## DATAPREPROCESSING

Total Monsoon Rainfall (TMRF) time series (𝑋𝑖) from the year 1993 to 2011 has been constructed for Mainpar region of Chhattisgarh, India. Since, transfer function sigmoid axon is used in the BPN model. The output of sigmoid axon has in close interval 0 to 1. Therefore, TMRF data time

series is normalized by using following equation 1 and obtained new normalized TMRF data time series (𝑅𝑖). The equation 2 is used to de-normalize thereafter for actual representation 𝑋𝑖. Data time series (𝑋𝑖) for the first 19 years (1993-2011) are used for developing the model. Remaining 11 years

(2012-2022) data time series (𝑋𝑖) is used to test the model independently for its acceptance. And

finally model is verified and operated for the year of 2023.

𝑅𝑖 = 𝑋𝑖 + min(𝑋𝑖)

𝑋𝑖 + max(𝑋𝑖)

𝑋𝑖 = (min(𝑋𝑖)− 𝑅𝑖 .max(𝑋𝑖))

(𝑅𝑖−1)

…(1)

…(2)

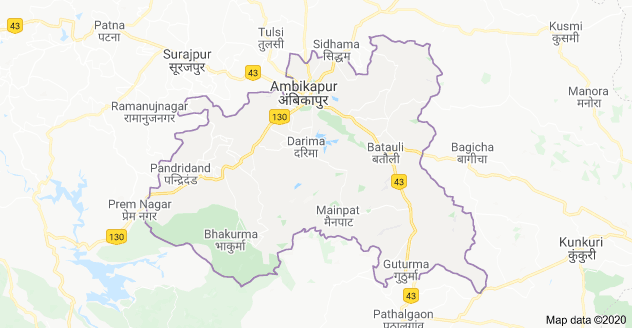
The MPS in parametric forecast is designed with twelve identified meteorological parameters. Most of those parameters have shown a very strong correlation with the TMRF. All these twelve parameters have different patterns, few of the current years parameters affecting the rainfall of same year wherein few previous years parameters affecting current years monsoon rainfall. MPS in parametric forecast is illustrated in **[Figure 1]**, where 12 parameters (p1…p12) as depicted in Table 2 is used as input to input layer, three neurons in hidden layer (z1…z3) and one neuron in output unit are used to observe TMRF. Thus a total of 43 trainable weights including biases have been used in the network. Output target value for the network is actual TMRF data. The model parameters are given in the following Table 1. The list of identified meteorological parameters for forecasting of monsoon rainfall is depicted in Table 2 along with requisite details.

## MPS IN PARAMETRIC FORECAST

MPS in parametric forecast is illustrated in **[Figure 2]**, where 12 parameters (*p*1…*p*12) as depicted in Table 2 is used as input to input layer, three neurons in hidden layer (*z*1…*z*3) and one neuron *Y*k in output unit are used to observe current year TMRF. The model parameters are given in the following Table 1. The list of identified meteorological parameters for forecasting of monsoon rainfall is depicted in Table 2 along with requisite details.

Thus a total of 43 trainable weights including biases have been used in the network. Output target value for the network is actual current year TMRF data. In fact, it has been found that MAD (% of mean) between actual and model predicted value is directly proportional to number of hidden neuron (*p*).

These networks have input layer at the bottom, one hidden layer at the middle, and one output layer at the top. It has been observed that one hidden layer is sufficient, using two hidden layer rarely improves the model, and it may introduce a greater risk of converging to a local minima. However it has been identified that the three neuron in hidden layer and twelve input vectors provided satisfactory result for all meteorological data. Increase in number of neurons in hidden layer increases MAD between actual and predicted value Karmakar et al. [14].



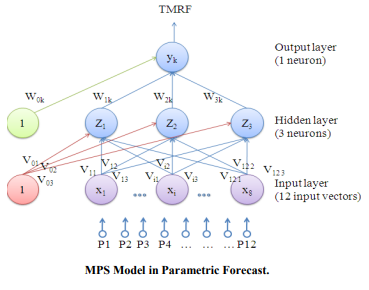
**Figure 1. Mainpat region, geographically located at** **22.8199°N longitude to 83.2828°E latitude.**

It is found that, a high learning rate ‘α’ leads to rapid learning but the weights may oscillate, while a lower learning rate leads to slower learning, Kumar (2007). If the momentum is added to the weight update formula, the convergence is faster. The main purpose of the momentum is to accelerate the convergence of error propagation algorithm. It is found that, if initial weight is too large the initial input signals to each hidden or output unit will fall in the saturation region where the derivative of sigmoid has very small value (f (network) = 0). If initial weights are too small, the network output to a hidden or output unit will approach zero, which then causes extremely slow learning. To get the best result the initial weights (and biases) are set to random number between 0 and 1. Rumelhart et al.(1986) first introduced BPN based on gradient descent method [16]. Being a gradient descent method, it minimizes the mean square error (MSE) of the output computed by the network during the feed-forward and back-propagation process.

As a result, three neurons in hidden layer have been selected and it provides most desirable result. Output target value for the network is actual data to be predicted. The neurons output can be obtained as *f* (*x*j). The neurons output is obtained as *f* (*x*j) where *f* is a transfer function typically the sigmoid function is used i.e.

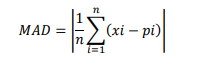
1

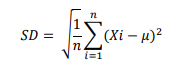
(𝑥) = 1 + 𝑒−𝛿𝑥+𝑛

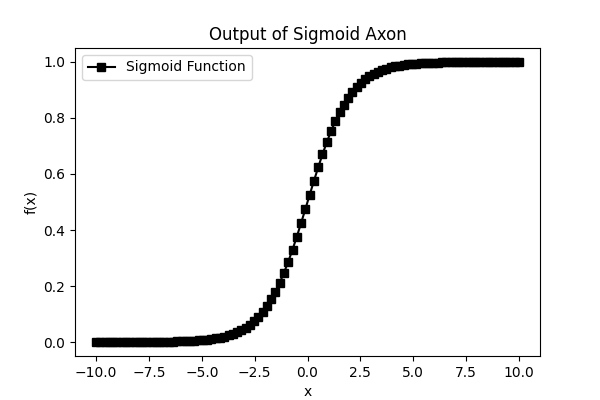
****

**Figure 2: MPS Model in Parameteric Forecast**

where *δ* determines the slope and *η* is the threshold. In the proposed model δ =1, η =0, the output of the neuron will be in close interval [0, 1] as shown in **[Figure 3]**. Learning rate α = 0.0364438 is used. It is found that, α = 0.0364438 is suitable for almost all of the cases and in this study. It is also observed that, for nearly all problems, one hidden layer is sufficient. Using two hidden layers rarely improves the model, and it may introduce a greater risk of converging to local minima and also there is no theoretical reason for using more than two hidden layers.







**Figure 3: Output of Sigmoid Axon**

**Table 1. Parameters of MPS Model**

|  |
| --- |
| Name of the model MPS in Parametric Forecast |

|  |  |
| --- | --- |
| 1. Number of layer | 03 |
|  |  |
| 2. Input vector (n) | 12 parameters + 1Dependant parameter |
|  |  |
| 3. Input data (independent parameter)  year | 12 parameters of current and previous |
| 4. Output neurons | One neuron |
|  |  |
| 5. Target (Dependent parameter) | TMRF |
|  |  |
| 6. Hidden layer | 01 |
|  |  |
| 7. Neurons in hidden layer (p) | 03 |
|  |  |
| 8. Transfer function (f) | Sigmoid. |
|  |  |
| 9. Learning rate (α) | 0.0364438 |
| 10. Momentum factor (μ) | 1.0 |
|  |  |
| 11. Weight optimization rule | Delta learning rule (Back Propogation ) |
|  |  |
| 12. Number of weights (wi) | 43 (Including biases). |
|  |  |
| 13. Epochs | 5000 |
|  |  |
| 16. Trained weights | Table 1. |
|  |  |
| 17. Performance | Excellent in training period, and good in the independent period . |

**Table 2 List of Identified Meterological Parameters**

|  |  |
| --- | --- |
| **Sl.**  **No.** | **Parameters** |

|  |
| --- |
| Highest Maximum Temperature |
| Lowest Minimum Temperature |
| Mean Maximum Temperature |
| Mean Minimum Temperature |
| Mean Relative Humidity |
| Mean Wind Speed |
| Mean Station Level Pressure |
| Mean Dew |
| Mean Wet Bulb |
| Mean Vapour Pressure |
| Heaviest Rainfall in 24 hour |
| Cloud Amt |

**Table 3. Initial Random Weights for MPS in Parametric forecast**

***Weights***

Hidden Layer Weights:

|  |  |  |
| --- | --- | --- |
| 0.863792 | 0.483292 | 0.782414 |
| 0.539032 | 0.428871 | 0.175198 |
| 0.98087 | 0.134237 | 0.895696 |
| 0.918302 | 0.376361 | 0.11424 |
| 0.16682 | 0.862062 | 0.866989 |
| 0.624043 | 0.010861 | 0.143803 |
| 0.836227 | 0.339984 | 0.203833 |
| 0.088325 | 0.574381 | 0.362452 |
| 0.806964 | 0.488113 | 0.606796 |
| 0.879906 | 0.026557 | 0.106788 |
| 0.275506 | 0.710538 | 0.522044 |
| 0.835395 | 0.281752 | 0.858415 |

Hidden Layer Bias

|  |
| --- |
| 0.135504 |
| 0.915212 |
| 0.578844 |

Output Layer Weights

|  |
| --- |
| 0.623661 |
| 0.739239 |
| 0.854091 |

Output Layer Bias

0.876588

Normalization

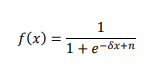


Denormalization



Activation Function

The neurons output is obtained as f (xj) where f is a transfer function typically the sigmoid function is used i.e



MAD,SD Formula

**[3] RESULTS & DISCUSSIONS**

To check the accuracy of the experiments and analyze the performances of MPS during training period, it is trained with 19 years (1993-2011) training dataset in parametric forecast.

While training of the network in parametric forecast it has shown a very high compatibility between the actual and predicted rainfall, in course of training the final CC between the actual and predicted rainfall was very high i.e. 0.830029 CC.

In order to check the performance of MPS during the testing / independent period with new data, it was tested with the 10 years (2012- 2022) of the test data and verified for 2023.

**Table 4 TMRF (Total Monsoon Rainfall) data for training period**

|  |  |  |
| --- | --- | --- |
| **YEAR** | **Predicted TMRF** | **Actual TMRF** |
| 1993 | 282.2228 | 171.09 |
| 1994 | 207.1296 | 172.18 |
| 1995 | 235.5964 | 158.02 |
| 1996 | 201.3776 | 162.78 |
| 1997 | 306.1561 | 161.55 |
| 1998 | 63.86442 | 119.61 |
| 1999 | 147.5526 | 125.12 |
| 2000 | 183.1023 | 160 |
| 2001 | 150.2657 | 138.7 |
| 2002 | 202.6334 | 162.36 |
| 2003 | 247.2671 | 160.41 |
| 2004 | 115.5908 | 130.91 |
| 2005 | 169.6282 | 166.44 |
| 2006 | 116.5974 | 188.44 |
| 2007 | 120.2133 | 135.25 |
| 2008 | 142.1412 | 143.7 |
| 2009 | 173.4022 | 168.62 |
| 2010 | 182.0057 | 150.2 |
| 2011 | 86.18304 | 177.81 |

**Table 5: TMRF (Total Monsoon Rainfall ) data for testing period**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **YEAR** | **Predicted TMRF** | **Actual TMRF** | | 2012 | 106.8722 | 95.15 | | 2013 | 158.8957 | 157.33 | | 2014 | 184.8759 | 140.83 | | 2015 | 201.5846 | 174.17 | | 2016 | 219.8377 | 192.54 | | 2017 | 249.6307 | 165.97 | | 2018 | 111.8397 | 100.61 | | 2019 | 132.5282 | 147.93 | | 2020 | 237.3531 | 200.57 | | 2021 | 193.8086 | 113.9 | | 2022 | 134.9287 | 108.43 |   **Table 6: TMRF(Total Monsoon Rainfall) for Verification Period 2023**   |  |  |  | | --- | --- | --- | | **YEAR** | **Predicted TMRF** | **Actual TMRF** | | 2023 | 130.5977 | 136.74 | |  |  |
| **Table 9. Statistics of Performance of MPS.**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Model** | **Training Period** | | | **Independent Period** | | | | **SD** | **MAD** | **CC** | **SD** | **MAD** | **CC** | | Parametric  forecast | 0.047373 | 0.027862 | 0.830029 | 0.017321 | 0.005128 | 0.80 | |  |  |
| C:\Users\PC\Desktop\CC 7 values\GRAPHS\graph2_metrics_plot.png  **Figure 4: Metrics for the performance of model** |  |  |
| C:\Users\PC\Desktop\CC 7 values\GRAPHS\graph 1.png **Figure 5: Performance of MPS in Parametric forecast on 5000 epochs of training and testing period** |  |  |
|  |  |  |
| IMD’s operational model is appropriate for LRF of monsoon rainfall over whole Central India. However it is inadequate in case of LRF over a very smaller region like Mainpat which is allocated in Central India. The IMD’s operational model is based on statistical power regression analysis which uses few global dynamic parameters (i.e., predictors). These similar parameters are not useful for prediction over this region. It is concluded that impact of global parameters (i.e., independent) on the monsoon rainfall (i.e., dependent) over the smaller region is irrelevant. Identification of physically connected global meteorological parameters for monsoon rainfall over smaller region is also extremely difficult. The ANN techniques are capable in identification of internal dynamics of chaotic time series data. Thus, MPS in parametric forecast is recommended for this type of study. |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**References:**

1. Rumelhart, D., Hinton, G. E., Williams, R. J.*,* Learning Internal Representation By Error Propagation, R J Parallel Distributed Processing: Exploration in the Microstructure of Cognition, MIT Press Cambridge, 1, 318-362, (1986).
2. Karmakar, S., et al., “Development of Artificial Neural Network Models for Long-Range Meteorological Parameters Pattern Recognition over the Smaller Scale Geographical Region”, IEEE Computer Society, Washington, DC, USA., 8-10 Dec., pp.1 – 6, (2008).
3. Kumar, S., Neural Network Computer Engineering Serie, The McGraw-Hill, New Delhi, , 104-152, (2007).
4. [A Pragmatic Study Based on Neural Network for Forecasting Long-range Monsoon Rainfall over Smaller Geographical Region](https://www.researchgate.net/publication/348183835_A_Pragmatic_Study_Based_on_Neural_Network_for_Forecasting_Long-range_Monsoon_Rainfall_over_Smaller_Geographical_Region?enrichId=rgreq-eb9e8ae92d7a9f06911e19b533c6f632-XXX&enrichSource=Y292ZXJQYWdlOzM0ODE4MzgzNTtBUzoxMDcwNzY5MjgzMDM1MTQyQDE2MzIzMDI2OTk0Nzc%3D&el=1_x_3&_esc=publicationCoverPdf) by Gyanesh Shrivastava