Extreme events in globally coupled chaotic maps

S Nag Chowdhury\(^1\), Arnob Ray\(^1\), Arindam Mishra\(^2\) and Dibakar Ghosh\(^1\,*\)

\(^1\) Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata-700108, India
\(^2\) Division of Dynamics, Technical University of Lodz, Stefanowskiego 1/15, 90-924 Lodz, Poland

* Author to whom any correspondence should be addressed.
E-mail: dibakar@isical.ac.in

Keywords: extreme events, prediction, generalized extreme value distribution, deep learning

Abstract
Understanding and predicting uncertain things are the central themes of scientific evolution. Human beings revolve around these fears of uncertainties concerning various aspects like a global pandemic, health, finances, to name but a few. Dealing with this unavoidable part of life is far tougher due to the chaotic nature of these unpredictable activities. In the present article, we consider a global network of identical chaotic maps, which splits into two different clusters, despite the interaction between all nodes are uniform. The stability analysis of the spatially homogeneous chaotic solutions provides a critical coupling strength, before which we anticipate such partial synchronization. The distance between these two chaotic synchronized populations often deviates more than eight times of standard deviation from its long-term average. The probability density function of these highly deviated values fits well with the generalized extreme value distribution. Meanwhile, the distribution of recurrence time intervals between extreme events resembles the Weibull distribution. The existing literature helps us to characterize such events as extreme events using the significant height. These extremely high fluctuations are less frequent in terms of their occurrence. We determine numerically a range of coupling strength for these extremely large but recurrent events. On-off intermittency is the responsible mechanism underlying the formation of such extreme events. Besides understanding the generation of such extreme events and their statistical signature, we furnish forecasting these events using the powerful deep learning algorithms of an artificial recurrent neural network. This long short-term memory (LSTM) can offer handy one-step forecasting of these chaotic intermittent bursts. We also ensure the robustness of this forecasting model with two hundred hidden cells in each LSTM layer.

1. Introduction

Human’s spontaneous endeavor for understanding the unknown drives innovation by surpassing the limitations. The growing fascination to predict uncertainty benefits us to construct flexible tools for generating feasible explanations behind those puzzles. Despite substantial progress of such forecasting schemes, accurate prediction of uncertain future remains a grand challenge. And this challenge accelerates to few times in case of understanding and predicting extreme events. Extreme events [1, 2], in general, occur far off from the mean state (central tendency) of the skewed distribution [3]. Consequently, extreme events appear in the tail of the distribution with comparatively less frequency. Thus, the scientific community has too often dealt with a small number of data extracting information about the unpleasant long-lasting outcomes of extreme events. Recently, dynamical systems [4, 5] are being recognized as a valuable contemporary theory to overcome the curse of limited data.

The dynamical systems perspective facilitates propagating a large amount of data by evolving the equations of motions of the system forward in time. The temporal evolution of a suitable observable exhibiting extremely large or small excursions from the long-term average has attracted considerable attention among interdisciplinary studies. These recurrent excursions may resemble extreme events, as their irregular occurrence in time is seemingly random. Motivated by these analogies, many investigations take place, shifting the focus to the...
single and coupled dynamical systems. Recently, Ray et al. [6] demonstrated the emergence of extreme events under the accumulated effect of distributed damping parameter and repulsive interaction in a globally coupled network of heterogeneous Josephson junctions. Earlier, the underlying generating mechanisms of extreme events in excitable systems of FitzHugh–Nagumo units with different coupling topologies, including global and small-world network, are investigated thoroughly in references [7–9]. Occurrence of the dragon-king-like extreme events due to occasional irregular unison firing in two coupled Hindmarsh–Rose bursting neurons is inspected in reference [10]. The role of edges in the origination of extreme events is scrutinized in references [11, 12]. However, the study of extreme events has been less explored in coupled maps.

The statistical properties of extreme events in deterministic dynamical systems are analyzed in the seminal work [13]. Ray et al. [14] successfully applied a threshold-activated coupling [15] to suppress extreme events in two coupled Ikeda maps. The sudden relatively rare and significantly large explosive growth of population densities is recently reported in a network of coupled chaotic Ricker maps [16]. Except for these few pioneering works, most of the earlier studies on extreme events deal with only continuous dynamical systems [17–22] to the best of our knowledge. In this article, our results reveal the emergence of extreme events in globally coupled chaotic maps due to on-off intermittency [23]. Cavalcante et al. [24] previously reported this mechanism of causing extreme events. On-off intermittency is also identified as the principal process for the genesis of extreme events in temporal networks of mobile agents [25, 26] under the influence of attractive-repulsive co-existing interactions [27–30].

Quite astonishingly, the coupled maps have fascinating, realistic applications, including stock markets [31], fluid dynamics [32], quantum field theories [33], optical systems [34], and many more. Coupled map lattices [35–40] can offer a plethora of emerging dynamical behaviors, ranging from complete synchronization [41–43], splay state [44], chimera [45, 46] to clustering [47, 48] and spatiotemporal intermittency [49] to name a but a few. We identify extreme events in the globally coupled maps near the synchronization manifold and utilize the benefits of the machine learning approach to predict the dynamics incorporating extreme events. Recently, model-free prediction of chaotic dynamics has become one of the mushrooming avenues of research that offers considerable benefits in forecasting extreme events without knowing the explicit equations of the model. Reservoir computing, one of those effective tools, can be applied to indicate early warning signals of extreme events successfully, as recently highlighted by Pyragas et al. [50]. The performance of the neural networks for predicting chaotic dynamics of the two-dimensional Hénon map is investigated rigorously in reference [51]. Several other data-driven methods enjoy widespread appreciation in the current literatures [52–58] for forecasting extreme events. Our study involves a special kind of recurrent neural networks (RNN) [59], viz long short-term memory (LSTM) [60–67] and unveils its excellent flexibility to capture the dynamics of upcoming extreme events optimally.

Elman RNNs [68] require repeated multiplication of the gradient with the weight matrix during back-propagation through time [69]. This leads to the vanishing or exploding gradients problem depending on the recurrent weight matrix’s negative or positive spectral radius. Besides, a vanilla RNN cannot remember long-term dependencies of data over a long time. The limitations of RNNs fuel the discoveries of LSTM [60], which can yield one-step-ahead of prediction, disregarding the problems of speed and stability in RNNs. Apart from multiple practical applications ranging from short-term traffic prediction [70] to handwriting [71] and speech [72] recognition, LSTM is recently found to be fruitful in detecting extreme events at Uber data [73] and flood forecasting [74]. In what follows, we are interested in exploring the emergence of extreme events in globally coupled chaotic maps. The necessary details of the model and the mechanism behind the formation of extreme events are the contents of section 2. Section 3 is devoted to the statistical studies of extreme events. We discuss the range of coupling strength to provide a clear separation between the non-extreme and extreme events in section 4. We then proceed with the accurate one-step-ahead prediction of extreme events in section 5, whereas we summarize and conclude with a discussion of our findings in the last section 6.

### 2. Mechanism involved in the formation of extreme events

The coupling through the mean-field in an ensemble of identical coupled oscillators can provide fascinating outcomes like explosive death [75] and synchronous state [76]. We consider the following equations

\[
x_{n+1}(i) = f(x_n(i)) + \varepsilon \left( f_0 - f(x_n(i)) \right), \quad i = 1, 2, \ldots, N
\]

(1)

describing the dynamics of \(N \geq 2\) globally coupled one-dimensional maps [77]. Here, \(x_n(i)\) reflects the one-dimensional real state vector of the \(i\)th map at the \(n\)th iteration. The dynamics of each map is given by the \(C^1\) smooth one-dimensional nonlinear function \(f: \mathbb{R} \to \mathbb{R}, f_0 = \frac{1}{N} \sum_{j=1}^{N} f(x_j(j))\) depicts the arithmetic mean. The coupling parameter \(\varepsilon\) lies within the compact interval [0, 1]. However, \(\varepsilon = 0\) describes \(N\) isolated maps. On the other hand, each map evolves synchronously after the first iteration for \(\varepsilon = 1\). To determine the range
of \( \varepsilon \) for the global stability of the synchronous state, we define the following function for any pair of \( k \)th and \( l \)th maps at the \( n \)th iteration as,

\[
V_n(kl) = (x_0(k) - x_0(l))^2. \tag{2}
\]

Clearly, this function is non-negative and it will vanish only when the \( k \)th and \( l \)th maps evolve to the same motion at the \( n \)th iteration. Thus, if we can show that this function is monotonically decreasing, then it will converge to the greatest lower bound (infimum), as the function (2) is bounded below by 0. The criteria of asymptotic global stability, hence, reduces to \( V_{n+1}(kl) < V_n(kl) \). For large \( N \), one can easily derive that the system (1) asymptotically synchronizes [40] if

\[
1 - \frac{1}{A} < \varepsilon < 1 + \frac{1}{A}, \tag{3}
\]

where \( A = \sup |f| \neq 0 \) represents the least upper bound (supremum) of \( f(x) \). We further perform linear stability analysis for the complete synchronization state. The Jacobian matrix is \( J = J_0f' \), where \( f' \) is the derivative at the synchronous state and \( J_0 = \text{circ} \left(1 - \varepsilon + \frac{1}{N}, \frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N}\right) \) is an \( N \times N \) circulant matrix. The eigen values of \( J_0 \) are \( \mu_1 = 1 \) and \( \mu_j = 1 - \varepsilon \) for \( j = 2, 3, \ldots, N \). The eigenvector \( \psi_j \) corresponding to the eigen value \( \mu_j \) represents the synchronization manifold and hence, the amplification of the perturbation inside the one-dimensional invariant manifold \( \{ x(1), x(2), \ldots, x(N) \}| x(1) = x(2) = \ldots = x(N) \} \) along this eigenvector does not destroy the coherence. The other \( N - 1 \) identical eigen values help in determining the stability of the coherent state, as their corresponding eigenvectors \( \psi_j \) for \( j = 2, 3, \ldots, N \) evolve transverse to the synchronization manifold. With the help of the Lyapunov exponent \( \lambda_1 \) of the uncoupled map, the linear stability analysis provides the critical condition [47] as follows,

\[
\varepsilon_c = 1 - \exp(-\lambda_1). \tag{4}
\]

Clearly, the synchronized state is stable for \( \varepsilon > \varepsilon_c \) after a sufficient number of iterations (transient). In order to verify our findings, we choose the classical logistic map \( f(x) = rx(1 - x) \), where \( r \in [0, 4] \) is the nonlinear parameter. The globally coupled map (1) settles down to the fixed point 0 for \( r \in [0, 1) \), and converges to the fixed point 1 for \( r \in (1, 3) \) independent of the choice of the coupling strength \( \varepsilon \). For the numerical simulations, we set \( f(x) \) in the chaotic regime with \( r = 4 \) and consider \( N = 200 \). As \( \varepsilon \) lies within \([0, 1]\), the specific choice of the logistic map with \( r = 4 \) leads to the global synchronization for \( \varepsilon \in \left(\frac{1}{4}, 1\right) \) as given in equation (3). On the other hand, \( \lambda_1 \) is a function of \( r \). For \( r = 4 \), the Lyapunov exponent of the single logistic map is \( \lambda_1 = \ln 2 \). Hence, the critical coupling strength \( \varepsilon_c \) becomes \( \frac{1}{4} \) as per the relation (4).

We plot figure 1 for \( \varepsilon = 0.4995 \) such that the difference between chosen \( \varepsilon \) and \( \varepsilon_c \) remains very small. All figures of this article, including this one, are plotted for a fixed set of initial conditions randomly distributed in the uniform interval \([0, 1]\). All necessary fundamental codes along with this set of initial conditions utilized in this article are available at [78]. However, we verify that the results will remain unchanged for any other random choice of initial conditions for each map within the interval \([0, 1]\). The choice of this \( \varepsilon = 0.4995 \) will not allow these \( N = 200 \) coupled maps to settle down to a complete synchronized state as expected through our stability analysis. Figure 1(a) reveals the partial synchronization [48, 79] state of the system, where the entire system (1) splits into two groups. Particularly for the chosen set of initial conditions, one group contains exactly \( N_1 = 3 \) elements, and the other one possesses \( N_2 = N - N_1 = 197 \) members. These numbers \( N_1 \) and \( N_2 \) depend on the choice of random initial conditions. All figures are inspected after \( 3 \times 10^5 \) iterations. We plot the dynamics of each logistic map in figure 1(b), which demonstrates the temporal chaotic behavior of each cluster. Careful scrutiny of this subfigure unfolds the complex behavior of these two clusters. The distance between these two groups is not same in the entire time domain.

The temporal dynamics of two clusters [48] can be described as

\[
y_{n+1}^{(1)} = (1 - \varepsilon)f(y_n(1)) + \varepsilon p_1 f(y_n(1)) + \varepsilon p_2 f(y_n(2)), \quad \text{and} \tag{5}
y_{n+1}^{(2)} = (1 - \varepsilon)f(y_n(2)) + \varepsilon p_1 f(y_n(1)) + \varepsilon p_2 f(y_n(2)),
\]

where \( p_j = N_j / N \) for \( j = 1 \) and 2 with \( N_1 + N_2 = N \). Here, without any loss of generality, we assume that the dynamical elements \( y_n(1) \) and \( y_n(2) \) satisfy the following relations of the synchronized oscillators after a viable relabeling as,

\[
x_0(1) = x_0(2) = \ldots = x_0(N_1) = y_0(1), \quad \text{and} \tag{6}
x_0(N_1 + 1) = x_0(N_1 + 2) = \ldots = x_0(N) = y_0(2),
\]
Figure 1. (a) Advent of two synchronized clusters in the globally coupled maps: we choose the coupling strength $\epsilon = 0.4995$ slightly lesser than the critical coupling strength $\epsilon_c = 0.5$. This $\epsilon_c$ is necessary for achieving complete synchronization of the system (1) with $N = 200$ as per equation (4). The system (1) settles down to a partial synchronized state. The number of elements in each cluster varies depending on the initial conditions. (b) Time evolution of each logistic map: we use the identical logistic maps $f(x) = rx(1-x)$ as the local dynamics with $r = 4$. Each map reflects a chaotic behavior. Although those oscillators occasionally collapse to a single trajectory, mostly they possess a two-cluster state, as observed in the snapshot at a specific iteration 447 855 in subfigure (a). (c) On-off intermittency of the observable $E$: we select two different maps, viz 1st and 50th oscillators from two distinct clusters using subfigure (a) and calculate the Euclidean distance $E$ between them. The temporal evolution of the observable $E$ contemplates chaotic intermittent bursts which cross the significant height $H_S = \langle E_n \rangle + 8\sigma$ from time to time. $\langle E_n \rangle$ is the sample mean of the local maxima $E_n$ of $E$ gathered over sufficiently long $10^8$ iterations after the initial transient. $\sigma$ is the standard deviation of the whole simulated data. We will assess an event $E_n$ as an extreme event if it crosses the red horizontal dashed line $H_S$. This figure is drawn for a particular set of randomly distributed initial conditions within the interval $[0, 1]$. We consider $3 \times 10^9$ iterations as the transient. For further details, please see the text.

When any one of $p_1$ or $p_2$ proceeds to the limit $0$, then the other one tends to $1$. For instance, if $p_1 \to 0^+$, then $p_2 \to 1^-$. Thus for $p_1 = 0$, equation (5) govern the following equations,

$$y_{n+1}(1) = (1 - \epsilon) f(y_n(1)) + \epsilon f(y_n(2)), \quad \text{and} \quad y_{n+1}(2) = f(y_n(2)).$$

(7)

Noticeably, $y(2)$ does not depend on the variable $y(1)$ for $p_1 = 0$, although $y(1)$ depends on both the variables $y(1)$ and $y(2)$. Thus, these equations in the asymmetric cluster limit possess the skew product structure [80].

We define a new variable $E = |y(1) - y(2)|$ describing the Euclidean distance between two clusters. Figure 1(c) displays the temporal dynamics of $E$ after the initial transient. $E$ remains near zero (under the numerical accuracy) most of the iterations revealing the synchronization of the two clusters. However, this off-state ($E \approx 0$) is interrupted occasionally, leaving the trait of large amplitude random bursts. This non-zero value of $E$ is so high that it crosses the pre-defined threshold $H_S$ (shown in dashed red line in sub figure 1(c)) infrequently. This $H_S = \langle E_n \rangle + 8\sigma$ is treated here as the extreme event qualifying threshold [26], where $E_n$ is the local maxima of $E$ after the initial transient of $3 \times 10^3$ iterations. We accumulate the values of $E$ over $10^8$ iterations. Then, we collect the values of $E_n$ excluding the transient data. $\langle E_n \rangle$ represents the sample mean of this gathered data of $E_n$, and $\sigma$ is the corresponding standard deviation. The choice behind this threshold $H_S$ is motivated by the fact that the extreme events qualifying threshold is generally taken as $\langle E_n \rangle + d\sigma$, where $d \in \mathbb{R}$ varies within the interval $[4, 8]$ [6, 10, 17, 19–21, 25, 26]. Evidently, we choose the upper bound 8 of $d$, as it definitely portrays more relatively scattered events with strikingly higher amplitude and lower frequency.
Moreover, reference [26] presented an analytical formulation for this particular choice of threshold, and designated this threshold as the significant height in analogy with the oceanic rogue waves [81]. We employ this significant height \( H_S \) borrowing the idea from reference [26] and whenever \( E_n \) crosses \( H_S \), we recognize it as an extreme event.

We rigorously investigate that these sporadic bursts of \( E \) are not due to the appearance of floating nodes [37]. Once a node joins a synchronized cluster, it will stay there forever and will never move to the other cluster. This intermittent bursts are completely due to the varying distance of the two clusters. We choose two distinct maps from two different clusters and study the temporal evolution of their Euclidean distance \( E \) in figure 1(c). For further evaluation of extreme events, we investigate their statistical properties in the following section.

3. Statistical information of extreme events

Generally, Gaussian (normal) distribution, being symmetrically distributed about the mean, indicates undoubtedly that the data near the central tendency are comparatively more frequent in occurrence than the data far from the expected value. We assemble the data of local maxima \( E_n \) of \( E \) over a considerable interval of \( 10^6 \) iterations after leaving the initial transient of \( 3 \times 10^3 \) iterations. We plot the PDF of \( E_n \) in figure 2(a). We choose the same set of initial conditions and coupling strength \( \varepsilon = 0.4995 \) as in figure 1. The red dashed vertical line \( H_S \) denotes the extreme events indicating threshold. The value of the threshold for this particular choice of initial conditions is \( H_S \approx 0.0693 \). The left-hand side of \( H_S \) contains all non-extreme events, while the right-hand side encompasses all extreme events. Clearly, the PDF of \( E_n \) follows the non-normal distribution as a fair portion of data deviates more than \( 8\sigma \) from the mean \( \langle E_n \rangle \). In addition, we use the statistical Kolmogorov–Smirnov test (K–S test) by assuming our simulated data are identical with the Gaussian distribution. As anticipated, the K–S test rejects the null hypothesis at the default 5\% significant level. These findings validate the signature of extreme events employed in references [25, 26, 82, 83]. More or less, the common notion of definition of extreme events in these references [25, 26, 82, 83] divulges the following characteristics, viz

- (a) These events are aperiodic, possessing amplitude higher than \( H_S = \langle E_n \rangle + 8\sigma \), and
- (b) These recurrent events occur less frequently, maintaining the non-Gaussian distribution.

In this article, we also maintain these two important properties in case of defining extreme events. We also represent the PDF of the amplitude of the extreme events in figure 2(b). We draw the subfigures (a)–(c) of figure 2 using the MATLAB distribution Fitter App. Our numerically simulated data (shown in cyan color) in figure 2(b) is the same data, lying in the right of the significant height \( H_S \) in figure 2(a). We utilize the Freedman–Diaconis rule [84] for choosing the width of the bins in this histogram. We find this data are well fitted with the PDF of the GEV distribution [85] given by

\[
G(x) = \frac{1}{\beta} \exp \left( - \left( 1 + \frac{x - \alpha}{\beta} \right)^{-\frac{1}{\gamma}} \right) \times \left( 1 + \frac{x - \alpha}{\beta} \right)^{-1} \tag{8}
\]

for \( \beta \neq 0 \) and \( 1 + \frac{\alpha}{\beta} > 0 \). Here, \( \alpha, \beta, \gamma > 0 \) stand for the location, scale, and shape parameters, respectively. As GEV distribution is a combination of three families (Gumbel et al), we get type-II extreme value distribution for \( \gamma > 0 \). In contrast, we obtain the type-III case for \( \gamma < 0 \). For \( \gamma = 0 \), we have

\[
G(x) = \frac{1}{\beta} \exp \left( - \exp \left( - \frac{x - \alpha}{\beta} \right) - \frac{x - \alpha}{\beta} \right) \tag{9}
\]

representing type-I case. Table 1 contains all the necessary information about the estimated parametric values of \( \alpha, \beta, \) and \( \gamma \) for figure 2(b). GEV distribution is applicable in various fields [85], including it is the only possible limiting distribution (if it exists) of an adequately normalized maximum of a sample of independent and identically distributed random variables.

We also study the statistics of the return intervals between extreme events in figure 2(c). Recurrence interval has a broad range of applications in hazard estimations like the interarrival packet times on internet traffic [86], earthquakes [87], floods [88], and many more. We select the extreme events from each burst and collect the return interval between two consecutive such extreme events. For numerical implementation, once we detect an extreme event, then we ignore the extreme events (if it exists) for the successive 100 iterations. The sample standard deviation of the gathered data is \( 1.0490 \times 10^4 \) and the sample mean is \( 8.3016 \times 10^3 \). Thus, the coefficient of variation (CV) is 1.2636. This dimensionless number being greater than unity reveals a high variance of the data. Moreover, the CV of an exponential distribution is precisely 1, which is less than the CV of our return intervals. Hence, we conclude that the return interval distribution of extreme events
Figure 2. (a) Non-normal distribution of the local maxima: we accumulate local maxima $E_n > 0$ of $E$ over sufficiently long $10^8$ iterations after repudiating the first $3 \times 10^6$ transients. We calculate $H_S$ over these data and draw this to identify the separation between non-extreme and extreme events. (b) Probability density function (PDF) of extreme events: generalized extreme value (GEV) distribution fits well with the data from the right-hand side of the significant height $H_S$ in subfigure (a). We use log likelihood to estimate the three parameters of the family given in table 1. (c) The distribution of return intervals for the extreme events within each burst: we numerically investigate the successive manifestation of extreme events within each burst. For numerical simulation, once we detect an extreme event, we throw away subsequent extreme events (if it exists) from next 100 iterations. This data is well fitted with Weibull distribution. The estimated scale parameter is $a = 6782.39$ and the estimated shape parameter is $b = 0.721932$. We also use the popular P–P (probability–probability) plot (not shown here) to confirm our claim. The subfigures (a)–(c) are drawn with $\epsilon = 0.4995$ for the $N = 200$ globally coupled identical logistic maps $f(x) = 4x(1 - x)$. (d) The probability of appearance of extreme events: we compute the probability of occurrence of extreme events for 7 different values of $\varepsilon$ using the relation (12) and successfully validate it by plotting the upper bound using the inequality (11). Here, $\varepsilon_c = 0.5$. We use the same set of random initial conditions drawn from $[0, 1]$ for all results.

Table 1. Estimated parameter values for GEV distribution.

| Parameter | Estimate     | Standard error |
|-----------|--------------|----------------|
| $\alpha$  | 0.0891089    | $8.7959 \times 10^{-5}$ |
| $\beta$   | 0.0212706    | $8.88201 \times 10^{-5}$ |
| $\gamma$  | 0.60792      | 0.00466071     |

The probability of appearance of extreme events does not follow the exponential distribution. Earlier studies [89–91] suggest these return intervals are generally approximated to any one of the following distributions, viz (i) exponentially distributed, (ii) stretched exponential, and (iii) Weibull distribution. Memoryless is a vital property of the exponential distribution of non-negative real numbers. In contrast, long-term memory usually gives rise to a stretched exponential distribution of the return intervals [90]. Recently, Weibull distribution is also found to be a good representative for the distribution of recurrence time intervals of long-range correlated data [89]. The PDF plotted in the semi-log scale resembles of a product of a stretched exponential and a power law. This PDF in figure 2(c) is well fitted by the Weibull distribution. Estimation of the scale and shape parameter is given in table 2.

Here, $a$ is the scale parameter and $b$ is the shape parameter. The flexibility of the Weibull distribution allows it to model in an astonishingly diverse array of data sets, ranging from weather forecasting to reliability
engineering and failure analysis. The PDF of the two parameter Weibull distribution is given by

$$W(x) = \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b}.$$  \hspace{1cm} (10)

Although $E$ remains zero for a large number of iterations as illustrated through figure 1(c), $E_n$, being the local maxima of $E$, is positive-valued observable. Evidently, $H_S = \langle E_n \rangle + 8\sigma$ is also positive quantity and varies with each different choice of initial conditions, even for fixed $\varepsilon$ and $N$. We come up with an upper bound for the probability of occurrence of extreme events using Markov’s inequality \cite{92} as follows

$$P(E_n > H_S) \leq \frac{\langle E_n \rangle}{H_S} \leq 1.$$  \hspace{1cm} (11)

We compute the probability of occurrence of extreme events over the interval $[0.4985, 0.4999]$ of the coupling strength $\varepsilon$ with the same set of initial conditions used in figure 1. To calculate its probability, we iterate the system 1 with $N = 200$ identical logistic maps for $10^8$ iterations and then discard the initial $3 \times 10^5$ data treating them as transient. Then, we calculate the number of local maxima $E_n$ of $E$ over that sufficiently long data. After that, we simulate the significant height $H_S$ in each case and count the number of extreme events, i.e. $E_n$ which cross $H_S$. Hence, we apply the classical definition of probability as follows

$$\text{Probability of occurrence of extreme events} = \frac{\text{Number of extreme events}}{\text{Number of } E_n}.$$  \hspace{1cm} (12)

We plot these values for $\varepsilon_c - \varepsilon$ in figure 2(d) along with the upper bound using the relation 11. Here, $\varepsilon_c = 0.5$ is the required critical coupling strength for the emergence of complete synchronization as calculated through the equation (4). One should notice that the upper bound proposed in the inequality 11 is not, by any means, the least upper bound. Even though we cannot observe any clear trend between the numerically calculated probabilities, one can anticipate that as the distance between $\varepsilon_c - \varepsilon$ increases, the chaotic bursts of $E$ become more intermittent due to the instability of synchronization invariant manifold $\{(x(1), x(2), \ldots, x(N)) | x(1) = x(2) = \ldots = x(N)\}$. Consequently, these sporadic bursts become more frequent and become less likely to cross the threshold $H_S$. We will discuss this result in detail in the next section.

4. Range of coupling strength for the emergence of extreme events

We contemplate the dynamical behavior of the system (1) of $N = 200$ identical logistic maps in figure 3 with the same initial conditions as used in all previous figures. This inspection will help us to determine the range of $\varepsilon$ corresponding to extreme events. Figure 3(a) reflects a one-dimensional parameter range of the coupling strength $\varepsilon$, where we identify numerically four distinct behaviors. For $\varepsilon < \varepsilon_{ch} \approx 0.433\,030\,12$ (dark blue interval of figure 3(a)), the system settles down to a two cluster state. The behavior of each map is periodic in time, and accordingly, the distance between two clusters $E$ is also periodic. Hence, we never expect the manifestation of extreme events in this range. The system falls in a chaotic two cluster state within the interval $(\varepsilon_{ch}, \varepsilon_c)$. The purple regime of $\varepsilon > \varepsilon_c$ in figure 3(a) gives rise to a coherent chaotic state, as anticipated from our stability analysis. However, we fail to provide an analytical estimation to decide the range of $\varepsilon$ exhibiting extreme events. That’s why we examine its range through numerics. Our numerical study suggests extreme events arise only within the interval $(\varepsilon_{EE}, \varepsilon_c)$, where $\varepsilon_{EE} \approx 0.496\,1745$. The interval $(\varepsilon_{ch}, \varepsilon_{EE})$ yields the two cluster chaotic attractor, but their Euclidean distance $E$ becomes too frequent to cross $H_S$. To reveal this nature, we choose $\varepsilon = 0.496$ from the non-extreme (orange) range of figure 3(a) and select the 1st and 2nd maps from two different clusters. The difference between these two maps is plotted in figure 3(b). Clearly, the intermittent random bursts of $E$ do not cross the extreme event qualifying height $H_S$ (red dashed horizontal line) in this figure. Similarly, we set $\varepsilon = 0.4999$ from the interval (black) of occurrence of extreme events in figure 3(a). The system again falls in a partially synchronized state, and we choose two different maps, viz 98th and 99th maps corresponding to two distinct clusters. The distance between these two maps occasionally crosses the significant height $H_S$ in figure 3(c). The aperiodic random bursts attest to the emergence of extreme events.

| Parameter | Estimate | Standard error |
|-----------|----------|----------------|
| $a$       | 6782.39  | 90.4573        |
| $b$       | 0.721932 | 0.00522843     |

Table 2. Estimated parameter values for Weibull distribution.
Figure 3. The range of coupling strength separating the non-extreme and extreme events: we split an interval of $\varepsilon$ over four distinct groups. Dark blue interval of $\varepsilon < \varepsilon_{ch}$ exhibits periodic two clusters, whereas the remaining interval contains a two-cluster chaotic attractor for $(\varepsilon_{ch}, \varepsilon_c)$. The orange interval $(\varepsilon_{ch}, \varepsilon_{EE})$ displays the interval encompassing non-extreme events. We come up with the subfigure (b) for $\varepsilon = 0.496$, where the red dashed horizontal line is the significant height $H_S = \langle E_n \rangle + 8\sigma$. Evidently, $E$ flunks to cross this threshold $H_S$. Nevertheless, $E$ crosses $H_S$ intermittently for $\varepsilon = 0.4999$ in subfigure (c). The black interval $(\varepsilon_{EE}, \varepsilon_c)$ of subfigure (a) incorporates the interval of extreme events. The purple interval of $\varepsilon > \varepsilon_c = 0.5$ presents the coherent chaotic attractor. For simulation, we consider $N = 200$ globally coupled maps with identical logistic maps $f(x) = 4x(1 - x)$ as the local dynamics. For further details, please see the text.

One should emphasize that although we plot the open interval $(\varepsilon_{ch}, \varepsilon_c)$ continuously by the black solid line in figure 3(a). However, one may find a set of Lebesgue measure zero within this interval, which contains few points of coupling strength that unable to produce extreme events. One such point is $\varepsilon = 0.49775$. For this choice of $\varepsilon$, we accumulate sufficiently long data of $E_n$ over $10^6$ iterations after discarding the initial transient $3 \times 10^5$. We observe that these long iterates of $E_n$ fail to cross the derived $H_S$. We draw this figure 3 with the same set of initial conditions as used in figure 2. However, we verify the results remain unchanged if we choose other random initial conditions of each element from the uniform interval $[0, 1]$. Although, other initial conditions may slightly change the value of $\varepsilon_{ch}$ and $\varepsilon_{EE}$.

Lastly, we want to conclude this section with the fact that this appearance of extreme events is a generic signature in globally coupled chaotic maps, and our choice of the identical one-dimensional logistic map as the chaotic local map is only a paradigmatic example for illustrating the phenomena. The intermittent transition [39, 47, 49] from the synchronized state to a two-cluster attractor always allows a range of coupling strength for the origination of extreme events near the synchronization manifold in the system (1) consisting of one-dimensional chaotic maps.

5. Short-term prediction of extreme events

After the statistical characterization of extreme events, we focus on their prediction utilizing LSTM, a deep learning architecture. The recurrent neural nets, viz LSTM networks used in this article is found to be capable of one-step-ahead forecasting of chaotic time series. The LSTM cell can process data sequentially and keeps its hidden state through time. Using the univariate time series, we create the regression data set $\{X, Y\}$ by taking a lag of 1 (for example, if $X_n$ be the actual time iteration, then $Y_n = X_{n+1}, n = 1, 2, \ldots$, sample size). And using the training data set $\{X, Y\}$, we build LSTM net so that we can deal with the vanishing gradient problem that can be encountered when training traditional RNNs. Thus, it is important to note that the dimension of the feature vector is one. So, sequence input with one dimension is passed in LSTM with 200 hidden units in a fully connected layer. Here, the loss function used to compute the weight matrix connecting different layers is a mean-squared error. We iterate the system (1) with $\varepsilon = 0.4995$ for $10^8$ iterations and collect the distance $E$ between the two clusters after deleting first $3 \times 10^5$ transient. Then we calculate the local maxima $E_n$ of this whole gathered data of $E$. We use the first 165,000 data points of $E_n$ consisting of few extreme events, out of which 90% data are used for the training. The remaining 10% data points are treated as a testing set. The
Figure 4. Architecture of long short term memory (LSTM): in subfigure (a), we present a detailed schematic design of an artificial recurrent neural network (RNN), which can avoid long-term dependency problems. Each layer contains 200 LSTM neurons globally connected to all the LSTM neurons of the next layer. For clear visualization, we show all LSTM neurons explicitly only in three layers. We train the RNNs with previous (n−1) data (90% of the accumulated data). Each LSTM cell consists of three different gates, as shown in subfigure (b). The forget gate decides whether the information from the cell state is required for the LSTM or not. The input gate helps in learning new relevant information from the input to the cell state. This gate also includes a sigmoid function like the forget gate. The output gate passes the updated filtered version of information from the current nth iteration to the following (n+1)th iteration. Each LSTM has to deal with three things, viz (i) the hidden state $h_{n−1}$, i.e. the output of the last cell, (ii) the input $E_n$, i.e. the new information available at the current time step, and (iii) the accessible information $c_{n−1}$ of the previous cell after the previous time step. Subfigure (c) displays an LSTM cell, where ellipses and circles stand for point-wise operations, while rectangles indicate neural network layer operations. We unveil a concise description of this diagram accompanying all the related details in the main text.

A schematic sketch of neural architectures is shown in figure 4(a), where each layer comprises 200 nodes for the LSTM. This number of hidden cells in each LSTM layer is determined using the cross-validation technique [93]. $\hat{E}_{n+1} = F_0^h(h_n)$ is the prediction for the observable $E_{n+1}$ depending on the input observable $E_n \in \mathbb{R}^{d_o}$. $F_0^h$ is the hidden to output mapping. $d_o$ and $d_h$ are the dimension of the input and hidden states respectively. This variant RNN consists of two internal states (i) cell state $c_n \in \mathbb{R}^{d_h}$, and (ii) hidden state $h_n = F_0^h(E_n, h_{n−1}) \in \mathbb{R}^{d_h}$, respectively. $F_0^h$ is the hidden to hidden mapping, which is given by the following equations

$$
\begin{align*}
\hat{c}_n &= \tanh (W_{\text{cell}} [h_{n−1}, E_n] + \xi_{\text{cell}}), \\
c_n &= g_{\text{forget}} \odot c_{n−1} + g_{\text{input}} \odot \hat{c}_n, \\
g_{\text{output}} &= \sum_{\text{hidden}} (W_{\text{hidden}} [h_{n−1}, E_n] + \xi_{\text{hidden}}), \\
h_n &= g_{\text{output}} \odot \tanh (c_n), \\
\hat{h}_{n+1} &= F_0^o(h_n) \\
\end{align*}
$$

Here, $\odot$ is the element-wise product (Hadamard product). Each LSTM is composed of three gates to protect and control the cell state, which carries the sequence input initially. $g_{\text{forget}}$, $g_{\text{input}}$, $g_{\text{output}} \in \mathbb{R}^{d_h}$ represent the activation vectors of respective gates. $\xi_{\text{forget}}$, $\xi_{\text{input}}$, $\xi_{\text{cell}}$ and $\xi_{\text{hidden}}$ are the bias vectors. $\sum_{\text{forget}}, \sum_{\text{input}}, \sum_{\text{hidden}}$...
are the sigmoid functions. The weight matrices \( W_{\text{forget}}, W_{\text{input}}, W_{\text{cell}}, W_{\text{hidden}} \in \mathbb{R}^{d_h \times (d_h + d_o)} \) along with the bias vectors need to be learned during training. The forget gate decides how much information should be appended into the final product of the cell state at each iteration \( n \). This final product, i.e. the hidden state, is treated as the input of the next layer. The average of these outcomes across all hidden states is the final predicted output. The second last relation of equation (13) suggest that depending on \( \Sigma_{\text{hidden}} \), we get an output gate’s activation vector \( g_{\text{output}} \). Hyperbolic tangent (tanh) function acts on the cell state generating a value between \(-1\) and \(1\). The last relation of equation (13) will produce the hidden state, which will give the prediction \( \hat{E}_{n+1} \). This forecast \( \hat{E}_{n+1} \) is a linear function \( \hat{F}_h \) given by the following relation

\[
\hat{E}_{n+1} = W_o h_n, \tag{14}
\]

where \( W_o \in \mathbb{R}^{d_o \times d_h} \). One such LSTM neuron is shown in figures 4(b) and (c). For a detailed description of LSTM networks, please see the pioneer work of Schmidhuber and Hochreiter [60].

This LSTM network can give rise to the one-step-ahead future prediction depending on each time step of the input sequence as illustrated in figure 4. We plot the predicted output and the test data simultaneously for a short horizon in figure 5(a). To determine the accuracy of the convergence between these two data sets, we calculate the arithmetic mean of the squares of the Euclidean distance between \( E \) and \( \hat{E} \) as follows

\[
\text{mse} \left( E, \hat{E} \right) = \frac{1}{l} e^T e = \frac{1}{l} \sum_{k=1}^{l} (e_k)^2, \tag{15}
\]

where \( e_k \) is the difference between the forecasted outcomes and the actual value. \( e \) is the \( l \times 1 \) matrix and \( l \) denotes the number of data points in the testing data. This loss function is a positive valued metric tending to zero for an accurate forecast. The mse for figure 5(a) with \( l = 16,499 \) is \( 0.000\,011\,589 \). However, reference [65] criticizes this mean squared error due to non-normalization for the variability of the data. Hence, we use a different measure as follows

\[
\text{nse} \left( E, \hat{E} \right) = 1 - \frac{\text{mse} \left( E, \hat{E} \right)}{\text{mse} \left( E, E \right)}, \tag{16}
\]
Figure 6. Goodness of prediction and robustness of the forecasting model: we perform 30 independent trials of the prediction method and calculate the mean squared error (mse), a statistical measure to quantify the prediction accuracy of event forecasting. We plot mean (⟨mse⟩) and fluctuation (σ⟨mse⟩) of mse over 30 realizations in this figure. We evaluate ⟨mse⟩ = 0.000010047, and σ⟨mse⟩ = 0.0000012476. The small value of ⟨mse⟩ signifies the good prediction as well as the small value of σ⟨mse⟩ indicates that our prediction scheme is robust over the realizations.

where ⟨E⟩ is the mean of the observed data. This measure is widely known as the Nash–Sutcliffe model efficiency coefficient being a normalized version of the mse, where mse (⟨E, ⟨E⟩⟩) is the data variance. This metric nse assumes the upper bound 1 for the ideal scenario of exact forecasting. The nse is 0.7907 for figure 5(a) determining the predictive skill of our setup. The existing literature suggests an architecture promises more predictive performance for nse → 1−. An enlarged portion of figure 5(a) is highlighted in figure 5(b).

In fact, we contemplate the robustness of the used deep learning technique in figure 6. We define the following equations

\[
\langle \text{mse} \rangle = \frac{1}{h} \sum_{i=1}^{h} \text{mse}^i, \quad \text{and}
\]

\[
\langle \text{nse} \rangle = \frac{1}{h} \sum_{i=1}^{h} \text{nse}^i.
\]

Here, h is the number of trials. mse^i and nse^i are the respective mse and nse at the i-th realization. We take independent h = 30 realizations and delineate the average ⟨mse⟩ with their corresponding standard deviation σ⟨mse⟩ in figure 6. ⟨nse⟩ for these 30 independent trials is 0.81856. It is evident that this ⟨nse⟩ is higher than the nse of figure 5. Figures 5 and 6 reflect the successful one-step forecasting of the local maxima En of E with sufficiently moderate accuracy. These data sets possess few extreme events crossing the dashed line HS as shown in figure 5. We find our trained model agrees well with the observed chaotic data consisting of extreme events.

6. Discussion and conclusion

Recent decades have seen many efforts to understand the mechanisms behind the recurrent emergence of extreme events in dynamical systems and complex systems. Attainment of giant extraordinary values in a short time of a relevant observable is generally known as an extreme event. These recurrent short-lasting events deviate significantly from the long-term average of their statistical distribution. Due to its practical importance, this topic gains widespread recognition among interdisciplinary researchers. In spite of such leaps of progress, these investigations provide only a partial understanding of the mechanisms of extreme events and their predictions. The scientific community of nonlinear dynamics brings few successful signs of progress by considering investigations through models following differential equations. However, most of these studies are performed by contemplating continuous dynamical systems. Theoretical investigations on maps for an understanding of the dynamics of extreme events are not well explored yet. Only a few studies are reported on discrete dynamical systems [13, 14, 16] along this line to the best of our knowledge.

Moreover, most of the existing studies on globally coupled networks require repulsive interaction [6, 25] for the appearance of extreme events. The co-existence of attractive-repulsive coupling of suitable strength is necessary for the emergence of extreme events in globally coupled identical Stuart–Landau oscillators as per the numerical investigations of reference [25]. Ray et al [6] also observed the signature of extreme events under repulsive mean-field interaction in a network of heterogeneous oscillators. Although, Ansmann et al
We have investigated the emergence of extreme events in a globally coupled network of one-dimensional chaotic maps. We have calculated the critical coupling strength $c_\epsilon$, independent of the system size $N$, for achieving a synchronized chaotic state. Before this coupling strength, the system settles down to an asymmetric two-cluster attractor. The distance $E$ between these two clusters occasionally travels far away and possesses some large values. These intermittent bursts of $E$ are so large that they infrequently deviate more than eight times the standard deviation from the mean of the local maxima $E_n$ of $E$ for a suitable range of coupling strength. These highly deviated values of $E$ are classified as extreme events if they cross the significant height $H_s$. The reason behind this specific choice of threshold is supplemented using the existing literature [26]. We have numerically determined the range of $\epsilon$ for the occurrence of extreme events in section 4. GEV distribution further corroborates the statistical signature of these events. We have evaluated the probability of the appearance of extreme events for some specific values of $\epsilon$ using equation (12) and the procedure given in section 3. For instance, this probability is 0.0027 for the system (1) with $N = 200$ coupled identical logistic maps $f(x) = 4x(1-x)$ and $\epsilon = 0.4995$. These results confirm that the appearance of extreme events is very fewer in frequency. The return time distribution of these extreme events resembles of Weibull distribution.

Understanding of extreme events brings to the limelight lots of open queries. The interdisciplinary research of complex systems plays a decisive role in providing valuable insights into these long-standing issues. We have contemplated the class of globally coupled one-dimensional chaotic maps hoping that our analysis may enhance the understanding of these non-equilibrium phenomena. This article has several folds with insightful results, including (i) the mechanisms triggering extreme events (see section 2), (ii) the statistical aspect of extreme events (see section 3), and (iii) the prediction of extreme events (see section 5). We analyze these three essential themes in the studies of extreme events by placing the paradigmatic one-dimensional chaotic identical logistic maps in each node of the globally coupled network. The mechanisms of formation of extreme events may unravel in gaining insightful information regarding the onset of intermittency transition for globally coupled chaotic maps. Additionally, this route of two-state intermittent nature may unfold the possibility of mitigation of extreme events within chaotic dynamical systems. By tuning the coupling parameter, one can easily avoid such occasional giant events (see section 4). And this mitigation strategy is only possible due to our precise knowledge of the mechanism viz self-organized on-off intermittency triggering extreme events. In addition, statistical studies are promising tools that help in drawing inferences regarding the frequency and probability of occurrence of extreme events from the samples [94]. Nevertheless, knowledge of the mechanism behind the generation of extreme events may not always help in predicting extreme events. That’s why we take advantage of the machine learning approach to predict the dynamics incorporating extreme events. There are several examples of model-free predictions such as one-day-ahead forecasting electricity consumption [95, 96], forecasting hourly solar irradiance [97], predicting El Niño events with a lead time of one-year [66] to name a few. The deep learning strategies are capable of learning the essential chaotic dynamics directly from data and can deal with this grand undesirable challenge. We are successfully able to predict the chaotic dynamics of $E_n$ consisting of extreme events with moderate accuracy. We hope that the perceived study may reveal a thorough understanding of this crucial phenomenon. The employed trained LSTM predictors have predictive power and practical applicability within the scope of viable forthcoming research and might help to comprehend the forecast of extreme events in realistic scenarios.
References

[1] McPhillips L E et al 2018 Earth's Future 6 441–55
[2] Ghil M et al 2011 Nonlinear Processes Geophys. 18 295–350
[3] Frolow N S et al 2019 Sci. Rep. 9 7243
[4] Lucarini V 2016 Extremes and Recurrence in Dynamical Systems (New York: Wiley)
[5] Nag Chowdhury S, Ray A, Dana S K and Ghosh D 2021 Phys. Rep. (submitted)
[6] Ray A, Mishra A, Ghosh D, Kapitaniak T, Dana S K and Hens C 2020 Phys. Rev. E 101 032209
[7] Ansmann G, Karnatak R, Lehnhertz K and Feudel U 2013 Phys. Rev. E 88 052911
[8] Karnatak R, Ansmann G, Feudel U and Lehnhertz K 2014 Phys. Rev. E 90 022917
[9] Ansmann G, Lehnhertz K and Feudel U 2016 Phys. Rev. X 6 011030
[10] Mishra A, Saha S, Vigneshwaran M, Pal P, Kapitaniak T and Dana S K 2018 Phys. Rev. E 97 062311
[11] Bröhl T and Lehnhertz K 2020 Chaos 30 073113
[12] Kumar A, Kulkarni S and Santhanam M S 2020 Chaos 30 043111
[13] Nicolis C, Balakrishnan V and Nicolis G 2006 Phys. Rev. Lett. 97 210602
[14] Ray A, Rakshit S, Ghosh D and Dana S K 2019 Chaos 29 043131
[15] Sinha S and Biswas D 1993 Phys. Rev. Lett. 71 2010
[16] Moitra P and Sinha S 2019
[17] Mishra A, Leo Kingston S, Hens C, Kapitaniak T, Feudel U and Dana S K 2020 Chaos 30 063114
[18] Ray A, Rakshit S, Basak G K, Dana S K and Ghosh D 2020 Phys. Rev. E 101 062210
[19] Suresh R and Chandrasekar V K 2020 Phys. Lett. A 383 125997
[20] Tayyebi A, Ray A, Saha S, Saha S, Saha S, Saha S, Saha S and Saha S 2019 Phys. Lett. A 383 125997
[21] Moitra P and Sinha S 2019
[22] Kaviya B, Suresh R, Chandrasekar V K and Balachandran B 2020 Int. J. Non-Linear Mech. 127 103596
[23] Bonatto C and Endler A 2017 Phys. Rev. E 96 012216
[24] Platt N, Spiegel E A and Tresser C 1993 Phys. Rev. Lett. 70 279
[25] de S Cavalcante H L D, Orí M, Sornette D, Ott E and Gauthier D J 2013 Phys. Rev. Lett. 111 198701
[26] Chowdhury S N, Majhi S and Ghosh D 2020 IEEE Trans. Netw. Sci. Eng. 7 3159–3169
[27] Chowdhury S N, Majhi S, Ozer M, Ghosh D and Perc M 2019 New J. Phys. 21 073048
[28] Majhi S, Chowdhury S N and Ghosh D 2020 Europhys. Lett. 123 20001
[29] Chowdhury S N, Ghosh D and Hens C 2020 Phys. Rev. E 101 022230
[30] Mishra A, Hens C, Bose M, Roy P K and Dana S K 2015 Phys. Rev. E 92 062920
[31] Chowdhury S N, Rakshit S, Buldh J M, Ghosh D and Hens C 2021 Phys. Rev. E 103 032210
[32] Ma W J, Hu C K and Amritkar R E 2004 Phys. Rev. E 70 026210
[33] Arsl H 1983 Ann. Rev. Fluid Mech. 15 345–359
[34] Beck C 2002 Spatio-Temporal Chaos and Vacuum Fluctuations of Quantized Fields vol 21 (Singapore: World Scientific)
[35] Perez G, Pando-Lambрусchi C, Sinha S and Cereide H A 1992 Phys. Rev. A 45 5469
[36] Kaneko K 2015 Chaos 25 097608
[37] Xie F and Cereide H A 1996 Phys. Rev. E 54 3235
[38] Jalan S, Amritkar R and Hu C K 2005 Phys. Rev. E 72 016211
[39] Amritkar R, Jalan S and Hu C K 2005 Phys. Rev. E 72 016212
[40] Kaneko K 1989 Phys. Rev. Lett. 63 219
[41] Jost J and Joy M P 2001 Phys. Rev. E 65 016201
[42] Chowdhury S N, Majhi S, Ghosh D and Prasad A 2019 Phys. Lett. A 383 125997
[43] Rakshit S, Ray A, Bera B K and Ghosh D 2018 Nonlinear Dyn. 94 785–805
[44] Chowdhury S N and Ghosh D 2019 Europhys. Lett. 125 10011
[45] Singh J and Gupta N 2016 Phys. Rev. E 94 052204
[46] Khaleghi L, Panahi S, Chowdhury S N, Bogomolov S, Ghosh D and Jafari S 2019 Physica A 536 122596
[47] Parastesh F, Jafari S, Azarnoush H, Shahriari Z, Wang Z, Boccaletti S and Perc M 2021 Phys. Rev. E 98 1–114
[48] Balmforth N J, Jacobson A and Provenzale A 1999 Chaos 9 738–754
[49] Popoyych O, Maistrenko Y and Mosekilde E 2002 Phys. Lett. A 302 171–181
[50] Kaneko K 1992 Chaos 2 279–82
[51] Pyragas V and Pyragas K 2020 Phys. Lett. A 384 126591
[52] Lellep M, Pirel J, Linkmann M and Eckhardt B 2020 Chaos 30 013113
[53] Ray A, Chakraborty T and Ghosh D 2021 arXiv:2106.08968
[54] Qi D and Majda A J 2020 Proc. Natl Acad. Sci. USA 117 52–9
[55] Narihi M, Salmela L, Toivonen J, Billet C, Dudley J M and Gentry G 2018 Nat. Commun. 9 4923
[56] Sapsis T P 2017 Phil. Trans. R. Soc. A 375 20170133
[57] Wan Z Y, Vlachos P, Koumoutsakos P and Sapsis T P 2018 PLoS One 13 e0197704
[58] Amil P, Soriano M C and Massoller C 2019 Chaos 29 113111
[59] Meiyazhagan J, Sudharsan S and Senthivelan M 2021 Eur. Phys. J. B 94 156
[60] Abiodun O I, Jantan A, Omolara A E, Dada K V, Mohamed N A and Arshad H 2018 Helion Phys. 4 e00938
[61] Schmidthuber J and Hochreiter S 1997 Neural Comput. 9 1735–80
[62] Vlachas P R, Byeon W, Wan Z Y, Sapsis T P and Koumoutsakos P 2018 Proc. R. Soc. A 474 20170844
[63] Qin H 2019 arXiv:1911.08414
[64] Guth S and Sapsis T P 2019 Entropy 21 925
[65] Vlachas P R, Pathak J, Hunt B R, Sapsis T P, Girvan M, Ott E and Koumoutsakos P 2020 Neurol. Netw. 126 191–217
[66] Sangiorzio M and Dercole F 2020 Chaos Solitons Fractals 139 110045
[67] Wang S, Mu L and Liu D 2021 Comput. Geosci. 149 104695
[68] Yanan G, Xiaoqun C, Baimian L and Kecheng P 2020 Chaotic time series prediction using LSTM with CEEMDAM J. Phys.: Conf. Ser. 1617 012094
[69] Elmam J L 1990 Cogn. Sci. 14 179–211
[70] Werbos P J 1988 Neural Netw. 1 339–56
[71] Zhao Z, Chen W, Wu X, Chen P Y and Liu J 2017 IET Intell. Transp. Syst. 11 68–75
[71] Graves A, Liwicki M, Fernández S, Bertolami R, Bunke H and Schmidhuber J 2008 IEEE Trans. Pattern Anal. Mach. Intel. 31 855–68
[72] Li X and Wu X 2015 Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition 2015 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP) (Piscataway, NJ: IEEE) pp 4520–4
[73] Laptev N, Yoshinski J, Li L and Smyl S 2017 Time-series extreme event forecasting with neural networks at Uber Int. Conf. Machine Learning vol 34 pp 1–5
[74] Le X H, Ho H V, Lee G and Jung S 2019 Water 11 1387
[75] Dixit S, Chowdhury S N, Ghosh D and Shrimali M D 2021 Europhys. Lett. 133 40003
[76] Dixit S, Chowdhury S N, Prasad A, Ghosh D and Shrimali M D 2021 Chaos 31 011105
[77] Kaneko K 1990 Phys. Rev. Lett. 65 1391
[78] All necessary fundamental codes along with the set of initial conditions utilized in this article are available at https://github.com/SayantanNagChowdhury/Sayantan-Nag-Chowdhury-Extreme-events-in-globally-coupled-chaotic-maps
[79] Panchuk A A and Maistrenko Y L 2002 Nonlinear Oscil. 5 326–37
[80] Glendinning P 2003 Stability of asymmetric cluster states in globally coupled maps Proc. Special Workshop on Nonlinear Dynamics of Electronic Systems (NDES’2003) (Scoul, Switzerland 18–23 May 2003) pp 89–92
[81] Massel S R 1996 Ocean Surface Waves: Their Physics and Prediction vol 11 (Singapore: World Scientific)
[82] Akhmediev N and Pelinovsky E 2010 Eur. Phys. J. Spec. Top. 185 1–4
[83] Pisarchik A N, Jaimes-Reátegui R, Sevilla-Escoboza R, Huerta-Cuellar G and Taki M 2011 Phys. Rev. Lett. 107 274101
[84] Freedman D and Diaconis P 1981 Z. Wahrscheinlichkeitstheor. Verwandte Geb. 57 453–76
[85] De Haan L, Ferreira A and Ferreira A 2006 Extreme Value Theory: An Introduction vol 21 (Berlin: Springer)
[86] Antoniou I, Ivanov V V, Ivanov V V and Zrelov P V 2002 Physica D 167 72–85
[87] Corral À 2004 Phys. Rev. Lett. 92 108501
[88] Santhanam M and Kantz H 2008 Phys. Rev. E 78 051113
[89] Bunde A, Eichner J F, Kantelhardt J W and Havlin S 2005 Phys. Rev. Lett. 94 048701
[90] Cohen J E 2015 Am. Stat. 69 5–7
[91] Goodfellow I, Bengio Y, Courville A and Bengio Y 2016 Deep Learning vol 1 (Cambridge, MA: MIT Press)
[92] Sapsis T P 2021 Annu. Rev. Fluid Mech. 53 85–111
[93] Devaine M, Gaillard P, Goude Y and Stoltz G 2013 Mach. Learn. 90 231–60
[94] Nowotarski J, Liu B, Weron R and Hong T 2016 Energy 98 40–9
[95] Mutavhatsindi T, Sigauke C and Mbuvha R 2020 IEEE Access 8 198872–85