University of Reading

Atmospheric predictability: revisiting the inherent finite-time barrier

Article

Accepted Version

Leung, Tsz Yan ORCID logoORCID: https://orcid.org/0000-0003-0056-284X, Leutbecher, M., Reich, S. and Shepherd, T. G. (2019) Atmospheric predictability: revisiting the inherent finite-time barrier. Journal of the Atmospheric Sciences, 76 (12). pp. 3883-3892. ISSN 1520-0469 doi: https://doi.org/10.1175/JAS-D-19-0057.1 Available at https://centaur.reading.ac.uk/86240/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1175/JAS-D-19-0057.1

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

University of Reading

Central Archive at the University of Reading

Reading's research outputs online

Atmospheric predictability: revisiting the inherent finite-time barrier

Tsz Yan Leung*

Department of Mathematics and Statistics, University of Reading, United Kingdom

Martin Leutbecher

European Centre for Medium-range Weather Forecasts, Reading, United Kingdom

Sebastian Reich

Institute of Mathematics, University of Potsdam, Germany

Theodore G. Shepherd

Department of Meteorology, University of Reading, United Kingdom

Corresponding author address: Department of Mathematics and Statistics, Whiteknights, PO Box 220, Reading RG6 6AX, United Kingdom

E-mail: tsz.leung@pgr.reading.ac.uk

ABSTRACT

The accepted idea that there exists an inherent finite-time barrier in deterministically predicting atmospheric flows originates from Edward N. Lorenz's 1969 work based on two-dimensional (2D) turbulence. Yet, known analytic results on the 2D Navier-Stokes (N-S) equations suggest that one can skilfully predict the 2D N-S system indefinitely far ahead should the initial-condition error become sufficiently small, thereby presenting a potential conflict with Lorenz's theory. Aided by numerical simulations, the present work re-examines Lorenz's model and reviews both sides of the argument, paying particular attention to the roles played by the slope of the kinetic energy spectrum. It is found that when this slope is shallower than -3, the Lipschitz continuity of analytic solutions (with respect to initial conditions) breaks down as the model resolution increases, unless the viscous range of the real system is resolved – which remains practically impossible. This breakdown leads to the inherent finite-time limit. If, on the other hand, the spectral slope is steeper than -3, then the breakdown does not occur. In this way, the apparent contradiction between the analytic results and Lorenz's theory is reconciled.

1. Introduction

Now an accepted fact in dynamical meteorology, the existence of an inherent finite-time barrier in predicting atmospheric flows was first conceptually shown by Lorenz (1969). Using a simple model, he estimated the predictability limit to be slightly over two weeks – a result echoed by recent studies with real-world operational models at major numerical weather prediction centres (Buizza and Leutbecher 2015; Judt 2018; Selz 2019; Zhang et al. 2019). Although advances in probabilistic prediction make it possible to extract predictable signals beyond this limit, the extended predictability mainly results from temporal averaging of the predicted fields, together with the slowly varying components of the climate system (Buizza and Leutbecher 2015). Moreover, the loss of information in probabilistic prediction is reflected in the growth of deterministic error, and under statistically stationary conditions, saturation of the error spectrum corresponds to the predicted probability distribution matching that of the climatology, so that the loss of deterministic and probabilistic predictability are matched. Indeed, empirical evidence suggests that the decay of forecast skill behaves broadly similarly across deterministic and probabilistic predictions (Buizza and Leutbecher 2015). Thus the study of deterministic error growth can be used to understand the mechanisms limiting the range of predictability to a finite horizon and the role of multi-scale interactions in error growth (Rotunno and Snyder 2008; Durran and Gingrich 2014; Sun and Zhang 2016). Such an analysis requires averaging over multiple cases to ensure robust results, which is a somewhat different notion of predictability compared to the fully probabilistic notion used today. However, since the earlier works adopted the deterministic approach with the averaging, we shall take the same approach as we revisit their works.

In his original work, Lorenz (1969) classified fluid systems into two categories:

- those whose error at any future time can be made arbitrarily small by suitably reducing the initial error, and

- those whose error at any future time cannot be reduced below a certain limit unless the initial error is zero.

Using an appropriate skill function, these systems can be equivalently characterised in terms of *range of predictability* (or simply *predictability*): the former category has an indefinite range and the latter has only an inherently finite range. The reader is referred to Appendix A for a motivation of the concept and more details about the skill function. By modelling atmospheric flows by the two-dimensional (2D) barotropic vorticity equation and assuming a $-\frac{5}{3}$ spectral slope along the inertial range of the kinetic energy (KE) spectrum of the unperturbed flow, he argued that such flows have an inherently finite range of predictability. (The inertial range is a continuous part of the spectrum where a specific power-law relationship is followed so that the flow restricted to such scales is self-similar. It is identified by a spectral slope which is the slope of the spectrum as appearing in a log-log plot.)

Although the barotropic vorticity equation with large-scale forcing produces a steeper spectral slope of -3, and unbalanced dynamics are required to produce a spectral slope of $-\frac{5}{3}$ in more realistic models (Sun and Zhang 2016), it has been shown that predictability is determined much more by the spectral slope than by the nature of the dynamics (Rotunno and Snyder 2008). Thus, it is appropriate to use the barotropic vorticity equation to study predictability with a range of spectral slopes, recognising that this only addresses one aspect of what limits atmospheric predictability in practice.

Closely related to this system are the incompressible 2D Navier-Stokes (2D N-S) equations, whose well-posedness (existence of a unique solution to the initial-value problem that depends

continuously on the initial conditions) was first rigorously shown by Ladyzhenskaya also in the second half of the twentieth century (Robinson 2001). As we will see in Section 4, it is not difficult to show that well-posedness implies an indefinite range of predictability in the sense of Lorenz.

The present paper aims to reconcile the difference between the inherently finite predictability result of Lorenz and the indefinite predictability corollary of Ladyzhenskaya's proof, in the context of incompressible 2D flows. Section 2 reviews Lorenz's argument of its inherent finite-time behaviour. In Section 3 we reproduce Lorenz's numerical results and discuss the predictability in the directly simulated 2D barotropic vorticity model. An account of the well-posedness and indefinite predictability of the incompressible 2D N-S equations is presented in Section 4, with which we reconcile Lorenz's result of inherently finite predictability in Section 5. The major findings are summarised in Section 6.

2. Lorenz's argument of inherently finite predictability

The model of Lorenz (1969) is based on the dimensionless 2D barotropic vorticity equation

$$\frac{\partial \theta}{\partial t} + J(\psi, \theta) = 0, \qquad \theta = \Delta\psi \tag{1}$$

where ψ is the velocity streamfunction (related to the velocity $\boldsymbol{u}$ by $\boldsymbol{u} = -\nabla \times (\psi\hat{\mathbf{k}})$), $\Delta = \nabla \cdot \nabla, \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$ and $J(A,B) = \frac{\partial A}{\partial x}\frac{\partial B}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial B}{\partial x}$. Assuming a doubly periodic domain, Lorenz expanded the variables ψ and θ in Fourier series and re-wrote the linearised error equation of (1) in Fourier components. Then he made various assumptions to an ensemble of error fields for the linearised error equation of (1), most notably homogeneity and a slight generalisation of the quasi-normal closure. The resulting equation was then passed into the large-domain and continuous-spectrum limit.

The derivation is arguably more straightforward if the domain is the whole $\mathbb{R}^2$ space and the variables are Fourier-transformed rather than expanded in Fourier series. We have checked that this method returns the same equation as the limiting equation of Lorenz, up to a constant multiplicative factor.

A further assumption of isotropy simplifies the equation, which was then discretised and numerically approximated. Depending on the specification of a KE spectrum for the unperturbed flow, a matrix of constant coefficients C was constructed so that the vector Z of error KE at different scales (each scale K collectively represents wavenumbers $k = 2^{K-1}$ to $k = 2^K$) evolves according to the linear model

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} Z = CZ, \qquad \text{or equivalently} \qquad \frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} Z \\ W \end{pmatrix} = \begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix} \begin{pmatrix} Z \\ W \end{pmatrix}, \tag{2}$$

where W is the first time-derivative of Z.

As Rotunno and Snyder (2008) mentioned, the computation of C involves computing integrals of nearly singular functions. We have been cautious about these integrations and have made sure that our integrations for C are accurate, some details of which are provided in Appendix B.

To time-integrate equation (2), it is necessary that the initial conditions for Z and W are specified. Lorenz did not explicitly give an initial condition for W, but as Rotunno and Snyder (2008) assumed $W(t = 0) \equiv 0$ in their predictability experiments, we shall prescribe the same for our numerical simulations (Section 3(a)). The non-linear effects were accounted for by removing the corresponding components of Z, W and C when the error KE saturated at a particular scale, whereupon an inhomogeneous forcing term was added to the right-hand-side of equation (2) to account for the saturated scale's contribution to the error growth at the unsaturated scales (details available in Appendix C). Time-integration with the resulting lower-dimensional system was carried on, until all scales became saturated. The evolution of the error KE spectrum in time is depicted in

Figure 1 (for illustrative purposes only; this is for a spectral slope of -3 whose simulation Lorenz excluded in his original work).

As Lorenz noted down the saturation times t_K of scale K, he found that the successive differences $t_K - t_{K+1}$ behaved approximately proportional to $2^{-\beta K}$ with β depending on the spectral slope. He therefore concluded that, given an initial error at an infinitesimally small scale, the range of predictability is inherently finite if and only if the telescoping series

$$t_K = \sum_{j=K}^{\infty} (t_j - t_{j+1}) = \sum_{j=K}^{\infty} 2^{-\beta j} \tag{3}$$

is summable, which is the case if and only if $\beta > 0$. By observing $\beta = \frac{2}{3}$ for the atmospherically relevant spectral slope of $-\frac{5}{3}$, he suggested inherently finite predictability for the atmosphere. Additionally, he found that $\beta = \frac{1}{3}$ for a hypothetical spectral slope of $-\frac{7}{3}$. Lorenz thus hypothesised by linear extrapolation that the range of predictability would be indefinite if the spectral slope were steepened to -3.

This result is echoed by arguments on dimensional grounds (Vallis 1985; Lilly 1990). Assuming that t_K in equation (1) depends only on the wavenumber k and the one-dimensional KE spectral density $E(k)$ of the background flow, one has $t_K \sim (k^3 E(k))^{-0.5}$ as this is the only way the physical units of k and $E(k)$ can combine to give the dimension of time. With $k \sim 2^K$, one obtains $t_K \sim 2^{-\beta K}$, with $\beta = \frac{2}{3}$, $\frac{1}{3}$ and 0 for the spectral slopes of $-\frac{5}{3}$, $-\frac{7}{3}$ and -3 respectively, the same result as Lorenz's.

3. Numerical simulations

We performed a series of numerical simulations, first on the Lorenz model (2) followed by a forced-dissipative version of the full 2D barotropic vorticity system (1), to see whether indefinite predictability is indeed achieved with a KE spectral slope of -3 as Lorenz hypothesised.

a. Lorenz's model

Rotunno and Snyder (2008) solved for the growth of the error KE spectrum for a background spectral slope of $-p$ where $p = 3$. In order to assess the range of predictability in Lorenz's framework, we extended their calculations to study the relationship between K and t_K.

Having computed the matrix C as in Rotunno and Snyder (2008), we solved the linear matrix system (2) explicitly, that is, by writing out the general solution in terms of the eigenvalues and eigenvectors of

$$\begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}$$

and projecting the initial condition onto such an eigenspace to determine the constants of the general solution. This exact approach is a good and easy alternative to the numerical schemes used by Lorenz (1969), Rotunno and Snyder (2008) and its extension by Durran and Gingrich (2014). Details of the solution procedure can be found in Appendix C.

Figure 1 shows the evolution of the error for the -3 spectrum as in Rotunno and Snyder (2008), and Figure 2 shows the saturation times t_K as a function of the scale K. Note that in Figure 2 t_K is plotted instead of $t_K - t_{K+1}$ against K, but the choice makes little difference when $\beta > 0$ since if $t_K - t_{K+1}$ is proportional to $2^{-\beta K}$ then so is t_K (cf. equation (3)). It is clear that the saturation times t_K scale as $2^{-\beta K}$ with a small but positive β (0.05) along the inertial range, so that the sum in equation (3) is still finite for $p = 3$, contrary to Lorenz's prediction. Indeed, arguing in the same way as Lorenz, our result indicates inherently finite predictability for a -3 spectrum which is contrary to Lorenz's hypothesis, although we acknowledge that $\beta = 0.05$ is just marginally away from the critical value of zero. We did, however, recover Lorenz's result for the case of a $-\frac{5}{3}$ spectrum (not shown).

b. Forced-dissipative 2D barotropic vorticity equation

The difference between our value of β and Lorenz's for the -3 spectrum deserves particular attention because it amounts to a qualitative contrast between inherently finite and indefinite ranges of predictability. To further investigate this, we performed direct numerical simulations (DNS) on this $p = 3$ spectrum in the form of identical-twin experiments (pairs of runs which only differ in the initial condition), and assessed the predictability following Lorenz's methodology with necessary adaptations.

First of all, equation (1) had to be restricted to a doubly periodic domain and be made forced-dissipative:

$$\frac{\partial \theta}{\partial t} + J(\psi, \theta) = f + d, \qquad \theta = \Delta\psi. \tag{4}$$

The forcing and dissipation, however small, are necessary for generating statistically stationary KE spectra in the DNS. To generate a -3 spectral slope, following standard practice (Maltrud and Vallis 1991), forcing was applied at the large scale: $f(t)$ was chosen to be an independent white-noise process for each 2D wavevector whose scalar wavenumber k falls in the narrow band ($\pm 10\%$) around $k = 20$. The dissipation d was a highly scale-selective hyperviscosity $d \sim -\Delta^6\theta$.

It is worth noting that equation (4) would also be the vorticity form of the incompressible 2D N-S equations

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\nabla p + f(x,t) + \nu\Delta\boldsymbol{u}, \qquad \nabla\cdot\boldsymbol{u} = 0 \tag{5}$$

if d were chosen to be $d = \nu\Delta\theta, \nu > 0$. We would have liked to run these DNS on the 2D N-S equations which will be discussed in Section 4, but the build-up of KE at the smallest scales as a numerical artefact was so strong that we had to either increase ν – which would substantially shorten the inertial range and thus reduce the reliability of our conclusions – or choose a more

scale-selective dissipation. We opted for the latter, as is standard practice in simulations of 2D turbulence (Maltrud and Vallis 1991).

We performed five pairs of identical-twin experiments on equation (4) by varying the random seed that generated the pre-perturbation (original) initial condition. Within each pair, notably, the realisations of the large-scale stochastic forcing $f(t)$ in the control and perturbed runs were identical. The model was pseudo-spectral with a truncation wavenumber of $k_t = 512$, in which the $J(\psi, \theta)$ term was computed in the physical domain via a pair of Fast Fourier Transforms with the spectral de-aliasing filter proposed by Hou and Li (2007). The original initial condition for each of the five cases was an already-developed homogeneous and (approximately) isotropic turbulence with a clean logarithmically corrected -3 spectrum in the inertial range (Figure 3), which has been shown to be a more accurate description of the large-scale-forced 2D turbulent spectrum for finite inertial ranges (Bowman 1996).

The perturbations were introduced spectrally at each of the 2D wavevectors $\boldsymbol{k}$ for a specified value of $k_p = |\boldsymbol{k}|$. A random phase shift independently drawn from a uniform distribution was applied on a pre-determined part $\gamma \in [0,1]$ of the spectral coefficients $\hat{\psi}(\boldsymbol{k})$ and thus $\hat{\theta}(\boldsymbol{k})$, where the hat indicates Fourier coefficients. It can be shown that $\gamma(k_p)$ and $\mathbb{E}(E_e(k_p))$, the expected value of the one-dimensional error KE spectral density at wavenumber k_p, are related by $\mathbb{E}(E_e(k_p)) = 2\gamma^2 E(k_p)$. By specifying the relative error $\frac{\mathbb{E}(E_e(k_p))}{E(k_p)}$, we could work out γ and thus generate the perturbation fields, to which we added the original initial conditions to obtain the perturbed initial conditions.

Since our truncation wavenumber $k_t = 512$ corresponds to $K = 9$ of Lorenz's paper, there would only be 8 values of $t_K - t_{K+1}$, among which only 4 or 5 would be in the inertial range. It would be inaccurate to determine β from such a few data points, so we have transformed Lorenz's argument

to incorporate information from all wavenumbers k, not only from the scale K as a collection of wavenumbers.

To transform the argument, recall that $t_K \sim 2^{-\beta K}$ when $\beta > 0$, and let $T(k)$ be the saturation time of wavenumber k. Since $k \sim 2^K$ and both t_K and T represent saturation times, we may write $T \sim k^{-\beta}$ and conclude that the T should vary with k as a power-law if Lorenz's results hold. The argument will break down when β becomes zero, that is, when the threshold for indefinite predictability is reached.

In this study, the perturbations were introduced at $k_p = 256$. The saturation threshold was chosen to be 1.315 times the KE spectral density of the control flow, or equivalently 0.6575 times the maximum permissible error energy, in accordance with Lorenz (1969). We applied sensitivity tests and found that the results are largely insensitive to the saturation threshold. Figures 4 and 5 show respectively the evolution of the error KE spectrum, and the saturation times T across different wavenumbers k which fit the $T \sim k^{-\beta}$ relationship for $\beta = 0.24$, averaged over the five cases. (The five cases exhibited very similar qualitative behaviour, showing that our results are robust to initial conditions, hence justifying the use of averaging to obtain smoother results.) Based on the transformed version of Lorenz's argument, our result also suggests inherently finite predictability for a (logarithmically corrected) -3 spectrum, this time with greater confidence as β is further away from zero.

4. Aspects from PDE theory: the incompressible 2D Navier-Stokes equations

A very different approach to the problem of inherently finite or indefinite predictability is via use of the analytic theory of partial differential equations (PDEs). The incompressible 2D N-S equations (5), where we shall drop the word 'incompressible' for the remainder of the paper, are always useful as a pedagogical first step towards understanding and modelling the motion of real

fluid flows in the atmosphere. As such, the analytical properties of the 2D N-S problem have been extensively studied. Building on these analytic results, we now consider their implications for predictability.

Well-posedness and implications on predictability

Unlike their three-dimensional counterpart whose regularity problem remains open, the initial-value problem for the 2D N-S equations on the torus (i.e. a doubly periodic domain) has been proven to be well-posed, by which we mean the existence of a unique solution that depends continuously on the initial conditions. Proofs of its well-posedness, for both strong and weak solutions respectively, can be found in the book by Robinson (2001). In the present paper we shall use his proof for weak solutions to demonstrate that the 2D N-S system is indefinitely predictable. To set the context, a summary of the uniqueness proof is provided below. Interested readers may refer to Robinson's book for a full proof.

First, the 2D N-S equations (5) are cast in the form of an ordinary differential equation in an appropriate function space depending on an arbitrary, fixed positive time T. An equation for the error velocity field $w = u - v$ of two solutions u and v is formulated, and its inner product with w itself is taken to obtain an equation for the time-evolution of the error energy $\frac{1}{2}\|w\|^2$, where $\|\cdot\|$ is the L^2 norm on the torus. This equation contains a term which can be bounded above by Ladyzhenskaya's inequalities (Robinson 2001) specific to the 2D N-S equations. After some work one uses Grönwall's inequality to show that

$$\|w(t)\|^2 \leq \exp\left(\int_0^t \frac{M}{\nu}\|\nabla u(s)\|^2 \,\mathrm{d}s\right)\|w(0)\|^2, \qquad t \in [0,T],$$

where M is a positive constant provided by Ladyzhenskaya's inequalities. Uniqueness follows by setting $w(0) = 0$. One can also show the continuous dependence on initial conditions, since

$$\|w(t)\| \leq \sqrt{\exp\left(\int_0^T \frac{M}{\nu}\|\nabla u(s)\|^2 \,\mathrm{d}s\right)}\|w(0)\| =: L(T)\|w(0)\|, \qquad t \in [0,T], \tag{6}$$

i.e. errors are Lipschitz in time.

As an immediate corollary to inequality (6), the 2D N-S system is indefinitely predictable (Palmer et al. 2014). Indeed, if a prediction is defined to lose its skill when $\|w(t)\| > \varepsilon$, then for any given time $T \in \mathbb{R}^+$, the prediction is skilful for at least up to T when the initial error $\|w(0)\|$ can be made sufficiently small, that is, smaller than $\frac{\varepsilon}{L(T)}$.

It is important to note that in the present Section the KE spectral slope plays no role in determining the inherent finiteness or indefiniteness of predictability of the 2D N-S equations. The above argument applies to 2D N-S systems of any spectral slope.

5. Reconciling the contradiction with Lorenz

At first glance, the indefinite predictability derived in Section 4 seems to contradict Lorenz's result in Section 2 for any $p < 3$. However, we have not discussed the role of the KE spectral slope in $L(T)$ which, as we will see in the following, reconciles the conflict.

Central to our argument is the inequality (6) presented above. For simplicity, suppose the real system has only one inertial range of slope $-p$ (without any logarithmic correction) in its KE spectrum so that $|\hat{u}(k)|^2 \sim k^{-p}$ (note the change of notation: the hat now represents Fourier coefficients in the space of one-dimensional wavenumbers k) between its large-scale cutoff wavenumber k_1 and small-scale cutoff wavenumber k_2. Then

$$\|\nabla u_s\|^2 = \int_0^\infty k^2|\hat{u}_s|^2 \,\mathrm{d}k = \int_0^{k_1} k^2|\hat{u}_s|^2 \,\mathrm{d}k + A_0\int_{k_1}^{k_2} k^{2-p} \,\mathrm{d}k + \int_{k_2}^\infty k^2|\hat{u}_s|^2 \,\mathrm{d}k \qquad (A_0 \text{ constant}), \tag{7}$$

where the subscript s distinguishes the system itself from a model for the system which we will denote with subscript m. The three terms on the right-hand-side of equation (7) represent contributions from the large scale, the inertial range and the viscous range respectively. Compared to the first two terms, the term representing the viscous range is assumed to be small. In particular, the integrand is assumed to decay rapidly enough so that $\|\nabla u_s\|^2$ remains finite (this is in fact part of the definition of the function space to which u_s belongs).

Now, suppose the model truncates at wavenumber $k_t \ll k_2$ and numerical dissipation kicks in at wavenumber $k_0 \in (k_1, k_t)$. For the model,

$$\|\nabla u_m\|^2 = \int_0^{k_t} k^2 |\hat{u_m}|^2 \, \mathrm{d}k = \int_0^{k_1} k^2 |\hat{u_m}|^2 \, \mathrm{d}k + A_0 \int_{k_1}^{k_0} k^{2-p} \, \mathrm{d}k + \int_{k_0}^{k_t} k^2 |\hat{u_m}|^2 \, \mathrm{d}k. \tag{8}$$

Again, we may neglect the contribution from the viscous range, so that

$$\|\nabla u_m\|^2 \sim \int_0^{k_1} k^2 |\hat{u_m}|^2 \, \mathrm{d}k + A_0 \int_{k_1}^{k_0} k^{2-p} \, \mathrm{d}k. \tag{9}$$

Because $k_0, k_t \ll k_2$, the second integral in relation (9) with $p < 3$ *appears* to diverge as the resolution (k_0, k_t) increases. Combining this with inequality (6), $L(T)$ – until k_2 is reached – grows exponentially with k_0, leading to a breakdown of the Lipschitz-continuous dependence on initial conditions in inequality (6). To keep the error $\|w(t)\|$ under control, the initial error $\|w(0)\|$ would have to decrease exponentially, but decreasing the scale of the initial error without changing its magnitude relative to the background KE spectral density (Lorenz's thought experiment) would only give a polynomial decrease. The corollary of indefinite predictability therefore fails to hold. Hence the range of predictability is inherently finite in practice, even though the system is indefinitely predictable, because indefinite predictability cannot be achieved without making the model resolution so high that its effective resolution k_0 (and the scale of the initial error) falls within the viscous range of the real system, let alone the large-scale error has to be constrained to zero (Durran and Gingrich 2014).

This concept, known as 'asymptotic ill-posedness', was put forward by Palmer et al. (2014) as they argued that whether the three-dimensional Navier-Stokes system is well-posed is practically irrelevant to the well-established theory of inherently finite predictability. We have now extended the discussion to the 2D N-S system and given a mathematical basis to the concept in our context.

When $p > 3$, the second integral in relation (9) does not appear to diverge as $k_0 \to k_2$. This means one may indeed approximate $\|\nabla u_s\|^2$ by the $\|\nabla u_m\|^2$ in relation (9) with a sufficiently large value of k_0. So would $L(T)$ of inequality (6) be approximated without regard to the model resolution, making it possible for $\|w(t)\| \leq \varepsilon$ by making $\|w(0)\|$ small enough in scale and thus achieving indefinite predictability.

So far our argument for the cases $p < 3$ and $p > 3$ are in harmony with Lorenz's result in Section 2. For the borderline case $p = 3$, our argument suggests practically inherently finite predictability, since $\|\nabla u_m\|^2 \sim \text{constant} + \int_{k_1}^{k_0} k^{-1}\, \mathrm{d}k = \text{constant} + \log \frac{k_0}{k_1}$ which appears to diverge as $k_0 \to k_2$. This disagrees with Lorenz. Even with the logarithmic correction

$$|\hat{u}(k)|^2 \sim k^{-3} \left[\log \left(\frac{k}{k_r} \right) \right]^{-\frac{1}{3}} \qquad (k_r > 0 \text{ constant}),$$

or more generally

$$|\hat{u}(k)|^2 \sim k^{-3} \left[A_1 \log \left(\frac{k}{k_r} \right) + A_2 \right]^{-\frac{1}{3}} \qquad (A_1, A_2, k_r > 0 \text{ constants}),$$

to the -3 spectrum (Bowman 1996), an easy calculation along the previous lines still suggests that the range of predictability is practically inherently finite. As such, we are unable to explain the disagreement and we leave the problem open.

For models and systems with multiple inertial ranges, only the range immediately before the viscous range pertains to our argument concerning the large-k_0 behaviour. This applies to the real atmosphere where $p = \frac{5}{3}$ (Nastrom and Gage 1985). Since k_t for atmospheric models is smaller than k_2 by 'at least seven or eight orders of magnitude' (Palmer et al. 2014), the crucial assumption

to our discussion ($k_t \ll k_2$) is satisfied and we conclude that atmospheric predictability is indeed practically inherently finite.

6. Conclusions

Half a century on since Lorenz's pioneering result of inherently finite atmospheric predictability, we revisited his original argument by (i) re-running his simplified model of the 2D barotropic vorticity equation, (ii) directly simulating the full model and (iii) comparing his conclusions with the well-posedness of the 2D N-S equations as proven by Ladyzhenskaya.

Although his main result – that atmospheric predictability is inherently finite because the KE spectral slope is shallower than -3 – has now become an 'accepted part of the canon of dynamical meteorology' (Rotunno and Snyder 2008), the details behind the conclusion were re-assessed. For the -3 spectrum, we saw a substantially different β in the DNS (Section 3(b)) than in the simplified model (Section 3(a)), which may be an indication that the model is inadequate in simulating error growth. In both cases, nevertheless, the hypothesis of indefinite predictability ($\beta = 0$) for $p = 3$ based on linear extrapolation (Section 2) was refuted.

The 2D N-S equations that closely relate to the 2D barotropic vorticity equation were used to address the predictability problem from a more rigorous perspective. The forced-dissipative system was shown to be indefinitely predictable regardless of the spectral slope (Section 4). However, we found that $p = 3$ serves as a cutoff between *practically* inherently finite and indefinite predictability by noting how quickly the initial error has to be brought down with increasing resolution in order to maintain the bound for the error at future times (Section 5). This echoes Lorenz's original conclusions except for the borderline case $p = 3$ itself, in which case our result of inherently finite predictability agrees with our own computations of Lorenz's model and the DNS.

Until recently, KE spectra in global weather forecast models had only resolved the synoptic-scale -3 range. As model resolutions start to extend into the $-\frac{5}{3}$ range, the strong constraints on the range of predictability envisaged by Lorenz will become visible (Judt 2018). However, the limits on predictability arising from initial errors on the large scales will also limit predictability in practice (Durran and Gingrich 2014), and the interplay between the two could be an interesting area to explore.

By providing another approach to attacking the problem of predictability (via the analytic theory of the 2D N-S equations), we look forward to similar results on more atmospherically relevant PDEs such as the surface quasi-geostrophic equations (Held et al. 1995), and a more active contribution from mathematicians on this topic.

Acknowledgments. Tsz Yan Leung is supported through a PhD scholarship awarded by the Engineering and Physical Sciences Research Council grant EP/L016613/1 'EPSRC Centre for Doctoral Training in the Mathematics of Planet Earth at Imperial College London and the University of Reading', with additional funding support from the European Research Council Advanced Grant 'Understanding the Atmospheric Circulation Response to Climate Change' (ACRCC), Project 339390, under Theodore G. Shepherd as the Principal Investigator. The authors thank Richard Scott for providing his code for modification for the numerical simulation in Section 3(b). The authors also wish to thank the three anonymous reviewers for their helpful and valuable comments on an earlier version of this manuscript.

APPENDIX A

Motivating the concept *range of predictability*

Standard theory in deterministic dynamical systems dictates that, for a perfect model of the system, the dynamics of the error ε can be completely described by the time t, the initial error ε_0 and the initial state of the system x, so that $\varepsilon = \varepsilon(t, \varepsilon_0, x)$. Suppose now that the skill S of a forecast, particular to an initial state, can be quantified by a continuously decreasing function of some norm $\|\cdot\|$ (such as the kinetic energy) of the error. In such a case, we can write $S = S(\|\varepsilon\|) = S(\varepsilon) = S(t, \varepsilon_0, x)$. Now, averaging over the initial states on some (non-trivial) attractor D of the system, we may define the overall skill $\overline{S} = \int_{x \in D} S(t, \varepsilon_0, x)\, \mathrm{d}x$ of a forecast. If we further assume that the error norm increases with t in the average, as is in the context of atmospheric predictions, we can infer that $\overline{S}(t, \varepsilon_0)$ is monotonically decreasing in time.

Let's say that a prediction loses its skill when $\overline{S} < \alpha$ (where α is a fixed threshold), which is typically realised in fluid flows by saturation of the error kinetic energy spectrum at specified scales. Perhaps a first question to the understanding of predictability can be formulated as follows: how long does it take for an initial error ε_0 to grow so that a prediction is no longer skilful? The answer $\tilde{T}$, known as the range of predictability, is the solution to $\overline{S}(t, \varepsilon_0) = \alpha$ for the specified ε_0. The monotonicity assumption of $\overline{S}$ guarantees the uniqueness of the solution $\tilde{T}$.

By formulating this question for different initial error fields we can regard $\tilde{T}$ as a function of ε_0. It is clear from the very definition of deterministic systems that $\varepsilon_0 = 0$ implies $\tilde{T}(\varepsilon_0) = \infty$. However, it is not quite obvious as to whether $\tilde{T}$ could be made arbitrarily large by reducing $\|\varepsilon_0\|$ to anything positive below a threshold, or equivalently whether the equality $\liminf_{\|\varepsilon_0\| \to 0} \tilde{T}(\varepsilon_0) = \infty$ holds, because $\tilde{T}$ may behave irregularly at small $\|\varepsilon_0\|$ – or at least appear to.

To see the equivalence, we unwrap the statement $\liminf_{\|\varepsilon_0\|\to 0} \tilde{T}(\varepsilon_0) = \infty$ to get

$$
\begin{aligned}
& \liminf_{\|\varepsilon_0\|\to 0} \tilde{T}(\varepsilon_0) = \infty \\
\Leftrightarrow \quad & \forall R \in \mathbb{R},\ \liminf_{\|\varepsilon_0\|\to 0} \tilde{T}(\varepsilon_0) \geq R \\
\Leftrightarrow \quad & \forall R \in \mathbb{R},\ \sup_{\varepsilon'>0} \inf_{\|\varepsilon_0\|\in(0,\varepsilon')} \tilde{T}(\varepsilon_0) \geq R \\
\Leftrightarrow \quad & \forall R \in \mathbb{R},\ \exists\, \varepsilon' > 0 \text{ such that } \inf_{\|\varepsilon_0\|\in(0,\varepsilon')} \tilde{T}(\varepsilon_0) \geq R \\
\Leftrightarrow \quad & \tilde{T} \text{ could be made arbitrarily large by reducing } \|\varepsilon_0\| \text{ to anything positive below a threshold.}
\end{aligned}
$$

With this in mind, a system is said to have an *indefinite range of predictability*, or be *indefinitely predictable*, if the range of predictability could be made arbitrarily large by reducing the initial error to a small yet positive value. Systems that fail to satisfy such a condition are referred to as *inherently finitely predictable*.

APPENDIX B

Some details regarding the computation of the matrix C

The integrations were performed using `scipy.integrate.nquad` on Python which returned a warning message '`IntegrationWarning: Extremely bad integrand behavior occurs at some points of the integration interval`' about the integrand's singular behaviour, even if the integration domain were confined to the support of the integrand so that resources were not wasted in integrating zero regions. The warning disappeared by applying a change of coordinates (from logarithmic to Cartesian) in the integrand and accordingly the integration limits, which sped up the wall-clock time of the computation by a factor of about 9 as well. The entries of C computed by these two methods differ by no more than 0.0025%. Based on these observations, we are confident that our computations are accurate.

The numerical code for these computations is available in the Supplemental Material.

APPENDIX C

Solution procedure of Lorenz's system

1. Set a time-step h; in this case, $h = 0.001$.

2. Initialise the run by setting time $t = 0$. Also initialise $t_0 = 0$. (t_0 is the time when the previous saturation occurs.)

3. Project the initial condition onto the eigenspace of the block matrix

$$\begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}$$

to determine the constants of the general solution.

4. Compute the solution at time $t - t_0$ and check if any of the scales K saturates by time t.

5. If none of the scales saturates, reset $t = t + h$ and repeat step 4.

6. If, by time t, the error energy at some scale exceeds the background energy at the same scale, then the clock (i.e. the variable t) is reset to the previous time-step $t - h$, and h is refined to 0.000001.

7. Repeat steps 4-5 with the new value of h until the error energy at some scale K exceeds the background energy at the same scale. The saturation time of scale K is determined as if the error energy increases linearly between times $t - h$ and t.

8. Reset $h = 0.001$ and set t_0 to be the current time t. Remove the row and column of the matrix C corresponding to scale K and the corresponding entries of Z, W and all the F_j terms (more

on the F_j terms below). The reduced-size system

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}Z = CZ + \sum_j F_j$$

(where the summation is performed over all saturated scales) has a new inhomogeneous term F_K which accounts for the contribution of scale K's saturated energy to the error growth at the remaining scales. (F_K equals the background energy of scale K multiplied by the removed column of C restricted to the rows corresponding to the remaining scales.) As the new system is equivalent to the first-order system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} Z \\ W \end{pmatrix} = \begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}\begin{pmatrix} Z \\ W \end{pmatrix} + \begin{pmatrix} 0 \\ \sum_j F_j \end{pmatrix}$$

(with the size of the identity matrix I accordingly reduced) and $\sum_j F_j$ is a constant vector, its solution $(Z,W)^T$ can be expressed as the sum of a particular solution $(\Delta Z, \Delta W)^T$ and a solution of of the homogeneous system $(\overline{Z},\overline{W})^T$. A particular solution to the differential equation can be obtained by solving

$$-\begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}\begin{pmatrix} \Delta Z \\ \Delta W \end{pmatrix} = \begin{pmatrix} 0 \\ \sum_j F_j \end{pmatrix}.$$

Re-calculate the eigenvalues and eigenvectors of the new

$$\begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}$$

and project the difference between the current solution at the unsaturated scales (treated as the initial condition in the variable $t - t_0$) and $(\Delta Z, \Delta W)^T$ onto such an eigenspace to determine the constants of $(\overline{Z},\overline{W})^T$ and thus the full solution.

9. Repeat steps 4–8 until all scales saturate.

The numerical code for the solution procedure is available in the Supplemental Material.

References

Bowman, J. C., 1996: On inertial-range scaling laws. *J. Fluid Mech.*, **306**, 167–181, doi:10.1017/S0022112096001279.

Buizza, R., and M. Leutbecher, 2015: The forecast skill horizon. *Quart. J. Roy. Meteor. Soc.*, **141**, 3366–3382, doi:10.1002/qj.2619.

Durran, D. R., and M. Gingrich, 2014: Atmospheric predictability: why butterflies are not of practical importance. *J. Atmos. Sci.*, **71**, 2476–2488, doi:10.1175/JAS-D-14-0007.1.

Held, I. M., R. T. Pierrehumbert, S. T. Garner, and K. L. Swanson, 1995: Surface quasi-geostrophic dynamics. *J. Fluid Mech.*, **282**, 1–20, doi:10.1017/S0022112095000012.

Hou, T. Y., and R. Li, 2007: Computing nearly singular solutions using pseudo-spectral methods. *Journal of Computational Physics*, **226**, 379–397, doi:10.1016/j.jcp.2007.04.014.

Judt, F., 2018: Insights into atmospheric predictability through global convection-permitting model simulations. *J. Atmos. Sci.*, **75**, 1477–1497, doi:10.1175/JAS-D-17-0343.1.

Lilly, D. K., 1990: Numerical prediction of thunderstorms – has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779–798, doi:10.1002/qj.49711649402.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, doi:10.3402/tellusa.v21i3.10086.

Maltrud, M. E., and G. K. Vallis, 1991: Energy spectra and coherent structures in forced two-dimensional and beta-plane turbulence. *J. Fluid Mech.*, **228**, 321–342, doi:10.10.1017/S0022112091002720.

Nastrom, G. D., and K. S. Gage, 1985: A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.*, **42**, 950–960, doi:10.1175/1520-0469(1985)042⟨0950:ACOAWS⟩2.0.CO;2.

Palmer, T. N., A. Döring, and G. Seregin, 2014: The real butterfly effect. *Nonlinearity*, **27**, R123–R141, doi:10.1088/0951-7715/27/9/R123.

Robinson, J. C., 2001: *Infinite-Dimensional Dynamical Systems: An Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors*. Cambridge University Press, 234–257 pp.

Rotunno, R., and C. Snyder, 2008: A generalization of Lorenz's model for the predictability of flows with many scales of motion. *J. Atmos. Sci.*, **65**, 1063–1076, doi:10.1175/2007JAS2449.1.

Selz, T., 2019: Estimating the intrinsic limit of predictability using a stochastic convection scheme. *J. Atmos. Sci.*, **76**, 757–765, doi:10.1175/JAS-D-17-0373.1.

Sun, Y. Q., and F. Zhang, 2016: Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *J. Atmos. Sci.*, **73**, 1419–1438, doi:10.1175/JAS-D-15-0142.1.

Vallis, G. K., 1985: Remarks on the predictability properties of two- and three-dimensional flow. *Quart. J. Roy. Meteor. Soc.*, **111**, 1039–1047, doi:10.1002/qj.49711146608.

Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077–1091, doi:10.1175/JAS-D-18-0269.1.

LIST OF FIGURES

Fig. 1. The evolution of the error KE spectrum (red, from bottom to top) in the Lorenz (1969) model, with a -3-slope background KE spectrum (appearing in this Figure as a -2 slope, as the KE per unit logarithmic wavenumber is plotted instead of the KE spectral density) as in Rotunno and Snyder (2008). The initial condition for this run is $Z(K=20)=2^{-40}$, $Z=0$ for all other K, and $W=0$ for all K. 25

Fig. 2. Saturation times of various scales (red) for the same model run as in Figure 1. The blue curve shows a line of fit with $\beta=0.05$. 26

Fig. 3. KE spectrum (averaged over the five cases) of the initial condition (red), and logarithmically corrected -3 reference spectra $E(k)\sim k^{-3}\left[\log\left(\frac{k}{k_r}\right)\right]^{-\frac{1}{3}}$ ($k_r=10$ in black, $k_r=20$ in green), where $E(\cdot)$ is the one-dimensional KE spectral density. 27

Fig. 4. Evolution of the error KE spectrum (magenta and blue, bottom to top) for an initial perturbation (blue dot) at $k_p=256$. The magenta curves are for $t=0.3,0.6,\ldots,2.7$ and the blue curves are for $t=3,6,\ldots,66$. The background KE spectra at $t=0,3,6,\ldots,66$, scaled by a factor of 2, are shown in red (top to bottom), with the reference spectra in black and green as in Figure 3. The spectra are averaged over the five cases. 28

Fig. 5. Saturation times T at different wavenumbers k (red) for an initial error at wavenumber $k_0=256$, averaged over the five cases. The black curve shows a line of fit with $\beta=0.24$. 29

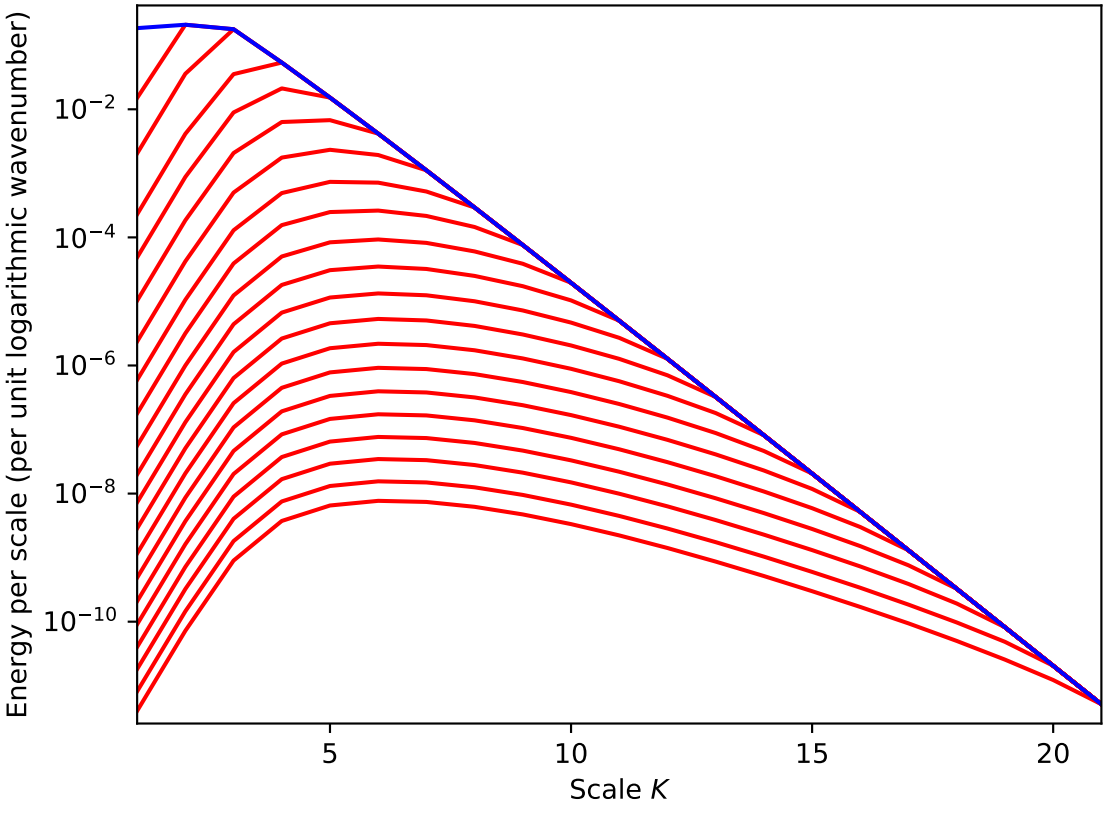


FIG. 1. The evolution of the error KE spectrum (red, from bottom to top) in the Lorenz (1969) model, with a -3-slope background KE spectrum (appearing in this Figure as a -2 slope, as the KE per unit logarithmic wavenumber is plotted instead of the KE spectral density) as in Rotunno and Snyder (2008). The initial condition for this run is $Z(K=20) = 2^{-40}$, $Z = 0$ for all other K, and $W = 0$ for all K.

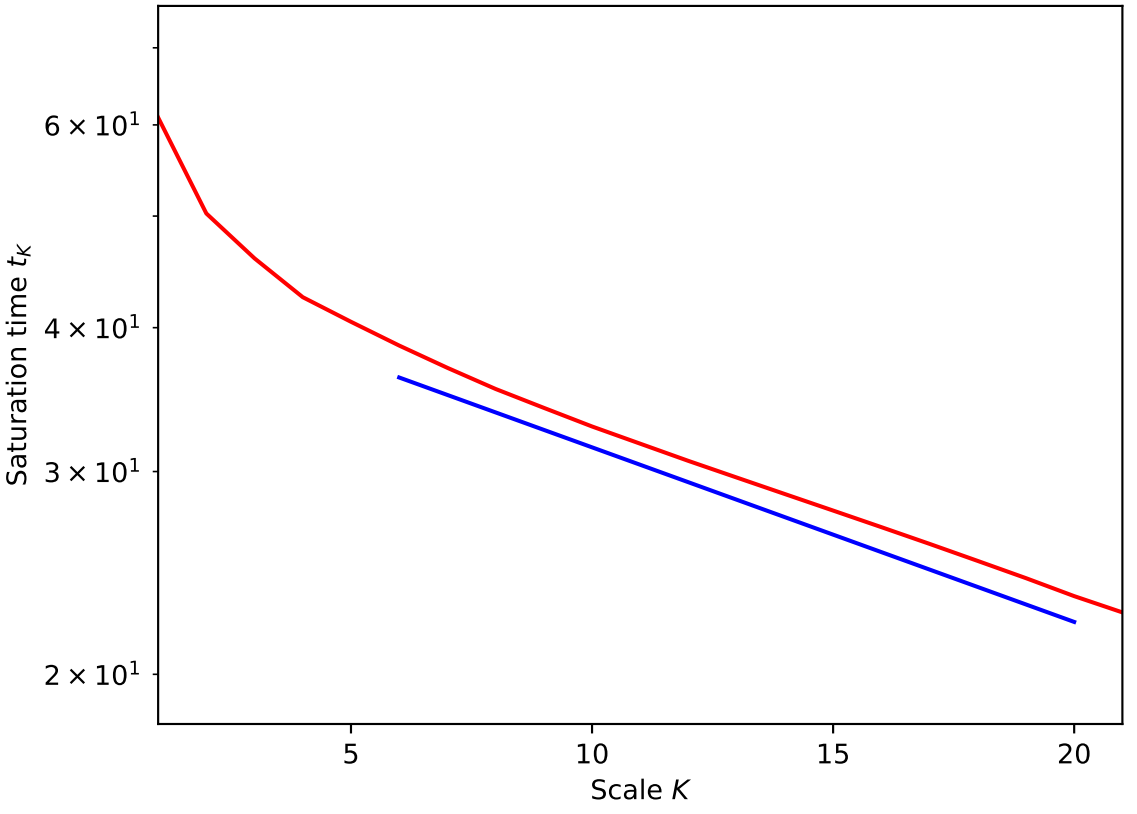
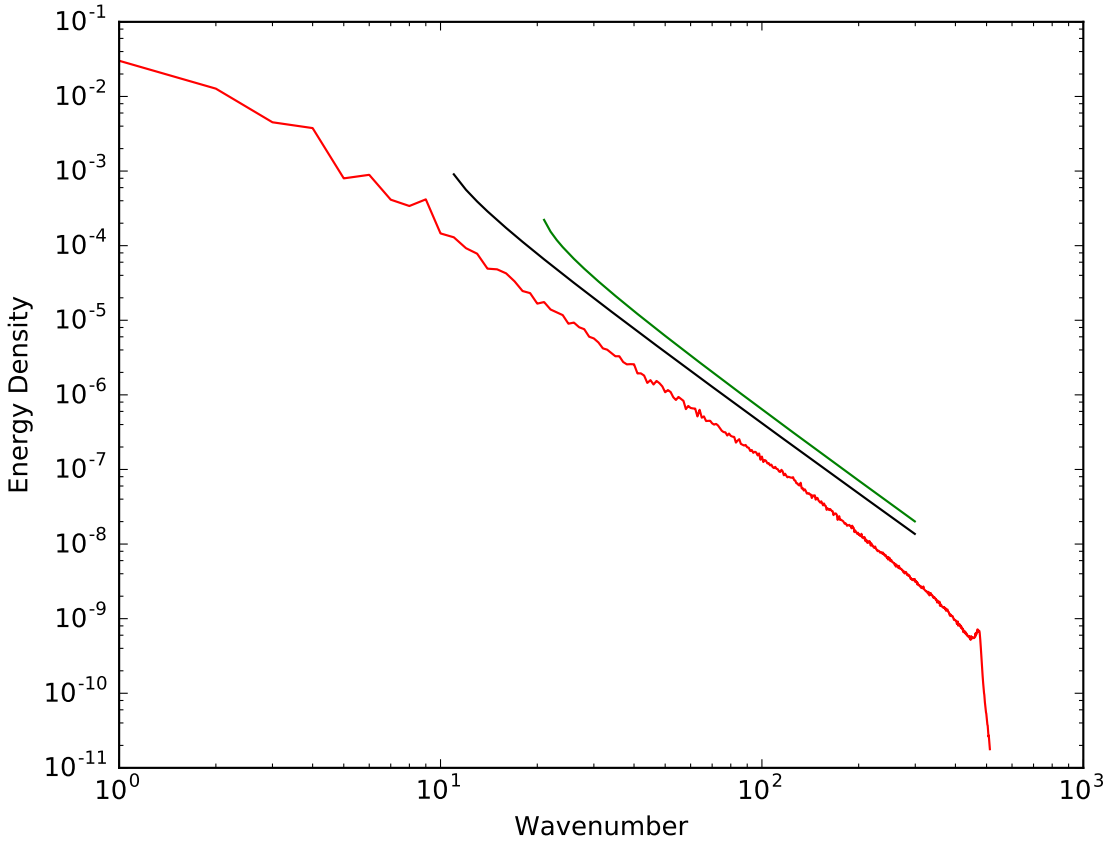


FIG. 2. Saturation times of various scales (red) for the same model run as in Figure 1. The blue curve shows a line of fit with $\beta = 0.05$.



480 FIG. 3. KE spectrum (averaged over the five cases) of the initial condition (red), and logarithmically corrected
481 -3 reference spectra $E(k) \sim k^{-3}\left[\log\left(\frac{k}{k_r}\right)\right]^{-\frac{1}{3}}$ ($k_r = 10$ in black, $k_r = 20$ in green), where $E(\cdot)$ is the one-
482 dimensional KE spectral density.

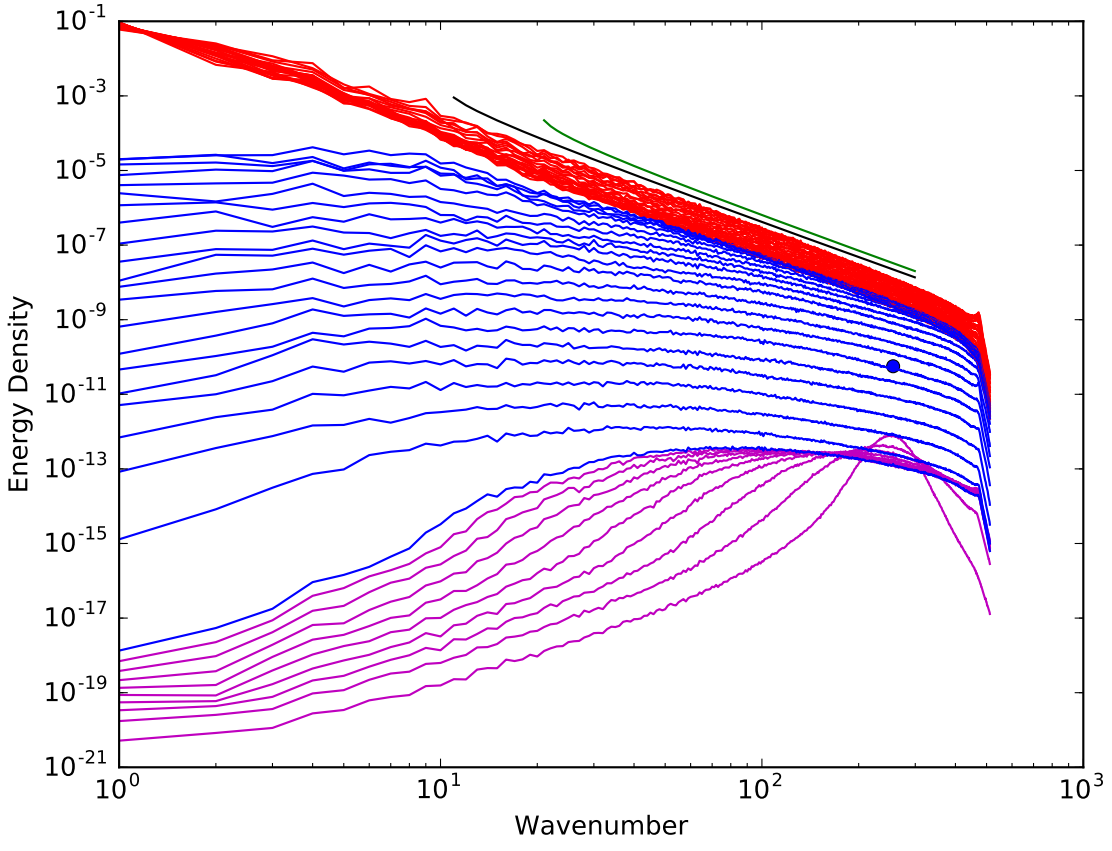


FIG. 4. Evolution of the error KE spectrum (magenta and blue, bottom to top) for an initial perturbation (blue dot) at $k_p = 256$. The magenta curves are for $t = 0.3, 0.6, \ldots, 2.7$ and the blue curves are for $t = 3, 6, \ldots, 66$. The background KE spectra at $t = 0, 3, 6, \ldots, 66$, scaled by a factor of 2, are shown in red (top to bottom), with the reference spectra in black and green as in Figure 3. The spectra are averaged over the five cases.

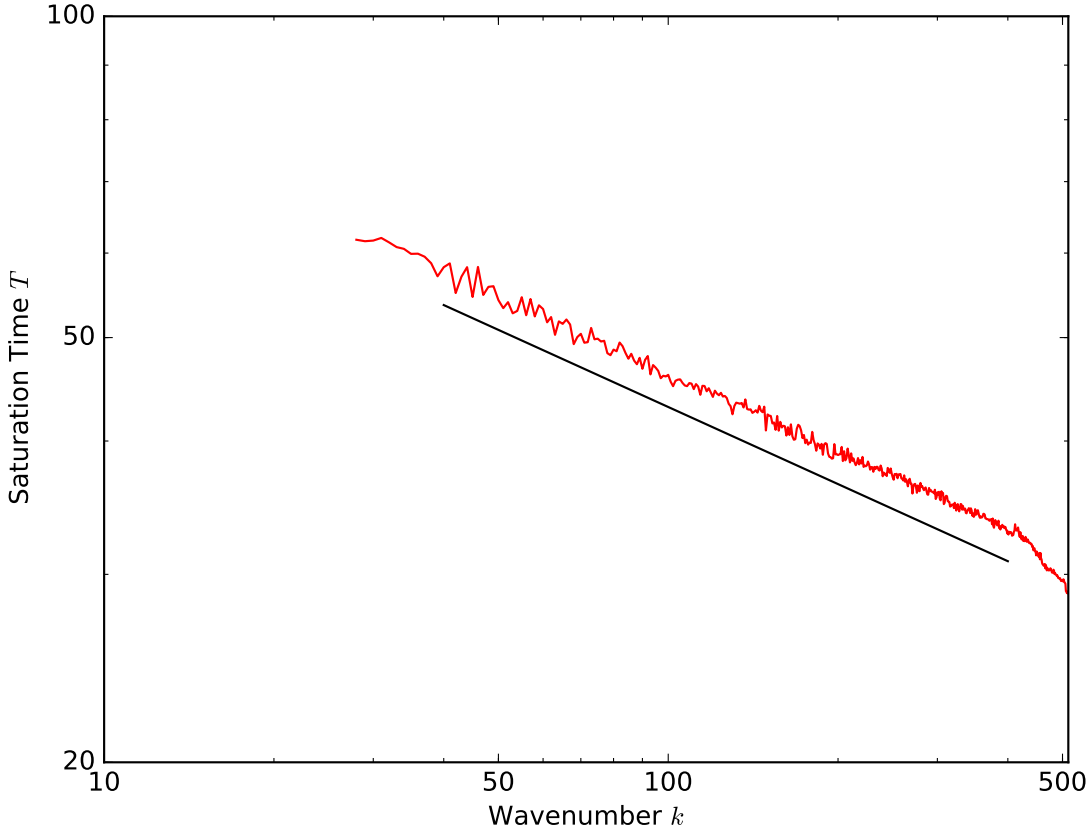


FIG. 5. Saturation times T at different wavenumbers k (red) for an initial error at wavenumber $k_0 = 256$, averaged over the five cases. The black curve shows a line of fit with $\beta = 0.24$.