Urban Drainage Systems modelling for Early Warning Service Using Data-Driven Modelling

Solomon Seyoum\textsuperscript{1,3}, Patrick Willems\textsuperscript{1,2} and Boud Verbeiren\textsuperscript{1}

\textsuperscript{1} Vrije Universiteit Brussel, Department of Hydrology and Hydraulic Engineering, Brussels, Belgium
\textsuperscript{2} Brussels Company for Water Management (SBGE/BMWB), Direction Exploitation, Brussels, Belgium
\textsuperscript{3} IHE Delft Institute for Water Education, Delft, Netherlands

Contact: s.seyoum@un-ihe.org / Boud.Verbeiren@vub.be
Problem statement

Pluvial flooding occurs frequently in Brussels capital region

6912 reports of flooding in total from 1992 to mid-2018 recorded (Brussels Environment)

**Pluvial floods** are typical the result of intense rainfall, triggering a fast hydrological response in cities.

No hydraulic model available to use for flood forecasting

Within the **FloodCitiSense** project we are exploring the use of data-driven models to forecast pluvial flooding for Brussels
We delineated the Brussels Capital Region into 54 subcatchments using:

- Digital Terrain Model
- Subcatchments from Brussels Environment
- looking at the existing drainage network
- Considering the flow gauging stations and reservoirs in the city
- local expert knowledge
Flood reports per subcatchment

- Flood reports from Brussels Environment for each sub catchment by date
- Flood reports total per subcatchment
Focus subcatchments

• 9 sub-catchments were selected based on flooding history
• 9 Random Forest models were developed
Why DDM

• Data availability

• Absence of hydraulic model
  • Hydraulic models (1D or 2D) require detailed data about the system
    • DTM, dimension of each network element such as slope, size, depth,
    • Require much longer run time (unsuitable for early warning system)

• DDM – relationship between input and output without the need to understanding the mechanism underlying the system
  • Require large amount of data for training and testing
  • quick runtime, suitable for early warning
Early flood warning system development

- On-going
Random Forest models

• For a model to perform adequately
  • determination of optimal model input and information in addition to the current time step is needed
  • Additional information can be derived from previous time step rainfall and runoff

• The Data driven models (Random Forest) use rainfall data of 5 most correlated rainfall stations (sum of RF values for the past 2 hours) and the flow data of the station before 2 hours to forecast the current flow

\[ Q_t = f \left( Q_{t-\text{lag}}, \sum_{j=t-\text{lag}}^{i=t} RF_{i,j} \right) \text{ for } i = 1 \text{ to } 5 \]

• After all the parameters of the DDM are tuned for each flow stations, the models can be used to forecast flows
Some results – station C02 – Good performance

C02 RandomForest Training

R² = 0.93

C02 RandomForest Test

R² = 0.84
Some results – station U17 – average performance
Some results – station U11 – low performance

**U11 Random Forest Training**

- $R^2 = 0.53$

**U11 Random Forest Test**

- $R^2 = 0.18$
Results and discussion

- For each flow station, RF models are being trained and tested.
- In most cases the DDMs perform well with R-squared values ranging from 0.55 to 0.98 for a 2-hour forecast horizon.
- Their performance increases 10 to 40% for a one-hour forecast horizon.
- Training the RF models are faster as they are classification models.
- The method can be used for other case studies if enough data is available (at least 2 to 3 years data with 5 minutes resolution).
- Large amount of data is needed to train and test DDMs.
- Can only forecast flow at measuring stations.
- Cannot be used to forecast depth and spatial extent of flooding.
Conclusion

• The use of DDM to forecast pluvial flooding was tested and shows promising results.

• DDM provide a means to forecast occurrence of pluvial flooding in absence of detailed hydraulic model for such a complex drainage system.

• Underestimation of peak flow is observed for RF models.

• Appropriate data transformation is said to improve the performance of such models (Sudheer, et al., 2003).

OUTLOOK:

• Include radar rainfall data as input for the training of the models

• Fine tune the DDM parameters