Robust Extremes in Chaotic Deterministic Systems

Renato Vitolo,* Mark P. Holland, and Christopher A.T. Ferro

School of Engineering, Computing and Mathematics, University of Exeter, Exeter, UK

(Dated: December 11, 2008)

Abstract

A chaotic deterministic system exhibits robust extremes when the statistics of its extreme values depend smoothly on the system's control parameters. Such robustness can be exploited to enhance the precision and accuracy of statistical estimators. It is conjectured here that robust chaos is sufficient for robust extremes. This is demonstrated for 1D Lorenz maps and illustrated numerically for the Lorenz63 model. It is also outlined how these ideas can be applied to more complex chaotic systems.

PACS numbers: 05.45.-a, 47.52.+j

Keywords: Extreme value statistics, deterministic chaos, Lorenz63 model

*Electronic address: `r.vitolo@exeter.ac.uk`

We consider how extreme values in chaotic deterministic systems respond to changes in the system's parameters. Extreme values can be associated with disruptive or catastrophic events (in the climate system or in financial markets for example) and so their behaviour is of both scientific and societal interest. Extreme-value theory (EVT) provides models for extremes in stationary stochastic processes and progress has been made recently in extending EVT to chaotic deterministic systems [1–5]. Understanding the response of deterministic systems to variation of control parameters leads to the study of bifurcations [6], typically associated with dramatic changes in behaviour [7]. Chaotic systems, however, often respond regularly to parameter variation. This phenomenon is called *robust chaos* [8] and is found in several systems, from neural networks [9] to (piecewise) smooth maps [10, 11]; see [12] and references therein. In this note we take initial steps towards understanding *robust extremes*, where the statistics of a system's extreme values vary smoothly with its parameters. We conjecture that robust chaos is sufficient for robust extremes and demonstrate that this is the case for 1D Lorenz maps.

Robust extremes were discovered recently in the 192-dimensional ordinary differential equation of [13]. This is a model for the atmospheric circulation at mid-latitudes in the Northern Hemisphere, providing a simplified representation of key physical processes (baroclinic conversion, barotropic stabilisation, thermal diffusion and viscous-like dissipation). The model has a robust chaotic attractor for suitable values of a parameter T_E representing baroclinic forcing [14]. The extremes of the system's total energy have a smooth power law in T_E.

The following numerical study illustrates the robustness of extremes for the Lorenz63 model [15]:

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{aligned} \tag{1}$$

derived from the Rayleigh equations for convection in a fluid layer between two plates. Here σ is the Prandtl and ρ the Rayleigh number; see [16] for a dynamical study. For the "classical" values $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$, (1) has a strange attractor [17], which is robust in the sense of [18]. This definition is different from that of [8] although the practical implications are similar.

For each $\rho_j = 27 + 0.05j$ $(j = 0, 1, \ldots, 20)$, time series of length 10^n units are generated from (1) for $n = 5, 8$. The time series consist of values of x, proportional to the intensity of the convective motion [15], recorded every 0.5 time units along an orbit [39]. Maxima from blocks of 1000 time units are extracted from each series [40] and their frequency distributions are modelled with the generalised extreme value (GEV) distribution function

$$G(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right], \tag{2}$$

where $\sigma > 0$, $\xi \neq 0$ and $w_+ = \max\{w, 0\}$. The GEV model is justified by EVT and its parameters (μ, σ, ξ) are estimated by maximum likelihood [19, 20].

The GEV parameter estimates converge to smooth functions of ρ for large n (Fig. 1 (a)-(c)) but estimates of ξ for $n = 5$ oscillate wildly around the "truth" obtained with $n = 8$. Accurate estimation of the tail index ξ is essential for extrapolating return levels to return periods which are larger than the record length [20]. Climatic records seldom contain more than 100 annual maxima [13], so fast convergence of estimators is desirable.

Robustness of extremes can be used to enhance the accuracy and precision of the GEV estimators. If functional forms can be assumed for the variation of the GEV parameters with ρ then the maxima $\{z_t^j\}$ when $\rho = \rho_j$ can be pooled across all j to estimate these functions. Consider $\mu(\rho) = \mu_0 + \mu_1\rho$, $\sigma(\rho) = \sigma_0 + \sigma_1\rho$ and $\xi(\rho) = \xi_0$ for example. The parameters $(\mu_0, \mu_1, \sigma_0, \sigma_1, \xi_0)$ are then estimated by maximising the *global* likelihood

$$\prod_{j,t} \frac{\partial G}{\partial x}(z_t^j; \mu(\rho_j), \sigma(\rho_j), \xi(\rho_j)). \tag{3}$$

Estimates from this pooled model for $n = 5$ (Fig. 1 (d)-(f)) are much closer to the "truth" than the former, independent fits (enhanced accuracy) and uncertainty is greatly reduced (enhanced precision).

In [21] a linear time trend was imposed on T_E in the model of [13] discussed previously. Statistical models like (3) can also be used to analyse extremes in such non-stationary systems. To justify this approach, an *adiabatic Ansatz* is required (beyond robustness of extremes): the trend timescale must be smaller than the sampling time for the upper tail of the energy distribution.

The above examples suggest that robust chaos may be a sufficient condition for robust extremes in deterministic systems. A formal discussion is now required. Let $f_\rho^t : \mathbb{R}^n \to \mathbb{R}^n$

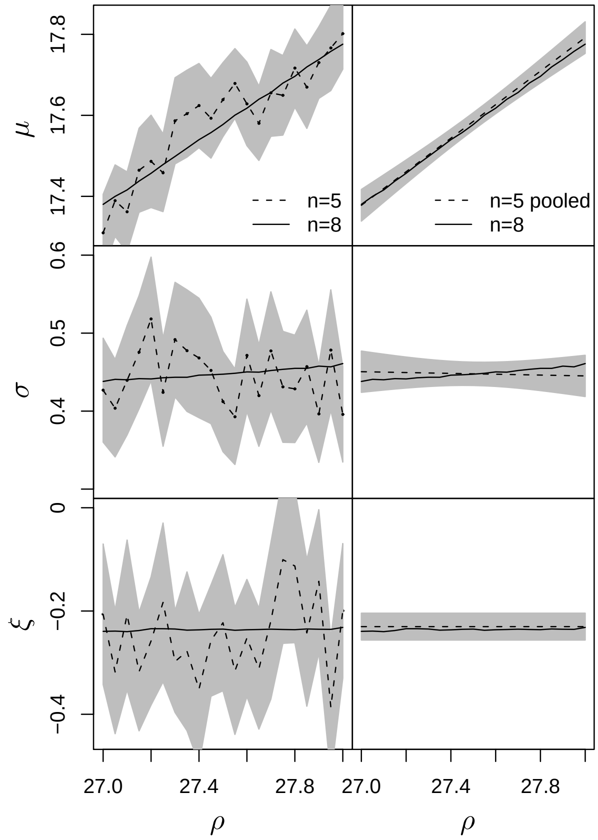

FIG. 1: a)-c) Estimates of the GEV parameters μ, σ, ξ (resp.) vs. the value ρ_j used in (1) for time series length 10^n (see legend) with approximate 95% confidence intervals for $n = 5$. d)-f) Estimates and confidence intervals obtained with (3) are labelled as "pooled".

be a parametrised family of continuous-time dynamical systems defined by the solution of a system of ordinary differential equations in the phase space $\mathbb{R}^n$ smoothly depending on the parameter ρ. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be an observable (the total energy of the system in [13], or the projection on x for Lorenz63). Suppose that for fixed ρ the system has an attractor Λ_ρ: this is a compact invariant subset of $\mathbb{R}^n$ which is transitive (there exists $x \in \Lambda_\rho$ with a dense orbit) and has an open trapping basin (there exists a neighbourhood U of Λ_ρ such that $f^t_\rho(U) \subset U$ for all $t > 0$ see [18]).

To define the statistical properties of the attractor, one **must** assume existence of a Sinai-Ruelle-Bowen (SRB) measure μ_ρ. This is a Borel probability measure in phase space $\mathbb{R}^n$ which is invariant under the flow f^t_ρ, is ergodic, singular with respect to the Lebesgue measure in $\mathbb{R}^n$ and its conditional measures along unstable manifolds are absolutely continuous [22]. Moreover, μ_ρ is a *physical measure*: there is a neighbourhood U of Λ_ρ such that for every continuous observable ϕ the frequency-limit

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \phi(f_\rho^t(x))dt = \int \phi d\mu_\rho \tag{4}$$

holds for orbits $\{f_\rho^t(x)\}_{t\geq 0}$ starting from Lebesgue-almost all $x \in U$. Invariance of the measure implies that the random variables $X_k = \phi \circ f_\rho^{(k-1)t_0}$, with $t_0 > 0$ and $k \in \mathbb{N}$, form a stationary stochastic process on the probability space (Λ_ρ, μ_ρ). Eq. (4) means that the first moment of X_1 (right hand side) can be computed by averaging the observable along a sufficiently long orbit.

If the normalised maxima of a stationary stochastic process converge to a non-degenerate limit distribution H and a mixing condition called $D(u_n)$ holds, then H is a GEV whose tail index is determined by regular variation of the marginal distribution [23]. We now sketch a proof of regular variation for 1D Lorenz maps, which are simplified geometric models for the Lorenz63 flow [6, 16, 24]. Let $T_\gamma : [-1,1] \setminus \{0\} \to [-1,1]$, with $T'_\gamma(x) = l(x)|x|^{\gamma-1}$ for $\gamma \in (0,1)$ and $l(x)$ slowly varying at 0 and Hölder continuous. The map T_γ admits an absolutely continuous invariant measure with density $\theta_\gamma(x) = d\mu_\gamma/dx$ of type *bounded variation* whose support has upper endpoint at $x^+ = 1$, where $\theta_\gamma(x)$ vanishes. For T_γ, the attractor Λ_γ is the whole interval $[-1,1]$. Given an observable ϕ, denote as ϕ_γ^+ the upper endpoint of the distribution function F of $X_1 = \phi$. For $g(u) = 1 - F(\phi_\gamma^+ - 1/u)$ and $t > 0$,

$$\frac{g(ut)}{g(u)} = \frac{\mu_\gamma[\phi(x) > \phi_\gamma^+ - 1/ut]}{\mu_\gamma[\phi(x) > \phi_\gamma^+ - 1/u]}. \tag{5}$$

We distinguish two cases, according to where ϕ is maximised. Assume first $\phi_\gamma^+ = \phi(x^+)$: loosely speaking, the maximum is "on the peel" of the attractor. Consider for example the observable $\phi(x) = x$. Using invariance of μ_γ and the Lebesgue differentiation theorem we obtain

$$\begin{aligned}\mu_\gamma[\phi(x) > \phi_\gamma^+ - 1/ut] &= \mu_\gamma[x > 1 - 1/ut] = \\ &= \mu_\gamma[T_\gamma^{-1}[x^+ - 1/tu, x^+]] = \theta_\gamma(0)h(1/tu)(tu)^{-1/\gamma} \\ &\quad + o\left(h(1/tu)(tu)^{-1/\gamma}\right) \text{ as } u \to \infty,\end{aligned} \tag{6}$$

where $h(1/u)$ is a slowly varying function as $u \to \infty$ and $h(0)$ is bounded between zero and infinity. By (6), Eq. (5) tends to $t^{-1/\gamma}$ as $u \to \infty$, provided $\theta_\gamma(0) \neq 0$ (which numerical experiments indicate is so). This shows that F is regularly varying with index $-1/\gamma$.

Furthermore, the $D(u_n)$ condition holds by decay of correlations in Lorenz maps [25, 26]. Therefore F is in the domain of attraction of a GEV with shape parameter $\xi = -\gamma$ [27].

An even simpler case is when the observable ϕ has a maximum at $x_0 \in [-1,1] \setminus \{0\}$, with $\theta_\gamma(x_0) \neq 0$. The latter condition means, loosely speaking, that x_0 is "within the attractor". Consider $\phi(x) = C - D|x - x_0|^\delta$, with $D, \delta > 0$. By an argument as above, for Eq. (5) we have

$$\frac{\theta_\gamma(x_0) + o(1)}{\theta_\gamma(x_0) + o(1)} t^{-1/\delta} \to t^{-1/\delta} \quad (u \to \infty). \tag{7}$$

The tail index $\xi = -\delta$ is thus constant in γ: it only depends on the "curvature" of the observable ϕ near the extremal point x_0. Both (6) and (7) can be generalised to observables with a strict absolute maximum at x_0 (where $x_0 = x^+$ for (6)) such that $\phi(x) = C - D|x - x_0|^\delta + o(|x - x_0|^\delta)$ as $x \to x_0$ with $\delta > 0$. In summary: if the maxima in the 1D Lorenz maps are GEV-distributed, then the tail index varies smoothly with the system parameter γ for large classes of observables.

Returning to the flow of (1): for $\rho \approx 28$ the attractor $\Lambda_\rho \subset \mathbb{R}^3$ supports a unique SRB measure μ_ρ [28]. Fig. 2 shows a histogram corresponding to the projected measure $\mu_\rho \circ \phi^{-1}$, for $\phi(x,y,z) = x$. A 1D Lorenz map P_γ is obtained by a smooth coordinate transformation S in a Poincaré section for (1). Here $\gamma = |\lambda_s|/\lambda_u$, where λ_u is the unstable and λ_s the weakly stable eigenvalue of the Lorenz63 flow at the origin [16, 26]. Therefore, $\xi = -\gamma$ smoothly depends on ρ for P_γ, for the case "on the peel". The tail index for (1) with $\phi(x,y,z) = x$ is equal to the value $-|\lambda_s|/\lambda_u$ obtained for P_γ with $\phi(x) = x$, multiplied by a factor depending on the coordinate transformation S: robustness of the tail index is preserved, since S is smooth in ρ. Note that the tail index for (1) is non-constant "on the peel" (Fig. 3): the choice of constant ξ and linear μ and σ used for Fig. 1 (d)-(f) is a tangent approximation (valid locally in ρ) for the shapes of the GEV parameters. These are, in general, non-parametric functions of ρ induced by the scaling of the attractor Λ_ρ.

The extreme statistics of Lorenz63 become non-robust under the observable $\phi(x,y,z) = x$ for large ρ, discontinuities become visible near $\rho = 60$. Indeed, (6) no longer holds for $\rho \gtrsim 32$: hyperbolicity is destroyed by the appearance of folds in the return map. The folds correspond to critical points in the 1D Lorenz map P_γ [24]. Therefore, the Loren63 attractor is no longer robust in this parameter regime. This is very different from the case $\rho \approx 28$, for which P_γ is hyperbolic (except at the origin). However, the prediction (7) (maximum "within the attractor") is insensitive to hyperbolicity loss and to the coordinate transformation used to

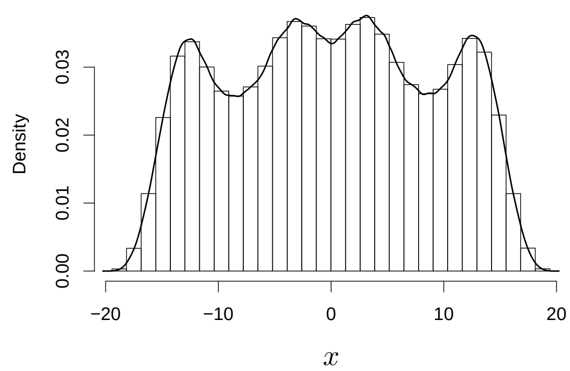

FIG. 2: Histogram and smoothed density from an orbit of Lorenz63 observed through $\phi(x, y, z) = x$ for $\rho = 28$.

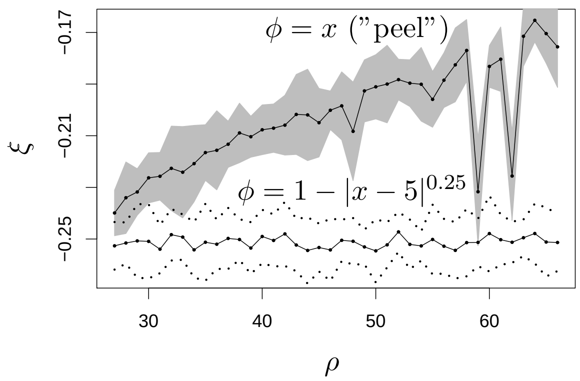

FIG. 3: Maximum likelihood estimates of the tail index ξ for two observables with 10^5 maxima. For each $\rho_j \in [25, 66]$, uncertainty is twice the standard deviation of the 10 estimates obtained with 10 "chunks" of 10^4 maxima.

obtain P_γ. For example, for the observable $\phi(x, y, z) = 1 - (x - 5)^{0.25}$ an average estimate of $\xi = -0.251$ with standard deviation 0.002 is obtained over $\rho \in [27, 66]$, as predicted (Fig. 3). This observable is maximised on the plane $\{x = 5\} \subset \mathbb{R}^3$, which intersects Λ_ρ for all $\rho \in [27, 66]$. In summary, the Lorenz63 model has both robust and non-robust extremes (depending on the parameter range), with constant or non-constant ξ (depending on the observable).

It is of considerable interest to extend these results to other systems and observables. If ϕ is maximised "within the attractor", the argument leading to (7) seems to be valid under weak conditions. The resulting constancy of ξ is a rather general property: for it to hold it is sufficient that the **un**projected invariant measure μ_ρ provides good sampling near the set where ϕ is maximised. Caution is required if this set is a point x_0: in general μ_ρ is supported on a fractal set in phase space and parameter variations may bring x_0 "outside the attractor". For robustness of extremes when ϕ is not maximised within Λ_ρ, the projected measure must be robust at its upper endpoint x^+, as in (6) (x^+ in general depends on ρ). Robustness at x^+ is therefore necessary for physical observables such as energy and vorticity [21, 29], which are unbounded in phase space. Such robustness does not always hold, as shown in Fig. 3. In fact, for low-dimensional systems one would often expect otherwise. Indeed, we have examined the extremes "on the peel" of other chaotic attractors: the Hénon map [6] $(x, y) \mapsto (1 - ax^2, by)$ with $b = 0.05$ and a ranging from 1.43 to 1.44 with step 0.0001; and the autonomous Lorenz84 model [30], varying the baroclinic forcing parameter. In both cases the observable is the projection on x and non-robust extreme statistics are found.

Robustness of extremes may be expected for Axiom-A attractors, since the SRB measure is differentiable in the parameter [31, 32]. The latter property has important physical consequences, such as the existence of Onsager-type or Kramers-Kronig dispersion relations [33]. Unfortunately, for many physical systems of interest it is exceedingly difficult to prove existence of an SRB measure, let alone its differentiability. This difficulty led to the *Chaotic Hypothesis* [34, 35], stating that a many particle system in a stationary state can be regarded, for the purpose of computing macroscopic properties, as a smooth dynamical system with a transitive axiom-A global attractor (a version exists for fluid dynamical systems [36]). In line with this, we conjecture that various types of system may display robust extremes at the experimental, *observable* level: robust chaotic systems [8, 12] (or robust attractors [18]), many-particle and fluid dynamical systems [34, 36], high-dimensional systems [37], geophysical flows [21, 29]. This may also explain the robustness which is observed in Fig. 3 for ρ between 32 and 58.

Are meteo-climatic extremes robust with respect to variations of e.g. atmospheric CO_2, or do they exhibit tipping points (bifurcations)? This issue is of great relevance for prediction and mitigation purposes, as well as for the rigorous quantification of trends in extremes.

RV acknowledges kind support by the Willis Research Network (www.willisresearchnetwork.com).

---

[1] P. Collet, Ergodic Theory Dynam. Systems **21**, 401 (2001).

[2] A. C. M. Freitas and J. M. Freitas (2007), arXiv/0706.3071.

[3] A. C. M. Freitas, J. M. Freitas, and M. Todd (2008), arXiv/0804.2887.

[4] M. Holland, M. Nicol, and A. Török, preprint (2008).

[5] A. C. M. Freitas and J. M. Freitas, Ergodic Theory Dynam. Systems **28**, 1117 (2008).

[6] J. Guckenheimer and P.Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer-Verlag, 1983).

[7] C. Grebogi, E. Ott, and J. A. Yorke, Phys. Rev. Lett. **48**, 1507 (1982).

[8] S. Banerjee, J. A. Yorke, and C. Grebogi, Phys. Rev. Lett. **80**, 3049 (1998).

[9] A. Potapov and M. K. Ali, Physics Letters A **277**, 310 (2000).

[10] M. Andrecut and M. K. Ali, Phys. Rev. E **64**, 025203(R) (2001).

[11] P. Kowalczyk, Nonlinearity **18**, 485 (2005).

[12] Z. Elhadj and J. Sprott, Frontiers of Physics in China **3**, 195 (2008).

[13] M. Felici, V. Lucarini, A. Speranza, and R. Vitolo, J. Atmos. Sci. **64**, 2137 (2007).

[14] V. Lucarini, A. Speranza, and R. Vitolo, Phys. D **234**, 105 (2007).

[15] E. Lorenz, J. Atmos. Sci. **20**, 130 (1963).

[16] C. Sparrow, IEEE Trans. Circuits and Systems **30**, 533 (1983).

[17] W. Tucker, C. R. Acad. Sci. Paris Sér. I Math. **328**, 1197 (1999).

[18] C. A. Morales, M. J. Pacifico, and E. R. Pujals, Proc. Amer. Math. Soc. **127**, 3393 (1999).

[19] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics (Springer, New York, 2001).

[20] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, *Statistics of Extremes: Theory and Applications* (John Wiley and Sons, Berlin, 2004).

[21] M. Felici, V. Lucarini, A. Speranza, and R. Vitolo, J. Atmos. Sci. **64**, 2159 (2007).

[22] L.-S. Young, J. Statist. Phys. **108**, 733 (2002).

[23] M. R. Leadbetter, Z. Wahrsch. Verw. Gebiete **65**, 291 (1983).

[24] S. Luzzatto and W. Tucker, Inst. Hautes Études Sci. Publ. Math. pp. 179–226 (2000) (1999).

[25] A. C. M. Freitas and J. M. Freitas, Statist. Probab. Lett. **78**, 1088 (2008).

[26] S. Luzzatto, I. Melbourne, and F. Paccaut, Comm. Math. Phys. **260**, 393 (2005).

[27] N. H. Bingham, J. Comput. Appl. Math. **200**, 357 (2007).

[28] V. Araujo, M. J. Pacifico, E. R. Pujals, and M. Viana, Trans. Amer. Math. Soc. (2008).

[29] R. Vitolo and A. Speranza, in preparation (2008).

[30] H. Broer, C. Simó, and R. Vitolo, Nonlinearity **15**, 1205 (2002).

[31] D. Ruelle, Ergodic Theory Dynam. Systems **28**, 613 (2008).

[32] D. Ruelle, Comm. Math. Phys. **187**, 227 (1997).

[33] G. Gallavotti and D. Ruelle, Comm. Math. Phys. **190**, 279 (1997).

[34] G. Gallavotti and E. G. D. Cohen, J. Statist. Phys. **80**, 931 (1995).

[35] F. Bonetto, G. Gallavotti, A. Giuliani, and F. Zamponi, J. Stat. Phys. **123**, 39 (2006).

[36] G. Gallavotti, *Foundations of Fluid Dynamics* (Springer Verlag, Berlin, 2002).

[37] D. J. Albers, J. C. Sprott, and J. P. Crutchfield, Phys. Rev. E **74**, 057201 (2006).

[38] À. Jorba and M. Zou, Experiment. Math. **14**, 99 (2005).

[39] The variable-order adaptive-stepsize Taylor method of [38] is used for integration of (1), with local absolute and relative truncation errors fixed at 10^{-8}.

[40] In meteo-climatic applications, where the block length is often 1 year, this is called the annual maximum method.