Journal of Hydrology 613 (2022) 128265

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Day-ahead inflow forecasting using causal empirical decomposition

Mojtaba Yousefi [a,*], Xiaomei Cheng [b], Michele Gazzea [a], August Hubert Wierling [a], Jayaprakash Rajasekharan [c], Arild Helseth [d], Hossein Farahmand [c], Reza Arghandeh [a]

[a] Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Science, Bergen, 5063, Norway
[b] Smart Innovation Norway, Halden, 1783, Norway
[c] Department of Electric Power Engineering, Norwegian University of Science and Technology, Trondheim, 7034, Norway
[d] Energy systems, SINTEF, Trondheim, 7034, Norway

ARTICLE INFO

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Fi-John Chang, Associate Editor.

Keywords:
Inflow forecasting
Hydropower scheduling
Time-series
Causal model
Multivariate decomposition
Machine learning
Information theory

ABSTRACT

It is essential to have accurate and reliable daily-inflow forecasting to improve short-term hydropower scheduling. This paper proposes a Causal multivariate Empirical mode Decomposition (CED) framework as a complementary pre-processing step for a day-ahead inflow forecasting problem. The idea behind CED is combining physics-based causal inference with signal processing-based decomposition to get the most relevant features among multiple time-series to the inflow values. The CED framework is validated for two areas in Norway with different meteorological and hydrological conditions. The validation results show that using CED as a pre-processing step significantly enhances (up to 70%) the forecasting accuracy for various state-of-the-art forecasting methods.

1. Introduction

Inflow forecasting improves operation and planning of hydropower stations while reducing the risk of flooding and reservoir rationing (Apaydin et al., 2020; Liao et al., 2020). Inflow forecasting approaches are categorized into two major groups: (1) physical models and (2) data-driven models. Physics-based methods for inflow forecasting partly address nonlinearity and non-stationary issues of inflow data using physical laws and catchment characteristics (Liao et al., 2020). However, the main challenge of such methods is their dependencies on initial conditions and input data (Bennett et al., 2016).

Data-driven methods for inflow forecasting include statistical and machine learning techniques such as linear regression (LR) (Kao et al., 2015), fuzzy inference systems (Lohani et al., 2014), spatial distribution-based model (Tsai et al., 2014), model tree (Jothiprakash and Kote, 2011), Multilayer perception (MLP) (Golob et al., 1998; Coulibaly et al., 2001; Cheng et al., 2015; Abdellatif et al., 2015), and support vector regression (SVR) (Luo et al., 2019; Moazenzadeh et al., 2018; Tongal and Booij, 2018). In Bordin et al. (2020) a comprehensive review paper is provided in which all machine learning techniques used for inflow forecasting are listed. Recently, deep learning methods such as long short term memory (LSTM) (Apaydin et al., 2020; Kao et al., 2020; Herbert et al., 2021; Cheng et al., 2021), recurrent neural network (RNN) (Ni et al., 2020), sequential to sequential (seq2seq) network (Han et al., 2021; Kao et al., 2021; Yin et al., 2021) and deep neural networks (DNNs) have been used for inflow forecasting problems due to their capability to capture the nonlinearities and long-term temporal dependencies learning (Yousefi et al., 2020). However, machine learning-based models are facing some challenges. For instance, their performance is biased by training data and may suffer from over-fitting, information saturation and under-fitting issues (Bai et al., 2016). This is more problematic, in the context of multivariate inflow forecasting problems, where these techniques are unable to improve the inflow accurately when used as standalone models without any preprocessing techniques (Apaydin and Sibtain, 2021).

Feature extraction and selection as pre-processing is key for any machine learning-based forecasting model to address the issues mentioned above. Some prominent pre-processing methods for time-series analysis are inspired by the signal processing tool sets such as Fast Fourier transform (FFT), wavelet transform (WT), Empirical Mode Decomposition (EMD), Multivariate Empirical Mode Decomposition(MEMD), etc. For example, WT has been used in hydrological studies for flood forecasting (Sehgal et al., 2014a) and river discharge forecasting (Sehgal et al., 2014b). However, such methods are time consuming and require extensive computation power (Roushangar and Alizadeh, 2018; Apaydin and

* Corresponding author.
E-mail address: Myou@hvl.no (M. Yousefi).

https://doi.org/10.1016/j.jhydrol.2022.128265
Received 22 November 2021; Received in revised form 5 May 2022; Accepted 24 July 2022
Available online 6 August 2022

Sibtain, 2021). In contrast to WT, EMD techniques work in time domain without the needing any preset basis functions or a mother wavelet. EMD can reduce the complexity of data set by breaking down the time-series (e.g. inflow historical values) into different sub-elements (Bai et al., 2016; Qi et al., 2019; Okkan and Serbes, 2013; Apaydin and Sibtain, 2021; Bai et al., 2015). For example, Apaydin and Sibtain (2021) proposes a multivariate streamflow forecasting model for inflow based on EMD. However, the authors suggested an ad-hoc process rather than an algorithmic approach to combine meteorological and hydrological data to forecast inflow. In multivariate inflow forecasting, decomposing original time-series to its intrinsic mode functions (IMFs), which carry different frequencies of the original series, create a large set of features. Using large-set of features without doing any feature selection techniques can introduce the risk of over-fitting and information saturation in the machine learning process. Therefore, feature selection techniques are usually used after decomposition techniques (Apaydin and Sibtain, 2021; Bai et al., 2016). In this paper, we propose a causal inference-based framework as a feature selection technique to find the most relevant features for inflow forecasting. Causal inference is an emerging concept within the machine learning community that characterizes the cause-and-effect relationship between distinct input and target variables based on directed acyclic graphs (Sriram et al., 2018).

This paper proposes a novel multivariate time-series decomposition framework powered by causal inference for short-term inflow forecasting application. The proposed framework is referred to as Causal Empirical Decomposition (CED). CED is a feature extraction and feature selection framework or a pre-processing step that can be used for any forecasting problem with multiple input time-series. Utilizing causal inference for time-series feature extraction enables better exploitation of the inherent underlying dynamic behavior of a hydrological system. To the best of the authors' knowledge, CED is the first attempt at causal-based feature selection in inflow forecasting for hydropower applications. The contribution of this paper are summarized as follows:

- In terms of **methodology**, a pre-processing feature extracting and selection framework powered by decomposition and causality analysis is proposed that is applicable to any time-series forecasting method.
- From an **application** point of view, the accuracy of short-term inflow forecasting is considerably improved by using the proposed CED pre-processing framework based on actual data from two different use cases in Norway.

2. Use case

All data used in this paper are collected from two main sources: Norway's regulatory agency (NVE) and the Norwegian Center for Climate Services (Anon, 0000a,b). Two study areas in this paper as presented in Fig. 1 are located in Vestlandet (Lærdal municipality) and Østlandet (Hemsedal Municipality) regions in Norway. The measurement stations and hydropower stations are presented with red and cyan circles, respectively. The first case study contains the Borgund power station, weather station (Filefjell–Kyrkjestølane), and inflow measurement stations (Sula with the elevation of 1200 (m)). Likewise, the second use case contains Gjuva hydropower station, inflow measurement station (Storeskar with the elevation of 850 (m)), and weather stations (Hemsedal ski center and Memesedal II). Borgund hydropower plant was commissioned in 1974 with an installed capacity of 212 (MW) and annual production of 985 (GWh). Gjuva hydropower plant located in Hemsedal municipality has been in operation since 1957 with a maximum performance of 10 (MW). Historical data from Sula and Storeskar inflow measurement stations are available from 1958. Other data sources, including meteorological and hydrological data, are available from 2010 and 2013 for first and second use cases, respectively.

The data set for both locations includes inflow (m^3/s), average temperature (°C), maximum temperature (°C), minimum temperature (°C), precipitation (mm), snow-depth (cm), and relative humidity (%). The historical inflow measurements for two use cases are presented in Fig. 2. The inflow peak occurs at the end of spring or the beginning of summer when the temperature rises and causes snow to melt in these geographical locations. There is also another peak at the beginning of fall due to increased precipitation. Furthermore, the inflow variation in the second use case is significantly more notable than Sula as its maximum inflow is 52 (m^3/s) while the maximum inflow for the Sula is 16.72 (m^3/s).

3. Methodology

The overall architecture of the proposed CED framework, together with the forecasting block, is depicted in Fig. 3. The CED framework is comprised of four major steps: decomposition, information transfer modeling, causal significant test, and forecasting algorithms.

In the decomposition block, the original time-series is breaking-down into three new features, i.e., stochastic, periodic and trend terms. In this paper, we have 7 time-series with a length of N=3379 samples applied to the decomposition block. The output is 21 new time-series with the same length as the original time-series. The output of the decomposition step is sent to the second step, which is a feature selection algorithm based on information theories. In this step, a greedy search algorithm is used to find the features' lag values with the maximum contribution to the current state of inflow. The objective function of the greedy search algorithm is to maximize the conditional mutual information and the transfer entropy. The obtained causal candidate set is sent to the third step, which is a pruning algorithm. This step guarantees that the minimal subset of the causal candidate set is selected (redundant variables are removed, and only synergies variables are kept). A causal network is obtained for each feature of inflow (stochastic term, periodic term, and trend term). These causal networks are used to forecast each feature of inflow. Once all three terms of inflow features are forecasted, they are bundled to form the forecasted inflow values. In the following subsections, each step is briefly explained.

3.1. Step 1: Time-series decomposition

This block utilizes a Multivariate Empirical Mode Decomposition (MEMD) method to split the original time-series into intrinsic mode functions (IMFs) while preserving all the original time-series information. MEMD is a self-adaptive technique designed for nonlinear and non-stationary data (Huang et al., 2020; Rehman and Mandic, 2010).

MEMD decomposes a multivariate time-series $X_{M\times N}$ including inflow, precipitation, snow-depth, relative humidity, maximum temperature, minimum temperature, and average temperature, where M is the number of variables, N is the length of series, and $X_{M\times N}$ is structured as follows:

$$X_{M\times N} = [X_1, X_2, \dots, X_M]^{\top} = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,N} \\ x_{2,1} & x_{2,2} & \dots & x_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{M,1} & x_{M,2} & \dots & x_{M,N} \end{bmatrix} \quad (1)$$

For the sake of better understanding, the MEMD decomposition is expressed for only one variable of $X_{M\times N}$ as follows:

$$X_1 = \sum_{j=0}^{m-1} (IMF_j^1 + R_m^1) = \sum_{j=0}^{m-1} \sum_{i=0}^{N} (imf_{j,i}^1 + r_{m,i}^1) \quad (2)$$

where $IMF^1 \in R^{m\times N}$ (the residual R_m is considered as the last IMF) is a two-dimension sub-matrix of a three-dimension IMF matrix, $IMF =$

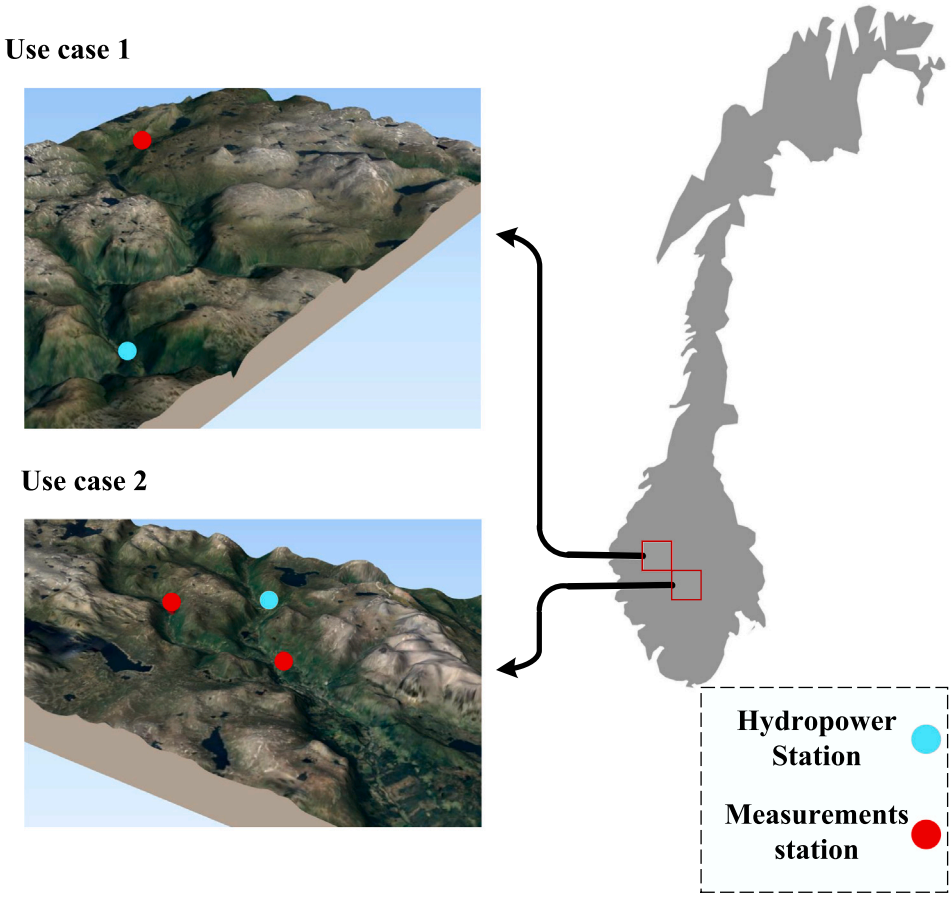

Fig. 1. Overview of the study area; measurement and hydropower stations are presented by red and cyan circles in both use cases.

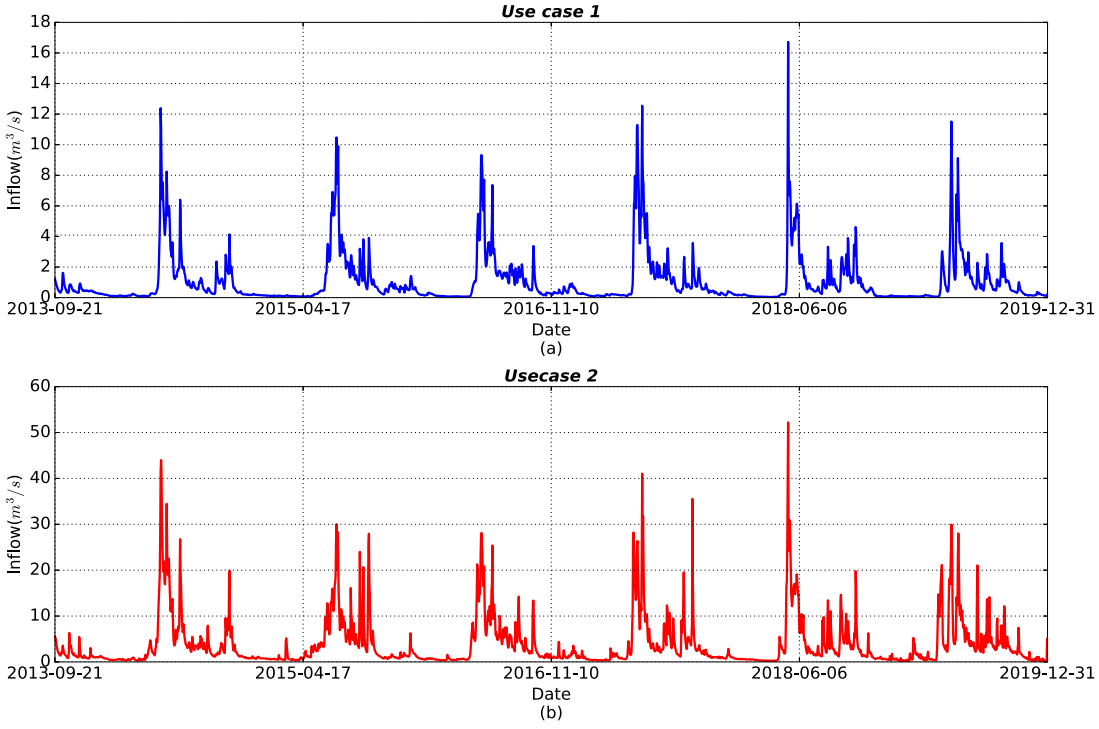

Fig. 2. Historical inflow measurement from 2011 to 2013; (a) inflow data collected from Sula use case 1; (b) inflow data collected from Storeskar.

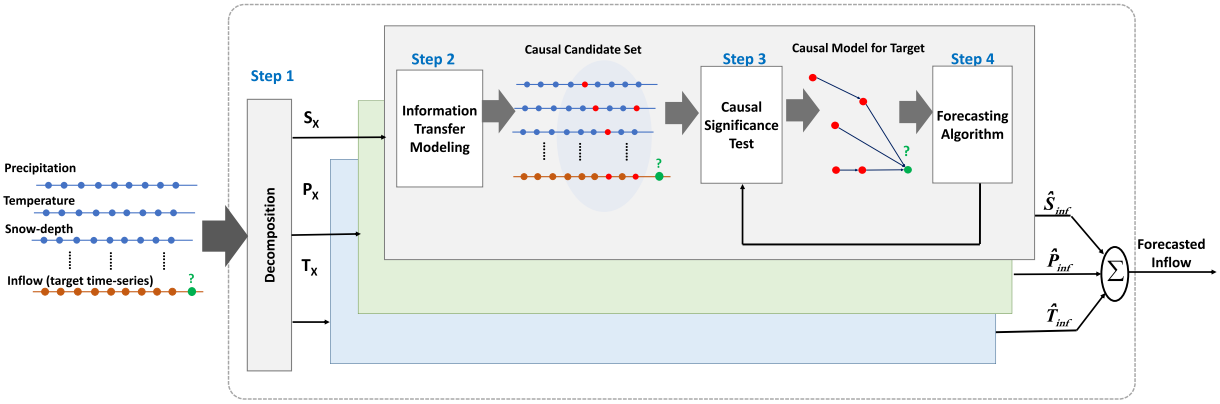

Fig. 3. Overview of the proposed CED framework; S_x, P_x, and T_x stand for stochastic, periodic, and trend terms of the original time-series and $\hat{S}_{inf}$, $\hat{P}_{inf}$, and $\hat{T}_{inf}$.

$[IMF^1, IMF^2, \ldots, IMF^M]^\top$, $IMF \in R^{M\times m\times N}$. Therefore, each variable (e.g. IMF^1) decomposed to m sub-series $IMF_1^1, IMF_2^1, \ldots, IMF_m^1$ with the length of N (same length with the original inflow).

After the decomposition of each variable, FFT analysis is applied on each of the IMFs to sort them by three major features, *periodic*, *stochastic*, and *trend* terms based on their frequencies. The obtained periodic, stochastic, and trend features are fed into the forecasting engine to improve time-series forecasting performance (Qi et al., 2019; Bai et al., 2016). Note that the obtained stochastic, periodic, and trend features are still in the time domain and not in the frequency domain. Each of these three features captures the underlying physical aspects. The *periodic* term contains slow dynamics of an inflow time-series depending on the seasonality and meteorological condition in a given geographical area. The *stochastic* term represents the fast dynamics due to abrupt changes in meteorological parameters such as temperature and precipitation. Finally, the *trend* term shows the average trend changes of the inflow during the decades, showing the effect of long-term phenomena such as global warming. Therefore, variable X_1 can be represented by these three features as below:

$$X_1 = S_{X_1} + P_{X_1} + T_{X_1} \tag{3}$$

where S_{X_1}, P_{X_1}, and T_{X_1} represent stochastic, periodic, and trend components of time-series X_1 which are obtained from FFT analysis as follows:

$$\begin{aligned} S_{X_1} &= \sum_{j=0}^{k} IMF_j^1, && k > 1, \text{high-frequencies} \\ P_{X_1} &= \sum_{j=k+1}^{K} IMF_j^1, && k < K < m-1, \text{mid-frequencies} \\ T_{X_1} &= \sum_{j=K}^{m-1} IMF_j^1 + R_m^1, && \text{nearly-zero frequencies} \end{aligned} \tag{4}$$

where k is the number IMF_j^1 that builds the stochastic term and $K-k$ is group of IMF_j^1 that creates the periodic feature, these parameters obtain from FFT analysis. As it is presented in the above equation, the primary sub-series of IMF^1 (e.g. IMF_1^1, IMF_2^1, and IMF_3^1) represents high dynamics or high-frequency components, and it is sorted as the stochastic term. The IMF^1 with nearly zero frequency are sorted as the trend term (e.g., IMF_{m-1}^1 and residual, R_m^1), and the rest of the sub-series in IMF^1 constitute the periodic term.

3.2. Step 2: Information transfer modeling

This block employs a greedy search algorithm with a objective to maximize the transfer entropy between input features obtained from step 1 to the current target value (e.g. Inflow).

In this paper, the greedy search algorithm searches among the decomposed features (components) with different time lags to find a causal candidate set for inflow features by maximizing information theory metrics such as Transfer Entropy (TE) as its objective function. TE is based on Conditional Mutual Information (CMI). CMI and TE expression are explained as follows:

- **Conditional Mutual Information (CMI)**: $I(X;Y|Z)$ is similar to Mutual Information (MI), where all variables in an MI equation are conditioned to a third variable named Z. The CMI $I(X;Y|Z)$ is an indicator that shows the uncertainty reduction in the variable Y (e.g., inflow) from observing another variable X (e.g., precipitation) given the third variable Z (e.g., snow-depth). The CMI formulation is presented as (Novelli et al., 2019):

$$I(Y;X|Z) = H(Y|Z) - H(Y|X,Z) \tag{5}$$

$$H(Y|Z) = -\sum_{y\in X, z\in Z} P(y,z)\log_2 \frac{P(y,z)}{P(z)} \tag{6}$$

  where $H(Y|Z)$ and $H(Y|X,Z)$ are the conditional entropy and $P(z)$ and $P(y,z)$ are the marginal probability of z and joint probability between y and z respectively.
- **Transfer Entropy (TE)**: $TE_{X\to Y}$ is the degree of uncertainty reduction of variable Y_t by past values of X and Y over and above the uncertainty reduction of Y by its own past values alone (Bossomaier et al., 2016). The expression of the $TE_{X\to Y}$ is as Novelli et al. (2019):

$$\begin{aligned} TE_{X\to Y} &= I(Y_t; X_{t-1:t-k}|Y_{t-1:t-k}) = \\ H(Y_t|Y_{t-1:t-k}) &- H(Y_t|Y_{t-1:t-k}, X_{t-1:t-k}) \end{aligned} \tag{7}$$

These metrics are used to identify the nonlinear and linear dependencies among different time-series. TE is the transfer information among different source variables and a target variable, i.e., inflow (Lizier, 2012). In other words, we find a minimal set of variables that collectively contribute to the target variable's next state.

The output of the second step is a subset of source variables and inflow with some time lags $SS_{<t}$ in a predefined searching space for the set of candidates, $CS_{<t}$ that contribute to the current state of the inflow while fulfilling the statistical significance requirements. Finding the global optimal set among all the possible subsets of inflow time-series together with other sources' lagged time-series is an NP-hard problem, which exponentially grows with the size of the searching space (Sun et al., 2015). Therefore, we use a sub-optimal greedy search algorithm (Sun et al., 2015; Lizier, 2012). For better clarification, the second step is illustrated in algorithm 1.

The greedy search algorithm starts with initializing $SS_{<t}$ and $CS_{<t}$. The algorithm selects inflow time-series lags in $Inf_{<t} \subset CS_{<t}$ to perform self-prediction (Wibral et al., 2013). Next, it maximizes the CMI $I(C; Inf_t|SS_{<t})$. At the beginning, when $SS_{<t}$ =[ ], the value of $I(C; Inf_t|SS_{<t})$ is equivalent to mutual information. By adding

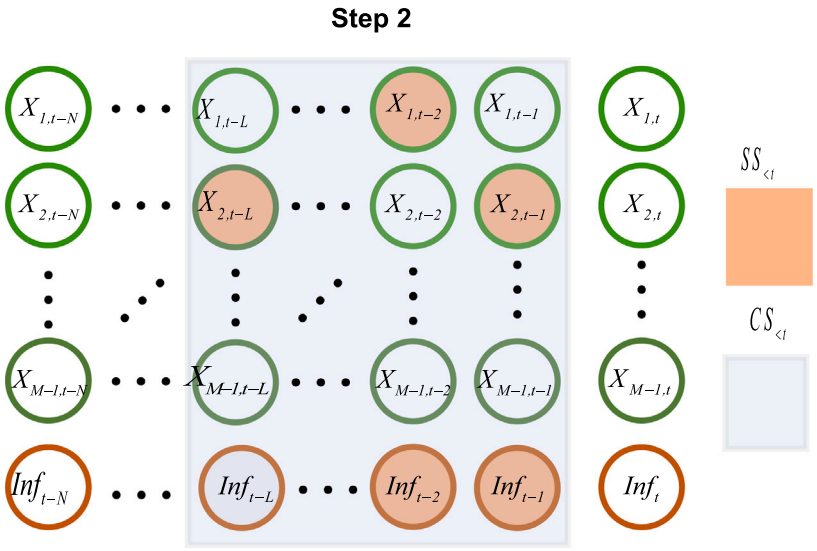

Fig. 4. The transfer modeling process of causal graph network.

one inflow lag to $SS_{<t}$, the $I(C; Inf_t|SS_{<t})$ will be equivalent to the CMI. Finally, adding lagged time-series from other sources turn the $I(C; Inf_t|SS_{<t})$ into TE. This procedure is explained in algorithm 1.

Algorithm 1 Information transfer modeling

initialize $SS_{<t}$ as an empty set
initialize $CS^c_{<t}$ based on the domain knowledge or empirically
while $CS_{<t}$ is not empty **do**
  if $inf_{<t}$ is not empty **then**
    select a candidate, C from $inf_{<t} \subset CS_{<t}$
  else
    select a candidate,C from $CS_{<t}$
  end if
  if the selected candidate maximize $I(C; Inf_t|SS_{<t})$ **then**
    check statistical significant test against $I(C; Inf_t|SS_{<t}) = 0$
    if the null hypothesis rejected **then**
      add the C to $SS_{<t}$
      remove C from $CS_{<t}$
    else
      remove C from $CS_{<t}$
    end if
  end if
end while

In Fig. 4, the procedure of obtaining $SS_{<t}$ set is illustrated using a toy model. In this example, the greedy algorithm searches within the search space of L lags and it establishes the set of Inf_{t-1}, Inf_{t-2}, $X_{2,t-1}$, $X_{1,t-2}$, and $X_{2,t-L}$ as the causal candidate set $SS_{<t}$ which contributes to the inflow current state Inf_t.

3.3. Step 3: Causal significance test

In this step, we check the causal strength of each variable in the causal candidate set, $SS_{<t}$, concerning the target variable (inflow in our case) by removing the influence of that variable in the set. Such a pruning process converges the inferred network to a causal network (Sun et al., 2015; Novelli et al., 2019). The causal strength test is explained in algorithm 2.

This process continues until all selected candidates in $SS_{<t}$ are tested. Moreover, an additional omnibus test have been done to secure the statistical significance against zero transfer entropy. Fig. 5 depicts

Algorithm 2 Causal Significant Test

for a selected C in $SS_{<t}$ **do**
if the conditionally excluded candidate minimize the CMI $I(C; Inf_t|SS_{<t}\backslash\{C\})$ **then**
  perform statistical significant test against conditional independent (null hypothesis)
  if the null hypothesis is rejected **then**
    keep the C in $SS_{<t}$
  else
    remove C from $SS_{<t}$
  end if
end if
end for

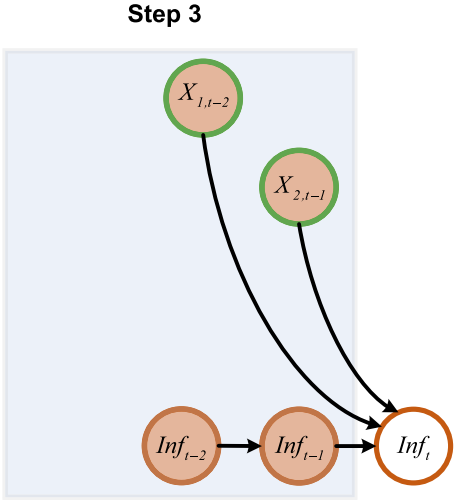

Fig. 5. The transfer modeling process of causal graph network.

the outcome of Step 3 on the CED framework. For example $X_{2,t-L}$ is removed from obtained $SS_{<t}$ in Fig. 5 because it is detected as a variable with non-significant causal strength.

Table 1
LSTM hyperparameters.

General hyperparameters for LSTM, RNN and MLP			
Activation function	Relu	Learning rate	0.001
Number of hidden layers	2	Optimizer	Adam
Loss function	MSE	Epoch	100
LSTM			
Neurons of first hidden layer	1000	Neurons of second hidden layer	150
Dropout rate for the first hidden layer	0.2	Dropout rate for the second hidden layer	0
Return sequence for the first hidden layer	True	Return sequence for the second hidden layer	False
Batch size	100	Neurons of output layer	1

3.4. Step 4: Forecasting

In step 4, the obtained causal graph is used for target-related (inflow-related) feature selection. Then, we feed significant causal features into a typical machine learning-based forecasting algorithm. In this paper, we consider four widely used algorithms in the literature for inflow forecasting, e.g., LR, MLP, RNN, and LSTM. To save the space in the paper, the description for each one of them is not provided here but can be find in Qi et al. (2019), Abdellatif et al. (2015), Tayebiyan et al. (2016), Apaydin et al. (2020). In addition, these are the models developed in Keras and Sciket learn libraries (Pedregosa et al., 2011; Chollet et al., 2015).

4. Results & discussions

The analysis is performed on a laptop with Core i7 Intel GPU, 32 GB RAM, and NVIDIA GeForce RTX 2080. The computation time of different scenarios ranges from less than one second for CED-LR to 250 s for LSTM with stand-alone decomposition preprocessing techniques for training each model. In addition, the grid search algorithm is used for finding the hyperparameters of each models like LSTM, MLP and RNN. For example, the grid-search set to the following ranges to find the optimal hyperparameters for the LSTM model: Learning rate= [0.1, 0.01, 0.001], activation functions= [Relu, Tanh, Sigmoid], first layer neurons=[500, 1000, 1500, 2000], second layer neurons=[100, 150, 200], batch-sizes= [30, 50, 100, 200, 500] and optimizer=[Adam, Adammax]. To save the space in the paper, only one table of hyperparameters related to the Sula area is presented (see Table 1).

In the following sub-sections, the results are provided for each block presented in Fig. 3. The effectiveness of the proposed CED framework is validated for two different use cases. Hence, the comparison is performed through two well-known performance indicator; the normalized root means square error (NRMSE) and the coefficient of determination as follows:

$$NRMES = \frac{\sqrt{\sum_{t=1}^{N}(Inf_t - \widehat{Inf}_t)^2}}{\overline{\text{Inf}}} \tag{8}$$

$$R^2 = 1 - \frac{\sum_{t=1}^{N}(Inf_t - \widehat{Inf}_t)^2}{\sum_{t=1}^{N}(Inf_t - \overline{\text{Inf}})^2} \tag{9}$$

where Inf_t and $\widehat{Inf}_t$ are actual inflow and forecasted inflow values at time t, respectively. $\overline{\text{Inf}}$ is the expected value or mean of Inf. NRMSE is a normalized form of the root mean square error (RMSE). The logic behind selecting RMSE or mean square error (MSE) is to penalize large error values, which often occur during the peak of inflow, exponentially more than many smaller errors. For example, 3 errors of [1, 1, 6] are significantly worse than [2, 2, 2], that lost value is 2 when a mean absolute error (MAE) is used. This property of assigning a higher loss value to larger errors is beneficial when more stability is wanted (Osberg, 2020). In this paper, for model training and validation, MSE is used to minimize large errors rather than minimizing small errors. In inflow forecasting problems, minimizing large errors have higher importance because of the cost of reservoir spillage and the environmental and social cost of the flood. Moreover, NRMSE and the coefficient of determination are used as the performance criteria for the validation of the models. The best possible scores for R^2 and NRMSE are 1 and 0, respectively. The closer R^2 is to 1 or NRMSE is to 0, the better trained the model is toward the unseen samples.

4.1. Multivariate decomposition

The MEMD method is employed to decompose input variables' time-series into their IMF sub-series. The projection number Num_d is a parameters of MEMD to extract proper number of IMF for a given time-series (Rehman and Mandic, 2010). We select $Num_d = 256$ based on our heuristic analysis and the literature (Rehman and Mandic, 2010). There are other parameters such as the stopping criteria condition for controlling the sifting process, which is selected as a vector $[0.05, 0.05, 0.05]$ as recommended in Rilling et al. (2003). The results of the multivariate decomposition block are presented for the inflow variable in Fig. 6. Applying MEMD on inflow and other variables results in 21 sub-series and a residual signal for the last IMF. As it can be seen in Fig. 6(c), (e) and (g), the inflow stochastic feature is a summation of the seven first IMF $S_{inf} = IMF_1^{inf} + IMF_2^{inf} \cdots + IMF_7^{inf}$, the Inflow's periodic feature is $P_{inf} = IMF_8^{inf} + IMF_9^{inf} \cdots + IMF_{16}^{inf}$, and the inflow's trend feature is $T_{inf} = IMF_{17}^{inf} + IMF_{18}^{inf} \cdots + R^{inf}$. Usually, the first IMF^{inf} includes the highest dynamics of a time-series while the last IMF (R^{inf}) contains the lowest dynamics.

The S_{Inf}, P_{Inf} and T_{Inf} in Fig. 6 are obtained from the FFT analysis by sorting the IMFs based on their similar frequencies to form stochastic, periodic, and trend features of the inflow. The FFT analysis was applied to all variables' time-series, but for visualization, only inflow's FFT conversion featured is presented in Fig. 6(b), (d), (f), and (h). It shows that high frequencies are sorted in the stochastic feature representing daily dynamics with dominant frequencies from 0.1 to 0.5 $(1/day)$ (such dynamics range from one to two days). The periodic features carry the main frequencies from 0.0027 to 0.03 $(1/day)$, representing time-series with yearly, half-yearly, seasonal, and monthly periods. The last feature is the trend which contains nearly zero frequencies.

Generally, forecasting trend T_X and periodic P_X features are not challenging for ML forecasting algorithms because they are less noisy. On the contrary, the most challenging part is forecasting stochastic features because they are nonlinear and noisy. Therefore, adding supplementary variables from meteorological and hydrological data such as temperature, snow-depth, or precipitation may reduce the stochastic inflow feature forecasting uncertainties. Hence, causality-based analysis is essential to find the best set of variables that can improve the inflow forecasting performance. A sensitivity analysis is provided in Section 4.3 in which the forecasting accuracy of inflow features are compared with CED and without CED in Table 3 and Fig. 9.

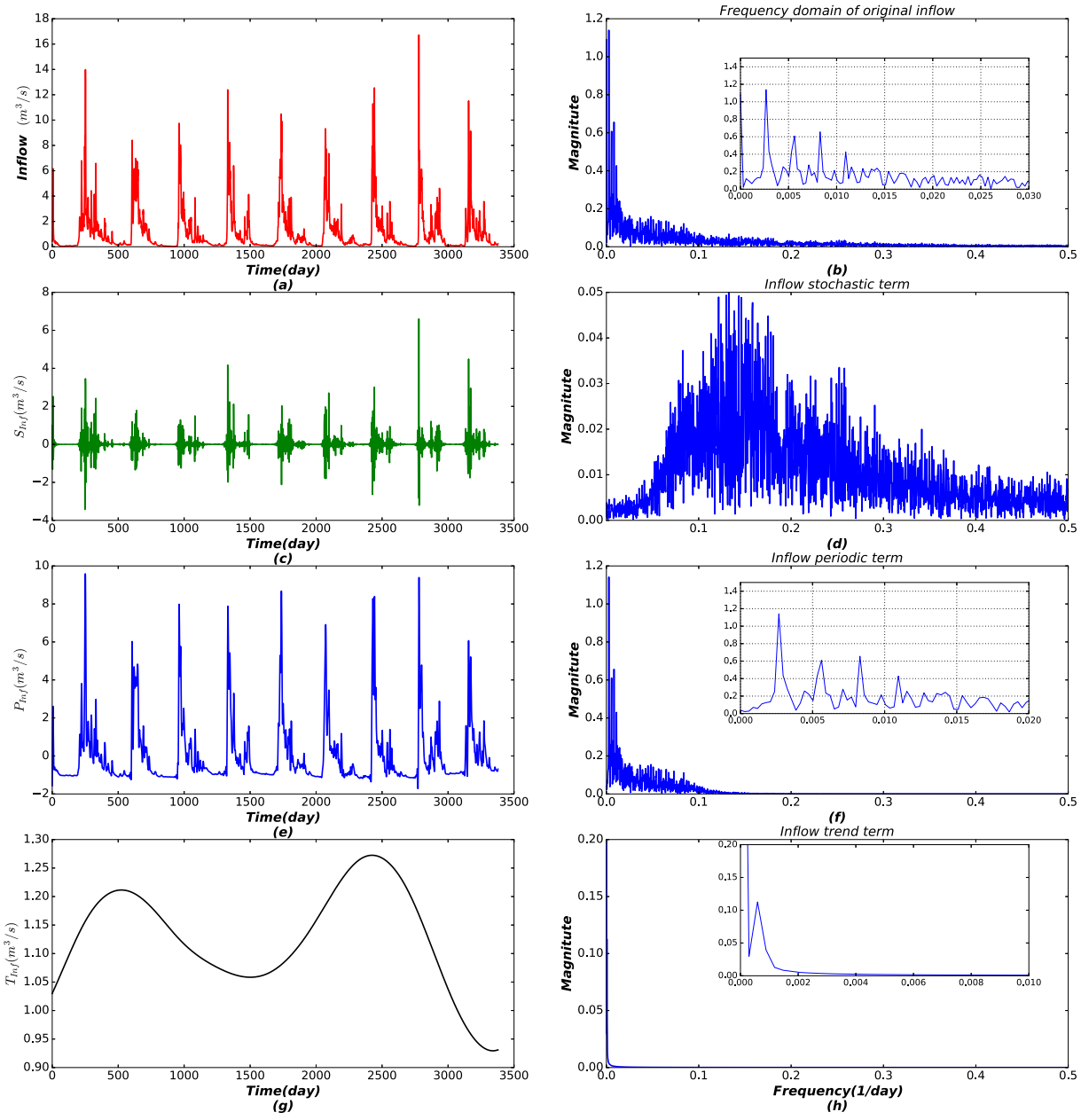

Fig. 6. Decomposition results of inflow variable from 2010 to 2019 for Sula area.

Table 2
Data variables with full name.

Name	Variable	Name	Variable
Inflow	Inf	Periodic of Tem_{max}	P_{max}
Stochastic of Inf	S_{inf}	Trend of Tem_{max}	T_{max}
Periodic of Inf	P_{inf}	Relative humidity	Rh
Trend of Inf	T_{inf}	Stochastic of Rh	S_{rh}
average-temperature	Tem_{ave}	Periodic of Rh	P_{rh}
Stochastic of Tem_{ave}	S_{ave}	Trend of Rh	T_{rh}
Periodic of Tem_{ave}	P_{ave}	Snow-depth	Sno
Trend of Tem_{ave}	T_{ave}	Stochastic of Sno	S_{sno}
Minimum-temperature	Tem_{min}	Periodic of Sno	P_{sno}
Stochastic of Tem_{min}	S_{min}	Trend of Sno	T_{sno}
Periodic of Tem_{min}	P_{min}	Precipitation	Pr
Trend of Tem_{min}	T_{min}	Stochastic of Pr	S_{pr}
Max-temperature	Tem_{max}	Periodic of Pr	P_{pr}
Stochastic of Tem_{max}	S_{max}	Trend of Pr	T_{pr}

4.2. Causal graph from inflow features

The feature decomposition block as explained in Section 3.1 creates a data set with 21 variables (see Table 2). Moreover, some of these variables with a certain delay can significantly impact the inflow forecasting problem. Therefore, these variables with their different lags need to be taken into account. Thereby, let us assume all variables are Markovian with $L = 30$ lags, a searching space with the size of 630 (30×21) can be created. However, feeding all these variables to the ML forecasting algorithms may cause over-fitting problem. Thus, finding the correct variables with the right lags which can maximize the improvement impact on inflow forecasting performance is a key that can be done by steps 2 and 3 in the CED framework. We assume 30 days for the maximum lag, L, which is based on our heuristic analysis and input from hydropower experts. As explained in steps 2 and 3 of the CED algorithm in the methodology, step 3 is a small set of variables

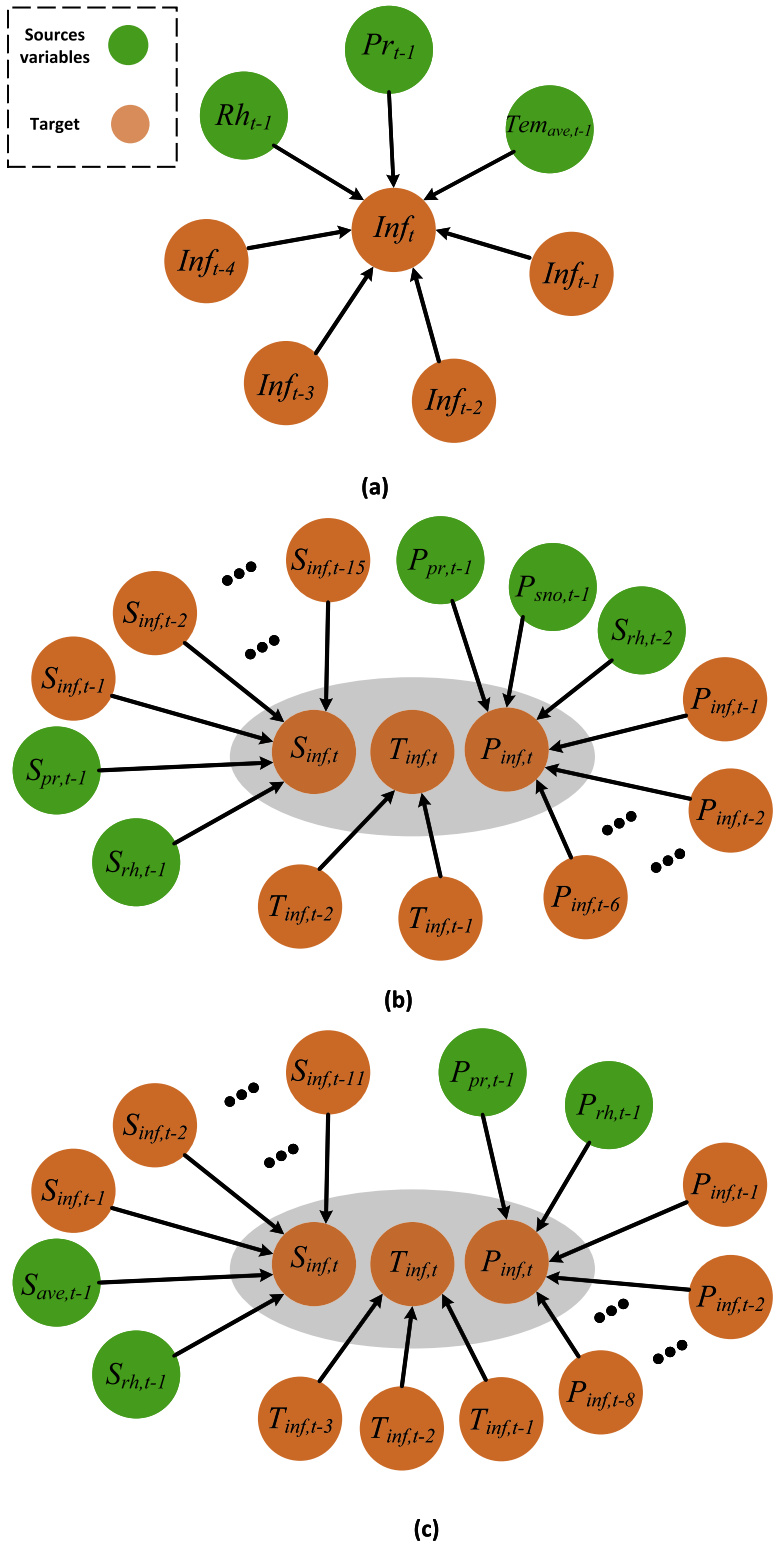

Fig. 7. Causal model for inflow values and its predictors: (a) original time-series values For Sula and Storeskar; (b) decomposed time-series values for Sula area; (c) decomposed time-series values for Storeskar area.

out of the 630 variables based on their significant causal relationship with the inflow.

We illustrate the resultant causal models for the Sula and Storeskar use-cases in Fig. 7. The dependencies between the original inflow time-series with its lag values and the lags of other time-series variables are presented in Fig. 7(a). It shows the average temperature, precipitation, and humidity time-series with their first lag (one day ago) values have the most contribution in forecasting the current state of the inflow. Moreover, the first four lags of the historical inflow can significantly contribute to the current state of the inflow.

Fig. 7(b) and (c) show the causal relationship between decomposed inflow features at the current time $S_{inf,t}$, $T_{inf,t}$ and $P_{inf,t}$ with their lags and lags of other resource variables for Sula and Storeskar areas

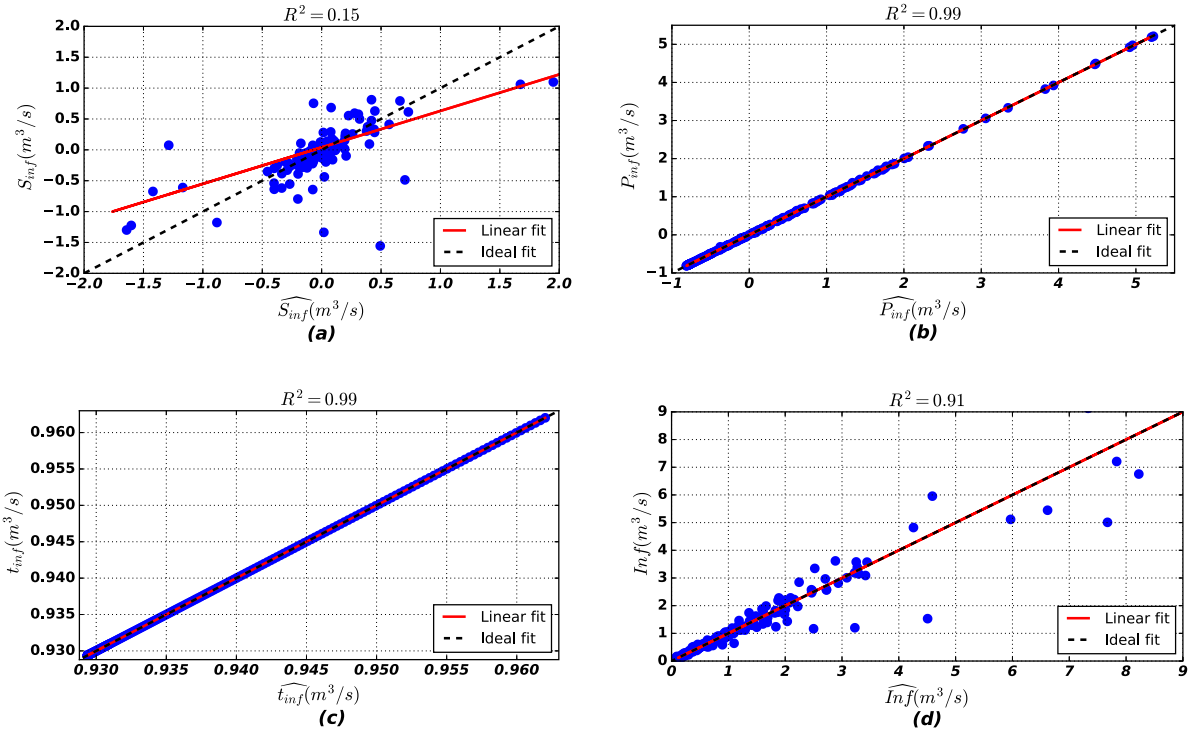

Fig. 8. CED-LR Forecasting results of inflow and its features for Sula area: (a) scatter plot of inflow stochastic term; (b) scatter plot of inflow periodic term; (c) scatter plot of inflow trend term; and (d) scatter plot of inflow.

respectively. Fig. 7(b) shows that the stochastic features of precipitation and humidity from the last day highly affect the stochastic inflow values at the current state. Moreover, the inflow's stochastic feature has significant dependencies on its own recorded values until 15 days ago. In addition, the periodic feature of inflow at the current time depends on the $P_{sno,t-1}$, $P_{pr,t-1}$ and $S_{rh,t-2}$ values. This is because the periodic pattern of inflow shows the seasonality behavior including snowy, snow-melting, rainy, and dry periods. Since the dynamic in the inflow's periodic feature is lower than the stochastic feature, its dependency on its past values is lower (depends on its six lags) than the stochastic feature. In addition, the trend feature of the inflow has almost no dynamic. So, it does not exhibit directed information with any exogenous variables in our data set. Instead, it only depends on its first two past values.

Another interesting observation from Fig. 7(b), is that decomposition of inflow and other weather-related enables us to explore more causal relationships between various components of input data and the current state of inflow. For example, Snow-depth was not found in causal graph without decomposition in Fig. 7(a), while the snow-depth periodic feature has an impact on the inflow periodic feature. The decomposition process denoises and detrends the original signals, strengthening steps 2 and 3 to discover more information-rich variables for the inflow forecasting problem. Moreover, the resultant causal model for decomposed values from the second use case, Storeskar area, is more distinguishable from our first use case, Sula area, see Fig. 7(c). However, the causal graph without decomposition step looks very similar for both use cases in Fig. 7(a).

In the next step, the obtained causal graphs for each feature of inflow are employed for training different ML tools to forecast the inflow features.

4.3. Inflow forecasting

In this step, forecasting techniques such as LR, MLP, RNN and LSTM are used to forecast S_{inf}, P_{inf}, and T_{inf} components. For training and testing of the forecasting algorithms, actual data from the case study areas (presented in Fig. 1) are divided into three parts; training, validation, and testing. The period from 01/06/2019 to 31/12/2019 is selected as the testing data and the remaining time is selected as training and validation data sets with 90% and 10% ratio respectively.

The inputs for LR, MLP, RNN and LSTM without CED is 30 past values of all the input variables presented in Section 2, while the inputs with CED is obtained in Section 4.2 and presented in Fig. 7 for each feature and use cases. After training the model, it is validated with the test data set. As an example, the results of inflow features forecasting are illustrated for the Sula area using the LR model in Fig. 8(a), (b), and (c). Fig. 8 shows that the obtained models accurately forecast the inflow features. The scatter plot also shows a good match between the realized values and the forecasted values, especially for the periodic and trend terms where $R^2 = 0.99$. The forecasted inflow is the summation of forecasted values of its stochastic, periodic, and trend terms, see Fig. 8(d). As it is observed, the forecast performance of CED-LR for Sula area is better than Storeskar area (look at their R^2 scores) because the inflow changes and nonlinearities in Storeskar is much higher than Sula area as explained and presented in Section 2 and Fig. 2 respectively. For the sake of better understanding the role of CED in improving forecasting accuracy of inflow features, a sensitivity analysis have been done for two scenarios with CED-MLP and MLP without-CED (stand alone MLP) to show which feature of inflow is more challenging. The results are provided in Table 3 and Fig. 9. According to Table 3, the accuracy of both trend and periodic features significantly improves by 66% and 97.5% for the RMSE when CED is used as a pre-processing step for MLP. As presented in Fig. 9, both trend and periodic futures are less nonlinear and noisy (represent nearly zero and slow inflow frequencies). The challenging feature is the stochastic feature which has the fastest dynamics of inflow. However, using CED can improve the accuracy of the stochastic feature by almost 37.5%.

The next step is to validate the performance of the proposed CED framework by comparing it with other forecasting algorithms.

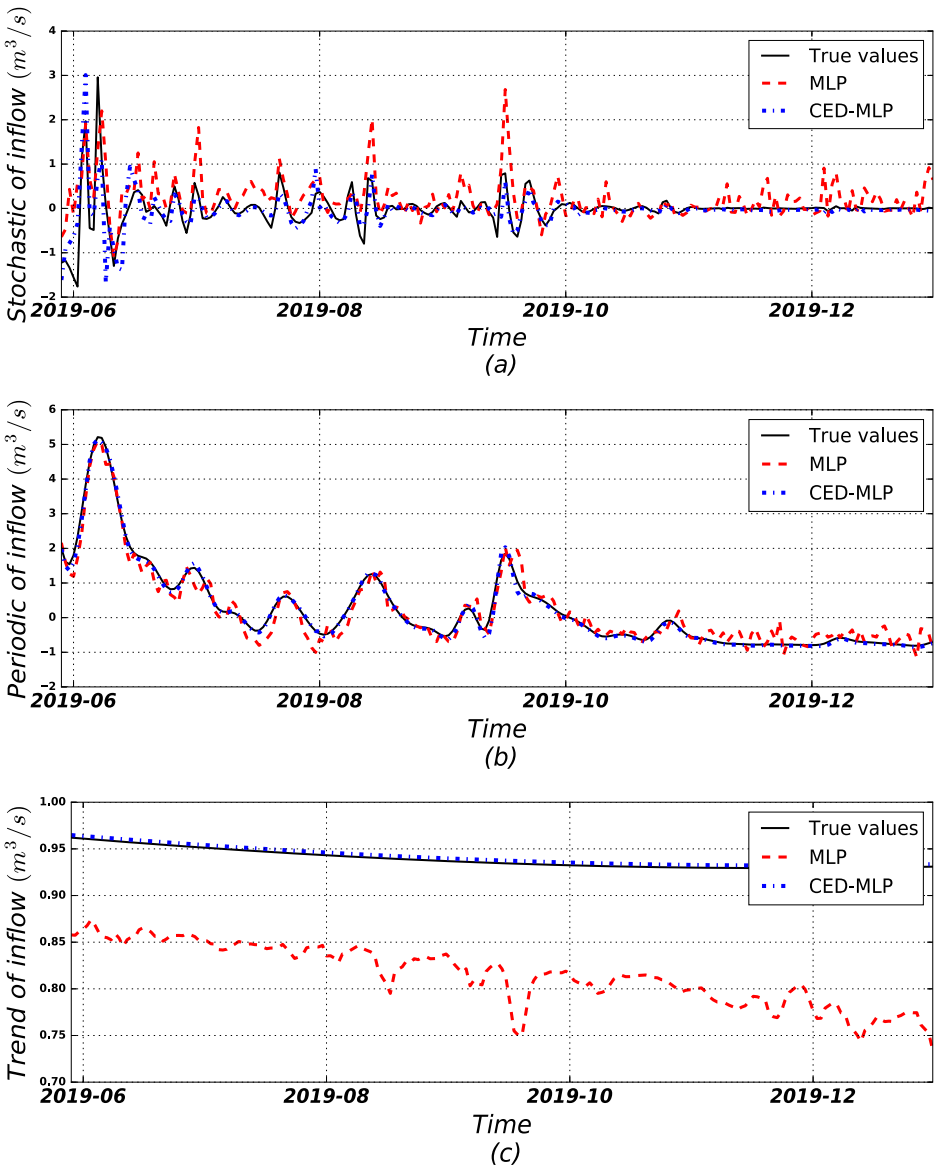

Fig. 9. Sensitivity analysis of inflow features forecasting with CED or without CED by MLP model from June 2019 to first of January 2020 at Sula area: (a) stochastic forecasting; (b) periodic forecasting; (c) trend forecasting.

Table 3
Sensitivity analysis of inflow features forecasting accuracy for two scenarios; with CED and without CED for MLP model in Sula area.

Inflow features	MLP		CED-MLP	
	RMSE (m^3/s)	R^2	RMSE (m^3/s)	R^2
S_{inf}	0.48	0.12	0.30	0.36
P_{inf}	0.28	0.94	0.07	0.99
T_{inf}	0.11	0.026	0.0028	0.91

4.4. CED validation

To validate the performance of our proposed CED framework for time-series feature selection, we use it as a pre-processing step with a number of widely used forecasting algorithms such as LR, MLP, RNN, and LSTM. Then the results are compared with the baseline (no pre-processing) case. The results in Table 4 show that CED pre-processing improves forecasting accuracy across different forecasting methods. However, in the case of using only decomposition pre-processing (removing steps 2 and 3 in Fig. 3) or only causal pre-processing (removing step 1 in Fig. 3), the inflow forecasting accuracy is less than using CED pre-processing. For example, the CED pre-processing with LSTM forecasting method results in 70%, 64%, 27% relative improvement in reducing forecasting error (NRMSE) in comparison to no pre-processing, decomposition pre-processing, and causal based pre-processing, respectively. The improvement trend is also similar for the R^2 score as the CED-MLP improves the forecasting accuracy by 30% compared with no-preprocessing MLP.

It may be surprising why the LR outperforms the other models in the first column of Table 4. Using deep learning methods in a multivariate inflow forecasting problem when input is the raw data without pre-processing (feature extraction or selection) increases the risk of over-fitting and information saturation which might result in poorly trained models. According to Table 4, when only Causality is used as the pre-processing step, the LSTM performance outperforms all the other models. However when only the decomposition technique is used LR outperforms all the others. Because the decomposition technique reduces the complexity and non-linearities of the forecasting problem by breaking down the original time-series into three terms stochastic, periodic and trend, while causality reduces the risk of over-fitting and

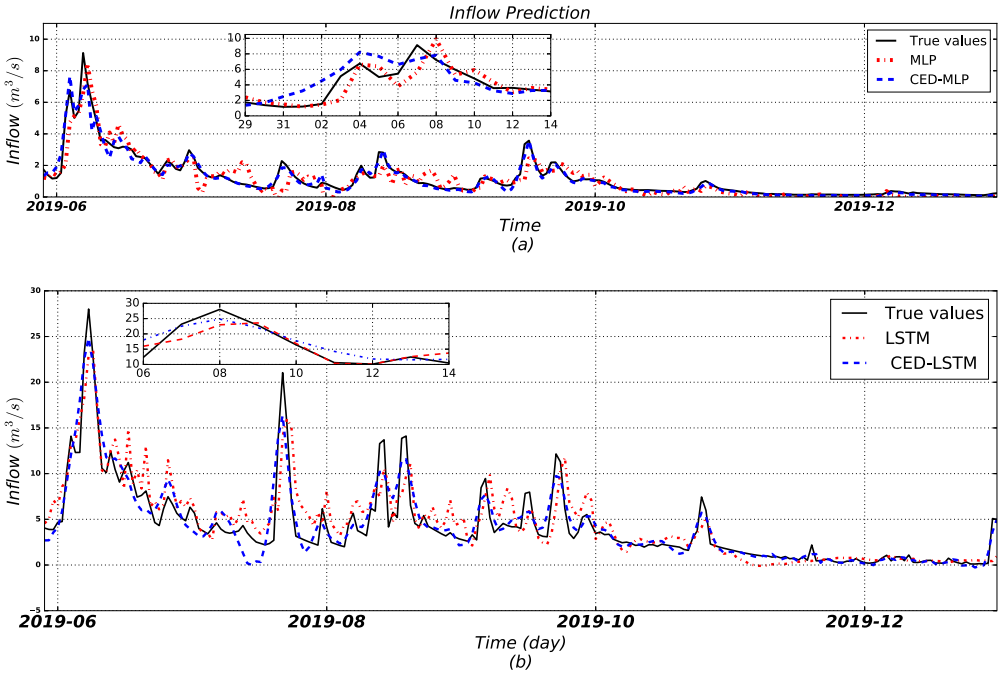

Fig. 10. Multivariate inflow forecasting comparison results from June 2019 to first of January 2020: (a) Sula area; (b) Storeskar area.

Table 4
Comparison analysis on different pre-processing techniques for inflow forecasting in Sula area.

Model	No pre-process		Causal pre-process		Decompose pre-process		CED	
	NRMSE	R^2	NRMSE	R^2	NRMSE	R^2	NRMSE	R^2
LR	0.40	0.81	0.32	0.819	0.16	0.90	0.1424	0.91
MLP	0.56	0.80	0.4	0.874	0.25	0.88	0.14	0.95
RNN	0.68	0.65	0.45	0.84	0.48	0.7	0.21	0.904
LSTM	0.5	0.7	0.198	0.883	0.4	0.77	0.143	0.906

Table 5
Comparison analysis on different pre-processing techniques for inflow forecasting in Storeskar area.

Model	No pre-process		Causal pre-process		Decompose pre-process		CED	
	NRMSE	R^2	NRMSE	R^2	NRMSE	R^2	NRMSE	R^2
LR	0.58	0.61	0.52	0.65	0.24	0.85	0.20	0.90
MLP	0.63	0.52	0.57	0.63	0.55	0.61	0.25	0.86
RNN	0.67	0.47	0.57	0.61	0.48	0.47	0.24	0.87
LSTM	0.64	0.60	0.35	0.65	0.33	0.68	0.19	0.90

information saturation of the forecasting problem by finding the best informative lag values of each source. Moreover, stand-alone causality techniques may fail to distinguish crucial exogenous parameters in input data for the forecasting purpose (See Section 4.2). Our proposed hybrid CED framework takes advantage of both techniques' strengths and improves the inflow forecasting model's performance. Last, but not least, the CED is implemented for the second use case to validate the CED scalability for the inflow forecasting problem. The results of the second use case are presented in Table 5. We observe a similar enhancement in inflow forecasting for the second area using a CED feature selection framework.

It is worth mentioning some of the interesting observations in the presented results in Tables 4 and 5 as follows:

- Feeding multiple time-series from different domains (e.g. meteorological, hydrological, etc.) into an inflow forecasting algorithm without pre-processing does not necessarily improve the forecasting results. Our proposed CED framework highly improves inflow forecasting results with down-selecting the most informative part of input data.
- We picked four different forecasting methods ranging from classical algorithms such as LR and MLP to more complex techniques like RNN and LSTM for CED validation. All the selected state-of-the-art forecasting methods exhibit relatively performance enhancement after using the CED pre-processing step. CED can serve as a complementary pre-processing step for any forecasting method.

Inflow forecasting results using CED-MLP and CED-lSTM framework in comparison with MLP and LSTM (no pre-processing) are shown in Fig. 10(a) and (b) for Sula and Storeskar use cases respectively. As it is presented, both CED-MLP and CED-LSTM have a much better performance in forecasting the real values of inflow for different use cases. Another interesting observation in Fig. 10 is that CED-MLP and CED-LSTM accurately forecast inflow's peak values, which is a critical and challenging task for state-of-the-art inflow forecasting methods and has significant consequences for short-term hydropower scheduling.

Another interesting aspect is to present the sensitivity of the CED algorithm to the size of training horizons and how often the model should be updated and retrained accordingly. The sensitivity result of the CED-MLP is presented in Fig. 11, with different lengths of historical data

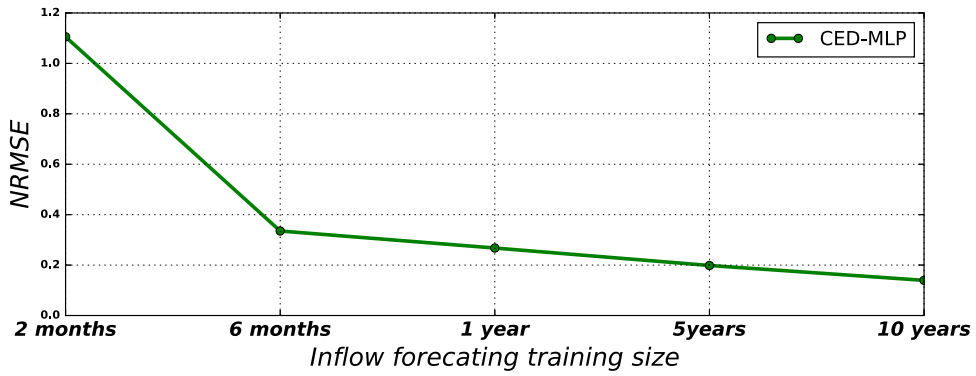

Fig. 11. Sensitivity analysis of the CED-MLP to the train size for the Sula area.

(2 months, 6 months, 1 year, 5 years, 10 years) for day-ahead inflow forecasting. It shows after 5 years, the forecast error improvement is marginal. Therefore, 5 to 10 years of historical data is good enough to make a causal model in our use case. It is worth mentioning that climate change may alter the historical temperature and precipitation patterns in different parts of the world. Therefore, from a practical point of view, it is worthwhile to update the causal models and retrain the model annually using a five to ten years historical horizon.

5. Practical consideration

From a practical point of view, accurate inflow forecasting models can significantly improve the hydropower scheduling performance, flooding control measures, water drinking and agriculture efficiency. In this paper, we focused on daily inflow forecasting mainly for short-term hydro-power scheduling problems with large-storage reservoirs in Norway.

Usually, most of the precipitation in Norway especially close to the glaciers, is in the form of snow. Looking at the historical values of inflow for two different use cases in Fig. 2, shows that the peak of inflow always happens at the end of spring or early summer when the temperature rise and snow starts to melt. Snow melting may occur in two weeks time period. Therefore, it is crucial to estimate these periods precisely to avoid the risk of spillage by providing enough space for the maximum harnessing of inflow to the reservoirs. Our proposed CED pre-processing framework enables the forecasting models to discover the dependencies of inflow to the lag values of other variables such as snow-dept, temperature, and precipitation. This feature improves the estimation of this period which is a critical factor in inflow forecasting in Norway. Fig. 10(a) and (b) show that using CED pre-processing with forecaster models LSTM and MLP improves the forecasting accuracy at peak values.

In addition to meteorological data that we used in this paper, other variables, such as topography, soil characteristics, etc., can impact the inflow forecasting. The proposed CED pre-processing framework is expandable to a higher number of input source time series (more input variables). As a direction for future works, we will consider expanding source values beyond what we used in this paper.

6. Conclusion

This paper proposes a Causal Multivariate Empirical mode decomposition (CED) framework for multiple time-series feature selection. In other words, CED is a pre-processing step that is complementary to any machine learning-based forecasting method. We specifically developed CED for the inflow forecasting problem by taking advantage of both decomposition and causal inference approaches for time-series. At the same time, we overcome the inherent weaknesses in both approaches. In other words, CED combines physics-based causal inference with signal processing-based decomposition to get the more relevant features among multiple time-series to the target variable (the inflow). Our proposed CED framework is validated for two different use cases related to hydropower reservoirs in Norway. The validation results show that using CED as a pre-processing step significantly enhances the forecasting accuracy for various forecasting methods. For example, adding CED as a pre-processing step to LSTM and LR improved forecasting results by almost 70%. We are trying to use more hydrological parameters as input and expand the CED framework's inflow forecasting horizon for future work.

CRediT authorship contribution statement

Mojtaba Yousefi: Writing – original draft, Writing – review & editing, Visualization, Methodology, Validation, Software, Investigation. **Xiaomei Cheng:** Collecting data, Sorting data and cleaning data. **Michele Gazzea:** Assisting in visualization. **August Hubert Wierling:** Writing – review & editing, Providing technical comments and supervising. **Jayaprakash Rajasekharan:** Writing – review & editing, Providing technical comments. **Arild Helseth:** Providing technical and domain knowledge. **Hossein Farahmand:** Project coordination, Project administration, Providing technical and domain knowledge, Writing – review & editing. **Reza Arghandeh:** Supervision, Providing technical comment, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is partly supported by The Research Council of Norway (grant number: 309997) and internal fund from Western Norway University of Applied Sciences (HVL university project number: 1030730).

References

Abdellatif, M.E., Osman, Y.Z., Elkhidir, A.M., 2015. Comparison of artificial neural networks and autoregressive model for inflows forecasting of roseires reservoir for better prediction of irrigation water supply in sudan. Int. J. River Basin Manage. 13 (2), 203–214.

Anon, 0000a. The Norwegian Water Resources and Energy Directorate.

Anon, 0000b. The Norwegian Climate Service.

Apaydin, H., Feizi, H., Sattari, M.T., Colak, M.S., Shamshirband, S., Chau, K.-W., 2020. Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting. Water 12 (5), 1500.

Apaydin, H., Sibtain, M., 2021. A multivariate streamflow forecasting model by integrating improved complete ensemble empirical mode decomposition with additive noise, sample entropy, gini index and sequence-to-sequence approaches. J. Hydrol. 603, 126831.

Bai, Y., Chen, Z., Xie, J., Li, C., 2016. Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models. J. Hydrol. 532, 193–206.
Bai, Y., Wang, P., Xie, J., Li, J., Li, C., 2015. Additive model for monthly reservoir inflow forecast. J. Hydrol. Eng. 20 (7), 04014079.
Bennett, J.C., Wang, Q., Li, M., Robertson, D.E., Schepen, A., 2016. Reliable long-range ensemble streamflow forecasts: Combining calibrated climate forecasts with a conceptual runoff model and a staged error model. Water Resour. Res. 52 (10), 8238–8259.
Bordin, C., Skjelbred, H.I., Kong, J., Yang, Z., 2020. Machine learning for hydropower scheduling: state of the art and future research directions. Procedia Comput. Sci. 176, 1659–1668.
Bossomaier, T., Barnett, L., Harré, M., Lizier, J.T., 2016. Transfer entropy. In: An Introduction to Transfer Entropy. Springer, pp. 65–95.
Cheng, X., Farahmand, H., Yousefi, M., Arghandeh, R., Wang, H., 2021. Inflow forecasting based on principal component analysis and long short term memory. In: 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress. DASC/PiCom/CBDCom/CyberSciTech, IEEE, pp. 589–596.
Cheng, C.-T., Feng, Z.-K., Niu, W.-J., Liao, S.-L., 2015. Heuristic methods for reservoir monthly inflow forecasting: A case study of xinfengjiang reservoir in Pearl River, China. Water 7 (8), 4477–4495.
Chollet, F., et al., 2015. Keras.
Coulibaly, P., Anctil, F., Bobee, B., 2001. Multivariate reservoir inflow forecasting using temporal neural networks. J. Hydrol. Eng. 6 (5), 367–376.
Golob, R., Štokelj, T., Grgič, D., 1998. Neural-network-based water inflow forecasting. Control Eng. Pract. 6 (5), 593–600.
Han, H., Choi, C., Jung, J., Kim, H.S., 2021. Deep learning with long short term memory based sequence-to-sequence model for Rainfall-Runoff simulation. Water 13 (4), 437.
Herbert, Z.C., Asghar, Z., Oroza, C.A., 2021. Long-term reservoir inflow forecasts: Enhanced water supply and inflow volume accuracy using deep learning. J. Hydrol. 601, 126676.
Huang, Y., Deng, C., Zhang, X., Bao, Y., 2020. Forecasting of stock price index using support vector regression with multivariate empirical mode decomposition. J. Syst. Inf. Technol..
Jothiprakash, V., Kote, A.S., 2011. Effect of pruning and smoothing while using M5 model tree technique for reservoir inflow prediction. J. Hydrol. Eng. 16 (7), 563–574.
Kao, I.-F., Liou, J.-Y., Lee, M.-H., Chang, F.-J., 2021. Fusing stacked autoencoder and long short-term memory for regional multistep-ahead flood inundation forecasts. J. Hydrol. 598, 126371.
Kao, S.-C., Sale, M.J., Ashfaq, M., Martinez, R.U., Kaiser, D.P., Wei, Y., Diffenbaugh, N.S., 2015. Projecting changes in annual hydropower generation using regional runoff data: An assessment of the United States federal hydropower plants. Energy 80, 239–250.
Kao, I.-F., Zhou, Y., Chang, L.-C., Chang, F.-J., 2020. Exploring a long short-term memory based encoder-decoder framework for multi-step-ahead flood forecasting. J. Hydrol. 583, 124631.
Liao, S., Liu, Z., Liu, B., Cheng, C., Jin, X., Zhao, Z., 2020. Multistep-ahead daily inflow forecasting using the ERA-interim reanalysis data set based on gradient-boosting regression trees. Hydrol. Earth Syst. Sci. 24 (5), 2343–2363.
Lizier, J.T., 2012. The Local Information Dynamics of Distributed Computation in Complex Systems. Springer Science & Business Media.
Lohani, A.K., Goel, N., Bhatia, K., 2014. Improving real time flood forecasting using fuzzy inference system. J. Hydrol. 509, 25–41.
Luo, X., Yuan, X., Zhu, S., Xu, Z., Meng, L., Peng, J., 2019. A hybrid support vector regression framework for streamflow forecast. J. Hydrol. 568, 184–193.
Moazenzadeh, R., Mohammadi, B., Shamshirband, S., Chau, K.-w., 2018. Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran. Eng. Appl. Comput. Fluid Mech. 12 (1), 584–597.
Ni, L., Wang, D., Wu, J., Wang, Y., Tao, Y., Zhang, J., Liu, J., 2020. Streamflow forecasting using extreme gradient boosting model coupled with Gaussian mixture model. J. Hydrol. 586, 124901.
Novelli, L., Wollstadt, P., Mediano, P., Wibral, M., Lizier, J.T., 2019. Large-scale directed network inference with multivariate transfer entropy and hierarchical statistical testing. Netw. Neurosci. 3 (3), 827–847.
Okkan, U., Serbes, Z.A., 2013. The combined use of wavelet transform and black box models in reservoir inflow modeling. J. Hydrol. Hydromech. 61 (2), 112–119.
Osberg, M.M., 2020. LSTM hybrid model for water reservoir inflow forecasting, a comparison between black box-and interpretable hybrid models. NTNU.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.
Qi, Y., Zhou, Z., Yang, L., Quan, Y., Miao, Q., 2019. A decomposition-ensemble learning model based on LSTM neural network for daily reservoir inflow forecasting. Water Resour. Manage. 33 (12), 4123–4139.
Rehman, N., Mandic, D.P., 2010. Multivariate empirical mode decomposition. Proc. R. Soc. A Math. Phys. Eng. Sci. 466 (2117), 1291–1302.
Rilling, G., Flandrin, P., Goncalves, P., et al., 2003. On empirical mode decomposition and its algorithms. In: IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing. Vol. 3. Citeseer, pp. 8–11.
Roushangar, K., Alizadeh, F., 2018. Entropy-based analysis and regionalization of annual precipitation variation in Iran during 1960–2010 using ensemble empirical mode decomposition. J. Hydroinform. 20 (2), 468–485.
Sehgal, V., Sahay, R.R., Chatterjee, C., 2014a. Effect of utilization of discrete wavelet components on flood forecasting performance of wavelet based ANFIS models. Water Resour. Manage. 28 (6), 1733–1749.
Sehgal, V., Tiwari, M.K., Chatterjee, C., 2014b. Wavelet bootstrap multiple linear regression based hybrid modeling for daily river discharge forecasting. Water Resour. Manage. 28 (10), 2793–2811.
Sriram, L.M.K., Gilanifar, M., Zhou, Y., Ozguven, E.E., Arghandeh, R., 2018. Causal markov elman network for load forecasting in multinetwork systems. IEEE Trans. Ind. Electron. 66 (2), 1434–1442.
Sun, J., Taylor, D., Bollt, E.M., 2015. Causal network inference by optimal causation entropy. SIAM J. Appl. Dyn. Syst. 14 (1), 73–106.
Tayebiyan, A., Mohammad, T.A., Ghazali, A.H., Mashohor, S., 2016. Artificial neural network for modelling rainfall-runoff. Pertanika J. Sci. Technol. 24 (2), 319–330.
Tongal, H., Booij, M.J., 2018. Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. J. Hydrol. 564, 266–282.
Tsai, M.-J., Abrahart, R.J., Mount, N.J., Chang, F.-J., 2014. Including spatial distribution in a data-driven rainfall-runoff model to improve reservoir inflow forecasting in Taiwan. Hydrol. Process. 28 (3), 1055–1070.
Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., Lizier, J.T., Vicente, R., 2013. Measuring information-transfer delays. PLoS One 8 (2), e55809.
Yin, H., Guo, Z., Zhang, X., Chen, J., Zhang, Y., 2021. Runoff predictions in ungauged basins using sequence-to-sequence models. J. Hydrol. 603, 126975.
Yousefi, M., Hajizadeh, A., Soltani, M.N., Hredzak, B., 2020. Predictive home energy management system with photovoltaic array, heat pump, and plug-in electric vehicle. IEEE Trans. Ind. Inf. 17 (1), 430–440.