Indian summer monsoon rainfall prediction

A comparison of ARIMA and ANN models

Dr. Abhishek Singh

Sunday, June 14, 2020

Introduction

- **Challenge:** accurate forecasting of Indian summer monsoon rainfall (ISMR) from June through September.

Data

- **Data Soure:** Indian Institute of Tropical Meteorology (IITM), Pune.

Data

- **Data Soure:** Indian Institute of Tropical Meteorology (IITM), Pune.
  - contains monthly rainfall data for **146 years** (1871-2016).

Data

- **Data Soure:** Indian Institute of Tropical Meteorology (IITM), Pune.
  - contains monthly rainfall data for **146 years** (1871-2016).
- The data were divided into:

Data

- **Data Soure:** Indian Institute of Tropical Meteorology (IITM), Pune.
  - contains monthly rainfall data for **146 years** (1871-2016).
- The data were divided into:
  - training (1871-1960), and

Data

- **Data Soure:** Indian Institute of Tropical Meteorology (IITM), Pune.
  - contains monthly rainfall data for **146 years** (1871-2016).
- The data were divided into:
  - training (1871-1960), and
  - testing (1961-2016) data sets

Methods

- The following methods were used to forecast the values of ISMR:

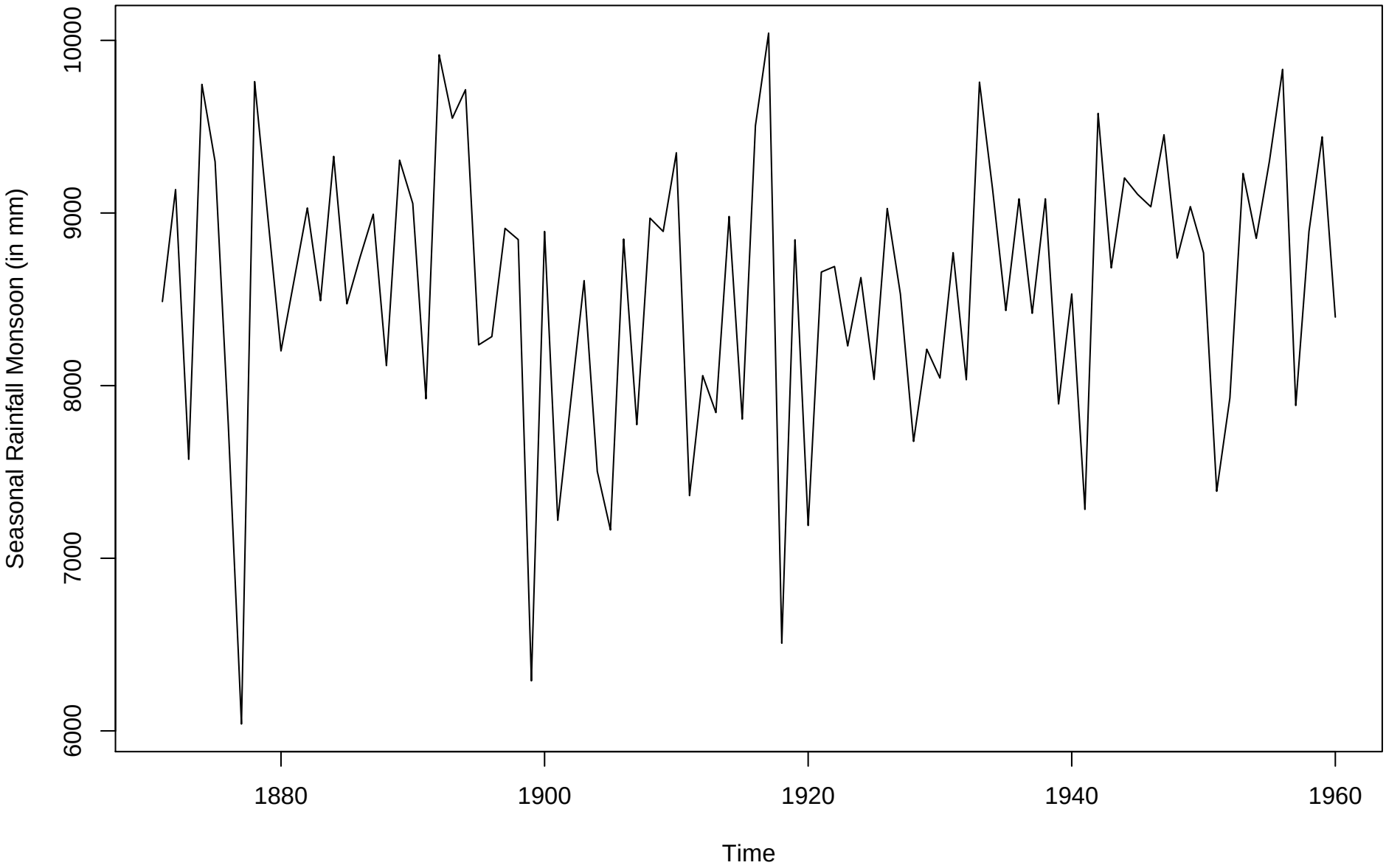
Methods

- The following methods were used to forecast the values of ISMR:
  - Autoregressive Integrated Moving Average (ARIMA) Model

Methods

- The following methods were used to forecast the values of ISMR:
  - Autoregressive Integrated Moving Average (ARIMA) Model
  - Non-linear Artificial Neural Network (NANN) Model

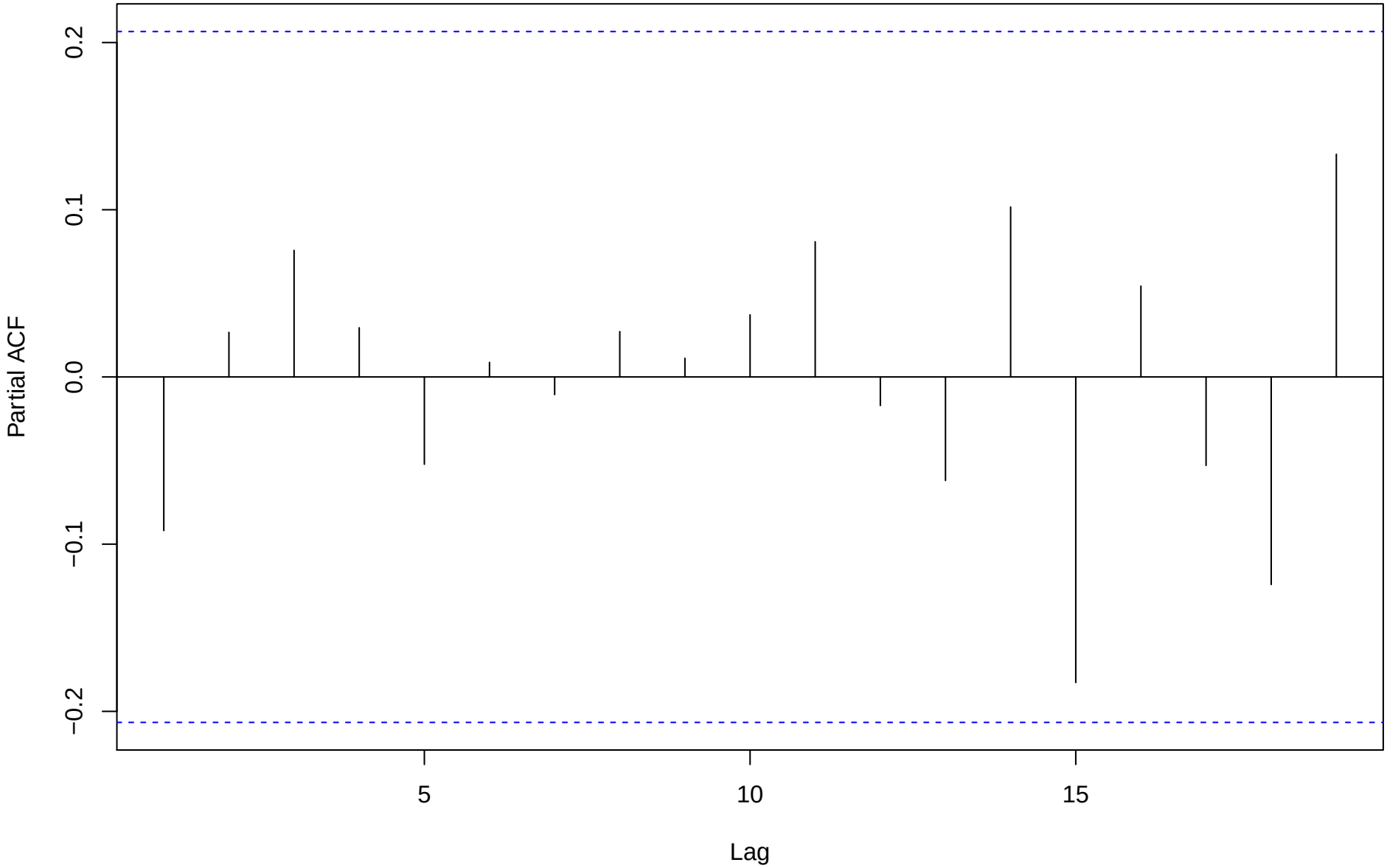
ISRM training data (1871-1960)



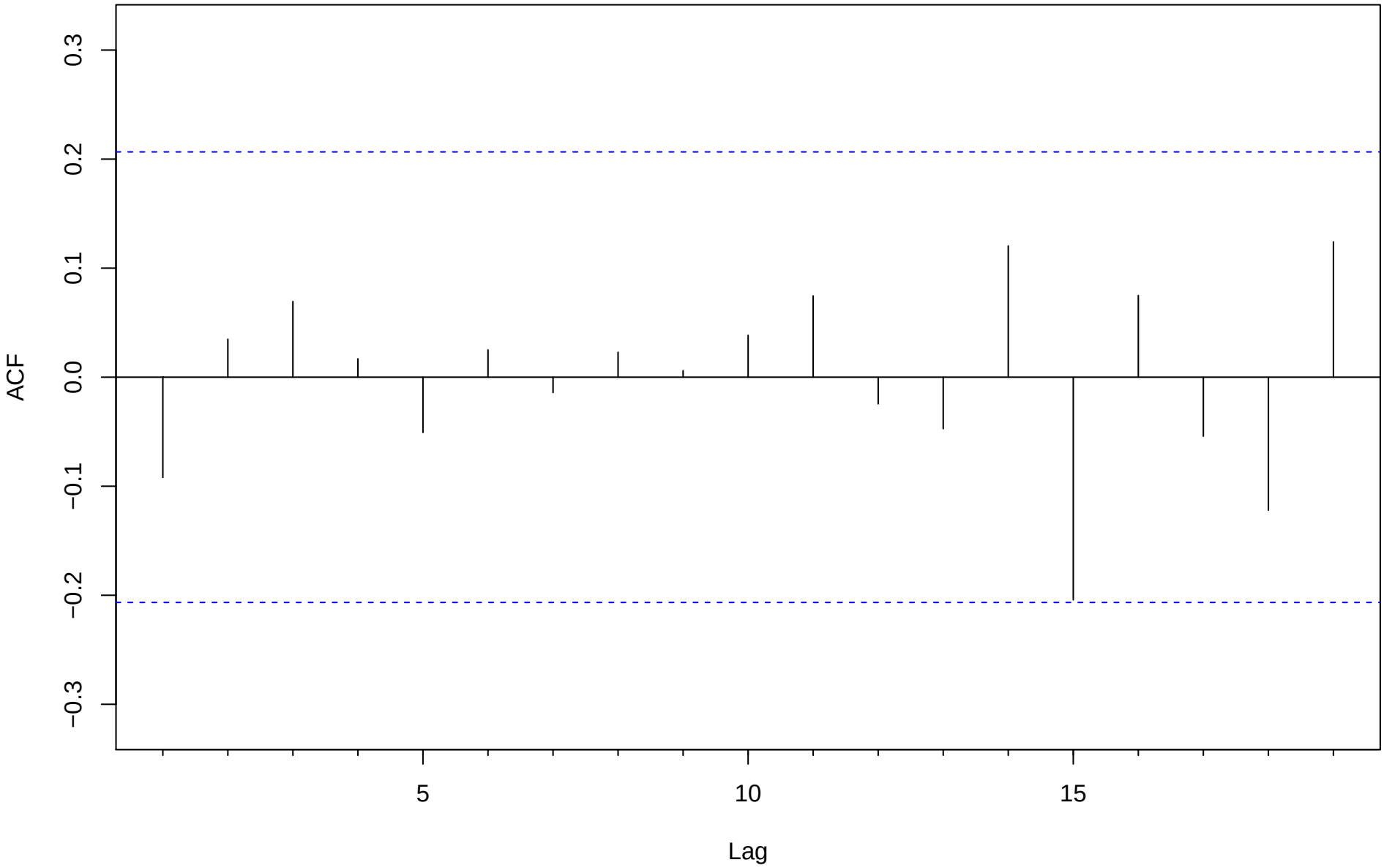
Augmented Dickey-Fuller Test (1871-1960)

```
    Augmented Dickey-Fuller Test

data:  trainingdata
Dickey-Fuller = -4.1077, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

Partial Autocorrelation Function (1871-1960)



Autocorrelation Function (1871-1960)

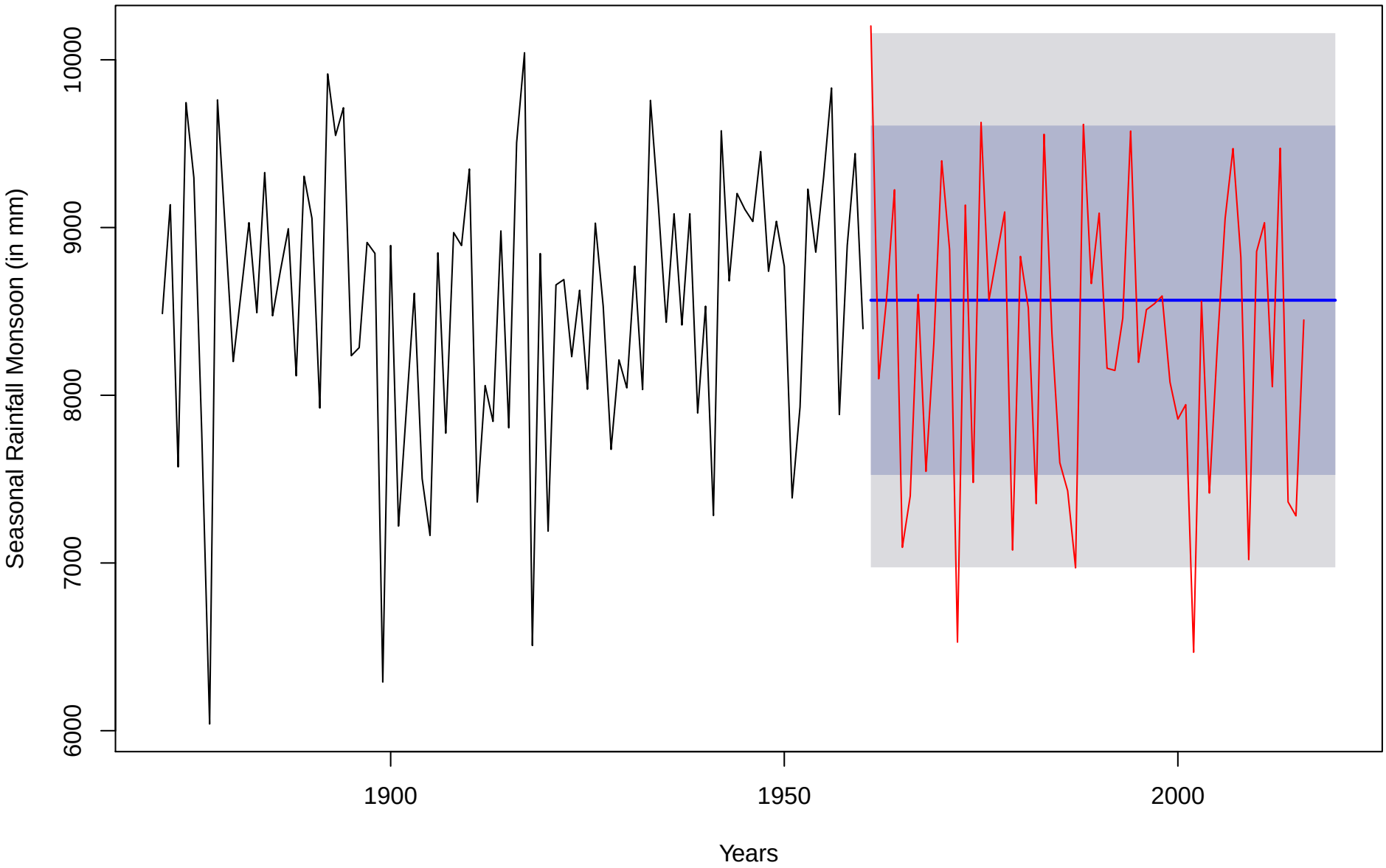


Parameter Estimation

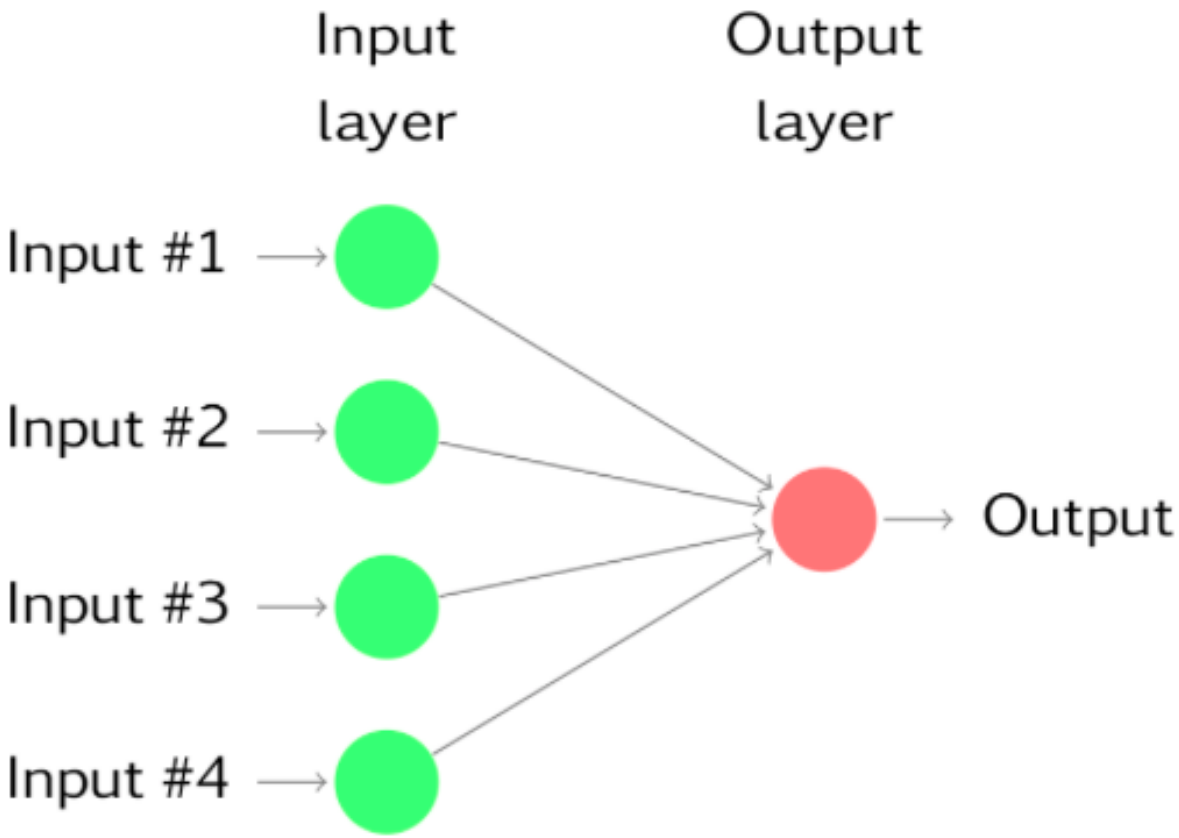
- Thus the proposed model is $ARIMA(0, 0, 0)$.

```
Series: trainingdata
ARIMA(0,0,0) with non-zero mean

Coefficients:
          mean
      8566.6444
s.e.    85.1768

sigma^2 estimated as 660294:  log likelihood=-730.22
AIC=1464.44   AICc=1464.58   BIC=1469.44
```

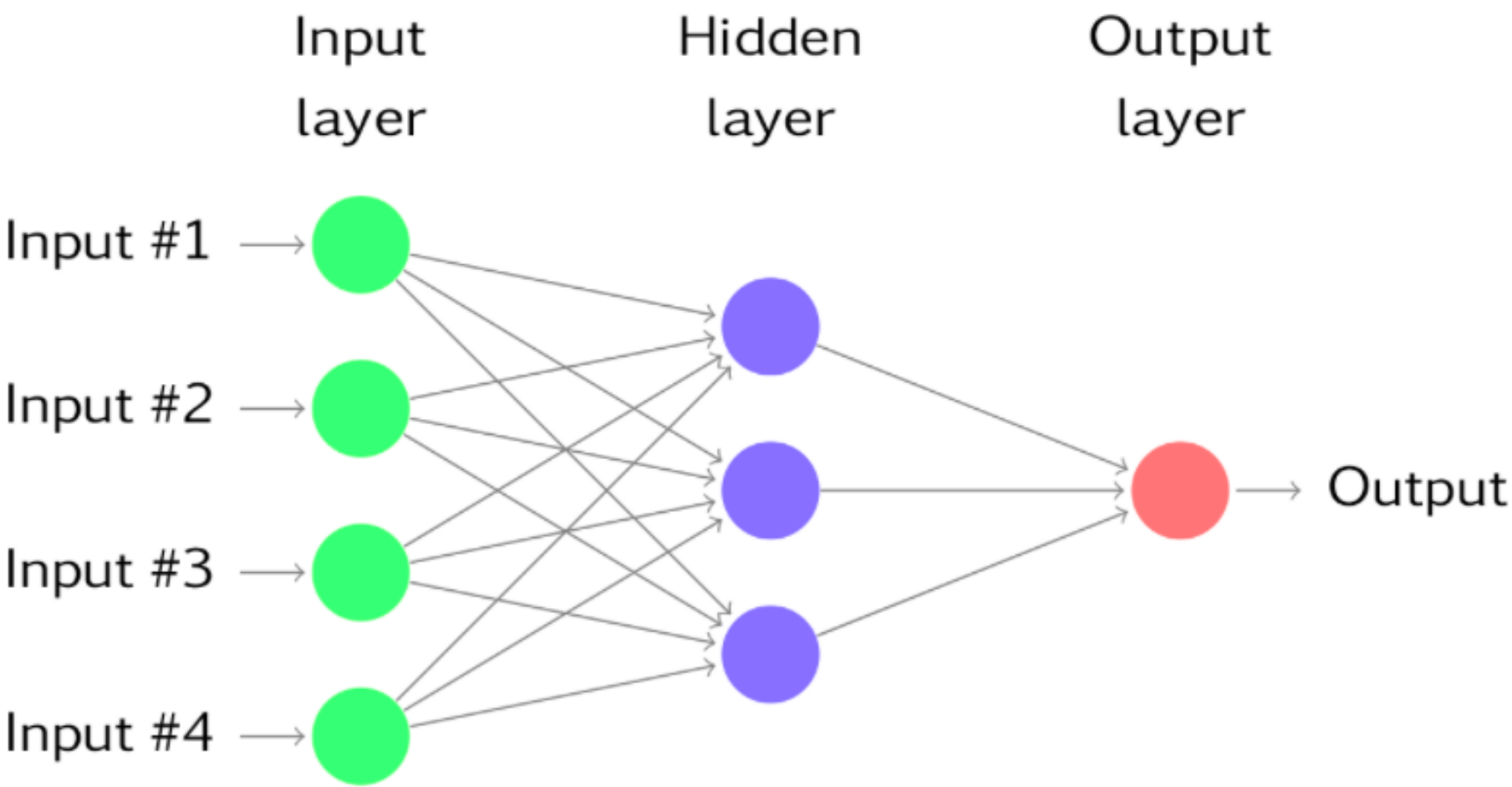
Forecasts from ARIMA(0,0,0)



Simple Neural Netwrok



Multilayer Feed-Forwad Neural Netwrok



Proposed NANN(p,k) Model

- The proposed model is $NANN(1,1)$.

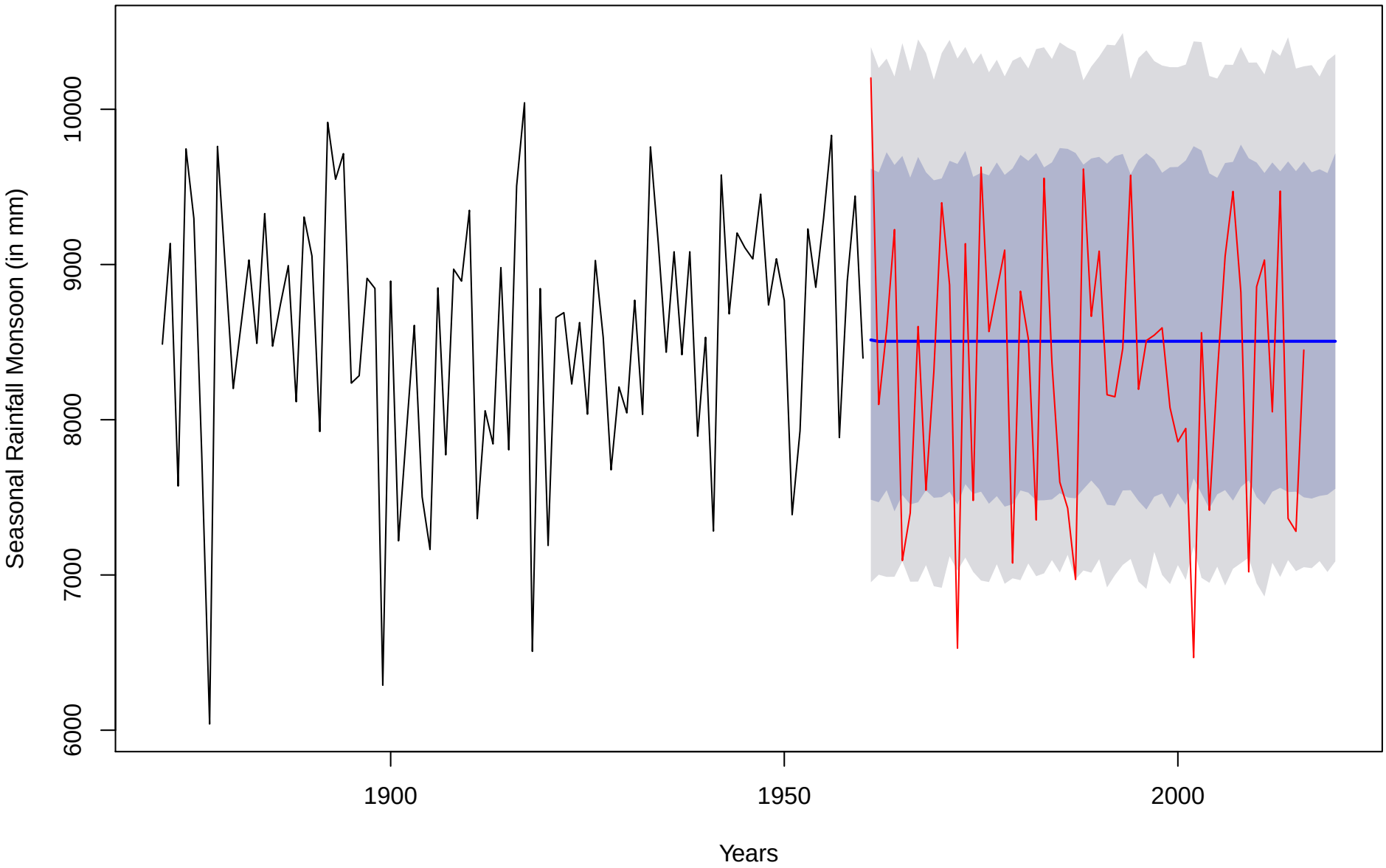
Proposed NANN(p,k) Model

- The proposed model is $NANN(1,1)$.
  - $p = 1$ is lagged input

Proposed NANN(p,k) Model

- The proposed model is $NANN(1,1)$.
  - $p = 1$ is lagged input
  - $k = 1$ is number of node(s) in hidden layer

Forecasts from NANN(1,1)



Conclusion

- Both of the methods $ARIMA(0,0,0)$ and $NANN(1,1)$ provide same level of prediction.

Conclusion

- Both of the methods $ARIMA(0,0,0)$ and $NANN(1,1)$ provide same level of prediction.
- The behavior of their predication intervals differ significantly.

Conclusion

- Both of the methods $ARIMA(0,0,0)$ and $NANN(1,1)$ provide same level of prediction.
- The behavior of their predication intervals differ significantly.
  - The possible reason may be that, neural networks are not based on a well-defined stochastic model.

Conclusion

- Both of the methods $ARIMA(0,0,0)$ and $NANN(1,1)$ provide same level of prediction.
- The behavior of their predication intervals differ significantly.
  - The possible reason may be that, neural networks are not based on a well-defined stochastic model.
  - Thus, prediction intervals were calculated from the simulation models.

Take Away

- Making interpretations and further policy implications using ANN we should be very cautious due to the limitations on prediction interval.

References

Box, George EP, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. 2015. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

IITM. 2020. "Indian Institute of Tropical Meteorology." https://www.tropmet.res.in.

Sivanandam, SN, and SN Deepa. 2007. *Principles of Soft Computing (with Cd)*. John Wiley & Sons.