EGU2020-6254
https://doi.org/10.5194/egusphere-egu2020-6254
EGU General Assembly 2020

Efficient simulation of flood events using machine learning

Jihane Elyahyioui[1,2], Valentijn Pauwels[1], Edoardo Daly[1], Francois Petitjean[3], and Mahesh Prakash[2]
[1]Monash University, Faculty of Engineering, Civil Engineering, Australia
[2]CSIRO Data61, Docklands, VIC 3008, Australia
[3]Monash University, Faculty of Information Technology, Centre for Data Science, Australia

Flooding is one of the most common and costly natural hazards at global scale. Flood models are important in supporting flood management. This is a computationally expensive process, due to the high nonlinearity of the equations involved and the complexity of the surface topography. New modelling approaches based on deep learning algorithms have recently emerged for multiple applications.

This study aims to investigate the capacity of machine learning to achieve spatio-temporal flood modelling. The combination of spatial and temporal input data to obtain dynamic results of water levels and flows from a machine learning model on multiple domains for applications in flood risk assessments has not been achieved yet. Here, we develop increasingly complex architectures aimed at interpreting the raw input data of precipitation and terrain to generate essential spatio-temporal variables (water level and velocity fields) and derived products (flood maps) by training these based on hydrodynamic simulations.

An extensive training dataset is generated by solving the 2D shallow water equations on simplified topographies using Lisflood-FP.

As a first task, the machine learning model is trained to reproduce the maximum water depth, using as inputs the precipitation time series and the topographic grid. The models combine the spatial and temporal information through a combination of 1D and 2D convolutional layers, pooling, merging and upscaling. Multiple variations of this generic architecture are trained to determine the best one(s). Overall, the trained models return good results regarding performance indices (mean squared error, mean absolute error and classification accuracy) but fail at predicting the maximum water depths with sufficient precision for practical applications.

A major limitation of this approach is the availability of training examples. As a second task, models will be trained to bring the state of the system (spatially distributed water depth and velocity) from one time step to the next, based on the same inputs as previously, generating the full solution equivalent to that of a hydrodynamic solver. The training database becomes much larger as each pair of consecutive time steps constitutes one training example.

Assuming that a reliable model can be built and trained, such methodology could be applied to build models that are faster and less computationally demanding than hydrodynamic models.

Indeed, in with the synthetic cases shown here, the simulation times of the machine learning models (< seconds) are far shorter than those of the hydrodynamic model (a few minutes at least). These data-driven models could be used for interpolation and forecasting. The potential for extrapolation beyond the range of training datasets will also be investigated (different topography and high intensity precipitation events).