



## High-quality observation of surface imperviousness for urban runoff modelling using UAV imagery

P. Tokarczyk<sup>1</sup>, J. P. Leitao<sup>2</sup>, J. Rieckermann<sup>2</sup>, K. Schindler<sup>1</sup>, and F. Blumensaat<sup>2,3</sup>

<sup>1</sup>Institute of Geodesy and Photogrammetry, ETH Zurich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland

<sup>2</sup>Swiss Federal Institute of Aquatic Science and Technology, Eawag, Überlandstrasse 133, 8600 Dübendorf, Switzerland

<sup>3</sup>Institute of Environmental Engineering, Chair of Urban Water Systems, ETH Zurich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland

Correspondence to: P. Tokarczyk (piotr.tokarczyk@gmail.com)

Received: 17 December 2014 – Published in Hydrol. Earth Syst. Sci. Discuss.: 27 January 2015

Revised: 23 September 2015 – Accepted: 27 September 2015 – Published: 21 October 2015

**Abstract.** Modelling rainfall–runoff in urban areas is increasingly applied to support flood risk assessment, particularly against the background of a changing climate and an increasing urbanization. These models typically rely on high-quality data for rainfall and surface characteristics of the catchment area as model input.

While recent research in urban drainage has been focusing on providing spatially detailed rainfall data, the technological advances in remote sensing that ease the acquisition of detailed land-use information are less prominently discussed within the community. The relevance of such methods increases as in many parts of the globe, accurate land-use information is generally lacking, because detailed image data are often unavailable. Modern unmanned aerial vehicles (UAVs) allow one to acquire high-resolution images on a local level at comparably lower cost, performing on-demand repetitive measurements and obtaining a degree of detail tailored for the purpose of the study.

In this study, we investigate for the first time the possibility of deriving high-resolution imperviousness maps for urban areas from UAV imagery and of using this information as input for urban drainage models. To do so, an automatic processing pipeline with a modern classification method is proposed and evaluated in a state-of-the-art urban drainage modelling exercise. In a real-life case study (Lucerne, Switzerland), we compare imperviousness maps generated using a fixed-wing consumer micro-UAV and standard large-format aerial images acquired by the Swiss national mapping agency (*swisstopo*). After assessing their overall accuracy, we perform an end-to-end comparison, in which they are used

as an input for an urban drainage model. Then, we evaluate the influence which different image data sources and their processing methods have on hydrological and hydraulic model performance. We analyse the surface runoff of the 307 individual subcatchments regarding relevant attributes, such as peak runoff and runoff volume. Finally, we evaluate the model's channel flow prediction performance through a cross-comparison with reference flow measured at the catchment outlet.

We show that imperviousness maps generated from UAV images processed with modern classification methods achieve an accuracy comparable to standard, off-the-shelf aerial imagery. In the examined case study, we find that the different imperviousness maps only have a limited influence on predicted surface runoff and pipe flows, when traditional workflows are used. We expect that they will have a substantial influence when more detailed modelling approaches are employed to characterize land use and to predict surface runoff. We conclude that UAV imagery represents a valuable alternative data source for urban drainage model applications due to the possibility of flexibly acquiring up-to-date aerial images at a quality compared with off-the-shelf image products and a competitive price at the same time. We believe that in the future, urban drainage models representing a higher degree of spatial detail will fully benefit from the strengths of UAV imagery.

## 1 Introduction

In the last century we have witnessed increased migration of people from rural areas to cities. Today, the majority of the human population live in cities, and this number is estimated to grow constantly and reach a level of 60 % (UN, 2013). The process of rapid urbanization required developing an infrastructure capable of coping with a constantly increasing number of its users. Accordingly, ensuring water supply for the people is important, but due to the increased hydrological extremes induced by climate change (Hirabayashi et al., 2013; Hall et al., 2014; Rojas et al., 2013), being able to safely direct stormwater away from populated areas, in order to avoid flooding, is not least a challenging task. It requires predicting the hydraulic behaviour of the given drainage infrastructure using reliable hydrological models (Arrighi et al., 2013). Those models, apart from detailed rainfall information, call for surface characteristics such as imperviousness.

Impervious surfaces reduce the infiltration of water into the soil. They can be directly related to a level of urbanization (Stankowski, 1972), because in urban environments, impervious surfaces dominate (e.g. rooftops or roads). Monitoring of the imperviousness level is important as it directly impacts many environmental processes. An increasing percentage of impervious surfaces increases surface runoff volume and peak discharge, and decreases soil moisture compensation and groundwater recharge. Moreover, increased peak runoff volumes together with an inefficient drainage network can not only cause urban floods, but also lead to an increased hydraulic stress and increase the risk of loading waterbeds with sediments, and its associated constituents (e.g. nutrients, contaminants and micro-pollutants).

An important step towards automation of the processes applied to map impervious areas was made as a consequence of remote sensing sensors and classification techniques development (for a detailed review of remote sensing methods used to map imperviousness, please refer to the Supplement). In general, most of the studies on extraction of impervious surfaces from remote sensing data focused on satellite images. During the last decade, a rapid improvement of imaging sensors gave the end-user an access to very high spatial resolution (VHR) imagery<sup>1</sup>. Satellite sensors like Ikonos (Chormanski et al., 2008) and QuickBird (Zhou and Wang, 2008) or VHR aerial images (Fankhauser, 1999; Nielsen et al., 2011) were quickly adopted for impervious surfaces mapping. Some studies suggest using highly accurate methods to quantify landscape changes (land-use and land-cover) using multi-sensor approaches (Forzieri et al., 2012a, b). In the context of urban hydrology, Ravagnani et al. (2009) attempted to use impervious surfaces extracted from VHR satellite and aerial imagery as an input to the urban drainage model, but they did not analyse pipe flow predictions, focus-

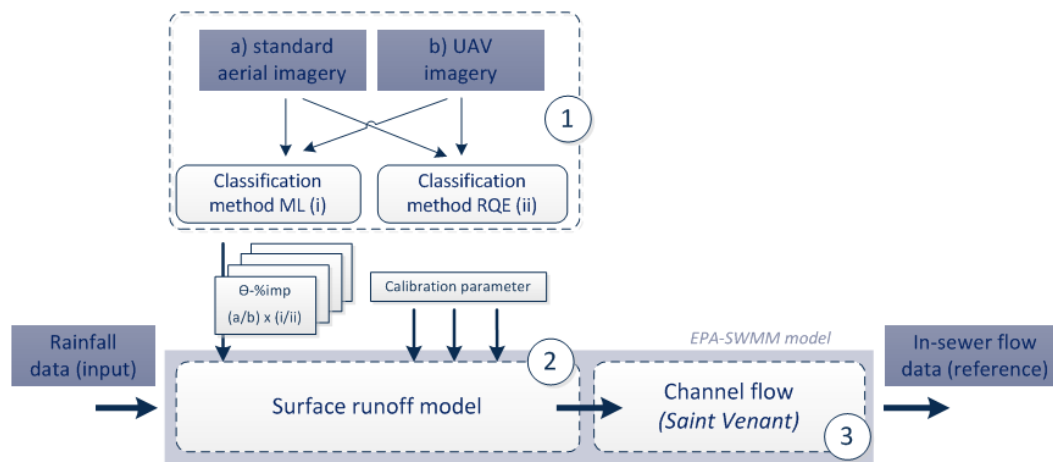
ing only on the surface runoff component. However, modern urban drainage modelling methods call for up-to-date and detailed input data, which could also be acquired in an efficient way. Even though VHR satellite images able to acquire fine-grained image information (WorldView-3 satellite can achieve up to 0.31 m GSD in panchromatic channel) and have short revisit periods, are still expensive and vulnerable to cloud cover. VHR aerial imagery on the other hand, although being able to acquire very detailed imagery, is usually being updated at most once a year, but usually every third year (swisstopo, 2010). Recently, imaging platforms based on UAVs became very popular, finding their application in the fields of photogrammetry, archeology or agriculture (Sauerbier and Eisenbeiß, 2010; Eisenbeiß, 2009; Zhang and Kovacs, 2012). More recently, Leitão et al. (2015) investigated the quality of digital elevation models (DEMs) generated using UAV imagery from urban drainage modelling applications. In the study, the authors show that the quality of UAV DEMs is comparable to that of conventional, off-the-shelf height data sets. However, to our best knowledge no studies exist, that used UAV-based imagery to extract imperviousness information, and to use it in the field of urban drainage modelling. In comparison to a standard, off-the-shelf satellite or aerial remote sensing imagery, UAVs demonstrate greater flexibility and are more efficient in terms of money and time. Yet, the classification of UAV VHR imagery, particularly in urban areas, is challenging, because in this level of detail, many small objects appear, and fine-grained texture details of larger objects emerge. Thus, describing an object class using only single raw pixel values is insufficient. Accurate classification needs additional image features, which would characterize the contextual information by describing an object's local neighbourhood. The value of such approach in classification of surface imperviousness has already been acknowledged (Moser et al., 2013). However, what is highly relevant, but currently unclear, is how to best exploit the rich information, i.e. the unprecedented level of detail and flexibility to acquire problem-specific images. And, whether it is feasible to use imagery acquired with UAVs for urban drainage modelling.

Specifically, we present three key aspects:

1. We evaluate whether land-use data based on UAV imagery can be used to assess the performance of urban drainage systems.
2. We propose a unique workflow based on a randomized quasi-exhaustive (RQE) feature bank and a boosting classifier<sup>2</sup>. The RQE feature bank consists of a multitude of multiscale textural features describing both, spectral and height information (Tokarczyk et al., 2015). The boosting classifier lends itself to the task to only

<sup>1</sup>We refer to a VHR image when sensor's ground sampling distance (GSD) is lower than 1 m.

<sup>2</sup>The boosting classifier used in conjunction with RQE features will be referred to as the "RQE method" in this paper.



**Figure 1.** Overall analysis approach ( $\Theta$ -%imp: model parameter “degree of imperviousness”; ML: maximum likelihood; RQE: boosting with randomized quasi-exhaustive feature bank).

choose the optimal features during training (for details, see below).

3. We perform end-to-end comparison of land use against high-quality sewer pipe flow data. Although important to correctly interpret the results, this is not routinely done in remote sensing literature.

The key idea of our study was not to solely base the assessment of the usefulness of UAV images for urban drainage applications on the performance of the classifiers. Thus, we demonstrate the usefulness of our approach by means of a case study in a small urban area in Lucerne, Switzerland, in two steps (see also Fig. 1): first, we compare the UAV data with standard airborne imagery using a maximum likelihood (ML) classifier and the RQE method on both image sources (1). Second, we use a hydrodynamic model to show the consequences of different land-use information for urban drainage performance indicators, here surface runoff (2) and in-sewer pipe flow (3).

The remainder of the paper is structured as follows: first we present a general approach and the case study catchment with related material, such as the hydrodynamic rainfall–runoff model, rainfall and runoff observations, and remote sensing data. Then we describe the applied methods, land-use classification, surface runoff and in-sewer flow modelling, as well as the suggested performance criteria. Finally we present results and discuss the potential and limitations of using UAV images in urban hydrology.

## 2 Materials and methods

### 2.1 Case study and data sets

#### 2.1.1 Case study

For our case study we used a residential area, called the Wartegg catchment, in the city of Lucerne, Switzerland (see Fig. 2). The catchment covers about 77 ha and is home for 6900 residents. It is typical of many suburban areas in Switzerland: high- to moderate-density population, and scattered single- to two-story housing embedded in a hilly landscape, including typical public infrastructure such as shopping centres and sports grounds.

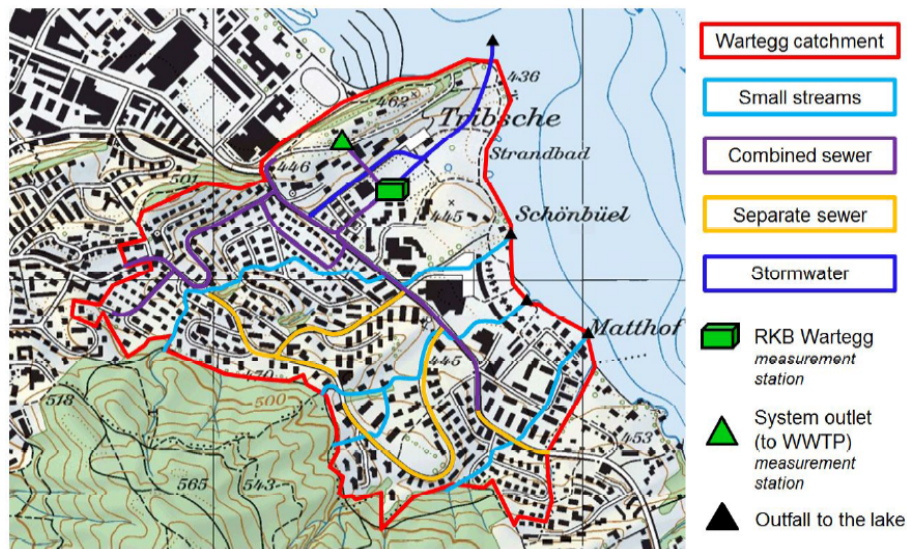
Stormwater and wastewater are drained by separate and combined sewers (see Fig. 2) with a total length of 11.2 km. An overflow structure connected to a small storage basin is installed to avoid hydraulic overload in case of heavy rainfall. Excess combined sewage is directly discharged to the lake; the carry-on flow travels by gravity to the wastewater treatment works. Three small creeks, to some extent culverted, cross the catchment and are interlinked with the stormwater network.

#### 2.1.2 Remote sensing data sets

##### Image data

In this study we used two image data sets. The first image data were acquired by *swisstopo*<sup>3</sup> in June 2013. It is a part of an aerial orthophoto mosaic (RGB channels) with a GSD of 0.0625 m, and consists of images acquired during leaves-on conditions. Although this data set was acquired on-demand (standard *swisstopo* orthophotos have a GSD of

<sup>3</sup>In this paper the “ortho” and “orthophoto” terms will be used interchangeably with *swisstopo* imagery.



**Figure 2.** Case study catchment area situated in Lucerne.

0.25 m), images acquired by *swisstopo* are publicly available, and this data source is, to our best knowledge, the standard for hydrological applications in Switzerland. Because *swisstopo* offers off-the-shelf image products, which are already orthorectified and georeferenced, one can avoid costly and time-consuming pre-processing of raw image data. On the other hand, image acquisitions are made at most once a year, usually every third year, and try to alternate between leaves-on and leaves-off periods (*swisstopo*, 2010). Thus, it might happen that one is not able to obtain up-to-date results.

The second data set was acquired with a Canon IXUS 127 HS digital consumer camera with 16Mpix sensor, mounted on a fixed-wing micro-UAV platform (Sensefly eBee; see Sect. S2 in the Supplement for details). The flight was performed during leaves-off conditions in March 2014. The custom processing software, which is shipped together with the UAV (cf. <http://www.senseFly.com>, based on the Pix4D technology, cf. <http://pix4d.com/products/>), was used to process the images. It is designed for use by non-experts and is highly automated; user interaction is limited to selecting input images, entering flight parameters (camera details and GPS/INS data) and measuring ground control points (GCPs). Orthophotos (RGB channels) generated from the acquired images have a GSD of 0.10 m. In the case of a small catchment, as in our study, a main advantage of UAVs, when compared to manned aircraft with large-format mapping cameras, lies in their flexibility in terms of deployment, and in their low cost. Conducting a standard photogrammetric flight campaign typically requires days of preparation and is more dependent on weather conditions. Note though: micro-UAVs are at present not suitable for large-area mapping, because of their low speed and limited battery capacity.

Prior to the classification, both data sets were downsampled to a GSD of 0.25 m in order to make the evaluation comparable to standard *swisstopo* imagery available on the market. Furthermore, this step reduces the time needed for training the classifier.

### Height model

In this study we used two different height models: (i) a DTM provided by *swisstopo* (*swisstopo*, 2014) was used to classify the *swisstopo* data set and to derive the catchment slope for the urban drainage model. This model features a grid size of 2 m, and for the land-use classification it was upsampled to the resolution of corresponding image data set; (ii) a nDSM<sup>4</sup>, created by subtracting a DSM extracted using dense image matching from a DTM provided by *swisstopo*, was used to classify the UAV data set.

#### 2.1.3 Rainfall

Precipitation data were collected from a meteorological station located 2 km away from the Wartegg catchment area, operated by the Swiss Meteorological Institute (MeteoSwiss). Recordings were taken in a 10 min interval using a tipping bucket rain gauge with a precision of 0.1 mm. Hourly precipitation was checked following the quality assurance criteria of MeteoSwiss. Additional quality checks were carried out to ensure that the 10 min data are reliable. Spatial rainfall vari-

<sup>4</sup>A digital terrain model (DTM) represents the bare ground surface; a digital surface model (DSM) represents the surface visible from the top, including buildings, trees etc.; the normalized digital surface model (nDSM) is obtained by subtracting the DTM from the DSM and shows the relative height of non-ground objects over the ground.

ability was not considered in the study due to the short distance between the meteorological station and the study area.

#### 2.1.4 Sewer flow reference data

Two flow data sets were obtained from in-sewer flow monitoring located at the outlet of the subcatchment (see Fig. 2). Over a period of 4 months (17 July 2014 to 18 November 2014) the in-sewer flow was continuously monitored with two different sensors: (i) an acoustic Doppler flow sensor (Sigma submerged AV sensor, HACH) – 1 min monitoring frequency – and (ii) a digital Doppler radar velocity sensor, along with ultrasonic level-sensing (FLO-DAR, Marsh Mc Birney) – 15 min monitoring frequency – to provide redundant flow rate information. Correlation analysis between the two reference signals shows a high agreement and confirms the solid quality of the data.

#### 2.1.5 Urban drainage model

Urban drainage models are tools to analyse the hydraulic behaviour of urban drainage systems, and to support risk analysis of urban flooding and receiving water pollution. Typically, these models include two main computing modules: the surface runoff (hydrological) and the in-sewer flow (hydraulic) model. The hydrological model estimates the time and space distribution of the direct runoff under consideration of initial precipitation losses (evaporation, wetting losses) and soil infiltration for pervious areas. The resulting runoff is then used as input for the hydraulic model to simulate the pipe flow in the sewer network.

In the present study we use the freely available Stormwater Management Model released and constantly developed by the US Environmental Protection Agency (SWMM, Release 5.1.006; US-EPA, 2010). SWMM is a widely used and well-accepted state-of-the-art 1-D dynamic rainfall–runoff model. We deliberately chose SWMM despite its limitations (lumped surface runoff model concept) as it represents a widely used standard application in urban drainage modelling, and we wanted to keep the modelling use case as simple as possible.

The description of the surface runoff is based on the MANNING approach, a simplifying, conceptual formulation of transport phenomena in the catchment assuming that the surface runoff starts after the rainfall volume has exceeded a representative value of the initial losses in the catchment. Rainfall losses are adjusted throughout the rainfall event according to the changes occurring in the infiltration process (pervious part of catchment surface) which is a function of the soil water saturation level. Impervious surfaces are those where no infiltration occurs; the catchment’s imperviousness degree and its spatial distribution are then expected to have a great impact on surface runoff and urban drainage system modelling results. Flow routing through a system of sewer pipes, storage basins and regulating devices is accomplished

by solving the Saint-Venant flow equations, whereas here we applied a type of diffusive wave approximation which neglects inertial terms from the momentum equation when flow becomes supercritical.

## 2.2 Methodology

### 2.2.1 Classification

Generally, supervised classification consists of three main steps: (i) extraction of the features from a raw input image, (ii) training the classifier using a small, manually annotated training set (not necessarily from the same image), and (iii) classification of all pixels in the area of interest, using the classifier trained in the previous step. In the following we describe two different types of supervised classifiers: (i) Gaussian maximum likelihood, and (ii) boosting.

#### Maximum likelihood

The maximum likelihood (ML) classifier is a popular classification method in the field of urban hydrology. It is a simple generative model which assumes that the image features within each target class follow a normal distribution. Under this assumption, each of the target classes can be described by its mean vector and covariance matrix. Given this information one can directly compute the statistical probability of particular pixel belonging to one of the target classes. A serious limitation of ML is that it is not well suited for high-dimensional data. Due to the “curse of dimensionality” (Hughes, 1968), its performance degrades typically beyond a dozen or so feature dimensions. For imagery with a medium spatial resolution imagery, where objects are usually spectrally consistent<sup>5</sup>, it might be enough to construct image features consisting only of single raw pixel values. However, the variability of the pixel values within an object class grows with the spatial resolution of the image, for example, when a roof consists of many pixels and substructures such as planted areas or roof gardens become visible. Therefore one should no longer rely on single pixel values, but has to consider contextual information and, for example, construct features that exploit the neighbourhood of a pixel (e.g. textural features). Such features expand the dimensionality of data, making generative classifiers inefficient. Here we classified two image data sets using a maximum likelihood classifier implemented in ArcGIS software (ESRI, 2013). As often done in conjunction with the ML method, we use only the single raw pixel values as features.

#### Boosting

As an alternative to ML we propose a multiclass extension (Benbouzid et al., 2012) of adaptive boosting (AdaBoost, Freund and Schapire, 1995). Unlike ML, boosting methods

<sup>5</sup>Meaning that they consist of pixels of similar values.

(and related discriminative classifiers) are better suited for very high-dimensional feature spaces, as they do not attempt to model the input data distribution. Boosting greedily learns an additive combination of many simple classifiers (in our case shallow decision trees). A useful property of the method is that it performs explicit feature selection as part of the classifier training. Thanks to this, we sidestep manual feature selection. Moreover, at test time only the selected features need to be computed, which significantly reduces the computational effort. Here, we classified the images using randomized quasi-exhaustive (RQE) feature banks (Tokarczyk et al., 2015), which are able to capture multiscale texture properties in a pixel's neighbourhood.

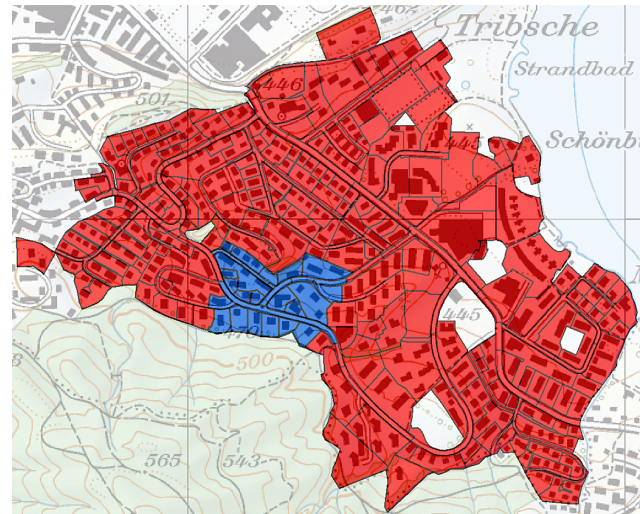
### Performance assessment of classification

To assess the performance of the two classifiers used in this study, we have manually labelled a subset (5 ha) of each of the image data sets (see Fig. 3). Hence, we were able to report the classification accuracy for all pixels in an extended area, which in our view is a lot more reliable than sparse, point-wise accuracy assessment. We selected either three (rooftops/streets/vegetation) or two (impervious/pervious) target classes, where in the two-categories case, the “impervious” class is an aggregation of the “rooftops” and “streets” classes. For the subsequent hydrological analysis, only the two-class maps were used.

Both classifiers were trained using randomly selected subsets of pixels (1, 2 or 5 %, which correspond roughly to 7000, 14 000 and 36 000 pixels). Thereby we can evaluate how the size of the training data has an influence on the overall classification accuracy. If satisfactory results can be obtained, then a lower number of training samples is preferable, since it reduces the training time and saves annotation effort. Similarly to experiments carried out in Tokarczyk et al. (2015), we trained the boosting classifier using decision trees with eight leaf nodes, and set the number of boosting rounds to 500. As a performance metric for the classification, we used the overall accuracy (OA), i.e. the fraction of correctly classified pixels.

#### 2.2.2 Assessing the importance of input data for surface runoff

To assess the influence of input data accuracy on the surface runoff, we predicted the surface runoff for a rain event of moderate intensity (total volume of 29.7 mm; peak rainfall intensity of  $2.9 \text{ L s}^{-1}$ ). Then, we analysed the runoff of the 307 individual subcatchments regarding the following attributes: (i) peak flow ( $Q_{\text{peak}}$ ) and (ii) volume of the peak ( $V_{\text{peak}}$ ). As it is very challenging to directly measure surface runoff that can be compared with model predictions, we first performed an exploratory analysis of the major influence factors. Second, we investigated interactions between the data



**Figure 3.** The Wartegg area containing 307 subcatchments (red polygons including blue polygons) overlaid on a topographic map. The performance of classifiers was assessed on a subset depicted in blue.

source and processing method by means of a regression analysis (see Sect. S3 in the Supplement for details).

### Performance assessment

#### *Exploratory data analysis of surface and surface runoff characteristics*

To summarize the important characteristics of the surface runoff, we visualized important aspects using boxplots and scatterplots (see Fig. 6). The main research questions were the following:

- Which differences in imperviousness ( $\Delta_{\text{Imp}}$ ) result for each subcatchment: (i) for the two data sources and (ii) for the two classification methods?
- Does the image source have a substantial influence on the predictions of surface runoff from a subcatchment? How does this depend on the processing method?

#### *Regression analysis of surface runoff characteristics*

To answer the second question, we constructed four regression models with indicator variables (Montgomery et al., 2012). This makes it possible to consider the individual effects of the data and the processing method. In addition, a model with an interaction term, unlike an additive model, could add a further adjustment for the “interaction” between the data source and the classification method. Specifically, we would like to explore whether the relationship between the image source and the imperviousness in the subcatchments, and their surface runoff characteristics, is different for each classifier. The model for a dependent variable  $y$  is

$$y_i = \beta_0 + \beta_1 I_i^{\text{Data}} + \beta_2 I_i^{\text{Method}} + \beta_3 I_i^{\text{Data} \cdot \text{Method}} + \epsilon_i, \quad (1)$$

where  $y_i$  is the  $i$ th observation of the dependent variable,  $I_i^{\text{Data}}$  an indicator variable which is 1 if  $y_i$  was computed from UAV images (UAV) and 0 from orthophotos, and  $I_i^{\text{Method}}$  is an indicator variable which is 1 if  $y_i$  was computed with the RQE method and 0 for the ML classifier (ML).  $\beta_0 \dots \beta_3$  are the parameters to be estimated and  $\epsilon_i$  is a random error term. If  $\epsilon_i$  is normally and independently distributed, i.e.  $\epsilon_i \sim N(0, \sigma^2)$ , this model is equivalent to a classical least square regression or to a three-way analysis of variance model with treatment contrasts (Montgomery and Runger, 2007).

The imperviousness is bounded between 0 and 1, whereas a linear model could easily predict values beyond this range, which is not admissible. To have a more plausible model, we therefore used a logit-transformation on the imperviousness (%imp):

$$z = 2 \cdot \operatorname{arctanh}(2 \cdot \text{Imp} - 1). \quad (2)$$

In addition, we analyse the results of this regression analysis on a qualitative basis only. With more correct and more complex models, which better represent the underlying process that generated the data,  $p$  values (see Tables S3–S5 in the Supplement) would tend to be larger. Here, however, we are not interested in the magnitude or statistical significance of the individual effect, but we would just like to see whether they are very different or not.

### 2.2.3 Prediction of pipe flows

To assess the model’s capability to predict the resulting in-sewer flow, we predicted stormwater flows at the catchment outlet for 36 independent rain events of different intensity and duration (see below) and compared them with flow data measurements (see Sect. 3.3). In particular, we compared measured and predicted volume of the total runoff as well as peak flows. The main driving questions for the analysis were the following:

- How do differences in imperviousness affect pipe flow predictions?
- To what extent may differences regarding input data, i.e. degree of imperviousness of subcatchment areas, be compensated by the model calibration procedure?

#### Model calibration

To address the latter question, we compared the results of the different model implementations prior to and after calibration. For the calibration/validation procedure we split the reference data set into a calibration (July to September 2014) and a validation period (September to November 2014). In

total, for both periods, 36 independent rain events of different intensity and duration were observed, which we consider sufficient to cover the inherent variability of rain events.

To analyse the effect of different input data and how this would be addressed by model calibration, we applied a genetically adaptive multi-objective calibration algorithm (AMALGAM, Vrugt and Robinson, 2007) to adjust the calibration parameters of the four implementations. The model input (two image data sources  $\times$  two different classifiers) is used to derive the “%imp” parameter. In the optimization, four different calibration parameters were adjusted to match three objective functions: (i) simulation bias (SB) and Nash–Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970), (ii) total flow balance, and (iii) peak flow deviation – all with respect to the flow at the catchment outlet. The input parameter imperviousness “%imp” is derived from orthophotos and is not subject to calibration. The calibration parameters are

- catchment width (m),
- HORTON maximum infiltration rate ( $\text{mm day}^{-1}$ ),
- decay constant for the HORTON curve ( $\text{day}^{-1}$ ), and
- size of a virtual subcatchment (ha), mimicking ground-water infiltration into the sewer pipe network.

#### Performance assessment: flow balance and flow dynamics

In a first step, we evaluated the match between modelled hydrographs and reference flow data using the SB and NSE. Both goodness-of-fit measures are well established in urban hydrology to cover deviations regarding the flow balance (bias) and flow dynamics (NSE). The simulation bias  $B$  is defined as follows:

$$B = (\bar{E} - \bar{M})^2, \quad (3)$$

whereas  $\bar{M}$  is the mean of measured (observed) values and  $\bar{E}$  is the mean of estimated (simulated) values. The bias ranges from  $-\infty$  until  $+\infty$  with an optimum at 0. The Nash–Sutcliffe efficiency NSE is defined as

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N |M_i - E_i|^2}{\sum_{i=1}^N |M_i - \bar{M}|^2}, \quad (4)$$

whereas  $M_i$  is the measured (observed) and  $E_i$  is the simulated value at the time  $i$ ,  $\bar{M}$  is the mean of measured (observed) values,  $\bar{E}$  is the mean of estimated (simulated) values, and  $N$  the number of paired data. NSE reaches 0 when the square of the differences between measured and estimated values is as large as the variability in the measured data. In case of negative NSE values the measured mean is a better predictor than the model.

**Table 1.** RQE vs. ML method: overall classification accuracies (in %). Boosting with RQE features after 500 iterations. Maximum likelihood classifier was trained with features consisting of single raw pixel intensities (all spectral channels).

Class. method/ % of train data	UAV			Orthophoto		
	1 %	2 %	5 %	1 %	2 %	5 %
Three classes						
ML	78.9	72.8	78.4	84.2	84.4	80.8
RQE	93.7	94.3	95.2	95.6	95.8	96.3
Two classes						
ML	87.7	81.6	84.3	90.9	90.8	88.4
RQE	95.5	95.6	96.2	96.6	97.0	97.4

To cover one of the key figures relevant for engineering urban drainage systems, we included an event-specific evaluation of peak flows in a second evaluation step. To this end, we extracted peak flows from observed and modelled hydrographs using an event filter that identifies independent rainfall–runoff events preceding a dry weather period by at least 6 h.

### 3 Results

#### 3.1 Classification

Table 1 presents per-pixel overall classification accuracy achieved using (i) two different data sets, (ii) two classification methods, and (iii) either two or three target classes. Figures 4 and 5 present visual classification results for a subset of each of the two data sets, together with a respective ground truth. We did not perform any pre- or post-processing of the data. Image pre-processing adds no information and typically does not help, except for physically meaningful reflectance calibration, which in our setting was not feasible. Post-processing of the imperviousness map might improve overall accuracy, but carries the danger of introducing unwanted biases.

#### 3.2 Prediction of surface runoff

##### 3.2.1 Exploratory analysis

We used boxplots and scatterplots to investigate the effect of combining two data sources and two processing methods on (i) the imperviousness and the surface runoff characteristics, (ii) peak flows, and (iii) runoff volumes (see Fig. 6).

- *Imperviousness (Imp)*: the boxplot shows that the overall distributions of imperviousness for 307 subcatchments do not differ much across the different image sources and classification methods. In general, the UAV

images seem to produce slightly lower values of imperviousness than the orthophoto, although this effect might also be dominated by the set of UAV image which was processed by the ML classifier. Regarding the classification methods, the boosting classification method delivers slightly larger imperviousness values for both data sources than the ML method.

- *Peak runoff (Peak)*: like for the imperviousness, the different image sources lead to very similar peak runoff values. In general, boosting leads to slightly higher peak flows, which also have a larger variance and slightly higher extreme values for a couple of subcatchments. Regarding the suitability of UAV images in rainfall–runoff modelling, there are no relevant differences between the image sources.
- *Runoff volumes (Volume)*: the exploratory analysis effectively suggests the same patterns for the runoff volume as for the peak flows: boosting leads to larger runoff volumes and the resulting variability of the rainfall runoff from the 307 subcatchments is slightly larger than for the ML classifier. Also, the UAV data seem to be associated with smaller runoff volumes. This is consistent, as this relates to the lower degree of imperviousness in the subcatchments.

In general, the relative differences between the different alternatives are very small, with average values of a few percent (see Fig. 6). For the imperviousness, there are only a few subcatchments which show rather large differences. These are even less relevant for the peak runoff and runoff volumes.

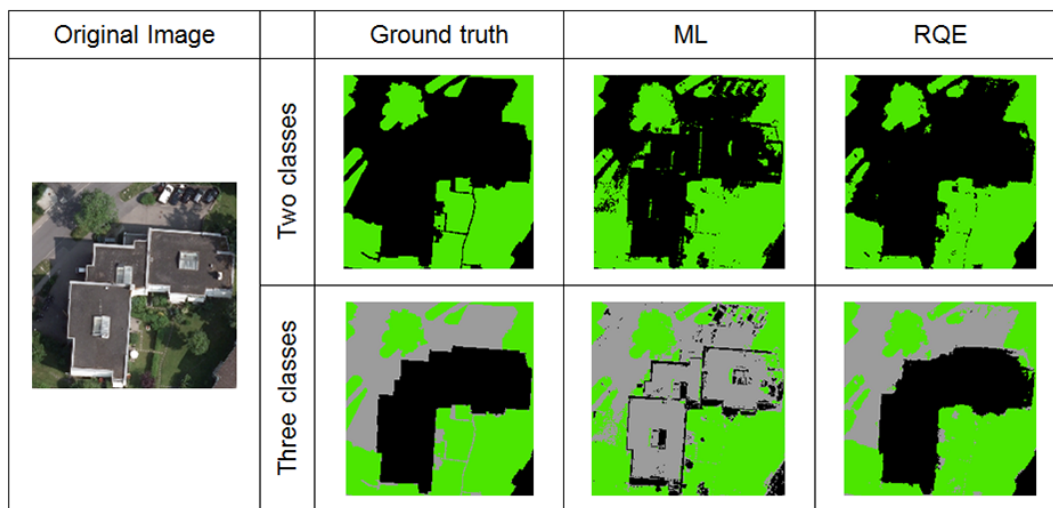
Furthermore, the scatterplots of the different explanatory and dependent variables suggest that there is not a substantial difference between the image sources or classification approaches for the modelled surface runoff in the different subcatchments (see Fig. S1 in the Supplement). For the boosting classifier, we observe a weak positive correlation with the degree of imperviousness (see Fig. S2 in the Supplement), which means that catchments which are rather impervious (or pervious) based on the ML classifier tend to be even more impervious (or pervious) for the boosting classifier. However, this is difficult to identify by means of visual analysis and is better explored by an analysis of variance or regression analysis.

##### 3.2.2 Regression analysis

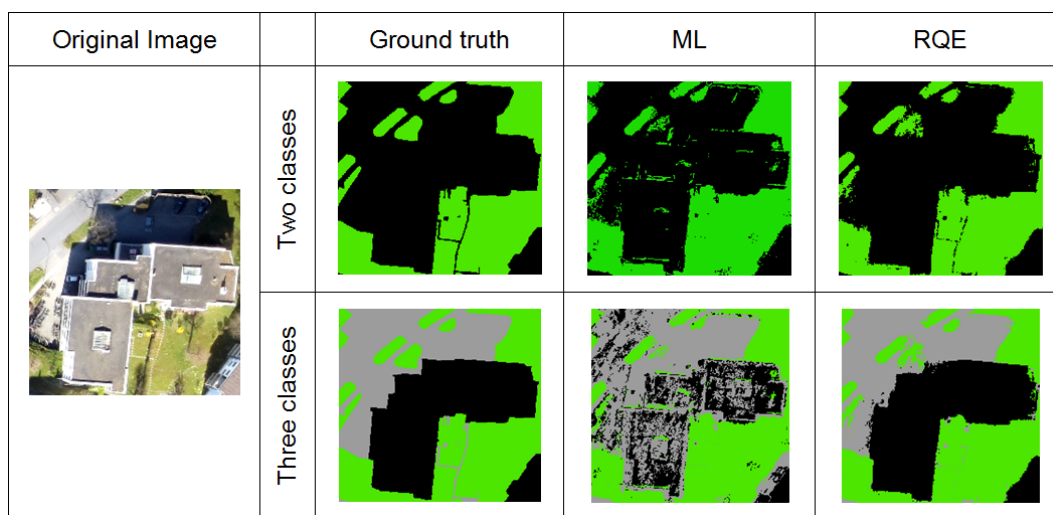
The results from the regression analysis are mainly the maximum likelihood estimates of the model parameters and an indicator of their importance (see Tables S3 and S4 in the Supplement).

For the *imperviousness*, as expected neither the image source nor the classifier is strongly correlated. The negative sign of the estimated slope parameter for the image source ( $\beta_1 = -0.16$ ) suggests that UAV images generally go to-





**Figure 4.** Cutouts of the swisstopo image: original image, manually labelled ground truth, and classification results using ML and RQE (two and three classes). In the case of two classes, impervious surfaces are black and pervious are green. In the case of three classes, rooftops are black, streets/sidewalks are grey and vegetation is green.



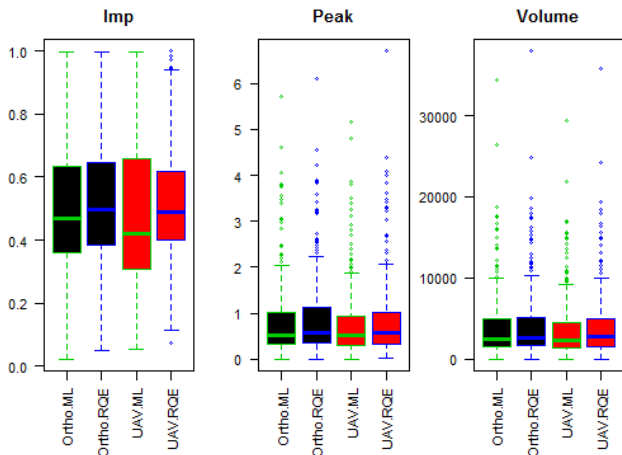
**Figure 5.** Cutouts of the UAV image: original image, manually labelled ground truth, and classification results using ML and RQE (two and three classes). In the case of two classes, impervious surfaces are black and pervious are green. In the case of three classes, rooftops are black, streets/sidewalks are grey and vegetation is green.

gether with a lower imperviousness. In addition, the influence of the image source seems to be larger than that of the classification method ( $\beta_2 = 0.003$ ), although the large  $p$  values for all parameters suggest that it is not very likely that the observed values of imperviousness were to have occurred under the given statistical model. Therefore, there is virtually no evidence that there are interactions between the image source and the classifiers.

For the *peak runoff*, neither the image source nor the classifier are strongly correlated. The negative sign of the estimated slope parameter for the image source ( $\beta_1 = -0.6$ ) suggests that UAV images correlate with a smaller peaks. Here,

the influence of the image source seems to be equally important as the classification method ( $\beta_2 = -0.6$ ), just with a different sign. Nevertheless, the high  $p$  values for all parameters again suggest that it is not very likely that the observed values of imperviousness were to have occurred under the given statistical model. Also, the interaction between the image sources and classifiers is not important.

For the runoff volume, the UAV data generally seem to be correlated with slightly lower runoff volumes ( $\beta_1 = -302$ ), whereas the RQE method shows a positive correlation ( $\beta_2 = 298$ ). Again, neither the two effects nor their interaction seem to be important.



**Figure 6.** Boxplots of the imperviousness and surface runoff characteristics (Imp (–), Peak ( $L s^{-1}$ ), Volume ( $m^3$ )) for the 307 sub-catchments for the four combinations of data sources and processing methods. Black = Ortho, red = UAV, green = ML, and blue = RQE.

In summary, the analysis suggests that surface runoffs predicted with SWMM are similar for the different data sources or classifiers. In addition, neither the imperviousness nor peaks nor volumes of the runoff are influenced by interactions between the image sources and the classification methods. As the data source and classifier alone do not represent the data generating process, the underlying statistical assumptions are not met and the numerical results should not be over-interpreted.

### 3.3 Prediction of in-sewer flow

The evaluation regarding sewer pipe flow is split into two parts: (1) model performance of uncalibrated implementations, and (2) calibrated implementations compared to reference data, i.e. flow measured at the outlet of the catchment.

1. Focusing on the results prior to calibration, it becomes clear that uncalibrated models, among each other, differ particularly regarding the peak flow performance (see boxplot in Fig. 7). This clearly corresponds to the findings of the surface runoff analysis (see Sect. 3.2) in which, for instance the implementation “UAV ML” with the lowest mean degree of imperviousness produces the lowest runoff peaks. The comparison with reference data through hydrological goodness-of-fit measures (see Table 2) underlines the moderate performance regarding flow dynamics (NSE), whereas already good agreement is achieved for the total flow balance (bias). The slightly improved performance of the implementation of which the imperviousness is derived from UAV data classified with the ML method (UAV ML) probably occurs by chance.

**Table 2.** Goodness-of-fit measures prior to and after calibration (both quantified for the validation period).

	Prior to calibration SB (–)/ NSE (–)	After calibration SB (–)/ NSE (–)
Ortho ML	2.0/0.54	$3.16 \times 10^{-5}$ /0.72
Ortho RQE	2.0/0.52	0.007/0.71
UAV ML	0.3/0.62	0.1/0.75
UAV RQE	2.0/0.53	1.38/0.73

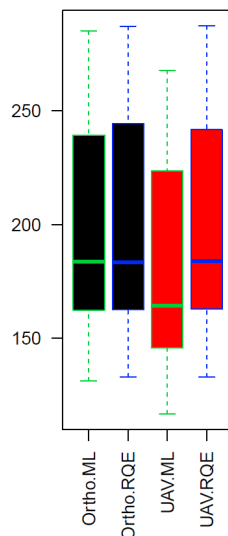
2. Results from calibrated models (see Fig. 8 and Table 2, right column) show that conducting a detailed calibration, as expected, leads to an improved model performance (NSE increase, bias reduction) and interestingly smooths out the land-use differences among the four implementations. This is visible in Fig. 8, where the hydrographs are practically the same. Even though the results from the UAV ML implementation after calibration still shows slightly different results (see Fig. 8, right panel), the differences in peak flow for the 13 most intense rain events are very similar (see Fig. 9).

Interestingly, the very similar performance is achieved with very different parameter estimates (see Fig. S6 in the Supplement). Particularly the parameter “width”, “maximum infiltration rate” and “Decay K” (influencing the peak flow) vary substantially. Results show that the calibrated runoff model should be fairly robust against variations of the perviousness map, since these can be compensated by changing other, more uncertain parameters, e.g. by different parameters defining the infiltration into pervious surfaces.

## 4 Discussion

### 4.1 Classification

In order to fully exploit the advantages coming with high spatial resolution of an image, one has to use the classification method tailored to the characteristics of a data set. Thus, the choice of the classifier has a substantial impact on the overall classification accuracy. While boosting achieves accuracies between 93.7 and 96.2 % for the UAV data set and 95.6 to 97.4 % for the *swisstopo* data set, maximum likelihood yields results which are up to 20 % worse. Furthermore, it can be seen that the number of target classes strongly influences the results of the ML method. Classification with three target classes is up to 9 % less accurate than with two. Moreover, the amount of data used to train the ML classifier gives inconclusive results. By increasing the number of training samples, overall accuracy should increase. However, in



**Figure 7.** Evaluation of peak flows ( $\text{L s}^{-1}$ ) for the 13 most intense rain events (prior calibration).

our case the training appears to be unstable, and the expected increase only materializes in a single case (see Table 1, orthophoto data set, three classes). A possible explanation is that the class distribution is not unimodal, and thus not appropriately captured by the Gaussian model.

In contrast to the ML method, the boosting classifier behaves in a stable manner. Differences in overall accuracy do not exceed 2.5 % per data set. The changes in boosting performance with varying amounts of training data are negligible: 1 % (7000 pixels) already yield satisfactory results, i.e. the effort for annotation as well as the training time remains low. The efficiency and robustness of boosting used together with features appropriate for VHR aerial imagery makes this approach a good choice for the task. Also, overall classification accuracy achieved using a boosting classifier together with UAV-based imagery shows that in terms of classification accuracy of impervious surfaces, this new imaging platform gives comparable results to the off-the-shelf aerial image products.

Moreover, our experiments show that at the level of surface runoff prediction, the differences between different imaging platforms and between different processing methods are small. Even though the classification accuracy between data sets and methods differs by up to 20 %, their influence on surface runoff characteristics lies within only a few percent on average. We believe that one of the reasons is the spatial size of our subcatchments. Each of them consists of hundreds of image pixels, but the hydrological model disregards the spatial information and only uses aggregated values, i.e. the sum of all impervious pixels belonging to one subcatchment. A further observation is that the differences in classification accuracy are much larger for the three-class case. This is in line with conventional machine learning wisdom (“only predict

what you need to know”); however, we have not yet constructed an end-to-end study with the three-class result as an input.

## 4.2 Prediction of surface runoff

### 4.2.1 Exploratory analysis of surface runoff

While there are substantial differences when the images are compared pixel-by-pixel (see Figs. 4 and 5), these are largely lost for the predicted surface runoff. In our view, this is again explained by the SWMM surface runoff model. It is a lumped model, which aggregates the pixels and thus smoothes out the differences already on this small scale. This tendency will be even more pronounced for a higher degree of spatial aggregation, e.g. when modelling larger urban areas, where the subcatchments equipped with flow measurements will also be larger. Future experiments that investigate the continuous spatial downsampling of images may reveal when differences fully disappear.

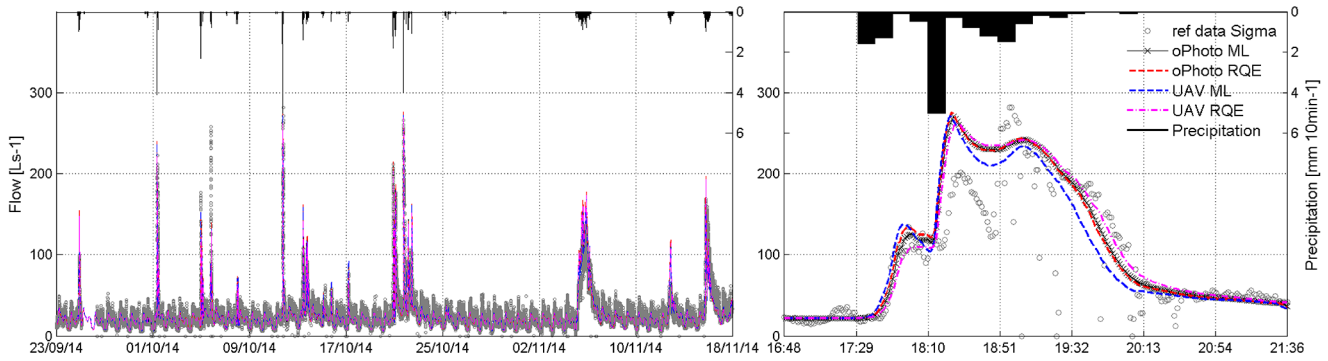
### 4.2.2 Model structure as a bottleneck?

Obvious differences in the input data may be smoothed out due to the simplified, conceptual representation of the surface runoff in SWMM. We do expect different results for more detailed representation of land use, e.g. with a separate “roof” land-use or modern pixel-based modelling approaches for surface runoff. In the future, this might be even more important considering the increasing popularity of coupled 2-D-overland/1-D-channel flow models including more detailed overland-flow modelling using raster/pixel-based approaches (cf. Leandro et al., 2009). Traditional models – as currently used in day-to-day engineering practice – will probably never be able to fully make use the amount of detail (pixel basis) provided by such aerial images.

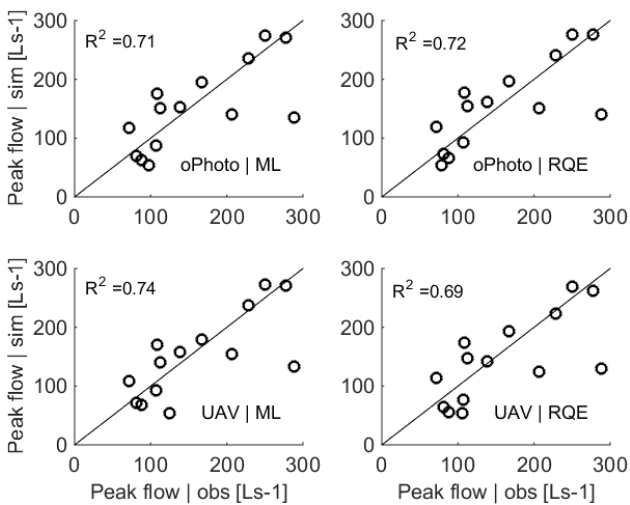
### 4.2.3 High-resolution images provide added value in urban drainage

The effect on surface runoff and pipe hydraulics using spatially aggregating models (two land-use classes) may not be as immense. However, in future investigations, models that allow differentiating between three or more land-use classes should be further investigated. This may be particularly relevant for pollutant load modelling, for which detail, accuracy and actuality of land-use characteristics are highly influential. Relevance of input data accuracy may even further increase due to the fact that obtaining adequate pollution load reference data is considered to be very difficult (cf. Dotto et al., 2014).

Also, other urban drainage tasks would greatly benefit from detailed land-use maps, e.g. precise and justified stormwater fees due to exactly delineated types of impervious areas (cf. Figs. 4 and 5). An improved feature (gully pots, sewer inlets, curbstone structures) identification is expected



**Figure 8.** Observed reference and simulations (prior calibration) for the full validation period September to November 2014 (left panels) and a selected event on 11 October 2014 (right panels).



**Figure 9.** Evaluation of the peak flows for the 13 most intense rain events in the validation period (after calibration).

to further provide valuable input data for network generation approaches (e.g. as outlined in Blumensaat et al., 2012) and the coupled 2-D surface runoff/1-D pipe flow model applications. For this, the RQE method seems to be most promising, although for the runoff analysis, a simpler method still seems to produce robust results.

The possibility of on-demand image acquisitions through UAV flights allows almost instantaneous response to land-use developments in dynamic urban environments. As land use changes become increasingly evident, keeping hydrological models up-to-date appears to be a key to effectively reduce the risk of urban flooding. We consider the flexibility of collecting high-quality images at almost any time (“on-demand”) for spatially pre-defined urban areas of manageable size as clear benefit, also with regard to cost efficiency.

### 4.3 Pipe flow predictions

The results from the model calibration show that input data deviations are nearly fully compensated by the calibration procedure, involving an adaption of four different calibration parameter sets. The analysis of the final calibration parameter values however reveals that the best fit for each of the implementations is achieved by differing parameter sets (see Fig. S6 in the Supplement). On the one hand side, this may indicate that, even though the full range of a priori defined parameter ranges is used during the auto-calibration procedure, for each implementation a different (local) optimum in the Pareto front is identified. On the other hand, it may underline that the given model structure is flexible enough to address different model inputs through different parameter settings. Here, it becomes clear that the compensation is achieved by adjusting parameters in a way that involves the risk that some parameters lose its physically based origin and turn into “conceptual handles”. The discussion on this particular question is certainly interesting and would need further analyses, but it cannot be accomplished in this paper contribution as it would blur the main focus of the paper.

### 5 Conclusions

In this study we investigated the possibility to automatically generate high-resolution imperviousness maps for urban areas from imagery acquired with UAVs, and for the first time assessed the potential of UAVs for high-resolution hydrological applications compared with a standard large-format aerial orthophotos. We proposed an automatic processing pipeline with modern classification methods to extract accurate imperviousness maps from high resolution aerial images, and presented an end-to-end comparison, in which the maps obtained from different sources and processed with different classification methods were used as input for urban drainage models.

The first part of our analysis indicates that, using a boosting classifier in conjunction with RQE features, we were able to classify UAV imagery with an accuracy comparable to standard aerial orthophotos. The proposed classification method yields more stable results, when compared with those produced using the maximum likelihood method. This improvement is even more apparent when classifying three instead of two classes of land use.

In the second part of our analysis we have demonstrated how model input data variations propagate in the course of the urban drainage modelling exercise, and how this is reflected in the surface runoff and sewer flow predictions. Results from uncalibrated model implementations actually show deviations in the predictions, which can be explained by input data variations. But still predictions are inaccurate. Conversely, after calibration the performance analysis shows that the calibration process attenuates variations in the input data, suggesting that model predictions are insensitive to these variations. However, the analysis of the resulting model parameter settings also reveals that apparent robustness is achieved by tweaking the parameter in a way which involves the risk of leaving valid parameter ranges.

Because model development and calibration in everyday practice is often based on less accurate information than used in our case study, it is important to underline reliable input data to reduce overall uncertainty in model predictions.

We note that the conclusions of the study are limited regarding (i) the small size of the case study catchment, (ii) the degree of detail in which the catchment has been described (more detail may show a more pronounced input error propagation, a more lumped description may absorb input deviations from the start), and (iii) the type of hydrological modelling concept used. Therefore we suggest conducting further research to evaluate the impact of the spatial scale, i.e. the degree of spatial aggregation linked to the hydrological model approach (ensemble modelling). In the case study presented here we chose a traditional and widely used urban drainage model (EPA SWMM) to deliberately demonstrate the effect of new image sources and processing methods for standard engineering practice.

Still, we suggest using imperviousness maps consisting of three land-use classes as more differentiated input for a more detailed hydrological model, i.e. a pollution load model, which makes better use of urban land-use differentiation. Because the proposed boosting classifier showed the largest accuracy gain for a three-class case, we strongly believe that introducing this additional information more clearly shows the potential of UAV data sets and advanced classification methods for more accurate urban drainage and pollution load modelling.

**The Supplement related to this article is available online at doi:10.5194/hess-19-4215-2015-supplement.**

*Acknowledgements.* This publication is an outcome of a fruitful interdisciplinary collaboration between photogrammetrists and hydrologists, which was triggered by joint supervision of Matthew Moy de Vitry's Master's thesis titled "Improving Urban Flood Management with Autonomous Mini-UAVs". We would like to thank Matthew for providing us with a UAV data set. Also, we are very grateful to MeteoSwiss and the city of Lucerne for providing us with the precipitation and infrastructure data, and the Engineering Consultants from HOLINGER AG, Bern for providing us with the Holinger company for providing us with the *swisstopo* aerial image data. Last but not least, we would like to thank Philippe Gerber for helping us with an automatic calibration of the pipe flow model.

Edited by: N. Romano

## References

- Arrighi, C., Brugioni, M., Castelli, F., Franceschini, S., and Mazzanti, B.: Urban micro-scale flood risk estimation with parsimonious hydraulic modelling and census data, *Nat. Hazards Earth Syst. Sci.*, 13, 1375–1391, doi:10.5194/nhess-13-1375-2013, 2013.
- Benbouzid, D., Busa-Fekete, R., Casagrande, N., Collin, F.-D., and Kégl, B.: MultiBoost: a multi-purpose boosting package, *J. Mach. Learn. Res.*, 13, 549–553, 2012.
- Blumensaat, F., Wolfram, M., and Krebs, P.: Sewer model development under minimum data requirements, *Environ. Earth Sci.*, 65, 1427–1437, 2012.
- Chormanski, J., Van de Voorde, T., De Roeck, T., Batelaan, O., and Canters, F.: Improving Distributed Runoff Prediction in Urbanized Catchments with Remote Sensing based Estimates of Impervious Surface Cover, *Sensors*, 8, 910–932, 2008.
- Dotto, C., Kleidorfer, M., Deletic, A., Rauch, W., and McCarthy, D.: Impacts of measured data uncertainty on urban stormwater models, *J. Hydrol.*, 508, 28–42, 2014.
- Eisenbeiß, H.: UAV photogrammetry, Institut für Geodäsie und Photogrammetrie, Eidgenössische Technische Hochschule Zürich, Zürich, 2009.
- ESRI: ArcMap 10.2.1. Environmental Systems Resource Institute, Redlands, California, 2013.
- Fankhauser, R.: Automatic determination of imperviousness in urban areas from digital orthophotos, *Water Sci. Technol.*, 39, 81–86, 1999.
- Forzieri, G., Battistini, A., and Catani, F.: ES4LUCC: A GIS-tool for Remotely Monitoring Landscape Dynamics, *Comput. Geosci.*, 49, 72–80, 2012a.
- Forzieri, G., Moser, G., and Catani, F.: Assessment of hyperspectral MIVIS sensor capability for heterogeneous landscape classification, *ISPRS J. Photogram. Remote Sens.*, 74, 175–184, 2012b.
- Freund, Y. and Schapire, R.: A decision-theoretic generalization of on-line learning and an application to boosting, in: *Computational Learning Theory*, in: vol. 904 of *Lecture Notes in Computer Science*, Springer Verlag, Heidelberg, 23–37, 1995.
- Hall, J., Arheimer, B., Borga, M., Brézdil, R., Claps, P., Kiss, A., Kjeldsen, T. R., Kriauciuniene, J., Kundzewicz, Z. W., Lang, M., Llasat, M. C., Macdonald, N., McIntyre, N., Mediero, L., Merz, B., Merz, R., Molnar, P., Montanari, A., Neuhold, C.,

- Parajka, J., Perdigão, R. A. P., Plavcová, L., Rogger, M., Salinas, J. L., Sauquet, E., Schär, C., Szolgay, J., Viglione, A., and Blöschl, G.: Understanding flood regime changes in Europe: a state-of-the-art assessment, *Hydrol. Earth Syst. Sci.*, 18, 2735–2772, doi:10.5194/hess-18-2735-2014, 2014.
- Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., and Kanae, S.: Global flood risk under climate change, *Nat. Clim. Change*, 3, 816–821, 2013.
- Hughes, G.: On the mean accuracy of statistical pattern recognizers, *IEEE T. Inform. Theory*, 14, 55–63, 1968.
- Leandro, J., Chen, A. S., Djordjević, S., and Savić, D. A.: Comparison of 1D/1D and 1D/2D coupled (sewer/surface) hydraulic models for urban flood simulation, *J. Hydraul. Eng.*, 135, 495–504, 2009.
- Leitão, J. P., Moy de Vitry, M., Scheidegger, A., and Rieckermann, J.: Assessing the quality of Digital Elevation Models obtained from mini-Unmanned Aerial Vehicles for overland flow modelling in urban areas, *Hydrol. Earth Syst. Sci. Discuss.*, 12, 5629–5670, doi:10.5194/hessd-12-5629-2015, 2015.
- Montgomery, D. C. and Runger, G. C.: *Applied statistics and probability for engineers*, Hoboken, NJ, Wiley, 2007.
- Montgomery, D. C., Peck, E. A., and Vining, G. G.: Introduction to linear regression analysis, in: vol. 821, Wiley, Hoboken, NJ, 2012.
- Moser, G., Serpico, S., and Benediktsson, J.: Land-Cover Mapping by Markov Modeling of Spatial-Contextual Information in Very-High-Resolution Remote Sensing Images, *Proc. IEEE*, 101, 631–651, 2013.
- Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I: A discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.
- Nielsen, N. H., Joergensen, A., and Larsen, A.: Use of spectral analysis in urban drainage modelling, in: International Conference on Urban Drainage, 11–16 September 2011, Porto Alegre, Brazil, 2011.
- Ravagnani, F., Pellegrinelli, A., and Franchini, M.: Estimation of Urban Impervious Fraction from Satellite Images and Its Impact on Peak Discharge Entering a Storm Sewer System, *Water Resour. Manage.*, 23, 1893–1915, 2009.
- Rojas, R., Feyen, L., and Watkiss, P.: Climate change and river floods in the European Union: Socio-economic consequences and the costs and benefits of adaptation, *Global Environ. Change*, 23, 1737–1751, 2013.
- Sauerbier, M. and Eisenbeiß, H.: UAVs for the documentation of archaeological excavations, *International Archives of Photogrammetry, Remote Sens. Spat. Inform. Sci.*, 38, 526–531, 2010.
- Stankowski, S. J.: Population density as an indirect indicator of urban and suburban land-surface modifications, US Geological Survey Professional Paper, US Geological Survey, Washington, D.C., USA, 800, 219–224, 1972.
- swisstopo: SWISSIMAGE: Das digitale Farbrothophotomosaik der Schweiz, Eidgenössisches Departement für Verteidigung, Bevölkerungsschutz und Sport VBS, <http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/images/ortho/swissimage.parsysrelated1.76752.downloadList.50684.DownloadFile.tmp/infosi201003deu.pdf> (last access: 18 October 2015), 2010.
- swisstopo: swissALTI3D – Ausgabebericht, Eidgenössisches Departement für Verteidigung, Bevölkerungsschutz und Sport VBS, <http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/height/swissALTI3D.html> (last access: 18 October 2015), 2014.
- Tokarczyk, P., Wegner, J., Walk, S., and Schindler, K.: Features, Color Spaces, and Boosting: New Insights on Semantic Classification of Remote Sensing Images, *IEEE T. Geosci. Remote*, 53, 280–295, 2015.
- UN: World Population Prospects 1950–2050, The 2012 Revision, Key Findings and Advance Tables, United Nations Population Division, [http://esa.un.org/unpd/wpp/Publications/Files/WPP2012\\_HIGHLIGHTS.pdf](http://esa.un.org/unpd/wpp/Publications/Files/WPP2012_HIGHLIGHTS.pdf) (last access: 18 October 2015), 2013.
- US-EPA: Storm Water Management Model (SWMM), Version 5.0.022. United States Environmental Protection Agency, <http://www2.epa.gov/water-research/storm-water-management-model-swmm> (last access: 18 October 2015), 2010.
- Vrugt, J. A. and Robinson, B. A.: Improved evolutionary optimization from genetically adaptive multimethod search, *P. Natl. Acad. Sci.*, 104, 708–711, 2007.
- Zhang, C. and Kovacs, J.: The application of small unmanned aerial systems for precision agriculture: a review, *Precis. Agricult.*, 13, 693–712, 2012.
- Zhou, Y. and Wang, Y. Q.: Extraction of Impervious Surface Areas from High Spatial Resolution Imageries by Multiple Agent Segmentation and Classification, *Photogram. Eng. Remote Sens.*, 74, 857–868, 2008.