25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

Study on the impact of ubiquitous sensing of rainfall data on hydrological simulation

J. Wang, S.T. Khu and Y.N. Mao

School of Environmental Science and Engineering, Tianjin University, Tianjin, China
Engineering Research Center of City Intelligence and Digital Governance, Ministry of Education of the People's Republic of China, Tianjin, China
Email: soon.thiam.khu@tju.edu.cn

Abstract: Rainfall is an important basic element for simulating the hydrological cycle system of a region or basin. Rainfall data with high accuracy and spatial-temporal resolution are of great significance for industrial and agricultural production, water conservancy development, drought and flood monitoring and prevention, which are the basis for understanding regional hydrological processes and exploring regional hydrological changes. How to obtain rainfall data with high spatial and temporal resolution is an important research topic in the field of hydrometeorology.

Conventional rainfall measurement methods include specialized gauged station monitoring and remote sensing retrieval. Although traditional ground gauged network can obtain accurate rainfall at each station, they have obvious limitations due to the uneven distribution of ground stations and spatial discontinuity of rainfall data. Radar rainfall measurement and satellite rainfall estimation respectively have high resolution and wide space coverage, while their monitoring data often have significant errors. The rapid development of intelligent monitoring technology has brought significant possibilities for ubiquitous sensing of rainfall data (such as microwave links, camera sensing, portable sensors, etc.), which can be distributed with high spatial/temporal resolution and covering ratio.

As we know, runoff simulation accuracy depends on the density of rain gauges and rainfall monitoring error quantity. Although ubiquitous sensing technology can effectively solve the problem of insufficient rainfall data, the accuracy of ubiquitous data is lower than measurements from dedicated rainfall gauges, which is also an important reason for its lack of application in practice. The challenge posed by ubiquitous sensing is a potential explosion of data collected by multiple groups for different purposes, with differing accuracy, precision and hence data quality. It is worth noting that existing research fell short of stating the need also to investigate what happens if the quality of data collected is inherently poor, should these large amounts of poor quality data be discarded (with huge effort of collection wasted)? Or should we spend effort on refining the collection methodology so that we can improve the data quality?

This study focuses on the impact mechanisms of ubiquitous rainfall sensing on hydrology simulation. We explored how the ubiquitous monitoring site density, error quantity of sensing data and a combination of both factors affect the runoff simulation. The results show that : (1) The increased density of monitoring sites can effectively improve runoff simulation accuracy, especially when the monitoring site density is more than 100 km^2/1 site; (2) The increase in error quantity increases the range of variability in runoff simulation performance, particularly for monthly runoff simulations; (3) the error quantity of ubiquitous sensing data should be controlled within 30%, and the density of the site over more than 20 km^2/1 site. Our work is dedicated to proposing the applicability and providing a basis for ubiquitous sensing data in hydrological simulation. The resolution achievable using today's sensors will be sufficient to improve hydrology models, an important continuation of this work.

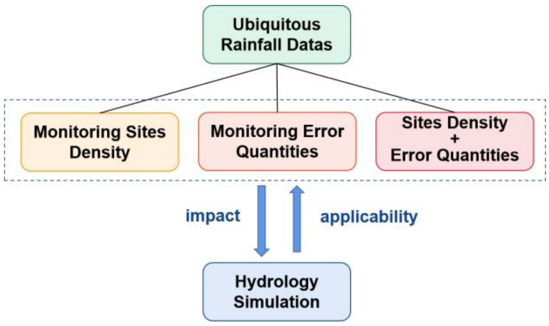


Figure 1. The impact study framework for ubiquitous sensing data on hydrological simulation

Keywords: *Ubiquitous sensing, rainfall data, hydrology model, applicability*

1. INTRODUCTION

With rapid urbanization and population growth, more problems in water resources make understanding the hydrological cycle a vital role in alleviating current and future water-related problems (Charles et al. 2000). Precipitation is the main driving force of the land part of the hydrological cycle, so accurate observation of precipitation is the key to these necessary improvements. However, recent years have seen a gradual decrease in investment in hydrological monitoring networks around the world, leading to problems such as insufficient monitoring data and irregular monitoring frequency, especially in remote areas or basins requiring monitoring of specific hydrological elements (Mazzoleni et al. 2017). So, precipitation observation and retrieval are still considered scientific challenges (Christian et al. 2019).

The traditional rainfall station has high construction, maintenance and management costs, and it is difficult to build stations in high-altitude, low-temperature and urban dense areas, and there are limited observation points that cannot truly reflect the spatial distribution of precipitation (Vladislav et al. 1999). Weather radar has a monitoring blind zone, obstacles affect the accuracy, and the low-altitude measurement error is large; The construction investment intensity is large, and the maintenance and management costs are high (Ulbrich et al. 1999). In terms of satellite remote sensing, satellite remote sensing can provide global rainfall observation data with different temporal and spatial resolutions, which provides a data basis for global change research and application, but there are also some problems, such as non-real-time rain measurement, low spatial resolution, 3 hours lower temporal resolution, poor inversion accuracy in middle and high latitudes and complex terrains (cities, mountainous areas, etc.). There is a serious lag that cannot meet the real-time needs of emergency processing (Hou et al. 2014; Gaona et al. 2017). Therefore, the high spatial and temporal resolution of rainfall data is still a challenge. In order to avoid the loss of hydrological information, it is necessary to explore cost-effective methods of data collection, especially in the context of diminishing resources where there is an urgent need to find new and reliable means of hydrological data collection to ensure the availability of extensive and continuous data sets.

With the advances in technology, small and low cost sensors are being embedded in every-day consumer products, including automobiles, hand-held devices (including cellular telephones), smart buildings, and traffic management systems (Hill et al. 2019). Sensors are becoming ubiquitous, generating data at an unprecedented rate and scale (Chen et al. 2022; Pan et al. 2017). The advances in ubiquitous sensing present an emerging opportunity to improve our capability to monitor the weather. Very recently, ubiquitous sensors, including smart phones (Yin et al. 2022), surveillance cameras (Lee et al. 2022), microwave communication links (Chwala and Kunstmann. 2019) and vehicle-based automatic windshield wiper sensors (Rabiei et al. 2016) have been explored for quantitative rainfall measurement. However, there are certain errors in the retrieval of rainfall through the ubiquitous sensing method. For example, based on the quantitative detection of rainfall intensity from images, that is, from images taken under rainy conditions, this method can be effectively applied to actual rainfall events, but the error is about 25%; The average absolute percentage error of the method for measuring rainfall intensity under real world conditions based on video collected by ordinary surveillance cameras is 21.8% (Allamano et al. 2015); Microwave link networks with different structures have different retrieval effects for different rainfall fields, the relative deviation of rainfall intensity retrieved by the microwave link rainfall retrieval model is mostly within 15% (Jiang et al. 2019). It is expected that the accuracy of ubiquitous sensor of rainfall measurements is lower than measurements from dedicated rainfall gauges. It can be seen that although the ubiquitous sensing means can make up for the shortcomings of the limited traditional measuring station and low resolution, there is still a certain range of errors, which is also an important reason for its lack of application in practice. It is worth noting that existing research fell short of stating the need also to investigate what happens if the quality of data collected is inherently poor, should these massive amounts of poor quality data be discarded (with huge effort of collection wasted)? Or should we spend effort on refining the collection methodology so that we can improve the data quality?

Therefore, this paper focuses on the impact of ubiquitous sensing of rainfall data on hydrological simulation, one that is cutting-edge research to advance the application of ubiquitous sensory data. We are intended to achieve two main objectives: firstly, to explore the mechanism by which errors in large amounts of ubiquitous sensing data affect the hydrological cycle simulation; and secondly, to propose the applicability in the monitoring process of ubiquitous sensing means.

2. METHOD

2.1. Implementation of the framework

In addition to a large amount of supplementary monitoring data, ubiquitous sensing of rainfall data also brings uncontrolled errors. The impact mechanism of the abundant sensing data with errors on hydrological simulation is unclear. Therefore, this study explores whether sensing data can contribute to improving simulations of hydrological models. Two variables of station density and ubiquitous sensing inversion error were selected for simulation by the SWAT model, to obtain the optimal distribution density and acceptable error range of the station. The detailed analysis method is summarized in the following steps and the flowchart is shown in Figure 2.

Step 1: Preparation of rainfall data sets. Design rainfall datasets with single factor variation and dual factor variation for monitoring site density and error quantity.

Step 2: Build the hydrology model. The hydrology model is employed for simulating the rainfall-runoff process. Meanwhile, the model parameters are calibrated using historical runoff data.

Step 3: Impact analysis. This paper analyzes the relationship between monitoring site density and runoff error, monitoring errors and runoff error, and site density and monitoring errors and runoff error by comparing simulated results.

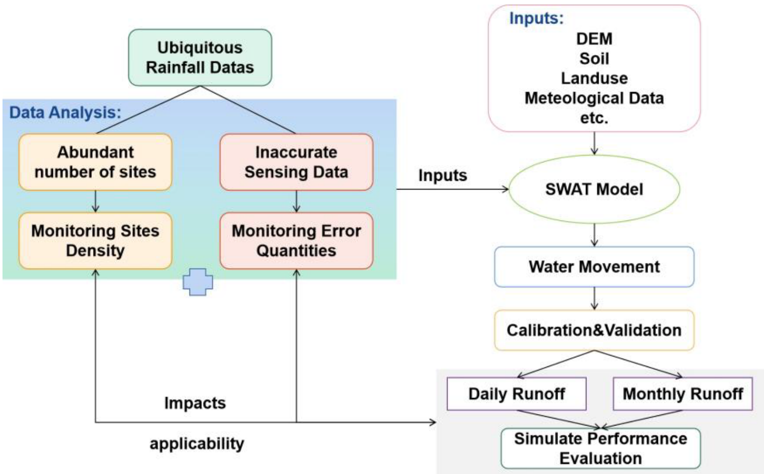


Figure 2. Flowchart of the influence study of ubiquitous sensing of rainfall data on hydrological simulation

Step 4: Applicability evaluation. The simulation will be used to explore the tradeoffs between increasing spatial density of the ubiquitous sensors and increased sensing errors.

2.2. Hydrology model

SWAT model is a semi-distributed physically based hydrological model, which was developed by the United States Department of Agriculture (USDA). The model can use the spatial data obtained by GIS and RS to conduct numerical simulations of large-scale and complex rivers under various soil, land use, climate conditions, and human activities. The advantages of this model are that it has strong physical mechanisms, high computational efficiency and high accuracy. The model needs to input digital elevation data (DEM), land use data and soil type, as well as meteorological and flow data of monitoring stations, and then automatically divide the sub-basin and hydrological response unit (HRU), and separately calculate and simulate each HRU.

SWAT water movement calculations include SCS and Green-Ampt models. Among them, the SCS runoff curve number method is used more, and the model has three basic assumptions: the existence of the maximum soil water storage capacity S; The ratio between the actual water storage F and the maximum water storage capacity S is equal to the ratio of the difference between runoff Q and rainfall P and initial loss I_a. The relationship between Ia and S is linear. The rainfall-runoff relationship is expressed as follows :

$$\frac{F}{S} = \frac{Q}{P - I_a}$$

$$I_a = aS$$

where a is a constant, which is generally taken as 0.2 in SCS models. It is noteworthy that relative parameters calibration of the SWAT model is carried out by SWAT-CUP based on the observed data.

2.3. Experimental protocol

The spatial distribution of rainfall monitoring sites largely determines the accuracy of rainfall simulation. After the study area is selected, the grid is divided according to the density of monitoring stations in the study area. Establishing a benchmark for spatiotemporal distribution of rainfall based on national rainfall gauges and the Kriging interpolation method. Then, N different monitoring station network densities and spatial distribution modes are set up to evaluate the difference of rainfall process under different monitoring station network density and spatial distribution mode, and clarify the impact of monitoring site density on runoff simulation.

At the same time, different ubiquitous sensing methods have different error ranges, which increases the uncertainty of the rainfall process. This paper considers *M* error modes and assumes they all follow normal distribution functions. Introduce random error into the benchmark rainfall data of *N* station schemes by the Monte Carlo sampling method, the new rainfall data as ubiquitous sensing data samples. These ubiquitous sensing sample data are without the limitation of system multi-dimensionality, multi-factor and other complexity. Finally, based on the simulation of the SWAT model, the optimal error range of the ubiquitous data is obtained.

3. STUDY AREA AND DATA

3.1. Study area

In this paper, Jianjiang River Basin is selected, which is a tributary of the Yangtze River, located in Duyun City, Guizhou province, and its watershed area is 2,158.8 km^2. The watershed altitude ranges from 589 to 1,938 *m* above mean sea level (as shown in Figure 3). The headwaters originate from Doupeng Mountain and flow 91.2 km through Duyun city. The area receives an average annual precipitation of 1,431 mm, with pronounced seasonality.

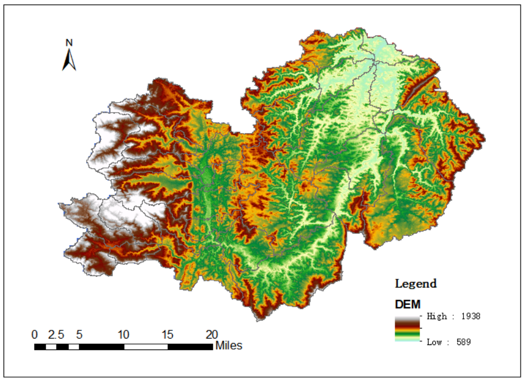


Figure 3. The DEM of Jianjiang River Basin in Guizhou province of China

The available streamflow and meteorological information of the Jianjiang River Basin is limited. The available climate data from the Duyun meteorological station including daily values of maximum and minimum air temperature, solar radiation, relative humidity, and wind speed were obtained from the National Science and Technology Infrastructure of China Meteorological Administration, and are used as meteorological input variables to the SWAT model. It should be noted that the calibration parameters of SWAT model of this study area can refer to the author’s previous research (Wang et al. 2019).

3.2. Rainfall data

In this paper, we take the 2012 rainfall data as an example to study the impact of ubiquitous sensing rainfall data on hydrological simulation. From Figure 4, the rainfall distribution in the basin gradually decreases from upstream to downstream, with a maximum difference in rainfall of nearly 150mm. However, there are only four rainfall gauged stations in the basin, which are concentrated in urban areas in the upper of the basin (the green points in Figure 4). As we know, the observed data from four gauged stations are not sufficiently representative. Further, we analyze whether the abundance of ubiquitous sensing data contributes to hydrological modelling.

Using data from surrounding gauged stations, the rainfall data for the basin was interpolated at 0.02° intervals between latitude and longitude by the Kriging interpolation method (as seen in Figure 5). The 498 points in Figure 5 are assumed ubiquitous monitoring sites which is the basic scheme. There are 12 ubiquitous monitoring site density gradients (T1～T12) set up based on the assumed monitoring sites network layout (in Table 1).

Considering the different error quantities of monitoring sites, five error schemes (E1~E5) are set in this paper, which are shown in Table 1. The random samplings of the rainfall datasets are based on the normal distribution of the E1~E5 scheme. To avoid accidental sampling, the number of iterations is 500. Finally, we explore the impact of abundant ubiquitous monitoring sites with different errors on hydrological simulation, which schemes combine the T1~T12 and E1~E5.

Table 1. The impacts of the density and error quantity of ubiquitous monitoring sites on runoff simulation

Schemes	Variables
Basic scheme	498 ubiquitous monitoring sites without error
Actual gauged scheme	4 actual gauged stations without error
The density of ubiquitous monitoring sites scheme	T1:498 sites, 4 km^2/sites; T2:248 sites, 9 km^2/sites; T3:123 sites, 18 km^2/sites; T4:99 sites, 22 km^2/sites T5:49 sites, 44 km^2/sites; T6:34 sites, 63 km^2/sites; T7:25 sites, 86 km^2/sites; T8:19 sites, 113 km^2/sites T9:12 sites,180 km^2/sites; T10:9 sites, 240 km^2/sites; T11:7 sites, 308 km^2/sites; T12:5 sites, 430 km^2/sites
Error quantity of ubiquitous monitoring sites scheme	E1: ubiquitous monitoring sites with 10% error; E2: ubiquitous monitoring sites with 20% error E3: ubiquitous monitoring sites with 30% error; E4: ubiquitous monitoring sites with 40% error E5: ubiquitous monitoring sites with 50% error
Dual variation scheme of density and error quantity	T1~T2 combined with E1~E5

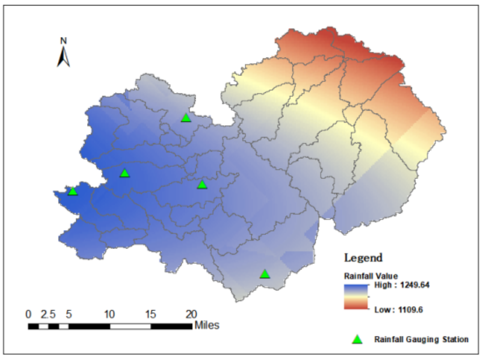


Figure 4. Spatial distribution of 2012 annual rainfall of Jianjiang River Basin

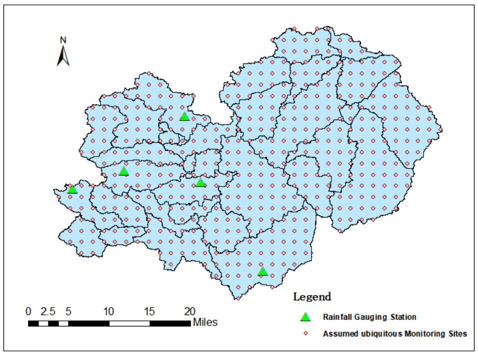


Figure 5. Ubiquitous monitoring network layout

4. RESULTS AND DISCUSSION

4.1. Effect of monitoring sites density on simulating runoff

The effect of the density of ubiquitous sensing data on model enhancement is analyzed by simulating T1 to T12 schemes without error. In this paper, we adopted the Nash-Sutcliffe model efficiency (NS) and Root Mean Squared Error (RMSE) to evaluate the performance and quantify the impact of the simulation. From Figure 6, it is not difficult to find that with the increase in monitoring density, the accuracy of runoff simulation continues to improve, especially for monthly runoff simulation. At the same time, it can be concluded that when the number of sites is 9 (density: about 250 km^2/sites), the monthly runoff simulation effect increases rapidly, and then the increase slows down; The turning point of daily runoff simulation is at the point where the number of monitoring sites is equal to 19 (density: about 100 km^2/sites).

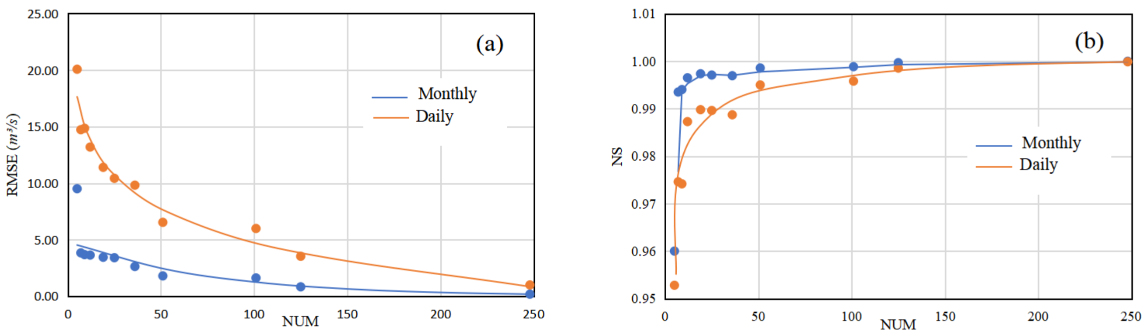
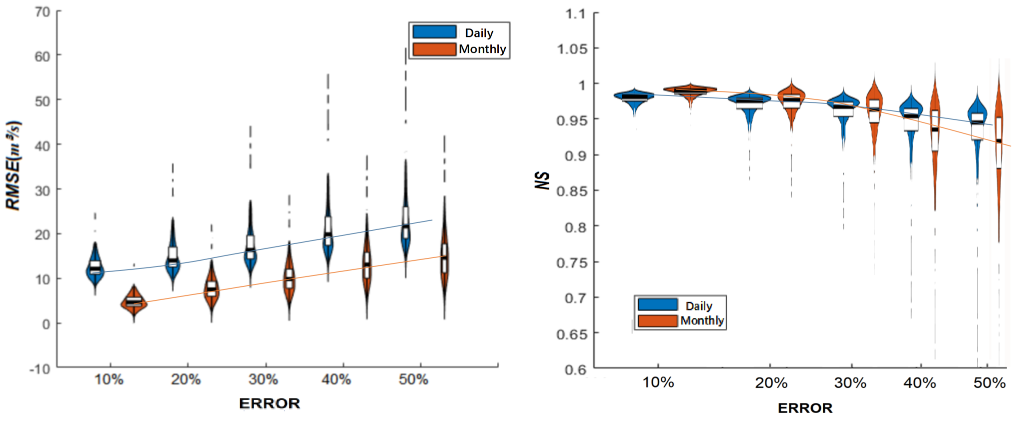


Figure 6. The relationship between monitoring density and runoff simulation performance (the X-axis and Y-axis represent the number of monitoring sites and performance evaluation indicators, respectively)

4.2. Effect of sensing rainfall errors on simulating runoff

Figure 7 shows an example of the different error schemes under 19 monitoring sites in Jianjiang River. As the amount of error increases, the performance variation range of runoff simulation continues to increase. For the RMSE index, the monthly runoff simulation performance is better than the daily runoff simulation under different error schemes. However, for NS, when the error quantity is greater than 30%, daily runoff simulation performs better than monthly runoff simulation. As can be seen the trend orange and blue lines from Figure 7 (b), it is better to control the error range at 30% (turning point).

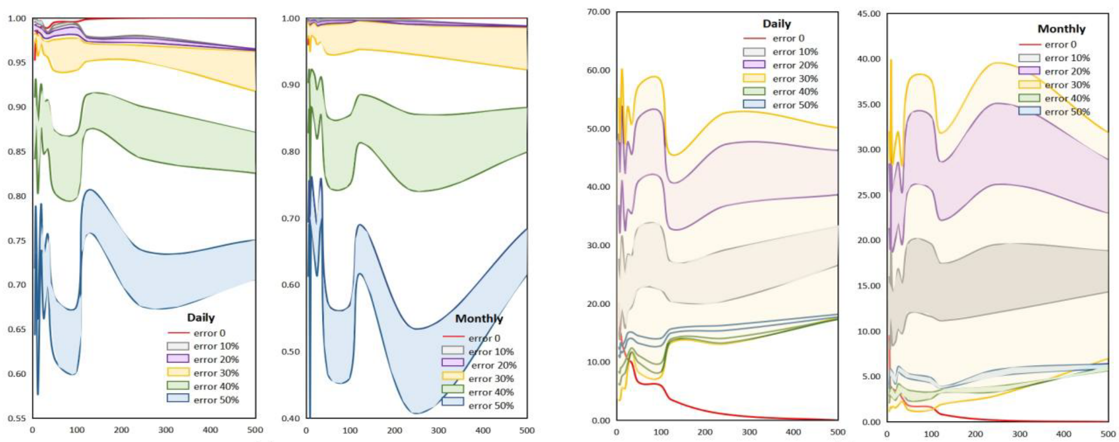


(a) (b)

Figure 7. The relationship between monitoring error quantity and runoff simulation performance

4.3. Effect of the coupling effect of monitoring sites density and sensing errors on simulating runoff

Based on 4.1 and 4.2 an analysis of how the dual variation in rainfall sensing data volume and error affects the hydrological simulations is presented, and thus the scope of application of ubiquitous sensing techniques. From Figure 8, when the number of monitoring sites is less than 100, the sensing error has a large random effect on the runoff simulation; Conversely, for monitoring sites at less than 20 km^2/site, the effect of error on runoff tends to stabilize. When the error quantity is greater than 30%, the range of influence on the range of NS variation is larger and NS performance decreases rapidly. In particular, for the RMSE metric, the range of RMSE variation is greatest when the error equals 30%. Therefore, in the practical application of ubiquitous sensing data, the error quantity needs to be controlled to within 30%, while the density of the monitoring sites should be less than 20 km^2/site.



(a) NS (b) RMSE

Figure 8. The impact of variation of monitoring density and error quantity on runoff simulation performance

5. CONCLUSION

This paper quantified the impact of different levels of error and different spatial densities of ubiquitous sensing data on a hydrology model. Through a case study conducted in the Jian-jiang River basin, the following conclusions can be drawn: (1) The increased density of monitoring sites can effectively improve runoff simulation accuracy, especially when the monitoring sites density is greater than 100 km^2/site; (2) The increase in error quantity increases the range of variability in runoff simulation performance, particularly for monthly runoff simulations; (3) the error quantity of ubiquitous sensing data should be controlled within 30%, and the site density of more than 20 km^2/site. This study is only a preliminary exploration, these conclusions may be limited by the random distribution of errors, uniform parameters, the length of the rainfall series and so on. Therefore, it is necessary to conduct more comprehensive research in the follow-up research. Our work is dedicated to proposing the applicability and providing a basis for ubiquitous sensing data in hydrological simulation. The resolution achievable using today's sensors will be sufficient to improve hydrology models, an important continuation of this work.

ACKNOWLEDGEMENTS

This research is partially supported by the National Natural Science Foundation of China (52079144).

REFERENCES

Allamano, P., Croci, A., Laio, F., 2015. Toward the Camera Rain Gauge. Water Resources Research 51(3), 1744-1757.

Chen, A.B., Goodall J.L., Chen T.D., Zhang, Z., 2022. Flood Resilience Through Crowdsourced Rainfall Data Collection: Growing Engagement Faces Non-Uniform Spatial Adoption. Journal of Hydrology 609, 127724.

Chwala, C., Kunstmann, H., 2019. Commercial Microwave Link Networks for Rainfall Observation: Assessment of the Current Status and Future Challenges. Wiley Interdisciplinary Reviews: Water 6(2), e1337.

Gaona, M.F.R., Overeem, A., Brasjen, A.M., Meirink, J.F., Leijnse, H., Uijlenhoet R., 2017. Evaluation of Rainfall Products Derived from Satellites and Microwave Links for the Netherlands. IEEE Transactions on Geoscience and Remote Sensing 55(12), 6849-6859.

Hill, D.J., Farzan, F., 2012. Downscaling Radar-Rainfall Data via Ubiquitous Sensing and Data Fusion. World Environmental and Water Resources Congress 2012: Crossing Boundaries: 729-737.

Hou, A.Y., Kakar, R.K., Neeck,S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The Global Precipitation Measurement Mission. Bulletin of the American Meteorological Society 95(5), 701-722.

Jiang, S.J., Babovic, V., Zheng, Y., Xiong, J., 2019. Advancing Opportunistic Sensing in Hydrology: A Novel Approach to Measuring Rainfall with Ordinary Surveillance Cameras. Water Resources Research 55(4), 3004-3027.

Lee, J., Byun, J., Baik, J., Jun, C., Kim, H.J., 2022. Estimation of Raindrop Size Distribution and Rain Rate with Infrared Surveillance Camera in Dark Conditions. Atmospheric Measurement Techniques Discussions 16(3), 707-725.

Mazzoleni, M., Verlaan, M., Alfonso, L., Monego, M., Norbiato, D., Ferri, M., Solomatine D.P., 2017. Can Assimilation of Crowdsourced Data in Hydrological Modelling Improve Flood Prediction? Hydrology and Earth System Sciences 21(2), 839-861.

Nešpor, V., Sevruk, B., 1999. Estimation of Wind-Induced Error of Rainfall Gauge Measurements Using a Numerical Simulation. American Meteorological Society 16(4), 450-464.

Rabiei, E., Haberlandt, U., Sester, M., Fitzner, D., Wallner, M., 2016. Areal Rainfall Estimation Using Moving Cars-Computer Experiments Including Hydrological Modeling. Hydrology and Earth System Sciences 20(9), 3907-3922.

Ulbrich, C.W., Lee, L.G., 1999. Rainfall Measurement Error by WSR-88D Radars due to Variations in Z-R Law Parameters and the Radar Constant. Journal of Atmospheric and Oceanic Technology 16(8), 1017-1024.

Vörösmarty, C.J., Green, P., Salisbury, J., Lammers, R.B., 2000. Global Water Resources: Vulnerability from Climate Change and Population Growth. Science 289(5477), 284–288.

Wang J., Zhang X.H., Xu C.Y., Wang H., Lei X.H., Wang X., Li S.Y., 2019. Development of load duration curve system in data-scarce watersheds based on a distributed hydrological model. Hydrology Research 50(3), 886:900.

Yang, P., Ng, Tze Ling, 2017. Gauging Through the Crowd: A Crowd-Sourcing Approach to Urban Rainfall Measurement and Storm Water Modeling Implications. Water Resources Research 53(11), 9462-9478.

Yin H, Zheng F, Duan H F, Savic, D., Kapelan, Z., 2022. Estimating Rainfall Intensity Using an Image-Based Deep Learning Model. Engineering. 2(21), 162-174.