Advances in Reliability, Safety and Security

ESREL 2024
Collection of Extended Abstracts

Improving Spatial Monitoring Of Rain Gauges For Flood Assessment

Chengjun Cao[a], Franck Schoefs[b]

[a] *Nantes Université, École Centrale Nantes, CNRS, GeM, UMR 6183, F-44000 Nantes, France*
[b] *Nantes Université, École Centrale Nantes, CNRS, GeM, UMR 6183, IUML, FR 3473, F-44000 Nantes, France*

Keywords: uncertainty, roc curve, cross validation, probability of detection, probability of false alarm

Flood assessment relies on reliable rainfall data. One of the main sources of rainfall data is the rain gauge network. Therefore, an assessment of the current performance of rain gauge stations can help to make recommendations for improving the accuracy of rainfall monitoring. In this study, the performance of different stations is judged by investigating the uncertainty of the rain gauge network in the Landro basin, located in the north of Spain.

Firstly, for the precipitation used in this study, a spatial interpolation method was utilized to combine rain gauge data with satellite data to obtain a spatially interpolated estimate of rainfall. This method allows the sparsely distributed but highly accurate rain gauge data to be fused with the closely spaced but less accurate satellite data to obtain a rainfall that is closer to the true rainfall. This method is now widely studied, (Sun et al., 2000) used Kriging method to merge rain gauges with radar data for optimal unbiased estimation. And (Sadat-Noori et al., 2013) obtained that the kriging with external drift method gives the most coherent results according to cross-validation.

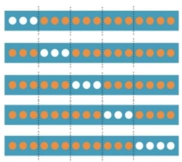

Fig. 1. Schematic diagram of cross-validation.

Cross-validation is one of the most commonly used resampling methods and an important tool in many practical applications of statistical learning processes. It can be used to estimate the test errors associated with a given statistical learning method in order to evaluate its performance or to select the appropriate level of flexibility. (Xu et al., 2018) compute the uncertainty of a new rain gauge network by the cross-validation method. This is also the principal method used nowadays to quantify the uncertainty of the network of rain gauge stations. (Marco et al., 2022) use cross-validation to test machine learning-based models used to predict flood risk. Whereas in this study firstly cross validation is utilized for error calculation then The Receiver Operating Characteristic (ROC) Curve is utilized for uncertainty study.

In the cross-validation process, the data is divided into two parts, i.e., training set and validation set. As shown in Figure 1. The white dots are the validation set and the yellow dots are the test set. In this study, k-fold cross-validation is used. The data is divided into k groups, one of which is set as the validation set and the remaining k-1 groups are the training set. The training set is then used to perform calculations by substituting it into the model to obtain the data corresponding to the validation set. This in turn enables the calculation of the error between the validation set and the results obtained from the model training. This loop is repeated, using a different group as the validation set in each loop, until all the data has been validated. Thus, the error situation of the whole set of data in the model is obtained.

The Receiver Operating Characteristic Curve (ROC) is an analytical tool for selecting the best detection signal (BOERO et al., 2009). It can give neutral and objective recommendations. The ROC curve is particularly valuable when dealing with binary classification problems, where the performance of a model in distinguishing between two classes needs to be thoroughly examined. (Rouhan and Schoefs, 2003) analyzed and made recommendations for the maintenance of coastal structures using ROC curves in their study. In this study, the ROC curve is used to analyze the performance of the rainfall station in the same way it is used for other NDT performance assessment (Schoefs et al., 2009). The largest the error on a small signal, the worst will be the decision. The ROC curve is a curve based on the Probability of Detection (PoD) and probability of false alarm (PFA). The probability of detection is derived from the signal that is synthesized in the sensor as an aggregate value of rainfall and disturbance errors. And the probability of false alarm originates from the error. As shown in Figure 2, the plot shows the probability density curves for the error and the combination of error and rainfall with threshold equal to -2 and 0. Then different thresholds are selected and for different thresholds there is a corresponding area under the curve. These areas are the PoD and PFA. Therefore, by selecting different thresholds, a series of PoDs and PFAs can be obtained and the corresponding ROC curves can be obtained. And for a particular event, the most desirable state of the model is (PoD,PFA)=(1,0). The closer to this point indicates that the model performs better. Meanwhile, when (PoD,PFA)=(0.5,0.5), it indicates that the results of the model present a random state. The probability of detection is equal to the probability of false alarm, which is not favorable to the model.

a)

b)

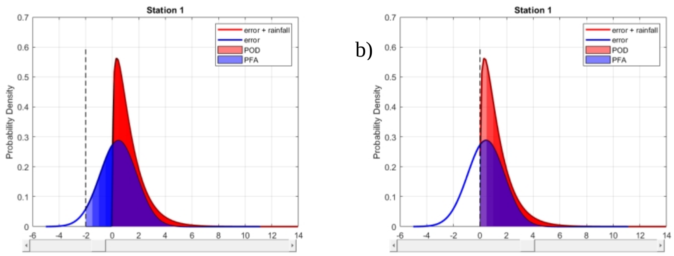

Fig. 2. (a) threshold = -2; (b) threshold = 0.

In this study, uncertainty was investigated by examining the rain gauge network consisting of ten rain gauge stations in the study area. The ten rain gauges were categorized into five groups for error calculation during the cross-validation phase. Then, the ROC images were plotted for five different storm events by taking the ten stations in the interval of thresholds -4 to 4. Finally, one of the stations was found to be underperforming and it is recommended that the monitoring capability should be strengthened here to provide a significant improvement in the detection capability of the rain gauge network.

Acknowledgements

Thanks to European Committee trough the European Regional Development Found for the support to AAFLOOD Interreg Atlantic Area Project (https://aafloods.eu/).

References

Boero, J., Schoefs, F., Capra, B., 2009. Effect on Risk Based Inspection of spatio-temporal dependence of ROC curves: Study of the corrosion of steel harbour structures.

Marco, Z., Elena, A., Anna, S., Silvia, T., Andrea, C., 2022. Spatio-temporal cross-validation to predict pluvial flood events in the Metropolitan City of Venice. J. Hydrol. 612, 128150. https://doi.org/10.1016/j.jhydrol.2022.128150.

Rouhan, A., Schoefs, F., 2003. Probabilistic modeling of inspection results for offshore structures. Struct. Saf. 25, 379-399. https://doi.org/10.1016/S0167-4730(03)00016-X.

Sadat-Noori, M., Ebrahimi, K., Liaghat, A., Hoorfar, A., 2013. Comparison of different geostatistical methods to estimate groundwater level at different climatic periods. Water Environ. J. 27. https://doi.org/10.1111/j.1747-6593.2012.00321.x

Schoefs, F., Clément, A., Nouy, A., 2009. Assessment of ROC curves for inspection of random fields. Struct. Saf. 31, 409-419. https://doi.org/10.1016/j.strusafe.2009.01.004.

Sun, X., Mein, R.G., Keenan, T.D., Elliott, J.F., 2000. Flood estimation using radar and raingauge data. J. Hydrol. 239, 4-18. https://doi.org/10.1016/S0022-1694(00)00350-4.

Xu, P., Wang, D., Singh, V.P., Wang, Y., Wu, J., Wang, L., Zou, X., Liu, J., Zou, Y., He, R., 2018. A kriging and entropy-based approach to raingauge network design. Environ. Res. 161, 61-75. https://doi.org/10.1016/j.envres.2017.10.038.