Drinking Water Engineering and Science Discussions

Drink. Water Eng. Sci. Discuss., author comment AC2
https://doi.org/10.5194/dwes-2021-7-AC2, 2021

Reply on RC2

Ahmad Ravanbakhsh et al.

Author comment on "Implementing and evaluating various machine learning models for pipe burst prediction" by Ahmad Ravanbakhsh et al., Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2021-7-AC2, 2021

1- L7 "accurate" should be "accurately"

Response:

Will be corrected in next revision

2- L25 "this" should be "these"

Response:

Will be corrected in next revision

3- L52. Is PBR determined of each individual pipe? Of subsets of pipes? Please explain. If it per pipe, then I would expect lots of 0 results (0 bursts / x length). How did you process that?

Response:

PBR is only calculated through failed pipes, not all pipes, so there are not many zeros, and we only have PBR for the number of pipes which has bursts.

For example, consider the length of a pipe made of asbestos cement is 131.18 meters and this pipe has had 1 failure during 8 years of study, so its PBR is equal to:

PBR= A/B

A= 1/8= 0.125 = Annual burst rate

B= Length to kilometers = 131.18/1000 = 0.13118 Km

So PBR=0.9528

4- L53. What is random sampling?

Response:

Suppose the total number of failure cases is 100, now 85 of them are provided to artificial intelligence for learning, and the remaining 15 are used to check the accuracy of the output given by the machine to predict the failure rate of the pipes. The choice of 85 out of 100 is completely random.

5- L240. What means inverse relation with length? Because 1/length is also in the PBR formula

Response:

Because in the PBR formula, the length of the pipe is the denominator. It is inversely related to the PBR, and because the length of the pipe is much more than the number of failure statistics, there is a large negative correlation between the length of the pipe and the PBR. According to Formula 1, the inverse relationship between length and PBR is clear, but the reason for mentioning it is the emphasis on the negative correlation between PBR and pipe length. Now, if this issue does not need to be mentioned in the opinion of the honorable referee, this sentence can be deleted in the next edition.

6- L251. "compare" must be "compares"

Response:

Will be corrected in next revision

7-1 L270. I miss the discussion. Why is RCNNSVR that much better than all other methods.

Response:

Because all models were measured with the evaluation criteria in the 2.2 section (Model performance assessment (L170)) and according to the results in Table 2, the RCNN-SVR method has the lowest error and the highest accuracy among the studied models.

For example, the RMSE criterion for PE pipes for MARS = 0.37 M5 '= 0.3 FCR = 0.38 LSSVR = 0.35 while for RCNN-SVR = 0.052 is obtained, which shows the high accuracy of this method for predicting of PBR. Figure 6 shows the high accommodation between the numbers obtained for PBR and the actual values taken from the real network of Jopar city in RCNN-SVR method.

7-2 What unique features make this method outperform all other methods by far.

Response:

The unique feature of this method is the high accuracy of PBR prediction compared to other methods.

7-3 Are the other methods that much simpler?

Response:

This paper does not examine the simplicity or complexity of the methods and only attempts to implement and evaluate the performance of a number of outstanding machine learning models in PBR prediction on a real water distribution network. The novelty of this paper is the use of RCNN-SVR method in water networks. Although this method has been used in other engineering sciences, but according to the research done by the authors, the use of this method in predicting PBR on a real network has not existed in any of the published articles. By examining this method and presenting its brilliant results, an efficient method has been introduced to the readers of this article.

7-4 And did you put equal effort in all methods in parameter tuning? Or was your goal to show that RCNNSVR is a preferred method?

Response:

Yes, attempts have been made to adjust the parameters in all models so that the best possible answer of each model was obtained. Each of the models was evaluated with the same evaluation criteria under the same conditions and it did not matter to the authors which method was better, but after the implementing and evaluating all models on real case study, it was found that the RCNN-SVR method was clearly better than the other methods.