UNIVERSITÀ
DEGLI STUDI
DELL'AQUILA

DISIM

Dipartimento
di Ingegneria e Scienze
dell'Informazione
e Matematica

Software Engineering for the Auto of Things

Prof. Davide Di Ruscio

A.Y.: 2024/2025

An AS to manage an Hydropower reservoir

Students:

- Marco Giarrusso, id: 231133, e-mail: marco.giarrusso@student.univaq.it
- Fabio Paolantonio, id: 293028, e-mail: fabio.paolantonio@student.univaq.it

Github project link:

https://github.com/oobooee/Univaq_Se4AS_Project.git

Introduction: How can autonomous systems help managing big infrastructures

In recent years, the growing demand for sustainable and renewable energy has made hydroelectric dams increasingly central to the global energy landscape. These facilities not only provide a clean and renewable source of electricity but also contribute to water resource management, and the prevention of catastrophic events such as floods.

With advancements in automation technologies and artificial intelligence (AI), it has become possible to develop advanced systems for dam management capable of operating autonomously and adaptively.

1. The SE4AS autonomous System project

The project is **inspired by infrastructures consisting of a water reservoir and a hydroelectric power plant located downstream**. **The downstream plant generates energy by rotating a turbine, powered by the force of falling water.**
Typically, this water is reintroduced into a lower reservoir, which can, in turn, serve as a reserve for another power plant located even lower in elevation.
Modern hydroelectric plants not only benefit from discharge water or tributary inflows, but they are also equipped with pumping systems that return water from the lower reservoir back to the upper reservoir.
Below, you can view an image of a real infrastructure located between Veneto and Friuli in Italy, featuring a traditional dammed reservoirs serving hydroelectric plants.

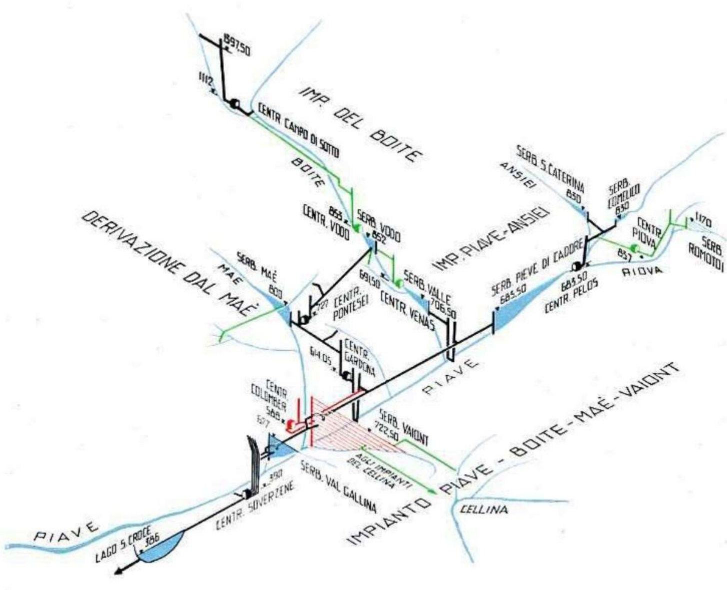


Fig 1: The piave, boite, mae', vajont system

The main objective is to **maximize** and maintain the **water level height** of the reservoir over time while simultaneously **maximizing the water flow directed toward the production plant**. **Maintaining height is critical**, as in real scenarios, the **efficiency of energy** production depends on the potential energy generated by the water's falling height.
Additional objectives include keeping the water level below the **maximum threshold** by using emergency mechanisms and **enabling gradual filling and outflow for production, even when the reservoir has not reached its maximum capacity.**
A lower-priority goal is to **minimize the use of emergency gates**, as the water released through them is considered non-reusable.
To achieve these objectives, a **predictive model** will be developed to estimate water inflows from **rivers** based on **historical data, ensuring plausible, sensor-like data instead of random values**. The same model will be used to forecast future inflows, helping to better address the defined goals.
Below an example of a hydropower that benefits of rivers and **pumps** to refill the reservoir

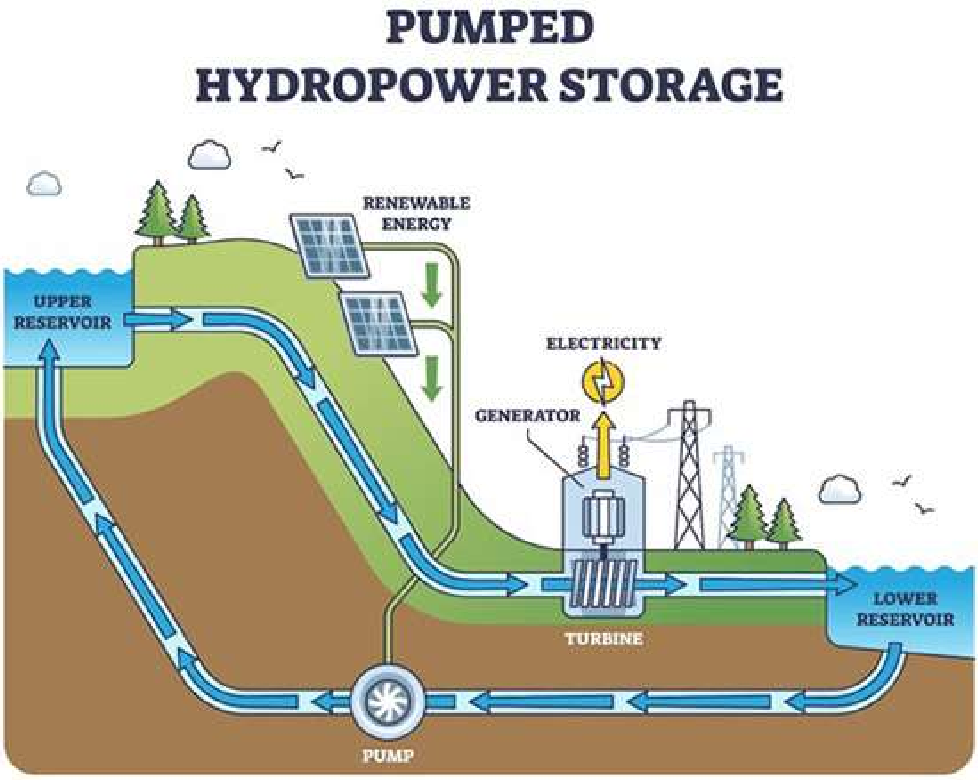


Fig 2: https://amigoenergy.com/blog/pumped-hydro-storage/

2. Goals

Below is a list of all the main objectives that the system must fulfill, ordered by priority, along with their respective justifications.

1. **Maximize the lake water level.**
   Priority: **High**
   Maintain the water level at the highest possible height to ensure a high-water reserve, both to prevent drought periods and to increase production capacity due to the formula where P, the value in watts, is directly proportional to the value h. The higher the height, the greater the power produced.
2. **Maximize the outflow from the lake to the turbine.**
   Priority: **High**.
   Maximize the amount of water flowing out to the turbine as soon as possible and in the largest possible quantity, even if the lake level is not optimal or still filling.
3. **Ensure the safety of the infrastructure.**
   Priority: **Medium-High**.
   Avoid allowing the water level to exceed the maximum height of the lake to prevent overflow or damage to the infrastructure and water waste.
4. **Adapt to environmental variability.**
   Priority: **Medium**
   Incorporate short- and long-term forecasts (daily, weekly, monthly) and to manage unexpected changes in rivers inflow due to rainfall, snowmelt, or seasonal variations.
5. **Minimize the use of emergency (spillway) gates**
   Priority: **Low**
   Assuming that the water released through the emergency gates is considered lost, the system aims to minimize its use.

3. Functional Requirements

The autonomous system must be able to:

- Monitor sensors serving the system and ensures data consistency, in particular:
  - Each inflow sensor (rivers, pumping stations, etc.)
  - The power gate outflow over the turbin
  - The power gate opening percentage
  - The emergency gates outflow
  - The emergency gates opening percentage
- Analyze, calculate and write, if necessary, the system data, specifically:
  - The amount of water present in the reservoir and the high level
  - The total amount of water inflows into the reservoir.

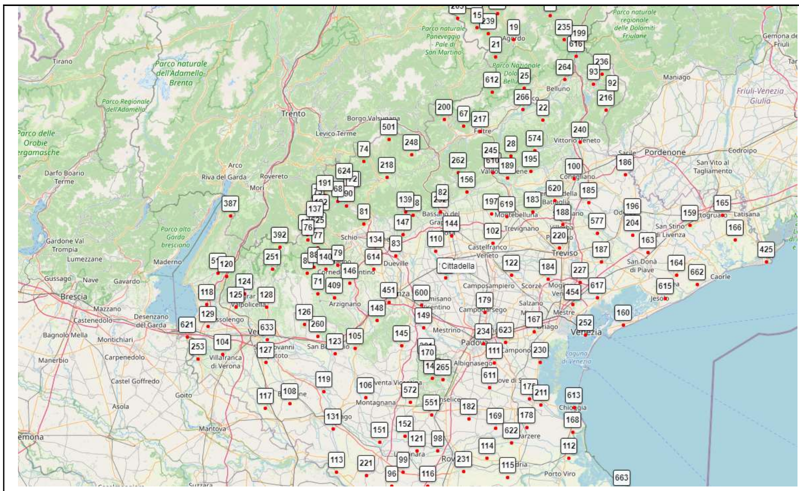
  - The total amount of water outflow from the reservoir.
  - Short-term and long-term forecasts of inflow from each river
- Plans strategies based on data and forecasts to achieve the objectives defined above, specifically:
  - The opening level of the gate leads to the turbine.
  - The opening level of the emergency gates.
- Executes the planned operations and is specifically responsible for implementing:
  - The opening of the gate leading to the turbine.
  - The opening of the emergency gates.

4. Non-Functional Requirements:

- Portability
  - The architecture supports modularity with Dockerized containers for each component, allowing independent scaling and deployment of services facilitating horizontal and vertical scalability.
- Scalability
  - The system allows customization of dam parameters, topics syntax, gate capacity through external configuration .env files. Adding pumps or sources will be very easy. The system automatically acquires data from the new deployed managed resource or manage them.

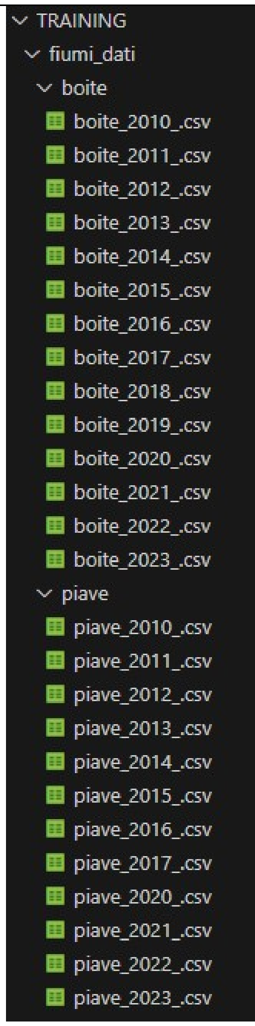
5. Training a predictive model to estimate future water flows from rivers.

To define strategies based on a proactive approach, we developed a predictive system to estimate water flow rates for the Piave and Boite rivers based on historical hourly data, supposing those rivers inflows are part of the managed resources (sensors) of our project demo.
The official data source is available here:
https://www.ambienteveneto.it/datiorari/
The process began with the aggregation and preprocessing of datasets, covering several years of measurements, including recent data and extending back to 2010.
We addressed inconsistencies and missing data by cleaning and normalizing the input files to ensure uniformity. Each dataset was carefully processed to calculate hourly averages, providing a solid foundation for building predictive models.
To capture seasonal patterns and time-based trends, we introduced engineered features such as sinusoidal transformations of days and hours, which allowed the models to better understand recurring patterns in water flow behavior.



For the predictive modeling, we initially employed a Random Forest Regressor, chosen for its robustness and ability to handle complex, nonlinear relationships.

We enhanced its performance by integrating a Grid Search optimization process, which systematically evaluated different hyperparameters to identify the best configuration. This step significantly improved the model's accuracy by fine-tuning parameters like the number of estimators, maximum tree depth, and minimum sample splits.
After training, the models were saved individually for the Piave and Boite rivers using the respective dataset, allowing for distinct predictions tailored to the specific characteristics of each river.



To test the models, we created a system capable of generating random test dates and hours across the year, simulating realistic conditions.
This approach allowed us to validate the models' performance outside of the training set and assess their ability to generalize predictions for unseen data.
Additionally, we implemented a method to estimate the total annual water volume based on the model's hourly predictions. This provided a practical metric to compare the predicted outputs against historical totals, offering insights into the accuracy and reliability of the forecasts.

Throughout the development, we focused on scalability by designing the code to automatically process data from multiple rivers stored in dedicated folders.

Each river's dataset was processed independently, ensuring flexibility for expanding the system to accommodate more datasets in the future. Finally, we emphasized modularity and maintainability, structuring the code to allow for easy modifications and improvements. By combining machine learning techniques with time-series analysis and seasonal feature engineering, the project demonstrates an adaptable framework for forecasting water flow dynamics in different rivers, with potential applications in water resource management and flood prediction systems.

6. The MAPE-K architecture for the system

We approached this project using the MAPE-K methodology that ensures adaptation solutions.

- **Managed resources**
  The system has many managed resources.
  To maintain a consistent condition, real-like data are generated:
  1) Flows:
     - Using the predictive model described below that, given the current time, return the instant inflow for each **river** (with a micro variability randomly generated to make it even more realistic)
     - Fixed or randomly generated if the flow is coming from a **pump**.
  2) Height and volume:
     - Calculated in real-time, as there is a close relationship between the height and volume data and the inflow and outflow rates. For example, it is assumed that the height can be measured using a **webcam** positioned to read images of the dam's height, or readed by a **hydrostatic level measurement** sensor but the data must be consistent. Since it cannot be generated randomly, it is calculated.

  In particular we have to consider those below as managed resources:
  - **Sensors:**
    - reading inbound water flows in m^3/s (rivers, pumps, etc.)
    - reading the level of the lake
    - reading outbound water flows
    - reading the opening state of the gates
  - **Actuators:**
    - changing the opening percentage of each gate
- **Monitor**

In the Monitor phase, we focus on gathering real-time data from all critical component's sensors tracking inflow, gate flows, and pump behavior.
This monitoring system is designed to handle changing conditions, like adding or removing pumps, by keeping track of active components and ensuring that outdated data is removed.
The inflow data reflects the operation of pumps driven by renewable solar energy, so it accounts for natural patterns like higher activity during daylight hours and inactivity at night.

- **Analyze**

The Analyze phase is where the system processes all incoming data.
It validates and combines raw inputs to calculate total inflow and outflow, detect anomalies, and generate insights into how the system is performing.
Predictions play a big role here, as we calculate and tore historical trends and real-time inputs to anticipate how water levels will fluctuate throughout the day or even over the coming weeks.

- **Plan**

The Planner in this project is the core of decision-making within the MAPE-K loop, responsible for generating actionable plans to manage the dam's resources effectively. Its primary focus is to ensure a balance between inflow and outflow while adhering to operational constraints such as critical water levels and energy considerations.

Using data stored in the InfluxDB knowledge base, the Planner processes real-time sensor inputs, including inflow rates from rivers or pumps and gate states. It also integrates predictions about future inflows based on pre-configured models and historical trends. The Planner evaluates these inputs to decide on actions such as opening or closing gates, adjusting the power gate flow, or initiating voluntary pumped storage operations. These decisions aim to prevent overflow, maintain a stable reservoir level, and ensure efficient resource utilization

The Planner adapts its actions to different scenarios, transitioning between adaptive states depending on the water level. For instance, in the FINAL state, it uses predictive inflow data to adjust the power gate percentage dynamically, ensuring long-term system stability.
This process is based on defined thresholds and weighted considerations of real-time and forecasted data, balancing short-term needs with longer-term objectives.

- **Execute**

The Executor takes these plans and puts them into action.
Gates are adjusted, pumps are activated or deactivated, and the system continuously checks whether the planned actions are having the desired effect. If something unexpected happens, it quickly loops back to adjust the plan, ensuring the system stays on track.

- **Knowledge**

The Knowledge component in this project serves as the foundation for informed decision-making within the MAPE-K loop. It is implemented using InfluxDB, which acts as a centralized repository for storing real-time sensor data, historical trends, and predictive data.
This database contains critical information such as inflow and outflow measurements, gate states, and forecasted inflow patterns over various time horizons (daily, weekly, monthly).
By maintaining a structured and easily queryable knowledge base, the system ensures that all components, especially the Planner and Analyzer, can access accurate and up-to-date

information.
This enables precise adjustments to the dam's operations, such as managing gate openings and balancing water levels efficiently. The dynamic nature of the Knowledge component allows for real-time updates, ensuring that the system remains responsive to sudden changes in environmental conditions or resource availability.

7. Architectural pattern

In this project, we have a single autonomic manager that governs the system, although it can be considered part of a more complex system that could also represent the autonomous management of the turbine. Therefore, we can consider the adaptation logic approach as fully centralized.

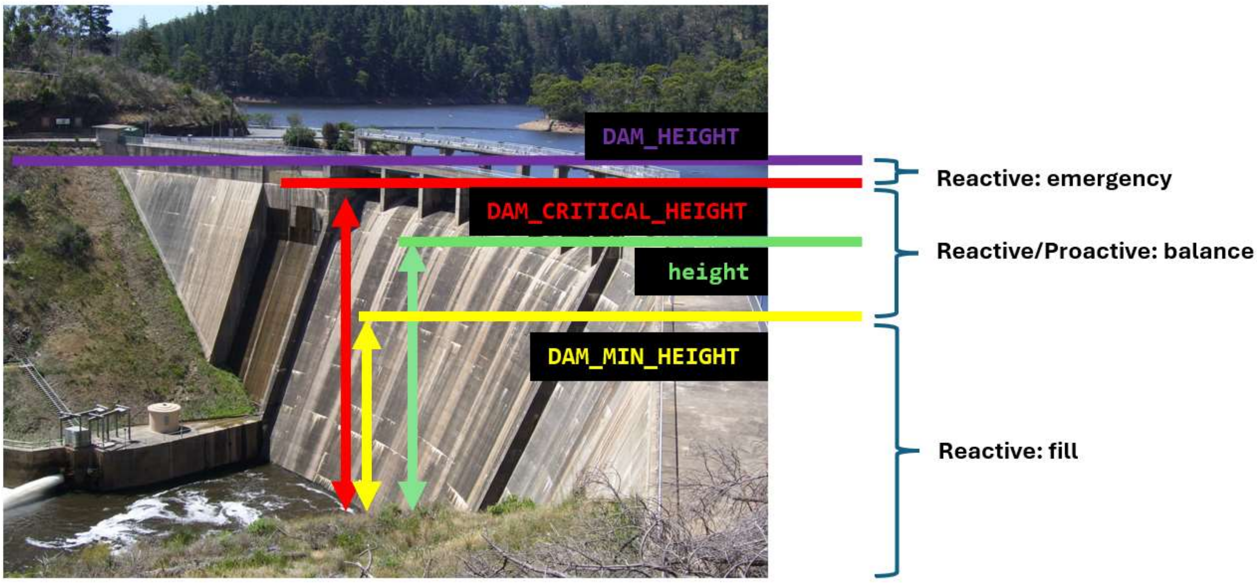
8. Adaptation goals of the autonomic manager

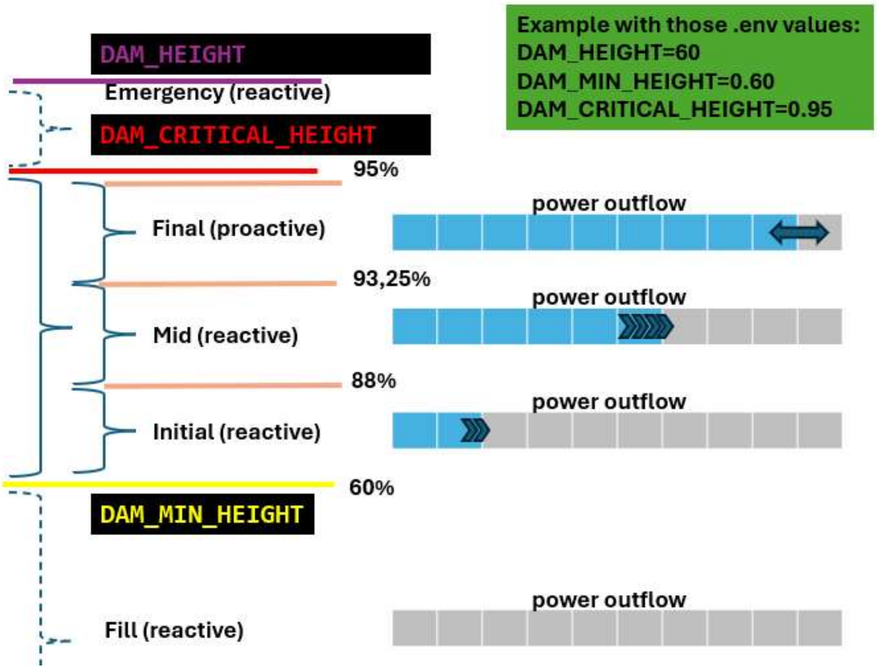
We will discuss now about the approach for determining the **decision functions** that the system must adopt.
The **first** and **second** goals of the project **conflict with each other**, as achieving the first would harm the second. For this reason, a **reactive** approach was used alongside a **proactive** approach to achieve both those goals.
In the following image, some of the variables used to implement the height management system are displayed. The **minimum height** is the threshold below which production cannot be started. **From this point up to the critical height, energy production is initiated at a minimum level, progressively increasing**, with the ultimate goal of reaching the maximum reservoir height.
When the system transitions from the filling phase to the balancing phase, three growth sub-levels are introduced, as shown in the next figure.
The **first sub-level, "initial", starts production at a minimum by slightly opening the gates to the turbine.** This stage maintains the primary goal of increasing the reservoir height while also considering the objective of increasing the flow to the turbine.
In the "mid" phase, production progressively increases while continuing to balance the water flow into the turbine with the inflow into the reservoir.
The ideal final state, "final", marks the reservoir's full capacity. At this stage, predictive data enables the system to adapt the outflow based not only on **short-term forecasts** and **immediate unexpected flows but also** on **long-term weighted predictions**. These predictions account for expected flows over the coming **weeks** and **months**, allowing the system to balance the outflow accordingly.
The outflow increase (the lake level decrease) during expected future high-flow periods, such as floods or snowmelt, and decrease (the lake level increase) during anticipated droughts. The figure below illustrates the weighting function, which considers both recent values and long-term forecasts. The generation of these forecasts is handled by a dedicated thread in the Analyzer module that once a day, generate database data (in the future) using the model predicted for each river

$$Q_{predicted} =$$

$$w_1 \times Q_{current} + w_2 \times Q_{daily}\ w_3 \times Q_{weekly} + w_4 \times Q_{month} + w_5 \times Q_{quarterly} + w_6 \times Q_{semiannual}$$

If the **maximum threshold is exceeded**, possibly due to flood or unexpected events, the system will switch to a purely **reactive** mode to **open the emergency gates**. This situation generally occurs when the inflow rate from tributaries, pumps, and external conditions surpasses the maximum outflow capacity of the turbine. In such cases, the opening of the emergency gates becomes **unavoidable**. When the inflow rate drops back below the physical limit of the gates, the system will return to a balanced state.





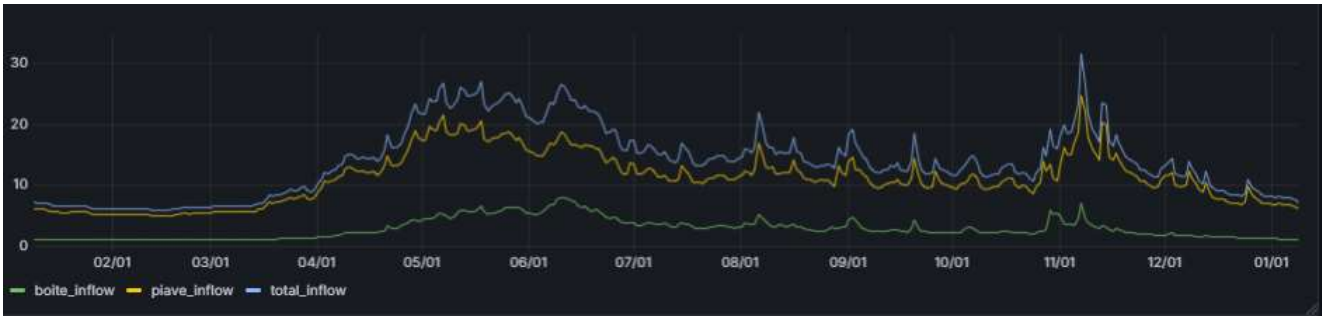
We will show now the evaluation metrics to the adaptation system for each goal

1. **Maximize the lake water level.** Priority: **High**	$$(0.95 \times H_{critical} < h_{current}(t_0) \leq H_{critical})$$ $$h_{current}(t_1) \geq h_{current}(t_0)$$ The system's highest priority is to maximize the lake water level to ensure the maximum power capacity and stocking skills, while maintaining it within safe limits. It ensures that the water level increases over time, aiming to approach the optimal threshold without exceeding safety constraints.
2. **Maximize the outflow from the lake to the turbine.** Priority: **High**	$$\frac{Q_{out}(t_i)}{Q_{\mathrm{POWER\_GATE\_OUTFLOW}}} \times 100 \;\forall\, i$$ $$Q_{out}(t_1) \geq Q_{out}(t_0)$$ The system dynamically maximizes the outflow from the lake to the turbine. The performance is evaluated as the ratio between the current outflow and the maximum gate capacity. Additionally, the system guarantees a progressive increase in outflow when conditions allow.
3. **Ensure the safety of the infrastructure.** Priority: **Medium-High**.	$$h_{current}(t_i) \leq H_{critical} \;\forall\, i$$ The system ensures that the dam's water level remains below the critical threshold opening the emergency gates if some extraordinary event like rainfall or snowmelt occurs or due to prediction failure.
4. **Adapt to environmental variability** Priority: **Medium**	$$Q_{out}(t_i) = \left(f\left(Q_{in}(t_i), Q_{predicted}(t_i)\right) \;\forall\, i\right)$$ The system dynamically adjusts the outflow by incorporating both real-time inflow and predicted inflow: short-term (daily), medium-term (weekly), and long-term (monthly) forecasts are used to anticipate changes due to seasonal variations. This allows lowering the lake level when periods of flooding or large water inflows are expected and maintaining the height at its maximum while reducing power output in anticipation of droughts.
5. **Minimize the use of emergency (spillway) gates** Priority: **Low**	$$h_{current} \leq H_{critical} = 0.95 \times H_{DAM\_HEIGHT}$$ To minimize the use of emergency gates, the system ensures that the water level remains below 95% of the critical height

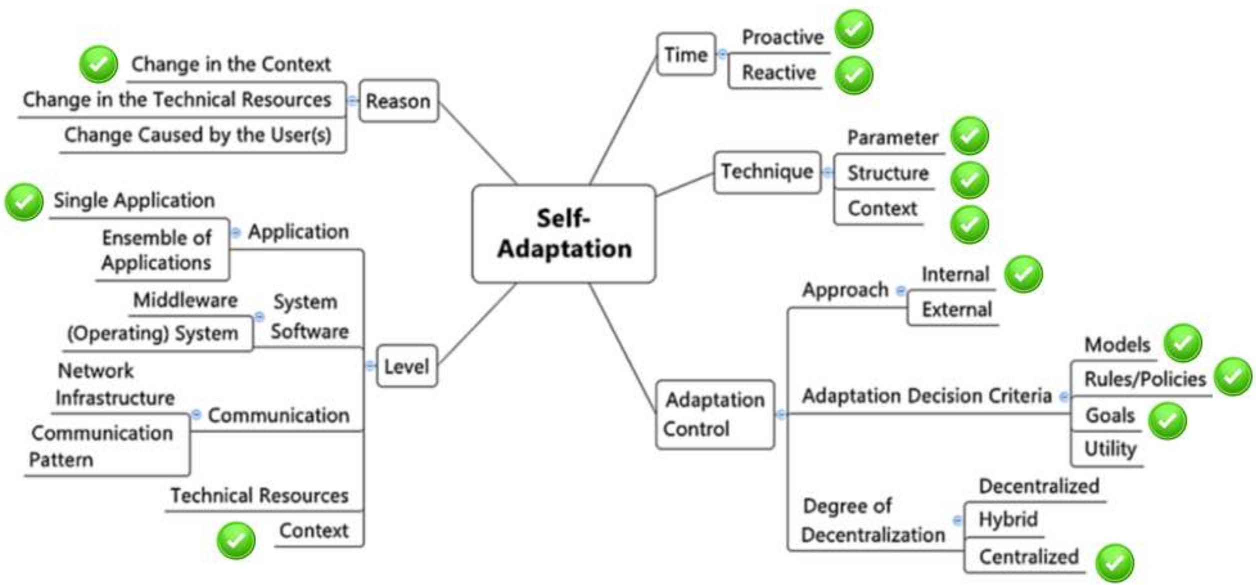
	which is defined as a percentage in the .env file respect to the DAM total height. This threshold provides an additional safety margin to prevent frequent activation of emergency gates.

In this image you can observe the data generated an then used for predict the flows from rivers:

$$Q_{predicted}(Day_i),\ \forall i,\ i = today, i\ \rightarrow 365$$



9. Self-adaptation



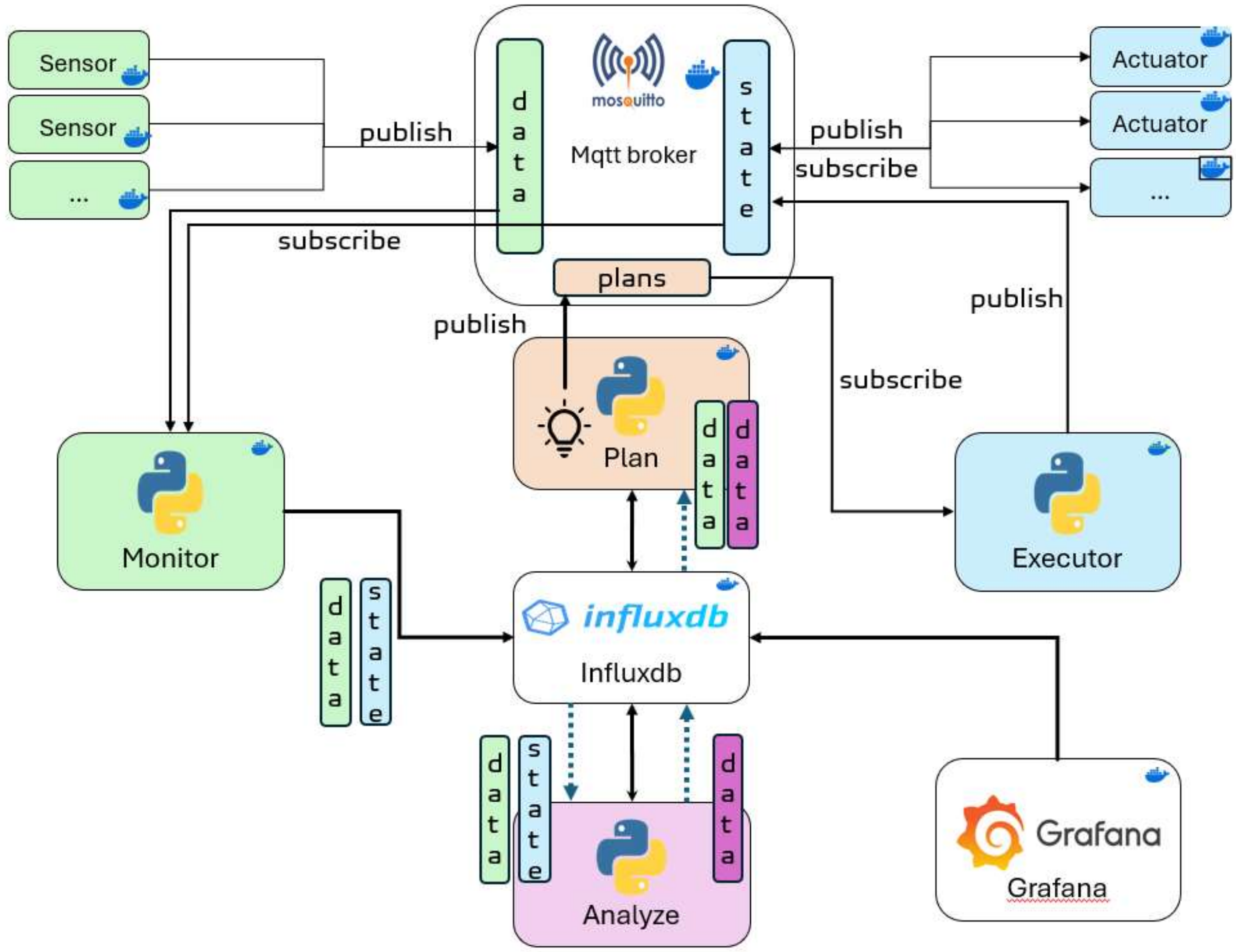
1) **Time**:
   - Proactive: The system uses forecasts (daily, weekly, monthly) to anticipate water flows and adapt its behavior accordingly.

- Reactive: It responds in real-time to critical events, such as sudden changes in water levels, or to planned actions if levels fall below predefined thresholds.

2) **Reason:**
   - Change in the context: The system adapts to environmental variations, such as river flows, climatic changes (e.g., seasonality), and solar pumps, by opening and closing the gates accordingly. If sensors (such as pumps or rivers) are added or removed, the system dynamically adjusts the management of data and resources.
3) **Level:**
   - Application: The system adapts at the application level (dam management), integrating all functionalities into a single application.

4) **Technique:**
   - Parameter: The adaptation is based on modifying parameters such as the target water level, the pumping rate, or the gate opening levels.
   - Structure: The system structure is dynamic, allowing pump sensors to be hot-swapped or removed without interrupting functionality.
   - Context: The system adapts by changing the behavior of components, such as actuators, in real-time.
5) **Adaptation control:**
   - Approach: The system primarily employs an internal approach, where the adaptation logic and managed resources (gates, sensors, turbines) are tightly integrated.
   - Adaptation decision criteria:
     - Models: The system uses predictive models based on historical data and forecasts.
     - Rules/Policies: Actions are driven by rules defined in the code (e.g., fill, balance, final).
     - Goals: The objectives include maintaining the water level, ensuring infrastructure safety, and maximizing energy production.
   - Degree of decentralization: L'intero processo di adattamento è governato da un autonomic manager centralizzato.

10. System architecture



11. Dashboards

The attached dashboard provides an overview of the hydroelectric dam's operational metrics over a selected period (last two days). It focuses on real-time and historical data regarding inflows, outflows, gate operations, and reservoir levels. The "Sensors Flows M^3/s" panel highlights individual contributions to the inflow from three sources: two rivers (Piave and Boite) and a solar-energy-powered pump.

These components contribute to a combined total inflow of 25.39 m^3/s, as shown in the "Global Flows" panel, compared to a total outflow of 17.70 m^3/s.

The "Gates Opening Percentage" panel displays the operational states of the Power Gate (59.04% open) and the two spillway gates (both closed at 0%). A time-series graph tracks the Power Gate opening state fluctuations, reflecting dynamic adjustments in response to real-time conditions.

The "Height and Volume" panel monitors the reservoir's water level, currently at 54.76 meters, approaching its critical threshold. This corresponds to an estimated volume of 912,718 m^3, a key metric for balancing energy production and safety. The "Height and Thresholds" graph visually represents the reservoir's water level against defined thresholds, ensuring that the system operates within safe limits.

This dashboard facilitates real-time monitoring, ensuring the system meets operational goals such as maximizing inflows, efficiently utilizing outflows for energy production, and maintaining reservoir safety. It also demonstrates the integration of environmental variability through data from rivers and renewable energy sources

12. Expansions and enhancement

In the future, the system can be expanded to include a complementary **autonomous manager dedicated to managing the turbine and grid demands**.

This manager would dynamically balance the water outflow to optimize power generation based on real-time energy demands, market prices, and grid conditions. Integrating such a subsystem would enhance the adaptability of the overall system, making it capable of responding to energy consumption peaks while ensuring efficient water usage.Another critical improvement involves implementing a **dynamic learning mechanism** for predictive models. This would allow the system to continuously **update its knowledge of the pump inflow patterns by analyzing historical and real-time data**. Machine learning techniques could be employed to fine-tune the weights used in predictive calculations, adapting to seasonal variations or unexpected changes in environmental conditions.