Investigating an association between sunspot numbers and summer-monsoo rainfall in India

Surajit Chattopadhyay
Pailan College of Management and Technology, Kolkata, India
Email: surajitchatto@outlook.com

3rd International Conference on
Hydrology & Meteorology
September 15-16, 2014, HICC, Hyderabad, India

ISMR-an overview

e summer monsoon during the months of June to September (JJAS) is the major rainfall season for
ia. Hence, the prediction of Indian summer monsoon rainfall (ISMR) is always in high demand as i
n variability has significant impact on economy of the country.

ring ISM the Western Ghats and the northeast region receives precipitation due to orographic
rplay and eastern India i.e. Orissa, Chhattisgarh and West Bengal receives precipitation due to
nsoon trough.

st of the statistical forecasting techniques available till dates have concentrated only in generating a
erministic prediction without any measure of its inherent uncertainty.

probabilistic predictions have the ability to quantify the uncertainty, it has more potential to decisi
kers than deterministic forecasts

ently some studies have raised the issues of making probabilistic forecast using GCMs.

Significant recent works on ISMR

harya, N., Kar, S.C., Mohanty, U.C., Kulkarni, M.A., Dash, S.K., 2011. Performance of GCMs seasonal prediction over India - a case study for 2009 monsoon. *Theor. Appl. Climatol*. 105, 50).

harya, N., Chattopadhyay, S., Kulkarni, M.A., Mohanty, U.C., 2012. A neurocomputing proach to predict Monsoon Rainfall in Monthly scale using SST anomaly as a predictor. *Acta ophys*. 60, 260–279.

harya, N., Mohanty, U. C., Sahoo, L., 2013. Probabilistic multi-model ensemble prediction of dian summer monsoon rainfall using general circulation models: A non-parametric approach. *mptes Rendus Geoscience*. 345, 126-135.

hilash, S., Sahai, A. K., Pattnaik , S., De, S., 2013. Predictability during active break phases of dian summer monsoon in an ensemble prediction system using climate forecast system. *J. Atm lar-Terres. Phys*. 100–101, 13–23.

gh, A., Acharya, N., Mohanty, U.C., Robertson, A. W., Mishra. G., 2012. On the predictability dian summer monsoon rainfall in general circulation model at different lead time. *Dynamics o nospheres and Oceans*. 58, 108-127.

The following pioneering work states……

he Effects of Solar
riability on the
rth's Climate

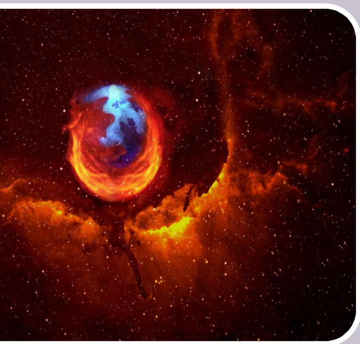
uthor(s): Joanna D.
aigh
ource: *Philosophical*
ansactions:
athematical,
ysical and
gineering Sciences,
ol. 361, No. 1802,
an. 15, 2003), pp. 95-
1

"The extent to which changes in solar activity affect climate has been the subject of considerable investigation over many years and has often been the cause of speculation and controversy.

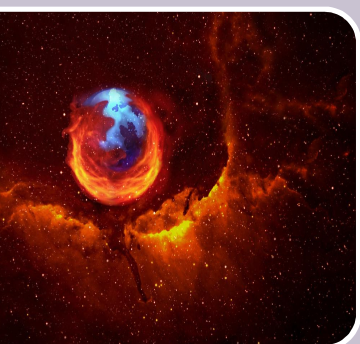
As observational and modelling techniques improve, and our understanding of the natural internal variability of the climate system advances, it is becoming more feasible both to detect solar signals in climate records and to investigate the mechanisms whereby the solar influence acts.

This subject is an interesting and complex scientific area to study, but it is now also of considerable practical importance in terms of differentiating natural and anthropogenic causes of climate change so that more reliable estimates can be made of the potential future impacts of human activities on climate."

A brief overview of Sunspot

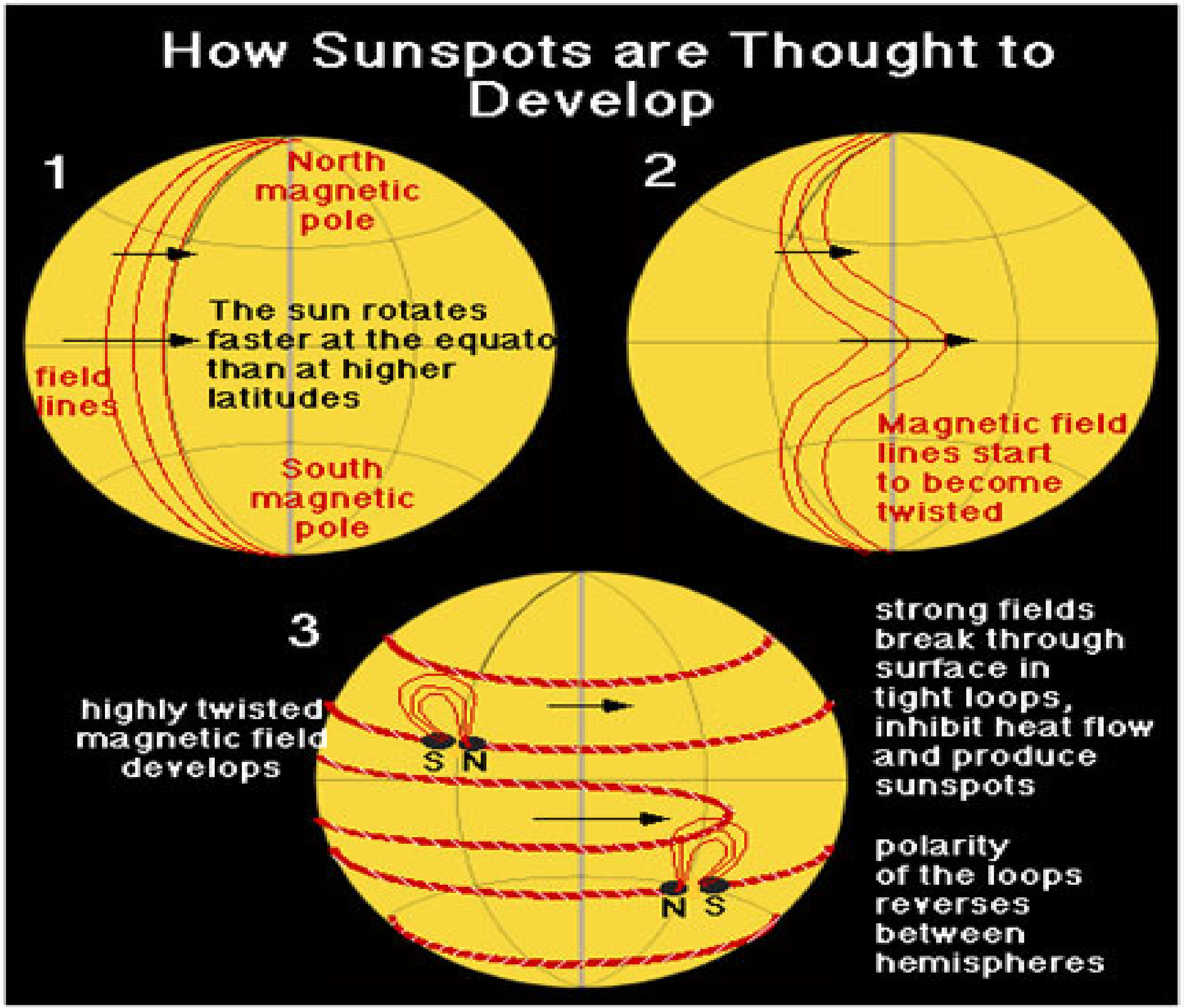


Although the details of sunspot generation are still a matter of research, it appears that sunspots are the visible counterparts o magnetic flux tubes in the Sun's convective zone that get "wou up" by differential rotation.



If the stress on the tubes reaches a certain limit, they curl up lik rubber band and puncture the Sun's surface. Convection is inhib at the puncture points; the energy flux from the Sun's interior decreases; and with it surface temperature.

Due to its link to other kinds of solar activity, sunspot occurrenc can be used to help predict space weather, the state of the ionosphere, and hence the conditions of short-wave radio propagation or satellite communications. Solar activity (and th sunspot cycle) are frequently discussed in the context of global warming.


How Sunspots are Thought to Develop
1
North magnetic pole
The sun rotates faster at the equato than at higher latitudes
field lines
South magnetic pole
2
Magnetic field lines start to become twisted
3
highly twisted magnetic field develops
S N
N S
strong fields break through surface in tight loops, inhibit heat flow and produce sunspots
polarity of the loops reverses between hemispheres

Motivation behind the study

GEOPHYSICAL RESEARCH LETTERS, VOL. 32, L05813, doi:10.1029/2004GL021044, 2005

association between Indian monsoon rainfall and solar activity

aryya and R. Narasimha
Centre For Advanced Scientific Research, Jakkur, Bangalore, India

y 2004; revised 1 December 2004; accepted 18 January 2005; published 11 March 2005.

e 120 y period (1871–1990) for which reliable
fall statistics are available, solar activity
xhibit nonstationarity. Taking this fact into
present here the results of an analysis of four
indices and seven major Indian monsoon
eries, over two distinct test periods respectively
high solar activity, each comprising three
ar cycles. It is found that the average rainfall
all seven rainfall indices during periods of
activity, at confidence levels varying from 75%
95% or greater in three of them. Using wavelet
is also found that the power in the 8–16 y band
riod of higher solar activity is higher in 6 of the
e series, at confidence levels exceeding 99.99%.
s support existence of connections between
l and solar activity. **Citation:** Bhattacharyya, S.,
mha (2005), Possible association between Indian
fall and solar activity, *Geophys. Res. Lett.*, 32,
0.1029/2004GL021044.

ction

ingly strong evidence for possible association
r processes and terrestrial climate indices [*Beer*
Lassen and Friis-Christensen, 1995; *Haigh*,
001; *Labitzke and van Loon*, 1997; *Mehta and*
s accumulated in recent years. In a recent study
l. [2001], strong coherence between solar
d the monsoon in Oman in the period between
.P. has been reported. Doubts about effects on
cales (of the order of 10 to 100y) seem however
arlier studies of possible connections between
on rainfall and solar activity [*Jagannathan and*
3; *Jagannathan and Parthasarathy*, 1973],
tion and power spectral analysis of the rainfall
n 48 meteorological stations spread all over
ed presence of the 11-year sunspot cycle at
evels of 95% or higher in only 5 of them.
e data on Indian rainfall are available for a
) y (1871–1990). The solar activity over this
arly not a stationary process (see Figure 2
ing relatively low activity during 1875 to
gnoring this nonstationarity, we compute a
correlation coefficient between all India sum-
rainfall and sunspot numbers, we get a value
le, only 0.11 (95% confidence band −0.07 to
n equal-tails test). To test for possible con-
veen rainfall and solar activity, one therefore
into account the phase of such multidecadal
similar point has been made in a different
context by *Labitzke and van Loon* [1997]. We tackle this problem by (i) comparing averages over periods of about the same duration over which solar activity is respectively low and high, and (ii) analysing the multidecadal variations of both solar and rainfall time series through wavelet techniques. Wavelet methods have earlier been used to analyse Indian rainfall by *Torrence and Webster* [1999], *Torrence and Compo* [1998], and *Narasimha and Kailas* [200].

[4] Th
flux sig
elsewhe
200], serves to provide further motivation for the present work.

2. The Data Analysed

[5] Seven annual area-weighted Indian rainfall time series for the period 1871–1990 have been considered for the analysis. These comprise the 6 homogeneous regions of northeast India (NEI), northwest India (NWI), central northeast India (CNEI), west central India (WCI), peninsular India (PENSI) and a so-called homogeneous Indian monsoon (HIM) region, and an overall time series called the all India summer monsoon (AISM) rainfall [*Parthasarathy et al.*, 1995].

[6] We have performed the analysis also over the spatially finer scale provided by the data in each of the 28 meteorological subdivisions of India [*Parthasarathy et al.*, 1995], but will not discuss this in detail as the overall conclusions from an analysis of the 7 major rainfall time series are not thereby significantly altered. However some striking features of the results for some subdivisions will be cited.

[7] The range of scales over which these rainfall data can provide useful information on temporal variability is limited by the relatively short length of data. For the present study we have found annual rainfall to be the most appropriate rainfall index to use. The solar indices under study are sunspot number index, group sunspot number, solar irradiance and sunspot area. The sunspot index data have been obtained from *Rai Choudhuri* [1999] and *Fligge et al.* [1999], and the data for group sunspot number from the NOAA ftp site ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA. The data for the time series of annual mean solar irradiance has been obtained from *Lean* [2004] [see also *Lean et al.*, 1995] available at the online resource of the World Paleoclimate Data Center website, although some reservations have been recently expressed over irradiance estimates [see *Foukal et al.*, 2004]. We will in the main cite results using the sunspot number as the primary index of solar activity, as the other indices do not lead to significantly different conclusions.

[8] One important difference between rainfall and sunspot data is that the former are cumulative, the latter are not:

04GL021044

It is found that the average rainfall is higher
seven rainfall indices during periods of greate
activity, at confidence levels varying from 7
99%, being 95% or greater in three of them.
wavelet techniques it is also found that the p
the 8–16 y band during the period of higher
activity is higher in 6 of the 7 rainfall time se
confidence levels exceeding 99.99%.

**These results support existenc
connections between Indian rai
and solar activity**

**Bhattacharyya, S., and R. Narasimha (2005), "Possib
association between Indian monsoon rainfall and sola
activity", *Geophys. Res. Lett.*, 32, L05813,
doi:10.1029/2004GL021044.**

Other works in this direction

Thomas, Robert G. "Rome rainfall and sunspot numbers." *Journal of atmospheric and solar terrestrial physics* 55.2 (1993): 155-164.

Seleshi, Yilma, G. R. Demarée, and J. W. Delleur. "Sunspot numbers as a possible indicator of annual rainfall at Addis Ababa, Ethiopia." *International journal of Climatology* 14.8 (1994): 911-923.

Bell, Gordon John. "Changes in sign of the relationship between sunspots and pressure, rainfall and the monsoons." *Weather* 32.1 (1977): 26-32.

Purpose of the present paper

In the present work we shall investigate whether there is any association between the SN time series and the summer monsoon (June–September; abbreviated as JJAS) rainfall time series over India.

We shall analyze the autocorrelation structure of the time series under consideration and subsequently execute a spectral analysis after Box–Cox transformation of both of the time series.

We shall also examine the possibility of using SN as a predictor fo the summer monsoon rainfall over India.

Autocorrelation function

We consider the summer monsoon (JJAS) rainfall time series over India

collected from the website of the Indian Institute of Tropical Meteorology, Pune, and the mean annual SN time series

collected by the National Oceanic and Atmospheric Administration (NOAA), Boulder, Colorado during the period from 1935 to 1998.

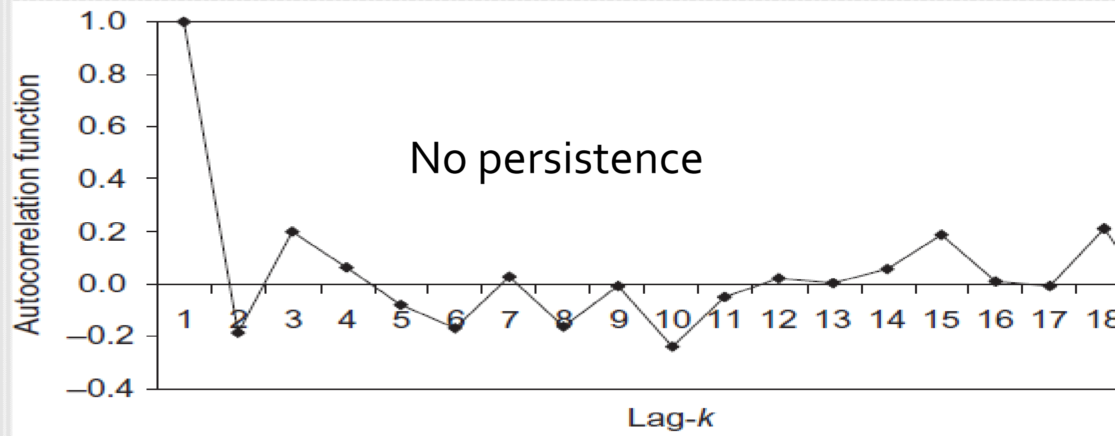
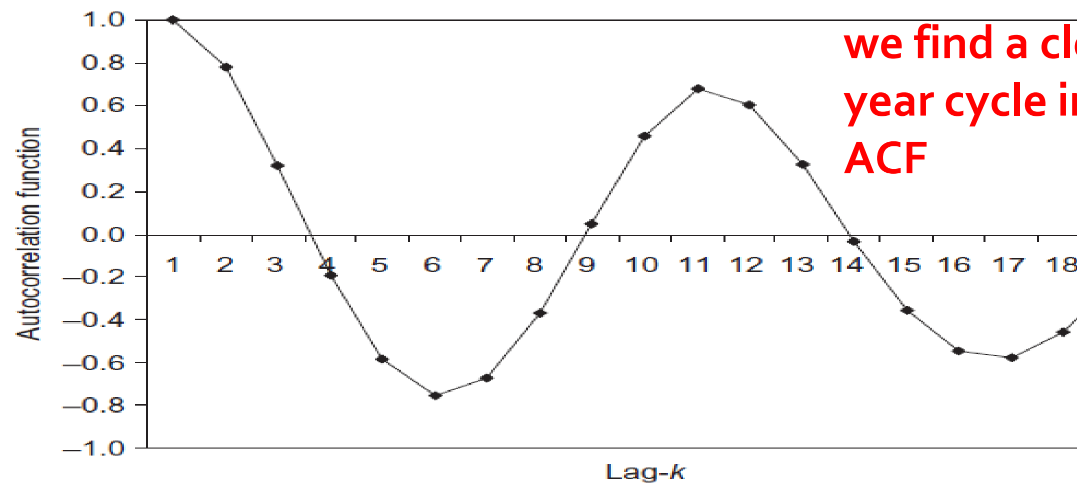


Figure 1. Autocorrelation function of the rainfall time series.



we find a cl
year cycle i
ACF

Figure 2. Autocorrelation function of the sunspot number (SN) time s

Box–Cox transformation to convert the time series to Normal i.e. skewness (μ_3/σ^3)=0 (the transformation parameter is computed using Newton Raphson method)

ewness (Pearson) for ISMR=0.049 , for l=0.477

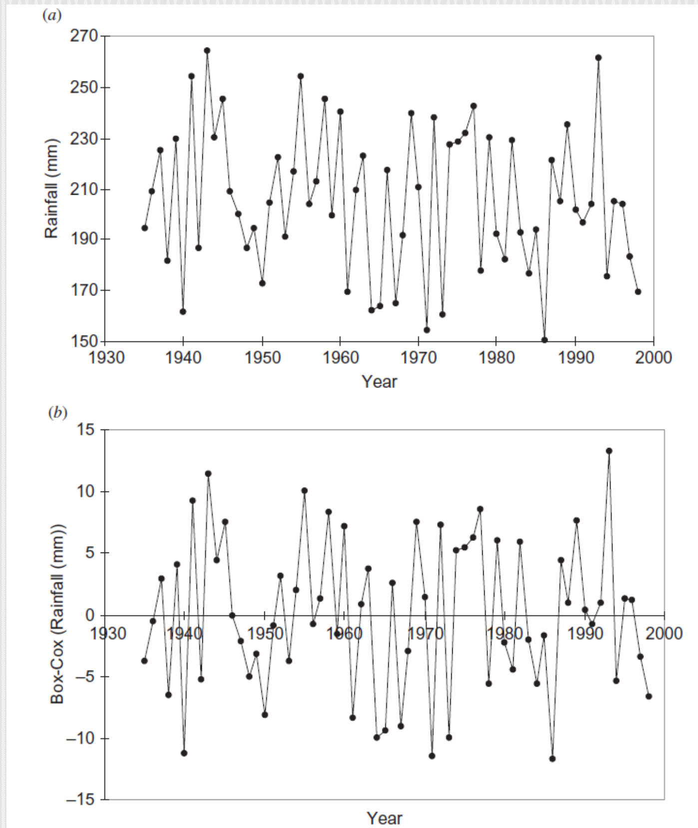
For the rainfall time series, *λ = 0.716 and G =203.8* *and for the SN time series λ = 0.366, and G =52.83*

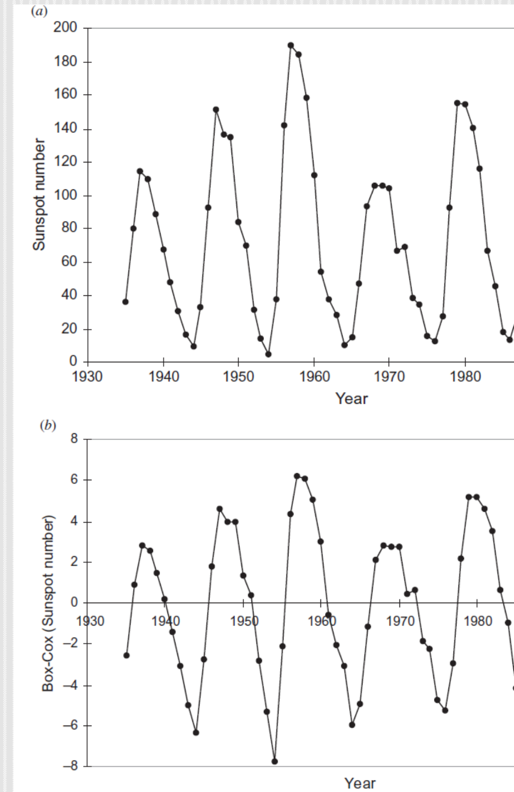
- It is found that both of the time series under consideration are positively skewed.

- Thus, we apply a Box–Cox transformation to the time series to stabilize the variance as follows

$$Z_t = \frac{X_t^{\lambda} - 1}{\lambda G^{\lambda - 1}}$$

where X_t is the original series, G is the sample geometric mean, λ is the transformation parameter, and Z_t is the transformed series.





Spectral Analysis

After transforming the time series by Box–Cox, we carry out a spectral analysis of the time series. Detailed theory of the spectral analysis is available in Wilks (2006, pp. 83–86).
Any data series consisting of n points can be represented by adding together a series of n/2 harmonic functions as:

$$y_t = \bar{y} + \sum_{k=1}^{n/2} \left\{ A_k \cos\left[\frac{2\pi kt}{n}\right] + B_k \sin\left[\frac{2\pi kt}{n}\right] \right\}$$

A_k and B_k are Fourier coefficients, y_t is the entry to the time series at time t. Spectral density is

$$C_k = \sqrt{A_k^{\,2} + B_k^{\,2}}$$

The advantage of this new perspective is tha it allows us to see separately the contributions to a time series that are made by processes varying at different speeds, tha is by processes operating at a spectrum of different frequencies

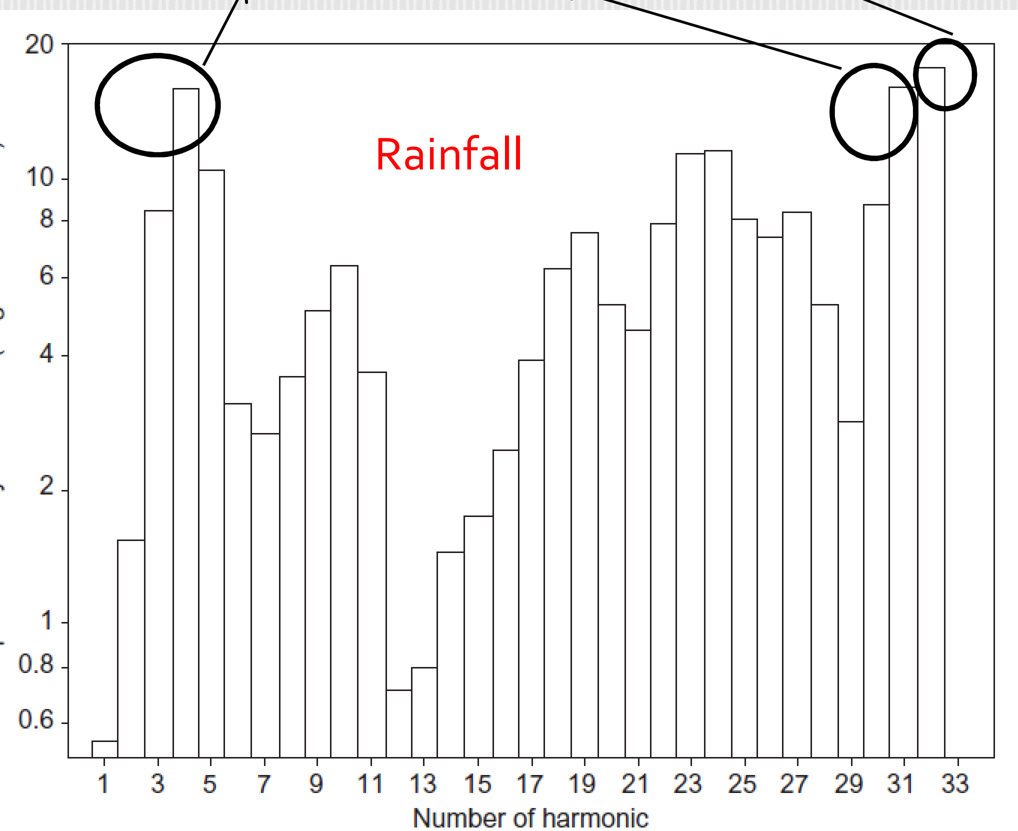
The horizontal axis of a spectrum can also b scaled according to the reciprocal of the frequency, or the period of the kth harmoni as:

$\tau_k = n/k$, where k=1, 2, ..., n/2=32.

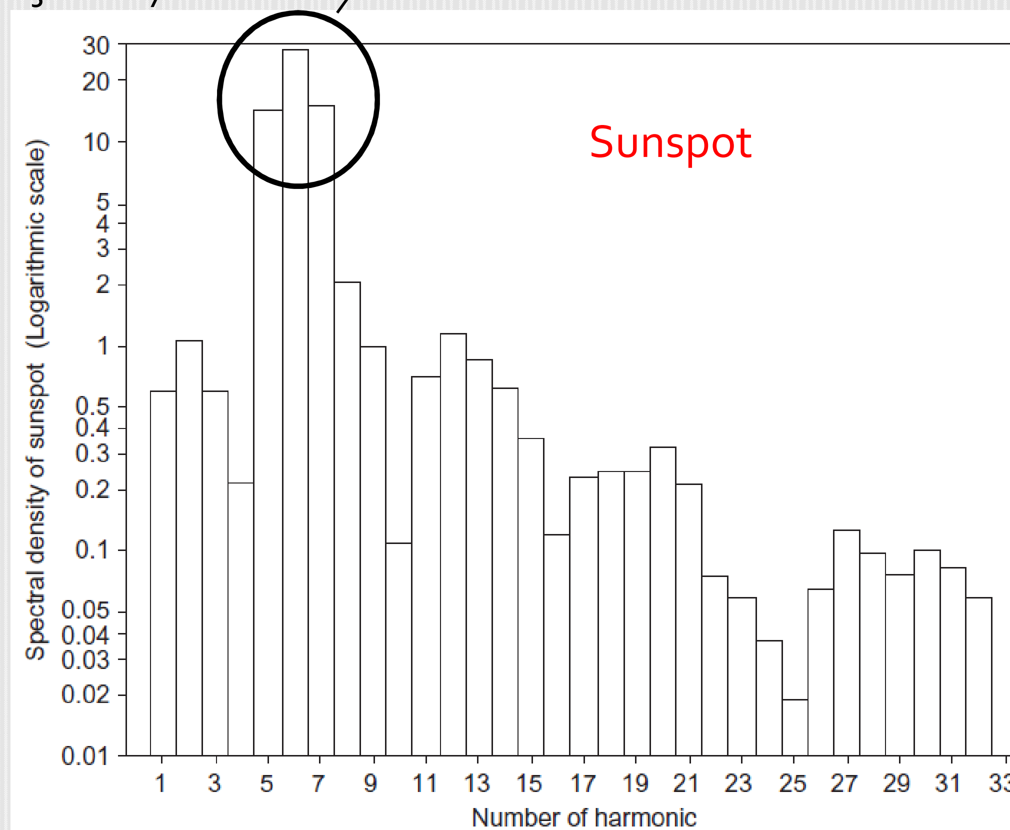
Thus, we understand that the most important harmonic in the two time series differ by (16–10.67)=5.33 years.

If we look at the spectral densities for the rainfall time series we find that there is a prominent spectral density at the harmonic τ_5. *Thus, τ_5 is a prominent harmonic in both of the time series.*

st important variations in the data series are represented by harmonics τ_4 *=64/4=16 and* τ_{31} *=64/31=2.06,* τ_{32} *=2.*



Most important variations in the data series are represented the harmonics τ_6 = 64/6 = 10.67. Other important harmonics τ_5 and τ_7.



Development of a predictive model for the monsoon rainfall using the SN as the single predictor

We revealed that τ_5 *is a prominent harmonic in both of the time series. Thus*, we can consider τ_5 *as a common harmonic.*

From this, we attempt to establish a predictor–predictand relationship between SN and monsoon rainfall.

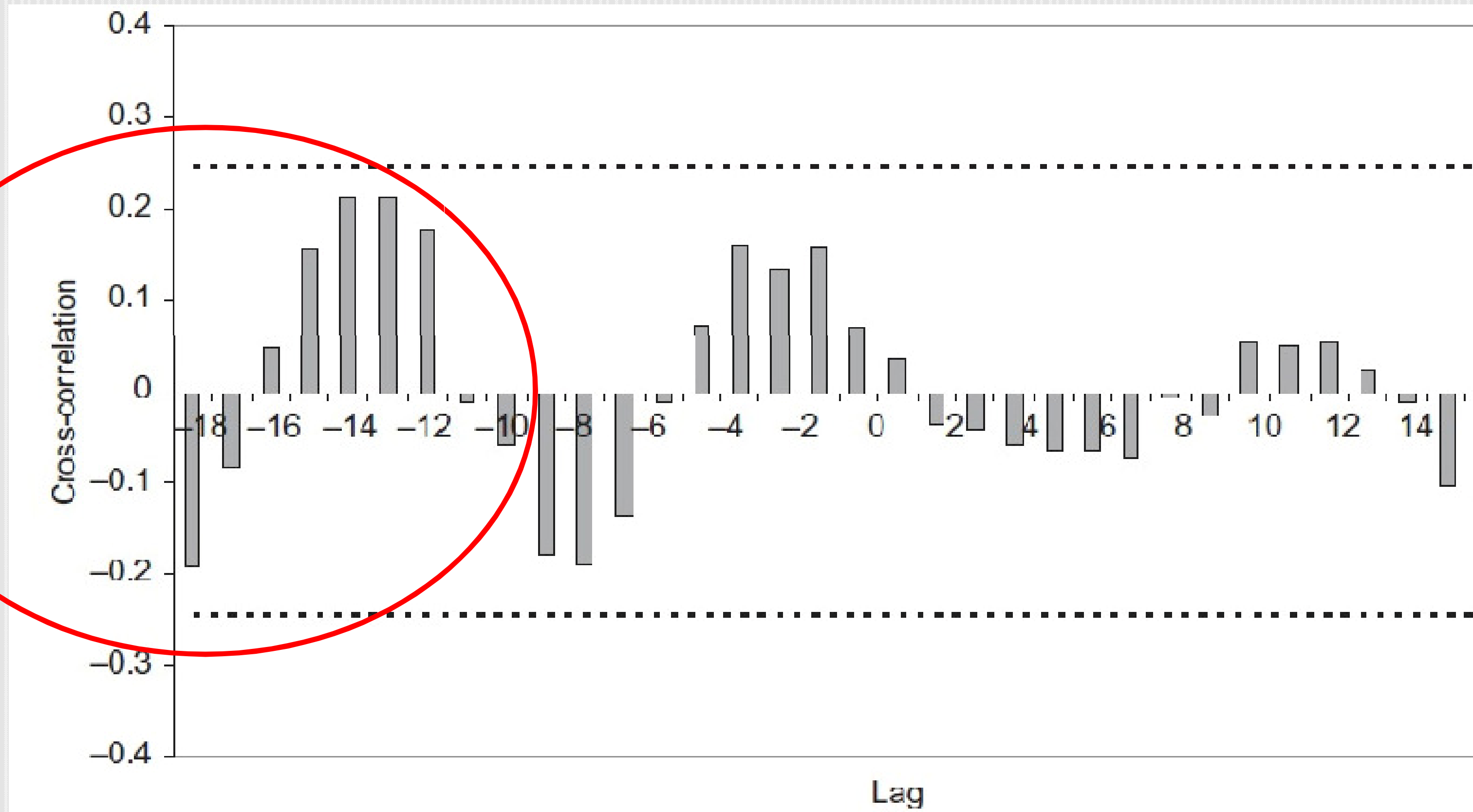
As the common harmonic contributes very little to the entire spectra, it is not logical to go for a linear regression.

Furthermore, the correlation between the two time series is 0.037, whic indicates the absence of any linear association between the two time series.

ss Correlation

Cross-correlation is a measure of similarity of two waveforms a function of a time-lag applied to one of them.

- present the cross-ation between SNs infall amounts.
- observed that the correlation nts always lie in nge [−0.2,0.2].
- further indicates sence of any linear ation between the time series under leration.



$$\rho_{xy} = E\left[\left(X_n - \mu_X\right)\left(Y_{n+m} - \mu_Y\right)\right] / \left(\sigma_X \sigma_Y\right)$$

As τ_5 has been revealed as common to both of the time series, we can try to determine some relationships within a period of 64/5=12.8 years.

In cross-correlation between SN and rainfall, we find that there are significant positive and negative spikes at lags separated by approximately 5.

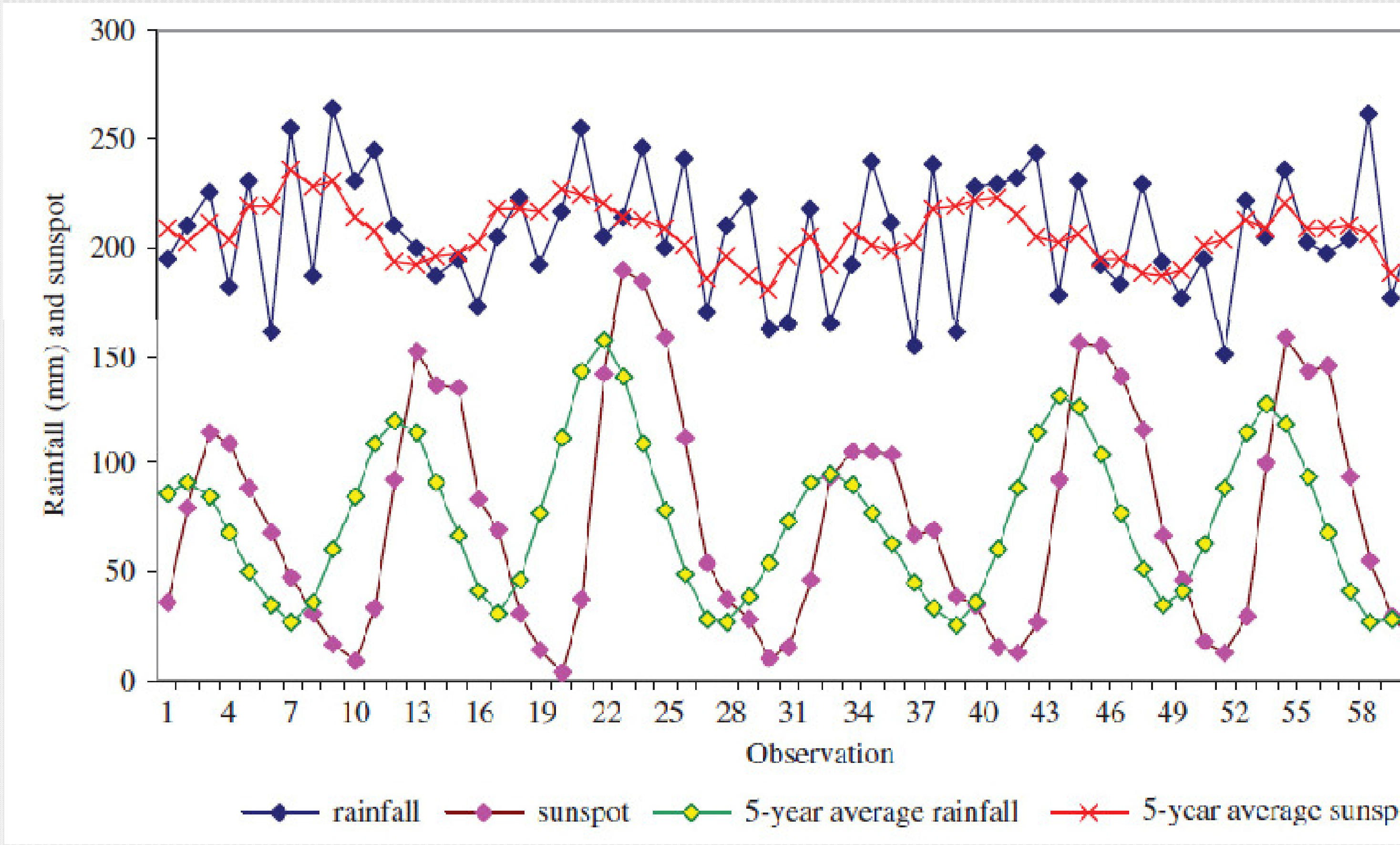
This means that, in five time steps, the correlation is changing from extremely negative to extremely positive.

Taking all these into account, we decided to test a model based on 5-year averages of the observations.

ng Average

that moving the average
rs gives more smoothness
me series. Thus, we try to
e the 5-year average
on rainfall amount over
ing the corresponding 5-
erage SNs as predictor.

ent the actual time series
se after taking the 5-year
. It can be seen that the
ries are becoming
er after taking the 5-year
s.



nodel generation

sider a time series
ng to the study period
998. Thus, we have 64
each time series. We
this series by x_1, x_2, x_3,
x_{64}.

e original time series,
ve a new time series
ntries are

$$\frac{x_1 + x_2 + \ldots + x_5}{5},$$

$$\frac{x_2 + x_3 + \ldots + x_6}{5}$$

$$_0 = \frac{y_{60} + y_{61} + \ldots + y_{64}}{5}$$

Advantages of ANN

(see: SILVERMAN, D. and DRACUP, J.A., 2000, Artificial neural networks and long-
precipitation prediction in California. *Journal of Applied Meteorology*, ***39*, pp. 57–66**

- 1. A priori knowledge of the underlying process is not required.
- 2. Existing complex relationships among th various aspects of the process under investigation need not be recognized.
- 3. Solution conditions, such as those requir by standard optimization or statistical mod are not preset.
- 4. Constraints and a priori solution structur are neither assumed nor enforced.

ine of an ANN

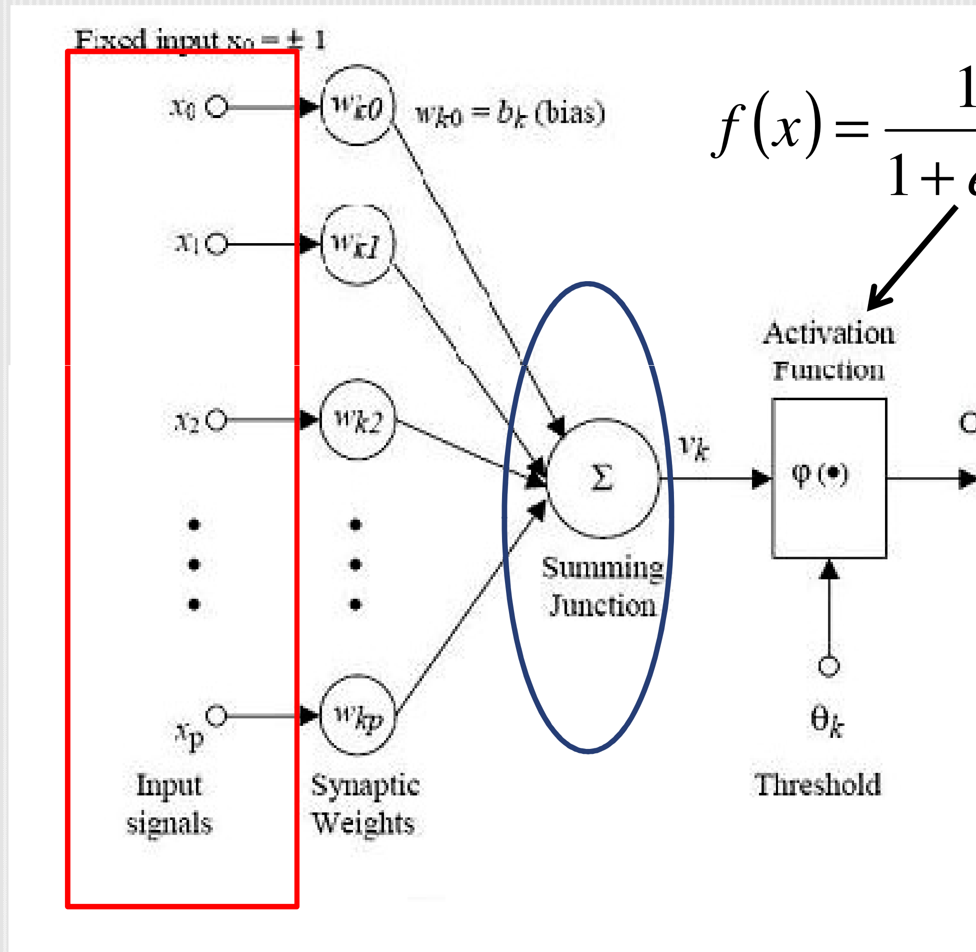
s of the theory of ANNs are available in Bishop Rojas (1996) and in many other Geophysical systems are generally characterized ous degrees of chaos and non-linearity .

ous examples are available where the ANN has mplemented in forecasting atmospheric mena. Some significant examples include er and Dorling (1998), Hsieh and Tang (1998).

r MW, Dorling SR. 1998. Multilayer Perceptron- a review of ions in atmospheric sciences. *Atmospheric Environment* –2636.

.W. and Tang, B., 1998, Applying neural network models ction and data analysis in meteorology and graphy. *Bulletin of the American Meteorological Society*, **79, 5–1870.**

Outline of a Multilayer Perceptron (MLP)



P learning

the advent of the
pagation algorithm a vast
of improvements has been
the technique for supervised
of MLP. The backpropagation
n looks for minimum error
in weight space using the
of gradient descent.

$$= w_k + \eta d_k$$

$$= -\nabla E(w_k)$$

$$) = -\gamma \frac{\partial E}{\partial w_k} + \alpha \Delta w_k (l-1)$$

Rojas R. 1996. *Neural Networks: A Systematic Introduction. Springer-* Verlag New York, Inc.: New York.

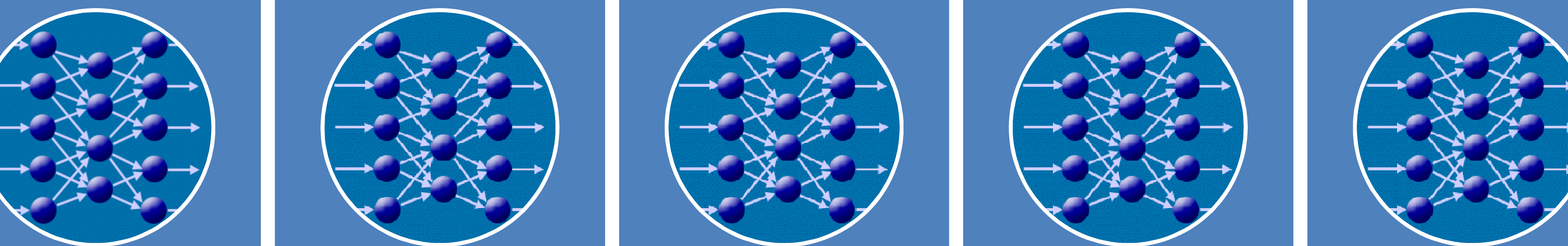
Medsker LR, Jain LC. 2000. *Recurrent Neural Networks: Design and Applications. CRC Press: Boca Raton, FL.*

This equation represents an iteration process that finds the opt weight vector by adapting the initial weight vector wo. This adaptation is performed by presenting to the network seque a set of pairs of input and target vectors.

The positive constant *η* is called the learning rate. The direction vector *the negative* gradient of the output error function *E.*

the i-th correction for weight w_k
γ and α are the learning and momentum rate respectively, and *E* denotes the error.

The input matrix for the generation of the ANN-based prediction model is of the order of
60 × 2), where the first column pertains to the predictor (i.e. the SN) and the second colun
pertains to the predictand (i.e. the monsoon rainfall).



m the data series,
% of the data are
sen as the training
and the remaining
% of the data are
sen as the test set.

An adaptive gradient learning in the cascade mode is used with a maximum allowable hidden layer size of 60.

The sigmoid function ($f(x) = (1 + e^{-x})^{-1}$) is taken as the activation function.

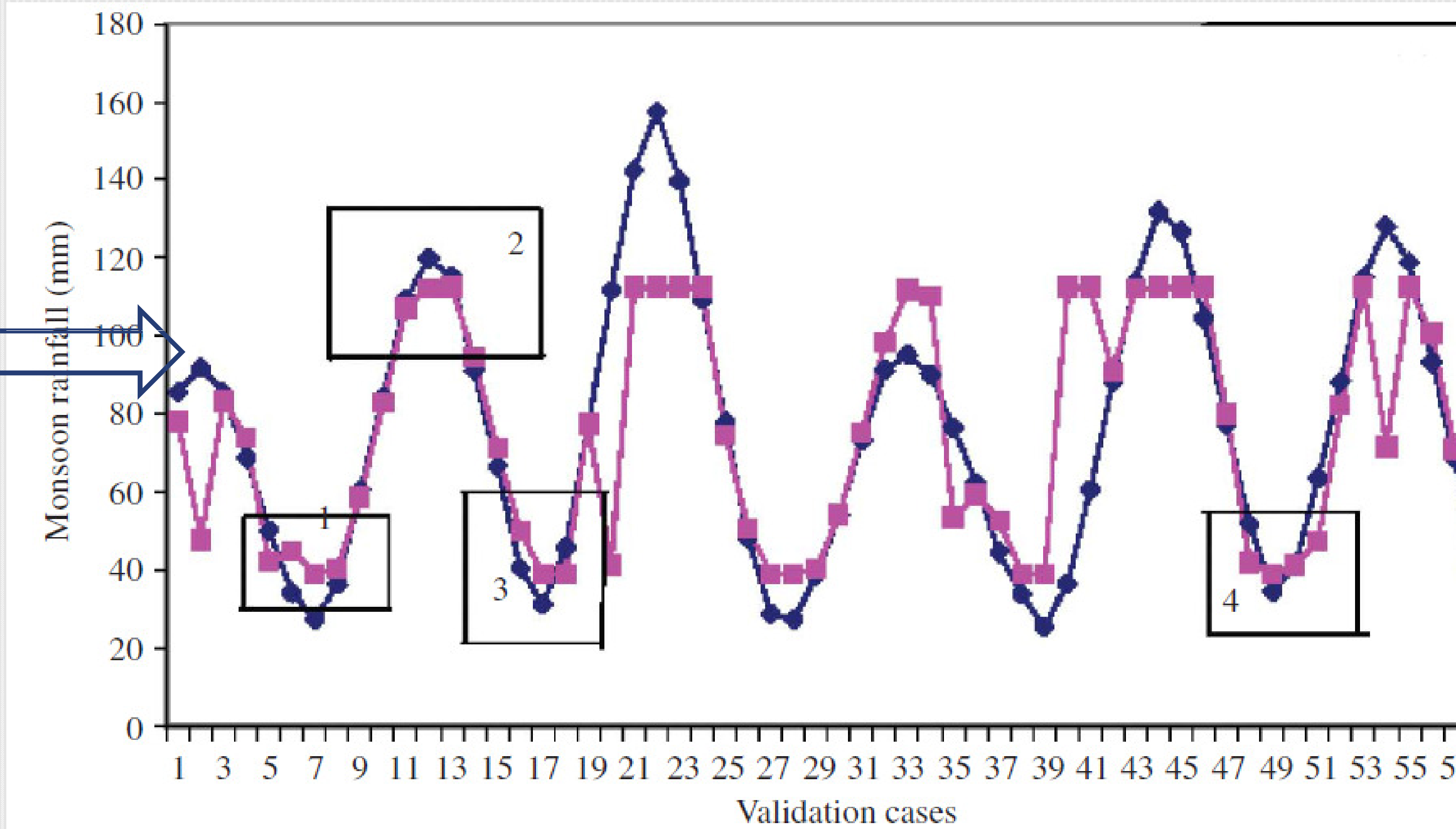
Minimization of the mean squared error is chosen as the stopping criterion.

After training an
testing the netwo
and validating ov
the entire set of
data, the optimu
network architect
is available with se
hidden nodes.

mott's Index

$$d = 1 - \left[\sum_i |P_i - O_i|^2\right]\left[\sum_i \left(|P_i - \overline{O}| + |O_i - \overline{O}|\right)^2\right]^{-1} \quad 0 \le d$$

nd that d =0.795 in
esent model. This
tes, in general, a
nodel.
nodel outputs are
ted here, where all
idation cases are
ted.
s figure, it is found
various validation
the actual and
ted values have
coincided.
ny cases, although
re not coincident, a
ssociation is visible.



fusion matrix

sion matrix was generated to
prediction capacity of the
odel.

sion matrix is a square matrix
ows and columns represent the
es for the real-world target
del outputs, respectively.

ue in the (i,j) position of the
s the number of records for
he real-world target output is
n subrange and whose real-
odel output is within the jth
e.

	Actual subranges				
Predictand subranges	25.61	51.98	78.34	104.7	131.07
25.61	16	2	1	1	0
51.98	2	10	1	1	0
78.34	0	1	7	0	0
104.7	2	1	2	9	4
131.1	0	0	0	0	0
Totals	20	14	11	11	4

This table shows that for the subrange 25.61–51.98 mm there are 20 observations and 16 predictions. Thus, there is 80% accuracy for this subrange. For the subrange 51.975–78.34 mm, there are 14 observation 10 predictions, which means that there is 71% accuracy for this subrang Thus, it can be interpreted that for lower rainfall amounts there is more possibility of an accurate forecast by ANN than in the case of higher rai amounts.

Concludi[n] remarks

analysed the autocorrelation structures, which showed a wave tern in the case of SN. However, in the case of the monsoon rainfall, such pattern was discerned.

Considering previous studies on the association between solar activities and climate, we tried to find some association despite the apparent non-linearity implied by the very low cross-correlation.

Implementing spectral analysis we obtained a common prominent harmonic between the two time series under consideration. The common harmonic was the fifth harmonic, which implies a common cycle of 12.8 years.

We decided to implement an ANN in the form of multilayer perceptrons with sigmoid non-linearity.

We derived new time series by averaging the observations for 5 years. Finally, it was revealed that by the implementation of the ANN, an average of 5-year SNs can give some estimate for the 5-year averaged summer monsoon rainfall over India.

Thank You