

THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Using theoretical physics to address complex dynamical problems in turbulence, geosciences and biology

Daniel Jack Brener



Doctor of Philosophy
The University of Edinburgh
2025

Abstract

Real-world systems are generally nonlinear, exhibiting complexity across a range of physical and temporal scales. In this thesis I tackle a series of real-world problems and questions which span the breadth of these different scales, from the microbiological to interplanetary. This is possible when each system is viewed from the perspective of theoretical physics as being reducible to the level of a dynamical system. Early on in my research, I realised that these approaches could be applied across many fields and chose to work in three of them. Turbulence remains one of the greatest challenges remaining in classical physics, and arises in our first dynamical system, the Navier-Stokes equations (NSEs). I investigate higher-dimensional versions of these equations using Direct Numerical Simulation studies to search for evidence of critical phenomena, using chaos as a measure. Using this together with a closure-approximated model, I investigate how increasing spatial dimensions appears to lead to a reduction in chaos, which might tentatively suggest a critical dimension of six for these equations. From the problem of turbulence, I turn to a related problem in ensemble numerical weather prediction, the signal-to-noise paradox. The paradox is that current ensemble systems seem to predict reality better than they predict themselves. By applying ergodic theory to ensemble forecasting, I show that using the ensemble mean as our best forecast of observations amounts to interpreting it as the most likely phase-space trajectory, which relies on the ergodic theorem. I argue that this fails in certain cases due to evidence of multi-modality, which can break the ergodic theorem, creating the paradox. Moving up a scale, the next problem I address is the practical challenge of contamination when trying to detect microbiology in Earth's upper atmosphere. Leveraging advances in aerospace technology, such as CubeSats and rocket-borne samplers, this research proposes a new technique called relative-velocity sampling, for capturing large particles in the understudied mesosphere and lower thermosphere. This technique reduces the contamination challenge to a simple one-dimensional dynamical system, obeying the NSEs. My ultimate problem lies on the largest scales where I apply the theory of island biogeography to interplanetary scales for planetary

protection. This theory models population dynamics for inter-island populations and has been suggested as applicable to interplanetary systems. This is used to show that although such an equilibrium theory generally breaks down when applied to interplanetary scales, the mean-time to extinction resulting from the combined effects of growth and death rates can be quantified. This is used to challenge the probabilistic model of planetary protection and suggest how the mean-time to extinction can instead be used to assess colonisation risk. The broader applicability of island biogeography to considering biotic transfer at the interplanetary scale is considered. This thesis concludes with a synthesis of the problems tackled, illustrating the capability of theoretical physics to advance understanding in different fields underpinned by dynamical systems, and the role of interdisciplinary research.

Lay Summary

This thesis explores how theoretical physics can deepen our understanding of complex real-world systems, ranging from microscopic organisms to planetary ecosystems. By understanding these diverse systems as related forms of dynamical systems—models that describe how things change over time—the research addresses challenges in fluid turbulence, weather forecasting, upper-atmosphere microbiology, and planetary protection. In the study of turbulence, it suggests that chaos in fluids decreases in higher spatial dimensions. This could change how we understand normal three-dimensional turbulence. Regarding weather prediction, the approach offers an explanation of a paradox where ensemble forecasts predict actual weather better than they predict themselves. This reveals that certain mathematical assumptions may fail for ensemble weather forecasting. The thesis also introduces a new method for collecting microorganisms from Earth's upper atmosphere using advanced aerospace technology, solving longstanding contamination issues. Finally, by applying ecological theories to space exploration, a new perspective is offered on how risk should be assessed for contaminating other planets with Earth-based life. Overall, the work demonstrates the power of theoretical physics to solve diverse problems and underscores the importance of interdisciplinary research.

Declaration

I declare that this thesis has been composed by myself, and that this body of work has not been submitted for any other degree or professional qualification. I further confirm that the work submitted is my own except where it is formed of publications produced through collaborative research. My own contribution to these published works and those of my co-authors are described below for each chapter.

The work presented in Chapter 2 shows the results of simulations which I, Daniel Jack Brener, designed, ran, maintained, and processed, and then analysed using my own code. The EddyBurgh simulation code was originally written by Dr Samuel R. Yoffe in 2012, and was further developed by Dr Richard Ho, Dr Daniel Clark and Dr Andrés Armúa. I wrote and tested, with support from Dr Andrés Armúa and Dr Daniel Clark, the function to test the invariant in six dimensions. With support from Dr Daniel Clark and the ARCHER2 Support Team, I further developed the code to use memory more efficiently. The EDQNM code was originally developed by Dr Andrés Armúa and Dr Daniel Clark. I used this code independently on the Edinburgh Eddie facility to run my own simulations. I wrote my own analysis code to process the output and plot the results. Arjun Berera (former supervisor) provided supervision for this work.

The publication in Chapter 3 originally appeared in Atmospheric Science Letters by me, Daniel J. Brener. I completed all aspects of this work independently.

The publication in Chapter 4 originally appeared in Astrobiology by Arjun Berera, Daniel J. Brener and Charles S. Cockell (supervisor). I designed the detector, did the calculations, and produced and led the initial draft. Arjun Berera and Charles S. Cockell provided supervision for this publication. Authorship list for this paper was alphabetical by surname. All authors contributed to interpretation, editing, response to reviewers and the final formatting of the publication.

The publication in Chapter 5 originally appeared in Journal of the Royal Society Interface by Daniel J. Brener and Charles S. Cockell (supervisor). I, Daniel J. Brener designed the study, did the calculations, and produced the initial draft. Charles S. Cockell provided supervision for this publication. All authors contributed to interpretation, editing, response to reviewers and the final formatting of the publication.

(*Daniel Jack Brener, 2025*)

Acknowledgements

My sincere gratitude and thanks to Professor Charles Cockell, for his invaluable guidance, support, and encouragement. We established a conversation on biogeography in the first year of my doctoral studies, and I am most grateful that he agreed to take me under his supervision for the final year of my PhD, stepping in during a challenging period to provide the support and mentorship I needed to bring this thesis to fruition.

Professor Jorge Peñarrubia, as my second supervisor, has been wonderfully encouraging and supportive of my theoretical work, especially with regard to ergodic theory and statistical physics. I have enjoyed many long meetings on the side of Blackford Hill, imagining complex Brownian processes and talking mathematics into the cold air.

I am grateful for the support of my supervisory team, Dr Anton Ilderton and Professor Luigi del Debbio, as well as Professor James Dunlop and Dr Mark Williams. I also acknowledge Professor Arjun Berera for his guidance and support during the early stages of my research career as an undergraduate student, and into my doctoral studies.

Professor Chris Stock, the Glenn Gould of condensed matter physics, is a friend, teacher and collaborator without whom my time in Edinburgh would not have been as bright, colourful, fun and interesting. I shall miss dropping in to put the world to rights. It was also a huge privilege and an honour to teach quantum mechanics with you. Which leads me to the many hundreds of quantum mechanics students I have taught over the past three years; through their curiosity I have had the privilege to deepen my understanding.

My esteemed colleagues across the School of Physics and Astronomy; in Particle Physics Theory Dr Andrés Armúa, Dr Daniel Clark, Dr Richard Ho with whom I had the pleasure to work with directly, and Sam Teale, William Lindved, Dogan Akpinar, Giorgio Frangi and Mile Vrbica of JCMB 4301, for distracting me with their far superior and interesting work in actual particle physics; Mr Neil Corsie, Health and Safety Manager for his kindness and professional support and the Servitors of the James Clerk Maxwell Building for allowing me to use blackboards in peculiar rooms at odd hours. I would also like to acknowledge and thank Dr Armúa for his kind support, encouragement and enthusiasm, as well as for being a most inspiring figure, by completing his doctorate whilst raising a wonderful family.

My journey to Edinburgh as a physicist began with my cousins Professor Alan Miller and Professor Malcolm Walkinshaw, who invited me aged 14, to visit their laboratories. It seems particularly apt that this thesis includes elements from both of their respective fields of physics and biology.

I am grateful to the UK Science and Technology Facilities Council for funding my doctoral research. The work presented in chapter 2 used the ARCHER2 UK National Supercomputing Service via the Director's Time budget and was also partially funded by the UK Engineering and Physical Sciences Research Council under an Access to HPC grant. During my studies I received support from the following individuals: Professor Mark Parsons, Director of the Edinburgh Parallel Computing Centre (EPCC) and responsible for ARCHER2, Dr David Henty of EPCC, and Dr Douglas Shanks of Hewlett Packard Enterprise. This support included specific work for an ambitious turbulence simulation project not presented in this thesis.

I would like to thank all my teachers and mentors over the years, and specifically Rosemary Moore, Olly and Lee-Anne Wedgewood, Dave Jeffers, Steven Whyte, Philip Dobson, Amanda Boyd and Dr David Homan. I also wish to acknowledge the following mentors in atmospheric science specifically: Dr Gill Ryall who, as Executive Head of Observations at the UK Met Office, saw something in me and invited me to visit the Office aged just 11, setting a goal which would see me through a decade; Professor Adam Scaife who kindly mentored me from my earliest interests in Sudden Stratospheric Warmings to introducing me to and encouraging me to tackle, the signal-to-noise paradox problem; Dr Juan Manuell Castillo and Dr José Rodríguez for their kindness and nurturing of my scientific curiosity and Spanish speaking at the Met Office.

None of the work here enclosed would have been possible without a strong, dear set of friends. Cheers to the "slightly perturbed gang", without whom I would not have survived my undergraduate studies, and thence this thesis would not have been possible. Dear friends Berenice Benjelloun, Scott Blum-Woodland, Tommaso Bruggi, Alexandra Burgar, Егор Смирнов, David Harris and Universino Sousa, Richie Hofmann, Daniel Isengart and Filip Noterdaeme, Benny Nemer, Dominic Sorrell, Alexander Tully and Sue Turner – you all know why you are inscribed here.

Last, but by no means least, my parents, for always going above and beyond, for believing in me where all others did not, and for always being there.

Contents

Abstract i

Lay Summary iii

Declaration iv

Acknowledgements v

Contents vii

List of Figures x

List of Tables xii

1 Introduction 1

1.1 Physical motivation 3

1.1.1 Turbulence and the Navier-Stokes equations 3

1.1.2 Numerical weather prediction 10

1.1.3 Planetary biospheres 12

1.2 Theoretical underpinning – dynamical systems theory 14

1.2.1 Chaos theory 17

1.2.2 Ergodic theory 18

1.2.3 Some specific forms of dynamical systems 20

2 Searching for a critical dimension in homogeneous isotropic turbulence using chaos measurements 23

2.1 Introduction 23

2.2 Methodology - simulations 25

2.2.1 Direct Numerical Simulation (DNS) 26

2.2.2 Eddy Damped Quasi-Normal Markovian (EDQNM) model.. 31

2.3 Results 34

2.4 Discussion 39

2.5 Conclusion 42

3 A hypothesis on ergodicity and the signal-to-noise paradox 43

3.1 Introduction 43

3.2 Research publication 43

3.3 Conclusion 53

4 Detecting microbiology in the upper atmosphere: relative-velocity filtered sampling 54

4.1 Introduction 54

4.2 Research publication 54

4.3 Conclusion 62

5 Rethinking planetary protection: an island biogeographical analysis 63

5.1 Introduction 63

5.2 Research publication 64

5.3 Conclusion 74

6 Conclusions 76

6.1 Theoretical physics as an approach..... 76

6.2 Role of dynamical systems theory 78

6.3 Future work..... 79

A Feynman's Letter 82

B DNS simulation parameters 84

List of Symbols 86

List of Acronyms 87

Glossary 88

Bibliography 94

List of Figures

1.1 Painting by the artist Berenice Benjelloun, 1969 titled "Con-joining fields". Photograph taken by the author and reproduced with kind permission of the artist. 3

1.2 Richardson cascade illustration. Big swirls break up into littler swirls until the smallest scales. 5

1.3 Illustration of the Kolmogorov cascade. 9

1.4 Illustration of the North Atlantic Oscillation (NAO) pressure anomalies. Left is negative, right is positive. H and L denote regions of higher than average and lower than average pressure, respectively. Note this does not necessarily mean regions of absolute high and low pressure. Designed by the author using imagery from NASA's Earth Observatory team, led by Reto Stöckli [114]. 11

1.5 Chapter order and scaling in this thesis. 16

1.6 Illustration of chaos in one-dimension (vertical axis). Divergence of initially close trajectories with time (horizontal axis). 17

2.1 Two different examples of graphical output used to assistance in the semi-automatic plus verification of the EDQNM simulation data to obtain Lyapunov exponents. 32

2.2 Lyapunov exponent λT versus integral scale Reynolds number, Re for multiple dimensions at low Re measured from DNS of NSEs on left and EDQNM on right. Residuals are for the linear regressions fitted to DNS with the error bars indicating the magnitude of the fluctuations in the FTLE. Data points are indicated by blue squares, orange circles, green crosses and pink triangles for 3, 4, 5 and 6 dimensions respectively. Fits to each of these are the same colour but dashed, dotted, dot-dashed and solid for 3, 4, 5 and 6 dimensions respectively. 35

2.3 Velocity derivative skewness versus Taylor Reynolds number, Re_λ for three, four, five and six dimensions. 38

2.4 Variation of energy normalised at the microscale, with integral scale Reynolds number in three, four, five and six dimensions. . . 39

5.1 *Ship ready for return.* Painting by Chesley Bonestell, reproduced from *The Conquest of Space*, The Readers Union, London, 1952. . 75

List of Tables

2.1 EDQNM simulation fixed parameters used for each dimension, from [5]. 34

B.1 $d = 5$ 84

B.2 $d = 6$ 84

B.3 $d = 3$ 85

B.4 $d = 4$ 85

Chapter 1

Introduction

The consequences of our actions are always so complicated, so diverse, that predicting the future is a very difficult business indeed. – Dumbledore [103]

Looking out over the vast vibrant landscape of the Cairngorm mountains in the Highlands of Scotland I feel myself immersed in the beauty of our planet. I'm out hiking in the wilderness for a few days to get away from the urban city drawl. At every scale, from the methanophile bacteria in the peat bog under my feet to the weather circulation patterns and winds blowing over, through to the dark night sky, I am confounded by complexity. Wherever I look there are complex, nonlinear and chaotic phenomena. How multiple systems connect and the relationships between them is something that has fascinated me. My thesis addresses a set of these diverse multidisciplinary problems ranging from turbulence in fluid dynamics to microbiological sampling in the Earth's upper atmosphere and space exploration. At first glance, these challenges might appear to belong to distinct disciplines. However, as I hope it will become clear, when one thinks about each of them in a particular way, they can become related and thus tackled using similar tools.

The mathematics underlying this thesis is dynamical systems theory, but this thesis is not *about* dynamical systems theory specifically. Rather, I wish to illustrate more broadly how theoretical approaches in general can be unifying for seemingly disparate fields which involve complex dynamical problems. One should understand theoretical physics as a way of thinking, a way of approaching problems and seeing the world, as opposed to an isolated discipline that tackles problems which are described according to fashion, as "fundamental".

I was influenced in this way of seeing theoretical physics by my scientific hero, the late Professor Richard Feynman. In reading his published correspondence, I came across a letter[1] to his former doctoral student Dr Koichi Mano [38]. Dr Mano had expressed his embarrassment that he was working on a "humble and down-to-earth" problem related to electromagnetic waves in turbulent atmospheres,

[1]Reproduced in the Appendix A as it conveys the philosophical essence of my thesis.

and felt unsuccessful as a theoretical physicist for not working on fundamental problems. Feynman replied,

> I have worked on innumerable problems that you would call humble...for example, the theory of turbulence (I have spent several years on it without success)... No problem is too small or too trivial if we can really do something about it.

This is pertinent to this thesis as turbulence underlies many of the problems tackled herein, and the final sentence is the spirit in which I work.

It is also worth remarking on the context of this thesis in relation to previous interdisciplinary theoretical physics theses and the current growing body of research which makes these kinds of connections. The closest example is the thesis of Dr Joel Tenenbaum in which he illustrates the use of theoretical physics approaches, particularly statistical ones, to three different disparate fields: seismic physics, econophysics and sociophysics [117]. My thesis, however, is unique in terms of the fields it covers and its focus on the dynamical systems point of view. It is only really in the last decade that this latter approach has begun to be used outside theoretical condensed matter physics. This is partly due to computational advances which allow for data-driven studies as these can thoroughly map out phase-spaces for complex systems [31]. There is a long tradition of the kind of "dynamical thinking" I advocate for in geosciences, but this has tended to focus on either non-linear behaviour [44] or the study of geophysical extreme events [36, 109].

Due to the interdisciplinary nature of the work enclosed and the credo I have just outlined, I have written the introductory chapter to this thesis assuming as little background knowledge on the specific fields and problems as possible. This includes the mathematics necessary to understand the work in the later chapters. Some of the mathematics and simulations exist beyond regular three-dimensional imagination and intuition, e.g. higher-dimensional solutions to the Navier-Stokes equations. When I encountered the painting *Con-joining Fields* (1969) by Berenice Benjelloun, I immediately thought of turbulence in higher dimensions, reproduced in Figure 1.1. Although there is no pleasant way to fully imagine these abstract spaces, I believe art sometimes offers a way to evoke the feeling of such phenomena better than simply plotting three-dimensional representations of six dimensional fields and saying "well, it must kind of look like this".

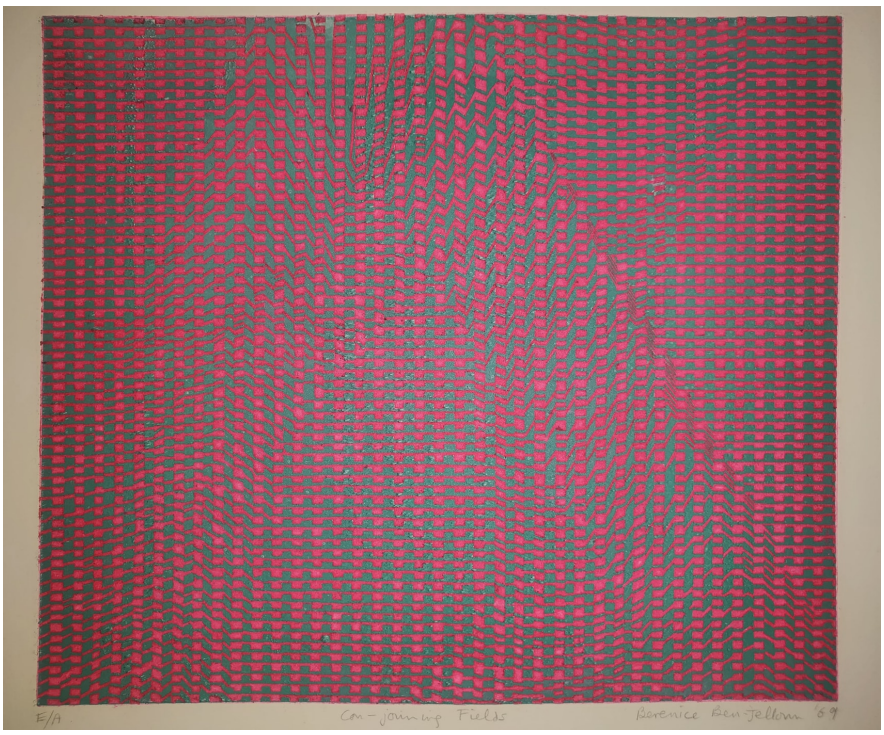


Figure 1.1 *Painting by the artist Berenice Benjelloun, 1969 titled "Con-joining fields". Photograph taken by the author and reproduced with kind permission of the artist.*

1.1 Physical motivation

In this section I will give an overview and introduce each of the physical domains, systems and associated problems tackled by the later research chapters.

1.1.1 Turbulence and the Navier-Stokes equations

The Navier-Stokes equations are an effective *field theory* (see section 1.2.3) for a infinite boundary-free incompressible fluid, satisfying spatial dimensions $\forall d \geq 2$, and in three-dimensional Eulerian coordinates are given by

$$\begin{aligned} \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} &= -\frac{1}{\rho} \boldsymbol{\nabla} P + \nu \nabla^2 \boldsymbol{u} \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} &= 0 \end{aligned} \tag{1.1}$$

where $\boldsymbol{u}(\boldsymbol{x}, t)$ is the velocity field, $P(\boldsymbol{x}, t)$ the pressure field and ν is the kinematic viscosity. Physically, the terms representing inertial forces are on the left hand side, the viscous term is the final term on the right hand side.

Reynolds number

Using the ratio of the inertial terms to the viscous terms, and then by introducing a characteristic or typical velocity, U, length, L and time scale T provides us with the following argument,

$$\frac{|\boldsymbol{u}\cdot\nabla\boldsymbol{u}|}{\nu\,|\nabla^2\boldsymbol{u}|}\sim\frac{U^2}{L}\left(\frac{\nu U}{L^2}\right)^{-1}=\frac{UL}{\nu}\equiv Re, \tag{1.2}$$

hence defining the *Reynolds number*, Re. The greater the Reynolds number the more turbulent the fluid is said to be, and hence this number characterises turbulence. Depending on the nature of the fluid and its environment, there is usually a range of Re for which the fluid remains laminar, before the transition to turbulence. In examining the Reynolds number we can thus re-write the Navier-Stokes equations in the following non-dimensional form,

$$\begin{aligned}\frac{\partial\boldsymbol{u}^*}{\partial t^*}+\boldsymbol{u}^*\cdot\nabla^*\boldsymbol{u}^*&=-\nabla^*P^*+\frac{1}{Re}\nabla^{2*}\boldsymbol{u}^*\\ \nabla^*\cdot\boldsymbol{u}^*&=0,\end{aligned} \tag{1.3}$$

where * denotes the non-dimensional definitions $\boldsymbol{u}^*=\frac{\boldsymbol{u}}{U}$, $P^*=\frac{P}{\rho U^2}$ and $t^*=t\frac{U}{L}$.

Note how Re is now the only *parameter* in these equations; meaning that without any additional forcing terms, any two flows with the same Reynolds number and the same initial/boundary conditions (i.e. geometry), should be equivalent. This is known as *Reynolds' Law of Similarity*.

Homogeneous isotropic turbulence (HIT)

In Chapter 2 of this thesis I will study homogeneous isotropic turbulence (HIT) in simulations. HIT is a carefully defined form of turbulence in which the velocity field probability density function (PDF) is

1. invariant under spatial translation and rotation, and
2. satisfies the property of ergodicity (see section 1.2.2), typically through being stationary.

Due to these constraints, HIT is a fluid which cannot have boundaries and their associated effects so in that sense it is not "practical" turbulence as we might

know it, in a pipe for example. However, HIT is very useful precisely because of its well-defined qualities, removing extraneous related turbulent phenomena except for "pure" turbulence itself.

One way of imagining HIT is to try to picture someone vigorously and quickly shaking a sieve through a trough of water, alternating directions very quickly and randomly. This was in fact how some of the earliest experiments were done on HIT, before the advent of computers [42]. One can also realise by picturing different kinds of flows that it is possible to have turbulent flows which are non-isotropic (e.g. the flow behind an asymmetric airfoil - turbulence changes with rotation) and or non-homogeneous (e.g. frying sausages in a real saucepan - no real saucepan actually heats everywhere uniformly so one must move the sausages around on the surface).

Theories of turbulence

The theories of turbulence are primarily concerned with correctly describing scaling behaviour. There are several theories which correctly predict parts of the scaling behaviour, but none of them work universally, according to numerical simulations and experimental tests [4, 8, 42, 111]. The first theory of turbulence, the Richardson cascade model [101], is phenomenological only and cannot be derived from the Navier-Stokes equations directly. Richardson describes turbulence as an energy cascade in which large eddies, created by mean flow fluctuations, gradually divide into smaller and smaller eddies until one reaches the viscous scale at which point viscous dissipation evaporates the remaining energy as heat. This I have illustrated in Figure 1.2.

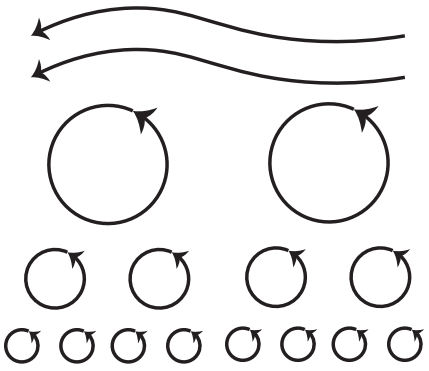


Figure 1.2 *Richardson cascade illustration. Big swirls break up into littler swirls until the smallest scales.*

The most successful quantifiable theory is due to Колмогоров (Kolmogorov) in his 1941 paper[2] (K41), which builds on the Richardson cascade [42]. The K41 theory postulates that the key macroscopic quantities which define a turbulent flow are the mean (energy) dissipation rate, $\bar{\epsilon}$ and kinematic viscosity, ν.

Since the Richardson cascade is assumed to be a self-similar process, where each eddy "forgets" the scale its parents were, the dissipation process becomes universal for all scales. This suggests the existence of a sub-range of scales (inertial range) for which the flow statistics are independent of the kinematic viscosity [42]. Through dimensional analysis one can construct the lower bound of this sub-range, known as the Kolmogorov microscale length,

$$\eta = \left(\frac{\nu^3}{\bar{\epsilon}}\right)^{1/4}. \tag{1.4}$$

This is the length scale at which the viscosity *dominates* the turbulence, which can be seen by substitution of the scale, as well as the characteristic velocity at the Kolmogorov scale ($U_\eta \sim (\nu\bar{\epsilon})^{1/4}$) into the Re definition, yielding $Re = 1$. Similarly, the Kolmogorov timescale is

$$\tau = \left(\frac{\nu}{\bar{\epsilon}}\right)^{1/2}, \tag{1.5}$$

which is the smallest evolution timescale for the turbulent dissipation.

In K41, Kolmogorov examined turbulence through looking at infinitesimal velocity differences in space at the same point in time [56]. Today this is referred to as an equal-time velocity correlation function or a longitudinal velocity *structure function*, the latter term being more commonly used when discussing theories of turbulence. These functions are not exact and can be expanded to multiple orders.

So, for $d > 2$, two length scales are dominant for HIT,

1. Kolmogorov microscale, η, where energy is dissipated by molecular viscosity, hence also known as the microscopic scale, and
2. macroscopic integral scale L at which energy is injected in the system.

In between these two scales is the inertial range within which energy is transferred

[2]It should be noted that K41 introduces the narrow definition of HIT, including the aforementioned ergodicity assumption.

towards the small scales just as in the Richardson cascade. Inside this sub-range the equal-time velocity correlation functions scale universally such that they behave as power laws with exponents.

Presently, these power laws accurately model the observed behaviour to low-order, but higher order correlation functions are observed to deviate from K41 theory [4, 8, 18, 42, 111]. For $d = 2$, the situation is different from all other cases because enstrophy (vorticity squared) must be conserved, as well as energy [42]. This results in an inverse cascade, where some energy is transferred from the integral scale to the larger scales until it reaches the system boundaries where it is dissipated [121]. In this thesis I am concerned only with $d > 2$.

The key result from K41 is the Kolmogorov spectrum[3], whereby the Fourier transform of the K41 second order structure function gives,

$$E(k) = C_3 \cdot (\bar{\epsilon})^{2/3} \cdot k^{-5/3} \tag{1.6}$$

valid in the range $1/L \ll k \ll 1/\nu$ (the inertial range), where C_3 is a universal constant, and $\bar{\epsilon}$ is the mean dissipation as before.

However, due to criticism from Ландау (Landau) [64], and on the suggestion of Óбухов (Oboukov), Kolmogorov modified his K41 theory to account for the observation that as Re increases, the bulk of energy dissipation shrinks such that it takes up a smaller and smaller proportion of the total flow volume [57]. This phenomenon is known as internal intermittency. He introduced this by replacing the mean energy dissipation, $\bar{\epsilon}$ in his K41 theory with the local spatial average,

$$\epsilon_r(\boldsymbol{x}, t) = \frac{3}{4\pi r^3} \int_{x' \leq r} \epsilon_r(\boldsymbol{x} + \boldsymbol{x}', t) d\boldsymbol{x}' \tag{1.7}$$

at a time t over a sphere of radius r. Then Kolmogorov introduced another hypothesis, that the distribution of $\epsilon_r(\boldsymbol{x}, t)$ is given by a log-normal function which is valid for large L/r, where L is the integral scale, as before [57]. Through this new definition of the dissipation, one can derive a new inertial range energy spectrum via Fourier transforming the second order structure function to give,

$$E(k) = C \cdot (\bar{\epsilon})^{2/3} \cdot k^{-5/3} \cdot (kL)^{-\mu} \tag{1.8}$$

where the definitions are the same as before and we introduce μ as an anomalous

[3]Although not actually found in his original paper.

exponent, which these days is often referred to as an *intermittency correction*. This result is known as the K62 theory. Experimental measurement of this spectrum has been unable to distinguish between the K41 and K62, implying that $\mu \ll 1$ [47].

The reason the term *intermittency* is used is that another way of seeing the shrinkage of the bulk energy dissipation with increasing Re, is that the dissipation is becoming more spatially intermittent. This happens such that eventually one sees separate regions of intense dissipation, surrounded by a sea of lower dissipation. This breaks the K41 assumption of statistical self-similarity (homogeneity), and in K62 theory Kolmogorov attempted to account for this.

However, major problems with K62 arise when examining the structure functions which are n-order moments of velocity increments along a separation distance r. For isotropic turbulence, direction does not matter, so we are free to choose the x-axis arbitrarily, and thus write the structure function as,

$$S_n(r) = \langle [u_x(x+r) - u_x(x)]^n \rangle \,. \tag{1.9}$$

Within the inertial range the structure function follows a power law, commonly written as

$$S_n(r) \sim r^{\zeta_n} \tag{1.10}$$

where ζ_n is the structure function exponent of order n. K41 predicts $\zeta_n = \frac{n}{3}$, but for $n > 3$, the theory fails. It is correctly predicted by K62 but, K62 still fails at higher n [4, 8]. At the time of writing, there is no theory for turbulence which produces good predictions for all n. This has led to almost an entire field of endeavour - creating corrections to intermittency. A particularly good review of this area is [111]. I have illustrated the complete Kolmogorov energy cascade in Figure 1.3.

An assumption core to K41 is that the Re is sufficiently large. How "large" is not defined satisfactorily, raising the argument that K41 is exact in the limit $Re \longrightarrow \infty$ and therefore the theory requires *finite* Re corrections to produce accurate predictions for the higher order anomalous exponents. This hypothesis has been examined in detail by Qian (see e.g. [98, 99]), who suggests that below some critical Re the K41 inertial range cannot exist. K41 has also been found to be correct in the regime of infinite Re [75]. However, in the regime where the Re is large enough but not infinite, corrections are still necessary. There is a wide range of literature on the topic of finite Re corrections [79].

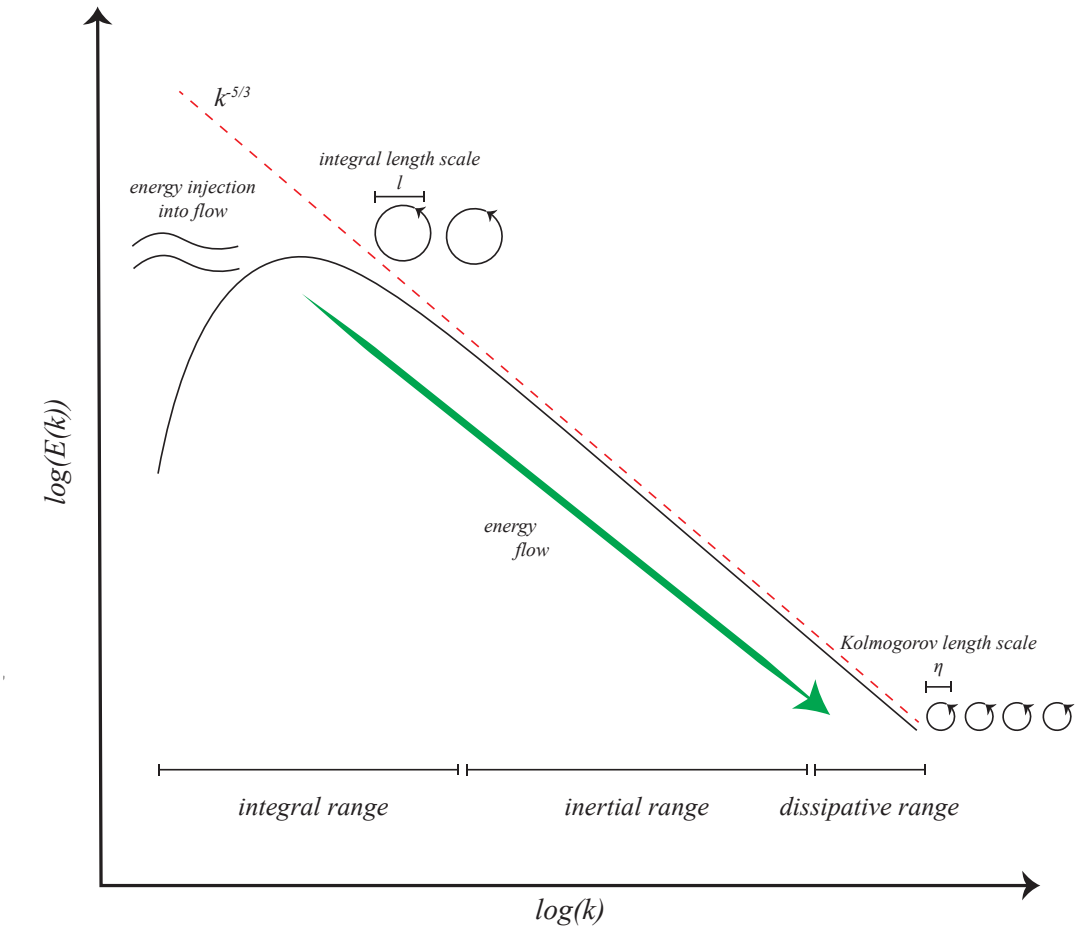


Figure 1.3 *Illustration of the Kolmogorov cascade.*

Correcting K41 for either intermittency or finite Reynolds number has been a significant focus for the field since the 1970s. However, the scale invariance of the interial range has led to analogies with critical phenomena (see e.g. [2, 16, 85]), particularly in relation to critical spatial dimension. Above critical spatial dimensions in equilibrium condensed-matter systems, critical exponents are given by mean field theories (see section 1.2.3). Since K41 makes use of mean energy transfer, some have suggested that K41 is a mean field theory, valid above some critical dimension [9, 110]. It is this hypothesis which I will examine by using simulations to exactly solve the Navier-Stokes equations at higher dimensions and using chaos as a measure in the next Chapter.

1.1.2 Numerical weather prediction

Weather forecasting is a theoretical physics problem, older than perhaps even that of turbulence, going back to the time of Aristotle [43]. The field has somewhat developed since then but as the atmosphere is a fluid, ultimately governed by the Navier-Stokes equations, it is fundamentally non-linear and chaotic. Even with modern advances in computational power, there remain challenges.

Ensemble forecasting

The models used to forecast the weather for the short-term, an hour ahead, the next-day etc. are the very same as those for the month, the season, the decade and even centuries ahead. In meteorology, such regions of time are often separated into: weather forecasting, long-range forecasting and climate prediction. In Chapter 3 of this thesis, long-range prediction is the area of primary concern. Due to the inherent uncertainty an entirely probabilistic, also known as stochastic, approach is taken to forecasting at these timescales. Whereas our more common weather forecasts are still sadly presented as largely deterministic, with perhaps the exception of precipitation.

Probabilities are generated using an *ensemble forecasting system*. The basic idea of an ensemble model is to generate a set of forecasts by making small-perturbations to the model state at initialisation which represent the uncertainty in the observations and in nature itself [35, 92]. Seeing how these evolve gives an indication of how chaotic the atmosphere is and consequently how uncertain one should be of any given forecast. There are many different ways one can generate these different initial conditions and/or perturbations. Some systems use an ensemble of entirely different models, known as a *multi-model ensemble*, or the same model but running slightly different physics schemes or different types of perturbations [66]. However it is done, the end result is that one has a handful of different realisations of what the weather might do.

The North Atlantic Oscillation

When predicting and simulating at the long-range, the larger scale phenomena dominate predictability as these are slow developing, regular or quasi-regular occurrences. In the Northern Hemisphere, the single most important factor for

year to year fluctuations is the state of the North Atlantic Oscillation, (NAO) [107]. The NAO is a statistical relationship, discovered by the English physicist Sir Gilbert Walker in 1924, between the annual mean sea-level pressure in the high North Atlantic (Icelandic Low) and the slightly East and equatorward pressure (Azores High) [127]. Seasonal variability over Europe is mediated by maritime air from the Atlantic westerlies. The pressure-gradient between the Azores High and Icelandic Low defines these westerlies. The fluctuation of the NAO is determined to be either positive or negative by the difference in pressure between the two regions. If the difference is high or low, then the NAO is positive or negative, respectively. In Figure 1.4 I have illustrated the two NAO phases as it features in Chapter 3.

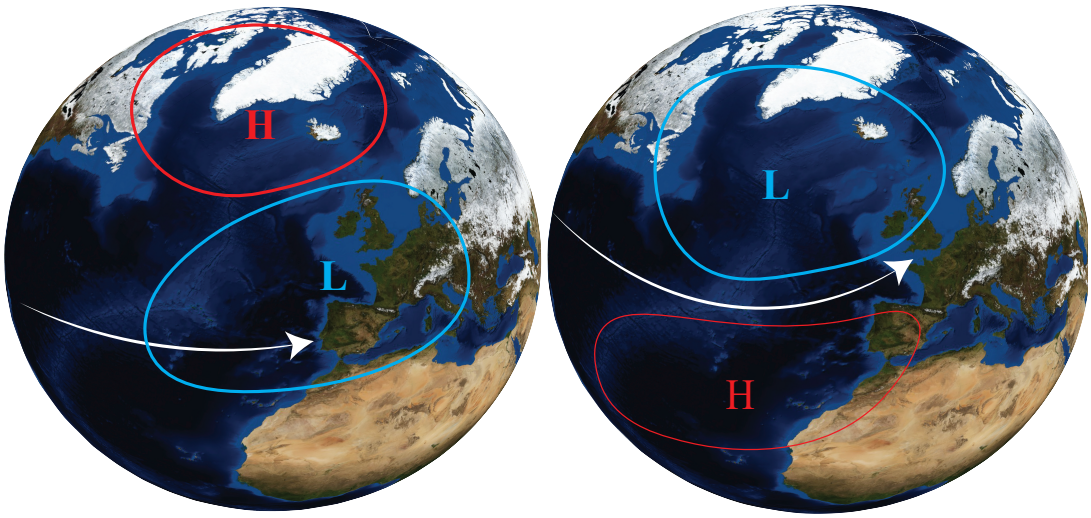


Figure 1.4 *Illustration of the North Atlantic Oscillation (NAO) pressure anomalies. Left is negative, right is positive. H and L denote regions of higher than average and lower than average pressure, respectively. Note this does not necessarily mean regions of absolute high and low pressure. Designed by the author using imagery from NASA's Earth Observatory team, led by Reto Stöckli [114].*

In the past decade, a phenomenon known as a signal-to-noise paradox has emerged in atmosphere-ocean coupled climate and long range prediction systems, particularly affecting predictions of the NAO. The paradox is that these models statistically appear to be better at predicting reality, than they are at predicting themselves [33, 63, 108]. This is puzzling indeed, as one might think that this is a *good* thing, as we want to predict reality well. So why is there a problem? It implies that the *real world* is more predictable than the *model world*, and if we want our models to be physically representative then their inherent predictability should match reality. This means that potentially we

could even improve the forecasts further, through resolving the paradox by fixing the model errors. It is now a widespread problem, beyond that of the NAO forecasts [19, 32, 107, 113, 129]. By taking a dynamical systems approach to the problem using ergodic theory, I will present an alternative hypothesis on the paradox as well as potential solutions.

1.1.3 Planetary biospheres

Moving up a scale, to consider the concept of a biosphere. The notion of a biosphere was first introduced by the geologist Eduard Suess in 1875 as the surface on the Earth where life can be found [115]. Russian polymath Вернадський (Vernadskii) later extended this to be any complete system which encapsulates life [124]; in other words, a biosphere has boundaries beyond which life cannot be found. Therefore it is at the boundaries of a biosphere that one finds life at the extremes. Beyond introducing the concepts, these seminal works offer little in the way of narrow formal definitions of the biosphere with regards to the atmosphere.

The existence of life in extreme conditions is a long established field of study, determining the extent of different biospheres. Research has established the existence of life at extreme temperatures within volcanoes and in the Arctic, at extreme pressures in the deep ocean, and in high radiation environments [23, 62, 100]. Certain life forms have even been shown to survive successfully in the space environment [55, 83]. In terms of the atmosphere, the question of what the biosphere definition might be was first discussed by Pasteur [93]. His focus was far more local than what we consider today, as he only discussed certain mountain ranges. The notion being that different peaks, which are above certain altitudes, harboured different microorganisms according to their ability to survive a combination of extremes of low temperatures and oxygen. However he stops short of defining the atmospheric biosphere. The first paper which mentions explicitly the idea of a well-defined biosphere for the atmosphere is a report of a rocket sampling campaign conducted in 1974 by Имшенецкий (Imshenetsky) et al. of the Institute of Microbiology, USSR Academy of Sciences [54]. Sounding rockets were fitted to capture microbiological samples in the mesosphere at altitudes between 48 km and 77 km. Subsequently they reported finding microscopic fungi in the form of spores. These findings remain to this day, the only published study of these altitudes, and therefore there are credible concerns about contamination.

Searches for life in the limits of the upper-atmosphere

The primary reason why little research has been conducted in the upper atmosphere is because of its specific inaccessibility. The mesosphere - lower thermosphere (MLT) region is often colloquially referred to as the "ignore-osphere" because it is inaccessible via balloons and conventional aircraft due to the low air density. It is also inaccessible to satellites, as there remains sufficient air for the drag to be non-negligible [112]. Thus until recently, the only viable sampling option has been to use sounding rockets. However, in the last decade new possibilities for sampling have arrived, such as small inexpensive CubeSats and low Earth-orbit (LEO) commercial space craft [106, 125].

The highest altitudes where life seems to have been plausibly established are the top of the troposphere and the lower stratosphere [29]. However, studies deeper into the stratosphere remain controversial due to the serious and credible questions of contamination. As with all research that intends to establish the existence of life in an extreme habitat, there are serious concerns around potential contamination of results. This is especially true for the upper atmosphere, as the discovery of even a single particle at such high altitudes is potentially significant, and yet the error size is considerably larger than one particle.

Several studies have shown that there are micrometre sized particles in concentrations of approximately 1 particle cm^{-3} in the stratosphere [49, 102, 120, 134, 136]. A key and relatively recent field campaign is [126], which focused on sampling the stratosphere, finding two different species of bacteria, *Bacillus simplex* and *Staphylococcus pasteuri* and a fungus, *Engyodonitium album* at 41 km [126]. Additional studies of the stratosphere which have found bacteria include [48] and [120]. Most recently there was a review by Dassarma in 2020 which calls for further stratospheric field campaigns in order to independently establish whether contamination is at play with their measurements. However, further to this, the transmission of microbes in the atmosphere has also been postulated as one of the mechanisms behind changing microbial diversity in Antarctica [94].

The upper-atmosphere sampling contamination problem is tackled in Chapter 4 by re-thinking the flow of air through a detector in terms of its motion relative to contaminants from the sampling craft. The aerobiology community now regards the atmospheric biosphere to be defined as *the maximum altitude at which life survives in the atmosphere*, so the issue of contamination is a pressing one.

Planetary protection

If we take a further step back, we can consider how different planetary biospheres might interact. One way of imagining two planets and their biospheres in space, is to draw analogy to separate islands in an ocean. The ocean and distance between the islands prevents certain species from colonising each other, acting as a dispersal filter, much like interplanetary space [24]. Such interactions are a complex dynamical system, and the variations of this, the density, populations dynamics and species in space and time, is known as the study of biogeography. Due to the far from equilibrium nature of planets as opposed to islands, it is not possible to draw exactly the same parallels, however there are applications to planetary protection.

Planetary protection is an endeavour to ensure space exploration is conducted responsibly, through preventing contamination of outbound space missions to other planetary bodies via colonisation by life originating from Earth. Likewise protecting the Earth from extraterrestrial life from return missions. Through applying the concept of phase-spaces from dynamical systems theory as well as aspects of a dynamical model developed by MacArthur and Wilson in their seminal text *The Theory of Island Biogeography* [76], a new approach to planetary protection is found. By reducing the problem to the dynamics of microbial growth and death, as opposed to simply survivability or habitability, practical steps are proposed both for the theorists and for experimentalists. The mathematics that underpins this is covered at the end of the next section.

1.2 Theoretical underpinning – dynamical systems theory

A dynamical system can be loosely thought of as something, be it a quantity or a collection of quantities, that evolves over time, according to some rules. The mathematical thread running through this thesis is that in each of the chapters a problem is tackled by reducing an immensely complicated physical system to essentially some form of the equation,

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = F(\boldsymbol{x}(t)) \text{ with } \boldsymbol{x} \in R^m \tag{1.11}$$

where $\boldsymbol{x}(t)$ can be abstractly referred to as the "state" of the system at a time t (i.e. this could be anything from temperature, pressure or any number of these), F is some differentiable function that defines the evolution of the state $\boldsymbol{x}$ with time. For all dynamical systems we can abstractly describe their properties by considering their phase-space (sometimes referred to as state-space). This space, here denoted R^m, is a geometric representation where each point represents a unique possible state of the system. A system evolves with time according to a trajectory through the phase-space, typically some continuous curve.

When one initially sets off from $t = 0$ there is usually some transient behaviour where the system gradually moves towards a set of states defining a motion on an "attractor" in the phase-space. The attractor is effectively a trajectory in this abstract space which the system on average prefers, independent of the initial conditions. For low dimensional or simple systems, it can be easy to characterise the geometry of the attractor e.g. periodic, fixed point etc but for real-world non-equilibrium complex systems these kinds of simplistic representations are no longer possible. Instead a more statistical approach is required in which we think carefully about time-averages and statistical descriptors of chaos such as maximal Lyapunov exponents.

Within this thesis, the phase-space concept is used through Chapter 2, 3 and 5. For example, in the first two Chapters we will use these trajectories to define the chaotic evolution of fluids, and in Chapter 5 we more abstractly use the notion of the phase-space to define sub-spaces of states or "allowable" trajectories to characterise colonisation success for a population. This illustrates the scale invariant nature of the dynamical systems approach. By mapping each problem onto an abstract space, we can conceptualise them in the same way and use similar tools. There is mathematical rigour behind this, although beyond the scope of this thesis, for instance Ornstein's isomorphism theorem indicates that seemingly unrelated complex systems can effectively be the same problem [90]. To tease out the scale invariance of the dynamical systems approach, the chapters are in order of smallest to largest scale. This is shown in Figure 1.5. I feel it is also worth noting for completeness that this dynamical systems reductionist approach to complex systems is not the only way to tackle such problems. Other approaches include focusing on statistical methods, often using large data or modeling or on only non-linear parts.

Finally, please note that the mathematics in this section is more formal than what is required knowledge to understand the individual research chapters. It is

included for those who require slightly more formal definitions to understand the concepts.

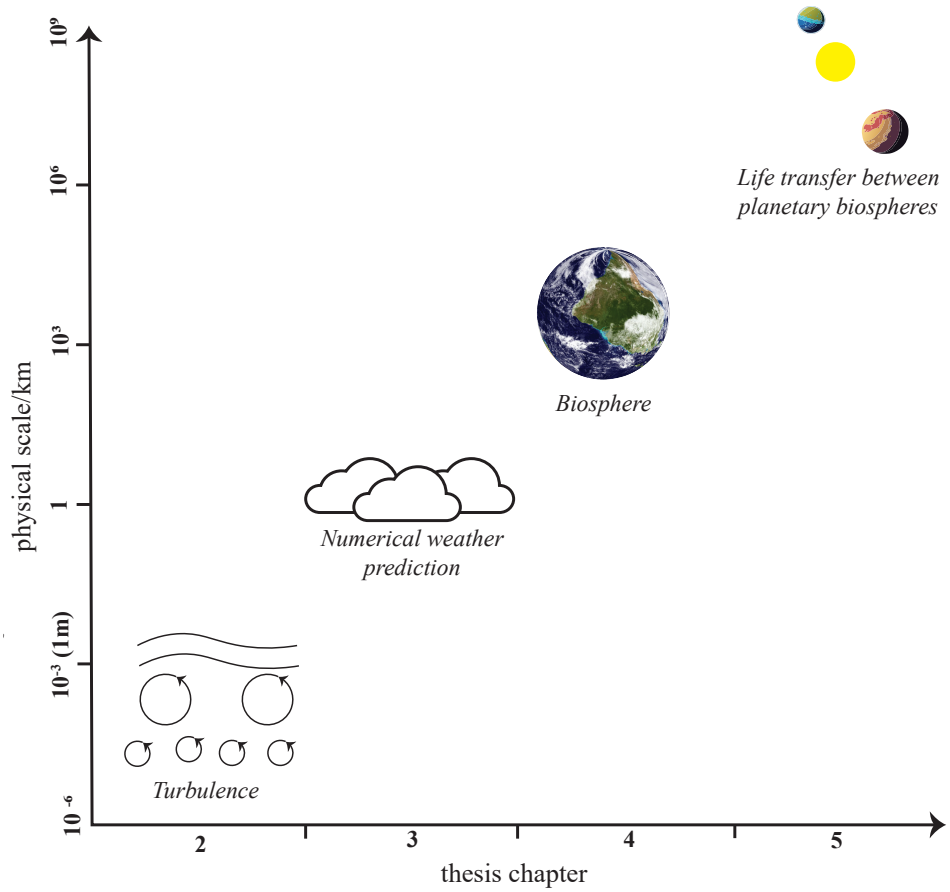


Figure 1.5 *Chapter order and scaling in this thesis.*

1.2.1 Chaos theory

"Chaos" in this thesis will refer only to the consequence of a system's *sensitive dependence on initial conditions*. Consider a deterministic system that evolves according to the ordinary differential equation 1.11. Then let $\boldsymbol{x}(t)$ be the solution of such an equation, where we have an initial condition $\boldsymbol{x}(0)$. If we change *very* slightly the position of this initial condition s.t.

$$\boldsymbol{x}(0) \to \boldsymbol{x}'(0) = \boldsymbol{x}(0) + \delta\boldsymbol{x}(0) \tag{1.12}$$

then the end result at a later time t will be different. This is illustrated in Figure 1.6. Very initially the difference might be quite small, but as t increases one will see this difference grow exponentially fast i.e. $\delta\boldsymbol{x}(t) \sim \delta\boldsymbol{x}(0)\exp(\lambda t)$, where λ is a measure of the rate of divergence between the two solutions. It is specifically this kind of growth that is commonly referred to as chaotic. Since the system is

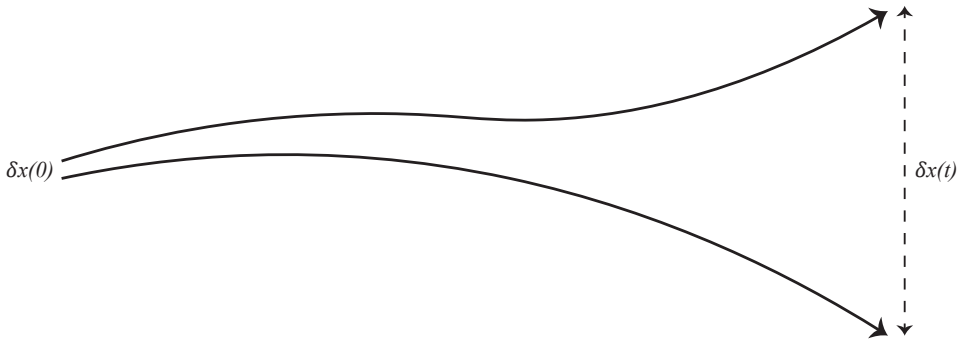


Figure 1.6 *Illustration of chaos in one-dimension (vertical axis). Divergence of initially close trajectories with time (horizontal axis).*

m-dimensional, we can write the evolution of such perturbations as

$$\frac{d}{dt}\left(\delta x_j(t)\right) = \sum_j^m \frac{\partial f_j}{\partial x_k}\left(\delta x_k(t)\right) \tag{1.13}$$

which via Oseledet's theorem (treated in next section 1.2.2) yields

$$\delta\boldsymbol{x}(t) = \sum_i^n \alpha_i \exp\left(\lambda_i t\right)\boldsymbol{e}_i \tag{1.14}$$

as the theorem allows for the evolution of the perturbation difference to be expressed in terms of an orthonormal basis called the *Lyapunov vectors* $\{\boldsymbol{e}_i\}$, and λ_i are known as the *Lyapunov exponents* which as a set $\{\lambda_i\}$ are the *Lyapunov*

spectrum. In the limit $t \to \infty$ if $\exists$ a single $\lambda_i > 0$, this will always dominate and consequently there will always be divergence. Hence one typically uses the *maximal Lyapunov exponent*,

$$\lambda_{\max} = \lim_{t\to\infty} \frac{1}{t} \left\langle \frac{|\delta \boldsymbol{x}(t)|}{|\delta \boldsymbol{x}(0)|} \right\rangle \tag{1.15}$$

where the braces denote the time-mean. For very simple systems, the Lyapunov exponent can be computed analytically but these are not encountered in this thesis. In the next Chapter a numerical method will be introduced and used to compute the maximal Lyapunov exponent for turbulence.

1.2.2 Ergodic theory

Ergodic theory is typically a very formal area of mathematics due to its abstract nature and the necessary complexity of the reasoning. My approach in this section will be to give reasonably formal definitions, without proofs, followed by more qualitative or physical description. For even more formal approaches, I refer the reader to [34, 105].

The ergodic theorem

Let ϕ be a continuous integrable function and ρ an invariant probability measure. Consider an initial condition $x_0 \in M$ and the path $\{f^i x_0\}_i \geq_0$. Then for almost all[4] x_0 and ρ $\exists$ the limit (time-mean) such that

$$\bar{\phi}\left(x_0\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi\left(f^i x_0\right). \tag{1.16}$$

If ρ is said to be *ergodic* $\implies$ $\bar{\phi}$ is almost the constant

$$\bar{\phi}\left(x_0\right) = \int \phi(x) \rho_0(dx) \tag{1.17}$$

over almost all the space (space-mean). If all the above conditions are true, then this has the consequence that the time-mean and space-mean are equivalent [105].

[4]For the uninitiated, "almost all" is mathematicians speak for "all elements of an infinite set, except for finitely many". So there may be points where the convergence does not hold, but these points have measure zero on ρ (negligible probabilistically).

To my mind, the most physically graspable consequence of this is in the statistical mechanics of Brownian motion. One can track the motion of individual colloidal particles to determine their mean square displacement with time. Tracking enough of these particles, one will find (for a "pure" Brownian motion system i.e. ergodic) that the mean square displacement for a single particle over very *large times* will converge on the average for a very *large number of particles* over a shorter time. For the mathematicians, the measure in this example is not the "pure" invariant probability measure that we have within the theorem but really a "physical measure" related to time averages [105]. Оселедец (Oseledet) generalised the ergodic theorem to the product of matrices, otherwise known as the multiplicative ergodic theorem.

The Oseledet multiplicative ergodic theorem

Let ρ be a probability measure in a space M, and $f : M \to M$ a measure preserving map[5] such that ρ is ergodic. Let also $T : M \to m \times m$ matrices be a measurable map. Define the matrix $T_x^n = T\left(f^{n-1}x\right)\ldots T(fx)T(x)$. Then, for ρ almost all x, the following limit exists,

$$\lim_{n\to\infty} \frac{1}{n} \log \|T_x^n v\| = \lambda_i \tag{1.18}$$

where v is an arbitrary non-zero vector $\in M$ that transforms under repeated transformations by the matrix T_x^n [105]. The rate of the growth or shrinkage due to the transformation is our Lyapunov exponent λ_i. Thus the Oseledet theorem guarantees the existence of the limit in 1.15.

The Ruelle relation: connecting Lyapunov exponent to Reynolds number

The French mathematician David Ruelle derived a link between turbulence and chaos, by relating the measure of chaos in a system, known as a Lyapunov exponent, with its Reynolds number [104]. Ruelle argued that the Lyapunov exponent should vary with the size of an eddy, l and that the largest characteristic exponent should vary with the smallest eddy size which are defined in K41 by the Kolmogorov microscale, η [104]. By simple dimensional analysis he shows that the Lyapunov exponent should go as, $\lambda_l \sim \bar{\epsilon}^{1/3} l^{-2/3}$. If one then applies the K41

[5]Here "map" is used in the strict mathematical sense that f is a function which uniquely takes us between two different places both of which reside in the space M.

microscale to the size of the largest eddy l,

$$\begin{aligned}\lambda_l &\sim \frac{(\bar{\epsilon})^{1/3}}{\eta^{2/3}} \\ &\sim \frac{(\bar{\epsilon})^{1/2}}{\nu^{1/2}} \\ &\sim \frac{1}{\tau}, \text{the Kolmogorov timescale.}\end{aligned}$$

Since this all relies directly on K41, one can then use the relation $\eta \sim LRe^{-3/4}$ to make the following connection,

$$\lambda \sim \frac{1}{T_0} Re^{\alpha} \tag{1.19}$$

where $\alpha = \frac{1}{2}$, and T_0 is the timescale of the integral length scale. It is equation 1.19 which I will refer to as *the Ruelle relation* in Chapter 2. This result was first verified using Direct Numerical Simulation (DNS) by [13]. They found that the Lyapunov exponent solely depended on the Reynolds number according to

$$\lambda \propto Re^{0.53}. \tag{1.20}$$

Further direct numerical simulation tests of this Ruelle relation yielded the same scaling [13, 14]. The difference in α from 0.5 is considered an intermittency effect [27].

1.2.3 Some specific forms of dynamical systems

Mean field theory

Whilst not often considered a formal part of dynamical systems theory in pure mathematics, the concept of the *mean field theory* most certainly pertains to such systems as it is a technique for solving complicated interacting dynamics. In general, mean field theories replace the complicated fluctuating or interacting part of the system with some *mean field* that represents the average of those interactions. In so doing, the fluctuations around the mean value are ignored. Thus, it is important that the fluctuations can be considered negligible.

One way of exploring complex systems is by scaling them by their dimension, this is the mathematical technique known was *renormalization*. As one goes to

higher and higher dimensions, sometimes the significance of the fluctuations in the system reduces until they are negligible relative to the mean, at this point the system behaves like the mean field theory. The dimension at which this occurs is known as the *upper critical dimension*, d_c [45]. As one scales the system, there can sometimes occur lines of points in the phase-space which remain unchanged, despite the scaling; these are known as *fixed lines*.

In the next chapter we will consider the Navier-Stokes equations which describe the dynamical evolution of fluids. In theoretical physics, such equations can be viewed as an "effective" mean field theory as one replaces the microscopic molecular fluctuations with the viscosity parameter. Likewise when we come to think about population dynamics, instead of exactly modelling the fluctuations with time, one just considers the average growth.

Population growth and ecology

A dynamical equilibrium, also known as a fixed point, occurs when the state of a time-evolving system does not change with respect to time. This naturally is of importance to biological systems modeling, where one typically has large populations which compete with each other for resources over time, reaching effectively a stationary-state. These theories break down in the limit that either of the increasing or decreasing rates becomes too small, or indeed that the overall numbers of units involved are too small as at these limits small fluctuations cease to be negligible. This is in much the same way that in the low-density limit, the Navier-Stokes equations cease to accurately model fluid behaviour.

Consider a net population at a time t of size $N = N(t)$, and let the birth and death rates be $\mu(t)$ and $\lambda(t)$ respectively. Then the balance between these rates giving the net population is

$$\frac{1}{N}\frac{dN}{dt} = \mu(t) - \lambda(t) \tag{1.21}$$

which in dynamical equilibrium $\frac{dN}{dt}|_{N=N^*} = 0$ gives an average population size N^*. In MacArthur and Wilson's 1967 *The Theory of Island Biogeography* this idealised model is adapted for modeling the numbers of different species on different islands. The number of species S varies as,

$$\frac{dS}{dt} = \mathcal{I} - \mathcal{E}, \tag{1.22}$$

where $\mathcal{I}$ and $\mathcal{E}$ are immigration and extinction rates, respectively. This theory is the inspiration for my final research chapter which uses the analogy of planets as being like islands on Earth, to arrive at a new approach to planetary protection procedures. The key mathematical concept to grasp is that some theories (i.e. Island biogeography) are reliant on high continuous transfer rates (e.g. between islands or between planets), to have any predictive utility as no meaningful equilibrium can be established. Secondly, MacArthur and Wilson's 1967 work introduced a fundamental model for the mean-time to extinction of a population of mammals on an island. The mean-time to extinction, for any species, is proposed as the key metric for successful colonisation of a planet (or planetary protection failure). Planetary protection traditionally has taken a probabilistic approach to determining acceptable contamination risks. This is challenged by arguing that growth is not a probabilistic process but rather a binary one.

Chapter 2

Searching for a critical dimension in homogeneous isotropic turbulence using chaos measurements

He could see the honey, he could smell the honey,
but he couldn't quite reach the honey. –A.A Milne [82]

2.1 Introduction

This chapter will consider the physics of fluid motion, which scales ever increasingly from the molecular size onward, as described by the Navier-Stokes equations (NSEs). Earlier, in the introductory chapter, we saw how the NSEs can be re-written such that the Reynolds number is the only governing parameter of the equations. However, this is not the whole story - those equations can also be written for any number of dimensions, hence dimension is another parameter besides Reynolds number. As a tool in theoretical physics, dimension is now often used to explore problems. There are many examples of this approach, particularly in areas such as quantum field theory [10, 11, 84] and in what is known as "critical phenomena", which broadly covers the phenomena which have critical points: sudden, instantaneous transitions in behaviour, often describable by mean-field theories. For example, Wilson and Fischer successfully predicted critical points in the Ising model of ferromagnetism by mathematically investigating the behaviour of fluctuations in higher dimensions [130]. The success of this technique, known as renormalization, was later applied to turbulence in the late 1970s [41, 58] with not much success, and then by Wuwell Liao in the late 1980s and early 1990s. Liao developed an approach, very similar to that of Wilson-Fischer, for deriving

critical phenomena in the freely decaying isotropic and homogeneous Navier-Stokes turbulence [69, 70]. Liao derived the Kolmogorov exponents for the energy spectrum of freely-decaying, fully-developed, near-incompressible turbulence, and found to first order in the expansion parameter ϵ, an IR stable fixed-line with upper critical dimension, $d_c = 6$. This result thus tentatively suggests that above six dimensions, turbulence can be described by a mean-field theory.

Prior to Liao, Nelkin proposed through a similar analogy to critical phenomena, the existence of a crossover dimension, above which the Kolmogorov scaling relations in K41 become exact [85, 86]. Following Nelkin, Kraichnan derived that temporal fluctuations of a passive scalar field acting as a tracer for an incompressible d-dimensional turbulent flow were suppressed by increasing dimension [59]. Around the same time, these ideas led to the use of dynamic renormalization group methods by Forster et al. [39, 40] and DeDomincis et al. [30]. Further analytical progress in this direction has proved challenging, leaving numerical approaches, in particular DNS, as the optimal methodology with which to explore fluid turbulence in higher dimensions [46, 53, 135].

Recently, [12] using DNS of the forced NSEs in four spatial dimensions found reduction in total energy fluctuations, increased velocity derivative skewness[1] and asymptotic dissipation rate compared to three dimensions. These findings are consistent indications of increased forward energy transfer, resulting in an extended inertial range with smaller Kolmogorov scale. Using numerical solutions to an eddy damped quasi-normal Markovian (EDQNM) approximation for the NSEs, the same trends were seen in higher dimensions [22]. In both studies, the reduction in fluctuations with increasing dimension was significantly larger than could be argued from the central limit theorem [22]. The EDQNM study found velocity derivative skewness to increase until approximately eight dimensions, before decreasing towards a potentially dimension independent constant and enstrophy production reached a maximum value near six dimensions. In six dimensions, no positive Lyapunov exponent was measured, even at very high Reynolds numbers. These results from EDQNM corroborate Liao's work, suggesting that six may be a critical dimension for the NSEs. However, such closure-based schemes do not account for intermittency and hence are not suitable for anomalous scaling measurements (Kolmogorov exponents). The only acceptable test is to use DNS measurements in six dimensions.

[1] A measure of the asymmetry of the distribution of velocity gradients, defined by equation 2.12.

In this chapter I present an initial simulation study of chaos in the Navier-Stokes equations (NSEs) with integer dimension ranging from $d = 3$ to $d = 6$. The objective of this study was to try to find signs that $d = 6$ is the critical dimension for the NSEs, specifically using chaos as a measure. This is because it is not computationally feasible to run simulations of sufficient resolution to measure anomalous scaling at dimensions higher than three. During the course of my research I did run the largest ever turbulence simulation in four dimensions (1024^4 box size), which was expected to have sufficient resolution to provide preliminary scaling measurements. However, the High Performance Computing (HPC) failed to cope with the size of the simulation. The simulation would fail persistently because it exceeded available memory during runtime, causing out-of-memory (OOM) errors and crashes on various nodes. This was due to upgrades in the underlying ARCHER2 system software that resulted in an increased memory overhead per node, making configurations that previously worked unstable. This prevented sufficient data collection. Instead, my focus turned to the chaotic study of the NSEs using DNS, which required less memory, and comparing these results to those from EDQNM.

2.2 Methodology - simulations

Simulating the NSEs with a sufficient resolution in six dimensions to finally test these theoretical predictions directly is not possible presently, due to such a simulation requiring exabytes of running memory. Even with all the world's largest supercomputers combined, one is at least an order of magnitude away from such capacity. Only on an exascale system with advanced memory handling might this be possible, but this is more than a few decades away [15, 68]. Following Dr Daniel Clarke's development of a six dimensional version of the eddyBurgh DNS code, the goal was to do the best test of the six dimensions hypothesis that is possible today. I hence present the first DNS results of the NSEs in six dimensions, and compare them to similar, in terms of Reynolds number, simulations in three, four and five dimensions, as well as EDQNM results at similar Reynolds numbers. I find that the general trends observed by EDQNM at both low and high Reynolds numbers, are also seen in DNS.

2.2.1 Direct Numerical Simulation (DNS)

This study examines turbulent statistics and chaos using forced DNS in three, four, five and six spatial dimensions. For each simulation, calculation of statistics is only done once a steady state is reached in the total energy fluctuations. Statistics are computed for multiple large eddy turnover times. For a like-for-like comparison to be possible over such a wide range of d, one is limited by computational maximum grid sizes of 512^3, 512^4, 128^5 and 64^6. This allows us to compare each of these dimensions over a similar lower Reynolds number range and extend our results to higher Reynolds number in the case of three and four dimensions.

The simulation code

The fully de-aliased pseudo-spectral eddyBurgh code [5, 20, 51, 138] was used to numerically integrate the d-dimensional NSEs in a periodic cube of length 2π. Periodic boundary conditions are imposed to ensure the simulation is homogeneous and isotropic. The d-dimensional NSEs are defined,

$$\begin{aligned} \partial_t u_i + u_j \Omega_{ji} &= -\partial_i \left(P + \frac{u^2}{2} \right) + \nu \nabla^2 u_i + f_i \\ \partial_i u_i &= 0 \text{ (known as the incompressibility condition)} \end{aligned} \tag{2.1}$$

where u_i and P_i are the velocity and pressure fields of integer dimension respectively, ν is the kinematic viscosity, $\Omega_{ij} = \partial_i u_j - \partial_j u_i$ is the vorticity 2-form, f_i is some external forcing function applied to the whole field, and the density has been set to unity. Here I use the Einstein Summation Convention (ESC) for dimensions over i, j and in this chapter I will examine integer dimensions between 3 and 6. Through imposing periodic boundary conditions one can also decompose the velocity field into a truncated Fourier series (Galerkin approximation), known as spectral decomposition.

By moving into spectral space (k-space), spatial derivatives in real-space become products in k-space, which improves numerical accuracy and efficiency for the linear terms of the NSEs. However, the non-linear term involves products of fields, which correspond to convolutions in k-space. To avoid the computational cost of direct convolution, the pseudo-spectral method computes the non-linear term in real-space, where these products are trivial, and then transforms this term

back to k-space. The necessary projection onto the divergence-free subspace also removes pressure explicitly, simplifying the evolution of the velocity field while also preserving incompressibility.

The Navier-Stokes equation can be re-written in spectral form as,

$$\left(\partial_t + \nu k^2\right) u_i(k_i, t) = W_i(k_i, t) - ik_i \mathcal{F}\left[p + \frac{u^2}{2}\right] \tag{2.2}$$

where $\mathcal{F}$ is the Fourier transform and $W_i(k_i, t) = \mathcal{F}\left[W(x_i, t)\right] = \mathcal{F}\left[\epsilon_{ijk} u_j \omega_k\right]$ i.e. the k-space version of the non-linear term. One can then use the k-space incompressibility condition $k_i u_i = 0$, by applying it to equation 2.2, leading to the LHS being zero such that,

$$0 = k_i \cdot W_i(k_i, t) - ik^2 \mathcal{F}\left[p + \frac{u^2}{2}\right]. \tag{2.3}$$

Substituting back into equation 2.2,

$$\begin{aligned}\left(\partial_t + \nu k^2\right) u_i(k_i, t) &= W_i(k_i, t) - ik_i \mathcal{F}\left[p + \frac{u^2}{2}\right] \\ &= W_i(k_i, t) - \frac{k_i k_j W_j(k_i, t)}{k^2}, \\ &= P_{ij}(k_i) W_j(k_i, t)\end{aligned}$$

by defining projection where $P_{ij}(k_i) = \delta_{ij} - \dfrac{k_i k_j}{k^2}$,

thus solving the evolution is now much simpler.

The code completes the following procedure to evolve the velocity field a single step forward in time:

1. Compute vorticity in k-space, $\omega_i(k, t) = \epsilon_{ijk} i k_j u_k(k, t)$.
2. Convert $\omega_i(k, t)$ and $u_i(k_i, t)$ into real-space velocity and vorticity via fast Fourier Transform (FFT).
3. Compute nonlinear term, $W_i(x_i, t)$, using real-space terms from previous step.
4. Convert non-linear term to real-space via FFT.
5. Project the non-linear term onto solenoidal subspace to remove divergence.

6. Compute the velocity at the next step $u_i(k_i, t + \delta t)$ using the divergence free non-linear term from previous step and a time-stepping procedure.

Complete details of the code can be found in [138]. However in short, the code is fully parallelised by dividing the simulated lattice into equal slabs, each computed by a single processor which communicate with each other via Message Passing Interface (MPI) and FFTs are computed by the PFFT library [96].

The problem of relaminarisation

A previous study found that when the NSEs are forced with a negative damping scheme, which is commonly used to study HIT in DNS, at low Reynolds number a relaminarisation process occurs analogous to that in parallel wall-bounded shear flows [71]. This adds a new timescale to the flow which results in Lyapunov exponents saturating rather than dropping to zero or less. Thus in this study a simple Gaussian force was used, which was set to ensure the energy dissipation was approximately equal to the average forcing rate,

$$\hat{f}_i(k_i, t) = \begin{cases} \sqrt{\frac{2\epsilon}{tk_f}}\hat{u}_i(k_i, t) & \text{if } k_i = 2, \\ 0 & \text{otherwise ,} \end{cases} \tag{2.4}$$

where in all the simulations the dissipation rate ϵ, was a parameter set a priori to 0.1. The field was initialised by random sampling a zero-mean Gaussian distribution for each spatial component of the velocity field for the complete set of points defining the field.

The simulations were well resolved, with $k_{max}\eta > 1$ and time step dt such that $dt < 0.05\tau$, where k_{max} is the largest wavenumber, η the Kolmogorov microscale and τ the Kolmogorov timescale. The former condition ensures that the simulation remains de-aliased. Aliasing errors occur when wavenumbers higher than the maximum resolvable wavenumber get misrepresented as lower wavenumbers, so the condition ensures these are truncated before that. The time step condition ensures that all the scales right down to the Kolmogorov microscale remain well resolved. To verify numerical stability and the persistence of turbulence, diagnostics such as energy spectra, total energy and Reynolds number were routinely monitored during runtime.

Testing invariants - code verification

In order to check that the code at each dimension was performing physically correct calculations, I computed inviscid invariants for the NSEs [12]. There are two sets of invariant, depending on the parity of the integer dimension. The invariants for the three and four dimensional versions of the code were tested before [51]. Since six dimensions has never been done before (the case for five dimensions can found in [46]) I will illustrate the even case. In d even dimensions, let $d = 2m$, where m is an integer, then it has been proved by both Овсиенко (Ovsienko) et al. and Хесин (Khesin) et al. [2] that,

$$I_m[f] = \int f\left((dv)^m\right) d^d x, \tag{2.5}$$

is conserved, where dv is the velocity 2-form, and $(dv)^m$ is the antisymmetric multiple product of order m, and f is some arbitrary function. The 2-form of the velocity field can be written, $\Omega_{ij} = \partial_i u_j - \partial_j u_i$, putting this together for the case of $d = 6$ s.t. $m = 3$ one has

$$I_3 = \int (\epsilon_{ijklmn}\Omega_{ij}\Omega_{kl}\Omega_{mn}) d^6 x, \tag{2.6}$$

which is a 15 term sum. Using the `FORM` Symbolic Manipulation System by [123] this yields,

$$\begin{aligned} I_3 = &+ \Omega_{01}\Omega_{23}\Omega_{45} - \Omega_{01}\Omega_{24}\Omega_{35} + \Omega_{01}\Omega_{25}\Omega_{34} - \Omega_{02}\Omega_{13}\Omega_{45} \\ &+ \Omega_{02}\Omega_{14}\Omega_{35} - \Omega_{02}\Omega_{15}\Omega_{34} + \Omega_{03}\Omega_{12}\Omega_{45} - \Omega_{03}\Omega_{14}\Omega_{25} \\ &+ \Omega_{03}\Omega_{15}\Omega_{24} - \Omega_{04}\Omega_{12}\Omega_{35} + \Omega_{04}\Omega_{13}\Omega_{25} - \Omega_{04}\Omega_{15}\Omega_{23} \\ &+ \Omega_{05}\Omega_{12}\Omega_{34} - \Omega_{05}\Omega_{13}\Omega_{24} + \Omega_{05}\Omega_{14}\Omega_{23} \end{aligned}$$

This summation calculation was implemented within the real-space part of the DNS code, at the point that the non-linear term is computed. When run with zero viscosity, in six dimensions, this invariant was found to be constant, although growing a little at very long-times, due to the eventual compounding floating-point error inherent in each sum. The lower dimensional invariants had been tested prior with similar results, as reported in [51].

[2] See Theorem 9.2 in [6].

Chaos measurement - Finite time Lyapunov exponent

In this study I use the Finite time Lyapunov exponent (FTLE) as the chaotic measure. FTLEs were used in previous studies using the same DNS code [14, 52], and as a method have been studied in detail (see e.g. [65, 137]). The Maximal Lyapunov exponent has been proved to be independent of the initial infinitesimal perturbation for infinite dimensional systems of equations such as the NSEs, and it is commonly assumed to be the case for finite dimensional ones [13, 91]. To account for the small dependence on the initialisation for finite perturbations in finite time, I estimate the Maximal Lyapunov exponent by averaging over an ensemble of different realisations of finite perturbations in finite times.

For some original field $u_i^{(1)}$ a perturbed field is generated, $u_i^{(2)} = u_i^{(1)} + \delta u_i$. This perturbed field is then evolved over a finite time step ΔT, and the difference from the original measured. This process is repeated at each time step. The field difference is defined,

$$\begin{aligned} \delta_t &= \sqrt{2E_{1,2}(t)} \text{ where } E_{1,2} \text{ is,} \\ E_{1,2} &= \frac{1}{2}\int dk_i \left| \hat{u_i}^{(1)}(k_i,t) - \hat{u_i}^{(2)}(k_i,t) \right|^2 \end{aligned} \tag{2.7}$$

where under ESC one sums over dimension i. The perturbation to the field at each time step is,

$$u^{(2)}(x_i, \Delta t) = u^{(1)}(x_i, \Delta t) + \frac{\delta_0}{\delta_{\Delta t}} \delta u(x_i, \Delta t) \tag{2.8}$$

where δ_0 is set to 10^{-3} and δT to 0.1 seconds. The FTLE is then determined by,

$$\tilde{\lambda} = \frac{1}{\Delta t} \ln \left(\frac{\delta_{\Delta t}}{\delta_0} \right) , \tag{2.9}$$

and consequently the average FTLE which is assumed to be representative of the Maximal Lyapunov exponent is defined by the average over all the FTLEs measured over the course of the simulation, $\lambda = \langle \tilde{\lambda} \rangle$.

The FTLE method has been studied in detail and found to compute Lyapunov exponents stable under the variation of both time step and lattice size [52]. It was also found that the FTLEs reached stability faster than other characteristic quantities, such as energy and dissipation rate. Hence, even though the DNS in six dimensions is at grid sizes too small to measure intermittency properties,

these chaos measurements using FTLEs are stable. See the review by [53], which illustrates the evidence for using FTLEs as stable measures of flows.

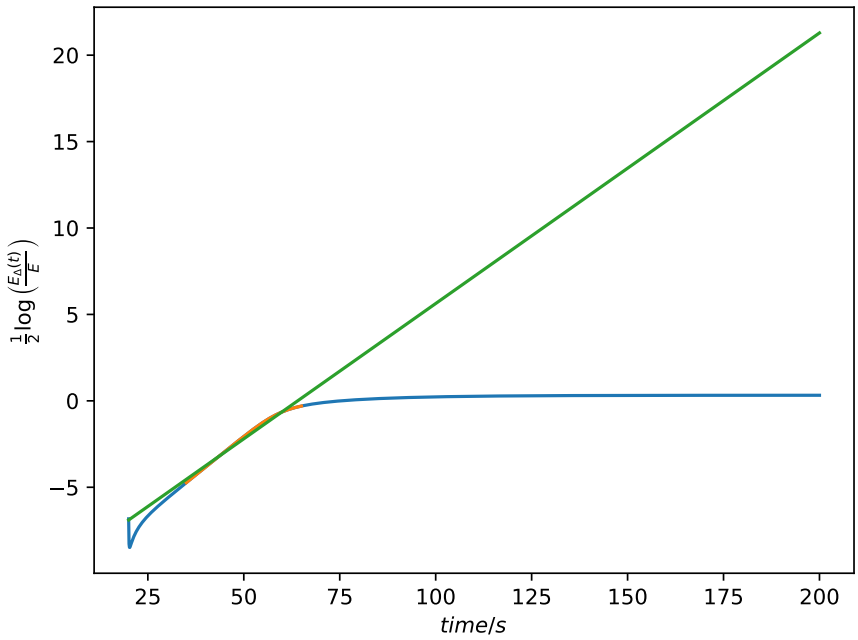
2.2.2 Eddy Damped Quasi-Normal Markovian (EDQNM) model

Due to the computational limitations imposed by increasing dimension up to six, and to provide stronger cross-validation of the general trends found in [21], the same EDQNM code as in that paper was used to simulate comparable low Reynolds numbers for the same dimensions, extending into higher Reynolds number to re-obtain the same results as [21]. EDQNMs are not capable of observing intermittency effects but they are suitable for examining first order statistical trends regarding chaos. The same procedure to find the Lyapunov exponent was used as in that work also.

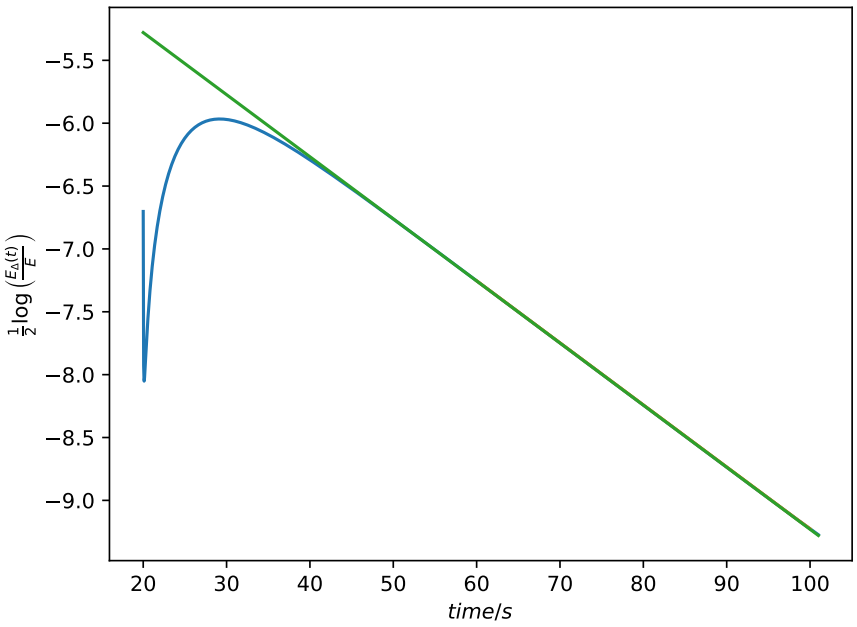
This involves manually locating the latest stage of exponential growth or decay, and then performing a linear fit of $(1/2)\log\left(E_\Delta(t)/E\right)$ vs t, where E is the total energy defined by the average of the half mean square velocity. To do this faster I developed a quasi-automated procedure in which the analysis code would initially attempt the best fit over a given time-range and present the result to the user. This could then be accepted or rejected as a correct fit. If rejected, the user could then manually key-in a more suitable time-range over which to compute the best-fit. This procedure was then iterated until each data point was computed and verified. Examples from this procedure are shown in Figure 2.1.

For a complete description of the EDQNM code see the theses of its original developers [5, 20]. No additional changes to the physics of this code were necessary to run the simulations apart from technical changes necessary to port it to a different HPC system. The most essential thing to understand about EDQNM is the nature of its approximation.

The utility of EDQNM is that it can successfully reproduce the main turbulence phenomenology e.g. K41 spectrum, without requiring massive computational expense. This is achieved by the development of a closure theory based on a statistical description of turbulence - a set of equations which allow us to compute the evolution of the energy spectrum over time, without needing to know or compute all the interactions to very high orders or resolve the velocity field exactly.



(a) *Example of slightly problematic fitting for $d = 5$ simulation data (blue) with viscosity 0.0006 Pa·s as the fit incorporates part of the steady state. Yellow highlighting aids user in understanding which part of the curve is used in the fitting (green line).*



(b) *Correctly fitted trend (green) for $d = 6$ simulation data (blue) with viscosity 0.0002 Pa·s. Here the Lyapunov exponent is negative.*

Figure 2.1 *Two different examples of graphical output used to assistance in the semi-automatic plus verification of the EDQNM simulation data to obtain Lyapunov exponents.*

This statistical approach to the velocity field involves writing it in terms of its statistical moments i.e. the second moment, which can represent energy as $\sim u^2$. For turbulent fluids one ideally wants to explore as high an order moment as possible, but one cannot explore infinitely many. One can close this never ending moments issue by the *quasi-normal approximation*, which assumes that fourth-order moments can be approximated by products of the second-order moments, thereby expressing all higher-order correlations in terms of lower ones. This of course, on its own, leads to mathematical problems due to singularities as injected energy builds over time (blow-ups), rather than dissipating to higher and higher moments.

To solve this issue a *damping mechanism* is introduced, known as eddy-damped closure. This effectively acts as a dissipation term for all scales, ensuring that energy can leave the system, preventing blow-ups. This dissipation process would be very difficult to replicate properly for all times, so a *Markovian assumption* is introduced, which means that previous states of the velocity field are ignored when calculating the dissipation i.e. the dissipation computed at the timestep excludes previous velocity moments.

In essence, the EDQNM model exactly reproduces the results of K41 theory, but there can be no guarantee that its statistical predictions remain close or even relevant to those of the full NSEs. For this reason it is sometimes referred to as an *uncontrolled approximation* to the NSEs.

To execute, the EDQNM code requires for each dimension d, the corresponding Kolmogorov constant C_d and a free parameter β to be set. This study used identical values to those in the previous research of [5], which have been well-tested. However, it is worth noting that values of C_d are found by running a fully developed DNS at each dimension and fitting the resulting energy spectrum, which is only possible for up to $d = 4$. For $d > 4$, the parameters were taken from Lagrangian renormalised approximation simulations (another type of closure approximation not reliant on extra parameters), thus these higher dimensions are considered slightly less reliable. From the values of C_d it is possible, using numerical integration, to compute β, the details of which can be found on page 76 of [5].

Table 2.1 *EDQNM simulation fixed parameters used for each dimension, from [5].*

d	C_d	β
3	1.72	0.49
4	1.33	0.366
5	1.16	0.28
6	1.08	0.23

2.3 Results

In total, I carried out 30 simulations in 3D for $6 \leqslant \mathrm{Re}_L \leqslant 700$ on $24^3 - 512^3$ grid points, 40 simulations in 4D with $7 \leqslant \mathrm{Re}_L \leqslant 235$ on $24^4 - 256^4$ grid points, 39 simulations in 5D with $6 \leqslant \mathrm{Re}_L \leqslant 98$ on $24^5 - 256^5$ grid points, and 19 simulations in 6D with $6 \leqslant \mathrm{Re}_L \leqslant 25$ on $24^6 - 48^6$ grid points. In Appendix B the complete sets of simulation parameters necessary to initialise and re-run the simulations are presented. It is worthwhile noting that these simulations include the first reported DNS of the NSEs in dimensions higher than five.

Firstly, in Figure 2.2 (left) I examine the dependence of the average maximal Lyapunov exponent, λ on the Reynolds number with dimension. One sees that as *dimension increases* the *Lyapunov exponents measured decreases* and that below a critical Reynolds number, different for each dimension, the average maximal Lyapunov exponent is negative. This qualitatively validates the same conclusions found previously using EDQNM [21, 22].

For six dimensions I do not find a positive Lyapunov exponent but the Lyapunov exponents were found to still increase with increasing Reynolds number. Due to the box size limiting the maximum Reynolds number in six dimensions, it is not possible to test directly the EDQNM result in Figure 2.2 (right), that the six dimensions chaos curve flattens off at high Reynolds number. As dimension increases I find that the critical Reynolds number at which the Lyapunov exponent is zero, increases. This is consistent with the results from the EDQNM [22]. Quantitatively for all dimensions I find that at low Reynolds numbers, $\lambda \sim \log(Re)$.

At high Reynolds number, for three and four dimensions the Ruelle relation (see section 1.2.2) begins to be recovered. The deviation from the Ruelle relation at low Reynolds number is expected as at low Reynolds number there is no longer a scaling range. Hence there ceases to be separation between the viscous and

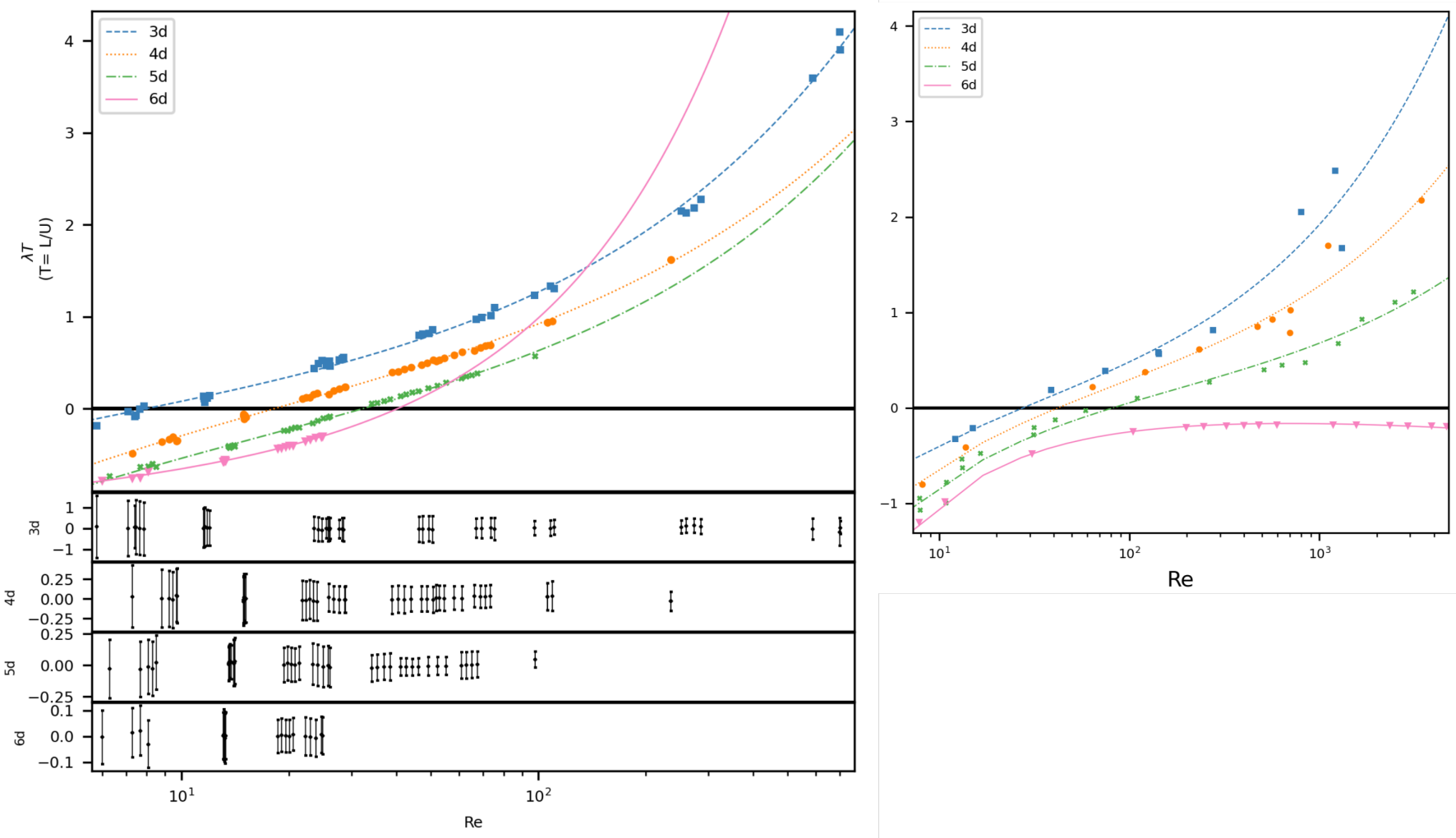


Figure 2.2 *Lyapunov exponent λT versus integral scale Reynolds number, Re for multiple dimensions at low Re measured from DNS of NSEs on left and EDQNM on right. Residuals are for the linear regressions fitted to DNS with the error bars indicating the magnitude of the fluctuations in the FTLE. Data points are indicated by blue squares, orange circles, green crosses and pink triangles for 3, 4, 5 and 6 dimensions respectively. Fits to each of these are the same colour but dashed, dotted, dot-dashed and solid for 3, 4, 5 and 6 dimensions respectively.*

large-eddy timescales which is required for Ruelle's prediction that the Lyapunov exponent goes as the inverse of *only* the viscous timescale.

Also noticeable is that this figure indicates a non-zero negative intercept implying that $\lambda \neq 0$ in the limit $Re \to 0$. I conjecture that this behaviour is expected in whichever way one examines the limit. The NSEs in the distinguished limit of $Re \to 0$ reduce to the Stokes equations which are a stable laminar flow [122], $\lambda < 0$ and at $\nu = \infty$ they reduce to a "solid" (static velocity field). I speculate that in theory such an object could take an infinitely negative Lyapunov exponent (division by zero in equation 1.15), and for a real system it would be negative finite with a large magnitude.

One way of interpreting the results in the low Reynolds number regime is to consider when Re is small, the Ruelle relation expanded as a power series to first order such that $\lambda \sim \alpha \log(Re)$. This is motivated by the fact that in the distinguished limit $Re \to 0$, the NSEs become the Stokes equations [122]. The main issue then is adjusting the Ruelle relation, such that one retains α close to 0.5 while also having $\lambda \neq 0$ in the limit $Re \to 0$.

To fit this data I embed the low Reynolds number non-chaotic, $\lambda < 0$ regime in the higher orders of the dimensionless dissipation rate as derived by [78]. Their expression for the dimensionless dissipation rate is derived by taking the Kármán–Howarth equation and performing an asymptotic expansion in the second and third-order structure functions in powers of the inverse Reynolds number. This gives,

$$C_\epsilon(Re) = C_{\epsilon,\infty} + \frac{C}{Re} + \mathcal{O}\left(\frac{1}{Re^2}\right) \tag{2.10}$$

where C is a dimensionless constant, independent of Re. Following the derivation in [52], this yields a new form of the Ruelle relation,

$$\lambda T_0 = \frac{2D_E}{3} Re^\alpha C_\epsilon^{(1)}(Re) \tag{2.11}$$

where D_E is a dimensionless constant and $C_\epsilon^{(1)}(Re)$ is truncated to first order, such that in the low Re limit $\lambda T_0 \sim log(Re)(C_{\epsilon,\infty} + \frac{C}{Re})$. Least squares fits of equation 2.11 are shown in figure 2.2 for both DNS and EDQNM data. Both $C_{\epsilon,\infty}$ and C were left as free parameters. Residuals shown only for the DNS data do not indicate significant systematic fitting bias. Given the low range of the Reynolds number, I would not want to overstate any numerical significance of

these fits. However, these fits obtain α of 0.54, 0.50, 0.58 and 0.88 in three to six dimensions respectively. The purpose of the fits is to illustrate the general trends more clearly.

The velocity derivative skewness, S_0 was also measured as

$$S_0 = -\frac{\left\langle (\partial_x u_x)^3 \right\rangle}{\left\langle (\partial_x u_x)^2 \right\rangle^{3/2}}, \tag{2.12}$$

defined as the negative for convenience. This was measured because it is a reliable microscale measure which relates the second- and third-order moments and thus relates the energy transfer and the energy dissipation ratio. It is easier to see these relationships by writing equation 2.12 using the structure function definition we used in the introduction, equation 1.9 giving,

$$S_0 = \lim_{r \to 0} \left(-\frac{S_3(r)}{S_2(r)^{3/2}} \right). \tag{2.13}$$

Higher values of S_0 are associated with increased vortex stretching, which provides a mechanism for the increased forward energy transfer in higher dimensions as found in previous studies using both DNS and EDQNM [12, 21, 22]. Velocity derivative skewness was not measured in the EDQNM model as this approximation is exactly compatible with K41; one simply observes the K41 theoretical result that it scales asymptotically with increasing Reynolds number.

In Figure 2.3 I show the dependence of S_0 on the Taylor Reynolds number Re_λ, which is defined by using the Taylor microscale. This characteristic length scale corresponds to the leading order quadratic term found when Taylor expanding the two-point velocity autocorrelation function. Taylor Reynolds number is commonly used with S_0 measurements because it characterises turbulence at intermediate scales where velocity gradients dominate, it being smaller than the integral scale and larger than the Kolmogorov microscale where dissipation occurs. My results shows that increased Taylor Reynolds number correlates with increased skewness in three and four dimensions. This is in keeping with previous studies at higher Reynolds number already mentioned [12, 21, 22]. In five dimensions skewness appears to plateau with higher Reynolds number, consistent with the K41 theory for scaling at high Reynolds numbers [56]. In six dimensions the data is more limited, but there are signs that the skewness may scale in a similar way to five dimensions, possibly plateauing at higher Reynolds numbers but without further data it is not possible to be definitive.

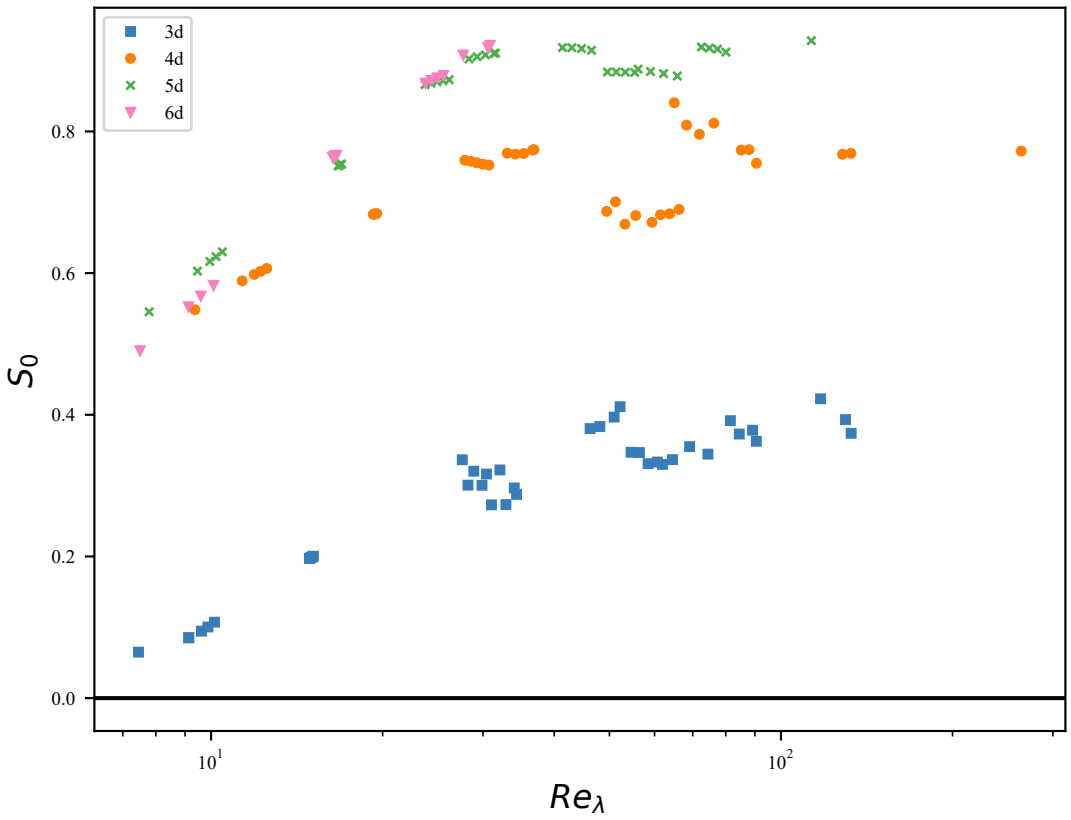


Figure 2.3 *Velocity derivative skewness versus Taylor Reynolds number, Re_λ for three, four, five and six dimensions.*

In critical phenomena, above the upper critical dimension, mean field theories become exact close to the critical point. Hence, if there exists such a dimension for the NSEs then within the inertial range one should find true scale invariance, with a reduction in fluctuations. Given that the skewness results indicate increased forward energy transfer, one would expect a reduction in the fluctuations with increasing dimension.

Figure 2.4 shows the variation of $\Delta E/(\epsilon\nu)^{1/2}$ with integral scale Reynolds number, where ΔE is the standard deviation of the total energy in time for each simulation. The magnitude of the fluctuations decreases with increasing dimension. The gradients decrease with increasing dimension, except in the case of six dimensions. Due to the log-log nature of the plot, the reduction in fluctuations with dimension appears to be greater than simply $1/d$, which would be the central limit expectation and also what was derived by Kraichnan [59]. However more data at higher Reynolds number would be needed to quantify this more accurately. I also checked that the reduction in fluctuations with dimension

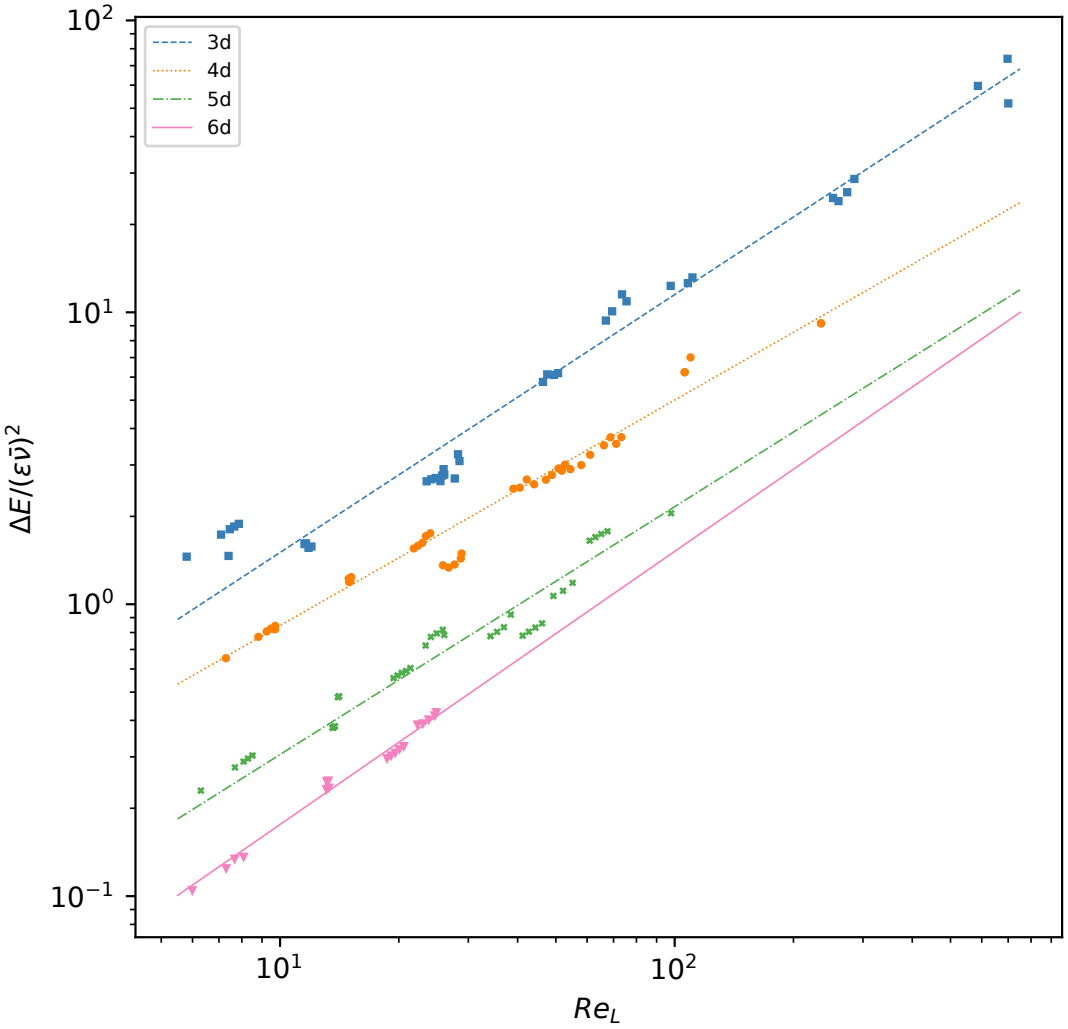


Figure 2.4 *Variation of energy normalised at the microscale, with integral scale Reynolds number in three, four, five and six dimensions.*

was not due to reduction in the total energy.

2.4 Discussion

This study is the first to report DNS results in six dimensions for the forced NSEs. These simulations are complemented by simulations in three, four and five dimensions, and by an EDQNM. The only previous studies to examine the relationship between chaos and dimension in the NSEs have been EDQNM studies. Previous DNS studies have been in four and five dimensions using free decay [46, 88, 116, 135] and one forced simulation of fully developed turbulence

[12]. Previous EDQNM results have indicated six dimensions as being an upper critical dimension for turbulence in the NSEs. This finding is consistent with the theoretical prediction by [70]. My results, especially for chaos, are consistent with these trends. The trends at low and high Reynolds numbers are consistent for all dimensions, except that in six I do not have high Reynolds number DNS data.

Recent work in holography showed that forced black hole horizons exhibit statistically steady turbulent spacetime dynamics consistent with K41 for infinite spacetime dimensions [3]. This is consistent with theoretical arguments for the existence of a critical dimension for K41 [41, 60, 87]. If Liao's prediction of $d_c = 6$ is correct, then it may be that at and above this dimension, K41 becomes an exact theory - analogous to the $d_c = 4$ crossover dimension of the Ising model [130]. Exploring the relationship between Liao's prediction and these holography results may shed light on the matter. The trends in our results from DNS and EDQNM show consistency with the Liao fixed point.

If one takes a broader dynamical systems view of the NSEs as advocated by Ruelle [105], then consider an alternative form of the NSEs

$$\frac{d\boldsymbol{u}}{dt} = X_\mu(\boldsymbol{u}) \tag{2.14}$$

where X_μ is a vector field on a manifold of finite dimension and μ is a control parameter which may be the Reynolds number, for example, and is restricted to some domain V. Ruelle points out that "the fluid motion is expected to become irregular and chaotic when μ increases", whilst this might be true if Reynolds number is varied, it is not the case for varying the system dimension according to [21]. Ruelle clarifies later more formally[3] that changing μ will change the structure of the attracting set for time evolution in the phase-space, which he thus conjectures will lead to increased chaos, which is consistent with this preliminary simulation study. Although Ruelle does not consider the possibility of a critical dimension, his point that changing the dimension will fundamentally change the attractor structure is congruent with this idea, if the attractor change involves some discontinuity at some critical dimension.

We have used chaos to measure the transition with increasing dimension. It has been shown that DNS measures of chaos do not vary strongly with lattice size

[3]"The attracting set thus changes its qualitative nature when μ crosses sets of V called bifurcation sets." [105]

[52]. EDQNM and DNS agree on chaos measures up to largest Reynolds numbers of DNS for which I have data in three, four and five dimensions, as in this work and in [21, 22]. My data shows consistency with both low and high Reynolds numbers especially in three and four dimensions, and signs of the same trends in five dimensions. This is interesting when one considers that there are no specific reasons why the DNS and EDQNM should chaotically present similar results. While the EDQNM approximation is useful, it is most definitely not the same dynamical system as the NSEs although it possesses many of the same features. There is presently no evidence that they occupy the same phase-space apart from phenomenological similarities such as the K41 spectrum. As it is closed at the third order, all higher moments and information within them is lost. It is this information which is important for resolving issues like intermittency effects. It is also worth noting that EDQNM is an "uncontrolled" approximation to the NSEs, in that there are no mechanisms to ensure that it is statistically close to the NSEs [5].

The results are a preliminary study as further work is required without using chaotic measures to be certain of a substantive difference at six dimensions. If there really exists a major change going from the lower dimenions to six, then there should be other evidence of this, for example in the spectral quantities. The spectral values of the energy difference field need to be explored, both in their geometry and growth with time. The difficulty is that these kinds of simulations require significantly more computing power than the FTLE method. A future direction could be to explore some of the additional findings in [21], such as that above a critical dimension (between five and six), the inverse error[4] cascade only reaches the dissipation scale, while below this dimension, errors cascade to larger scales. Using DNS, it might be possible to observe this effect by not using FTLEs and instead measuring the difference field over long times. This would require further code development so as to extract error spectra (how error is distributed across different scales) and transfer terms[5].

[4]"Error" is a slightly confusing use of terminology in the turbulence community here - it is not really an error in the conventional experimental sense, but rather just a measure of how sensitive the different components of the energy are to small perturbations i.e. sensitivity to initial conditions.

[5]These describe the exchange of energy or error between scales, telling us whether energy or error moves toward smaller scales (direct cascade) or larger scales (inverse cascade) within the turbulent flow. In other words, which direction one moves on Figure 1.2.

2.5 Conclusion

In my view, physically, these results in tandem with [12] give the following picture. When one increases dimension, for angular momentum to be conserved, the vortex stretches in the additional dimensions, resulting in reduced vorticity. This transfers the fluctuation energy away via the nonlinear term until the point when the velocity gradient is sufficiently large that the fluctuations are dissipated by viscosity. One would expect FTLEs to be especially sensitive to this process, as they act as Lagrangian tracers at the smallest scales of the simulation. This would explain why one sees changes in energy fluctuations and chaos without needing to capture a complete inertial range.

Individually, the results are inconclusive but the aggregate evidence begins to point to some type of change of behaviour in six dimensions. This gives new impetus to notions of some upper critical dimension for fluid turbulence. Results consistent with this outcome have been found in earlier EDQNM studies. In six dimensions the DNS results do not reach sufficient Reynolds number to confirm a transition. However, there is good agreement found between the DNS and EDQNM results at the lower dimensions, and there is no reason to expect divergence from this at higher dimensions. Since there are no theoretical methods to effectively study the idea of an upper critical dimension for turbulence or any other phase changes, it has waited for sufficient computational power to test directly. This chapter establishes the first such test, the results of which motivate more conclusive computational experiments in the coming decades once sufficient resources are available. In the next chapter, I continue to explore the use of ensembles in dynamical systems, taking a higher level view of the most well-known turbulent system: the atmosphere.

Chapter 3

A hypothesis on ergodicity and the signal-to-noise paradox

It was one of those heavy, sultry afternoons when nature seems to be saying to itself, 'Now, shall I, or shall I not, scare the pants off these people with a hell of a thunderstorm?' –P. G. Wodehouse [131]

This publication originally appeared in Atmospheric Science Letters Volume 25, Issue 11 on 4th November 2024 DOI:10.1002/asl.1265.

3.1 Introduction

In an ideal world, with a very big supercomputer, one would use exactly the same DNS I have introduced in the previous Chapter to forecast the weather, as the atmosphere is just a fluid. As we saw though, even a well-resolved small box in three-dimensions requires enormous computing power, hence such an approach is unfeasible computationally. Instead, we must rely on coarser grain models. However, one can still use the same perturbation approach to generate an ensemble to estimate how chaotic and thus how uncertain a particular forecast is. An outcome of the ensemble systems approach to weather prediction is that the long-range forecasts, particularly for the North Atlantic region, suffer from a signal-to-noise paradox [108, 128]. This is where it has been shown that the models are better at predicting reality than themselves. This chapter introduces my hypothesis that ergodicity might explain the paradox. In applying ergodic theory to ensemble forecasting systems, I show that using the ensemble mean as our best forecast of observations means that one interprets it as the most likely phase-space trajectory. Such an interpretation relies on the ergodic theorem, which can fail for ensemble systems under three conditions. Systematically addressing each of these, I argue that they can fail for North Atlantic Oscillation forecasts, potentially giving rise to the paradox.

Received: 8 December 2023 | Revised: 16 July 2024 | Accepted: 19 July 2024

DOI: 10.1002/asl.1265

RESEARCH ARTICLE

Atmospheric Science Letters RMetS

A hypothesis on ergodicity and the signal-to-noise paradox

Daniel J. Brener

The Higgs Centre for Theoretical Physics, The University of Edinburgh, Edinburgh, UK

Correspondence
Daniel J. Brener, The Higgs Centre for Theoretical Physics, The School of Physics & Astronomy, The University of Edinburgh, Edinburgh, EH9 3FD, UK.
Email: daniel.brener@ed.ac.uk

Funding information
Science and Technology Facilities Council, Grant/Award Number: ST/W507441/1

Abstract

This letter raises the possibility that ergodicity concerns might have some bearing on the signal-to-noise paradox. This is explored by applying the ergodic theorem to the theory behind ensemble weather forecasting and the ensemble mean. Using the ensemble mean as our best forecast of observations amounts to interpreting it as the most likely phase-space trajectory, which relies on the ergodic theorem. This can fail for ensemble forecasting systems if members are not perfectly exchangeable with each other, the averaging window is too short and/or there are too few members. We argue these failures can occur in cases such as the winter North Atlantic Oscillation (NAO) forecasts due to intransitivity or regime behaviour for regions such as the North Atlantic and Arctic. This behaviour, where different ensemble members may become stuck in different relatively persistent flow states (intransitivity) or multi-modality (regime behaviour), can in certain situations break the ergodic theorem. The problem of non-ergodic systems and models in the case of weather forecasting is discussed, as are potential mitigation methods and metrics for ergodicity in ensemble systems.

KEYWORDS

ensembles, ergodicity, NAO, signal-to-noise paradox, seasonal forecasting

1 | INTRODUCTION

The equations governing the atmosphere were shown in the seminal papers of Lorenz (1963) and Lorenz (1969) to exhibit sensitive dependence on initial conditions, commonly known as chaos. Given the impossibility of perfectly knowing these conditions, predictability of the atmosphere is inherently limited, even with a perfect model. To address this uncertainty, the method of ensemble weather prediction was developed, see Epstein (1969). Using multiple realizations of a numerical model with tiny variations in initial conditions, these small perturbations give rise to divergent predictions due to chaos. The average value of an observable over several ensemble members can sometimes be interpreted as the best estimate of the said state of the observable at or over a particular time. The mean and standard deviation of predictions across ensemble members provide estimates of an observable's state and uncertainty, despite the deterministic equations underlying the models, as discussed in Murphy and Palmer (1986) and Palmer and Hagedorn (2006).

Scaife and Smith (2018) reviewed what has come to be known as a signal-to-noise paradox: that atmosphere–

Abbreviations: GloSea5, Met Office Global Seasonal Forecast System, version 5; NAO, North Atlantic Oscillation; PDF, probability density function; RPC, ratio of predictable components.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2024 The Author(s). *Atmospheric Science Letters* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

ocean coupled climate and long range prediction ensemble models are better at predicting reality, than they are at predicting themselves. This was first raised as possibility by Kumar (2009) and consequently demonstrated by Eade et al. (2014). They derived a statistic for comparing ensemble models to observations known as the "ratio of predictable components" (RPC) for several atmospheric parameters,

$$\mathrm{RPC}^2 = \frac{r^2_{\overline{\mathrm{E}}\mathrm{O}}}{r^2_{\overline{\mathrm{E}}\mathrm{E}_i}}, \tag{1}$$

where $r_{\overline{\mathrm{E}}\mathrm{O}}$ is the Pearson correlation between the model ensemble mean and the observations, and $r_{\overline{\mathrm{E}}\mathrm{E}_i}$ is the average correlation between the model ensemble mean and a single ensemble member. It was found that the correlation between the ensemble mean and observations is often much greater than the average correlation between the ensemble mean and a single ensemble member $(\mathrm{RPC} > 1)$, referred to as an anomalous RPC. Hence, a signal-to-noise paradox - the model predicts reality better than it predicts itself. This has since been reported by many different groups such as Scaife et al. (2014) Stockdale et al. (2015), Charlton-Perez et al. (2019), Weisheimer et al. (2019) and Dunstone et al. (2023).

Over the past few years, a great deal of effort from the community has been put into finding different resolutions to the paradox, and different scenarios where it arises. Strommen et al. (2023) showed that for the North Atlantic Oscillation (NAO), the paradox may be interpreted as a probabilistically under-confident forecast for occurrences of high magnitude NAO. A possible explanation for this is that the model has reduced persistence of particular atmospheric regimes, especially in the Northern Hemisphere according to Strommen and Palmer (2018) and Strommen (2020). In terms of model dynamics, weak atmosphere eddy feedback was suggested by Scaife et al. (2019) and Hardiman et al. (2022), or weak ocean–atmosphere coupling in models by Ossó et al. (2020). Early approaches to the paradox hoped that enhancing model physics could directly improve the correlation between the model and the ensemble mean, thereby increasing forecast skill. These improvements have not yet been achieved. A detailed summary of the current state of understanding and avenues to tackle the problem can be found in Weisheimer et al. (2024). Taking a different angle, Bröcker et al. (2023) argued that in some cases, that the signal-to-noise paradox should not be considered paradoxical due to the assumption that the forecast error is related to the correlation of the ensemble mean with the observations, which is not necessarily always the case. It is in this more statistical direction that this Letter takes us.

We suggest that part of the paradox could be due to a violation of the ergodic theorem which is relied upon for the statistical moments comprising the RPC. This is because using the ensemble mean as our best forecast of observations amounts to interpreting it as the most likely phase-space trajectory, which relies on the ergodic theorem. We argue that this can fail in cases such as the winter NAO forecasts due to intransitivity/multi-modality which can in certain situations break the ergodic theorem. The ideas presented are not intended to resolve the paradox entirely as there are other contributing issues related to both modelling and the definitions used in the computation of the RPC (e.g. Bröcker et al. (2023), Zhang and Kirtman (2019), Knight et al. (2022) and O'Reilly et al. (2019)). We hope this article will act as a catalyst for others to consider how the ergodicity assumption might be tested more rigorously, and serve as a reminder of it.

2 | ERGODIC THEORY

2.1 | The ergodic theorem

Consider a dynamical system, x_t evolving with time t on some phase space $X = \mathbb{R}^d$, where d is the system dimension. Suppose there exists some observable $f(x_t) \in X$, which can be integrated such that a measure μ on X is preserved. Let T be a unique time-evolution operator such that an initial state x_0 can evolve to a discrete later time x_l via $T^l x_0 = x_l$, $l \in \mathbb{N}$. If T preserves the invariant measure μ, then T is said to be ergodic. Our observable f has a *time mean of a single path* on X going from an initial position (x_0, t_0) to a later position (x_l, t_l) given by,

$$\bar{f}(x_0) = \lim_{l\to\infty} \frac{1}{l} \sum_{k=0}^{l-1} f\left(T^k x_0\right). \tag{2}$$

For the same observable f, one can also find the *space mean*,

$$\bar{f}(x_{t'}) = \int \mu f(x_{t'}) dx \tag{3}$$

at an instant in time t'. Under certain conditions on the system, the ergodic theorem holds that: the time and space means are the same for all possible initial conditions x_0. This is also known as the ergodic hypothesis and was proved, along with the existence of the averages individually, by Birkhoff (1931). See Ollagnier (1985) for a

formal approach to the subject. In the next section, we will see how this works for ensemble models.

2.2 | Application to ensemble modelling

In numerical weather prediction, x_t is the state of the atmosphere at time t, evolving in some phase space R^d, where d is very large. The operator T represents the evolution of the atmosphere forward in time according to all the equations of physics. The phase space of the various atmospheric/climatic states is some strange attractor, $\mathcal{A}$ see Tsonis and Elsner (1989) for an introduction. T takes us between states on $\mathcal{A}$, whilst preserving $\mathcal{A}$. If T is chosen to be ergodic, then T will visit all parts of $\mathcal{A}$ over time; for a rigorous treatment see Eckmann and Ruelle (1985) part 2E.

The ensemble mean of an observable, $E(f(x_t))$ at a time t is

$$E(f(x_t)) = \frac{1}{n}\sum_{i=1}^{n} f^i(x_t) \tag{4}$$

where $f^i(x_t)$ is the observable from the i^{th} member of the ensemble of size n. The significance of the ensemble size, n has been well studied on all temporal and spatial scales; see Palmer and Hagedorn (2006) and Leutbecher (2019). Under a "signal + noise" model, the idea is that forecast skill grows with n due to the suppression of unpredictable noise.

In using the RPC metric, we want to interpret the ensemble mean as being the *most likely trajectory* in the phase space X, and hence our best guess for what the observations will do. This is an ergodicity assumption as we want to equate the space average (ensemble mean) with the most likely trajectory (time average). An ensemble is initialised as a set of perturbed initial states $\{x_0^i\}$. There must exist a unique time-evolution operator T_i^k for each member that evolves each initial state over a time $\Delta T \equiv T_{k=l} - T_{k=0}$. So our ensemble mean can be given by,

$$E(f(x_l)) = \frac{1}{n \cdot l}\sum_{i=1}^{n}\sum_{k=0}^{l-1} f\left(T_i^k x_0^i\right) \tag{5}$$

In other words, running our ensemble simulation is equivalent to computing the evolution of the initial states by the operator T. Since T is unique, it is more convenient to think of each ensemble member as having their own continuous time-dependant probability density function (PDF), $P^i(x_t)$, that evolves according to unique T_i^l, so $T_i^l P^i(x_0) = P^i(x_l)$. This gives the ensemble mean as

$$E(f(x_t)) = \frac{1}{n \cdot \Delta T}\int_{T_0}^{T} f^1(x_t)P^1(x_t) + f^2(x_t)P^2(x_t) + \ldots + f^n(x_t)P^n(x_t)dt, \tag{6}$$

so that the observable as measured in each member is weighted by the evolving PDF unique to each member. This average will only represent the most likely trajectory if the following conditions hold:

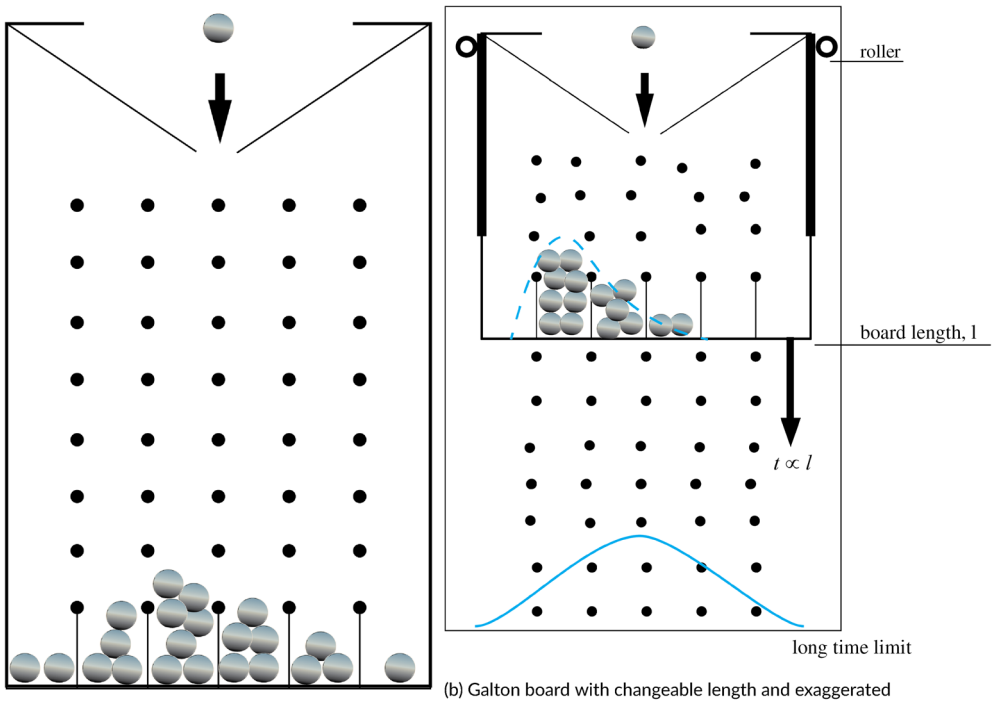
$$\begin{aligned}
&\lim n \to \infty; \text{infinite ensemble size,}\\
&\lim \Delta T \to \infty; \text{infinitely long averaging window,}\\
&\frac{dP^i}{dt}\Big|_{t=\Delta T, \forall i} = 0; \text{distribution function for each ensemble}\\
&\qquad\qquad \text{member is stationary as } T^i \text{ must be}\\
&\qquad\qquad \text{unique over period } \Delta T.
\end{aligned}$$

The final condition is equivalent to saying that the ensemble members must be *exchangeable*, which is a common assumption for ensemble prediction systems and will be discussed in a later section. We will next interrogate these concepts using a Galton Board for Gedankenexperimente.

2.3 | Ergodicity in the Galton board: Winding a classic model

A Galton board is a device invented by Galton (1889) that consists of a vertical board with an array of pegs. At the top, there is an entry point from which small balls, typically marbles, are released. The Galton board is often used to visually illustrate how randomness, combined with a large number of trials (balls), leads to a Gaussian distribution. In meteorology, it has been used to teach the principles of ensemble forecasting and has also been used to develop conceptual arguments in this field. We illustrate a normal Galton board in Figure 1a where after multiple individual balls we begin to approximate a Gaussian.

For our purposes, each individual ball trajectory represents the time evolution of a single ensemble member and the statistical distribution of the final positions represents the ensemble average for the model. As one observes more and more balls passing through the board, the distribution of final positions at the bottom converges to a pattern that resembles a Gaussian distribution. This convergence is analogous to the ergodic theorem, where the time average (individual ball's trajectory) converges to the ensemble average (statistical distribution of final positions) over a large number of trials. If we imagine an *infinitely long board*, the long-time average trajectory for



(a) Fixed length classic Galton board.

(b) Galton board with changeable length and exaggerated peg-misalignment.

FIGURE 1 Illustration of our two Galton boards. Not drawn to scale. The right-hand board has adjustable length using a roller. Adjusting the length corresponds to adjusting the time the balls have to explore the system. The longer the board, the longer the time the balls spend in the system and the better representative the ensemble mean and variance will be of the underlying probability density function (PDF), even if the space is multi-modal. Blue solid line in (b) indicates the Gaussian outcome of the long-board (time) limit, in contrast to distribution of balls (blue dashed) with some fixed short board length (time). Note that the distribution of balls is not instantaneous, but an average over a number of different balls (ensemble members) over a time proportional to the length they are allowed to drop.

a single ball will be straight down; equivalent to the position of the ensemble mean. It is also worth noting from Kumičák (2000) and Hoover and Moran (1992) that the classic Galton board has been found to possess an ergodic strange attractor, and from Judd (2007) that it is not simply a random-walk phenomena. Hence, it has useful parallels to weather and climate prediction.

We can play with this idealised ensemble model to examine what can happen if the ergodic assumption fails. The way this can happen is if the system contains multi-modality, in other words, there exist potential wells via certain trajectories into which the balls can stumble from which they are less likely to recover in a certain time-frame. This does not have to be dependent on the different initial conditions. In simulations of Galton boards such potential wells have been found to exist, see Ahmed et al. (2022).

Consider a Galton board of *finite length* where we are able to modify this length using a roller, thus change the length of time that the balls must spend in the system. Now let us suppose that the gaps between the pegs are not identically uniform, so at certain points the balls path is impeded slightly more. Locally these act as potential wells, creating multi-modal features. The longer we make the board, the more time the balls will have to explore, enter and escape these potentials. With a sufficiently long board (time), the resulting distribution, on average, will be our Gaussian again. However, the shorter the board, the more the multi-modal features dominate the picture. One can imagine all kinds of situations that could occur, such as where the balls get stuck in roughly the same region of the board and so do not explore all the possible trajectories adequately. The longer we make the board, hence the longer the time, the more Gaussian our distribution of balls will become. We also need to use enough balls (large enough ensemble) to explore the board adequately. The key point is that the underlying distributions for the space/ensemble and time averages will be different when ergodicity fails.

Our model can also be extended to incorporate external forcings. For instance, a light fan blowing across the board would shift the distributions, similar to the conceptual model of Palmer (1999). The main difference between our Galton board and the picture of Palmer (1999) is that the latter constrains the system to a bimodal structure which is examined in the long time-

limit, whereas our system allows us to imagine any number of modes (depending on the board "resolution" and complexity) and our roller lets us examine long and short time limits. One can also think about exchangeability, as if one were to drop different kinds of balls (e.g. size/mass/stickiness), the final distributions would not make sense as a predictor for any one particular ball.

We can now see that without enough balls and or if the averaging time is too short, the ergodic theorem is not satisfied. In the case of the NAO, which evolves roughly daily, then for a 90-day season in which the flow exhibits multimodality, 90 days will not always be sufficient to explore the complete phase space. This is because such ensemble statistics reflect only those states the system has visited, highlighting the impact of daily variations on the dynamics that govern long-term averages.

3 | FAILURE OF ERGODICITY AND INTERPRETATION OF THE RPC

We will now examine how failure of the ergodic assumption could lead to anomalous RPC values and hence the signal-to-noise paradox. To do this, we draw on the statistical framework developed by Bröcker et al. (2023). Let $V(e)$ and $V(o)$ be the respective variances of the ensemble and the observations, then it has been shown that,

$$\sqrt{\frac{V(e)}{V(o)}} < r_{\overline{\mathrm{EO}}} \Rightarrow \mathrm{RPC} > 1, \tag{7}$$

where $r_{\overline{\mathrm{EO}}}$ is the Pearson correlation between the observations and the ensemble mean. Bröcker et al. (2023) showed that when we have situations of small correlations, something highly likely for seasonal forecasting, then even minor differences between the observations and ensemble variance gives rise to a signal-to-noise paradox via the above inequality.

We make the conjecture that failure of ergodicity in an ensemble can in some cases lead to these differences. If the ergodic assumption fails for the mean, as defined by Equation 6, then it also fails for *all* statistical moments, including the variance $V(e)$. Mathematically, this is a strong statement as it renders such a variance as meaningless, as without ergodicity, the distribution will be different from sampling either from a long time-series of a single member or from a complete ensemble. Like in our Galton board Figure 1b, the long-time distribution is a Gaussian, but with too short a board we are more likely to get some skewed distribution. The magnitude of the difference is something which can only be determined by direct computation of the distributions. However, simply failing to respect ergodicity does not mean that $V(e)$ should systematically underestimate $V(o)$.

Let us try a thought experiment with our windable Galton board in Figure 1b. Suppose the observational reality is that the ball falls straight down, round a zero-mean Gaussian. We could conceive of a situation where our ensemble of balls gets stuck in a potential well such that they clump together relatively close to the centre. This would result in $V(e) < V(o)$ with a large $r_{\overline{\mathrm{EO}}}$, hence an anomalous RPC value. From the weather regime perspective of the NAO, if one has multiple persistent regimes then a single 90 day season will not always be sufficient for a finite ensemble to adequately sample the space, necessary for the ergodic theorem. The members will appear to become "stuck" in certain regimes. Consequently, each member on average then underestimates the total variance of a given season, leading to $V(e) < V(o)$.

The spatial inhomogeneity of the anomalous RPC values (see e.g. Weisheimer et al. (2019) for detailed study), especially over the North Atlantic and Arctic as reported by Cottrell et al. (2024) strongly suggest multimodality is involved, as these regions correlate with well-studied multi-modal behaviour. Strommen and Palmer (2018) point out that a multi-modal or regime approach provides an alternative perspective on the paradox. Using a toy bimodal Markov-chain model, they obtain anomalous RPC values when the persistence of the modes is underrepresented. They note that the potential wells for the regimes may be too shallow in the models. Additionally, Falkena et al. (2022) found that using more regimes improved the representation of wintertime Euro-Atlantic sector dynamics. We posit that this multimodality, in breaking the ergodic assumption necessary for the ensemble mean and variance, would consequently invalidate interpretation of those statistics, such as RPC. In fact, it leads one to reach the same conclusion as Bröcker et al. (2023), the signal-to-noise "paradox" is not really paradoxical; because ergodically we cannot expect the ensemble mean/variance to be the most likely phase-space trajectory.

Whilst our approach using the ergodic theorem to ensembles is new to this problem, we note that Lorenz discussed transitive and intransitive systems and climate, see Lorenz (1968), Lorenz (1976) and Lorenz (1990). Ergodicity and transitivity are essentially the same mathematical phenomenon (see pedagogical discussion by Shalizi (2007)). Lorenz's discussion of transitivity sets up in more general terms the concepts for climatic regime states which were later more clearly defined by Palmer (1999). The ergodic theorem therefore provides further supporting rigour and a different perspective on the multi-modality paradigm.

One mitigation strategy for the multi-modality issue is to increase the ensemble size. However, even with a relatively large ensemble/sample, the paradox has been found to persist, for example see Cottrell et al. (2024) and Shi et al. (2015). There is weak evidence that the paradox exists on longer than multi-decadal timescales Scaife and Smith (2018). The reason for this may be understood from an ergodic perspective, as to define a mean climate state the interval ΔT over which we compute the average, is sufficiently long that the distributions P^i converge Tantet (2016). One could increase the averaging window to see how this affects the RPC. However one cannot increase it too much else the definition of the observable would be changed. In other words, for very large averaging windows, the observable simply becomes a different kind of climatic average: sub-seasonal to seasonal, seasonal to annual, etc.

Another way in which failure of ergodicity can occur is if the ensemble members are not exchangeable with each other. This can be seen from Equation 6, as one cannot factorise a single $P^i(x_t)$ in the integrand unless they are all *perfectly* exchangeable. According to Leutbecher (2019), the ensemble members of the operational ECMWF model, which has reported anomalous RPCs, are not truly exchangeable because its initial perturbations have a plus-minus symmetry (i.e., initial perturbation of member $2k$ is minus the perturbation of member $(2k-1)$), introducing systematic differences between ensemble members. Other forecasting systems which are known to not satisfy exchangeability include the Meteo-France global ensemble, and Canadian ensemble prediction system, see Leutbecher et al. (2017). Also note that Siegert et al. (2016), see section 4b, raised the possibility that the Met Office's GloSea5 climate prediction system ensemble members might not be exactly exchangeable. As reported by Weisheimer et al. (2024), it is understood that exchangeability of observations and ensemble members provides a strong criterion for the signal-and-noise paradox. What is not recognised explicitly is that the members themselves must also be exchangeable to avoid failure of the ergodic theorem when interpreting the ensemble mean/variance as the most likely trajectory and then using these for the RPC. Combined with multi-modality, non-exchangeability could make it more likely for situations of $V(e) < V(o)$ to occur as we need to correctly weight each member according to its uniquely evolved PDF. Members might appear "stuck" - in contrast to their evolved PDFs, which if used as weights would make the ensemble mean and variance better estimates of the most likely phase-space trajectory.

If we look at other simpler physical systems we readily find cases where failure of the ergodic assumption leads to significant incorrect statistical measures. Examples of such systems include supercooled liquids to glass transitions (see Thirumalai et al. (1989)), diffusive processes within the plasma membrane of living cells (see Weigel et al. (2011)) or general noise processes (see Mangalam and Kelty-Stephen (2022)). In each of these cases, handling ensemble averages without due attention to the ergodic assumption leads to in simple terms, wrong answers. Peters and Klein (2013) demonstrates an example of this in geometric brownian motion where nonergodicity can lead to the ensemble mean growing exponentially, whilst simultaneously any individual trajectory decays exponentially according to its time average.

4 | SUGGESTIONS FOR FUTURE STUDIES

Future work to examine the hypotheses in this Letter could include testing the Thirumalai–Mountain effective ergodic convergence metric, $\Omega(t)$ measures the effective ergodicity by the difference between the time average of an observable and its ensemble average over the entire system. So for an ensemble of size n,

$$\Omega(t) = \frac{1}{n}\sum_{i=1}^{n}\left[f(t)_i - E(f(t))\right]^2, \tag{8}$$

where the ensemble average is as defined by Equation 4. In the long-time limit, $\Omega \to 0$ for an ergodic system. The rate at which this occurs can be used to indicate timescales over which an ergodic approximation might be appropriate. This metric has, as far as we are aware, not been tested before in ensemble prediction systems. However, it is utilised for earthquake forecasting see for example Tiampo et al. (2003) and Tiampo et al. (2007), and studies of fluids, see de Souza and Wales (2005). For other metrics related to ergodicity, see Mathew and Mezić (2011).

Given sufficient computing resources, one could develop an "ensemble-of-ensembles" approach to test perfect exchangeability. Once could generate at each time step small perturbations to each member generating an ensemble per member, so as to compute finite time Lyapunov exponents. From this, one could compute the Lyapunov spectrum per main member. If the spectrum varies between the ensemble members then this would imply that fundamentally the members evolve into different state spaces, and thus are not exchangeable, hence failure of the ergodic mean to represent the most likely trajectory.

An approach to relax the ergodicity constraint but still make use of ensembles is to enforce dynamical invariants. These are quantities that are conserved along the phase-space path of the system. For example, in the probability theory for non-equilibrium gravitational systems developed by Peñarrubia (2015), dynamical invariants are used with the explicit purpose of not relying ergodicity assumptions. For turbulent fluids, the most obvious dynamical invariant would be the Lyapunov spectrum and fractal dimension, besides vorticity. Recently Platt et al. (2023) developed such a scheme for a numerical weather model machine learning training method. They tested their model on the classic Lorenz 1996 chaotic dynamical system, and found including ergodic constraints improved forecast skill.

5 | SUMMARY

The signal-to-noise paradox, assessed by the RPC metric, relies on interpreting ensemble mean and variance as the most probable phase-space trajectory. However, this interpretation relies on the ergodic theorem, which faces potential failure due to insufficient ensemble size, short averaging windows, and non-exchangeability of members. In the cases like winter NAO forecasts, evidence of multi-modal regime behaviour or intransitivity between NAO phases, irrespective of multi-modality, could exacerbate this failure, leading to hindered exploration of the complete phase space. Consequently, the ensemble mean and variance cannot reliably represent the most likely trajectory, resulting in the paradox when the ensemble becomes stuck – correlating with observations but underestimating the variance.

AUTHOR CONTRIBUTION

Daniel J. Brener: Conceptualization; formal analysis; investigation; methodology; visualization; writing – original draft preparation, writing – review & editing; project administration.

ACKNOWLEDGEMENTS

D.J.B. was funded by the UK Science and Technology Facilities Council. He was particularly grateful to Professor Adam Scaife who introduced him to the problem, as well as Dr Jose Rodriguez and Dr Leon Hermanson of the Met Office and Professor Jorge Peñarrubia of The University of Edinburgh, for giving their time to insightful tête-à-tête discussions. He also thanks Alexander Tully of The University of Liverpool for comments on the manuscript. D.J.B. expresses his gratitude to two anonymous reviewers and Associate Editor Dr. Steven Hardiman for their invaluable comments and feedback.

FUNDING INFORMATION

Science and Technology Facilities Council, UK.

CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

DATA AVAILABILITY STATEMENT

No data was used or produced in the course of this research.

ORCID

Daniel J. Brener https://orcid.org/0000-0003-3340-0177

REFERENCES

Ahmed, J., Chumley, T., Cook, S., Cox, C., Grant, H., Petela, N. et al. (2022) Dynamics of the no-slip Galton board. Available at: https://arxiv.org/abs/2208.07790

Birkhoff, G.D. (1931) Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 17, 656–660 Available at: http://www.jstor.org/stable/86016

Bröcker, J., Charlton–Perez, A.J. & Weisheimer, A. (2023) A statistical perspective on the signal-to-noise paradox. *Quarterly Journal of the Royal Meteorological Society*, 149, 911–923. Available from: https://doi.org/10.1002/qj.4440

Charlton-Perez, A., Bröcker, J., Stockdale, T.N. & Johnson, S.J. (2019) When and where do ECMWF seasonal forecast systems exhibit anomalously low signal-to-noise ratio? *Quarterly Journal of the Royal Meteorological Society*, 145, 3466–3478. Available from: https://doi.org/10.1002/qj.3631

Cottrell, F.M., Screen, J.A. & Scaife, A.A. (2024) Signal-to-noise errors in free-running atmospheric simulations and their dependence on model resolution. *Atmospheric Science Letters*, 25, e1212. Available from: https://doi.org/10.1002/asl.1212

de Souza, V.K. & Wales, D.J. (2005) Diagnosing broken ergodicity using an energy fluctuation metric. *The Journal of Chemical Physics*, 123, 134504. Available from: https://doi.org/10.1063/1.2035080

Dunstone, N., Smith, D.M., Hardiman, S.C., Hermanson, L., Ineson, S., Kay, G. et al. (2023) Skilful predictions of the summer north Atlantic oscillation. *Communications Earth & Environment*, 4, 409. Available from: https://doi.org/10.1038/s43247-023-01063-2

Eade, R., Smith, D., Scaife, A.A., Wallace, E., Dunstone, N., Hermanson, L. et al. (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophysical Research Letters*, 41, 5620–5628. Available from: https://doi.org/10.1002/2014gl061146

Eckmann, J.P. & Ruelle, D. (1985) Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57, 617–656. Available from: https://doi.org/10.1103/RevModPhys.57.617

Epstein, E.S. (1969) Stochastic dynamic prediction. *Tellus*, 21, 739–759.

Falkena, S.K., de Wiljes, J., Weisheimer, A. & Shepherd, T.G. (2022) Detection of interannual ensemble forecast signals over the North Atlantic and Europe using atmospheric circulation regimes. *Quarterly Journal of the Royal Meteorological Society*, 148, 434–453. Available from: https://doi.org/10.1002/qj.4213

Galton, F. (1889) *Natural inheritance*. No. v. 42; v. 590. London: Macmillan.

Hardiman, S.C., Dunstone, N.J., Scaife, A.A., Smith, D.M., Comer, R., Nie, Y. et al. (2022) Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions. *NPJ Climate and Atmospheric Science*, 5, 57.

Hoover, W.G. & Moran, B. (1992) Viscous attractor for the Galton board. *Chaos*, 2, 599–602.

Judd, K. (2007) Galton's quincunx: random walk or chaos? *International Journal of Bifurcation and Chaos*, 17, 4463–4469.

Knight, J.R., Scaife, A.A. & Maidens, A. (2022) An extratropical contribution to the signal-to-noise paradox in seasonal climate prediction. *Geophysical Research Letters*, 49, e2022GL100471. Available from: https://doi.org/10.1029/2022GL100471

Kumar, A. (2009) Finite samples and uncertainty estimates for skill measures for seasonal prediction. *Monthly Weather Review*, 137, 2622–2631.

Kumičák, J. (2000) Galton board as a model for fluctuations. In *Unsolved Problems of Noise and Fluctuations: UPoN'99: Second International Conference*, vol. 511 of *American Institute of Physics Conference Series*, 144–149.

Leutbecher, M. (2019) Ensemble size: how suboptimal is less than infinity? *Quarterly Journal of the Royal Meteorological Society*, 145, 107–128. Available from: https://doi.org/10.1002/qj.3387

Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S.T.K., Balsamo, G., Bechtold, P. et al. (2017) Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, 143, 2315–2339. Available from: https://doi.org/10.1002/qj.3094

Lorenz, E.N. (1963) Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20, 130–141.

Lorenz, E.N. (1968) *Climatic determinism*. Boston, MA: American Meteorological Society, pp. 1–3. Available from: https://doi.org/10.1007/978-1-935704-38-6_1

Lorenz, E.N. (1969) The predictability of a flow which possesses many scales of motion. *Tellus*, 21, 289–307.

Lorenz, E.N. (1976) Nondeterministic theories of climatic change. *Quaternary Research*, 6, 495–506 Available at: https://www.sciencedirect.com/science/article/pii/0033589476900223

Lorenz, E.N. (1990) Can chaos and intransitivity lead to interannual variability? *Tellus A*, 42, 378–389. Available from: https://doi.org/10.1034/j.1600-0870.1990.t01-2-00005.x

Mangalam, M. & Kelty-Stephen, D.G. (2022) Ergodic descriptors of non-ergodic stochastic processes. *Journal of the Royal Society Interface*, 19, 20220095. Available from: https://doi.org/10.1098/rsif.2022.0095

Mathew, G. & Mezić, I. (2011) Metrics for ergodicity and design of ergodic dynamics for multi-agent systems. *Physica D: Nonlinear Phenomena*, 240, 432–442 Available at: https://www.sciencedirect.com/science/article/pii/S016727891000285X

Murphy, J. & Palmer, T. (1986) Experimental monthly long-range forecasts for the united-kingdom. 2. A real-time long-range-forecast by an ensemble of numerical integrations. *Meteorological Magazine*, 115, 337–349.

Ollagnier, J.M. (1985) *Ergodic theory and statistical mechanics, vol. 1115 of lecture notes in mathematics*. Berlin, Heidelberg: Springer.

O'Reilly, C.H., Weisheimer, A., Woollings, T., Gray, L.J. & MacLeod, D. (2019) The importance of stratospheric initial conditions for winter north Atlantic oscillation predictability and implications for the signal-to-noise paradox. *Quarterly Journal of the Royal Meteorological Society*, 145, 131–146. Available from: https://doi.org/10.1002/qj.3413

Ossó, A., Sutton, R., Shaffrey, L. & Dong, B. (2020) Development, amplification, and decay of Atlantic/European summer weather patterns linked to spring north Atlantic sea surface temperatures. *Journal of Climate*, 33, 5939–5951.

Palmer, T. & Hagedorn, R. (2006) *Predictability of weather and climate*. Cambridge: Cambridge University Press.

Palmer, T.N. (1999) A nonlinear dynamical perspective on climate prediction. *Journal of Climate*, 12, 575–591 Available at: https://journals.ametsoc.org/view/journals/clim/12/2/1520-0442_1999_012_0575_andpoc_2.0.co_2.xml

Peñarrubia, J. (2015) A probability theory for non-equilibrium gravitational systems. *Monthly Notices of the Royal Astronomical Society*, 451, 3537–3550. Available from: https://doi.org/10.1093/mnras/stv1146

Peters, O. & Klein, W. (2013) Ergodicity breaking in geometric brownian motion. *Physical Review Letters*, 110, 100603. Available from: https://doi.org/10.1103/PhysRevLett.110.100603

Platt, J.A., Penny, S.G., Smith, T.A., Chen, T.-C. & Abarbanel, H.D.I. (2023) Constraining chaos: enforcing dynamical invariants in the training of reservoir computers. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33, 103107. Available from: https://doi.org/10.1063/5.0156999

Scaife, A.A., Arribas, A., Blockley, E.W., Brookshaw, A., Clark, R.T., Dunstone, N. et al. (2014) Skillful long-range prediction of European and north American winters. *Geophysical Research Letters*, 41, 2514–2519. Available from: https://doi.org/10.1002/2014gl059637

Scaife, A.A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M. et al. (2019) Does increased atmospheric resolution improve seasonal climate predictions? *Atmospheric Science Letters*, 20, e922.

Scaife, A.A. & Smith, D. (2018) A signal-to-noise paradox in climate science. *NPJ Climate and Atmospheric Science*, 1, 28. Available from: https://doi.org/10.1038/s41612-018-0038-4

Shalizi, C. (2007) Ergodicity and metric transitivity. Available at: https://www.stat.cmu.edu/∼cshalizi/754/notes/lecture-25.pdf. [Accessed 5 July 2024]

Shi, W., Schaller, N., MacLeod, D., Palmer, T.N. & Weisheimer, A. (2015) Impact of hindcast length on estimates of seasonal climate predictability. *Geophysical Research Letters*, 42, 1554–1559. Available from: https://doi.org/10.1002/2014GL062829

Siegert, S., Stephenson, D.B., Sansom, P.G., Scaife, A.A., Eade, R. & Arribas, A. (2016) A Bayesian framework for verification and recalibration of ensemble forecasts: how uncertain is NAO predictability? *Journal of Climate*, 29, 995–1012. Available from: https://doi.org/10.1175/jcli-d-15-0196.1

Stockdale, T.N., Molteni, F. & Ferranti, L. (2015) Atmospheric initial conditions and the predictability of the arctic oscillation. *Geophysical Research Letters*, 42, 1173–1179. Available from: https://doi.org/10.1002/2014gl062681

Strommen, K. (2020) Jet latitude regimes and the predictability of the north Atlantic oscillation. *Quarterly Journal of the Royal Meteorological Society*, 146, 2368–2391.

Strommen, K., MacRae, M. & Christensen, H. (2023) On the relationship between reliability diagrams and the "signal-to-noise

paradox". *Geophysical Research Letters*, 50, e2023GL103710. Available from: https://doi.org/10.1029/2023GL103710

Strommen, K. & Palmer, T. (2018) Signal and noise in regime systems: a hypothesis on the predictability of the north Atlantic oscillation. *Quarterly Journal of the Royal Meteorological Society*, 145, 147–163. Available from: https://doi.org/10.1002/qj.3414

Tantet, A.J.J. (2016) Ergodic theory of climate: variability, stability and response. Ph.D. thesis, Utrecht University.

Thirumalai, D., Mountain, R.D. & Kirkpatrick, T.R. (1989) Ergodic behavior in supercooled liquids and in glasses. *Physical Review A*, 39, 3563–3574. Available from: https://doi.org/10.1103/PhysRevA.39.3563

Tiampo, K.F., Rundle, J.B., Klein, W., Holliday, J., Sá Martins, J.S. & Ferguson, C.D. (2007) Ergodicity in natural earthquake fault networks. *Physical Review E*, 75, 66107. Available from: https://doi.org/10.1103/PhysRevE.75.066107

Tiampo, K.F., Rundle, J.B., Klein, W., Martins, J.S.S. & Ferguson, C.D. (2003) Ergodic dynamics in a natural threshold system. *Physical Review Letters*, 91, 238501. Available from: https://doi.org/10.1103/PhysRevLett.91.238501

Tsonis, A.A. & Elsner, J.B. (1989) Chaos, strange attractors, and weather. *Bulletin of the American Meteorological Society*, 70, 14–23 Available at: https://journals.ametsoc.org/view/journals/bams/70/1/1520-0477_1989_070_0014_csaaw_2_0_co_2.xml

Weigel, A.V., Simon, B., Tamkun, M.M. & Krapf, D. (2011) Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 6438–6443.

Weisheimer, A., Baker, L.H., Bröcker, J., Garfinkel, C.I., Hardiman, S.C., Hodson, D.L.R. et al. (2024) The signal-to-noise paradox in climate forecasts: revisiting our understanding and identifying future priorities. *Bulletin of the American Meteorological Society*, 105, E651–E659 Available at: https://journals.ametsoc.org/view/journals/bams/105/3/BAMS-D-24-0019.1.xml

Weisheimer, A., Decremer, D., MacLeod, D., O'Reilly, C.H., Stockdale, T.N., Johnson, S.J. et al. (2019) How confident are predictability estimates of the winter north Atlantic oscillation. *Quarterly Journal of the Royal Meteorological Society*, 145, 140–159. Available from: https://doi.org/10.1002/qj.3446

Zhang, W. & Kirtman, B.P. (2019) Understanding the signal-to-noise paradox with a simple Markov model. *Geophysical Research Letters*, 46, 13308–13317. Available from: https://doi.org/10.1029/2019gl085159t

How to cite this article: Brener, D. J. (2024). A hypothesis on ergodicity and the signal-to-noise paradox. *Atmospheric Science Letters*, e1265. https://doi.org/10.1002/asl.1265

3.3 Conclusion

In some respects, this chapter returns to the early foundations of ensemble weather forecasting, as considering ergodicity is mathematically related to the work of Ed Lorenz on transitivity [72–74]. The central argument of the hypothesis lies on there existing different flow states in phase-space, which can be more or less transitive. One can imagine different ensemble members getting stuck into different flow states in the phase-space, each of which has relatively high persistence, without these different flow states being necessarily clearly distinct modes. For example, some members might get stuck in a more NAO-negative state and others in a more NAO-positive state, and this could result in some amount of intransitivity irrespective of what the distribution of the NAO index looks like (i.e. whether it's multimodal or not). In the chapter I also used the language of regimes, as it is more approachable, which is important for working in an interdisciplinary way. A more accurate description would be "state space flow tendencies" but the purpose of this research was to inspire practical testing of my hypothesis, as well as awareness of the utility of my dynamical systems approach.

There are several avenues for future work based on this hypothesis. One of which is exploration of the Thirumalai–Mountain effective ergodic convergence metric. In the long-time limit, $\Omega = \lambda/t$ if the ergodic theorem is holding, where λ is a constant for the rate of convergence. For a large ensemble it would be interesting to test this, as one should see the metric grow linearly in time. I suspect there might be bumps in the trend indicating times where the ensemble is not satisfying the theorem. The metric is a good indicator as slower convergence means weaker ergodicity, and no convergence at all is non-ergodicity. Given the hypothesis outlined, one would expect to find weak convergence, and the weaker it is, the more likely $RPC > 1$. Other simulation experiments should include changing the definition of the time-mean, as this in effect changes the phase-space. It would be interesting to see how $V(e)$ changes with respect to $V(o)$ and how $r_{\bar{e}o}$ changes, when the time-average is gradually reduced or increased for the NAO from a 90 day DJF-mean all the way up to a single mid-day in January, and likewise adding more weeks either side of the DJF-mean.

In terms of further theoretical work, it would be useful to derive the exact rate of convergence as a function of ensemble size and averaging length. Though these would be very formal mathematics problems.

Chapter 4

Detecting microbiology in the upper atmosphere: relative-velocity filtered sampling

If anyone should come to the top of the air...lift his head above it and see...he would recognize that that is the real heaven. –Plato from [97]

This publication originally appeared in Astrobiology, 31st March 2023, DOI: 10.1089/ast.2022.0045

4.1 Introduction

In this research chapter we remain in the atmosphere, rising to its penultimate outer layers, where complex, non-equilibrium processes dominate. This is due to the mesosphere-lower thermosphere encompassing the transition from continuous turbulent fluid to more gaseous interactions as the air density reduces substantially. This region is otherwise known as the "ignore-osphere" due to its inaccessibility by balloons, conventional aircraft and satellites as there is still enough air for drag to be non-negligible. Here the problem of measurement contamination in searches for biological material is addressed by examining the dynamical system of the contaminant particle in a flow generated by the motion of the detector. From this point of view, we propose a new detector design that filters for particles based on their size and relative velocity to the detector in-motion. Advances in very low Earth orbit CubeSat technology are adopted, based on theoretical calculations, to the proposed detector design. This research underlines the point that simple calculations based on sound theoretical principles can be powerful tools where the pioneering nature of the study is too early to warrant the development of costly computational simulations, such as those in Chapter 2.

ASTROBIOLOGY
Volume 23, Number 4, 2023
© Mary Ann Liebert, Inc.
DOI: 10.1089/ast.2022.0045

Hypothesis Article

Open camera or QR reader and scan code to access this article and other resources online.

Detecting Microbiology in the Upper Atmosphere: Relative-Velocity Filtered Sampling

Arjun Berera,[1] Daniel J. Brener,[1] and Charles S. Cockell[2]

Downloaded by Edinburgh University (Roslin Inst) from www.liebertpub.com at 05/03/24. For personal use only.

Abstract

The purpose of this article is to reopen from a practical perspective the question of the extent in altitude of Earth's biosphere. We make a number of different suggestions for how searches for biological material could be conducted in the mesosphere and lower thermosphere, colloquially referred to as the ''ignore-osphere'' because it has been generally ignored in the meteorological community compared to other regions. Relatively recent technological advances such as CubeSats in very low Earth orbit or more standard approaches such as the rocket-borne MAGIC meteoric smoke particle sampler are shown as potentially viable for sampling biological material in the ignore-osphere. The issue of contamination is discussed, and a potential solution to the problem is proposed by means of a new detector design that filters for particles based on their size and relative velocity to the detector. Key Words: Thermosphere—Mesosphere—Biosphere and upper atmosphere sampling. Astrobiology 23, 469–475.

1. Introduction

The physical pervasiveness of life and its associated biological structures around Earth has been an open and much-discussed question for centuries. The idea that Earth might be a source of living material in the Solar System, and beyond, has both practical and philosophical implications. In terms of pure ecology, the maximum altitude at which life might be found is an important question in its own right as it defines the Earth system's physical extent as a habitat. An example of a practical application is in exploring the upper atmosphere as a possible transport mechanism for pathogenic microorganisms around Earth (Brown and Hovmøller, 2002; Bebber *et al.,* 2014). Furthermore, it has been suggested that if biological particles can be found in the thermosphere, then hypervelocity space dust has sufficient momentum to facilitate the planetary escape of such particles, which leads to astrobiology implications in the form of the transport of biological material beyond Earth (Berera, 2017). There has been no concerted research effort placed on the mesosphere and thermosphere by the astrobiology community, despite this interest (Smith, 2013; Dassarma *et al.,* 2020; Šantl-Temkiv *et al.,* 2022). The concept of the biosphere was first introduced by the geologist Eduard Suess in 1875 as the surface on Earth where life can be found (Suess, 1875). This was later extended by the Russian polymath Vladimir I. Vernadsky to be any complete system which encapsulates life (Vernadsky, 1998). As the word implies, a biosphere has boundaries beyond which life cannot be found. Therefore, it is at the boundaries of the biosphere that one finds life at the extremes. Research has established the existence of life at extreme temperatures within volcanoes and in the arctic, at extreme pressures in the deep ocean, and in high-radiation environments (Clarke *et al.,* 2013; Krisko and Radman, 2013; Richard *et al.,* 2020). Certain life-forms have even been shown to survive in the space environment (Moissl-Eichinger *et al.,* 2016; Kawaguchi *et al.,* 2020). The extent of the biosphere in the upper atmosphere, that is, the existence of life at extreme altitudes, remains comparatively little explored.

In certain meteorological circles, the mesosphere and lower thermosphere (MLT) is colloquially referred to as the ''ignore-osphere'' because its unique position makes it inaccessible via balloons and conventional aircraft due to the low air density. It is also inaccessible for satellites as there is still sufficient air for drag to be non-negligible (Stern, 2010).

[1]The Higgs Centre for Theoretical Physics, University of Edinburgh, Edinburgh, UK.
[2]UK Centre of Astrobiology, University of Edinburgh, Edinburgh, UK.

Thus, until recently, the only viable sampling option has been to use sounding rockets. In the last decade, new possibilities for sampling have arrived, such as small inexpensive CubeSats and low Earth orbit commercial spacecraft (Virgili-Llop *et al.,* 2014; Saeed *et al.,* 2020).

The highest altitudes where life seems to have been plausibly established are the top of the troposphere and the lower stratosphere (Bryan *et al.,* 2019; Dassarma *et al.,* 2020; Šantl-Temkiv *et al.,* 2022). However, studies deeper into the stratosphere remain controversial due to questions of contamination. In a recent review, Šantl-Temkiv *et al.* (2022) emphasized that caution is required when interpreting findings in studies where decontamination measures are not fully reported. Earlier studies do not fully report decontamination measures or evaluate the issue compared to modern standards. However, there is a consistent literature dating to the present which reports micrometer-sized particles in the stratosphere, for example Rosen (1964), Heintzenberg *et al.* (2003), Xu *et al.* (2003), Yin *et al.* (2012), Ursem (2016), and Šantl-Temkiv *et al.* (2022).

Additional studies found that these particles include bacteria (Wainwright *et al.,* 2003; Griffin, 2004; Ursem, 2016; Šantl-Temkiv *et al.,* 2022). The most recent observational campaign of the stratosphere found two bacterial species, *Bacillus simplex* and *Staphylococcus pasteuri,* and a fungus, *Engyodonitium album,* at 41 km (Wainwright *et al.,* 2003).

The purpose of this article is to reopen the question of what the upper limit of Earth's biosphere is with practical calculations and suggestions for field campaigns. Our article builds on the preexisting observations of the lower atmosphere and on recent work that has identified strong vertical winds, among other mechanisms such as volcanic eruptions, as capable of ejecting nanometer-sized particles into the MLT (Berera, 2017; Berera and Brener, 2022).

The highest altitude mission that sought to determine the extent of the biosphere was conducted in 1974 by Imshenetsky *et al.* of the Institute of Microbiology, USSR Academy of Sciences (Imshenetsky *et al.,* 1978). They used sounding rockets fitted to capture microbiological samples up to 77 km in the mesosphere. They reported finding fungal spores. These findings are the only published study of these altitudes, and therefore there remain concerns about contamination. We propose a new approach to address these questions. Only through multiple and independently run studies will it be possible to settle concerns around contamination, as was done for determining the existence of life in the deep ocean (Gronstal *et al.,* 2009; Escudero *et al.,* 2018).

Samples from the surface of the International Space Station have been found to have DNA from several kinds of bacteria that were genetically similar to those found in the Barents and Kara Seas' coastal zones (Grebennikova *et al.,* 2018). It was proposed that the wild land and marine bacteria DNA could transfer from the lower atmosphere into the ionosphere-thermosphere using the ascending branch of the global electric circuit. However, a more likely explanation would be contamination due to spacecraft passing through the lower atmosphere. Atmospheric transmission of microbes has also been postulated as one of the mechanisms behind Antarctic microbial diversity changes (Pearce *et al.,* 2009).

2. MLT Sampling Methodologies

2.1. Standard approaches

Since the late 1940s, there have been rocket sampling missions to high altitudes (Martell, 1970). These initially involved the use of evacuated steel bottles with an altitude-triggered opening and closing mechanism. Vacuum tubes are a valid method for directly sampling the upper atmosphere, but they are limited in terms of temporal and volumetric capacity.

With the advent of the Cold War, it became imperative to sample the upper atmosphere for radioactive fallout during nuclear weapons tests. To capture enough of these particles, larger volumes needed to be sampled, which led to the development of the first cryogenic samplers such as the ENCAR-1 (Denton, 1970; Martell, 1970). Prior contamination of such a sampling device is readily detectable via a Geiger counter, and there are no large natural sources of radioactive material in the atmosphere. Hence, these earlier studies had little with which to contend with respect to contamination (Martell, 1970).

The only direct searches for biological material in the MLT were conducted by Imshenetsky *et al.,* who used a specially designed sampler located in the nose cone of a sounding rocket (Imshenetsky *et al.,* 1978). These authors assumed that the rocket's outer casing would experience temperatures > 1000 K that would act as an ascent sterilization. While this is correct in principle, aerodynamically modeling the heating profile for the rocket is the only way to be certain. The assumption by Imshenetsky *et al.* (1978) that the heating was uniform and therefore the sterilization would also be is a flaw in their contamination strategy.

Since the missions by Imshenetsky *et al.* (1978), there have been no further samplings of the MLT conducted that looked directly for biological particles. This led the aerobiology community to conclude that 77 km is the maximum empirically defined extent of the biosphere in altitude (Horneck *et al.,* 2010; Smith, 2013). The review by Smith (2013) criticized this conclusion as Imshenetsky *et al.* (1978) neglected to give details on contamination precautions other than the aforementioned drag heating of the rocket exterior or the sealing of the samples in flight (Smith, 2013). It is crucial that future field campaigns report clearly and transparently all procedures used, even those considered routine.

The MLT has been studied by those trying to understand meteoric smoke and the formation of noctilucent clouds. The most relevant to our work is the MAGIC meteoric smoke particle sampler (Hedin *et al.,* 2014). Meteoric smoke particles are neutrally charged and nanometer in size, making them challenging to detect. Berera and Brener (2022) hypothesized that MLT large vertical winds could potentially be strong enough to push particles of this size to around 120 km in rare events.

This problem of the air density is a question of whether the flow is in the continuum limit. Physically, this corresponds to whether the particles are dominated by individual molecular collisions, such as in Brownian motion, or by the larger-scale fluid dynamics. This is determined by the number density of the air molecules with respect to the characteristic length scale of the object in the flow. MLT flow conditions range from the transition region between continuum and noncontinuum to the free molecular flow regime (noncontinuum). The rarefaction of a gas that interacts with

Downloaded by Edinburgh University (Roslin Inst) from www.liebertpub.com at 05/03/24. For personal use only.

Downloaded by Edinburgh University (Roslin Inst) from www.liebertpub.com at 05/03/24. For personal use only.

an object with characteristic size L can be described by the Knudsen number, $\mathrm{Kn}=\frac{\lambda}{L}$ where λ is the mean free path of the gas. For continuum flow conditions, one must have Kn << 0.1 and for free molecular flow Kn > 10.

In the MLT, Kn ranges from 0.1 to 10,000; hence, any direct flow measuring instrument must be capable of dealing with both transition and free molecular flow, unless separate missions are conducted per regime. Flow conditions and momentum transfer also change depending on the sizes of particle considered. In the MLT, particles with $L \sim 10$ nm interact enough with the surrounding rarefied gases that their interaction can be considered continuous (Hedin *et al.,* 2007a). Most viruses have a radius of around 20 nm; therefore, if one only wanted to search for whole viruses and larger bacteria or fungi, simple continuum modeling may be appropriate. However, accurate modeling of the effect of the gas on heating the detector would need to be carried out.

2.2. New approach: Relative-velocity filtered sampling

Here, we propose a new technique for reducing the risks of contamination that affects sampling at high altitudes. A detector needs to be designed that can sample high volumes over long periods of time in order to have a reasonable chance of encountering a biological particle. Detectors that impede the airflow are not appropriate, as this will increase drag, reducing the flight time and thence the total sampled volume per mission.

For a detector that continuously samples a flow, the particles move with respect to the detector rest frame. We propose that most contaminants from craft interior and exterior will likely be approximately at rest with respect to the detector. Therefore, by eliminating particles which satisfy that criteria, one reduces an aspect of the contamination problem.

In terms of detector design, it is easier to apply this thinking by considering the velocities of the particles. Contaminants from the payload delivery system will have a velocity smaller than particles entering the detector as part of the flow generated by the motion of the rocket or orbiting craft. Contaminants are most likely to come from the exterior of the craft's surface, and due to the low density at these altitudes, they will take longer to reach velocities close to the mean flow.

Therefore, if the detector moves sufficiently fast, greater than an order of magnitude relative to the rarefied gas flow, then any particle that detaches itself from the exterior of the craft that enters the detector will have a slower velocity. This could allow for such a detector to be placed at a position other than at the head of the rocket. However, careful simulation and modeling would be required to ensure the correct relative velocity is selected for an accurate trigger action.

Figure 1 depicts a high-level schematic of the triggering process for taking a sample. Once the craft is at the selected sampling altitude, the sampling probe should open to allow the air to pass through it. After some short time, the triggering mechanism is activated. To detect a particle of interest to capture, a small lidar system could be used to scan the incoming flow in the tube detecting both particle size and velocity. The triggering computer will have two options. It can either allow the flow to continue through the detector or it can close a valve, forcing the flow to pass into a sampling chamber where particles can be captured. The decision tree is

(1) IF $v < u$: valve open, particle is rejected from analysis
(2) IF $v \sim u$ or $v > u$ AND size > 10.0 nm: valve close, particle accepted for analysis,

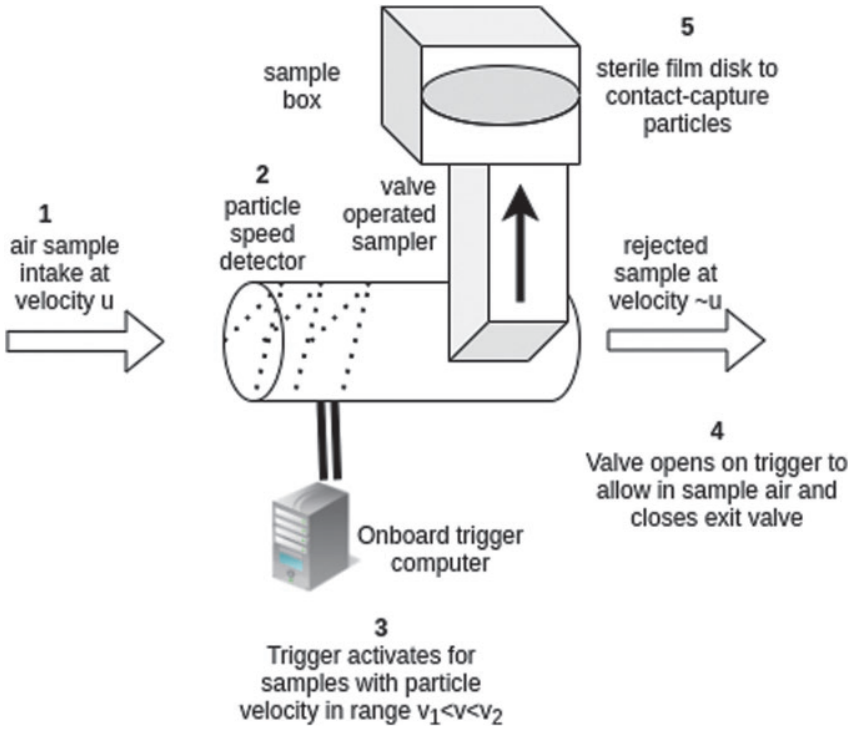


FIG. 1. Trigger mechanism key stages.

where u is the relative velocity between the analyzer and the air flow, which would approximately be Mach 3, and v is the speed of the particle detected by the lidar relative to the analyzer.

For an order of magnitude estimate of the scanning frequency and trigger speed requirements of such a detector, we assume the flow into the detector travels at 1000 m/s (approximately Mach 3, a typical sounding rocket velocity). This means that, for a particle with characteristic size of 1 nm, a detector of length 10 cm would have $\sim 100\,\mu s$ to detect it, which is within the operating capabilities of circuitry.

Dyes that bind to specific proteins or nucleic acids (*e.g.,* Sybr GREEN; Vitzthum *et al.,* 1999) could be used to illuminate biological material. Currently, these take time to bind and illuminate, but in principle one could develop an instantaneously binding dye. The air could be sampled, sprayed, and then accepted/rejected for analysis based on the detection of fluorescence. An alternative procedure could be to spray the samples once they have been filtered by relative velocity, as a means of highlighting regions of interest on the film discs. A similar idea has been proposed not for detection of life but to eliminate contamination concerns by painting the exterior rocket surface with fluorescent beads to trace and identify contamination pathways (Juanes-Vallejo *et al.,* 2011; Smith, 2013).

There are a number of engineering challenges, however, that would need to be dealt with. The crucial problem is that the detector must function for the transition and free molecular regimes. To do this, Direct Simulation Monte Carlo (DSMC) models of the detector would be required. This technique was used for the MAGIC detector experiments as well as the development of other MLT sounding rocket instrumentation (Hedin *et al.,* 2007a). Secondly, the scanning device itself will be a challenge to engineers such that it can operate at high frequency and be relatively compact. There are a number of similar devices already developed for industry, as detecting small particles in a flow is a standard problem (see, *e.g.,* Sudo *et al.,* 2007).

However, this technique to mitigate contamination during sampling is only as good as the weakest point in a mission, from craft assembly through to post-sampling processing. These are specific challenges in microbiology due to the wide range of biomasses that need to be analyzed. This problem was tackled by Luhung *et al.* (2021), where the authors developed a four-stage ultralow biomass pipeline: amassment, storage, extraction, and nucleic acid analysis. They used several decontamination procedures, including a negative control, where they processed a sample collected with zero airflow into their detector for 1 min. Such a procedure should be implemented for our proposed technique.

The problem of contamination in low microbial biomass microbiome studies was thoroughly reviewed by Eisenhofer *et al.* (2018), and we strongly support the recommendations to minimize the influence of contaminant DNA in atmospheric sampling missions (Eisenhofer *et al.,* 2018). Another approach to contamination control would be to deliberately introduce artificial contaminants, such as polystyrene spheres, onto the surfaces of locations from which contamination might occur. The detection of these spheres by the capture device would allow for contamination to be quantified and identified. Similar approaches are regularly used in deep subsurface microbiological drilling (*e.g.,* Cockell *et al.,* 2009), whereby fluorescent polystyrene spheres are added to the drilling mud to control for the internal contamination of core material. In summary, contamination can never be completely ruled out, as with other low biomass studies; but similarly to deep drilling and core recovery in low biomass environments, there are a number of ways to control and quantify the potential for contamination in the instrument, which would be a critical part of instrument design and development.

3. Mission Strategies

Firstly, we consider the use of sounding rockets, which could be used simply to repeat the experiments by Imshenetsky *et al.* (1978). Using a sounding rocket is beneficial as it provides a high sampling volume for relatively low cost. To zeroth order, even in the free molecular regime, the sampled volume will be given by the product of the flow rate through or over the surface of the detector with the sampling time. The flow rate to zeroth order is the surface area of the detector orthogonal to the flow product with the mean speed of the flow relative to the detector.

Assuming a relative velocity of 1000 m/s and a circular detector area of $0.007\,m^2$, corresponding to a diameter of 10 cm, one finds a flow rate of $\sim 7\,m^3/s$. A typical sounding rocket flight through the MLT lasts around 80 s, which would result in $\sim 560\,m^3$ of sampled volume. To find one particle of biological material in such a volume would be a significant result in the MLT. The key factor in the total volume sampled is the time the rocket can spend in the MLT. Even just 20 min of horizontal flight at a fixed altitude would yield a sample volume of $\sim 8400\,m^3$. There are groups working on sampling rockets capable of such flights (Strelnikov, 2022). While these calculations are extremely crude, they nonetheless demonstrate real potential for defining the biosphere even for extremely low particle concentrations.

One could devise a series of sounding rocket missions where the rocket cruises at a fixed altitude while sampling, at intervals of 10 km from 50–150 km. These should be launched months apart to avoid inter-launch contamination. Geospatial separation of the launch sites would also be ideal as Berera and Brener (2022) identified the poles as more likely to have biological material, since these areas experienced stronger vertical winds due to geomagnetic activity.

Of vital importance when considering the use of rockets is the aerodynamic design of the rocket and detector itself. This is to ensure that a steady, representative flow passes through the device and that when particles are selected for capture, blow-back is not an issue. Future work should examine ways to ensure that the particles captured at high speed have their momentum sufficiently arrested so that they stick to the sampler. These are crucial problems that have been around for some time in rocket-based sampling (Bird, 1988; Horányi *et al.,* 1999; Gumbel, 2001; Hedin *et al.,* 2007b).

Alternatively, CubeSats represent a significant advance in technology in the last decade. We propose an ambitious potential mission program that would use multiple CubeSats launched months apart to sample the upper thermosphere down to the lower thermosphere, prior to reentry. The commercial space industry has already developed CubeSats that can orbit for short periods in what is called the extremely low Earth orbit (ELEO) region, and there have already been field campaigns using multiple satellites to explore this region (Millan *et al.,* 2019).

One example of such a CubeSat is the TSAT Globalstar ELaNa-5 ELEO satellite (ELEO-SAT), which has dimensions of $10 \times 10 \times 20$ cm and an orbit velocity of ~ 7 km/s

Downloaded by Edinburgh University (Roslin Inst) from www.liebertpub.com at 05/03/24. For personal use only.

Downloaded by Edinburgh University (Roslin Inst) from www.liebertpub.com at 05/03/24. For personal use only.

(Voss *et al.,* 2014). Using the same simple calculation method as before, this has the potential to sample $\sim 4200\,m^3$ in 1 h. Without a propulsion system, it is not possible to control the altitude of the CubeSat, and any such system would add too much additional weight that would reduce the sampling time. Hence, the CubeSat rapidly decreases in altitude and, thus, cannot be used to sample a single altitude accurately.

However, there are two key benefits to CubeSats over conventional sounding rockets: firstly, CubeSats can sample for hours and even weeks, whereas rockets are limited to periods of minutes. Secondly, they sample the atmosphere globally, as they are in orbit. The ELEO-SAT is designed for atmospheric sampling between 120 and 325 km with over 600 orbits during the 5 weeks of flight time. To isolate altitude ranges, we suggest that a series of CubeSats should be deployed to sample different 100 km bands (*e.g.,* 200–100 km).

In Fig. 2, we have illustrated the mission strategy. In phase 1, the CubeSat is launched and deployed by rocket. The CubeSat could be projected out of the rocket forward by an explosive charge so that it remains upstream of the rocket on its first orbit ($\sim$90 min). For phase 2, the CubeSat will sample a region of atmosphere consistent with its orbital resonance as CubeSats can rise and fall by as much as 50 km in a single orbit (see, *e.g.,* Fig. 16 of Voss *et al.,* 2014).

Alternatively, a single ELEO-SAT released from 400 km, traveling on average at 7 km/s for 325 km down to 120 km over 5 weeks (840 h), would have come into contact with approximately $10^5\,m^3$ of air. Due to the high mixing rates at these altitudes, it is extremely improbable that the satellite could sample identical volumes of air.

CubeSats usually burn up in the mesosphere at around 110 km; however, it is necessary that the craft can be fully recovered. This could be achieved by using a commercially available system called a CubeSat Deorbit and Recovery System (DRS) (Carandente, 2014). The CubeSat DRS is a module that can be attached to the CubeSat containing a 1.2 m tension cone heat shield and parachute, all weighing <1.5 kg. This module is triggered to deploy the shield immediately prior to reentry (Carandente, 2014).

4. Discussion

It will never be possible to entirely rule out contamination from a human-launched craft. We would argue that this is not an issue for defining the upper limits of the biosphere because the increased distribution of microbiological material in the upper atmosphere and possibly into space caused by human activities is now most likely a fact. Even if we were to only consider this human-derived biological material, we could still ask: What is the extent of this material? Does the microbiology survive, and does it continue to replicate there?

If naturally occurring biology is to be found in the MLT, it is most likely to be on the scale of viruses, prokaryotes (bacteria and archaea), or possibly fungal spores (Imshenetsky *et al.,* 1978; Berera and Brener, 2022). If one recovers viable biological material, then one could use such samples, together with the flight data, to better quantify the proportion of particles that are viable as a function of altitude. This could provide valuable data on microbial viability, survival, and destruction in the upper atmosphere.

Relative-velocity filtered sampling might also be applicable to searches for evidence of life on other planets. Conventional approaches require the landing of probes on the planet's surface, excavating the soils and performing analysis on samples (Anderson *et al.,* 2012; Clark *et al.,* 2021). Our proposal solves two problems: significantly reducing a major part of the contamination concerns and removing the requirement of landing a probe in order to sample for life. Of course, a land-based

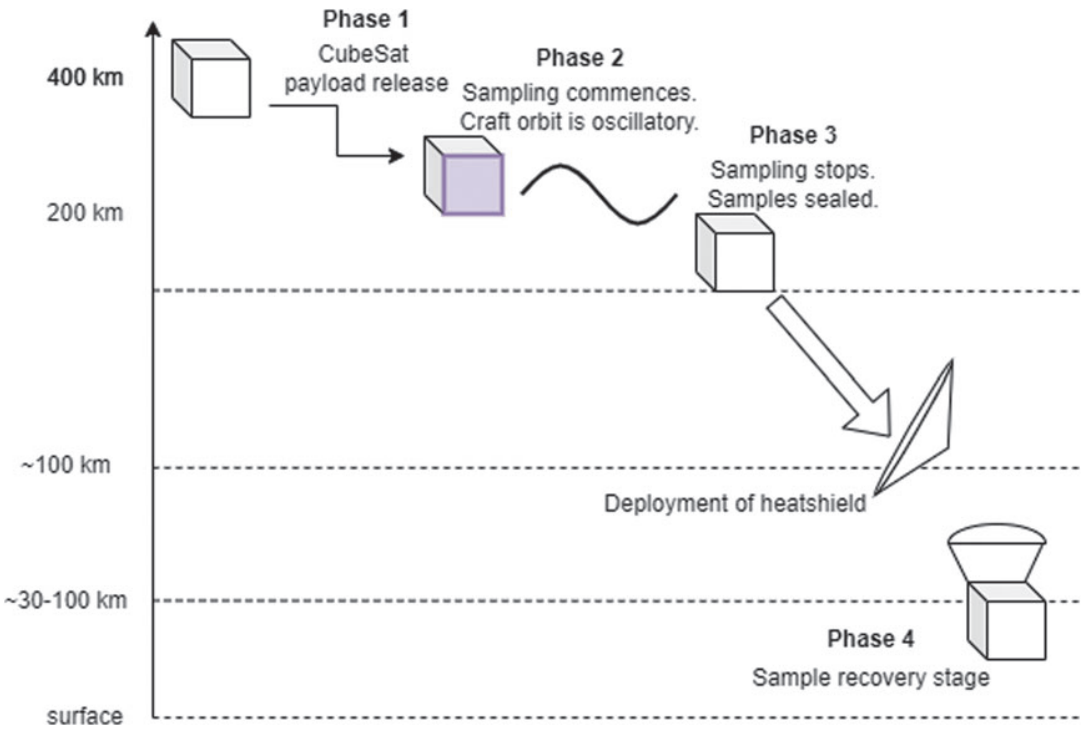


FIG. 2. CubeSat mission strategy core phases.

relative-velocity filter device would not eliminate the need for landing, although it could potentially reduce the need to move the probe around the planet. Using the natural winds to sample the biosphere is, in our opinion, likely to be a more efficient and lower-risk strategy than is currently achieved with rovers. An added advantage is that being at ground level, if there is any biology on the planet, it is more likely to be found closer to the surface, assuming the planet has a similar atmospheric structure to Earth (Smith, 2013; Berera and Brener, 2022).

5. Concluding Remarks

Defining the biosphere's extent in altitude is of interest to the microbiology and aerobiology communities and has implications for ideas in astrobiology. Few field campaigns have been conducted to empirically determine this upper extent of the biosphere, largely due to the difficulty of sampling the MLT. As with previous microbiological campaigns that have looked for life at the extremes, the contamination problem must be foremost in any effort to sample these remote atmospheric regions. We propose that a sampling device could detect the relative velocity and size of incoming particulates, thereby filtering sampled air to mitigate against self-contamination.

The main underlying feature of the relative-velocity-filtered sampling method we propose is that the particle of interest that is to be measured should be at a different relative velocity to the detector, thus eliminating contamination that may persist on the detector. The proposal in the present study is to utilize this method to look for biological particles in the upper atmosphere, above 100 km, which would be a novel discovery. The detector should be on a moving craft, such as a sounding rocket, satellite, or even a drone, which intakes the surrounding air at high velocity and then detects the particles of interest in flight.

We have investigated how such a device could be designed around a CubeSat craft in very low Earth orbit or in the nose cone of a rocket to enable sampling of the MLT. The ideas involve several engineering challenges, specifically around the rapid detection of particles using a lidar that is lightweight, small, and of sufficient temporal and spatial resolution. Other technologies yet to be realized may be more appropriate or in the future be developed, but the principle of using the relative velocity of individual particles in a flow will in general always be applicable.

Acknowledgment

The authors were supported by the United Kingdom Science and Technology Facilities Council, with Charles S. Cockell under grant ST/V000586/1.

References

Anderson RC, Jandura L, Okon AB, *et al.* Collecting samples in Gale Crater, Mars; An overview of the Mars Science Laboratory Sample Acquisition, Sample Processing and Handling System. *Space Sci Rev* 2012;170(1):57–75; doi: 10.1007/s11214-012-9898-9.

Bebber DP, Holmes T, Gurr SJ. The global spread of crop pests and pathogens. *Glob Ecol Biogeogr* 2014;23(12):1398–1407; doi: 10.1111/geb.12214.

Berera A. Space dust collisions as a planetary escape mechanism. *Astrobiology* 2017;17(12):1274–1282; doi: 10.1089/ast.2017.1662.

Berera A and Brener DJ. On the force of vertical winds in the upper atmosphere: Consequences for small biological particles. *Proc Math Phys Eng Sci* 2022;478(2257):20210626; doi: 10.1098/rspa.2021.0626.

Bird GA. Aerodynamic effects on atmospheric composition measurements from rocket vehicles in the thermosphere. *Planet Space Sci* 1988;36(9):921–926; doi: https://doi.org/10.1016/0032-0633(88)90099-2.

Brown JK and Hovmøller MS. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* 2002;297(5581):537–541; doi: 10.1126/science.1072678.

Bryan NC, Christner BC, Guzik TG, *et al.* Abundance and survival of microbial aerosols in the troposphere and stratosphere. *ISME J* 2019;13(11):2789–2799; doi: 10.1038/s41396-019-0474-0.

Carandente V. New concepts of deployable de-orbit and re-entry systems for CubeSat miniaturized satellites. *Recent Patents on Engineering* 2014;8:2–12; doi: 10.2174/1872212108666140204004335.

Clark BC, Kolb VM, Steele A, *et al.* Origin of life on Mars: Suitability and opportunities. *Life* 2021;11(6); doi: 10.3390/life11060539.

Clarke A, Morris GJ, Fonseca F, *et al.* A low temperature limit for life on Earth. *PloS One* 2013;8(6):e66207; doi: 10.1371/journal.pone.0066207.

Cockell CS, Gronstal AL, Voytek MA, *et al.* Microbial abundance in the deep subsurface of the Chesapeake Bay Impact Crater: Relationship to lithology and impact processes. In *The ICDP-USGS Deep Drilling Project in the Chesapeake Bay Impact Structure: Results from the Eyreville Core Holes*. Geological Society of America: Boulder, CO, 2009; doi: 10.1130/2009.2458(40).

Dassarma P, Antunes A, Simões MF, *et al.* Earth's stratosphere and microbial life. In *Astrobiology: Current, Evolving, and Emerging Perspectives.* (Antunes A. ed.) Caister Academic Press: Poole, England, 2020; doi: 10.21775/9781912530304.08.

Denton EH. Rocket-borne cryogenic air samplers. *Science Progress* 1970;58(232):539–561.

Eisenhofer R, Minich JJ, Marotz C, *et al.* Contamination in low microbial biomass microbiome studies: Issues and recommendations. *Trends Microbiol* 2018;27(2):105–117; doi: 10.1016/j.tim.2018.11.003.

Escudero C, Oggerin M, Amils R. The deep continental subsurface: The dark biosphere. *Int Microbiol* 2018;21(1):3–14; doi: 10.1007/s10123-018-0009-y.

Grebennikova TV, Syroeshkin AV, Shubralova EV, *et al.* The DNA of bacteria of the World Ocean and the Earth in cosmic dust at the International Space Station. *The Scientific World Journal* 2018;1–7; doi: 10.1155/2018/7360147.

Griffin D. Terrestrial microorganisms at 20,000 meters in Earth's atmosphere. *Aerobiologia* 2004;20:135–140; doi: 10.1023/B:AERO.0000032948.84077.12.

Gronstal AL, Voytek MA, Kirshtein JD, *et al.* Contamination assessment in microbiological sampling of the Eyreville Core, Chesapeake Bay Impact Structure. In *The ICDP-USGS Deep Drilling Project in the Chesapeake Bay Impact Structure: Results from the Eyreville Core Holes.* Geological Society of America: Boulder, CO, 2009; doi: 10.1130/2009.2458(41).

Gumbel J. Aerodynamic influences on atmospheric *in situ* measurements from sounding rockets. *J Geophys Res Space Physics* 2001;106(A6):10553–10563; doi: 10.1029/2000JA900166.

Hedin J, Gumbel J and Rapp M. On the efficiency of rocket-borne particle detection in the mesosphere. *Atmos Chem Phys* 2007a;7(14):3701–3711; doi: 10.5194/acp-7-3701-2007.

Downloaded by Edinburgh University (Roslin Inst) from www.liebertpub.com at 05/03/24. For personal use only.

Downloaded by Edinburgh University (Roslin Inst) from www.liebertpub.com at 05/03/24. For personal use only.

Hedin J, Gumbel J, Waldemarsson T, *et al.* The aerodynamics of the MAGIC meteoric smoke sampler. *Adv Space Res* 2007b; 40(6):818–824; doi: https://doi.org/10.1016/j.asr.2007.06.046.

Hedin J, Giovane F, Waldemarsson T, *et al.* The MAGIC meteoric smoke particle sampler. *Journal of Atmospheric and Solar-Terrestrial Physics* 2014;118:127–144; doi: 10.1016/j.jastp.2014.03.003.

Heintzenberg J, Hermann M, Theiss D. Out of Africa: High aerosol concentrations in the upper troposphere over Africa. *Atmos Chem Phys* 2003;3(4):1191–1198; doi: 10.5194/acp-3-1191-2003.

Horányi M, Gumbel J, Witt G, *et al.* Simulation of rocket-borne particle measurements in the mesosphere. *Geophys Res Lett* 1999;26(11):1537–1540; doi: 10.1029/1999GL900298.

Horneck G, Klaus DM, Mancinelli RL. Space microbiology. *Microbiol Mol Biol Rev* 2010;74(1):121–156; doi: 10.1128/MMBR.00016-09.

Imshenetsky AA, Lysenko SV, Kazakov GA. Upper boundary of the biosphere. *Appl Environ Microbiol* 1978;35(1):1–5; doi: 10.1128/aem.35.1.1-5.1978.

Juanes-Vallejo CM, Grama VV, Katramados I, *et al.* Cranfield Astrobiological Stratospheric Sampling Experiment (Cass.E): Overview of flight hardware configuration, implemented planetary protection and contamination control procedures and preliminary post-flight results. In *20th Symposium on European Rocket and Balloon Programmes and Related Research.* (Ouwehand L. ed.) ESA Special Publication: Noordwijk, the Netherlands, 2011; pp 467–472.

Kawaguchi Y, Shibuya M, Kinoshita I, *et al.* DNA damage and survival time course of *Deinococcal* cell pellets during 3 years of exposure to outer space. *Front Microbiol* 2020;11: 2050; doi: 10.3389/fmicb.2020.02050.

Krisko A, Radman M. Biology of extreme radiation resistance: The way of *Deinococcus radiodurans*. *Cold Spring Harb Perspect Biol* 2013;5(7):a012765; doi: 10.1101/cshperspect.a012765.

Luhung I, Uchida A, Lim SBY, *et al.* Experimental parameters defining ultra-low biomass bioaerosol analysis. *NPJ Biofilms Microbiomes* 2021;7(1):37; doi: 10.1038/s41522-021-00209-4.

Martell EA. High altitude air sampling with a rocketborne cryocondenser. *Journal of Applied Meteorology* 1970;9(1):170–177.

Millan RM, von Steiger R, Ariel M, *et al.* small satellites for space science: A COSPAR Scientific Roadmap. *Adv Space Res* 2019;64(8):1466–1517; doi: https://doi.org/10.1016/j.asr.2019.07.035.

Moissl-Eichinger C, Cockell C, Rettberg P. Venturing into new realms? Microorganisms in space. *FEMS Microbiol Rev* 2016;40(5):722–737; doi: 10.1093/femsre/fuw015.

Pearce DA, Bridge PD, Hughes KA, *et al.* Microorganisms in the atmosphere over Antarctica. *FEMS Microbiol Ecol* 2009; 69(2):143–157; doi: 10.1111/j.1574-6941.2009.00706.x.

Richard L, Bizzoco W, Kelley S. Volcanic steam vents: Life at low pH and high temperature. In *Extremophiles as Astrobiological Models.* (Seckbach J, Stan-Lotter H. eds.) Wiley-Scrivener: Hoboken, NJ, 2020; pp 1–20; doi: 10.1002/9781119593096.ch1.

Rosen JM. The vertical distribution of dust to 30 kilometers. *J Geophys Res* 1964;69(21):4673–4676; doi: 10.1029/JZ069i021p04673.

Saeed N, Elzanaty A, Almorad H, *et al.* CubeSat communications: Recent advances and future challenges. *IEEE Communications Surveys & Tutorials* 2020;22(3):1839–1862; doi: 10.1109/COMST.2020.2990499.

Šantl-Temkiv T, Amato P, Casamayor EO, *et al.* Microbial ecology of the atmosphere. *FEMS Microbiol Rev* 2022;46(4): fuac009; doi: 10.1093/femsre/fuac009.

Smith D. Microbes in the upper atmosphere and unique opportunities for astrobiology research. *Astrobiology* 2013; 13(10):981–990; doi: 10.1089/ast.2013.1074.

Stern A. *Next-Generation Suborbital Researchers Conference: February 18-20, 2010*. Boulder, CO; 2010.

Strelnikov B. Private Communication, Leibniz-Institute of Atmospheric Physics, 2022.

Sudo S, Miyasaka Y, Nemoto K, *et al.* Detection of small particles in fluid flow using a self-mixing laser. *Opt Express* 2007;15(13):8135–8145; doi: 10.1364/OE.15.008135.

Suess E. *Die Entstehung Der Alpen*. Braumüller; 1875.

Ursem B. Climate shifts and the role of nano structured particles in the atmosphere. *Atmospheric and Climate Sciences* 2016; 06:51–76; doi: 10.4236/acs.2016.61005.

Vernadsky VI. The biosphere: An envelope of the Earth. In *The Biosphere.* Springer: New York, 1998; pp 91–102; doi: 10.1007/978-1-4612-1750-3_8.

Virgili-Llop J, Roberts P, Hao Z, *et al.* Very low Earth orbit mission concepts for Earth observation: Benefits and challenges. In *Proceedings of the Reinventing Space Conference*. Royal Society: London, 2014.

Vitzthum F, Geiger G, Bisswanger H, *et al.* A quantitative fluorescence-based microplate assay for the determination of double-stranded DNA using SYBR Green I and a standard ultraviolet transilluminator gel imaging system. *Anal Biochem* 1999;276(1):59–64; doi: 10.1006/abio.1999.4298.

Voss H, Dailey J, Crowley J, *et al.* TSAT Globalstar ELaNa-5 Extremely Low-Earth Orbit (ELEO) Satellite. In *28th Annual AIAA/USU Conference on Small Satellites.* Utah State University: Logan, UT, 2014; SSC14-WK-6.

Wainwright M, Wickramasinghe NC, Narlikar JV, *et al.* Microorganisms cultured from stratospheric air samples obtained at 41 km. *FEMS Microbiol Lett* 2003;218(1):161–165; doi: 10.1111/j.1574-6968.2003.tb11513.x.

Xu L, Shi G, Zhang L, *et al.* Number concentration, size distribution and fine particle fraction of tropospheric and stratospheric aerosols. *China Particuology* 2003;1(5):201–205; doi: https://doi.org/10.1016/S1672-2515(07)60142-0.

Yin Y, Chen Q, Jin L, *et al.* The effects of deep convection on the concentration and size distribution of aerosol particles within the upper troposphere: A case study. *J Geophys Res Atmospheres* 2012;117(D22); doi: 10.1029/2012JD017827.

Address correspondence to:
Daniel J. Brener
The Higgs Centre for Theoretical Physics
University of Edinburgh
James Clerk Maxwell Building
Peter Guthrie Tait Road
Edinburgh
EH9 3FD
UK

E-mail: daniel.brener@ed.ac.uk

Submitted 11 April 2022
Accepted 11 December 2022
Associate Editor: Don Cowan

Abbreviations Used

ELEO = extremely low Earth orbit
ELEO-SAT = TSAT Globalstar ELaNa-5 ELEO satellite
MLT = mesosphere and lower thermosphere

4.3 Conclusion

This chapter has re-opened the question of how we might be to able to define the biosphere's extent in altitude by tackling two key problems. First and foremost, the issue of contamination through the design of a sampling detector that determines the relative velocity and size of incoming particulates, and thence filtering sampled air to mitigate against contaminants. By careful launch procedures, trajectory and orbit design, other sources of contamination could also be managed. The key theoretical insight to the problem is that the particle of interest that is to be measured should be at a different relative velocity to the detector. Hence through filtering out such particles one removes self-contamination from the detector.

To make calculations, careful consideration of the flow characteristics were required by computing the Knudsen number for different particulate sizes. Ideally, one would accurately model the flow solution to take care of the non-equilibrium rarefied gas flow through the detector. However to zeroth order, I found that most viruses and larger biological particles have sufficient size to be in continuum flow conditions, which was verified by previous simulation studies of similar flows. In other words, the Navier-Stokes equations for laminar flow, which are effectively Newton's Second Law, are approximately suitable.

The next step will be to develop a complete geometric detector design and subject it to simulation flow studies. This would not be a trivial undertaking as it would require the development of a simulated flow that can span the continuum, transition and rarefied gas conditions. A study comparing a continuum-based Navier–Stokes and a stochastic Boltzmann solver for rarefied nitrogen gas (non-atmospheric) flow showed that both need to be coupled to efficiently simulate rarefied gases [95]. Using DNS such as in Chapter 2 is not feasible due to the high number of degrees of freedom and detector scale. Solving the fundamental problem of turbulence, as in Chapter 2, would hopefully enable less computationally expensive models i.e. via development of digital-twins. The second key problem was the difficulty of sampling in the MLT. Mission strategies for the use of CubeSats in ELEO which can sample a wide altitude range above 100 km were designed. Discovering life at any altitude above even the upper-stratosphere would be a significant finding. Next I consider the problem of contamination in a different setting, through the dynamical system of microbial birth, death and how this relates colonisability. Can any contaminant colonise?

Chapter 5

Rethinking planetary protection: an island biogeographical analysis

The sky was so starry, so bright that, looking at it, one could not help asking oneself whether ill-humoured and capricious people could live under such a sky.
–Ф. Достоевский (F. Dostoyevskii), translated by Constance Garnett in 1918

This publication originally appeared in Journal of the Royal Society Interface Volume 22, Issue 227 on 25th June 2025 DOI:10.1098/rsif.2025.0079.

5.1 Introduction

In the last chapter, solutions were developed to the detection of microbiology in the upper atmosphere through, as in the previous chapters, reducing this complex problem to simple consideration of a fluid dynamical system. In this chapter, I focus more on the dynamical systems aspect although not for fluids, instead the system is that of microbial growth. The issue of planetary protection, both in terms of forward and back contamination is revisited from this perspective, as inspired by island biogeography. The current approach to planetary protection relies on probabilistic models to estimate the likelihood of contamination during space exploration [1, 26, 80]. These models neglect the question as to whether microbial life can sustain a long-term population successfully.

In defining successful colonisation as a microbial population in an environment conducive to a large mean-time to extinction, we propose a new methodology for calculating planetary protection risk. We identify that the birth-death rate ratio, uniquely determined by genomics and cellular capabilities, is the key variable for determining whether net population growth with long mean-time to extinction is viable in multi-parameter extreme environments. This corresponds

to determining theoretically and experimentally the phase-space of the microbial population dynamical system for birth and death.

INTERFACE

royalsocietypublishing.org/journal/rsif

CC BY

Research

Cite this article: Brener DJ, Cockell CS. 2025 Rethinking planetary protection: an island biogeographical analysis. *J. R. Soc. Interface* **22**: 20250079.
https://doi.org/10.1098/rsif.2025.0079

Received: 21 January 2025
Accepted: 25 April 2025

Subject Category:
Life Sciences–Physics interface

Subject Areas:
astrobiology, biogeography, biophysics

Keywords:
planetary protection, biogeography, interplanetary, life, population dynamics

Author for correspondence:
Daniel J. Brener
e-mail: : daniel.brener@ed.ac.uk

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

Rethinking planetary protection: an island biogeographical analysis

Daniel J. Brener[1] and Charles S. Cockell[2]

[1]Higgs Centre for Theoretical Physics, University of Edinburgh, Edinburgh, UK
[2]UK Centre for Astrobiology, School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK

DJB, 0000-0003-3340-0177; CSC, 0000-0003-3662-0503

We reconsider the problem of planetary protection using, by the analogy of planets as islands, the theory of island biogeography. We show that although the notion of equilibrium populations that emerge from the effects of immigration and extinction generally breaks down when applied to interplanetary scales, the mean-time to extinction resulting from the combined effects of growth and death rates can be quantified. We reconsider the probabilistic model of planetary protection, discuss how mean-time to extinction can instead be used to assess contamination risk, and we propose a research direction for planetary protection based on these ideas. We discuss more broadly the applicability of island biogeography to considering biotic transfer at interplanetary scales.

1. Introduction

Can life be successfully transferred between planets? This question has been explored in both natural contexts (panspermia) [1] and human-aided transfer, whether deliberate [2] or unintentional. The potential for unintentional transfer has led to discussions on planetary protection [3,4], focusing on the risk of contaminating other planetary environments (forward contamination) and the potential harm to Earth's biosphere from extraterrestrial life (back contamination).

Numerous attempts have quantified the factors involved in interplanetary life transfer. In planetary protection, the favoured approach is probabilistic, assigning probabilities to organisms' survival and growth in various transfer phases, such as surviving the interplanetary journey and successful growth in the destination environment. This approach can place quantitative numerical limits on microbial loads, for example, on specific missions [5]. For instance, the Committee on Space Research (COSPAR) policy on planetary protection, followed by NASA [6], directs that the probability that a planetary body will be contaminated during any one exploratory mission to Mars shall be no more than 1×10^{-3} [7]. In the case of the Europa Clipper mission, it was estimated that the probability of contaminating the moon was 2.25×10^{-5} [5].

The most common method for calculating these probabilities is the Coleman–Sagan equation, which gives the probability of contamination, P_c:

$$P_c = n_0 R P_s P_I P_R P_g, \tag{1.1}$$

where P_s, P_I, P_R and P_g are the probabilities that the microorganisms reach the planetary surface, the spacecraft impacts the planet, the microorganisms are released to the environment post-landing and for growth, respectively, and n_0 and R are the number of microorganisms initially on the spacecraft and their rate of reduction due to conditions on the spacecraft before and after launch, respectively [8]. The terms within this approximation capture the same phases of transfer that can be estimated for the natural transfer of life between planets [1].

However, critiques of the probabilistic method have emerged, notably in the National Academies of Sciences (NAS) report on planetary protection for icy

© 2025 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

THE ROYAL SOCIETY PUBLISHING

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

Table 1. Planets as islands—a comparison in biogeography.

	physical systems		
biogeography mechanism	terrestrial islands	planetary islands	effect on population
dispersal mechanism	prevailing winds; ocean current; self-propelled travel	gravitational potentials; gravitational perturbations; rock/dust fields; tidal evolution	increase/decrease immigration rate
dispersal filter	mountain ranges; island separation	interplanetary journey time (distance); atmospheric re-entry; shock pressure and temperature during ejection	reduces immigration rate
changing geophysical dynamics	plate tectonic shift; climatic changes	gravitational drift; plate tectonic shift; climatic changes	over long times can increase or decrease immigration rate

solar system bodies [9]. The NAS report highlighted substantial limitations of probabilistic models, such as uncertainties in the magnitude of bioload reduction factors and the problematic assumption that individual contamination factors are independent and can be estimated with sufficient statistical robustness to be useful in prediction.

Probabilistic parameters provide quantitative bounds on the chance of an organism reaching a planet. However, once the organism is there, the chance of colonization is not probabilistic, but rather determined by whether the environment is suitable for replication and, if it is, whether the growth and replication rate can exceed the rate of processes causing loss of viability and death, resulting in net population increase or, at the minimum, a state of viable dormancy. The factors that lead to stable population establishment in new isolated environments have been considered for many decades and were famously synthesized in the work of MacArthur & Wilson [10] on island biogeography. The applicability of island biogeography and some of its ideas to the interplanetary (even interstellar) transfer of life has been considered before [11–13], but it has not been critically examined, and surprisingly, there has been no attempt to apply its principles to planetary protection. In this paper, we use island biogeography to reconsider the probabilistic approach to planetary protection, proposing establishment of a population as the primary criterion.

2. Theory

2.1. Island biogeography

MacArthur & Wilson's [10] *The Theory of Island Biogeography* unites ideas in both biogeography and ecology to give quantitative estimations for inter-island population dynamics [10]. The credo of the theory is that the number of species on an island is determined by both the size of the island and its distance from the mainland or other islands. The balance between these factors provides a dynamic equilibrium of immigration I and extinction ϵ, such that one has a net population size with a given number of species S,

$$\frac{\mathrm{d}S}{dt} = I - \mathcal{E}. \tag{2.1}$$

The step of viewing planets as analogous to islands in terms of biogeography is a simple one; both planets and islands are physical locations with a relatively high degree of separation, and in both cases, there are phenomena that can carry small physical objects between them. As locations, they may both have habitable and uninhabitable environments for specific biota. Thus, on face value, equilibrium theory would be applicable to understanding interplanetary population dynamics. In table 1, we illustrate comparisons between the island biogeography theory and the interplanetary application, focusing on the physical analogies that control the rate of immigration of organisms. In figure 1, we revise the classic result from MacArthur & Wilson [10] for speciation in a multiplanet system.

One aspect where island biogeography for planets breaks down is that, unlike small Earth islands, propagules arriving on a planet generally have a lower probability of landing near previous immigrants due to the vast planetary surface area, scaling with the square of the planetary radius. For most realistic scenarios, this would reduce the effective immigration rate to the point where it can be considered irrelevant in the equilibrium population. For Earth, assuming propagules randomly fall across the surface, the chance that any two land within 10 km is about 10^{-6}.

MacArthur and Wilson define successful colonization by a long mean-time to extinction for an immigrant population. This idea is the basis for our suggestions on planetary protection. Let μ and λ denote the birth and death rates per unit time of a single-species population, respectively. According to their model, the mean-time to extinction t_k for an equilibrium population of size k is approximately $\mu/(\mu - \lambda)$. This assumes that the population is free to grow until reaching some maximum size, determined by the environment through its resources, known as the environment carrying capacity, k. In figure 2, we illustrate from [10] how the birth–death rate ratio μ/λ can limit the mean-time to extinction. There are other models of population growth but that of MacArthur and Wilson comes from first principles such that these concepts are the general case in most models [14,15].

MacArthur and Wilson's original theory was developed specifically for terrestrial island ecosystems. Consequently, we cannot directly apply their framework to planetary-scale environments without appropriate modifications. For example, although the same power-law relation for species with area, i.e. $S \sim A^{\text{constant}}$, has been found for both microbes and mammals, a study found

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

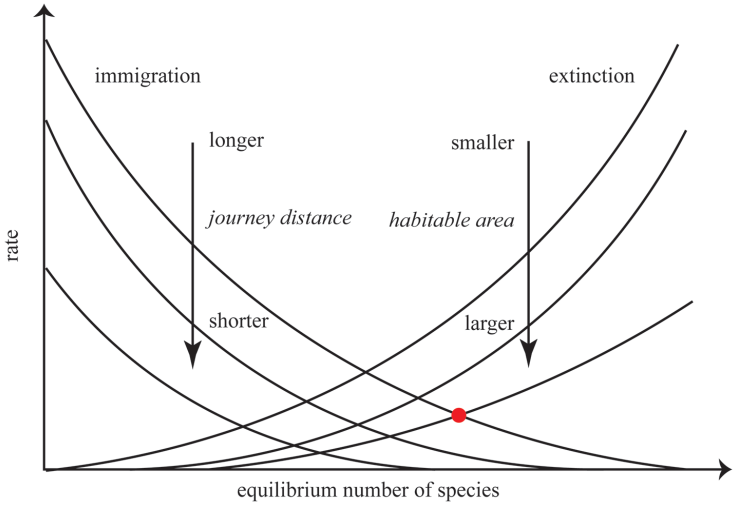


Figure 1. Equilibrium model for multiple planets separated by different journey times from a planet with a productive biosphere and different habitable area sizes. The intersection of the immigration and extinction curves gives the average number of species present on the colonized planet. The red point indicates the maximal species situation. This figure assumes that new species are not being created by speciation events and only by the balance of immigration and extinction. Adapted from [10].

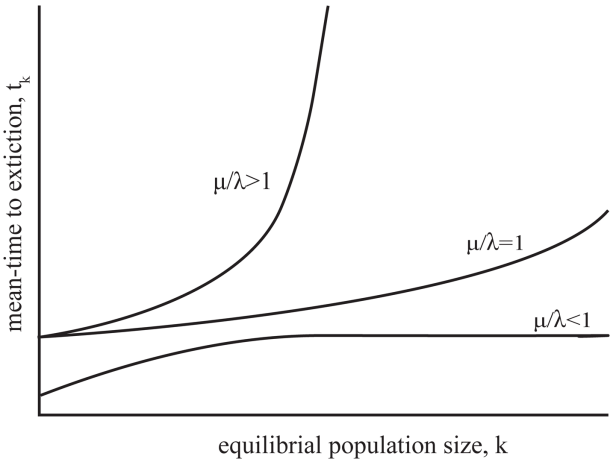


Figure 2. Mean-time to extinction curves for different birth–death rate ratios with population size. For $\mu/\lambda \geq 1$, t_k is unbounded. Note that to be consistent with the notation in microbiology with growth as μ, we denote death rates by λ, opposite to that of MacArthur and Wilson, where they use λ and μ as the rates of birth and death, respectively.

that microbial speciation was more strongly driven by environmental heterogeneity [16]. Our framework begins with the analogy between planets and islands, which we argue gives good grounds to expect the general trends of immigration and extinction rates with area and distance to remain. What we develop more specifically is the theory for the mean-time to extinction, as this can be applied to microbial species.

Also of importance is that we are considering the transfer of microorganisms through human agency, which is different from the natural processes originally considered by MacArthur and Wilson. The introduction of microorganisms to other planetary bodies, in the case of planetary protection concerns, is linked to spacecraft construction, deployment and exploration. On Earth, anthropologists and archaeologists have, for decades, examined analogous processes using island biogeography, assessing how human-mediated settlement impacts previously pristine ecosystems on Earth. These historical and archaeological studies document extensive ecological transformations, including numerous extinctions and substantial alterations to native species composition resulting from both intentional and accidental introductions of biota [17–22]. However, whether these same concerns, which have largely focused on macrofauna, apply to microbial transfer across interplanetary space is the question that must be addressed.

2.2. Microbial birth and death rates

To illustrate how environmental extremes influence the birth–death rate and thus colonization success, we can choose any parameter relevant to microbial growth. For example, consider temperature and how it may bound $\frac{\mu}{\lambda} > 1$, necessary for an unbounded mean-time to extinction. The general functional form for the thermal growth (birth) rate has been well studied (see [23] for a review). We consider a generalized form describing the effect of temperature on microbial growth over the entire range:

$$\mu(T) = AT\mathrm{e}^{-\frac{a}{T}}\left(1 - \mathrm{e}^{b(T-T_{\max})}\right), \tag{2.2}$$

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

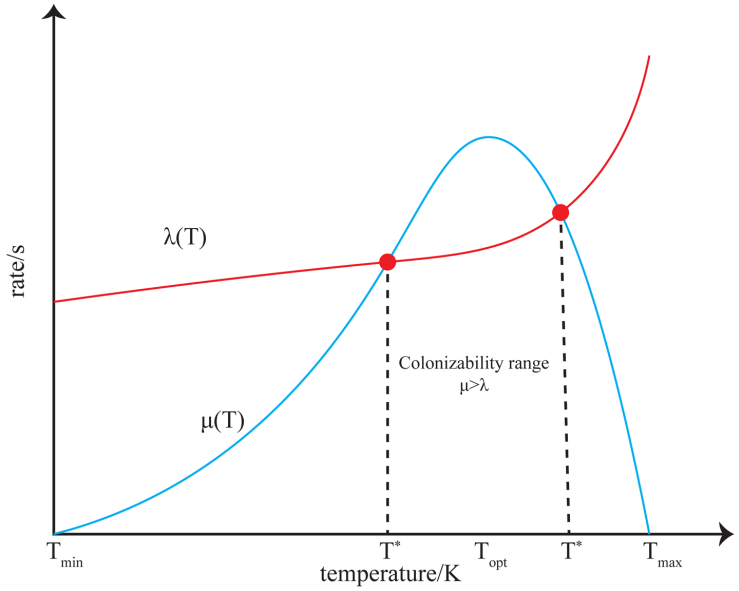


Figure 3. Birth (blue line) and death (red line) rates as a function of temperature. Red points mark the zero net population growth thresholds, where the death rate becomes higher than the birth rate, which results in eventual extinction.

where T is temperature, $\{A, a, b\}$ are a set of real coefficients and T_{max} is the first highest temperature at which the birth rate is zero [24]. The death rate with temperature has not been as well studied, and there are no theoretical models available for the complete temperature range. Here, we use an approximation for the sake of demonstrating the conceptual idea (figure 3), but we can physically motivate some possible functional forms.

Although vitrification may improve survival at low temperatures, organisms still lose viability [25], so there is always background death at any temperature. Below the minimum temperature for birth rate, T_{min}, extinction is inevitable. Therefore, death rates might be approximately constant or increase monotonically as temperature decreases. At high temperatures, thermal inactivation kinetics and Arrhenius-type reactions dominate, leading to deaths; thus, an exponential dependence is expected. In the mid-range, growth processes dominate, making the death rate comparatively negligible and its variation unclear. Various functional forms satisfy these constraints, but generally one would expect

$$\lambda(T) \sim f(T) \cdot g(T, T_{opt}, T_{min}) + e^{\beta T}, \tag{2.3}$$

where f, g are functions that take care of low temperature and intermediate regions, respectively.

Planetary protection concerns the points where $\mu(T^*) = \lambda(T^*)$, with T^* being the equilibrium temperature above or below which the population eventually becomes extinct (figure 3). These transition regions are not commonly investigated since disentangling microbial death and birth is experimentally challenging [26]. Generally, microbiologists measure the lag phase and doubling time when quantifying microbial populations. However, these measurements, although sometimes simple to make, are a combination of the underlying metabolic and physiological factors that lead to microbial replication and the factors that are leading to loss of microbial viability and death, whereby a net population increase observed in an experiment occurs when the former dominates over the latter. As we discuss here, since the functions that describe the factors that lead to replication and those that lead to loss of viability and death are not necessarily simple or correlated to one another, ideally one would disentangle them in order to be able to both understand the biochemical basis of the net population growth and model it for predictive purposes.

Cells can also enter a dormant state, remaining viable but ceasing replication. If relocated or conditions change, they may replicate again, posing concerns for planetary protection. There is no universal set of physiological requirements that define dormancy [27], but if considered as a balance between maintenance processes and loss of viability, dormancy corresponds to cells near the intersection of birth and death rates, especially at low temperatures (see figure 3) and/or limitations of energy or nutrients. Again, a grasp of the mathematical functions of the birth and death curves could allow us to predict the intersections of those curves and thus where these borderline dormancy states may lie.

2.3. Rethinking successful colonization

Defining *success* for a planetary protection strategy depends on whether one is concerned with the contamination of scientific results or of other planets' environments, or both. For the former, a probabilistic framework is appropriate as only the transfer process and the chemical integrity of the contaminant matter are appropriate [5]. Since we can never be certain what microbes may be on a spacecraft, a key initial inquiry should be: Can any known microbe achieve prolonged dormancy or net population growth in the destination conditions? If yes, one would make a probabilistic assessment of its transfer likelihood; otherwise, such assessment is not warranted. Thus, probability is relevant until the point of arrival on the extraterrestrial body. Methods to calculate probabilities, such as stochastic binomial models, were originally designed for evaluating spacecraft sterilization procedures

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

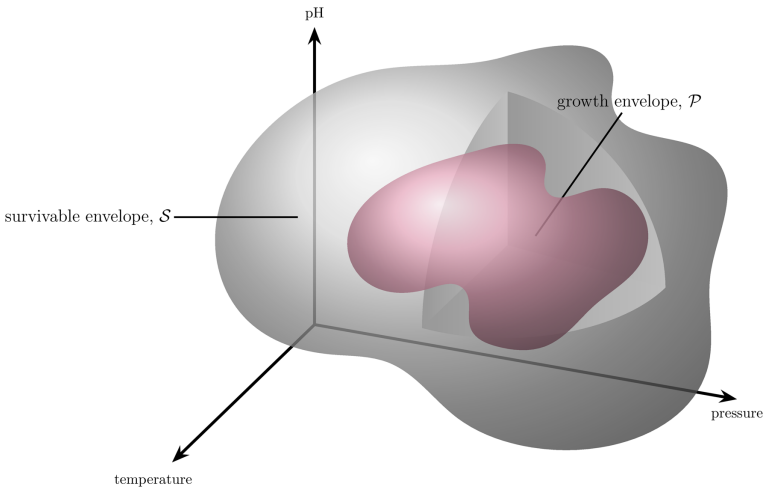


Figure 4. Three-dimensional example phase space for the mean-time to extinction growth threshold. Purple sub-space represents the parameter space beyond which the mean-time to extinction is too short for colonization. Grey sub-subspace is the habitable zone, beyond which the organism will die.

[28], and not for predicting successful colonization or population establishment in new environments. We illustrate the difference by considering the phase space of the mean-time to extinction function.

We define successful colonisation as

$$t_k \geq t_c, \tag{2.4}$$

where t_c is a critical time beyond which the species is considered established—this could be the planet's habitable lifetime t_{hab} or a mission timeframe. The mean-time to extinction is a multivariate function

$$t_k = t_k(T, p, \mathrm{pH}, k, \lambda, \mu, M, x), \tag{2.5}$$

where T,p,pH are the temperature, pressure, and acidity of the environment with carrying capacity k, M is the mutation rate and x is the initial population size. Determining t_k is a highly multidimensional, nonlinear problem since microbes are capable of rapid reproduction, short-term physiological adaptation, and beneficial genetic alterations (e.g. horizontal gene transfer). Consider a d-dimensional phase space where each dimension corresponds to a parameter in t_k. Restricting t_k to three dimensions for simplicity, the phase space is

$$\mathcal{P} = \left\{(T, p, \mathrm{pH}) \in \mathbb{R}^3 \mid t_k(T, p, \mathrm{pH}) \geq t_c\right\}, \tag{2.6}$$

which defines a geometric surface beyond which colonization is not possible (figure 4). Within the survivability envelope lies the space for net population growth. A microbe might initially *survive* but not *replicate*. The idea to tackle these issues using concepts from dynamical systems theory has, as far as we are aware, not before been considered. In the next section, we elucidate further how these spaces help us see colonizability as binary.

Microbiological life can establish populations without going extinct in extreme conditions far outside those *optimal* for growth. For example, cyanobacteria in extreme cold environments like the Arctic and Antarctic have a higher optimal growth temperature than their habitat [29]. Despite slow growth, their mean-time to extinction is large, so the population is well-established. Thus, although growth is small, $\mu/\lambda > 1$ remains. A small body of literature seeks to define these phase spaces for microbiology, but data are limited [30–32].

Various non-environmental and non-species-specific mathematical models provide a mean-time to extinction function [10,33–36]. These could be used to construct parametrized models from experimental data on how environmental parameters influence growth and death. This task is complex as models have varied functional forms, e.g. exponential versus power-law relationships for environmental versus demographic stochasticity [35]. These functions depend on carrying capacity, which fluctuates and is driven by complicated processes (e.g. [31]). All these features vary between planetary environments.

3. Planetary protection: a new approach

Currently, planetary protection risk is assessed through an estimated 'probability for growth', P_g. However, successful growth of an organism is not probabilistic; either the birth rate exceeds the death rate, leading to a net population increase, or it does not. Similarly, at the intersection of these rates, conditions either allow for prolonged dormancy over some defined time period or they do not. Thus, a simpler approach is to consider whether the environment lies outside the colonizability domain (short mean-time to extinction) for the microorganism's phase space. Each species has a definitive survivable envelope outside of which $P_g = 0$ as the organism is dead, but within which there is a zone defined by parameters sufficient for dormancy or replication to maintain or establish a population (long mean-time to extinction), where $P_g = 1$.

These spaces define a binary option: either the planetary conditions permit colonization or they do not. At extremes, we already consider the survival of an organism to be binary; beyond certain temperature thresholds on spacecraft surfaces, we conclude no

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

life can survive. Conversely, in locations with liquid water, P_g is typically set to unity [8], as the presence of liquid water is sometimes assumed (erroneously) to imply habitable conditions. Many other variables, such as salinity and pressure, define the edges of survival and microbial growth/reproduction. A probabilistic argument is misleading, as it implies that with a sufficiently large n_0, if $P_g > 0$ (which presently seems arbitrarily assumed [8]), some proportion of the contaminant organisms will survive *and* establish a population. For microbes outside their optimal growth conditions, the probabilistic approach may be misleading since even outside the optimum, P_g is simply 1 no matter how slow growth is.

The generally non-probabilistic growth of microorganisms is evident in the science of microbiology. If growth under specific conditions were probabilistic, then since the emergence of microbiology in the seventeenth century, microbiologists would have empirically created tables of growth probabilities in given medium formulations. Every time they wished to grow an organism, they would consult these tables and prepare a given number of flasks to ensure that one or two of them had successful growth. However, this is not how the science operates. Instead, they use a single medium; if growth fails, they assume conditions are unsuitable and try different media. This reflects a largely binary problem.

This is because reproduction cannot partially occur; a cell either divides or it does not. Division depends on whether nutrients, energy, physico-chemical conditions, inter alia, are sufficient for replication. Net population growth depends on the replication rate exceeding the death rate. Stochastic events (e.g. local variations in conditions or the metabolic state of a cell operating at its thermodynamic limits) may introduce some probability into successful growth, especially with small initial populations. Probabilistic models can also describe exit from dormant states like bacterial spores [37], predicting the number of spores that germinate. While we acknowledge probabilistic models, we suggest that once an actively metabolizing cell is in an environment, a single probability term for successful colonization is inadequate. It seems parsimonious to assume that if conditions support growth, stochastic events will not prevent it, allowing us to consider the mean-time to extinction based on birth and death rates to assess colonization. Our approach, in *assuming* that microbial life will reach extraterrestrial bodies, is arguably also a more cautious one. These arguments apply to assessing whether environmental and energetic conditions allow for the persistence of dormant cellular states over some defined time period.

For planetary protection, we propose an alternative procedure aiming to determine whether the mean-time to extinction is shorter than a reasonable timescale (e.g. mission or exploration lifetime); if so, no planetary protection risk is concluded. Detailed environmental studies must evaluate parameter ranges for key microbiological variables (e.g. temperature, pressure, acidity, medium content), indicating which microorganisms find the environment habitable. We then determine whether $\mu \geq \lambda$ for given planetary parameters, in which case colonization is inevitable. If $\mu < \lambda$, we need to estimate experimentally the mean-time to extinction to see if it is long enough to be considered colonization, i.e. similar to geophysical timescales or longer than mission windows.

For many microbes, birth and death rate data are unknown, even for single stressors like temperature. On planets like Mars, numerous stressors—salinity, pressure, radiation exposure, perchlorate toxicity [38,39]—create different genetic pressures on microorganisms. Therefore, we need a theoretical method, constrained by experiments, to estimate the mean-time to extinction for populations on other celestial bodies. Planetary protection research cites many knowledge gaps, but this issue has not been previously identified.

3.1. Problem of measuring microbial death

Death in microbes, particularly at low temperatures, can be challenging to define. In the absence of replication, dormant states may still allow sufficient metabolic activity to overcome inescapable background rates of detrimental processes such as racemization of amino acids, depurination of DNA and thermal degradation over time [40]. Yet, there is always a background level of population decline, albeit potentially very slow. This is consistent with the MacArthur and Wilson model: for $\mu/\lambda < 1$, the mean-time to extinction does not scale indefinitely with population size, though it may still be long, set by the doubling time, yet death remains present.

While birth and death rates both depend on cellular metabolism, they must be isolated to model them effectively and determine which processes, metabolic, environmental or otherwise, define the birth curve (i.e. processes that maintain a cell in a viable state and potentially allow for replication) or lead to death (i.e. processes that cause loss of viability) [41]. More in-depth studies of metabolic and bioenergetic pathways, molecular biology and the interactions of cells with their environments could allow for better quantification of these curves and thus lead to better predictive models under stated environmental conditions. Most experiments are conducted under fixed laboratory conditions, but environments like Mars present diurnal temperature fluctuations [42] that introduce stochasticity into population dynamics. These temporal and environmental changes must be considered when assessing microbial population dynamics [43]. Stochasticity in genome-expressed pathways makes predictions extremely challenging. However, with sufficient computing resources and more metabolic data, it should be possible to combinatorially check gene combinations to assess optimal situations, similar to [44].

3.2. Colonizability: an example

Currently, we lack separate replication and death rate curves across full temperature ranges. For example, figure 5 presents the growth rate of the thermophile *Geobacillus stearothermophilus* as a function of temperature from [45]. While growth rates are measured during the exponential phase—where death rates are usually considered negligible—this assumption breaks down near critical temperatures ($\mu(T^*) = \lambda(T^*)$). In these regions, growth rate measurements may not accurately reflect replication rates due to significant death rates. By including a crude death rate curve following MacArthur and Wilson's approach, we highlight the

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

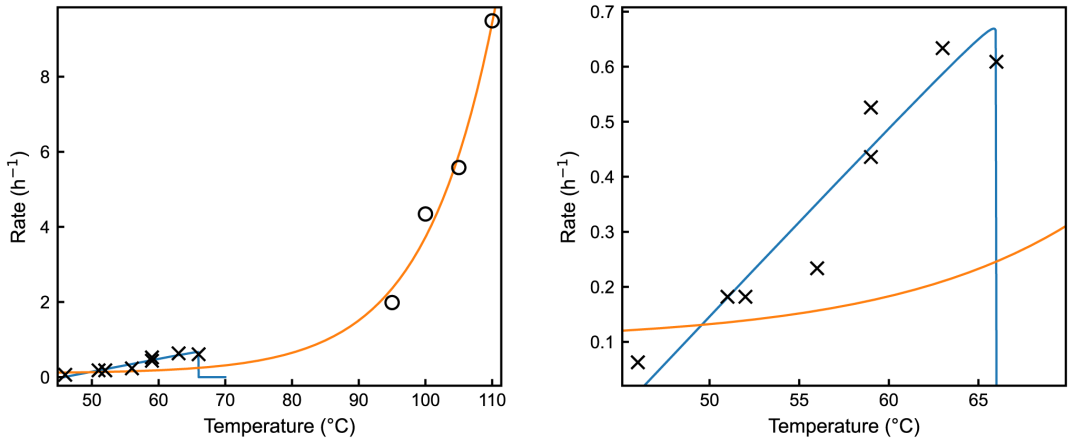


Figure 5. Left: growth rate of *Geobacillus stearothermophilus* as a function of temperature from [45] (black crosses), fitted with the Cardinal model (blue line). Thermal death rate (open black circles) from [46] fitted with an Arrhenius-type model (orange line). Right: same plot as on the left, enlarged to show intersections of the death and growth rate curves. Either side of the intersections, the mean-time to extinction would be finite and will not positively co-vary with the population's growth rate.

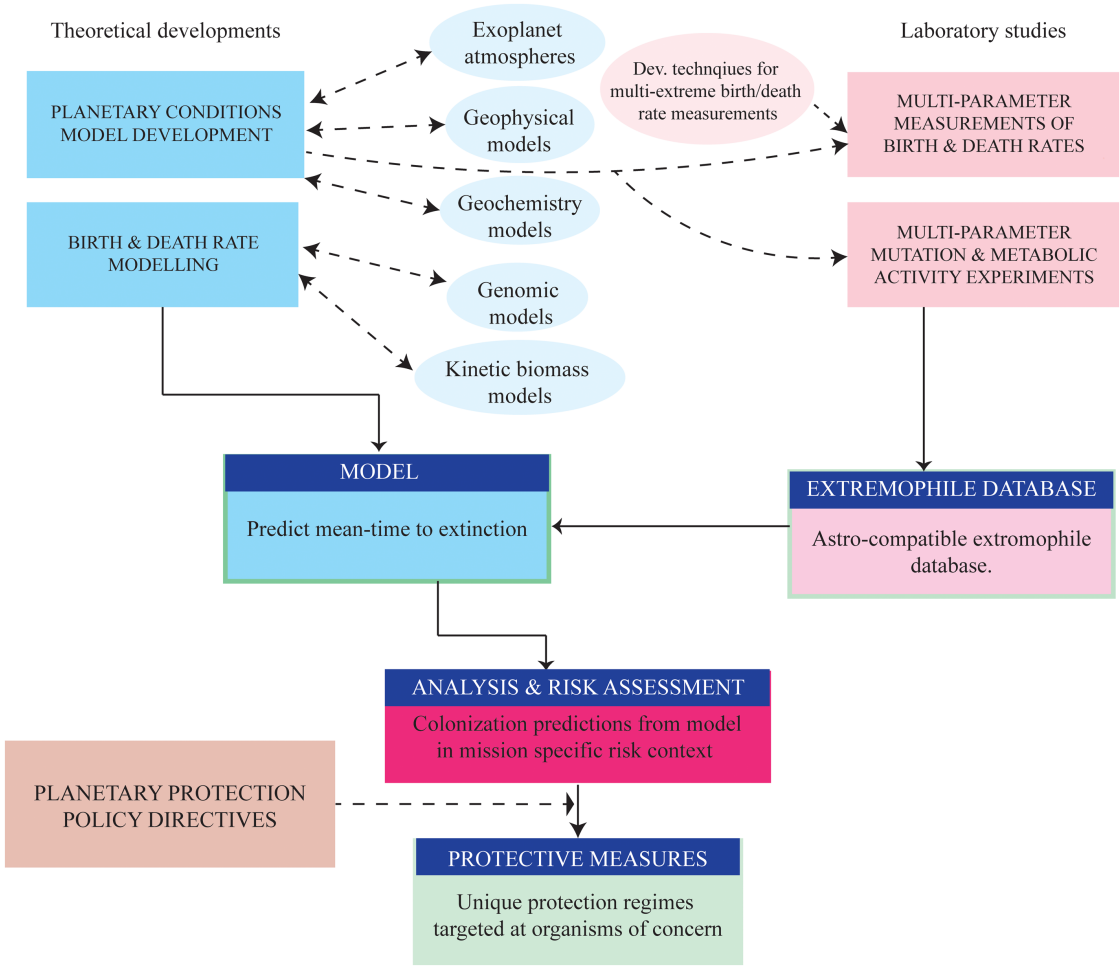


Figure 6. A roadmap to a new planetary protection framework.

uncertainties in extrapolating to low temperatures with limited data. This illustrates how replication and death processes are distinct and can lead to different mean times to extinction near critical temperatures, even when resources are sufficient. Clearly, without separate models for replication and death processes, our ability to make accurate predictions is limited, especially at the intersections of these curves, near which cells may experience dormant states.

Theoretical and laboratory experiments should eventually enable us to predict whether a microbe with specific characteristics would exceed a certain mean-time to extinction under specific planetary conditions. Alternatively, since we never know the full inventory of organisms on a spacecraft, one could collect physical and chemical data from an extraterrestrial environment and develop a database of known organisms capable of achieving net population growth or dormancy under those conditions. This *extremophile database* would be useful for two reasons. First, it defines organisms of concern—if they are mission contaminants, they could colonize or persist in the destination environment. Second, these organisms could be used for deliberate introduction

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

Table 2. Planetary protection programme of microbial research, including techniques.

	objective	techniques	references
modelling development	microbial colonization models based on mean-time to extinction, incorporating environmental stressors	Arrhenius-based models for microbial growth and death rates under extreme conditions, phase-space diagrams to visualize colonization limits	[23,49]
	death rate models for microbial populations under multiple stressors	kinetic inactivation models	[46,51]
	create extremophile database cataloguing survival profiles across diverse planetary environments	multivariate extinction models, parametrize theoretical models	[51,52]
	dynamic model development	stochastic factors such as diurnal temperature fluctuations	[42,52]
laboratory studies	measurement of microbial birth and death rates across stressors range (e.g. temperature, pressure, acidity)	continuous culture techniques, fluorescence microscopy	[23,25]
	measurement of microbial death processes (e.g. apoptosis) in extreme planetary environments using simulations of e.g. Mars-like conditions	environmental chambers, real-time high-content live-cell imaging	[26,31,38,53]
	bioenergetics studies of microbial replication and death under stress to distinguish between survival and reproductive processes	metabolomics, mass spectrometry, respirometry, time-lapse microscopy	[41]
	validation of colonization thresholds for various microbes using extreme environment simulations to estimate the mean-time to extinction	Mars simulation chambers, environmental chambers, phase-space analysis	[26,30,39]

to a planetary environment one aims to inoculate or terraform. It may also be possible to use theoretical considerations, such as environmental energy availability, to predict which organisms might achieve net population growth [47,48]. Figure 6 illustrates a roadmap for developing such a framework with practical aims for the planetary protection community. Such predictive capabilities, linked to policy, already exist in the food safety community, where laboratory experiments and theoretical models predict microbial contamination [49,50]. Our suggestions are summarized in table 2, including specific techniques.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. The growth rate data presented in figure 5 were provided to us on request by the author of the original study [45]. The death rate data in figure 5 were taken directly from table 2 in the paper [46].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. D.J.B.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, visualization, writing—original draft, writing—review and editing; C.S.C.: conceptualization, formal analysis, investigation, methodology, project administration, resources, supervision, visualization, writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. The authors were supported by the United Kingdom Science and Technology Facilities Council, with C.S.C and D.J.B. under grants ST/V000586/1 and ST/W507441/1, respectively.

References

1. Clark BC. 2001 Planetary interchange of bioactive material: probability factors and implications. *Orig. Life Evol. Biosphere* **31**, 185–197. (doi:10.1023/A:1006757011007)
2. Tepfer D. 2008 The origin of life, panspermia and a proposal to seed the universe. *Plant Sci*. **175**, 756–760. (doi:10.1016/j.plantsci.2008.08.007)
3. DeVincenzi D, Stabekis P, Barengoltz J. 1983 A proposed new policy for planetary protection. *Adv. Space Res*. **3**, 13–21. (doi:10.1016/0273-1177(83)90167-9)
4. Coustenis A *et al*. 2023 Planetary protection: an international concern and responsibility. *Front. Astron. Space Sci*. **10**, 1172546. (doi:10.3389/fspas.2023.1172546)
5. McCoy KJ, DiNicola M, Everline C, Burgoyne H, Reinholtz K, Clement B. 2021 Europa clipper planetary protection probabilistic risk assessment summary. *Planet. Space Sci*. **196**, 105139. (doi:10.1016/j.pss.2020.105139)
6. NASA. 2022 *NASA Technical Standard: implementing planetary protection requirements for space flight*. Baseline version. Washington, DC: NASA.
7. COSPAR Policy on Planetary Protection. 2023 COSPAR Panel on Planetary Protection.
8. Kazarians GA *et al*. 2017 The evolution of planetary protection implementation on Mars landed missions. In *2017 IEEE Aerospace Conf., Big Sky, MT, USA*. (doi:10.1109/AERO.2017.7943576)
9. Council NR. 2012 *Assessment of planetary protection requirements for spacecraft missions to icy solar system bodies*. Washington, DC: The National Academies Press.
10. MacArthur RH, Wilson EO. 1967 *The theory of island biogeography*. Princeton, NJ: Princeton University Press.
11. Lingam M, Loeb A. 2017 Enhanced interplanetary panspermia in the TRAPPIST-1 system. *Proc. Natl Acad. Sci. USA* **114**, 6689–6693. (doi:10.1073/pnas.1703517114)
12. Cockell CS *et al*. 2007 Interplanetary transfer of photosynthesis: an experimental demonstration of a selective dispersal filter in planetary island biogeography. *Astrobiology* **7**, 1–9. (doi:10.1089/ast.2006.0038)

Downloaded from https://royalsocietypublishing.org/ on 24 June 2025

13. Cockell CS. 2008 The interplanetary exchange of photosynthesis. *Orig. Life Evol. Biospheres* **38**, 87–104. (doi:10.1007/s11084-007-9112-3)
14. Soulé ME (ed). 1987 *Viable populations for conservation*, pp. 1–10. Cambridge, UK: Cambridge University Press.
15. Mangel M, Tier C. 1993 A simple direct method for finding persistence times of populations and application to conservation problems. *Proc. Natl Acad. Sci. USA* **90**, 1083–1086. (doi:10.1073/pnas.90.3.1083)
16. Horner-Devine MC, Lage M, Hughes JB, Bohannan BJM. 2004 A taxa–area relationship for bacteria. *Nature* **432**, 750–753. (doi:10.1038/nature03073)
17. Finney BR. 1988 Solar system colonization and interstellar migration. *Acta Astronaut*. **18**, 225–230. (doi:10.1016/0094-5765(88)90103-8)
18. Finney BR, Jones EM. 1985 Voyagers into ocean space. Interstellar migration and the human experience (eds RF Ben, MJ Eric), pp. 164–179. Berkeley, CA: University of California Press.
19. Keegan WF, Diamond JM. 1987 Colonization of islands by humans: a biogeographical perspective. In *Advances in archaeological method and theory* (ed. MB Schiffer), pp. 49–92. San Diego, CA: Academic Press. (doi:10.1016/B978-0-12-003110-8.50005-0)
20. Fitzpatrick SM. 2022 'Detritus of a coming world': the colonization of islands as microcosms for human impacts on an interplanetary scale. In *Speciesism in biology and culture: how human exceptionalism is pushing planetary boundaries* (eds B Swartz, BD Mishler), pp. 65–93. Cham, Switzerland: Springer. (doi:10.1007/978-3-030-99031-2_4)
21. Leppard TP, Fitzpatrick SM. 2025 Island archaeology provides ecological and behavioral analogs for off-planet exploration and colonization. *Futures* **166**, 103544. (doi:10.1016/j.futures.2025.103544)
22. Holcomb JA, O'Leary BL, Fairén AG, Mandel RD, Wegmann KW. 2024 The emerging archaeological record of Mars. *Nat. Astron*. **8**, 1490–1492. (doi:10.1038/s41550-024-02439-w)
23. Noll P, Lilge L, Hausmann R, Henkel M. 2020 Modeling and exploiting microbial temperature response. *Processes* **8**, 121. (doi:10.3390/pr8010121)
24. Ratkowsky DA, Lowry RK, McMeekin TA, Stokes AN, Chandler RE. 1983 Model for bacterial culture growth rate throughout the entire biokinetic temperature range. *J. Bacteriol.* **154**, 1222–1226. (doi:10.1128/jb.154.3.1222-1226.1983)
25. Clarke A, Morris GJ, Fonseca F, Murray BJ, Acton E, Price HC. 2013 A low temperature limit for life on earth. *PLoS One* **8**, e66207. (doi:10.1371/journal.pone.0066207)
26. Wu R *et al*. 2024 Bacterial killing and the dimensions of bacterial death. *NPJ Biofilms Microbiomes* **10**, 87. (doi:10.1038/s41522-024-00559-9)
27. McDonald MD *et al*. 2024 What is microbial dormancy? *Trends Microbiol*. **32**, 142–150. (doi:10.1016/j.tim.2023.08.006)
28. Fredrickson AG. 1966 Stochastic models for sterilization. *Biotechnol. Bioeng*. **8**, 167–182. (doi:10.1002/bit.260080114)
29. Tang EPY, Tremblay R, Vincent WF. 1997 Cyanobacterial dominance of polar freshwater ecosystems: are high-latitude mat-formers adapted to low temperature? *J. Phycol.* **33**, 171–181. (doi:10.1111/j.0022-3646.1997.00171.x)
30. Harrison JP, Gheeraert N, Tsigelnitskiy D, Cockell CS. 2013 The limits for life under multiple extremes. *Trends Microbiol*. **21**, 204–212. (doi:10.1016/j.tim.2013.01.006)
31. Qiu Y, Zhou Y, Chang Y, Liang X, Zhang H, Lin X, Qing K, Zhou X, Luo Z. 2022 The effects of ventilation, humidity, and temperature on bacterial growth and bacterial genera distribution. *Int. J. Environ. Res. Public Health* **19**, 15345. (doi:10.3390/ijerph192215345)
32. Harrison JP, Dobinson L, Freeman K, McKenzie R, Wyllie D, Nixon SL, Cockell CS. 2015 Aerobically respiring prokaryotic strains exhibit a broader temperature–pH–salinity space for cell division than anaerobically respiring and fermentative strains. *J. R. Soc. Interface* **12**, 20150658. (doi:10.1098/rsif.2015.0658)
33. Thakur B, Meyer-Ortmanns H. 2023 Controlling the mean time to extinction in populations of bacteria. *Entropy* **25**, 755. (doi:10.3390/e25050755)
34. Wylie CS, Shakhnovich EI. 2012 Mutation induced extinction in finite populations: lethal mutagenesis and lethal isolation. *PLoS Comput. Biol*. **8**, e1002609. (doi:10.1371/journal.pcbi.1002609)
35. Lande R. 1993 Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat*. **142**, 911–927. (doi:10.1086/285580)
36. Ovaskainen O, Meerson B. 2010 Stochastic models of population extinction. *Trends Ecol. Evol*. **25**, 643–652. (doi:10.1016/j.tree.2010.07.009)
37. Barker G, Malakar P, Peck M. 2005 Germination and growth from spores: variability and uncertainty in the assessment of food borne hazards. *Int. J. Food Microbiol*. **100**, 67–76. (doi:10.1016/j.ijfoodmicro.2004.10.020)
38. Matarredona L, Camacho M, Zafrilla B, Bonete MJ, Esclapez J. 2020 The role of stress proteins in haloarchaea and their adaptive response to environmental shifts. *Biomolecules* **10**, 1390. (doi:10.3390/biom10101390)
39. Archer PD *et al*. 2019 Perchlorate on Mars — overview and implications In *9th Int. Conf. on Mars, 22–25 July 2019*, Pasadena, CA, USA.
40. Price PB, Sowers T. 2004 Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proc. Natl Acad. Sci. USA* **101**, 4631–4636. (doi:10.1073/pnas.0400522101)
41. Bae SY, Guan N, Yan R, Warner K, Taylor SD, Meyer AS. 2020 Measurement and models accounting for cell death capture hidden variation in compound response. *Cell Death Dis*. **11**, 255. (doi:10.1038/s41419-020-2462-8)
42. Atri D, Abdelmoneim N, Dhuri DB, Simoni M. 2022 Diurnal variation of the surface temperature of Mars with the Emirates Mars mission: a comparison with curiosity and perseverance rover measurements. *Mon. Not. R. Astron. Soc*. **518**, L1–L6. (doi:10.1093/mnrasl/slac094)
43. Huete-Stauffer TM, Arandia-Gorostidi N, Díaz-Pérez L, Morán XAG. 2015 Temperature dependences of growth rates and carrying capacities of marine bacteria depart from metabolic theoretical predictions. *FEMS Microbiol. Ecol*. **91**, fiv111. (doi:10.1093/femsec/fiv111)
44. Sun D, Liu Y, Zhang XS, Wu LY. 2018 CEA: combination-based gene set functional enrichment analysis. *Sci. Rep*. **8**, 13085. (doi:10.1038/s41598-018-31396-4)
45. Molinaro C, Bénéfice M, Gorlas A, Da Cunha V, Robert HML, Catchpole R, Gallais L, Forterre P, Baffou G. 2022 Life at high temperature observed *in vitro* upon laser heating of gold nanoparticles. *Nat. Commun*. **13**, 5342. (doi:10.1038/s41467-022-33074-6)
46. Ayeni AO, Samuel IT, Adekeye BT, Agboola O, Nwinyi OC, Oladokun O, Ayoola AA, Elehinafe FB. 2022 Inactivation kinetics and thermodynamics assessments of *Geobacillus stearothermophilus* during thermal sterilization for products safety. *South Afr. J. Chem. Eng*. **42**, 223–228. (doi:10.1016/j.sajce.2022.09.003)
47. Rogers KL, Amend JP, Gurrieri S. 2007 Temporal changes in fluid chemistry and energy profiles in the Vulcano Island hydrothermal system. *Astrobiology* **7**, 905–932. (doi:10.1089/ast.2007.0128)
48. Hoehler TM. 2007 An energy balance concept for habitability. *Astrobiology* **7**, 824–838. (doi:10.1089/ast.2006.0095)
49. Leporq B, Membré JM, Dervin C, Buche P, Guyonnet JP. 2005 The 'Sym'Previus' software, a tool to support decisions to the foodstuff safety. *Int. J. Food Microbiol.* **100**, 231–237. (doi:10.1016/j.ijfoodmicro.2004.10.006)
50. Membre J, Leporq B, Vialette M, Mettler E, Perrier L, Thuault D, Zwietering M. 2005 Temperature effect on bacterial growth rate: quantitative microbiology approach including cardinal values and variability estimates to perform growth simulations on/in food. *Int. J. Food Microbiol*. **100**, 179–186. (doi:10.1016/j.ijfoodmicro.2004.10.015)
51. Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH. 2012 Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol*. **10**, 497–506. (doi:10.1038/nrmicro2795)
52. Martiny JBH *et al*. 2006 Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol*. **4**, 102–112. (doi:10.1038/nrmicro1341)
53. Gelles JD, Chipuk JE. 2016 Robust high-throughput kinetic analysis of apoptosis with real-time high-content live-cell imaging. *Cell Death Dis*. **7**, e2493. (doi:10.1038/cddis.2016.332)

5.3 Conclusion

One of the key outcomes of this chapter is a clearer understanding of the birth-death dynamics that shape microbial populations in extreme environments. Unlike terrestrial ecosystems where species are frequently replenished, planets present a unique biogeographical challenge due to the vast distances and barriers to dispersal. Hence the survival of microorganisms on these planetary "islands" depends not only on their initial arrival but also on the delicate balance between birth and death, which is largely dictated by the environment and the adaptive genetic capabilities of the species to those conditions.

Drawing on concepts from dynamical systems theory, we illustrate how planetary protection can be reformulated as mapping population dynamic phase-spaces, specifically for growth and death. This helps to determine a binary condition for planetary protection rather than a probabilistic one. The process of determining the phase-space requires both theoretical developments and laboratory experiments. We propose that a database should be created which can return a list of organisms that fit certain colonisability profiles, based on mean-time to extinction. This extromophile database would have two applications: planetary protection and deliberate planetary terraforming.

In the paper we considered planetary protection in terms of forward contamination of other planets. The same theory also applies to backwards contamination of Earth, as well as any planetary body with a widespread, diverse biosphere. From biogeography theory, planetary protection for the Earth centres on the ability of immigrants to out-compete the Earth's existing biota for resources; a fight over k. This is complex because for the Earth's present-day biosphere, even if a single foreign organism landed in an area rich with sufficient organics and energy, it would face competition from the existing flora and fauna. Whereas for the early Earth, with a less well-established biosphere, the rate of impacts was higher [7, 25, 28, 61, 77, 118, 119], hence we might hypothesize that a theoretically higher rate of transfer of material into areas where competition may have been lower, could have led to a higher chance of colonisation. This has consequences for back-contamination planetary protection objectives.

Future work could seek to establish immigration and extinction rates using stochastic methods during the late heavy bombardment period in the early Solar system. Studies have found that transfer of rock capable of carrying life has

likely occurred from both Earth and Mars to all the terrestrial planets in the Solar system [81, 132]. An ultimate goal would be to integrate early Solar system models with a stochastic island biogeography theory model to estimate various biogeographic scenarios [89]. As human space exploration progresses, it is likely there will eventually be permanent bases on Mars with frequent interplanetary transfer of microbiological material. Future planning of such missions will need to consider the fundamental points made within this chapter to avoid unintended consequences or potentially to aid terraforming of Mars. Hence, part of future work may require revisions to planetary protection policy.



Figure 5.1 *Ship ready for return. Painting by Chesley Bonestell, reproduced from The Conquest of Space, The Readers Union, London, 1952.*

Chapter 6

Conclusions

Words can be meaningless. If they are used in such a way that no sharp conclusions can be drawn. –R.P. Feynman [37]

In this thesis I have presented theoretical investigations and (potential) solutions to problems which have to contend with complex phenomena over a wide range of scales, in a broad range of fields. Firstly, I have tackled problems related to turbulence, both in pure and applied contexts concerning the Earth's atmosphere; secondly, I have addressed problems related to planetary biospheres. This chapter will ultimately act to summarise the main findings, making clear the relations between the fields.

6.1 Theoretical physics as an approach

The major contributions in this thesis are new ideas and specific insights, which originate from a theoretical physics perspective. What this means is that each problem is addressed by reducing it to a set of physical and mathematical principles, so that a new solution or point of view is arrived at by theoretical argument and or calculations.

In the problem of turbulence, of Chapter 2, by using a suite of simulations I have given the first *reliable* indications that the Navier-Stokes equations may have an upper critical dimension of six. This study was the first systematic direct numerical simulation investigation of dimension using chaos as a measure of the Navier-Stokes equations from three to six dimensions. Due to the small box sizes and hence low Reynolds numbers, it is difficult to be more conclusive but this study is sufficient evidence to strongly motivate the development of costly, higher resolution simulations in more than three dimensions.

Then relatedly in Chapter 3, I have presented an entirely original hypothesis on one of the most significant problems faced by long-range ensemble weather prediction systems in the last decade, the signal-to-noise paradox. Taking

a dynamical systems perspective on ensemble forecasting, I argued that this paradox could arise as a result of interpreting the ensemble mean to be our best forecast of observations which relies on the ergodic theorem. Such an approach interprets the mean as the most likely phase-space trajectory. I show that the theorem results in three major constraints: the ensemble members must be perfectly exchangeable; the averaging window must be sufficiently long; and the ensemble size must be sufficient. To illustrate and argue this, I developed a set of Gedankenexperimente using a modified Galton board. Using my hypothesis, I have attempted to reconcile many different aspects of the paradox that have been proposed in the literature, including providing specific suggestions for future testing.

Moving into the upper-atmosphere, a potential solution to the contamination problem which has impinged on determining the extent of the Earth's biosphere in altitude is presented. We suggested a new method, relative-velocity filtered sampling, based on the observation that the particle of interest that is to be measured will have a different velocity relative to the detector. Thus one can eliminate contamination originating from the detector by filtering for the relative velocity. We illustrated the potential efficacy of an upper-atmosphere sampling campaign using very low Earth orbiting CubeSats which have sufficient velocity such that the intake of the surrounding air would have high enough velocity to sample meaningful volumes and ensure sufficiently high relative-velocities. This research, in proposing a contamination strategy, re-opens the important question of the maximum altitude at which life might be found on Earth.

This directly relates to the final research chapter where we introduce a new perspective on another contamination problem: planetary protection, through utilising ideas from island biogeography. Island biogeography theory provides principles for analysing how immigrants to new isolated ecosystems or environments can become established and some of its ideas have been applied to interplanetary transfer of life. We argue for a shift towards analyzing the mean-time to extinction of microbial populations as a more reliable metric for colonisation risk. This is in contrast to the probabilistic assessment paradigm in planetary protection, which occurs by our distinguishing between survival and colonisation. Probability is useful in assessing risk of transfer to another planet (survival), but once it arrives, colonisation is not probabilistic but deterministic, governed by the balance between microbial birth and death rates. Hence we reframe planetary protection as a binary issue, where a combination of biological

and environmental conditions either allow for microbial growth and long-term survival or lead to extinction. We have outlined a road map for developing theoretical models and experimental techniques that incorporate these factors to predict colonisation potential.

6.2 Role of dynamical systems theory

The first two research chapters are connected mathematically by dynamical systems theory, as Chapter 2's use of the Maximal Lyapunov exponent relies on the more general form of the ergodic theorem, the multiplicative ergodic theorem, which is central to Chapter 3. In fact, the finite time Lyapunov exponent technique to measure chaos in effect uses the principle of ensembles. I suggest this Lyapunov exponent ensemble-of-ensembles technique as a potential investigative tool to resolve the signal-to-noise paradox. Chapter 3 also connects weather forecasting with a host of other fields which all experience ergodicity breaking phenomena. These connections are a direct result of both problems being tackled from the perspective dynamical systems theory.

More broadly, dynamical systems theory provides an abstract framework with which one can transfer, visualise and solve seemingly different problems. When considering each of the systems, one can remove extraneous details and concentrate on critical points where the systems behaviour is problematic or changes in a particular way. In the final chapter, this was manifest in the realisation that the balance of birth and death are the definitive control on whether long-term colonisation is possible. Prior to reducing the system to this, one had a massively complex problem where inestimable quantities, such as local immigration rates and probabilities of growth, clouded the picture. Rather than the notion that the birth and death rates effectively provide turning points for the establishment of a population. Likewise with ergodicity breaking, a phenomenon which occurs at a critical point for certain dynamical systems, be them physically real-world or computational models.

As the very structure of this thesis has illustrated, the broad range of physical scales to which dynamical systems theory is applicable is indicative of its underlying scale invariance. This arises because dynamical systems theory is really a set of pure mathematics axioms which are applied on the principle that

the system evolves with time on some compact manifold[1], and as such, does not "care" about the physics and its associated scales. This applicability occurs for many physical systems as strong interactions with high numbers of degrees of freedom often generates scale invariance, especially close to critical points which are usually of most interest, as these are where behaviour changes [50]. It is also worth noting that dynamical systems often fall into universality classes. These are groupings of systems which share at least one scale invariant limit under the renormalisation technique [45]. In the case of Navier-Stokes equations, if they do turn out to have an upper critical dimension of six, then this could place them in the *undirected percolation* class for a mean field theory [67]. However, the relationship between spatial constraint, translation symmetry and the Navier-Stokes equations needs to be determined, as $d_c = 6$ alone is not necessarily sufficient for this universality class determination [133].

6.3 Future work

In the course of my research, I have identified several major directions for further work. In terms of turbulence, the next stage is very much in the realms of technological development. My pilot study gives good grounds to develop direct numerical simulations capable of resolving intermittency phenomena and therefore measure scaling in higher dimensions. This should just about be technically possible for four dimensions but the memory required for such a simulation is at the very limits of that available with national supercomputing infrastructure. Given my results, I think it would also be worthwhile to formally revisit the predictions of Liao [69, 70], taking his calculations to a higher order. If Navier-Stokes is a mean-field theory with a critical dimension, then this might enable development of simpler turbulence models with less intensive computing requirements through the development of digital-twins.

I have been made aware of a few research groups who are considering my hypothesis on the signal-to-noise paradox and how to test it for the North Atlantic Oscillation. Naturally a next step is also to see how failure of the ergodic theorem may affect other metrics that rely on first and second order ensemble statistics. Besides examining it in the data, further theoretical development should seek to extract how the ensemble size and time averaging lengths vary with the

[1] A space that locally looks Euclidean and does not extend off to infinity (non-infinitely bounded).

convergence of ergodicity. It would also be worthwhile to examine the use of ergodicity in earthquake forecasting to see potential connections and analogies which might be useful for weather forecasting on seasonal and longer timescales.

In terms of Chapter 4, future work will be similar in nature to some of that for Chapter 2, as both require further code development for fluids simulations. Testing the theoretical principle of velocity filtered sampling in a rarefied gas flow simulation would make the case for the development of the physical detector. Although, as I noted, development of such a simulation is not a trivial undertaking, so with assistance from a good engineer a prototype wind-tunnel (rarefied gas) test might more sensible.

Given the clear theoretical case made in the penultimate chapter, the next steps for planetary protection are extremely practical ones. Measurements of bacterial replication and death rates for a single stressor, most likely temperature as this seems to be the easiest control, need to be completed to make the first estimates of the critical points. This will require the development of more sophisticated techniques to measure microbial death. There are signs in the literature that such processes may be available but they do need consolidating. Once we have a good set of experimental data, it will be possible to develop better models of the microbial death processes. In tandem to this, I think further studies will need to clearly illustrate colonisability as a binary process, as my instinct is that this new concept will prove a challenge to the community. The theoretical argument made is solid, but it is conceptual and therefore it will take some time for it to be properly understood. This will be helped by the experiments of course, but also as the theoretical models for the various metabolic pathways for replication and death become established. As at present, there are no good models of the bacterial death process at low temperatures, which are important for the space environment.

Lessons for interdisciplinary research

Each of the research chapters within this thesis is situated in related, but distinctly different disciplines. I therefore feel it prudent to give some reflection in this final chapter on how using theoretical physics as an approach to these disciplines has been achieved, as there is some evidence of this within the research chapters themselves.

Throughout the majority of the research chapters, visualisations have been required to communicate, at times nuanced, theoretical arguments. These are not necessarily trivial actions. In the case of the ergodic theory, it was necessary to develop a whole new argument, thought experiment and reasoning based on the Galton board. What is particularly key here is that this analogy is physically and mathematically rigorous yet enables the idea to be communicated. Whilst in dealing with planetary protection, the notion of visualising the phase-space directly as a three-dimensional hypersphere was important for communicating both the complexity of the underlying dynamical equations and the idea that colonisability is a binary situation embedded within a broader space of survivability. Likewise in Chapter 4, illustrating the detector directly allows for those not familiar with the complex details of the upper-atmosphere to grasp the essential functions of the proposed detector.

Remaining humble more than being right, is, I would argue, an approach that is particularly important when coming into a different discipline, from theoretical physics. Different research communities, even within the same discipline, sometimes use different language to communicate the same ideas. Therefore repeating definitions, even if tautological, is not only necessary but essential for the communication of scientific ideas. From birth versus replication, to structure function versus first order longitudinal velocity difference, such examples can be seen throughout this thesis. In addition, careful understanding of the experimental or computational procedures, sufficient to build an accurate physical picture, is vital for then reducing the problem to the core physical principles.

The prevalence of scale invariant phenomena across the various fields I have studied and worked in suggests that this perspective could be a productive approach for future research. While the concept of scale invariance is common in theoretical physics, it is less widely recognised in geoscience and biology. My thesis suggests it would be useful for these fields to interface more regularly, as scale invariant phenomena often allow for the application of similar techniques, even if only in an analogous manner. Beyond the methods outlined in this thesis, such techniques could include statistical approaches to experimental data, such as data-driven dynamical schemes [31], or entirely theoretical frameworks, such as Koopman operator theory [17].

Appendix A

Feynman's Letter

Dear Koichi,

I was very happy to hear from you, and that you have such a position in the Research Laboratories. Unfortunately your letter made me unhappy for you seem to be truly sad. It seems that the influence of your teacher has been to give you a false idea of what are worthwhile problems. The worthwhile problems are the ones you can really solve or help solve, the ones you can really contribute something to. A problem is grand in science if it lies before us unsolved and we see some way for us to make some headway into it. I would advise you to take even simpler, or as you say, humbler, problems until you find some you can really solve easily, no matter how trivial. You will get the pleasure of success, and of helping your fellow man, even if it is only to answer a question in the mind of a colleague less able than you. You must not take away from yourself these pleasures because you have some erroneous idea of what is worthwhile.

You met me at the peak of my career when I seemed to you to be concerned with problems close to the gods. But at the same time I had another Ph.D. Student (Albert Hibbs) whose thesis was on how it is that the winds build up waves blowing over water in the sea. I accepted him as a student because he came to me with the problem he wanted to solve. With you I made a mistake, I gave you the problem instead of letting you find your own; and left you with a wrong idea of what is interesting or pleasant or important to work on (namely those problems you see you may do something about). I am sorry, excuse me. I hope by this letter to correct it a little.

I have worked on innumerable problems that you would call humble, but which I enjoyed and felt very good about because I sometimes could partially succeed. For example, experiments on the coefficient of friction on highly polished surfaces, to try to learn something about how friction worked (failure). Or, how elastic properties of crystals depend on the forces between the atoms in them, or how

to make electroplated metal stick to plastic objects (like radio knobs). Or, how neutrons diffuse out of Uranium. Or, the reflection of electromagnetic waves from films coating glass. The development of shock waves in explosions. The design of a neutron counter. Why some elements capture electrons from the L-orbits, but not the K-orbits. General theory of how to fold paper to make a certain type of child's toy (called flexagons). The energy levels in the light nuclei. The theory of turbulence (I have spent several years on it without success). Plus all the "grander" problems of quantum theory.

No problem is too small or too trivial if we can really do something about it.

You say you are a nameless man. You are not to your wife and to your child. You will not long remain so to your immediate colleagues if you can answer their simple questions when they come into your office. You are not nameless to me. Do not remain nameless to yourself – it is too sad a way to be. Know your place in the world and evaluate yourself fairly, not in terms of your naïve ideals of your own youth, nor in terms of what you erroneously imagine your teacher's ideals are.

Best of luck and happiness.

Sincerely,

Richard P. Feynman.

From [38].

DNS simulation parameters

Table B.1 $d = 5$

N	ν	$\bar{Re}_L$	$\bar{\epsilon}$
24	0.0306	14	0.0322
24	0.0307	14	0.0322
24	0.0308	14	0.0323
24	0.047	9	0.0371
24	0.048	8	0.0373
24	0.049	8	0.0375
24	0.051	8	0.0379
24	0.0595	6	0.0394
32	0.0305	14	0.0307
32	0.0306	14	0.0306
32	0.0307	14	0.0307
32	0.0308	14	0.0307
32	0.0309	14	0.0308
42	0.0205	21	0.0267
42	0.021	21	0.027
42	0.0215	20	0.0272
42	0.022	20	0.0275
42	0.0225	19	0.0278
48	0.0169	26	0.0238
48	0.017	26	0.0238
48	0.0176	25	0.0242
48	0.0182	24	0.0245
48	0.0188	23	0.025
64	0.0115	38	0.0205
64	0.012	37	0.0211
64	0.0125	36	0.0215
64	0.013	34	0.0219
72	0.0096	46	0.0188
72	0.01	44	0.0192
72	0.0104	43	0.0196
72	0.0108	41	0.02
84	0.008	55	0.0172
84	0.0085	52	0.0178
84	0.009	49	0.0184
96	0.0065	68	0.0152
96	0.0067	65	0.0156
96	0.0069	63	0.0159
96	0.0072	61	0.0163
128	0.0043	98	0.012

Table B.2 $d = 6$

N	ν	$\bar{Re}_L$	$\bar{\epsilon}$
24	0.047	8	0.0351
24	0.049	8	0.0355
24	0.051	7	0.0358
24	0.0595	6	0.0371
32	0.0305	13	0.0299
32	0.0306	13	0.0299
32	0.0307	13	0.03
32	0.0308	13	0.03
32	0.0309	13	0.0301
42	0.0205	21	0.0259
42	0.021	20	0.0262
42	0.0215	20	0.0265
42	0.022	19	0.0267
42	0.0225	19	0.027
48	0.0169	25	0.0232
48	0.017	25	0.0231
48	0.0176	24	0.0235
48	0.0182	23	0.0239
48	0.0188	22	0.0243

Table B.3 $d = 3$

N	ν	$\bar{Re}_L$	$\bar{\epsilon}$
24	0.047	8	0.039
24	0.048	8	0.0392
24	0.049	7	0.0394
24	0.051	7	0.0397
24	0.0595	6	0.0408
32	0.0305	12	0.0236
32	0.0306	12	0.0227
32	0.0307	12	0.0228
32	0.0308	12	0.0228
32	0.0309	12	0.0237
42	0.0205	26	0.0325
42	0.021	26	0.0326
42	0.0215	25	0.0327
42	0.022	24	0.0332
42	0.0225	24	0.0337
48	0.0169	28	0.0269
48	0.017	28	0.0259
48	0.0176	28	0.0273
48	0.0182	26	0.027
48	0.0188	26	0.0286
72	0.0096	51	0.0197
72	0.01	49	0.0196
72	0.0104	48	0.0203
72	0.0108	46	0.0205
96	0.0065	76	0.0176
96	0.0067	74	0.0192
96	0.0069	69	0.0185
96	0.0072	67	0.0193
128	0.0043	111	0.0154
128	0.0045	108	0.0159
128	0.005	98	0.0166
256	0.0017	285	0.0112
256	0.0018	274	0.0117
256	0.0019	260	0.0119
256	0.002	252	0.0124
512	0.0007	697	0.0052
512	0.0008	700	0.0072
512	0.0008	587	0.0062
512	0.047	7	0.0335

Table B.4 $d = 4$

N	ν	$\bar{Re}_L$	$\bar{\epsilon}$
24	0.047	10	0.059
24	0.048	9	0.0594
24	0.049	9	0.0597
24	0.051	9	0.0604
24	0.0595	7	0.0627
32	0.0305	15	0.0446
32	0.0306	15	0.0444
32	0.0307	15	0.0445
32	0.0308	15	0.0446
32	0.0309	15	0.0448
42	0.0205	24	0.0411
42	0.021	23	0.0415
42	0.0215	23	0.042
42	0.022	22	0.0424
42	0.0225	22	0.0428
48	0.0169	29	0.0353
48	0.017	29	0.0355
48	0.0176	28	0.036
48	0.0182	27	0.0366
48	0.0188	26	0.0371
64	0.0115	44	0.0329
64	0.012	42	0.0335
64	0.0125	40	0.0341
64	0.013	39	0.0349
72	0.0096	53	0.0304
72	0.01	51	0.0311
72	0.0104	49	0.0316
72	0.0108	47	0.0323
84	0.008	61	0.0255
84	0.0085	58	0.0263
84	0.009	54	0.027
84	0.0095	52	0.0279
96	0.0065	73	0.0211
96	0.0067	71	0.0217
96	0.0069	69	0.0222
96	0.0072	66	0.0225
128	0.0043	110	0.0183
128	0.0045	106	0.0185
128	0.047	10	0.0582
256	0.002	235	0.0117

List of Symbols

C_d Kolmogorov constant at dimension d.

$E(\ldots)$ expected value/mean of ....

L macroscopic integral scale.

P_g probability of microbial growth.

Re Reynolds number, usually at the integral scale.

Re_L integral scale Reynolds number.

Re_λ Taylor Reynolds number.

S total number of species.

$V(\ldots)$ variance of ....

ϵ energy dissipation rate.

η Kolmogorov microscale length.

λ Lyapunov exponent (chapter 2), population death rate (chapter 5).

$\mathcal{E}$ extinction rate.

$\mathcal{I}$ immigration rate.

μ population death rate.

τ Kolmogorov timescale.

k wavenumber (chapter 2), total population size (chapter 5).

Nota bene: This list is not exhaustive; featuring only key symbols and those which have dual uses.

List of Acronyms

ARCHER2 Advanced Research Computing High End Resource 2 (UK national supercomputing service), based on a HPE Cray EX supercomputer.

DNS direct numerical simulation.

EDQNM eddy damped quasi-normal Markovian.

ELEO extremely low Earth orbit.

ESC Einstein summation convention.

FFT fast Fourier transform.

FTLE finite-time Lyapunov exponent.

HIT homogeneous isotropic turbulence.

HPC high performance computing.

K41 Kolmogorov 1941 theory of turbulence.

K62 Kolmogorov 1962 theory of turbulence.

LEO low Earth orbit.

MLT mesosphere and lower thermosphere.

MPI message passing interface.

NAO North Atlantic Oscillation.

NSE Navier–Stokes equation.

OOM out of memory.

PDF probability density function.

RPC ratio of predictable components.

Glossary

2-form An antisymmetric rank-2 covariant tensor. It takes two vectors and returns a scalar, changing sign if you swap the vectors.

almost all All except a set of measure zero, from pure mathematics measure theory.

anomalous scaling Deviation from scaling laws in turbulence according to Kolmogorov theories.

attractor A set of trajectories in phase-space toward which a dynamical system evolves over time.

autocorrelation function A measure of how a signal correlates with itself over time lag. It quantifies the similarity between data points at different points in time.

back-contamination Biological contamination of Earth by potential extraterrestrial life forms or substances returned from space.

biogeography The study of the distribution of species and ecosystems in geographic space.

chaos Sensitive dependence on initial conditions in deterministic dynamical systems.

closure model A method to truncate an infinite sequence of equations by approximating higher-order terms.

colloidal A microscopic substance in suspension throughout another substance.

colonisability profile A set of organism traits and environmental conditions determining viability for microbial colonisation.

critical dimension Generally is the dimension at which the behaviour of a physical system changes but in critical phenomena, is usually the dimension at which a physical system when renormalized, is exactly describable by a mean-field theory.

crossover dimension A dimension at which the scaling behaviour of a system changes, e.g. from anomalous scaling to K41 theory (mean-field) predicted scaling in turbulence.

CubeSat A class of miniaturised satellites used for space research, typically in low Earth orbit.

diurnal A daily cycle.

divergence A vector operator measuring the magnitude of a source or sink of flow at a point. It is the resultant rate of stretching of a fluid given by the sum of the velocity gradients in all dimensions.

divergence free A field with zero divergence, used in this thesis to indicate fluid incompressibility.

eddy A localised and coherent swirling motion in a fluid.

ensemble forecasting A method using multiple weather forecast model simulations to estimate uncertainty in predictions.

enstrophy A measure of the intensity of vorticity in fluid flow.

ergodicity A property where long-term time averages of a system equal averages over its phase-space ensemble.

exascale Computational systems capable of performing at least 10^{18} operations per second, also associated with exabyte systems which have an operational memory equivalent to one quintillion bytes or a thousand petabytes.

extinction time The expected time until a population reaches zero.

fixed line A continuum of fixed points in the renormalization group flow.

Fourier transform A mathematical transformation from time or space to frequency domain. In this thesis typically from physical space (real-space) to frequency space (k-space).

Gedankenexperiment A thought experiment used to explore the consequences of a theory.

habitability The potential of an environment to support life.

homogeneous Translationally invariant, i.e. the same everywhere. Turbulence is said to be homogeneous if it statistically the same under spatial translation.

inertial range A range of scales in turbulent flow where energy cascades without (significant) dissipation.

integrable function A function for which the integral over its domain finitely exists.

integral scale The largest scale of turbulent motion in a flow, i.e. the biggest swirls.

intermittency This refers to the observed phenomena that as turbulence increases (increased Reynolds number), the bulk of energy dissipation occupies an ever smaller proportion of the total flow volume such that it becomes more spatially intermittent. It is also known as "internal intermittency" .

inviscid A theoretically ideal fluid which has no viscosity.

island biogeography theory A theory describing the number and distribution of species on an island as set out by [76].

isomorphism A mapping between two mathematical objects which can be exactly reversed due to the fact that the two objects have the same structure.

isotropic Invariant under rotation i.e. the same in all directions .

Kolmogorov microscale The smallest scale in turbulence at which viscous dissipation occurs.

Kolmogorov spectrum An energy spectrum characterising the inertial range in isotropic turbulence, i.e. $E(k) \sim k^{-5/3}$.

Lyapunov exponent A measure of the average rate of divergence or convergence of nearby trajectories in the phase-space of a dynamical system, hence a measure of chaos. The higher the Lyapunov exponent, the more chaotic the system. A negative Lyapunov exponent indicates a truly non-chaotic system.

Lyapunov spectrum The full set of Lyapunov exponents characterising the rates of separation along different directions in phase-space.

mean-time to extinction The expected time for a population to go extinct under given birth and death rates.

measure A mathematical function that assigns a non-negative size to subsets of a space. It is a mathematicians generalisation of the physical measures of length, area etc, and in this thesis it usually refers to a *probability measure* which returns values in the range $[0, 1]$ such that for the whole space it assigns a value of 1..

negative damping scheme A direct numerical simulation forcing scheme where energy is added to a system by injecting energy to the lowest modes in the velocity field at a set rate.

passive scalar A quantity advected by the flow without affecting it e.g. red dye transported in a flow.

periodic boundary conditions Boundary conditions where opposite edges of the domain are connected directly so that when an object reaches the edge of the domain, it transfers to the opposite side with the same state.

phase-space A space in which all possible states of a system are represented, also known as state space.

planetary protection Measures to prevent biological contamination of other planets and Earth.

pseudo-spectral A numerical method, which like the usual spectral method, uses Fourier transforms for spatial derivatives so that these can be more easily computed as products in k-space, except for the non-linear terms which are computed in real-space. Thus it combines the high accuracy of spectral methods with the efficiency of the finite-difference method.

relaminarisation The transition of turbulent flow back to laminar state.

renormalization A mathematical technique of rescaling a system to understand its behaviour at different length or energy scales.

Ruelle relation A theoretical result linking the chaos of a fluid via the Lyapunov exponent λ and how turbulent a fluid is via its Reynolds number Re, $\lambda \sim \frac{1}{T_0} Re^{\alpha}$ where T_0 is the integral timescale and α is some constant. For K41, $\alpha = \frac{1}{2}$.

scale invariance A property of systems that do not change under rescaling.

signal-to-noise paradox A phenomenon where ensemble weather forecasts are better at predicting reality than themselves (their ensemble mean).

solenoidal subspace A subspace of vector fields with zero divergence.

sounding rocket A rocket designed to take measurements and perform scientific experiments during sub-orbital flight.

state space The set of all possible states of a dynamical system. Also known as phase-space.

structure function A statistical measure of variability between points in a turbulent field.

survivability Whether an organism lives or dies in a given environment.

Taylor microscale An intermediate length scale in a fluid that characterises the scale at which viscous dissipation begins to dominate velocity fluctuations. It is larger than the Kolmogorov scale but smaller than the integral scale.

Taylor Reynolds number A Reynolds number defined by a characteristic length of the Taylor microscale.

terraforming The process of modifying a planet's environment to make it habitable for Earth life.

Thirumalai–Mountain metric A metric used to test the ergodicity of a system based on convergence of time-averaged observables.

transitivity In dynamical systems, the property that any state can evolve to any other, given enough time.

universality class A group of systems that share the same critical exponents and scaling behaviour near a phase transition.

velocity derivative skewness A measure of asymmetry in the probability distribution of velocity gradients.

vorticity A vector field representing the local spinning motion of a fluid.

Bibliography

[1] "COSPAR Policy on Planetary Protection.", 2024. URL `https://cosparhq.cnes.fr/assets/uploads/2024/07/PP-Policy_SRT_220-July-2024.pdf`. COSPAR Panel on Planetary Protection.

[2] Aji, Vivek, and Nigel Goldenfeld. "Fluctuations in Finite Critical and Turbulent Systems." *Physical Review Letters* 86 (2001): 1007–1010.

[3] Andrade, Tomas, Christiana Pantelidou, Julian Sonner, and Benjamin Withers. "Driven black holes: from Kolmogorov scaling to turbulent wakes." *Journal of High Energy Physics* 2021, 7 (2021): 63.

[4] Anselmet, F., Y. Gagne, E. J. Hopfinger, and R. A. Antonia. "High-order velocity structure functions in turbulent shear flows." *Journal of Fluid Mechanics* 140 (1984): 63–89.

[5] Armúa, Andrés. *Chaos, statistics and inverse cascades in turbulent flows.* Ph.D. thesis, The University of Edinburgh, 2023.

[6] Arnold, Vladimir I., and Boris A. Khesin. *Topological Methods in Hydrodynamics.* Springer International Publishing, 2021.

[7] Ash, R. D., S. F. Knott, and G. Turner. "A 4-Gyr shock age for a Martian meteorite and implications for the cratering history of Mars." *Nature* 380 (1996): 57–59.

[8] Atta, C. W. Van, and W. Y. Chen. "Structure functions of turbulence in the atmospheric boundary layer over the ocean." *Journal of Fluid Mechanics* 44, 1 (1970): 145–159.

[9] Bell, Thomas L., and Mark Nelkin. "Time-dependent scaling relations and a cascade model of turbulence." *Journal of Fluid Mechanics* 88, 2 (1978): 369–391.

[10] Bender, Carl M., and Stefan Boettcher. "Dimensional expansion for the Ising limit of quantum field theory." *Physical Review D* 48 (1993): 4919–4923.

[11] Bender, Carl M., Fred Cooper, G. S. Guralnik, and David H. Sharp. "Strong-coupling expansion in quantum field theory." *Physical Review D* 19 (1979): 1865–1881.

[12] Berera, A, R D J G Ho, and D Clark. "Homogeneous isotropic turbulence in four spatial dimensions." *Physics of Fluids* 32 (2020): 085,107. Doi: 10.1063/5.0022929.

[13] Berera, Arjun, and Richard D. J. G. Ho. "Chaotic Properties of a Turbulent Isotropic Fluid." *Physical Review Letters* 120, 2 (2018): 24,101.

[14] Boffetta, G., and S. Musacchio. "Chaos and Predictability of Homogeneous-Isotropic Turbulence." *Physical Review Letters* 119 (2017): 054,102.

[15] Boureima, Ismael, Manish Bhattarai, Maksim Eren, Erik Skau, Philip Romero, Stephan Eidenbenz, and Boian Alexandrov. "Distributed out-of-memory NMF on CPU/GPU architectures." *The Journal of Supercomputing* 80, 3 (2024): 3970–3999.

[16] Bramwell, S. T., P. C. W. Holdsworth, and J. F. Pinton. "Universality of rare fluctuations in turbulence and critical phenomena." *Nature* 396, 6711 (1998): 552–554.

[17] Brunton, Steven L., Marko Budišić, Eurika Kaiser, and J. Nathan Kutz. "Modern Koopman Theory for Dynamical Systems." *SIAM Review* 64, 2 (2022): 229–340.

[18] Canet, Léonie, Bertrand Delamotte, and Nicolás Wschebor. "Fully developed isotropic turbulence: Nonperturbative renormalization group formalism and fixed-point solution." *Physical Review E* 93, 6 (2016): 063,101.

[19] Charlton-Perez, Andrew, Jochen Bröcker, Timothy N. Stockdale, and Stephanie J. Johnson. "When and where do ECMWF seasonal forecast systems exhibit anomalously low signal-to-noise ratio?" *Quarterly Journal of the Royal Meteorological Society* 145, 725 (2019): 3466–3478.

[20] Clark, Daniel. *Effect of spatial dimensionality on the chaotic properties of turbulent flow.* Ph.D. thesis, The University of Edinburgh, 2022.

[21] Clark, Daniel, Andres Armua, Richard D.J.G. Ho, and Arjun Berera. "Critical transition to a non-chaotic regime in isotropic turbulence." *Journal of Fluid Mechanics* 930 (2022): A17.

[22] Clark, Daniel, Richard D.J.G. Ho, and Arjun Berera. "Effect of spatial dimension on a model of fluid turbulence." *Journal of Fluid Mechanics* 912 (2021): A40.

[23] Clarke, Andrew, G. John Morris, Fernanda Fonseca, Benjamin J. Murray, Elizabeth Acton, and Hannah C. Price. "A Low Temperature Limit for Life on Earth." *PLOS ONE* 8, 6 (2013): 1–11.

[24] Cockell, Charles S., André Brack, David D. Wynn-Williams, Pietro Baglioni, Franz Brandstätter, René Demets, Howell G.M. Edwards,

Aaron L. Gronstal, Gero Kurat, Pascal Lee, Gordon R. Osinski, David A. Pearce, Judith M. Pillinger, Claude-Alain Roten, and Suzy Sancisi-Frey. "Interplanetary Transfer of Photosynthesis: An Experimental Demonstration of A Selective Dispersal Filter in Planetary Island Biogeography." *Astrobiology* 7, 1 (2007): 1–9.

[25] Cohen, B. A., T. D. Swindle, and D. A. Kring. "Support for the lunar cataclysm hypothesis from lunar meteorite impact melt ages." *Science* 290 (2000): 1754–1756.

[26] Coustenis, Athena, Nathan Hedman, Peter T. Doran, Omar Al Shehhi, Eleonora Ammannito, Masaki Fujimoto, Olivier Grasset, Frank Groen, Alexander G. Hayes, Vitaly Ilyin, Krishnaswamy P. Kumar, Catherine-Evans Morisset, Craig Mustin, Karen Olsson-Francis, Jing Peng, Oscar Prieto-Ballesteros, François Raulin, Petra Rettberg, Sarah Sinibaldi, Yasushi Suzuki, Ke Xu, and Michael Zaitsev. "Planetary protection: an international concern and responsibility." *Frontiers in Astronomy and Space Sciences* 10 (2023): 1172,546.

[27] Crisanti, A., M. Jensen, A. Vulpiani, and G. Paladin. "Intermittency and predictability in turbulence." *Physical Review Letters* 70, 2 (1993): 166–169.

[28] Dalrymple, G. B., and G. Ryder. "40Ar/39Ar age spectra of Apollo 15 impact melt rocks by laser step-heating and their bearing on the history of lunar basin formation." *Journal of Geophysical Research* 98 (1993): 13,085–13,095.

[29] DasSarma, Priya, André Antunes, Marta Filipa Simões, and Shiladitya DasSarma. "Earth's Stratosphere and Microbial Life." *Current Issues in Molecular Biology* 38, 1 (2020): 197–244.

[30] DeDominicis, C, and P C Martin. "Energy spectra of certain randomly-stirred fluids." *Physical Review A* 19 (1979): 419.

[31] Dong, Chenyu, Davide Faranda, Adriano Gualandi, Valerio Lucarini, and Gianmarco Mengaldo. "Revisiting the predictability of dynamical systems: a new local data-driven approach.", 2024. URL `https://arxiv.org/abs/2409.14865`.

[32] Dunstone, Nick, Doug M. Smith, Steven C. Hardiman, Leon Hermanson, Sarah Ineson, Gillian Kay, Chaofan Li, Julia F. Lockwood, Adam A. Scaife, Hazel Thornton, Mingfang Ting, and Lei Wang. "Skilful predictions of the Summer North Atlantic Oscillation." *Communications Earth & Environment* 4, 1 (2023): 409.

[33] Eade, Rosie, Doug Smith, Adam A. Scaife, Emily Wallace, Nick Dunstone, Leon Hermanson, and Niall Robinson. "Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?" *Geophysical research letters* 41, 15 (2014): 5620–5628.

[34] Eckmann, J. P., and D. Ruelle. "Ergodic theory of chaos and strange attractors." *Rev. Mod. Phys.* 57 (1985): 617–656.

[35] Epstein, Edward S. "Stochastic dynamic prediction." *Tellus* 21, 6 (1969): 739–759.

[36] Faranda, D., G. Messori, T. Alberti, C. Alvarez-Castro, T. Caby, L. Cavicchia, E. Coppola, R. V. Donner, B. Dubrulle, V. M. Galfi, E. Holmberg, V. Lembo, R. Noyelle, P. Yiou, B. Spagnolo, D. Valenti, S. Vaienti, and C. Wormell. "Statistical physics and dynamical systems perspectives on geophysical extreme events." *Physical Review E* 110 (2024): 041,001.

[37] Feynman, Richard P. *The Meaning of It All: Thoughts of a Citizen-Scientist.* Basic Books, 2005.

[38] Feynman, R.P., M. Feynman, and T. Ferris. *Perfectly Reasonable Deviations from the Beaten Track: The Letters of Richard P. Feynman.* Basic Books, 2008.

[39] Forster, D, D R Nelson, and M J Stephen. "Long-time tails and the large-eddy behavior of a randomly stirred fluid." *Physical Review Letters* 36 (1976): 867.

[40] ———. "Large-distance and long-time properties of a randomly stirred fluid." *Physical Review A* 16 (1977): 732.

[41] Fournier, J-D, U Frisch, and HA Rose. "Infinite-dimensional turbulence." *Journal of Physics A: Mathematical and General* 11, 1 (1978): 187.

[42] Frisch, Uriel. *Turbulence: The Legacy of A. N. Kolmogorov.* Cambridge University Press, 1995.

[43] Frisinger, H. Howard. "Aristotle and his "Meteorologica"." *Bulletin of the American Meteorological Society* 53, 7 (1972): 634 – 638.

[44] Ghil, Michael. "A Century of Nonlinearity in the Geosciences." *Earth and Space Science* 6, 7 (2019): 1007–1042.

[45] Goldenfeld, Nigel. *Lectures on phase transitions and the renormalization group.* CRC Press, 2019.

[46] Gotoh, T, Y Watanabe, Y Shiga, T Nakano, and E Suzuki. "Statistical properties of four-dimensional turbulence." *Physical Review E* 75 (2007): 016,310.

[47] Grant, H. L., R. W. Stewart, and A. Moilliet. "Turbulence spectra from a tidal channel." *Journal of Fluid Mechanics* 12, 2 (1962): 241–268.

[48] Griffin, Dale. "Terrestrial microorganisms at 20,000 meters in Earth's atmosphere." *Aerobiologia* 20 (2004): 135–140.

[49] Heintzenberg, J., M. Hermann, and D. Theiss. "Out of Africa: High aerosol concentrations in the upper troposphere over Africa." *Atmospheric Chemistry and Physics* 3, 4 (2003): 1191–1198.

[50] Henkel, Malte. "From dynamical scaling to local scale-invariance: a tutorial." *The European Physical Journal Special Topics* 226, 4 (2017): 605–625.

[51] Ho, Richard D. J. G. *Effects of macroscopic variables on turbulent evolution.* Ph.D. thesis, The University of Edinburgh, 2019.

[52] Ho, Richard D. J. G., Andres Armua, and Arjun Berera. "Fluctuations of Lyapunov exponents in homogeneous and isotropic turbulence." *Physical Review Fluids* 5 (2020): 024,602.

[53] Ho, Richard D. J. G., Daniel Clark, and Arjun Berera. "Chaotic Measures as an Alternative to Spectral Measures for Analysing Turbulent Flow." *Atmosphere* 15, 9 (2024): 1053.

[54] Imshenetsky, A A, S V Lysenko, and G A Kazakov. "Upper boundary of the biosphere." *Applied and environmental microbiology* 35, 1 (1978): 1–5.

[55] Kawaguchi, Yuko, Mio Shibuya, Iori Kinoshita, Jun Yatabe, Issay Narumi, Hiromi Shibata, Risako Hayashi, Daisuke Fujiwara, Yuka Murano, Hirofumi Hashimoto, Eiichi Imai, Satoshi Kodaira, Yukio Uchihori, Kazumichi Nakagawa, Hajime Mita, Shin-ichi Yokobori, and Akihiko Yamagishi. "DNA Damage and Survival Time Course of Deinococcal Cell Pellets During 3 Years of Exposure to Outer Space." *Frontiers in Microbiology* 11 (2020): 2050.

[56] Kolmogorov, A. "The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers." *Proceedings of the USSR Academy of Sciences* 30 (1941): 301–305.

[57] Kolmogorov, A. N. "A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number." *Journal of Fluid Mechanics* 13, 1 (1962): 82–85.

[58] Kraichnan, R H. "The structure of isotropic turbulence at very high Reynolds numbers." *Journal of Fluid Mechanics.* 5 (1959): 497.

[59] ———. "Convection of a passive scalar by a quasi-uniform random straining field." *Journal of Fluid Mechanics.* 64 (1974): 737.

[60] ———. "Intermittency and attractor size in isotropic turbulence." *Physics of Fluids* 28 (1985): 10.

[61] Kring, D. A., and B. A. Cohen. "Cataclysmic bombardment throughout the inner solar system 3.9–4.0 Ga." *Journal of Geophysical Research* 107.

[62] Krisko, Anita, and Miroslav Radman. "Biology of extreme radiation resistance: the way of Deinococcus radiodurans." *Cold Spring Harbor perspectives in biology* 5, 7 (2013): a012 765.

[63] Kumar, Arun. "Finite samples and uncertainty estimates for skill measures for seasonal prediction." *Monthly Weather Review* 137, 8 (2009): 2622–2631.

[64] Landau, Lev Davidovich, and Evgeny Mikhailovich Lifshitz. *Fluid mechanics*. Pergamon, 1959.

[65] Lapeyre, G. "Characterization of finite-time Lyapunov exponents and vectors in two-dimensional turbulence." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 12, 3 (2002): 688–698.

[66] Leutbecher, Martin, Sarah-Jane Lock, Pirkka Ollinaho, Simon T. K. Lang, Gianpaolo Balsamo, Peter Bechtold, Massimo Bonavita, Hannah M. Christensen, Michail Diamantakis, Emanuel Dutra, Stephen English, Michael Fisher, Richard M. Forbes, Jacqueline Goddard, Thomas Haiden, Robin J. Hogan, Stephan Juricke, Heather Lawrence, Dave MacLeod, Linus Magnusson, Sylvie Malardel, Sebastien Massart, Irina Sandu, Piotr K. Smolarkiewicz, Aneesh Subramanian, Frédéric Vitart, Nils Wedi, and Antje Weisheimer. "Stochastic representations of model uncertainties at ECMWF: state of the art and future vision." *Quarterly Journal of the Royal Meteorological Society* 143, 707 (2017): 2315–2339.

[67] Li, Ming, Run-Ran Liu, Linyuan Lü, Mao-Bin Hu, Shuqi Xu, and Yi-Cheng Zhang. "Percolation on complex networks: Theory and application." *Physics Reports* 907 (2021): 1–68.

[68] Li, Sheng, Ke Chen, Mingyu Hsieh, Naveen Muralimanohar, Chad Kersey, Jay Brockman, Arun Rodrigues, and Norman Jouppi. "System implications of memory reliability in exascale computing." In *Proceedings of 2011 SC - International Conference for High Performance Computing, Networking, Storage and Analysis*. 2011, 46.

[69] Liao, W. "Some ideas on the freely decaying Navier-Stokes turbulence." *Journal of Physics A: Mathematical and General* 23, 4 (1990): L159.

[70] ———. "Kolmogorov exponents for near-incompressible turbulence from perturbative quantum field theory." *Journal of Statistical Physics* 65 (1991): 1.

[71] Linkmann, Moritz F., and Alexander Morozov. "Sudden Relaminarization and Lifetimes in Forced Isotropic Turbulence." *Physical Review Letters* 115 (2015): 134,502.

[72] Lorenz, Edward N. *Climatic Determinism*, Boston, MA: American Meteorological Society, 1968, 1–3.

[73] ———. "Nondeterministic theories of climatic change." *Quaternary Research* 6, 4 (1976): 495–506.

[74] ———. "Can chaos and intransitivity lead to interannual variability?" *Tellus A* 42, 3 (1990): 378–389.

[75] Lundgren, Thomas S. "Kolmogorov turbulence by matched asymptotic expansions." *Physics of Fluids* 15, 4 (2003): 1074–1081.

[76] MacArthur, R.H., and E.O. Wilson. *The Theory of Island Biogeography.* Monographs in population biology. Princeton University Press, 1967.

[77] Marchi, S., N. Drabon, T. Schulz, L. Schaefer, D. Nesvorny, W. F. Bottke, C. Koeberl, and T. Lyons. "Delayed and variable late Archaean atmospheric oxidation due to high collision rates on Earth." *Nature Geoscience* 14, 11 (2021): 827–831.

[78] McComb, W D, A Berera, S R Yoffe, and M F Linkmann. "Energy transfer and dissipation in forced isotropic turbulence." *Physical Review E* 91 (2015): 043,013.

[79] McComb, W. David. *Homogeneous, isotropic turbulence: phenomenology, renormalization and statistical closures.* International series of monographs on physics; 162. Oxford University Press, 2014.

[80] McCoy, Kelli J., Michael DiNicola, Chester Everline, Hayden Burgoyne, Kirk Reinholtz, and Brian Clement. "Europa Clipper planetary protection probabilistic risk assessment summary." *Planetary and Space Science* 196 (2021): 105,139.

[81] Mileikowsky, Curt, Francis A. Cucinotta, John W. Wilson, Brett Gladman, Gerda Horneck, Lennart Lindegren, Jay Melosh, Hans Rickman, Mauri Valtonen, and J.Q. Zheng. "Natural Transfer of Viable Microbes in Space: 1. From Mars to Earth and Earth to Mars." *Icarus* 145, 2 (2000): 391–427.

[82] Milne, A. A. *Winnie-the-Pooh.* Methuen, 1926.

[83] Moissl-Eichinger, Christine, Charles Cockell, and Petra Rettberg. "Venturing into new realms? Microorganisms in space." *FEMS Microbiology Reviews* 40, 5 (2016): 722–737.

[84] Moshe, Moshe, and Jean Zinn-Justin. "Quantum field theory in the large N limit: A Review." *Physics Reports* 385, 3-6 (2003): 69–228.

[85] Nelkin, Mark. "Turbulence, critical fluctuations, and intermittency." *Physical Review A* 9 (1974): 388–395.

[86] ———. "Scaling theory of hydrodynamic turbulence." *Physical Review A* 11 (1975): 1737.

[87] ———. "Does Kolmogorov mean field theory become exact for turbulence above some critical dimension?", 2001. URL https://arxiv.org/abs/nlin/0103046.

[88] Nikitin, N. "Four-dimensional turbulence in a plane channel." *Journal of Fluid Mechanics.* 680 (2011): 67.

[89] Ontiveros, Vicente J., Jose A. Capitan, Rohan Arthur, Emilio O. Casamayor, and David Alonso. "Colonization and Extinction Rates estimated from Temporal Dynamics of Ecological Communities: The island R Package." *Methods in Ecology and Evolution* 10, 7 (2019): 1108–1117.

[90] Ornstein, Donald. "Bernoulli shifts with the same entropy are isomorphic." *Advances in Mathematics* 4, 3 (1970): 337–352.

[91] Oseledets, Valery Iustinovich. "A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems." *Trudy Moskovskogo Matematicheskogo Obshchestva* 19 (1968): 179–210.

[92] Palmer, Tim, and Renate Hagedorn. *Predictability of Weather and Climate.* Cambridge University Press, 2006.

[93] Pasteur, L. "Mémoire sur les corpuscules organisés qui existent dans l'atmosphère, examen de la doctrine des générations spontanées." *Comptes Rendus de l'Académie des Sciences* 2, 8 (1861): 1142–1143.

[94] Pearce, David A., Paul D. Bridge, Kevin A. Hughes, Birgit Sattler, Roland Psenner, and Nick J. Russell. "Microorganisms in the atmosphere over Antarctica." *FEMS Microbiology Ecology* 69, 2 (2009): 143–157.

[95] Peravali, Surya Kiran, Vahid Jafari, Amit K. Samanta, Jochen Küpper, Muhamed Amin, Philipp Neumann, and Michael Breuer. "Accuracy and performance evaluation of low density internal and external flow predictions using CFD and DSMC." *Computers & Fluids* 279 (2024): 106,346.

[96] Pippig, Michael. "PFFT: An Extension of FFTW to Massively Parallel Architectures." *SIAM Journal on Scientific Computing* 35, 3 (2013): C213–C236.

[97] Plato. *Euthyphro apology crito Phaedo Phaedrus*, volume 1. Harvard University Press, 1966. Translator: Harold North Fowler.

[98] Qian, J. "Inertial range and the finite Reynolds number effect of turbulence." *Physical Review E* 55 (1997): 337–342.

[99] ———. "Slow decay of the finite Reynolds number effect of turbulence." *Physical Review E* 60 (1999): 3409–3412.

[100] Richard, L., Weiss Bizzoco, and Scott T. Kelley. *Volcanic Steam Vents: Life at Low pH and High Temperature*, John Wiley & Sons, Ltd, 2020, chapter 1, 1–20.

[101] Richardson, Lewis F. *Weather prediction by numerical process.* Cambridge University Press, 1922.

[102] Rosen, James M. "The vertical distribution of dust to 30 kilometers." *Journal of Geophysical Research (1896-1977)* 69, 21 (1964): 4673–4676.

[103] Rowling, J.K. *Harry Potter and the Prisoner of Azkaban.* Harry Potter. Bloomsbury, 1999.

[104] Ruelle, David. "Microscopic fluctuations and turbulence." *Physics Letters A* 72, 2 (1979): 81–82.

[105] ———. *Chaotic Evolution and Strange Attractors.* Lezioni Lincee. Cambridge University Press, 1989.

[106] Saeed, Nasir, Ahmed Elzanaty, Heba Almorad, Hayssam Dahrouj, Tareq Y. Al-Naffouri, and Mohamed-Slim Alouini. "CubeSat Communications: Recent Advances and Future Challenges." *IEEE Communications Surveys Tutorials* 22, 3 (2020): 1839–1862.

[107] Scaife, Adam A., Alberto Arribas, Edward W. Blockley, Anca Brookshaw, Robin T. Clark, Nick Dunstone, Rosie Eade, D. Fereday, Chris K. Folland, Margaret Gordon, Leon Hermanson, Jeff Knight, D. J. Lea, Craig MacLachlan, A. Maidens, Matthew Martin, A. Peterson, Doug Smith, Michael Vellinga, Emily Wallace, Jennifer Waters, and Amy J. Williams. "Skillful long-range prediction of European and North American winters." *Geophysical Research Letters* 41, 7 (2014): 2514–2519.

[108] Scaife, Adam A, and Doug Smith. "A signal-to-noise paradox in climate science." *npj Climate and Atmospheric Science* 1, 1 (2018): 28.

[109] Sharma, A. Surjalal, Daniel N. Baker, Archana Bhattacharyya, Armin Bunde, Vijay P. Dimri, Harsh K. Gupta, Vijay K. Gupta, Shaun Lovejoy, Ian G. Main, Daniel Schertzer, Hans von Storch, and Nicholas W. Watkins. *Complexity and Extreme Events in Geosciences: An Overview*, American Geophysical Union (AGU), 2012, 1–16.

[110] Siggia, Eric D. "Origin of intermittency in fully developed turbulence." *Physical Review A* 15 (1977): 1730–1750.

[111] Sreenivasan, K. R., and R. A. Antonia. "The phenomenology of small-scale turbulence." *Annual Review of Fluid Mechanics* 29, 1 (1997): 435–472.

[112] Stern, Alan. *Next-Generation Suborbital Researchers Conference: February 18-20, 2010, Boulder, Colorado.* Lunar and Planetary Institute, 2010.

[113] Stockdale, Timothy N., Franco Molteni, and Laura Ferranti. "Atmospheric initial conditions and the predictability of the Arctic Oscillation." *Geophysical Research Letters* 42, 4 (2015): 1173–1179.

[114] Stöckli, Reto, Eric Vermote, Nazmi Saleous, Robert Simmon, and David Herring. "The Blue Marble Next Generation: A true color earth dataset including seasonal dynamics from MODIS.", 2005. URL `https://neo.gsfc.nasa.gov/archive/bluemarble/`.

[115] Suess, Eduard. *Die Entstehung der Alpen.* Braumüller, 1875.

[116] Suzuki, E, T Nakano, N Takahashi, and T Gotoh. "Energy transfer and intermittency in four-dimensional turbulence." *Physics of Fluids* 17 (2005): 081,702.

[117] Tenenbaum, Joel. *Interdisciplinary applications of statistical physics to complex systems: seismic physics, econophysics, and sociophysics.* Ph.D. thesis, Boston University, 2012.

[118] Tera, F., D. A. Papanastassiou, and G. J. Wasserburg. "Isotopic evidence for a terminal lunar cataclysm." *Earth and Planetary Science Letters* 22 (1974): 1–21.

[119] Turner, G., P. H. Cadogan, and C. J. Yonge. "Argon selenochronology." In *Proceedings of the Fourth Lunar Science Conference.* Lunar and Planetary Institute, 1973, volume 2, 1889–1914.

[120] Ursem, Bob. "Climate Shifts and the Role of Nano Structured Particles in the Atmosphere." *Atmospheric and Climate Sciences* 06 (2016): 51–76.

[121] Vallis, Geoffrey K. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation.* Cambridge University Press, 2006.

[122] Van Dyke, Milton. *Perturbation Methods in Fluid Mechanics.* Parabolic Press, 1975.

[123] Vermaseren, J. A. M. "New features of FORM.", 2000. URL `https://arxiv.org/abs/math-ph/0010025`.

[124] Vernadsky, Vladimir I. "The Biosphere: An Envelope of the Earth." In *The Biosphere*, 91–102. Springer New York, 1998.

[125] Virgili-Llop, Josep, Peter Roberts, Zhou Hao, Laia Ramio, and Valentin Beauplet. "Very Low Earth Orbit mission concepts for Earth Observation. Benefits and challenges." In *Reinventing Space Conference.* Royal Society, 2014.

[126] Wainwright, M, N C Wickramasinghe, J V Narlikar, and P Rajaratnam. "Microorganisms cultured from stratospheric air samples obtained at 41 km." *FEMS microbiology letters* 218, 1 (2003): 161–165.

[127] Walker, GT, and EW Bliss. "World weather V." *Memoirs of the Royal Meteorological Society* 4 (1932): 53–84.

[128] Weisheimer, Antje, Laura H. Baker, Jochen Bröcker, Chaim I. Garfinkel, Steven C. Hardiman, Dan L. R. Hodson, Tim N. Palmer, Jon I. Robson, Adam A. Scaife, James A. Screen, Theodore G. Shepherd, Doug M. Smith, and Rowan T. Sutton. "The Signal-to-Noise Paradox in Climate Forecasts: Revisiting Our Understanding and Identifying Future Priorities." *Bulletin of the American Meteorological Society* 105, 3 (2024): E651 – E659.

[129] Weisheimer, Antje, Damien Decremer, David MacLeod, Christopher H. O'Reilly, Timothy N. Stockdale, Stephanie J. Johnson, and Tim Palmer. "How confident are predictability estimates of the winter North Atlantic Oscillation." *Quarterly Journal of the Royal Meteorological Society* 145, S1 (2019): 140–159.

[130] Wilson, Kenneth G., and Michael E. Fisher. "Critical Exponents in 3.99 Dimensions." *Physical Review Letters* 28 (1972): 240–243.

[131] Wodehouse, P.G. *Much Obliged, Jeeves.* Jeeves. Barrie & Jenkins, 1971.

[132] Worth, R.J., Steinn Sigurdsson, and Christopher H. House. "Seeding Life on the Moons of the Outer Planets via Lithopanspermia." *Astrobiology* 13, 12 (2013): 1155–1165.

[133] Wu, Zhenhua, Cecilia Lagorio, Lidia A. Braunstein, Reuven Cohen, Shlomo Havlin, and H. Eugene Stanley. "Numerical evaluation of the upper critical dimension of percolation in scale-free networks." *Phys. Rev. E* 75 (2007): 066,110.

[134] Xu, Li, Guangyu Shi, Li Zhang, Jun Zhou, and Yasunobu Iwasaka. "Number concentration, size distribution and fine particle fraction of tropospheric and stratospheric aerosols." *China Particuology* 1, 5 (2003): 201–205.

[135] Yamamoto, T, H Shimizu, T Inoshita, T Nakano, and T Gotoh. "Local flow structure of turbulence in three, four, and five dimensions." *Physical Review E* 86 (2012): 046,320.

[136] Yin, Yan, Qian Chen, Lianji Jin, Baojun Chen, Shichao Zhu, and Xiaopei Zhang. "The effects of deep convection on the concentration and size distribution of aerosol particles within the upper troposphere: A case study." *Journal of Geophysical Research: Atmospheres* 117 (2012): D22 202.

[137] Yoden, Shigeo, and Masako Nomura. "Finite-Time Lyapunov Stability Analysis and Its Application to Atmospheric Predictability." *Journal of Atmospheric Sciences* 50, 11 (1993): 1531 – 1543.

[138] Yoffe, Samuel R. *Investigation of the transfer and dissipation of energy in isotropic turbulence.* Ph.D. thesis, The University of Edinburgh, 2013.