ROLE OF PREDICTORS IN STATISTICAL DOWNSCALING SURFACE TEMPERATURE FOR MALAPRABHA BASIN

Nagraj S. Patil[1], Soumya S. Bankapur[2]

[1]*Associate Professor, VTU PG Center, Belagavi, 590018, India*
[2]*2nd year M. Tech, VTU PG Center, Belagavi, 590018, India*

Abstract

A change in the statistical dissemination of weather patterns or arrangements which lasts for a prolonged time is called climate change. Researchers vigorously work to figure out past and future climate by using theoretical and observations models. Based on the physical sciences GCMs, are generally used in theoretical approaches to match past climate data, projection of future data, and to associate with the causes and consequences in climate change. As Global Climate Models are only accessible or available at coarse resolution the downscaling technique has been acknowledged as an essential component for the assessment of climate change impacts. In this paper, a surface temperature data is considered as a predictand and has been downscaled for the Malaprabha basin.The purpose of this study is to performance different methods of sensitivity analysis for selecting appropriate predictors influencing the predictand and to study the predictor-predictand relationship. Artificial Neural Network(ANN) methodology is utilized to downscale CanCM4 GCM surface temperature to local scale on monthly time series. The sensitivity analysis method PCA provides the most probable predictors strongly influencing the performance of the ANN model. Hence, this approach shows the importance of sensitivity analysis which helps in the performance of model for downscaling

Keywords: *Climate Change; Statistical Downscaling; Artificial Neutral Network; Global Climate Model; CanCM4 GCM.*

---***---

1. INTRODUCTION

A change in the statistical dissemination of weather patterns or arrangements which lasts for a prolonged time is called climate change. Researchers vigorously work to figure out past and future climate by using theoretical and observations models. Based on the physical sciences GCMs, are generally used in theoretical approaches to match past climate data, projection of future data, and to associate with the causes and consequences in climate change.As Global Climate Models are only accessible or available at coarse resolution the downscaling technique has been acknowledged as an essential component for the assessment of climate change impacts.

In this paper, a surface temperature data is considered as a predictand and has been statistically downscaled for the Malaprabha basin using ANN model.

The main purpose of this study is to performance different methods of sensitivityanalysis for selecting appropriate predictors influencing thepredictand and to study the predictor-predictand relationship.

Artificial Neural Network(ANN) methodology is utilized to downscale CanCM4 GCM surface temperature to local scale on monthly time series and provide future projection for the time period 2006-2035 for the considered stations.

2. STUDY AREA AND DATA EXTRACTION

2.1. Description of Study Region

Malprabha is a right bank tributary of Krishna River. The Malprabha catchment lies between North scopes 15° 45' and 16° 25' and east longitudes 74° 00' and 75° 55'. The Malprabha River begins from the Chorla Ghats, a segment of the Western Ghats, at a height of around 792m around 35m

south-west of Belgaum District of Karnataka. The mean month to month Tmax in the catchment fluctuates from 25 to 34 °C what's more, mean yearly Tmax is 28 °C. The mean month to month Tmin ranges from 17 to 21 °C. The Malaprabha bowl is one of the significant helps for the dry areas of north Karnataka (potentially the biggest dry area in India outside the Thar desert). Malaprabha store supplies water for watering system to the regions of northern Karnataka with an irrigable zone of 11549 sq.km. The area of the study areais delineated using SWAT is shown in Fig.1. below.

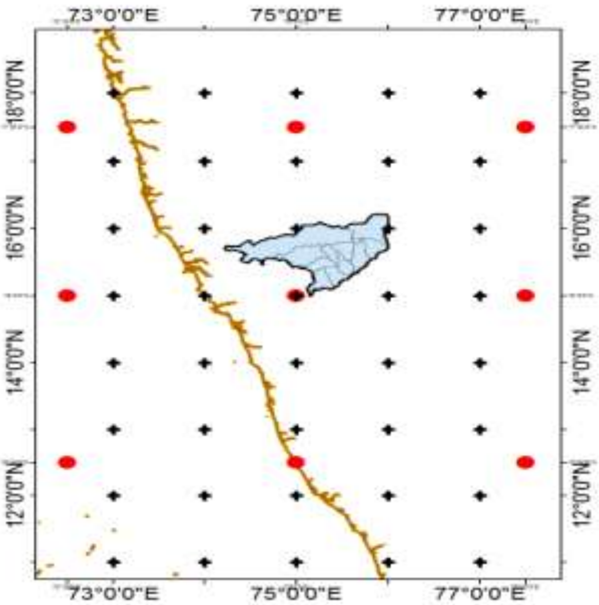

Fig.1. Delineation of Malaprabha Basin with GCM gird points

2.2. Data Utilized

GCM, atmosphere information's are utilized from Canadian Center for Climate Modeling and Analysis Canada (CCCM). CanCM4 the Fourth Generation coupled model creates the Historical(1961-2005) and the Future(2006-2035) atmosphere information. The dataset, accessible in NetCDF organization was perused in MATLAB. The information gives long haul month to month implies that are accessible for various variables which can be chosen relying upon the prerequisite of the study. A determination of variable (or Predictors) is itself an exploration.

2.3. Station Data

Station information was utilized for the institutionalization and acceptance of the downscaling model with the GCM atmosphere learning. Since measurable downscaling was performed at a station level/site-level; precipitation, least temperature and most extreme temperature was required for each station. The everyday information of the higher than communicated variables were given by the Indian meteorological Department (IMD), for a time of 37 years from 1969 to 2005. However 2 years(1969-70) was considered as a cradle and consequently IMD information for the four diverse station was made accessible from 1971-2005 and was changed over into mean month to month information.

3. SELECTION OF PREDICTOR VARIABLES

The choice of proper indicators for downscaling predictands is a standout amongst the most essential strides in a downscaling exercise (Hewitson and Crane, 1996; Cavazos and Hewitson, 2005). The decision of indicators could fluctuate from district to area contingent upon the qualities of the extensive scale air course and the predictand to be downscaled. Any sort of variable or record can be utilized as indicator the length of it is sensible to expect that there exists a relationship between the indicator and the predictand (Wetterhall et al., 2005). Regularly, in atmosphere sway concentrates, such indicators are picked as variables that seem to be:

(1) dependably reproduced by GCMs and are promptly accessible from documents of GCM yield and reanalysis datasets,

(2) unequivocally associated with the predictand and

(3) taking into account past studies.

Arrangements of CanCM4 model climatic variables considered for the study:

Table 1 CanCM4 model variables

Variable names	Short names	Units
Surface upward latent heat flux	hfls	W/m^2
Surface upward sensible heat flux	hfss	W/m^2
Precipitation	pr	$Kg/m^2/s$
Sea level pressure	psl	Pa
Air temperature	Ta	K
Surface temperature	Ts	K
Eastward wind	Va	m/s
Northward wind	Ua	m/s
Geopotential height	Zg	M

4. SENSITIVITY ANALYSIS

The sensitivity analysis strategies are performed with a specific end goal to depict how touchy the result variables i.e. indicators are to the variety of individual information parameters i.e. predictand.

Numerical models, for example, indicator models, are characterized by conditions, information components, parameters, and variables meant to portray the procedure being explored. The info is liable to numerous wellsprings of vulnerability including mistakes of estimation, nonappearance of data and poor or fractional comprehension of the main thrusts and systems. Affectability examination permits to portray the vulnerability connected with a model. This answers questions like "what information calculate should be explored in more detail?" or "what information elements can be evacuated?".

The following methods explained in the sections below are used for the screening of most probable predictor variables for the predictand considered for the study.

4.1. Pearson`s Product Moment Correlation Coefficient

Pearson's connection coefficient is a measure of the quality of the straight relationship between two such variables.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

$\bar{x}$ = *arithmetic mean of* $x(i)$

$\bar{y}$ = *arithmetic mean of* $y(i)$

n = *sample size*

Pearson’s r is between –1 and +1,
+1 = complete positive linear relation;
0 = no relation;
–1 = complete negative linear relation.

4.2. Spearman`s Rank Correlation Coefficient

In insights, Spearman's rank connection coefficient or Spearman's rho, named after Charles Spearman and regularly indicated by the Greek letter (rho) or as , is a nonparametric measure of factual reliance between two variables. It evaluates how well the relationship between two variables can be depicted utilizing a monotonic capacity.

Computation of the Spearman rank relationship coefficient (ρ) should likewise be possible as takes after:

$$\rho = 1 - \frac{6 \cdot \sum_{i=1}^{n}(r_i - s_i)^2}{n^3 - n}$$

r_i = *rank of* $x(i)$

r_i = *rank of* $y(i)$

n = *sample size*

Spearman’s ρ is between –1 and +1,
+1 = complete positive monotonic relation;
0 = no relation;
–1 = complete negative monotonic relation.

4.3. Principal Component Analysis

It is a method for recognizing designs in information, and communicating the information so as to highlight their similitude and contrasts. Since examples in information can be elusive in information of high measurement, where the advantage of graphical representation is not accessible, PCA is a capable device for examining information. The quantity of main parts is not exactly or equivalent to the quantity of unique variables. The other principle point of preference of PCA is that once you have found these examples in the information, and you pack the information, i.e. by decreasing the quantity of measurements, without much loss of data.

Here, Principal component analysis is executed utilizing an include as a part of programming for exceed expectations called NUMXL and OriginPro 2016 where the most likely indicator variables are been screened for further examination.

4.4. Analysis of Variance (ANOVA)

Ronald Fisher. In the ANOVA setting, the watched difference in a specific variable is divided into segments inferable from various wellsprings of variety. In its least complex structure, ANOVA gives a measurable test of regardless of whether the method for a few gatherings are equivalent, and accordingly sums up the t-test to more than two gatherings. ANOVAs are helpful for looking at (testing) three or more means (gatherings or variables) for measurable noteworthiness. It is reasonably like various two-example t-tests, however is less preservationist (results in less sort I blunder) and is in this way suited to an extensive variety of down to earth issues.
Hence, one-way ANOVA is carried out using OriginPro 2016 software for selecting the probable predictors for the study area.

5. ARITIFICAL NEURAL NETWORK (ANN)

An artificial neuron is a computational model motivated in the common neurons. Common neurons get signals through neurotransmitters situated on the dendrites or film of the neuron. At the point when the signs got are sufficiently solid (surpass a specific limit), the neuron is actuated and radiates a sign however the axon. This sign may be sent to another neurotransmitter, and might enact different neurons.

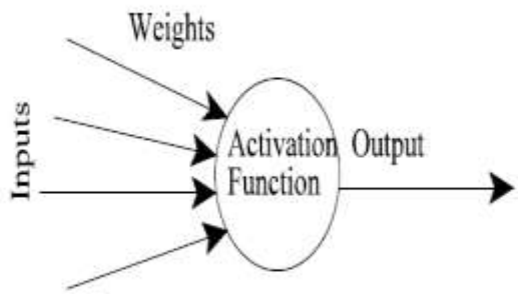

Fig.2. Artificial Neural Network Architecture

The higher a weight of a manufactured neuron is, the more grounded the info which is duplicated by it will be. Weights can likewise be negative, so we can say that the sign is restrained by the negative weight. Be that as it may, when we have an ANN of hundreds or a huge number of neurons, it would be very convoluted to discover by hand all the essential weights. Be that as it may, we can discover calculations which can change the weights of the ANN keeping in mind the end goal to get the wanted yield from the system. This procedure of changing the weights is called learning or preparing.

A normal counterfeit neuron and the showing of a multilayered neural framework are spoken to in Fig.2. Suggesting, the sign stream from inputs x1, x2, x3,… xn is thought to be unidirectional, which are appeared by bolts, and w1, w2,w3 ... wn are connected weights to the looking at inputs. The complete framework of ANN model is represented in the Fig.3.

5.1 Bias Correction

As indicated by Wilby et al(2004), Bias adjustment should be possible as quartile based revisions or re-gridding or by institutionalization of the information. The numerically illuminated essential conditions in GCM contain certain methodical mistakes (known as predisposition), that should be adjusted taking into account the observed information.

Standardization has been performed by subtracting the mean from every worth and isolating by the standard deviation. The markers for the reproduced period and the future period have similarly been organized in light of the standard period. The standard period was taken as a 30-year set which is viewed as adequate to build up the solid climatologically drift.

5.2 Model Training and Validation

For the present work, the model has been aligned for a long time (1971-2005) on a normal for every station. GCM CanCM4 Historical information removed and introduced in exceed expectations group utilizing MATLAB. In a worksheet, the IMD station information for the above notice is organized in agreement and a nonlinear relationship is built up utilizing ANN instrument director. Subsequently, the model is adjusted.

5.3 Evaluation of downscaled model ANN

The execution of the trained ANN model is evaluated utilizing the accompanying statistical measures :

1. Nash Sutcliffe Coefficient

$$E = 1 - \frac{\sum_{i=1}^{n} (X_{obs,i} - X_{model})^2}{\sum_{i=1}^{n} (X_{obs,i} - \overline{X_{obs}})^2}$$

Here,
Xobs observed values and
Xmodel are demonstrated qualities.
The N-S Coefficient ranges between - ∞ to 1.

2. Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_{obs,i} - X_{model,i})^2}{n}}$$

Here,
Xobs observed values and
Xmodel are demonstrated/reproduced values.

6. RESULTS AND DISCUSSION

6.1. Screened Potential Predictor Variables

The most potential predictor variables for the considered station are screened using the four sensitivity analysis methods mentioned in Section 4. The Table.2 represents the most probable predictors screened for the considered study area.

6.2. Evaluation of Downscale Model ANN

The evaluation of downscale model ANN is determined using N-S coefficient and RMSE as explained in Section 6.3. The assessment is done for each sensitivity analysis technique for Station S1 and S2 individually and the result shows the most efficient sensitivity analysis method for theefficient performance of the statistical downscaling model ANN. Hence, the Table 3 shows the results of the evaluation of downscale model.

6.3. Future Projection

The future GCM information is separated from CanCM4_RCP45 and the same standardization process by maxima and minima capacity is executed for the downscaling of future GCM information for the period 2006-2035. The standardized information is utilized as a contribution for ANN model tool stash where the model is prepared and the downscaled qualities are given as a yield. The annual maximum temperature for the time period 2006-2035 for the Malaprabha basin S1 is represented in the Fig.4.

7. CONCLUSION

The primary aim of this study is to analyse the role of predictors in the statistical downscaling surface temperature for the Malaprabha basin using the ANN model. The future projections for the considered station is obtained for the time period from 2006-2035. The performance of model is more efficient by using the predictors screened by Principal component analysis method.

ACKNOWLEDGMENTS

I am greatly thankful to my guide **Dr.Nagraj S. Patil** for his guidance, support, motivation and encouragement His readiness's for consultation at all times, his educative comments, his concern during this period has been invaluable.

I would like to thank **ALMIGHTY** for making us as wearer. In particular, I would like to take this opportunity to express my Honour, Respect, Deepest Gratitude and Genuine regards to all my Professors of Department of Water and Land Management, VTU, Belagavi for giving me all guidance required for my paper apart from being a constant source of inspiration and motivation.

REFERENCES

[1] Aavudai Anandhi,V. V. Srinivas, D. Nagesh Kumar and Ravi S. Nanjundiah.2009. Role of predictors in downscaling surface temperature to river basin in India for IPCC SRES scenarios using support vector machine. *Int. J. Climatol.* 29: 583–603. DOI: 10.1002/joc.1719

[2] Anandhi A, Srinivas VV, Nanjundiah RS, Kumar DN. 2008. Downscaling precipitationto river basin in India for IPCC SRES scenarios using support vector machine. Int. J.Climatol. 28: 401–420, DOI: 10.1002/joc.1529.

[3] Deepashree Raje and P. P. Mujumdar.2011. A comparison of three methods for downscaling daily precipitation in the Punjab region. Hydrol. Process.DOI: 10.1002/hyp.8083.

[4] Gavin C. Cawley, Malcolm Haylock, Stephen R. Dorling,Clare Goodess and Philip D. Jones(2003).Statistical Downscaling with Artifcial Neural Networks ESANN'2003 proceedings - European Symposium onArtificial Neural Networks Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 167-172.

[5] Hugo G. Hidalgo, Michael D. Dettinger, and Daniel R. Cayan.2008. Downscaling With Constructed Analogues: Daily Precipitation And Temperature Fields Over The United States. California Climate Change Centre Report Series Number 2007-027. CEC-500-2007-123

Table 3. Evaluation of downscale model for station S1

Parameter	Value – S1	Method
Nash Sutcliffe Coefficient	0.802	Pearson`s Correlation Coefficient
R^2- Coefficient of determination	0.815	
Nash Sutcliffe Coefficient	0.761	Spearman`s Rank Correlation Coefficient
R^2- Coefficient of determination	0.793	
Nash Sutcliffe Coefficient	0.892	Principal Component Analysis
R^2- Coefficient of determination	0.90	
Nash Sutcliffe Coefficient	0.844	One-way ANOVA
R^2- Coefficient of determination	0.852	

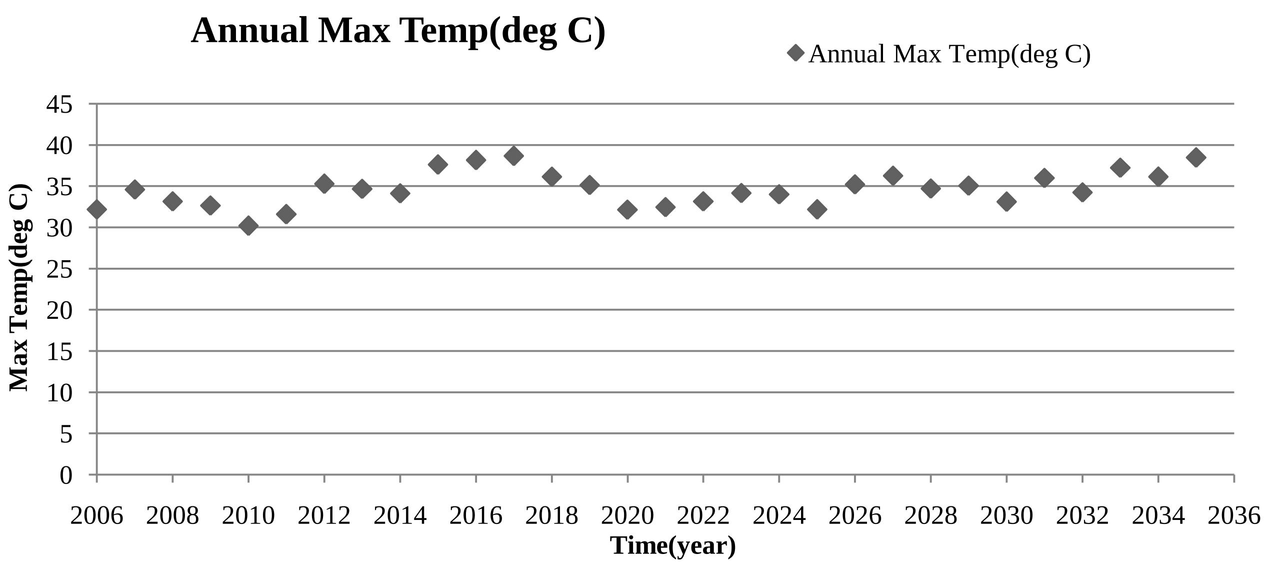

Fig.4. Downscaled Annual Max. Temperature for the period(2006-2035) for Station S1