Layer for providing

data sources used to make analyses. The data source manager also enforce the access rights permission to the data sources.

4.2.3. Service Layer Repository

4.2.3.1. Description

The service layer repository is made up by the following components:

- *Non-spatial RDBMS*.
- *Spatial RDBMS*.
- *Spatial Fileserver*.
- *Data source manager.* On the top of the previous three components there is the data source manager that is responsible to manage the requests from the chart manager for providing data sources used to make charts. The data source manager also enforce the access rights permission to the data sources.

4.2.4. Data Pre-Processing

The data pre-processing component, when necessary, will transform data coming from the acquisition stage into a standard and homogeneous format capable of describing all the entities and interactions of the variables assessed.

Data pre-processing will leverage on two main types of resources.

- Pre-processing scripts: scripts aim to extract from downloaded files only the information that is useful for analysis. For text files, those involving time series measurements, this process focuses on removing unusable headers, formatting dates in a correct and accepted format and extracting information from the text written within the file such as the unit of measurement used. In some cases, given measurements, it is necessary to change the units to match the standard used. Null or inaccurate measurements must be removed, and finally all actual data are organised by day and value and entered into the database. For geospatial files, all available meta-information (srs, bbox, height, width) must be extracted from them and then the raster images are loaded as a store and exposed as layers in GeoServer.
- Data models: harmonizing data with respect to commonly accepted data models is an important step when it is necessary to share the data to other stakeholders. We consider FIWARE Data Models and in particular data models managed in the Smart Data Model Initiative[1] to be used for this purpose. In particular, before publishing data to the Data

[1] https://smartdatamodels.org

Context Broker, it will be necessary to transform and harmonized the ingested data into the appropriate NGSI entities.

4.2.5. Context Data Broker

This component will be represented by an implementation of FIWARE's Publish/Subscribe Context Broker Generic Enabler and will provide NGSI interfaces. The component is responsible for aggregating and exposing data from all sensors / devices of interest, using the Publish / Subscribe paradigm. Clients can access this data as NGSI entity attributes representing devices distributed in a territory. For example, a consumer (e.g. web application) can obtain data in different ways:

- By sending requests to the broker;
- Subscribing to data updates that correspond to specific conditions (e.g. change of an entity);

If an IoT device is directly communicating with the Context Broker, consumers can also send commands to devices by updating the attributes to the related entity, provided they have access rights for that operation.

At this moment we do not have evidence of requirements for connecting IoT devices directly to the Context Broker. The Context Broker will be fed from data ingested and pre-processed from the components described above.

4.3. Analytics Layer

4.3.1. Water Balance and Water Quality Dynamics

4.3.1.1. Description

This module makes use of geochemical modelling and data analytics from data loggers and water sampling information to assess the natural attenuation capacity (NAC) of the groundwater bodies, and thus calculate real-time aquifer remediation efficiency. Specifically, a set of remediation indicators will be used depending on the type of aquifer contamination.

This component will be formed by three different submodules, two submodules will be to obtain data for water balance and the other module will be to obtain data on water quality.

GROUNDWATER BALANCE DECISION SUPPORT SYSTEM

This submodule will need an initial set of parameters (aquifer material, aquifer functioning, location, etc.) which the model will process, recommending the most adequate water balance method (chloride mass balance - CMB, soil water balance - SWB, water table fluctuation method - WTFM, etc.) and use it to estimate groundwater recharge and piezometric level. The data could be show in different graphs or generate a file data like a JSON or CSV.

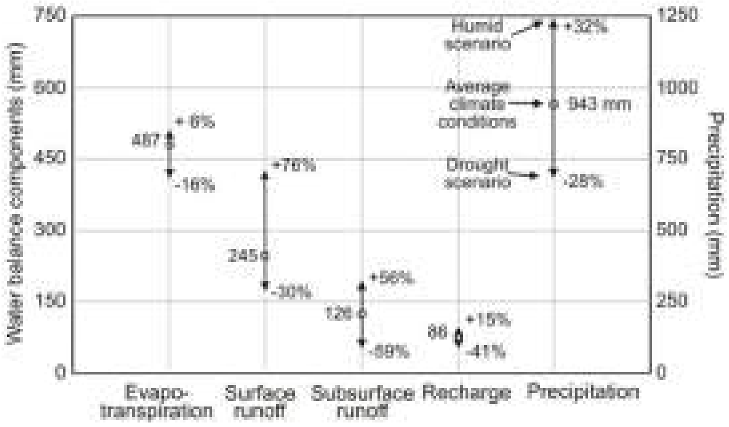

Figure 16. Example of graphical result of the groundwater balance decision support system module.

GROUNDWATER POTENTIAL INDEX AND MAPPING

Using coordinates and geospatial data (rainfall map, textured soil map, DEM analysis, satellite data) this submodule uses geospatial algorithms to overlay and weight GIS layers. These GIS layers will be displayed on a map with which the user could interact.

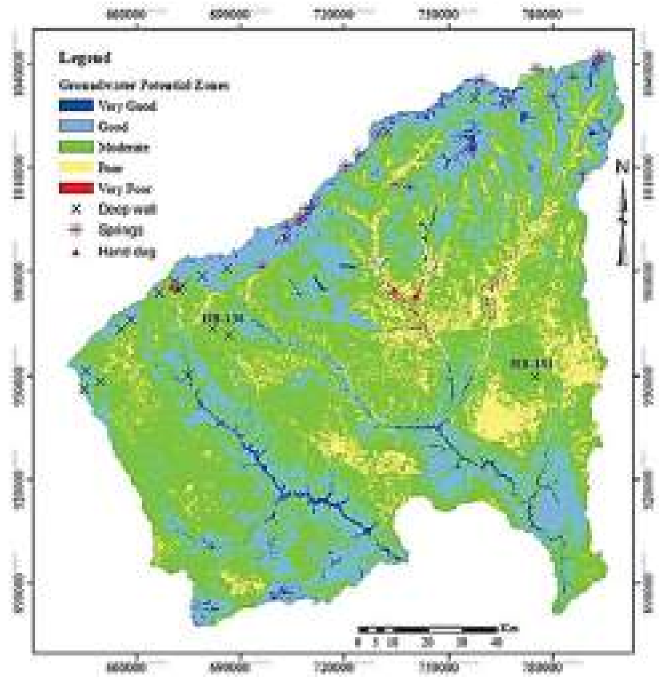

Figure 17. Example of groundwater potential mapping.

STATISTICAL MULTIVARIATE ANALYSIS

In this submodule the user must import data series with water quality data (electrical conductivity, temperature, pH, major ions concentration, etc.). The model will make a statistical multivariate analysis which will identify the main geochemical process occurring in the aquifer and classify the water samples in homogeneous groups. These data will be displayed in different graphs.

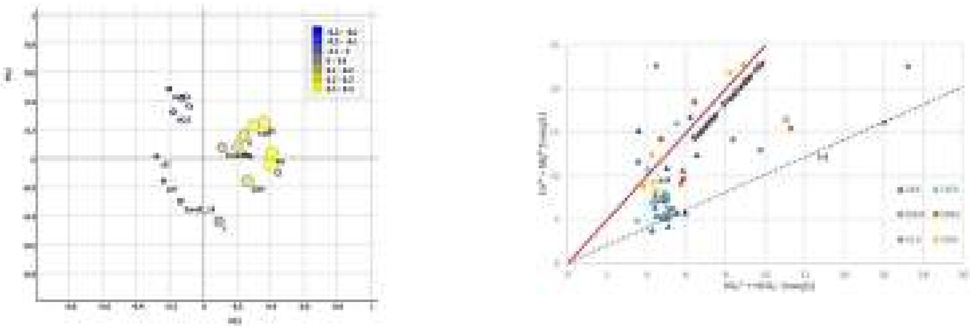

Figure 18. Example of groundwater samples classification based on statistical methods.

4.3.2. Water Availability and Demand Forecasting

4.3.2.1. Description

This module will predict different water scenarios and assess their impact on groundwater quality and quantity status. This module is composed of two submodules: water availability and water demand.

Water availability submodule will be used to forecast the evolution of different time series at short-term. Making use of machine learning-based models it is intended to forecast groundwater recharge, water tables, electrical conductivity or turbidity. Those generated data provided guidance on groundwater trend assessment. The information generated will be used in the component Drought Monitoring and Early Warning System.

DROUGHT MONITORING AND EARLY WARNING SYSTEM

This component uses different time series (streamflow and stored water in reservoirs, groundwater level, rainfall, temperature, evapotranspiration, water demands) as input. These data are used by models, which generate estimated data, which is intended to develop an early warning system, generating a warning when data is generated that does not belong to a specified range. In addition, these data will be shown in graphs.

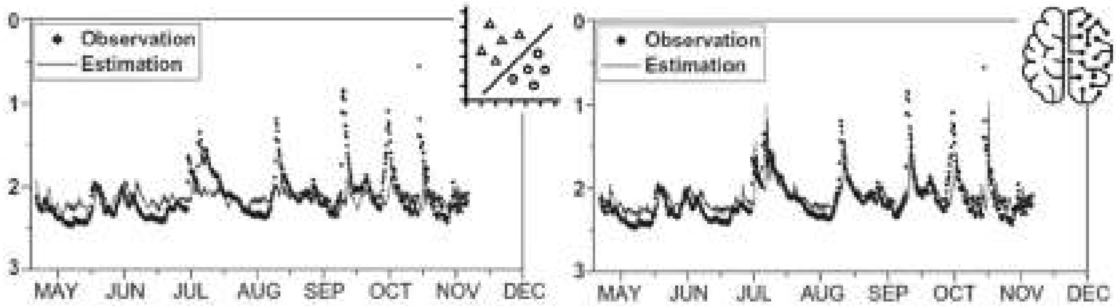

Figure 19. Results of modelling and comparison with observed data.

The water demand sub-module uses remote sensing-based models to estimate water demand for agriculture. A web application will be developed in which data on standard agriculture and greenhouse agriculture will be available.

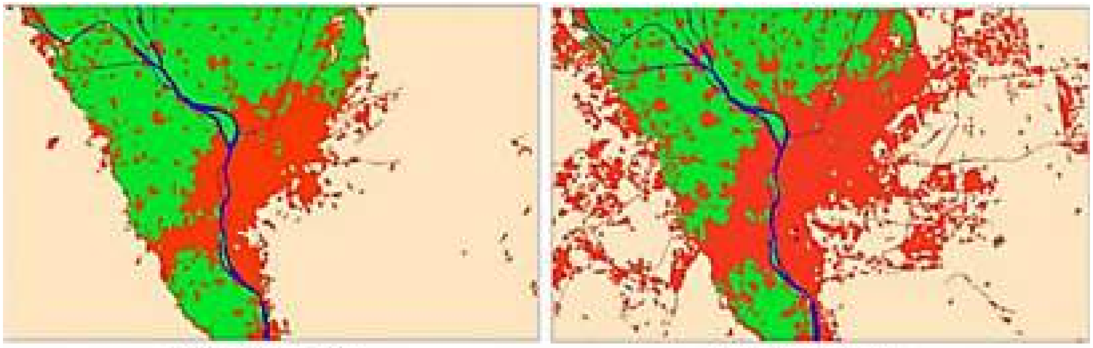

Figure 20. Example of temporal land use changes.

4.3.3. Agro-economic Module

4.3.3.1. Description

This module will allow simulating the effect of different economic instruments, such as water tariff structures, water markets contexts and incentives for water savings (water demand management) and assessing the economic use values and trade-offs between users in alternative resource allocation scenarios.

The agro-economic module will integrate the fields of agronomy (crop water use), engineering (capacity expansion measures), hydrogeology and hydrology to involve different agricultural stakeholder's perspectives and objectives. All possible technical and economic measures such as water reuse, groundwater recharge, agricultural and urban water saving incentives, and economic instruments regarding markets and pricing will be assessed.

This component will contain a submodule in order to optimize water rights exchange.

OPTIMIZATION OF WATER RIGHTS EXCHANGE AT IRRIGATION DISTRICT SCALE

Optimized water rights exchange between both end-water users and public administration by using Blockchain technology will be assessed.

Figure 21. Conceptualization of the application of blockchain technology to the agricultural sector.

The functionality will show the available water coins and water concession for a specific end-user, as well as a graphical evolution of price variation depending on the type of resource (surface water, groundwater, regenerated and desalinated water) and the latest transfer contracts carried out by the different water users. Based on Blockchain principles, each transaction will be stored in a fair, safe and secure way.

4.3.4. Managed Aquifer Recharge and Aquifer Remediation

4.3.4.1. Description

This module will be used for evaluating the feasibility of managed aquifer recharge (MAR) schemes as a result of multicriteria analysis (QGIS environment), including hydrogeological, economic, and chemical (water quality) indicators, as well as regulatory restrictions. The module will contain two different components in order to complete its purpose.

MANAGED AQUIFER RECHARGE (MAR)

It will compare economic framework and water management factors, evaluating the regional status of the aquifer, costs of recharged water and the final use of recovered water, among others. These aspects are crucial for water managers to know the feasibility of MAR under the given constraints.

Secondly, the hydrogeological component will allow evaluating the suitability of the geological formations to support water-recharge activities, taking into account relevant parameters such as specific yield, transmissivity or hydraulic conductivity. A third component will analyse water quality aspects (recharge water vs native groundwater), differentiating two types of parameters:

1. Physicochemical and biological indicators characterizing water quality impairment
2. Chemical components liable to modify groundwater equilibrium.

This MAR feasibility tool will include the possibility to simulate geochemical reactions within the aquifer (PhreeqXcel), as a result of different water recharge volumes and water qualities.

AQUIFER REMEDIATION

Using economic (CAPEX and OPEX), technical and environmental constraints for several groundwater remediation alternatives (bio-remediation, in situ denitrification, air sparging, in-well air stripping, pump & treat methodologies, chemical oxidation or thermal treatment) will be evaluated, as well as their potential effect on the aquifer in terms of water quality. The remedial design decision-making will be developed by combining elements from Latin-Hypercube simulation of contaminant transport, economic risk-cost analysis, and regional sensitivity analysis.

4.3.5. Groundwater Bodies' Response

4.3.5.1. Description

The effects of different water strategies and water governance policies in the status of groundwater bodies will be assessed. As a combination of groundwater availability and demand, this module will be able to evaluate the pros and cons of several scenarios and establish the most likely temporal horizon to achieve the good status of the groundwater body.

4.3.5.2. Interactions

The user should introduce some variables, then a series of some machine learning algorithms will produce groundwater indexes and finally, by the evaluation of these indexes, the model will return the groundwater body status.

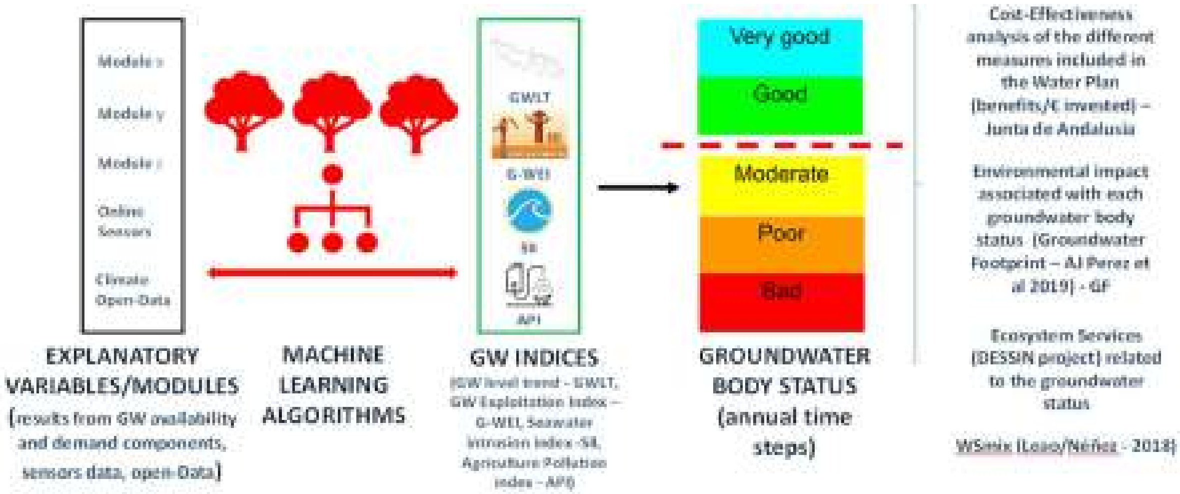

Figure 22. Conceptualization of the groundwater bodies' response approach to be applied in GOTHAM project.

4.3.6. User's Engagement

4.3.6.1. Description

This module will be used to ease the definition of water priorities (water boundary conditions) by water users, taking into consideration water resources to meet water demands. The best water allocation scenario will not only need to take into account technical criteria such as the volume of water resources available, future demands, expected crop prices or MAR impact, but it will also have to incorporate boundary conditions of the water system since they can significantly determine the feasible optimal water resources distribution. The boundary conditions will relate to:

1. Water network and infrastructure: existing and potential water connections.
2. Water quality thresholds: they will depend on final water use.
3. Economic constraints: maximum water price the user is willing to pay.
4. Volume of water needed: a water source might meet the quality and price requirements but be insufficient.
5. Temporal needs: demand usually varies along the hydrological year, which might compromise water allocation for all users in specific moments.

4.4. Service Layer

4.4.1. Dashboard Manager

4.4.1.1. Description

The Dashboard Manager is responsible for the creation and management of dashboards that include the charts created through the chart manager and other graphic and text components.

It is also responsible for dashboard publication to a group of decision makers identified by a specific role.

This component is made up by two subcomponents:

- *Dashboard Manager BE*. Contains the business logic of the Dashboard Manager.
- *Dashboard Manager FE*. The frontend of the Dashboard Manager.

4.4.1.2. Interactions

In the following figure can be seen the main interaction of the dashboard manager with user (dashboard administrator) and other architectural components.

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

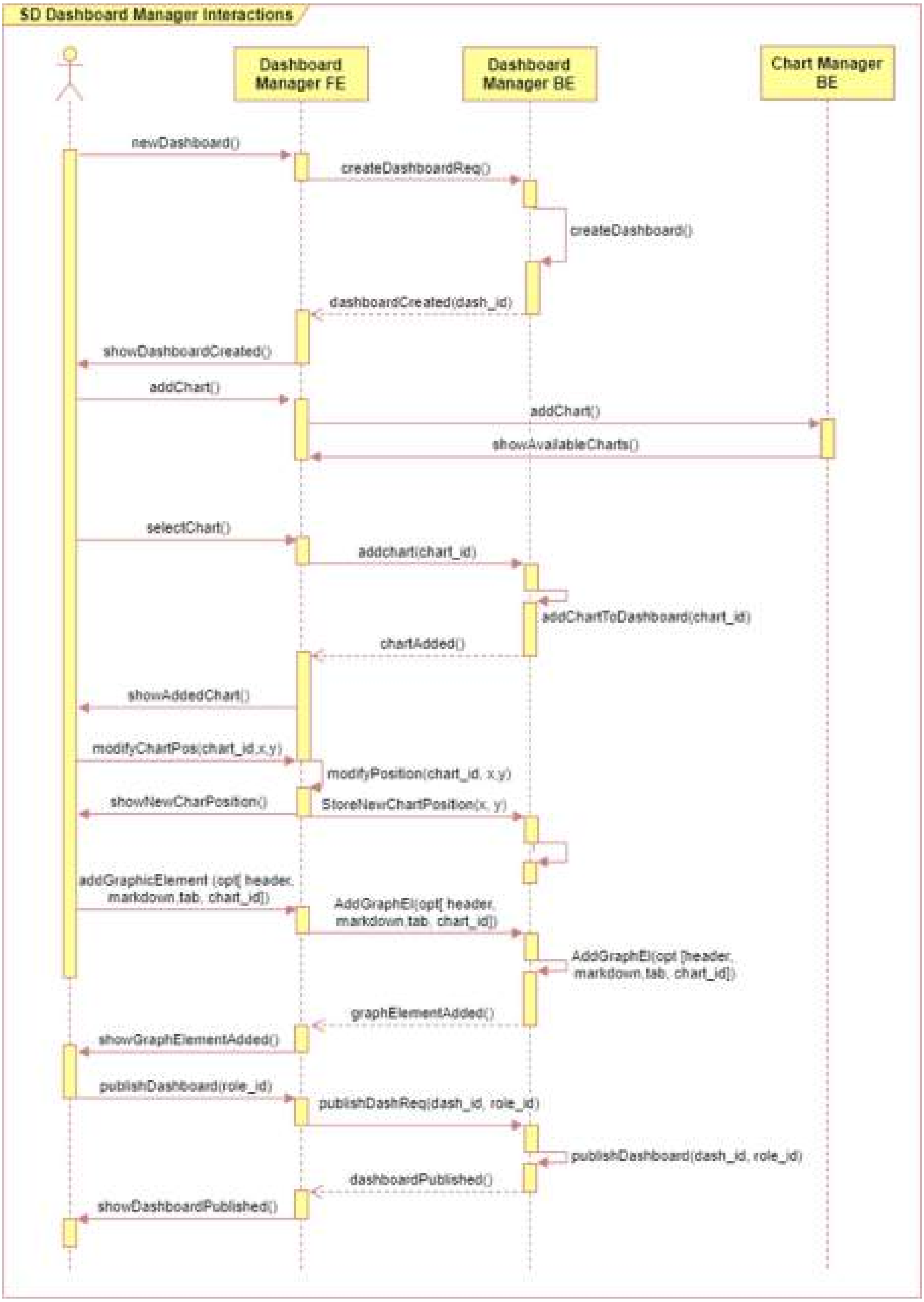

Figure 23 - Sequence Diagram: Dashboard Manager Interactions

4.4.2. Chart Manager

4.4.2.1. Description

Chart Manager is the component that deals with the creation and management of charts. It can support different types of charts (see functional requirements).

This component is made up by two subcomponents:

- *Chart Manager BE*. Contains the business logic of the Chart Manager.
- *Chart Manager FE*. The frontend of the Chart Manager.

4.4.2.2. Interactions

In the following figure can be seen the main interaction of the chart manager with users (dashboard administrator) and other architectural components.

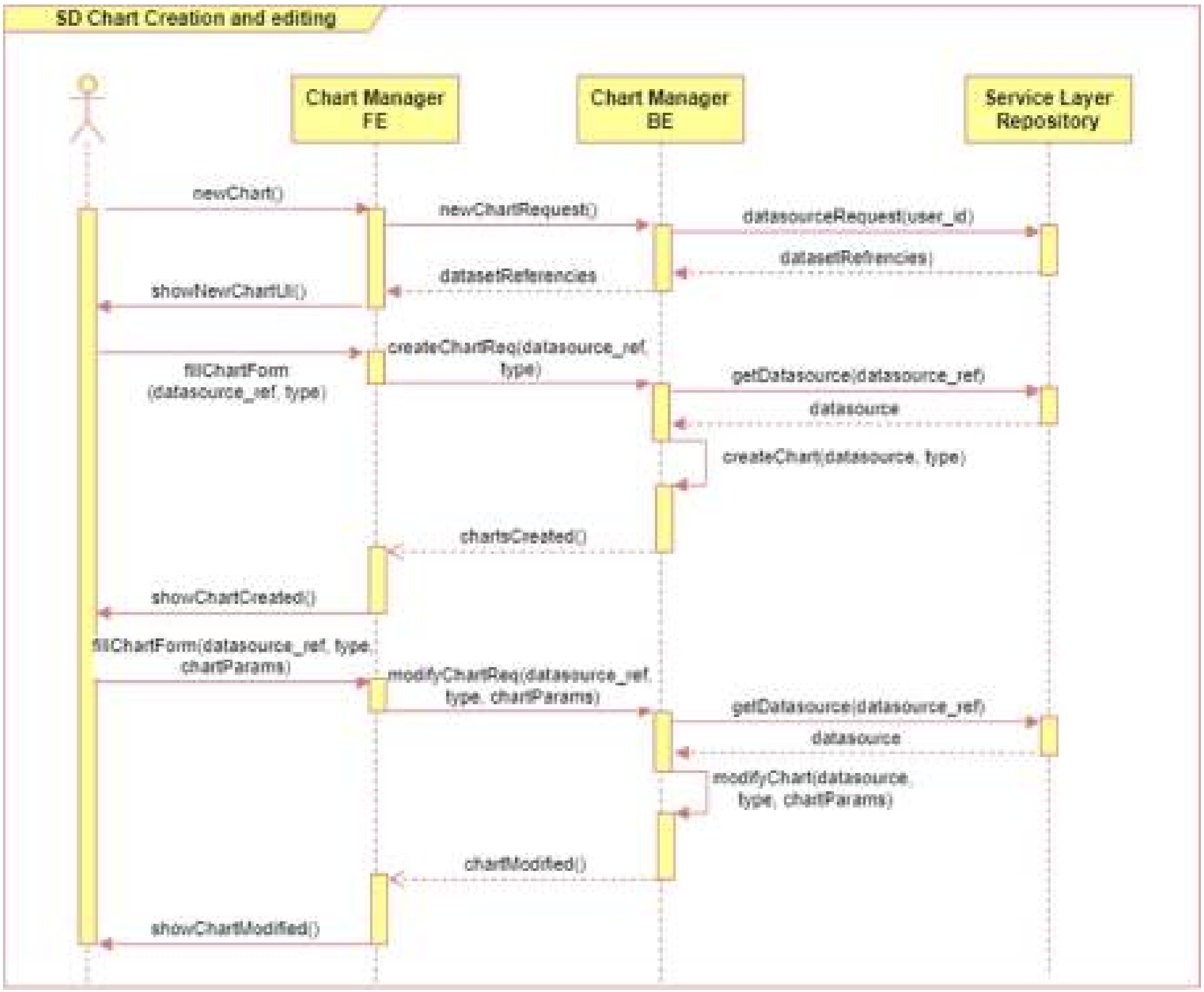

Figure 24 – SD Chart Creation and Editing: Interaction among components

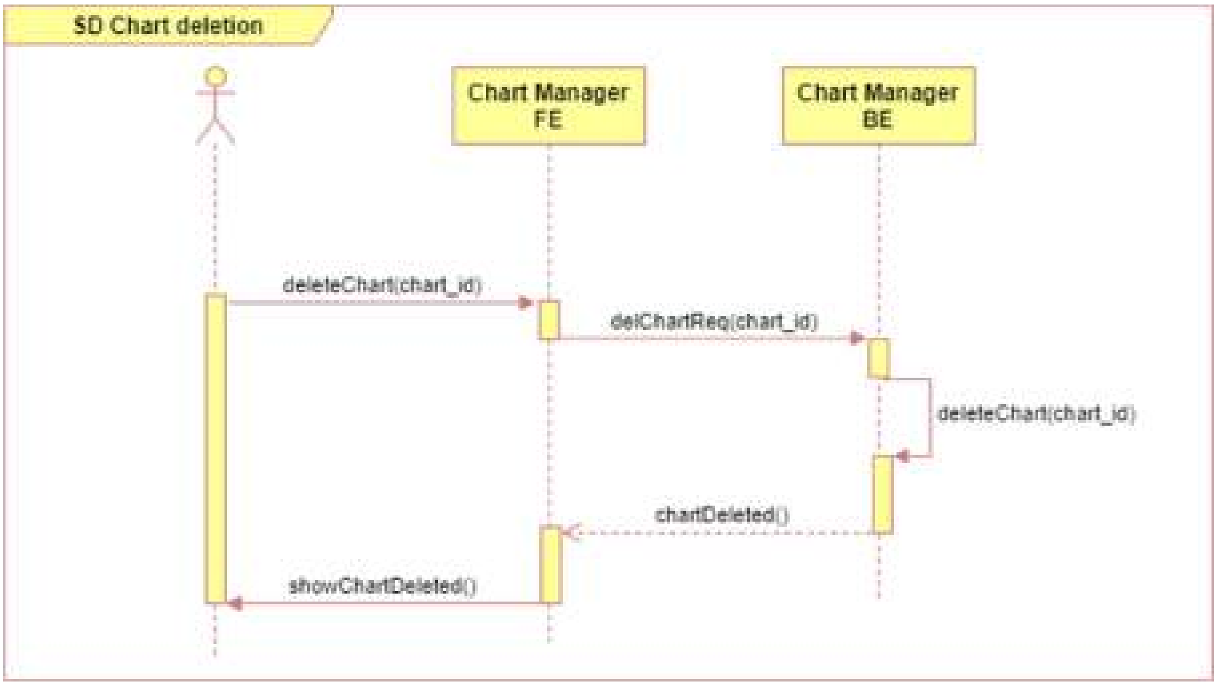

Figure 25 – SD Chart deletion : interaction among components

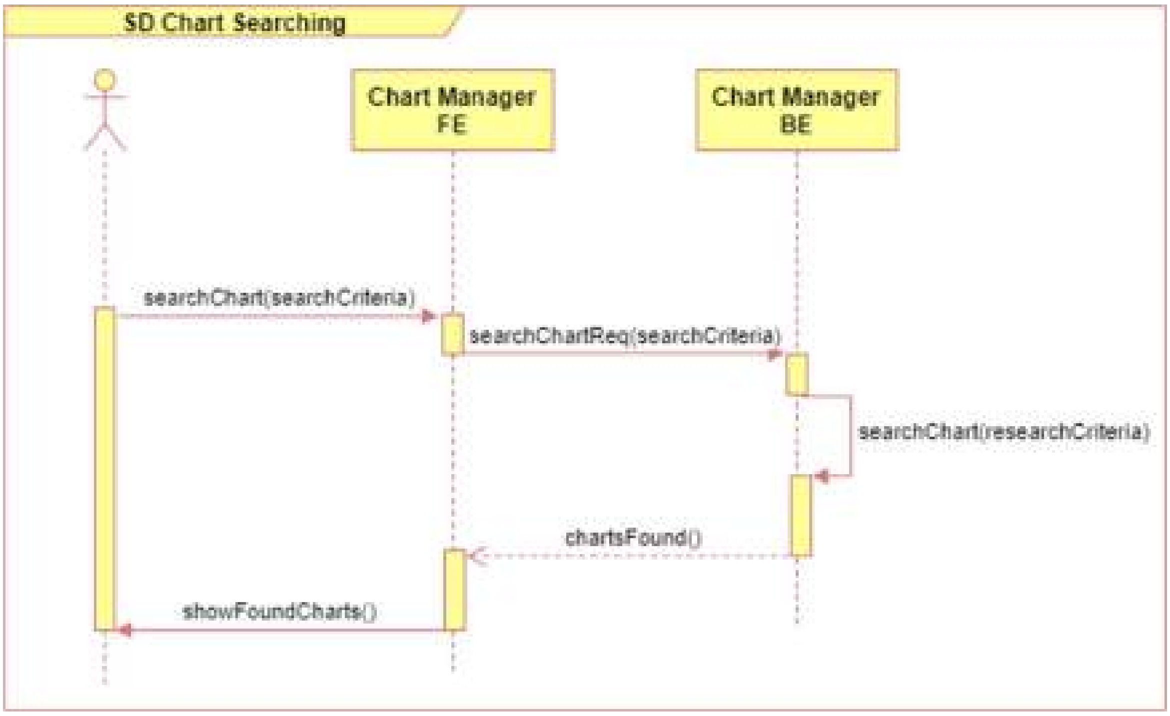

Figure 26-SD Chart Searching. Interaction among components.

4.4.3. Dashboard Frontend

4.4.3.1. Description

This component is used by decision makers in order to view/navigate dashboard. In particular it displays the dashboard frontend interface with the list of dashboards belonging to the user. And the user can View/navigate a specific dashboard; Search for a Dashboard specifying several search criteria; View the details of a dashboard (metadata).

4.4.4. Open API

4.4.4.1. Description

This component will allow third parties to access data and specific functionalities provided by the GTool and to use it to develop applications and services. APIs will expose results of the analysis processes and, if deemed appropriate, specific data directly derived from the data sources (e.g. sensors measurements).

5. Available data sources for the Data layer

5.1. Time series

Time series are typically reported as tabular files in the form of comma separated values (CSV) or another textual format. For the purpose of the GTool we collected a preliminary list of data sources that may be used by the different Analytical Modules.

5.1.1. SAIH HIDROSUR

SAIH HIDROSUR[2] (*Sistema Automático de Información Hidrológica*) is a Spanish regional online database that publishes data generated by a network of remote stations strategically distributed to obtain, in real time, information on the hydrometeorological incidents that occur in each of them, in order to be able to take the appropriate measures, both in anticipation of floods in order to prevent and minimize their damage, and in exploitation of water resources. Measurements include hydrometeorological parameters such as water levels, flows, rainfall, snow levels, gate position sensors, storm detectors, etc. Data can be filtered with respect to different parameters, including dates, region, station and sensor. For each sensor a table is displayed with the measurements performed every hour in the time interval selected by the user. The data can be exported as CSV file.

[2] http://www.redhidrosurmedioambiente.es/saih/datos/a/la/carta

GOTHAM is part of the PRIMA Programme supported by the European Union. The PRIMA programme is supported under Horizon 2020 the European Union's Framework Programme for Research and Innovation. Grant Agreement number: 1922

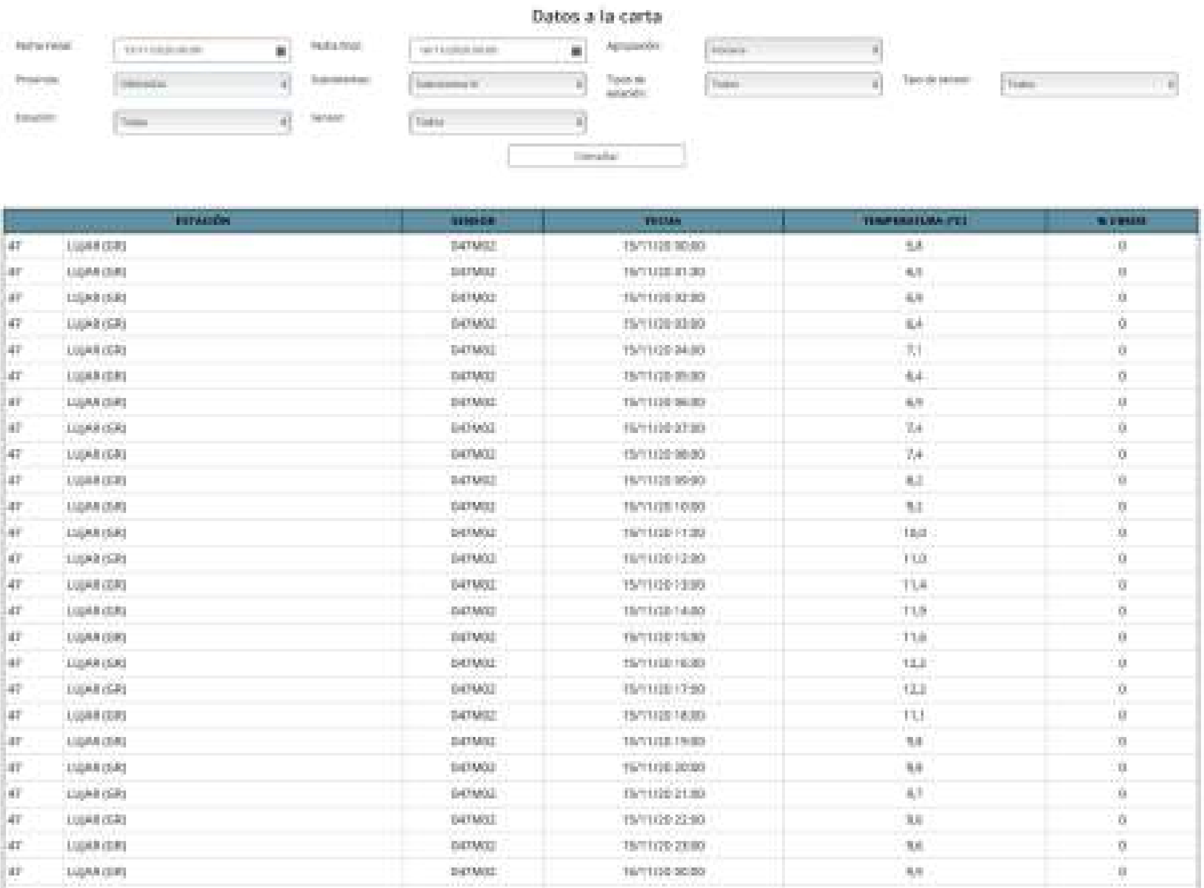

Figure 27- The user interface of the SAIH HIDROSUR database

This data source can be used for feeding the Water Balance and Water Quality Dynamics Module.

5.1.2. EUROPEAN CLIMATE ASSESSMENT & DATASET

The European Climate Assessment & Dataset[3] (ECA&D) project provides information on changes in weather and climate extremes, publishing daily dataset needed to monitor and analyse these extremes. The data can be filtered with respect to country, location and element for which you wish to examine or process a dataset. The list of elements can be seen in the following figure. The measurements can be downloaded in a .txt format with included comma separated values.

3 https://www.ecad.eu

Figure 28- The user interface of the ECA&D database

This data source can be used for feeding the Water Balance and Water Quality Dynamics Module.

5.2. Remote sensing and regional GIS databases

Free and open databases covering the three pilot sites of the project have been foreseen for the data gathering phase:

- European Union Copernicus platform
- SPEI Global Drought Monitor
- Earth Engine Evapotranspiration Flux (EEFlux)
- SENTINEL Hub
- Google Earth Engine

Copernicus is the European Union's Earth Observation Programme. It provides information services based on Earth observation satellite and in-situ (non-space) data. The information services provided are free and open to its users. Copernicus has its own set of satellites (the Sentinel families), complemented by participating missions (other commercial or public satellites in orbit). Copernicus also collects information from in-situ systems, such as ground stations, which provide data obtained by a multitude of terrestrial, marine and airborne sensors. Services provided by Copernicus include Land Monitoring, Emergency Management, Atmosphere Monitoring, Maritime Environment Monitoring, Climate Change and Security. Data from Copernicus will be used for developing the Water Balance and Water Quality Dynamics Module and the Water Availability and Demand Forecasting Module.

The Standardised Precipitation-Evapotranspiration Index (SPEI) is a multiscale drought index based on climatic data. It can be used for determining the onset, duration and magnitude of drought conditions with respect to normal conditions in a variety of natural and managed systems such as crops, ecosystems, rivers, water resources, etc. The SPEI fulfils can measure drought severity according to its intensity and duration, and can identify the onset and end of drought episodes. One of the services derived from the SPEI is the SPEI Global Drought Monitor. The latter offers near real-time information about drought conditions at the global scale, with a 1 degree spatial resolution and a monthly time resolution. SPEI time-scales between 1 and 48 months are provided. The calibration period for the SPEI is January 1950 to December 2010. The starting date of the dataset is 1955 in order to provide common information across the different SPEI time-scales. Data from the SPEI Global Drought Monitor will be used for developing the Water Availability and Demand Forecasting Module.

Earth Engine Evapotranspiration Flux, also known as EEFlux, is based on the operational surface energy balance model "METRIC" (Mapping ET at high Resolution with Internalized Calibration), and is a Landsat-image-based process. Landsat imagery supports the production of ET maps at resolutions of 30 m, which is the scale of many human-impacted and human-interest activities including agricultural fields, forest clearcuts and vegetation systems along streams. ET over extended time periods provides valuable information regarding impacts of water consumption on Earth resources and on humans. Reference ET is calculated using the ASCE Penman-Monteith and GridMET weather data sets. EEFlux will be freely available to the public and includes a web-based operating console. EEFlux will be used in the Water Availability and Demand Forecasting Module.

Sentinel Hub is an engine for processing of petabytes of satellite data. It is opening the doors for machine learning and helping hundreds of application developers worldwide. It makes Sentinel, Landsat, and other Earth observation imagery easily accessible for browsing, visualization and analysis. It lets the user scale the system globally with an intuitive and user-friendly interface, without any hassle. The satellite imagery distribution service is based on a cloud GIS web-based platform which is suitable for large scale spatial deployments. Sentinel Hub uses data from the following satellites and services: Sentinel-1, Sentinel-2, Sentinel-3, Sentinel-5P, Landsat, Planet Labs, Envisat Meris, DEM (digital elevation models), MODIS and other commercial data. Data from Sentinel Hub will be used to develop Water Availability and Demand Forecasting Module.

Google Earth Engine is a platform for scientific analysis and visualization of geospatial datasets. It hosts satellite imagery and stores it in a public data archive that includes historical earth images going back more than forty years. The images, ingested on a daily basis, are then made available for global-scale data mining. Earth Engine also incorporates APIs and other tools to enable the analysis of large datasets, providing easy, web-based access to an extensive catalogue of satellite imagery and other geospatial data in an analysis-ready format. Services offered by Google Earth Engine include Earth surface temperature, climate models, atmospheric data, weather datasets including precipitation, temperature, humidity, and wind, high-resolution imagery, digital elevation models, land cover and cropland data, among other. Google Earth Engine will be used in the Water Availability and Demand Forecasting Module.

5.3. IoT Devices

Two in situ measuring units with incorporated telemetry system have been deployed in the Spanish case study. These sensors are capable of measuring the groundwater electrical conductivity, temperature and depth every hour, and to upload these data to a server. Users can access this server through an interface called vanwaltCONNECT which let them view, manipulate and download the data from the onsite sensors as well as other functions such as setting alarms or assessed the performance of the batteries, for example. Data is downloaded as CSV files.

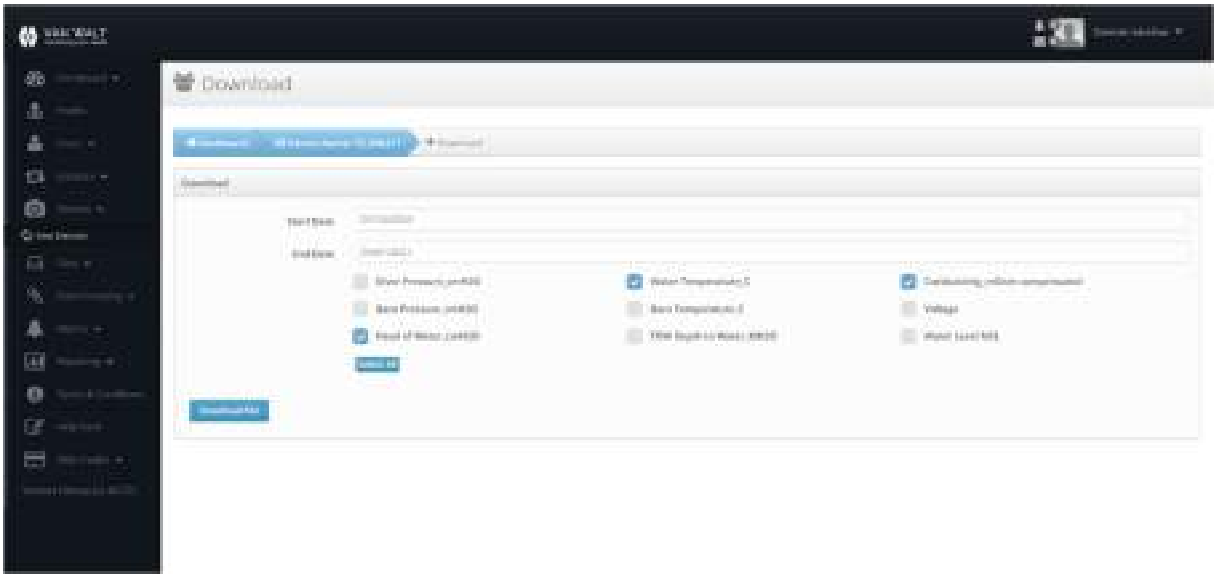

Figure 29- The vanwaltCONNECT user interface

This data source can be used for feeding the Water Balance and Water Quality Dynamics Module.

6. Conclusion

The aim of this report is to define the initial functional requirements and the reference architecture of the GTool in order to define the functional building blocks of the software to be developed during the project. The main users have been identified and their use cases defined according to the functional requirements. On the one side technically skilled people who aim at collecting data and performing analysis on it with advanced tools and algorithms. On the other side end users, which have been identified already at proposal time as water producers, water suppliers, water end-users and water regulators. They are not interested in advanced analytical tools but in accessing their results and up-to-date data and information.

Based on such use cases and on the preliminary architectural layers recognised at proposal time, the overall reference architecture to be implemented in the rest of the project has been designed, including its main components and their interaction.

This report will feed software development and integration tasks in WP4, specifically: T4.2 Data Layer, T4.3 Analytics Layer, T4.4 Service Layer, T4.5 GTool integration and deployment for the Spanish case study.

It is to be noted that some of the functionalities and use cases described in this report may be updated if new requirements emerge from the pilot cases and from the stakeholders.