Available online at www.dergipark.org.tr/en

e-ISSN: 2618-575X

INTERNATIONAL ADVANCED RESEARCHES and ENGINEERING JOURNAL

Journal homepage: www.dergipark.org.tr/en/pub/iarej

International Open Access

Volume 06
Issue 01

April, 2022

Research Article

A novel approach for prediction of daily streamflow discharge data using correlation based feature selection and random forest method

Levent Latifoğlu [a,*]

[a]*Erciyes University, Engineering Faculty, Civil Engineering Dept., Kayseri, Turkey*

ARTICLE INFO

Article history:
Received 25 August 2021
Accepted 22 February 2022
Published 15 April 2022

Keywords:
Correlation-based feature selection
Forecasting
Random Forest
Streamflow discharge

ABSTRACT

The accurate methods for the forecasting of hydrological characteristics are significantly important for water resource management and environmental aspects. In this study, a novel approach for daily streamflow discharge data forecasting is proposed. Streamflow discharge, temperature, and precipitation data were used for feature extraction which were systematically employed for forecasting studies. While the correlation-based feature selection (CFS) was used for feature selection, Random Forest (RF) model is employed for forecasting of following 7 days. Moreover, an accuracy comparison between the RF model and CFS-RF model is drawn by using streamflow discharge data. Acquired results confirmed the accuracy of CFS-RF model for both, middle and extended forecasting times compared to RF model which had similar accuracy values for the closer forecasting times. Moreover, the CFS-RF model proved to be much robust for extended forecasting durations.

1. Introduction

The world population has increased dramatically in the latter half of previous century. However, the scarcity of water is the most pressing issue that sustainable survival of human civilization is facing. Against this backdrop, conservation of water resources needs urgent attention. On the other hand, accurate estimation of water characteristics such as precipitation, flow, evaporation, runoff, land use, and basin characteristics is critical to manage water resources. Furthermore, these estimates have a significant role in minimizing the fallouts of natural disasters such as drought and floods. Above all, one of the most crucial parameter is the streamflow discharge [1]. Moreover, there is an essential role of streamflow data in dam project design, basin management, hydroelectric energy capacity determination, flood control projects etc. Irregularities in water flow may result in significant economic losses and permanent damage to the environment around the river. The factor causing variation in river flows include change in climate, the greenhouse gases emissions, and meteorological and hydrological features [1, 2].

Streamflow discharge processes are challenging to forecast due to their dynamic nature, complexity, non-stationarity, non-linearity. Therefore, the challenges involved in forecast accuracy of streamflow discharge made it an attractive area of research among hydrologists.

Traditionally, time series have been evaluated using models such the Autoregressive Integrated Moving Average (ARIMA) using linear approach which is a parametric approach [3,4]. However, due to the non-stationary and non-linear nature of streamflow discharge data, Artificial Intelligence (AI) techniques have been introduced [5-7].

Furthermore, while predicting streamflow discharge Artificial Neural Network (ANN) based models provided better forecast accuracy. In addition, for estimation of streamflow discharge, models such as generalized regression, radial basis neural networks along with meteorological data are available now [8, 9]. Likewise, in water resource management and hydrological prediction, support vector machines are widely used. [10]. In previous literature, Streamflow discharge data were analyzed as signals, and decomposed using Discrete Wavelet Transform (DWT), Singular Spectral Analysis (SSA), Empirical Mode Decomposition (EMD), and Fourier Transform. The previous values of streamflow discharge data and the sub-band components of the signals are estimated using ANN, Support Vector Machines, and other techniques [6, 7, 11]. In addition, forecasting

* Corresponding author. Tel.: +90 352 207 6666; Fax: +90 352 437 5784.
E-mail addresses: latifoglu@erciyes.edu.tr *(L. Latifoğlu)*
ORCID: 0000-0002-2837-3306 *(L. Latifoğlu)*
DOI: 10.35860/iarej.987245

building energy, streamflow discharge, day-ahead load, an hour ahead wind power and electrical short-term load, Random Forest (RF) is also used in literature [12-15].

The present study used an ensemble of CFS and RF as a novel approach to forecast daily streamflow discharge data. The main contributions of the present study are:

- It proposes a forecasting method for streamflow discharge data with better accuracy by using previous data (seven days before the first forecast day).
- Systematically extracted features from the daily streamflow discharge, precipitation, and temperature data were used using CFS-RF model for training and validation of the proposed method.

2. Materials and Methods

The input data used in this study is based on streamflow discharge (Q), precipitation (P) and temperature (T_{max}, T_{min}) values. The streamflow discharge data (16527 datapoints) was recorded over a period of 46 years from Kootenay River near Skookumchuck basin in British Columbia, Canada. The exact coordinates of the drainage basin location at which data was collected are reported as 49°54'38" N, 115°44'08" W (latitude: 49.91056061, longitude: 115.7355576) as shown in Figure 1. The data was obtained from CANOPEX database [16, 17] from a drainage area of 7196.93 km^2.

Figure 2 shows how the above-mentioned input data was used for the forecast model by utilizing seven days previous data for predicting the discharge data for the following seven days. While the previous seven days are denoted as one-to-seven previous data, the following seven days are mentioned as one-to-seven ahead data. One-to-seven data (denoted as t-n while n takes the value of 1 to 7) for each input (i.e., Q, P, T_{max}, T_{min}, a total of 28 dataset values). Correlation-based method (CFS) was used for selecting the features from the above mentioned 28 datasets for forecasting of one-to-seven ahead data. Also, Figure 2 shows the overall methodology as well as data constituents used for the development of CFS-RF model.

This model is based on 3 steps. The first step is distinguishing between the training and testing data. The second step is feature extraction from the input datasets of 28 values as shown in Table 1. The third step comprises of the application of selected features using RF model and this way a forecasting model was achieved. Later the accuracy of the trained forecasting model was evaluated using the testing data.

More specifically, 70% of the datapoints belonging to all input datasets (P, Q, T_{max}, T_{min}) were used for training the forecasting model. The remaining 30% of the data was used for checking the accuracy of the model.

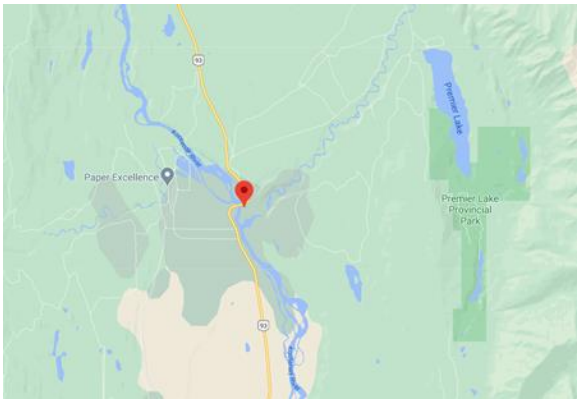


Figure 1. Map of Kootenay River Near Skookumchuck

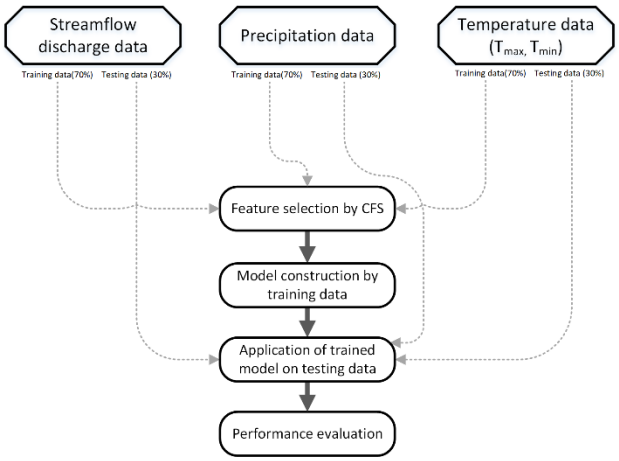
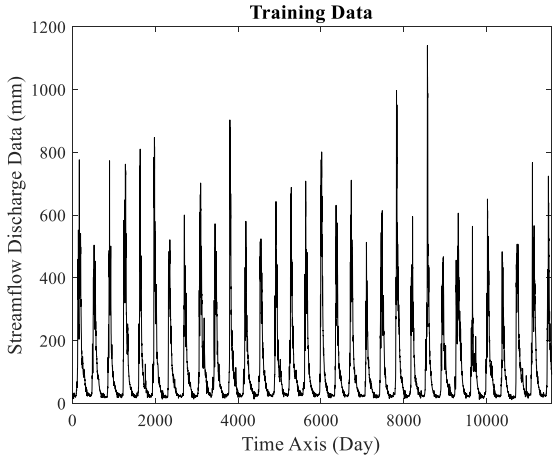


Figure 2. The proposed CFS-RF forecasting model for daily streamflow discharge data

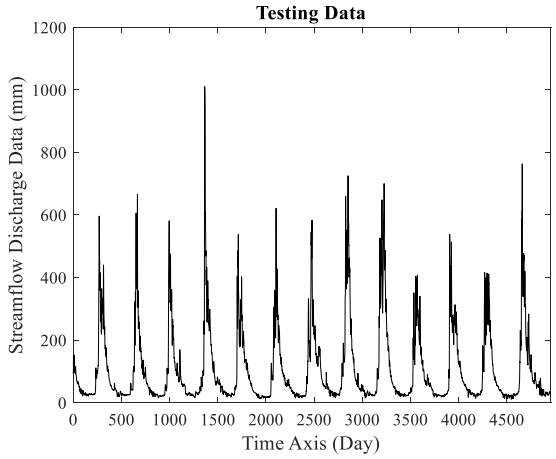
Table 1. Selected features for forecasting of Q(t)

Total features	
Tmax(t-1), Tmax(t-2), Tmax(t-3), Tmax(t-4), Tmax(t-5), Tmax(t-6), Tmax(t-7)	
Tmin(t-1), Tmin(t-2), Tmin(t-3), Tmin(t-4), Tmin(t-5), Tmin(t-6), Tmin(t-7)	
P(t-1), P(t-2), P(t-3), P(t-4), P(t-5), P(t-6), P(t-7)	
Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5), Q(t-6), Q(t-7),	
Selected Features for One Ahead Forecast	**Output**
Tmax(t-2), Tmin(t-5), Tmin(t-7), P(t-1), Q(t-1)	Q(t)
Selected Features for Two Ahead Forecast	**Output**
Tmax(t-2), Tmin(t-1), P(t-1), P(t-6), Q(t-1)	Q(t+1)
Selected Features for Three Ahead Forecast	**Output**
Tmax(t-3), Tmin(t-1), Tmin(t-1),P(t-1), P(t-2), P(t-7), Q(t-1)	Q(t+2)
Selected Features for Four Ahead Forecast	**Output**
Tmax(t-1), P(t-5), P(t-6), P(t-7), Q(t-1)	Q(t+3)
Selected Features for Five Ahead Forecast	**Output**
Tmax(t-1), P(t-4), P(t-5), P(t-6), P(t-7), Q(t-1)	Q(t+4)
Selected Features for Six Ahead Forecast	**Output**
Tmax(t-1), P(t-4), P(t-5), P(t-6), P(t-7), Q(t-1)	Q(t+5)
Selected Features for Seven Ahead Forecast	**Output**
Tmax(t-1), P(t-4), P(t-5), P(t-6), P(t-7), Q(t-1)	Q(t+6)

Training and testing streamflow discharge data used in this study are shown in Figure 3. Correlation coefficient for all datasets until t (the forecast day) to t-7 (seven days before the forecast day) is shown in Table 2.



a)



b)

Figure 3. Daily streamflow data measured on Kootenay River Near Skookumchuck basin a) Training Data, b) Testing Data

Table 2. Correlation Coefficient until (t) to (t-7)

Training Correlation Coefficient (t) to (t-7)	Testing Correlation Coefficient (t) to (t-7)
Streamflow Discharge Data	
0.9887, 0.9656, 0.9423, 0.9219, 0.9043, 0.8894	0.9877, 0.9628, 0.9372, 0.9139, 0.8931, 0.8739
Precipitation Data	
0.3259, 0.0897, 0.0599, 0.0513, 0.0454, 0.0473	0.3171, 0.0679, 0.0477, 0.0369, 0.0484, 0.0775
Maximum Temperature Data	
0.9623, 0.9184, 0.8863, 0.8645, 0.8487, 0.8365	0.9586, 0.9111, 0.8766, 0.8528, 0.8358, 0.8243
Minimum Temperature Data	
0.9325, 0.8678, 0.8250, 0.7956, 0.7745, 0.7595	0.9368, 0.8691, 0.8232, 0.7893, 0.7613, 0.7378

2.2 Feature Selection

The feature selection is a machine learning preprocessing stage that reduces the dimensionality of the data, removes irrelevant data, improves learning accuracy, and result comprehensibility. CFS method is a peculiar approach used for regression of datasets by evaluating the classification capabilities of the inherent features. This model prefers non-contradicting features from the datasets by examining its relationship with the expected classification criteria. The CFS model uses the entropy based information theory. The definition of entropy is shown in Equation (1).

$$H(x) = -\sum_{x \in X} P(x_i) log2(P(x_i)) \quad (1)$$

The entropy of variable x is calculated using conditional probability, as shown in Equation 2 based on the input values of y.

$$H(x/y) = -\sum_{x \in X} P(y_j) \sum_{y \in Y} P(x_i/y_j)\, log2(P(x_i/y_j)) \quad (2)$$

Where $P(x_i)$, $P(x_i/y_j)$ are the prior probability for all x values and the posterior probability for x and y values, respectively. Mutual information is defined as the amount by which the entropy of x as a result of additional information about x given by y shown in Equation 3.

$$gain(x/y) = H(x) - H(x/y) \quad (3)$$

If the gain(x/y) > gain(z/y), we can infer that feature y is more associated with feature x than to feature z.

Equation 4 shows the symmetrical uncertainty coefficient (SUC), an important metric which shows the relationship between the features

$$SUC = 2x(\frac{gain}{H(y) + H(x)}) \quad (4)$$

The SUC shows its tendency towards the relationship and has a normalized value with the range [0,1]; 1 denotes that one's knowledge completely predicts the value of the other, whereas 0 denotes that x and y are unrelated. It symmetrically handles a pair of attributes [18].

2.3 Random Forest Algorithm

The random forest (RF) is an ensemble approach that combines the predictions of numerous decision trees into a single forecast and can be used for both, regression and classification problems. Leo Breiman [19] invented the RF technique in 2001. The main principle is bagging, which involves randomly selecting a sample of size m from the training set and fitting it to a regression tree. This is known as a bootstrap sample, and it is picked using replacement, meaning that the same observations may appear many times [20]. The RF algorithm is applied as follows:

- With the Bootstrap method, n size data set is selected. This data set is split into two sections: training data and test data.
- The largest decision tree (CART) is generated using the training dataset, and this decision tree is not pruned. When dividing each node in this tree, m out of a total of p estimator variables are chosen randomly. The condition m<p must be satisfied in this situation because it is undesirable to see the tree grow too fast and adapt too soon. The highest knowledge gain among the m estimators chosen is used for branching. The value of this variable is decided by the Gini index. This process is repeated until there are no more branches to be created for each node.
- Each leaf node is assigned a class. The test data set is then at the top of the tree, and each observation in this data set is assigned to a class.
- All stages from 1st to 3rd step are repeated N times.
- The tree is evaluated using observations that were not used during the development process. The repeat number of observations is used to classify the data.
- A majority of votes is used to decide class assignments for each observation, tree set.

Random forest parameters in Table 3 were established by trial and error throughout the model's creation, taking into consideration calculation time and predicting performance.

2.4 Performance Evaluation

In this study, The Root Mean Square Error (RMSE), The Mean Absolute Error (MAE), The Correlation Coefficient (R) and The Determination Coefficient (R^2) were used to show the performance of the proposed method [21].

The differences between observed time series data and forecasted data by the proposed model are measured by average absolute error. MAE is described as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|X_{observed,i} - X_{estimated,i}| \quad (5)$$

RMSE is calculated by root of squared the average difference across the time series data. RMSE is denoted by the following equation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{X_{observed,i}} - \mathrm{X_{estimated,i}})^2} \quad (6)$$

The R value shows the magnitude, direction, and significance of the relationship between measured and forecasted time series data. The R represents the correlation coefficient, which has a value between [-1, 1]. The R value is determined as shown in Equation 7:

Table 3. Random Forest Parameters used in the forecasting study

	Trials	**The best result for forecasting**
Number of iterations	100, 200, 300, 400	300
Number of attributes to randomly investigate	0, 1, 2	0
Number of folds for backfitting	0 ,1	0
Size of each bag, as a percentage of the training set size	70, 80, 100	100
Seed for random number generator	Yes, No	Yes
The desired batch size for batch prediction	70, 80, 100	100

$$\mathrm{R} = \frac{1}{\mathrm{N}-1}\sum_{\mathrm{i=1}}^{\mathrm{N}}\left(\frac{\mathrm{X_{observed,i}} - \mu_{\mathrm{X}}}{\sigma_{\mathrm{X}}}\right)\left(\frac{\mathrm{X_{estimated,i}} - \mu_{\mathrm{Xe}}}{\sigma_{\mathrm{Xe}}}\right) \quad (7)$$

While $\boldsymbol{X_{observed,i}}$ shows the measured data, $\boldsymbol{\mu_X}$ is the average, $\boldsymbol{X_{estimated,i}}$ is predicted data and $\boldsymbol{\sigma_X}$ is the standard deviation of the measured data, $\boldsymbol{\mu_{Xe}}$ shows the average of the predicted data is and the standard deviation $\boldsymbol{\sigma_{Xe}}$.

The R^2 coefficient is widely used to measure the predictability of hydrological models and is described as shown in Equation 8. This statistical criterion takes the value between -∞ and 1. The closer the R^2 value is to 1, higher is the forecast accuracy. [22]. The R^2 value is calculated using the following equation:

$$\mathrm{R^2} = 1 - \frac{\sum_{\mathrm{i=1}}^{\mathrm{N}}[\mathrm{X_{observed,i}} - \mathrm{X_{estimated,i}}]^2}{\sum_{\mathrm{i=1}}^{\mathrm{N}}[\mathrm{X_{observed,i}} - \mu_{\mathrm{X}}]^2} \quad (8)$$

Weka and MATLAB software packages are used to perform all the required calculations.

3. Results and Discussion

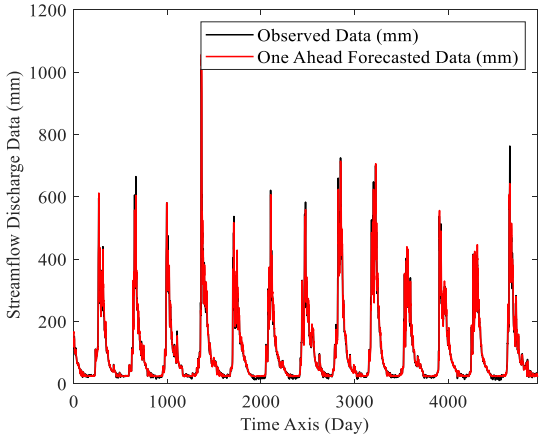
The results of the feature selection are given in Table 1. Figures 4, 5, and 6 respectively show the graphical representation of the trained model for the forecasting of t, t+1, t+2 forecasts for streamflow discharge data, out of total t+7 forecasts.

Table 4 on the other hand shows the numerical values of the total forecast from t to t+6.

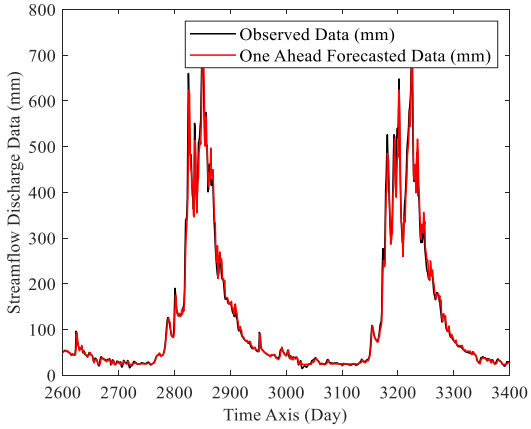
Q(t-1) feature was applied to obtain Q(t) value for one ahead forecasting, Q(t-1) feature was applied to obtain Q(t+1) value for two ahead forecasting, Similarly, Q(t-1), data was applied as input features for three (Q(t+2)), four (Q(t+3)), five (Q(t+4)), six (Q(t+5)) and seven (Q(t+6)) ahead forecasting. Obtained performance parameters using Q(t-1) feature and RF model are given in Table 5.

Table 4. 1-7Ahead Forecasting performance of CFS-RF Model

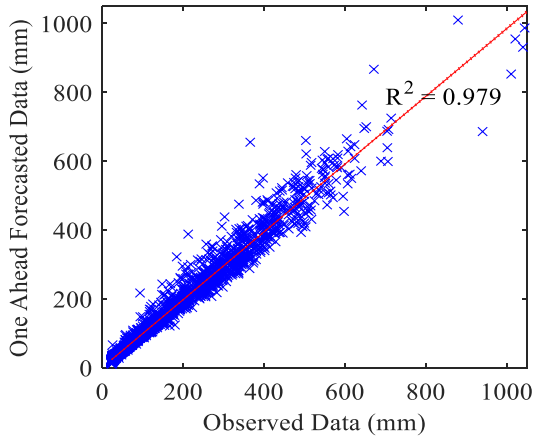
	RMSE	MAE	R	R^2
One Ahead	17.9864	8.1520	0.9895	0.9791
Two Ahead	30.1717	13.3113	0.9703	0.9414
Three Ahead	41.2998	19.0428	0.9442	0.8914
Four Ahead	49.0520	23.0490	0.9212	0.8486
Five Ahead	54.5374	26.2710	0.9021	0.8138
Six Ahead	59.0526	28.7231	0.8852	0.7837
Seven Ahead	63.1710	30.7558	0.8685	0.7542



(a)



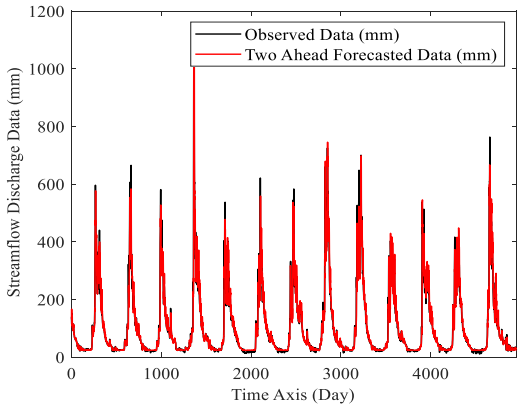
(b)



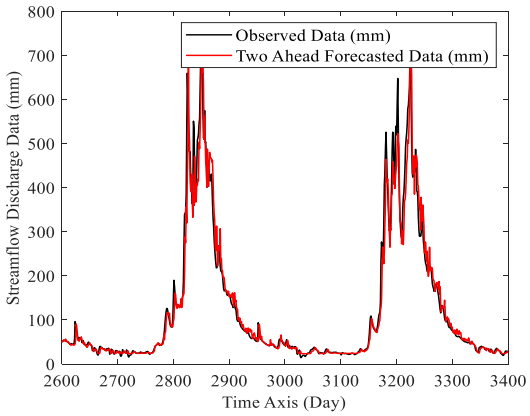
(c)

Figure 4. a) One-ahead forecasting of daily streamflow discharge data using CFS-RF Model, b) Zoomed graphic, c) Scatter plot

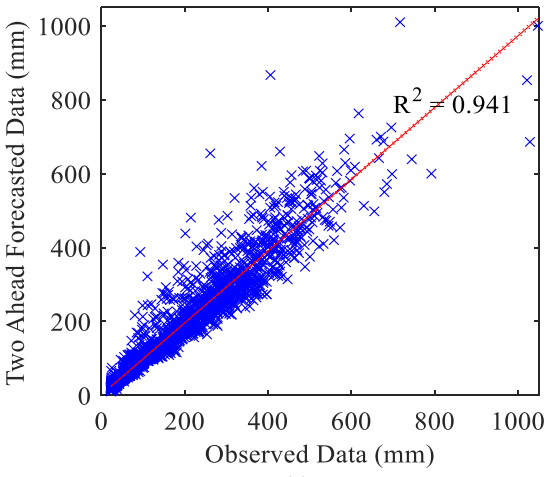
In the area of water resources planning and management, efficient water resource utilization demands accurate and successful streamflow discharge data forecasting [23]. A forecasting framework was established in this study to investigate the efficacy of the CFS-RF model with a novel approach. An estimation of daily streamflow discharge data for the upcoming seven days was performed using the CFS-RF model. Daily streamflow discharge data was divided into training and testing data. A comparative analysis was drawn between the forecasting capabilities of single RF model and the CFS-RF model combined.



(a)



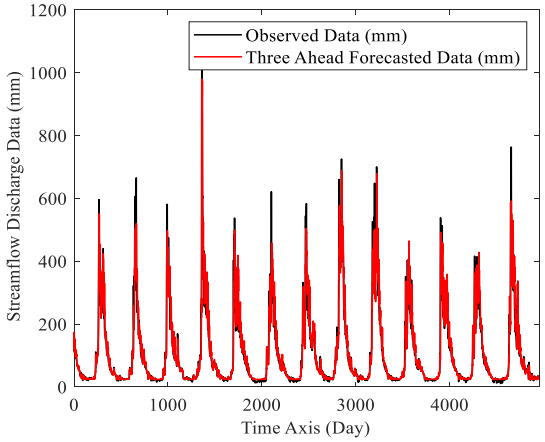
(b)



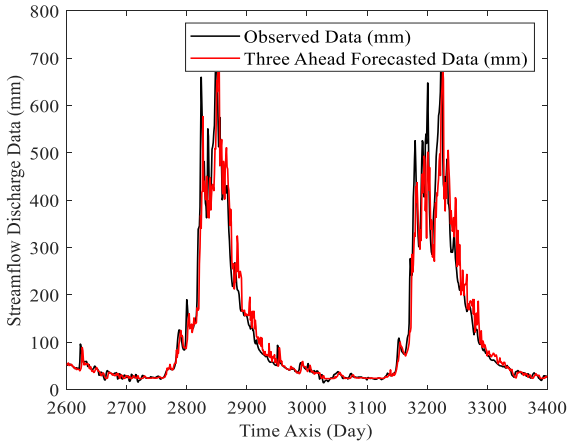
(c)

Figure 5. a) Two-ahead forecasting of daily streamflow discharge data using CFS-RF Model, b) Zoomed graphic, c) Scatter plot

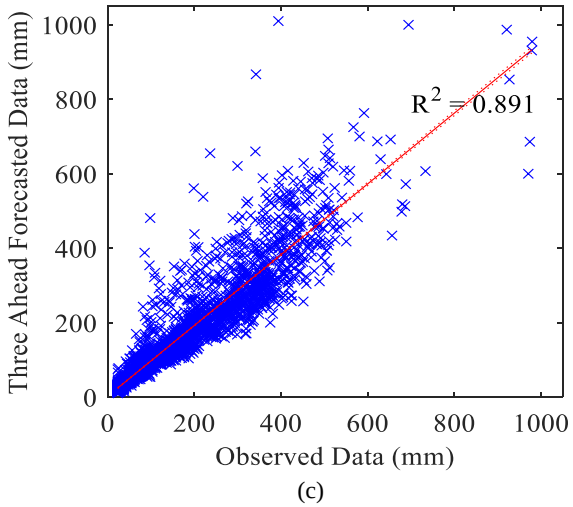
It can be seen from the Table 4 that the R^2 value for the t to t+4 is above 0.80 compared to much later forecast days such as t+5 and t+6 for which the R^2 value is slightly above 0.75. Moreover, the value of R^2 for much earlier forecasts (t and t+1) are above 0.94 as shown in Figure 4 and 5 as well. As significantly linear relation between the forecasted and actual (real) values can be seen (see Figure 4 and 5).



(a)



(b)



(c)

Figure 6. a) Three-ahead forecasting of daily streamflow discharge data using CFS-RF Model, b) Zoomed graphic, c) Scatter plot

Table 5. 1-7 Ahead Forecasting performance of RF Model

	RMSE	MAE	R	R^2
One Ahead	20.7351	9.3833	0.9859	0.9720
Two Ahead	35.5437	15.7612	0.9584	0.9184
Three Ahead	45.6941	20.5596	0.9307	0.8661
Four Ahead	77.7896	38.8902	0.7922	0.6276
Five Ahead	81.1282	41.0384	0.7725	0.5968
Six Ahead	84.3759	43.1028	0.7535	0.5678
Seven Ahead	87.0211	44.9170	0.7366	0.5426

The forecast study results acquired by CFS-RF model using only streamflow discharge data are given in in Table 5 in order to compare the reliability of CFS-RF and only RF model. It can be clearly seen, after inspecting the numerical values of the Table 4 and 5, that although CFS-RF model given much accurate results for mediocre (t+3 and t+4 forecasting) and extended (t+5, t+6) forecasting, it has low accuracy values for much recent forecasting. Similar scenario is valid for the RF model for closer forecasting time; however, its accuracy is slightly lower than the CFS-RF model.

4. Conclusion

Using temperature and precipitation data instead of only streamflow discharge data increases the forecast performance. Also, the selection of input features plays an important role for model performance and the accuracy of the results. In this study, it is recommended that temperature data, precipitation data in addition to streamflow discharge data should be used to obtain the best input variable combination for forecasting of streamflow discharge data.

Based on the findings, it is established that the RF model combined with CFS model shows an inherent superior capability of streamflow discharge forecasting for the river of Kootenay, Canada.

The proposed model appears to be an important tool that can be used in forecasting studies.

Declaration

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

L. Latifoğlu developed the methodology, performed the analysis and wrote the whole article.

References

1. Sharma, P. and D. Machiwal, *Advances in streamflow forecasting: from traditional to modern approaches*. 2021, USA: Elsevier, Inc.
2. Peters, R.L., *The greenhouse effect and nature reserves*. Bioscience, 1985. **35**(11): p.707-717.
3. Rojas, I., O. Valenzuela, F. Roja, A. Guillén, L.J. Herrera, H. Pomares, L. Marquez, and M. Pasadas, *Soft-computing techniques and ARMA model for time series prediction*. Neurocomputing, 2008. **71**(4-6): p. 519-537.
4. Khandelwal, I., R. Adhikari, and G. Verma, *Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition*. Procedia Computer Science, 2015. **48**: p. 173-179.
5. Yaseen, Z. M., A. El-Shafie, O. Jaafar, H.A. Afan, and K.N. Sayl., *Artificial intelligence based models for stream-flow forecasting: 2000–2015*. Journal of Hydrology, 2015. **530**: p. 829-844.
6. Kisi, O., L. Latifoğlu, and F. Latifoğlu, *Investigation of empirical mode decomposition in forecasting of hydrological time series*. Water Resources Management, 2014. **28**(12): p. 4045-4057.
7. Latifoğlu, L., O. Kişi, and F. Latifoğlu, *Importance of hybrid models for forecasting of hydrological variable*. Neural Computing and Applications, 2015. **26**(7): p. 1669-1680.
8. Meshram, S.G., C. Meshram, C.A.G. Santos, B. Benzougagh, and K.M. Khedher, *Streamflow prediction based on artificial intelligence techniques.* Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2021. p. 1-11.
9. Nourani, V., N.J. Paknezhad, and H. Tanaka, *Prediction interval estimation methods for artificial neural network (ANN)-based modeling of the hydro-climatic processes, a Review*. Sustainability, 2021. **13**(4): p. 1633.
10. Adnan, R. M., X. Yuan, O. Kisi, and Y. Yuan, *Streamflow forecasting using artificial neural network and support vector machine models*. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 2017. **29**(1): p. 286-294.
11. Saraiva, S. V., F. de Oliveira Carvalho, C.A.G. Santos, L.C. Barreto, and P.K.D.M.M Freire, *Daily streamflow forecasting in Sobradinho Reservoir using machine learning models coupled with wavelet transform and bootstrapping*. Applied Soft Computing, 2021. **102**: p.107081.
12. Pham, L. T., L. Luo, and A. Finley, *Evaluation of random forests for short-term daily streamflow forecasting in rainfall-and snowmelt-driven watersheds*. Hydrology and Earth System Sciences, 2021. **25**(6): p. 2997-3015.
13. Li, X., J. Sha, and Z.L. Wang, *Comparison of daily streamflow discharge forecasts using extreme learning machines and the random forest method.* Hydrological Sciences Journal, 2019. **64**(15): p. 1857-1866.
14. Lahouar A. and J.B.H. Slama, *Day-ahead load forecast using random forest and expert input selection*. Energy Conversion and Management, 2015. **103**: p. 1040-1051.
15. Huo, J., T. Shi and J. Chang., *Comparison of random forest and SVM for electrical short-term load forecast with different data sources*, in 7th IEEE International conference on software engineering and service science (ICSESS), 2016, Beijing: China. p. 1077-1080.
16. Canopex hydrometeorological watershed database. [cited 2020 1 December]; Available from: http://canopex.etsmtl.net/
17. Arsenault, R., R. Bazile, C. Dallaire-Ouellet, and F. Brissette, *CANOPEX: A Canadian hydrometeorological watershed database*. Hydrological Processes, 2016. **30**(15): p. 2734-2736.
18. Gopika, N. and A. Kowshalaya M.E, *Correlation based feature selection algorithm for machine learning*, in 3rd International Conference on Communication and Electronics Systems (ICCES), 2018, Coimbatore: India. p. 692-695.
19. Breiman L., *Random forests*. Machine Learning, 2001, **45**: p. 5–32.
20. Liu Y., Y. Wang, and J. Zhang, *New machine learning algorithm: Random forest*, in International Conference on Information Computing and Applications, 2012, Chengde: China. p. 246-252.
21. Samanataray, S., and A. Sahoo, *A Comparative study on prediction of monthly streamflow using hybrid ANFIS-PSO approaches*. KSCE Journal of Civil Engineering, 2021. **25**(10): p. 4032-4043.
22. Ali, M.H. and I. Abustan, *A new novel index for evaluating model performance*. Journal of Natural Resources and Development, 2014. **4**: p. 1-9.
23. Kumbur, H., V. Yamaçlı, and A. Küçükbahar, *Mersin province water projections and water information and management system: Erdemli district model*. International Advanced Researches and Engineering Journal, 2018. **2**(3): p. 261-266.