EPQ

Investigating the Sensitivity of Chaos Theory and the Lorenz System in Atmospheric Modelling

---

Candidate Name: Haolin Guo
Candidate Number: [0000]
Centre Name: [Your school or college]
Centre Number: [00000]
Supervisor: Lisa Bray
Examination Board: [AQA / Edexcel / OCR / CAIE]
Word Count: 3,120
Submission Date:

Contents

1. Introduction

1.1 The Laplacian Assumption

In 1814, Pierre-Simon Laplace imagined an intellect vast enough to know all the data of the universe at a single instant. For this intellect, he claimed, “the future, just like the past, would be present before its eyes”[1]. On this view every state of the universe fully determines every other, and any failure of prediction reflects only measuring devices that are too coarse and computers that are too slow, problems which technological development can solve.

1.2 The Puzzle: Deterministic Physics, Failing Forecasts

Weather should be the perfect testing ground for that assertion. Modern meteorology monitors the atmosphere with satellites and simulates it on some of the world’s most powerful supercomputers, yet forecast skill deteriorates markedly after about ten days, and beyond fourteen it is no better than the seasonal average[2]. Substantial improvements in instruments and computing power have extended that horizon by only about one day per decade[3]. Is the two-week limit a technological barrier or a fundamental one?

1.3 Defining Chaos: Sensitivity to Initial Conditions

Chaos theory offers one answer. Imperfect models, limited computing and uncertain initial data can be improved; sensitivity to initial conditions cannot, because no measurement is exact. A chaotic system is deterministic but highly sensitive to its starting state : two almost identical states can diverge exponentially even though no randomness is introduced. This is the intuition behind Lorenz's butterfly image[4]. Started twice from exactly the same state a chaotic system repeats itself exactly, but that exact state can never be achieved[5]. Chaos converts microscopic error into macroscopic uncertainty.

1.4 Why the Lorenz System

[1] Laplace, A Philosophical Essay on Probabilities (1814).
[2] Zhang et al., ‘What Is the Predictability Limit of Midlatitude Weather?’ (2019).
[3] Bauer, Thorpe and Brunet, ‘The Quiet Revolution of Numerical Weather Prediction’ (2015).
[4] Lorenz, Predictability: Does the Flap of a Butterfly’s Wings in Brazil Set Off a Tornado in Texas? (1972).
[5] Lorenz, The Essence of Chaos (1993).

The Lorenz system is a suitable vehicle for three reasons. It is small — three variables and seven terms, only two of them nonlinear — so its behaviour can be traced to specific terms rather than to obscured complexity. It is derived from atmospheric convection rather than invented as a curiosity[6]. And it is the most studied chaotic system in the literature, which matters practically: an unfamiliar system may be equally chaotic yet offer no published exponent against which a correct implementation can be distinguished from a merely believable one.

1.5 Research Question and Scope

Accompanying this report is an interactive explorer that runs the Lorenz system twice from almost identical initial conditions and measures the rate at which the runs diverge. It asks what the system shows about the way microscopic uncertainty limits long-term prediction, and what that principle can and cannot suggest about the two-week horizon of weather forecasting. The artefact is not a weather model: it measures error growth in an idealised system. An artefact is necessary because the claim is counter-intuitive and easier to show than to explain. The equations contain no stochastic term, so a reader meeting them expects a predictable system and, on being told otherwise, suspects that randomness has been smuggled in elsewhere. Watching two trajectories that are indistinguishable for twelve time units come apart, with nothing added at the moment they part, removes that suspicion in a way prose cannot. Sections 2–4 establish the evidence and method, Section 5 reports the results, and Sections 6–7 evaluate and conclude.

2. Literature Review

The literature is used as an argument: the first sources establish the model and its mathematics, while the final sources qualify what an idealised result can say about the real atmosphere.

2.1 Deterministic Nonperiodic Flow

Lorenz's “Deterministic Nonperiodic Flow” provides the three-equation convection model, the standard parameters, and the original description of sensitive dependence. Its stability analysis supplies the critical value of ρ used later. It is the primary source for what the model

[6] Saltzman, ‘Finite Amplitude Free Convection as an Initial Value Problem—I’ (1962); Lorenz, ‘Deterministic Nonperiodic Flow’ (1963).

is and what was originally claimed; Gleick is used only for the later popular history of its discovery[7]. As a peer-reviewed primary source it carries the highest authority on both points[8], and its age is not a weakness: part of its role here is historical, since it is the discovery itself.

2.2 Historical Context and Significance

Resler situates Lorenz's paper within its historical context and follows its impact beyond meteorology. This justifies the historical approach of the project and the choice of the 1963 model as a case study[9].

2.3 The Mathematical Toolkit: MIT OpenCourseWare

The established ideas of phase space, attractors and Lyapunov exponents are from Chumakova notes from MIT[10]and Strogatz[11]. Details on the Lorenz system are from Sparrow[12], and the measurement of trajectory separation used here is from Wolf et al.[13].

2.4 The Case for an Intrinsic Limit

Zhang et al. repeat Lorenz's approach[14], comparing pairs of global-model simulations that differ only by a tiny initial error. Their distinction between practical and intrinsic predictability motivates the twin-run design used here, though their result remains model-based rather than a direct consequence of this artefact. Their headline finding is that even with near-perfect models and initial conditions, a tenfold reduction in initial error buys only about five additional days of forecast skill[15].

2.5 The Modern Debate

Shen, Pielke Sr. and Zeng show that the Lorenz 1963 and 1969 models[16] imply finite predictability for different reasons[17]. Palmer, Döring and Seregin likewise distinguish

---

[7]Gleick, Chaos: Making a New Science (1987).

[8]Lorenz, 'Deterministic Nonperiodic Flow' (1963).

[9]Resler, 'Edward N Lorenz's 1963 Paper, "Deterministic Nonperiodic Flow": Its History and Relevance to Physical Geography' (2016).

[10]Chumakova, 18.353J Nonlinear Dynamics I: Chaos — Lecture Notes (2012).

[11]Strogatz, Nonlinear Dynamics and Chaos, 2nd edn (2015).

[12]Sparrow, The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors (1982).

[13]Wolf et al., 'Determining Lyapunov Exponents from a Time Series' (1985).

[14]Lorenz, 'Atmospheric Predictability Experiments with a Large Numerical Model' (1982).

[15]Zhang et al., 'What Is the Predictability Limit of Midlatitude Weather?' (2019).

[16]Lorenz, 'The Predictability of a Flow Which Possesses Many Scales of Motion' (1969).

[17]Shen, Pielke Sr. and Zeng, 'One Saddle Point and Two Types of Sensitivities within the Lorenz 1963 and 1969 Models' (2022).

exponential error growth from an absolute finite-time barrier[18]. The artefact therefore measures only the Lorenz-system exponent and quotes atmospheric conclusions from the literature. This is the single most important constraint on the artefact's design: it is why the program computes an exponent and a doubling time for the Lorenz system alone, and why any figure for the real atmosphere is cited rather than calculated here.

3. Mathematical Framework

The Lorenz system is three coupled ordinary differential equations: $\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma(y - x)$, $\frac{\mathrm{d}y}{\mathrm{d}t} = x(\rho - z) - y$, $\frac{\mathrm{d}z}{\mathrm{d}t} = xy - \beta z$. The variables describe a convecting fluid at one instant: x is the intensity of the overturning motion, y the temperature difference between the ascending and descending currents, and z the departure of the vertical temperature profile from linear. There is no random variable in the system, which makes it deterministic, and the two product terms xz and xy make it nonlinear. This matters because a linear system cannot produce chaos, so any chaotic behaviour must arise from those terms.

At any given time, the state is a single point in three-dimensional phase space with coordinates x, y and z, and as time progresses the point moves along a trajectory. Determinism implies that only one trajectory can pass through any given point, so trajectories cannot intersect. The set of states onto which the system settles after transients have died out is its attractor. The Lorenz attractor is a strange attractor: bounded, roughly shaped like a butterfly, and never repeating. The largest Lyapunov exponent λ measures the sensitivity to initial conditions, which causes two trajectories with initial separation δ_0 to diverge roughly as $\delta_0 e^{\lambda t}$; a positive λ is indicative of chaos, and errors double about every $\frac{\ln 2}{\lambda}$ time units.

This section is intentionally brief. The fuller treatment — the derivation from Rayleigh–Bénard convection, the stability analysis that fixes the critical value of ρ, and the full account of phase space and attractors — is delivered by the artefact beside the figure it explains. What remains here is only what the methodology requires.

4. Methodology

[18] Palmer, Döring and Seregin, ‘The Real Butterfly Effect’ (2014).

This method can be viewed as a series of design choices: language, integrator, parameters, experiment, measurement, validation, and numerical error.

4.1 Justification of Java

Python was used only to convert completed CSV files into PNG figures and to re-plot the fitted line; it was not the numerical solver, so it cannot accurately be blamed for a solver failure. Java performs the simulation so that the fixed-step RK4 update, the common time grid and the floating-point arithmetic are explicit in the source. Java's double-precision arithmetic conforms to the same IEEE 754 standard as Python's, so the language itself does not alter the numbers.

4.2 The Runge–Kutta Method and the Fixed Step

The integrator is the classic fourth-order Runge–Kutta (RK4) method[19]. It evaluates the derivative at the start of the step, at two estimated midpoints and at an estimated endpoint, weighting the midpoints twice. The local truncation error is $O(dt^5)$ and the error accumulated over a fixed interval is $O(dt^4)$, so halving the step reduces the accumulated error by a factor of sixteen rather than two. Embedded Runge–Kutta methods such as Dormand–Prince[20] offer a useful alternative, but not the shared grid required here. The comparison that justifies the cost is with Euler's method, which evaluates the derivative only at the start of the step and accumulates error at first order: four derivative evaluations per RK4 step buy far more accuracy than four times as many Euler steps.

The step is fixed by the measurement, not by efficiency. An adaptive solver can compare the trajectories, but it will choose different steps for each run, so the separation must then be aligned or interpolated onto a common grid — possible, but it adds a processing stage and an interpolation error to $\delta(t)$. With $dt = 0.001$ both trajectories are recorded at identical instants by construction, at the cost of unnecessary work in smooth stretches.

4.3 Selection of Parameters and Step Size

[19] Press et al., Numerical Recipes: The Art of Scientific Computing, 3rd edn (2007).
[20] Dormand and Prince, ‘A Family of Embedded Runge–Kutta Formulae’ (1980).

To allow comparison with published values, the original parameters of Lorenz[21] are retained: $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$. For these values of σ and β the non-zero fixed points become unstable at about $\rho = 24.74$, so $\rho = 28$ places the system in the chaotic regime.

The choice $dt = 0.001$ was verified by halving it and observing no significant change in the exponent. The 10^{-5} perturbation in x satisfies three order-of-magnitude conditions: it is small against an attractor of order ten, so it represents a measurement error rather than a substantially different initial condition; it is large against the 10^{-16} round-off noise of double-precision arithmetic, so what grows is the perturbation rather than the round-off; and it leaves about five orders of magnitude of uncorrupted growth before saturation. A similar result was obtained when perturbing y, which is unsurprising, since a small error aligns quickly with the fastest-growing direction.

4.4 Experimental Design

Two trajectories are run from the same state on the attractor, with x in B increased by 10^{-5}. Their Euclidean separation $\delta(t)$ is recorded at every step. Following Wolf et al.[22], the three-dimensional distance is used so that the result does not depend on a single coordinate. Spin-up occurs before the measured run so that the exponent is not dominated by the approach to the attractor. Figure 1 shows the attractor that spin-up is designed to reach.

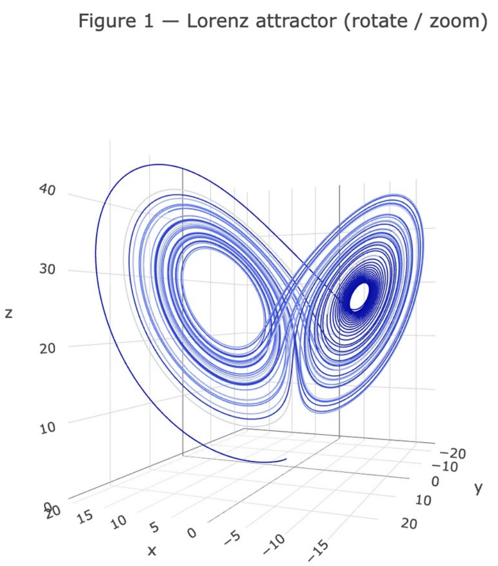

Figure 1. The Lorenz attractor as visualised in the artefact's browser viewer. The trajectory never leaves a bounded region of phase space, yet never repeats and never crosses itself. Spin-up is complete once the state lies on this surface, which is the condition Section 5.1 depends on.

[21] Lorenz, ‘Deterministic Nonperiodic Flow’ (1963).
[22] Wolf et al., ‘Determining Lyapunov Exponents from a Time Series’ (1985).

4.5 Measuring the Exponent

The exponent is the slope of ln δ against time, but the curve is not linear over its whole range, so the fit excludes the early alignment transient and the flat saturation tail. The rule was fixed in advance: begin at the first point after $t = 2$ and stop permanently at the first point where $\delta \geq 1$. The measurement is therefore reproducible rather than selected by eye.

4.6 Testing and Validation

Five checks were used: derivatives at the non-zero fixed points, convergence on halving dt, comparison with Sprott's published value[23], R^2 for every fit, and a repeat with the perturbation in y. Their numerical outcomes are reported in Section 5 rather than repeated here. Two are worth stating because they could have failed badly. The fixed-point check is evaluated and printed on every run, so a mistyped equation is caught before any trajectory is calculated; and reporting R^2 for every fit is what exposed the spin-up problem numerically, rather than letting it pass as a plausibly small exponent.

4.7 Numerical Error

Rounding and discretisation errors increase for the same reason as the deliberate perturbation, and a smaller step only postpones the loss of an exact trajectory. The result does not depend on any one path remaining exact, but on the rate of separation and the shape of the attractor; it therefore reports an exponent and a doubling time rather than forecasting a state at long lead time.

5. Results

5.1 A representative experiment

Member 1 starts from (1, 1, 1) with a spin-up of 30 time units and a perturbation of 10^{-5} in x, so the initial $\delta_0 = 1.000 \times 10^{-5}$. The predetermined fit runs from $t = 2.001$ to $t = 11.915$, using 9,915 of the 50,001 recorded points. A least-squares fit of ln δ against t yields a slope of $\lambda = 0.972$ with a coefficient of determination $R^2 = 0.912$, so the doubling time is $\ln 2/\lambda = 0.71$ model time units. After about $t = 12$, δ saturates, reaching 31.2 at $t = 50$, the order of the attractor's own width. Figure 2 shows the two trajectories for x, and Figure 3 the separation on a logarithmic scale.

[23] Sprott, Chaos and Time-Series Analysis (2003).

As a negative control, the same analysis without spin-up was performed over the much longer period t = 2.001 to 21.556, starting from (1, 1, 1) as before. This yielded λ = 0.441 with R^2 = 0.751. The lower value and the poorer straight-line fit indicate that this early portion of the run has not yet reached the attractor but is still approaching it. Reporting R^2 therefore turns a misleading result into a detectable flaw in the experimental design.

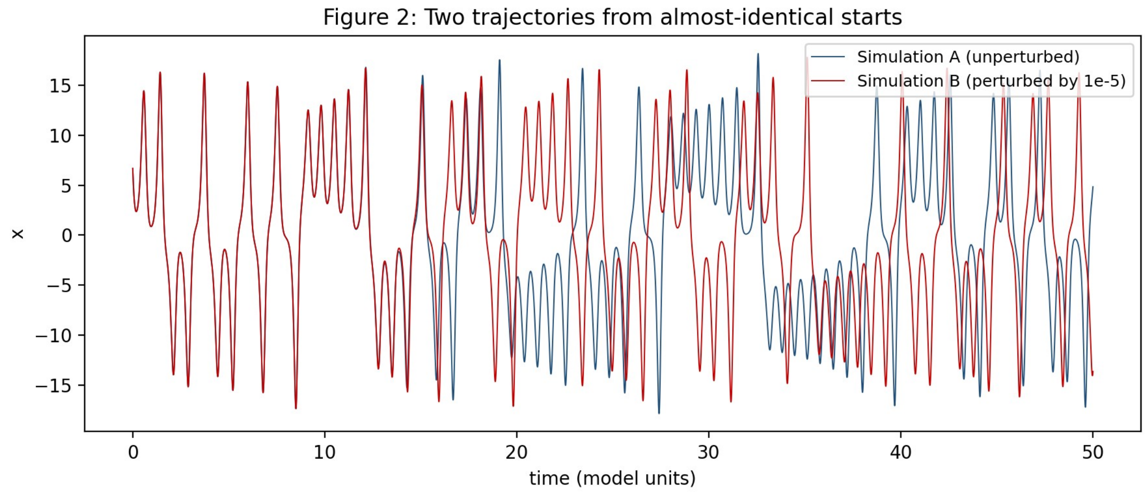

Figure 2. The two x trajectories in the representative spun-up run. The small initial difference is not apparent until after exponential growth.

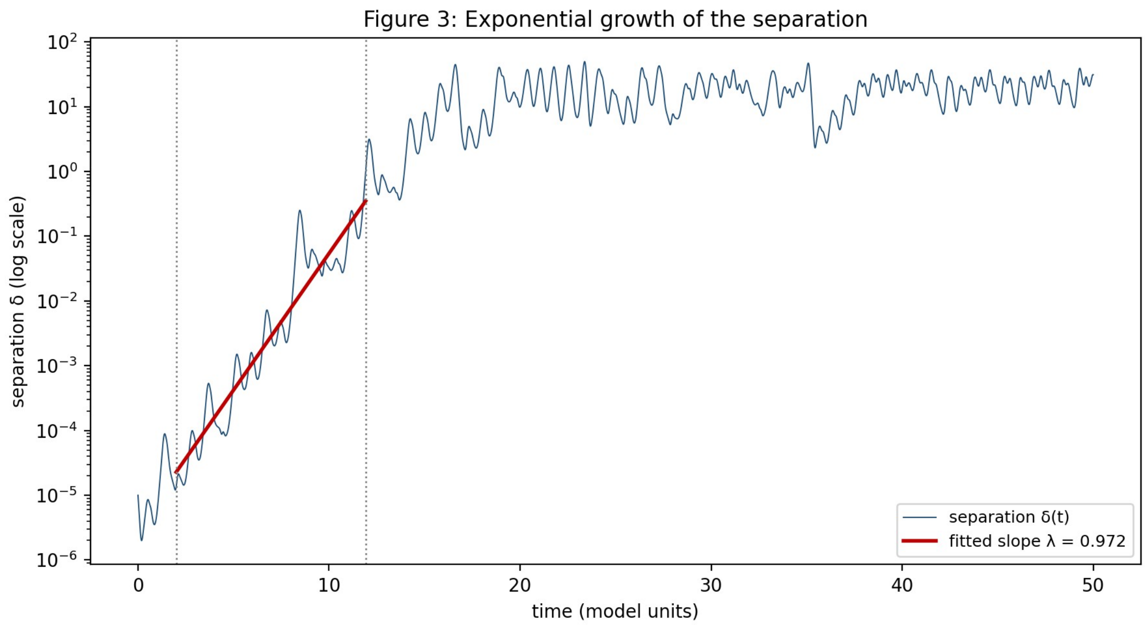

Figure 3. Separation of the representative pair in logarithmic scale. A straight segment is fitted from t = 2.001 to 11.915 before saturation.

5.2 The full ensemble of eight

Member	Spin-up (t.u.)	λ	R^2	Doubling time
1	30	0.972	0.912	0.71

2	35	0.853	0.866	0.81
3	40	0.910	0.880	0.76
4	45	0.827	0.870	0.84
5	50	0.866	0.846	0.80
6	55	0.681	0.746	1.02
7	60	0.615	0.742	1.13
8	65	0.937	0.936	0.74
Mean		0.833 ± 0.124		0.83

The range of the eight members is 0.615 to 0.972, so no single run can be taken as the measurement. The mean is 0.833 with a sample standard deviation of 0.124, giving a standard error of 0.044. Sprott's published value[24] of 0.906 lies 0.59 standard deviations from the mean, and so agrees within the spread of this ensemble. At the mean rate, a measurement error of 10^{-5} propagates to $\delta = 1$ in about 13.8 model time units. The mean doubling time is reported as ln 2 divided by the mean exponent rather than as the mean of the eight individual doubling times, which is slightly larger at 0.85.

5.3 What the spread is made of

The spread is not without structure. Across the eight members the fitted exponent and the quality of the fit move together almost exactly: the Pearson correlation between λ and R^2 is 0.96, and the two members with the lowest exponents, 0.615 and 0.681, also have the lowest R^2, 0.742 and 0.746. The six members with R^2 above 0.80 have exponents ranging from 0.827 to 0.972.

This suggests that part of the fifteen per cent spread is a measurement artefact rather than a real difference in the attractor's local stretching rate: where the log-separation curve is least linear, the single-slope approximation fits worst and the resulting slope is depressed. Restricting the ensemble to those six members yields $\lambda = 0.894 \pm 0.055$, both tighter and closer to Sprott's 0.906[25]. The eight-member mean is nevertheless retained as the reported result, because excluding members on a threshold chosen after the answers were known would be exactly the after-the-fact selection that a fixed fit rule exists to prevent. The comparison is reported only to show that the spread has a traceable cause, and to motivate the renormalisation procedure described in Section 7.3.

6. Evaluation

[24] Sprott, Chaos and Time-Series Analysis (2003).
[25] Sprott, Chaos and Time-Series Analysis (2003).

6.1 Strengths

The main strengths are reproducibility and honesty of scope. The measurement rule is fixed beforehand; the Java output is saved as CSV; the viewer and plotting script can re-analyse it; and uncertainty is reported rather than hidden. The viewer exposes the fit window, fitted value, and recorded run configuration so that a result can be checked against its input data.

6.2 Limitations

The ensemble spread is about fifteen per cent of the mean, so agreement with the published value is a weak test; a thirty-two-member ensemble would make it a reasonably strong one. Only one exponent method is used, and only the largest exponent is measured — the full spectrum would also give the dimension of the attractor. The three-variable Lorenz system cannot represent the atmosphere's effectively infinite degrees of freedom.

Zhang et al. argue that the midlatitude limit is intrinsic[26], but Shen, Pielke Sr. and Zeng show that idealised models can reach finite predictability by different mechanisms[27]. What transfers from this artefact is the principle of exponential error growth, not its numerical value.

7. Conclusion

7.1 Determinism Without Predictability

The Lorenz system is deterministic, but it is not predictable beyond a horizon set by its own dynamics. The ensemble exponent, 0.833 ± 0.124, gives a doubling time of about 0.83 model time units, so an initial error of 10^{-5} becomes of order one after about fourteen model time units. Determinism and predictability are therefore distinct claims.

7.2 Real-World Application

The error doubles on a fixed timescale, so halving the initial error gains one doubling time rather than doubling the forecast horizon; even a thousandfold improvement in measurement gains only about ten doubling times. The atmospheric figures make this concrete: Lorenz estimated the doubling time of small errors in operational forecasts at about 2.4 days[28], and Zhang et al. find that even with near-perfect models a tenfold reduction in initial error

[26] Zhang et al., ‘What Is the Predictability Limit of Midlatitude Weather?’ (2019).
[27] Shen, Pielke Sr. and Zeng, ‘One Saddle Point and Two Types of Sensitivities within the Lorenz 1963 and 1969 Models’ (2022).
[28] Lorenz, ‘Atmospheric Predictability Experiments with a Large Numerical Model’ (1982).

extends useful skill by only about five days[29]. This is the reasoning behind ensemble prediction, in which many closely neighbouring initial states are run and their spread reported, rather than a single apparently precise forecast being taken at face value. Ensemble prediction is this artefact's twin-run experiment at industrial scale.

7.3 Future Improvement

A renormalising procedure as suggested by Wolf et al.[30]would avoid saturation and provide an independent estimate of the exponent. Running the Lorenz 1969 model alongside the 1963 model would compare different routes to finite predictability, and a second independent implementation would strengthen the validation. None of these would alter the central conclusion: the limit belongs to the dynamics and not merely to the measuring instrument.

Bibliography

Bauer, P., Thorpe, A. and Brunet, G. (2015) 'The quiet revolution of numerical weather prediction', Nature, 525, pp. 47–55. doi:10.1038/nature14956.

Chumakova, L. (2012) 18.353J Nonlinear Dynamics I: Chaos — Lecture Notes. MIT OpenCourseWare. Available at: https://ocw.mit.edu/courses/18-353j-nonlinear-dynamics-and-chaos-fall-2012/ (Accessed: 10 August 2026).

Dormand, J. R. and Prince, P. J. (1980) 'A family of embedded Runge–Kutta formulae', Journal of Computational and Applied Mathematics, 6(1), pp. 19–26. doi:10.1016/0771-050X(80)90013-3.

Gleick, J. (1987) Chaos: Making a New Science. New York: Viking.

[29]Zhang et al., 'What Is the Predictability Limit of Midlatitude Weather?' (2019).
[30]Wolf et al., 'Determining Lyapunov Exponents from a Time Series' (1985).

Laplace, P.-S. (1814) A Philosophical Essay on Probabilities. Translated by F. W. Truscott and F. L. Emory. New York: Dover.

Lorenz, E. N. (1963) 'Deterministic Nonperiodic Flow', Journal of the Atmospheric Sciences, 20(2), pp. 130–141.

Lorenz, E. N. (1969) 'The predictability of a flow which possesses many scales of motion', Tellus, 21(3), pp. 289–307. doi:10.1111/j.2153-3490.1969.tb00444.x.

Lorenz, E. N. (1972) Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas? Address to the American Association for the Advancement of Science, Washington DC, 29 December. Reprinted in Lorenz (1993), Appendix 1.

Lorenz, E. N. (1982) 'Atmospheric predictability experiments with a large numerical model', Tellus, 34(6), pp. 505–513. doi:10.1111/j.2153-3490.1982.tb01839.x.

Lorenz, E. N. (1993) The Essence of Chaos. Seattle: University of Washington Press.

Palmer, T. N., Döring, A. and Seregin, G. (2014) 'The real butterfly effect', Nonlinearity, 27(9), pp. R123–R141. doi:10.1088/0951-7715/27/9/R123.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2007) Numerical Recipes: The Art of Scientific Computing. 3rd edn. Cambridge: Cambridge University Press.

Resler, L. M. (2016) ''Edward N Lorenz's 1963 paper, 'Deterministic nonperiodic flow': Its history and relevance to physical geography'', Progress in Physical Geography, 40(1), pp. 175–180. doi:10.1177/0309133315623099.

Saltzman, B. (1962) 'Finite Amplitude Free Convection as an Initial Value Problem—I', Journal of the Atmospheric Sciences, 19(4), pp. 329–341.

Shen, B.-W., Pielke Sr., R. A. and Zeng, X. (2022) 'One Saddle Point and Two Types of Sensitivities within the Lorenz 1963 and 1969 Models', Atmosphere, 13(5), 753. doi:10.3390/atmos13050753.

Sparrow, C. (1982) The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors. New York: Springer-Verlag.

Sprott, J. C. (2003) Chaos and Time-Series Analysis. Oxford: Oxford University Press.

Strogatz, S. H. (2015) Nonlinear Dynamics and Chaos. 2nd edn. Boulder, CO: Westview Press.

Wolf, A., Swift, J. B., Swinney, H. L. and Vastano, J. A. (1985) 'Determining Lyapunov exponents from a time series', Physica D, 16(3), pp. 285–317. doi:10.1016/0167-2789(85)90011-9.

Zhang, F., Sun, Y. Q., Magnusson, L., Buizza, R., Lin, S.-J., Chen, J.-H. and Emanuel, K. (2019) 'What Is the Predictability Limit of Midlatitude Weather?', Journal of the Atmospheric Sciences, 76(4), pp. 1077–1091. doi:10.1175/JAS-D-18-0269.1.