Statistical Downscaling and Future Projection of Temperatures of Bhuntar, Himachal Pradesh: Intercomparison of Machine Learning, ANN and SDSM

Progress Report

In fulfillment of the requirements for the

NU 302 R&D Project

At NIIT University

Submitted by

Rahul Ghosh, Siddharth Bisht, Hridaya Annuncio,Omkar Vuddanti, C.V. Koushik

Area

NIIT University

Neemrana

Rajasthan

CERTIFICATE

This is to certify that the present research work entitled "Statistical Downscaling and Future Projection of Temperatures of Bhuntar, Himachal Pradesh: Intercomparison of Machine Learning, ANN and SDSM" being submitted to NIIT University, Neemrana, Rajasthan, in the fulfillment of the requirements for the course at NIIT University, Neemrana, embodies authentic and faithful record of original research carried out by Rahul Ghosh, Hridaya Annuncio, Omkar Vuddanti, C.V. Koushik, Siddharth Bisht student/s of B Tech (Artificial Intelligence and Data Science) at NIIT University, Neemrana,. She /He has worked under our supervision and that the matter embodied in this project work has not been submitted, in part or full, as a project report for any course of NIIT University, Neemrana or any other university.

Dr. M. Anul Haq　　　　　　　　　　　　　　　　　　Dr. Nidhi Chahal

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1. Rationale

Climatic models have predicted that the temperature of Earth is going to rise by about 0.2° Celsius in the next two decades due to the increase in greenhouse gases.This increase in temperature will cause the melting of glaciers and permafrost which may consequently lead to hazards. Thus the prediction of future temperatures is important so as to know the movement of glaciers and the spatial distribution of permafrost.GCM(General Circulation model) data contains predicted temperatures of regions based on the concentration of greenhouse gases. It does this at a lower resolution. Hence, downscaling is required so as to predict local temperatures. In this project we have made a comparison amongst five statistical downscaling methods. Our project requires the prediction of temperatures. Therefore, we are using Regression Techniques amongst all Machine Learning methods. ANN(Artificial Neural Networks-A deep learning method) is known for greater efficiency and has thus been used as one of the five methods. SDSM(Statistical Downscaling model) is well known for high efficiency for the process of statistical downscaling and is hence the 5th method used.

2. Literature Review

According to **UCAR**, climate models have predicted that the Earth's global average temperature will increase by about 0.2° Celsius in the next two decades.This has been caused due to the increase in the emissions of greenhouse gases. The increase in greenhouse gases and the consequent rise in temperatures has led to the melting of glaciers and degradation of **permafrost**.
The prediction of these rising temperatures is critical to predict the changes in the retreats of glaciers and the spatial distribution of permafrost.

GCM(General Circulation Model) is a numerical model that represents physical processes in the atmosphere, ocean, cryosphere and land by simulating the impact of increasing concentrations of Greenhouse Gases(GHGs) on the global climate system . The study of the impact of climate change on temperatures rely on the future climate change scenario projected by GCMs[6][16]. In this paper, the GCM data helps to understand the impact of the concentrations of greenhouse gases on the future predicted temperatures of Bhuntar.
The GCMs are run at coarse grid resolution and can thus not directly be used to predict local temperatures. Here, we used the Hadley Couple Model version 3(HADcm3) which runs

at a resolution of 2.5 degree latitude and 3.75 degree longitude.

This leads to the use of a technique called downscaling. This helps connect the gap between the results given by the GCM and its impact on local temperatures.
According to **IPCC**," downscaling is a method that derives local-to-regional-scale(10 - 100 km) information from large-scale models or data analysis" .
Downscaling is broadly divided into two categories which are dynamic downscaling and statistical downscaling.
In dynamic downscaling a Regional Climate model (RCM) is used. RCM is an atmospheric, physics based model to which "boundary conditions" are given with the output of a GCM . While this is an efficient model, it has two major drawbacks. First being that it is a highly complex model and the second being that it has a high computation cost.

Statistical downscaling is a technique that aims to learn a statistical relationship between coarse scale climate variables and high resolution observations . There are 3 approaches to statistical downscaling.
Weather generators are used for temporal downscaling rather than spatial. Weather typing searches for similar historical coarse resolution climate state that closely represents the current state. Transfer functions or regression methods are used by learning the functional relationship between historical precipitation and climate variables to project high resolution precipitation .
The third method is what we have used as our approach to statistical downscaling.

The Statistical Downscaling Model (SDSM) is a tool based on Multiple Linear Regression. It is used to generate future scenarios so as to be able to predict the impact of climate change in the future.This method can capture "inter-annual variability" better than other statistical downscaling methods like weather generators.

NCEP(National Centers of Environmental Predictions) data provides large-scale atmospheric variables' (predictors) re-analysed datasets which have horizontal resolution 2.5 degree by 2.5 degree for the entire world. NCEP has been providing the data of 26 parameters.

The NCEP data is required to be trained with the local temperatures of Bhuntar so as to be able to predict Bhuntar temperatures given the NCEP parameters' values.

3. Objectives

a. To predict the future temperatures of Bhuntar using Machine Learning(Regression), Artificial neural networks and Statistical Downscaling Model.

b. To compare the accuracy of projection of various techniques by computing the R^2 value and Mean Square Error.

4. Methodology

4.1 Machine Learning Methods

4.1.1 Simple Linear Regression

Here from the 26 parameters of NCEP we have just chosen one, i.e., temperature.

The dataset has been divided into the training set and test set with a 80:20 ratio. Here we have used the tensorflow library.
Only 1 hidden layer is used as the layers are fully connected.

Hypothesis Function:

$$h_\theta(x) = \theta_0 + \theta_1 x$$

Cost Function:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Gradient Descent:

$$\begin{aligned} &\text{repeat until convergence } \{ \\ &\quad \theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \\ &\qquad (\text{for } j = 1 \text{ and } j = 0) \\ &\} \end{aligned}$$

The model details: [8]

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 64)                128
_________________________________________________________________
dense_1 (Dense)              (None, 64)                4160
_________________________________________________________________
dense_2 (Dense)              (None, 1)                 65
=================================================================
Total params: 4,353
Trainable params: 4,353
Non-trainable params:
_________________________________________________________________
```

Table 1: Description of the model used

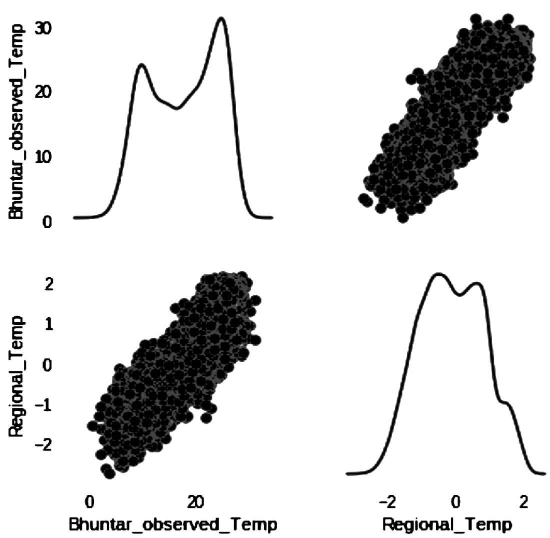
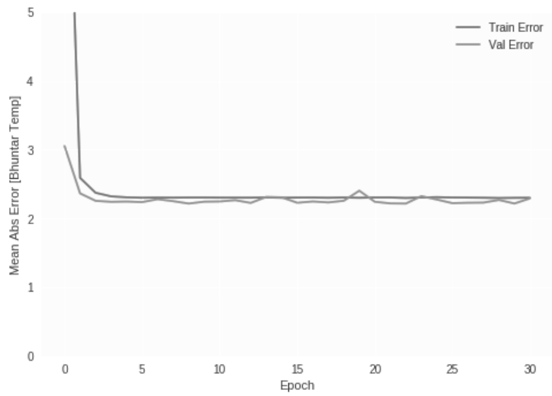


Fig 2: Correlation between NCEP temperature variable and local Bhuntar data.



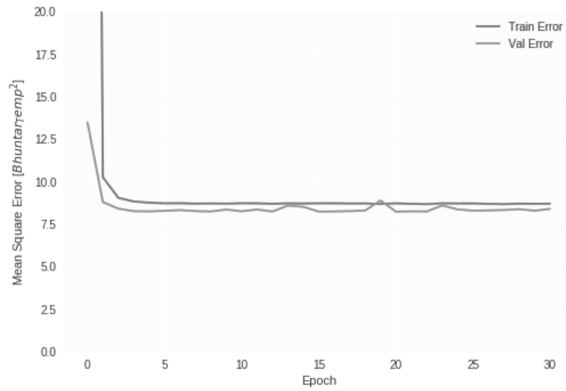
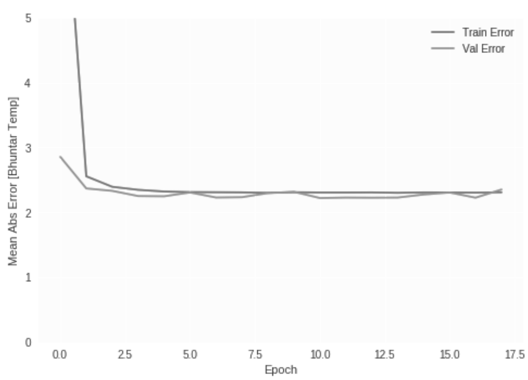


Fig 3.1 **Fig 3.2**

Training of model for 1000 epochs
Fig 3.1: Change in mean absolute error with increasing number of epochs, Fig 3.2: Change in mean square error with increasing number of epochs



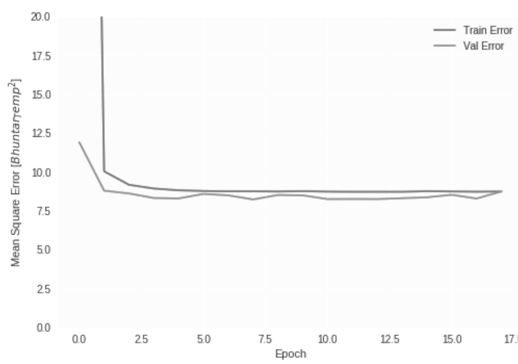


Fig 4.1 **Fig 4.2**

The patience parameter(the amount of epochs to check for improvement) being analysed.
Fig 4.1: Change in mean absolute error with increasing number of epochs, Fig 4.2: Change in mean square error with increasing number of epochs

4.1.2 Multiple Linear Regression

Here along with temperature, we chose five more parameters of the 26 parameters of NCEP: fas,vas, 500as, 5thas and ncephumas. These variables were chosen as they had the highest correlations with the local Bhuntar temperature.[11]

Hypothesis Function:

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

Cost Function:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Gradient Descent:

$$\begin{aligned}&\text{repeat until convergence } \{\\ &\quad \theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)\\ &\qquad \text{(for } j = 1 \text{ and } j = 0)\\ &\}\end{aligned}$$

The model details:[8]

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 64)	448
dense_1 (Dense)	(None, 64)	4160
dense_2 (Dense)	(None, 1)	65

Total params: 4,673
Trainable params: 4,673
Non-trainable params: 0

Table 2: Description of model used

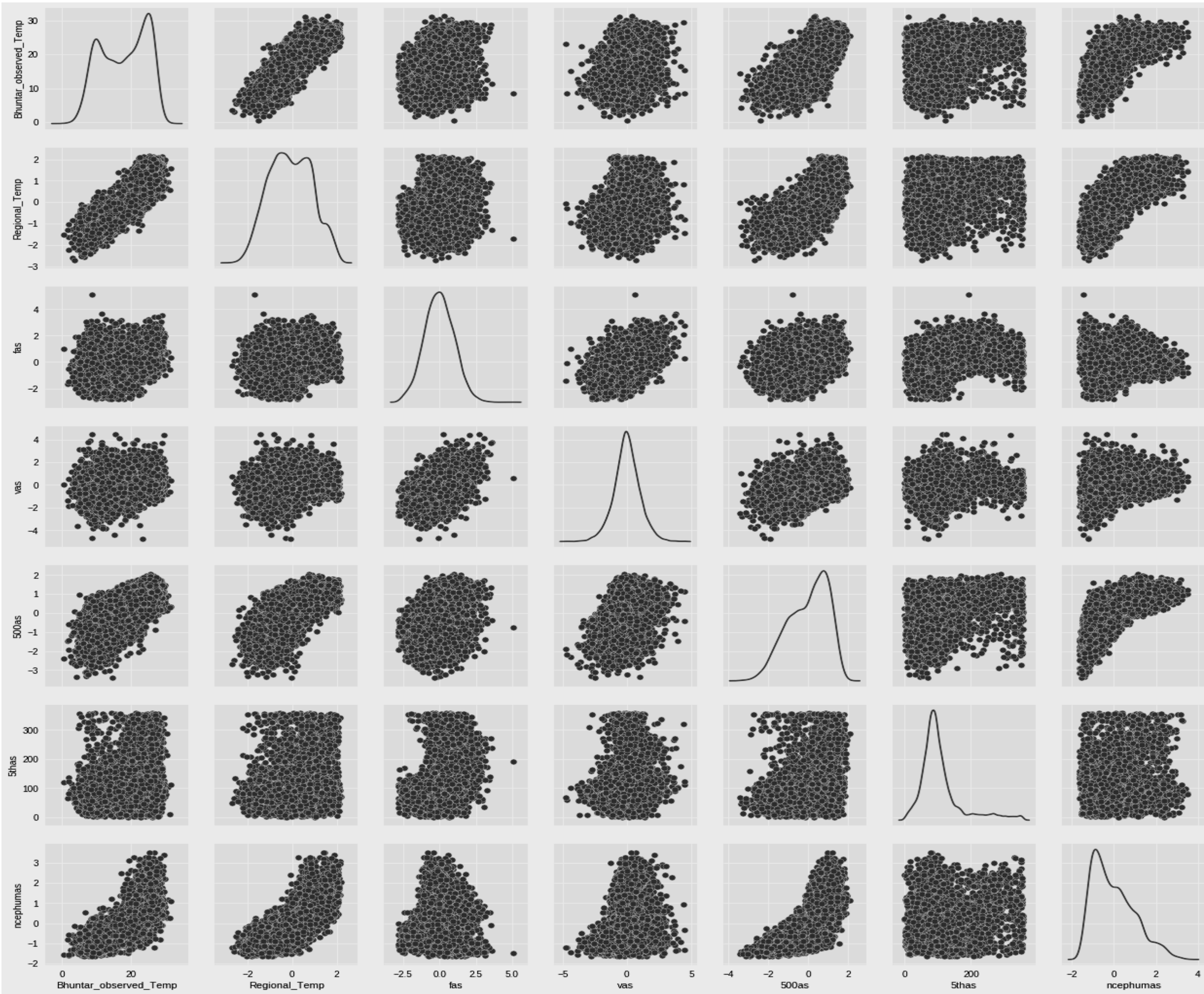


Fig 5: Correlation between all 6 chosen NCEP variables and Bhuntar local data

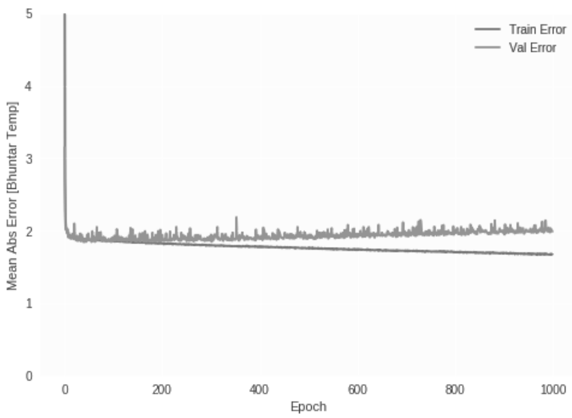


Fig 6.1

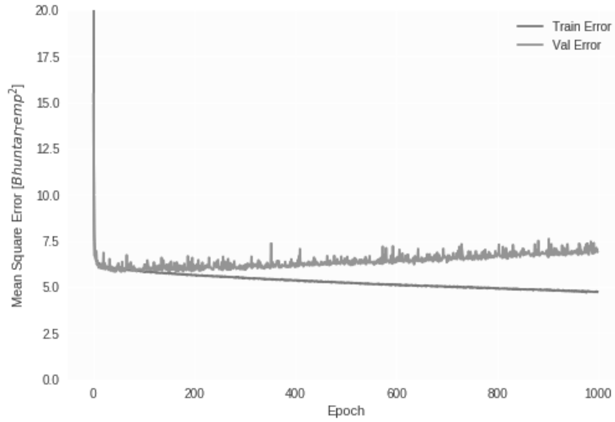


Fig 6.2

Training of model for 1000 epochs.
Fig 6.1: Change in mean absolute error with increasing number of epochs, Fig 6.2: Change in mean square error with increasing number of epochs

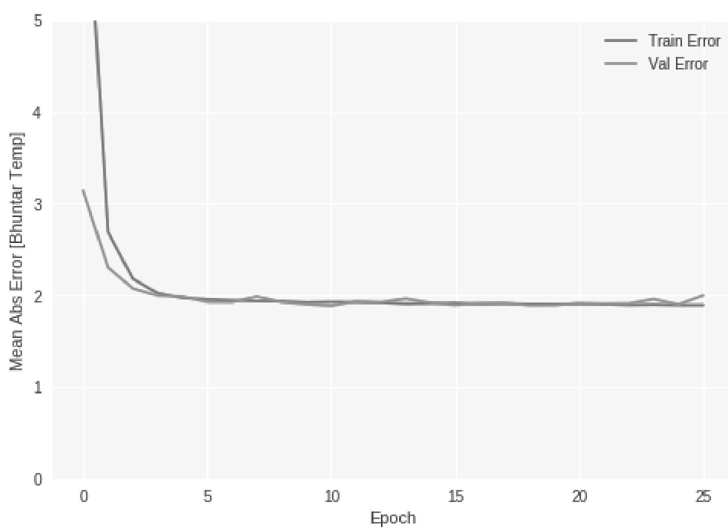


Fig 7: Change in mean absolute error on decreasing the number of epochs required

4.2 Non-Linear Regression

This is a form of regression that provides the best fitting curve for the relation between the sets of data.

The highest correlation between the temperatures of Bhuntar was found with the six NCEP variables tempas,fas,vas, 500as, 5thas and humas. Hence they were used for the study. XLSTAT was the software used for this purpose. Models with highest values of power of 2,3,5 and 7 were used.

The basic equation was:

Function: $Y = p_0 + p_{11}(x_1)^1 + p_{12}(x_1)^2 + p_{13}(x_1)^3 + p_{14}(x_1)^4 + \ldots\ldots. + p_{1m}(x_1)^m + p_{21}(x_2)^1 + p_{22}(x_2)^2 + p_{23}(x_2)^3 + p_{24}(x_2)^4 + p_{25}(x_2)^5 + \ldots\ldots.. + p_{n1}(x_n)^1 + p_{n2}(x_n)^2 + p_{n3}(x_n)^3 + p_{n4}(x_n)^4 + \ldots\ldots. + p_{nm}(x_n)^m$

This pattern continues till the power specified by the model is reached.
Here **p** means **parameter**. x_i represents one variable of NCEP.
x_1 : ncepshumas
x_2 : ncepp__fas
x_3 : ncepp500as
x_4 : ncepp5_vas
x_5 : ncepp5thas
x_6 : tempas

NLR 2:

Function: $Y_2 = p_0 + p_{11}(x_1)^1 + p_{12}(x_1)^2 + p_{21}(x_2)^1 + p_{22}(x_2)^2 + p_{31}(x_3)^1 + p_{32}(x_3)^2 + p_{41}(x_4)^1 + p_{42}(x_4)^2 + p_{51}(x_5)^1 + p_{52}(x_5)^2 + p_{61}(x_6)^1 + p_{62}(x_6)^2$

NLR 3:

Function: $Y_3 = p_0 + p_{11}(x_1)^1 + p_{12}(x_1)^2 + p_{13}(x_1)^3 + p_{21}(x_2)^1 + p_{22}(x_2)^2 + p_{23}(x_2)^3 + p_{31}(x_3)^1 + p_{32}(x_3)^2 + p_{33}(x_3)^3 + p_{41}(x_4)^1 + p_{42}(x_4)^2 + p_{43}(x_4)^3 + p_{51}(x_5)^1 + p_{52}(x_5)^2 + p_{53}(x_5)^3 + p_{61}(x_6)^1 + p_{62}(x_6)^2 + p_{63}(x_6)^3$

NLR 5:

Function: $Y_5 = p_0 + p_{11}(x_1)^1 + p_{12}(x_1)^2 + p_{13}(x_1)^3 + p_{14}(x_1)^4 + p_{15}(x_1)^5 + p_{21}(x_2)^1 + p_{22}(x_2)^2 + p_{23}(x_2)^3 + p_{24}(x_2)^4 + p_{25}(x_2)^5 + p_{31}(x_3)^1 + p_{32}(x_3)^2 + p_{33}(x_3)^3 + p_{34}(x_3)^4 + p_{35}(x_3)^5 + p_{41}(x_4)^1 + p_{42}(x_4)^2 + p_{43}(x_4)^3 + p_{44}(x_4)^4 + p_{45}(x_4)^5 + p_{51}(x_5)^1 + p_{52}(x_5)^2 + p_{53}(x_5)^3 + p_{54}(x_5)^4 + p_{55}(x_5)^5 + p_{61}(x_6)^1 + p_{62}(x_6)^2 + p_{63}(x_6)^3 + p_{64}(x_6)^4 + p_{65}(x_6)^5$

NLR 7:

$$\textbf{\textit{Function: }} Y_7 = p_0 + p_{11}(x_1)^1 + p_{12}(x_1)^2 + p_{13}(x_1)^3 + p_{14}(x_1)^4 + p_{15}(x_1)^5 + p_{16}(x_1)^6 + p_{17}(x_1)^7 + p_{21}(x_2)^1 + p_{22}(x_2)^2 + p_{23}(x_2)^3 + p_{24}(x_2)^4 + p_{25}(x_2)^5 + p_{26}(x_2)^6 + p_{27}(x_2)^7 + p_{31}(x_3)^1 + p_{32}(x_3)^2 + p_{33}(x_3)^3 + p_{34}(x_3)^4 + p_{35}(x_3)^5 + p_{36}(x_3)^6 + p_{37}(x_3)^7 + p_{41}(x_4)^1 + p_{42}(x_4)^2 + p_{43}(x_4)^3 + p_{44}(x_4)^4 + p_{45}(x_4)^5 + p_{46}(x_4)^6 + p_{47}(x_4)^7 + p_{51}(x_5)^1 + p_{52}(x_5)^2 + p_{53}(x_5)^3 + p_{54}(x_5)^4 + p_{55}(x_5)^5 + p_{56}(x_5)^6 + p_{57}(x_5)^7 + p_{61}(x_6)^1 + p_{62}(x_6)^2 + p_{63}(x_6)^3 + p_{64}(x_6)^4 + p_{65}(x_6)^5 + p_{66}(x_6)^6 + p_{67}(x_6)^7$$

4.3 ANN (Artificial Neural Networks)

ANNs are known to provide much better results for prediction than machine learning methods. The structure of an ANN is such that it has layers called hidden layers between the input and the output. Each layer has a set of neurons. The output of the input layer acts as input to the first hidden layer. This process continues until the output layer is reached. The edges connecting neurons of different layers are assigned different weights. The different features are given as inputs and not all are required by each neuron.

This is what makes ANN more efficient.

In our model we are training all 6 variables of NCEP i.e. ncepshumas, ncepp__fas, ncepp500as, ncepp5_vas, ncep5thas and tempas.

This model has 3 hidden layers with 6 neurons each and ‘ReLu’ Activation function applied to it. ReLu (Rectified Linear Unit) is the most common activation function used in deep learning.

Function:

R(z)=max(0,z)

This formula says that all negative values are given the value of 0.

4.4 SDSM (Statistical Downscaling Model)

This model requires the local observed data of Bhuntar (predictand) and larger scale data of different atmospheric variables(NCEP and GCM). This tool is used to find a statistical relation between the above two types of data.[9]

The atmospheric variables have been chosen for the study based on their correlation with the observed local Bhuntar data. SDSM produces results for two future scenarios i.e. HadCM3 A2 and HadCM3 B2.

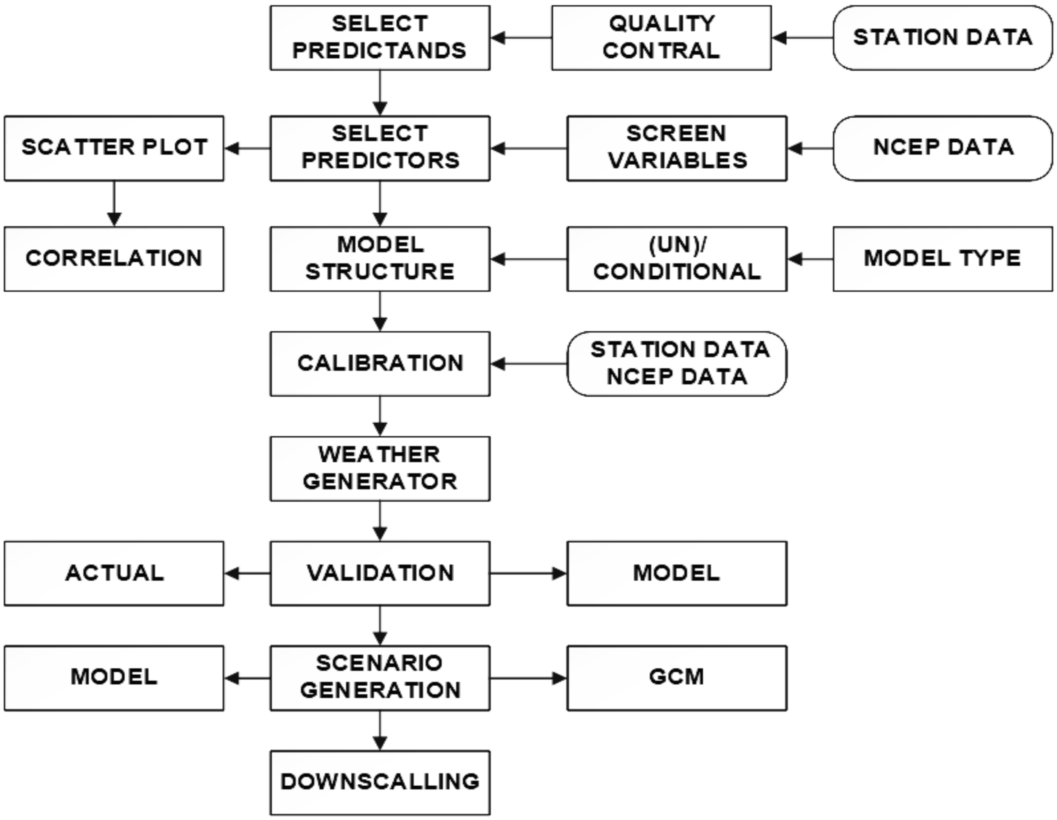


Fig 8: SDSM scenario generation of HadCM3

5. Results

5.1 Simple Linear Regression

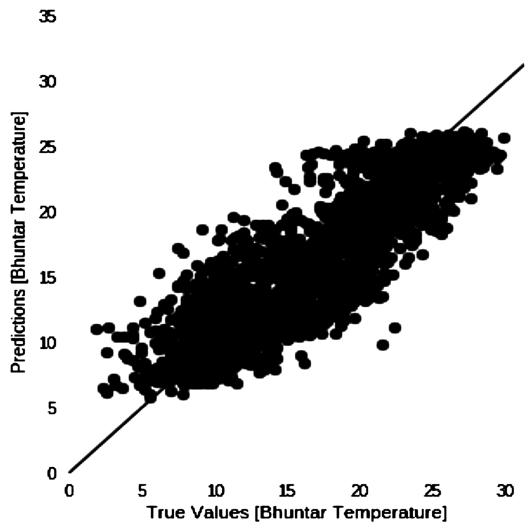


Fig 9: Correlation between Predicted values and actual values of Bhuntar temperature

The correlation between the predicted values and the true values was checked :
R^2 value : 0.7607966
Testing set Mean Absolute Error : 2.37
Testing set Loss Error : 8.90

5.2 Multiple Linear Regression

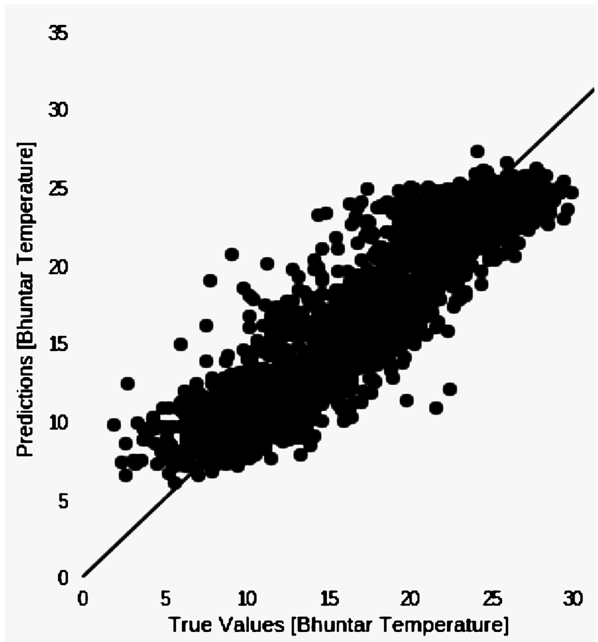


Fig 10: Correlation between Predicted values and actual values of Bhuntar temperature

R^2 value : 0.82
Testing set Mean Absolute Error : 2.00
Testing set Loss Error : 6.36

5.3 Non Linear Regression

NLR2:

Observations	11323.000
DF	11310.000
R^2	0.850
SSE	73977.797
MSE	6.541
RMSE	2.558

Table 3: Results of correlation

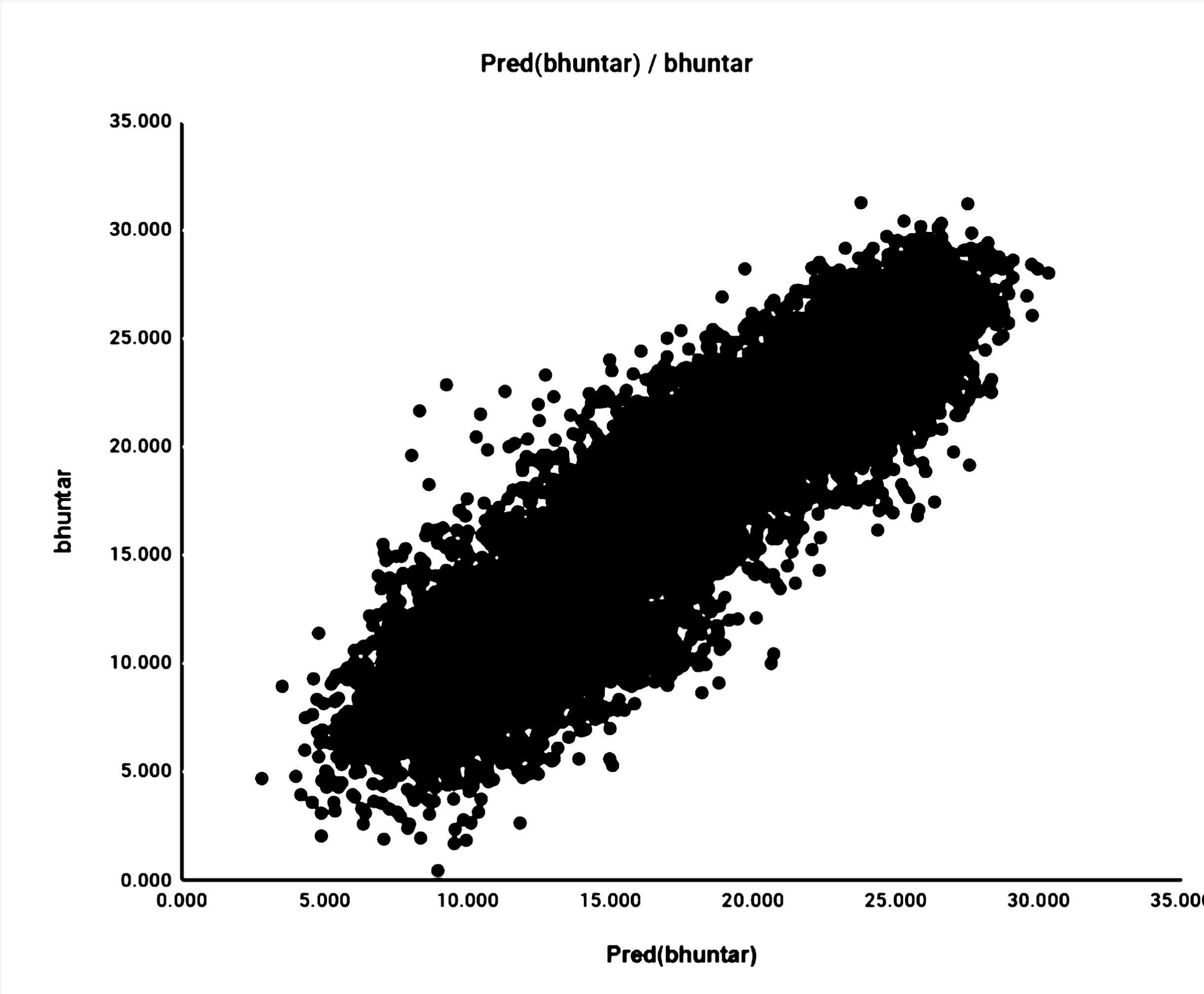


Fig 11 : Correlation of actual and predicted values of temperatures of Bhuntar

NLR3:

Observations	11323.000
DF	11304.000
R^2	0.855
SSE	71583.864
MSE	6.333
RMSE	2.516

Table 4: Results of correlation

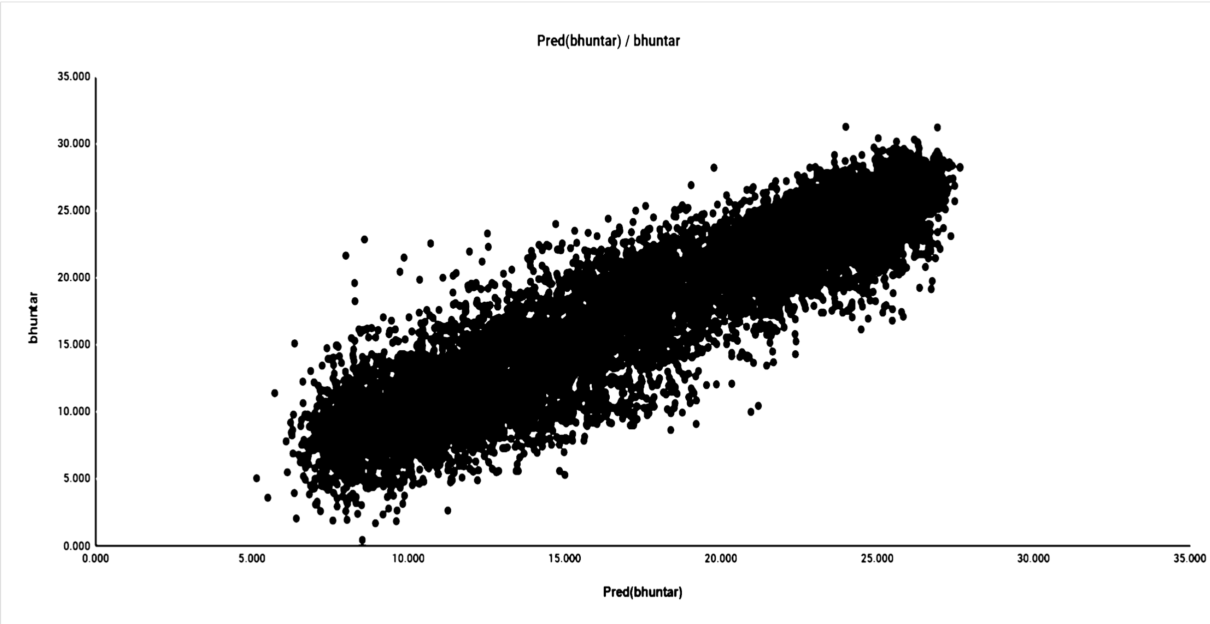


Fig 12 : Correlation of actual and predicted values of temperatures of Bhuntar

NLR5:

Observations	11323.000
DF	11292.000
R^2	0.857
SSE	70788.339
MSE	6.269
RMSE	2.504

Table 5: Results of correlation

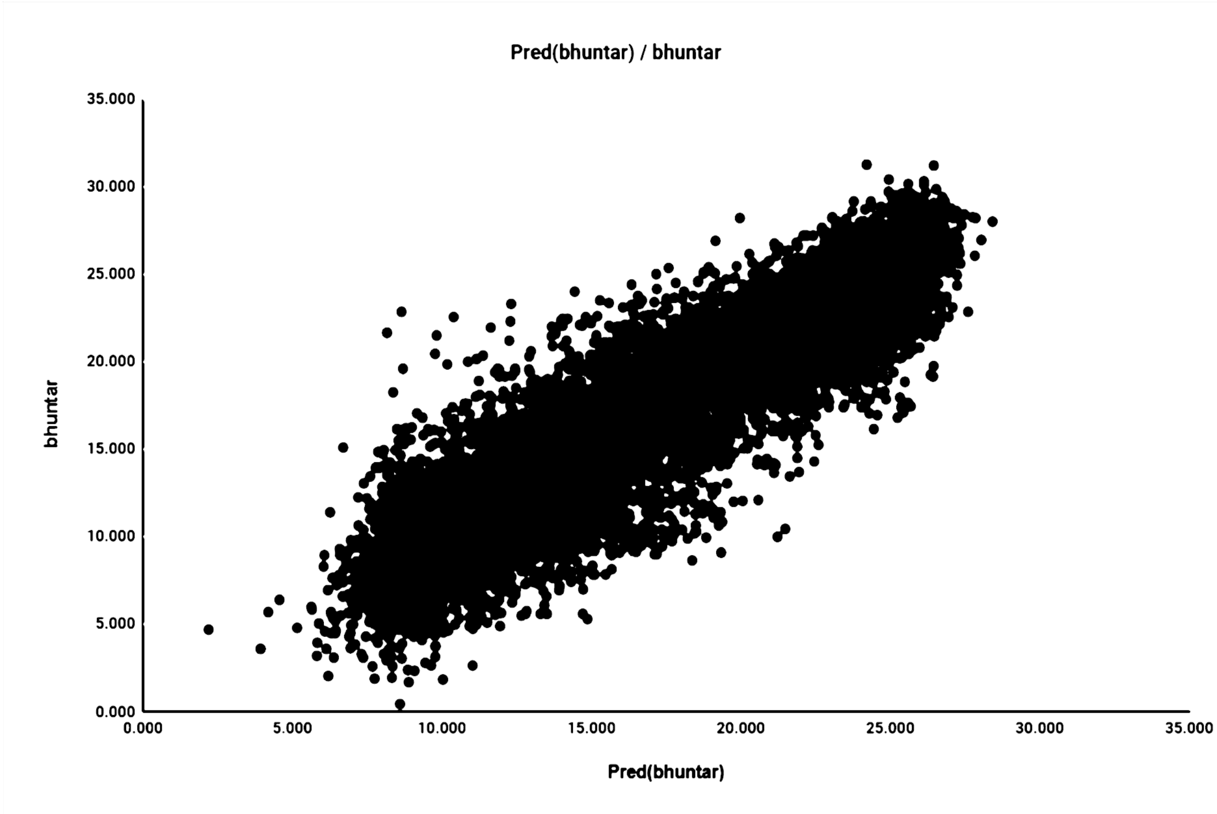


Fig 13 : Correlation of actual and predicted values of temperatures of Bhuntar

NLR7:

Observations	11323.000
DF	11280.000
R²	0.857
SSE	70592.884
MSE	6.258
RMSE	2.502

Table 6: Results of correlation

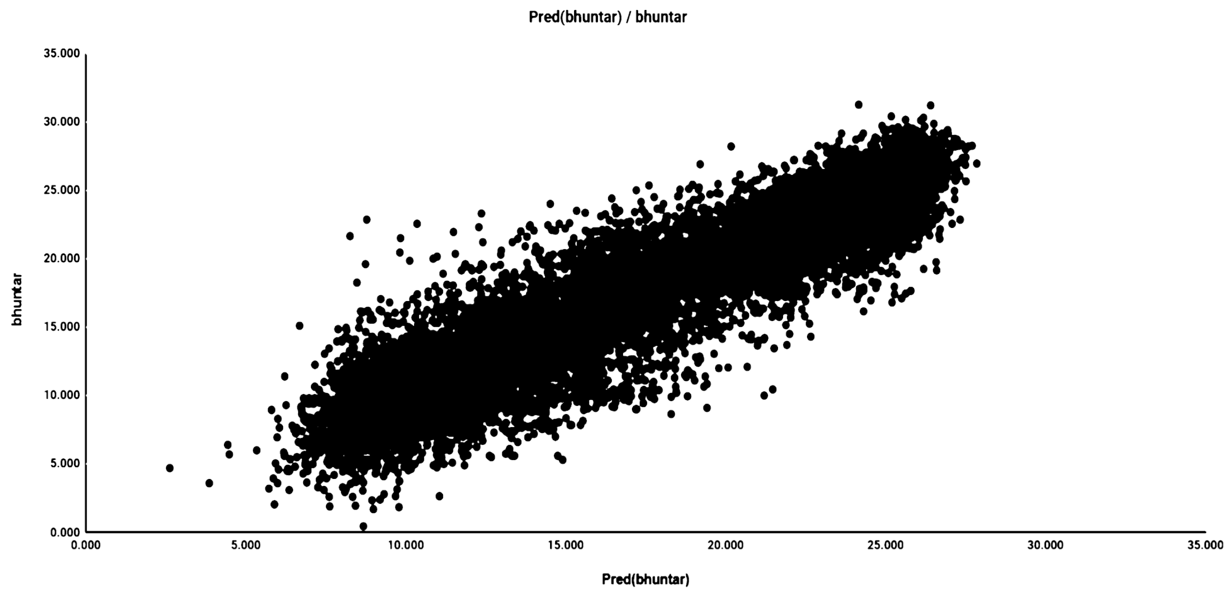


Fig 14 : Correlation of actual and predicted values of temperatures of Bhuntar

Discussion:

The R² value increases as the maximum power of the model increases. In NLR5 and NLR7 the result is the maximum and the same. Hence we can conclude and say that 5 is the highest power required to get the best correlation between the predicted and actual values of temperatures of Bhuntar.

5.4 ANN

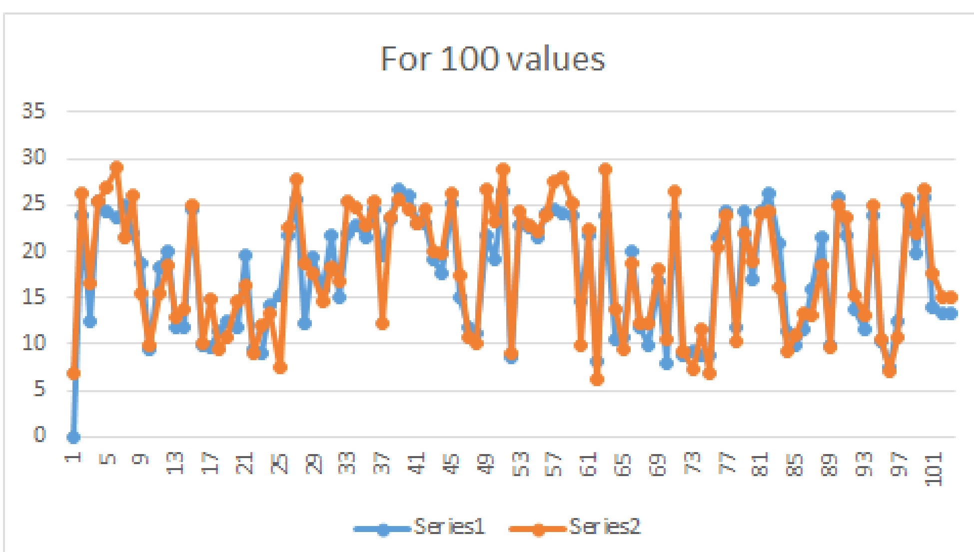


Fig 15 : Comparison between actual and predicted values of temperatures of Bhuntar using ANN,

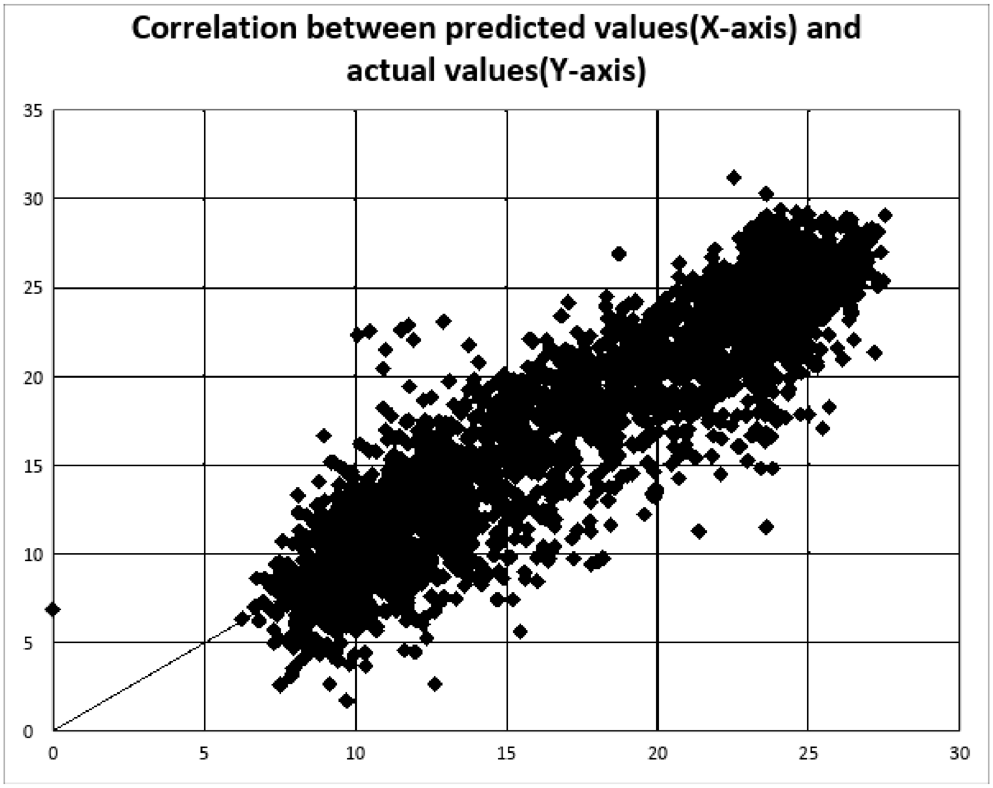
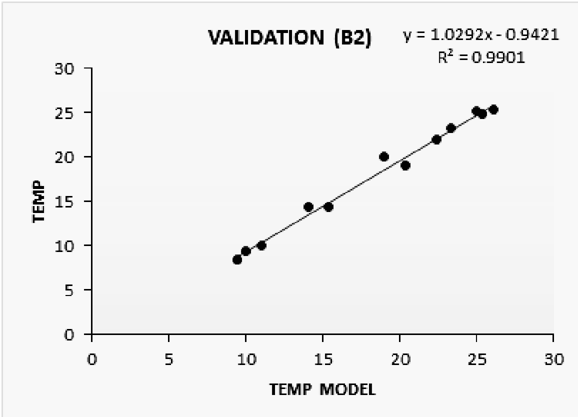


Fig 16: Correlation between actual and predicted values of Bhuntar using ANN

R^2 value: 0.84

5.5 SDSM



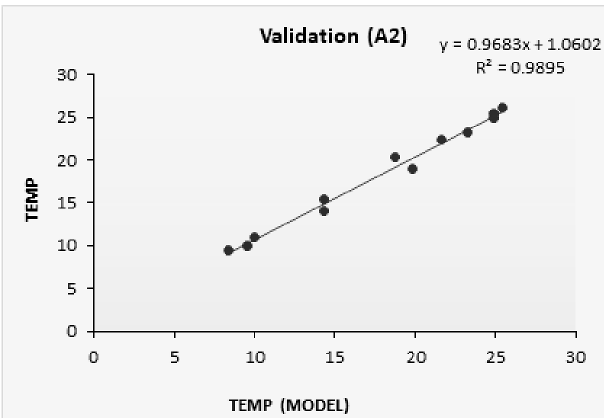


Fig 17.1 **Fig 17.2**

Fig 17.1: Average annual predicted temperatures of Bhuntar plotted with actual temperatures in B2 scenario, Fig 15.7: Average annual predicted temperatures of Bhuntar plotted with actual temperatures in A2 scenario

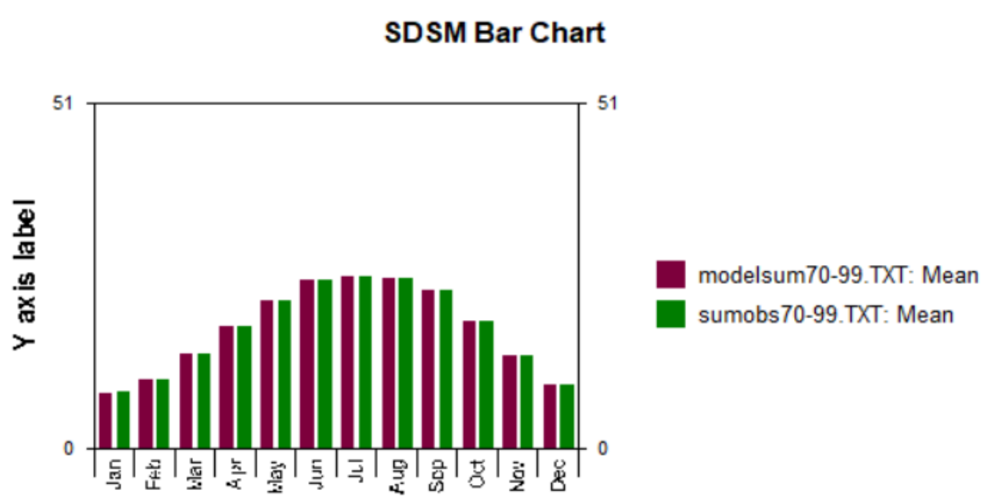


Fig 18: Bar graph comparison of predicted and actual Bhuntar data

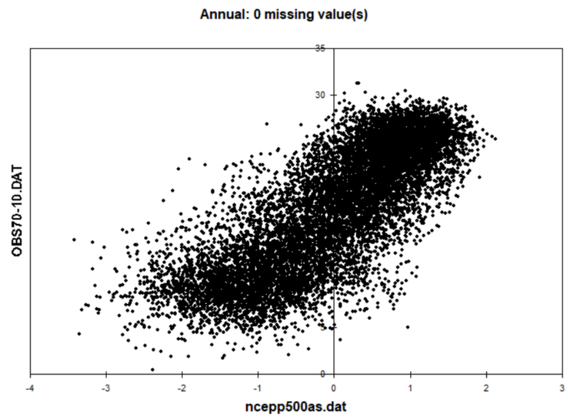


Fig 19.1

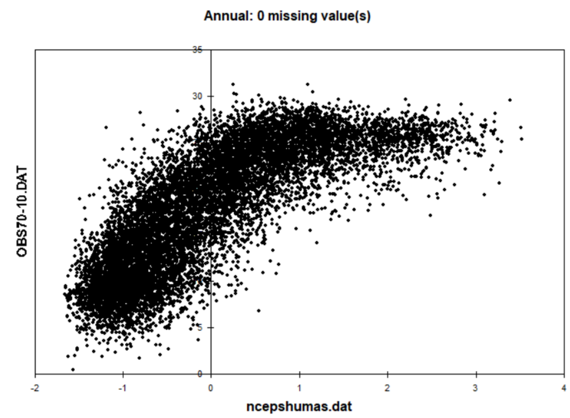


Fig 19.2

Fig 19.1: Scatter plot description of correlation of ncepp500as and Bhuntar temperature,
Fig 19.2: Scatter plot description of correlation of ncepshumas and Bhuntar temperature

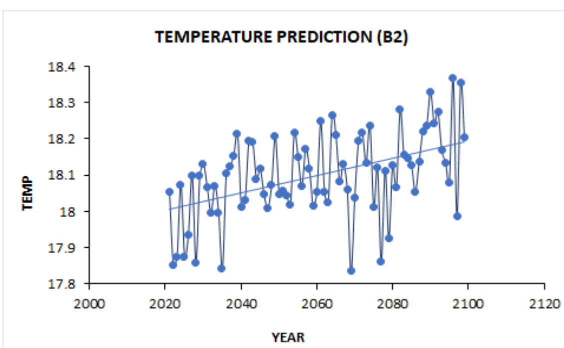


Fig 20.1

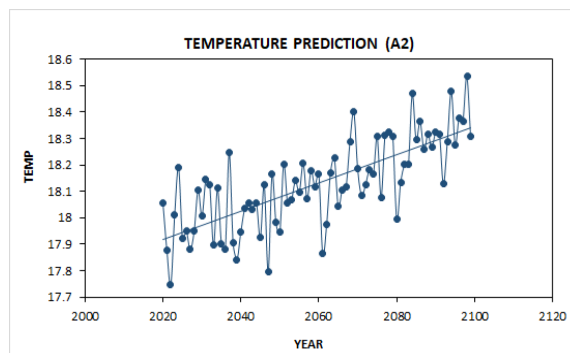


Fig 20.2

Fig 20.1: A2 scenario of HadCM model, Fig 20.2: B2 scenario of HadCM model

Mann Kendall’s test for checking trends(Yearly):

Value of significance level taken as 5%, thus value of alpha is 0.05.

H0: No trend in the series

Ha: There is a trend in the series

A2 scenario:

Month	Variable	Observations	Min.	Max.	Mean	Std. deviation	Kendall's tau	S	Var(S	p-value (Two-tailed)

January	8.57	79	7.85	8.63	8.32	0.160	0.111	341.0	55800.3	0.150
February	10.19	79	9.68	10.43	10.09	0.176	0.194	597.0	55800.3	0.012
March	14.50	79	13.98	14.97	14.44	0.177	0.199	613.0	55800.3	0.010
April	19.01	79	17.78	20.70	19.14	0.453	0.224	689.0	55800.3	0.004
May	22.61	79	20.24	22.69	21.79	0.540	0.087	267.0	55800.3	0.260
June	25.45	79	24.17	26.10	25.25	0.299	0.316	975.0	55800.3	< 0.0001
July	25.29	79	25.06	25.75	25.35	0.130	-0.031	-95.0	55800.3	0.691
August	24.74	79	24.49	25.17	24.81	0.136	-0.307	-947.0	55800.3	< 0.0001
September	23.31	79	22.95	24.37	23.64	0.323	0.547	1685.0	55800.3	< 0.0001
October	19.41	79	19.24	20.96	20.14	0.450	0.399	1229.0	55800.3	< 0.0001
November	13.77	79	13.17	16.56	14.85	0.817	0.474	1461.0	55800.3	< 0.0001
December	9.81	79	9.44	10.17	9.73	0.181	0.462	1423.0	55800.3	< 0.0001
Yearly	18.06	77	17.74	18.54	18.13	0.170	0.557	1630.0	55800.3	< 0.0001

Table 7: Results of Mann Kendall's test in A2 scenario

We can see that the p-value calculated for the yearly scenario is less than the value of alpha.

Thus here the null hypothesis **H0**, is rejected and the alternative hypothesis **Ha**, is accepted.

Similarly, when checking trends on monthly basis, leaving January, May and July, the p-values of all other months are less than the value of alpha. Hence in each, Ha is accepted.

Amongst the months wherein Ha has been accepted, only August has a decreasing trend while the other eight months have an increasing trend.

Based on the monthly and yearly trend observations we can conclude that there is an increase in the local temperatures through the years. This hence proves that the concentration of Greenhouse gases is impacting the temperatures of Bhuntar.

B2 scenario:

Month	Variable	Observations	Min.	Max.	Mean	Std. deviation	Kendall's tau	S	Var(S	p-value (Two-tailed)
January	8.31	78	7.92	8.67	8.31	0.153	0.087	261.0	53720.3	0.262
February	10.13	78	9.74	10.53	10.09	0.137	-0.034	-101.0	53720.3	0.666

March	14.27	78	14.04	14.82	14.44	0.172	-0.030	-91.0	53720.3	0.698
April	19.14	78	18.38	20.18	19.09	0.353	0.112	337.0	53720.3	0.147
May	21.57	78	20.66	22.84	21.95	0.534	0.034	101.0	53720.3	0.666
June	24.98	78	24.49	25.98	25.28	0.302	0.358	1075.0	53720.3	< 0.0001
July	25.27	78	25.11	25.65	25.37	0.141	0.015	45.0	53720.3	0.849
August	25.05	78	24.60	25.20	24.83	0.132	-0.233	-699.0	53720.3	0.003
September	23.51	78	22.86	24.06	23.45	0.231	0.341	1023.0	53720.3	< 0.0001
October	20.31	78	19.16	21.03	20.04	0.373	0.151	453.0	53720.3	0.051
November	14.87	78	13.01	16.30	14.68	0.644	0.193	581.0	53720.3	0.012
December	9.27	78	9.23	10.16	9.67	0.163	0.301	903.0	53720.3	< 0.0001
Yearly	18.05	78	17.84	18.37	18.10	0.119	0.321	963.0	53720.3	< 0.0001

Table 7: Results of Mann Kendall's test in B2 scenario

We can see that the p-value calculated for the yearly scenario is less than the value of alpha.
Thus here the null hypothesis **H0**, is rejected and the alternative hypothesis **Ha**, is accepted.
Similarly, when checking trends on monthly basis, the p-values of the months June, August, September, October, November and December are less than the value of alpha. Hence in each Ha is accepted. In the others there is no trend present.Out of the months that accept Ha, only August has a decreasing trend whereas the other five months have an increasing trend.
Based on some monthly observations and the yearly observations we can conclude that an increasing trend is seen in the temperature values of Bhuntar. This proves that there is an impact of the increasing concentrations of greenhouse gases on the temperatures of Bhuntar.

Conclusion

Amongst all three machine learning methods, non-linear regression has outperformed simple linear regression(R^2 =0.76) and multiple linear regression(R^2 =0.853) with an R^2 value of 0.857.Non Linear regression performed better than ANN too which had an R^2 value of 0.84. These values of R^2 represent the correlation between the predicted and actual temperature values of Bhuntar. The limitation of machine learning is that previous training cannot be used to project future temperatures of Bhuntar without the input of future NCEP data.
SDSM has outperformed non-linear regression with an R^2 value of 0.9901 in B2 scenario and 0.9895 in A2 scenario. Also, it can predict future temperatures of Bhuntar, based on

previous training, using GCM values. The results obtained from SDSM have also shown an overall increase in the future predicted temperature values of Bhuntar in both A2 and B2 scenarios. Hence, proving that the concentrations of greenhouse gases is affecting the climate (temperatures).

The above given pointers help conclude that SDSM has proven to be the most efficient model in every aspect.

6. Summary

Climatic models have predicted that the temperature of Earth is going to rise by about 0.2° Celsius in the next two decades due to the increase in greenhouse gases.This increase in temperature will cause the melting of glaciers and permafrost which may consequently lead to hazards. Thus the prediction of future temperatures is important so as to know the movement of glaciers and the spatial distribution of permafrost.

GCM(General Circulation model) data contains predicted temperatures of regions based on the concentration of greenhouse gases. We use statistical downscaling so project these temperatures at a higher resolution.

This paper describes different methods of statistical downscaling used to downscale the temperatures of the city of Bhuntar in Himachal Pradesh. We have used NCEP data also. The Statistical downscaling has been done in three main ways. First,three machine learning methods have been used. Simple Linear Regression just uses Temperature as a predictor and Multiple Linear regression uses the NCEP parameters fas,vas, 500as, 5thas and humas too. Non-linear regression was performed in XLSTAT using the same six variables. Amongst the machine learning methods, non-linear regression got the best correlation with an R^2 of 0.857. Second, ANN(Artificial Neural Networks), a deep learning method adopted, gave an R^2 value of 0.84.Third, SDSM was used, which is a tool that is based on Multiple Linear Regression for prediction . It gave an R^2 value of 0.9895 in the A2 scenario and 0.9901 in B2 scenario.This tool gives the highest value of correlation amongst all the methods performed. SDSM has also predicted future values of Bhuntar based on given GCM values. Hence,SDSM has proven to be most efficient.

7. Future Work

We have achieved significant R^2 values for Non-linear Regression, and Artificial Neural Networks but SDSM continues to be the most efficient of them all. One reason for this is its modelling and scenario generation. Apart from SDSM, WRF, the Weather Research and Forecast Model is another complex model that is also being used but for Dynamic Downscaling. Unlike SDSM which only has 26 parameters, WRF supports more than a hundred parameters for its prediction. We plan to integrate our results with WRF and research further on temperature downscaling to devise new and efficient Machine-Learning based and Deep-Learning based models..

8. References

[1] UCAR's Website: https://scied.ucar.edu/longcontent/predictions-future-global-climate
https://ral.ucar.edu/solutions/products\/weather-research-and-forecasting-model-wrf

[2] IPCC's website: https://www.ipcc-data.org/guidelines/pages/gcm_guide.html

[3] Vidya R. Saraf, Dattatray G. Regulwar (d 13 January 2016) Assessment of Climate Change for Precipitation and Temperature Using Statistical Downscaling Methods in Upper Godavari River Basin, India

[4] Thomas Vandal∗1, Evan Kodra†2, and Auroop R Ganguly‡1(February 15 2017) Intercomparison of Machine Learning Methods for Statistical Downscaling: The Case of Daily and Extreme Precipitation

[5] https://www.mapsofindia.com/maps/himachalpradesh/districts/kullu.htm

[6] Application of Statistical Downscaling Model for Prediction of Future Rainfall in Bhudhabalanga River Basin, Odisha (India)
Satyapriya Behera, Deepak Khare, Prabhash Kumar Mishra, Sangitarani Sahoo

[7] DOWNSCALING OF CLIMATIC DATA USING STATISTICAL TECHNIQUES IN COASTAL SAURASHTRA, GUJARAT Vikrant Vijay Singh1, Anupma Sharma2 and P.C. Joshi

[8] Tensorflow documentation: https://www.tensorflow.org/tutorials/keras/basic_regression

[9] Statistical downscaling of daily mean temperature, pan evaporation and precipitation for climate change scenarios in Haihe River, China by J. T. Chu & J. Xia & C.-Y. Xu & V. P. Sing

[10] Downscaling of Future Temperature and Precipitation Extremes in Addis Ababa under Climate Chang by Getnet Feyissa 1,* ID , Gete Zeleke 1, Woldeamlak Bewket 2 and Ephrem Gebremariam 1 ID

[11] MLR Based Statistical Downscaling of Temperature and Precipitation in Lidder Basin Region of India YasirAltaf1*, Ahanger Manzoor Ahmad1 and Fahimuddin Mohd2

[12] Assessment of Climate Change for Precipitation and Temperature Using Statistical Downscaling Methods in Upper Godavari River Basin, India by Vidya R. Saraf1,2, Dattatray G. Regulwar1

[13] Downscaling the Impacts of Large-Scale LUCC on Surface Temperature along with IPCC RCPs: A Global Perspective Xiangzheng Deng 1,2,*, Chunhong Zhao 1,2,3, Yingzhi Lin 4, Tao Zhang 1,2, Yi Qu 1,2,3, Fan Zhang 4, Zhan Wang 1,2 and Feng Wu 5 1

[14] Statistical Downscaling of Daily Temperature in Central Europe by RADAN HUTH from Institute of Atmospheric Physics, Prague, Czech Republic

[15] Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine by Aavudai Anandhi,a V. V. Srinivas,a Ravi S. Nanjundiahb and D. Nagesh Kumara*
[16] SVM-PGSL coupled approach for statistical downscaling to predict rainfall from GCM output Subimal Ghosh1 Received 13 November 2009; revised 10 July 2010; accepted 28 July 2010; published 16 November 2010.