Journal of
Water & Climate Change

IWA PUBLISHING

Journal of Water and Climate Change Vol 15 No 2, 370 doi: 10.2166/wcc.2024.320

Predicting the peak flow and assessing the hydrologic hazard of the Kessem Dam, Ethiopia using machine learning and risk management centre-reservoir frequency analysis software

Elias Gebeyehu Ayele *, Esayas Tesfaye Ergete and Getachew Bereta Geremew

Faculty of Hydraulic & Water Resources Engineering, Water Technology Institute, Arba Minch University, P.O.B. 21, Arba Minch, Ethiopia
*Corresponding author. E-mail: elias.gebeyehu@amu.edu.et

EGA, 0000-0002-8150-1610

ABSTRACT

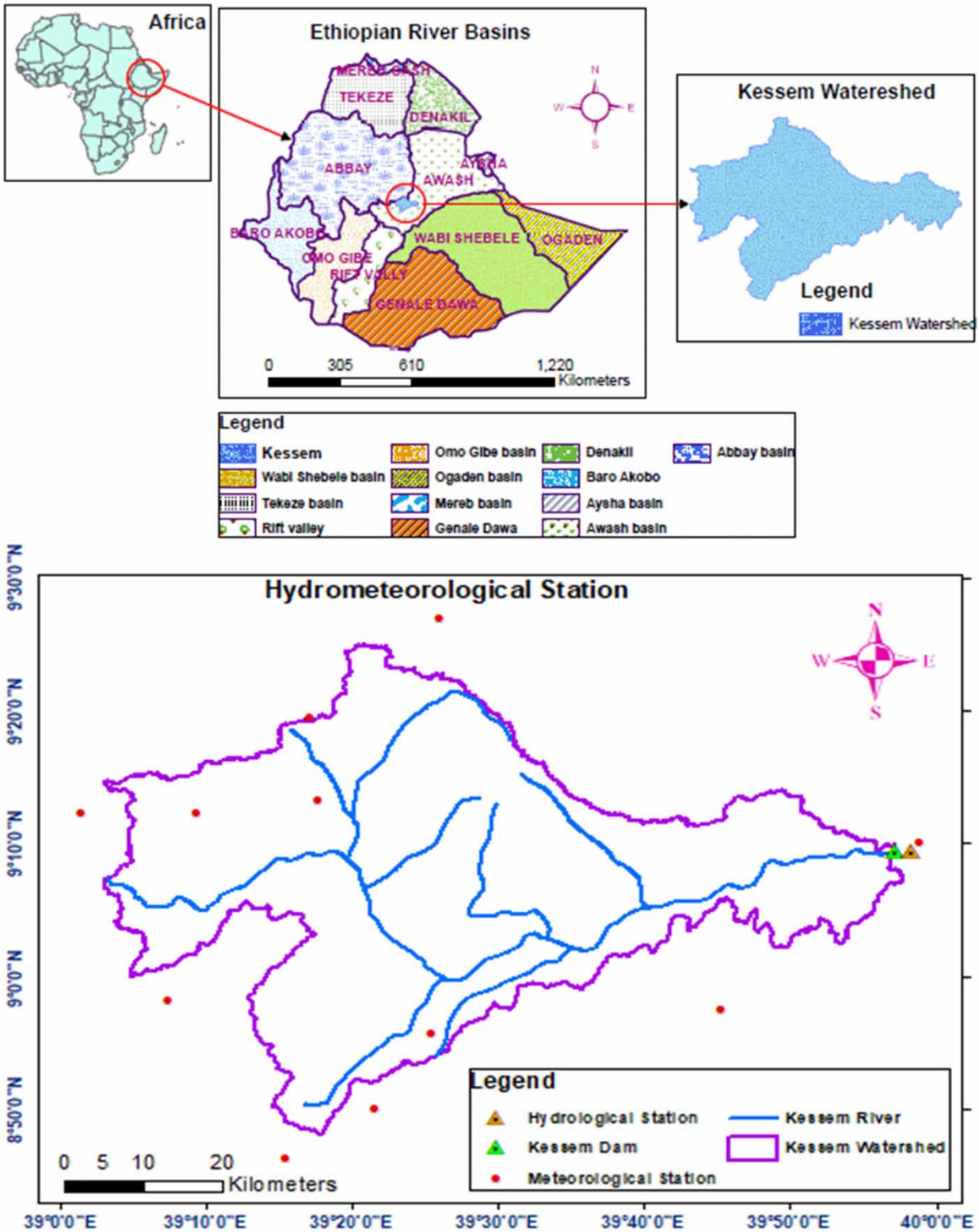
Flooding due to overtopping during peak flow in embankment dams primarily causes dam failure. The Kessem River watershed of the Awash basin in the Rift Valley of the Afar region in Ethiopia was studied intricately to predict the causes of the Kessem Dam safety using machine learning predictive models and risk management centre-reservoir frequency analysis. Recently developed recurrent neural network predictive models with hybrid Soil Conservation Service Curve Number (SCS-CN) were used for simulation of river flow. Peak daily inflow to the reservoir is predicted to be 467.72, 435.88, and 513.55 m^3/s in 2035, 2061, and 2090, respectively. The hydrologic hazard analysis results show 2,823.57 m^3/s and 935.21 m; 2,126.3 m^3/s and 934.18 m; and 11,491.1 m^3/s and 942.11 m peak discharge and maximum reservoir water level during the periods of 2022–2050, 2051–2075, and 2076–2100, respectively, for 0.0001 annual exceedance probability. The Kessem Dam may potentially be overtopped by a flood with a return period of about 10,000 years during the period of 2076–2100. Quantitative hydrologic risk assessment of the dam is used for dam safety evaluation to decide whether the existing structure provides an adequate level of safety, and if not, what modifications are necessary to improve the dam's safety.

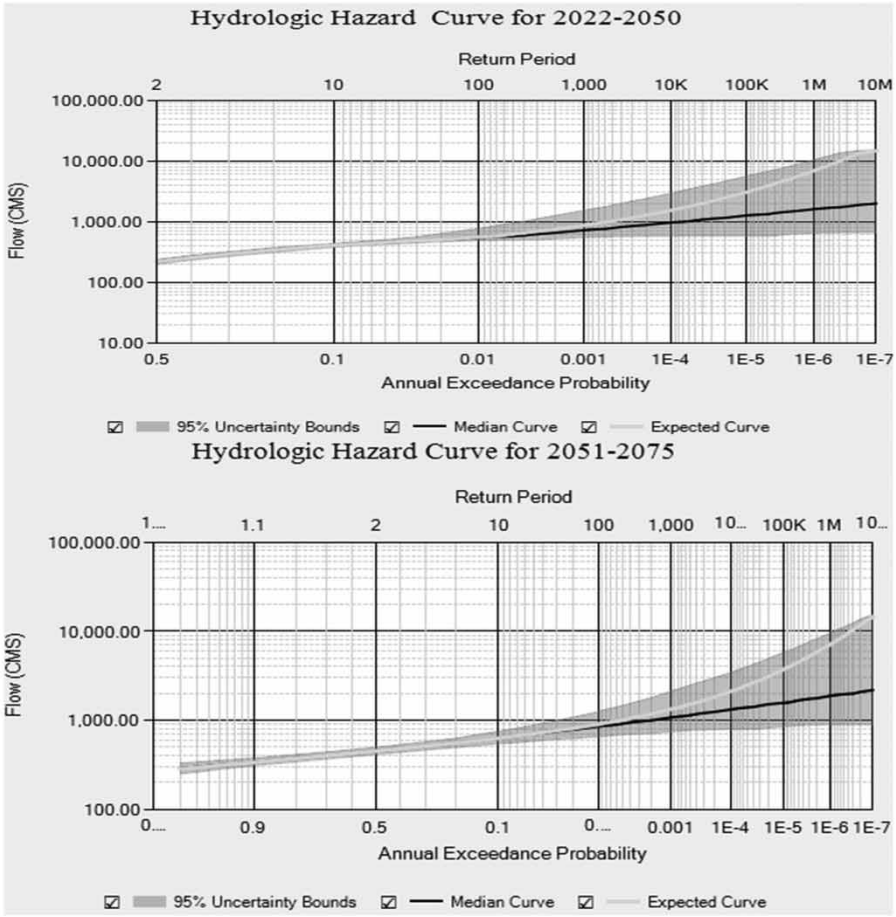
Key words: dam safety, hydrologic hazard analysis, Kessem Dam, machine learning, peak flow prediction

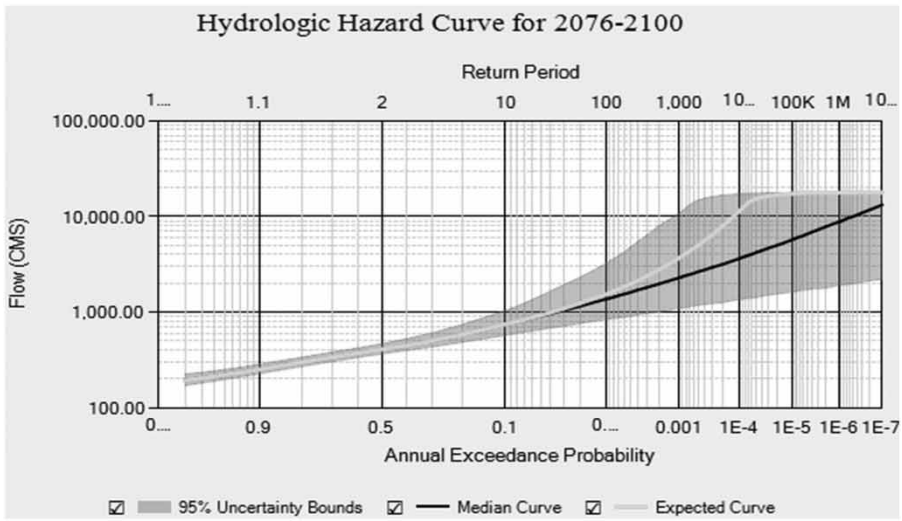
HIGHLIGHT

- The importance of this study will contribute to enhancing the knowledge gap and can provide a piece of information for federal and regional policymakers and water resources implementing agencies about the upcoming flood frequency and its magnitude in the Kessem watershed on the infrastructure to control unexpected floods.

GRAPHICAL ABSTRACT







1. INTRODUCTION

Dams are built to deliver water for human consumption, irrigation and water supply, or to be used in hydropower generation, flood control, recreation, navigation, and sediment control (Jackson & Brown 2023). Dams in Ethiopia are used mainly for hydropower generation (Grand Ethiopian Renaissance Dam, Gibe-I and -III, Tekeze, Koysha, Aba Samuel, Genale Dawa III, Genale Dawa VI, and Melka Wakena), water supply (Geferesa, Legedadi, etc.), and irrigation (Kessem, Tendaho, Arjo

Didessa, Gidabo, Megech, Amerti, Alwero, Geba, Rib, etc.). The Kessem Dam is an embankment dam constructed for irrigation purpose to irrigate 20,000 ha of land for sugarcane agriculture.

Flood is one of the most vulnerable natural disasters caused by peak runoff from the river during the period of extremely high precipitation ravaging the inundated areas (Subramanya 2008; Getahun & Gebre 2015; Khalaf *et al.* 2018), especially affecting the community living close to the river. Changes caused due to increased human activities within the river system, climate change, and changes in land use/land cover in the upstream part of the watershed lead to changes in the watershed along the downstream (Sivapalan *et al.* 2005; Mahmood *et al.* 2010; Vo *et al.* 2016). The main reason for dam overtopping and river channel overflow is the peak flow occurring in the watershed, which increases the reservoir and river channel water level. The erosive action of water flow during overtopping is the primary cause of dam body failure. To effectively control and reduce the effects of these events, the frequency and magnitude of peak flow need to be predicted using advanced techniques (Le *et al.* 2019). Floods in recent decades have occurred recurrently in various parts of Ethiopia (Assefa 2018). For instance, a disastrous flash flood downstream was caused by an unexpected peak flow supplemented by excessive storage beyond the designed capacity in the Awash basin, which overwhelmed Meteka town in Ethiopia (Shumie 2019).

Mosavi *et al.* (2018) discussed flood prediction using machine learning (ML) models, which focused on the state-of-the-art ML models in flood prediction and gave insight into the most appropriate models. They particularly investigated to provide an extensive overview of the various ML algorithms used in the field; finally, they concluded that it is the most auspicious prediction method for both long-term and short-term floods. In addition, machine learning models are the most simple, powerful, robust, and predictive models in water systems associated with data-driven techniques (Hosseiny *et al.* 2020) for mapping the nonlinear relationship between rainfall and runoff, even though they cannot represent the physical process of the catchment (Kratzert *et al.* 2018). These models establish the relationship between input and output based on data driving techniques and contain no physical transformation function to relate input to output. Thus, sometimes, the direct use of data-driven models does not provide appropriate results as the physical processes are ignored in the model. Those ignored characteristics are the moisture content (AMC) and physical characteristics of the catchment, such as geology, soil, slope, and land use/land cover (LULC) conditions (Mahmood *et al.* 2010). To overcome the above lacunae, a hybrid physics-based (SCS-CN) model (which is able to represent the spatial variability of land surface characteristics such as LULC and soil type) has been incorporated to increase the performance of machine learning models in predicting the peak flow.

In addition, in recent years, models based on machine learning techniques have been developed/analyzed for forecasting reservoir inflow, estimating instantaneous peak flow, and analyzing hydrological risk (Jimeno-Sáez *et al.* 2017; Gabriel-Martin *et al.* 2019; Hong *et al.* 2020; Huang *et al.* 2022; Tian *et al.* 2022). Senent-Aparicio *et al.* (2019) estimated the instantaneous peak flow by combining the Soil and Water Assessment Tool (SWAT) and machine learning models. This study includes addressing the issue of accurate prediction of peak flow with uncertainty reduction and flood risk analysis on the Kessem Dam, which was carried out using machine learning models (such as recurrent neural network (RNN) models) for accurate flood prediction, and assessing the risk of these events on the dam's safety by uploading the predicted flood on the Kessem reservoir and then assessing the risk by using advanced reservoir simulation techniques (such as Risk Management Centre-Reservoir Frequency analysis (RMC-RFA) software) at present as well as future times. This study finding will be used as an input parameter for good decision-makers to plan and design flood control structures and to take proper flood control measures in the watershed during the flood season.

2. DATA AND METHODOLOGY

2.1. Description of the study area

The Awash River Basin is one of the most important river basins in Ethiopia; the Kessem River is the most significant tributary of this basin at the middle section, which is the southern end of the Afar rift in the Afar regional state of Ethiopia, 225 km east of Addis Ababa and 40 km northwest of Metehara town. The Kessem Dam, located at 9°8′45″ latitude and 39°55′31″ longitude, stores half a billion cubic meter of water in the reservoir for irrigation purposes to irrigate 20,000 ha of land under the Kessem Dam irrigation project. Figure 1 shows the location of the study area.

2.2. Data collection

Hydro-meteorological data (stream flow data, precipitation, and temperature data), digital elevation model (DEM), land used/cover (LULC), and soil type map data were collected for this study to predict flood amount and assessment of the risk on dam

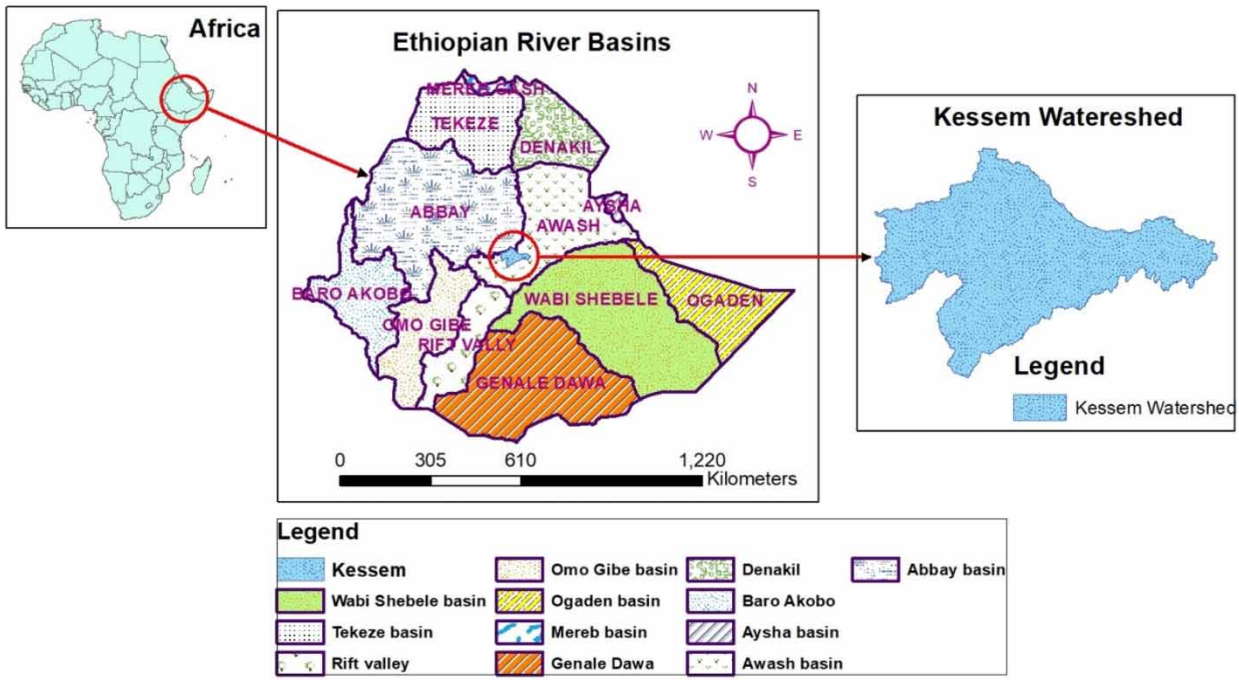


Figure 1 | Topographical map of the study area.

safety. Daily precipitation data from 1988 to 2018 years of 15 stations in and around the Kessem watershed were collected from Ethiopian Meteorological Institute (EMI). However, 11 rainfall stations were selected for this study based on the availability and continuity of data that contribute to the Kessem River. Climate data such as maximum and minimum temperature for six stations were collected from EMA. Twenty-one years from 1990 to 2009 and 4 years from 2010 to 2013 of daily stream flow data for Kessem River at Aware Melka and Kessem Dam stations, respectively, were procured from the Ministry of Water, Irrigation and Energy (MoWIE) hydrology department. Recorded water levels for the Kessem Dam reservoir from 2018 to 2021 were collected from the Ethiopian Construction Works Cooperation (ECWC) dam and irrigation development centre. 1 arc second (30 m × 30 m) resolution Advanced Space borne Thermal Emission and Reflection Radiometer Global-DEM (ASTER GDEM) was downloaded from the United States Geological Survey (USGS) Earth Explorer data (https://earthexplorer.usgs.gov/), the Kessem sub-basins of the middle part of Awash were extracted from Awash DEM following the procedural steps of watershed processing to be used for further analysis (see Figure 2).

2.3. Materials/tools used

ArcGIS software was used to analyze spatial data for selected models, artificial intelligence (AI), and machine learning models for prediction of a flow/flood, Statistical Downscale Model version of SDSM 4.2.9 (Wilby & Dawson 2007) to downscale climate information from coarse-resolution of GCMs to local or site level, Python 3.9, and Jupyter notebook for writing and running the code, Matplotlib (Hunter 2007) for data visualization, Hydrostats packages for evaluation of model performance (Roberts *et al.* 2018), RMC-RFA software for reservoir routing and analysis of hydrologic hazard on dam, SFE_IFC MATLAB toolbox to determine the flood characteristics (start and end date, peak and duration) and to develop flood hydrograph (Zhang *et al.* 2021) and Hydrological Engineering Centre-Statistical Software Packages (HEC-SSP 2.2) software for volume frequency analysis (Bartles *et al.* 2019).

2.4. Methodology

The observed climate data and flow used to calibrate and validate the selected models were collected from National meteorological service agency of Ethiopian and Ministry of Water and Energy. DEM data and Landsat images were used as input for reclassified LULC approach processed by Arc GIS, Arc hydro tools for the catchment delineation and estimation of catchment characteristics. The course climate data (GCM) downloaded from the Canadian Climate Data and Scenarios

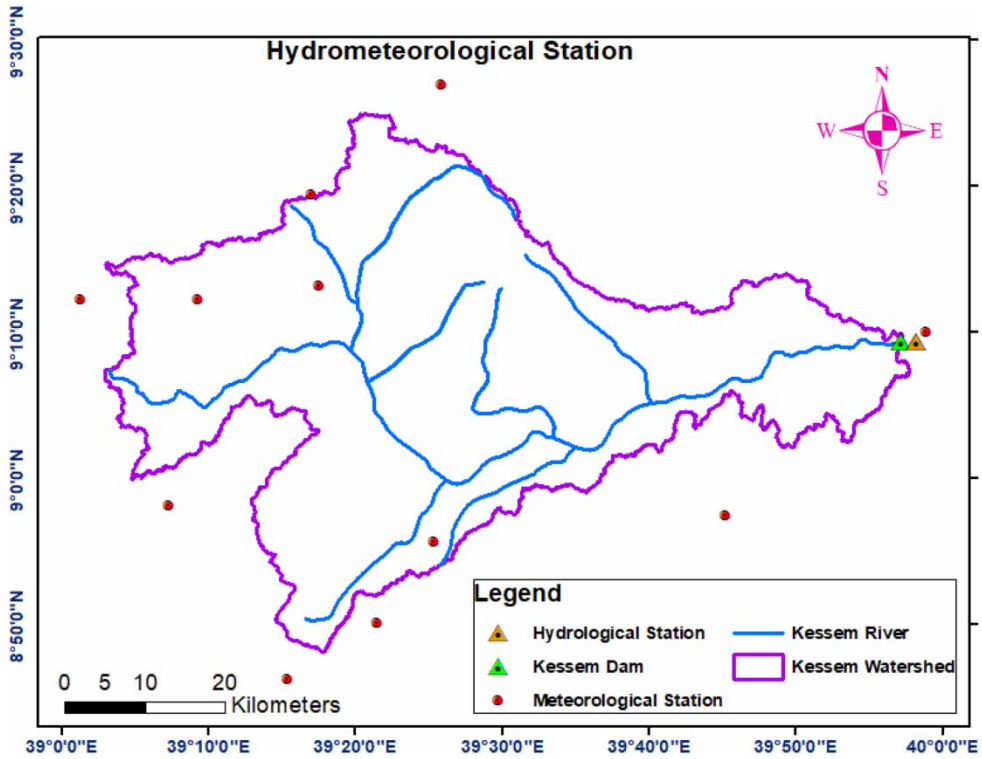


Figure 2 | Hydro-metrological stations.

(CCDS) portal CanESM2 model outputs for the study area were downscaled into finer spatial resolution at watershed level by bias correction through SDSM statistical approach and by using the selected potential predictors projected the future climate (precipitation and temperature). The Landsat images with different bands were downloaded from USGS for the study area and reclassified the LULC class in watershed used supervised classification techniques (Gebeyehu *et al.* 2019a). The climate projected data under the climate change scenario and LULC were used as input for the SCS-CN model to estimate runoff at each sub-watershed outlet. The output of SCS-CN model was used as input for ML models to predict flow at the Kessem Dam watershed outlet and estimated flood event within future three-time horizons. The results were imported to RMC-RFA software to assess future hydrological risk of the selected flood events on Kessem reservoir for dam safety evaluation. Moreover, the overall accuracy and kappa coefficient were calculated by the following equations (Stehman 1996):

$$\text{Overall Accuracy (\%)} = \frac{\text{Total Number of Correctly Classified Pixels (Diagonal)}}{\text{Total Number of Reference (Ground Truth) Pixels}} * 100 \quad (1)$$

$$\text{Kappa Coefficient (K)} = \frac{(\text{TS} * \text{TCS}) - \sum(\text{Column Total} * \text{Row Total})}{\text{TS}^2 - \sum(\text{Column Total} * \text{Row Total})} \quad (2)$$

Effective rainfall, potential evapotranspiration (Mahmood *et al.* 2010), and stream flow data are the main input datasets that were used in ML model. All these datasets used are observation data and estimated data based on observation to calibrate and validate the models.

The effective rainfall was estimated by using SCS-CN model considering the characteristics of sub-watershed of the Kessem River. The PET was also computed using Hargreaves methods based on the observed maximum and minimum temperature data at each available station. The complete stream flow data for the Kessem River at the Kessem Dam during the observed period of 1990–2013 were transformed by using area ratio transform techniques from the Aware Melka station to the Kessem Dam during the period of 1990–2009 and the observed stream flow at the Kessem Dam during the period of 2010–2013. Based on the climate projection data and LULC scenario, future flow was predicted by using the performed ML models

with a hybrid SCS-CN model. Using the daily time series data, the model was constructed using Keral Tensor Flow packages in python 3. Training and testing were performed for the period 1990 to 2013, for which observed discharge data are available. In the network modeling, out of the total data, 70% (January 1990–October 2006) were selected for training and 30% (November, 2006–December, 2013) each for testing.

3. RESULTS AND DISCUSSIONS

3.1. Climate projection and LULC changes

3.1.1. Climate projection at the future

The climate data (temperature and precipitation) projection in future within the watershed have been studied using CanESM2 climate model for RCP2.6, RCP4.5, and RCP8.5 climate scenarios from coupled model inter-comparison project-5 (CMIP5) experiments which have been downscaled by the statistical downscaling model (SDSM). After trial and error to get the highest model performance by changing the values of bias correction and variance inflation in the SDSM model for precipitation, maximum temperature, and minimum temperature, the statistical results were estimated (Tables 1–3), and the mean values of the graphical results were drawn (Figures 3–5). The selected potential predictors for calibrating the model were ncepp8_ugl, ncepp8_thgl, nceps500gl, ncepshumgl, and nceptempgl, with 1.356 bias correction and 12 values of variance inflation for precipitation, ncepp1_ugl, ncepp1thgl, nceps500gl, nceps850gl, ncepshumgl, and nceptempgl predictors with the values of bias correction and variance inflation of 1 and 12, respectively, were used for the model calibrated for minimum temperature, and

Table 1 | The performance results from the SDSM model for downscaled precipitation after taking different trials and errors

Period		RMSE	NSE	*R*
Calibration		3.795	0.319	0.584
Validation		1.529	0.309	0.597
Testing	RCP8.5	3.446	0.324	0.625
	RCP4.5	3.429	0.331	0.613
	RCP2.6	3.371	0.353	0.601

Table 2 | The performance results of the SDSM model for downscaled minimum temperature after taking different trials and errors

Period		RMSE	NSE	*R*
Calibration		1.586	0.509	0.722
Validation		1.514	0.582	0.779
Testing	RCP8.5	1.722	0.312	0.662
	RCP4.5	1.697	0.332	0.670
	RCP2.6	1.678	0.346	0.676

Table 3 | The performance results from the SDSM model for downscaled maximum temperature after taking different trials and errors

Period		RMSE	NSE	*R*
Calibration		1.265	0.489	0.715
Validation		1.179	0.522	0.780
Testing	RCP 8.5	1.443	0.114	0.624
	RCP 4.5	1.418	0.144	0.646
	RCP 2.6	1.429	0.129	0.628

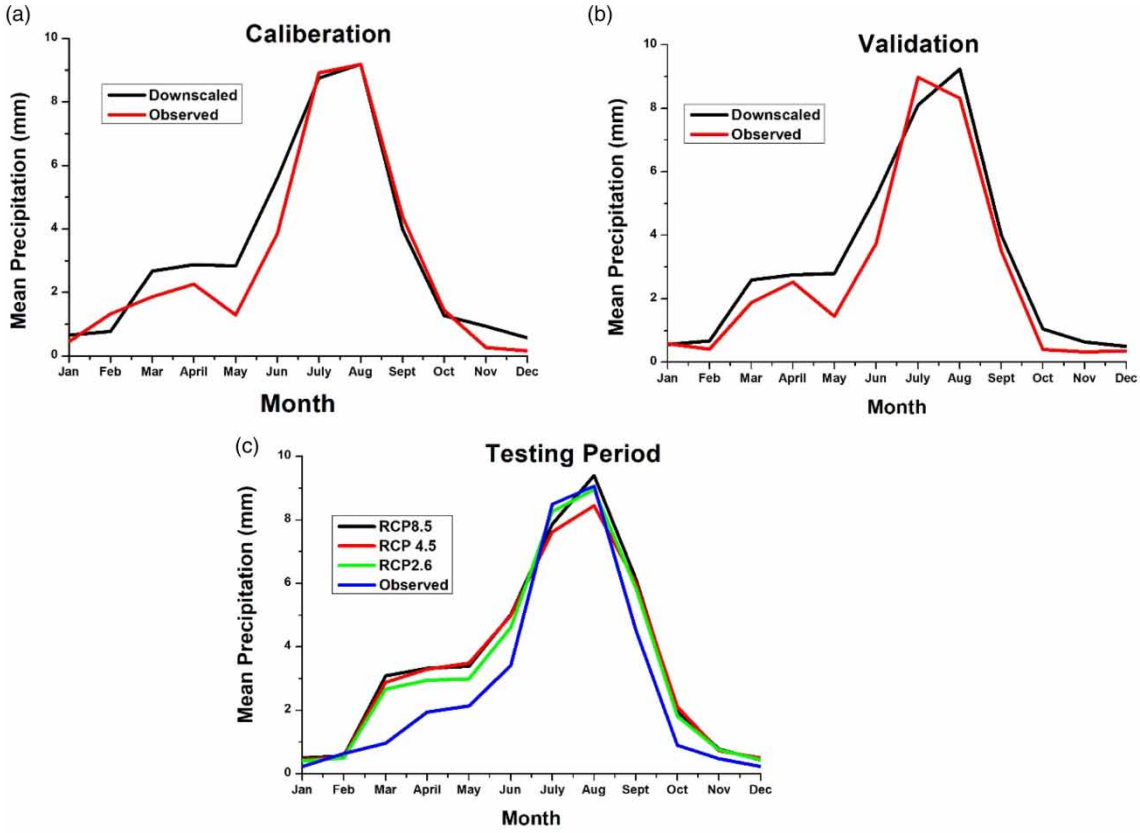


Figure 3 | Downscaled vs. observed mean precipitation: (a) for model calibration period (1988–1999), (b) for validation period (2000–2005), and (c) for testing period (2006–2018).

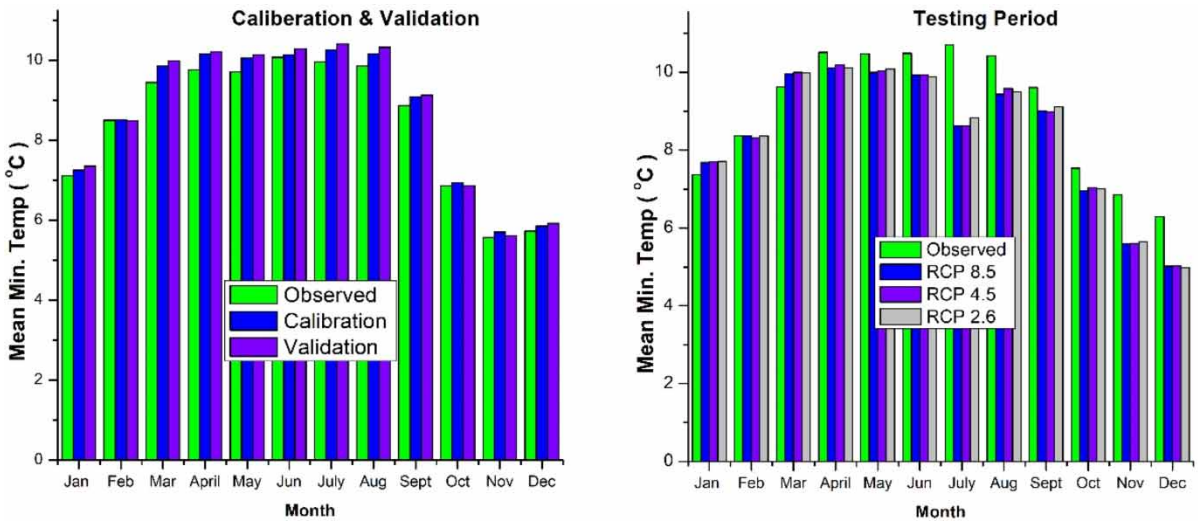


Figure 4 | Downscaled vs. observed mean minimum temperature for model calibration period (1988–1999), validation period (2000–2005), and testing period (2006–2018).

nceps500gl, nceps500gl. ncepp1_zgl, ncepp5_fgl, ncepp5_vgl, ncepp500gl, ncepp5thgl, ncepp8_vgl, ncepp8_zgl, and nceptempgl predictors with the values of bias correction and variance inflation of 1 and 12, respectively, were used for model calibrating, validating, and testing for maximum temperature (Gebeyehu *et al.* 2019b).

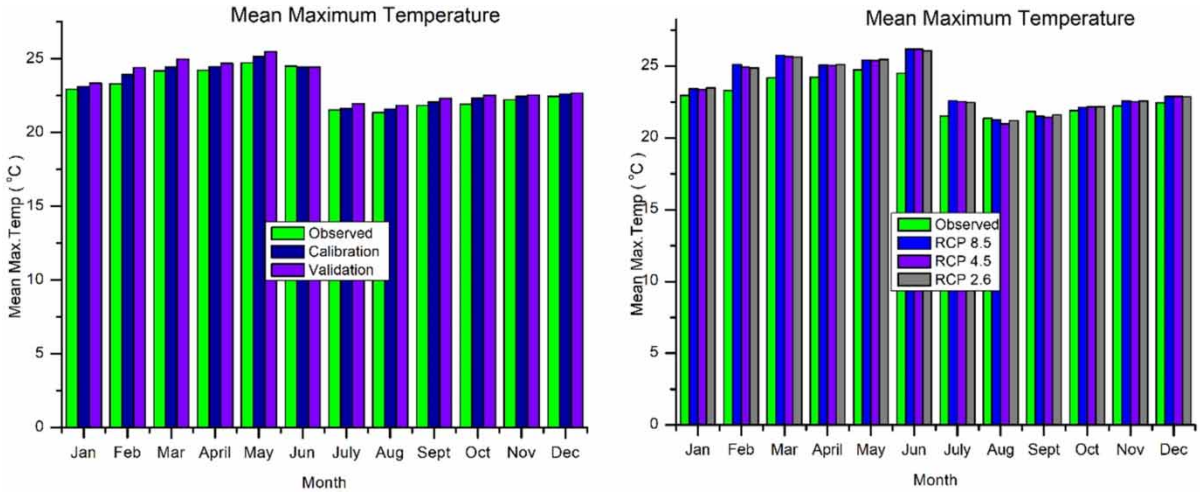


Figure 5 | Downscaled vs. observed mean maximum temperature for model calibration period (1988–1999), validation period (2000–2005), and testing period (2006–2018).

After calibrating and validating the model, statistical evaluated values of RMSE, NSE, and *R* were found to be 3.446, 0.324 and 0.625, 3.429, 0.331 and 0.613, 3.371, 0.353, and 0.601, respectively, for model performance downscaled precipitation during testing period under RCP8.5, RCP4.5, and RCP2.6 climate scenarios, respectively. The SDSM model super performed during testing period under the RCP2.6 climate scenario to downscaled precipitation, however, this scenario used to project the precipitation data at future time horizons during 2022–2050, 2051–2075 and 2076–2100 in the watershed varies a lot.

The statistical evaluated values of RMSE, NSE, and *R* were 1.722, 0.312, and 0.662, respectively under RCP8.5, 1.697, 0.332, and 0.670, respectively, under RCP4.5, and 1.678, 0.346, and 0.676, respectively, under the RCP2.6 climate scenario (Table 2) during testing period. From analysis, the SDSM model has super performed during testing period under the RCP2.6 climate scenario to downscaled minimum temperature, however, this scenario used to project the minimum temperature data at future time horizons from 2022 to 2050, 2051 to 2075, and 2076 to 2100.

The evaluated values of RMSE, NSE, and *R* were 1.443, 0.114, and 0.624, respectively, under RCP8.5, 1.418, 0.144, and 0.646, respectively, under RCP4.5, and 1.429, 0.129, and 0.628, respectively, under the RCP2.6 climate scenario (Table 3) for the model performance of downscaled maximum temperature during testing period. From the analysis, the SDSM model has super performed during testing period under the RCP4.5 climate scenario to downscaled maximum temperature, therefore, this scenario was used to project the maximum temperature data at future time horizons from 2022 to 2050, 2051 to 2075, and 2076–2100.

3.1.2. LULC changes and scenario at the future

We analyzed the LULC map in ArcGIS 10.5 by using the Landsat 8 and Landsat 7 images downloaded from USGS for path 168, row 54 with different bands at different acquired years (2000, 2010, and 2020). The download images were declassified by using the supervised classification method to seven different LULC types namely, agricultural lands, bare lands, forests areas, grasslands, settlement areas, shrub lands, and water bodies for each acquired year (Figure 6).

Based on the collected sample data confusion matrix, the total sample points (TS) are 71, and the total corrected classified (TCS) value is 57, the sum of the product values in the total ground truth column and in the total user row is 1,101. Substituting those values into Equations (1) and (2), the overall accuracy and kappa coefficient are 80.3% and 0.75, respectively. This implies 80.3% of land use and land cover classes are correctly classified.

A LULC change detection study was performed by the supervised classification method using the maximum likelihood classifier algorithm in ArcGIS 10.5 software during the period 2000–2020. Table 4 shows the changing area covering each LULC class for the past 20 years.

Quantitative analysis of the overall LULC changes decreases and increases in each class between 2000 and 2020. A considerable decrease in forest (0.6%), grassland (3.85%) and bare land area (1.23%), and shrub lands (0.56%) per year was observed during this period (Table 5). On the other hand, there is an increase in agriculture lands (4.42%), settlement

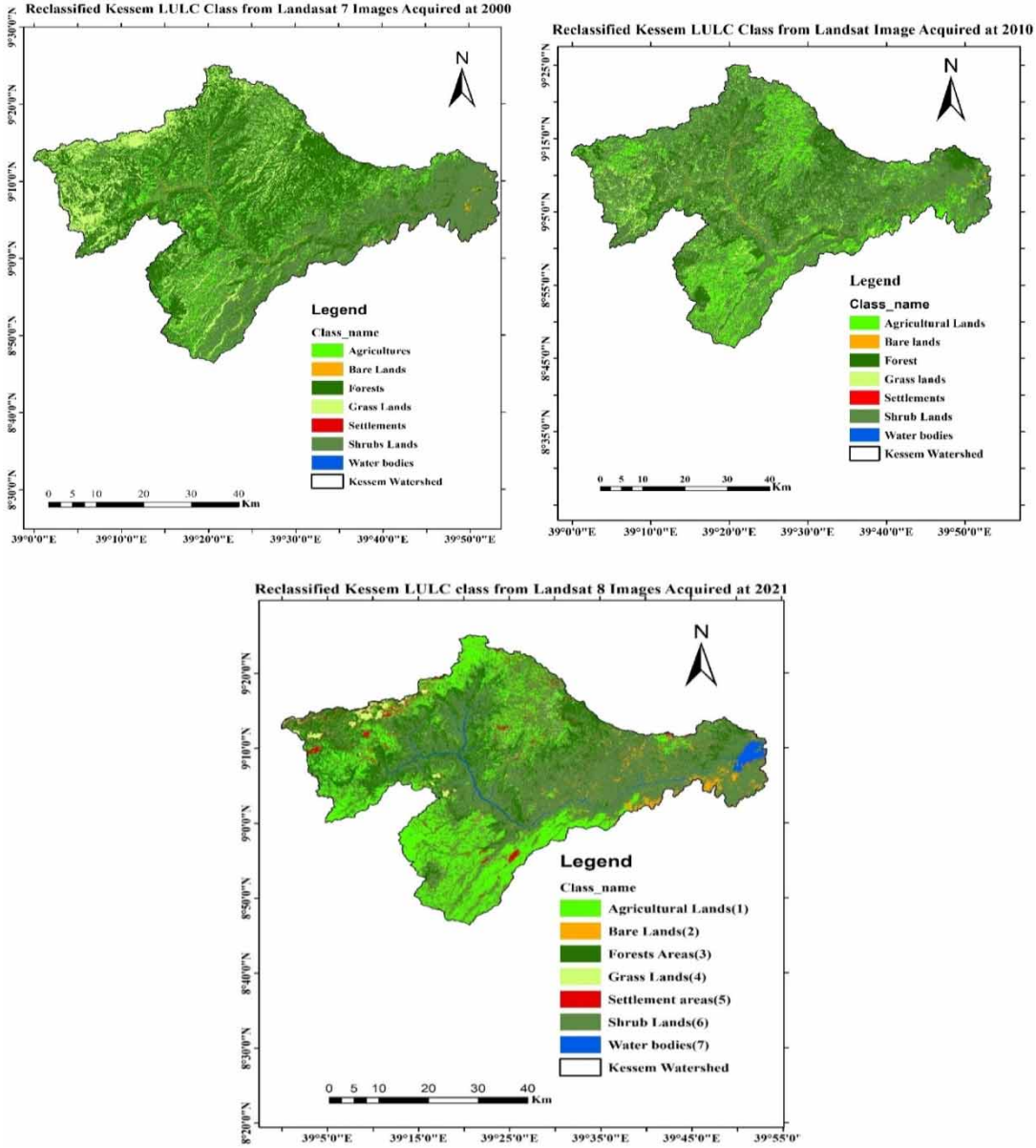


Figure 6 | Reclassified Kessem watershed LULC class based on the used Landsat Image for the periods 2000, 2010, and 2021.

Table 4 | Changes in LULC from 2000 to 2020

Class name	Areas covered in 2000 (km^2)	Areas covered in 2010 (km^2)	Areas covered in 2020 (km^2)	Change in %/year 2000–2010	Change in %/year 2010–2020	Change in %/year 2000–2020
AG	422.3	541.23	795.97	2.82	4.71	4.42
BL	16.46	15.681	12.414	−0.47	−2.08	−1.23
F	613.6	455.23	540.33	−2.58	1.87	−0.6
GL	241.6	219.87	55.768	−0.9	−7.46	−3.85
S	68.97	80.38	93.197	1.65	1.59	1.76
SL	1,614.6	1,664.6	1,432.9	0.31	−1.39	−0.56
WB	0.004	0.516	46.926	14.24	8.99	6.52

Table 5 | Optimal hyper-parameter networks

Hyper-parameter	Values
Neuron	12
Optimization	Adam
Learning rate	0.001
Activation function	Sigmoid and Tanh
Max Epoch	4,000
Batch size	128

areas (1.76%), and surface water bodies (6.52%) for the same time period. Based on the analysis, the future LULC change scenarios in the Kessem watershed for each class were decided as follows:

Scenario 1: Forest, bare lands, and shrub lands area have been reduced and all grassland areas were covered by agriculture lands, settlement areas and surface water bodies during the period 2022–2050.

Scenario 2: Under this scenario, further reductions have been made in forest, bare lands, and shrub lands area for the period 2051–2075. These will be afterwards covered by agriculture lands, settlement areas and surface water bodies.

Scenario 3: Reduction has been predicted on the area covered by agriculture lands resulting in the formation of bare land during the period 2076–2100 and in addition forest and shrub lands will get reduced, increasing the settlement area and surface water bodies.

3.2. Flow/peak flow prediction at the future time horizons

Three different deep-learning methods (LSTM, Bi-LSTM, GRU) for flow prediction during the historical period were implemented. Prediction models were subsequently applied to predict flow for calibration and validation period, and their performance was measured. Daily stream flow to the Kessem Dam reservoir in Kessem watershed was simulated using various deep-learning models. The historical observation stream flow data were compared with the computed stream flow from RNN models (LSTM, Bi-LSTM, and GRU) using 30 lag days. A network was attempted to predict outcomes as accurately as possible. The value of this precision in the network is obtained by the cost function, which tries to penalize the network when it fails. The optimal output is the one with the lowest cost. For the applied networks of Mean Square Error (Shoaib *et al.* 2014), the cost function is used. A repetition step in training generally works with a division of training data named as a batch size. The number of samples for each batch is a hyper-parameter, which is normally obtained by trial and error. The value of this parameter in all models is 128 in the best mode. In each repetition step, the cost function is computed as the mean MSE of these 128 samples of observed and predicted stream flow. The number of iteration steps for neural networks is named as an epoch and in each epoch, the stream flow time series is simulated by the network once. Like other networks, neurons or network layers can be selected arbitrarily in recurrent networks. For the purpose of the comparison of models with each other, the structures of all recurrent network models are created identically. In each network, a double hidden layer is used so that there are 12 units in each for the first layer and the second layer. The last layer output of the network at the final time step is linked to a dense layer with a single output neuron. Between the layers, a dropout equal to 10% is used. The structure of the neural network is also used in two hidden layers. The first and second layers have 12 neurons each. In all networks, the sigmoid activation function is applied for the hidden layer. The main advantage of using sigmoid is that, for all inputs greater than 0, there is a fixed derivative. This constant derivative speeds up network learning. Each method is run with different epoch numbers. After different trials and errors, the optimal hyper-parameter networks provide the details (Table 5).

The optimized model results were evaluated using Hydrostats packages with statistical error assessment techniques. In Hydrostats, the statistical as well as graphical evaluations are made using error metrics function between observed and simulated flow. Graphically by plotting the predicting and observed flow (Figure 7). Several descriptive statistics can be used for evaluation of predictive models. In this study, RMSE, NSE, R^2 have been purposively used, and the results of different methods based on the evaluation criteria are presented in Table 6. Among the RNN methods, Bi-LSTM performed the best.

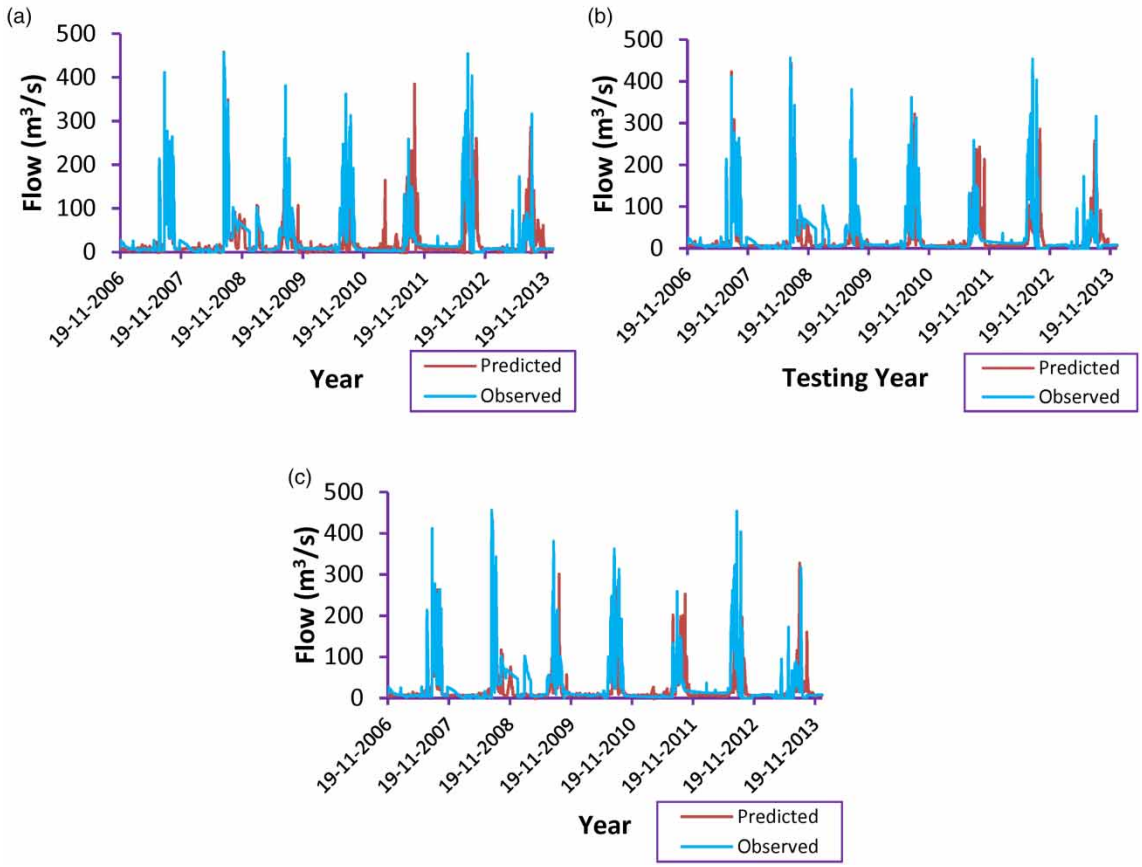


Figure 7 | The performance of predictive models results for training and testing periods: (a) bi-LSTM model, (b) LSTM model, and (c) GRU model.

Table 6 | Statistical evaluation performance of ML models

	Training period (1990–2006)				Testing (2006–2013)			
ML models	**RMSE**	**NSE**	**R^2**	**KGE**	**RMSE**	**NSE**	**R^2**	**KGE**
Bi-LSTM	3.873	0.968	0.994	0.702	17.547	0.744	0.749	0.832
LSTM	4.276	0.962	0.974	0.777	19.878	0.672	0.727	0.690
GRU	4.003	0.966	0.982	0.743	20.703	0.644	0.681	0.712

The calculated discharges match well with the observed, as indicated by the high NSE and the small RMSE values for the overall evaluation of three ML models, revealing that the Bi-LSTM models outperform than LSTM and GRU. Therefore, in this study the result of the Bi-LSTM model to predict the flow of the Kessem River at the Kessem Dam within three-time horizons is used.

3.3. Hydrologic hazard analysis for the Kessem Dam

3.3.1. Inflow hydrograph shape

During 2022–2100 the future period of inflow to the Kessem Dam, the selected three events on September 2035 from 2022 to 2050 time horizon, September 2061 from 2051 to 2075 time horizon, and September 2090 from time horizon of 2076–2100 are the largest peak flow events that may occur in the Kessem Dam watershed within the next 100 years.

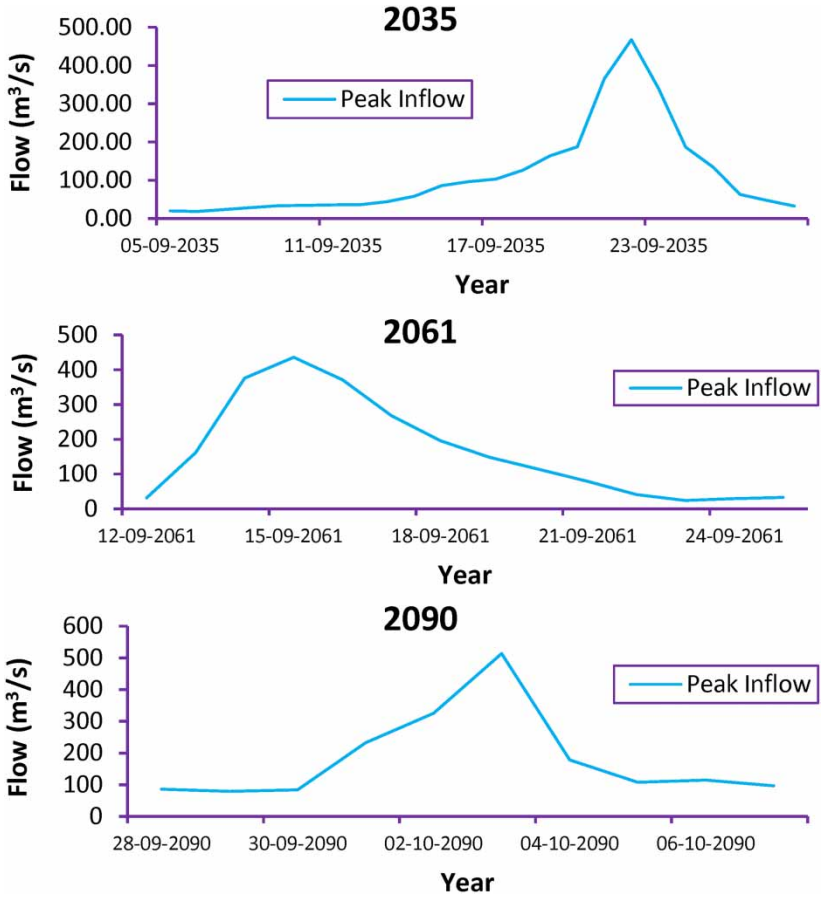


Figure 8 | Peak inflow hydrograph of the 2035 flood event during 2022–2050, the 2061 flood event during 2051–2075, and the 2090 flood event during 2076–2100.

Inflow hydrograph shapes for the three-time horizons shown in Figure 8 were derived using the results from the predicted ML model. September 2035, September 2061, and September 2090 events were used for rescaling the sampled inflow flood events.

The PMF hydrograph was developed based on 9,237 m^3/s (Q_p) of the design of the peak inflow PMF for a 10,000-year return period of the Kessem Dam by using SCS dimensionless methods with a time to peak (T_p) of 33.39 h, including the watershed characteristics ($L = 136.64$ km $= 448{,}294$ ft, Sl $= 1.238\%$, CN $= 79.41$ and $T_c = 51.016$ h) to compute discharge Q and the corresponding time t (Figure 9), which depicts the PMF hydrograph that represent the computed value of discharge

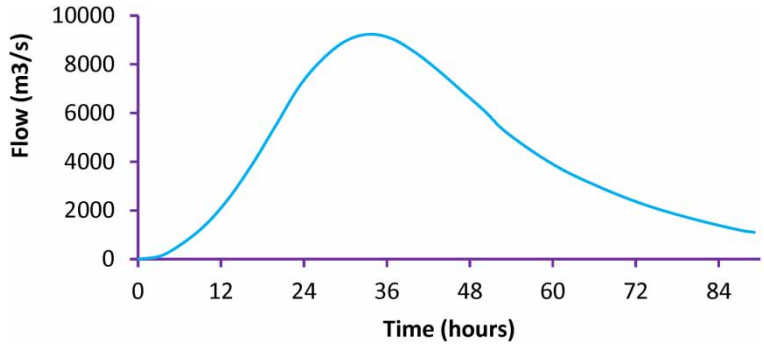


Figure 9 | The PMF hydrograph.

Q versus time t. This hydrograph was used to compare with the results of peak discharge and stage frequency from hydrologic hazard analysis and to determine whether future flood events on the Kessem Dam are at risk or not.

3.3.2. Inflow volume frequency curve

The developed volume frequency curve of the Kessem Dam is based on the Log Pearson Type III distribution with mean, standard deviation, skew coefficient, and the effective record length values for the future within three-time horizons. For the volume frequency analysis, the Bulletin 17C with EMA analysis was performed using HEC-SSP (Table 7).

Based on the result of volume frequency analysis, the volume frequency curve within 90% uncertainty bounds were computed for each of the corresponding future time horizons (Figure 10).

3.3.3. Flood seasonality analysis

For the threshold value of 206.4 m^3/s, the frequency sample size during the period 2022–2050, 2051–2075, and 2076–2100 is 35, 31, and 37, respectively. Those are adequate sample size to analyze the flood seasonality for each time horizons. The flood seasonality histogram developed for this analysis for each time frame is presented in Figure 11. According to the result analysis, the annual flows normally from October through May, with June–September as the wettest months but flood seasonal month is August during the period of 2022–2050, while the annual flows normally flow from November through June, with July–October as the wettest months but flood seasonal month is September during both the periods 2051–2075 and 2076–2100.

3.3.4. Reservoir starting-stage duration analysis

Initial reservoir levels and associated exceedance probabilities were estimated from daily reservoir elevation estimates for the period of record. The duration curve results indicate that a median reservoir elevation for the June through October period is approximately 926 m, with a quartile range (25 to 75 percentiles) from about 922 to 928 m (Figure 12). This reservoir elevation range was considered as initial reservoir water surface elevations for routing the hydrographs. From the results, August produces the lowest pool duration curve. From the flood seasonality analysis section it has been inferred that floods are most likely to occur in August and September. However, the dam is operated with consideration of this flood seasonality. Therefore, large events are most likely to occur in August and September, but they are also most likely to have low reservoir starting pools, mitigating some of the risk for large peak stage events in the summer season.

3.3.5. Hydrologic hazard curve of the Kessem Dam

Once RMC-RFA is computed, it automatically creates the Stage Frequency Curve and Hydrologic Hazard Curve plots (Figures 13 and 14). The median curve represents the uncertainty in stage frequency and peak discharge frequency due to natural variability. The 95% uncertainty bounds represent the uncertainty in stage and peak discharge frequency due to knowledge uncertainty, whereas the expected curve represents the combined uncertainty due to both natural variability and knowledge uncertainty. Those curves are used for semi-quantitative risk analysis for the Kessem Dam.

3.4. Discussion

Hydrological hazard analysis produces the expected peak discharge and the corresponding peak stage for 100 to 1,000,000 year of return period for each time horizons (Tables 8 and 9), respectively.

The Kessem Dam has a spillway discharge capacity of 6,180 m^3/s at the maximum water surface elevation of 939.5 m. Comparing this value with the stage and peak discharge frequency curve it indicates that the spillway is capable of passing a flood with a return period of 100–100,000 years for the future time horizons (2022–2075).

Table 7 | The result of statistics from volume frequency analysis for each future time horizons

Period/Statistics	Mean (of log)	Standard deviation (of log)	Skew (of log)	Effective record length
2022–2050	2.446	0.125	0.654	29
2051–2075	2.375	0.107	0.450	25
2076–2100	2.461	0.174	1.087	25

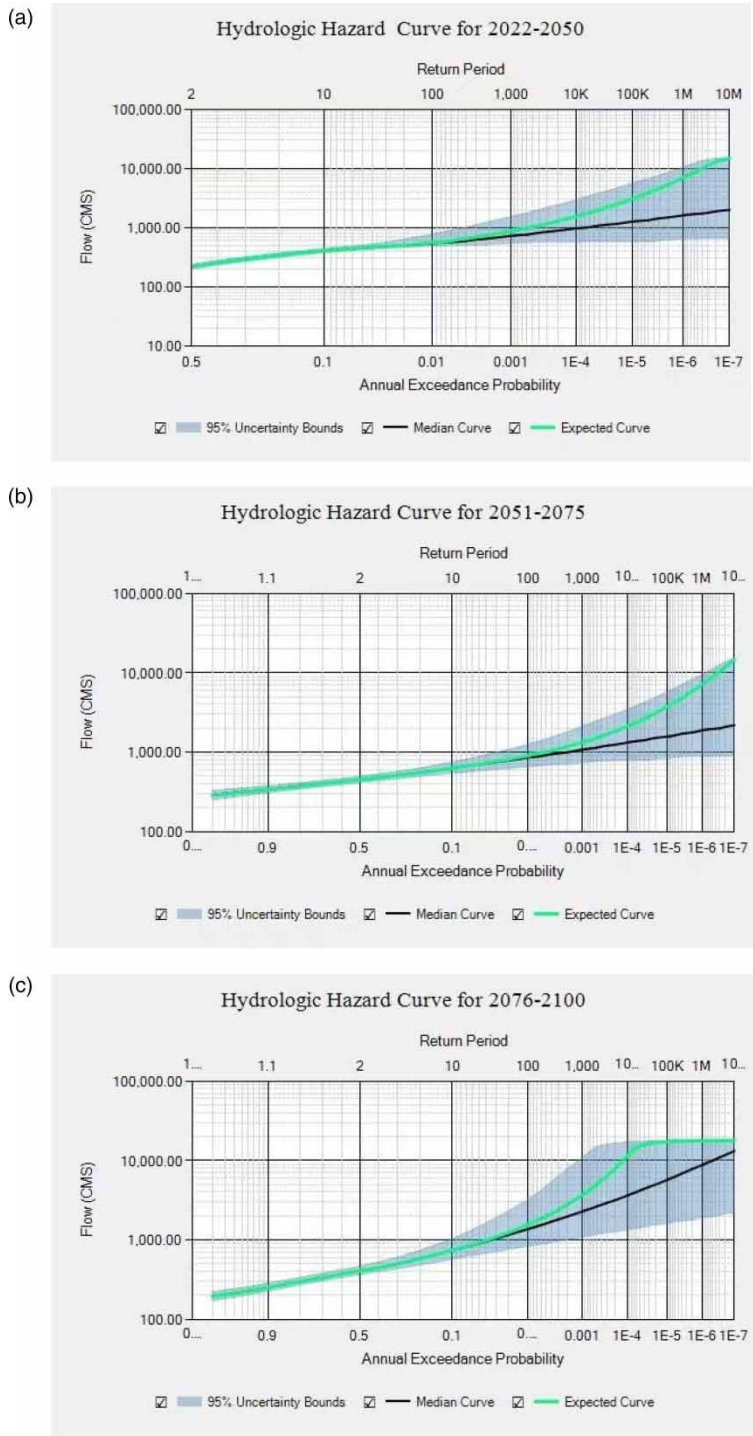


Figure 10 | Volume frequency curve during the period: (a) 2022–2050, (b) 2051–2075, and (c) 2076–2100.

During the period of 2022–2050, the expected peak discharge for 1/10,000 AEP was equal to 2,823.57 m^3/s. The 10,000-year peak discharge at 95% confidence upper and lower limits is 987 and 5,430.6 m^3/s, respectively, from the hydrological hazard analysis. The corresponding expected peak stage for 0.0001 APE is 935.21 m, and the lower and upper 95% of bounds values are 932.1 and 938.4 m, respectively. It has not exceeded the PMF discharge of 6,180 m^3/s and the maximum water surface elevation of 939.5 m.

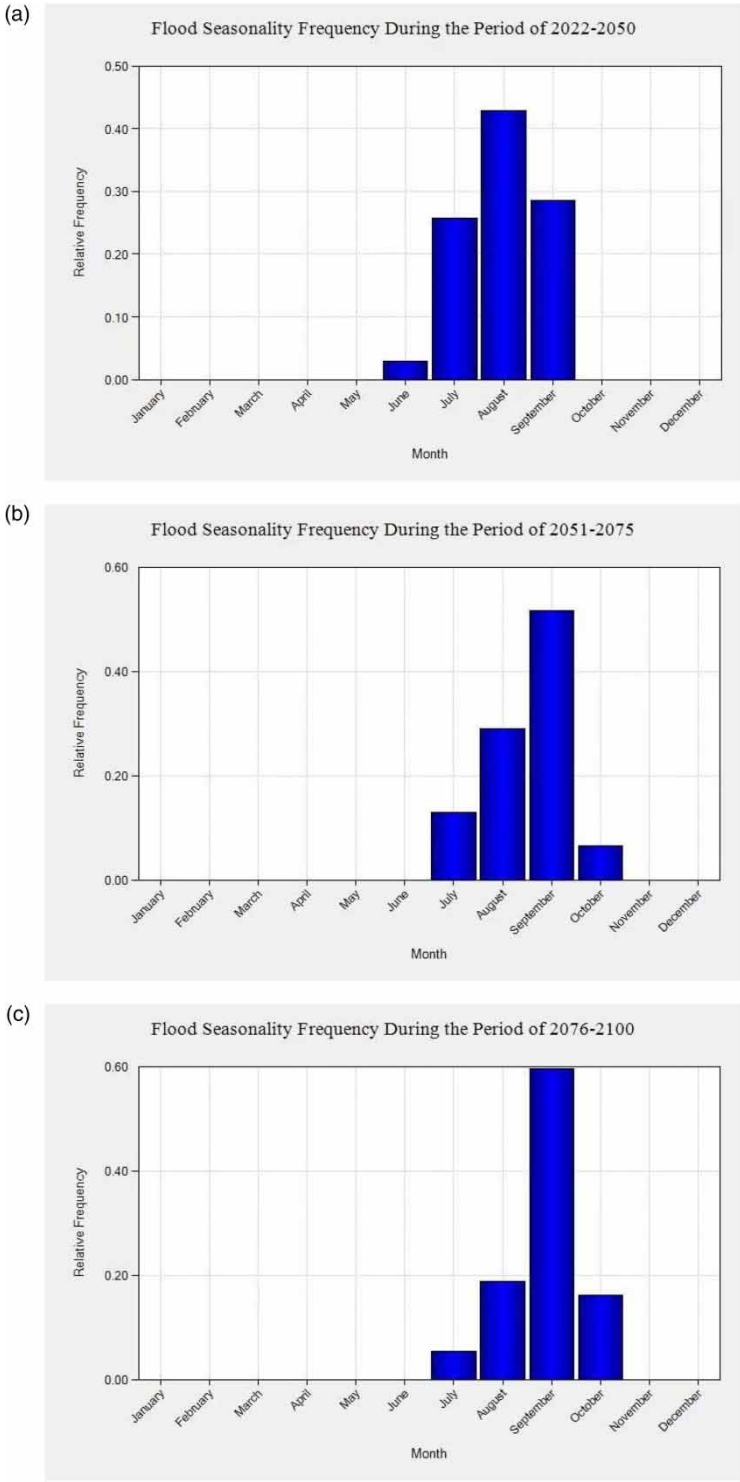


Figure 11 | Flood seasonality analysis during the period: (a) 2022–2050, (b) 2051–2075, and (c) 2076–2100.

During the period of 2051–2075, the expected peak discharge for 1/10,000 AEP is equal to 2,126.3 m^3/s. The peak discharge at 95% confidence upper and lower limits is 790 and 3,468 m^3/s, respectively. The corresponding peak stage for expected value and 95% lower and upper bounds values of 0.0001 APE is 934.18, 931.7, and 936.1, respectively. The PMF discharge of 6,180 m^3/s and maximum water surface elevation of 939.5 m are also not exceeded by those values.

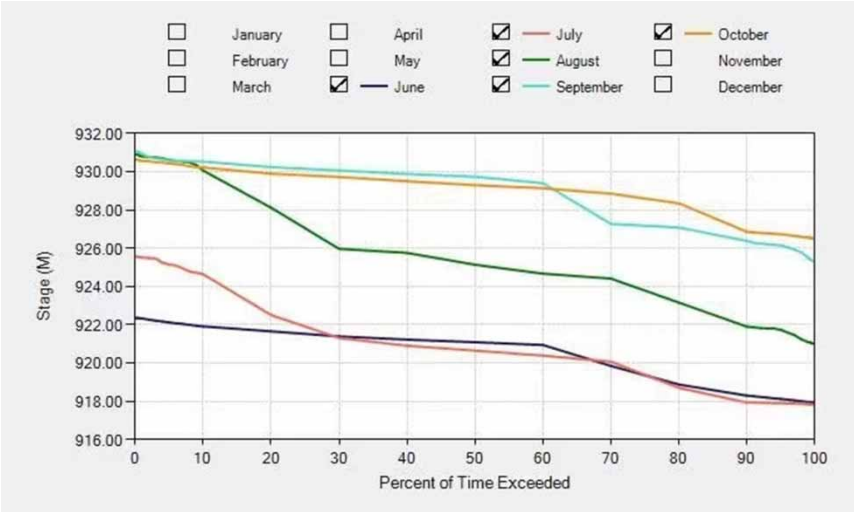


Figure 12 | Reservoir starting-stage duration analysis for the Kessem Dam during the flood season.

During the period of 2076–2100, the expected peak discharge for a 1/10,000 AEP was equal to 11,491.1 m^3/s. The 10,000-year peak discharge at 95% confidence upper and lower limits is 1,357 and 17,217 m^3/s, respectively, from the hydrological hazard analysis. The corresponding expected peak stage for 0.0001APE is 942.11 m, and the lower and upper 95% of bounds values are 932.8 and 943 m, respectively. It exceeds that from the PMF discharge of 6,180 m^3/s and the maximum water surface elevation of 939.5 m.

The results from this initial hydrologic hazard curve characterization and flood hydrograph (Figure 14) routing indicate that the Kessem Dam may potentially be overtopped by a flood with a return period of about 10,000 years during the period of 2076–2100. However, this indicates that the Kessem Dam does not meet reclamation hydrologic hazard criteria for overtopping because it does not pass through a PMF for 2076–2100 future time horizons. Therefore, the dam requires further risk analysis study and dam safety modification in order to control this probable failure mode during the period of 2076–2100.

Flood frequency estimation is currently being undertaken using models based on the basic concept of a stationary historical record. However, observed data represent the past, and these same data might no longer provide all the required information needed concerning the future statistical behavior of extreme precipitation and floods. Accurate forecasting of flood/peak flow is an essential requirement for reducing the risk of flooding and is important for planning and managing water resource systems. However, accurate forecasting of river discharge is a difficult problem because river flood stage analysis is a complex, dynamic process characterized by spatial and temporal variations. Besides, the river flow process is nonlinear and influenced by many factors such as river basin surface covering, the rainfall process, as well as riverbed terrain and climatic characteristics (Le *et al.* 2019). There are several findings and analyses of flood frequency and predictions of the magnitude of extreme events (Hamed & Rao 2000; Subramanya 2008). Some of the previous studies discuss the problem of estimating the relationship between the magnitude and frequency of unexpected floods (Bezak *et al.* 2014). Currently, there are two approaches for flow prediction (Le *et al.* 2019). The first method consists of mathematical models that simulate the hydrodynamic process of the water's flow. It is a widely used method because the mathematical models are based on concepts of hydraulics and hydrology. These models generally tend to require a large amount of input data (i.e., rainfall forecasts and topography data), which may not always be available or could be difficult to obtain. Also, the parameters of the model need to be tested and evaluated carefully as these parameters are regionally dependent, and sometimes it is difficult to estimate or calibrate the parameters that fit the model. As a result, models do not gain good performance, especially in areas where available data is limited. Moreover, process-based methods have limitations regarding flood warnings because the runtime of these models is usually quite long. The second approach for predicting river flow consists of data-driven methods, based on the statistical relationship between input and output data. One of the most common data-driven models is the machine learning model. The amount of data required for the models, as well as the ability of the models to handle both linear and nonlinear systems without the need to make any assumptions. The deep-learning models are more accurate predictors than

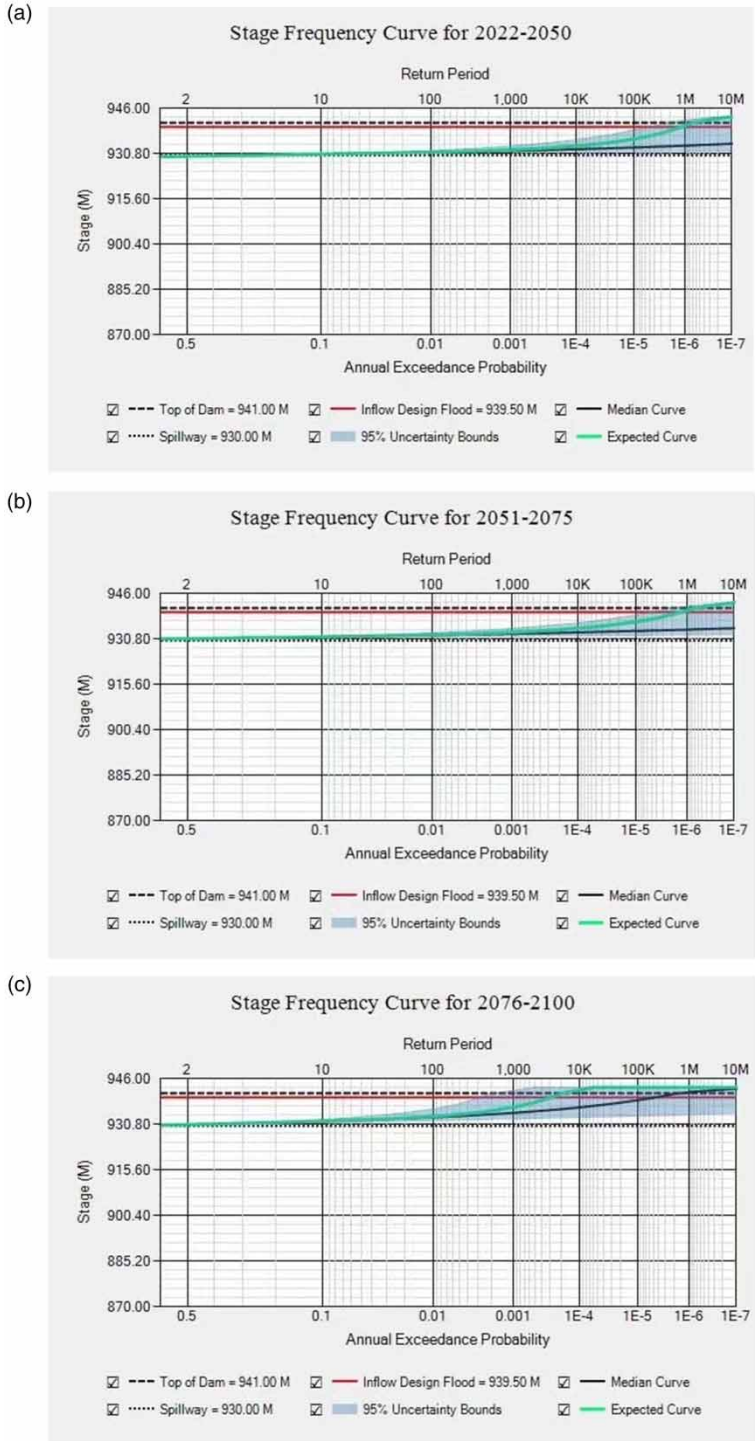


Figure 13 | Stage frequency curve during the period: (a) 2022–2050, (b) 2051–2075, and (c) 2076–2100.

the others. This advanced predictive model was used, in order to reduce the uncertainty during the predictive process (Mich 2022).

Furthermore, to mitigate the impact of this event on various assets, there should be an analysis and assessment of the relationship between flood magnitude, design period, and the probability of its occurrences, as well as the risk of failure of water projects (such as a dam) using different flood risk analysis and assessment methods, approaches, and techniques

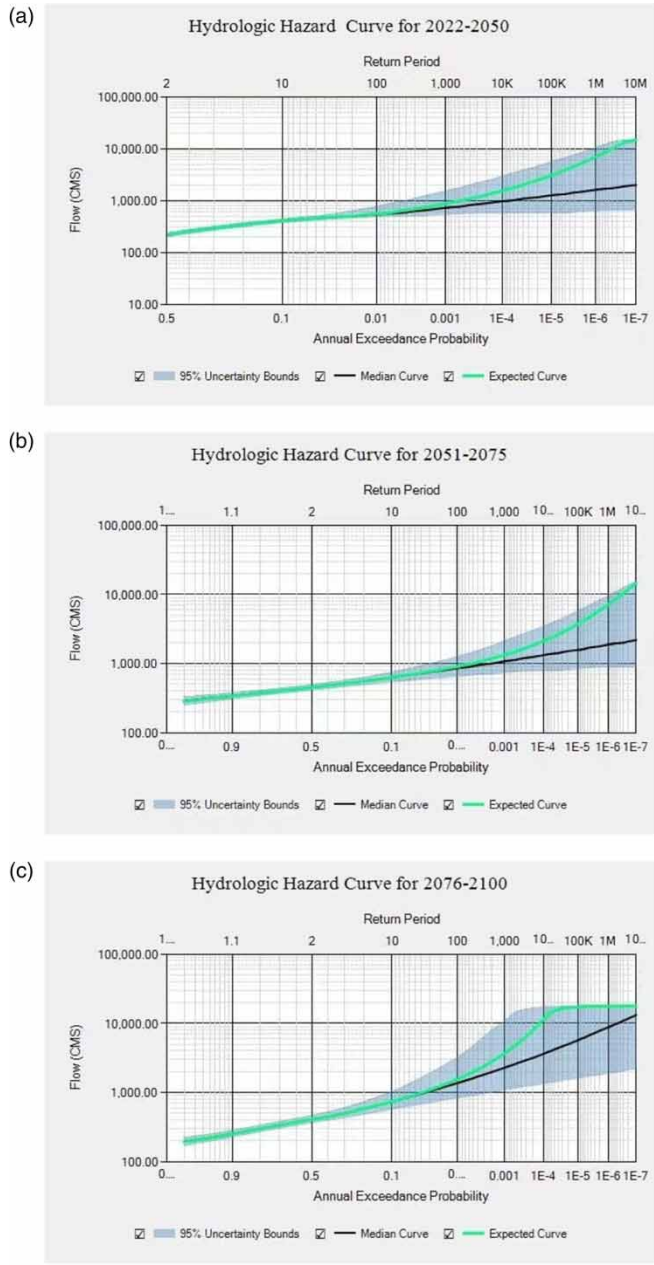


Figure 14 | Hydrologic hazard curve during the period: (a) 2022–2050, (b) 2051–2075, and (c) 2076–2100.

Table 8 | Expected probable peak discharge (m^3/s) for each future time horizons

	2022–2050		2051–2075		2076–2100	
Return period	**Expected**	**95% Bounds**	**Expected**	**95% Bounds**	**Expected**	**95% Bounds**
100	867.67	669.3–1,168	809.4	621–1,082	1,230.9	759–2,308
1,000	1,383.95	835–2,294.8	1,178.9	705–1,796	2,770.7	1,003–8,013
10,000	2,823.57	987–5,430.6	2,126.3	790–3,468	11,491.1	1,357–17,217
100,000	5,686.2	1,115–10,429.3	3,738.5	851–5,811	17,292.3	1,636–17,726
1,000,000	12,916.72	1,147–15,231.7	7,104.67	887–9,625	17,733.7	1,900–17,777

Table 9 | Expected probable peak stage (m) for each future time horizons

	2022–2050		2051–2075		2076–2100	
Return period	Expected	95% Bounds	Expected	95% Bounds	Expected	95% Bounds
100	931.92	931.4–932.6	931.77	931.2–932.4	932.7	931.6–934.5
1,000	932.94	931.7–934.5	932.54	931.5–933.7	935.13	932.2–940.9
10,000	935.21	932.1–938.4	934.18	931.7–936.1	942.11	932.8–943
100,000	938.59	932.2–941.8	936.37	931.8–938.7	943	933.3–943
1,000,000	942.42	932.4–943	940.1	931.9–941.6	943	933.7–943

that are used to check dam safety under these extreme conditions, and it helps for good decision-makers to take proper measures and for good flood management planning in the water (Hamed & Rao 2000).

4. CONCLUSION

The research predicts the peak flow/flood that may load the reservoir in the future using a machine learning model and develop the HHC for the dam to assess the hydrologic risk for dam safety evaluation of the Kessem Dam using the RMC-RFA software. Machine learning models were used to simulate future stream flow and peak inflow hybrids with the SCS-CN model using the future LULC scenario and the climate projection under the RCP emission scenario. Based on the novel and powerful ML predictive models' approach to stream flow simulation based on recurrent neural networks an inflow volume frequency analysis, the hydrological hazard was assessed using RMC-RFA software for three future time horizons (2022–2051, 2051–2075, and 2076–2100). Here, 70% of the flow data is trained on RNNs, including Bi-LSTM, LSTM, and GRU and the remaining 30% of the data were used for performance evaluation. For each of the networks, double hidden layer was used. The numbers of neurons in the hidden layers, the learning rate, and the number of iterations have a basic role in modeling accuracy, the optimal values of which were obtained using trial and error. Using the dropout function in the network structure prevents network over fitting. The efficiency of the proposed predictive approach is evaluated for the simulation of daily inflow to the Kessem Dam reservoir in the Kessem watershed. Results of the study point to the better performance of Bi-LSTM compared to the other RNN architectures. The Bi-LSTM network performed the best and can estimate stream flow with fairly good accuracy. These networks were used as an effective method of stream flow modeling and predict the peak flow. The performance of the Bi-LSTM model to simulate stream flow for the baseline period (1990–2013) was acceptable and it outperformed other proposed ML models. During the periods of 2022–2100 and 2051–2075, the peak discharge does not exceed the PMF and the overtopping does not occur during this period. The outcome of the semi-quantitative risk analysis showed that for all peak flood potential loading, the results indicate that the Kessem Dam may potentially be overtopped by a flood with a return period during the period of 2076–2100, since the peak discharge does not pass through a PMF for 2076–2100 future time horizons and does not meet reclamation hydrologic hazard criteria for overtopping. The dam required further risk analysis regarding its consequences and damage, and how to modify dam safety in order to control this probable failure mode, but the dam is safe within the period of 2022–2075. But all the critical upcoming peak flows pass through the dam spillway, which means that the existing spillway capacity is sufficient to surcharge the excess water without overtopping. This research contributes advanced scientific information regarding the future flood to the Kessem Dam reservoir and assesses its hazard on dam safety and risk assessment.

ACKNOWLEDGEMENTS

The authors would like to thank the Ethiopian Ministry of Water and Energy, Ethiopian National Meteorological Agency, ECWC and United States Geological Survey (USGS) for providing essential data for this study. The authors are also thankful to Arba Minch University who has provided all the logistic support in completing this research work.

FUNDING

No funding has been received from any source to conduct the research work.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Assefa, T. 2018 Flood risk assessment in Ethiopia. *Civ. Environ. Res.* **10**, 35–40.

Bartles, M., Brunner, G., Fleming, M., Faber, B., Karlovits, G. & Slaughter, J. 2019 *HEC-SSP Statistical Software Package User's Manual*. US Army Corps of Engineers Institute for Water Resources Engineering Center, Davis CA.

Bezak, N., Brilly, M. & Šraj, M. 2014 Comparaison entre les méthodes de dépassement de seuil et du maximum annuel pour les analyses de fréquence des crues. *Hydrol. Sci. J.* **59**, 959–977. https://doi.org/10.1080/02626667.2013.831174.

Gabriel-Martin, I., Sordo-Ward, A. & Garrote, L. 2019 Granados hydrological risk analysis of dams: The influence of initial reservoir level conditions. *Water* **11**, 1–17.

Gebeyehu, A. E., Chunju, Z. & Yihong, Z. 2019a Assessment and mapping of land use change by remote sensing and GIS: A case study of Abaya Chamo Sub-basin, Ethiopia. *Nat. Environ. Pollut. Technol.* **18**, 549–554.

Gebeyehu, A. E., Chunju, Z., Yihong, Z. & Yang, Z. 2019b Assessment of future temperature change scenario by statistical downscaling model (CANESM2) in Abaya Chamo sub basin, Ethiopia. *Int. J. Curr. Res.* **11**, 100–109. https://doi.org/10.24941/ijcr.33786.01.2019.

Getahun, Y. S. & Gebre, S. L. 2015 Flood hazard assessment and mapping of flood inundation area of the Awash River Basin in Ethiopia using GIS and HEC-GeoRAS/HEC-RAS model. *J. Civ. Environ. Eng.* **05**. https://doi.org/10.4172/2165-784x.1000179.

Hamed, K. & Rao, A. R. 2000 *Flood Frequency Analysis*, 1st edn. CRC Press, Boca Raton.

Hong, J., Lee, S., Bae, J. H., Lee, J., Park, W. J., Lee, D., Kim, J. & Lim, K. J. 2020 Development and evaluation of the combined machine learning models for the prediction of dam inflow. *Water (Switzerland)* **12**, 1–18. https://doi.org/10.3390/w12102927.

Hosseiny, H., Nazari, F., Smith, V. & Nataraj, C. 2020 A framework for modeling flood depth using a hybrid of hydraulics and machine learning. *Sci. Rep.* **10**, 1–14. https://doi.org/10.1038/s41598-020-65232-5.

Huang, I. H., Chang, M. J. & Lin, G. F. 2022 An optimal integration of multiple machine learning techniques to real-time reservoir inflow forecasting. *Stoch. Environ. Res. Risk Assess.* **36**, 1541–1561. https://doi.org/10.1007/s00477-021-02085-y.

Hunter, J. D. 2007 Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95. https://doi.org/10.1109/MCSE.2007.55.

Jackson, D. C. & Brown, J. G. 2023 Dam. *Encyclopaedia Britannica*. Available from: https://www.britannica.com/technology/dam-engineering (accessed 8 July 2023).

Jimeno-Sáez, P., Senent-Aparicio, J., Pérez-Sánchez, J., Pulido-Velazquez, D. & María Cecilia, J. 2017 Estimation of instantaneous peak flow using machine-learning models and empirical formula in Peninsular Spain. *Water (Switzerland)* **9**, 347. https://doi.org/10.3390/w9050347.

Khalaf, M., Hussain, A. J., Al-Jumeily, D., Baker, T., Keight, R., Lisboa, P., Fergus, P. & Kafri, A. S. A. 2018 A Data Science Methodology Based on Machine Learning Algorithms for Flood Severity Prediction. In: *2018 IEEE Congress on Evolutionary Computation (CEC)*, Rio de Janeiro, Brazil. https://doi.org/10.1109/CEC.2018.8477904

Kratzert, F., Klotz, D., Brenner, C., Schulz, K. & Herrnegger, M. 2018 Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* **22**, 6005–6022. https://doi.org/10.5194/hess-22-6005-2018.

Le, X. H., Ho, H. V., Lee, G. & Jung, S. 2019 Application of long short-term memory (LSTM) neural network for flood forecasting. *Water (Switzerland)* **11**, 1387. https://doi.org/10.3390/w11071387.

Mahmood, R., Pielke, R. A., Hubbard, K. G., Niyogi, D., Bonan, G., Lawrence, P., McNider, R., McAlpine, C., Etter, A., Gameda, S., Qian, B., Carleton, A., Beltran-Przekurat, A., Chase, T., Quintanar, A. I., Adegoke, J. O., Vezhapparambu, S., Conner, G., Asefi, S., Sertel, E., Legates, D. R., Wu, Y., Hale, R., Frauenfeld, O. W., Watts, A., Shepherd, M., Mitra, C., Anantharaj, V. G., Fall, S., Lund, R., Treviño, A., Blanken, P., Du, J., Chang, H. I., Leeper, R., Nair, U. S., Dobler, S., Deo, R. & Syktus, J. 2010 Impacts of land use/land cover change on climate and future research priorities. *Bull. Am. Meteorol. Soc.* **91**, 37–46. https://doi.org/10.1175/2009BAMS2769.1.

Mich, L. 2022 *Artificial Intelligence and Machine Learning*. Springer. https://doi.org/10.1007/978-3-030-48652-5_25.

Mosavi, A., Ozturk, P. & Chau, K. W. 2018 Flood prediction using machine learning models: Literature review. *Water (Switzerland)* **10**, 1–40. https://doi.org/10.3390/w10111536.

Roberts, W., Williams, G. P., Jackson, E., Nelson, E. J. & Ames, D. P. 2018 Hydrostats: A Python package for characterizing errors between observed and predicted time series. *Hydrology* **5**, 66. https://doi.org/10.3390/hydrology5040066.

Senent-Aparicio, J., Jimeno-Sáez, P., Bueno-Crespo, A., Pérez-Sánchez, J. & Pulido-Velázquez, D. 2019 Coupling machine-learning techniques with SWAT model for instantaneous peak flow prediction. *Biosyst. Eng.* **177**, 67–77. https://doi.org/10.1016/j.biosystemseng.2018.04.022.

Shoaib, M., Shamseldin, A. Y. & Melville, B. W. 2014 Comparative study of different wavelet based neural network models for rainfall-runoff modeling. *J. Hydrol.* **515**, 47–58. https://doi.org/10.1016/j.jhydrol.2014.04.055.

Shumie, M. C. 2019 Evaluation of potential reservoir deficiency due to climate change, Kesem Kebena dam, Ethiopia. *J. Environ. Geogr.* **12**, 33–40.

Sivapalan, M., Blöschl, G., Merz, R. & Gutknecht, D. 2005 Linking flood frequency to long-term water balance: Incorporating effects of seasonality. *Water Resour. Res.* **41**, 1–17. https://doi.org/10.1029/2004WR003439.

Stehman, S. 1996 Estimating the kappa coefficient and its variance under stratified random sampling. *Photogramm. Eng. Remote Sensing* **62**, 401–407.

Subramanya, K. 2008 *Engineering Hydrology*, 3rd edn. Tata McGraw-Hill Publishing, New Delhi.

Tian, D., He, X., Srivastava, P. & Kalin, L. 2022 A hybrid framework for forecasting monthly reservoir inflow based on machine learning techniques with dynamic climate forecasts, satellite-based data, and climate phenomenon information. *Stoch. Environ. Res. Risk Assess.* **36**, 2353–2375. https://doi.org/10.1007/s00477-021-02023-y.

Vo, N. D., Gourbesville, P., Vu, M. T., Raghavan, S. V. & Liong, S. Y. 2016 A deterministic hydrological approach to estimate climate change impact on river flow: Vu Gia-Thu Bon catchment, Vietnam. *J. Hydro-Environ. Res.* **11**, 59–74. https://doi.org/10.1016/j.jher.2015.11.001.

Wilby, R. L. & Dawson, C. W. 2007 SDSM 4.2– A decision support tool for the assessment of regional climate change impacts, User Manual. Department of Geography, Lancaster University, UK, pp. 1–94.

Zhang, Q., Zhang, L., She, D., Wang, S., Wang, G. & Zeng, S. 2021 Automatic procedure for selecting flood events and identifying flood characteristics from daily streamflow data. *Environ. Model. Softw.* **145**, 105180. https://doi.org/10.1016/j.envsoft.2021.105180.

First received 21 May 2023; accepted in revised form 22 September 2023. Available online 18 January 2024