Flood Forecasting Iberdrola (PoC)  
*Functional Documentation*

Table of Contents

[1. Introduction 3](#_Toc191888368)

[2. General Project Objective 3](#_Toc191888369)

[3. Functional Process Description 4](#_Toc191888370)

[3.1. Data Reception and Validation 4](#_Toc191888371)

[3.2. Calculation of Individual Risks 5](#_Toc191888372)

[3.3. Combination of Factors and Modifiers for Probability Calculation 6](#_Toc191888373)

[3.6. Integration with LLM and Document Database 7](#_Toc191888374)

[3.6. Response 7](#_Toc191888375)

[3.7. Output 8](#_Toc191888376)

[4. Assumptions and Configurable Parameters 8](#_Toc191888377)

[4.1. Thresholds and Weights 8](#_Toc191888378)

[4.2. Operational Assumptions 9](#_Toc191888379)

# 1. Introduction

In the context of risk management and operational continuity of infrastructures, this project aims to assess the likelihood of flooding for Iberdrola’s Transformation Centers (TCs). TCs are essential facilities that, when affected by extreme weather events such as floods, can disrupt power supply and cause operational disturbances.

To carry out this assessment, multiple data sources are integrated:

* **TC Data:** Geographic information (latitude and longitude) and structural characteristics indicating whether the facility is underground and if it is considered flood-prone based on direct observation by the company's technicians.
* **Weather forecast (German Weather Service DWD, using the Icon prediction model):** Precipitation data expressed in millimeters (liters/hour per square meter), provided in hourly intervals. Rainfall forecasts are available for any municipality up to six days in advance.
* **Historical Rainfall Accumulation:** For any future day, predictions are available for the sum of accumulated precipitation in 12-, 24- and 48-hour periods.
* **Cartographic Information from CNIG (National Center for Geographic Information) of flood maps:** Indicates whether a location is in a flood-prone area and, if so, the altitude that the water would reach during a flooding event.
* **Copernicus Flood Alert System:** Indicates the probability of river flooding based on satellite data.
* **AEMET Alert System:** Indicates risk levels (yellow, orange, and red) based on forecasts of heavy rainfall or storms for a specific area.
* **Manual Flood Risk Information for a Transformer Center (TC):** Information provided by the company for those TCs with an extraordinary flood risk due to their specific characteristics.

These data are used to calculate a flood risk index for each TC at any given time, allowing for the prioritization of mitigation measures and early warning of potential critical events.

In the future, integration with the Civil Protection alert system for climate emergency notifications is planned, as well as collaboration with Hydrographic Confederations (CH), which manage the safety of dams. However, as of today, this information is not yet available for consultation through an application.

# 2. General Project Objective

The main objective is to develop an application that, using rainfall forecast data, historical precipitation accumulation, and flood zone data, automatically calculates the flood risk probability for a Transformation Center (TC) over the next four days. This probability will be obtained through the weighted combination of the various factors mentioned above and will allow for:

* Establishing an early warning system.
* Facilitating decision-making in the management and maintenance of infrastructure.
* Serving as a foundation for future expansions that integrate new risk parameters.

For this Proof of Concept, the study area will be limited to the province of Alicante (Spain), which has multiple flood-prone areas and allows for the inclusion of information related to coastal flood hazard maps.

Additionally, we will limit the evaluation of flood risk to 50 TCs.

The project's architecture has been designed with a modular approach using AWS in combination with SAP BPT, to enable the development of a scalable project, considering the potential future incorporation of new risk parameters that will enhance the model’s robustness and expand its reach.

# 3. Functional Process Description

### 3.1. Data Reception and Validation

* **Daily connection to SAP to retrieve information related to Iberdrola’s Transformer Centers (TCs):** Connection to the SAP BPT API to retrieve information related to the TCs.
* **Daily connection, every 12 hours, to Open-Meteo and the DWD ICON model to receive precipitation forecasts for the next four days:** Connection to the Open-Meteo API and data formatting for application consumption. Data is received for the latitude and longitude coordinates of each TC. Scheduled twice a day.
* **Daily connection, every 12 hours, to Open-Meteo and the Copernicus system to receive flood probability data:** Connection to the Open-Meteo API to receive river flood probability data from the Copernicus system for the geographic locations of the TCs. Scheduled twice a day.
* **Daily connection to AEMET, every 4 hours, to receive weather alert data:**

Connection to the AEMET API to receive alerts for weather events. The geographic points of the TCs are linked to the alert-affected zones. Scheduled six times a day.

* **Processing of CNIG maps related to river flood risk areas:** Extraction of map coordinates along with water height data in case of flooding. Executed only once.
* **Processing of CNIG maps related to coastal flood risk areas:** Extraction of map coordinates along with water height data in case of flooding. Executed only once.
* **Processing of Excel sheets with information obtained through direct observation of flood-prone facilities:** If available, information will be stored for those TCs considered especially prone to flooding. Executed only once.
* **Documentation related to preventive and reactive actions to be taken:** Uploading to the document database of the company’s action protocols in case of flooding, failure, or risk situations.

### 3.2. Calculation of Individual Risks

#### 3.2.1. Calculation of precipitation risk for a given day.

* **Rain forecast for each TC:** Precipitation forecasts are obtained via API based on the coordinates of each TC.

Two independent calculations are made for rain:

* **Risk based on rain forecast:** The total forecasted rainfall (in mm) is mapped to a value between 0 and 1 using a predefined threshold per province (see section 4.1 for thresholds).
* **Risk based on rain accumulation:** A similar mapping is applied for accumulated rainfall over the past 48 hours, in 12-, 24-, and 48-hour intervals, using a different threshold (see section 4.1 Thresholds).

Both risks are combined using configurable weights, resulting in a global value associated with precipitation.

#### 3.2.2. Calculation of flood-prone area risk.

Based on the information from CNIG:

• It is determined whether the CT is located in a flood-prone area. If it is not, the distance to the nearest flood-prone area is measured. A value between 0 and 1 is assigned using a predefined threshold (see 4.1 thresholds).

• If the CT is in a flood-prone area, the potential water altitude is used and normalized against a threshold (see 4.1 thresholds) to obtain a risk value between 0 and 1.

#### 3.2.3 Calculation of flood risk for a facility based on direct observation.

If the CT appears in the Excel lists provided by the company as “flood-prone by direct observation,” an additional risk value is assigned.

This value is applied as a multiplier to the final result (see 4.1 thresholds) to reinforce the probability of flooding. It will be applied to facilities where, despite CNIG maps indicating they are not in a flood-prone area, the company has confirmed recurring flooding issues.

### 3.3. Combination of Factors and Modifiers for Probability Calculation

The following key factors are combined:

• **Precipitation Risk:** Resulting from the combination of forecast and accumulation.

• **Flood-Prone Area Risk:** Calculated based on the potential water altitude.

• **Direct Observation of a Flood-Prone Facility:** A risk multiplier applied based on manual information provided by the company or the facility’s characteristics.

The combination is performed using a weighted formula (see **w** values in 4.1 thresholds).

### 3.5. Response

Once the global risk index (a value between 0 and 1) is calculated, the application returns a response in **JSON format**, including:

• **CT data.**

• **Forecasted and accumulated precipitation values.**

• **Flood-prone area information.**

• **Intermediate values and the calculated global risk as a percentage out of 100.**

### 3.6. Integration with LLM and Document Database

Once the flood risk has been calculated, and if it exceeds the minimum threshold (see section 4.1 Thresholds) and reaches a yellow risk level, the following processing flow is triggered to generate recommended actions based on Iberdrola’s response protocols:

An Excel sheet provided by the company is accessed, containing specific preventive and reactive measures for a given Transformer Center (TC).

If no specific actions are available for that TC, a **structured prompt** is sent to the LLM with:

A **structured prompt** is sent to the **LLM**, including:

• **CT location.**

• **Calculated risk index.**

• **Contributing factors** (rain, water accumulation, flood-prone area, etc.).

• **A request to the LLM** to determine the appropriate alert type and any necessary preliminary actions.

**Integration of the LLM with the Document Database**

The **Iberdrola technical and operational documentation database** is queried to retrieve:

• **Preventive protocols** for flooding based on the specific type of installation.

• **Emergency procedures** in case the CT is affected.

The document database returns relevant information to the **LLM**, allowing it to configure the **Action Plan**, which includes:

• **Preventive actions:** Measures to take before flooding occurs.

• **Reactive actions:** Procedures to follow in case the CT is flooded.

### 3.7. Output

A web interface will display a **map of Spain**, allowing users to select:

• **CTs names, characteristics, and locations.**

• **Flood risk** value for a CT on a specific date, represented by a risk color.

• **Intermediate values considered** (flood-prone area, type of installation, etc.).

• **Information retrieved from the LLM**, including work orders and action recommendations.

* **Charts** showing risk evolution over time for a CT

# 4. Assumptions and Configurable Parameters

### 4.1. Thresholds and Weights

**Rain Forecast (current):**

* **Threshold:** 40 mm (values equal to or greater indicate maximum risk).
* **Weight:** 0.5 within the overall precipitation calculation.

**Accumulated Rainfall (48 hours):**

* **Threshold:** 80 mm (values equal to or greater indicate maximum risk).
* **Weight:** 0.5 within the overall precipitation calculation.

**Global Precipitation Combination:**

* **Total weight applied:** 60% for precipitation and 40% for flood-prone area information.

**Flood-Prone Area:**

* **Water altitude threshold:** Maximum threshold: 1 meter, minimum threshold: 0.1 meters.

**Distance to the Nearest Flood-Prone Point:**

* **10 meters or less:** Considered 1 (maximum risk).
* **200 meters or more:** Considered 0 (minimum risk).

A maximum of a 10% multiplier is applied to the overall risk.

**Flood-Prone Facility by Direct Observation:**

* If the facility appears in a company-provided list as flood-prone, a 20% multiplier is applied to the calculated probability value.

### 4.2. Operational Assumptions

**Data Availability and Accuracy:**

* It is assumed that the weather forecast and historical rainfall accumulation data provided by AEMET are accurate and up to date.
* It is assumed that the flood-prone area maps provided by CNIG are accurate and up to date.

**Scalability:**

* The architecture has been designed in a modular way to allow the future incorporation of new risk parameters without requiring a system restructuring.