Journal of Hydrology: Regional Studies 45 (2023) 101305

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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

# Calibration of ECMWF SEAS5 based streamflow forecast in Seasonal hydrological forecasting for Citarum river basin, West Java, Indonesia

Dian Nur Ratri [a,b,d,*], Albrecht Weerts [a,c,**], Robi Muharsyah [d], Kirien Whan [b], Albert Klein Tank [a,e], Edvin Aldrian [f], Mugni Hadi Hariadi [a,b,d]

[a] Wageningen University and Research, The Netherlands
[b] Koninklijk Nederlands Meteorologisch, The Netherlands
[c] Deltares, Delft, The Netherlands
[d] Meteorological, Climatological, and Geophysical, Indonesia
[e] Met Office Hadley Centre, United Kingdom
[f] The National Research and Innovation Agency, Indonesia

ARTICLE INFO

*Keywords:*
ECMWF SEAS5
Bias-corrected
Streamflow forecast verification/skill
Wflow_sbm model
Hydrological ensemble forecast
Citarum river basin

ABSTRACT

**Study region:** Citarum river basin, West Java — Indonesia.
**Study focus:** We look at the skill of Empirical Quantile Mapping corrected ECMWF SEAS5 (SEAS5 EQM bias-corrected) based streamflow forecasts in the Citarum river basin. We focus on July to October because these are agriculturally important months in Java. We use a high-resolution hydrologic model (wflow_sbm) with data for the period 1989–2009.
**New hydrological insights for the region:** Water users and agricultural practitioners commonly need monthly to seasonal hydrological forecasts. The forecasts should be sufficiently skillful and provide information that is relevant to the decisions makers in order to have practical value to them. We assess if skilful SEAS5 EQM bias-corrected based seasonal forecasts are available with the purpose to support rice production. In this streamflow forecast calibration, we look at different aggregation days and different lead times. For the verification, we use the Continuous Ranked Probability Skill Score (CRPSS), Brier Skill Score (BSS), and Mean Average Error (MAE). We also look at the correlation, the Root Mean Square Error (RMSE), and the Receiver operating characteristic Skill (ROCS). The LT1 and LT2 forecast show higher skills than longer lead times. Meanwhile, streamflow based on the aggregated forecast at 30 to 60 days aggregation days is more skillful than larger aggregations. In Indonesia, this study is a study that initiates using a hydrological model with inputs from a seasonal rainfall forecast.

## 1. Introduction

National development programs in Indonesia have led to increasing demand for water for uses such as safe drinking water and irrigation. The water demand in Jakarta and Bandung, two big cities in Indonesia depend on the Citarum river as its water resources (Loebis and Syariman, 1993; Fares and Yudianto, 2003; Sahu et al., 2017). Loebis and Syariman (1993) mentioned that 80% of the water demands in Jakarta are supplied by the Citarum river basin, the longest and largest river basin in West Java. This

* Correspondence to: Dept. of Environmental Sciences, Droevendaalsesteeg 3a, Gebouw 100, 6708 PB, Wageningen.
** Corresponding author at: Wageningen University and Research, The Netherlands.
*E-mail addresses:* dianuratri@gmail.com (D.N. Ratri), albrecht.weerts@wur.nl (A. Weerts).

https://doi.org/10.1016/j.ejrh.2022.101305
Received 2 February 2022; Received in revised form 9 December 2022; Accepted 12 December 2022
Available online 21 December 2022

river basin has a significant influence on the daily life of the population in the province of West Java, not only for electricity and water supply but also for agriculture, fisheries, and industrial sectors.

In the agricultural sector, the Indonesian government has a target of development to achieve self-sufficiency in rice production (Pasaribu, 2010). Adequacy of the supply of rice will increase food security for Indonesia. According to Suwarno (2016), one of the indicators of national food security is the availability of foodstuffs. The Indonesian Ministry of Agriculture encouraged rice farmers to increase rice productivity in rain-fed rice fields to ensure abundant national rice stock. To harmonize government programs, the water availability of rice fields needs to be considered. Rain-fed rice fields rely on rainfall, so farmers require skillful forecasts of rainfall and water availability. Meanwhile, in irrigated paddy fields, if the water is available throughout the year, farmers most likely can continuously run rice cultivation. However, this has a negative side in terms of the development of pests and diseases plant (Rusastra et al., 2019). Besides this problem, the farmers often face rice production risk and harvesting failure due to the occurrence of floods and droughts.

Farmers need to plan strategically to avoid risks associated with varying water availability. To support this, a skillful seasonal forecast of streamflow is needed. The streamflows are affected by not only morphological but also anthropogenic factors including topography and land-use changes besides rainfall. Sahu et al. (2012) mention that the relationship between river streamflow and variation of climate sometimes still unclear though the topography factor has a direct impact on rainfall variability. A timely forecast of climate-related natural hazards such as floods and droughts can help farmers and decision-makers to increase preparedness and reduce the impacts of these hazards. Adjusting planting time, selecting commodities, and knowing the climatic conditions before and during the growing season are important (Surmaini and Syahbuddin, 2016). In addition, a lot of research has been invested in the provision of skillful seasonal forecasts, which is motivated by the development of streamflow forecasts effort towards developing streamflow forecasting (Crochemore et al., 2016). For the adequate management of water resources systems, precise forecast streamflows are needed (Masselot et al., 2016). The streamflow forecasts' accuracy relies on a proper understanding of meteorological and hydrological uncertainties (Brown et al., 2012; Siddique et al., 2015). Zhao et al. (2011) mentioned that hydrological forecasting are sensitive to biases in atmospheric forcing. In addition, streamflow postprocessors often assume unbiased forcing. Accordingly, research in hydrological forecasting have to develop and enhanced forecasting systems in order to improve the skill and reliability of short to long range streamflow forecasts (Sharma et al., 2018).

This study will focus on July, August, September, and October. The period of July until September is important because these are dry months when farmers have to carefully decide their action whether they will grow a third rice crop or not (Ratri et al., 2019). The farmers should have knowledge about whether they will have enough rain in the region or have enough water to irrigate all the fields on time for the growth of rice. In addition, October will also be the focus month of the analysis in this study because farmers commonly divide a year into three rain-dependent rice cultivation seasons, and the prime cultivation starting in October (Kumalasari and Bergmeier, 2014; Naylor et al., 2001). The main planting season comprises the months between October and December (Naylor et al., 2001). The rice planting pattern in Java follows the marked seasonality of rainfall and has been explained by Naylor et al. (2001).

Forecasting skill is valuable to many sector because it gives the ability to make informed decisions and develop driven strategies. We need skillful and reliable rainfall forecasts as the input to any hydrological model. A statistical post-processing is needed to correct the errors in order to increases streamflow forecast skills across lead times and spatialscales (Siddique and Mejia, 2017; Yang et al., 2017). Yang et al. (2017), Sharma et al. (2017) implemented a regional hydrological ensemble prediction system (RHEPS) to improve the quality of ensemble streamflow forecasts in the U.S. mid-Atlantic region. Yang et al. (2017) assessed the potential of Bayesian Model Averaging (BMA) and heteroscedastic censored logistic regression (HCLR) to postprocess precipitation ensembles from the 11-member GEFSRv2 dataset, and found that generally the postprocessors perform similarly.. In this study We use ECMWF SEAS5 rainfall forecasts that have been bias-corrected using EQM as the inputs to the hydrological model. We use seasonal precipitation forecasts that have been corrected using empirical quantile mapping (SEAS5 EQM bias-corrected) as input to the hydrological model that produces seasonal streamflow forecasts for the Citarum river basin. This produces SEAS5 EQM bias-corrected streamflow forecasts. Crochemore et al. (2017) notes that precipitation forecasts are an integral part of hydrological forecasting systems both at sub-seasonal and at seasonal time scales. Moreover, taking into account climate information in seasonal streamflow forecasts increases streamflow predictability (Wood et al., 2016). In other words, to increase the skill of the hydrological forecasts, we need to improve the skill of the meteorological forecasts to achieve a better representation of the initial conditions and simulated flows.

Skillful and reliable forecasts are important. With advances in computing power, seasonal ensemble forecasts from global climate models (GCMs) have become available operationally, and these are used to force hydrological models for hydrological predictions. Models contain errors. The systematic bias in ensemble forecasts of seasonal precipitation from the European Centre for Medium-Range Weather Forecasts (ECMWF) System 5 (Johnson et al., 2019) for Java has been corrected by Ratri et al. (2019). ECMWF issues forecasts with SEAS5 up to 7-months lead times, however, in this study, we will focus only up to the 4-months lead time because the rice crop needs between 90–120 days to grow, starting from nurseries to harvest (Naylor et al., 2001). In this study, we use corrected rainfall as the input parameter for hydrological model to demonstrate that improved basic inputs for the hydrological modeling system increase skill and facilitate human and computational efforts for step-by-step verification.

In Section 2, the location of the study area is further described. In Section 3 of this article, the models and data in this study will be explained. In Sections 4 and 5, we present the verification methods as well as results and discussion, respectively. To look at the skill, we will verify the streamflow forecasts using several verification measures, e.g brier skill score (BSS), continuous ranked probability skill score (CRPSS), and mean absolute error (MAE). Furthermore, this is followed by the conclusions and future works in Section 6.

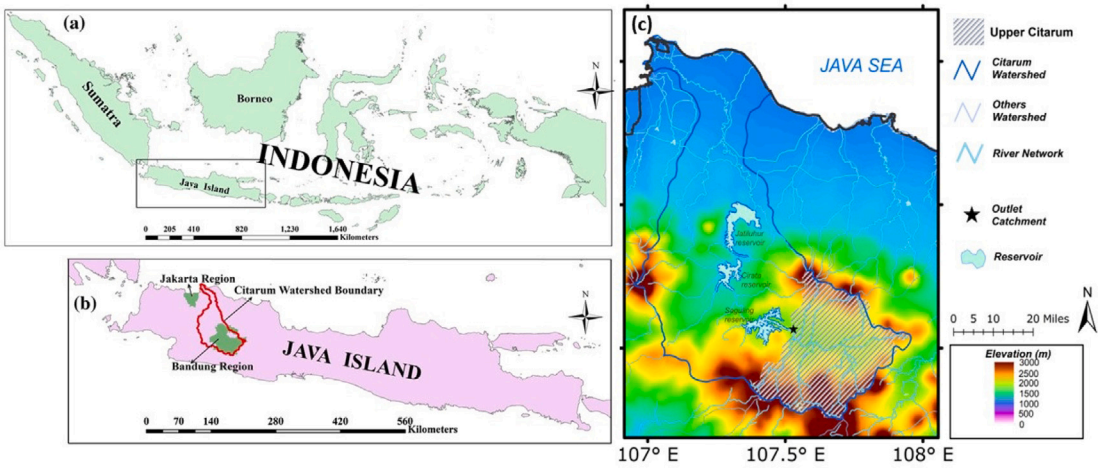


**Fig. 1.** Study area (Sahu et al., 2017).

## 2. Location of the case study area

The Citarum River basin (Fig. 1) is located in West Java, Indonesia. Geographically, this basin is located between 6°43'S 107°22'E and 7°15'S 107°57'E. The Citarum River is not only the longest but also the largest in West Java (Sahu et al., 2017). The Citarum river basin has a very central role in agricultural developments in West Java and its surrounding. Citarum river basin supplies water to 10 million residents in Jakarta (D'Arrigo et al., 2011a) and irrigates 240.000 hectares of rice field, (Hatmoko et al., 2020). The Citarum River basin is topographically surrounded by the mountainous area around the basin periphery (Dwi Dasanto et al., 2014). The highest elevation of this river basin is 2,500 m above sea level (ASL). This basin has three dams. They are Jatiluhur in the north part, Cirata in the middle part, and Saguling in the south part (Mayrowani et al., 2006). The water in the main reservoir in West Java, namely Saguling Reservoir was supplied by Citarum basin (Syahputra, 1987).

The peak of precipitation events in Java is mostly during the December–February. The rainfall variability in the Java Sea is affected by spatial and temporal monsoon wind patterns (Aldrian and Dwi Susanto, 2003). Satyawardhana and Susandi (2015) mentioned that the western part of Java has higher variability in rainfall during June–August and sometimes there is still rainfall over this region during the dry season. Meanwhile, The annual rainfall of the Citarum basin is between 2000 to 3000 mm per year, as shown in Fig. 2. The year is divided into two seasons, a rainy season (November–April) with a monthly mean rainfall of approximately is between 200–300 mm per month, and a dry season (May–October) with a monthly mean rainfall of approximately is between 50–200 mm per month (Fig. 3).

## 3. Models and data

### 3.1. Hydrological model

In this study, we use a high-resolution conceptual hydrologic model (wflow_sbm). This model has been used in broad applications on hydrological modeling (Lopez, 2018; Rusli et al., 2021). Wflow_sbm model (Schellekens et al., 2018) is used to simulate the hydrological cycle in the Citarum basin. The wflow_sbm is built on and Python language PCRaster (Karssenberg et al., 2010). Vertessy and Elsenbeer (1999) mentioned that the primary strength of this model is parameters that mostly represent physical characteristics. These strengths of the model lead to the easiness of interpreting and correlating the values of the model with physical catchment properties. An overview of the simulated processes and fluxes within the model framework can be seen in Fig. 4. See Imhoff et al. (2020) and Eilander et al. (2021) for more information on the model setup and parameterizations. Here we used the same model parameter choice and set up as Rusli et al. (2021)

Fig. 4 shows the concept of model wflow_sbm. The soil part of this model follows the same ideas as the topog_sbm model, which is designed to simulate fast runoff processes during discrete storm events in small catchments (Vertessy and Elsenbeer, 1999). However, wflow_sbm can be applied on a wide variety of catchments (Wannasin et al., 2021; Sperna Weiland et al., 2021; Imhoff et al., 2020). Rainfall interception in this model is based on the analytical approach by Gash (1979). Every single bucket represents the soil in every model grid cell. It is divided into two parts; a saturated and unsaturated store. Soil infiltration depends on the soil infiltration capacity and the fraction of paved and unpaved areas. The transfer of water from the unsaturated store to the saturated store depends on the saturated hydraulic conductivity at the water table and the saturation deficit 1 of the whole soil profile. Transpiration is first derived from the saturated store if roots intersect with the saturated store and then from the unsaturated store. Capillary rise from the saturated store also results in influx from the saturated store to the unsaturated store. The Darcy equation is applied to route water in the saturated zone laterally. The kinematic wave is used to route water over the surface and subsurface, and through rivers, based on the D8 method, which was used to extract the drainage pattern.

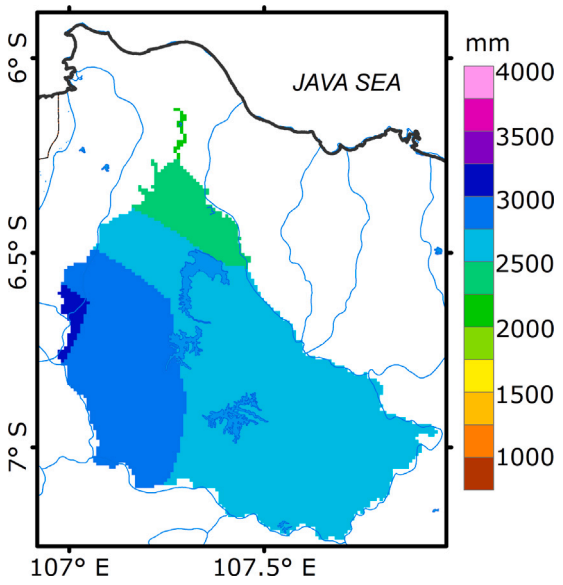


**Fig. 2.** Annual precipitation in Citarum.

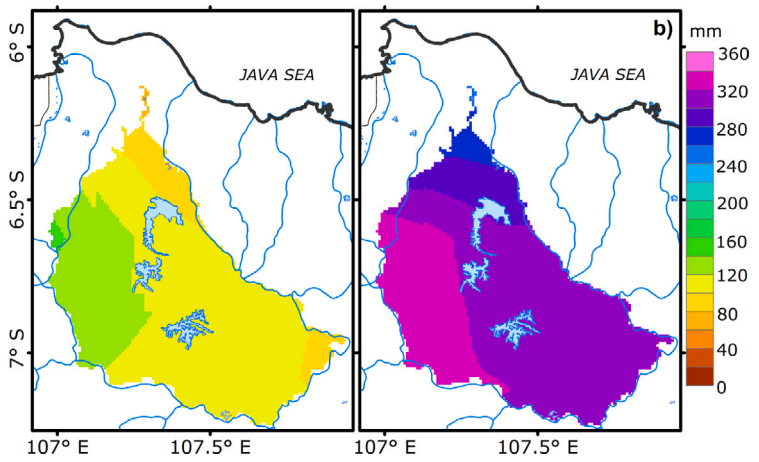


**Fig. 3.** Mean monthly rainfall (mm/month) for dry season (left plot) and wet season (right plot).

### *3.2. Estimating hydrological states: 1989–2009 simulation run*

Wflow_sbm is run in continuous simulation mode meaning that at any time an estimate of the current hydrological conditions is made: the 'hydrologcial state' of the mode. This state constitutes the starting point for a hydrological forecast run. States are estimated by using observed precipitation and temperature inputs to the model, which is then run in 'simulation mode' (as opposed to 'forecasting mode'). The simulation run uses observations as the inputs. This hydrological model was run in simulation mode for 20 years (1989–2009), resulting in what we call the 'state updating run'. Observed meteorological data originates from daily high-resolution land-only observational gridded (SA-OBS) dataset (Van Den Besselaar et al., 2017) as this data was also used to bias correct ECMWF S5 forecast (Ratri et al., 2019, 2021). The SA-OBS dataset covers the period 1981–2014. This precipitation dataset covers the Southeast Asia region and the resolution is 0.25° and 0.5°. The SA-OBS dataset is constructed based on daily station time series collected in cooperation with meteorological services in the region.

### *3.3. SEAS5 EQM bias-corrected data*

In this study, we use EQM bias-corrected precipitation at 1 to 4-month lead time as the base input (Ratri et al., 2019). The EQM bias-corrected data are bias-corrected using empirical quantile mapping of daily precipitation reforecasts from SEAS5 (Ratri et al., 2019).

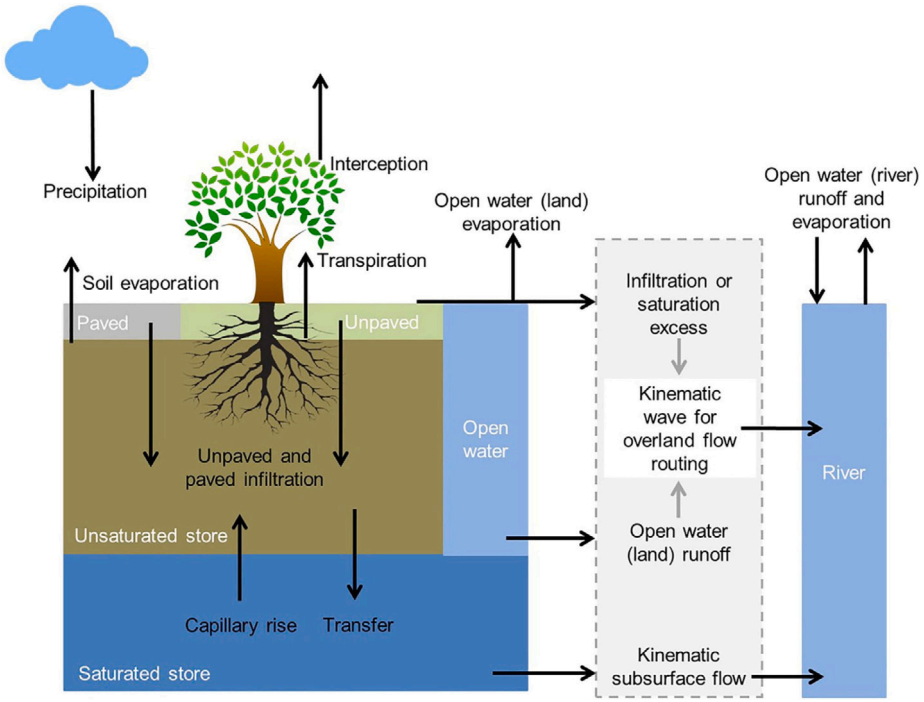


**Fig. 4.** An overview of the different processes and fluxes by the spatially distributed wflow_sbm (Schellekens et al., 2018).

In bias correcting the seasonal precipitation reforecasts, Ratri et al. (2019) used daily total precipitation SEAS5 as the raw precipitation data for the period of 1981–2010. The resolution of the reforecasts data is 36 km horizontal.SEAS5 has 51 ensemble members for the real-time forecasts, and 25 ensemble members for the hindcasts (Johnson et al., 2018).

### *3.4. Aggregated forecasts and simulation time series*

The forecasts and the simulation data set have a daily time step. However, for verification purposes, both of them are aggregated across timescales. The aggregations of 30, 60, 90, and 120 days were derived by averaging the forecast and the simulation values over the aggregation duration. These aggregation durations are relevant for different types of drought impacts and early warning systems. This is also extensively used for agricultural activities purposes.

### *3.5. Forecast lead times*

This study aims to understand how hydrological forecasting skill varies with month and season. It is important to make it possible to meet farmers' needs at their most preferred lead time of 1 month before the farming season (Nyadzi et al., 2019). Besides the aggregation, following the needs assessment, the skill assessment of the discharge forecast was performed on different lead times (1, 2, 3, and 4 months). The discharge was evaluated by validating streamflow forecast with simulation data to determine their performance in the study area.

## 4. Verification methods

To look at the skill of the wflow_sbm model derived by Rusli et al. (2021) when forced with SA-OBS for the period 1989–2009 before it is further used for this study, a metric is calculated according to the observed discharge. The observed daily discharge data are provided by the Water Resources Research and Development Center (PUSAIR). Kling–Gupta Efficiency (KGE) value is used to measure the hydrological model performances. Vis et al. 2015 mentioned that KGE has been introduced as an improvement of the used Nash–Sutcliffe efficiency (NSE), in which NSE considers different types of model errors (error in the mean, the variability, and the dynamics). The difference between KGE and NSE is that KGE focuses more on the simulation of flow variability, while the NSE focuses more on high flows (Knoben et al., 2019; Qi et al., 2022; Mizukami et al., 2019). Topography affects the KGE skill score (Qi et al., 2020) and the KGE of the hydrological model used in this study is 0.32, a range suitable to the Citarum basin and indicates that the hydrological model improves upon the mean flow benchmark (Knoben et al., 2019; Rusli et al., 2021). For comparison, Rusli et al. (2021) obtained when forcing the model with CHIRPS for the period 2005–2015, a KGE of 0.18 (and a KGE

value for GLOFAS of −0.03) for the same discharge station. In Supplementary Figure 1, we show a time-series plot (Figure S1.a) and scatterplot (Figure S1.b) of the simulated discharge against the observed discharge.

The hydrological model that uses forecast information as inputs is verified against the simulation that uses observations as inputs. This hydrological model can generate stream flow estimates over a long period and can be used to generate future yields. Simulations comprise a time series produced by the wflow_sbm model with a one-day time step. The forecasts are aggregated to forecasts comprising time series with timesteps of 30, 90, and 120 days. Those numbers of aggregation are chosen for the purpose of irrigation, especially for rice-field.

This study aims to investigate whether using EQM bias-corrected as the input of the streamflow forecast in the Citarum river basin can have good skill. We measure the quality of the deterministic and probabilistic forecasts. For the deterministic forecast quality, we look at the correlation, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) of the ensemble mean. The calculations are based on Wilks (2006).

We also measure the probabilistic forecast quality with the Brier's probability skill score (BSS), Continuous Ranked Probability Skill Score (CRPSS) and the Relative Operating Characteristic score (ROCS). The BSS is based on the Brier score (BS):

$$BS = \frac{1}{N}\sum_{t=1}^{N}\left(F_t - O_t\right)^2, \tag{1}$$

This formula above measures the root mean squared error of the forecasts in probability space. $N$ is the observed events. The forecast for event $F_t$ are expressed as a number between 0 and 1. The observation $O_t$ is indicated by either 0 (for non-occurrence of the event) or 1 (for event occurrence). Then, the calculation of the skill score based on BS of reference forecast (i.e. climatology) is as below:

$$BSS = 1 - \frac{BS}{BS_{ref}} \tag{2}$$

For the BSS, the forecast is perfect when the value is 1. Meanwhile, the is called no improvement compared to the climatology if the value is 0. A negative value means the forecast quality is poorer than the climatology. The BSS values in this study will be calculated as a function of two; less than 25th percentile and more than 75th percentile.

The next verification metric is CRPSS. Firstly, for each forecast, the difference between a cumulative ensemble forecast and an observational step function is determined, which is subsequently averaged over all forecasts considered. The calculation of CRPSS is defined as follows (e.g. Wilks, 2006):

$$CRPS = \int_{-\infty}^{\infty}\left[F(y) - F_o(y)\right]^2 dy \tag{3}$$

where

$$F_o(y) = \begin{cases} 0, y < observed\ value \\ 1, y \geq observed\ value \end{cases} \tag{4}$$

$F$(y) represents the forecast's CDF (Cumulative Distribution Function). *Fo*(y) is a stepwise CDF of the observed value. Score 0 is a perfect CRPS. The formula above defined the CRPS calculation for a single case. Meanwhile, for several cases, in practice, the CRPS is averaged (Wilks, 2006). The CRPSS compares the CRPS of a rainfall forecast to the observed climatology as a reference forecast. It measures reliability and resolution. The corresponding CRPSS can be written as follows (Wilks, 2006) :

$$CRPSS = 1 - \frac{CRPS}{CRPS_{ref}}, \tag{5}$$

The range value of CRPSS is $-\infty$ to 1. If the value $> 0$ means that the forecast improves on the reference and vice versa, it is worse than the reference if the value is >0. The $CRPS_{ref}$ is the CRPS of the reference, i.e. climatology.

The relative operating characteristic (ROC) curve is considered as a measure of the probabilistic and ensemble-based forecast potential usefulness (Mason, 1982). The ROC measures the quality of probability forecasts that relates the hit rate to the corresponding false-alarm rate (Kharin and Zwiers, 2003), to see the ability of the forecast in discriminating between events and non-events (Mason, 1982; Mason and Graham, 2002). This score does not inform about reliability because this score is not sensitive to forecast bias. The benefit of ROC is it allows non-calibrated (have not been bias corrected) forecasts to be compared (Anctil and Ramos, 2019). in The calculation, the event was defined here as the streamflow less than 25th and streamflow exceeding the 75th percentile of the sample climatology. Using a probability threshold, one has to 'decide' whether a probabilistic forecast would be interpreted as an 'event forecast' or 'event not forecast' occurrence. ROCS comprises the AUC of the main forecasting system normalized by the AUC of the climatological forecast, i.e. 0.5.

The stationary block bootstrap (Politis and Romano, 1994) is applied in order to quantify the sampling uncertainties of the verification measurements. Adjacent pairs of blocks are randomly sampled from a number of available pairs and overlapping blocks are permitted. The size of the block is 21 (length of the series data) × 25 (number of ensemble members). The average length of each block is calculated by the sample data that is being autocorrelated. This resampling action is repeated up to 1000 times to produce 1000 means for the bootstrap sample. We compute the verification metrics from each sample. For the probabilistic verification, the forecast probability was calculated for each bootstraps sample. The threshold that is measured for each sample is the discharge with a probability of more than 0.90 quantiles.

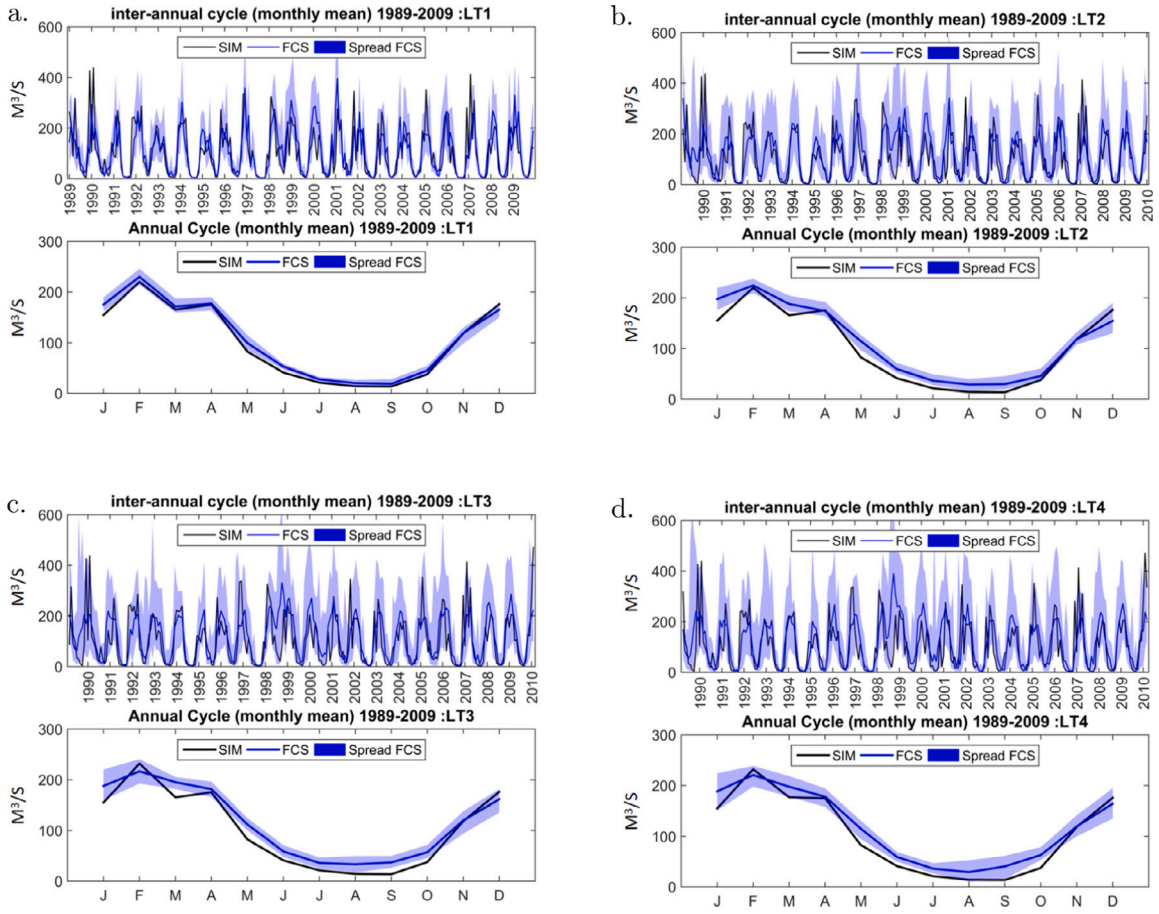


**Fig. 5.** The interannual and annual cycles of monthly mean discharge in Citarum river basin for the period 1989–2009 for (a) 1-month, (b) 2-month, (c) 3-month, and (d) 4-month lead time. The black line is the time series of the simulation data, the blue line is the mean of the forecasts ensemble member, and the shade blue areas is the forecast ensemble spread.

## 5. Results and discussion

Fig. 5 shows the interannual variability and annual cycle of monthly mean discharge of simulated and ensemble streamflow forecasts for the Citarum river basin for the period of 1989–2009 at four different lead times. We find that the mean of the forecast ensemble member shows acceptable performance in simulating discharge data. Comparing the simulated discharge to forecast discharge time series, we generally see that the temporal dynamics are covered reasonably well by the hydrological model. Comparison of the inter-annual cycle of 1 month lead time (LT1) to longer lead times (LT2–LT4) shows that the forecast ensemble spread of 1 month lead time is narrower than other lead times. As is expected, the ensemble is more confident at shorter lead times.

In addition, the annual cycle of streamflow cycles is also presented in Fig. 5, for 1 month lead time to 4 month lead time. The monsoonal pattern is observable from the figure. The annual cycle of monthly mean discharge from November to April (wet season) is between 50–250 m/s, and from May to September (dry season) is below 100 m/s to nearly zero. For LT1, the forecast follows closely the simulation data set. The forecast still follows the simulation at the longer lead times, but we can see that there are some periods of obvious over prediction such as January and March. Also, at longer lead times, we see a gap between simulation and forecast data in the dry season.

### *5.1. Verification results based on the aggregated forecast*

As described in the method section, several verification metrics were used to determine the accuracy and the skill of the streamflow forecast. This section analyzes the forecast's predictive power at the four different aggregation periods; 30d, 60d, 90d, and 120d. Identification of skillful temporal aggregations of forecast information is valuable to water supply managers (McInerney et al., 2020).

Table 1 shows the value metrics that describe the accuracy of the forecast to the simulations for different aggregation days, both deterministic and probabilistic verification measures. Correlation, RMSE, and MAE are measures of the deviations of the

**Table 1**
Score of model evaluation measures for streamflow at different aggregation days in Citarum river basin.

| Metrics | Month | 30d | 60d | 90d | 120d |
|---|---|---|---|---|---|
| Correlation | Jul | 0.65 | 0.68 | 0.61 | 0.47 |
| | Aug | 0.36 | 0.34 | 0.41 | 0.53 |
| | Sep | 0.81 | 0.59 | 0.53 | 0.51 |
| | Oct | 0.59 | 0.62 | 0.47 | 0.58 |
| RMSE | Jul | 22 | 29 | 38 | 50 |
| | Aug | 28 | 37 | 47 | 48 |
| | Sep | 17 | 30 | 46 | 44 |
| | Oct | 41 | 55 | 53 | 45 |
| MAE | Jul | 14 | 16 | 20 | 26 |
| | Aug | 14 | 18 | 24 | 29 |
| | Sep | 9 | 20 | 35 | 35 |
| | Oct | 27 | 41 | 42 | 36 |
| CRPSS | Jul | 0.43 | 0.34 | 0.04 | −0.08 |
| | Aug | 0.21 | −0.05 | 0.04 | 0.40 |
| | Sep | 0.53 | 0.42 | 0.48 | 0.42 |
| | Oct | 0.50 | 0.57 | 0.46 | 0.53 |
| BSS (prob<P 25th) | Jul | 0.65 | 0.82 | 0.60 | 0.71 |
| | Aug | 0.78 | 0.63 | 0.59 | 0.62 |
| | Sep | 0.63 | 0.85 | 0.65 | 0.52 |
| | Oct | 0.71 | 0.62 | 0.45 | 0.53 |
| BSS (prob<P 75th) | Jul | 0.75 | 0.63 | 0.55 | 0.55 |
| | Aug | 0.49 | 0.58 | 0.66 | 0.66 |
| | Sep | 0.57 | 0.61 | 0.63 | 0.53 |
| | Oct | 0.70 | 0.77 | 0.49 | 0.45 |
| ROCS (prob<P 25th) | Jul | 0.85 | 0.96 | 0.70 | 0.90 |
| | Aug | 0.95 | 0.81 | 0.74 | 0.73 |
| | Sep | 0.73 | 0.95 | 0.74 | 0.61 |
| | Oct | 0.73 | 0.79 | 0.51 | 0.54 |
| ROCS (prob<P 75th) | Jul | 0.70 | 0.73 | 0.65 | 0.64 |
| | Aug | 0.76 | 0.60 | 0.70 | 0.78 |
| | Sep | 0.60 | 0.54 | 0.66 | 0.40 |
| | Oct | 0.79 | 0.88 | 0.24 | 0.28 |

deterministic forecasts values from the values of the simulation. The correlation coefficient varies considerably with the aggregation period, with optimum values attained when data is aggregated to 30 to 60 days, except August in terms of the correlation value. Mostly, the error metrics reduce by the higher aggregation days.

From Table 1, it shows the increase in the RMSE as a function of time aggregation for all months (July–October). In addition, MAE measures the average magnitude of the errors in a set of forecasts. However, it does not consider their direction. Moreover, from Table 1, for all months, the MAE increases with the increasing aggregation days. Table 1 also shows the probabilistic or quantitative model checking values. They are CRPSS, and both the BSS and ROCS for the probability of streamflow less than 25th and more than 75th percentile. From the scores, streamflow at 30 to 60 days aggregation days is more skillful than larger aggregation.

In July (Supplementary Figure 2), based on the RMSE and MAE, the error is the lowest for a 30-day aggregation and increases with the increasing level of aggregation, reaching the largest value for a 120-day aggregation. In general, the more aggregations the larger the error. This is due to its seasonal variability. This happens because when a value is aggregated over a longer time, it will cause the number of seasons involved to have more varied rainfall. To provide information tailored to the user's needs, aggregation over longer (longer than 30-days) periods is needed (Bohn et al., 2010). Temporal aggregation of 30 days is typically used in seasonal forecasting services (Apel et al., 2018). In addition, probabilistic seasonal forecasts are essential for water-intensive activities requiring long-term planning. Table 1 also shows that, even if the skill for the forecast period decreases when aggregating of the days is higher, the forecasts mostly remain skillful (CRPSS are mostly greater than 0). The BSS value for the threshold more than the 75th percentile also shows the values are greater than zero.

In addition, Supplementary Figure 2 also shows that the skill of the forecast when predicting the streamflow less than the 25th percentile decreases by the increasing number of aggregations even though the different values of BSS are not significantly different for the 60 to 120 days aggregation. Meanwhile, for the CRPSS, the skill after the 30-days aggregation decreases and reaches around zero but then increases again for the higher aggregation days. The streamflow forecast skill may be affected by the precipitation forecast skill. The variability of the CRPSS values of SEAS5 bias-corrected model forecast in the western part of Java is probably because of its high rainfall variability that is linked to topography (Ratri et al., 2019)

The forecast model skill is expected to decay with the increase in lead time. However, this is not always the case in August (Supplementary Figure 3). However, if we look at the error, both RMSE and MAE, we can see there is a rate of increase of errors with lead time as expected. The skill of the streamflow forecast is affected by the skill of rainfall forecast. In Indonesia, the skill also depends on the region, month and season, and the distribution of land and sea (Aldrian et al., 2007).

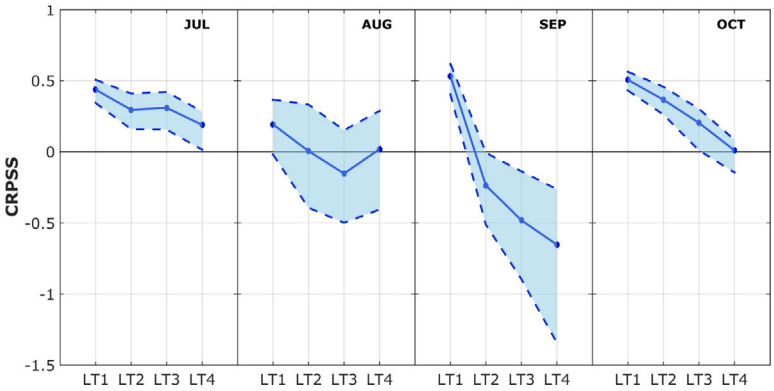


**Fig. 6.** CRPSS values for July to October at LT1 to LT4.

In September (Supplementary Figure 4), except for the BSS and ROCS with a threshold less than the 25th percentile, the other metrics such as RMSE, MAE, CORR, CRPSS are almost as expected. It shows that 30 to 60 days aggregation has higher performance. This is also the same as in October (Supplementary Figure 5). The validation metrics for October show that 30 to 60 days of aggregation have higher performance than higher aggregation days. During September–October, streamflow variability in the Citarum river basin is linked to the IOD and ENSO variations. The streamflow over this region and its link to climate has been analyzed by D'Arrigo et al. (2011b)

### *5.2. Verification results based on the lead time forecasts*

The verification scores are also investigated based on the lead times. Fig. 6 presents the CRPSS values of the streamflow forecast for LT1 to LT4. The figure shows that all forecasts for LT1 are skillful. In addition, in July and October, all forecasts until LT4 are skillful, while in August and September skill is lost in the second lead time.

For all months, the median CRPSS values of LT1 and LT2 are higher than longer lead times (shown by blue solid lines in Fig. 6). The CRPSS values up to about 0.5 for July, September, and October in the LT1 except for August, the CRPSS values decrease by lead times. In July and October, most of the CRPSS values for all lead times are above zero; however, in August and September the CRPPS values for LT2–LT4 are below zero.

Nyamekye et al. (2021) mentioned that seasonal forecast information provided at an early lead time significantly informed farmer decision-making compared to longer lead times. Forecasts of precipitation and streamflow are closely related to each other and it can potentially assist with planning of managed reservoir releases provided forecasts are sufficiently accurate at lead times necessary for management decisions (Turner et al., 2017). Nyamekye et al. (2021) added that communicating forecast information at an appropriate lead time has the potential to help users especially farmers manage risks and improve decision-making.

Pappenberger et al. (2015) mentioned that the CRPSS is one of the most recommended skill scores if we want to evaluate overall hydrological ensemble forecast performance. Fig. 11 shows that for all the months, the CRPSS values tend overall to decline with increasing lead time. This is because the longer the lead times, the weather uncertainties tend to grow and become more dominant (Siddique and Mejia, 2017). Across all lead times, the CRPSS values vary approximately from about −1 to 1 (the median CRPSS showed by the solid line). The CRPSS value at LT1 for all months has positive skill. For LT2, the median of the CRPSS is still nearly 0, however for longer lead times especially for August and September the CRPSS have slightly negative skills. Even though in general there are decreasing CRPSS values by lead times, there is a relatively narrow range of the lead times. Sharma et al. (2019) found that after postprocessing is applied into a streamflow forecast skill, looking at the CRPSS values, the models have comparable skill across lead times.

Generally, as would be expected, the MAE increases slightly by increasing the lead times. This indicates decreasing skill with increasing lead time (Fig. 7). In all months the SEAS5 EQM bias-corrected based streamflow forecasts are more skillful in early (LT1–LT2) compared to longer (LT3–LT4) lead times. Furthermore, Fig. 7 shows the variability of MAE among months and lead times. Looking at the accuracy error statistics like MAE will allow us to see how large the bias and errors are. The variability and median values of MAE show agreement with the CRPSS values. This indicates there are more variability and uncertainty as the streamflow forecast extends further in time. Meanwhile, among those four months, for October, the majority of the grid cells have larger MAE than other months for all lead times. This is probably because October is generally wetter than the other months.

The performance of the forecasts is good for the shorter lead time but may become worse as the lead times increases. This can be also observed in Figs. 8 and 9, which show that correlation (RMSE) is generally high (small) for early lead times but increases (decreases) thereafter.

Boxplots of BSSs are shown in Fig. 10. In the wet months (September and October), better performance of streamflow ensembles is obtained for early LT, both for threshold less than 25th percentile and more than the 75th percentile. BSS values vary significantly with lead time. However, in dry months (July and August) better scores are generally obtained when considering high discharge

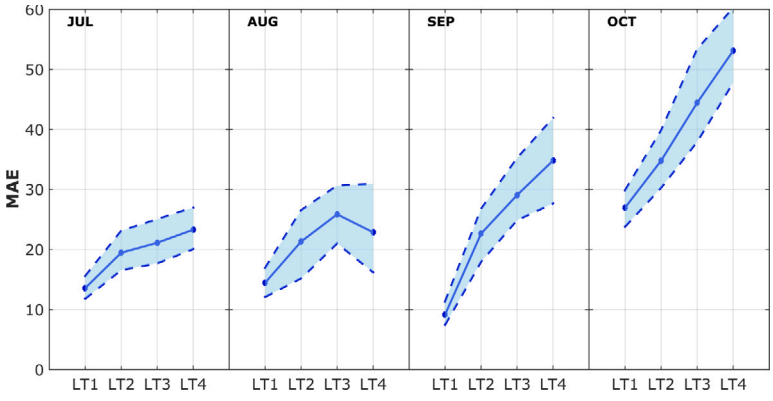


**Fig. 7.** MAE values for July to October at LT1 to LT4.

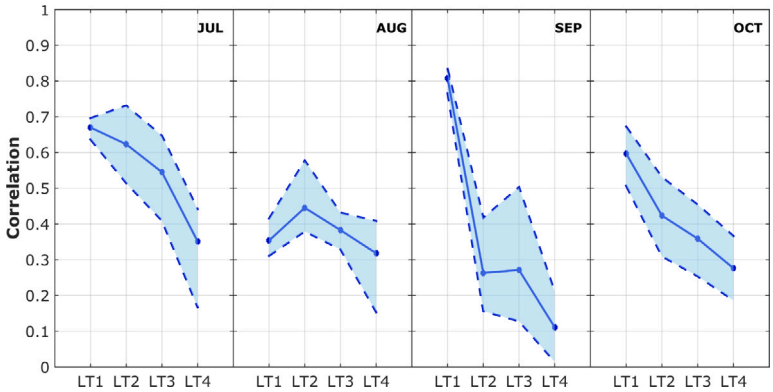


**Fig. 8.** CORRELATION values for July to October at LT1 to LT4.

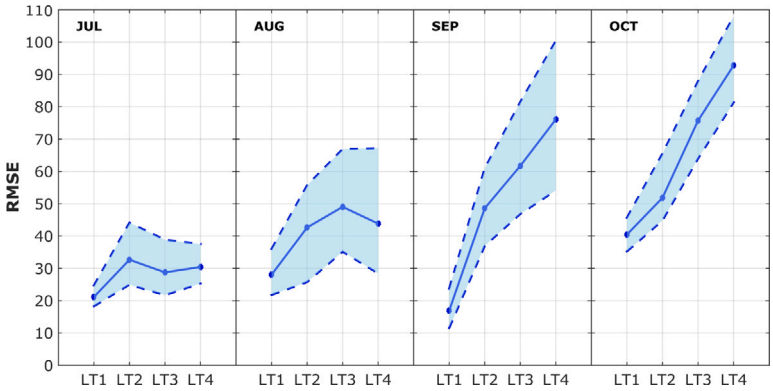


**Fig. 9.** RMSE values for July to October at LT1 to LT4.

thresholds (exceedances of the percentile Q75). This is especially observed in the case of the hydrological models that do not make use of an updating procedure during forecasting. The boxplot distributions of BSS shown in Fig. 10 indicate positive values and therefore a better performance of the forecasting systems compared to the climatology. Except for August, the better skill was shown in earlier lead times.

Considering the growth of rice crops is between 90–120 days which is related to the early lead times, the forecast information with the positive skill of BSS is sufficient to adjust critical agricultural decisions. To improve the efficiency of agricultural management and to ensure food and livelihood security, forecast information with a sufficient lead time to adjust critical agricultural decisions is needed (Apipattanavis et al., 2010).

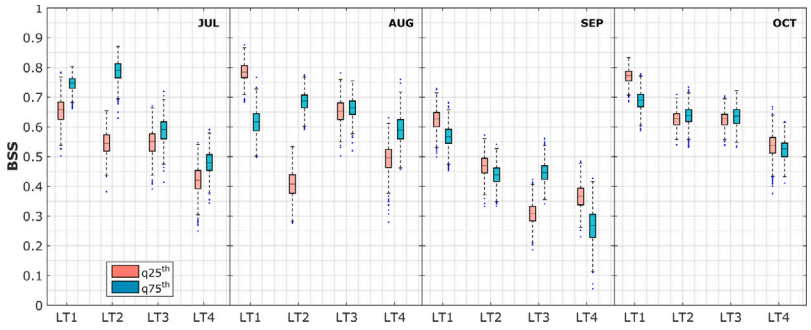


**Fig. 10.** Box-plots of BSS for July, August, September, and October at 1-month lead to 4-month lead times. The box-plots showing results for BSS with threshold less than 25th percentile (shown in red) and more than the 75th percentile (shown in blue). The middle horizontal line in the box is the median over all values, the upper and lower limits of the box are the 75th and 25th percentile, respectively, and the dots indicate values that are outside this range.

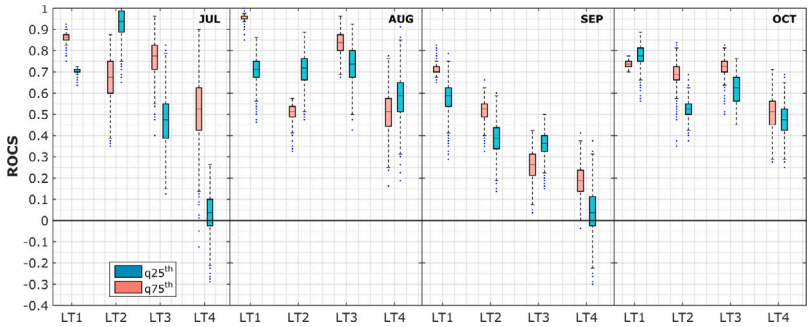


**Fig. 11.** Box-plots of ROCS for July, August, September, and October at 1-month lead to 4-month lead times. The box-plots showing results for ROCS with threshold less than 25th percentile (shown in red) and more than the 75th percentile (shown in blue). The middle horizontal line in the box is the median over all values, the upper and lower limits of the box are the 75th and 25th percentile, respectively, and the dots indicate values that are outside this range.

Fig. 11 shows the ROC skill score as a function of forecast lead time. The evaluation that is conducted in this study is also based on the ROC skill score at a threshold less than the 25th percentile and more than the 75th percentile. The no-skill value for ROCSS is 0, so from Fig. 11, it is shown that especially in the wet months (September and October), for both thresholds, the longer lead times are the less skillful. Meanwhile, for the early LT, there is no regular pattern of skill decreasing by lead times.

## 6. Conclusion and future works

Enhancing hydrological forecasting accuracy is still a challenging issue. By applying the EQM bias-correction approaches to ECMWF SEAS5 before generating the streamflow forecasts, exploring four different aggregation days, and four different lead times, we find that the model skills reach the most efficient at 30 or 60 aggregation days. Meanwhile, based on the different lead times, the performances of LT1 and LT2 are better than other lead times.

BSS indicates that the model approach is reasonable for early lead times, as the skill score shows positive value. This evaluation performance of the hydrological model is good and can help farmers to adjust critical agricultural decisions. LT1 and LT2 of forecast information are more useful for agricultural rice practitioners than longer lead times as they relate to the rice plants' cycle.

So far, given information related to these issues is still lacking, this methodology is now good enough to be used for advice to farmers. This study, especially in Indonesia, is a study that initiates using a hydrological model with inputs from a bias-corrected seasonal rainfall forecast. When forecasts are accurate, then the information has the potential as a significant tool in water resource management (Alemu et al., 2011). On the other hand, unskilled forecasts can result in unwanted impacts. To allocate resources and meet demands, water resource managers and agricultural practitioners require streamflow predictions that are accurate and timely.

For example, there will be unnecessary drought mitigation actions happen if the forecasts predict lower flows than actually occur. Conversely, improper drought mitigation actions will be called for which could result in a drought if the forecasts predict higher flows than actually occur.

For now, this paper only describes the performance, skill, and uncertainty of SEAS5 EQM bias-corrected based streamflow prediction. However, the improved skill compared to climatology cannot be achieved for all forecast lead times. In a future study, we could look at post-processing the streamflow forecast itself. In addition, a more advanced statistical postprocessing approach for the ensemble precipitation forecast is may be needed to further improve seasonal forecast skills. So, the performance of SEAS5 Advanced statistical bias-corrected based streamflow forecast skill can be studied. Adding climate indices (e.g: dipole mode index (DMI), Madden–Julian oscillation (MJO) indices, sea surface temperature (SST)) as potential predictors in the analysis for future study can be another sight of future work concerns.

## CRediT authorship contribution statement

**Dian Nur Ratri:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Albrecht Weerts:** Conception and design of study, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Robi Muharsyah:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft. **Kirien Whan:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft. **Albert Klein Tank:** Writing – review & editing. **Edvin Aldrian:** Writing – review & editing. **Mugni Hadi Hariadi:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data & editing.

## Declaration of competing interest

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

## Data availability

Data will be made available on request.

## Acknowledgments

The authors are grateful to anonymous reviewers for their comments on an of the manuscript, which have helped to improve it. This research is also significantly supported by Koninklijk Nederlands Meteorologisch Instituut (KNMI) and Deltares Netherlands, as well as Indonesian Meteorology, Climatology, and Geophysical Agency (BMKG) by providing access to the available data and softwares. All authors approved the version of the manuscript to be published. The work was partly conducted with a grant of Dutch Space Office (NSO) in the G4AW program XI-IATI-NSO-G4AW13009.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ejrh.2022.101305.

## References

Aldrian, E., Dwi Susanto, R., 2003. Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. International Journal of Climatology: A Journal of the Royal Meteorological Society 23 (12), 1435–1452. http://dx.doi.org/10.1002/joc.950.

Aldrian, E., Gates, L.D., Widodo, F.H., 2007. Seasonal variability of Indonesian rainfall in ECHAM4 simulations and in the reanalyses: The role of ENSO. Theor. Appl. Climatol. 87 (1), 41–59. http://dx.doi.org/10.1007/s00704-006-0218-8.

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

Anctil, F., Ramos, M., 2019. Verification metrics for hydrological ensemble forecasts. In: Handbook of Hydrometeorological Ensemble Forecasting; Springer: Berlin/Heidelberg, Germany. pp. 893–922. http://dx.doi.org/10.1007/978-3-642-40457-3.

Apel, H., Abdykerimova, Z., Agalhanova, M., Baimaganbetov, A., Gavrilenko, N., Gerlitz, L., Kalashnikova, O., Unger-Shayesteh, K., Vorogushyn, S., Gafurov, A., 2018. Statistical forecast of seasonal discharge in Central Asia using observational records: development of a generic linear modelling tool for operational water resource management. Hydrol. Earth Syst. Sci. 22 (4), 2225–2254. http://dx.doi.org/10.5194/hess-22-2225-2018.

Apipattanavis, S., Bert, F., Podestá, G., Rajagopalan, B., 2010. Linking weather generators and crop models for assessment of climate forecast outcomes. Agricult. Forest Meteorol. 150 (2), 166–174. http://dx.doi.org/10.1016/j.agrformet.2009.09.012, URL https://www.sciencedirect.com/science/article/pii/S0168192309002287.

Bohn, T.J., Sonessa, M.Y., Lettenmaier, D.P., 2010. Seasonal hydrologic forecasting: Do multimodel ensemble averages always yield improvements in forecast skill? J. Hydrometeorol. 11 (6), 1358–1372. http://dx.doi.org/10.1175/2010JHM1267.1.

Brown, J.D., Seo, D.-J., Du, J., 2012. Verification of precipitation forecasts from NCEP's short-range ensemble forecast (SREF) system with reference to ensemble streamflow prediction using lumped hydrologic models. J. Hydrometeorol. 13 (3), 808–836.

Crochemore, L., Ramos, M.-H., Pappenberger, F., 2016. Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts. Hydrol. Earth Syst. Sci. 20, 3601–3618. http://dx.doi.org/10.5194/hess-20-3601-2016.

Crochemore, L., Ramos, M.-H., Pappenberger, F., Perrin, C., 2017. Seasonal streamflow forecasting by conditioning climatology with precipitation indices. Hydrol. Earth Syst. Sci. 21 (3), 1573–1591. http://dx.doi.org/10.5194/hess-21-1573-2017.

D'Arrigo, R., Abram, N., Ummenhofer, C., Palmer, J., Mudelsee, M., 2011a. Reconstructed streamflow for Citarum River, Java, Indonesia: linkages to tropical climate dynamics. Clim. Dynam. 36 (3–4), 451–462. http://dx.doi.org/10.1007/s00382-009-0717-2.

D'Arrigo, R., Abram, N., Ummenhofer, C., Palmer, J., Mudelsee, M., 2011b. Reconstructed streamflow for Citarum River, Java, Indonesia: linkages to tropical climate dynamics. Clim. Dynam. 36 (3), 451–462.

Dwi Dasanto, B., Boer, R., Pramudya, B., Suharnoto, Y., 2014. Simple method for assessing spread of Flood Prone Areas under historical and future rainfall in the upper citarum watershed. EnvironmentAsia 7 (2).

Eilander, D., Van Verseveld, W., Yamazaki, D., Weerts, A., Winsemius, H.C., Ward, P.J., 2021. A hydrography upscaling method for scale-invariant parametrization of distributed hydrological models. Hydrol. Earth Syst. Sci. 25 (9), 5287–5313. http://dx.doi.org/10.5194/hess-25-5287-2021.

Fares, Y., Yudianto, D., 2003. Assessment of hydrological characteristics in the upper Citarum catchment, West Java. WIT Trans. Ecol. Environ. 60, URL https://www.witpress.com/Secure/elibrary/papers/RM03/RM03016FU.pdf.

Gash, J., 1979. An analytical model of rainfall interception by forests. Q. J. R. Meteorol. Soc. 105 (443), 43–55. http://dx.doi.org/10.1002/qj.49710544304.

Hatmoko, W., Diaz, B., et al., 2020. Comparison of rainfall-runoff models for climate change projection–case study of citarum river basin, Indonesia. In: IOP Conference Series: Earth and Environmental Science, vol. 423, no. 1. IOP Publishing, 012045, URL https://iopscience.iop.org/article/10.1088/1755-1315/423/1/012045/meta.

Imhoff, R., van Verseveld, W., van Osnabrugge, B., Weerts, A., 2020. Scaling point-scale (pedo) transfer functions to seamless large-domain parameter estimates for high-resolution distributed hydrologic modeling: An example for the Rhine river. Water Resour. Res. 56 (4), http://dx.doi.org/10.1029/2019WR026807, e2019WR026807.

Johnson, S., Ferranti, L., Balmaseda, M., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S., Mogensen, K., Zuo, H., Monge-Sanz, B., 2018. SEAS5: The new ECMWF seasonal forecast system. Geosci. Model Dev. Discuss. 1–44. http://dx.doi.org/10.5194/gmd-2018-228.

Johnson, S.J., Stockdale, T.N., Ferranti, L., Balmaseda, M.A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., et al., 2019. SEAS5: The new ECMWF seasonal forecast system. Geosci. Model Dev. 12 (3).

Karssenberg, D., Schmitz, O., Salamon, P., de Jong, K., Bierkens, M.F., 2010. A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. Environ. Model. Softw. 25 (4), 489–502. http://dx.doi.org/10.1016/j.envsoft.2009.10.004.

Kharin, V.V., Zwiers, F.W., 2003. On the ROC score of probability forecasts. J. Clim. 16 (24), 4145–4150. http://dx.doi.org/10.1175/1520-0442(2003)016<4145:OTRSOP>2.0.CO;2.

Knoben, W.J., Freer, J.E., Woods, R.A., 2019. Inherent benchmark or not? Comparing Nash–sutcliffe and Kling–Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23 (10), 4323–4331.

Kumalasari, N.R., Bergmeier, E., 2014. Effects of surrounding crop and semi-natural vegetation on the plant diversity of paddy fields. Agric. Food Secur. 3 (1), 15. http://dx.doi.org/10.1186/2048-7010-3-15.

Loebis, J., Syariman, P., 1993. Reservoir operation conflict in citarum river basin management. IAHS PUBLICATION 455.

Lopez, P., 2018. Application of global hydrological datasets for river basin modelling (Ph.D. thesis). Utrecht University.

Mason, I., 1982. A model for assessment of weather forecasts. Aust. Meteor. Mag 30 (4), 291–303.

Mason, S.J., Graham, N.E., 2002. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. Quart. J. R. Meteorol. Soc. J. Atm. Sci. Appl. Meteorol. Phys. Oceanogr. 128 (584), 2145–2166. http://dx.doi.org/10.1256/003590002320603584.

Masselot, P., Dabo-Niang, S., Chebana, F., Ouarda, T.B., 2016. Streamflow forecasting using functional regression. J. Hydrol. 538, 754–766. http://dx.doi.org/10.1016/j.jhydrol.2016.04.048.

Mayrowani, H., Nurmanaf, A., Kundarto, M., 2006. Assessment OF environmental multifunctions OF PADDY FARMING IN citarum RIVER basin, WEST java, INDONESIA. URL https://balittanah.litbang.pertanian.go.id/ind/dokumentasi/prosiding/mflp2003/fahmuddin_agus01.pdf.

McInerney, D., Thyer, M., Kavetski, D., Laugesen, R., Tuteja, N., Kuczera, G., 2020. Multi-temporal hydrological residual error modeling for seamless subseasonal streamflow forecasting. Water Resour. Res. 56 (11), http://dx.doi.org/10.1029/2019WR026979, e2019WR026979.

Mizukami, N., Rakovec, O., Newman, A.J., Clark, M.P., Wood, A.W., Gupta, H.V., Kumar, R., 2019. On the choice of calibration metrics for "high-flow" estimation using hydrologic models. Hydrol. Earth Syst. Sci. 23 (6), 2601–2614.

Naylor, R.L., Falcon, W.P., Rochberg, D., Wada, N., 2001. Using El Nino/Southern Oscillation climate data to predict rice production in Indonesia. Clim. Change 50 (3), 255–265. http://dx.doi.org/10.1023/A:1010662115348.

Nyadzi, E., Werners, E.S., Biesbroek, R., Long, P.H., Franssen, W., Ludwig, F., 2019. Verification of seasonal climate forecast toward hydroclimatic information needs of rice farmers in northern ghana, weather, climate, and society. Weather Clim. Soc. 11 (1), 127–142. http://dx.doi.org/10.1175/WCAS-D-17-0137.1.

Nyamekye, A.B., Nyadzi, E., Dewulf, A., Werners, S., Van Slobbe, E., Biesbroek, R.G., Termeer, C.J., Ludwig, F., 2021. Forecast probability, lead time and farmer decision-making in rice farming systems in Northern Ghana. Clim. Risk Manag. 31, 100258. http://dx.doi.org/10.1016/j.crm.2020.100258, URL https://www.sciencedirect.com/science/article/pii/S2212096320300486.

Pappenberger, F., Ramos, M.-H., Cloke, H.L., Wetterhall, F., Alfieri, L., Bogner, K., Mueller, A., Salamon, P., 2015. How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. J. Hydrol. 522, 697–713.

Pasaribu, S.M., 2010. Developing rice farm insurance in Indonesia. Agric. Agric. Sci. Procedia 1, 33–41.

Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. J. Amer. Statist. Assoc. 89 (428), 1303–1313. http://dx.doi.org/10.1080/01621459.1994.10476870.

Qi, W., Chen, J., Li, L., Xu, C.-y., Li, J., Xiang, Y., Zhang, S., 2020. A framework to regionalize conceptual model parameters for global hydrological modeling. Hydrol. Earth Syst. Sci. Discuss. 1–28.

Qi, W.-y., Chen, J., Li, L., Xu, C.-Y., Li, J., Xiang, Y., Zhang, S., 2022. Regionalization of catchment hydrological model parameters for global water resources simulations. Hydrol. Res. 53 (3), 441–466.

Ratri, D.N., Whan, K., Schmeits, M., 2019. A comparative verification of raw and bias-corrected ECMWF seasonal ensemble precipitation reforecasts in Java (Indonesia). J. Appl. Meteorol. Climatol. 58 (8), 1709–1723. http://dx.doi.org/10.1175/JAMC-D-18-0210.1.

Ratri, D.N., Whan, K., Schmeits, M., 2021. Calibration of ECMWF seasonal ensemble precipitation reforecasts in Java (Indonesia) using bias-corrected precipitation and climate indices. Weather Forecast. 36, 1375–1386. http://dx.doi.org/10.1175/WAF-D-20-0124.1.

Rusastra, I.W., Saliem, H.P., et al., 2019. Prospek pengembangan pola tanam dan diversifikasi tanaman pangan di Indonesia. URL http://ejurnal.litbang.pertanian.go.id/index.php/fae/article/view/4075.

Rusli, S., Weerts, A., Taufiq, A., Bense, V., 2021. Estimating water balance components and their uncertainty bounds in highly groundwater-dependent and data-scarce area: An example for the upper citarum basin. J. Hydrol. Regional Stud. 37, 100911. http://dx.doi.org/10.1016/j.ejrh.2021.100911, URL https://www.sciencedirect.com/science/article/pii/S2214581821001403.

Sahu, N., Behera, S.K., Yamashiki, Y., Takara, K., Yamagata, T., 2012. IOD and ENSO impacts on the extreme stream-flows of Citarum river in Indonesia. Clim. Dynam. 39 (7), 1673–1680.

Sahu, N., Robertson, A.W., Boer, R., Behera, S., DeWitt, D.G., Takara, K., Kumar, M., Singh, R., 2017. Probabilistic seasonal streamflow forecasts of the Citarum River, Indonesia, based on general circulation models. Stochastic Environ. Res. Risk Assess. 31 (7), 1747–1758. http://dx.doi.org/10.1007/s00477-016-1297-4.

Satyawardhana, H., Susandi, A., 2015. Proyeksi awal musim di jawa berbasis hasil downscaling conformal cubic atmospheric model (CCAM)(season onset projection in Java based on CCAM downscaling output). J. Sains Dirgantara 13 (1).

Schellekens, J., Verseveld, W.v., Euser, T., Winsemius, H.C., Thiange, C., Bouaziz, L., 20 Tollenaar, D., Vries, S.d., Weerts, A.H., 2018. https://Github.Com/Openstreams/Wflow.

Sharma, S., Siddique, R., Balderas, N., Fuentes, J.D., Reed, S., Ahnert, P., Shedd, R., Astifan, B., Cabrera, R., Laing, A., et al., 2017. Eastern US verification of ensemble precipitation forecasts. Weather Forecast. 32 (1), 117–139.

Sharma, S., Siddique, R., Reed, S., Ahnert, P., Mejia, A., 2019. Hydrological model diversity enhances streamflow forecast skill at short-to medium-range timescales. Water Resour. Res. 55 (2), 1510–1530. http://dx.doi.org/10.1029/2018WR023197.

Sharma, S., Siddique, R., Reed, S., Ahnert, P., Mendoza, P., Mejia, A., 2018. Relative effects of statistical preprocessing and postprocessing on a regional hydrological ensemble prediction system. Hydrol. Earth Syst. Sci. 22 (3), 1831–1849.

Siddique, R., Mejia, A., 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. http://dx.doi.org/10.1175/JHM-D-16-0243.1.

Siddique, R., Mejia, A., Brown, J., Reed, S., Ahnert, P., 2015. Verification of precipitation forecasts from two numerical weather prediction models in the Middle Atlantic Region of the USA: A precursory analysis to hydrologic forecasting. J. Hydrol. 529, 1390–1406.

Sperna Weiland, F.C., Visser, R.D., Greve, P., Bisselink, B., Brunner, L., Weerts, A.H., 2021. Estimating regionalized hydrological impacts of climate change over europe by performance-based weighting of CORDEX projections. Front. Water 3, 143. http://dx.doi.org/10.3389/frwa.2021.713537, URL https://www.frontiersin.org/article/10.3389/frwa.2021.713537.

Surmaini, E., Syahbuddin, H., 2016. Kriteria Awal Musim Tanam: Tinjauan Prediksi Waktu Tanam padi di Indonesia. J. Penelitian Pengembangan Pertanian 35 (2), 47–56. http://dx.doi.org/10.21082/jp3.v35n2.2016.p47-56.

Suwarno, S., 2016. Meningkatkan produksi padi menuju ketahanan pangan yang lestari. J. Pangan 19 (3), 233–243. http://dx.doi.org/10.33964/jp.v19i3.150.

Syahputra, M.R., 1987. Verification of upper citarum river discharge prediction using climate forecast system version 2 (CFSv2) output. In: AIP Conference Proceedings, vol. 20040, no. 2018.

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

Van Den Besselaar, E.J., Van Der Schrier, G., Cornes, R.C., Iqbal, A.S., Klein Tank, A.M., 2017. SA-OBS: a daily gridded surface temperature and precipitation dataset for southeast Asia. J. Clim. 30 (14), 5151–5165. http://dx.doi.org/10.1175/JCLI-D-16-0575.1.

Vertessy, R.A., Elsenbeer, H., 1999. Distributed modeling of storm flow generation in an amazonian rain forest catchment: Effects of model parameterization. Water Resour. Res. 35 (7), 2173–2187. http://dx.doi.org/10.1029/1999WR900051.

Wannasin, C., Brauer, C., Uijlenhoet, R., van Verseveld, W., Weerts, A., 2021. Daily flow simulation in Thailand part I: Testing a distributed hydrological model with seamless parameter maps based on global data. J. Hydrol. Regional Stud. 34, 100794. http://dx.doi.org/10.1016/j.ejrh.2021.100794, URL https://www.sciencedirect.com/science/article/pii/S2214581821000239.

Wilks, D.S., 2006. Comparison of ensemble-MOS methods in the Lorenz'96 setting. Meteor. Appl. 13 (3), 243–256. http://dx.doi.org/10.1017/S1350482706002192.

Wood, A., Pagano, T., Roos, M., 2016. Tracing the origins of ESP. HEPEX Blog Available At: https://Hepex.Irstea.Fr/Tracing-the-Origins-of-Esp/, (Accessed 28 July 2017).

Yang, X., Sharma, S., Siddique, R., Greybush, S.J., Mejia, A., 2017. Postprocessing of GEFS precipitation ensemble reforecasts over the U.S. mid-atlantic region. Mon. Weather Rev. 145 (5), 1641–1658. http://dx.doi.org/10.1175/MWR-D-16-0251.1, URL https://journals.ametsoc.org/view/journals/mwre/145/5/mwr-d-16-0251.1.xml.

Zhao, L., Duan, Q., Schaake, J., Ye, A., Xia, J., 2011. A hydrologic post-processor for ensemble streamflow predictions. Adv. Geosci. 29, 51–59.