A deep learning technique-based data-driven model for accurate and rapid flood prediction

Qianqian Zhou, Shuai Teng, Xiaoting Liao, Zuxiang Situ, Junman Feng, Gongfa Chen

School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou,

5 510006, China

Correspondence to: Qianqian Zhou (qiaz@foxmail.com)

Abstract. An accurate and rapid urban flood prediction model is essential to support decision-making on flood management, especially under increasing extreme precipitation conditions driven by climate change and urbanization. This study developed a deep learning technique-based data-driven flood

10 prediction model based on an integration of LSTM network and Bayesian optimization. A case study in north China was applied to test the model performance and the results clearly showed that the model can accurately predict flood maps for various hyetograph inputs, meanwhile with substantial improvements in computation time. The model predicted flood maps 19,585 times faster than the physical-based hydrodynamic model and achieved a mean relative error of 9.5%. For retrieving the spatial patterns of

15 water depths, the degree of similarity of the flood maps was very high. In a best case, the difference between the ground truth and model prediction was only 0.76% and the spatial distributions of inundated paths and areas were almost identical. The proposed model showed a robust generalizability and high computational efficiency, and can potentially replace and/or complement the conventional hydrodynamic model for urban flood assessment and management, particularly in applications of real

20 time control, optimization and emergency design and plan.

1 Introduction

Flooding has been one of the most frequent and disturbing disasters in many urban areas, especially under impacts of climate change and urbanization (Arnone et al., 2018; Zhou et al., 2019; Kaspersen et al., 2017; Mahmoud and Gan, 2018). Prediction of flooding plays a key role in urban flood evaluation

25 and management and can provide effective decision aid tools to reduce flooding impacts on both society and environment (Lowe et al., 2017; Xie et al., 2017; Hou et al., 2021a). Establishing rapid and accurate flood prediction methods is thus essential, however, a complicated and challenging task (Guo et al.,

2021). Conventionally, hydrodynamic models have been employed for applications such as flood inundation simulations, assessment of mitigation and adaptation measures (Wolfs and Willems, 2013;

30 Wang et al., 2021; Li et al., 2019). Despite the fact that the physically-based models can well simulate the drainage and surface flooding processes, they require a large number of inputs to describe the model structure and parameters, and are often computational intensive, especially with the two-dimensional calculations (Yin et al., 2020; Jamali et al., 2018; Ziliani et al., 2019; Hou et al., 2021b). Meanwhile, there is an inevitable need for conceptualization and simplification in the physical model, and the

35 relevant calibration and validation procedures are also quite challenging (Davidsen et al., 2017; Coulthard et al., 2013; Wu et al., 2018).

To solve the bottleneck problems on complex model construction, long computation time and high cost of the hydrodynamic models, the potential of novel deep learning techniques in capturing and predicting flooding processes have been increasingly explored in recent years to alleviate the burden on

40 physical modelling (Han et al., 2021; Guo et al., 2021; Hou et al., 2021a). The deep learning (DL) methods harbor intelligent learning mechanisms and can extract learning data features from historical knowledge. The DL can find the relationships between input and output data with much lower computational cost, in particular with high-performance computers. It has been demonstrated that these DLs have excellent generalization capabilities so that even complex data features (e.g., flood pattern and

45 tendency) can be automatically learned with a high prediction accuracy and computation efficiency (Lecun and Bengio, 1995; Rawat and Wang, 2017; Guo et al., 2021; Yosinski et al., 2014). With proper data provided, the methods can learn the flood patterns through data features and eliminate the analysis of the actual physical processes. The high computation efficiency is essential, especially for modelling flooding impacts in urban areas with complex local conditions and high spatial resolutions.

50 A number of early attempts on deep learning applications are found in the field of drainage and flood condition detection and assessment. Bhola et al. (2019) employed computer vision algorithms and edge detection techniques to identify water levels from images as supplementary validation data for flood forecasting. Moy De Vitry et al. (2019) used a deep convolutional neural network (DCNN) approach for scalable flood level trend monitoring with data from surveillance camera systems. Han et al. (2021)

55 proposed a YOLO-based deep learning framework to automatically monitor the urban road inundation under both dry and wet conditions. Hou et al. (2021c) performed an experimental flooding test using surveillance cameras to obtain flood images and employed a Mask R-CNN (mask region-based

convolutional neural network) to detect and segment the inundated areas in river channels. Guo et al.
(2021) adopted a DCNN-based approach for urban flood prediction and was reported to achieve

60   satisfying prediction accuracy and computation efficiency. Zhu et al. (2020) developed a probabilistic
long short-term memory (LSTM) network coupled with Gaussian process (GP) to improve the
streamflow forecasting in the upper Yangtze River. Note that different from other popular deep learning
algorithms, the LSTM network allows inputs of unequal dimensions/ lengths, which is especially suitable
for processing time-series data, such as traffic flow (Xia et al., 2021) and power systems (Ciechulski and

65   Osowski, 2021). All these studies have demonstrated remarkable capabilities of DL in automated data
feature learning with high prediction accuracy and efficiency. The reliability of the methods was also
verified in the various types of applications.

        Despite the advances of the studies, most of which focused on relatively large spatial scales and
required several types of input data (e.g., rainfall, terrain, flow depth) for model predictions. So far, no

70   study has explored the automated prediction of urban-scale flood inundation using the LSTM-based deep
learning techniques. This goal of this study is to provide a novel end-to-end method for a dynamic, rapid
and accurate urban flood prediction for real-time evaluation and emergency decision-making. Given the
uncertainty/unknown of rainfall events and the advantages of LSTM, we present a deep learning-based
technique with an integration of LSTM network and Bayesian optimization. The inundation areas and

75   water depths can be forecasted with only rainfall inputs. The method is tested in a case study in northern
China with various rainfall conditions. The developed method showed very promising prediction
accuracy and low computation cost and is thus of great value to be used as decision making aids in urban
flood evaluation and management.

2 Methodology and data

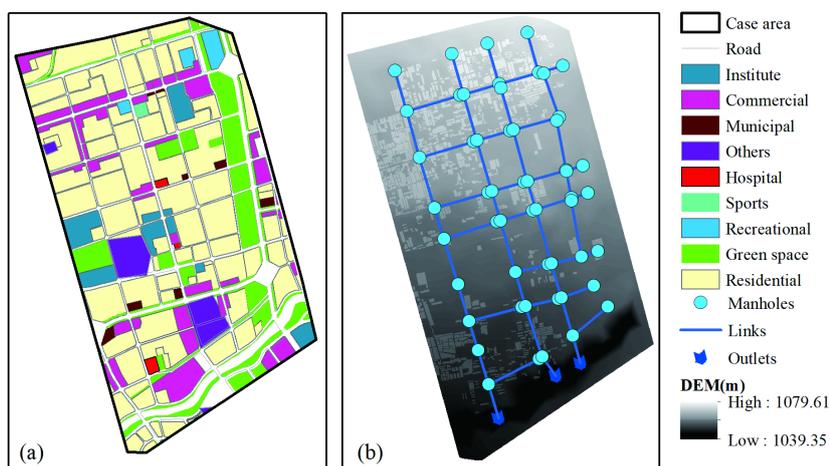
80   To examine the performance of the proposed approach, we firstly selected a case study and obtained the
relevant data describing the rainfall inputs, local topography and drainage systems. A coupled 1D-2D
hydrodynamic model was employed to simulate the inundation areas and water depths under various
design rainfall events. Then the DL technique-based prediction model was established and trained based
on the simulated flood maps and tested with random rainfall inputs to examine the relevant prediction

85   accuracy and computation cost.

2.1 Case study area

A portion of the city Hohhot, the capital of the Inner Mongolia Autonomous Region, was used as the case study to test the performance of proposed method. The city is located in Northern China and within a cold semi-arid climate zone. The winters are dry but the summers can be very hot and rainy. The average

90   annual rainfall was approximately 396 mm, with majority of which concentrated from July to August (Zhou et al., 2018; Zhou et al., 2016). The detailed landuse is shown in Fig. 1a and mainly consists of residential areas, commercial districts, institutes, green spaces and other landuse. The terrain is high in the north and lower in the south (see Fig. 1b) and thus the runoffs generally flow in a north to south direction. The service level of the drainage system was rather low and the original design return period

95   was below once a year (Zhou et al., 2018). In recent years, flooding has occurred more frequently in the area. Nevertheless, there is a lack of accurate historical data on flood areas and depths and thus simulations of flood events are performed with a 1d/2D coupled hydrodynamic model (to be introduced in the following sections) under various design rainfalls.

100   **Figure 1: case study (a) landuse and (b) drainage system with DEM descriptions.**

The input rainfall hydrographs for model training and validation were calculated using the regional Storm Intensity Formula (SIF) ($q=635\times(1+0.841\times lg(P))/t^{0.61}$, where q is the storm intensity ((L/s)/hm^2), p is the design return period (a) and t is the rainfall duration (minutes), respectively) (Zhang and Guan, 2012; Zhou et al., 2016). The rainfall calculation follows the national code for design of

105   outdoor drainage (Mohurd, 2016) and the design principles of Chicago Design Storms (Berggren et al., 2014; Panthou et al., 2014; Zhou et al., 2012). The detailed procedures in applying the regional SIF to

obtain CDSs are outlined in the national Technical Guidelines for Establishment of Intensity-Duration-Frequency Curve and Design Rainstorm Profile (Mohurd, 2014). In this study, we adopted in total 90 rainfall events, with return periods ranging from 2 to 100 years and rainfall duration of

110   2, 4 or 6 hours, respectively. All rainfall inputs were generated with a temporal resolution of 10 minutes.

2.2 Physically-based hydrodynamic model

The overland flooding maps under the 90 rainfall inputs were simulated using the 1D-2D coupled hydrodynamic model, namely Mike Urban & Mike Flood (Mike by Dhi, 2016a). The hydrological inputs and pipe flows were simulated using the 1D Mouse model and the overland flows were calculated using

115   the Mike 21 model. With the precipitation inputs, the runoff model was described by the general catchment data, such as locations, areas, imperviousness, and time of concentration. The shape of the runoff hydrograph was computed by the 'Time-Area' method (Mike by Dhi, 2016b). The calculation of unsteady flow in the pipe network was done by solving the Saint Venant equations, that are the vertically integrated equations of conservation of continuity (i.e., continuity equation) and momentum (i.e.,

120   momentum equation) (Mike by Dhi, 2016c). The surface flow model was described by the MIKE 21 rectangular grid component and links between the 1D and 2D models were established to simulate the flow interactions between the pipeline and overland flows. Especially, the flow in the links is governed by an orifice equation (Mike by Dhi, 2016d). When the underground drainage is surcharged, the excess water will flow to the surface and conduct surface inundation calculations under the context of extreme

125   precipitation. On the surface, the water typically flows along buildings or streets based on a description of the local digital elevation/topography (Mark et al., 2004; Leandro et al., 2009).

       Model outputs include overland flow paths, extents, depths and velocities at different time steps. One of the most commonly used outputs is the flood maps describing the maximum water depths caused by the given rainfall inputs (Kaspersen et al., 2017; Mike by Dhi, 2016a; Zhou et al., 2012). These flood

130   maps can be further integrated with vulnerability data for an assessment of flood risk levels at different spatial scales (Sampson et al., 2014; De Moel et al., 2009; Ashley et al., 2007). In doing so, critical areas with higher levels of flood risks can be identified and allocated with priorities in mitigation and adaptation plans (De Moel et al., 2015; Zhou et al., 2012). As shown in Fig. 2 that changes in input rainfalls lead to variations in simulated flood maps. Increases in flood extents and depths are seen with

135   rainfalls of larger return periods in the case study. Specifically, there were in total 90 flood maps output

as the dataset for deep learning model. Among that, 90% of the flood maps were randomly selected for model training and validation, and the rest 10% for testing. That means, for the deep learning model, all the tested hyetographs were the inputs and all the simulated flood maps were ground truth (GT) data to train the model network. After the training, the randomly sampled 10% flood maps were used to test the

140   model prediction performances.

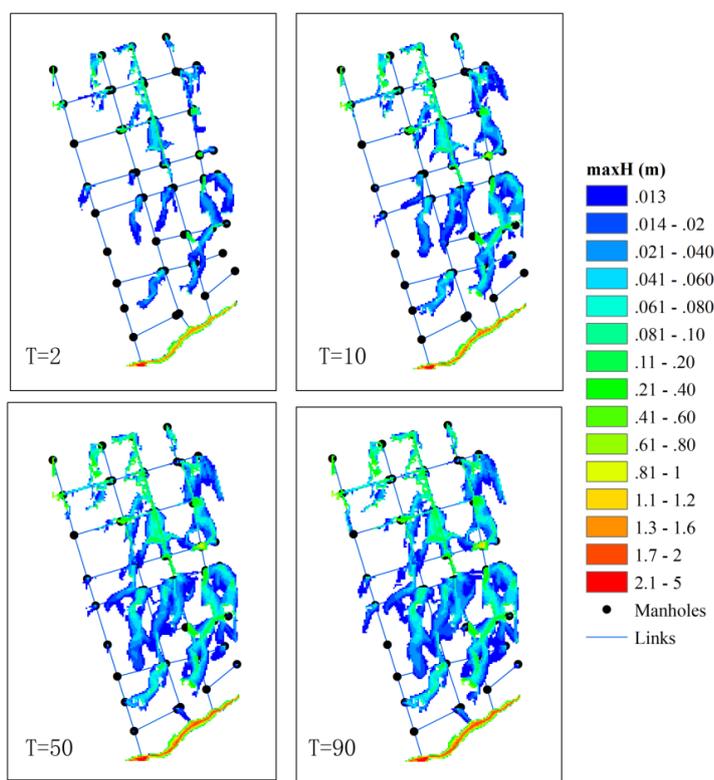



Figure 2: Maximum flood maps simulated using hydrodynamic models for four return periods.

2.3 Deep learning model for flood prediction

2.3.1 LSTM (Long Short Term Memory) network

145   The LSTM network has advantages in processing time-series data, especially for the long-term memory of data. As shown in Fig. 3 that the LSTM network is used to predict the flood maps, with the rainfall intensity series used as the network input. Ideally, the network can predict the flood depth distributions in the region as close as the real values/ ground truth (GT) values. The relative error between the output and GT values is calculated and used as a priori condition for the Bayesian optimization (BO). Finally, an

150 optimal network model (e.g., with an appropriate number of layers) is obtained through the iterative BO

process.

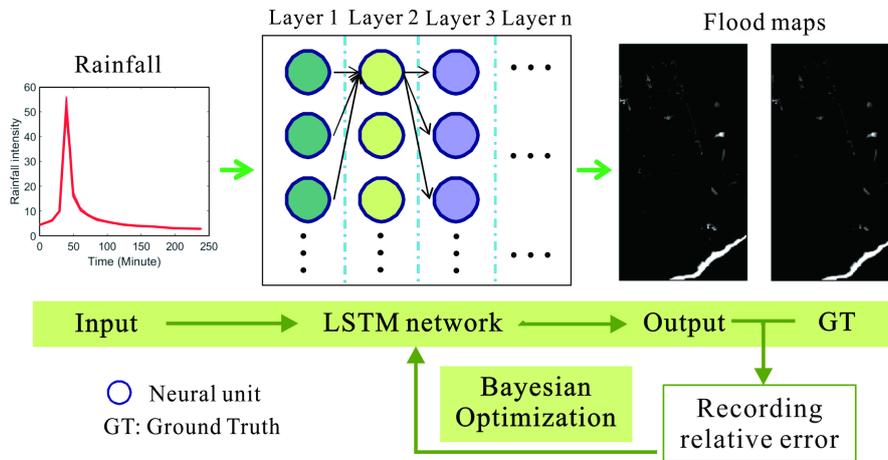



Figure 3. Method implementation steps.

A benchmark LSTM network structure is shown in Fig. 4. With the input data (rainfall intensity),

155 the LSTM gets the output (water depth) through a series of functional layers, including a LSTM layer

(containing N neural units), a Leaky ReLU activation function (Eq. (1)), and a fully connection (FC)

layer. In the LSTM layer, the rainfall is input to N neural units and N outputs (i.e., h_0, h_1, h_2, …, h_N) are

obtained. The outputs of these neural units are then transformed nonlinearly by the Leaky ReLU

activation function and enter the FC layer. Eventually, the FC layer delivers the output of the network.

160 Specifically, the LSTM network is trained through the adaptive moment estimation (Adam) optimizer

(Eqs. (2-4)). Meanwhile, two performance indicators were used to reflect the network training, which are

the loss (Eq. (5)) and root mean square error (RMSE, Eq. (6)) respectively.

$$f(x) = \begin{cases} x, & x \geq 0 \\ scale \times x, & x < 0 \end{cases} \tag{1}$$

$$m_l = \beta_1 m_l + (1 - \beta_1)\nabla E(\theta_l) \tag{2}$$

$$v_l = \beta_2 v_{l-1} + (1 - \beta_2)[\nabla E(\theta_l)]^2 \tag{3}$$

$$\theta_{l+1} = \theta_l + \frac{\alpha m_l}{\sqrt{v_l + \varepsilon}} \tag{4}$$

$$Loss = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)} \tag{6}$$

Where, x and $scale$ are input and scale factor (0.01), receptively. Any input value that is less than zero

is multiplied by a fixed scale factor. β_1 and β_2 are the Gradient Decay Factor (0.9) and Squared

165   Gradient Decay Factor (0.999), respectively. $E(\theta)$ is the loss function, m and v are the momentum terms,

and $\varepsilon=10^{-8}$. n is the number of samples, y_i and $\hat{y}_i$ are the predicted and real results, respectively.

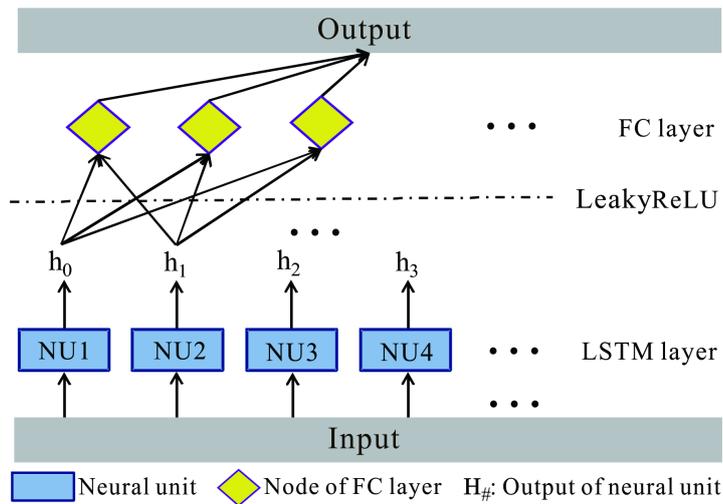



Figure 4: Framework of the LSTM network.

The neural unit is a key component of the LSTM network and the structure of a single neural unit is

170   shown in Fig. 5, including a forget gate (Eq. (7)), an input gate (Eqs. (8-10)) and an output gate (Eqs.

(11-12)). The forget gate determines how many unit states at time (t-1) are retained until time (t). The

input gate determines the update of the unit states. The output of the LSTM neural unit state is

determined by the nonlinear activation function (Sigmoid, in Eq. (13)) and the output gate. In general, an

input (x) passes through a neural unit to get an output (h). Specifically, the calculation process of a single

175   LSTM neural unit is shown as follows:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \tag{7}$$

$$i_t = \sigma(\boldsymbol{W}_i[h_{t-1}, x_t] + b_i) \tag{8}$$

$$c'_t = \tanh(\boldsymbol{W}_c[h_{t-1}, x_t] + b_c) \tag{9}$$

$$c_t = f_t c_{t-1} + i_t c'_t \tag{10}$$

$$o_t = \sigma(\boldsymbol{W}_o[h_{t-1}, x_t] + b_o) \tag{11}$$

$$h_t = o_t \times \tanh(c_t) \tag{12}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{13}$$

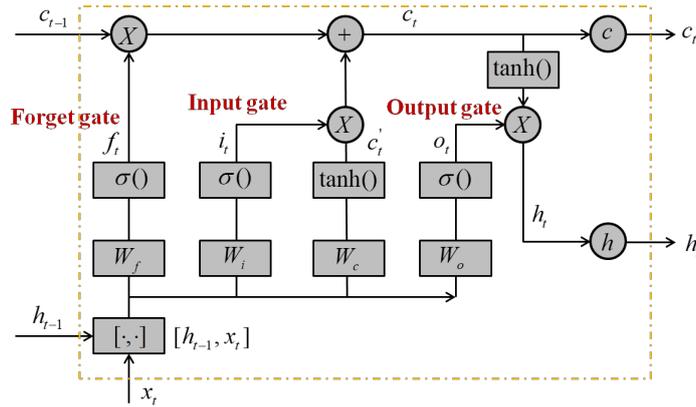
Where, f_t is the output of the forget gate, $\boldsymbol{W}_f$ and b_f are the weight matrix and bias of the forget gate, and h_{t-1} and x_t are the output of the previous neural unit (time (t-1)) and the current input (time (t)), respectively. i_t is the output of the input gate, c'_t and c_t are the unit state of the current input and current time, respectively. o_t is the output of the output gate, h_t is the neural unit output of time (t).

180

Figure 5: The structure of single neural unit.

In this paper, the LSTM network was built in the MATLAB 2021a (MathWorks Inc, Natick, MA, US). The input was the rainfall intensity varying with time, and the output was the flood map of the case study site. The training platform was performed on a computer with NVIDIA RTX GeForce 2060 GPU,

185 Intel Core i7-4790 @ 3.60 GHz CPU, Windows 10.

2.3.2 Bayesian optimization

One problem with the aforementioned LSTM network is that its structure layers, learning rate, number of training epochs, mini-BatchSize and number of neural units were all unknown. To start from scratch, it can be very difficult and time consuming to manually select and fine-tune these hyper-parameters.

190 Bayesian optimization (BO) is an algorithm that can automatically search for the optimal hyper-parameter combinations. The BO is a continuously updated probability model (Eq. (14)) and assumes that the probability of occurrence of Event A under the a priori condition of Event B is directly proportional to the probability of occurrence of the a posteriori condition of Event B. That is, for successively occurring events, the latter events are related to all previous events. It is a potential

195 hyper-parametric optimization scheme, which means the most likely parametric combination is inferred through a number of a priori attempts (i.e., training network models with different structures).

The posterior probability of the optimization function is updated through a number of evaluations of objective function to obtain the optimal parameter combination. It can provide reference for the subsequent tried models according to the a priori conditions (i.e., historical evaluation records, which are

200 the mean relative errors of the tried network model in this paper). When selecting the next group of parameter combinations, the algorithm made full use of the previous evaluation information to reduce the search time of the parameters. Specifically, we designed a variety of search ranges of the hyper-parameters and BO algorithm automatically took the values from the search ranges and constantly tried the network models with different structures, and then recorded the errors. In this paper, the

205 hyper-parameters to be optimized included the number of LSTM layer, learning rate, Epoch, mini-BatchSize, and number of hidden units. The search ranges of these five parameters were set to [1-5], $[10^{-4}\text{-}1]$, [0-600], [0-100] and [0-100], respectively. Finally, BO inferred the possible optimal network combination according to the historical error information. The selection process is shown in Eq. (15).

$$P(A|B) \propto P(B|A)P(A) \tag{14}$$

$$x^* = \arg\min_{x \in \chi} f(x) \tag{15}$$

Where, $P(A|B)$ and $P(A)$ are the posterior and prior probabilities of Event A, respectively, and

210 $P(B|A)$ is the observation point probability obtained from the previous events. $f(x)$ is the objective

function (i.e., the mean relative error (Eq. (16)), see the next section), x^* is the optimal parametric combination, and χ is the value range of parameters.

2.3.3 Performance indicators

215   In order to evaluate the reliability of the proposed method, five indexes were employed to evaluate the prediction results, focusing on estimating the differences in flood depths and the spatial patterns of the flood distributions. First of all, the mean relative error (Mre) was used to calculate the depth error between the prediction results (PR) and the ground truth (GT). Next, the 2-D correlation coefficient (2D-CC) and structural similarity (SS) were used to evaluate the correlation and similarity of images (distributions of flood areas), respectively. Further, the Bhattacharyya distance (BD) and Histogram

220   Intersection Distance (HID) measure the similarity of two discrete or continuous probability distributions. They were adopted to measure the amount of overlap between two statistical samples or images (distribution of water area).

$$Mre = \sum \frac{|(PR - GT)|}{GT} \tag{16}$$

$$2D - CC(I,J) = \frac{\sum_m \sum_n (I_{mn} - \bar{I})(J_{mn} - \bar{J})}{\sqrt{(\sum_m \sum_n (I_{mn} - \bar{I})^2)(\sum_m \sum_n (J_{mn} - \bar{J})^2)}} \tag{17}$$

$$SS(I,J) = \frac{(2\mu_I \mu_J + C_1)(2\sigma_{IJ} + C_2)}{(\mu_I^2 + \mu_J^2 + C_1)(\sigma_I^2 + \sigma_J^2 + C_2)} \tag{18}$$

$$BD(I,J) = -\ln(\sum_{x \in X} \sqrt{p(x)q(x)}) \tag{19}$$

$$HID(I,J) = \frac{\sum_{x \in X} \min(p(x), q(x))}{\sum_{x \in X} p(x)} \tag{20}$$

Where, $\bar{I}$ and $\bar{J}$ are the average pixel values of Image I and J, respectively, μ_x, μ_y, σ_x σ_y and σ_{xy} are the pixel local mean, standard deviation and cross covariance of Image I and J, respectively. C_1 and C_2

225   were 6.5 and 58.5 respectively. $p(x)$ and $q(x)$ are probability distributions of pixels of Image I and Image J, respectively. X is the domain of $p(x)$ and $q(x)$.

3 Results and discussion

An illustration of the mean relative error of the testing dataset obtained from the 100 Bayesian optimizations is shown in Fig. 6a. The range of mean error is between 0.095 and 0.73 and the details on

230   the error value of each number of BO optimization can be seen by the size of the bubble plot. Especially, one of the networks, with a mean relative error value of 0.095, worked best in learning the flood map features. Figure 6b shows the RMSE and loss of the model with the best performance identified from the Bayesian optimization. It is shown that the loss curve stably decreased along the network training and the model achieved a convergence status after the 100 iterations with a small loss value. This implies that the

235   deep learning network is very robust and trained well with the input data. The detailed network structure of the optimized model is shown in Fig. 7. The learning rates, Epoch, mini-BatchSize, and number of hidden units were 0.0146, 385, 59 and 94 respectively.

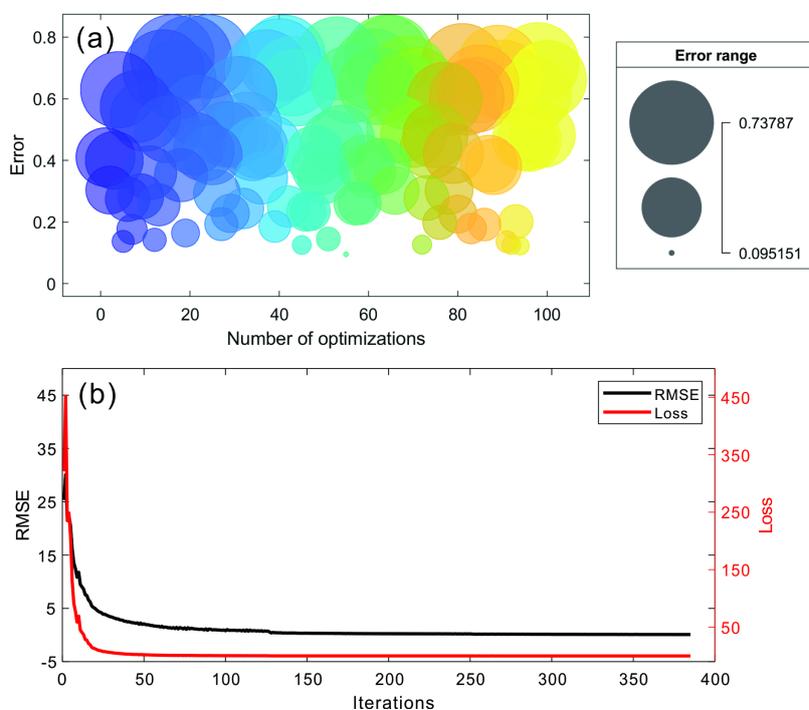



Figure 6: (a) the mean relative errors obtained from the 100 Bayesian optimizations, and (b) the RMSE and
240   **loss achieved by the model with the best performance.**

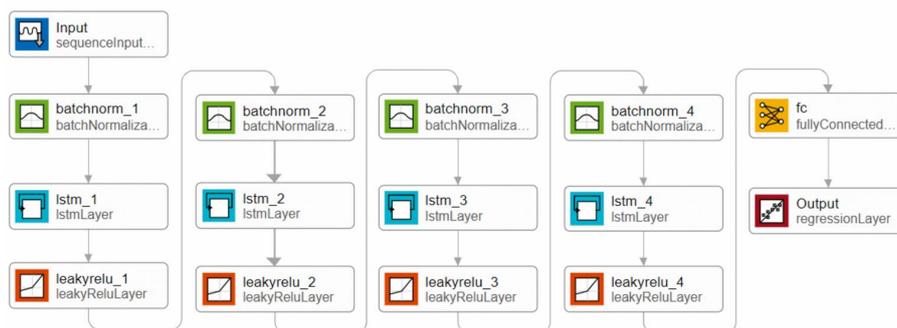



Figure 7: The optimized model structure of the LSTM network. Batchnorm_#: Batch normalization layer, normalize the network training data (mapping raw data to [0, 1]) to speed up the training speed.

We further analyzed the statistics of the performance indicators of the best performing model.
245   Boxplots of the specific relative error of each testing flood map is shown in Fig. 8a. As shown previously, the deep learning model obtained satisfying results with a mean relative error of 9.5%. The achieved minimum RE of a single prediction is 0.76%, which implies the predicted flood map (both the inundation locations and depths) is very close to the ground truth map for validation. Through statistics, the degree of similarity is illustrated by the four types of parameters in Fig. 8b. First of all, the Bhattacharyya

250   distances of the testing dataset were all close to zero, which means that the spatial distributions of the ground truth and predicted flood hazard maps were very similar and a majority of the two map populations were overlapped. The good results were further verified by the Histogram intersection distance, structural similarity and 2D correlation coefficient as their values were all close to one. This implies that the spatial similarity of the predicted maps was very high. On the whole, Figure 8a and 8b

255   indicate that the model is superior in learning and predicting the flood maps with different hyetographs.

The computation times of the hydrodynamic model and the deep learning model are compared in Fig. 8c. The average computation time of the hydrodynamic model was 153.2s, while the mean time of the prediction model was significantly reduced, with a value of 0.038s. It is shown in Fig. 8d that the hydrodynamic model took almost 19,585 times (i.e., mean value) the simulation time of the deep learning

260   model. In the worst case, the hydrodynamic model simulates the flood map more than 36,600 times slower. Note that in fact, the computation time of the hydrodynamic model was even longer, as the model needs to run the hydrological and pipe-network+2D simulations separately and a manual integration of the two simulations were not taken into account. The results show that with proper model training, the

deep learning model is accurate and much more computational efficient, which can provide important

265 support tools for real-time/rapid forecasting of urban flooding and emergency decision-making.

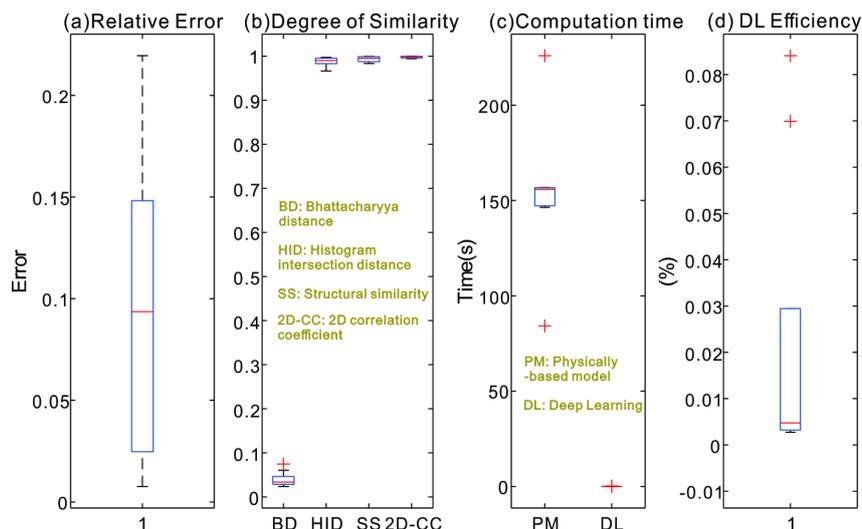



Figure 8: (a) relative error, (b) degree of similarity, (c) computation time and (d) deep learning (DL) efficiency (i.e., computation time of DL model divided by computation time of hydrodynamic model) achieved with the testing datasets.

270 In visual quality, Figure 9 illustrated the inundated areas of the ground truth and the predicted flood

maps with the best model performance (i.e., with the minimum relative error). In total there are 27,183

grids in the flood map. It can be seen that the deep learning model successfully retrieved the depths and

spatial patterns of the inundated areas. The two maps were almost identical and it is very difficult to tell

the difference without looking into further statistic details. Figure 10a shows the corresponding spatial

275 distributions of the relative errors of the flood map. The differences between the two maps were almost

negligible except the highlighted regions near the water bodies. The predicted flood map could identify

all the flow paths and local depressions as the ground truth map. Meanwhile, the spatial distributions of

the mean relative errors of the testing dataset are shown in Fig. 10b. Results indicate that there is a good

agreement between the series of predictions and the ground truth maps. The distribution of the mean

280 errors is even in most areas, with an error below 1%. We also note that the errors are greater where there

are higher water depths and more flow volumes. In the maximum case, the relative error of predicted

water depths near the water bodies can reach above 10% (Fig. 10c).

Ground Truth   Model Prediction

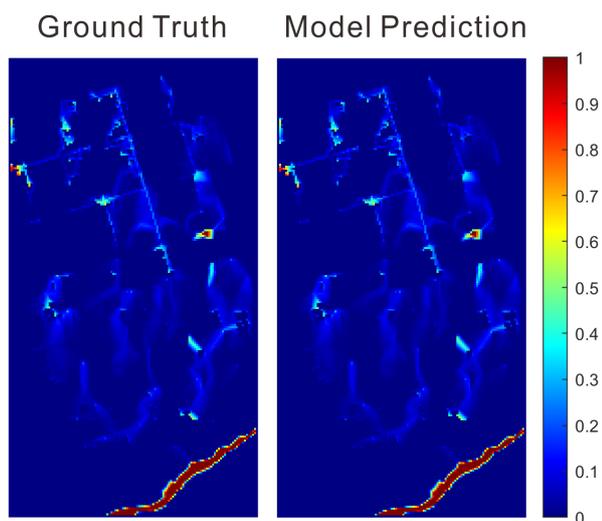



Figure 9: Flood inundation maps simulated based on Mike Urban & Mike Flood.

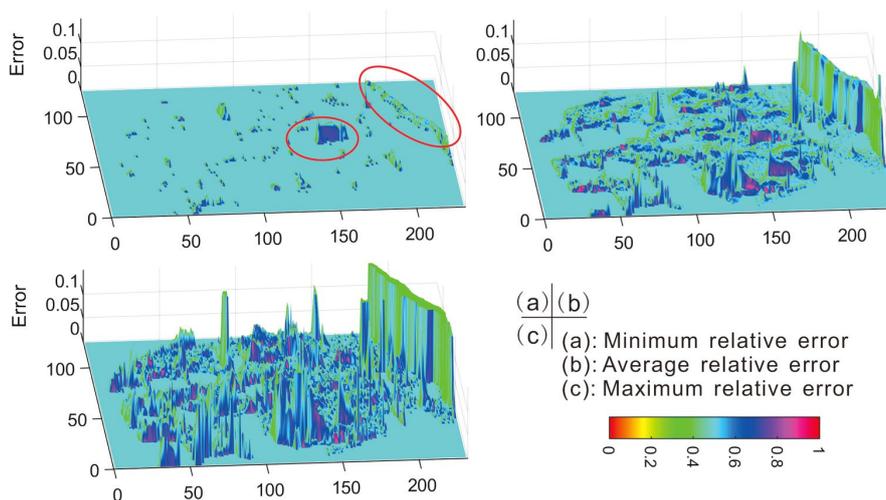
(a)|(b)
(c)| (a): Minimum relative error
        (b): Average relative error
        (c): Maximum relative error

285

Figure 10: Flood inundation maps simulated based on Mike Urban & Mike Flood.

The prediction accuracies of the deep learning model are further examined as a function of water depths in Fig. 11. Results show that the flood map dataset is imbalanced as a majority of the results contain no and shallow water. Results show that for water depth below 3m, the model performed well and most errors were below 2%. The errors tended to increase under extreme conditions, with water

290   depths above 3.5m. Figure 11b shows that the predicted water depths are basically consistent with the ground truth water depths. These results clearly indicate that the deep learning model generalizes well

with the different hyetograph variations and can produce very accurate flood results even with only rainfall inputs.

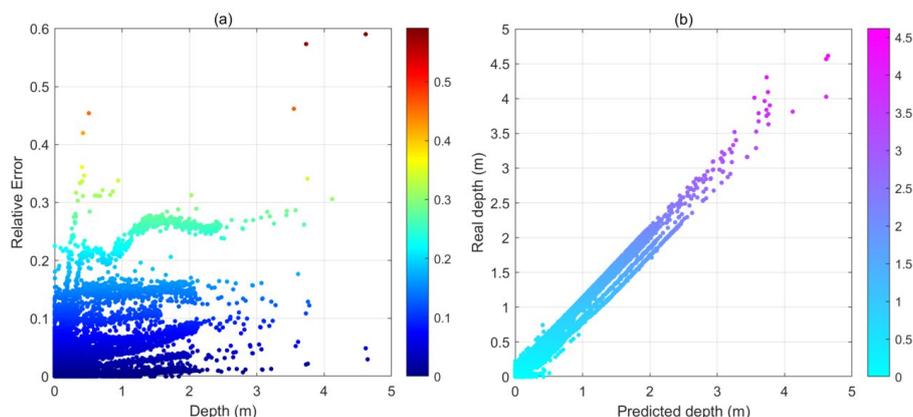



Figure 11: (a) Relative errors of predicted flood maps as a function of water depths, and (b) ground truth water depths as a function of predicted water depths.

4 Conclusions

A rapid, accurate and dynamic flood prediction tool is of great significance to urban water management to protect people, social assets and environment from flood hazards. This study proposed a deep learning technique-based data-driven flood prediction approach, employing an integration of the LSTM technique and Bayesian optimization approach. Results clearly show that model can accurately produce flood maps for various hyetograph inputs with much lower computation costs. The presented model showed a robust generalizability and predicted the flood maps 19,585 times faster than the hydrodynamic model. The achieved mean relative error in water depths is 9.5% and the degree of similarity of flood maps was very high. Specifically, in a best case, the difference between the ground truth and model prediction was only 0.76% and the spatial patterns of the two types of maps were almost identical. In conclusion, the accuracy and efficiency of the proposed method is satisfying.

We acknowledge some limitations in this study and discuss directions of future work. First of all, the current training and testing data were obtained from hydrodynamic modelling due to a lack of detailed field site data. In future work, we consider adopting image capture techniques for data supplement, such as deep learning techniques for automated detection, acquisition and evaluation of water depths from camera images. In doing so, there will be more real case/field survey dataset for model

315 training and testing. Meanwhile, data augmentation is useful in enhancing the quantity and quality of input data, which will be tested in future investigations. In addition, the deep-learning model currently only predicts the maximum water depths caused by the rainfall inputs. This means that the temporal changes in water depths during the rainfall events were not considered. Further research to predict the dynamic changes in both temporal and spatial scales is of great interest.

320 Despite the limitations, this work with its advances can well contribute to a better understanding of the deep learning techniques for urban flood mapping. The proposed methodology predicts water depths with only rainfall inputs, without further requirements of e.g., local terrains and geographical conditions. The approach can be easily adjusted or adopted for other types of applications in water management field. More importantly, the proposed method can potentially replace and/or complement the conventional detailed hydrodynamic model for urban flood assessment and management, particularly in applications

325 of real time control, optimization and emergency design and plan.

Data availability

The dataset that support the findings of this study are available from the corresponding author upon reasonable request.

Author contribution

330 QZ conceived the idea and acquired the project and financial support. QZ and ST designed the study. XL and JF collected and preprocessed the data. QZ and ST conducted all the experiments and analyzed the results. QZ wrote the first draft of the manuscript with contributions from ST. XL, JF, ZS and GC provided feedback on results and edited the manuscript.

Competing interests

335 The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

340 This research was funded by the National Natural Science Foundation of China (Grant No. 51809049), the National College Students Innovation and Entrepreneurship Training Program (Grant No. 202111845038).

References

Arnone, E., Pumo, D., Francipane, A., La Loggia, G., and Noto, L. V.: The role of urban growth,
345 climate change, and their interplay in altering runoff extremes, Hydrol. Process., 32, 1755-1770, 10.1002/hyp.13141, 2018.

Ashley, R., Garvin, S., Pasche, E., Vassilopoulos, A., and Zevenbergen, C.: Advances in Urban Flood Management, 2007.

Berggren, K., Packman, J., Ashley, R., and Viklander, M.: Climate changed rainfalls for urban drainage
350 capacity assessment, Urban Water Journal, 11, 543-556, 10.1080/1573062x.2013.851709, 2014.

Bhola, P. K., Nair, B. B., Leandro, J., Rao, S. N., and Disse, M.: Flood inundation forecasts using validation data generated with the assistance of computer vision, Journal of Hydroinformatics, 21, 240-256, 10.2166/hydro.2018.044, 2019.

Ciechulski, T. and Osowski, S.: High Precision LSTM Model for Short-Time Load Forecasting in
355 Power Systems, Energies, 14, 2983, 2021.

Coulthard, T. J., Neal, J. C., Bates, P. D., Ramirez, J., de Almeida, G. A. M., and Hancock, G. R.: Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: implications for modelling landscape evolution, Earth Surface Processes and Landforms, 38, 1897-1906, 10.1002/esp.3478, 2013.

360 Davidsen, S., Lowe, R., Thrysoe, C., and Arnbjerg-Nielsen, K.: Simplification of one-dimensional hydraulic networks by automated processes evaluated on 1D/2D deterministic flood models, Journal of Hydroinformatics, 19, 686-700, 10.2166/hydro.2017.152, 2017.

de Moel, H., van Alphen, J., and Aerts, J. C. J. H.: Flood maps in Europe - methods, availability and use, Natural Hazards and Earth System Sciences, 9, 289-301, 2009.

365   de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E., and Ward, P. J.: Flood risk assessments at different spatial scales, Mitigation and Adaptation Strategies for Global Change, 20, 865-890, 10.1007/s11027-015-9654-z, 2015.

Guo, Z. F., Leitao, J. P., Simoes, N. E., and Moosavi, V.: Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks, Journal of Flood Risk Management, 14,

370   10.1111/jfr3.12684, 2021.

Han, H., Hou, J. M., Bai, G. G., Li, B. Y., Wang, T., Li, X., Gao, X. J., Su, F., Wang, Z. F., Liang, Q. H., and Gong, J. H.: A deep learning technique-based automatic monitoring method for experimental urban road inundation, Journal of Hydroinformatics, 23, 764-781, 10.2166/hydro.2021.156, 2021.

Hou, J. M., Zhou, N., Chen, G. Z., Huang, M. A. S., and Bai, G. B.: Rapid forecasting of urban flood

375   inundation using multiple machine learning models, Natural Hazards, 108, 2335-2356, 10.1007/s11069-021-04782-x, 2021a.

Hou, J. M., Ma, Y. Y., Wang, T., Li, B. Y., Li, X., Wang, F., Jin, S. L., and Ma, H. L.: A river channel terrain reconstruction method for flood simulations based on coarse DEMs, Environmental Modelling & Software, 140, 10.1016/j.envsoft.2021.105035, 2021b.

380   Hou, J. M., Li, X., Bai, G. G., Wang, X. H., Zhang, Z. X., Yang, L., Du, Y. E., Ma, Y. Y., Fu, D. Y., and Zhang, X. G.: A deep learning technique based flood propagation experiment, Journal of Flood Risk Management, 14, 10.1111/jfr3.12718, 2021c.

Jamali, B., Löwe, R., Bach, P. M., Urich, C., Arnbjerg-Nielsen, K., and Deletic, A.: A rapid urban flood inundation and damage assessment model, Journal of Hydrology, 564, 1085-1098,

385   https://doi.org/10.1016/j.jhydrol.2018.07.064, 2018.

Kaspersen, P. S., Ravn, N. H., Arnbjerg-Nielsen, K., Madsen, H., and Drews, M.: Comparison of the impacts of urban development and climate change on exposing European cities to pluvial flooding, Hydrology and Earth System Sciences, 21, 4131-4147, 10.5194/hess-21-4131-2017, 2017.

Leandro, J., Chen, A. S., Djordjevic, S., and Savic, D. A.: Comparison of 1D/1D and 1D/2D Coupled

390   (Sewer/Surface) Hydraulic Models for Urban Flood Simulation, Journal of Hydraulic Engineering-Asce, 135, 495-504, 10.1061/(asce)hy.1943-7900.0000037, 2009.

LeCun, Y. and Bengio, Y.: Convolutional networks for images, speech, and time series, The handbook of brain theory and neural networks, 3361, 1995, 1995.

Li, W., Lin, K., Zhao, T., Lan, T., Chen, X., Du, H., and Chen, H.: Risk assessment and sensitivity

395   analysis of flash floods in ungauged basins using coupled hydrologic and hydrodynamic models,

Journal of Hydrology, 572, 108-120, https://doi.org/10.1016/j.jhydrol.2019.03.002, 2019.

Lowe, R., Urich, C., Domingo, N. S., Mark, O., Deletic, A., and Arnbjerg-Nielsen, K.: Assessment of

urban pluvial flood risk and efficiency of adaptation options through simulations - A new generation of

urban planning tools, Journal of Hydrology, 550, 355-367, 10.1016/j.jhydrol.2017.05.009, 2017.

400   Mahmoud, S. H. and Gan, T. Y.: Urbanization and climate change implications in flood risk

management: Developing an efficient decision support system for flood susceptibility mapping,

Science of The Total Environment, 636, 152-167, https://doi.org/10.1016/j.scitotenv.2018.04.282, 2018.

Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, S. B., and Djordjevic, S.: Potential and

limitations of 1D modelling of urban flooding, Journal of Hydrology, 299, 284-299, 2004.

405   MIKE by DHI: MIKE by DHI software, Release Note_MIKE URBAN, 2016, 2016a.

MIKE by DHI: MOUSE Reference Manual (Runoff) reference manual, 2016b.

MIKE By DHI: MOUSE Reference Manual (Pipe Flow) reference manual, MIKE By DHI Software,

Horsolm, Denmark., 2016c.

MIKE By DHI: MIKE FLOOD User Manual, MIKE By DHI Software, Horsolm, Denmark., 2016d.

410   MOHURD: Technical Guidelines for Establishment of Intensity-Duration-Frequency Curve and Design

Rainstorm Profile (In Chinese), Ministry of Housing and Urban-Rural Development of the People's

Republic of China and China Meteorological Administration. Accessed on October 2018 from

http://www.mohurd.gov.cn/wjfb/201405/W020140519104225.pdf, 2014.

MOHURD: Code for design of outdoor wastewater engineering (GB 50014—2006) (2016 revised

415   version)(In Chinese), Ministry Of Housing And Urban-Rural Development and Ministry Of National

Quality Standard Monitoring Bureau. Beijing, China 2016.

Moy de Vitry, M., Kramer, S., Wegner, J. D., and Leitão, J. P.: Scalable flood level trend monitoring

with surveillance cameras using a deep convolutional neural network, Hydrol. Earth Syst. Sci., 23,

4621-4634, 10.5194/hess-23-4621-2019, 2019.

420   Panthou, G., Vischel, T., Lebel, T., Quantin, G., and Molinie, G.: Characterising the space-time

structure of rainfall in the Sahel with a view to estimating IDAF curves, Hydrology and Earth System

Sciences, 18, 5093-5107, 10.5194/hess-18-5093-2014, 2014.

Rawat, W. and Wang, Z.: Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review, Neural Computation, 29, 2352-2449, 10.1162/neco_a_00990, 2017.

425   Sampson, C. C., Fewtrell, T. J., O'Loughlin, F., Pappenberger, F., Bates, P. B., Freer, J. E., and Cloke, H. L.: The impact of uncertain precipitation data on insurance loss estimates using a flood catastrophe model, Hydrology and Earth System Sciences, 18, 2305-2324, 10.5194/hess-18-2305-2014, 2014.

Wang, N., Hou, J. M., Du, Y. G., Jing, H. X., Wang, T., Xia, J. Q., Gong, J. H., and Huang, M. S.: A dynamic, convenient and accurate method for assessing the flood risk of people and vehicle, Science of

430   the Total Environment, 797, 10.1016/j.scitotenv.2021.149036, 2021.

Wolfs, V. and Willems, P.: A data driven approach using Takagi-Sugeno models for computationally efficient lumped floodplain modelling, Journal of Hydrology, 503, 222-232, 10.1016/j.jhydrol.2013.08.020, 2013.

Wu, X., Wang, Z., Guo, S., Lai, C., and Chen, X.: A simplified approach for flood modeling in urban

435   environments   Hydrology Research, 49, 2018.

Xia, D., Zhang, M., Yan, X., Bai, Y., Zheng, Y., Li, Y., and Li, H.: A distributed WND-LSTM model on MapReduce for short-term traffic flow prediction, Neural Computing and Applications, 33, 2393-2410, 2021.

Xie, K., Ozbay, K., Zhu, Y., and Yang, H.: Evacuation Zone Modeling under Climate Change: A

440   Data-Driven   Method,   Journal   of   Infrastructure   Systems,   23,   04017013, 10.1061/(ASCE)IS.1943-555X.0000369, 2017.

Yin, D., Evans, B., Wang, Q., Chen, Z., Jia, H., Chen, A. S., Fu, G., Ahmad, S., and Leng, L.: Integrated 1D and 2D model for better assessing runoff quantity control of low impact development facilities   on   community   scale,   Science   of   The   Total   Environment,   720,   137630,

445   https://doi.org/10.1016/j.scitotenv.2020.137630, 2020.

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H.: How transferable are features in deep neural networks?, Proceedings of the Advances in Neural Information Processing Systems, Montreal, 3320–3328, 2014.

Zhang, B. and Guan, Y.: Watersupply & Drainage Design Handbook, China Construction Industry

450   Press, Being, China. ISBN: 9787112136803, 2012.

Zhou, Q., Leng, G., Su, J., and Ren, Y.: Comparison of urbanization and climate change impacts on urban flood volumes: Importance of urban planning and drainage adaptation, Science of The Total Environment, 658, 24-33, https://doi.org/10.1016/j.scitotenv.2018.12.184, 2019.

Zhou, Q., Mikkelsen, P. S., Halsnaes, K., and Arnbjerg-Nielsen, K.: Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits, Journal of Hydrology, 414, 539-549, 10.1016/j.jhydrol.2011.11.031, 2012.

Zhou, Q., Ren, Y., Xu, M., Han, N., and Wang, H.: Adaptation to urbanization impacts on drainage in the city of Hohhot, China, Water Science and Technology, 73, 167-175, 10.2166/wst.2015.478, 2016.

Zhou, Q. Q., Leng, G. Y., and Huang, M. Y.: Impacts of future climate change on urban flood volumes in Hohhot in northern China: benefits of climate change mitigation and adaptations, Hydrology and Earth System Sciences, 22, 305-316, 10.5194/hess-22-305-2018, 2018.

Zhu, S., Luo, X., Yuan, X., and Xu, Z.: An improved long short-term memory network for streamflow forecasting in the upper Yangtze River, Stoch. Environ. Res. Risk Assess., 34, 1313-1329, 10.1007/s00477-020-01766-4, 2020.

Ziliani, M. G., Ghostine, R., Ait-El-Fquih, B., McCabe, M. F., and Hoteit, I.: Enhanced flood forecasting through ensemble data assimilation and joint state-parameter estimation, Journal of Hydrology, 577, 123924, https://doi.org/10.1016/j.jhydrol.2019.123924, 2019.