



## Assessing Climate Change Impacts on the Environment: Comparative Analysis of Temperature Predictors Using Various Statistical Methods

KESHAV KUMAR<sup>1\*</sup>, MANI BHUSHAN<sup>2</sup>, MADAN KUMAR<sup>3</sup>,  
AVINASH KUMAR<sup>4</sup> and NIRAJ KUMAR SINGH<sup>5</sup>

<sup>1</sup>Department of Civil Engineering, Nalanda College of Engineering Chandi (Nalanda),  
Department of Science, Technology and Technical Education, Government of Bihar, India.

<sup>2</sup>Principal-in-Charge, Government Engineering College Khagaria, Department of Science,  
Technology and Technical Education, Government of Bihar, India.

<sup>3</sup>Lecturer, Department of Civil Engineering, Government Polytechnic Muzaffarpur, Bihar-842001, India.

<sup>4</sup>Department of Mechanical Engineering, Nalanda College of Engineering Chandi, Bihar-803108, India.

<sup>5</sup>Government Polytechnic Jamui, Department of Science, Technology and Technical Education,  
Government of Bihar, India.

\*Corresponding author E-mail: keshav151008@nitp.ac.in

<http://dx.doi.org/10.13005/ojc/410323>

(Received: November 11, 2024; Accepted: April 01, 2025)

### ABSTRACT

To effectively assess climate change modeling, it is crucial to carefully select predictor variables such as geopotential height (GH), relative humidity (RH), specific humidity (SH), u-wind (UW), precipitation, and temperature to establish robust relationships among them. This study employs several statistical techniques, including factor analysis and correlation coefficient, to identify the optimal predictors for downscaling the predictand. By meticulously choosing predictors, the study aims to enhance model performance, computational efficiency, and overall output accuracy. The research focuses on the Bagmati River Basin, originating from the Himalayan region in Bihar, India, an area known for its flood susceptibility. During heavy rainfall, the Bagmati River frequently overflows, leading to substantial water discharge and debris transfer from the erodible Himalayan slopes. This basin was chosen to study the impact of climate change on the environment by selecting temperature predictors using a variety of statistical methods. The study utilizes 33 years of data (1988-2021) of reanalysis data from NCEP/NCAR and observed temperature data from the IMD Pune. Among the statistical techniques applied, factor analysis demonstrated superior performance.

**Keywords:** Climate change, Statistical techniques, Factor analysis, Bagmati river.

### INTRODUCTION

The impact of climate change on temperature

has profound and far-reaching consequences for our environment. As human activities like fossil fuel combustion and deforestation drive global



temperatures upward, observing a series of cascading impacts on our planet. These effects include more frequent and severe heatwaves, altered precipitation patterns, shifting ecosystems, and the accelerated melting of polar ice caps and glaciers.

It is quite challenging to choose the variables or elements that impact temperature since there are so many factors that affect temperature, which is unpredictable. Numerous data are produced by various factors. Today's real-world applications generate a lot of data, which necessitates the use of multivariate approaches for data analysis. Taking into account diverse factors lengthens the computing process and decreases the model's effectiveness, output, and performance. The model's efficiency is increased and calculation time is decreased via the use of statistical approaches. Variable selection strategy and data reduction are often required to reduce the impact of such noisy variables<sup>1</sup>.

The process of creating a link between the local climatic variables such as rainfall, evaporation, dew point, temperature, etc. and the coarse resolution GCM (General Circulation Model) variables such as relative humidity (RH), geopotential heights (GH), air temperature (AT), specific humidity (SH), u-wind (UW), and v-wind (VW) is known as downscaling. Statistical or dynamical downscaling may do this<sup>2</sup>. RCM (Regional Climate Model), which directly utilises the GCM output and provides information about the area's more complicated topography at a greater resolution, is used in dynamic downscaling. This approach had a restricted use since the complexity and inaccuracy associated with dynamical downscaling were greater than those with statistical downscaling. Although statistical downscaling requires a lot of data, it is simpler to use and has lower inaccuracy than other methods. Weather types, weather generators, and transfer functions may all be used for statistical downscaling<sup>3-8</sup>. The transfer function-based downscaling method is regarded as one of the most effective approaches for temperature downscaling. This method establishes a non-linear relationship between predictors, such as, data collected from the National Centers for Environmental Prediction (NCEP), and the predictands, which are the observed temperature data. It is particularly well-suited for capturing complex, non-linear, and time-varying relationships

between the input and the output variables, making it a powerful tool for temperature prediction.

Before statistical downscaling can commence, it is essential to establish the optimal link between predictors and predictands (such as temperature and precipitation)<sup>9-11</sup>. This study identifies the best predictors for downscaling the predictand using statistical techniques like factor analysis and correlation coefficients. Factor analysis allows us to combine sets of observed variables, providing a more accurate measure of a given factor by confirming that these variables measure the same underlying component, albeit with varying reliability. Factor analysis (FA) encompasses a range of techniques used to examine how responses to various measured variables are influenced by underlying constructs. FA includes two main types:

- (A) **Confirmatory Factor Analysis:** This method is used to test the factor structure of a set of observed variables. It is statistical method to examine whether the predicted responses of a given set of constructs are affected.
- (B) **Exploratory factor analysis:** This method is used to verify the underlying structure of a large set of variables. It is statistical method which aims of discovering the nature of the constructs that influence a set of responses.

#### DATA collection and Study area

This study is carried out on Bagmati River Basin. It originated from Himalaya region and runs continuously across both the countries, Nepal and India, particularly in North Bihar. The Shivpuri ranges hills situated in Nepal which is approximately 16 km north-east of Kathmandu, located in longitude 85°17'E and latitude 27°47'N, and at an elevation of fifteen hundred meter above mean sea level (MSL), are the source of the Bagmati River. The Kathmandu Valley situated in Nepal, which is mostly rice-Pattie in sloping hillsides, is 10 kilometres long and runs south-westerly following the Bagmati River. The Pashupatinath Temple is built atop one of the several impassive rock surfaces that prevent the valley's flow below. The river travels south through a plain after passing a temple, where it is joined by the Monahara River and turns westward. The river runs through numerous historic, significant areas in Kathmandu.

The Bagmati river basin have four stations, Dheng, Kamtaul, Benibag, and Hayaghat, that provided monthly temperature data for thirty-four years, or from 1988 to 2021, shown in Fig. 1. From the NCEP/NCAR

REANALYSIS project, another set of mean monthly data for thirty-four years, or from 1988 to 2021, is collected for the region under study at various pressure levels from 1000 mb to 100 mb. In order to determine the prospective predictor, a connection is made between observed Indian meteorological department (IMD) data which make up the predictant data<sup>2</sup> and NCEP data utilizing the factor analysis approach and the product moment correlation calculation<sup>12</sup>. These potential predictors used in this study are the atmospheric factors at various pressure levels that most closely align with the collected observed data.

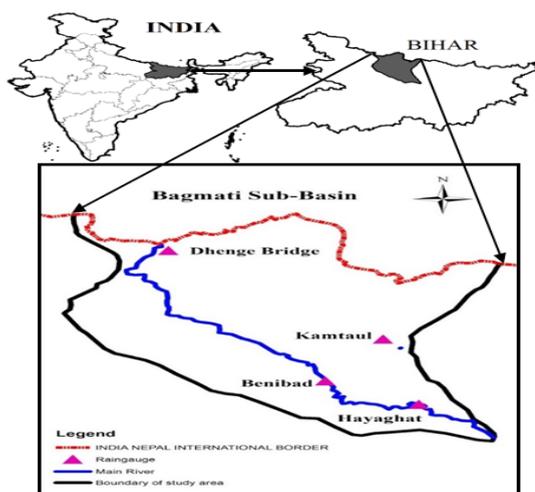


Fig. 1. Selected Locations of Bagmati River Basin

## Methodology

### Selection of potential predictor

It is exceedingly difficult to narrow down the numerous unexpected elements that impact temperature since there are so many. Numerous data are produced by various factors. Today's real-world applications generate a lot of data, which necessitates the use of multivariate approaches for data analysis. Taking into account diverse factors lengthens the computing process and decreases the model's effectiveness, the output, and the performance of the model. The predictor model's efficiency is increased and calculation time is decreased via the use of statistical approaches. Variable selection strategy and data reduction are often required to reduce the impact of such noisy variables. For the optimum association to be established between the predictors (such as RH, GH, AT, UW, VW, SH, etc.), the predictors must be carefully chosen for statistical downscaling (like precipitation, temperature etc). Fig. 2 shows the flow chart diagram of methods used in this study. The best temperature downscaling predictions

are discovered using two statistical approaches. One approach is the Pearson Product Moment Correlation (PPMC) while factor analysis is an another approaches.

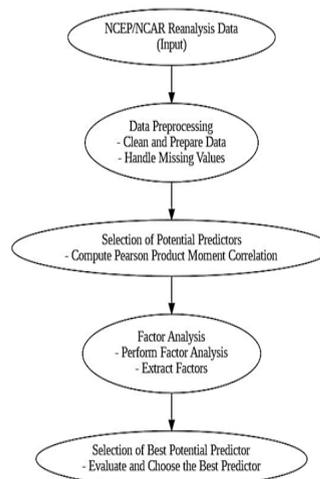


Fig. 2. Flow chart of the methodology used in this study

### Pearson product moment correlation

The PPMC formula is used to determine the coefficient of correlation between these two sets of data, as shown below:

$$\rho_{xy} = \frac{\sum_{n=1}^N (x_n - \bar{x})(y_n - \bar{y})}{N\sigma_x\sigma_y} \quad (1)$$

Where,

$x_1$  = predictant,  $y_1$  = predictor,  $\rho_{xy}$  = correlation coefficient,  $\sigma_x$  = standard deviation of predictant,  $\sigma_y$  = standard deviation of predictor,  $N$  = total number of data.

### Factor analysis

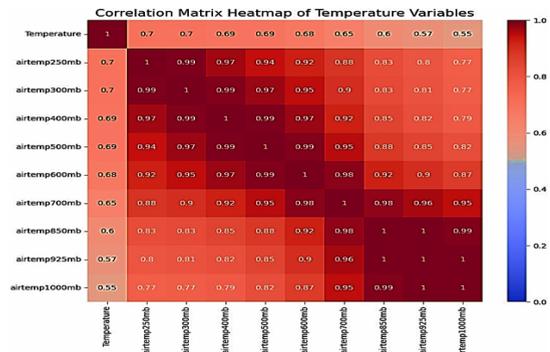
One statistical method for reducing a big dataset to its essential components is factor analysis. To do this, it takes each variable and adds together their maximum shared variation to get a composite score. Using this score as a measure of all elements allows for further investigation. As with the general linear model (GLM), the factor analysis technique relies on a number of assumptions made during the study. These methods include the correlation between variables and factors considered. The existence of a linear relationship, the absence of multicollinearity, the inclusion of relevant variables, and the reality of correlation are obtained.

## RESULT AND DISCUSSION

### Potential predictor selection

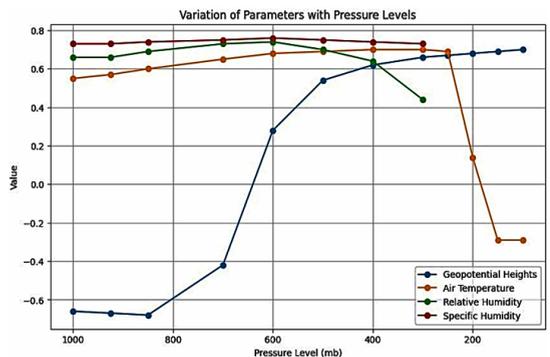
Potential predictors from the NCEP/NCAR

REANALYSIS model, such as GH, RH, AT and SH at various pressure levels, are connected with the observed mean monthly temperature collected from IMD, Pune. These variables are chosen as possible temperature predictors because they have values in the range of 0.5 to 1.0, which is regarded to be a good connection. Fig. 3 displays the calculations for the temperature's potential predictor.



**Fig. 3. Matrix of correlation between air temperature at different pressure level with observed temperature**

Figure 3 illustrates the correlation matrix between air temperature at various pressure and observed temperature. Despite the generally strong correlation across these levels, the highest correlation is found between air temperature and temperature at the 250 and 300 mb pressure levels, with correlation values of 0.6998 and 0.6987, respectively. At 500 mb pressure, there is little correlation between temperature and air temperature (0.5534). The correlation coefficients between each predictor and temperature at different pressure levels, ranging from 100 mb to 1000 mb, are shown in Fig. 4. It shows that geopotential height and temperature have strong correlations at pressure levels of 100 mb, 150 mb, 200 mb, 250 mb, 300 mb, and 400 mb. It also shows moderate correlations at pressure levels of 1000 mb, 925 mb, 850 mb, and 700 mb.



**Fig. 4. Potential Predictor for Temperature**

**Selection of potential predictors using factor analysis**

This study examines potential predictors from the NCEP/NCAR reanalysis model, such as GH, RH, SH, and AT at various pressure levels, in conjunction with observed temperatures. The goal is to identify the factors that influence temperature. Table 1 displays the eigenvalues, their associated variability, and cumulative data.

**Table 1: Eigenvalues, variability and cumulative data for different factors**

	Eigenvalue	Variability (%)	Cumulative %
f1	8.19	91.03	91.03
f2	0.21	2.33	93.36
f3	0.00	0.05	93.41

Among the three components presented in Table 1 factors f2 and f3 have eigenvalues less than 1; whereas, factors f1 has observed an eigenvalue larger than 1. The factor f1, variability is observed as 91.03%, whereas the variability of other components is just 8.98%. Therefore, the factor whose variance is closest to 100% and whose eigen value is bigger than 1 has been chosen. Fig. 5 shows the factor pattern for temperature and specific humidity at different pressure levels. Fig. 5 illustrates the factor distribution of all the predictors by taking into account factor f1. According to Fig. 5, the specific humidity at 600 mb and 700 mb has a value of 0.9963 and 0.9918, which is extremely near to 1. It demonstrates how these two variables have a greater impact on temperature.

The findings of a similar factor analysis on additional predictors are shown in Figure 5.

In Fig. 5, the geopotential height (GH) with values of 0.9934 and 0.9968 is identified as a potential predictor due to its proximity to 1. The values of specific humidity (SH) at the pressure levels of 600 mb and 700 mb, which are 0.9961 and 0.9913 respectively, are considered as predictors due to their closeness to 1. Similarly, in the case of relative humidity (RH) at these pressure levels, with values of 0.9550 and 0.9675, is selected as a potential predictor because these values are relatively near 1. Additionally, for the air temperature (AT) at the 600 mb and 700 mb levels, with values of 0.9901 and 0.9892, is identified as a highly potential predictor due to its proximity to 1.

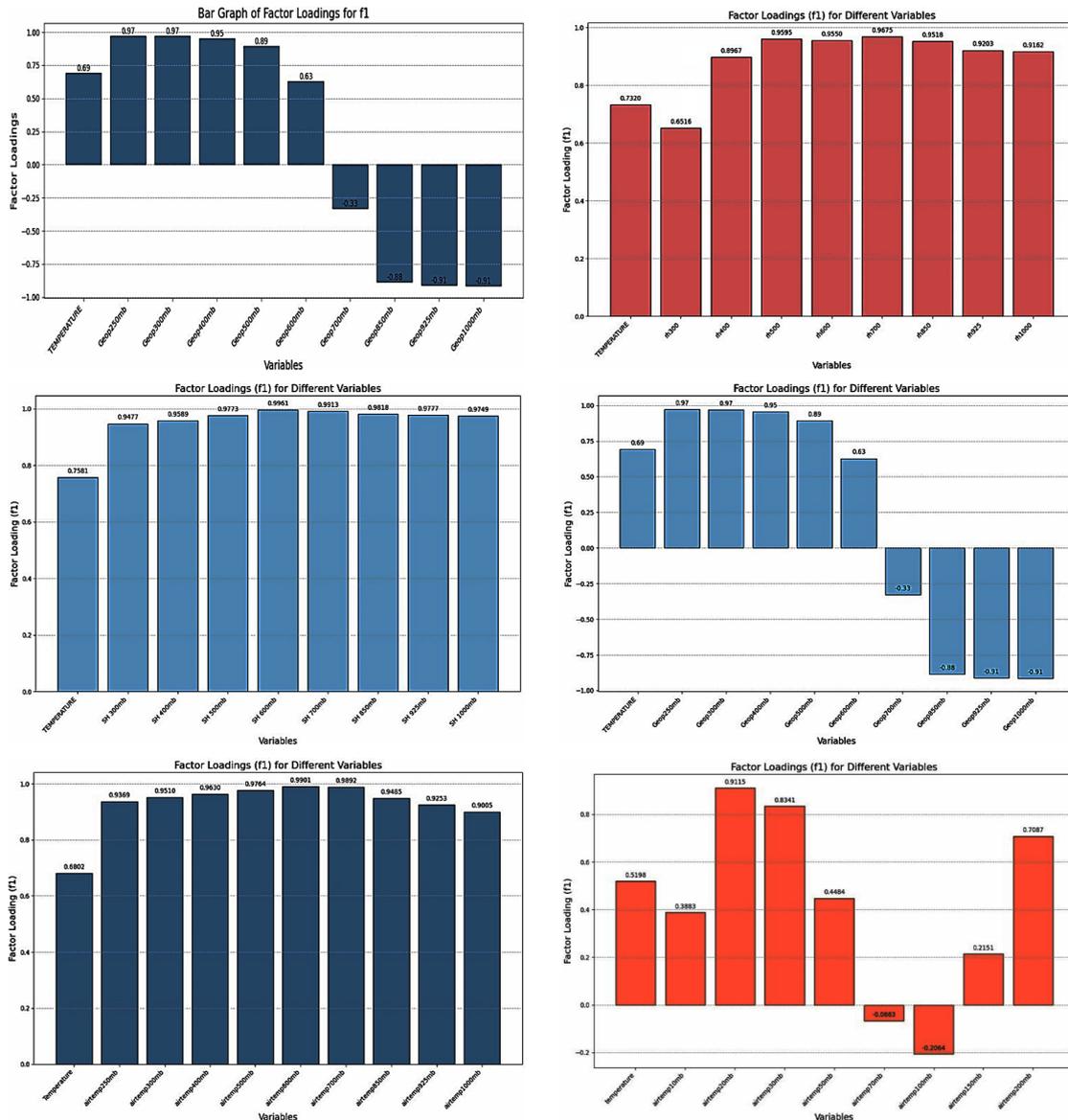


Fig. 5. Potential Temperature Predictor for by factor analysis at different pressure level

**CONCLUSION**

Factor analysis was employed to refine the selection of predictors for temperature based on initial positive correlations observed using PPMC across various pressure levels (100 mb to 1000 mb). Among the predictors GH, RH, AT, and SH correlation values ranged from 0.5 to 0.75. Factor analysis method identified three factors (f1, f2 and f3). Among three factors, one factor (f1) having an eigenvalue greater than 1 and explaining 91.02% of the variability, indicating its suitability for predicting temperature.

In conjunction with this factor, geopotential height at 50 mb and 70 mb were observed as 0.993 and 0.996 respectively. Hence, it was selected as an optimal predictor due to its correlation values close to 1. Specific humidity at 600 mb and 700 mb were observed as 0.996 and 0.991, respectively, and relative humidity at the same levels (0.9550 and 0.9675) were also chosen for their strong correlation values. These selections were considered by factor analysis, which provided clearer insights compared to the initial Pearson product moment correlation approach.

**ACKNOWLEDGEMENT**

The authors are thankful to Nalanda College of Engineering, Chandi (Nalanda); Government Engineering College, Khagaria; Government Polytechnic, Muzaffarpur and Government Polytechnic, Jamui, Bihar, India, for providing facilities for conducting this research work.

**Conflict of interest**

All authors declare that they have no conflict of interest concerning this article.

**Funding Sources**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**REFERENCES**

1. Kumar, K.; Singh, V., & Roshni, T., *Water Practice & Technology*, **2022**, 17(12), 2554-2569.
2. Wetterhall, F.; Halldin, S. and Xu, C.Y., *Journal of Hydrology*, **2005**, 306, 136–174.
3. B.M, F., *International Journal of Water Sciences*, **2012**, 1.
4. Clement Tisseuil.; Mathieu Vrac b.; Sovan Lek a.; Andrew J. Wadec., *Journal of Hydrology*, **2010**, 385, 279– 291.
5. Colin Harpham and Robert L. Wilby., *Journal of Hydrology*, **2005**, 312, 235–255.
6. David Hunter.; Hao Yu.; Michael S. Pukish, Janusz Kolbusz, and Bogdan M. Wilamowski., *IEEE Transactions On Industrial Informatics*, **2003**, 8, 228-240.
7. David Veitch.; Dissertation submitted for the Msc in Data Analysis, Networks And Non-linear Dynamics., **2005**.
8. Deepashree Rajee and P.P. Mujumdar; *Journal of Hydrology*, **2009**.
9. Ghosh, S. and P. P. Mujumdar., *Advances in Water Resources*, **2008**, 31(1), 132-146.
10. Skelly, W.C and Henderson-Sellers, A.N.N., *International Journal of Climatology*, **1996**, 079-1086.
11. Sylwia Trzaska and Emilie Schnarr; *African And Latin American Resilience To Climate Change Center for International Earth Science Information Network (CIESIN)*, **2014**, 56
- [12] Chitra NR. And Santosh G.Thampi., *International Journal of Innovative Research in Science, Engineering and Technology*, 2013, 2, 6-11