Title
Machine learning and optimization applied to the assessment of flood hazard magnification by reservoir operation

Journal
Water Resources Management, 40(7)

ISSN
0920-4741

Authors
Zarei, Manizhe
Bozorg-Haddad, Omid
Loáiciga, Hugo A

Publication Date
2026-05-01

DOI
10.1007/s11269-026-04562-x

Peer reviewed

Scientific Reports
https://doi.org/10.1038/s41598-026-66479-0

Article in Press

Forensic engineering framework for evaluating meteorological forecasting models applied to flood control

Received: 24 June 2026

Accepted: 7 August 2026

Published online: 11 August 2026

Cite this article as: Zeinali M., Bozorg-Haddad O. & Loáiciga H.A. Forensic engineering framework for evaluating meteorological forecasting models applied to flood control. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-66479-0

Masoumeh Zeinali, Omid Bozorg-Haddad & Hugo A. Loáiciga

Forensic engineering framework for evaluating meteorological forecasting models applied to flood control

Masoumeh Zeinali[1]; Omid Bozorg-Haddad[2]; and Hugo A. Loáiciga[3]

[1] Ph.D. Candidate, Department of Irrigation & Reclamation Engineering, Faculty of Agricultural Engineering & Technology, College of Agriculture & Natural Resources, University of Tehran, Karaj, Tehran, Iran. E-mail: Masoumeh.Zeinalie@ut.ac.ir

[2] Distinguished Professor, Department of Irrigation & Reclamation Engineering, Faculty of Agricultural Engineering & Technology, College of Agriculture & Natural Resources, University of Tehran, Karaj, Tehran, Iran. E-mail: OBHaddad@ut.ac.ir

[3] Distinguished Professor, Department of Geography, University of California, Santa Barbara, California 93106; E-mail: hloaiciga@ucsb.edu.

Abstract

Accurate precipitation forecasting is fundamental to flood prediction, reservoir operation, and disaster risk reduction. However, systematic biases in numerical weather prediction (NWP) models often reduce the reliability of hydrological simulations. This study proposes a forensic engineering framework to evaluate the performance of two widely used NWP models, which are the Global Forecast System (GFS) and the Weather Research and Forecasting (WRF) model, for flood forecasting in the Karkheh watershed, southwestern Iran. Support Vector Regression (SVR) was employed to improve precipitation forecasts as a post-processing bias-correction technique, and the corrected forecasts were subsequently used as inputs to a rainfall–runoff model for reconstructing the 2019 flood event. The forecasting performance of the original and bias-corrected precipitation datasets was evaluated using statistical performance indicators, and their impacts on runoff simulation were assessed through a forensic engineering analysis of the 2019 flood. This work's results demonstrate that SVR-based bias correction substantially enhanced the forecasting accuracy of both the GFS and WRF models and significantly improved the agreement between predicted and observed runoff volumes. Bias correction applied to the first and second events reduced runoff prediction errors by 73.4, 136.6, and 78.9 million m^3 for the WRF model and by 74.6, 162.0, and 50.4 million m^3 for the GFS model with respect to the 1-, 2-, and 3-day lead forecasting times, respectively. These improvements resulted in more reliable runoff estimates and closer agreement with the observed runoff volumes. Overall, the proposed forensic engineering framework demonstrates that integrating SVR-based bias correction with numerical weather prediction models considerably improves precipitation and runoff forecasting accuracy. The proposed methodology provides a practical decision-support framework for flood forecasting, reservoir operation, and flood risk management in watersheds exposed to extreme hydrological events.

Keywords: Numerical prediction models; Forensic engineering; Flood management and control; Bias correction; Flood events.

1. Introduction

A survey of natural disasters indicates that floods have been the most frequent and destructive natural hazard worldwide since 2005 (UNDRR, 2019; World Bank and United Nations, 2010). Their increasing frequency and intensity, driven by climate change and land-use alterations, have significantly increased the challenges associated with flood risk management (Bezada, 2009; National Research Council, 2026). Consequently, improving the understanding of flood-generating processes and enhancing forecasting capabilities have become essential for reducing flood impacts and supporting effective water resources management (Tingsanchali, 2012).

Numerical weather prediction (NWP) models have become indispensable tools because they simulate atmospheric processes by numerically solving the governing physical equations. Widely used NWP models, such as the European Centre for Medium-Range Weather Forecasts (ECMWF), Icosahedral Nonhydrostatic (ICON), Global Forecast System (GFS), and the Weather Research and Forecasting (WRF) model, provide valuable meteorological information for flood forecasting over a wide range of spatial and temporal scales. Previous studies have consistently demonstrated the applicability of NWP models for flood forecasting in different hydro-climatic regions. For example, Li et al. (2017) evaluated regional numerical weather prediction models for flood warning systems in southern China, while Yu et al. (2016) employed ensemble NWP forecasts with a 10-km spatial resolution for flood prediction in Japan. Other studies have also assessed the capability of different NWP models to improve precipitation forecasting and hydrological prediction under diverse climatic conditions (Liguori and Rico-Ramirez, 2012; He et al., 2013; Hsiao et al., 2013; Yu et al., 2015; Gebremichael et al., 2022; Da Silva et al., 2022; Moghim and Takallou, 2023). Collectively, these studies confirm the considerable potential of NWP models for flood forecasting while also indicating that their predictive performance remains highly dependent on regional climatic conditions, forecast lead time, and systematic errors in precipitation prediction.

Several studies have evaluated the performance of numerical weather prediction (NWP) models for precipitation forecasting in Iran, particularly the Weather Research and Forecasting (WRF) model. Overall, these studies indicate that WRF can reasonably reproduce rainfall patterns; however, its performance is strongly influenced by topography, forecast lead time, and model resolution. For example, Taghavi et al. (2012) reported satisfactory simulation of precipitation patterns but lower accuracy over mountainous regions. Similarly, Arkian et al. (2013) found acceptable forecasting skill for rainfall events exceeding 10 mm, while Amini et al. (2013) showed that a finer spatial resolution (3 km) substantially improved forecasting accuracy compared with coarser resolutions. These findings suggest that, despite the capability of the WRF model for precipitation forecasting in Iran, further improvements are still required to reduce systematic forecasting errors before its application in hydrological and flood forecasting studies.

Despite the continuous development of numerical weather prediction (NWP) models, their forecasts are still affected by systematic errors and uncertainties (Kalnay, 2003). Therefore, post-processing and bias correction have become essential steps before applying meteorological forecasts to hydrological modelling (Ajaaj et al., 2015; Yuan et al., 2015; Schepen et al., 2017; Zhao et al., 2017). Various bias-correction techniques have been proposed, including Quantile Mapping (QM), Delta, Linear Scaling, Bayesian approaches, and machine learning methods such as Support Vector Regression (SVR) and Artificial Neural Networks (ANN). These techniques assume that the statistical relationship between forecasts and observations can be used to reduce systematic forecasting errors (Khajehei and Moradkhani, 2017; HEPEX, 2018).

Previous studies have demonstrated the effectiveness of post-processing techniques for reducing systematic errors in precipitation forecasts generated by numerical weather prediction (NWP) models under different climatic conditions. For example, Ngai et al. (2017), Amin Yavari et al. (2019), Lazoglu et al. (2020), Medina and Tian (2020), and Kolachian et al. (2019) reported that post-processing methods generally improved the accuracy of precipitation forecasts, although their performance varied depending on the correction technique, meteorological variable, and study region. Despite these advances forecasting uncertainties remain, particularly for extreme precipitation events and their subsequent hydrological applications. Therefore, bias correction has become an essential step before applying NWP forecasts to operational flood forecasting and water resources management. Overall, these findings highlight that, although bias correction is an essential component of modern numerical weather forecasting, identifying robust post-processing approaches capable of further improving precipitation forecasts and enhancing the reliability of hydrological simulations remains an important research challenge.

Forensic hydrology is a branch of forensic engineering that aims at identifying the causes of failures and extreme events in water systems through systematic analysis of hydrological, meteorological, and operational data (Loáiciga, 2001; Witte et al., 2019; Zarei et al., 2023). Forensic hydrology has been increasingly applied to investigate flood events, reservoir operation, droughts, sediment transport, groundwater depletion, and other water-related hazards (Zarei et al., 2021). Previous studies have demonstrated the value of forensic approaches for reconstructing flood events and evaluating water management practices. For example, Aparicio et al. (2009), Branstret et al. (2018), Delpasand et al. (2021), and Zarei et al. (2023) successfully applied forensic engineering concepts to analyze flood events and assess reservoir operation under extreme hydrological conditions. Collectively, these studies highlight the capability of forensic hydrology to improve the understanding of flood mechanisms and support more effective water resources management.

There are many previous published studies that evaluate water resources systems in present and future periods (Beygi et al. 2014; Bozorg-Haddad and Mariño 2011; Bozorg-Haddad et al. 2007, 2009; Fallah-Mehdipour et al. 2011, 2013; Karimi-Hosseini et al. 2011; Orouji et al. 2014; Sabbaghpour et al. 2012; Shokri et al. 2013; Soltanjalili et al. 2011). The present study aims to investigate the past conditions of a flood-control system.

The 2019 Iranian flood was one of the most severe hydrological disasters in recent decades, causing large human, economic, and environmental losses across large parts of the country.

Prolonged and intense precipitation in the Karkheh River Basin generated exceptionally high runoff, highlighting the critical role of accurate precipitation forecasting for flood management. The close relationship between extreme precipitation and flood generation makes this event a representative case study for evaluating the performance of bias-corrected numerical weather prediction (NWP) models. Forensic hydrology has been widely applied to flood reconstruction and reservoir operation; howeverits integration with NWP models and machine learning-based bias correction for flood forecasting remains limited. This study addresses this research gap and proposes an integrated forensic engineering framework that combines GFS and WRF precipitation forecasts with Support Vector Regression (SVR)-based bias correction and rainfall–runoff modelling to investigate the 2019 Karkheh flood. The proposed framework evaluates the impact of bias correction on hydrological prediction and flood management by comparing simulated and observed flood volumes under different forecasting scenarios. This work's findings demonstrate how improvements in precipitation forecasting enhance the reliability of hydrological prediction and support more informed reservoir operation and flood risk management. Furthermore, the proposed framework provides a transferable methodology for evaluating bias-corrected NWP models in regions exposed to extreme hydrological events.

2. Methods and materials

Iran features a dry and semi-arid climate; yet it experiences occasional destructive floods, one of which occurred in 2019. This flood affected Iran's western and southwestern regions during heavy rainfall. The first rainfall event occurred during 24–26 March 2019, whereas the second occurred during 31 March–2 April 2019 and, in total, an average of 190 mm of rain fell on the Karkheh catchment area. As a result, five provinces, 200 cities, and 4300 villages were affected, and nearly 80 people died (SRCIF, 2019a).

Accurate and timely precipitation forecasting is one of the fundamental components of effective flood management and forms the basis of the present study, which focuses on the 2019 Iranian flood. The proposed methodology first evaluates the performance of GFS and WRF precipitation forecasts before and after Support Vector Regression (SVR)-based bias correction. The corrected forecasts are then incorporated into rainfall–runoff modelling to assess how improvements in precipitation forecasting influence runoff simulation and forensic engineering analysis. Figure 1 provides an overview of the proposed methodological framework, which comprises five sequential stages: data acquisition, data preprocessing, bias correction, model performance evaluation, and forensic engineering analysis. This workflow establishes the overall structure of the study, while the individual methodological steps are described in detail in the following sections.

2.1. The study area

The study area is the Karkheh Basin, located in southwestern Iran. This basin has an area of 51643 square kilometers, of which 55.5% is mountainous. The Karkheh basin is located between coordinates of 46º06´ to 49º10´ east longitude and 30º58´ to 35º04´ north latitude. Figure 2 depicts the study area along with the location of synoptic meteorological stations.

The Karkheh Basin is one of the most important river basins in Iran, playing a vital role in water supply, agriculture, hydropower generation, and flood control. The basin is characterized by complex topography, considerable spatial variability in precipitation, and the presence of the Karkheh Dam, which significantly influences flood management. During the extreme flood event of 2019, the basin experienced widespread flooding that caused substantial economic and environmental damage. This event, highlights the importance of improving precipitation forecasting and flood prediction in this region.

2.2. Observational data

Daily rainfall data were collected from 22 synoptic stations operated by the Iran Meteorological Organization (IMO). The characteristics of the stations are presented in Table 1.

2.3. Data related to weather-forecasting models

The GFS and WRF models were used to forecast rainfall in the studied area. These two weather forecasting models are described below.

2.4. The GFS numerical meteorological model

The GFS Global Forecast System is a spectral global weather forecasting model with advanced dynamics and physics (Kanamitsu, 1989; Kalnay et al., 1990; Kanamitsu et al., 1991; Moorthi et al., 2001; Durai et al., 2010; Saha et al. et al. 2010). It is a well-known global meteorological model. It conforms to a dynamical framework known as the Earth System Modeling Framework (ESMF). Its code was rewritten to have many options for updated dynamics and physics (Sridevi et al., 2020). This model is developed and runs four times a day (00:00, 06:00, 12:00, 18:00 UTC) in real-time operational mode at the US National Center for Environmental Prediction (NCEP). Its outputs are available to the user community through an FTP server. Its outputs are updated every 6 hours and forecasts up to 16 days in advance. The GFS model consists of 4 separate models of the atmosphere, ocean, surface soil, and sea ice, which are run simultaneously. The minimal horizontal resolution of this model network is 28 km, and it covers up to 8 km of the vertical column of the atmosphere. A wide range of atmospheric and surface variables (e.g., soil moisture, temperature, wind, precipitation) from the GFS's simulations are available to users. The new version of this model has a horizontal resolution of 13 km (0.13 degrees). The GFS features an annual update that maintains its current with actual global weather phenomena (Khajehei and Moradkhani, 2017). Many operational weather forecasting agencies in the world use the GFS model output for real-time weather forecasting.

2.5. The WRF numerical meteorological model

The WRF model performs numerical simulations of weather and atmospheric conditions intended for research and operational applications. The WRF model was developed by various agencies, such as the United States National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Center of America (NOAA), the Center for Environmental Prediction (NCEP), and the Meteorological Center of the Air Force Weather Agency (AFWA) (Powers et al., 2017). The spatial range of application of this model varies from several meters to thousands of kilometers (Skamarock et al., 2008).

The software structure of the WRF model is shown in Figure 3. This structure includes parts of physics, chemistry, dynamic cores, and programs for initiation and data guarding. The model includes two different dynamic cores: the advanced research dynamic core (ARW) developed by NCAR and the dynamic core of the non-hydrostatic mesoscale model (NMM) developed by NCEP. In addition to the dynamic core, physical and chemical processes are simulated separately in the WRF model. The physics component of the model is dedicated to the parameterization of the physical processes in the atmosphere on grid and sub-grid scales. This model component includes various processes related to cloud microphysics, cumulus convection, surface physics, planetary boundary layer physics, and long-wavelength and short-wavelength atmospheric radiation (Skamarock et al., 2008). Figure 3 displays the WRF model structure.

2.5.1 Schemes available in the WRF model

Radiation parameterization schemes, cumulus cloud parameterization, surface layer, surface models, planetary boundary layer, and microphysics schemes are used in the parameterization of physical variables in the WRF model. The WRF model has three configurations, each implemented based on different physical schemes. Table 2 presents the microphysics schemas of the WRF model used in this study. Figure 4 displays the feedbacks that exist among different physical schemes of the WRF model. The latter model is implemented in two domains: the first domain has a horizontal resolution of 28 km, and the second domain has a horizontal resolution of 8 km. This study used the data of the second domain with a spatial resolution of 8 km.

Observed rainfall data from the 22 synoptic stations were used to evaluate the WRF simulations. Rainfall forecasts corresponding to the station locations were extracted from the WRF outputs and compared with the observations. Table 3 lists the general characteristics of the numerical models.

2.6. Statistical evaluation indicators

The rainfall forecasts from the weather models were compared with the observational data gathered at the meteorological stations to evaluate the predictive skill of the numerical meteorological models. A set of statistical evaluation indicators of goodness-of-fit were applied in the evaluation, which are described below.

The statistical criteria implemented in this paper are the correlation coefficient [Equation (1)], the root mean squared error [Equation (2)], the bias percentage [Equation (3)], and the Nash-Sutcliffe criterion [Equation (4)]:

Correlation coefficient:

$$R = \frac{\sum_{i=1}^{N}(O_i - \overline{O})(F_i - \overline{F})}{\sqrt{\sum(O_i - \overline{O})^2 \sum(F_i - \overline{F})^2}} \tag{1}$$

The root mean squared error:

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(F_i - O_i)^2}{N}} \tag{2}$$

The Bias Percentage:

$$pBias = \frac{\sum_{i=1}^{N}(F_i - O_i)}{\sum_{i=1}^{N} O_i} \times 100 \tag{3}$$

The Nash-Sutcliffe criterion:

$$NSE = 1 - \frac{\sum_{i=1}^{N}(O_i - F_i)^2}{\sum_{i=1}^{N}(O_i - \overline{F})^2} \tag{4}$$

where in Equations (1) to (4), F_i = the forecasted value of the rainfall variable; O_i = the corresponding observed value; $\overline{F}$ Average forecasted values of rainfall, and N = the number of observations.

2.7. Post-processing of numerical forecasts of precipitation in the basin

Raw precipitation forecasts generated by numerical weather prediction (NWP) models generally contain systematic errors and biases (Wu et al., 2014). Bias is the systematic difference between forecasted and observed data. Therefore, bias correction methods have been considered in many studies to post-process the results of prediction models and remove their bias (Tao et al., 2014). In this study, Support Vector Regression (SVR) was adopted as a statistical post-processing method to correct systematic biases in the GFS and WRF precipitation forecasts.

For this purpose, the rainfall forecasts from GFS and WRF models from 13 March 2019 through 11 September 2022 and the forecast horizon of 1 to 3 days in the future were extracted at the investigated stations' locations. The paired forecast–observation datasets were subsequently used to develop SVR models for bias correction at each station.

2.7.1 Support Vector Regression (SVR)

The support vector machine (SVM) was developed by Vapnik (1995). SVM is used for supervised learning methods in data analysis and pattern recognition (Araghinejad, 2014).

The decision-making space is divided into hyper-pages. The best separation is done so that the set of vectors belonging to each class is separated without error, and the closest point of each class has the most significant possible distance to the separation plane (Gunn, 1998). The second feature reduces the misclassification error when new information is added and is the primary difference between support vector machines and other classification methods (Araghinejad, 2014). This research used the support vector regression (SVR) method to identify the relationship between forecasts and observational data. For this purpose, paired forecast and observed precipitation data at each station were entered in the SVR modeling stage. The data were divided into calibration and validation data sets, with an 80% to 20% ratio. The SVR model was fitted on these data sets. For this purpose, 50 SVR models were fitted in each station, and the best model among the 50 models was selected based on the NSE criterion. The selected SVR model was then used to correct the precipitation forecasts at the corresponding station. This work was repeated for each station separately for the WRF and GFS models and for each forecast horizon. Due to the improvement of the performance of this method by considering more data the entire daily time series of data were used. The chosen kernel function is the RBF function, and the parameters are determined by trial and error. The SvmRadial method with train control based on the cross-validation method was implemented for developing the SVR model in the R platform. Lastly, the model was run 50 times; 80% of the data were considered for training, and 20% were considered for testing.

This calibration and validation procedure was performed independently for each meteorological station, each forecast horizon (1-, 2-, and 3-day ahead), and for both the GFS and WRF forecasting models. The optimal SVR model was selected according to the highest NSE value obtained during the validation stage and subsequently used for bias correction of the corresponding precipitation forecasts.

2.8. Forensic engineering analysis of flood events

Forensic engineering is used to determine the contribution of different factors to infrastructure failures and flood damages, including the role of available meteorological forecasts in flood management and reservoir operation. This study applied a forensic engineering framework to evaluate the performance of the GFS and WRF rainfall forecasting models before and after bias correction using the Support Vector Regression (SVR) method. The objective was to investigate how forecast uncertainty propagates into runoff estimation during the March 2019 flood events in the Karkheh Basin and to assess the potential benefits of bias correction for future flood management. Rainfall–runoff modelling was performed after evaluating the precipitation forecasts in order to convert the observed and forecasted rainfall into runoff within each sub-basin. The post-processing of raw numerical weather prediction outputs through SVR-based bias correction was subsequently assessed in terms of its influence on runoff prediction accuracy. The forensic engineering analysis was conducted for the two major flood events that occurred during 24–26 March and 31 March–2 April 2019 (SRCIF, 2019b). The Karkheh Basin was divided into nine sub-basins according to the 2019 Special Flood Committee hydrology and water resources report (Figure 5). The characteristics and runoff coefficients of the investigated sub-basins are presented in Table 4. The runoff coefficients were assumed to remain constant throughout the study period (24 March–2 April 2019) and ranged from 0.18 in the Seimareh sub-basin to 0.69 in the Doab Kashkan sub-basin (SRCIF, 2019b).

The runoff coefficients adopted in this study were obtained from the hydrological assessment reported by the 2019 Special Flood Committee (SRCIF, 2019b) and represent event-based values estimated specifically for the March–April 2019 flood events. Therefore, these coefficients were kept constant throughout the analysis to ensure consistency in the comparison of runoff estimates derived from the raw and bias-corrected precipitation forecasts.

For each forecast horizon (1, 2, and 3 days), the observed precipitation and the forecasts generated by the GFS and WRF models, both before and after bias correction, were spatially averaged over each sub-basin using the Thiessen polygon method. The spatially averaged precipitation was then converted into runoff using the corresponding runoff coefficient for each sub-basin. Cumulative runoff volumes were estimated separately for each sub-basin and forecast horizon, and the percentage difference between the observed and simulated runoff volumes was calculated. These percentage differences were used in the forensic engineering framework to quantify the effect of precipitation forecast uncertainty on runoff prediction and to evaluate the effectiveness of bias correction in improving flood forecasting performance. Therefore, improved precipitation forecasts reduce uncertainty in runoff estimation, thereby providing more reliable information for reservoir operation, flood warning, and emergency management. Consequently, more accurate runoff predictions can support timely management decisions and contribute to reducing the potential impacts of flood events.

3. Results and discussion

3.1. Evaluation of numerical rainfall forecasting models before and after bias correction

Raw precipitation forecasts generated by numerical weather prediction models contain systematic errors that reduce their accuracy. The support vector machine method was applied to correct the bias of the forecasts. The performance of the WRF and GFS models was evaluated with data from March 13 2019 through September 11 2022. The forecast horizon of the WRF model is up to the next 3 days. For this reason, the inter-model comparison is limited to a 3-day forecasting horizon. The forecast horizon of the GFS model is up to the next 16 days.

Table 5 summarizes the correlation coefficient (R) obtained for the GFS and WRF precipitation forecasts over lead times of one to three days before and after SVR-based bias correction. Overall, bias correction consistently improved the correlation between forecasted and observed precipitation at all stations and for all forecasting horizons, demonstrating the effectiveness of the proposed post-processing approach. The R values before bias correction ranged from 0.23 to 0.91, whereas after correction they increased to a range of 0.44–0.98, indicating a substantial enhancement in forecast reliability. The magnitude of improvement varied among stations and forecasting horizons. For the GFS model, the largest increases were observed at the Bostan station for the 1- and 2-day forecasts, where the R coefficient increased from 0.34 to 0.63 and from 0.23 to 0.53, respectively. The greatest improvement corresponding the 3-day forecast occurred at the Kermanshah station, where the R value increased from 0.70 to 0.90. Similarly, the WRF model showed marked improvements after bias correction, particularly at the Dereh Shahr, Harsin, and Ravansar stations, where the R coefficient increased from 0.68 to 0.90, from 0.56 to 0.87, and from 0.48 to 0.84, respectively. A comparison between the two numerical weather prediction models indicates that SVR-based bias correction enhanced the performance of both GFS and WRF forecasts, although the magnitude of improvement depended on the station and forecast lead time. These results suggest that the correction method effectively reduced systematic forecast errors and strengthened the agreement between simulated and observed precipitation. The consistent improvement across all stations is particularly important for flood forecasting because more reliable precipitation estimates provide a stronger basis for rainfall–runoff simulations and subsequent forensic evaluation of flood events.

Table 6 presents the Root Mean Square Error (RMSE) values for the GFS and WRF precipitation forecasts over lead times of one to three days before and after SVR-based bias correction. Overall, bias correction reduced RMSE across most stations and forecasting horizons, with the overall range decreasing from 5.85–15.25 mm before correction to 3.30–13.21 mm after correction, indicating a clear improvement in forecast accuracy. For the GFS model, the largest reductions in RMSE were observed at the Sarableh, Noorabad Lorestan, and Bostan stations corresponding to the 1-, 2-, and 3-day forecasts, respectively. Similarly, the WRF model showed its greatest improvements at the Dareh Shahr station for the 1-day forecast and at the Sarableh station for the 2- and 3-day forecasts. Overall, the results indicate that the SVR-based bias correction effectively reduced precipitation forecast errors for both numerical weather prediction models, although the magnitude of improvement varied among stations and forecast lead times. These improvements provide more reliable precipitation inputs for rainfall–runoff modelling and subsequent flood forecasting.

Table 7 summarizes the percent bias (pBias) values for the GFS and WRF precipitation forecasts over lead times of one to three days before and after SVR-based bias correction. Since the optimal pBias value is zero, the results indicate that bias correction substantially reduced systematic errors in both numerical weather prediction models across most stations and forecasting horizons. Overall, the pBias range decreased from −42.5% to 42.5% before correction to −23% to 14% after correction, demonstrating a marked reduction in forecast bias. The largest improvements for the GFS model were observed at the Sonqor station for the 1- and 2-day forecasts and at the Asadabad station for the 3-day forecast, whereas the WRF model exhibited its greatest improvements at the Dareh Shahr station for the 1- and 2-day forecasts and at the Sarableh station for the 3-day forecast. These results demonstrate that the SVR-based bias correction effectively reduced systematic overestimation and underestimation of precipitation forecasts, thereby providing more reliable inputs for hydrological modelling and flood forecasting.

Table 8 presents the Nash–Sutcliffe Efficiency (NSE) values for the GFS and WRF precipitation forecasts over lead times of one to three days before and after SVR-based bias correction. Since an NSE value closer to one indicates better model performance, the results demonstrate that bias correction substantially improved the predictive skill of both numerical weather prediction models across most stations and forecasting horizons. Overall, the NSE values increased from a range of −1.85 to 0.77 before correction to 0.14–0.95 after correction, indicating a considerable enhancement in forecast performance. The largest improvements for the GFS model were observed at the Sonqor station for the 1-day forecast and at the Bostan station for the 2- and 3-day forecasts. Similarly, the WRF model showed its greatest improvements at the Harsin station for the 1-day forecast and at the Bostan station for the 2- and 3-day forecasts. Overall, the results indicate that SVR-based bias correction consistently enhanced the agreement between forecasted and observed precipitation, providing more reliable inputs for rainfall–runoff modelling and flood forecasting.

The observed improvements following SVR-based bias correction can be explained by the ability of machine learning algorithms to capture nonlinear relationships between numerical weather prediction outputs and ground observations. Raw precipitation forecasts produced by numerical weather prediction models are affected by systematic errors arising from simplified representations of atmospheric processes, uncertainties in initial conditions, and spatial resolution limitations. , The SVR model learned these systematic discrepancies from historical observations and effectively reduced forecast bias and improved the agreement between simulated and observed precipitation across different forecasting horizons. The larger improvements observed for longer lead times indicate that statistical post-processing becomes increasingly valuable as the uncertainty of numerical weather prediction models increases with forecasting horizon.

3.2. Examining the performance of models during flood events

The two flood events of 2019 occurred on March 24-26 and March 31-April 2, at which time the results of the GFS global model and the WRF regional model were available. This section deals with the issue of how much the forecasting of these two models made without bias corrections can be modified and applied in the management and forecasting of floods. Therefore, for the two events in question the cumulative rainfall forecast values associated with the 3 days of each event were drawn as bar charts for the 1-, 2-, and 3-days forecast horizons at the event location. Data from rain gauge stations were compared to the observed values (see Figure 6 and Figure 7).

Figures 6 and 7 compare the observed precipitation with the forecasts produced by the raw and bias-corrected GFS and WRF models during the first and second flood events, respectively. Overall, both numerical weather prediction models successfully detected the occurrence and temporal evolution of the rainfall events. On the other had, before bias correction, both models systematically underestimated precipitation intensity and were unable to reproduce the observed peak rainfall accurately, particularly during extreme precipitation periods. For the first flood event (Figure 6), the benefits of SVR-based bias correction are clearly evident. At the Noorabad Lorestan station, where the observed peak precipitation reached 141.2 mm, the raw GFS forecasts for the 1-, 2-, and 3-day lead times were 93.43, 70.65, and 55.12 mm, respectively, while the corresponding WRF forecasts were 79.00, 90.12, and 74.65 mm. The estimated peak precipitation increased substantially after bias correction and became much closer to the observations. The GFS-BC model provided the most accurate estimates for the 1- and 2-day forecasts, predicting peak rainfall of 140.9 and 134.5 mm, respectively, whereas the WRF-BC model achieved the best performance for the 3-day forecast with an estimated peak of 117.5 mm. These results indicate that bias correction considerably reduced the underestimation of heavy rainfall, although forecast accuracy generally decreased as the prediction horizon increased.

A similar pattern was observed during the second flood event (Figure 7). No precipitation data were available at the Poldokhtar station on the final day of the event, and for this reason this station was excluded from the analysis. At the Dareh Shahr station the observed maximum precipitation was 191 mm and both raw models substantially underestimated the rainfall peak. The uncorrected GFS forecasts for the 1-, 2-, and 3-day lead times were 96.7, 64.15, and 62.6 mm, respectively, whereas the corresponding WRF forecasts were 66.21, 69.4, and 35.12 mm. The GFS-BC model produced peak estimates of 118.7, 91.5, and 104.5 mm for the three forecasting horizons following bias correction, thus achieving a clear improvement over the raw forecasts. Similar improvements were observed for the WRF-BC model at the Noorabad Lorestan, Kangavar, and Nahavand stations, although peak precipitation remained underestimated during the most extreme rainfall conditions. A comparison of the forecasting scenarios demonstrates that SVR-based bias correction consistently improved the performance of both GFS and WRF models by reducing systematic underestimation of precipitation peaks. The relative performance of the two models varied among stations and forecasting horizons, and the WRF-BC model generally provided the most reliable precipitation estimates during the two flood events. Nevertheless, the remaining underestimation of the highest rainfall peaks, particularly for the 3-day forecasts, indicates that uncertainty increases with lead time and remains a challenge for extreme precipitation forecasting. Overall, the improved precipitation estimates obtained after bias correction provide more reliable inputs for rainfall–runoff simulations and strengthen the application of the proposed forensic engineering framework for flood forecasting and flood management.

3.3. Forensic engineering investigation of the flood events

The evaluation presented in the previous section demonstrated that bias correction is necessary to improve the precipitation forecasts produced by the numerical weather prediction models. This section evaluates the effect of correcting the precipitation forecasts bias on the volume of generated runoff. This evaluation assesses the value of the bias correction for predicting flood events in the study area with a forensic engineering approach (i.e., in a hydrological forensic engineering sense). The observed and forecasted precipitation values were converted to runoff in each sub-basin of the study area based on Figure 5 and the runoff coefficients listed in Table 4. The observed and forecasted precipitation was spatially averaged for each prediction horizon using the Thiessen polygon method in each sub-basin. Next, the runoff was estimated by applying the runoff coefficient to the forecasted precipitation in each sub-basin. The percentage difference between the observed and forecasted runoff was obtained for each sub-basin.

3.3.1. The first flood event

Tables 9–11 summarize the percentage differences between the observed cumulative runoff and the runoff predicted using the GFS and WRF models before and after SVR-based bias correction for the 1-, 2-, and 3-day forecasting horizons. Overall, bias correction consistently reduced runoff prediction errors across all sub-basins. This demonstrates that improvements in precipitation forecasts were successfully propagated into the hydrological simulations.

The 1-day forecasts (Table 9) indicate the bias-corrected GFS model generally produced the smallest percentage differences from the observed runoff. The improvement was particularly evident in the Polchehr, Holeilan, Doab Kashkan, Cham Anjir, Seimareh, Jelogir, and Paypol sub-basins, indicating that the GFS forecasts benefited more from bias correction than the WRF forecasts at the shortest lead time. The 2-day forecasts (Table 10) demonstrate the the relative model performance changed. The WRF-BC model showed greater reductions in runoff prediction errors across most sub-basins, i.e., Poldokhtar Kashkan, Afarineh Kashkan, Doab Kashkan, Jelogir, and Paypol, while both models exhibited comparable performance in the Cham Anjir and Seimareh sub-basins. These results suggest that the regional WRF model provided more reliable runoff estimates than the GFS model after bias correction for intermediate forecasting horizons. A similar trend was observed for the 3-day forecasts (Table 11). Bias correction generally improved both forecasting models. Specifically, WRF-BC model achieved better agreement with the observed runoff in almost all sub-basins except Cham Anjir. The improvements obtained for the GFS and WRF models were generally comparable; however, the WRF-BC model produced more accurate runoff estimates over most of the study area. Considering the Paypol sub-basin, which represents the entire Karkheh watershed (45,616.4 km²), the forensic engineering analysis demonstrates a clear dependence of model performance on forecasting horizon. The GFS-BC model yielded the smallest runoff error for the 1-day forecast (−13%), whereas the WRF-BC model achieved the best performance for the 2-day (−25%) and 3-day (−21%) forecasts. Despite these improvements, all forecasting scenarios continued to underestimate the observed flood runoff, which signals that uncertainties associated with extreme rainfall events remain even after bias correction. Nevertheless, the consistent reduction in runoff prediction errors confirms that the proposed SVR-based bias correction substantially enhances the reliability of rainfall–runoff simulations and strengthens the application of the proposed forensic engineering framework for flood forecasting and flood management.

Figure 8 compares the observed cumulative runoff during the first flood event with the runoff simulated by the GFS and WRF models before and after SVR-based bias correction for the 1-, 2-, and 3-day forecasting horizons. Overall, all forecasting scenarios underestimated the observed runoff in every sub-basin. However, bias correction consistently reduced the runoff prediction error, resulting in simulated runoff volumes that were substantially closer to the observations, in agreement with the results presented in Tables 9–11.

The relative performance of the forecasting models varied with the prediction horizon. For the 1-day forecasts, the GFS-BC model produced the most accurate runoff estimates across all sub-basins, followed by the WRF-BC model, whereas the raw GFS and WRF models showed larger deviations from the observed runoff. For the 2-day forecasts, the WRF-BC model generally outperformed the other forecasting scenarios in most sub-basins, while the GFS-BC model ranked second. The raw GFS model performed slightly better than the raw WRF model in the Poldokhtar Kashkan, Afarineh Kashkan, and Cham Anjir sub-basins, while the opposite trend was observed in the Polchehr, Holeilan, and Seimareh sub-basins, and both models exhibited similar performance elsewhere. The WRF-BC model's 3-day forecasts provided the closest agreement with the observed runoff in most sub-basins, i.e., in the Poldokhtar Kashkan, Afarineh Kashkan, Doab Kashkan, Jelogir, and Paypol sub-basins, indicating that the regional WRF model benefited more from bias correction for longer forecasting horizons. This finding is further illustrated in the Paypol sub-basin, which represents the entire study area (45,616.4 km^2). The observed cumulative runoff during the first flood event was 1258.9 million m^3, while the GFS-BC model estimated 1097.4 million m^3 for the 1-day forecast. For the 2- and 3-day forecasts, the WRF-BC model produced runoff estimates of 941.0 and 990.5 million m^3, respectively. Overall, these results demonstrate that bias correction substantially improved runoff prediction accuracy and that its effectiveness depended on both the forecasting horizon and the numerical weather prediction model, thereby strengthening the reliability of the proposed forensic engineering framework for flood forecasting.

3.3.2. The second flood event

Tables 12–14 present the percentage differences between the observed cumulative runoff and the runoff simulated by the GFS and WRF models before and after SVR-based bias correction for the second flood event, corresponding to the 1-, 2-, and 3-day forecasting horizons. Overall, bias correction consistently reduced runoff prediction errors across all sub-basins, demonstrating that the improvements achieved in precipitation forecasts were successfully transferred to the hydrological simulations. The 1-day forecasts (Table 12) show that the WRF-BC model generally provided more accurate runoff estimates than the GFS-BC model, with the exception of the Cham Anjir sub-basin. Bias correction improved both forecasting models, with the reduction in percentage error being more pronounced for the GFS model in the Polchehr, Holeilan, Doab Kashkan, Cham Anjir, Seimareh, and Jelogir sub-basins. This demonstrates that that the GFS forecasts benefited more from bias correction at the shortest lead time. The 2-day forecasts (Table 13) indicate the WRF-BC model consistently outperformed the GFS-BC model across all sub-basins. The largest improvements following bias correction were observed in the Holeilan, Poldokhtar Kashkan, Afarineh Kashkan, Doab Kashkan, Seimareh, Jelogir, and Paypol sub-basins and demonstrate that the regional WRF model produced more reliable runoff estimates for intermediate forecasting horizons. A similar pattern was observed for the 3-day forecasts (Table 14). Both forecasting models benefited from bias correction. The WRF-BC model achieved better agreement with the observed runoff in all sub-basins except Seimareh. The greatest improvements were obtained in the Polchehr, Poldokhtar Kashkan, Afarineh Kashkan, Doab Kashkan, and Cham Anjir sub-basins and confirm the effectiveness of bias correction for longer forecasting horizons. The Paypol sub-basin, which represents the entire Karkheh watershed (45,616.4 km^2), provides an overall measure of model performance at the basin scale. The WRF-BC model produced the smallest runoff differences for the 1-, 2-, and 3-day forecasts with percentage errors of −9%, −1%, and −13%, respectively. Despite these improvements all forecasting scenarios continued to underestimate the observed runoff. This demonstrates that uncertainty remains in predicting extreme flood events. Nevertheless, the consistent reduction in runoff prediction errors demonstrates that integrating SVR-based bias correction with numerical weather prediction models substantially enhances the reliability of rainfall–runoff simulations and strengthens the proposed forensic engineering framework for operational flood forecasting and flood management. Figure 9 compares the observed cumulative runoff during the second flood event with the runoff simulated by the GFS and WRF models before and after SVR-based bias correction for the 1-, 2-, and 3-day forecasting horizons. Overall, all forecasting scenarios underestimated the observed runoff across most sub-basins. However, applying SVR-based bias correction consistently reduced the runoff prediction errors and produce runoff estimates that were substantially closer to the observations that are in agreement with the results presented in Tables 12–14. The relative performance of the forecasting models varied with the prediction horizon. The 1-day forecasts show that the WRF-BC model generally provided the most accurate runoff estimates in most sub-basins, followed by the GFS-BC model, whereas the raw WRF and GFS models exhibited larger deviations from the observed runoff. The raw GFS and WRF models showed comparable performance across most sub-basins and prove that the primary improvement resulted from the bias-correction procedure rather than from differences between the original numerical weather prediction models. The 2-day forecasts demonstrate the WRF-BC model maintained the best overall performance, while the GFS-BC model ranked second in this respect. The raw GFS model slightly outperformed the raw WRF model in most sub-basins. However, the WRF model produced more accurate runoff estimates in the Polchehr sub-basin which highlights the spatial variability in model performance across the watershedee The 3-day forecasts show the superiority of the

WRF-BC model became more evident as it provided the closest agreement with the observed runoff in all sub-basins. The ranking of model performance from most to least accurate was WRF-BC, WRF, GFS-BC, and GFS. This shows the regional WRF model retained higher predictive skill than the global GFS model at longer lead times following bias correction. The Paypol sub-basin, representing the entire Karkheh watershed (45,616.14 km^2) further illustrates these findings. The observed cumulative runoff during the second flood event reached 1478.6 million m^3, whereas the WRF-BC model estimated runoff volumes of 1341.3, 1282.9, and 1141.7 million m^3 for the 1-, 2-, and 3-day forecasting horizons, respectively. The runoff remained underestimated, yet, the bias-corrected WRF model substantially reduced the prediction error and demonstrated the highest reliability among all forecasting scenarios. These findings confirm that integrating SVR-based bias correction with numerical weather prediction models significantly improves runoff simulation and enhances the practical applicability of the proposed forensic engineering framework for flood forecasting and reservoir operation.

The propagation of improved precipitation forecasts into rainfall–runoff simulations demonstrates the strong sensitivity of hydrological predictions to precipitation uncertainty. Runoff generation is a nonlinear process in which relatively small improvements in precipitation estimates can produce substantial reductions in runoff prediction errors, particularly during extreme flood events. All forecasting scenarios continued to underestimate the observed runoff but the bias-corrected forecasts consistently reduced these errors across nearly all sub-basins. This finding highlights the importance of incorporating statistical post-processing into operational flood forecasting systems before hydrological modelling is undertaken.

The results obtained in this study are consistent with previous investigations reporting that statistical post-processing methods improve the performance of numerical weather prediction models. Similar improvements following bias correction have been reported by Ngai et al. (2017), Amin Yavari et al. (2019), Medina and Tian (2020), and Kolachian et al. (2019), who demonstrated that statistical correction techniques substantially enhance precipitation forecasting accuracy. However, unlike previous studies that primarily evaluated meteorological performance using statistical indices, the present study further quantifies how improvements in precipitation forecasts propagate through rainfall–runoff simulations within a forensic engineering framework. This integrated evaluation provides a more comprehensive assessment of forecasting performance by linking meteorological uncertainty directly to flood prediction accuracy and flood management. This paper’s results show the advantages and of the application of bias-corrected numerical weather prediction models.

4. Concluding remarks

This study proposed an integrated forensic engineering framework to evaluate the performance of the GFS and WRF numerical weather prediction models before and after Support Vector Regression (SVR)-based bias correction for flood forecasting in the Karkheh watershed, southwestern Iran. Rainfall forecasts were evaluated against observations from rain gauge stations, and the bias-corrected precipitation forecasts were subsequently used to estimate runoff in the watershed sub-basins. The results demonstrated that applying SVR-based bias correction substantially improved the performance of both numerical weather prediction models and enhanced the accuracy of rainfall–runoff simulations.

The evaluation based on the Nash–Sutcliffe Efficiency (NSE) index confirmed the effectiveness of the proposed bias-correction approach. For the GFS model, the average NSE values for the 1-, 2-, and 3-day forecasting horizons increased from 0.38, 0.38, and 0.26 before bias correction to 0.72, 0.63, and 0.55, respectively, after bias correction. Similarly, for the WRF model, the average

NSE values improved from 0.37, 0.43, and 0.30 to 0.70, 0.67, and 0.59, respectively. These results demonstrate that raw numerical weather prediction outputs contain systematic errors and should be statistically post-processed before being used for operational flood forecasting.

The forensic engineering analysis further showed that the improvements in precipitation forecasting were successfully propagated into runoff simulations. The observed runoff in the Paypol sub-basin (45,616.4 km^2), which represents the entire study area, reached 1258.9 million m^3 during the first flood event. The GFS-BC model produced the most accurate runoff estimates for the 1-day forecasting horizon, whereas the WRF-BC model achieved the best performance for the 2- and 3-day forecasting horizons. The reduction in runoff prediction error between the raw GFS and GFS-BC forecasts for the 1-day horizon reached 73.4 million m^3, while the corresponding improvements achieved by the WRF-BC model for the 2- and 3-day horizons were 136.6 and 78.9 million m^3, respectively.

The observed runoff in the Paypol sub-basin increased to 1478.6 million m^3 during the second flood event. The WRF-BC model consistently outperformed the other forecasting scenarios for the 1-, 2-, and 3-day forecasting horizons associated with this event. Bias correction reduced runoff prediction errors by 74.6, 162.0, and 50.4 million m^3 for the 1-, 2-, and 3-day forecast horizons, respectively, compared with the raw WRF forecasts. These values represent a reduction in prediction error relative to the raw forecasts rather than the total flood runoff volumes. These results demonstrate the considerable benefit of statistical post-processing for operational flood forecasting.

These findings are consistent with the 2019 Special Flood Committee meteorological report (SRCIF, 2019a), which concluded that the underestimation of precipitation forecasts adversely affected flood management during the 2019 Iranian flood. The present study demonstrates that applying SVR-based bias correction substantially improves precipitation and runoff forecasting. This reduces one of the key sources of uncertainty that contributed to flood management failures during this event. Overall, this study demonstrates that integrating numerical weather prediction models, machine learning-based bias correction, and forensic engineering provides a comprehensive framework for improving flood forecasting accuracy. The proposed methodology has the potential to support operational reservoir operation, flood early-warning systems beyond the 2019 Karkheh flood, and water resources management in other regions exposed to extreme hydrological events. Future studies may investigate the applicability of the proposed framework using additional flood events, different climatic regions, and alternative machine-learning bias-correction techniques to further assess its generality and operational robustness.

Acknowledgement

The authors thank Iran's National Science Foundation (INSF) for its support for this research.

Ethics approval

All authors abide by all ethical standards.

Consent to participate

All authors consent to participate.

Consent for publish

All authors consent to publish.

Authors' contributions

Masoumeh Zeinali; Software, Formal analysis, Writing - Original Draft

Omid Bozorg-Haddad; Conceptualization, Supervision, Project administration

Hugo A. Loáiciga; Validation, Writing - Review and Editing

Funding

The authors thank Iran's National Science Foundation (INSF) for its support for this research.

Competing interests

There is no conflict of interest.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The codes that support the findings of this study are available from the corresponding author upon reasonable request.

5. References

Ajaaj, A., Mishra, A. K., Khan, A. A. (2015). Comparison of bias correction techniques for GPCC rainfall. Stochastic Environmental Research and Risk Assessment, 30(6):1659-1675.

Amini, L., Parhizgar, D., and Khakian, Gh. (2013). The role of the WRF numerical model in numerically predicting heavy rains in Isfahan province with a resolution of 27, 9 and 3 km. the second national conference on flood management and engineering with the approach of urban floods (In Persian).

Aminyavari, S., Saghafian, B., and Delavar, M. (2019). Evaluation of TIGGE Ensemble Forecasts of Precipitation in Distinct Climate Regions in Iran. Advances in Atmospheric Sciences, 35(4): 457–468.

Aparicio, J., Martinez-Austria, P. F., Guitron, A., and Ramirez, A.I. (2009). Floods in Tabasco, Mexico: a diagnosis and proposal for courses of action. Journal of Flood Risk Management, 2(2), 132–138.

Araghinejad, Sh. (2014) Data-driven modeling: Using MATLAB® in water resources and environmental engineering. Water Science and Technology Library 67.

Arkian, F., Mashatan, N., Katiraei Boroujerdi, P. S., and Mirzaei Hajibaqlou, E. (2013). Evaluation of WRF regional model rainfall forecast for heavy daily rainfall in Qazvin province during the period 2011 to 2019. Marine Science and Technology Research, 9(4), 88-102 (In Persian).

Beygi, S., Bozorg-Haddad, O., Fallah-Mehdipour, E., and Mariño, M.A., (2014). “Bargaining models for optimal design of water distribution networks”, Journal of Water Resources Planning and Management, 140 (1), 92-99, DOI: 10.1061/ (ASCE)WR.1943-5452.0000324.

Bezada, M. (2009). Natural hazards and human-induced disasters triggered by intense and episodic tropical rains in the Venezuelan Mountains. Developments in Earth Surface Processes, 13, 115-129.

Bozorg-Haddad, O. and Mariño, M.A. (2007). “Dynamic penalty function as a strategy in solving water resources combinatorial optimization problems with honey-bee mating optimization (HBMO) algorithm”, Journal of Hydroinformatics, 9 (3), 233-250, DOI: 10.2166/hydro.2007.025.

Bozorg-Haddad, O., Moradi-Jalal, M., Mirmomeni, M., Kholghi, M.K.H., and Mariño, M.A. (2009a). “Optimal cultivation rules in multi-crop irrigation areas”, Irrigation and Drainage, 58(1), 38-49, DOI: 10.1002/ird.381.

Bozorg-Haddad, O. and Mariño, M.A. (2011). “Optimum operation of wells in coastal aquifers”, Proceedings of the Institution of Civil Engineers: Water Management, 164 (3), 135-146, DOI: 10.1680/wama.1000037.

Bronstert, A., Agarwal, A., Boessenkool, B., Crisologo, I., Fischer, M., Heistermann, M., Kohn-Riech, L., Lopez-Tarazon, J. A., Maron, T., Ozturk, U., Reinhardt-Imjela, C., and Wendi, D.

(2018). Forensic hydro-meteorological analysis of an extreme flash flood: the 2016-05-29 event in Braunsbach, SW Germany. Science of the Total Environment, 630(1), 991-997.

Da Silva, F. P., da Silva, A. S., da Silva, M. G. A. J., and Pires, G. D. (2022). Assessment of WRF numerical model forecasts using different lead time initializations during extreme precipitation events over Macaé city, Rio de Janeiro (Brazil). Natural Hazards, 110(1), 695-718.

Delpasand, M., Fallah-Mehdipour, E., Azizipour, M., Jalali, M., Safavi, H. R., Saghafian, B., Loáiciga, H. A., Singh Babel, M., Savic, D., Bozorg-Haddad, O. (2021). Forensic engineering analysis applied to flood control. Journal of Hydrology, 594, doi.org/10.1016/j.jhydrol.2021.125961

Durai, V. R., Bhowmik, S. R., and Mukhopadhyay, B. (2010). Performance Evaluation of precipitation prediction skill of NCEP Global Forecasting System (GFS) over Indian region during Summer Monsoon 2008. Mausam, 61(2), 139-154.

Gebremichael, M., Yue, H., Nourani, V., and Damoah, R. (2022). The skills of medium-range precipitation forecasts in the Senegal River basin. Sustainability, 14(6), 3349.

Fallah-Mehdipour, E., Bozorg-Haddad, O., Beygi, S., and Mariño, M.A. (2011). "Effect of utility function curvature of Young's bargaining method on the design of WDNs", Water Resources Management, 25(9), 2197-2218, DOI: 10.1007/s11269-011-9802-5.

Fallah-Mehdipour, E., Bozorg-Haddad, O., and Mariño, M.A. (2013). "Extraction of multicrop planning rules in a reservoir system: Application of evolutionary algorithms", Journal of Irrigation and Drainage Engineering, 139 (6), 490-498, DOI: 10.1061/ (ASCE)IR.1943-4774.0000572.

Gunn, S. R. (1998). Support vector machines for classification and regression. ISIS technical report, 14(1), 5-16.

He, S., Raghavan, S. V., Nguyen, N. S., and Liong, S. Y. (2013). Ensemble rainfall forecasting with numerical weather prediction and radar-based nowcasting models. Hydrological Processes, 27(11), 1560-1571.

Hydrologic Ensemble Prediction Experiment. HEPEX-SIP, Topic:Post-processing (1/3). (2018). Available: https://hepex.org.au/hepex-sip-topic-post-processing-13. [Online].

Hsiao, L. F., Yang, M. J., Lee, C. S., Kuo, H. C., Shih, D. S., Tsai, C. C., ... and Hong, J. S. (2013). Ensemble forecasting of typhoon rainfall and floods over a mountainous watershed in Taiwan. Journal of hydrology, 506, 55-68.

Kalnay, E. (2003). Atmospheric modeling, data assimilation and predictability. Cambridge university press.

Kalnay, E., Kanamitsu, M., and Baker, W. E. (1990). Global numerical weather prediction at the National Meteorological Center. Bulletin of the American Meteorological Society, 71(10), 1410-1428.

Kanamitsu, M. (1989). Description of the NMC global data assimilation and forecast system. Weather and forecasting, 4(3), 335-342.

Kanamitsu, M., Alpert, J. C., Campana, K. A., Caplan, P. M., Deaven, D. G., Iredell, M., ... and White, G. H. (1991). Recent changes implemented into the global forecast system at NMC. Weather and Forecasting, 6(3), 425-435.

Karimi-Hosseini, A., Bozorg-Haddad, O., and Mariño, M.A. (2011). “Site selection of raingauges using entropy methodologies”, Proceedings of the Institution of Civil Engineers: Water Management, 164 (7), 321-333, DOI: 10.1680/wama.2011.164.7.321.

Khajehei, S., and Moradkhani, H. (2017). Towards an improved ensemble precipitation forecast: a probabilistic post-processing approach. Journal of Hydrology, 546, 476-489.

Kolachian, R., Saghafian, B., and Moazzami, S. (2019). Evaluating the effectiveness of post-processing and skew correction methods on monthly precipitation and temperature forecasts in the Karun basin. Iran Water Resources Research, 16(4), 98-111. (In Persian).

Lazoglou, G., Zittis, G., Anagnostopoulou, C., Hadjinicolaou, P., and Lelieveld, J. (2020). Bias correction of RCM precipitation by TIN-copula method: a case study for historical and future simulations in Cyprus. Climate, 8(7), 85. https://doi.org/10.3390/CLI8070085

Li, J., Chen, Y., Wang, H., Qin, J., Li, J., and Chiao, S. (2017). Extending flood forecasting lead time in a large watershed by coupling WRF QPF with a distributed hydrological model. Hydrology and Earth System Sciences, 21(2), 1279.

Liguori, S., and Rico-Ramirez, M. A. (2012). Quantitative assessment of short-term rainfall forecasts from radar nowcasts and MM5 forecasts. Hydrological Processes, 26(25), 3842-3857.

Lim, K. S. S., and Hong, S. Y. (2010). Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models. Monthly Weather Review, 138(5), pp.1587-1612.

Loáiciga, H. A. (2001). Flood damages under changing flood-plain conditions: a forensic-hydrology case study, Journal of the American Water Resources Association 37(2), 467-478. https://doi.org/10.1111/j.1752-1688.2001.tb00982.x

Medina, H., and Tian, D. (2020). Comparison of probabilistic post-processing approaches for improving numerical weather prediction-based daily and weekly reference evapotranspiration forecasts. Hydrology and Earth System Sciences, 24(2), 1011-1030.

Milbrandt, J.A., and Yau, M.K. (2005). A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. Journal of the Aatmospheric Sciences, 62(9), pp.3065-3081.

Moghim, S., and Takallou, A. (2023). An integrated assessment of extreme hydrometeorological events in Bangladesh. Stochastic Environmental Research and Risk Assessment, 1-21.

Moorthi, S., Pan, H. L., Caplan, P. (2001). Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system. NWS Technical Procedures Bulletin, vol 484, p14. http://www.nws.noaa.gov/om/ tpb/484.htm

National Research Council. 2026. Attribution of Extreme Weather and Climate Events and Their Impacts. Washington, DC: The National Academies Press. https://doi.org/10.17226/28590

Ngai, S. T., Tangang, F., and Juneng, L. (2017). Bias correction of global and regional simulated daily precipitation and surface mean temperature over Southeast Asia using quantile mapping method. Global and Planetary Change, 149, 79-90.

Orouji, H., Bozorg-Haddad, O., Fallah-Mehdipour, E., and Marino, M.A., (2014). "Extraction of decision alternatives in project management: Application of hybrid PSO-SFLA", Journal of Management in Engineering, 30 (1), 50–59, DOI: 10.1061/ (ASCE)ME.1943-5479.0000186.

Powers, J. G., J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia, D. O. Gill, J. L. Coen, D. J. Gochis, R. Ahmadov, S. E. Peckham, G. A. Grell, J. Michalakes, S. Trahan, W. Wang, C. S. Schwartz, G. S. Romine, Z.-Q. Liu, C. Snyder, F. Chen, M. J. Barlage, W. Yu, and M. G. Duda, (2017). The Weather Research and Forecasting (WRF) Model: Overview, system eorts, and future directions. Bull. Amer. Meteor. Soc., 98, 1717-1737.

Sabbaghpour, S., Naghashzadehgan, M., Javaherdeh, K., and Bozorg-Haddad, O. (2012). "HBMO algorithm for calibrating water distribution network of Langarud city", Water Science and Technology, 65 (9), 1564-1569, DOI: 1564-1569; 10.2166/wst.2012.045.

Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., ..., and Goldberg, M. (2010). The NCEP climate forecast system reanalysis. Bulletin of the American Meteorological Society, 91(8), 1015-1058.

Schepen, A., Zhao, T., Wang, Q. J., and Robertson, D. E. (2017). A new method for post-processing daily sub-seasonal to seasonal rainfall forecasts from GCMs and evaluation for 12 Australian catchments. Hydrology and Earth System Sciences, Discuss, 1-27. https://doi.org/10.5194/hess-2017-380-RC1

Shokri, A., Bozorg-Haddad, O., and Mariño, M.A. (2013). "Algorithm for increasing the speed of evolutionary optimization and its accuracy in multi-objective problems", Water Resources Management, 27 (7), 2231-2249, DOI: 10.1007/s11269-013-0285-4.

Skamarock, W. C., and Klemp, J. B. (2008). A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. Journal of computational physics, 227(7), 3465-3485.

Soltanjalili, M., Bozorg-Haddad, O., and Mariño, M.A. (2011). "Effect of breakage level one in design of water distribution networks", Water Resources Management, 25(1), 311-337, DOI: 10.1007/s11269-010-9701-1.

SRCIF. (2019a). Report of Meteorological and Climatology Working group, Special Reporting Committee on Iran Floods 2019. URL: https://nfr.ut.ac.ir/

SRCIF. (2019b). Report of Hydrology and Water Resources Working group, Special Reporting Committee on Iran Floods 2019. URL: https://nfr.ut.ac.ir/

Sridevi, C., Singh, K. K., Suneetha, P., Durai, V. R., and Kumar, A. (2020). Rainfall forecasting skill of GFS model at T1534 and T574 resolution over India during the monsoon season. Meteorology and Atmospheric Physics, 132, 35-52.

Taghavi, F., Nistani, A., and Qader, S. (2012). Evaluation of short-term rainfall forecasts of the WRF numerical model in the Iranian region in a period of one month. Earth and Space Physics, 39(2), 145-170. doi: 10.22059/jesphys.2013.35196. (In Persian).

Tao, Y., Duan, Q., Ye, A., Gong, W., Di, Z., Xiao, M., and Hsu, K. (2014). An evaluation of post-processed TIGGE multimodel ensemble precipitation forecast in the Huai river basin. Journal of hydrology, 519, 2890-2905.

Tao, Y., Duan, Q., Ye, A., Gong, W., Di, Z h., Xiao, M. and Hsu, K. (2014). An evaluation of postprocessed TIGGE multimodel ensemble precipitation forecast in the Huai river basin. Journal of hydrology, 519(27), 2890-2905.

Tingsanchali, T. (2012). Urban flood disaster management. Procedia engineering, 32, 2537.

Toth, Z., (2001). Ensemble forecasting in WRF. Bulletin of the American Meteorological Society.Vol. 82, 695-697.

UNDRR. (2019). Global Assessment Report on Disaster Risk Reduction. United Nations for Disaster Risk Reduction: Geneva, Switzerland.

Vapnik, V. N. (1995). The nature of statistical learning theory. Springer, New York, ISBN 0-387-98780-0

Witte. M, J. P., Zaadnoordijk, W. J., and Buyse, J. J. (2019). Forensic hydrology reveals why groundwater tables in the province of Noord Brabant (the Netherlands) decreased more than expected. Water, 11(478), 1-14.

World Bank and United Nations. (2010). Natural Hazards, Unnatural Disasters: The Economics of Effective Prevention. The International Bank for Reconstruction and Development, The World Bank: Washington, DC, USA.

Wu, J., Lu, G., and Wu, Z. (2014). Flood forecasts based on multi-model ensemble precipitation forecasting using a coupled atmospheric-hydrological modeling system. Natural hazards, 74, 325-340.

Yu, W., Nakakita, E., and Jung, K. (2016). Flood Forecast and Early Warning with High-Resolution Ensemble Rainfall from Numerical Weather Prediction Model. Procedia Engineering, 154, 498-503.

Yu, W., Nakakita, E., Kim, S., and Yamaguchi, K. (2015). Improvement of rainfall and flood forecasts by blending ensemble NWP rainfall with radar prediction considering orographic rainfall. Journal of Hydrology, 531, 494-507.

Yuan, X., Wood, E. F., and Ma, Z. (2015). A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development. Wiley Interdisciplinary Reviews: Water, 2(5), 523-536.

Zarei, M., Bozorg-Haddad, O., and Singh, V. P. (2021). Forensic engineering. Water Resources: Future Perspectives, Challenges, Concepts and Necessities, 287. https://doi.org/10.2166/9781789062144_0287.

Zarei, M., Bozorg-Haddad, O., Loaiciga, H.A. (2023). A framework for the forensic-engineering assessment of reservoir operation during floods based on a new standard operation policy. Journal of Hydrology, https://doi.org/10.1016/j.jhydrol.2023.129774.

Zhao, T., Bennett, J. C., Wang, Q. J., Schepen, A., Wood, A. W., Robertson, D. E., and Ramos, M. H. (2017). How suitable is quantile mapping for postprocessing GCM precipitation forecasts?. Journal of Climate, 30(9), 3185-3196.

Ziegler, C.L. (1985). Retrieval of thermal and microphysical variables in observed convective storms. Part 1: Model development and preliminary testing. Journal of the Atmospheric Sciences, 42(14), pp.1487-1509.

Table 1. Characteristics of the weather stations in the studied area.

No.	Name	Latitude (North degree)	Longitude (East degree)	Elevation (m)	Established year
1	Ravansar	34.71	46.65	1380	1998
2	Kermanshah	34.35	47.15	1319	1942
3	Kangavar	34.50	47.98	1468	1987
4	Nurabad (Lorestan)	34.05	48.00	1859	2000
5	Malayer	34.25	48.86	1777	1991
6	Alashtar	33.82	48.25	1567	1996
7	Eslamabad-e-Gharb	34.12	46.47	1349	1987
8	Khorramabad	33.44	48.28	1148	1949
9	Poldokhtar	33.15	47.72	714	1998
10	Bostan	31.71	48.01	8	1986
11	Darrehshahr	33.14	47.41	670	2002
12	Toyserkan	34.55	48.43	1783	2003
13	Nahavand	34.14	48.41	1678	1992
14	Asadabad	34.77	48.12	1552	2010
15	Kamyaran	34.80	46.93	1404	2000
16	Sonqor	34.78	47.58	1700	2000
17	Harsin	34.27	47.55	1546	2006
18	Kuhdasht	33.52	47.65	1198	1987
19	Loumar	33.56	46.83	850	2006
20	Sarableh	33.79	46.58	1045	2006
21	Romeshkan	33.28	47.50	1089	2011
22	Sararud	34.33	47.29	1362	1989

Table 2. Some of the investigated microphysics schemes available in the WRF model.

mp physics	Schema name	Reference	Mass mixing ratio calculation
1	WDM	Lim and Hong (2010)	Cloud water, rain, snow, ice and snowflakes
2	NSSL	Ziegler (1985)	Cloud water, rain, snow, ice, snowflakes and hail
3	MYDM-7	Milbrandt and Yau (2005)	Cloud water, rain, snow, ice, snowflakes and hail

WDM: WRF Double-Moment 6-class scheme
NSSL: National Severe Storms Laboratory
MYDM-7: Milbrandt-Yau Double-Moment 7-class scheme

Table 3. Characteristics of the weather forecasting models.

Model	Data center	Spatial resolution (km)	Forecast horizon (days)
GFS	NCEP	13	16
WRF	WRI	8	3

Table 4. Characteristics of different sub-basins of the study area.

The name of the subbasin	Runoff coefficient	Area (km^2)
Cham Anjir	0.54	1773.07
Doab Kashkan	0.69	4498.09
Afarineh-Kashkan	0.58	7685.29
Poldokhtar-Kashkan	0.64	9825.98
Polchehr	0.225	12032.27
Holeilan	0.23	21528.7
Seimareh	0.18	27064.25
Jelogir	0.30	41903.59
Paypol	0.31	45616.4

Table 5. R index values for the WRF and GFS forecasting models with 3-days ahead forecasts without and with bias correction.

Station name	GFS						WRF					
	Lead Time 1		Lead Time 2		Lead Time 3		Lead Time 1		Lead Time 2		Lead Time 3	
	R	R-bc	R	R-bc	R	R-bc	R	R-bc	R	R-bc	R	R-bc
Alashtar	0.75	0.92	0.75	0.81	0.68	0.75	0.75	0.86	0.84	0.91	0.81	0.81
Asadabad	0.82	0.90	0.75	0.72	0.70	0.87	0.87	0.97	0.85	0.93	0.66	0.79
Bostan	0.34	0.63	0.23	0.53	0.28	0.44	0.32	0.47	0.37	0.49	0.31	0.53
Darrehshahr	0.75	0.91	0.80	0.88	0.72	0.76	0.68	0.90	0.79	0.89	0.75	0.83
Eslamabad-E-Gharb	0.72	0.89	0.66	0.80	0.76	0.86	0.72	0.77	0.67	0.74	0.73	0.76
Harsin	0.73	0.91	0.61	0.85	0.62	0.51	0.52	0.69	0.56	0.87	0.64	0.83
Kamyaran	0.85	0.88	0.81	0.94	0.70	0.83	0.78	0.81	0.84	0.92	0.61	0.87
Kangavar	0.84	0.89	0.75	0.79	0.67	0.82	0.76	0.80	0.75	0.70	0.71	0.83
Kermanshah	0.71	0.74	0.70	0.74	0.70	0.90	0.72	0.86	0.75	0.86	0.59	0.79
Khorramabad	0.64	0.70	0.70	0.74	0.65	0.66	0.58	0.70	0.66	0.74	0.67	0.61
Kuhdasht	0.74	0.82	0.74	0.82	0.67	0.69	0.68	0.83	0.76	0.90	0.59	0.78
Loumar	0.74	0.91	0.78	0.93	0.76	0.86	0.73	0.92	0.71	0.77	0.74	0.81
Malayer	0.56	0.70	0.65	0.86	0.51	0.54	0.60	0.77	0.77	0.87	0.71	0.81
Nahavand	0.64	0.89	0.71	0.73	0.58	0.65	0.81	0.95	0.86	0.94	0.65	0.86
Nurabad (Lorestan)	0.78	0.92	0.85	0.98	0.61	0.72	0.84	0.94	0.91	0.97	0.71	0.85
Poldokhtar	0.85	0.87	0.70	0.75	0.63	0.73	0.86	0.95	0.68	0.83	0.78	0.80
Ravansar	0.84	0.87	0.79	0.90	0.68	0.81	0.73	0.81	0.78	0.81	0.48	0.84
Romeshkan	0.77	0.90	0.76	0.84	0.73	0.79	0.73	0.95	0.82	0.87	0.75	0.95
Sarableh	0.72	0.94	0.67	0.69	0.70	0.84	0.75	0.92	0.65	0.90	0.76	0.93
Sararud (Kermanshah)	0.81	0.94	0.76	0.72	0.75	0.91	0.83	0.91	0.79	0.61	0.61	0.56

Sonqor	0.78	0.82	0.71	0.86	0.67	0.51	0.83	0.88	0.75	0.73	0.41	0.69
Toyserkan	0.84	0.88	0.80	0.87	0.68	0.86	0.73	0.79	0.75	0.84	0.65	0.64

bc: the index representing results obtained after applying the bias correction.

Table 6. RMSE index values for the WRF and GFS forecasting models with a 3-days ahead forecasting horizon without and with bias correction.

Station name	GFS						WRF					
	Lead Time 1		Lead Time 2		Lead Time 3		Lead Time 1		Lead Time 2		Lead Time 3	
	RMSE	RMSE-bc	RMSE	RMSE-bc	RMSE	RMSE-bc	RMSE	RMSE-bc	RMSE	RMSE-bc	RMSE	RMSE-bc
Alashtar	10.2	5.4	9.7	8.4	11.7	9.1	8.8	6.9	7.9	5.9	7.6	7.5
Asadabad	7.5	5.8	8.5	9.8	9.6	6.3	6.3	3.3	7.5	5.5	8.5	7.2
Bostan	9.8	7.6	11.8	7.7	15.3	8.2	10.2	7.9	12.4	7.8	12.8	8.0
Darrehshahr	11.7	6.9	10.1	7.9	11.9	11.2	12.7	7.0	12.6	8.8	13.5	11.2
Eslamabad-E-Gharb	8.8	4.9	8.7	6.5	7.8	5.6	6.9	5.0	8.3	7.6	7.4	7.2
Harsin	10.4	4.8	11.6	6.2	9.2	9.7	10.2	6.3	10.8	5.8	8.9	6.4
Kamyaran	6.4	5.2	6.6	3.8	8.0	6.1	7.1	6.1	6.1	4.5	7.4	4.6
Kangavar	6.9	5.6	8.6	7.5	9.3	7.4	6.9	6.3	8.8	9.5	8.9	7.3
Kermanshah	8.9	7.8	8.7	7.6	8.3	4.8	9.3	6.2	8.3	6.1	8.7	6.4
Khorramabad	10.8	9.2	9.5	8.4	11.8	9.4	10.8	9.1	10.5	9.1	9.1	9.4
Kuhdasht	8.1	5.6	7.3	5.7	7.5	6.9	8.4	6.2	7.1	4.6	9.5	6.5
Loumar	10.2	5.5	8.7	4.9	9.1	7.2	7.6	4.2	10.3	9.5	10.6	9.0
Malayer	12.6	10.4	11.1	7.4	13.1	12.2	11.1	9.3	10.9	8.3	8.1	6.4
Nahavand	13.9	8.2	11.8	11.7	14.1	12.5	9.5	4.9	10.5	6.1	9.9	7.5
Nurabad (Lorestan)	11.1	7.1	9.6	3.9	14.3	13.2	9.6	5.9	9.1	4.8	9.9	7.6
Poldokhtar	7.9	5.6	9.1	7.8	11.7	7.8	6.3	3.9	9.3	6.8	8.1	7.9
Ravansar	6.2	5.1	6.6	4.6	8.0	6.4	7.2	5.3	6.9	6.5	9.2	5.6
Romeshkan	9.8	5.9	9.0	7.6	9.8	8.2	9.6	4.2	8.3	7.2	9.6	4.7
Sarableh	12.1	5.8	12.2	12.4	11.9	9.0	9.6	5.3	13.2	7.7	11.7	6.5
Sararud (Kermanshah)	7.4	4.1	8.0	7.4	6.7	4.2	8.3	4.6	8.6	9.2	7.3	7.2
Sonqor	7.8	4.2	7.3	3.5	5.9	6.0	6.0	3.4	7.1	5.0	5.9	3.6
Toyserkan	6.1	5.1	6.6	5.1	7.9	5.4	7.6	6.0	8.1	6.5	8.9	8.4

bc: the index representing the model results with the bias correction.

Table 7. pBias index values for the WRF and GFS forecasting models with a 3-day ahead forecasting horizon without and with bias correction.

Station name	GFS						WRF					
	Lead Time 1		Lead Time 2		Lead Time 3		Lead Time 1		Lead Time 2		Lead Time 3	
	PBias	PBias-bc	PBias	PBias-bc	PBias	PBias-bc	PBias	PBias-bc	PBias	PBias-bc	PBias	PBias-bc
Alashtar	2.5	0.4	-4.6	-3.1	-2.5	-2.0	-11.0	-9.0	-11.3	-4.9	-1.9	-1.4

Asadabad	-21.0	-2.7	-17.5	-12.7	-31.5	-3.9	8.7	-4.4	12.4	6.4	7.4	-4.6
Bostan	-23.3	-11.6	-17.6	-13.8	-12.8	-9.1	-29.9	-23.0	-7.9	-5.2	-21.9	-12.3
Darrehshahr	-6.6	-5.3	-16.6	-12.3	-16.1	-12.2	-42.5	-4.7	-38.2	-11.4	-42.4	-18.3
Eslamabad-E-Gharb	-2.2	-1.7	-11.0	-7.4	-24.0	-15.3	-7.2	-5.2	-13.1	-10.0	-28.0	-17.3
Harsin	9.1	-5.8	11.7	6.6	-11.2	-7.9	0.5	-0.3	-3.1	-2.2	-7.0	-5.6
Kamyaran	4.1	-2.7	0.8	-0.6	-5.1	-3.7	11.5	6.7	0.1	-0.1	-1.4	-1.1
Kangavar	-12.0	-7.7	-5.3	-3.5	-28.3	-8.7	-8.8	-6.7	-13.6	-11.0	-19.3	-15.5
Kermanshah	2.4	-1.8	-3.9	-2.6	-20.9	-6.0	1.6	-1.1	-6.3	-5.2	-25.2	-10.6
Khorramabad	-12.8	-8.7	-15.8	-10.3	-10.1	-6.8	-32.0	-11.6	-29.6	-17.5	-18.8	-15.2
Kuhdasht	25.1	11.9	19.8	10.0	6.1	-3.7	9.1	-5.1	10.0	0.0	20.0	12.3
Loumar	0.4	-0.3	-7.2	-5.5	-20.1	-8.7	-28.4	5.9	-24.7	-21.3	-38.9	-19.6
Malayer	-8.4	-6.0	-13.5	-11.9	-16.7	-11.5	-17.5	-13.7	-25.9	-18.3	-10.0	-7.2
Nahavand	15.2	8.8	1.7	-1.2	2.3	-1.6	-24.2	-8.5	-26.1	-10.3	-22.5	-13.0
Nurabad (Lorestan)	-5.9	1.7	-10.7	-1.6	-16.1	-12.8	-13.7	-1.0	-13.6	-4.3	-4.5	-3.3
Poldokhtar	4.9	-3.3	2.2	-1.4	3.5	-2.3	-15.7	-2.7	-10.8	-2.5	6.1	-3.2
Ravansar	-6.3	-4.6	-6.8	-5.1	-9.7	-7.2	7.4	-4.6	3.7	-2.7	0.7	-0.5
Romeshkan	7.1	-5.9	1.2	-0.8	-6.8	-4.7	-7.6	-5.9	-9.4	0.8	-4.6	-3.6
Sarableh	-12.1	-11.3	-21.7	-16.4	-30.2	-18.9	-7.6	-5.0	-18.3	-12.9	-33.2	-6.0
Sararud (Kermanshah)	17.1	10.9	14.9	7.7	-4.7	0.7	25.8	5.8	28.7	13.0	11.8	6.1
Sonqor	34.3	14.7	42.5	7.2	13.1	6.6	22.3	12.8	29.8	14.8	17.3	10.2
Toyserkan	1.6	-1.0	0.2	-0.1	-11.3	-7.4	28.4	13.4	23.3	5.5	23.9	11.2

bc: the index representing the model results obtained with the bias correction.

Table 8. NSE index values for the WRF and GFS forecasting models, with a 3-day forecasting horizon without and with bias correction.

Station name	GFS						WRF					
	Lead Time 1		Lead Time 2		Lead Time 3		Lead Time 1		Lead Time 2		Lead Time 3	
	NSE	NSE -bc	NSE	NSE -bc	NSE	NSE -bc	NSE	NSE -bc	NSE	NSE -bc	NSE	NSE -bc
Alashtar	0.43	0.84	0.48	0.62	0.26	0.55	0.54	0.72	0.70	0.83	0.65	0.65
Asadabad	0.65	0.79	0.55	0.40	0.43	0.75	0.75	0.93	0.69	0.84	0.41	0.57
Bostan	-0.21	0.28	-0.74	0.26	-1.85	0.18	-0.41	0.14	-1.06	0.17	-0.90	0.26
Darrehshahr	0.50	0.83	0.62	0.77	0.49	0.55	0.35	0.80	0.49	0.75	0.41	0.59
Eslamabad-E-Gharb	0.30	0.78	0.31	0.61	0.45	0.72	0.24	0.59	0.42	0.51	0.48	0.51
Harsin	0.14	0.81	-0.07	0.70	0.33	0.25	-0.38	0.47	0.15	0.76	0.37	0.68
Kamyaran	0.65	0.76	0.62	0.88	0.44	0.68	0.52	0.64	0.70	0.84	0.33	0.74
Kangavar	0.67	0.78	0.50	0.62	0.42	0.64	0.56	0.64	0.55	0.48	0.49	0.66
Kermanshah	0.36	0.52	0.39	0.54	0.45	0.82	0.40	0.73	0.51	0.74	0.30	0.62
Khorramabad	0.24	0.45	0.41	0.54	0.10	0.43	0.25	0.46	0.38	0.53	0.38	0.34
Kuhdasht	0.28	0.65	0.42	0.64	0.39	0.48	0.30	0.62	0.55	0.81	0.16	0.60
Loumar	0.43	0.83	0.58	0.87	0.54	0.71	0.48	0.84	0.47	0.55	0.43	0.59
Malayer	0.23	0.48	0.40	0.74	0.18	0.28	0.33	0.54	0.51	0.72	0.43	0.65

Nahavand	0.28	0.75	0.47	0.48	0.26	0.41	0.62	0.90	0.64	0.88	0.38	0.64
Nurabad (Lorestan)	0.61	0.84	0.71	0.95	0.36	0.45	0.68	0.88	0.77	0.94	0.51	0.71
Poldokhtar	0.52	0.76	0.36	0.53	-0.05	0.54	0.72	0.89	0.43	0.69	0.56	0.58
Ravansar	0.65	0.76	0.60	0.81	0.40	0.63	0.34	0.65	0.59	0.64	0.15	0.69
Romeshkan	0.47	0.81	0.54	0.67	0.47	0.63	0.46	0.90	0.67	0.75	0.55	0.89
Sarableh	0.42	0.87	0.41	0.38	0.43	0.68	0.45	0.84	0.40	0.79	0.50	0.84
Sararud (Kermanshah)	0.45	0.83	0.36	0.45	0.55	0.82	0.40	0.81	0.35	0.27	0.23	0.26
Sonqor	-0.36	0.60	-0.20	0.73	0.22	0.20	0.25	0.76	0.00	0.51	-0.44	0.46
Toyserkan	0.66	0.76	0.59	0.76	0.43	0.73	0.31	0.57	0.47	0.66	0.31	0.39

bc: the index representing the model results with the bias correction.

Table 9. The percentage difference between the forecasted runoff and the observed runoff for a 1-day ahead horizon.

Subbasin	Runoff coefficient	Area (km2)	Observed flood volume (mcm)	Day1			
				GFS	GFS-BC	WRF	WRF-BC
Polghehr	0.225	12032.3	219.7	-29	-20	-29	-24
Holeilan	0.23	21528.7	379.6	-19	-11	-19	-13
Poldokhtar-kashkan	0.64	9826.0	717.5	-19	-17	-33	-28
Afarineh-kashkan	0.58	7685.3	510.8	-23	-18	-33	-27
Doab-kashkan	0.69	4498.1	364.1	-24	-9	-33	-24
Cham Anjir	0.54	1773.1	105.8	-48	-33	-33	-33
Seimareh	0.18	27064.3	359.2	-18	-10	-18	-13
Jelogir	0.3	41903.6	1072.0	-18	-12	-26	-21
Paypol	0.31	45616.4	1258.9	-19	-13	-28	-22

Table 10. The percentage difference between the forecasted runoff and the observed runoff for a 2-days ahead horizon.

Subbasin	Runoff coefficient	Area (km2)	Observed flood volume (mcm)	Day2			
				GFS	GFS-BC	WRF	WRF-BC
Polghehr	0.225	12032.3	219.7	-48	-37	-30	-24
Holeilan	0.23	21528.7	379.6	-33	-26	-22	-18
Poldokhtar-kashkan	0.64	9826.0	717.5	-46	-34	-52	-37
Afarineh-kashkan	0.58	7685.3	510.8	-49	-36	-53	-37
Doab-kashkan	0.69	4498.1	364.1	-47	-30	-45	-28
Cham Anjir	0.54	1773.1	105.8	-56	-49	-66	-53
Seimareh	0.18	27064.3	359.2	-30	-26	-20	-16
Jelogir	0.3	41903.6	1072.0	-34	-28	-33	-24
Paypol	0.31	45616.4	1258.9	-36	-30	-36	-25

Table 11. The percentage difference between the forecasted runoff and the observed runoff for a 3-days ahead horizon.

Subbasin	Runoff coefficient	Area (km2)	Observed flood volume (mcm)	Day3			
				GFS	GFS-BC	WRF	WRF-BC
Polghehr	0.225	12032.3	219.7	-59	-47	-24	-20
Holeilan	0.23	21528.7	379.6	-50	-38	-14	-11
Poldokhtar-kashkan	0.64	9826.0	717.5	-47	-43	-45	-30
Afarineh-kashkan	0.58	7685.3	510.8	-47	-45	-46	-30
Doab-kashkan	0.69	4498.1	364.1	-49	-40	-42	-27
Cham Anjir	0.54	1773.1	105.8	-43	-32	-53	-38
Seimareh	0.18	27064.3	359.2	-47	-37	-13	-10
Jelogir	0.3	41903.6	1072.0	-47	-36	-25	-19
Paypol	0.31	45616.4	1258.9	-48	-38	-28	-21

Table 12. The percentage difference between the forecasted runoff and the observed flood volume for a 1-day ahead horizon.

Subbasin	Runoff coefficient	Area (km2)	Observed flood volume (mcm)	Day1			
				GFS	GFS-BC	WRF	WRF-BC
Polghehr	0.225	12032.3	327.6	-41	-17	-20	-10
Holeilan	0.23	21528.7	549.3	-26	-11	-11	-9
Poldokhtar-kashkan	0.64	9826.0	537.6	-10	-5	-9	3
Afarineh-kashkan	0.58	7685.3	385.6	-17	-6	-11	3
Doab-kashkan	0.69	4498.1	315.5	-33	-10	-18	-2
Cham Anjir	0.54	1773.1	60.4	14	3	4	16
Seimareh	0.18	27064.3	512.0	-15	-10	-6	-8
Jelogir	0.3	41903.6	1283.5	-14	-11	-10	-7
Paypol	0.31	45616.4	1478.6	-16	-13	-14	-9

Table 13. The percentage difference between the forecasted runoff and the observed runoff for a 2-days ahead horizon.

Subbasin	Runoff coefficient	Area (km2)	Observed flood volume (mcm)	Day2			
				GFS	GFS-BC	WRF	WRF-BC
Polghehr	0.225	12032.3	327.6	-46	-24	-29	-10
Holeilan	0.23	21528.7	549.3	-30	-21	-24	-8

Poldokhtar-kashkan	0.64	9826.0	537.6	-14	-10	-10	-1
Afarineh-kashkan	0.58	7685.3	385.6	-17	-12	-12	-1
Doab-kashkan	0.69	4498.1	315.5	-28	-13	-17	-1
Cham Anjir	0.54	1773.1	60.4	7	-10	1	-2
Seimareh	0.18	27064.3	512.0	-24	-19	-21	-8
Jelogir	0.3	41903.6	1283.5	-23	-18	-21	-10
Paypol	0.31	45616.4	1478.6	-26	-21	-24	-13

Table 14. The percentage difference between the forecasted runoff and the observed runoff for 3-days ahead horizon.

Subbasin	Runoff coefficient	Area (km2)	Observed flood volume (mcm)	Day3			
				GFS	GFS-BC	WRF	WRF-BC
Polghehr	0.225	12032.3	327.6	-65	-50	-40	-24
Holeilan	0.23	21528.7	549.3	-46	-39	-21	-21
Poldokhtar-kashkan	0.64	9826.0	537.6	-32	-29	-34	-20
Afarineh-kashkan	0.58	7685.3	385.6	-37	-34	-34	-20
Doab-kashkan	0.69	4498.1	315.5	-51	-45	-38	-23
Cham Anjir	0.54	1773.1	60.4	-6	-12	-25	-12
Seimareh	0.18	27064.3	512.0	-35	-33	-11	-17
Jelogir	0.3	41903.6	1283.5	-35	-31	-21	-20
Paypol	0.31	45616.4	1478.6	-37	-32	-26	-23

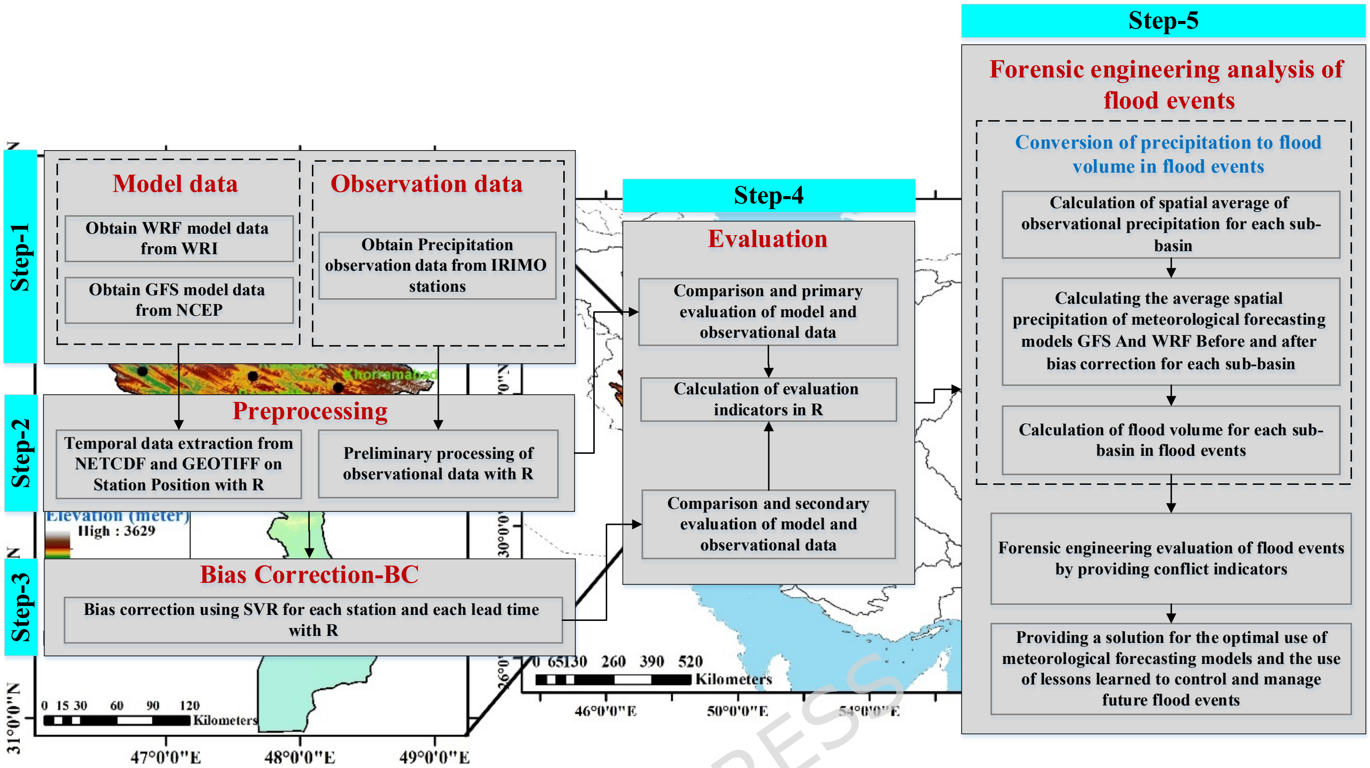


Figure 1. Flowchart for this paper's methodology.

Figure 2. The study area and the location of the weather stations.

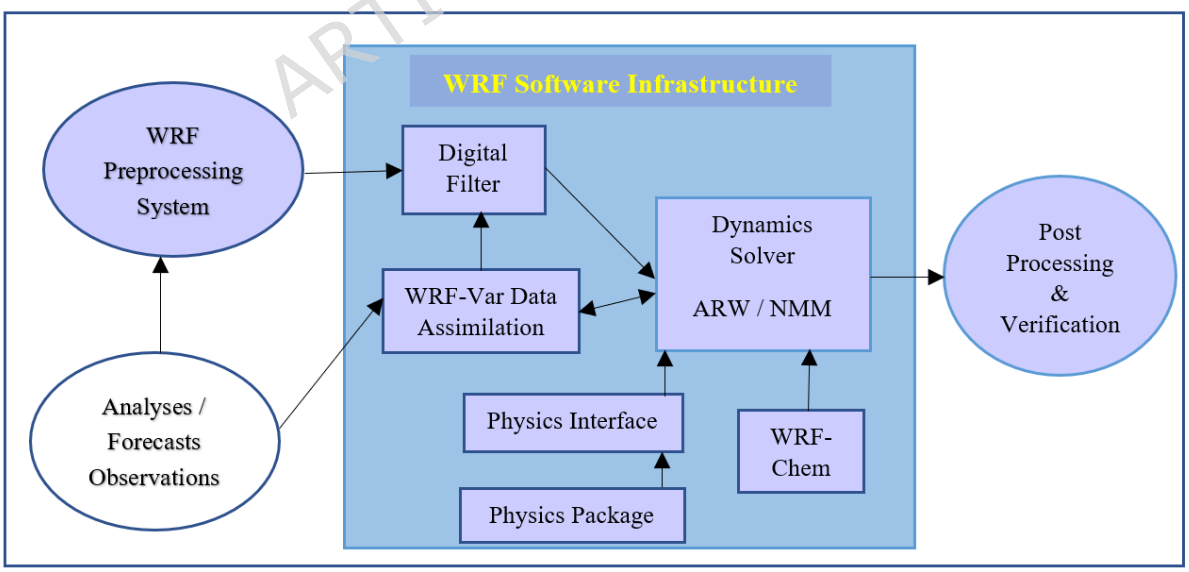


Figure 3. Software structure of WRF model (Skamarock et al. 2008)

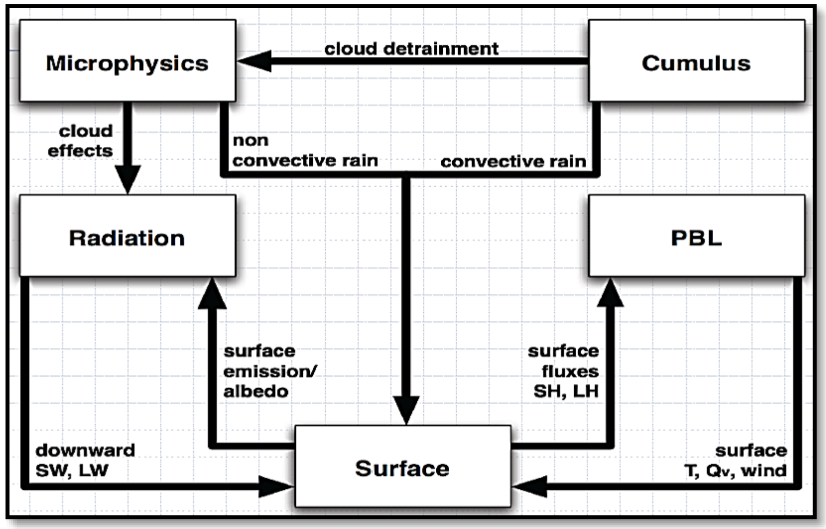


Figure 4. The relationship between different schemes of the WRF model (Skamarock et al. 2008)

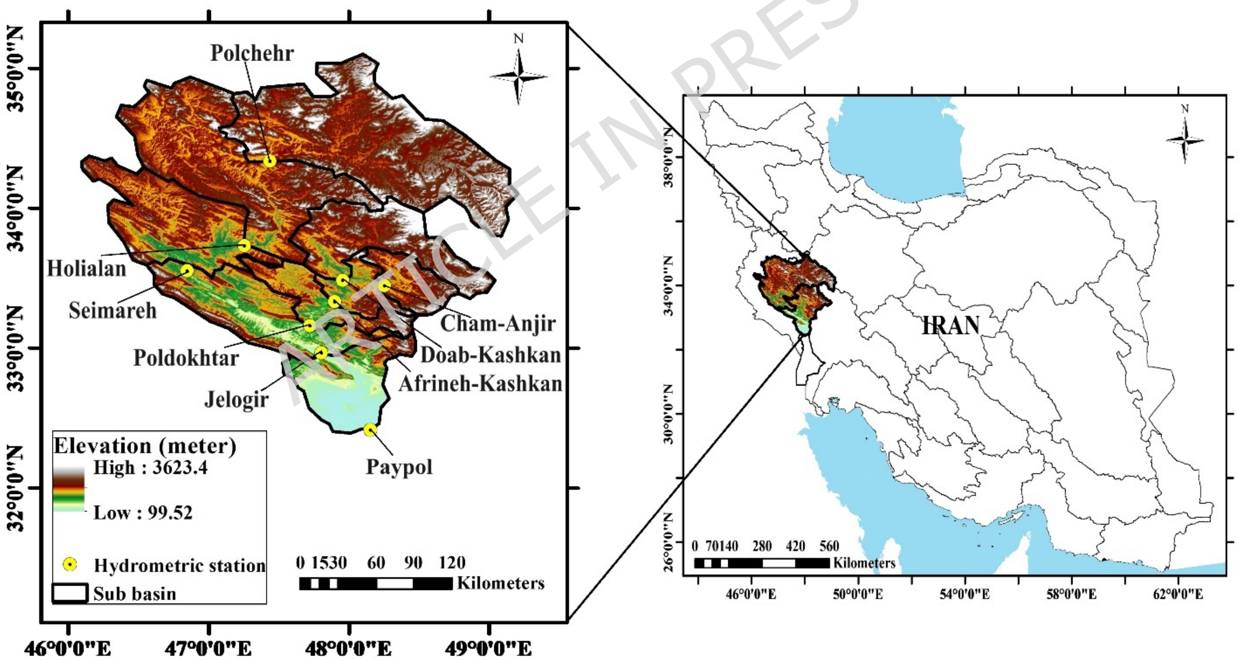
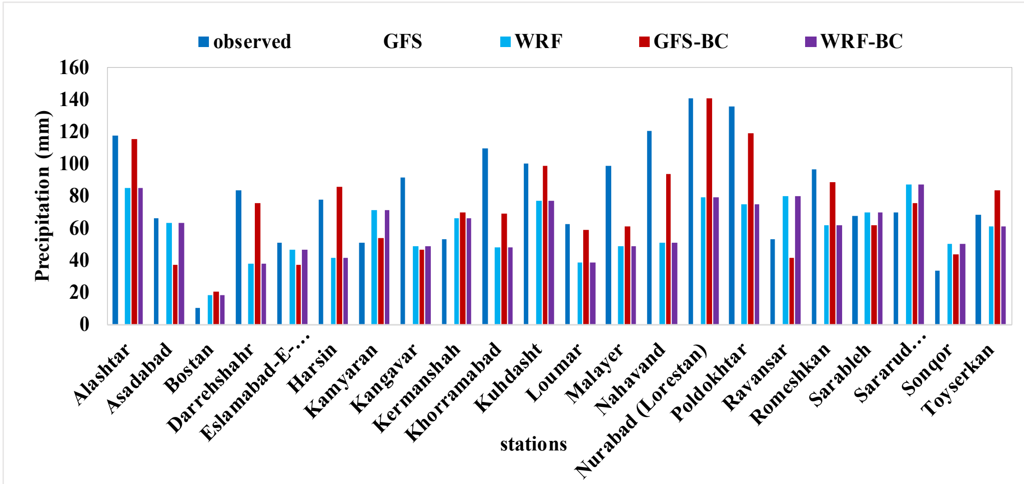
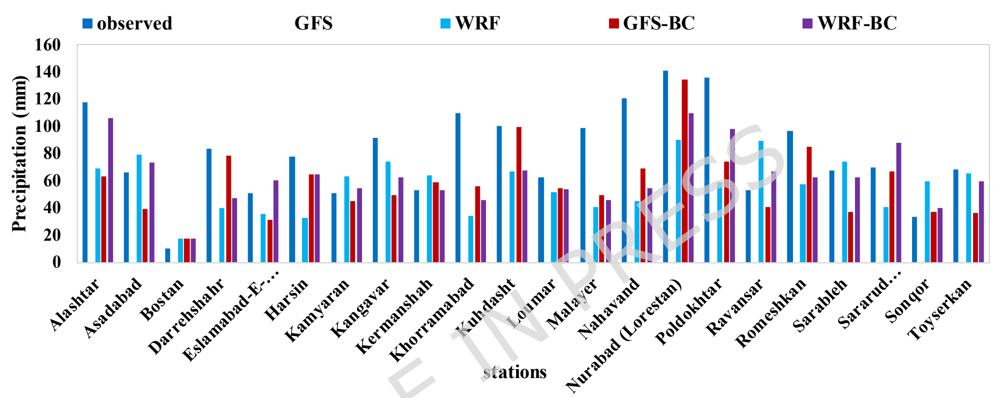


Figure 5. Division of the studied basin into several sub-basins and the location of hydrometric stations.

Accumulation 1 day



Accumulation 2 days



Accumulation 3 days

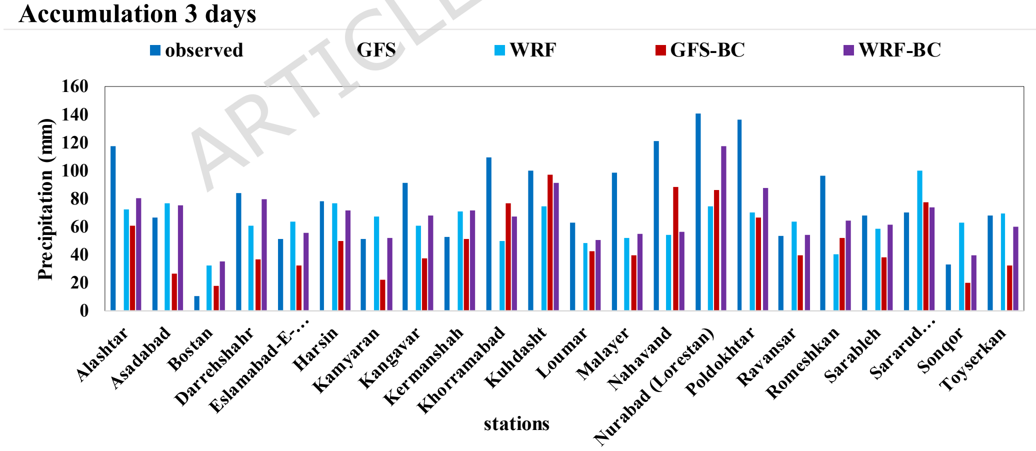
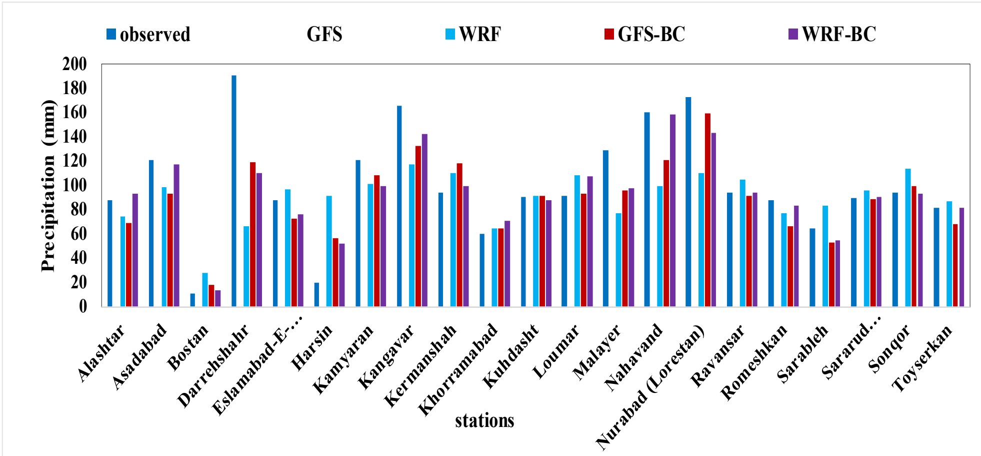
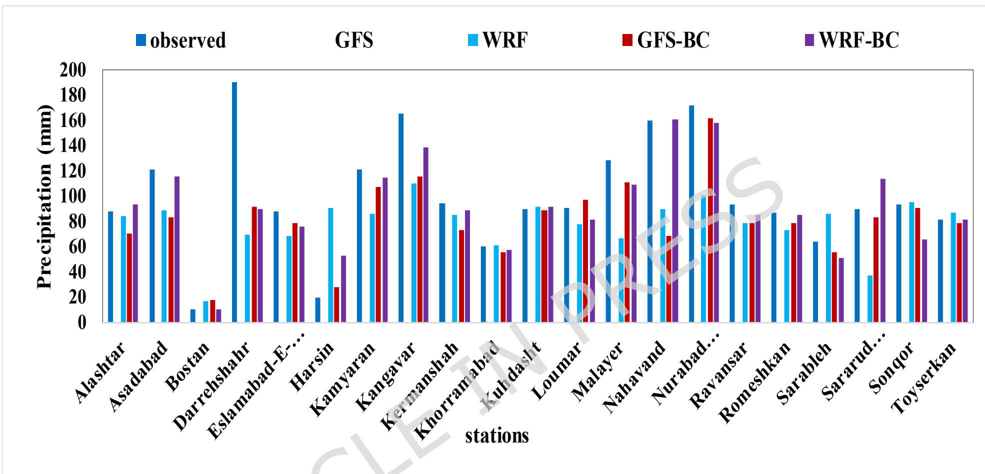


Figure 6. Observed and forecasted rainfall at the stations for 1-, 2-, and 3-days ahead horizons (first flood event).

Accumulation 1 day



Accumulation 2 days



Accumulation 3 days

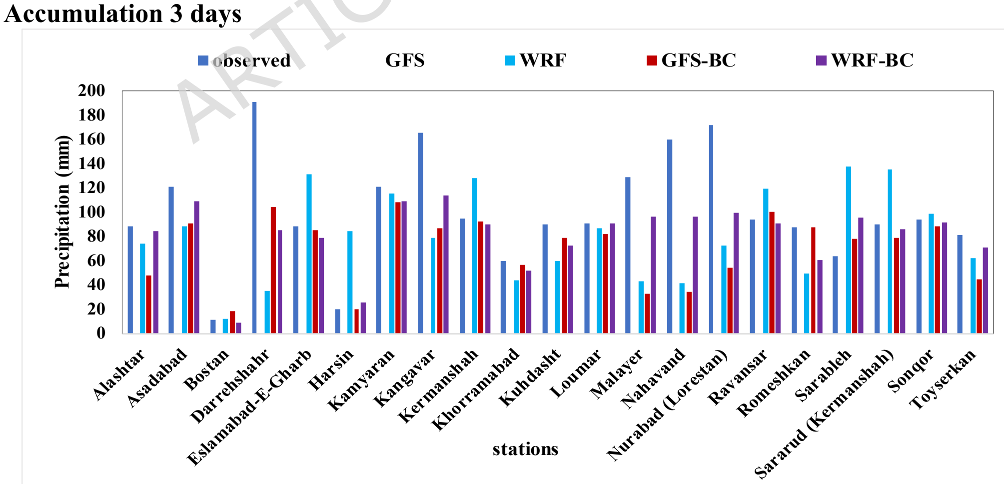
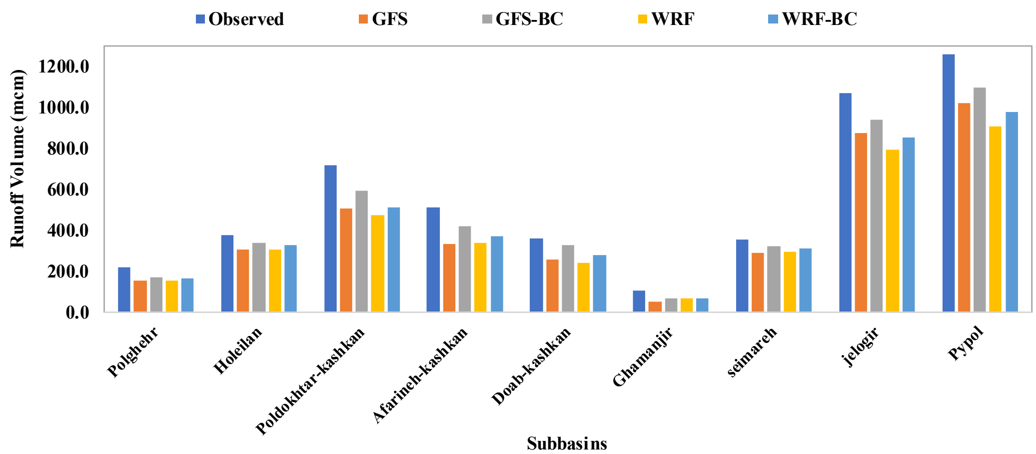
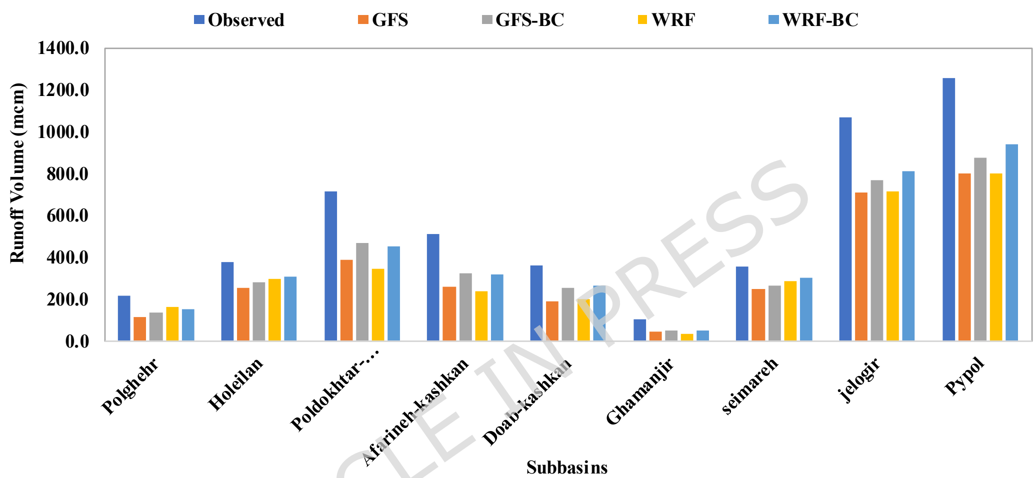


Figure 7. Observed forecasted rainfall at the stations for the 1-, 2-, and 3-days ahead horizons (second flood event)

Day1



Day2



Day3

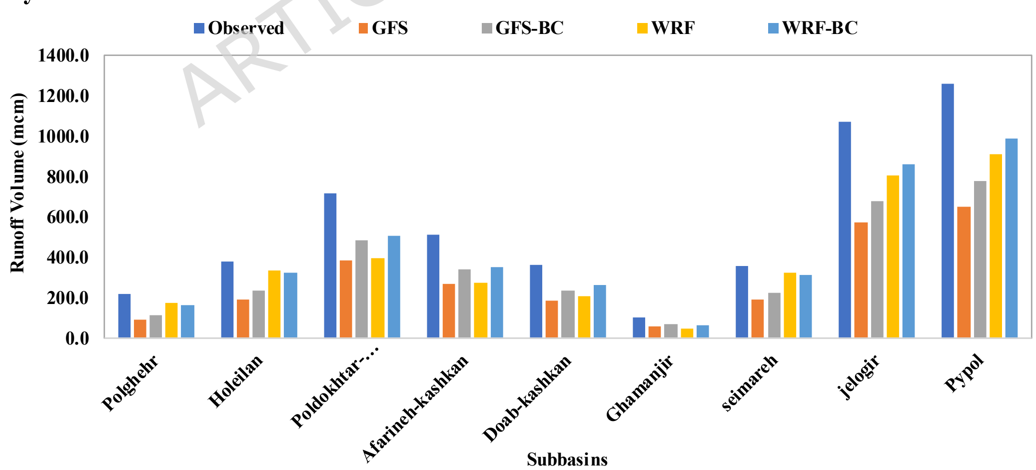


Figure 8. Comparison of observed and forecasted runoff in the study sub-basins for the 1-, 2-, and 3-days ahead horizons (first flood event).

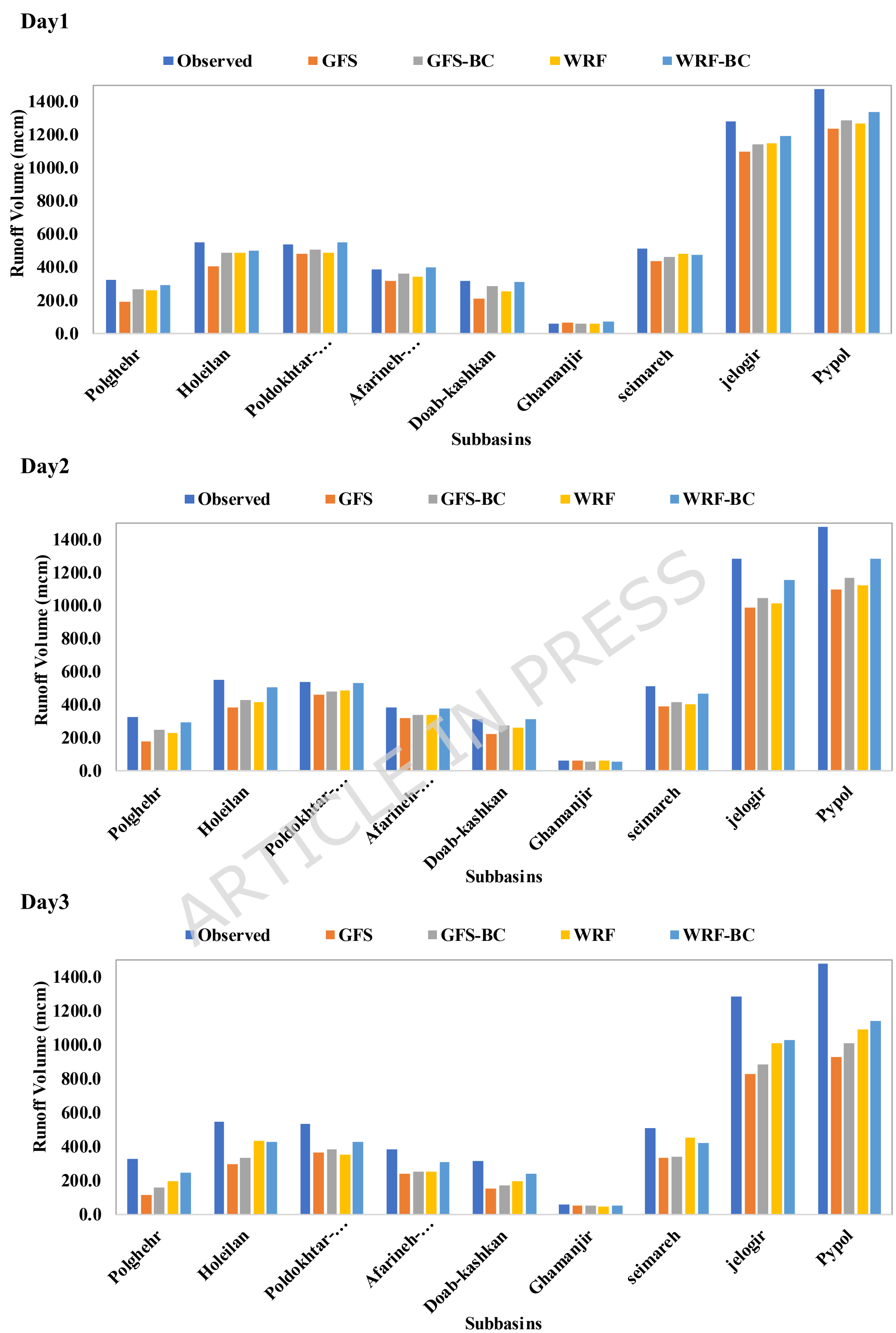


Figure 9. Comparison of the observed and forecasted runoff in the study sub-basins for the 1-, 2, and 3-days horizons (second flood event).