Satellite Images for Physical Risk Modelling and Economic Activity

Columbia Engineering
The Fu Foundation School of Engineering and Applied Science

Konstantina Mavrogianni – Seoyeon Tai – Vikram Mathew

Advisors: Prof. Ali Hirsa, Miao Wang, Dr. Oleg Reichmann, Dr. Giuseppe Bonavolontà

European Central Bank

European Investment Bank

Abstract

This project is newly developed with aim to use alternative data, in particular, satellite images to obtain a better understanding of economic phenomena. The case study of the project is the flood analysis and damage estimation of the recent flooding event of 29th October 2024, in Valencia, Spain. The data used are obtained from the Sentinel-2 satellite mission and the images are freely available from Copernicus. The project consists of four important parts: (a) the exploration of the Copernicus dataset and selection of the proper data for the flood case study, (b) running the super-resolution algorithm to achieve the necessary higher resolution, (c) the cloud removal algorithm to achieve clear dataset without noisy data and (d) the quantification of the flood, damage and cost estimation.

I. Introduction

Climate change is increasing the frequency and severity of floods, creating urgent challenges for physical risk modeling. For the European Investment Bank (EIB) and European Central Bank (ECB), understanding these risks requires robust dataset that can both capture the physical extent of disasters and link them to firm-level outcomes.

II. Dataset

Satellite images have several advantages compared to traditional data sources since (a) they are typically available in a timely manner and thus allow for assessments almost without a reporting lag and (b) they are available with a high degree of spatial granularity which allows for a precise estimation of relevant quantities.

Sentinel-2 mission consists of two identical satellites Sentinel-A (2A) and Sentinel-B (2B), which fly on the same orbit with 180 degrees difference to provide more frequent data. The two satellites cover the same coordinates and have identical sensors each one acquiring 13 spectral bands with spatial resolution 10m, 20m, and 60m, depending on the band. The data frequency is approximately every 5 days since each satellite covers the same coordinates every 10 days, combined they decrease the revisit time to 5 days providing more frequent data set.

We focus on analyzing pre and post flood data. Both satellites have the same tiling system, MGRS (Military Grid Reference System), and each tile covers approximately 110x110 km². We selected the pre-flood data of 1st October 2024 from T30SYJ tile because it offers clearer, and without clouds image of the area before the flood event. For the post-flood data, we selected the same tile and the earliest available date after the flood event on 31st October 2024 [2].

Pre-Flood:
- 01 October 2024
- S2B_MSIL2A_20241001T104759_N0511_R051_T30SYJ_20241001T133447.SAFE
- TILE ID: T30SYJ

Post-Flood:
- 31 October 2024
- S2B_MSIL2A_20241031T105109_N0511_R051_T30SYJ_20241031T133016.SAFE
- TILE ID: T30SYJ

III. Super Resolution Algorithm

Our implementation followed Lanaras et Al. [5], where Sentinel-2 Level-2A images were processed with the pre-trained DSen2 models to upscale the 20m (e.g., B5, B6, B12) and 60m bands (B1, B9) to 10m resolution, while retaining the native 10m bands (B2, B3). The outputs were merged into a unified 10m GeoTIFF stack which was then visualized in QGIS.

Figure 1: Original Sentinel-2 image of Ford Plant at native resolution, prior to the application of the super-resolution neural network.

Figure 2: Super-resolved Sentinel-2 image of Ford Plant obtained after applying the deep learning model, showing enhanced spatial detail.

The DSen2 convolutional neural network was originally trained on synthetically degraded Sentinel-2 data, learning to reconstruct high-resolution detail from lower-resolution inputs. It generates high-resolution outputs with higher spatial detail than the original images. We validated these results by visualizing the super-resolved bands in Python, confirming more accurate water bodies and infrastructure, and then applied the workflow to both pre- and post-flood scenes over Valencia, Spain.

IV. Cloud Removal

Cloud removal is incorporated within two models: the cloud masking model S2Cloudless [7] and the cloud removal neural network DSen2-CR [6]. These two models work in a complementary way: S2Cloudless detects cloud-covered areas and generates an external mask, while DSen2-CR applies this mask to inpaint and reconstruct the underlying terrain that clouds would otherwise hide. This approach ensures that we do not accidentally count clouds in the water quantification.

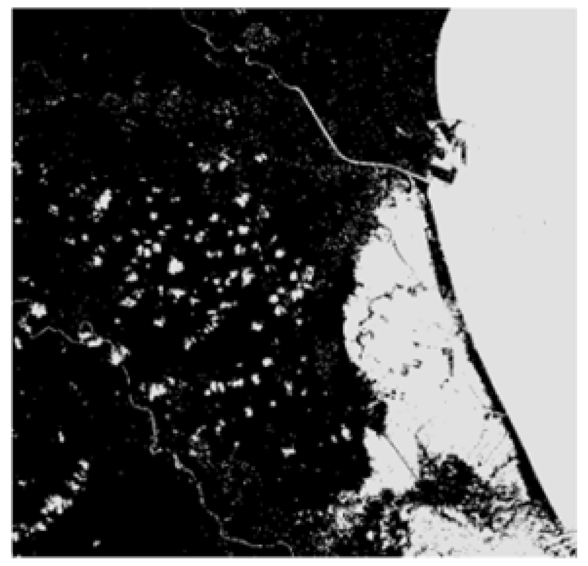

Figure 3: Sentinel-2 (NDWI layer) image with clouds, before applying the S2PixelCloudDetector neural network.

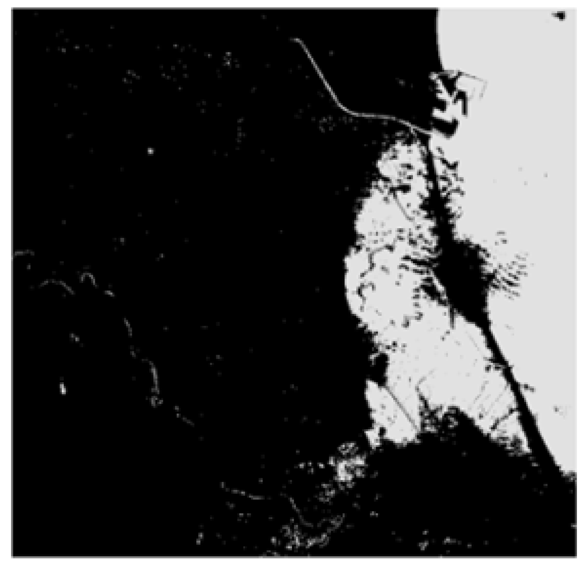

Figure 4: Sentinel-2 (NDWI layer) image with the clouds masked, using the S2PixelCloudDetector neural network.

A key difficulty in cloud removal is reconstructing the area beneath the clouds. Floods are rare events, and images from such scenarios are typically absent from the datasets on which the model was trained. As a result, predictions of the underlying area often lack accuracy. While cloud masking provides a reliable way to identify cloud-covered regions, incorporating an inpainting model is an important step to achieve accurate water quantification.

V. Water Quantification

This floodwater depths analysis is based on the paper by Sagy Cohen, *"Estimating floodwater depths from flood inundation maps and topography"* [1], and performs comprehensive flood volume analysis using geospatial raster data derived from Sentinel-2 satellite imagery. The methodology integrates two critical datasets:

1. Modified Normalized Difference Water Index (MNDWI) raster for flood extent detection. [3]

$$MNDWI = \frac{\rho_{\text{Green}} - \rho_{\text{SWIR}}}{\rho_{\text{Green}} + \rho_{\text{SWIR}}}$$

Where:
- ρ_{Green} = surface reflectance in the green band B03
- ρ_{SWIR} = surface reflectance in the shortwave infrared band

2. Digital Elevation Model (DEM) for topographic information.

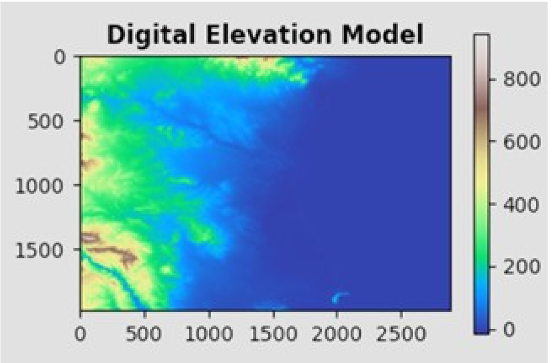

Figure: Digital Elevation Model of the Valencia region

The displayed chart is a DEM of the Valencia region, visualizing land surface elevation in meters. Color gradients represent altitude: lower elevations appear in deep blue, mid-range elevations in green and yellow, and higher terrain in brown. The methodology applies the FwDET (Floodwater Depth Estimation Tool) to estimate flood depths and volumes within a specified Region of Interest (ROI).

The pseudo-code methodology for the water depth estimation:

Algorithm 1 Flood Water Quantification (FWQ) Pseudo-Code

```
Given DEM, MNDWI, AOI bounds, pixel size s
Mask flooded pixels: M = [MNDWI > θ]    where θ is water threshold
Identify boundary pixels: B = M − erode(M)
for each flooded pixel p ∈ M do
    Find closest boundary pixel b ∈ B
    Assign water elevation: h_w(p) = h(b)
    Compute depth: d(p) = max(h_w(p) − h(p), 0)
end for
Compute total volume: V = Σ_{p∈M} d(p) s^2
Compute flooded area: A = |M| s^2
Aggregate depth statistics (mean, max, classify by ranges)

Where:
▷ h(p): DEM elevation at pixel p
▷ h_w(p): water surface elevation from nearest boundary
▷ d(p): depth at pixel p
▷ s: pixel size (meters)
▷ M: mask of flooded pixels
▷ B: mask of boundary pixels
```

Figure 5: Sentinel-2 image of the Valencia region (5km extent) showing baseline land conditions prior to flooding.

Figure 6: Sentinel-2 image of the Valencia region (5km extent) capturing the flooded extent after the event on 29th of October 2024.

VII. Ford Plant Damage Assessment

As the final stage of the project, we estimated the flood damage for the Ford Valencia Body & Assembly (VB&A) Plant, one of Spain's largest car manufacturing facilities. The region of interest (ROI) was defined using Copernicus satellite coordinates of the plant which has a built-up area of $\sim 300{,}000\,\text{m}^2$.
The analysis was performed on the super-resolved tiff files $10 \times 10\,\text{m}$ tiles covering $100\,\text{m}^2$ each.

Table: Maximum Structural Damage Analysis

ROI	Flooded Area (ha)	m^2	Euro damage/m²	Damage Cost
Ford Plant	15.4	154,000	€319	€49.1 million

Using an industrial damage rate of €319 per square meter [4], the maximum damage cost was estimated at €49,126,000.
A secondary damage estimate was computed based on flood depth:

Table: Depth Specific Structural Damage Analysis

Depth Category	Mega Liters	ML (1000 m³)	Damage Ratio	Max Cost (€)
Very Shallow **(0.1m - 1m)**	3.92	3920	0.15	1.069.090,91
Shallow **(1m - 2m)**	6.74	6740	0.4	1.797.333,33
Medium Shallow **(2m - 3m)**	7.29	7290	0.61	1.778.760,00
Medium Deep **(3m - 4m)**	4.67	4670	0.77	1.027.400,00
Deep **(4m - 5m)**	5.36	5360	0.92	1.095.822,22
Very Deep **(5m+)**	4.92	4920	1	820.000,00
Total				**7.588.406,46**

Damage estimation by applying depth-specific damage ratios to reflect the severity of impact on different structures and equipment. [4]

Future Work

- Incorporate temporal priors for cloud removal. Combine current images with observations from ∼1 week before and after the flood event to the inpainting model, Dsen2-CR to improve reconstruction of cloud-covered regions.
- Flood is considered a rare event and therefore the trained DL cloud removal model is lacking accuracy. A way to address this is to create synthetic data using GANs and train the model with these data to increase the prediction and simulation of the underneath region area.

References

[1] Sagy Cohen et al. "Estimating Floodwater Depths from Flood Inundation Maps and Topography". In: *Journal of the American Water Resources Association (JAWRA)* (2017).

[2] Copernicus. *Copernicus Browser*. 2025.

[3] Yun Du et al. "Water Bodies' Mapping from Sentinel-2 Imagery". In: *Remote Sensing* 8.4 (2016).

[4] Jan Huizinga, Hans de Moel, and Wojciech Szewczyk. *Global flood depth-damage functions database*. European Commission, Joint Research Centre (JRC). 2017.

[5] Charis Lanaras et al. "Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network". In: *ISPRS Journal of Photogrammetry and Remote Sensing* (2018).

[6] Andreas Meraner et al. "Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion". In: *ISPRS Journal of Photogrammetry and Remote Sensing* (2020).

[7] Aljaz Zupanc and et al. *Sentinel-2 Cloudless: Global cloud mask for Sentinel-2 imagery*. https://github.com/sentinel-hub/sentinel2-cloud-detector. 2017.