RiverCure Portal: Collaborative GeoPortal for Curatorship of Digital Resources in the Water Management Domain

Marta Gonçalves Barcia Gonzalez
Instituto Superior Técnico (IST), Universidade de Lisboa, Portugal
marta.gonzalez@tecnico.ulisboa.pt

***Abstract* — Nowadays, there is a huge quantity of information available regarding water resources, although with different structures, formats, languages, vocabulary, which lead to a lack of consistency, accuracy and increased the time spent modifying such information. On the other hand, flood events are relatively frequent. The previously stated problems underline the importance of a well-designed water resources information system. Without solving them, the knowledge that several sources have, cannot be linked or shared properly. This inhibits the development of accurate monitoring and alert systems.**

To explore these problems, the Design Science Research Methodology was applied. The first two steps focused on the identification of the problem, motivation and objectives of this work. In the third step, the system was designed and developed. Afterwards, the system was demonstrated using a case study, the flood event in Águeda in 2016.

The RiverCure portal is a web-based system, that will allow its users to simulate flood events and to access information regarding water resources captured by sensors from distinct sources. This system intends to improve flood forecasting and the management of water resources and supporting decision-makers to make better decisions.

This research intends to preliminary discuss and define the requirements specification of the RiverCure portal using a rigorous language, namely the Requirements Specification Language. To achieve this goal, it was necessary to specify the systems' requirements of the main inputs of the portal, namely the Portuguese monitoring and surveillance system and the simulation software since these systems are important features of the portal.

Keywords: RiverCure, Water Resources, Information Management, Requirements Engineering

1. Introduction

Flood events have been increasing in Portugal, for instance, between 2011 and 2018 the Agência Portuguesa do Ambiente (APA) reported 306 flood events and the number one cause identified was strong precipitation [1]. The fact that several basins, namely the Tagus and Douro river, cross Spain's border hinders the assessment of the flood risk and the characterization of these risks. Another factor that must be taken into consideration is climate change, which again according to APA can have a significative impact since the forecast for Portugal is that in general, the annual mean precipitation will decrease, but the number of short-period precipitation events is expected to increase and with more intensity compared with current days [1].

Nowadays, with the evolution of technology, the size of information collected, discovered and available has been increasing rapidly [2]. New problems have thus emerged with it, such as data not properly documented, difficult access to data, due to ownership or cost and inconsistent data between different sources, making the use of it more difficult [3]. To overcome these problems, data should be rigorously defined. In the current research different languages were analyzed, so that concepts could be accurately defined and connected. Taking that into account, the Unified Modelling Language (UML) [4], the Requirements Specification Language (RSL) language [5] and consequently the ITLingo project were researched, so that the requirements specification of the systems linked to the project could be written in the best possible way. Overall, RSL includes a process to transform requirements written in a natural language into formal structured requirements [5, 6].

The main problems and challenges regarding the subject are the difficulty in managing the water resources data, the existence of several stakeholders, and the presence of some guidelines to standardize water-related information. All these issues are some of the reasons why forecasting natural events, such as floods have been a challenge. To solve some of these issues the RiverCure project was proposed.

The RiverCure project has the objective of improving the forecasting systems focusing on the impact of floods, to manage water resources and habitat protection. In order to achieve that it is necessary to assimilate data from official agencies (e.g. APA) and acquire new data that is shared on the internet, such as images of such events shared in Instagram or Flickr, among other social media platforms, to help tune the forecasting capabilities of the mathematical models. High performance version of Strong Transients in Alluvial Valleys (HiSTAV) is the simulation tool applied, to check the consequences inland of a flood [4]. The three main outputs expected from the RiverCure project are (1) an IT platform, named RiverCure portal (RCP), that articulates model forecasts and data from authorities (e.g. APA, Instituto Português do Mar e da Atmosfera (IPMA)) and crowdsourcing from social networks (e.g. Flickr); (2) a new-generation river model; and (3) a curated database that combines authoritative and crowdsourced data. The project is funded through a Fundação Ciência e Tecnologia (FCT) grant. There are three main partners involved in the project, Information and

Decision Support Systems' laboratory of the Instituto de Engenharia de Sistemas e Computadores - Investigação e Desenvolvimento (INESC-ID), IST-ID (Association of Instituto Superior Técnico for Research and Development) and APA.

This dissertation is linked to the RiverCure project and aims to:

• Define the as-is model of the SNIRH, SVARH (Sistema de Vigilância e Alerta de Recursos Hídricos) and HiSTAV (High performance version of Strong Transients in Alluvial Valleys);

• Define a requirements specification system for the RiverCure portal.

To achieve the objectives defined, the Design Science Research Methodology [7] was applied, due to the existence of several guidelines, stakeholders and information systems. The first step was the identification of the problem and the motivation of this work, which was followed by the second step, setting the objectives. The third step that was focused on the design and development of the system, for that to happen, knowledge regarding the system had to be acquired. Upon the system has been developed had to be demonstrated, for that a case study was used and then tested so that the conditions could be checked (in case they were not fulfilled the system had to be designed and developed again, returning this way to the third step). The flood event at 12 of February 2016 in Águeda, Aveiro, Portugal was the chosen case study. Since it reached historical water levels according to the data obtained from the City Hall and the Sistema Nacional de Informação de Recursos Hídricos (SNIRH), which is owned by APA. The last step is to write a research paper, for the community to be aware of the developments made regarding water resources management.

2. Background

Requirements models, such as those defined with UML, are commonly used to produce conceptual models, that contain the requirements of a system [8]. These models allow a much better understanding of the information compared with descriptions written with just natural languages.

Unified Modeling Language (UML) is a language that allows specification, construction, visualization and documentation of objects in an information system. It was created in 1996 after an attempt to unify the most significant object-oriented modelling languages at the time, namely the Boosh, the OMT (Object Modelling Language) and the OOSE (Object Oriented Software Engineering). One of the advantages of this language is that it has extension mechanisms which allow future modifications. Other benefits include the fact that it can be applied to a lot of different domains, merges many distinct diagrams from different languages, and is independent in terms of process and of modulation tools [9]. The most recent version is the UML 2.5.1 and it was launched in December 2017. The model's elements can be categorized in three groups, classifiers, which describes a set of objects, events, that describes a set of situations which will cause consequences to the system and behaviour, which describes a set of possible actions that can lead to the occurrence of events [10]. In the UML 2 there are diagrams for behavioural modelling (e.g. use cases, activity) and for structural modelling (e.g. class, packages).

ITLingo is a long term initiative aiming to research, develop and apply rigorous specification languages in the Information Technology area, namely in technical documentation. It focusses on three subjects, Requirements Engineering, with the RSL (Requirements Specification Language); Testing Engineering, with the TSL (Testing Specification Language); and Project Management, with the PSL (Projects Specification Language [5].

RSLingo RSL is a Systems Requirements Specification (SRS) language. This means that RSLingo is an organized method to collect the requirements of a system. RSL is supported by ITLingo-Studio, i.e. an Eclipse-based tool [11].

RSL is a language in which the goal is to improve the definition, production and documentation of requirements specifications, making them more consistent, clear and rigorous. An SRS with RSL can have distinct writing styles and requirements types, such as business goals, systems goals, functional requirements, user stories and use cases.

Throughout the whole process, the Requirement Engineer and the Business Stakeholder are the two major participants. They have distinct roles: the first one acts as a facilitator with the business stakeholder so that the best requirements can be found, and the process takes less effort. The second player, in turn, is essential to develop a proper project, since it is a major source of information and an excellent evaluator of requirements in order to validate them. Together, they are responsible for the textual inputs (Requirements Specs and Glossary) [12].

3. Bibliographic review

Considering the importance of water resources in human life, several systems were developed to help manage. In this chapter, it is presented an overview of some water resources management systems.

WaterML 2.0-GeoServer is a combination of the WaterML 2.0 standard and the Commonwealth Scientific and Industrial Research Organization's (CSIRO) GeoServer Web Feature Services (WFS), that aims to publish datasets related to water resources on the web [13]. The data is structured in four parts in this framework: time series, geographical information and geometry, details from the data provider and stations' details. A web-based Flood Information System (FIS), called Some Mare FIS Portal, was created using the WaterML 2.0-GeoServer, allowing the access to flood information, the participation of citizens, by discussing related issues and reporting floods, and raising awareness to flood risk management. This portal has many components, such as, the "Data Access" component, which contains spatial and timeseries data, and the "Hydrometeorological Data", that leads to a map-based interface, that shows the monitoring stations and their information regarding precipitation, discharge or temperature. The downside of this implementation is the vast quantity of configurations needed to reach the functionalities.

Sava GIS Geoportal is a web application for sharing data regarding water resources and water management in the Sava River Basin [14]. This platform has an interactive map with several layers that contain different types of information regarding river basin management, flood risk management, and hydrological and meteorological data. A major advantage of this platform is that can interoperate between software since it follows open industry standards and protocols. There are five types of users defined, the public users, which can export areas of the map, view public spatial data, attributes and features without login in. The user's permissions can be restricted in terms of data access, which can be limited to a specific set from a specific country, and according to the user role specified. The administrator is responsible for managing the users, which includes the definition of the users' permissions on the geoportal. Connected to this geoportal, there is another webpage linked, the Sava HIS, which is a portal that presents on an interactive map real-time data acquired from hydrological and meteorological stations from the Sava river basin [14]–[16].

Iowa Flood Information System (IFIS) is an online platform that allows checking the weather, flood inundation maps, real-time flood conditions, flood forecasts, flood related data, information, applications and interactive visualizations for the population. Several system requirements were defined, such as the platform must be user-friendly, interactive, accessible through many different devices and it must not require technical skills or external applications. Data in real-time from several sensors and organizations is available on the web platform for ten days. The sensors have four flood levels (action, flood, moderate flood and major flood) that are triggered when the values registered are higher than the threshold defined by the National Weather Service and the Iowa Flood Center. The system is also able to present graphs of the time series and data from the Weather Surveillance Doppler Radars. Every twenty minutes hydrological models are run, and the results uploaded to IFIS, providing this way several flood inundations scenarios. One major advantage of this system is its adaptive capability and integrated structure, which allows to apply it to other geographical areas and different environmental domains. With this platform, the users are able to monitor the community for floods through a single webpage, creators are encouraging the data exchange and collaboration, decision-makers can make more informed decisions, with the help of real-time data and forecasting models [17]–[19].

Hydrologic Engineering Center Data Storage System (HEC-DSS) by the US Army Corps of Engineers is a database system that stores data, especially sequential data (e.g. time series). This system has an MS-Excel add-in that permits to read and write regular interval time series and supports modelling software, such as the HEC-HMS [20].

The Observations Data Model (ODM) by Consortium of Universities for the Advancement of Hydrologic Science Inc (CUAHSI) [3] is a relational database schema that aims to store and retrieve water data, specifically time series of observations acquired from experimental sites. The data stored from observation points are saved and associated with its metadata. Tools can be associated allowing to export data and metadata, to visualize the data in plots, to modify data, to edit the controlled vocabulary used and to use web services.

CUAHSI Hydrologic Information System (HIS) is an internet-based system, that allows different entities to share data, connect servers, databases and applications. This system is composed by three parts: i) HIS Central: functions as a catalogue that contains the metadata associated with the time series, which is linked to the owners' server, ii) HydroServer: saves and organizes and publishes data, iii) Client (e.g.HydroDesktop): an interface to access the data, for it to function it requires to collect metadata from the HIS Central and data from the HydroServer. The HydroServer has several components, such as the WaterOneFlow Web Services [21].

WaterOneFlow (WOF) is a web service that aims to transmit and receive information from other computers via the web, using applications [22]. Several applications allow importing data directly to their systems, such as MS-Excel, Matlab and ArcGIS. This web service can be used to access the ODM Database [22].

3.1. Discussion

Each system presented has a distinct set of features, but some of the features can be found in more than one system, which reflects their importance and the impact of the benefits associated.

The ability to present real-time data acquired from monitoring stations is a feature of the Sava GIS Geoportal and IFIS systems have. With this feature available it is possible to better forecast floods and to manage water resources more efficiently.

The association of other tools, systems or applications, is a feature that allows adding new features, that can improve the service provided and the users' experience. For instance, HEC-DSS allows adding HEC-HMS as an extension., which allows the user to use a simulate a scenario using the data in the database.

HEC-DSS and IFIS are able to support simulation scenarios, which is an important feature because it can help define proper plans to mitigate the consequences of a flood and prevent major consequences in the communities.

By following a standard, such as WaterML, it becomes must easier to understand a system, to interoperate between software and to improve the system.

Both, the IFIS and the Sava GIS Geoportal provide tools that allow the public to analyze and visualize data regarding water resources, which can bring major benefits, such as raising a higher level of awareness among the population regarding floods.

The open-source feature is present in many of the systems previously shown, which proves that it is highly important. With this feature, it is possible to raise awareness to the topic and help improve the systems faster [23].

4. Related systems

To better understand the specifications of the RiverCure portal it is important to comprehend the functioning of the systems that are the main inputs of the RCP, the SNIRH (section 4.1), SVARH (section 4.2) and HiSTAV (section 4.3). These systems requirements were specified using both RSL and UML languages. In this section is presented a brief introduction of these systems.

4.1. SNIRH

SNIRH, was divulgated on October the 1st, 1995 by Instituto da Água (INAG) and is a water resources monitorization system, that acquires hydro-meteorological data from a set of stations, saves it in APA's Structured Query Language (SQL) Server database and releases the information in a portal. The portal [24] gets 600 visits per day, between teachers, students, researchers journalists and public administration personal and is divided into three subsystems: SNIR Lit, SNIRH Júnior and Sistema de Vigilância e Alerta de Recursos Hídricos (SVARH) [24].

There are of two main types of automatic stations in Portugal: Hydrometric or Meteorological [24]. The Hydrometric stations record data from rivers, lakes or reservoirs. This data includes the water level, flow, transport and deposit of sediments, temperature and other physical, chemical and biological properties of the water. The Meteorological stations can obtain meteorological data, such as precipitation, temperature, air humidity, wind velocity and others. It can be divided into two sub types, Udometric and Climatological. If the station only measures precipitation it is called Udometric or Udographic. If it also measures air and wind properties it is called Climatological.

In Portugal, there are a total of approximately 931 stations, 311 hydrometric and 620 meteorological [25]. For them to work, the stations must have sensors, a datalogger and a power supply, which includes a battery and a solar panel. Some stations are also equipped with teletransmission, which means that the data is sent directly from the station to APA's central. Otherwise, the data must be collected during the maintenance of the station by the maintenance company and uploaded in the maintenance company server afterwards.

For the data from the stations to become available to the regional authorities and on the web, it must go through three stages, data acquisition, data processing and data release. The processes required to acquire data from the stations vary depending on the station communication capabilities and their ownership. If the data is provided by the Spanish hydrographic authorities (e.g. Confederação Hidrológica do Tejo (CHT)) or by hydroelectric power plants' companies, known as Companhia Portuguesa de Produção de Eletricidade (CPPE) (e.g. EDP), the data is transferred to the SNIRH's server using a script. For an APA's station without teletransmission, the data will be recorded, and the parameters defined using a software called Geolog. Then it will be uploaded to the maintenance company server, where APA's technicians will download it from. Upon these steps the data is run in the application INAG SIF, that will generate a file with all the parameters that will be uploaded to the SNIRH's server and published on the web.

4.2. SVARH

SVARH is a system that allows analyzing in real-time the hydric state of the rivers and dams all over Portugal. Meteorological information is also collected to allow a more precise prediction and understand better the data collected. For the system to work a network of stations with teletransmission was installed in critical points, where the risk of floods exists.

The only users are APA's technicians, Regional Environmental Departments, Regional Water Authorities, National Fire Department and Civil Protection Service [26].

The SVARH has three core modules, as it can be seen in Figure 1 [27]:

- First module: Data acquisition (Acquisition and conversion of data from the stations with teletransmission and from partners);
- Second module: Data processing (Upload of the data to the databases);
- Third module: Data release (Dissemination of the data in real-time using Rios application).

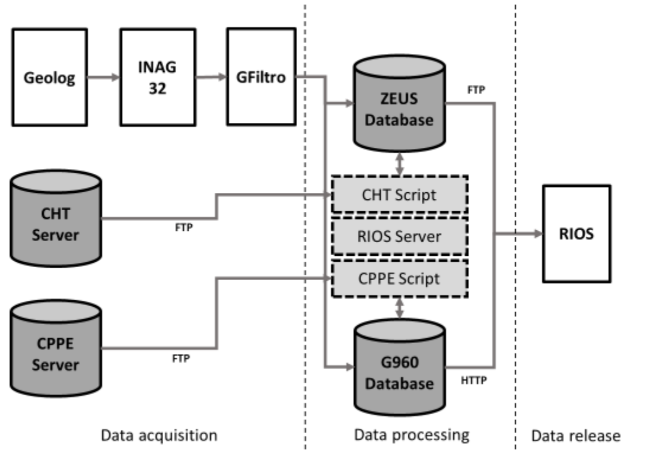

Figure 1: Schema of SVARH's processes (adapted from [27]).

The management of the alarms sent from the automatic stations (e.g. change of interrogation rhythm) and the alteration of the parameters remotely are some of the functions that the Geolog software is responsible for.

INAG32 software is responsible for converting the data from the Geolog format to American Standard Code for Information Interchange (ASCII), creating one or more files. This application works in collaboration with the Geolog. This allows to define the number of days that the output will contain, and it deletes older data [27].

GFiltro software is used to convert the ASCII files obtained from the INAG32 to files in the SVARH format and it also uploads them into the database [27].

This system has two servers, Zeus and G960, both can store the data recorded by the automatic stations. The Zeus server uses the File Transfer Protocol (FTP) to transfer the data and the G960 server uses Hypertext Transfer Protocol (HTTP) (Figure 1). The data after being transferred can be consulted using RIOS app. Another functionality of the G960 server is the ability to send the alarms from the stations via text message or email [27].

Besides the information recovered from APA's stations, the CPPE, and the CHT, also send their data in exchange for APA's information (Figure 1) [26]. The data is transferred every hour from both servers using specific scripts that follow the FTP transference protocol [27].

RIOS is the application used to present the data in a user-friendly manner, considering that the access to the data is based on web services and SQL servers. RIOS desktop is a software that can create charts and is able to export data to MS-Excel. Besides the RIOS desktop, there is also the RIOS editor, that is used to manage users and to edit the screen layouts, which includes designing and defining rivers, basins and stations. When defining a station, the parameters, alarm levels and past events are specified. RIOS SMS application is also an important service that allows the reception of text messages (SMS) alarms on the mobile phone for selected incidents [26].

When an alarm is triggered, a notification is sent to the control centre automatically and it also changes the frequency of the data storage and transfer. For the warning to be sent, a specific parameter that is being measured must have reached a previously defined limit. From Figure 2 it is possible to understand that Rios has two levels of alarm (yellow and red) defined and that the colour of the station symbol, basin and district matches the alarm level. Note that the limits of each level change between the stations, since they are specifically defined for each site. All the records from the area are considered by an APA technician when deciding the alarm limits.

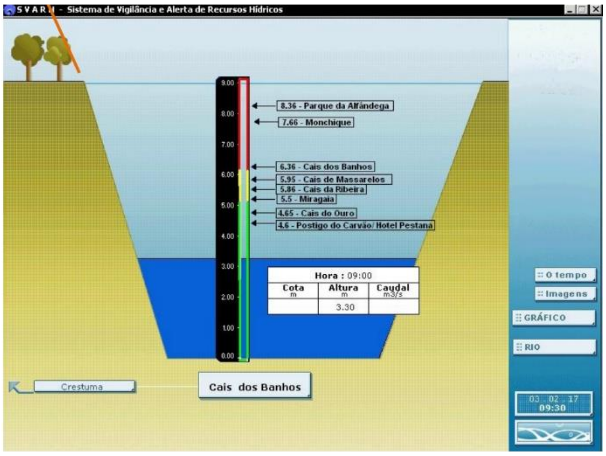

Figure 2: Alarm levels example in Rios (extracted from [25]).

4.3. *HiSTAV*

HiSTAV, a high performance version of Strong Transients in Alluvial Valleys model, is a high resolution simulation tool that uses a two dimensional horizontal mathematical model based on shallow water equations featuring dynamic bed geometries and sediment transport to simulate fluvial and estuarine flows. The possibility of forecasting an overland tsunami and the development of a system that could assist in the decision making process regarding water related hazards were some of the motivations for the creation of this model [28]. One important feature of this tool is that it can track uncoupled materials between consecutive timesteps, namely large debris, such as vehicles and containers. This property allows detecting areas where the debris are more likely to accumulate, thus creating a new obstacle [29]. It is important to state that this model is faster than real-time tsunami forecasts [28].

The HiSTAV has some subprograms, that have the function of supporting the user when defining the geometric entities needed for the creation of the mesh and refining them with the assist of a GIS tool [30]. There are two types of utilities, the pre processing, which includes a GIS software (e.g. ArcMap) and the mesh refinement, and the post processing, which consists of the resampler and raster converter.

The simulation process normally starts with the preparation of all the data that is going to be applied in the model. To accomplish that a pre-processing tool is utilized, which contains a mesh generator tool. To generate a mesh, it is necessary to first define the domain and the boundary conditions using QGIS [31]. The boundary conditions (inlet or outlet) are defined by the boundary lines and points. To each point, it is assigned a specific timeserie and to each line is specified the parameter (e.g. flow, height) present in the timeseries. The pre-processor, imports shapefiles and raster files, such as, DTM, domain, obstacles and boundaries, in order to obtain the pre-processing outputs, which includes the mesh file and the files with the initial conditions.

Then, after the control files (time, physics and numeric) are created and the input data is selected, the data processing stage can start, which is the execution of the HiSTAV. Lastly, upon obtaining the simulation output, it can be viewed in the ParaView software [32].

5. RiverCure Portal

This chapter describes the RCP system, all its stakeholders, components and requirements. The specifications presented in this chapter are the result of the information and feedback collected during team meetings from several members. These meetings and the case study allowed to evaluate the specifications.

5.1. *Introduction*

This system has four main data inputs from different sources, the weather radar information from the IPMA website, the data from SVARH, the photos uploaded by the users or collected from the social networks, and the simulations obtain from the HiSTAV software.

From Figure 3 is possible to understand that the data has three stages of transformation, acquisition, processing and view and that the process that each data pass-through depends on the source and the type of data. Data from the IPMA web and SVARH server are obtained using an importation script that stores the data in the RiverCure server, so it can become available on the portal. The process to obtain the simulations results depends on the inputs available, if the mesh, initial conditions and boundary conditions are not defined then the data must be prepared using the QGIS and the Pre-processor, otherwise, the files can be selected from the RiverCure server and the HiSTAV can be executed immediately. Note that data from SVARH, IPMA and photos can be inputs of the HiSTAV and are available in the RiverCure server. To view the outputs of the simulations the portal uses the ParaView software [32] that reads the results files stored in the RiverCure server. Additionally, after acquiring and storing the photos, the Social Flood application gathers the photo metadata and the estimated water height, that are saved in the RiverCure server, so it can become available on the RCP.

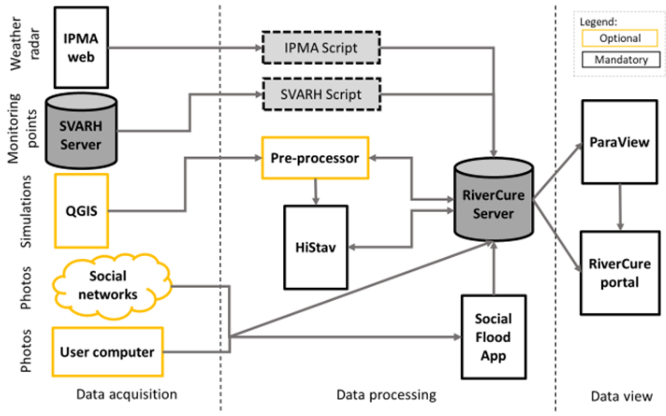

Figure 3: RCP data transformation schema.

5.2. *Stakeholders*

This project has several stakeholders, each has a different impact on the project, and it can be impacted differently.

The RiverCure project is sponsored by FCT and it has three partners, IST, INESC ID and APA, so all of these organizations are stakeholders of the RCP.

There are several people involved in the project, some are part of the team RiverCure and others belong to the partners' organization (e.g. APA responsible). The project manager is responsible for managing the development of the RiverCure project, the team and the communication between the stakeholders.

HiSTAV and Social flood are considered as system internal since the two tools are incorporated into the system. SVARH will transfer its data to the RiverCure server, so it is considered an external system.

5.3. Actors

The RiverCure portal has three types of actors, external system, user and timer. The SVARH and all the user types are actors of the system and each one has a specific stakeholder associated with it, as it can be seen in Specification 1.

In the RCP, a user is someone that has a user account. SuperAdmin, Manager, Technician and Citizen are the four main types of users, each with different access permissions.

A RiverCure portal user has a main role (SuperAdmin, Manager, Technician or Citizen) and if the main role is "Technician" it can have one or more secondary roles. There are six types of secondary roles:

- Radar technician: Person that is specialized in radar images (e.g. IPMA employee);
- Numeric technician: User that knows about numeric models (e.g. Researcher);
- Udometric technician: Specialized in udometric stations and consequently data (e.g. APA responsible);
- Spatial technician: GIS is the area of expertise of this person (e.g. Researcher, APA or IPMA employee);
- Crowdsourcing technician: User that is specialized in the crowdsourcing topic (e.g. Researcher);
- Hydrometric technician: Specialist in hydrometric stations and data (e.g. APA employee).

```
Actor aS_SVARH "SVARH" : ExternalSystem [stakeholder RCPortal_Business.stk_SVARH]

Actor aU_User "User" : User [stakeholder RCPortal_Business.stk_user]
Actor aU_Citizen "Citizen" : User [isA aU_User stakeholder RCPortal_Business.stk_citizen]
Actor aU_Technician "Technician" : User [isA aU_User stakeholder RCPortal_Business.stk_technician]
Actor aU_Radar "Radar technician" : User [isA aU_Technician stakeholder RCPortal_Business.stk_technician]
Actor aU_Num "Numeric technician" : User [isA aU_Technician stakeholder RCPortal_Business.stk_technician]
Actor aU_Udo "Udometric technician" : User [isA aU_Technician stakeholder RCPortal_Business.stk_technician]
Actor aU_Spatial "Spatial technician" : User [isA aU_Technician stakeholder RCPortal_Business.stk_technician]
Actor aU_Crowd "Crowdsourcing technician" : User [isA aU_Technician stakeholder RCPortal_Business.stk_technician]
Actor aU_Hydro "Hydrometric technician" : User [isA aU_Technician stakeholder RCPortal_Business.stk_technician]
Actor aU_Manager "Manager" : User [isA aU_User stakeholder RCPortal_Business.stk_manager]
Actor aU_SuperAdmin "SuperAdmin" : User [isA aU_User stakeholder RCPortal_Business.stk_SuperAdmin]

Actor aT_NotifyAlarm "Notify active alarm" : Timer [description "Notify user that alarm is active"]
Actor aT_AccountNotValidated "Account not validated after 24h" : Timer [description ...
```

Specification 1: RiverCure portal actors in RSL.

The permissions of user roles accumulate, hereby, for instance, the SuperAdmin actor has accumulated the permissions of the manager, that can perform the tasks of the technicians and citizens. However, managing alarms has a restriction, a Hydrometric/Udometric technician can only edit and delete the sensor alarms if the legislation of the country in question allows it.

5.4. Data Entities

The data entities of this information system were defined using a UML diagram and an RSL specification.

It was defined eight main data entities packages in the RiverCure system, user, simulation, sensor, event, interactive map, about, spatial and feature. Part of the structure of these entities was inspired by the SVARH data entity structure since it is one of the main contributors of this portal.

To describe the spatial location of an object, five data entities were defined, "Spatial location", "Parish", "Municipality", "District" and "Country".

In the feature package, is defined the "Hydrometric feature" data entity, which is used to describe a hydrometric object, such as a river, a dam or a basin. Each hydrometric feature is associated with one or more municipalities since it can cross several municipal frontiers. For instance, the Vouga river crosses Oliveira de Frades, Vouzela and many other municipalities.

From Figure 4 it is possible to understand that the sensor entity is characterized by code, name, type and it measures a specific parameter. There are three types of sensors, the fixed in situ, which are sensors that register values regarding an exact, established location (e.g. SVARH sensors), the weather radar, which is a fixed sensor that can capture data remotely about several places (e.g. weather radars from IPMA) and the photo, that it is a mobile sensor since the camera operator can shift location. Fixed in situ sensors and weather radars have more than one user responsible for it, but the photo sensor normally only has one.

A monitoring point is composed of one or more sensors, that register several data timeseries, which are a series of time value pairs. The "e_MonitoringPoint" data entity states the equipment altitude, "descriptionReference", which is a list of documents related to the monitoring point, code, name, status (e.g. active, suspended) and the type of monitoring point (e.g. Hydrometric, Meteorological), which depends on the type of sensor associated. Each sensor has a specific location set and at least one hydrometric feature associated. For example, a meteorological sensor can state the name of the basin to which is associated and a hydrometric indicates the basin and river name. A recording rhythm is defined for each sensor, which means that a value is recorded every 15 minutes, for instance. The hydrometric sensor has a special attribute, the zero level scale, which indicates the level where zero of the scale is set.

A fixed in-situ sensor allows to define up to three alarms for a parameter. Each one with a specific colour (red, yellow or green), name, description and action, which can be to send a notification to a certain user, for example. A low and up threshold must be defined for each alarm and to do so the historical measurements ("Historical value") of each location can be consulted, in case there is any. The "District" and the "River basin" data entities are also associated with the "Alarm" data entity since the alarm of sensors can be represented using those entities.

As can be seen in Figure 4, the weather radar is a generalization of the "e_Sensor". This data entity has the "Spatial location" entity associated, which states the latitude and longitude of the equipment. The "e_RadarObservation" is the data entity that defines the data obtained from the radar. Since the data obtained is in image format, the only attributes are the ones associated with it such as date, time, ID, file identifier and version.

The ”Photo source” data entity, states if the photo was uploaded or obtained from social networks and it is linked to the “Photo observation” entity, which defines the main attributes of the photo, namely its ID, date, time, the value obtained from the social flood application and last modifications made, also known as history. Each photo observation has a specific location and it can have a hydrometric feature associated.

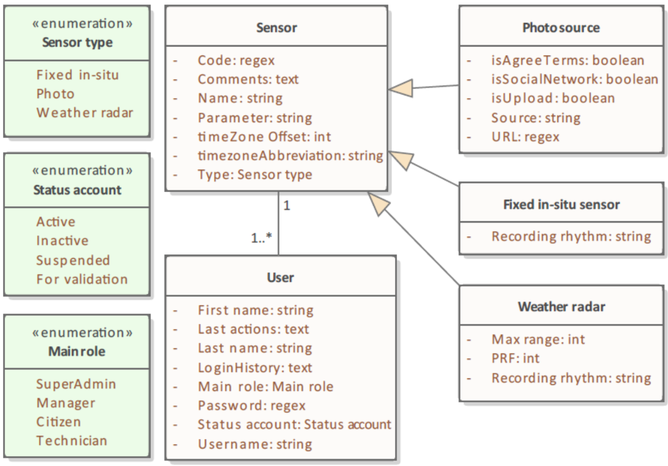

Figure 4: UML diagram of the RCP DataEntities: Sensor simple package.

An event is created by a single user, that is responsible for defining all its attributes, namely a code, a description, the state (if it is occurring or over), the type of event (e.g. Flood, Hydrological drought), and the start and end date-time. Each event has one or more fixed in-situ sensor alarms, which means that the “Alarm” and the “Fixed in-situ” data entities are linked to the “Event” data entity, and it can also have photos associated.

For a simulation to work it is necessary to properly define the following entities, physical constants, time and numeric information and input files. There are two types of simulations, the simulation and the scenario, which is a type based on the main simulation in which the “Time information” is the only data entity that is modified. The simulation results are a set of several output files, each one with a specific name and folder. Both simulation and input file have a user associated, that corresponds to the owner of the file. Each simulation is associated with one or more basins and districts, which allows an easier search process since it is not required to open each file to discover the location of a simulation.

The “e_InitialCondition”, “e_Mesh” and “e_BoundaryCond” data entities are types of input files of the simulation. Photos and weather radar observations regarding the Águeda case study were available, so they were also selected as inputs.

The interactive map is composed of layers, that show different types of information (e.g. simulation layer, sensor layer, event layer, alarm layer). On the map data regarding the last days or a time interval can be presented and the basemap can be altered. Each type of layer has a specific symbol, colour and set of attributes associated. For example, the event layer can show events that are occurring or that occurred in the last 48 or 72 hours and the simulation layer can present simulations or scenarios.

In the “About” package there are two data entities, the “RiverCure project” data entity, that describes the RiverCure project and its partners, and the “Feedback” entity that defines the attributes required for the user to provide feedback regarding the portal. The “e_RCProject” data entity is associated with the “e_Entity” and “e_Contact” entities. The feedback requires the username of the user, so it is associated with the “e_User” data entity.

To define the user’s profile information several data entities were created, as it can be seen in Specification 2. The “User” entity is connected to the “e_Responsible Party” data entity, which describes the position that a user has in an entity and to the “Contact” data entity which states the email and phone number of the user. An “Entity” can be for example a company or an association and it must have an official address linked. An “Entity” has many employees, each with a specific position, which means that an “Entity” has several users associated with it.

```
DataEntity e_User "User" : Master [
    attribute firstName "First name" : String [isNotNull isEncrypted]
    attribute lastName "Last name" : String [isNotNull isEncrypted]
    attribute mainRole "Main role" : String [isNotNull]
    attribute pass "Password" : Regex [isNotNull isEncrypted]
    attribute Status "Status account" : DataEnumeration de_StatusAccount [isNotNull]
    attribute username "Username" : String [isNotNull isUnique]
    attribute lastActions "Last actions": Text [isReadOnly]
    attribute loginHistory "Login history" : Text [isReadOnly]
    primaryKey (username)]

DataEnumeration de_MainRole "Main role" : Simple [values ...

DataEnumeration de_StatusAccount "Status account" : Simple [values
"Active", "Inactive", "Suspended", "For validation"]

DataEntity e_Contact "Contact" : Reference [...

DataEntity e_ResponsibleParty "Responsible party" : Reference [
    attribute positionName "Position name" : String [isNotNull]
    attribute username "Username" : String [isNotNull isUnique]
    attribute entityName "Entity name" : String [isNotNull isUnique]
    primaryKey (positionName, username)
    foreignKey e_User (username)
    foreignKey e_Entity (entityName)]

DataEntity e_Entity "Entity" : Reference [...

DataEntity e_Adress "Adress" : Reference [...
```

Specification 2: RSL specification of RCP DataEntities: User main.

To define the user settings, it is required to associate them with a user, so the username is the foreign key used to link the two data entities. The “e_UdoHydroSettings” entity describes special attributes for udometric and hydrometric technicians and the “e_TechnicianSettings” data entity specifies particular attributes for the technician users, which means that this entity is a generalization of the first.

The user permissions depend on the user role and the river basins associated with it, which means that the user can only access the information regarding a specific set of river basins. In terms of alarm permission, some users can access all the alarms, others can only access alarms created by the authorities, others depend on the secondary role can or cannot access the alarms and some do not have access to any alarm.

5.5. *Requirements*

The use cases of the RiverCure portal are specified using a UML diagram and an RSL specification. The UML allows the user to have a quicker overview and the RSL provides more details regarding the use cases, namely the data entities associated with it.

To demonstrate the sequence of actions that each user type can do and the information that it requires, RiverCure portal mock-ups were made.

Figure 5 shows the home page mock-up of the RiverCure portal and as it can be seen, it has five main tabs on the top of the page, data, photos, simulations, events and about. The information presented on this page can be changed according to the user's preferences. The last actions made by the user on the platform, the last events added, and the interactive map are the information displayed in the mock-up. The user permissions can also be visualized on the top right section of the page, below the username, and the settings and the log out button. The icon located next to the mock-up indicates the users' roles that can access the mock-up view.

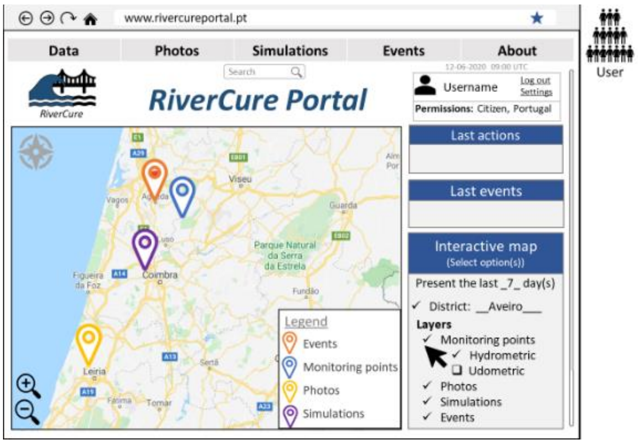

Figure 5: Home page mock-up of the RCP.

The data tab allows the user to access the interactive map tool, information regarding SVARH monitoring points and radar observations. The interactive map tool, allows the user to view the data selected on the control panel. Each point of information is represented with a distinct symbol and it has a set of data associated that pops up by clicking on it. Note that the information available on this tool varies according to the user role.

A user can consult the interactive map by selecting the layers that correspond to the information that the user wishes to view. This tool also allows the user to edit the basemap (e.g. satellite, administrative) and manage areas of interest, which includes adding, deleting, editing areas selected on the map.

From Figure 6 is possible to understand that to consult the monitoring points data the user must choose one of three views, by basin, by district or by monitoring point. By selecting the option by basin or by district, a schema of a monitoring point in that area opens showing the last data record stored. The option by monitoring point shows a list of all the points with data available and the user can select which one it wishes to read the data records. The user can also view the data in an alternative mode, by creating a graphic with the data.

Hydrometric and udometric technicians have some use cases specific to them regarding the monitoring points and the interactive map. These user roles can check real-time information from the monitoring points and to flag errors in any monitoring point or data.

A user can consult radar observations, by selecting a radar, a parameter and a date on the search control. The radar observations can also be exported.

The "Photos" tab of the RiverCure mock-up has two selection options, the first allows the user to upload photos in the portal and the second is the gallery, which is where all the photos are displayed. For a user to upload a photo in the portal it is required to read and agree with the terms and conditions defined.

All users have access to the photo gallery and can select a photo in order to read its information, which includes its metadata and the estimated water height. When consulting a photo, a user can export, select, flag, view the next or the previous image without any restrictions, but it can only delete the photo or edit its metadata if it is owned by the user. The spatial and crowdsourcing technicians have use cases specific to them, such as editing the photo's metadata and estimated height, consulting the modifications made to the photo's information and deleting a photo.

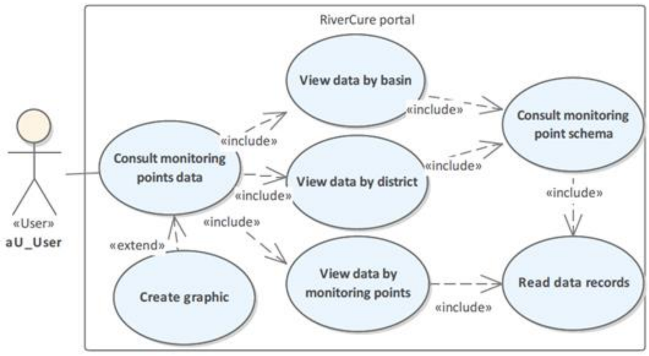

Figure 6: UML diagram of the RCP UseCases: Monitoring points package.

The simulator tab has two versions, one for the general users and other for the numeric and spatial technician since these last two roles have access to special features, such as creating a new simulation from scratch.

To create a new scenario the user must define its input data previously, which includes the area of study (also known as domain), and the start and end date-time. While the simulator is running, a progress indicator is shown on the screen, so the user can track the simulation status. During this time the portal can suggest data related to the simulation domain, such as photos and data from monitoring points.

Besides being able to create a new simulation scenario, a user can also consult the simulation records. A user can choose to save or not the simulation scenario obtained, and it can opt to export those results. In the simulation records tab is available all the simulations saved by the user and the ones published by specialized users, like spatial technicians, for instance. Simulation scenarios owned by the user can be edited and deleted by the user.

As formerly mentioned, the numeric and spatial technicians can create new simulations, but to understand how the simulation tool works it is important to read the instructions. For this tool to work, input files (e.g. DTM, Domain) are required, so in the simulations tab was created a sub-tab, "Input files". This sub-tab allows the user to view the files, manage the files owned and upload new files, which may encourage information

Spatial and numeric technicians have specific use cases regarding simulations. The "Create new simulation" use case has an extension point, "Run the Pre processor" since this use case is only required if a mesh of the area that is being studied is not created or if it can be updated to a better version. To develop a new simulation, the user must first configure the model and then execute the simulator tool. To configure the simulation model a user must perform certain actions (e.g. define time and numeric information). However, it can also choose to execute other actions, such as defining distributed conditions, if feasible. Data from the radar observations, photos and monitoring points can be used to configure the simulation model. These technicians can choose to share or not the results of the simulations performed.

The user's events tab has only one division, that presents the events records, but the udometric and hydrometric technicians events tab version has two, "Open events" and "Events records".

From Specification 3 it is possible to understand that in an event record a user can export its data and consult the photos associated with it, in case there are any. Besides being able to execute da same actions as a user, a udometric and hydrometric technician can also flag event data errors and consult events that are still occurring. The use case "uc_16_ConsultOpenEvents" shows that it has three extension points "EPConsultPhoto", "EPFlagEventDataError" and "EPExportEventData".

The "About" tab is the same for all the users and it has two main divisions, "Feedback" and "RiverCure project", where a user can read about the project and the partners involved.

A user by clicking on the settings can access a secondary panel, that varies according to the user role. A citizen can view its profile information, which includes its permissions and personal data, and it can also manage its account settings, which include the notifications and preferences settings.

Technicians have certain actions regarding the management of the account settings that are specific for them, like be able to view the portal as a citizen would. A technician can select its favourite parameters and monitoring points, which means that the most relevant data on the portal, for this user, is more easily accessible.

Besides the previously mentioned use cases, the hydrometric and udometric technicians are also able to edit the alarms notifications by choosing to be notified for all red and/or yellow alarms and if they wish to be alerted via SMS or email.

Since the alarm feature is only available to certain user roles, SuperAdmin, manager, udometric and hydrometric technician, these are the only actors of the use cases regarding alarm settings. The portal offers the user two alarm modes, from which the user must select one. The first alarm mode follows the alarm definitions, such as alarm levels and thresholds, set by the authorities, which means that the user cannot edit them. The second mode, also known as "My alarms", allows the user to specify its alarms, by defining a monitoring point, parameter, threshold and alarm level for each alarm. However, this mode is not available for all users and countries, since the legislation in some countries does not allow it.

The manager is responsible for managing the information, which can include a variety of use cases, such as managing the monitoring points, events and radar observations. This actor has access to all the data flagged by other users, which means that "View flagged data", "View records flagged" and "View list of flagged photos" are use cases that can be executed by the manager. Besides managing all the data regarding the monitoring points, radar observations, events and simulations, the user can view the portal analytics and the user accounts logs.

The SuperAdmin is responsible for managing users, entities (e.g. companies, universities), the communication with the partners and the portal's updates, such as the model of the photos. Managing users includes validating the users' accounts, managing user roles, which allows this user to create, edit and delete user roles, and managing permissions, which allows to edit the permissions associated with each user and user role. When an account is not validated after 24h the SuperAdmin receives a notification automatically from the system.

```
UseCase uc_15_ConsultEventsRecords "Consult events records" : EntitiesBrowse [
    actorInitiates aU_User
    dataEntity ec_Event
    actions aFilter, aRead, aSelect
    extensionPoints EPExportEventData, EPConsultPhoto, EPFlagEventDataError
    description "Consult the records from past events"]

UseCase uc_15_1_ExportEventData "Export event data" : EntitiesInteropExport [...

UseCase uc_15_2_FlagEventDataError "Flag event data error" : EntitiesOther [
    actorInitiates aU_Hydro
    actorParticipates aU_Udo
    dataEntity ec_Event
    actions aFlag, aSelect
    extends uc_15_ConsultEventsRecords onExtensionPoint EPFlagEventDataError
    extends uc_16_ConsultOpenEvents onExtensionPoint EPFlagEventDataError]

UseCase uc_16_ConsultOpenEvents "Consult open events" : EntitiesBrowse [
    actorInitiates aU_Hydro
    actorParticipates aU_Udo
    dataEntity ec_Event
    actions aFilter, aRead, aSelect
    extensionPoints EPConsultPhoto, EPFlagEventDataError, EPExportEventData
    description "Consult the events that are occuring"]
```

Specification 3: RSL specification of RC UseCases: Events.

6. CONCLUSIONS AND FUTURE WORK

In this work, it was defined the SNIRH's and SVARH's models in both RSL and UML, which allowed to understand how these systems function. The HiSTAV model was also defined in RSL and UML so that the integration of this system in the portal could be done properly, by considering all its specifications. The specification of the RiverCure portal system was made using UML diagrams and RSL so that in a first stage the specification would be easily understandable and in a second it would assure that the requirements were written in a rigorous and structured way. Mock-ups of the RiverCure portal were also developed, to demonstrate the specifications of the portal and to receive more feedback, especially from team members that are not experts in requirements specification.

The RiverCure portal is a web platform that integrates data from different entities (e.g. official agencies), flood photos with interactive maps and a flood simulation tool. With the extensive data resources and capabilities, this portal can bring many benefits such as improving the flood forecasting systems and supporting in the decision-making process. By providing a user-friendly system, stakeholders can make smarter decisions. It can also support the development of plans to prevent and mitigate the impacts of floods. Consequently, it can improve the time and quality of the authority's responses during a flood since the allocation of the means of assistance would be done more effectively. The fact that a single platform can integrate and present data from several sources, it shall allow to better manage the water resources.

The future work should focus on creating the portal through the implementation of the system requirements preliminary defined in this dissertation. Afterwards, the RCP must be tested with a specific dataset, by a group of controlled users. To evaluate the system in terms of features and usability, a survey must be made to all actors.

Implementing new features, such as a flood risk index for infrastructures, with several levels (e.g. low, medium, critical), can improve this portal and consequently help the decision-maker to have more confidence in its choice.

ACKNOWLEDGEMENTS

This work was supported by national funds through FCT – Fundação para a Ciência e a Tecnologia, under the PTDC/CTA-OHR/29360/2017 project.

REFERENCES

[1] Agência Portuguesa do Ambiente, "Avaliação Preliminar dos Riscos de Inundações em Portugal Continental," Relatório Técnico, 2018.

[2] T. Miksa, J. Cardoso, and J. Borbinha, "Framing the scope of the common data model for machine-actionable Data Management Plans," in *Proceedings of the 2018 IEEE International Conference on Big Data*, 2018.

[3] J. S. Horsburgh, D. G. Tarboton, D. R. Maidment, and I. Zaslavsky, "A relational model for environmental and water resources data," *Water Resour. Res. Wiley*, vol. 44, no. 5, May 2008.

[4] D. A. S. Conde, M. A. V. Baptista, C. Sousa Oliveira, and R. M. L. Ferreira, "A shallow-flow model for the propagation of tsunamis over complex geometries and mobile beds," *Nat. Hazards Earth Syst. Sci. Copernicus Publ.*, vol. 13, no. 10, 2013.

[5] A. R. Silva, "Rigorous Requirements Specification with the RSL Language: Focus on Uses Cases," in *Proceedings of ISD' 2019, AIS*, 2019.

[6] A. R. Silva, "Linguistic Patterns and Linguistic Styles for Requirements Specification (I)," in *Proceedings of the 22nd European Conference on Pattern Languages of Programs (EuroPLoP '17), ACM*, 2017.

[7] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A Design Science Research Methodology for Information Systems Research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, 2007.

[8] C. Pohl, Klaus; Rupp, *Requirements Engineering Fundamentals: a study guide for the certified professional for requirements engineering exam, foundation level*, 2nd ed. Rocky Nook, 2015.

[9] A. R. Silva and C. Videira, *UML, Metodologias e Ferramentas CASE*, 2nd ed., vol. 1. Centro Atlântico Editora, 2005.

[10] OMG, "OMG Unified Modeling Language (OMG UML)." OMG, 2017.

[11] L. P. Gonçalves and A. R. Silva, "Towards a Catalogue of Reusable Security Requirements, Risks and Vulnerabilities," in *Proceedings of ISD' 2018, AIS*, 2018.

[12] D. de Almeida Ferreira and A. R. Silva, "RSLingo: An Information Extraction Approach toward Formal Requirements Specifications," in *Proceedings of 2nd IEEE International Workshop on Model-Driven Requirements Engineering (MoDRE), IEEE Computer Society*, 2012.

[13] A. Almoradie, I. Popescu, A. Jonoski, and D. Solomatine, "Web Based Access to Water Related Data Using OGC WaterML 2.0," *Int. J. Adv. Comput. Sci. Appl.*, vol. 3, no. Building a Regional Observation System in the Black Sea Catchment, 2013.

[14] "International Sava River Basin Commission - Sava GIS Metadata Catalogue." [Online]. Available: http://savagis.org/metadatacatalogue/srv/eng/search#fast=index&from=1&to=50. [Accessed: 12-Jun-2020].

[15] International Sava River Basin Commission, "Metadata Catalogue Public User Manual." 2016.

[16] International Sava River Basin Commission, "Data and Information Exchange - Sava GIS & Sava HIS," Presented for the Western Balkans Investment Framework.

[17] I. Demir and W. F. Krajewski, "Towards an integrated Flood Information System: Centralized data access, analysis, and visualization," *Environ. Model. Softw.*, vol. 50, Dec. 2013.

[18] I. Demir, E. Yildirim, Y. Sermet, and M. A. Sit, "FLOODSS: Iowa flood information system as a generalized flood cyberinfrastructure," *Int. J. River Basin Manag.*, vol. 16, no. 3, Jul. 2018.

[19] W. F. Krajewski *et al.*, "Real-Time Flood Forecasting and Information System for the State of Iowa," *Bull. Am. Meteorol. Soc.*, vol. 98, no. 3, Mar. 2017.

[20] D. Ford, N. Pingel, and J. J. DeVries, "Hydrologic Modeling System HEC-HMS: Applications Guide." U.S. Army Corps of Engineers Hydrologic Engineering Center, Washington, 2008.

[21] "CUAHSI-HIS." [Online]. Available: http://his.cuahsi.org/index.html. [Accessed: 15-Jun-2019].

[22] I. Zaslavsky, D. Valentine, and T. Whiteaker, "CUAHSI WaterML," Open Geospatial Consortium Discussion Paper 07-041r1, 2007.

[23] D. P. Ames, J. S. Horsburgh, Y. Cao, J. Kadlec, T. Whiteaker, and D. Valentine, "HydroDesktop: Web services-based software for hydrologic data discovery, download, visualization, and analysis," *Environ. Model. Software, Elsevier*, vol. 37, 2012.

[24] "Sistema Nacional de Informação sobre Recursos Hídricos (SNIRH)." [Online]. Available: https://snirh.apambiente.pt. [Accessed: 13-May-2019].

[25] M. Saramago, "Redes de Monitorização Hidrometeorológicas," *Rev. Recur. Hídricos, Assoc. Port. dos Recur. Hídricos*, vol. 38, no. 1, 2017.

[26] M. Saramago, "Hydrologic Surveillance System Using Wireless Technologies," in *Proceedings of 4th International Conference on Experiences with Automatic Weather Stations (ICEAWS)*, 2006.

[27] R. Gomes, M. Saramago, and R. Rodrigues, "SVARH - Sistema de Vigilância e Alerta de Recursos Hídricos," Relatório Técnico, Instituto da Água, 2003.

[28] D. A. S. Conde, "High-Performance Modelling of Tsunami Impacts on Built Environments," PhD Thesis, Universidade de Lisboa, Instituto Superior Técnico, 2018.

[29] D. A. S. Conde, M. J. Telhado, M. A. Viana Baptista, and R. M. L. Ferreira, "Severity and exposure associated with tsunami actions in urban waterfronts: the case of Lisbon, Portugal," *Nat. Hazards*, vol. 79, no. 3, 2015.

[30] D. A. S. Conde, "A Tsunami in Lisbon Where to run ?," MSc Thesis, Instituto Superior Técnico, Universidade de Lisboa, 2012.

[31] B. C. Almeida, A. P. M. Saliba, and D. A. S. Conde, "Retroanálise da propagação decorrente da ruptura da barragem do Fundão através do modelo numérico HISTAV," in *Proceedings of the III SIAAR - Simpósio Íbero-Afro-Americano de Riscos*, 2019, vol. 3, no. June.

[32] "ParaView." [Online]. Available: https://www.paraview.org/. [Accessed: 14-Sep-2020].