Assessment of predictability in Downscaling GEFS Precipitation Forecasts

Smit Chetan Doshi[1,1,1] and Tirthankar Roy[2,2,2]

[1]University of Nebraska - Lincoln
[2]University of Nebraska-Lincoln

November 30, 2022

Abstract

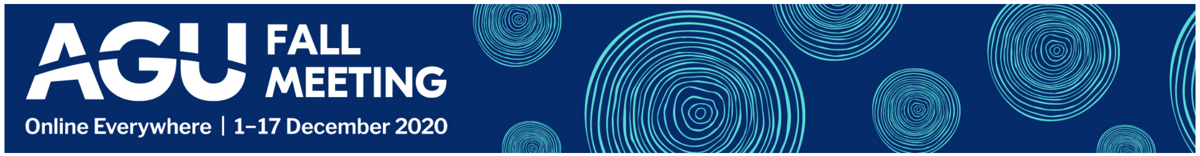
The NOAA Physical Sciences Laboratory produces the Global Ensemble Forecasting System (GEFS) which comprises 11 ensemble members (1 control and 10 perturbation runs) for over a 36-year period (December 1984 to present), with forecasts initialized every day for the next 16 days (first 8-day forecasts obtained from a high-resolution grid and the next 8-day forecasts from a low-resolution grid). The system provides 36 variables related to a wide range of hydrometeorological processes. In this study, we assess the predictability of precipitation within the context of statistical downscaling using a minimum set of predictor variables (precipitation and temperature). We use feedforward backpropagation neural networks with a suite of training algorithms to determine which variables (features) are of most relevance at different forecast lead times. The outcome of this study will significantly benefit short-term flood forecasting using GEFS data.

Assessment of Predictability in Downscaling GEFS Precipitation Forecasts

Smit Chetan Doshi , Tirthankar Roy

University of Nebraska - Lincoln, Nebraska, United States

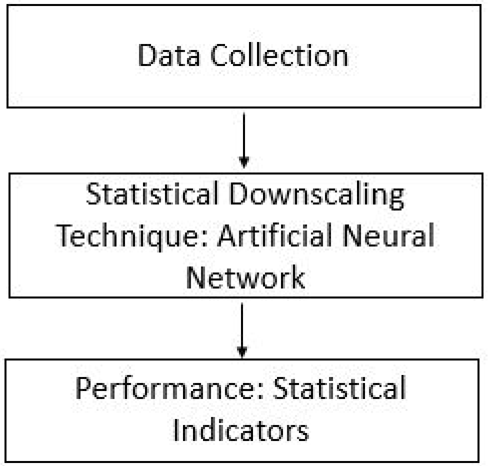
PRESENTED AT:



OVERVIEW

- In this study, we carried out the statistical downscaling of GEFS forecasts.
- The forecasts skill was assessed across a wide range of lead times.
- The point-scale rain gauge measurements were used as the targets to match.

Figure 1: Framework

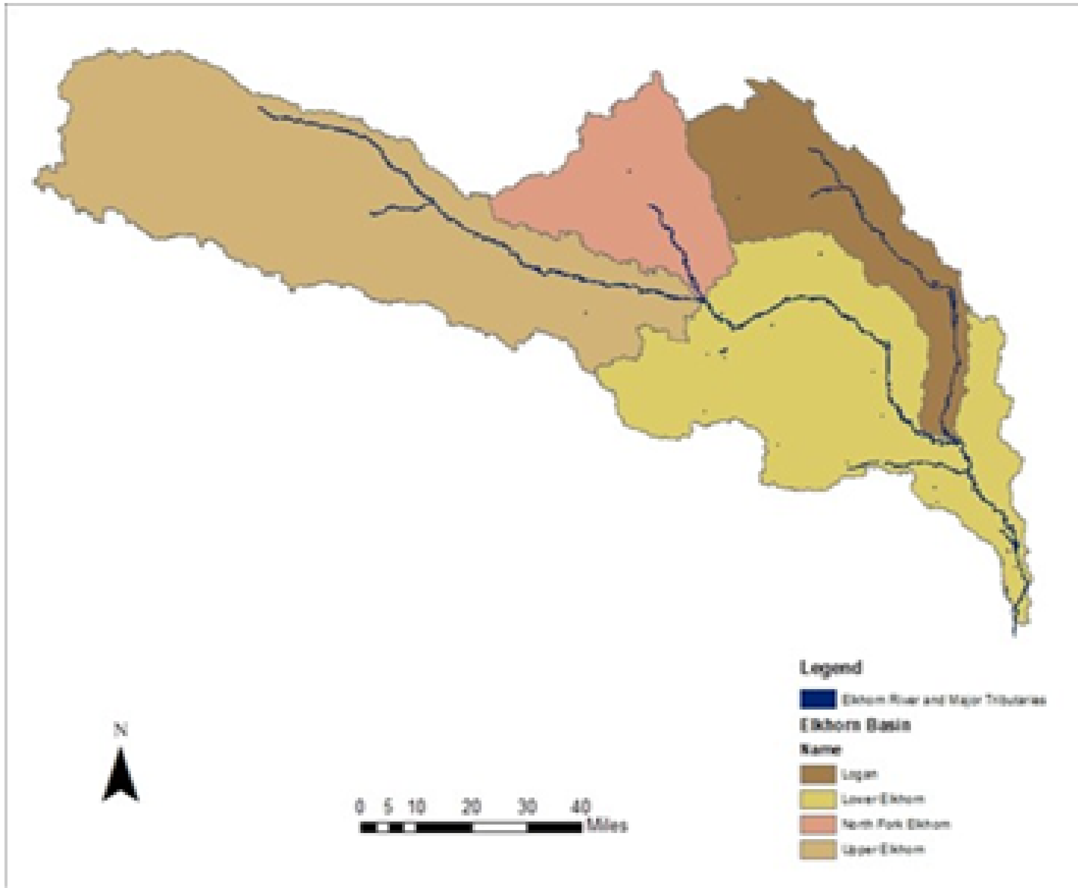


STUDY AREA AND DATA

Study Area:

- Catchment: Elkhorn river basin
- Sub-catchments: Upper Elkhorn, North Fork Elkhorn, Logan, and Lower Elkhorn.
- Location: Northeast and north-central Nebraska
- Area of catchment: 17,871 km^2
- Length of Elkhorn river: 466.71 km

Figure 2: Elkhorn river basin

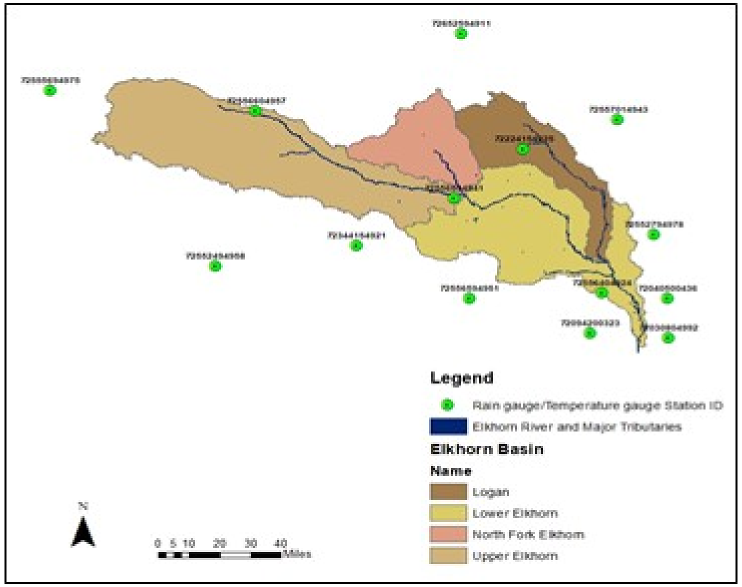


Data Collection:

1. Climatological Data:

- For ground-Based Station: National Water Information System: USGS Water Resources (NWIS, 2020)

Figure 3: Meteorological stations in Elkhorn river basin

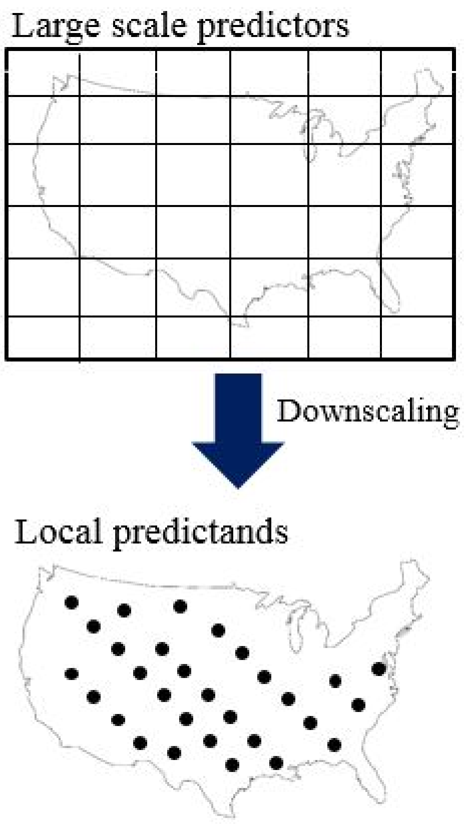


- For Ensemble Data: Earth System Research Laboratories (ESRL): Global Ensemble Forecast System (GEFS-Reforecast-V2) (GEFS,2020)

METHODOLOGY

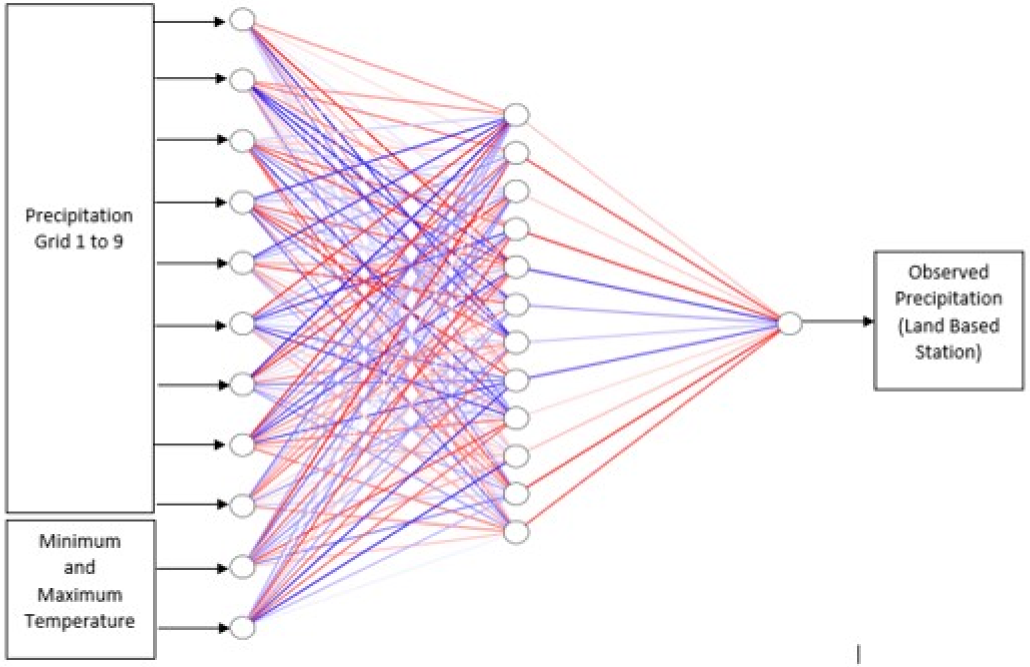
Statistical Downscaling:

Figure 4: Statistical Downscaling Technique:



Artificial Neural Network:

Figure 5: Neural Network diagram



- The feedforward backpropagation network was developed with precipitation, minimum temperature, and maximum temperature as input variables, which was trained using the Levenberg-Marquardt algorithm.
- To capture the spatial variability, precipitation values from nine adjacent grid cells, with the gauge station located in the middle cell, were used as inputs.
- Output variable: observed precipitation from ground-based stations.
- The sigmoid transfer function was used between the input and hidden layers, whereas the linear transfer function was used between the hidden and output layers.
- Number of iterations: 500; number of neurons: 11; number of hidden layer: 1.
- In the calibration data set of 2009 to 2016, the training was carried out with 70%, validation with 15%, and testing with 15% of the data, and the performance was further validated with the validation data set of 2017 to 2019.

PRELIMINARY RESULTS

Result 1:

Figure 6: Station 1: Correlation coefficient between the ground-based precipitation target variable and GEFS precipitation and temperature input variables for all leads (Day 0 to 15)

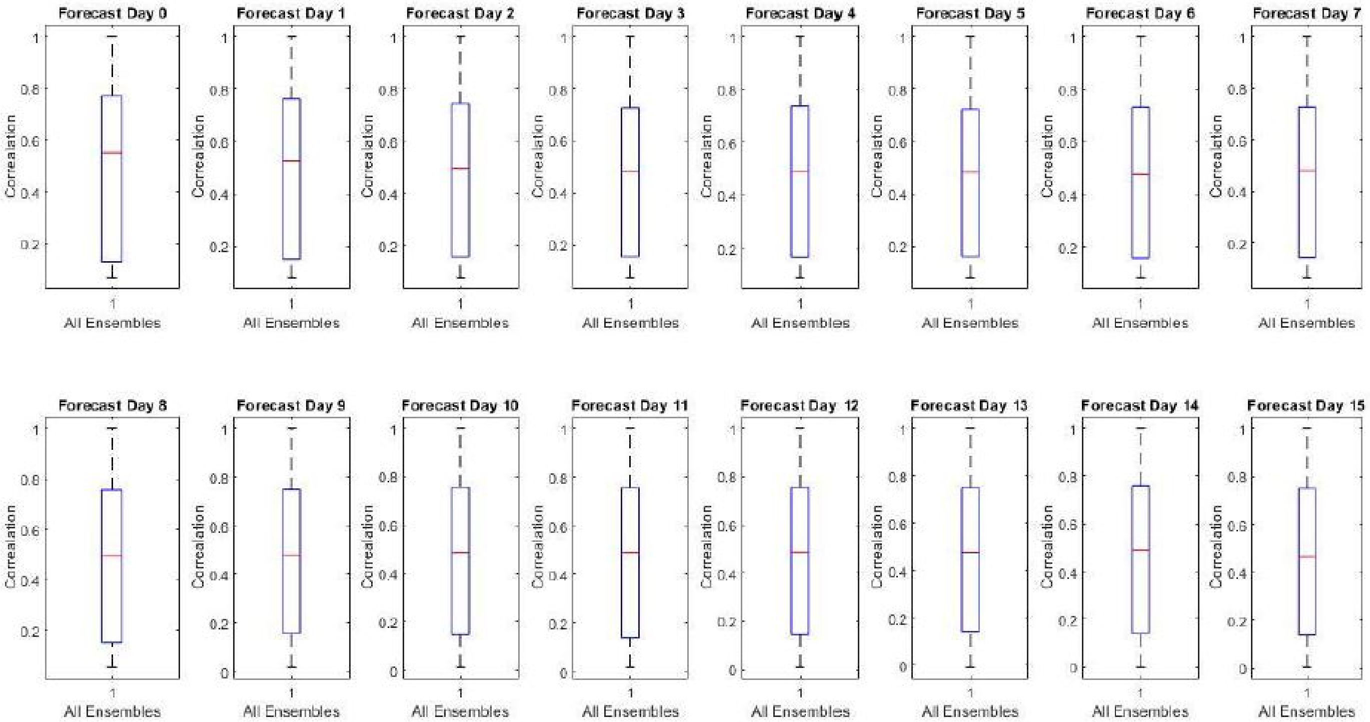


Table 1: Range of correlation for all ensembles, forecasts, and stations

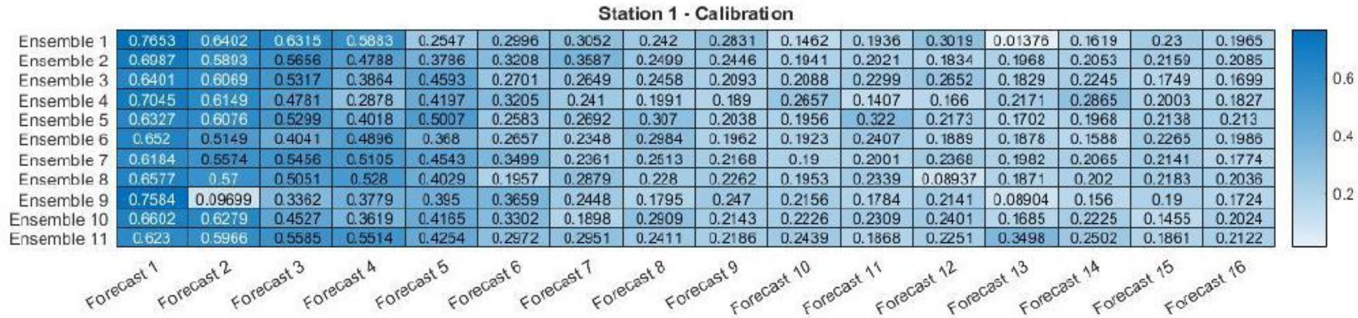
Forecast		Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15
Station 1	Min	0.0702	0.0822	0.078	0.0747	0.0884	0.0813	0.0851	0.0669	0.0542	0.0161	0.0116	0.0172	0.0088	-0.0097	-5.47E-04	0.0041
	Max	0.8864	0.8855	0.8739	0.9049	0.9028	0.8992	0.8374	0.8711	0.8931	0.8465	0.8682	0.8825	0.8762	0.8653	0.8926	0.841
Station 2	Min	0.1173	0.1439	0.1417	0.134	0.1356	0.1051	0.0649	0.0631	0.0361	0.0188	0.011	0.0017	-0.0038	0.0046	0.0185	0.0123
	Max	0.8876	0.9034	0.8811	0.9034	0.8705	0.911	0.852	0.8746	0.9035	0.8882	0.9082	0.9067	0.8944	0.9175	0.9118	0.8875
Station 3	Min	0.1064	0.0991	0.0776	0.0618	0.0412	0.0325	0.0189	0.0114	-0.0038	-0.0039	-0.0048	-0.0057	-0.0106	0.002	-0.0183	-0.018
	Max	0.901	0.8789	0.8949	0.8834	0.894	0.8828	0.8761	0.8758	0.9252	0.9294	0.893	0.9163	0.8971	0.917	0.9136	0.9141
Station 4	Min	0.1418	0.1584	0.1567	0.1373	0.1532	0.1303	0.0692	0.0605	0.0355	0.0174	0.0097	0.0059	0.009	0.0099	0.0018	0.0078
	Max	0.8896	0.9052	0.8899	0.8849	0.9009	0.8622	0.8425	0.8798	0.893	0.8977	0.8797	0.8844	0.8773	0.9166	0.8833	0.8949
Station 5	Min	0.1325	0.1555	0.1463	0.1408	0.1259	0.1231	0.0769	0.0399	0.0329	0.0161	0.0085	7.23E-04	0.0013	0.009	-0.0058	-0.0022
	Max	0.8701	0.8605	0.8252	0.8856	0.8776	0.8889	0.8713	0.8897	0.8979	0.906	0.891	0.8818	0.8728	0.8921	0.8897	0.8678
Station 6	Min	0.1325	0.1555	0.1463	0.1408	0.1297	0.12	0.0752	0.0386	0.0338	0.0169	0.0168	5.57E-05	2.07E-05	-0.0045	0.0011	0.0098
	Max	0.8701	0.8605	0.8252	0.8856	0.8776	0.8889	0.8713	0.8897	0.8979	0.906	0.891	0.8818	0.8728	0.8921	0.8897	0.8678
Station 7	Min	0.1331	0.1522	0.154	0.1476	0.1413	0.1225	0.0928	0.0676	0.0406	0.0247	0.0206	0.0192	-9.86E-05	-0.0045	0.0061	0.0043
	Max	0.8918	0.8743	0.8541	0.8943	0.8733	0.8829	0.8665	0.8898	0.9059	0.9119	0.8899	0.9179	0.9051	0.9144	0.9033	0.8937
Station 8	Min	0.1467	0.1777	0.1648	0.1608	0.1498	0.144	0.0791	0.0563	0.0458	0.0258	0.0116	0.0103	0.0091	0.0024	-0.0144	0.0053
	Max	0.8758	0.8508	0.8996	0.8802	0.8778	0.8907	0.8711	0.87	0.9169	0.9101	0.9101	0.9074	0.9268	0.9068	0.9211	0.8821
Station 9	Min	0.1467	0.1793	0.1648	0.1223	0.1362	0.101	0.0639	0.065	0.0405	0.0326	0.0106	0.0112	0.0064	0.0052	-0.0016	-1.63E-04
	Max	0.8758	0.8508	0.8996	0.8802	0.8778	0.8907	0.8711	0.87	0.9169	0.9101	0.9101	0.9074	0.9268	0.9068	0.9211	0.8821
Station 10	Min	0.0965	0.1097	0.1118	0.1072	0.1151	0.0928	0.0932	0.053	0.0332	0.0367	0.0255	0.0109	0.0073	0.0057	0.0012	0.0025
	Max	0.8741	0.8702	0.8612	0.8719	0.864	0.91	0.8843	0.8468	0.9046	0.8807	0.9222	0.9055	0.8729	0.8971	0.903	0.8899
Station 11	Min	0.1331	0.1656	0.1541	0.1476	0.1413	0.0975	0.0603	0.0492	0.0288	0.0281	0.0071	0.0031	-0.0054	0.007	-0.0176	0.0056
	Max	0.8918	0.8743	0.8541	0.8943	0.8733	0.8829	0.8665	0.8898	0.9059	0.9119	0.8899	0.9179	0.9051	0.9144	0.9033	0.8937

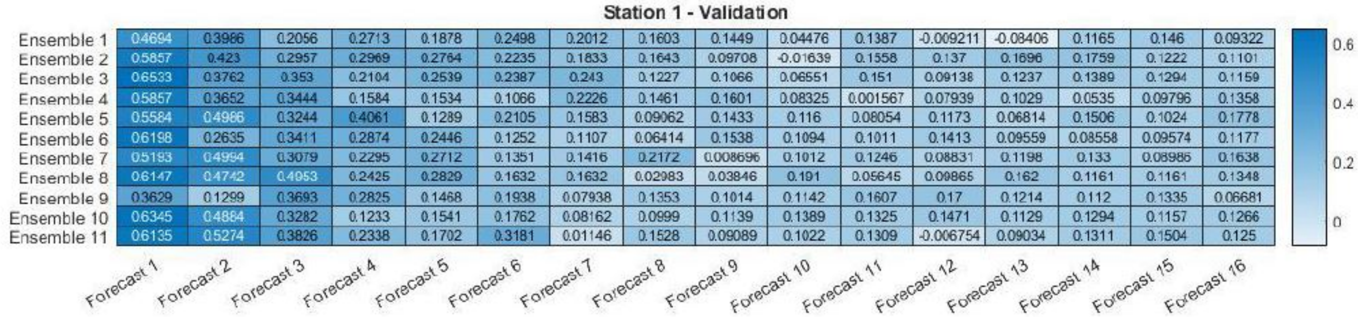
Discussion:

- Input variables from GEFS are correlated with the ground-based stations.
- All the correlation is within the mean range of [0.4, 0.6].
- The range of minimum and maximum correlation vary highly for each day forecast.
- The analysis also suggests that input variables derived from the low-resolution grid (Day +8 to +15) is poorly correlated in comparison with a high-resolution grid (Day 0 to +7).

Result 2:

Figure 7: Station 1: Performance of calibrated and validated results of the trained neural network





Discussion:

- The correlation coefficient was obtained from the statistical analysis of the output of the neural network by comparing predictors and predictands.
- For any given forecast day 0 to +3 showed a better correlation coefficient in comparison to the later forecast days.
- The model performance is dependent on the correlation between input variables and the output variables.

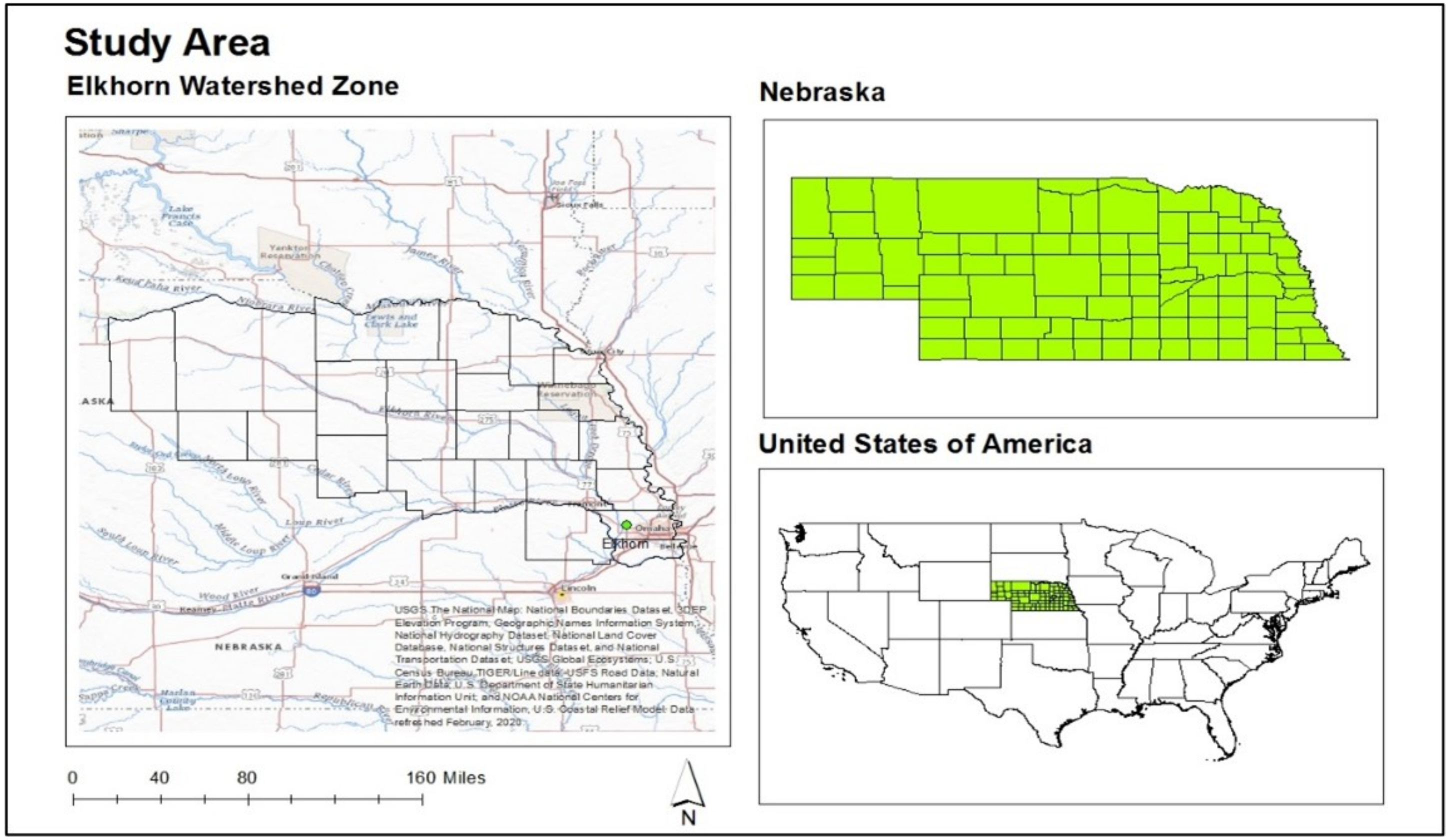
OBJECTIVE & RESEARCH QUESTION

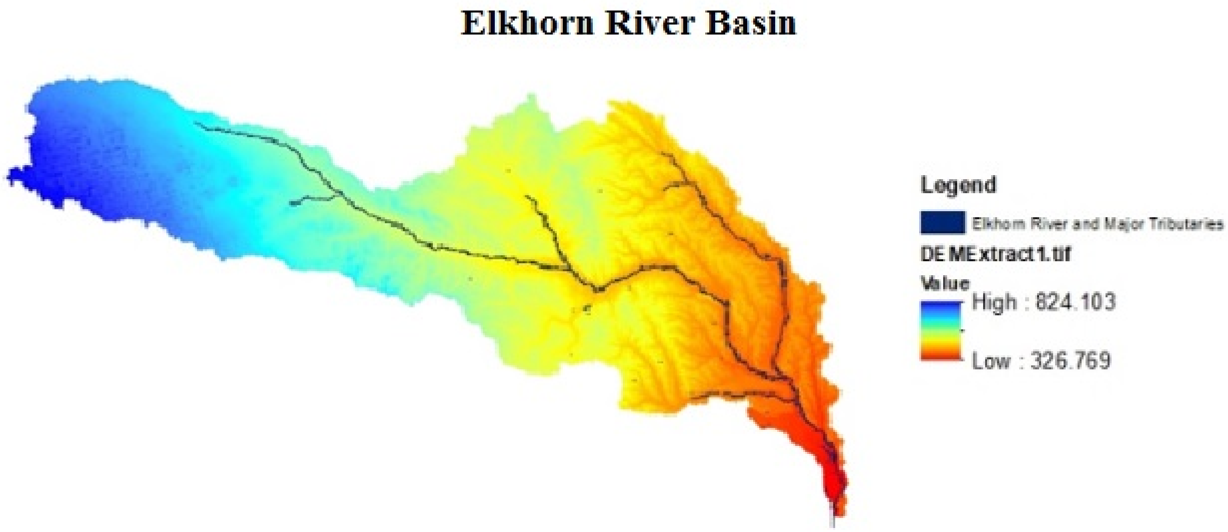
Objective:

Statistical downscaling using Artificial Neural Network technique of precipitation of various ensembles from GEFS forecasts.

Research question:

How does the predictability change with lead times while downscaling the ensemble precipitation forecasts from the GEFS system?




Elkhorn River Basin
Legend
Elkhorn River and Major Tributaries
DEMExtract1.tif
Value
High : 824.103
Low : 326.769

CONCLUSION

- Statistical downscaling technique using artificial neural network showed good performance for the first few days forecast.
- Further analysis of optimum network architecture needs to be carried out.
- Results of the correlation plot could be used to study how predictability varies along with the forecast lead time. More in-depth analysis is needed to better understand predictability change.
- This is a preliminary assessment. Additional variables can be included to test if the performance can be improved.

DISCLOSURES

Station ID	Station Number	Station	Start Date	Latitude (m)	Longitude (m)	Elevation (m)
72555694975	1	AINSWORTH MUNICIPAL ARPT	1/1/06	42.58	-100.00	787.60
72344154921	2	ALBION MUNICIPAL AIRPORT	1/1/06	41.73	-98.05	548.34
72040500436	3	BLAIR MUNICIPAL AIRPORT	1/1/09	41.41	-96.11	395.94
72652594911	4	CHAN GURNEY MUNICIPAL ARPT	1/1/06	42.88	-97.36	357.23
72556594951	5	COLUMBUS MUNICIPAL AIRPORT	1/1/06	41.43	-97.35	441.05
72556404924	6	FREMONT MUNICIPAL AIRPORT	1/1/06	41.45	-96.52	366.67
72556014941	7	KARL STEFAN MEMORIAL AIRPORT	1/1/73	41.99	-97.44	472.74
72557014943	8	SIOUX GATEWAY/COL BUD DAY FIELD AP	1/10/42	42.39	-96.38	333.76
72552794978	9	TEKAMAH MUNICIPAL AIRPORT	1/1/06	41.76	-96.18	313.33
72556604957	10	THE O'NEILL MUNI-JOHN L BAKER FIELD AIRPORT	1/1/06	42.47	-98.69	619.05
72224154925	11	WAYNE MUNICIPAL AIRPORT	1/1/06	42.24	-96.98	434.04

AUTHOR INFORMATION

Smit Doshi

Smit worked in the Roy research group from March 2020 to August 2020. He received the prestigious Erasmus Mundus Scholarship for his master's degree. His master's thesis work was carried out here at UNL on the topic of probabilistic flood inundation mapping in the Elkhorn River basin, Nebraska. He attended two European Universities for his master's coursework, Universitat Politècnica de Catalunya in Spain and Newcastle University in the UK. Smit has a bachelor's degree in Civil Engineering from the Maharaja Sayajirao University of Baroda, India, where he was a gold medalist. His research interests include flood modeling and risk assessment, hydrology, hydraulics, irrigation engineering, machine learning, and stormwater management.

Tirthankar Roy (Advisor)

Tirthankar Roy joined the Department of Civil and Environmental Engineering at the University of Nebraska-Lincoln (UNL) in the fall of 2019. Prior to joining UNL, he was a Postdoc in the Department of Civil and Environmental Engineering at Princeton University (2017-2019). He holds a Ph.D. in Hydrology from the University of Arizona (2017), an M.Tech. in Civil Engineering from the Indian Institute of Technology Kanpur (2012), and a B.Tech. in Agricultural Engineering from Bidhan Chandra Krishi Viswavidyalaya, which is a state agricultural university in West Bengal, India. During his M.Tech., he received the DAAD Scholarship to work on his thesis at Technische Universität Dresden, Germany. He serves on the Early Career Committee of the International Association of Hydrological Sciences and the Hydrological Uncertainty Technical committee of the American Geophysical Union. His research interests include satellite remote sensing applications in hydrology, hydrologic extremes, catchment hydrology, land-atmospheric interactions, statistics and machine learning, water and the society, and water resources management.

ABSTRACT

The NOAA Physical Sciences Laboratory produces the Global Ensemble Forecasting System (GEFS) which comprises 11 ensemble members (1 control and 10 perturbation runs) for over a 36-year period (December 1984 to present), with forecasts initialized every day for the next 16 days (first 8-day forecasts obtained from a high-resolution grid and the next 8-day forecasts from a low-resolution grid). The system provides 36 variables related to a wide range of hydrometeorological processes. In this study, we assess the predictability of precipitation within the context of statistical downscaling using a minimum set of predictor variables (precipitation and temperature). We use feedforward backpropagation neural networks with a suite of training algorithms to determine which variables (features) are of most relevance at different forecast lead times. The outcome of this study will significantly benefit short-term flood forecasting using GEFS data.

REFERENCES

1. Hamill, T. et al. (2020) ‘A Description of the 2nd Generation NOAA Global Ensemble Reforecast Data Set’, NOAA Earth System Research Lab, Physical Sciences Division Bouder, Colorado, USA, p. 10.
2. NWIS (2020) USGS Water Resources.
3. Schoof, J. T. and Pryor, S. C. (2001) ‘Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks’, International Journal of Climatology, 21(7), pp. 773–790. doi: 10.1002/joc.655.
4. GEFS (2020) NOAA PSL GEFS Reforecast Version 2.
5. Sharma, V. K. et al. (2017) ‘Satellite data planning for flood mapping activities based on high rainfall events generated using TRMM, GEFS and disaster news’, Annals of GIS. Taylor & Francis, 23(2), pp. 131–140. doi: 10.1080/19475683.2017.1304449.
6. Hsu, K. -l, Gupta, H. V. and Sorooshian, S. (1995) ‘Artificial Neural Network Modeling of the Rainfall-Runoff Process’, Water Resources Research, 31(10), pp. 2517–2530. doi: 10.1029/95WR01955