Assessing the Value of a Regional Climate Model's Rainfall Forecasts in Improving Dry-Season Streamflow Predictions

Hui Wang, M.ASCE[1]; Tirusew Asefa, F.ASCE[2]; Vasubandhu Misra[3]; and Amit Bhardwaj[4]

Abstract: Rainfall is a critical input variable of statistical streamflow forecasting models at subseasonal to seasonal time scales. This study presents a framework for evaluating the utility of a high-resolution experimental winter seasonal climate reforecasts for Florida (CLIFF) in improving streamflow forecasts. The CLIFF forecasts were coproduced through a scientist–stakeholder group of the Florida Water and Climate Alliance. The framework consists of a statistical streamflow generation model, four different sets of rainfall inputs, and distinct metrics for evaluating the resulting streamflow forecasts. The four sets of rainfall inputs include rainfall climatology, observed rainfall, NOAA-based seasonal rainfall forecasts, and CLIFF-based rainfall forecasts. Because NOAA ensemble precipitation forecasts were not available in this study, NOAA-based categorical precipitation outlooks were postprocessed via a hidden Markov chain model to obtain the corresponding NOAA-based seasonal rainfall forecasts. Streamflow forecasts based on rainfall climatology served as a reference. Different evaluation metrics, including Spearman correlation, mean absolute percent error (MAPE), and rank probability skill score (RPSS), were employed to evaluate model performance. The framework was demonstrated for streamflow forecasts for two rivers in the southwest of Florida, serving as a major source of a regional water supply agency. A retrospective streamflow forecasting model was designed for the dry season [November, December, January, and February (NDJF) months] for each of the 20 years from 2000 to 2019. Results revealed that CLIFF-based streamflow forecasts are a promising alternative to NOAA-based forecasts. Deterministic streamflow forecasts based on CLIFF rainfall have a smaller mean absolute percent error (MAPE) compared with the NOAA-based streamflow forecasts. Although NOAA-based probabilistic streamflow forecasts outperformed CLIFF-based probabilistic streamflow forecasts for the winter forecasting periods of November, December, and January, the latter forecasts performed better for the forecasting period of February. Thus, the two probabilistic forecasts are complementary. Although the results are limited to the study area, it has general application for evaluating the utility of different rainfall forecasts in providing deterministic/probabilistic streamflow forecasts. **DOI: 10.1061/(ASCE)WR.1943-5452.0001571.**

Author keywords: Seasonal rainfall; Climate reforecasts for Florida (CLIFF); Subseasonal to seasonal forecasts; Water supply.

Introduction

Over the last 20 years, operational seasonal hydrologic forecasts of rainfall and streamflow have gained increasing attention in the research community and water practitioners (Gong et al. 2010; Alemu et al. 2011; Lu et al. 2017; Sikder et al. 2016; Lopez and Haines 2017). Much progress has been made to address issues such as how to incorporate streamflow uncertainty into water resources management and how to best interpret probabilistic outcomes for decision-making (Zhao et al. 2011, 2012; Chen et al. 2016; Turner et al. 2017), which are considered as major challenges in the use of operational streamflow forecasts (Pagano et al. 2014; Wang et al. 2015, 2020). Operational hydrologic forecasting has found widespread application because of several reasons: (1) improved skills of seasonal and subseasonal climate/weather forecasting models (Kirtman et al. 2014; Vitart et al. 2017), (2) availability of forecast products in near-real-time (Kirtman et al. 2014; Bhardwaj et al. 2021), (3) advances in computational resources for both climate and hydrologic modeling (Demargne et al. 2014; Maidment 2016), and (4) the need to explore what-if scenarios to understand how a water supply system performs under uncertain streamflow conditions. (e.g., Wang et al. 2020).

In many cases, it is challenging to quantify the value of operational forecasts in terms of economic benefit (Chiew et al. 2003). Instead, it can be evaluated in terms of the potential enhancement of the system performance of a water supply system. For instance, the application of operational forecasts could potentially lead to enhanced system reliability, reduced water deficit, lower flooding risk, and increased hydropower generation, among others. Case studies in the literature have shown the benefit of seasonal streamflow forecasts in irrigation management (Chiew et al. 2003), drought management for water supply (Golembesky et al. 2009), optimal reservoir operation for different objectives (Alemu et al. 2011; Block 2011; Wang and Liu 2013; Steinschneider and Brown 2012;

[1]Principal Water Resources System Engineer, Dept. of System Decision Support, Tampa Bay Water, 2575 Enterprise Rd., Clearwater, FL 33763 (corresponding author). Email: hwang@tampabaywater.org

[2]Manager, Dept. of System Decision Support, Tampa Bay Water, Clearwater, FL 33763; Courtesy Professor, Patel College of Global Sustainability, Univ. of South Florida, Tampa, FL 33620. Email: tasefa@tampabaywater.org

[3]Professor, Dept. of Earth, Ocean, and Atmospheric Science, Florida State Univ., Tallahassee, FL 32306; Center for Ocean-Atmospheric Prediction Studies, Florida State Univ., Tallahassee, FL 32306; Florida Climate Institute, Florida State Univ., Tallahassee, FL 32306. Email: vmisra@fsu.edu

[4]Researcher, Center for Ocean-Atmospheric Prediction Studies, Florida State Univ., Tallahassee, FL 32306; Florida Climate Institute, Florida State Univ., Tallahassee, FL 32306; Numerical Weather Prediction, India Meteorological Department, Ministry of Earth Sciences, New Delhi 110003, India. Email: amit.bhardwaj.iitd@gmail.com

Note. This manuscript was submitted on November 2, 2021; approved on March 4, 2022; published online on April 6, 2022. Discussion period open until September 6, 2022; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

Ashbolt and Perera 2018) and facility maintenance scheduling (Wang et al. 2020). Insights from those studies suggest that the skills of operational streamflow forecasts are critical in supporting decision-making.

Efforts to improve operational forecasts can be summarized into two categories: (1) improved hydrologic models, e.g., model structure (Siddique and Mejia 2017), model parameters (Vrugt et al. 2006), and model averaging techniques (Schepen and Wang 2015) in simulating the hydrologic response; and (2) enhanced model input, e.g., initial soil moisture conditions and rainfall forecasts (Harpold et al. 2017; Oubeidillah et al. 2019). Putting hydrologic modeling issues aside, the skill of rainfall forecasts often dictates the skill of streamflow forecasts (Cuo et al. 2011), at least in rainfall-dominated river systems (Pagano et al. 2014). The study reported in this paper falls into the second category.

Due to improved forecasting skills from general circulation models (GCMs) and regional climate models (RCMs), output from those models have been employed to improve operational streamflow forecast. Using statistically downscaled output from GCMs, Landman et al. (2001) developed real-time categorical streamflow forecasts at the inlets of 12 dams of the Vaal and Upper Tugela River catchments in South Africa. Block et al. (2009) integrated GCMs, multiple RCMs, and numerous hydrologic models to improve streamflow forecasting for the Jaguaribe Basin in northeastern Brazil. Sikder et al. (2016) used the rainfall output from the North American multimodel ensemble (NMME) to develop streamflow forecasts for two large river basins at multiple time scales ranging from monthly to annual. Retrospective gridded rainfall from European Centre/Hamburg Model (ECHAM) version 4.5 has been used in different studies to develop streamflow forecasts with a lead time of up to 6 months (Sinha and Sankarasubramanian 2013; Wang and Liu 2013; Wang et al. 2015). Additional treatment of outputs from GCMs such as statistical downscaling required owing to their coarse spatial resolution before it is used in developing ensemble streamflow forecasts. Such treatments may be avoided if GCMs outputs are produced at a finer spatial resolution that could be applied directly to developing streamflow forecasts.

The motivation of this study was to fine-tune seasonal streamflow forecasts for decision-making at Tampa Bay Water, a regional water supply agency on the west coast in Florida. The existing streamflow forecasts used by Tampa Bay Water utilize National Oceanic and Atmospheric Administration (NOAA) rainfall forecasts (Wang et al. 2020). Promising alternatives to NOAA rainfall forecasts would provide water practitioners with flexibility in developing operational forecasts and potentially lead to improved streamflow forecasts. This study is an outcome of a knowledge coproducing research project through collaboration among climate scientists and water practitioners under the Florida Water and Climate Alliance (Misra et al. 2021). The sustained collaborative effort among scientists and practitioners resulted in coproducing a high-resolution experimental winter seasonal climate reforecasts for Florida (CLIFF) at 10-km grid spacing (Bhardwaj et al. 2021). The model was configured for the Florida region. For example, CLIFF ensemble members are uniquely designed to sample uncertainty in the lateral boundary conditions that forces the regional model and samples the uncertainty in the parameterizations used in the regional model. In essence, CLIFF attempts to best resolve the noise in the forecast system from its 30 ensemble members. This resolution of the noise at 10-km grid resolution is important to avoid or mitigate erroneous confidence in wrong forecasts.

The CLIFF archive carries the output of climate variables for 20 years of retrospective forecasts of the winter season, i.e., November, December, January, and February (NDJF), from 2000 to 2019. Ensemble rainfall forecasts were tested by Tampa Bay Water, a regional water supply agency, to develop ensemble streamflow forecasts that will be used for seasonal water resources allocation. This paper reports the evaluation of the CLIFF rainfall forecasts in developing streamflow forecasts. The objectives of this study were to (1) develop an evaluation framework assessing streamflow forecasting scheme that uses different sets of rainfall forecasts; and (2) evaluate and compare forecasting skills of the CLIFF-based streamflow forecasts and streamflow forecasts based on other types of rainfall inputs, which include rainfall climatology and NOAA rainfall forecasts. The NOAA rainfall forecasts are being currently used by Tampa Bay Water to develop decision-making tools.

This paper is arranged as follows. Right after the "Introduction" section, the "Data and Methodology" section briefly describes the study area, the CLIFF model, the rainfall-driven streamflow forecasting model, and the streamflow evaluation framework. Evaluation criteria for both the deterministic forecasts and probabilistic forecasts are also introduced. In the "Results and Discussion" section, evaluation results of streamflow forecasts for two surface river flows are offered. Implications of the CLIFF-based streamflow forecasts for the dry season are also discussed. Finally, concluding remarks are provided in the "Conclusions" section.

Data and Methodology

Study Area

Tampa Bay Water, the largest wholesale water provider in the southeast US, provides drinking water to its six member governments; three cities including New Port Richey, St. Petersburg, and Tampa; and three counties including Hillsborough, Pasco, and Pinellas. The total service population is about 2.5 million residents. Over the last 20 years, the agency has built an integrated water supply system that includes a surface water system, groundwater wells, and a seawater desalination plant. This has enabled the agency to shift from being 100% reliant on groundwater to a mixture of sources with an increasing reliance on surface waters (Wang et al. 2020). This shift to having substantial surface water in the agency's supply portfolio has resulted in increased hydrologic risks. This necessitated the development and implementation of seasonal forecasting tools.

Tampa Bay Water has water use permits from two surface rivers, namely the Hillsborough and Alafia Rivers (Fig. 1). The normal dry season extends from October through May. Streamflow at both rivers has strong seasonality, with the highest monthly flows in the summer rainy season and the lowest monthly flows typically in the spring dry season. Seasonal multivariate linear regression models (Asefa et al. 2014; Wang et al. 2020) are currently used to generate monthly flow for the two primary surface water sources given rainfall forecast at three rainfall stations, including Saint Leo, Cypress Creek, and Plant City (Fig. 1). Details of the streamflow simulation model have been given by Asefa et al. (2014).

Coproduction of Seasonal Climate Reforecasts for Florida

As part of a 3-year National Aeronautics and Space Administration (NASA) stakeholder-scientist coproduction pilot project, customized high-resolution experimental winter seasonal climate reforecasts for Florida were produced for the years 2000–2019 (Bhardwaj et al. 2021). There are two major features of these winter seasonal reforecasts. First, these are dynamically downscaled from an atmospheric general circulation model (AGCM) run at the spectral truncation of T62 (~210 grid spacing) (Kanamitsu et al. 2002; Misra et al. 2014). Second, the regional atmosphere model, namely

J. Water Resour. Plann. Manage., 2022, 148(6): 04022029

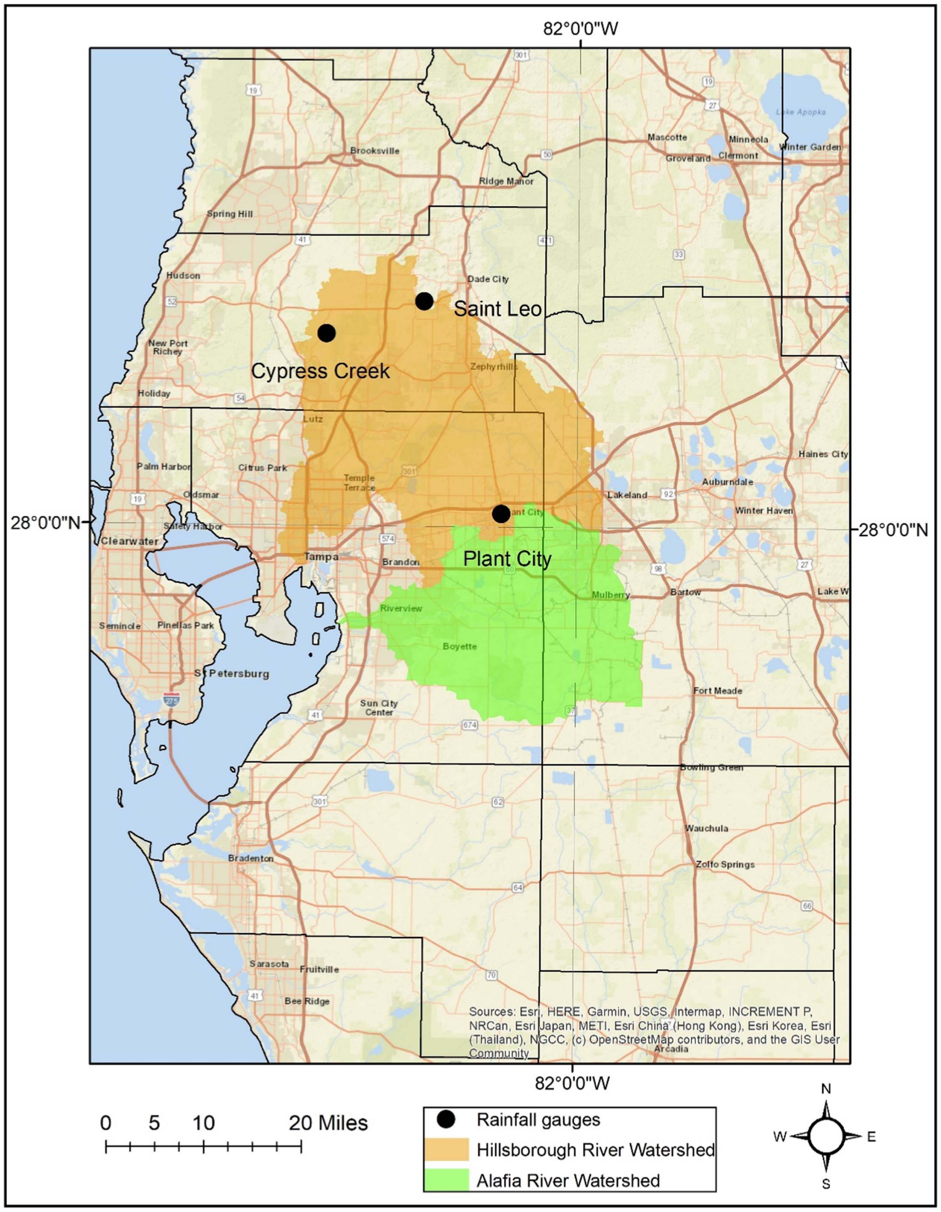


Fig. 1. Two watersheds in the study area and the three rainfall gauges that are used in generating operational streamflow forecasts. [Sources: Esri, HERE, Garmin, USGS, Intermap, INCREMENT P, NRCan, Esri Japan, METI, Esri China (Hong Kong), Esri Korea, Esri (Thailand), NGCC, © OpenStreetMap contributors, and the GIS User Community.]

the Regional Spectral Model (Juang and Kanamitsu 1994; Misra et al. 2019), was forced with sea surface temperatures (SSTs) obtained from one of the global models in the North American NMME. The horizontal resolution of CLIFF was set at 10-km grid spacing, and it provides 30-member ensemble forecasts for the winter NDJF months. Details of the CLIFF setup and model configurations of the AGCM and Regional Spectral Model (RSM) have been given by Bhardwaj et al. (2021). CLIFF outputs were prepared in the Network Common Data Form (NETCDF) format. Although results reported in this study are based on retrospective forecasts using CLIFF, the model is in place to produce operational dry-season rainfall forecasts for Florida.

Evaluation Framework of Streamflow Forecasts

Besides CLIFF rainfall forecasts, there were three other candidates of rainfall data considered in this study, including rainfall climatology, NOAA forecasts, and observed rainfall. The rainfall climatology for a particular month was obtained by averaging monthly total rainfall received at each station over the years 1991–2020. Retrospective NOAA rainfall forecasts were obtained from the Climate Prediction Center. Categorical forecasts, such as above normal, normal, and below-normal rainfall for the study region, were obtained from the forecast archive available at the Climate Prediction Center (Monthly and Seasonal Forecast Archive 2021). Forecasts for the NDJF season issued on November 1 were used to ensure a fair comparison with CLIFF.

Categorical rainfall forecasts were then converted to probabilistic forecasts using a Hidden Markov chain model (Wang et al. 2020) with a predetermined occurrence of different states in the winter months. Because NOAA ensemble precipitation forecasts were not available in this study, NOAA-based categorical precipitation was postprocessed via a hidden Markov chain model to derive the corresponding ensemble precipitation forecasts. If NOAA ensemble precipitation forecasts were available or different

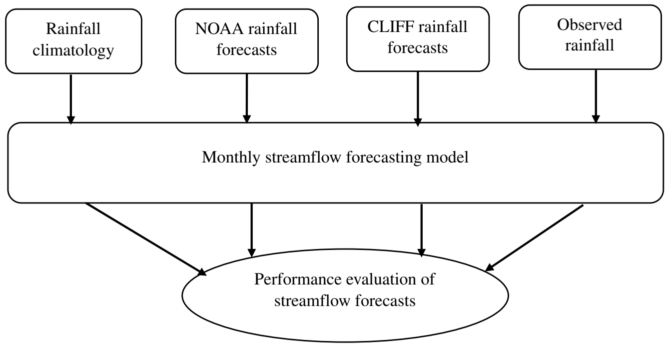


Fig. 2. Evaluation framework used in this study to examine the performance of streamflow forecasts based on different rainfall inputs.

postprocessing techniques were applied, then the results reported in this study and the conclusions drawn could be different. Rainfall climatology was used as input to the streamflow, the output of which represented the lower bound for the streamflow forecasting skills. The observed rainfall was used as input to the streamflow model, the output of which represented the upper bound of the streamflow forecasting skills. These extreme theoretical conditions binding the streamflow forecasting skills were included in this study for comparison purposes.

Fig. 2 describes the framework used in this study to compare streamflow forecasting skills based on different rainfall inputs. For each winter season during the years 2000–2019, four different sets of rainfall forecasts for the NDJF season were used to develop streamflow forecasts. The streamflow forecasting model is a monthly scale seasonal multilinear regression model, which has been used to provide monthly operational forecasts for the study area (Asefa et al. 2014; Wang et al. 2020). Asefa et al. (2014) has provided a detailed description, the mathematical formulation of these techniques, and their validation for the study area. For each rainfall input, 1,000 ensemble members of stochastic monthly flow forecasts were generated for each month of the NDJF season during the years 2000–2019. The ensemble members of the streamflow forecasts were then used to derive probabilistic forecasts. At the same time, ensemble median values were treated as deterministic forecasts.

Different evaluation metrics were applied in this study to evaluate the performance of deterministic/probabilistic forecasts. For deterministic forecasts, Spearman correlation between forecasts and observed streamflow for each month in the winter season was calculated to evaluate if the forecasts capture the interannual variabilities. The mean absolute percent error (MAPE) was defined to recognize the difference in magnitude of the streamflow values in distinct months, as shown in Eqs. (1) and (2). For simplicity, an index for different months in the winter season is skipped, and the equations were applied to calculate the MAPE for each of the forecasts

$$e_{t,j} = \frac{|f_{t,j} - o_t|}{o_t} \times 100\% \tag{1}$$

$$\overline{e_j} = \frac{1}{n}\sum_{t=2000}^{t=2019} e_{t,j} \tag{2}$$

where $f_{t,j}$ = ensemble mean of streamflow forecasts for year t ($t = 2000, 2001, \ldots, 2019$) from the jth rainfall input ($j = 1$, 2, 3, or 4, representing the four different rainfall forecasts in Fig. 2); o_t = observed streamflow for year t; $e_{t,j}$ = absolute percent error of streamflow forecasts for year t ($t = 2000, 2001, \ldots, 2019$) from the jth rainfall input; n ($n = 20$) = total number years; and $\overline{e_j}$ = mean absolute percent error of streamflow forecasts using the jth rainfall input. Essentially, the MAPE metric normalizes the deviation of the forecasts from the observed streamflow.

To evaluate probabilistic forecasts derived from the 1,000 ensemble members at each month, rank probability skill (RPS) was used to evaluate the tercile forecasts (Wilks 2006). Streamflow data between 1991 and 2020 were used to calculate the 34th and 67th percentile values. Streamflow within the range of 34th and 67th percentiles was defined as normal; flow less than the 34th percentile fell into the below-normal category, and the flow above the 67th percentile belonged to the above-normal flow category. Ensemble forecasts were first used to calculate the probability of each flow category, which were then used to calculate the RPS. RPS essentially summarizes the sum of the square of errors in the cumulative probabilities between a given categorical forecast and observation.

To compare RPS between streamflow forecast derived from different rainfall inputs, the rank probability skill score (RPSS) was used. In calculating the RPSS for different probabilistic forecasts, streamflow derived from rainfall climatology was used as the baseline forecasts. Both evaluation metrics, RPS and RPSS, are widely used in the literature (Wilks 1995; Wang et al. 2013; Devineni et al. 2008; Yuan and Wood 2012). A positive RPSS indicates the candidate model outperforms the reference model, and the larger the RPSS value is, the better the candidate model performs. Wang et al. (2013) has given a detailed description of RPS and RPSS.

Results and Discussion

Fig. 3 shows the comparison of forecasted streamflow and observation at the Alafia River for the month of November during the years 2000–2019. All four sets of streamflow forecasts generally captured the interannual variabilities of November streamflow. Fig. 3(a) shows the boxplots of observed rainfall driven streamflow forecasts compared with observed streamflow. Observed streamflow was within the range of the 25th to the 75th percentiles of the boxplots for most of the years. This is better than the streamflow forecasts based on rainfall climatology [Fig. 3(b)], CLIFF [Fig. 3(c)], and NOAA operational forecasts [Fig. 3(d)], constituting the upper bound on forecast accuracy. For instance, only the interquartile range of the observed rainfall-based streamflow forecasts contained observed streamflow for November 2014, which was one of the high flow months.

A couple of observations can be made based on Fig. 3. First, NOAA and CLIFF-based streamflow forecasts are generally better than climatology. Second, CLIFF-based forecasts seem to be more tightly bounded than NOAA-based forecasts, but show a tendency to underpredict, November 2014 for example. This will be discussed in subsequent sections.

Fig. 4 displays a comparison between observed streamflow and different rainfall-based streamflow forecasts for February in the years 2000–2019. The forecasting lead is 3 months because all retrospective forecasts were issued at the beginning of November each year. Hence, any streamflow forecasting errors in the prior months could propagate to later months. Even with observed rainfall in the NDJF months, there are times that observed values were out of the interquartile range of the boxplots [Fig. 4(a)]. The climatology-based forecasts [Fig. 4(b)] overforecast for a few years, where observed flow was less than the 25th percentile of the forecasted values. Similar observations can be seen for the NOAA-based forecasts, where the magnitudes of the forecasted values

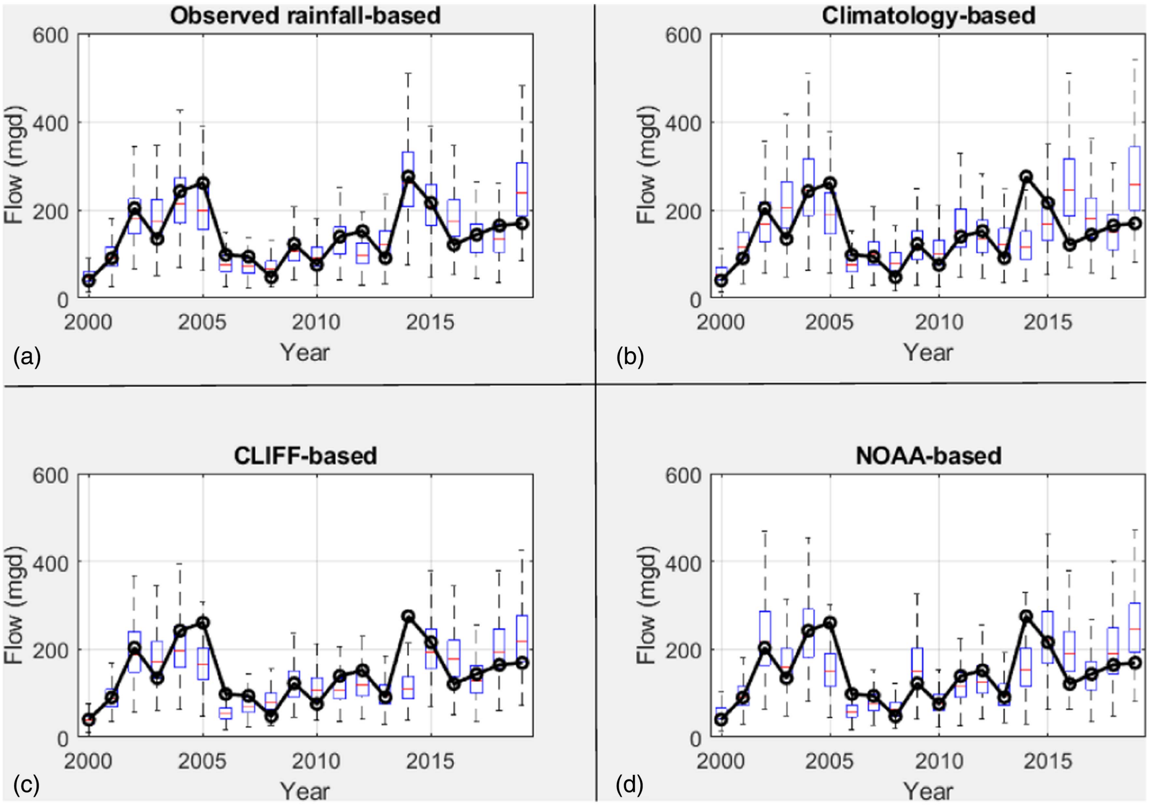


Fig. 3. Comparison between ensemble streamflow forecasts and observed data of the Alafia River for the month of November over the years of the forecasting period of 2000–2019: (a) observed rainfall-based streamflow forecasts; (b) climatology rainfall-based forecasts; (c) CLIFF-based streamflow forecasts; and (d) NOAA-based forecasts.

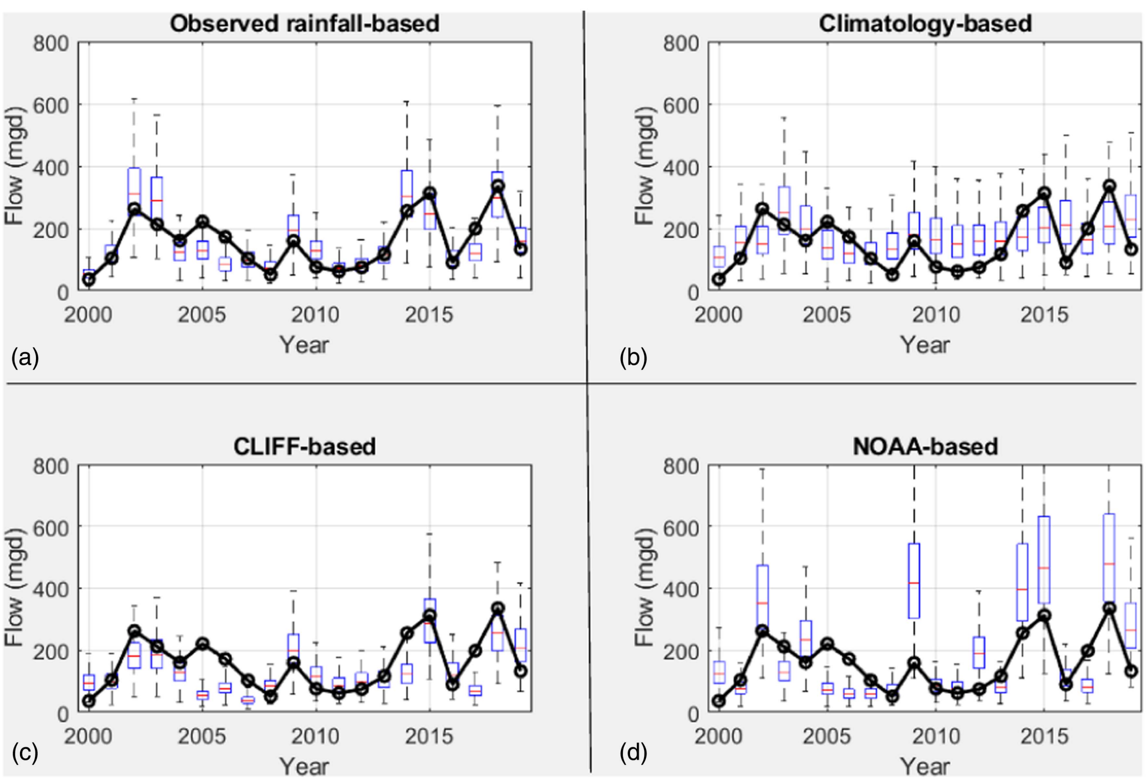


Fig. 4. Comparison between ensemble streamflow forecasts and observed data of the Alafia River for the month of February over the years of the forecasting period of 2000–2019: (a) observed rainfall-based streamflow forecasts; (b) climatology rainfall-based forecasts; (c) CLIFF-based streamflow forecasts; and (d) NOAA-based forecasts.

Table 1. Evaluation of winter-season NDJF streamflow forecasts based on different rainfall inputs

Streamflow forecasts	Forecasting month	Correlation		MAPE (%)		RPSS	
		Alafia	Hillsborough	Alafia	Hillsborough	Alafia	Hillsborough
Observed rainfall-based forecasts	November	0.87	0.83	21.02	22.49	0.36	0.39
	December	0.98	0.99	16.64	20.40	0.65	0.50
	January	0.87	0.88	28.41	28.53	0.69	0.46
	February	0.84	0.85	25.78	24.83	0.73	0.74
CLIFF-based forecasts	November	0.66	0.45	25.36	31.62	0.44	0.33
	December	0.64	0.67	28.19	38.16	0.48	0.39
	January	0.54	0.46	37.18	39.07	0.53	0.42
	February	0.62	0.59	44.44	40.88	0.38	0.51
NOAA-based forecasts	November	0.73	0.52	21.50	33.46	0.64	0.50
	December	0.70	0.66	31.71	49.08	0.79	0.62
	January	0.67	0.63	55.49	60.46	0.74	0.64
	February	0.71	0.79	64.39	62.67	−0.48	0.26
Climatology-based forecasts	November	0.56	0.42	30.80	31.92	N/A	N/A
	December	0.30	0.32	43.16	51.58	N/A	N/A
	January	0.20	0.21	60.32	62.38	N/A	N/A
	February	0.41	0.55	65.02	74.06	N/A	N/A

Note: Correlation and MAPE are used to evaluate the performance of deterministic forecasts by comparing them with the observed streamflow. RPSS is used to evaluate the performance of tercile forecasts using the climatology-based forecasts as a reference model, the RPSS of which is hence not available (N/A).

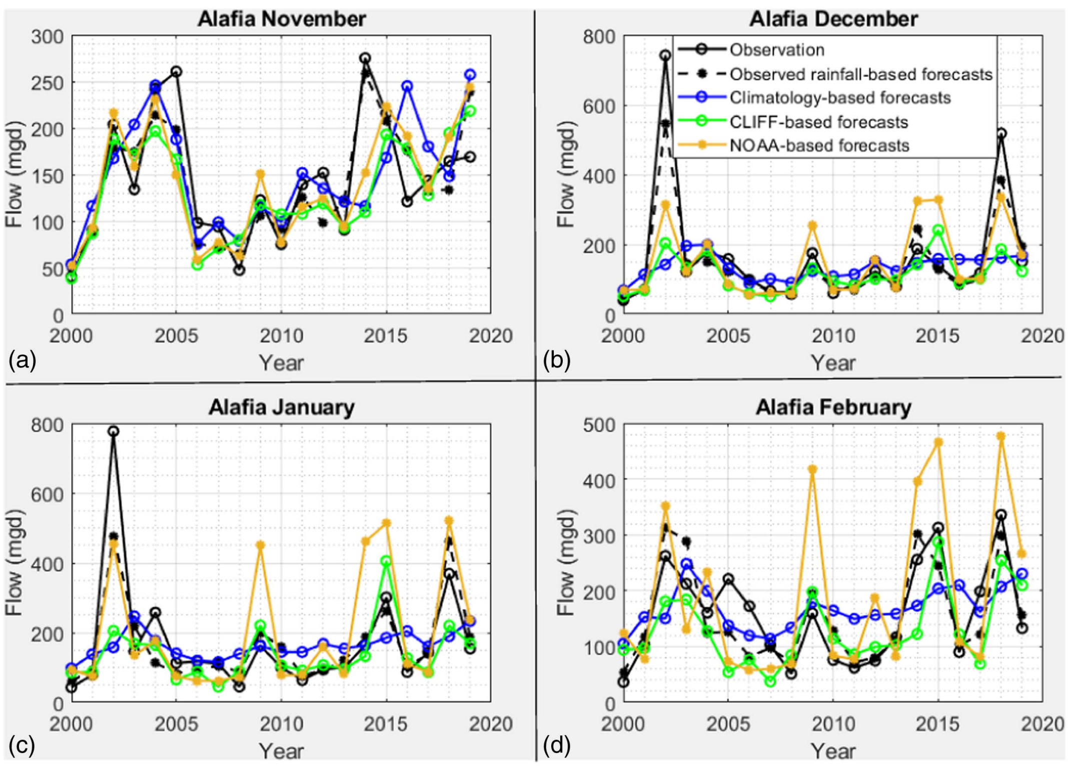


Fig. 5. Comparison between the median of ensemble streamflow forecasts and observed data of the Alafia River for the 4 months of (a) November; (b) December; (c) January; and (d) February over the years of the forecasting period of 2000–2019.

were much higher than observed flow for a few years [Fig. 4(d)]. On the other hand, underforecasting of the CLIFF-based forecasts [Fig. 4(c)] was observed for years including 2005–2007 and 2016. Comparison between different streamflow forecasts and observed flow at the Morris Bridge gauge station at the Hillsborough River show edsimilar results (figures not shown, but evaluation criteria are given in Table 1).

Median forecasts, corresponding to the median value of the boxplots shown in the preceding figures, were used here as deterministic forecasts for comparison purposes. Figs. 5 and 6 display the comparison between median forecasts and observation for all forecasting periods at the two river gauges. A few notable observations can be drawn based on Fig. 5. First, the observed rainfall-based streamflow forecasts had the least deviation from observed

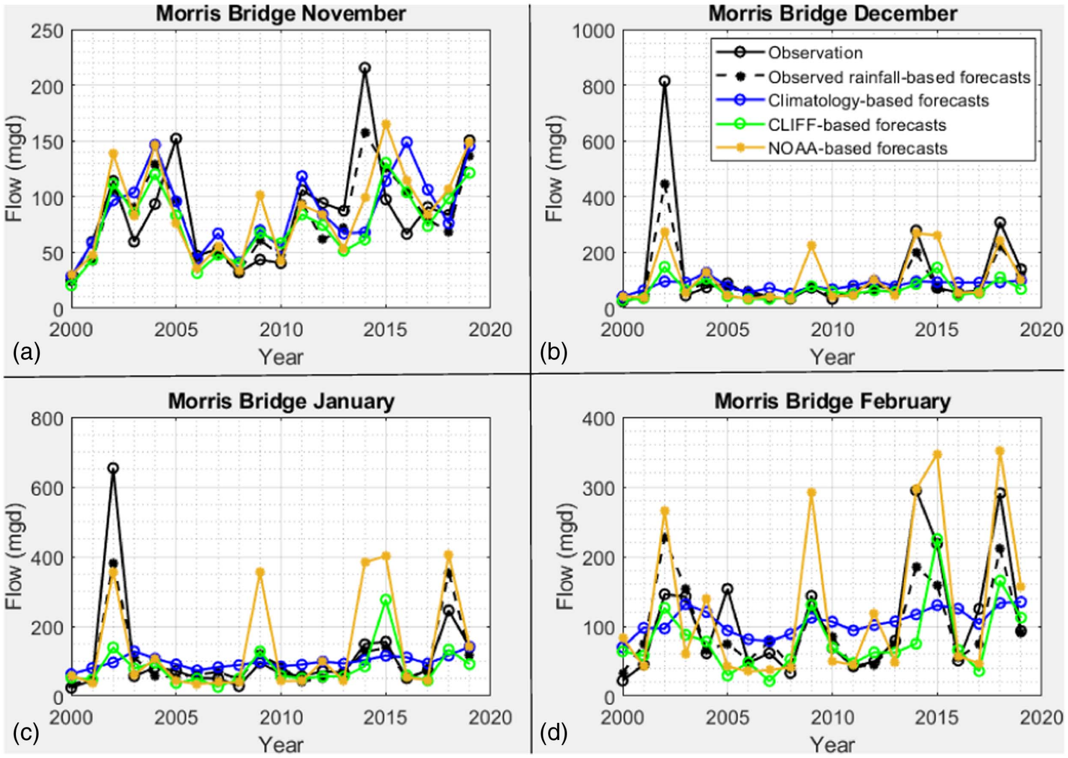


Fig. 6. Comparison between the median of ensemble streamflow forecasts and observed data at the Morris Bridge gauge of the Hillsborough River for the 4 months of (a) November; (b) December; (c); January; and (d) February over the years of the forecasting period of 2000–2019.

streamflow. This was true for all the forecasting lead times ranging from zero to 3 months for both river gauges. The benefit of accurate rainfall forecasts can be clearly seen at times when high flow occurred, such as November 2014 and December and January 2002.

Second, the performance of the streamflow forecasts decayed as the lead time of the forecast increased. This was noted from the increase in the deviation between the forecasts and the observed data in later months from the start of the forecast, e.g., errors in February were larger compared with prior months. Overforecasting for the NOAA-based forecasts in January and February for several years can clearly be observed. For instance, when a low-flow condition was observed in November 2009 at the Morris Bridge gauge station, the median of NOAA-based forecasts was higher than the observed flow. The gap between the two increased in December 2009 and was further enhanced in January and February of 2010. This is due to error propagation of rainfall forecasts and streamflow forecasts from the prior months.

Fig. 7 shows Spearman correlation between observed streamflow and different forecasts for Alafia River [Fig. 7(a)] and Hillsborough River [Fig. 7(b)]. Correlation values are given in Table 1. It was highest for the observed rainfall-based forecasts, ranging between 0.82 and 0.95. It was lowest for the climatology-based forecasts. NOAA-based forecasts had a higher correlation with observed data than the correlation between CLIFF-based forecast observations. This is probably due to the strong underforecasting of large events for the CLIFF-based forecasts. There was also a difference in the forecasts for the two river gauges. For both the CLIFF-based and NOAA-based forecasts, correlation values were higher for the Alafia River than for the Hillsborough River. This is driven by the difference in the physical hydrologic process at the two watersheds. The Hillsborough River watershed is more conditionally connected and has a relatively larger watershed memory than the Alafia River. Hence, streamflow at the Alafia River is more dominated by rainfall amount in terms of runoff generation mechanisms.

Fig. 8 compares the MAPE between different streamflow forecasts for both flow gauges, the values of which are given in Table 1.

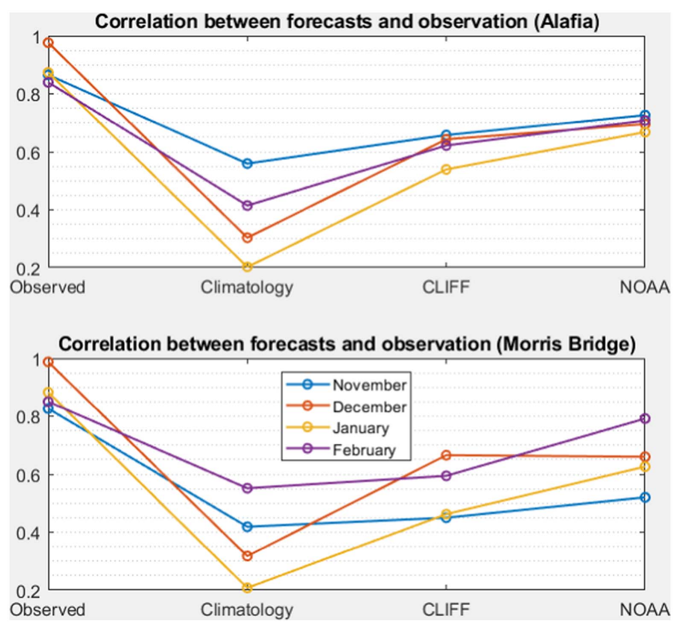


Fig. 7. Correlation between the median of ensemble forecasts and observed streamflow for the 4 months of November, December, January, and February over the years of the forecasting period of 2000–2019 at both the Alafia River and Hillsborough River.

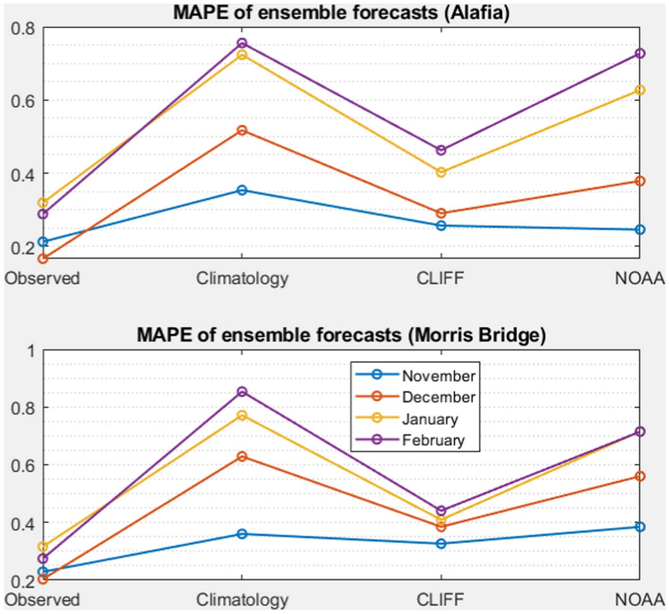


Fig. 8. MAPE of ensemble forecasts for the 4 months of November, December, January, and February over the years of the forecasting period of 2000–2019 at both the Alafia River and Hillsborough River.

The lower the MAPE value, the smaller the discrepancy between forecasts and observed values. Consistent with what has been revealed in Fig. 7, observed rainfall-based forecasts had the lowest MAPE, whereas climatology-based forecasts had the largest. Forecasting errors were lower in the first two forecasting periods, i.e., November and December, compared with January and February. This is consistent with what has been found previously in examining the boxplots. CLIFF-based forecasts, however, have lower forecasting error compared with the NOAA-based forecasts. The difference in the MAPE values for the first forecasting period, i.e., November, was smaller compared with later months.

To further evaluate different forecasts, all ensemble members were used to derive probabilistic forecasts as described in the "Methodology" section. Fig. 9 presents ensemble forecasts for the NOAA-based and CLIFF-based method for the forecasting period of February 2020 at the Hillsborough River. The histogram shows the relative frequency of different flow intervals, and the filled circle on the X-axis represents the observed value. The RPS for the NOAA-based and CLIFF-based forecasts was 0.44 and 0.10, respectively. The is primarily because the CLIFF-based forecast had a tighter distribution around the observed value. The RPSS can then be calculated using the climatology-based forecasts as the reference model. Unlike the RPS, a higher RPSS indicates a better forecasting skill. Given that retrospective forecasts were provided for each month in the NDJF season for the years 2000–2019, there were a total of 20 RPSS values for each forecasting period.

Boxplots of the RPSS values for the three different rainfall-based forecasts at the Alafia River and Hillsborough River are shown in Fig. 10. When the RPSS value is greater than zero, it indicates the candidate model performs better than the reference model, which is climatology-based streamflow forecasts. The median values of boxplots for all forecasting periods were above zero except the NOAA-based forecasts for February at Alafia River. As expected, the observed rainfall-based model performed the best with the largest median values for nearly all the forecasting periods. In addition, it had the smallest interquartile ranges, indicating the least interannual variabilities in its forecasting performance. The NOAA-based forecasts had greater RPSS than the CLIFF-based forecasts for the first three forecasting periods,

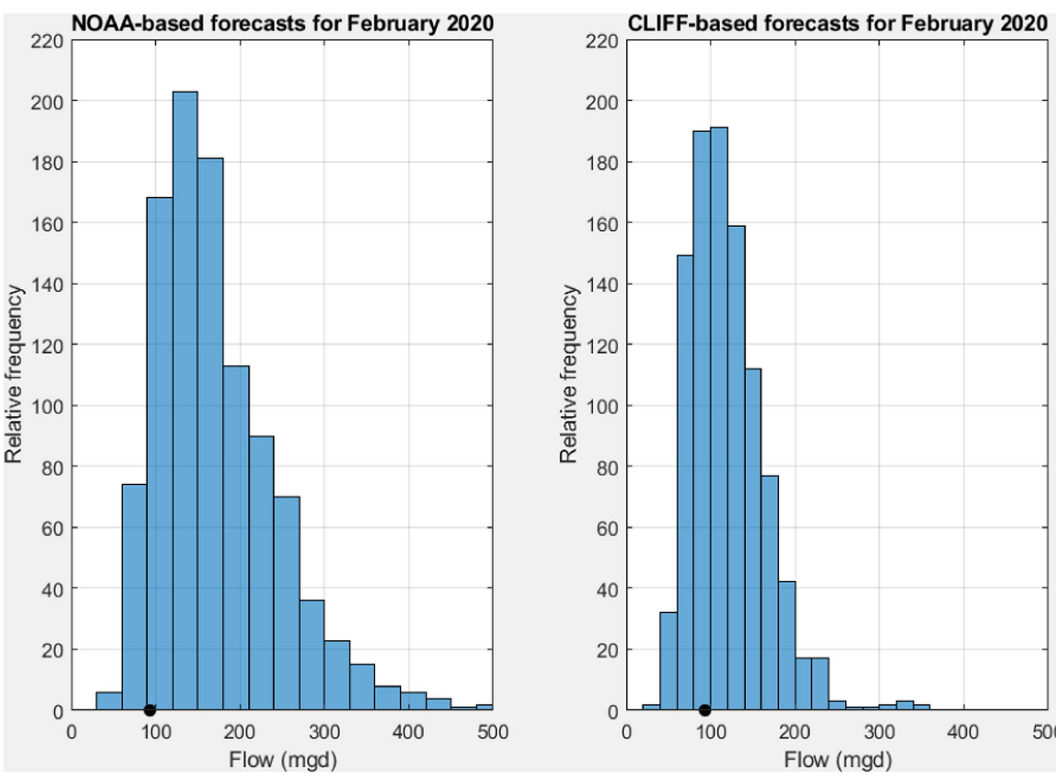
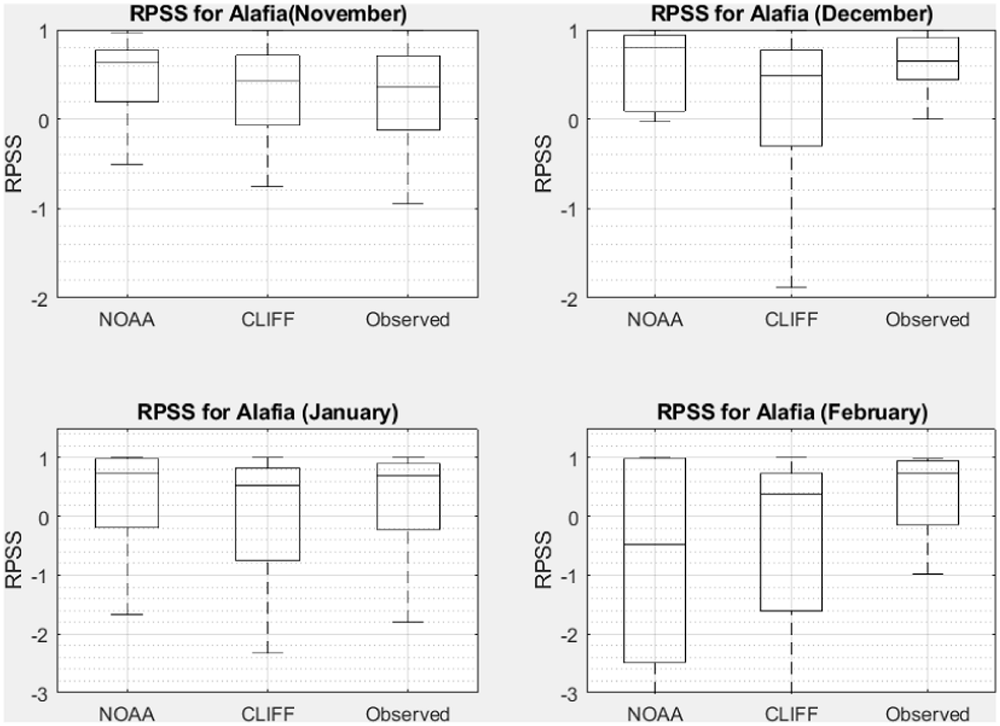
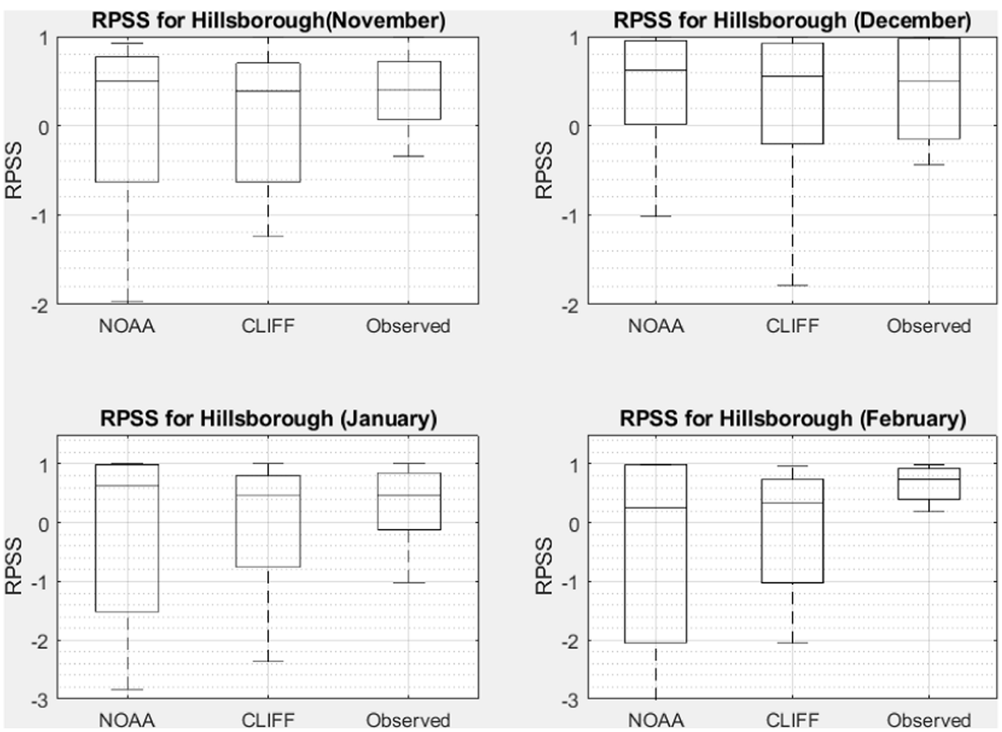


Fig. 9. NOAA-based and CLIFF-based streamflow forecasts for Morris Bridge gauge at the Hillsborough River for February 2020.



(a)



(b)

Fig. 10. Boxplots of RPSS for different ensemble forecasts at winter months for (a) Alafia River; and (b) Morris Bridge gauge at the Hillsborough River.

including November, December, and January, but excepting February. The median of RPSS values for all the forecasting periods are given in Table 1.

Accurate winter season streamflow forecasts, either deterministic or ensemble forecasts, have implications for water resources management. The NOAA-based model is currently used to facilitate decision-making at the seasonal time scale (Wang et al. 2020) and determine water shortage stages (Wang et al. 2019) to trigger potential water shortage mitigation measures. This study found that CLIFF-based forecasts are a promising candidate for winter seasonal flow forecasting. In terms of deterministic forecasts, although CLIFF-based forecasts have a smaller correlation with observed streamflow, they have smaller forecasting errors than the NOAA-based forecasts. In terms of probabilistic forecasts, the CLIFF-based forecasts have better RPSS than the NOAA-based forecasts for the month of February.

There are a few sources of the difference between streamflow forecasts based on the postprocessed NOAA categorical forecasts and streamflow forecasts based on CLIFF ensemble forecasts. The first is that CLIFF is a regional climate model that is configured and calibrated for the Florida region. For example, CLIFF ensemble members are uniquely designed to sample uncertainty in the lateral boundary conditions and uncertainty in the model parameterizations. The second is that postprocessing of the NOAA categorical forecasts might have enhanced the uncertainty of the ensemble precipitation, which further propagated to the streamflow forecasts. Evaluation of the utility of these forecasts in the decision-making processes of seasonal resources allocation is beyond the scope of this study because it requires a systems model that these streamflow forecasts feed into, which will be pursued in a separate study.

Conclusions

This study evaluated the potential use of rainfall forecasts from high-resolution experimental winter seasonal climate reforecasts that was coproduced through scientist-stakeholder group at Florida Water and Climate Alliance (Florida Water and Climate Alliance 2021). Through sustained interaction of its members, a pilot study was borne to produce customized and actionable climate forecasts, e.g., seasonal climate forecast data sets (CLIFF), whose utility were explored in this study. In this study, CLIFF rainfall forecasts were used to assess if improvements in operational streamflow forecasts could be achieved. The evaluation framework entailed using four distinct rainfall data sets, a streamflow generation scheme, and evaluation metrics for deterministic and probabilistic forecasts.

Four different sets of rainfall data were examined, including rainfall climatology, observed data, NOAA-based rainfall forecasts, and CLIFF rainfall forecasts. Because NOAA ensemble precipitation forecasts were not available in this study, NOAA-based categorical precipitation was postprocessed to obtain NOAA-based seasonal rainfall forecasts. Different evaluation metrics, including Spearman correlation, mean absolute percent error, and rank probability skill score, were applied to evaluate the performance of deterministic and probabilistic forecasts. The framework was applied on streamflow forecasts generated for the winter season, i.e., NDJF months, at the two surface water supply sources for Tampa Bay Water during the years 2000–2019. The NOAA-based rainfall is currently used in the study area to develop operational streamflow forecasts.

Results found that the performance of all retrospective forecasts decayed with forecasting period increases, whereas observed rainfall-based streamflow forecasts had the best performance, and the climatology-based had the lowest score. It was also found that CLIFF-based streamflow forecasts are a promising alternative to NOAA-based forecasts. The CLIFF-based streamflow forecasts, however, seem to underforecast large events. Deterministic forecasts of CLIFF-based streamflow had a smaller MAPE compared with the NOAA-based streamflow forecasts. Further study is needed to evaluate the use of operational streamflow forecasts in decision-making processes through systems-based models, which is outside the scope of this study. Although the results are limited to the study area, they have general application for evaluating the utility of different rainfall forecasts in providing deterministic/probabilistic streamflow forecasts and the value of coproducing customized actionable data sets.

Data Availability Statement

Some or all data, models, or code generated or used during the study are available from the corresponding author by request (including the rainfall output from the CLIFF model and code for simulating streamflow).

Acknowledgments

This work is partially supported by Grant NNX17AG72G from NASA. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the funding agency. The authors thank the editor, anonymous associate editor, and reviewers for the constructive comments that improved the manuscript.

References

Alemu, E. T., R. N. Palmer, A. Polebitski, and B. Meaker. 2011. "Decision support system for optimizing reservoir operations using ensemble streamflow predictions." *J. Water Resour. Plann. Manage.* 137 (1): 72–82. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000088.

Asefa, T., J. Clayton, A. Adams, and D. Anderson. 2014. "Performance evaluation of a water resources system under varying climatic conditions: Reliability, resilience, vulnerability and beyond." *J. Hydrol.* 508 (Jan): 53–65. https://doi.org/10.1016/j.jhydrol.2013.10.043.

Ashbolt, S. C., and B. J. C. Perera. 2018. "Multiobjective optimization of seasonal operating rules for water grids using streamflow forecast information." *J. Water Resour. Plann. Manage.* 144 (4): 05018003. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000902.

Bhardwaj, A., V. Misra, B. Kirtman, T. Asefa, C. Maran, K. Morris, E. Carter, C. Martinez, and D. Roberts. 2021. "Experimental high-resolution winter seasonal climate reforecasts for Florida." *Weather Forecasting* 36 (4): 1169–1182. https://doi.org/10.1175/WAF-D-21-0004.1.

Block, P. 2011. "Tailoring seasonal climate forecasts for hydropower operations." *Hydrol. Earth Syst. Sci.* 15 (4): 1355–1368. https://doi.org/10.5194/hess-15-1355-2011.

Block, P. J., F. A. Souza Filho, L. Sun, and H.-H. Kwon. 2009. "A streamflow forecasting framework using multiple climate and hydrological models." *JAWRA J. Am. Water Resour. Assoc.* 45 (4): 828–843. https://doi.org/10.1111/j.1752-1688.2009.00327.x.

Chen, L., V. P. Singh, W. Lu, J. Zhang, J. Zhou, and S. Guo. 2016. "Streamflow forecast uncertainty evolution and its effect on real-time reservoir operation." *J. Hydrol.* 540 (Sep): 710–726. https://doi.org/10.1016/j.jhydrol.2016.06.015.

Chiew, F. H. S., S. L. Zhou, and T. A. McMahon. 2003. "Use of seasonal streamflow forecasts in water resources management." *J. Hydrol.* 270 (1–2): 135–144. https://doi.org/10.1016/S0022-1694(02)00292-5.

Cuo, L., T. C. Pagano, and Q. J. Wang. 2011. "A review of quantitative precipitation forecasts and their use in short- to medium-range streamflow forecasting." *J. Hydrometeorol.* 12 (5): 713–728. https://doi.org/10.1175/2011JHM1347.1.

Demargne, J., et al. 2014. "The science of NOAA's operational hydrologic ensemble forecast service." *Bull. Am. Meteorol. Soc.* 95 (1): 79–98. https://doi.org/10.1175/BAMS-D-12-00081.1.

Devineni, N., A. Sankarasubramanian, and S. Ghosh. 2008. "Multimodel ensembles of streamflow forecasts: Role of predictor state in developing optimal combinations." *Water Resour. Res.* 44 (9): W09404. https://doi.org/10.1029/2006WR005855.

Florida Water and Climate Alliance. 2021. "Uncertainties and risks associated with climate change, climate variability, and sea level rise pose complex challenges to Florida's water sector." Accessed October 1, 2021. http://www.floridawca.org/.

Golembesky, K., A. Sankarasubramanian, and N. Devineni. 2009. "Improved drought management of falls lake reservoir: Role of multimodel streamflow forecasts in setting up restrictions." *J. Water Res. Plann.*

Manage. 135 (3): 188–197. https://doi.org/10.1061/(ASCE)0733-9496(2009)135:3(188).

Gong, G., L. Wang, L. Condon, A. Shearman, and U. Lall. 2010. "A simple framework for incorporating seasonal streamflow forecasts into existing water resource management practices." *JAWRA J. Am. Water Resour. Assoc.* 46 (3): 574–585. https://doi.org/10.1111/j.1752-1688.2010.00435.x.

Harpold, A. A., K. Sutcliffe, J. Clayton, A. Goodbody, and S. Vazquez. 2017. "Does including soil moisture observations improve operational streamflow forecasts in snow-dominated watersheds?" *JAWRA J. Am. Water Resour. Assoc.* 53 (1): 179–196. https://doi.org/10.1111/1752-1688.12490.

Juang, H.-M., and M. Kanamitsu. 1994. "The NMC nested regional spectral model." *Mon. Weather Rev.* 122 (1): 3–26. https://doi.org/10.1175/1520-0493(1994)122<0003:TNNRSM>2.0.CO;2.

Kanamitsu, M., et al. 2002. "NCEP Dynamical Seasonal Forecast System 2000." *Bull. Am. Meteorol. Soc.* 83 (7): 1019–1038. https://doi.org/10.1175/1520-0477(2002)083<1019:NDSFS>2.3.CO;2.

Kirtman, B. P., et al. 2014. "The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction." *Bull. Am. Meteorol. Soc.* 95 (4): 585–601. https://doi.org/10.1175/BAMS-D-12-00050.1.

Landman, W. A., S. J. Mason, P. D. Tyson, and W. J. Tennant. 2001. "Statistical downscaling of GCM simulations to streamflow." *J. Hydrol.* 252 (1–2): 221–236. https://doi.org/10.1016/S0022-1694(01)00457-7.

Lopez, A., and S. Haines. 2017. "Exploring the usability of probabilistic weather forecasts for water resources decision-making in the United Kingdom." *Weather Clim. Soc.* 9 (4): 701–715. https://doi.org/10.1175/WCAS-D-16-0072.1.

Lu, M., U. Lall, A. W. Robertson, and E. Cook. 2017. "Optimizing multiple reliable forward contracts for reservoir allocation using multitime scale streamflow forecasts." *Water Resour. Res.* 53 (3): 2035–2050. https://doi.org/10.1002/2016WR019552.

Maidment, D. R. 2016. "Open water data in space and time." *JAWRA J. Am. Water Resour. Assoc.* 52 (4): 816–824. https://doi.org/10.1111/1752-1688.12436.

Misra, V., T. Irani, L. Staal, K. Morris, T. Asefa, C. Martinez, and W. Graham. 2021. "The Florida water and climate alliance (FloridaWCA): Developing a stakeholder–scientist partnership to create actionable science in climate adaptation and water resource management." *Bull. Am. Meteorol. Soc.* 102 (2): E367–E382. https://doi.org/10.1175/BAMS-D-19-0302.1.

Misra, V., H. Li, Z. Wu, and S. DiNapoli. 2014. "Global seasonal climate predictability in a two tiered forecast system. Part I: Boreal summer and fall seasons." *Clim. Dyn.* 42 (5): 1425–1448. https://doi.org/10.1007/s00382-013-1812-y.

Misra, V., A. Mishra, and A. Bhardwaj. 2019. "A coupled ocean-atmosphere downscaled climate projection for the peninsular Florida region." *J. Mar. Syst.* 194 (Jun): 25–40. https://doi.org/10.1016/j.jmarsys.2019.02.010.

Monthly and Seasonal Forecast Archive. 2021. "CPC monthly and seasonal forecast archive." Accessed October 1, 2021. https://www.cpc.ncep.noaa.gov/products/archives/long_lead/llarc.ind.php.

Oubeidillah, A., G. Tootle, and T. Piechota. 2019. "Incorporating antecedent soil moisture into streamflow forecasting." *Hydrology* 6 (2): 50. https://doi.org/10.3390/hydrology6020050.

Pagano, T. C., et al. 2014. "Challenges of operational river forecasting." *J. Hydrometeorol.* 15 (4): 1692–1707. https://doi.org/10.1175/JHM-D-13-0188.1.

Schepen, A., and Q. J. Wang. 2015. "Model averaging methods to merge operational statistical and dynamic seasonal streamflow forecasts in Australia." *Water Resour. Res.* 51 (3): 1797–1812. https://doi.org/10.1002/2014WR016163.

Siddique, R., and A. Mejia. 2017. "Ensemble streamflow forecasting across the US Mid-Atlantic region with a distributed hydrological model forced by GEFS reforecasts." *J. Hydrometeorol.* 18 (7): 1905–1928. https://doi.org/10.1175/JHM-D-16-0243.1.

Sikder, S., X. Chen, F. Hossain, J. B. Roberts, F. Robertson, C. K. Shum, and F. J. Turk. 2016. "Are general circulation models ready for operational streamflow forecasting for water management in the Ganges and Brahmaputra River Basins?" *J. Hydrometeorol.* 17 (1): 195–210. https://doi.org/10.1175/JHM-D-14-0099.1.

Sinha, T., and A. Sankarasubramanian. 2013. "Role of climate forecasts and initial conditions in developing streamflow and soil moisture forecasts in a rainfall–runoff regime." *Hydrol. Earth Syst. Sci.* 17 (2): 721–733. https://doi.org/10.5194/hess-17-721-2013.

Steinschneider, S., and C. Brown. 2012. "Dynamic reservoir management with real-option risk hedging as a robust adaptation to nonstationary climate." *Water Resour. Res.* 48 (5): W05524. https://doi.org/10.1029/2011WR011540.

Turner, S. W. D., J. C. Bennett, D. E. Robertson, and S. Galelli. 2017. "Complex relationship between seasonal streamflow forecast skill and value in reservoir operations." *Hydrol. Earth Syst. Sci.* 21 (9): 4841–4859. https://doi.org/10.5194/hess-21-4841-2017.

Vitart, F., et al. 2017. "The subseasonal to seasonal (S2S) prediction project database." *Bull. Am. Meteorol. Soc.* 98 (1): 163–173. https://doi.org/10.1175/BAMS-D-16-0017.1.

Vrugt, J. A., H. V. Gupta, B. Nualláin, and W. Bouten. 2006. "Real-time data assimilation for operational ensemble streamflow forecasting." *J. Hydrometeorol.* 7 (3): 548–565. https://doi.org/10.1175/JHM504.1.

Wang, H., T. Asefa, D. Bracciano, A. Adams, and N. Wanakule. 2019. "Proactive water shortage mitigation integrating system optimization and input uncertainty." *J. Hydrol.* 571 (Apr): 711–722. https://doi.org/10.1016/j.jhydrol.2019.01.071.

Wang, H., T. Asefa, N. Wanakule, and A. Adams. 2020. "Application of decision-support tools for seasonal water supply management that incorporates system uncertainties and operational constraints." *J. Water Resour. Plann. Manage.* 146 (6): 05020008. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001225.

Wang, H., E. D. Brill, R. S. Ranjithan, and A. Sankarasubramanian. 2015. "A framework for incorporating ecological releases in single reservoir operation." *Adv. Water Resour.* 78 (Apr): 9–21. https://doi.org/10.1016/j.advwatres.2015.01.006.

Wang, H., and J. Liu. 2013. "Reservoir operation incorporating hedging rules and operational inflow forecasts." *Water Resour. Manage.* 27 (5): 1427–1438. https://doi.org/10.1007/s11269-012-0246-3.

Wang, H., A. Sankarasubramanian, and R. S. Ranjithan. 2013. "Integration of climate and weather information for improving 15-day-ahead accumulated precipitation forecasts." *J. Hydrometeorol.* 14 (1): 186–202. https://doi.org/10.1175/JHM-D-11-0128.1.

Wilks, D. S. 1995. Vol. 59 of *Statistical methods in the atmospheric sciences: An introduction*. San Diego: Academic.

Wilks, D. S. 2006. *Statistical methods in the atmospheric sciences*. 2nd ed. London: Academic Press.

Yuan, X., and E. F. Wood. 2012. "Downscaling precipitation or bias-correcting streamflow? Some implications for coupled general circulation model (CGCM)-based ensemble seasonal hydrologic forecast." *Water Resour. Res.* 48 (12): W12519. https://doi.org/10.1029/2012WR012256.

Zhao, T., X. Cai, and D. Yang. 2011. "Effect of streamflow forecast uncertainty on real-time reservoir operation." *Adv. Water Resour.* 34 (4): 495–504. https://doi.org/10.1016/j.advwatres.2011.01.004.

Zhao, T., D. Yang, X. Cai, J. Zhao, and H. Wang. 2012. "Identifying effective forecast horizon for real-time reservoir operation under a limited inflow forecast." *Water Resour. Res.* 48 (1): W01540. https://doi.org/10.1029/2011WR010623.