Project Seminar:
Flood prediction at area German Corner in Koblenz

Long Phan

ldphan@uni-koblenz.de

Supervisor: Prof. Dr. Thomas Goetz
University Koblenz, Department Mathematics

July 20, 2017

Overview

WARNING, Flood!

Understanding Data

Observation

- Flood events in years
- Notice about water level
- Location of container
- Rate of change of water level

Data we have

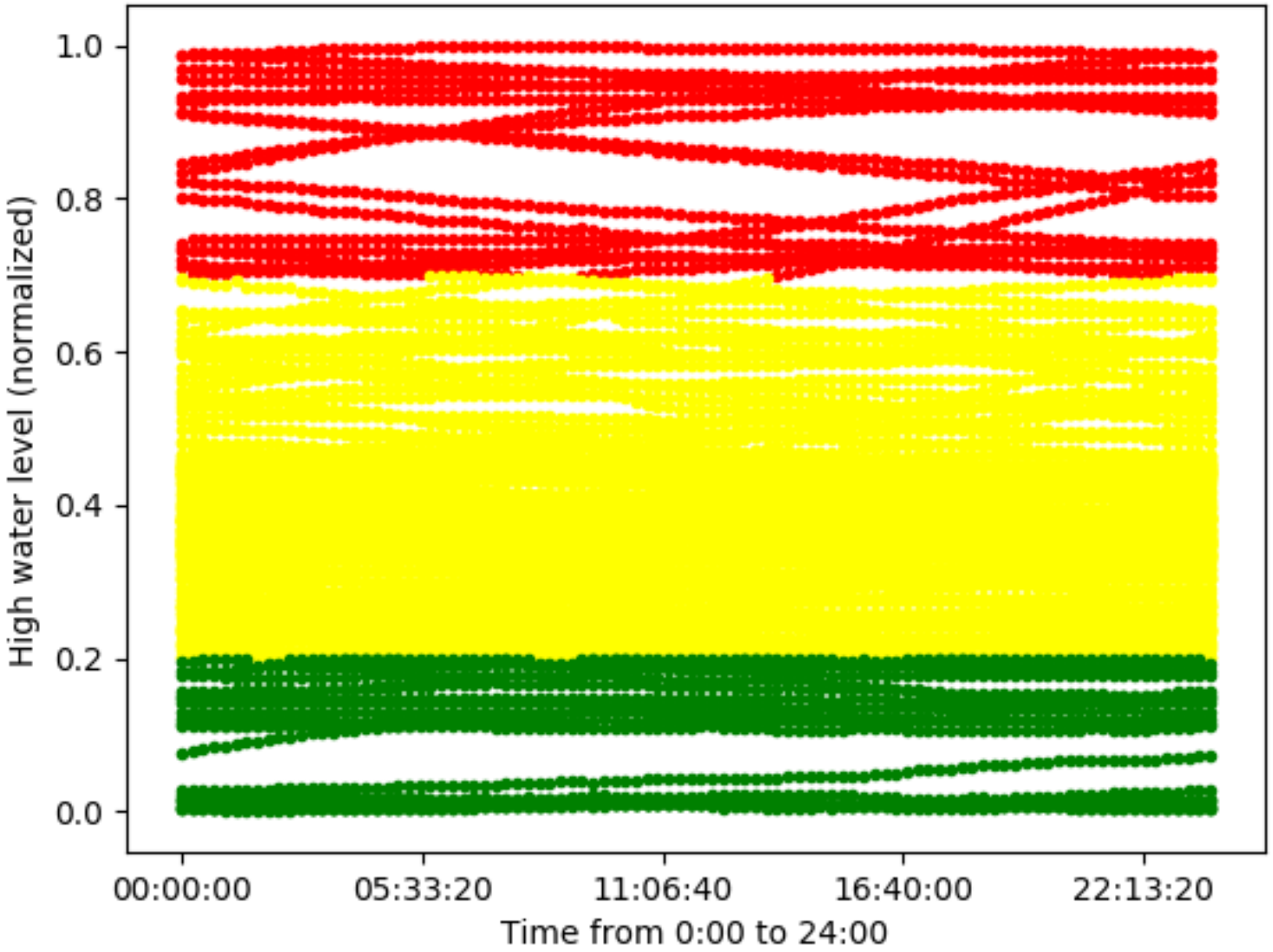
- water level in 4 years 1993, 1995, 2013, 2016
- timing from 0:00 to 24:00
- step time 15 minutes

Understanding Data

Data classification- Why and How?

- Why?
  - Reduce big dataset to small dataset, transform unstructured data to structured data
  - Recognize pattern from dataset to achieve information
  - Choose the right data for the calculation, reduce massive computation, speed-up calculation performance etc.
- How?
  Normalize data: method (Min-Max)
  Classify water level in 3 status: red, yellow, green
  Apply method of classification (nearest neighbour):
  - water-max: highest popular water level,
  - water-min: lowest popular water level,
  - derive average water level and domain of red, yellow, green.

Understanding Data



Design Solution

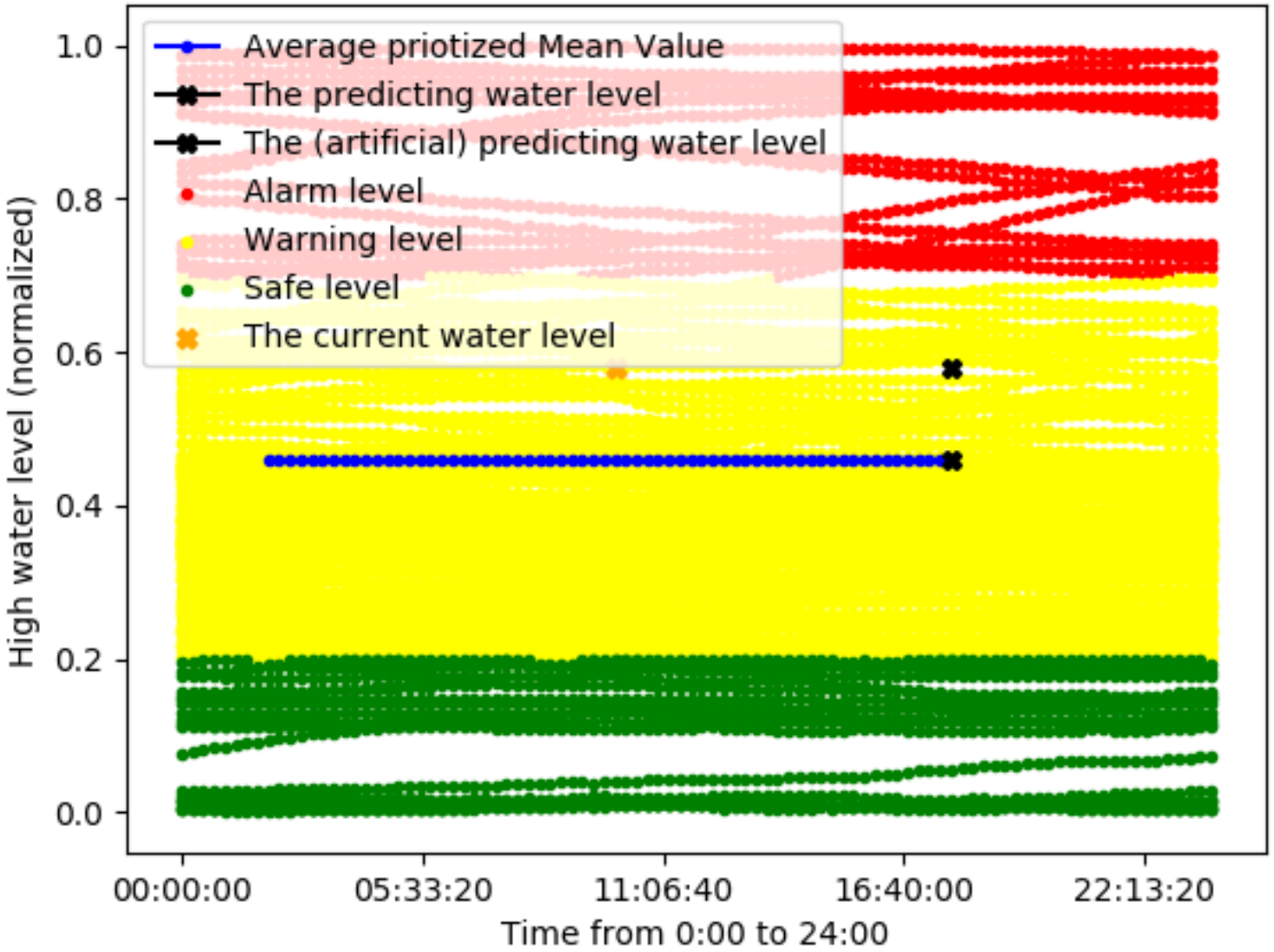
Approach 1: calculate at the time point given by user **(17043 water points)**

- Problem (issue): the classified (history) dataset need to get mean-value. At every time point there might exists at least one of 3 water levels

Solution

- 1. Reduce dataset at every single time point 15 min (from 0:00 to 24:00) by calculating combined coefficient mean-value w.r.t current water level
- 2. Transform the calculated mean-value of history data nearby close possibly to the value of current happening water level
- 3. Apply polynomial data (curve) fit method, e.g. (Linear) Least Square to the dataset
- 4. Predict water level based on the known equation and choose the best result among the calculated results (smallest error estimate)

Design solution & approach 1



Design Solution

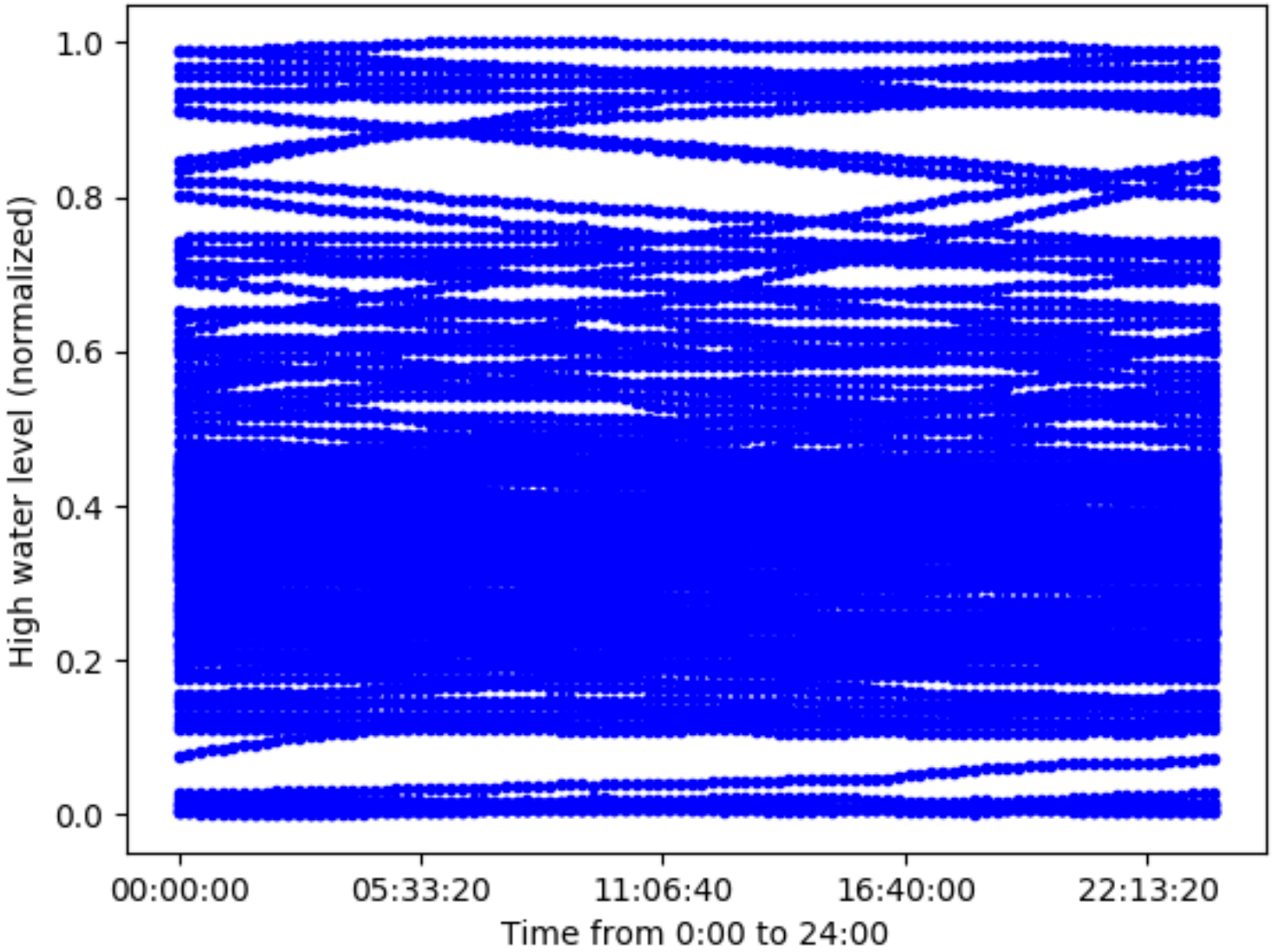
Approach 2: Find and Sort under certain conditions **(181 days)**

- calculate and sort out rate of change of all days from 0:00 to 24:00
- calculate and sort out all days with water level close possibly to the current water level given by user
- Visualization
- Decision can be made based on the result (probability to happen 'Worst Case' and 'Best Case') from the visualized history data

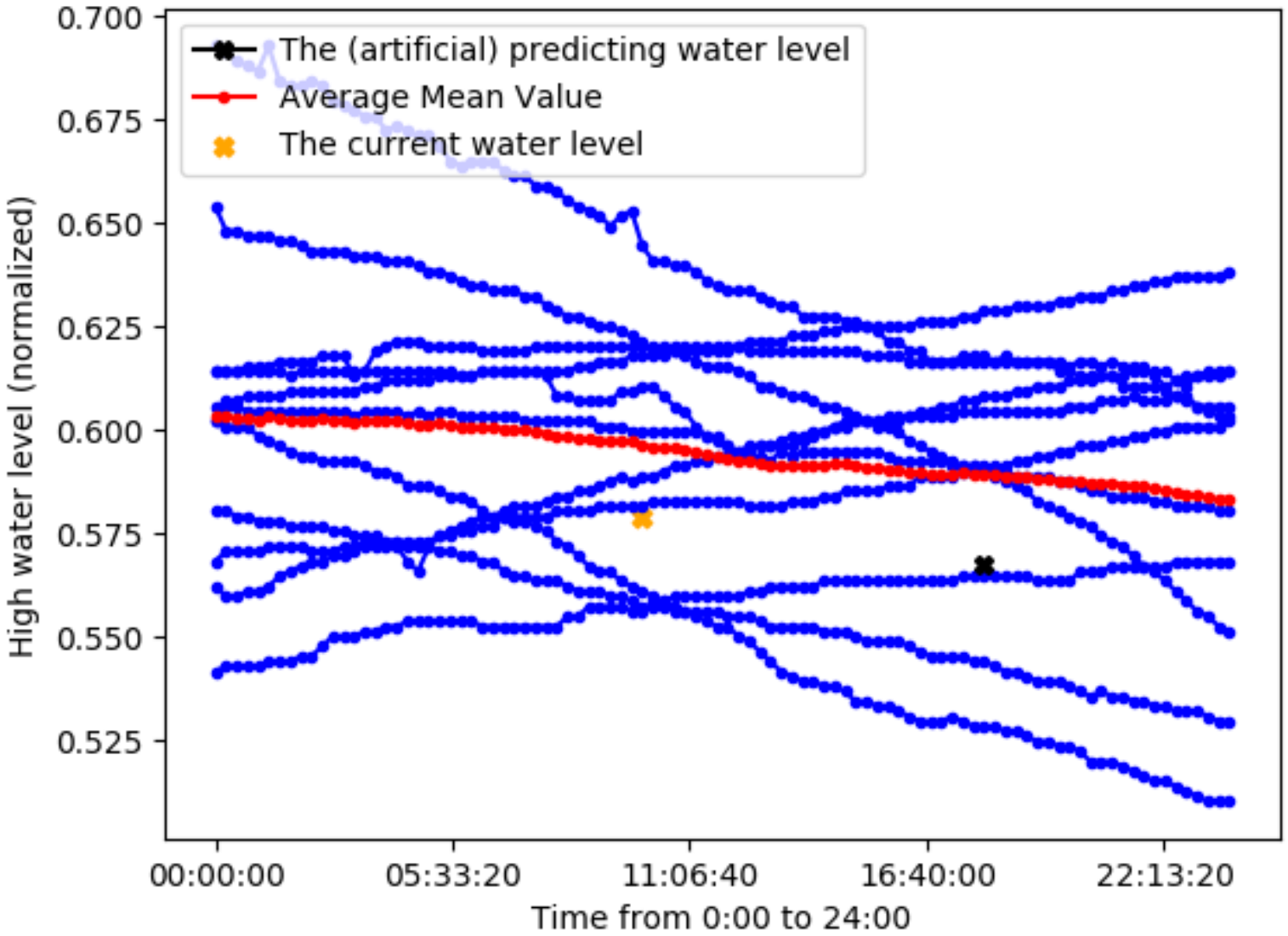
Remaining issues

- data need to be delivered in 96 points measured from 0:00 to 23:45
- However, data can be updated via Excel-file manually.

Design solution & approach 2



Design solution & approach 2



Conclusion

In order to get the best possibly decision

- User starts both solutions (via CLI)
- User should start again at later time with new input-information (current water level)
- User can improve the result easily by updating more data (via Excel-File) in the future

Opening points

- Try other methods e.g. Cubic spline smoothing, and ideas e.g. calculate path of water flow and water velocity, identify water intensity at different locations on map, etc.
- Performance: integrate this program into the system e.g. develop different UIs, build database system, resilient control system etc.

Techs

CLI

Input parameter:
- Current water level,
- Current time,
- Predict hours and choose methods

Output result:
- in logging file & print out
- Figures to visualize result

Applied Open Source Technologies

Python3 (Pandas, Numpy, Scipy, Matplotlib)
Linux (Windows)
Open Source License GPL-v2 on Github
https://github.com/lphan/utilkit/tree/master/floodpred

THANK YOU!

Weihnachtshochwasser 1993 in Koblenz