Machine learning and HEC-RAS integrated models for flood inundation mapping in Baro River Basin, Ethiopia

Habtamu Tamiru*

Department of Water Resources and Irrigation Engineering, Wollega University, P.O. Box 395, Ethiopia

Abstract

This paper presents the integrated machine learning and HEC-RAS models for flood inundation mapping in Baro River Basin, Ethiopia. A predictive rainfall-runoff and spatially distributed river simulation models were developed using Artificial Neural Networks (ANNs) and HEC-RAS respectively. Daily rainfall and temperature data of 7-yrs and Topographical Wetness Index (TWI) with a spatial resolution of 50 x 50m were used to train the ANN in R studio. The integration of the spatial and temporal variability in this paper improved the accuracy of the predictive models integrated with ANN and HEC-RAS. The predictive ANN model was tested with the observed daily discharge of the same temporal resolution and the rainfall-runoff result obtained from the tested ANN model was used as input for the HEC-RAS. The flood event of 2005 was used to verify the accuracy of flood generated in the HEC-RAS model by implementing the Normal Difference Water Index (NDWI). The comparison was made between the flood inundation map generated by HEC-RAS and flood events of different periods based on coverage percentage areas and a good agreement was reached with 96 % overlapped areas. The performance of ANN and HEC-RAS models were evaluated with 0.86 and 0.88 values at the training and testing period respectively. Finally, it was concluded that the integration of a machine learning approach with the HEC-RAS model in developing a flood inundation mapping is an appropriate tool to warn residents in this river basin.

Keywords **|** ANNs, Flood inundation, HEC-RAS, machine learning, NDWI, TWI

1 Introduction

The severity of flooding is very high in a region where there is no sufficient structural affordability due to financial limitations. Flood risk is increasing due to the growth in people and property living in flood plains[1], and the possible increase in flood hazard associated with climate change. The accuracy in flood modeling highly relies on the quality of data and the appropriateness of the method implemented. Space and time uncertainties in flood inundation are common in flood risk management due to the different factors. The improvement in the capability of models to get more accurate results in space and time is dramatically increasing day-to-day.

A neural network is a machine learning algorithm that focuses on an information processing technique based on the model of a human neuron[2]. The application of data-driven models such as Artificial Neural Networks (ANNs) for hydrological modeling reveals a timely accurate result to warn river beach residentials[3]. The integration of different models in the areas of hydrologic and hydraulic models is getting global attention and has a paramount role in flood risk management strategies [4];[5]. Flood inundation mapping is a difficult task that needs a combination of high quality and observed data to verify the performance of the models [6];[7]. The modification of space and time uncertainties[8] is the novelty of this paper. The application of machine learning is which is explored by integrating a machine learning and two-dimensional analysis in HEC-RAS. The application of machine learning (ANN) in areas of hydrologic processes is a current[9] and has been applied in rainfall-runoff modeling[10], daily water supply demand[11], streamflow computation[12],[13], extreme hydrologic event analysis[14] and generation of unit hydrograph[15]. A feedforward ANN network structure (Fig.1) is commonly used in the hydrologic process is the one-way computation in which inputs are pushed forward until the rough result is obtained[16].

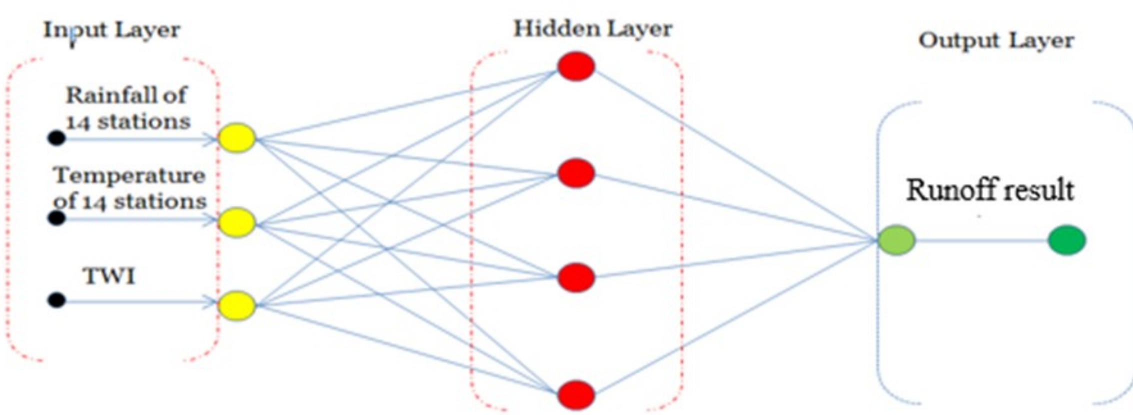


Figure 1: ANN feedforward structure

2 Method and Data

2.1 Study area

The Baro Akobo basin is located in the southwestern part of Ethiopia. It covers approximately an area of 74,100km^2. The River basin (Fig.2) is the fourth largest basin in the country. The western, northwestern, and southwestern side of the basin is bordered with Sudan, while in the northern and northeast it is bordered by the Abay river basin, the east and southeast it is bordered by the Omo-Gibe river basin. Geographically, it is located, between latitudes 5^0 31" and $10^0$54" north and longitude 33^0 and $36^0$17" east. The River originates from the highlands in the southwest part of Ethiopia and flows across the low-lying plains. The most recent flood period in 2015 in the river basin forced around 2,000 peoples out of their homes[17].

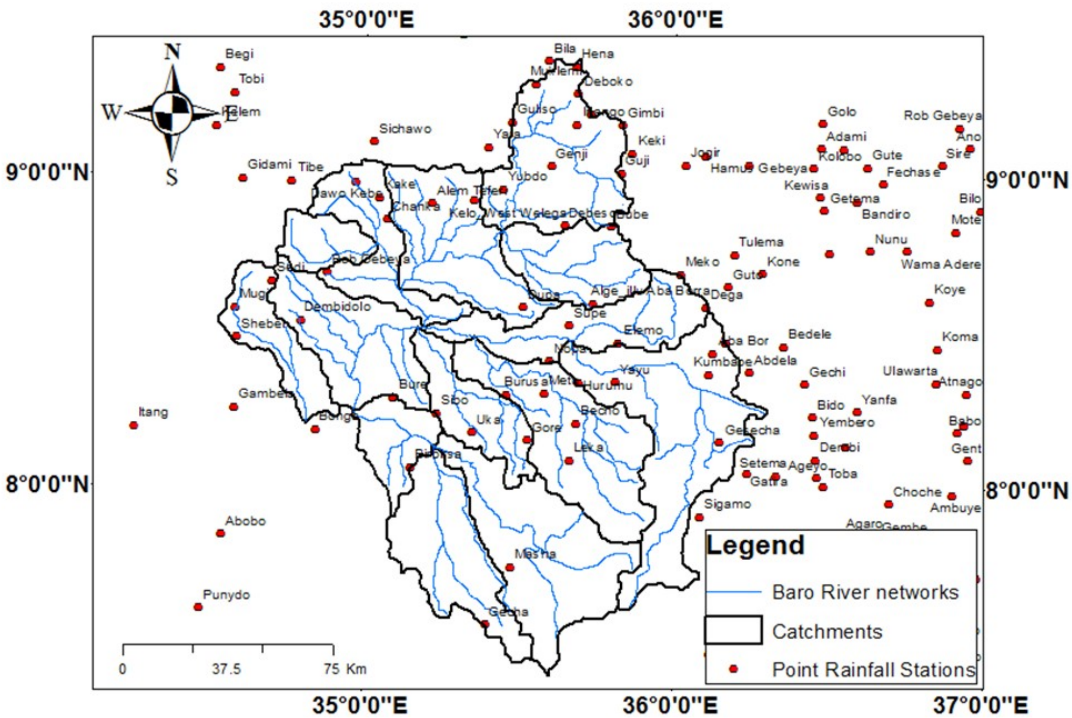


Figure 2: Map of the study area

2.2 Data and software used

In this study, ArcGIS 10.4, R-studio, and HEC-RAS 5.0.1, were used to generate an inundation map, develop ANN hydrological model, simulate a hydraulic model respectively. All packages are supported by student license and open-source privileges. For ANN hydrological model training, both temporal (7-years daily Rainfall and Temperature) of point data (Table) and spatial

(Topographical wetness Index) data were used. The spatial resolution of 50 x 50 m pixel size was implemented and all input parameters were prepared based on the fixed grid.

Table 1: Point rainfall and temperature stations in the basin

Station name	Longitude (deg)	Latitude (deg)	Elevation(m)	Year	Annual Rainfall (mm)
Abdela	36.25	8.37	1859.90	1999-2005	2009.68
Alge	35.74	8.59	1807.02	1999-2005	1828.75
Bila	35.59	9.37	1911.82	1999-2005	1945.74
Bonga	34.85	8.18	519.23	1999-2005	1186.85
Bure	35.10	8.28	1600.60	1999-2005	1706.10
Dusta	36.18	7.75	2328.71	1999-2005	1936.07
Gambela	34.59	8.25	517.49	1999-2005	1095.85
Gatira	36.24	8.05	2203.02	1999-2005	2221.12
Gecha	35.40	7.56	2203.40	1999-2005	2091.86
Gimbi	35.83	9.16	1940.37	1999-2005	1897.69
Gore	35.53	8.15	1802.98	1999-2005	2080.38
Guliso	35.48	9.17	1606.56	1999-2005	1645.91
Metu	35.59	8.30	1736.55	1999-2005	1832.52
RobGebya	34.88	8.69	1791.29	1999-2005	1652.00

2.3 Artificial Neural Networks (ANNs) and HEC-RAS Models

2.3.1 Artificial Neural Networks (ANNs) Model

Once the inputs reached the hidden nodes, then the activation function usually (sigmoid function) is applied to convert the inputs into normalized values to give the rough result in the output layer. The feedforward porocess is ended when the first rough result is obtained in the output layer. The backpropagation (the process from the output node to the hidden nodes) will take place to minimize the error in the rough result and this process continues until the network learns from mistakes[18][12]. This backpropagation is aimed to adjust the weights in each layer and minimize the error by propagating back into the networks until the difference between targeted and simulated results is optimized.

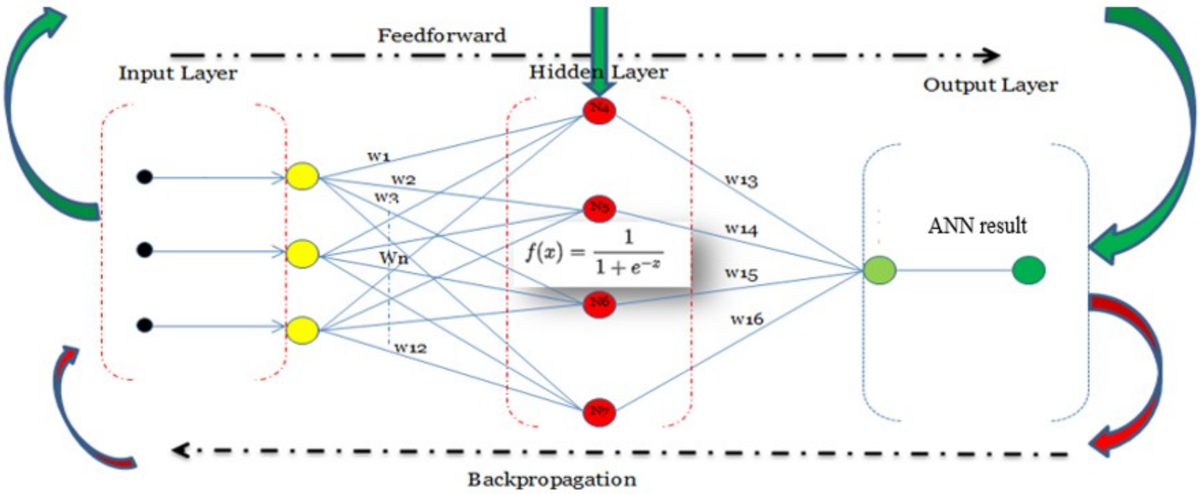


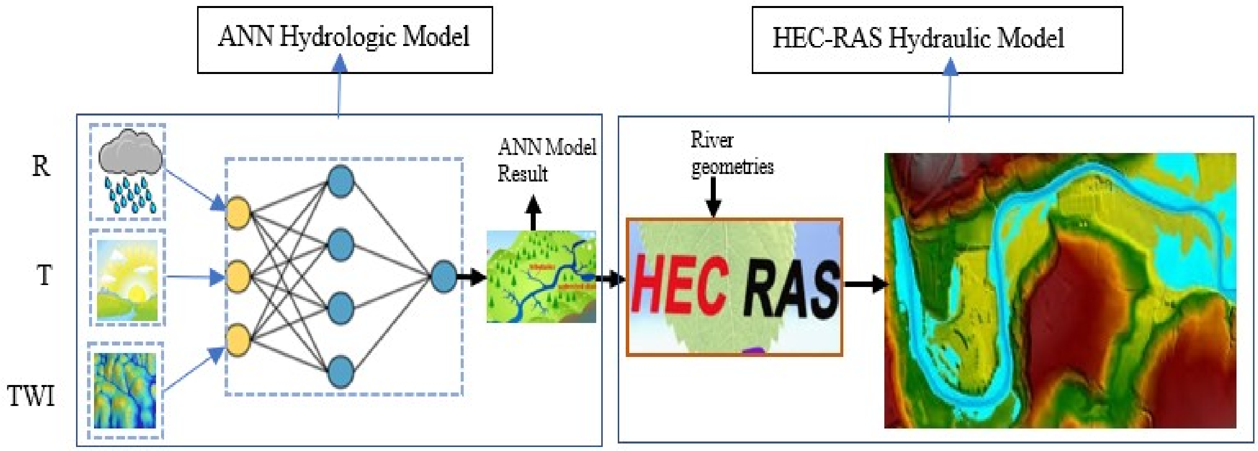
Figure 3: Back propagation process in ANN hydrological model

2.3.2 HEC-RAS Models

A flood inundation map provides information on the spatially distributed depth of flooding and prone areas[19]. A coupled 1D and 2D model has capable of generating the depth and prone or flood extent of a given area[20]. HEC-RAS model receives the result from the hydrological model (ANN model result) and gives the information on the spatial extent and depth of flooding along a river.

2.4 Integrated ANN and HEC-RAS Models

The hydrologic and hydraulic models are integrated in such a way that the tested values of runoff results from the ANN model are linked to the HEC-RAS model. The river geometries set up in the 1D analysis are associated with 2D river simulation to identify the flood spatial extent. The temporal (Rainfall and Temperature) and spatial (Topographical wetness Index) basin characteristics are linked together to get a better result of flooding.



R = Rainfall, T= Temperature, TWI = Topographical Wetness Index

Figure 4: A conceptual framework of ANN and HEC-RAS models

3 RESULTS AND DISCUSSIONS

3.1 ANN Hydrological model result

A Feedforward network is selected in this study. For this class of ANN architecture, input data at input nodes propagate in one direction to the output layer and a rough result is reached. The input data used were Rainfall (R), Temperate (T), and Topographical wetness Index (TWI), and these data were converted into gridded (50 x 50m) datasets to reduce the lumped characteristics of the basin. Inverse Distance Weights (IDW) in ArcGIS was implemented to generate initial weights based on the point rainfall and altitude of the stations. The first training process begun with the generated weight values.

$$Weights = \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix} = \begin{bmatrix} 0.17 & 0.15 & 0.31 & 0.09 \\ 0.68 & 0.40 & 0.07 & 0.80 \\ 0.67 & 0.11 & 0.31 & 0.95 \\ 0.22 & 0.99 & 0.14 & 0.83 \end{bmatrix}$$

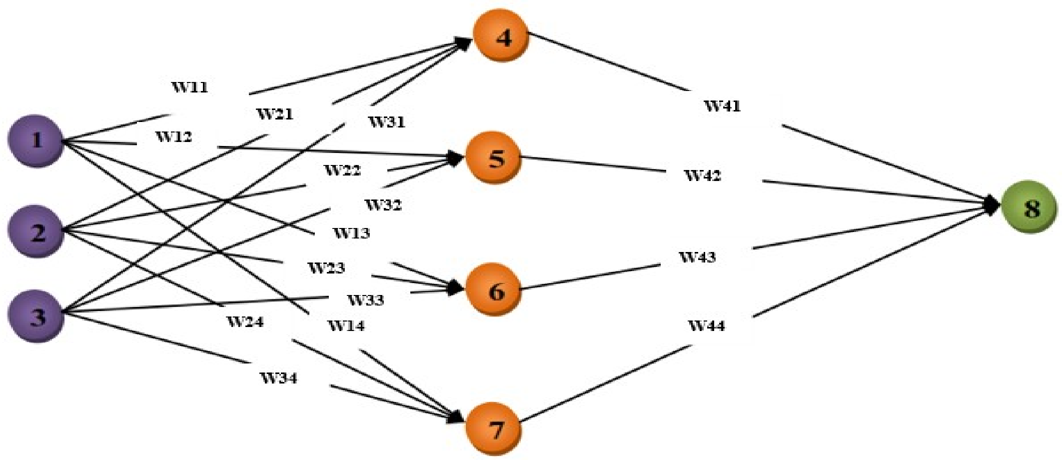


Figure 5: Assigned initial weights in the ANN networks

The ANN predictive hydrologic model was trained with rainfall (R), Temperature (T), and topographical wetness Index (TWI). The initial weight values were used to start the training processes, then the updated values were re-entered into the networks after the errors were propagated back into the process. The feedforward process undergoes the training process with the updated weight values and this will continue until the difference between the model result and observed value is minimized (Fig.6). Model performance was checked by implementing both graphical and statistical (NSE) evaluation techniques and evaluated at 0.86 and 0.88 values for training and testing periods[21], [19][8][22]. The ANN and observed hydrograph (Fig.6) developed in R studio with an acceptable training process and performance evaluation revealed that the application of ANN in the hydrological modeling is better than any other methods for the basin where flooding is damaging human lives and properties due to an appropriate method selection. The updated values obtained after training is reached, are used as fixed values in ANN model testing of 2006-2008 periods. The observed daily discharge data of 2006 – 2008 was used to validate the result and a good result (Fig. 7) was obtained.

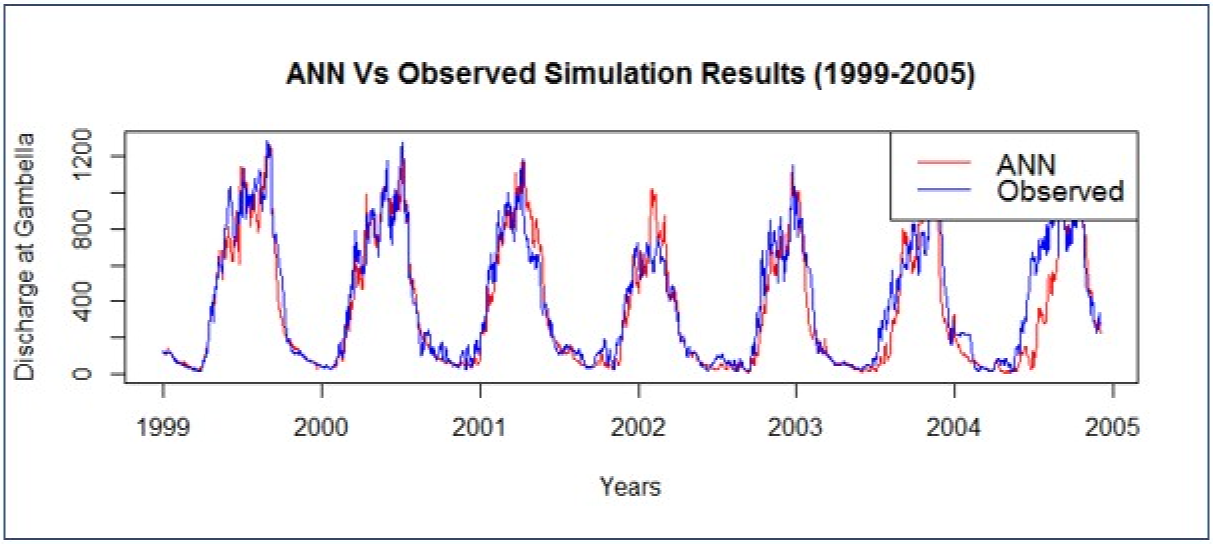


Figure 6: ANN hydrologic trained model result (1999-2005)

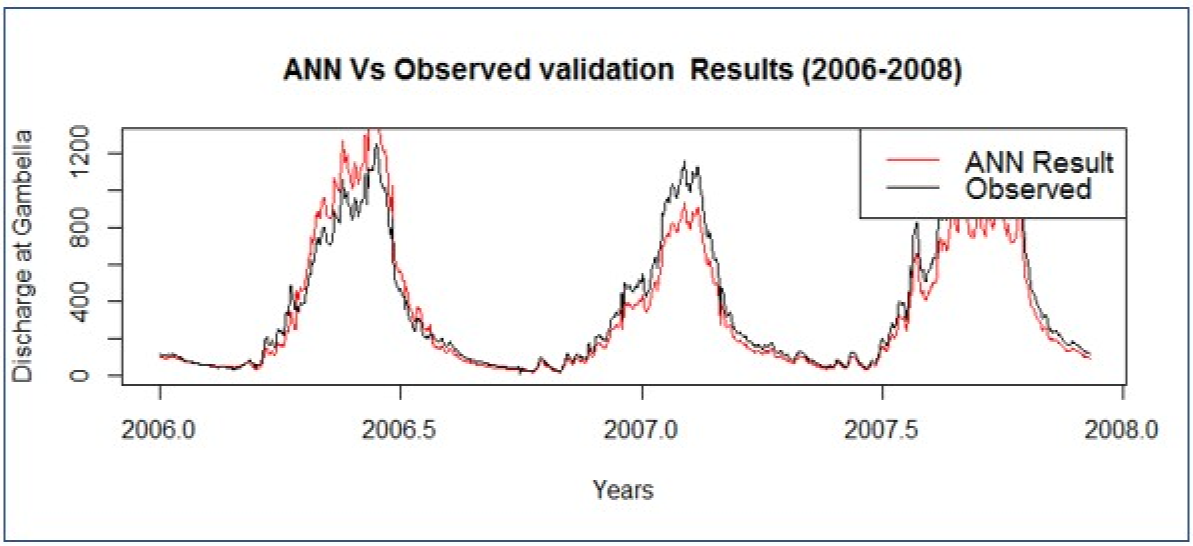


Figure 7: ANN hydrologic tested model result (2006-2008)

3.2 HEC-RAS Model result

Digital Elevation Models (DEM) of 12.5 x 12.5 m pixel resolution was used to generate 1D river geometries such as centerlines, bank lines, flow paths, and cross-sectional lines, and a flood-prone area in the 2D analysis (Fig.8). RAS Mapper window of the HEC-RAS 5.0.1 version was implemented to link the 1D and 2D simulation processes and exported to ArcGIS 10.4 version to overlay administrative and flood inundation maps together.

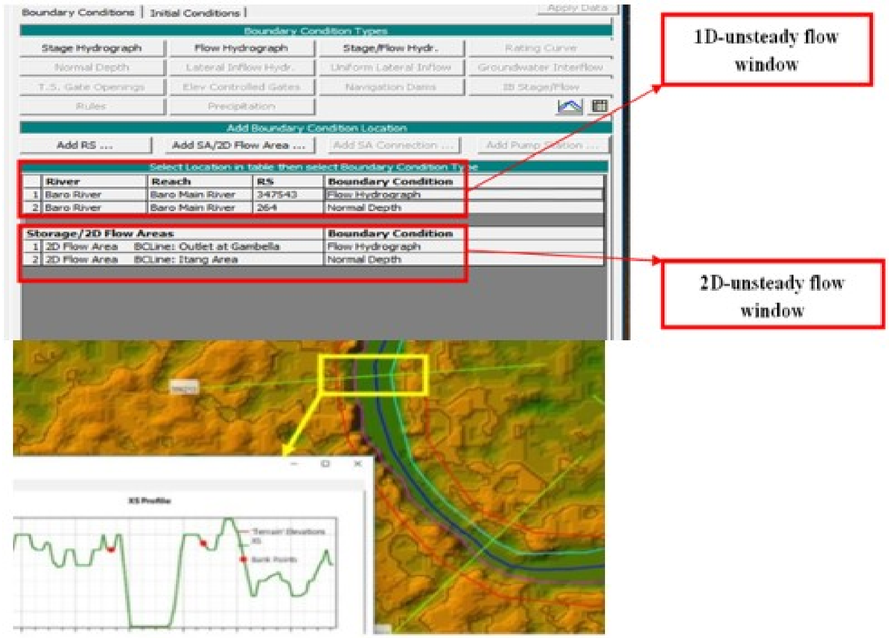


Figure 8: Baro river geometries (1D and 2D)

The developed ANN model in R was trained and tested before using it as input in HEC-RAS. The tested ANN model result (runoff values) was used as input in the hydraulic model (HEC-RAS) and routed in the model to generate the flood depths. Two areas were identified based on the variation depths (one varies from 0.00-0.82 cm and the other varies from 0.00-0.51cm) of flooding along the river (Fig.9).

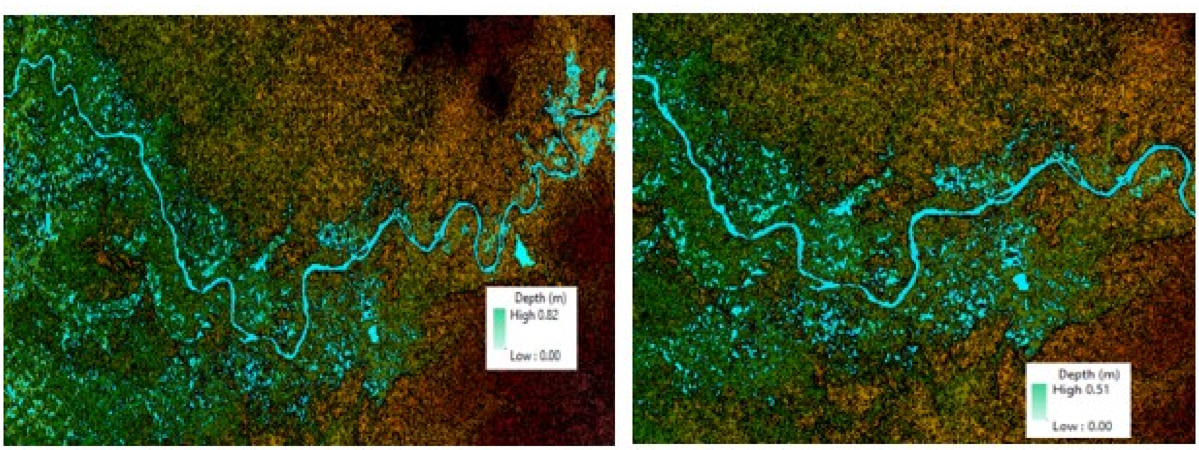


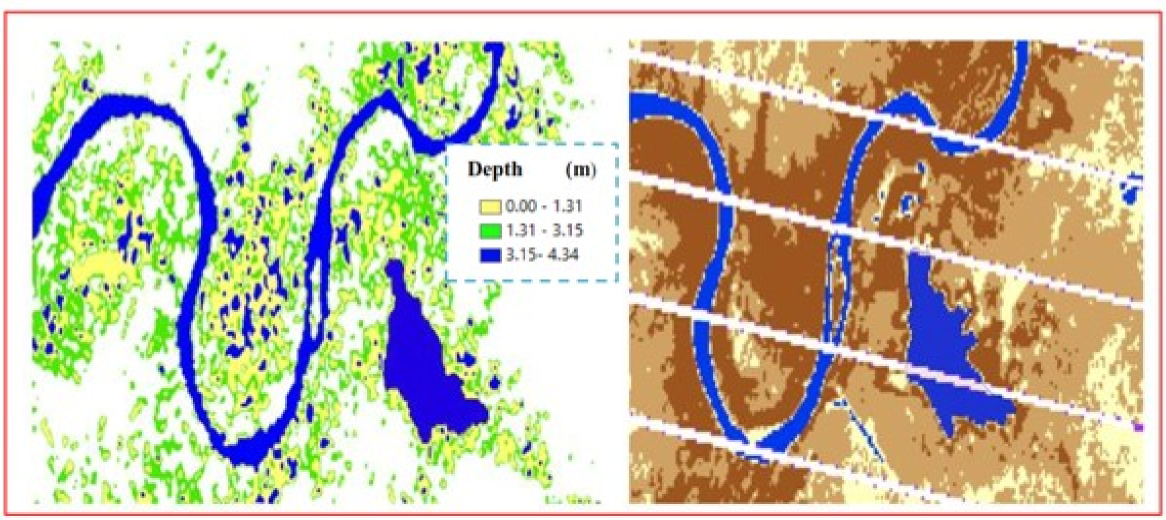
Figure 9: Inundation map along Baro River

3.3 Training and testing HEC-RAS model

The flood events of 2005 and 2008 were extracted using Normalized Difference Water Index (NDWI) from remotely sensed LANDSAT images. NDWI uses Green (Band-2) and near infra-red (Band-4) bands of remote sensing images to extract a water body in which near-infrared (NIR) and short-wave infrared (SWI) are used as the main input.

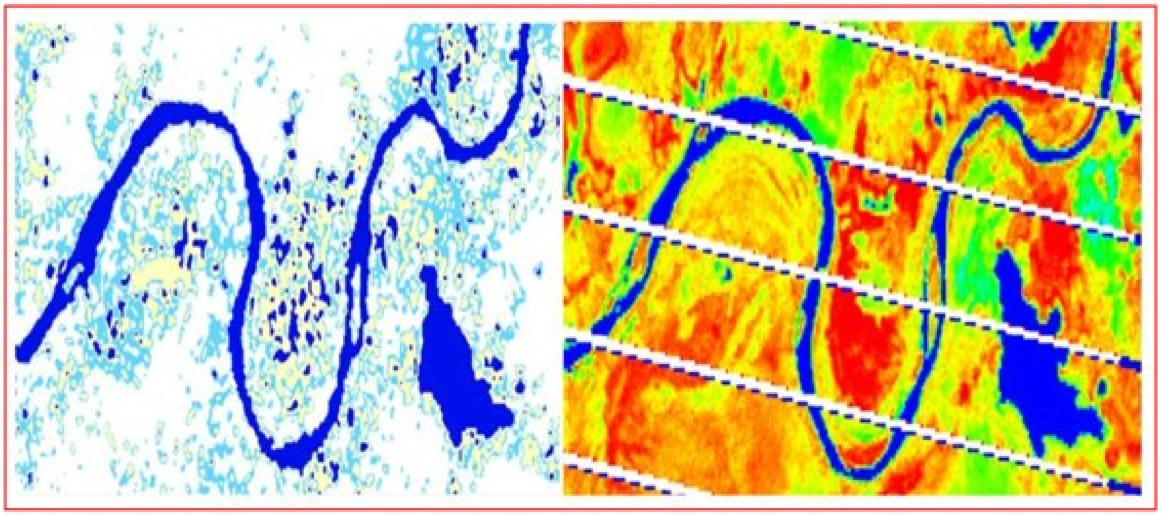
$$NDWI = \frac{NIR - SWI}{NIR + SWIR}$$

The area comparison between the NDWI and inundation maps as made and the overlapping percentage areas are crossed-checked in ArcGIS for training and testing periods. The 2005 flood event was used to train the inundation map meaning the ANN hydrologic model result of 1999-2005 was frequently run until an acceptable agreement is made between the inundation map and extracted NDWI map (Fig.10). Accordingly, the inundation map of 2006-2008 generated from the ANN hydrologic model result was validated using the 2008 flood event extracted in NDWI (Fig.11). The results in both periods were amazing that almost the area percentage revealed that the generated inundation map is accurate and about 96 % of areas are overlapped.



a) Inundation map generated in HEC-RAS (1999-2005) b) Flood event extracted in NDWI (2005)

Figure 10: Trained (calibrated) inundation map result in HEC-RAS (1999-2005)



a) Inundation map generated in HEC-RAS (200-2008) b) Flood event extracted in NDWI (2008)

Figure 11: Tested (validated) inundation map result in HEC-RAS (2006-2008)

4 CONCLUSION

In this study, machine learning and HEC-RAS models were integrated for flood inundation mapping in Baro River Basin, Ethiopia. A spatially distributed predictive rainfall-runoff model and river simulation models were developed using Artificial Neural Networks (ANNs) and HEC-RAS respectively. An acceptable temporal (five years daily rainfall and temperature data of 1999-2005) and spatial resolution of 50 x 50 m were used and the input parameters were prepared based on these resolutions. Daily rainfall (R) and temperature (T) of fourteen (14) point stations and topographical wetness Index (TWI) were used as input and converted into areal using ArcGIS version 10.4 with the same spatial resolution. The ANN hydrologic model was trained with the prepared temporal and spatial normalized data. Initial weights for the feedforward ANN networks were generated using Inverse Distance weighted (IDW) IDW by taking into account the point rainfall and temperature data and the corresponding altitudes of the stations due to the spatial variability of climate and topographic data. The integration of the spatial and temporal variability in this paper improved the accuracy of the predictive models which was not reached in the previous researches. The predictive ANN model was tested with the observed daily discharge of the same temporal resolution and the rainfall-runoff result obtained from the tested ANN model was used as input for the HEC-RAS. The flood event of

2005 was used to verify the accuracy of flood generated in the HEC-RAS model by implementing the Normal Difference Water Index (NDWI). The comparison of the prone areas generated in HEC-RAS and flood events delineated in NDWI made a good agreement with 96% overlapped areas. The performance of ANN and HEC-RAS models were evaluated with 0.86 and 0.88 values at the training and testing period respectively. Finally, it was concluded that the integration of a machine learning approach with the HEC-RAS model in developing a flood inundation mapping is an appropriate tool to warn residents in this river basin.

5 REFERENCES

[1] S. W. Abaya, N. Mandere, and G. Ewald, "Floods and health in Gambella region, Ethiopia: A qualitative assessment of the strengths and weaknesses of coping mechanisms," *Glob. Health Action*, vol. 2, no. 1, pp. 1–10, 2009, doi: 10.3402/gha.v2i0.2019.

[2] I. N. Daliakopoulos and I. K. Tsanis, "Comparison of an artificial neural network and a conceptual rainfall – runoff model in the simulation of ephemeral streamflow," *Hydrol. Sci. J.*, vol. 61, no. 15, pp. 2763–2774, 2016, doi: 10.1080/02626667.2016.1154151.

[3] S. H. Elsafi, "Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan," *Alexandria Eng. J.*, vol. 53, no. 3, pp. 655–662, 2014, doi: 10.1016/j.aej.2014.06.010.

[4] F. J. Chang, Y. M. Chiang, and L. C. Chang, "Multi-step-ahead neural networks for flood forecasting," *Hydrol. Sci. J.*, vol. 52, no. 1, pp. 114–130, 2007, doi: 10.1623/hysj.52.1.114.

[5] S. W. Abaya, "Floods and Health in Gambella region , Ethiopia : An Assessment of the strength and weakness of the coping mechanism," *Lund Univ. Cent. Sustain. Stud.*, pp. 1–44, 2008.

[6] M. S. Jami, N. A. Kabashi, I. A. F. Husain, and N. Abdullah, "Improvement of artificial neural network model for the prediction of wastewater treatment plant performance," *Proc. IASTED Int. Conf. Environ. Manag. Eng. EME 2011*, no. August, pp. 67–74, 2011, doi: 10.2316/P.2011.736-050.

[7] A. K. Lohani, R. Kumar, and R. D. Singh, "Hydrological time series modeling: A comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques," *J. Hydrol.*, vol. 442–443, no. January, pp. 23–35, 2012, doi: 10.1016/j.jhydrol.2012.03.031.

[8] T. A. Dessalegn, M. A. Moges, D. C. Dagnew, and A. Gashaw, "Applicability of Galway River Flow Forecasting and Modeling System (GFFMS) for Lake Tana Basin, Ethiopia,"

J. Water Resour. Prot., vol. 09, no. 12, pp. 1319–1334, 2017, doi: 10.4236/jwarp.2017.912084.

[9] C. Engineering, “Review on Rainfall Forecasting Using,” pp. 2901–2902, 2016, doi: 10.15680/IJIRCCE.2016.

[10] S. Riad, J. Mania, L. Bouchaou, and Y. Najjar, “Rainfall-runoff model usingan artificial neural network approach,” *Math. Comput. Model.*, vol. 40, no. 7–8, pp. 839–846, 2004, doi: 10.1016/j.mcm.2004.10.012.

[11] M. K. Akhtar, G. A. Corzo, S. J. Van Andel, and A. Jonoski, “River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information : case study of the Ganges river basin,” pp. 1607–1618, 2009.

[12] J. Veintimilla-Reyes, F. Cisneros, and P. Vanegas, “Artificial Neural Networks Applied to Flow Prediction: A Use Case for the Tomebamba River,” *Procedia Eng.*, vol. 162, pp. 153–161, 2016, doi: 10.1016/j.proeng.2016.11.031.

[13] V. Poonia, “Hydrological Analysis by Artificial Neural Network : A Review,” *Int. J. Adv. Res. Ideas Innov. Technol.*, vol. 4, no. 3, pp. 265–270, 2018.

[14] F. I. Evaluation, “Short-Term Load Forecasting Using EMD-LSTM,” 2017, doi: 10.3390/ en10081168.

[15] D. T. Mengistu, S. A. Moges, and A. Sorteberg, “Revisiting Systems Type Black-Box Rainfall-Runoff Models for Flow Forecasting Application,” no. January, pp. 65–83, 2016.

[16] M. GOSWAMI and K. M. O’C, “Application of Artificial Neural Networks for river flow simulation in three French\rcatchments,” *fourth Inter-Cletic Colloquim Hydrol. Manag. Water Resour.*, no. July 2005, p. 14, 2005.

[17] A. T. Haile, M. Dessalegn, I. Researcher, E. Alemu, and S. W. Corporation, “Household Adaptation to Flood in Itang District of Gambella Region in,” no. January, 2014.

[18] L. A. Dar, "RAINFALL-RUNOFF MODELING USING ARTIFICIAL NEURAL NETWORK RAINFALL-RUNOFF MODELING USING ARTIFICIAL NEURAL NETWORK TECHNIQUE," no. December, 2017.

[19] P. K. Parhi, "HEC-RAS Model for Mannnig's Roughness: A Case Study," *Open J. Mod. Hydrol.*, vol. 03, no. 03, pp. 97–101, 2013, doi: 10.4236/ojmh.2013.33013.

[20] A. Enea *et al.*, "Remote sensing , gis and hec-ras techniques , applied for flood extent validation , based on landsat imagery , lidar and hydrological data . Case study : Baseu river , Romania LANDSAT IMAGERY , LIDAR AND HYDROLOGICAL DATA .," no. November, 2018.

[21] C. Massmann, "Identification of factors influencing hydrologic model performance using a top-down approach in a large number of U.S. catchments," *Hydrol. Process.*, vol. 34, no. 1, pp. 4–20, 2020, doi: 10.1002/hyp.13566.

[22] F. Karamage *et al.*, "Modeling rainfall-runoffresponse to land use and land cover change in Rwanda (1990-2016)," *Water (Switzerland)*, vol. 9, no. 2, 2017, doi: 10.3390/w9020147.