Evaluating the predictability of central Indian rainfall on short and long timescales using theory of nonlinear dynamics

Uday Pratap Singh, Ashok Kumar Mittal, Suneet Dwivedi
and Anurag Tiwari

ABSTRACT

The theoretical and practical understanding of projected changes in rainfall is desirable for planning and adapting to climate change. In this study, finite size Lyapunov exponents (FSLE) are used to study error growth rates of the system at different timescales. This is done to quantify the impact of enhanced anthropogenic greenhouse gas emissions on the predictability of fast and slow varying components of central Indian rainfall (CIR). The CIR time series for this purpose is constructed using the daily gridded high-resolution India Meteorological Department (IMD) dataset and Coupled Model Inter-comparison Project phase 5 (CMIP5) output for historical run and three representative concentration pathways (RCP2.6, RCP4.5, and RCP8.5) from the HadGEM2-ES, IPSL-CM5A-LR, CCSM4, BCC-CSM1.1, and MPI-ESM-LR models. The analyzed CIR dataset reveals a low dimensional chaotic attractor, suggesting that CIR requires a minimum of 5 and maximum of 11 variables to describe the state of the system. FSLE analysis shows a rapid decrease in the Lyapunov exponent with increasing timescales. This analysis suggests a predictability of about 2–3 weeks for fast varying components at short timescale of the CIR and about 5–9 years for slow varying components at long timescales.

Key words | CIR predictability, CMIP5, embedding dimension, finite size Lyapunov exponents, nonlinear time series analysis

Uday Pratap Singh (corresponding author)
Suneet Dwivedi
Anurag Tiwari
K. Banerjee Centre of Atmospheric and Ocean
   Studies,
University of Allahabad,
Allahabad,
India
E-mail: *udaykbcaos@gmail.com*

Ashok Kumar Mittal
Department of Physics,
University of Allahabad,
Allahabad,
India

Suneet Dwivedi
M. N. Saha Centre of Space Studies,
University of Allahabad,
Allahabad,
India

INTRODUCTION

The predictability study of rainfall or another atmospheric variable has its own importance. The projection from the CMIP5-based model outputs for historical as well as future scenarios has enabled us to utilize these data to study the predictability of the system for historical as well as future scenarios. This quantification of the impact of the increase in greenhouse forcing into the atmospheric predictability can be possible with these long-term data. The influence of anticipated changes in greenhouse gases on climate predictability over timescales of several decades to centuries can be explored using global models (Meehl *et al.* 2009).

The Lyapunov exponent of a dynamical system measures the average rate at which nearby trajectories of a system diverge. The positive exponent provides evidence that the underlying dynamical system is chaotic and there is a limit on the predictability of the dynamical system, i.e., the dynamical system may be predictable over some finite time periods. The nonlinear dynamical method – finite size Lyapunov exponent (FSLE) – is used for quantifying the predictability. It has several advantages over the traditional methods, since nonlinear dynamical methods like largest Lyapunov exponent (LLE) (Rosenstein *et al.* 1993; Kantz 1994) and Kolmogorov–Sinai entropy quantify the error growth rate exhibited by a dynamical system (Boffetta *et al.* 2002). The predictability analysis based on LLE holds for infinitesimal perturbations, which are linked with

Corrected Proof

2   U. P. Singh *et al.* | Evaluating predictability of central India rainfall on short/long time scales   **Journal of Water and Climate Change** | in press | 2019

errors in the smallest, fastest-evolving scales of a dynamical system. However, the smallest errors, if finite, soon grow to nonlinear saturation, thereafter not playing any role in the error growth (Boffetta *et al.* 1998b; Cencini *et al.* 1999; Siqueira & Kirtman 2014) of fast scale dynamics. On the other hand, relatively larger errors have slower error growth rates on larger timescales. Thus, LLE cannot estimate the predictability on longer timescales (Toth & Kalnay 1993; Boffetta *et al.* 1998a, 1998b). This advantage of FSLE over LLE motivated us to use the method for quantifying the predictability of both long and short timescales.

Aurell *et al.* (1997) introduced the FSLE as a measure of the divergence of nearby trajectories, to quantify predictability of processes developing on different timescales. Boffetta *et al.* (1998a) further advanced the concept of FSLE and investigated its applicability to detect large-scale (slow varying) dynamical properties of a measured time series characterized by different timescales. They showed that the large-scale error growth rate due to slow varying components of a dynamical system can be estimated by FSLE, even though data points are limited and embedding dimension is moderate. The FSLE has been used to estimate predictability on both fast and slow timescales (Boffetta *et al.* 1998a; Cencini & Vulpiani 2013; Siqueira & Kirtman 2014). The climate system is supposed to be a system with many characteristic timescales. In a study by Boffetta *et al.* (1998a), dynamics of a system with two different timescales were analyzed using the FSLE method. It was found that the system provides two plateaus, one for short-scale perturbation, which reveals FSLE of fast varying component λ_f and the second plateau at large perturbation, which gives FSLE of slow varying component λ_s. They also investigated the role of short-scale parameterization by replacing the true dynamics of the fast-varying component with the stochastic process and found that even though the fast dynamics is not resolved it does not have any effect on correct computation of FSLE of slow varying components. It is also pointed out in their study that, even with a smaller number of data points and also with a moderate embedding dimension, the dynamics of slow varying components can be practically captured. Siqueira & Kirtman (2014) investigated the predictability of ENSO using FSLE on the fast timescale. On the other hand, Mitchell & Gottwald (2012) suggested that FSLE can well approximate the predictability at the slow

timescale. It is suggested in their study that predictability at the slower timescale can be much higher than the local LLE λ_{max}, since short-scale instabilities grow faster but become nonlinearly saturated at a much smaller time than large-scale instabilities. Thus, in the present study, the strength of FSLE over LLE to efficiently extract predictability at both fast and slow timescales motivated us to quantify the predictability of the CIR on both fast and slow time. The scale at small perturbation is referred to as short timescale and the scale at large perturbation, where the value of λ_s saturated (plateau) is referred to as long timescale.

Studies about the impact of global warming on the Indian monsoon at different timescales are very limited. Mani *et al.* (2009) studied the impact of global warming over the medium range prediction. They showed that the medium range prediction of Indian monsoon will be less predictable due to global warming (Mani *et al.* 2009). They addressed the impact of climate change over medium range prediction but did not study its impact over long timescales. The short-scale predictability of CIR is about 2–3 weeks, but the clear limit on the predictability and role of climate change on long timescales is not well established. Therefore, in the present study, an attempt is made to quantify the predictability of CIR for historical as well as future projections using the theory of nonlinear dynamics on both short and long timescales.

There are several other studies, both on toy model and real-world scenarios, which suggest that the inclusion of forcing increases the predictability of the system (Lorenz model – Palmer 1993; Dwivedi *et al.* 2007; Singh *et al.* 2015). Kirtman *et al.* (2013a) have pointed out that external forcing like aerosol concentration, greenhouse gases, volcanic eruption, etc. provides predictability to the system, ranging from sub-seasonal to decadal timescales. Latif (2013) states that, on longer lead times of several decades, external forcing will cause weakening in Atlantic meridional overturning circulation, which will give rise to predictability. Luo *et al.* (2011) inferred that the influence of global warming results in an increase in predictability of the system on seasonal to inter-annual timescales.

The climate system is a combination of several timescales and the longer timescale components play an important role in climate change (Bigg *et al.* 2003). Climatic variables like rainfall also have variability from sub-seasonal

Corrected Proof

3 U. P. Singh et al. | Evaluating predictability of central India rainfall on short/long time scales | Journal of Water and Climate Change | in press | 2019

to multidecadal timescales. Climate patterns such as the El-Niño Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), etc. reveal climate variability (Stocker 2013) and the Indian monsoon is found to be influenced by ENSO and the decadal oscillations like PDO and AMO (Krishnamurthy & Krishnamurthy 2016, 2017). The ocean surface variables play a vital role in the climate variability on shorter timescales, for example, SST changes leading to El Niño and La Niña events. On the other hand, the slow varying nature of deep ocean currents produces little influence on climate variability at shorter timescales (Kirtman et al. 2013a). However, on longer timescales, these deep ocean currents may play a significant, even a dominant role (Bigg et al. 2003; Marshall & Zanna 2014).

The fast-varying components of the atmosphere effect predictability at short timescales and the slow varying components of the atmosphere effect predictability at long timescales. The study of predictability at slow timescale is important because future emission scenarios have less influence over near-term predictions than long-term projections (Kirtman et al. 2013b). Over longer periods, these effects can accumulate to cause a critical change (Doblas-Reyes et al. 2006). For example, subtle yearly changes in the thickness of ocean ice in polar locales (influencing the surface temperature of the ice and its albedo) may cause long-term cumulative impacts driving climate change. Therefore, slow processes, which do not influence variability on seasonal to inter-annual timescales, may become important in influencing long-term climate variability and hence predictability (Hurrell et al. 2010; Solomon et al. 2011). At long timescales, the predictability of the system is primarily determined by the ocean, because it has large thermal capacity and longer time for adjustment (Slingo & Palmer 2011).
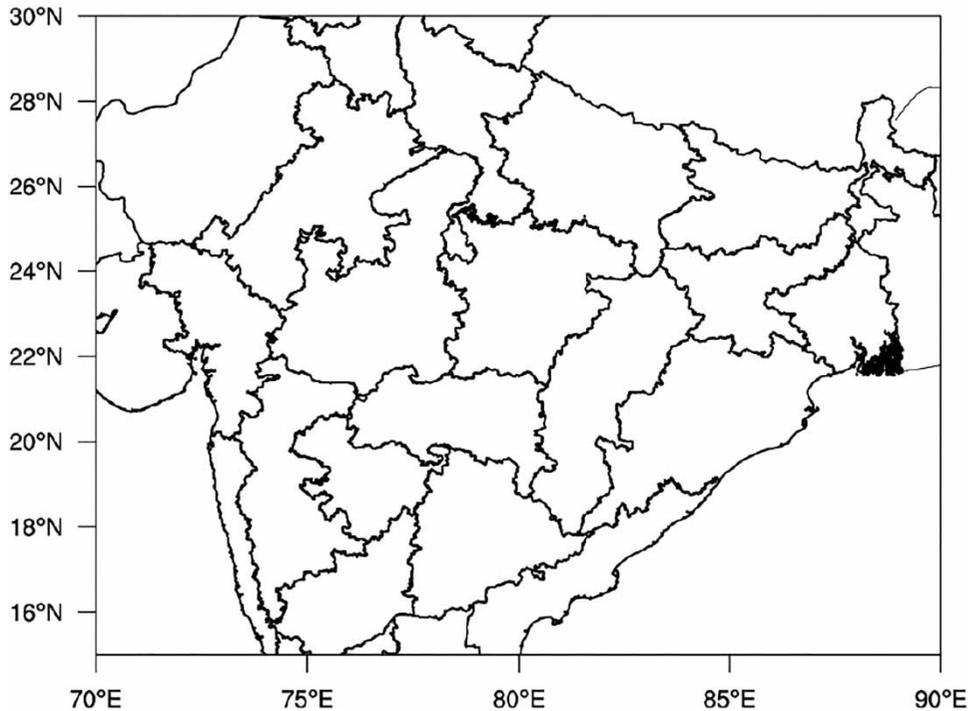
The predictability horizon depends on the initial condition and on the variable chosen (Anderson et al. 2009). The predictability of fast-varying components is generally governed by the internally generated atmospheric variability, i.e. weather, which loses its memory in around a month (Lorenz 1963; Shukla 1981). On the other hand, the predictability of slow-varying components varies at decadal timescales (Tang et al. 2008). However, the absolute limit on the predictability of slow-varying components is not well established. The Hasselmann mechanism suggests that the 'scale of predictability is restricted to the order of the

autocorrelation time of the slower component of the system, usually the ocean' (NRC 1998). That is, the long-term memory of the ocean offers some hope for initialized decadal-scale prediction and predictability (Li et al. 2017). There are a large number of other studies which support the presence of predictability of atmospheric variations on multiyear timescale (Griffies & Bryan 1997a, 1997b; Collins et al. 2006; Boer & Lambert 2008; Meehl et al. 2009, 2014; DelSole et al. 2013; Ding et al. 2016; Wu et al. 2016).

The rainfall observation data from IMD and the data from IPCC CMIP5 model output of five models have been used. The main aim of this study is to assess the impact of increase in anthropogenic forcing on the predictability of short and long timescales of CIR using FSLE. Our result for the short timescale is in agreement with the result of Mani et al. (2009). The predictability knowledge of rainfall under future emission scenarios will empower preparedness for optimal water resource usage and disaster management. This paper is organized as follows. Below, the study region and datasets used are described, followed by the methodology adopted. Then, the next section shows the results of the calculated embedding dimensions and FSLE for both fast and slow timescales of IMD observation and CMIP5 model data. The final section explains the conclusions drawn from the results.

STUDY REGION AND DATASETS

Time series of the daily central Indian rainfall (CIR) is constructed by area (spatial) averaging the daily gridded (0.25×0.25) IMD dataset over the monsoon trough region, (Figure 1) [15N–30N; 70E–90E]. A lowpass filter of 10 days is applied in each case in order to filter high frequency components having frequency up to 10 days. The choice of 10-day filter is made in order to preserve the information with the smallest loss and produce clearer information from the data, so that the high frequency noise component from the data is filtered. On the other hand, without applying the filter, due to the presence of high frequency components in the data, correlation exponent (v) vs embedding dimension (m) is increasing with the increase in m, which is a behavior of random noise. Therfore, choice of the filtering high frequency components from the data is

Corrected Proof

4   U. P. Singh *et al.* | Evaluating predictability of central India rainfall on short/long time scales   **Journal of Water and Climate Change** | in press | 2019

Figure 1 | Map of study region (15N–30N; 70E–90E) central India (monsoon trough region).

justified. The time series are also created using similar methodologies for Historical and RCP Scenarios of the IPCC AR5 CMIP5 models (Taylor *et al.* 2012), namely, HadGEM2-ES, IPSL-CM5A-LR, CCSM4, BCC-CSM1.1, and MPI-ESM-LR. Historical simulations are based on solar and volcanic forcing, land use, observed concentrations of greenhouse gases, and reconstructed aerosol emissions. Future projections are based on the three representative concentration pathways (RCPs) (Moss *et al.* 2010). In the RCP2.6 pathway the peak in radiative forcing reaches 3 W m^{-2} before the year 2100 and afterwards it declines to 2.6 W m^{-2} in the year 2300. RCP4.5 represents the pathways for which radiative forcing reaches 4.5 W m^{-2} in the year 2300. Similarly, RCP8.5 is the pathway for which radiative forcing reaches 8.5 W m^{-2} by the year 2300. Most of the GCMs are unable to simulate the mean, intra-seasonal variability of summer monsoon rainfall. According to Ashfaq *et al.* (2017), among the GCMs under study, only the MPI-ESM-LR model is able to simulate both rainfall maxima in July and less than 25% bias in precipitation for the south Asian summer monsoon during the historical time 1970–1999. However, the model

selection is done on the basis of availability of long-term data to get the required data for time series reconstruction. Tsonis *et al.* (1993) have shown that $10^{(2+0.4D)}$ number of data points are sufficient to correctly estimate the correlation dimension, therefore, the number of data points listed in Table 1 are adequate for faithful computation on correlation dimension. Therefore the data up to the year 2300 is used, which is a combination of RCP up to 2100 and after that, extended RCP up to 2300; for simplicity it is termed RCP throughout this paper. Table 1 lists the number of data points and durations that have been used and Table 2 lists the model description.

Table 1 | Number of data points used in the CIR time series and time duration

Scenario	Duration	Years	Data points
Historical	1860–2005	146 years	53,290
RCP2.6	2006–2300	295 years	107,675
RCP4.5	2006–2300	295 years	107,675
RCP8.5	2006–2300	295 years	107,675
IMD observation	1901–2013	113 years	41,273

Corrected Proof

5 | U. P. Singh *et al.* | Evaluating predictability of central India rainfall on short/long time scales | Journal of Water and Climate Change | in press | 2019

Table 2 | Model descriptions

Modeling center (or group)	Model
Met Office Hadley Centre	HadGEM2-ES
Institute Pierre-Simon Laplace	IPSL-CM5A-LR
National Center for Atmospheric Research	CCSM4
Beijing Climate Center, China Meteorological Administration	BCC-CSM1.1
Max Planck Institute for Meteorology	MPI-ESM-LR
Indian Metrological Department	IMD observation

METHODOLOGY

From a measured time series, phase space reconstruction is done by using the method of time delay as described in the section immediately below. The time delay embedding was done for the total appended time series. Appropriate time delay is estimated as described in the section Mutual information. The method used for determining the embedding dimension (m) that efficiently accommodates the attractor is then described. This is followed by a section describing Determination of finite size Lyapunov exponents. Figure 2 shows the work flow of the methods shown below.

Phase space reconstruction

The phase space reconstruction of a time series is done by embedding it in an m-dimensional space. In the case of numerical simulations of mathematical models, the number of variables required for constructing the phase space of the system is known. However, this is not known for experimental data obtained from a real physical dynamical system, e.g., fluid dynamics (Ding *et al.* 2014). The reconstructed trajectory, X, can be expressed by a matrix, where each row is a phase space vector (Rosenstein *et al.* 1993), i.e.:

$$X = [X_1, \ X_2, \ \ldots, X_t, \ \ldots, X_m]^T \tag{1}$$

where X_t is the state of the system at discrete time t. Let x_t, $t = 0, 1, 2, \ldots, N$ represent the single variable time series of precipitation data. An m-dimensional phase space may be reconstructed from this time series using the method of time delay (Packard *et al.* 1980; Takens 1981). Let:
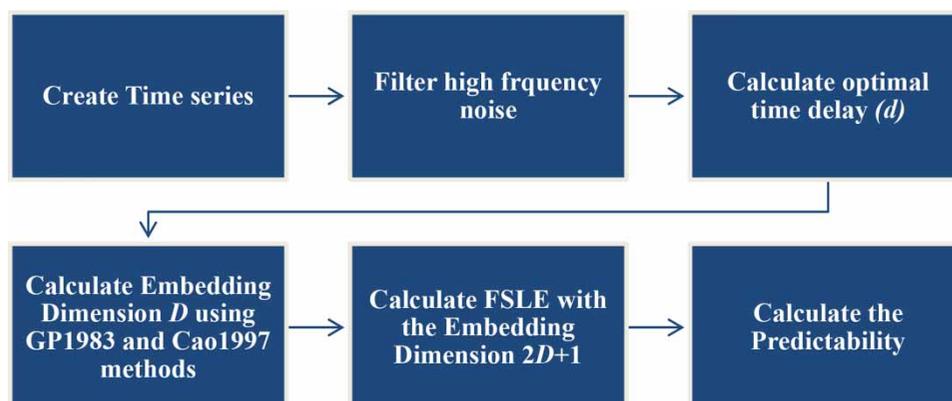
$$X_t(m) = (x_t, \ x_{t+d}, x_{t+2d}, \ \ldots, x_{t+(m-1)d}) \tag{2}$$

where d is the appropriately chosen time delay and m is the embedding dimension. Thus, X is an $M \times m$ matrix, where

$$M = N - (m - 1) \times d \tag{3}$$

Mutual information

The limitation of the autocorrelation method for calculating time delay is that it is a linear statistic and does not account for nonlinear correlations. The mutual information method is then chosen for the calculation of suitable time delay since it captures nonlinear correlation. The first minimum of the mutual information is then used as an optimal time

Figure 2 | Flow diagram for the method described in the Methodology section.

Corrected Proof

6 U. P. Singh *et al.* │ Evaluating predictability of central India rainfall on short/long time scales │ Journal of Water and Climate Change │ in press │ 2019

delay (d) (Fraser & Swinney 1986; Siqueira & Kirtman 2014). The mutual information is calculated by using the formula:

$$I(t) = \sum_{i=1}^{N} \sum_{j=1}^{N} P_{ij}(t) \ln P_{ij}(t) - 2 \sum_{i=1}^{N} P_i \ln P_i \qquad (4)$$

where P_i is the probability that $X(t)$ is in interval i and P_{ij} is the joint probability that the time series $X(t)$ is found to be in interval i and after Δt time $X(t + \Delta t)$ is found in interval j (Battiti 1994; Sprott 2003). Figure 3 shows mutual information $I(t)$ vs steps *(k)* for IMD rainfall time series over central India, and the inset panel shows $d = 72$ (first minima of $I(t)$).

Embedding dimension

The calculation of embedding dimension m to acquire the minimum number of variables sufficient to model the dynamics of the system is done by using two methods, those of Grassberger & Procaccia (1983) (termed GP1983) and Cao (1997) (termed Cao 1997). Details about both methods are given in the following two subsections, respectively.

Embedding dimension using GP1983

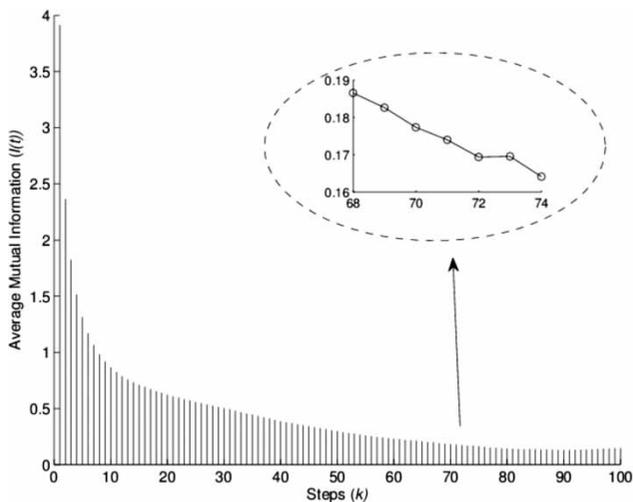The minimum number of variables required to map the system dynamics is estimated by using correlation

Figure 3 │ Mutual information $I(t)$ vs steps *(k)* for IMD rainfall time series over central India. Inset panel shows delay = 72 (first minima of $I(t)$).

dimension. Consider a single variable time series x_i, $i = 1, 2, …, N$ of length N. The method described earlier is used for phase space reconstruction for different values of m. The correlation sum $C(r)$ in an m-dimensional phase space is defined as the ratio of the number of pairs of embedded data points having distance less than r to the total number of pairs (GP1983).

Mathematically:

$$C(r) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{i,j=1}^{\infty} \theta(r - |X_i - X_j|) \qquad (5)$$

where $\theta(x)$ denotes the Heaviside step function, which is equal to zero for negative values and one for other values. This method counts the number of pairs (X_i, X_j) having distance $|X_i - X_j|$ less than r and $N(N-1) \approx N^2$ is the total number of pairs for N data points. The limit $N \to \infty$ ensures the characterization of the entire attractor. The correlation dimension is defined by:

$$D = \lim_{r \to 0} \frac{\log C(r)}{\log r} \qquad (6)$$

This mathematical definition cannot be used for experimental data. However, it is often found that $C(r)$ estimated from a reasonably large dataset of size N satisfies a scaling law $C(r) \sim r^\nu$ for a range of r. The scaling exponent ν is called the correlation exponent and for a particular embedding dimension m can be calculated from slope of the scaling region of the $\log C(r)$ vs $\log r$ graph.

The correlation exponent ν is computed for different embedding dimensions m. A graph of ν vs m is plotted. If the value of ν saturates to a constant value ν_s as m is increased, the time series is considered to be chaotic. In this case, the dynamics of the system can be said to reside on a low-dimension attractor of correlation dimension $D = \nu_s$. On the other hand, if ν increases without saturation, as m is increased, the analyzed time series is stochastic because a stochastic process embedded in an m dimensional space always fills that space (Tsonis & Elsner 1990). The data size, sampling frequency, and noise present in the data are the main factors that affect estimation accuracy of the correlation dimension.

The smallest integer which is greater than the correlation dimension D gives the smallest number of suitably

Corrected Proof

7   U. P. Singh *et al.* │ Evaluating predictability of central India rainfall on short/long time scales   **Journal of Water and Climate Change** │ **in press** │ **2019**

chosen variables that can map the dynamics. The value of the embedding dimension m_s at which the value of the correlation exponent begins to saturate gives the minimum number of dimensions in which the attractor can be embedded completely (Fraedrich 1986). Abarbanel *et al.* (1990) has pointed out that although Taken's reconstruction theorem requires $D > 2D_s + 1$ to ensure faithful representation of the D_s dimensional attractor as seen in D-dimensional embedding space, in practice, $D > D_s$ is usually sufficient.

Minimum embedding dimension using Cao 1997

Cao (1997) introduced a new method based on false nearest neighbor (FNN) which determines the minimum embedding dimension of a scalar time series. This method is also free from the restriction of the number of data points. This method uses Equation (2) for phase space reconstruction and can be defined as follows:

$$a(i, m) = \frac{||X_i(m + 1) - X_{n(i,m)}(m + 1)||}{||X_i(m) - X_{n(i,m)}(m)||} \tag{7}$$

where $i = 1,2,3\ldots\ldots M$, and $||\ldots||$ is measure of Euclidean distance and mostly chosen to be maximum norm, which can be given as:

$$||X_k(m) - X_i(m)|| = \max_{0 \leq j \leq m-1} |x_{k+j\times d} - x_{i+j\times d}| \tag{8}$$

$X_i(m + 1)$ is the ith reconstructed vector in the $m + 1$ embedding dimension and $n(i.m)$ is an integer in the range $1 \leq n(i,m) \leq N\text{-}m \times d$, such that $X_{n(i,m)}(m)$ is the nearest neighbor of $X_i(m)$ in the m dimensional reconstructed phase space. The integer $n(i,m)$ in Equation (7) is the same as that of the denominator. The second nearest neighbor is used, if $X_{n(i,m)}(m) \approx X_i(m)$. The mean value of all $a(i,m)$ is denoted by a quantity $E(m)$, which is defined as:

$$E(m) = \frac{1}{N - m \times d} \sum_{i=1}^{N-m\times d} a(i, m) \tag{9}$$

This averaging ensures the elimination of the subjectivity involved with the FNN method (Kennel *et al.* 1992). This is due to the fact that the quantity $E(m)$ depends only

on the dimension m and lag d. The variation E_1 of this quantity E from m to $m + 1$, can be given by:

$$E_1(m) = E(m + 1)/E(m) \tag{10}$$

If $E_1(m)$ saturates at any m_0 then $m_0 + 1$ is the minimum embedding dimension, which can efficiently accommodate the attractor. To distinguish deterministic signals from stochastic signals the following quantity is calculated:

$$E^*(m) = \frac{1}{N - m \times d} \sum_{i=1}^{N-m\times d} |x_{i+m\times d} - x_{n(i,m)+m\times d}| \tag{11}$$

here, $n(i,m)$ has the same meaning defined above for Equation (7). Its variation from m to $m + 1$ can be given as:

$$E_2(m) = E^*(m + 1)/E^*_m \tag{12}$$

In the case of the time series from some random process, the $E_1(m)$ will not saturate as m is increased, which is due to the limited data samples. Thus, practically, it may be difficult to determine whether $E_1(m)$ is changing slowly or has stopped changing. In such a situation, $E_2(m)$ will be useful because for random data future values are independent of the past values and $E_2(m)$ will be equal to 1 for any m, whereas for deterministic signals there exists some m such that $E_2(m) \neq 1$.

Finite size Lyapunov exponent (FSLE)

The finite size Lyapunov exponent $\lambda(\delta)$ is based on error growth time $\tau_r(\delta)$, which is the smallest time it takes for a perturbation of initial size δ to grow by a factor r (>1, Cencini *et al.* 1999). The ratio r should not be too large, in order to avoid the growth through different scales. In this study $r = 2$ is chosen for each case, and the error doubling time is denoted by τ. The finite size Lyapunov exponent is then defined by:

$$\lambda(\delta) = \frac{\ln 2}{\langle \tau(\delta) \rangle} \tag{13}$$

where τ is the time in which the separation between neighboring trajectories grows from δ to 2δ and $\langle \ldots \rangle$ is the average over many realizations, so that:

$$\langle \tau(\delta) \rangle = \frac{\sum_{i=1}^{n} \tau_i(\delta)}{n} \tag{14}$$

Corrected Proof

8 U. P. Singh *et al.* | Evaluating predictability of central India rainfall on short/long time scales | Journal of Water and Climate Change | in press | 2019

The direction of the initial perturbation must align with the most unstable direction in the phase space. This can be ensured by considering two neighboring trajectories that have separation very much less than δ. The two trajectories are followed until at time denoted by $t = 0$ the separation is δ and at time $t = \tau_1$ the separation is 2δ. It can be shown (Aurell *et al.* 1997) that:

$$\frac{1}{\langle \tau(\delta) \rangle} = \langle \frac{1}{\tau(\delta)_t} \rangle \tag{15}$$

where $\langle \ldots \rangle_t$ denotes the natural measure along the trajectory by putting $A = 1/\tau$ in

$$\langle A_t \rangle = \frac{1}{T} \int_0^T A(t)dt = \frac{\sum_i A_i \tau_i}{\sum_i \tau_i} = \frac{\langle A\tau \rangle}{\langle \tau \rangle} \tag{16}$$

The FSLE $\lambda(\delta)$ were computed from the time series datasets using the method described by Boffetta *et al.* (1998a, 1998b). The computation of FSLE from measured time series dataset is done by sequentially applying subsequent steps. First, a phase space reconstruction is performed using the method described earlier and the time delay is calculated with the method described in the section Mutual information. Second, a threshold series of perturbation scales $\delta_n = r^n \delta_0$ is taken. A value $\delta_{min} \ll \delta_0$ is chosen so that perturbations of size less than δ_{min} align with the most unstable direction in the phase space. Then, for each reconstructed vector X_n, its nearest neighbors are obtained. Lastly, from these nearest neighbors, only the nearest neighbors (X_m), which have separation $\delta = |X_n - X_m|$, smaller than δ_{min}, are chosen. The trajectories starting from reconstructed vector X_n and its nearest neighbor X_m are used to calculate the error doubling time $\tau(\delta)$ by measuring the growth time of perturbation from δ_n to δ_{n+1}. The growth of perturbation is followed from initial size less than δ_{min} to the maximum value δ_{max} in the threshold series. Finally, The FSLE $\lambda(\delta_n)$ at perturbation size δ_n is obtained by averaging the error doubling time $\tau(\delta_n)$ for several nearest neighbors X_m and using Equation (13). A measure of the predictability of a system is the error doubling time. It is thus proportional to the inverse of the FSLE (Mitchell & Gottwald 2012).
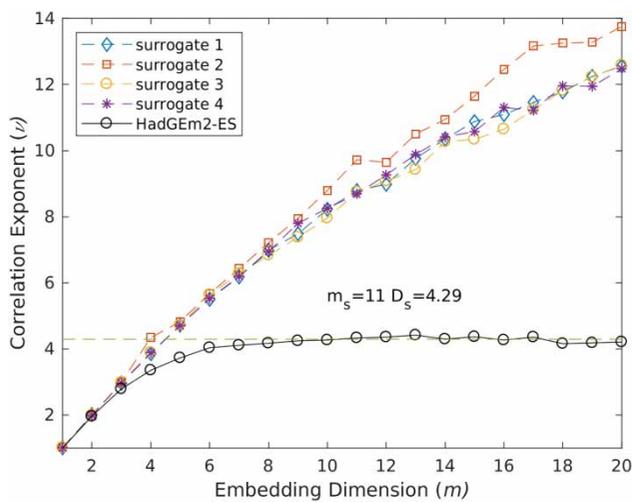
RESULTS

The nonlinear time series analysis of different datasets – CIR time series for IMD observation, historical, and various RCP scenarios of CMIP5 model outputs over central India – is performed using the method described in the section Methodology. The method described in Equations (5) and (6) (GP1983) is used for calculation of the correlation dimension D to obtain embedding dimension (m) for each of the datasets. In Figure 4(a), the plot of log $C(r)$ vs log(r) is shown for IMD CIR data with the embedding dimension m ranging from 1 to 20 (from left to right). The slope of the straight line within the appropriately chosen scaling region for each m gives the correlation exponent v corresponding to that m. In Figure 4(b), the plot between correlation exponent v and embedding dimension m shows that D_s saturates to a value of 4.46 at about $m_s = 11$. Thus, the CIR data reveal the existence of a low dimensional chaotic attractor having saturation dimension $D_s = 4.46$. Hence, the next positive integer above D_s, i.e., 5, is the minimum number of variables sufficient to describe the dynamics of the system (Fraedrich 1986). On the other hand, to confirm the deterministic features in data the surrogate data test is performed (Theiler *et al.* 1992). The four ensembles of surrogate time series were constructed by
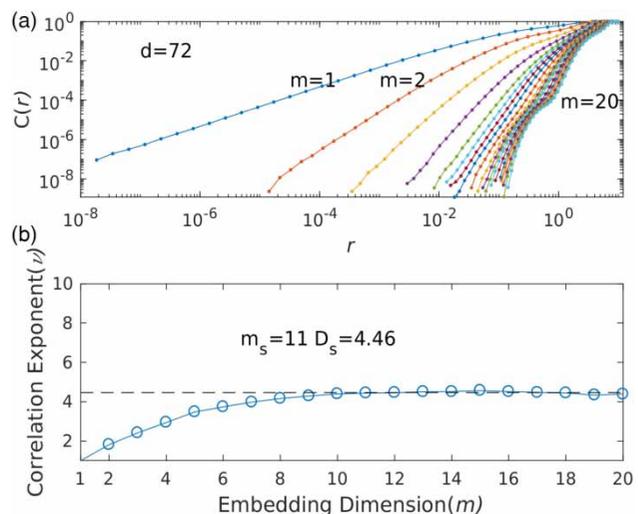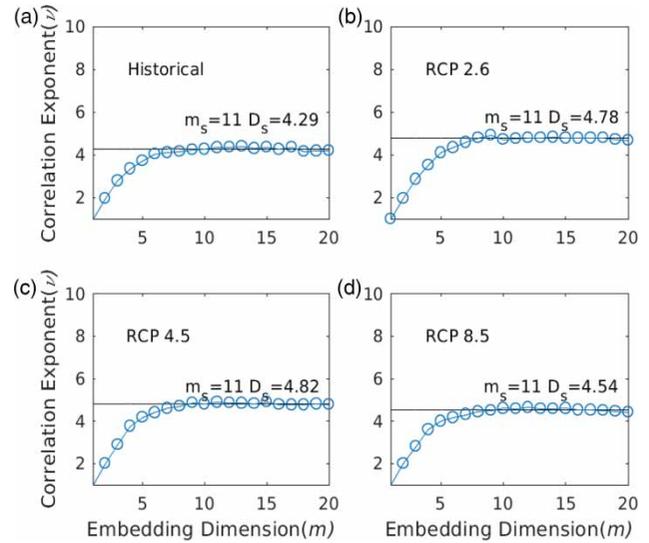

Figure 4 | (a) *C(r)* vs *r* plot of IMD rainfall time series over central India. (b) Correlation exponent (*v*) vs embedding dimension (*m*) plot of IMD rainfall time series over central India. The next positive integer above saturation dimension D_s, i.e., $D = 5$, is the minimum number of variables sufficient to describe the dynamics of the system.

shuffling randomly the indices of HadGEM2-ES model historical time series data and then the method of correlation dimension on ensemble surrogate time series is applied. There is no saturation found in these cases, as shown in Figure 5. We also performed surrogate data tests on the rest of the data and similar results were found (figure not shown). Thus, the convergence of v with m is a genuine characteristic of the data and not an artifact of the method used, confirming the deterministic features of the data. The GP1983 method is also chosen for calculation of correlation dimensions of all the other CMIP5 models listed in Table 2. Figure 6 shows the plot of the correlation exponent (v) vs embedding dimension (m) for historical as well as all the RCP scenarios for the HadGEM2-ES model of CMIP5 (figures for the rest of the models are not shown). The fractal dimension of the historical and RCP scenarios for different models of CMIP5 have been calculated and listed in Table 3. From Table 3, it is clear that similar results have been obtained for all IPCC AR5 CMIP5 models. In the present study, D from Table 3 is found to be 5 for all the models. To validate the correct computation of minimum dimension that is sufficient to represent the dynamics of the system the Cao1997 method is used. Figure 7 shows E_1 and E_2 quantities and E_1 gives the number of minimum variables sufficient to represent the dynamics of the system and is

Figure 6 | Correlation exponent (v) vs embedding dimension (m) plot of HadGEM2-ES for all four scenarios. Figure indicates a minimum of 5 and maximum of 11 variables are sufficient to efficiently accommodate the attractor.

found to be 5 (saturation value, $m_0 + 1$) for the HadGEM2-ES model. It is similar to embedding dimension m calculated from the GP1983 method (Table 3). The E_2 quantity is also calculated, as mentioned in the Cao1997 method (represented by diamonds in Figure 7) to check whether the datasets are the outcome of the deterministic or stochastic processes. E_2 is also found to be less than 1 for each dimension m, which ensures the dataset under study results from a deterministic process. To check whether the Cao1997 method is able to capture stochasticity in the surrogate dataset, the method is also applied on the surrogate dataset obtained from the method described above for the HadGEM2-ES RCP2.6 scenario. Figure 8 clearly shows that the $E_2 \geq 1$ as m increases, so the method is able to capture stochasticity of the surrogate dataset and it is also evident from the figure that the E_1 also do not saturate, which is expected for the stochastic dataset. The embedding dimension for the rest of the models listed in Table 2 is also calculated using the Cao1997 method and similar results are evident as in the GP1983 method, i.e., embedding dimension $m = 5$ and the E_2 quantities for these models also suggest the deterministic nature of datasets since $E_2 \leq 1$ for each case (figure not shown).

The calculation of the FSLE $\lambda(\delta)$ is done by using Equation (13) to be on the safe side embedding dimension equal to $2D + 1$, which is 11 for all datasets (as mentioned

Figure 5 | Correlation exponent (v) vs embedding dimension (m) plot of four ensemble surrogate data formed by shuffling the indices of HadGEM2-ES historical scenario rainfall time series over central India. The value of v increases as m increases for all the four surrogates; on the other hand, for HadGEM2-ES historical time series D_s saturates to a value of 4.29. Thus, the convergence of v with m confirms the deterministic features of the data.

Corrected Proof

10   U. P. Singh *et al.* │ Evaluating predictability of central India rainfall on short/long time scales   Journal of Water and Climate Change │ in press │ 2019

Table 3 │ Correlation dimensions of IMD and various scenarios of CMIP5 model

|	Correlation dimension									
	HadGEM2-ES		IPSL-CM5A-LR		CCSM4		BCC-CSM1		MPI-ESM-LR	
Scenario	D_s	m_s	D_s	m_s	D_s	m_s	D_s	m_s	D_s	m_s
Historical	4.29	11	4.47	11	4.51	11	4.32	11	4.27	11
RCP2.6	4.78	11	4.58	11	4.71	11	4.82	11	4.75	11
RCP4.5	4.82	11	4.65	11	4.81	11	4.87	11	4.82	11
RCP8.5	4.54	11	4.68	11	4.85	11	4.85	11	4.91	11
IMD	4.46			11						
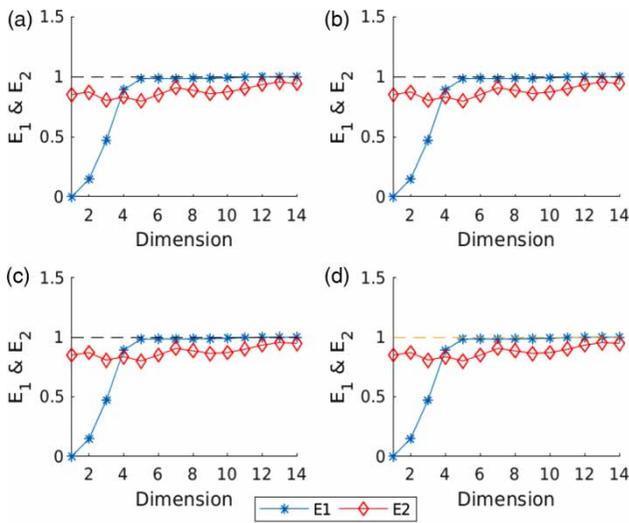
Figure 7 │ E_1 and E_2 vs dimension plot of HadGEM2-ES model for all four scenarios: (top left panel: historical; top right panel: RCP2.6; bottom left panel: RCP4.5; bottom right panel: RCP8.5). The quantity E_1 (represented by star) starts to saturate at $m_0 = 4$ so that $m = 5$ is the minimum embedding dimension. It is also evident that $E2 < 1$ (represented by diamond) for each case, thus the time series are deterministic.
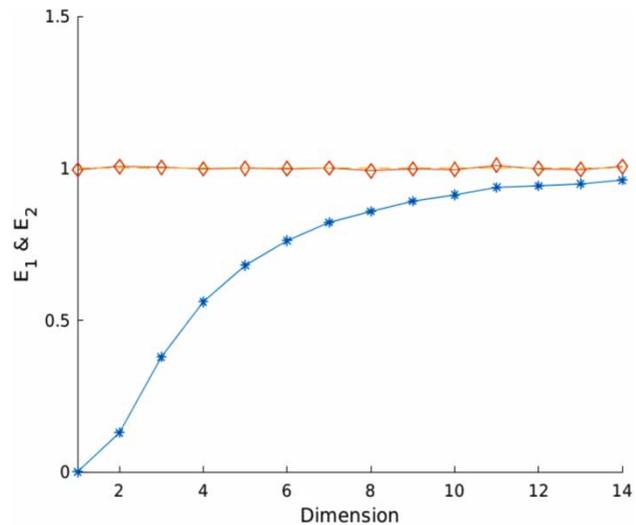
Figure 8 │ E_1 and E_2 vs dimension plot for surrogate data of HadGEM2-ES RCP2.6 scenario. The quantity E_1 (represented by star) is not saturated as m increases and it is also evident that $E2 \geq 1$ (represented by diamond) for each case, thus the method is able to capture stochasticity in the dataset.

before, D is the next positive integer above saturation dimension D_s), is used as the embedding parameter for the calculation of FSLE. The FSLE can give information about how a finite initial uncertainty δ grows with time. The perturbation threshold $\delta_{min} = 0.07$ is selected. The k-nearest neighbor method is used for the computation of nearest neighbors and the Euclidian distance is used as a distance parameter. The 20 nearest neighbors corresponding to each point in the reconstructed phase space were calculated and the nearest neighbors which have distance less than δ_{min} are chosen for calculation of FSLE (for instance, 23,305 nearest neighbors found below δ_{min} in the case of

HadGEM2-ES historical scenario). It is also verified that increasing the number of nearest neighbors greater than ten, corresponding to each point in the reconstructed phase space, does not contribute to an increase in the number of nearest neighbors below δ_{min}, thus it is the optimal choice for the current analysis. The FSLE $\lambda(\delta)$, as a function of δ for the IMD rainfall observed data, is shown in Figure 9. The FSLE $\lambda(\delta)$ for the short scale, i.e., fast-varying component (λ_f) of IMD rainfall observed data saturates to a value 0.0343 (Table 4). The error doubling time obtained by using Equation (13) can be calculated from the formula $\boldsymbol{ln}(2)/(\lambda_f)$, and is around 20.02 days which is

Corrected Proof
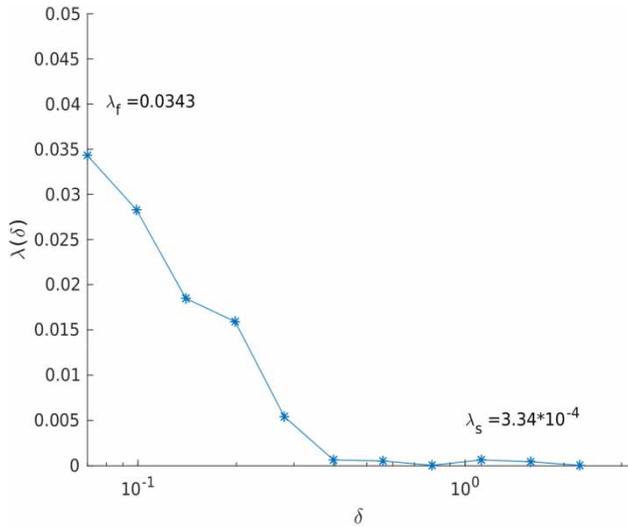
Figure 9 | The FSLE, $\lambda(\delta)$, as a function of scale δ for the IMD observation. The λ_f and λ_s (plateau) are the values of FSLE at short timescale and long timescale, respectively.

Table 4 | FSLE of IMD and various scenarios of CMIP5 model for small perturbations

	Finite size Lyapunov (λ_f) exponent for small perturbations				
Scenarios	HadGEM2-ES	IPSL-CM5A-LR	CCSM4	BCC-CSM	MPI-ESM-LR
Historical	0.0352	0.0461	0.0373	0.0359	0.0324
RCP2.6	0.0373	0.0425	0.0411	0.0345	0.0484
RCP4.5	0.0360	0.0393	0.0445	0.0351	0.0480
RCP8.5	0.0315	0.0338	0.0324	0.0352	0.0398
IMD	0.0343				

Table 5 | FSLE of IMD and various scenarios of CMIP5 model for large perturbations

	Finite size Lyapunov (λ_s) exponent at large perturbations ($\times 10^{-4}$)				
Scenarios	HadGEM2-ES	IPSL-CM5A-LR	CCSM4	BCC-CSM1	MPI-ESM-LR
Historical	3.78	3.60	3.67	3.41	3.62
RCP2.6	3.61	3.00	3.12	2.73	3.38
RCP4.5	3.11	2.47	3.14	2.58	2.43
RCP8.5	2.81	2.12	2.96	2.32	2.47
IMD	3.34				
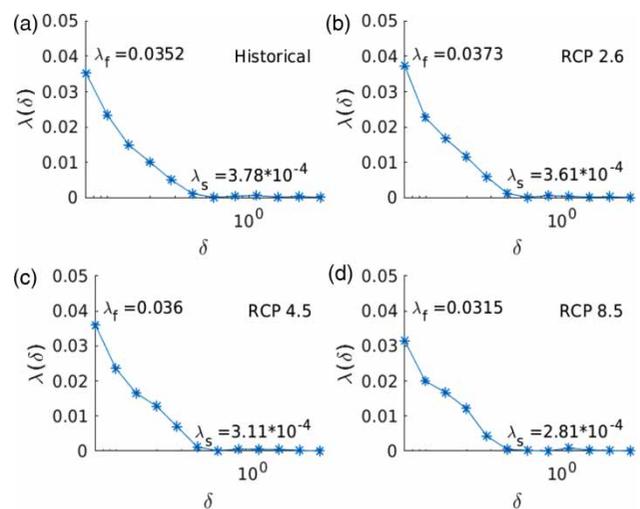
Figure 10 | The FSLE, $\lambda(\delta)$, as a function of scale δ for all the four scenarios of HadGEM2-ES. The λ_f and λ_s are the values of FSLE at short timescale and long timescale, respectively.

in agreement with the weather predictability found by Dwivedi (2012). This quantifies the predictability time for fast variables. The FSLE $\lambda(\delta)$ for large perturbations, that evolve according to slow-varying components of IMD rainfall dynamics, saturate to a value of 3.34×10^{-4} (Table 5), which indicates a predictability time of 5.7 years (Table 7) by using formula $\mathbf{\mathit{ln}}(2)/(365 \times \lambda_s)$. Our analysis of the rainfall data suggests that the data can be the outcome of two weakly coupled dynamical systems, one having a fast timescale and the other having a slow timescale. The FSLE $\lambda(\delta)$ for the historical and all RCP scenarios from all five models have been calculated by using the same method as described above for IMD rainfall observed data and, for brevity, only one figure for computation of FSLE $\lambda(\delta)$ for the HadGEM2-ES model is shown in Figure 10. (Figures

for the calculation of FSLE for the rest of the models are shown in Figures S1 to S4.)

The FSLE of historical and of future RCP projections for HadGEM2-ES, IPSL-CM5A-LR, CCSM4, BCC-CSM1.1, and MPI-ESM-LR models at small perturbations are shown in Table 4. The FSLE results for all the models and IMD observation for larger perturbations are listed in Table 5. The positive values of the FSLE confirm the chaotic nature of CIR time series (Hilborn 2000; Sprott 2003). Tables 6 and 7 show the estimated predictability times for different CIR datasets on fast and slow timescales, respectively. The predictability of the fast-varying component of CIR for all the datasets examined was found to be between 14 and 22 days. On the other hand, the predictability for

Table 6 | Predictability (in days) of the fast-varying components (short timescale), calculated by using formula $ln(2)/(\lambda_f)$, where values of λ_f are listed in Table 4

	Predictability (in days) of the fast-varying component				
Scenarios	HadGEM2-ES	IPSL-CM5A-LR	CCSM4	BCC-CSM1	MPI-ESM-LR
Historical	19.7	15.0	18.6	19.3	21.4
RCP2.6	18.6	16.3	16.9	17.0	14.3
RCP4.5	19.3	17.6	15.6	18.6	14.4
RCP8.5	22.0	20.5	21.4	16.0	17.4
IMD	20.2				

Table 7 | Predictability (in years) of the slow varying components (long timescale), calculated by using, $ln(2)/(365 \times \lambda_s)$, where values of λ_s are listed in Table 5

	Predictability (in years) of the slow-varying component				
Scenarios	HadGEM2-ES	IPSL-CM5A-LR	CCSM4	BCC-CSM1	MPI-ESM-LR
Historical	5.0	5.3	5.2	5.6	5.2
RCP2.6	5.3	6.3	6.1	7.0	5.6
RCP4.5	6.1	7.7	6.0	7.4	7.8
RCP8.5	6.8	9.0	6.4	8.2	7.7
IMD	5.7				

slow-varying component of CIR for IMD observation and all historical scenarios was found to be around five to six years while the predictability of the CIR under the RCP scenarios was around six to nine years.
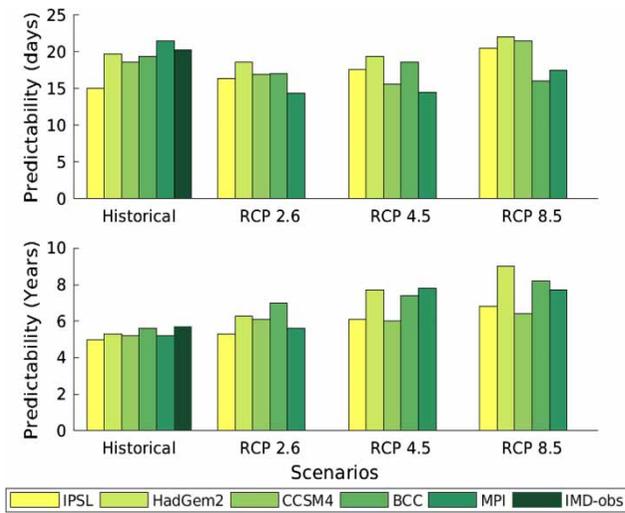
DISCUSSION AND CONCLUSIONS

India is an agroeconomic country. Most of the areas in the central Indian region are rainfed. Therefore, production and yield of Kharif (April to September (summer)) and Rabi season (October to March (winter)) crops (like rice, wheat, etc.) majorly depend on the CIR. Nowadays, groundwater is heavily used in several places in central India for irrigation, but too much exploitation of groundwater resources has caused diminution of this finite resource. The increase in population, food demand, urbanization, and industrialization have caused some serious concerns over fresh water availability (Mall *et al.* 2007), and fresh water resources are likely to decrease due to climate

change (Gosain *et al.* 2006). Climate change is supposed to increase extreme events, which could potentially alter the rainfall distribution in terms of time and space. The increase in aridity due to climate warming is likely to add to irrigation water demand (Hanasaki *et al.* 2013; Leng *et al.* 2015). These factors will cause serious concerns about water-related disasters over central Indian regions. Therefore, revealing the potential impact of climate change on the predictability of CIR has great significance.

The correlation dimensions for the different datasets of the study are determined by the technique described in the section Embedding dimensions. It is evident that a low dimensional chaotic attractor exists for all the datasets and a minimum of 5 and a maximum of 11 variables will be able to simulate the system (Table 3). Similarly, low dimensional attractors from rainfall data have been found by several investigations in the range of 4–9 (Dhanya & Kumar 2010; Jothiprakash & Fathima 2013). The results from Table 3 do not indicate any clear increasing or decreasing trend in the values of correlation dimensions as a result of climate change. In other words, the correlation dimensions obtained for different RCP scenarios of a particular CMIP5 model do not present any clear influence of increase in radiative forcing on the correlation dimension. It is clear from the table that IMD, historical as well as all future scenarios can be considered as residing on a strange attractor of non-integer dimensions (Mandelbrot 1982; Tsonis & Elsner 1989; Hilborn 2000).

The FSLE for small perturbations is substantially in the same range for all the datasets (Table 4). The limit on predictability for the fast-varying component, obtained from the growth rate of small perturbations, is found to be about 2–3 weeks for IMD rainfall, historical scenarios and the RCP scenarios of all the CMIP5 models (Table 6). The predictability estimates for the fast component are similar to those reported in the literature in the range of 2–3 weeks (Goswami & Xavier 2003; Waliser *et al.* 2003; Webster & Hoyos 2004; Dwivedi *et al.* 2006; Rai & Krishnamurthy 2011; Dwivedi 2012; Abhilash *et al.* 2014). Figure 11 (top panel) shows the potential predictability of the fast-varying components of CIR in the historical scenario and IMD observation is greater than the RCPs. This implies a short timescale will be less predictable due to global warming and this result is consistent with the result of Mani *et al.* (2009), although some models suggest increased

Corrected Proof

13   U. P. Singh *et al.* │ Evaluating predictability of central India rainfall on short/long time scales   Journal of Water and Climate Change │ in press │ 2019

Figure 11 │ Top panel: Predictability in days for fast-varying component of the CIR. Figure shows a decrease in the predictability in RCP2.6 and 4.5 than in the historical scenario but a few models in RCP8.5 suggest a slight increase in the potential predictability. Bottom panel: Predictability in years for slow-varying component of the CIR. Figure shows an increase in the potential predictability from historical to RCP8.5 due to anthropogenic forcing and longer memory of the oceans.

predictability of a few days in RCP8.5 for the fast-varying components. A clear reason for this could not be established.

The FSLE obtained for the slow-varying component of CIR (Table 5) for all the historical data and the IMD observation data are significantly larger than the FSLE obtained for each of the RCP scenarios. Most of the models suggest a reduction in the value of λ_s with the increase in radiative forcing. It can be concluded from the results that the predictability of slow-varying component of CIR for RCP scenarios is in the range of six to nine years (Table 7), whereas for historical data and for IMD observational data, the range is between five and six years. Thus, the lower panel of Figure 11 shows that the increase in radiative forcing from 2.6 W m^{-2} to 8.5 W m^{-2} will increase the predictability of CIR. Therefore, the CIR under the RCPs will be more predictable than the historical scenarios in long timescales for central India. This increase in the predictability of slow-varying component under RCPs might be because of the increase in the CO_2 concentration from RCP2.6 to RCP8.5. This will lead to increased anthropogenic forcing, which could be the possible source of the potential predictability as suggested by Kirtman *et al.* (2013a), that external forcing like aerosol concentration, greenhouse gases, volcanic eruption, etc. provide predictability to the system, ranging from

sub-seasonal to decadal timescales. Using a conceptual forced Lorenz model, Palmer (1993) demonstrated that inclusion of forcing results in increase in the predictability, in the regime residency time of the Lorenz attractor. This means that the dynamical system is confined to a specific region in phase space for a longer time, leading to increased predictability and suggesting that the climate response will be more predictable even in the presence of weak forcing. Dwivedi *et al.* (2007) and Singh *et al.* (2015) also used a conceptual forced Lorenz model and found that the forcing causes increase in predictability. This result from conceptual models is supported by studies on real-world data which suggest that external forcing such as aerosols and greenhouse warming may contribute to the long-term predictability (Meehl *et al.* 2010; Jia & DelSole 2011; Kirtman *et al.* 2013a). Beyond the monthly timescales, the atmospheric predictability arises from the forcing induced by slowly varying boundary conditions like SST, soil moisture, etc. (Wu *et al.* 2016).

However, these results should be taken with the caveat that the climate models are still in their evolutionary stage and substantial improvement is still required for better climate change information and improved ability to track the time evolution of the inherent long-scale variability in combination with the forced response.

ACKNOWLEDGEMENTS

REFERENCES

Abarbanel, H. D. I., Brown, R. & Kadtke, J. B. 1990 Prediction in chaotic nonlinear systems: methods for time series with

Corrected Proof

14   U. P. Singh *et al.* | Evaluating predictability of central India rainfall on short/long time scales   Journal of Water and Climate Change | in press | 2019

broadband Fourier spectra. *Physical Review A.* **41** (4), 1782–1807. https://doi.org/10.1103/PhysRevA.41.1782.

Abhilash, S., Sahai, A. K., Pattnaik, S., Goswami, B. N. & Kumar, A. 2014 Extended range prediction of active-break spells of Indian summer monsoon rainfall using an ensemble prediction system in NCEP Climate Forecast System. *International Journal of Climatology* **34**, 98–113. https://doi.org/10.1002/joc.3668.

Anderson, D. L. T., Doblas-Reyes, F. J., Balmaseda, M. & Weisheimer, A. 2009 *Decadal Variability: Processes, Predictability and Prediction*. ECMWF Technical Memorandum No. 591, 47 pp, ECMWF, Reading, UK.

Ashfaq, M., Rastogi, D., Mei, R., Touma, D. & Leung, R. 2017 Sources of errors in the simulation of south asian summer monsoon in the CMIP5 GCMs. *Climate Dynamics* **49**, 193–223. https://doi.org/10.1007/s00382-016-3337-7.

Aurell, E., Boffetta, G., Crisanti, A., Paladin, G. & Vulpiani, A. 1997 Predictability in the large: an extension of the concept of Lyapunov exponent. *Journal of Physics A: Mathematical and General* **30**, 1–26.

Battiti, R. 1994 Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks* **5** (4), 537–550. https://doi.org/10.1109/72.298224.

Bigg, G. R., Jickells, T. D., Liss, P. S. & Osborn, T. J. 2003 The role of the oceans in climate. *International Journal of Climatology* **23** (10), 1127–1159.

Boer, G. J. & Lambert, S. J. 2008 Multi-model decadal potential predictability of precipitation and temperature. *Geophysical Research Letters* **35**, L05706. https://doi.org/10.1029/2008GL033234.

Boffetta, G., Crisanti, A., Paparella, F., Provenzale, A. & Vulpiani, A. 1998a Slow and fast dynamics in coupled systems: a time series analysis view. *Physica D: Nonlinear Phenomena* **116** (3), 301–312.

Boffetta, G., Giuliani, P., Paladin, G. & Vulpiani, A. 1998b An extension of the lyapunov analysis for the predictability problem. *Journal of Atmospheric Science* **55** (23), 3409–3416.

Boffetta, G., Cencini, M., Falcioni, M. & Vulpiani, A. 2002 Predictability: a way to characterize complexity. *Physics Reports* **356** (6), 367–474.

Cao, L. 1997 Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D* **110** (1), 43–50. http://dx.doi.org/10.1016/S0167-2789(97)00118-8.

Cencini, M. & Vulpiani, A. 2013 Finite size Lyapunov exponent: review on applications. *Journal of Physics A: Mathematical and Theoretical* **46** (25), 254019.

Cencini, M., Falcioni, M., Vergni, D. & Vulpiani, A. 1999 Macroscopic chaos in globally coupled maps. *Physica D: Nonlinear Phenomena* **130** (1), 58–72. http://dx.doi.org/10.1016/S0167-2789(99)00015-9.

Collins, M., Botzet, M., Carril, A. F., Drange, H., Jouzeau, A., Latif, M., Masina, S., Otteraa, O. H., Pohlmann, H., Sorteberg, A., Sutton, R. & Terray, L. 2006 Interannual to decadal climate predictability in the North Atlantic: a multimodel-ensemble study. *Journal of Climate* **19**, 1195–1203. https://doi.org/10.1175/JCLI3654.1.

DelSole, T., Jia, L. & Tippett, M. K. 2013 Decadal prediction of observed and simulated sea surface temperatures. *Geophysical Research Letters* **40**, 2773–2778. doi:10.1002/grl.501859/2007GL032838.

Dhanya, C. T. & Kumar, D. N. 2010 Nonlinear ensemble prediction of chaotic daily rainfall. *Advances in Water Resources* **33** (3), 327–347.

Ding, H., Dong, W. & Wu, D. 2014 Chaotic features identification and analysis in Liujiang River runoff. In: *Intelligent Computing Theory* (D. S. Huang, V. Bevilacqua & P. Premaratne eds). ICIC 2014. Lecture Notes in Computer Science, Vol. 8588. Springer, Cham, Switzerland. doi: 10.1007/978-3-319-09333-8_71.

Ding, R., Jianping, L., Zheng, F., Feng, J. & Liu, D. 2016 Estimating the limit of decadal-scale climate predictability using observational data. *Climate Dynamics* **46**, 1563. https://doi.org/10.1007/s00382-015-2662-6.

Doblas-Reyes, F. J., Hagedorn, R., Palmer, T. N. & Morcrette, J. 2006 Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts. *Geophysical Research Letters* **33** (7). doi:10.1029/2005gl025061.

Dwivedi, S. 2012 Quantifying predictability of Indian summer monsoon intraseasonal oscillations using nonlinear time series analysis. *Meteorologische Zeitschrift* **21** (4), 413–419.

Dwivedi, S., Mittal, A. K. & Goswami, B. N. 2006 An empirical rule for extended range prediction of duration of Indian summer monsoon breaks. *Geophysical Research Letters* **33** (18), L18801.

Dwivedi, S., Mittal, A. K. & Pandey, A. C. 2007 Effect of averaging timescale on a forced Lorenz model. *Atmosphere-Ocean* **45** (2), 71–79.

Fraedrich, K. 1986 Estimating the dimensions of weather and climate attractors. *Journal of the Atmospheric Sciences* **43** (5), 419–432.

Fraser, A. M. & Swinney, H. L. 1986 Independent coordinates for strange attractors from mutual information. *Physical Review A* **33**, 1134. doi: 0.1103/PhysRevA.33.1134.

Gosain, A. K., Rao, S. & Basuray, D. 2006 Climate change impact assessment on hydrology of Indian river basins. *Current Science* **90**, 346–353.

Goswami, B. N. & Xavier, P. K. 2003 Potential predictability and extended range prediction of Indian summer monsoon breaks. *Geophysical Research Letters* **30** (18), 1966.

Grassberger, P. & Procaccia, I. 1983 Characterization of strange attractors. *Physical Review Letters* **50** (5), 346.

Griffies, S. M. & Bryan, K. 1997a Predictability of North Atlantic multidecadal climate variability. *Science* **275** (5297), 181–184.

Griffies, S. M. & Bryan, K. 1997b A predictability study of simulated North Atlantic multidecadal variability. *Climate Dynamics* **13** (7–8), 459–487.

Hanasaki, N., Fujimori, S., Yamamoto, T., Yoshikawa, S., Masaki, Y., Hijioka, Y., Kainuma, Y., Kanamori, Y., Masui, T. & Takahashi, K. 2013 A global water scarcity assessment under

Corrected Proof

15   U. P. Singh *et al.* | Evaluating predictability of central India rainfall on short/long time scales   **Journal of Water and Climate Change** | in press | 2019

shared socio-economic pathways–Part 2: water availability and scarcity. *Hydrology and Earth System Sciences* **17** (7), 2393–2413.

Hilborn, R. C. 2000 *Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers*. Oxford University Press, New York, USA.

Hurrell, J. W., Delworth, T., Danabasoglu, G., Drange, H., Drinkwater, K., Griffies, S., Holbrook, N., Kirtman, B., Keenlyside, N., Latif, M., Marotzke, J., Murphy, J., Meehl, G., Palmer, T., Pohlmann, H., Rosati, T., Seager, R., Smith, D., Sutton, R., Timmermann, A., Trenberth, K., Tribbia, J. & Visbeck, M. 2010 Decadal climate prediction: opportunities and challenges. In: *Proceedings of OceanObs&Rsquo;09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison and D. Stammer (Eds.), ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.45

Jia, L. & DelSole, T. 2011 Diagnosis of multiyear predictability on continental scales. *Journal of Climate* **24**, 5108–5124.

Jothiprakash, V. & Fathima, T. A. 2013 Chaotic analysis of daily rainfall series in Koyna reservoir catchment area, India. *Stochastic Environmental Research and Risk Aassessment* **27** (6), 1371–1381.

Kantz, H. 1994 A robust method to estimate the maximal Lyapunov exponent of a time series. *Physics Letters A* **185** (1), 77–87.

Kennel, M. B., Brown, R. & Abarbanel, H. D. 1992 Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A* **45** (6), 3403. doi: https://doi.org/10.1103/PhysRevA.45.3403.

Kirtman, B., Anderson, D., Brunet, G., Kang, I. S., Scaife, A. A. & Smith, D. 2013a Prediction from weeks to decades. In: G. R. Asrar & J. W. Hurrrell (eds), *Climate Science for Serving Society*, Springer, Dordrecht, the Netherlands, pp. 205–235. doi:10.1007/978-94-007-6692-1_8.

Kirtman, B., Power, S. B., Adedoyin, J. A., Boer, G. J., Bojariu, R., Camilloni, I., Doblas-Reyes, F. J., Fiore, A. M., Kimoto, M., Meehl, G. A., Prather, M., Sarr, A., Schär, C., Sutton, R., Oldenborgh, G. J., Vecchi, G. & Wang, H. J. 2013b Near-term climate change: projections and predictability. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex & P. M. Midgley, eds). Cambridge University Press, Cambridge, UK and New York, USA.

Krishnamurthy, L. & Krishnamurthy, V. 2016 Teleconnections of Indian monsoon rainfall with AMO and Atlantic tripole. *Climate Dynamics* **46**, 2269. https://doi.org/10.1007/s00382-015-2701-3.

Krishnamurthy, L. & Krishnamurthy, V. 2017 Indian monsoon's relation with the decadal part of PDO in observations and NCAR CCSM4. *International Journal of Climatology* **37**, 1824–1833. doi:10.1002/joc.4815.

Latif, M. 2013 The ocean's role in modeling and predicting decadal climate variations. In: *International Geophysics* (G. Siedler, S. M. Griffies, J. Gould & J. A. Church, eds). Academic Press, San Diego, CA, Ch. 25, Volume 103, pp. 645-665. http://dx.doi.org/10.1016/B978-0-12-391851-2.00025-8.

Leng, G., Huang, M., Tang, Q. & Leung, L. R. 2015 A modeling study of irrigation effects on global surface water and groundwater resources under a changing climate. *Journal of Advances in Modeling Earth Systems* **7**, 1285–1304. doi:10.1002/2015MS000437.

Li, S., Zhang, L. & Wu, L. 2017 Decadal potential predictability of upper ocean heat content over the twentieth century. *Climate Dynamics* **49**, 3293. https://doi.org/10.1007/s00382-016-3513-9.

Lorenz, E. N. 1963 Deterministic nonperiodic flow. *Journal of Atmospheric Science* **20**, 130–141. doi:10.1175/1520-0469(1963)020 < 0130:DNF > 2.0.CO;2.

Luo, J. J., Behera, S. K., Masumoto, Y. & Yamagata, T. 2011 Impact of global ocean surface warming on seasonal-to-interannual climate prediction. *Journal of Climate* **24** (6), 1626–1646.

Mall, R. K., Bhatia, R. & Pandey, S. N. 2007 Water resources in India and impact of climate change. *Jalvigyan Sameeksha* **22**, 157–176.

Mandelbrot, B. B. 1982 *The Fractal Geometry of Nature*. WH Freeman and Co, San Francisco, CA, USA.

Mani, N. J., Suhas, E. & Goswami, B. N. 2009 Can global warming make Indian monsoon weather less predictable? *Geophysical Research Letters* **36**, L08811. doi:10.1029/2009GL037989.

Marshall, D. P. & Zanna, L. 2014 A conceptual model of ocean heat uptake under climate change. *Journal of Climate* **27** (22), 8444–8465.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E. D., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D. & Stockdale, T. 2009 Decadal prediction: can it be skillful? *Bulletin of the American Meteorological Society* **90** (10), 1467. https://doi.org/10.1175/2009BAMS2778.1.

Meehl, G. A., Hu, A. & Tebaldi, C. 2010 Decadal prediction in the Pacific region. *Journal of Climate* **23** (11), 2959–2973.

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Reyes, F., Hawkins, E., Karspeck, A., Kimoto, M., Kumar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M., Rosati, T., Schneider, E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G. J., Vecchi, G. & Yeager, S. 2014 Decadal climate prediction: an update from the trenches. *Bulletin of the American Meteorological Society* **95**, 243–267. https://doi.org/10.1175/BAMS-D-12-00241.1.

Mitchell, L. & Gottwald, G. A. 2012 On finite-size lyapunov exponents in multiscale systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **22** (2), 023115(1-9). doi: 10.1063/1.4704805.

Corrected Proof

16   U. P. Singh *et al.* | Evaluating predictability of central India rainfall on short/long time scales   **Journal of Water and Climate Change** | **in press** | **2019**

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P. & Wilbanks, T. J. 2010 The next generation of scenarios for climate change research and assessment, *Nature* **463**, 747–756. https://doi.org/10.1038/nature08823.

National Research Council (NRC) (1998) *Decade-to-Century Scale Climate Variability and Change: A Science Strategy*. National Academy Press, Washington, DC, USA, 146 pp. (http://www.nap.edu).

Packard, N. H., Crutchfield, J. P., Farmer, J. D. & Shaw, R. S. 1980 Geometry from a time series. *Physical Review Letters* **45** (9), 712.

Palmer, T. N. 1993 Extended-range atmospheric prediction and the Lorenz model. *Bulletin of the American Meteorological Society* **74** (1), 49–65.

Rai, S. & Krishnamurthy, V. 2011 Error growth in Climate Forecast System daily retrospective forecasts of South Asian monsoon. *Journal of Geophysical Research: Atmospheres* **116** (D3), D03108.

Rosenstein, M. T., Collins, J. J. & De Luca, C. J. 1993 A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena* **65** (1), 117–134.

Shukla, J. 1981 Dynamical predictability of monthly means. *Journal of Atmospheric Science* **38**, 2547–2572.

Singh, U. P., Mittal, A. K., Dwivedi, S. & Tiwari, A. 2015 Predictability study of forced Lorenz model: an artificial neural network approach. *Discovery* **40** (181), 27–33. doi: 10.13140/RG.2.1.1985.7368.

Siqueira, L. & Kirtman, B. 2014 Nonlinear dynamics approach to the predictability of the Cane–Zebiak coupled ocean–atmosphere model. *Nonlinear Processes in Geophysics* **21** (1), 155–163.

Slingo, J. & Palmer, T. 2011 Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **369**(1956), 4751–4767. doi: 10.1098/rsta.2011.0161.

Solomon, A., Goddard, L., Kumar, A., Carton, J., Deser, C., Fukumori, I., Greene, A. M., Hegerl, G., Kirtman, B., Kushnir, Y., Newman, M., Smith, D., Vimont, D., Delworth, T., Meehl, G. A. & Stockdale, T. 2011 Distinguishing the roles of natural and anthropogenically forced decadal climate variability: implications for prediction. *Bulletin of the American Meteorological Society* **92** (2), 141.

Sprott, J. C. 2003 *Chaos and Time-Series Analysis, Vol. 69*. Oxford University Press, Oxford, UK.

Stocker, T. F. (Ed.) 2013 *Climate Change 2013: the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK & New York, USA. doi:10.1017/CBO9781107415324.

Takens, F. 1981 *Detecting Strange Attractors in Turbulence*, Springer, Berlin, Heidelberg, Germany, pp. 366–381. doi:10.1007/BFb0091924.

Tang, Y., Deng, Z., Zhou, X., Cheng, Y. & Chen, D. 2008 Interdecadal variation of ENSO predictability in multiple models. *Journal of Climate* **21** (18), 4811–4833.

Taylor, K. E., Stouffer, R. J. & Meehl, G. A. 2012 An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society* **93** (4), 485–498.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. 1992 Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena* **58** (1–4), 77–94. ISSN 0167-2789, http://dx.doi.org/10.1016/0167-2789(92)90102-S.

Toth, Z. & Kalnay, E. 1993 Ensemble forecasting at NMC: the generation of perturbations. *Bulletin of the American Meteorological Society* **74** (12), 2317–2330.

Tsonis, A. A. & Elsner, J. B. 1989 Chaos, strange attractors, and weather. *Bulletin of the American Meteorological Society* **70** (1), 14–23.

Tsonis, A. A. & Elsner, J. B. 1990 Comments on dimension analysis of climatic data. *Journal of Climate* **3** (12), 1502–1505.

Tsonis, A. A., Elsner, J. B. & Georgakakos, K. P. 1993 Estimating the dimension of weather and climate attractors: important issues about the procedure and interpretation. *Journal of the Atmospheric Sciences* **50** (15), 2549–2555.

Waliser, D. E., Stern, W., Schubert, S. & Lau, K. M. 2003 Dynamic predictability of intraseasonal variability associated with the Asian summer monsoon. *Quarterly Journal of the Royal Meteorological Society* **129** (594), 2897–2925.

Webster, P. J. & Hoyos, C. 2004 Prediction of monsoon rainfall and river discharge on 15–30-day time scales. *Bulletin of the American Meteorological Society* **85** (11), 1745.

Wu, Y., Latif, M. & Park, W. 2016 Multiyear predictability of Northern Hemisphere surface air temperature in the Kiel Climate Model. *Climate Dynamics* **47**, 793. https://doi.org/10.1007/s00382-015-2871-z.