Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-570-AC2, 2019


Hydrology and Earth System Sciences
Open Access
Discussions



# *Interactive comment on* "Scalable Flood Level Trend Monitoring with Surveillance Cameras using a Deep Convolutional Neural Network" *by* Matthew Moy de Vitry et al.

**Matthew Moy de Vitry et al.**

matthew.moydevitry@eawag.ch

Received and published: 5 April 2019

## 1 Reply to general comments from Referee # 2

*This paper proposed an approach to monitor flood level trend using DCNN. The topic is very interesting. However, in my opinion, it could be difficult for modelers, decision-makers and city planners to use the Static Observer Flooding Index (SOFI) directly. The authors should clearly explain what the direct or specific application scenarios of SOFI are. If SOFI or visible area of the flooding can be converted into water depth value or even class information on water depth, it would make this approach more*



Printer-friendly version

Discussion paper

*useful and valuable. Unfortunately, this paper lacks this attempt, so I suggest rejecting this paper but encourage resubmission after the improvements have been made.*

We thank Referee 2 for the overall positive review, for the useful recommendations, and for the questions which will allow us to further improve the manuscript.

**On the value of flood level trend information**

While we agree with Referee 2 that decision-makers and city planners may not be accustomed to flood trend information, modelers in research have already successfully used trend data for model calibration (e.g. van Meerveld et al 2017), albeit in the field of natural catchment hydrology. In that publication, stream level class data are used as trend information in a conventional calibration scheme, with the minor adaptation being the use of the Spearman rank correlation coefficient as an objective function. Similarly, Wani et al. (2017) showed that even simple binary information of a combined sewer overflow in an urban catchment can reduce model parameter uncertainty. Making the link to our manuscript, binary flooding information can easily be derived from the flood trend information that SOFI provides, by defining for each camera a baseline SOFI value below which no flooding is assumed.

We agree with Referee 2 that our manuscript could better explain how the utility of SOFI is supported by literature. To do so, we will augment our manuscript with details and background information for the examples listed above. We will make these changes in the introduction, discussion, and conclusion of the manuscript.

**On the advantages of SOFI over water depth estimation**

Referee 2 also suggests converting SOFI into water depths (either values or classes) if possible. Although methodologies for determining flood water levels are of value, we argue that the concept of flood trend monitoring with SOFI has advantageous characteristics that would be lost if SOFI were to be converted into a water depth.

1. **Higher scalability**. The obtention of water depth information from a segmented flooding image requires knowledge of the dimensions of reference objects in the camera scene. As shown by the literature suggested by Referee 2, there are different ways of obtaining this information. For example, one can rely on the visibility of ubiquitous objects of known dimensions (Jiang et al. 2019). This approach is limited to scenes in which such objects are visible, and fails if the reference object is hidden or partially hidden by mobile objects during a flood. Another way of obtaining reference measurements is with a survey of a large feature in each scene. Bohla et al. (2018) demonstrate this method nicely with bridges in urban stream settings. Needless to say, this method requires time, effort, and field knowledge for each of the sites surveyed, making it expensive and time-consuming to deploy at scale. SOFI was conceived as a metric that has few prerequisites and relies on minimal on-site knowledge, making it applicable to a broader range of scenes than any alternative method known to the authors.

2. **Fewer errors**. In stream or river settings, direct conversion of SOFI to a water depth is most probably possible e.g., by determining a "SOFI-water depth" curve analogous to an H-Q curve. However, in an urban environment, the movement of objects (e.g., cars, people) constantly change the extent of visible water. It follows that the "SOFI-water depth" curve will be inherently wrong. By setting lower ambitions, SOFI also relies on fewer assumptions about the scene monitored, and thereby introduces fewer possible sources of error.

3. **Lower risk of overvaluation**. Although a conversion of SOFI into water depth (either absolute value or class) could make the data more interpretable, it would

dissimulate uncertainty about the information delivered and could risk being valued at the same level of trust as data from conventional sensors. We therefore see value in keeping SOFI an adimensional, qualitative metric.

## Summary

In summary, literature supports the argument that the flood level trend information reflected in SOFI is valuable in itself to improve urban flood analysis, and do not necessarily need to be converted into water depth. Furthermore, the conversion of SOFI into a water depth would not only reduce the ease of applicability, but also introduce additional error sources and lead to possible overvaluation. Based on these arguments, we do not see the need to transform SOFI into a water level estimation method. However, we understand that the value of SOFI can be clarified and we will carry out the changes described above to this regard. We will also present the advantageous characteristics of SOFI in an extended comparison with alternative methods, as requested by Referee 2 in Comment # 2.13.

## References

Bhola, Punit Kumar, Bhavana B. Nair, Jorge Leandro, Sethuraman N. Rao, and Markus Disse. 2018. "Flood Inundation Forecasts Using Validation Data Generated with the Assistance of Computer Vision." *Journal of Hydroinformatics* 21 (2): 240–56. https://doi.org/10.2166/hydro.2018.044.

Jiang, Jingchao, Junzhi Liu, Changxiu Cheng, Jingzhou Huang, and Anke Xue. 2019. "Automatic Estimation of Urban Waterlogging Depths from Video Images Based on Ubiquitous Reference Objects." *Remote Sensing* 11 (5): 587. https://doi.org/10.3390/rs11050587.

van Meerveld, H. J. I., Vis, M. J. P., and Seibert, J.: Information content of stream level class data for hydrological model calibration, Hydrol. Earth Syst. Sci., 21, 4895-4905, https://doi.org/10.5194/hess-21-4895-2017 , 2017.

Wani, Omar, Andreas Scheidegger, Juan Pablo Carbajal, Jörg Rieckermann, and Frank Blumensaat. 2017. "Parameter Estimation of Hydrologic Models Using a Likelihood Function for Censored and Binary Observations." *Water Research* 121 (September): 290–301. https://doi.org/10.1016/j.watres.2017.05.038 .

## 2 Replies to specific comments from Referee # 2

### 2.1 Comment # 2.1

*Page 1, Line 16* "*The results suggest that the approach can be used with almost any surveillance footage* "*. I think this conclusion may be too strong.*

Thank you for pointing this out. This conclusion is indeed too strong given the limited number of videos on which the method was demonstrated (even though we chose videos that are diverse from a quality and scenery standpoint).

**Changes**: We will instead write that the results confirm that the method is versatile and can be applied to a variety of camera models and flooding situations.

### 2.2 Comment # 2.2

*Some new related references about automatic water level monitoring with surveillance images should be included to strengthen this paper. Wang, R. Q., Mao, H., Wang, Y., Rae, C., & Shaw, W. (2018). Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. Computers & Geosciences, 111, 139-147. Jiang, J., Liu, J., Cheng, C., Huang, J., & Xue, A. (2019). Automatic Estimation of Urban Water-*





Printer-friendly version

Discussion paper

*logging Depths from Video Images Based on Ubiquitous Reference Objects. Remote Sensing, 11(5), 587. Bhola, P. K., Nair, B. B., Leandro, J., Rao, S. N., & Disse, M. (2019). Flood inundation forecasts using validation data generated with the assistance of computer vision. Journal of Hydroinformatics, 21(2), 240-256.*

Thank you for referring us to these recent publications, which both highlight the need for flood monitoring data and illustrate the complexity of extracting absolute water level information from images.

**Changes**: We will add references to these publications to Sections 1.1 and 1.3 of the manuscript

## 2.3 Comment # 2.3

*Page 4, Figure 1. There should be a "surveillance images" box between "camera" box and "deep conv. network" box.*

Thank you for this input. We had several versions of Figure 1, some of which contained the box proposed, but left it out in the final version.

**Changes**: we will reintroduce a "surveillance video" box between "camera" box and "deep conv. neural network" box.

## 2.4 Comment # 2.4

*Page 4, Line 18. Why U-net was selected for water segmentation?*

U-Net is a very well known DCNN architecture for semantic segmentation, as the method has nearly 5000 citations in Google Scholar. It is well-suited to the flooding segmentation problem because of its relatively compact size compared to more recent state-of-the-art architectures (such as Mask-RCNN). The smaller size makes it both easier to train with small datasets (which we have) and faster to run, which is useful for flood monitoring.

**Changes**: We will include these reasons in the manuscript.

## 2.5 Comment # 2.5

*Page 6, Lines 18-20. In the augmented strategy, how many images were used to train the U-net model? Is the augmented data the same amount as the basic data? Moreover, Page 7, Table 1, the augmentation steps in the augmented strategy should be clearly explained.*

Thank you for this question. For the augmented strategy, we use the same images as for the basic strategy. Like for the basic strategy, each training image is fed into the network up to 200 times during training (fewer if the training is completed faster). For the augmented strategy, however, each image is first randomly transformed before being fed into the DCNN.

**Changes**: We will add information about the augmented strategy in Section 2.1.3. Fur-



Printer-friendly version

Discussion paper

thermore, we will provide more details about the augmentation transformations applied.

## 2.6 Comment # 2.6

*Page 7, Lines 5-6. It is not accurate enough. How many seconds or minutes does each model training take?*

**Changes**: At the line indicated, instead of a range we will indicate the approximate average training time in minutes for the basic and augmented strategies separately.

## 2.7 Comment # 2.7

*Page 8, Table2. The total frames or minutes and the resolutions of surveillance images should be given. How to define the quality of surveillance footage should be explained.*

It is true we could have provided more information here about the videos and the quality terms used.

**Changes**: We will add the requested information to Table2 and define quality in a table note.

## 2.8 Comment # 2.8

*Please add the average segmentation time of a single frame for the testing as a performance, since the real-time monitoring is usually important.*

Thanks for this request; it is true that this figure could be of interest to readers.

**Changes**: We will provide average segmentation time in Section 3.1

## 2.9 Comment # 2.9

*Page 9, Lines 16-17. The authors should explain more extensively how to compute the rank of each signal value.*

Thank you for your interest. The computation of the Spearman rank-order correlation coefficient, including signal rank of each value, was performed using the pandas Python library (McKinney, 2010), which is open source.

**Changes**: We will add a reference to this library in Section 2.3.3. However, Referee 1 commented that the description of the methodology is already very detailed, so we will not further develop on the signal rank computation in the manuscript.

McKinney, Wes. 2010. “Data Structures for Statistical Computing in Python.” In *Proceedings of the 9th Python in Science Conference*, edited by Stéfan van der Walt and Jarrod Millman, 51–56.

## 2.10 Comment # 2.10

*Page 9, Line 20. "the two signals" should be clearly identified as the SOFI signal and the reference signal.*

Thank you for pointing this out.

**Changes**: We will identify the two signals as suggested.

## 2.11 Comment # 2.11

*Page 12, Figure 6. Only the case of video FloodXCam5 was given, please add cases of parking lot and park that are the typical scenes of urban flooding.*

It is true that these other cases could be of interest to readers. We had left them out for space reasons, but they can easily be added.

**Changes**: We will add the two additional cases to Figure 6.

## 2.12 Comment # 2.12

*Page 14, Line 2. "the visually estimated flooding intensity" and "Flooding intensity" in Figure 8 should be consistent with "the visually estimated water level" in the caption of Figure 8.*





Printer-friendly version

Discussion paper

Thank you for pointing this out.

**Changes**: We will change the figure caption accordingly.

### 2.13 Comment # 2.13

*In the "Discussion" section, more comparisons with other existing methods for urban flooding information extraction from surveillance images should be added.*

Thank you for this comment. We were originally not sure how much to develop comparisons, but are happy to do add a few more.

**Changes**: We will add additional comparisons to existing methods for urban flood information extraction from surveillance images, showing both advantages and disadvantages of the alternatives. The publication of Jiang et al. (2019) proposed in comment # 2.2 will be included.

### 2.14 Comment # 2.14

*The accuracy of this approach could be affected by camera inclination and manual labeling. Topography can also affect the accuracy, e.g., in flat areas, the visible area of the flooding varies greatly, but the change of water depth value may be very small, while in low-lying areas, the change of visible area of the flooding is small, but the change of water depth value may be very big. These error sources should be discussed in the "Discussion" section.*

Thank you for this comment. It is true that a variety of factors can affect the accuracy of our approach. Although some error sources are already mentioned in the Discussion, it is true that we could have developed this further.

**Changes**: we will add an additional discussion of influencing factors after Section 4.1.

2.15 Comment # 2.15

*In the "Conclusions" section, the authors should explain the direct applications and managerial implications of this approach.*

Thank you for this comment. We agree that a broader perspective was lacking in the Conclusions section.

**Changes**: In the Conclusions, we will relate how our method could be applied in practice and how it could change the management of urban water systems.



Printer-friendly version

Discussion paper