https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

A Bayesian updating framework for calibrating hydrological parameters of road network using taxi GPS data

Xiangfu Kong[1,2], Jiawen Yang[2], Ke Xu[3], Bo Dong[1], Shan Jiang[4,5]

[1] Research Center for AI Social Governance, Zhejiang Lab, Hangzhou, Zhejiang 311100, China
[2] Shenzhen Graduate School, Peking University, Shenzhen, Guangdong 518055, China
[3] Zhejiang Development and Planning Institute, Hangzhou, Zhejiang 310012, China
[4] Department of Urban and Environmental Policy and Planning, Tufts University, Medford, MA 02155, USA
[5] Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155, USA

Correspondence to: Xiangfu Kong (kongxf@zhejianglab.com)

Abstract. Hydrological parameters should pass through a careful calibration procedure before aiding decision-making. However, great difficulties are encountered when applying calibration methods to regions where runoff data are inadequate. To fill the gap of hydrological calibration for the ungaged road network, we proposed a Bayesian updating framework to calibrate hydrological parameters based on taxi GPS data. Hydrological parameters are calibrated by adjusting their values such that the runoff generated by the acceptable parameter sets could yield the road disruption period during which no taxi points are observed. The method is validated through 10 flood-prone roads in Shenzhen, and the result reveals that the trends of runoff could be correctly predicted for 8 out of 10 roads. This study shows that integration of hydrological model and taxi GPS data suggests viable alternative measures for the model calibration, and provides actionable insights for flood hazard mitigation.

1 Introduction

Under the background of climate change and increased urbanization, flooding is posing far-reaching threats to the urban road network of coastal metropolis (Balistrocchi et al., 2020). In Australia, around 53% of flood-related drowning deaths were the result of driving into flooded waters. Indirect losses caused by flooding, such as cancelled commutes, mandatory detours, and travel time delays, even outweigh direct losses (Kasmalkar et al., 2020). Quantifying the impact of flooding exposure requires the prediction of surface runoff over the road and computation of road disruption induced by the runoff, which are critical to flooding mitigation, traffic resilience improvement, and risk early warning.

Public concerns about road flooding hazards created the pressing need to develop fine-grained and accurate models for hydrological simulation. The hydrological modelling, as a quite well-established theory, provides an approximation of the real-world hydrological system, and has been widely used in many road-related studies (Versini et al., 2010; Yin et al., 2016; Safaei-Moghadam et al., 2022). As the hydrological modelling is subject to uncertainty, which arises from the false

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

reflection of the hydrological system, the initial and boundary conditions, and the lack of true knowledge, parameters of hydrologic models should be carefully calibrated before applying to solve practical problems, so that the model is capable of closely matching the historical evidences (Gupta et al., 1998). An uncalibrated model is indefensible and sterile, so few models documented in the literature have been applied without any calibrations (Beven, 2012).

During the last four decades, numerous studies have been devoted to the development of calibration methods. Methodologies of model calibration range from the simple trial-and-error, which adjusts one parameter value each turn until the difference between predicted and observed value is satisfactory, to the Bayesian updating framework, which rejected the idea that there is single correct solution. No matter what kinds of methods, hydrological models are basically calibrated based on the runoff data alone (Dembélé et al., 2020), so the success of model calibration is, to a great extent, dominated by the availability of field-observed runoff data. However, runoff data are generally gathered at only few sites, and some cities even never measured runoff data in the built-up regions (Gebremedhin et al., 2020). Even though the runoff data could be effectively collected by administration departments, they had no motivations to share the data to the public. For example, among China's top 10 largest cities[1], only Shenzhen has shared runoff-related data on the open data platform. As for the model calibration for the road scale, the runoff data are even more difficult to acquire, because a road network is far denser than a river network and flood gauges are only located in a few flood-prone roads considering the high measurement cost, leaving most of roads ungaged. As pointed out by Beven (2012, p:55), "the ungauged catchment problem is one of the real challenges for hydrological modellers."

The lack of hydrological data prompted researchers to seek extra data sources to support flood-related decision-making. In response to this need, big data, owing to the advance of mobile telecommunication technologies, are emerging as alternative sources of information for coping with flood risks (Gebremedhin et al., 2020; Paul et al., 2018; Li et al., 2018). Citizens voluntarily or passively acting as human sensors generate georeferenced data to improve flood monitoring. Typical studies involve the use of crowdsourcing social media data (Brouwer and Eilander, 2017; Li et al., 2018), mobile phone data (Yabe et al., 2018; Balistrocchi et al., 2020), and taxi GPS data (She et al., 2019; Kong et al., 2022). However, most of previous works concentrated on using big data either for flood mapping or mining spatiotemporal patterns (Restrepo-Estrada et al., 2018), and it remains an open question of how to calibrate parameters based on big data for ungauged roads.

This study differentiates from our previous study (Kong et al., 2022) by going one step further than simply recognizing the flooding roads. We propose a calibration method for road-related hydrological parameters using the taxi GPS data. Many studies have shown that vehicle-related information during the rainfall, such as vehicle volume, speed, and trajectory, is critical to flooding road detection (Qi et al., 2020; Yao et al., 2020; Zhang et al., 2019). When a road segment is inundated by the heavy rainfall, the vehicle volume may present a sharp or gradual drop, depending on the intensity of the rainfall event. Conversely, the abnormal drop of vehicle volume during the rainfall implies that the road may experience some rainfall-induced inundations. This motivated us to use a traffic-related data source to calibrate hydrological parameters. In

[1] Rank by the resident population in 2021.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

this study, we develop a transformation process which converts the rainfall time series to the time series of the probability that no taxis drive through the road for every hydrological parameter set, and then assign a probability to every parameter set by integrating the no-taxi-passing probability with the observed taxi GPS data. We not only outline a generalized taxi-data-driven calibration framework but also realize the framework with specific hydrologic and transportation models.

2 Methodology

2.1 A Bayesian updating procedure

Observed data are not always as informative as expected and may be inconsistent with other data sources, so hydrologists usually adopt the Bayesian framework to update hydrological parameters, which provides a generalized formalism that integrates prior probability representing the prior knowledge with the likelihood that reflects how well the presumed model can reproduce the observation to form the posterior probability. Suppose we have several hydrological models, each with different sets of parameters, the purpose of the Bayesian updating procedure in this study is to assign a posterior probability to every hydrological parameter set as new taxi data become available.

Two components are critical for the Bayesian updating procedure. One is the prior probability, and the other is the likelihood function. For the prior probability, Beven and Binley (1992) stated in their famous calibration model, generalized likelihood uncertainty estimation (GLUE), that all the parameter combinations are considered equally probable before extra information is introduced. After the first updating, the prior probability of each updating run could be replaced by the posterior probability of the latest updating run. The likelihood, as a measurement of how well the given model conforms to the observed taxi behaviour, is not as easy to compute as the prior probability, because the parameter set to be estimated is hydrology-related while the observed evidence is taxi-based. The question then arises as how to construct a taxi-based proxy whose probability equals to that of the associated hydrological parameter and construct a function enabling the transform from the hydrological parameter to the taxi-related proxy.

The selected proxy in this study is the time series of probability that no taxi drives through the road in a given time interval (short for no-taxi-passing probability). Figure 1 illustrates a generalized procedure of converting the precipitation process to the time series of no-taxi-passing probability for each hydrological parameter. The procedure consists of three steps. First, a hydrological model is used to convert the rainfall process to the hydrograph, which is a graph showing runoff with respect to time past a specific point. Second, a runoff-disruption function, which relates the runoff to the probability that the road is blocked, is used to transition the hydrograph to the time series of road disruption probability. Third, a taxi arriving rate is combined with the time series of road disruption probability to yield the time series of no-taxi-passing probability. Note that the hydrological model and the taxi arriving rate are considered unique for every road and invariable in a short period, while the runoff-disruption function is identical for all roads.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

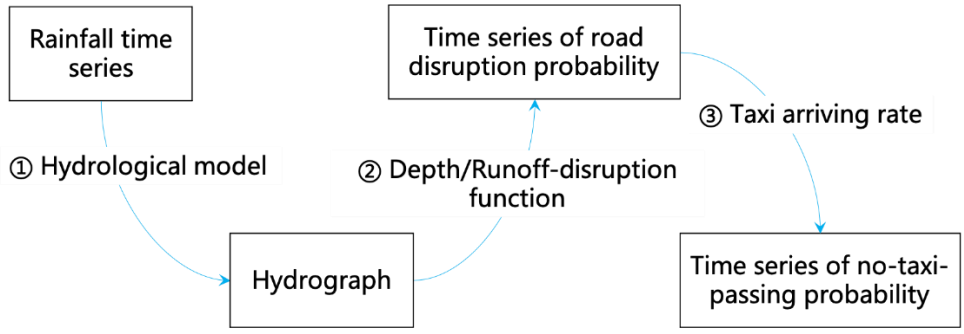


Figure 1 A generalized procedure of converting the rainfall time series to the time series of no-taxi-passing probability

Integrating the three-step process with the Bayesian equation enables us to compute the posterior probability of the parameter set based on the taxi data. For a specific road, suppose there are *N* hydrological parameter sets to be estimated. As the runoff-disruption function and the taxi arriving rate are assumed to be invariant for the road, we can construct a composite function converting the *i*th parameter set, denoted as $\theta^{(i)}$, to the time series of no-taxi-passing probability, denoted as $\Omega^{(i)}$. Therefore, the probability of $\theta^{(i)}$ to be optimal equals to the probability of $\Omega^{(i)}$ to be true:

$$P(\theta^{(i)}) = P(\Omega^{(i)}) \quad (1)$$

where $P(\theta^{(i)})$ and $P(\Omega^{(i)})$ are the prior probability of $\theta^{(i)}$ and $\Omega^{(i)}$ respectively. As taxi observations become available, $P(\theta^{(i)})$ (or $P(\Omega^{(i)})$) can be updated using the Bayes Theorem:

$$P(\theta^{(i)}|X) = P(\Omega^{(i)}|X) \propto P(\Omega^{(i)})\mathscr{L}(X|\Omega^{(i)}) \quad (2)$$

where X is the taxi observation, $P(\theta^{(i)}|X)$ and $P(\Omega^{(i)}|X)$ are posterior probability of $\theta^{(i)}$ and $\Omega^{(i)}$ under the condition of taxi observation, X. The $\mathscr{L}(X|\Omega^{(i)})$ is the likelihood of X given $\Omega^{(i)}$. The optimal parameter set is the one that derives $\Omega^{(i)}$ that best fits the observed taxi data.

The solution of Eq.(2) involves the calculation of $P(\Omega^{(i)})$ and $\mathscr{L}(X|\Omega^{(i)})$. According to Eq.(1), $P(\Omega^{(i)})$ can be replaced with $P(\theta^{(i)})$, which is the prior probability of parameter sets. Beven and Binley (1992) suggested that prior to the introduction of any quantitative and qualitative information, any parameter set combination should be considered equally likely. This implied that the parameter set is drawn from a uniform distribution:

$$P(\Omega^{(i)}) = P(\theta^{(i)}) = 1/N \quad (3)$$

Next, $\mathscr{L}(X|\Omega^{(i)})$, as a likelihood function, describes the joint probability of the observed taxi data, X, as a function of the chosen $\Omega^{(i)}$. Consider a rainfall event that is broken into T 5 min intervals. From the taxi data, we obtained a sequence of taxi-related observations, denoted as $\vec{X} = \{X_1, X_2, ..., X_T\}$, where $X_t = 1$ if the road is observed with at least one taxi passing by in the *t*th 5 min, and $X_t = 0$ otherwise. The $\Omega^{(i)} = \{\omega_1^{(i)}, \omega_2^{(i)}, ..., \omega_T^{(i)}\}$ is also a *T*-dimensional vector, where $\omega_t^{(i)}$ is the no-taxi-passing probability at the *t*th 5 min taking $\theta^{(i)}$ as the parameter set. Note that $\Omega^{(i)}$ is only determined by the chosen

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

hydrological parameter and the rainfall process, and is not measured from the observed data. Considering that the arrival of taxis is independent with respect to time, $\mathscr{L}(X|\Omega^{(i)})$ can be formulated as:

$$\mathscr{L}(X|\Omega^{(i)}) = \prod_{t=1}^{T} (1-\omega_t^{(i)})^{X_t} (\omega_t^{(i)})^{1-X_t} \quad (4)$$

By substituting Eq.(3) and Eq.(4) back into Eq.(2), the Eq.(5) is obtained:

$$P(\theta^{(i)}|X) \propto \frac{1}{N} \prod_{t=1}^{T} (1-\omega_t^{(i)})^{X_t} (\omega_t^{(i)})^{1-X_t} \quad (5)$$

Equation (5) is the proposed Bayesian updating model to calibrate the hydrological parameter based on the taxi data, where X could be directly measured and $\Omega^{(i)}$ is calculated through the three-step process shown in Fig.1, which will be discussed in detail in the next section. Having chosen an updating model, the optimum parameter for one period of observations may not be optimal for another period. As the model may have a continuing input of new taxi observation, the posterior probability for $\theta^{(i)}$ should be updated as new evidence becomes available. For the second update, the posterior from the first observation becomes the prior for the second observation, and the posterior probability for $\theta^{(i)}$ is recursively updated as:

$$P(\theta^{(i)}|\ X_2\ ) \propto \mathscr{L}(X_2|\ \theta^{(i)}) P(\theta^{(i)}|\ X_1\ ) \quad (6)$$

where X_1 and X_2 are the first and the second taxi observations.

2.2 Instantization of the three-step procedure

Section 2.1 presents a generalized three-step procedure which converts the rainfall time series to the time series of no-taxi-passing probability. In this section, we specialize the process by integrating existing theories with our model. Three conceptualized steps shown in Fig.1 are substituted with three more concrete submodels. Firstly, a SCS unit hydrograph is used to turn the rainfall excess to the hydrograph of the road. Secondly, an empirical runoff-disruption function, whose data extracted from various experimental, observational, and modelling studies, is applied to convert the hydrograph to the time series of the road disruption probability. Thirdly, a Poisson distribution, representing the distribution of taxi arriving rate, is combined with the road disruption probability to yield the no-taxi-passing probability.

Step 1: Convert rainfall to runoff based on the SCS unit hydrograph

Not all rainfall will produce runoff because storage from soils can absorb light shower. While in the urbanized area, only a small proportion of rainfall infiltrates into the soil or retained on the land surface, most of them flow across the urban surface and becomes the direct runoff. The rainfall that yields the direct runoff is termed rainfall excess. The Natural

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

Resources Conservation Service (NRCS)[2] developed a method to estimate the rainfall excess based on the soil types and land use using the curve number equation:

$$P_e = \begin{cases} (P_a - 0.2S)/(P_a + 0.8S) & P_a > 0.2S \\ 0 & P_a \leq 0.2S \end{cases} \quad (7)$$

where P_e is the accumulated rainfall excess in cm, P_a is the accumulated rainfall in cm, and S is the potential retention after runoff begins, which is a function of the curve number:

$$S = 2.54 \times (1000/CN - 10) \quad (8)$$

where CN is the curve number. For urban and residential land, the curve number varies from 65 to 85, depending on the impervious areas. For sake of brevity, the curve number will not be regarded as a parameter to be calibrated in this study but as a given parameter with the value of 85.

Next, the rainfall excess derived by Eq.(7) is input to the unit hydrograph to produce the runoff. The unit hydrograph is a commonly used rainfall-runoff model that converts rainfall excess to direct runoff. First proposed by Sherman in 1932, the unit hydrograph is defined as the hydrograph resulting from one unit of rainfall excess distributed uniformly over a catchment. It assumes the rainfall is uniform over the catchment and that runoff increases linearly with the rainfall excess. Although under most conditions these assumptions cannot be perfectly satisfied, the results obtained from the unit hydrograph are generally acceptable for most practical uses. The model was originally designed for larger watersheds, but it has been found applicable to some catchments less than 5,000 m^2 (Chow et al., 1988).

The unit hydrograph applies only for the watershed where the runoff data were measured. The paucity of the runoff data sparkled the idea of the synthetic unit hydrograph (SUH) concept. The term "synthetic" in SUH denotes the unit hydrograph derived from watershed characteristics rather than empirical rainfall-runoff relationship. In this study, we utilized the SCS unit hydrograph, which is a dimensionless SUH proposed by the NRCS. For the dimensionless SUH, the discharge (i.e. *y*-axis) is expressed as the ratio of discharge q to the peak discharge q_p and the time (i.e. *x*-axis) as the ratio of time t to the peak time t_p. Therefore, the SCS unit hydrograph, rigorously speaking, is not a SUH itself, but a useful tool for constructing a SUH.

The shape of SCS unit hydrograph is totally determined by the peak rate factor. A standard value of 2.08 for the peak rate factor is recommended and commonly used by the NRCS (Fig.2). To construct a SUH from the SCS unit hydrograph, the *x*-axis of the SCS unit hydrograph is multiplied by t_p and the *y*-axis by q_p. The values of q_p and t_p are functions of the catchment area and the time of concentration:

$$t_p = 0.6t_C + D/2 \quad (9)$$

$$q_p = 2.08A/t_p \quad (10)$$

[2] The NRCS used to be called the US Soil Conservation Service (SCS).

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

where t_C is the time of concentration in hour, A is the catchment area in km^2, D is the duration of unit rainfall excess in hour, which is 1/12 h (i.e. 5 min) in this study. As can be seen, the catchment area and time of concentration are required to construct a SUH, and they are the two hydrological parameters we would calibrate based on the taxi data. For sake of simplicity, the peak rate factor would not be calibrated and be fixed as 2.08, although some studies have showed that it has a much wider range from 0.43 for steep terrain to 2.58 for very flat terrain (Chow et al., 1988). After t_C and A are chosen, a SUH could be constructed, and then we use it to convert the rainfall excess to the runoff by applying the discrete convolution equation. The detailed computation process of the discrete convolution equation can be found in most hydrological textbooks (e.g., see Chow et al., (1988) pp: 211-213), and will not be discussed here. To be clear, a graphic workflow in Fig.3 shows how the rainfall time series is transformed to the hydrograph for every parameter set.

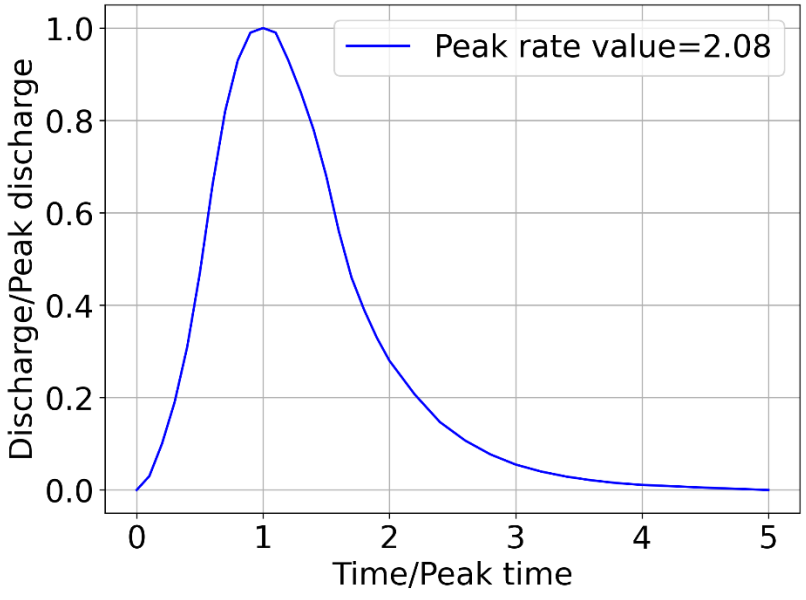


Figure 2 The standard SCS unit hydrograph. Data provided by the NRCS (2007)

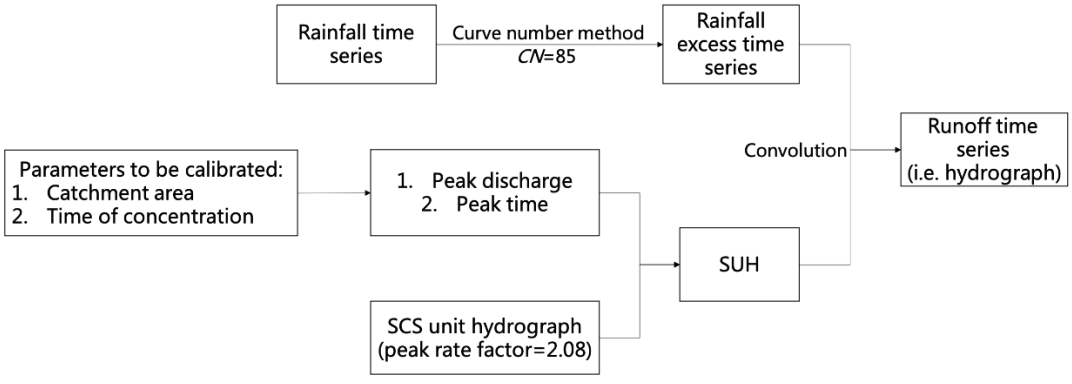


Figure 3 Workflow of the SCS unit hydrograph to convert rainfall to runoff.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

Step 2: Derive the road disruption probability using the runoff-disruption function

The goal of Step 2 is to convert the hydrograph generated in Step 1 to the time series of road disruption probability, or more specifically, the probability that a taxi driver chooses to turn the car when arriving at a flooded road. Most models in the literature assume a road is either open or closed, which usually does not correspond to the empirical evidence that many drivers take risk to drive along the rood even though it is inundated. In order to transition from a binary view of a flooded road being considered "open" or "closed", Pregnolato et al. (2017) proposed to use a curve that relates the depth of floodwater to a reduction in vehicle speed to show the probability of road disruption, and such idea is soon followed by Contreras-Jara et al. (2018) and Nieto et al. (2021).

A driver will turn around when he believes that the flow rate overcomes the vehicle configuration. From this perspective, the road disruption probability equals to the probability that the vehicle performance is lower than the flow rate perceived by a driver. However, it is a difficult task to quantify the common belief of what guide people's willingness to drive through a flooded waterway, and is also difficult to obtain the precise knowledge of all taxi-flood intersections. Alternatively, to ensure the vehicle stability in flood flows, guidelines are usually recommended based on the limiting values of depth times velocity, and many studies have carried out laboratory testing on the stability of different kinds of vehicle models exposed to different combinations of depth and velocity (Merz and Thieken, 2009; Shah et al., 2018). Suggested by Pregnolato et al. (2017), we constructed the runoff-disruption function by integrating data from reviewed literatures and some authoritative guidelines. In this study, the road disruption probability is defined as the probability that the product of flow velocity and flow depth is higher than the stability limits extracting from existing studies, which are shown in Table 1 and plotted in Fig.4. The expression of the fitting curve is:

$$y = \left[1 + \exp\left(-16.6\left(x - 0.48\right)^2\right)\right]^{-1} \quad (11)$$

According to Eq.(11), a road has a disruption probability of 50% when the product of flow velocity and flow depth is 0.47 $m^2\ s^{-1}$, and is totally disrupted when the product is higher than 0.80 $m^2\ s^{-1}$. Applying the fitting curve, we can easily convert the flood runoff to the disruption probability:

$$P(Disrupt)_t^{(i)} = \left[1 + \exp\left(-16.6\left(q_t^{(i)}/W - 0.48\right)^2\right)\right]^{-1} \quad (12)$$

where $P(Disrupt)_t^{(i)}$ and $q_t^{(i)}$ are the road disruption probability and discharge at the tth 5 min derived by the hydrological model with parameter set $\theta^{(i)}$, and W is the road width.

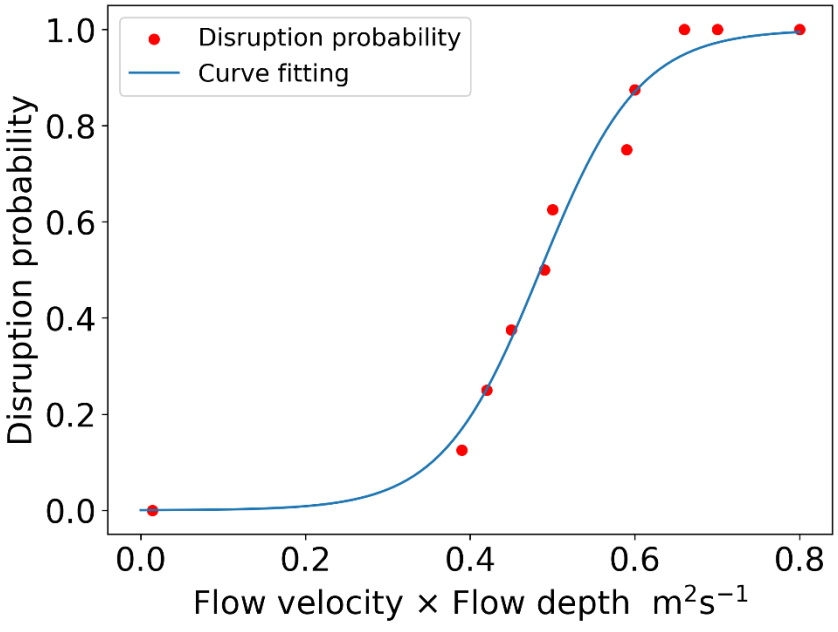
Table 1 Guidelines recommended by existing literatures.

Reference	Vehicle type	Feature	Recommended limits for vehicle stability ($m^2\ s^{-1}$)
Shah et al. (2018)	Volkswagen Scirocco	Flow direction =0°	velocity×depth<0.014
Al-Qadami et al. (2022)	Perodua Viva	Ground clearance =0.18 m	velocity×depth<0.39
Calculated according to Kramer et al. (2016)	VW Golf III	/	velocity×depth<0.42

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

Shand et al. (2016)	Large passenger	Ground clearance >0.12 m	velocity×depth<0.45
Martínez-Gomariz et al. (2017)	Mini Cooper	Ground clearance =0.12 m	velocity×depth<0.49
Martínez-Gomariz et al. (2017)	BMW i3	Ground clearance =0.10 m	velocity×depth<0.49
Martínez-Gomariz et al. (2017)	BMW 650	Ground clearance =0.08 m	velocity×depth <0.50
Martínez-Gomariz et al. (2017)	Mercedes GLA	Ground clearance =0.17 m	velocity×depth <0.59
Moore and Power (2002)	All but very small cars	/	velocity×depth <0.60
Calculated according to Xia et al. (2014)	Honda Accord	/	velocity×depth <0.65



215 **Figure 4** Empirical runoff-disruption function derived from existing literatures

Step 3: Derive the time series of no-taxi-passing probability

A road has no taxis passing by in a fixed time step if the road has no taxis visiting or every taxi that arrives at the road turns around, so the no-taxi-passing probability can be inferred by the following equation:

$$\omega_t^{(i)} = \sum_{n=0}^{\infty} P(Arrival\_taxi{=}n) \times \left(P(Disrupt)_t^{(i)}\right)^n \quad (13)$$

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

where $\omega_t^{(i)}$ is the no-taxi-passing probability in the tth 5 min, and $P(Arrival\_taxi=n)$ is the probability that n taxis arrives at the road segment in the tth 5 min. Equation (13) indicates that if every taxi arrives at the road segment makes a turn because of the flooded waterway, taxi volume of the road will be zero. In this study, $P(Arrival\_taxi=n)$ is assumed to follow the Poisson distribution:

$$P(Arrival\_taxi=n) = e^{-\lambda}\lambda^n/n! \tag{14}$$

where λ is the average number of taxis arriving at the road. Substitute Eq.(14) into Eq.(13), we derive:

$$\omega_t^{(i)} = \sum_{n=0}^{\infty}(e^{-\lambda}\lambda^n/n!)\times\left(P(Disrupt)_t^{(i)}\right)^n \tag{15}$$

Applying $e^x = \sum_{n=0}^{\infty} x^n/n!$, Eq.(15) can be further converted to:

$$\omega_t^{(i)} = e^{-\lambda}\sum_{n=0}^{\infty}\left(P(Disrupt)_t^{(i)}\lambda\right)^n/n! = \exp\left(\lambda\left(P(Disrupt)_t^{(i)}-1\right)\right) \tag{16}$$

Equation (16) indicates that $\omega_t^{(i)}$ is totally determined by λ and $P(Disrupt)_t^{(i)}$. Since $P(Disrupt)_t^{(i)}$ is given through Step 2, what is left to determine is the value of λ. As the rain gets heavier, experienced taxi drivers will avoid flood-prone roads in advance, which means that λ, strictly speaking, is a decreasing function of rainfall intensity. However, fitting the rainfall-λ curve requires substantial taxi GPS trajectories to inspect the route choices of taxi drivers under heavy rain, which is currently unfeasible in this study. We assume that λ is a constant quantity which keeps unchanged with respect to rainfall. The value of λ can be calculated by averaging all 5 min taxi volume using the historical taxi GPS data.

Finally, Table 2 lists all the submodels and parameters of the three-step process. The core principle of the three-step process is to calculate the time series of no-taxi-passing probability, $\Omega^{(i)}$, for each parameter set, $\theta^{(i)}$. As the best choice of a model is often data-specific, it is probable that the model combination proposed in this study is not optimal for other studies. To apply the calibration method in practical use, one must specify the submodel in the three-step process according to the available data, prior knowledge, and accuracy requirement.

Table 2 Specific submodels and parameters of the three-step process.

Purpose of the step	Specific model	Parameter	Source of parameters
Step 1: Convert the rainfall to the hydrograph	Curve number equation	1. Curve number	Existing literature
	SCS unit hydrograph	2. peak rate factor	Existing literature
		3. Catchment area 4. Time of concentration	Parameters to be calibrated
Step 2: Convert the hydrograph to the time series of disruption probability	Empirical runoff-disruption function	5. Limit of product of flow velocity and depth	Existing literatures
Step 3: Convert the time series of disruption probability to the time series of no-taxi probability	Taxi arrival rate follows the Poisson distribution	6. Average taxi volume in 5 min	Taxi GPS data

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

3 A working example

The method outlined above was tested on an intersection located in Xinzhou Road and Hongli Road in Shenzhen, which is recognized as a waterlogging point by the Water Authority of Shenzhen Municipality. Recall that parameters to be calibrated are catchment area, A, and time of concentration, t_c. The range of parameters should be wide enough to encompass most possible values. After several rounds of testing, the maximum value for A is set as 0.5 km^2, and the maximum value for t_c is 5 h. Optimal parameter sets for most roads would fall into the region enclosed by the maximum parameter sets. Table 3 shows the detail information of parameter sets to be calibrated, which totally form 3,000 possible combinations.

Table 3 Detail information of parameter sets to be calibrated.

Parameter	Annotation	Minimum	Maximum	Incremental	Number of possible parameter values
Catchment area	A	0.01 km^2	0.5 km^2	0.01 km^2	50
Time of concentration	t_c	1/12 h	5 h	1/12 h	60

The taxi GPS data collected during two storm events occurring on 9 May 2015 and 23 May 2015 were used to calibrate the parameter sets of the intersection. Rainfall time series and taxi observations under two storms are shown in Fig.5. Each taxi observation contains two time series. One is the time series of 5 min taxi volume, and the other is the 5 min road status, which is derived from the taxi volume, with the value to be 1 if the taxi volume is higher than 0 and 0 if the taxi volume is 0.

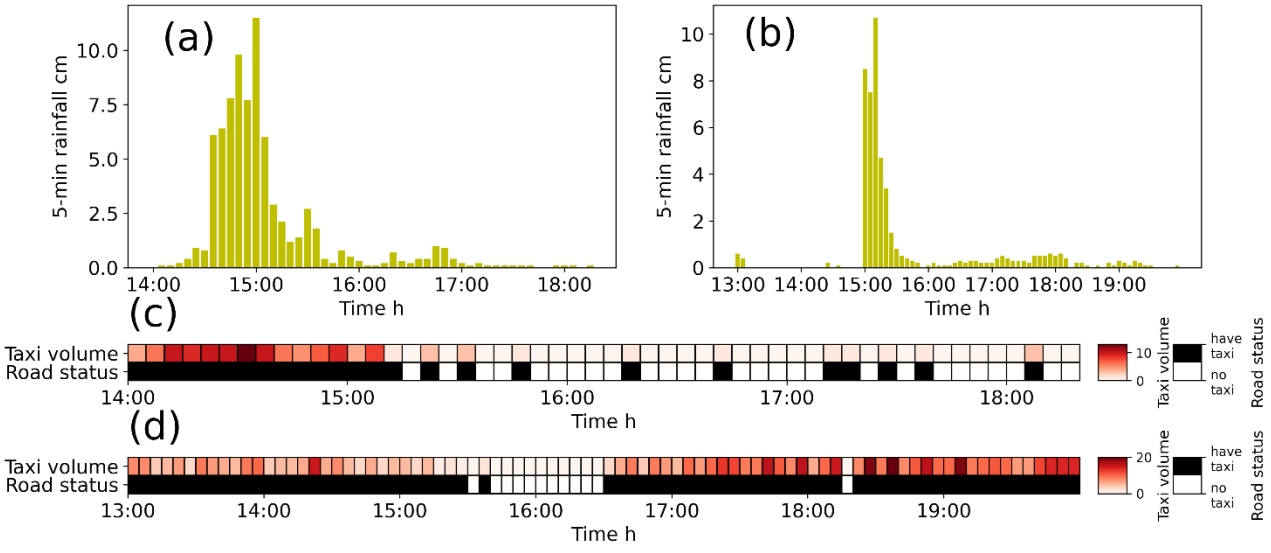


Figure 5 Rainfall and taxi observations used to calibrate the hydrological parameters. **(a)** 5 min rainfall time series on 9 May 2015. **(b)** 5 min rainfall time series on 23 May 2015. **(c)** Taxi observations on 9 May 2015. **(d)** Taxi observations on 23 May 2015.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

Given the rainfall on 9 May 2015, we should calculate the time series of no-taxi-passing possibility for each parameter combination. According to the three-step process, the first step is to convert the original rainfall to the rainfall excess using the curve number method (Fig.6a). Then, for each combination of A and t_c, we construct a SUH. As there are 3,000 parameter sets, we can construct 3,000 different SUHs. For simplicity, we only chose the 1,170th parameter set, i.e. A=0.2 km^2 and t_c=2.75 h, as examples to show the calibration works. Using Eq.(9) and Eq.(10), the peak discharge q_p and peak time t_p can be calculated as:

$$t_p = 0.6 \times 2.75 + 1/(2 \times 12) \approx 1.69 \text{ h}$$
$$q_p = 2.08 \times 0.2/1.69 \approx 0.24 \text{ m}^3/\text{s}$$

The SUH is derived by multiplied by t_p on the *x*-axis and by q_p on the *y*-axis of the standard SCS unit hydrograph (Fig.6b). Next, the rainfall excess shown in Fig.6a is combined with the derived SUH to yield the hydrograph through the convolution (Fig.6c).

In the second step, the runoff is transformed to the time series of road disruption probability based on the runoff-disruption function (Fig.6d). Note that the runoff-disruption function takes the production of water depth and velocity (in the unit of m^2 s^{-1}) as input. Therefore, the original runoff (in the unit of m^3 s^{-1}) produced by the first step should be divided by the road width before inputting to the runoff-disruption function.

In the third step, the time series of road disruption probability (Fig.6e) is converted to that of no-taxi-passing probability using Eq.(16) (Fig.6f). According to the historical taxi GPS data, the average number of taxis arriving at the road, λ, is 10.0 taxi per 5 min. The derived time series of no-taxi-passing possibility is shown in Fig.6g.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

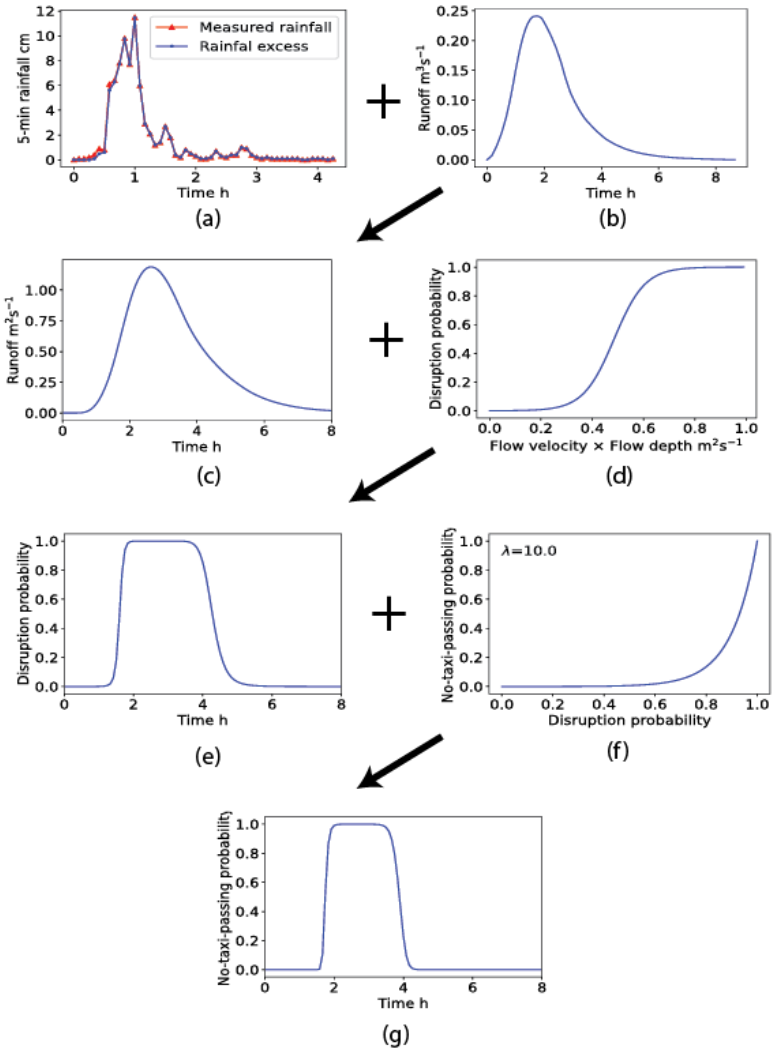


Figure 6 An example of how the rainfall time series is transformed to the no-taxi-passing probability using the three-step procedure for the 1170th parameter set. **(a)** Time series of rainfall and rainfall excess. **(b)** SUH constructed using the 1170th parameter set. **(c)** Derived runoff. **(d)** Empirical runoff-disruption function. **(e)** Derived time series of disruption probability. **(f)** Disruption-no-taxi-passing probability function. **(g)** Derived no-taxi-passing probability.

After the time series of no-taxi-passing probability for every parameter set is derived, we can calculate the degree of belief that a given parameter set is optimal by integrating it with the taxi observations on 9 May 2015. According to Eq.(5), the posterior probability of the 1,170th parameter set is calculated as:

$$P(\theta^{(1170)}|X) \propto P(\Omega^{(1170)})\,\mathcal{L}(X \mid \Omega^{(1170)}) = \frac{1}{3000}\prod_{t=1}^{T}(1-\omega_t^{(1170)})^{X_t}(\omega_t^{(1170)})^{1-X_t}$$

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

where $\mathcal{L}(\theta^{(1170)}|X)$ is the likelihood that the 1,170th parameter set is optimal conditioning on X, which is the taxi observation on 9 May 2015 shown in Fig.5c. The $P(\Omega^{(1170)})$ is the prior probability of the 1,170th parameter to be optimal, and its values is 1/3000 because there are 3,000 possible combinations.

Following the above process, we can calculate the posterior probability for every parameter set. Furthermore, the posterior probability distribution of parameter set could be updated using the taxi observation and rainfall data on 23 May 2015:

$$P(\theta^{(i)}|\ X_2\ ) \propto\ \mathcal{L}(X_2|\ \theta^{(i)})P(\theta^{(i)}|\ X_1\ )$$

where $P(\theta^{(i)}|\ X_1\ )$ is the original posterior probability distribution calculated based on the storm on 9 May 2015, and $P(\theta^{(i)}|\ X_2\ )$ is the updated posterior distribution after the data of storm on 23 May 2015 are added. Fig.7 illustrated the evolution of the probability distribution with the availability of more taxi data. The posterior distribution dominates the uniform prior distribution after the first updating, and the distribution is refined a little bit after the second updating.

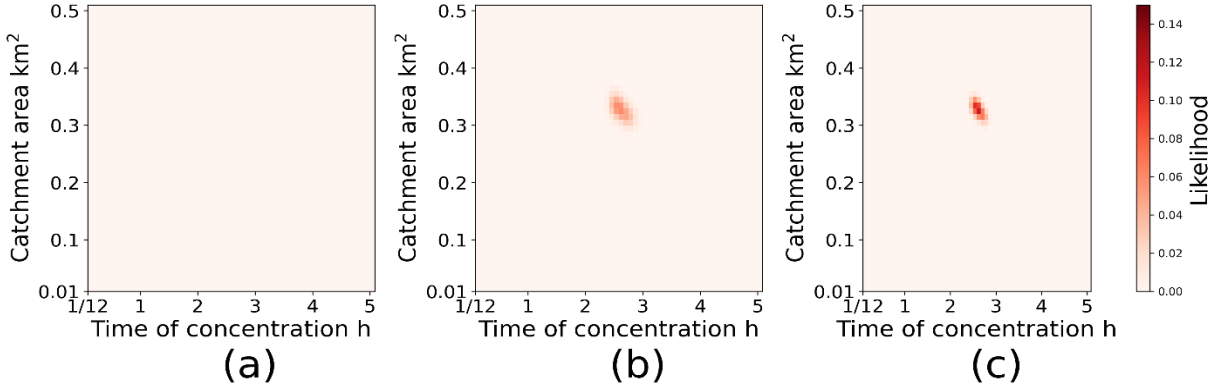


Figure 7 Evolution of the posterior probability distribution of hydrological parameter sets. **(a)** Prior distribution before updating. **(b)** Posterior distribution after the first updating. **(c)** Posterior distribution after the second updating.

4 Validation and result

4.1 Method validation

The proposed method is validated upon flood-prone roads located in Shenzhen, China, which is a coastal city frequently hit by extreme storms in summer. Another reason that Shenzhen is chosen is that only Shenzhen, as far as we known, has shared the runoff-related data to the public in China. Three data sources, which are taxi GPS data, rainfall data, and authoritative water level data, are used to validate the calibration method. Hydrological parameters are calibrated using the first two data sources, and the water level data acts as the ground truth to validate the method. Taxi GPS data are anonymized and aggregated to the road every 5 min. The rainfall data, which are also collected every 5 min, are measured at 115 gaging stations citywide, and are mapped to the road network throughout Shenzhen using the Ordinary Kriging spatial

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

interpolation algorithm. The water level data are only measured at some waterlogging points, with a dynamic sampling interval ranging from 5 min when rainy to 1 h when rainless. The calibration method was validated by checking the hydrograph derived from the calibrated hydrological model and authoritative water level for 10 selected roads. Detailed information of three data sources are listed in Table 4.

Table 4 Detailed information of three data sources.

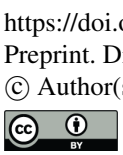
Item	Taxi GPS data[1]	Rainfall data[1]	Water level data[2]
Source	Transport Commission of Shenzhen Municipality	Meteorological Bureau of Shenzhen Municipality	Shenzhen Municipal Government Data Open Platform[1]
Record	Taxi volume of each road	5 min accumulative rainfall	Water level
Data collection period	May 2015	2015 and 2019	2019
Data collection interval	5 min	5 min	5 min -1 h
Location	Citywide	115 rainfall gaging stations	171 flooding gaging sites

[1] The complete taxi GPS data and rainfall data are not openly accessible due to the requirement of data policy. To validate the research findings, we uploaded necessary data in Zenodo (Kong, 2022).

[2] Openly available at the site: https://opendata.sz.gov.cn/data/dataSet/toDataDetails/29200_01403147

Two storm events, occurred on 9 May 2015 and 23 May 2015 are treated as calibration events and the storm occurred on 11 June 2019, is retained for testing. Obviously, there is a 4 year span between the calibration data and validation data due to the data availability. To reduce the validation error caused by the time difference, roads to be validated should be vulnerable to flooding on both 2015 and 2019 so that hydrological parameters of these roads have higher chance to remain unchanged. Therefore, in total of 10 flood-prone roads, which were labelled as flood-prone roads on both the List of 2015 Flood-prone Roads in Shenzhen (Water Authority of Shenzhen Municipality, 2015) and the List of 2019 Flood-prone Roads in Shenzhen (Water Authority of Shenzhen Municipality, 2019), were carefully selected (Fig.8).

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023



Hydrology and Earth System Sciences
Discussions

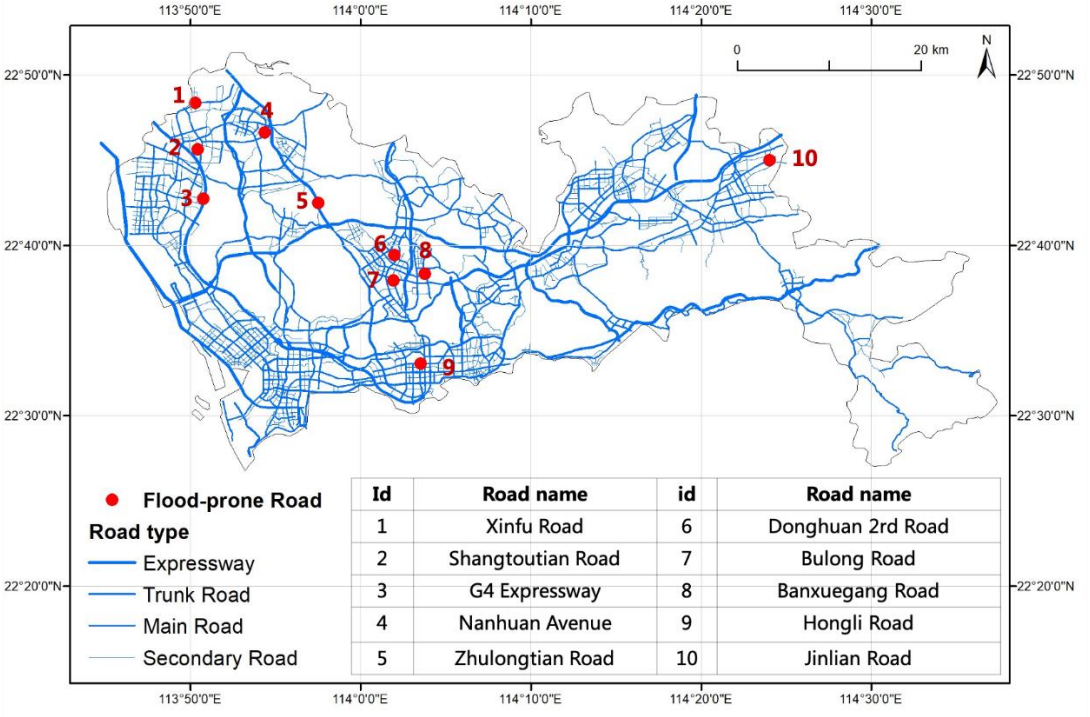
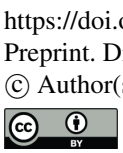


Figure 8 Spatial distribution of 10 flood-prone roads in Shenzhen.

Next, the posterior probability of parameter sets after calibration for the 10 roads are illustrated in Fig.9. As shown in Fig.9, the posterior probability distribution of parameter sets for most flood-prone roads are clustered around the optimal parameter set after two runs of updating, indicating that the uncertainty of parameter sets is refined to a much smaller area when taxi observations are added. It should also be noted that the posterior probability of parameter set for the Jinlian Road (Road ID=10) is evenly distributed on a triangular region (Fig.9j). By examining the taxi data of the road, we found that the taxi volume was greater than 0 for most 5 min intervals during two storms, indicating that the road was not disrupted during two storms. As hydrological parameters are calibrated by adjusting their values such that the runoff generated by the acceptable parameter sets could yield the disruption period during which no taxi points are observed, the lack of no-taxi-passing period would provide less information for calibration compared with when no-taxi-passing period is observed. This explains why the posterior probability is not refined to a small-area domain. However, we can still get some valuable information from Fig.9j. First, the catchment area for the Jinlian Road should not be too large to generate the runoff which may cause the road disruption. Second, the catchment area is highly intercorrelated with time of concentration. As the catchment area gets larger, the time of concentration is more likely to increase so that the high runoff volume could not converge in a short time.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023



Hydrology and Earth System Sciences
Discussions

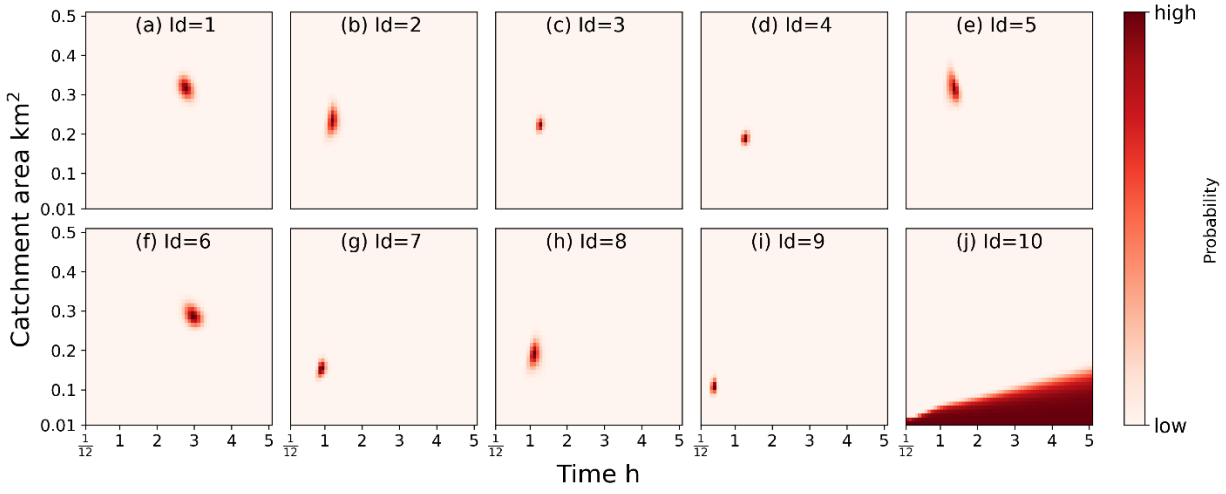


Figure 9 Posterior probability distribution of hydrological parameter sets after the first updating for 10 flood-prone roads. Subplots (a)-(j) represent the probability distribution for Road 1-10.

After the parameter sets were calibrated, they were combined with the SCS unit hydrograph to construct the SUH, which were further combined with the rainfall data occurring on 11 June 2019 to produce the predicted hydrograph. As the posterior probability associated with each parameter set can be regarded as a fuzzy measure reflecting the degree of belief that the parameter set is true, the weighted runoffs for each parameter set were summed to produce the final predicted runoff:

$$Q = \sum_{i=1}^{N} P(\theta^{(i)}|X) Q^{(i)} \tag{17}$$

where Q is the final predicted runoff, $Q^{(i)}$ is the simulated runoff derived from the ith parameter set, and $P(\theta^{(i)}|X)$ is the posterior probability of the ith parameter set, acting as the weight.

The output of the calibrated hydrological model is runoff (with the unit of $m^3\ s^{-1}$), whereas the validation data is water level (with the unit of m). As the calibration data and validation data arise from multiple sources and have different units, conventional error-based statistics such as the mean absolute error (MAE) are not suitable in this study. Most often, the discharge of stream is rarely measured directly. Instead, streamflow is typically determined by converting measured water depth (i.e. stage) into discharge through a rating curve, which provides a functional relationship between stage and discharge at a specified point (Le Coz et al., 2014). Inspired by the application of the rating curve, we validate the method by developing the rating curve for every road, and then estimate the goodness-of-fit of those rating curves. Comparisons between the observed water depth and the simulated runoff for 10 selected roads are shown in Fig.10, and rating curves constructed by fitting the runoff-stage scatter plot are shown in Fig.11. We use the Pearson correlation coefficient, which measures the linear correlation between two variables, as the goodness-of-fit indicator. The result shows that 8 of 10 roads have rating curves with significant positive Pearson coefficient, indicating that the runoff and water have similar and consistent variation process.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

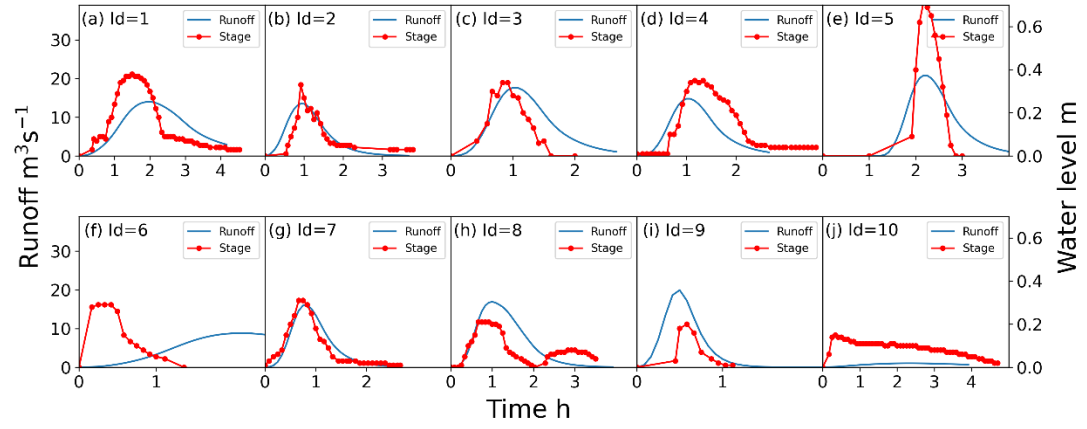


Figure 10 Comparisons between the observed water depth and the simulated runoff for Road 1-10. The maximum value is 30 m^3 s^{-1} of the left *y*-axis (i.e. runoff) and 0.6 m of the right *y*-axis (i.e. stage) for every subplot.

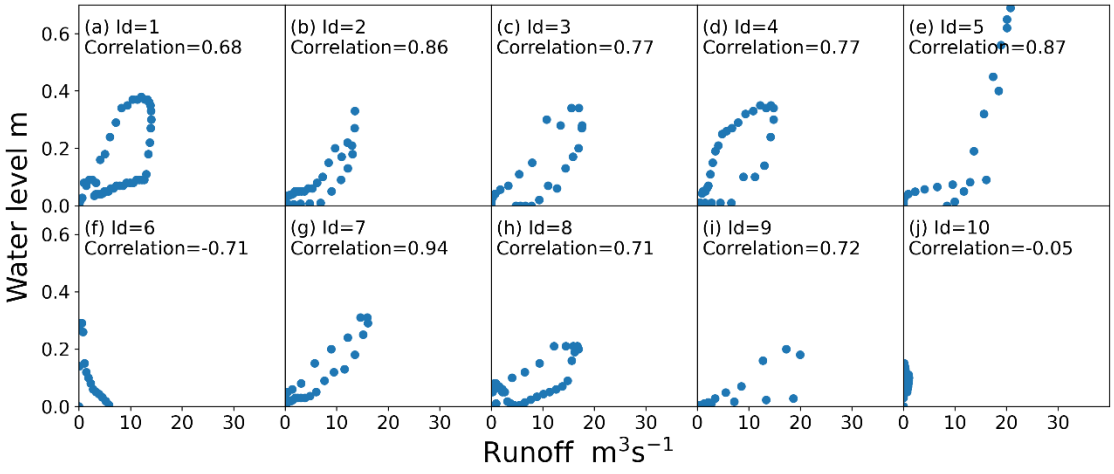


Figure 11 Scatter plots (a-j) of the observed water depth and the simulated runoff for Road 1-10.

It is worth noting that goodness-of-fit solely describe the degree of correlation between the observed and simulated data and may contain validation bias. As suggested by Legates and McCabe (1999), correlation-based statistic is insensitive to additive and proportional differences between the simulations and observations. Therefore, fitting of rating curve only reveals part of the validation truth, and the usefulness of the calibration method needs further inspection.

4.2 Application of the method to plot flooding maps in Shenzhen

Based on the proposed calibration method, we simulated how the road network experiences flooding for different rainfall return periods. Three storm events of different return periods (T = 2, 10, and 50 years) were designed according to the Rainfall Intensity Formula of Shenzhen (Meteorological Bureau of Shenzhen Municipality, 2015). Each storm lasts 3 hours, with an accumulative rainfall amount of 159 mm, 230 mm, and 283 mm for the 2, 10, and 50 year return period.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

375 Hydrological parameters of high-level flood-prone roads, including expressway, main road, and secondary road, are calibrated using the taxi data on 9 May 2015 and 23 May 2015. The flood-prone roads are identified based on the algorithm proposed in our previous studies (Kong et al., 2022). The road discharges under different rainfall return periods are simulated by inputting the designed rainfall to the calibrated hydrologic models. As an example, Fig.12 shows the spatio-temporal evolution of simulated discharge of parts of the road network, which locate in Baoan District, Shenzhen, for
380 different return periods. With the return periods rising from 2 year to 50 year, the average peak discharge for flood-prone roads increases by 80.6%, with the value from 13.9 m^3 s^{-1} to 25.1 m^3 s^{-1}. Inputting the simulated runoff to the empirical runoff-disruption function, expressed in Eq.(12), the time series of disruption probability for every road could be derived[3]. To facilitate discussion, we temporarily define the disruption period as the time when disruption probability is higher than 0.5. The average disruption period for flood-prone roads increases from 1.67 h to 3.15 h as the return period increases from 2
385 year to 50 year.

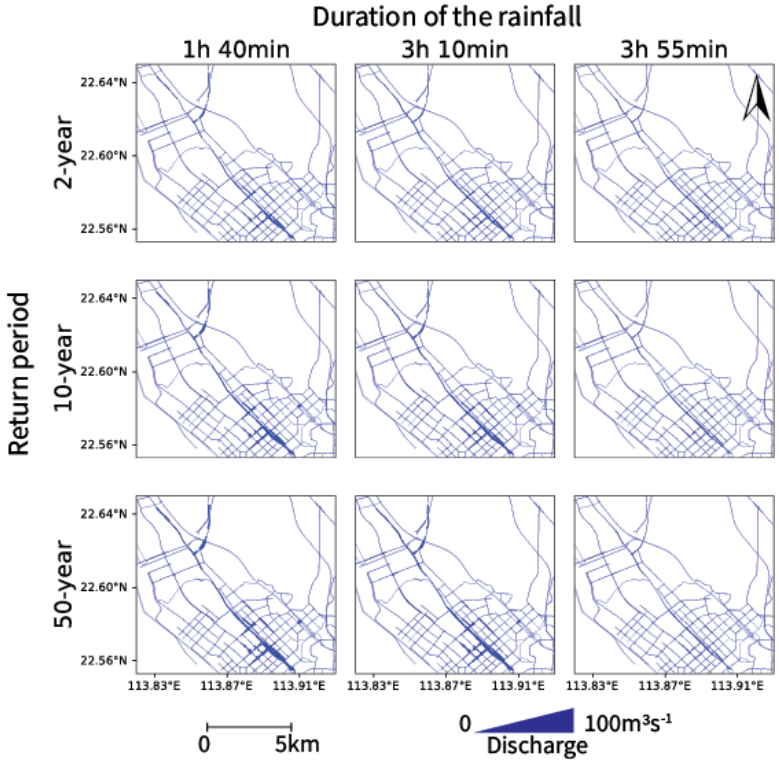


Figure 12 Spatio-temporal evolution of simulated runoff for different return periods in Baoan District, Shenzhen.

[3] Original runoff should be divided by the road width before inputting to the empirical runoff-disruption function.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023

Hydrology and Earth System Sciences
Discussions

5 Discussion

Three points are worth discussing about the proposed calibration method. The first is that, although the validation results support the use of taxi GPS data to calibrate hydrological parameters for poorly gauged road networks, the method is more applicable to the road which is frequently visited by taxis. Uncertainty increases as the taxi volume of a road decreases. A road is passable when at least one taxi GPS points are observed during the time interval, while we cannot assert that the road is disrupted when the taxi volume is zero. When a road with taxis frequently passing by is observed with no taxi GPS points during the storm, it is highly probable that the road is disrupted by the flooding, which provides relatively reliable information for parameter calibration. Conversely, when a road with few taxis visiting has no taxi points during the storm, there is a great chance that the road remains passable and just has no taxis as usual. The calibration method thus becomes relatively unreliable considering that the "no-taxi-passing period" is no longer a good proxy of the "disruption period" for the taxi-data-sparse road. To compensate for the shortage of the taxi GPS data, extra data sources, such as ride-hailing data and bus data, should be incorporated in the future work.

Secondly, the disruption of one road may cause cascading failure so that the disruption may be rapidly propagating from the inundated road to the adjacent non-inundated roads under the constraint of the road connectivity. For a road which is disrupted but not inundated by the storm, the implementation of calibration method may be subject to structural errors. Assume there are two connected roads, namely Road 1 and Road 2, which are both disrupted during a storm, and taxi volume of two roads are therefore zero (Fig.13). The difference lies in that Road 1 is disrupted by the flooding, while Road 2 is disrupted due to connecting to the disrupted road, i.e. Road 1. If taxi data is the only data source used for calibration, posterior distributions of hydrological parameters for Road 1 and Road 2 should be identical after calibration, because sequences of taxi volume are identical for both roads. Clearly, we know that hydrological parameters for two roads could not be the same, otherwise Road 1 and Road 2 should be both flooded. Just like we cannot simply treat the "no-taxi-passing period" as the "disruption period", we cannot confuse the "disruption period" with the "flooded period." In the future work, an algorithm enabling to distinguish the flooding-induced disruption and the connectivity-induced disruption should be developed.

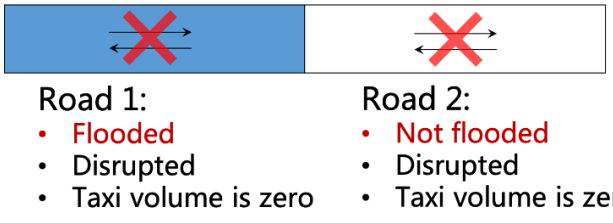


Figure 13 A graphic representation to show the difference between the "disruption period" and the "no-taxi-passing period."

Thirdly, the specific three-step process, which consists of the SCS unit hydrograph, the empirical runoff-disruption function, and the Poisson distribution, performs as a realization of the generalized framework shown in Fig.1. Submodels of the three-step process are not deterministic, and can be flexibly substituted by other submodels according to the needed complexity and data availability. For example, an alternative to the SCS unit hydrograph is the distributed hydrological

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

model. Compared with the SCS unit hydrograph, the distributed hydrological model partitions a watershed into physically homogeneous units and captures the complex spatial variation induced by human activity in high resolution, which may be more applicable to the urbanized environment, such as the road network. However, considering that some critical data including the road drainage data and land use data are missing, as well as the calibration procedure will become extremely computationally intensive, we did not use the distributed hydrological model in this study.

6 Conclusion

An urban flooding model requires various types of data for calibration. In this study, we proposed a Bayesian calibration framework for the hydrological parameters of the road network based on the taxi GPS data. A three-step procedure, consisting of a rainfall-runoff model, a runoff-disruption model, and a disruption-no-taxi-passing probability model, enables us to transform the given rainfall time series to the time series of no-taxi-passing probability for each parameter set, which is key to the taxi-data-driven model calibration. The calculated no-taxi-passing probability, acting as a proxy of the associated hydrological parameter set, is further compared with the observed taxi data through the Bayes equation to assess the posterior probability of the hydrological parameter set. The calibration method is instantiated by combing some classical hydrological and traffic flow models, and is validated on 10 flood-prone roads in Shenzhen. The validation results show that trends of runoff could be correctly predicted for 8 roads, indicating a good performance for hydrological parameter calibration.

This study illustrates the great potential of integrating transportation-related data with hydrological theory in transportation resilience improvement and flood risk management for the road network. We hope that our study provides a flexible calibration framework for countries which are short of runoff data but rich of taxi data. We accept that the application of the method is currently limited by the heterogeneous spatial distribution of taxis citywide and the cascading effect of road inundation, but expect this to change with the increasing availability of vehicle data and continuously optimization of modelling.

Code and data availability

The data and code used to validate the method are available at Zenodo (https://doi.org/10.5281/zenodo.7294849).

Author contribution

JY conceptualised the article and collected the field data. XK designed the methodology and was responsible for the code compilation. KX plotted the figures and revised the manuscript. BD manged the implementation of research activities.

445 SJ discussed the results and contributed to the method validation. XK wrote the final version of the article with contributions from all co-authors.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgments

450 This research was supported by the National Key Research and Development Program of China (grant number 2022YFC3303100).

Reference

Al-Qadami, E. H. H., Mustaffa, Z., Al-Atroush, M. E., Martinez-Gomariz, E., Teo, F. Y., and El-Husseini, Y.: A numerical approach to understand the responses of passenger vehicles moving through floodwaters, J. Flood. Risk. Manag., 455 https://doi.org/10.1111/jfr3.12828, 2022.

Balistrocchi, M., Metulini, R., Carpita, M., and Ranzi, R.: Dynamic maps of human exposure to floods based on mobile phone data, Nat. Hazard. Earth. Syst., 20, 3485–3500, https://doi.org/10.5194/nhess-20-3485-2020, 2020.

Beven, K.: Rainfall-Runoff Modelling: The Primer, 2nd Edition, 2nd ed., Wiley-Blackwell, 2012.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol. 460 Process., 6, 279–298, https://doi.org/10.1002/hyp.3360060305, 1992.

Brouwer, T. and Eilander, D.: Probabilistic flood extent estimates from social media flood observations, Nat. Hazard. Earth. Syst., 17, 735–747, 2017.

Chow, V. T., Maidment, D. R., and Mays, L. W.: Applied Hydrology, McGraw-Hill Book Company, 1988.

Contreras-Jara, M., Echaveguren, T., Vargas Baecheler, J., Chamorro Giné, A., and de Solminihac Tampier, H.: 465 Reliability-based estimation of traffic interruption probability due to road waterlogging, J. Adv. Transport., 2018, 1–12, https://doi.org/10.1155/2018/2850546, 2018.

Dembélé, M., Hrachowitz, M., Savenije, H. H. G., Mariéthoz, G., and Schaefli, B.: Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite data sets, Water Resour. Res., 56, 1–26, https://doi.org/10.1029/2019WR026085, 2020.

470 Gebremedhin, E. T., Basco-Carrera, L., Jonoski, A., Iliffe, M., and Winsemius, H.: Crowdsourcing and interactive modelling for urban flood management, J. Flood. Risk. Manag., 13, https://doi.org/10.1111/jfr3.12602, 2020.

Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and

noncommensurable measures of information, Water Resour. Res., 34, 751–763, https://doi.org/10.1029/97WR03495, 1998.

Kasmalkar, I. G., Serafin, K. A., Miao, Y., Bick, I. A., Ortolano, L., Ouyang, D., and Suckale, J.: When floods hit the road: resilience to flood-related traffic disruption in the San Francisco Bay Area and beyond, Sci. Adv., 6, eaba2423, https://doi.org/10.1126/sciadv.aba2423, 2020.

Kong, X.: Data and code used in the article titled " A Bayesian updating framework for calibrating hydrological parameters of road network using taxi GPS data," Zenodo [dataset], https://zenodo.org/record/7294880#.Y2nLv3ZByUk, 2022.

Kong, X., Yang, J., Qiu, J., Zhang, Q., Chen, X., Wang, M., and Jiang, S.: Post-event flood mapping for road networks using taxi GPS data, J. Flood. Risk. Manag., 15, e12799, https://doi.org/10.1111/jfr3.12799, 2022.

Kramer, M., Terheiden, K., and Wieprecht, S.: Safety criteria for the trafficability of inundated roads in urban floodings, Int. J. Disast. Risk. Re., 17, 77–84, https://doi.org/10.1016/j.ijdrr.2016.04.003, 2016.

Le Coz, J., Renard, B., Bonnifait, L., Branger, F., and Le Boursicaud, R.: Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach, J. Hydrol., 509, 573–587, https://doi.org/10.1016/j.jhydrol.2013.11.016, 2014.

Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35, 233–241, https://doi.org/10.1029/1998WR900018, 1999.

Li, Z., Wang, C., Emrich, C. T., and Guo, D.: A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods, Cartogr. Geogr. Inf. Sc, 45, 97–110, https://doi.org/10.1080/15230406.2016.1271356, 2018.

Martínez-Gomariz, E., Gómez, M., Russo, B., and Djordjević, S.: A new experiments-based methodology to define the stability threshold for any vehicle exposed to flooding, Urban. Water. J., 14, 930–939, https://doi.org/10.1080/1573062X.2017.1301501, 2017.

Merz, B. and Thieken, A. H.: Flood risk curves and uncertainty bounds, Nat. Hazards., 51, 437–458, https://doi.org/10.1007/s11069-009-9452-6, 2009.

Meteorological Bureau of Shenzhen Municipality: Rainfall Intensity Formula of Shenzhen, http://weather.sz.gov.cn/qixiangfuwu/qihoufuwu/qihouguanceyupinggu/baoyuqiangdugongshi/, 2015.

Moore, K. A. and Power, R. K.: Safe buffer distances for offstream earth dams, Australasian Journal of Water Resources, 6, 1–15, https://doi.org/10.1080/13241583.2002.11465206, 2002.

Natural Resources Conservation Service: Chapter 16 Hydrographs, in: National Engineering Handbook, United States Department of Agriculture, 3, 2007.

Nieto, N., Chamorro, A., Echaveguren, T., Sáez, E., and González, A.: Development of fragility curves for road embankments exposed to perpendicular debris flows, Geomat, Nat, Haz, Risk., 12, 1560–1583, https://doi.org/10.1080/19475705.2021.1935330, 2021.

Paul, J. D., Buytaert, W., Allen, S., Ballesteros-Cánovas, J. A., Bhusal, J., Cieslik, K., Clark, J., Dugar, S., Hannah, D.

M., Stoffel, M., Dewulf, A., Dhital, M. R., Liu, W., Nayaval, J. L., Neupane, B., Schiller, A., Smith, P. J., and Supper, R.: Citizen science for hydrological risk reduction and resilience building: Citizen science for hydrological risk reduction and resilience building, Wires. Water., 5, e1262, https://doi.org/10.1002/wat2.1262, 2018.

Pregnolato, M., Ford, A., Wilkinson, S. M., and Dawson, R. J.: The impact of flooding on road transport: A depth-disruption function, Transport. Res. D-Tr. E., 55, 67–81, https://doi.org/10.1016/j.trd.2017.06.020, 2017.

Qi, Y., Zheng, Z., and Jia, D.: Exploring the spatial-temporal relationship between rainfall and traffic flow: a case study of brisbane, Australia, Sustainability, 12, 5596, https://doi.org/10.3390/su12145596, 2020.

Restrepo-Estrada, C., Andrade, S. C., Abe, N., Fava, M. C., Mendiondo, E. M., and Albuquerque, J. P.: Geo-social media as a proxy for hydrometeorological data for streamflow estimation and to improve flood monitoring, Comput. Geosci., 111, 148–158, 2018.

Safaei-Moghadam, A., Tarboton, D., and Minsker, B.: Estimating the likelihood of roadway pluvial flood based on crowdsourced traffic data and depression-based DEM analysis, Natural Hazards and Earth System Sciences (preprint), https://doi.org/10.5194/nhess-2022-77, 2022.

Shah, S. M. H., Mustaffa, Z., and Yusof, K. W.: Experimental Studies on the threshold of vehicle instability in floodwaters, Jurnal Teknologi, 80, https://doi.org/10.11113/jt.v80.11198, 2018.

Shand, T. D., Cox, R. J., Blacka, M. J., and Smith, G. P.: Australian rainfall and runoff project 10: appropriate safety criteria for vehicles, Australian Rainfall & Runoff, 2016.

She, Zhong, Fang, Zheng, and Zhou: Extracting Flooded Roads by Fusing GPS Trajectories and Road Network, ISPRS. Int. Geo-Inf., 8, 407, https://doi.org/10.3390/ijgi8090407, 2019.

Versini, P.-A., Gaume, E., and Andrieu, H.: Application of a distributed hydrological model to the design of a road inundation warning system for flash flood prone areas, Nat. Hazard. Earth. Sys., 10, 805–817, https://doi.org/10.5194/nhess-10-805-2010, 2010.

Water Authority of Shenzhen Municipality: List of 2015 Flood-prone Roads in Shenzhen, http://swj.sz.gov.cn/ztzl/ndmsss/yldzl/zrrxxb/content/post_2918436.html, 2015.

Water Authority of Shenzhen Municipality: List of 2019 Flood-prone Roads in Shenzhen, https://opendata.sz.gov.cn/data/dataSet/toDataDetails/29200_01403146, 2019.

Xia, J., Falconer, R. A., Xiao, X., and Wang, Y.: Criterion of vehicle stability in floodwaters based on theoretical and experimental studies, Nat Hazards, 70, 1619–1630, https://doi.org/10.1007/s11069-013-0889-2, 2014.

Yabe, T., Tsubouchi, K., and Sekimoto, Y.: Fusion of Terrain Information and Mobile Phone Location Data for Flood Area Detection in Rural Areas, in: 2018 IEEE International Conference on Big Data (Big Data), 2018 IEEE International Conference on Big Data (Big Data), 881–890, https://doi.org/10.1109/BigData.2018.8622156, 2018.

Yao, Y., Wu, D., Hong, Y., Chen, D., Liang, Z., Guan, Q., Liang, X., and Dai, L.: Analyzing the effects of rainfall on urban traffic-congestion bottlenecks, IEEE J. Sel. Top. Appl., 13, 504–512, https://doi.org/10.1109/JSTARS.2020.2966591, 2020.

https://doi.org/10.5194/hess-2023-7
Preprint. Discussion started: 24 January 2023


Hydrology and Earth System Sciences
Discussions

Yin, J., Yu, D., Yin, Z., Liu, M., and He, Q.: Evaluating the impact and risk of pluvial flash flood on intra-urban road network: A case study in the city center of Shanghai, China, J. Hydrol., 537, 138–145, https://doi.org/10.1016/j.jhydrol.2016.03.037, 2016.

Zhang, W., Li, R., Shang, P., and Liu, H.: Impact Analysis of Rainfall on Traffic Flow Characteristics in Beijing, Int. J. ITS Res., 17, 150–160, https://doi.org/10.1007/s13177-018-0162-x, 2019.