A prelude to emulation for *flood prediction*

March 12, 2018

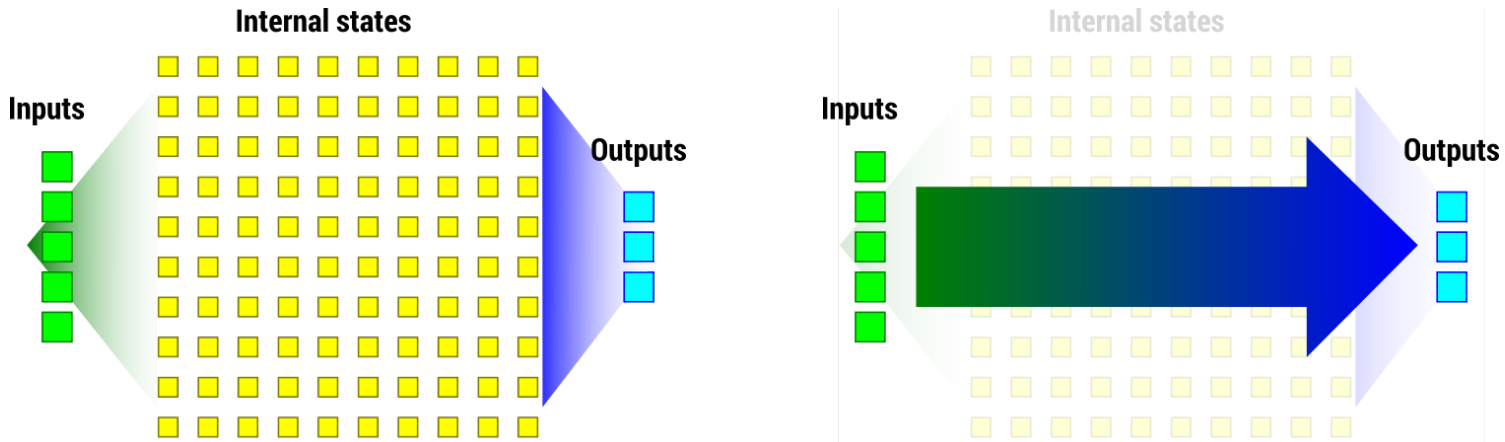
Sebastiano Rusca
Eawag: Swiss Federal Institute of Aquatic Science and Technology

Outline

- What is emulation?
- Why to emulate?
- When to emulate?
- What are the tools for emulating?
- Case studies
  - A mechanistic emulator: *fitting the weir equation*
  - A hydrological emulator: *estimating the time-to-threshold*
- Conclusion & outlook

What is emulation?

Emulating means building an *ad hoc* data-driven surrogate model which closely mimics the behavior of the simulator on which it is based



(a) Simulator (b) Emulator

source: *"Model Order Reduction and Emulation." EmuMore's blog, March 9, 2018*

Why to emulate?

Simulation has great "power" **but**:

- Often all of this power is not actually needed
- The price to pay is a high computational cost

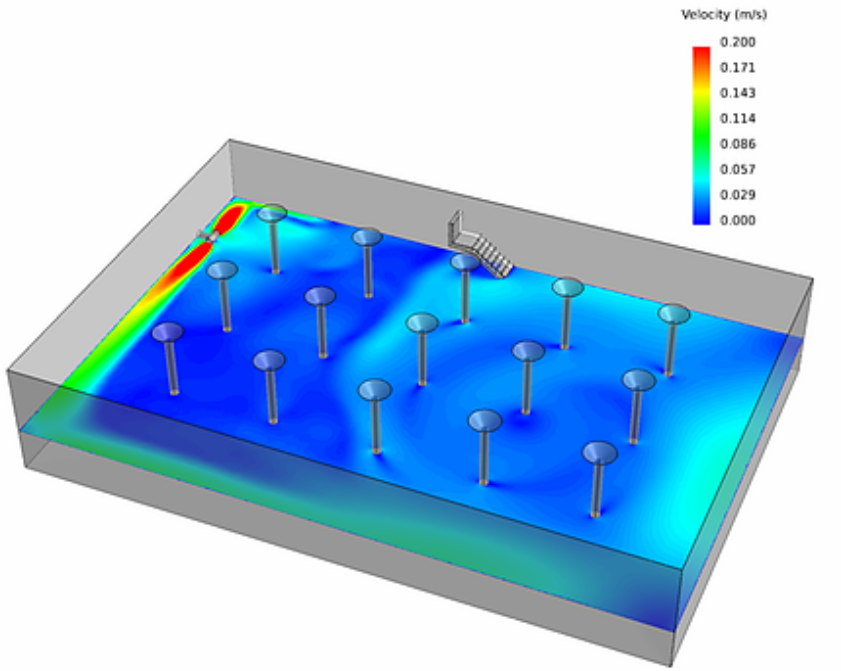
Emulation of *shallow water equation* can reach a speedup of $\mathbf{10^5}\times$

- *Real-time* applications can be done
- Very efficient for e.g. *uncertainty estimation*

When to emulate?

When we do not need to exploit all of the simulator's power

An example: outflow from a reservoir

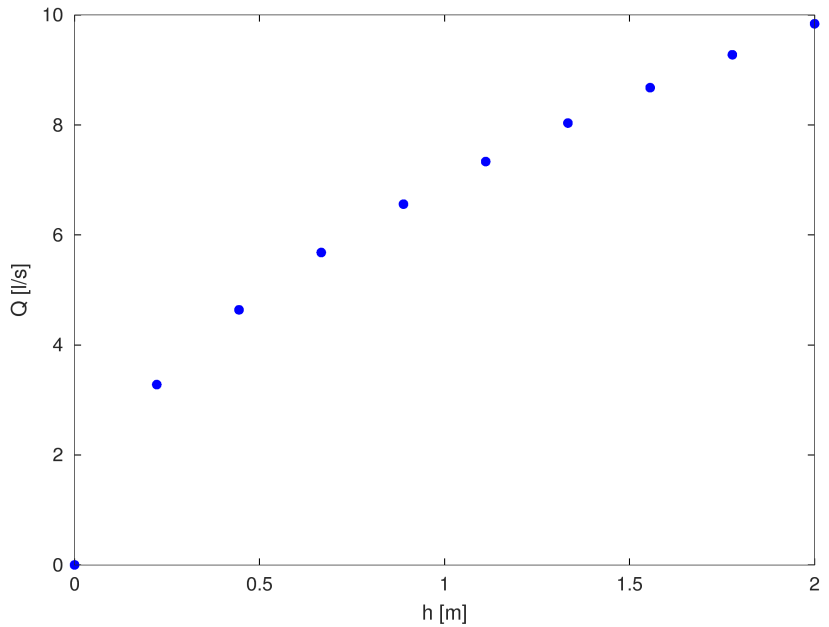


source: *https://www.fluidflowltd.com/projects*

When to emulate?

We plot the results:

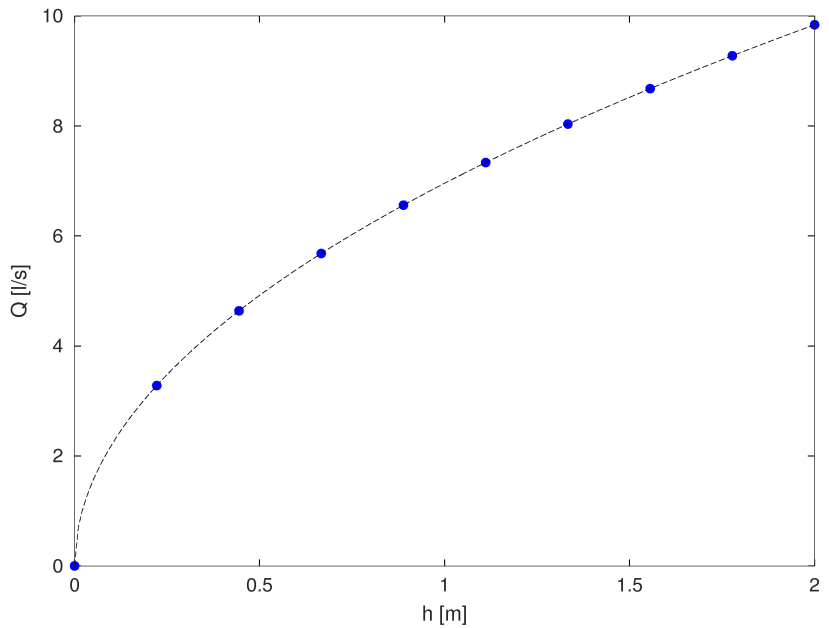
$$Q = f(h)\,?$$



When to emulate?

From the plot we manage to establish:

$$Q = c \cdot A_{outlet} \cdot \sqrt{2g \cdot h}$$



What are the tools for emulation?

- A simulator for generating the data
  - used: *FullSWOF_2D*
- A programming language for extracting the needed data
  - used: *GNU Octave*
- A data regression method
  - used: *linear regression, Gaussian processes*

Case studies

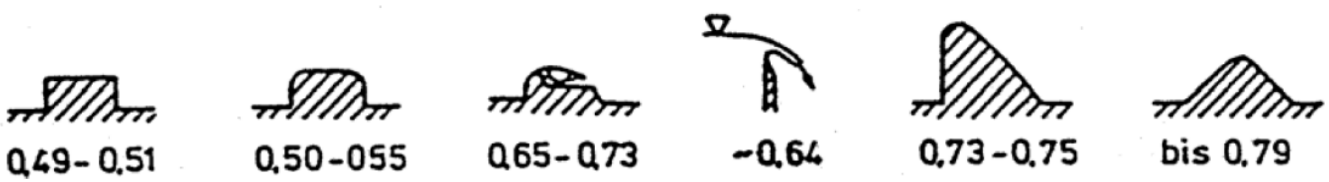
Mechanistic emulator: *the weir equation*

Goal: fit the weir equation to simulated data

Weir equation

$$Q = \frac{2}{3}\,\mu\,B_w\,\sqrt{2g}\,h_w^a, \quad usu.\ a = 3/2$$

μ :



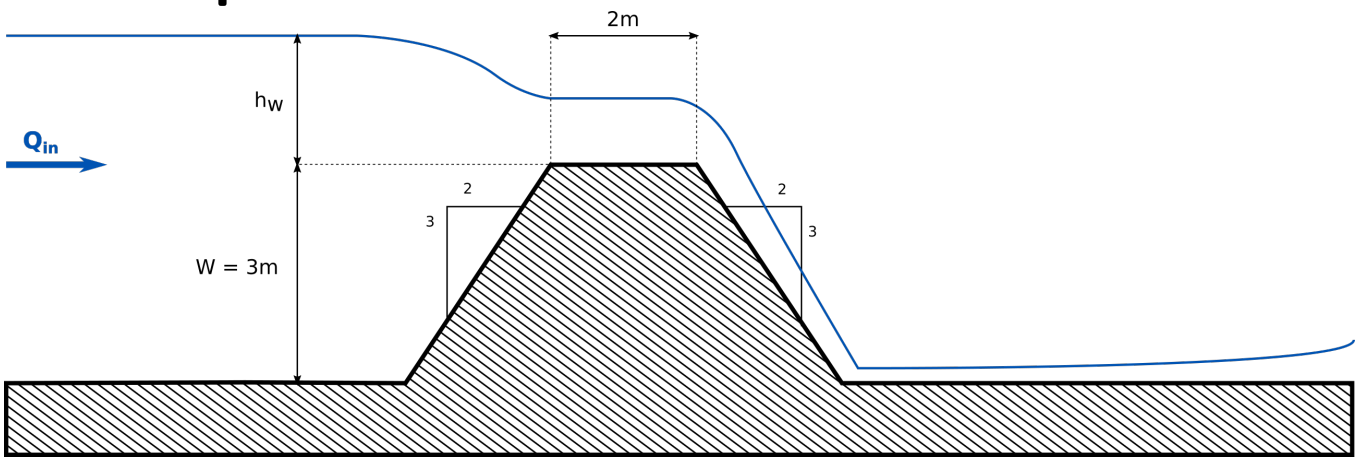
source: *Boes, Robert. "Wasserbau - Vorlesungsmanuskript." ETH Zürich - VAW, 2016.*

Simulations set-up

Input discharge

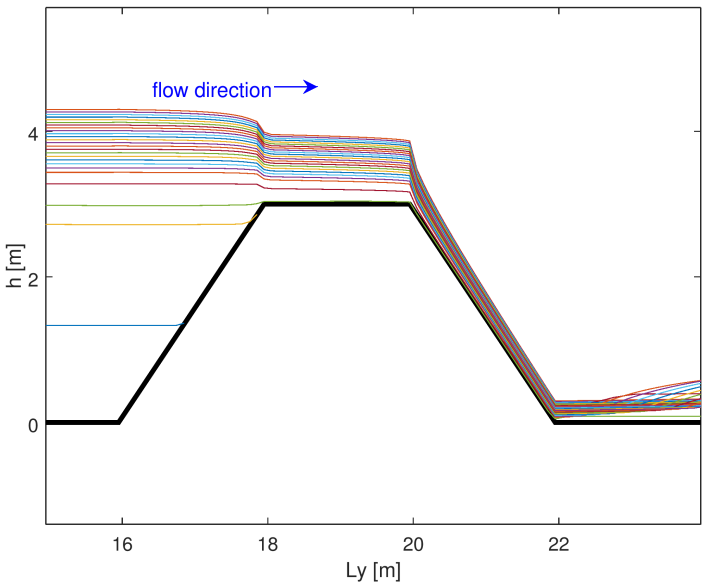
```
nQ   = 25; #number of experiments
Qin = linspace (0.1, 10, nQ); #Qin values [m3/s]
```

Channel set-up

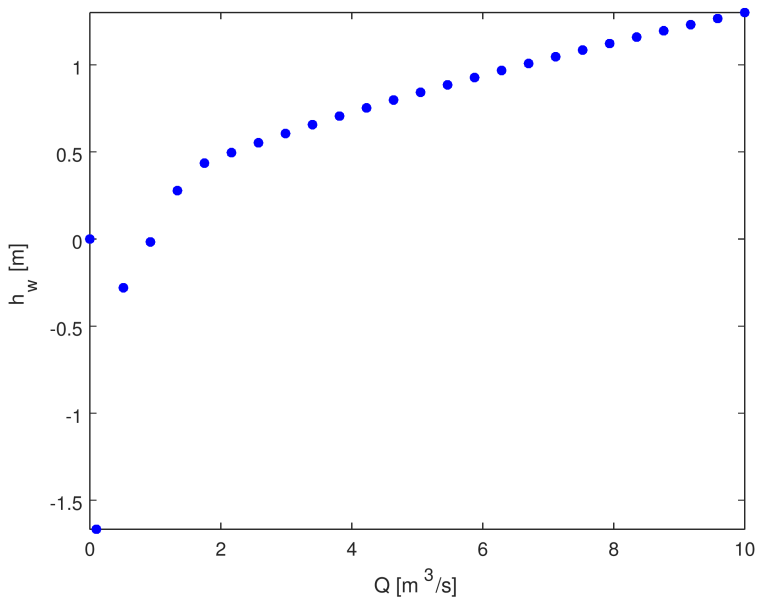


Simulations results

Experiments free-surfaces

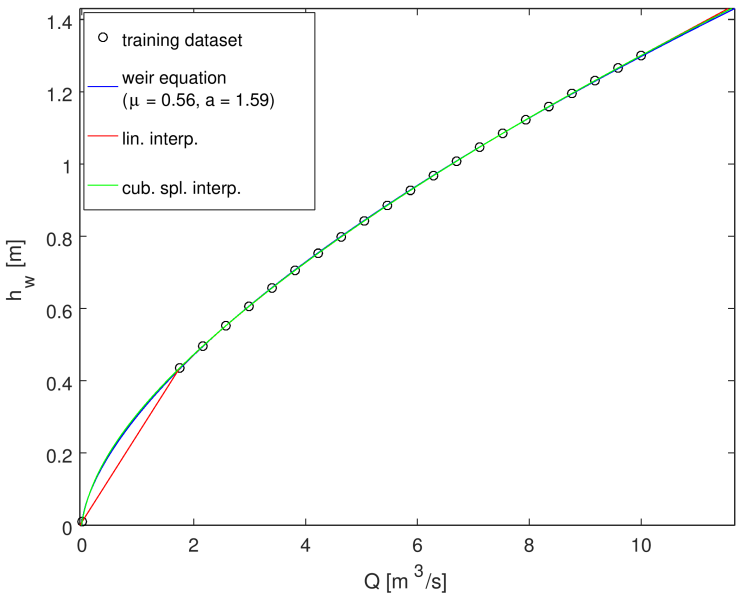


Experiments h_w vs. Q

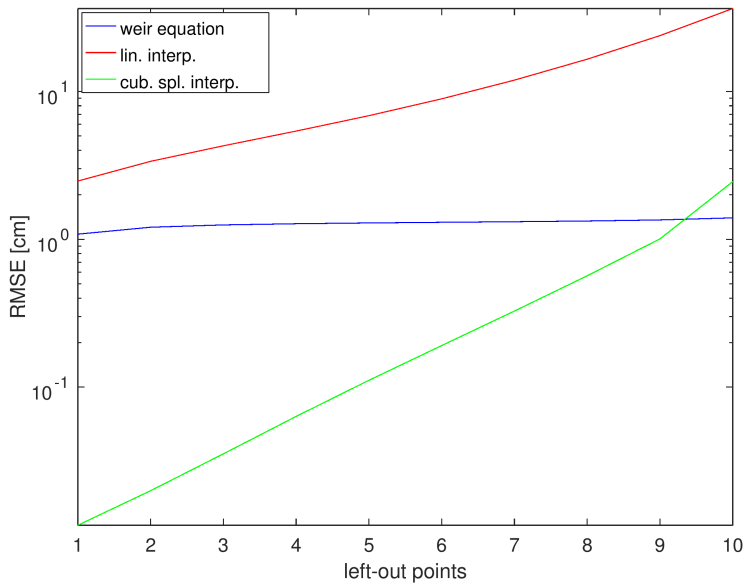


Fitting results

Fitting different models

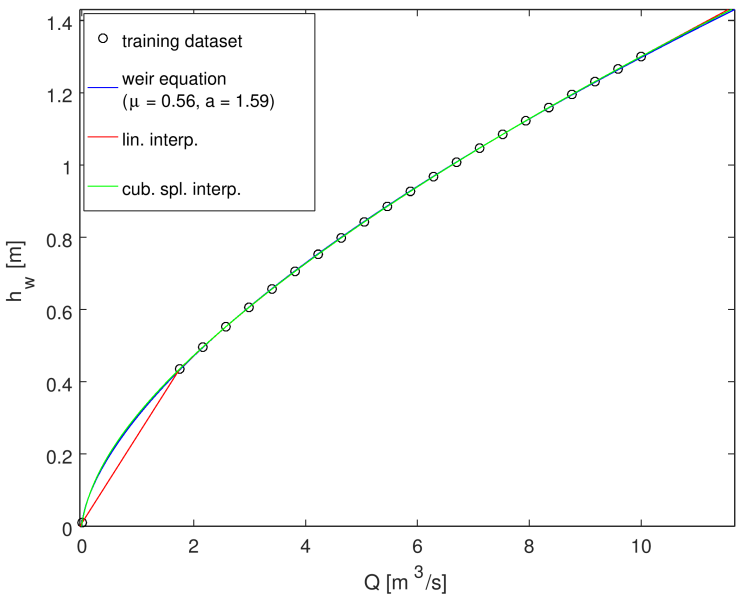


Fitting performance

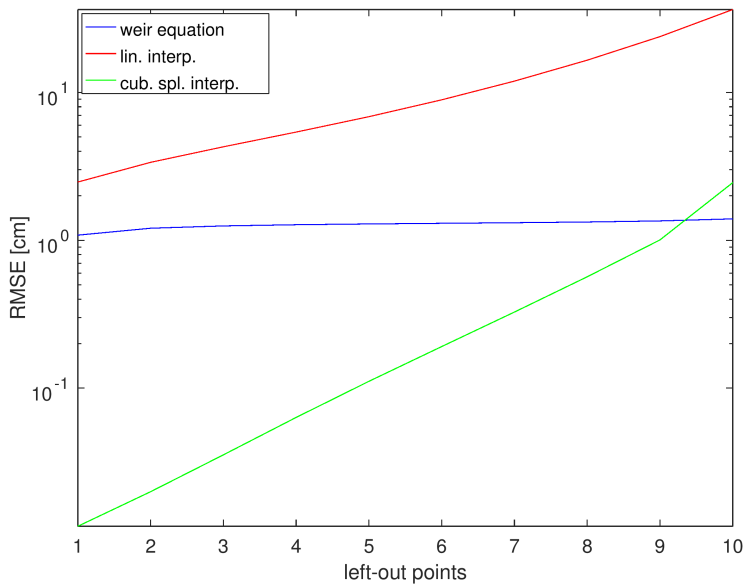


Fitting results

Fitting different models



Fitting performance



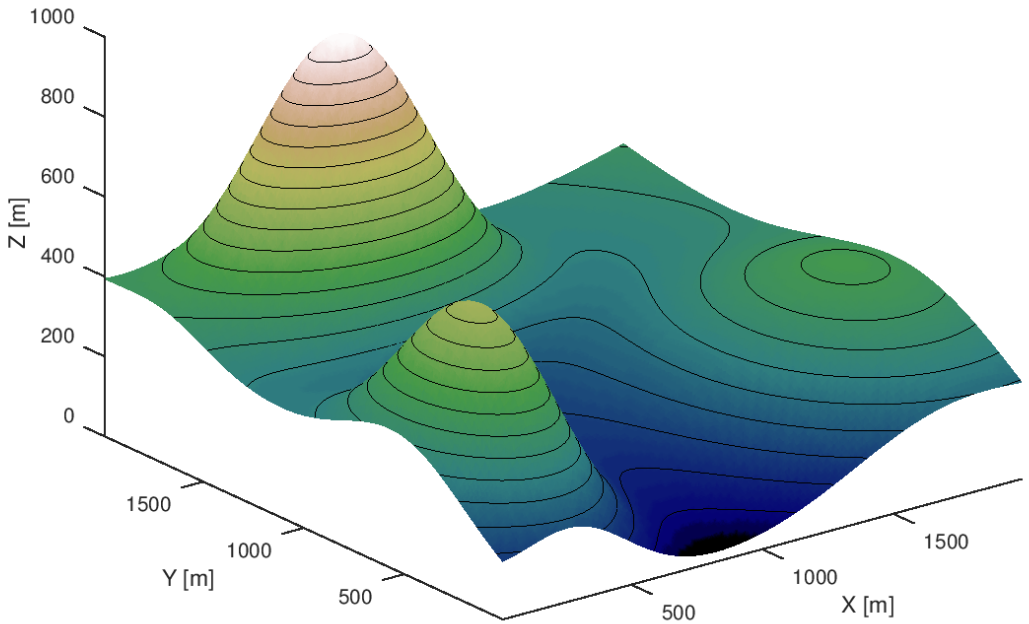
Take home message: if an accurate model is identified there is no need of much data for doing good predictions

Hydrological emulator: *time-to-threshold*

Goal: build an emulator to estimate the *time-to-threshold*

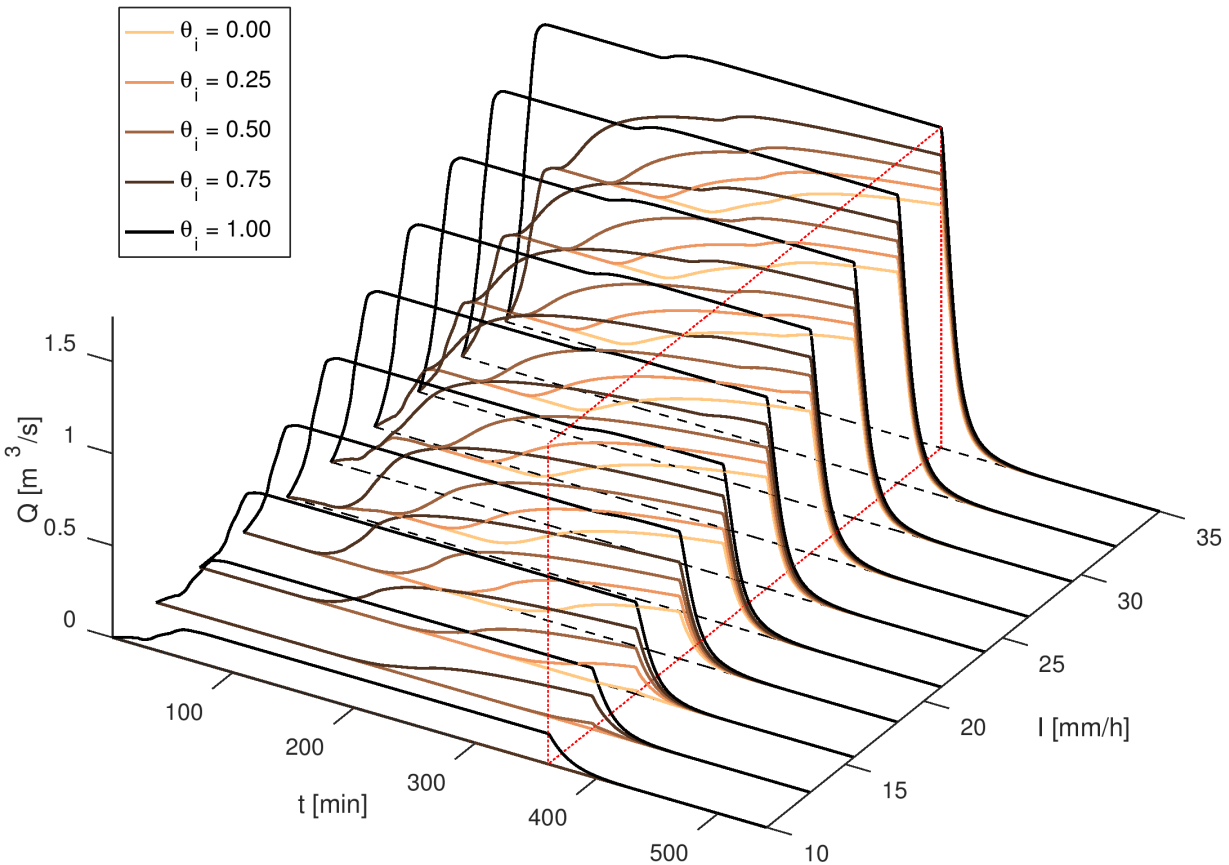
source: *https://www.dailyrecord.co.uk/news/scottish-news/storm-frank-rescuers-brave-storm-7096067*

Simulation set-up

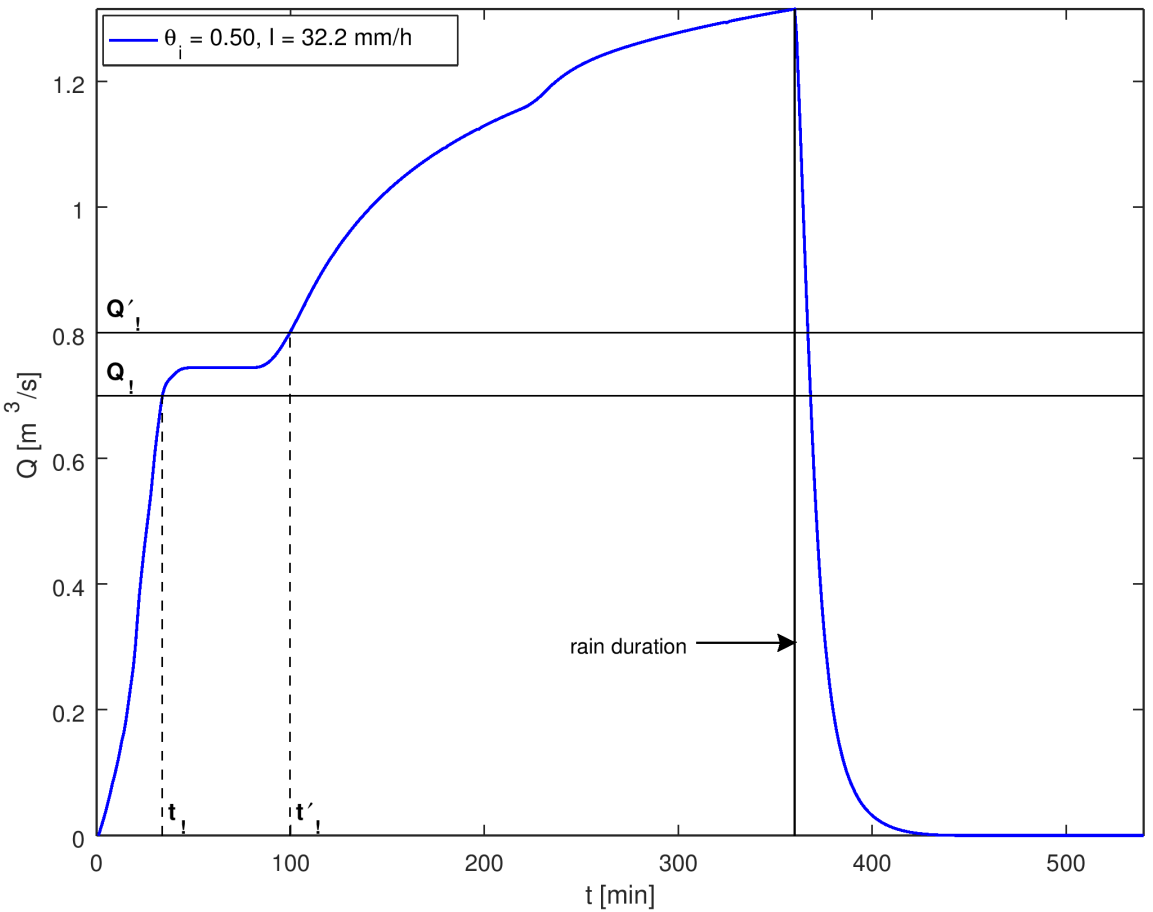



```
rain_intensities    = linspace (10, 35, 10); #[mm/h]
initial_saturations = linspace (0, 1, 5); #[-]
simulation_duration = 9; #[h]
rain_duration       = 6; #[h]
```

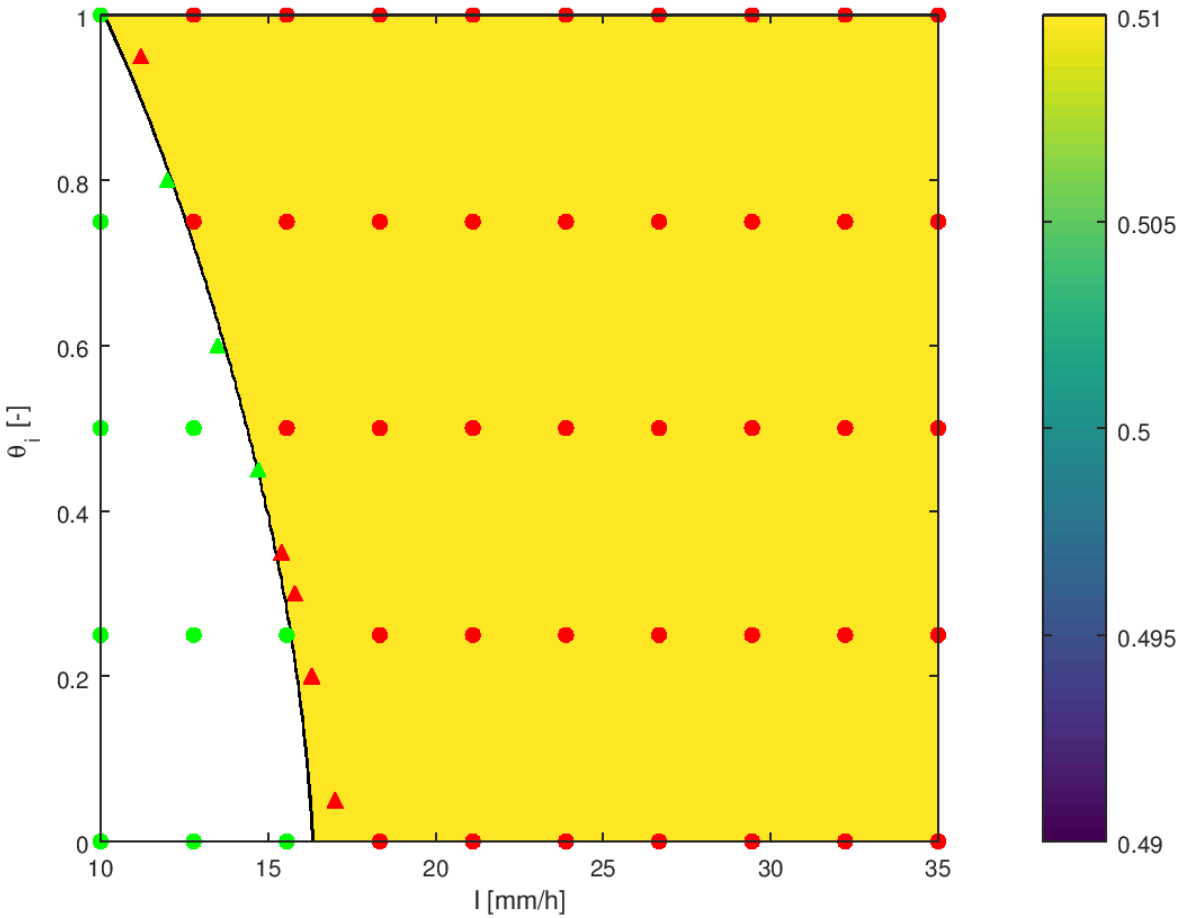
Simulations results



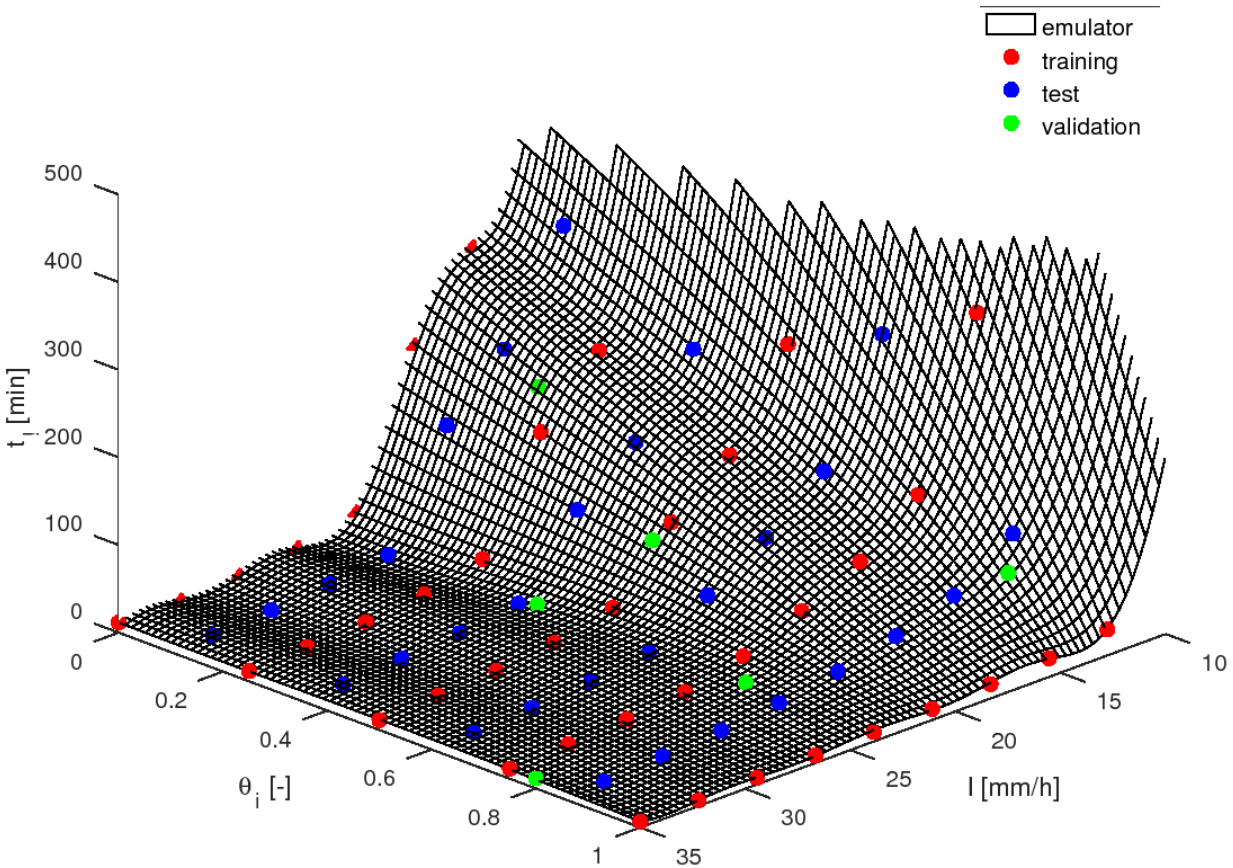
Emulator dataset extraction



Rain events classification



Time-to-threshold emulator



Time-to-threshold emulator

Validation performance

MAE [min]	RMSE [min]
21.8	10.8

Duration one simulation: $\approx 30\, min$
Duration one emulator-evaluation: $\approx 0.012\, s$

speedup factor: $1.5 \cdot 10^5$

Conclusion & Outlook

- Emulation approach is intrinsic to science
- Gaussian processes are a useful tool to:
  - incorporate prior knowledge in the modeling process
  - learn something about the functional relationship between the response and the predictors
- Case studies illustrate a subset of the wide range of potential applications of emulation
- Good accuracies at reproducing the simulator behavior can be reached but how good do we reproduce reality?
- Open science and open source software are the way to proceed if we want to given an answer to this question

THE END

Links to my repositories

https://bitbucket.org/binello7/fswof2d
https://bitbucket.org/binello7/master_thesis
https://bitbucket.org/binello7/master_thesis/wiki/Home