Summary

The paper presents a novel approach for calibrating an urban rainfall-runoff model using taxi GPS data. This is an original idea that seems to have potential as demonstrated in this study. I also commend the authors for making available their data and code.

Response:

We thank the reviewer for the constructive comments to help us improve the manuscript. We are pleased that the manuscript aroused the reviewer's interest and are thankful for the positive feedback. Our responses to your comments are provided below. The data and code used to validate the method are available at Zenodo (https://doi.org/10.5281/zenodo.7894921).

Major comments

-What's the reason for modeling the taxi data as pass/no pass instead of directly modeling the number of taxis passing? The latter somehow seems more obvious since the original data are taxi counts, while your approach first requires converting taxi counts to 0/1 values, which introduces a potential loss of information. Please better justify this modeling choice.

Response:

We attempted to utilize taxi counts as an indicator of road status. However, establishing a relationship between precipitation or runoff and taxi count proved to be challenging. In situations where the road is not entirely disrupted by runoff, a stable quantitative relationship between taxi count and precipitation or runoff is hard to capture, as most taxis do not alter their travel routes when the water level is not too high. Consequently, estimating $\mathscr{L}(X|\Omega^{(i)})$ becomes difficult when X represents taxi volume. In contrast, when the road is disrupted, the taxi volume should be zero by the definition of disruption, and typically greater than zero when the road is open, simplifying the estimation of $\mathscr{L}(X|\Omega^{(i)})$. Overall, if the method calibrates parameters based on the taxi count, it may select the optimal parameter that best corresponds to the observed taxi count rather than the road disruption status, thereby introducing unnecessary errors.

-An alternative approach would be to use the Poisson distribution to directly model the number of taxis passing (rather than arriving). Have you considered this? This would be more like a Poisson regression model, but perhaps leveraging your road disruption function to model lambda instead of the usual Poisson link function.

Response:

We appreciate your suggestion. As stated in our previous response, this approach requires establishing a numerical relationship between the road disruption function and lambda, which denotes the mean of the 5 min taxi volume. Since most taxis do not modify their travel routes during less intense runoff, the effect of rainfall on taxi volume becomes difficult to capture.

-Did you check whether the Poisson model for the number of taxis arriving at a road is a good assumption for your data?

Response:

In the revised manuscript, a Chi-square goodness-of-fit test is conducted to check whether the frequency distribution adheres to a Poisson distribution. The squared differences between the observed and expected 5 min taxi frequencies, as predicted by the Poisson distribution, are computed to construct the Chi-square statistic. This statistic is then compared to the critical value corresponding to a significance level of 0.05. If the Chi-square statistic exceeds the critical value, the null hypothesis—that the probability distribution follows a Poisson distribution—is rejected; otherwise, the null hypothesis is accepted, suggesting the distribution may conform to a Poisson distribution.

Ten roads illustrated in Figure 8 of the manuscript were selected, and the frequency distribution of 5 min taxi volume for each period on each road was derived. Each frequency distribution consists of 31 samples, representing the taxi volume during a specific 5 min period collected from May 1, 2015, to May 31, 2015, for a specific road. The Chi-square goodness-of-fit test was applied to each frequency distribution, and the proportion of periods following a Poisson distribution for each road was calculated. The test results are presented in Table 1. According to these results, the frequency distribution of the 5 min taxi volume during a specific period adheres to the Poisson distribution for more than 50% of the periods in 7 out of the 10 roads. Consequently, the Poisson model appears to be a suitable assumption in this study.

Table 1 Chi-square goodness-of-fit test of Poisson distribution for 10 roads.

Road ID	1	2	3	4	5	6	7	8	9	10
Proportion of periods that follow Poisson distribution %	12.5	95.8	79.2	97.5	85.8	27.5	96.7	48.3	73.3	95.8

-A limitation is that all variables are treated as discrete random variables whereas the hydrological model parameters are continuous. Why discretize the parameters?

Response:

The reason to discretize parameters stems from the challenges associated with solving optimal problems. While continuous parameters may yield more accurate estimations, it is often arduous to obtain an analytical or numerical solution for $\theta^{(i)}$ that maximizes $P(\theta^{(i)})\mathscr{L}(X|\theta^{(i)})$ from a continuous parameter space. For instance, finding the analytical solution which maximizes $P(\theta^{(i)})\mathscr{L}(X|\theta^{(i)})$ necessitates differentiating $\mathscr{L}(X|\theta^{(i)})$, which may not always be feasible.

Given a catchment area and a time of concentration, constructing a synthetic unit hydrograph (SUH) based on the SCS unit hydrograph is straightforward. However, determining these two parameters from a given SUH poses a significant challenge. Similarly, generating runoff by combining rainfall with the SUH through a convolution formula is easy, but deriving the SUH from rainfall and runoff is difficult. Chow et al. (1988) demonstrated deconvolution methods such as matrix calculations or linear programming to derive the SUH, but these approaches are complex and cannot provide explicit functions that input runoff and output SUH. To circumvent the backward parameter-solving process, we discretize continuous parameters and calculate $P(\theta^{(i)})\mathscr{L}(X|\theta^{(i)})$ for each parameter set using a forward calculation process, which is more convenient in this study. It is important to note that the exclusion of continuous parameters in this study is due to the complexity of differentiating the proposed three-step procedure, not an indication that they are inapplicable to other procedures.

-Does the model account for other (non-flooding) factors that may affect the number of taxis in a road, e.g. time of day (rush hour)?

Response:

The time-of-day variation may affect the taxi volume in a road. In the previous version of our manuscript, we did not account for variations in taxi volume concerning the time-of-day or day-of-week. We assumed that the average number of taxis arriving on the road was constant, and the no-taxi-passing probability is given by:

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

where λ is the average taxi volume per 5 min interval, calculated by averaging all 5 min taxi volumes using historical taxi GPS data for a specific road.

However, the value of λ fluctuates according to the time of day, exhibiting higher taxi volume during congested periods and lower volume during non-congested periods. In the revised version, we incorporated the time-of-day variation in taxi volume when computing the no-taxi-passing probability:

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

where λ_t is the 5 min taxi volume during the tth period, calculated by averaging the taxi volume of the tth period from May 1, 2015, to May 31, 2015. Compared with λ, λ_t has smaller deviance because it excludes more non-flooding factors.

-Curve number CN is kept fixed even though it is also uncertain.

Response:

Thank you for pointing this out. We acknowledge that fixing the curve number as 85 is not realistic as it is influenced by various factors in urban areas, such as impervious surface percentage and soil type. Therefore, we have revised the manuscript to include curve number as one of the parameters to be calibrated. In total, we calibrate three parameters: catchment area, time of concentration, and curve number.

Figure 1 presents the probability distributions of three parameters after calibration. Each row in Fig.1 represents a different road, and each column represents the curve number. Furthermore, each subplot shows the joint probability distributions of catchment area and time of concentration given the curve number. The depth of colour in Fig.1 represents the magnitude of probability. Following two iterations of updating, the posterior probability distribution for both catchment area and time of concentration converges around the optimal parameter sets for most flood-prone roads. This demonstrates that incorporating taxi observations has substantially narrowed the uncertainty associated with catchment area and time of concentration. The probability typically attains its maximum value when the curve number is either 55 or 60. Moreover, each subplot presents a salient cluster with higher probability than other regions, suggesting that there may be multiple parameter sets which can effectively represent the acceptable ones.

Additionally, it is observed that the optimal catchment area under a given curve number diminishes as the curve number rises, and the optimal time of concentration under a given curve number increases in relation to the curve number. This is logical, as a higher curve number corresponds to increased rainfall excess given identical rainfall conditions, necessitating a decrease in catchment area to maintain the runoff that best aligns with the observed taxi-related road conditions. Similarly, the increase in time of concentration compensates for the additional runoff generated by a higher curve number, also preserving the optimal runoff status.

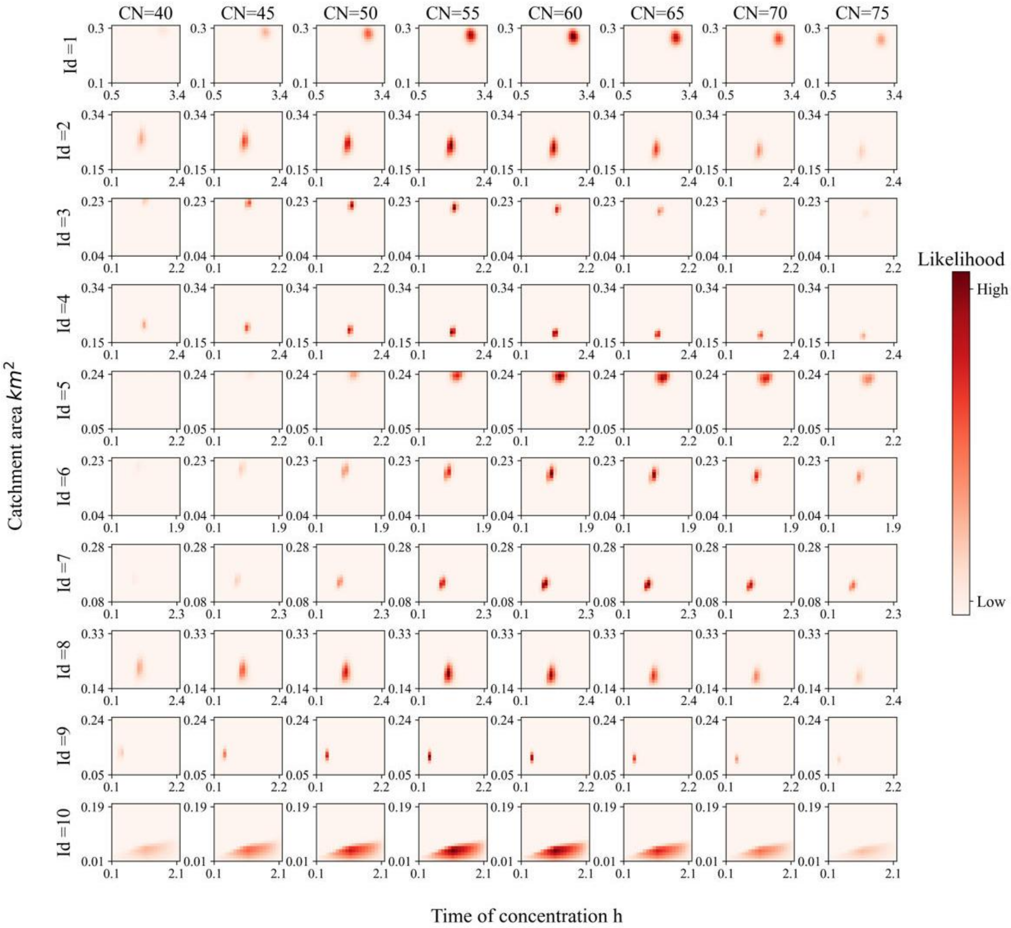

Figure 1 Posterior probability distributions of hydrological parameter sets for 10 flood-prone roads after calibration. The prior probability distributions are derived from the DEM and additional prior knowledge.

In addition, we plotted the marginal distributions of three parameters for ten roads before and after calibration in Fig.2. Upon examining Fig.2, the marginal posterior distributions of curve number post-calibration appear relatively similar to the marginal prior distributions of curve number. It seems that the method employing taxi data offers limited information about the distribution of curve number after calibration in comparison to catchment area and time of concentration. This outcome may be ascribed to the range and discretization granularity of the parameter spaces. Catchment area and time of concentration encompass 20 and 30 possible values, respectively, whereas the curve number includes only eight potential values. The smaller parameter space of curve number diminishes the search space, and its impact on the no-taxi-passing probability is comparatively lower than that of catchment area and time of concentration.

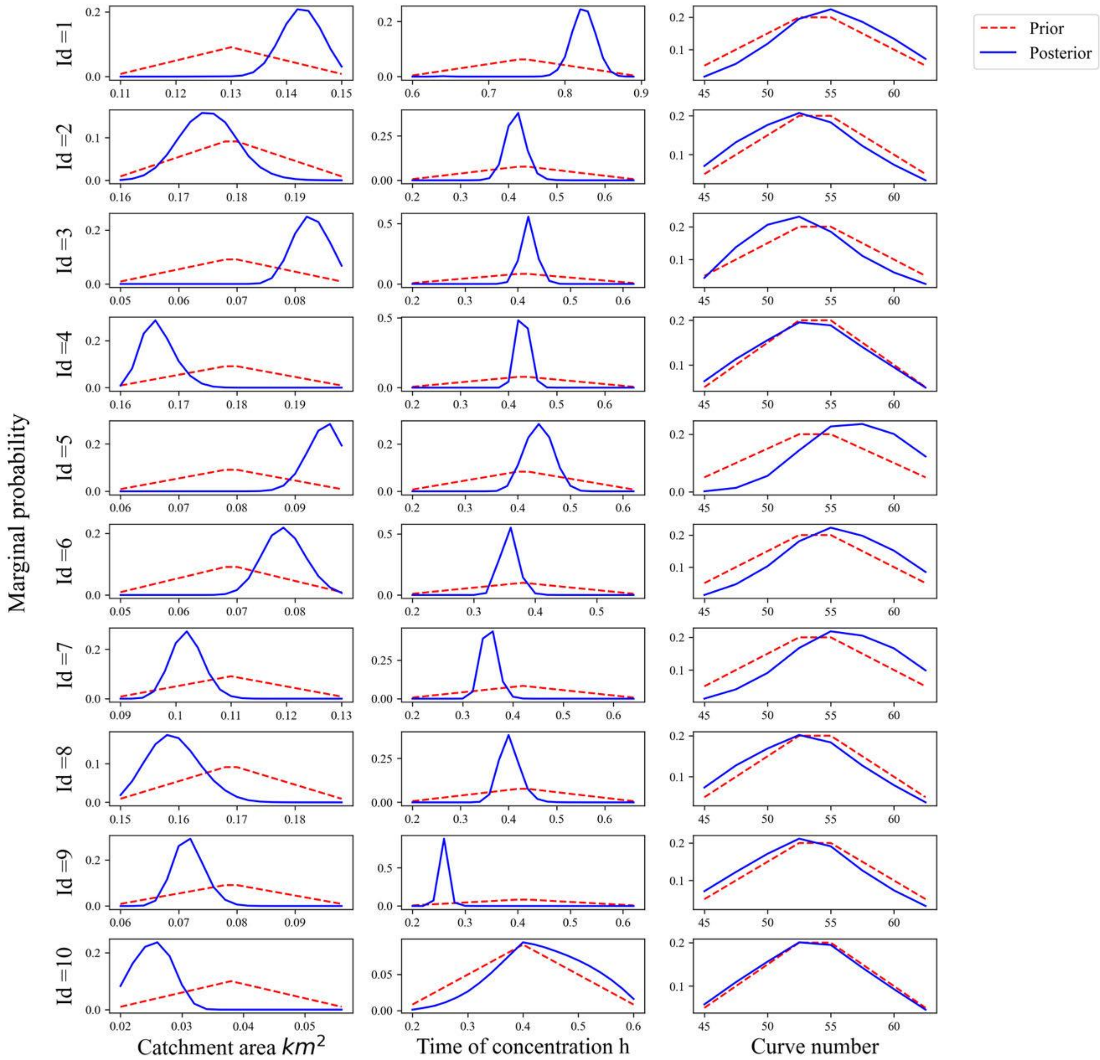

Figure 2 Marginal prior and posterior probability distribution of curve number for 10 flood-prone roads.

For instance, for road ID=6, the optimal parameter set maximizing the no-taxi-passing probability consists of a catchment area of 0.19 km², a time of concentration of 0.9 hour, and a curve number of 55. To investigate the effects of these parameters on the hydrograph and the time series of no-taxi-passing probability, we held two parameters constant at their optimal values and observed the impact as the third parameter varied. Our findings, illustrated in Fig.3, demonstrate that when the catchment area varies from 0.04 km² to 0.23 km², the maximum no-taxi-passing probability ranges from 20% to 100%, and the duration of no-taxi-passing probability exceeding 0.5 hour extends from 0.0 to 1.3 hours. Similarly, when the time of concentration fluctuates from 0.1 to 1.9 hour, the peak time of no-taxi-passing probability spans from 0.5 to 1.8 hour. In contrast, when the curve number varies from 40 to 75, the maximum no-taxi-passing probability is fixed at 100%, the duration of no-taxi-passing probability excessing 0.5 hour extends from 1.1 to 1.3 hours, and the peak time of no-taxi-passing probability remains fixed at the 1.1 hour. The smaller fluctuations in the time series of no-taxi-passing probability interpret why the distribution of curve number remains relatively stable after calibration compared to catchment area and time of concentration.

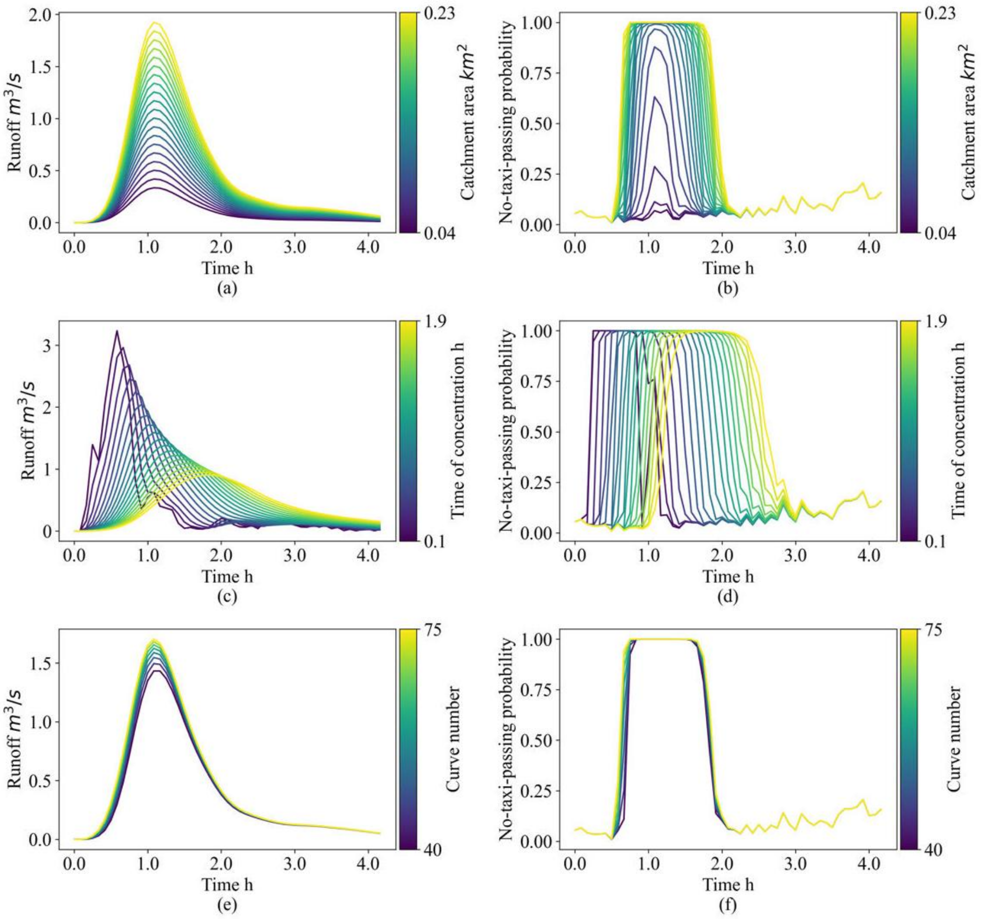

Figure 3 Impacts of three parameters on the variation of time series of runoff and no-taxi-passing probability. (a) Catchment area on the runoff. (b) Catchment area on the no-taxi-passing probability. (c) Time of concentration on the runoff. (d) Time of concentration on the no-taxi-passing probability (e) Curve number on the runoff. (f) Curve number on the no-taxi-passing probability

-Section 2.1: a more common/general way is to write Bayes equation directly in terms of parameters theta, as in p(theta|X) \propto p(theta)*p(X|theta) or p(theta|X) \propto p(theta)*L(theta|X). The likelihood on the rhs of eq. 4 in the paper would then be written in terms of a function omega(theta) given by your eq. 16.

Response:

Thanks for your suggestion. The Bayes equation is rewritten as:

$$P(\theta^{(i)} \mid X) \propto P(\theta^{(i)}) \mathscr{L}(X \mid \theta^{(i)})$$

Editorial comments

-eq. 11: please define x and y

Response:

The expression of the fitting curve is:

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

where x is the product of flow velocity and flow depth, and y is the disruption probability.

-L23: metropolis --> metropolises or metropolitan areas

Response:

Modified as suggested.

-L40: "calibrated on runoff data alone" - there are many studies that calibrate on other data as well

Response:

Response:

It is not rigorous to say that "No matter what kinds of methods, hydrological models are basically calibrated based on the runoff data alone." Thus, that sentence is removed.

-L47: ungaged vs ungauged: pick one spelling

Response:

Modified to "ungaged" as suggested.

-L83 (and other places): equals to --> equals

Response:

Modified as suggested.

-L90: arriving --> arrival

Response:

Modified as suggested.

-L99: does index i refer to road i?

Response:

No, the index i refers to the ith parameter set.

-L132: instantization --> instantiation

Response:

Modified as suggested.

-suggest to proofread entire manuscript to fix issues with use of English

Response:

We thank the reviewer for pointing this out. As suggested, the manuscript is thoroughly proofread, and the grammar, clarity, and overall readability is also improved.