floodX: urban flash flood experiments monitored with conventional and alternative sensors

Matthew Moy de Vitry, Simon Dicht, and João P. Leitão

1Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland
2Institute of Civil, Environmental and Geomatic Engineering, ETH Zurich, 8093 Zurich, Switzerland

Correspondence to: Matthew Moy de Vitry (matthew.moydevitry@eawag.ch)

Received: 1 February 2017 – Discussion started: 23 February 2017
Revised: 24 July 2017 – Accepted: 25 July 2017 – Published: 4 September 2017

Abstract. The data sets described in this paper provide a basis for developing and testing new methods for monitoring and modelling urban pluvial flash floods. Pluvial flash floods are a growing hazard to property and inhabitants’ well-being in urban areas. However, the lack of appropriate data collection methods is often cited as an impediment for reliable flood modelling, thereby hindering the improvement of flood risk mapping and early warning systems. The potential of surveillance infrastructure and social media is starting to draw attention for this purpose. In the floodX project, 22 controlled urban flash floods were generated in a flood response training facility and monitored with state-of-the-art sensors as well as standard surveillance cameras. With these data, it is possible to explore the use of video data and computer vision for urban flood monitoring and modelling. The floodX project stands out as the largest documented flood experiment of its kind, providing both conventional measurements and video data in parallel and at high temporal resolution. The data set used in this paper is available at https://doi.org/10.5281/zenodo.830513.

1 Introduction

1.1 The need for comprehensive urban flood data

Urban pluvial floods, which occur when precipitation cannot be fully assimilated by a city’s drainage system, can cause substantial economic disruption and claim lives every year. In Switzerland, the damages caused by surface water runoff to urban infrastructure is around CHF 45 million per annum,1 and the figure is proportionally similar in the United Kingdom (Evans, 2004). Numerical modelling is an essential tool for urban flood risk assessment and flood risk management, allowing flood events to be simulated so that drainage system weaknesses can be identified and possible corrective solutions can be evaluated. Additionally, such models make it possible to create detailed forecasts and optimize drainage network operation with model predictive control (Raso and Malaterre, 2016). Currently, urban pluvial flood models are most often calibrated with monitoring data from sensors installed in the underground drainage network, if at all. This information, while sufficient for modelling urban drainage in normal weather conditions, does not directly inform on the situation above ground during flood events. The lack of surface flood information leaves model parameters such as ground roughness with large uncertainty (Hunter et al., 2008), and makes it difficult to estimate the predictive performance of the model (Dottori and Todini, 2013). This deficiency is regularly brought up in urban flood modelling research (Fewtrell et al., 2011; Hénonin et al., 2015; Sampson et al., 2012; Schmitt et al., 2004). Additionally, the lack of data limits the detail in which events can be modelled (Ciervo et al., 2015). To further illustrate the need for urban flood calibration data, Leandro et al. (2011) proposed a method for circumventing the issue by using detailed physical models to generate virtual overland flow data.

Because of the unfavourable conditions encountered at the street level, contactless measurement methods have an advantage when compared to conventional sensors. In particular, optical methods such as large-scale particle image ve-

---

1Etablissement Cantonal d’Assurance, Switzerland, personal communication, 2016

Published by Copernicus Publications.
locimetry (LSPIV) can be used to estimate flood discharges, and researchers have started leveraging social media and crowdsourcing to collect data for this purpose (Le Bourhiscaud et al., 2016; Le Coz et al., 2016; Dramais et al., 2011). The potential of such data will grow as the availability and pervasiveness of image-capturing devices such as smartphones, surveillance cameras, and unmanned aerial vehicles (Perks et al., 2016) increases.

In this context, the floodX project was launched to create a comprehensive urban flood data set in which conventional sensors are complemented with surveillance cameras in a controlled environment. Three key points make the floodX project stand out. First, it is the largest-scale controlled experiment examining the urban pluvial flood phenomenon (compare Fraga et al., 2015; Hakiel and Szydowski, 2016; Testa et al., 2007). Second, the experiment has a relatively high density of sensors providing information at major storage nodes and flow channels (Fig. 1) so a comprehensive picture of the flooding dynamics can be gained. Third, the sensors used are a combination of state-of-the-art conventional sensors, such as a magnetic-inductive pipe profiler and radar-based flow measurement, and novel sources of data (surveillance cameras) for monitoring water levels, overland flow velocities, and manhole overflows.

In summary, the floodX data will be used for two distinct but related areas of urban flood research: (i) automatic interpretation of image data into useful information for flood monitoring and model validation, and (ii) development of flood model calibration methods with overland flow data to improve the predictive power of the models.

1.2 Paper structure

The following sections contain information essential for understanding and using the data produced from the conducted flash flooding experiments. In Sect. 2, the hydraulic network of the experimental setup is presented and its general hydraulic behaviour is described. In Sect. 3, an overview of the sensor network is given, and in Sect. 4 the experimental procedure and data preparation are presented. In Sect. 5, the accuracy of the data with regard to temporal shifts, drifts, and anomalies is described. In Sect. 6, the organization, availability, and licences of the data are presented. In Sect. 7, an outline is given of the potential applications of the data for urban pluvial flood research.

2 Hydraulic network

The experiments took place at a flood response training facility (Fig. 1) at a military training area in the Canton of Bern, Switzerland. This facility was used for a period of 5 days, during which the whole experimental setup was installed, used, and dismantled. Water comes from a 450 m³ reservoir 10 m higher than the facility and is continually replenished by surrounding groundwater at low but stable rate.

The facility consists of a floodable area of around 500 m² with a maximum elevation difference of 2.9 m. It contains a small construction with a basement and has a configurable drainage system with multiple drainage points and manholes. The floodX Documentation package (Moy de Vitry et al., 2017a) contains the detailed construction plans of the facility.

The existing hydraulic network was configured in a way that would instigate dynamic behaviour and increase the response time of the facility, despite its limited size. To this end, sandbag walls were used to dam and channel water, and valves of the sewer network were configured so that the system would display the following phenomena relevant to urban flooding: indoor and outdoor ponding of water, drainage into sewer inlets, overflows of a surcharged manhole, and overland flow of one-dimensional and two-dimensional character. The resulting hydraulic network, including the reservoir, is depicted schematically in Fig. 2. The locations and characteristics of the hydraulic components can be found in the floodX Documentation package. The measurement of flow at the inlet and outlet pipes of the system (p1 and p6, Fig. 2) was a challenge because of the non-laminar flow conditions caused by a turn in the pipe or the presence of valve v4, respectively. For this reason, the flow meters were installed in specially designed pipe extensions as documented in the Supplement.

The flood experiments were conducted in the following way. Flooding is initiated when valve v1 is partially opened, allowing water to flow into the facility from reservoir s1 through pipe p1 into a small shaft s2. Rapidly, shaft s2 fills with water and overflows (weir w4) onto an open channel c1 that leads to a storage element s3. Storage s3 is caused by a dam constructed with sandbags (see Fig. 1), over which wa-
ter can spill at weir w1 and into an open channel c2. Two orifices (r1 and r2) drain storage s3 into manhole m1, from which the water can flow to manhole m2 through pipe p3. The natural exit of manhole m2 is pipe p4, but valve v2 at the extremity of pipe p4 was closed during the experiments, causing manhole m2 to overflow through its opening r3 into a small storage area s4. Storage s4 is also the outlet of channel c2, which originates from the dam overflow w1. Storage s4 drains through open channel c3 (well-defined channel walls) and channel c4 (no walls) into shaft s5, at the base of which there are two orifices. The first orifice (r4) leads out of the facility through pipe p6, valve v4 and pipe p8. The second orifice (r6) leads to a manhole (m3) in the basement of the small building through pipe p5, valve v3 and pipe p7. Because valve v4 was partially closed, water could build up in the exit shaft s5 and cause manhole m3 to overflow through its opening r5 and into the basement s6. These hydraulic components are graphically represented in Fig. 2 and their characteristics can be found in the floodX Documentation package.

3 Sensor network

A total of 21 sensor systems were installed to monitor flooding-relevant variables (Table 1). For readability, the complete list of sensors, including their mounting location and other relevant information can be found in the floodX Documentation package. The sensors used were not only state-of-the-art equipment used in urban drainage monitoring, but also security cameras and temperature probes for monitoring surface flooding. Naturally, some sensor systems were composed of multiple sensors – for example a Radar flow measurement system not only measures surface velocity but also water level. The naming convention used for data sources guarantees clear distinction between sensors (e.g. p1_q_mid_endress_logi designates data from pipe p1 representing discharge measured with a magnetic-inductive sensor produced by Endress+Hauser and logged with a Logitech data logger).

4 Experimental procedure and data preprocessing

4.1 Initial conditions of the hydraulic network

To ensure that manhole m2 would overflow, the valve v2 was closed for each experiment. Thus, to enter the basement, the water has to flow over the surface (channels c3 and c4) to the shaft s5 and flow back through pipes p5 and p7, and out the manhole m3. However, this also meant that water could not drain from pipes p3 and p4 at the end of a flood event. Most often, before the start of an experiment, valve v2 was opened and pipes p3 and p4 were allowed to drain. However, in a few cases, the water was intentionally left in pipes p3 and p4. The consequence of pipes p3 and p4 being full is that the system would respond much faster to inflow, thus resulting in manhole m2 overflowing very rapidly. Valve v4, which allows the whole facility to be drained, was fully opened between each experiment.

4.2 Experiment execution

Artificial flash floods were created by manually opening valve v1 at the entry of the flood facility. The hydrographs...
Table 1. Sensor systems installed for flood monitoring.

<table>
<thead>
<tr>
<th>Sensor system</th>
<th>Variables measured</th>
<th>Maximum temporal resolution</th>
<th>Number of sensor systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure sensor</td>
<td>Water level</td>
<td>10 s&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3</td>
</tr>
<tr>
<td>Temperature sensor</td>
<td>Manhole overflow</td>
<td>1 s</td>
<td>4</td>
</tr>
<tr>
<td>Ultrasonic distance sensor</td>
<td>Water level</td>
<td>200 ms</td>
<td>4</td>
</tr>
<tr>
<td>Radar system</td>
<td>Surface flow velocity and water level</td>
<td>5 s&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2</td>
</tr>
<tr>
<td>Ultrasonic flow sensor</td>
<td>Flow and pressure in partially filled and filled pipes</td>
<td>5 s&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Security camera</td>
<td>Surface flow velocity and water level</td>
<td>15–25 frames s&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>5</td>
</tr>
<tr>
<td>Magnetic-inductive pipe profiler</td>
<td>flow in filled pipe</td>
<td>1 s&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
</tr>
</tbody>
</table>

<sup>a</sup> The data recorded directly by the sensor data logger have a resolution of 1 min, but a greater temporal resolution can be acquired by interpreting images of the data logger screen. <sup>b</sup> Two systems were installed at the same location. One has a temporal resolution of 1 min and the other of 5 s. See floodX Documentation package for more information. <sup>c</sup> Data from one of the two sensors were omitted from the final data sets due to quality issues.

used for the experiments were based on simple step-like functions rather than realistic-looking hydrographs. There are several justifications for this. First, the valve controlling the inflow to the system had to be controlled manually and therefore the flow commands had to be simply defined. Second, well-defined shapes make the experiments more reproducible, and third, the presence of a high-storage node (s3) at the very start of the hydraulic network dampens high-frequency variability of flow into the system. The extra effort to produce such variability was therefore not justified.

The individual experiments (Table 2) were planned to be as expedient as possible, so as to optimize the experimentation time. For this reason, experiments were stopped after the main dynamics of the system stabilized, even though water often remained in some pipes as well as in the exit shaft s5. The experiments range from a scale of 3.5 to 64.1 m<sup>3</sup> discharge volume, and last between 6 and 23 min. The hydrographs of the flood events were defined through trial and error to produce a range of system responses (e.g. occurrence of dam overflow, flooding in basement). The environmental conditions in which the experiments were carried out include overcast skies, direct sunlight, windy conditions, and night so the robustness of video-based flood monitoring can be tested.

4.3 Variable quality of flood experiment data

The experiments conducted produced data with varying degrees of quality (Table 2). Three levels of experiment data quality have been defined in order to facilitate reuse of the data:

- Insufficient quality experiments have issues with sensor performance and experimental layout. These experiments are omitted from the preprocessed data sets. However, they are still available in the raw data and documented in the metadata provided in the floodX Raw Data, Metadata, and Preprocessing Code package (Moy de Vitry et al., 2017h).

- Monitoring quality experiments provide data that can be used for investigating novel flood monitoring approaches with video material. The data provided are correct but the hydraulic system (e.g. valve configuration) is not necessarily the same between experiments, and so the data are not suited for model calibration research. The data set for these experiments is available in the floodX Preprocessed Monitoring Data package (Moy de Vitry et al., 2017g).

- Calibration quality experiments are a subset of the monitoring quality experiments and can be used to calibrate and validate urban flooding and drainage models. The hydraulic system is the same for each experiment, only the hydrograph differs. The data set for these experiments is available in the floodX Preprocessed Calibration Data package (Moy de Vitry et al., 2017f).

4.4 Preprocessing of data

The data provided in the monitoring and calibration data sets have been preprocessed to facilitate reuse of the data. The raw data and the Python code used to perform the preprocessing are provided in the floodX Raw Data, Metadata, and Preprocessing Code package.

The following transformations were applied to preprocess the data: consolidation of multiple data files, chronological sorting, reformating of date, time, and null values, correction of temporal offsets, removal of extreme and impossible values, and segmentation of data into a separate file for each experiment. In order to reach higher sampling rates for certain sensors, images and videos<sup>2</sup> of the sensor displays were made and analysed with optical character recognition (OCR). To further enhance the usability of the data, the code is available on GitHub<sup>3</sup> and will be continually adapted as new ap-

<sup>2</sup>floodX Data Logger Images and floodX Data Logger Videos packages (Moy de Vitry et al., 2017b, c)
<sup>3</sup>https://github.com/Eawag-SWW/floodx_data_preprocessing
Table 2. Selection of high-quality experiments conducted, including the duration of the flooding, the total volume of water introduced in the system, and the experiment quality. The experiments are sorted by their total flood volume.

<table>
<thead>
<tr>
<th>No.</th>
<th>Duration</th>
<th>Volume (m$^3$)</th>
<th>Description</th>
<th>Lighting</th>
<th>Experiment quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>00:09:50</td>
<td>3.5</td>
<td>small event</td>
<td>night + infrared</td>
<td>calibration</td>
</tr>
<tr>
<td>14</td>
<td>00:07:00</td>
<td>3.8</td>
<td>small event</td>
<td>overcast</td>
<td>monitoring</td>
</tr>
<tr>
<td>35</td>
<td>00:15:00</td>
<td>3.8</td>
<td>small, empty sewers, manhole cover on m2</td>
<td>overcast</td>
<td>monitoring</td>
</tr>
<tr>
<td>16</td>
<td>00:05:00</td>
<td>4.0</td>
<td>small event, full sewers at start</td>
<td>direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>30</td>
<td>00:05:15</td>
<td>4.2</td>
<td>small event</td>
<td>night + flood light</td>
<td>calibration</td>
</tr>
<tr>
<td>26</td>
<td>00:07:00</td>
<td>4.8</td>
<td>small event, full sewers at start</td>
<td>evening sun</td>
<td>calibration</td>
</tr>
<tr>
<td>17</td>
<td>00:06:00</td>
<td>5.0</td>
<td>two small events, full sewers at start</td>
<td>direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>25</td>
<td>00:06:00</td>
<td>5.3</td>
<td>small event</td>
<td>evening sun</td>
<td>calibration</td>
</tr>
<tr>
<td>27</td>
<td>00:07:00</td>
<td>6.6</td>
<td>two small events</td>
<td>evening sun</td>
<td>calibration</td>
</tr>
<tr>
<td>15</td>
<td>00:07:00</td>
<td>7.0</td>
<td>two small events</td>
<td>overcast; direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>34</td>
<td>00:20:00</td>
<td>12.8</td>
<td>multiple small, slow events</td>
<td>overcast</td>
<td>calibration</td>
</tr>
<tr>
<td>31</td>
<td>00:16:00</td>
<td>24.2</td>
<td>simple dam overflow with pre-impulse</td>
<td>night + flood light</td>
<td>calibration</td>
</tr>
<tr>
<td>18</td>
<td>00:15:00</td>
<td>25.3</td>
<td>simple dam overflow</td>
<td>direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>21</td>
<td>00:17:00</td>
<td>26.5</td>
<td>simple dam overflow with pre-impulse</td>
<td>direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>28</td>
<td>00:23:00</td>
<td>27.6</td>
<td>simple dam overflow</td>
<td>evening sun</td>
<td>calibration</td>
</tr>
<tr>
<td>23</td>
<td>00:19:00</td>
<td>31.4</td>
<td>slow event</td>
<td>overcast; direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>33</td>
<td>00:20:40</td>
<td>32.2</td>
<td>large event followed by a small one</td>
<td>overcast</td>
<td>calibration</td>
</tr>
<tr>
<td>20</td>
<td>00:16:00</td>
<td>33.3</td>
<td>multiple large events</td>
<td>overcast; direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>22</td>
<td>00:20:40</td>
<td>34.5</td>
<td>large event followed by a small one</td>
<td>direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>19</td>
<td>00:15:00</td>
<td>44.9</td>
<td>multiple large events</td>
<td>direct sun</td>
<td>monitoring</td>
</tr>
<tr>
<td>32</td>
<td>00:22:00</td>
<td>64.1</td>
<td>multiple large events</td>
<td>night + infrared</td>
<td>calibration</td>
</tr>
</tbody>
</table>

5 Data accuracy

5.1 Temporal accuracy

For a majority of the sensors, an offset in the time records was observed and was corrected for in the preprocessed data sets of this paper. However, the following sensor systems have a more complex temporal offset that is worth mentioning.

5.1.1 Security cameras

The internal clocks of the cameras CAM1, CAM3, CAM4, and CAM5 displayed variable temporal offsets, whereas CAM2 was verified to be correctly synchronized. Because the camera recordings were started and stopped at the same moment in time (no more than a second), the relative misalignment of the camera’s internal clocks could be quantified by comparing the video file times of the recordings. Visual investigation of moments in the video material confirms the related misalignments. The offsets are provided in a text file of the floodX Flooding Videos package (Moy de Vitry et al., 2017e).

5.1.2 Overland flow sensors at channel c3 for water depth and velocity (c3_h_us_nivus and c3_v_radar_nivus)

A comparison of the video material from camera c3_cam3_instar with measurements from overland flow measurements from sensors c3_h_us_nivus and c3_v_radar_nivus revealed a discrepancy in the time of the logger of these two devices. The logger time is consistently found to be between 5 and 16 s ahead of reference time. Given the logging frequency of 5 s of the Nivus PCM Pro logger used for c3_h_us_nivus and c3_v_radar_nivus, it is assumed that the variability of the offset is largely due to sampling errors. A fixed offset of 12 s was assumed for preprocessing the data.

5.2 Sensor accuracy

5.2.1 Drift in ultrasonic water level measurements

The sensors s3_h_us_maxbotix, s5_h_us_maxbotix_2, and c3_h_us_nivus displayed instability or drift in their measurements. For example, at times when the dam s3 is empty, the value provided by s3_h_us_maxbotix varies in a slow and stable manner (Fig. 3). These ultrasonic sensors use the speed of sound, which is temperature-dependent, to estimate distance. The drift is possibly linked to direct solar radiation on the sensor body, which would raise the internal temperature of the sensor and cause it to overestimate the ambient air temperature for which there is a compensation. No corrective action was taken.
5.2.2 Other characteristics and anomalies of the data
- s1_h_us_maxbotix: waves on the surface of the water can be seen in the signal. The amplitude of these waves increases as the water level in the reservoir decreases. The data were not modified to suppress the waves.
- c3_v_radar_nivus: in the absence of flow and when a person walked near the measurement location, this sensor produced flow estimations. When flow is present, the passage of people on the crossing does not appear to affect the measurement. The false measurements have been removed from the data sets.
- p1_q_mid_endress_logi: outliers can be found in the flow data logged by the webcam. These data values were read using optical character recognition from images of the flow meter display. Invariably, there are errors in the interpretation, especially when the image was taken at the same moment at which the value was being updated on the display. The most critical outliers have been removed from the data but some remain.
- s6_h_us_maxbotix: a curious artefact is visible after flooding occurs in the basement (Fig. 4). For a few minutes, the water level appears to be negative, before returning to normal. No corrective action was taken.

5.3 Data omitted from data sets
Over all experiments, the discharge measured by the ultrasonic discharge sensor p6_q_us_nivus at the facility outlet deviates substantially from the discharge measured at the inlet of the flood facility in pipe p1. Different hypotheses to explain the volume differences were brought forward, such as residual water in the system and pipe leaks, but investigation of the data and the facility plans led to the conclusion that such factors could not fully explain the discrepancy. The only remaining explanation is that the constrained measurement conditions, especially the short stabilization distance before and after the sensor, the presence of a valve at the end of the pipe, and the frequent regime changes between full and partially filled pipe, caused the measurement to be erroneous. This conclusion is corroborated by the discovery of artefacts in the discharge data. Since the volume differences are larger than the expected measurement error for the technology (DWA, 2011), the data were judged to be of insufficient quality and were therefore removed from the data sets.

5.4 Incomplete control of facility infrastructure
In the data, the water level in manhole m2 can be seen sinking (from 17:42 in Fig. 5), despite there being no expected outlet, since all valves are supposed to be closed. When modelling the facility, leakages in the pipes and valves must be represented.
5.5 Gaps in the data

Equipment manipulation errors and data management issues led to a few gaps in the data. This has been observed for the Nivus radar system mounted in channel c3 in experiment 7, sensors m1_h_p_endress and m2_h_p_endress for experiments 14 to 23, and the temperature sensors in experiment 24. This information is also available in the experiment metadata file.

5.6 Discussion

The floodX experiments illustrated the importance of certain aspects of experimental design and execution. In our case, certain issues could have been avoided if the following points were given more attention:

- Synchronization of data loggers must be performed manually and at an appropriate temporal resolution.
- Environmental disturbances like wind and solar radiation, which can influence experiment execution and sensor performance, should be anticipated and planned for.
- Checklists should be used to ensure that equipment is functioning correctly before each experiment.
- Performance of data loggers can be as important as that of sensors, depending on the timescale of an experiment’s dynamics.

6 Potential applications of data sets

Monitoring data for historic urban pluvial floods is typically limited to the (underground) drainage network because most sensors are designed specifically for that setting. Overland flow and accumulation is of much more interest than the drainage network when modelling urban pluvial floods, but the lack of suitable sensors means that flood hydrologists must often calibrate and validate their models with very limited or partial information. The data collected in the floodX project will be used to develop and investigate both image-based flood monitoring methods and flood model calibration schemes that can assimilate non-standard overland flow data. The tools developed in these two lines of research are necessary for the long-term vision of utilizing social media images and surveillance videos from real flood events to obtain more reliable flood models.

A trove of overland flow information lies in existing surveillance infrastructure and social media in the form of images and videos. With appropriate processing methods, flow and water depth could be automatically estimated from these data, and the floodX data sets provide an opportunity to research such methods. For example, the measurement of shallow overland flows with large-scale particle image velocimetry (LSPIV) could be investigated in channel c3 with the two cameras and two radar systems (Fig. 6). For the moment, LSPIV has been investigated in urban settings only for large flows and without direct validation data (Guillén et al., 2017; Perks et al., 2016), or for seeded flows (Branislavljević and Prodanović, 2006). Another example is the use of deep learning to estimate flood water levels through semantic scene interpretation, e.g. by interpreting the immersion level of objects of known dimensions in snapshots and videos (Fig. 7).

The floodX data can also be used for urban pluvial flood model calibration research since the flood monitoring setup was contained within a hydraulic system comparable to an urban catchment. Thanks to this unique setup and the sub-

4 © floodX Raw Data, Metadata, and Preprocessing Code package

5 © floodX Preprocessed Monitoring Data, floodX Flooding Videos, floodX Flooding Images (Moy de Vitry et al., 2017d).
Table 3. Data packages associated with the floodX project.

<table>
<thead>
<tr>
<th>Package name</th>
<th>DOI</th>
<th>License</th>
</tr>
</thead>
<tbody>
<tr>
<td>floodX Raw Data, Metadata and Preprocessing</td>
<td><a href="https://doi.org/10.5281/zenodo.830505">https://doi.org/10.5281/zenodo.830505</a></td>
<td>Creative Commons Zero 1.0*</td>
</tr>
<tr>
<td>floodX Preprocessed Monitoring Data</td>
<td><a href="https://doi.org/10.5281/zenodo.830511">https://doi.org/10.5281/zenodo.830511</a></td>
<td>Creative Commons Zero 1.0</td>
</tr>
<tr>
<td>floodX Preprocessed Calibration Data</td>
<td><a href="https://doi.org/10.5281/zenodo.830513">https://doi.org/10.5281/zenodo.830513</a></td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>floodX Documentation</td>
<td><a href="https://doi.org/10.5281/zenodo.830506">https://doi.org/10.5281/zenodo.830506</a></td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>floodX Flooding Videos</td>
<td><a href="https://doi.org/10.5281/zenodo.830451">https://doi.org/10.5281/zenodo.830451</a></td>
<td>Creative Commons Zero 1.0</td>
</tr>
<tr>
<td>floodX Flooding Images</td>
<td><a href="https://doi.org/10.5281/zenodo.830501">https://doi.org/10.5281/zenodo.830501</a></td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>floodX Data Logger Videos</td>
<td><a href="https://doi.org/10.5281/zenodo.830504">https://doi.org/10.5281/zenodo.830504</a></td>
<td>Creative Commons Zero 1.0</td>
</tr>
<tr>
<td>floodX Data Logger Images</td>
<td><a href="https://doi.org/10.5281/zenodo.830507">https://doi.org/10.5281/zenodo.830507</a></td>
<td>Creative Commons Zero 1.0</td>
</tr>
</tbody>
</table>

* The preprocessing code included in this package is under the MIT license.

set of calibration-quality data, it will be possible to test model calibration concepts capable of assimilating the overland flooding data delivered by the novel monitoring methods. Specific questions that need to be addressed include the choice of appropriate objective functions and of an appropriate weighting strategy for multiple objective functions.

In summary, the research made possible by the floodX data will both contribute to urban flood monitoring innovation and improve the reliability of urban flood modelling, thereby increasing the effectiveness of urban flood management services such as flood forecasting, response, and risk management.

7 Data availability

Eight data packages are provided through the Zenodo data repository as described below, in open-access and with liberal licences (see Table 3 for a summary). It is important to note that while the time series data have been preprocessed, the videos and images of the flooding are provided as is. A brief description of the data packages follows.

floodX Raw Data, Metadata, and Preprocessing Code

This package contains all collected data that is available in a text format, as well as useful code for preprocessing data.

– metadata: this directory contains metadata for the sensors used, the data sources (sensor + location + data logger), flooding experiments and image files for optical character recognition.

– data_raw: this directory contains the raw text format data, organized by sensor.

– code: this directory contains processing code for making the raw data more usable for visualization and modelling, including the code for interpreting sensor readouts with OCR. The code is maintained in the following repository: https://github.com/eaw-sww/floodx_data_preprocessing.git.

– data_ocr_result: this directory contains data read from sensor display images using OCR.

floodX Preprocessed Monitoring Data

This package contains a preprocessed subset of the flooding experiments for which the data quality is appropriate for urban flood monitoring research. The package includes text-based data only – videos and images of flooding can be found in the floodX Flooding Videos and floodX Flooding Images packages, respectively.

floodX Preprocessed Calibration Data

This package contains a preprocessed subset of the flooding experiments for which the data quality is appropriate for urban flood model calibration and validation research. The package includes text-based data only – videos and images of flooding can be found in the floodX Flooding Videos and floodX Flooding Images packages, respectively.

floodX Flooding Videos

This package contains archives of videos of the flooding taken with surveillance cameras. The videos are grouped by camera and by recording sessions.

floodX Flooding Images

This package contains pictures taken during the flooding experiments. The cameras used to take the pictures had shifted timestamps, so a text file provides the time shifts of the cameras to the reference times.

floodX Documentation

This package contains an archive with material docu-
menting the flood facility and the sensors that were used in the experiments.

- Facility construction plans (provided as is, without any guarantee for geometric accuracy).
- Plan of floodable area, including hydraulic components and sensors.
- Information regarding individual sensors and dimensions of hydraulic elements such as pipes and storage elements.
- Photos of the experimental layout.

**floodX Data Logger Images**

This package contains archives of pictures taken of sensor displays in order to record measurements at a higher frequency than certain data loggers allowed. The images for the pressure data is stored in multiple archive files that must be assembled during unpacking. The images are grouped together because sometimes the camera had to be moved between experiments, thereby changing the position of the display in the image. Each image group contains a settings file that indicates the location of the display(s) in the images of the group. With this information, the sensor reading can be automatically extracted from the images.

**floodX Data Logger Videos**

This package contains videos that constitute an alternative to the data logger images. The videos provide high-quality readings of the logger displays, and values can be read more frequently. The image quality may be superior to that of the data logger images.

## 8 Conclusions

The flood experiments described in this paper stand out from similar urban flood experiments thanks to the relatively controlled conditions and the diversity of sensor systems involved. While not void of blemishes, the data generated in the floodX project hold significant potential to support urban flooding research, especially for the exploration of alternative measurement strategies for the quantification of urban flood phenomena. In particular, the published data sets constitute a valuable starting point for investigating large-scale particle image velocimetry in urban environments and computer-vision-based water level estimation. Indeed, the need for flood monitoring data has often been pointed out by researchers, and alternative data sources could be the key to providing such data. Additionally, thanks to the large number of experiments, the data sets provide a unique opportunity to explore methods for calibrating urban flood models with alternative data sources. In the face of climate change, the potential improvements to urban flood monitoring and modelling expedited by these data will contribute to the integrity of urban infrastructure and the populations that rely on them.

The Supplement related to this article is available online at https://doi.org/10.5194/essd-9-657-2017-supplement.

**Author contributions.** Matthew Moy de Vitry was lead for the project design and execution, as well as for drafting the paper. Simon Dicht was the lead technician, responsible for the acquisition and installation of instrumentation. João P. Leitão was principal investigator of the project, providing valuable support in the orientation and coordination of project execution and paper drafting. All authors were involved in reviewing the paper.

**Competing interests.** This project was financed by the Swiss National Science Foundation under grant #169630. The authors declare that they have no conflict of interest.

**Acknowledgements.** The authors wish to thank the handling editor and the three anonymous referees, whose detailed comments greatly improved the quality of this paper. The authors would like to thank the following people for their valuable counsel for conceptual design of the experiments: Tobias Doppler, Frank Blumensaat, Andreas Scheidegger, and Kris Villez. Additionally, great assistance was provided in execution of the experiments by Christian Ebi, Alex Hunziker, Lena Mutzner, Joerg Rieckermann, Andreas Scheidegger, Luis M. de Sousa, and Omar Wani. The project was made possible thanks to the cooperation of armasuisse and the training village facility managers Roland Rickli and Michael Gehrig. Furthermore, the authors thank Stebatec AG and Nivus AG for their collaboration in providing sensor equipment.

Edited by: David Carlson
Reviewed by: three anonymous referees

**References**


Dramais, G., Le Coz, J., Camenen, B., and Hauet, A.: Advantages of a mobile LSPIV method for measuring flood discharges and


