Revisiting Lorenz's and Lilly's Empirical Formulas for Predictability Estimates

by

Bo-Wen Shen[1]*, Roger A. Pielke Sr.[2], and Xubin Zeng[3]

[1]Department of Mathematics and Statistics, San Diego State University, San Diego, California

[2]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado

[3]Department of Hydrology and Atmospheric Science, The University of Arizona, Tucson, Arizona

*bshen@sdsu.edu

(Submitted: July 2024)

(Revised: January 2025)

Abstract

Recent studies have reiterated that the two-week predictability limit was originally estimated using a doubling time of five days from the Mintz-Arakawa model in the 1960s. However, this two-week predictability limit has conventionally been viewed as one of Lorenz's major findings from his 1969 studies. The limit has been presumably attributed to the mechanism involving the insignificant contributions of unresolved scales smaller than 38 meters. To understand the discrepancies in the origin of the two-week limit and to validate the mechanism in addressing the dependence of finite predictability on the atmospheric spectrum, we revisit Lorenz's studies, Lilly's work, and related research from the 1960s and early 1970s.

We first review how Lilly applied turnover time in turbulence theory to construct a convergent series that appears mathematically similar to the original Lorenz series. We then reexamine how Lorenz observed regularity in a sequence of saturation times over 21 selected wave modes and used the regularity to construct a convergent series, illustrating the negligible contribution of unresolved small-scale processes to predictability enhancement.

Our reanalysis does not support the claim that Lorenz’s and Lilly’s formulas are mathematically identical or physically comparable. Major discrepancies and inconsistencies include the use of different physical time scales in Lorenz's and Lilly's studies and the lack of a common factor of $2^{-2/3}$ that can be robustly determined from Lorenz's data. This falsifies the assumption that saturation time difference and turnover time are linearly proportional over the selected wave modes. Additionally, given the $-5/3$ power spectrum, we demonstrate that the convergence properties of Lorenz's or Lilly's series depend on spectral discretization. These issues, along with the highly simplified features of the Lorenz 1969 model, indicate that an upper bound for the predictability limit has not been robustly determined in Lorenz's and Lilly's studies. This view is consistent with Lorenz’s updated review in the 1990s and 2000s. Therefore, caution should be exercised when applying Lilly's formula to conclude the dependence of finite predictability on the slopes of spectra. This perspective suggests opportunities to explore larger predictability and extend weather forecasts using various approaches, including sophisticated theoretical, real-world, and artificial intelligence-powered models.

1. Introduction

Recent advancements in numerical models, data assimilation systems, and numerical approaches (e.g., the application of ensemble runs) have demonstrated promising, long-range simulations surpassing the previously considered two-week predictability limit. These advancements have been achieved using physics-based models (e.g., Shen et al. 2010, 2011; Buizza and Leutbecher 2015; Bretherton and Khairoutdinov 2015; Judt 2018, 2020), hybrid dynamical and artificial intelligence (AI) systems (e.g., Bach et al. 2024), and AI-based systems (e.g., Li et al. 2024; Lang et al. 2024; See a review in Shen et al. 2024).

Such a result leads to an intriguing question regarding the apparent gap between current modeling capabilities and the theoretical predictability limit. To address this question, several review articles regarding Lorenz's predictability studies have recently been completed by the authors. For example, Shen et al. (2023a, 2024) presented Lorenz's perspective on predictability limits, while Shen et al. (2023b) and Shen (2023) examined the major features of Lorenz's models spanning 1960 to 2008. These studies reaffirmed that a predictability limit of two weeks was indeed established, based on a doubling time of 5 days obtained using the Mintz-Arakawa model (e.g., Charney et al. 1966; GARP 1969; Lorenz 1969a, b, c, d, e; Lewis 2005). Two-week predictability was later attributed to the findings of Lorenz's studies during the 1960s, likely due to the following: (1) Lorenz's chaotic or multiscale models demonstrated a qualitatively finite predictability (Lorenz 1963, 1969d); (2) Lorenz's study in 1969 reported an estimated predictability limit of 16.8 days at the largest wavelength (Lorenz 1969d); and (3) In the early 1970s, surprising similarities were determined between Lorenz's saturation-time-based formula (Lorenz 1969d) and Lilly's turnover-time based formula (Lilly 1972, 1973), which was applied in order to project the impact of unresolved small-scale processes on predictability. To address the above question, Shen et al. (2021a, b, 2022a, 2023a, b, 2024) previously reexamined the validity of the first two points by presenting new insights of Lorenz's models, including the linear feature of the Lorenz 1969 model. This study specifically revisits Lorenz's and Lilly's formulas for the validity of the third point outlined above.

During the 1960s, Lorenz conducted a series of studies that documented the existence of non-periodic solutions and examined their dependence on initial conditions. These findings were fundamental in establishing chaos theory and in defining the objective of determining predictability limits (Lorenz 1962, 1963, 1965, 1969a, b, c, d, e; 1993). A review of Lorenz's studies by Shen et al. (2023a, b, 2024) indicated that Lorenz and others attempted to quantitively estimate predictability limits using results obtained from three approaches (including general circulation models, natural analogues, and a theoretical model in five of Lorenz's 1969 studies). The well-cited study Lorenz (1969d) applied the concept of saturation time to report a predictability of 16.8 days at the largest wavelength, close to the estimated well-acknowledged predictability limit estimated using a doubling time of 5 days by Charney et al. (1966). However, it remains unclear regarding the relationship between the saturation time and doubling time. In particular, Lorenz obtained a doubling time of ~2-3 days, different from the doubling time of 5 days.

Lorenz (1969d) is, indeed, the most well-known study amongst his five studies published during 1969. The Lorenz (1969d) study has often been cited together with Lorenz (1963), in particular, in meteorology. In Lorenz (1969d), the Lorenz 1969 (L69) model, consisting of a system of 21, 2nd-order ordinary differential equations (ODEs), was proposed for estimating predictability at multiple scales. As reanalyzed by Shen et al. (2022a), the L69 model is a closure-based, physically multiscale, mathematically linear, and numerically ill-conditioned system. The overlooked feature that Lorenz 1969 model is linear was indeed mentioned in Lorenz's own study (e.g., Lorenz 1969d; 1984). Such a feature is acknowledged in Sakai and Yorke (2023) and by Prof. Tim Palmer (personal communication, May 2024).

One unique feature of the L69 model is the application of a non-uniform spectral discretization that allows illustrations of predictability over a wide range of wavelengths, from 38 m to 40,000 km. As defined in Eq. (1a), $k_j = 2^{j-1}k_L,$ the wavenumber k is an exponential function of integer j, where $k_L$represents the wavenumber for the largest scale. In fact, the feature of the non-uniform grid and its strength and weakness has also been overlooked. Numerical results obtained from the L69 model yield a sequence of *saturation times* for estimating predictability at different wavelengths (i.e., spatial scales). As shown in the excerpts in Figure 1, from Lorenz (1969d), such a sequence (i.e., saturation time at a specific scale) is associated with energy growth at different scales.

By observing *regularity* in the finite sequence of *saturation time differences*, Lorenz suggested an empirical, convergent series for estimating the predictability limit at the system scale. In this study, such a series is referred to as Lorenz's formula (or Lorenz's predictability series). Based on the series, which is a geometric series and can be extended to become an infinite series, Lorenz made it feasible to examine the impact of unresolved scales on the contribution of predictability. Compared to coarse-resolution general circulation models from the 1960s, which rely on parameterizations due to their lack of fine resolution, the uniqueness of Lorenz's (1969d) study rests in its utilization of a non-uniform grid and the identification of regularity within the sequence of saturation times. However, some issues are reported below.

Later, during the early 1970s, Lilly applied a turbulence theory-based turnover time in order to formulate the energy transfer described in Figure 1. Lilly (1972) discovered similar regularity in the sequence of turnover times at non-uniform grids under the same spectral discretization as that in Lorenz (1969d). Based on turnover times over an infinite set of wavenumbers, a series was formulated for estimating predictability. The turbulence-theory-based formula is referred to as Lilly's formula (or Lilly's predictability series). Lilly's work linked the two concepts of turnover time and saturation time differences although he did not provide physical justifications. As defined in Section 3.1, the concepts of turnover time and saturation time are physically different.

By considering turnover times for various atmospheric spectra over the non-uniform grid, Lilly's formula is effective for illustrating the dependence of "finite" or "infinite" predictability on the slopes of the kinetic energy spectrum as well as the dimensionality of turbulence (i.e., two vs. three dimensions). Lilly's studies produced results that appear consistent with the findings of Lorenz, and, thus, Lorenz's and Lilly's formulas have been jointly applied to explain the predictability limit of two weeks (e.g., Lorenz 1969d; Lilly 1972, 1973, 1990). However, such an approach that also appeared in follow-up predictability studies (e.g., Palmer et al. 2014; Durran and Gingrich 2014) will be reexamined in this study.

From a closer study of Table 3 we can infer what the result would have been if much smaller scales of motion had been included. Except for the smallest scales retained, where the effect of omitting still smaller scales is noticeable, and the very largest scales, where X_k does not conform to a $-\frac{5}{3}$ law, successive differences $t_k - t_{k+1}$ differ by a factor of about $2^{-\frac{2}{3}}$. If one chooses to reevaluate t_1 by summing the terms of the sequence $t_1 - t_2$, $t_2 - t_3$, ..., one is effectively summing a truncated geometric series. If n had been chosen larger, the series would simply contain additional terms. Even with $n = \infty$, this series would converge to a value only about 2 minutes greater than its value for $n = 20$. It thus appears that with an arbitrarily small initial error, confined to an arbitrarily small scale, the range of predictability of the present model is still about 16.8 days. If we can trust the various assumptions used in deriving and solving the equations, we must conclude that the system falls in the third category previously enumerated, and possesses an intrinsic finite range of predictability.

Figure 1: An excerpt from page 14 of Lorenz (1969d). Note that symbols “k” and “n" in the above Figure are, indeed, “n” and N, respectively, in this study.

While today’s atmospheric predictability studies are usually based on the growth of root-mean-sq~~a~~uare error and/or anomaly correlation coefficients (ACC), the above studies and associated concepts (e.g., scale-dependent error-doubling time and turnover time, energy cascade, predictability limit) remain foundational in our understanding of atmospheric predictability and predictions. Using a real-world model, recent results by Lloveras, Tierney, and Durran (2022, hereafter referred to as LTD22) reported discrepancies in the fundamental concepts applied to derive Lorenz’s and Lilly’s formulas. We applaud the results of LTD22. In this study, we further provide a mathematical analysis in order to show discrepancies in Lorenz’s and Lilly’s formulas and, thus, illustrate unrealistic features produced by the two formulas. Assuming that discrepancies in the two formulas could be ignored, we apply Lilly's formula in order to reveal the dependence of series convergence (i.e., finite predictability in his studies) on spectral discretization, and, then, re-examined its dependence on the slopes of a spectrum. Our analysis not only supports LTD22's result but also provides aid regarding proper interpretations for Lilly's formula. Due to the dependence on spectral discretization, implying different impacts using different multiscale interactions, we, furthermore, provide comments on the validity of the two formulas in the support of the two-week predictability limit.

This paper, which combines a summary of a recent review by the authors, as well as new mathematical analyses, is organized as follows. Section 2 documents a summary of a recent review regarding Lorenz's view on the predictability limit during the 1990s and 2000s. We additionally provide a brief review of 2D and 3D turbulence. In Section 3, we first compare the similarities and differences of Lilly's and Lorenz's formulas, which were originally proposed based on different physical time scales, and report the mathematical discrepancy of the two formulas. We then apply~~generalize~~ Lilly's formula ~~and apply it in order to illustrate~~ to reiterate the dependence of finite predictability on different spectral discretization for a non-uniform, stretching grid and a uniform grid. Section 4 provides a summary. Appendix A includes discussions regarding the impact of different discretization on the convergent property of a simple function of $1/k$. Appendix B provides the mathematical details of the Lilly's formula over the uniform and non-uniform grids. Supplementary Materials provide a review for the integral test and convergent properties of the so-called p-series (Stewart 2014). Additionally, Supplementary Materials (e.g., Eqs. S10a and S10b) analyze the integral over a non-uniform grid in Leith and Kraichhan (1972) to demonstrate that, as a result of the variable change, a scaling factor in the form of a Jacobian should be taken into account.

2. A Review of Related Studies

2.1 Lorenz's View on the Predictability Limit

Based on the content of Lorenz's studies during the 1990s and 2000s (Lorenz 1993, 1996, 2006) and his responses in an interview from 2007 (Reeves, 2014), Shen et. al. (2023a) summarized Lorenz's view on predictability, as follows:

A. The essence of a finite predictability limit within a chaotic system (e.g., atmosphere) was (qualitatively) revealed using the Lorenz 1963 model. However, the 1963 model did not quantitively determine a limit for the predictability of the atmosphere.
B. During the 1960s, the so-called two-week predictability limit was originally estimated based on a doubling time of five days in real world models. Since that time, such a finding was documented in Charney et al. (1966) and has become a consensus.

While Lorenz's major predictability estimates for the 2-week limit were reported in the 1960s, different estimates of predictability limits for an approximate 3-week limit were also discovered during the 1970s and 1980s (Lorenz 1970, 1972, 1984, 1985) using the same 1969 model and included a spectral gap (e.g., Figure 2 of Shen et al. 2023a). Interestingly, based on our literature review, none of Lorenz's five studies from 1969 were cited in the 1993 book that presented a historical perspective for choosing two weeks as the basis for the predictability limit. Additionally, the fact that the Lorenz 1969 model is not chaotic has been overlooked. As a result, our study sought to provide insight on whether and how Lorenz's or Lilly's formula could quantitively determine an intrinsic limit of two weeks for the atmosphere. Our study was specifically designed to understand the relationship of the two formulas and, thus, to examine their validity in quantitatively or qualitatively revealing the role of small processes in contributing predictability.

2.1 A Brief Review of 2D and 3D Turbulence

In contrast to finite-dimensional chaotic systems, high-dimensional irregular turbulent systems also appear within the atmosphere. Both 3D turbulence (Kolmogorov, 1941, 1962) and 2D turbulence (Kraichnan 1967; Kraichnan and Montgomery 1980) have been applied for decades in order to understand atmospheric dynamics and predictability (e.g., Lilly and Petersen, 1983; Nastrom and Gage 1983; Lindborg 1999; Lindborg and Alvelius 2000). A focus has been on nonlinear multiscale interactions (or transfer across scales) within inertial ranges (e.g., Tribbia and Baumhefner 2004), where nonlinear processes dominate (as compared to dissipation). Major features in 2D and 3D turbulence include the following: (1) 3D turbulence has a kinetic energy (KE) -5/3 power law for its inertia range, where a direct cascade of KE occurs (Figure 2a); and (2) 2D turbulence possesses two inertia ranges (Figure 2b) - one inertial range with an inverse cascade of energy possesses a KE -5/3 power law and the other inertia range with a direct cascade of enstrophy possesses a KE -3 power law. The direct and inverse cascades, respectively, indicate transfers to smaller and larger scale processes. Enstrophy is proportional to vorticity squared. Here, in both Figures 2a and 2b, as discussed in Section 3.7, we draw the readers' attention to the existence of a dissipation layer range.

Based on theoretical turbulence and observation studies, Figure 2c (e.g., the Figure of Larsen et al. 1982; Gage and Nastrom 1986) displays a composite picture for the spectra of atmospheric turbulence. Region (I), with a -3 power law, indicates 2D turbulence at synoptic scales. Region (II), with a -5/3 power law, suggests 2D turbulence within the mesoscale (Gage 1979). While both Regions (I) and (II) are associated with 2D turbulence, Figure 2c suggests different energy sources for these two regions. Region (III) in Figure 2c indicates either a -5/3 power law associated with 3D turbulence or a -3 power law associated with 2D turbulence. We additionally added Region (IV) for a dissipation layer. For scales larger than those in Region (I), a power law of -5/3 was reported by Lilly (1969) and cited by Pedlosky (1987).

Since vertically propagating gravity waves (Lilly 1983) and vertical convection also appear in the atmosphere, whether (or not) the theory of 2D turbulence may be applicable to the atmosphere has been discussed (e.g., Lilly 1983; Zilitinkevich et al. 2021) and "new" types of turbulence (e.g., stratified turbulence and convective turbulence) have been suggested. Recent studies (e.g., Pouquet and Marino, 2013; De Wit et al., 2022; Boffetta, 2023) that explored the bistability of coexistence between 2D and 3D flows, the duality of both direct and inverse cascades, and dimensional transitions between 2D and 3D turbulence, have illuminated the intricacies of turbulence, weather, and climate. These studies warrant a cautious interpretation of the findings from earlier studies conducted in the 1960s and 1970s.

~~However, S~~since our focus is on the relationship between Lorenz's and Lilly's formulas, discussions regarding stratified and convective turbulence as well as bistability and duality are omitted. We simply emphasize that while weather possesses both turbulent and non-turbulent components, predictability in Lorenz's and Lilly's studies is associated with stationary turbulence (e.g., Lorenz 1969d; Leith 1971; Leith and Kranchnan 1972; Lilly 1972, 1973, 1990) instead of decaying turbulence (Metais and Lesieur 1986).

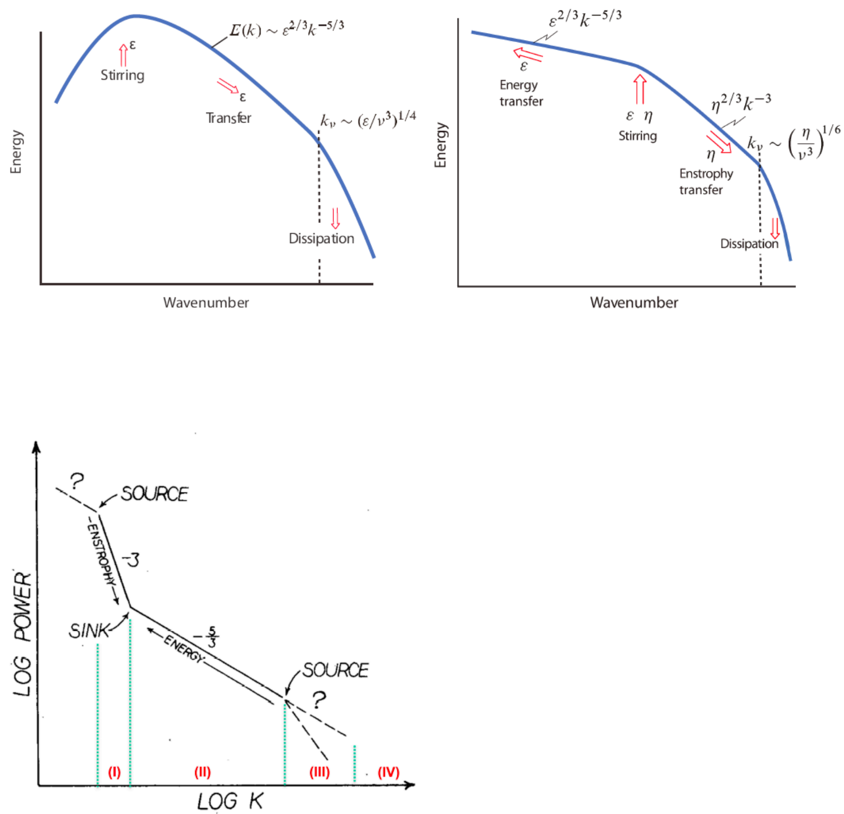

Figure 2: (a) Spectra of the 3D turbulence with a direct cascade of energy. (b) Spectra of 2D turbulence with an inverse cascade of energy in one inertia subrange and a direct cascade of enstrophy in the other inertia subrange. (c) A compositive spectra of the atmosphere (courtesy of Vallis 2007 for panels (a) and (b) and of Larsen et al. (1982) for panel (c).

3. Discussion

In this section, we first provide definitions for various time scales and introduce two types of numerical discretization, document major features of the L69 model, provide equations that define Lorenz's and Lilly's predictability formulas, reanalyze Table 3 of Lorenz (1969d), compare the discrete and continuous versions of Lilly's formula, and extend Lilly's formula for a different data grid. We briefly discuss differences in energy transfer across scales and spaces near the end of the section. Detailed mathematical discussions of the Lilly integral are presented in Appendix B.

3.1 Definitions of Various Time Scales

To date, various types of time scales have been used for determining predictability horizons (e.g., Lorenz 1996; Rotunno and Snyder 2008; Shen et al. 2022a, b). For example, in Lorenz (1969d), the predictability horizon was estimated using the saturation time (as well as saturation time difference). To facilitate discussions, major assumptions from Lorenz (1969d) are listed in Figure 1. Definitions for the various time scales are provided below. In this study, k represents the wavenumber of the Fourier mode for the background or the perturbation KE, which is consistent with turbulence theory. As listed in Table 1 from Lorenz (1969d), the original discretization scheme, $k = 2^{n-1}k_L$, is applicable to a non-uniform grid. Here, n is an integer and k_L represents the smallest wavenumber. For comparison, we also applied a different discretization scheme, $k = nk_L$, for a uniform grid.

Time scales related to predictability estimates include:

1. The saturation time (t_n), which is defined as the time for the perturbation at a particular wavenumber to become saturated (i.e., reaching the value of background KE).
2. The saturation time difference (SDT_n), which is computed by subtracting saturation times for perturbations at two successive wave number indices (i.e., $SDT_n = t_n - t_{n+1}$, as discussed in Table 1).
3. The turnover time (τ_k), which is the time for a parcel with velocity v_k to move a distance of $1/k$, with v_k being the velocity associated with wavenumber k (e.g., Vallis 2006). The turnover time is further used to indicate the time that an error at one wavenumber spreads to another wavenumber, a movement within the spectral space (e.g., LTD22).
4. The e-folding time or doubling time which represents the time for a specific mode with a growth rate (that depends on wavenumber k) to increase by a factor of e (≈ 2.71828) or two.

Based on the above, we can describe the turnover time as the time for perturbation transfer across scales, and saturation time as the maximal time interval for the growth of a perturbation at a specific scale. Thus, the turnover time and saturation time are physically different. In addition, the concepts of energy transfer across scales and spatial spaces are different, suggested by the following overlooked feature (Castelvecchi 2017):

Kolmogorov's picture implies that energy spreads from large swirls to smaller eddies nearby, rather than spreading to farther distances

In Charney et al. (1966), a doubling time (e.g., e-folding time) was applied to estimate the predictability horizon in a general circulation model. ~~In literature, by comparison,~~ Recently, the sum of e-folding times associated with various growth rates at different scales has been suggested for qualitatively illustrating a system predictability horizon within the L69 model (Shen et al. 2022a). Mathematically, error growth that involves e-folding time (or doubling time) and saturation time can be illustrated using a linear ordinary differential equation (ODE) and the Logistic ODE with a quadratic term, respectively (See a concise review by Shen 2024b).

By comparison, in chaotic systems, Lyapunov exponents were computed to measure the long-term averaged rate of divergence of nearby trajectories. Therefore, the largest Lyapunov exponent can be roughly interpreted as the long-term averaged growth rates of errors. It has been applied to predictability horizons. However, the time-averaged properties caution against the proper interpretation of the estimated predictability (e.g., Shen 2024b). Furthermore, it has been demonstrated that the largest Lyapunov exponent is related to the turnover time at the smallest scale within the system (Ruelle 1979; Aurell, Boffetta et al. 1996). However, it has never been proven that a sequence of turnover times (within a range of wavenumbers) and a "sequence" of Lyapunov exponents are proportional.

In Lorenz (1969d), the saturation time (t_n) determines the predictability horizon at wavenumber $k = 2^{n-1}k_L$ within a non-uniform grid. As shown in the third column of Table 1, the saturation time at the smallest wavelength of 38 m for $k = 2^{20}k_L$ (i.e., $n = 21$) is 2.9 minutes. As indicated in Figure 1 (i.e., from Lorenz 1969d), the sum of (estimated) saturation time differences over 21 selected wavenumbers on a non-uniform grid is applied in order to represent a predictability horizon at the largest scale. Such a concept has effectively promoted research to estimate the impact of unresolved scales and to address questions of whether (or not) predictability is finite. New insights on this approach are provided in this study.

In contrast, based on the turbulence theory, Lilly (1972, 1973) computed the sum of turnover times over a set of wavenumbers on the same non-uniform grid used for estimating the predictability horizon. As discussed, turnover time and saturation time are physically different. Turnover time represents the time for energy transfer across scales rather than energy growth. More importantly, an implicit assumption for Lilly's approach is that once a specific scale is influenced (or contaminated), it immediately loses predictability. Thus, Lilly's and Lorenz's approaches are compared below.

~~In literature, by comparison, the sum of e-folding times associated with various growth rates has been suggested for qualitatively illustrating a system predictability horizon within the L69 model (Shen et al. 2022a). Mathematically, error growth that involves e-folding time (or doubling time) and saturation time can be illustrated using a linear ordinary differential equation (ODE) and the Logistic ODE with a quadratic term, respectively (See a concise review by Shen 2024).~~

As outlined in the Introduction, Lorenz (1969d) first introduced a sequence to estimate the impact of unresolved scales on system predictability. In the early 1970s, Lilly further developed Lorenz's concept by providing a mathematical formulation, making it more verifiable. Accordingly, this study begins by exploring the mathematical foundations of these formulations, focusing on the convergent properties of Lilly's formula and its dependence on different grid discretization methods. Following this, Lorenz's formula is reexamined by analyzing his original tables and comparing them with Lilly's formulation to evaluate its validity.

3.2 Two Types of Discretization and Their Impact

The aforementioned two discretization schemes generate grid points using exponential and linear functions of the wavenumber index, denoted as:

$$k_j = 2^{j-1}k_L, \quad (1a) \qquad \text{and} \qquad k_j = (j)k_L. \quad (1b)$$

In this study, these schemes are referred to as non-uniform and uniform discretization. Here, j is an integer for a sequence or series within this section and represents a real number for an integral in Section 3.6 and Appendix B. As briefly mentioned earlier, the choice in Eq. (1a) is the same as Lorenz's choice in Table 1. The specific, non-uniform grid, covering a wide range of scales from

38 m to 40,000 km, has been also utilized in other studies (e.g., Lorenz 1969d; Lilly 1972, 1973; Palmer et al. 2014; LTD22). By comparison, in this study, we additionally consider a different set of wavenumbers in Eq. (1b) for a uniform grid. While the set of wavenumbers for a uniform grid does not always represent a superset of wavenumbers for a stretching, non-uniform grid, Eq. (1b), indeed, is a superset of Eq. (1a).

The choice in Eq. (1) possesses the following features. First, within the chosen set of 21 wave modes (Lorenz 1969d; Shen et al. 2022a), each mode can interact with all selected wave modes, resulting in a coefficient matrix with 21 x 21 elements for the L69 model. As a result, the notation "$k_L \rightarrow 2k_L \rightarrow 4k_L \rightarrow \cdots 2^{n-1}k_L$ ... ", which signifies sequential cascade and, consequently, ~~may~~ provides ~~lead to~~ misleading information about~~regarding~~ scale interactions, is no longer applied for designating the selected wave modes.

Secondly, regarding the new discretization, we argue that adopting a linear function $k = nk_L$ for a uniform grid is more realistic, especially for large scales. This is the case because the nonlinear function, $k = 2^{n-1}k_L$, which excludes certain wavenumbers such as 3, 5, 7, 9, 10, 11, 12, etc., cannot accurately resolve baroclinic waves with a dominant wavelength of approximately 4,000 km (i.e., $k = 10$). Below, we compare differences between uniform and non-uniform grids. Detailed results can be found in Sections 3.3 and 3.6.

Before we examine the impact of different discretization on the integral of the turnover time. Here, we consider the function $f(k) = 1/k$ for a simple illustration. The function is representative as discussed in Section 3.3. First, we compute the sum of the function $f(k)$ over non-uniform and unform grids in Eqs. (1a) and (1b), written as follows:

$$\frac{1}{k_L}\lim_{N\to\infty}\sum_{j=1}^{N}\left(\frac{1}{2^{j-1}}\right) \quad (2a) \qquad and \qquad \frac{1}{k_L}\lim_{N\to\infty}\sum_{j=1}^{N}\left(\frac{1}{j}\right), \quad (2b)$$

respectively. Both can be expressed as follows:

$$\frac{1}{k_L}\lim_{N\to\infty}\sum_{j=1}^{N} f\left(k_j\right)\Delta j.$$

Equation (2a) is a geometric series. As discussed in Appendix A, Eqs. (2a) and (2b) represent a convergent and divergent series, respectively. This simple case illustrates the dependence of divergence and convergence on discretization. For the choice in Eqs. (1a) and (1b), we can additionally point out that when the sum over a subset of wave modes in Eq. (1a) is convergent, the sum over a superset of wave modes can be divergent. Below, we further consider integrals.

Here, we apply Riemann sums in order to "construct" (or approximate) integrals of the function $1/k$. The following discussions illustrate *that the non-uniform grid could* potentially yield a series with different convergent properties, as compared to the uniform grid. From Eq. (1a), we compute the derivative of k with respect to j:

$$\frac{dk}{dj} = \ln(2)\, k,$$

yielding:

$$\Delta j = \frac{1}{\ln(2)} \frac{\Delta k}{k}.$$

Thus, a fixed value of Δj $(e.g., \Delta j = 1)$ in the above equation requires a constant of $\frac{\Delta k}{k}$. Thus, as k changes, Δk varies. As a result, the expression $f(k_j)\Delta j$ in Eq. (1) is now approximated by $f(k)\Delta k/k/ln(2)$. Thus, the integral of the function $1/k$ over the non-uniform grid becomes $\int \frac{1}{k} \frac{dk}{ln(2)k}$. In a similar manner, we can show that the integral of the function $1/k$ over the uniform grid is $\int \frac{1}{k} dk$. After computing both integrals, we know that they produce different convergent properties.

In fact, the above scaling factor of $1/k$ for a non-uniform grid can be easily illustrated using well-established calculus. Below, we first review the concept of the Jacobian. A well-known example is given by a double integral that can be evaluated in Cartesian or Polar Coordinates. Considering an area within a grid box, we have $dA = dxdy$ in Cartesian coordinates and $dA = rdrd\theta$, where r is the Jacobian. For a single variable function, a Jacobian simply represents a derivative, representing a ratio of increments between old and new variables. Given an integral $\int F(j)d(j)$, after a variable transformation $j = g(k)$, the integral can be evaluated in the new coordinate, as follows:

$$\int F(k)|J|dk. \quad (3)$$

Here, the scale factor $J = \frac{dg}{dk}$ is called the Jacobian. Given the non-uniform grid $k = 2^{j-1}k_L$, we have $j = 1 + \; log_2\,(k/k_L) = g(k)$ and the Jacobian

$$J = \frac{dg(k)}{dk} = \frac{d}{dk}\frac{\ln\left(\frac{k}{k_L}\right)}{\ln(2)} = \frac{1}{\ln(2)}\frac{d(\frac{k}{k_L})/dk}{k/k_L} = 1/(ln(2)k).$$

A factor of $1/k$ appears. Thus, Eq. (3) becomes:

$$\int F(k)\,\frac{1}{ln(2)k}dk = \int F(k)d\,log_2(k). \quad (4a)$$

In contrast, for the uniform grid, $k = j\,k_L$, we have $j = k/k_L$ and the Jacobian $J = 1/k_{L.}$ Here, the scale factor is $1/k_L$. With the Jacobian, Equation (3) becomes:

$$\int F(k)\,\frac{1}{k_L}dk. \quad (4b)$$

As a brief summary, the above discussions suggest that *the sum of a function $F(j)$ over a uniform grid can be approximated using an integral with respect to k for the linear function $k = j\,k_L$. In comparison, the sum of the function $F(j)$ over the non-uniform grid can be approximated using an integral with respect to $log_2(k)$ for an exponential function $k = 2^{j-1}k_L$. The convergent properties of the integrals in Eqs. (4a) and (4b) are mathematically consistent with those of the series in Eqs. (2a) and (2b), respectively.* The appearance of a scale factor of $1/\mathrm{k}$ (i.e., the Jacobian) suggests that the specific non-uniform grid could potentially change the power-law properties of the sum of a function, as compared to the sum over a uniform grid. In Supplementary Materials (e.g., Eqs. S10a and S10b), we additionally analyze the integral over a non-uniform grid in Leith and Kraichhan (1972) to demonstrate that a scaling factor in the form of a Jacobian should be taken into account.

Considering the exponential function $k = e^{j-1}k_L$, which is similar to Eq. (1a) with $k = 2^{j-1}k_L$, the corresponding sum is approximated using an integral with respect to $ln(k)$ ~~for~~. The two nonlinear transformations differ by their bases (i.e., *2* vs. *e*) for exponential functions, as well as logarithm functions (i.e., $log_2(k)$ vs. $ln(k)$). See details in the Supplementary Materials.

Table 2: An illustration of a series for the function 1/k over a non-uniform grid and uniform grids (e.g., in the 5th column.) The corresponding integrals (in the 2nd column) contain different scale factors, yielding convergent and divergent integrals, $\int_1^\infty \frac{1}{k^2}dk$ and $\int_1^\infty \frac{1}{k}dk$. The former integral can be re-written as $\int_1^\infty \frac{1}{k}dln(k)$. Based on the integrals, we can construct a Series using a different Transformation (TR) and $\Delta j = 1$. The original table is prepared as a simpler illustration in Table 4.

Discretization	Integral	TR	Integral	Series	Convergent
Uniform	$\int_1^\infty \frac{1}{k}dk$	$k = j$	$\int_0^\infty \frac{1}{j}dj$	$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\cdots$	No
Non-uniform	$\int_1^\infty \frac{1}{k}dlog_2(k)$	$k = 2^j$	$\int_0^\infty \frac{1}{2^j}dj$	$1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}\cdots$	Yes
Non-uniform	$\int_1^\infty \frac{1}{k}dln(k)$	$k = e^j$	$\int_0^\infty \frac{1}{e^j}dj$	$1+\frac{1}{e}+\frac{1}{e^2}+\frac{1}{e^3}\cdots$	Yes

3.3 Lilly's Formula for Predictability Estimates

One of the major differences in Eqs. (1a) and (1b) is: only the linear transformation in Eq. (1b) possesses the original power law properties of $f(k) = 1/k$. Below, we illustrate how such a difference yields different divergent/convergent properties for the "sum" of turnover times associated with the KE -5/3 power law. Here, we first provide a review in order to construct Lilly's mathematical formula.

In turbulence theory, an eddy turnover time is given by the following formula (e.g., Vallis 2006):

$$\tau(k) \sim k^{-\frac{3}{2}} E(k)^{-\frac{1}{2}}, \quad (5)$$

where $E(k)$ is the background KE density. Although Figure 2c suggested different KE power laws at different scales, the KE -5/3 power law is generally analyzed in predictability studies. By plugging $E(k) \sim k^{-\frac{5}{3}}$ into Eq. (5), we obtain the following:

$$\tau(k) = C_0 k^{-\frac{2}{3}}. \qquad (6)$$

Here, to simplify the expression, a constant C_0 is introduced. The constant C_0 is a function of the viscous coefficient. Equation (6) implies a -2/3 power law for the turnover times that correspond to the -5/3 kinetic energy power law. Below, we mainly show that the features of the integral of $f(k) = 1/k$ appear in the integral of Eq. (6). (In fact, this can be also illustrated using the p-series, which is provided in the supplementary materials.)

Table 1: An analysis of Table 3 from Lorenz (1969). The first and third columns are taken from Table 3 of Lorenz (1969), while the second column for wavelengths (λ) is from Table 1 of Lorenz (1969). Here, t_n indicates the saturation time for the perturbation at wavenumber $k = 2^{n-1}$. The 3rd column represents the saturation time in minutes. The 4th column displays the saturation time differences (STD), defined as $STD_n = t_n - t_{n+1}$. The 5th column is computed to determine the ratio of successive $STDs$, defined as $m_n = -ln(STD_{n+1}/STD_n)/ln(2)$. A constant $m = 2/3$ (i.e., $STD_{n+1} = STD_n 2^{-2/3}$) was applied for estimating a predictability horizon by Lorenz (1969). The 7th column is computed for verification, as follows: $STD_{n+1} = STD_n 2^{-m_{n+1}}$, $n = 1, 2, \cdots, 19$. A higher-precision m_n is used for verification. In the 8th column, the relative error is defined as $|\frac{m_n - 2/3}{2/3}|$.

n	λ	t_n	t_n (min)	STD_n	m_n	Verification	Relative Error (%)
21	38 m	2.9 min	2.9				
20	76	3.1	3.1	0.2	2.170	0.2	225.489
19	153	4.0	4.0	0.9	0.918	0.9	37.631
18	305	5.7	5.7	1.7	0.667	1.7	0.114
17	610	8.4	8.4	2.7	0.769	2.7	15.301
16	1,221	13.0	13.0	4.6	0.666	4.6	0.061
15	2,441	20.3	20.3	7.3	0.693	7.3	3.922
14	4,883	32.1	32.1	11.8	0.687	11.8	3.082
13	9,766	51.1	51.1	19.0	0.502	19	24.759
12	19,531	1.3 hr	78	26.9	1.005	26.9	50.803
11	39 km	2.2	132.0	54.0	0.637	54	4.386
10	78	3.6	216.0	84.0	0.652	84	2.188
9	156	5.8	348.0	132.0	0.750	132	12.503
8	312	9.5	570.0	222.0	0.745	222	11.711
7	625	15.7	942.0	372.0	0.787	372	18.091
6	1,250	1.1 day	1584	642.0	0.651	642	2.372
5	2,500	1.8	2592.0	1008.0	1.000	1008	50.0
4	5,000	3.2	4608.0	2016.0	0.778	2016	16.641
3	10,000	5.6	8064.0	3456.0	0.907	3456	36.034
2	20,000	10.1	14544.0	6480.0	0.574	6480	13.865
1	40,000	16.8	24192.0	9648.0			

To illustrate the impact of different spectral discretization, we consider both the non-uniform and ~~a~~ uniform grids. Plugging Equations (1a) - (1b) into Eq. (6) yields the following turnover times:

$$\tau(k_j) = \tau(k_L)\left(2^{-\frac{2}{3}}\right)^j \; (7a) \quad \text{and} \quad \tau(k_j) = \tau(k_L)(j)^{-\frac{2}{3}}, \quad (7\text{b})$$

on grid points, respectively. Here, $\tau(k_L) = C_0 k_L{}^{-\frac{2}{3}} \equiv \tau_o$ is the turnover time for the largest scale. The formula in Eq. (7a), as shown, for example, in Figure 3, has been used in predictability studies (e.g., Lilly 1972, 1973; LTD22; Palmer et al. 2014). Note that j *can be viewed as a new variable. Thus, Eq. (7a), for a non-uniform grid, does not maintain the original power law and only Eq. (7b), for a uniform grid with a linear variable transformation, still possesses the same -2/3 power law as Eq. (6).* Details are provided below.

Similarly to. Section 3.2, we now consider the sum of turnover times $\tau(k_j)$ over the selected data points in Eq. (7), as follows:

$$\lim_{N\to\infty}\left(\sum_{j=0}^{N} \tau(k_j)\,\Delta j\right). \quad (8)$$

Below, $\Delta j = 1$, which was implicitly assumed in earlier studies, is explicitly added for a comparison of Lilly's series to his integrals in Section 3.6. The above formula in Eq. (8) is referred to as Lilly's formula.

Plugging the turnover times $\tau(k_j)$ in Eqs. (7a) - (7b), for both grids, into Eq. (8) yields:

$$\tau(k_L)\left(\frac{1}{1-2^{-2/3}}\right) = 2.7 * \tau(k_L) \quad (9a) \quad \text{and} \quad \tau(k_L)\lim_{N\to\infty}\sum_{j=0}^{N}\left(\frac{1}{j^{2/3}}\right) = \infty, \quad (9b)$$

respectively. Thus, Eqs. (9a) and (9b) represent Lilly's formula that applies the KE $(-5/3)$ spectrum over a non-uniform and uniform grid, respectively, representing different complexities of multiscale interactions. As originally derived by Lilly (1973), Eq. (9a) that yields a convergent series was applied to suggest a finite predictability horizon of 2.7 $\tau(k_L)$. This same equation was applied by LTD22 and Palmer et al. (2014). The validity of Eq. (9a) in determining the predictability horizon is examined below.

In comparison, as discussed, the choice of the uniform grid in Eq. (1b) includes wavenumbers 3, 5, 6, 7, 9, 10, etc., and, thus, is more realistic than the choice of the non-uniform grid in Eq. (1a). When the uniform grid in Eq. (1b) is considered, the corresponding sum of turnover times $\tau(k_j)$ in Eq. (9b) produces a divergent series. *Thus, Eqs. (9a) and (9b) collectively indicate the dependence of convergence (or "finite predictability") on spectral discretization.* Additional details regarding the different discretization are~~is~~ discussed in Section 3.6.

As mentioned in Section 3.2, the original Lorenz 1969 model was constructed by computing the mode-mode interactions, where each selected mode interacts with all the other selected modes. Consequently, when the two discretization schemes were applied, the resulting models could have vastly different complexities. While the validity of the findings obtained using the non-uniform discretization is being re-examined, the effectiveness of incorporating a wider range of scales using a non-uniform grid (i.e., a stretching grid) is acknowledged. From a broader perspective of applications, as the self-attention mechanism in the AI transformer technology for ChatGPT (e.g., Shen 2024a, c) computes the attention scores between any pair of words, similar to the mode-mode interactions in the L69 model, it is worth introducing the concept of a stretching grid to save computing costs for very long sequences of words, which is beyond the scope of this study.

When the non-uniform discretization in Eq. (1a) is applied, a “the magic factor” of $2^{-\frac{2}{3}}$ appears in Eq. (7a). Such a factor was first documented by Lorenz (e.g., in Figure 1 from Lorenz ~~(~~1969d) and is discussed in Section 3.4.

Lilly (1972)	$\tau(k) = A^{-1/2} k^{(n-3)/2}.$ (2.2) $\sum_{n=0}^{p} \tau(2^n k) = \frac{1-2^{(n-3)p/2}}{1-2^{(n-3)/2}} \tau(k).$ (2.3)	Eqs. (2.2)-(2.3) of Lilly (1972)
Lilly (1973)	$\sum_{n=0}^{\infty} \tau(2^n k) = \frac{\alpha^{-1/2} \varepsilon^{-1/3} k^{-2/3}}{1 - 2^{-2/3}} \sim 2.7\alpha^{-1/2} \varepsilon^{-1/3} k^{-2/3} = 2.7\tau(k)$. (9.3)	Eq. (9.3) of Lilly (1973)
Lilly (1990)	$T_e(k) = 1/[kV(k)] = [k^3 E(k)]^{-1/2}.$ (2.1) $T_p(k) = \int_k^{k_r} (T_e/k)\, dk = \int_k^{k_r} [k^5 E(k)]^{-1/2}\, dk.$ (2.2) Lorenz assumed $E(k) \sim k^{-5/3}$ for wavelengths much less than the earth's circumference, leading to the relation $T_p \sim 1{\cdot}5 T_e$. The scaling velocity $V(k)$ of thunderstorm	Eqs. (2.1) and (2.2) of Lilly (1990)

Figure 3: The evolution of Lilly's formulas for determining a predictability horizon. All the above formulas are consistent. Notably, the presence of the factor 1/k in the integral simplifies the understanding of the effect of the non-uniform discretization.

3.4 A Re-examination of Lorenz's Empirical Formulas

Lilly's formula in Section 3.3 was constructed based on turnover time within the specific spectral grid. In Eq. (9a), the appearance of a common factor of $2^{-2/3}$ has been viewed as evidence for the relationship between Lilly's and Lorenz's formulas. Below, we reanalyze Table 1 in order to determine the condition under which a common factor of $2^{-2/3}$ may appear within the successive saturation time differences.

3.4.1 The Reconstruction of Lorenz's Formula

Table 1, derived from Table 3 of Lorenz (1969d), lists saturation time and saturation time differences (STD_n) in the 3rd and 5th columns, respectively. As discussed in Figure 1, Lorenz

(1969d) “observed” that saturation time differences differ by a common factor ($2^{-\mathrm{m}}$)~~and $m = 2/3$, namely~~as follows:

$$\frac{STD_{n+1}}{STD_n} = 2^{-\mathrm{m}}. \quad \cancel{and \quad m = .} \qquad (10)$$

Lorenz explicitly suggested the value of $m = 2/3$. The above common factor makes it feasible to construct an infinite series for predictability estimates over an infinite set of wavenumbers (Lorenz 1969d). However,~~ it should be noted~~ we will first show that *such an infinite series is convergent for all positive* m. ~~Here, to maintain flexibility within~~ Additionally, to examine the validity of Eq. (10), we apply~~ Lorenz’s formula~~ the left-hand side of Eq. (10) to recompute~~,~~ m ~~can be empirically re-estimated~~ from data in~~(e.g.,~~ Table 1~~)~~.

From Lorenz’s “observation” for saturation time differences in Eq. (10), we first assume that two successive “estimated” saturation time differences, denoted as EST_{k+1} and EST_k, also hold the same ratio. Given a common factor ($2^{-\mathrm{m}}$)~~-~~, as shown in Table 3, the estimated saturation time differences EST_n~~, denoted as ES, are~~ can be computed~~,~~ as follows:

$$EST_n = 2^{-\mathrm{m}} EST_{n-1} = (2^{-\mathrm{m}})^{n-1} EST_1 = (2^{-\mathrm{m}})^{n-1} STD_1. \qquad (11)$$

Here, the first estimated saturation time difference is $EST_1 = STD_1 = t_1 - t_2$.

Based on the above formula in Eq. (11), the sum of estimated saturation time differences (i.e., EST_n) produces an estimated predictability horizon at the largest scale:

$$T_{est} = \lim_{N\to\infty}\left(\sum_{n=1}^{N} EST_n \,\Delta n\right) = \lim_{N\to\infty}\sum_{n=1}^{N} STD_1((2^{-m})^{n-1}). \qquad (12)$$

For a comparison with Lilly’s formula in Eq. (8), Eq. (12) is referred to as Lorenz’s formula. Both formulas represent geometric series. Overall, Eq. (12) is convergent for $2^{-m} < 1 \; as \; m > 0$ (i.e., any positive m) and divergent for $2^{-m} \geq 1 \; as \; m \leq 0$.

Table 3: An empirical formula for estimating a predictability horizon (Lorenz 1969). n is the index for wavenumber $k = 2^{n-1}$, while n^* is an index representing an average of two successive scales. The 2nd row for t_n indicates the saturation time for the perturbation at index n. The 4th row displays the saturation time differences (STD), defined as $STD_{n^*} = t_n - t_{n+1}$. By assuming that successive saturation time differences differ by a common factor of 2^{-m} with $m = 2/3$, the 5th row represents an estimate (EST_{n^*}) of $STDn^*$.

n	1		2		3		4	$\cdots$	N	
t_n	t_1		t_2		t_3		t_4	$\cdots$	t_N	
n^*		1		2		3		$\cdots$		N
STD_{n^*}		$t_1 - t_2$		$t_2 - t_3$		$t_3 - t_4$		$\cdots$		$t_N - t_{N+1}$
EST_{n^*}		STD_1		$2^{-m}STD_1$		$2^{-m}STD_2$		$\cdots$		$2^{-m}STD_{N-1}$

.

As suggested by Lorenz (1969d), plugging a special value of m = 2/3 into the above leads to:

$$T_{est} = \lim_{N\to\infty} \sum_{n=1}^{N} STD_1 \left(\left(2^{-2/3}\right)^{n-1}\right) \Delta n = STD_1 \left(\frac{1}{1 - 2^{-2/3}}\right) = 2.7 * (STD_1). \quad (13)$$

The above Lorenz predictability series appears to be the same as Lilly's formula in Eq. (9a) within the same non-uniform grid except for the factor of STD_1. For scientific accuracy, whether STD_1 is the same as $\tau(k_L)$ in $(9a)$ remains physical justifications.

The infinite series (both in Lorenz's and Lilly's formulas) has been applied in order to project the contribution of unresolved scales to predictability and to determine whether (or not) predictability is finite. Due to STD_1 = 9,648, as listed in the 4th column of Table 1, we obtained T_{est} = 2.7 * 9648 (minutes) ~ 18.1 (days) using Eq. (13), which is very close to $t_1 = 16.8$ days at the largest time scale. From Eq. (13), the contribution of unresolved scales (for $n \geq 22$) to the predictability horizon becomes:

$$\lim_{N\to\infty} \sum_{n=22}^{N} STD_1 \left(\left(2^{-2/3}\right)^{n-1}\right) = \left(2^{-2/3}\right)^{21} STD_1 \left(\frac{1}{1 - 2^{-2/3}}\right) = 1.65 * 10^{-4} * STD_1, (14)$$

which is 1.65*$10^{-4} * 9648$ (minutes) = 1.59 (minutes), which is negligible. As a result of Eq. (14), Eq. (13) with the 21 selected modes (i.e., the original 1969 study) largely represents a predictability limit in Eq. (13) with infinitely many modes.

The above reproduced Lorenz's findings. Here, we emphasize that the assumption in Eq. (10) always produces a convergent geometric series that yields a finite value and, thus, a finite predictability. Additionally, although the Lorenz's and Lilly's formulas display "mathematical" similarity for the common factor of $2^{-2/3}$, no physical foundation has been rigorously provided for establishing the linear relationship between saturation time differences and turnover times. For example, a quick, "physical" check is, as follows: while Lilly's equation of turnover time contains a constant coefficient "A" (e.g., Figure 3) proportional to the rate of the viscous dissipation of enstrophy, the L69 model does not include viscous dissipation.

More importantly, below, a mathematical reanalysis challenges the existence of the common factor of $2^{-2/3}$ in the sequence of the saturation time difference and suggests that the linear relationship between the sequences of two physical times (e.g., turnover times and saturation time differences) cannot be accurately established.

3.4.2 Reexamination of the Common Factor of $2^{-2/3}$

To examine Lorenz's discovery for the common factor of $2^{-2/3}$ (i.e., m =2/3) in the sequence of saturation time differences, as suggested in Figure 1, we computed $m = m(n)$ using the following formula, which is derived from Eq. (10)~~:~~

$$m(n) = -\frac{ln\left(\frac{STD_{n+1}}{STD_n}\right)}{\ln(2)}, \qquad (15)$$

and data from Table 1. Computed values for ($m(n)$) are provided in the *6th* column of Table 1. The computed values that vary between 0.502 and 2.170 are not exactly the same as the common factor of m = 2/3 = 0.667, as discovered and reported by Lorenz (1969d). To illustrate the discrepancies, relative errors, defined as $\left|\frac{m_n - 2/3}{2/3}\right|$, are displayed in the 8th column of Table 1. Amongst 19 relative errors for different wavenumbers, nine are larger than 15%.

To further illustrate the deviations of computed m_n from the "hypothetical" value of m = 2/3, we applied a least squares method in order to fit the computed values of m_n to the curve: $m(n) = \alpha\, n + \beta$. Parameters α and β are often called the slope and the intercept, respectively. Two fitted curves are provided in the top and bottom panels of Figure 4. The first curve applied all 19 data points, while the second curve used the first 18 data points without the "outlier" $m(20) = 2.170$. For the first case, a positive slope of $\alpha = 0.02$ and an average of the predicted values of 0.819 (denoted as $\bar{m} = 0.819$) were obtained. For the second case that excluded the outlier of $m(20)$, a negative slope of $\alpha = -0.002$ and an average $\bar{m} = 0.744$ were determined. As shown in Figure 4, the above results indicate that Table 1 (from Lorenz 1969d) does not support the idea of a common factor of $2^{-\frac{2}{3}}$ that requires m = 2/3 = 0.667, raising a concern as to whether (or not) Lorenz's findings can be explained using the Lilly's formula that is based on the turnover times in turbulence theory.

As a result, t~~T~~he above results invalidate the assumption that the saturation time difference for a given wavenumber, k, is proportional to the eddy turnover time. Our results complement the findings of LTD22. Consequently, by applying the concept of turnover time (which is based on turbulence theory) for analyzing saturation time differences and illustrating scale interactions within the L69 model (which is not a turbulence model) becomes questionable. Without such a common factor for constructing an infinite (geometric) series, estimating the contribution of new modes (or unresolved modes) to the predictability horizon within the L69 study becomes challenging. When the sum of infinitely many terms is considered, strict accuracy is required in order to determine whether (or not) such a sum is a finite number.

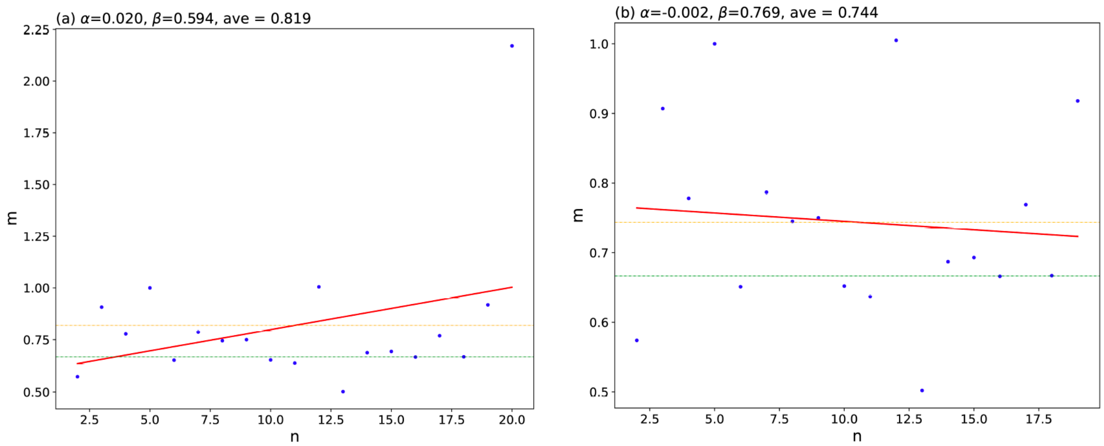

Figure 4: An analysis of m_n in the 6th column of Table 1 (blue dots). A linear least square method was applied to fit the values of m_n into a curve, $m = \alpha * n + \beta$, in red. (a) $(\alpha, \beta) = (0.02, 0.594)$ for all of 19 data points. (b) $(\alpha, \beta) = (-0.002, 0.769)$ for 18 data points without the outlier of $m(20) = 2.170$. The green dotted line represents the value of $m = 2/3$, while the orange dotted line indicates the averaged value of predicted m using the fitted curve.

3.4.3 The Impact of Different Common Factors in Eq. (10)

Under the assumption in Eq. (10), Eq. (12) indicates the existence of a geometric series with a common ratio 2^{-m} between adjacent terms. Eq. (12) is convergent when the common ratio is less than one (i.e., $m > 0$). Thus, the assumption in Eq. (10) with a positive $m > 0$ guarantees a convergent series, producing a finite predictability. Furthermore, ~~by~~ applying any common factor with $m > 0$, different from a factor of $2^{-2/3}$, simply produces a different rate of convergence, yielding a different degree of contribution from smaller scales to a predictability limit. For example, based on Figure 4, a common factor may be $2^{-0.819}$ or $2^{-0.744}$, both of which are smaller than $2^{-2/3}$. A geometric series with a smaller common factor produces a smaller value for the sum. As a result, *Table 3 of Lorenz (1969d) could still suggest a convergent series and, thus, a finite predictability*, as long as a common factor 2^{-m} with a positive m is assumed.

However, a series with a value of $m \neq 2/3$ is not the same as the one suggested by turbulence theory,. Namely, the revised Lorenz formula (with the updated $\overline{m}$) is different from Lilly's formula over the 21 selected wavenumbers. Without such similarity, we do not have a foundation for applying the turbulence theory (i.e., with turnover times) to understand the features of the Lorenz 1969 *model (e.g., for saturation times)* over the selected or unresolved wave modes.

3.5 An Illustration for Unrealistic Features

In addition to the above inconsistency between the Lorenz's findings and Lilly's formulation, we report another issue using a (-3) KE spectrum. We can consider the following discussions as illustrative of issues between the concepts of turnover time and saturation time differences (e.g., Eq. 9a and Eq. 13). When a (-3) KE spectrum is applied, Eqs. (5) and (6) collectively yield a constant turnover time for all wavenumbers. "Assuming" that saturation time differences (STD_n) are proportional to the turnover time, STD_n are constants and we have the ratio of two consecutive saturation time differences as follows:~~:~~

$$\frac{STD_n}{STD_{n+1}} = \frac{t_n - t_{n+1}}{t_{n+1} - t_{n+2}} = 1. \qquad (16)$$

Here, again, t_n represents a saturation time. Equation (16) suggests that the saturation time for a specific scale is written as $t_{n+1} = (t_n + t_{n+2})/2$, yielding the following general solution:

$$t_n = C_1 + C_2 n. \qquad (17)$$

Both C_1 and C_2 are constants. Applying $t_1 = C_0$ and $t_2 = r\, t_1$, and $r < 1$, we can determine $C_1 = (2-r)C_0$ and $C_2 = (r-1)C_0$. However, the presented sequence becomes

$$t_n = C_0(2-r) - C_0(1-r)n. \qquad (18)$$

~~in Eq. (1)~~which is not realistic. For example, when $r = 2/3$, we have $t_4 = 0$, which results in the saturation time being zero at n = 4 ~~.~~ Therefore, we believe that Eq. (16) falsifies the assumption that the saturation time difference is proportional to the turnover time for all wavenumbers.

In addition to physical justifications between the saturation time differences and turnover times, below, we present a mathematical issue of the continuous version of Lilly's formula (i.e., the integral form of the formula) ~~over continuous wavenumbers~~) regarding the dependence on spectral discretization below.

3.6 The Continuous Version of Lilly's Empirical Formula (Note that the original version was moved into Appendix B)

As discussed above, if discrepancies reported in Figure 4 can be ignored, Lorenz's and Lilly's formulas can be mathematically comparable. Nevertheless, the problems arising from the uniform and non-uniform grid discretization persist. These issues are further reiterated using the integral form of Lilly's formula. The mathematical concepts underlying the discrete and continuous forms of Lilly's formulas are similar, and the discrete forms have already been discussed. Thus, while a detailed mathematical analysis of Lilly's integrals for both discretization methods is presented in Appendix B, a brief summary is provided in this section.

In Appendix B, based on the Lilly's work in 1990, we first extend the concept of "summation of turnover times over selected wavenumbers" in Lilly's formula~~ is consistently extended~~ to the concept of "an integral of the turnover time with respect to rescaled wavenumbers". Here, the term "with respect to rescaled wavenumbers" means "with respect to $ln(k)$ and k/k_L (i.e., dk/k and dk/k_L) for non-uniform and uniform discretization, respectively. We show that Lilly's series over a non-uniform grid ($k = 2^n k_L$) and Lilly's integral with respect to ln(k) (i.e., a varying scale factor of 1/k) are consistent. Then, Lilly's integral formulas for the two discretization methods are compared to emphasize their distinct characteristics.

As summarized in Table 2, the integrals of $1/k$ with respect to the above two rescaled wavenumbers can be illustrated using $\int_1^\infty \frac{1}{k^2} dk$ and $\int_1^\infty \frac{1}{k} dk$. In Appendix B~~Below~~, we present a continuous version of Lilly's empirical formula and show that it is consistent with the discrete form in Eq. (9a) for a non-uniform grid. Based on Lilly's original integral, we further ~~then~~ generalize the integral for the uniform grid, consistent with the discrete form in Eq. (9b). By applying Lilly's two integral formulas, we reiterate ~~address~~ the dependence of the integral's convergence on the two spectral discretization methods for the KE -5/3 spectrum.

In fact, after helpful discussions with reviewers, the following mathematical analysis can better illustrate the impact of the Jacobian scale factor and, thus, provide additional support to our findings. To facilitate discussions, we use the turnover time $\tau(k) = C_0 k^{-\frac{2}{3}}$ in Eq. (6) and compute the derivative and integral of the turnover time with respect to (ln(k)), as follows:

$$\frac{d\tau(k)}{dln(k)} = \frac{d\tau(k)}{\frac{1}{k}dk} = k\frac{d\tau(k)}{dk} \sim kC_0\left(-\frac{2}{3}\right)k^{-\frac{5}{3}} = -\frac{2}{3}\tau(k) \quad (19)$$

$$\int_{k_o}^{k_1} \tau d\ln(k) = C_o\left(\int_{k_o}^{k_1} k^{-5/3}dk\right) = -\frac{3}{2}\left((\tau_k(k_1) - \tau_k(k_0)\right) \quad (20)$$

The equations above suggest that "*the derivative and integral of turnover time with respect to the natural logarithm of x (i.e., ln (x)) are proportional to the turnover time itself.*" They follow the same power laws. In fact, this property in the quote remains valid as long as the turnover time follows a power function of wavenumber k, with the exception of constant functions.

As a result, for the KE -5/3 power law in Eq. (20), the convergent property of the integral is determined by the turnover time at the highest wavenumber, leading to a finite value of $\int_{k_o}^{k_1} \tau d\ln(k) = \frac{3}{2}\tau_k(k_0)$ as $k_1 \to \infty$.

3.7 The Validity of Lilly's Formula in Determining the Predictability Limit

Although Lorenz's and Lilly's approaches have effectively promoted related investigations, misunderstandings and misinterpretations appear and cause some issues that explicitly or implicitly inhibit research. Important issues concerning the validity of Lilly's formulas in determining the predictability limit are summarized as follows:

1. Saturation and turnover times are physically different.
2. The convergence of the Lilly's formula depends on both spectral slopes and spectral discretization. Over the uniform grid, the integral of turnover times for both -5/3 and -3 power laws are divergent when the largest wavenumber approaches infinity.
3. When a frictional layer appears over the largest wavenumbers, the interval of the inertia subranges should be finite. As a result, the integral of turnover times over the finite

interval should be finite, suggesting convergent integrals for both -5/3 and -3 power laws.

Additional details are provided below.

First, whether (or not) saturation time differences in the Lorenz's formula and turnover times in the Lilly's formula can have a linear relationship for each wavenumber is not clear. Each sequence of "saturation time differences" and "turnover times" at various wavenumbers can be viewed as a "vector" with infinitely many components. From a mathematical perspective, showing that the two vectors are "parallel" is challenging. Recently, a study by LTD22, who applied a real-world model in order to perform a predictability study, indicated that it is not appropriate to assume that the error-growth time scale for a given wavenumber k is proportional to the eddy turnover time.

Secondly, a uniform spectral discretization in Eq. (1b) is more general than the non-uniform discretization in Eq. (1a). However, as discussed in Section 3.3, Section 3.6, and Appendix B, the convergence of the Lilly's formula displays dependence on discretization.

Thirdly, the above turnover-time-based discussions within inertia range(s) implicitly indicate how the impact of a dissipation layer should be considered. Namely, the above discussions are valid within inertia ranges where nonlinear interactions dominate, as compared to dissipations. If a dissipation layer exists, *as indicated in Figure 2c, the upper bound of the wavenumber for the inertia range is finite, thus, the integral of turnover time with respect to k/k_L or ln(k) within the inertia ranges should be finite.* Can such a result be applied in order to determine finite intrinsic predictability for the atmosphere?

Lastly, the assumption of homogeneity and isotropy, as often applied in turbulence theory, cannot be universally applied to weather in all places for all time scales. When applying findings from classical turbulence theory, one must take into consideration the fact that real weather contains both fully turbulent and non-turbulent components, thus providing different environments for perturbations to grow and transfer across space and time. In contrast, as briefly discussed in Section 2.1, recent advancements in turbulence research have enabled the application of novel concepts, such as bistability, to effectively illustrate the complexities of weather and climate. These new concepts should be taken into account when determining the predictability limits.

3.8 The Validity of Lorenz's Formula in Determining the Predictability Limit

While the concept of turnover time appears within systems that contain dissipations, the L69 model was originally derived from a conservative partial differential equation that conserves vorticity. Additionally, a recent study (Shen et al. 2022a) suggested that the L69 model is a closure-based, physically multiscale, mathematically linear, and numerically ill-conditioned system. The linear feature of the L69 model is also recognized by Saiki and Yorke (2023). Thus, the L69 is not a turbulence model nor a chaos model. The concept of turnover time cannot be directly applied to examine the findings from the L69 models.

Other than the above, the following common ground shared by one of the reviewers, provides additional support to our analysis:

The reviewer appreciates these efforts and would like to highlight some common ground that both the authors and the reviewer agree upon in the revision.

1) *L69 is not a model for the real atmosphere, and it is not "chaotic" under the author's definition of chaos. However, it could still provide valuable insights into the error growth for a multi-scale system.*
2) *L69 proposed the two-week predictability limit, which was a revolutionary insight. While the predictability limit of the real atmosphere remains unknown, this limit has been verified by many complex global cloud-resolving systems, especially for mid-latitudes. It is also acknowledged that the predictability limit could differ for different regions, for example, it could be longer in the tropics, where the circulation scale (MJO) is much larger.*
3) *It is important to keep in mind that any results obtained from L69 may not always hold true for the real atmosphere. Thus, showing that L69 is inappropriate or based on strong assumptions does not necessarily mean that the real atmosphere has a longer predictability limit.*

In response to the above third comment, we simply point out that the appearance of the Madden-Julia Oscillation (MJO) in the 2nd comment suggests the potential for longer predictability. In a recent study using an AI-powered model, remarkable 30-day ensemble simulations of MJO were presented (e.g., Figure 11 of Lang et al. 2024)

Although predictability estimates within Lorenz (1969d) have been highly cited, the role of a spectral gap in extending predictability horizons was also illustrated by Lorenz himself during the 1970s and 1980s (e.g., Lorenz 1970; 1972c; 1984; 1985), which have been overlooked. Thus, the two-week predictability limit was not robustly determined by the Lorenz 1969 model. The findings of estimated predictability within the Lorenz's formula were not robustly supported by the Lilly's formulas. Readers with interest in features of the L69 model and Lorenz's updated view on the predictability limit are referred to Lorenz (1993), Reeves (2014), and our recent studies (Shen et al. 2022a, 2023a, b, 2024)

3.9 Differences in Energy Transfer Across Scales and Spaces

As discussed above, the saturation time and the turnover time are different. The first is associated with perturbation growth, while the second is associated with the energy transfer of perturbation across scales. Here, we emphasize that energy transfers across scales and spaces are different, as also suggested Castelvecchi (2017). In Exp-A of Lorenz (1969d), which focuses on the impact of an initial perturbation at a small scale, a perturbation is provided at a specific, small wavelength (i.e., a large wavenumber) within the spectral (or wavenumber) space. As a result, the perturbation, indeed, represents a periodic signal for the entire physical world, yielding spatially periodic "butterfly flaps". Here, *energy transfer across physical space is automatically complete*. However, to have a non-negligible impact on the real world at larger spatial scales and distances, such a perturbation must grow, requiring an energy source.

In comparison, when a perturbation is prescribed as a Dirac delta function (or as a localized signal) within the physical space, the initial perturbation automatically appears to have the same amplitude for all selected wavelengths (or for a wide range of wavelengths). *Energy transfer across all (or many) scales is automatically completed*, and perturbations at all scales can immediately grow. A future study will address how perturbations at different scales can grow at different growth/decay rates to form or impact spatially-coherent weather systems. In fact, the dependence of predictability horizons on different types of initial errors (i.e., periodic or local type) was documented as early as the 1960s (e.g., Charney et al. 1966).

4. Concluding Remarks

In his 1969 studies, Lorenz utilized the Lorenz 1969 (L69) model alongside the saturation time to illustrate how predictability depends on different scales, estimating a predictability of 16.8 days for the largest wavelength. He also proposed a geometric series based on the sequence of saturation times at different scales to estimate the contribution of small-scale processes to predictability enhancement, known as Lorenz's 1969 formula. Inspired by the factor of $2^{-2/3}$ in Lorenz's formula, Lilly applied turbulence theory in the early 1970s to develop a series summing turnover times to reconstruct Lorenz's series. Although Lorenz's and Lilly's formulas appear similar (e.g., Eq. 9a and Eq. 13), our study revisited their consistency and found that they differ both physically and mathematically.

Based on our analysis and a literature review, the major discrepancies and inconsistencies are as follows:

1. Different Physical Time Scales: Lorenz's and Lilly's empirical formulas were derived using different physical time scales, including saturation time differences and turnover times over the 21 selected wave modes. Saturation time is the scale for the growth of energy, while turnover time is the scale for energy transfer across scales.
2. No Common Factor of $2^{-2/3}$ in Saturation Time Differences: Our revisit of Lorenz's results indicates that successive saturation time differences do not follow a common factor (i.e., $2^{-2/3}$). Consistent with LTD22's findings, our results do not support the assumption that saturation time difference is linearly proportional to turnover time for each selected mode.
3. Geometric Series Assumption: Lorenz's formula's assumption in Eq. (10), involving 2^{-m}, produces a geometric series that guarantees a convergent series for any positive m. This assumption should be applied with caution.
4. Convergent Properties and Discretization: Ignoring the discrepancy between the two formulas, we demonstrate that the "same" formula displays dependence of convergent properties on spectral discretization for the KE -5/3 power law. The new uniform grid discretization is more realistic compared to the original non-uniform discretization that misses certain wavenumbers (3, 5, 6, 7, 9, 10, etc.). Our results, summarized in Tables 4

and 5, imply that whether or not the predictability horizon is finite cannot be robustly determined based on the integral's convergent property of turnover time.

5. Unrealistic Saturation Time Sequence: Assuming both formulas are the same, they produce an unrealistic sequence of saturation time differences for the KE -3 power law.

Furthermore, our recent review of Lorenz's 1969 model and Lorenz's updated view on predictability reveals the following:

- The L69 model is closure-based, physically multiscale, mathematically linear, and numerically ill-conditioned. It is not a turbulence model or chaotic system as it lacks dissipative terms.
- Although the L69 model suggested a predictability of 16.8 days, Lorenz’s later studies in the 1970s and 1980s indicated that the presence of a spectral gap could extend predictability up to three weeks (e.g., Lorenz 1972, 1985).
- The two-week predictability limit was not robustly established by the L69 model and Lorenz's formula. Instead, it was estimated using a doubling time of 5 days from the Mintz-Arakawa model in the 1960s (Charney et al. 1996; GARP 1969; Shen et al. 2023a, 2024). This history is documented in Lorenz's book "The Essence of Chaos" and a review titled "Edward Lorenz Revisiting the Limits of Predictability and Their Implications" (Lorenz 1993; Reeves 2014).
- Lorenz's 1993 book attributes the two-week predictability limit to Charney's 1966 report but does not discuss any of his five studies from 1969 or Lilly's studies from 1972 and 1973.
- The differences in physical processes between the L69 model (without thermodynamic feedback) and real-world models make direct comparisons challenging.

Acknowledgment: **We express our gratitude to the Editor and two reviewers of NPG for their insightful and valuable comments.** We initially wrote the information contained herein as a Comment (e.g., Shen, Pielke, and Zeng, unpublished). During the review process, two reviewers suggested that we should submit our findings as a stand-alone article. Valuable comments provided by reviewers are greatly appreciated. The first author additionally and sincerely thanks Prof. Dale Durran and Mr. Daniel Lloveras, Prof. Geoffrey Vallis, Dr. Richard Rotunno, and Dr. Yongqiang

Sun for valuable discussions. Their open-mindedness is greatly appreciated. The first author is grateful for comments on an earlier version of the draft by Dr. Jerome Gilles, and for the verification of derivations by Mrs. Wei Paxson. Some details regarding consistency issues within the scientific literature were incrementally added in order to help readers, in particular, young researchers, view the predictability study from a different angle.

Data Availability: Once the manuscript is accepted and published, data will be available on the lead author's website (https://bwshen.sdsu.edu/).

Appendix A: Dependence of Convergence Property on Discretization

To simplify discussions, the following is provided in order to illustrate the series over a superset and a subset. Considering the integral $\int \frac{1}{x} dx$, we can construct the following two series:

$$\sum_{n=1} \frac{1}{n} \quad and \quad \sum_{j=1} \frac{1}{2^j}.$$

As shown in Table A1 below, the first represents a divergent series over a uniform grid with a superset $(x = n)$, while the second is a convergent series over a non-uniform grid with a subset $((x = 2^j)$. On the other hand, please note that a different subset may also lead to a divergent series. This is an interesting feature of a divergent series. (Thus, within the main text, a specific "subset" and "superset" were chosen for facilitating discussions.) Please see details in the Supplementary Materials.

Table A1: Three series derived from the integral of $\frac{1}{x}$, labeled as S-A, S-B, and S-C, respectively.

S-A	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{11}$	$\frac{1}{12}$	$\frac{1}{13}$	⋯	$\sum_{n=1} \frac{1}{n}$	divergent
S-B	1	$\frac{1}{2}$		$\frac{1}{4}$				$\frac{1}{8}$						⋯	$\sum_{j=0} \frac{1}{2^j}$	convergent
S-C		$\frac{1}{2}$	$\frac{1}{3}$		$\frac{1}{5}$		$\frac{1}{7}$				$\frac{1}{11}$		$\frac{1}{13}$	⋯	$\sum_{p\ prime} \frac{1}{p}$	divergent

Stated alternatively, within the revised draft and in the Supplementary Materials, based on both the discrete and continuous versions of Lilly's formulas (e.g., series and integral), we show that a specific discretization ($k = 2^{n-1} k_L$) was applied by both Lorenz (1969d) and Lilly (1972, 1973, 1990) in order to obtain a convergent series for the KE -5/3 power law.

Appendix B: Lilly’s Integrals for The Non-Uniform and Uniform Discretization

In this section, we extend Lilly’s formula from a non-uniform grid to a uniform grid and examine how the convergence of Lilly’s formula depends on the two discretization methods.

B1. Lilly's Integral for the Non-uniform Discretization

Since a non-uniform discretization ($k = 2^{n-1}k_L$) was applied, Lorenz's formula was built on the sum of saturation time differences over a non-uniform grid. Thus, Lilly’s series and integral were originally proposed based on the same grid system (e.g., Lilly 1972, 1973, 1990). In this subsection, we first present Lilly’s original integral, which is an integral of turnover time with respect to $ln(k)$ (e.g., Lilly 1990; Vallis 2006). We then show that the scale factor of $1/k$ is, indeed, a Jacobian associated with a variable transformation and that Lilly's integral is consistent with the discrete version in Eq. (9a) over a non-uniform grid. In the next subsection, we extend Lilly's original integral to the integral of turnover time with respect to k, consistent with the discrete version in Eq. (9b) for the uniform grid.

Below, we begin with Eq. (8.80) of Vallis (2006), consistent with Lilly's equation in Figure 3, as follows:

$$T = \left(\int_{k_o}^{k_1} \tau d\ln(k) \right). \qquad (B1)$$

The above formula that represents the integral of turnover time with respect to $ln(k)$ is referred to as Lilly’s (integral) formula. By plugging the turnover time in Eq. (6), with the $k^{-5/3}$ KE spectrum, into Eq. (B1), we obtain the following:

$$T = C_o \left(\int_{k_o}^{k_1} k^{-2/3} d\ln(k) \right). \qquad (B2a)$$

The above can be rewritten, by introducing a new variable, $y = \ln(k)$, as follows:

$$T = C_o \left(\int_{y_o}^{y_1} e^{-(2/3)y}\, dy \right), \qquad (B2b)$$

where its lower and upper bounds are $y_o = \ln(k_o)$ and $y_1 = \ln(k_1)$. The derivations from Eq. (B2a) to Eq. (B2b) only introduce a new variable y and do not pose any assumption.

The above discussions can be illustrated using a Jacobian and Eq. (3). Now, from Eq. (B2b), we derive Eq. (B2a). Considering the integral in Eq. (B2b) with a variable transformation $y = ln(k)$, the corresponding integral in the new variable k space is written, as follows:

$$T = C_o\left(\int_{k_o}^{k_1} e^{-(2/3)y(k)} J \, dk\right). (B3)$$

Here, $k_o = e^{y_o}$ and $k_1 = e^{y_1}$. J represents a Jacobian. The integrand of $e^{-(2/3)y(k)}$ is indeed $k^{-2/3}$, and the Jacobian is $J = |dy/dk| = 1/k$. By plugging the integrand and the Jacobian into Eq. (B3), we obtain Eq. (B2a). Once again, the scale factor $1/k$ appears in the integral for a non-uniform grid.

To be directly compared to the discrete version in Eq. (9a) for a non-uniform grid, we replaced $ln(k)$ in Eq. (B2a) by $log_2(k)$ to obtain:

$$T = C_o\left(\int_{k_o}^{k_1} k^{-2/3} d\log_2(k)\right). \quad (B4a)$$

We now introduce a new variable $y = \log_2(\mathrm{k})$ and turn Eq. (B4a) into the following:

$$T = C_o\left(\int_{y_o}^{y_1} 2^{-(2/3)y} \, dy\right). \qquad (B4b)$$

Eq. (B4b) can be applied in order to construct a series for a constant value of Δy, producing a series consistent with the one shown in Eq. (9a).

B2. Lilly's Integral for the Uniform Discretization

Below, motivated by Eq. (B1), we define a time scale as "the integral of turnover time with respect to wavenumber" divided by the reference wavenumber k_L, as follows:

$$T_{uniform} = \frac{1}{k_L}\left(\int_{k_o}^{k_1} \tau\, d\, k\right). \quad (B5)$$

As discussed below, $T_{uniform}$ represents the sum of turnover times over the entire set of wavenumbers for a uniform grid. For comparison, Eq. (B5) and Eq. (B1) are rewritten, respectively, as follows:

$$T_{uniform} = \left(\int_{k_o}^{k_1} \tau \frac{dk}{k_L}\right), \quad (B6a)$$

and

$$T = \left(\int_{k_o}^{k_1} \tau\, \frac{dk}{k}\right). \quad (B6b)$$

Both integrals have the same unit. Thus, both can be viewed as integrals of turnover time with respect to the "rescaled" wavenumbers. The difference between the two integrals is that the first integral has a constant rescaling factor while the second integral contains a k-dependent rescaling factor. As discussed below, the constant rescaling factor for the uniform grid can also be illustrated using the Jacobian.

As compared to Lilly's series in Eq. (1b) for a uniform grid, the following expression is considered:

$$k = (j)k_L. \quad (B7)$$

In general, the new variable j is a real number. By plugging Eq. (B7) into Eq. (6), we obtain the following turnover time:

$$\tau(j) = C_0\left(j^{-\frac{2}{3}}\right)k_L^{-\frac{2}{3}} = \tau(k_L)\left(j^{-\frac{2}{3}}\right), \quad (B8)$$

where $\tau(k_L) = C_0\, k_L^{-\frac{2}{3}}$. Using Eq. (B8), Eq. (B5) yields:

$$T_{uniform} = \tau(k_L)\left(\int_{j_1}^{j_2}\left(j^{-\frac{2}{3}}\right)d(j)\right). \quad (B9)$$

In Eq. (B9), since the quantity within the parentheses is dimensionless, $T_{uniform}$ has the same unit as the turnover time, $\tau(k_L)$. Similar to the previous subsection, from Eq. (B9) with a variable transformation of $j = k/k_L$, we can convert the integral in Eq. (B9) to Eq. (B5), where $1/k_L$ represents a Jacobian, $J = |dj/dk| = 1/k_L$. Since Eq. (B9) is equivalent to Eq. (9b), the integral in Eq. (B7) that applies a uniform discretization in Eq. (B7) is consistent with the series in Eq. (9b)

B3. The Dependence of Convergence on Discretization

Discussions in the previous subsections suggested that the continuous version of Lilly's formulas for different grids, Eq. (B1) for a non-uniform grid and Eq. (B5) for a uniform grid, produce different convergent properties. Such findings are consistent with those obtained using the corresponding discrete version, Eq. (9a) for the non-uniform grid and Eq. (9b) for the uniform grid.

While the discrete version in Eqs. (9a) and (9b) is based on "a sum of turnover times" over all data points, the continuous version in Eqs. (B5) and (B7) represents the integral of turnover time with respect to rescaled wavenumbers. Rescaled wavenumbers that can be determined by the Jacobian are $ln(k)$ and k/k_L for the non-uniform and uniform discretization, respectively. Table 4 provides a summary for the discrete and continuous versions of Lilly's formulas for the two discretization.

Table B1: A comparison of the continuous (Column 2) and discrete (Column 4) versions of Lilly's formulas for non-uniform and uniform grids. Eq. (B4a) is the original Lilly's integral with respect to $\log_2(k)$, and Eq. (B5) is Lilly's integral with respect to k. This table can be compared to Table 3.

Grid	Integral	Transformation	Series	Convergent
Non-uniform	Eq. (B4a)		Eq. (9a)	Yes
	$\frac{1}{ln(2)}\int_{k_o}^{k_1} k^{-\frac{2}{3}}\frac{dk}{k}$	$k = 2^j k_L$	$\approx \lim_{N\to\infty}\sum_{j=0}^{N}\left(\left(2^{-\frac{2}{3}}\right)^j \Delta j\right)$	
Uniform	Eq. (B5)		Eq. (9b)	No
	$\int_{j_o}^{\infty} k^{-\frac{2}{3}}\frac{dk}{k_L}$	$k = jk_L$	$\approx \lim_{N\to\infty}\sum_{j=0}^{N}\left(j^{-2/3}\,\Delta j\right)$	

If we agree that the sum of turnover times over the non-uniform grid ($k = 2^j k_L$) can represent the predictability of the atmosphere, the sum over the uniform grid ($k = j\,k_L$), including wavenumbers 3, 5, 6, 7, 9, 10 etc., should better represent atmospheric predictability. As discussed above, given the same KE -5/3 power spectrum, integrals of turnover time two different grids are either convergent or divergent, respectively. Lilly's formula cannot definitely determine whether (or not) atmospheric predictability with a KE -5/3 power spectrum is finite.

Below, we consider a general case in order to support the claim that the given KE -5/3 power law, the convergent property for integrals of turnover time over non-uniform and uniform grids, can be illustrated using the integrals of $1/k$ over the two grids as discussed in Section 3.2. For the general KE $-s$ power laws, the background KE energy and the corresponding turnover time are written:

$$E(k) = C_o k^{-s} \text{ and } \tau(k) = \tau_o k^{-(3-s)/2},$$

respectively. τ_o is defined in Section 3.3 and $-s$ indicates the power of KE energy Table 5 provides convergent properties for Lilly's integral with respect to (k/k_L) or $ln(k)$. The two integrals with respect to k/k_L and $ln(k)$ are divergent when s ≥ 1 and s ≥ 3, respectively. Thus, based on Lilly's integral for a non-uniform grid, the integral of turnover time is convergent for KE -5/3 power law but divergent for the KE -3 power law. These were reported in Lilly's studies. However, based on the Lilly's integral formula for a uniform grid, the integrals of turnover time are divergent for both the KE -3 and -5/3 power laws. *As a result, our analysis indicates the dependence of convergence on not only the slopes of the KE spectra but also on spectral discretization.*

Table B2: A summary of two integrals, representing different discretization, for the turnover time $\tau(k) = \tau_o k^{-(3-s)/2}$. Here, $-s$ represents the power of the KE spectrum, namely $E(k) = C_o k^{-s}$ "CON" and "DIV" represent convergent and divergent, respectively.

	integral	$p = \frac{3-s}{2}$	s	remarks
uniform		CON if $p > 1$	CON if $s < 1$	divergent for

grid	$\frac{1}{k_L}\int_1^{\infty} \tau(k)\, dk$	DIV if $p \leq 1$	**DIV if $\boldsymbol{s} \geq 1$**	both $s = 5/3$ and $s = 3$
non-uniform grid	$\int_1^{\infty} \tau(k) dln(k)$	COV if $p > 0$	COV if $ss < 3$	convergent for $s = 5/3$ but divergent for $s = 3$
		DIV if $p \leq 0$	**DIV if $\boldsymbol{s} \geq 3$**	

References:

Aurell, E.; Boffetta, G.; Crisanti, A.; Paladin, G.; Vulpiani, A. Predictability in Systems with Many Characteristic Times: The Case of Turbulence. Phys. Rev. E 1996, 53, 2337–2349.

Bach, E., et al. (2024). Improved subseasonal prediction of South Asian monsoon rainfall using data-driven forecasts of oscillatory modes. Proceedings of the National Academy of Sciences. Vol. 121 No. 15 e2312573121 https://doi.org/10.1073/pnas.2312573121

Boffetta, G. 2023: Dimensional Transitions in Turbulence: The Effects of Rotation and Stratification. *Atmosphere* 2023, 14, 1688. https://doi.org/10.3390/atmos14111688

Bretherton, C. S., and M. F. Khairoutdinov, 2015: Convective self-aggregation feedbacks in near-global cloud-resolving simulations of an aquaplanet. J. Adv. Model. Earth Syst., 7, 1765–1787, https://doi.org/10.1002/2015MS000499.

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

Castevecchi, D., 2017: On the trail of turbulence. Nature; London Vol. 548, Iss. 7668, (Aug 24, 2017): 382-383.

Charney, J. G., R. G. Fleagle, V. E. Lally, H. Riehl, and D. Q. Wark, 1966: The feasibility of a global observation and analysis experiment. Bull. Amer. Meteor. Soc., 47, 200–220.

De Wit, X.M.; Van Kan, A.; Alexakis, A., 2022: Bistability of the large-scale dynamics in quasi-two-dimensional turbulence. J. Fluid Mech. 2022, 939, R2.

Durran D. and Gingrich, 2014: Atmospheric predictability: Why atmospheric butterflies are not of practical importance. J. Atmos. Sci. 2014, 71, 2476–2478.

Gage, K. S., 1979: Evidence for a $k^{-5/3}$ law inertial range in mesoscale two-dimensional turbulence. J. Atmos. Sci., 36, I 950-1954.

Gage, K. S., and G. D. Nastrom, 1986: Theoretical Interpretation of Atmospheric Wavenumber Spectra of Wind and Temperature Observed by Commercial Aircraft During GASP. J. Atmos. Sci. 43, 7. 729–740.

GARP, 1969: GARP topics. Bull. Amer. Meteor. Soc., 50, 136–141.

Judt, F., 2020: Atmospheric predictability of the tropics, middle latitudes, and polar regions explored through global storm-resolving simulations. J. Atmos. Sci., 77, 257–276, https://doi.org/10.1175/JAS-D-19-0116.1.

Judt, F., 2018: Insights into Atmospheric Predictability through Global Convection-Permitting Model Simulations, Journal of the Atmospheric Sciences, 75(5), 1477-1497.

Kolmogorov, A. N., 1941. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. Dokl. Acad. Sci. USSR, 30, 299–303.

Kolmogorov, A. N., 1962. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds numbers. J. Fluid Mech., 13, 82–85.

Kraichnan, R., 1967. Inertial ranges in two-dimensional turbulence. Phys. Fluids, 10, 1417–1423.

Kraichnan, R. and Montgomery, D., 1980. Two-dimensional turbulence. Rep. Prog. Phys., 43, 547–619.

Lang, S., M. Alexe, M. C.A. Clare, et al: 2024: *AIFS-CRPS: Ensemble Forecasting Using A Model Trained With A Loss Function Based On The Continuous Ranked Probability Score. https://doi.org/10.48550/arXiv.2412.15832*

Larsen, M. F., M. C. Kelley, and K. S. Gage, 1982: Turbulence Spectra in the Upper Troposphere and Lower Stratosphere at Periods Between 2 Hours and 40 Days. J. Atmos. Sci. 39, 1035–1041. https://doi.org/10.1175/1520-0469(1982)039<1035:TSITUT>2.0.CO;2

Leith, C.E., 1971: Atmospheric predictability and two-dimensional turbulence. J. Atmos. Sci. 28, 145–161.

Leith, C.E. and R. H. Kraichnan, 1972: Predictability of turbulent flows. J. Atmos. Sci. 29, 1041–1058.

Lewis, J. 2005: Roots of ensemble forecasting. Mon. Weather. Rev., 133, 1865-1885.

Li, H. et. al. 2024: A machine learning model that outperforms conventional global subseasonal forecast models. https://doi.org/10.21203/rs.3.rs-3776375/v1

Lilly, D. K., 1969. Numerical simulation of two-dimensional turbulence. Phys. Fluid Suppl. II, 12, 240–249.

Lilly, D.K., 1972: Numerical simulation studies of two-dimensional turbulence: II. Stability and predictability studies. Geophys. Fluid Dyn. 1972, 4, 1–28.

Lilly, K.D., 1973: Lectures in Sub-Synoptic Scales of Motions and Two-Dimensional Turbulence Dynamic Meteorology; Morel, P., Ed.; Reidel: Boston, MA, USA, 1973; pp. 353–418.

Lilly, K.D., 1983: Stratified Turbulence and the Mesoscale Variability of the Atmosphere. J. Atmos. Sci. 40, 749-761.

Lilly, D. K. and E. L. Petersen, 1983: Aircraft measurements of atmospheric kinetic energy spectra, Tellus A: Dynamic Meteorology and Oceanography, 35:5, 379-382, DOI: 10.3402/tellusa.v35i5.11448

Lilly, K.D., 1990: Numerical prediction of thunderstorms-has its time come? J. R. Meteorol. Soc. (1990), 116, 779-798.

Lindborg, E., 1999. Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? J. Fluid Mech., 388, 259–288.

Lindborg, E. and Alvelius, K., 2000. The kinetic energy spectrum of the two-dimensional enstrophy turbulence cascade. Phys. Fluids, 12, 945–947.

Lloveras, D.J., L. H. Tierney, and D. Durran, 2022: Mesoscale Predictability in Moist Midlatitude Cyclones Is Not Sensitive to the Slope of the Background Kinetic Energy Spectrum. Journal of the Atmospheric Sciences, 79(1), 119-139.

Lorenz, E.N. 1962: The statistical prediction of solutions of dynamic equations. In Proceedings of the International Symposium on Numerical Weather Prediction, Tokyo, Japan, 7–13 November 1962; pp. 629–635.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. J. Atmos. Sci., 20, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Lorenz, E.N., 1965: A study of the predictability of a 28-variable atmospheric model. Tellus, 17, 321-333.

Lorenz, E. N., 1969a: Studies of atmospheric predictability. [Part 1] [Part 2] [Part 3] [Part 4] Final Report, February, Statistical Forecasting Project. Air Force Research Laboratories, Office of Aerospace Research, USAF, Bedford, MA, 145 pp.

Lorenz, E. N., 1969b: Three approaches to atmospheric predictability. Bull. Amer. Meteor. Soc., 50, 345-351.

Lorenz, E. N., 1969c: Atmospheric predictability as revealed by naturally occurring analogues. J. Atmos. Sci., 26, 636-646.

Lorenz, E. N., 1969d: The predictability of a flow which possesses many scales of motion. Tellus, 21, 19 pp.

Lorenz, E. N., 1969e: How much better can weather prediction become? MIT Technology Review, July/August, 39-49.

Lorenz, E. N., 1969f: The nature of the global circulation of the atmosphere: a present view. The Global Circulation of the Atmosphere, London, Roy. Meteor. Soc., 3-23.

Lorenz, E. N., 1970: Progress report on atmospheric predictability. Never printed. (available from https://eapsweb.mit.edu/sites/default/files/Progress_Report_on_Atmospheric_Predictability_1970.pdf)

Lorenz, E., 1972a: Investigating the predictability of turbulent motion. Statistical Models and Turbulence, Proceedings of symposium held at the University of California, San Diego, July 15-21, 1971, Springer-Verlag, pp. 195-204. https://eapsweb.mit.edu/sites/default/files/Investigating_predictability_1972.pdf

Lorenz, E., 1972b: Low-order models representing realizations of turbulence. J. Fluid Mech., 55, 545-563.

Lorenz, E., 1972: Limits of meteorological predictability. Prepared for the American Meteorological Society, February. (unpublished, available from https://eapsweb.mit.edu/sites/default/files/Limits_of_Meteorological_Predictability_Feb1972.pdf)

Lorenz, E., 1982: Atmospheric predictability experiments with a large numerical model. Tellus, 34, 505-513.

Lorenz, E., 1984: Estimates of atmospheric predictability at medium range. Predictability of Fluid Motions (G. Holloway and B. West, eds.), New York, American Institute of Physics, 133-139.

Lorenz, E. N., 1985: The growth of errors in prediction. In *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, Soc. Italiana di Fisica, Bologna, Italy, 243-265.

Lorenz, E.N. The Essence of Chaos; University of Washington Press: Seattle, WA, USA, 1993; 227p.

Lorenz, E. N., 1996: "Predictability – A problem partly solved". Seminar on Predictability, Vol. I, ECMWF.

Lorenz, E.N., 2006: Predictability – A problem partly solved, in T. Palmer, R. Hagedorn (eds.), Predictability of Weather and Climate). (Cambridge University Press, Cambridge, 2006), pp. 40–58. https://doi.org/10.1017/CBO9780511617652.004

Metais O. and M. Lesieur, 1986: Statistical Predictability of Decaying Turbulence. J. Atmos. Sci., 43 (9), 857-870.

Nastrom. G. D. and Gage. K. S. 1983. A first look at wavenumber spectra from GASP data. Tellus 35A. 383-388.

Reeves, R.W. 2014: Edward Lorenz Revisiting the Limits of Predictability and Their Implications: An Interview From 2007. Bull. Am. Meteorol. Soc. 2014, 95, 681–687.

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.

Pouquet A., and R. Marino, 2013: Geophysical Turbulence and the Duality of the Energy Flow Across Scales. Physical Review Letters. 111, 234501.

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

Pedlosky, J. 1987: Geophysical Fluid Dynamics, 2nd ed.; Springer: New York, NY, USA, 1987; 710p.

Reeves, R. W., 2014: Edward Lorenz Revisiting the Limits of Predictability and Their Implications: An Interview From 2007. BAMS, 681-687.

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.

Ruelle, 1979: *Microscopic Fluctuations and Turbulence. Physics Letters, Vol.72A, No.2, pp. 81-82.*

Saiki, Yoshitaka, and James A. Yorke. 2023. "Can the Flap of a Butterfly's Wings Shift a Tornado into Texas—Without Chaos?" Atmosphere 14, no. 5: 821.

Selz, T., M. Riemer and G. Craig, 2022: The transition from practical to intrinsic predictability of midlatitude weather. J.Atmos. Sci. 79, 2013-2030.

Shen, B.-W., 2024a: Exploring Downscaling in High-Dimensional Lorenz Models Using the Transformer Decoder. Mach. Learn. Knowl. Extr. 2024, 6, 2161 2182. https://doi.org/10.3390/make6040107

Shen, B.-W., 2024b: Revisiting Lorenz's Error Growth Models: Insights and Applications. *Encyclopedia* **2024**, *4*(3), 1134-1146; https://doi.org/10.3390/encyclopedia4030073 (registering DOI)

Shen, B.-W., 2024c: Using Transformers to Investigate Downscaling in the Generalized Lorenz Model. Course Materials for Math340. San Diego State University. Available from ResearchGate: https://doi.org/10.13140/RG.2.2.11127.84642 (a short link to the pdf file on RG https://bit.ly/40cx3Fi)

Shen, B.-W., 2023: A Review of Lorenz's Models from 1960 to 2008. International Journal of Bifurcation and Chaos. Vol. 33, No. 10, 2330024 (2023). https://doi.org/10.1142/S0218127423300240.

Shen, B.-W., 2019: Aggregated Negative Feedback in a Generalized Lorenz Model. International Journal of Bifurcation and Chaos, Vol. 29, No. 3 (2019) 1950037 (20 pages). https://doi.org/10.1142/S0218127419500378

Shen, B.-W., R. A. Pielke Sr., X. Zeng, and X. Zeng, 2024: Exploring the Origin of the Two-Week Predictability Limit: A Revisit of Lorenz's Predictability Studies in the 1960s. *Atmosphere* **2024**, *15*(7), 837; https://doi.org/10.3390/atmos15070837 (registering DOI). Available from ResearchGate: https://doi.org/10.13140/RG.2.2.13760.30727.

Shen, B.-W., R. A. Pielke Sr., and X. Zeng 2023b: 50th Anniversary of the Metaphorical Butterfly Effect since Lorenz (1972): Special Issue on Multistability, Multiscale Predictability, and Sensitivity in Numerical Models. Atmosphere 2023, 14(8), 1279; https://doi.org/10.3390/atmos14081279 (22 journal pages)

Shen, B.-W., R. A. Pielke Sr., X. Zeng, and X. Zeng, 2023a: Lorenz's View on the Predictability Limit. Encyclopedia 2023, 3(3), 887-899; https://doi.org/10.3390/encyclopedia3030063Shen, B.-W., R. A. Pielke Sr., X. Zeng, 2022a: One Saddle Point and Two Types of Sensitivities Within the Lorenz 1963 and 1969 Models. Atmosphere, 13, 753. https://doi.org/10.3390/atmos13050753

Shen, B.-W.*, R. A. Pielke Sr., X. Zeng, J. Cui, S. Faghih-Naini, W. Paxson, R. Atlas, 2022b: Three Kinds of Butterfly Effects Within Lorenz Models. Encyclopedia 2, no. 3: 1250-1259. https://doi.org/10.3390/encyclopedia2030084

Shen, B.-W., Pielke, R.A., Sr., Zeng, X., Baik, J.-J., Faghih-Naini, S., Cui, J., Atlas, R. 2021a: Is weather chaotic? Coexistence of chaos and order within a generalized Lorenz model. Bull. Am. Meteorol. Soc. 2021, 2, E148–E158. Available online: https://journals.ametsoc.org/view/journals/bams/102/1/BAMS-D-19-0165.1.xml

Shen, B.-W., Pielke, R.A., Sr., Zeng, X., Baik, J.-J., Faghih-Naini, S., Cui, J., Atlas, R., Reyes, T.A., 2021b: Is Weather Chaotic? Coexisting Chaotic and Non-Chaotic Attractors within Lorenz Models. In Proceedings of the 13th Chaos International Conference CHAOS 2020, Florence, Italy, 9–12 June 2020; Skiadas, C.H., Dimotikalis, Y., Eds.; Springer Proceedings in Complexity; Springer: Cham, Switzerland, 2021

Shen, B.-W., W.-K. Tao, and B. Green, 2011: Coupling Advanced Modeling and Visualization to Improve High-Impact Tropical Weather Prediction (CAMVis), IEEE Computing in Science and Engineering (CiSE), vol. 13, no. 5, pp. 56-67, Sep./Oct. 2011, doi:10.1109/MCSE.2010.141

Shen, B.-W., W.-K. Tao, and M.-L. Wu, 2010: African Easterly Waves in 30-day High-resolution Global Simulations: A Case Study during the 2006 NAMMA Period. Geophys. Res. Lett., 37, L18803, doi:10.1029/2010GL044355.

Stewart, J., 2014: Calculus: Early Transcendentals, Eighth Edition. Cengage Learning. Boston, MA, USA. 1182 pp.

Tribbia J. J. and D. P. Baumhefner, 2004: Scale Interactions and Atmospheric Predictability: An Updated Perspective. MWR, 132. 703–713. https://doi.org/10.1175/1520-0493(2004)132<0703:SIAAPA>2.0.CO;2

Vallis, G. [2006] Atmospheric and Oceanic Fluid Dynamics. Cambridge. 745 pp.

Zilitinkevich, S. E. Kadantsev, I. Repina, E. Mortikov, and A. Glazunov, 2021: Order out of Chaos: Shifting Paradigm of Convective Turbulence. J. Atmos. Sci., 78, 3925–3932.

Supplementary Materials for the Manuscript Entitled:

“Revisiting Lorenz’s and Lilly’s Empirical Formulas for Predictability Estimates"

by

Bo-Wen Shen[1]*, Roger A. Pielke Sr.[2], and Xubin Zeng[3]

[1]Department of Mathematics and Statistics, San Diego State University, San Diego, California

[2]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado

[3]Department of Hydrology and Atmospheric Science, The University of Arizona, Tucson, Arizona

*bshen@sdsu.edu

https://bwshen.sdsu.edu

(Last Updated: 2024/05/31)

Citation:

Shen, B.-W., R. Pielke, Sr., and X. Zeng, 2023: Supplementary Materials for the Manuscript Entitled “Revisiting Lorenz’s and Lilly’s Empirical Formulas for Predictability Estimates" Available from ResearchGate: https://doi.org/10.13140/RG.2.2.10629.83683

1. Introduction

These supplementary materials were prepared to help examine the convergence of the infinite series and the integrals proposed by Lilly (1972, 1990), and to provide an illustration for the different impact of spectral discretization (that produces different "subsets" of wave modes). Within the supplementary materials, to facilitate discussions, since we focus on the convergent properties of integrals and series, variables are "non-dimensional". For example, *while* "k" *and* "K" *represent non-dimensional and dimensional variables*, respectively, in this supplementary document, "k" is a dimensional variable in the main text.

Section 2 provides a brief review of materials for the integral test, as well as the so-called p-series and differences between the two integrals $\int \frac{1}{x} dx$ and $\int \frac{1}{x} dln(x)$. The latter is the same as $\int \frac{1}{x^2} dx$. Based on Section 2, Section 3 discusses differences for the integrals of turnover time $\tau(k)$ with respect to k and $ln(k)$. Given the specific KE -5/3 spectra that yields $\tau(K) = C_o K^{-\frac{2}{3}}$ (i.e., Eq. 6 in the main text), we show that the different convergent properties between $\int_1^{\infty} (k^{-\frac{2}{3}}) d(k)$ and $\int_1^{\infty} (k^{-\frac{2}{3}}) dln(k)$ can be illustrated by revealing the different properties of $\int \frac{1}{x} dx$ and $\int \frac{1}{x} dln(x)$ that are, respectively, divergent and convergent.

Finally, we suggest that (1) the discrete and continuous forms of Lilly's formulas are consistent and (2) a proper discretization is crucial for determining whether or not the integral of turnover time is convergent, yielding a finite or infinite predictability limit. To help readers, a page break is added at the end of each subsection in Section 3. In Section 4, a succinct approach using a Jacobian is provided to obtain scale factors of $1/k$ in Eq. (18) and $1/k_L$ in Eq. (22) for the non-uniform and uniform discretization, respectively, in the manuscript.

2. The Integral Test and Properties of the p-series

For the specific type of series in this study, an integral test, which is an effective way to test whether a series is convergent, is first reviewed (e.g., Stewart, 2014), as follows:

The Integral Test Suppose f is a continuous, positive, decreasing function on $[1, \infty)$ and let $a_n = f(n)$. Then the series $\sum_{n=1}^{\infty} a_n$ is convergent if and only if the improper integral $\int_1^{\infty} f(x)\,dx$ is convergent. In other words:

(i) If $\int_1^{\infty} f(x)\,dx$ is convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent.

(ii) If $\int_1^{\infty} f(x)\,dx$ is divergent, then $\sum_{n=1}^{\infty} a_n$ is divergent.

The above theorem can be stated as follows: if the integral is convergent (or divergent), the corresponding series is convergent (or divergent). Next, we consider the integral of $1/x^p$ with respect to x. After performing the integral and plugging in the lower and upper bounds, the following properties can be obtained:

$$\int_1^{\infty} \frac{1}{x^p}\,dx \quad \text{is convergent if } p > 1 \text{ and divergent if } p \leqslant 1. \tag{S1}$$

Below, we first illustrate the properties of $\int \frac{1}{x^2} dx$ and $\int \frac{1}{x} dx$ in order to understand the differences between integrals with respect to ln(x) and (x). The two integrals, which are convergent and divergent, respectively, are compared in the following excerpt from Steward (2014). <u>Thus, one may state that the integrals of 1/x with respect to ln(x) and (x) are convergent and divergent, respectively.</u>

Let's compare the result of Example 1 with the example given at the beginning of this section:

$$\int_1^\infty \frac{1}{x^2}\,dx \text{ converges} \qquad \int_1^\infty \frac{1}{x}\,dx \text{ diverges}$$

Geometrically, this says that although the curves $y = 1/x^2$ and $y = 1/x$ look very similar for $x > 0$, the region under $y = 1/x^2$ to the right of $x = 1$ (the shaded region in Figure S1a) has finite area whereas the corresponding region under $y = 1/x$ in Figure S1b has infinite area. Note that both $1/x^2$ and $1/x$ approach 0 as $x \to \infty$ but $1/x^2$ approaches 0 faster than $1/x$. The values of $1/x$ don't decrease fast enough for its integral to have a finite value.

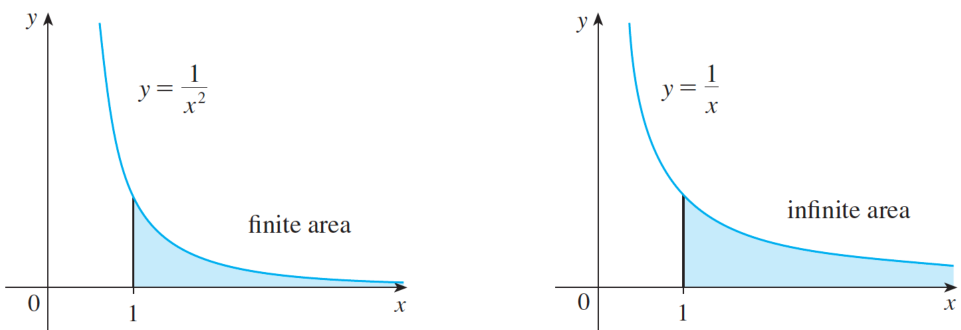

Figure S1a

$\int_1^\infty (1/x^2)\,dx$ converges

Figure S1b

$\int_1^\infty (1/x)\,dx$ diverges

Figure S1: Convergent $\int_1^\infty \frac{1}{x^2}dx$ and Divergent $\int_1^\infty \frac{1}{x}dx$. The former can be written as $\int_1^\infty \frac{1}{x}d(\ln(x))$.

Equation S1 suggests that if an integral is performed with respect to $d\ln(x)$ (i.e., dx/x), we obtain the following:

$$\int_1^\infty \frac{1}{x^p} d(\ln(x)) \text{ is convergent if } p > 0 \text{ and divergent if } p \le 0. \qquad \text{(S2)}$$

A comparison between Eqs. (S1) and (S2) is provided below. The original integrand (x^{-p}) has a power of $-p$. When its integral is performed with respect to dln(k), the power of the "effective" integrand becomes $-(p+1)$. Thus, the effective integrand ($x^{-(p+1)}$) approaches zero faster than the original integrand (x^{-p}). Now, the KE -5/3 power law is considered and the turnover time is given in Eq. 6 of the manuscript (i.e., $\tau(k) = C_0 k^{-\frac{2}{3}}$). After a straightforward computation, we know that two integrals with respect to (k) and ln(k) (i.e., $\int_1^\infty (k^{-\frac{2}{3}}) d(k)$, and $\int_1^\infty \left(k^{-\frac{2}{3}}\right) dln(k)$), are divergent and convergent, respectively. In fact, differences between the two integrals can be illustrated using the above two cases (i.e., $\int \frac{1}{x} dx$ and $\int \frac{1}{x^2} dx$). Namely, while the integral of $\tau(k)$ with respective to k is divergent (as shown in the right panel of list S1), the integral with respect to ln(k), where the effective integrand approaches 0 faster, may be convergent (as shown in the left panel of Figure S1). The above statement is true for $\tau(K) = C_0 K^{-\frac{2}{3}}$. Table S1 provides Eqs. (S1) and (S2) for integrals of turnover time ($\tau(K)$) with KE $E(K) = C_o K^{-m}$.

Table S1: A summary of Eqs. (S1) and (S2) for integrals of turnover time ($\tau(K)$) with a KE $E(K) = C_o K^{-s}$, yielding $\tau(k) = \tau_o k^{-(3-s)/2}$. <u>Since scale invariance is one attribute of power laws</u>, one can rescale the wavenumber to have a non-dimensional wavenumber k, and τ_o is a reference value for turnover time. "CON" and "DIV" represent convergent and divergent, respectively.

		$p = \frac{3-s}{2}$	s	remarks
Eq. (S1)	$\int_1^\infty \tau(k)\, dk$	CON if $p > 1$	CON if $s < 1$	divergent for $s = 5/3$ and $s = 3$
		DIV if $p \leq 1$	**DIV if $s \geq 1$**	
Eq. (S2)	$\int_1^\infty \tau(k)\, dln(k)$	COV if $p > 0$	COV if $s < 3$	convergent for $s = 5/3$ but divergent for $s = 3$
		DIV if $p \leq 0$	**DIV if $s \geq 3$**	

Based on the integral test and properties of the integral of $1/x^p$ with respect to x, we have the following properties for the p-series:

$$\text{The } p\text{-series } \sum_{n=1}^{\infty} \frac{1}{n^p} \text{ is convergent if } p > 1 \text{ and divergent if } p \leqslant 1. \tag{S3}$$

EXAMPLE 3

(a) The series

$$\sum_{n=1}^{\infty} \frac{1}{n^3} = \frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{4^3} + \cdots$$

is convergent because it is a p-series with $p = 3 > 1$.

(b) The series

$$\sum_{n=1}^{\infty} \frac{1}{n^{1/3}} = \sum_{n=1}^{\infty} \frac{1}{\sqrt[3]{n}} = 1 + \frac{1}{\sqrt[3]{2}} + \frac{1}{\sqrt[3]{3}} + \frac{1}{\sqrt[3]{4}} + \cdots$$

is divergent because it is a p-series with $p = \frac{1}{3} < 1$. ■

3. An Illustration on the impact of "discretization" (i.e., the choice of the subset of wave modes)

3.1 A Discrete Form (for a Series $\sum \mathbf{1/n}$)

Based on discussions within the previous section, the integral of $\frac{1}{x}$ with respect to x (i.e., $\int_1^\infty \frac{1}{x} d\,x$) and the series $\sum \frac{1}{n}$ share the same divergent/convergent properties. Below, we first apply this series to illustrate the "impact" of the (grid) discretization (that yields a different subset) and then illustrate that an integral with respect to ln(x) represents an integral with respect to x over a different subset. The sum of the series ($\sum \frac{1}{n}$) on the uniform grid (i.e., $x = n$; referred to as a full set) is expressed in the 1st row of Table A1 in Appendix A (e.g., case S-A).

Compared to the above sequence, we apply a non-uniform grid (i.e., $x = 2^n$) to construct a sequence containing 1, 1/2, 1/4 , 1/8, etc. to form a new series, which is listed in the 2nd row (e.g., case S-B in Appendix A). Since the new series represents a geometric series, it is convergent. Thus, the new convergent series (with a specific subset of numbers for the non-uniform grid) cannot possess divergent properties of the original series over the entire set (for the uniform grid). Namely, the properties of the two series S-A and S-B are different. As discussed in Section 3.2 in the main text, the series in case S-B can be constructed from the convergent integral $\int_1^\infty \frac{1}{x} d\big(log(x)\big)$ with $x = 2^j$.

On the other hand, since the original series in case S-A (i.e., the original integral $\int_1^\infty \frac{1}{x} d\,x$ is divergent), it is possible to construct a new divergent series by selecting a different subset of elements. The 3rd row provides such a choice for a new divergent series (e.g., case S-C in Appendix A).

3.2 A Discrete Form (For an Integral of a Simple Function $1/k$ With Respect To $\log(k)$)

In Sections 3.2 and 3.3 of the main text, we discuss the properties of an integral of $f(k)$ with respect to ln(k) (or log(k)). Here, the function "ln" indicates a natural logrithm function, while "log" represents a logrithm function with a base of 2. As a simple illustration, we begin with the following integral with respect to log(k):

$$\int_{2k_r}^{\infty} \frac{1}{k} d\big(log(k)\big). \qquad (S4)$$

The answer to the above integral is:

$$\frac{1}{\ln(2)}\frac{1}{2k_r} \approx \frac{1.442}{2k_r}. \qquad (S5)$$

Below, as listed in Table S3, Riemann sums are constructed in order to compare the integral $\int \frac{1}{k} dk$ and the above integral $\int \frac{1}{k} d\big(log(k)\big)$ in Eq. (S4). To approximate an integral using a Riemann sum, we need to (1) first build a grid system by performing discretization for an "independent" variable; (2) evaluate the integrand (i.e., the function) at each of the grid points; (3) multiply the functions by the interval of two neighboring grid points, yielding an area for each grid interval; and (4) sum areas for all of the grid integrals. As indicated in Table S3, we use a constant increment of $log(k)$ in Eq. (S4).. A constant of $dlog(k)$ indicates a fixed $\frac{\Delta k}{k}$. That indeed yields a non-uniform grid (for k), as follows:

$$k = 2k_r 2^{y}. \qquad (S6)$$

The above is consistent with the choice in Lilly's formula (e.g., Eq. 1a in the main text) and Lorenz's formula (e.g., the 2nd column of Table 1 in the main text). Here, k_r just represents a reference wavenumber. By plugging Eq. (S6) into Eq. (S4), we have:

$$\int_{0}^{\infty} \frac{1}{2k_r}\frac{dy}{2^y}, \qquad S(7)$$

which yields $\frac{1}{\ln(2)}\frac{1}{2k_r} = \frac{1.44}{2k_r}$ (which is the same as the above answer in Eq. S5). Next, we discretize y into $y = n\Delta y$, where n is an integer. Thus, we have:

$$k_n = 2k_r 2^{n\Delta y}. \qquad (S8)$$

The above leads to $\Delta k_n = k_{n+1} - k_n = k_n$, yielding $\frac{\Delta k_n}{k_n} = 1$ (i.e., a constant of 1 for $d\big(log(k)\big)$ in Eq. S4), as shown in Figure S2. Thus, when Eq. S(8) with $\Delta y = 1$ is applied, Eq. S(4) is approximated by:

$$\frac{1}{2k_r}\sum_{n=0}^{\infty}\frac{1}{2^n}, \qquad (S9)$$

which is approximately $\frac{2}{2k_r}$, which is close to Eq. (S5). More importantly, it is a finite number. Note that when $\Delta y = 3/2$ (which is larger than $\Delta y = 1$) we are able to obtain a better solution with a smaller error. However, for the integral with respect to log(k), convergent properties of the Riemann sum with different "resolutions" (i.e., different values of Δy) are beyond the scope of this study.

Table S3: Integrals of $1/k$ with respect to k and $\log(k)$, which are, respectively, proportional to $\sum_{n=1}\frac{1}{n}$ and $\sum_{n=1}\frac{1}{2^n}$.

Integral	$\int_{2k_r}^{\infty}\frac{1}{k}dk$	$\int_{2k_r}^{\infty}\frac{1}{k}d\big(log(k)\big)$
Riemann sum	$\sum height \times width$	$\sum height \times width$
$(height, width)$	$(\frac{1}{k}, dk)$	$(\frac{1}{k}, dlog(k))$
discretization	dk =constant	$dlog(k)$ =constant
Property	divergent	convergent

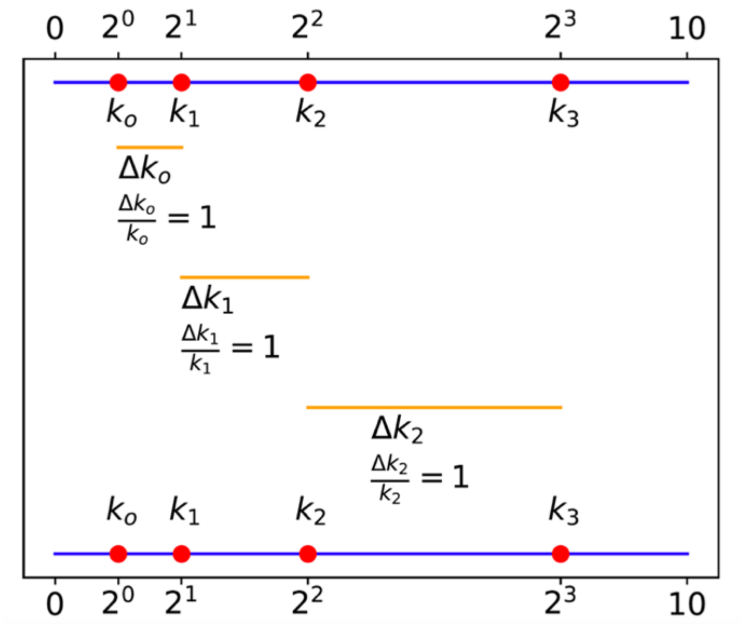

Figure S2: A grid system for an integral of $f(k)$ with respect to $log(k)$. To have a uniform grid with $d\big(log(k)\big) = constant$, we may choose $k_n = k_o 2^n$. Here, $k_o = 2k_r$. Such a choice yields $\Delta k_n = k_{n+1} - k_n = k_n$ and, thus, $\frac{\Delta k_n}{k_n} = 1$. Please note that Δk_n is not a constant.

Below, we provided an example of the "non-unform" grids from the study by Leith and Kraichhan (1972).

We choose discrete wavenumbers $k_l = 2^{l/F}$, $l = 0, 1, \ldots, L$, whose logarithm is distributed uniformly. The integer $F \geqslant 2$ measures the fineness of the mesh by specifying the number of mesh intervals per factor of 2. The statistical functions are defined at these meshpoints as $U_l = U(k_l)$ and $\Delta_l = \Delta(k_l)$. By a further subdivision at points $k_{l+\frac{1}{2}} = 2^{(l+\frac{1}{2})/F}$ we associate with each wavenumber k_l a surrounding interval $k_{l-\frac{1}{2}} < k < k_{l+\frac{1}{2}}$. The length of such an interval is βk_l with the constant $\beta = 2^{(\frac{1}{2})/F} - 2^{-(\frac{1}{2})/F}$. An integral of any function $f(k)$ over k will then be estimated by the rectangular approximation

$$\int f(k)dk \approx \beta \sum_l f_l k_l,$$

where $f_l = f(k_l)$.

Here, we first provide derivations for the last formula in the above image. Then, we show that βk_l represents the Jacobian.

Consider:

$$dk \approx \Delta k = k_{l+\frac{1}{2}} - k_{l-\frac{1}{2}} = 2^{(l+\frac{1}{2})/F} - 2^{(l-\frac{1}{2})/F} = 2^{l/F}(2^{(\frac{1}{2})/F} - 2^{(\frac{-1}{2})/F}) = \beta k_l,$$

here $\beta = (2^{(\frac{1}{2})/F} - 2^{(\frac{-1}{2})/F})$. Thus, we have:

$$\int f(k)dk \approx \sum_l f(k_l)\Delta k = \sum_l f(k_l)\beta k_l. \qquad (S10a)$$

Equation (S10a) is the same as that in the image. Secondly, we consider the change of variables, $k_l = 2^{l/F}$, and compute its Jacobian (J) as follows:

$$J = \frac{dk_l}{dl} = \frac{ln(2)}{F}k_l.$$

Then, we have the integral for the new variable as follows:

$$\int f(k)dk \approx \int f(l)\,|\,J\,|dl = \int f(l)\frac{ln(2)}{F}k_l\,dl \qquad (S10b).$$

Equation (S10b) and (S10a) are the same, as discissed using the Jacobian below. Lastly, we show $\beta\, k_l \approx\ |\,J\,|$ as follows. Consdier the following Taylor series approximation:

$$2^{(\frac{1}{2F})} \approx\ 1\ +\ ln(2)/2F,$$

and

$$2^{(\frac{-1}{2F})} \approx\ 1 -\ ln(2)/2F,$$

yielding

$$\beta \approx 2^{(\frac{1}{2F})} - 2^{(\frac{-1}{2F})}\ \approx\ ln(2)/F = J/k_l.$$

The above demonstrates that, as a result of the variable change, a scaling factor in the form of a Jacobian should be taken into account. Specifically, for the integral with respect to k, a scaling factor of k (i.e., k_l) appears on the right-hand side in Eqs. (S10a, b). In contrast, when considering an integral with respect to $ln(k)$ (e.g., in Lilly's formula), the factor of $1/k$ associated with $ln(k)$ is cancelled out by the Jacobian k following the variable change.

3.3 A Discrete Form (For an Integral of a General Function $\boldsymbol{f(k)}$ With Respect To $\boldsymbol{log(k)}$)

Below, we consider the integral of $f(k)$ with respect to $ln(k)$ and $log(k)$. In Table S4, the case "I-base" represents a Rieman sum of an integral with respect to k (i.e., $\int f(k)dk$). When $\Delta k = 1$, the case I-base yields the series in case I-A in Table S4 that represents a Rieman sum of the integral with respect to k within a "full" set for discrete k. Below, we first show that the two series in cases I-B and I-C, respectively, represent integrals with respect to log(k) and ln(k). From cases I-A and I-B, we then show that the series in case I-B represents a Rieman sum within a "subset" of k as compared to the series in case I-A.

As a result of relatively simplicity, we first begin with case I-C. Let's consider an integral with respect to ln(k):

$$\int f(k)d\big(ln(k)\big), \qquad (S11a)$$

yielding:

$$\int \frac{f(k)}{k}dk. \qquad (S11b)$$

We now consider the following subset to have a fixed $d\big(ln(k)\big)$, which is shown in case I-C in Table S4:

$$k = e^m k_o. \qquad (S12)$$

Here, k_o is a reference wavenumber and can be the smallest wavenumber. Note that the integral with respect to ln(k) in Eq. (S11a) was applied in Lilly (1990). The integral is (mathematically) consistent with the discrete form in Lilly (1972). The choice in Eq. (S12) yields a collection of selected wavenumbers $k_o, ek_o, e^2k_o, \cdots e^n k_o, etc.$, which may not be "physically" intuitive. Here, no attempt is made to discuss the impact of fractal dimensions. A "physically" intuitive choice of $k = 2^m k_o$ is later discussed. From the above, we have:

$$\frac{dk}{dm} = e^m k_o,$$

and, thus,

$$dk = e^m k_o dm. \qquad (S13)$$

Plugging Eqs. (S12) and S(13) into Eq. (S11), we obtain:

$$\int f(e^m k_o)dm. \qquad (S14)$$

When we select $m = n$ and when n presents an integer, we obtain the series for case I-C. Namely, the series in case I-C represents an integral with respect to ln(k) (e.g., Eq. S11a).

Table S4: Integrals of $f(k)$ with respect to k, log (k), or $\ln(k)$ (i.e., over a full or a subset of k).

I-base	Δk	$2\Delta k$	$3\Delta k$	$4\Delta k$	$5\Delta k$	$6\Delta k$	$7\Delta k$	$8\Delta k$	$\cdots$	$n\Delta k$	$\cdots$	$\sum_{n=1} f(n\Delta k)\,\Delta k$
I-A ($\Delta k = 1$)	1	2	3	4	5	6	7	8	$\cdots$	n	$\cdots$	$\sum_{n=1} f(n)$
I-B		2		4				8	$\cdots$		$\cdots$	$\sum_{n=1} f(2^n)$
I-C		$\mathrm{e}k_o$					$\mathrm{e}^2 k_o$		$\cdots$			$\sum_{n=1} f(e^n k_o)$

We can similarly show that the series in case I-B with the specific subset of wave modes represents an integral with respect to log(k). Mathematical details are provided below. We first consider the following integral with respect to log(k):

$$\int f(k)d\big(log(k)\big), \qquad (S15)$$

yielding:

$$\int \frac{f(k)}{kln(2)}dk. \qquad (S16)$$

We now consider the following subset, listed in case I-B of Table S4:

$$k = 2^m k_o. \qquad (S17)$$

From the above, we have:

$$\frac{dk}{dm} = \ln(2)2^m k_o,$$

and, thus,

$$dk = \ln(2)2^m k_o dm. \qquad (S18)$$

Plugging Eqs. (S17) and (S18) into Eq. (S16), we obtain:

$$\int f(2^m k_o) dm. \qquad (S19)$$

When we select $m = n$ and when n represents an integer, Eq. (S19) yields the series in case I-B. Namely, the series in case I-B with a specific subset of wave modes represents an integral with respect to log(k) (e.g., Eq. S15). The series in case I-B represents the sum over a subset of k as compared to the series in case I-A.

From the above two cases in Eqs. (S11a) and (S15), we may consider a more general case, as follows:

$$\int f(k) d(\log_b(k)). \qquad (S20)$$

Here, the logrithm function has a base of b. To constrcut a grid for a constant increment of $\log_b(k)$, we may choose:

$$k = b^m k_o. \qquad (S21).$$

The general convergent properties of Eq. (S20), with the choice of Eq. (S21), is beyond the scope of this study.

4. Change of Variables in Single and Double Integrals

In fact, the scale factors of $1/k$ in Eq. (18) and $1/k_L$ in Eq. (22) can be simply obtained using the concept of Jacobian for changes of variables. Below, we first review the concept of changes of variables in single and double integrals and then apply the concept to obtain the scale factors of $1/k$ and $1/k_L$.

For a change of variable in a single integral, Formula 2 indicates a scale factor of dx/du, as shown below (Stewart, 2014).

> In one-dimensional calculus we often use a change of variable (a substitution) to simplify an integral. By reversing the roles of x and u, we can write the Substitution Rule (5.5.6) as
>
> **1** $$\int_a^b f(x)\,dx = \int_c^d f(g(u))\,g'(u)\,du$$
>
> where $x = g(u)$ and $a = g(c)$, $b = g(d)$. Another way of writing Formula 1 is as follows:
>
> **2** $$\int_a^b f(x)\,dx = \int_c^d f(x(u))\,\frac{dx}{du}\,du$$

To provide an additional illustration, the change of variables in a double integral is listed below.

> **9 Change of Variables in a Double Integral** Suppose that T is a C^1 transformation whose Jacobian is nonzero and that T maps a region S in the uv-plane onto a region R in the xy-plane. Suppose that f is continuous on R and that R and S are type I or type II plane regions. Suppose also that T is one-to-one, except perhaps on the boundary of S. Then
>
> $$\iint_R f(x, y)\,dA = \iint_S f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du\,dv$$

A scale factor is indicated by the Jacobian which is defined as follows:

7 Definition The **Jacobian** of the transformation T given by $x = g(u, v)$ and $y = h(u, v)$ is

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}$$

Based on the above Formula 2, Table S5 is constructed as follows:

Step 1 Series	Step 2	Step 3 (Jacobian)	Step 4 Integral
$\lim_{N\to\infty}\sum_{j=0}^{N}(F(e^j k_L))\Delta j$ $k = e^j k_L$	$\int F(e^j k_L)dj$	$k = e^j k_L$ $j = \ln\left(\frac{k}{k_L}\right)$ $\frac{dj}{dk} = \frac{1}{k}$	$\int F(k)\frac{1}{k}dk$
$\lim_{N\to\infty}\sum_{j=0}^{N}(F(jk_L))\Delta j$ $k = jk_L$	$\int F(jk_L)dj$	$k = jk_L$ $j = \frac{k}{k_L}$ $\frac{dj}{dk} = \frac{1}{k_L}$	$\int F(k)\frac{1}{k_L}dk$

References:

- Leith, C.E. and R. H. Kraichnan, 1972: Predictability of turbulent flows. J. Atmos. Sci. 29, 1041–1058.
- Stewart, J., 2014: Calculus: Early Transcendentals, Eighth Edition. Cengage Learning. Boston, MA, USA. 1182 pp.