

Generalizability of deep-learning-based emulation of hydrodynamic flood models

Rongxin Li^a, Lachlan Tychsen-Smith^a, Fazlul Karim^b, David J. Penton^b
and Mohammed Ali Armin^a

^a CSIRO Data61, Canberra, Australia

^b CSIRO Environment Business Unit, Canberra, Australia

Email: ron.li@csiro.au

Abstract: Scenarios of flood inundation are traditionally simulated by numerically solving the partial differential equations (PDEs) that govern fluid dynamics, with the initial and boundary conditions derived from observational data. While proven highly valuable for improving flood management and risk mitigation (Karim et al., 2023), such hydrodynamic simulations are limited in its applicability in near-real-time or emergency management due to its high demand on time and hardware resources.

Drastically improved efficiency can result from leveraging the power of deep artificial neural networks (ANNs). In a recently proposed methodology known as physics-informed machine learning (PIML, Karniadakis et al., 2021), ANNs are trained to obtain numerical solutions based on the PDEs and limited observational data. A more mature and realistic category of methods, on the other hand, employs ANNs to emulate traditional hydrodynamic simulators. That is, an ANN is trained to predict the output of a physics-based simulator given its inputs. This category, often referred to as surrogate models, includes algorithms that predict flood inundation at specific future time points, as well as those that generate series of predictions.

Compared with PIML, a key concern regarding ANN-based surrogate models is their generalizability. That is, if the simulation data used to train the surrogate model are insufficiently diverse (in terms of the covered ranges of scenarios and conditions) or inadequately representative, whether the model is able to effectively learn the underlying physics in the absence of physics-based constraints.

Recently, we proposed an ANN-based framework for rapid inundation modelling in large regions (Tychsen-Smith, et al., 2023). This framework takes into account the local elevation, surface roughness, river inflows and the current water heights, and predicts the water heights at the subsequent time point. The prediction is performed at grid points in a rolling forward fashion. In this paper we focus on case studies on the generalizability of this framework. Specifically, we aim to answer the following questions: How well does a model based on our framework perform on unseen floods in the same location? How well does a model that was trained in one location perform in a different location? How much can the performance on unseen floods in unseen locations improve with an additional simulation in an additional location? Finally, we propose measures aimed at improving the generalizability of an ANN-based model.

REFERENCES

- Karim F, Armin MA, Ahmedt-Aristizabal D, Tychsen-Smith L, Petersson L. A Review of Hydrodynamic and Machine Learning Approaches for Flood Inundation Modeling. *Water*. 2023; 15(3):566. <https://doi.org/10.3390/w15030566>
- Karniadakis, G.E., Kevrekidis, I.G., Lu, L. et al. Physics-informed machine learning. *Nat Rev Phys* 3, 422–440 (2021). <https://doi.org/10.1038/s42254-021-00314-5>
- Tychsen-Smith, L., Armin, M. A. and Karim, F., Large-Scale Inundation Modelling with Machine Learning Methods. <http://dx.doi.org/10.2139/ssrn.4372800>

Keywords: *Flood inundation, inundation modelling, machine learning, deep learning, hydrodynamic modelling*