Multi-agent System for Flood Forecasting in Tropical River Basin

by

José Antonio Simmonds Sheppard

A dissertation submitted by in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Computer Science and Technology

Universidad Carlos III de Madrid

Advisor(s):

Agapito Ledezma Espino
Juan Antonio Gómez Herrera

Léganes, 2022

To My Mother and Siblings,
for your faith, support, and patience during these years

Abstract

It is well known, the problems related to the generation of floods, their control, and management, have been treated with traditional hydrologic modeling tools focused on the study and the analysis of the precipitation-runoff relationship, a physical process which is driven by the hydrological cycle and the climate regime and that is directly proportional to the generation of floodwaters. Within the hydrological discipline, they classify these traditional modeling tools according to three principal groups, being the first group defined as trial-and-error models (e.g., "black-models"), the second group are the conceptual models, which are categorized in three main sub-groups as "lumped", "semi-lumped" and "semi-distributed", according to the special distribution, and finally, models that are based on physical processes, known as "white-box models" are the so-called "distributed-models". On the other hand, in engineering applications, there are two types of models used in streamflow forecasting, and which are classified concerning the type of measurements and variables required as "physically based models", as well as "data-driven models".

The Physically oriented prototypes present an in-depth account of the dynamics related to the physical aspects that occur internally among the different systems of a given hydrographic basin. However, aside from being laborious to implement, they rely thoroughly on mathematical algorithms, and an understanding of these interactions requires the abstraction of mathematical concepts and the conceptualization of the physical processes that are intertwined among these systems. Besides, models determined by data necessitates an a-priori understanding of the physical laws controlling the process within the system, and they are bound to mathematical formulations, which require a lot of numeric information for field adjustments. Therefore, these models are remarkably different from each other because of their needs for data, and their interpretation of physical phenomena. Although there is considerable progress in hydrologic modeling for flood forecasting, several significant setbacks remain unresolved, given the stochastic nature of the hydrological phenomena, is the challenge to implement user-friendly, re-usable, robust, and reliable forecasting systems, the amount of uncertainty they must deal with when trying to solve the flood forecasting problem. However, in the past decades, with the growing environment and development of the artificial intelligence (AI) field, some researchers have seldomly attempted to deal with the stochastic nature of hydrologic events with the application of some of these techniques.

Given the setbacks to hydrologic flood forecasting previously described this thesis research aims to integrate the physics-based hydrologic, hydraulic, and data-driven models under the paradigm of Multi-agent Systems for flood forecasting by designing and developing a multi-

agent system (MAS) framework for flood forecasting events within the scope of tropical watersheds.

With the emergence of the agent technologies, the "agent-based modeling" and "multi-agent systems" simulation methods have provided applications for some areas of hydro base management like flood protection, planning, control, management, mitigation, and forecasting to combat the shocks produced by floods on society; however, all these focused on evacuation drills, and the latter not aimed at the tropical river basin, whose hydrological regime is extremely unique.

In this catchment modeling environment approach, it was applied the multi-agent systems approach as a surrogate of the conventional hydrologic model to build a system that operates at the catchment level displayed with hydrometric stations, that use the data from hydrometric sensors networks (e.g., rainfall, river stage, river flow) captured, stored and administered by an organization of interacting agents whose main aim is to perform flow forecasting and awareness, and in so doing enhance the policy-making process at the watershed level.

Section one of this document surveys the status of the current research in hydrologic modeling for the flood forecasting task. It is a journey through the background of related concerns to the hydrological process, flood ontologies, management, and forecasting. The section covers, to a certain extent, the techniques, methods, and theoretical aspects and methods of hydrological modeling and their types, from the conventional models to the present-day artificial intelligence prototypes, making special emphasis on the multi-agent systems, as most recent modeling methodology in the hydrological sciences. However, it is also underlined here that the section does not contribute to an all-inclusive revision, rather its purpose is to serve as a framework for this sort of work and a path to underline the significant aspects of the works.

In section two of the document, it is detailed the conceptual framework for the suggested Multiagent system in support of flood forecasting. To accomplish this task, several works need to be carried out such as the sketching and implementation of the system's framework with the (Belief-Desire-Intention model) architecture for flood forecasting events within the concept of the tropical river basin. Contributions of this proposed architecture are the replacement of the conventional hydrologic modeling with the use of multi-agent systems, which makes it quick for hydrometric time-series data administration and modeling of the precipitation-runoff process which conveys to flood in a river course. Another advantage is the user-friendly environment provided by the proposed multi-agent system platform graphical interface, the real-time generation of graphs, charts, and monitors with the information on the immediate event taking place in the catchment, which makes it easy for the viewer with some or no background in data analysis and their interpretation to get a visual idea of the information at hand regarding the flood awareness.

The required agents developed in this multi-agent system modeling framework for flood forecasting have been trained, tested, and validated under a series of experimental tasks, using the hydrometric series information of rainfall, river stage, and streamflow data collected by the hydrometric sensor agents from the hydrometric sensors.

Resumen

Como se sabe, los problemas relacionados con la generación de inundaciones, su control y manejo, han sido tratados con herramientas tradicionales de modelado hidrológico enfocados al estudio y análisis de la relación precipitación-escorrentía, proceso físico que es impulsado por el ciclo hidrológico y el régimen climático y este esta directamente proporcional a la generación de crecidas. Dentro de la disciplina hidrológica, clasifican estas herramientas de modelado tradicionales en tres grupos principales, siendo el primer grupo el de modelos empíricos (modelos de caja negra), modelos conceptuales (o agrupados, semi-agrupados o semi-distribuidos) dependiendo de la distribución espacial y, por último, los basados en la física, modelos de proceso (o "modelos de caja blanca", y/o distribuidos). En este sentido, clasifican las aplicaciones de predicción de caudal fluvial en la ingeniería de recursos hídricos en dos tipos con respecto a los valores y parámetros que requieren en: modelos de procesos basados en la física y la categoría de modelos impulsados por datos.

Los modelos basados en la física proporcionan una descripción detallada de la dinámica relacionada con los aspectos físicos que ocurren internamente entre los diferentes sistemas de una cuenca hidrográfica determinada. Sin embargo, aparte de ser complejos de implementar, se basan completamente en algoritmos matemáticos, y la comprensión de estas interacciones requiere la abstracción de conceptos matemáticos y la conceptualización de los procesos físicos que se entrelazan entre estos sistemas. Además, los modelos impulsados por datos no requieren conocimiento de los procesos físicos que gobiernan, sino que se basan únicamente en ecuaciones empíricas que necesitan una gran cantidad de datos y requieren calibración de los datos en el sitio. Los dos modelos difieren significativamente debido a sus requisitos de datos y de cómo expresan los fenómenos físicos. La elaboración de modelos hidrológicos para el pronóstico de inundaciones ha dado grandes pasos, pero siguen sin resolverse algunos contratiempos importantes, dada la naturaleza estocástica de los fenómenos hidrológicos, es el desafío de implementar sistemas de pronóstico fáciles de usar, reutilizables, robustos y confiables, la cantidad de incertidumbre que deben afrontar al intentar resolver el problema de la predicción de inundaciones. Sin embargo, en las últimas décadas, con el entorno creciente y el desarrollo del campo de la inteligencia artificial (IA), algunos investigadores rara vez han intentado abordar la naturaleza estocástica de los eventos hidrológicos con la aplicación de algunas de estas técnicas.

Dados los contratiempos en el pronóstico de inundaciones hidrológicas descritos anteriormente, esta investigación de tesis tiene como objetivo integrar los modelos hidrológicos, basados en la física, hidráulicos e impulsados por datos bajo el paradigma de Sistemas de

múltiples agentes para el pronóstico de inundaciones por medio del bosquejo y desarrollo del marco de trabajo del sistema multi-agente (MAS) para los eventos de predicción de inundaciones en el contexto de cuenca hidrográfica tropical.

Con la aparición de las tecnologías de agentes, se han emprendido algunos enfoques de simulación recientes en la investigación hidrológica con modelos basados en agentes y sistema multi-agente, principalmente en alerta por inundaciones, seguridad y planificación de inundaciones, control y gestión de inundaciones y pronóstico de inundaciones, todos estos enfocado a simulacros de evacuación, y este último no dirigido a la cuenca tropical, cuyo régimen hidrológico es extremadamente único.

En este enfoque de entorno de modelado de cuencas, se aplican los enfoques de sistemas multi-agente como un sustituto del modelado hidrológico convencional para construir un sistema que opera a nivel de cuenca con estaciones hidrométricas desplegadas, que utilizan los datos de redes de sensores hidrométricos (por ejemplo, lluvia , nivel del río, caudal del río) capturado, almacenado y administrado por una organización de agentes interactuantes cuyo objetivo principal es realizar pronósticos de caudal y concientización para mejorar las capacidades de soporte en la formulación de políticas a nivel de cuenca hidrográfica.

La primera sección de este documento analiza el estado del arte sobre la investigación actual en modelos hidrológicos para la tarea de pronóstico de inundaciones. Es un viaje a través de los antecedentes preocupantes relacionadas con el proceso hidrológico, las ontologías de inundaciones, la gestión y la predicción. El apartado abarca, en cierta medida, las técnicas, métodos y aspectos teóricos y métodos del modelado hidrológico y sus tipologías, desde los modelos convencionales hasta los prototipos de inteligencia artificial actuales, haciendo hincapié en los sistemas multi-agente, como un enfoque de simulación reciente en la investigación hidrológica. Sin embargo, se destaca que esta sección no contribuye a una revisión integral, sino que su propósito es servir de marco para este tipo de trabajos y una guía para subrayar los aspectos significativos de los trabajos.

En la sección dos del documento, se detalla el marco de trabajo propuesto para el sistema multi-agente para el pronóstico de inundaciones. Los trabajos realizados comprendieron el diseño y desarrollo del marco de trabajo del sistema multi-agente con la arquitectura (modelo Creencia-Deseo-Intención) para la predicción de eventos de crecidas dentro del concepto de cuenca hidrográfica tropical. Las contribuciones de esta arquitectura propuesta son el reemplazo del modelado hidrológico convencional con el uso de sistemas multi-agente, lo que agiliza la administración de las series de tiempo de datos hidrométricos y el modelado del proceso de precipitación-escorrentía que conduce a la inundación en el curso de un río. Otra ventaja es el entorno amigable proporcionado por la interfaz gráfica de la plataforma del sistema multi-agente propuesto, la generación en tiempo real de gráficos, cuadros y monitores con la información sobre el evento inmediato que tiene lugar en la cuenca, lo que lo hace fácil para el espectador con algo o sin experiencia en análisis de datos y su interpretación para tener una idea visual de la información disponible con respecto a la cognición de las inundaciones.

Los agentes necesarios desarrollados en este marco de modelado de sistemas multi-agente para el pronóstico de inundaciones han sido entrenados, probados y validados en una serie de tareas experimentales, utilizando la información de la serie hidrométrica de datos de lluvia, nivel del río y flujo del curso de agua recolectados por los agentes sensores hidrométricos de los sensores hidrométricos de campo.

Acknowledgments

On an opportune time like this one, I want to take the advantage to seize the occasion and show my warmest thanks toward my thesis research director Dr. Agapito Ledezma Espino, for his constant supervision, patience, and incentive during this experimental approach. in the absence of his instructions and effective encouragement, this experimental approach may have not come to fruition. I also extended my indebtedness to my Co-director Dr. Juan Antonio Gómez Herrera for his huge assistance, valuable guide, encouragement, and counsel.

Toward Universidad Carlos III de Madrid and to all the staff from the pertinent Departments who had their administrative involvement, be it directly or indirectly.

I also acknowledge "Secretaría Nacional de Ciencia, Tecnología e Innovación (SENACYT)", on behalf of the grants received and for providing much support in conducting this research.

Many thanks to the CAOS group, who were helpful to me in many ways, be it knowingly or unknowingly. To my colleagues, I made during my stay in and out of campus.

Finally, extend my wholehearted thanks to my siblings, especially to my mother, sisters, and my dearest acquaintances, for providing persistent reassurance, empathy, and generously supporting me in the course of this Ph.D. endeavor.

Publications

This section summarizes a list of accepted publications resulted during this research:

- **J. A. Simmonds**, J. A. Gómez, and A. Ledezma, "Forecasting sea level changes applying data mining techniques to the Cristobal Bay time series, Panama," J. Water Clim. Chang., vol. 8, no. 1, pp. 89-101, 2017.

- **J. Simmonds**, J. A. Gómez, and A. Ledezma, "Data Preprocessing to Enhance Flow Forecasting in a Tropical River Basin." In: Boracchi G., Iliadis L., Jayne C., Likas A. (eds) Engineering Applications of Neural Networks. EANN 2017. Communications in Computer and Information Science, vol 744, pp.429-440. Springer, Cham, 2017.

- **J. Simmonds**, J. A. Gómez, and A. L. Espino, "Knowledge Inference from a Small Water Quality Dataset with Multivariate Statistics and Data-Mining," AACC'17 2017 Adv. Inf. Commun. Technol. Adapt. Agric. to Clim. Chang., vol. 1, pp. 1-15, 2017.

- **J. Simmonds**, J. A. Gómez, and A. Ledezma, "Statistical and Data Mining Techniques for Understanding Water Quality Profiles in a Mining-Affected River Basin," Int. J. Agric. Environ. Inf. Syst., vol. 9, no. 2, pp. 1-19, 2018.

- **J. Simmonds**, J. A. Gómez, and A. Ledezma, "The role of agent-based modeling and multi-agent systems in flood-based hydrological problems: a brief review". Journal of Water and Climate Change, 2019.

- **Jose. Simmonds**, J. A. Gómez, and A. Ledezma, "Fuzzy Logic Based Agent for Flood-Awareness". XIX Conferencia de la Asociación Española para la Inteligencia Artificial 2021.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This first heading gives an overview of the exploratory settings, followed by Section 1.2 is the approach to the issue for this activity, and Section 1.3 is its motivation. The general approach is presented in Sections 1.4 and 1.5 respectively. From a Hydroinformatics perspective, the problem statement and multi-agent system approach presented in this work are focused on a broad overview, however, the experimental method carried out applies to flow forecasting from the information obtained by hydrometric sensors. Finally, Section 1.7 describes the arrangement of the rest of the document.

1.1 Research Viewpoint

Hydrologic related problems observed worldwide on the changes in environmental conditions are becoming increasingly frequent and pose major threats to human lives and economic development [1, 2]. In addition, the number of communities being affected by climate change and its related disasters is increasing, especially in highly dense populations, and an increase in industrialization and rapid urbanization is in demand [3, 4].

From a river basin perspective, hydrological related problems, especially those induced by floods, have had severe impacts beyond geographical boundaries. For example, in territories of Asia and the Pacific, floods of 2014 accounted for damages of almost $26.8 billion in economic impacts and resulted in 3,559 deaths as reported in [5] and in 2015 accounted for two-fifths of all disasters in the region, causing over $11 billion in economic damage [6].

The worldwide increase in weather-related phenomena, as global warming continues to observed rising temperatures, and similarly, causing river flood risk distribution to increase unevenly, throughout regions like Asia, the US, and Europe [7] has become evident. Thus, it is clear that water-related disasters are increasing exponentially and present serious concerns to governments. It is important to notice that flooding is not restricted to specific countries and often takes no preference for people it affects, as is the case for recent flooding events in the USA, Australia, India, Iraq, Thailand, and Cambodia [8]. For these reasons, the need for efficient and precise river flow forecasting tools has rapidly increased during the past ten years. Conversely, the understanding of future conditions on surface water resources is a valuable asset for the development and management of proper flood risk mitigation and sustainability of water resources management. Traditional flood management systems composed of hydraulic structural protection measures such as terracing, flood-ways, inundation ponds, dams, weirs, levees, barriers, dikes, embankment, and other structures make up the most common structural solutions to reduce the flood peak, stages, and extent of inundation. One concern of structural measures is that although they reduce flood risk, they are expensive, and sometimes they are not a guaranteed effective solution. Besides, structural measures are not practical for installation in some areas (e.g., inaccessible areas, inaccessible roads, and mountainous regions), and may not be effective for all flood processes and they can also generate unfavorable environmental effects [9]. Of the two, non-traditional flood mitigation methods, the first is of environmental nature (e.g., replanting, soil management, bank stabilization, and re-vegetation, river training, and flood plain restoration and administration). Notwithstanding, measures like these are far-reaching and costly and do not necessarily reduce flood loss. The second group involves implementing river flood warning systems.

Subsequently, researchers such as Gleick [10], Brooks et al. [11], Opperman [12], and Son et al. [13] suggested shifting from traditional to non-traditional defense methods to ease the hazards of flooding, involving the governance of land development and enhance flood prediction, in densely populated areas, specifically with high risks of flooding. Therefore, they describe a flood event as a condition, whether incomplete or complete, in which the flow continuously exceeds 80 percent of flows observed during a time. Floods can be of various categories and dimensions. They represent the building block for the different designs and operations of inundation prediction techniques. Flood prediction systems have been in operation for various scale domains, which include global [14], continental, basin-scale [15], and urban [16].

Traditionally, the flood problematic, as being dealt with the use of rainfall-runoff modeling tools, can be classified in three main categories: a) empirical (or so-called black-box models), b)conceptual (or so-called lumped, semi-lumped or semi-distributed), and c) physics-based process models (PBPMs) (or white box models, also known as distributed models) as suggested by Sitterson et al. [17]. In this sense, they classify river flow forecasting applications in water resources engineering into two types regarding the values and parameters needed in physics-based process models and the models of the data-driven modeling (DDM) type [18]. PBPM describe the physics involved in the processes occurring within the catchment by mathematical equations, by connecting empirical and physics-based mathematical formulations. Besides, DDMs do not require knowledge of the governing physical processes; but rely solely on empirical equations that need lots of data and require on-site data calibration. Generally, the two models differ in their data requirements and the manner in which they express the physical phenomena [19].

The number of catastrophic floods that are occurring around the world today is attributed to climate change, this motivates the need to develop sturdier and "intelligent" flood protection systems to enhance operational flood forecasting that is resulting in many projects for protecting, coastal zones from high water surges events which can hinder maritime operations [20–24], urban areas [25–28], infrastructure and the population [29]. These problems call for the design of more sophisticated early warning systems (EWS) and robust databases (DB). Yet, they can be very expensive and one of the most challenging tasks to undertake for flood prevention and disaster management.

An example of tools for implementing EWS is the Urban Flood project, a European initiative that implemented a framework for EWS with the purpose of connecting sensors to real-time models for flood forecasting and warning; however, validation for this system was only for dikes. Some of these EWSs implemented are localized, custom-designed and need high computational resource [30].

As the world is becoming more complex, the hydrological processes in parallel with changes in the global climate are also increasing in complexity, leading to weigh the systems already existing, to optimize, and investigate for the implementation of new user-friendly modeling prototypes with a predictive capacity which can deal with mayor complexities concerning the interactions with the increasing complications of the hydrological system. This global change in the increase in complexities could also mean that conventional hydrologic models may not be equally relevant to model these complex changes.

1.2 Problem Remarks

Today's climatic conditions worldwide are increasing dramatically and resulting particularly in hurricanes and extreme precipitation, the latter being the triggering for severe flooding problems worldwide. This situation is of great concern for governments, specifically in river basins in the tropical regions. A recent survey on weather-related natural catastrophes, showed that the number of weather-related damages, namely of (meteorological, hydrological, and climatic) in nature has been increasing globally for the last 36 years [31].

In Panama, there are several professionals whose decisions are limited because they lose sight of many of the physical aspects governing some natural phenomena that cause disasters. There is also a limitation of information and availability of hydrometric data because of the poorly gauged river basins. For these limitations, engineering professionals require a resource that links engineering information with understanding the language desirable for decision-makers (citizens, governments, and NGOs, etc.) in spite of the limitations mentioned.

This thesis will explore new methods for the implementation of a hybrid system useful for water flow forecasting based on traditional hydrologic modeling methods in combination with techniques from Artificial Intelligence (AI), and with particular attention to the use of agent technology. Therefore, confronting such a task will demand resources directly supplying the requirements necessary to comprehend and provide answers.

Then it is addressed this task following the paradigms provided by hydroinformatics techniques (viz. the fusion of water resources engineering (WRE) and information and communications technologies (ICTs), and its crossing point with hydrologic modeling and "Computer Science" (CS) disciplines (e.g., "artificial intelligence" (AI), "data mining" (DM), besides others), applied to water resources engineering, flood management, and forecasting has several tools, based on information technology, and can meet the needs of knowledge among decision-makers and enable engineers to identify those needs.

1.3 Motivation

Previously, in Sections 1.1 and 1.2, it was introduced this doctoral research proposal manuscript, reporting in a very concise way, the current global situation that has been taking place concerning the evident meteorological phenomena and climate change in the world. So much so that the research motivation is based on that region of the planet that is bound by the "cyclone and hurricane belts" and yes indeed, because it is situated on the border of the equator, the "tropics". The United Nations Office for the Coordination of Humanitarian Affairs, OCHA, reported that it focused on this particular region and declared the countries in this region "Natural disasters in Latin America and the Caribbean" as the region with the second most vulnerability to disasters in the world, and there is a growing trend in the severity and occurrence of stronger storms that is affecting the region, and as a consequence originating intensified precipitation and augmented high flood [32]. The report disclosed, that at least, for the past 20 years, countries like Mexico, Cuba, and Haiti, had over some 110 cloudbursts, with deaths amounting to about 5,000, with 29 M victims losing their goods and some $39 billion in losses. Unfortunately, Haiti had the highest death toll, with 85%. However, these events as natural as they can be, are recurring and the poorest of the population suffers the most.

As the sources have been reporting, it can be observed that one of the most expensive natural catastrophes occurring is categorized as floods. As presented in the introduction, they are responsible for causing a wide variety of calamities and much of the damage extent to property, buildings infrastructure, destruction of natural habitats, grazing and agricultural land, the economic goods, human loss of properties, and lives and impairment to environmental health. However, as small as a flood event can be, the devastating force, in addition to timing and uncertainty is exerted by the worse kind of floods, which is known as "Flash Floods". Flash floods are so destructive and deadly because they occur with almost no previous warning and eventually take the population by surprise and unaware.

According to major studies by the World Bank on natural disaster sites, in many Latin American countries, the population is exposed to multiple threats, and most of it is due to a lack of mechanisms for development, monitoring, planning, and low compliance with building regulations, and land use is some factors identified as aggravating the vulnerability of these countries to disasters [3, 4].

Most of the information on climate change is compelling, and it is telling us that there is a higher risk of flood-related problems, as shown in a study by Hirabayashi et al. [33] in which they analyzed the simulation results of eleven climate models. Results showed that the observed changes in streamflow and the flood plains by several climate change simulations may cause rising inundation in Southeastern, and African countries and vast regions of Central and South American. However, fluvial flooding is not the only climate-related problem in the tropics, but also a high-water surge affecting coastal shorelines, coastal maritime infrastructure, and low land areas densely populated [24, 34].

The severity of a flood, along with the degree of preparedness and national capacity for self-assessments to react to the problem, will entirely affect how well the forecasting system responds, giving enough time to managers and first responders if needed to issue evacuation plans, and in this way, save lives and reduce impacts. To some extent, flood warning systems have shown a great deed in the reduction of hazards and damages caused by inundation episodes. Disappointingly, in many of the basins in Latin America, there are no governmental organizations that are dedicated to the sole task of permanent surveillance regarding the forecast of river flooding. This is possible because of a lack of resources, computer and electronic materials, and data availability. In most cases, these countries only have or make use of meteorological information on the web that provides generalized services on flood management without themselves carrying out hydrological and hydrodynamic studies of their rivers with the data. However, if they do any study at all, most of it is on rainfall information. In some circumstances, they consider the implementation of hydrological models and analysis for flood forecasting a technical activity, isolated only for the specialist, and this implies the lack of contact with government entities and society. Moreover, forecasting the future of floods is very important for decision-making.

Taking into consideration these problems, in this research study, it is intended to investigate the use and implementation of MAS model based on an ABM framework in the application of flood forecasting in the humid tropic domain in place of standard hydrologic models to facilitate the support of decision-making of water resources professionals with little or no knowledge on hydrologic modeling.

1.4 Objectives

The goals of this research are aimed at designing and building a "Multi-agent System" (MAS) configuration for forecasting flood events within the context of the humid watershed. The research entails the following specific objectives:

1. To setup and run the hydrological and hydraulic model simulation with the selected catchment hydrometric data.

2. To design the architecture for a domain-specific agent-based concept framework for supporting stream-flow predictions in the context of a tropical watershed domain.

3. To develop a MAS framework from the designed architecture on the selected agent platform, and implement the system using an integrated approach of hydrological, hydraulic and data-driven modeling, and artificial intelligence techniques for flood forecasting.

4. To implement the belief-desire-intention (BDI) organizational structure for modeling of cognitive agent actions, communication, and their interactions within the MAS to manage the data obtained from the framework tool application.

5. To simulate, calibrate, and verify the MAS outputs against the results from the hydrologic and hydrodynamic models.

6. To test and validate the model by simulations and field deployment.

1.5 Research Questioning

Grounded on the problematic previously described and the need to create systems to deal with it, the current doctoral project raises the following research question:

¿How to integrate physics-based hydrologic, hydraulic, and data-driven models under the paradigm of "Multi-Agent Systems"?

A "multi-agent" conceptual framework established on the knowledge of standard hydrologic modeling techniques, DDM, and AI techniques and algorithms is recommended to cope with this problem, will allow to effective management and use of field-based hydrometric sensor data, for flood forecasting by water managers and professionals, and the layperson in the context of tropical River Basins without them being subjected to the computational costs and hassles of hydrologic modeling development and deployment.

1.6 Contributions

The thesis contributes to developing a flood forecasting framework solution, as shown in Figure 3.1 for a tropical river basin by bridging the domains of traditional hydrologic simulation, data-driven modeling, artificial intelligence, and agent technology like the "agent-based modeling and simulation (ABMS)" and the MAS which will have the capacity of using the acquired data from hydrometric sensors and their processing and allowing the visualization in real-time of flooding in a river basin on a GIS (geographic information system) environment. Therefore, some core facets of the solutions include the use of hydrometric sensor's time-series data for the simulation of flood forecasting and flood-awareness (FA) level inferenc with the implemented MAS. The different contributions related to this research endeavor are:

1. building the MAS simulation on a GIS enable ABM environment of choice and its displayed agents to recreate flood forecasting.

2. Implementing the MAS data acquisition for the displayed hydrometric sensors.

3. Hydrometric sensor data pre-processing.

4. Development of an optimal flood forecasting machine learning model and Artificial Intelligence (AI) to predict the risk of flooding.

5. Flood forecasting and risk assessment with fuzzy expert systems.

1.7 Outline

To provide a thorough insight into the research problem, existing approaches, besides projected objectives, this thesis covers the following topics:

- **Chapter 2:** Status of current research. This section outlines recent advances in agent-based technologies as promising modeling paradigms for solving complex problems, as well as traditional hydrologic modeling techniques, data-driven models (DDM), and existing approaches in the literature review to AI.

- **Chapter 3:** MAS Framework for Flood Forecasting. Explains the abstract MAS framework built for streamflow forecasting, like the solution designed to address the flooding problem. It also describes the MAS framework, that is built on the GAMA platform, upon which the agents implemented become instantiated (agent organization), communicate, and interact within the problem domain; also describes how agents execute internally (agent behavior), following the BDI architecture and the FIPA-ACL specifications communication protocol.

- **Chapter 4:** Preliminary Validation scenario: Rio del Medio Sub-watershed. Presents the experimental setup of the ABM simulation scenarios on the GAMA platform used to validate the system (the Medio River Sub-watershed). The simulation domain code is written in the GAML language, which is specific to the GAMA platform. To verify the correct execution of the code in hydrological terms, the performance was subject to analysis.

- **Chapter 5:** Hydrologic Modeling with MAS for Flood Forecasting. Introduces the developmental tasks of the MAS exemplary framework, with the BDI enabled agents, with behavioral skills, for example, such as data preprocessing, machine learning, fuzzy logic, and SQL with hydrometric time-series data as input files for lead-time flow forecasting and its application in an abstract watershed implemented on a GIS-enabled agent-based modeling platform.

- **Chapter 6:** Key Conclusions, and perspective for future work. Gives a summarized feedback on the outcomes of the research, model simulations, results, and findings, and suggestions for future work.

Chapter 2

Status of Current Research

2.1 Background: Analogous Work

This chapter, in Section 2.1, begins briefly by explaining the hydrological phenomena as the main driving force responsible for the impact of floods on society, affecting the economy, agriculture, civilian life, and live-stock, and its relationship with the influences of climate change. Afterward, it presents the views of flood management strategies and forecasting adopted by several European countries, Canada and the USA with a detailed discussion about measurements used to counteract flooding. Last, Sections 2.2 through 2.4 starts with background recounts on some research in the area, and end with a presentation on existing approaches concerning standard methodologies applied in flood forecasting, methods from artificial intelligence, and recent attempts with agent technologies.

2.1.1 Hydrological Processes

The hydrological process is a natural phenomenon that expresses the physical interactions between the different phases of the water cycle. These processes can take place in systems such as lakes, rivers, seas, and oceans, and interact at discrete spatio-temporal dimensions. There are two methods for modeling the hydrologic system; one is by physics-based models and the other by data-driven modeling approaches. For physics-based models, the details of the dynamics of their behavior are given by the governing mathematical equations of the system, which require a detailed description of the initial values. This type of model is founded on abstractions of the generalized forms of "Newton's laws" (i.e., inertia, momentum) and applications of the first law of barodynamics, which define the generic equations of conservation in their abstracted form. Therefore, the input minus the output quantities of a system may be expressed in mathematical terms as:

$$O\left(\tau\right) = \Omega\left[I\left(\tau\right)\right] \tag{2.1}$$

then, the term τ is the period of duration of the phenomenon, $O(\tau)$ is a conservative magnitude leaving the system, and $\Omega[I(\tau)]$ represents the transform function that relates the system entry $I(\tau)$ concerning its output $O(\tau)$. The transform function Ω denotes the system's properties (e.g., linear, nonlinear, random, and non-random, allocated, and non-allocated). In the former models, Ω is centered on applying the knowledge of physics that governs the environment domain. Whereas, the later models implement empirical knowledge about Ω to gather experience from the "performance of the data", i.e. gain experience based on intake and exit information.

Of those events caused by the hydrologic phenomena (e.g., geomorphological, climatological, floods, and landslides) flooding events handle significant economic damage worldwide. Hence, in catchment management, the issues of flooding and prediction are earning increased attention as inundation problems and their extremes are increasing and becoming gradually frequent across the world. This situation has triggered major concern for many governments given the increasing events related to disastrous flooding. Therefore, the necessity of establishing mechanisms for early warning to support flood forecasting is in demand. Since then, several researchers have claimed that in the years to come, the events unfolding are going to become worsened because of climatic fluctuations [35].

For managing river flooding, prediction models of hydrological flows are essential, and as such are of paramount concern for managers and hydrological researchers. Notwithstanding, in most cases, it is almost difficult to avert their effects, so the implementation of simulations tools capable of forecasting the ever-occurring threats of floods needs to be reliable and accurate. In this respect, the assessments should be composed of early forecast systems that can provide knowledge about the scope of the event.

Then again, as was noted earlier, severe flooding is also considered responsible for a great deal of loss of both human and animal lives, lands, homes, properties, and damage to the economy, specifically in areas that are densely populated. Furthermost, flooding events are attributed to be the greatest damaging events of all-natural risks, on civilians across the globe. On the other hand, unfavorable effects of flood disasters encompass such problems as the spreading of waterborne infections; landscape devastation, farmed lands, cattle grazing lands, homes, and impairment of the water quality, water pollution by sedimentation of water streams; losses to the flora and fauna habitat. Besides, the increasing variability in rainfall extremes, is responsible for the impairment of stream water pollution, as runoff coveys polluted substances from overland and agricultural point sources [36], [37].

2.1.2 Flood Phenomena Ontology

In the development of software structures, an essential topic is the definition and application of ontologies [38, 39], and in computational sciences, it is an idea that emerged from AI to transfer information and experience as showed by Gruber [40].

In solving any information problem, the three main criteria of "Semantic Web Technology" are "Protocol and RDF Query Language (SPARQL)" and "Resource Description Framework (RDF)" which form the basis of "RDF Query Language" and the widely used "Web Ontology Language (OWL)". Hence, comprehensive exposure to the different approaches used in ontology development is available in the works of [41–45].

The ontologies offered within the domain of the "flood phenomena" found in the literature, except for a few; not oriented necessarily to the flood forecasting task, since most of these flood-related ontologies have their classes and properties focused on flooding as an environmental hazard, risk, and management. Notwithstanding, these ontologies, although built from scrap or reused from other ontologies, and presented in various types, have their foundation on well-known ontologies like the "Semantic Web for Earth and Environment Technology Ontology (SWEET)" [46], the "Semantic Sensor Network (SSN)" [47], the "Time Ontology" [48], and some non-generalizable "Hydraulic Ontologies" and several others. However, for a comprehensive and systematic analysis of flood ontologies see Sinha et al. [49].

From these ontologies, only a few [50–56]) have examples of classes and properties specific to the domain of interest (i.e., forecasting and sensor network), so they can be reusable in this work. Therefore, it is not intended to implement from scrap new ontologies regarding the flood forecasting task.

2.1.3 Flood Management and Forecasting

State-of-the-art flood management and containment suggest the procedures and techniques aimed at reducing or preventing the hazardous consequences of floods. However, releasing floodwaters (flood relief) is an engineering approach (i.e., applying structural or non-structural measures) employed for the control of flood surges.

European countries lying low below sea levels like the Netherlands and Belgium are at the top front with the construction of such infrastructures. In this sense, many efforts by several research institutions, governments, and academia have started projects like the "FLOOD RELIEF" [57], "FLOODsite" [58–61], "UrbanFlood FP7" [62] and "FloodControl 2015" [63, 64], just to mention a few, are some efforts carried out as flood "Early Warning System (EWS)" implementations for protection against flood. In Canada and the US, it can be found the construction of a diverse set of flood control and relief structures such as floodways, dams, and levees as the main structural means of defense used. In Latin America, with flood management and forecasting, it is most likely to see the absence of structural defense measures due to their high cost of construction.

Implementing EWS possesses many challenges and deficiencies. Despite these deficiencies, climate change is an observed growing trend, and the vulnerability of the water resources is becoming more affected and the assessment measures for containment, prevention and mitigation are some of the ultimate tasks faced by river flood management today. In this sense, the implementation of new EWS prototypes is the core of the challenges faced for tomorrow's ever-changing climate, and this thought is continually attracting the interest of water engineers, scientists, and managers.

2.2 Related Research

This work is intended to apply agent-based technology, specifically "multi-agent systems", to support simulations for forecasting floods in tropical watersheds. Using agent-based technologies is a highly active research paradigm within the social, geographical, and ecological sciences where a great deal on the subject has already been written. Notwithstanding, as noted earlier in the previous chapter concerning traditional hydrologic methods for dealing with flood-related problems and which are well documented, it begins the discussion of the related research by addressing the flood forecasting issue with a short explanation of the state-of-the-current-research using conventional hydrologic modeling, its data-driven, and artificial intelligence counterparts. Finally, the last two subsections discuss the use of agent technologies (ABM and MAS) as a new approach in the role of flood forecasting simulations.

2.2.1 Hydrologic Modeling

On a watershed level, various types of hydrologic (also known as catchment/watershed) models are tools applied to simulate overland flow and river flow routing produced by precipitation. So, in the process of developing these models earlier, today more are being developed, there will be various ways and approaches taken to classify hydrological models. Thus, it is presented a brief classification of these models that assumed their governing physical rule has shown by Jain et al. [9]. Table 2.1 and Figure 2.1 shows a depiction. According to their applications, it classifies them based on process, spatial representation, or randomness, and degree of detail, which means there will be various ways and approaches taken to classify hydrologic models. From the existing classification, some researchers have proposed several alternative classifications in the literature [65–72]. A brief classification of these models assumes the physical rules which govern and categorized in many types depending on the modeling approach used, their function, goals, their structure, and level of spatial discretization as an empirical, conceptual, and physics-based process models subdivided into various specialized sub-fields (Table 2.1).

As the special distribution of a catchment is important when selecting and developing hydrological simulation models (Figure 2.1), a twofold methodology can be applied to implement such simulations. In the first methodology, the main idea is to run through and feed the collected dataset having the nature of time series through probabilistic models, by which the variables of interest are forecasted. In the latter method, it is assumed that the variables, parameters, and constants that are intervening in both hydrological and hydraulic processes are known. Therefore, this information becomes crucial for the formulation and implementation of models, as can be noted in [73] and improved by [74].

As noted, simulation prototypes for hydrological processes are diverse, and in the same manner, can be presented from simple to the most complex models. Of course, their use varies according to the process being studied and the problem to be solved. In addition, there is the fact that, because of their great diversification, they can couple it to other systems, which make them capable of even simulating the dynamic processes of water resources quality pollutants in catchments [75]. Below it can be seen in more detail the classification of these models.

Table 2.1: A nomenclature categorization in hydrological tools. (Adapted from: [9].)

Principle	**Classification**		
Geographical location	Clustered: black box setup	Quasi-allocated model	Totally allocated model
Watershed Simulation Method	Determinist *Trial/Error *Theoretical +Clustered +Allocated *Physics-oriented +Mesh-oriented +"HRU" oriented +Subwatershed oriented	Data Driven *Stochastic	Data Driven *ANN *Fuzzy
data feed and watershed area	Riverine route models *Hydrological routing *Hydraulic routing	Catchment models	Combined catchment and routing models
Rainfall prediction and revision	No rain prediction No revision	Rainfall prediction Model revision	Radar-nowcast Model revision

2.2.2 Deterministic Models in Flood Forecasting

They characterize deterministic models as models that use mathematical formulations and consider the correlation between what is entering and leaving the modeling system. An example is a regional model for simulating the "River Severn" catchment, which could incorporate both surface and groundwater components, as shown by Liddament and Oakes [77]. They applied this model to the operational control for regulating the stream. The researchers argued that because of the complexity of the river flow, the model is complex, requiring large computational time and others may render it costly to build and difficult to work with. One of the characteristics of deterministic models is that the non-linear partial differential equations used for describing hydrological processes can be implemented. In this sense, it should be pointed out that solutions to the analytical operations cannot be solved by the equations. There is a benefit offered by deterministic approaches to exhibiting an insight into the internal processes that agree with a broad understanding of hydrological processes. Deterministic approaches usually yield for any model input, one unique output. Table 2.1.described a categorization of these models.

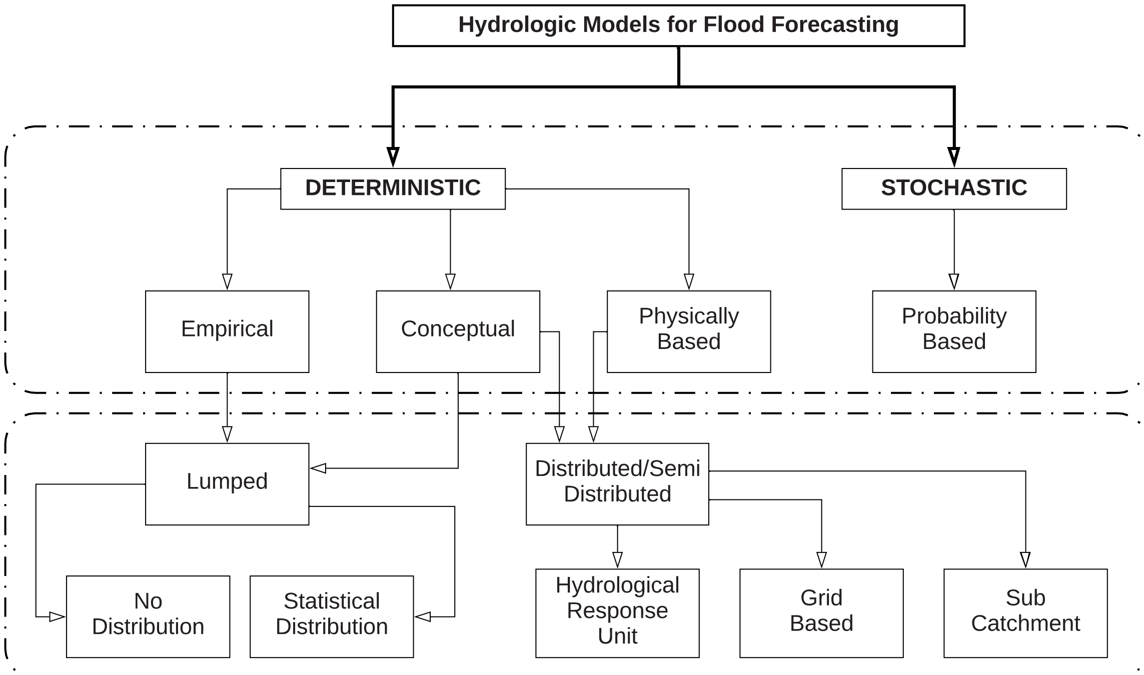


Figure 2.1: Taxonomy for hydrological simulation prototypes. Source: [76].

Following the classification indicated in the Table 2.1, deterministic approaches can also be considered as empirical representations, which becomes useful when there is no knowledge about the system to be modeled; this means, the catchment is considered as a whole (a lumped system) and the approximation made of the physical values have resulted from a grouped modeling approach over the entire catchment. In other words, this methodology entails the so-called "black box" modeling scheme, as hardly any information on the dynamics of the precipitation-runoff process is thoroughly known [78, 79]. Therefore, these models rely entirely on observed data, their validity, and precision [80]. Thus, given their high requirements for data availability, empirical models are also called data-driven models [81, 82]. However, in continuation with the discussion on hydrologic models, the simulation of its dynamics can be accomplished, aside from the approaches formerly mentioned, be it by deterministic, or empirical means, the models can also take on a conceptual or physics-based approach, and these resulting from the synergy of combining several approaches, they can obtain an amalgamation or sub-models, designed and suitable for simulating the hydrologic process considered as significant as shown by Bergström and Forsman [83], Kirby and Beven [84], Jakeman and Hornberger [85], Scharffenberg et al. [86], and in [87–89] respectively.

2.2.3 Data Driven Flood Forecasting Modeling

Among the models used for simulating the rainfall-runoff process, data-driven modeling (DDM) is a modeling method that connects the state variables present in the process (e.g., external, internal input/output parameters, like the rainfall-runoff relationship) and builds a model with the corresponding field-collected data [90, 91]. This doesn't imply that initial knowledge about the details and processes that govern the system's behaviors is not required; notwithstanding, data-driven models are regularly very useful when this knowledge is unavailable *a priori*. Data-driven methods also encompass other simpler theoretic probabilistic techniques that are not founded on complicated physical laws. However, some approaches in this line of analysis are present in earlier techniques of the linear time series, "auto-regressive (AR)", and "mixed-auto-regressive with moving average (ARMA)" as reported in [92], and in techniques from computational intelligence like "artificial neural networks (ANN)", "support vector machines (SVM)", "fuzzy logic (FL)", and hybrids, like the "adaptive neural Fuzzy Inference Systems (ANFIS)". Contrary to its "hard" probabilistic technique's counterpart, computational intelligence techniques are unbiased with inaccuracy and improbability to converge to better results [93].

Stochastic models are constructed from one sample known in time as initial condition (i.e., precipitation time series) which are based on minimum criteria of the errors, and from which estimates are derived from its past values [92]. Therefore, they rely mainly on the time series analysis and forecasting methodology, which use in hydrology is very popular [94]. Since stochastic models are shown to handle random processes, this quality makes them useful for modeling precipitation data, given their random nature. In this sense, to deal with the stationarity or non-stationarity properties of precipitation data; approaches such as the ARMA and ARIMA models can help elucidate the characteristics of the physics that engage the rainfall-runoff process at large spatio-temporal scales. Therefore, one of the reasons for adopting this alternative in hydrological simulation is due to its numerical simplicity, low cost of its computational requirements, as well as its ability for hydrographs approximation, and are useful in parameter estimation and correcting error giving the uncertainty related to other models. However, their limitations show problems with extrapolation, they require adequate and reliable data, and cannot reproduce any changes in the system.

Despite their ease of implementation, mathematical simplicity, and abilities with hydrograph approximation, stochastic models cannot deal with the non-linearity that is common in the hydrologic process [95]. To cope with this problem, researchers have pursued to implement models that integrate non-linearity of the hydrologic system. Such quests are found in other examples of data-driven models for rainfall-runoff simulations, like "Artificial Neural Networks (ANNs)" as documented in [96, 97]. The theoretical approach of ANNs modeling is nonparametric (i.e., data-driven) and as such uses a self-adaptive method to deal with information inputs and can represent complex nonlinear relationships [81, 98]. ANNs can learn from input data, employ gradient-based learning algorithms, generalize the behavior of data, and handle noise. A complete survey on ANNs applied in rainfall-runoff modeling is reported in [96, 97, 99, 100]. An interesting issue observed when conducting rainfall-runoff simulations in real-time over an entire basin, with the basin responding under the modeling of both use of conceptual and ARMA models, is the performance of the data-driving modeling approach in optimizing forecasts results. In this sense, in [101] is found an example of such model coupling between ARIMA and ANNs to significantly enhance streamflow predictions and profile revision.

Even though ANN is a reliable approach for modeling hydrologic processes, it still has some disadvantages. However, when dealing with hydrological time series that consists of extreme missing and impaired data instances, they can reduce the forecasting performance of the ANN [102]. However, with high non-stationary rainfall time-series, without the proper data treatment, the concealed frequencies in the rainfall data may inhibit the ability of the ANN's performance [103].

Among the DDM's, rainfall-runoff modeling has found a place with the use of "Fuzzy Logic" and "Fuzzy Set" applications. While the former DDM techniques rely only on numeric values, fuzzy logic and fuzzy set approach provide the feasibility for working with the linguistic nature of non-numeric values, as they extend this feature as an expression of instructions and proofs. In general terms, this expression of instructions (rules) and proofs (facts) [104], which fuzzy systems can take on intervals between 0 and 1 [93, 105], finds an appealing solution for engineering applications in water resources engineering, where for example, a representation of the rainfall-runoff process as a real-world phenomenon is pursuing. Furthermore, to explore the potential of fuzzy logic and sets, researchers have come up with a series of methods and integration of these with other methodologies [106]. Yu and Chen [107] proposed a method adapted from the "fuzzy-rule-based" system like a technique for actualizing data and that could predict errors to improve real-time streamflow prediction for 1 and 4 units of time ahead. Kalayathankal and Singh [108] developed a fuzzy scheme for flood alarm forecasting. Concerning disaster evaluation produced via inundations, He et al. [109] presented an evolutionary algorithm based on an improved logistic map and fuzzy clustering iterative model.

The versatility of fuzzy logic in rainfall-streamflow relations has taken several steps further with hybridization, where fuzzy logic and sets when combined with one or more DDM techniques, for example, ANNs enhances the proficiencies of the modeling task. In this sense, an exhaustive survey on ANFIS can be seen by Nayak et al. [110]. An implemented "ANFIS-based" modeling system for monthly predictions of inflow to the reservoir with precipitation data as inputs to assess inflow and periodic weather forecasts to support the operational management of the reservoir was implemented in [111]. The authors argued that the implemented system when examined under situations where forecast information was or was not available, shows the model efficiency achieved when using both past and observed data, with the outputs from the weather forecasting information are better than when employing solely the previously data past information. Similarly, using an ANFIS model for a 24-hour river flow forecasting in a Turkish watershed (e.g., Great Menderes River) using antecedent flows, Firat and Güngör [112] observed that the ANFIS model was successfully good for simulating the task after validating the system against measured data. With the same line of work, Akrami et al. [113] applied a model coupled with ANFIS and wavelets for dealing with the minimization of errors in rainfall data processing overtime for rainfall forecasting purposes for the Klang River basin in Malaysia.

In summary, the statistical significance of DDM approaches in rainfall-runoff modeling is that they offer a generalization of the rainfall-runoff process in a simplified manner, without having to rely thoroughly on the physical principles of the system [114]. However, several examples show that they have several limitations such as deficiencies for parameterization [115], lack of integration of catchment dynamics in the models, absence of certain physical quantities, parameters and model simulation tends toward overestimation; and they are data demanding, just to mention a few. Nevertheless, they represent an expedient modeling choice, because of the simplicity of the modeling implementation framework in the absence of spatiotemporal details yet remaining a tool of choice among water resources management [116, 117].

2.2.4 Artificial Intelligence in Flood Forecasting

Artificial Intelligence (AI) in its broadest sense describes how a machine or object can execute similar kinds of roles that characterize human thought [118]. Moreover, for a comprehensive study in this area and some of its applications, an effective review reported in [119] offers a thorough review of AI applied to flood forecasting from 2000 through 2015. The use of AI has also been proposed as an alternative to standard methods for ecosystem simulation [120]. In the following subsections, a brief description without entering into details is provided, as this work is not solely on DDM methods for flow forecasting and the variety of methods offered by Artificial Intelligence is widely presented in several extensive reviews on the topic [9], [119], and [121].

2.2.4.1 Artificial neural network

There are several connotations for artificial neural networks (ANNs). However, among the many suggestions, they can be viewed as a theoretical representation of a numerical modeling approach that receives inputs from the external environment as signals (patterns) and images (vectors) which routed through a processing function maps information similarly to the biological neurons would. As neuron imitators, ANNs contain numerous parallel concomitant neuron units that work independently connecting each other by weighted links [122]. ANNs are an advanced simplification of numerical approaches that simulate the biological cognitive function of hominid intelligence. As an algorithmic function, in the process of training ANNs for learning patterns, the back-propagation algorithm is used. This algorithm allows changing the mathematical expression of weights in the constructed network as a means for calculating the gradient of the error function by chained differentiation.

An artificial neural network may have very specialized architectures that result in the different types of ANNs of which the most common configuration is the feedforward [123], [124]. Figure 2.2 shows schematically a simple ANN with its various components, and details of its functionality can be found in [96]

With the arrival of ANN techniques, flood hydrology entered a new facet to streamflow forecasting [125–127]. Amid the different ANN techniques; as reported in [96, 97] artificial neural networks embrace a fundamental part as an effective tool for flow forecasting in water

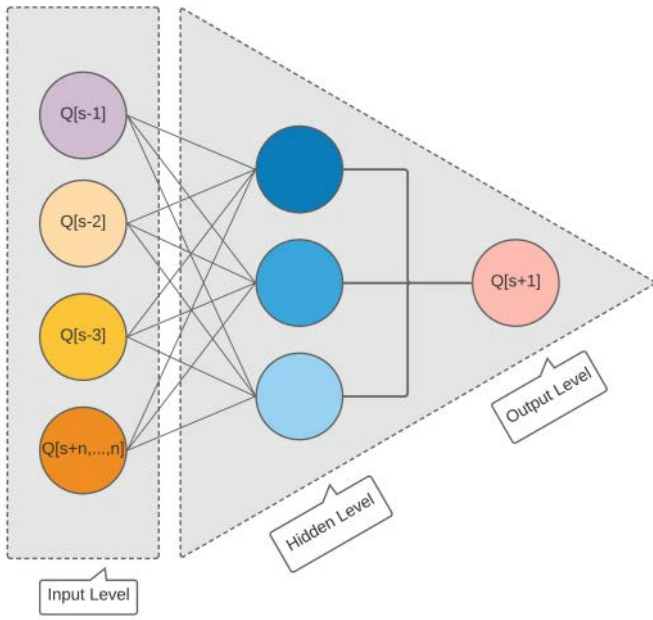


Figure 2.2: Schematic diagram of a forward propagation ANN: Representing a time series with Input Level as streamflow (Q) values for lags and lead time values, of time $\sigma =$ Q[$\sigma - 1$], Q[$\sigma - 2$], Q[$\sigma - 3$] and Q[$\sigma + n, ..., n$], and Output Level with predicted flow at lead time Q[$\sigma + 1$].

resources engineering applications. Since its beginning, the advertisement of ANN has grown and has advanced as the branch of AI whose potential for modeling the hidden processes of the complex task of rainfall-runoff cannot be ignored [128, 129]. Therefore, from the simple neuron theory, ANN as evolved into many various forms of architectures, been for example the feed-forward-back-propagation (FFBP) the configuration of choice among practitioners' engineers who seeks solutions for solving the nonlinear approximation problem [130]. The majority of the development of some of the research on ANN architecture configuration can be found in papers by [131–133]. Contrasting with its conceptual and physics-based model counterparts, ANNs are capable of reproducing streamflow observations without the need for the mathematical descriptions that define the dynamics of the process. A study by Sudheer and Jain [134] demonstrated that ANN models in river modeling were able to capture the flow duration curve during simulation time. Toth et al. [135] focused on shorter river forecast lead time with two methods other than ANN. They noted the success of hydrological forecasting to improve with ANN than compared with the other two approaches. Chang et al. [136] applied a recurrent neural network (RNN) implementation of two steps ahead for streamflow simulation. Subsequently, they took this initiative ahead updating the previous approach to forecast several steps ahead [137]. In a similar paper, Yonaba et al. [138] investigated the multi-step forward flow forecasting by exploring with several activation-functions.

Osanai et al. [139] used RBF artificial neural network to correct precipitation records that failed to issue warnings.

In flood forecasting, of crucial importance is the setting of a long-lead-time forecast that can enable the reduction and correction of the bias in the forecasting task. Flow simulation with physics-based models are subjected to errors and inaccuracies, as these phenomena are complicated, therefore, lead time forecast of environmental catastrophes should have lead-time of almost 6 to 2 weeks calendar day.

Kia et al. [140] suggested a flow forecasting model coupling artificial neural networks and geospatial techniques.

Besaw et al. [141] point out the advantages of applying ANN in the case of un-monitored catchments. The authors stressed the fact that with the chosen ANNs always lead to convergence, avoided non-linear learning schemes, and they are useful in the small catchment where gauging is relatively low or absent. Danandeh Mehr et al. [142] used a FFBP network approach as a search algorithm and later combined a GRNN with RBF network for flow forecasting in an un-gauged river basin.

Di et al. [143] introduced a machine learning technique based on KNN to treat impaired sensor data for an extended lead time of extreme rainfall. Comparably Napolitano et al. [144] suggested a conceptualization of streamflow forecasting modeling and ANN approach to deal with similar issues in un-gauged watersheds. Recently, Taormina et al. [145] proposed a methodology of optimization named "LUBE" to deal with the problems of prediction intervals (PIs) faced by artificial neural networks for operational streamflow forecasting.

2.2.4.2 Fuzzy logic methods

In the former subsection, they explored some of the applications offered by ANNs in flood forecasting. In this subsection, it explore some research with the use of "fuzzy logic" in hydrological sciences and engineering with an emphasis on engineering problems caused by floods. Why fuzzy logics in hydrology? Fuzzy logic offers the ability to deal with questions concerning uncertainties in hydrologic data, mainly on its stochastic and deterministic descriptions [105, 146, 147] For example, Klir, Klir and Yuan [148, 149] presented an effort for clarifying the description of stochasticity and determinism in uncertainties concerning the intricacies present at the time of simulations. Subsequently, in dealing with uncertainties in the models, many other outlooks on the topic have been presented in studies by other researchers, [150–156].

Typically, with Fuzzy logic and sets theory, they apply set rules to handle inaccurate and incomplete instances. This criterion has been universally accepted as a useful approach for approximating data that are fuzzy and uncertain which precedence is from intricate environments [157]. In set theory, there may be the belonging of an object to a set or not; however, when it comes to the fuzziness of fuzzy sets, these can take on true or false values, that is, values between 0 and 1. Hence diffuse models can describe ambiguous connotations similar to those existing in natural language [158]. The fuzzification procedure implicates an exact "crisp" transforms of input variables into "fuzzy sets" [159]. A fuzzy logic model is developed from preceding instructions, in combination with fuzzified values through fuzzification, extrapolation, and configuration processes, that result in the fuzzified outturn values transformed into real signal [160]. These methods comprise maxima, center of sums, weighted average, weighted sum combination, center, and centroid of area. Figure 2.3 depicts an illustration of a typical "Mandami fuzzy logic" scheme.

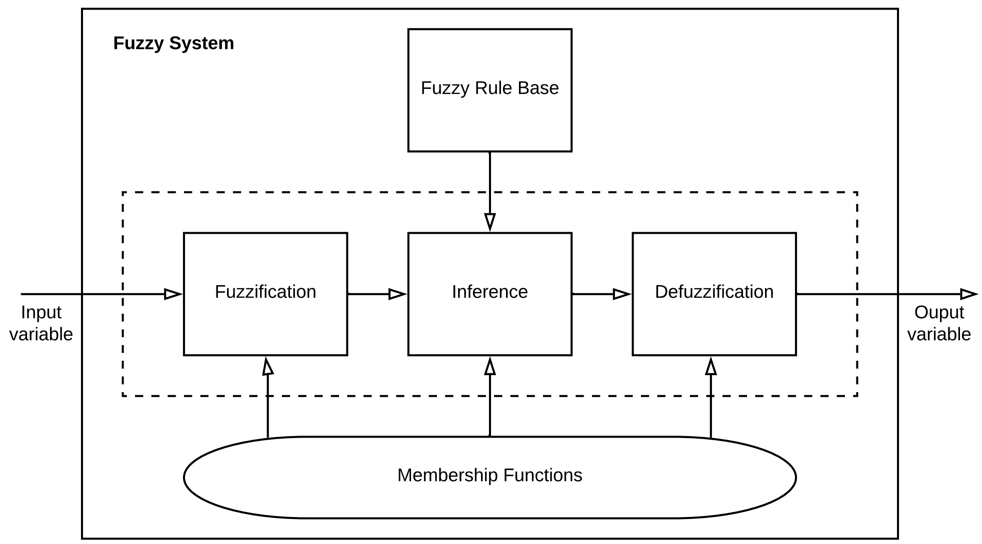


Figure 2.3: Schematic of a typical mandami fuzzy model scheme.

In Chapter Two it was emphasized the issue of the hydrological phenomena to be a complex process, therefore this refers to the nature of uncertainty and ambiguity of its behavior. So far, it has been observed that the success in hydrological forecasting models relies on how close the models make estimates for extensive forecast time. Till now, several researchers attempt multiple-step-ahead forecasting schemes, with the aim of forecasting events some time steps ahead in time. However, multi-step-ahead forecasting presents itself as a challenging approach. Nayak et al. [161]. evaluated the proficiencies of a real-time precipitation-streamflow model with a fuzzy-based computing approach. Abrishamchi et al. [162] developed a fuzzy inference system based on "IF-THEN" implemented an ANFIS model based on "IF-THEN" detailed rules for applying predictions on reservoir flux operation. Performance evaluation of the operation model was obtained from calculating different performance criteria like reliability, resiliency, and liability. Results showed that the application of this scheme in extracting knowledge from an informative data set having imprecise and highly nonlinear structures would be helpful and have advantages over traditional procedure techniques such as standard operating plans and ordinary least-squared return rules constructed based on the results of optimization models. Çimen and Saplioğlu [163] demonstrated the use of a fuzzy built paradigm for streamflow prediction at the upper and lower part of a river gauging station.

Toprak et al. [164] used a fuzzy logic approach implemented with the Mamdani algorithm for stream-flow estimation in ungauged or poorly gauged river basins. Due to the limited data, the authors had to use some stream and time coefficients as feedback data. Other findings reported by authors show the simulation was dependent upon other numerical traits, including the errors, and mapping scheme. In conclusion, they found the fuzzy model to produce good results. A study by Katambara and Ndiritu [165] demonstrated that fuzzy systems work very well in basins where flow data availability is sparse. Al-Zu'bi et al. [166] proposed a rule-based system based on the "Takagi-Sugeno" system fuzzy model for flow prediction. Firat and Turan [167] tested the performance of an ANFIS model with ANNs and conventional time series analysis and forecasting models of AR schemes.

2.2.4.3 Evolutionary Computing

Evolutionary computing (EC) techniques for the last past decades have attracted the attention of various researchers and as developed as an effective means in solving hydrologic related problems [168–173]. As other data-driven available models, there are several types of schemes within the evolutionary computing techniques [174]. In this sense, the classification is as follows: "evolutionary programming (EP)" [175], "genetic algorithms (GAs)" [176], [177], "evolution strategies (ES)" [178], "genetic programming (GP)" [179], and "gene expression programming (GEP)", a branch under GP.

In streamflow simulation and forecasting, evolutionary computing has shown to outperform previous computational intelligence approaches [119]. Wang et al. [180] employed TDNN (time-delay-neural-network) and GA to forecast overland flow in a catchment. From this approach, the researchers concluded that combining both methods, did significantly improve forecasting capacity. Makkeasorn et al. [181] applied genetic programming and artificial neural network for streamflow forecasting. A comparison between the two techniques shows genetic programming to accurately predict the streamflow at a monthly a lead-time. Chen and colaborators [182] highlighted NLTS (nonlinear time series) models to reproduce the fluctuating dynamics of streamflow. Guven [183] evaluated the efficiency of predicting river inflow with linear genetic programming and neural network techniques. The results indicated the linear genetic scheme to outperform that of the neural network counterparts.

Ni et al. [184] considering meteorological variations and performed yearly streamflow forecast through the genetic programming simulation with other DDM schemes (i.e., multilinear regression and Greys's model) and assessed that genetic programming outperformed.

GP model can also catch the complex correlation concerning the forecasters and the predictable outcome (streamflow). To illustrate this method, a study is provided by Kashid et al. [185]. In the study, the researchers applied a modeling approach with precipitation data resulted from predictions from the "El Niño Southern Oscillation (ENSO)" data files. This Precipitation information was then loaded into a GP model for consequently predicting streamflow. The authors reported good performance of the prediction of streamflow with the implemented approach.

Keshtegar et al. [186] reported the importance of the need for accurate and reliable streamflow forecasting models in hydrologic related assets scheduling and administration. Therefore, they investigated, the so-called "High-Order Response Surface (HORS)" process applying some improvements on the model using "high-order polynomial functions" to predict streamflow. The monthly chronological streamflow information of the "Aswan High Dam (AHD)" was reviewed by the authors, from which they concluded that the examination of several higher-order polynomial functions, of the 2^{nd}, 3^{rd}, 4^{th}, and 5^{th}-order, are better-case scenario to replicate streamflow predictions. From the analysis, statistical measures indicated the intended scheme performance with the fifth-order polynomial function to outperform the monthly streamflow forecasting at AHD.

2.2.4.4 Hybrid systems

The complexity of the hydrological phenomenon urges us toward the implementation of other robust hydrological simulation applications, with forecasting capabilities accuracy at least superior to those that already exist. Given these reasons, the hybridization of systems has become a field of motivation for various researchers involved in the hydrological field. The rationale behind the application of hybrid models in hydrological forecasting is that hybrid models represent an ensemble of multi-models approximations that are applied to escalate the exactitude and correctness of sole models [120]. Wang et al. [187] evaluated streamflow prediction only from discharge data in the absence of precipitation with the implementation of a model resulted from the fusion of three types of ANNs hybrids. Results indicated that the hybrid fusion with further analysis of the discharge data (i.e., detrending and differentiation) rendered the forecast effective. Jain and Kumar [188] combined artificial neural networks with a probabilistic approach to investigate hydrologic prediction. Partal [189] studied streamflow prediction with a hybrid approach employing wavelets-based networks.

In Section 2.2.4.2 it was observed how Fuzzy systems are adapted to deal with inaccurate and ambiguous data. From its principles, the knowledge derived can be used in expert systems. Besides, it was learned that Fuzzy systems are expert methods and means predicated on fuzzy rules and inference, that is established on two rules IF and THEN. In addressing flow estimation below a river reach with streamflow data located at the upper part of the gauging station, Pramanik and Panda [190] investigated the application of ANN and ANFIS to estimate the streamflow rate. They concluded the approach with the ANFIS to outperform that of the ANN technique, albeit taking into account that the selection of the correct algorithm scheme plays an essential role not to be underestimated. Dastorani et al. [191] in a study with ANNs and ANFIS, addressed the problems with station missing flow data and concluded the ANFIS technique to satisfactorily forecast the missing data over the results obtained with the ANNs. Adnan et al. [192] proposed a back-propagation neural network with an Extended Kalman Filter at the network exit to enhance model results accuracy.

It was shown earlier that for flow forecasting, the lead time is of great importance, in this sense, Guimarães Santos and Silva [193] presented a hybrid model for daily streamflow forecasting centered on WDT and ANNs. The models could forecast streamflow for lead times from one day, up to a week in advance, constructed on the bandwidth frequencies of the initial elements. Results suggested a hybridization scheme to outperform the typical ANNs model. Humphrey et al. [194] explored hybridization methods by fusing the numerical hydrologic model with Bayesian Artificial Neural Network (BANN) for discharge prediction. Yaseen et al. [195] combined ANFIS and EC techniques and came up with a new hybrid solution to perform monthly discharge predictions. Zaini et al. [196] applied a hybridization approach and built a model based on SVM and optimization scheme for 24 hours streamflow forecasting in Malaysian watershed. Results show the hybrid approach (SVM-PSO) to give better results than forecasting only with SVM.

2.3 Agent Technology

The initial studies in the agent-based modeling area, have their beginning around environments with complex arrangements [197–200], initiated with artificial intelligence and computer science with its many forms of one-agent systems, such as intelligent assistants and service robots [201] but nowadays it is being developed in other areas in academic research and industry (vide infra Table 2.2). In this sense, it should be noted that according to O'sullivan and Haklay [202] and Gimblett [203] the prospective for sociological applications lies around the knowledge of techniques as MAS. Besides, in fields such as geography, the physical components that entail complex systems like vegetation, fauna and flora, physiography, climate, and hydrological component are often unconnected from the socio-economic factors, such as demography, culture, economy, and policy [204].

In the 70s, John Conway built a two dimensional (2D) cellular automata model which he coined the term "Game of Life" (Figure 2.4). This model architecture consisted of a cell layout having two conditions, alive or dead; in this respect, the condition of one cell depended on the one of its neighbor's previous time step. Conway's game triggered the interest at the beginning of complexity from simple instructions.

The evolution of ABM development continued throughout the 1990s as can be witnessed through the emergence of diverse applications means, such as Swarm, and the early beginnings of ABM programs [205] and NetLogo, which was first known as "StarLogoT" in the middle 90s [206], besides others like "Repast" [207] then "AnyLogic" with is initial release in 2000 by the former XJ Technologies known today as The AnyLogic Company [208].

During that same period, [209] implemented Sugarscape, an AI agent-based social simulation model (ABSS), which adapts the fundamental concepts of social sciences. This prototype was composed of a system of naive procedures which formed the basis for the creation of other procedures that enhanced supplementary remarkable outcomes. Sugarscape is an example that presented how basic procedures might produce composite organization in a bottom-to-top approach, this means with local instructions being at the bottom and ascending to adaptive behaviors and capabilities of the interconnected structure terminating on the top. Culminating the 90s, the power of computers progressed considerably, and agent-based models became well-known.

The term "ABM" has been coined with many other terminologies as shown in the literature, these are "agent-based systems (ABS)","individual-based modeling (IBM)", and "multi-agent system (MAS)" are commonly applied names with their abbreviations, and which use will be common throughout this thesis. Nevertheless, in Sections 2.3.1 and 2.3.3 the ABM and MAS concepts would be addressed in-depth, respectively.

Table 2.2: Areas of Agent-Based Modeling Applications (Adapted from: [210]).

Field	Applications	Field	Applications
Commercial and Institutional	• Industrial Processes • Resource chain • Supply chain management • Supply and demand • Customer marketplaces • Manufacturing handling	Society and Culture	• Early societies • Civilian insubordination • Associated terrorist attacks • Administrative systems
Economics	• Computational economic marketplaces • Commercial nets • Computational economics • Economical ecology	Military	• Authority & jurisdiction • Enforcement
Infrastructure	• Energy power markets • Transportation • Hydrogen infrastructure • Oil and Gas industry	Biology	• Civilian displacement • Natural systems • Ethology • Operation of automated cell-like systems and subsystem implementation • Ecology
Multitudes	• Circulation of people • Withdrawal simulation	Technology	• Energy technology • Dynamics of the systems • Technical control • Hydro base technology • geographical information systems modeling

2.3.1 ABM in Flood Forecasting

ABM can be defined fairly as a novel methodology for modeling and simulating complicated domains, in which cooperative, independent agents, that can exist in space and time interacts among each other [212, 213]. Many authors in this research community claim that ABM

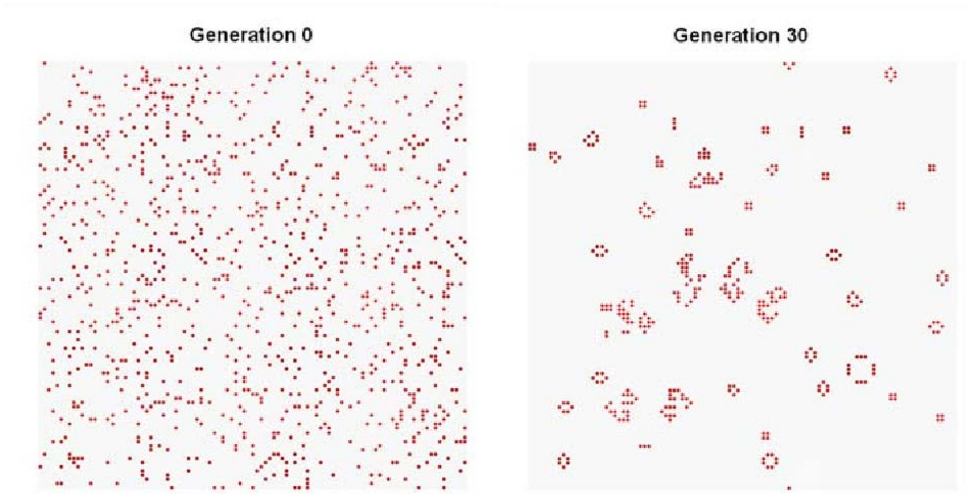


Figure 2.4: The "Game of Life". Example of a "cellular automata machine". Adapted from: Conway [211].

represents a new paradigm for simulation [214], however, they disagree among themselves on how to define the agent concept since there are many points of view [215–220]. So, what is an agent? (Figure 2.5). As mentioned previously, the agent description is not well-defined, since an "agent" could be any components forming part of a system when different objectives in different paradigms are studied. For example, some would regard it as any type of notable parts integrating a model, a system, or a subsystem in a given program environment. Likewise, an agent could be considered as any class of autonomous entity, such as an organization, institution, or an individual person. As the idea of an agent is crucial for the understanding and implementation of the ABM systems, most researchers point to the following characterization by [221] for ABM as they find it useful. ABMs are computer simulations for modeling the behaviors and collaborations between agents which display some sort of autonomous behavior in assessing their co-operation on a system. Additionally, they integrate features from other computational simulation paradigms and schemes. For instance, to introduce repeated randomness into a system, computational algorithms like the Monte Carlo methods are used. According to [222] in an ecological concept, agent-based models, are known as "individual-based models (IBMs)". A review by [223] shows the applications of ABMs in non-computing domains (e.g., biology, ecology, and social science).

Another particular property of ABMs is that they can be displayed as a microscale model [224] in which the synchronized procedures and relationships of several agents to replicate and forecast the appearance of complicated trends is modeled. Therefore, the process would be developed as a micro and macro plane of interconnected structures. Therefore, a significant concept is that from simple behavioral rules more complex behaviors can be generated. Another key principle is the synergy displayed among agents working together as one. Individual agents are usually portrayed as reasoning entity, acting with independence, and having behavior rules varying from basic reaction decisions to more complex adaptive

Artificial Intelligence [210]. Bonabeau [225] argued that in ABM agents are autonomous entities that may experience "learning", adaptation, and reproduction.

In the literature review, the use of ABMs has been shown to become increasingly popular in social sciences, given its elegance and explicitness to show objects, environment, and relations between them [226–231]. Despite many ABMs turned out to be built to model the complicated social phenomena, a few are being developed for stream-flow or flood forecasting.

Another important aspect of ABMs is their consideration as experimental tools for theoretical research on complex social experiences for understanding the interactions of the agents and a giving environmental scenario. Therefore, as a new approach for analyzing complex systems that emerge through interactions among autonomous agents [213], and the trends observed in ABM is the manner in which it simulates the system in a bottom-up approach, suggests, patterns, structure, and behaviors that could be perceived.

In respect to their structure, an agent-based model involves three basic components: agents, the agent's association with other agents, and the agent's setting (i.e., "objects" and their surroundings). In other words, agents denote the active part of the system, the objects are represented by passive elements and the agent's environment represents the interactions among the different components in the systems.

To offer an intellectual validation specifically why ABMs methodologies are appropriate for solving complexities in computational domains, a qualitative analysis on the topic can be seen in a paper by [217]. Hence, the reasons why today this idea is largely accepted among researchers in the ABM society is that the technology proposes innovative and often new applicable directions for building complex procedures, particularly accessible and active environments [232]. Briefly shown under this section, it introduced the ABM as a useful intelligent, and complex software system that can have many interacting components and parts. Therefore, agents are useful perspective and appropriate for modeling, as they reduce the time in coupling various systems given the abstraction offered through independence, besides their robust, reactive, and pro-active skills [233].

2.3.2 ABM Architectures

For the planning and developmental stages of an ABM, an assembly of agents is needed for complex system during the planning phase. Thus, often times it is required to combine various AI methods and several agents. On the contrary, since diverse methods can simply be coupled into a hybrid system using a connecting type of agent frame, various complicated glitches are resolved within a smaller length of time. Likewise, given the diversity of complementary techniques and approaches to deal with these problems, these become coalesced, and solutions with higher quality have resulted from such schemes. Of course, it is imminent that every one of these smart systems has its potentials and disadvantages, and they do not represent the norm for solving every difficulty. In this respect, many investigations on this subject, are addressed further by several research. For example, the "The MIX Multi-Agent Architecture" introduced by Hilario et al. [234], and Iglesias et al. [235] is an architecture aimed at developing plans and kits for combining neuronal and representative pieces of

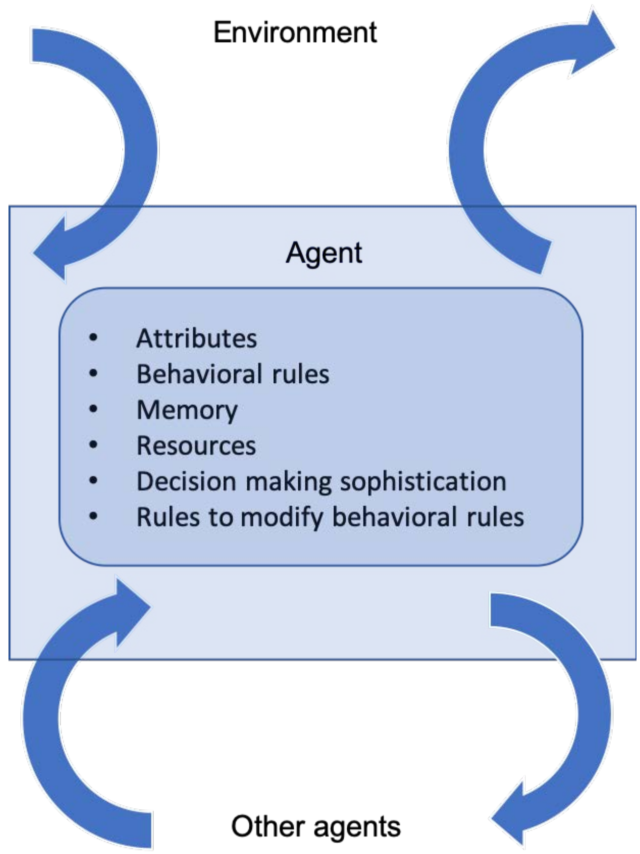


Figure 2.5: A characteristic agent. Adapted from: [210].

knowledge. The testing skills of this architecture are done on the basis of an allocated procedure on multiple systems composed of collaborating diverse agents. The MIX architecture involves a multi-agent toolbox with a basic agent array, facilities, and rules for agents to be able to communicate. Hence, the development of agents is explicit for diverse sorts of neuronal arrangements unlike other tasks such as reasoning systems. A similar attempt is the "PREDICTOR" system by Scherer and Schlageter [236] detailed how an allocated system of AI methodologies can be used by coalescing neuronal and knowledge methods. To present this system, the authors have argued that such a design is centered on a slate construction and is validated in the financial and prediction areas. The researchers developed this method aiming to offer solutions to forecasting problems in the financial system.

The "IMAHDA Architectural Setup", an architecture introduced by Khosla and Dillon [237] is a computer design known as the "Intelligent Multi-Agent Hybrid Distributed Architecture" (IMAHDA). In IMAHDA, an arrangement of the execution of its functioning and knowledge base is composed of four levels. These levels are: objects, computer agents, intelligent agents, and agents for the resolution of problems. This system can be built from general-purpose agent software, standard smart agents and from other general AI systems.

A so-called "Multi-agent and Fuzzy Modeling Architecture" is a more recent attempt and was introduced by Delgado and Gómez-Skarmeta [238]. In this methodology, the authors have proposed a hybridized knowledgeable prototype which is an amalgamation of fuzzy logic techniques and MAS. The arrangement of this type of system allows the participation of several agents with their tasks within a context of fuzzy logic for the solutions of specific problems. The prototype involves four types of agents namely, service, which acts as a directory, classifier agents, resources, and finally those that control tasks.

The Generic architecture suggested by Jacobsen [239] was a standard agent-based design for a mixed platform of intelligent agents, that was built on the framework already introduced in [240]. Consequently, they showed two abstractions typical of this mixed platform, which on the one hand is a reinforced- architecture linked to fuzzy logic, and on the other hand, in conjunctions with knowledge-based systems, and neural-fuzzy techniques they were able to experimentally validate their designs.

In summary, a thorough insight into the proficiency and limitations of these five agent-based hybrid platforms or systems can be found in [241].

2.3.3 Multi-agent Systems in Flood Forecasting

Just like ABMs, MAS is also part of AI, it is a methodology that aims to pair artificial intelligence techniques with collaborative environments, computer science, and applications as one system [242, 243] and is also part of AI, it is a methodology that aims to couple artificial intelligence techniques with distributed system, and software engineering in a single discipline system. Generally, MASs are composed of autonomous agents grouped and sharing the same ordered environment [244]. They are displayed as a network that interacts with each other with the aim to accomplish common goals [245]. An agent can be viewed as a software component that contains coding instructions and data [246]. Within multi-agent systems MAS(s), the agents that are composed of the network are likely incapable of solving the assigned problems on their own [247].

Agent communications in a given environment are performed while they are working autonomously and are in coordination among peers. Then, in a collaborative society, each agent is equipped with a series of skills that are their own and that together allows them as a team to solve problems. In this sense, the messaging that occurs between agents is feasible thanks to the language protocol for agents provided by the Agent Communication Language (ACL), they can exchange information and demand assistance among themselves in a negotiable manner [246].

The most useful ACL protocol among agents is the "Knowledge Query and Manipulation Language" (KQML). It is a messaging level that involves a low-level layer of variables like send, recipient, and identify. The messaging level is the one that specifies the performative and the interpretation procedure and the information about the performative is at the content level [247]. The coordination between agents is imperative since the couriers are given over relays and these messages are not dispatched at regular intervals between the agents. From there, the coordination will depend on the architecture of the system which determines that the flow of information will flow smoothly [246]. The most simple architecture is a "point-to-point (P2P)" network in which agents speak openly to each other. This is effective to solve situations in which immediate solutions are needed without creating conflicting situations. Two other platforms are shared between the MAS, and its associated "agent", besides the so-called "black-boards". An associated agent network is one that facilitates an agent as a mediator among other agents. While a network of black-board agents is integrated a central coordinator who directs the actions of coordination of the activities of data sharing [246]. MAS has a conceptual and well-defined historical background, in this sense, an example of this is illustrated in "game theory". "Game theory" (GT) [248] is a mathematical modeling paradigm in which a cooperative environment is simulated between non-thinking and thinking agents in which the agents interacts among themselves to resolve conflicts. An illustration using GT to address a hydrological related problem can be found in [249].

In an investigation performed by Montalvo et al. [250] an integrated MAS architecture was proposed with an optimization algorithm approach, to correct the difficulties for the improvements of drinking water distribution systems. The approach, the author's comment, is apt to train managers and decision makers when offering proposals for improvements to facilities.

There are various areas in which MAS techniques have been applied [251]. Recently, MAS has shown to find a place in the assessment of ecosystems [252–255]. Some of these examples also include rangeland management [256], fish farming [257], urban catchments [258] and irrigation farming systems [259]. The approach of the MAS is practical for these areas due to the representation they offer to the environment with its great complexity of details and that they allow the possibility to integrate human actors, to find solutions and to contribute to improvements in the assessment of methodologies.

2.3.3.1 Multi-agent Systems Architectures and Agents Types

Having previously discussed what is an agent, in constructing MASs it is important to define the agent internal organization and processes that details how agents pursue and gain their preferred intentions. According to Wooldridge [260], an agent design can be abstract or concrete. The structure of an abstract concept is defined by the components and its engine structure (e.g., roles, permissions, responsibilities, activities, and protocols). On the other hand, the concrete concept is determined by the assignation of types to each component and executing each function instruction of the engine. In summary, architectures define how an agent perceives data by external sensors, and its internal state determines the actions it will perform and its future behavior [261, 262].

As suggested by Wooldridge and Jennings [263] there is generally three known taxonomies of agent's typology, which are classified concerning their, internal configuration, modular components, how they behave and interact as: i) "reactive architectures", ii) "deliberative architectures", and iii) "hybrid or layered architecture".

Reactive agents, as the name implies, react directly to changes in the environment, they are constantly adapting their internal status. This direct response to environmental stimuli (e.g., sensor signal) in whatsoever direction, given the order to change into acting is also known as the agent role [214]. In contrast to reactive architectures, where there is no thinking on choosing the next step of action, in deliberative architectures, agents can think and decide before choosing the next action. The commonly known "Belief-Desire-Intention (BDI)" architecture known for its behavior capabilities and typical interaction model that can display a rational behavior and realistic reasoning is an example in this regard. [215]. The BDI architecture was proposed originally by Rao and Georgeff [264] and its main components are described in [265]. Concluding, hybrid and layered agent architectures are a type of architecture that shares these other two classes, combines multiple agent mechanisms to use a particular benefit, and operates abstracting over different levels of the surrounding environment [260].

2.4 ABM and MAS Applications in Hydrologic Problems

Tables 2.3 and 2.4.2 provide several articles on ABM and MAS dealing with various hydrologic problems. Both tables also show a brief description of the approaches employed by the researchers and the flood-based hydrologic problem-solution type addressed.

2.4.1 ABM Applications in Hydrologic Problems

As shown in the previous chapter, ABM focuses on whether agents (though, not necessarily "intelligent") are obeying designated rules that formed their behaviors. Giving the novelty and ongoing research in the ABM area. Briefly, Table 2.3 shows the areas in which ABM is in active research.

Table 2.3: Agent-based modeling in flood problems.

Method	**Reference**	**Description**	**Hydrologic Solution**
ABM	Brouwers & Boman [266]	Merging ABM and GIS and hydrologic modeling for societal modeling.	Flood Management Strategies
ABM	Anantsuksomsri & Tontisirin [267]	A comprehensive examination of ABM and Disaster Planning.	Disaster Management

Table 2.3 – *Continued from previous page*

Method	Reference	Description	Hydrologic Solution
ABM, Virtual Geographic Environment (VGE)	Coates et al. [268]	Fusion of agent-based modeling and "VGE" to evaluate hazards caused by flooding on businesses infrastructure.	Flood Recovery and Response
ABM, Geospatial Services	Tan et al. [269]	Applies ABM and online GIS for real-time charting to assist in the reaction to inundation episodes.	Flood Responding
ABM	Berglund [270]	Demonstration of the applicability of ABM in water supplies issues to comprehend the complexities involved.	integrated operation of water governance
ABM, Hydrologic Simulation	Sunde et al. [271]	Studied the effects of impervious catchments toward development using agent-based model and catchment modeling with ("SWAT").	Preparation and Security for Flooding
ABM, "Complex Adaptive System CAS"	Medina et al. [272]	Applied "CAS" and agent-based model in drill assessment for the planning of demographic response to coastal surges.	Flood Evacuation Strategies
ABM	Du et al. [273]	An agent-based frame to examine the impact of managing inundation alerts by different actors.	Flood Warning
ABM	Jenkins [274]	Agent-based approach in climate change simulation triggering inundation hazard	Preparation and Security for Flooding
ABM	Yang et al. [275]	Proposed an ABM for simulating individuals responding to flood hazards.	Flood Recovery and Response

End of table

Some ten years ago, in an attempt to elucidate the progress in the agent-based field, a survey of some 279 manuscripts from ninety-two research papers was done by Heath et al. [276] in search for authors that had implemented and evaluated an agent-based model, with the need to constantly assess the up-to-date knowledge of the current advances in ABMs and detect cases in which the systems needs improvement. The authors identified "six improvements needed to advance ABM as an analysis tool": i) implementation of the agent-based-modeling tool kits particular to the problem domain and are software independent, ii) ABM development as a unique specialty with a language shared and extended to other domains, iii) An ABM system that is equivalent to the intended purposes, iv), An ABM system that is simulation descriptive and provides results that are reproducible, v) They should be totally adequate for validation, and vi) That ABM model validation statistic metrics be specific to ABM. Other interesting findings made by the researchers showed the agent-based-modeling methodology is a recent method and that the experimental models supporting agent-based-modeling as an effective simulation environment have not yet been approved. They also noted that in agent-based modeling, the absence of development and typical standards reflects the absence of verified prototypes. Finally, the authors recommended that a solution that the ideas, methods, and techniques used for ABM must be acquired from other modeling prototypes or implemented towards agent-based modeling. A complete study on the potentials of ABM is also referenced by Macal and North [210] in which a comprehensive insight on the rationale of ABM is introduced with the idea to address its present settings. The author reexamines several issues of ABM, characterizes this paradigm as a novel methodology, considers the limits of over-employing and under-employing this AI sub-discipline and envisions its potential as a prospective research tool in complex systems modeling.Additionally, the study identifies the essential features of ABM properties, research, and groups. The study also suggests some corresponding meanings for ABM, grounded on experience, proposed for determining a unique lexicon to be employed. The author concludes by recommending research and challenges in this area to advance its growth and potential in the coming years.

Despite the active research of ABM in domains like economics, social science, biology, military, public policy, ecology and traffic, and with the growing interest in engineering applications, the literature review shows, however, that ABM applied to stream-flow or flood forecasting problems are relatively limited and show not to be thoroughly enriched in that field. Notwithstanding, it observed some examples given by some researchers (Table 2.3) that have used such systems for addressing other related hydrologic problems, as a problem-solving approach in flood safety and planning, recovery, response, evacuation policies, control, and management.

Brouwers and Boman [266] designed and implemented a single ABM to discover the preferences of individuals toward assessment on inundation administration plans, for a geographically explicit flood simulation model under situations in communities with spatial accretion. However, the researchers commented that for this model to function, the results must be available and practically outreaching to the entire community.

In a recent summary of the agent-based modeling approach for solving problems in the area of disaster management, a study was published by Anantsuksomsri and Tontisirin [267]. In this study, the authors conducted a review on ABM applied to disaster management. From their review, they explain the development of such systems and define ABM and give insights on some software toolkits used for building ABM systems. The study also contributed to the present modeling of issues in law models as theoretical testing whereas robbery and driving behavior models are chosen as the implications of ABM in urban planning. The article also discusses the use of ABM on natural disaster policies and management, drills for inundations retreat, and liability assessment. Of particular notice is the references of the authors to two papers related to agent-based modeling coupled to a hydrodynamic simulation model, in either of these cases, the ABM uses the ready hand information of the hydrodynamic model water levels of address the issues of evacuating the population rather than the ABM system to forecast the water surges.

Interestingly, Coates et al. [268] introduced a study in which they implemented with the use of geospatial systems and ABM coupled with flood forecasting estimates model a system that could identify commercial properties prone to suffer from flood damage. The rationale behind the geospatial system consisted of the development of layers (e.g., Topography, Integrated Transport Network, and Address) from Ordinance Surveys Master Map to be used by the ABM system. The study showed that integrating the geospatial system with flood estimates layers enhanced the reliability in modeling flooding occurrences. Finally, the authors proposed the idea that in order to improve operational response and business continuity, it is necessary to build prototypes that can replicate companies throughout and after flood events.

A common tool that complements to the simulation of ABM in water resources engineering is the Integration of Geographic Information Systems (GIS), as reported by [269]. They studied how effective the integration of GIS data features and ABM could be. They installed the integration approach on the web, and its task was to gather and handle vast amounts of generated geospatial data and in this manner gain reliable geospatial services information that could respond more effectively to floods. This integration showed that the proposed method is ideal to avoid the transfer of a massive amount of geospatial information.

Berglund [270] provided a complete and comprehensive presentation on ABM to the water resources community, with the aim of exploring the use of ABM in the hydro-sector giving the complexities common to this domain. The author presented two descriptive overviews implemented with ABM which were explained. From the analysis, he demonstrated the applicability of the agent paradigm for simulating scenarios like water resources planning problems. Additionally, in the study, the author also observed the restraints in employing agent-based modeling to this simulation domain. To comprehend stream-flow on impervious surfaces, with a hydrologic simulation coupled to an ABM technology, Sunde et al. [271] employed a "Pixel-based increased impervious surface" dataset projected from 2011 to 2031. This dataset was linked to the "Soil and Water Assessment Tool (SWAT)", a hydrological simulation application. The simulation runs were focused to study the potential of the hydrologic effects on upcoming urban development.

Medina et al. [272] used "complex adaptive system" theories and ABMs to undertake the challenges of testing massive withdrawal policies in coastline communities prone to flooding events. The authors showed the viability of agent-based modeling approach in this scenario to test evacuation policies for coastline communities withdrawals due to severe hydrometeorological incidents. They also stressed its perspective as a toolkit for effective hazard managing.

Du et al. [273] developed an ABM architecture to evaluate the effects of the diversity in human reactions to inundation threats, domestic concentration, and the benefits of inundation alerts. The modeling scheme consisted of the implementation of an ABM coupled to a transportation simulation to model mass departures in a highway grid with different flooding alert cases. Simulation outcomes indicated that if the population behaved subject to the stressing effects of threats, particularly areas with the high-density population the marginal benefit related to efficient flooding warnings becomes significantly constrained. Results also showed the profits of inundation alerts to notably impact human behavioral heterogeneity and from thence the meaning of seeing human reactions to inundation alert simulation routines. Finally, the authors recommended the development of more accurate models on social reactions and conduct, to inundation alerts, and to increase the number of domestic spaces and elements for better assessing and improving the advantages of inundation alert systems.

Concerning the future views on meteorological conditions, Jenkins et al. [274] presented a novel ABM, which they feed with information from an inundation hazard scenario analysis. The agent-based model was proposed to evaluate the relationship amid distinct modification alternatives; it could propose that a decline in risk may be achieved by proprietors and authorities.

Yang et al. [275] proposed an agent-based model that could perform flood response simulations in regards to the choices and actions taken by every house owner to ease flood damages. The model implements a framework for individual response in which agents evaluate different flooding situations concurring to inundation alert systems, collect and select if and what amount they would spend in reaction procedures to ease latent flood losses. The researchers observed that estate worth, forewarning communication, and heavy rainfall situations altogether influence housing damages, located at lowland areas of the watershed, and highly populated zones are most likely to be susceptible. The ABM also demonstrated to be useful for analyzing housing damages to large scale flooding and reactions in storm-water flooding episodes.

2.4.2 MAS Applications in Hydrologic Problems

The previous section, briefly examined how various agent-based models were implemented to simulate human behavior in response to complex phenomena such as disasters that could be climatic or of other sources, as part of disaster management. On the other hand, it could be noted that most of the modeling approaches with ABM are not focused directly for streamflow forecasting, but rather for simulations on population dynamics responses to risk and hazardous situations, as well as for the design, implementation, and evaluation of risk policies, flood warning, safety, and planning. Besides, as an increase in the amount of risk and hazards around the world escalates, as they become more frequent and severe due to weather-related events (Figure 2.6), especially in regions where the precipitation regime can be even higher than the annual global rainfall, the continuous need for the development of more sophisticated agent related technologies that can be applied to hydrologic-related-problems and water-dynamics forecasters are the order of the day. Hydrologic- related problems, which in this case is given special attention to events such as floods are triggered by severe weather conditions like heavy rainfall.

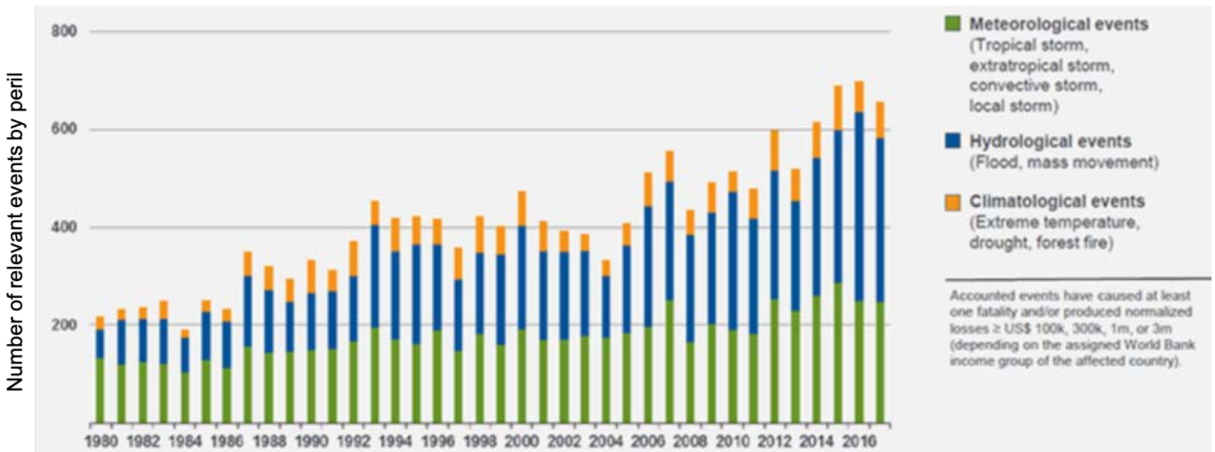


Figure 2.6: World weather-related natural phenomenon, years 1980 to 2017. Source: [277].

In the urban and catchment levels, floods are of importance given the damage they cause on the impacted domain (human settlements, animals, environment, agriculture, infrastructure, and economy). To monitor these weather or climatic variables that are responsible for flooding, it is necessary to have at hand and deploy a network of hydrometric sensors (Figure 2.7) for data collection to allow us to study and forecast flooding events. In response to the later, a MAS is an option that is best suitable for addressing a network of distributed sensors [278]. This owes precisely to the fact that a multi-agent system comprises a network of agent displaying intelligence, a framework that involves collective features and individual thinking for the solution of complex issues. The MAS approach to investigate and assess stream-flow, flood, or high-water level forecasting and its association with management in the civil society is a recent effort.

From a careful review of the related literature, the applications of MASs have been successfully applied to several problem domains, including energy market forecasting, monitoring, system analysis, and corrective actions [279–282].

Within the field of hydrology, the community of MAS researchers has also made some progress, therefore, to get some understanding regarding the performance of MAS techniques. In Table 2.4.2 a condensation of MAS main techniques is provided. Some of these techniques include, but or not limited to single MAS methods in real-time simulation for flood forecasting, monitoring and warning, ANNs, GIS, DM, fuzzy logic, case-based reasoning (CBR), mobile communications systems using "Very Small Aperture Terminal (VSAT)", and expert systems.

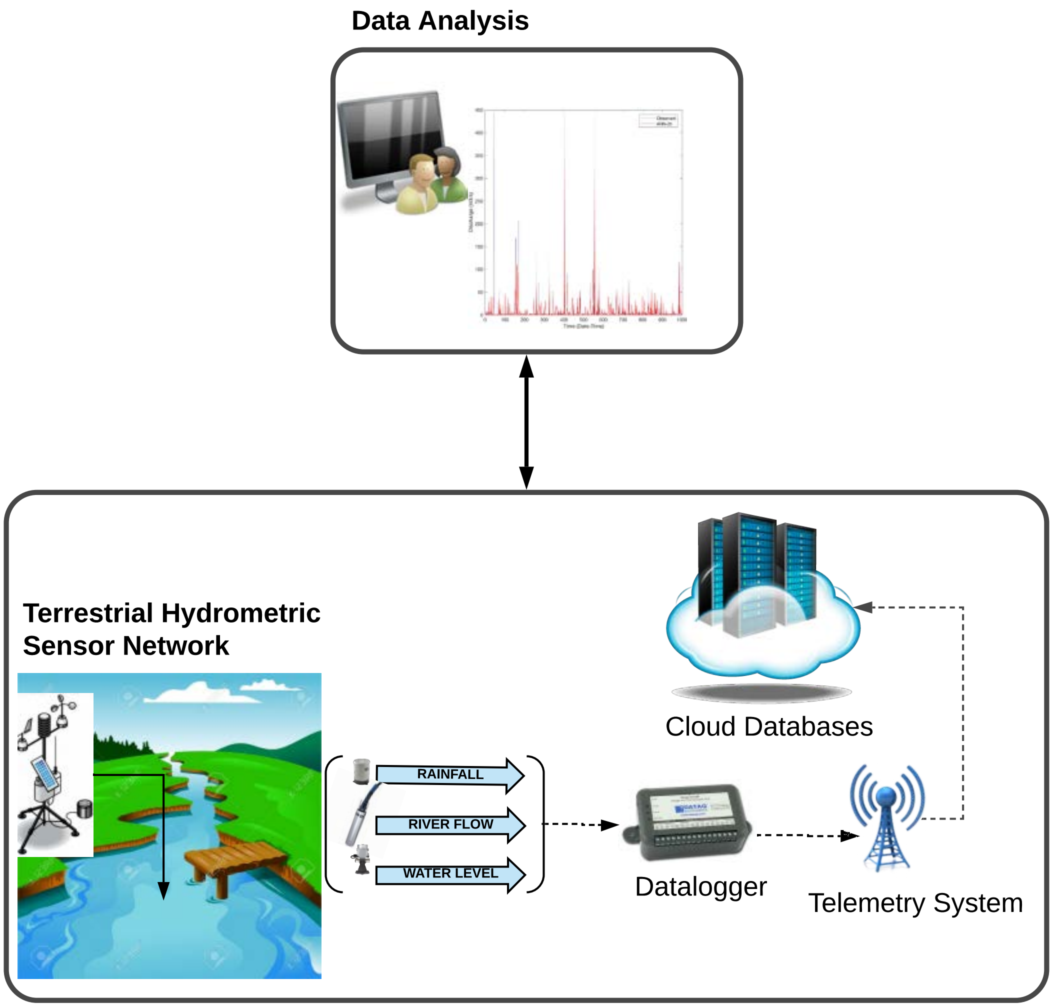


Figure 2.7: Deployment of a hydrometric sensor networks.

Table 2.4: Multi-agent systems modeling in flood problems.

Method	Researchers	Outline	Hydro-Solution
MAS	George et al. [283]	Used a MAS approach implemented with two stages. In stage one, the fluctuations in water level are computed during 1-hour. In stage two is computed the sum of the weights of agents.	Simultaneous modeling for Flooding Prediction

Table 2.4 – *Continued from previous page*

Method	**Researchers**	**Outline**	**Hydro-Solution**
MAS, GIS Grid	De Roure et al. [284]	A global computerized infrastructure with stationary and moving agents, knowledge engineering, and a GIS grid for river stage surveillance.	Gauge Data Administration and River Stage Surveying
MAS	Matei [285]	Employs disseminated frameworks.	Surveillance with Analysis of Watershed
MAS, ANN	López et al. [286]	Applies NNs types assemblies and intelligent agents with mobile devices.	Early Warning Against Floods
MAS, DM	Mabrouk et al. [287]	Combines MAS and DM for flood forecasting and warning.	Decision support for Predicting River Flood and Issue Alerting
MAS	Marouane et al. [288]	Data retrieval and administration of cordless equipment with MAS for online streamflow forecasting.	Flood Forecasting
MAS, "VSAT"	Iqbal et al. [289]	Apply MAS with mobile agents and algorithms coupled to "VSAT technology".	Prediction of Inundation Area
MAS, CBR	Linghu and Chen [290]	Use an "agent-based" and "case-based reasoning" approach for the prediction of inundation hazards.	Flood Disaster Forecasting
MAS, GIS, DM	Al-azzam et al. [291]	Applies a MAS modeling approach based on DDM techniques and GIS as a system for supporting flood risk management.	Flood Prediction and Risk Assessment
MAS	Bao et al. [292]	Applies a MAS modeling concept to evaluate runoff.	Urban Water-log Simulation and Prediction
MAS, Ontology, Fuzzy Logic	Aris et al. [293]	A MAS integrated, with ontology concepts and uncertainty modeling.	Flood Warning Prediction

Table 2.4 – *Continued from previous page*

Method	Researchers	Outline	Hydro-Solution
MAS, Knowledge based systems	Mabrouk et al. [294]	Employs "multi-agent systems" with knowledge based system for instantaneous inundation warning with prediction for various time-span intervals.	Flood Forecasting with Warning
MAS	Mabrouk and Gaou [295]	An expert system for real-time processing of data acquisition, classification, and collecting for inundation nowcasting and warning.	Flood Forecasting/Warning

End of table

George et al. [283] built a platform for the actual simulation that is coupled to a model for estimation of floods, with a two-level self-adapting MAS, was implemented. In this system, the objective of each level is well defined, such that at the top, the variation in water level is measured for an hour period and in order to achieve this, it adjusts the weights in the lower stage. The adaptiveness of this model is really acquired by adjusting the weights, carried by the agents that cooperate among themselves. This adjustment of the weights renders the model generic and allows to improve its performance. The authors analyzed several use cases (e.g., noisy and missing data, and number of upstream stations) where the use of hydrologic simulation is inappropriate. Lastly, they studied the features of the knowledge system taking into account classical measures and related works.

De Roure et al. [284] implemented a ubiquitous computing system of static and mobile agents, in which the static agents are designed to have complex functionalities, for example using an expert-system to manage available sensors on the network and use their information to monitor the river level and determine if the network is storing data that differ significantly from the standard, feed the data into a grid-based flood predictor model, and issue an alert. The main functionality for mobile agents is to ensure they are lightweight with actual specific functions to perform, like data discovery routes across the network and deliver sensor data. The authors concluded in this study that the use of intelligent static agents employing simple mobile agents in combination to perform assignments on their behalf results in the effective, self-organizing ubiquitous computer for a simulated network of nodes.

Matei [285] developed and evaluated a multi-agent system for inundation forecast and water level observing. The MAS is fed by an automatic hydrometric data collection module. The author concluded by identifying some advantages and disadvantages of the system. As the MAS is implemented on independent units, the modular arrangement of such and the collaboration among these units permit the assessment in hydrologic problems in an easy manner.

The multi-agent system, suggested in [286] uses a Counter-Propagation neural networks and intelligent agents to analyze and assess flood risk caused by rainfall. Additionally, they implemented agents, on mobile devices, for the dissemination of early warnings on floods. Other features of the system are that it can display flood forecast and enable messages in order to administer a hydropower plant basin with the aim to prevent damage to the infrastructure, it offers other smart means for messaging and broadcasting of massive warnings to the population.

A MAS proposed by Mabrouk et al. [287] coupled with a data-driven approach was developed to improve flood prediction and alerts for a decision support system and management at the catchment scale. Their study was aimed at providing technical support for the flood control and warning division. This approach involved the applications of a data mining tree algorithm (i.e., "C4.5") to construct a real-time flood prediction and early awareness model. Besides to couple the decision tree to the multi-agent system, an algorithm developed and which they named the "ANYtime Multi-Agent System" (ANYMAS) was used to obtain the coupling between to two systems.

Marouane Mabrouk et al. [288] presented an intelligent system that gathers and handles hydrometric data from sensors. The system detects errors in the data and classified them, so as to render them reliable and acceptable for their storage in a database where they undergo data-processing. The data once processed are further used for performing inundation estimation, and the results are distributed in collaboration with the MAS and mobile agents for the dissemination of results to the central station.

Iqbal et al. [289] developed a MAS model implemented with "Very Small Aperture Terminal" (VSAT) technology that optimized the agents in mobile-phone communication. The team of researchers rendered the system to be useful in any type of critical conditions of floods, of course, because the signal strength is enhanced by the technology. Nevertheless, although the cellular signal could be optimized, a set back is that for the implemented agents on the server to be proficient in taking the decision of flooded zones, it depends on the available river historical flow data.

Linghu and Chen [290] presented an innovative example of MAS coupled with "Case-Based Reasoning" (CBR) for inundation disaster prediction. The authors implemented an algorithm for this particular task and they concluded that the algorithm implemented could estimate the river stage correctly and realized the prediction offset of the algorithm was smaller than the existing scheme.

Al-azzam and collegues. [291] presented a MAS architecture linked to a geographic system setup on a virtual environment and data-mining techniques to assist in flood forecasting and risk evaluation. Although the objectives of the implementation were met, the authors observed that during the implementation phase they could discover issues that are of concern when building such systems. They emphasized on the trials that arise whenever building systems like these and such a predicament need to be considered especially when classifying and aggregating erroneous hydro-data from wireless sensors.

Bao and collaborators [292] proposed a MAS to simulate and estimate water level in an urban watershed surface domain. The system design was configured in such a way for simulating water level saturation using the MAS approach instead of hydrologic modeling. This approach was done by using a method that was able to simulate the uncertainties in the flow-regime and to calculate the flow depth at any point in time and optimize reservoir storing capacity to allocate water excess. Hydrometric data are used as inputs to the MAS. To validate the system, the authors compared the results of the MAS with outputs from cellular automata modeling. The authors concluded the system to improve, given the methodology applied, is feasible for mitigating the loss of life and infrastructure damage by high water levels.

Aris, Hamdan, Pa, and Nazeer [293] proposed a flood forecasting theoretical model that was based on agents ontology and fuzzy systems. In this system, the function of agents is to offer reports on flooding situations according to the river stage and precipitation via alerts warnings sent to the population. The ontological setup was meant to categorize the inundation awareness to aid in the agent interaction. Fuzzy techniques are used to forecast weather undefined conditions.

An innovative knowledgeable scheme for flood estimation with three chronological terms (e.g., short, medium, and long) and awareness, was implemented by El Mabrouk et al. [294]. Given the authors concerns to work with a system that could integrate many components into one, they realized that the multi-agent systems had the benefits and the advantages in terms of the distributed artificial intelligence, and for allowing the system to cope with rule-based features, they made use of expert systems due to the advantage of the concept of logical coding and the theories of proofs and directions.

Most recently, El Mabrouk and Gaou [295] proposed a smart system that could pre-process data before performing instantaneous flood prediction and awareness. The system is composed of some levels that supervise the cordless instruments and their accurate performance, to ensure the incoming information from the sensors are of the best quality, store this incoming information in a database from which the instances would are feed into the model future streamflow prediction tasks. Finally, it can be concluded from the study that wireless sensor networks, could be used for a distributed and auto-organized method of information managing in a dispersed system and have considerably upgraded with the attendance and the application of MAS is in agreement with the studies by Guijarro and Fuentes-Fernández [296], and Hamzi et al. [297].

2.4.3 Conclusions

Summarizing, this chapter presented an overview of the literature on some methods used for flood forecasting. It presented both the theoretical and mathematical aspects of hydrologic flood forecasting using such a standard approach as the deterministic and the stochastic variants and observed some drawbacks and difficulties while applying any of these traditional methods. However, given the setbacks in the former, it witnessed the merging of these methods to overcome the setbacks with the DDM and AI paradigms. Although the DMM and AI provided workable solutions in addressing some issues previously discussed, it showed certain setbacks are still latent, specifically the DMM and AI approach requiring high-quality historical time series data and the fact these do not capture entirely that hydrologic processes modeling. Therefore, hydrologists continue to battle these limitations to improve the requirements demanded from hydrologic flood forecasting in a changing climate, that is becoming more complex.

With the agent technology (ABM/MAS) discussed above, it showed the possibilities of the potentials of its applications to hydrologic modeling for flood forecasting and water management; and exhibited some examples of ABM techniques in undertaken issues in flood hazard, flood approach to the management, flood reparation, evacuating strategies, and flood alert, yet a very few cases concerned with flood forecasting, as well as some limitations it might pose [217], which need to be addressed and can open fields to new research solutions.

From the literature review, they acknowledged that, in the mechanisms for examining MAS as a promising tool, they noted that this theory relates to different species of agents. This shows its applications in hydrologic modeling are possible with different species of agents. Therefore, in the administration of hydro base systems, the application of the ABM theory in simulation problems, the representative approach is established on rational agents' species; however, in this context, seldom agents pose greater skills and abilities.

As argued previously, clearly, the merging of hydrologic models and an agent-based concept discloses understandings about the social aspects and related strategies toward managing flood events at regional [298] and overall level [299], an aspect not retrievable from a common standard hydrologic model. Then again, ABM can investigate different pieces of information, principles, and the heterogeneous behaviors that differentiate an environment [300]. Besides, agents are cognitive, diversified in their qualities and behaviors, they can adapt to the environment and are perceptive to the historic account of their decisions, they can be ubiquitous across network nodes or distributed geographically. To this extent, it has been shown the complications with traditional hydrologic models to deal with the complications of the complexities of the real world; therefore, it is recommended the use of the technology, as it allows the abstraction of states simulated with ABM/MAS more characteristic of the physical world.

Chapter 3

MAS Framework for Flood Forecasting

3.1 Introduction

This chapter formulates the methodical approaches and provide a thorough description of the intrinsic configuration and inner components of the proposed MAS model for flood forecasting within the context of the humid watershed. Each participating agent described here, has unique goals and tasks, and deliberately fulfills the tasks organized in the system coherently with other agents of the different levels, regarding the information generated in the river basin, captured by hydrometric monitoring field sensors deployed along the river reach stored and available in a datalogger, and transmitted via telemetry to a central server. Changes in surface water elevation (flood stage) or rainfall intensity between the sites covered by different agents in the river basin will be the most important situation in this administration process. The information extracted from the datalogger collected and fused in databases and controlled and managed by the different level nodes presented here.

The framework considered for the design and development of the MAS and computational intelligent tool to forecast and simulate flow scenarios for evaluating the consequences of critical flood surges is a complex task. Therefore, to achieve the building of this MAS, it will be necessary to follow some guidelines or directives that guide us in the elaboration of the scheme and its implementation. In this sense, among the various existing methodologies in the literature, it is chose to follow the one presented by Magid et al. [301] and apply the conditions, pertinent to the analysis for the procedures, the development of the organizational structure given by the selected Belief-Desire-Intention (BDI) model [233] for each agent individual behavior, for example, as how it is structured in the GAMA platform [302] respectively.

The next section details the proposed MAS model framework as a flood forecasting approach. Subsequently, Section 3.2 is introducing the developmental stages in the setting up of the MAS model for flood forecasting. Hereafter, Section 3.3 provides a detailed view concerning the methods provisioned on intelligent agents for the implementation of the MAS model agent's behavior and evaluation. Finally, Section 3.4 highlights the integration of the BDI concepts into the flood forecasting model.

3.2 MAS Platform Development for Flood Forecasting

It is known from the literature (section 2.3) that there are several agent-based development platforms for deploying agent-based simulations concerning the elaboration of the proposed MAS for flood prediction. However, this search noted the Generic Agent-Based Modeling Architecture (GAMA) platform [303] as an agent simulation platform that offers the potentials for both micro and macro-model simulations, it is GIS oriented and it includes the feasibility to implement the BDI model. The overall setup of the agents conveys an arrangement in which agents can be connected to the deployed hydrometric sensor network or a database. The agents fetch knowledge of the physical conditions of the river reach through the sensors it links them to (e.g., rainfall, surface water elevation, and discharge sensors). As there are several options available for implementing this system, hardly the other platforms offer specifically the capabilities of the "agentification" of the catchment components as GAMA. It is provided a schematic illustration in Figure 3.1 which illustrates the main idea for the flow of this information among the agents. The GAMA platform has been in development since 2007 by the "MSI research team", whose headquarters is situated at the "Institut de la Francophonie pour l'Informatique (IFI) in Hanoi", which is part of the Programme Doctoral International (IRD) and the UPMC which is an International Research Unit (UMMISCO) [302].

Given that the river basin domain is a very complex system, it is proposed the use of a hierarchical aggregated structure with a five-level architecture to enable the scalability and modularity in the MAS platform to be chosen. Therefore, the principal awareness of the agent's sensor verification, information pre-processing, and storage, system classifier, and user interface must be able to be offered by the agent-based platform of choice, to provide the management of data received from the hydrometric sensors network deployed in the river basin using multi-agents to capture, filter, control, and administrate the data with minimum human intervention for the instantaneous prediction of inundation surges.

In regions of tropical river basins, the problem with hydrometric data is that they contain many possible cases of data scarcity [304], hence, this issue would entail, the initial processing of missing values, outliers, redundant or even rendered impaired, so as the purpose of the framework is a real-time prediction of streamflow, and the problem can be dealing mainly with the scarcity of data, then the need is to have data ready at hand so they can be able to manage the threats and disaster that floods inflict on society, then they must ensure that the availability, integrity and the quality of this data are met, to put to advantage this knowledge and to provide good decision support for expert knowledge.

To achieve this task, it is proposed an intelligent model for fetching, pre-processing, gathering, and cataloging the hydrometric monitoring data. The proposed model as it receives data, stores it in a historical database (i.e., historic data storage agent) from which the agents (e.g., river flow agent, rainfall agent, and water level agent) access it, and begins data pre-processing. It renders the data captured reliable and uses the MAS for the task. Therefore, the resulting MAS can be organized with seventeen entities displayed throughout five levels of interaction.

3.2.1 Formulation of Agents Definition

This section describes the agents of the MAS conceptual framework that are used in this system setup presented here as a smart agent-based solution for flood forecasting, hydrometric data management, and flow inference information extraction. The information entering the system is obtained from the environment by the field stations equipped with hydrometric sensors through intelligent sensing agents.

The system comprises five principal levels, with each one displaying certain specialization and integrating the intelligent agent components for its best operation:

1. Hydrometric sensor network level (HSnL)
2. Sensor data preprocessing level (SDPPL)
3. Historic data storage level (HDSL)
4. System classifier level (SCLAL)
5. User interface level (UIL)

Figures 3.1 and 3.2 illustrate the agents that configure and collaborate within the MAS model.

3.2.1.1 Level I: Hydrometric Sensor Network:

- Hydrometric sensor agents (**HSn):** These agents are represented by the hydrometric station sensors deployed in the field. Their roles are described below.
    1. Rainfall sensor agent (**AgentRNSn):** The AgentRNSn role is to obtain, aggregate, and forward the real-time incoming rain data readings obtained from the field rain gauge sensor connected to the hydro-station datalogger which would be used by the agent DPP, Data2Lags, {AgentFCST1, . . . , AgentFCST8}, and FL.
    2. Water level sensor agent (**AgentWLSn):** The AgentWLSn role is to obtain, aggregate, and forward the real-time incoming river surface water level data obtained from the field water level sensor connected to the hydro-station datalogger which would be used by the agents DPP, Data2Lags, forecasters = {AgentFCST1, . . . , AgentFCST8}, and FL.

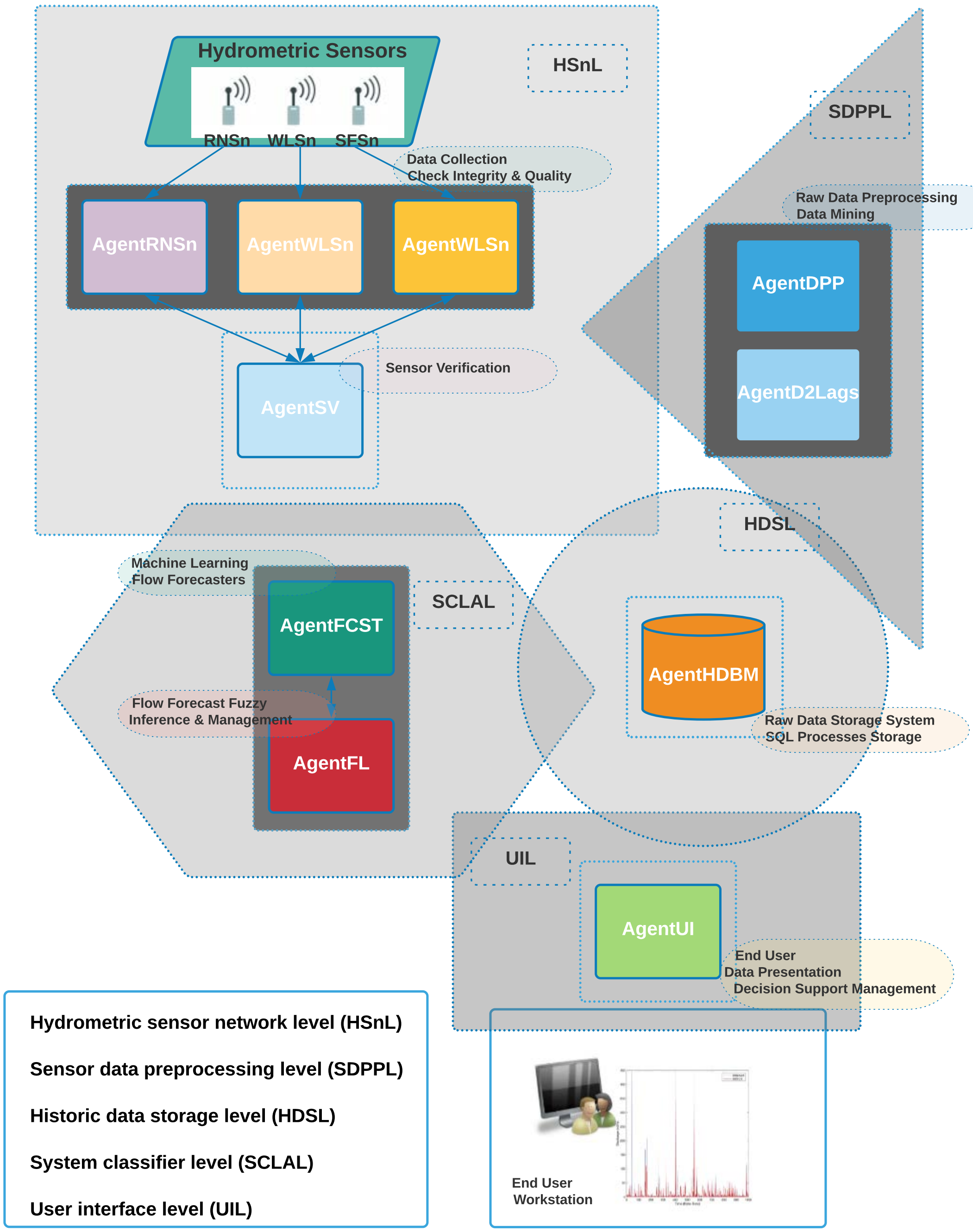


Figure 3.1: An abstraction of the proposed MAS model architecture for flood forecasting showing each of the five overlapping levels of operations that integrates the flood forecasting and inference management process.

3. Streamflow sensor agent (**AgentSFSn):** The AgentSFSn role is to obtain, aggregate, and forward the real-time incoming flow discharge data obtained from the field flow meter sensor connected to the hydro-station datalogger which would be used by the agents DPP, Data2Lags, {AgentFCST1, . . . , AgentFCST8}, and FL.

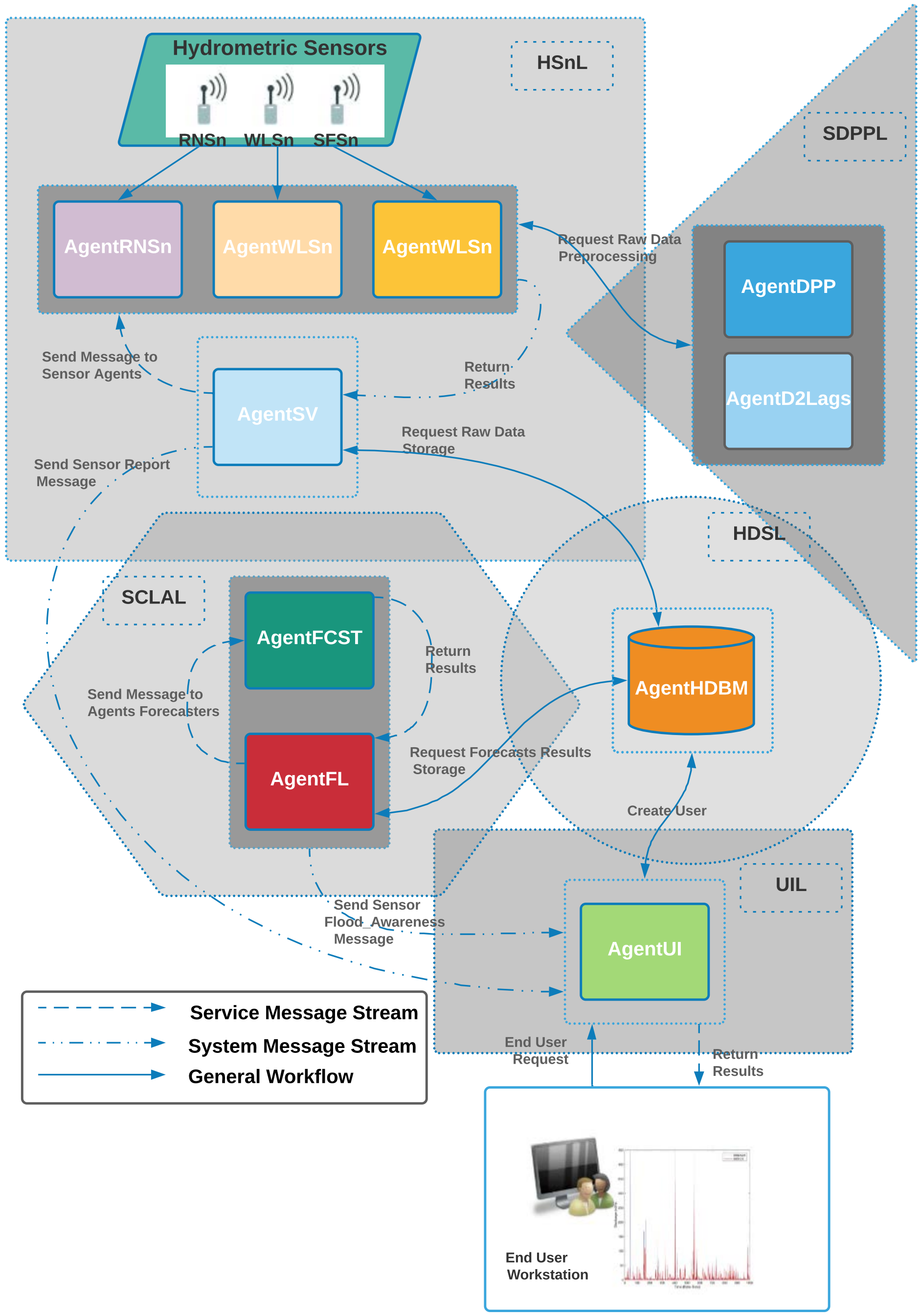


Figure 3.2: Schematic diagram showing the MAS model agent deliberations and collaborations. The communication between the hydro agents on each of the system's levels of operations shows the direction of the flow and the exchange of information from the agents for one level to another in a bidirectional manner.

- Sensor verification agent (**AgentSV):** Its central roles are to verify the operational status of the hydrometric sensors, then inform the agent (AgentUI) about the sensor's functionality and operability. This agent also as communications among the agents AgentHDBM, and of the hydrometric sensors (e.g., AgentRNSn, AgentWLSn, and AgentSFSn respectively).

3.2.1.2 Level II: Sensor Data Preprocessing:

- Sensor data preprocessing agents (**AgentDPP) and (AgentData2Lags):** The AgentDPP agent has among its roles the preparing of the hydrometric sensor data obtained from the HSn agents, by performing data preprocessing (i.e., data treatment, data imputation) of the rainfall, water stage, as well as streamflow. At the same time, the AgentData2Lags role (functionality) is to perform the transformation of the captured data, into a matrix form with their respective lags and lead-time. This is the required format to perform supervised machine learning by the forecasters agents. The preprocessed data is stored by the AgentHDBM and will serve as input data that the AgentCLA's would use for flow forecasting and assessment by AgentFL.

3.2.1.3 Level III: Historic Data Storage:

- Hydrodatabase management agent (**AgentHDBM):** This agent is responsible for the administration and storage of all DataStream through the MAS environment, for logging communications and storing sensor readings and share this information with another agent or agents. The AgentHDBM can function as a service agent for agents and at the user interface level.

3.2.1.4 Level IV: System Classifiers:

- Machine Learner Agents (**AgentFCST) and (AgentFL):** In the proposed MAS model configuration, this level is composed of nine agents. These are the agents based on machine learning skills used for performing multi-step streamflow prediction and fuzzy inference for the hydrometric parameters. Below is a brief description of their roles.

    1. Flow forecast agents (**AgentFCST):** The flow forecasting agents is a group of agents formed by the coalition of eight agents which have been named (e.g., AgentFCST1,..., AgentFCST8, respectively). Therefore, these are N stationary agents, where N = 1, 2, 3,..., and in this case N = 8 of machine learning agents (ML) which have been endowed with ML algorithms (e.g., "random forest (RF)", "support vector regression (SVR)"). In this sense, the arrangement of these eight agents is defined by four agents with a specific ML behavior and the other four with another ML behavior for computing each of which will produce the flow regression task at 1, 2, 3, and 4 hours lead time in the MAS model for streamflow forecast estimates. The roles of these eight agents are initialized immediately after

the subsequent time step that the HSn agents have begun to collect hydrometric data, and that the roles preprocessing and storage of the other collaborating agents (e.g., AgentDPP, AgentData2Lags) as put into effect in parallel to the data collection, the eight Agent Forecasters will initiate the corresponding regression models training and forecasting based on the features and hyperparameters of each ML algorithm.

2. Flow inference agent (**AgentFL):** A decision-making agent, the AgentFL, for "fuzzy inference system (FIS)", is charged with the role for rule extraction, and information assessment on flood-awareness (FA). Similarly, with the available forecast that he receives from each of the eight agent forecasters' results, the actions of the agent fuzzy logic involve the assessments of the forecasts from which he normalizes and creates a single forecast period by merging the pairs of the corresponding time frame and computing the means to obtain a final one, two, three and four-hour lead time result. From this information and with the actual readings of flow data, water level, and rainfall, categories are defined for the fuzzification of the hydrometric data elements, and defuzzifying data, to quantify fuzzy variables, the AgentFL makes inferences on the hourly flow forecast. Finally, the overall responsibility of the classifier level is to provide flood forecasting and awareness on a time horizon of one, two, three, and four hours lead time and report flood warning levels.

3.2.1.5 Level V: User interface:

- User interface agent (**AgentUI)**: The AgentUI assists a user that access the flood forecasting results from information provided by the system to a computer input interface unit. This agent also has the task of communicating with other agents as well as users. It permits activities such as altering resultant forecast output or reporting inconsistencies between forecasts, user subscription, receives flood-awareness information from AgentFL and other related field issues such as sensor failures that are announced by the AgentSV.

3.2.2 Agents Communication

Among agents, messaging is accomplished in two ways, the first is through direct messaging [305] and the second through indirect messaging [306]. In the MAS model, the communication between the agents will be carried out following the direct communication principle, which is based on communications theory and compliance with the "Agent Communication Language (ACL)" collected in the FIPA-ACL protocol [307].

3.2.3 Agents Interactions and Relationships

The notion of agents as "autonomous entities", should not isolate their abilities for interacting with other agents for cooperatively accomplishing their different tasks within a given environment. In this sense, agents can sense and communicate with their surroundings and decide within their capabilities which actions to take for a given problem.In complex environments, decisions are complex. Hence, agents communicate with control agents and inform the requirements. In Figure 3.2, is illustrated an instance where the AgentSV informs the user interface agent about the actual status of a sensor upon request. If there is missing information on a sensor, the user interface agent will complain about this missing field. The sensor data verification agent gets the status, validates the request made and performs the action of re-checking the actual status of the sensor to corroborate that is malfunctioning or down.

In the GAMA platform for example, the communication among agents is accomplished through messages. Typically, the messages are delivered via message loops internally from conversation executed among the agents. Hence, conversations are handled according to FIPA messaging configuration.

Implementing these conversations can be challenging. However, the platform offers rich support for several of the most commonly used interaction protocols defined by FIPA. The platform also provides a devoted modeling language that facilities the non-programmer to implement models through high-level primitives committed to ABM. As the interest is in the hydrologic/hydraulic systems models, the GAMA platform as many advantages to implement hydrologic/hydraulic systems environments over the other platforms like JADE, JADEX, and others. Besides, it is an adequate platform for this kind of purpose has it support geospatial information, and imported as GIS data layers into ABMs and facilitates the required spatial operations as needed.

The agents at the Hydrometric Sensors Level handle the functionality of the sensors and the data retrieval task, such that the AgentSV carries out a verification of the sensors to ensure that the sensors are functional or non-functional and the data acquisition process by the three HSn agents. The AgentSV sends a message about a sensor's integrity to the AgentUI upon requests, and he requests the AgentHDBM to store the retrieved hydrometric data within the hydrometric database (Figure 3.2). If a sensor is not functional, corrective actions are requested from the user interface level. The data that is captured by the HSn agents and later stored, is fetched also by both agents' AgentDPP and AgentData2Lags in Level II from the hydrometric database for preprocessing, which is later stored into the hydrodatabase by the AgentHDBM. The hydrodatabase management agent manages to store the information transmitted from either the AgentDPP and AgentData2Lags and the processes performed at levels IV and V in a standardized format within the hydrodatabase and sharing this information with the with the AgentUI upon request. For the administration to carry out flood forecasting, the Sensor Data Pre-Processing level agent will be responsible for pre-processing, the raw incoming data measurements from the different sensors, as it be available to the Classifier Agents Level (i.e., Forecaster Agents, and AgentFL) to be used for flow forecasting, and the issuing of flood-awareness levels and when completed the process will be sent to the AgentUI, upon request, for verification, analysis, validation and reporting. These agents display reactive behavior described by rules, plans and perform the data analysis and forecasting. After establishing this information, the User Interface agent can request this information from the hydrodatabase agent, who stores the forecasting results. The User Interface agent also can instruct the AgentHDBM about collected information stored about a sensor status, notified by the sensor verification agent, allowing a user-specification for conducting a follow-up task-specification namely (i) the target malfunctioning sensor, (ii) data that needs retrieving on-field physically by the user and (iii) the results. If a user accesses a certain interface application from a device (e.g., pc or mobile application and web application) to view a specific report, let us say on forecasting results, the AgentUI can issue a query to the AgentHDBM, and the request can be delivered if available. Communication among the different agents controlled here is through the *"simple_bdi"* architecture built-in function available in the GAMA platform as a plug-in, which facilitates the definition of behavior using the BDI architecture in GAMA, from the abstracted idea of "Behavior with Emotions and Norms", known as the "BEN architecture" based on the work of Svennevig [308].

3.3 Overview of the Agents Behavior Implementation

Previously in Section 3.2.1 it was defined the agents definition, this section introduces cognition (e.g., perceptions, feelings, emotions, and knowledge) and behavioral (e.g., planning, capacities, skills, and decision making) states of each of the agents in this MAS administration.

The behaviors of the agents are developed in reference to the type of communicative actions that they will use to implement in the system.

For the implementation of the cognitive capabilities within agents, the rationale will adapt to the "Belief-Desire-Intention (BDI) model" proposed by [309] and that was later adapted and updated to a model that is much suitable for MAS application by Rao and Georgeff [310] is implemented in each of the agents within the MAS framework . The BDI model is amid the various "deliberative agent architectures" that are in use currently and extensively by researchers in the multi-agent community.

That is why, in this MAS task, it is proposed the adoption of the "BDI-architecture model" for the agents involved in this multi-agent system as follows.

- **Beliefs**

  For the agents, the belief system represents consciousness of their surroundings, as the information a hydrometric sensor-agent as about the environment, in this case, the values of the hydrometric parameters they are monitoring, and the information they have about other known agents responsible for carrying out other specialized tasks, other than monitoring, and in this way they all share information.

- **Desires**

The goal of hydrometric sensor-agents is to correctly capture the values of the hydrometric sensors, so that they are engaged in the permanent monitoring of these registers. The corresponding monitoring plan is being briefly explained: sensor data verification-agent permanently register information measurements from the hydrometric sensors that reach that datalogger. The data collection management agents as the goal to administrate this information collected from the monitoring activities and it runs a scanning process on this data until it finds inconsistencies in the data and stops if it is no longer possible to find these inconsistencies. The other agents in the system, as the desire to make this information be in the right format, free of missing values and outliers and suitable for the classifier-agents to work classification tasks on it.

- **Intentions**

  Intentions represent action plans to accomplish a specific goal. Actions are basically of two types: outward and inward performance. By outward performance, they are referring to the performance of the messaging skills among agents with distinctive supportive negotiations, whereas inward performance comprises instructions given to other agents of the organization (i.e., monitoring system), and/or the hydrometric sensors.

3.3.1 Agents Individual Behavior

It has been discussed in previous sections the roles and the components of the framework, for example, the sensor data verification agent and all the other specialized agents that form the MAS organization. In this section, the architecture proposed has been selected and implemented as a GAMA plug-in provided in the works of Taillandier et al. [311] as it is the most updated BDI model engine adaptation and is much more optimized in terms of memory and computational expenditure.

The GAMA BDI plug-in offers an easy environment for implementing the data structure and statements (i.e., it extends the platform's language (GAML)) for developing agents whose behavior is designed using the BDI model. As noted, it offers an architecture, known as *simple_bdi*, that can be implemented in an agent (species) and that allows modelers to organize in their agents the advantages of the features, reflexes, and functionalities provided by GAMA agents as part of its BDI reasoning mechanism. In addition, it is indicated the type of communicative actions that agents will use to implement in the system in conformity with FIPA protocols.

Depending on the platform of choice for agent development and programming, the "belief-desire-intention" architecture is platform-specific, which means that it generally contains the features and routines necessary for allowing the implementation of the agent's cognition. Therefore, the agents' knowledge is established on the BDI bases for agents' behaviors that are based on perceptions, rules, and plans. Subsequently, it can be recalled that beliefs represent the state that a given agent is at a particular moment in time, about its surroundings, desires are the intentions an agent wishes to fulfill, and intentions are what the agent has decided to do. It must be pointed out that for agents to achieve their intentions, a series of actions (plans) are required. Therefore, in the GAMA platform, the idea of knowledge in agents is represented in a data type known as *predicates*. Predicates form the basis of a structure in which beliefs, desires, and intentions are defined. In the succeeding sections, it presents the pathway that the agents' knowledge and behaviors are represented by in the architecture.

3.3.1.1 Data Administration Agents Behavior

The behavior of the agents has been defined by adapting the FIPA Request and Query Interaction Protocol (SC00026H and SC00027H, respectively). This protocol manages the communication between a requester and a respondent agent. The requester agent issues a request and expects respondent agent to carryout the requested task (see Figure 3.3).

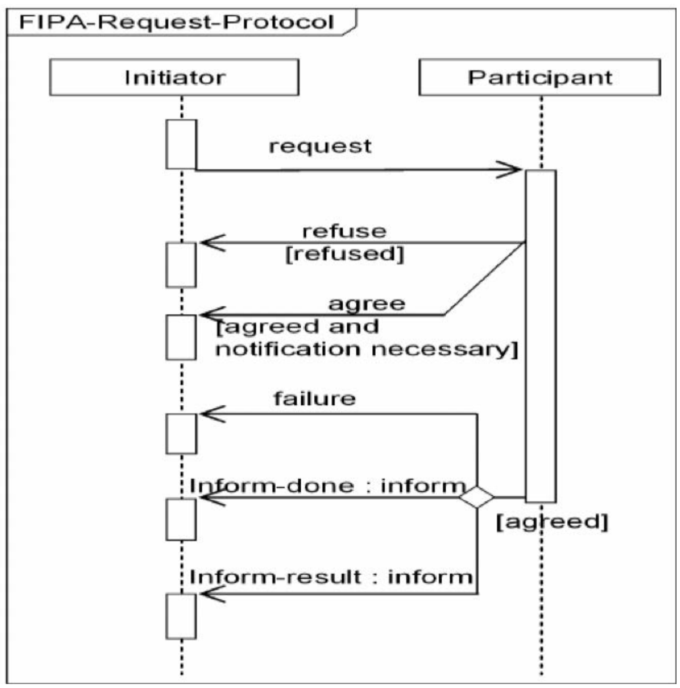


Figure 3.3: FIPA request interaction protocol: Adapted from [312].

- Sensor verification agent

  The behavior of this agent is to keep track of the operation of the hydrometric sensors deployed in the field, by communicating with the hydrometric sensors to ensure from them that there is data inflow, request the storage of this information to the AgentHDBM, as well as to send the information regarding the functionality of a given sensor as it coordinates with the AgentUI to improve the quality of this information. This "dynamic-behavioral" task is depicted in the Figure 3.4 below.

- Hydrometric sensor agents

  These are the hydrometric data monitoring sensor agents whose role is to capture the rainfall, river flow, and water level data generated from the sensors deployed at specific hydrometric stations. This data is produced on time frame intervals by the field sensors, are accumulated in a datalogger device, then transmitted in actual time, via telecommunications modes (e.g., Radio, GSM/GPRS, INSAT Radio).

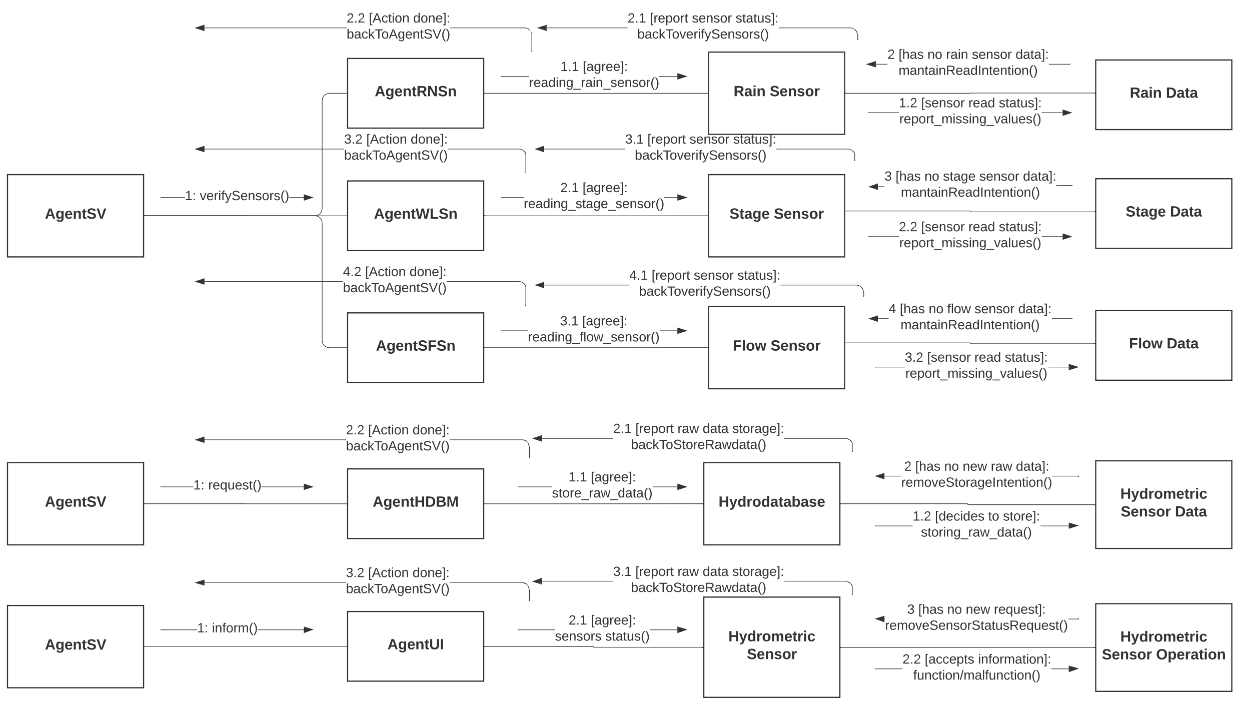


Figure 3.4: Sensor verification agent communication protocol diagram.

3.3.1.2 Hydrometric data Preprocessing Agents Behavior

The hydrometric data preprocessing agent behavior has also been defined by following FIPA Protocols SC00026H and SC00027H.

- Data pre-processing agent

  The information captured from the hydrometric monitoring field sensors is ready for the pre-processing task, such as searching for data issues, filtering, outliers, and imputation of missing instances, and treats the file accordingly for these anomalies by performing the corresponding data treatment. On the other hand, the sole task of the AgentData2Lags is to convert the hydrometric sensor variables to a lagged data matrix, which is the usual format for supervised ML learning task performed by the classifier agents to use the information. On completion of these task, the AgentDPP and AgentData2Lags request the AgentHDBM to store the pre-processed data files, in conformity with FIPA request interaction protocol (Figure 3.3), and also sends a proposal message to the classifier agents that files preprocessed files is ready and can be available for performing classification and inference tasks (Figure 3.5).

3.3.1.3 Data Storage Agent Behavior

- Database management agent The AgentHDBM performs the action of storing and managing the hydrometric sensors data files, flow classification results, system information, reports of sensor failures, data shared among agents and other warning events within the MAS organization as per requested by the Agent' SV, DPP, and CLA's. A query-if (Figure 3.6)is also performed by the agent AgentUI to request information on the system and then it act as a responder, returning information on sensor status and the results of forecast and flood-awareness for inspection and analysis by the human user.

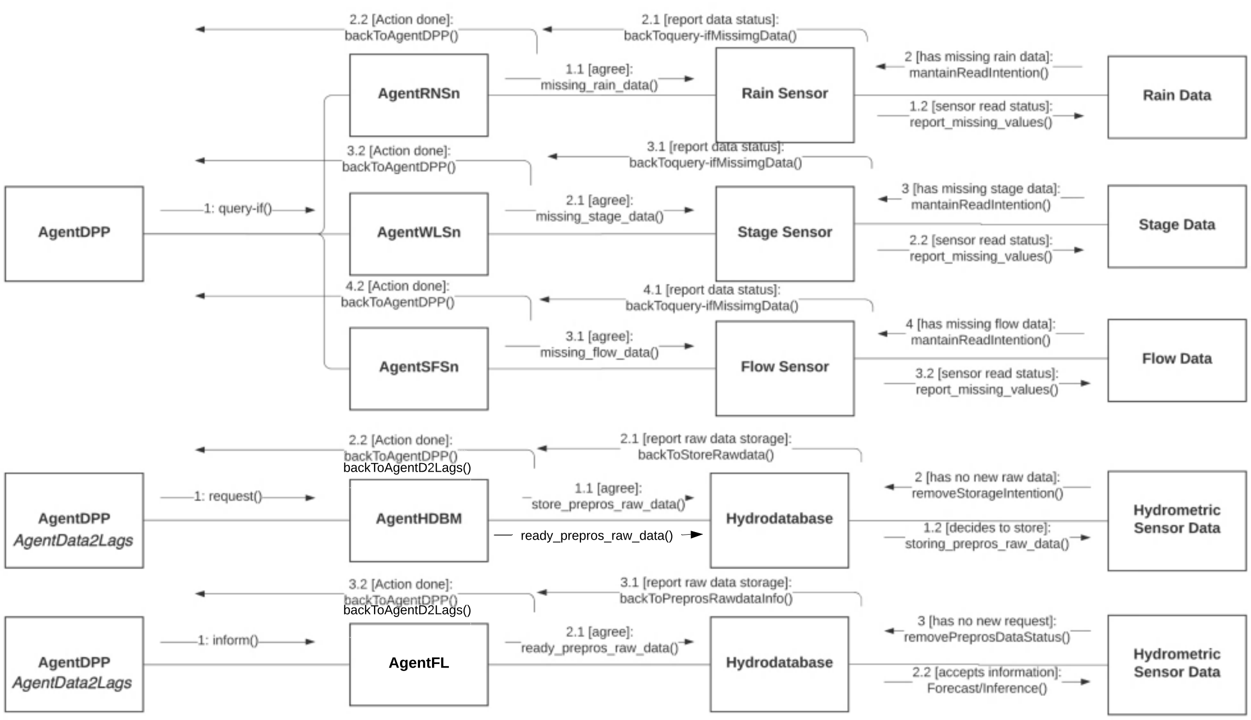


Figure 3.5: Data preprocessing agents communication protocol diagram.

3.3.1.4 Classifier Agents Behavior

- Agent's ML and AgentFL

  The agents of this level, are accountable for handling the streamflow prediction task and processes from the hydrometric stations as described under Subsection 3.2.1.1.

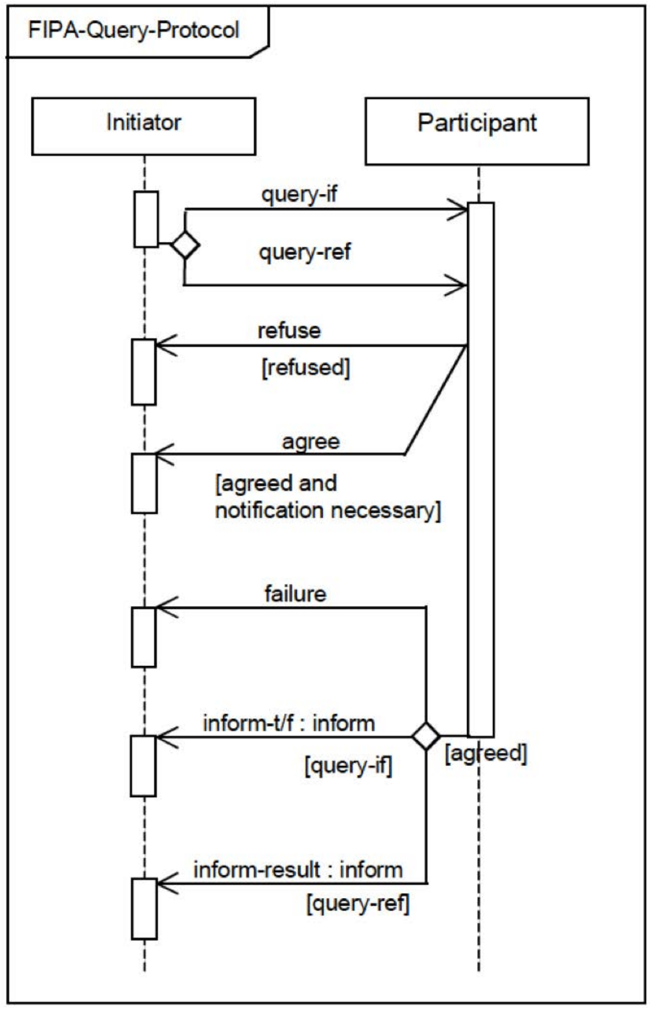


Figure 3.6: FIPA query interaction protocol: Adapted from [312].

They represent the following classes: machine learning agent' behavior (AgentML), a fuzzy rule agent behavior (AgentFL). The forecaster agents takes the incoming data directly from the HSn agents or data that is stored for performing flow forecasting. The forecast results, upon request by the AgentFL is used in the inference process tasks, and can if necessary make use of hydrometric forecast information derived from external data sources in conformity with FIPA Protocols SC00026H and SC00027H (Figure 3.3, 3.6). When this goal is completed, the AgentFL informs to AgentUI on flood-awareness levels as depicted in Figure 3.7.

3.3.1.5 Management Interface Agent Behavior

- User interface

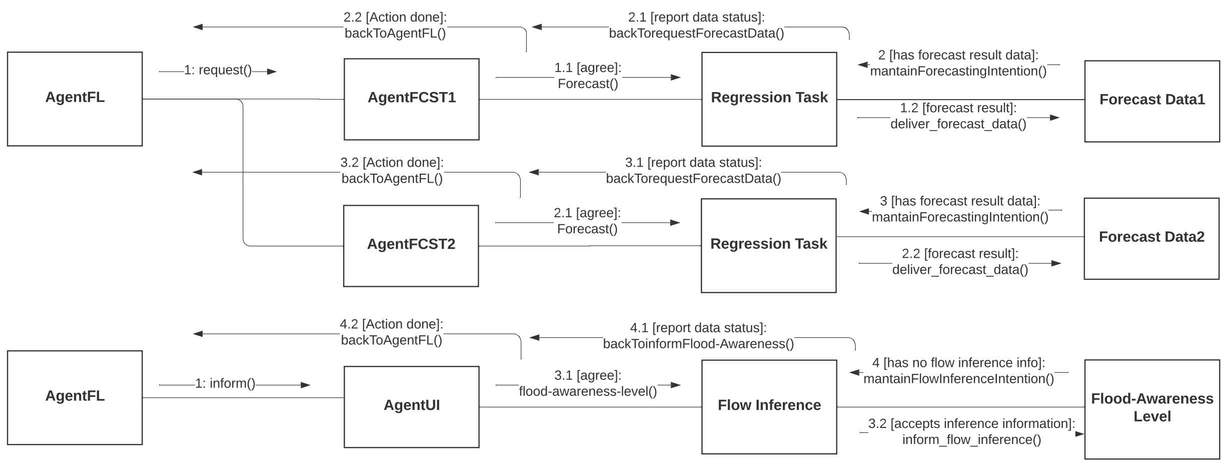


Figure 3.7: Fuzzy logic agent communication protocol diagram.

The user interface agent (AgentUI) acts as the liaison officer of the MAS organization and the external users. Therefore, it acts as a consumer gateway that makes available the information of the MAS to users. Its behaviors are defined by adapting the FIPA Request and Query Interaction Protocol (SC00026H and SC00027H). The services offered by this agent include providing users with real-time flood forecasting information. It has the responsibility of subscribing users to the system, in allowing users to control the feed of external hydrologic/hydraulic models information files, to provide user-specific information, outputs of sensor reports, of the MAS, retrieve relevant and conditioned data, modify and/or include new rules, check reluctant forecast results and customizing the outputs to user-specific needs as shown in its communication protocol (Figure 3.8). It coordinates with other agents within the MAS organization and the graphical user interface.

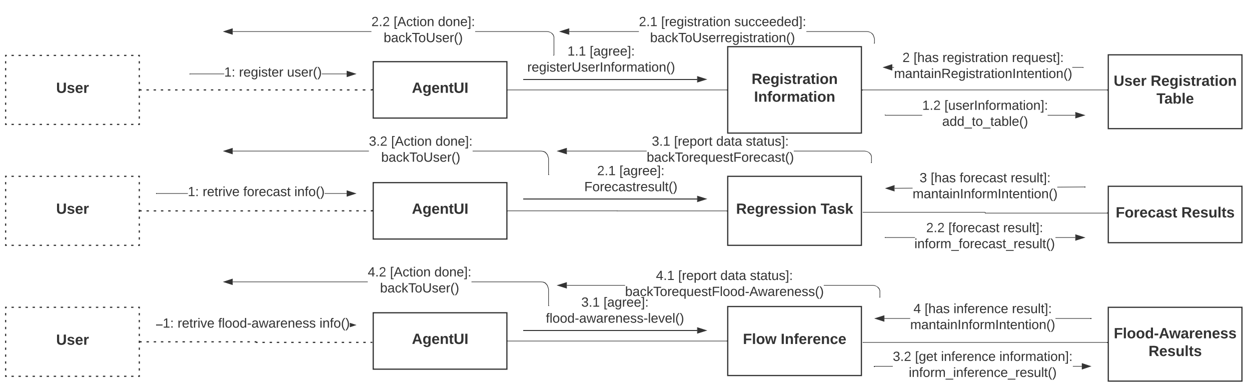


Figure 3.8: User interface agent communication protocol diagram.

Chapter 4

Experimental Setup

In this Chapter, the evolution of the ABM formulation used to achieve the objectives described in Chapter 1 is presented, the purpose of which is to confer an easy-to-use software toolkit for water managers, professionals, and the layperson, capable of delivering flood forecasting within the tropical watershed domain. Moreover, to address the initial setup necessary for building the proposed agent-based approach, that will satisfy catchment modeling, the following must be attained:

- Provide the hydrometric data of the catchment.
- Have the necessary GIS shapefiles and "Digital Elevation Model (DEM)" Grids of the watershed.
- Have physiographic parameters and constants of the catchment and the hydraulic data information of the river.
- Provide the necessary machine learning algorithms for supporting data collection, preprocessing, flood forecasting, and inference.
- Guarantee the MAS model with the presence of agents capable of reasoning about data quality, flood forecast, and inference.

In order to simulate the flood concerns explained in earlier chapters, the time series information collected from a hydrometric station in a Panamanian basin was used, and from this information was the simulation of the flood environment for a tropical basin recreated, with the GAMA agent-based modeling platform. The selection of this watershed was because it has been heavily exploited by mining activities and other development projects that are affecting its flood plains and drainage capacity. However, for a broader understanding of the problems therein, more details about this experimental basin and the nature of the data related to it are presented below in Subsections 4.1.1 and 4.1.2. The remaining portions of the chapter are outlined as follow: first, Section 4.1 through 4.3 it is described the experimental data used, secondly, Section 4.4; is presented the conceptual framework upon which the problem domain was implemented, and the data analyzed, Section 4.5 defines the basis for the ABM extension with the BDI architecture on the MAS model concept and is finalized under Section 4.6 with several preliminary findings and subsequent works are made concerning the modeling results of the rainfall-runoff process obtained using HECHMS and the ABM approach with the GAMA platform.

4.1 Model Domain, Study Area and Hydrologic Data

4.1.1 Problem Statement of the Model Domain

As it is known, a catchment, watershed, or river basin is an essential ecosystem with unique physical characteristics of urban hydrology as it is the source of natural resources within an eco-zone, delimited by river courses, and composed of the flora, fauna, villages, and neighborhoods. Notwithstanding, as a watershed can provide a reservoir for drinking water for both people and wildlife, on occasions when there is too much of this resource, water can represent an imminent threat to human and animal lives, agriculture, pollution of the environment, and the economy, as mentioned earlier. Therefore, as "floodwaters" can represent hazardous situations to any community, need for its management is necessary for assessing the sociological, economic, environmental, and safety aspects of the catchment. Figure 4.1 shows the watershed model that was implemented with the GAMA platform tool which allows using geospatial information, digital elevation models (DEMs), satellite imagery, orthophotos, and vector data from geographic information systems to create models that can agree with the domain of study.

4.1.2 Domain Description and Hydrologic Dataset

The Medio River subcatchment, (see Figure 4.2) is part of the major Caimito's river catchment located in the Republic of Panama in the Donoso region. The Medio River flows predominantly northward and combines, along with other tributaries to form the Caimito River about 7 km from the Caribbean coast. At present, there is a major ongoing project of copper mining extraction components located at the upper sub-basin, including the mill, waste rock storage facility, and the tailings management facility (TMF) and it is scheduled for construction of other future facilities. The hydrometric data employed in this experiment was obtained from two real-time data acquisition system that monitors, log, and transmit 15 minute and 1-hour interval rainfall, river water level, and river flow data during low and high flow periods from installed rain gauges, flow meters, and surface water elevation radars. The data recorded span from March 31st, 2012, to December 31st, 2016. Table 4.1 it is summarized the mean annual rainfall associated with each of the long-term regional climate stations in the Donoso region. Both H3 and H4 stations were installed in early March of 2012, so they are relatively new, as such, they are not mentioned in the tables below, and therefore, they do not possess an extensive hydrometric data record. Consequently, based on this information, the mean annual rainfall in the Donoso region varies from approximately 5,000 mm at the

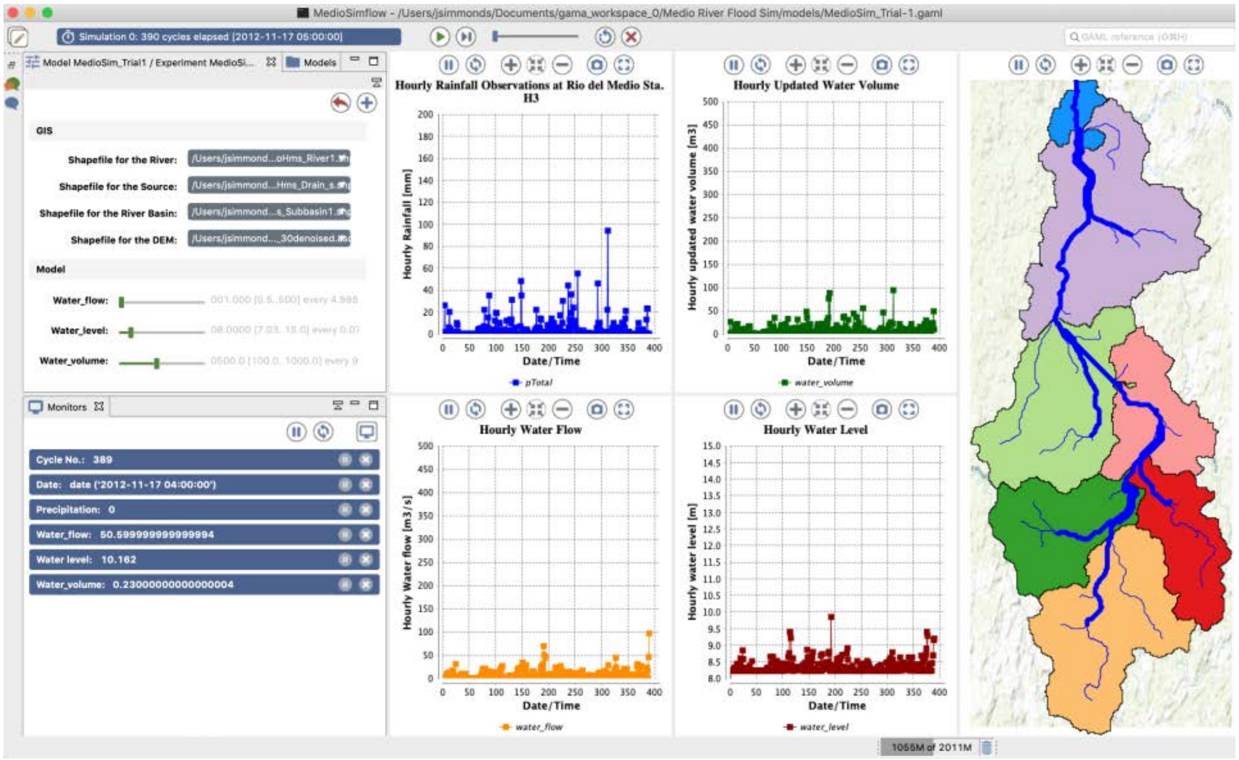


Figure 4.1: GAMA platform interface showing the flooded Medio River Subcatchment.

coast to 3,200 mm inland. Annual rainfall amounts associated with extreme wet and dry conditions are shown in Table 4.2. The stations also have the capability for monitoring turbidity, temperature, and conductivity data. The installed monitoring stations (Station H3 and H4) as labeled in the environmental assessment study are located at the upper and lower parts of the Medio River Catchment (08°52′07.2”N80°39′57.1”W and 08°55′58.0”N80°40′07.6”W).

Table 4.1: Mean Annual Rainfall in the Donoso Region, Panama. Source: MPSA [313].

Station	No. Years of Record	Approximate Distance from Coast [km]	Elevation [m]	Mean Annual Rainfall [mm]
Cocle del Norte	43	0	2	4,989
San Lucas	43	10	30	4,716
Boca de Toabré	43	20	30	4,413
Coclesito	33	30	60	3,171

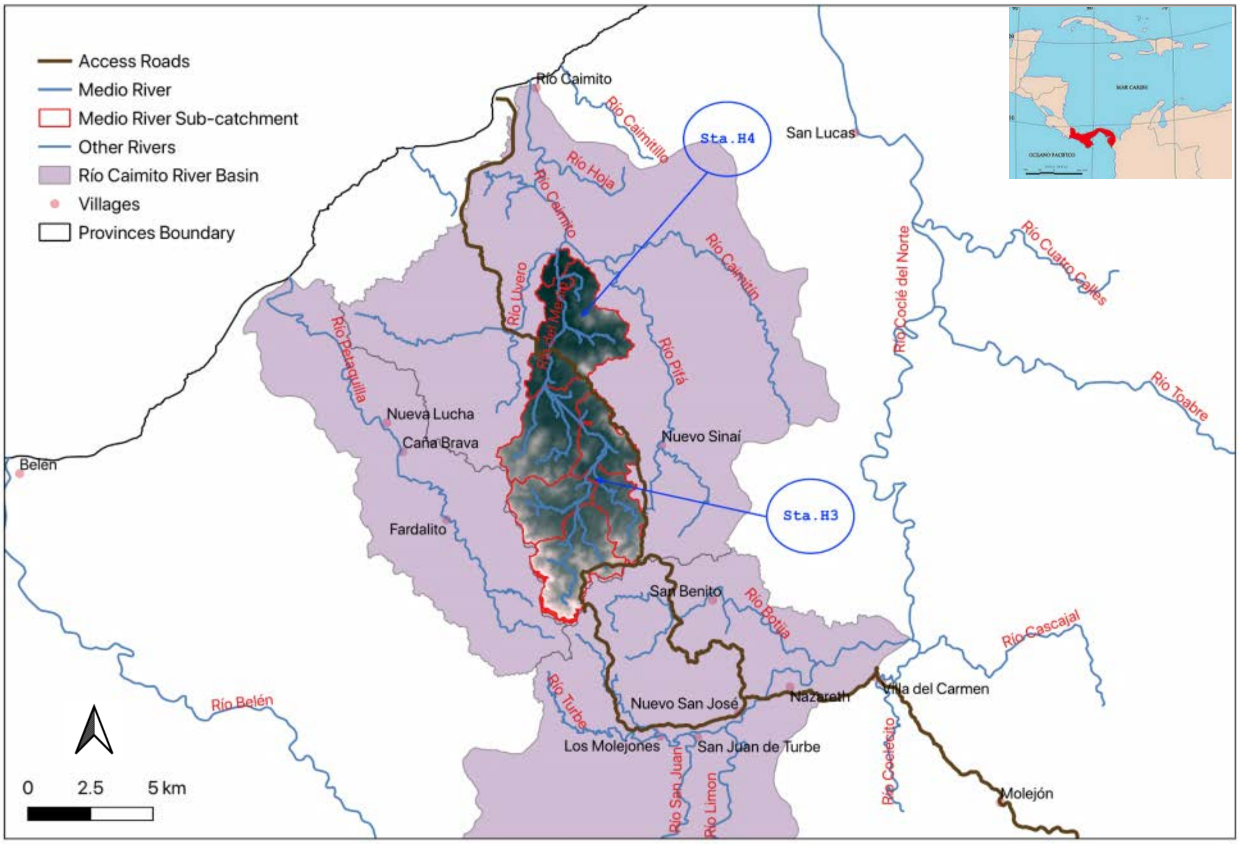


Figure 4.2: Medio River Sub-catchment in Upper Caimito River Basin.

Table 4.2: Extreme Annual Rainfall in the Donoso Region, Panama. Source: MPSA [313].

Return Period	**Cocle del Norte**	**Annual Rainfall [mm]**		
		San Lucas	Boca de Toabré	Coclesito
Number Years of Record	33	40	39	33
Highest Recorded	8,836	6,715	6,239	5,195
Average	4,989	4,716	4,416	3,171
Lowest Recorded	3,164	3,420	2,990	2,491

4.1.2.1 Catchment Soil Type Profile

At the catchment scale, soil classification types are fundamental for calculating the infiltration and runoff rates as they determine the different hydro-soil groups. Each soil profile has a specific infiltration rate. The upstream region of the Río Medio catchment is a mountainous region with steep terrain whereas the lower part of the catchment lies in flat plains with mild slopes. The dominant soil type type [314] in the catchment is haplic nitisols, acrisols, and vitric andosols. A high percentage of clay, composed of various mixtures (i.e., clay-loam, sandy-clay-loam, and sandy-loam) as a major soil type is present throughout the soil of this entire catchment area. Soil types in the catchments are shown inTable 4.3.

Table 4.3: Soil Type Composition in the Medio River Catchment. Source: FAO-HWSD [315].

	Dominant Soil Type	**Associated Soils and Inclusions**	
Order	1	2	3
Land Group Label (FAO 90)	Haplic Nitosols	Haplic Acrisols	Vitric Andisols
UpperLand Grain	Medium	Medium	Medium
Land Source Deep (cm)	100	100	100
Catchment Type (0-0.5% slope)	Moderately Well	Moderately Well	Moderately Well
UPPERLAND ("Sand Fraction") (%)	45	48	66
UPPERLAND ("Silt Fraction") (%)	24	23	29
UPPERLAND ("Clay Fraction") (%)	31	29	5
UPPERLAND "USDA" Grain Category	clay loam	sandy clay loam	sandy loam

4.1.2.2 Catchment Land Use Profile

The entire western and northern parts of the catchment area are dominated by dense forest and evergreen broadleaf forest and some croplands on the lower northern areas. The catchment's eastern and south-eastern zones are characterized by allocating land that is dominated by woody savannas, savannas, grasslands, with some development of the mining industry, and on the south-eastern side, there is the presence of permanent wetlands.

4.1.2.3 Rainfall Data Profile

The Medio River catchment precipitation data is composed of short hydrometric records from both stations H3 and H4, containing several cases of missingness in the series, as shown in the gaps displayed in Figure 4.3. Nevertheless, the data for station H4 is extremely impaired in comparison to that of station H3, the reason why it was decided only to use the H3 station data for this research proposal. The issues related to missing values and gaps are known to arise for many different reasons including but not limited to instrument damage that results in data gaps, data inconsistencies, incorrect logging of timestamp, duplicates, data loss, and station vandalism (as occurred in the H4 station downstream). The field technician may also attribute the data gaps to several factors such as the malfunctioning and damages to meteorologic instrumentation, negligence in managing field readings, climate disasters such as floods, cyclones, hurricanes, bush fires, and anthropogenic intervention perpetrated by the willful destruction of gauging stations and instruments by bystanders, passerby, or strollers. Because of the former, these problems make up the factors that hinder the proper assessment in water-related management. Consequently, the lack of consistent and complete hydrometric data can represent the loss of valuable and necessary information to carry out models of hydrological processes in any stage of planning and construction of hydraulic works and for the implementation of decision support systems for the prevention of floods and risk assessment in highly vulnerable zones. Thus, given those reasons, to have data available for use in the experiments in this section, it was necessary to deal with the missing data impairment of the hydrometric series; therefore, in the following Section 4.2, it is shown the data treatment task that was aimed at improving the existing time series.

4.1.2.4 Streamflow Data Profile

During the dry season, the river water input comes mostly from base flow; however, the Sub- catchment, as a flow regime dominated primarily by rainfall. The maximum occurs between May and November; however, September and October are the rainiest months. Therefore, monthly discharge can vary according to the hydrological regime of a year, as subjected to El Niño/Niña-Southern Oscillation (ENSO) effects, and other meteorological causes such as flash floods produced by frontal systems, tropical cyclones, and mesoscale convective systems [316]. But river floods mainly occur in the rainy season in the form of long-lasting precipitation, flash floods, and thunderstorms. Despite the rain data used in the simulation, there are limited real-time hydrometric datasets for the Medio River catchment. Therefore, the only two available hydrometric stations, from which data could be fetched were from station H3 that is located upstream of the catchment, specifically in subcatchment-1, and station H4 located downstream, at subcatchment-4. Unfortunately, the only workable dataset with some degree of missingness in the data is the datasets for station H3, as station H4 datasets were found incomplete and impaired as was mentioned earlier in the text. Station H3 began recording data on March 31st of 2012, and the last recorded data available was for the year 2016. The length of the station H3 datasets is roughly five years. Figure 4.4 shows the raw data series with gaps in station H3 hydrograph. However, it is important to mention that to model the catchment, it was necessary to reconstruct the time series for Station H3 using data imputation techniques, which is discussed thoroughly in Section 4.2.

4.1.3 Flood Forecasting Ontology

A domain ontology should have the purpose of explaining the different components that make up a specific modeling domain, along with the relationships that occur between the different components involved. With this said, the ontology of the phenomenon of study should capture the concepts, attributes, and interactions occurring in the system while considering its constraints, rules, and boundaries.

Building an ontology is an engineering task that requires much effort. They have documented several development processes in several works [317–320], as design methods and guidelines, to ease in the development process, their execution, assessment, substantiation, reuse, and life cycle until its deployment. However, none of the existing guidelines is a fixed standard, and given the type of ontology, the designer can freely adopt any method or even reuse a combination of these, or whatever is convenient for the study domain ontology [321, 322].

Ontologies, although the thought of being labeled as vocabularies and taxonomies, their principal purpose is to share and reuse knowledge through request. This means that on-

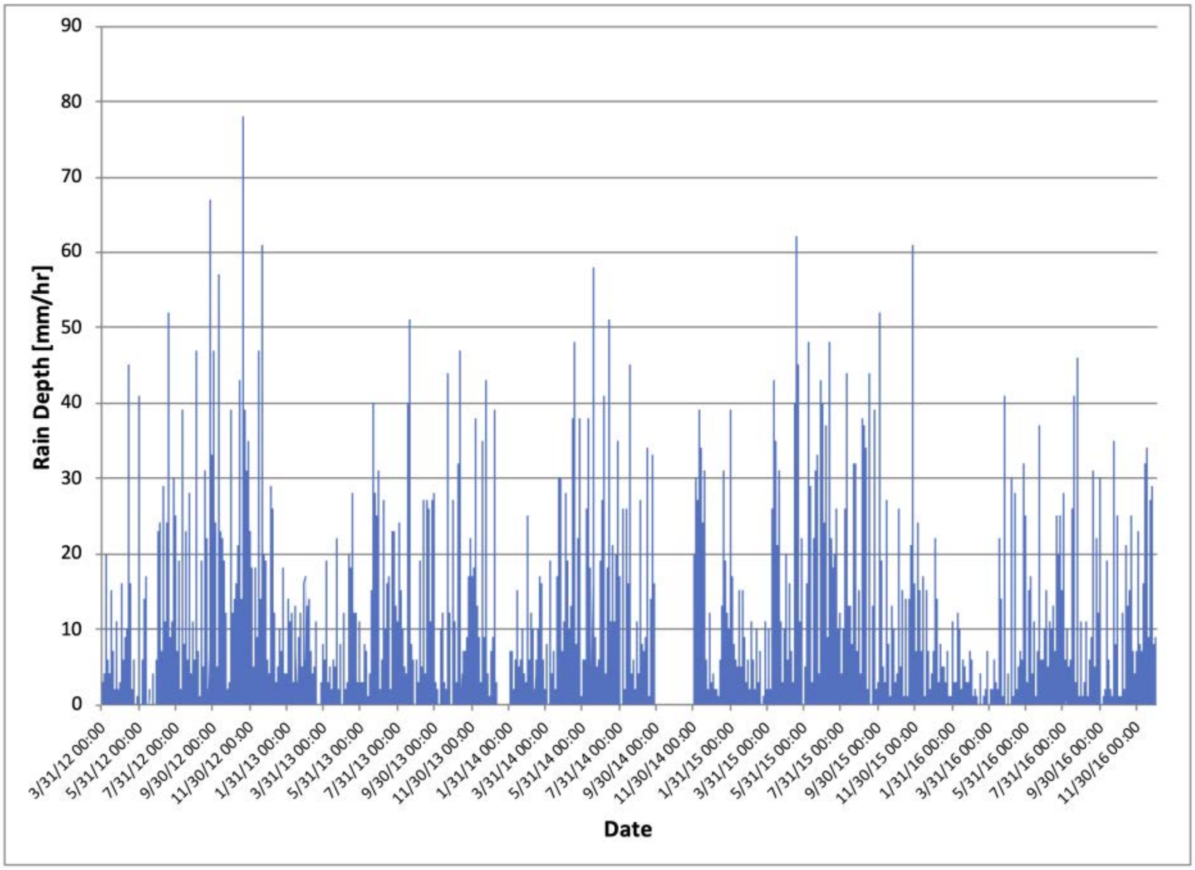


Figure 4.3: Hyetograph at Medio River Subcatchment in Upper Caimito main River Basin.

tologies can provide a characterization of the topics and existing connections present in a specific domain which can become distributed and reused between rational agents and their uses [323]. To share and reuse ontologies requires an understanding of the domain terminologies, requirements, descriptions, and compatibilities of the model prototypes implemented in each of the agents and modules integrating the domain of study [324]. Therefore, to facilitate the flow of knowledge, reuse, and sharing of ontologies, the ontology should be of a specific domain.

Some examples of ontology sharing, and reusing can be found in disciplines such as in earth and environmental sciences [325, 326], building construction and automation control [327, 328], process and systems engineering [329, 330], Internet of Things (IoT) [331], e-learning [332], biomedical engineering [333], production factories [334], and multi-agent systems development [335–337], just to mention a few.

4.1.3.1 Selection of a Flood Ontology

How to select and reuse an adequate flood domain ontology, considering the complexities of the flood modeling problem-solving task? As previously noted, there is no consensus for the development of ontology engineering; in the same way, it is emphasized that for their reuse,

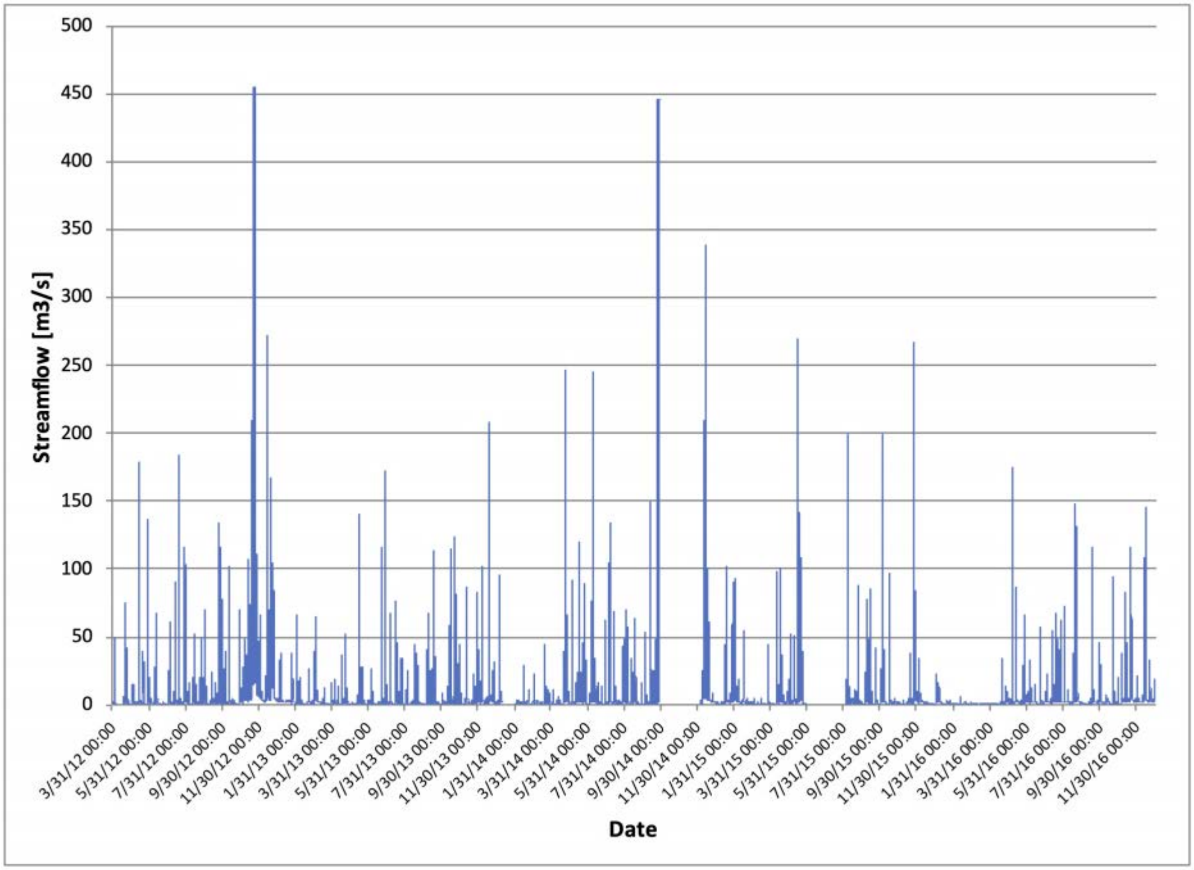


Figure 4.4: Hydrograph at Medio River Subcatchment in Upper Caimito main River Basin.

limitations also exist. However, the methodology suggested in "METHONTOLOGY" [338], Neches et al. [324], and Suarez et al. [339], for the design of ontologies, suggest the reuse of existing vocabularies to avoid redundancy between terminologies, given the heterogeneities that exist within them. Among the existing flood ontologies identified in the literature, a few have components that have a relationship to flood forecasting class and are mainly focused on flood mitigation and risk management.

The guidelines of some methods were necessary to follow, to resolve the selection of one or several of the flood ontologies identified in the literature review and it was the basis to solve the query at the start of the preceding paragraph.

To reuse an ontology, it must follow a reasoning and systematic order. According to Ruy et al. [340] there are two ways for referencing the structure in an ontology intended for reuse purposes; one is by analogy and the other by extension. According to Rittgen [341]. In ontology reuse by analogy, involves the search for structures that define the knowledge linked to a particular domain problem that exists now. After this structure has been identified, the next step is to distinguish which topics within the ontology for reuse are like those in the domain ontology and then replicate the structure of this blueprint in the domain ontology. Whereas, by extension, the blueprint ontology is inserted into the ontology that is been developed, which can be further extended, updated, or include new concepts, relationships, and properties of the classes.

The reuse of the concepts of the vocabulary that are closely related to the present domain of interest, and to decrease time in the development and avoid repetitiveness in their adaptation, it is chosen the reuse by analogy method, explained in [341]. Therefore, it was decided on reusing some of the existing flood ontologies found previously in the literature review, such as those implemented in [52] and [55, 56] as they included useful concepts related to hydrology, hydraulics, and sensor networks, in adapting and aligning domain vocabularies and extracting the usual core knowledge that represents the classes relevant to flood forecasting.

4.1.3.2 Ontology Implementation

In this subsection, it is described the Flood Forecasting Ontology, proposed with reuse method, and adapting terms from other flood-related ontologies. Therefore, instead of creating a new flood ontology from scrap, it is reused the concepts from the existing flood ontologies (e.g., [52], [55], and [56]) wherever available. It is observed these ontologies to gather the knowledge, concepts, and the relationships of hydrologic and hydraulic domains, along with hydrometric data capturing stations for reuse but required it be prearranged in a manner suitable for an application domain.

As referenced priorly in the literature, ontology development can be a complex task. Despite the most sophisticated ontology building tools, a major setback in ontological engineering lies in the languages, platforms, ease of use, learning curve, and methodologies they use. Based on software engineering techniques, many ontology tools exist, and these can vary according to their different formats, which may need conversion into one format to use them. However, for a detailed study on some of these tools, see the works of Alatrish [342] and Slimani [343].

Given the feasibility of its application, the ontology tool of choice is "GenMyModel" [344] a web-based modeling platform that uses the "Unified Modeling Language (UML)". UML enables a usual and dynamic way for visually modeling and designing a process as behavior and structural illustrations [345]. One of the features of "GenMyModel" is the capability to represent ontologies in a parsable intermediate language according to OMG standards via the "Extensible Markup Language (XML)". Besides, the "Metadata Interchange (XMI) standard" allows UML classes, with their variables, fields, and attributes to be expressed through the "Extensible Markup Language (XML)" [346].

Following the guidelines and procedures for the reuse of ontologies described in [340] and [341], the ontology for flood forecasting is implemented under three main ontological categories (e.g., Climate Events, Hydrometric Instrumentation, and Environmental Catastrophes) which is defined below.

Climatic or Atmospheric Events are understood as the "difference between the value of a climatic element in a specific place and the average value of that element averaged over the circle of latitude through that region" [347]. These events are the causes for the major disasters that affect the human population, taking both human and animal lives, causing damage to infrastructure, and to the economy [1, 2]. Regarding the climate and given the scope, nature, and magnitude that climatological events can cause, this category has extended into three important groups, associated in either a direct or indirect manner with disasters caused by floods; however, it is emphasized on those caused by hydrological events as seen in Figure 4.5. In the ontology, under the Flood concept, it can be realized those related elements that make up this concept, such as, for example, river floods, coastal floods, storm surges, overland floods, and flash floods. The possibility of being able to extend the ontology lies in its ability to expand the link of the semantic relationship that exists between the concepts related to climatic events, since, for example, in those areas that have been devoid of vegetation because of forest fires, observed an increase in the amount of runoff overland, that agrees with the findings in [348].

Figure 4.6, which constitutes the Category for the Hydrometric Instruments, represents the hydrometric instrumentation necessary for monitoring the variables responsible for the flood phenomenon and its outlook. Monitoring and compilation of hydrometric data are of utmost importance in the study and understanding of the climatic phenomena for implementing models that represent the rainfall-runoff process. Hence some examples of hydrometric instrumentation are Streamflow Sensor, Rain Gauge, Rain Radar, and Water Level Gauge, or Radar.

Defined under the Environmental Catastrophes category (Figure 4.5) are other environmental concepts involved throughout the life process of the flood phenomena. This category is subject to the domain of climatic events, is in close relation with the Hydrological, Meteorological, and Socio-environmental subcategories. In this sense, when satisfying the ontology, it needs to integrate the three categories defined through their semantic relationship for flood forecasting. Table C.1 below, shows the glossary of terms that make up the ontology for flood forecasting.

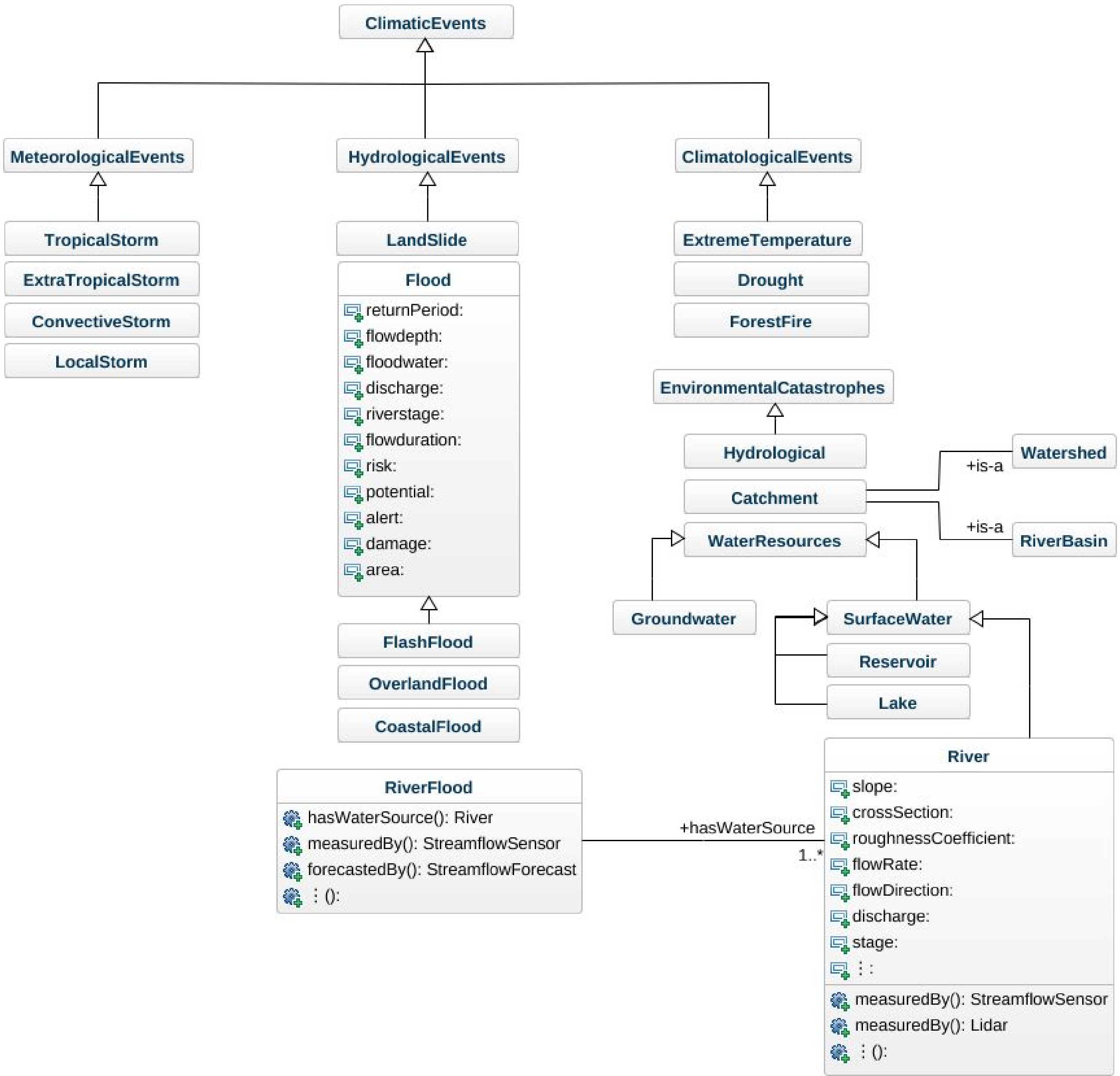


Figure 4.5: Climatic Events and Related Environmental Phenomena Categories of the Flood Ontology with Respect to River Flooding.

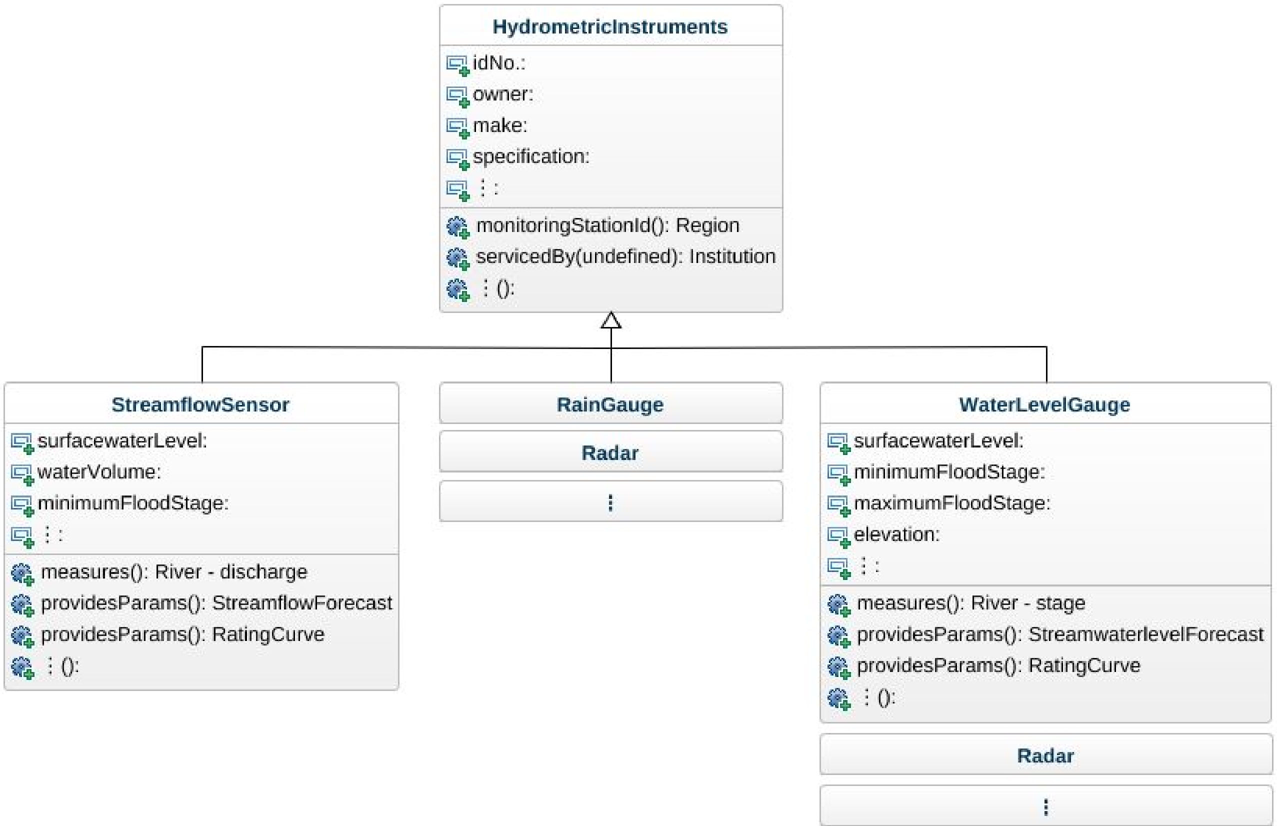


Figure 4.6: UML Diagram Representation of the Hydrometric Instrumentation Category for the Flood Ontology.

4.2 Hydrometric Data Enhancement with Imputation

This section details the data imputation task that was undertaken and necessary for offering a solution to correct and reconstruct the missing data problem that was addressed earlier in Section 4.1.2.4 concerning the Medio catchment Station H3 time series before the data could be fit and used in the modeling phases. Missing data problems can be found across several branches of knowledge such as the sociological, behaviorist, health care, and technological disciplines. For some time now, much research aimed at fixing incomplete and impaired data, specifically those cases that present missing instances, has been an object of intense studies by several scholars and investigators in the field, to find a solution for improving impaired datasets. However, nearly all these methods are strongly dependable on certain knowledge a priori regarding the missingness in the data, which, if one does not know, can render inconvenient which technique to use. Despite being unpopular [349, 350] a group still favored these methods, the reason they might find a space amongst researchers [351, 352].There are three mechanisms of missing data [353]: 1) when most of the data is entirely missing at random in the sample, MCAR; 2) this case appears when several instances in a dataset may be missing randomly, MAR, and 3) if the data does not show clear signs of being of either of the prior mechanism, it is data that is not randomly absent, that is

"MNAR". Data under this category cannot be handled straightforwardly, so it requires very good guesses on the structure of data (Figure 4.7. Multiple imputation algorithms (MI) can handle all three types of missing data situations. However, packages that do MI are usually not designed for MNAR cases, as it is a more complicated mechanism. In this study the dataset pattern suggests the assumption that the data are not MAR or MCAR but are of missingness that depends on the unobserved predictors. Table 4.4 depicts the overall missing pattern of the data record before imputation, in which each row corresponds to a missing data pattern (1 = observed, 0 = missing). It sorts rows and columns in increasing amounts of missing information. The last column and row contain a row and column counts, respectively. Table 4.4 shows that there are 147,086 rows out of 166,756 rows that are complete. There are 95 rows for which only rainfall (RN) data is missing, 8,241 rows are eventually missing for the three variables and there are 5,778 rows for which only RN is known. The variable with the most missing values is for the water level (WL) and together with the streamflow data (Q), they both account for approximately 20% of the data missingness (Figure 4.4), although 87% of the samples are not missing.

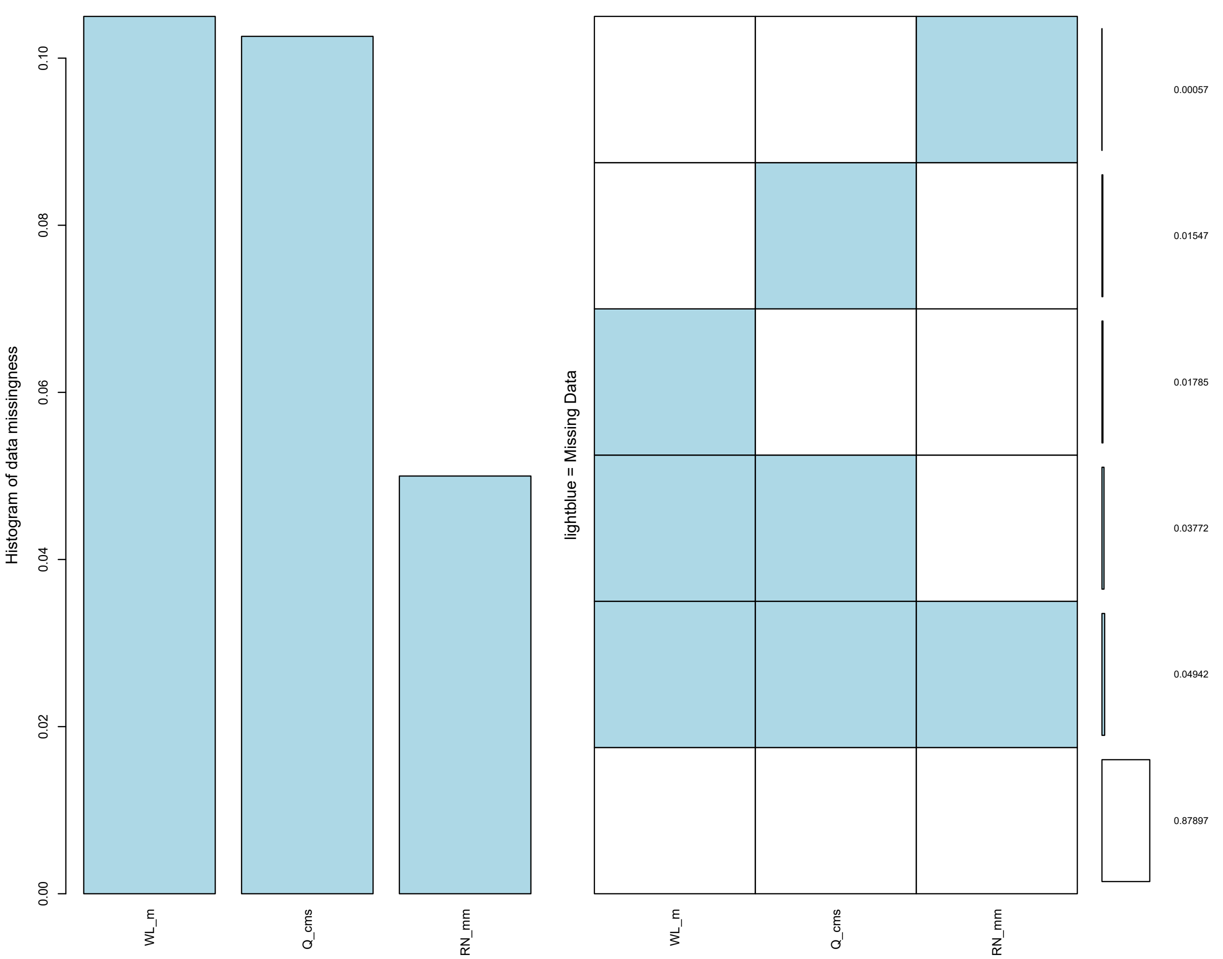


Figure 4.7: Histogram depicting the percentage of missing data at Medio River Station H3.

Table 4.4: Missing data pattern at the H3 hydrometric station.

No. of instances	RN	Q	WL	Missing
147086	1	1	1	0
95	0	1	1	1
2977	1	1	0	1
2579	1	0	1	1
5778	1	0	0	2
8241	0	0	0	3
Total of missing	8336	16598	16996	41930

4.2.1 Data Imputation Results

To enhance the impaired hydrometric record, missing data were imputed with the MICE package [354] from the R statistical computing language [355] within RStudio IDE, version 1.2.5033 [356]. The multiple imputations were done using four of the built-in univariates imputation methods; Predictive mean matching (pmm), Bayesian linear regression (norm), linear regression, non-Bayesian (norm.nob) and Random sample from the observed data (sample). Furthermore, a complete review of these functions can be found in [357]. After the process of data pruning and identification of possible outliers, the hydrometric raw dataset as a matrix table was imported into RStudio, to run MI on the data, to produce the imputations for the RN, WL, and Q values. The imputation setup consisted of selecting the MICE parameters (e.g., number of imputations = 10, number of iterations = 50, and seed = 500) and the four imputation methods mentioned previously for the imputing process, with that said, MICE also allow the selection of the set of predictors one needs to be used in the imputation process for each variable with incomplete instances. Upon termination of the imputation process, the selection of the imputation was done by pooling the results of the imputation methods by fitting a model to each of the imputed datasets and then combining the results per the methodology from the MICE package, by this process the results for pmm method was selected (Figure 4.8). This new imputed dataset is used as inputs to the hydrologic modeling in HEC-HMS and GAMA. The linear fit with the pmm model resulted in an adjusted R^2 of 0.8524 in contrast to that of the sampling method that resulted second during the imputation process ($R^2 = 0.6701$). The summary statistics for the pmm model fit also showed the response variable water level as influenced by the rainfall and streamflow predictors with $P(= 2.2\text{e-}16)$. This explains a significant interaction between the two predictor variables exits at the significance level 0; so, the relationship between the rainfall predictor and the response variable (water level) depends on the streamflow predictor. Therefore, the imputation model for predicting the WL is $WL_{(t)} = 8.23 + 0.0018 * RN_{(t)} + 0.020 * Q_{(t)}$.

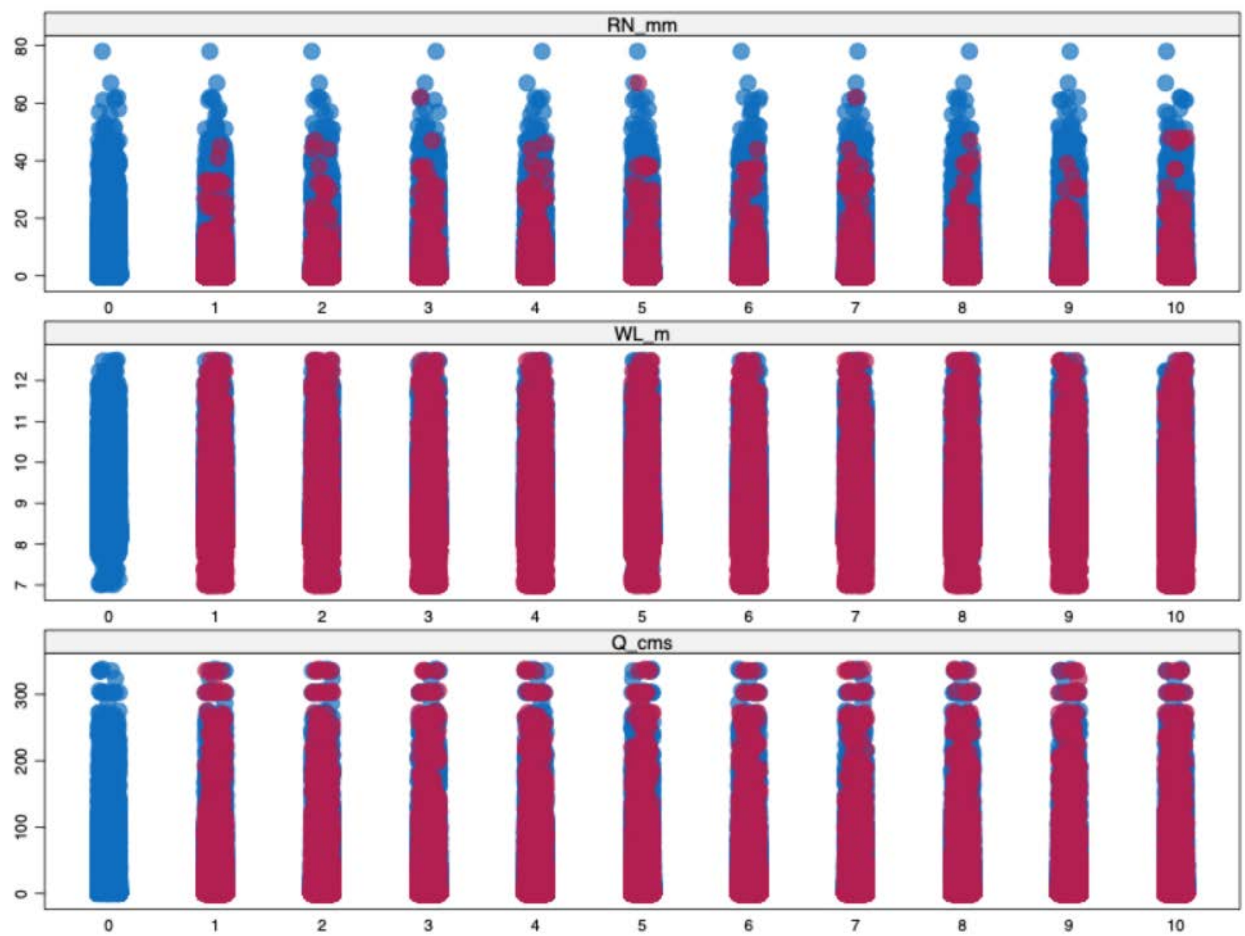


Figure 4.8: Stripplot showing the distribution of the observed versus the imputed datasets with 10 imputations for pmm method for three variables. Observed data in blue, imputed data in red.

4.3 Selection of Storm Episodes

The following section, as the purpose to identify the periods in the hydrometric series that constitutes cases under which storm episodes can be used to model flooding making use of the agent paradigm; therefore, the historical imputed hydrometric time-series data that correspond to a flood event are required (See Figures 4.9, 4.10 and 4.11) for an observation of the extreme flows in the hydrograph and rating curve (river stage) related to the flow rate at station H3. Hence, the focus was on the periods in a given year in which months showed extreme rainfall events and used for calibration and model tweaking purposes.

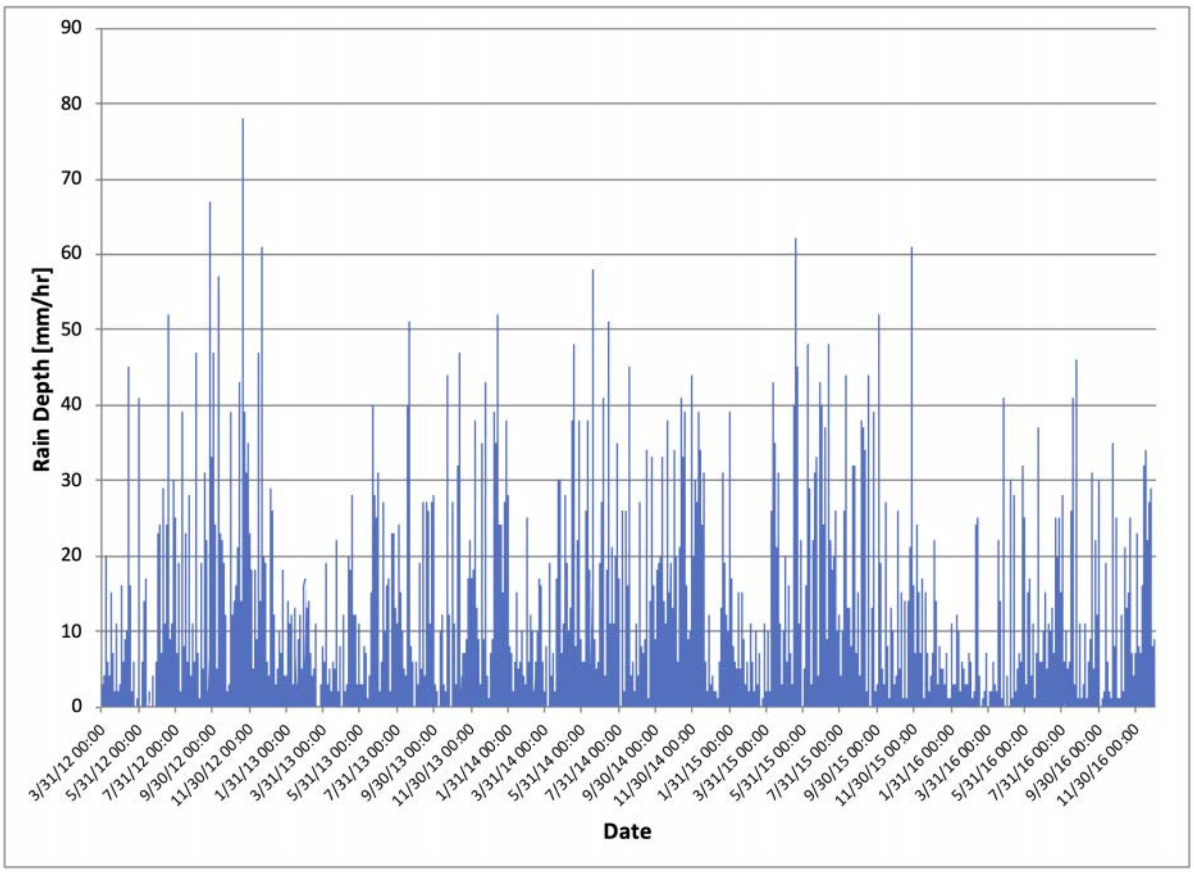


Figure 4.9: Hyetograph with imputed data for Medio River at Station H3 (2012-2016).

As commented above, to select the periods in which a specific rainstorm is useful in model calibration, it is required to detect periods in the series that show an increased rainfall and streamflow trend (isolated peaks corresponding to days of heavy rainfall). This is attainable by graphing of the rainfall and streamflow data over time and identifying outstanding peaks in the graphs, taking into account the idea that most likely a peak in the rainfall data may either correspond to a peak in the streamflow. Though an increased rainfall event is the cause of flooding and consequently the cause of an increase in streamflow rate; the streamflow rate data can be used to identify potential rainstorms, because observed streamflow, is thought to be more consistent reference than rainfall observation and a rainstorm may not essentially

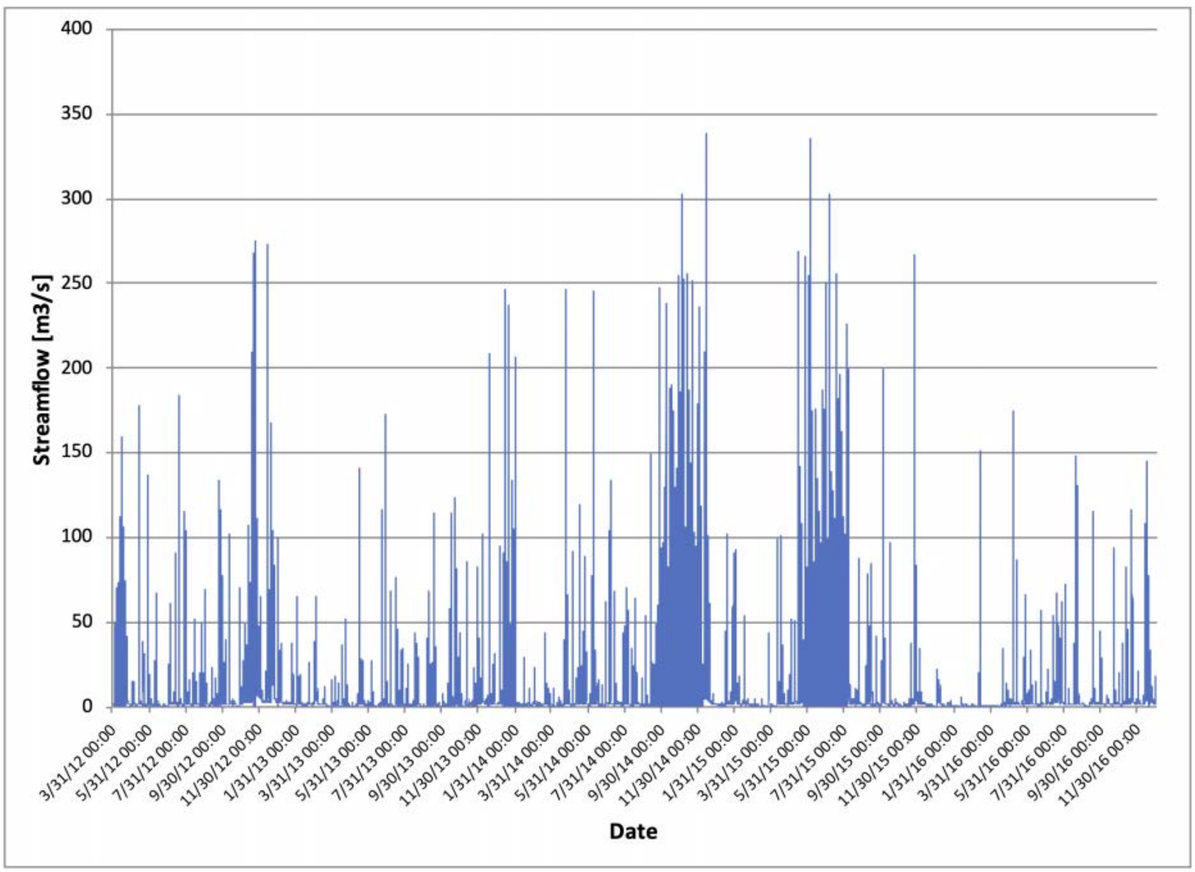


Figure 4.10: Hydrograph with imputed data for Medio River at Station H3 (2012-2016).

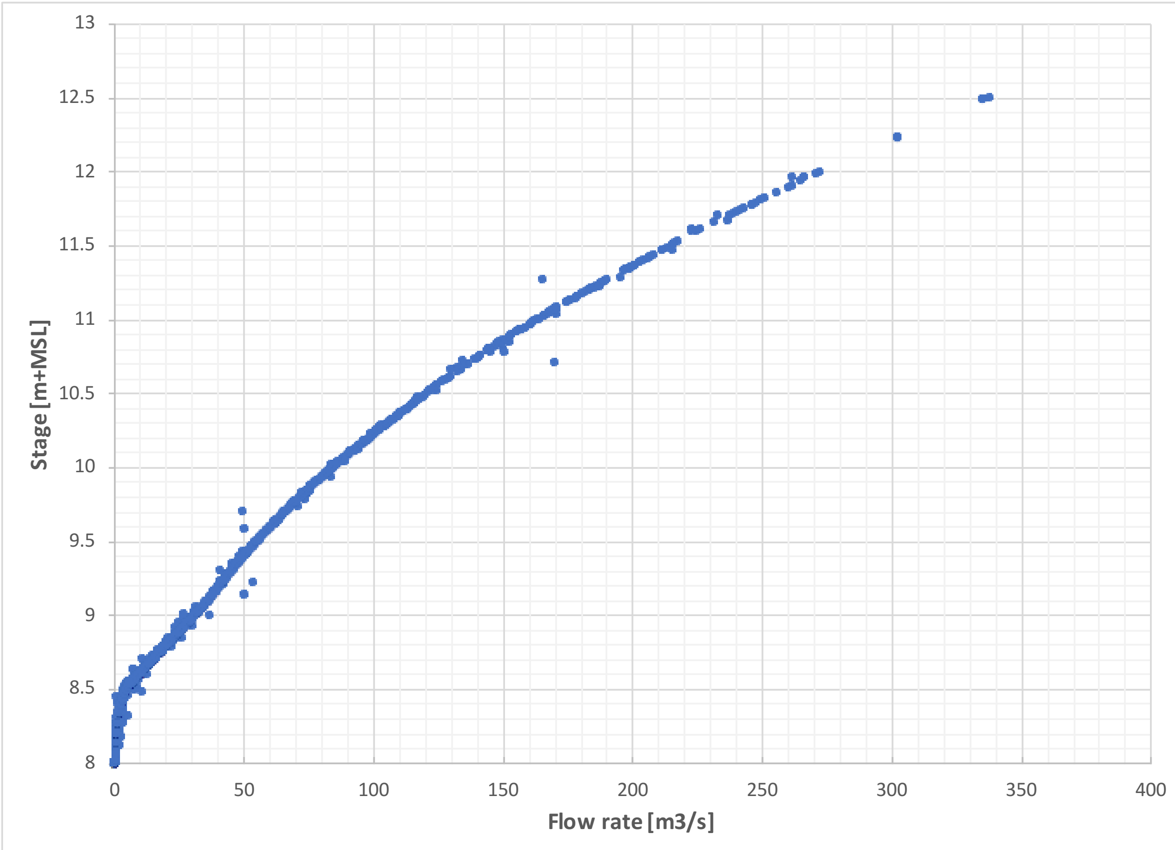


Figure 4.11: Rating Curve with imputed data for Medio River at Station H3 (2012-2016).

account for a flood, therefore, not every rainstorm event is prone to cause a flood. On the contrary, it stands that whenever there is a notable increase in the streamflow of a river, it is probable that an inundation of the river will follow.

While several criteria for categorizing critical rainfall, threshold exists [358–362], the International Meteorological Vocabulary defines heavy rainfall as the incremental precipitation with a depth ratio surpassing the threshold quantity of "7.6 $mm \cdot h^{-1}$" [363]. In the Americas, the Canadian Atmospheric Environment Service describes extreme storms as rain events that exceed a cut-off value >25 $mm \cdot h^{-1}$ [364]. In countries like Indonesia, a tropical country with an annual precipitation regime closely related to that of Panama, the "Badan Meteorologi Klimatologi dan Geofisika (BMKG)" [365] as developed an acute precipitation limit which has been categorized into two classes where class one is the limit for "daily operational definition of precipitation (DODRE)" and class two the "monthly operational definitions of precipitation events (MODRE)" as can be seen in Table 4.5. In Panama, there is no knowledge of such a limit; the reason why it was decided on adapting the Canadian cut-off value to this study. However, when applying the Canadian cut-off limit to the H3 gauging station precipitation values recorded for each month's flooding events, show the exceeding maximum of 25 $mm \cdot h^{-1}$ for a day, as seen listed in Tables 4.6, 4.7, 4.8 and 4.9, respectively. Therefore, for the identification of theoretically extreme rainfall events, it is only considered the constraint to the rainy season.

Table 4.5: Acute Precipitation Cut-off limit. Source: [365].

Precipitation Class	**Precipitation depth**	
	Class I $[mm \cdot h^{-1}]$	**Class II** $[mm \cdot d^{-1}]$
Light precipitation	1 - 5	5 - 20
Average precipitation	5 - 10	20 - 50
Intense precipitation	10 - 20	50 - 100
Very intense precipitation	>20	>100

A common practice in hydrological sciences is to designate discrete rainfall events by periods of rainfall recesses exceeding a designated duration known as the inter-event time (T_i) the gaps that are noticeable in a hyetograph. Nonetheless, in the literature, many criteria for isolating rainstorms applying fixed "Inter-Event" periods have been presented, having quantities ranging between 3 minutes to 24 hours [366, 367]. Based on the autocorrelation function time in the rainfall records, in this study, it have been selected an inter-event time of $T_i = 2$ hours.

To determine the lag time for the entire rainfall series, and the individual months selected separately, it is applied the Pearson correlation coefficient analysis to the time series. An approximation of the autocorrelation functions for the entire seasons at the H3 station is shown with the lag time ranging between 1, 2, 3 and roughly 4 hours as shown in Figure 4.12. This result was similar for the rain data for every individual month analyzed, the reason why it was determined that $T_i = 2$ hours as appropriate lag time in the rainfall series.

Now, recall previously that a rainstorm may not account for a flooding event, and not every rainstorm event is a guarantee for producing a flood; in selecting periods with significant flood events, the identification of peaks in the streamflow data that correspond to a flood needs to follow some other guidelines or specifications besides those already mentioned in the preceding paragraphs. Consequently, the data needs to be trimmed first to select the periods for simulations to the most recent rating curve extracted from the whole series. Then, to decide on which peaks in the streamflow data corresponded to floods, it is considered the following guidelines described beneath:

1. Storm Depth: The storms considered must be of an intensity strong enough to produce flooding. However, since the actual hydro-record is presently scarce, information that would associate specific streamflow records with the time of a consequent flood, may not be straightforward; therefore, it is taken into account that only the observed highest peaks caused floods.

2. Storm Duration:

Table 4.6: Days with hourly rainfall over 25 $mm \cdot h^{-1}$ for Nov and Dec 2012 and Dec 2013.

Period	24 Hr total	Period	24 Hr total	Period	24 Hr total
11/1/12	62	12/1/12	32	12/1/13	48
11/2/12	22	12/2/12	73	12/2/13	120
11/3/12	14	12/3/12	41	12/3/13	0
11/4/12	108	12/4/12	44	12/4/13	28
11/5/12	87	12/5/12	22	12/5/13	117
11/6/12	103	12/6/12	6	12/6/13	325
11/7/12	100	12/7/12	16	12/7/13	60
11/8/12	74	12/8/12	0	12/8/13	78
11/9/12	78	12/9/12	60	12/9/13	6
11/10/12	134	12/10/12	0	12/10/13	0
11/11/12	212	12/11/12	26	12/11/13	24
11/12/12	6	12/12/12	24	12/12/13	40
11/13/12	204	12/13/12	104	12/13/13	56
11/14/12	38	12/14/12	618	12/14/13	10
11/15/12	80	12/15/12	202	12/15/13	2
11/16/12	19	12/16/12	14	12/16/13	10
11/17/12	177	12/17/12	19	12/17/13	130
11/18/12	624	12/18/12	39	12/18/13	512
11/19/12	243	12/19/12	56	12/19/13	42
11/20/12	707	12/20/12	230	12/20/13	68
11/21/12	254	12/21/12	602	12/21/13	22
11/22/12	218	12/22/12	118	12/22/13	42
11/23/12	170	12/23/12	84	12/23/13	10
11/24/12	522	12/24/12	200	12/24/13	130
11/25/12	669	12/25/12	90	12/25/13	78
11/26/12	161	12/26/12	20	12/26/13	18
11/27/12	4	12/27/12	18	12/27/13	2
11/28/12	106	12/28/12	16	12/28/13	6
11/29/12	52	12/29/12	4	12/29/13	4
11/30/12	212	12/30/12	20	12/30/13	4
12/1/12		12/31/12	14	12/31/13	36

- Rainstorm must be apparent from the hydrometric dataset to reproduce it as a storm event.
- Streamflow conditions of preceding and antecedent flows must be low enough and be attributable to that of base flow for these periods and in this respect discarding antecedent soil moisture.

Table 4.7: Days with hourly rainfall over 25 $mm \cdot h^{-1}$ for Sep and Dec 2014.

Period	24 Hr total	Period	24 Hr total
9/1/14	0	12/1/14	22
9/2/14	8	12/2/14	142
9/3/14	4	12/3/14	124
9/4/14	110	12/4/14	182
9/5/14	0	12/5/14	188
9/6/14	40	12/6/14	16
9/7/14	10	12/7/14	108
9/8/14	0	12/8/14	58
9/9/14	0	12/9/14	92
9/10/14	0	12/10/14	168
9/11/14	48	12/11/14	310
9/12/14	2	12/12/14	84
9/13/14	54	12/13/14	395
9/14/14	208	12/14/14	269
9/15/14	0	12/15/14	24
9/16/14	70	12/16/14	15
9/17/14	2	12/17/14	290
9/18/14	2	12/18/14	19
9/19/14	0	12/19/14	128
9/20/14	34	12/20/14	76
9/21/14	0	12/21/14	24
9/22/14	32	12/22/14	10
9/23/14	50	12/23/14	0
9/24/14	70	12/24/14	0
9/25/14	4	12/25/14	8
9/26/14	2	12/26/14	80
9/27/14	38	12/27/14	26
9/28/14	0	12/28/14	8
9/29/14	0	12/29/14	8
9/30/14	0	12/30/14	4
10/1/14		12/31/14	8

- Validation of flood days extent with the assistance of the surrounding citizens and rainfall records.

3. Storm Period: The storms should as per the available data and be able to be traced back to the computed rating curve shown in Figure 4.11. This guarantees accuracy in modeling scheme and that it closely replicates the current hydraulics of the river (i.e.,

Table 4.8: Days with hourly rainfall over 25 $mm \cdot h^{-1}$ for May and Nov 2015.

Period	24 Hr total	Period	24 Hr total
5/1/15	34	11/1/15	2
5/2/15	92	11/2/15	0
5/3/15	100	11/3/15	58
5/4/15	0	11/4/15	2
5/5/15	10	11/5/15	2
5/6/15	16	11/6/15	16
5/7/15	10	11/7/15	0
5/8/15	16	11/8/15	0
5/9/15	100	11/9/15	74
5/10/15	6	11/10/15	20
5/11/15	52	11/11/15	0
5/12/15	12	11/12/15	0
5/13/15	0	11/13/15	2
5/14/15	5	11/14/15	16
5/15/15	291	11/15/15	56
5/16/15	372	11/16/15	68
5/17/15	10	11/17/15	0
5/18/15	31	11/18/15	2
5/19/15	225	11/19/15	0
5/20/15	256	11/20/15	34
5/21/15	198	11/21/15	32
5/22/15	2	11/22/15	2
5/23/15	236	11/23/15	58
5/24/15	86	11/24/15	82
5/25/15	2	11/25/15	206
5/26/15	14	11/26/15	742
5/27/15	172	11/27/15	72
5/28/15	4	11/28/15	104
5/29/15	22	11/29/15	102
5/30/15	0	11/30/15	40
5/31/15	0	12/1/15	

channel geometry).

4. Data Accuracy: All data in the hydrometric database are electronically recorded, as such, there may be errors in the values in the form of outliers, intermittent recordings, data gaps, and missing values. Therefore, each instance of the hydrograph should correspond to the dynamics of undergoing physical processes.

Table 4.9: Days with hourly rainfall over 25 $mm \cdot h^{-1}$ for May and Dec 2016.

Period	24 Hr total	Period	24 Hr total
5/1/16	8	12/1/16	82
5/2/16	0	12/2/16	50
5/3/16	4	12/3/16	0
5/4/16	0	12/4/16	18
5/5/16	0	12/5/16	0
5/6/16	0	12/6/16	2
5/7/16	320	12/7/16	18
5/8/16	0	12/8/16	54
5/9/16	0	12/9/16	24
5/10/16	0	12/10/16	82
5/11/16	2	12/11/16	2
5/12/16	64	12/12/16	0
5/13/16	124	12/13/16	2
5/14/16	178	12/14/16	171
5/15/16	2	12/15/16	99
5/16/16	2	12/16/16	265
5/17/16	6	12/17/16	6
5/18/16	12	12/18/16	62
5/19/16	2	12/19/16	4
5/20/16	14	12/20/16	46
5/21/16	58	12/21/16	48
5/22/16	4	12/22/16	10
5/23/16	2	12/23/16	106
5/24/16	28	12/24/16	10
5/25/16	18	12/25/16	12
5/26/16	84	12/26/16	112
5/27/16	2	12/27/16	28
5/28/16	68	12/28/16	38
5/29/16	90	12/29/16	14
5/30/16	0	12/30/16	92
5/31/16	34	12/31/16	82

From the information shown in Figure 4.9, though it is possible to observe the peaks, to select them is not an easy task (Guideline 1). This is the reason why it was important to trim the data and plot the separated hydrographs and select the various peaks that constituted extreme storm events (Guidelines 1 and 2 (points 1 and 2)). From the information displayed in the rainfall and streamflow data, fourteen initial storms were selected; both the dates and intensity are shown in Figures 4.13, 4.14, and 4.15, respectively. The storms are assessed

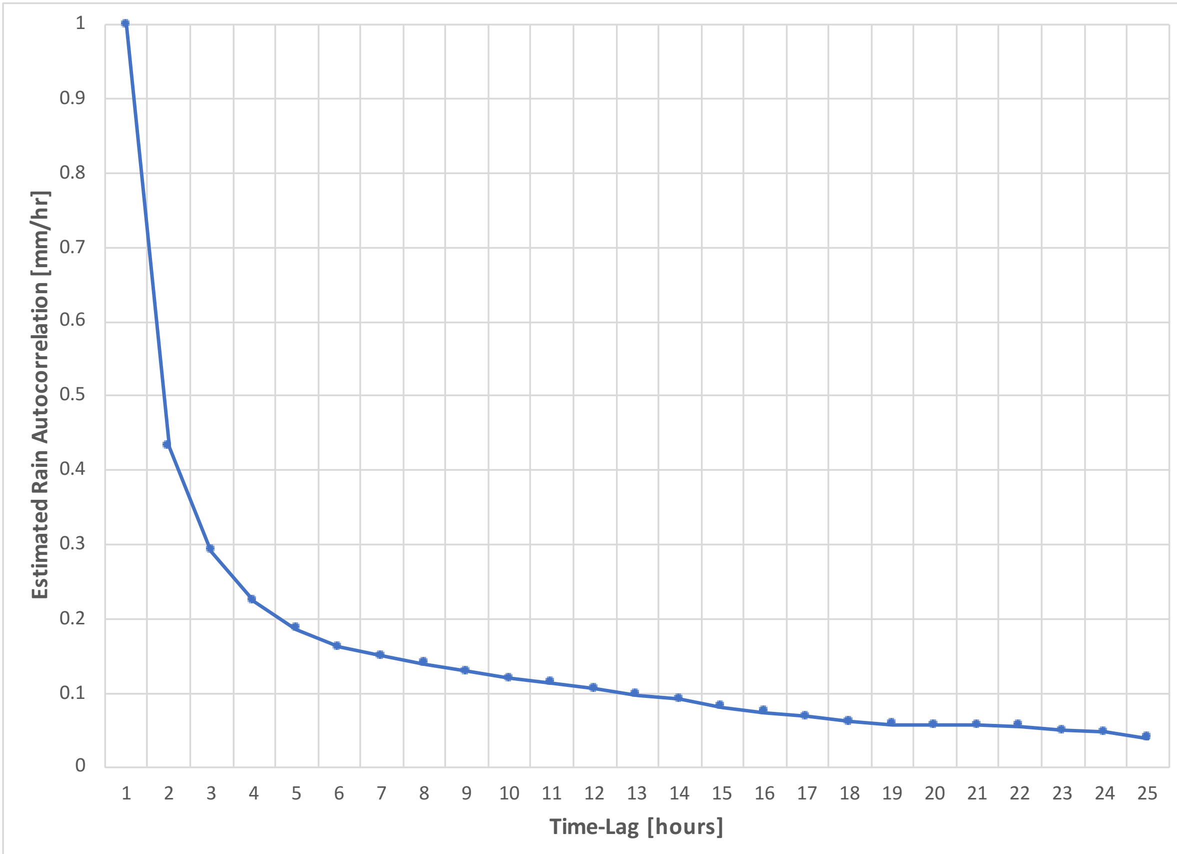


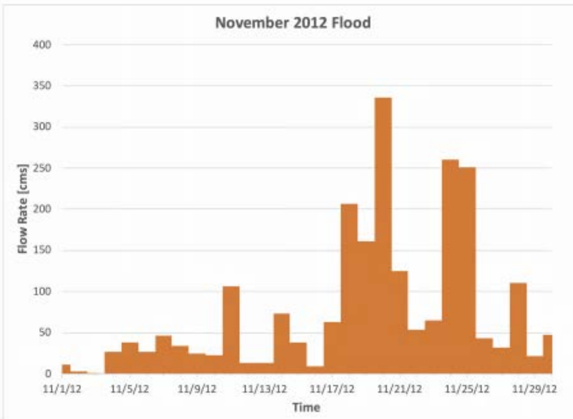
Figure 4.12: Autocorrelogram for the H3 Hydrometric Station with lags of an hour.

individually to decide on their relevance.

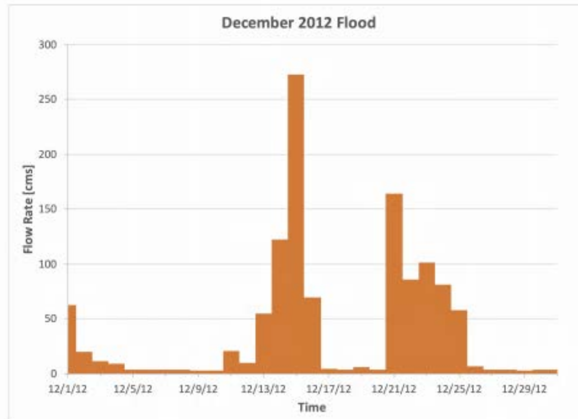
The individual hydrographs depicted here at this resolution elucidates a clearer view of the rainstorms and flood peaks, given us some better criteria to decide on which of these peaks in the streamflow corresponds to a flood and making it able to consider which of these to use in the hydrologic modeling framework. Nevertheless, as can be seen in the graphs, making a selection can be an overwhelming task since another important aspect that can also be observed is the localized sudden floods, which are characteristically due to heavy rainfall (i.e., flash flood).

Following this methodology, the rainstorms of December 2013, September 2014, May 2016 and December 2016 were disregarded since they do not conform to guidelines 1 and 2 presented. The peak streamflow rate of each of these rainstorms, though they seem high in the graphs, did not present the necessary flow duration to guarantee a flood occurred (Guideline 2). Additionally, the rainstorms exhibit a rapid drop in streamflow rate during the peak, which is likely due to flash floods.

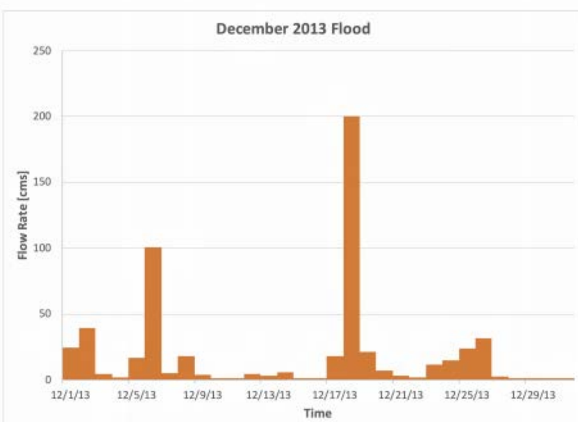
On the other hand, the November and December 2012, December 2014, May, and November 2015 storms were compared to their corresponding rainfall and streamflow data (Guidelines 1, 2, and 4, respectively).



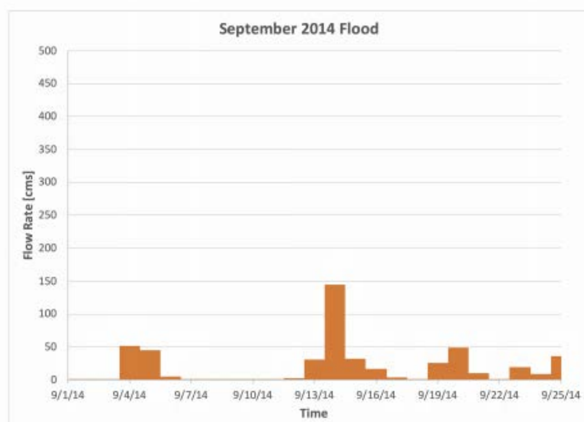
(a)



(b)

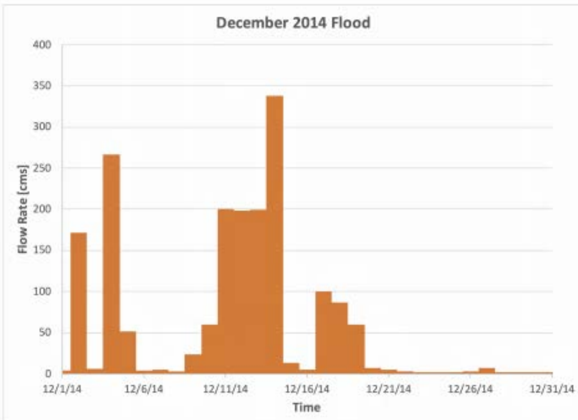


(c)

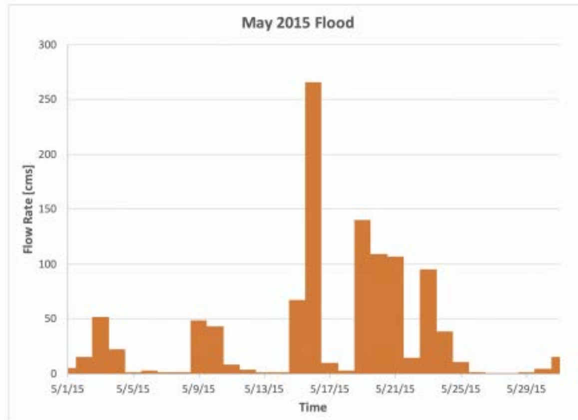


(d)

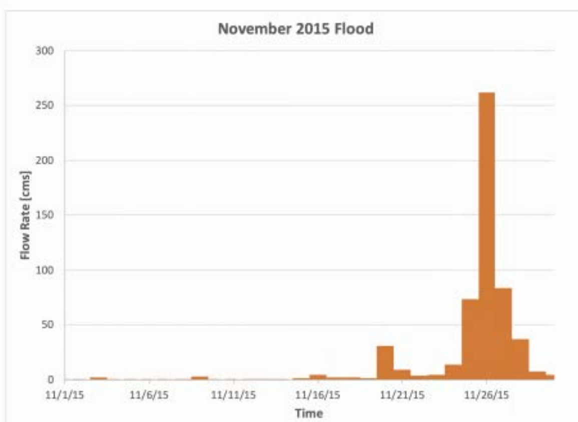
Figure 4.13: Hydrographs at Medio River Sta. H3 with Selected Storm Events: (a) Nov 2012, (b) Dec 2012, (c) Dec 2013 and (d) Sep 2014.



(a)

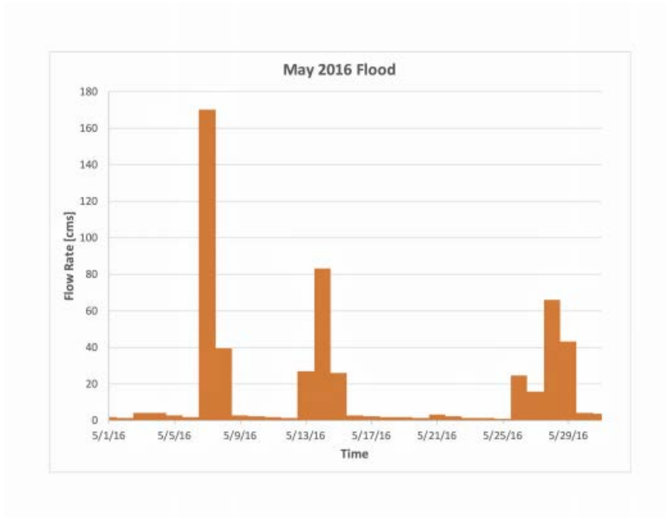


(b)

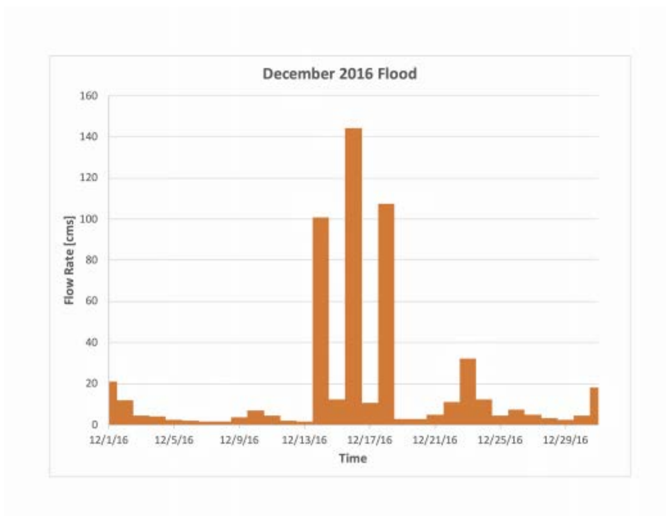


(c)

Figure 4.14: Hydrographs at Medio River Sta. H3 with Selected Storm Events: (a) Dec 2014, (b) May 2015, and (c) Nov 2015.



(a)



(b)

Figure 4.15: Hydrographs at Medio River Sta. H3 with Selected Storm Events: (a) May 2016 and (b) Dec 2016.

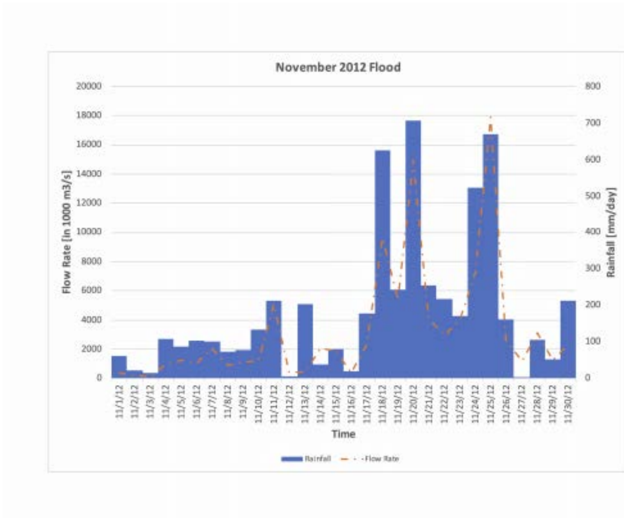
Figures 4.16, and 4.17 show day-to-day relationship between the rainfall and streamflow data of each month throughout a year with a given rainstorm event. It can be also observed the changing patterns that appear in the graphs, where some trends are predictable, while others are quite perplexing. This shows the complex relationship between some of the basic dynamics of the water cycle. Consequently, a closer look at several sectors of the complete charts reveals that the Medio river response to rainfall is highly variable. Continuing with the analysis of Figure 4.16 the rainstorm event that initiated on November 1st shows the precipitation to gradually increase from 26 mm to 94 mm, and later decreasing to approximately 10 mm on day 14th. This precipitation was accompanied by a moderate increase in river discharge that reflected an oscillating recording between 2.56 and 105.70 $m^3 \cdot s^{-1}$, on day 11^{th}. Notwithstanding, for this same period, which was practically a wet month, its maximum hourly precipitation registered at 170 mm on days 18^{th} through 22^{nd} and a recorded stream discharge of 227.38 $m^3 \cdot s^{-1}$; while on the other hand, of particularly notable is the smaller hourly precipitation event of days 24^{th} through 26^{th} which is accompanied by a larger increase in river discharge of 258.63 $m^3 \cdot s^{-1}$. In a similar manner to the November 2012 rainstorm, two observed peak river discharges or characterized by the December 2012 hourly rainstorms and the resulted streamflow profiles, which may probably indicate that the data for that period is in agreement with observed precipitation data. The increase in flow is in agreement with the increase in precipitation, and as was the case with the November 2012 storms, the same patterns are observable where a lighter rain shows a higher stream discharge than a larger storm event. On December 2014, three storm events were visible, being the events of the days 7^{th} through 12^{th} and 13^{th} through 15^{th} the most significant. In this period, it is noticed an increase in streamflow for both rainstorm events, and as was shown in the earlier periods, there is a similar discrepancy with the rainfall that occurs on days 11^{th} through 14^{th} that totaled approximately an hourly registered rainfall on day 14^{th} of 1498 mm but is accompanied by a larger streamflow value of 335.53 $m^3 \cdot s^{-1}$, but there is still visible a much smaller increase in river discharge of 206.17 $m^3 \cdot s^{-1}$ on day 11^{th} with 106 mm rainfall.

In the May 2015 storm, there are two observable peaks in both rainfall and streamflow. The precipitation and the increased streamflow correspond, as in the other previous storms, and a similarly observed higher river discharge with lower rainfall depth. Finally, in the November 2015 storm, that extended from days 23^{rd} through 30^{th}, on day 26^{th} the maximum hourly rain of 196 mm was recorded and the largest streamflow of approximately 264.28 $m^3 \cdot s^{-1}$.

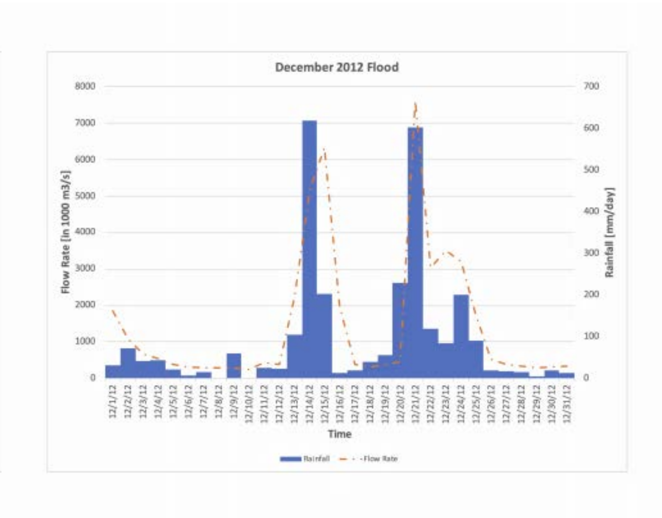
In the previous paragraphs, it was observed from each rainstorm so far, the tendency in which a small and prolonged rainfall event is capable of conveying with it a considerable increase in the discharge rates. This aspect gives us a clear indication of the precipitation-streamflow relationship in the Medio river catchment and how the stream responds to these precipitation magnitudes. The fact that the Medio river catchment is localized in an area of high precipitation regime in the tropics and in which the hourly precipitation on a given day can at least exceed 25 $mm \cdot h^{-1}$ are just two of the dynamics that combined leads to the streamflow-precipitation relationship a complex one. However, these factors should not be emphasized, have their own set of factors, as some natural and anthropogenic related disturbances have directly correlated that influence the streamflow-precipitation relationship.

With the aforesaid, many causes affect the way in which a river course responds to a rainstorm event. Some of these causes will vary temporarily and spatially within one catchment. For example:

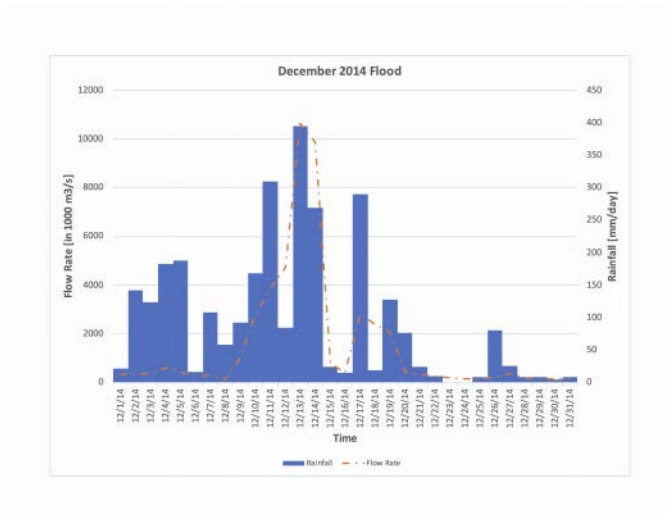
- rainstorm intensity
- rainstorm duration
- air temperature
- wind speed



(a)



(b)

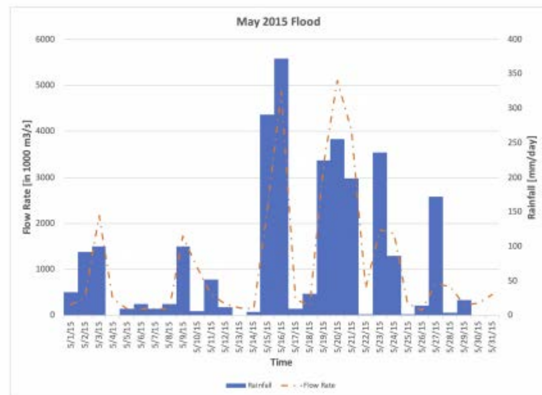


(c)

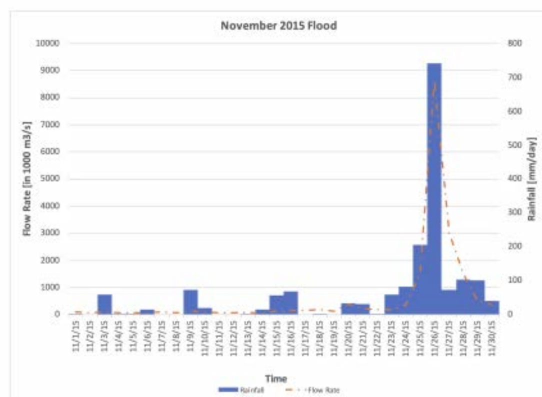
Figure 4.16: Rainfall-Streamflow Relationship for Potential Storm at Medio River Sta. H3. (a) Nov 2012, (b) Dec 2012, and (c) Dec 2014.

Nevertheless, there are other aspects to consider, and these may not probably not vary from season to season, but which can influence the precipitation-streamflow relationship. These factors will also vary within one catchment to another. These comprise:

- soil type
- land use
- basing slope
- river slope



(a)



(b)

Figure 4.17: Rainfall-Streamflow Relationship for Potential Storm at Medio River Sta. H3. (a) May 2015 and (b) Nov 2015.

- channel geometry
- geographic and climatic differences
- vegetation cover and percentage of impervious surfaces

In summary, for a rainstorm to be practical in the modeling implementation, several adjustments, and calibration tasks need to be carried out to the recorded measured rainfall dataset, therefore, for this process to be effective, the information must be of quality and in correspondence with the reproduced hydrograph.

The nature and profiles of the large magnitude of streamflow during November and December 2012 rainstorms, December 2014, the May and November 2015 rainstorms are indicative that a flood likely occurred in the Medio stream. The available hydrometric records and rating curve also indicate that a flood occurred in these periods, and with it affecting the population in the catchment. Therefore, the streamflow-rainfall analysis of the actual hydrometric data provides the insight that the prominent rainfall depth and streamflow rate did cause a flooding event.

4.4 Hydrologic Simulation Environment Setup

HEC-HMS (Hydrologic Modeling System) is the standard hydrological model used for modeling the precipitation-streamflow process for this research as mentioned in Section 2.2.2. The chapter does not pretend to be a treatise on the models, and more information on the model can be consulted [86]. However, it should be kept in mind that GAMA is the software of choice that is used for this flood simulation task. Recalling on previous chapters, HEC-HMS falls under the category of a numerical, physics-based model, as it uses the equations of flow hydrodynamics, rainfall and evapotranspiration schemes and physical parameters of the river basin to model the rainfall and its yielded flow in river networks within a given catchment [86]. As such, HEC-HMS can perform simulations scenarios such as lumped or semi-lumped and distributed paradigms. HEC-HMS as a free Open Source model is the hydrologic software of choice used by many institutions, universities, researchers, and engineering professionals for researching flow forecasting, reservoir, and spillway design, flood control, and regulation, flood risk assessment, among others [86].

To build a model in HEC-HMS, users are required to provide a feed of information that needs to be given to several of its model compartments, called the "Basin Model", the "Meteorological Model", "Control Specifications" and last, the "Data Manipulation" suit component that is managed by the "DSSVue" [368].

Besides the use of standard hydrologic models, it is analyzed in this research the use of agent-based systems to realize flood forecasting as an alternative to the non-conventional hydrologic simulation approach. Specifically, it allows the building of dynamically driven agent environment prototypes that can adapt a representation of the modeled systems. From the available choice of the types of modeling environments containing agents, the GAMA agent-based modeling platform is the perfect choice because it can generally add support to the data-driven modeling paradigm, allows the creation of agents from different types of databases and products, comprising geospatial databases, social, environmental database, and it permits the implementation of simulation of well-built large-scale models (i.e., millions of agents with high degree of complexity). Nevertheless, for a thorough understanding of the potentials of the platform, comprehensive details of its various features are provided by Patrick and Drougoul [369]. The GAMA architecture is developed on a high-level agent-oriented language (GAML), which offers a clear and straightforward syntax that renders it easy for the non-programmer to build models. As GAMA is an open source software, it allows modelers to develop new facets and functionalities in the form of plugins, using Java, and in this way, users can meet their necessities for their modeling projects. Given these characteristics mentioned earlier makes GAMA suitable for implementing hydrologic simulations and forecasting of floods. Therefore, in this chapter, it is presented the approach of the hydrologic modeling of several rainstorms registered in the Medio river catchment for different periods, overall model results, and analysis of the rainfall-runoff simulations over the catchment resulted with HEC-HMS and GAMA platform.

4.4.1 HEC-HMS: Simulation Setup

The following section summarizes a brief hydrologic investigation of the Medio River Subcatchment and how it responded to each of the rainstorm occurrences recorded at the upstream H3 Hydrometric station using HEC-HMS hydrologic system version 4.3. The Analysis included a hydrologic study related to rainfall events that could be considered significant to produce enough water to cause floods and for conveying solid flows, most of which flows into the Río Caimito catchment.

The Medio catchment morphologic profile is characterized by soils with high clay content and is classified under the hydrologic group of soils according to the water absorption capacity in "C/D". The soils under group C, have a granulometric (texture) composition is sandy clay loam is characterized by small soil penetration grade when fully saturated and is mainly composed of soils with stratigraphic layers resistant to water from moving down into the soil. The soils of group D are composed of a heterogeneous mixture of clay [314] the surface is covered by grasses, shrubs, grasslands, and lush secondary forest, just as was discussed under Section 4.1.2. Almost 50% of the catchment is a mountainous region.

To perform rainfall-runoff simulations in HEC-HMS, it is required first to configure the Basin Submodel, since here it is determined the methodology to be used regarding basin losses. In this catchment submodel, it is adopted the SCS curve number method (SCS CN) for the losses. The SCS method operates by making increments to the assigned curve number in the catchment or each sub-catchment. The curve numbers assigned to every subcatchment are taken from soil hydrologic groups tables, as previously shown. Afterwards, the losses that are obtained initially, are calculated from the assigned curve numbers. Next, to carry out the rainfall-runoff transformation, it is selected the transformation method based on the SCS unit hydrograph (SCS UH), which requires certain initial values such as the basin lag time. As the base flow is not considered, it is not necessary to make use of the recession component. Once these and some other parameters have been established, model runs are carried out, and retrieved knowledge about the analysis of the hyetographs, hydrographs, the rating curves from different periods in the series, five storm events have been observed and selected. One of the periods (November 2012) used for calibration and the other selected storm periods of concern for analysis, as specified in Section 4.3.

In HEC-HMS the catchment is depicted as a "dendritic" system (i.e., an interconnected network of streams) with various hydrographic components (Figure4.18). These components are combined to model real-world catchment dynamics.

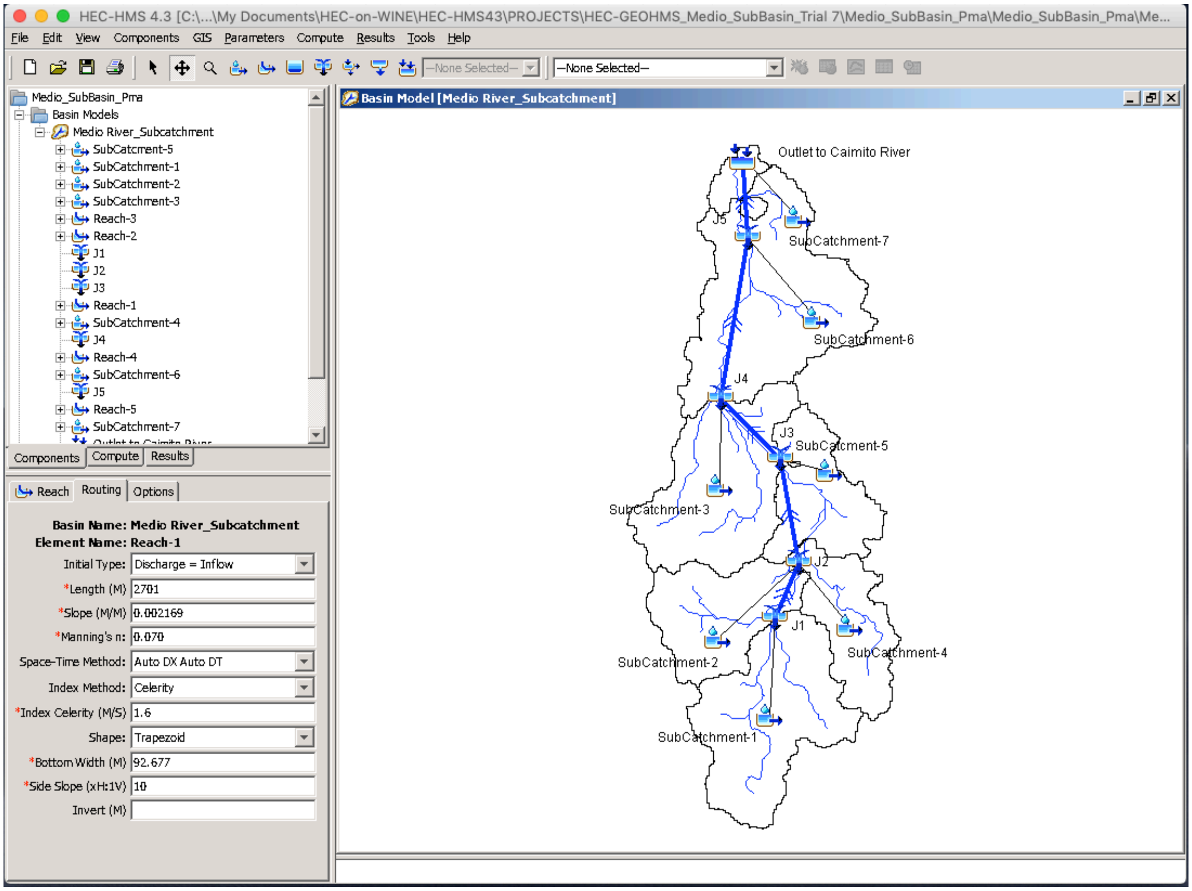


Figure 4.18: Schematic View of the HEC-HMS Graphic User Interface for the Medio River Subcatchment Model.

4.4.1.1 Required Geospatial Information and Data Features

- Digital Elevation Model (DEM):

  HEC-HMS requires geospatial information as a part of is modeling requirements. The DEM file that is used in the modeling framework was a fundamental GIS data product for the Medio river catchment. DEM files are important assets in hydrological modeling since they provide information on elevation, that is useful for defining flow direction patterns, and for computing the runoff. An "ASTER GDEM" (Global Digital Elevation Model) V003 in Geo TIFF format was downloaded at [370] and processed in QGIS to extract the area of interest (AOI).

- River Basin and River Shapefiles:

  The catchment geometry, physiographic features, and the river network model parameter values were extracted from each shapefile attribute tables and the attributes of the raw DEM, which was denoised earlier with the Geospatial Analysis Tools, "WhiteboxGAT" version 3.4 [371]. Other features and maps created for importation into HEC-HMS and GAMA were created in QGIS version 3.8 [372]. Other required GIS inputs for the catchment (i.e., river network and catchment delineation) were created with QGIS via available plug-ins (i.e., RiverGIS and Profiling Tool) and running the respective basing delineation task, from which was extracted 7 hydrological subcatchments. In Table 4.10 below is shown some of the Medio river HEC-HMS catchment basin model parameters.

Table 4.10: HEC-HMS Catchment Model Parameters for Medio River.

Model	**Technique**	**Variable**
Loss rate	SCS CN	I_a [mm], CN, imperviousness [%]
Excess transform	SCS UH	Basin lag [min]
Flow routing	Muskingum-Cunge	Specified discharge $[m^3 \cdot s^{-1}]$

- Hydrometric data:

  The meteorological and hydrometric data defined under Section 4.2.1 represent the necessary metric information to be used as input data to the HEC-HMS hydrologic modeling system. These are the instantaneous data obtained from rain, stage, and water flow (discharge) sensors, and monthly evaporation estimates located at station H3 in the Medio catchment.

4.4.1.2 HEC-HMS: Input Parameters

- Roughness coefficient:

  The roughness coefficient values of the river, commonly called Manning's n, were calculated from tables, following the guidelines for natural channels given in [373]. This data is also available online.

- Loss method:

  From all the precipitation that may fall over a catchment in a given time, not all ends up in a stream, but a portion of it represents losses accounted from the processes of infiltration, overland flow, and other processes interacting at the surface level. HEC-HMS provides some twelve different loss methods, from which is adopted the SCS Curve Number (SCS CN) method. This method requires the following initial parameters:

  1. This variable describes the volume of rain that must precipitate previously to surface excess results. This value is calculated from the following equations (4.1) and (4.2).

$$I_a = 0.2S \tag{4.1}$$

$$S = \left(\frac{25400}{CN}\right) - 254 \tag{4.2}$$

where I_a is the initial abstraction, this is a 20% loss of the direct runoff and S is maximum abstraction, and CN the runoff curve number.

2. Catchment CN values for each subcatchment were calculated from curve number tables [374] taking into account the soil type and land use activities.
3. The percentage of impermeable surface for each subcatchment was left at its default value (0.0%) because the catchment is characterized by different land uses as was defined previously and estimation of this value would probably render improper.

- Rainfall Excess Transform method:

  Estimation of lag periods ($\Delta\tau$) for the subcatchments were calculated using equation (4.4), but first, it was necessary to find the time of concentration (T_c) as per defined by equation (4.3) which is given in hour and then converted to minutes to find the lag time (T_L).

$$T_c = 0.161\left(\frac{L}{S_c^{0.5}}\right)^{0.64} \tag{4.3}$$

$$T_L = 0.6T_c \tag{4.4}$$

then, in equation (4.3) L denotes the length of main stream (km) and (S_c) refers to the sine of channel slope angle (dimensionless). Equation (4.3) is based on Chow [375] and equation (4.4) from the Bransby-Williams method, based on the SCS. Table 4.11 below shows the estimated SCS variables for the Loss and Transform criterion.

Table 4.11: Estimated SCS Loss and Transform coefficients criterion.

SubCatchment ID	CN	S [mm]	Ia [mm]	Lag [min]
1	89	31.4	6.3	171.0
2	83	52.0	10.4	90.0
3	83	52.0	10.4	145.0
4	83	52.0	10.4	99.0
5	93	19.1	3.8	90.0
6	83	52.1	10.4	100.0
7	83	52.1	10.4	90.0

- Routing criteria for a river reach:

  Infiltration rates, rainfall excess transformation to runoff, and flow routing are computed by methods such as the "Green-Ampt", "SCS Unit hydrograph", and the "Muskingum-Cunge" among others presented in the technical reference manual [376]. From the derived information in the DEM (e.g., river length, river slope, basin slope, and ground elevations) and knowledge of the catchment physiographic characteristics, the selected routing method of choice was the Muskingum-Cunge [376]. To apply this methodology, information on the length of the river reach (m), gradient (m/m), roughness coefficient, bottom width (m), and side gradient needs to be known. The Muskingum-Cunge formulation is implemented in two ways as referred to in the Technical Manual: i) using a standard configuration, in which it is viewed the channel as a trapezoidal cross-section. This is the simplest representation provided, given the channel roughness, slope gradient, and reach are specified. Channel reach and coarseness can be extracted from cartographic inputs, remote sensing, on-site measurements, or even DEM file products. Energy slope can be abstracted from channel slope bed and so on and ii) by an 8-point cross-section configuration, this is the case if the geometry of the channel does not illustrate the optimum cross-sectional configurations, then it is recommended to adapt the configuration as defined in the Technical Manual. An outlook of the selected Muskingum-Cunge method parameters for each channel's reach is in Table 4.12 below:

Table 4.12: Muskingum-Cunge Channel Reach Hydraulic Routing Parameters.

	R. 1	**R. 2**	**R. 3**	**R. 4**	**R. 5**
R. Length [m]	2,701.30	2,861.67	2,632.34	1,760.53	1,152.16
R. Slope [m/m]	0.002169	0.00169	0.001436	0.000972	0.000000
Manning's [n]	0.070	0.063	0.042	0.050	0.050
Invert [m]	-	-	-	-	-
R. Shape	Trapezoid	Trapezoid	Trapezoid	Trapezoid	Trapezoid
R. Bottom Width [m]	15	30	7	5	20
R. Side Slope [xH:1V]	10	10	10	10	10

4.4.2 HEC-HMS: Simulation, Calibration, and Selected Flood Cases

Results of the simulations, calibration task and the scenario run setup for the selected storm periods and hydrometric parameters defined under Section 4.4.1 are presented in this section. Since several of these parameters are estimates of observed field data and GIS calculated basin physical products, the simulated results may not likely be in close agreement with actual observed values. Therefore, this observation requires the process of parameter tweaking (e.g., model calibration, exploration, and sensitivity analysis) for correcting and lowering the bias in the estimated parameters applied in the simulation process of the observed hydrograph. Thus, in identifying the new parameter estimates, the modeling process is then re-run to study the selected November 2012 rainstorm event, in an attempt to understand the flood hydrograph of the Medio river sub-catchment and to provide an instrument which facilitates in the assessment of flood forecasting and warning to decision-makers and stakeholders at the tropical river basin scale.

4.4.2.1 HEC-HMS: Hydrological Modeling

Terminated the simulation setup for the HEC-HMS model, with all corresponding hydrometric data and GIS features a simulation is run applying the framework as defined in Section 4.4.1. Then, to explore the model's ability to replicate the hydrograph, the simulated hydrograph of the November 2012 rainstorm is compared to the historically observed hydrograph at the Medio river H3 Hydrometric station. A graphical illustration of both hydrographs is presented below in Figure 4.19.

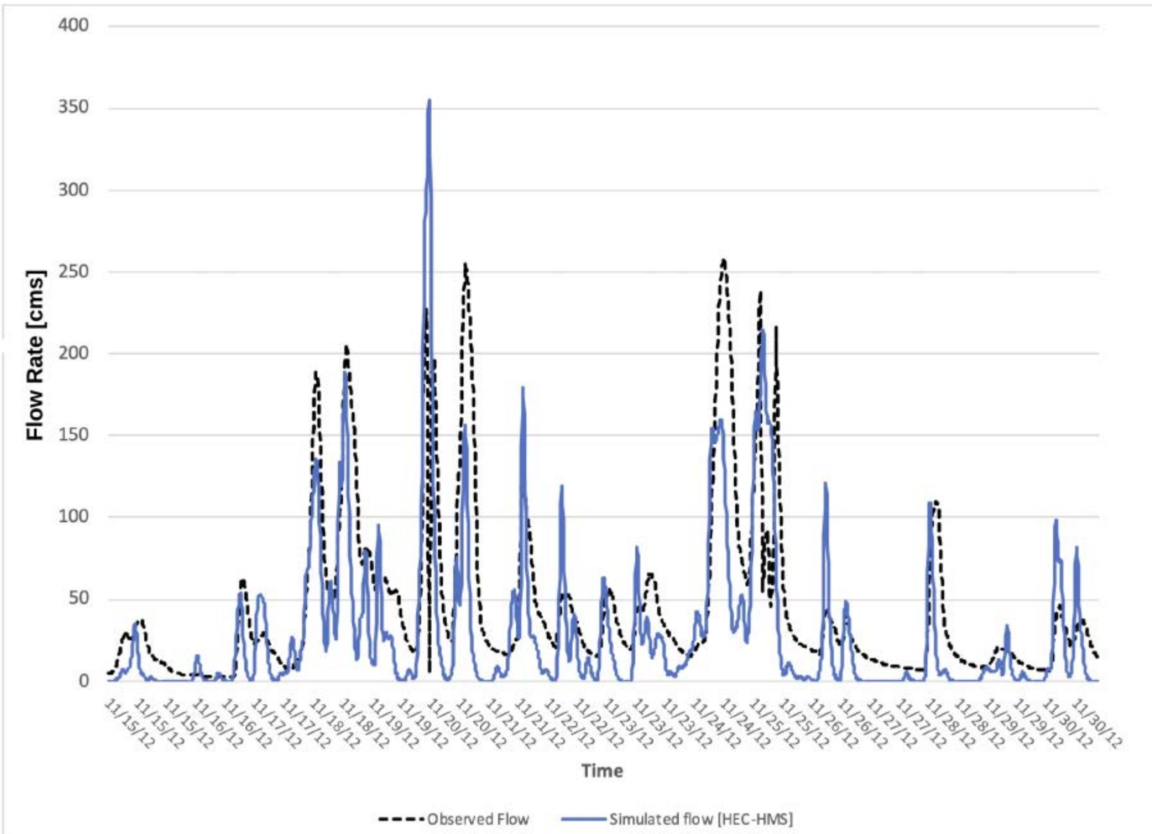


Figure 4.19: Observed vs Simulated Hydrograph for November 2012 flood.

It should be noted that both for the observed hydrograph series as well as for the simulated hydrograph, are in a time resolution of 1-hour interval. This facilitates the comparison of the two hydrographs at each 10-minute time step and consequently allows, the simulated results to be evaluated by relating the overall shape of the hydrographs, and also the size and timing of the peak discharge.

In flood hydrology, several elements contribute to characterize flooding, herewith the two main ones are rainfall intensity and the duration; next, to landscape, soil characteristics, subsurface cover, and land use likewise contribute significant aspects. Despite that, the hydrometric variables which are usually selected to describe floods are the peak discharge Q_p ($m^3 \cdot s^{-1}$), the maximum mean discharge Q_{max}(d) ($m^3 \cdot s^{-1}$), for a specified duration d, the volume of flood (m^3), Time to Peak (TTP) (t_p) in hours, and the flow rate ($m \cdot s^{-1}$) as reported in [377]. As the peak discharge is directly related to the maximum downstream river-stage (flooding), this parameter is the most significant in this study for analysis. In this respect, the observed discharge peaks measured at station H3 that are considerably less than the maximum may have a relationship with increased water levels, though they may not contribute to significant flooding. It can also be observed from the results of the simulated hydrograph as depicted in Figure 4.19 a mild similarity between this modeled hydrograph curve concerning the measured hydrograph curve, however, an overestimation of the peak flow can be noticed by the model. Several factors could have contributed to this divergence of the overestimation of the modeled hydrograph to the measured hydrograph, a possible reason could be attributed to variations in actual flow measurement by the flow sensor for a given time of the day. This is a situation that is likely portrayed by real-time continuous sensors deployed in the field. Submerged sensors function by measuring the amount of water flowing through a section of the stream and fluctuations in streamflow by an obstruction (e.g., debris, stones, and branches) can impair readings. Actual stage readings could have been higher than the average observed value of the day and may have contributed to the high values of the simulated peak discharge. Other factors could revolve around the estimated catchment parameters and as a result, generate higher simulated runoff conveyance than measured.

It should be noted that this flood hydrograph modeling implementation approach has not been done before in the Medio catchment, which implies that no previous antecedent works exist that could have been referenced; therefore, it is important to state that it is assumed the measured hydrograph to correctly describe the peak discharge that occurred on a particular day. In place of this actual description, further adjustments to the model should greatly enhance the simulated hydrograph, decrease overfitting and approximate as closely as possible the peak discharge of the actual hydrograph, the model can then become a plausible estimate that represents the flood hydrology of the Medio river catchment. To attain this goal, it is required to adjust model parameters through the process of model calibration by manual or automatic means.

4.4.2.2 HEC-HMS: Simulation Calibration Process

Previously, it was emphasized on the need to adjust the parameters used for simulating the actual streamflow series to approximate the observed peak flow. Here it is presented the calibration process of adjusting the required parameters that cannot be related to observable physiographic characteristics of the catchment to create a simulated hydrograph that approximates an observed hydrograph as accurately as possible. As it may be recalled, the H3 gauging station only as measurements for three hydrometric parameters (rainfall, streamflow, and water level). Therefore, calibration is necessary to obtain reasonable values of non-existing hydrologic, hydraulic, and the catchment physical parameters for the whole catchment area over each sub-catchment.

For the modeler, obtaining consensus from the simulated hydrographs is essential for ensuring an accurate representation of the catchment with the least bias between the observed and estimated hydrographs [376]. By exploring a set of initial parameter estimates, the model calibration was carried out following the methodology in [376] to find the appropriate values of the estimated parameters that would generate estimates of the maximizing and/or decreasing selection criterion leading to better approximations between modeled and measured flows. This process is challenging, as it is an iterative process, since it requires many simulations experiments to be run, with each experiment using a different batch of parameters, intending to identify a group of parameters that introduces the least bias between the hydrographs [376].

Following the methodology in the HEC-HMS User Manual [86], HEC-HMS presents seven different ways to obtain the objective function. These are "search methods" that are ideally discussed in the HEC-HMS Technical Reference Manual as "search methods". Therefore, the following search methods used for this simulation task are the "Univariate Percent Error in peak discharge" (PEP-Q), and the "Univariate Peak-Weighted RMSE" (PW-RMSE), with the "Simplex Percent Error in Peak Discharge" and the "Simplex Peak-Weighted RMSE" methods used as minimization functions of the objective functions. The selection of the two search functions is straightforward as they contribute greater weight when approximating the peak of the hydrograph. The Univariate PEP-Q method only provides the measure of the right size between the modeled and measured hydrographs; perhaps, no description of the biases in the volume of flowrate or time of peak discharge is given. Otherwise, the Univariate PW-RMSE procedure, considers the peak flows, flow-rates capacity, and the time to peak discharge.

When implementing hydrologic models, it is not possible to tweak all parameters through calibration. Certain hydrometric parameters like the river reach are difficult to calibrate without the knowledge of the hydrographs at both ends of the reach. Moreover, precautions must be taken when tweaking individual parameters dependent on each other. Therefore, the only parameters suitable for adjustment are those that can be fine-tuned throughout the whole catchment by a single scale factor; these parameters are the curve number (CN) and the initial abstraction (I_a).

Finally, four different calibration trials are performed with the selection of two objective functions and two minimization techniques of the objective function, and with the objective function tolerance set at 0.001 and the maximum number of iterations set at 500.

4.4.2.3 HEC-HMS: Simulation Calibration Results

After finalizing the calibration procedures with the use of the four calibration methods detailed in the previous section for the Medio river HEC-HMS model, a close inspection of Figure 4.20 reveals the shape of the calibrated hydrograph to depend predominantly on the objective functions (PEP-Q versus PW-RMSE). The calibrated hydrograph curve also shows that if the objective function is kept constantly, small changes can be observed when changing the objective function minimization method. Results showed the magnitude of the observed peak to be satisfactorily approximated by the goal function PEP-Q, while the PW-RMSE approximation shows to relate to the general volume of the hydrograph. As

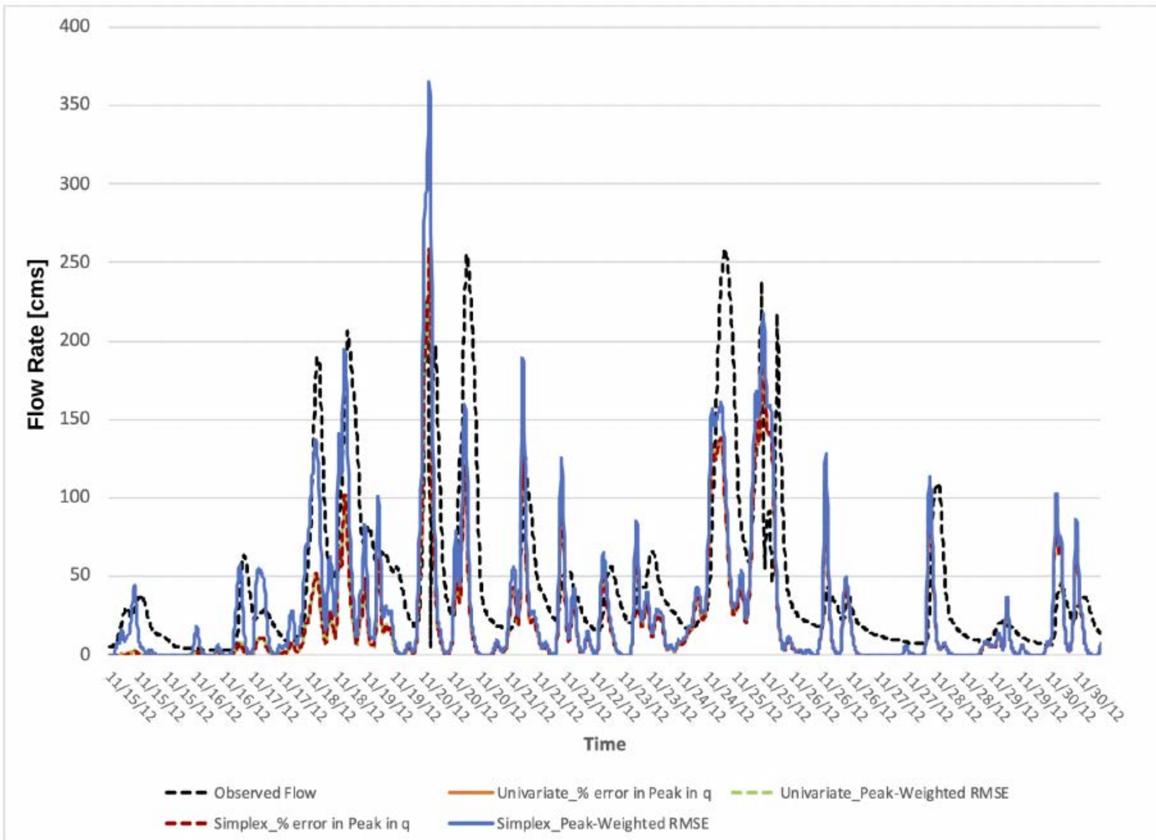


Figure 4.20: HEC-HMS Calibration Methods Results for Station H3 Hydrograph.

the calibration target in the approach of this research is to approximate the peak of the hydrograph flows, it was chosen the objective function "Percent Error in Peak Discharge" for these purposes. Calibration results obtained with the objective function Percent Error in Peak Discharge are summarized in Table 4.13.

Table 4.13: Model Parameter Scaling Factors with Percent Error in Peak Discharge Method.

		Scaling Factors		**Goal Function Value**
Search Method	**Minimization Statistics**	CN	Ia	
Univariate	PEP-Q	1.0108	0.0119	0.02
Simplex	PEP-Q	1.0116	1.14490	0.00

As noted in Section 4.4.2.2, the Univariate object function method only modifies the curve number by approximating the hydrograph, while the Simplex minimization method concurrently modifies the curve number and initial abstraction. Due to initial conditions of soil wetness or dryness before a storm event, the initial abstraction value is expected to be optimized. From the calibration results, the adjustment of this parameter gave a smaller estimate for the selected goal function, therefore, the Simplex minimization PEP-Q goal function method was the method of choice for model calibration. This lowest value, applied in combination with the similarity of the observed peak and simulated calibrated hydrograph, shows the simulation with calibrated hydrograph can approximate the precipitation-streamflow relationship of the actual measured peak discharge at the H3 Station for the November 2012 rainstorm.

The adjustment of the initial values of the basin parameters presented in Section 4.4.1.2, were obtained with the scalar factors found during the calibration process, and the calibrated hydrograph at the Medio River H3 Station gauge location for the November 2012 rainstorm event can be seen below across Table 4.14 and the resulted hydrograph in Figure 4.21.

Table 4.14: Catchment Adjusted Variables After Calibration.

SubCatchment ID	Initial CN	Adjusted CN	Initial I_a	Adjusted I_a
1	89	90	6.3	7.21
2	83	84	10.4	11.91
3	83	84	10.4	11.91
4	83	84	10.4	11.91
5	93	94	3.8	4.35
6	83	84	10.4	11.91
7	83	84	10.4	11.91

An initial simulation of the November 2012 rainstorm event showed an overestimation of peak streamflow at Station H3 that traveled downstream through the entire catchment. For this reason, because of several impairments of the hydrometric data record, intensive data treatment was performed, and the model required calibration in other to adjust the peak flow to the observed value as closest as possible. However, since there is only one hydrometric Station located upstream of subcatchment-1, it must be kept in mind, as was referred by Sherman [378] the assumption that precipitation was relatively and uniformly distributed throughout the entire catchment, both in-depth and in time (Figure 4.16a); so it is necessary to maintain the assumption that, each subcatchment experienced about the same rainfall intensity. This uniformly distributed precipitation caused a rise in river stage along the entire length of the Medio River because each subcatchment contributed flow and therefore resulted in a single upstream flood wave moving downstream.

4.4.2.4 HEC-HMS: Simulating Selected Storm Events

In this section, it is presented the results and discussion for the implemented and calibrated HMS model performance on each of the selected rainfall-runoff storm events. In this respect, it should be noted, that after performing a simulation with a model that was previously calibrated, the following procedure is to perceive the overall modeling performance of the simulation. However, for this, it is necessary to have current hydrometric data of the catch-

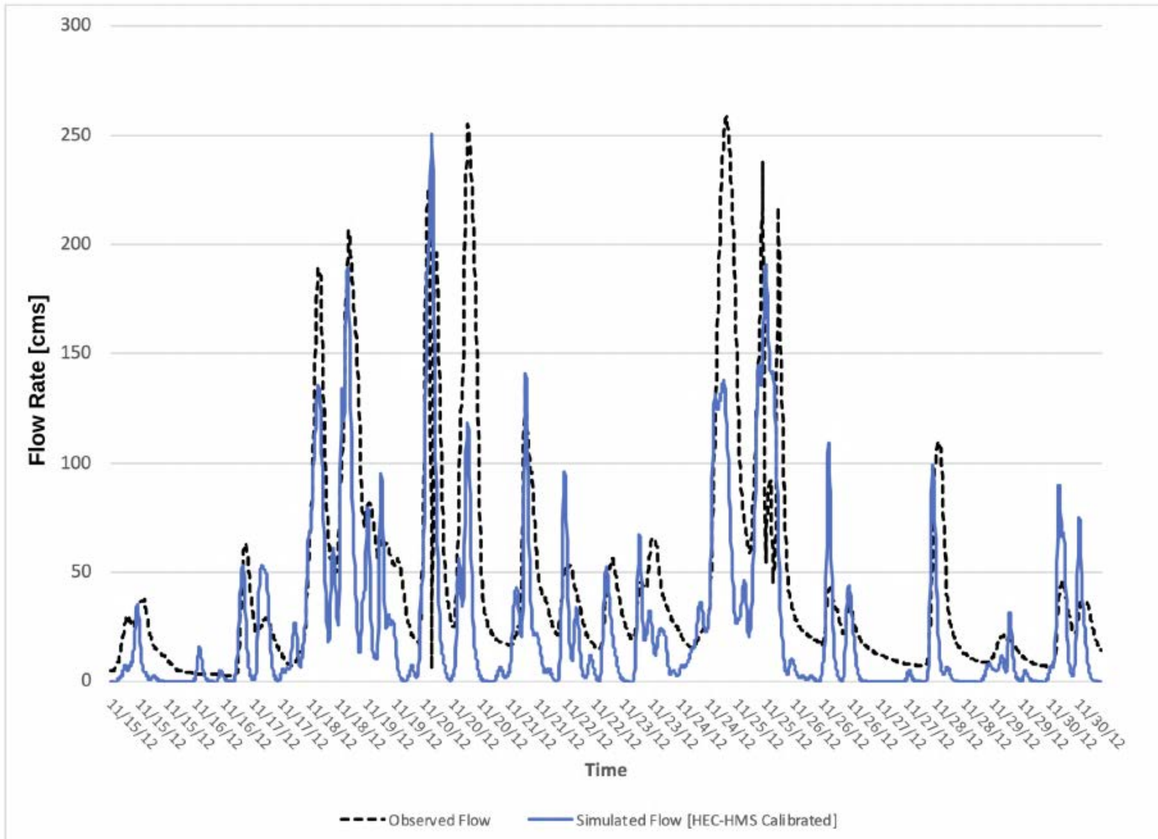


Figure 4.21: Calibration of Observed vs HEC-HMS Simulated Hydrograph for November 2012 flood.

ment and for longer periods. As the former concern is of importance, they selected those precipitation events corresponding to varying rainfall depths observed for some periods in the entire hyetograph of Station H3, like the December 2012 and 2014, May and November 2015 records (Figures 4.16b through 4.17b, respectively) that are likely to produce flood cases. These were selected as simulation case scenarios to assess the model performance with different storm intensities as can be seen below across Table 4.15 and the resulted hydrographs are compared to the observed streamflows for each simulated period.

For the storm events, all flow originates upstream the catchment and the rate is approximately the same. The magnitude of flow over the entire hydrograph is smaller than that of the November 2012 storm as the total volume of precipitation used is different, though it is concentrated over the same area of the subcatchment-1 gauging station, resulting in less initial abstraction, and therefore less runoff.

The statistical metrics observed, allows analyzing the goodness of fit for the model after calibration on estimations of the peak flow surges under different rainfall intensities. However, from Figure 4.22, although the simulation has approached the shape of the observed hydrograph curve, it presented some difficulties in approximating the highest peak flow (Q_{pk})

of the period, as can be seen from the simulated curve, where for an actual $Q_{pk} = 266.4$ $m^3 \cdot s^{-1}$, the model estimated Q_{pk} discharge was 183.2 $m^3 \cdot s^{-1}$, which corresponds to a value of the peak runoff error of 31.2% and RMSE of 14.7 $m^3 \cdot s^{-1}$. However, the model agreement is also assessed by the simulation correlation coefficient of 0.94, which demonstrated the calibrated model to be practical for flood hydrology estimation. Besides these results, the model attained an accuracy of 88% in forecasting the flood hydrograph, with p<.001 at $\alpha = 0.05$ which shows significantly good. On the other hand, model performance with the December 2014 Storm simulation can be observed in Figure 4.23, despite the graph bears some resemblance to the pattern of the observed hydrograph at station H3, it presented difficulties toward the peak flow approximation. Nevertheless, the overall efficiency of the simulation task, in general terms proved purposeful as shown by the hydrograph simulation statistics (Table 4.15) with a percentage error in peak flow of 66.5% between model simulation and observed streamflow (112.4 and 335.5 $m^3 \cdot s^{-1}$, respectively), and the standard distance between the HMS forecasted and observed values of 29.6 $m^3 \cdot s^{-1}$. In addition, the model showed to similarly agreed with the observed data as shown by the correlation coefficient (r = 0.82) and with much of the variability explained by the observed data only accounting for 67%, attesting also for the strong relationship between HMS simulated model and the observed data, and this relationship significant given a p<.001 at $\alpha = 0.05$.

Table 4.15: Statistical metrics of the performance between HEC-HMS simulations and the observed flood hydrographs for various Storm Events.

Validation Storm	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	Obs. Q_{pk} [$m^3 \cdot s^{-1}$]	HMS Sim. Q_{pk} [$m^3 \cdot s^{-1}$]
December 2012	0.94	0.88	14.7	31.2	266.4	183.2
December 2014	0.82	0.67	29.6	65.5	335.5	112.4
May 2015	0.82	0.67	15.9	37.6	257.1	160.5
November 2015	0.76	0.58	11.6	17.5	264.4	309.9

The graphical information in Figure 4.24 for the storm simulation event of May 2015, reveals little differences with simulated results concerning the similarity with the observed hydrograph. However, it is observed one to three overestimated low peaks, and the model also undergoing difficulties to reproduce the Q_{pk} of 257.1 $m^3 \cdot s^{-1}$ for the period with a simulated flow value of approximately 160.5 $m^3 \cdot s^{-1}$. Nonetheless, in statistical terms for model (Table 4.15) efficiency, the model as acceptable similarity agreement with the observed data given the resulting correlation coefficient (r = 0.82), with model accuracy accounting for 67% (p<.001) and the errors of spread and peak flow between model and observed values of 15.9 $m^3 \cdot s^{-1}$, and 37.6%, respectively.

Finally, in Figure 4.25, it is observed that for November 2015 Storm the results from modeling the entire month, although there are days with precipitation lower than 100 mm (Figure 4.17b), the HMS did estimate runoff hydrograph for these days as can be seen in Figure 4.25. However, for some reason, the observed flow at station H3 doesn't follow this trend accordingly. Conversely, the results between model simulation ($Q_{pk} = 309.9$ $m^3 \cdot s^{-1}$) and observed values ($Q_{pk} = 264.35$ $m^3 \cdot s^{-1}$) for the period, shows the HMS to overestimate the

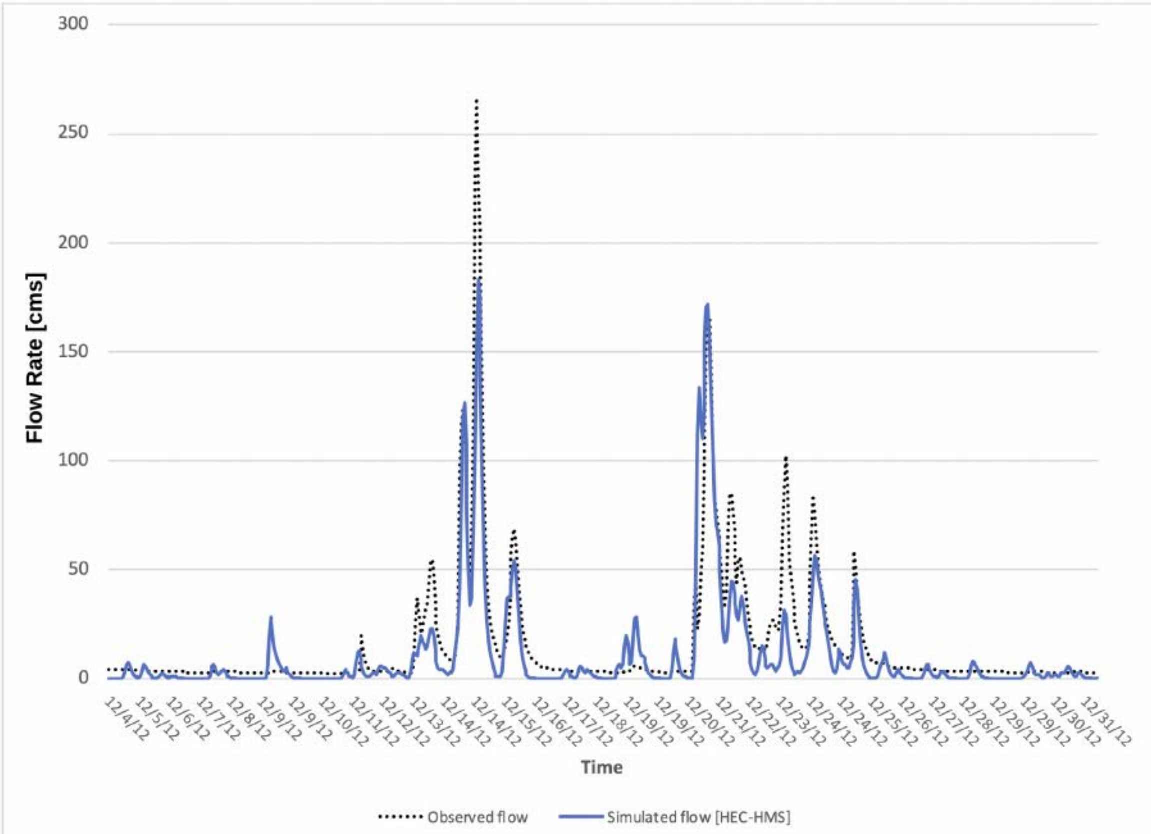


Figure 4.22: Observed vs HEC-HMS Simulated Hydrograph for December 2012 Flood Event.

measured runoff, which in statistical terms represents an error of the spread between modeled and observed data of 11.6 $m^3 \cdot s^{-1}$. Besides the comparison between model-simulated and observed flow shows an error in peak flow of 17.5% and a lower correlation value (0.76), as compared to the other simulated periods. The results also showed, the HMS model on this data, the observed data to only explain 58% of the variability concerning the modeled outputs; despite that, it could be argued the acceptance of the model, since, as the results indicated the relationships in the model are statistically significant with the observed data (p<.001, at $\alpha = 0.05$).

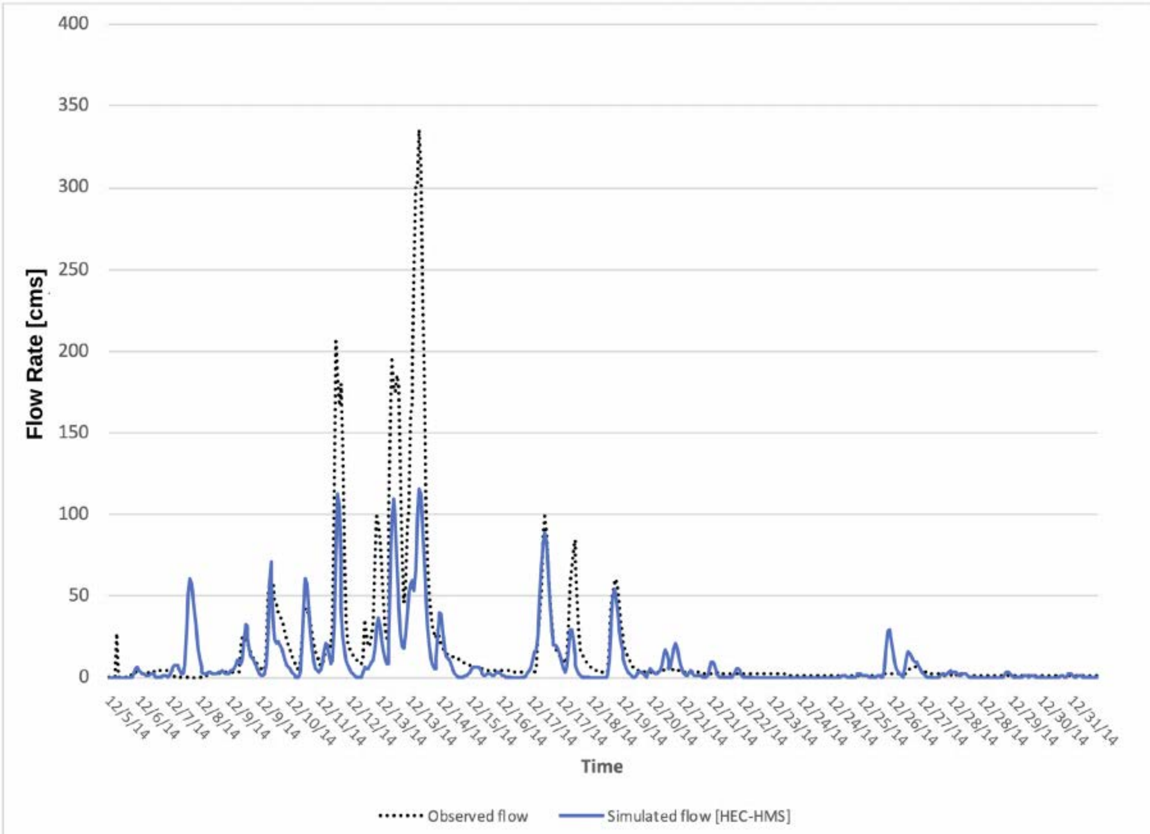


Figure 4.23: Observed vs HEC-HMS Simulated Hydrograph for December 2014 Flood Event.

4.4.3 Conclusions

In this section, using the standard hydrologic model "HEC-HMS" a typical flood hydrograph software package, was used to performed hydrologic modeling, which in hydrological studies represents the analysis of the simulation for "rainfall-runoff" processes and analysis to assess the occurrences of floods, for flood hazards assessment, water resources management, and engineering design purposes. The model was developed to serve as a standard benchmark hydrologic modeling approach and reference point from which can be build an ABM prototype that can offer the same functions as the hydrologic simulation for flood forecasting.

Once the setup of the model is achieved, and by set up, it is meant that all components (e.g., GIS, meteorological model, data preparation, control specification) of the HMS basing model are in place, the model for the Medio River is run for calibration of the sensitive basin parameters as was observed in Section 4.4.2.2 and 4.4.2.3. After this calibration was achieved, the model was validated on four storm events (December 2012, December 2014, May 2015, and November 2015) of the Medio River watershed, from which outputs of the HEC-HMS were contrasted to the observed hydrographs as reported in Section 4.4.2.4. The main clues drawn from the analysis are:

- Modeling with HEC-HMS provided insights into the possibility to perform simulation of

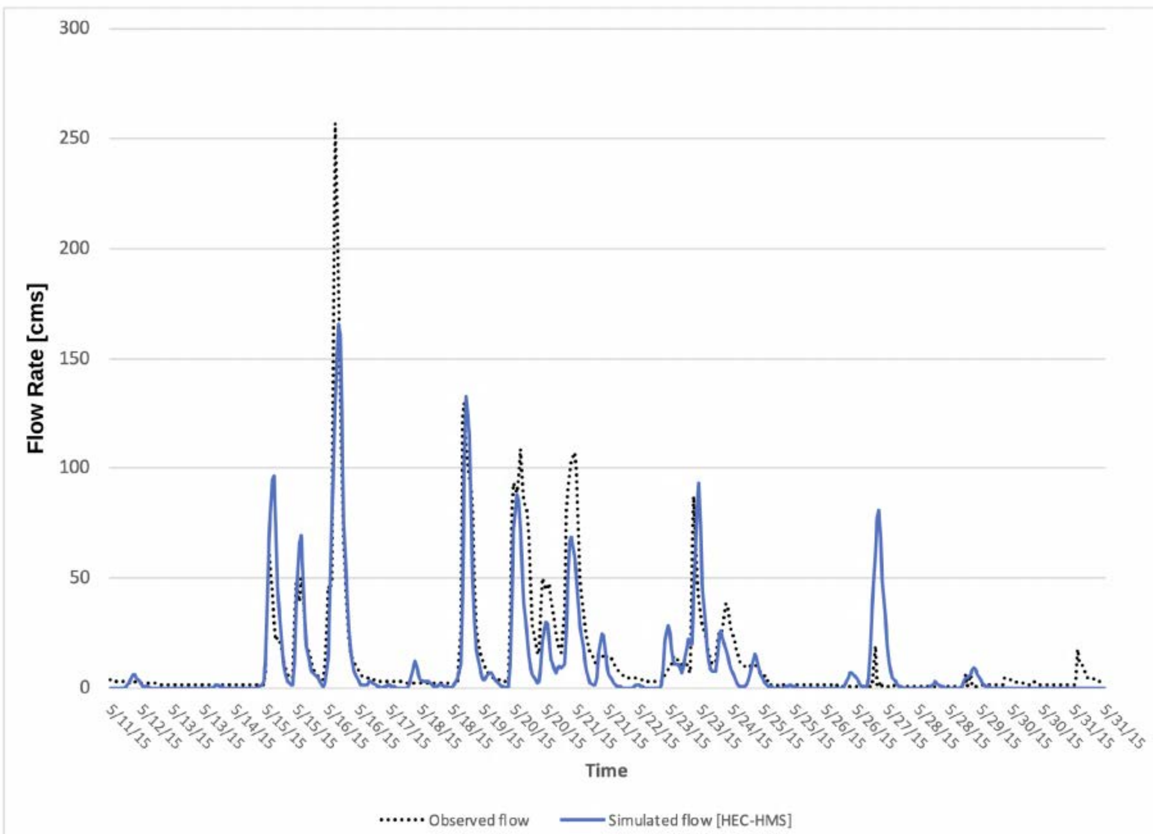


Figure 4.24: Observed vs HEC-HMS Simulated Hydrograph for May 2015 Flood Event.

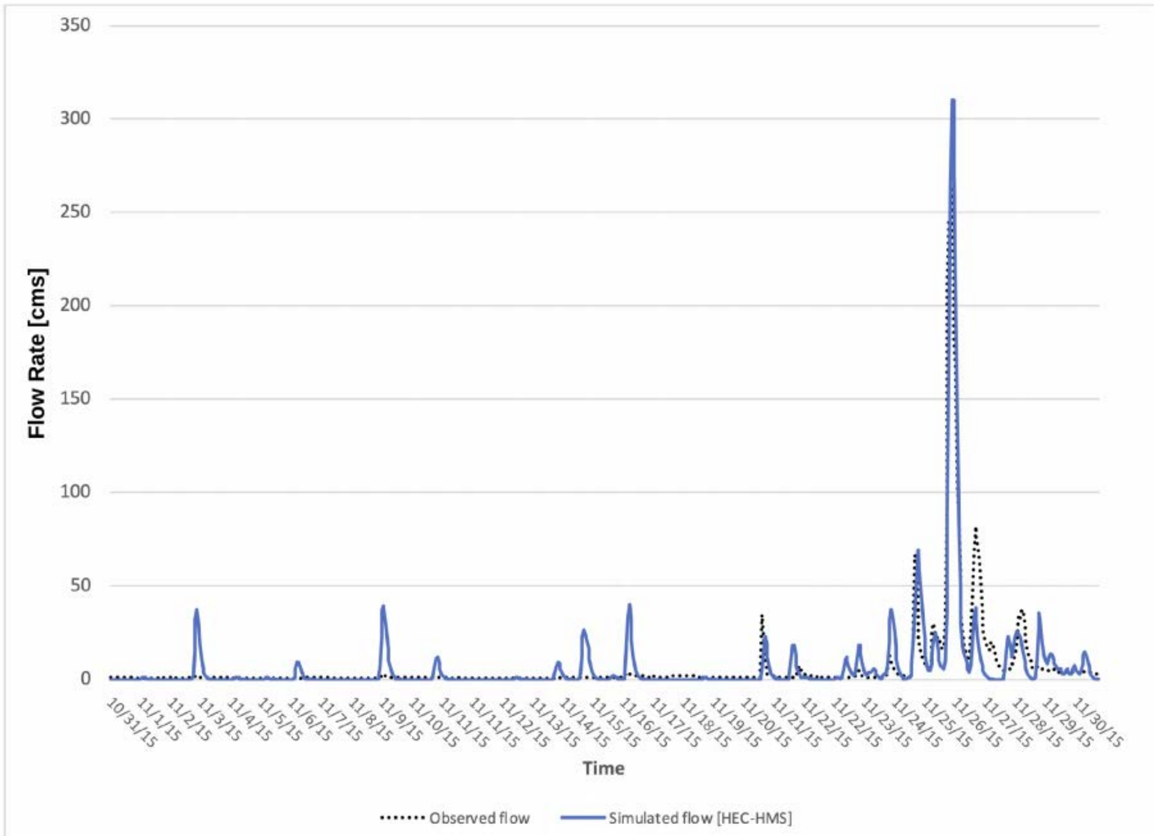


Figure 4.25: Observed vs HEC-HMS Simulated Hydrograph for November 2015 Flood Event.

the precipitation-runoff modellng for the Medio River catchment. Despite some cases of misestimation of the simulation outputs concerning the observed peak flows, the resulting objective function, and simulated hydrographs after calibration were found to be good estimates.

- The model was simulated under the lumped concept, this could have probably led to the limitations in the peak flow simulations. However, this can be improved by optimizing the model further under the quasi-allocated concept.

- Upon the continuous optimization of the model it is more likely it can be useful in other tropical watersheds of other sizes.

4.4.4 GAMA: Simulation Setup

As shown in Figure 4.1, one of the motivational objectives of this research is to recreate the simulation of hydrological floods in a tropical catchment using the ABM paradigm environment rather than the standard hydrological model archetype such as the HEC-HMS software model as was done in the preceding sections. From the agent development platform review and some relative comments shown on the topic under Chapter 3, it was found that there are several platforms for developing the agent-based modeling task: some of these for example "NetLogo", "JADE", and its extension "JADEX", "MACSim" [379] software, developed to allow communicating JADE via Simulink are among the most popular, and the recent GAMA platform. Despite these and other choices, and since much of the inputs used in this simulation depend on a large part of georeferenced information, the GAMA platform is the agent simulation environment of choice selected because this platform offers the possibility to use as input the GIS data products and making it possible to represent the domain data as mappings, the variation of the river volumes, visible streamflows, and the water levels (stage height). Furthermore, given that it is highly GIS oriented, it allows assigning different values to the various configuration parameters of the environment domain by modifications of the main code.

GAMA platform can discretize these items into different spatial layers and represent them independently, allowing the interaction of these layers clearly and easily. In this manner, each subcatchment in a particular catchment of study can be considered a specific agent with its characteristics and attributes. In Figure 4.26 for example, it can be visualized the format and location of each subcatchment, as well as add the color attribute to each of them and facilitate their distinction. GAMA allows the user to relate the existing information through the position and topology of the objects, thus generating new information. In this case, considering the juncture of information in Figure 4.26a (subcatchment) and Figure 4.26b (rivers), it can be obtained the representation of the image presented in Figure 4.26c (rivers and subcatchment), from which the platform can extract new information and attributes for each object, such as associating a river with a given subcatchment according to its location.

4.4.4.1 Framework Setup for ABM for Hydrologic Flows Simulation

This section provides the details of the agent classes (species) for the ABM Medio catchment hydrologic modeling framework. The Medio catchment is composed of seven subcatchment already described under Section 4.4.1.1 and here in (Figure 4.27a), with distinct characteristics, as presented as an example in Table 4.16 the attributes of the AgentCatchment. For this agent-based two level architecture, the framework arrangement (Figure 4.28) consists of the following agents forming the basis of this system and some of which can be identified as static and non-static agents:

- Hydrometric sensor agents (**HSn):** Three agents that don't exhibit any mobility during simulation cycles and as described in Section 4.4.1.1 they are represented by the hydrometric station sensors deployed in the field. Their roles are described below except for some roles that would be implemented in their BDI versions. It should be noted that these agents are set up to connect to the sensors deployed in the field; however, it should be stated the fact that due to technical issues and availability for onsite station access at the time of development of this project, in this simulation and the simulation undertaking consequently in Chapter 5, the HSn captures the hydrometric data basically from file folders which are emulated as the field-deployed datalogger. Therefore, no experiments are conducted with data captured in real-time. For now, the experiments have been done as explained, and works for coupling the model prototype to a real-time field-based station remain for future work.

1. Rainfall sensor agent (**AgentRNSn):** The AgentRNSn role is to capture, aggregate, and distribute the real-time incoming rain data readings to the river agent sources.
2. Water level sensor agent (**AgentWLSn):** The AgentWLSn role is to obtain, aggregate, and distribute the real-time incoming river surface water level data to the river agent.
3. Streamflow sensor agent (**AgentSFSn):** The AgentSFSn role is to obtain, aggregate, and forward the real-time incoming flow discharge data obtained from the field flow meter sensor data to the river agent.

- Environment domain agents (**EDA):** Four agents making up the catchment environment along with the GAMA generated default global agent.

1. Catchment agent (**AgentCatchment):** This agent represents the real-world Medio River catchment (River Basin), so it is considered an individual static agent, has specific values for the size of its area, catchment order, neighboring sub-catchments, and its drainage outlet into an adjacent channel and the main Medio River channel. It is therefore a layer upon which the rivers and the monitoring stations are displayed. This layer provides what would be the morphometric characteristics of the catchment, which is essential along with the grid agent, to provide the gradient behavior on the catchment and to directly act on the river agent in the conveyance and exchange of water through the channels by gravity.
2. Water source agent (**AgentSource):** Represents the hydrological agent responsible for providing to the rivers an amount of water from the flow and precipitation series inputs. The AgentSource at the beginning of a simulation has the same amount of water; however, this amount is changed over time given the location and special extent of the rivers they affect and the changes in input precipitation values. Source agents are linked to a river inlet. At simulation start time, given the frequency of action that is given to the source, the global agent will ask the source agent to give water from either a set input volume, the flow, and/or precipitation inputs to the river. In this case, this is set to every hour for one simulation scenario and consecutively it can be set to two, three, four hours, ten or twenty minutes, or even days, months, depending on the desired time length. Nevertheless, care should be taken when doing this setting, as it can impair visualization of the flow of water in the river.
3. River network agent (**AgentRiver):** The river agent is also a static agent distributed between the different subcatchments according to their geospatial location. The river agent receives and shares information about incoming water from the source agent, conveying the water as a volume in the rivers from uphill to downhill the catchment. This amount of water within the river is generated from the runoff that comes from the precipitation series information, or it can be given

initially a set value, or it can be set to a range that would change over time. Additionally, the routing of flows is gained through water exchange between the river reach segments across a bordering catchment, this is also dependent on the precipitation amount and frequency. The water that enters the rivers from the source is routed through as a volume, as discussed in detail in [380]. Therefore, the AgentRiver has the role of computing these volumes, generating the flow of water, updating water volumes, and computing the water levels through actions, each requested and managed by the global agent. In the beginning, before initializing a simulation, the volume of river water will increase and decrease over time according to the precipitation rate of the subcatchment in which the river is inserted. Each river belongs to a single subcatchment or may cross another. Flood water recedes after a precipitation event as occurred over the catchments or if lower than the precipitation rate of return, the tendency is that the volume of floodwater in the rivers will grow over time and the subcatchment will become flooded conveying their waters to the main Medio Channel as can be seen in (Figure 4.27c). Neighboring subcatchment experience flooding of their streams according to the precipitation distribution and contribute as lateral inflows into the Medio Catchment main channel. The capacity to convey water is established by its hydraulic characteristics, from thence a maximum volume (vol_{max}) value of the water, so, the more the water volume of the subcatchment exceeds this value (flooding), the higher the water level and the discharge at the Medio river outlet will be.

4. Terrain elevation agent (**AgentDEM):** The DEM that is being "agentified", it is a special type of agent species portraying a grid topology. It has no mobility during simulation time, is a static agent. It represents the terrain elevation of the catchment. This agent has the role of representing the gradient profile distributed across the entire catchment.

- Global agent (**AgentGlobal):** Like the grid species, the global species is a special type of agent. It is an automatically created agent in GAMA. Its instance is the world and it inherits automatically from the default variables and actions within GAMA. In the AgentGlobal, all variables, parameters, attributes, actions and behaviors that governs the world, can be described. The global agent (AgentGlobal) manages all the agents in the system; besides that, it is also responsible for the saving of the information generated during the simulations. Therefore it can be seen as a supervisor agent.

4.4.4.2 Required Datasets and Features

- The GIS data:

  In this model, the GIS data are presented as a database with geographic information, found in a 30m digital elevation model (DEM), two shapefile, one containing geospatial information for the Medio river catchment and subcatchment, and the other archives describe the spatial qualities of vectors (points, lines, and polygons) to represent the

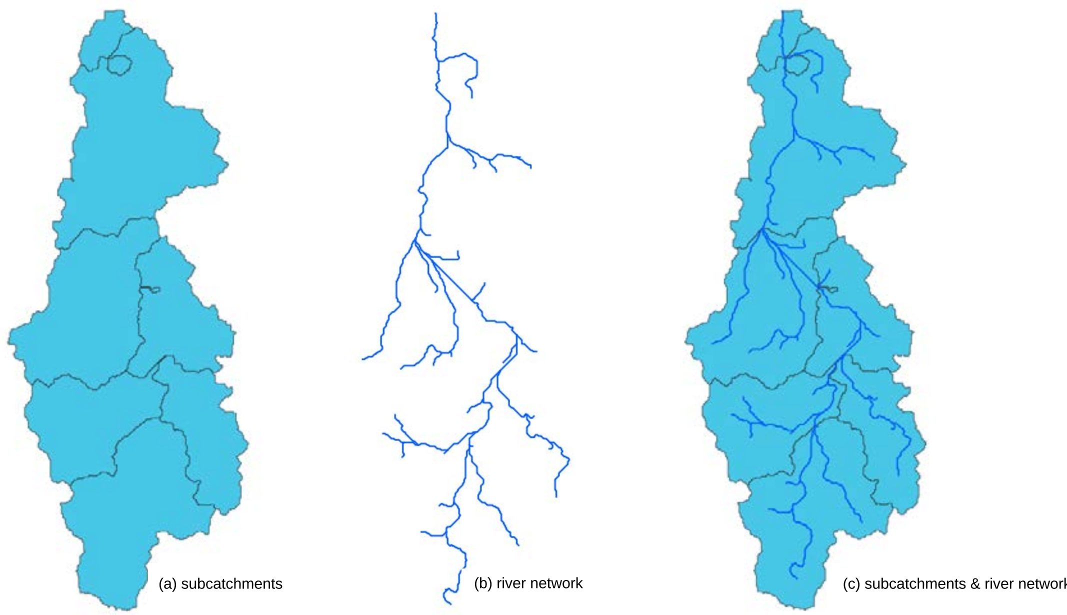


Figure 4.26: Medio catchment with mini-catchment GIS data sources graphical representation.

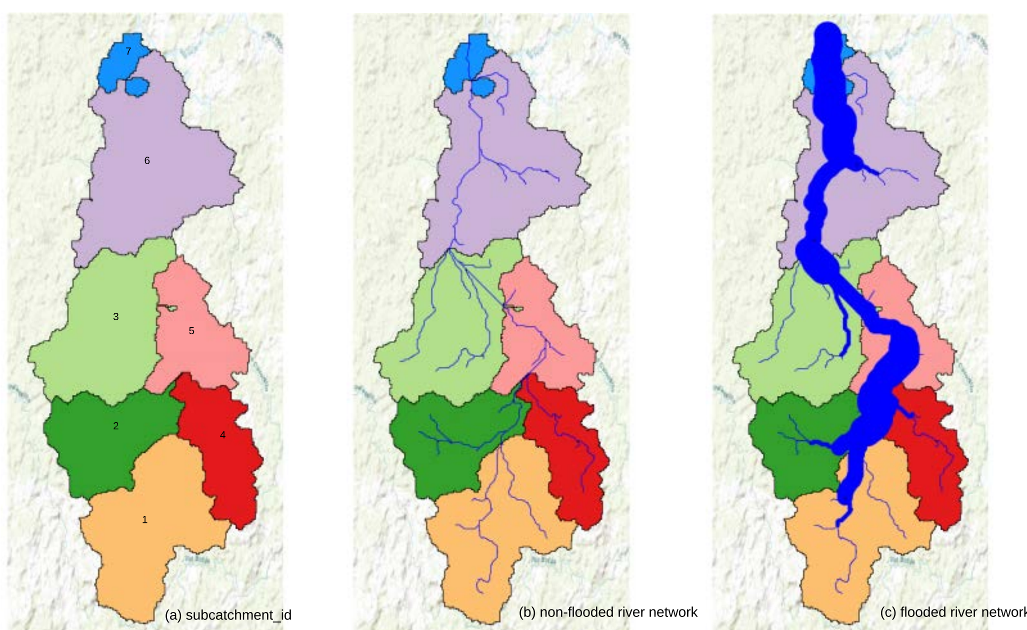


Figure 4.27: Medio catchment with mini-catchment id GIS data sources graphical representation.

Table 4.16: Definition of spatial attributes of the AgentCatchment.

Subcatchment_ID	**Area [km2]**	**Order**	**Catchment outlet_ID**
1	9.9	3	4
2	6.3	2	4
3	9.6	3	4
4	4.9	1	8
5	5.3	1	8
6	12.8	2	8
7	1.3	1	8

subcatchment and the streams composing the Medio catchment, with their attributes that describe them (e.g., stream length, basin elevation, area, and drainage outlet).

- The Base Map:

  In the model, it represents the simulation environment of the location of the Medio River Catchment. This map was extracted from part of the Donoso District in Colon City, from OpenStreetMap using the QGIS to import and get the shapefile, for use in the GAMA platform to recreate the area of the catchment as depicted in Figure 4.1.

- The Precipitation:

  The precipitation, a meteorological 1hr interval time series dataset that is one of the core elements of this modeling experiment, as it is responsible for producing the inundations in the Medio river main channel. When it is applied the precipitation regime, flood waves are systematically generated from adjacent rivers as lateral flows towards the direction of the Medio river course. This, flood wave is then conveyed to the Medio outlet which forms a confluence with the larger Caimito River. The main features that distinguish this flows conveyance, are the excessive inundation intensity of the stream banks and flood plain areas due to the high stage that is produced by the fast-rising flood waves. Among the other hydrometric elements, it is observed surface water elevation and streamflow time series files.

4.4.4.3 GAMA: Input Parameters

In GAMA, a model is built by providing the platform with GIS data, this represents the dynamics of the capabilities of the platform to read and write GIS data, and in using this data in models. In GAMA, the input data is represented by the integration of vector data to the simulation's environment, and in turn, returning this information as the resulting modeling system. In addition to GIS information, GAMA can also accept as input, other data files such, images, data feeds, databases, and time-series information in a text or csv format. However, a description of the data requirements used in the platform is detailed in Section 4.4.3.1.

4.4.5 GAMA: Simulation, Calibration, and Selected Flood Case

Some of the settings of the input data for the GIS environment initially created and used in the previous hydrologic modeling with the HEC-HMS physical-and-equation based hydrologic model had set the pace for some of the input data required and used in this GAMA

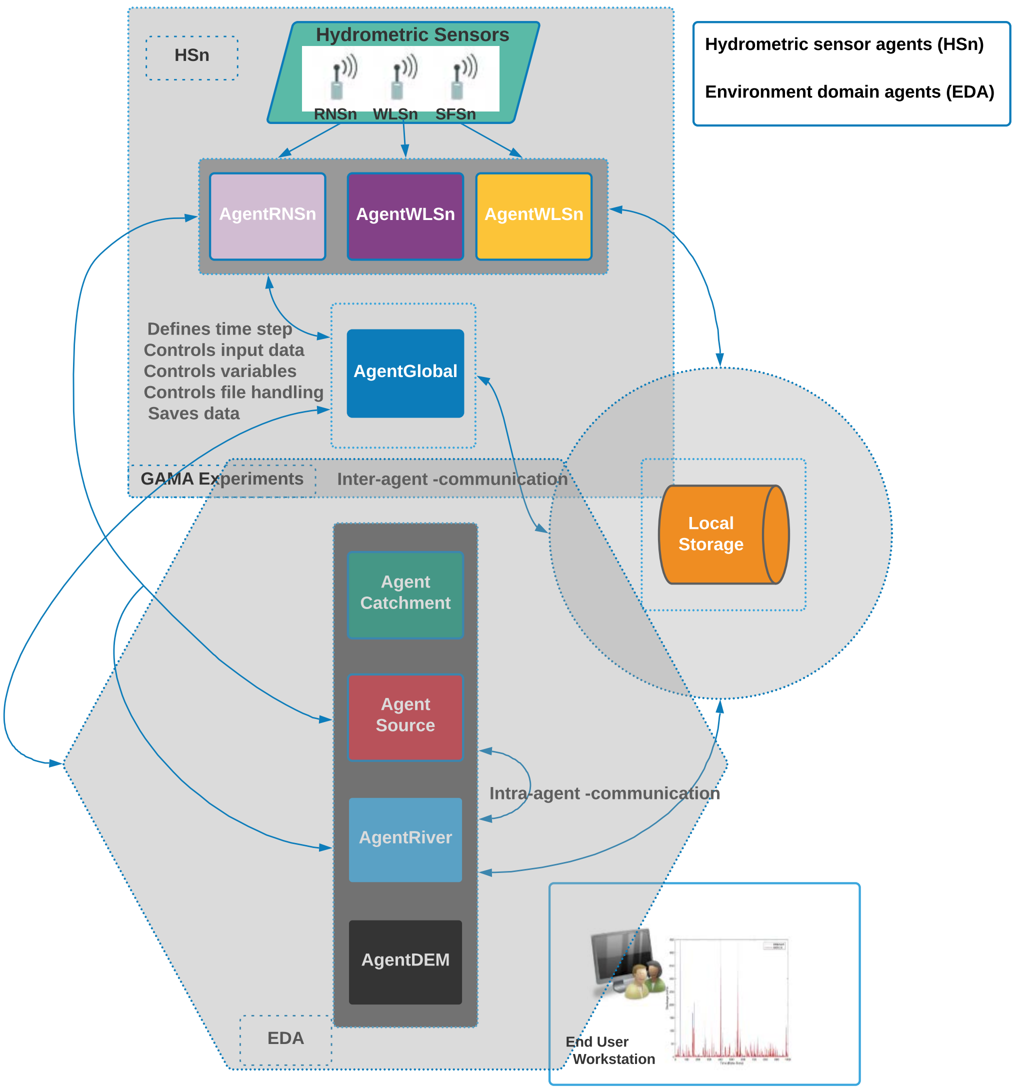


Figure 4.28: Schematic diagram of the ABM showing the overlapping levels hydrometric sensor level (HSn) and the environment domain level (EDA) containing the GAMA automatically created global agent (AgentGlobal) along with the other agent actors' implementations for the system comprises the experiment environment. The diagram also shows the role of the global agent as the lead agent, who practically controls the types of agents, be they static or mobile, and the interactions among agents at an intra and inter-agent-communication level that is needed at the initial run time when they execute a simulation.

hydrologic simulation. The aforementioned is of utmost importance for the simple reason that the implementation of hydrological models requires a large number of physical and physiographic coefficients of the basin, which makes its implementation a complex task. These variables vary from terrain data, elevation, precipitation, evaporation, land use, infiltration, soil-water interactions, overland flow, just to mention a few of the data requirements.

As specified before, the term hydrologic model refers to the governing equations that intend to approximate the response of the hydrologic cycle. Therefore, to model such system response, information about the physical states of the system must be known or at least estimated to allow the implementation of the modeling scheme specific to a particular computer application by mathematical equations. If these equations result too complex to solve, then translation into computer code and algorithm is needed.

The adoption of such modeling approach is nowadays been implemented thanks to the concept of ABM and MAS, which reach as not only taken a full hold on the social sciences; and with very few researches in the hydrological field, is growing attention especially in the flood prediction and management environment [381].

Hydrologic modeling is now possible at the agent scale, because of the use of GIS data integration into the agent paradigm, in other words, "agentification of GIS data" [369].

This agentification capability of GIS data can allow the creation of an agent's implementation with behaviors of precipitation, catchment characteristics, elevation grid, water flow, water volume, water level, and river network, by simulations using the GAMA platform. In GAMA these data are called species.

4.4.5.1 Hydrologic Simulation with GAMA

Frequent storms with high flood water production that cause losses to human lives, livestock, and damage farmland, and the economy is a common phenomenon in tropical regions. In the tropical river basin, which oftentimes is ungauged; flood warning and management systems are seldom present and are most likely not available to stakeholders and the inhabitants of those regions as ready-hand tools and applications. In this sense, knowledge about the time and origin and distribution of runoff entering locally into the stream during a storm would be valuable hours ahead before matters get worse and it leaves no time for an evacuation. On the other hand, at the catchment level, government officials, managers, stakeholders, and the public do not understand the physical properties and dynamics of catchment hydrology.

Another difficulty is that presented by most distributed hydrological models, which can mean a costly purchase, and which usability and the learning curve can become very difficult, and time demanding, besides they are data-greedy, hard to implement and computationally time-consuming. This is even true for the open-source counterpart.

With some of these facts about distributed, physics-and-equation based hydrologic models, it is the motivation to embark on the quest for the exploration, experimentation, and implementation with the GAMA agent platform as an alternative tool in hydrological flood simulation, a platform which GIS capabilities as been announced in the previous sections.

In theory, GAMA is not a standard hydrologic simulation software created and dedicated solely for such purposes per se. Therefore, as shown previously the simulation process in GAMA is different from how it is carried out in the HEC-HMS standard hydrologic application. GAMA platform offers the experienced modeler as well as the novice with limited programming abilities a modeling complex environment for linking Agent-based simulation and GIS features. The GAML language emulates the "object-oriented programming (OOP)" concepts, as such the idea of "class" is abstracted in the concept for specifying objects as agents, with their attributes, actions, behaviors, and other properties of the agents' population, extracted from the GIS data which can be presented as dynamic maps, graphs, and charts [382, 383]. An advantage within the GAMA simulation environment is the ability to allow the user to pause a model during execution as commenced, and in such a way the user can perform adjustments to the code or variables. Such that, when a simulation session is completed, the results obtained are useful for the scrutiny and examination of its efficiency. These outputs will permit the evaluation of the performance of the modeled system. Therefore, it is presented in this section the simulation schemes, the calibration tasks, and finally the evaluation of parameters.

As it was with the HMS simulations, the November 2012 1-hour interval time series was used to run the simulation scenarios and calibration process for the ABM hydrologic model preliminary simulation in GAMA just as it is discussed in Section 4.4.2.1. Whereas in HEC-HMS simulation time step ($\Delta\tau$) must not exceed 29% of the lag time ([376]), simulation is carried out each 10-min and in the case of GAMA (See [303]), the time step must correspond to the time frame of the time series data, in which case $\Delta\tau = 1$-hour simulation step and so on. In this simulation, the approach for the environment setup had the following configuration for the initial global parameter values for some hydrometric state variables such as the observed rainfall, water stage, and streamflow. As an example, they can define initial conditions for the precipitation, water level, and the flow, from an input time series or the values, which can be defined randomly between a ranging boundary, as per example the flow in the range [0.1, 1000 m^3/s], stage in the range [7.03, 15.0 m], and the volume of flow in the range [100, 500 hm], and so on, before initiating a simulation, which would be started at the model initialization. The other elements that comprise this environment are the agents comprising the HSn and the EDA levels (Figure 4.28), for example, the rainfall time series data is captured by the AgentRNSn, the water level and streamflow time series is captured by the AgentWLSn and AgentSFSn agents, the river network, represented by the AgentRiver, the sub-catchment agent (AgentCatchment), the water source (AgentSource), and the grid agent (AgentDEM). These agents all of which are static, reactive agents. Perhaps, as there is only one rain gauge available (Upstream at Station H3), it is taken on the assumption of the hydrologic lumped model concept, as it is considered the rainfall as uniform distribution across the basin.

The simulations were performed over the entire November 2012 storm period and were considered as a tune-up for a first flood hydrograph simulation with the ABM hydrologic model. At the end of the simulations, the flood hydrograph approximation resulting with the ABM shows the simulated versus the observed hydrograph (Figure 4.29) to present some difficulties to closely replicate the shape of the observed flood hydrograph curve, as it showed to either overestimates or underestimates some low and high peak flows, the performance metrics for this tune-up run produced a correlation coefficient (r = 0.40) in contrast to the value obtained from the HEC-HMS run before calibration (r = 0.68), the RMSE = 56.9 $m^3 \cdot s^{-1}$ when with the case of the HEC-HMS (RMSE = 43.8 $m^3 \cdot s^{-1}$), and finally, a percent error of peak discharge of 72.2%, distinct to 47.6% obtained with the HEC-HMS counterpart. Generally, it is considered the low correlation values below 70% between model simulations and observed hydrograph as models that are probably not capable of providing consistent simulations of rainstorm-induced runoff for flood predictions. However, the interpretations of the correlation coefficient can be biased, since it is not known if the nature of a dataset had contributed to the actual metrics resulting from the analysis. Moreover, to further investigate, this low correlation coefficient, the next step would require the calibration task, to see if the results can outperform the approximation resulting from the tune-up run (r = 0.40) with a model accuracy of 32% and $p < 0.001$ at $\alpha = 0.05$; and therefore, in the next section it will be used the other storm events to verify the overall goodness of the model.

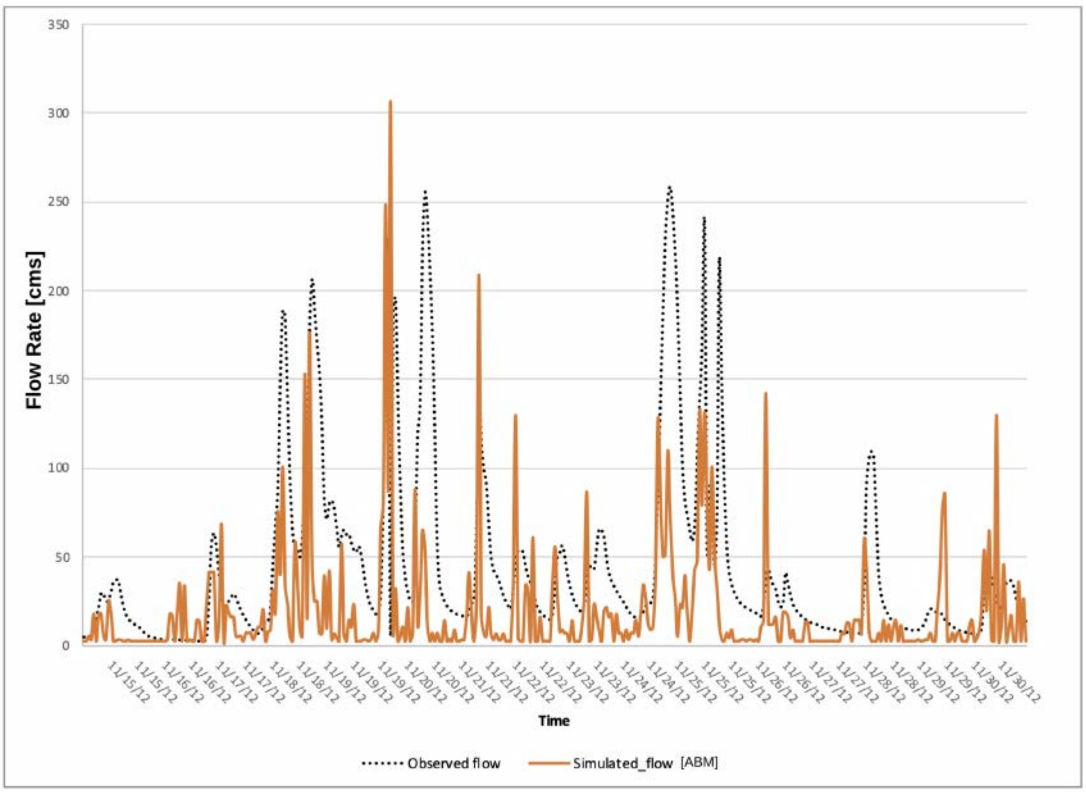


Figure 4.29: Observed vs ABM Simulated Hydrograph for November 2012 flood.

4.4.5.2 GAMA: Model Calibration Process

The calibration process for the GAMA rainfall-runoff simulation differs from the one performed in the HEC-HMS simulation setup. As discussed above in Section 4.4.2.2 the HEC-HMS uses different objective functions and a set of parameters for model calibration. In contrast, to a model in GAMA that can be calibrated by exploration through the introduction of "batch experiments". By setting a set of formulations, models can be evaluated to analyze the sensitivity toward "stochasticity", and an experiment can also be added to explore the impact of a variable on the model outcome, and experiments implemented for calibration. These additions will simply add a new experiment to the model in batch form. However, more aspects of these experiments are detailed in GAMA documentation [303].

In principle, it is necessary to emphasize the fact that none of the formulations offered by the literature in GAMA and in general have defined guidelines on how a flow forecasting simulation model based on agents should be calibrated and verified. It is known from the literature that the process of calibration and verification of systems based on agent simulation is a subject of challenges, specifically when they must deal with complex, large-scale domains. integrated, and hybrid systems, that involve large volumes of data entering and exiting the process, thus limiting simulation time [383–388]. Despite the calibration and verification challenges that these systems may face, they still require certain procedures to facilitate calibration and verification. In this sense, in what concerns the calibration of simulation models, many authors have used certain procedures such as multi-objective optimization (MOO) [389–391], genetic algorithms (GA) [392–394], automatic methods for calibration (AMCMAS) [395–398], statistical estimators (SEST) [399, 400] among others. Finally, in dealing with the question on calibration and verification within the GAMA platform, Truong [401] proposed a methodology employing an index of agreement for "calibration and validation of an agent-based simulation model applied to business intelligence databases". However, this project is based on the domain of hydrological modeling. In this sense, the aforesaid methodology is not appropriate, but rather one that follows an understanding of trial-and-error [402, 403] while the calibration and validation processes based on the hydrological principles properly established for the analysis of flood hydrographs are carried out.

As it was done priorly in the experimental setup for the calibration task of the previous HMS simulation, in GAMA it is also carried out with the November 2012 storm data, since it represented the wettest period of that year followed by December 2012. For this storm dataset, a full simulation was run for each of the experimental scenarios implemented, and the corresponding initial global inputs variables were adjusted accordingly. Then, the results obtained from the calibrated ABM experiments were evaluated based on their efficiency using the usual statistical estimators (e.g., "correlation coefficient (r)", "root mean squared error (RMSE)", and "percentage error in peak runoff ($Q_{pk}\%$)") commonly used in hydrologic modeling to analyze the performance between the simulated hydrograph outputs concerning the observed hydrograph outputs, in this case, represented by the ABM hydrologic model that was implemented in the GAMA platform.

After the tune-up runs simulations had been completed, recall that the results indicated, the model showed to have problems approximating the shape of the curve of the observed hydrograph, and it either under or overestimated the flows, even before calibration. Then, this was indicative that the model from the tune-up runs required some sort of calibration tasks to enhance the efficiency of the simulation outputs. What followed was the experimental setup of the environments created from initially calibrated ABM simulations that were done with four implemented experimental cases and are described as follows: i) in this experiment the only source of water input in the catchment is from the observed streamflow series and no rain input, ii) using a varying volume as the only source of water input in the catchment, without precipitation, iii) using the observed streamflow series and the precipitation as the only sources of water in the catchment, and iv) the observed streamflow series, the precipitation and an initial discharge value set to 1.5 $m^3 \cdot s^{-1}$. Besides these input scenarios, the tweaking of the GAML script for the modeling was also a very fundamental factor in the calibration endeavor.

4.4.5.3 GAMA: Model Calibration Results

The following subsection presents and discusses the verification results obtained from the calibration process performed using the four experimental approaches that were explained previously in the model calibration rationale above.

It is important to observe that flood hydrograph analysis has its settings in the peak flows, which is used in the assessment of flood events. Therefore, to objectively calibrate and verify an ABM hydrological prototype it is required some guidelines to simulate, calibrate, and analyze flood hydrographs simulation efficacy through validation. Thus, the accomplishments of the former requirements and their statistical metrics, reporting the simulation results of peak flows, their errors of spread to the measured data, the coefficient of similarity, and determination for each experiment simulation scenario executed are shown in Table 4.17, along with the Figures representing the model and measured flood hydrographs are presented below.

In calibration and validation of hydrologic modeling, the selection of the flood events is critical [404], just as it was discussed in section 4.3. Therefore, a storm is selected for calibration and validation, considering the effects, nature, and the characteristics of the sensitive variables used in the process such as the morphometric properties of the catchment and of the climate conditions that drive the rainfall-streamflow characteristics, as well as the data availability required in hydrograph analysis.

From the observed and simulated peak flows shown in Table 4.17 in Figures 4.30, 4.31, 4.32, and 4.33, the four scenarios reproduced, by all simulation outputs from the ABM model are shown to accurately reflect the shape of the observed hydrograph recorded at Station H3 for the storm period of study. However, for scenario 2, despite a reasonable result for the "r", daily flows were overestimated, and the simulated peak discharge was overestimated sextuple (1675 $m^3 \cdot s^{-1}$), while the other scenarios have seemed to have doubled the measured peak flow of 258.6 $m^3 \cdot s^{-1}$. In any case, the calibration tasks performed, have been shown to have greatly contributed to the model being improved; however, on average, an error of $Q_{pk}\%$ = ±81.3% overestimations of peaks were still observed. However, the Authors in [405, 406] suggested a criterion that the model performs well if ±50%≤Q_{pk}<±100%, then the model could be accepted as satisfactory whether the simulated value was less or larger than the measured peak flow. Nevertheless, the information shown in Table 4.17, suggests the ABM simulated peak flows were within 100% of the error criteria for simulated peak flows that are larger than the measured values. It is noticed also that the four simulation scenarios agreed with these criteria (Table 4.17). Of course, the lower this value, the lower the overall bias in the model. On average, the error of spread that was observed between the measured and the ABM simulated flows was 89.2 $m^3 \cdot s^{-1}$; however, must the size for this error was inputted by scenario 2. The effectiveness of the model's accuracy in predicting the flows was in rage [60, 65%]. The observed peak for this calibration storm occurred in November 24$^{\text{th}}$ at approximately 22:00 hours and the average simulated peak flow was 768.4 $m^3 \cdot s^{-1}$, keeping in mind the very high overestimation of the Q_{pk} by scenario 2. The mean correlation coefficient observed for this calibration storm period can be accepted as significantly correlated (r>0.7). In general, the calibration task showed to have contributed satisfactorily to the simulation between the ABM hydrologic model estimated and the measured flows at Medio River Station H3 according to the average results of the $Q_{pk}\% = \pm 81.3\%$, $R^2 = 0.63$, RMSE = 89.2 $m^3 \cdot s^{-1}$.

Table 4.17: Statistical metrics for the observed and simulated (ABM) flood hydrographs after calibration process.

Simulation Scenario	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	ABM Sim. Q_{pk} [$m^3 \cdot s^{-1}$]
1	0.78	0.60	42.3	80.0	465.5
2	0.80	0.65	230.2	84.6	1675.9
3	0.79	0.63	40.3	80.8	465.5
4	0.80	0.64	43.9	80.5	466.7

Measured $Q_{pk} = 258.6\ m^3/s$ on November 24$^{\text{th}}$ at 22:00

Simulated flood hydrographs of the

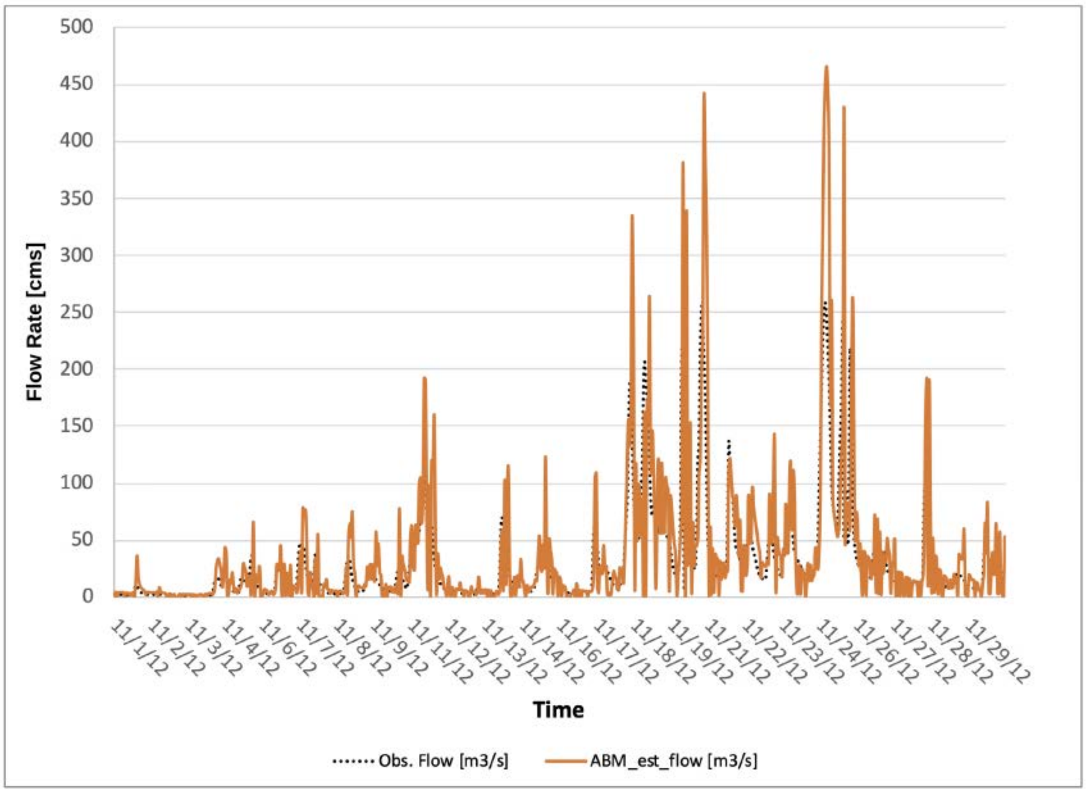


Figure 4.30: Scenario 1: Measured vs ABM simulation output for November 2012 storm after calibration process.

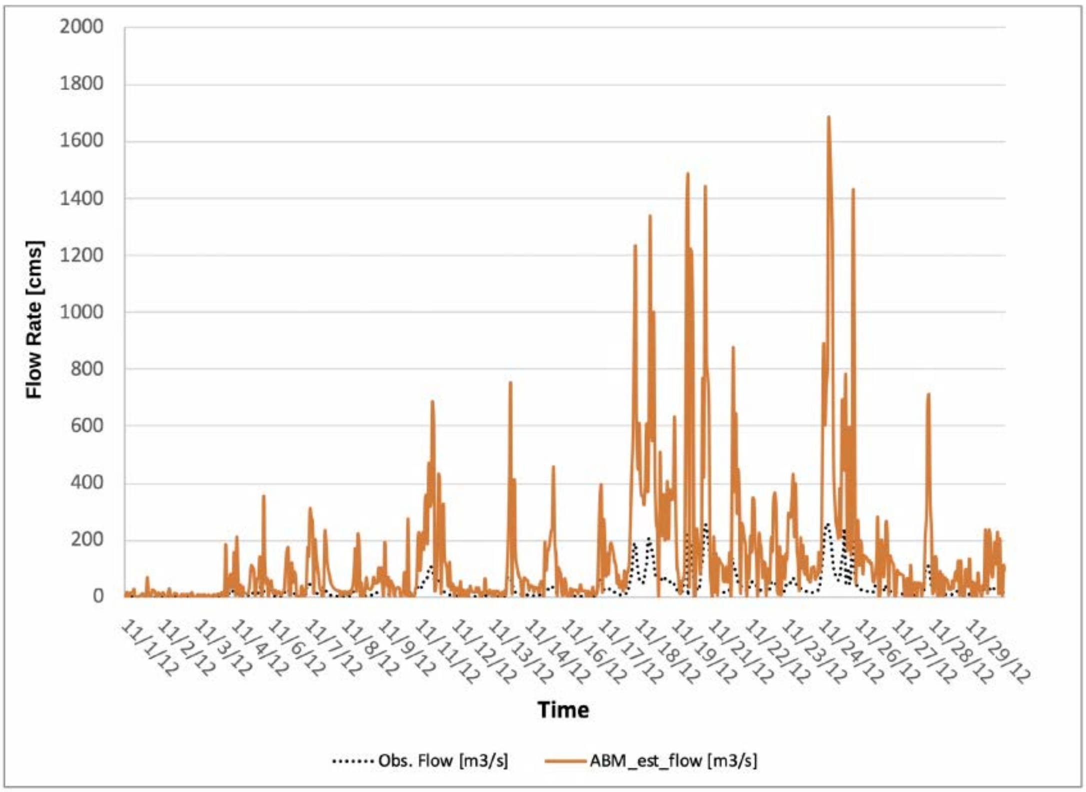


Figure 4.31: Scenario 2: Measured vs ABM simulation output for November 2012 storm after calibration process.

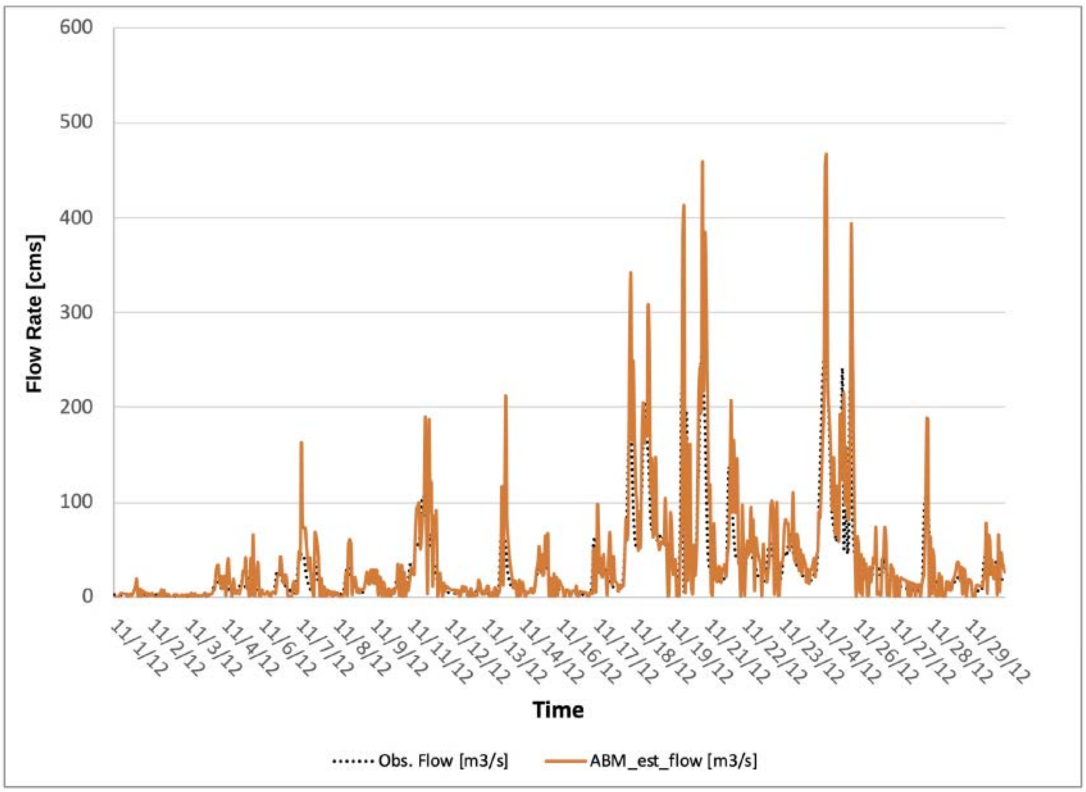


Figure 4.32: Scenario 3: Measured vs ABM simulation output for November 2012 storm after calibration process.

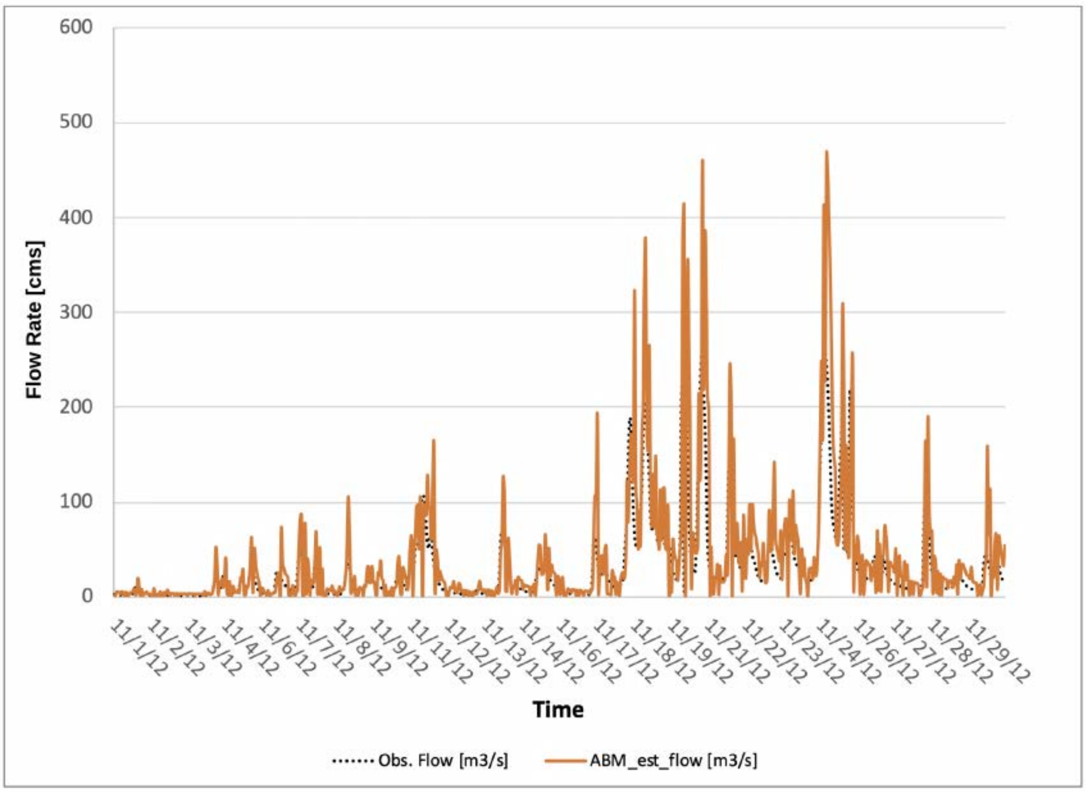


Figure 4.33: Scenario 4: Measured vs ABM simulation output for November 2012 storm after calibration process.

4.4.5.4 GAMA: Modeling Selected Storm Events with ABM Approach

Validation of the ABM hydrologic model were tested with simulations using the other selected storm periods as inputs, just as it is defined in Section 4.3. Technically speaking, this represents what is known as the validation process, since the implemented model is presented with new data and its performance can be finally evaluated to determine the overall goodness of the model. In this respect and for the sake of time, the data set used under Section 4.4.2.4 is used herein as input data. An inspection of the measured and the ABM simulated hydrograph outputs along with the statistical measures can be seen in Figures 4.34, 4.35, 4.36, and 4.37, and the performance metrics are listed in Table 4.18, respectively.

Four validation storms were chosen and used in this process. For example, the third case with the simulation of the May 2015 storm, although it marginally overestimated the observed peak flow produced satisfactory results with predicted peak discharge at the H3 Station about 329.3 $m^3 \cdot s^{-1}$; in comparison to the observed streamflow (266.4 $m^3 \cdot s^{-1}$) represented an error of the spread between the two and an error of the peak discharge of 20.1 $m^3 \cdot s^{-1}$ and 28.1%, respectively. However, the correlation coefficient (r = 0.76) was the lowest and the accuracy of the prediction accounted for only 58%. Additionally, the verification on all cases showed the errors were substantially lower than the calibration results, except for the $Q_{pk}\%$ for the November 2015 verification storm. Figures 4.34 to 4.37 shows the various events of the simulated ABM model versus observed storms used in the validation task. Table 4.18 lists the performance metrics on the remaining 3 storms addressed in the validation process.

Summarizing, results indicated all simulated validation storm hydrographs events by the ABM model to satisfactorily approximate the shape of the observed storm hydrographs. This implies, there is good agreement in obtaining acceptable performance of the ABM hydrologic model in forecasting the flows with time series of tropical watersheds. Complementarily, the experiments also showed that with further model supplementary adjustment, it can augment the correlation coefficient and gain substantial reduction of errors, it can be good implementation alternative to HEC-HMS or other standard hydrologic models.This assumption is based on the fact that it is generally believed that a good value for the correlation coefficient should be at least 70%. However, though this may be the desired case for any correlation analysis done on a certain dataset, it does not hold for several reasons, as Goodwin and Leech [407] indicated there are six reasons why a correlation size can be compromised:

1. Data with large variability
2. Distinct shape of the data distribution
3. Non-linearity

4. Outliers

5. Uniqueness of the sample characteristics

6. F-measures

Table 4.18: Statistical metrics performance between GAMA simulations and the observed flood hydrographs for various Storm Events.

Validation Storm	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	Obs. Q_{pk} [$m^3 \cdot s^{-1}$]	ABM Sim. Q_{pk} [$m^3 \cdot s^{-1}$]
December 2012	0.86	0.74	25.1	51.7	266.4	404.0
December 2014	0.79	0.62	41.7	88.7	335.5	633.3
May 2015	0.76	0.58	20.1	28.1	257.1	329.3
November 2015	0.90	0.82	23.0	80.0	264.3	475.7

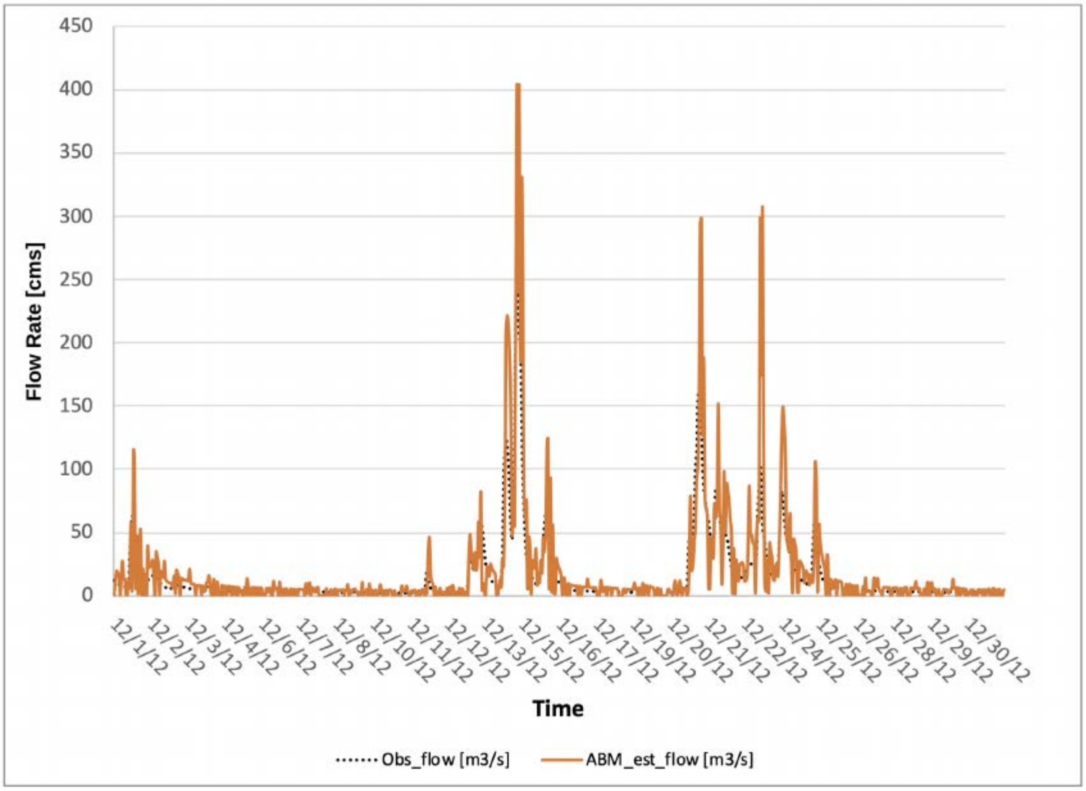


Figure 4.34: Observed vs ABM simulated hydrograph of selected December 2012 validation storm.

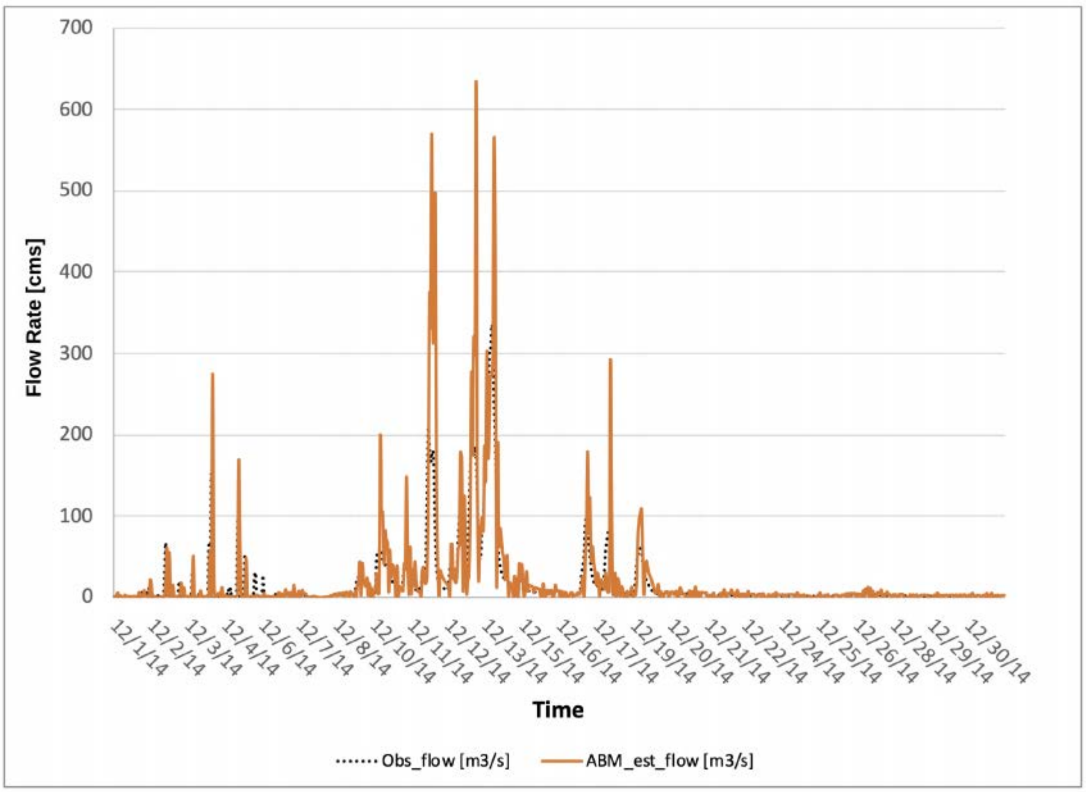


Figure 4.35: Observed vs ABM simulated hydrograph of selected December 2014 validation storm.

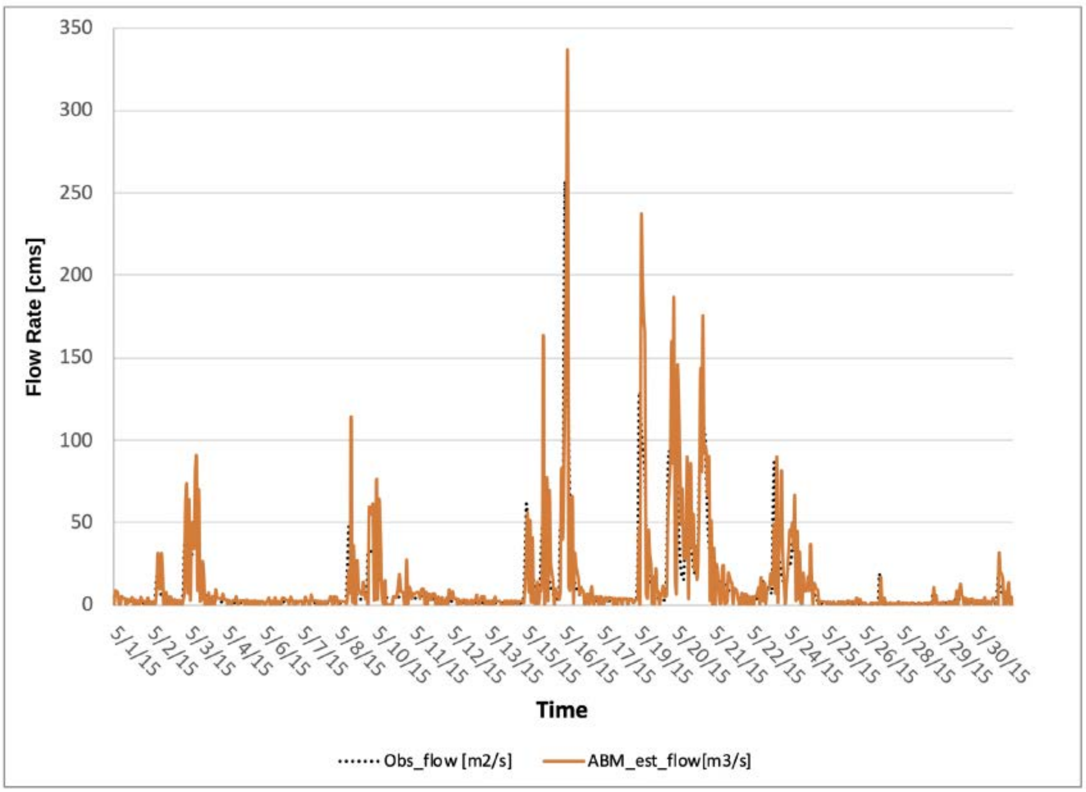


Figure 4.36: Observed vs ABM simulated hydrograph of selected May 2015 validation storm.

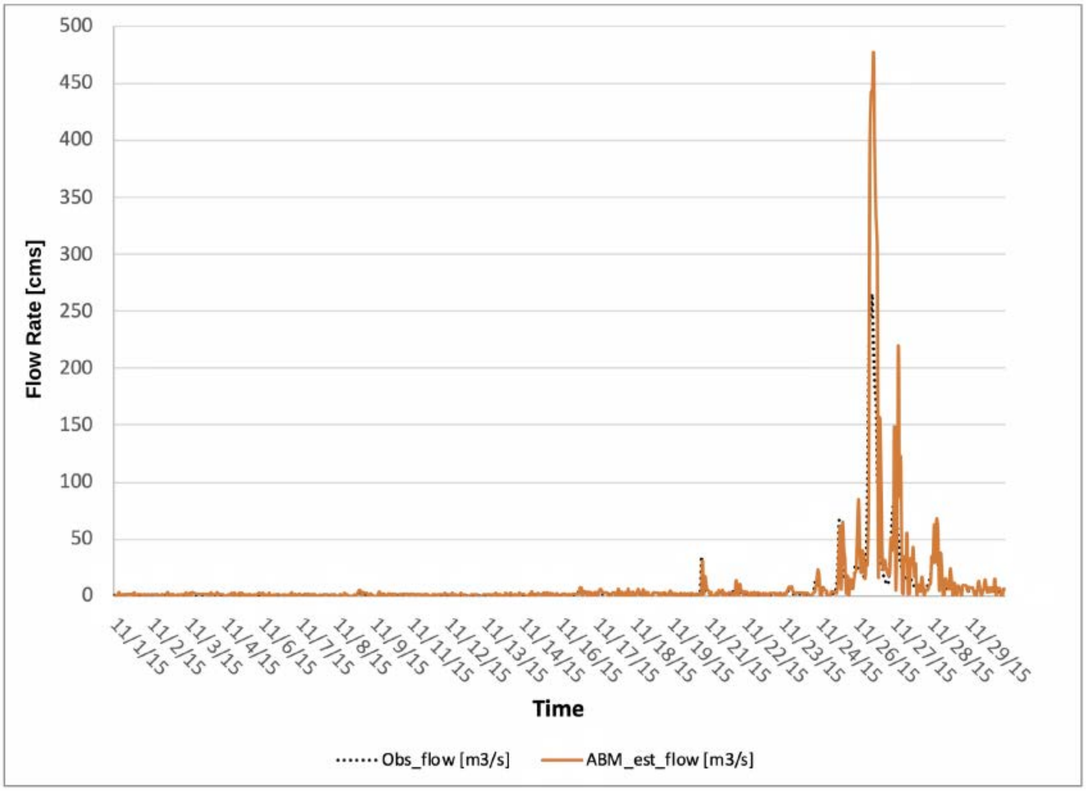


Figure 4.37: Observed vs ABM simulated hydrograph of selected November 2015 validation storm.

4.4.6 Conclusions

At this time, an initial agent-based modeling concept was built with which was realized flood forecasting with the Medio River catchment hydrometric time-series data from Station H3, have trained, calibrated, and validated the model on the observed case study flood storms. This first approximation serves as a benchmark scenario upon which is abstracted the conceptualization of river flood simulation in a humid catchment. Therefore, the key purpose of this benchmark simulation presented in this section constitutes the baseline on which to adapt the proposed intelligent multi-agent model for flood forecasting into the GAMA BDI Concept. Consequently, the simulation performed here would be compared to other simulation scheme outputs. Therefore, the following sections are dedicated to the task of upgrading the ABM modeling approach with BDI agents that can be equipped with DDM and AI skills, and in this manner seek to solve and maximize model capabilities for flood forecasting tasks.

4.5 Adapting A Flood Forecasting Model into GAMA BDI Concept

Up to this point, under the previous section, they simulated flood forecasting with the implemented ABM model, mostly with reflective agents; however, the proposal includes the implementation of the MAS model, through BDI architecture. In addition, it was emphasized in Section 3.3, the suggested mental and behavioral states of each of the agents represented across the multi-agent organization. Anyhow, it is acknowledged the importance to suffice how each of these agents would refer to one another, so it required the task to abstract the Belief-Desire-Intention (BDI) assumptions to the MAS architecture since there are no requirements in practice that can work as instructions for realizing agents that describes the instantaneous conjunction of complex hydrometric and geospatial data for the modeling of hydrologic behaviors in watercourses. From this opinion, this section is focused on the BDI model plan available in GAMA while seizing advantage of its agent-based and GIS capabilities to construct a resilient, responsive, and convenient BDI-driven MAS model for static hydrometric sensors. So, in this section and in the subsections that follow, it is explained the do about for provisioning the implementation, with intelligent agents in the MAS model with the BDI capabilities implemented in the GAMA platform as a solution for the problem domain.

4.5.1 The BDI Rationale and Hydrometric Sensors

As was stated earlier in Section 2.4.2, the development of MAS applied to problems related to hydrological engineering is a recent approach, and relatively scarce, and although there are indeed some applicable works in the areas of flood management, planning, hazards, and forecasting, most of the work done with flood forecasting in river basins with multi-agent systems and artificial intelligence is for non-tropical river basins.

In principle, they have applied the BDI rationale in the social, socio-environmental, environmental risk assessment, evacuation plans, and hazard mitigation studies. Therefore, from the specifications presented here in Chapter 3 and the arguments supported in Chapter 2, it is integrated the hydrologic domain with the BDI agent architecture, representing the rationale to manage and effectively use field-based hydrometric sensor data for flood forecasting.

In hydrologic monitoring, the role of hydrometric sensors is to monitor the dynamics of water resources. When there is precipitation as rainfall or snowmelt, the recorded hydrometric quantities support streamflow forecasting, evaluation of the quality of the water, and records the discharge or water level in rivers, stream, and flow channels and in this same manner, must be capable of the management of the hydrometric flow information.

Recalling what was presented in Section 3.3, the BDI architecture enables the agents in a multi-agent system to carry out certain roles by simulating naturally the management of this data acquisition and assessment with a certain abstraction.

4.5.2 Agents Knowledge Base

4.5.2.1 Declaring BDI Agents in GAMA

GAMA's reasoning engine (Figure 4.38) adapts to the BDI architecture and in this way allowing the programming of agents with cognitive capabilities using the GAML language. As a plugin, the BDI architecture, which is an implementation of the Behavior with Emotions and Norms (BEN) model introduced briefly in Section 3.2.3, is an architecture that together with the default reflexes, provides social agents with cognitive capabilities. The BEN architecture can be available during the definition of agents using the ***simple_bdi*** control.

Initialized a simulation period, an agent can use the architecture to decide on its next obligations. This means that each agent is instantiated with an action of this decision-making process. The BEN architecture configuration incorporates an arrangement of four components that are embedded within the belief and identity of the agent's cognition system.

The sections in blue, which are depicted in the diagram represent transformations that are systematically executed. If a modeler requires defining some processes manually, this is also possible, as shown in the sections shaded pink. The block in solid lines is processes controlled by the architecture, the block with dotted lines is optional. One advantage of such modular design is that it permits the modeler to make use of the items that apply to his/her domain of study and in this way reducing computational strain on the simulation.

The GAMA BDI engine enables the execution of the interaction "agent-system-environment". They implement this interaction in the GAMA BDI platform as specific behaviors, such as **Perceptions**, **Rules**, and **Plans**. Perceptions are behaviors that are executed at every iteration. According to to [311], perceptions provide agents with capabilities to perceive the environment and neighboring agents according to the events, and the information that comes into the system or that leaves the system.

To show the use of perceptions is provided an example using the GAML syntax. On this occasion, the sensor verification agent wants to perceive the status of a faulty sensor at a given hydrometric station location in a radius of 2000m and consequently updates its belief about the existence of a station in such proximity. Upon perceiving a sensor status, it stops the intention for verification, though keeping the intention to verify.

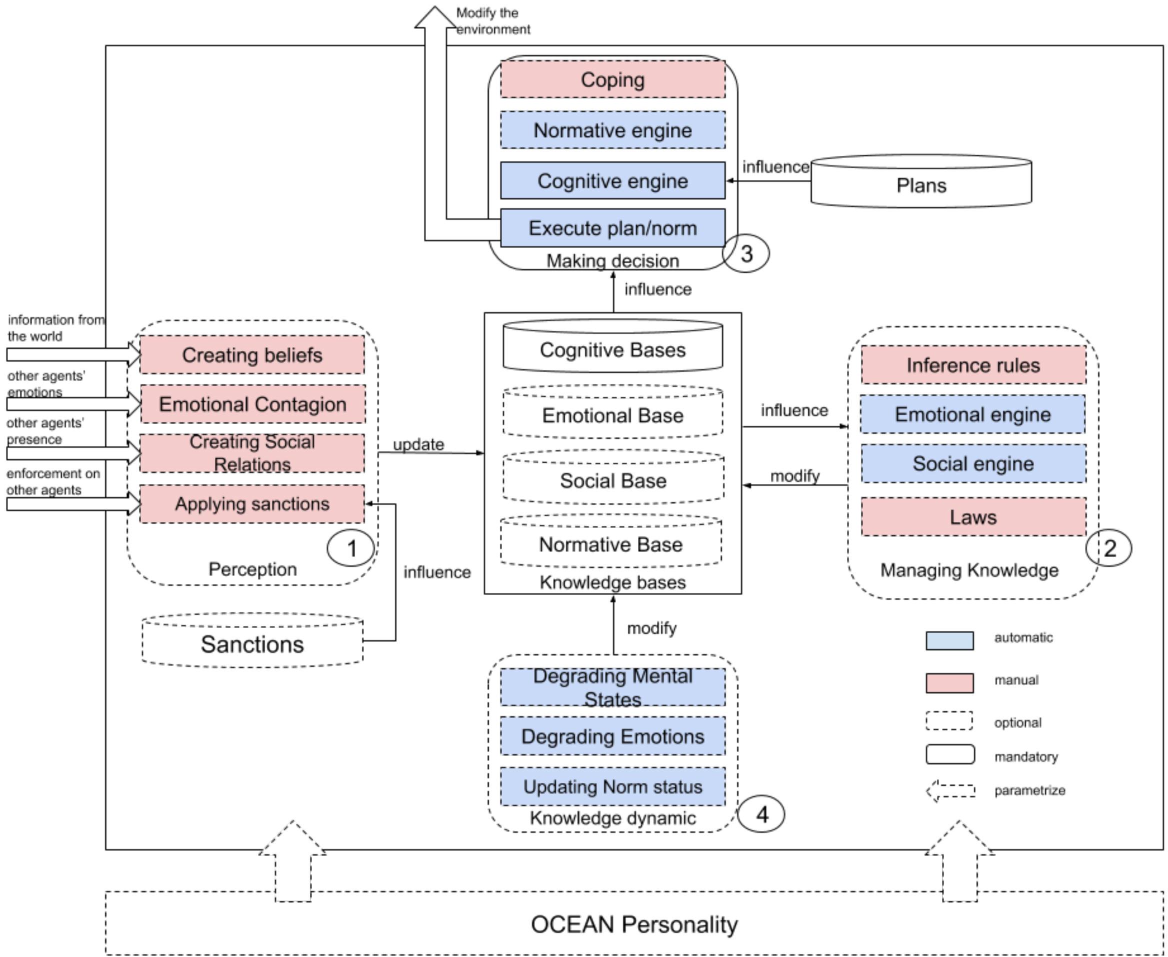


Figure 4.38: GAMA's abstract architecture and its reasoning engine. [Obtained from GAMA User's Guide, Ver 1.8.1].

perceive hydroSta target: subbasin in: hms_siting {
focus id: "location_hms" var: location strength: 10.0;
ask myself {**do remove_intention**(`verify_sensor_status_desire`, true);}}

In GAMA, the reasoning engine used to select the next plan to be executed by each of the agents is based on rules that are a specific behavior of the GAMA BDI engine. With the rule engine, agents can create desires from beliefs at each iteration. In the following illustration the idea that an agent has no connection to the hydrodatabase, will be automatically added to its desire base to get connected.

rule belief: new_predicate (`"no_DBConnection"`) **new_desire:** `DBConnection`;

Plans in GAMA are defined by the modeler. As explained previously, plans are also behaviors that are used by the GAMA reasoning engine and that are executed in reply to an intention.

The actions taken by an agent in response to his plan base are dependent upon the information he has of a specific state of the environment; therefore, such a plan should be deliberate to the agent when deciding. This capability facilitates several plans to answer the same intention, while enabled from different circumstances.

Below is an example that illustrates the definition of a plan to verify the intention of knowing the status of a sensor.

plan verifyingSensor intention: verify_sensor_status_desire {
ask AgentRNSn {**do add_desire**(`read_rain_sensor_data`);}}

The sequence in which each block is enabled is facilitated by the BDI engine plugin features, as shown in Figure 4.39. The following subchapters describe in detail how each process in each block is performed.

4.5.2.2 The Agent's Beliefs

The agent's belief stored within their knowledge base or belief system characterizes the information that each agent species holds internally. As presented earlier in Section 3.3, this knowledge base system represents beliefs about the information they have gathered from the sensors regarding the external environment in which they are present, localized, of the fixed information, and information they have concerning themselves and other agents in their surroundings via messages.

In GAMA, a set of ***predicates*** that expresses the internal awareness of the agent and its environment defines the belief base of an agent. This belief base forms part of the BDI agent's memory, it allows the agent to update its belief system during a simulation by adding or dropping a belief. This feasibility provides the agent with the capabilities of updating their knowledge base. Once an agent in GAMA is defined with the control simple_bdi, all other agents that compose this species, achieve new understanding and behaviors. In this sense, the agent knowledge base comprises the Belief-Desire-Intention rationale, and the new behavioral patterns comprise the perceptions, rules, and plans. This mechanism has the feasibility to be updated in rules, perceptions, reflexes, and plan statements.

An example of how the predicate data type function in GAMA can be represented by

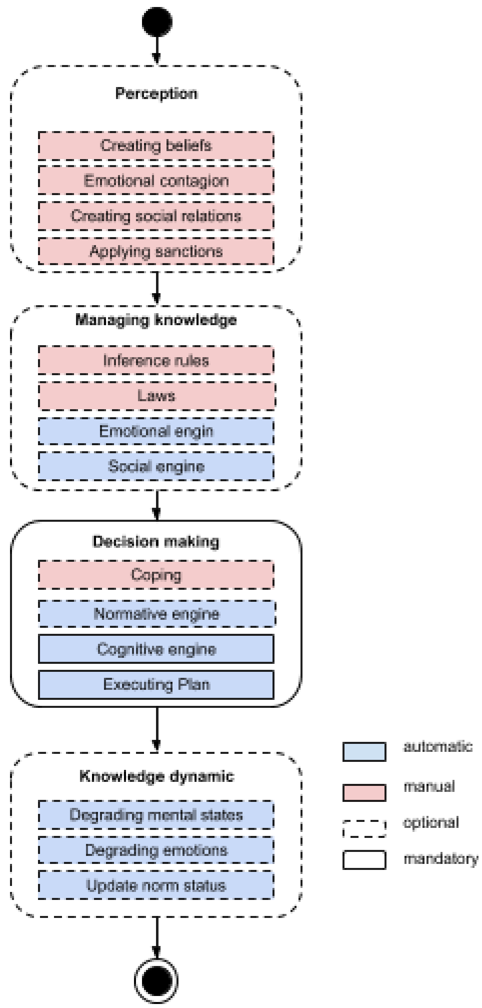


Figure 4.39: GAMA's abstract architecture showing the order of execution for each block in the reasoning engine. [Obtained from GAMA UserGuide, Ver 1.8.1].

the agent AgentSV. This agent can either add or update knowledge to its belief system by using the predicate data type with the following examples of the GAML language syntax:

rnSensor__predicate

if the agent is aware of the rain sensor. Once this belief has been added, it will be stored in its belief system as:

do add__belief(`verify_sensor_status`);

or otherwise removed from the belief base with the following syntax:

do remove_belief(verify_sensor_status);

4.5.2.3 The Agent's Desires

In GAMA, agent desires are implemented as objectives (e.g., Goals, Aims). In this respect, an agent's aim to achieve a goal would be defined by his actual state. Desires contain predicates that can be planned with the following GAML syntax:

do add_desire(find_user_registration);

There are two conditions in which desires are met within GAMA, the first condition is when an agent believes an objective to be true, whereas in the second condition a desire or intention can be removed manually from its knowledge base. For example, in the following syntax, a desire is added or removed.

do add_belief(`rnSensor_predicate`);

do remove_desire(`no_rnSensor_predicate`);

The GAMA BDI architecture allows concurrency between desires in an ordered manner, so desires can have sub-desires.

With sub-desires, agents can designate intermediate objectives. For example, when the sensor verification agent has an objective of checking a specific hydrometric sensor (i.e., water level sensor), but he does not have a connection to that sensor, he can ask the respective water level sensor agent to confirm if the sensor is in operation or not. In such a situation, to fulfill its primary objective, the agent can add an intermittent desire that it will try to accomplish first.

To prioritize a desire, it can be assigned a priority value to desires so that the agent can select a new intention from its desire base. The priority of a desire depends on the priority of the related desire.

4.5.2.4 The Agent's Intentions

Within the BDI model and particularly in GAMA, to achieve its goals an agent executes plans, which comprise GAML code. In GAMA, the agent's intentions are selected from its desires. If an agent has several desires, when he decides on choosing one of his desires, that new desire (intension) becomes its new priority and is assigned a high priority value. Therefore, this new intention will define which plan(s) should be realized. However, to avoid conflicting priorities, should there be any, GAMA allows the use of a Boolean parameter for defining whether the priority is deterministic or probabilistic by simply adjusting this value. This functionality gives way to a hierarchical arrangement of the plans.

As with sub-desires, plans in GAMA are not just a sequence of executed actions, but an agent can have sub-intentions that can be added to his intention base. When adding new sub-intentions, the agent would halt his actual intention, which will be put on a hold status, allowing the execution of his new intentions to proceed. An illustration of this action is shown in the piece of GAML syntax below.

do add_subintention(no_qSensor_predicate),

subintentions: ask myself {**do remove_intention**(`verify_sensor_status_desire`, true)});

do `current_intention_on_hold()`;

4.5.2.5 Agent's Use-Case Diagram Graphical Depiction

Figure 4.40, shows the "UML use case diagram" for the MAS organization of 17 members, head node of HSn (Hydrometric Sensor Agent) cluster, 8 nodes of learning agent (e.g., Agents Forecasters 1 to 8), and decision-making node (AgentFL). The secondary actors are represented by the agent's nodes, AgentSV, AgentDPP, agentData2Lags, AgentHDBM, and AgentUI respectively. Every node has its functionality with some different functions: participant node and cluster head-node, and under certain circumstances may share the same functionality, like, for example, the database connectivity. Head-nodes will capture sensor data, generate information on the data and about sensors functionality, send reports, and they convey workable paths to send the data. Participant nodes will retrieve data from different cluster head-nodes, to pre-process, store locally and into the database, and aggregate that data to send to the AgentFL node. Participant nodes will also check for messages regarding data access, sensors functionality, forecasts, and warnings.

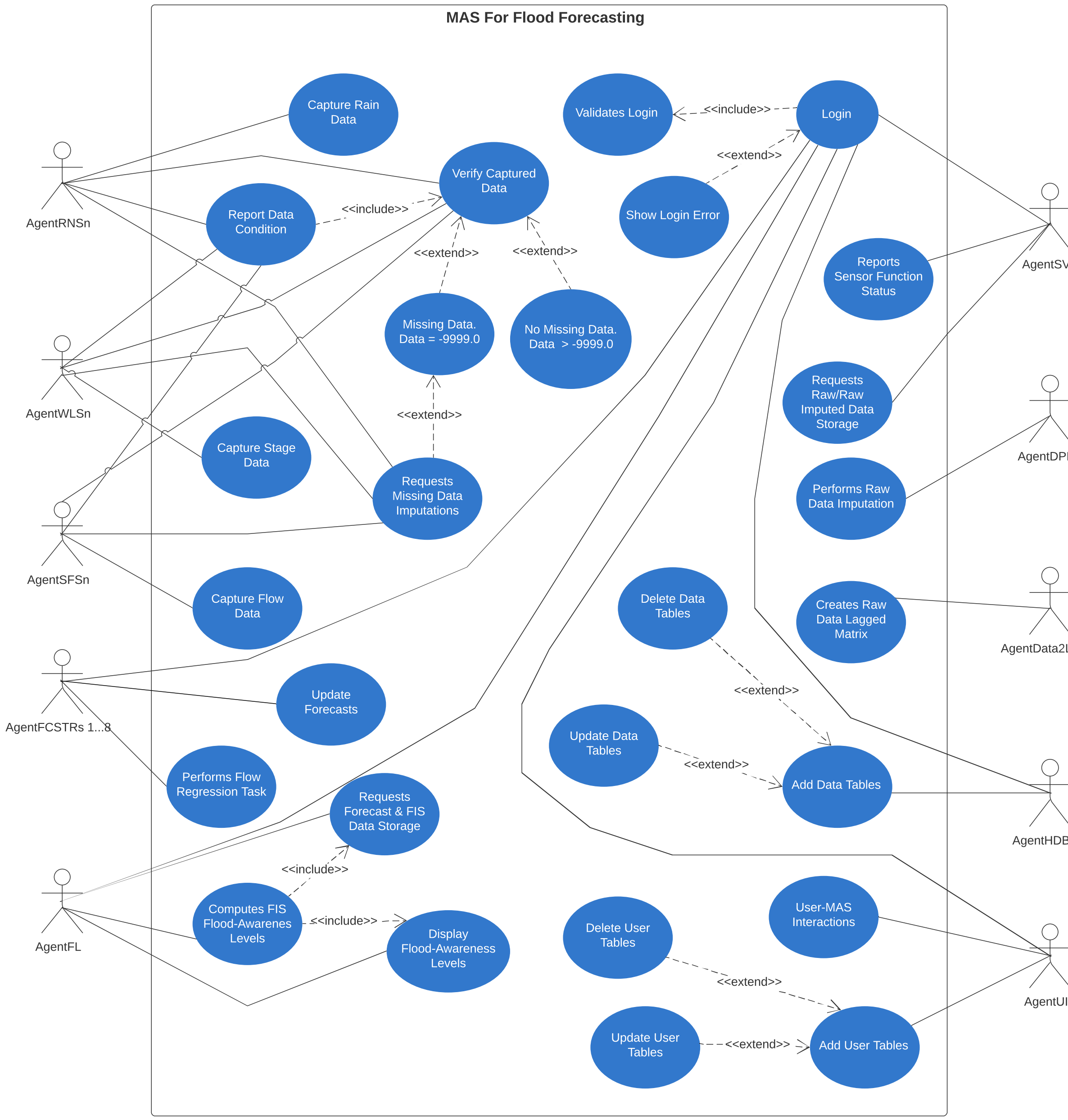


Figure 4.40: Use-Case diagram representing each Agent's Role.

4.6 Conclusions and Future Perspectives

Contrary to the modeling engine in HEC-HMS, as a hydrologic numerical model where a series of compartmental models exists specific for adding components for meteorology, basin, control specifications and numeric data (e.g., precipitation, flows, and evapotranspiration) as seen in earlier chapters; in GAMA, these components don't exist implicitly in its simulation engine, as GAMA is not in the strict sense of the word a hydrologic model per se. However, aside from allowing adding data as inputs, it facilitates the introduction of GIS data that can be agentized within a modeling implementation and initial conditions are possible. Although GAMA and HEC-HMS may have some of the same similar input data requirements, there are still some main significant differences between the two, especially when the purpose is hydrologic modeling implementation. As a fact, both applications require the catchment's daily hydrometric data (e.g., rainfall, streamflow). The computational modeling structure in both models is significantly different. There are several algorithms selected for surface runoff in HEC-HMS such as the SCS UH method, which was shown earlier that are not implemented in GAMA. Other examples are the Snyder and the Clark unit hydrograph used also for hyetograph transformation methods. For river routing the Cunge or the Muskingum-Cunge methods (that are established on conservative, massive, and diffusion-momentum laws) constitute the hydrodynamic and physics-based mathematical equations not necessarily present in the GAMA simulation engine but can be implemented with proper coding. For example, in the catchment modeling platform "Soil and Water Assessment Tool (SWAT)" the "volume" of water is routed as flow (discharge) through the channels, the GAMA engine similarly routes water volumes as flow through the river reach elements, and as such this assumes that the main channels, or reaches, are trapezoidal as defined in [380]. However, given the uncertainties and the need for a large number of parameters required by the SWAT routing engine, the authors in [408–416] have reported several issues with the seasonal underestimations and overestimation of streamflow routing in SWAT, similarly, this seems to be the case with the simulation scenarios in GAMA. Nevertheless, although GAMA may employ the SWAT routing routines, for routing flow, despite these issues, the Medio River catchment, like most of the catchments in Panama, is ungauged; therefore, hydrometric data is an issue of scarcity and data impairment. For the Medio River catchment, only Station H3 has had a complete dataset, although, with some data quality issues, it was the station with data that could be used for this approach. Station H4 dataset as it was mentioned earlier, the data record was severely impaired; besides, these stations were installed for an environmental assessment baseline. The status of the current dataset assumes a dedicated work of data repair and enhancement had to be undertaken. This is probably another motive for the dissimilarity concerning modeled and measured flood peaks, as even the standard hydrological model such as HECHMS has had difficulty estimating flood peaks. Notwithstanding, both models achieved acceptable values of the assessment metrics. Thus, this hydrologic modeling approximation task was not done to compare the GAMA modeling platform with the "HEC-HMS", but to present the potentials of the agent paradigm as an approach to undertake with few modeling complications setups required of the standard hydrologic models for streamflow forecasting, to provide a tool for the water community in the evaluation and oversee of flood hazards in tropical watersheds. For these reasons, the present work outlook is to look forward to optimizing the ABM hydrologic model implemented in the GAMA platform for flow estimation codes and capabilities already at hand and obtain, if possible, although the results are not compromising an increase of the actual performance statistics used in the model assessment (Table 4.17).

Chapter 5

MAS for Flood Forecasting

In Chapter 4 it was performed streamflow forecasting within the Medio catchment domain using conventional hydrologic modeling with the HEC-HMS, and a non-conventional approach with an ABM model implemented on the GAMA platform, and the results from the flood hydrographs simulation were compared and reported. Moreover, this chapter constitutes the basis of the proposed MAS hydrologic modeling framework for streamflow prediction within the tropical watersheds in reference to the ontological settings discussed previously under Chapter 4, in Section 4.1.3.

So, under this chapter, it is outlined the "BDI-architecture model" behavior for each of the agent's present in this MAS hydrologic modeling framework: with hydrometric sensor's agents in the administration of the hydrometric data, the preprocessing of the hydrometric data, data storage, the making of forecasts, and the delivery of inferences concerning flood-awareness levels from a MAS perspective.

To recreate the hydrologic modeling with the proposed MAS model framework an arrangement of tests was formulated built on the previous agent-based model framework which was upgraded with the addition of BDI supportive agent's species and "Machine Learning (ML)" algorithms (Section 3.3.1). Therefore, the rationale of the series of experiments that were executed in this chapter takes on as inputs to the MAS the information derived from the ABM simulation outputs.

5.1 Climatic Events and Hydro-Agents Interaction

In this subsection, the ideas and experiments performed are for coalescing the multi-agent concept approach through agent-based concepts to simulate the flood hydrographs with forecasting and decision-making agents regarding flood-awareness levels in the Medio River catchment using the "`simple_bdi`" control feature that makes up the BDI architecture capabilities already integrated into GAMA.

Following the designated roles (Section 3.3.1) assigned to each of the hydrometric agents comprising the hydrometric station network, it is built the corresponding multi-agent system model, and through which is represented the hydrometric sensors and other components of the hydrometric monitoring network as agents having behaviors based on the BDI format.

The case study dataset, represented by the ABM simulation outputs (see Sections 4.4.5.3 and 4.4.5.4) along with a description of the source and metadata used to specify the behavior of the agents comprising the MAS were taken from the information formerly described in Section 4.1.

5.1.1 BDI-driven Conceptual Model for Flood Forecasting

When building a scalable system, it should be noted that such design requires using components that search for means of solving the problems autonomously [398, 417] and which are self-organized for addressing such issues of monitoring and control [418]. However, as a reminder, these design requirements do not imply the management of cooperation in the multi-agent concept [419].

A flood event is a dynamic system, in such an active system, agents with different behavior and goals interact mutually, and as their objective is for each to attain their respective goals or set of goals, the realization of such can affect one of the other.

Again, the main purpose of a hydrometric sensor network is to monitor hydrologic quantities in catchments and deliver accurate information on what is occurring in such environments. Therefore, to carry out such a monitoring task, especially in vast watersheds, a single monitoring station or a sensor is not sufficient; Thus, it is required to include a network with complete instrumentation for monitoring station (i.e., small catchments).

A description of the abstraction of the BDI reasoning conceptual model for each agent is shown in Figure 5.1. It presents the UML class diagrams for the agents in the flood forecasting model and their specifications using the GAMA built-in BDI framework, these agents along with the representation of the hydrometric station(s) deployed in the Medio catchment are defined with their respective fields, actions, predicates, set of beliefs, desire, intentions, and plans as defined by the GAML language syntax used in the GAMA platform.

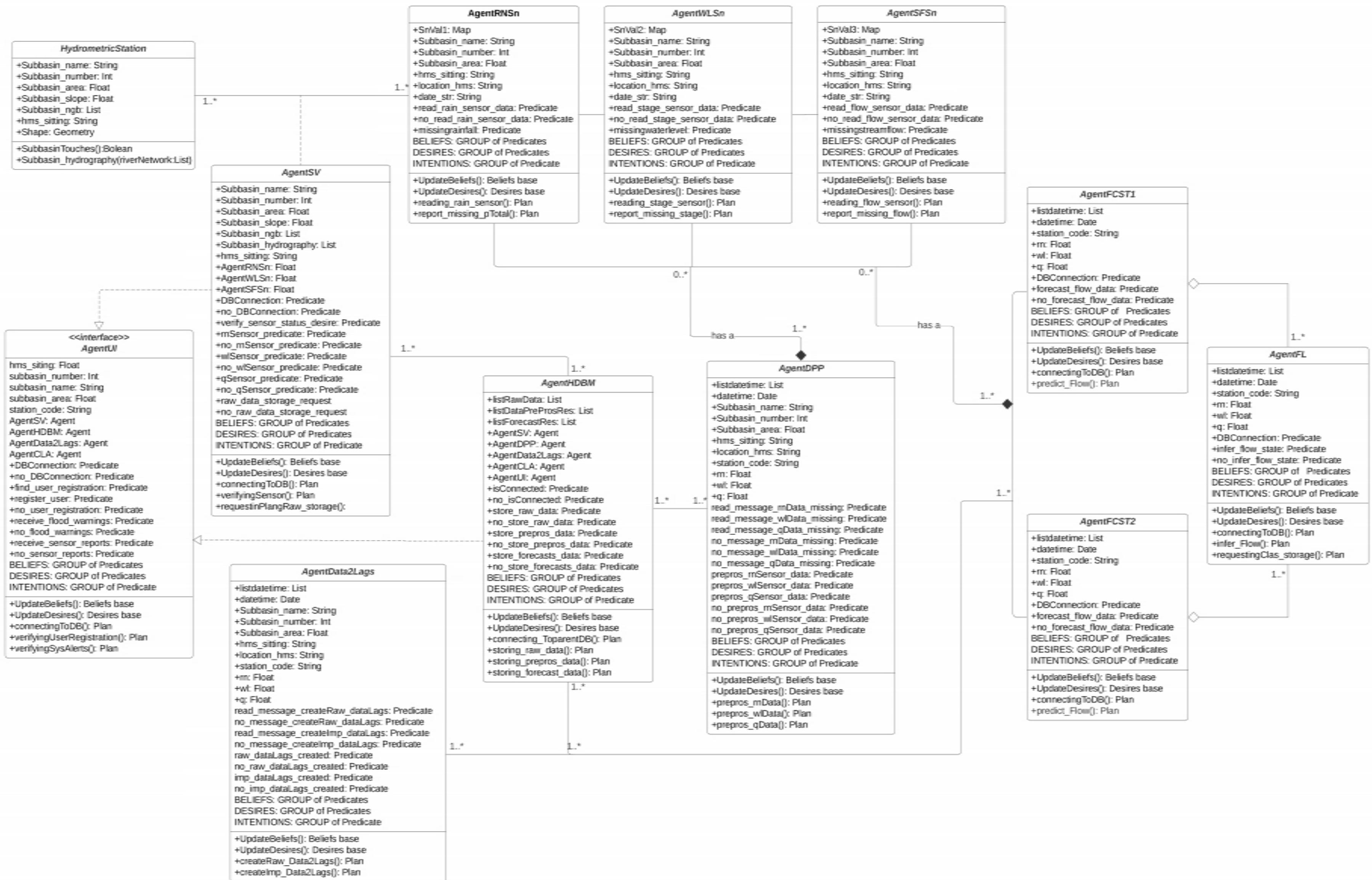


Figure 5.1: General overview of different agent classes used in the BDI Medio river flood forecasting model and their relationships. The world agent (non-bdi) is the hydrometric station located in the catchment and is only used in the simulation, it represents the physical world and the sensors network.

5.1.2 The Agents Belief Knowledge Base

As shown in Figure 4.38, for each of the agents, the belief knowledge base is updated by the "`update_beliefs()`" function, which is a reflex activated immediately upon initialization of the agent and after each simulation time step. The belief or set of beliefs assigned at initialization for an agent are executed based on parameter attributes that are necessary to trigger the agent's belief about its environment or its relationship to a cooperating agent. So, an agent belief represents the knowledge of their capabilities, the knowledge of the capabilities of the neighboring agents with whom they interact, and the knowledge of the information of the environment (i.e., the catchment) collected from the hydrometric network instrumentation. In this context, the parameters of the belief for each agent are defined, along with their set of fixed state terms required to allow the activation of the beliefs, if the state is true, as summarized in Table 5.1. For instance, with the case of the sensor verification agent, the first belief is triggered when he validates the need to check if indeed there is a connection to the hydrodatabase. The second is his general belief to check for the functionality of the sensors in the hydrometric network based on perceptions regarding the information on precipitation, the river water level, and the streamflow collected by the hydrometric sensor agents. Last is the belief in the need to have this incoming data to be stored in the hydrodatabase. For the hydrometric sensor agents (i.e., AgentRNSn, AgentWLSn, AgentSFSn), they define two beliefs, in the first belief, the sensor agents should read the sensors' incoming data, and secondly, that there are no missing instances in the readings. The data preprocessing agents, (i.e., AgentDPP and AgentData2Lags) as the beliefs concerning messages about data that require preprocessing, be it the imputation of missing instances and/or the creation of lagged data matrices. The beliefs of the hydrodatabase manager agent at initialization time, are four, being the first its perception of its uninterrupted connection to the hydrodatabase, and the other three corresponds to the notions that there are for example raw, preprocessed forecast, and warning flags that need to be imported into the database. The classifier agents, which includes the eight agents with machine learning capabilities for performing lead-time regressions on flows and the decision agent responsible for computing the flood-awareness levels from the actual and forecasted results, all believe that they are connected to the database and that there is either raw and/or imputed raw data that has been previously pre-processed, by either of the two data preprocessing agents if there was the need to. Finally, the user interface agent, as five beliefs, being the first that it is connected to the database, the second and third that there is a request for a user subscription and to register the user, and the other two are beliefs that are instantiated if there are reports on sensor status and flood awareness levels from the sensor verification and/or decision-making agent.

Table 5.1: Hydrometric Network Agents Beliefs base and states triggered at initialization.

Agent	Belief Variable	Fixed State
SV	DBConnection	Access to the hydrometric database
	verify_sensor_status	Hydrometric sensors are functioning
	raw_data_storage	Requests hydrometric data for storage
RNSn	read_rain_sensor_data	Read incoming rain data
	missingrainfall	If such is the case
WLSn	read_waterlev_sensor_data	Read incoming stage data
	missingwaterlevel	If such is the case
SFSn	read_flow_sensor_data	Read incoming flow data
	missingstreamflow	If such is the case
DPP Data2Lags	read_message_rnData_missing	Incoming message if rain data is missing
	read_message_wlData_missing	Incoming message if stage data is missing
	read_message_qData_missing	Incoming message if flow data is missing
	read_message_createRaw_dataLags.	Incoming message create lags from raw
	read_message_createImp_dataLags.	Incoming message create lags from imp
HDBM	isConnected	Access to the hydrometric database
	store_raw_data	If a request is issued
	store_prepros_data	If a request is issued
	store_forecasts_data	If a request is issued
FCST1	DBConnection	Access to the hydrometric database
FCST2	forecast_flow_data	Available hydrometric incoming data
FL	infer_flow_state	Access to forecasts
UI	DBConnection	Access to the hydrometric database
	find_user_registration	User request to register
	register_user	If a request is issued
	receive_flood_warnings	If such is the case
	receive_sensor_reports	If such is the case

5.1.3 The Agents Desire Knowledge Base

In the BDI agent model, what they define as desires are the functions that capture the agents' desires. Within this context, for each of the agents in this MAS organization, targets (goals) represent the desires to accomplish, and which are predicated on its beliefs. In GAMA, desires can be prioritized in the sense as to allow an agent to choose between an intention using the function "`with_priority()`". Based on the system domain, each agent has their desires; however, some may share, for example, the same set of desires, as with the desire of having a connection to the hydrodatabase. As it was for the case with the belief knowledge base, for each agent the function call "`update_desires()`" updates the desire knowledge base. In Table 5.2 it can be found for each agent defined their respective desires.

Table 5.2: Hydrometric Network Agents Beliefs-Desires relationships.

Agent	Beliefs	Desires/Intentions	Plans
SV	no_DBConnection no_rnSensor no_wlSensor no_qSensor no_raw_data_storage	DBConnection verify_sensor_status raw_data_storage_request	connectingToDB() verifyingSensor() requestingRaw_storage()
RNSn WLSn SFSn	no_read_rain_sensor_data no_read_stage_sensor_data no_read_flow_sensor_data	read_rain_sensor_data read_waterlev_sensor_data read_flow_sensor_data	reading_rain_sensor() report_missing_pTotal() reading_stage_sensor() report_missing_stage() reading_flow_sensor()
DPP Data2Lags	no_message_rnData_missing no_message_wlData_missing no_message_qData_missing prepros_rnSensor_data prepros_wlSensor_data prepros_qSensor_data no_prepros_rnSensor_data no_prepros_wlSensor_data no_prepros_qSensor_data no_message_createRaw_dataLags no_message_createImp_dataLags no_raw_dataLags_created no_imp_dataLags_created	rain_data_for_dpp waterlev_data_for_dpp flow_data_for_dpp raw_data_to_lag Imp_data_to_lag	prepros_rnData() prepros_wlData() prepros_qData() createRaw_Data2Lags() createImp_Data2Lags()
HDBM	no_isConnected no_store_raw_data no_store_prepros_data no_srore_forecasts_data	test_parentDB_connection store_all_hydro_data	connecting_ToparentDB() storing_raw_data() storing_prepros_data() storing_forecast_data()
FCST1 FCST2 FL	no_DBConnection no_forecast_flow_data no_infer_flow_state	DBConnection do_flow_predict do_flow_inference clas_data_storage_request	connectingToDB() predict_Flow() infer_Flow() requestingClas_storage()
UI	no_DBConnection no_user_registration no_flood_warnings no_sensor_reports	DBConnection find_user_registration view_system_msg	connectingToDB() verifyingUserRegistration() verifyingSysAlerts()

In the following paragraphs, the desires of each of the agents in the MAS are defined.

- **AgentSV**. The sensor verification agent desires to check periodically if it is connected to the hydrodatabase, it is his initial desire, and it should be noted that this desire is also shared by other agents in the system (e.g., classifier agents, and the user interface). This agent also has two other desires which are to continuously verify the status of the sensors and to request the hydrometric database management agent to store the raw data. At initialization time, the first two desires are added to the knowledge base when the agent believes that it does not have a connection to the database and/or that a sensor in the network may be down (out of service). The GAML syntax which defines each of these desires and their corresponding intention is shown in Table 5.2 above.

- **Agents RNSn, WLSn, and SFSn**. The desire of all hydrometric sensor agents is the continuous capturing and reading of the real-time data in their environment, that is rainfall, water level, and the river flow in the catchment. This is their general desire base, and it is their initial state activated at simulation time. If any believe that it has missing cases, they add a temporary desire to request an estimation on the instance(s) be undertaken by the data preprocessing agent.

- **AgentDPP, and AgentData2Lags**. The data preprocessing agent's (AgentDPP) desire to perform approximation estimates on the values for the missing data regarding the information received about the hydrometric data by all the sensor agents, providing the best possible estimate given the estimates computed on each or any of the hydrometric quantities of the rainfall, water level, and flow data. Besides, the AgentData2Lags, continuously as the role to transform either the complete or incomplete imputed raw data into a lagged matrix form, suitable for supervised ML tasks.

- **AgentHDBM**. The "`test_parentDB_connection`", desire relates exclusively to this agent who believes that his connection to the hydrodatabase must be permanently uninterrupted, while other agents have only a temporary connection to the database whenever there is a need for a connection to perform a specific task, can lose the connection afterward. In the beginning, the first general desire of the hydrodatabase management agent is to check his connection, whereas his second desire is the storing of the information generated from the processes carried out in the MAS.

- **Agents FCST1, FCST2 and FL**. The first two agents are a group of eight machine learning agents hosted under the terminology of Agent forecasters 1 and 2 for convenience, and the latter is the fuzzy logic skills implemented agent. Their initial general desire, as in the case of both forecaster agent's groups, is to perform forecast computations of the river flow, whereas the decision-maker agent desires to make inferences about flood-awareness levels based on the actual hydrometric data stream and of receiving flow forecast results data from the forecaster agents. In addition, the second desire in the agents belief-base is the awareness of database connectivity (`DBConnection`), which ensures they can have access to the hydrometric database. At least one agent has the capability for requesting that the classification and inference results be stored in the hydrodatabase.

- **AgentUI**. This agent has three desires, which are added to his desire base engine at the beginning of the simulation; the first desire is to verify his database connection. This is followed by his desire to register, delete, or update user data, and finally, the exchange of information with the decision-making agent, to the user about the flood-awareness level status alarms, and the sensor verification agent concerning reports on the functionality of the hydrometric sensors.

5.1.4 The Agents Intention Knowledge Base

Intensions are the last feature found in the BDI model; these are defined as the procedures that an agent carries out to accomplish their desires during several simulation time steps if its past condition be achieved or the intention in question is dropped from the intention base. In GAMA, the actual intention of an agent determines which plan would be selected. This view of the intentions is shown by the existing relationship among the features of the BDI model as was defined previously in Table 5.2. An idea of this feature can be illustrated from the following scenario with the sensor verification agent who when having the intention to issue a request for data storage (`raw_data_storage_request`) from the hydrodatabase management agent, can't get the request to proceed (`no_raw_data_storage`), notices that it is because it does not have an actual connection to the database (`no_DBConnection`), which would activate its database connection plan (`connectingToDB()`), and therefore, drops the request for raw data storage (`requestingRaw_storage()`) plan.

5.1.5 The Agents Plans

In GAMA, the reasoning engine used to select the next plan to be executed by each of the agents is based on rules, as presented previously in Section 3.4.2.1. Therefore, to accomplish their intentions, the agents in the hydrometric system need to execute various plans accompanying their motivations as specified in Table 5.2. These plans represent behaviors put into effect in any frame of reference in reaction to the agent's desire. In the BEN architecture, a plan pertaining to an agent i is denoted by the function "$Pl_i(Int, Cont, Pr, B)$".

where Pl_i refers to the name of the plan, Int the intention that activated the plan, $Cont$ is the perspective in which the plan is applicable, Pr implies a priority value to help the agent to choose amongst various plans, and B relates to the behavior performed by the agent if it chose a particular plan.

For the agents to accomplish their intentions, they need to commit to certain plans that are related to their intentions, as described in Table 5.2. Hence, for each agent in the hydrometric organization, there is a set of plans by which they must abide, these are breakdown into a sequence of several tasks as described in the activity diagrams shown below in the following passages:

- **Plans for Agent's RNSn, WLSn, SFSn, SV, DPP, and Data2Lags.** As shown in the activity diagram in Figure 5.2, the intentions of capturing and generating data products (e.g., imputed raw data, creation of lagged data matrix), and ensuring sensor performance are all activities executed on an hourly basis in parallel by the agents of Levels I and II of the MAS organization as described in Section 3.2.1. Each member has the intention of messaging information before sending data to other segments of the organization. As depicted in Figure 5.2, the HSn agents, being primary actors have different functionalities than the secondary agents as defined in the use-case diagram in Section 3.4.2.5. As cluster head-nodes, the HSn agents have the intention to collect the incoming data from the field hydrometric sensors, each by following their respective plan of `reading_rain_sensor()`, `reading_stage_sensor()` and `reading_flow_sensor()` data, checking for missing instances and quality of the data (e.g., instances labeled as -9999.0), and if they identify an instance or instances of the data with readings of -9999.0 as true or false, they drop their intention to read the incoming data and take on their intention to report about the conditions of the sensor and the collected data to the SV agent, who will inform the UI agent about the actual situation. The HSn agents, in case of sensor data missingness, will trigger the intention "`report_missing()`" to report missing instances of any of the hydrometric variables of interest in the data stream. The triggering of these plans motivates the intentions of the data pre-processing agents, and in this case, the AgentDPP intention to impute the respective variable that has missing instances by any of the corresponding plans for (e.g., `prepros_rnData(), prepros_wlData()` or `prepros_qData()`), otherwise if there is no need to impute data, he drops that current intention. However, concerning the AgentData2Lags behavior, he intends to continue and in parallel with the AgentDPP, the execution of its plans (`createRaw_Data2Lags()` and `createImp_Data2Lags()`) to create the lagged data matrix from the raw and/or raw imputed data, regardless of the AgentDPP dropped his plans of imputing or not imputing data. On the other hand, the SV agent, upon the information issued by the HSn agents regarding the condition of the data, triggers its plan to report the issues concerning the status of the sensors and upon request puts into execution the intention of storing the collected raw and imputed raw data as per the selected plans outlined in Table 5.2.

- **Plans for AgentHDBM.** The plan (`connectingToDB()`), for example, is executed only once at the start of every simulation initialization in which the AgentHDBM is conscious of its intention to connect to the database, to satisfy this intention he executes this plan. Given the duties of the AgentHDBM to ensure the storing of all data products both locally and specifically into the hydrometric database server and to provide the services as data accessibility for other agents, it needs to have uninterrupted login status to the database, connection speed, and be able to perform all required SQL

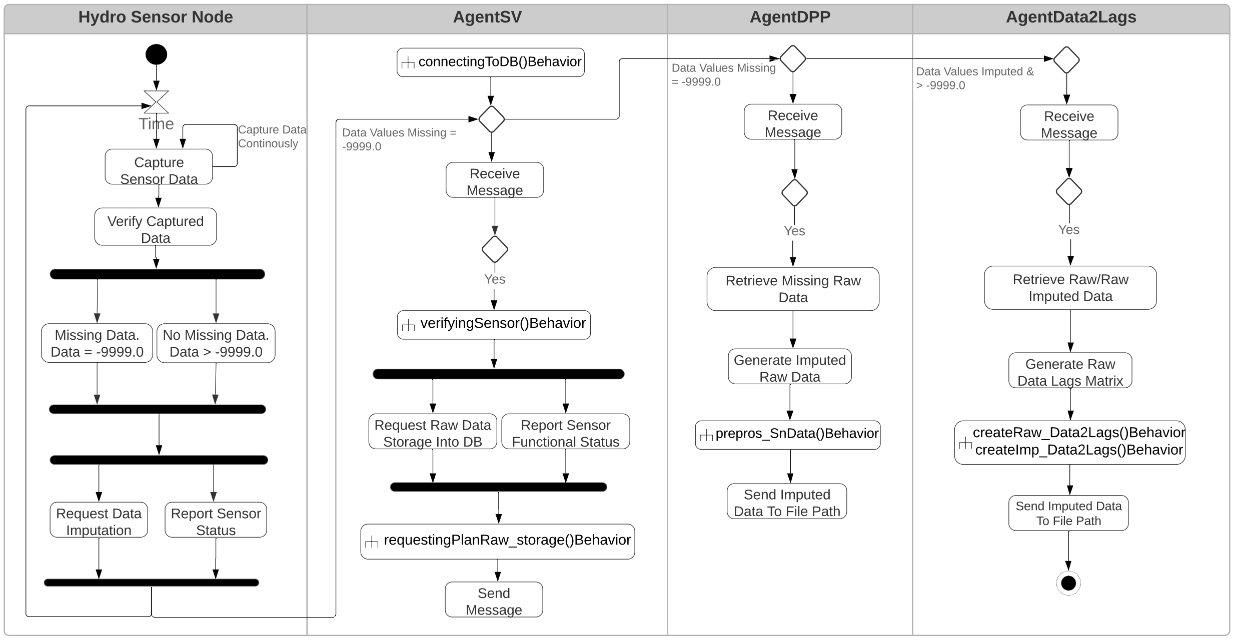


Figure 5.2: Activity diagram for agent's node: HSn, SV, DPP and Data2Lags.

actions. This feature is accessible in GAMA through the parent "AgentDB" a skill that can be inherited by an agent species. However, whereas other agents, for example, the AgentSV, AgentDPP, AgentFL, or AgentUI, can also access the database, their login capabilities are only temporary, and they will have access whenever they need to query information from the database. In this respect, they will require fulfilling three steps: 1) connect to the database, 2) perform SQL statements, and 3) disconnect from the database. The execution of such steps is time-consuming and imposes an expenditure of simulation resources. The activity diagram of the AgentHDBM plans for fulfilling its roles of data storing, querying, table creation is presented in Figure 5.3.

- **Plans for Agents FCST1, FCST2, and FL.** The activity diagram for the flow forecasting learners and decision-making agents is presented in Figure 5.4. The corresponding illustration constitutes a general flow of the AgentFCSTR's and AgentFL behaviors considering their decisions. Behaviors assigned in this activity diagram are relevant to the following intentions: connectingToDB()Behavior – the intention of connecting to the database for the retrieval of the lagged raw data matrix from AgentData2Lags (generator of the actual data); the plan to UpdateBeliefs()Behavior – represents the intention to update and train models and the `predict_Flow()`Behavior – is the intention to conduct numerical computation of flow predictions with the most suitable models based on the "Random Forest" and "Support Vector Regression" algorithms; the intention to send forecast results behavior – sending flow forecast results data to the decision-making agent, AgentFL; who also executes the plan connecting-

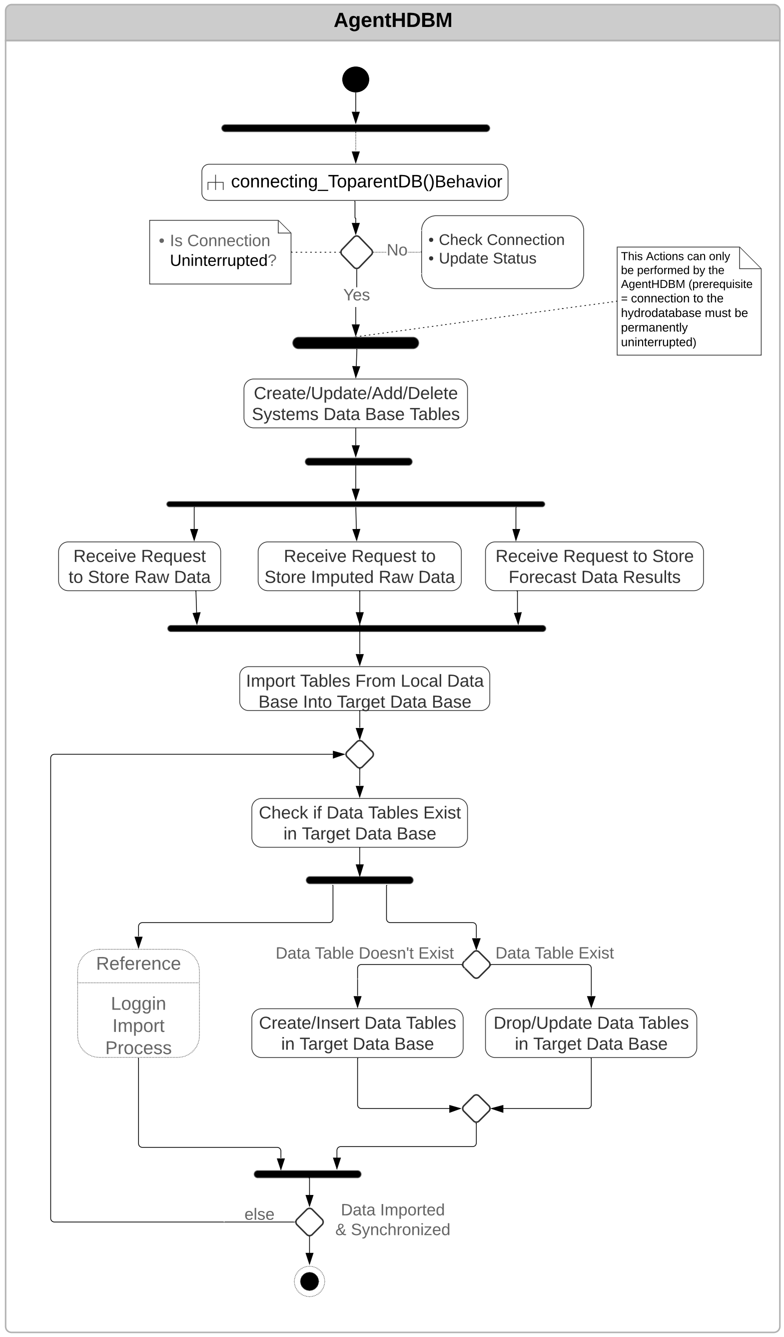


Figure 5.3: Activity diagram for AgentHDBM node.

ToDB()Behavior – the current intention to retrieve flow forecast results, revise the forecast; the intention to `infer_Flow()`Behavior – the plan to perform computation of flow-awareness levels of the actual and hourly forecast data; and the intention to `requestingClas_storage()` – request from the AgentHDBM the storing of the classification and inference data results.

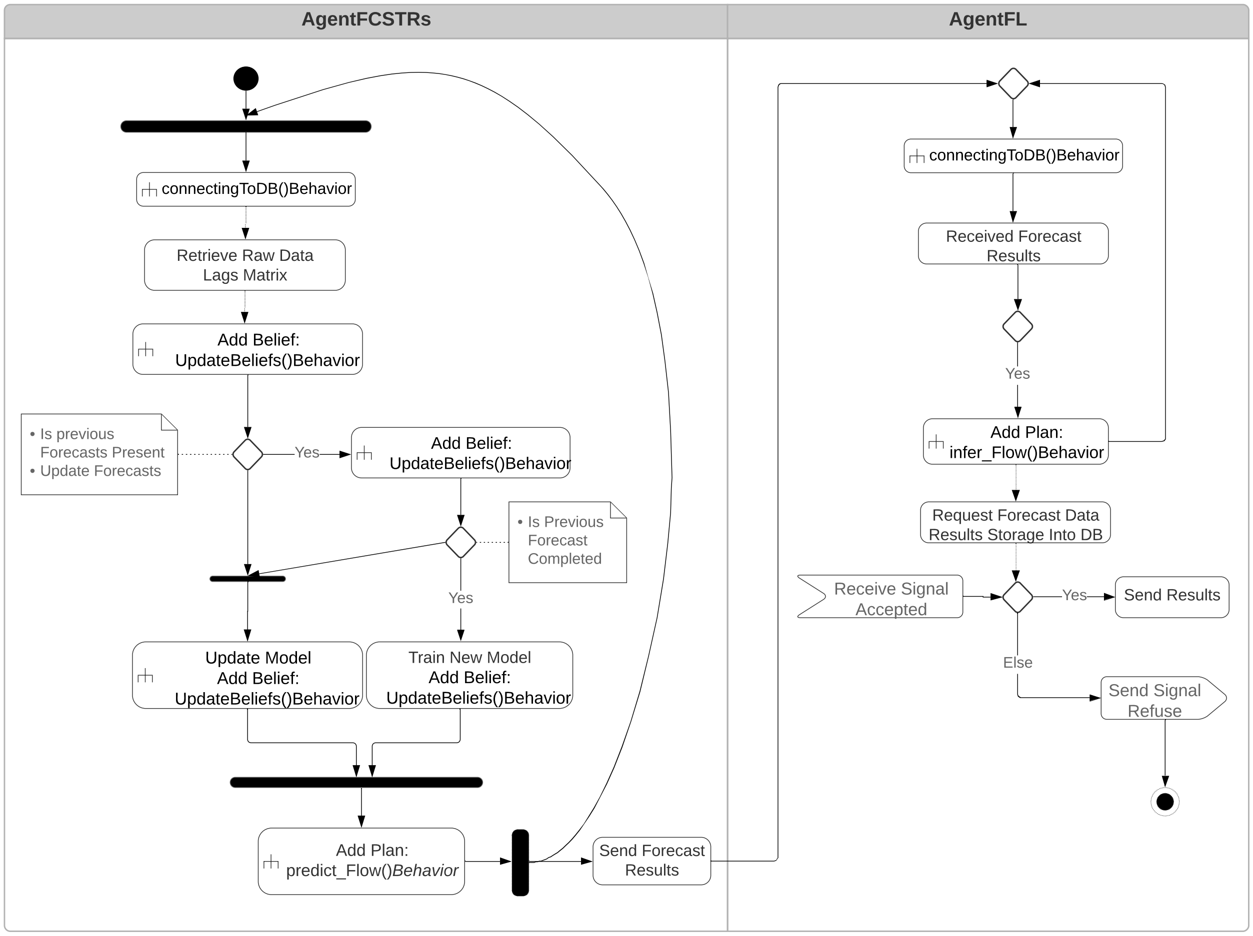


Figure 5.4: Activity diagram for Agent's FCSTR's and AgentFL nodes.

- **Plans for AgentUI.** The AgentUI behavior flows, as for the other agents are organized in the activity diagram displayed in Figure 5.5. These behaviors correspond to the following tasks: the intention connectingToDB()Behavior – the intention of connecting to the database to send and check for system messages and flood-awareness warnings that he receives from other agents in the system (e.g., AgentSV, AgentFL); the intention to become acquainted with the latest accounts in the system, by which he executes the plan UpdateBeliefs()Behavior; – the intensions to register users and verify system alert by activations of the plans verifyingUserRegistration() and verifyingSysAlerts(), respectively.

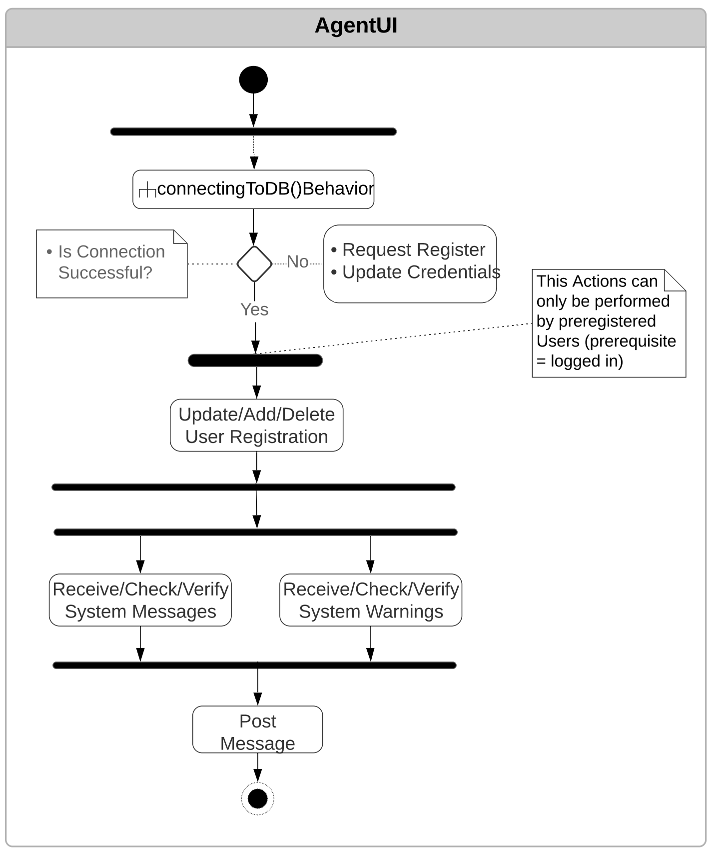


Figure 5.5: Activity diagram for AgentUI node.

5.2 Machine Learning Algorithm for Agent Species

The use of ML over its numerical counterpart is that it offers much power in its ability to deal with small datasets with missing instances and other data integrity-related problems [420–423], that are common to environmental time series; its ease at identifying trends and patterns, and data with high complexity; they offer great efficiency and accuracy with data that has grown over time, just to mention a few of their characteristics. Despite these advantages, they still suffer from several disadvantages when it comes to data acquisition or can be time-consuming to implement, among other factors. Though it is not intended to go deeper into the topic of ML, given that Section 2.2.4 briefly discusses some of the approaches of AI technology in flood forecasting.

5.2.1 Experimental Setup: ML Algorithm for Agent Species

To select the ML regression classifier algorithm for the classifier agents (i.e., AgentFCST1, and AgentFCST2), the key agents for the forecasting objective, eight ML algorithms that are used for regression tasks were evaluated.

To implement the experiments for the evaluation of the algorithms, the RStudio [356] Statistical tool and WEKA [424] (version 3.9.5) data mining software package were used. Both workbenches offer a suite of machine learning algorithms (e.g., Random Forest, XGBoost, SVM) to construct the models for flow predictions.

They are two main reasons for using both workbenches, first because GAMA offers an API that is fully functional for running both RScripts and WEKA implementations through R/Java code implemented modules, and secondly because the bridge between R and GAMA when it comes to ML implementations, is quite much developed than the WEKA. However, there are also advantages between WEKA and RStudio, and this feature offers the use of both R packages and scripts to be implemented in WEKA and vice versa.

5.2.2 Description of the Regression Learner Algorithms

In ML, there are three main ways to go about the learning process. These learning approaches are outlined in the following list.

- **Supervised Learning**
  In supervised learning, models learn by experience, from information that is fed to them during the learning process, and consequently use the learned information to generate new information from the previous experience. As ML is suitable for learning

complex systems, their relationship to agents modeling technologies is closely related, as discussed in [425].

- **Unsupervised Learning**
  Unlike supervised learning which requires learning from data during the learning process, unsupervised approaches try to learn an instance pattern even in the absence of a recognized output, even if there are no available recognized outputs or like in the case of reinforcement learning, on the feedback of the structure of the output. So, without the need for human involvement, these algorithms can disclose information that is hidden in groups of data. Some examples of this type of learning are "clustering", "principal component analysis (PCA)", "K-means clustering", and "self-organizing maps (SOM)".

- **Reinforcement Learning**
  Being a subdivision of ML, this paradigm focuses on the notion of considering the agent as a machine. Therefore, this machine (the agent) is present in an environment in which he perceptively interacts by conceiving the state of the elements of such environment in a vectorized manner. Therefore, in this scheme, a series of decisions must be taken by the agent before the achievement of his goal(s), and these are rewarded upon fulfillment instead of informing if the agent's performance along the path has been for better or for worse. Examples include the Markov Decision Process and Monte Carlo Methods.

With the WEKA workbench, the algorithms were evaluated over one data domain (the synthetic data) delivered from the ABM simulation outputs with the supervised learning approach, as the problem task to solve at hand is of regression type.

The learning processes involved the use of eight algorithms that are available in R and WEKA through the RPlugin API. The algorithms are used for solving regression problems and are detailed below:

- **Elastic-Net**. Elastic-Net forms part of the Generalized Linear Models (glmnet) regression learner of the machine learning package (mlr) function in R [426]. It is one of the three commonly used methods of regularization techniques used in linear models. The other two are the "Ridge Regression" and "LASSO (Least Absolute Shrinkage and Selection Operator)". Regularization (also known as shrinkage) is a method that reduces overfitting. It prevents overfitting in the parameters of a model from becoming excessively large by shrinking them toward zero. This befits the model while making predictions on new data, as the models as less variance. The Elastic-Net is viewed as an extension of both "Ridge Regression" and "LASSO" [427].

- **NNET**. A supervised "feed-forward neural network" learner [428]. It allows fitting a single hidden layer. Feedforward networks only allow the information to travel unidirectionally, which means at the network entrance level across the unseen level toward the exit level, given that loops are not present.

- **SVM**. Support Vector Machine is an algorithm that is a function in R and forms part of the mlr package. It is available in WEKA through the RPlugin. The philosophy of SVM resides in that it has dual usage in problem-solving (e.g., as a classifier and as a regressor). However, it is used more often for classification analysis. SVM works well for data which distribution is unknown. The package source of the SVM algorithm in R is available through the interface of the LIBSVM in the e1071 package [429].

- **Random Forest**. Based on Breiman [430] is a type of regression tree that adjusts several decision tree classification schemes on various sub-samples of a dataset. An ensemble of techniques such as Bagging and Random Space Method, using decision trees as base learners.

- **RPart**. An algorithm implementation based on CART, by Breiman et al. [431] focused on recursive partitioning and decision trees. Although any decision tree can be implemented for regression tasks, the RPart algorithm offers the most accurate implementations. RPart builds general structured models for classification and regression using a procedure of two steps, in which the generated models are defined by binary trees.

- **XGBoost**. A work by Chen Guestrin [432] is an implementation based on the "Gradient Boosting Framework" by Friedman [433]. The XGBoost algorithm is available as an R package, and in other data analysis application languages like Python, Java, and Julia, just to mention a few. XGBoost allows automated computations in parallel on one PC allowing a speed of processing than its boosting counterpart. It can also support regression, and classification tasks, and various objective functions.

- **MLP**. Is a function based on the work of Rosenblatt [434] that creates and trains a Multilayer Perceptron. They are the more basic feedforward networks. The train is done by the error of the backpropagation algorithm and related schemes.

- **MLR**. A Simple Regression Linear Model is a regression classifier from the mlr methods in the R implemented by Bischl et al. [435].

5.2.3 ML Algorithm Performance Metrics and Selection Workflow

The methodology for comparing several ML techniques across various data domains is fairly known, and of which examples can be found in [436, 437], and a recent work reported in [438]. However, the literary information on assessing ML evaluation over a single dataset is not as fully documented, except for a recent survey on this topic done by Basha and Rajput [439] who evaluated several machine learning algorithms for regression and classification tasks over a single dataset for each case.

There are various proposed statistics convenient for measuring the regression aim's efficiency. Notwithstanding, for the quantitative performance of each algorithm evaluated, it is applied the "RMSE", the estimator that shares the same units as the observed and estimated data, to measure the errors between the results, and "R-coefficient" as the estimator of similarity between two variables.

For the selection of the algorithms, it conveyed the following workflow:

- **The Data**. The dataset used in the elaboration of the ML models were created from the ABM simulated available storm hydrographs already detailed in Chapter 4. Then, the simulation outputs from the ABM became the synthetic data used both in the selection of the algorithms, the training of forecasting agent learners, and the validation of the proposed MAS for streamflow forecasts.

- **Predictive Variables**. The selected ABM synthetic storm data which chronologically corresponds to November 2012, is created the lagged and lead time data that corresponds to the time domain of the forecasting periods. From this data set after performing the necessary feature engineering on the data, is selected as the dataset for testing the algorithms to be selected. Table 5.3 shows the characteristics of the dataset with 716 instances that was prepared with actual (δ) and lagged values ($\delta - k$) resulting from the autocorrelation function for the variable rainfall (rn), and water level(wl), as independent variables, and with actual values (δ) for the streamflow (Q), and lag ($\delta - k$), and lead time ($\delta + k$) for the target streamflow discharge 1-hour ahead.

- **Algorithms Tuning Process**. With the WEKA data-mining workbench, the workflow used for testing each of the eight ML algorithms described above was executed within the Experimenter environment. This selection was since the WEKA suite offers a good deal of ML algorithm selection, and those that are not present can easily be obtained from its new features via APIs and plugins, allowing such algorithms from the R or Julia in WEKA.

The object of the experimental simulations in the Experimenter was to identify the prediction correctness and rigor of the algorithm schemes and in this way decide on the best possible prototype model. Therefore, the experiments were performed by tuning the sensitive parameters (Table 5.4) of each algorithm, running the set of experiments with the commonly known "stratified 10-fold cross-validation", like this providing for the synthetic hydrometric dataset to be portioned in random folds of ten between training and testing.

This process of leaven one out at each repeated simulation permits model fitness until the last fold is completed. WEKA experimenter in the process allows for allows model evaluation, in the case of prediction, using the RMSE, the correlation coefficient, and a "Paired Corrected t-Test" is available, as selected here and described in the subsequent paragraphs.

Table 5.3: A representation of the dataset used to evaluate the classifier algorithm.

Variable	Abr.	Instances	Actual, Lag and Lead Attributes	Unit
Rainfall	rn	716	rn_{δ}, $rn_{\delta-1}$, $rn_{\delta-2}$, $rn_{\delta-3}$, rn_{-4}	mm
Water Level	wl	716	wl_{δ}, $wl_{\delta-1}$, $wl_{\delta-2}$, $wl_{\delta-3}$, $wl_{\delta-4}$	m
Streamflow	Q	716	Q_{δ}, $Q_{\delta-1}$, $Q_{\delta-2}$, $Q_{\delta+1}$	$m^3 \cdot s^{-1}$

5.2.4 Experimental Results: ML Algorithm for Agent Species

The selection of algorithms suitable for machine learning problem solving is a very important task. In this experimental task, both R and WEKA were used, with which was implemented eight ML algorithms for evaluation on the dataset with 716 instances, thirteen independent variables, and target variable (i.e., $Q_{\delta+1}$) and their performance reported.

The experimental setup follows the building of the regression models (see equation (5.1)) with 10-fold cross-validation given its importance as an adequate tool for evaluating classification and regression methods [440–442] with the tuning of hyperparameters for each algorithm as can be seen in Table 5.4.

$$Q\left(\delta+k\right)=F\left(rn_{\delta},rn_{\delta-k},wl_{\delta},wl_{\delta-k},Q_{\delta},Q_{\delta-k},Q_{\delta+k}\right) \tag{5.1}$$

where δ is the timeframe of the phenomenon, k is the time step variable, rn and wl the rainfall and water level variables, and F is a function that usually defines a large flow rate (Q) of the time series.

Completed the process of simulation runs in the WEKA Experimenter environment, the overall analysis revealed by the performance metrics for the numerical predictions given by the algorithms are summarized in Table 5.5 and 5.6 respectively.

In general terms, the algorithms to have lower values of the RMSE are SVM (17.56 $m^3 \cdot s^{-1}$), Elastic-Net (18.22 $m^3 \cdot s^{-1}$), LM (18.84 $m^3 \cdot s^{-1}$ and RF (19.71 $m^3 \cdot s^{-1}$), with SVM and RF both displaying the highest correlation coefficient (R = 0.92). The remaining four algorithms have higher error values, been XGBoost with the highest (32.83 $m^3 \cdot s^{-1}$), followed by NNET (29.70 $m^3 \cdot s^{-1}$), MLP (24.51 $m^3 \cdot s^{-1}$), and Rpart (22.07 $m^3 \cdot s^{-1}$) with

Table 5.4: Machine learning algorithms and Tuned Hyperparameters.

ML Algorithm	Hyperparameter to Tune	Tuned Hyperparameter	Best Tune Hyperparameter
Elastic-NET	α: A parameter used for shuffling the data, with $0 \leq \alpha \leq 1$. In LASSO and Ridge regression $\alpha = 1$, and $\alpha = 0$ a: Elastic-NET mixing parameter with $0 \leq \alpha \leq 1$. The penalty is defined as a lasso penalty for $\alpha = 1$ and a ridge penalty for $\alpha = 0$. λ: User-specified sequence for λ. Normally, the program should calculate the λ sequence based on $n\lambda$ and $\lambda.min.ratio$. If the value λ is specified, it overrides this.	$\alpha = 0.01, 0.05, 0.1$. $\lambda = 0.10, 0.21, 0.15$. tune length = 10, 20, 15	$\alpha = 0.1$. $\lambda = 0.21$. tune length = 10.
NNET	size: No. of hidden layers. Decay: Is the weight decay. Default=0.	size = 10, 6, 20. Decay = 0.05, 0.1, 1, 2	size = 10. Decay = 0.05
SVM	C: The penalizing factor.	C = 5, 3, 2	C = 2
Random Forest	ntree: The amount of three to build. Assures inputs to be predicted. mtry: The number of random samples per splits.	ntree = 500, 200, 1000, 1200. mtry = 5, 3, 2, 6	ntree = 1000. mtry = 2
Rpart	cp: The list of complex values.	cp = 0.05, 0.01, 0.1, 0.5	cp = 0.01
XGBoost	η: Controls the learning-rate: The input normalization factor $[0 \leq \eta \leq 1]$. Default: 1.	$\eta = 0.01, 0.1, 0.5$. $\gamma = 3, 5, 2, 1$. max_depth = 10, 5, 6, 8. min_child_weight = 5, 3, 10. subsample = 0.1, 0.01, 0.5.	$\eta = 0.5$. $\gamma = 5$. max_depth = 10, min_child_weight = 5, subsample = 0.5
MLP	seed: A random sample generator. Momentum: Update weights. hiddenlayers: No. of hidden layer for the neural net. decay: Decreases the learningrate. learningrate: Update weights.	seed = 500, 100, 200, 300. Momentum = 0.05, 0.1, 0.01, 0.02. hiddenLayers = 10, 20, 30. decay = 0.05, 0.01, 0.1. learningRate = 0.01, 0.05, 0.1	seed = 300. Momentum = 0.05 hiddenLayers = 10. decay = 0.05. learningRate = 0.01
MLR	N.A	None	None

the R coefficient in the range [0.85, 0.89] respectively, except for NNET which had the lowest correlation value (0.53).

Now, on a significance level fixed at $\alpha = 0.05$ when regarding the performance of each algorithm as a case base scheme, in the case of Elastic-Net, Table 5.5 shows that the results of NNET, Rpart, XGBoost, and the LM are statistically improved over Elastic-Net. When defining NNET as the base algorithm, it turns out that Elastic-Net, SVM, RF, Rpart, and LM, are statistically worse than NNET. If it is SVM the baseline, then NNET, RF, Rpart, and XGBoost are statistically better.

Table 5.5: A Paired Corrected t-Test Using the RMSE Measure for the ML Algorithms. The entries in row M and column N show the results for the RMSE performance measures of M over N and the ratio of the number of wins:lost:difference (base on a 5 x 10 t-test). The markers α and β, represents the algorithms that were statistically better or worse than the other when acting as base line algorithm.

	Elas-Net	NNET	SVM	RF	Rpart	XGBoost	MLP	LM	Total
Elas-Net		29.70α	17.56	19.71	22.07α	32.83α	24.51	18.84α	0:4:-4
NNET	18.22β		17.56β	19.71β	22.07β	32.83	24.51	18.84β	6:1:5
SVM	18.22	29.70α		19.71α	22.07α	32.83α	24.51	18.84	0:3:-3
RF	18.22	29.70α	17.56β		22.07	32.83α	24.51	18.84	0:4:-4
Rpart	18.22	29.70α	17.56β	19.71		32.83α	24.51	18.84β	4:2:2
XGBoost	18.22β	29.70	17.56β	19.71β	22.07β		24.51β	18.84β	7:0:7
MLP	18.22	29.70	17.56	19.71	22.07	32.83α		18.84	1:2:-1
LM	18.22β	29.70α	17.56	19.71	22.07α	32.83α	24.51		1:3:-2

Continuing with the same evaluation methodology, when evaluating the algorithms on the regression task and using RF as a test baseline algorithm, it was noticed that only NNET and XGBoost are statistically best over RF, whereas SVM performs worse. For Rpart, similarly, NNET and XGBoost perform statistically better over Rpart, and just as SVM performed badly with RF, the same outcome was observed for Rpart, in addition to the LM that also resulted statistically worse. Interestingly, XGBoost as test base, with a high value of the RMSE, resulted in no tie, and no losses, as it portrays seven wins, given there are six algorithms that performed statistically worse than it. Only NNET lags behind XGBoost with six wins and one loss. For the MLP, only XGBoost showed to be statistically improved, whereas, for the LM algorithm, NNET, Rpart, and XGBoost were labeled as statistically better, with Elastic-Net performing significantly worse.

So far, as can be seen in both Tables 5.5 and 5.6 though, the magnitude of the bias, and concerning the R coefficient, it seems like SVM is better than RF. The table also reveals that SVM has even lower errors than Elastic-Net, and although RF has a slightly higher

Table 5.6: A Paired Corrected t-Test Using the R Measure for the ML Algorithms. The entries in row M and column N show the results of the R performance measure of M over N and the ratio of the number of wins:lost:difference (base on a 5 x 10 t-test). The markers α and β, represents the algorithms that were statistically better or worse than the other when acting as base line algorithm.

	Elastic-Net	NNET	SVM	RF	Rpart	XGBoost	MLP	LM	Total
Elastic-Net		0.53β	0.92	0.92	0.86	0.86	0.85	0.89	2:0:2
NNET	0.89α		0.92α	0.92α	0.86α	0.86α	0.85α	0.89α	0:7:-7
SVM	0.89	0.53β		0.92	0.86β	0.86β	0.85	0.89	3:0:3
RF	0.89	0.53β	0.92		0.86β	0.86β	0.85	0.89	3:0:3
Rpart	0.89α	0.53β	0.92α	0.92α		0.86	0.85	0.89	1:3:-2
XGBoost	0.89	0.53β	0.92α	0.92α	0.86		0.85	0.89	1:2:-1
MLP	0.89	0.53β	0.92	0.92	0.86	0.86		0.89	1:0:1
LM	0.89	0.53β	0.92	0.92	0.86	0.86	0.85		1:0:1

RMSE error than Elastic-Net, it has a higher value of R. Both SVM and RF algorithms are followed in performance by Elastic-Net and the LM with the same value of the R = 0.89, although Elastic-Net reported having the lowest error. A dispute between the performance of Rpart and MLP can also be noted, where the correlation coefficient of Rpart is only one unit higher than that of MLP; however, based on the appropriate performance measure (for example, RMSE) that runs the significance test, Rpart outperforms MLP on the RMSE. Finally, NNET and XGBoost have the highest errors, but XGBoost has a better correlation coefficient than NNET.

5.2.5 Conclusions

As mentioned earlier in this section, the aim for evaluating several ML algorithms was not meant to conduct a thorough study and examination of the algorithms per se, but for selecting, that algorithm or set of algorithms that could be adequate in the implementation of the MAS classifier agents' species whose goal implies the forecasting of flows. Eight different ML models were developed using the WEKA package software. To verify the complete baseline dataset and minimize the overestimation of the model for flow predictions during training, it was applied a 10-fold cross-validation approach. The algorithms showed significant differences in their performance, SVM and RF tested to be the most relevant algorithm for the forecasting task. For the most part, they achieved better results for the hydrometric dataset relating to the RMSE and the R coefficient metrics. Although the corrected paired t-test on the RMSE confirmed XGBoost, NNET, and RPart with the greatest number of wins, they have the highest RMSE and the lowest R. However, of notice, in ML, it should be noted that the best learning scheme for a given dataset, doesn't ensure to perform accurately on new data with attributes that are slightly varying or different. So, the question is not which algorithm outperforms which, but is goodness in fitting a particular data domain.

5.3 Fuzzy Logic Model for Agent Species

In Subsection 2.2.4.2, it was explored in the literature review certain soft computing techniques like in the "applications of fuzzy logic in hydrology" and "water resources engineering and management" to cope with the engineering problems caused by floods. This section, though, not intended to be a treatise about the fuzzy system, which is comprehensibly submitted in the dedicated published writings, describes the settings used for the implementation of the decision agent (i.e., AgentFL) that form part of the MAS model framework, is endowed with "FUZZY LOGIC" capabilities and skills, who based on the actual hydrometric information and flow forecasts provided by the forecaster agents perform the flood-awareness levels inferences at each established time horizon.

5.3.1 Fuzzy Inference System: A General View

Inference means What? The inference is a process whereby an assertion made from some premise is true, proceed to another assertion because of a structure of rules, that leads to a second assertion being true. However, according to [443] inference implies a reasoning object and facilitates ins (data) and outs (results) on account of rules meant for this aim. These rules are what is known in fuzzy set theory, as "fuzzy inference rules".

Generally, FIS are systems from which is built "fuzzy expert systems". A FIS is constructed from a group of functions known as membership functions (MFs) and is rule-based. One of the characteristics of FIS is that they offer reasonable estimates regarding the data, on which the inference based on rules is applied. In any expert system, it is likely to find two important functions, where the first is, the problem-solving function able to exploit specific fields of information, and the interaction function, operating at the user's level, and which explains the system's intention at the initial and ending of the problem-solving cycle. They often expected the expert system to handle unclear and insufficient data. However, an expert system is an operator-interaction configuration, as shown in Figure 5.6, which comprises three components:

- **Knowledge base:** This component considers knowledge that is unique to the domain of application, with information concerning the domain and rules that explain correlations in the domain. Hereby, "IF-THEN" rules are the most common form of imparting knowledge.

- **Interference engine:** In this component, knowledge of the system is actively used to perform reasoning to generate answers for a user query.
- **User interface:** This component provides communication to flow between users and the system, besides supplying the user with understanding into the problem-solving cycle by the use of inference (reasoning).

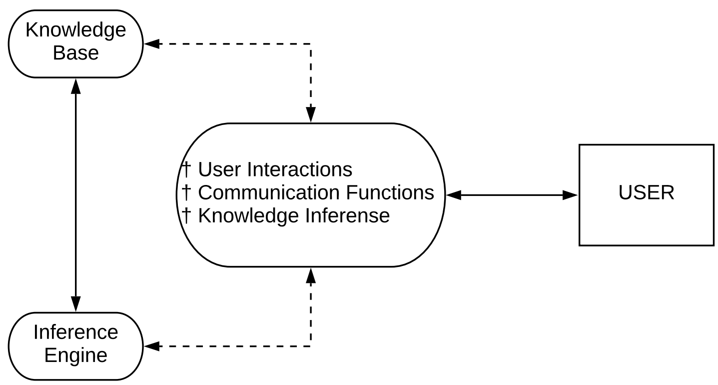


Figure 5.6: Structure of a knowledge-base setting.

A FIS is presented in four widely known operating methods; However, here there is focus on two: the Mamdani [444, 445] and the Takagi and Sugeno (TSK) [446] fuzzy models. Both Mamdani and TSK are a part of the "MATLAB Toolbox" for Fuzzy Logic applications development.

5.3.2 Experimental Setup: FL Model for Agent Species

The initial setup for the fuzzy logic agent can be seen portrayed in Figure 5.7, which is the schematic representation of the recommended rationale for the fuzzy logic prototype that is presented here as a component (i.e., at the classification level) within the suggested hydrologic MAS framework for streamflow prediction in the tropical basin. Therefore, this fuzzy logic skill is an element that is embedded in the GAMA AgentFL species, which portrays the role of the decision agent. In this arrangement, both Agents groups of forecasters (FCST1 and FCST2), and the AgentFL are only employing the hydrometric data information captured directly from standalone files and not data fed into the system in real-time. For example, the rainfall agent, stage (water level) agent, and river flow agent, are connected to the field sensors in real-time, each of which senses the surroundings of a hydrometric station for perceptions on precipitation, river water level, and flow.

The crisp intervals here are the input (actual) variables collected by the hydrometric sensor's agents, and the flow forecasts generated by the AgentFCST group, these become conceptualized because of the fuzzification mechanisms into linguistic quantities which are in direct relation to the input linguistic variable quantity. The fusion of the rainfall, stage, streamflow, and flow forecasts data will become earmarked for the decision agent to infer the flow regime and deliver warnings and disclosure of information to the users concerning the measures to be adopted.

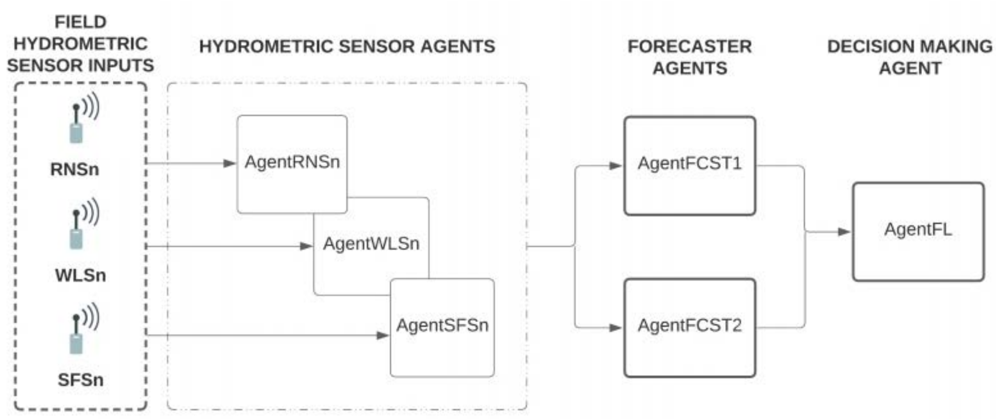


Figure 5.7: Schematic of the fuzzy logic model rationale, showing the field hydrometric sensor data input level, the hydrometric sensor agents' level, classifier level, and the corresponding decision-making level.

5.3.3 The Fuzzy Model Layout

The design for this fuzzy logic task is implemented with the three input hydrometric variables, and the AgentFCST results which were initially divided into four linguistic variables depending on the type of variable (see Table 5.7). In this sense, for these hydrometric variables, the linguistic partitions are as follows: i) rainfall, with linguistic variables Light, Average, Intense, VeryIntense, ii) variable stage, Normal, Average, High, VeryHigh, and iii) for both variables streamflow and AgentFCST, Low, Average, High, and VeryHigh. For each of these, its degree of membership is also identified within the fuzzy set. There are several membership functions (MFs) that can build fuzzy inference systems such as the Triangular, Trapezoidal, Gaussian, Sigmoidal [447]. The triangular membership function was chosen(see Figures 5.8 and 5.9) whose mathematical description is determined by equation (5.2) below, given its ease of use in handling accurate ranges and its evaluation is much simpler than the other contra parts, that even the lay and the non-technical user can easily understand it. The calibration and tuning process of fuzzy rule systems, along with the identification and selection of suitable MFs, although they are several insights and approaches offered in the literature to conduct such a task, can be a complex one and depend on the problem to be solved. However, fuzzy rule-based systems can be enhanced with adequate calibration techniques that allow the enhancement of the model's output performance. Perhaps, using the tuning process, the fuzzification process, with the conjunction of an expert's input, helps formulate the symbolic logic. The tuning process depends on the set rule base defined for the system and the database. The tuning process that was followed here is the well-known *ad hoc* data covering technique proposed in [448]. An iterative method which as a low time-consuming process for learning the fuzzy rules by covering the criteria of the data being modeled.

Table 5.7: Input and output variables for the flood awareness level measure.

Hydro Variable	Linguistic Tag	Range
Rain (mm)	Light	[1 – 50]
	Average	[25 – 75]
	Intense	[50 – 150]
	VeryIntense	[100 – 250]
Stage (m)	Normal	[8.0 – 9.0]
	Average	[8.5 – 10.0]
	High	[9.5 – 11.0]
	VeryHigh	[10.5 – 12.0]
Streamflow (cms)	Low	[0.5 – 100]
	Average	[50 – 200]
	High	[100 – 300]
	VeryHigh	[200 – 400]
AgentFCST (cms)	Low	[0.5 – 400]
	Average	[200 – 500]
	High	[400 – 600]
	VeryHigh	[500 – 1000]
Flood-Awareness Levels	LOW	[0 – 4]
	CAUTION	[2 – 6]
	ALARM	[4 – 8]
	RISK	[6 – 10]

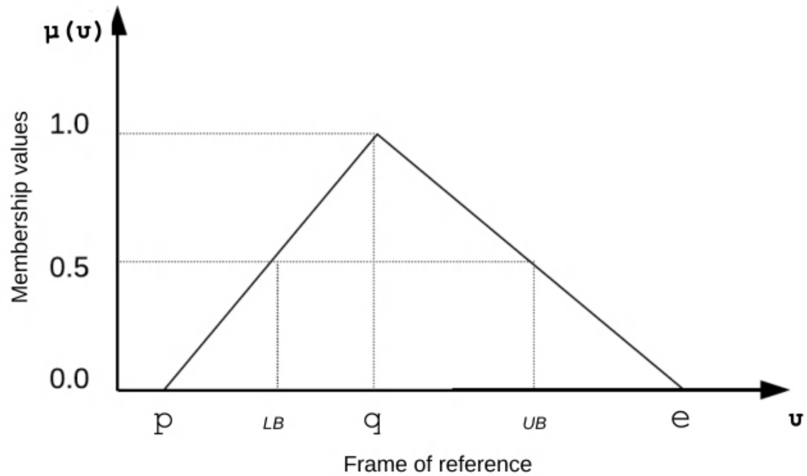


Figure 5.8: Triangular membership function. Where p is a lower limit, an upper limit e, and a focal point q, LB & UB, are cross over points, and $p < q < e$.

$$\mu_A(\nu) = \begin{cases} 0 & \nu \leq p; \\ \dfrac{\nu - p}{q - p} & p < \nu \leq q; \\ \dfrac{e - \nu}{e - q} & q < \nu < e; \\ 0 & \nu \geq e \end{cases} \tag{5.2}$$

5.3.4 Experimental Results: FL Model for Agent Species

Once the configurations of the fuzzy system have been set up with the input variables of interest, and the rules have been established, the previous are duly merged with their convenient linguistic variables and hence the inference engine, using the Fuzzy Logic Toolbox of the MATLAB computational package (2020b, Academic version) accesses the fuzzy rule base to compute and generate both the average and the output linguistic values (Figure 5.10). Hence, before this process can take place, there are two important steps to complete which lead to these processes. These steps are the aggregation and composition phases. With aggregation, the output for each fuzzy set rule is integrated with the inference phase. In simpler terms aggregation computes the IF portion of the rule whereas in the composition phase, the THEN operation is computed.

The fuzzy inference structure is the central unit of a fuzzy rule implementation from linguistic variables upon which experts base their knowledge on decision making. It relies on the "IF-THEN" rules alongside the "OR" or "AND" connectors for representing necessary decision rules. In this experiment, a set of one hundred twenty rules were created (see Figure 5.10). The rules used in the decision process were generated from the three hydrologic input parameters (rainfall, stage level, and streamflow) and according to these is decided the output flood-awareness level value. The system was simulated using the min-max or maximum means (MOM) Mamdani's model and the centroid (COG) method for defuzzification. Therefore, if the actual number of active rules are p^n, where "p" represents the highest quantity that the membership functions or inclusions can have, "n" is total inputs, then the configuration for the fuzzy system based on the three mentioned inputs above is 3 inputs with 4 membership each, which results in a total of $3 \times 4 = 12$ rules. However, to enhance the system further, and safeguard a workable setup of fuzzy rules and membership functions be valid and practical in computational terms, especially for real-time operation purposes as suggested for the system, four more rules were added based on expert opinion and knowledge of the precipitation regime and conditions of local flow for the river system within the catchment of study.

During the defuzzification mechanism, the values corresponding to the outputs of the linguistic variables are transformed into their crisp format. Thus, in the arrangement for each linguistic term, its maximum value, in turn, represents the maximum (center) values of each membership function. this suggests that the median of the interval gives the maximum value of the interval of a membership function. Now, regarding the final decision concerning the flood-awareness state that the system should disclose, the usual values for the linguistic terms are given by "LOW", "CAUTION", "ALERT", and "RISK", as described in Figure 5.10 and 5.11 respectively. From this perspective, it was calculated the best combinations of the input values using the weighted average and considering their degrees of membership. Once these calculations are carried out, the decision agent (i.e., fuzzy logic agent) based on the results of the linguistic terms oversees announcing to the user interface agent about the flood-awareness levels defined in Table 5.7. As the three hydrometric agents are constantly perceiving the hydrologic conditions of the environment, to evaluate the performance of the fuzzy logic decision-maker agent, the situation of the simulation of the changes in flooding severity is shown by the mapping from streamflow, stage, and rainfall to the flood-awareness level (Figure 5.11).

Once the design and implementation of the fuzzy system were complete, which purposes and functions are part of the component for the agent associated with deciding about flood alert levels within the proposed MAS model, its integration into the MAS model, allows afterwards the consecution of a series of experimental and validation tests to verify its efficiency for the assigned task. Below in the following paragraphs, is described the experiments performed and comment on the eventual results.

For the evaluation of the fuzzy model, it was used the synthetic data generated by the ABM outputs, as the case study to build and test the system, since it is a period that presented the highest rainfall, stage height, and discharge. So, to test the system, for example, is evaluated on five rule cases, as shown in Figure 5.12 according to the MATLAB simulation, the outputs consider the accumulated flood volume in the channel, giving the input precipitation and the water level to hydraulically reproduce the final decision of the fuzzy agent model, as it reproduces the flood-awareness as expected (see Table 5.8). Therefore, it can be seen that when the system is simulated for a given range of input values, say for example, if the value of rainfall is = 23 mm, which lies within the range [1, 50 mm] as Light, and stage = 11.6 m, which lies in the range [10.5, 12.0 m] as VeryHigh, and the observed and 1 hour forecast streamflow = 404.0 and 410.0 $m^3 \cdot s^{-1}$, which lies within the ranges [200, 400 $m^3 \cdot s^{-1}$] as High and [200, 500 $m^3 \cdot s^{-1}$] as VeryHigh, the resulted flood-awareness level is likely 5.0 (Figure 5.11). Therefore, high values of flood-awareness, for high values of rainfall, streamflow, and water level, and low values of the flood-awareness for low values of rainfall, water level, and streamflow when contrasting the results in the Table with the graphics (see Table 5.8, Figures 5.12, and 5.13) as shown below.

Table 5.8: Outputs on flood awareness inference based on the hydrometric data inputs, ABM averaged simulated flows and the lead time forecasts.

Cases	Rain [mm]	Stage [m]	ABM Obs. Q_{pk} $[m^3 \cdot s^{-1}]$	MAS Forecasted. Q_{pk} $[m^3 \cdot s^{-1}]$	FA
1	23	11.6	404.0	410.0	5.0
2	1	8.5	432.8	462.9	1.6
3	16	9.0	586.5	539.6	2.0
4	1	8.5	358.5	434.4	1.6

5.3.5 Conclusions

In this section, it was presented the design, implementation and testing of the fuzzy logic component integrated into the proposed MAS hydrologic model that can be useful in the humid catchment for inferring on different flood-awareness levels. This model implements the features for decision-making by the AgentFL, based on a fuzzy rule with capabilities to infer flow conditions in the river, also comprises other modules with capabilities for managing and storing the hydrometric forecast data. The experimental results show that the hydrometric variables of rainfall, water level, and streamflow, the generated hourly forecast values, once implemented with their membership function and crisp inputs and subjected to the process of fuzzification and defuzzification methods and validated by certain rules that express the participatory capacity building of an expert domain, represents a highly effective tool when dealing with data that present high uncertainty. Hence, for each condition of the actual hydrometric and flow forecast values, the system could infer the conditions of the flow regime in the channel as was shown in Table 5.8.

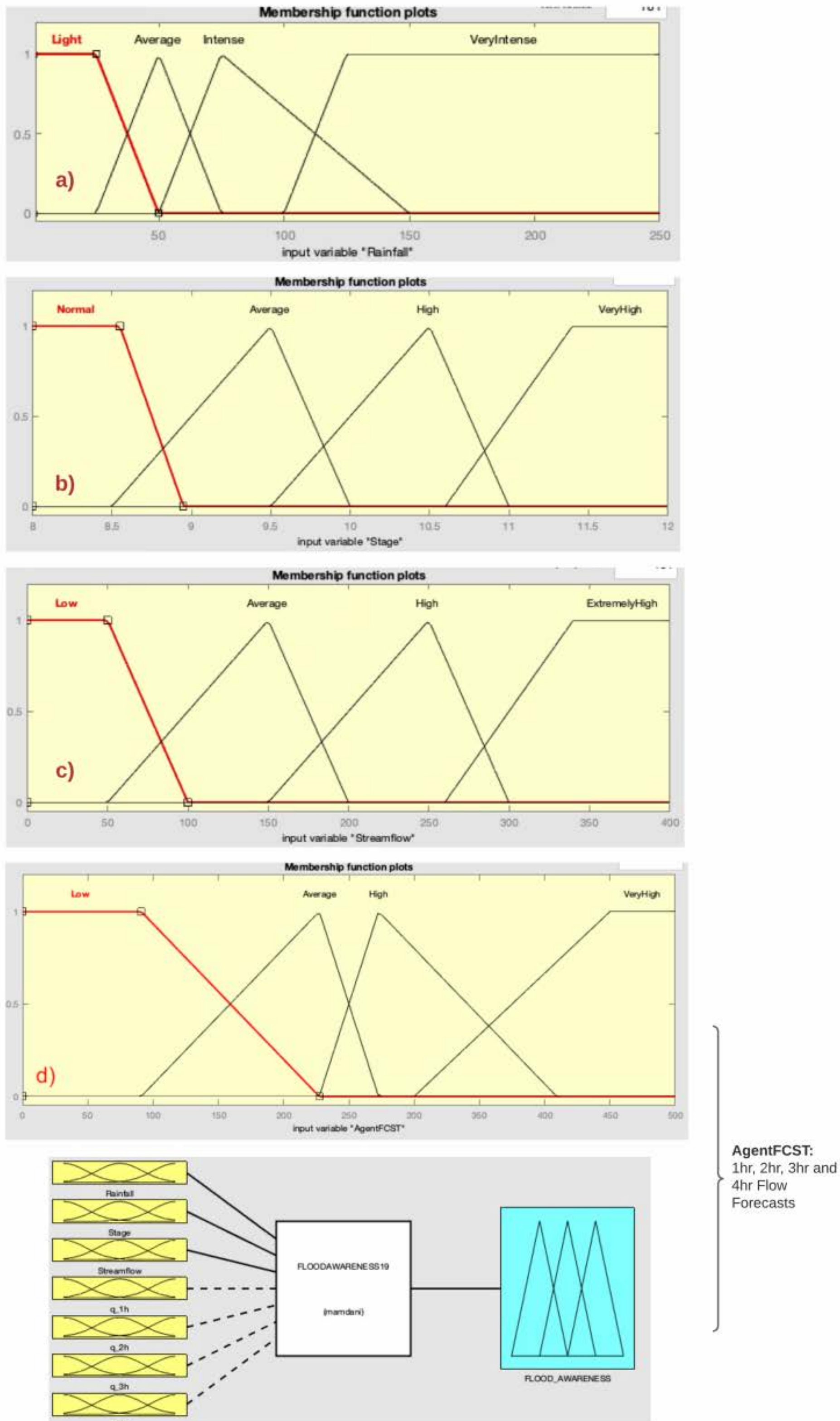


Figure 5.9: Triangular membership function for three hydrometric and the Agent Forecasters input values: a) rainfall [mm], b) stage [m], c) streamflow [m3/s], d) AgentFCST's (implies the 4hr lead time forecasted variables (q1h,...,q4h, accordingly)).

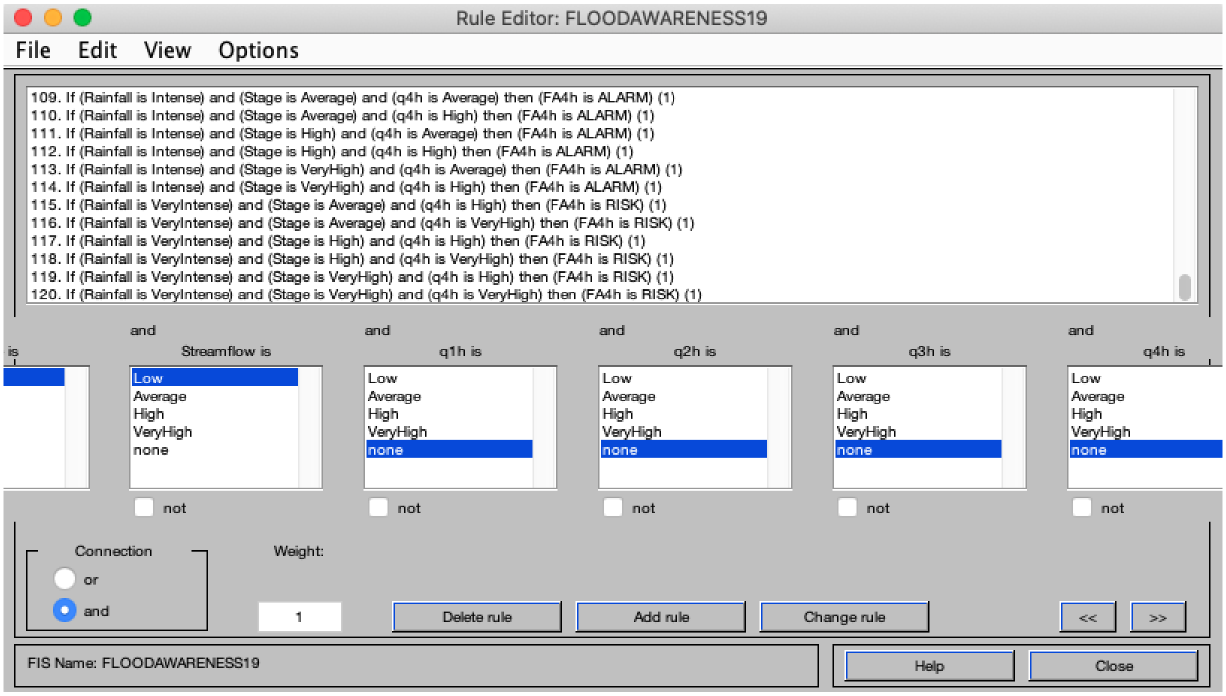


Figure 5.10: Fuzzy rules implementation for the MAS decision-making capability.

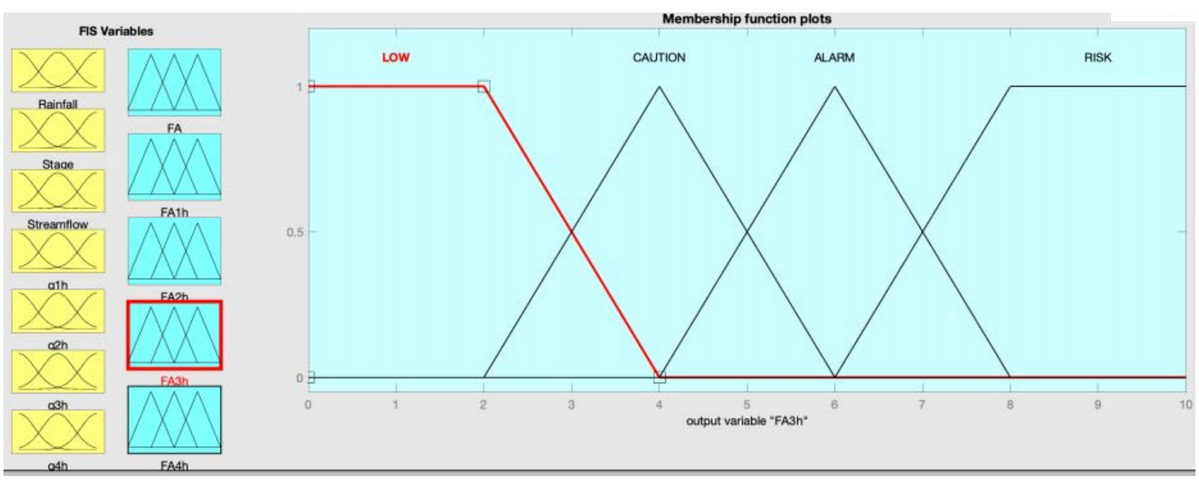


Figure 5.11: Flood-Awareness universe of values.

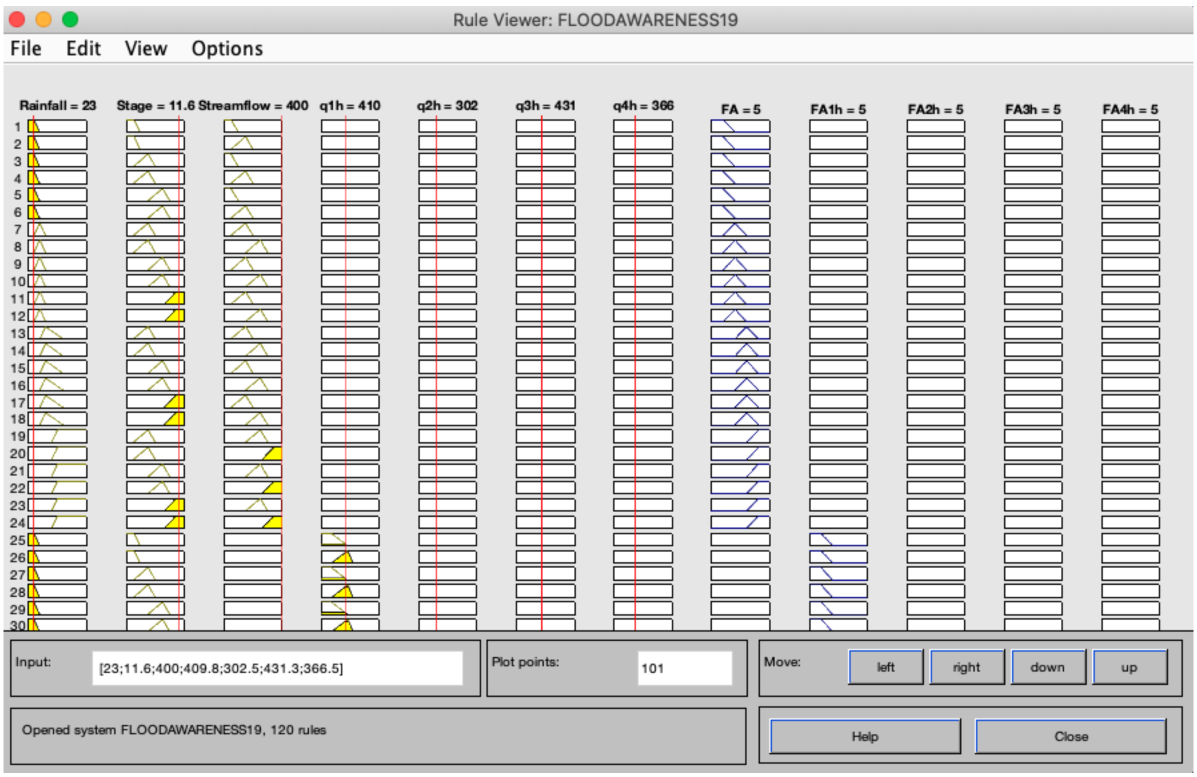


Figure 5.12: Fuzzy inference simulation output corresponding to the input variables of the rainfall [mm], stage [m], and the observed ABM streamflow [m3/s] output, concerning the MAS lead time flow forecasts (q1h,...,q4h) and the respective delivered hourly flood-awareness (FA1h,...,FA4h).

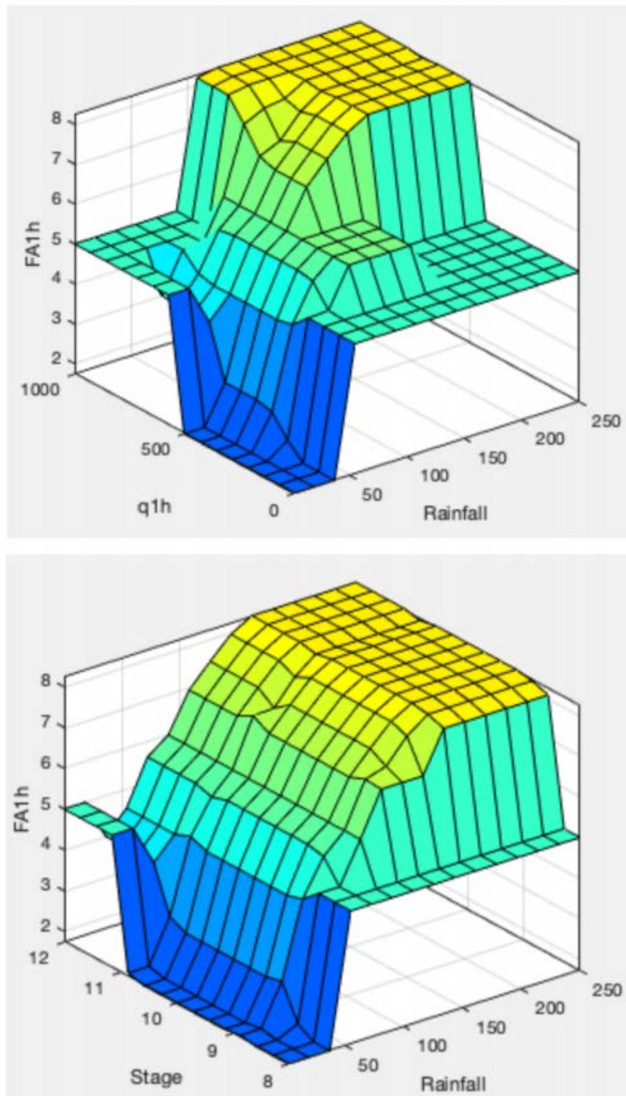


Figure 5.13: A mapping of the spatio-temporal representation of the results of the simulated membership functions related to three inputs and the fuzzy inference outputs: a) mapping the rainfall [mm], the 1 hour flow forecast (q1h) to the corresponding 1 hour flood-awareness (FA1h) level and, b) mapping rainfall [mm] and stage level [m] for 1 hour lead time flood-awareness (FA1h).

5.4 Hydrologic Simulation with MAS

In this section, it is provided an attempt to demonstrate the performance with an extension to the initially built ABM prototype described in Section 4.4.3, once this model had been calibrated and validated, now it is extended with the addition of machine learning approaches and the BDI architecture, represents the foundations for the MAS model approach that is built, trained, tested, and validated. Some of the agent species were provided with ML algorithms, cognitive skills, and capabilities for data-mining tasks such as data preprocessing, regression for forecasting flood hydrographs, and fuzzy logic skills for performing inferences on the forecasts. Therefore, the proposed MAS framework for flood forecasting described previously in Chapter 3 is validated when compared its outputs against the simulations of the ABM flood hydrographs outputs. This was done by conducting a series of experiments simulating the real-world storm events that occurred within the Medio River catchment.

5.4.1 Experimental Setup: MAS Hydrologic Model Setup

In, this experimental section what was done is laying the foundations for a first approximation of the flood hydrograph forecast using the MAS model that is deploying cognitive agents on the previous ABM model setup (e.g., the hydrometric sensors, the catchment, the DEM grid, the river network), the ABM model is extended with new capabilities including other agents with cognitive abilities (i.e., "BDI-architecture model") for some type of agents to exhibit certain knowledge and on that premise, solve the flood forecasting problem as they interact with each other and perform certain specialized tasks with machine learning capacities (e.g., collect data, store data, make predictions, data-mining task, pre-process data, decision-making, etc.) as was formulated in Section 3.2.1.

Therefore, to carry out these experiments with the implemented MAS model, and remake the previous steps carried out in the other modeling follow-up sections, the process begins with the selection of the benchmark data, which in this case is represented by the synthetic dataset that was delivered from the ABM simulations of the observed flood hydrographs as the data feed to the MAS framework. Once this process had started, what followed is the completion of the training and testing of the machine learner agents, whose roles are to make the forecasts. As a reminder, the problem to be solved is to perform flow forecasting of the Medio River channel at various lead-time (e.g., 1, 2, 3, and 4 hrs. ahead) to ensure the anticipated flood conditions of the river (i.e., flood-awareness levels). Hence, the administration of this task is performed by the interactions of BDI hydrometric agents endowed with actions, skills, and behaviors defined in Sections 3.3 and 5.1. The rationale for the experiments performed with the MAS model implemented to simulate overland flow in the Medio watershed and test outputs of storm scenarios are described below in the next section.

5.4.1.1 Hydrologic Agent Forecaster Learners

As was defined, the agents who within the MAS model will carry out the forecasting tasks for the unknown flood hydrographs at $Q_{\delta+1}$ for certain time intervals into the future denoted by $\delta+k$, and k represents the next time step already selected from the information collected from the previous autocorrelation analysis of the hydrometric series, it is then carried out the enabling activities of these agents that have been equipped with ML algorithm capabilities for regression tasks.

Based on the ML algorithms chosen under Section 5.2 (e.g., "random forest (RF)" and "support vector regression (SVR)") and which were implemented in the ML agents, the forecaster agent's capabilities were trained, tested, and validated on the ABM model simulated flood hydrographs outputs for the duration of time (δ) which means that the outcome is the forecasted value. The input variables (rainfall, water level, and streamflow), are the values captured from the sensors by the hydrometric sensor agents. However, this captured data needs to be preprocessed by the agents DPP and Data2Lags before the forecaster agents can make proper use of the data to make the respective forecasts. Since the aim of the implemented models is for prediction, it must be highlighted that the output value is projected as opposed to the actual data since an implementation for forecasting flow ($Q_{\delta+1}$) at a given time ($\delta+k$) is only valid till time (δ) as expressed previously in equation (5.1).

Recalling some hints described under Section 3.2.1.4 (System Classifier Level), for the forecasting and inference of flow (flood-awareness levels), the MAS model integrates two groups of agents whose roles are to perform the regression task, and in the MAS model, and for the sake of name handling, given that they are composed of eight individual agents, the groups have been labeled as AgentFCST1 and AgentFCST2. Both AgentFCST1 and AgentFCST2, consisted of four agents each that were implemented with the RF and SVR algorithm to predict flows (Q) at time horizons $Q_{\delta+1}$, $Q_{\delta+2}$, $Q_{\delta+3}$, and $Q_{\delta+4}$, accordingly. Thence, a total of eight agents resulting from such implementation, predicts the flow at different lead times individually. The respective lead hours are analyzed, then paired, and averaged to produce the final 1, 2, 3, and 4-hour forecast predictions, which are accessible to the decision agent (AgentFL) to perform inference for each period about the flood-awareness conditions in the river.

5.4.1.2 Implementation of RF and SVR Agent Forecaster Leaners

From the ML algorithm selection experiments results described in Section 5.2, it was implemented 4 standalone RF and SVM(=SVR) regression models in R-Scripts and which were embedded into the 8 forecaster agent's species. The use of R scripts enables an agent species with an "RSKILL" behavior, a feature implemented in GAMA as an API (RCaller package) that allows calling R from the GAML code (GAMA User's Guide, Ver 1.8.1, p. 627). Regarding the selection of the hyperparameters for both models, RF models utilize 2000 trees and the SVR with the linear kernel and best tuning parameters of sigma = 0.015 and C = 2 for each prediction lead time test. These forecaster agents containing one model each for each lead-time receive the lagged data inputs and at initialization begins to perform the simulation of one to four hours ahead based on the direct multi-step ahead forecast scheme documented in [449].

The input data to these agent learners are created from the ABM model flood hydrographs simulation outputs time series, which, after being preprocessed by the AgentDPP are then reshaped by the AgentData2Lags as a supervised learning matrix data set according to the information identified in the autocorrelation analysis (Figure 4.12) showed to significantly correlate at 2, 3, and marginally 4 hours of rainfall, and water level on the streamflow at a given time and up to 2 hours of runoff lag, as described in Section 4.3. Feature engineering allowed the creation of a data matrix that consisted of 17 variables composed of the actual values and 4-h lags for the rainfall and water level variables, 2-hour lags, and 4-hour lead times for the streamflow variable (Table 5.3). Then all dataset matrices were partitioned by the agent learners into a proportion of 80:20 for training and testing instances, a process carried out internally within the agent forecasters.

To facilitate stability in the predictions, each experiment is run with a "10-fold cross-validation" approach with 5 repetitions and reported the mean over these results. Although the RF algorithm is good at dealing with model overfitting [450, 451], the adequate selection of parameters identified for 100 samples was 10 each, with the tuning parameter (mtry = 2). The extension of the SVM (used mainly in classification) has been extended to regression problems [452, 453], therefore making it suitable for forecasting numeric data (SVR), allowing the SVM's to model extremely complex data interactions.

The data set previously described is collected by the hydrometric sensor agents (AgentRNSn, AgentWLSn, and AgentSFSn) described in Section 3.2.1.1. However, for this data to be handy for use by the forecaster agents, it must be stored followed by a message issued by the AgentSV to the AgentHDBM. From thence, at the follow-up on a messaging request on data availability, the data must be processed in a supervised matrix form. This means, from the actual values, their lags and lead time must be created. Such a task is performed by the AgentData2Lags, as it receives information on the available stored data stream issued by the AgentHDBM. However, in the case of missing instances, these are estimated beforehand by the AgentDPP, who imputes the missing instances. Then, the AgentData2Lags proceeds on creating the lags on this imputed data set. Once this data flow and agent interactions are achieved (Figure 5.1), for the train and test sessions, both agent's forecaster group performs a forecast of the lead time flows that are compared to the observed ABM model flow hydrograph outputs.

5.4.1.3 Training the Forecaster Agents on the Train Dataset

Following the initial testing of the functioning of the system, which is initiated at simulation time with the hydrometric sensors having the belief that there is hydrometric data available at the sensors level (recall here sensors refers to the time series locally stored in files in csv format), and as such, they have the desire (their initial desire) to collect such data from the sensors, immediately, it is triggered, their plan to collect the data (see Table 5.1). Once this data is being collected it is requested for storage by the AgentSV to the AgentHDBM. However, if this is not the case, it means the data doesn't require preprocessing by the AgentDPP, this agent would drop the desire to preprocess data, then it would be further collected by the AgentData2Lags, whose role is to perform the creation of a lagged data matrix from the dataset, which is the appropriate form for the time series to be represented for supervised ML tasks. Next, this stored matrix is obtained and used as input data by the eight forecaster agents to perform a one to four-hour ahead lead time forecast of the streamflow variable. Therefore, the data matrix generated from the previous agent's interaction leads to the construction of a substantial dataset for the RF and SVR models, the implemented standalone RF and SVR algorithms embedded within the forecaster agents that forecast the absolute streamflow for the relevant time horizons (i.e., target variable q = $\delta + k$).

To investigate the proficiency of the forecaster agent's hourly streamflow forecasts from 1 to 4-hours lead-time, training is performed by the simulation with all interacting agents in the GAMA model that was built initially under Chapter 4. As mentioned earlier and referenced previously in [449], from the four existing strategies for "multi-step-ahead" forecasting, the "direct multi-step-ahead" strategy was chosen as the method of forecasting for the forecaster agents, given its simplicity of implementation, though it is required of this method to build separate models for each forecast time horizon, and it may seem as a computational burden, contrary to the recursive approach which uses a "one-step-ahead" model several times; however, a setback of this approach as it relies on the previous forecast to predict the next time step is the inclusion of error propagation, which can degrade forecasting accuracy as the prediction time step increases over time. Another reason for not employing this method is to avoid the "blackboard" pitfalls, which would have allowed the forecasting task to be reduced to a single agent that could have been implemented with this method.

Upon completion of a simulation cycle, some eight records are generated containing the values of the flow forecasts made by each of the agents belonging to the two groups of forecasting agents for each of the assigned time horizons. This results in a duplicate of each lead-time forecast, which adds up to a total of eight forecast files that are then merged and averaged by the AgentFL to produce the corresponding final and single flood forecast for $Q_{\delta+1}$ to $Q_{\delta+4}$ ahead along with the computed flood-awareness level for each forecasted horizon.

A comparative analysis of the statistical metrics results on the performance of the forecasts made by both forecasting agent's groups (Table 5.9), for the training cycles revealed that for both agent's groups implementation, the forecasting efficiency of the models in both cases depends on the duration and breadth of the prediction horizon, rather than the data entry, as the information shows the further the lead-time is extended the forecasts proficiency decreases. Hence, this peculiarity of time series has been well documented by some researchers who had reported the cost that some models can undergo when forecasting at very extended periods, weather extremes, and data shortage [454–457]. However, it can be noted from this information, that no more than three hours lead-time is an appropriate forecast window, provided the previous lag values should be applied for this scenario, as it was seen later in the results for the simulations runs with the other storm cases used in the MAS hydrologic model validation task.

The information that is shown in Table 5.9 and the scatterplots (Figures 5.16 and 5.17) between the observed ABM model flood hydrograph and the forecasted by the MAS model agent's forecasters have a high level of connection. The MAS model agent forecasters were capable of satisfactorily simulating the hourly streamflow (Figures 5.14, and 5.15) with values of the R^2 for the AgentFCST1 group in the range [0.78, 0.82], and the lowest values for the group AgentFCST2 in the range [0.58, 0.69], respectively. The correlation for the 4-hour lead-time forecaster agent under group 2 showed the lowest value (r = 0.76), and that about a forecast accuracy of 58%, and an underestimation of the ABM observed peak flow (Q_{pk} = 465.5 $m^3 \cdot s^{-1}$) of 25.5%. In contrast, the forecaster agents under group 1 exhibited a strong positive correlation (r = 0.90 and relatively good $R^2 = 0.82$) for the prediction at that same hour, and for almost all hours the correlations were not <0.8. In terms of model proficiency, the RMSE, which measures the spread between the observed and calculated values, was 33.6% higher for the AgentFCSTR2 than it was for the group AgentFCSTR2.

Summarizing on the training session both forecaster's agent's groups showed best forecasting accuracy for the first three hours of forecast with values in the range [67, 82%], and shows the 4-hour lead-time streamflow peak forecast to be less biased for the forecaster agents of group 2. Regarding the simulated flood hydrograph, both groups of agent's forecasters portrayed an acceptable approximation to the shape of the flood discharge curve (see Figures 5.14 and 5.15); although, for the instances, where they revealed some difficulties in either achieving or slightly coming close to the value of the ABM observed extreme flow (Q_{pk} = 465.5 $m^3 \cdot s^{-1}$), nonetheless, there is an exception for the agent's forecasters of group 2, that significantly showed to fall short some 26% of the magnitude of the observed Q_{pk}. This corroborate to show after all that the SVR algorithm is good at the prediction of extremely complex data, has compared to those of group 1.

Table 5.9: Performance metrics on train data for agent forecast models at 1 to 4-hour lead forecast simulations.

	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	ABM Obs. Q_{pk} [$m^3 \cdot s^{-1}$]	MAS Sim. Q_{pk} [$m^3 \cdot s^{-1}$]
AgentFCST1						
Simulation Period						
q1h	0.90	0.81	30.2	-40.3	459.1	274.1
q2h	0.90	0.80	31.5	-38.5	465.5	286.9
q3h	0.88	0.78	32.1	-42.1	465.5	269.5
q4h	0.90	0.82	21.9	-32.2	258.6	175.5
AgentFCST2						
Simulation Period						
q1h	0.83	0.69	36.6	-32.6	465.5	313.7
q2h	0.83	0.69	37.9	-31.3	465.5	320.0
q3h	0.82	0.67	36.6	-37.4	465.5	291.3
q4h	0.76	0.58	43.4	-25.5	465.5	346.8

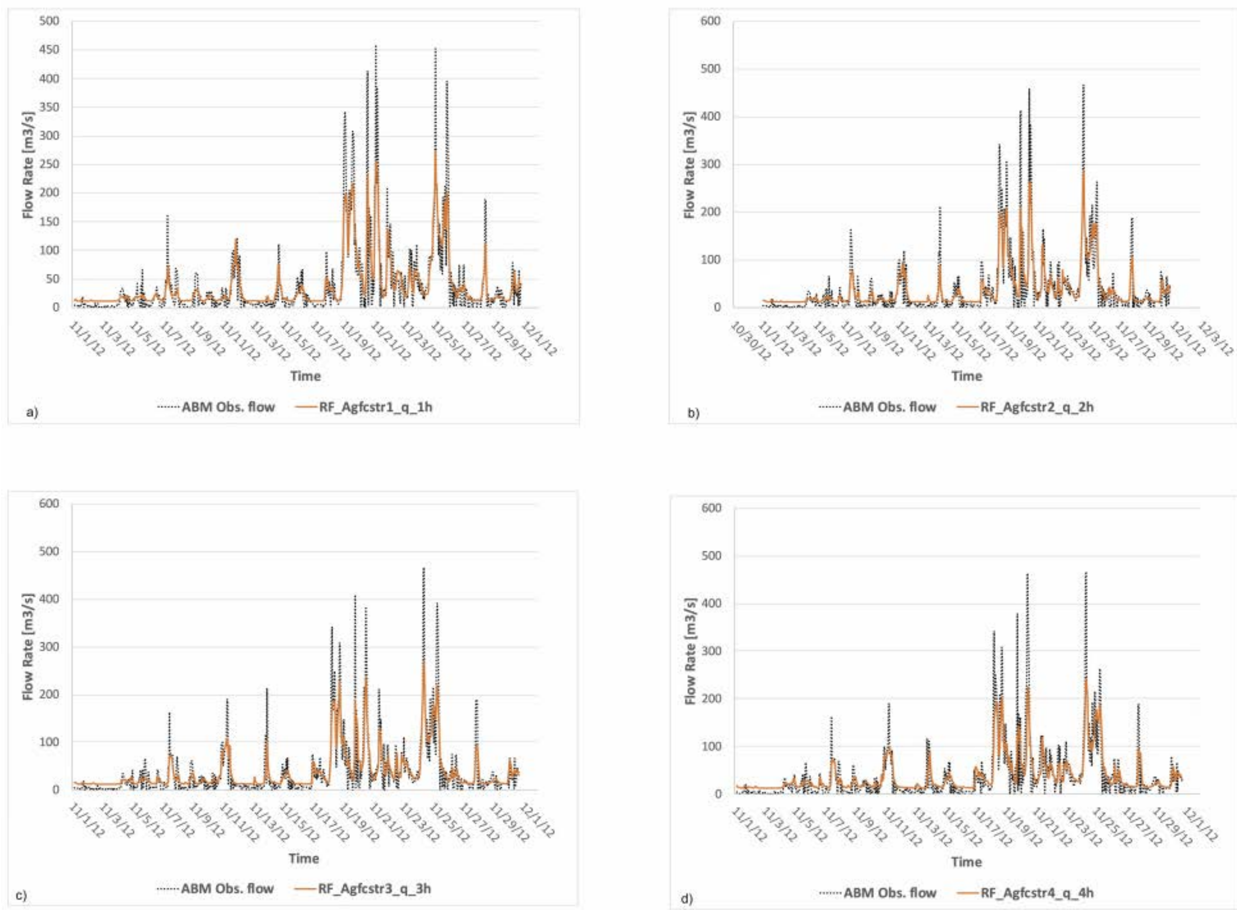


Figure 5.14: Comparison of the training session for the MAS AgentFCST1 lead-time flood forecasting simulated vs ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

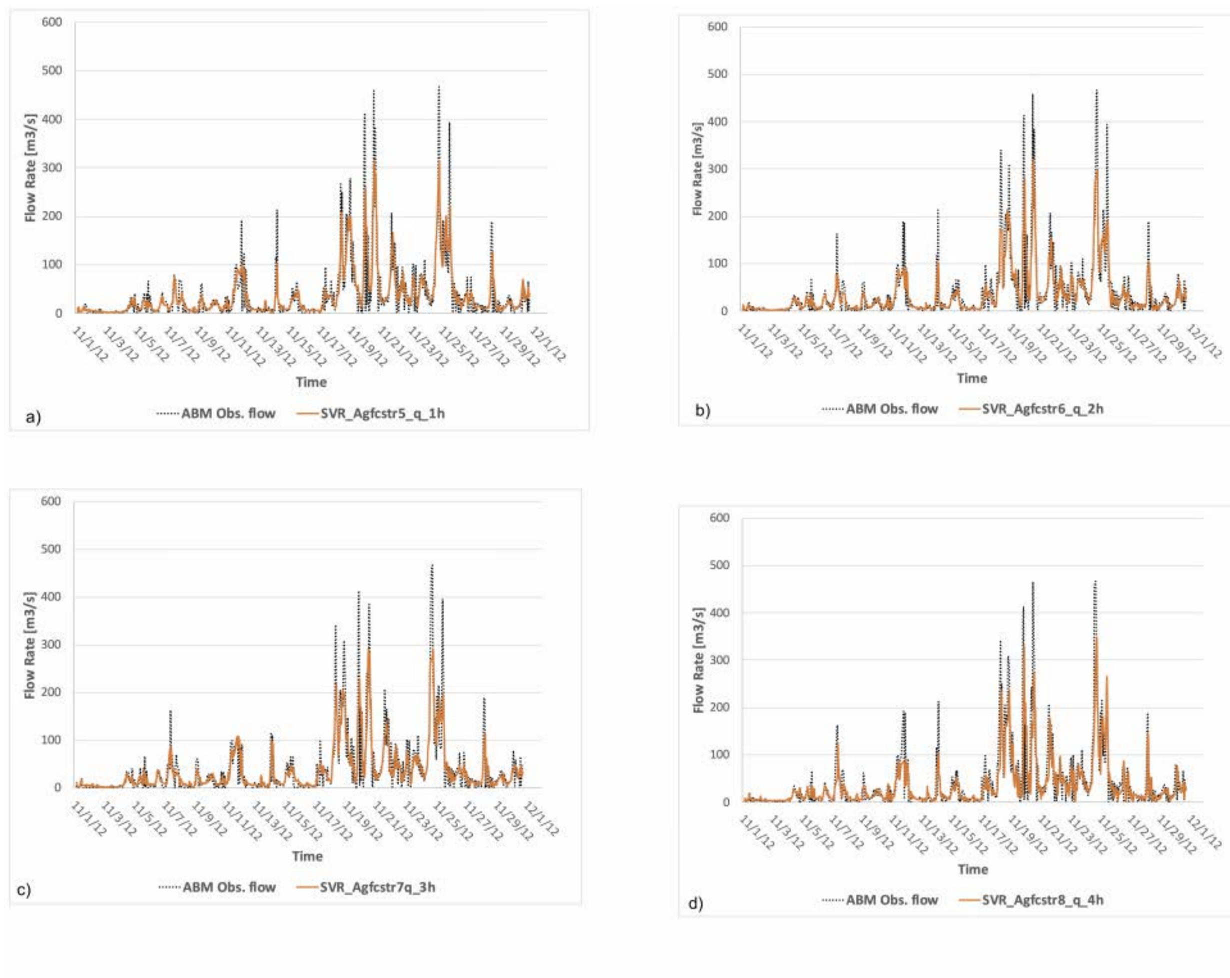


Figure 5.15: Comparison of the training session for the MAS AgentFCST2 lead-time flood forecasting simulated vs ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

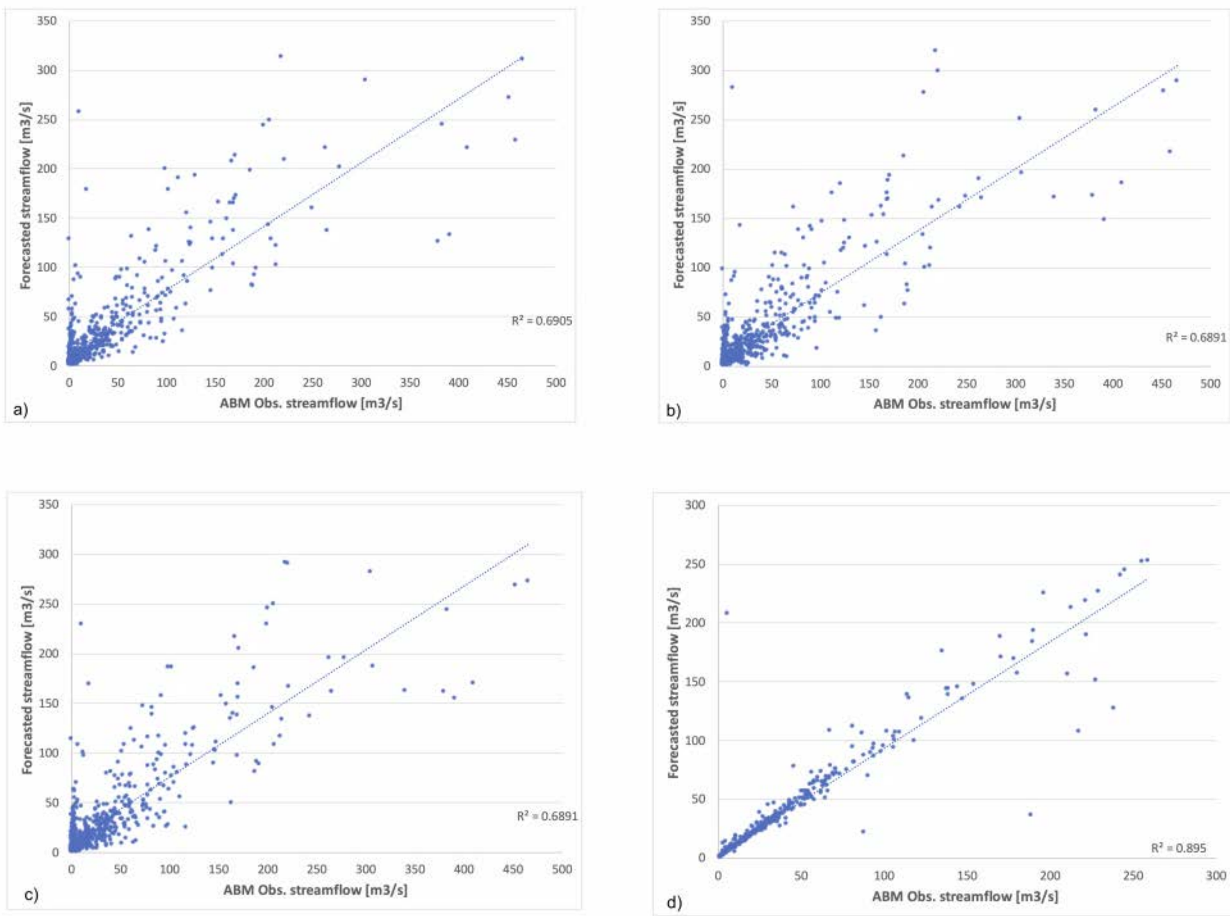


Figure 5.16: Scatterplot diagrams with fitted regression line of training session for the MAS AgentFCST1 lead-time flood forecasting simulated and ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

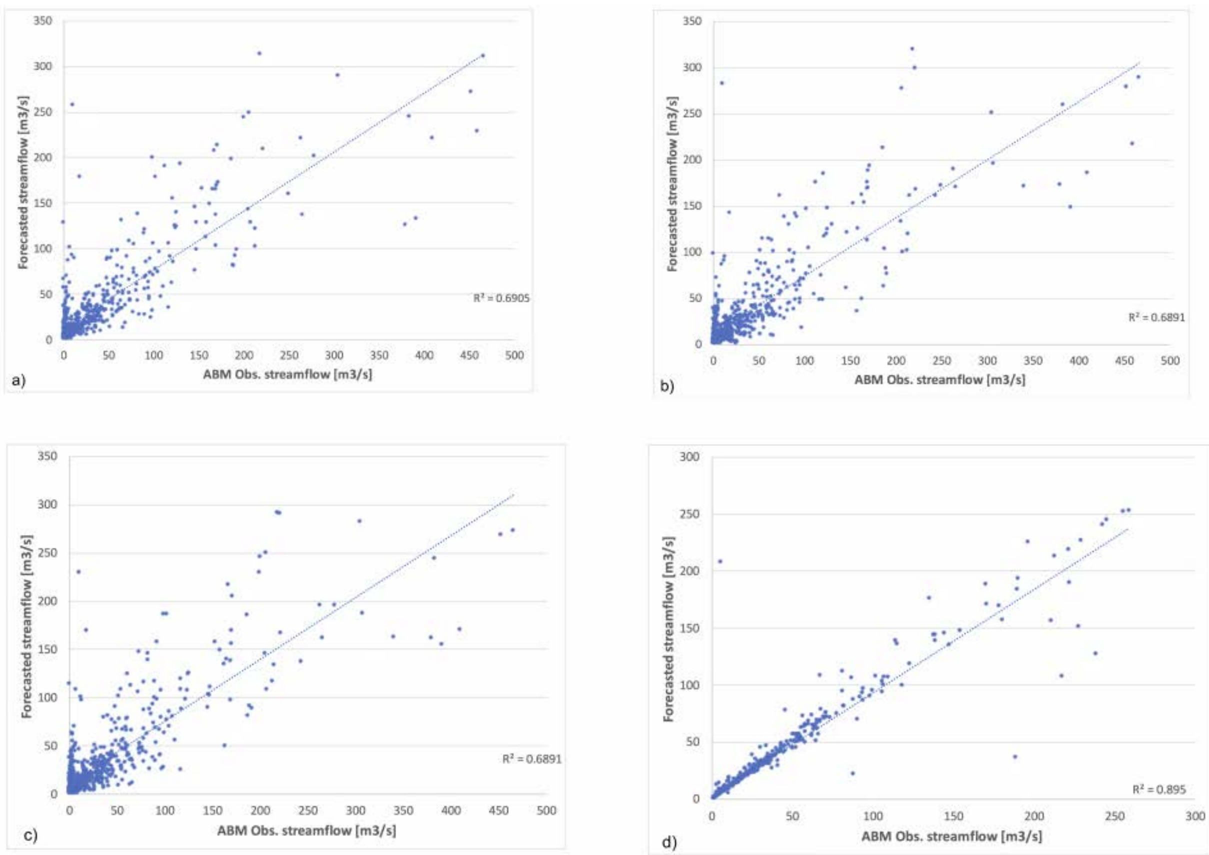


Figure 5.17: Scatterplot diagrams with fitted regression line of training session for the MAS AgentFCST2 lead-time flood forecasting simulated and ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

5.4.2 Experimental Results: MAS Hydrologic Model Setup

This section discusses the verification results for the implemented standalone RF and SVR algorithms that were embedded into every four sets of agents grouped under the identifier agents Forecaster1 (AgentFCST1) and Forecaster2 (AgentFCST2) as previously explained in Section 3.2.1.4. In the previous experimental sessions, with the MAS model forecaster agents were trained on the ABM synthetic data resulting from the simulated flood hydrographs outputs, which were partitioned into training and testing dataset; therefore, in this occasion to test the efficiency of the agents to make a forecast on unknown data for one up to four hours ahead it is used the remaining test data partition.

5.4.2.1 Testing the Forecaster Agents on the Test Dataset

The testing was performed following the same methodological approach that was done in the training sessions, with the sole difference that it used a separate data set unknown to each model and it was reported the performance metrics values of these tests. The statistical analysis of the AgentFCST2 on the unknown data (Table 5.10) showed that about 64% of the prediction accuracy with a RMSE = 32.6 $m^3 \cdot s^{-1}$ was for the 3-hour forecast, while 72% with a RMSE = 35.4 $m^3 \cdot s^{-1}$ was for the AgentFCST1 for the 3-hour lead-time forecast. However, for both forecasters groups, 52% of the prediction accuracy was for the 1-hour forecast with a 45% error increase of the RMSE by the AgentFCST2 more than the AgentFCST1. In addition, the results for the AgentFCST2 were much worse, with about 41% of the prediction accuracy and with an RMSE = 42.7 $m^3 \cdot s^{-1}$ for the forecast at lead-time 4-hours.

Critiquing, these results indicate that the agent group AgentFCST1 outperforms the group AgentFCST2, at long-range predictions, with the AgentFCST1 compensating to produce the best calibration results for the agent learners.

5.4.3 Conclusions

Recapping the section, all models displayed excellent performance during the training session. This shows that model efficiency is improved the closer R^2 tends to unity. However, this assumption does not always hold to be true since this coefficient is a performance statistic to be handled with precaution, as it can probably induce it to 1 by the mere addition of terms to the model. For the test session, the agent's forecasters of group 2 have shown their ability to fall short a few magnitudes to the ABM model observed Q_{pk} for each resulting test data combination, albeit underestimating this flow. An illustration of the forecasted streamflow hydrograph is shown in Figures 5.18 and 5.19 show the forecasted streamflow hydrograph and their generated scatterplots (Figures 5.20 and 5.21) for the eight forecaster agents.

Table 5.10: Performance metrics on test data for agent forecaster models at 1 to 4-hour lead forecast simulation.

	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	ABM Obs. Q_{pk} [$m^3 \cdot s^{-1}$]	MAS Sim. Q_{pk} [$m^3 \cdot s^{-1}$]
AgentFCST1						
Simulation Period						
q1h	0.72	0.52	34.1	16.7	206.0	240.4
q2h	0.85	0.72	35.4	-50.5	452.1	223.7
q3h	0.75	0.56	44.3	-51.3	465.5	226.9
q4h	0.77	0.59	32.0	-30.8	339.9	235.3
AgentFCST2						
Simulation Period						
q1h	0.72	0.52	49.3	-38.1	452.1	280.1
q2h	0.73	0.54	41.2	-39.7	465.5	280.5
q3h	0.80	0.64	32.6	-13.3	339.9	294.8
q4h	0.64	0.41	42.7	-18.1	339.9	278.4

Finalizing, all forecasts agent models accuracy, for both train and test sessions, range in [41, 82%] of the variability in the test dataset, being the highest and lowest percentage reported for the 4-hour lead-time forecast simulation that corresponds to the agentforecaster4 of group 1 and 2, respectively. This could infer that no more than four hours of lead time is adequate for these models.

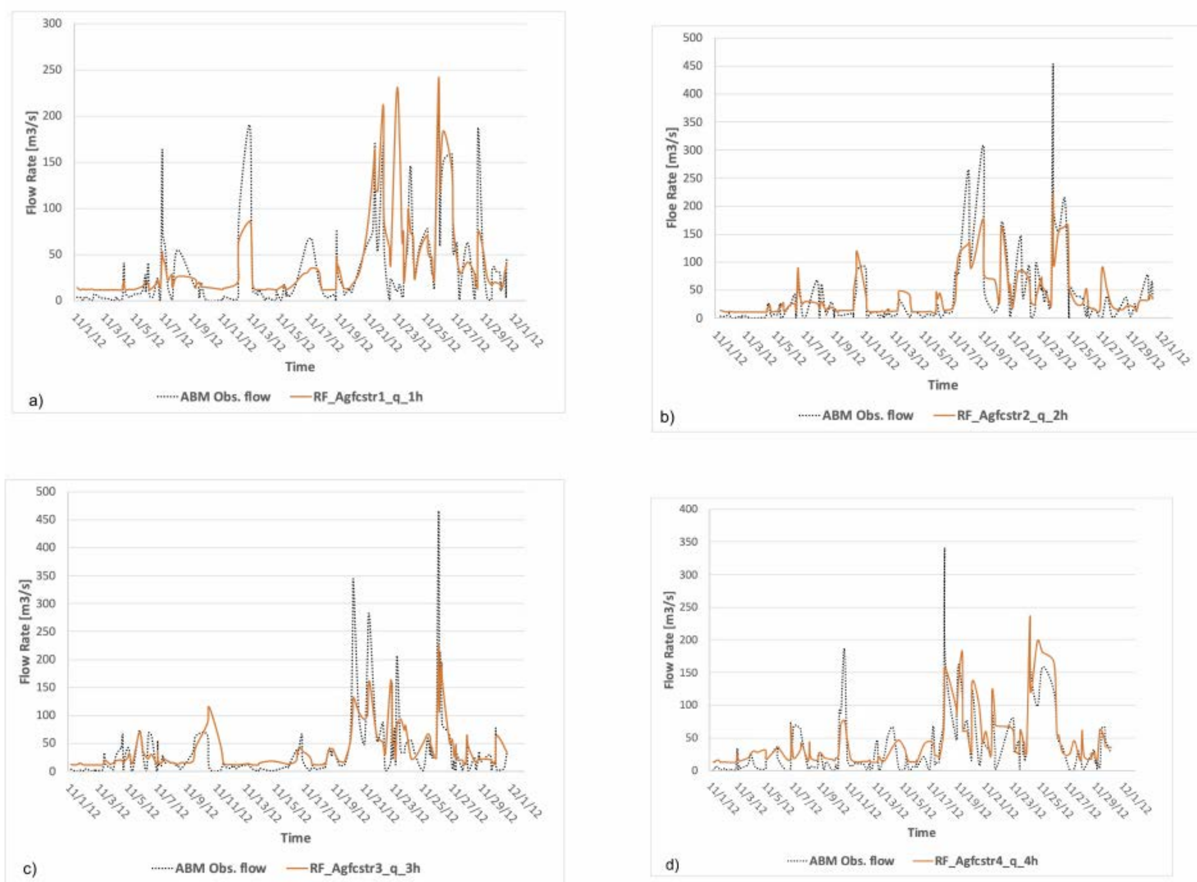


Figure 5.18: Comparison of the testing session for the MAS AgentFCST1 lead-time flood forecasting simulated vs ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

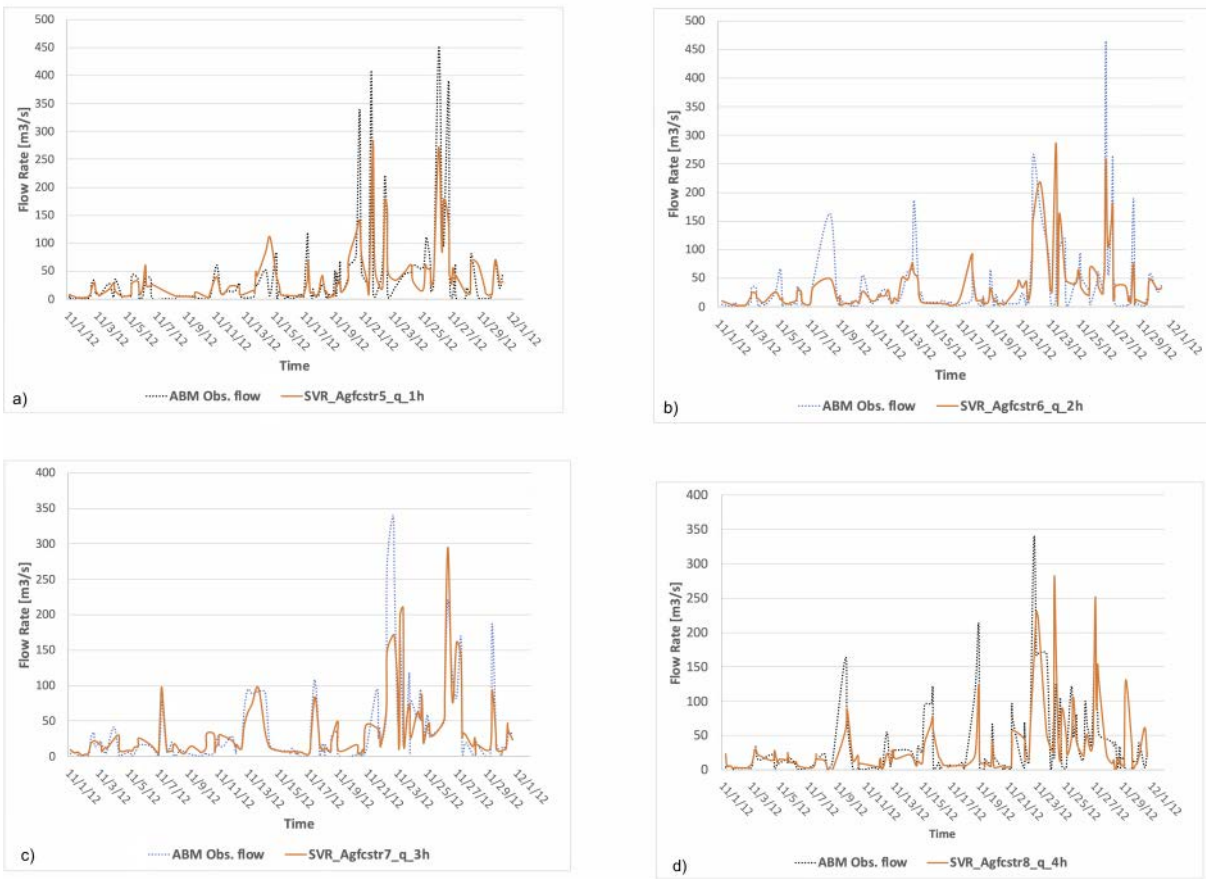


Figure 5.19: Comparison of the testing session for the MAS AgentFCST2 lead-time flood forecasting simulated vs ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

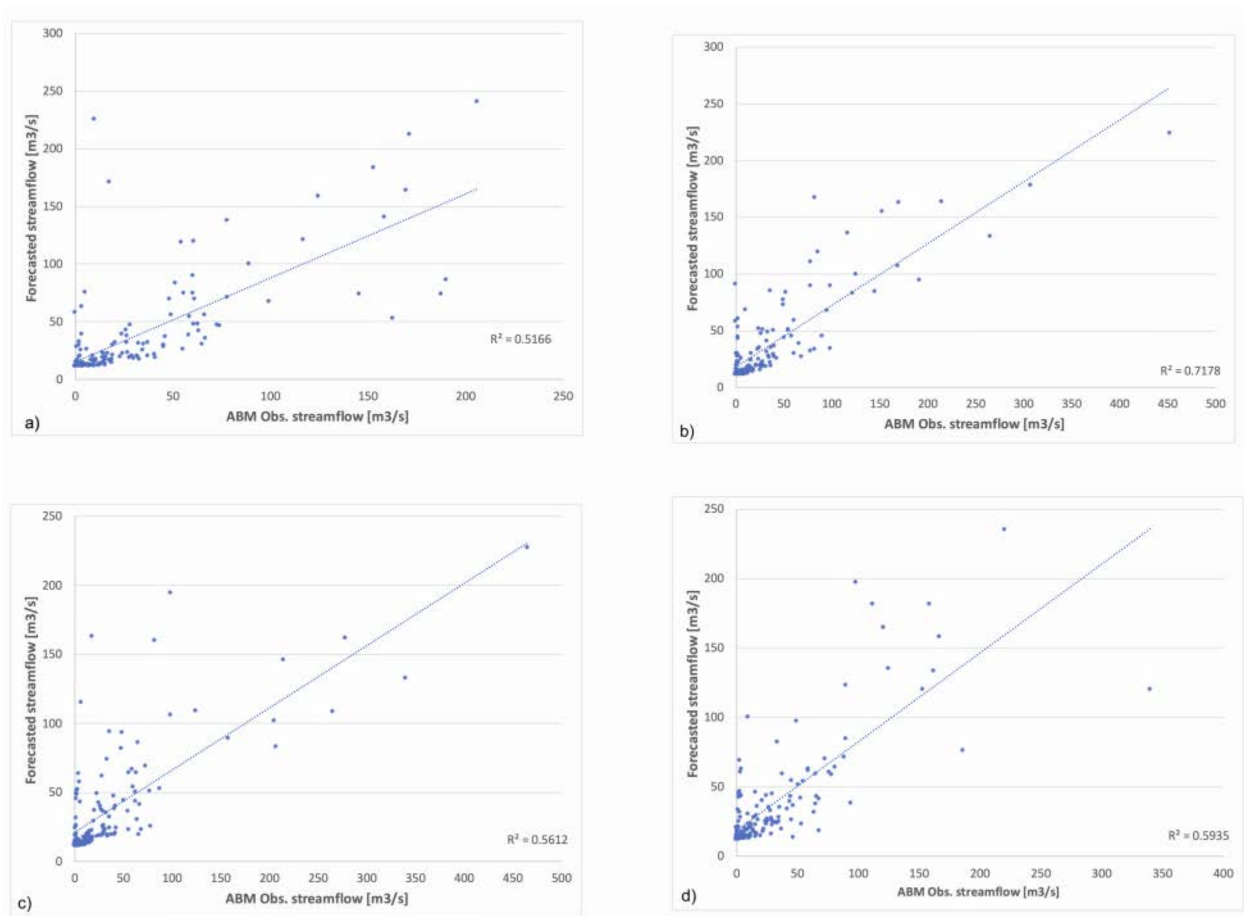


Figure 5.20: Scatterplot diagrams with fitted regression line of testing session for the MAS AgentFCST1 lead-time flood forecasting simulated vs ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

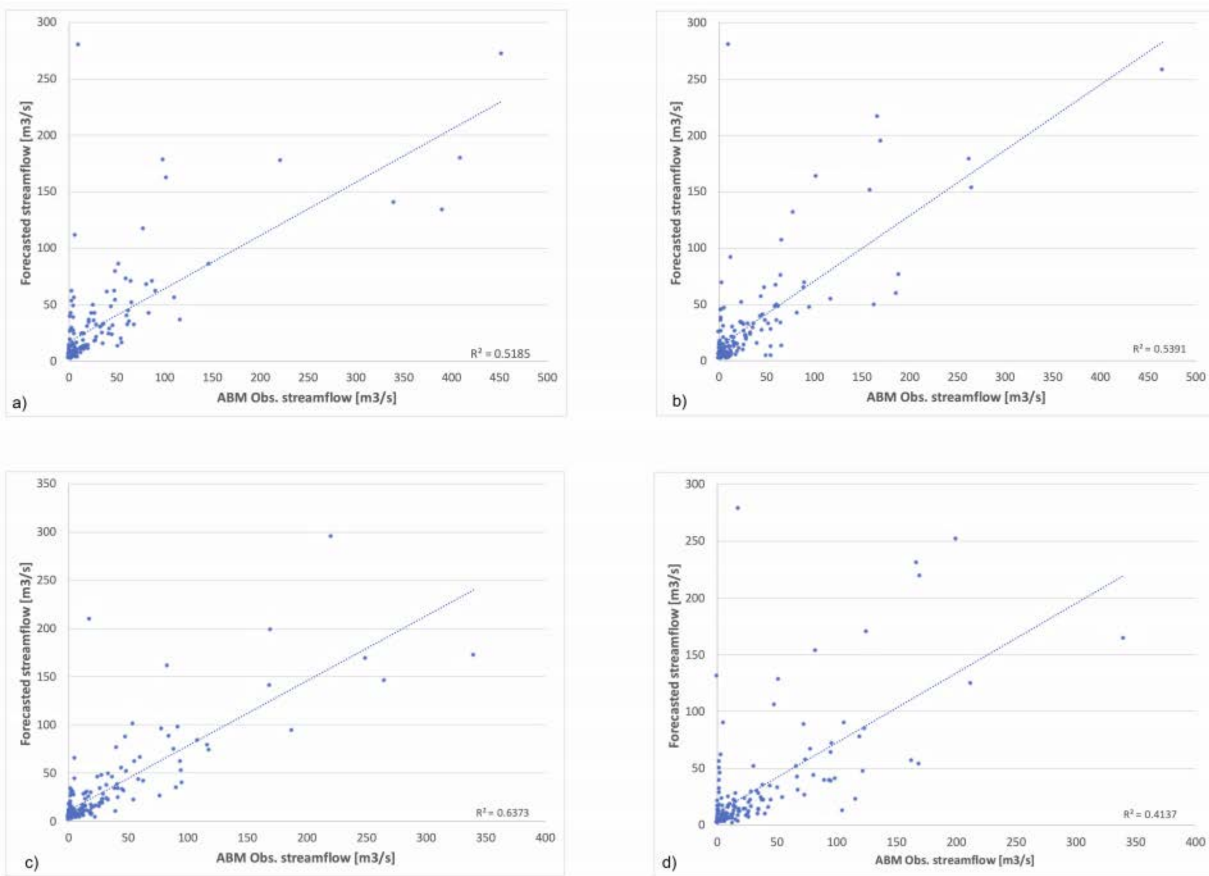


Figure 5.21: Scatterplot diagrams with fitted regression line of testing session for the MAS AgentFCST2 lead-time flood forecasting simulated vs ABM observed hydrograph, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

5.5 Validation of the BDI Hydrologic MAS Model

In the present section, it is provided the results for the experimental runs to conduct an evaluation of the MAS model framework for flood forecasting. In the current framework, the forecasting of flood hydrographs along with the corresponding level of the computed flood-awareness with lead times from one and up to four hours ahead is simulated with the entire MAS hydrologic model setup comprising of the the hydrometric sensor agent's, data preprocessing agent's, hydrometric sensor verifier agent, database management agent, forecaster agent's, user interface agent and a decision-making agent with four validation storm scenario events. The experimental evaluation was undertaken to contrast the outputs of the synthetic data resulting from the ABM simulated validation storm hydrographs for December 2012, December 2014, May 2015, and November 2015 with those of the MAS hydrologic model.

According to Fortino et al. [458], in the MAS development, the goal of simulation corresponds to the validation phase in which the model is tested before the deployment of its end purpose. During this phase, the qualitative or quantitative outputs of the model can serve as information that can be used by the modeler to correct the model and provide additional fine-tuning to improve its proficiency. However, despite the little progress made for the last few years in the application of ABM and MAS applicable to hydrologic related problems, there is still limited information regarding studies in this area by the hydrologic community and that, furthermore, one of the various reasons for these limitations revolved around the lack of consensual and objective methods and methodologies for the validation of such models, as some might argue [381].

In the Medio River hydrological network, there are two hydrometric stations with three sensors (e.g., rain, water level, and streamflow), one upstream, and the other downstream of the catchment. However, only the upstream station was used for the application because the downstream station was severely vandalized. In principle, no hydrological models were found to be applied because the essential information about the river evolution upstream and downstream are missing.

The, the Medio River flood forecasting process within the MAS hydrologic model can be detailed in the subsequent sequence of stages:

1. Capture hydrometric data. This is the upper level of the MAS when at the initialization of this sequence of flow forecast dynamics, the hydrometric sensor agents have the intention `read_rain_sensor_data`, `read_waterlev_sensor_data`, and `read_flow_sensor_data` at each hour. To guarantee the end-point control of the captured data, the agents must execute their algorithms to check and consequently verify for missing data, which if true, will trigger their plans to report at each hour the current condition

to the AgentSV, who with the intention `verify_sensor_status` inquire from the sensor agents and then issues a message to the AgentUI, with intention `view_system_msg` who will conduct a follow.

2. Data Pretreatment. Data pretreatment is carried out locally and in parallel at each data pre-processing agent, the AgentDPP who performs data preprocessing and imputations, and the AgentData2Lags agent who creates a matrix with hourly lagged values, and the resulting post-processed data is stored locally.

3. Data storage. While the incoming raw data is been captured and/or preprocessed, reformatted and inserted into the database, the AgentSV, request from the AgentHDBM whose initial intention is to `store_all_hydro_data` store the incoming raw post-processed data into the database server. The AgentHDBM stores all data products produced by the other agent's process.

4. Forecasting. Through the intention (`do_flow_predict`) the flow forecasting task is launched by the eight AgentFCSTR's equipped with the ML models to perform an hourly forecast of 1 to 4 hours using the resulting hourly matrix that was created and stored in the data pre-processing stage. The forecaster agents compare their previous forecast with the last one and refresh the previous forecast each time step (e.g., $F_{\delta+1}$ with F_{δ}).

5. Decision-making. With the creation of the AgentFCSTR's and the results of the hourly forecasts that they deliver, the intention `do_flow_inference` is carried on by the decision-making agent. Once he had satisfied his intention (`DBConnection`) to ensure database connectivity, at every simulation step, the agent interacts with the AgentFCSTR's and the database agent to receive the desired forecast data, evaluates the data received, makes final decisions based on the forecasts, and conducts the following tasks:

    - Aggregates the forecasts and computes the averages between each group's forecasts to produce an absolute value of results for each forecast lead-time.
    - Computes the fuzzy inference from the observed and forecasted flows at each lead-time.
    - With the method (`clas_data_storage_request`), execute the intention to request the forecast and inference results to be stored into the database server by the database agent.
    - Display the actual flow and the corresponding lead-time flow forecast for each flood-awareness levels.

6. Reporting. With the `view_system_msg` intention, the AgentUI reports to the user the information on the state of the conditions of the sensors received from the AgentSV messages, and the flood-awareness levels received from the decision-making agent messages.

As a reminder, in Chapter 4, it was conducted an exhaustive analysis with two different hydrological modeling approaches as an initiative to predict the phenomenon of floods produced by the effects of climatic alterations of the hydrological processes in tropical catchments. Therefore, for the first of these approaches, hydrological modeling was used through conventional methods, using the the HEC-HMS modeling tool. The second approach was applied with an implemented agent-based model (ABM) to perform the same hydrological modeling task as part of this research aim and as a more viable modeling alternative for decision-makers and the less expert, with less of the physically oriented and mathematical complications encountered in the implementation of the conventional modeling tools. From the experiments executed, it was noted the results from the assessment between the two-modeling approaches to show both approaches to give acceptable results, especially for the case of the ABM which in principle is not a standard hydrologic modeling tool per se, and that it forms the bases for the main idea of this endeavor. However, they both showed to approximating the shape of the curve of the observed hydrographs at the moment of each model's execution, and it was found also they either underestimated or overestimated the observed Q_{pk}. Despite that, in either of the cases, based on the results, both models showed capable of dealing with the flood hydrograph approximation.

From this point on, it is set the research goal with a series of experiments undertaken within the MAS model framework as a hydrologic modeling alternative and with which when adding cognitive agents and endowing them with decision-making capabilities, can maximize the expected results. Generally, the assessment on the agent-based concept is pointed toward evaluating the behavior of each agent species, system organization, and agent interactions. Nonetheless, although the former holds valuable, the type of domain of study, given its degree of complexity in which an attempt with ABM/MAS is used, does not occasionally conform to strict mathematical and physical laws that govern the system and consequently must not be taken for granted during the evaluation phase. Therefore, in the modeling of environmental processes, the analyst should be concerned as well with the understanding of the abstraction of the effectiveness of the ABM or MAS performance in simulating such physical environments.

At the initialization of the sequences of the flow forecasting dynamics, each agent in the MAS arrangement is set with an initial state of desire(s). The agent "AgentSV" for example is initialized with its three-desire state in the following order: `DBConnection`, `verify_sensor_status` (as his general desire), and `request_raw_data_storage`. However, the dynamic starts after the creation of the three hydrometric sensor agents which at the beginning of the simulation are initialized with the desire to do `read_rain_sensor_data`, `read_waterlev_sensor_data`, and `read_flow_sensor_data`, which are their general desire. Subsequently, the `report_missing_sensor_data` and `no_missing_data` are initialized. Thereafter, the AgentHDBM with is first general desire `test_parentDB_connection`, AgentUI set with initial general desire `DBConnection`, the data pre-processing agents with general desires `impute_raw_data` and `create_lagged_raw_data`, forecaster agents with `DBConnection` and `do_flow_predict`, decision-making agent with `DBConnection` and `do_flow_inference`, accordingly.

In the execution of the experiments carried out for all storm cases, in GAMA the process of initialization of the agents are the same, except for the cases when their initial desires are dropped and replaced by another depending on the current circumstances (e.g., when the need arises for agents to perform some job given to a request if not their perceptions of the environment have changed) and weights and priorities can also be assigned to the desires.

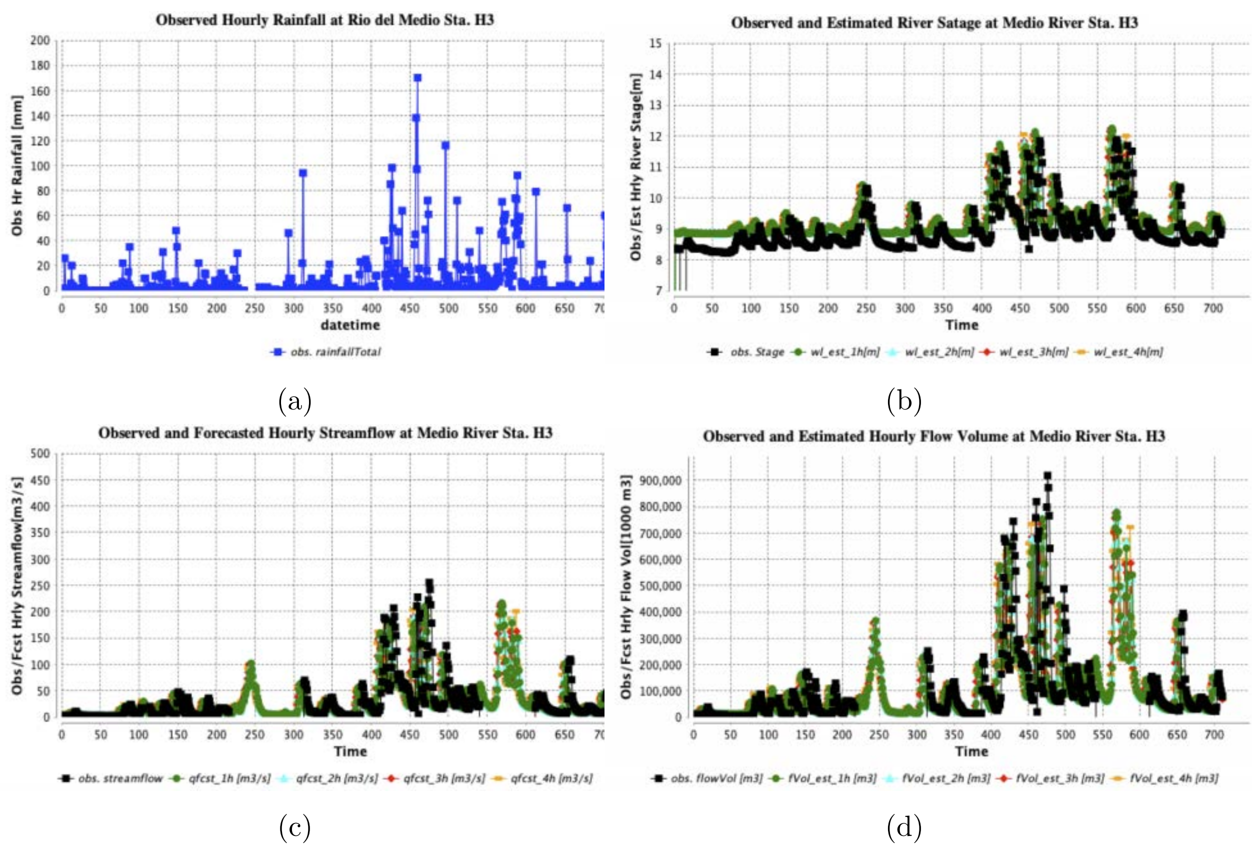
5.5.1 MAS vs ABM Synthetic Data of December 2012 Storm

Overall, in this and every other simulation runs, the hydrometric sensor agents of both modeling approaches perceive and process the information of the environment, namely the precipitation, the stage, and the flow variables, and begin to read this information.

The sensor agents in the ABM model, simply read the precipitation data, whereas, in the MAS model, they do not only read and capture this information but can identify errors and missing instances in the data, and consequently, can reason if to execute some behavior to deal with the problems in the data. In this sense, they report the data incongruences and can coordinate with another agent in the system whose role could oversee and correct the erroneous instances (i.e., AgentDPP).

To test the capabilities of the MAS hydrologic model, for example, the hydrometric sensor agents when in communications with the data imputation agent, can coordinate with each other when the intentions of the sensor agents are reporting missing instances, then messages a request to the data pre-processing agent to impute those missing instances. Figure 5.22 shows for example where several periods of missing instances have been imputed and therefore given way to the data lag agent to create the lagged matrix for the forecasting agents to provide a forecast for those periods, and along this sequence, provide the decision agent to make inferences despite the missing periods (see Figure 5.23). These capabilities are not present in the hydrometric sensor agents of the ABM model approach, therefore giving way to the possibility for that model to provide data forecasted on noisy instances (missing periods).

The hydrometric sensor agents in the MAS model being cluster head-nodes behaves as an autonomous process that can reason about the hydrometric variables captured from the environment and make corrective decisions in cooperation with other agents concerning the quality of the data before this one is dispatched to the other members of the node. This means, that the active real-time verification of the quality of the data by each sensor agent, implies an hourly checking by which data issues could be possibly corrected before reaching the flow forecasting and the inference phase.



(a) (b) (c) (d)

Figure 5.22: Example of an output showing the MAS model on a 4-hour forecast simulated hydrograph with imputed instances for November 2012 Storm Events: (a) Observed hourly hyetograph with missing periods [mm], (b) Observed and estimated hourly river stage [m], (c) observed and forecasted streamflow $[m^3 \cdot s^{-1}]$ with missing and imputed periods and (d) Observed and estimated flow volume $[m^3]$.

In hydrologic modeling, the common practice to analyze and validate a hydrograph resulting from a model is to apply standard metrics such as amplitude errors and their associated functions; examples are the "Root Mean Square Error (RMSE)", "Nash-Sutcliffe efficiency (NSE)" [459], or metrics such as the "percentage error in peak discharge (Q_{pk})" [86] which measures asymmetry between the observed and estimated flows, that is by how much percentage thus the model overestimates or underestimates the observed peak discharge.

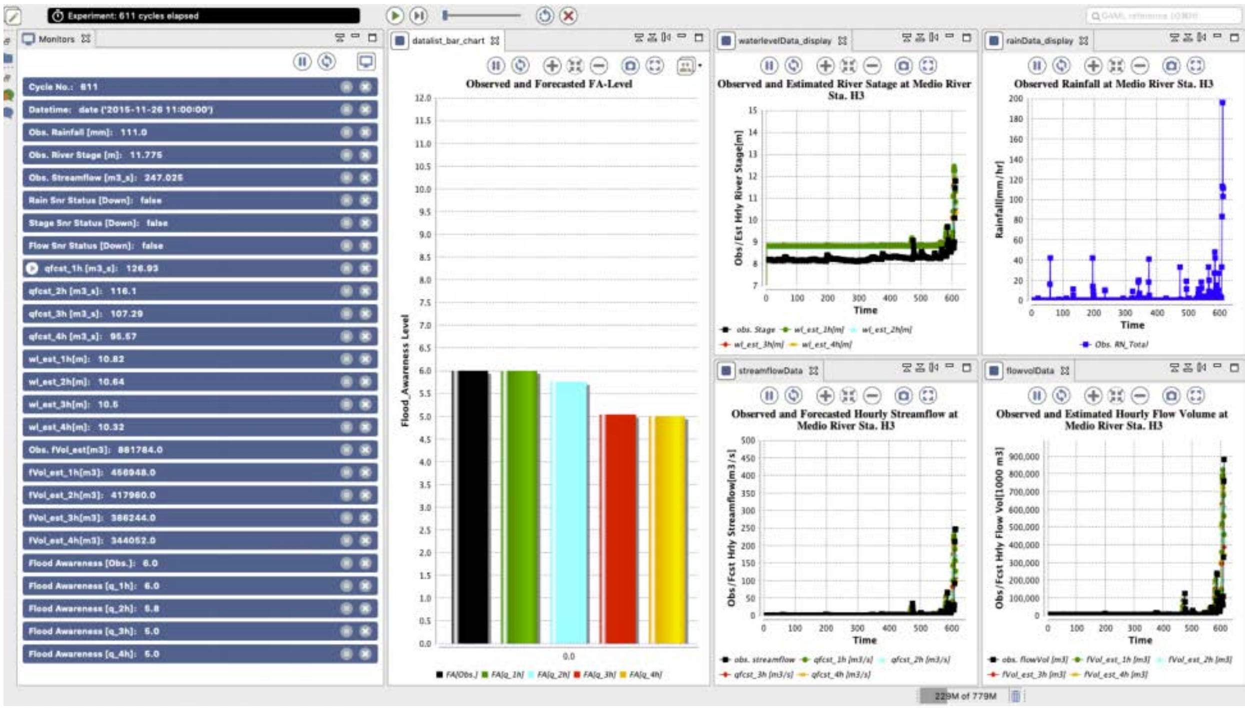


Figure 5.23: An example of a screenshot showing the numeric outputs from the forecasters and decision-making agents as they are interacting on the November 2015 storm validation session: (left) Simulation output variables, (center) Flood-Awareness Levels, and (right) observed and estimated stage [m], rainfall [mm], observed and forecasted flows [$m^3 \cdot s^{-1}$] and estimated flow volume [m^3].

Therefore, the approach that was used to validate the MAS hydrologic model is by assessing the variations between the ABM simulation outputs and the output results from the MAS. This process comprises the use of the Pearson Correlation Coefficient (r) as the similarity measurement and the RMSE as the measure of the differences between the MAS model simulations and the ABM observed outputs. In addition, to check the significance of the asymmetry between forecasting results from the MAS model and those from the ABM, it was performed a "t-test" analysis assigning alpha at the 0.05 level of significance. So, if $t > 0.05$ df and $p \leq 0.05$, hence the level of significance is high, and the MAS model is good.

Provided the information captured from each hydrometric sensor by the sensor agents, the system is started, and it can run a complete simulation scenario of the precipitation-run-off water process within the abstraction of a catchment environment. The resulting flood hydrographs obtained from simulating the input data of the ABM simulated December 2012 storm hydrographs outputs with the MAS hydrologic model, are then compared as shown in Figure 5.24 summarizes the statistical measures of these comparisons.

For this validation scenario, the table shows the metrics that resulted from the grouping and the computation of the average of the agent forecaster's simulations forecast in the MAS model by the decision agent to generate each of the final lead-time forecasts. It can be noted for the 3-hour forecast, strong important coefficient of similarity (r= 0.89), with p <.001 between the observed ABM and the MAS simulated peak flow, subsequently it can be seen good coefficient of determination R^2 = 0.80. This R^2 explains that only 80% of the total variability in the MAS forecasts, forecasted by the agent forecasters can be accounted for by the variation in the outputs of the ABM for that lead-time. Therefore, for this time frame, the highest accuracy of the MAS model forecast was only 80%.

As shown by Figure 5.24 the MAS simulated hydrograph show to conforms pretty well to the overall shape of the ABM output hydrograph curve, when compared, except for the visible instances of the mismatches shown by the biases in the peaks, it shows that there is some marginal correlation of the shapes for each of the simulated time frames. For the total storm episode, the amplitude of the errors between the two models, show ranging [26.2, 51.2 $m^3 \cdot s^{-1}$], being the highest value recorded for the 1-hour ahead forecast, while it recorded the smallest value for the 3-hour ahead forecast.

For an observed ABM model value of the peak flow at each forecast period, the MAS model was less at the 2-hour and 3-hour forecast and larger at the 1-hour and 4-hour forecast. The average error in peak discharge was ±13.0%, indicating that it agreed with the error criteria. The proportion of variance explained by (R^2) between observed (ABM) Q_{pk} and simulated (MAS) Q_{pk} is in the range [51, 80%], and with (r) in the range [0.71 – 0.89, p <.001] implying that the MAS model performed significantly good during this storm period as compared to the ABM, and as shown by the p-value. Likewise, for the observed and forecasted flows of this period, for each lead-time forecast, the MAS generated the level of awareness for flooding, on average observed overall for a value of 5.0, each hour on the verge of intercepting with the possibility for moving from the "CAUTION" to the "ALARM" level as shown in Figure 5.11.

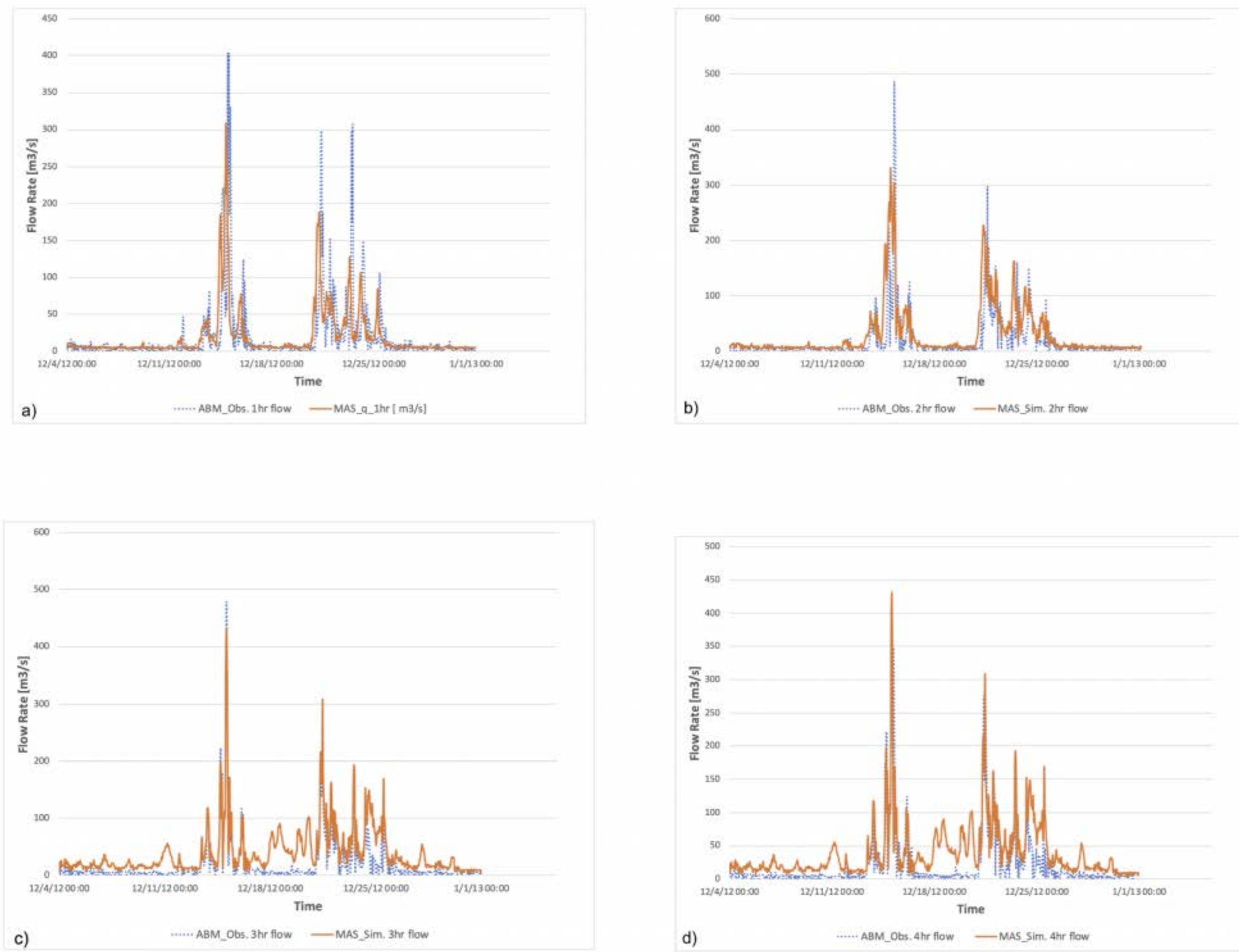


Figure 5.24: Observed (ABM) and simulated (MAS) hydrographs comparisons with the ABM synthetic data of December 2012 storm for time-horizons, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

Table 5.11: Verification metrics for 4-hour lead-time forecast between the observed ABM synthetic December 2012 storm hydrograph and simulated MAS hydrograph, along with the computed flood-awareness.

	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	Obs. (ABM) Q_{pk} [$m^3 \cdot s^{-1}$]	Sim. (MAS) Q_{pk} [$m^3 \cdot s^{-1}$]	FA
December 2012							
Simulation Period							
q1h	0.79	0.62	51.2	4.1	404.0	420.8	5.0
q2h	0.75	0.56	39.4	-31.1	479.5	328.7	5.0
q3h	0.89	0.8	26.2	-10.1	479.5	431.3	5.0
q4h	0.71	0.51	36.2	6.8	404.0	431.3	5.0

5.5.2 MAS vs ABM Synthetic Data of December 2014 Storm

Following the same approach for the analysis of the flood hydrograph carried out in the previous section to validate the MAS model simulations against that of the ABM outputs, similar tests have been performed with the December 2014 validation storm event. As usual, at simulation start-up, the information captured from each hydrometric sensor by the sensor agents initiates the system and it can run a complete simulation of a storm-water event with the hourly data within the catchment representation, with each of the agent's nodes implemented.

A statistical analysis of the information presented in Table 5.12 and Figure 5.25, for this storm episode, reveals the absolute measure of fit resulting between the hourly flood hydrograph forecasts of the MAS model forecasts concerning the ABM hydrograph output, varied between 34.6. and 49.8 $m^3 \cdot s^{-1}$, although these values did not show any significant changes. Also, for this validation storm, for each of the ABM models measured Q_{pk}, the MAS model overestimated once (e.g., at the 1 hour time-horizon). When contrasted with the ABM observed flows for each of the time horizons, it showed overall with the percentage errors of Q_{pk} varying between ±41.4% and ±15.6%, with the range of these values agreeing with the error criteria.

The results of the correlation metrics (Table 5.12) between the ABM observed and the MAS model flood hydrographs, and by interpreting the statistical significance seen from the linear relationship estimator which can be seen across this storm scenario fluctuating between (r = 0.75 and 0.89, p <.001), which in terms of the coefficient of determination explained variability of the correlation between the MAS simulated Q_{pk} to the ABM observed Q_{pk} raging between 53 and 79%, support the MAS model acceptance.

Table 5.12: Verification metrics for 4-hour lead-time forecast between the observed ABM synthetic December 2014 storm hydrograph and simulated MAS hydrograph, along with the computed flood-awareness.

	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	Obs. (ABM) Q_{pk} [$m^3 \cdot s^{-1}$]	Sim. (MAS) Q_{pk} [$m^3 \cdot s^{-1}$]	FA
December 2014							
Simulation Period							
q1h	0.89	0.79	49.8	15.6	633.3	734.3	5.3
q2h	0.81	0.66	34.6	-52.1	544.7	256.1	5.1
q3h	0.73	0.53	45.0	-41.4	604.0	354.2	5.3
q4h	0.75	0.56	42.5	-43.4	564.0	319.0	5.3

Like the synthetic data of December 2012 for this storm period, the flow-awareness value inferred by the decision agent according to each lead-time flow forecasted by the MAS forecaster agents was between the CAUTION and the ALARM ranges, yet the ALARM status shows some degrees more likely. However, an interesting pattern to notice is the flood-awareness value triggered during the 1-hour lead-time forecast, which could suggest given the magnitudes of the observed and simulated peak flows a warning trigger at least within the ALARM region. However, it is important to recall that the triggering of the alarms is controlled by the rules implemented in the fuzzy inference engine, which means that the state variables of rain and water level also play a significant role in the final output FA as was shown earlier in Table 5.8. Flows can be high, but the rainfall and stage level may be low at a given instance in time, given way for lower FA values. Then, rainfall is the determining component in the generation of overland flows, and indeed, the high water level is a consequence of high flows in a channel. Then, high levels of rainfall and stage are also determining factors in the weighing of the computed FA. To corroborate the value of the FA for that period, a closer look into the time series revealed a rainfall and stage value of 16 [mm] and 9.07 [m], accordingly. Both values are in the Light for rain and Normal for the stage (Table 5.7)).Hence, these findings significantly show the FA = 5.3 computed by the decision agent to be working as expected. Notwithstanding, as hydrometric data are a type of non-stationary data highly influenced by changes in the climate, it suggests the fuzzy rules would need to undergo periodic revisions and changes to allow the decision agent to be up to date.

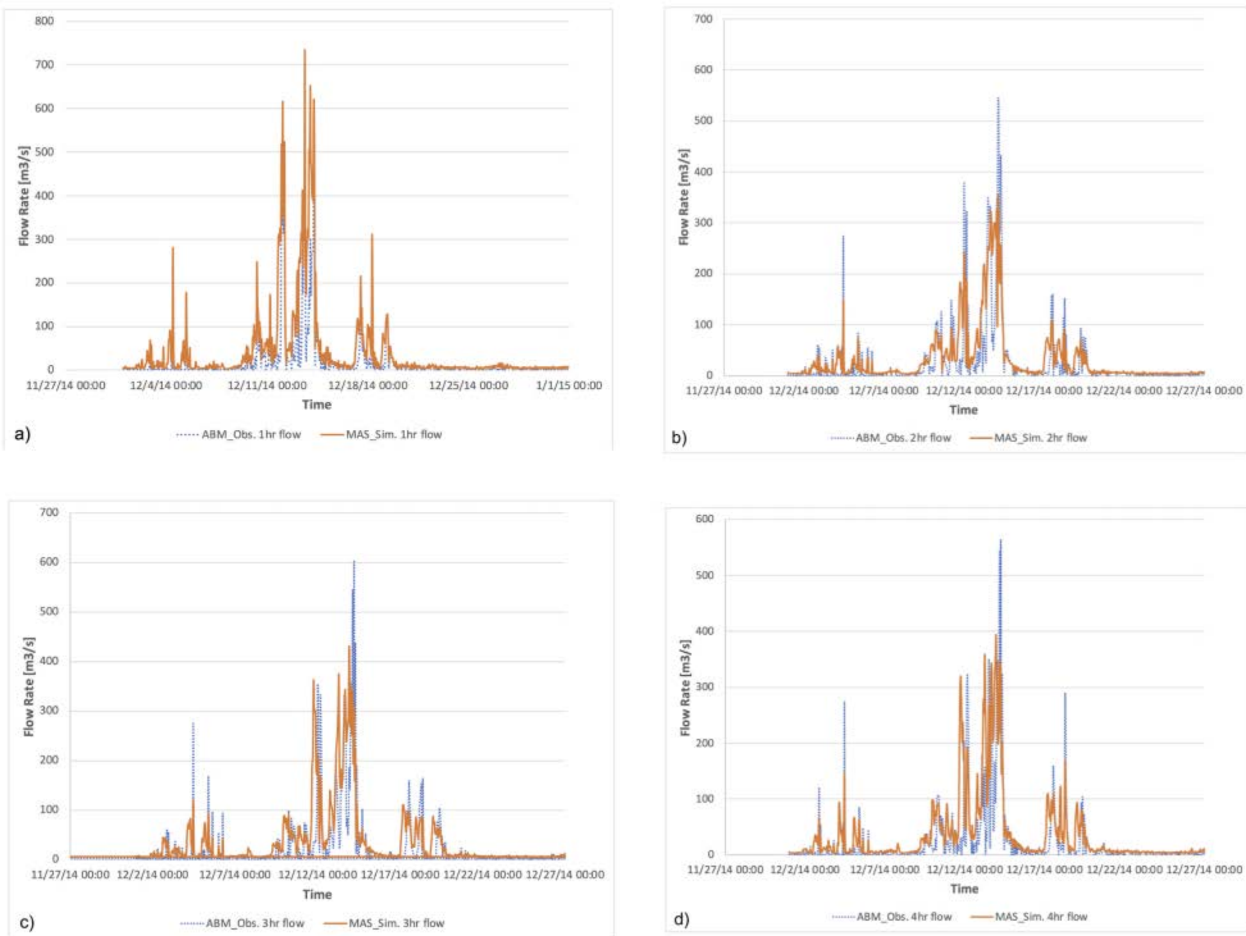


Figure 5.25: Observed (ABM) and simulated (MAS) hydrographs comparisons with the ABM synthetic data of December 2014 storm for time-horizons, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

5.5.3 MAS vs ABM Synthetic Data of May 2015 Storm

In continuation with the experiments aimed at the validation of the MAS model, it can be discussed that for May 2015 storm there were visible around seven isolated storm occurrences as can be recognized in Figure 5.26, where four are minor events producing flood waves of less than 100 $m^3 \cdot s^{-1}$ and the other three register flow values above 100 $m^3 \cdot s^{-1}$ for certain days of the month. However, it may be noticed the isolated trends in this storm episode. For the ABM model measured peak flows at the 3-hour and 4-hour lead times of 312.7 and 462.7 $m^3 \cdot s^{-1}$, respectively, the MAS model overestimated. The 1-hour lead-time MAS model has an RMSE of 22.1, whereas the 2-hour lead-time forecast of 23.8, the 3-hour ahead MAS model forecaster agent is forecasting the flow values with RMSE of 22.8, and additionally the 4-hour ahead forecast agent has an RMSE of 23.2 $m^3 \cdot s^{-1}$. This trend continues to show so far that on average, the three and four-hour forecast agents are forecasting the streamflow at a lesser error. Moreover, the mean absolute percentage error of all estimated Q_{pk} by the MAS model tends to over and underestimate the ABM observed $Q_{pk}(\%)$ with values in the range [±4.5, ±37.8%], and as can be seen from Table 5.13, it is observed that the 3-hour and 4-hour ahead agent forecaster of the MAS model estimates peaks with ±4.5% and ±5.0%, whereas the 1-hour lead-time agent has ±37.8% error and the 2-hour lead-time agent forecasters have ±35.7%.

A comparison of the resulting lead-time forecasts for hydrographs for both models is shown in Figure 5.26. In addition, the r-value showing statistic relevance between the ABM and the MAS model ranged (r = 0.78 to 0.91, p <.001, accordingly), and would give a coefficient of determination R^2 varying in the range [60, 83%], meaning that only 60 to 83% of the variability noted in the MAS model hydrographs forecasts could be accounted for by the ABM forecasts.

Table 5.13: Verification metrics for 4-hour lead-time forecast between the observed ABM synthetic May 2015 storm hydrograph and simulated MAS hydrograph, along with the computed flood-awareness.

	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	Obs. (ABM) Q_{pk} [$m^3 \cdot s^{-1}$]	Sim. (MAS) Q_{pk} [$m^3 \cdot s^{-1}$]	FA
May 2015							
Simulation Period							
q1h	0.79	0.62	22.1	-37.8	329.3	204.8	1.6
q2h	0.78	0.60	23.8	-35.7	329.3	211.6	1.6
q3h	0.90	0.81	22.8	4.5	312.7	326.6	1.6
q4h	0.91	0.83	23.2	5.0	462.7	485.6	1.6

For the ABM observed and simulated MAS flows, this period showed lower values for the computed flood awareness by the decision agent for each of the time horizon forecasts. Generally, each time had a computed FA = 1.6, which places it within the LOW alert zone. However, although this value is low, it does not mean there is no danger in the next hours. If the precipitation persists, a condition favorable for the increase in streamflows and water levels is maintained or increased, proving for a possible flood. The ideal rationale for the suggested structure of the MAS model is that it not only predicts river flows, but also induces subsequent FA levels for each expected flow and is delivered at least 4 hours in advance. On the other hand, if the average lead time of 4 hours is calculated, which for the observed values of the ABM is 358.5 and the forecast of the MAS 307.4 $m^3 \cdot s^{-1}$, according to the Table 5.7 both values would be placed in the very high and average flow levels, respectively. This is information that would be available at least four hours before a flood disaster by decision-makers.

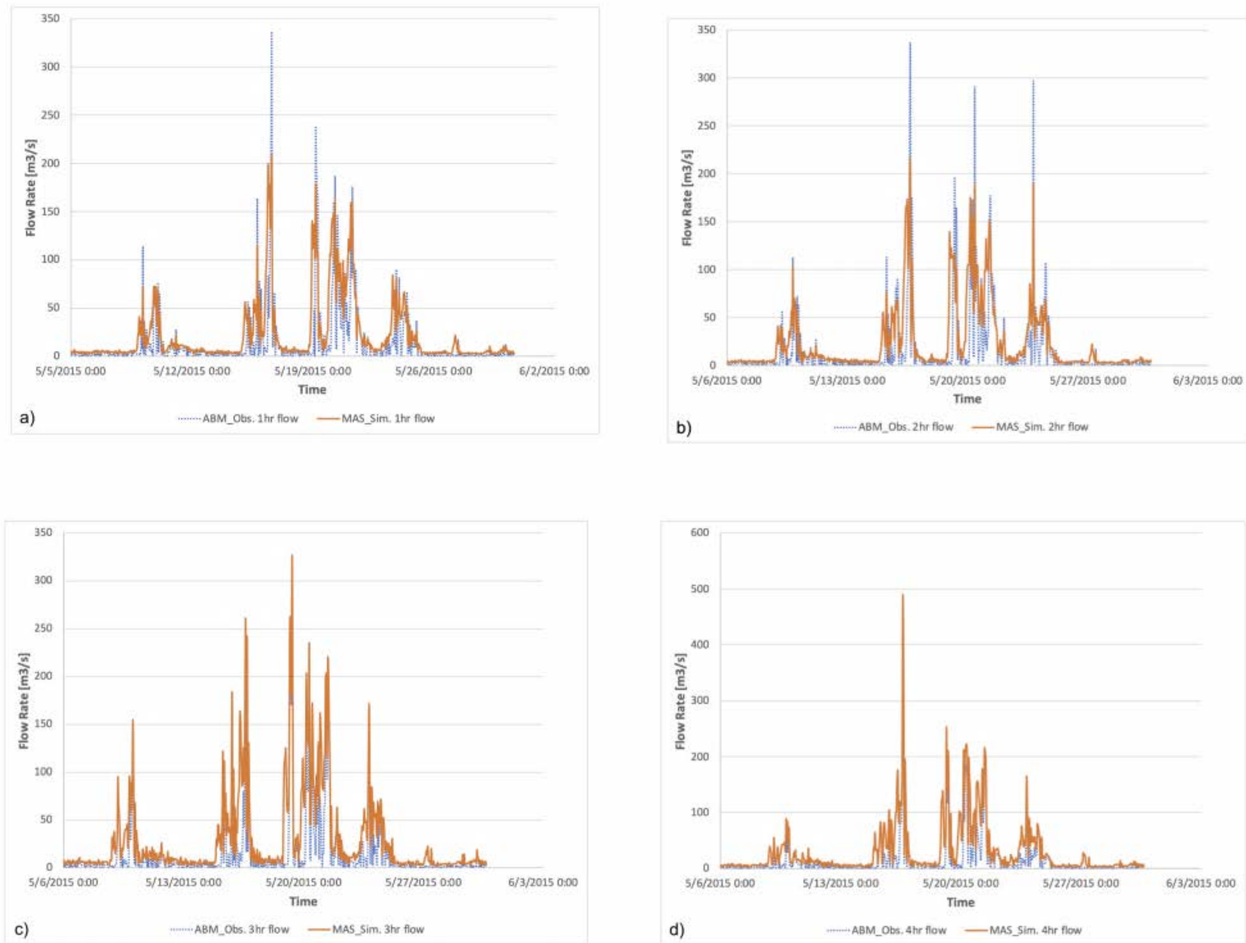


Figure 5.26: Observed (ABM) and simulated (MAS) hydrographs comparisons with the ABM synthetic data of May 2015 storm for time-horizons, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

5.5.4 MAS vs ABM Synthetic Data of November 2015 Storm

For the November 2015 storm event, a statistical comparison of the results about the performance between the MAS model simulated flood hydrograph, and the ABM outputs (Table 5.14) show that the MAS model forecasts with shorter lead-time have better predictions performance compared to the longer lead-time prediction. Notice also, in Figure 5.27 the MAS model simulated hydrographs versus the ABM observed hydrographs, was seen three times to overestimate the ABM flood hydrographs peak flows.

One significant storm event of only five days characterized this modeling period. Across the table, the difference in predictions of the flood hydrographs between the MAS model simulations and the ABM was seen fluctuating between 33.9 and 39.1 $m^3 \cdot s^{-1}$, with the highest error reported at the 2-hour lead-time and the lowest for the 4-hour lead-time forecast.

Likewise, for this period, the $Q_{pk}(\%)$ ranged in [±36.9, 52.2%] between the two models, with the smallest percent for the 4-hour forecast and the highest for the 1-hour forecast.

Conclusively, any contrast of the observed ABM and MAS model simulated hydrographs requires an assessment or rule on what to compare. As seen from across the table, with the November 2015 storm dataset, between the ABM and the MAS model simulated flood hydrographs, the models show significantly correlated, at all lead times forecast in the range [0.77, 0.88] with the value of p<.001 at $\alpha = 0.05$ again suggesting the acceptance and the good performance of the MAS hydrologic model. From the correlation analysis, the coefficient of determination would yield that the ABM model could account for 59 to 79% of the variability detected in the results of the observed flood hydrographs of the MAS model. Concerning the decision agent inferences on the hourly forecasted flows, the system computed a FA that was on average 2.4, LOW, which was seen to intercept with the CAUTION level, to a lesser degree. Overall, this period had a similar trait to that which was observed for the ABM synthetic storm of May 2015 which also presented on average a low value of the FA.

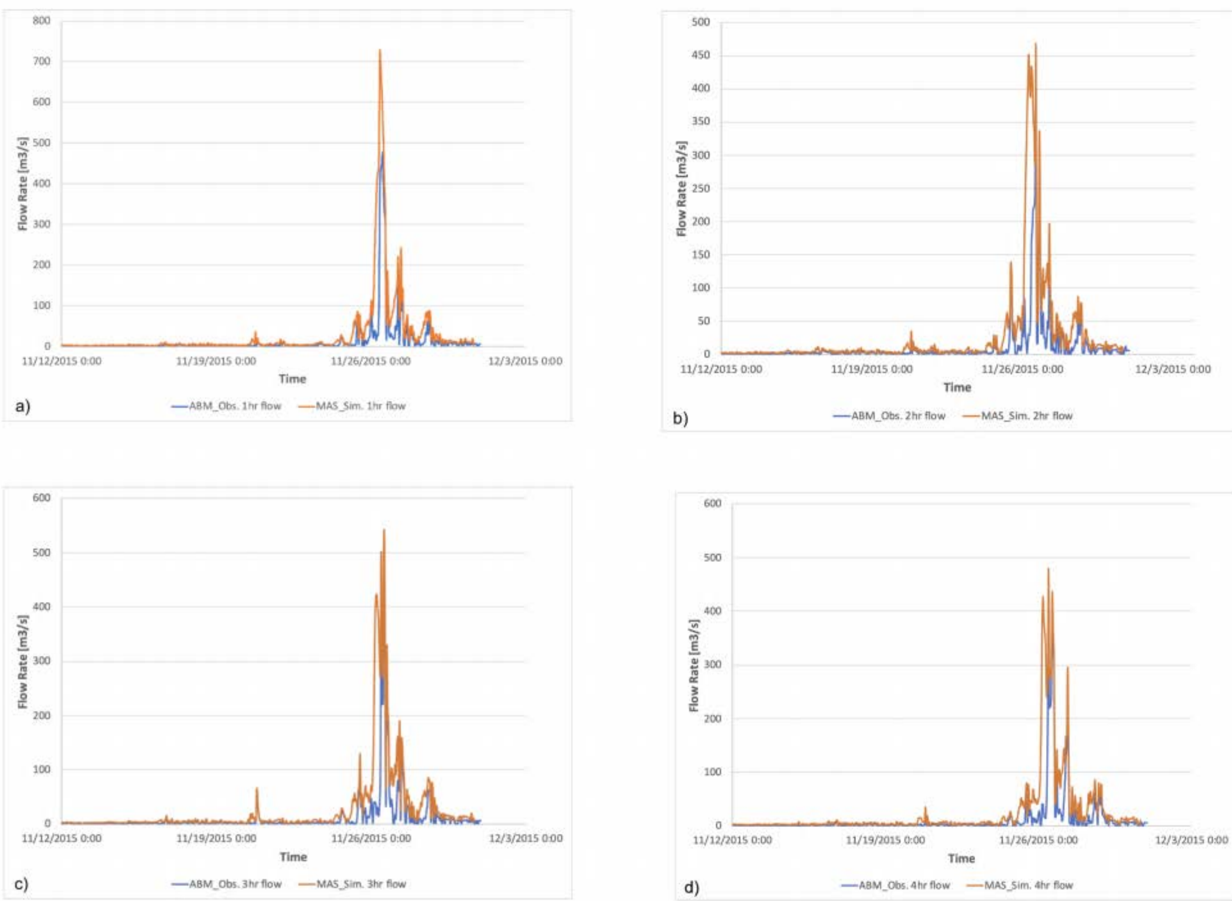


Figure 5.27: Observed (ABM) and simulated (MAS) hydrographs comparisons with the ABM synthetic data of November 2015 storm for time-horizons, where a) one-hour, b) two-hour, c) three-hour, and d) four-hour forecast.

Table 5.14: Verification metrics for 4-hour lead-time forecast between the observed ABM synthetic November 2015 storm hydrograph and simulated MAS hydrograph, along with the computed flood-awareness.

	Cor. Coef. [r]	Coef. of Det. [R^2]	RMSE [$m^3 \cdot s^{-1}$]	Percent Error in Q_{pk} [%]	Obs. (ABM) Q_{pk} [$m^3 \cdot s^{-1}$]	Sim. (MAS) Q_{pk} [$m^3 \cdot s^{-1}$]	FA
November 2015							
Simulation Period							
q1h	0.88	0.79	37.4	52.2	475.7	724.3	2.5
q2h	0.77	0.59	39.1	15.8	389.9	451.8	2.5
q3h	0.79	0.63	36.7	-10.9	475.7	423.8	2.5
q4h	0.80	0.64	33.9	9.5	389.9	427.2	2.2

5.5.5 Conclusions

This section introduced the implementation of the MAS hydrologic model with the necessary BDI-based agents to carry out the task of flood hydrograph forecasting and decision making with the proposed MAS hydrologic model organization with the purpose to recreate the forecasting of streamflows in the humid tropics, by simplifying the ordeals of the conventional mathematical, physically oriented hydrologic modeling setup..

Several experiments were run to validate the MAS model simulation of flood hydrographs for a 1 to 4-hour lead-time by comparing the simulated flood hydrographs outputs of the MAS hydrologic model with those of the ABM hydrologic model outputs. The investigation of ML algorithms, among several others, such as "random forest (RF)" and "support vector regression (SVR)", was selected as the schemes with which the MAS hydrologic model forecaster agents were facilitated with the task of forecasting the hourly flows, at lead-times 1, 2, 3 and 4-hours ahead with the multi-step-ahead strategy. For these agents to carry out their forecasting task, they required the cooperation of several agents within the MAS model organization, e.g. the hydrometric sensor agents (e.g., AgentRNSn, AgentWLSn, and AgentSFSn) who oversee capturing the incoming data from the hydrometric sensors, the AgentSV, who verifies that the sensors that capture data are functional, the AgentHDBM who manages and stores sensor raw data on behalf of the AgentSV, preprocessed data generated by the AgentDPP and AgentData2Lags, the assessment of flow forecast results and inference, on behalf of the AgentFL. The AgentDPP has overseen data imputation preprocessing and the AgentData2Lags the creation of the lagged data matrix, from the raw data for use by the agent forecasters to make the hourly forecasts.

The results from the training and testing sessions of the MAS model forecasting agents, for flow forecasting, with the RF and SVR agent's implementation, indicates the SVR forecasters agents during the training and testing sessions to be capable of dealing with extreme complex values, whereas those with the RF implementation concerning the lower values of the RMSE, and the percentage error in the simulated Q_{pk} performance good at flow estimation in general. Overall, the SVR implementation, despite overestimation, shows at some level of degree to approximate the value of the observed ABM Q_{pk}. However, for both agents group implemented, the lowest measures of the correlation coefficients, were observed at the 3 and 4 hour time periods.

To further enhance and validate the forecasting capacity of the agents in the MAS, each forecast produced by the RF and SVR group of agent forecaster models implementation for 1, 2, 3 and 4-hour ahead forecast is merged and averaged by the decision-making agent (AgentFL) to aggregate the final absolute values of each lead-time forecast that is used by the decision agent to performer the flood-awareness levels inference computations.

The validation data was based on the simulations of four selected validation storm events, obtained from the ABM hydrologic model simulation outputs on which the MAS hydrologic model forecaster agents made lead-time predictions and the decision agent inferred the flood-awareness level for each forecasted flow at lead-time 1, 2, 3, and 4-hour respectively; however, it should be noted that this parameter (flood-awareness level) is not used in the validation phase of MAS simulations against the ABM outputs.

The experiments in this section show that the proposed MAS hydrologic model when tested against the observed ABM simulated validation storm hydrographs despite running with the entire simulation process from capturing the data by the sensor agents to dealing with data preprocessing by the agent performing imputations and the management of the forecast results delivered by the agent forecasters to the decision-making agent, allows simulating the flood hydrographs reasonably well. However, although the purpose of the validation experiments was contrasting the MAS model simulation to that of the ABM observed outputs, the The RMSE and the coefficient of similarity "r" statistics were utilized equally like quantitative indicators to assess the modeling outcomes between the MAS model flood hydrograph modeling output and the observed ABM flood hydrograph output., which suggests that the MAS model implementation is efficiently good in assuming the proposed task and shows that these experiments will be useful both to the hydrologic and water resources community. In addition, it is important to recall that despite an initial start-up of a hydrologic simulation at the catchment level through was carried with the use of a tool whose paradigm, in theory, is not specifically that of a tool for hydrological simulation studies per se, it has demonstrated through its characteristics of the "agentification of GIS data" of the components of a specific domain and that after the calibration and validation process it will demonstrate a satisfactory approximation to the curves of certain flood hydrographs as indicated by the results of the evaluation metrics (Table 4.17).

Therefore, the initial implementation of a blueprint of an agent-based concept for a hydrologic model to simulate the flood hydrographs in the context of a tropical basin domain, its calibration, and validation phases were successfully implemented in the GAMA agent platform, and that it requires further work to enhance and improve this feature. However, it was found the literature review falls significantly short of being sufficient in containing verification and validation procedures for multi-agent models, since they are difficult to achieve, as some researchers may argue [460–462]. There may exist several papers on MAS, but most of them seldom address the calibration and verification topic, it is just simply ignored. If this is the case, there is the need to implement methods for watershed and hydrologic simulations, calibration, and validation. Fortunately, this blueprint had placed the sketch that allowed its extension with the BDI-concept to build upon the MAS hydrologic model framework.

Chapter 6

Key Conclusions

This research study and experimentation represented a challenge for the formulation of a practical MAS for the forecasting of streamflows implemented with experimental hybrid, computational learning paradigms through practical hydroinformatics skills. To solve the social issues related to hydrologic generated flood problems, researchers and the hydrologic engineer may resort to conceptual models or numerical models that are customary physically based, which are usually not simple to use, are complex to implement because of the large number of resources required and most times, due to lack of experience, the results are difficult to interpret and this not to mention the professional water resources manager whose knowledge or use of these tools are few or absent. The ideal scenario with computational intelligence models in contrast to their conventional hydrological model's counterpart is that they do not require most of the physical variables of the latter, making them useful for their ability to extract information from hydrometric time series, even from those with poor data. It was explored the application of several artificial intelligence approaches, including those of data-driven modeling, soft computing, agent-based, and multi-agent conceptions. Given these circumstances, the Medio catchment, Panama was selected as a test case. Besides, like most all catchments in the Republic of Panama, the Medio catchment is not thoroughly studied, data availability is a problem and the precipitation regime in the area is very high, and it is highly prone to flooding events. That being the case six objectives were put in place to design, develop, and validate a multi-agent conceptual framework to deal with the inundation problems within humid watersheds. Therefore, the first of these six objectives, which was completed, was to comprehend the hydrology of the Medio catchment by performing physical modeling of the environment. To comply with this task, it was utilized the non-commercially available "Hydrologic Modeling System (HEC-HMS)", the "River Analysis System (HEC-RAS)", the "Quantum Geographic Information System (QGIS)" application, and the "Whitebox Geospatial Analysis Tools (Whitebox-GAT)".

The second objective was to design the MAS architecture for the domain-specific MAS framework for supporting flood forecasting in the tropical river basin domain. To accomplish this objective, it is realized that a very few of the exiting MAS platforms (e.g., JADE, JADEX, NetLogo, Mesa, just to mention the most used) could not fill the expectations, and after the careful exploration of such, though some would provide the BDI architecture, they lacked the features of the GIS that was needed to set up a unique catchment emulating environment capability until the GAMA simulation platform was identified and it was able to implement the domain knowledge on this system following the flood ontology-based approach in Section 4.1.3.

The third objective was to identify the specifications for such a catchment domain MAS architecture. So, to comply with this objective, it was necessary to set up the MAS framework from the designed architecture on top of the selected agent platform (in this case GAMA) and implement the system using an integrated approach of hydrological, hydraulic, and data-driven modeling, and artificial intelligence techniques for flood forecasting.

The fourth objective addresses an integral part of the forecasting framework behavioral capabilities of each of the agents composing the system. Therefore, to fulfill this endeavor, implementing the BDI model organizational structure for agents' behavior, communication, and interactions skills that compose the MAS to manage the data obtained by the framework tool application had to be properly identified and abstracted.

Then, the fifth and sixth objective deals with the computer-based simulation of the MAS to provide the necessary verification of the functionality of the system that was implemented via the second, third, and fourth objectives and its future deployment to a real-world scenario as the proposed hourly flood forecasting and flood-awareness tool provided the appropriate means. Therefore, experimental tasks such as simulation, calibration, and verification of the MAS outputs against the results from the hydrologic and hydrodynamic models had to be conducted and completed appropriately as can be referenced from Chapters 4 and 5 of this thesis.

6.1 Summary

The motivations that started the goal of this research thesis design has been to address the problems of inundation by flooding in the tropical river basin by proposing, on the one hand, a cognitive approach regarding the BDI theory to embody simulation of water surges caused by river flooding and to represent decision actions processes by human actors from the warnings issued by the multi-agent system models and, by the other hand, a validation of this approach in a complete watershed domain (simulation environment) model in which flood forecasting of this type have not been simulated.

The subsequent points summarize the conclusions of this research;

- Efficient hydrologic modeling and assessment with conventional tools require a great deal of physically-oriented information types and given their degree of complexities, can represent a complicated task to implement. On the other hand, artificial intelligence and its various paradigms can offer expedient and proper solutions to hydrologic modeling and flood assessment. However, although they may not require huge physical data, the length and quality of the time series data they require are essential for the excellent performance of their forecasting models.

- Why MAS for flood forecasting and not just simply conventional models? As distributed systems in the case of streamflow forecasting and assessment of flooding, AI implementations are better because many computational tasks before modeling (e.g., data cleaning, data imputation, data preparation) that would require a dedicated person to accomplish, can be left unattended and performed by specialized agents with cognitive abilities and skills.

- Besides, the rainfall-runoff events that trigger river flooding are not usually localized, so of relevance is the information upstream of the catchment to the information downstream of the catchment which suggests the information may not be uniformly distributed, available, or steady. With the MAS approach, the process of decision-making amid localized or non-localized events can generate results of the same reliability, even if a section of the hydrometric sensor node is unavailable at the time.

- Last, though the agents of this administration team up for the same aim, they do not have different objectives or isolation of their own, each entity that takes part has different capabilities, their representation, and their operation in real-time may vary as this is best reflected with agents and their roles.

6.2 Future Perspectives

The experimental findings of this research demonstrate and provide insight on the capacity to make use of agent technologies, combined and implemented with artificial intelligence and hybridization alternatives for flood forecasting. However, as it is not until recently that the agent's paradigms and even the adoption of principles from computational intelligence, such as soft computing (i.e., Fuzzy Logics) techniques by the hydrological community for the solution of problems related to floods still have gaps for research and development especially with the constant changes in the current climate regime.

The research has also disclosed the advantages of the multi-agent systems as a good surrogate for real-time hydrologic modeling at the catchment level, given its capacity as a distributed system, decentralized philosophy, and the capacity of communication through a high-level messaging protocol, and interactions using skilled cognitive agents, a characteristic not typical of the conventional physically based models.

Model training, testing, calibration, and validation with storm scenarios were completed with data that at some point had to be reconstructed due to the severity of missing instances in the actual data record. Therefore, the data feed for model simulation was not done in real-time, as is intended the system must do directly from the hydrometric sensors, but data was feed directly into the models from the hydro files.

Many were the challenges encountered, been the main one of the datasets from two available hydrometric stations that measure precipitation, river stage, and streamflow in real-time, only one was currently used, the other was severely impaired and GIS information on the catchment is scarce, so all river hydraulic, and GIS features, physiographic quantities, and constants had to be estimated and treated from the DEM downloaded. Additionally, the existence in the literature of adequate methods and methodologies for validating multi-agent system models for hydrologic problem solving is either scarce or non-available and surely represents a challenge for research and development to take better advantage of the agent technology.

Finally, all simulations and model development show to efficiently model the problems aforementioned. Therefore, it is advisable as future works to obtain new and complete datasets, an updated DEM, and new hydraulic measurements of the current physical conditions of the river's course for updating the MAS model and with such provide deployment of the system for in situ testing.

Bibliography

[1] Eric K Noji and CY Lee. "Disaster preparedness." In: *Environmental health: from global to local, 1st edn. Jossey-Bass, San Francisco* (2005), pp. 745–780.

[2] Yoganath Adikari and Junichi Yoshitani. "Global trends in water-related disasters: an insight for policymakers." In: *World Water Assessment Programme Side Publication Series, Insights. The United Nations, UNESCO. International Centre for Water Hazard and Risk Management (ICHARM)* (2009).

[3] Maxx Dilley et al. *Natural disaster hotspots: a global risk analysis.* The World Bank, 2005.

[4] Banco Mundial Colombia. "Global Facility for Disaster Reduction and Recovery." In: *Análisis de la gestión del riesgo de desastres en Colombia: un aporte a la construcción de políticas públicas* (2012).

[5] UN ESCAP. "Disasters in Asia and the Pacific: 2014 year in review." In: *United Nations report. Economic and social commission for Asia and the Pacific* (2015).

[6] UN ESCAP. "Disasters in Asia and the Pacific: 2015 year in review." In: *United Nations report. Economic and social commission for Asia and the Pacific* (2016).

[7] Lorenzo Alfieri et al. "Global projections of river flood risk in a warmer world." In: *Earth's Future* 5.2 (2017), pp. 171–182.

[8] Rob Garner. *NASA Earth Images.* 2013. URL: https://www.nasa.gov/content/nasa-earth-images.

[9] Sharad Kumar Jain et al. "A Brief review of flood forecasting techniques and their applications." In: *International journal of river basin management* 16.3 (2018), pp. 329–344.

[10] Peter H Gleick. "Global freshwater resources: soft-path solutions for the 21st century." In: *Science* 302.5650 (2003), pp. 1524–1528.

[11] David B Brooks, Oliver M Brandes, and Stephen Gurman. *Making the most of the water we have: The soft path approach to water management.* Earthscan, 2009.

[12] J Opperman. *A Flood of benefits: Using green infrastructure to reduce flood risks. The Nature Conservancy, Arlington, Virginia*. 2014.

[13] Cheol-Hee Son et al. “The effects of mitigation measures on flood damage prevention in Korea.” In: *Sustainability* 7.12 (2015), pp. 16866–16884.

[14] Beatriz Revilla-Romero et al. “On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions.” In: *Remote Sensing* 7.11 (2015), pp. 15702–15728.

[15] Muhammad Azam, Hyung San Kim, and Seung Jin Maeng. “Development of flood alert application in Mushim stream watershed Korea.” In: *International journal of disaster risk reduction* 21 (2017), pp. 11–26.

[16] Joost Dewelde et al. “Real-time flood forecasting systems in Flanders.” In: (2014).

[17] J Sitterson et al. *An Overview of Rainfall-Runoff Model Types*. 2017.

[18] Wen Wang. *Stochasticity, nonlinearity and forecasting of streamflow processes*. Ios Press, 2006.

[19] Rajesh Raj Shrestha and Franz Nestmann. “Physically based and data-driven models and propagation of input uncertainties in river flood prediction.” In: *Journal of Hydrologic Engineering* 14.12 (2009), pp. 1309–1319.

[20] Jorge L Pousa et al. “Environmental impacts and simultaneity of positive and negative storm surges on the coast of the Province of Buenos Aires, Argentina.” In: *Environmental earth sciences* 68.8 (2013), pp. 2325–2335.

[21] Paolo Alfredini et al. “Exposure of Santos Harbor metropolitan area (Brazil) to wave and storm surge climate changes.” In: *Water Quality, Exposure and Health* 6.1-2 (2014), pp. 73–88.

[22] Sheng Dong et al. “A storm surge intensity classification based on extreme water level and concomitant wave height.” In: *Journal of Ocean University of China* 14.2 (2015), pp. 237–244.

[23] EF Adam et al. “A systematic assessment of maritime disruptions affecting UK ports, coastal areas and surrounding seas from 1950 to 2014.” In: *Natural Hazards* 83.1 (2016), pp. 691–713.

[24] José A Simmonds, Juan A Gómez, and Agapito Ledezma. “Forecasting sea level changes applying data mining techniques to the Cristobal Bay time series, Panama.” In: *Journal of Water and Climate Change* 8.1 (2016), pp. 89–101.

[25] Michael Elliott, Nicholas D Cutts, and Anna Trono. “A typology of marine and estuarine hazards and risks as vectors of change: a review for vulnerable coasts and their management.” In: *Ocean & Coastal Management* 93 (2014), pp. 88–99.

[26] N Quinn et al. “Assessing the temporal variability in extreme storm-tide time series for coastal flood risk assessment.” In: *Journal of Geophysical Research: Oceans* 119.8 (2014), pp. 4983–4998.

[27] Tim Webster et al. “Integrated river and coastal hydrodynamic flood risk mapping of the lahave river estuary and town of Bridgewater, Nova Scotia, Canada.” In: *Water* 6.3 (2014), pp. 517–546.

[28] Thomas Prime, Jennifer M Brown, and Andrew J Plater. “Physical and economic impacts of sea-level rise and low probability flooding events on coastal communities.” In: *PLoS One* 10.2 (2015), e0117030.

[29] George T Raber and Zach Ferdaña. “Coastal resilience: an ecosystem-based coastal and marine spatial planning framework.” In: *Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications*. ACM. 2011, p. 57.

[30] MK Akhtar et al. *Ganges River Flood Forecasting Using Spatially Distributed Rainfall from Satellite Data and Artificial Neural Networks*. Tech. rep. Water Mill Working Paper Series, 2008.

[31] Inc. Insurance Information Institute. *Facts + Statistics: Global catastrophes*. 2018. URL: http://https://www.iii.org/fact-statistic/facts-statistics-global-catastrophes%20[Accessed%2013%20Feb.%202018]..

[32] NU OCHA. “Natural Disasters in Latin America and The Caribbean.” In: (2019).

[33] Yukiko Hirabayashi et al. “Global flood risk under climate change.” In: *Nature climate change* 3.9 (2013), pp. 816–821.

[34] Sanne Muis et al. “A global reanalysis of storm surges and extreme sea levels.” In: *Nature communications* 7.1 (2016), pp. 1–12.

[35] TN Palmer and Jouni Räisänen. “Quantifying the risk of extreme seasonal precipitation events in a changing climate.” In: *Nature* 415.6871 (2002), p. 512.

[36] Water Research Foundation. “Effects of Climate Change on Public Water Suppliers.” In: *Awwa Research Foundation* (2008), pp. 1–5. URL: https://www.waterrf.org%20[Accessed%2013%20Feb.%202018].

[37] Credit Valley Conservation. “Rising to the challenge: a handbook for understanding and protecting the Credit River Watershed.” In: *Retrieved October* 23 (2009), p. 2014.

[38] F John. “Sowa, Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks Cole Publishing Co.” In: (2000).

[39] Brian HW Guo and Yang Miang Goh. “Ontology for design of active fall protection systems.” In: *Automation in Construction* 82 (2017), pp. 138–153.

[40] Thomas R Gruber. “A translation approach to portable ontology specifications.” In: *Knowledge acquisition* 5.2 (1993), pp. 199–220.

[41] Nikos Bikakis et al. “The XML and semantic web worlds: technologies, interoperability and integration: a survey of the state of the art.” In: *Semantic hyper/multimedia adaptation*. Springer, 2013, pp. 319–360.

[42] Hayet Brabra et al. “Semantic web technologies in cloud computing: a systematic literature review.” In: *2016 IEEE International Conference on Services Computing (SCC)*. IEEE. 2016, pp. 744–751.

[43] Pieter Pauwels, Sijie Zhang, and Yong-Cheol Lee. “Semantic web technologies in AEC industry: A literature overview.” In: *Automation in Construction* 73 (2017), pp. 145–165.

[44] Jesper Jensen. “A systematic literature review of the use of semantic web technologies in formal education.” In: *British Journal of Educational Technology* 50.2 (2019), pp. 505–517.

[45] Mahboubeh Dadkhah, Saeed Araban, and Samad Paydar. “A systematic literature review on semantic web enabled software testing.” In: *Journal of Systems and Software* 162 (2020), p. 110485.

[46] Rob Raskin and Michael Pan. “Semantic web for earth and environmental terminology (sweet).” In: *Proc. of the Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data*. Vol. 25. 2003.

[47] Michael Compton et al. “The SSN ontology of the W3C semantic sensor network incubator group.” In: *Journal of Web Semantics* 17 (2012), pp. 25–32.

[48] S Cox et al. *Time ontology in OWL. W3C Recommendation, 19 October 2017*. 2018.

[49] Prasant Kumar Sinha and Biswanath Dutta. “A Systematic Analysis of Flood Ontologies: A Parametric Approach.” In: *KO KNOWLEDGE ORGANIZATION* 47.2 (2020), pp. 138–159.

[50] Kwok-Wing Chau. “An ontology-based knowledge management system for flow and water quality modeling.” In: *Advances in Engineering Software* 38.3 (2007), pp. 172–181.

[51] Julián Garrido, Ignacio Requena, and Stefano Mambretti. “Semantic model for flood management.” In: *Journal of Hydroinformatics* 14.4 (2012), pp. 918–936.

[52] Annalisa Agresta et al. “An ontology framework for flooding forecasting.” In: *International Conference on Computational Science and Its Applications*. Springer. 2014, pp. 417–428.

[53] Yulin Ding, Qing Zhu, and Hui Lin. "An integrated virtual geographic environmental simulation framework: a case study of flood disaster simulation." In: *Geo-spatial Information Science* 17.4 (2014), pp. 190–200.

[54] Jie Sun et al. "Intelligent flood adaptive context-aware system: How wireless sensors adapt their configuration based on environmental phenomenon events." In: *Sensors & Transducers* 206.11 (2016), p. 68.

[55] Chao Wang et al. "A hydrological sensor web ontology based on the SSN ontology: A case study for a flood." In: *ISPRS International Journal of Geo-Information* 7.1 (2018), p. 2.

[56] Yusuf Sermet and Ibrahim Demir. "Towards an information centric flood ontology for information management and communication." In: *Earth Science Informatics* 12.4 (2019), pp. 541–551.

[57] Klaas-Jan Douben. "Characteristics of river floods and flooding: a global overview, 1985–2003." In: *Irrigation and Drainage: The journal of the International Commission on Irrigation and Drainage* 55.S1 (2006), S9–S21.

[58] Integrated Flood Risk Analysis and Management Methodologies. *UrbanFlood European Seventh Framework Programme project*. 2009. URL: http://www.floodsite.net/[Accessed%2013%20Feb.%202018]..

[59] PG Samuels et al. "Advances in flood risk management from the FLOODsite project." In: (2008).

[60] Paul Samuels. "Language of risk: project definitions." In: *T32-04-01* (2009).

[61] F Klijn. "Flood risk assessment and flood risk management; an introduction and guidance based on experiences and findings of FLOODsite (an EU-funded integrated project)." In: *T29-09-01* (2009).

[62] EU 7th framework Programme. *UrbanFlood European Seventh Framework Programme project*. 2009. URL: http://urbanflood.eu/Pages/default.html%20[Accessed%2013%20Feb.%202018]..

[63] Kees Vermeer et al. *Flood Control 2015: Five years of innovation in flood risk*. 2012.

[64] HRGK Hack et al. "Strength of peat dykes evaluated by remote sensing (Gebiedsdekkende dijksterkte bepaling met remote sensing), Pilot project: RSDYK2008, Program Flood Control 2015." In: *Flood Control* (2015).

[65] RT Clarke. "A review of some mathematical models used in hydrology, with observations on their calibration and use." In: *Journal of hydrology* 19.1 (1973), pp. 1–20.

[66] HS Wheater, AJ Jakeman, and KJ Beven. "Progress and directions in rainfall-runoff modelling." In: (1993).

[67] Vijay P Singh. *Computer models of watershed hydrology. Rev.* 1995.

[68] GE Hecker et al. “Hydrology Handbook: ASCE Manuals and Reports on Engineering Practice No. 28.” In: *American Society of Civil Engineers, New York* (1996).

[69] Eurydice 92 and Bernard Chocat. *Encyclopédie de l’hydrologie urbaine et de l’assainissement.* Tech. & Doc., 1997.

[70] Jens Christian Refsgaard. “Validation and intercomparison of different updating procedures for real-time forecasting.” In: *Hydrology Research* 28.2 (1997), pp. 65–84.

[71] B Ambroise. “Genèse des débits dans les petits bassins versants ruraux en milieu tempéré: 2-modélisation systémique et dynamique [Streamflow generation within small rural catchments in a temperate environment: 2 - Systemic and dynamic modelling].” In: *Revue des sciences de l’eau/Journal of Water Science* 12.1 (1999), pp. 125–153.

[72] IG Pechlivanidis et al. “Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications.” In: *Global NEST journal* 13.3 (2011), pp. 193–214.

[73] Vincent Fortin et al. “Revue bibliographique des méthodes de prévision des débits [Agricultural extension and advisory research: A review of international literature].” In: *Revue des sciences de l’eau/Journal of Water Science* 10.4 (1997), pp. 461–487.

[74] Dongkyun Kim and Francisco Olivera. “Improving Stochastic Rainfall Generators.” In: May 2010, pp. 2324–2327. ISBN: 978-0-7844-1114-8. DOI: 10.1061/41114(371)240.

[75] Anna Kauffeldt et al. “Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level.” In: *Environmental Modelling & Software* 75 (2016), pp. 68–76.

[76] W WMO. *Manual on flood forecasting and warning.* Tech. rep. 2011.

[77] MW Liddament and DB Oakes. “The Use of a Groundwater Model in the Design, Performance Assessment, and Operation of a River Regulation Scheme.” In: *Logistics and Benefits of Using Mathematical Models of Hydrologic and Water Resource Systems* (1978), p. 149.

[78] Keith J Beven. *Rainfall-runoff modelling: the primer.* John Wiley & Sons, 2011.

[79] Francesco Granata, Rudy Gargano, and Giovanni de Marinis. “Support vector regression for rainfall-runoff modeling in urban drainage: A comparison with the EPA’s storm water management model.” In: *Water* 8.3 (2016), p. 69.

[80] Teemu Kokkonen, Harri Koivusalo, and Tuomo Karvonen. “A semi-distributed approach to rainfall-runoff modelling-a case study in a snow affected catchment.” In: *Environmental Modelling & Software* 16.5 (2001), pp. 481–493.

[81] Mamdouh A Antar, Ibrahim Elassiouti, and Mohamed N Allam. “Rainfall-runoff modelling using artificial neural networks technique: a Blue Nile catchment case study.” In: *Hydrological Processes: An International Journal* 20.5 (2006), pp. 1201–1216.

[82] Rao S Govindaraju and Adiseshappa Ramachandra Rao. *Artificial neural networks in hydrology*. Vol. 36. Springer Science & Business Media, 2013.

[83] S Bergström and A Forsman. “Development of a conceptual deterministic rainfall-runoff mode.” In: *Nord. Hydrol* 4 (1973), pp. 240–253.

[84] MJ Kirkby and KJ Beven. “A physically based, variable contributing area model of basin hydrology.” In: *Hydrological Sciences Journal* 24.1 (1979), pp. 43–69.

[85] AJ Jakeman and GM Hornberger. “How much complexity is warranted in a rainfall-runoff model?” In: *Water resources research* 29.8 (1993), pp. 2637–2649.

[86] Bill Scharffenberg et al. *Hydrologic Modeling System HEC-HMS User’s Manual - Version 4.3. U.S.* Tech. rep. 2018.

[87] MB Abbott et al. “An introduction to the European Hydrological System-Systeme Hydrologique Europeen,"SHE", 2: Structure of a physically-based, distributed modelling system.” In: *Journal of hydrology* 87.1-2 (1986), pp. 61–77.

[88] K Beven, A Calver, and EM Morris. “The Institute of Hydrology distributed model.” In: (1987).

[89] JT Croton and DA Barry. “WEC-C: a distributed, deterministic catchment model-theory, formulation and testing.” In: *Environmental Modelling & Software* 16.7 (2001), pp. 583–599.

[90] A Luchetta and S Manetti. “A real time hydrological forecasting system using a fuzzy clustering approach.” In: *Computers & geosciences* 29.9 (2003), pp. 1111–1117.

[91] Nor Irwan Nor, Sobri Harun, and Amir Hashim Kassim. “Radial basis function modeling of hourly streamflow hydrograph.” In: *Journal of Hydrologic Engineering* 12.1 (2007), pp. 113–123.

[92] G Box and G Jenkins. “Time series analysis: Forecasting and control.” In: (1970).

[93] Lofti A Zadeh. “Fuzzy logic, neural networks, and soft computing.” In: *Communications of the ACM* 37.3 (1994), pp. 77–85.

[94] George EP Box et al. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[95] Kuo-lin Hsu, Hoshin Vijai Gupta, and Soroosh Sorooshian. “Artificial neural network modeling of the rainfall-runoff process.” In: *Water resources research* 31.10 (1995), pp. 2517–2530.

[96] ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. "Artificial neural networks in hydrology. I: Preliminary concepts." In: *Journal of Hydrologic Engineering* 5.2 (2000), pp. 115–123.

[97] Rao S Govindaraju et al. "Artificial neural networks in hydrology. II: hydrologic applications." In: *Journal of Hydrologic Engineering* 5.2 (2000), pp. 124–137.

[98] NJ De Vos and THM Rientjes. "Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation." In: *Hydrology and Earth System Sciences Discussions* 2.1 (2005), pp. 365–415.

[99] Holger R Maier et al. "Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions." In: *Environmental modelling & software* 25.8 (2010), pp. 891–909.

[100] Robert J Abrahart et al. "Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting." In: *Progress in Physical Geography* 36.4 (2012), pp. 480–513.

[101] A Brath, A Montanari, and E Toth. "Neural networks and non-parametric methods for improving real-time flood forecasting through conceptual hydrological models." In: *Hydrology and Earth System Sciences Discussions* 6.4 (2002), pp. 627–639.

[102] Jose Simmonds, Juan A Gómez, and Agapito Ledezma. "Data Preprocessing to Enhance Flow Forecasting in a Tropical River Basin." In: *International Conference on Engineering Applications of Neural Networks*. Springer. 2017, pp. 429–440.

[103] Barbara Cannas et al. "Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning." In: *Physics and Chemistry of the Earth, Parts A/B/C* 31.18 (2006), pp. 1164–1171.

[104] Bijaya P Shrestha, Lucien Duckstein, and Eugene Z Stakhiv. "Fuzzy rule-based modeling of reservoir operation." In: *Journal of water resources planning and management* 122.4 (1996), pp. 262–269.

[105] Lotfi A Zadeh. "Fuzzy sets." In: *Information and control* 8.3 (1965), pp. 338–353.

[106] Tanja Dubrovin, Ari Jolma, and Esko Turunen. "Fuzzy model for real-time reservoir operation." In: *Journal of water resources planning and management* 128.1 (2002), pp. 66–73.

[107] Pao-Shan Yu and Shien-Tsung Chen. "Updating real-time flood forecasting using a fuzzy rule-based model/mise a Jour de Prevision de Crue en Temps Reel Grace a un Modele a Base de Regles Floues." In: *Hydrological Sciences Journal* 50.2 (2005).

[108] Sunny Joseph Kalayathankal and G Suresh Singh. "A fuzzy soft flood alarm model." In: *Mathematics and Computers in Simulation* 80.5 (2010), pp. 887–893.

[109] Yaoyao He et al. "A fuzzy clustering iterative model using chaotic differential evolution algorithm for evaluating flood disaster." In: *Expert Systems with Applications* 38.8 (2011), pp. 10060–10065.

[110] PC Nayak, KP Sudheer, and SK Jain. "Rainfall-runoff modeling through hybrid intelligent system." In: *Water Resources Research* 43.7 (2007).

[111] Deg-Hyo Bae, Dae Myung Jeong, and Gwangseob Kim. "Monthly dam inflow forecasts using weather forecasting information and neuro-fuzzy technique." In: *Hydrological Sciences Journal* 52.1 (2007), pp. 99–113.

[112] Mahmut Firat and Mahmud Güngör. "River flow estimation using adaptive neuro fuzzy inference system." In: *Mathematics and Computers in Simulation* 75.3-4 (2007), pp. 87–96.

[113] Seyed Ahmad Akrami, Ahmed El-Shafie, and Othman Jaafar. "Improving rainfall forecasting efficiency using modified adaptive neuro-fuzzy inference system (MANFIS)." In: *Water resources management* 27.9 (2013), pp. 3507–3523.

[114] Robert J Abrahart, Linda M See, and Dimitri P Solomatine. *Practical hydroinformatics: computational intelligence and technological developments in water applications*. Vol. 68. Springer Science & Business Media, 2008.

[115] Alex J Cannon and Paul H Whitfield. "Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models." In: *Journal of Hydrology* 259.1-4 (2002), pp. 136–151.

[116] F Melone et al. *Review and selection of hydrological models-Integration of hydrological models and meteorological inputs*. Tech. rep. 2005.

[117] Yi Zheng and Arturo A Keller. "Stochastic Watershed Water Quality Simulation for TMDL Development–A Case Study in the Newport Bay Watershed 1." In: *JAWRA Journal of the American Water Resources Association* 44.6 (2008), pp. 1397–1410.

[118] Soteris A Kalogirou. "Artificial intelligence for the modeling and control of combustion processes: a review." In: *Progress in energy and combustion science* 29.6 (2003), pp. 515–566.

[119] Zaher Mundher Yaseen et al. "Artificial intelligence based models for stream-flow forecasting: 2000–2015." In: *Journal of Hydrology* 530 (2015), pp. 829–844.

[120] Serena H Chen, Anthony J Jakeman, and John P Norton. "Artificial intelligence techniques: an introduction to their use for modelling environmental systems." In: *Mathematics and computers in simulation* 78.2-3 (2008), pp. 379–400.

[121] Amir Mosavi, Sina Faizollahzadeh Ardabili, and Shahabodin Shamshirband. "Demand prediction with machine learning models: State of the art and a systematic review of advances." In: *Demand Prediction with Machine Learning Models; State of the Art and a Systematic Review of Advances* (2019).

[122] Simon Haykin. *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994.

[123] C Stergiou and D Siganos. *Neurel Networks*. 1996.

[124] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. *Learning internal representations by error propagation*. Tech. rep. California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[125] Michaela Bray and Dawei Han. "Identification of support vector machines for runoff modelling." In: *Journal of Hydroinformatics* 6.4 (2004), pp. 265–280.

[126] D Han, L Chan, and N Zhu. "Flood forecasting using support vector machines." In: *Journal of hydroinformatics* 9.4 (2007), pp. 267–276.

[127] D Han et al. "River flow modelling using fuzzy decision trees." In: *Water Resources Management* 16.6 (2002), pp. 431–445.

[128] RJ Abrahart and LM See. "Neural network modelling of non-linear hydrological relationships." In: (2007).

[129] Dimitri P Solomatine and Avi Ostfeld. "Data-driven modelling: some past experiences and new approaches." In: *Journal of hydroinformatics* 10.1 (2008), pp. 3–22.

[130] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators." In: *Neural networks* 2.5 (1989), pp. 359–366.

[131] David S Broomhead and David Lowe. *Radial basis functions, multi-variable functional interpolation and adaptive networks*. Tech. rep. Royal Signals and Radar Establishment Malvern (United Kingdom), 1988.

[132] Donald F Specht. "A general regression neural network." In: *IEEE transactions on neural networks* 2.6 (1991), pp. 568–576.

[133] Hikmet Kerem Cigizoglu. "Application of generalized regression neural networks to intermittent flow forecasting and estimation." In: *Journal of Hydrologic Engineering* 10.4 (2005), pp. 336–341.

[134] KP Sudheer and Ashu Jain. "Explaining the internal behaviour of artificial neural network river flow models." In: *Hydrological Processes* 18.4 (2004), pp. 833–844.

[135] E Toth, A Brath, and A Montanari. "Comparison of short-term rainfall prediction models for real-time flood forecasting." In: *Journal of hydrology* 239.1-4 (2000), pp. 132–147.

[136] Li-Chiu Chang, Fi-John Chang, and Yen-Ming Chiang. "A two-step-ahead recurrent neural network for stream-flow forecasting." In: *Hydrological Processes* 18.1 (2004), pp. 81–92.

[137] Fi-John Chang, Yen-Ming Chiang, and Li-Chiu Chang. "Multi-step-ahead neural networks for flood forecasting." In: *Hydrological sciences journal* 52.1 (2007), pp. 114–130.

[138] H Yonaba, F Anctil, and V Fortin. "Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting." In: *Journal of Hydrologic Engineering* 15.4 (2010), pp. 275–283.

[139] Nobutomo Osanai et al. "Japanese early-warning for debris flows and slope failures using rainfall indices with Radial Basis Function Network." In: *Landslides* 7.3 (2010), pp. 325–338.

[140] Masoud Bakhtyari Kia et al. "An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia." In: *Environmental Earth Sciences* 67.1 (2012), pp. 251–264.

[141] Lance E Besaw et al. "Advances in ungauged streamflow prediction using artificial neural networks." In: *Journal of Hydrology* 386.1-4 (2010), pp. 27–37.

[142] A Danandeh Mehr et al. "Successive-station monthly streamflow prediction using different artificial neural network algorithms." In: *International Journal of Environmental Science and Technology* 12.7 (2015), pp. 2191–2200.

[143] Yahui Di et al. "Developing machine learning tools for long-lead heavy precipitation prediction with multi-sensor data." In: *2015 IEEE 12th International Conference on Networking, Sensing and Control*. IEEE. 2015, pp. 63–68.

[144] G Napolitano et al. "A conceptual and neural network model for real-time flood forecasting of the Tiber River in Rome." In: *Physics and Chemistry of the Earth, Parts A/B/C* 35.3-5 (2010), pp. 187–194.

[145] Riccardo Taormina and Kwok-Wing Chau. "ANN-based interval forecasting of streamflow discharges using the LUBE method and MOFIPS." In: *Engineering Applications of Artificial Intelligence* 45 (2015), pp. 429–440.

[146] EJ Plate and L Duckstein. "Reliability in hydraulic design." In: *Engineering reliability and risk in water resources*. Springer, 1987, pp. 27–60.

[147] Lotfi A Zadeh. "Fuzzy sets as a basis for a theory of possibility." In: *Fuzzy sets and systems* 100 (1999), pp. 9–34.

[148] GJ Klir. "Methodological principles of uncertainty in inductive modelling: a new perspective." In: *Maximum-entropy and Bayesian methods in science and engineering.* Springer, 1988, pp. 295–304.

[149] George J Klir and Bo Yuan. "Fuzzy sets and fuzzy logic: theory and applications." In: *Possibility Theory versus Probab. Theory* 32.2 (1996), pp. 207–208.

[150] M Bruce Beck. "Water quality modeling: a review of the analysis of uncertainty." In: *Water Resources Research* 23.8 (1987), pp. 1393–1442.

[151] Keith Beven. "Prophecy, reality and uncertainty in distributed hydrological modelling." In: *Advances in water resources* 16.1 (1993), pp. 41–51.

[152] K Mizumura. *Application of fuzzy theory to snowmelt-runoff.* New uncertainty concepts in hydrology and water resources, 1995.

[153] Keith Beven. "How far can we go in distributed hydrological modelling?" In: (2001).

[154] Thorsten Wagener, Matthew J Lees, Howard S Wheater, et al. "A toolkit for the development and application of parsimonious hydrological models." In: *Mathematical models of small watershed hydrology* 2 (2001), pp. 1–34.

[155] J KINDLER and S TYSZEWSKI. "7 On the value of fuzzy concepts in hydrology and." In: *New Uncertainty Concepts in Hydrology and Water Resources* (2007), p. 126.

[156] W Feluch. "10 Nonparametric estimation of multivariate density and nonparametric regression." In: *New uncertainty concepts in hydrology and water resources* (2007), p. 145.

[157] Vladik Kreinovich, Hung T Nguyen, and Yeung Yam. "Fuzzy systems are universal approximators for a smooth function and its derivatives." In: *International Journal of Intelligent Systems* 15.6 (2000), pp. 565–574.

[158] Constantin Negoita. *Expert systems and fuzzy systems.* Benjamin-Cummings Pub. Co., Menlo Park, CA, 1985.

[159] Iphigenia Keramitsoglou, Constantinos Cartalis, and Chris T Kiranoudis. "Automatic identification of oil spills on satellite images." In: *Environmental modelling & software* 21.5 (2006), pp. 640–652.

[160] Guanrong Chen and Trung Tat Pham. *Introduction to fuzzy systems.* Chapman and Hall/CRC, 2005.

[161] PC Nayak, KP Sudheer, and KS Ramasastri. "Fuzzy computing based rainfall-runoff model for real time flood forecasting." In: *Hydrological Processes: An International Journal* 19.4 (2005), pp. 955–968.

[162] A Abrishamchi et al. "Stream flow forecasting and reservoir operation models using fuzzy inference systems." In: *Operating Reservoirs in Changing Conditions.* 2006, pp. 373–382.

[163] Mesut Cimen and Kemal SAPLIOĞLU. "Strem flow forecasting by fuzzy logic method." In: *International congress on river basin management.* 2007, pp. 612–620.

[164] Zeynel Fuat Toprak et al. “Modeling monthly mean flow in a poorly gauged basin by fuzzy logic.” In: *Clean–Soil, Air, Water* 37.7 (2009), pp. 555–564.

[165] Zacharia Katambara and John Ndiritu. “A fuzzy inference system for modelling streamflow: Case of Letaba River, South Africa.” In: *Physics and Chemistry of the Earth, Parts A/B/C* 34.10-12 (2009), pp. 688–700.

[166] Y Al-Zu’bi, A Sheta, J Al-Zu’bi, et al. “Nile River flow forecasting based Takagi-Sugeno fuzzy model.” In: *Journal of Applied Sciences* 10.4 (2010), pp. 284–290.

[167] Mahmut Firat and M Erkan Turan. “Monthly river flow forecasting by an adaptive neuro-fuzzy inference system.” In: *Water and environment journal* 24.2 (2010), pp. 116–125.

[168] Aytac Guven and Ali Aytek. “New approach for stage–discharge relationship: gene-expression programming.” In: *Journal of Hydrologic Engineering* 14.8 (2009), pp. 812–820.

[169] Liang-Cheng Chang, Chih-Chao Ho, and Yu-Wen Chen. “Applying multiobjective genetic algorithm to analyze the conflict among different water use sectors during drought period.” In: *Journal of Water Resources Planning and Management* 136.5 (2010), pp. 539–546.

[170] Taher Rajaee. “Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers.” In: *Science of the total environment* 409.15 (2011), pp. 2917–2928.

[171] Hazi Mohammad Azamathulla et al. “Gene-expression programming for the development of a stage-discharge curve of the Pahang River.” In: *Water resources management* 25.11 (2011), pp. 2901–2916.

[172] Ozgur Kisi, Jalal Shiri, and Bagher Nikoofar. “Forecasting daily lake levels using artificial intelligence approaches.” In: *Computers & Geosciences* 41 (2012), pp. 169–180.

[173] Ozgur Kisi, Jalal Shiri, and Mustafa Tombul. “Modeling rainfall-runoff process using soft computing techniques.” In: *Computers & Geosciences* 51 (2013), pp. 108–117.

[174] Vladan Babovic and Maarten Keijzer. “Rainfall runoff modelling based on genetic programming.” In: *Hydrology Research* 33.5 (2002), pp. 331–346.

[175] Lawrence J Fogel, Alvin J Owens, and Michael J Walsh. “Intelligent decision making through a simulation of evolution.” In: *Behavioral science* 11.4 (1966), pp. 253–272.

[176] John Henry Holland et al. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.* MIT press, 1992.

[177] John Holland. “Adaptation in natural and artificial systems: an introductory analysis with application to biology.” In: *Control and artificial intelligence* (1975).

[178] Hans-Paul Schwefel. *Numerical optimization of computer models.* John Wiley & Sons, Inc., 1981.

[179] John R Koza. “Genetic programming II: Automatic discovery of reusable subprograms.” In: *Cambridge, MA, USA* 13.8 (1994), p. 32.

[180] XK Wang et al. “Using time-delay neural network combined with genetic algorithms to predict runoff level of Linshan Watershed, Sichuan, China.” In: *Journal of Hydrologic Engineering* 12.2 (2007), pp. 231–236.

[181] A Makkeasorn, Ni-Bin Chang, and Xiaobing Zhou. “Short-term streamflow forecasting with global climate change implications–A comparative study between genetic programming and neural network models.” In: *Journal of hydrology* 352.3-4 (2008), pp. 336–354.

[182] Chang-Shian Chen, Chin-Hui Liu, and Hui-Chen Su. “A nonlinear time series analysis using two-stage genetic algorithms for streamflow forecasting.” In: *Hydrological Processes: An International Journal* 22.18 (2008), pp. 3697–3711.

[183] Aytac Guven. “Linear genetic programming for time-series modelling of daily flow rate.” In: *Journal of earth system science* 118.2 (2009), pp. 137–146.

[184] Qingwei Ni et al. “Evolutionary modeling for streamflow forecasting with minimal datasets: a case study in the West Malian River, China.” In: *Environmental Engineering Science* 27.5 (2010), pp. 377–385.

[185] SS Kashid, Subimal Ghosh, and Rajib Maity. “Streamflow prediction using multi-site rainfall obtained from hydroclimatic teleconnection.” In: *Journal of Hydrology* 395.1-2 (2010), pp. 23–38.

[186] Behrooz Keshtegar et al. “Optimized river stream-flow forecasting model utilizing high-order response surface method.” In: *Water Resources Management* 30.11 (2016), pp. 3899–3914.

[187] Wen Wang et al. “Forecasting daily streamflow using hybrid ANN models.” In: *Journal of Hydrology* 324.1-4 (2006), pp. 383–399.

[188] Ashu Jain and Avadhnam Madhav Kumar. “Hybrid neural network models for hydrologic time series forecasting.” In: *Applied Soft Computing* 7.2 (2007), pp. 585–592.

[189] Turgay Partal. “River flow forecasting using different artificial neural network algorithms and wavelet transform.” In: *Canadian Journal of Civil Engineering* 36.1 (2008), pp. 26–38.

[190] Niranjan Pramanik and Rabindra Kumar Panda. “Application of neural network and adaptive neuro-fuzzy inference systems for river flow prediction.” In: *Hydrological Sciences Journal* 54.2 (2009), pp. 247–260.

[191] Mohammad T Dastorani et al. “Application of ANN and ANFIS models for reconstructing missing flow data.” In: *Environmental monitoring and assessment* 166.1-4 (2010), pp. 421–434.

[192] Ramli Adnan et al. “New Artificial Neural Network and Extended Kalman Filter hybrid model of flood prediction system.” In: *2013 IEEE 9th International Colloquium on Signal Processing and its Applications*. IEEE. 2013, pp. 252–257.

[193] Celso Augusto Guimarães Santos and Gustavo Barbosa Lima da Silva. “Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models.” In: *Hydrological Sciences Journal* 59.2 (2014), pp. 312–324.

[194] Greer B Humphrey et al. “A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network.” In: *Journal of Hydrology* 540 (2016), pp. 623–640.

[195] Zaher Mundher Yaseen et al. “Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model.” In: *Journal of Hydrology* 554 (2017), pp. 263–276.

[196] Nuratiah Zaini et al. “Daily River Flow Forecasting with Hybrid Support Vector Machine–Particle Swarm Optimization.” In: *IOP Conference Series: Earth and Environmental Science*. Vol. 140. 1. IOP Publishing. 2018, p. 012035.

[197] Mitchell M Waldrop. *Complexity: The emerging science at the edge of order and chaos*. Simon and Schuster, 1993.

[198] S Chan. “Complex adaptive systems. ESD. 83 Research seminar in engineering systems.” In: *Complex adaptive systems, ESD* 83 (2001).

[199] Andrei Borshchev and Alexei Filippov. “From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools.” In: *Proceedings of the 22nd international conference of the system dynamics society*. Vol. 22. Citeseer. 2004.

[200] Brian L Heath. “The history, philosophy, and practice of agent-based modeling and the development of the conceptual model for simulation diagram.” In: (2010).

[201] Philip Agre and Stanley J Rosenschein. *Computational theories of interaction and agency*. Mit Press, 1996.

[202] David O’Sullivan and Mordechai Haklay. “Agent-based models and individualism: is the world agent-based?” In: *Environment and Planning A* 32.8 (2000), pp. 1409–1425.

[203] H Randy Gimblett. *Integrating geographic information systems and agent-based modeling techniques for simulating social and ecological processes*. Oxford University Press, 2002.

[204] Anette Reenberg. “Agricultural land use pattern dynamics in the Sudan-Sahel-towards an event-driven framework.” In: *Land Use Policy* 18.4 (2001), pp. 309–319.

[205] Nelson Minar et al. “The swarm simulation system: A toolkit for building multi-agent simulations.” In: (1996).

[206] Seth Tisue and Uri Wilensky. “Center for Connected Learning and Computer-Based Modeling Northwestern University, Evanston, Illinois.” In: ().

[207] Nicholson Collier, T Howe, and M North. “Onward and upward: The transition to Repast 2.0.” In: *Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference*. Vol. 122. 2003, p. 136.

[208] The AnyLogic Company. *AnyLogic simulation software*. 2019. URL: https://www.anylogic.com%20[Accessed%2018%20May%202019].

[209] Joshua M Epstein and Robert Axtell. *Growing artificial societies: social science from the bottom up*. Brookings Institution Press, 1996.

[210] Charles M Macal and Michael J North. “Tutorial on agent-based modeling and simulation part 2: how to model with agents.” In: *Proceedings of the 38th conference on Winter simulation*. Winter Simulation Conference. 2006, pp. 73–83.

[211] Charles M Macal and Michael J North. “Agent-based modeling and simulation: ABMS examples.” In: *2008 Winter Simulation Conference*. IEEE. 2008, pp. 101–112.

[212] John H Holland and John H Miller. “Artificial adaptive agents in economic theory.” In: *The American Economic Review* 81.2 (1991), pp. 365–370.

[213] Charles M Macal and Michael J North. “Tutorial on agent-based modelling and simulation.” In: *Journal of Simulation* 4.3 (Sept. 2010), pp. 151–162. DOI: 10.1057/jos.2010.3. URL: https://www.tandfonline.com/doi/full/10.1057/jos.2010.3.

[214] Paul Davidsson. “Multi agent based simulation: beyond social simulation.” In: *International workshop on multi-agent systems and agent-based simulation*. Springer. 2000, pp. 97–107.

[215] Stan Franklin and Art Graesser. “Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents.” In: *International Workshop on Agent Theories, Architectures, and Languages*. Springer. 1996, pp. 21–35.

[216] Michael Wooldridge. “Intelligent agents.” In: *Multiagent systems* 35.4 (1999), p. 51.

[217] Nicholas R Jennings. “On agent-based software engineering.” In: *Artificial intelligence* 117.2 (2000), pp. 277–296.

[218] Michael Luck and Mark d’Inverno. “A conceptual framework for agent definition and development.” In: *The Computer Journal* 44.1 (2001), pp. 1–20.

[219] Patrick Riley. “MPADES: Middleware for parallel agent discrete event simulation.” In: *Robot Soccer World Cup*. Springer. 2002, pp. 162–178.

[220] Charles M Macal and Michael J North. “Tutorial on agent-based modeling and simulation.” In: *Proceedings of the Winter Simulation Conference, 2005.* IEEE. 2005, 14–pp.

[221] Michael Wooldridge. “Agent-based software engineering.” In: *IEEE Proceedings-software* 144.1 (1997), pp. 26–37.

[222] Grimm Volker and SF Railsback. *Individual-based modeling and ecology*. 2005.

[223] Muaz Niazi and Amir Hussain. “Agent-based computing from multi-agent systems to agent-based models: a visual survey.” In: *Scientometrics* 89.2 (2011), p. 479.

[224] Leif Gustafsson and Mikael Sternad. “Consistent micro, macro and state-based population modelling.” In: *Mathematical biosciences* 225.2 (2010), pp. 94–107.

[225] Eric Bonabeau. “Agent-based modeling: Methods and techniques for simulating human systems.” In: *Proceedings of the national academy of sciences* 99.suppl 3 (2002), pp. 7280–7287.

[226] Mitchel Resnick. *Turtles, termites, and traffic jams: Explorations in massively parallel microworlds*. Mit Press, 1997.

[227] Jacques Ferber and Gerhard Weiss. *Multi-agent systems: an introduction to distributed artificial intelligence*. Vol. 1. Addison-Wesley Reading, 1999.

[228] Joshua M Epstein. *Generative social science: Studies in agent-based computational modeling*. Princeton University Press, 2006.

[229] Nigel Gilbert. *Agent-based models*. 153. Sage, 2008.

[230] Jean-Pierre Treuil, Alexis Drogoul, and Jean-Daniel Zucker. *Modélisation et simulation à base d’agents: exemples commentés, outils informatiques et questions théoriques.* Dunod, 2008.

[231] Steven F Railsback and Volker Grimm. *Agent-based and individual-based modeling: a practical introduction.* Princeton university press, 2019.

[232] Michael Luck, Peter McBurney, and Chris Preist. “A manifesto for agent technology: Towards next generation computing.” In: *Autonomous Agents and Multi-Agent Systems* 9.3 (2004), pp. 203–252.

[233] Lin Padgham and Michael Winikoff. *Developing intelligent agent systems: A practical guide.* Vol. 13. John Wiley & Sons, 2005.

[234] Melanie Hilario, Christian Pellegrini, and Frédéric Alexandre. *Modular integration of connectionist and symbolic processing in knowledge-based systems.* Centre de Recherche en Informatique de Nancy, 1994.

[235] Carlos A Iglesias, José C González, and Juan R Velasco. “MIX: A general purpose multiagent architecture.” In: *International Workshop on Agent Theories, Architectures, and Languages*. Springer. 1995, pp. 251–266.

[236] Andreas Scherer and Gunter Schlageter. *A Multi Agent Approach for the Integration of Neural Networks and Expert Systems*. 1994.

[237] R Khosla and T Dillon. *Engineering intelligent hybrid multi-agent systems*. Boston: Kluwer Academic Publishers, 1997.

[238] Miguel Delgado et al. “A multiagent architecture for fuzzy modeling.” In: *International Journal of Intelligent Systems* 14.3 (1999), pp. 305–329.

[239] Hans-Arno Jacobsen. “A generic architecture for hybrid intelligent systems.” In: *Deep fusion of computational and symbolic processing*. Springer, 2001, pp. 145–172.

[240] Stuart Russel and P Norwig. *A modern approach to artificial intelligence*. 1995.

[241] Zili Zhang and Chengqi Zhang. *Agent-based hybrid intelligent systems: an agent-based framework for complex problem solving*. Vol. 2938. Springer, 2004.

[242] Chong Hon Lim, Patricia Anthony, and Liau Chung Fan. “Applying multi-agent system in a context aware smart home.” In: *Learning* 24 (2009), pp. 53–64.

[243] Alex Rogers, Daniel D Corkill, and Nicholas R Jennings. “Agent technologies for sensor networks.” In: *IEEE Intelligent Systems* 2 (2009), pp. 13–17.

[244] Danny Weyns et al. “Environments for multiagent systems state-of-the-art and research challenges.” In: *International Workshop on Environments for Multi-Agent Systems*. Springer. 2004, pp. 1–47.

[245] Victor R Lesser. “Multiagent systems: An emerging subdiscipline of AI.” In: *ACM Computing Surveys (CSUR)* 27.3 (1995), pp. 340–342.

[246] Lael Parrott, René Lacroix, and Kevin M Wade. “Design considerations for the implementation of multi-agent systems in the dairy industry.” In: *Computers and electronics in agriculture* 38.2 (2003), pp. 79–98.

[247] A Roberto et al. “Towards a standardization of multi-agent system frameworks.” In: *Journal of the Association for Computing Machinery* 5.4 (1999).

[248] Roger B Myerson. *Game theory*. Harvard university press, 2013.

[249] Chengguang Lai et al. “A fuzzy comprehensive evaluation model for flood risk based on the combination weight of game theory.” In: *Natural Hazards* 77.2 (2015), pp. 1243–1259.

[250] Idel Montalvo et al. “Water distribution system computer-aided design by agent swarm optimization.” In: *Computer-Aided Civil and Infrastructure Engineering* 29.6 (2014), pp. 433–448.

[251] Katia P Sycara. “Multiagent systems.” In: *AI magazine* 19.2 (1998), pp. 79–79.

[252] François Bousquet and Christophe Le Page. “Multi-agent simulations and ecosystem management: a review.” In: *Ecological modelling* 176.3-4 (2004), pp. 313–332.

[253] Dino Borri, Domenico Camarda, and Anna De Liddo. “Mobility in environmental planning: an integrated multi-agent approach.” In: *International Conference on Cooperative Design, Visualization and Engineering*. Springer. 2005, pp. 119–129.

[254] Vladimir V Nikolic, Slobodan P Simonovic, and Dragan B Milicevic. “Analytical support for integrated water resources management: a new method for addressing spatial and temporal variability.” In: *Water resources management* 27.2 (2013), pp. 401–417.

[255] Sondoss Elsawah et al. “A methodology for eliciting, representing, and analysing stakeholder knowledge for decision making on complex socio-ecological systems: From cognitive maps to agent-based models.” In: *Journal of environmental management* 151 (2015), pp. 500–516.

[256] JE Gross et al. “Australian rangelands as complex adaptive systems: a conceptual model and preliminary results.” In: *Environmental Modelling & Software* 21.9 (2006), pp. 1264–1272.

[257] David Batten. “Are some human ecosystems self-defeating?” In: *Environmental Modelling & Software* 22.5 (2007), pp. 649–655.

[258] Claudia Pahl-Wostl. “Information, public empowerment, and the management of urban watersheds.” In: *Environmental Modelling & Software* 20.4 (2005), pp. 457–467.

[259] Olivier Barreteau et al. “Suitability of Multi-Agent Simulations to study irrigated system viability: application to case studies in the Senegal River Valley.” In: *Agricultural Systems* 80.3 (2004), pp. 255–275.

[260] Michael Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, 2009.

[261] Pattie Maes. “The agent network architecture (ANA).” In: *Acm sigart bulletin* 2.4 (1991), pp. 115–120.

[262] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited, 2016.

[263] Michael Wooldridge and Nicholas R Jennings. “Intelligent agents: Theory and practice.” In: *The knowledge engineering review* 10.2 (1995), pp. 115–152.

[264] Anand S Rao, Michael P Georgeff, et al. “BDI agents: from theory to practice.” In: *ICMAS*. Vol. 95. 1995, pp. 312–319.

[265] Gerhard Weiss. *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press, 1999.

[266] Lisa Brouwers and Magnus Boman. "A computational agent model of flood management strategies." In: *Regional Development: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2012, pp. 522–534.

[267] Sutee Anantsuksomsri and Nij Tontisirin. "Agent-based Modeling and Disaster Management." In: *Journal of Architectural/Planning Research and Studies* 10.2 (2013), pp. 1–14.

[268] G Coates et al. "Agent-based modelling and inundation prediction to enable the identification of businesses affected by flooding." In: *WIT Transactions on Ecology and the Environment* 184 (2014), pp. 13–22.

[269] Xicheng Tan et al. "Agent-and Cloud-Supported Geospatial Service Aggregation for Flood Response." In: *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 2.4 (2015).

[270] Emily Zechman Berglund. "Using agent-based modeling for water resources planning and management." In: *Journal of Water Resources Planning and Management* 141.11 (2015), p. 04015025.

[271] Michael Sunde et al. "Forecasting streamflow response to increased imperviousness in an urbanizing Midwestern watershed using a coupled modeling approach." In: *Applied geography* 72 (2016), pp. 14–25.

[272] Neiler Medina et al. "Agent based models for testing city evacuation strategies under a flood event as strategy to reduce flood risk." In: *EGU General Assembly Conference Abstracts*. Vol. 18. 2016.

[273] Erhu Du et al. "Impacts of Human Behavioral Heterogeneity on the Benefits of Probabilistic Flood Warnings: An Agent-Based Modeling Framework." In: *JAWRA Journal of the American Water Resources Association* 53.2 (2017), pp. 316–332.

[274] Katie Jenkins et al. "Assessing surface water flood risk and management strategies under future climate change: Insights from an Agent-Based Model." In: *Science of the Total Environment* 595 (2017), pp. 159–168.

[275] Liang Emlyn Yang et al. "Assessment of flood losses with household responses: Agent-based simulation in an urban catchment area." In: *Environmental Modeling & Assessment* 23.4 (2018), pp. 369–388.

[276] Brian Heath, Raymond Hill, and Frank Ciarallo. "A survey of agent-based modeling practices (January 1998 to July 2008)." In: *Journal of Artificial Societies and Social Simulation* 12.4 (2009), p. 9.

[277] RE Munich. *NatCatSERVICE: natural catastrophe know-how for risk management and research*. Tech. rep. Munich RE, 2018.

[278] Alex Rogers. “Agent technologies for sensor networks.” In: *The Computer Journal* 54.3 (2011), pp. 307–308.

[279] Tiago Pinto et al. “Multi-agent based electricity market simulator with VPP: Conceptual and implementation issues.” In: *2009 IEEE Power & Energy Society General Meeting*. IEEE. 2009, pp. 1–9.

[280] Lorena A Bearzotti, Enrique Salomone, and Omar J Chiotti. “An autonomous multi-agent approach to supply chain event management.” In: *International Journal of Production Economics* 135.1 (2012), pp. 468–478.

[281] Spyros Skarvelis-Kazakos et al. “Multiple energy carrier optimisation with intelligent agents.” In: *Applied energy* 167 (2016), pp. 323–335.

[282] Gabriel Villarrubia et al. “Combining multi-agent systems and wireless sensor networks for monitoring crop irrigation.” In: *Sensors* 17.8 (2017), p. 1775.

[283] Jean-Pierre Georgé et al. “Real-time simulation for flood forecast: an adaptive multi-agent system staff.” In: *Proceedings of the AISB*. Vol. 3. 2003, pp. 109–114.

[284] David De Roure et al. “FloodNet–Improving Flood Warning Times using Pervasive and Grid Computing.” In: (2005).

[285] ALEXANDRA MARIA Matei. “Multi-agent system for monitoring and analysis prahova hydrographical basin.” In: *Buletinul Institutului Politehnic din Iaşi, Automatic Control and Computer Science Section* 57 (2011), pp. 9–19.

[286] Vivian F López, Santo L Medina, and Juan F de Paz. “Taranis: Neural networks and intelligent agents in the early warning against floods.” In: *Expert Systems with Applications* 39.11 (2012), pp. 10031–10037.

[287] Marouane El Mabrouk et al. *An Improved Flood Forecasting and Warning System using Data Mining*. Tech. rep. 11. 2013, p. 2277. URL: www.ijarcsse.com.

[288] El Mabrouk Marouane, Ezziyyani Mostafa, and Essaaidi Mohamed. “Intelligent data classification and aggregation in wireless sensors for flood forecasting system.” In: *Mediterranean Microwave Symposium* 2015-April (2014).

[289] RS Iqbal et al. “A mobile agent-based algorithm for prediction of inundation area.” In: *Research Journal of Recent Sciences* 3.1 (2014), pp. 72–77.

[290] Bin Linghu and Feng Chen. “An intelligent multi-agent approach for flood disaster forecasting utilizing case based reasoning.” In: *2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications*. IEEE. 2014, pp. 182–185.

[291] Omar Al-Azzam et al. “Flood Prediction and Risk Assessment Using Advanced Geo-Visualization and Data Mining Techniques: A Case Study in the Red-Lake Valley, Applied Computational Science.” In: *Proceedings of Applied Computational Science* (2014).

[292] Shitai Bao et al. "Urban Water-log Simulation and Prediction based on Multi-Agent Systems." In: 2015.

[293] Teh Noranis Mohd Aris et al. "An agent-based ontology fuzzy logic conceptual model for flood warning prediction." In: *Academics World 12th International Conference, 20 Dec. 2015, Singapore* (2015), pp. 36–39.

[294] Marouane El Mabrouk et al. "New Expert System for Short, Medium and Long-Term Flood Forecasting and Warning." In: *Journal of Theoretical & Applied Information Technology* 78.2 (2015).

[295] Marouane El Mabrouk and Salma Gaou. "Proposed Intelligent Pre-Processing Model of Real-Time Flood Forecasting and Warning for Data Classification and Aggregation." In: *International Journal of Online Engineering* 13.11 (2017).

[296] M Guijarro and R Fuentes-Fernández. "A multi-agent system architecture for sensor networks. " In Multi-Agent Systems-Modeling, Control, Programming." In: *Simulations and Applications* (2011). URL: https://www.intechopen.com/download/pdf/14507.

[297] Abdelhakim Hamzi et al. "Multi-Agent Architecture for the Design of WSN Applications." In: *Wireless Sensor Network* 05.02 (2013), pp. 14–25.

[298] Y Liu et al. "Agent-based flood evacuation simulation of life-threatening conditions using vitae system model." In: *Journal of Natural Disaster Science* 31.2 (2009), pp. 69–77. URL: https://www.jstage.jst.go.jp/article/jnds/31/2/31_2_69/_article/-char/ja/.

[299] Richard J. Dawson, Roger Peppe, and Miao Wang. "An agent-based model for risk-based flood incident management." In: *Natural Hazards* 59.1 (Oct. 2011), pp. 167–189. ISSN: 0921-030X. DOI: 10.1007/s11069-011-9745-4. URL: http://link.springer.com/10.1007/s11069-011-9745-4.

[300] R Lempert. "Agent-based modeling as organizational and public policy simulators." In: *Proceedings of the National Academy of Sciences* 99.3 (2002), pp. 7195–7196. URL: https://www.pnas.org/content/99/suppl_3/7195.short.

[301] Magid Nikraz A, Giovanni Caire B, and Parisa A. Bahri A. *A Methodology for the Analysis and Design of Multi-Agent Systems using JADE*. 2006.

[302] Edouard Amouroux et al. "GAMA: an environment for implementing and running spatially explicit multi-agent simulations." In: *Pacific Rim International Conference on Multi-Agents*. Springer. 2007, pp. 359–371.

[303] Patrick Taillandier et al. "Building, composing and experimenting complex spatial models with the GAMA platform." In: *GeoInformatica* 23.2 (2019), pp. 299–322.

[304] David RA Carruth and Francis P Smith. "Assessing Flood Hydrology in Data Scarce Tropical regions: a Congo (ROC) case study." In: (2017).

[305] Claude Elwood Shannon. "A mathematical theory of communication." In: *Bell system technical journal* 27.3 (1948), pp. 379–423.

[306] Alexis Drogoul. "De la simulation multi-agents a la resolution collective de problemes: une etude de l'emergence de structures d'organisation dans les systemes multi-agents." PhD thesis. Paris 6, 1993.

[307] Bernard Burg and Vice President FIPA. "Foundation for intelligent physical agents." In: *Official FIPA presentation, Lausanne, February* (2002).

[308] Jan Svennevig. *Getting acquainted in conversation: A study of initial interactions*. Vol. 64. John Benjamins Publishing, 2000.

[309] Michael Bratman. "Intention, Plans, and Practical Reason." In: (1987).

[310] Anand S Rao and Michael P Georgeff. "BDI Agents: From Theory to Practice." In: 95 (1995), pp. 312–319. URL: www.aaai.org.

[311] Patrick Taillandier et al. "A BDI agent architecture for the GAMA modeling and simulation platform." In: *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer. 2016, pp. 3–23.

[312] Foundation for Intelligent Physical Agents. *FIPA Standard Specifications*. 2002. URL: http://www.fipa.org/repository/standardspecs.html[Accessed%2013%20Feb.%202018]..

[313] Minera Panama S.A. MPSA. *Environmental Impact Assessment Study*. 2010.

[314] Victor Mockus. *National engineering handbook*. Vol. 4. Section, 1964.

[315] WR Wieder et al. "Regridded harmonized world soil database v1. 2." In: *ORNL DAAC* (2014).

[316] Kevin Sene. *Flash floods: forecasting and warning*. Springer Science & Business Media, 2013.

[317] Michael Uschold and Martin King. *Towards a methodology for building ontologies*. Citeseer, 1995.

[318] Michael Uschold, Michael Gruninger, et al. "Ontologies: Principles, methods and applications." In: *Technical Report-University of Edinburgh Artificial Intelligence Applications Institute AIAI TR* (1996).

[319] Antonio De Nicola, Michele Missikoff, and Roberto Navigli. "A software engineering approach to ontology building." In: *Information systems* 34.2 (2009), pp. 258–275.

[320] Ricardo de Almeida Falbo. "SABiO: Systematic Approach for Building Ontologies." In: *ONTO. COM/ODISE@ FOIS*. 2014.

[321] Nicola Guarino. *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy*. Vol. 46. IOS press, 1998.

[322] Natalya F Noy, Deborah L McGuinness, et al. *Ontology development 101: A guide to creating your first ontology*. 2001.

[323] Dragan Gaševic, Dragan Djuric, and Vladan Devedžic. *Model driven architecture and ontology development*. Springer Science & Business Media, 2006.

[324] Robert Neches et al. "Enabling technology for knowledge sharing." In: *AI magazine* 12.3 (1991), pp. 36–36.

[325] Manas Gaur et al. "empathi: An ontology for emergency managing and planning about hazard crisis." In: *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*. IEEE. 2019, pp. 396–403.

[326] Maroua Masmoudi et al. "MEMOn: Modular Environmental Monitoring Ontology to link heterogeneous Earth observed data." In: *Environmental Modelling & Software* 124 (2020), p. 104581.

[327] Ruben Costa et al. "Facilitating knowledge sharing and reuse in building and construction domain: an ontology-based approach." In: *Journal of Intelligent Manufacturing* 27.1 (2016), pp. 263–282.

[328] Walter Terkaj, Georg Ferdinand Schneider, and Pieter Pauwels. "Reusing domain ontologies in linked building data: the case of building automation and control." In: *8th International Workshop on Formal Ontologies meet Industry*. Vol. 2050. 2017.

[329] Nikolaos Trokanas and Franjo Cecelja. "Ontology evaluation for reuse in the domain of process systems engineering." In: *Computers & Chemical Engineering* 85 (2016), pp. 177–187.

[330] Hendrik Walzel et al. "An Approach for an Automated Adaption of KPI Ontologies by Reusing Systems Engineering Data." In: *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE. 2019, pp. 1693–1696.

[331] Amelie Gyrard et al. "Reusing and unifying background knowledge for internet of things with LOV4IoT." In: *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)*. IEEE. 2016, pp. 262–269.

[332] Tatyana Ivanova. "E-Learning resource reuse, based on bilingual ontology annotation and ontology mapping." In: *International Journal of Advanced Computer Research* 9.45 (2019), pp. 351–364.

[333] Sabrina Azzi, Michal Iglewski, and Véronique Nabelsi. "COMPETENCY QUESTIONS FOR BIOMEDICAL ONTOLOGY REUSE." In: *ProBernard Burg and Vice President FIPAa Computer Science* 160 (2019), pp. 362–368.

[334] Patrick Klein, Lukas Malburg, and Ralph Bergmann. “FTOnto: A Domain Ontology for a Fischertechnik Simulation Production Factory by Reusing Existing Ontologies.” In: *LWDA*. 2019, pp. 253–264.

[335] Cheah Wai Shiang et al. “Ontology reuse for multiagent system development through pattern classification.” In: *Software: Practice and Experience* 48.11 (2018), pp. 1923–1939.

[336] Zheng Ma et al. “The application of ontologies in multi-agent systems in the energy sector: a scoping review.” In: *Energies* 12.16 (2019), p. 3200.

[337] Javier Cuenca, Felix Larrinaga, and Edward Curry. “DABGEO: A reusable and usable global energy ontology for the energy domain.” In: *Journal of Web Semantics* (2020), p. 100550.

[338] Mariano Fernández-López, Asunción Gómez-Pérez, and Natalia Juristo. “Methontology: from ontological art towards ontological engineering.” In: (1997).

[339] Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, and Mariano Fernández-López. “The NeOn methodology for ontology engineering.” In: *Ontology engineering in a networked world*. Springer, 2012, pp. 9–34.

[340] Fabiano B Ruy et al. “From reference ontologies to ontology patterns and back.” In: *Data & Knowledge Engineering* 109 (2017), pp. 41–69.

[341] Peter Rittgen. “Translating Metaphors into Design Patterns.” In: *Advances in Information Systems Development*. Springer, 2006, pp. 425–436.

[342] Emhimed Salem Alatrish. “Comparison of ontology editors.” In: *eRAF Journal on Computing* 4 (2012), pp. 23–38.

[343] Thabet Slimani. “Ontology development: A comparing study on tools, languages and formalisms.” In: *Indian Journal of Science and Technology* 8.24 (2015), pp. 1–12.

[344] Michel Dirix, Alexis Muller, and Vincent Aranega. “Genmymodel: an online uml case tool.” In: (2013).

[345] Hassan Gomaa. *Software modeling and design: UML, use cases, patterns, and software architectures*. Cambridge University Press, 2011.

[346] XML OMG. *XML Metadata Interchange (XMI) Specification*. 2015. URL: https://www.omg.org/spec/XMI/2.5.1/%20[Accessed%2008%20Sep.%202020].

[347] World Meteorological Organization. *International Meteorological Vocabulary: Vocabulaire Météorologique International. Vocabulario Meteorológico Internacional*. Secretariat of the World Meteorological Organization, 1966.

[348] Philip C Saksa et al. “Fuels treatment and wildfire effects on runoff from Sierra Nevada mixed-conifer forests.” In: *Ecohydrology* 13.3 (2020), e2151.

[349] Roderick JA Little and Donald B Rubin. *Statistical analysis with missing data*. Vol. 793. John Wiley & Sons, 2019.

[350] L Wilkinson et al. “Task Force on Statistical Inference. Statistical methods in psychology journals.” In: *American Psychologist* 54.3 (1999), pp. 594–604.

[351] Todd E Bodner. “Missing data: Prevalence and reporting practices.” In: *Psychological Reports* 99.3 (2006), pp. 675–680.

[352] James L Peugh and Craig K Enders. “Missing data in educational research: A review of reporting practices and suggestions for improvement.” In: *Review of educational research* 74.4 (2004), pp. 525–556.

[353] Donald B Rubin. “Inference and missing data.” In: *Biometrika* 63.3 (1976), pp. 581–592.

[354] Stef van Buuren and Karin Groothuis-Oudshoorn. “mice: Multivariate Imputation by Chained Equations in R.” In: *Journal of Statistical Software* 45.3 (2011), pp. 1–67. URL: https://www.jstatsoft.org/v45/i03/.

[355] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2019. URL: http://www.R-project.org/.

[356] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc. Boston, MA, 2015. URL: http://www.rstudio.com/.

[357] Stef van Buuren and Karin Groothuis-Oudshoorn. “mice: Multivariate Imputation by Chained Equations in R.” In: *Journal of Statistical Software, Articles* 45.3 (2011), pp. 1–67. ISSN: 1548-7660. DOI: 10.18637/jss.v045.i03. URL: https://www.jstatsoft.org/v045/i03.

[358] Sally J Priest et al. “Assessing options for the development of surface water flood warning in England and Wales.” In: *Journal of environmental management* 92.12 (2011), pp. 3038–3048.

[359] AP Hurford et al. “Validating the return period of rainfall thresholds used for E xtreme R ainfall A lerts by linking rainfall intensities with observed surface water flood events.” In: *Journal of Flood Risk Management* 5.2 (2012), pp. 134–142.

[360] Janice Blanc et al. “Enhanced efficiency of pluvial flood risk estimation in urban areas using spatial–temporal rainfall simulations.” In: *Journal of Flood Risk Management* 5.2 (2012), pp. 143–152.

[361] Chad Shouquan Cheng et al. “Climate change and heavy rainfall-related water damage insurance claims and losses in Ontario, Canada.” In: *Journal of Water Resource and Protection* 4.02 (2012), p. 49.

[362] Jiun-Huei Jang. "An advanced method to apply multiple rainfall thresholds for urban flood warnings." In: *Water* 7.11 (2015), pp. 6056–6078.

[363] WMO Słownik. *International meteorological vocabulary, 1992*. Tech. rep. WMO/OMM/IMGW, 182, Geneva.

[364] WMO. *Guidelines on cross-border exchange of warnings*. Tech. rep. 2003.

[365] I Sofiati and A Nurlatifah. "The prediction of rainfall events using WRF (weather research and forecasting) model with ensemble technique." In: *IOP Conference Series: Earth and Environmental Science*. Vol. 374. 1. IOP Publishing. 2019, p. 012036.

[366] David Dunkerley. "Rain event properties in nature and in rainfall simulation experiments: a comparative review with recommendations for increasingly systematic study and reporting." In: *Hydrological Processes: An International Journal* 22.22 (2008), pp. 4415–4435.

[367] David L Dunkerley. "How do the rain rates of sub-event intervals such as the maximum 5-and 15-min rates (I5 or I30) relate to the properties of the enclosing rainfall event?" In: *Hydrological Processes* 24.17 (2010), pp. 2425–2439.

[368] Bill Scharffenberg et al. *Hydrologic Engineering Center, 2009. HEC-DSSVue User's Manual - Version 2.0. U.S.* Tech. rep. 2009.

[369] T Patrick and A Drogoul. "From GIS Data to GIS Agents Modeling with the GAMA simulation platform." In: *TF SIM* (2010).

[370] NASA/METI/AIST/Japan Spacesystems and U.S./Japan ASTER Science Team (2019). *ASTER Global Digital Elevation Model V003 [Data set]. NASA EOSDIS Land Processes DAAC*. 2019. URL: https://doi.org/10.5067/ASTER/ASTGTM.003/%20[Aug.%2025%202019]..

[371] John Lindsay. "The Whitebox Geospatial Analysis Tools project and open-access GIS." In: Apr. 2014, pp. 1–8.

[372] QGIS Development Team. *QGIS Geographic Information System*. Open Source Geospatial Foundation. 2009. URL: http://qgis.osgeo.org.

[373] George J Arcement and Verne R Schneider. *Guide for selecting Manning's roughness coefficients for natural channels and flood plains*. 1989. URL: https://dpw.lacounty.gov/lacfcd/wdr/files/WG/041615/Guide%20for%20Selecting%20n-Value.pdf.

[374] R Cronshey. *Urban hydrology for small watersheds. Technical release 55 (TR-55). Washington, DC: US Department of Agriculture, Soil Conservation Service, Engineering Division*. 2018.

[375] V. T. Chow. *Hydrologic determination of waterway areas for the design of drainage structures in small drainage basins*. Tech. rep. 462. 1962, p. 104.

[376] Arlen D. Feldman. *Hydrologic Modeling System HEC-HMS Technical Reference Manual*. Tech. rep. 2000.

[377] André Musy, Benoit Hingray, and Cécile Picouet. *Hydrology: a science for engineers*. CRC press, 2014.

[378] LeRoy K Sherman. "The relation of hydrographs of runoff to size and character of drainage-basins." In: *Eos, Transactions American Geophysical Union* 13.1 (1932), pp. 332–339.

[379] Peter Mendham and Tim Clarke. "Macsim: A simulink enabled environment for multi-agent system simulation." In: *IFAC Proceedings Volumes* 38.1 (2005), pp. 325–329.

[380] Susan L Neitsch et al. *Soil and water assessment tool theoretical documentation version 2009*. Tech. rep. Texas Water Resources Institute, 2011.

[381] Jose Simmonds, Juan A Gómez, and Agapito Ledezma. "The role of agent-based modeling and multi-agent systems in flood-based hydrological problems: a brief review." In: *Journal of Water and Climate Change* (2019).

[382] Alexis Drogoul et al. "Gama: A spatially explicit, multi-level, agent-based modeling and simulation platform." In: *International Conference on Practical Applications of Agents and Multi-Agent Systems*. Springer. 2013, pp. 271–274.

[383] Patrick Taillandier et al. "GAMA: a simulation platform that integrates geographical information data, agent-based modeling and multi-scale control." In: *International Conference on Principles and Practice of Multi-Agent Systems*. Springer. 2012, pp. 242–258.

[384] Andrew Crooks, Christian Castle, and Michael Batty. "Key challenges in agent-based modelling for geo-spatial simulation." In: *Computers, Environment and Urban Systems* 32.6 (2008), pp. 417–430.

[385] Franziska Klügl. "A validation methodology for agent-based simulations." In: *Proceedings of the 2008 ACM symposium on Applied computing*. 2008, pp. 39–43.

[386] Hazel Parry. "Agent Based Modeling, Large Scale Simulations." In: *Computational Complexity: Theory, Techniques, and Applications* (Nov. 2013). DOI: 10.1007/978-0-387-30440-3_9.

[387] Andrew Crooks and A.J. Heppenstall. "Introduction to Agent-Based Modelling." In: Jan. 2012, pp. 85–105. DOI: 10.1007/978-90-481-8927-4_5.

[388] Duc-An Vo, Alexis Drogoul, and Jean-Daniel Zucker. "An operational meta-model for handling multiple scales in agent-based simulations." In: *2012 IEEE RIVF International Conference on Computing & Communication Technologies, Research, Innovation, and Vision for the Future*. IEEE. 2012, pp. 1–6.

[389] A. Rogers and P. V. Tessin. "Multi-Objective Calibration For Agent-Based Models." In: *Proceedings 5th Workshop on Agent-Based Simulation. H. Coelho, B. Espinasse, eds. (c) SCS Europe BVBA, 2004 ISBN 3-936150-31-1 (book) / 3-936150-32-X (CD).* 2004.

[390] Linda See et al. "Calibration and validation of agent-based models of land cover change." In: *Agent-based models of geographical systems*. Springer, 2012, pp. 181–197.

[391] Soon Thiam Khu and Henrik Madsen. "Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration." In: *Water Resources Research* 41.3 (2005).

[392] Omar Baqueiro Espinosa. "A genetic algorithm for the calibration of a micro-simulation model." In: *arXiv preprint arXiv:1201.3456* (2012).

[393] Enrique Canessa and Sergio Chaigneau. "Calibrating Agent-Based Models Using a Genetic Algorithm." In: *Studies in Informatics and Control* 24 (Mar. 2015), pp. 79–90. DOI: 10.24846/v24i1y201509.

[394] Chase Cockrell and Gary An. "Genetic Algorithms for model refinement and rule discovery in a high-dimensional agent-based model of inflammation." In: *bioRxiv* (2019), p. 790394.

[395] Philippe Caillou. "Automated Multi-Agent Simulation Generation and Validation." In: vol. 7057. Nov. 2010, pp. 398–412. DOI: 10.1007/978-3-642-25920-3_29.

[396] Manuel Moyo Oliveros and Kai Nagel. "Automatic calibration of agent-based public transit assignment path choice to count data." In: *Transportation Research Part C: Emerging Technologies* 64 (2016), pp. 58–71.

[397] Alessio Lomuscio, Hongyang Qu, and Franco Raimondi. "MCMAS: an open-source model checker for the verification of multi-agent systems." In: *International Journal on Software Tools for Technology Transfer* 19.1 (2017), pp. 9–30.

[398] Jérémy Boes and Frédéric Migeon. "Self-organizing multi-agent systems for the control of complex systems." In: *Journal of Systems and Software* 134 (2017), pp. 12–28.

[399] Zhengchun Liu et al. "A simulation and optimization based method for calibrating agent-based emergency department models under data scarcity." In: *Computers & Industrial Engineering* 103 (2017), pp. 300–309.

[400] Jan Drchal, Michal Čertickỳ, and Michal Jakob. "Data driven validation framework for multi-agent activity-based models." In: *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer. 2015, pp. 55–67.

[401] Minh Thai Truong. "To Develop a Database Management Tool for Multi-Agent Simulation Platform." PhD thesis. 2015.

[402] PB Duda et al. “BASINS/HSPF: Model use, calibration, and validation.” In: *Transactions of the ASABE* 55.4 (2012), pp. 1523–1547.

[403] André Fonseca et al. “Watershed model parameter estimation and uncertainty in data-limited environments.” In: *Environmental Modelling & Software* 51 (2014), pp. 84–93.

[404] Arekhi Saleh, Rostamizad Ghobad, and Rostami Noredin. “Evaluation of HEC-HMS methods in surface runoff simulation (Case study: Kan watershed, Iran).” In: *Advances in Environmental Biology* (2011), pp. 1316–1322.

[405] W Harry Doyle and Jeffrey E Miller. *Calibration of a distributed routing rainfall-runoff model at four urban sites near Miami, Florida.* Vol. 80. 1. US Geological Survey, Water Resources Division, Gulf Coast Hydroscience, 1980.

[406] Patricia J Shade. *Hydrology and sediment transport, Moanalua Valley, Oahu, Hawaii.* Vol. 84. 4156. US Geological Survey, 1984.

[407] Laura D Goodwin and Nancy L Leech. “Understanding correlation: Factors that affect the size of r.” In: *The Journal of Experimental Education* 74.3 (2006), pp. 249–266.

[408] TW Chu et al. “Evaluation of the SWAT modelâÂÂÂÂ™s sediment and nutrient components in the Piedmont physiographic region of Maryland.” In: *Transactions of the ASAE* 47.5 (2004), p. 1523.

[409] Bryan A Tolson and Christine A Shoemaker. “Watershed modeling of the Cannonsville basin using SWAT2000: model development, calibration and validation for the prediction of flow, sediment and phosphorus transport to the Cannonsville Reservoir.” In: *Rep. Prepared for Cornell Univ* (2004).

[410] Daniel N Moriasi et al. “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations.” In: *Transactions of the ASABE* 50.3 (2007), pp. 885–900.

[411] M Larose et al. “Hydrologic and atrazine simulation of the Cedar Creek watershed using the SWAT model.” In: *Journal of environmental quality* 36.2 (2007), pp. 521–531.

[412] Ali H Ahmed Suliman et al. “Comparison of semi-distributed, GIS-based hydrological models for the prediction of streamflow in a large catchment.” In: *Water resources management* 29.9 (2015), pp. 3095–3110.

[413] Donizete dos R Pereira et al. “Hydrological simulation in a basin of typical tropical climate and soil using the SWAT model part I: Calibration and validation tests.” In: *Journal of Hydrology: Regional Studies* 7 (2016), pp. 14–37.

[414] Jungang Gao et al. “Impacts of alternative climate information on hydrologic processes with SWAT: A comparison of NCDC, PRISM and NEXRAD datasets.” In: *Catena* 156 (2017), pp. 353–364.

[415] Giovanni Francesco Ricci et al. “Identifying sediment source areas in a Mediterranean watershed using the SWAT model.” In: *Land Degradation & Development* 29.4 (2018), pp. 1233–1248.

[416] Won Seok Jang, Bernard Engel, and Jichul Ryu. “Efficient flow calibration method for accurate estimation of baseflow using a watershed scale hydrological model (SWAT).” In: *Ecological engineering* 125 (2018), pp. 50–67.

[417] Giovanna Di Marzo Serugendo, Marie-Pierre Gleizes, and Anthony Karageorgos. “Self-organising Systems.” In: *Self-organising Software: From Natural to Artificial Adaptation*. Ed. by Giovanna Di Marzo Serugendo, Marie-Pierre Gleizes, and Anthony Karageorgos. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 7–32. ISBN: 978-3-642-17348-6. DOI: 10.1007/978-3-642-17348-6_2. URL: https://doi.org/10.1007/978-3-642-17348-6_2.

[418] A. Vaccaro et al. “A Self-Organizing Architecture for Decentralized Smart Microgrids Synchronization, Control, and Monitoring.” In: *IEEE Transactions on Industrial Informatics* 11.1 (2015), pp. 289–298. DOI: 10.1109/TII.2014.2342876.

[419] Filiberto Fele et al. “Coalitional control for self-organizing agents.” In: *IEEE Transactions on Automatic Control* 63.9 (2018), pp. 2883–2897.

[420] Ying Zhang and Chen Ling. “A strategy to apply machine learning to small datasets in materials science.” In: *Npj Computational Materials* 4.1 (2018), pp. 1–8.

[421] Hidetaka Taniguchi, Hiroshi Sato, and Tomohiro Shirakawa. “A machine learning model with human cognitive biases capable of learning from small and biased datasets.” In: *Scientific reports* 8.1 (2018), pp. 1–13.

[422] Andrius Vabalas et al. “Machine learning algorithm validation with a limited sample size.” In: *PloS one* 14.11 (2019), e0224365.

[423] Cesar F Caiafa et al. *Machine Learning Methods with Noisy, Incomplete or Small Datasets*. 2021.

[424] Mark Hall et al. “The WEKA data mining software: an update.” In: *ACM SIGKDD explorations newsletter* 11.1 (2009), pp. 10–18.

[425] Johannes Dahlke et al. “Is the Juice Worth the Squeeze? Machine Learning (ML) In and For Agent-Based Modelling (ABM).” In: *arXiv preprint arXiv:2003.11985* (2020).

[426] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. “Regularization paths for generalized linear models via coordinate descent.” In: *Journal of statistical software* 33.1 (2010), p. 1.

[427] Hui Zou and Trevor Hastie. "Regularization and variable selection via the elastic net." In: *Journal of the royal statistical society: series B (statistical methodology)* 67.2 (2005), pp. 301–320.

[428] Brian D Ripley. *Pattern recognition and neural networks*. Cambridge university press, 2007.

[429] David Meyer et al. "Package âÂÂÂÂe1071âÂÂ™." In: *The R Journal* (2019).

[430] Leo Breiman. "Random forests." In: *Machine learning* 45.1 (2001), pp. 5–32.

[431] L Breiman et al. "CART." In: *Classification and Regression Trees, Wadsworth and Brooks/Cole, Monterey, CA* (1984).

[432] Tianqi Chen et al. "Xgboost: extreme gradient boosting." In: *R package version 0.4-2* 1.4 (2015).

[433] Jerome H Friedman. "Stochastic gradient boosting." In: *Computational statistics & data analysis* 38.4 (2002), pp. 367–378.

[434] Frank Rosenblatt. "The perceptron: a probabilistic model for information storage and organization in the brain." In: *Psychological review* 65.6 (1958), p. 386.

[435] Bernd Bischl et al. "mlr: Machine Learning in R." In: *The Journal of Machine Learning Research* 17.1 (2016), pp. 5938–5942.

[436] Janez Demšar. "Statistical comparisons of classifiers over multiple data sets." In: *The Journal of Machine Learning Research* 7 (2006), pp. 1–30.

[437] Nathalie Japkowicz and Mohak Shah. *Evaluating learning algorithms: a classification perspective*. Cambridge University Press, 2011.

[438] Sebastian Raschka. "Model evaluation, model selection, and algorithm selection in machine learning." In: *arXiv preprint arXiv:1811.12808* (2018).

[439] Syed Muzamil Basha and Dharmendra Singh Rajput. "Survey on Evaluating the Performance of Machine Learning Algorithms: Past Contributions and Future Roadmap." In: *Deep Learning and Parallel Computing Environment for Bioengineering Systems*. Elsevier, 2019, pp. 153–164.

[440] Sylvain Arlot, Alain Celisse, et al. "A survey of cross-validation procedures for model selection." In: *Statistics surveys* 4 (2010), pp. 40–79.

[441] Christoph Bergmeir and José M Benítez. "On the use of cross-validation for time series predictor evaluation." In: *Information Sciences* 191 (2012), pp. 192–213.

[442] Marcel Neunhoeffer and Sebastian Sternberg. "How cross-validation can go wrong and what to do about it." In: *Political Analysis* 27.1 (2019), pp. 101–106.

[443] Claire Wagner-Rémy. “La pensée dirigée: Traité sur le raisonnement et les logiques [Guided Thinking: A Treatise on Reasoning and Logic].” In: *Edition: BoD–Books on Demand, Paris* (2016).

[444] Ebrahim H Mamdani. “Application of fuzzy algorithms for control of simple dynamic plant.” In: *Proceedings of the institution of electrical engineers*. Vol. 121. 12. IET. 1974, pp. 1585–1588.

[445] Ebrahim H Mamdani and Sedrak Assilian. “An experiment in linguistic synthesis with a fuzzy logic controller.” In: *International journal of man-machine studies* 7.1 (1975), pp. 1–13.

[446] Tomohiro Takagi and Michio Sugeno. “Fuzzy identification of systems and its applications to modeling and control.” In: *IEEE transactions on systems, man, and cybernetics* 1 (1985), pp. 116–132.

[447] SN Sivanandam, Sai Sumathi, SN Deepa, et al. *Introduction to fuzzy logic using MATLAB*. Vol. 1. Springer, 2007.

[448] L-X Wang and Jerry M Mendel. “Generating fuzzy rules by learning from examples.” In: *IEEE Transactions on systems, man, and cybernetics* 22.6 (1992), pp. 1414–1427.

[449] Gianluca Bontempi, Souhaib Ben Taieb, and Yann-Aël Le Borgne. “Machine learning strategies for time series forecasting.” In: *European business intelligence summer school*. Springer. 2012, pp. 62–77.

[450] Fabian Pedregosa et al. “Scikit-learn: Machine learning in Python.” In: *the Journal of machine Learning research* 12 (2011), pp. 2825–2830.

[451] Leo Breiman et al. *Classification and regression trees*. Routledge, 2017.

[452] Vladimir Vapnik, Steven E Golowich, Alex Smola, et al. “Support vector method for function approximation, regression estimation, and signal processing.” In: *Advances in neural information processing systems* (1997), pp. 281–287.

[453] Alexander J Smola and Bernhard Schölkopf. “On a kernel-based method for pattern recognition, regression, approximation, and operator inversion.” In: *Algorithmica* 22.1 (1998), pp. 211–231.

[454] George M Lady. “Evaluating long term forecasts.” In: *Energy Economics* 32.2 (2010), pp. 450–457.

[455] Ian T Jolliffe and David B Stephenson. *Forecast verification: a practitioner’s guide in atmospheric science*. John Wiley & Sons, 2012.

[456] Shannon R Fye et al. “An examination of factors affecting accuracy in technology forecasts.” In: *Technological Forecasting and Social Change* 80.6 (2013), pp. 1222–1231.

[457] Bahram Choubin et al. “Precipitation forecasting using classification and regression trees (CART) model: a comparative study of different approaches.” In: *Environmental earth sciences* 77.8 (2018), pp. 1–13.

[458] Giancarlo Fortino, Alfredo Garro, and Wilma Russo. “An integrated approach for the development and validation of multi-agent systems.” In: *International Journal of Computer Systems Science & Engineering* 20.4 (2005), pp. 259–271.

[459] J Eamonn Nash and Jonh V Sutcliffe. “River flow forecasting through conceptual models part Iâ”A discussion of principles.” In: *Journal of hydrology* 10.3 (1970), pp. 282–290.

[460] Nicolas Becu et al. “Agent based simulation of a small catchment water management in northern Thailand: description of the CATCHSCAPE model.” In: *Ecological modelling* 170.2-3 (2003), pp. 319–331.

[461] Maja Schlüter and Claudia Pahl-Wostl. “Mechanisms of resilience in common-pool resource management systems: an agent-based model of water use in a river basin.” In: *Ecology and Society* 12.2 (2007).

[462] Georg Holtz and Claudia Pahl-Wostl. “An agent-based model of groundwater overexploitation in the Upper Guadiana, Spain.” In: *Regional Environmental Change* 12.1 (2012), pp. 95–121.

[463] Nicolas G Adrien. *Computational hydraulics and hydrology: an illustrated dictionary*. CRC Press, 2004.

[464] Mark D Licker. *Dictionary of Earth Science*. 2003.

[465] Choi Chuck Lee. *Environmental engineering dictionary*. Government Institutes, 2005.

[466] Andy D Ward and Stanley W Trimble. *Environmental hydrology*. Crc Press, 2003.

Referenced material only

[30] MK Akhtar et al. *Ganges River Flood Forecasting Using Spatially Distributed Rainfall from Satellite Data and Artificial Neural Networks*. Tech. rep. Water Mill Working Paper Series, 2008.

[76] W WMO. *Manual on flood forecasting and warning*. Tech. rep. 2011.

[86] Bill Scharffenberg et al. *Hydrologic Modeling System HEC-HMS User’s Manual - Version 4.3. U.S.* Tech. rep. 2018.

[116] F Melone et al. *Review and selection of hydrological models-Integration of hydrological models and meteorological inputs*. Tech. rep. 2005.

[124] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. *Learning internal representations by error propagation*. Tech. rep. California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[131] David S Broomhead and David Lowe. *Radial basis functions, multi-variable functional interpolation and adaptive networks*. Tech. rep. Royal Signals and Radar Establishment Malvern (United Kingdom), 1988.

[277] RE Munich. *NatCatSERVICE: natural catastrophe know-how for risk management and research*. Tech. rep. Munich RE, 2018.

[287] Marouane El Mabrouk et al. *An Improved Flood Forecasting and Warning System using Data Mining*. Tech. rep. 11. 2013, p. 2277. URL: www.ijarcsse.com.

[363] WMO Słownik. *International meteorological vocabulary, 1992*. Tech. rep. WMO/OMM/IMGW, 182, Geneva.

[364] WMO. *Guidelines on cross-border exchange of warnings*. Tech. rep. 2003.

[368] Bill Scharffenberg et al. *Hydrologic Engineering Center, 2009. HEC-DSSVue User's Manual - Version 2.0. U.S.* Tech. rep. 2009.

[375] V. T. Chow. *Hydrologic determination of waterway areas for the design of drainage structures in small drainage basins*. Tech. rep. 462. 1962, p. 104.

[376] Arlen D. Feldman. *Hydrologic Modeling System HEC-HMS Technical Reference Manual*. Tech. rep. 2000.

[380] Susan L Neitsch et al. *Soil and water assessment tool theoretical documentation version 2009*. Tech. rep. Texas Water Resources Institute, 2011.

Appendix A

Acronyms

ABM Agent-based Modeling

ABMS Agent-Based Modeling And Simulation

ABS Agent-Based Systems

ABSS Agent-Based Social Simulation

ACL Agent Communication Language

AHD Aswan High Dam

AI Artificial Intelligence

AMCMAS Automatic Methods for Calibration

ANFIS Adaptive Neural Fuzzy Inference Systems

ANN Artificial Neural Network

ANYMAS ANY Time Multi-Agent System

API Application Programming Interface

AR Auto-Regressive

ARIMA Autoregressive Integrated Moving Average

ARMA Mixed-Auto-Regressive with Moving Average

ASTER Advanced Spaceborne Thermal Emission and Reflection Radiometer

BANN Bayesian Artificial Neural Network

BDI Belief Desire Intention

CA Cellular Automata

CART Classification and Regression Trees

CAS Complex Adaptive Systems

CBR Case Based Reasoning

CLA Classifier

CS Computer Science

Data2Lags Data to Lags

DB Database

DEM Digital Elevation Model

DM Data Mining

DMM Data-Driven Model

DPP Data Pre-Processing

EC Evolutionary Computing

EDA Environment Domain Agents

ENSO El Niño Southern Oscillation

EP Evolutionary Programming

ES Evolution Strategies

EWS Early Warning Systems

FA Flood-Awareness

FCST Forecaster

FFBP Feed-Forward-Back-Propagation

FIPA The Foundation for Intelligent Physical Agents

FIS Fuzzy Inference Systems

FL Fuzzy Logic

GAs Genetic Algorithms

GAMA Generic Agent-based Modeling Architecture

GAML Gama Modeling Language

GDEM Global Digital Elevation Model

GEOTIFF Geographic Tag Image File Format

GEP Gene Expression Programming

GIS Geographic Information System

GT Game Theory

GRNN General Regression Neural Network

GP Genetic Programming

HDBM Historic Database Management

HDSL Historic Data Storage Level

HEC-HMS Hydrologic Engineering Center-Hydrologic Modeling System

HORS High-Order Response Surface

HSn Hydrometric Sensor

HSnL Hydrometric Sensor Level

H3 Hydrometric Station No.3

IBM Individual-Based Model

ICTs Information and Communications Technologies

IMAHDA Intelligent Multi-Agent Hybrid Distributed Architecture

IoT Internet of Things

JADE Java Agent Development Environment

JADEX Java Agent Development Environment Extension

KNN k-Nearest Neighbors

KQML Knowledge Query and Manipulation Language

LASSO Least Absolute Shrinkage and Selection Operator

LIBSVM A Support Vector Machine Library

LM Linear Model

MACSim Multi-Agent Control Simulation

MAR Missing Randomly

MAS Multi-Agent System

MCAR Entirely Missing at Random

MFs Membership functions

MICE Multiple Imputation by Chained Equations

MI Multiple imputation

ML Machine Learning

MLP Multilayer Perceptron

MNAR Data Not Randomly Absent

MOO Multi-Objective Optimization

NGOs Non-Governmental Organization

NLTS Nonlinear Time Series

NNET Neural Net

OCHA United Nations Office for the Coordination of Humanitarian Affairs

OMG Object Management Group

OOP Object Oriented Programming

OWL Web Ontology Language

PBPMs Physics-Based Process Models

PIs Prediction Intervals

PSO Particle Swarm Optimization

P2P Point-to-Point

QGIS Q Geographic Information System

R, r Correlation Coefficient

R^2 Coefficient of determination

RBF Radial Basis Function

RDF Resource Description Framework

RF Random Forest

RMSE Root Mean Square Error

RNN Recurrent Neural Network

RNSn Rain Sensor

RPart Recursive Partitioning And Regression Trees

SCLAL System Classifier Level

SCS Soil Conservation Service

SDPPL Sensor Data Preprocessing Level

SEST Statistical Estimators

SFSn Streamflow Sensor

SOM Self-Organizing Maps

SPARQL Protocol and RDF Query Language

SQL Structured Query Language

SSN Semantic Sensor Network

SVM Support Vector Machines

SV Sensor Verification

SVR Support Vector Regression

SWAT Soil and Water Assessment Tool

SWEET Semantic Web for Earth and Environment Technology Ontology

TDNN Time-Delay-Neural-Network

TTP Time to Peak

UH Unit Hydrograph

UI User Interface

UIL User Interface Level

UML Unified Modeling Language

VGE Virtual Geographic Environment

VSAT Very Small Aperture Terminal

WDT Wavelet Denoising Technique

WEKA Waikato Environment for Knowledge Analysis

WLSn Water Level Sensor

WRE Water Resources Engineering

XGBoost Optimized Distributed Gradient Boosting

XMI Metadata Interchange

XML Extensible Markup Language

Appendix B

Used Symbols

τ The Period of Duration of a Phenomenon
$\Omega [I(\tau)]$ Transform Function that Relates a System Input/Output
$I(\tau)$ A Conservative Magnitude Entering a System
$O(\tau)$ A Conservative Magnitude Leaving a System
Ω Transform Function that Denotes System Properties
$RN_{(t)}$ Rainfall at Time t
$WL_{(t)}$ River Stage at Time t
$Q_{(t)}$ River Discharge, Streamflow, Flow at Time t
T_i Inter-Event Time
I_a Initial Abstraction
CN Runoff Curve Number
S Maximum Abstraction
$\Delta\tau$ Estimation of Lag Period
T_c Time of Concentration
T_L Lag Time
S_c Sine of the Channel
L Length of Main Stream
Q_p, Q_{pk} Peak Discharge
vol_{max} Maximum Volume
$Q_{pk}\%$ Percentage Error in Peak Runoff
Pl_i Plan Name(BDI)
Int Plan Intention(BDI)
$Cont$ Plan Perspective(BDI)
Pr Plan Priority(BDI)
B Agent Behavior(BDI)
$q1h$, $q2h$, $q3h$, $q4h$ Forecasted Lead Time Streamflow

$FA1h$, $FA2h$, $FA3h$, $FA4h$ Computed Lead Time Flood-Awareness

Appendix C

Glossary

The important terminologies addressed in this thesis that apply to the modeling problem domain.

Table C.1: A Glossary of Technical Terms and Concepts used in the Flood Ontology Domain.

Item	Word	Synonym	Definition	Source
1	Alert	alarm, warning, flood warning, storm warning, weather warning	Meteorological message issued to provide appropriate warnings of hazardous weather conditions.	WMO [347]
2	Catchment	river basin, watershed, drainage basin	Same as Item No. 48	
3	Climatic Events	weather, atmosphere, atmospheric conditions, environment, setting, climatic anomaly	Are changes occurring among the environmental variables at a precise location and differ regarding the averaged values over two regions or space of the globe.	WMO [347]
4	Climatological Events	Weather Events	Same as Item No. 1	
5	Coastal Flood	storm surge, wind-induced surge, high flood, hurricane surge,	Same as Item No. 8 and 24	
6	Convective Storm	thunderstorms	Thunderstorm produced by a convective cloud.	WMO [347]

Table C.1 – *Resumed from the last sheet*

Item	Word	Synonym	Definition	Source
7	Cross-section	cross sectional area, transverse section, cross sectional shape	In a natural watercourse or man-made channel, the cross-section is the perpendicular section to the direction of the main flow.	Adrien [463]
8	Damage	flood destruction, flood devastation, flood stage	The water level at which the overtopping of a water course induces destructive consequences overland.	Licker [464]
9	DEM	digital elevation model	A GIS map-like product stored in digital file format; usually contain grid points with locations (x), elevations (y), or depth (z) variables used in many catchments' hydrologic studies.	Adrien [463]
10	Discharge	flow, outflow, flow rate, current	Mass of water flowing along the cross-sectional area of a waterway at different intervals.	WMO [347]
11	Drought	aridity, dryness, water shortage	The lack of occurring hydrological imbalance produced by the absence of rainfall due to extended periods of drought.	WMO [347]
12	Environmental Catastrophes	adverse environmental conditions, environmental disaster, ecological disasters, ecological catastrophes, environmental emergencies	Incidents involving the release (or potential release) of hazardous materials into the environment which require immediate corrective actions.	Lee [465]
13	Extratropical Storm	storm surge	The change in the observed stage caused by a turbulent climatic state with strong breezes.	WMO [347]
14	Extreme Temperature	severe, highest, maximum (temperature)	Is the intervals of high and low temperature occurring throughout a particular period.	WMO [347]

Table C.1 – *Resumed from the last sheet*

Item	Word	Synonym	Definition	Source
15	Flash Flood	alluvion, flood, deluge, inundation, torrent, high water, flooding, tsunami, overflow, tidal wave, overflow, surge	A rapidly occurring flood, with seldom any warning, typically resulting from an ice melt, dam break or a period of extreme precipitation covering a small area of land.	WMO [347]
16	Flood	deluge, inundation, outpouring, submerge, flooding, surge	Is the inundation of water occurring overland and in low depressions of the earth by flood waters.	WMO [347]
17	Flood Water	deluge, flood, flooding, inundation, surge	Same as Item No. 8	
18	Flow Depth	depth of discharge, depth of flow, critical depth, flow profile, regime, region, zone	The depth of water in a water course at which the smallest energy of flow is at the bed.	Licker [464]
19	Flow Direction	through-flow, travel direction, direction of flow, flow path	The pathway water would follow given the rainfall-runoff event. This water in excess would flow downstream on the steepest path of the catchment.	Adrien [463]
20	Flow Duration	flow period, flow rate, rate of flow, flow duration curve	The amount of time a flow event of a certain magnitude occurs.	Ward [466]
21	Flow Rate	discharge, streamflow, river flow, rate of flow, flowrate,	The rate of water conveyance through a river or canal ("EPA-40CFR146.3"), measured in units of volume into time.	Adrien [463]
22	Forest Fire	wildfire, wild-land fire, brush fire, smoke pall	A compressed smoke screen occasionally rises from the burning of organic plant material in the wild, or a large city, or an industrial area..	WMO [347]
23	GeoSpatial	geographic, geographical, land, regional, territorial, region, spatial, local	Area or space.	Ward [466]

Table C.1 – *Resumed from the last sheet*

Item	Word	Synonym	Definition	Source
24	GIS	geographic information system	Computer database management system for spatially distributed attributes.	Ward [466]
25	Groundwater	underground water, water table, aquifer, subsoil water, subsurface, subterranean	The source of water extending under the earth's ground.	Ward [466]
26	Hydrograph		The representation of flow rate with respect to time in graphic or table format.	Ward [466]
27	Hydrological	hydrological cycle, water, aquatic, hydrologic, hydrology, hydrographic, hydro, hydropower	It is the consecutive phases of the dynamics of the water cycle.	WMO [347]
28	Hydrological Events	hydrologic, hydrological cycle, hydrographic, hydrographical, water	The commonly known hydrological cycle, as the process the drives the exchange of energy from the great water masses, to the atmosphere, and back to the great water masses.	Licker [464]
29	Hydrometric Sensor	hydrometry, hydrometrical, measurement, measuring, gauging (gaging), hydrometric station	Same as Item No. 44, 45 and 46	
30	Hyetograph		The representation of precipitation intensity with respect to time in graphic format.	Ward [466]
31	Lake	pond, lagoon, lacustrine, reservoir,	A landlocked body of water, with or without an outlet on earth's surface, formed by natural processes of geo-ecological successions, the melting of enormous masses of ice or man-made.	Licker [464]
32	Land Slide	landslips, landfall, mud slides, mud flows, upheavals	The movement of enormous masses of land produced by the earth's gravitational processes.	Licker [464]

Table C.1 – *Resumed from the last sheet*

Item	Word	Synonym	Definition	Source
33	Local Storm		A mesometeorological deluge of localized scale effects.	Licker [464]
34	Meteorological Events	meteorological phenomena, weather events, weather phenomena, climatic events, weather conditions, weather patterns, meteorological, weather, climate, adverse weather conditions	Of or pertaining to meteorology or weather.	Licker [464]
35	Overland Flood	surface runoff	That part of the precipitation which flows on the ground surface.	WMO [347]
36	Rain Gauge	pluviometer, pluviograph	A manual or electronic instrument that is used to collect, measure and record precipitation amounts at single sampling point intervals in time.	Adrien [463]
37	Raster	grid, screen, frame, array	A GIS image-like product that stores watersheds elementary data of rivers (lines), solid objects (points), and water bodies (polygons) in a raster or vectorized format, just as "pixels" in a picture.	Ward [466]
38	Rating Curve	discharge curve, discharge rating curve, stage-discharge relationship	The relationship between the stage and the streamflow in a channel, displayed in a graphic or tabular form.	Adrien [463]

Table C.1 – *Resumed from the last sheet*

Item	Word	Synonym	Definition	Source
39	Reservoir	pool, lake, store, reserve, pond, tank, repository, source	A natural (e.g., lake, pond, basin) or man-made (e.g., dam, tank, reservoir) area containing water. In water resources management, these confinements serve as sources of potable water distribution, irrigation and recreational purposes, electrical power generation, as well as in engineering flood control measures.	Adrien [463]
40	Return Period	period of return, recurrence period, recurrence interval	The expected yearly event of a hydrological episode that occurs with an excess beyond the expectations.	Ward [466]
41	Risk	danger, hazard, threat, peril	In Climatology, is the degree of probability that unfavorable weather will occur over a certain period.	WMO [347]
42	River	stream, creek, water course, water way, tributary, flood, flow	A natural large body of fresh water extending over a large extension on the earth's surface, and usually occurs seasonally with its flow moving downhill approaching another body of water.	Licker [464]
43	River Flood	river inundation, river overflow, outpouring, flooding	Same as Item No. 8	
44	River Stage	river water level, river water hight	Same as Item No. 42	
45	Roughness Coefficient	roughness factor, friction coefficient, channel bed rugosity	A property of watercourses and canals, related to the frictional forces excerpted on water flowing over the bottom and across the features of the hydraulic geometry of the conduits.	Adrien [463]
46	Shapefile		A geographic information system (GIS) file in ArcView.	Adrien [463]

Table C.1 – *Resumed from the last sheet*

Item	Word	Synonym	Definition	Source
47	Slope	gradient, inclination	A zone on the land with a tilted part.	Licker [464]
48	Stage	level, water level, flood stage, river level, river stage	The surface water in a stream elevating above or below a referenced gauge point.	WMO [347]
49	Streamflow Forecast	flood forecasting	Forecasting of water level height, *discharge,* the monitoring of an *inundation event,* from its initial to its final stage, specifically the maximum flow that results from *precipitation,* or *melting of snow.*	WMO [347]
50	Streamflow Sensor	flow meter, current meter, fluid meter, venturi, manometer, magnetic flowmeter	A device that measures the quantity of water passing through a fixed station in rivers streams over determined periods.	Adrien [463]
51	Surface Water	run-off water, run off,	The occurrence of water courses either flowing or contained in depressions overland.	Ward [466]
52	Tropical Storm	tropical cyclone, tropical depression, cyclone, typhoon, hurricane, storm, tempest, tropic thunder, vortex, tornado, cyclonic, windstorm	Is the hazard produced by the gathering of non-frontage extensive and convective water movements and surface wind circulations in the tropics and below tropics zones.	WMO [347]
53	Vector	direction, line, heading	Another commonly used format for storing GIS data.	Ward [466]
54	Water Level Gauge	stage gauge, level radar	A gauge providing the stage of the water surface elevation in a river at specified points.	Adrien [463]
55	Water Resources	aquatic resources, water related, water supplies, hydro resources, water reserves	The naturally existing bodies of water covering an extension of areas upon the face of the earth because of the driving dynamics of the hydrologic cycle.	Lee [465]

Table C.1 – *Resumed from the last sheet*

Item	**Word**	**Synonym**	**Definition**	**Source**
56	Water Volume	body of water, quantity of water, channel storage, water quantity, volume of water	The sum of all the retained waters, including the detection waters and stored in the rivers, excluding those stored in depressions, this sum has as its route the area of outflow defined in a basin.	Adrien [463]
57	Watershed	catchment, drainage basin, hydrological basin, river basin.	Surface area drained by a portion or the totality of one or several given watercourses.	WMO [347]

End of table