Surrogate modelling and user-in-the-loop experimentation for urban flood prediction: the EXTREMEXP approach

Pauline Delporte[1], Gwendoline Stéphan[1], Yasmine Boulfani[1], George Papastefanatos[2], Vincent Gaudissard[1]
[1] CS GROUP – France, Toulouse, France
[2] Athena Research Center – Greece, Athena

Context and objectives

[1] https://extremexp.eu/

ExtremeXP project [1]

Decisions based on data-driven insights can be vital and have a phenomenal impact on the environment, society and business. Many critical domains such as crisis management, predictive maintenance, mobility, public safety, and cyber-security become increasingly disrupted by new means to harness the extreme proliferation of data for effective decision making. Generating data-driven insights that can be used and trusted by decision-makers is, however, still far from trivial.

ExtremeXP proposes a new paradigm for data analytics, which we call experimentation-driven analytics. The main contribution is to put the end user at the center of complex analytics processes from data discovery to novel interactions, proposing a human-in-the-loop. The framework will integrate interactive visualization and explainability techniques to increase the trustworthiness of not only the outcomes but also of the process to reach such outcomes.

Use Case: flash flood in the city of Nîmes

The Cévennes Moutains in the South of France is prone to severe flood events mainly occurring in Fall, also known as Cévenols events. In particular, the city of Nîmes suffered multiple very localized storm which caused floods events (1988, 2002, 2005, 2014) during the last 4 decades, with lots of damages and even several deaths.
Since the 1988 event the city council of Nîmes has launched a vast program to prevent and mitigate those events.

This use case for **ExtremeXP** aims to setup a surrogate method to replace physics methods to lead to faster predictions of flood events.

Objective: Setup a surrogate model for the prediction of flash flood events in urban areas, especially for the city of Nîmes.

Methodology

Dataset

Topographical data
- High resolution Digital Elevation Model from IGN [1].
- Building footprints from IGN or OpenStreetMap [2].

Meteorological data
- Rainfall information from CALAMAR system, a software that evaluate the rain activity with radars and gauges [3].

AI model architecture

The deep learning model selected is a UNet associated with a temporal attention layer. This particular layer encodes the temporal aspect for our data and propagate the information through the decoder [4].

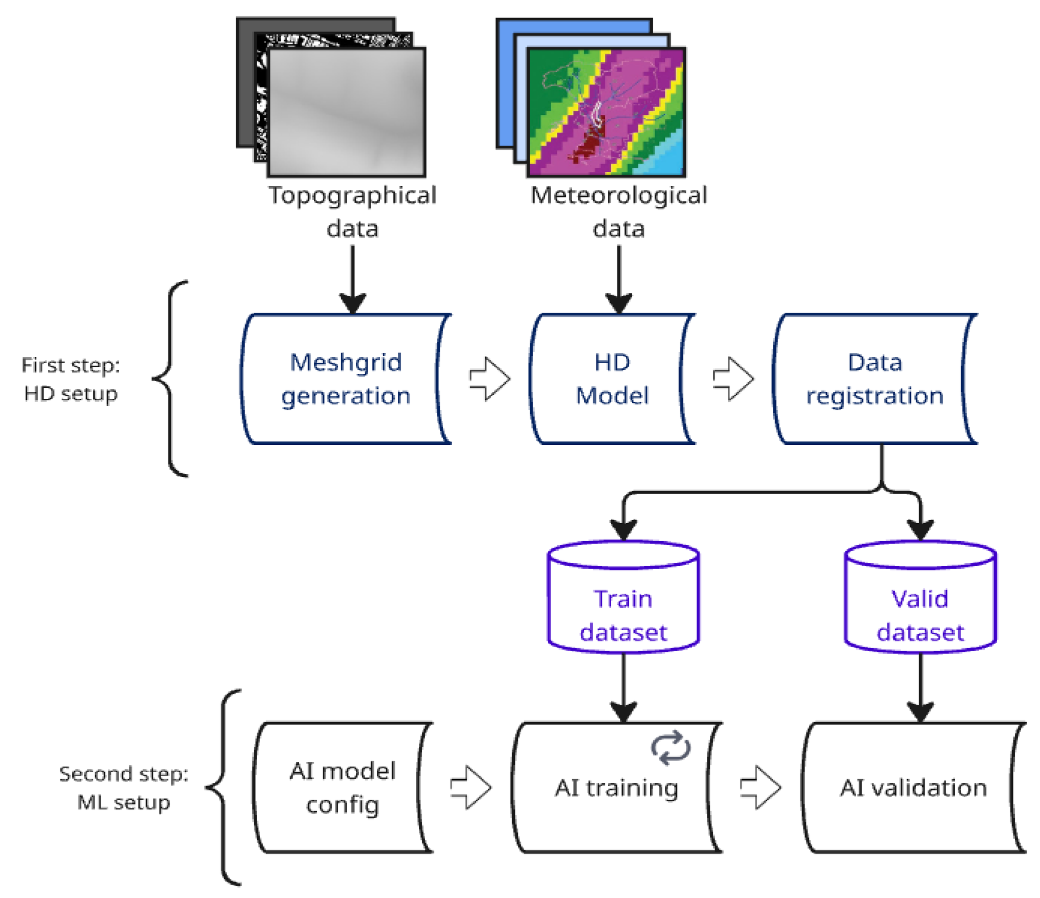
[2] https://geoservices.ign.fr/rgealti
[3] https://www.openstreetmap.org
[4] http://wikhydro.developpement-durable.gouv.fr/index.php/Calamar_(HU)
[5] Chaudhary, P.; Leitão, J.P.; Schindler, K.; Wegner, J.D. "Flood Water Depth Prediction with Convolutional Temporal Attention Networks". Water 2024.

General methodology

Hydrodynamic model setup [5][6]
- Using publicly available datasets for the meshgrid
  - OpenStreetMap (building and streets footprints).
  - LIDAR High Resolution DEM.
- Semi-automated workflow for the generation of input data
  - Mesher dedicated to geospatial datasets.
  - Computation of the boundary conditions.

AI model setup
- Model and training configuration
- Training step with dataset from HD method
- Validation against physical model

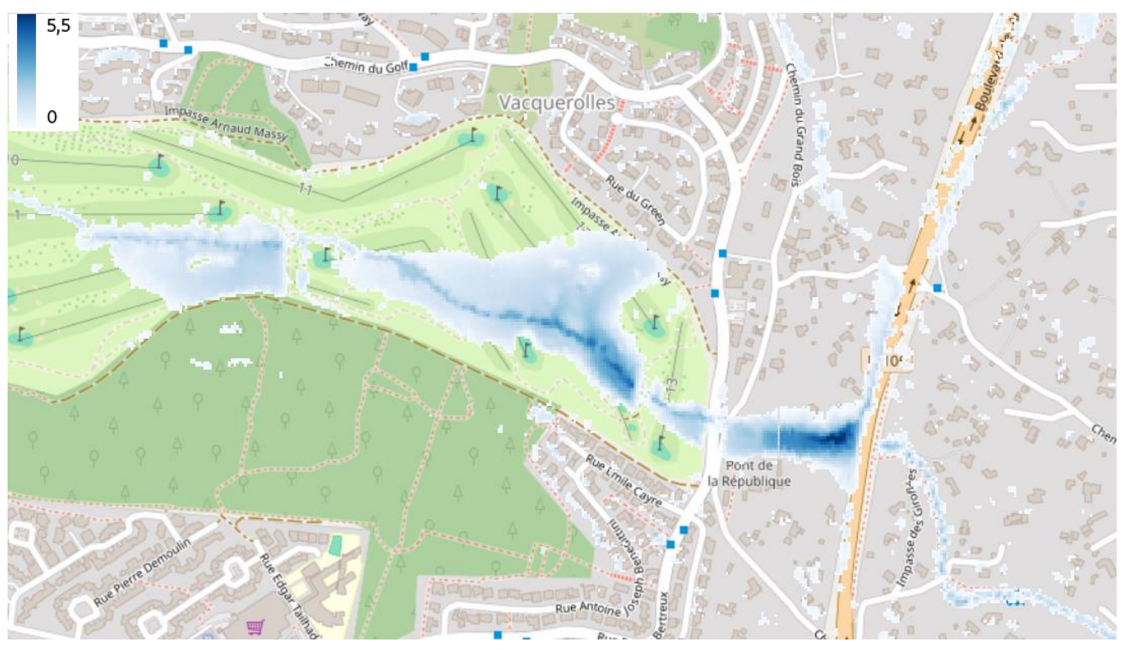
[6] J.R Shewchuk, "Triangle: Engineering a 2D quality mesh generator and Delaunay triangulator", 1996
[7] Jonathan Richard Shewchuk, "Delaunay Refinement Algorithms for Triangular Mesh Generation", "Computational Geometry", 2002.



General methodology

Metrics & Visualization

Example of 2D Visualization in 2005 event



Prediction map

Ground truth map

Integration in the ExtremeXP plateform

The UC is integrated inside the ExtremeXP plateform, with the goal to put the user at the center of the experimentation.
The framework with its different modules (visualization, explainability, etc) allows the user to give feeback in order to improve at the next iteration of the model.

2 types of metrics:
- **Classification metrics**: recall, precision and CSI
- **Regression metrics**: MSE, MAE

Name	Recall$^{0.20}$	Precision$^{0.20}$	CSI$^{0.20}$	MSE	Accuracy$^{0.05}$
Nîmes 256	0.86	0.91	0.80	0.03	0.94
Nîmes 512	0.91	0.90	0.82	0.01	0.97

Metrics Table

Conclusion and perspectives

Conclusion
- This work presented a flash flood prediction UC developed within the ExtremeXP platform, focusing on the city of Nîmes and the events of 2002, 2005 and 2014. By combining hydrodynamic simulations with a deep learning surrogate model, the system predicts water depth 30 minutes ahead, showing encouraging performance.
- The ExtremeXP platform has supported the full lifecycle of experimentation, from data integration to visualization, emphasizing user involvement and explainability.

ExtremeXP framework perspectives
- Integrated tools to help decision-making by giving them the right information for a better trust in the model.

AI model setup perspectives
- Generalization to other cities and different conditions (lake of data, other resolution, different time prediction).
- Developed methods to measure the uncertainty of the model and the data.

Outlook and recommendations

EO products for the validation of the inundation dynamics
- Nowadays it is difficult to derive flood inundation extents during the event. Most of the studies use watermasks computed after the event, mainly because they use optical sensors.
- Even if possible to derive flood inundation extents the repetitivity of existing EO missions are not sufficient for flash floods. However, some commercial missions dedicated to this subject exist.

The ExtremeXP project is co-funded by the European Union Horizon Program HORIZON-CL4-2022-DATA-01-01, under Grant Agreement No. 101093164