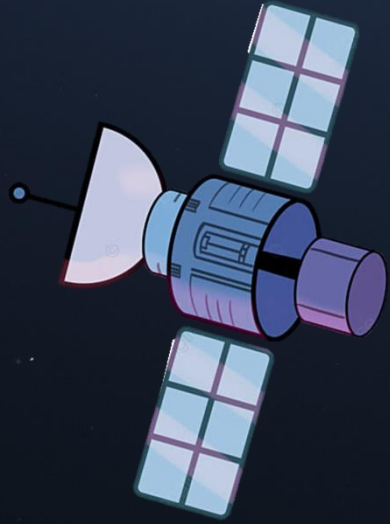




# FLOODCATCH

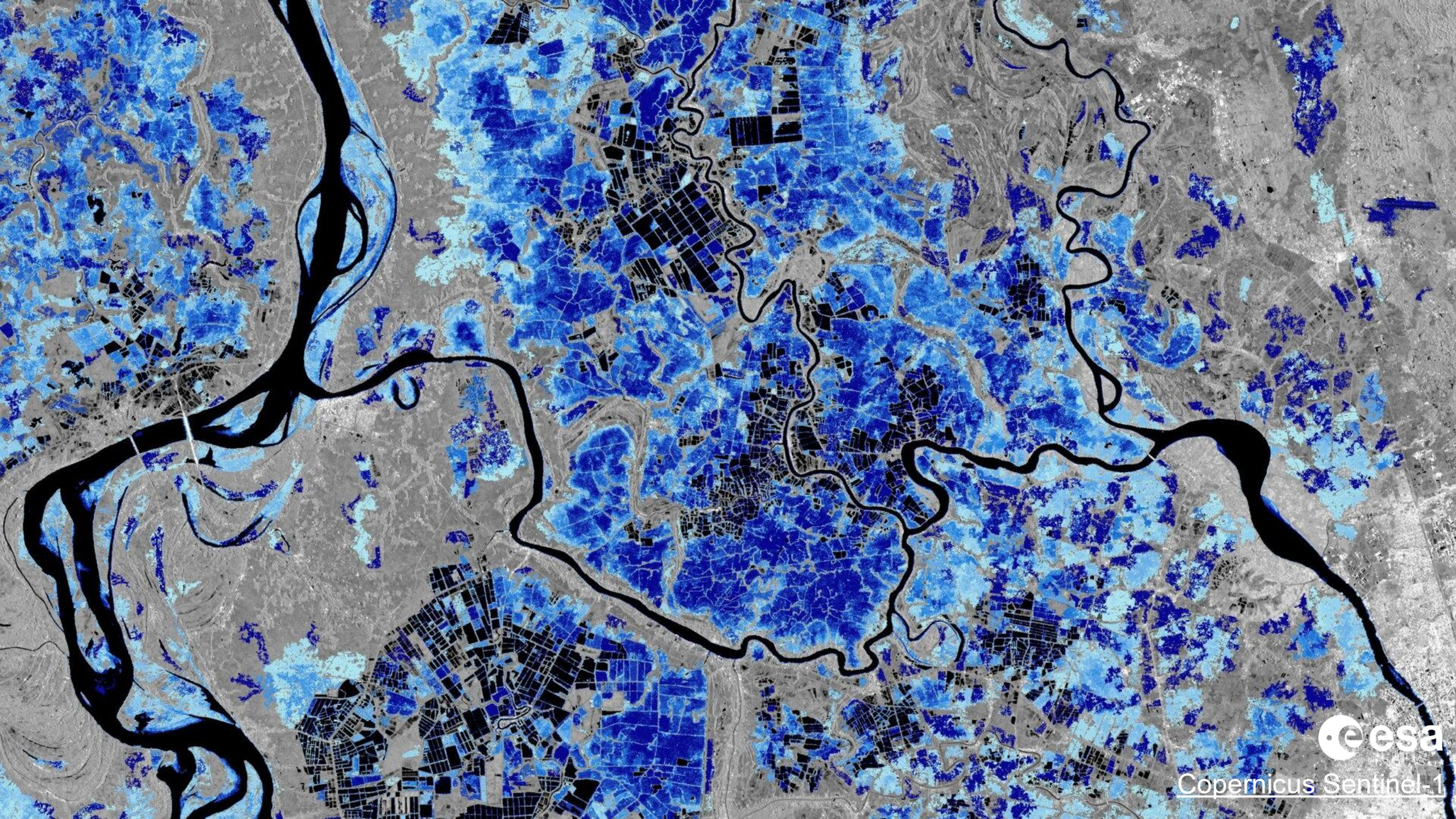
## AI-Driven Continuous Flood Dynamics Monitoring from Satellite Data

Jiayi Cai\*, Marin Heckmann, Julien Elkhiaati, Martin Chauvin, Kevin De Vasconcelos, Sofiane Schaack\*  
Capgemini Invent Lab

Contact: [jiayi.cai@capgemini.com](mailto:jiayi.cai@capgemini.com) ; [sofiane.schaack@capgemini.com](mailto:sofiane.schaack@capgemini.com)  
Eau & IA | Grenoble | March 5, 2026

## AGENDA

- 01 Project overview
- 02 Technical approach
- 03 Conclusion & Next steps





# PROJECT PRESENTATION

## ACTUAL SITUATION AND TARGETED PRODUCT AND SERVICE

The FloodCatch project aims to develop an **AI-based solution** for detecting and forecasting flooded areas across French agricultural plots. While Sentinel-1 provides high-resolution SAR imagery, its **6–12-day revisit cycle leaves critical gaps** in flood monitoring. FloodCatch addresses this limitation by **bridging these gaps through AI-driven imputation and forecasting**

### Daily flood detection

- Use Sentinel-1 data to create masks that distinguish between flooded areas and non-flooded areas
- Correlate flood masks with agricultural maps to detect affected plots

### Short term flood prediction

- Use past flood masks and weather forecasts to anticipate future floods
- Generate flood prediction masks for upcoming days

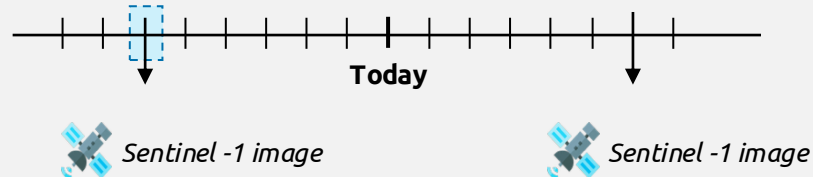
### Dashboarding for monitoring

- Visualize flood-affected areas daily between satellite passes and the +N day period
- Identify impacted agricultural plots through geographical maps and dashboards

#### Actual situation

*The satellite revisit time does not allow daily flood detections*

#### Punctual water detection

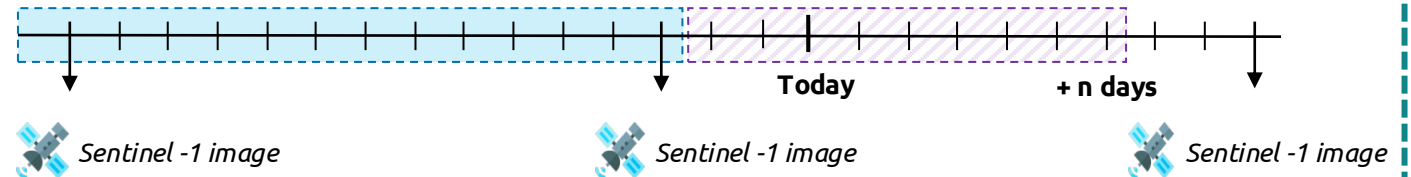


#### FloodCatch project

*The models provide daily and short term prediction of flooded areas*

#### Flood detection between 2 satellites passes

#### FORECASTING MODE



## AGENDA

01 Project presentation

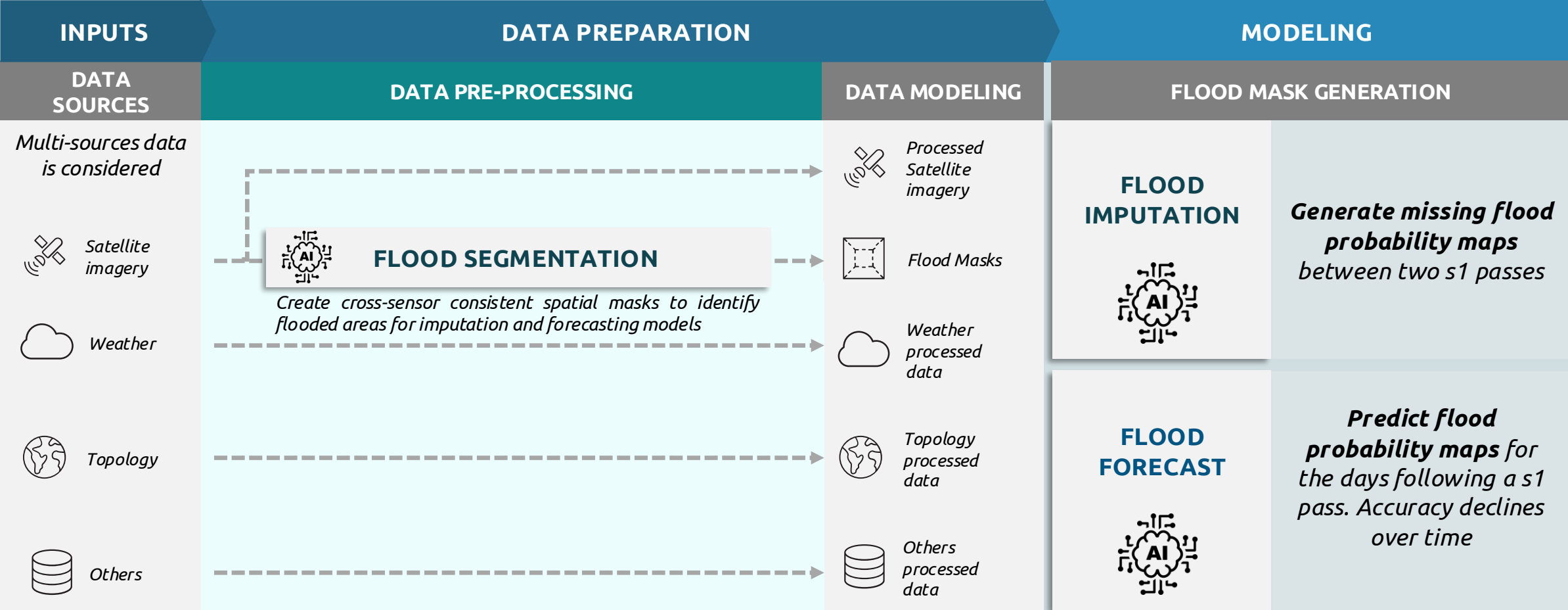
02 Technical approach

03 Conclusion & Next steps



# TECHNICAL APPROACH OVERVIEW

The technical implementation of the project’s features follows a **logical sequence of steps**, and the success of **imputation** and **forecasting** outputs relies on effective data **preprocessing** and a well-functioning **segmentation model**

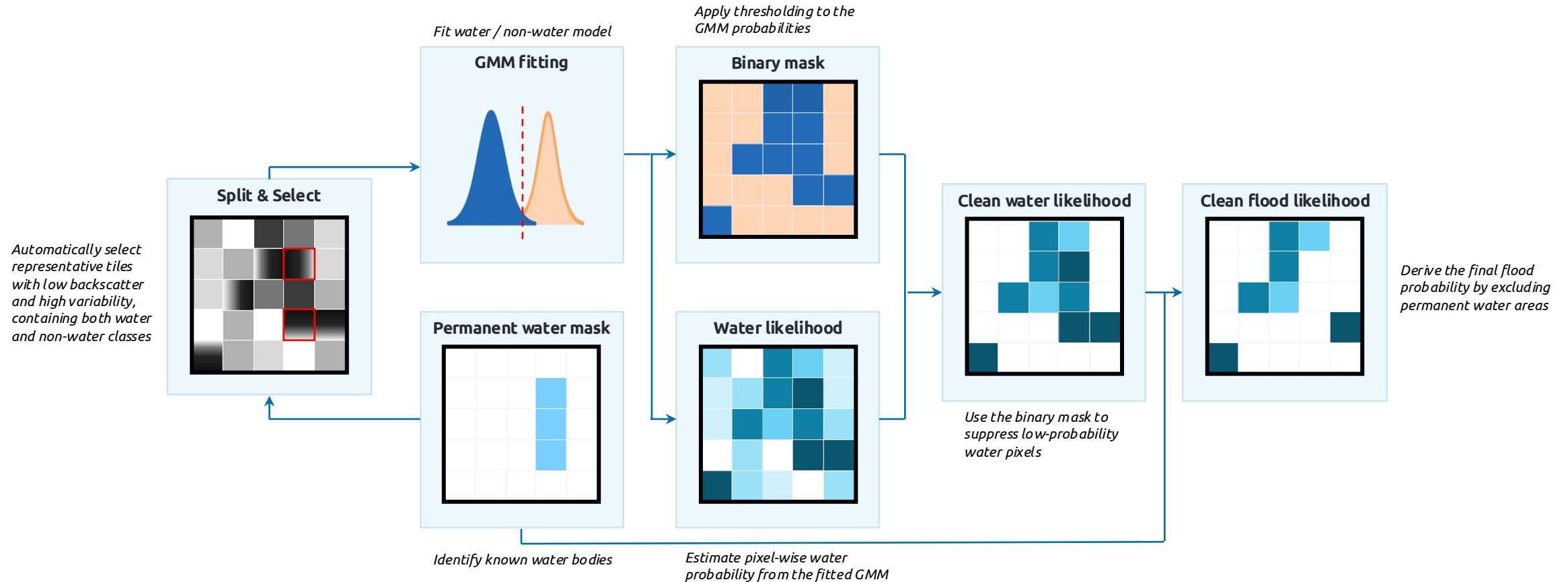




## TECHNICAL APPROACH

# FLOOD SEGMENTATION MODEL: SPLIT-BASED GAUSSIAN MIXTURE MODEL (SBGMM)

We adopt an **unsupervised split-based segmentation strategy** [1] that selects representative low-backscatter tiles for GMM model fitting.



Local **split-based** modeling improves GMM fitting, while **unsupervised learning** ensures cross-sensor robustness to **generate consistent masks for the imputation model**.

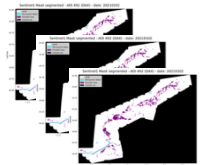
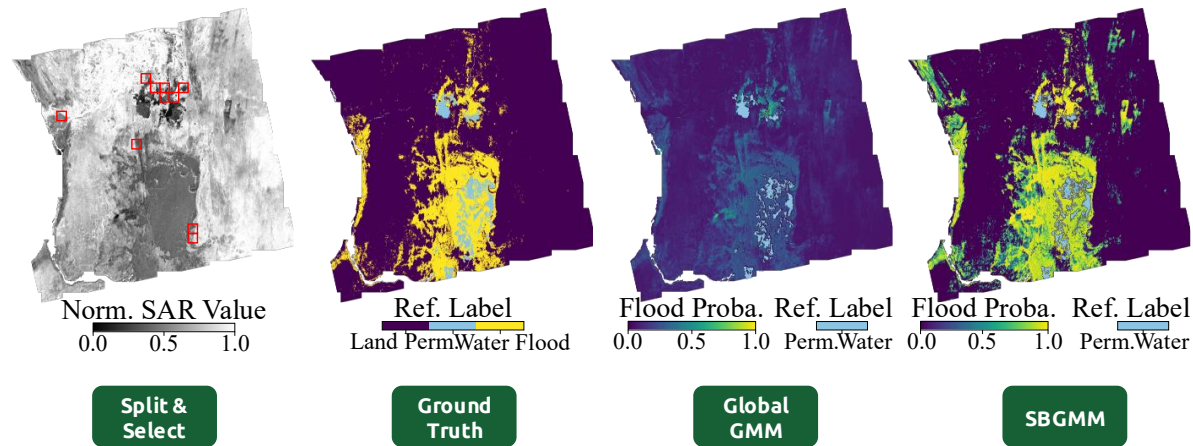
[1] Martinis et al. (2009), NHESS



# TECHNICAL APPROACH FLOOD SEGMENTATION RESULTS

## Benchmark validation – KuroSiwo (Sentinel-1)

- SBGMM shows competitive benchmark performance, with a marked relative improvement over the global GMM approach.



RESULT

≈ 0.56

IoU average  
Comparable to supervised methods  
0.54 – 0.76 [1]

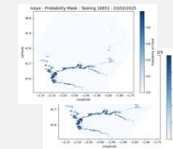
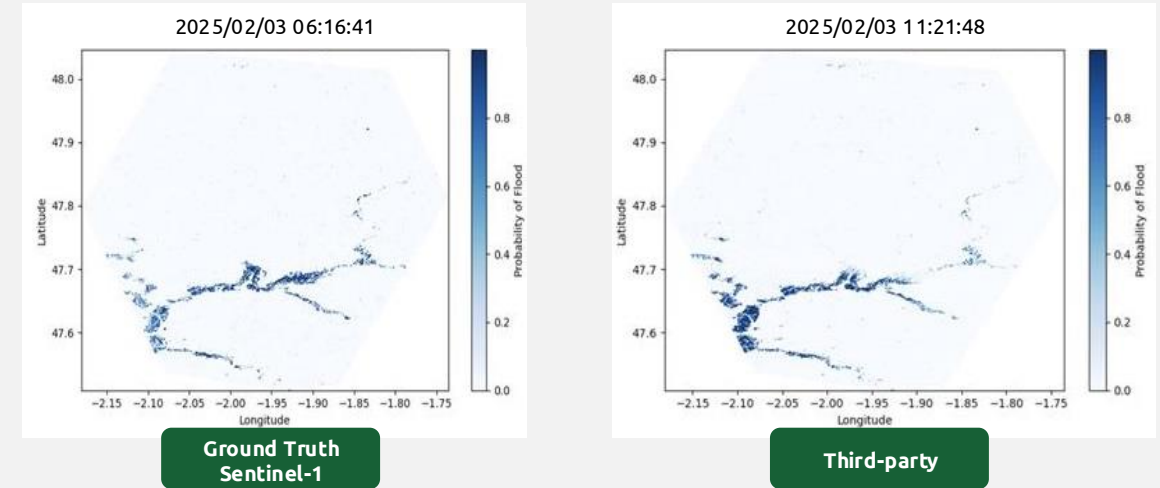
RESULT

≈ 0.65

Soft-Dice average  
+27% relative improvement  
over global GMM

## Cross-sensor validation – Sentinel-1 & Third-party

- Following fine-tuning, the third-party segmentation results showed consistent performance



RESULT

≈ 0.56

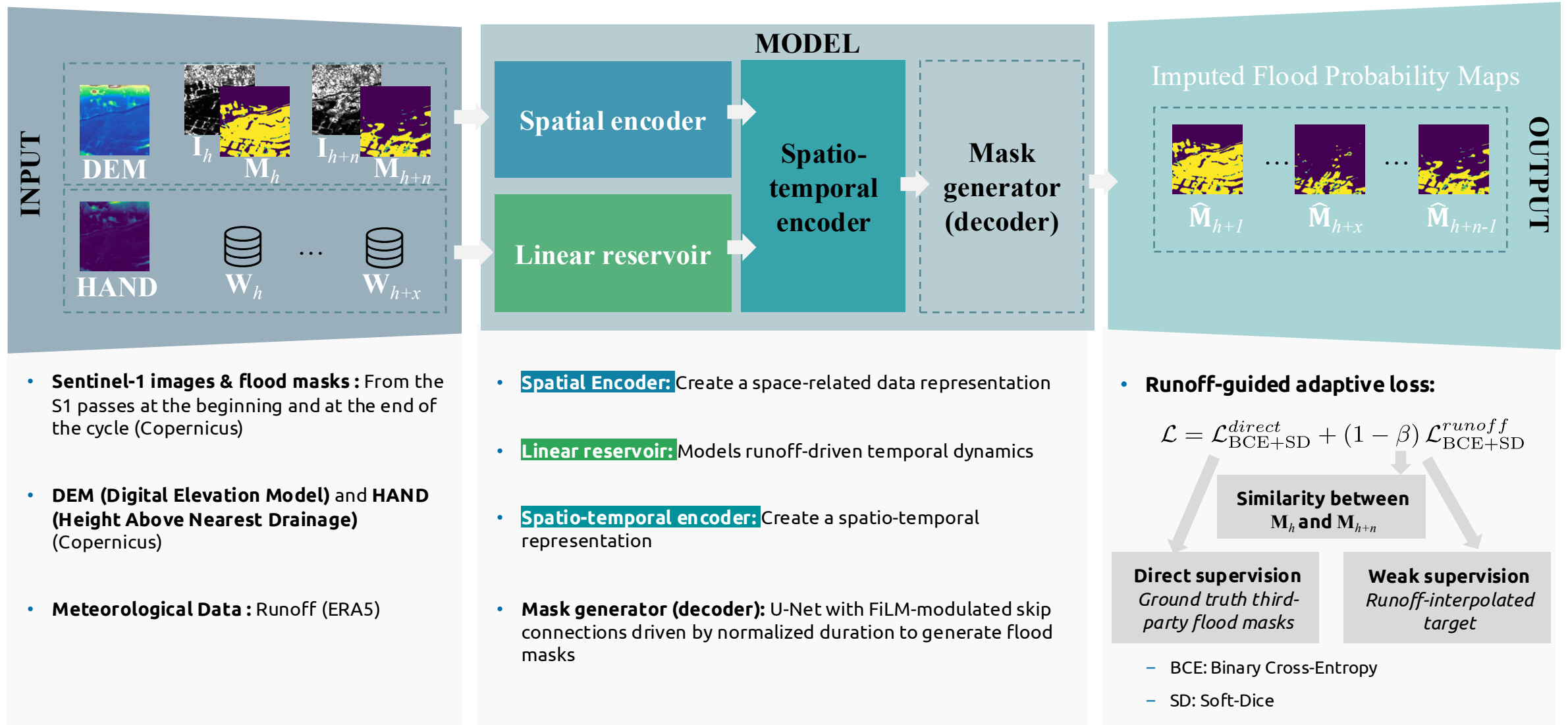
Soft-Dice average  
on 15 pairs\*

\*Image pairing : S1 and third-party  
images acquired on the same day,  
were paired for comparison

**SBGMM** improves flood detection by fitting the GMM on representative local tiles, achieving **competitive performance and cross-sensor robustness**.

[1] Bountos et al. (2024), arXiv

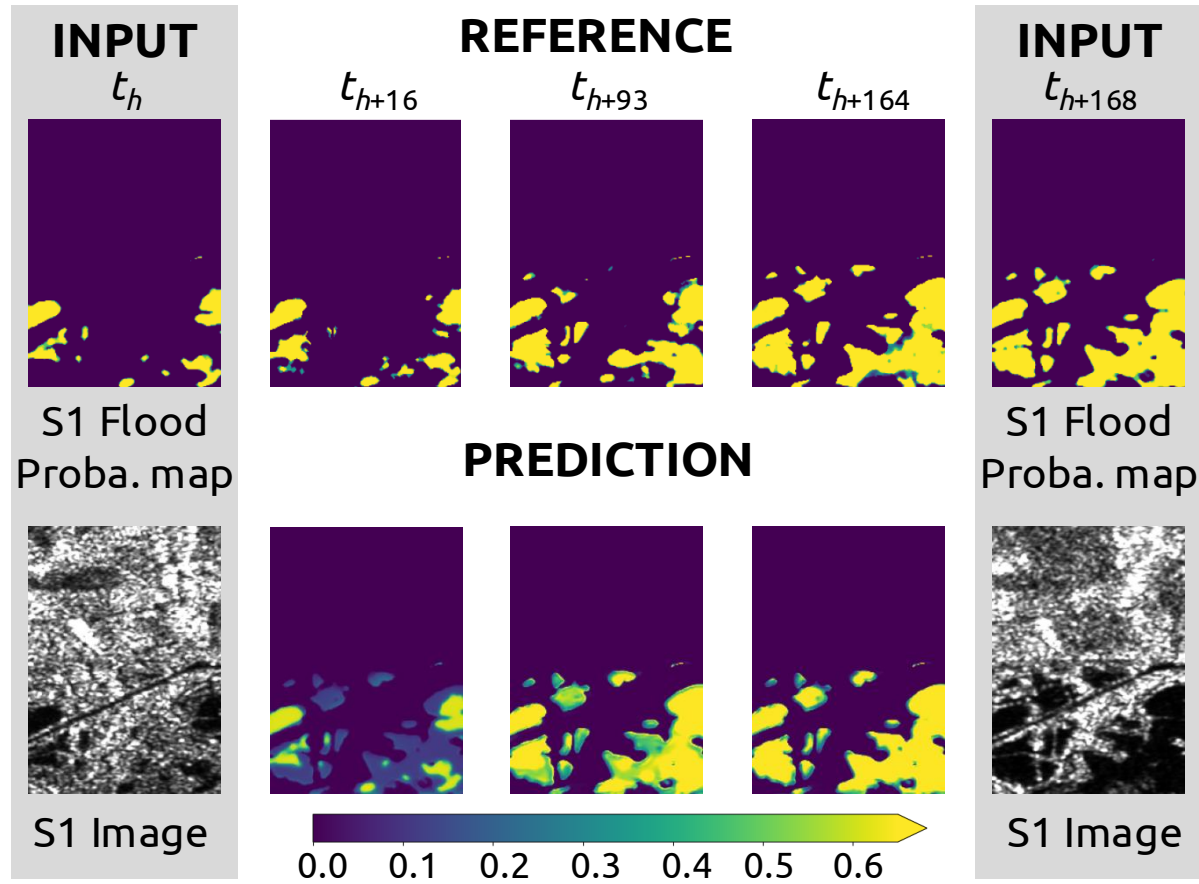
# TECHNICAL APPROACH FLOOD IMPUTATION MODEL





# TECHNICAL APPROACH FLOOD IMPUTATION RESULTS

## RESULTS



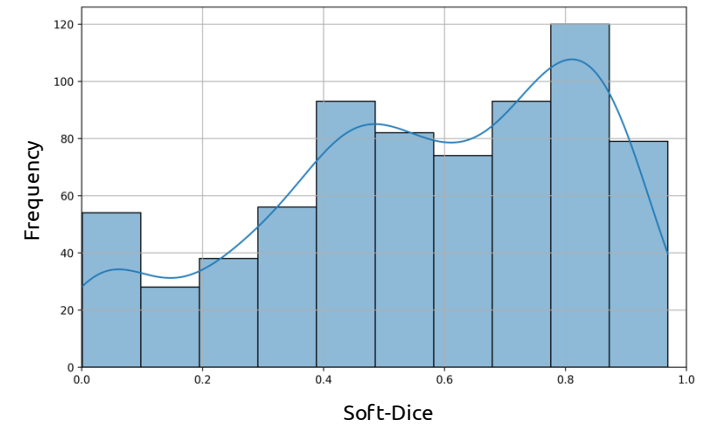
## EVALUATION METRICS IN THE LAST TRAINING

### imputation performance on the test set

$\approx 0.53$

Soft-Dice between flood masks and Ground Truth

*\* Average DICE compared directly with the reference*



The Soft-Dice distribution indicates a trend toward higher values, with many scores close to 0.8, demonstrating **good results**.

The imputation model shows **stable performance**, with Soft-Dice distributions indicating **consistent reconstruction of flood dynamics** across interpolation cycles.

## AGENDA

01 Project presentation

02 Technical approach

03 Conclusion & Next steps



# FLOODCATCH: TOWARD CONTINUOUS DAILY FLOOD MAPPING FROM SPARSE SAR OBSERVATIONS

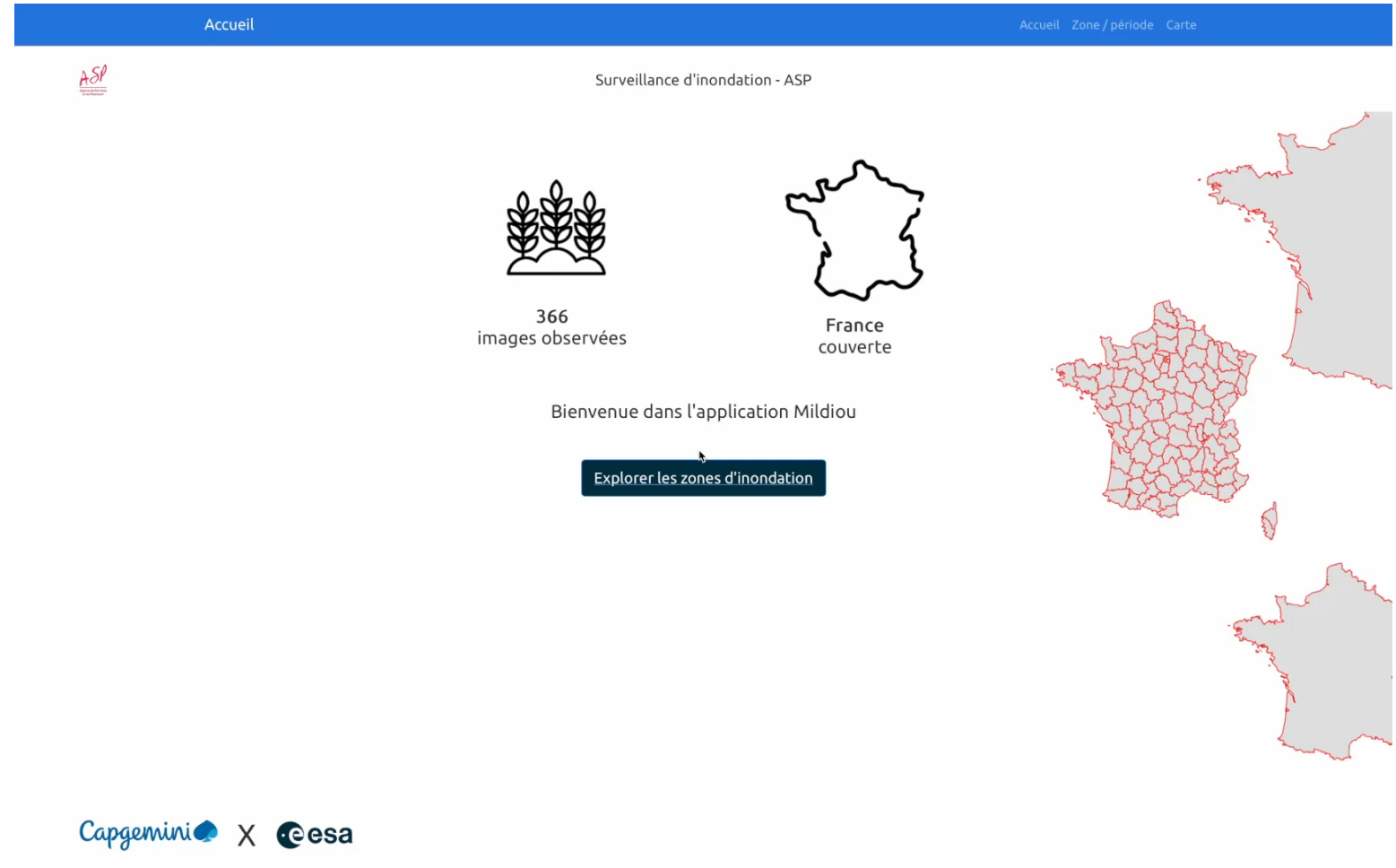
## CONCLUSION & NEXT STEPS

### Unsupervised SAR segmentation via Split-Based Gaussian Mixture Model

- Achieves **competitive performance** against supervised methods on the KuroSiwo benchmark
- Produces consistent flood probability maps across **heterogeneous SAR acquisitions**

### Physics-guided spatio-temporal imputation

- Integrates spatial encoder-decoder with **runoff-driven temporal dynamics and runoff-guided adaptive loss**
- Enforces **hydrologically consistent** temporal evolution
- Reconstructs **intermediate flood probability maps** between sparse S1 acquisitions



Capgemini X esa

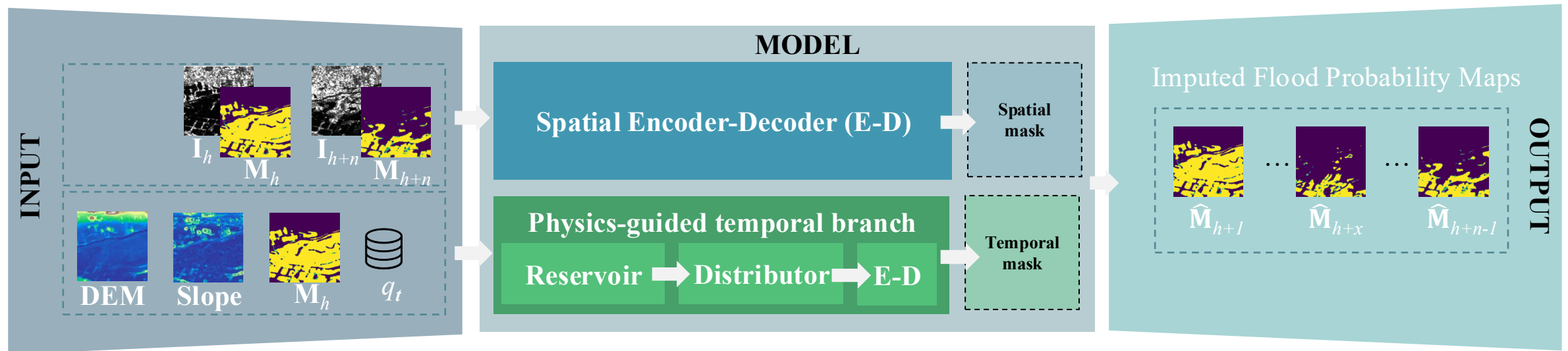
Capturing flood dynamics between SAR observations requires a **careful balance** between **data-driven spatial learning** and **physics-guided temporal modeling**.

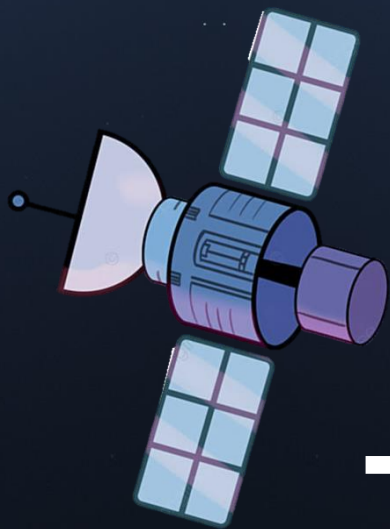
# FLOODCATCH: TOWARD CONTINUOUS DAILY FLOOD MAPPING FROM SPARSE SAR OBSERVATIONS

## CONCLUSION & NEXT STEPS

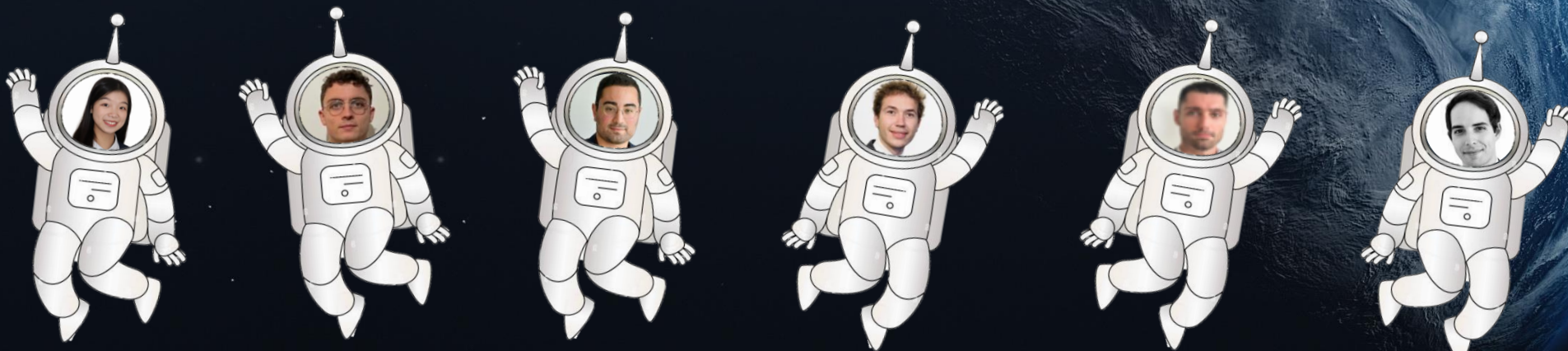
### Next steps

- Enhance **fine-scale spatial delineation** via a topography-aware distributor (DEM, slope and water proximity)
- Extend evaluation across **diverse climatic regions and flood regimes**
- Integrate **rainfall-runoff prediction** for short-term flood forecasting





# THANK YOU!



Jiayi Cai\*, Marin Heckmann, Julien Elkhiti, Martin Chauvin, Kevin De Vasconcelos, Sofiane Schaack\*  
Capgemini Invent Lab

Contact: [jiayi.cai@capgemini.com](mailto:jiayi.cai@capgemini.com) ; [sofiane.schaack@capgemini.com](mailto:sofiane.schaack@capgemini.com)  
Eau & IA | Grenoble | March 5, 2026



# DASHBOARD APPLICATION APP DEMONSTRATION

Accueil

Accueil Zone / période Carte



Surveillance d'inondation - ASP



366  
images observées



France  
couverte

Bienvenue dans l'application Mildiou

Explorer les zones d'inondation



Capgemini X eesa

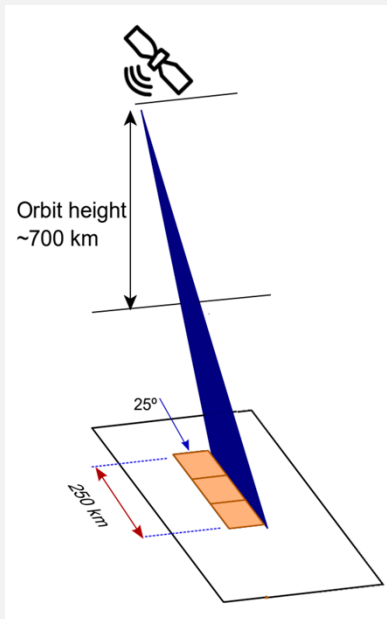


# APPROACH AND MODELISATION

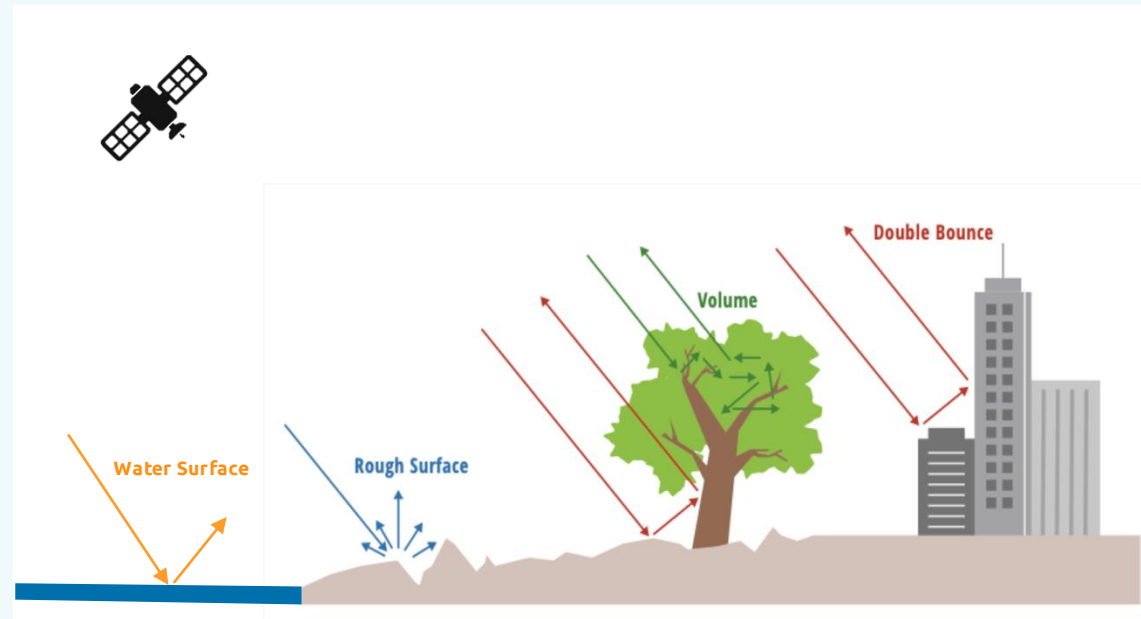
## WHAT IS SENTINEL-1 SAR IMAGERY ?

**S1-SAR Imagery** emits a radar wave over **the area of interest (Aoi)**. The returned **backscatter** varies with **surface roughness and moisture**, ranges from low over water to higher over rough terrain. The **overall result** is expressed in **dB of intensity**, producing a **black-and-white image** representing the radar return.

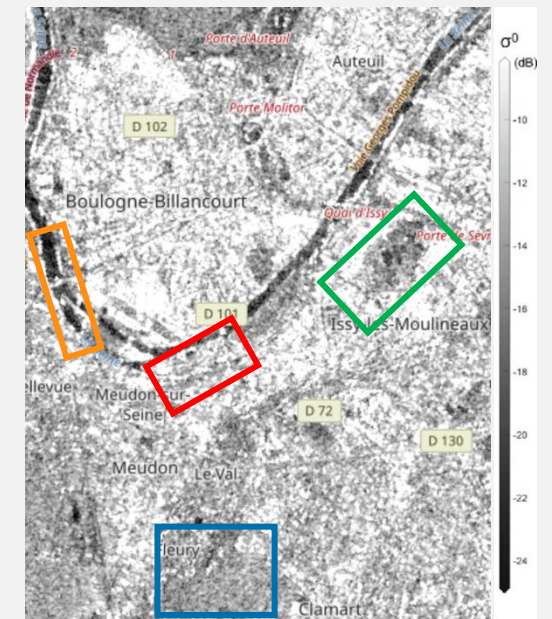
### S1 RADAR EMISSION



### BACKSCATTER RESPONSE DEPENDING ON SURFACE TYPE



### RESULT (in dB)



# EXPLORATORY DATA ANALYSIS - KEY FINDINGS

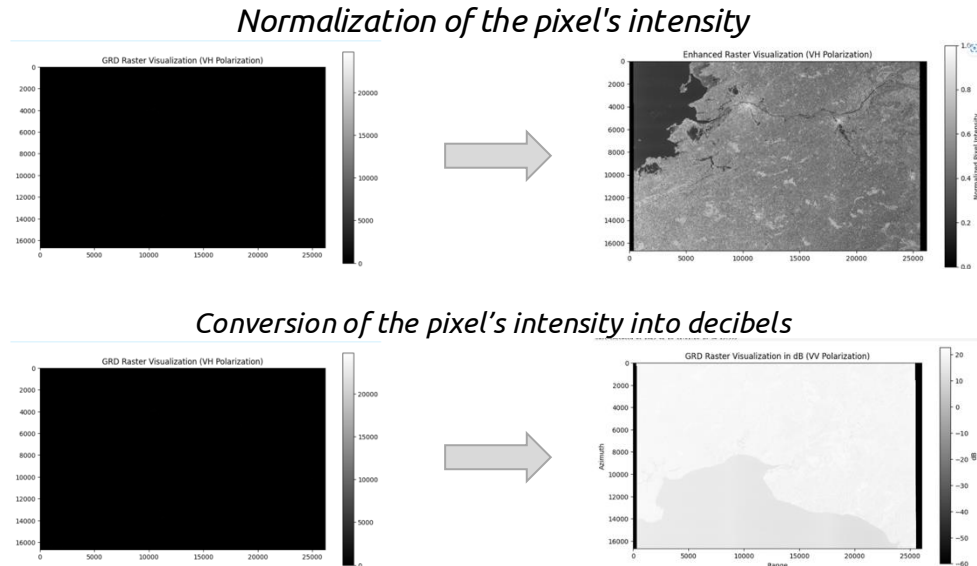
## SENTINEL-1

### Relevant Insights

- The geographical coverage of the image spans a width ranging from 0 to 26,166 pixels and a height from 0 to 16,680 units. Each pixel represents a resolution of 10mx10m. Additionally, these pixels are currently defined in pixel-based positions and must be converted to coordinate references for accurate geospatial analysis
- Each file has two polarizations : VH and VV – 1 image for each

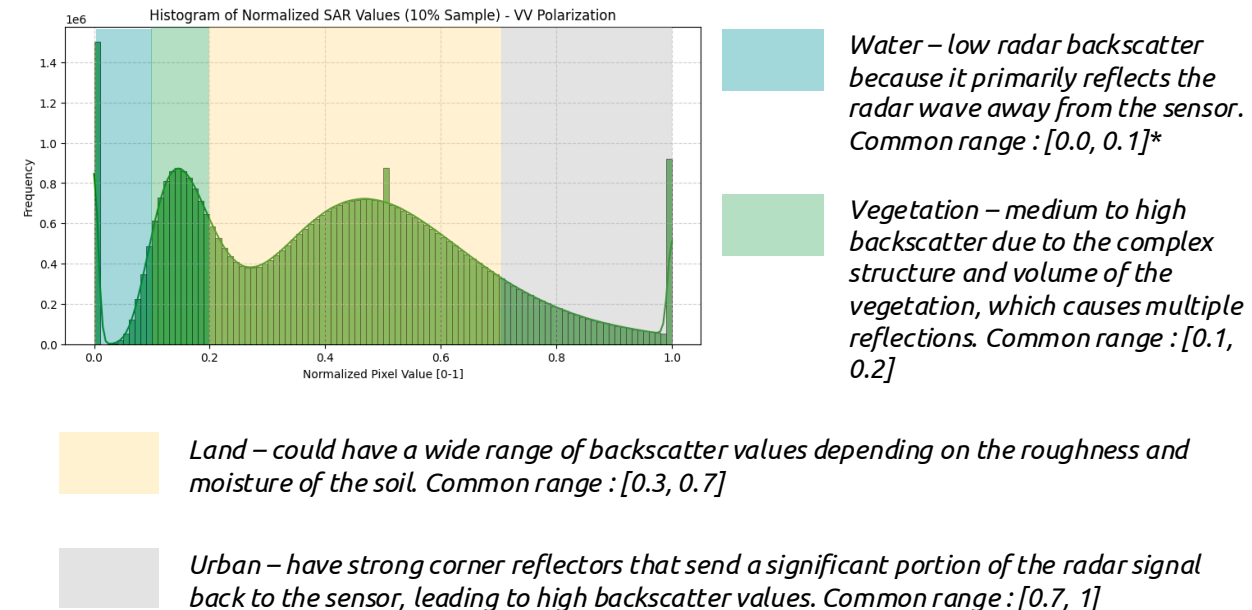
### Pre-processing is needed for Data Analysis

In the raw data, the pixels are not normalized, making interpretation challenging due to the lack of contrast. To effectively analyze the S1 images and extract insights. **We applied basic transformation for SAR images with relevant result**



### Pixels intensity distribution analysis

We built a histogram that shows the distribution of the pixels normalized intensity values. **This histogram aligns with our expectations, indicating that water can be distinctly identified**





# EXPLORATORY DATA ANALYSIS - KEY FINDINGS

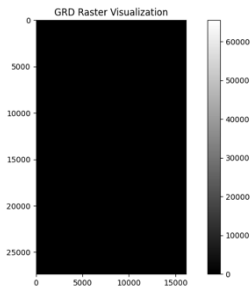
## THIRD-PARTY

### Relevant Insights

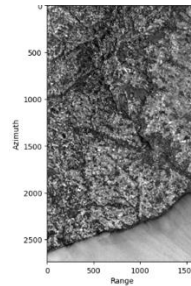
- The geographical coverage of the image spans a width ranging from 0 to 16 130 pixels and a height from 0 to 27 361 units. Each pixel represents a resolution of 3m x 3m. Additionally, these pixels are currently defined in pixel-based positions and must be converted to coordinate references for accurate geospatial analysis
- Each file has only one polarization : VV

### Basic pre-processing is insufficient due to the lack of contrast

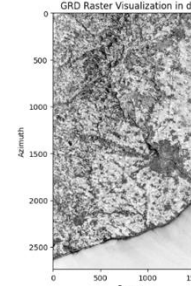
Like Sentinel-1, pixels are not normalized, which makes it challenging to manipulate raw data without pre-processing. However, unlike Sentinel-1, we need to apply additional processing to get sufficient contrast between the water bodies and others. We could also notice that sea water is not dark, due to third-party processing. Converting to decibels extends the spectrum of low backscatter values, enhancing the contrast and making it easier to distinguish water bodies from other features



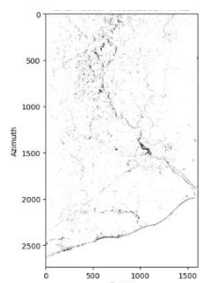
*Normalization of the pixel's intensity and contrast enhancement*



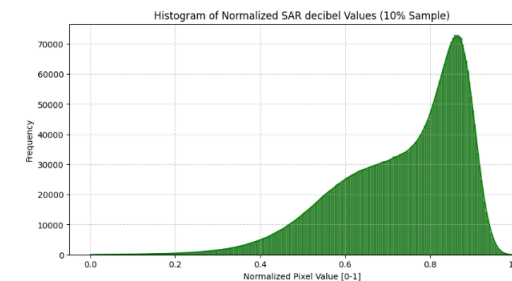
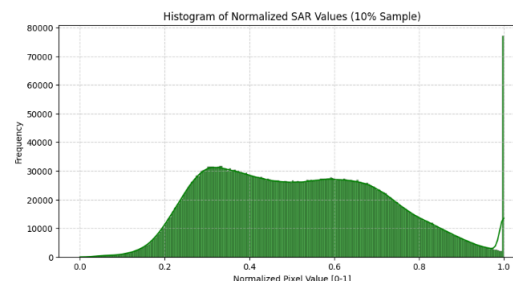
*Conversion of the pixel's intensity into decibels, then normalization of the pixel's intensity, then contrast enhancement*



*Searching for the right threshold between water bodies and others*



*Histograms showing the pixel values distribution corresponding to the images above*

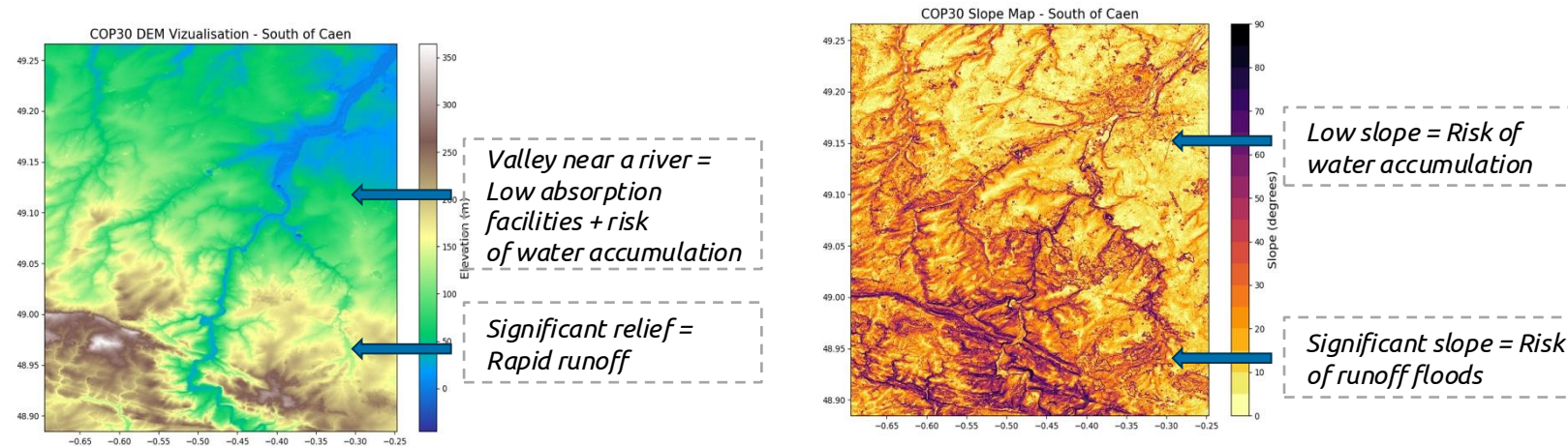


# EXPLORATORY DATA ANALYSIS - KEY FINDINGS

## DIGITAL ELEVATION MODEL



### Exploratory Analysis on the Caen region, which has experienced floods



#### DEM map

- The DEM image highlights the low-lying areas adjacent to the rivers, which represent the floodplains
- The presence of significant relief suggests that heavy rainfall in those areas will quickly flow downhill toward the rivers.

#### Slope map

- The yellow areas indicate low slope, zones of potential water accumulation and runoff-driven flooding
- The steep slopes in the southern part of the Area of Interest indicate areas where rapid runoff will occur during rainfall events.

### Sentinel-1 image during a flood



**Interpretation:** The Sentinel-1 image visually confirms areas of water accumulation that correspond to regions identified as high-risk based on the DEM analysis

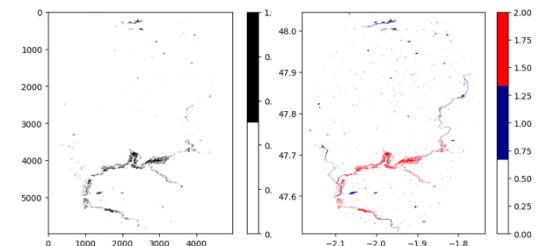
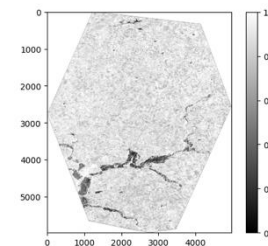
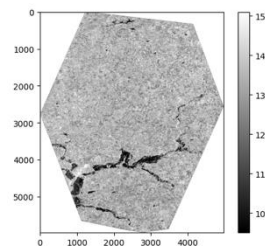
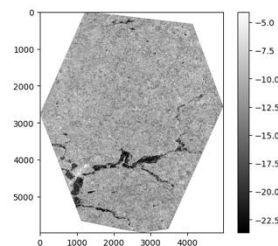
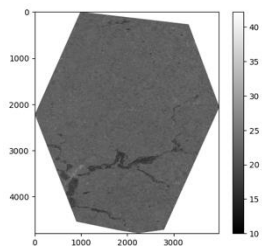
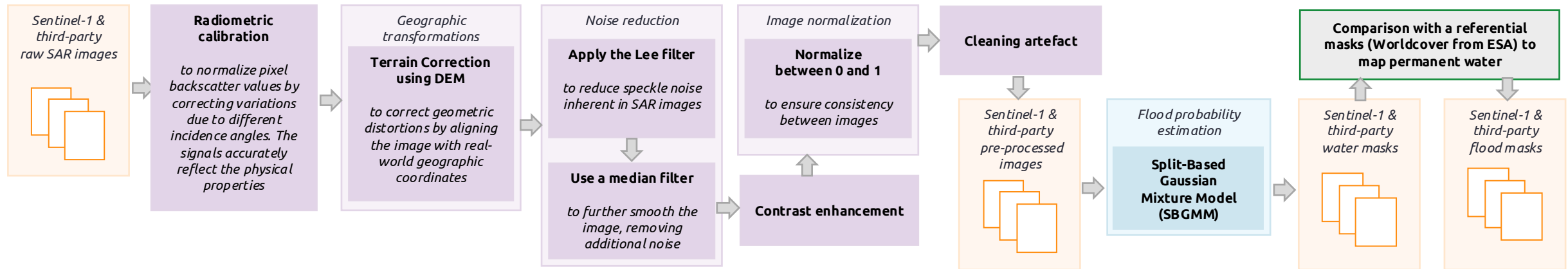
**The combination of significant relief and extensive floodplains along the rivers makes the Caen region susceptible to both runoff floods and river overflow flooding.**



# TECHNICAL APPROACH

## SAR PREPROCESSING & SEGMENTATION PIPELINE

We apply several **preprocessing and segmentation steps** to the raw satellite images. After that, we compare the segmented output with a **permanent water mask** to produce a **flood mask**.



Pre-processing steps

Segmentation steps

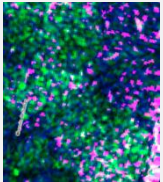


## APPROACH AND MODELISATION

# MODELS OVERVIEW TECHNICAL METHODOLOGY

At the root of the modeling, flood segmentation generates spatial masks that **outline flooded areas**. These masks can then be used as **input** (from S1 images) and **reference sets** (from third-party images) for imputation and forecasting purpose.

### INPUT



SAR images

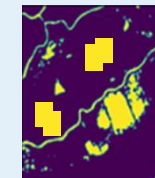
### PHYSICS-BASED SEGMENTATION

The segmentation process combines signal processing and unsupervised learning model to generate spatial masks of water bodies and flooded areas from SAR data. Each pixel is classified as water or non-water with an associated likelihood. These masks are computed for both input data (Sentinel-1) and reference ones (*third-party*) for both imputation and forecasting models training



*An alternative approach would be to use deep learning models to generate flood masks from SAR image data. However, open-source models using this methodology are not trained with third-party data and could be costly*

### WATER MASK



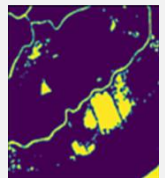
All water bodies

### PERMANENT WATER MASK



Permanent water reference

### FLOOD MASK



Flood detection



Select same area of interest during the driest months or using permanent water map reference



## TECHNICAL APPROACH

### RUNOFF-GUIDED ADAPTIVE LOSS

- **Runoff-guided adaptive loss:**

$$\mathcal{L} = \mathcal{L}_{\text{BCE+SD}}^{\text{direct}} + (1 - \beta) \mathcal{L}_{\text{BCE+SD}}^{\text{runoff}}$$

Similarity between  
 $\mathbf{M}_h$  and  $\mathbf{M}_{h+n}$

**Direct supervision**  
*Ground truth third-party flood masks*

**Weak supervision**  
*Runoff-interpolated target*

- BCE: Binary Cross-Entropy
- SD: Soft-Dice

Training employs a Binary Cross-Entropy (BCE) + Soft Dice (SD) [3] loss with a runoff-guided temporal regularization term. We define a cycle-dependent similarity coefficient:

$$\beta = \left\langle \frac{\mathbf{M}_h}{\|\mathbf{M}_h\|_2}, \frac{\mathbf{M}_{h+n}}{\|\mathbf{M}_{h+n}\|_2} \right\rangle. \quad (1)$$

Weak temporal supervision is introduced through a runoff-interpolated target,

$$\mathbf{M}_{h+x}^{\text{runoff}} = (1 - r) \mathbf{M}_h + r \mathbf{M}_{h+n}, \quad r = \frac{\sum_{t=h}^{h+x} q_t}{\sum_{t=h}^{h+n} q_t}. \quad (2)$$

The final loss is given by

$$\mathcal{L} = \mathcal{L}_{\text{BCE+SD}}(\hat{\mathbf{M}}_{h+x}, \mathbf{M}_{h+x}) + (1 - \beta) \mathcal{L}_{\text{BCE+SD}}(\hat{\mathbf{M}}_{h+x}, \mathbf{M}_{h+x}^{\text{runoff}}). \quad (3)$$

This formulation emphasizes the runoff-guided constraint for large flood changes, while relying primarily on direct supervision when the endpoint maps are similar.