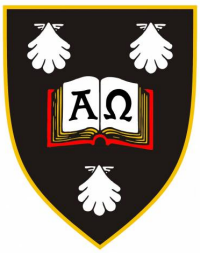
On the variability and forced response of atmospheric regime systems

Joshua C.F. Dorrington

Linacre College



University of Oxford

A thesis presented for the degree of

Doctor of Philosophy

Michelmas 2021

I dedicate this thesis to Olivia, my fiancée and best friend, who has supported me unwaveringly through difficult times, and makes the good times shine brighter.

Acknowledgements

When I started my DPhil in September 2017, I had no formal training in atmospheric physics or fluid dynamics and in those early months I had frequent trouble remembering the difference between my barotropic and baroclinic flows. If we are indeed the sum of everyone we have ever spoken to, then this thesis is the summed contribution of everyone who has chatted, educated, challenged, and collaborated with me over the past four years. Without their support this work could never have been written, and I would like to acknowledge them now.

I would like to start by thanking my supervisors, Tim Palmer and Antje Weisheimer, for their advice and guidance throughout the length of my DPhil and for their insightful comments that have significantly strengthened the work in this thesis. While keeping me on the right path they have given me space to develop my own ideas and so made me a more confident and independent researcher. I also owe considerable thanks to Kristian Strommen, whose idea it was that regimes might be obscured by jet speed and with whom I've collaborated closely on developing the geopotential-jet regime framework. Kristian's support and advice on academic and career issues has been invaluable. I would also like to thank Federico Fabiano with whom I am collaborating on the analysis of CMIP regimes, and who has shared his insight on identifying regimes in a changing climate. While the work has not ultimately made it into this thesis, I owe thanks to Isla Finney and Lake Street Consulting Ltd for hosting me as a summer intern in 2019, from which I learned a great deal about operational forecasting and which resulted in my first academic publication.

Fenwick Cooper was a font of good advice at the beginning of my research, and his repeated refusals to accept any definition of a regime system I suggested was very formative and encouraged me to think critically. I am greatly appreciative of Glen Shutts for taking the time to discuss the physical interpretation of the CdV system with me, and of

Frank Kwasniok and Dan Crommelin, with whom I discussed the mathematical aspects. Hannah Christensen made the YoTC data available to me, which was a substantial help in analysing the role of stochastic parameterisations in the SPHINX project. Adam Paxton introduced me to the Wasserstein distance and happily answered all my questions about its use and interpretation. Valerio Lucarini has also helped me understand the current state of the art in dynamical systems approaches to understanding blocking and regimes, which has helped broaden my perspective. I have had valuable discussions with many others in the Predictability Group and in AOPP more broadly, including Oscar Dimdore-Miles, Matt Chantry, Dave MacLeod, Daniel Befort, Tim Woollings, and Mat Patterson. I would also like to thank my aunt, Hilary Cadman, for her help with copy-editing, which helped me to catch a litany of typos, misconstructed sentences and missing commas in an earlier draft of this thesis.

Finally, I would like to thank NERC for funding my DPhil as part of the Environmental Research Doctoral Training Program (award number NE/L002612/1).

Abstract

The nonlinear, eddy-driven dynamics of the Euro-Atlantic troposphere are notoriously complex. The regime hypothesis – the idea that the continuous range of atmospheric flow states can be meaningfully partitioned into a small number of large scale 'regimes' – provides a conceptual framework from which to understand the impacts of internal variability and external forcing on the Euro-Atlantic.

In reanalysis and realistic model output, regime structure is noisy and often difficult to detect, confounded by the high dimensionality of the system and by model deficiencies. Therefore much of our understanding of regime systems comes from simple, conceptual models, where the mechanisms of regime dynamics can be studied cleanly. This dialogue across the hierarchy of models has been very fruitful, but it can be challenging to translate the behaviour of simple systems into testable hypotheses for the behaviour of the vastly more complex systems studied by the climate modeller or weather forecaster.

This thesis contributes to bridging this divide by working across the model hierarchy to compare the behaviour of simple and complex regime systems. In Chapter 1 we explain the phenomena of stochastically induced regime persistence using the Charney deVore model, and frame our results in physical terms. To assess the impact of stochasticity on more complex regime structure, we find we must first address the large sampling variability in regime statistics. We introduce a novel regime framework in chapter 2 which provides very stable regime structure, and use it to compare stochastic and deterministic climate model simulations. In chapter 3 we characterise historical regime variability and assess the regime response of European climate change in CMIP6. We explicitly test the long-standing hypothesis that anthropogenic forcing will primarily change the frequency of regime occurrence rather than the regime patterns themselves. We show the hypothesis broadly holds and then extend the simple regime forcing model to better capture climate model trends.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AR . . . . . Atlantic Ridge

BLK . . . . . [European/Scandinavian] Blocking

CMIP . . . . . Coupled Model Intercomparison Project

CdV . . . . . Charney deVore

DJF . . . . . Boreal Winter (December, January, February)

ENSO . . . . . El-Niño-southern oscillation

EOF . . . . . Empirical orthogonal function

HMM . . . . . Hidden Markov model

MTU . . . . . Model time unit

NAO . . . . . North Atlantic oscillation

NWP . . . . . Numerical weather prediction

PCA . . . . . Principal component analysis

SKEB . . . . . Stochastic kinetic energy backscatter

SPPT . . . . . Stochastic perturbed parameterisation tendency

SSP . . . . . Shared socio-economic pathway

T850 . . . . . 850 hectopascal Temperature

U850 . . . . . 850 hectopascal zonal wind speed

UPO . . . . . Unstable periodic orbit

YoTC . . . . . Year of tropical convection

Z500 . . . . . 500 hectopascal geopotential height

0 | Introduction

0.1 The atmospheric regime hypothesis

In this thesis we will study the dynamical features of regime systems in atmospheric data and in models of varying complexity – but what exactly is meant by the term 'regime', and what can adopting a regime approach lend to our understanding of the atmosphere? In this first short chapter we will discuss the concept and historical origin of regime systems in atmospheric physics, and provide a review of the current state of the art in the analysis of regime systems.

The fundamental idea behind the regime concept involves categorising the continuously evolving flow of a dynamical system – such as the atmosphere or some sub-region of it – into a discrete number of representative states with the aid of a clustering algorithm or other objective measure. The roots of this approach in fact stretch back to the earliest Western scientific studies of meteorology. In the late nineteenth and early twentieth century, prior to the development of numerical weather prediction or global observation networks, analogue forecasting was the primary tool at a meteorologist's disposal. This involved comparing the current synoptic flow state to (often limited) records of historical flow states, under the assumption that tomorrow's weather would likely be similar to the weather following historical analogue days. This task of searching for analogues was simplified by the subjective grouping of synoptic weather patterns into a more manageable number of *weather types* or *Grosswetterlagen* restricted to local regions of interest, and was applied in the continental US [Bowie and Weightman, 1914], the British Isles [Gold, 1920], and central Europe [Baur et al., 1944].

A natural philosophical question raised by the analogue approach is what meaning, if any, should be ascribed to the particular weather types used? As the question was posed by Franz Baur writing only a year after the first computerised forecasts [Baur, 1951]:

"Can the observed longer periods of persistent cold or warm, dry or wet weather be attributed to some major variable influences, or are they merely the consequence of the so-called "persistence tendency" in connection with random developments and the annual variation of meteorological elements?"

That is, are weather types purely artificial, deriving from the human predilection to seek patterns in auto-correlated randomness, or do they rather capture some 'major variable influences'; slow evolving dynamics that are intrinsic to the climate system and that might provide predictability beyond the synoptic range?

The assertion that an appropriately chosen discretisation of the atmospheric flow into a small number of representative patterns can isolate genuinely distinct dynamics of the circulation, identifying states that are visited recurrently and in preference to other arbitrary flow configurations, is the heart of the *regime hypothesis*. It remains a hypothesis, as despite the 70 years of progress in statistical methodologies, computing power, data availability and basic science since Baur first posed the question, there is still some disagreement over whether the atmosphere possesses identifiable regimes at all [Stephenson et al., 2004, Fereday et al., 2008, Fereday, 2017].

Even without accepting that the atmosphere has a strongly structured attractor supporting unique regimes, it is still possible to apply clustering algorithms to atmospheric data in order to generate a nonlinear basis of clusters from which the state dependence of any property of the atmosphere can be studied. These are normally referred to as 'weather types', following on from the early forecasting tradition, and are distinct philosophically from the concept of regimes, as we illustrate schematically in figure 1. This definitional distinction between regimes and weather types is present in the literature but is not always explicit, and the two perspectives are sometimes confused. Regime studies normally look for smaller numbers of distinct states and make efforts to justify the number of regimes used, whereas weather typing can involve large numbers of clusters (10+) and tends to choose the precise cluster number more arbitrarily.

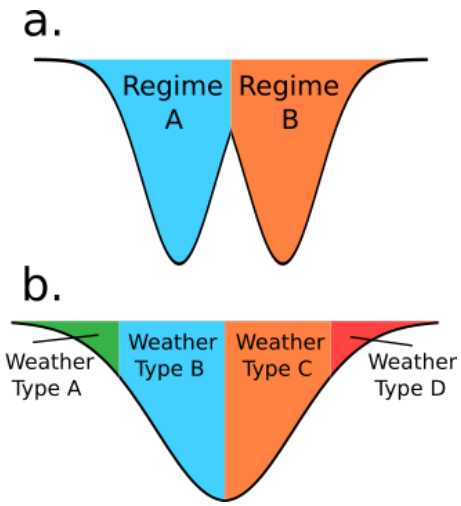


Figure 1: We draw a definitional distinction between *regimes*, which we consider as partitions of a variable's distribution that isolate distinct structures or features, and *weather types* which provide a more arbitrary discretisation. We illustrate this point in terms of a one-dimensional potential well structure, the minima of which represent quasi-stable equilibria states. System a) has two potential minima, and so naturally permits a description in terms of two regimes. System b) has only a single equilibria, but nevertheless it may be useful to arbitrarily divide the system into weather types.

Despite the continued use of weather typing as a meteorological tool, with movement towards more objective classification methods [Lund, 1963], theoretical studies did not start to utilise the concept of atmospheric multimodality until after the introduction of the Lorenz '63 system in [Lorenz, 1963]. Providing a concrete mathematical example of a chaotic system with a clear bimodal regime behaviour, the Lorenz '63 system remains a paradigmatic toy model for exploring regime systems, and it demonstrated how multimodal regime dynamics could be related to invariant measures of a dynamical attractor.

The Lorenz '63 system features unstable fixed points - the simplest type of invariant measure - at the centres of its regimes. More complex dynamical behaviour derives from these fixed points via a series of bifurcations, first producing stable periodic orbits around each fixed point, and then forming a chaotic strange attractor as the basins of attraction of the two fixed points coincide in a boundary crisis bifurcation. Therefore, the bimodal regime dynamics of the Lorenz '63 system can be ultimately related back to a simple fixed point analysis.

The idea that multiple atmospheric equilibria states might help explain the multiple weather types observed by forecasters, and in particular the persistence of anticyclonic blocking flows, was suggested by [Rossby, 1940] but not explored in detail until [Charney and DeVore, 1979], where two stable fixed points in a low-order barotropic spectral model were found to correspond to stable blocked and zonal flow states. This was later extended into a simple baroclinic setting [Charney and Straus, 1980], into a higher-resolution form [De Swart, 1988] and into a fully spherical geometry [Kallen, 1981]. Theoretical work soon followed showing how nonlinear barotropic resonances could produce multiple blocking equilibria with different topographic profiles [Trevisan and Buzzi, 1980, Malguzzi and Speranza, 1981], and found similar nonlinear responses in observational data [Benzi et al., 1986], while [Legras and Ghil, 1985] linked equilibria states to persistent zonal and blocked regimes in a 20-mode global model.

However, work in more complex baroclinic settings revealed challenges in directly relating large scale fixed points to the extended persistence of eddy-driven flows. In [Reinhold and Pierrehumbert, 1982] the introduction of a linearly unstable baroclinic wave was found to shift the model attractor away from fixed points, rendering them 'unrealisable', while for more fully converged truncations, much of the regime behaviour seen in low-order spectral models was found to vanish [Cehelsky and Tung, 1987]. An alternate quasi-empirical approach was introduced in [Vautard and Legras, 1988], where instead of analytically tracking the impact of small-scale instabilities on large-scale fixed points, they instead defined regimes in a statistical sense: as large scale flow configurations for which the average eddy tendency vanishes. This approach, later shown in [D'Andrea, 2002] to identify truly persistent states in a fairly realistic quasi-geostrophic model, amounts to taking a Poincare section through the subspace large-scale flow modes, and so provides an efficient heuristic for identifying almost-stable invariant measures more complex than a simple fixed point.

While the reproduction of regime dynamics in barotropic annulus experiments [Weeks

et al., 1997] suggests that the dynamics of fixed point regimes could be found even in untruncated flows, a more sophisticated mathematical perspective was needed to handle baroclinic dynamics. Progress in this line came from [Itoh and Kimoto, 1996] and [Itoh and Kimoto, 1997], which used results from a quasigeostrophic model to argue for weather regimes being linked to chaotic itinerancy. In this perspective, different regimes are associated with different attractors coexisting in phase space that merge into a single attractor, just as is now understood to occur in the Lorenz '63 system. The importance of preferred and avoided transitions between the regimes was also emphasised, as also seen in observations by [Kimoto and Ghil, 1993], and in later modelling studies [Kondrashov et al., 2004, Selten and Branstator, 2004].

Attractor merging is associated with the generation of more complex invariant measures such as homoclinic orbits, and can provide more insight into the mechanisms of regime *transitions*, only heuristically commented on in earlier, essentially steady state, theories. Less formally, a similar understanding of the role unstable equilibria could play was captured in the notion of 'ghost equilibria' introduced in [Ghil and Childress, 1987], leading to an understanding that even unrealisable equilibria of the large-scale flow could feasibly influence eddy-driven regime dynamics. With these theoretical developments in mind, [Crommelin et al., 2004] revisited the simple Charney deVore 1979 model and showed that for a specific range of parameters it possessed homoclinic orbits that gave rise to chaotic regime dynamics, related to transitions between the neighbourhoods of different unstable fixed points of the system.

The importance of small-scale modes, and transient dynamics, fundamentally derives from the ability of the large-scale flow to organise eddy activity, inducing secondary circulations that can alter the equilibria state. Not only can persistent atmospheric regimes exist in these eddy-driven flows, there is in fact much evidence that regimes in the real world rely on organised eddy activity to persist. [Shutts, 1983] introduced a theoretical model of eddy-straining to explain the anomalous persistence of blocking dynamics in

a simple chanel model. He found that as eddies were advected onto the upstream flank of a diffluent blocking anomaly, they would strain meridionally, and produce a net upstream flux of vorticity into the block, reinforcing it against dissipation. Evidence for the eddy-straining mechanism has been shown in conceptual [Haines and Marshall, 1987] and realistic models [Nakamura and Wallace, 1993], as well as in observations [Mullen, 1987]. While the specifics of the eddy-straining mechanism have more recently been questioned (such as in [Yamazaki and Itoh, 2013] which emphasises the importance of asymmetric vortex-vortex interactions in blocking persistence), the fundamentally multi-scale quality of the atmosphere is now well understood.

More recent mathematical approaches to understanding regimes have embraced this view of the atmosphere as inherently non-equilibrium, and indeed it is now known that chaotic attractors can be understood quite generally as being made up of a continuum of unstable periodic orbits (UPOs) [P Cvitanovic, 1991], all of which will exhibit variability in their small-scale motions. UPOs were used to identify blocking events in [Lucarini and Gritsun, 2020], with 'bundles' of UPOs acting as a generalisation of the more traditional regime framework [Maiocchi, C. C. and Lucarini, 2021].

Perhaps more successful than theoretical attempts to understand regime systems have been data-driven statistical approaches to identifying regimes both in model data and observations. Early methods searched for hemispheric regime patterns in the mid-tropospheric 500 hPa geopotential height field (Z500) by tracking persistent deviations from the mean state [Dole and Gordon, 1983, Shukla and Mo, 1983, Mo, 1986], isolating the now well-known Pacific-North-American pattern, the Arctic oscillation, and the Northern Annular Mode. A reduced-dimension approach based on empirical orthogonal functions (EOFs) was introduced in [Mo and Ghil, 1988], which allowed for robust estimates of multi-modality with more robust statistical significance testing [Molteni et al., 1990, Toth, 1993, Xinhua Cheng and Wallace, 1993]. It also recognised the importance of evaluating sectoral regimes [Kimoto et al., 1993], which could ultimately be linked back to

hemispheric wave activity in some cases [Molteni et al., 1990].

An impactful methodological refinement was introduced in [Michelangeli et al., 1995], which first proposed using K-means clustering for identifying sectoral regime patterns, arguing it was more robust than previous pdf-estimation methods, although high variability still makes statistically significant regime identification challenging [Molteni and Corti, 1998]. [Michelangeli et al., 1995] also first presented the commonly used four Euro-Atlantic circulation regimes, capturing two phases of the North Atlantic Oscillation (NAO), an Atlantic Ridge and a European or Scandinavian blocking pattern. Other clustering frameworks have been used in a sectoral regime context, such as Gaussian mixtures [Kondrashov et al., 2004, Hannachi and Hannachi, 2007], hierarchical methods including self-organised maps [Hewitson and Crane, 2002] and nonlinear PCA [Monahan et al., 2001], although K-means remains a common choice owing to its simplicity. Increasingly sophisticated significance testing has been used to justify choices of regime number, as in [Straus et al., 2007] and [Strommen et al., 2019], after the ability of autocorrelated [Sura et al., 2005] and skewed [Christiansen, 2007] linear systems to produce suggestive, but ultimately spurious, multimodality was documented.

Whereas circulation regimes identified in Z500 data aim to capture the persistent dynamics of blocking or wave resonance directly, the past decade has seen a valuable parallel approach develop via the concept of Euro-Atlantic jet regimes found in the unidimensional jet latitude index of [Woollings et al., 2010]. Hidden Markov model [Franzke et al., 2011] and Gaussian mixture [Hannachi et al., 2012] approaches have been used to identify a robustly trimodal regime structure, which can be partially mapped to equivalent circulation regimes as in [Madonna et al., 2017] and our own Chapter 2.

Though clustering approaches to regime identification are still the most common, exciting mathematical tools are emerging that offer complementary methods, including operator theoretic constructions such as dynamical mode decomposition [Gottwald and Gugole, 2020], the topological approach of [Strommen et al., 2021], and the extremal indices of

[Faranda et al., 2016], leveraging the modern theoretical understanding of regime systems. There has also been progress in embedding a persistence constraint directly into the K-means algorithm through the use of linear programming [Falkena et al., 2020].

The theory of hemispheric persistent waves is being revisited in the context of Northern hemisphere summer heatwaves by [Kornhuber et al., 2017, Kornhuber et al., 2019], and tropical variability modes such as ENSO and the Madden-Julian Oscillation can also be broadly understood as kinds of regime systems. However the most unambiguous and most studied regimes are those of the wintertime Euro-Atlantic circulation and its corresponding jet dynamics.

Euro-Atlantic circulation regimes have tremendous utility in helping to understand the flow dependent predictability of the atmosphere [Frame et al., 2013, Ferranti et al., 2015a, Matsueda and Palmer, 2018], in developing a holistic picture of model biases [Dawson et al., 2012, Fabiano et al., 2020] and in modulating the impact of remote teleconnections, such as from the Madden-Julian Oscillation [Cassou, 2008] or the stratosphere [Charlton-Perez et al., 2018] [Beerli and Grams, 2019]. Approaches generalised to cover the entire year have also been used to understand flow dependent impacts in applied settings such as energy generation [Grams et al., 2017a, Van Der Wiel et al., 2019, Garrido-Perez et al., 2020].

More abstracted lessons can also be learned from considering the Euro-Atlantic circulation as a regime system. For example, much has been made in recent years of the signal-to-noise problem [Eade et al., 2014, Dunstone et al., 2016, Scaife and Smith, 2018], where seasonal forecasts of the NAO prove more skillful than their ensemble spread would suggest they should be according to linear theory. However [Strommen and Palmer, 2018] showed that if the NAO is modelled as a bimodal regime system, the paradox of low signal-to-noise could be explained as a result of insufficiently persistent regimes in the model.

Additionally, no issue in atmospheric science is more pressing than understanding how

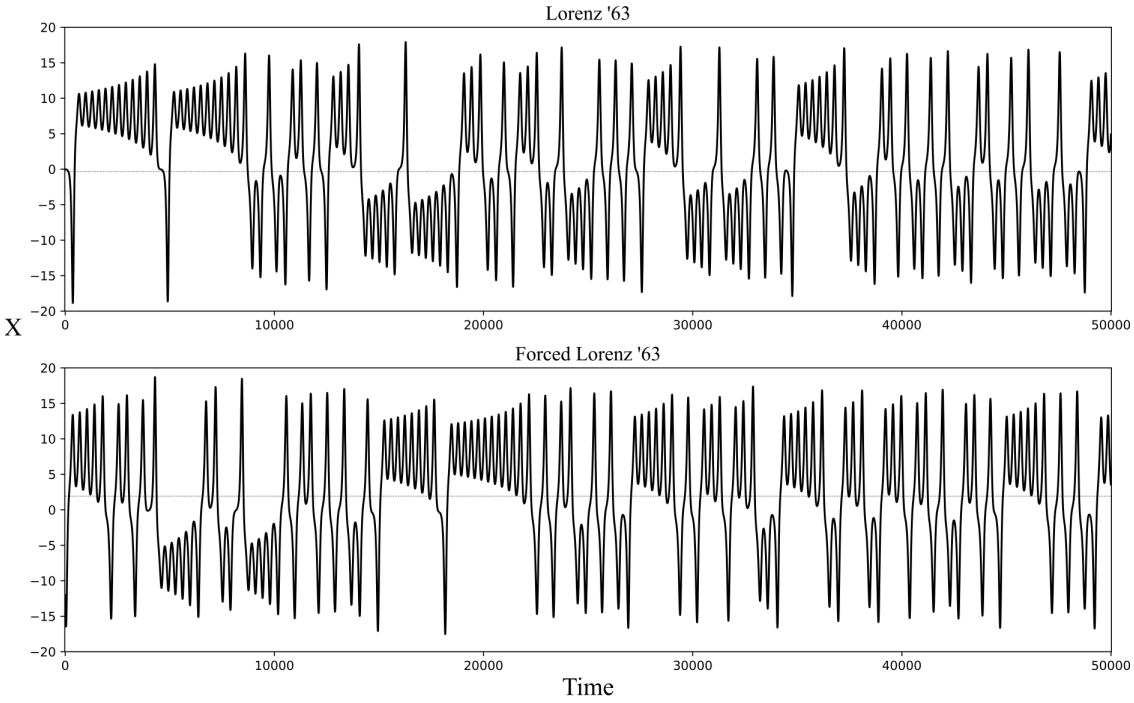


Figure 2: An illustration of the non-linear "temporal" response of simple regime systems to applied forcing. A short integration of the Lorenz '63 system (top) shows symmetrical regime dynamics. When a forcing term is applied to the x-variable (bottom) the position of the regimes does not change, but the positive-x regime is visited more frequently and becomes more persistent.

the climate will change as a result of anthropogenic global warming, especially in specific regional detail. Theoretically, a regime system may respond very differently to an applied forcing than a linear system would, as demonstrated in figure 2 for the simple Lorenz '63 system. Following the procedure of [Palmer, 1993], the Lorenz system has been integrated both with and without a constant additive forcing term added to the x-variable tendency. Whereas in a linear system we might expect the whole attractor to shift towards higher x-values, we instead find that the two wings of the attractor stay essentially where they were in the unforced system, but now the positive-x regime is more persistent and occurs more often. If we interpret our additive forcing as the influence of climate change on this toy model, then we are led to the hypothesis of [Palmer, 1993], expanded on in [Palmer, 1999], that the impact of climate change on European weather should be to shift the likelihood of certain circulation patterns occurring, rather than generate qualitatively new weather states. Schematically, as shown in figure 3, we can summarise this

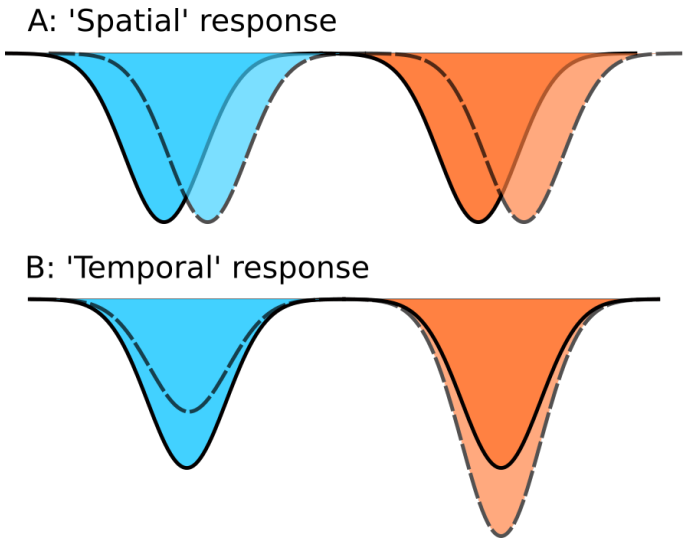


Figure 3: A system variable with multiple potential wells might respond to an applied forcing in two ways. Either the equilibria may shift, resulting in a (phase) spatial shift of the variable (A), or the time the system spends in the vicinity of the different equilibria may shift (B). Combinations of the two responses are also possible.

as the distinction between a linear 'space-like' response, where the attractor is shifted in phase space, or a non-linear 'time-like' response, where the frequency of occurrence and lifetime of the regimes changes. The hypothesis of [Palmer, 1999] has been influential, and yet it has never been explicitly tested - although see our Chapter 3 for just such a test.

There are many open questions and gaps in our knowledge of regime systems, especially in complex models and in observations, where the curse of dimensionality makes it difficult to apply simple mathematical tools, and where the sheer complexity of the mid-latitude circulation has so far stymied attempts to build up a clear physical picture of why persistent regimes occur.

This incomplete understanding makes it difficult to understand how to improve model representation of regimes, or to understand *why* certain model changes - such as the introduction of stochastic parameterisation schemes - improve regime representation. Since improving model regime structure and predictability is an important step in developing skilful subseasonal and seasonal European forecasts, this is an issue of considerable importance. Further, there is still dispute about the significance of circulation regimes even

in the Euro-Atlantic, a situation which is at odds with the visually apparent trimodality of the jet picture, and which does have meaning beyond the merely philosophical - whether the atmospheric circulation is in a strongly non-linear domain or not has important impacts on both predictability and forced response. Synthesising the different perspectives on mid-latitude variability, strengthening the statistical foundations of regime approaches and understanding how simple models of regimes relate to complex models are all important challenges to be tackled in the coming years, and we contribute to that challenge in this thesis.

0.2 Outline of thesis

In this thesis we will tackle several challenging open questions concerning Euro-Atlantic circulation regimes. To do so, we will make use of low-order modelling, analysis of observational regime dynamics, and assessment of regime representation in complex climate models.

In Chapter 1 we aim to understand how stochasticity can counter-intuitively act to increase the persistence of regimes, as has been noted in several simple dynamical systems, through a detailed investigation of the Charney deVore model. Once we obtain a mathematical picture of the persistence mechanism we leverage the physical basis of the model to frame our results in interpretable terms.

In Chapter 2 we begin by tackling the question of regime stability: why can it be so hard to find regime patterns that are robust to small shifts in time period and in different datasets? We introduce a geopotential-jet regime framework where we regress out the influence of jet speed prior to clustering and find substantial improvements in regime stability. We then provide a multi-reanalysis assessment of 20th century regime variability, providing a robust yardstick for model evaluation. We then use this framework to explore the impacts of stochasticity on climate model regimes in the EC-Earth model. We find it significantly improves regime patterns, but changes in regime occurrence and persistence

are mixed and hard to understand clearly.

Finally, in Chapter 3, we analyse CMIP5 and CMIP6 regimes, providing an assessment of improvements and persistent biases in the historical representation of Euro-Atlantic variability in the latest generation of climate models before moving on to analysing the projected trends in regime occurrence, persistence and pattern changes under climate change. We show that future regime changes are primarily temporal, in agreement with the hypothesis of [Palmer, 1999], and generalise the forcing profile used there to provide a closer match to the CMIP6 trends.

0.3 Statement of originality

Parts of Chapter 2 are based on a joint-first author paper [Dorrington and Strommen, 2020]. Kristian Strommen processed raw data and produced plots of the Z500 phase space for that work, while cluster analysis and other figures were prepared by Joshua Dorrington. Text in that paper was the contribution of both authors, and so is not included verbatim in this thesis, although Sections 2.1 and 2.4 are based on it.

The work in Chapter 3 is part of a collaboration with Kristian Strommen and Federico Fabiano, who helped to process the raw CMIP data. All text is the author's own, and where a figure has been produced by a collaborator it is explicitly noted.

1 | Regime dynamics in the Charney-deVore 1979 model

1.1 Introduction

In chapter 0 we introduced the history and motivation behind the concept of regimes, from the early empirical observations of recurring patterns in European weather charts to the paradigmatic Lorenz '63 model with its striking 'butterfly wing' bimodality. We have discussed how these developments have given rise to the modern theories of atmospheric circulation regimes and nonlinear dynamical systems, and seen how regime dynamics can generate counter-intuitive flow behaviours and forcing responses.

There are, of course, radical differences between the simple regime systems that interest applied mathematicians, and the 'kitchen sink' Earth system and numerical weather prediction (NWP) models used by climate scientists and meteorologists. Fundamentally, these are differences of complexity and state-space dimension. While recent cutting edge applications of unstable periodic orbit analysis have brought the tools of dynamical systems theory to the scale of models with $\mathcal{O}(10^3)$ degrees of freedom [Lucarini and Gritsun, 2020], this is still a far cry from the $\mathcal{O}(10^7)$ – $\mathcal{O}(10^9)$ degrees of freedom found in CMIP6 or operational NWP models. Even these are something of an outlier, with many simple regime models having only a few dozen degrees of freedom at most.

These simple models have many benefits, namely that they allow us to understand in detail every aspect of the dynamics, and they can capture fundamental behaviours of interest with as few extraneous features as possible. Nevertheless, just as a map is not the territory it represents, neither is (as an example) the bimodality of the Lorenz '63 system the bimodality observed in the North Atlantic oscillation; we only imagine it might have a similar structure. It can be challenging to evaluate whether the regimes

in simple models have the same origins and same underlying mathematical structure as the regimes found in complex models, because the tools we use to study them are very different. We know how bifurcations and invariant measures can induce regimes in simple models, but we cannot find those features in complex models, owing to the curse of dimensionality. Similarly, we know that Atlantic regime dynamics have links to the eddy driven jet and to persistent blocking, but simple models do not have enough resolution to truly replicate these spatially localised dynamics.

In this first chapter we attempt to narrow this divide, through in-depth analysis of the stochastically forced Charney-deVore (CdV) model. This simple spectrally truncated model, first introduced in [Charney and DeVore, 1979] (hereafter CdV79), has an important role in the history of regimes, as it provided the first clear evidence of multiple stable flow states in a simplified atmospheric model. We will discuss the equations, their derivation and their physical meaning in detail in Section 1.2, but for now we simply say that the six-equation model describes a large-scale barotropic flow in a mid-latitude beta-plane channel in the presence of orography.

CdV79 showed that the flow in their model could equilibrate either by reaching a stable, approximately zonally symmetric, 'high index' flow, or by reaching another stable 'low index' flow, defined by a resonant wave and anticyclonic ridging. There, it was suggested that the impact of the unresolved small-scale modes (in this idealised model, "small" includes spatial scales of $\sim \mathcal{O}(1000)$ km) was to destabilise these equilibria states, causing unpredictable shifts from one state to another and introducing an erratic vacillation between large scale weather patterns.

More recently, [Crommelin et al., 2004] (C04) reformulated the CdV79 model, using different parameter values that were at least as realistic, if not more so, than those used originally. The benefit of this reformulation was to demonstrate that the simple six-mode model was able to generate chaotic dynamics with multimodal regime behaviour, freely vacillating between a long-lived blocking state and a more zonal flow state, without ex-

ternal forcing or requiring unresolved small-scale instabilities. This reformulated model was used by [Kwasniok, 2014] (K14) to demonstrate that adding a stochastic forcing term into the system could act to stabilise regime dynamics, increasing the mean lifetime of the blocking state, quite contrary to the noise induced *de*stabilisation suggested by CdV79. It is this reformulated chaotic system, owing to C04, that we study in this chapter, not the original version of CdV79.

The idea of noise-induced regime stabilisation runs quite contrary to intuition; after all, we do not tend to think of fast varying, random processes as encouraging a predictable, slow evolution. However, stochasticity has been shown to reduce climatological model biases [Palmer and Weisheimer, 2011, Berner et al., 2012] and to improve regime predictability in truncations of the Lorenz '96 model [Düben et al., 2014, Christensen et al., 2015a] and regime representation in climate models [Dawson and Palmer, 2015]. Despite these observations, the details of how exactly stochasticity impacts regimes remain unclear. It is therefore important to explore the mechanisms of stochastic regime persistence, in order to better understand the physics at play, and to assess whether these results should generalise to models of different origin and complexity.

But why study the impact of stochastic regime forcing in the CdV model, rather than, say, the Lorenz '63 model, which also showed stochastic regime stabilisation in K14, or a much more realistic model? Because the CdV model has both few enough degrees of freedom that it can be understood entirely, and a firm rooting in barotropic vorticity dynamics, so that any features it exhibits can be understood in terms of physically meaningful mechanisms. Whether these extremely simplified physical mechanisms are relevant in the real world is a different question, but by being able to translate between the languages of dynamical systems (fixed points and bifurcations) and the language of atmospheric physics (vorticity budgets and Rossby waves), we can produce testable hypotheses about how complex regime systems might behave. One long-standing hypothesis to explain the persistence of atmospheric blocking events is that the straining

of transient eddies on the upstream flank of a block serves to inject vorticity into the large scales, and so oppose dissipation [Shutts, 1983]. It is quite conceivable that eddy-forcings might be intensified by stochasticity, and we shall explore the relevance of this eddy-straining mechanism in CdV.

In Section 1.2 we first derive the Charney deVore model from the shallow water equations, emphasising the physical basis of the model and taking note of the approximations we are introducing. Section 1.3 then discusses the behaviour of the model in both stochastic and deterministic formulations and introduces the vacillating regime dynamics, explaining their dynamical origin as a result of collapsed heteroclinic cycles between unstable equilibria. Section 1.4 discusses the hidden Markov model algorithm used to objectively partition the flow into regimes and the changes in regime statistics caused by increased stochastic forcing. In addition to making regimes more persistent, stochasticity impacts the predictability of the model, and we demonstrate a conditional increase in regime predictability in 1.5, as well as non-monotonic increases in ensemble spread with increasing noise amplitude.

After taking full account of the model dynamics, we then begin to address the 'why' of increased persistence in Section 1.6, where we introduce a simple mathematical argument that demonstrates the ability of stochasticity to destabilise low dimensional flow trajectories and to suppress instability growth. Section 1.7 continues this theme, now using a physically motivated vorticity analysis to show that eddy forcings in fact have little role in the regime persistence, and that instead the ability to access deterministically inaccessible weak Westerly flow states is the key to stochastic persistence. We show that this ability is amplified by a phase locking of the wave modes during blocking which confines the flow to the neighbourhood of the weakly unstable blocking fixed point. We then comment on the difficulties of generalising the model to higher resolution in Section 1.8, before summarising our analysis in Section 1.9, where we discuss lessons learned and possible extensions to the work presented here.

1.2 Derivation of the model

The CdV o.d.e. system can be derived directly from the basic physics of fluid flows, admittedly with a large number of approximations and simplifications introduced. Here we derive the system from the shallow water equations, partly to emphasise this physical basis, and also because the full derivation is in fact not present anywhere in the existing literature. Whenever a new assumption or approximation is introduced it will be **emphasised in boldface**. The starting point is the single-layer shallow water system, which we will then reformulate in terms of vorticity and simplify to produce the barotropic vorticity equation in the presence of orography. We then introduce the spectral truncation, which transforms the p.d.e to the six-equation o.d.e system derived in CdV79. Finally, we discuss the appropriate way to decompose the vorticity budget in the model.

1.2.1 Derivation from the shallow water equations

The shallow water system is a simple but highly useful model of fluid flow. Consider a **single layer** of fluid, with **constant density** ρ, **negligible viscosity**, a vertical pressure gradient deriving solely from gravity (i.e. the fluid is in **hydrostatic equilibrium**), and which, as the name implies, has a representative **horizontal length scale** L **much larger than the vertical length scale** H. The dynamics of this system are completely determined by the momentum equation and the mass-conservation equation.

The momentum equation in a rotating frame is given by

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{v} + \mathbf{f} \times \mathbf{v} = -g\boldsymbol{\nabla} H \tag{1.1}$$

while the mass-conservation equation is

$$\frac{\mathrm{D}\, H}{\mathrm{D}\, t} + H\boldsymbol{\nabla} \cdot \mathbf{v} = 0 \tag{1.2}$$

Where $\mathbf{v}$ is velocity, $\mathbf{f}$ represents Coriolis forces, g is simply the gravitational constant (**assumed constant**) and H is the depth of the fluid column. Here the mass-conservation simply expresses that if velocities converge at a point, the height of the fluid column must rise in response. Application of the identity $(\mathbf{a}\cdot\mathbf{b})\mathbf{c} = -\mathbf{a}\times(\mathbf{b}\times\mathbf{c})+(\mathbf{a}\cdot\mathbf{c})\mathbf{b}$ to equation 1.1 gives us

$$\frac{\partial \mathbf{v}}{\partial t} + [(\boldsymbol{\nabla}\times\mathbf{v})\times\mathbf{v} + \frac{1}{2}\boldsymbol{\nabla}v^2] + \mathbf{f}\times\mathbf{v} = -g\boldsymbol{\nabla}H \tag{1.3}$$

and by introducing the absolute vorticity $\xi = \boldsymbol{\nabla}\times\mathbf{v}$, and taking the curl of both sides, we get the shallow water vorticity equation:

$$\frac{\partial \boldsymbol{\xi}}{\partial t} + \boldsymbol{\nabla}\times((\boldsymbol{\xi}+\mathbf{f})\times\mathbf{v}) = 0 \tag{1.4}$$

Again, we apply the triple product identity and the chain rule and find:

$$\frac{\partial \boldsymbol{\xi}}{\partial t} + (\mathbf{v}\cdot\boldsymbol{\nabla})(\boldsymbol{\xi}+\mathbf{f}) + (\boldsymbol{\xi}+\mathbf{f})(\boldsymbol{\nabla}\cdot\mathbf{v}) = 0 \tag{1.5}$$

Where $\boldsymbol{\nabla}\cdot\boldsymbol{\xi} = \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{v}) = 0$ by definition and $(\boldsymbol{\xi}\cdot\nabla)\mathbf{v} = 0$ because $\boldsymbol{\xi}$, as a curl of horizontal wind, does not project onto the variability of $\mathbf{v}$. Indeed, at this point we note that both $\boldsymbol{\xi}$ and $\mathbf{f}$ have no horizontal components, and so we focus only on the scalar equation for the z-component, letting $\xi := \xi_z$ and $f := f_z$ for convenience. The divergent velocity term can be rewritten with the aid of equation 1.2 by rearranging:

$$\boldsymbol{\nabla}\cdot\mathbf{v} = \frac{-1}{H}\frac{\mathrm{D}\,H}{\mathrm{D}\,t} \tag{1.6}$$

and substituting into equation 1.5. This leaves us with

$$\frac{\mathrm{D}\,(\xi+f)}{\mathrm{D}\,t} - \frac{(\xi+f)}{H}\frac{\mathrm{D}\,H}{\mathrm{D}\,t} = 0 \tag{1.7}$$

which can be put into a clearer form using the chain rule:

$$\frac{\mathrm{D}}{\mathrm{D}\,t}\left[\frac{(\xi+f)}{H}\right]=0 \tag{1.8}$$

That is, the potential vorticity $q=\frac{\xi+f}{H}$ is materially conserved. If the thickness of the fluid increases, then the absolute vorticity must likewise increase, due to the stretching of the flow.

We divide our fluid height H into some average fluid depth D and an anomaly component $h(x,y)$ describing the height of surface orography, so that $H=D-h(x,y)$. We assume **shallow orography** so we may take the limit where $\frac{h}{D}<<1$ and Taylor expand H, truncating at first order so that $\frac{1}{H}\approx\frac{1}{D}+\frac{h}{D^2}$:

$$\frac{\mathrm{D}}{\mathrm{D}\,t}\left[\xi+f+\frac{(\xi+f)h(x,y)}{D}\right]=0 \tag{1.9}$$

In the midlatitude atmosphere this is a reasonable approximation at the largest scales. For example, if we let D be the height of the tropopause – around 8 km – and we let h be the average elevation of Europe or North America – between 500 and 800 metres – then the limit holds. We also choose to place ourselves in the **quasigeostrophic limit** of low Rossby number, where $\xi<f$. Finally, we take a β **plane approximation** to the Coriolis force, and this brings us to the barotropic vorticity equation in the presence of shallow orography:

$$\frac{\mathrm{D}}{\mathrm{D}\,t}\left[\xi+\beta y+\frac{f_0}{D}h(x,y)\right]=0=\frac{\mathrm{D}\,q}{\mathrm{D}\,t} \tag{1.10}$$

We define $\gamma=\frac{f_0}{D}$ and introduce the streamfunction Ψ, given by $u=-\frac{\partial\Psi}{\partial y}$, $v=\frac{\partial\Psi}{\partial x}$, $\xi=\nabla^2\Psi$. We also generalise slightly by introducing an external source (or sink) of vorticity F. Writing our material derivative out explicitly, we now have:

$$\frac{\partial \nabla^2 \Psi}{\partial t} - \frac{\partial \Psi}{\partial y} \cdot \frac{\partial}{\partial x}[\nabla^2 \Psi + \gamma h + \beta y] + \frac{\partial \Psi}{\partial x} \cdot \frac{\partial}{\partial y}[\nabla^2 \Psi + \gamma h + \beta y] = F \tag{1.11}$$

Or more concisely

$$\frac{\partial \nabla^2 \Psi}{\partial t} + \mathcal{J}(\Psi, \nabla^2 \Psi + \gamma h) - \beta \frac{\partial \Psi}{\partial x} = F \tag{1.12}$$

Where the Jacobian operator is given by

$$J(A, B) := \frac{\partial A}{\partial x}\frac{\partial B}{\partial y} - \frac{\partial B}{\partial x}\frac{\partial A}{\partial y} \tag{1.13}$$

Taking our **forcing to be a linear relaxation** towards a background state, we now have our explicit equation for the evolution of Ψ:

$$\frac{\partial}{\partial t}\boldsymbol{\nabla}^2 \Psi = -J(\Psi, \boldsymbol{\nabla}^2 \Psi + \gamma h) - \beta \frac{\partial \Psi}{\partial x} - C(\Psi - \Psi^*) \tag{1.14}$$

This is the CdV79 equation, before spectral truncation, where Ψ^* is our background streamfunction. In section 1.7 we will revisit this equation in its time-averaged vorticity formulation, but for now we will think in terms of the streamfunction. We consider the flow in a **channel geometry**, with periodic boundary conditions in the zonal direction, and with no-penetration and no-slip conditions in the meridional direction. Of course the real atmosphere has no hard meridional boundaries; however the impacts of the subtropical and midlatitude jets have been shown to produce a waveguiding effect that meridionally confines low-frequency variability in a similar way, both in low-order models [Hoskins and Ambrizzi, 1993, Yang et al., 1997] and in observations [Branstator, 2002], which makes this a reasonable first approximation.

We consider a channel of dimensionless zonal length $x = [0, 2\pi]$, and meridional length $y = [0, b\pi]$, so that b then is a free parameter controlling the shape of the channel.

1.2.2 Discretising the model

To convert this infinite dimensional partial differential equation into a system of ordinary differential equations, we can use a Galerkin projection. This involves expanding our spatially dependent variables (Ψ, Ψ^*, and h) in terms of an infinite linear sum of complete spatial basis functions, labelled $\{\Phi(x,y)_i\}$:

$$\Psi(x,y,t) = \sum_i^\infty \Psi_i(t)\Phi_i(x,y) \tag{1.15}$$

$$\Psi^*(x,y) = \sum_i^\infty \Psi_i^*\Phi_i(x,y) \tag{1.16}$$

$$h(x,y) = \sum_i^\infty h_i\Phi_i(x,y) \tag{1.17}$$

We start by defining this basis abstractly, requiring the basis functions to be eigenfunctions of the Laplacian operator and the zonal derivative:

$$\nabla^2\Phi_i = -a_i^2\Phi_i \tag{1.18}$$

$$\frac{\partial\Phi}{\partial x} = b_i\Phi_i \tag{1.19}$$

The Jacobian term is the thorniest point in the spectral analysis, as it non-linearly mixes spectral modes together. It is the impact of this term that prevents an arbitrary spectral truncation of equation 1.14 being a solution of the full equation. In full generality, we simply consider that the Jacobian of each pair of basis functions has some projection on every other basis function:

$$J(\Phi_j, \Phi_k) = c_{ijk}\Phi_i \tag{1.20}$$

The projection coefficient of the jth and kth modes on the ith mode, c_{ijk}, is computable simply through a direct pointwise product of the ith mode and the Jacobian, integrated over the domain and appropriately normalised:

$$c_{ijk} = \frac{\int_0^{2\pi}\int_0^{b\pi} \Phi_i J(\Phi_j, \Phi_k) dydx}{\int_0^{2\pi}\int_0^{b\pi} dydx} \tag{1.21}$$

Under these constraints, we can now rewrite equation 1.14 projected onto the ith basis function as:

$$\frac{\partial}{\partial t}\Psi_i = \beta a_i^{-2} b_i \Psi_i - C(\Psi_i - \Psi_i^*) - \sum_{j,k} a_i^{-2} c_{ijk}(a_k^2 \Psi_j \Psi_k - \gamma \Psi_j h_k) \tag{1.22}$$

Where the double sum indexes over all basis functions. We make an explicit choice of basis functions now, taking a Fourier basis defined as:

$$\begin{cases} \sqrt{2}\cos\left(\frac{ny}{b}\right) & \text{if } m = 0 \\ \sqrt{2}e^{imx}\sin\left(\frac{ny}{b}\right) & \text{else} \end{cases}$$

where the meridional mode $n > 0$ and the zonal mode m can take any non-integer value. It is easy to verify that these basis functions satisfy the boundary conditions and eigenfunction constraints defined above. To actually solve these spectrally transformed equations, it will be necessary to introduce a truncation, disregarding wave modes exceeding a cutoff wavenumber. Here we choose to introduce a zonal threshold of $|m| \leq 1$ and a meridional threshold of $n \leq 2$, **restricting us to just 6 basis functions**, as in CdV79, C04 and K14. Later, we will also consider a higher zonal threshold of $|m| \leq 2$, which gives us a 10 equation system introduced in [De Swart, 1988].

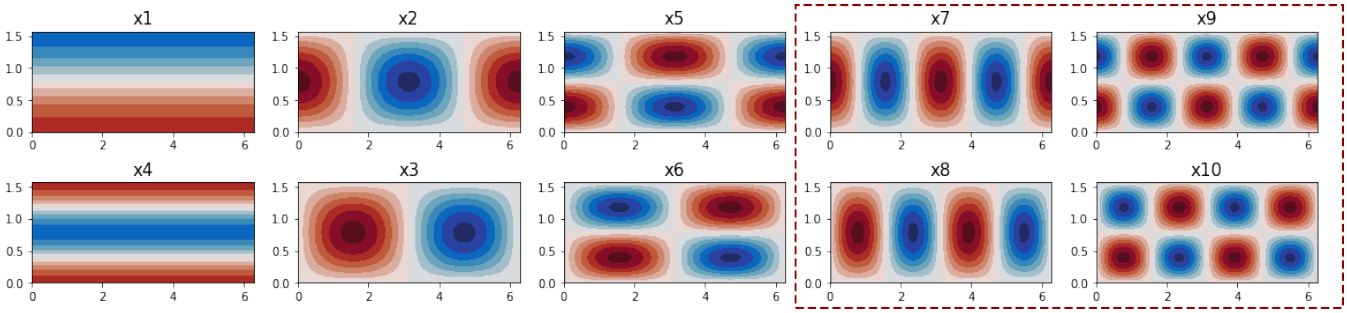


Figure 1.1: The real-valued basis functions used in the spectrally truncated CdV79 system, plotted as streamfunctions. The modes contained within the dotted box are introduced as part of the de Swart 1988 10-mode model.

Although up to this point our approximations have been well motivated physically, we cannot rigorously justify introducing such a drastic truncation without resorting to self-interest: such a system is much easier to study, and can be understood in more detail than a high-dimensional model. Nonetheless, we may reason that any phenomena observed in this truncated model, which will have to derive solely from the dynamics of the largest scales, may have analogues in the true atmospheric circulation, although there they will be shaped and influenced by coupling to smaller scales. The resulting basis functions are shown in figure 1.1.

Concretely, while the true spectral equation for the largest wave modes, equation 1.22, contains an infinite number of nonlinear coupling terms, after truncation we are reduced to only 36 such terms, representing only interactions between the resolved modes: the ability for small scales to feed vorticity into the untruncated large scales is neglected.

We choose our **linear relaxation to be zonally symmetric**, representing the forcing effect of the pole-to-equator temperature gradient. We choose our orography to include a single peak straddling the periodic boundary, with a trough in the centre of the domain. $h(x, y) = \gamma\sqrt{2}\cos(x)\sin\left(\frac{y}{b}\right)$ (this is shown in figure 1.2).

Obtaining the coefficients in 1.22 is now simply a matter of calculating integrals, but is very involved, and is best performed using symbolic computation software (an implementation using sympy [Meurer et al., 2017] can be found at `https://github.com/`

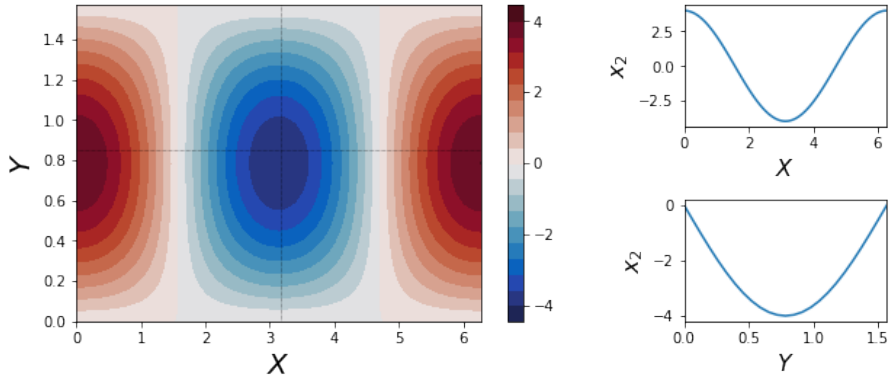


Figure 1.2: The orographic profile used in the CdV79 system, with a trough in the centre of the domain and a ridge over the periodic boundary. Sub-panels show cross sections across the domain, following the dotted black lines.

joshdorrington/Thesis_materials/blob/main/10d_mode_derivation.ipynb). After computing coefficients, we now diverge slightly from CdV79, transforming our variables to a real basis, as in C04, given by:

$$
\begin{aligned}
x_1 &= \tfrac{1}{b}\Phi_{0,1} & x_4 &= \tfrac{1}{b}\Phi_{0,2} \\
x_2 &= \tfrac{1}{\sqrt{2}b}(\Phi_{1,1} + \Phi_{-1,1}) & x_3 &= \tfrac{i}{\sqrt{2}b}(\Phi_{1,1} - \Phi_{-1,1}) \\
x_5 &= \tfrac{1}{\sqrt{2}b}(\Phi_{1,2} + \Phi_{-1,2}) & x_6 &= \tfrac{i}{\sqrt{2}b}(\Phi_{1,2} - \Phi_{-1,2})
\end{aligned}
$$

This transformation is purely cosmetic; the behaviour of the system is unchanged. At the end of this process we obtain an explicit six-equation system:

$$
\begin{aligned}
\dot{x}_1 &= -C(x_1 - x_1^*) + \tilde{\gamma}_1 x_3 \\
\dot{x}_2 &= -Cx_2 + \beta_1 x_3 - \alpha_1 x_1 x_3 - \delta_1 x_4 x_6 \\
\dot{x}_3 &= -Cx_3 - \beta_1 x_2 + \alpha_1 x_1 x_2 + \delta_1 x_4 x_5 - \gamma_1 x_1 \\
\dot{x}_4 &= -C(x_4 - x_4^*) + \epsilon(x_2 x_6 - x_3 x_5) + \tilde{\gamma}_2 x_6 \\
\dot{x}_5 &= -Cx_5 + \beta_2 x_6 - \alpha_2 x_1 x_6 - \delta_2 x_4 x_3 \\
\dot{x}_6 &= -Cx_6 - \beta_2 x_5 + \alpha_2 x_1 x_5 + \delta_2 x_4 x_2 - \gamma_2 x_4
\end{aligned}
\tag{1.23}
$$

Where colours encode the linear relaxation, Coriolis forces, orographic impacts and non-

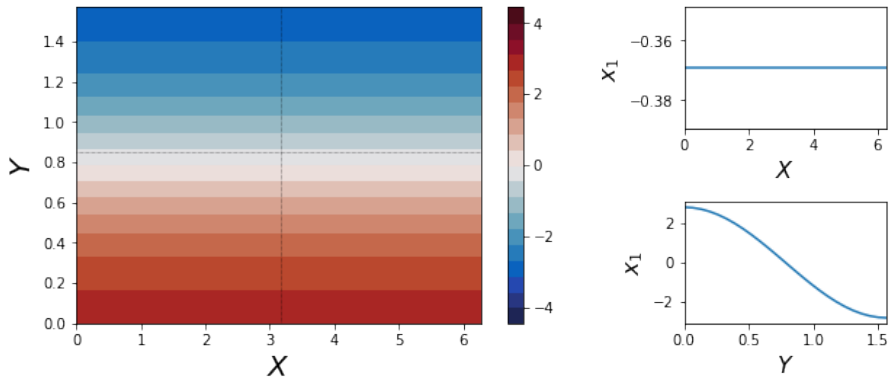


Figure 1.3: The background state in the Crommelin '04 formulation of the CdV system. It is zonally symmetric and represents the impact of the meridional temperature gradient. Again, sub-panels show cross sections across the domain, following the dotted black lines.

linear advection. We see that there are far less than the 36 promised non-linear terms per equation: most quadratic interactions integrated over the domain vanish due to symmetry considerations. $\{\alpha_n\}, \{\beta_n\}, \{\delta_n\}, \{\gamma_n\}, \{\tilde{\gamma_n}\}$ and ϵ are simple geometric constants of $\mathcal{O}(1)$ given in C04.

There are now a number of dimensionless free parameters in the model, which we set equal to the values used in K14 and C04:

- x_i^*: The forcing vector, representing thermal relaxation towards a background state. We use a zonally symmetric forcing profile with $x_1 = 0.95, x_4 = -0.76095$, all other terms zero (shown in figure 1.3).
- C: The thermal relaxation timescale. Set to $C = 0.1$, describing a damping time of $\sim$10 days.
- γ: The orographic amplitude. Set to $\gamma = 0.2$, corresponding to a 200-metre amplitude.
- β: The Coriolis parameter. Set to $\beta = 1.25$, defining a central latitude of 45°.
- b: Defines the channel half-width. Set to $b = 0.5$, which gives a channel of 6300 km x 1600 km

1.2.3 Vorticity formulation of CdV79

To analyse the vorticity budget of the full untruncated p.d.e, we would start by writing equation 1.14 in terms of the vorticity ξ, take a Reynolds decomposition in time into a mean component $\overline{\xi}$ and an anomaly ξ', and then time average to obtain:

$$\overline{\frac{\partial \xi'}{\partial t}} = -\overline{\mathbf{u}} \cdot \nabla \overline{\xi} - \left[\gamma \overline{\mathbf{u}} \cdot \nabla h + C(\overline{\xi} - \xi^*)\right] - \overline{\mathbf{u}' \cdot \nabla \xi'} \tag{1.24}$$

where the terms on the RHS describe the mean vorticity advection, mean linear forcings, and eddy advection, respectively. However, this budget is not conserved in our truncation, due to the nonlinear advection terms, which mix up the resolved and unresolved spatial modes. Therefore, we should look instead at the vorticity budget as it appears directly in the truncated equation 1.23, which is indeed conserved. To construct this, we see that equation 1.23 can be rewritten in more general terms as:

$$\dot{\mathbf{x}} = \underline{\underline{L}}\mathbf{x} + \mathbf{x}^* + \sum_{i,j} x_i x_j \mathbf{c}_{ij} \tag{1.25}$$

This simply splits the system of equations into a linear component, a constant forced component and a set of quadratic nonlinearities, where i and j index over the six resolved modes. If in analogy to before we decompose into a time mean flow $\overline{\mathbf{x}}$ and an anomaly $\mathbf{x}'$, then time average equation 1.25, we get:

$$\overline{\dot{\mathbf{x}}'} \equiv 0 = \underline{\underline{L}}\overline{\mathbf{x}} + \mathbf{x}^* + \sum_{i,j} \overline{x_i x_j} \mathbf{c}_{ij} + \sum_{i,j} \overline{x_i' x_j'} \mathbf{c}_{ij} \tag{1.26}$$

where the four terms on the RHS are the mean linear, forced, mean nonlinear, and eddy nonlinear terms, respectively. Although this equation is formulated in terms of the wave amplitudes $\mathbf{x}$, if we remember that our basis functions were chosen as eigenfunctions of the Laplacian, we see that equation 1.26 represents the truncated model's equivalent

of vorticity conservation. An explicit vorticity formulation can be found through the substitution $x_i \rightarrow a_i^{-2} x_i$, where $\mathbf{a}^{-2} = [-4, -5, -5, -16, -17, -17]$. This vorticity weighting therefore gives added importance to the amplitude of the wave-2 modes, and will be revisited in section 1.7.

1.3 Dynamics of the model

Having derived the CdV equations, we now turn our attention to their dynamics. We integrate equation 1.23 using an Euler-forward scheme at a numerical timestep of dt=2e-4 model time units (MTU) and sample the system every 1 MTU, approximately corresponding to one day in dimensionful coordinates. The first-order Euler-forward integration scheme has been used, as it is equivalent to the stochastic Euler-Maruyama scheme used later on, in the limit where the noise amplitude goes to zero.

Figure 1.4 shows the various two-dimensional projections of the model attractor resulting from a long, well-converged integration (T=2,000,000 MTU). The attractor is clearly highly structured, with a number of strongly preferred trajectories through phase space (note should be taken of the logarithmic colour scheme, without which many of the less frequently visited states would not be visible). This low dimensional attractor shows no obviously separable phase space regions such as the twin lobes of the Lorenz '63 system, and we see that the trajectories are highly twisted together, producing a scroll-like geometry.

To simplify the visualisation of the phase space, we can project the attractor into the 2D subspace that explains the most variance; that is the space spanned by the first two empirical orthogonal function (EOFs) of the system, as shown in figure 1.5. This two-dimensional space captures 94% of the system's variability, and is visually similar to the $x_2 - x_3$ and $x_5 - x_6$ sub-spaces. The EOF patterns themselves represent meridionally asymmetric waves – slightly distorted versions of the x_5 and x_6 modes, with EOF1 approximately antiphased to EOF2. We therefore see that circular movements in the phase

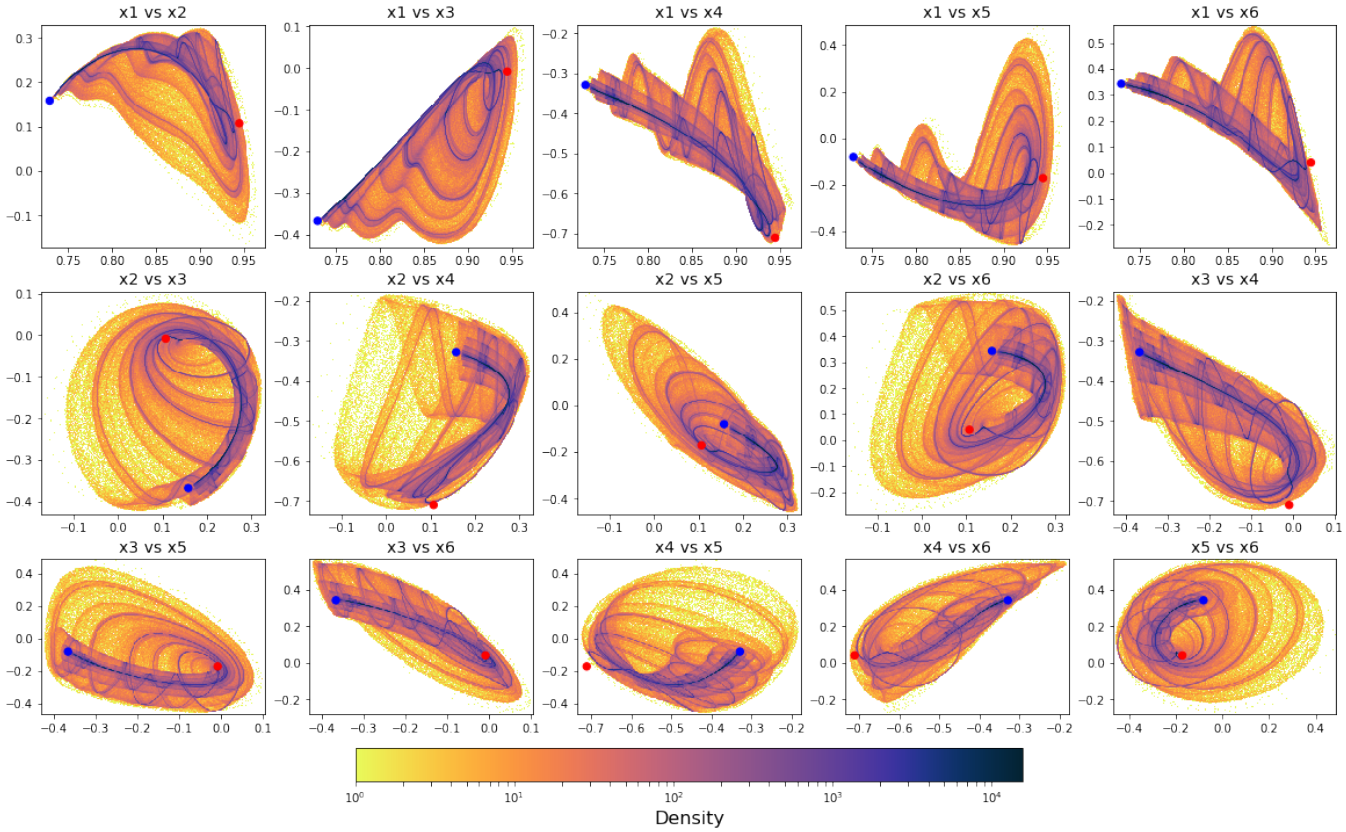


Figure 1.4: Log-density of the deterministic CdV79 attractor, showing the different two-dimensional projections of the 6-dimensional attractor. Computed using 200 bins on each axis. Blue and red dots show the position of two unstable equilibria corresponding to blocked and zonal flow configurations respectively.

space corresponds roughly to a freely propagating wave in the channel.

A representative sample of the model's temporal evolution is shown in figure 1.6. Multimodality is readily apparent, with predictable, slowly evolving dynamics and phase speeds close to zero corresponding to persistent blocked flow patterns, punctuated by highly chaotic, fast evolving flows that contain both zonal states and transient wave activity. This dynamical structure is therefore quite asymmetrical, unlike the Lorenz '63 or '84 systems whose regimes are essentially equivalent to eachother.

1.3.1 Bifurcations in Charney deVore

Some of the dynamical systems terminology less frequently used in atmospheric science is informally defined in appendix A.1.

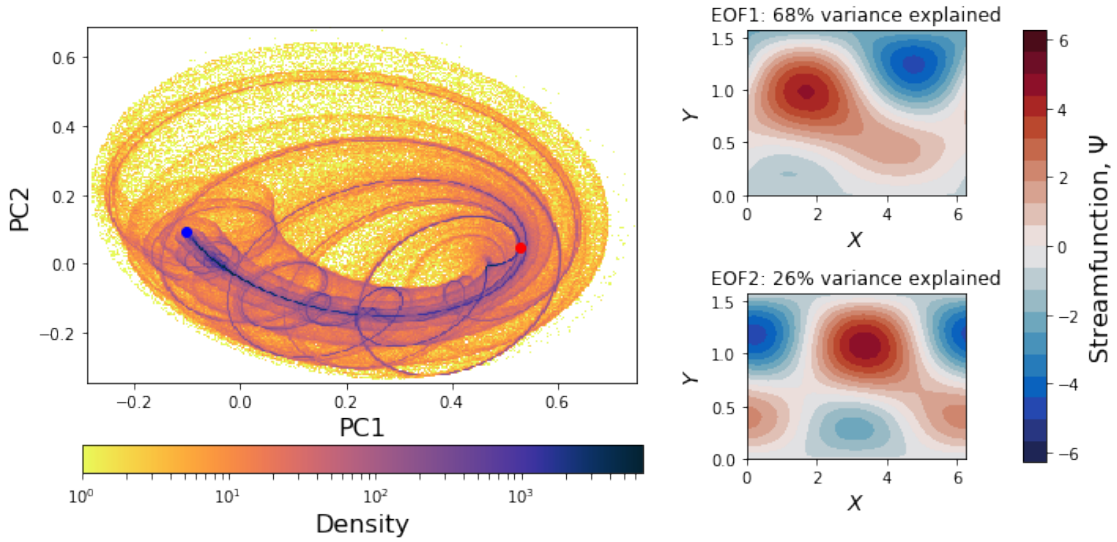


Figure 1.5: Left: Projection of the deterministic attractor into the space of the leading EOFs, capturing 92% of the total variance. Right: The flow patterns corresponding to the first and second EOFs of the system. Blue and red dots show the positions of fixed points corresponding to blocked and zonal flows respectively.

While we are not conducting a detailed bifurcation analysis in this thesis (as done in considerable detail already for the Charney deVore system in C04 and [De Swart, 1988]), our understanding of the system's dynamics will be enriched by a brief consideration of some of the relevant bifurcations and model invariants that give rise to the regime structure. Bifurcations are borders in parameter space, dividing parameter sets that produce qualitatively different dynamics. Common bifurcations include the creation or destruction of invariant measures, or switches in the stability of an invariant measure.

Two different basic instability mechanisms are at play in the CdV system. The first is an orographically induced saddle-node bifurcation – a transition from one to multiple fixed-point equilibria – as identified in the original CdV79 paper. When the zonally symmetric wave-2 mode, x_4 (and its forcing, x_4^*) is set to 0, this is the only instability that can manifest. However, when the wave-2 modes can be non-zero then a barotropic instability can be induced, associated with a Hopf bifurcation – the transitioning of a stationary fixed point to a periodic limit cycle. This manifests as the onset of travelling zonal waves in the model domain. Both of these bifurcations are codimension 1, meaning that they exist

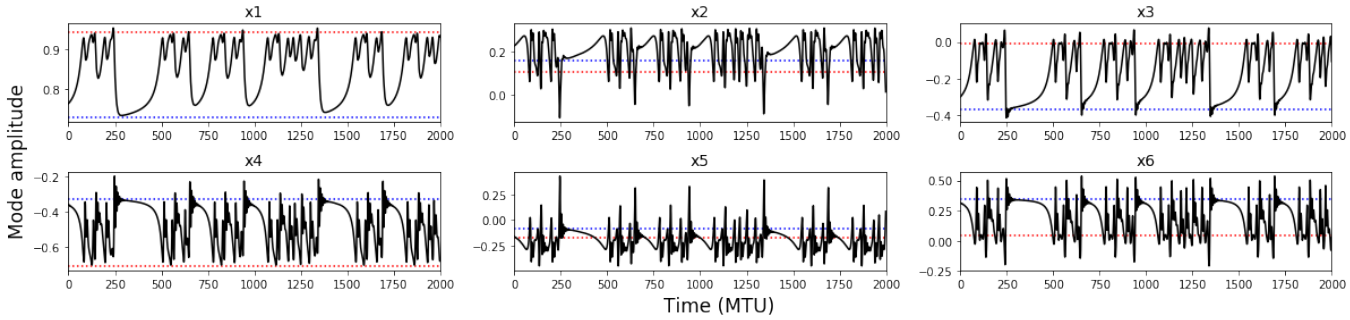


Figure 1.6: A 2000 MTU integration of the CdV79 system, showing the evolution of each mode amplitude. The dynamics are chaotic but weakly so, showing clear quasiperiodic behaviour. Two regimes of behaviour can be seen; fast chaotic oscillations, divided by periods of slow evolution with irregular duration. The mode values at the blocked and zonal fixed points are shown with dashed blue and red lines.

along a surface with one less dimension than the parameter space, and so can in general be found by tuning 1 model parameter. Stable steady state and periodic wave solutions to equation 1.23 are common for arbitrary sets of model parameters (not shown), while chaotic – and especially multimodally chaotic – solutions are harder to find.

The model parameters chosen here are close to a parameter set where the barotropic Hopf and orographic saddle-node bifurcations merge into a single fold-Hopf bifurcation. There, one fixed point bifurcates into two fixed points and a limit cycle. This fold-Hopf point has codimension 2, meaning two model parameters must be tuned to obtain it. High codimension bifurcations seem to be of reduced relevance to the real world, representing exceptional states in parameter space. However, the fold-Hopf bifurcation is known to produce a range of lower codimension, and so more relevant, bifurcations at nearby parameter values, including paths to chaotic dynamics [Champneys and Kirk, 2004]. Most relevant to us is the generation of heteroclinic and homoclinic orbits, a type of bifurcation where looping orbits stretch from one fixed point to another, or from a fixed point back to itself, respectively. These orbits can lead to fully fledged chaotic dynamics and the creation of a strange attractor through intersection with an unstable toroidal limit cycle. This creates a 'whirlpool repeller' [Shilnikov et al., 1995] and exactly the scrolling structure seen in figure 1.5. At the parameter values chosen, our strange attractor closely shadows

collapsed heteroclinic orbits that connected the fixed points.

The first key point here is that the global dynamics of the system can be influenced by the features of unstable fixed points that do not necessarily have to be near in phase space to the attractor. The second is that the multimodal dynamics we explore in this system are confined to a relatively small area of parameter space: for the low-dimensionality of our system we pay the price of a certain fragility. The final point is that the structure of the attractor and the multimodal variability of the CdV system is rooted in the dynamics of unstable homoclinic orbits, with close approaches to the unstable fixed point leading to periods of slow evolution, and then rapid 'bursting' to chaotic, more turbulent flow. This has been seen in simple atmospheric models [Crommelin, 2002] including the Lorenz 84 model [Shilnikov et al., 1995], and in the dynamics of nonlinear circuits and brain neurons [Pusuluri and Shilnikov, 2018, Izhikevich, 2006].

We should emphasise that this type of regime behaviour – or intermittency as it is often termed in the dynamical systems literature – is of a fundamentally different origin and type to that of the archetypal Lorenz '63 system. In Lorenz '63, crisis-induced intermittency creates regime dynamics when the basin of attraction of two attractors merge together, triggering movement between two distinct regimes of fully chaotic behaviour. In the Charney deVore system, we find Pomeau-Maneville intermittency where the system experiences almost deterministic dynamics before drifting away and entering a transient chaotic regime. At an unpredictable later time, the system will once again return to the neighbourhood of the periodic orbit, and reenter the slow-evolving regime. For a recent discussion of the co-existence of predictable and chaotic flow states in atmospheric systems see [Shen et al., 2021].

The two unstable fixed points that give structure to the CdV system are shown in phase space in figures 1.4 and 1.5, and the flow configurations they correspond to are shown in figure 1.7. These are broadly analogues of the stable fixed points found in CdV79, one blocked and one more zonally symmetric. The attractor tends to lie between these two

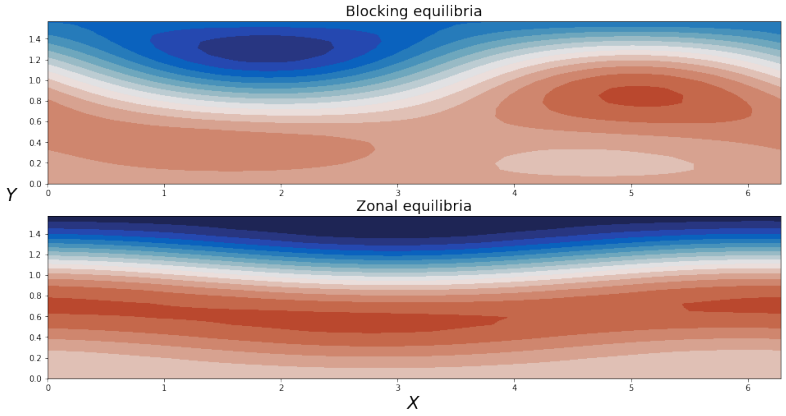


Figure 1.7: The flow patterns corresponding to the two unstable equilibria states marked in figures 1.4 and 1.5. The blocking state has a downstream ridge and a reduced meridional gradient, implying weaker zonal winds. The zonal state is almost – but not entirely – zonally symmetric. Neither of these states can be realised in a numerical integration, but they nevertheless influence the chaotic dynamics of the system.

points, shadowing unstable homoclinic orbits that bring the system into the blocked or zonal neighbourhood of the corresponding fixed point.

1.3.2 Stability properties of Charney-deVore

We can quantify the asymmetrical dynamics of the CdV system with linear stability analysis. By finding a linear approximation to the model's tendency equation at a given point in phase space, we can understand whether infinitesimally perturbed trajectories will exponentially converge or diverge in different directions. Calculating the linearised tendency at a point $\mathbf{x}_0$ is simple in the CdV system, as most of our terms are already linear and only the quadratic nonlinearities must be addressed. These are easily dealt with by letting $ax_ix_j \rightarrow a(x_{i,0}x_j + x_{j,0}x_i)$. This linearised tendency can now be written as a matrix equation, and the eigenvalues of that matrix are the local Lyapunov exponents, λ_i, evaluated at the chosen point $\mathbf{x_0}$.

Positive Lyapunov exponents lead to exponential divergence, while negative Lyapunov exponents lead to convergence. Figure 1.8a) shows the largest Lyapunov exponent calculated at each point on the model attractor, giving a measure of instability growth. We see that, almost everywhere, this leading exponent is positive, confirming that this is indeed

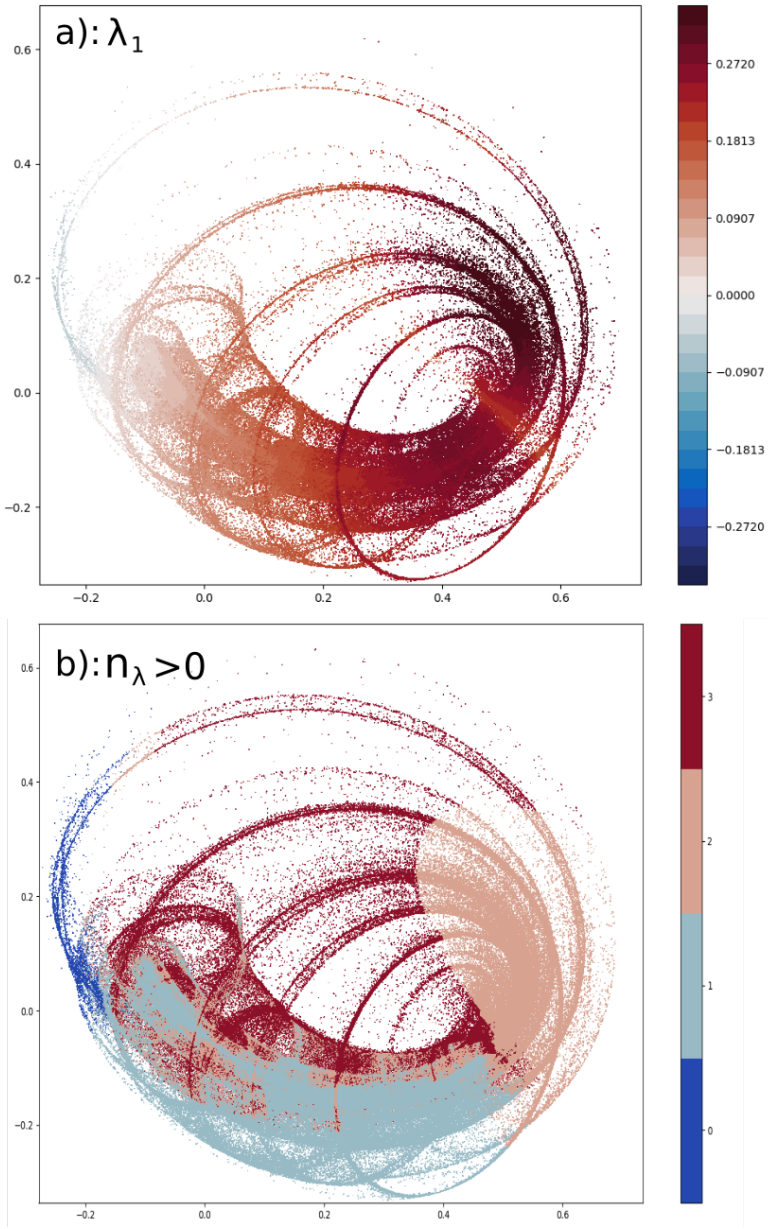


Figure 1.8: a) The real part of the leading local Lyapunov exponent and b) the number of positive local Lyapunov exponents, calculated at each point along the deterministic model attractor, projected into the EOF1-EOF2 plane. Strong variations in the degree of stability and number of unstable dimensions can be seen, indicating a non-hyperbolic system.

a chaotic system: trajectories diverge exponentially. Intuitively, the largest exponents are found where the different preferred orbits rapidly separate, at large values of PC1.

Figure 1.8b), meanwhile, shows the total number of unstable dimensions, which changes across the attractor, moving between zero and three positive Lyapunov exponents. The general chaotically evolving flow has two or three unstable dimensions, but this changes in the vicinity of the fixed points and in the portion of phase space between them for lower values of EOF2. We hypothesise that this low dimensionality of the unstable manifold is due to a close shadowing of the remnant heteroclinic connection, but the global bifurcation analysis required to concretely prove this conjecture is well beyond the scope of this thesis, and indeed the author's skills.

This variability in unstable dimension means we might expect to see quite substantial flow-dependent predictability in this system, a topic we will pick up in section 1.5. As is well-known to mathematicians, a system with a changing number of unstable dimensions must be non-hyperbolic [Araújo and Viana, 2009]. Uniformly hyperbolic systems have no Lyapunov exponents equal to zero at any point in phase space, implying that dimensions cannot switch from being stable or unstable. This is of interest to us because there is no proof that nonhyperbolic dynamical systems must be structurally stable; that is, a small perturbation in the parameters of a nonhyperbolic system may cause discontinuous changes in the invariant measures of the system, and approximate solutions to the true equations may not shadow any true solution of the system [Dawson et al., 1994]. While this nonhyperbolicity property is a common feature of many physical flows, the existence and successful application of numerical weather forecasts indicate that, in practice, approximate solutions to such flows seem to do just fine. Nonetheless, it is important to be aware of the challenges this property introduces to our underlying system, a property which is not shared by (for example) the singularly-hyperbolic Lorenz '63 system.

1.3.3 Stochastic forcing in Charney deVore

Now that we are well acquainted with the deterministic system, we consider how it is modified by stochastic forcing. Following the procedure of K14, we introduce an additive white noise term into the CdV model, so that:

$$\dot{x}_i = \dot{x}_{i,\text{Deterministic}} + \sigma\xi_i \tag{1.27}$$

where σ is the noise amplitude, setting the standard deviation of the applied stochastic forcing, and ξ_i is a random variable drawn from a normal distribution of mean 0 and standard deviation 1.

We integrate this stochastic CdV equation, now using the Euler-Maruyama scheme, for 200,000 MTU at a range of values of σ, in order to build up estimates of the attractor in each case. While our system is no longer deterministic, and so it is not formally accurate to speak of a model attractor, we still continue to do so as a shorthand for the well-converged probability density function of the model.

Figure 1.9 shows the impact of steadily increasing noise amplitude on the attractor structure, and shows a clear breakdown of the low-dimensional preferred trajectories for even the lowest noise amplitudes considered. As amplitude increases, the fundamental cyclical nature of the dynamics persists but probability density increasingly shifts to the blocking region, with a density maxima shifting closer and closer to the location of the unstable blocking fixed point.

In addition to altering the long-term structure of the attractor, the stochastic forcing has impacts on the temporal statistics of the system. This is well captured in figure 1.10, which shows Fourier transforms of the x_1 mode for a number of different noise amplitudes. The deterministic spectrum possesses a number of clear peaks, tied to the quasiperiodic nature of the dynamics, with each peak loosely associated to a preferred

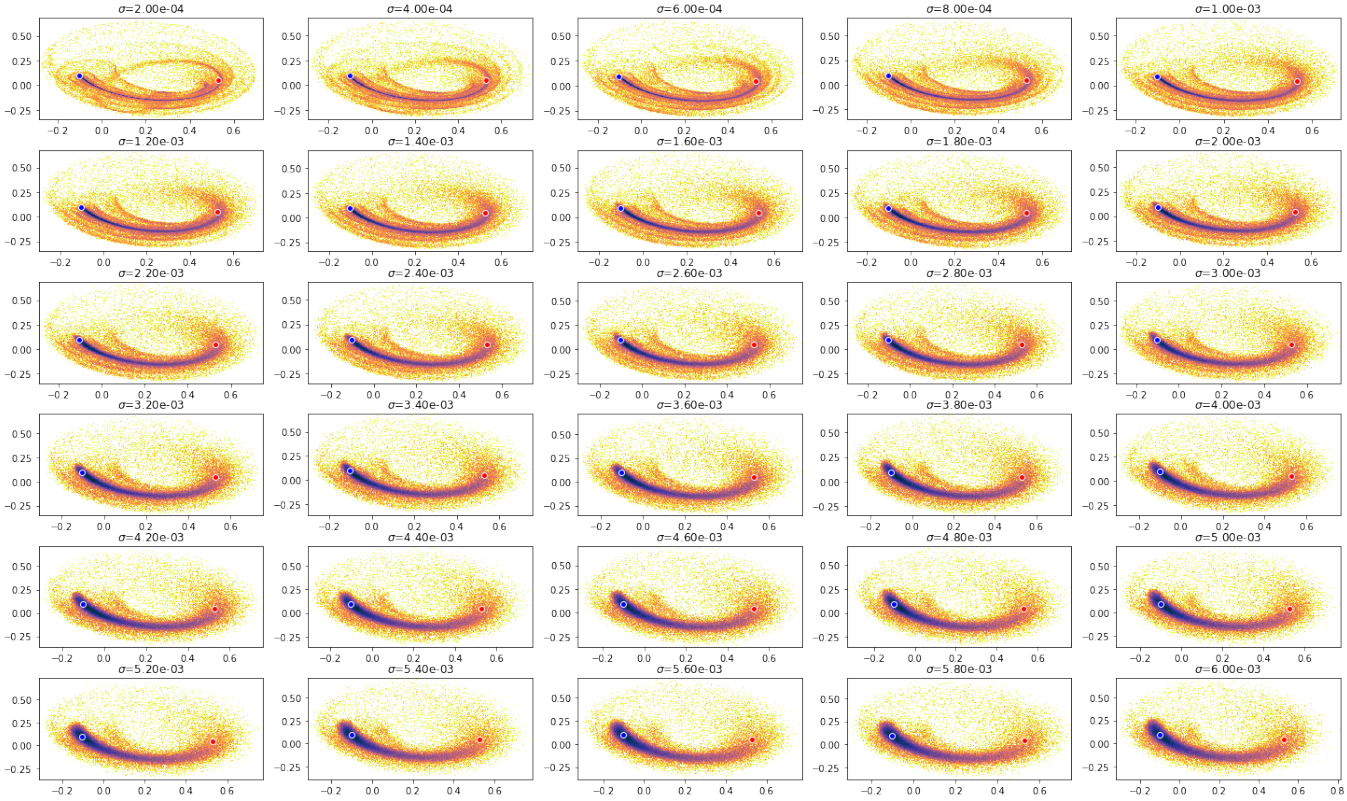


Figure 1.9: The impact of stochastic forcing on the model attractor, shown in the space of leading (deterministic) model EOFs. Each subplot shows the attractor for different noise amplitudes. The blocked and zonal fixed points are also shown as in figure 1.5. At large noise values, the attractor shifts towards the blocking fixed point.

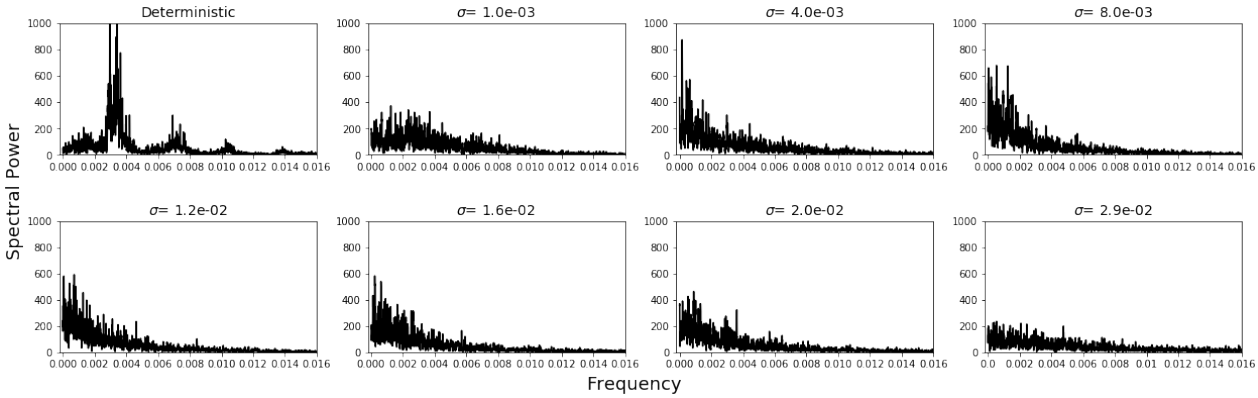


Figure 1.10: The power spectral density of the x_1 mode, for a range of stochastic forcing amplitudes, calculated from 200,000 MTU integrations.

trajectory through phase space. As would be expected, the application of low level stochastic forcing destroys this quasiperiodic structure, producing a much flatter spectra. However, as stochastic forcing increases up to $\sigma = 0.008$, increased power appears in the spectra at low frequencies – far more than in the deterministic case. This indicates that stochasticity is generating new, long-time scale behaviour in the system. As noise amplitude increases beyond this threshold, the spectrum flattens again, eventually reaching a flat white noise spectrum as the stochastic term completely overshadows the deterministic dynamics. Spectra are completely comparable for the other five modes (not shown).

This is in excellent agreement with the results of K14, which showed increasing regime persistence with noise amplitude that peaked at $\sigma = 0.008$ and then decreased. We show here that this is associated with a genuine slowing of the continuous modes visible in the Fourier spectrum, and that the stochastic persistence seen there was not simply an artefact of the regime identification approach used. With our confidence bolstered and an initial understanding of the model developed, we can move on to our own regime analysis.

1.4 Regimes in CdV

1.4.1 Regime classification

When studying a multimodal system, bulk metrics such as the mean or variance can be misleading and obscure rich dynamical structure. Therefore it is natural to use a regime framework as a way of partitioning phase space into dynamically distinct regions of interest. However, defining a metric or filter that isolates these regimes is in our case made more difficult by the complex structure of the attractor. In many clustering applications, the relatively straightforward K-means clustering algorithm is used, which we will discuss in more detail in chapter 2. However, K-means simply groups together nearby points, ignorant of temporal relationships, which is overly simplistic for a highly struc-

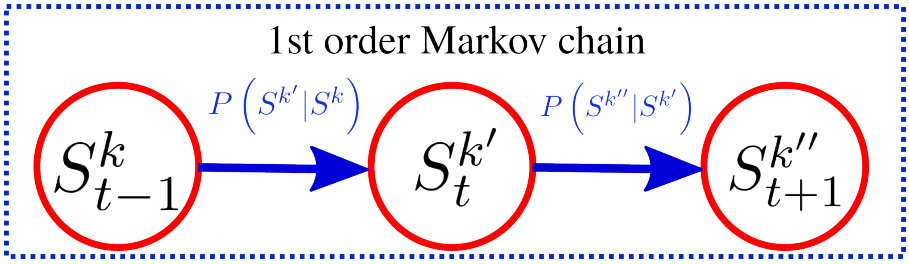


Figure 1.11: Graphical structure of a Markov chain, with conditionality indicated by arrows. S_t^k denotes that the system is in a state $S = k$ at time t. The probability of this occurring depends only on the previous state S_{t-1}.

tured attractor where orbits can pass close to each other but remain dynamically distinct. Therefore, following the approach of K14, we use an unsupervised machine learning technique, the hidden Markov model (HMM), to identify the qualitatively different dynamical regimes.

The HMM – and indeed often cluster analysis more broadly – makes use of the concept of a first-order Markov chain, a system discrete in time that can occupy one of a finite number of states S, and where the probability of occupying a particular state k at a time t depends only on the state at the previous time, $t - 1$. This is illustrated graphically in figure 1.11.

An HMM models an observed system by assuming it possesses a pre-specified number of states K, where movement between states can be modelled as a first order Markov chain. Each state is then associated with a phase space distribution (in our case a multivariate Gaussian distribution) which the model considers to have been sampled in order to generate the observational time series (see figure 1.12). The observations are related to the states only through this intermediary distribution, and so the states themselves are 'hidden', giving the algorithm its name. Fitting an HMM therefore involves using the observations to infer both the hidden states and the parameters of their associated distribution functions (here, the mean and covariance of each Gaussian distribution). Appendix A.2 includes a full exposition of how this is achieved algorithmically. Because this model makes use of temporal information (i.e. it knows which data points come after

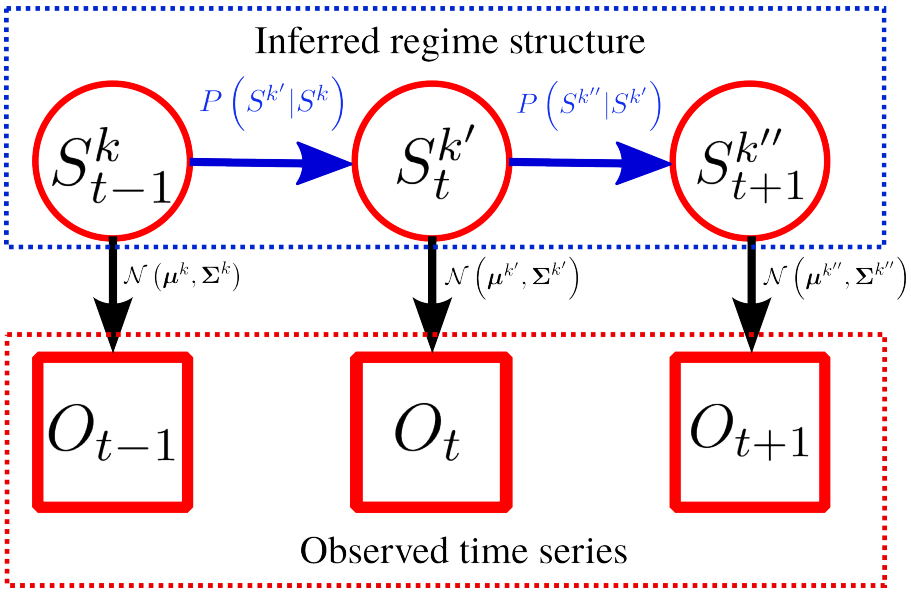


Figure 1.12: Graphical structure of a hidden Markov model. This consists of a 1st order Markov chain (enclosed in the blue box) as in figure 1.11 describing the evolution of an unobserved state variable S. Each hidden state, $S = k$, is equipped with a Gaussian probability distribution, $\mathcal{N}(\mu^k, \Sigma^k)$. The observations at each timestep, O_t, are drawn randomly from the distribution corresponding to the current hidden state. Fitting a hidden Markov model to data involves inferring the hidden states, and their distributions, from observations.

which) it is better able to handle the complex spatial structure of the CdV system.

In K14, the HMM is applied only to the x_1-x_4 subspace. We apply the algorithm to the full six-dimensional space in order to ensure that we account for any important variability that might be contained in the wave modes. As for most non-hierarchical clustering methods, we must make an *a priori* choice of regime number. We use three regimes, again as in K14, for no more principled reason than it best isolates the known structure of the model discussed in section 1.3.

This trimodal regime assignment produces a blocking regime characterised by slow-evolving flow, high amplitude wave modes and weak zonal winds, a zonal regime with strong zonal winds, and a transitional regime that features the largest phase velocities, transient wave activity and is linked to transitions from the zonal state back into blocking.

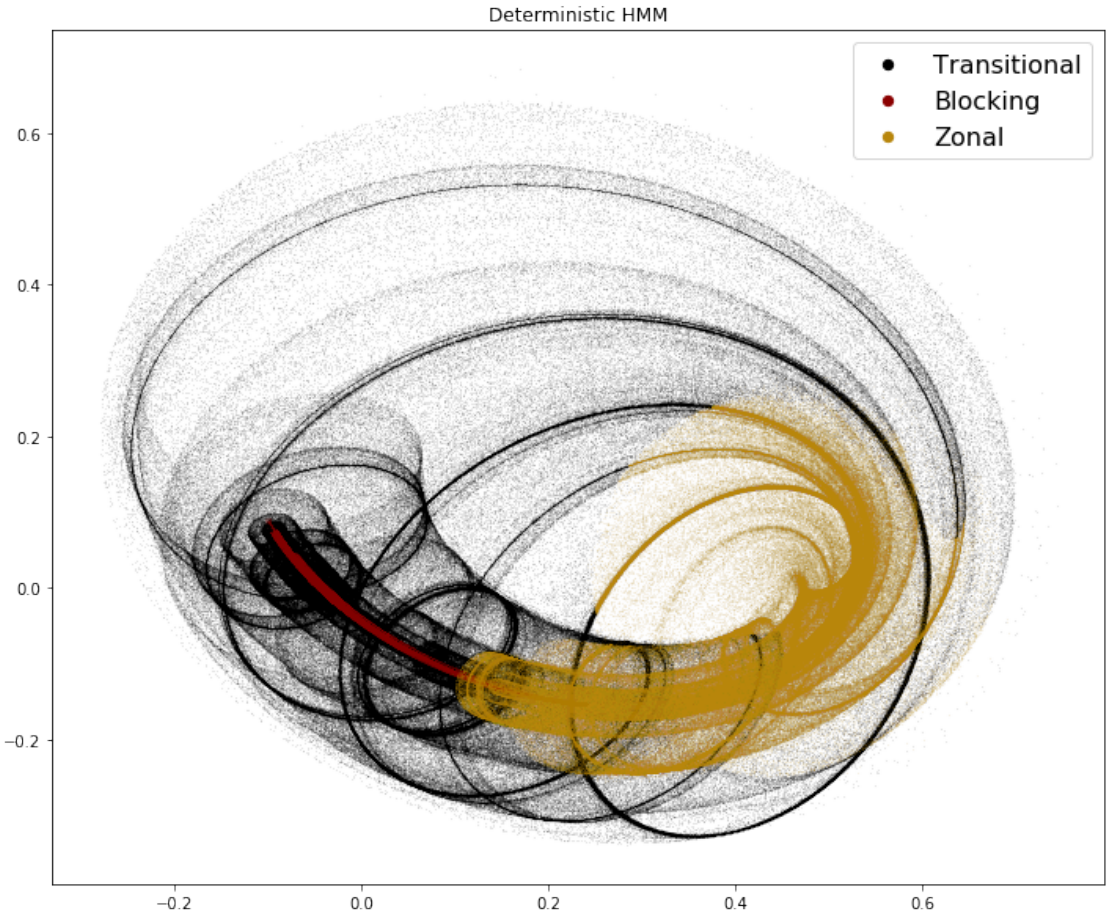


Figure 1.13: The assignment of points on the deterministic model attractor to the three hidden Markov model regimes projected into the EOF1-EOF2 plane.

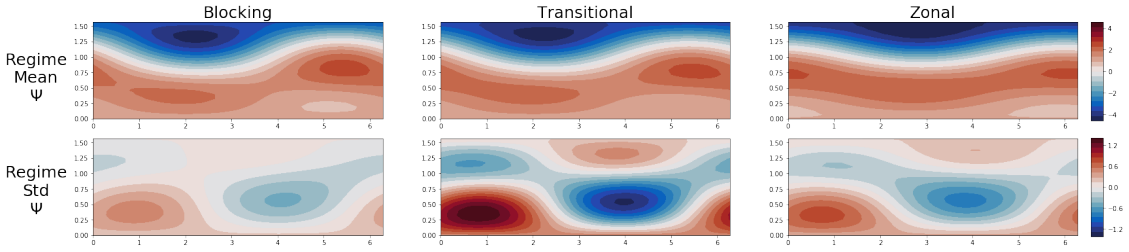


Figure 1.14: The mean (top) and standard deviation (bottom) of flow states assigned to each of the three hidden Markov model regimes.

Figure 1.13 shows the division of points on the deterministic attractor into the three regimes in the leading EOF space. The blocking regime is small in phase space extent, resembling a filament, and is focused on the tightly spiralled trajectories that start close to the blocking fixed point. It clearly separates from the swirling trajectories assigned to the transitional regime, which are seemingly close by in phase space (although we must remember this is a six-dimensional space and so any low dimensional projection distorts distance). The zonal state contains the very chaotic 'branching' of the attractor into different orbits, and contains the early parts of these orbits until their transition back to blocking becomes unavoidable. While it is not an exact match, a comparison of figures 1.13 and 1.8 shows that the blocking regime tends to correspond to flow states with one unstable dimension, and the zonal regime with two. The transitional regime is more diverse, but is generally the most chaotic, with three unstable dimensions at the majority of points. It was by no means guaranteed that even an approximate equivalence between unstable dimension and regime should manifest, so this is a pleasing result, as it indicates a physical significance to the regimes. This regime assignment also has some similarities to those found in [Strommen et al., 2021], which used persistent homology to identify regimes in the CdV system and found a persistent component similar to our blocking state.

The physical flow patterns corresponding to these regimes are presented in figure 1.14, with the average of wave amplitudes in each regime shown. There is a broad similarity between the blocking and zonal regimes and the equilibria shown in figure 1.7, although

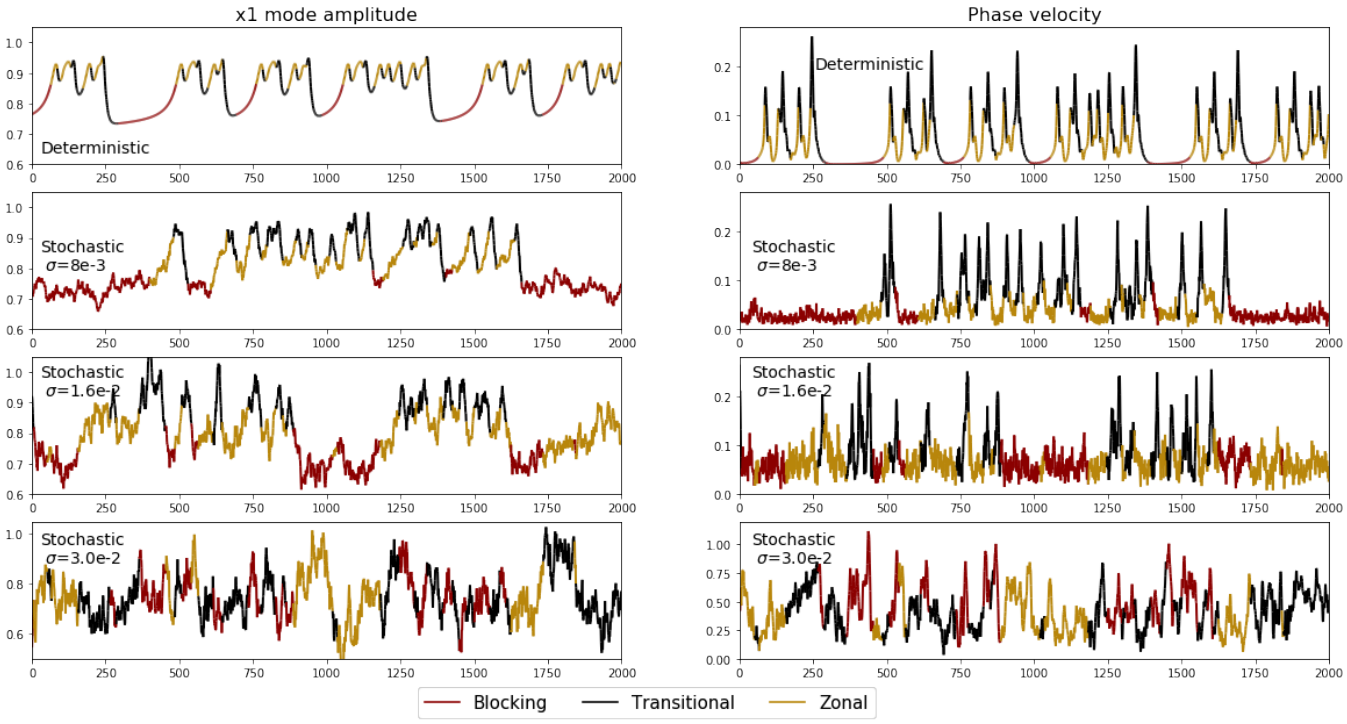


Figure 1.15: Left: Time evolution of the x_1 mode in deterministic and stochastic cases, with points coloured by their regime assignment. Right: The corresponding phase velocities.

the zonal regime shows a wavier flow. The transitional regime does not look radically different from the blocking regime, and we might question whether it is different at all. However, the transitional regime in particular is distinguished more by its fast evolution, and it contains points from a wide range of phase space, which makes the mean pattern less representative. This is captured by the standard deviations of these flow states, with much higher variability seen within the transitional regime. The reader who is still unconvinced, or seeks greater intuition for the behaviour of the model, should check the short video of the flow evolution found here: joshdorrington.github.io/cdv_animation/

For the stochastic simulations, the same approach is used, with a different HMM trained for each σ value. Attempts to simply assign datapoints in the stochastic runs to states in the deterministically-trained HMM were unsatisfying, with some regimes containing no points at all at higher noise amplitudes, due to mean state drift in the higher PCs.

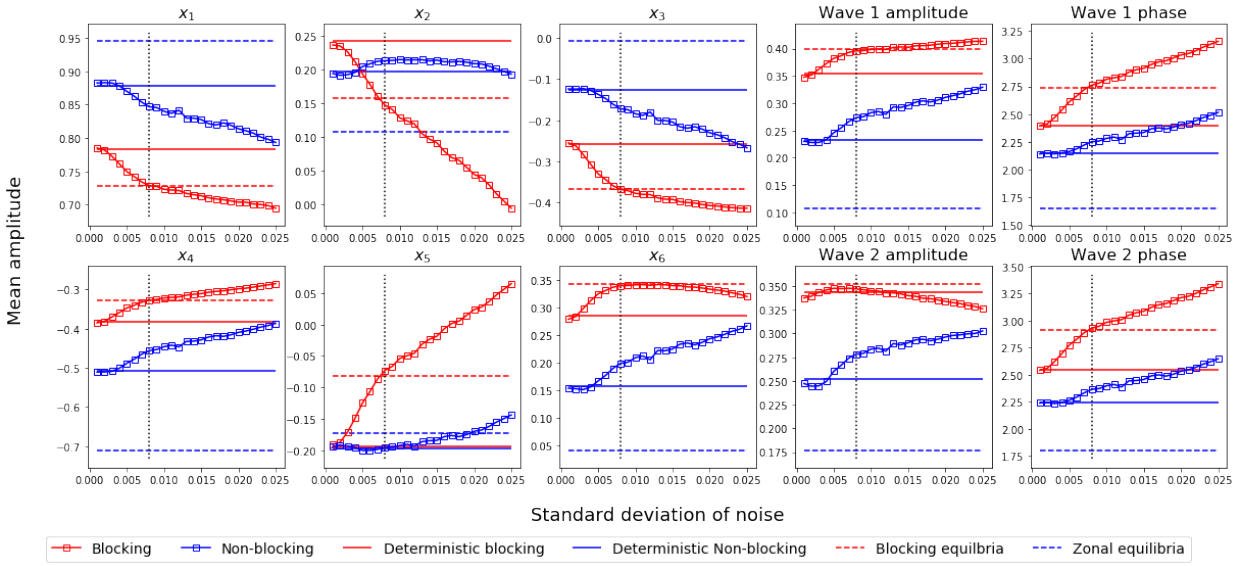


Figure 1.16: The changing mean state of each mode of the CdV system as stochastic forcing is increased, estimated from a 200,000 MTU run in each case, and separated out into blocked and unblocked states as assigned by the HMM.

We want to know over what range of stochastic forcing the fundamental characteristics of the flow states remain robust; that is, we still have clear chaotic zonal and persistent blocking type flows. There will, of course, be some noise threshold above which the chaotic dynamics we are interested in will have no relevance, playing second fiddle to the stochastic term, and the assignment of regimes will become meaningless. Figure 1.15 shows a snippet of the evolution of the CdV system at four different noise amplitudes, showing both the evolution of the zonally symmetric x_1 mode and the phase velocity, $v_t^2 = \frac{1}{4}\sum_{i=1}^{6}(x_{i,t+1} - x_{i,t-1})^2$. We see that the characteristic of the regimes are broadly preserved even up to σ=1.6e-2: the blocking state has low x_1 values and low phase velocities, the zonal state has high x_1 values and relatively low phase velocities, while the transitional regime has high x_1 values and very high phase velocity. However, by the time σ=3.0e-2, it is apparent that there is no meaningful regime structure remaining, and the phase velocity never drops below 0.2, indicating a loss of quasistationarity.

In addition to the temporal variability, we should also consider the mean state changes. In figure 1.16 we show the mean mode values using a bimodal decomposition, grouping the transitional and zonal states together as non-blocking states. This simplifies our analysis

and focuses attention on the most dynamically interesting behaviour; the shift between predictable and chaotic dynamics. We have also marked the values of the two relevant fixed points, and included the mean wave amplitudes and phases defined as:

$$\text{wave-1 amplitude} = \sqrt{x_2^2 + x_3^2}$$

$$\text{wave-2 amplitude} = \sqrt{x_5^2 + x_6^2}$$

$$\text{wave-1 phase} = \tan^{-1} \frac{x_2}{x_3}$$

$$\text{wave-2 phase} = \tan^{-1} \frac{x_5}{x_6}$$

which provide an alternate, more interpretable perspective on the changing wave dynamics. As the stochastic forcing increases, the mean value of each mode drifts. For the zonally symmetric modes, this is essentially a monotonic drift to lower absolute amplitude as stochastic forcing increases, with no difference in trend between blocked and nonblocked states. For the wave modes there is a trend to increasing phase, shifting the waves upstream. At peak persistence, the mean value of all modes during blocking passes close to the equilibria state, indicating a likely link between the two, which we explore in more detail in section 1.7. We also see that for very large noise amplitudes, the "nonblocking" state has actually drifted very close to our initial deterministic blocking state, and we see that the paradigm we have explored in the previous sections is no longer relevant.

To summarise, the regime structure is fundamentally the same in stochastic runs with moderate amplitudes, including the domain of peak persistence, but for larger forcings, $\sigma \gtrsim 0.016$, we enter the less interesting noise-dominated scenario.

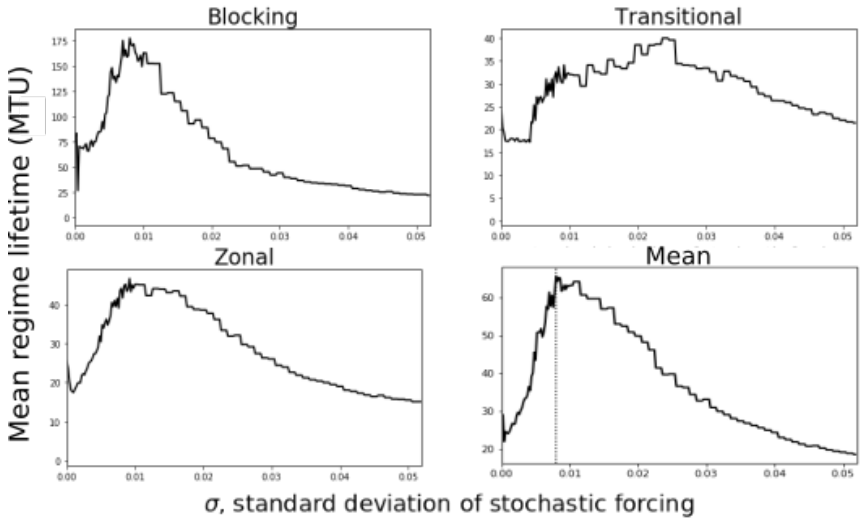


Figure 1.17: The mean regime lifetime for increasing σ calculated from a 200,000 MTU simulation in each case, for each of the HMM regimes, and their occurrence weighted mean. Vertical dotted line in the mean plot shows the value of peak persistence, $\sigma = 0.008$

1.4.2 Impact of stochasticity on regimes

Since the HMM regime framework isolates the dynamically distinct regions of the attractor and retains the same qualitative regime behaviour for moderate noise forcings, it provides a sensible way to analyse the amplified low frequency variability seen in figure 1.10. In figure 1.17, which shows the mean lifetime for each regime, we see that the most pronounced impact of stochasticity is present in the blocking regime. The zonal and transitional regimes show notable increases in persistence, with an approximate doubling of the average lifetime, peaking at σ=0.010 for the zonal regime, and σ=0.026 for the transitional regime, before declining as amplitude increases further. However, the blocking regime shows an almost trebling of regime lifetime, peaking much more sharply, and at a lower forcing amplitude of σ=0.008. Averaging over the three regimes, the overall persistence peak is at σ=0.008, just as for the blocking regime. This is perfectly in agreement with the results of K14, showing no sensitivity of the results to the inclusion of all six wave modes in the HMM algorithm.

We can gain more insight by looking at the full regime lifetime distribution. However, as

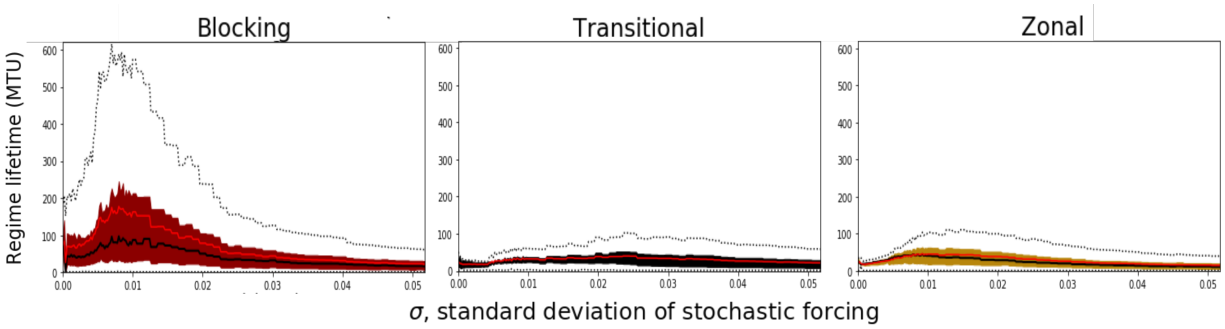


Figure 1.18: The distribution of blocking regime lifetimes, as a function of stochastic forcing amplitude. Shaded regions mark the interquartile range, with the thick black line showing the median, and the red line the mean. Dotted black lines show the 0th and 95th percentiles of the distribution.

in figure 1.18. The blocking regime shows a substantial skewing of the lifetime distribution as the peak noise amplitude is reached, with the 95th percentile of lifetime reaching over 600 MTU. While the other regimes do show a general increased variance in lifetime as we might expect from a stochastically forced process, the generation of a very fat-tailed distribution is unique to the blocking regime. This is an interesting result both because it gives us a hint that a stochastic hitting-time process may be at play as we will pick up in section 1.7, but also because it has been observed that blocking events in the real atmosphere have skewed lifetimes which are not always reproduced in current models [Masato et al., 2009].

1.5 Flow dependent predictability in CdV

We have shown that stochastic forcing, as long as its magnitude is not too high, can amplify low frequency variability in the CdV system, and that this results in increased regime persistence. Low frequency variability is often viewed as indicating possible long time-scale predictability, and as such this is where much of the interest in real-world regimes comes from. However it is not always the case that the persistence equates to predictability, and we might especially question this implication in our simple model where we inject purely random forcing so directly into the tendency equation. Therefore we directly examine the impact of stochastic forcing on the predictability of the CdV

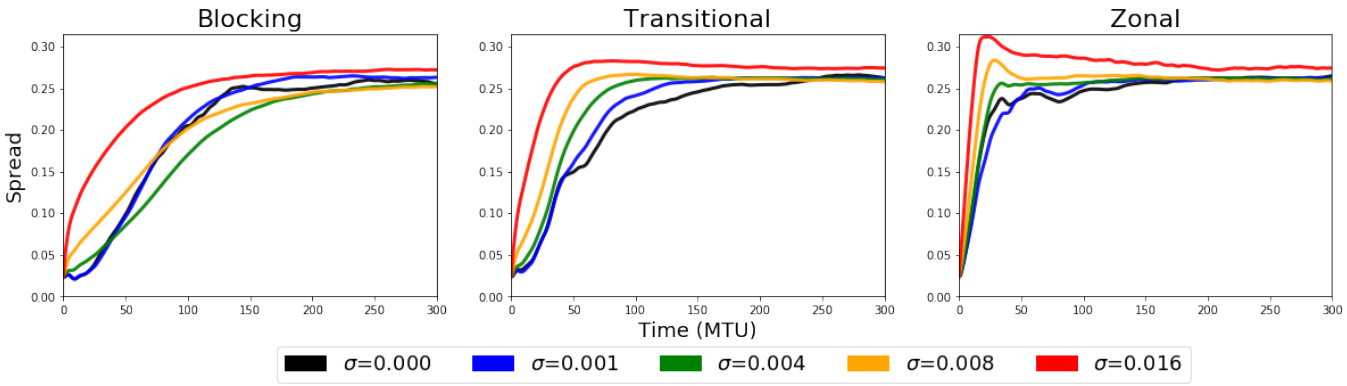


Figure 1.19: Ensemble spread as a function of lead time for simulations with different noise amplitudes, separated by the regime of the initial condition. Thick lines indicate the mean spread, over initial conditions.

system, first in terms of the impact on ensemble spread, and then in terms of the coarse grained regime predictability, to see whether this does indeed hold.

1.5.1 Ensemble predictability

In order to assess ensemble spread, integrations were performed using 5 noise amplitudes σ=[0.0,1e-3,4e-3,8e-3,1.6e-2]. In each case 5000 initial conditions along the attractor were sampled and initial gaussian perturbations of standard deviation 1e-3 were applied to each variable, to generate 50 ensemble members for each initial condition. These were then integrated forward with the corresponding noise amplitude for 300 MTU. This allows us to assess the divergence of initial conditions under a perfect model assumption. An extra 5000 initial conditions were run for the deterministic case to confirm that the spread distributions were well converged (not shown).

Figure 1.19 shows the resulting ensemble spread in each of the three hidden Markov model regimes. In the transitional and zonal regimes the relationship is very straightforward: the modest increases in regime persistence seen do not translate to increased predictability, instead increasing noise amplitude increased the growth of ensemble spread. In the blocking regime, the picture is similar, with the ensemble spread in the deterministic runs considerably lower than in any of the stochastic runs, and at lead times less than

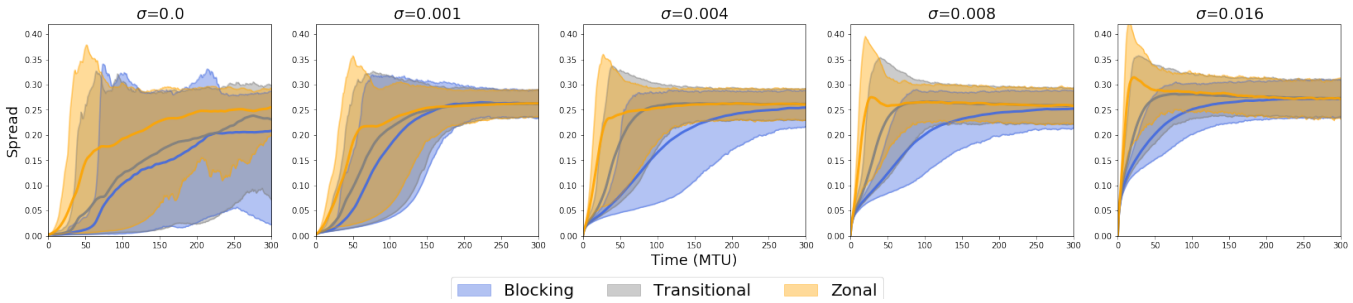


Figure 1.20: Ensemble spread as a function of lead time, separated out by simulation. Thick lines show the mean, while shaded regions show the interquartile range.

70 MTU we once again see that more strongly forced runs have larger spreads. However at longer leads the σ=4e-3 run has lower spread than the σ=1e-3 run, as does the σ=8e-3 run after 120 MTU. A selected number of extended simulations run to 1000 MTU show that for lead times greater than 300 MTU all spread distributions gradually tend to nearly the same climatological level of 0.27 (not shown).

On short timescales the diffusive effects of noise dominate, acting to increase the ensemble spread, but the increased persistence of the blocking state can partially overcome this on intermediate timescales, although never to the extent of creating an actual net decrease in spread relative to the deterministic model. It is not entirely apparent why the blocking persistence peak seen at $\sigma = 8e - 3$ does not give the smallest spread in the stochastic runs. Conjecturally, it might be that the increasingly large region of phase space assigned to the blocking region at peak persistence – as it drifts towards the fixed point – allows for larger spread between ensemble members, all of which are still assigned to the blocking regime.

Figure 1.20 shows the same spread distributions arranged now to highlight the degree of flow dependent spread in each simulation. We see that in all cases the zonal regime shows the largest spread, and the blocking regime the least. This runs counter to the linear stability analysis shown in figure 1.8, where the transitional regime showed the largest number of unstable dimensions and the largest Lyapunov exponents, and so would be expected to have the highest spread. A close analysis of the spread from each model initial condi-

tion in the deterministic case (not shown) highlighted that the largest ensemble spreads were seen for zonal initial conditions that were already close to entering the transitional regime. Small perturbations to these borderline initial conditions could cause the trajectories to diverge much more significantly than perturbations to initial conditions already well into the transitional regime, and so will soon converge along different trajectories towards the blocking regime.

We also see that because of the asymmetrical action of stochastic forcing on the regimes, the predictability of the σ=4e-3 and 8e-3 systems becomes increasingly flow dependent. That is, the variability in spread based on initial state increases: while flows in all regimes become more dispersive, flows in the nonblocking regimes become much more dispersive.

1.5.2 Regime predictability

Perhaps given the obviously dispersive effect of random noise, and the low dimension of our model, it is not that surprising that stochastic regime persistence doesn't translate directly to decreased spread. However in practice a coarser measure of predictability might suffice and we might consider a flow predictable if we can qualitatively estimate its behaviour, that is its prevailing regime, even if the details of the flow are uncertain. The concept of transition matrix entropy, as presented in K14, formalises this concept.

To wit, we ask what is the probability of being in a given regime k' at a time t, given the system is in a regime, k at time t=0? Again, with this question we coarse-grain the system's dynamics completely, and consider predictability only of the regime state, not the precise values of the 6 modes. High regime predictability essentially tells us that the regimes possess preferred (equivalently, avoided) transitions, and so knowing the current state provides more information about future states than simply considering the climatological occurrence of each regime. Phrasing the key idea informally, while knowing the last roll on a roulette wheel was '24 Red' does nothing to help one predict

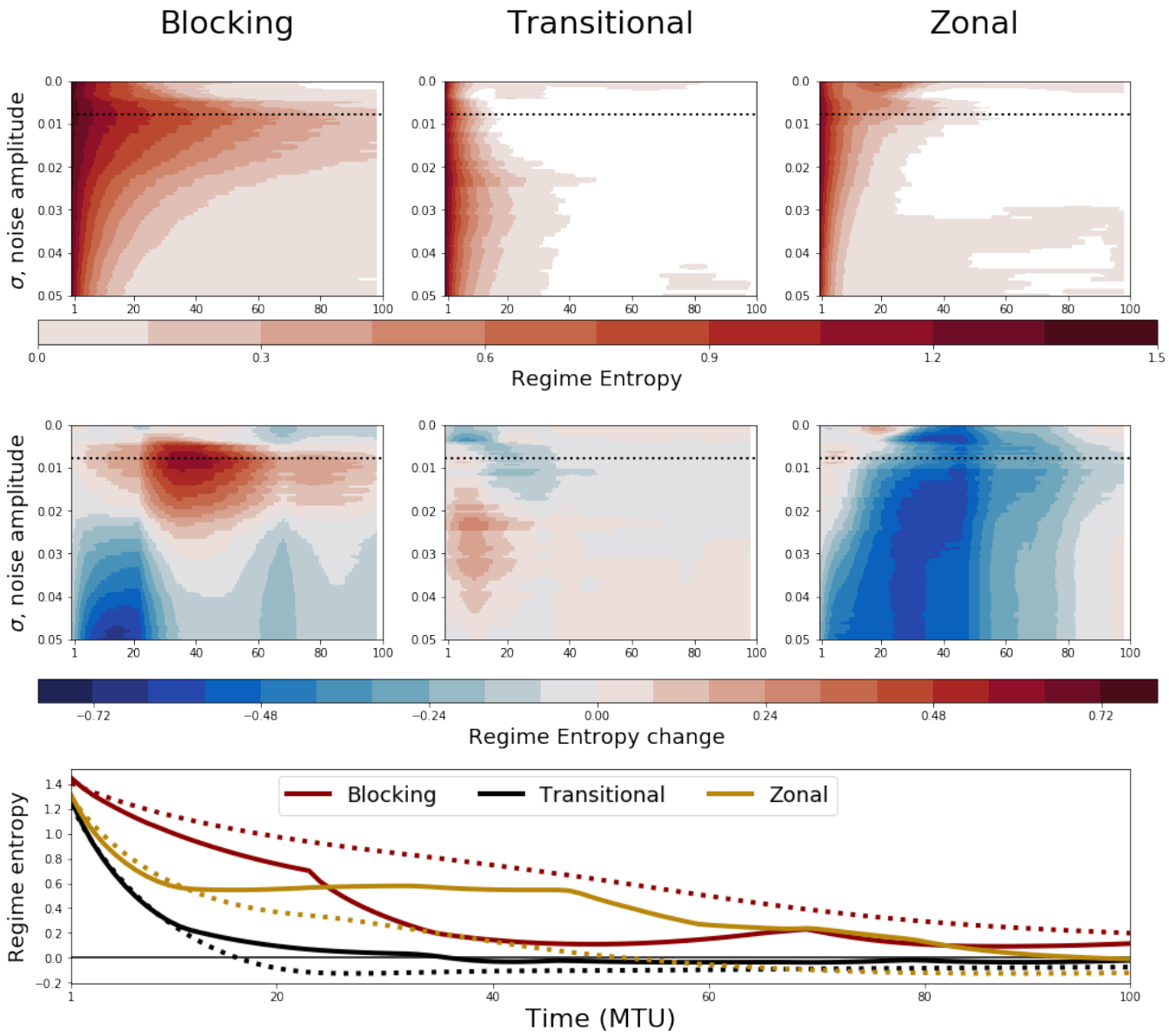


Figure 1.21: Top: The empirically estimated regime entropy as a function of lead time and noise amplitude for each regime. The dotted line indicates the noise amplitude at which peak blocking persistence was observed. Middle: As Top, but as an anomaly with respect to entropy of the deterministic simulation. Bottom: A horizontal cross section of the regime entropy heatmaps, showing the entropy of the deterministic system with solid lines, and the simulation with $\sigma = 0.008$ with dotted lines.

the next roll, knowing that 3 queens and a king have been dealt at the blackjack table carries substantial information about future hands.

The regime entropy, $\Omega_{k,t}$ is defined as:

$$\Omega_{k,t} = \sum_{k'} P_{t,k\to k'} \cdot \log_2\left(\frac{1}{P_{t,k\to k'}}\right) - P(k') \cdot \log_2\left(\frac{1}{P(k')}\right) \tag{1.28}$$

where $P_{t,k\to k'}$ is the probability of a system currently in regime k being in regime k' after t discrete time steps, and $P(k')$ is the climatological probability of being in the regime k', both estimated from a long integration of the system.

We estimate $\Omega_{k,t}$ for t=[1,2,...,100] MTU from a 200,000 MTU integration of the deterministic and each of the stochastic runs. These are shown in figure 1.21, with increasing lead time on the x-axis and increasing stochastic forcing amplitude on the y-axis.

The most striking feature is an increased regime entropy for the blocking regime between noise amplitudes in the range σ=[0.004,0.020] and centred at σ=0.008 – the amplitude of peak persistence seen previously. This is a direct result of the greatly increased regime lifetime. Given the extremely long lived blocking events seen in these stochastic model runs, persistence is an excellent forecast: knowing you are in the blocking regime means you will with high probability be in the blocking regime at a later time.

However there is not an exact equivalence between persistence and predictability as seen for the other regimes, which while slightly increased in persistence show a general decline in entropy. This is because the deterministic system has avoided transitions – from transitional to zonal and from zonal to blocking- which become more common in the stochastic case, and this acts to reduce the entropy overall. The sharp change in the blocking entropy anomaly at 23 MTU is due to a corresponding sharp decrease in the entropy of the deterministic system, as visible in the bottom subplot.

In summary, we see that stochasticity can increase the regime predictability of the block-

ing regime, primarily through the fairly direct effect that if a very persistent blocking is present at one time, it will probably be there at a later time. The nonblocking regimes generally have lower predictability, except for isolated pockets of quite small predictability increases at very small noise amplitudes. Further, for all regimes the ensemble spread is increased by increased noise amplitude, although again asymmetrically and not necessarily monotonically. We have seen that increasing the amplitude of noise can reduce spread at intermediate timescales through a regime mechanism, but not in this case to a level matching the deterministic system.

1.6 Stochastic suppression of convex instability – a geometric mechanism for increased regime persistence

We have seen clearly that there is stochastically amplified regime persistence in the CdV model, but *why* does stochasticity have this effect? And why asymmetrically; primarily in the blocking regime but less so in the others? Stochasticity is fundamentally a dispersive process, and so there must be some organisation of the flow that opposes that random dispersion. Here we present a very simple theoretical mechanism in which the ability of stochastic forcing to delay the growth of instabilities is key. Consider a one-dimensional system experiencing intermittent dynamics, such as in the schematic figure 1.22, moving between a highly chaotic state and a slowly evolving, predictable state. We consider that immediately upon entering the regime of slow evolving dynamics, the system is at a point $x = x_0$, and will stay in the quasi-stationary regime until it reaches a value $x = x_c$. We assume that x grows monotonically with the time t in this regime, so we can parameterise x in terms of t or vice versa: $f : x \to t$, $g : t \to x$.

While the CdV system has six dimensions, we recall from figure 1.8 that during the blocking events only one dimension was unstable, and so we can restrict our attention to that one unstable dimension, safe in the knowledge that perturbations into the stable

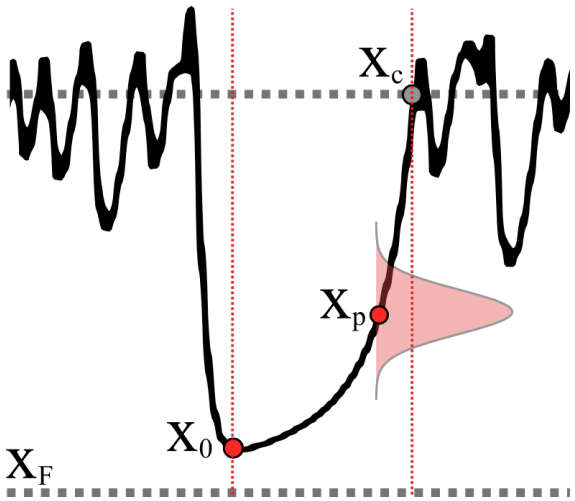


Figure 1.22: A schematic showing one variable of a chaotic process, x, which moves between periods of chaotic variability and periods of predictable, monotonic growth. The predictable regime is entered at a random position x_0, in the neighbourhood of an unstable fixed point x_F, and once it reaches an amplitude x_c it will re-enter the chaotic regime. At a position x_p, a random perturbation is applied to the x variable, assumed to have vanishing support outside the interval $[X_0, X_c]$.

manifold will be damped away.

To proceed in the clearest way let us now assume that x grows exponentially with time (although we soon show we can adopt a looser constraint):

$$x = f(t) = x_0 e^{\lambda t} \tag{1.29}$$

Here λ is a Lyapunov-like measure of how unstable the intermittent state is, and at the beginning of the intermittent period, $x(t = 0) = x_0$. We can easily find the inverse relation:

$$t = g(x) = \frac{1}{\lambda} \log\left(\frac{x}{x_0}\right) \tag{1.30}$$

What would be the impact of applying a single stochastic perturbation to this system at a time t_p (equivalently, when it has reached an amplitude x_p)? This stochastic perturbation would have the effect of moving the system from the point x_p to $x_p + \Delta x$. From our

bijection between x and t this perturbation also represents a corresponding movement from the point t_p to $t_p + \Delta t$; that is the perturbation moves us forward or backward in the lifetime of the intermittent period, shortening or lengthening it respectively. As a simple matter of notation we let $X := x_p + \Delta x$ and $T := t_p + \Delta t$, representing the state of the system after random perturbation.

Let us consider a Gaussian perturbation $\mathcal{N}(0, \sigma)$ in x. From the definition there will in expectation be no change in x:

$$\langle \Delta x \rangle = \langle X - x_p \rangle = 0 \tag{1.31}$$

But given that we have a nonlinear relationship between t and x, this symmetric perturbation introduces a non-zero change in the lifetime of the intermittent period:

$$\langle \Delta t \rangle = \langle T - t_p \rangle = \langle \frac{1}{\lambda} \log\left(\frac{X}{x_0}\right) - \frac{1}{\lambda} \log\left(\frac{x_p}{x_0}\right) \rangle = \langle \frac{1}{\lambda} \log\left(\frac{X}{x_p}\right) \rangle \tag{1.32}$$

Because we have made some nice choices of functions to work with, we can actually calculate this average shift in lifetime Δt explicitly. To start we use the definition of the expectation:

$$\langle \Delta t \rangle = \int_{x_0}^{x_c} \Delta t(X) \cdot p(X) dX \tag{1.33}$$

$$= \int_{x_0}^{x_c} \frac{1}{\lambda} \log\left(\frac{X}{x_p}\right) \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X - x_p}{\sigma}\right)^2} dX \tag{1.34}$$

We transform our variables to $w = \frac{\Delta x}{\sigma} = \frac{X - x_p}{\sigma}$, which then implies that $\frac{X}{x_p} = \frac{x_p + \sigma w}{x_p} = 1 + \frac{\sigma}{x_p} w$ and $dX = \sigma \cdot dw$. Meanwhile our limits of integration become $w_0 = \frac{x_0 - x_p}{\sigma} < 0$ and $w_c = \frac{x_c - x_p}{\sigma} > 0$.

If σ is taken to be small relative to a typical length scale of x in the intermittent period (that is the perturbation is vanishingly unlikely to force the system into the chaotic regime directly) then these limits will be very large in absolute magnitude. Given this and the

exponentially decaying tails of the Gaussian distribution, we approximate $w_0 \approx -\infty$, $w_c \approx +\infty$. This gives us:

$$\langle \Delta t \rangle = \frac{1}{\lambda\sqrt{2\pi}} \int_{-\infty}^{+\infty} \log\left(1 + \frac{\sigma}{x_p} w\right) \cdot e^{-\frac{w^2}{2}} dw \tag{1.35}$$

This isn't directly integrable, but we can Taylor expand the logarithm. Firstly we recognise that as the Gaussian distribution is even in w, then all odd terms in the Taylor expansion will vanish. Secondly we note that $\frac{\sigma}{x_p} w = \frac{\Delta x}{x_p}$ will be small if, as above, σ is small, so we can drop terms of 4th order and higher. Thus we end up with:

$$\langle \Delta t \rangle = \frac{-\sigma^2}{2\lambda x_p^2\sqrt{2\pi}} \int_{-\infty}^{+\infty} w^2 \cdot e^{-\frac{w^2}{2}} dw \tag{1.36}$$

Which from any book of identities we can find is simply:

$$\langle \Delta t \rangle = \frac{-\sigma^2}{2\lambda x_p^2\sqrt{2\pi}} \cdot \sqrt{2\pi} = \frac{-\sigma^2}{2\lambda x_p^2} \tag{1.37}$$

This quantity is negative definite and so on average our perturbation will move the system backwards in time; *the Gaussian perturbation acts to persist the slowly evolving regime*.

From this analytical expression we can see that larger σ corresponds to longer persistence as expected (remembering our overall weak noise approximation), as does small λ, indicating that nearly stable intermittent dynamics can be most substantially persisted. We also see smaller x_p leads to larger drift, which indicates that perturbations when the system has just entered an intermittent state are most impactful.

We chose a Gaussian perturbation and an exponential form for f, but these are not in fact central to the above result, merely allowing an analytical Δt to be obtained.

If we more generally let $x = f(t)$ be any convex (i.e. has positive curvature), monotonic function, then $t = g(x) = f^{-1}(x)$, and so g is therefore a concave function with negative

curvature. Further let $P(\Delta x)$ be any symmetric probability distribution function with vanishing tails. In this case, we have in perfect analogy to before:

$$\langle \Delta t \rangle = \int_{-\infty}^{\infty} G(\Delta x) \cdot P(\Delta x) d\Delta x \tag{1.38}$$

Where $G(\Delta X) := g(x) - g(x_p)$ is just a translation of g and so is still concave.

Because $P(\Delta x)$ is even, then if we Taylor expand G we will again lose all odd powers on integration and be left with:

$$\langle \Delta t \rangle = \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{1}{(2n)!} \frac{d^{2n} G}{d(\Delta X)^{2n}} \cdot P(\Delta x) d\Delta x \tag{1.39}$$

Again if ΔX is small, then we can drop high order terms:

$$\langle \Delta t \rangle = \int_{-\infty}^{\infty} \frac{1}{2} \frac{d^2 G}{d(\Delta X)^2} \cdot P(\Delta x) d\Delta x \tag{1.40}$$

Because G is concave then this derivative is negative and so we are left with

$$\langle \Delta t \rangle = -\int_{-\infty}^{\infty} \frac{1}{2} \left| \frac{d^2 G}{d(\Delta X)^2} \right| \cdot P(\Delta x) d\Delta x = -I \tag{1.41}$$

Where I is positive definite. Therefore any convex perturbation growth will be delayed (i.e. intermittency persisted) by small symmetric random perturbations, although we can not clearly extract the dependence on σ, λ or x_p in this general case. The case for a negatively oriented variable with concave curvature is precisely equivalent, if we simply multiply by -1.

Let us summarise: for a system with a monotonically growing, convex and one-dimensional instability, a symmetric stochastic perturbation will, in expectation, act to stabilise the system and so prolong intermittency, as long the perturbation is not so large as to im-

mediately move the system outside the neighbourhood of the intermittent dynamics. While we've considered only a single perturbation, this easily generalises to a continuous stochastic process, where now the $\langle \Delta t \rangle$ term we have calculated is a spatially dependent drift term. Our argument is similar in many ways to the consideration of a Poincaré recurrence or first-return map taken along the unstable manifold, where the stochastic variable plays the role of the chaotic mapping.

This mechanism, while abstract, is very general as it based solely on geometric arguments. Any low dimensional, convexly growing instability in the climate system should be susceptible to the same stochastic suppression. We also see why only the blocking regime showed dramatic persistence increases, as it is the only regime where the variables evolve with a constant sign of curvature, and where there is only one unstable dimension.

1.7 Vorticity dynamics of CdV

In the previous section we have identified how stochastic forcing produces prolonged persistent regime dynamics in the Charney deVore model, but in quite abstracted terms. One of the reasons to consider the CdV system was that while it is simple enough to study from a dynamical systems perspective, it also has enough physical grounding that we can hope to interpret our findings in the language of atmospheric phenomena. We take up this challenge in this section, aiming to frame the results we have obtained in a physically meaningful way.

We will do this by considering the vorticity budget of the CdV system, making use of equation 1.26. As mentioned in section 1.1, there is no agreed upon theory as to why atmospheric blocks can persist as long as they sometimes do. Fundamentally, a blocking anticyclone is a region of anomalous negative vorticity, and all else being equal that vorticity should be dissipated by friction. It was proposed in [Shutts, 1983] that blocking events might be maintained against dissipation by the injection of vorticity by eddy

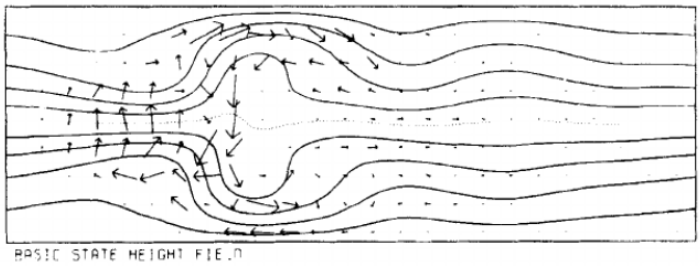


Figure 6(c). Time-mean streamlines with residual vorticity flux vectors, $(\overline{q'\mathbf{V}'})_\psi$, superimposed.

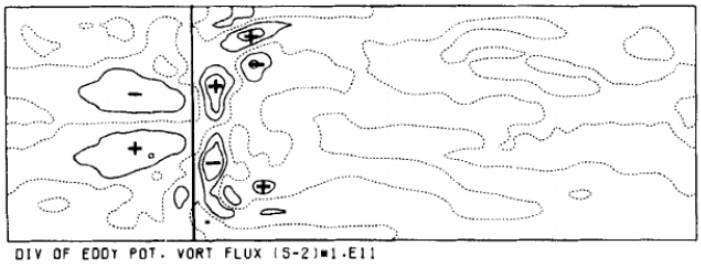


Figure 6(d). Eddy vorticity flux divergence, $\mathrm{div}(\overline{q'\mathbf{V}'})$. Contour interval: $5 \times 10^{-11}\,\mathrm{s}^{-2}$.

Figure 1.23: The proposed Shutts '83 eddy-straining mechanism for blocking persistence. Eddies stretching on the upstream flank of the block are proposed to inject vorticity into the blocking dipole in order to sustain the anomaly. Figure reproduced from [Shutts, 1983]

activity. Specifically, the deformation of eddies as they encounter the upstream side of the blocking anticyclone injects vorticity into the blocking region, and so amplifies the meridional vorticity dipole (see 1.23).

If this mechanism is indeed important for maintaining blocking events, then that means the lifetime of large-scale blocking events is contingent on small-scale, quasirandom variability. This in turn suggests that one of the reasons many atmospheric models struggle to produce adequately persistent blocks [Davini and D'Andrea, 2020] is that they are missing vital variability from sub-grid scale eddies. Might it be that the stochastically induced regime persistence seen in the CdV model is also a consequence of increased eddy activity? Of course here we must be aware that there is not a significant scale separation in our model, as we have specifically truncated to the large modes, and cannot resolve the spatially localised fluxes seen in figure 1.23. Nevertheless we can consider that the fast stochastic forcing applied to our model might be similar to the impact that resolved eddies might have on the large scale flow in a higher resolution model.

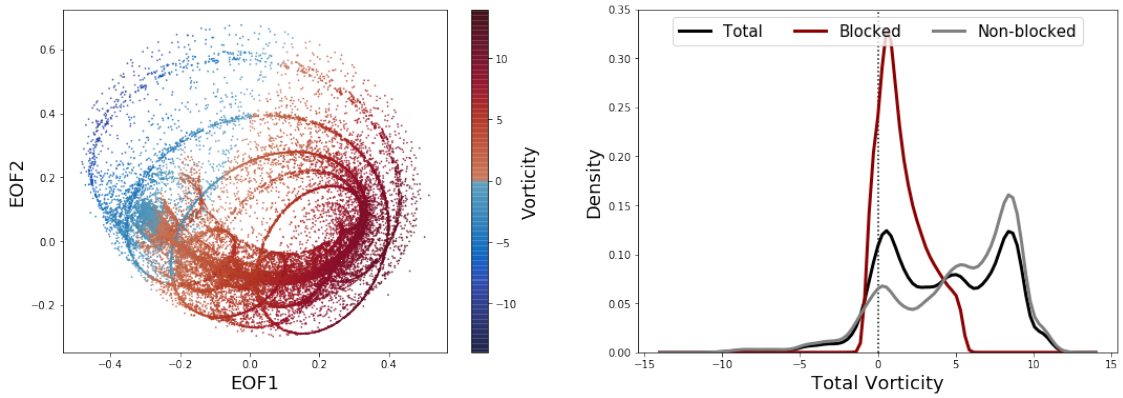


Figure 1.24: Left: The total vorticity of the resolved flow at different points along the deterministic attractor. Right: The distribution of resolved flow vorticities in different regions of the attractor.

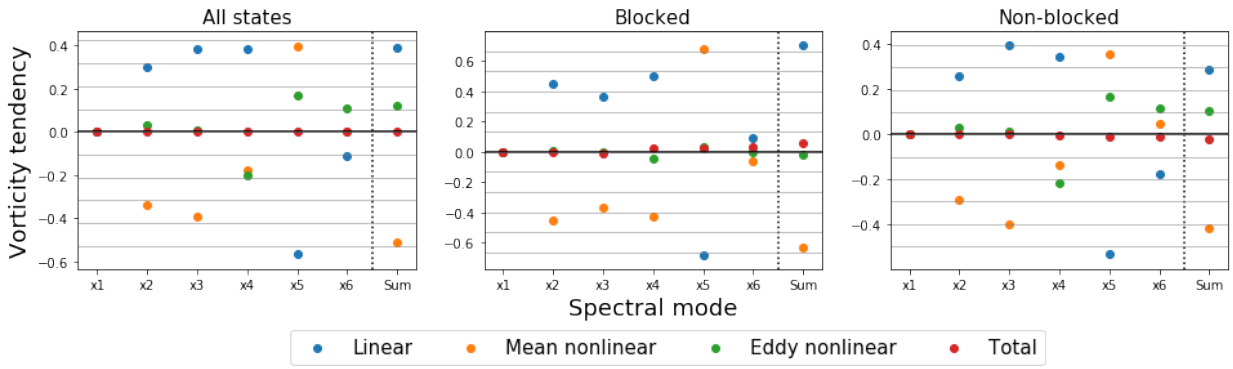


Figure 1.25: Left: The Reynolds balance of mean vorticity tendencies in each mode of the deterministic CdV simulations, including their total sum. The same is shown for only points assigned to the blocking regime (middle) and the non-blocking regimes (right) in order to show the flow dependent balances.

If we wish to assess whether this is indeed the case, the first thing to do is to understand the vorticity budget in the deterministic CdV model. Figure 1.24 shows that the total vorticity captured in the six resolved modes is predominately positive, and larger in the more zonal flow states. The exception is the transition into, and the onset of, the persistent blocking flows, where the net vorticity in the domain is negative. The dynamic is therefore a continuous vascillation between a low-vorticity blocking flow and a high vorticity zonal flow, and so the budget of vorticity tendencies will tell us much about the overall regime dynamics.

These tendencies are shown in figure 1.25, for the system as a whole, and also disaggregated by whether or not a flow state is assigned to the blocking regime by the HMM. In each case the flow has been Reynolds decomposed, with 'eddies' defined simply as temporal anomalies with respect to the mean over the subset of data being considered, i.e, anomaly terms in the blocked state are temporal anomalies with respect to the mean blocked state. For each spectral mode, we consider the average vorticity tendency deriving from the mean linear Coriolis, orographic and damping terms, the mean nonlinear advection term, and the eddy advection term, and their sum. The eddy linear term is of course identically zero.

Considering the model as a whole, there is no net gain of vorticity over time, just as we would expect for a stable model, with a balance between linear terms tending to inject vorticity, and mean advection terms removing it. This overall budget reflects the budgets of the larger scale wave , x_2 and x_3, and the zonally symmetric x_4 mode, with a reversed effect in the smaller x_5 and x_6 modes. The role of eddy terms is modest but important: this is an unsteady flow and eddy vorticity fluxes are needed to support the amplitudes of the zonal wave-2 modes. The blocked flow states show a net increase of vorticity, in agreement with figure 1.24, with linear forcings acting to drag the system out of blocking, opposed by mean advection. Eddies are less important here than in the non-blocked states, acting to prolong blocking (by opposing vorticity growth) but with a

negligible magnitude.

Since blocking persistence is our interest here, we ask what are the main physical terms maintaining the blocks in our model? We can now understand this as asking: which terms oppose vorticity growth? The x_1 mode during blocking is being constantly exhorted to amplify by the applied zonal forcing, and in the absence of nonlinearities this must be balanced by the orographic drag. We now see why orography is so crucial to the multimodality in this model, and have some insight into the saddle-node bifurcation discussed above: multiple (quasi-)equilibria manifest when drag from orography and the zonal forcing are able to (approximately) balance in different ways.

The other zonally symmetric mode, x_4, does experience non-linearities in the form of coupled wave interactions, and the mean impact of this interaction term balances 90% of the linear forcing on this mode, with eddy terms balancing the rest. At first glance, there is no special significance to this nonlinear term $\epsilon \cdot (x_2 x_6 - x_3 x_5)$, which balances the linear forcing. However if we suggestively rewrite this as:

$$\epsilon \cdot (x_2 x_6 - x_3 x_5) = \epsilon \cdot x_3 x_6 \cdot \left(\frac{x_2}{x_3} - \frac{x_5}{x_6} \right)$$

and note that in the blocking regime, x_2 and x_5 are generally of the same size or smaller than x_3 and x_6 respectively, then we can use the small angle formula to obtain

$$\begin{aligned} \epsilon \cdot (x_2 x_6 - x_3 x_5) &= \epsilon \cdot x_3 x_6 \cdot \left(\frac{x_2}{x_3} - \frac{x_5}{x_6} \right) \\ &\approx \epsilon \cdot x_3 x_6 \cdot \left(\arctan \frac{x_2}{x_3} - \arctan \frac{x_5}{x_6} \right) \\ &:= \epsilon x_3 x_6 \cdot (\theta_1 - \theta_2) = \epsilon x_3 x_6 \Delta\theta_{1,2} \end{aligned}$$

That is, it is a term directly proportional to the phase difference between the waves with meridional wavenumbers 1 and 2. In order to maintain the blocking state against dissipation, this term must precisely balance the linear terms, and so a phase difference between the two wave modes must be maintained, generating an approximate phase locking. This

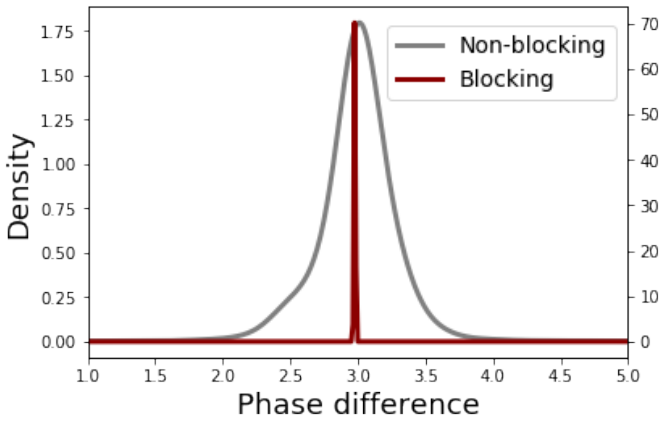


Figure 1.26: Distribution of phase differences between the wave-1 and wave-2 modes, both in the blocking and non-blocking states, calculated as $\tan^{-1}\frac{x_2}{x_3} - \tan^{-1}\frac{x_5}{x_6}$. The blocking regime shows a precise phase locking between the two wave modes.

extra constraint reduces the degrees of freedom in the system and gives us some insight into the low attractor dimension during blocking shown in figure 1.8.

Admittedly, the rigour of this small angle approximation is in question, as the x_2 and x_5 modes aren't always *that* small compared to the x_3 and x_6 modes. We fall back on the empirical fact that this suggested phase locking is indeed present during blocking, as shown in figure 1.26, even if we can't explain its origin precisely.

1.7.1 Vorticity budgets of equilibria states

We have seen that the fixed-point equilibria of the CdV model, while unstable and so never actually visited by the system, nonetheless shape the system's evolution. It is therefore valuable to understand their properties for the interpretation of the stochastically forced system. Figure 1.27 shows the vorticity tendencies at the two fixed points, where of course there is no net tendency in any mode, and no eddy terms. The vorticity budget of the blocking equilibria is very similar to the overall blocking regime: linear terms (dominated by damping) attempt to increase the vorticity, opposed by nonlinear terms, with only small adjustments in the overall balance to account for the absence of eddies. The magnitude of the terms is generally larger at the fixed point which follows intuition –

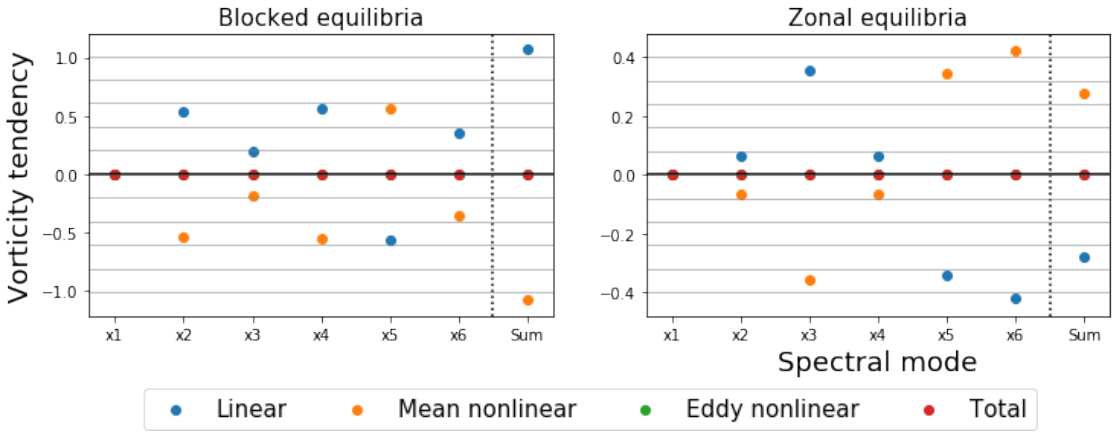


Figure 1.27: As in figure 1.25 but showing the vorticity balances at the zonal and blocked fixed points relevant to the CdV dynamics.

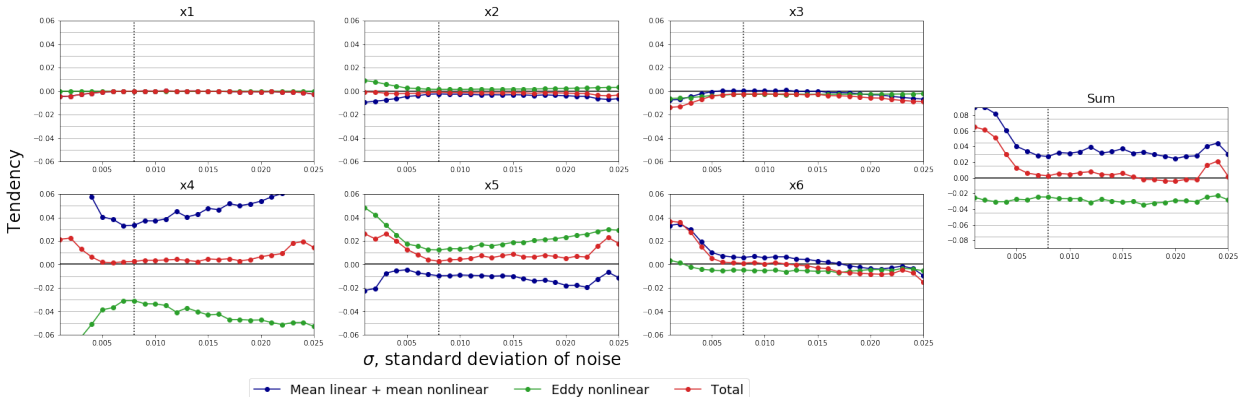


Figure 1.28: The changing Reynolds balance during blocking in each mode as the amplitude of stochastic forcing increases. The amplitude of peak blocking persistence σ=0.008 is indicated with a vertical line. The total vorticity tendency goes towards zero as the peak persistence amplitude is approached as a result of mean state changes.

we would expect a precise cancellation between large terms to produce an unstable state. The zonal fixed point meanwhile has quite different tendencies to the average nonblocked flow state where eddies were playing a significant role in maintaining the vorticity budget. It represents, qualitatively, the opposite of the blocked equilibria, with advection seeking to increase vorticity while now linear terms dissipate it.

If the Shutts eddy straining mechanism has any relevance in this model, we will expect to see an increase in the magnitude of nonlinear eddy terms with stochastic forcing, acting to oppose vorticity growth during the blocking state (i.e. acting against linear dissipation), presumably reaching a maximum at the noise amplitude of peak persistence, σ =8e-3.

Figure 1.28 shows the change in the vorticity tendencies as a function of stochastic forcing, where we have grouped the two mean terms together in order to better highlight the role of eddies. Contrary to the eddy straining hypothesis, we in fact see reduced eddy activity in almost all modes as the amplitude of stochastic forcing approaches its peak persistence. The overall impact of the eddy terms, as shown in the Sum panel, is constant regardless of noise amplitude, and it is in fact a reduction in the positive vorticity tendency of the mean terms that leads to the overall quasistationarity of blocking. That is, the mean state of the system changes under stochastic forcing, in such a way that supports longer lived blocks: the eddy-straining mechanism is not relevant here.

We have already seen that stochastic forcing induces a mean drift towards the blocking equilibria (c.f. figure 1.16). This prompts us to look in more detail at how the blocking mean state is being changed by stochastic forcing which, after all has no embedded drift term. Part of the answer is that additive noise serves to decorrelate eddies, changing the balance of tendencies in each mode, and so driving secondary adjustments in the circulation. However this does not capture the full picture, as we see that the changes in circulation are larger in the blocking regime than the other regimes, even though eddies were least relevant there. Figure 1.29 shows the average distance of every point in the stochastic simulations from its closest neighbour on the deterministic attractor. This measures to what extent novel regions of the phase space are being sampled. Naturally, the stronger the forcing the larger the average distance, but we see that at noise amplitudes exceeding 0.003, the average distance for points in the blocking regime starts to diverge rapidly from the non-blocked regimes. Now we see why the impact of stochasticity is asymmetric between the different regimes; the blocked regime is sampling new regions of phase space to a more significant degree than the other regimes. Figure 1.30 shows what these deterministically inaccessible blocked states look like. We sample the 10% of blocking points from the σ=8e-3 simulation that are furthest from the attractor and composite them. They represent upstream shifted versions of the mean blocking state, with weakened westerlies, very similar in fact to the unstable blocking equilibria. These states

however are not persistent or self sustaining, featuring very high net vorticity growth focused primarily on restoring faster Westerlies, and that they are visited at all is due to chance excursions driven by stochastic effects. That is, it is only through the chance stochastic reduction of westerly winds that these states can occur. But why these states in particular; what causes this preferential exploration of strongly blocked flows?

There are two contributing factors that lead to this preferred direction: the first is the presence of the fixed point, which can "trap" the flow in its vicinity, and the second is suppression of perturbations in many directions by the phase locking of the wave modes.

Consider the scenario where, by chance, the stochastic flow passes close to the blocking fixed point, which we remember has one very weakly unstable dimension and 5 highly stable dimensions. In this region of phase space the deterministic tendency is very close to zero, and so the stochastic terms dominate entirely: the dynamics will be very close to six-dimensional Brownian motion. How long, in expectation, will it take for this Brownian motion to move the system far away from the fixed point once again, and return to a dynamics-dominated flow state? This is the question considered in the hitting time problem: at what time will a random process first reach a distance a from its initial condition? For Brownian motion it has been mathematically proven that in fact the average time taken is infinite: the distribution of hitting times is fat-tailed and so the mean diverges [Kallenberg, 2017] – although any individual process will only take a finite hitting time. Now we can understand the heavily skewed blocking lifetime distribution of figure 1.18, which has been influenced by these fat-tailed hitting time dynamics.

This reasoning requires the system to approach the fixed point with some regularity. Why should random perturbations in a six-dimensional space approach a particular point on a recurrent basis? We saw that the number of unstable dimensions during blocking is far lower than during nonblocking configurations, which can be in part explained by the phase locking of the wave 1 and wave 2 modes. While stochastically perturbing the strength of the Westerlies will trigger a restoring force in the model, this is primarily

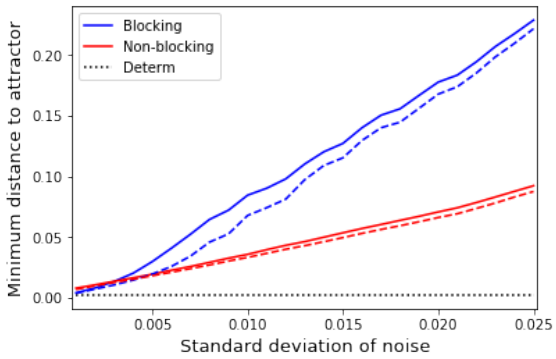


Figure 1.29: The mean (solid lines) and median (dashed lines) euclidean distance between points in the stochastic simulation and their closest neighbour on the deterministic attractor, shown for the blocking and non-blocking regimes. The blocking distance increases rapidly and skews for noise amplitudes exceeding 0.003, indicating the blocked flow visits many novel flow states.

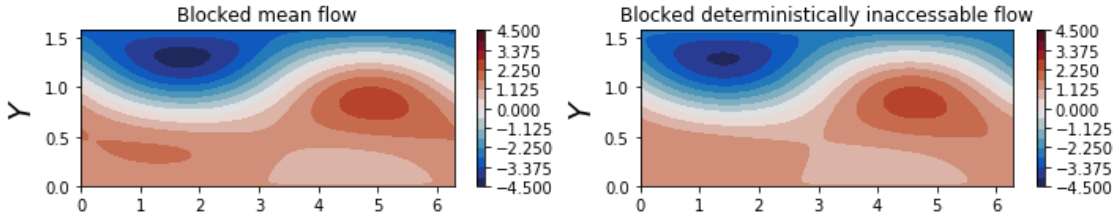


Figure 1.30: Left: Streamfunction composites of the 10% of blocking points in the σ=0.008 simulation that are furthest from the deterministic attractor. Right: The mean blocked state in the deterministic simulations, reproduced for comparison.

dominated by the linear relaxation term, and occasionally the orographic term, also linear. However perturbations in any of the wave modes will cause changes in the phase difference. This will produce quadratic restoring forces from induced anomalous advection, which have been observed in simulations to be larger than the entire linear tendency (not shown). This means that perturbations to the waves during blocking will quickly damp out, with the associated vorticity anomaly projecting onto the x_4 mode. It is this contrast between linearly and quadratically damped directions of instability that create a preferred direction of stochastic forcing, driving the system towards the unstable fixed point and the low associated deterministic tendencies.

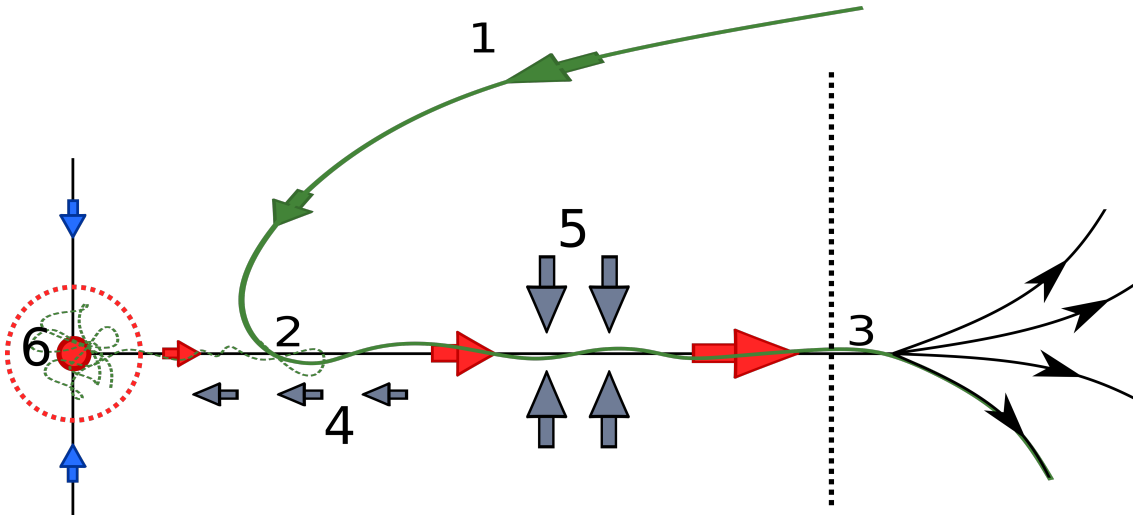


Figure 1.31: The stochastic regime persistence mechanism of the CdV system. A deterministic trajectory (shown with a solid green line) coming from the transitional regime (1) intersects the one-dimensional unstable manifold (2) at a point far from the fixed point(6) which is inaccessible to the deterministic attractor. It then remains in the blocking regime, accelarating away from the fixed point until the number of unstable dimensions increases (3) and the system enters the chaotic zonal regime. A corresponding stochastic trajectory however (shown with a dotted green line) experiences a drift towards the fixed point that weakens the instability (4), and so blocking lifetime increases. Any stochastic perturbations off the unstable manifold are strongly damped by quadratic nonlinearities (5). While there is still a net acceleration away from the fixed point, some stochastic trajectories will pass very close to the fixed point (6). Here the deterministic tendency is very close to zero, and the average time taken to leave the blocked state diverges, producing a fat-tailed lifetime distribution.

1.7.2 The stochastic persistence mechanism

We are now in a position to fully explain the impact of stochasticity on the dynamics of the CdV system. Introducing additive stochastic forcing produced a modest increase in the persistence of all regimes, but by far the most pronounced impact is on the blocking regime.

The mechanism of increased blocking persistence is summarised in the schematic figure 1.31. In the deterministic model, trajectories in the highly chaotic transitional regime are defined by high-amplitude travelling wave modes. These eventually settle down into a long-lived blocking event with a downstream ridge. This persistent blocking state has low vorticity, and experiences strong relaxations to the zonally symmetric background state that attempt to increase vorticity. These relaxations are balanced by orographic drag, and by nonlinear wave interaction terms that necessitate an approximate phase locking of the wave 1 and wave 2 modes.

The balance is only approximate, and there is a slow increase in vorticity, meaning that eventually the blocking event decays and the flow returns to a high vorticity, chaotic, zonal flow. No matter how weak the growth of vorticity during blocking, it is always positive in magnitude, and so the block is always weakening during its lifetime. Longer lived blocks enter the blocking regime closer to the unstable blocking fixed point, where the initial growth of vorticity is smaller, and so the lifetime of the event is set by the precise details of the chaotic trajectory that brings the system into blocking.

In the stochastic simulations, the transition into blocking is similar; in the transitional regime the deterministic tendencies are large and so the additive stochastic terms are less impactful. Within the blocking regime, the system experiences stochastic perturbations both in the zonally symmetric modes that set the strength of the model Westerlies, and in the wave modes. Perturbations in the wave modes will almost always violate the phase-locking and so are damped away by strong quadratic tendencies. Meanwhile, perturba-

tions in the westerly winds can either hasten the dissipation of blocking or prolong the blocking event, depending on whether they strengthen or weaken the zonally symmetric flow respectively. However as discussed in section 1.6, we have shown that a stochastic perturbation to an exponentially growing instability will on average delay the growth of instability, and so effectively a stochastic drift term is introduced that acts to lengthen blocking lifetime.

On top of this, some stochastic blocking events will reach the vicinity of the deterministically inaccessible blocking fixed point, which requires a reduction of the Westerlies that allows the block to move upstream. Once close to the fixed point the deterministic tendency is essentially negligible, and so the system is experiencing near-Brownian motion. The time taken for Brownian motion to leave a region diverges in expectation, producing a highly skewed, and long-lived blocking lifetime. This means that even if very few stochastic trajectories scatter far enough to reach the blocking fixed point, those that do will dominate the statistics. These effects are not seen in the zonal regime, which also features a fixed point close to the deterministic attractor, for two reasons. Firstly, the dimension of the chaotic flow is higher in the zonal regime, meaning there is less restriction on the direction of stochastic perturbations, and their effect is less coherent. Secondly, the chaotic zonal flow does not feature a clearly growing instability with constant sign of curvature which would be amenable to the persistence mechanism discussed. Importantly, the CdV system moves between chaotic and predictable regimes, not two chaotic regimes, and it is the predictable blocking regime that is most sensitive to stochasticity. For very large noise amplitudes the regime structure falls apart and an essentially perpetual block sets in, in a fairly uninteresting stochasticity-dominated regime.

1.8 Challenges of increasing resolution in CdV

One of the most tantalising promises of stochastic parameterisations in atmospheric modelling is the idea that cheap stochastic forcing may be able to mimic some of the benefits

of a computationally expensive resolution upgrade. Testing whether stochastic forcing in simple models can emulate a more complex variant of the model is therefore a very interesting question, which might offer theoretical insights useful in the model development process. The CdV model seems like a likely candidate for such an experiment, as we need only reduce the severity of our spectral truncation to find a natural extension of our six mode model. It would then be possible to assess to what degree the increased regime persistence and predictability seen here with stochasticity emulates the behaviour of the extended model. However as discussed in section 1.3, the range of parameters for which the six equation CdV model maintains self-generating chaotic dynamics is relatively small, and the system is fragile to structural perturbations.

Here we briefly describe efforts made to compare the six-mode model with a ten-mode extension originally derived by [De Swart, 1988] (DS88) and the problems encountered along the way.

In order to extend the CdV model, a natural starting point is to introduce one additional mode. This could be done in either the meridional direction, producing a 9-equation system, or in the zonal direction producing a 10-equation system. As we are most interested in blocking and localised ridging, increasing zonal resolution is therefore the natural choice. Following the same steps as in discretising our 6-mode system, we acquire the

	Meaning	Crommelin '04	deSwaart '88
C	Linear damping	0.1	0.1
g	Orographic height	0.2	1
b	Channel half-width	0.5	1.6
β	Coriolis parameter	1.25	2
x_1^*	meridional wave-1 forcing	0.95	2.5
x_4^*	meridional wave-2 forcing	$-0.801 \cdot x_1^*$	$-2 \cdot x_1^*$

Table 1.1: The CdV parameter values used in the six-mode formulation of Crommelin '04, and the ten-mode formulation of deSwaart '88, both of which produce chaotic dynamics.

following, ten equation o.d.e. system:

$$
\begin{aligned}
\dot{x}_1 &= -C(x_1 - x_1^*) + \tilde{\gamma}_1 x_3 \\
\dot{x}_2 &= -Cx_2 + \beta_1 x_3 - \alpha_1 x_1 x_3 - \delta_1 x_4 x_6 - \rho_{11}(x_5 x_6 - x_6 x_7) \\
\dot{x}_3 &= -Cx_3 - \beta_1 x_2 + \alpha_1 x_1 x_2 + \delta_1 x_4 x_5 - \gamma_1 x_1 + \rho_{11}(x_5 x_7 + x_6 x_8) \\
\dot{x}_4 &= -C(x_4 - x_4^*) + \epsilon(x_2 x_6 - x_3 x_5) + \tilde{\gamma}_2 x_6 + \epsilon_2(x_7 x_{10} + x_8 x_9) \\
\dot{x}_5 &= -Cx_5 + \beta_2 x_6 - \alpha_2 x_1 x_6 - \delta_2 x_4 x_3 + \rho_{11}(x_2 x_8 - x_3 x_7) + \gamma'_{12} x_8 \\
\dot{x}_6 &= -Cx_6 - \beta_2 x_5 + \alpha_2 x_1 x_5 + \delta_2 x_4 x_2 - \gamma_2 x_4 - \rho_{11}(x_2 x_7 + x_3 x_8) - \gamma'_{12} x_7 \\
\dot{x}_7 &= -Cx_7 + \beta_{21} x_8 - \alpha_{21} x_1 x_8 - \delta_{21} x_4 x_{10} - \rho_{21}(x_2 x_6 + x_3 x_5) + \gamma'_{21} x_6 \\
\dot{x}_8 &= -Cx_8 - \beta_{21} x_7 + \alpha_{21} x_1 x_7 + \delta_{21} x_4 x_9 + \rho_{21}(x_2 x_5 - x_3 x_6) - \gamma'_{21} x_5 \\
\dot{x}_9 &= -Cx_9 + \beta_{22} x_{10} - \alpha_{22} x_1 x_{10} - \delta_{22} x_4 x_8 \\
\dot{x}_{10} &= -Cx_{10} - \beta_{22} x_9 + \alpha_{22} x_1 x_9 + \delta_{22} x_4 x_7
\end{aligned}
\tag{1.42}
$$

The terms shown in purple are wave triad interactions, a new kind of quadratic nonlinearity that includes only wave modes, not present in our six equation system. We see that the six wave modes that we had previously have picked up additional terms, representing their interactions with the newly introduced modes. This changes the stability properties of the modes, and so importantly, solutions to the six mode sub-system are not generally solutions of the ten mode system.

In DS88 it was shown that this extended system possesses free-evolving chaotic dynamics and in fact there they incorrectly stated that the six mode model considered in this chapter did not possess such dynamics. The differences in the model parameters of C04 and DS88 are shown in table 1.1. In the DS88 configuration, the zonal forcings are stronger, with a dominant x_4^* component, and the orography is 5 times higher. The channel geometry, controlled by b and β, is also significantly different: the meridional length of the channel is close to the zonal length, producing a more synoptic length scale compared to the more planetary scale used in this work, and the value of β implies a channel centred at 55 degrees latitude. The increased meridional width allows for an increased range of bifurcation behaviour as discussed in both DS88 and C04. Clearly, the two models are in very different regions of parameter space, and so there dynamics are not immediately comparable.

Unfortunately, the structural instability of these simple models means that for neither parameter set do both truncated systems possess chaotic dynamics. Figure 1.32 shows the power spectra and x_1-x_4 subspaces of the deterministic 6D model, the stochastically perturbed 6D model (using the peak persistence noise amplitude of σ=0.008) and the deterministic 10D model at both parameter sets. While the deterministic 10D model of DS88 has a power spectrum similar to the deterministic 6D model of C04, with high power at low frequencies, the 10D model produces a simple limit cycle at the C04 parameters, while the deterministic and stochastic models produce periodic and quasiperiodic dynamics respectively using the DS88 parameters. These huge qualitative shifts in dynamics are of course not what we think about in a complex model context, where we are used to increasing resolution representing an iterative improvement and, at least roughly, a convergence to the real world climate.

Without resorting to 10-equation numerical continuation of the bifurcation points - well beyond the scope of this thesis - it has not proven possible to find a parameter set where both the six and ten mode systems exhibited chaotic dynamics. A large number ($\sim$

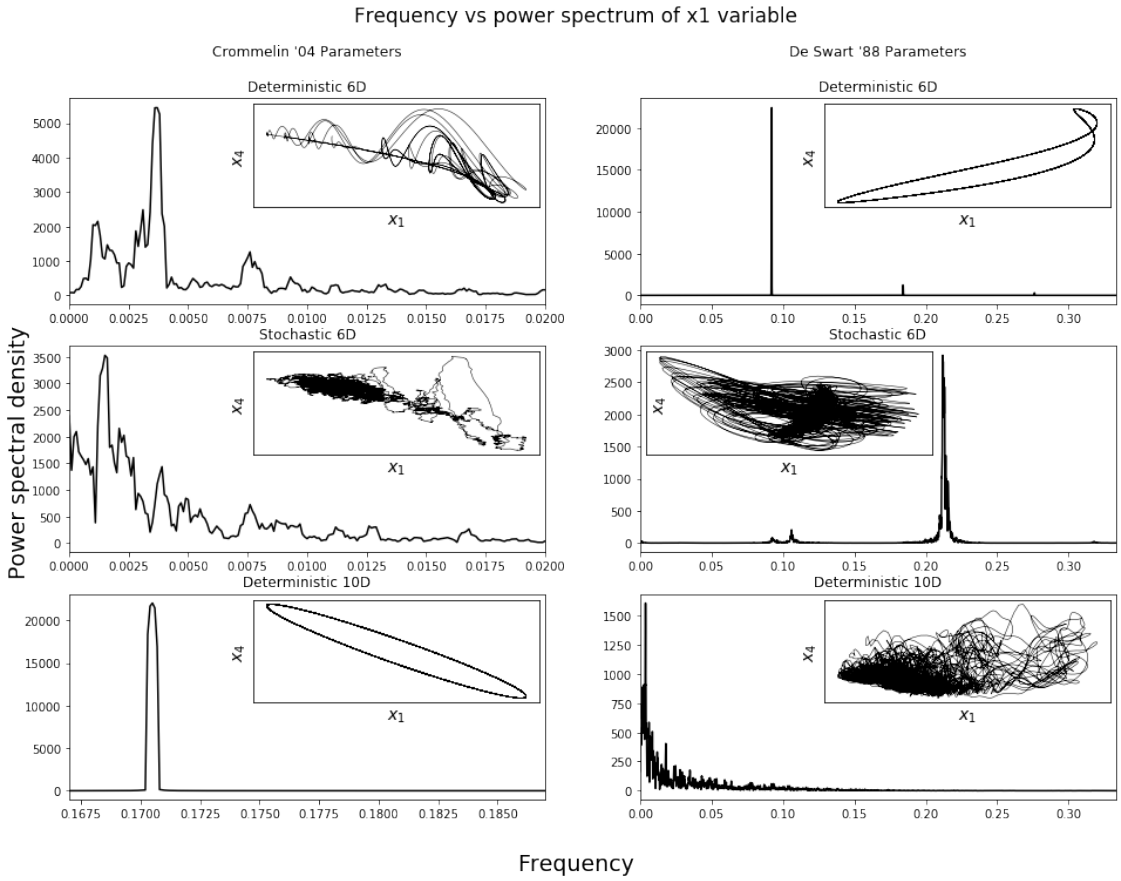


Figure 1.32: The power spectra of the x_1 mode for 200,000 MTU simulations of Top: the deterministic CdV model, Middle: the stochastically forced CdV model, and Bottom: The 10-mode deSwaart model, at different parameter sets. Embedded panels show the full corresponding simulations projected in the x_1-x_4 subspace.

$\mathcal{O}$(1e6)) of brute force integrations of both systems, varying g, b and C were performed in order to find a shared chaotic parameter set, but almost all combinations resulted in periodic or stationary solutions for both resolutions.

A final attempt at comparing the two systems was to introduce a coupling, or homotopy, parameter h, varied in the range [0,1], such that all the terms in the tendencies of modes 1-6 that involved modes 7-10 were multiplied by h, and so could be gradually 'turned on'. e.g., using the x_5 mode as an illustrative example:

$$\dot{x}_5 = -Cx_5 + \beta_2 x_6 - \alpha_2 x_1 x_6 - \delta_2 x_4 x_3 + h\left[\rho_{11}(x_2 x_8 - x_3 x_7) + \gamma'_{12} x_8\right] \tag{1.43}$$

However, this also proved unsuccessful, as the properties of the attractor radically changed as h was varied, and no regime dynamics were recovered for non-zero h for a wide range of different parameter values considered.

1.9 Discussion and conclusions

In this chapter we have taken a detailed look at the CdV model, in its fully chaotic formulation first derived by C04. This paradigmatic model in its original formulation introduced the concept of linking atmospheric regimes to fixed points of the underlying large scale flow, and we see that the newer formulation of this model is useful for understanding how more complex invariant measures – in this case unstable homoclinic orbits – might relate to self sustaining regime dynamics.

We have expanded on the work of K14, deriving the same three regimes that they obtained by applying a hidden Markov model to the full phase space. We demonstrate the dynamical importance of these regimes by showing they broadly correspond to regions of the attractor with differing numbers of unstable dimensions.

We show that the introduction of additive stochastic forcing into the model not only in-

creases the persistence of the regimes – primarily the blocking regime – but also produces demonstrable shifts in the Fourier spectra of the model modes to low frequencies. With the aid of ensemble forecasts we showed that while this increased stochastic persistence does not also increase the predictability as measured via ensemble spread, it does result in an increase in regime predictability within an intermediate range of lead times. This means that while we have less certainty as to the exact coordinates of the stochastic simulation at a later time than we do in the deterministic simulation, we may sometimes have a more skilful probabilistic estimate of what regime the stochastic simulation will be in. Even for the spread measure of predictability we noted a non-monotonic increase in spread with stochastic forcing; for example after 70MTU, simulations initialised in the blocking regime in the σ=0.004 simulation had lower ensemble spread than those in the σ=0.001 simulation.

We observe that the CdV system contains a vacillation between chaotic and ordered flow states, with a turbulent zonal and travelling-wave dominated dynamics punctuated by quasi-stationary periods in a strongly blocked state. Similarities can be seen to the behaviour of the 1-D Pomeau-Maneville map which demonstrates the same intermittently chaotic dynamics, and while the parameter tuning required to produce these dynamics in CdV are fragile, this is not a purely artificial construction. Similar observations of movements between chaos and order in atmospheric systems have been recently documented in [Shen et al., 2021], and the dynamics we see in our simple six-equation model bear remarkable similarities to the conclusions of [Faranda et al., 2016], who apply cutting-edge dynamical systems metrics to observational data and concludes that persistent atmospheric blocking in the Atlantic is due to the influence of an unstable fixed point in a single basin chaotic attractor of prevailing zonal flow.

Beyond simply describing the model dynamics, we have now explained the origin of this stochastically amplified persistence – and its asymmetric impact on the blocking state – both in mathematical and physically-minded terms. We have introduced a simple geomet-

ric argument with broad general applicability, which explains how a zero-mean stochastic perturbation will act to delay the growth of one-dimensional instabilities and so can prolong intermittent periods of order in a chaotic dynamical system. In this case, this serves to oppose the monotonic increase in vorticity deriving from linear forcing dominating over orographic drag and wave advection. We have also shown that the changing dimension of the unstable manifold across the attractor is key in explaining the asymmetry of the stochastic persistence. The reduction in the number of unstable dimensions during blocking can be partially understood in terms of a phase-locking which develops between the wave modes, and which we demonstrate only occurs in the blocked flow state.

This stochastic persistence is intensified, and a heavy tail introduced into the lifetime of blocking events, through the very weak instability of the blocking fixed point. This fixed point, representing a strong block shifted westward, is deterministically inaccessible as the imposed zonal wind forcing prevents blocks forming so far upstream. Once stochastic forcing is added however, the winds can be temporarily reduced long enough for the wave modes to adjust and support and stabilise the weak Westerlies. Because the growth of vorticity close to the fixed point is so weak, the dynamics in this region of phase space are almost entirely noise dominated, and the system will remain close to the fixed point until it scatters some critical distance away, where the deterministic tendency starts to become significant again. We have discussed how this scenario connects us to the theory of hitting time problems, where it can be proven that the distribution of times taken for Brownian motion to leave a radius a of an initial point is fat tailed. These chance reductions in the Westerlies could be thought of as loosely analogous to the influence of unresolved wave breaking removing momentum from the flow, or bursts of high form drag from interactions with unresolved topography.

Ultimately, if stochastic forcing is increased past σ=0.008, the blocking regime starts to become less persistent. This is not clearly linked to a resonance effect as suggested in K14 and C04, but is instead a result of entering a completely noise-dominated dynam-

ics; the zonal regime resembles the blocked regime and the blocked regime resembles a very strongly blocked regime. This effect is of little relevance for practical applications of stochastic parameterisation to models, as the stochastic terms will almost always be second-order compared to deterministic tendencies.

These mechanisms of stochastic persistence are entirely a result of mean state changes, and the net impact of eddy terms on the vorticity balance of the model did not change as stochasticity was increased. Therefore the Shutts '83 eddy-straining mechanism does not come into play here. This does not rule out the possibility of the mechanism playing an important role in more complex models, as here we do not have a clear range of spatial scales, and cannot resolve the spatially localised dipoles of eddy-vorticity flux seen in figure 1.23.

Another interesting aspect of stochastic parameterisation we cannot address in this simple model is the ability of stochastic models to emulate the performance of higher resolution deterministic models. Despite the seeming suitability of this system for such analysis, as we can simply adjust our spectral truncation scale, the sensitivity of the parameter tuning caused problems. Clearly there is a difference of kind between moving from six modes to ten modes, and moving from six thousand modes to ten thousand, and at these extreme truncations we have considered, the behaviour of the system is far from convergence to the 'true' infinite mode solution. One might reasonably ask whether the fragility of the CdV system to parameter changes and new modes in any way invalidates the results we have obtained. However, it is not necessarily surprising that a six equation model must be carefully tuned to reproduce a phenomena that is poorly captured even in very complex models. Further, even if the model itself does not generalise well, the dynamical building blocks and the mechanisms of stochasticity it contains should indeed have more general applicability.

There are multiple ways this analysis could be extended and continued. One line of research would be to move to slightly more complex models, such as the 15-mode spherical

QG model of [Kwasniok, 2007]. In such a model it would be possible to explore whether the phase-lockings, role of weakly unstable fixed points, and chaos-to-order transitions play the same role, in a similar way as we have done here. Another approach is to move directly to the high-resolution operational models that are of applied interest, and search for connections to the mechanisms seen here in more realistic regime systems, subject to teleconnections, ocean fluxes and other parts of the fully coupled Earth system. While this will indeed make drawing links harder, it will help evaluate the true usefulness of our simplified results more directly.

2 | The impact of stochastic perturbations on geopotential-jet regimes

2.1 Introduction

In this chapter we aim to understand the role of stochastic forcing on regime dynamics, not in a simple low order model – as in chapter 1 – but in fully coupled climate simulations. We found in our first chapter that introducing a moderate degree of additive stochastic forcing into the Charney-deVore model produced increased regime persistence, particularly in the blocking state, by allowing the system to explore deterministically inaccessible regions of the phase space. These newly visited states had lower zonal winds, and were close to an unstable equilibrium, where orographic drag and advection perfectly balanced thermal relaxation to produce a strong blocked state. Can we also detect robust increases in regime persistence in stochastically forced climate models, and if so, can the causes be in any way linked to the mechanisms of the simple Charney-deVore system?

We will aim to answer this question using EC-Earth simulations run as part of the SPHINX project [Davini et al., 2017b], which consist of long ensembles of deterministic and stochastically forced model simulations.

However, identifying robust regimes in the high-dimensional, under-sampled phase space of a climate model is more challenging than in a simple low-order system. Therefore the first part of this chapter is dedicated to a new methodological contribution to the literature, namely the introduction of 'geopotential jet regimes'. We deem these more suitable for regime comparisons between models and for analysing long simulations, than the classical circulation regime approach. We will start by explaining why a new regime framework is warranted in section 2.2, and after discussing the data we use in section 2.3, we will detail our motivation and develop our methodology using the ERA20C reanalysis

in section 2.4. In section 2.5 we will then summarise and validate our new approach using a number of other reanalyses, before discussing the regime variability in the historical record and the complexity of setting a real-world benchmark against which to evaluate models.

Only then, once we are satisfied with our ability to identify stable and reproducible regime dynamics in the historical record, do we move on to analysis of the SPHINX runs in section 2.7. We will show that stochastic parameterisation increases the fidelity of the geopotential jet regime patterns in a statistically robust manner, and that the biases in the Atlantic ridge regime are eliminated entirely. However we will also show decreased persistence in other regimes and an exaggeration of occurrence biases in the NAO- regime. We analyse mean state changes in the circulation as well as proxy parameterisation tendencies taken from IFS forecasts in order to understand the origin of some of these changes.

2.2 Why reinvent the wheel? The regime stability problem

Figure 2.1: A cautionary comic, reminding us that we must think carefully about whether we really need to introduce a new clustering methodology. Modified from xkcd.com/927

As has been discussed in chapter 0, there is an extensive literature developed over many

decades on the analysis of regimes. There are nearly as many frameworks and variant methodologies for isolating European circulation regimes as their are authors who have worked on the topic, each with their own particular advantages. Why then, do we want to try yet another methodology for identifying regimes? Are we falling into the trap suggested by the admonitory figure 2.1, and simply introducing a new clustering method for the sake of it?

We believe not. The most common regime perspective used in the Euro-Atlantic sector is to take a small number of the principal components of the Z500 field (typically 4-20 components are used), and identify regimes in this reduced phase space using K-means clustering. The idea here is to detect deviations from Gaussian dynamics in the large scale dominant modes of variability of the North Atlantic circulation. The geopotential height field contains similar information to the potential vorticity field, and so captures the highs and lows that define midlatitude weather, while the approximately barotropic nature of the large scale flow motivates analysing a mid-tropospheric height. From now on we refer to this standard approach as the *classical circulation regime* framework. This typically produces 4 regimes: the positive and negative phases of the North Atlantic Oscillation (NAO), an Atlantic ridge and a Scandinavian/European blocking pattern [Michelangeli et al., 1995, Cassou, 2008, Dawson et al., 2012, Strommen et al., 2019]. Confusions sometimes enter into the literature due to the fact that many related, but slightly different, phenomena are referred to as the NAO. For clarity, we note that the 'Negative NAO' regime features a high pressure anomaly centered over southern Greenland, displaced westward from the canonical centre of the NAO dipole over Iceland, and so is sometimes referred to as a Greenland blocking regime.

This framework is perfectly valid, and indeed has proven useful for model evaluation [Fabiano et al., 2021], analysis of flow-dependent predictability and impacts [Ferranti et al., 2015b, Grams et al., 2017b] and the study of tropical-european teleconnections [Cassou, 2008, Lee et al., 2020, Riddle et al., 2013]. Nonetheless there are some unsatis-

factory features of this framework. Firstly, there is no clear standout method with which to pick regime number, and the choice of exactly 4 regimes depends on complicated statistical significance testing, which is typically highly sensitive to the inclusion or removal of small numbers of points. While this is a particularly serious complication when evaluating regimes on centennial timescales, a shift in a thirty-year data-set by as little as four years (i.e. with 86% of the data remaining the same) can mean the difference between statistical significance or not, even when restricting to the satellite era [Strommen and Palmer, 2019].

Secondly, the 4 regimes obtained are themselves highly unstable, exhibiting significant decadal fluctuations in their spatial patterns. While this may ostensibly be a feature of the climate system, with slow-forcings and multi-decadal variability modifying or replacing the relevant regime patterns in different time periods, this would run counter to the intuition from simple regime systems that forcing should primarily modify regime occurrence and lifetime rather than the pattern itself [Palmer, 1999]. Pragmatically, it would also be much harder to understand such a non-stationary regime system, as variability in both spatial patterns and temporal statistics would have to be considered in tandem.

There is also evidence that a large part of this variability is in fact a simple consequence of sampling variability. Consider for example figure 2.2, which shows classical circulation regime patterns found in the satellite era, in 30 years of ERA5 data. A shift of one year creates a substantial change in regime patterns. The presence of frequent discontinuous jumps such as this in regimes suggest a non-physical origin for the variability.

Now we are hopefully convinced there is a problem to address, where does the issue come from, and what, if anything, can we do about it? Indeed, we may find that the atmosphere possesses no centennially-stable regime patterns, which would imply that spatial regime variability is a fundamental part of the regime dynamics. We will begin with a case study of the regime structure in the ERA20C reanalysis, which will help us intuit this complex issue and identify the root cause, before validating our approach with a broader range of

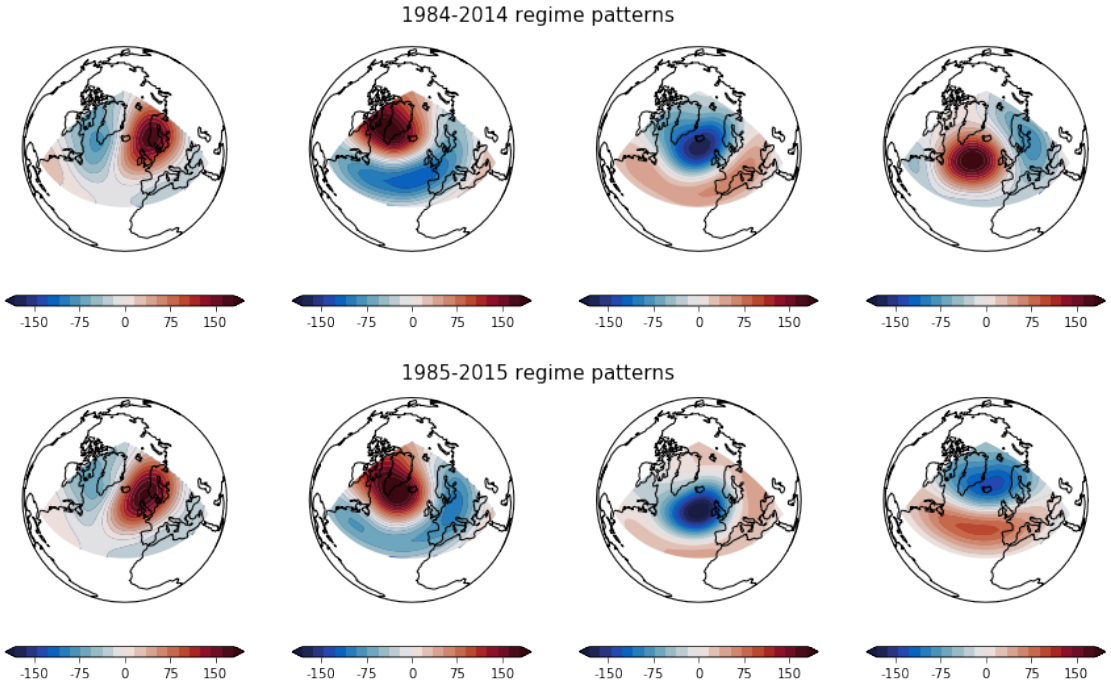


Figure 2.2: An illustrative example of the instability in identifying classical circulation regimes, using ERA5. The first 4 principal components of Z500 have been clustered into 4 regimes using K-means clustering, to produce two regime composites of geopoential height anomalies. The two datasets have 29 years in common, but have produced quite different regime patterns.

reanalyses. Ultimately, we will in fact find that all reanalyses we consider possess stable regime patterns.

Reanalysis name	**Generating Centre**	**Year of production**	**Time period used**	**Observations included**	**Assimilating Model**	**Resolution (xy,z,t)**
20CRv2	NOAA-CIRES	2011	1900-2010	surface observations	GFS (2008 experimental)	(210km, 28 levels, 6 hrs)
ERA20C	ECMWF	2016	1900-2010	surface pressure and marine wind observations	IFS Cy38r1	(125km, 91 levels, 3 hrs)
CERA20C	ECMWF	2018	1901-2010	surface, wind and ocean observations	IFS Cy41r2	(125km, 91 levels, 3 hrs)
ERA5	ECMWF	2020	1950-2010	surface, satellite and upper air observations	IFS Cy41r2	(31km, 137 levels, 1hr)
20CRv3	NOAA-CIRES-DOE	2021	1900-1980	surface observations	GFS 14.01	(60km, 64 levels, 6 hrs)

Table 2.1: A summary of the 5 reanalysis products considered in this chapter. For references to full descriptions of the products, see main text.

2.3 Data and methods

2.3.1 Reanalysis data

In order to assess the uncertainty in the historical record we use a number of different reanalysis products in this work. In the case of reanalyses with multiple ensemble members, we always use the first member, as the ensemble mean is unsuitable for analysis of blocking, a highly nonlinear quantity. We do not include any additional ensemble members, instead preferring to emphasise the spread between different reanalyses. We make use of boreal winter (DJF) geopotential height at 500 hPa (Z500) and zonal wind speed data at 850 hPa (U850), at a daily temporal resolution and on a 1 degree grid, from ERA20C [Poli et al., 2016], the extended ERA5 [Hersbach et al., 2020], CERA20C [Laloyaux et al., 2018], 20CRv2 [Compo et al., 2011], and 20CRv3 [Slivinski et al., 2021]. All available data covering the time period 1900-2010, as of Jan 2021, was used.

A summary of each product is given in table 2.1. Of these only ERA5 assimilates satellite data. ERA5 also uses the highest resolution assimilating model, therefore representing our best guess as to the true state of the atmosphere from 1950 onwards. The other centennial reanalyses use surface observations only, to ensure historical continuity with the pre-satellite era. Of these 20CRv3 is the most recent and uses the highest resolution. At the time this work was performed, the full reanalysis had not yet been produced, and so only data up until 1980 was available for use. CERA20C differs from ERA20C in that it is a coupled reanalysis and as a more recent product, has various methodological refinements. 20CRv2 was the first centennial reanalysis produced and has significantly lower resolution than more recent products.

2.3.2 SPHINX data

The climate SPHINX (Stochastic Physics HIgh resolutioN eXperiments) project [Davini et al., 2017b] aimed to explore the impacts of resolution and stochastic perturbations on

the climate of a state-of-the-art coupled GCM, namely the EC-Earth model [Hazeleger et al., 2012]. As a part of this project 3-member ensembles of the EC-Earth model were run both for a deterministic configuration, and for a stochastically forced configuration. The model resolution was T255 – roughly 80km – with 91 vertical levels, coupled to a 1 degree NEMO ocean model. The ensembles were integrated over the period 1850-2100 using historical CMIP5 forcings from 1850-2005 and the worst case RCP8.5 scenario from 2005-2100.

This dataset is unique in that it provides a clean comparison between a deterministic and stochastic model, with all other factors held constant, while also generating 465 years of historical Euro-Atlantic variability in each case. By analysing regime representation in the two experiments we will be able to make a robust assessment of the impact of stochasticity on regimes.

There are two types of stochastic perturbations present in the EC-Earth model: a stochastic kinetic energy backscatter (SKEB) scheme, and a stochastic perturbed parameterisation tendency (SPPT) scheme. SKEB is broadly analogous to the type of stochasticity we introduced to the CdV79 model in chapter 1, in that it is an additive perturbation to the streamfunction tendency designed to mimic neglected eddy feedbacks that reinject energy into the large scale modes. The SPPT scheme functions quite differently, addressing the unknown errors associated with parameterising subgrid scale processes, with a multiplicative perturbation. This perturbation is applied to the sum of all parameterised tendencies, and is not a white noise process, but is correlated on quite large spatial and temporal scales. It is not therefore a pure representation of subgrid scale variability itself, as is sometimes suggested, as such uncorrelated grid scale perturbations are dissipated without impacting model climate [Buizza et al., 1999]. It is instead a representation of the uncertainty in spatially extended phenomena such as cloudbanks and land processes. In this chapter, when we refer to the stochastic SPHINX runs, we mean these runs with both SKEB and SPPT active, while the deterministic SPHINX runs have neither active.

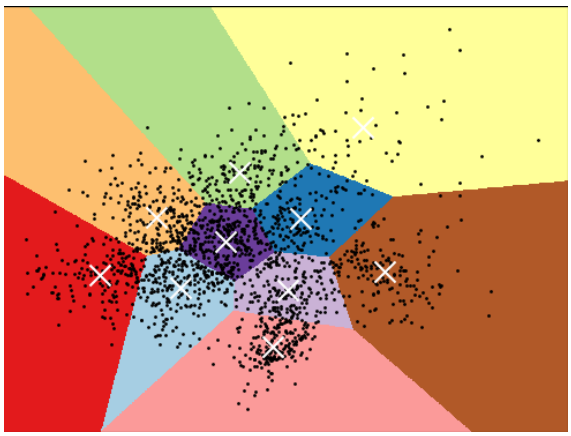


Figure 2.3: A schematic of the K-means clustering process for K=10, reproduced from https://scikit-learn.org/stable/modules/clustering. Data points (in black) are divided into clusters, the centroids of which are shown with white crosses. Each point is closest to the centroid of the cluster it belongs to, a constraint amounting to division of the space with planes.

2.3.3 K-means clustering

As discussed, K-means clustering is one of the most common methods used in atmospheric physics to identify regimes. It falls within the class of non-hierarchical clustering methods, meaning that it divides input data into a (in this case pre-specified) number of clusters with no structured relationship between them. This clustering is done using an iterative expectation-maximisation approach (see [Bishop, 2006] for details) such that each point is assigned to the cluster with the nearest mean value, using a Euclidean norm. Geometrically, grouping a dataset into K clusters in this way amounts to dividing the space with K-1 hyperplanes, as illustrated in figure 2.3. This also highlights why this approach was too coarse for the highly structured phase space of the CdV79 system analysed in chapter 1, where the regimes could clearly not be captured by splitting the space into straight edged boxes.

After clustering, we will have assigned each day of our dataset of length T to a given cluster, generating a state sequence $S = [s_1, s_2, ..., s_T]$ where $s_t \in [1, ..., K]$. Regime patterns can be found from the state sequence by compositing the Z500 anomaly field (or any other field for that matter) for all days assigned to a given regime. Note that under

this approach, the low-dimensional phase space of principal components is used only to identify the clusters, and to label each day in the dataset. Those labels are then used to produce composites in the original full-field datasets.

From this state sequence we can define *regime occurrence* as the fraction of days in S which are in the state k, and *transition probability* as the fraction of days, $\{s_t\} \subseteq S$, that are in state k_1 while $\{s_{t+1}\}$ are in state k_2.

In the case that $k_1 = k_2$ this tells us the Markovian probability of staying in the same cluster from one sample of the time series to the next, and so is called *regime persistence*, as in [Matsueda and Palmer, 2018, Falkena et al., 2020]. This metric would not be representative if regime lifetimes were strongly non-Markovian and so day 1 to day 2 persistence, say, was very different to day 5 to day 6 persistence. However in practice we find regime lifetime to be very close to Markovian for reanalyses and models (not shown), just as can be seen in [Straus and Molteni, 2004, Dawson and Palmer, 2015]. As a matter of terminology we will refer to the arbitrary divisions of the phase space produced by K-means as 'clusters', reserving the term 'regime' for those clusters for which we have obtained some confidence in their physical relevance. The 'state' of the atmosphere on a given day will generally refer to the active regime.

2.3.4 Jet diagnostics

We will often be considering the relationship between anticyclonic regime structure and the dynamics of the Atlantic jet. Here we are interested in the low-level eddy-driven jet which is dynamically distinct – even if often only theoretically – from the upper level subtropical jet, which is thermally driven.

The structure of the eddy driven jet can be summarised through the jet latitude and jet speed indices introduced in [Woollings et al., 2010], and here we compute them with the simplified method of [Parker et al., 2019]. In brief, the jet speed index is defined as the maximum (oriented Eastward) of latitudinally averaged 850 hPa zonal wind speed,

smoothed to remove synoptic fluctuations, over the Atlantic domain [60W-0W, 15-75N]. The smoothing timescale varies in the literature, but here we adopt a 5-day smoothing. The jet latitude index is complementarily defined as the latitude at which the jet speed is maximum on a given day. Note that this methodology makes no account for split jets, or an insignificant jet: a unique latitude and speed is defined for every day. Neither does it capture the tilt of the jet.

2.3.5 Stability and fidelity metrics

To quantify the robustness of regime patterns found in different subsamples of a dataset, we introduce a stability metric. Given a dataset D, containing a total of T 500 hPa geopotential height fields, we cluster the data (after reducing dimensionality using EOFs) into K clusters, with each day of the dataset assigned to one cluster. From this we define K cluster composites, R, which represent the average flow field associated with each regime, and using the full Z500 field as detailed above. As these regime composites are found by clustering the full dataset D, we term them the 'reference composites'. We also define a number of subsamples of D, $\{d_1, d_2, ..., d_N\}$ (which, for example, might be constructed by taking different 30 year windows of data from D), and define sub-sampled cluster composites $\{r_1, r_2, .., r_N\}$ by clustering and compositing each d_n independently. If the clusters are very stable, we would expect these subsampled composites to show very similar flow patterns to the reference composites, whereas if sampling variability or physically based non-stationarity is significant, they might show quite different patterns.

To quantify this, we compute the area-weighted pattern correlation between the K reference composites, R, and the K subsample composites of each r_n, producing a $K \times K$ matrix of correlations. The optimal 1-to-1 mapping between r_n and R which maximises the average pattern correlation of matched pairs is found using the linear sum assignment algorithm. The resulting length-K vector of pattern correlations between reference and subsample composites is then averaged over all subsamples. This vector then repre-

sents the average correlation between reference and subsampled regime patterns for each particular regime. This vector can be averaged once again to give the mean correlation across all patterns. This scalar value is termed the *cluster stability* and will be 1 iff clusters are perfectly unchanging. A visual schematic of this process is shown in figure 2.7 for 4 classical circulation regimes.

At times we will want to compare how well regimes in a subsampled dataset resemble the regimes found in a *different* dataset, such as when comparing model regimes to those seen in reanalysis. In this case the procedure is the same, except that we draw the reference patterns from a dataset D_1, while the subsamples $\{d_i\}$ are drawn from a second dataset D_2. In this case the average pattern correlation is termed *cluster fidelity*.

2.4 Geopotential jet regimes in ERA20C

Note: The work in this section is partially based on [Dorrington and Strommen, 2020]. The results presented in that work support the same qualitative conclusions as in this section, but differ in the details due to a change in methodology. In that work 10 principal components were used whereas here we use 4 principal components, in order to align our methodology with that which we will later use for CMIP6 data in chapter 3. This choice was in turn made in order to align with existing work on regimes in CMIP6 such as [Fabiano et al., 2021]. The robustness of our key findings to the choice of phase space dimension is a parenthetically reassuring result.

Why do classical circulation regimes suffer from such instability of their spatial patterns? If we consider the lower left panel of figure 2.4, which shows the first three principal components of ERA20C DJF Z500 over the domain [80W-40E,30-90N], we can see that the distribution of the system is not visually distinguishable from a linear, Gaussian noise system. Statistical testing as in [Dawson et al., 2012] and [Strommen et al., 2019] shows that this space is in fact significantly non-Gaussian, but arguably the subtlety of the multimodality, which is not visually apparent, is the cause of the ambiguity in regime

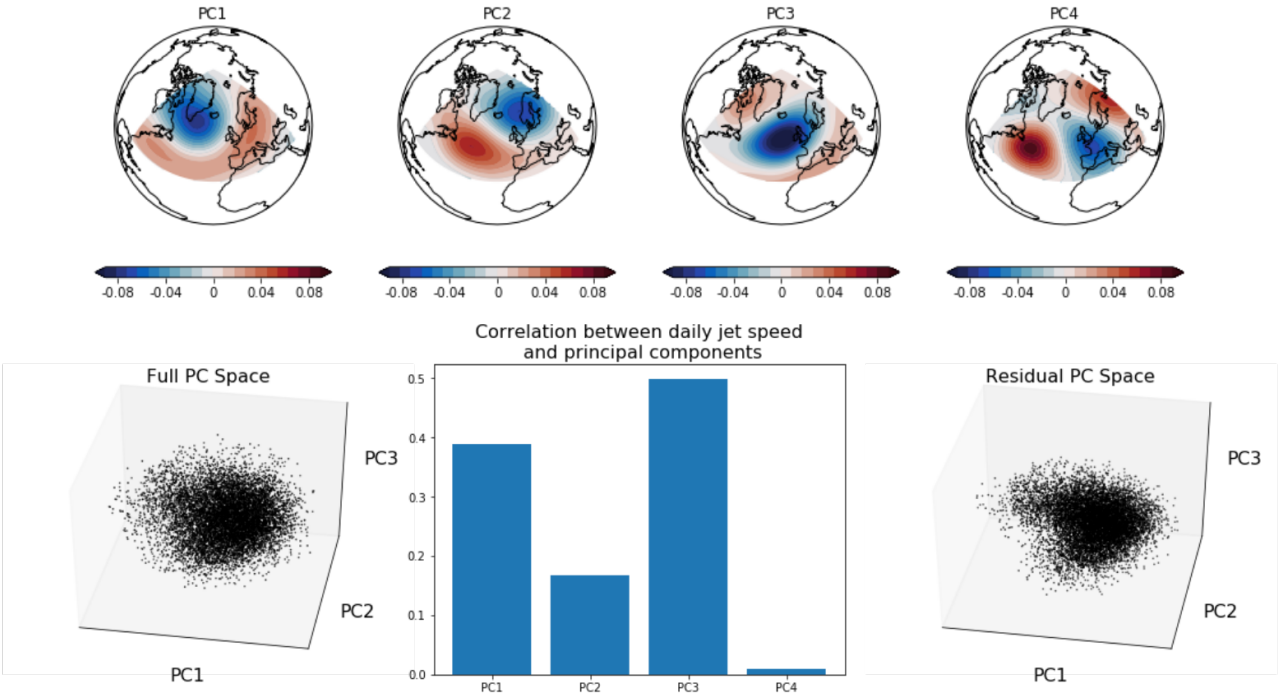


Figure 2.4: Top: The first 4 principal components of DJF Z500 in ERA20C. Lower left: The point cloud distribution of the first three principal components of ERA20C, showing a near Gaussian distribution. Lower centre: The correlation coefficients of jet speed with the 4 leading components of ERA20C. Lower right: As lower left, but showing points in the residual phase space, computed by regressing out the influence of jet speed from each principal component of Z500.

patterns.

This is in stark contrast to the regime behaviour of the eddy-driven Atlantic jet. In figure 2.5, we see the joint distribution of the jet speed and the jet latitude – again from ERA20C DJF data. While the jet speed is close to a linear Gaussian system, the jet latitude shows an unambiguous trimodality, featuring a southern, central and northern jet latitude peak. No detailed significance testing is needed to observe these regimes, and they are notably stable, showing little sensitivity to the choice of time-period used or subsetting of data. The unimodal jet speed and trimodal jet latitude are nearly independent modes of variability in a linear sense, with a very small correlation $\sim$-0.05, although the central jet latitude is associated with generally higher jet speeds.

It is somewhat puzzling that the jet dynamics should be so clearly a regime system,

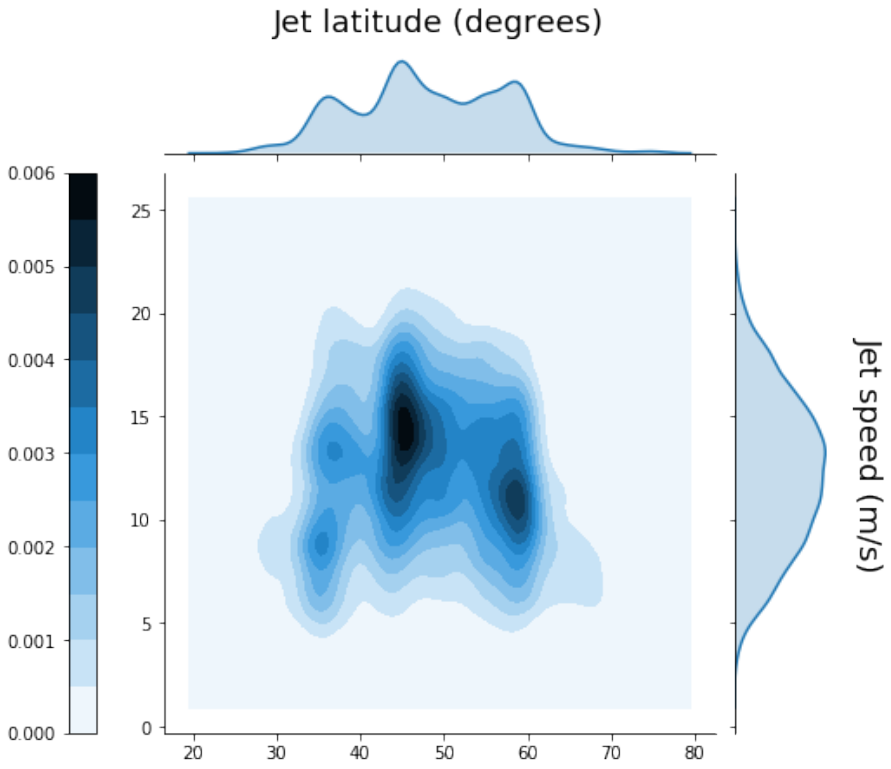


Figure 2.5: The joint and marginal probability distributions of daily DJF jet speed and jet latitude in the ERA20C reanalysis 1900-2010. The jet latitude shows a clear trimodality, while the jet speed has an essentially Gaussian distribution. The two indices are approximately orthogonal, although the central jet regime is in general associated with higher jet speeds.

while circulation regimes are so subtle, even when allowing for the different numbers of degrees of freedom between the Z500 space and the single jet latitude index: after all there is a close relation between the Rossby wave breaking that produces anticyclonic blocks, and the shifting jet latitude [Woollings et al., 2010, Masato et al., 2012]. This is also a practical concern as, while the jet picture is useful, it is insufficiently detailed to help explain forecast impacts in the way that circulation regimes can. So how do we reconcile these perspectives? The work in [Madonna et al., 2017] went a long way towards reconciling jet dynamics with classical circulation regimes, by identifying the 3 jet regimes with 3 of the 4 circulation patterns. It was showed that the 'missing' 4th regime (Scandinavian blocking) corresponds to either a tilted or split jet, not captured by the strictly zonal jet latitude index. However, these could only be cleanly separated by extending to 5 regimes, seemingly at odds with most studies supporting only 4 robust regimes.

The central conjecture we make here is that *the confounding factor obscuring Euro-Atlantic regime structure is the influence of jet speed.*

It is well known that the NAO index – the leading PC of Z500 during DJF – features strong linear correlation with the jet speed, which as we have shown is Gaussian in its distribution. This was also found by [Parker et al., 2019] who also showed the significant interdecadal variability in jet speed [Woollings et al., 2014] is essentially unpredictable on medium range timescales. Therefore it seems quite plausible that the linear jet speed variability could be making it difficult for clustering algorithms to identify the underlying multimodality. Concretely, we consider decomposing each of the principal components of geopotential height into a component linearly related to jet speed variations, and a residual which we hypothesise to capture the bulk of the nonlinear variability:

$$\mathrm{PC}_n(t) = A_n * u_{\mathrm{jet}}(t) + \mathrm{PC}_{n,\mathrm{resid}}(t) + c_n \tag{2.1}$$

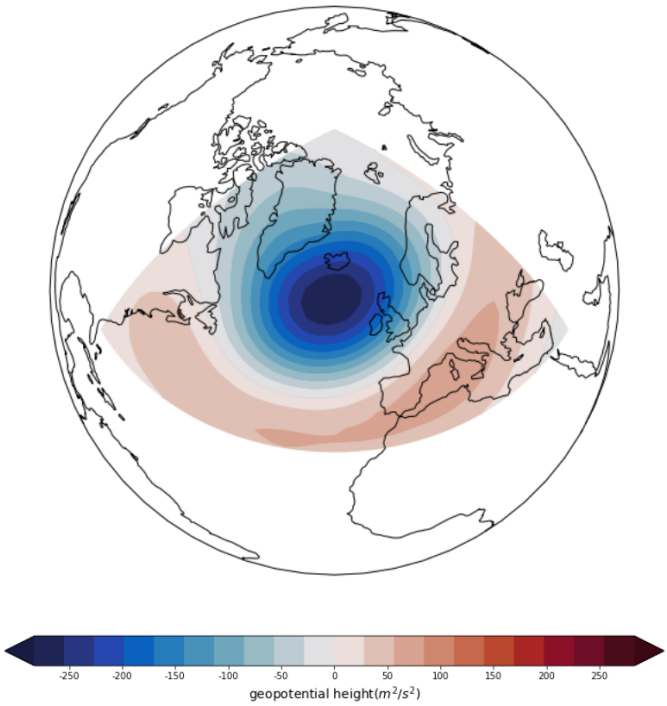


Figure 2.6: The regression pattern of jet speed onto the large-scale geopotential height field in ERA20C, as captured by the first 4 principal components.

Where $\mathrm{PC}_n(t)$ is the nth principal component of Z500, $u_{\mathrm{jet}}(t)$ is jet speed, and A_n and c_n are a slope and intercept obtained by a linear best fit. In general we should be cautious about regressing out a variable, even a linear one, from a nonlinear system. However, echoing the orthogonality of the jet speed and jet latitude, we posit an orthogonality between the blocking states and the jet speed. That is not to say that the zonal wind speed has no impact on blocking, as theory tells us it must do, but only that the speed of the jet core has no non-local impact on the blocking dynamics.

This conjecture leads us to introduce *geopotential-jet regimes* as shown in figure 2.4. We find there is a strong linear relationship between daily DJF jet speed and several of the leading PCs of Z500. The PCs corresponding to the first and third EOFs, representing an NAO dipole and an Atlantic ridge/trough dipole respectively, show the strongest correlations. This follows theory, where a stronger zonal wind speed is associated with a strengthened meridional pressure gradient and less wavy flow. The more zonally oriented blocking dipole of EOF2 also shows some small but significant correlation, while

the purer wave train of EOF4 is essentially uncorrelated to the jet speed. The overall linear projection of jet speed variability into this space of leading principal components is shown in figure 2.6. We see that a positive wind speed anomaly induces a steepening of the climatological height gradient, and is approximately zonally symmetric, although the poleward low is strongest in the mid-Atlantic. The projection of jet speed into the full Z500 field is visually indistinguishable (not shown), indicating that the jet influence is well captured within the leading PCs.

When this linear impact of jet speed is regressed out of the leading principal components, the phase space of the first three principal component residuals becomes noticeably non-linear, although still noisy, adopting a "croissant-like" shape, as seen in the lower right panel of figure 2.4. As discussed in section 2.3, regimes found in this regressed space are termed geopotential jet regimes because of the use of both geopotential height and jet speed data in their identification.

Has this regression procedure had the desired effect of making regime patterns more stable? To answer this we assess the cluster stability of classical circulation and geopotential jet regimes in ERA20C. We assume no specific cluster number at this time, keeping an open mind to K between 2 and 10. To find our reference clusters we use the full ERA20C dataset including all DJF seasons from 1900-2010. We produce 9 subsamples by using 30 year windows with ten years of separation between them: [1900-1930, 1910-1940,...,1980-2010]. This is similar in spirit to the resampling approach for assessing regime significance suggested in [Michelangeli et al., 1995].

As we will be discussing stability often in the rest of this thesis, it bears fruit to emphasise exactly what we are doing here. Therefore figure 2.7 explicitly shows the process of calculating the cluster stability from the full Z500 field when using four classical circulation regimes. We see that the overall stability of only 0.86 is quite low: this metric has quantified and summarised the interdecadal variability visible in the classical regime patterns. While here we are seeking to identify highly stable regimes, for the reasons mentioned in

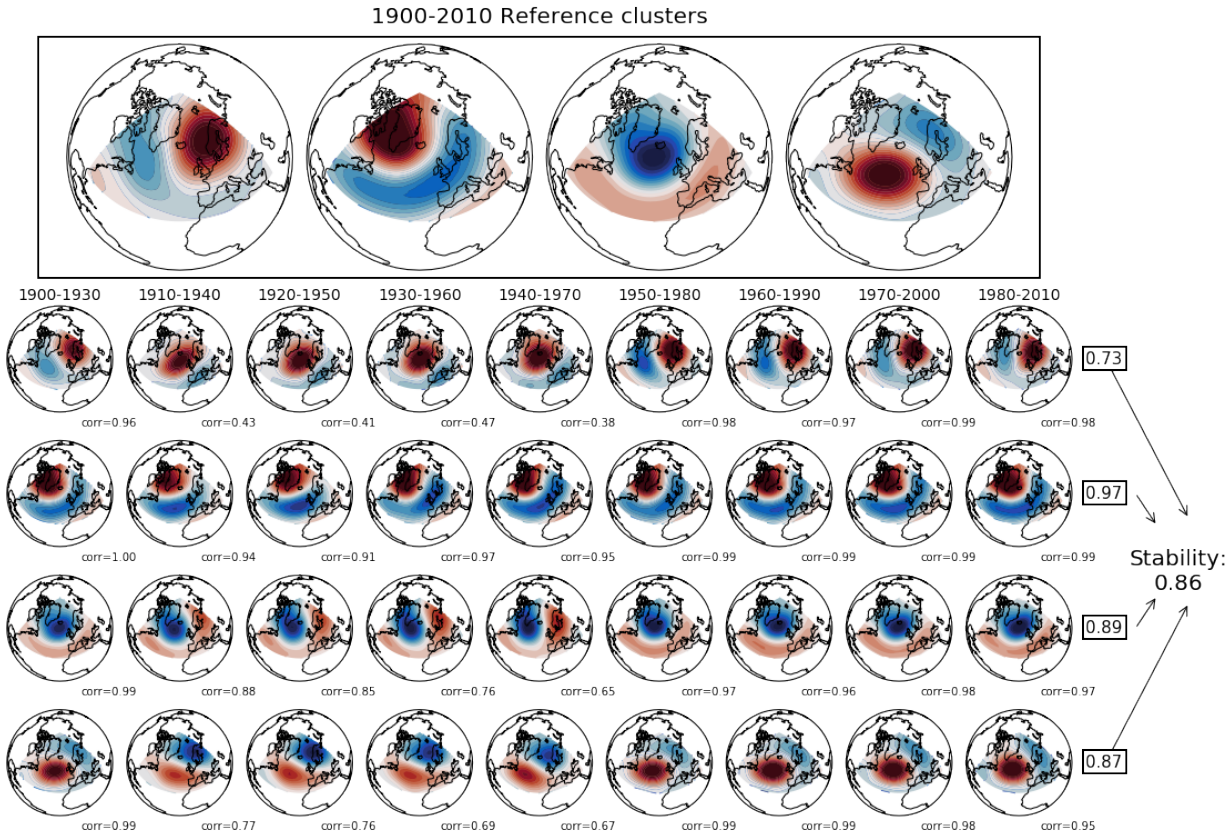


Figure 2.7: The process of calculating cluster stability for four classical circulation regimes in ERA20C. Reference composites are computed using the full 110 years of data. Each subsample is independently clustered, its 4 clusters are matched to the reference clusters, and the pattern correlation is computed. The average of these over all subsamples gives the stability for each pattern, which are then averaged again to get the total regime stability. Using 4 classical circulation regimes, a random cluster composite selected from a random subsample will, on average, have a pattern correlation with its corresponding reference cluster of 0.86.

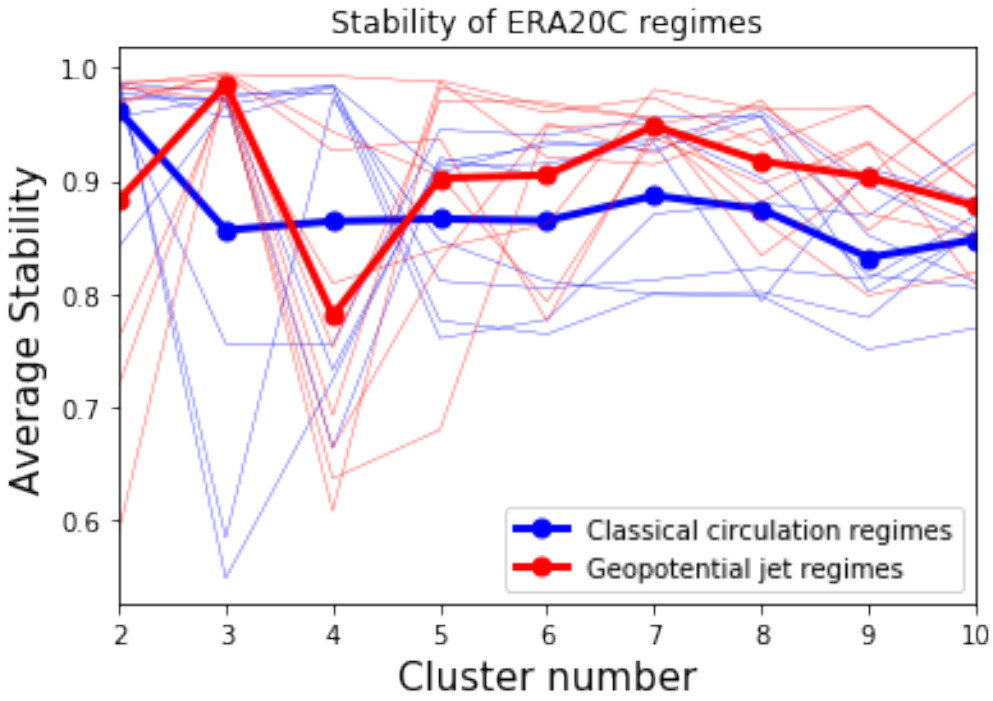


Figure 2.8: The cluster stability of classical circulation regimes and geopotential jet regimes found in the ERA20C reanalysis, as a function of cluster number. Thick lines show the average of all subsamples, while thin lines show the individual subsamples' correlation to the reference patterns.

Section 2.2, and so will be particularly interested in regime frameworks where stability is maximised, there are other possible applications for this metric. For example, this metric also helps reveal which regimes are most subject to interdecadal variations – in this case mainly the BLK regime, while the NAO- regime proves particularly stable.

Returning now to the more general case, figure 2.8 shows the stability metric for both clustering frameworks for all cluster numbers. We see that for classical circulation regimes, 2 clusters – which represents the two NAO states – gives the most robust results with a stability of 0.96. The common choice of 4 circulation regimes does not stand out as particularly stable compared to other cluster numbers. Geopotential jet regimes meanwhile show generally higher stability, with 3 geopotential jet regimes showing a clear maximum stability at 0.99. It bears emphasising just how stable these patterns are: these three clusters are even more robustly present in the data than the classical NAO!

If we choose to use stability as a metric to identify the number of clusters we wish to use, then the choice for geopotential jet regimes in ERA20C is clear – there are three regimes. Choosing cluster number in this way is both pragmatic, increasing the reproducibility of results and enabling comparison of regime dynamics found in different datasets, and also principled in that it ensures the regimes are some genuine persistent property of the atmosphere. We should emphasise that we have not imposed this stability criterion onto the regimes – if the regime dynamics really did fundamentally change over the 20th century then we would simply not observe any high stability regimes, and would have to resign ourselves to considering both spatial and temporal regime variability.

What are the characteristics of these new regimes? In figure 2.9 we show the full-field Z500 and U850 anomalies associated with these three regimes, as well as the conditional distributions of jet latitude and speed. The geopotential anomalies should look familiar to us, for they are in fact close analogues of the three anticyclonic regimes found when using four classical circulation regimes. The only difference is we have lost the cyclonic positive NAO regime. This is perhaps not totally surprising as we have regressed out the jet speed which strongly projects onto the NAO pattern (while the negative NAO regime, as stated, is really dominated by anomalies over Greenland, which project less closely on the jet speed). It also supports the perspective put forward by [Woollings et al., 2008] which suggested the NAO+ does not itself represent a true non-linear regime state, but is merely a default state from which deviations into persistent regimes can occur as a result of Rossby wave breaking.

We also see some similarities between these regimes and the three jet latitude regimes. The Atlantic ridge pattern is normally coincident with a northern jet, while the NAO- regime is bimodal in jet latitude but predominately linked to a southern jet. The blocking regime is largely decoupled from the jet latitude but significantly prefers a central jet. Looking at the U850 anomalies we see that the AR and NAO- do indeed link to a zonally symmetric displacement of the jet North and South respectively, while the blocking

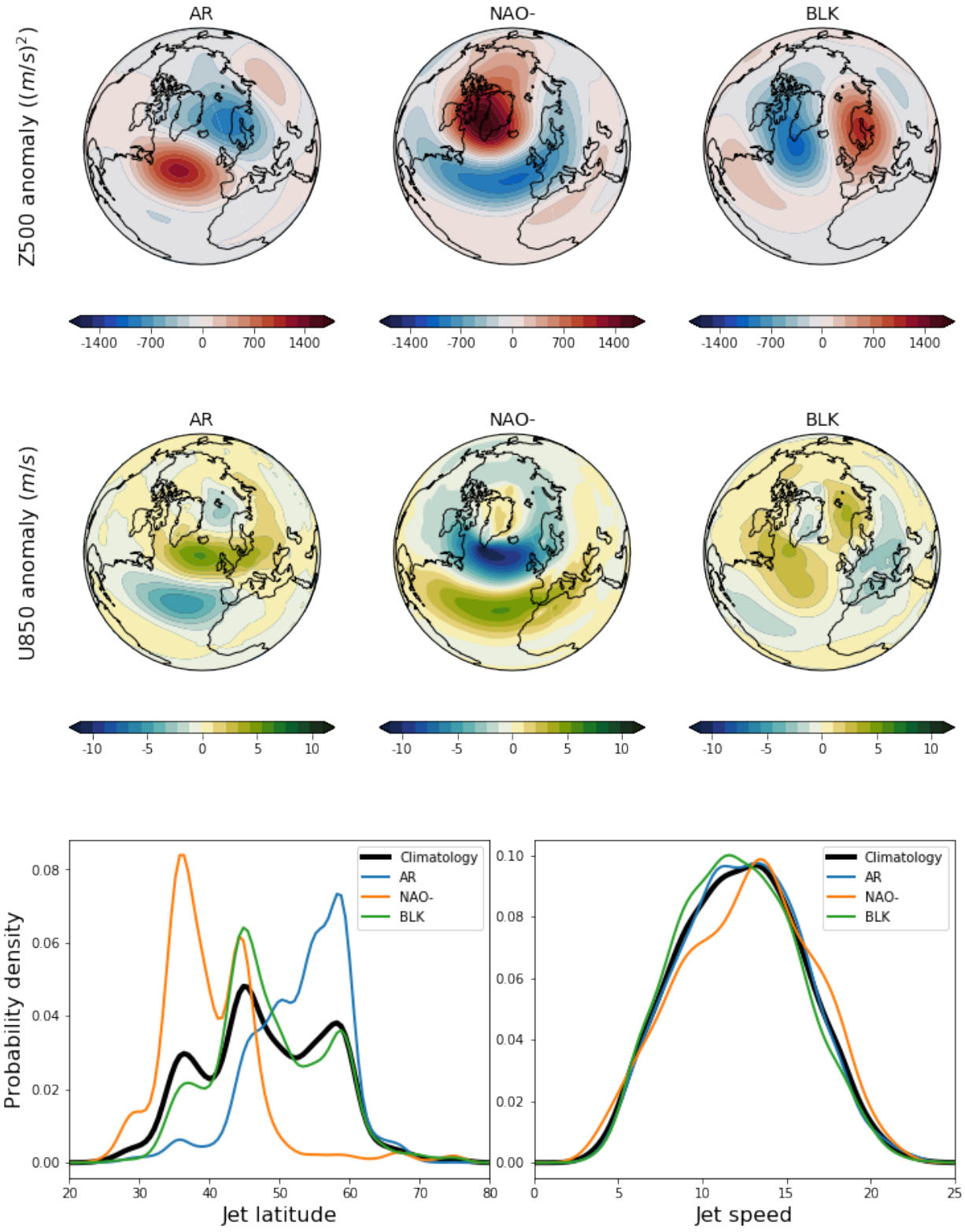


Figure 2.9: Regime composites for 3 geopotential jet regimes in ERA20C. Top: 500 hPa geopotential height anomalies. Middle: 850hPa zonal wind speed anomalies. Bottom left: distribution of jet latitude. Bottom right: distribution of jet speed.

regime is linked to a tilted jet, angled south-west to north-east, over Scandinavia. Despite having regressed out the linear influence of jet speed there are statistically significant, although modest, differences in jet speed distribution for some regimes as well. The NAO- state is associated with high jet speeds, which can be understood by the preference for strong Southerly jets merged with the upper level subtropical jet. The blocking regime conversely is associated with lower wind speeds, as is typical for a meandering jet.

To reiterate an important point, while the influence of jet speed has been removed from the phase space prior to clustering, the resulting regimes are by no means ignorant of the zonal wind field, which of course has important impacts on European weather and must affect the geopotential height field. Moving past the subtle jet speed differences between the regimes, there are large negative zonal wind speed anomalies in the local vicinity of the anticyclonic high in all three regimes. There is therefore no contradiction with the work of, say [Benzi et al., 1986], which emphasise the importance of *local* zonal wind speed for setting the conditions of blocking persistence. However, at the same time we acknowledge that we have no concrete theoretical model for why the influence of jet speed can be factored out – the conceptual connection between the jet and anticyclonic blocking is still imperfectly understood, and our approach has been primarily guided by observational data.

2.4.1 Introducing neutrality

Before we move forward with the validation and analysis of the three geopotential jet regimes which we have found to be stable, we wish to add one more refinement to the regime classification. It must be admitted that, despite their very high stability, three is a rather small number of patterns, and so a three-state understanding of the atmosphere cannot fully explain the synoptic variability of European climate. Indeed, not every day demonstrates anticyclonic blocking anomalies!

When it comes time to consider the temporal variability of regimes – their occurrence

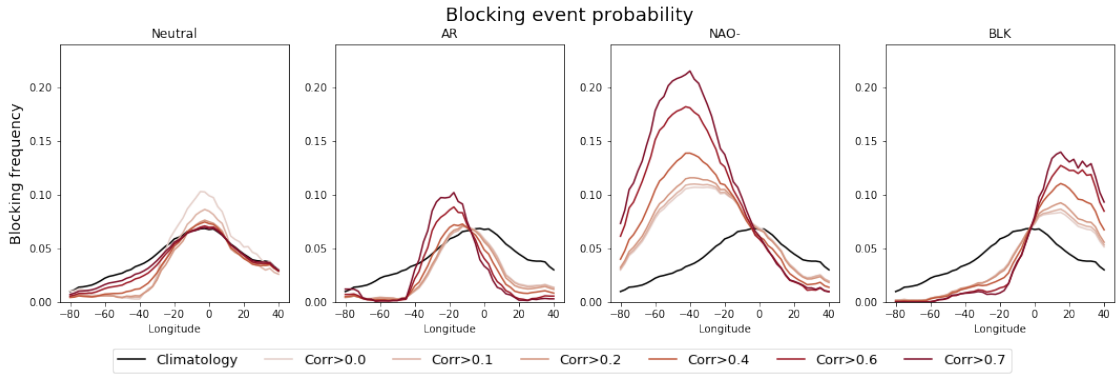


Figure 2.10: The black line shows the climatological latitudinal DJF average of the Davini blocking event index, over the domain [30-75N]. Coloured lines show the same, composited by geopotential jet regime, for a range of different neutral state thresholds. More stringent thresholds amplify the blocking frequency anomalies in the active regimes.

and persistence – it will be useful to introduce a neutral state: an assignment for days which show little similarity with any of the three regimes we have found. A skeptical reader could conceivably view this approach as quite ad-hoc, but in fact it is well rooted conceptually. Persistent blocks are, almost definitionally, rather exceptional events, and we should not expect their non-linearity to be apparent on all days, nor can we expect a clustering algorithm to separate out the continuum of linear, transient variability cleanly into a neutral state for us.

Therefore to define our neutral state, we compute the area weighted pattern correlation of every day's Z500 field with the composite of the regime it has been assigned to. If this correlation falls beneath a threshold that we must choose *a priori*, we consider the day to not truly resemble the supposed regime pattern, and assign it to the neutral state.

How should such a threshold be defined? There are many conceivable options, but here we choose to use blocking as a focus. Fundamentally we want our regimes to capture blocking events, and so we would like to set our threshold such that there are significantly more blocking events in the active regimes than in the neutral state.

We use the two-dimensional blocking event index of [Davini et al., 2012] to identify blocks in the full ERA20C dataset. This index detects blocks as reversals of the merid-

ional geopotential height gradient, with restrictions requiring a reversal to demonstrate temporal persistence and a large spatial extention before it is classified as a block. This acts to exclude transient Rossby waves and most ridging events. It has been used very satisfactorily to study blocking in a number of studies (e.g. [Drouard et al., 2021], [Davini et al., 2021] and [O'Reilly et al., 2016]). Figure 2.10 shows the distribution of blocking events in each of the three active regimes and in the neutral state, as a function of the correlation threshold used. Note that even for a threshold of 0.0, there are some days still in a neutral state, as around 2% of days have a negative correlation with all regimes. We see that the blocking regime and the negative NAO regime show levels of blocking much higher than climatology for all thresholds, whereas the Atlantic ridge does not for low thresholds. This is because it is a "ridge" not an Atlantic block, and the reversal of Z500 used to define the index is not always achieved. Nevertheless, we see that for thresholds exceeding 0.4, the Atlantic ridge regime also shows more blocking than climatology at longitude $\sim$10W, and the neutral state has a lower than average level of blocking at all longitudes. A threshold of 0.4 assigns 46% of days in ERA20C to the neutral regime, and we choose this value as our cutoff. Including the Neutral definition when analysing spatial patterns has little impact, increasing the amplitude of anomalies in regime composites, but not their shape, while the Neutral regime itself has no significant anomalies upon compositing. This provides additional reassurance that there is not an important regime state we have missed from our classification, as we would then expect this to show up in our neutral composite.

2.4.2 Links to existing regime frameworks

In order to quantify the relationship between the three geopotential jet regimes we find here with other regime frameworks we can make use of the coincidence matrix, C. The element $C_{i,j}$ of this matrix encodes the number of days assigned to a regime i under one clustering framework which are then assigned to a regime j under a different set of regime classifications. We start by looking at the coincidence with respect to the 4

classical circulation regimes both with and without a neutral state included:

$$\begin{array}{c} \\ BLK \\ NAO- \\ AR \end{array} \begin{array}{c} \begin{array}{cccc} BLK & NAO+ & NAO- & AR \end{array} \\ \begin{bmatrix} 2103 & 1785 & 65 & 229 \\ 264 & 128 & 1772 & 125 \\ 102 & 1045 & 200 & 1992 \end{bmatrix} \end{array}$$

$$\begin{array}{c} \\ Neutral \\ BLK \\ NAO- \\ AR \end{array} \begin{array}{c} \begin{array}{cccc} BLK & NAO+ & NAO- & AR \end{array} \\ \begin{bmatrix} 1265 & 1754 & 448 & 1139 \\ 1063 & 819 & 1 & 18 \\ 96 & 39 & 1539 & 16 \\ 1 & 404 & 58 & 1239 \end{bmatrix} \end{array}$$

We see that the naming convention used for our geopotential jet regimes is accurate, as each regime shares the most days in common with its corresponding classical circulation regime. Days classically assigned to the NAO+ regime are now primarily included in the BLK and AR patterns, the majority of which are then filtered out into the neutral state. We see that the neutral state also filters out many of the disagreements in regime assignment between the classical circulation regimes and the geopotential jet regimes, such as the 65 days assigned to NAO- in the 4 regime framework but to BLK in the 3 regime framework. A large number of classical BLK and AR days are also placed in the neutral regime: these are days whose circulation patterns did not resemble any of the 4 classical regimes particularly strongly.

We can do the same for the jet latitude regimes which are defined as Southern if jet

latitude < 40 degrees, Northern if latitude > 52 degrees, and Central otherwise:

	Central	*Southern*	*Northern*
BLK	2008	596	1317
NAO−	886	1483	110
AR	1156	159	2184

	Central	*Southern*	*Northern*
Neutral	2151	784	1671
BLK	1004	229	668
NAO−	460	1182	46
AR	435	43	1224

Again we see strong correspondences between the two frameworks, just as we saw in figure 2.9, which are enhanced by the introduction of the neutral state.

Despite these strong similarities the frameworks are certainly not fully equivalent to eachother. Importantly the neutral regime is not simply the NAO+ regime, or central jet regime, renamed. It has no clear Z500 anomaly pattern associated with it, and while it has a preference for central jet states, its probability distribution is still trimodal (not shown).

2.5 Geopotential jet regimes in the historical record

In the previous section we have, using the ERA20C reanalysis, explored the idea that removing the influence of jet speed from the Z500 field prior to clustering might allow us to find more robust regime patterns. We found that this hypothesis has borne fruit in the form of 3 anticyclonic geopotential jet regimes that can be found in the whole of the 20th century. However we must be careful when considering historical records of synoptic

variability, as the production of centennial reanalyses is a relatively new development, and observations prior to the satellite era can be quite sparse. Therefore, we would like to test whether the regimes we have found were merely an artefact belonging to the ERA20C reanalysis, by considering other long reanalysis products. As detailed in section 2.3 we look at 4 additional reanalysis products from two centres: ERA5 and CERA20C from ECMWF and 20CRv2 and 20CRv3 from NOAA-CIRES-DOE.

We will compute geopotential jet regimes for these additional reanalyses using the method that we have developed through study of ERA20C. Summarising this process, for each reanalysis we shall:

- Download daily Z500 over the domain [80W-40E,30-90N], and U850 over the domain [80W-0W,15-75N], both at 1 degree horizontal resolution.
- Compute daily jet speed from U850 as described in section 2.3.
- Compute the first four PCs of Z500, and regress out the jet speed from each PC, leaving four residual PCs.
- Use K-Means clustering to compute clusters in the space of residual PCs for different cluster numbers.
- Label each day of the Z500 and U850 fields with the cluster assignment found using the residual PCs.
- Use these labels to compute mean cluster composites of the full fields.
- Assign any days whose Z500 field have a correlation lower than 0.4 with their corresponding regime composite to the Neutral regime, when computing occurrence and persistence statistics.

  Therefore, despite identifying regimes in a very truncated residual phase space, by returning to the full space to produce composites we ensure the interpretability of the regime patterns we identify. To compute subsampled regime composites - in

order to assess regime stability - this entire procedure is repeated separately for each subset of Z500 and U850 data.

2.5.1 Stability of regimes

We start by calculating stability metrics for all the reanalyses just as we did for ERA20C, using 30-year windows and a 10-year sampling rate. This means that ERA5 has 4 subsamples of data, 20CRv3 has 6, and CERA20C, 20CRv2 and ERA20C have 9 subsamples. Maintaining the same degree of overlap between different samples in all reanalyses was prioritised over producing more subsamples for the shorter reanalyses. The fact that the different reanalyses cover slightly different time periods is not a problem here, because the basic principle of the metric is to assess the stability of regime patterns to different time periods of data.

In the upper panels of figure 2.11 we see the results for both classical circulation and geopotential jet regimes, where now we have left off the individual correlation of each subsample for clarity of presentation. Looking at the mean stability over all reanalyses, we see broadly similar results to our analysis of ERA20C. The NAO dipole obtained with K=2 is clearly the most stable classical circulation regime, with a stability of 0.96. A second less pronounced maxima appears at K=7, which could be seen in figure 2.8 but which is now much clearer. However for geopotential jet regimes we see the stability is in general much higher, with a minima for K=4 of 0.87, and clear peaks at K=3 and K=7. The three regime approach reaches a stability in excess of 0.99, but even the seven regime stability is equal to 0.95, comparable to the NAO dipole for classical circulation regimes. We see that there are considerable variations in stability between the reanalyses for different K values, with ERA5 being a particular outlier in many cases. However there is remarkably little disagreement between reanalyses in the high stability of 3 geopotential jet regimes, which are the most stable in all cases.

As ERA5 is the only reanalysis used to only cover 1950 onward, its outlying behaviour

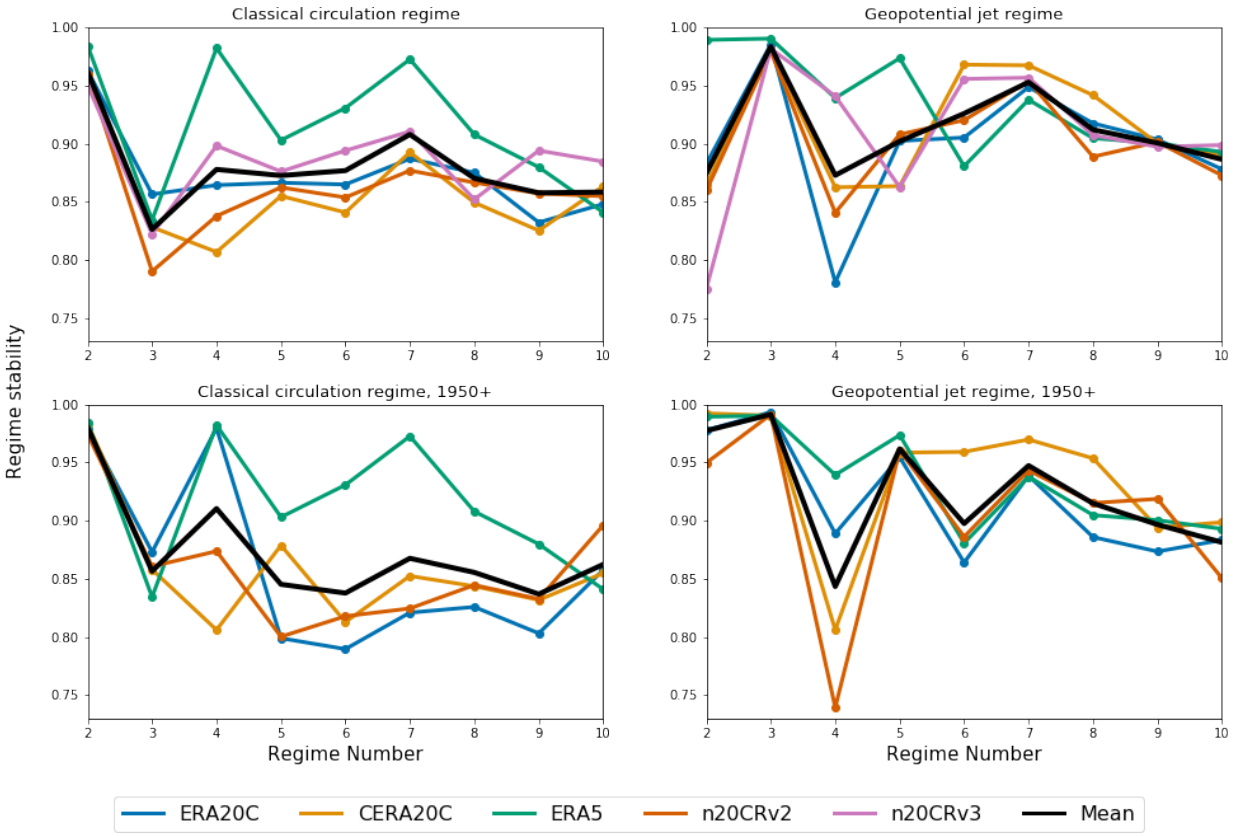


Figure 2.11: Multi-reanalysis regime stability calculated using 30-year subsamples with ten-year intervals. Upper panels show stability calculated using all available windows for each reanalysis. The lower panels show stability found using only the 4 30-year subsamples between 1950-2010 for each reanalysis, and excludes n20CRv3.

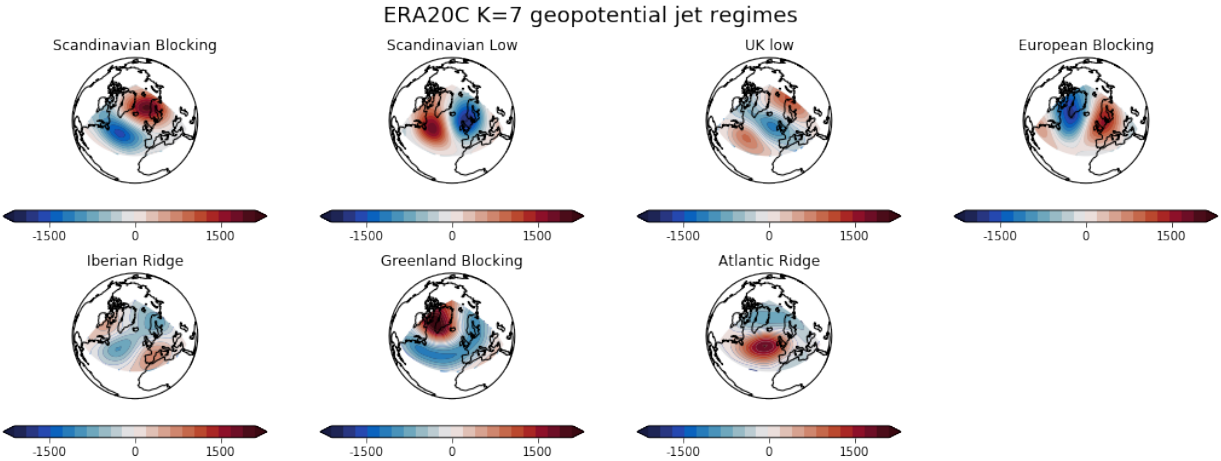


Figure 2.12: Geopotential height anomaly composites for 7 geopotential jet regimes.

bears closer examination. In order to determine whether this behaviour is simply a consequence of the smaller sample size, or is perhaps capturing a change in atmospheric stability in the recent period, stability was recalculated for all reanalyses, excluding n20CRv3, using only data for shared period 1950-2010. This is shown in the lower panels of figure 2.11. ERA5 is still an outlier compared to the other reanalyses in this shared period, showing a much greater stability for classical circulation regimes between K=4 and K=8, and does not exhibit as substantial a decrease in stability for four geopotential jet regimes, although it is now more closely mirrored by ERA20C. Overall the reanalyses are less coherent in this period, due to the smaller sample size, but the high stability of K=2 classical circulation regimes, and K=3 geopotential jet regimes remains. Five geopotential jet regimes also seem more stable in this period, as was found in [Dorrington and Strommen, 2020]. These results rule out a radical difference in regime stability between the two half-centuries. Sampling variability may play a factor, as we see from ERA20C and n20CRv2 for 4 classical and 4 geopotential jet regimes respectively, that large changes in stability can occur when decreasing sample size. However another factor might be the greater quantity of observations integrated by ERA5 than the other reanalyses considered.

The geopotential anomaly composites associated with the seven geopotential jet regimes, are shown in figure 2.12 for ERA20C (the composites for other reanalyses are visually

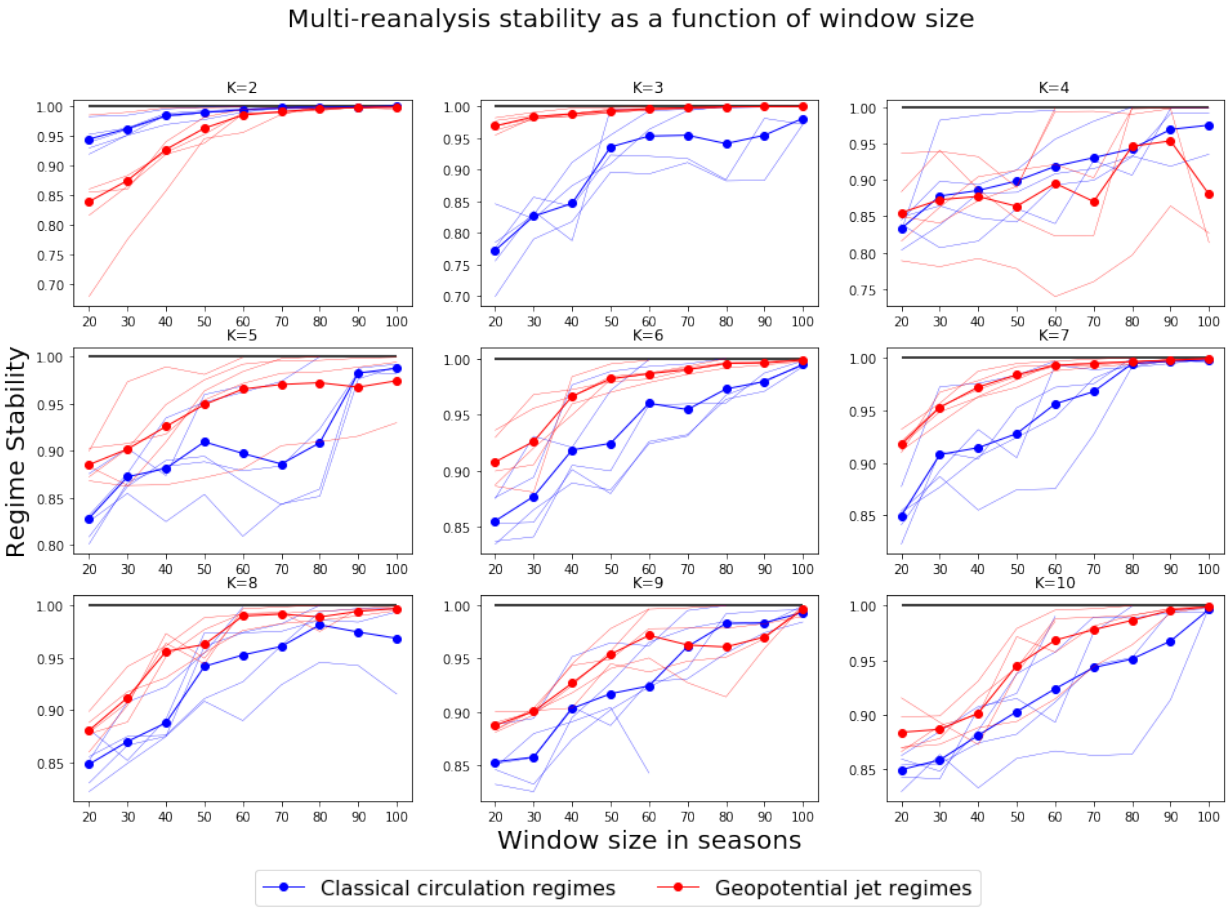


Figure 2.13: Multi-reanalysis regime stability presented as a function of window size. Thin lines show the stability for each reanalysis product, while thick dotted lines show the mean of all reanalysis products with at least as many years of data as the size of the window.

indistinguishable). Five of these seven regimes match quite closely to the year-round seven regime picture presented in [Grams et al., 2017b], with the Iberian ridge and UK low not well mapping onto the Atlantic trough and Zonal regime found there. There are also similarities to the six regimes of [Falkena et al., 2020], but again with no real analogue of their zonally symmetric cyclonic states AR- and NAO+. In this thesis we will focus on the more stable and conceptually simpler 3 regimes, but the increased detail a 7-regime picture provides might be more useful for evaluating impacts, potentially without the need for a neutral state, and will be revisited in future work.

In obtaining these stability results we have made a choice of using 30-year windows, which was done to provide as many completely independent subsamples as possible

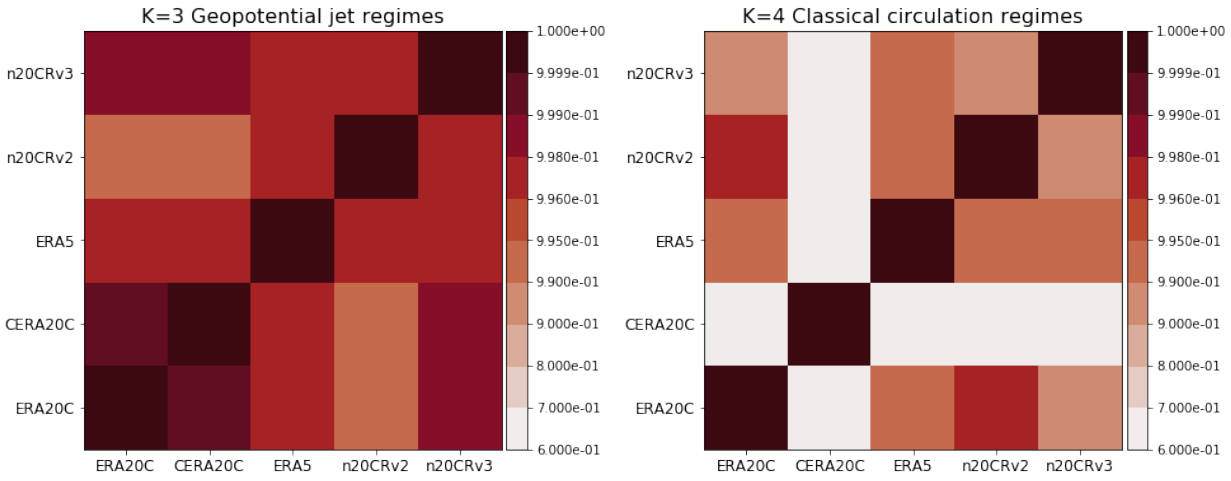


Figure 2.14: Regime pattern correlations between the different reanalyses in the shared period 1950-1980.

while maintaining a long enough number of DJF seasons to form a climatological sample. We analyse the sensitivity of our results to this choice, by repeating the stability analysis for subsample sizes of 20 years through to 100 years, always with a 10 year subsample rate. This means that for the 110 year ERA20C reanalysis, we have 10 20-year subsamples, and only 2 100-year subsamples. Figure 2.13 shows that, as we might expect, increasing the size of subsamples in general increases regime stability. For some cluster numbers there is some non-monotonicity in the stability change with window size, especially at high window sizes where the number of samples is lower. However for those cluster numbers considered to represent regimes (K=2 and K=4 for classical circulation regimes, and K=3 and K=7 for geopotential jet regimes) we see the results are much clearer. This makes intuitive sense, as if a regime framework is physically meaningful, we would expect it to be less subject to random sampling variability. The most striking result is that even with 20 year subsamples, K=3 geopotential jet regimes show a stability of 0.96, which even the 100 year samples fail to reach for 4 classical circulation regimes!

Finally for our treatment of stability, we restrict our attention to only the time period shared between all 5 reanalyses, 1950-1980, to understand how similar the regime patterns in the five reanalyses are to each other in the same historical window. We focus only on the 3 geopotential jet, and 4 classical circulation regimes, and look at the pat-

tern correlation of each reanalyses' regimes with each of the others, averaged across all regime patterns. The results are shown in figure 2.14 where we note the colour bar has been extended in order to distinguish high correlations from each other. We see that the regime pattern correlations are not necessarily closest between reanalyses from the same centre, and that as we may now have expected, the geopotential jet patterns are much more similar than the circulation regime patterns, with a minimal correlation of 0.994 between n20CRv2 and ERA20C – completely indistinguishable visually. In contrast, CERA20C wildly disagrees with the other reanalyses on the correct classical circulation regime patterns, with correlations of only 0.6-0.7.

We may now feel quite comfortable that the 3 geopotential jet regimes are indeed radically more stable than the classical circulation regimes, and will be well suited for studying differences between models. However, we are not quite done, as there has indeed been variability in the 20th century North Atlantic atmosphere: in the NAO index and predictability [Weisheimer et al., 2017], in Greenland blocking frequency [Hanna et al., 2016], and in rates of sudden stratospheric warmings [Butler et al., 2017] among other metrics. In our framework, since we have no spatial variability, then these documented changes must result in changes in the temporal variability of regimes – changes in regime persistence and occurrence. We therefore finish this section by analysing the temporal regime variability in reanalysis. This has never been done in a multi-reanalysis context before and so will be very useful for understanding the uncertainties in the historical record, and will inform our assessment of model regime variability.

2.5.2 Temporal regime statistics

Regime occurrence and persistence are computed for the three geopotential jet regimes, using the neutral state introduced in section 2.4, for all reanalyses as defined in section 2.3. In order to understand the variability of these metrics, they have been computed in 30-year rolling windows, which allows decade-to-decade variations to be easily identi-

fied. We also compute regime occurrence and persistence for 30 years randomly sampled from each reanalysis, which is repeated 400 times in order to build a bootstrapped distribution of 30-year metrics. These distributions represent the null hypothesis that there is no year-to-year memory in the temporal statistics, and so any variability in the historical record outside these distributions' bounds can be considered evidence of low-frequency interannual regime dynamics.

Figure 2.15 shows the evolution of regime occurrence over the 20th century, with the bootstrapped distributions shown with boxplots, while figure 2.16 shows the evolution of regime persistence.

A first key point to note is that the reanalyses can disagree with each other quite substantially. Looking at occurrence, this is particularly noticeable in the early century where ERA20C estimates more Neutral days and less Atlantic ridge days than other reanalyses. It is also apparent right up to present day, where between 1980-2010 ERA5 estimates a BLK occurrence approximately 4% higher than ERA20C. The impact on the early century is even more pronounced in persistence, with n20CRv3 showing very low NAO-persistences in the early century that then rapidly increase, reaching parity with other reanalyses by 1940, and ERA20C showing a Neutral persistence about 3% higher than other reanalyses between 1900 and 1950. It is somewhat intuitive that we see these disagreements more strongly in persistence due to the known model deficiencies in maintaining persistent blocking events, and the stronger reliance on the model in the more poorly constrained early period. The disagreements are primarily an issue of mean differences in each reanalysis, with the actual occurrence and persistence time series highly correlated from 1950 onwards (i.e. after the anomalous trends are reduced). Therefore it seems feasible that the multireanalysis mean should represent a good approximation of truth. However this uncertainty suggests we must be careful about over-interpreting small differences in regime persistence and occurrence, both in model vs reanalysis and model vs model contexts, as they are clearly not well constrained by current state-of-the-art

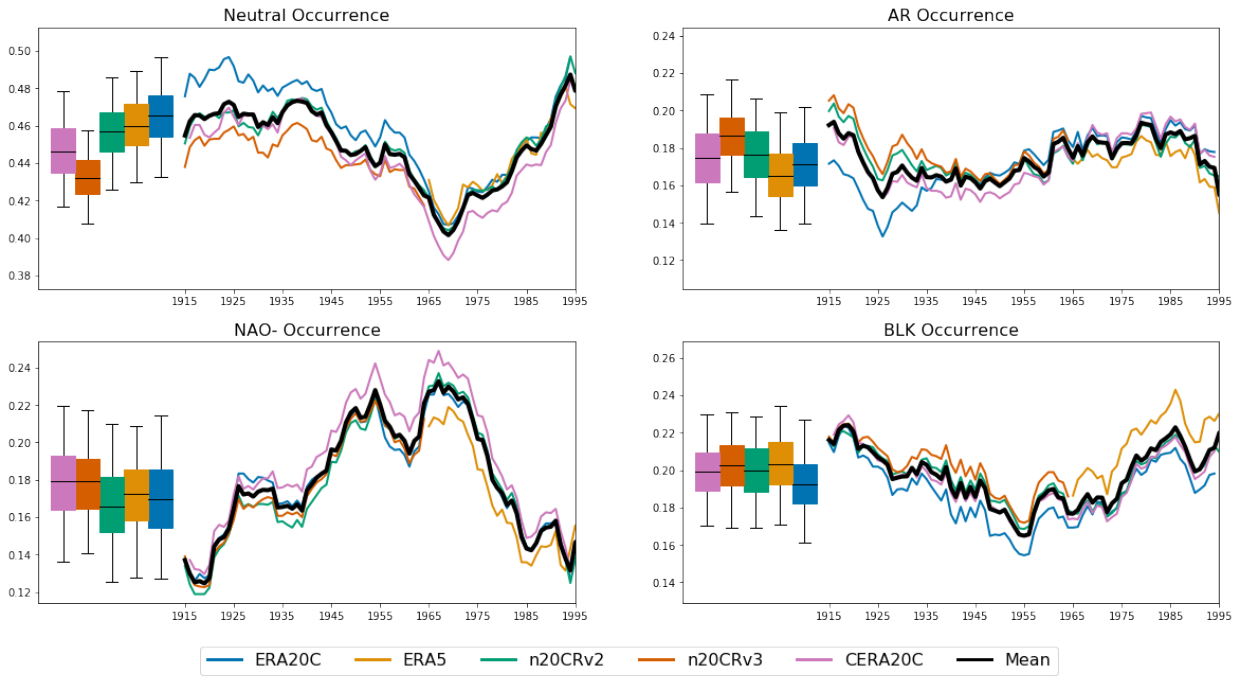


Figure 2.15: 30-year rolling windows of regime occurrence, computed for the three geopotential jet regimes and their neutral state for a range of reanalyses. Dates mark the central year of the rolling window, e.g. 1995 is a mean from 1980-2010. The black line shows the mean of reanalysis occurrences for all reanalyses with data for the full 30-year window. Boxplots show bootstrapped distributions of regime occurrence from randomly sampled years (see main text for details), with whiskers marking the 2.5th and 97.5th percentiles. All subplots have the same range on the y-axis of 14%, so magnitude of variations are comparable.

reanalysis.

Putting aside now the differences between the reanalyses, and focusing on commonalities, we see a strong peak of NAO- occurrence in the mid-century (central dates of 1948 - 1976) which is in agreement with [Weisheimer et al., 2017] and [Hanna et al., 2016], and which falls outside of the stationary occurrence distributions for an extended period of time, indicating significant low-frequency variability. There is also a dip in Neutral occurrence, but which occurs slightly later; significant between 1960 and 1980, indicating a period of generally more active regimes. No such significant variations can be seen in the AR or BLK regimes, although there is a general decrease in BLK during the mid century. With the exception of the non-stationary persistence in the early century already discussed, there is no significant low frequency variability in persistence.

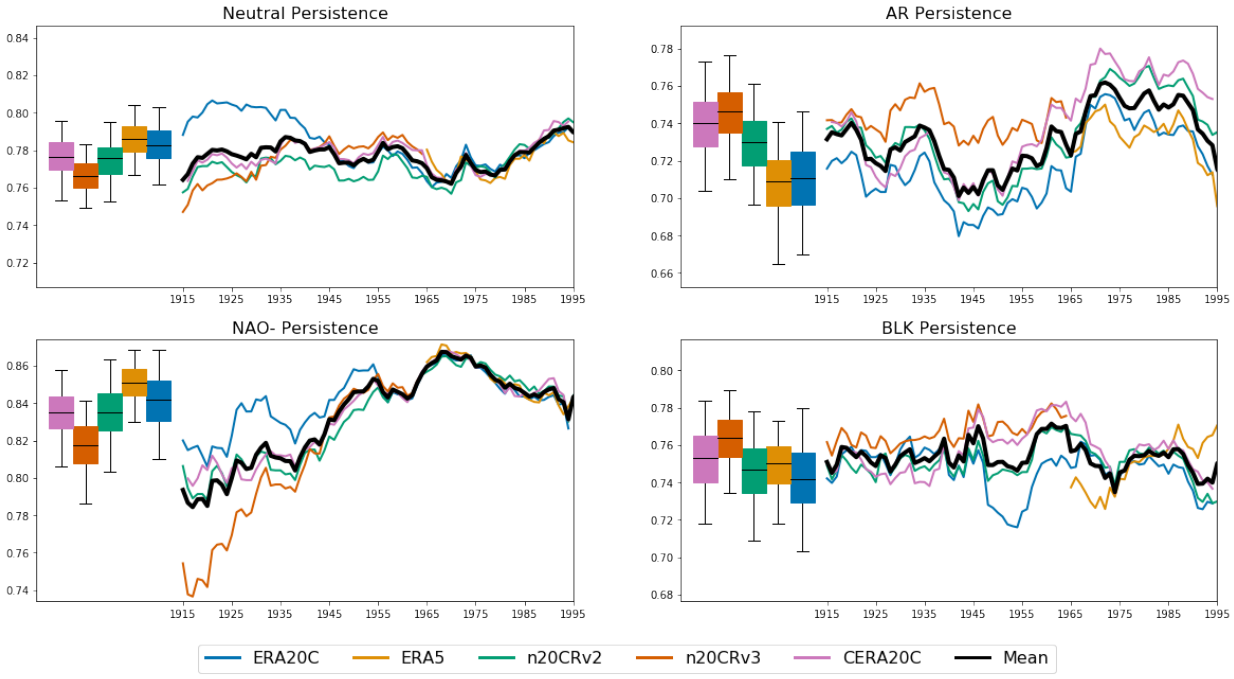


Figure 2.16: As in figure 2.15 but showing regime persistence.

2.6 Summary of geopotential jet regimes in reanalysis

In the preceding half-chapter we have tackled the thorny issue of identifying stable and reproducible regime structure in the Euro-Atlantic circulation. We have essentially solved this issue. Our hypothesis was that jet speed variability has no impact on regime dynamics and yet nevertheless introduces linear variability into the principal components of Z500, confounding regime identification. By regressing out the jet speed prior to clustering we saw dramatic increases in regime stability, and we have tested the robustness of that result to different cluster numbers, subsample sizes and reanalysis products. Terming these "geopotential jet regimes", we have identified a clearly preferred number of regimes, K=3, as well as identifying a secondary preference for a more expansive set of 7 regimes, which might be particularly useful for applications.

We have shown how the three geopotential jet regimes relate to classical circulation regimes and to the trimodal jet stream, as well as demonstrating that they capture periods of enhanced blocking frequency, when a suitable neutral state is introduced. We

have characterised the historical regime variability in the 20th century and shown that there are considerable uncertainties in persistence and occurrence statistics. In the early period, prior to 1950, there are disquieting trends in quantities such as NAO- persistence, which differ between reanalyses. In the satellite era the variability of regime dynamics is broadly agreed on by all reanalyses, but the mean values of regime persistence especially can diverge by $\sim$2% percent. We there have shown that extraordinary care must be taken when asserting differences in regime statistics between reanalysis and models, or between historical periods. However due to the high regime stability now demonstrated, comparisons between the spatial patterns of geopotential jet regimes in different datasets should be very robust. The stability behaviour of ERA5 was found to be quite different to the other reanalyses, and this could not be explained by the different time period it covered alone. An open question remains as to whether this is a chance feature, or a consequence of assimilating a greater range of observational data. Future work would benefit from including additional reanalyses that include satellite and upper atmosphere data for comparison.

2.7 Geopotential jet regimes in SPHINX

Now that we have introduced and developed the concept of geopotential jet regimes using reanalysis data, we can make use of this new regime framework to analyse regime dynamics in model data. Specifically, we are now in a strong position to analyse the impact of stochastic parameterisations on the regime representation of climate models, confident that any differences we observe are not a result of instability in the clustering methodology.

As discussed in section 2.3, the SPHINX project produced 3-member ensemble climate runs for both deterministic and stochastic versions of the EC-Earthv3 model, run from 1850-2100. Here we limit our attention to the historical period 1900-2010, providing a total of 330 years of model data, ignoring the first 50 years in order to ensure the model

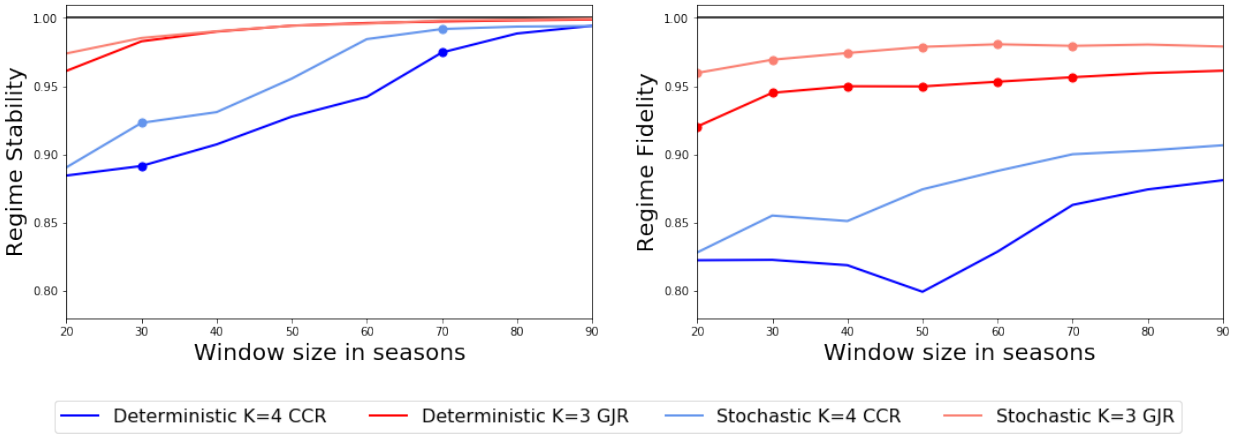


Figure 2.17: Left: Regime stability and Right: regime fidelity of SPHINX data as a function of subsample window size for both classical circulation regimes (CCR) and geopotential jet regimes (GJR). Dots mark statistically significant differences between deterministic and stochastic configurations at the 5% level, calculated using a Kolmogorov-Smirnov test on the distribution of subsamples.

is well converged to its climate attractor, and allowing comparison to reanalysis regime statistics starting in 1900. As we have seen, interdecadal variability in regime dynamics can be quite large, and so this unique experiment provides an excellent opportunity to cleanly compare the impact of stochasticity on regimes with samples sufficiently large to average out slow climate variability.

2.7.1 Spatial representation of regimes

We start by looking at the representation of regime spatial patterns in SPHINX, both in terms of their stability and also now their fidelity: the degree to which the regime patterns in the model resemble the patterns found in reanalysis data. Due to the high pattern correlations between different reanalysis regime patterns shown in figure 2.14, the fidelity calculation provides equivalent results regardless of which reanalysis product we compare against. Therefore, for simplicity, all fidelity results shown have been calculated relative only to ERA5.

To compute the metrics all three ensemble members of the stochastic and deterministic runs are concatenated together to provide a 330-year time series, from which subsamples

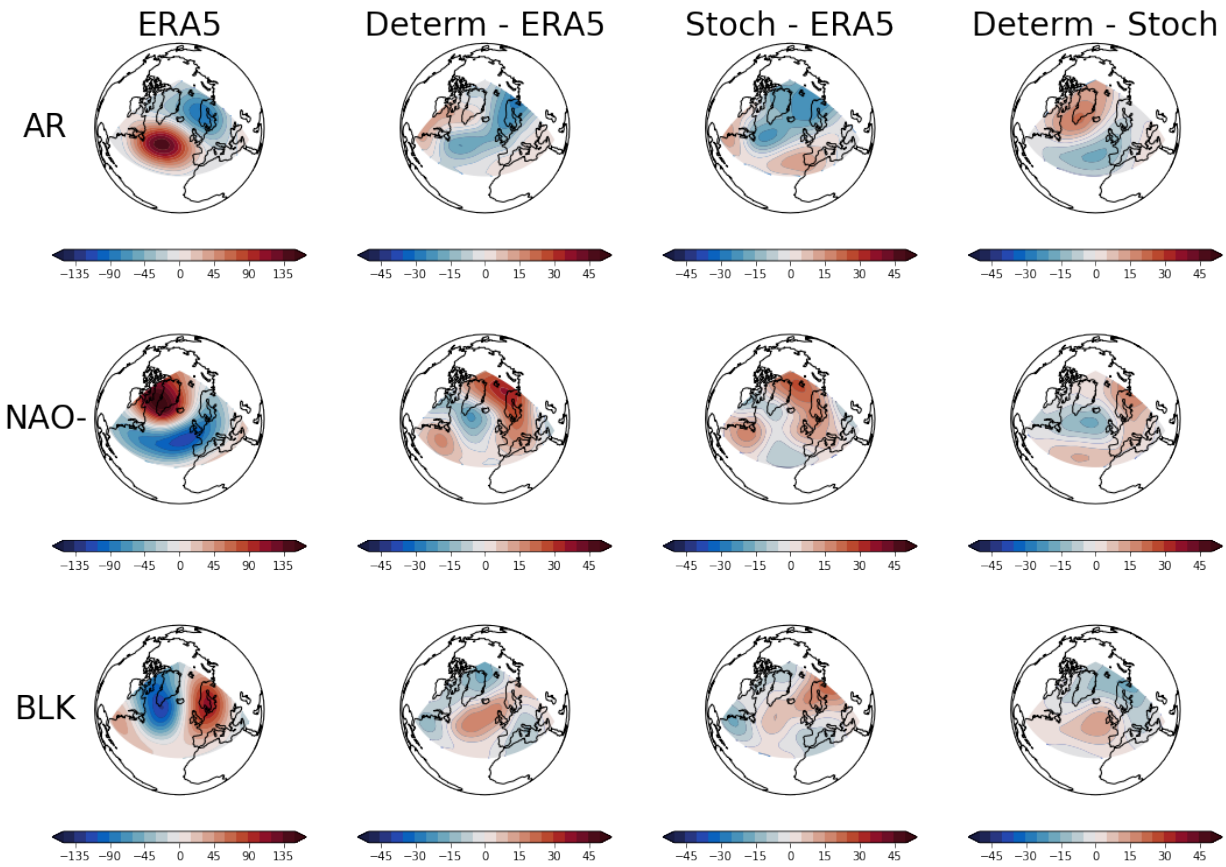


Figure 2.18: Left: Z500 regime composites of ERA5 in units of metres. Middle: difference between SPHINX composites and ERA5 composites. Right: difference between deterministic and stochastic SPHINX composites.

are extracted, just as for reanalysis in the previous sections. Figure 2.17 shows the resulting regime stability and fidelity for a range of subsample window sizes, both for the 4 classical circulation regimes and the 3 geopotential jet regimes. Significance at the 5% level was assessed for each window size separately, using the Kolmogorov-Smirnov test to compare the stochastic and deterministic distributions of subsample pattern correlations.

We can clearly see that in the model, just as in reanalysis, the geopotential jet regimes are more stable than the classical circulation regimes. There is no consistently significant increase in stability for either regime paradigm as a result of introducing stochasticity, although for 30 and 70 year window sizes, marginally significant increases in stability are seen for the classical circulation regimes.

The fidelity of the geopotential jet regimes shows a very clear improvement, with significantly higher pattern correlations between the stochastic model and ERA5 regime patterns for a wide range of window sizes, by $\sim$0.02 in most cases, and as much as $\sim$0.05 for the shortest 20 year window. This is in agreement with the improved spatial representation of classical circulation regimes in stochastic simulations found in figure 8 of [Dawson and Palmer, 2015] (hereafter D15), now validated using a much longer time period than the 45 years of data used there. We have also explicitly assessed the impact of interdecadal variability and subsample size, adding confidence to our conclusions.

The fidelity of the classical circulation regimes also seems notably higher but does not quite reach significance at the 5% level, as the spread of fidelity between different subsamples (not shown) is quite high for the 4 regimes. This does not contradict the finding of D15 as though they reported improved regime patterns there, they were unable to quantify the sampling uncertainty of their results.

As the high fidelity would suggest, the actual biases in regime pattern are quite small, even in deterministic SPHINX. The anomalies between Z500 composites of ERA5 and SPHINX (now calculated using the full 330 years of each dataset), as well as between

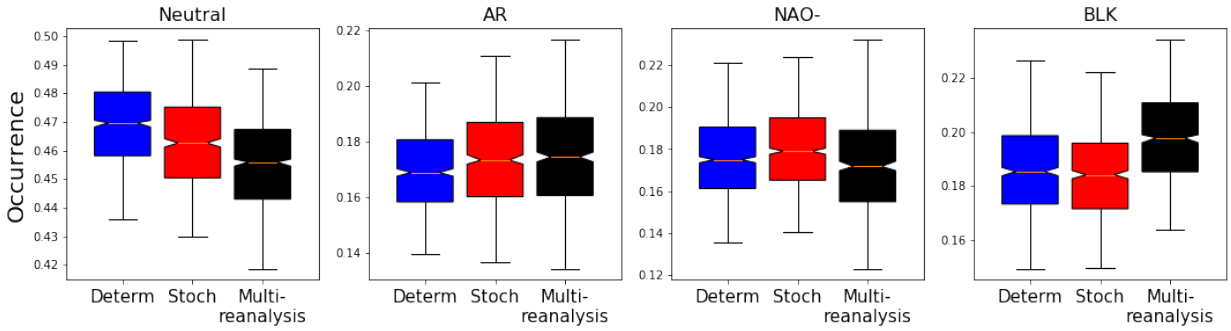


Figure 2.19: Regime occurrence for SPHINX and reanalysis. Boxplot distributions are produced from 1000 30-year bootstrap samples. Multireanalysis boxplots are produced through bootstrapping of the reanalysis ensemble mean annual occurrence, restricted to 1950-2010. Whiskers mark the 2.5th and 97.5th percentiles, notches indicate the 95% confidence interval of the distributions' median.

deterministic and stochastic SPHINX runs, are shown explicitly in figure 2.18. The deterministic runs suffer from an insufficiently strong and less zonally localised anticyclonic high in both the Atlantic ridge and NAO- regimes, as shown by negative anomalies over the anticyclones flanked by weaker positive anomalies. Further, the dipole of the BLK regime, oriented almost perfectly zonally in ERA5 is rotated and displaced slightly South West in deterministic SPHINX runs. The nature of the biases is qualitatively the same in the stochastic runs, but in general even more subtle – small shifts towards the ERA5 regime patterns can be seen for the NAO- and BLK patterns in particular.

Clearly then, the introduction of stochastic perturbations is detectably improving regime structure, through small refinements in the orientation and extent of the geopotential anomalies.

2.7.2 Temporal representation of regimes

Having seen that stochasticity improves the regime patterns, it is now natural to ask whether it also improves the regimes' temporal statistics. We show the regime occurrence statistics in figure 2.19 and the daily transition probabilities in figure 2.20 for the stochastic and deterministic runs as well as for reanalysis. Distributions have been computed through 1000 bootstrap samples of 30 DJF seasons, in order to demonstrate how

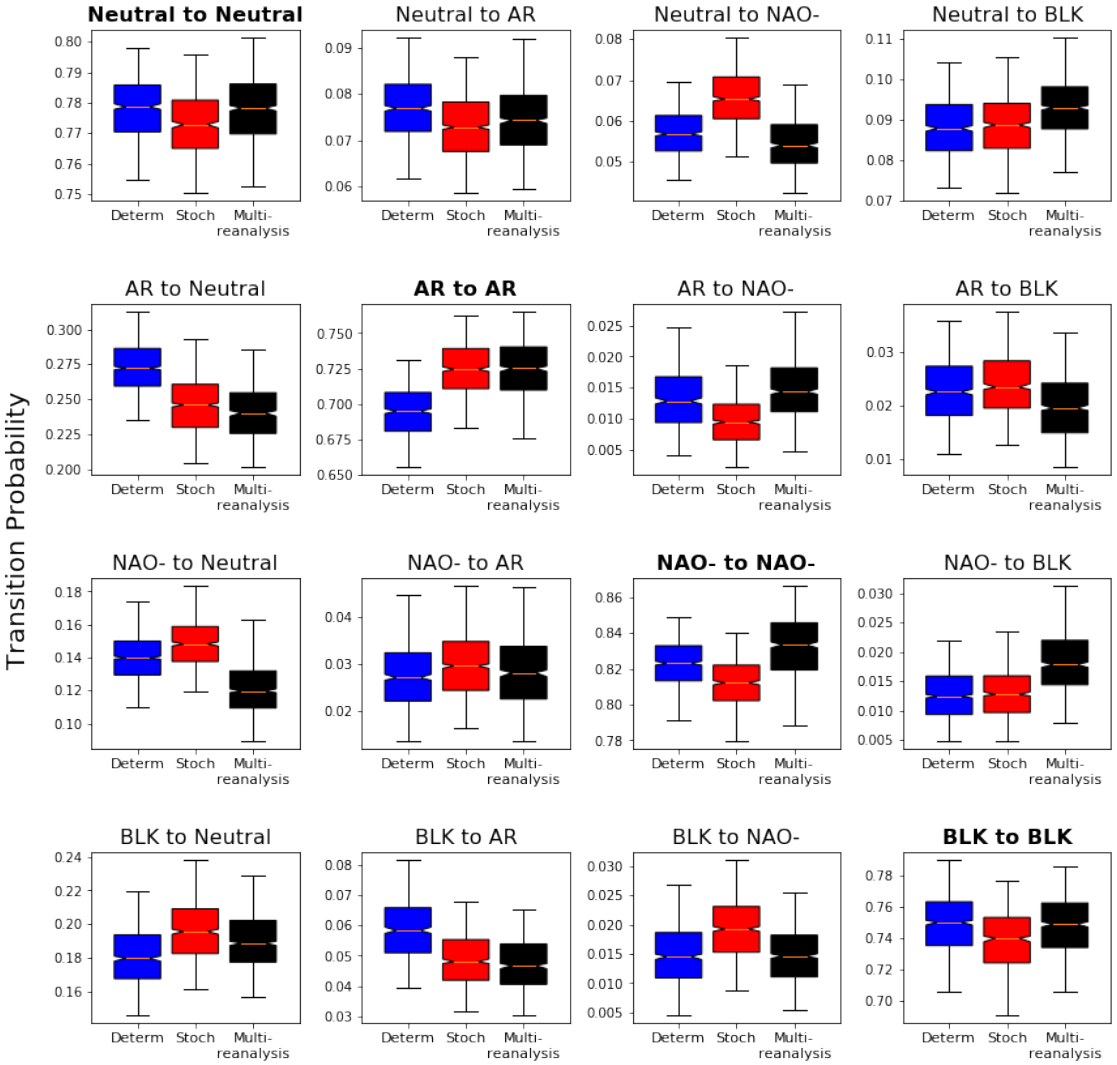


Figure 2.20: As figure 2.19 but showing daily regime transition probabilities.

Regime	Pattern changes	Occurrence	Persistence	Transition probabilities
Neutral	-	Decreases (towards reanalysis but still significantly higher)	Decreases (away from reanalysis)	Transitions to NAO- increase substantially (away from reanalysis)
AR	Ridge strengthened and increased meridional orientation of dipole	Increases (matching reanalysis)	Increases (matching reanalysis)	Transitions to NAO- and Neutral less often (away from/matching reanalysis respectively)
NAO-	Amplitude of anomaly increased and more zonally localised high	Increases (away from reanalysis)	Decreases (away from reanalysis)	Transitions to Neutral and AR increase (away from reanalysis)
BLK	Dipole is rotated, increasing zonal orientation	No significant change	Decreases (away from reanalysis)	Transitions to Neutral and NAO- increase (away from reanalysis) while transitions to AR decrease (matching reanalysis)

Table 2.2: A summary of regime changes in the stochastic SPHINX runs relative to the deterministic runs.

if only a small time period was used, as is sometimes the case in the literature, different results might be obtained. Our reanalysis distribution is based on the mean of all reanalyses for a given year, restricted to 1950+. This restriction has been made as reanalyses are in closer agreement in this period (c.f. figure 2.15), and so we have more confidence in our real-world reference. However, plots made using the full 1900-2010 reanalysis period are qualitatively identical.

Significant changes in regime occurrence emerge in the stochastic runs. In the case of the Neutral and AR regimes this is a movement towards reality, with the overfrequent Neutral bias reduced – implying a more active regime structure overall – and the underfrequent AR bias eliminated entirely. However, not all is to the good, with amplified biases in NAO-, which occurs too frequently, and with no improvement in the too infrequently occurring BLK regime.

In terms of regime lifetime, the deterministic SPHINX runs have a very underpersistent AR regime, a bias which is eliminated completely in the stochastic runs. However, unlike in the work of chapter 1 we see that stochasticity does not always serve to stabilise the regimes, exacerbating the NAO- underpersistence bias and even introducing such a bias to the Neutral and BLK regimes. There are likewise significant shifts in the transition probabilities, which on the whole move the system away from the reanalysis probabilities as often as they do towards them.

These results are summarised in the table 2.2. Clearly the impact of stochasticity on regimes is complex and asymmetrical between different regime patterns, which is not necessarily surprising in a complex nonlinear system. Stochasticity *can* improve regime persistence, as seen for the Atlantic ridge, but is by no means guaranteed to do so, and here the general trend is towards decreased regime persistence in the stochastic runs. We note the importance of the large time period of data used here, as in most cases there is substantial overlap between the model and reanalysis boxplots. Broadly speaking, we can see that differences of 2-3% in occurrence or persistence found in 30-year data samples should not be viewed as robust evidence of model bias. Note that this error compounds the uncertainty in the true observational occurrence and persistence discussed above.

These results are also consistent with D15, which found equally mixed impacts of stochasticity on the persistence of classical circulation regimes. They found unchanged or slightly degraded regime persistence in stochastic high resolution T511 simulations, while they saw a less persistent NAO+ but more persistent BLK and NAO- regime when adding stochasticity into a lower resolution T159 simulation. Our SPHINX runs, at T255, fall roughly halfway between these two resolutions, and see only slightly less pronounced impacts of stochasticity than for T159. Notably however the AR regime is largely unchanging in persistence in D15, whereas here it is the only substantially improved regime.

It seems likely that the impacts of stochastic perturbations on regime representation are highly model- and resolution-dependent; a consequence of the stochasticity being integrated through parameterised processes, which are heterogeneous across models. In a coupled GCM, it can be very difficult to trace back exactly why changes in the model state occur, especially if, as here, the changes are quite small overall. However by considering where parameterisation schemes might be most active, and through analysis of the changing Euro-Atlantic mean state, shifts in the jet, and the plausible remote impacts of tropical dynamics, we will build some conjectural arguments to explain the observed regime changes, and consider whether any similarities to the mechanism of stochastic

persistence identified in the CdV model can be seen.

2.7.3 Parameterisation schemes in the mid-latitudes

One impediment to tracking the direct impact of stochastic perturbations in the SPHINX runs is that the individual parameterisation tendencies perturbed by the stochastic scheme have not been archived. While the SKEB scheme perturbs the rotational wind field directly, in the IFS atmosphere model SPPT is by far the dominant stochastic term, and so to first order regions with large net parameterised tendencies will be those regions experiencing the direct impact of stochasticity. Without this tendency output we cannot directly assess which regions or processes are driving changes in the stochastic model's circulation, and which are merely responding to them. Nor can we see in which locations the direct impact of stochastic parameterisations is largest.

However, as part of the Year of Tropical Convection (YoTC) [Waliser et al., 2012], parameterised tendencies from a small number of IFS operational forecasts *were* saved, and these archived tendencies provide a best guess for the direct role of parameterised processes in the SPHINX EC-Earth runs.

The YoTC dataset we analyse here consists of 3-hourly accumulated tendencies of T850 for a number of parameterisation schemes, produced for the first 24 hours of 19 operational forecasts, using operational IFS cycle cy36r1 for DJF 2009/2010. This is a relatively similar model cycle to the cy36r4 used in EC-Earth, and so is a reasonable comparison point, although parameterisations are tuned slightly differently in the two cycles. There are some limitations in using this dataset for our purposes. Firstly, there are only 114 tendency fields available, all sampled from a single winter season, so sampling bias could potentially be significant and interannual variability is neglected completely. Secondly, the output is from an initialised atmosphere-only forecast and so is close to the real-world climate attractor, while the coupled and free-running SPHINX runs are on their own biased model attractor. Thirdly, T850 is not the optimal variable to consider for

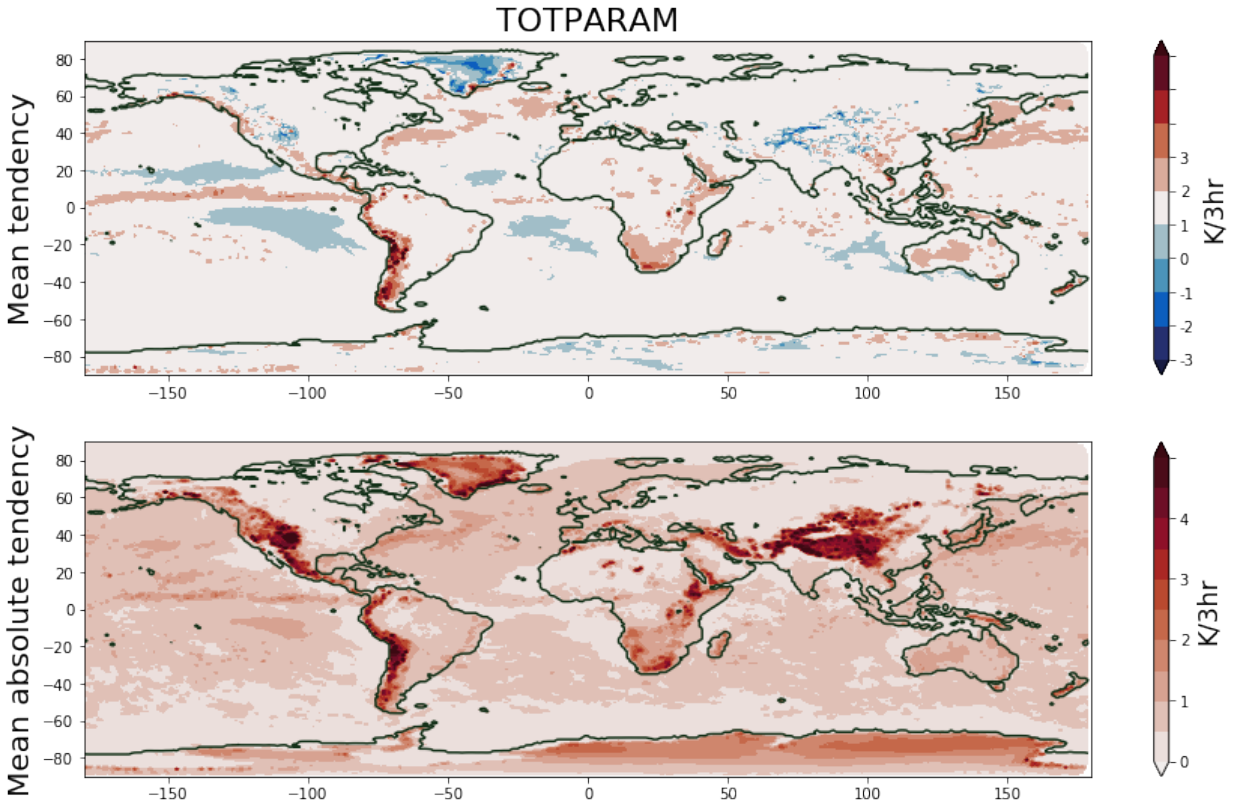


Figure 2.21: Top: The mean parameterised 3hr tendency of T850 from the YoTC project (see main text for details). Bottom: As top, but showing the mean absolute tendency.

understanding the forcing of the large scale circulation. However, these are the only such data readily available, and we may still be able to glean qualitative insights from them. For any future analysis of the impacts of stochasticity on atmospheric models it would be very valuable to archive these parameterised tendencies as a matter of course.

As described in [Christensen et al., 2015b], the IFS contains 4 families of parameterisations, the details and subcomponents of which are fully documented in [ECMWF, 2010]. CONV represents the impact of unresolved atmospheric convection, TGWD represents the turbulent boundary layer and orographic gravity wave drag, LSWP represents large scale wet processes – namely clouds –, and RDTN captures all radiative tendencies.

In figure 2.21 we show both the mean total tendency field across the 114 archived fields, as well as the mean absolute tendency, which helps differentiate between inactive parameterisations and active parameterisations with means close to zero. We see that the

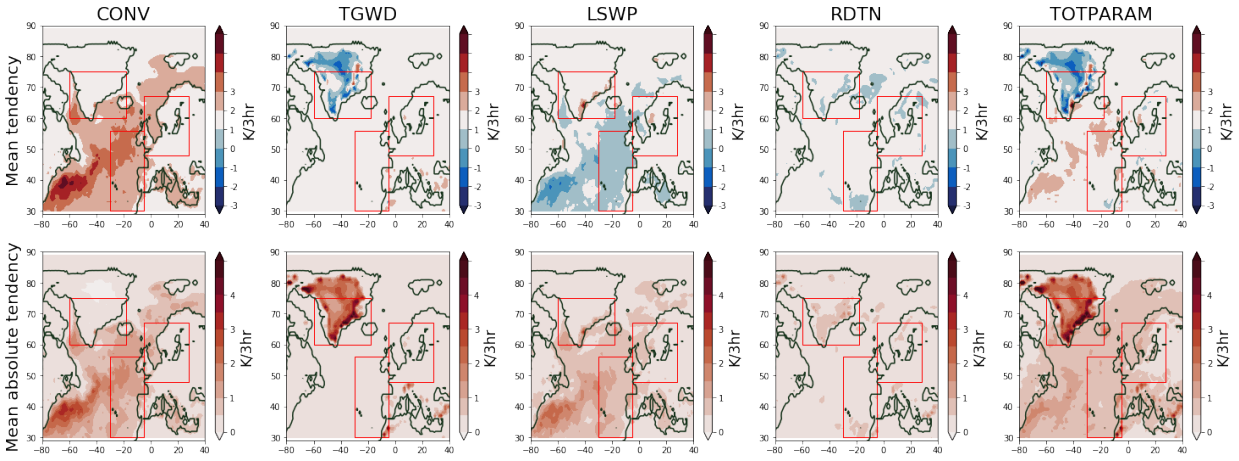


Figure 2.22: As figure 2.21 but showing the tendency and absolute tendency for each component scheme of the IFS model. Boxes mark the approximate locations of the anticylonic highs associated with the three geopotential jet regimes.

Euro-Atlantic sector has quite active parameterised processes, especially over Greenland and the Atlantic storm track. In an absolute amplitude sense, orographic processes clearly are dominant, as can be seen by the clear outline of the Rockies, Andies and the Himalayas, and outside of mountainous regions parameterisations are more active over the ocean than over land. A strong El Niño event occurred during 2009-2010 [Kim et al., 2011] and the corresponding cold-warm-cold tripole can be seen over the equatorial Pacific here. From these fields it is not obvious that the extratropical circulation should be less altered by stochasticity than the tropical circulation, as the tendencies and absolute tendencies are if anything slightly larger in the extratropics.

Figure 2.22 shows the Euro-Atlantic in more detail, now breaking down the total tendency into its constituent contributions. The rough regions associated with each geopotential jet regime's anticyclonic anomaly have been marked for convenient reference.

The impact of RDTN is fairly small, with slight cooling tendencies scattered across the domain. The collective impact of clouds, captured in LSWP and CONV, have large but broadly cancelling impacts, primarily over the ocean regions, and these are the only processes with significant impacts in the South of the domain. TGWD is weak in most of the domain, but very strongly active over Greenland where it introduces a substantial

cooling. Overall, the total parameterised tendency – which is what is perturbed by SPPT – is dominated by the TGWD term with smaller impacts from CONV prevailing over LSWP.

We see that the total tendencies are largest in the west of the domain, with the CONV tendencies warming the gulf stream region and Atlantic storm track, and TGWD tendencies cooling over Greenland. The net tendencies over Western Europe are close to zero, but the TWGD and LWSP schemes produce moderate amplitude parameterised tendencies over the UK, France and Spain.

If these tendencies from YOTC are representative of those in SPHINX then we might expect quite noticeable impacts on the NAO- regime due to modifications to the TGWD scheme, especially as momentum-dissipating turbulent processes should impact the wind variables more strongly than temperature. Enhanced convection in the storm track may also impact the AR regime, perhaps by increasing the injection of eddy momentum into the upstream side of the ridge. However, as emphasised, we can not make any firm conclusions solely on the basis of this proxy data, and we can use it only to inform our analysis of changes in the SPHINX model circulation, which we will consider now.

2.7.4 Mean state changes

A natural starting point for understanding the impact of stochastic perturbations in regimes is to consider changes in the Z500 and U850 fields, and in the jet indices. We have already determined that these different perspectives on the mid-latitude circulation are important for understanding the Euro-Atlantic regime system, and through their combined analysis we can gain a holistic understanding of the circulation changes.

We will also consider the upper-level meridional divergence of eddy-momentum flux, $\partial_y \langle U'_{250} V'_{250} \rangle$, where we have used a 2-7 day bandpass filter to extract the fast eddies from the wind fields, as in [Woollings et al., 2010]. This tells us whether momentum is being injected or removed from a given latitude band, and has important consequences

for jet positioning. It could also inform our understanding of regime persistence changes, tying into the consideration of Shutts type eddy-induced block maintenance explored in chapter 1. The rationale for using 250hPa wind as opposed to the 850hPa wind used to compute jet diagnostics is that most wave forcing is found at the upper levels of the atmosphere. However in the upper levels, the eddy-driven and subtropical jets become entangled (see [Li and Wettstein, 2012]), complicating the identification of trimodality, and so lower level winds are used for the jet indices.

Prior analysis of the SPHINX runs has demonstrated that the inclusion of stochastic perturbations can create quite large adjustments in the model mean state, but reported changes have so far primarily been limited to the tropics. For example [Yang et al., 2019] reports marked increases in the amplitude and low-frequency variability of ENSO – in line with observations- partially through reduction in Pacific SST biases. The low frequency variability and intensity of tropical rainfall is also improved [Watson et al., 2017], and total cloud liquid water in the tropics is increased [Strommen et al., 2019]. However [Davini et al., 2017b] reported no change in the extratropical circulation, and while [Meccia et al., 2020] found a delayed collapse of polar ice in the stochastic runs, no changes in 20th century sea ice coverage are observed.

However, with careful analysis we do indeed find statistically significant changes in the mean state across the Euro-Atlantic domain. Deterministic SPHINX has a small bias in mean 500 hPa geopotential height, featuring too weak a meridional gradient, as shown in figure 2.23. Stochastic perturbations slightly improves this, reducing geopotential heights everywhere but especially at the poles and in a particular region over the Labrador sea and North-Eastern seaboard. This results in a steepening of the meridional gradient on the Southern flank of the anomaly, reflected in an increased mean zonal wind speed there, and correspondingly a weaker gradient with decreased zonal wind speed visible on the Northern flank, as shown in the final column of figure 2.24.

The change in mean wind speed averaged over the southern Greenland area is primarily

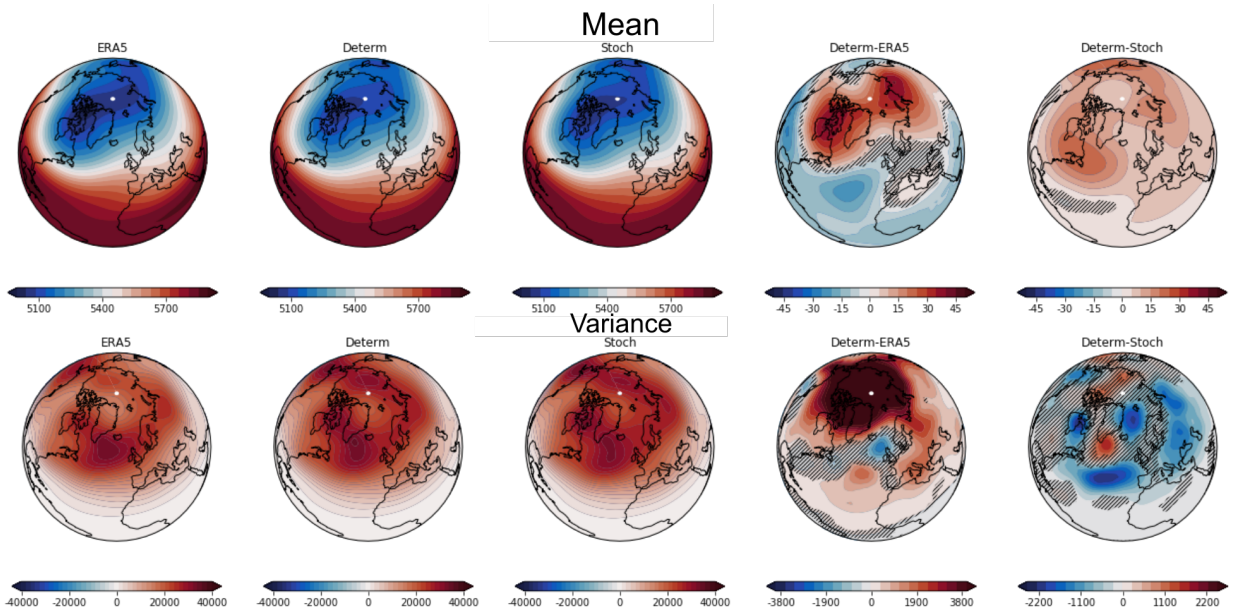


Figure 2.23: The mean and variance of DJF Z500 in ERA5 and in SPHINX, as well as differences between them. Areas where differences are not significant at the 5% level according to a bootstrap estimate are hatched out.

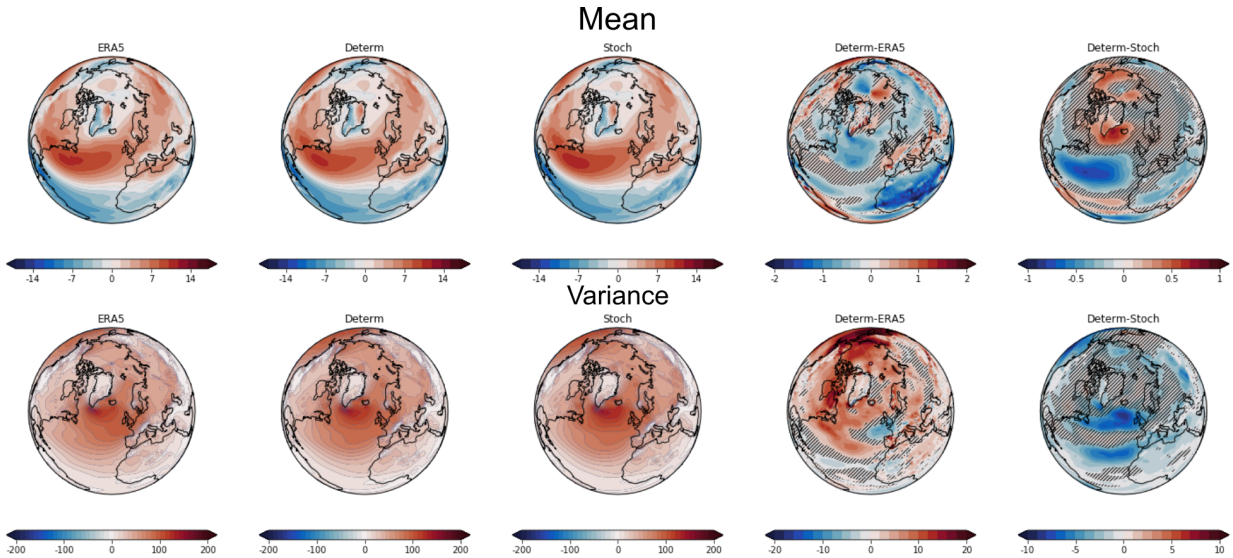


Figure 2.24: As in figure 2.23, but for U850

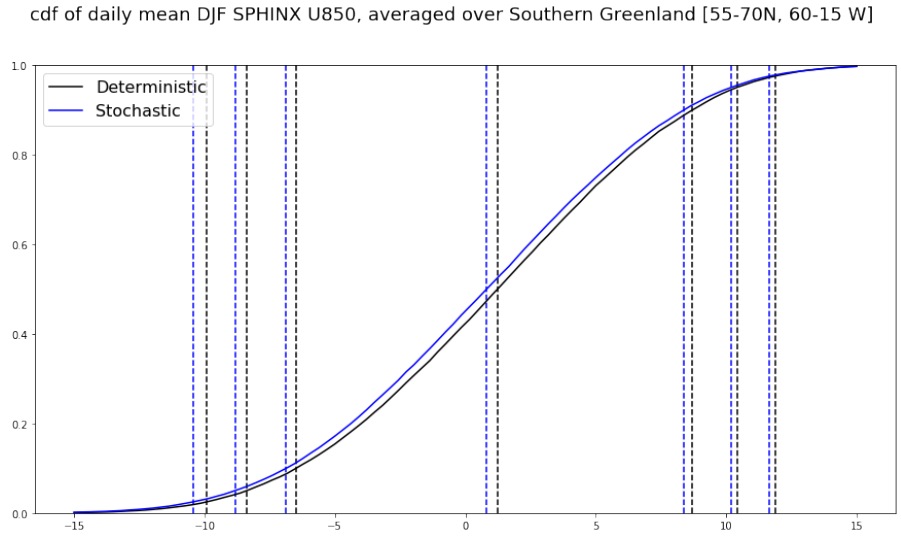


Figure 2.25: The cumulative density function of DJF U850 in the SPHINX runs, averaged with an area-weighted mean over Southern Greenland. Dashed lines mark the 2.5th, 5th, 10th, 50th 90th, 95th, and 97.5th percentiles of each distribution, highlighting the increased skew towards strong Easterlies in the stochastic runs.

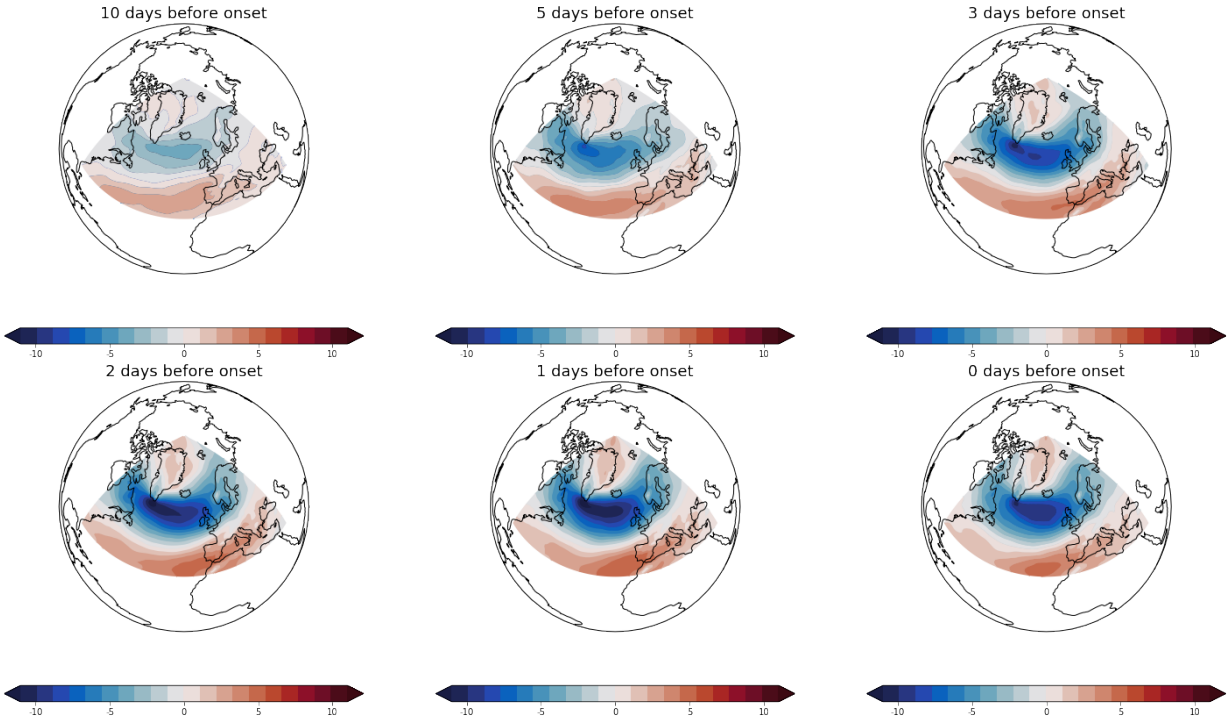


Figure 2.26: Composites of U850 wind anomalies in the days before the onset of an NAO- event, where the neutral state has been included in the regime classification. The Easterly wind anomalies are strongest prior to the onset of the NAO- event, indicating a possible causal link.

due to an increase in the number of strong Easterly days, as seen in the cumulative distribution function presented in figure 2.25, and the region is associated with a localised decrease in Z500 variance (c.f. figure 2.23). While the most likely hypothesis is that increased NAO- occurrence drives the variance reduction, the decreased Westerly winds present a viable mechanism to drive the increased NAO- occurrence. This is supported by figure 2.26 which shows composites of the zonal wind speed anomaly in stochastic SPHINX, at various lag times before the onset of an NAO- regime event. Composites produced using the deterministic ensemble members are qualitatively identical. The negative wind anomaly is larger in the three days before onset than at the beginning of the event, establishing the causal link, and suggesting the low wind speeds support the establishing of the NAO- regime, rather than the winds being a symptom of the increase.

There is a similarity here to the CdV system, except whereas there reductions in zonal wind speed served to persist the regime, here they seem to be responsible for increased occurrence, but do not increase persistence. It is in fact not easy to understand why a quite substantial decrease in NAO- persistence *is* seen, while the number of long lived blocks over Greenland modestly increases (analysis of persistence probabilities at longer than daily lead times suggests the Markovian hypothesis is not to blame). One possibility is that the stochastic model experiences more long-lived blocks and more transient highs over Greenland, such that overall the NAO- persistence decreases. No significant increase in stochastic regime persistence could be seen when restricting to only NAO- events that lasted 5+ days (not shown) however, so the picture is not fully coherent across different metrics. This emphasises the importance of analysing the very complex mid-latitude circulation from many different perspectives in order to gain a more holistic understanding.

Considering the changes in jet dynamics, the stochastic runs can not be said to reduce the bias in either latitude or speed, as shown in figure 2.27, in fact showing a stronger bias towards jet speeds >15 m/s. There is a bias towards southern jet days in both SPHINX configurations which is slightly strengthened in the stochastic runs, in agreement with

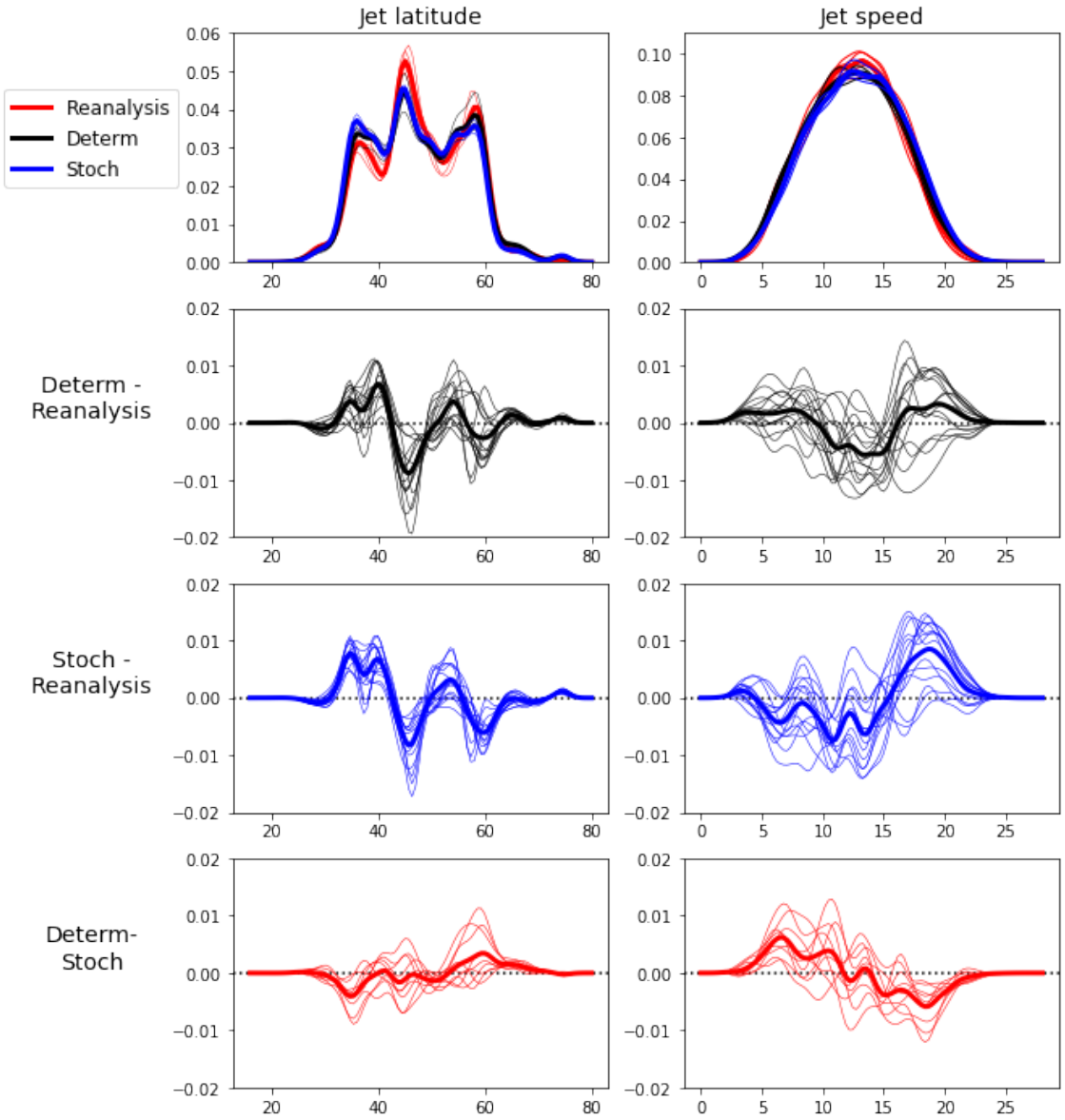


Figure 2.27: Top: The probability distribution of daily jet speed and jet latitude in reanalysis and SPHINX data. Thin lines show individual ensemble member distributions for SPHINX and each reanalysis product, while thick lines show averages. Lower: Anomalies between distributions. Thick lines are the anomaly of the ensemble means, while thin lines show the anomaly between every pair of ensemble members.

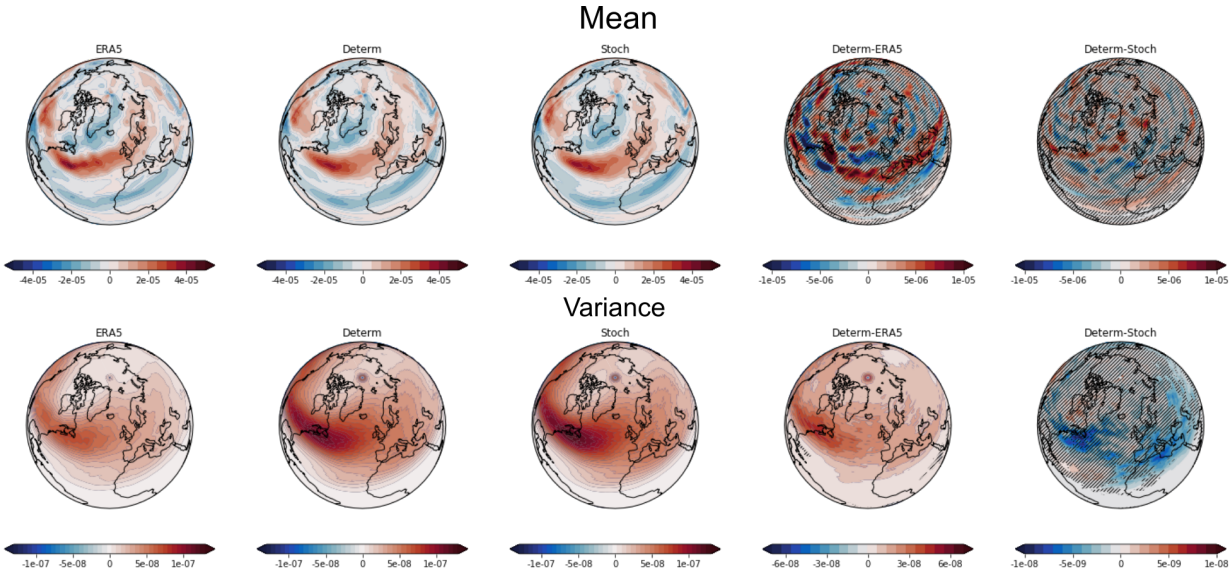


Figure 2.28: As in figure 2.23 but for the 2-7 day bandpassed meridional eddy momentum divergence.

the southward shift of U850 in the west of the domain. A faster jet does not clearly correspond with the decreased occurrence of the Neutral regime although it may help explain the reduced persistence of the neutral regime, as [Woollings et al., 2018] suggested faster jets may be less stable. Further, a southerly shift in jet latitude does not correspond with an increase in Atlantic ridge occurrence, which is linked to northerly jets, however the mapping between blocking regimes and jet regimes is by no means one-to-one and so there is space in our understanding for such conflicting changes to occur.

Despite shifts in the jet and mean U850, the upper level meridional divergent momentum flux is indistinguishable from reanalysis in both deterministic and stochastic runs (see figure 2.28). Both model variants demonstrate a much higher flux variance than seen in reanalysis, almost twice as high at its maximum latitude, around 40N. Stochasticity has an essentially negligible impact on this variance, but does increase it slightly in some locations. This suggests eddy momentum divergence is not playing an important role in controlling stochastically driven regime changes.

2.7.5 Can ENSO improvements explain regime changes

We have found very slight changes in the Euro-Atlantic circulation with stochastic perturbations, but there was no clear way of understanding how these might lead to the regime changes we've seen, with the possible exception of the increased NAO- regime occurrence.

Given that much of the prior documented impact of stochastic perturbations on model climate has been found in the tropics, we might wonder if the improved tropical climate is playing a role in regime changes. This seems particularly plausible for the changes in the Atlantic ridge regime, which is both the most improved in the stochastic model and features a block at the lowest latitude, exposing it most directly to changes in the tropics. There is a range of evidence showing the ability for tropical signals to modulate Euro-Atlantic regimes. For example, [Cassou, 2008] showed that the Madden-Julian oscillation influenced the relative occurrence of the four classical circulation regimes, while [O'Reilly et al., 2019] showed a JJA teleconnection between ENSO and the Euro-Atlantic circulation emerging in recent decades as a likely result of climate change. Other notable studies include [Zhang et al., 2019],[Mezzina et al., 2020] and [King et al., 2018].

While an in-depth investigation of tropical regime drivers is beyond the scope of this thesis, given that the ENSO improvements reported in [Yang et al., 2019] were quite significant, we shall briefly consider whether ENSO improvements can be seen to play a role here.

First we establish that the temporal statistics of geopotential jet regimes are susceptible to ENSO forcing in the real world. Figure 2.29 shows the correlation between the ERA20C DJF mean Niño3.4 index, and DJF mean regime occurrence and persistence over the period 1980-2010. The Niño3.4 index, defined as a 5-month running mean of sea surface temperature averaged over [5N-5S,170W-120W], is the most commonly used metric of ENSO variability. We compare to the regime metrics for all four reanalyses covering that

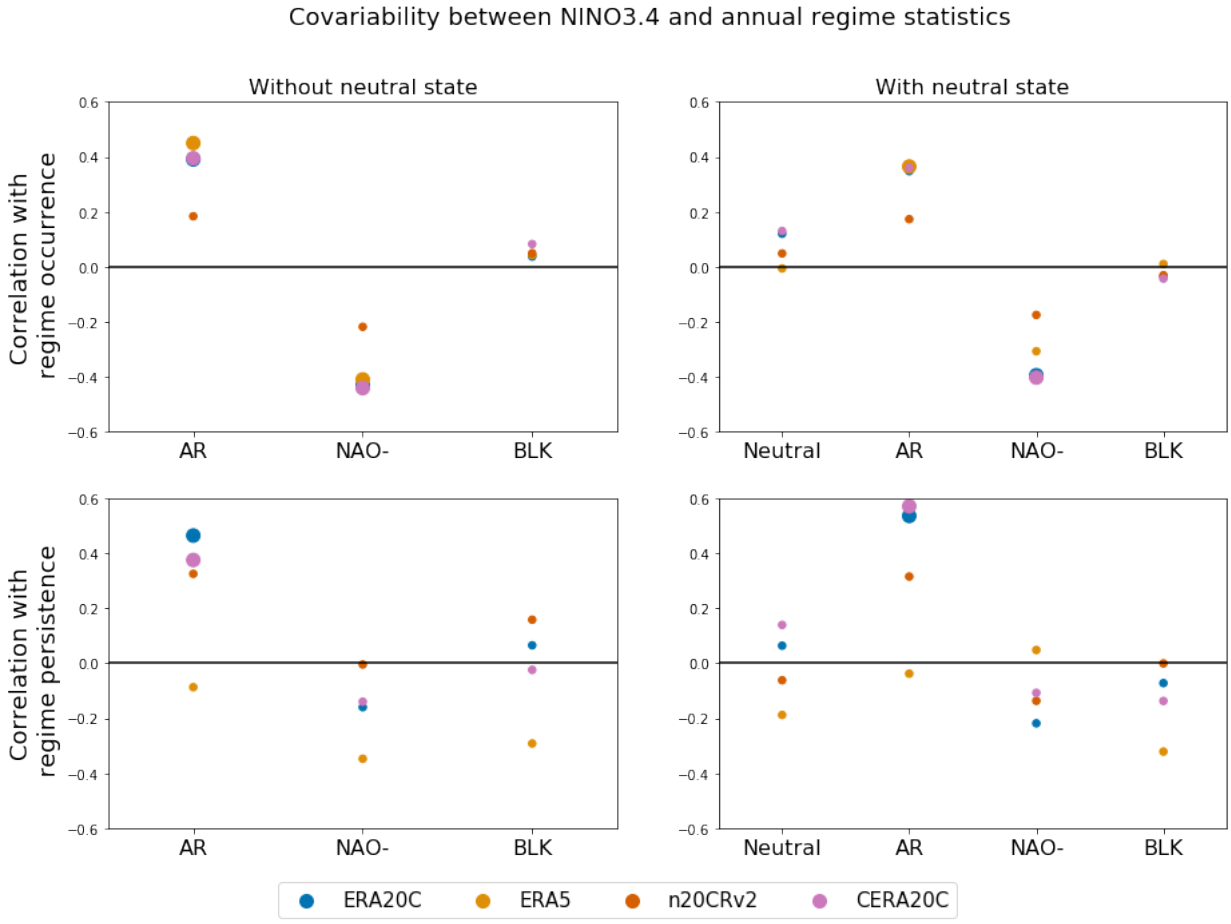


Figure 2.29: The correlation between DJF mean Niño3.4 computed from ERA20C and DJF mean regime metrics computed for several reanalyses, over the time period 1980-2010. Correlations that are significant at the 5% level are shown with large dots, while insignificant correlations are shown with small dots.

period, presenting results both with and without including the neutral state, and find that in both cases at least half the reanalyses show a statistically significant positive correlation between the niño3.4 index and the AR occurrence and persistence, and a significant negative correlation with NAO- occurrence. When using a longer historical period, no statistically significant correlations were found, suggesting a similar non-stationary connection to that found in [O'Reilly et al., 2019].

The asymmetrical impact on the Atlantic ridge is quite suggestive, in that it shows that positive equatorial pacific SST anomalies (El Niño events) are linked to increased Atlantic ridge persistence and occurrence, the same shifts seen in the stochastic SPHINX runs. We might now hypothesise that the improved ENSO representation in the stochastic runs is driving the changes in the Atlantic ridge.

However, when we repeat the analysis using each SPHINX configuration, first computing niño3.4 indices for each ensemble member, we find there are no easy answers available to us, as shown in figure 2.30. The deterministic SPHINX runs partially capture the recent teleconnection signal as long as the neutral state is included, showing AR persistence and occurrence correlated with ENSO. They fail to catch the NAO- correlation despite having three ensemble members and therefore three times as much data as in each reanalysis product. However, the stochastic runs do not show the teleconnection pattern at all, with no significant correlations between any regime metric and the niño3.4 index. Clearly, if the teleconnection is not captured by the model, it can not explain the changes in regimes. Nor is it the case that the stochastic model has an increase in equatorial Pacific SSTs, which would be undetectable from the centred niño index but which would produce the same effect. Instead there are no significant differences in SSTs between the two model configurations.

Taking a more heuristic line, we suggest that even though the stated teleconnection is absent in the stochastic runs, its existence in reanalysis and the deterministic runs does demonstrate a susceptibility of the Atlantic ridge regime's statistics to modification by

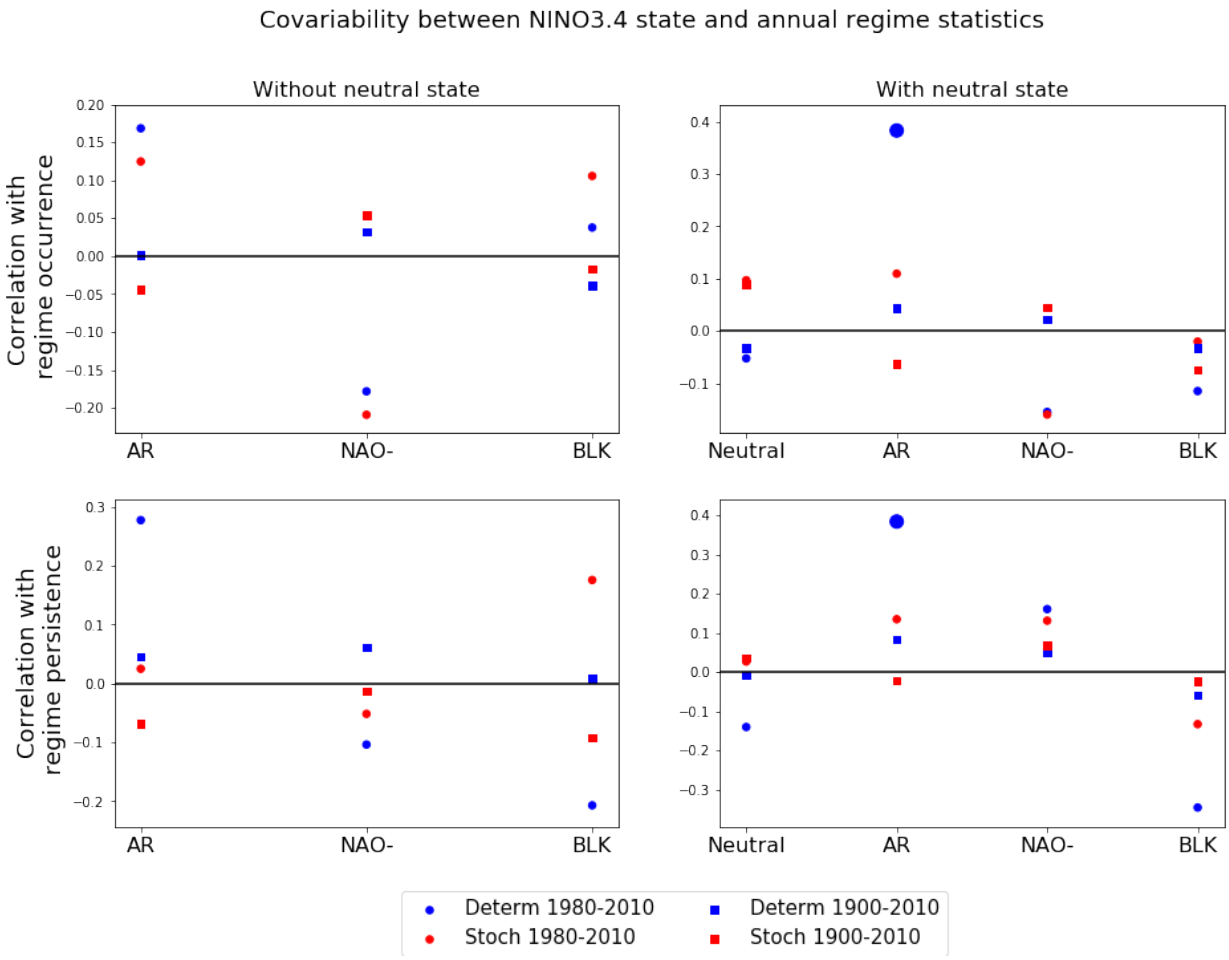


Figure 2.30: Correlations between DJF mean Niño3.4 indices and regime statistics in the SPHINX runs. Significant differences at the 5% level are indicated with enlarged markers.

tropical dynamics, greater than that of other regimes. We therefore conjecture that general improvements in tropical circulation already noted are possibly responsible for the improvement of the AR regime, but we lack a strong body of evidence for this claim. There is clearly much more work to be done on understanding the origin of changes in the Euro-Atlantic circulation.

2.7.6 Summary of SPHINX

In this section we have analysed a very unique set of climate simulations that, for the first time, allow the impacts of stochasticity on the climate system to be cleanly examined in a long-running, fully-coupled ensemble model. Even in this clean experiment, understanding is hard-won, and the overall picture of how regimes change is muddy.

Perhaps the clearest result found is that the stochastic perturbations have produced a significant improvement in the regime fidelity, moving towards the regime structure found in reanalysis. There are also generally improvements in the distribution of blocking events, although not at all locations, contrasted with a lack of improvement in the jet speed or latitude biases found in the deterministic model. This may in part be down to the fact that the deterministic biases are in any case relatively small – the EC-Earth model has a fairly well represented Euro-Atlantic sector. It is not easy to understand why these small changes in regime structure have occurred, but as discussed in [Berner et al., 2017, Sura et al., 2005], and our own section 1.7 stochasticity allows a model to explore broader regions of phase space, and to shift the mean value of nonlinear quantities, potentially drifting closer to quasi-stationary invariants.

By considering a sample of parameterised tendencies produced by operational IFS forecasts, we gained a perspective on the importance of parameterised processes – and thus on the direct impact of stochastic parameterisations – in different parts of the domain. While the stochastic SPHINX runs also include stochastic backscatter, in practice this is tuned away to have almost no impact, and SPPT dominates. We found the most pro-

nounced impact over Greenland, where turbulent processes and orographic gravity wave drag produced substantial cooling tendencies in the T850 field. Other processes had smaller impacts, with a net warming impact of convection in the west Atlantic.

If these sampled tendencies are representative, then they help explain the Z500 anomaly and shift of U850 winds which are strongest in the west of the domain. We might also expect the impact of turbulent processes over Greenland to be even larger on the wind variables than on temperature. Physically we would expect friction to dissipate momentum and so oppose the Westerly flow, which would match with the negative U850 anomalies seen over south Greenland.

We established that Easterly winds in this region were a precursor of the NAO- regime in the SPHINX model, and so explained the increased NAO- occurrence, although we could not explain its reduced persistence. We also developed no clear hypothesis for the reduced persistence of the BLK and Neutral regimes.

For the AR regime, we found stochasticity increased persistence and occurrence, eliminating the deterministic model bias completely. Ultimately we could not explain this in terms of other circulation changes. We reflected on the possibility of the AR regime shifts being driven by tropical teleconnections, on the basis of similar shifts being in reanalysis during El Niño events. Even though the teleconnection was not present in the stochastic model, we conjecture that due to the southerly location of Atlantic ridge events and the large changes in tropical circulation seen in SPHINX, that improvements in tropical dynamics are responsible for the extratropical improvement also.

When comparing to our analysis of the Charney-deVore system in chapter 1, we see many differences, and are reminded of how many aspects of the true climate are not captured by that simple model. There is no representation of tropical forcing for example that might impact regime persistence, nor is there enough spatial resolution to allow for different regime responses in different parts of the domain as seen here. However, there are some similarities. In particular, when we consider the NAO- regime and blocking

over Greenland, we see that stochasticity reduces wind speeds – a process to which the orography of Greenland seems likely to be crucial – supporting the onset of a NAO-event. There is a clear qualitative link here to the reduction of wind speeds allowing orographically-supported blocks to persist, seen in CdV. The key result of chapter 1 however, stochastically increased regime persistence, is only seen for the AR regime, where there are very few similarities to the CdV regime system. Of course in a GCM context there are many more processes than captured in a simple o.d.e. system, and real world regime structure is entangled with thermodynamic and remote forcings. While not commented on here, other aspects of the CdV system such as phase locking of waves during blocking and stochastic drift towards highly novel blocking states, did not seem to have any clear analogue in the SPHINX system, nor did the power spectrum of the Z500 PCs differ between the simulations.

Future work on understanding the impact of stochastic perturbations on regimes may benefit from applying stochastic perturbations regionally, only in the tropics and only in the Euro-Atlantic for example, in order to understand which regime impacts are local, and which are remotely driven. As a parenthetical but practically relevant final point, we have emphasised the large variability in regime statistics across different 30-year periods, which can derive solely from chance. Regime studies working with only short data windows should be very careful about assigning importance to shifts of a few percent in regime persistence or occurrence.

3 | Geopotential jet regimes in CMIP6

Note on authorship: Work in this chapter was done in collaboration with Kristian Strommen and Federico Fabiano. All text is written by the author. A small number of figures generated by collaborators are included with permission, and this is appropriately noted in the corresponding figure captions.

3.1 Introduction

The wintertime atmospheric circulation in the Euro-Atlantic region is highly complex, featuring non-linear Rossby wave breaking [Woollings et al., 2008]; complex orography [White et al., 2019]; variability in the position, speed and angle of the low-level jet; and external forcings from ocean heat fluxes [Delworth and Zeng, 2016, Delworth et al., 2017], Arctic sea ice [Blackport and Screen, 2019], stratospheric signals [Domeisen et al., 2020] and tropical teleconnections [Rodríguez-Fonseca et al., 2016, Jiménez-Esteve and Domeisen, 2018]. In the absence of a strong tropical jet, as in the North Pacific, the importance of eddy variability is amplified [Eichelberger and Hartmann, 2007, Lee and Kim, 2003], and the meridional tilt of the storm track induced by the precise details of the regional stationary wave pattern adds additional complexity [Brayshaw et al., 2009, Gerber and Vallis, 2009]. As discussed in chapter 2, decomposing the mid-latitude flow in terms of its non-linear variability modes – the jet and the preferred circulation regimes – provides a framework with which to understand the flow-dependent properties and model representation of this challenging geographical domain.

Prior analysis of the latest generation of climate models, as featured in the 6th round of the Climate Model Intercomparison Project (CMIP6), has shown evidence for a continued bias towards an overly zonal jet stream [Oudar et al., 2020], although with notable improvements in jet structure and storm track representation relative to earlier models [Harvey et al., 2020]. Issues with resolving sufficiently long-lived blocking events also

remain [Davini and D'Andrea, 2020], as also seen from a regime perspective in [Fabiano et al., 2020], where NAO- events are insufficiently persistent and where a model bias towards too frequent Atlantic ridging is seen. Understanding how the changes in jet and blocking regime structure relate to each other is an important aspect of correctly resolving the North Atlantic circulation, yet little work has been done in a multimodel context to understand how model errors in these different components of the circulation relate to one another.

Another particular issue of importance is trying to understand not just how to capture the present day circulation in models, but how to project the future climate trends over Europe associated with anthropogenic global warming. Europe is already facing considerable socio-economic risks associated with climate change that are predicted to only intensify [Forzieri et al., 2016], and while there is an agreement between models that Europe is likely to become hotter and experience more intense rainfall in the 21st century [Coppola et al., 2021] as a result of thermodynamic changes, the precise regional details are very uncertain, largely due to uncertainties in the dynamical response associated with future circulation changes [Shepherd, 2014].

In CMIP5, wintertime blocking events, associated in turn with extreme coldsnap events [Brunner et al., 2018], were found to become both less frequent and less intense in a warming climate [Masato et al., 2013, Masato et al., 2014]. This is coherent with projected changes in the jet to become increasingly zonal, confined within a narrower range of latitudes and less wavy overall [Barnes and Polvani, 2013, Peings et al., 2017]. While some analyses of blocking in CMIP6 future warming scenarios report an insignificant change in blocking [Bacer et al., 2021], analysis from a regime perspective reports an increased prevalence in cyclonic flow states, with anticyclonic blocked states becoming correspondingly less persistent in both a single- [Rousi et al., 2021] and multi-model context [Fabiano et al., 2021].

In this chapter we make use of the geopotential-jet regime framework developed in chap-

ter 2 to study both the historical and future dynamics of the Euro-Atlantic atmosphere in climate models. Using the multi-model CMIP5 and CMIP6 ensembles, we provide a robust assessment of model bias that accounts for interdecadal variability, and establish links between the representation of jet and regime dynamics in models. We also assess future trends in the dynamics of geopotential jet regimes, showing that changes in the spatial structure of regimes are small, but that the persistence of regimes is likely to decrease in the future. Finally, we return to the forced Lorenz 63 system and show that the results of [Palmer, 1999] can be extended to more closely capture the type of forced regime response seen in CMIP6 models.

Section 3.2 summarises the CMIP5 and CMIP6 models and experiments analysed. In Section 3.3 we show that the spatial structure of regimes and the eddy-driven jet is well captured by CMIP6, and to a lesser extent CMIP5 as well. In Section 3.4 we analyse the biases in regime persistence and occurrence of CMIP models and then proceed to considering the forced response under the SSP5-8.5 climate change scenario in Section 3.5. Section 3.6 presents the simple model results, and we then summarise our key results in Section 3.7.

3.2 Data

We analyse model data from the 5th (CMIP5) and 6th (CMIP6) phases of the coupled model inter-comparison project: multi-centre ensembles of Earth system models representing the state-of-the-art in global climate modelling in 2011 and 2020 respectively. We analyse the historical experiments for 31 single-member CMIP6 models [Eyring et al., 2016] detailed in table 3.1, and a total of 71 ensemble members from 28 distinct CMIP5 models [Taylor et al., 2012] detailed in table 3.2. These historical experiments consist of coupled uninitialised climate runs forced with the correct historical greenhouse gas and aerosol forcings over the 20th century, after a spin-up from a free-running pre-industrial control run. We also analyse model output from the SSP5-8.5 climate change scenario

for 20 of the CMIP6 models, covering the years 2010-2100 and providing climate projections associated with a supposed worst-case, 'business-as-usual' economic trajectory.

Regimes were identified in the four leading principal components of Z500 anomalies, computed for each model. For the SSP5-8.5 experiments a cubic fit to the area-averaged daily Z500 field over [80W-40E,30N-90N] was removed to detrend each model prior to taking principal components, as principal component analysis can only be applied to stationary data. Removing a spatially constant Z500 field is the most principled approach to detrending from a dynamical perspective, as it does not alter the implied geostrophic wind anomaly associated with the circulation pattern. Jet speed and jet latitude were computed for each model from U850 without detrending as their absolute values have dynamical relevance, although both series were seen to be very close to stationary for all models.

3.3 Climatological representation of regimes

3.3.1 CMIP6 geopotential jet regimes

The spatial representation of regime patterns in the CMIP5 and CMIP6 models is assessed in terms of the regime stability and regime fidelity metrics, as they were introduced in Section 2.7 for analysis of the SPHINX models. Again, for all data sets, subsamples were constructed by taking 30-year windows of DJF data at 10-year intervals, such that 4 subsamples would be computed for a 60-year dataset, and 9 subsamples for a 110-year dataset. This sampling set-up was chosen in order to provide a good number of subsamples for each model while also maintaining a reasonable degree of independence between the data in each sample. Contiguous windows were used rather than bootstraps in order to correctly sample any interdecadal variability in regime patterns.

The stability and fidelity of the CMIP5 and CMIP6 model ensembles, with reanalysis

Model name	Ensemble member	Historical dates	SSP5-8.5?
ACCESS-CM2	r1i1p1f1	1950-2010	Yes
AWI-ESM-1-1-LR	r1i1p1f1	1900-2010	No
BCC-CSM2	r1i1p1f1	1950-2010	Yes
BCC-ESM1	r1i1p1f1	1950-2010	No
CanESM5	r1i1p1f1	1900-2010	Yes
CESM2-FV2	r1i1p1f1	1900-2010	No
CESM2	r1i1p1f1	1900-2010	No
CESM2-WACCM-FV2	r1i1p1f1	1900-2010	No
CESM2-WACCM	r1i1p1f1	1900-2010	No
CNRM-CM6-1	r1i1p1f2	1950-2010	Yes
CNRM-CM6-1-HR	r1i1p1f2	1900-2010	Yes
CNRM-ESM2	r1i1p1f2	1950-2010	Yes
EC-Earth3	r1i1p1f1	1900-2010	Yes
FGOALS-f3	r1i1p1f1	1950-2010	No
FGOALS-g3	r1i1p1f1	1900-2010	Yes
GFDL-CM4	r1i1p1f1	1910-2010	Yes
GISS-E2-1-G	r1i1p1f1	1910-2010	No
HadGEM3-GC31-LL	r1i1p1f3	1900-2010	Yes
HadGEM3-GC31-MM	r1i1p1f3	1900-2010	Yes
INM-CM4-8	r1i1p1f1	1900-2010	Yes
INM-CM5-0	r1i1p1f1	1900-2010	Yes
IPSL-CM6A-LR	r1i1p1f1	1950-2010	Yes
MIROC6	r1i1p1f1	1900-2010	Yes
MPI-ESM1-2-HAM	r1i1p1f1	1900-2010	No
MPI-ESM1-2-HR	r1i1p1f1	1900-2010	Yes
MPI-ESM1-2-LR	r1i1p1f1	1910-2010	Yes
MRI-ESM2-0	r1i1p1f1	1950-2010	Yes
NorESM2-LM	r1i1p1f1	1950-2010	Yes
NorESM2-MM	r1i1p1f1	1950-2010	Yes
TaiESM1	r1i1p1f1	1900-2010	No
UKESM1-0-LL	r1i1p1f2	1900-2010	Yes

Table 3.1: CMIP6 models used in this chapter, with the range of historical data available, and whether their SSP5-8.5 future scenario experiments were available.

Model name	Ensemble members	Historical dates
ACCESS1-0	r1i1p1	1950-2005
ACCESS1-3	r1i1p1	1950-2005
BCC-CSM1-1	r1i1p1	1950-1998
BCC-CSM1-1-m	r1i1p1	1950-2005
BNU-ESM	r1i1p1	1950-2005
CanESM2	r[1,2,3,4,5]i1p1	1950-2005
CCSM4	r6i1p1	1950-2005
CMCC-CESM	r1i1p1	1950-2005
CMCC-CM	r1i1p1	1950-2005
CMCC-CMS	r1i1p1	1950-2005
CMCC-CM5	r1i1p1	1950-2005
EC-EARTH	r[1,2,7,9,12]i1p1	1950-2005
FGOALS-g2	r[1,3]i1p1	1950-2005
GFDL-CM3	r[1,2,3]i1p1	1950-2005
HadCM3	r[1,2,3,4,5,6,7,8,9,10]i1p1	1960-2005
HadGEM2	r[1,2,3]i1p1	1960-2005
IPSL-CM5A-LR	r[1,2,3,4,5,6]i1p1	1950-2005
IPSL-CM5A-MR	r[1,2,3]i1p1	1950-2005
IPSL-CM5B-LR	r1i1p1	1950-2005
MIROC5	r[1,2,3,4,5]i1p1	1950-2005
MIROC-ESM-CHEM	r1i1p1	1950-2005
MIROC-ESM	r[1,2,3]i1p1	1950-2005
MPI-ESM-LR	r[1,2,3]i1p1	1950-2005
MPI-ESM-MR	r[1,2,3]i1p1	1950-2005
MPI-ESM-P	r[1,2]i1p1	1950-2005
MRI-CGCM3	r1i1p1	1950-2005
MRI-ESM1	r1i1p1	1950-2005
NorESM1-M	r[1,2,3]i1p1	1950-2005

Table 3.2: CMIP5 models used in this chapter, and the corresponding ensemble members and historical date range available.

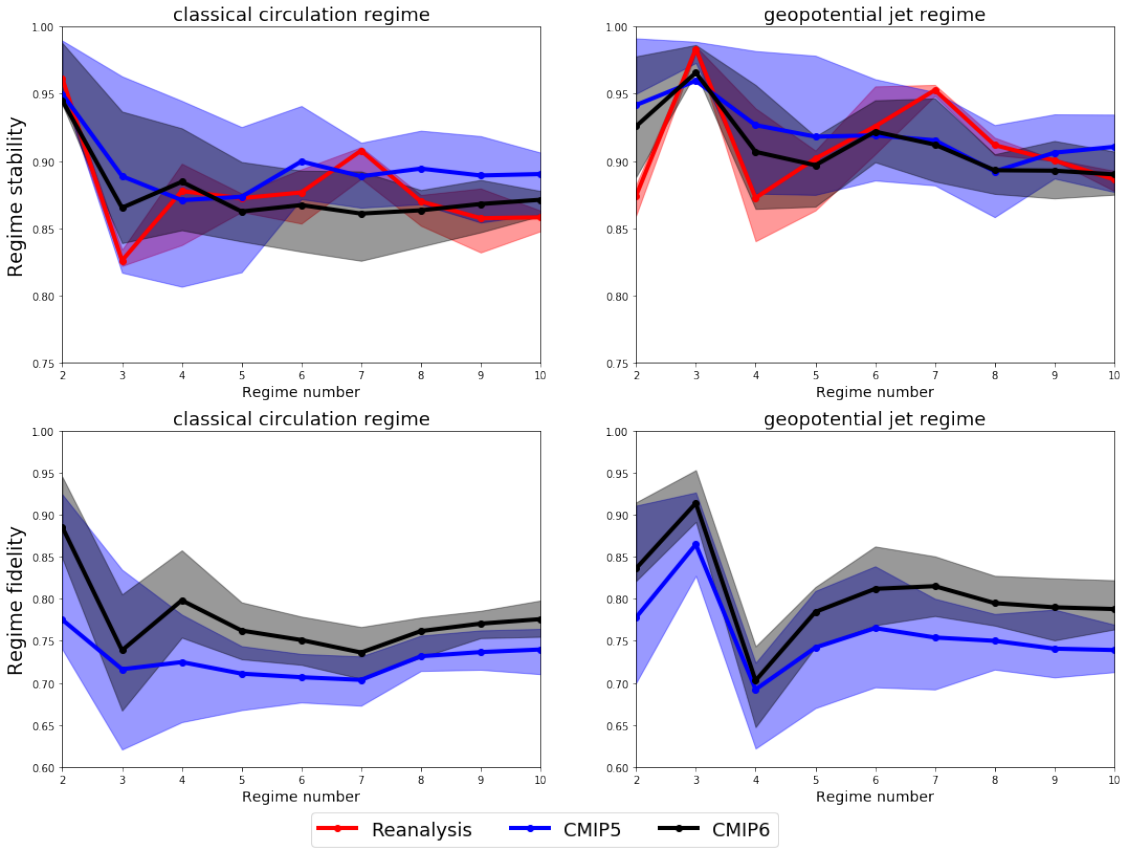


Figure 3.1: Regime stability and fidelity, calculated with 30-year subsamples, as a function of regime number, for both the classical circulation and geopotential jet regime frameworks. Lines show the ensemble mean value for each data set, while shading shows the ensemble interquartile range.

results reproduced as a visual reference, is shown in figure 3.1. Results are shown for both the classical circulation regime and geopotential jet regime frameworks, and for cluster numbers varying from 2 to 10.

Just as was found in reanalysis in Section 2.5 and in [Dorrington and Strommen, 2020], the regime stability in the CMIP5 and CMIP6 ensembles is highest for 3 geopotential-jet regimes, and in most cases the geopotential jet regimes are more stable than their classical circulation counterparts, with the only exception being the two-regime case, where the classical circulation regimes capture the NAO dipole very robustly.

That the high stability for three geopotential-jet regimes is seen not only in reanalyses but in a large number of uninitialised climate models is a strong indicator that these regimes are a property inherent to the structure of the Euro-Atlantic circulation, rather than a chance feature of 20th century. The alternative, that these stable regimes are indeed a chance feature but appear in models due to an overfitting to historical variability, seems far less probable. We note that the high reanalysis stability seen for 7 geopotential-jet regimes is not captured by the models.

Between the CMIP5 and CMIP6 ensembles there are no statistically significant differences in the regime stability, but CMIP6 models have robustly higher regime fidelity than CMIP5, showing a clear progression towards a more accurately resolved Euro-Atlantic circulation. There is also a convergence of model behaviour, as the spread in CMIP6 fidelity and stability is lower than in CMIP5.

The increase in regime fidelity between CMIP5 and CMIP6 is largest for those cluster numbers previously suggested to represent meaningful regimes: 2 and 4 clusters for classical circulation regimes, and 3 and 7 clusters for geopotential jet regimes. This is what we would expect if these regimes really do have a physical basis – clusters found for a non-physical cluster number will be to some extent arbitrary, and so we would not expect a model to reproduce the reanalysis clusters particularly well. In contrast, for physically motivated regimes, a model with improved physics and resolution should be expected to

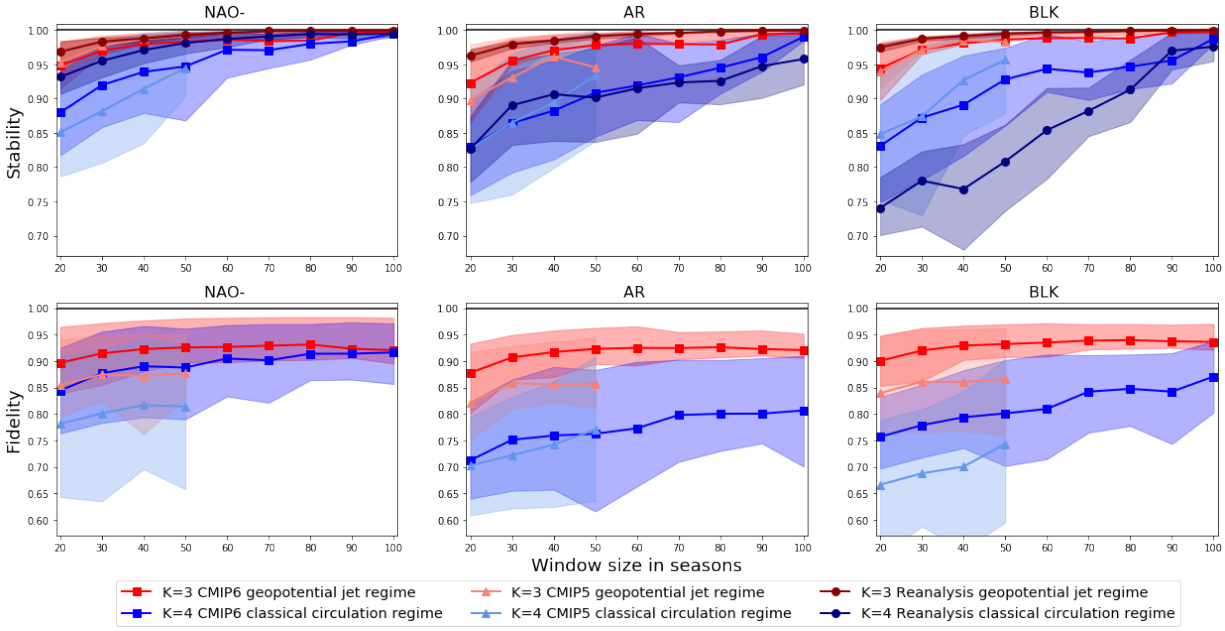


Figure 3.2: Regime stability and fidelity as a function of subsample window size. Lines show ensemble mean values for each atmospheric regime pattern as identified using both 4 classical circulation regimes and 3 geopotential jet regimes. Shading indicates the 1st standard deviation of the ensemble.

more accurately reproduce the real-world regime structure. This therefore gives us further reason to consider these particular regime frameworks as meaningful partitions of the phase space.

The fidelity and stability of the three geopotential jet regimes (AR, NAO- and BLK) taken individually are contrasted with their K=4 classical circulation regime equivalents in figure 3.2, where we also assess the sensitivity of our results to the subsample window size used, which is varied from 20 to 100 years while maintaining the same 10-year sampling interval. The geopotential jet regimes are robustly more stable and have higher fidelity in all datasets for all sampling windows. This is particularly pronounced in regime fidelity for the AR regime, and in both metrics for the BLK regime, while the classical circulation regime framework already captures the NAO- regime relatively well. For the geopotential jet regimes, both CMIP ensembles have slightly lower stability than seen in reanalysis, most notable in the 20-year subsamples, and the fidelity of CMIP6 is uniformly $\sim 5\%$ higher than in CMIP5, regardless of regime or subsample window size

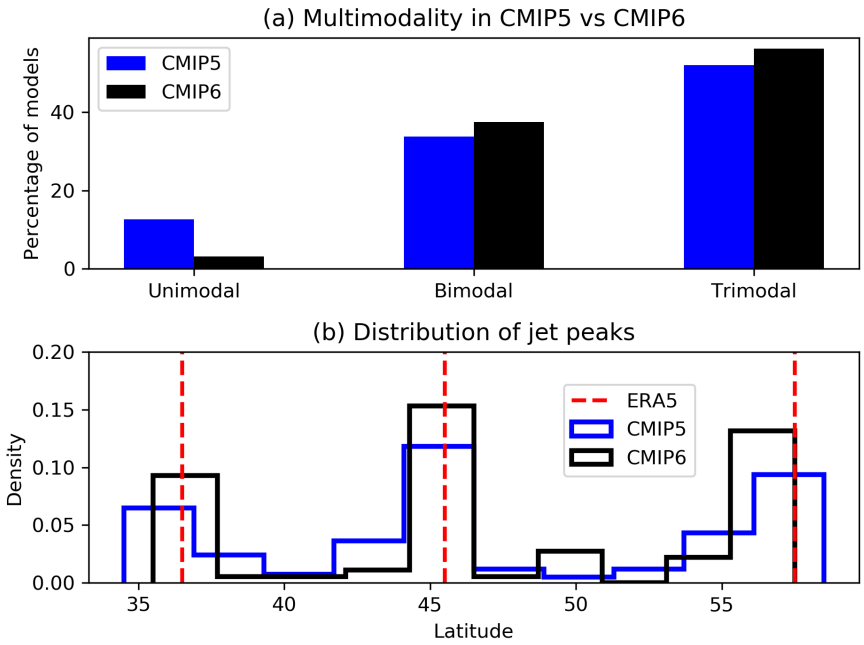


Figure 3.3: a) The percentage of climate models identified as having one, two, or three peaks in their DJF jet latitude distributions. b) Histograms of the latitude at which climate models produce jet latitude peaks, with vertical lines showing the true location of peaks found in ERA5. Created by Kristian Strommen and reproduced from [Dorrington et al., 2021] with permission from the authors.

considered.

There are also substantial improvements in the representation of the jet latitude regimes in CMIP6. Figure 3.3 shows the fraction of models in both CMIP ensembles with one, two, or three jet latitude peaks, as well as the average central latitude at which those peaks occur. While a larger fraction of CMIP6 models properly resolve the trimodal jet latitude distribution than in CMIP5, the main improvement comes from a large reduction in the number of models with unimodal jet latitude distributions. The positioning of the Southern and Central peaks is improved, and the number of peaks occurring at latitudes avoided in reanalysis is reduced. There are still issues in correctly resolving the Northern jet latitude peak however (associated with the AR regime), and many of the models with bimodal jet latitude distributions are missing this peak.

In order to holistically quantify the error in the jet latitude and jet speed distributions in models, we make use of the Wasserstein distance. Also known as the "Earth-mover

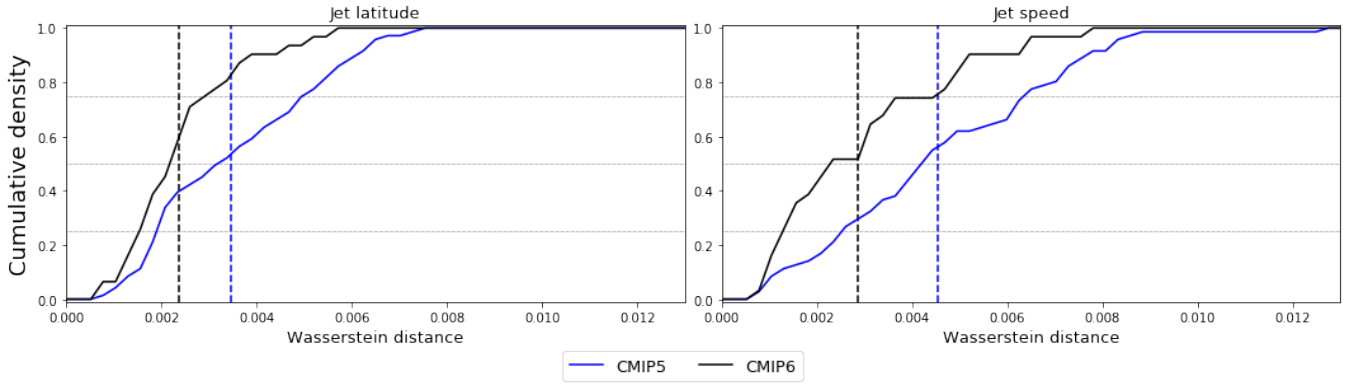


Figure 3.4: Cumulative density functions over model ensembles of the Wasserstein distance between model jet latitude and jet speed distributions and the multi-reanalysis mean distributions. Dashed vertical lines indicate the ensemble mean distances.

distance", this metric provides a natural way to compare two probability distributions, and can be understood informally as a measure of how much of the probability density of one distribution must be shifted to transport it into another. Formally, the Wasserstein distance (as used here) is defined as the integrated difference between two distributions cumulative density functions. This distance was introduced to the atmospheric literature in [Ghil, 2015] and has since been applied to the analysis of both simple climate attractors [Robin et al., 2017] and fully coupled climate models [Vissio et al., 2020], the latter of which uses it to evaluate model error, just as we do here. Like any distance measure, values close to zero indicate smaller errors. The cumulative density function of the jet latitude and jet speed Wasserstein distances between each model and the multi-reanalysis mean jet distributions are shown in figure 3.4. The average Wasserstein distance is 33% and 38% lower in CMIP6 than in CMIP5 for the jet latitude and jet speed distributions respectively, and for jet latitude this is particularly due to a reduction in the number of CMIP6 models with large errors as visible in the diverging of the c.d.fs for distances exceeding 0.002. To reiterate, this decreased distance can be directly interpreted in a holistic decrease in the error of the models' probability distributions.

The quality of these different aspects of the model regime climatology are not independent of each-other, and models that do better at representing one aspect of the Euro-

Atlantic circulation generally do better at representing other aspects. Figure 3.5 a) shows a positive correlation between the jet speed and jet latitude Wasserstein distances, significant at the 1% level, a relationship that is seemingly stronger in the CMIP6 ensemble. Panel b) shows that a smaller jet latitude Wasserstein distance is in turn predictive of a model having high fidelity in the 3 geopotential jet regime patterns with strong negative correlations between the two measures again significant at the 1% level. Jet speed errors are weakly negatively correlated to regime fidelity (correlations between $=-0.25$and -0.4), but this is entirely due to the covariability with jet latitude errors; when we regress out the jet latitude distance from the jet speed distance and correlate the residual with regime fidelity as in panel c), we see that there is no correlation whatsoever. The reverse approach – computing the residual of jet latitude distance after regressing out jet speed – produces a residual which is still strongly correlated with regime fidelity (not shown). This strongly suggests that the jet speed really is orthogonal to the regime dynamics we are interested in and that biases in the model jet speed have no relation to the quality of the anticyclonic regime structure, although of course we can not rigorously assert causal relationships from a post-hoc analysis. Results are very similar when correlating jet errors with the fidelity of the 4 classical circulation regimes (not shown), except the negative correlation between jet latitude errors and fidelity is weaker.

3.4 Historical regime variability

Having established the adequate representation of geopotential-jet regime patterns in CMIP models in the previous section, we can now turn to the representation of the temporal characteristics of those regimes – their persistence and occurrence probabilities – and their historical variability.

Regime occurrence and persistence have been computed in 30-year rolling windows for all reanalyses and all CMIP models, in order to highlight the interdecadal climatic variations in the regimes. The ensemble-mean 30-year regime occurrence for each dataset is

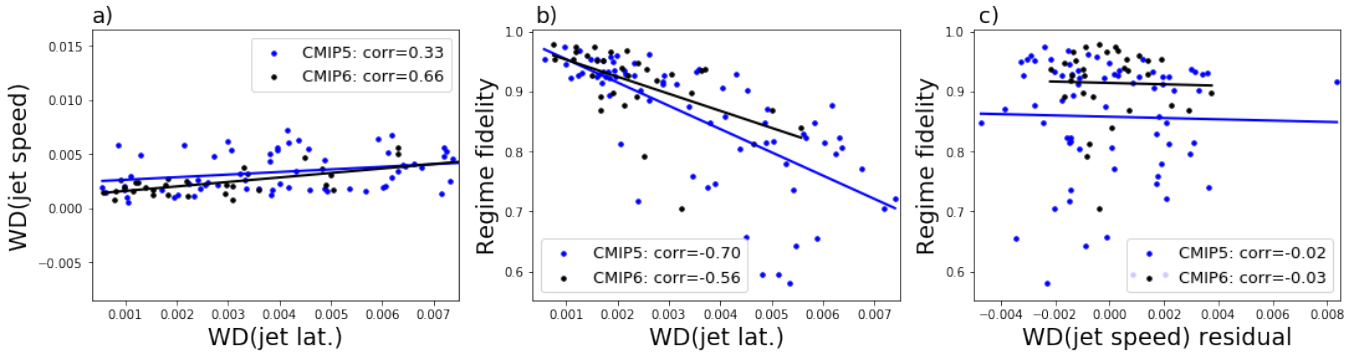


Figure 3.5: a) Correlations between the Wasserstein distances in jet latitude and jet speed in each climate model simulation. b) The same but showing correlation between Wasserstein distance in jet latitude and the fidelity of the three geopotential jet regimes. c) the correlation between regime fidelity and the residual of the jet speed distance after covariability with the jet latitude distance has been regressed out. Correlations in panels a) and b) are significant at the 1% level, while no significant correlations are present in panel c).

shown in figure 3.6, where all models/reanalyses that have data for each full 30-year period are included in the average. The full ensemble spread is shaded, and for the CMIP6 and reanalysis datasets, long-term mean values are indicated.

Observational variability in regime occurrence is quite substantial, and is markedly higher for the neutral and NAO- regime than for the AR and BLK regimes, which show less pronounced variation. The previously documented mid-century peak in NAO- occurrence that has been related to reductions in forecast skill in [Weisheimer et al., 2017] is visible here, and we now also document a clear decrease in occurrence of the Neutral regime from 1970-1990. Between the reanalyses there is a fair amount of spread in the exact occurrence statistics, which is notable given the almost perfectly equivalent regime patterns seen in the reanalyses. This is most pronounced in the early century Neutral and AR regimes where reanalyses disagree by as much as 3-4%, but can also be seen right into the modern era, such as in the case of the BLK regime.

The biases in CMIP6 are similar to those in CMIP5: Neutral flow states are too common and the BLK regime occurs too infrequently, both by $\sim$2%. There is no noticeable bias in the NAO- or AR occurrence, which in the NAO- case represents a slight improvement over CMIP5. The coupled CMIP models do not reproduce the actual historical variability

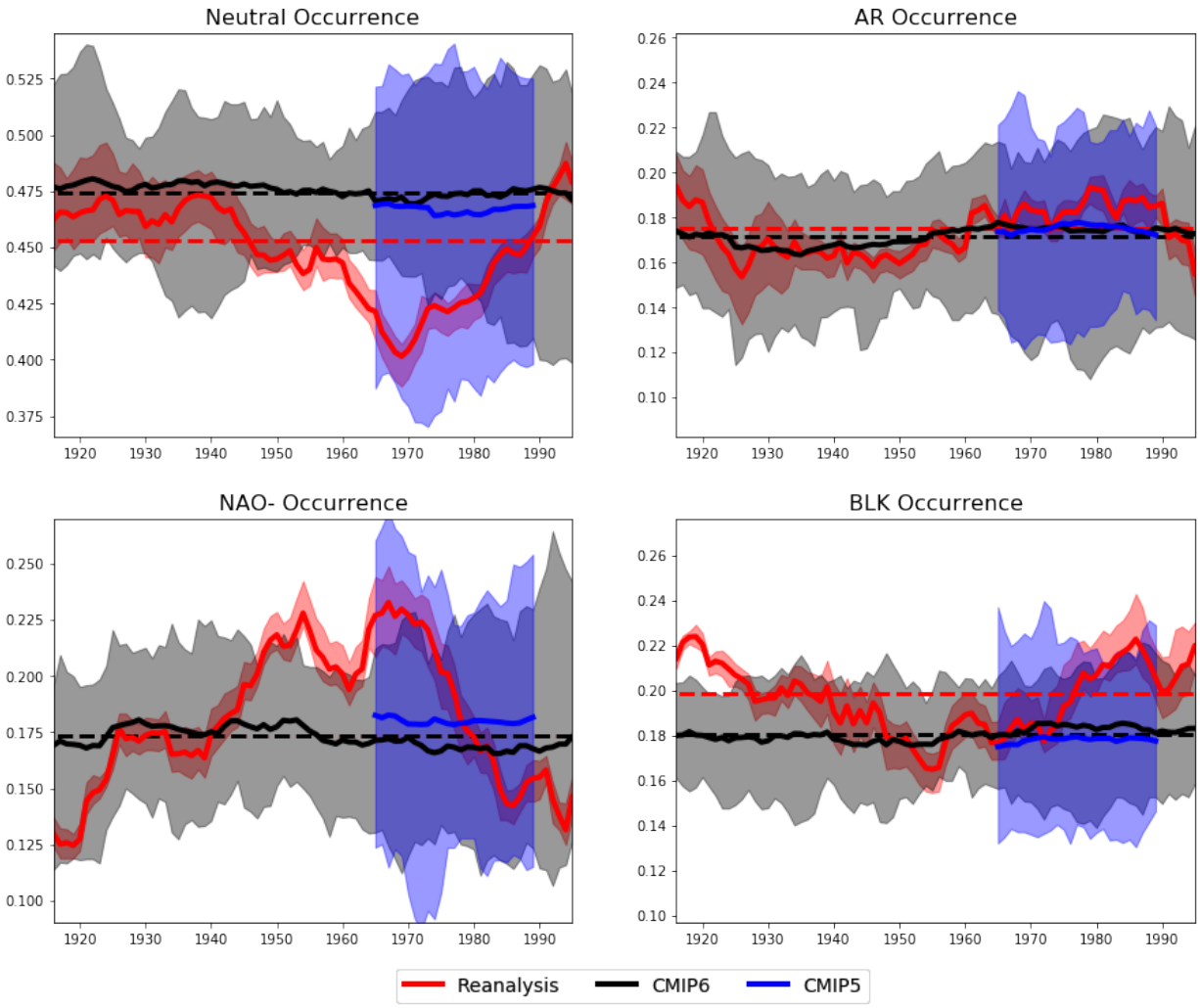


Figure 3.6: 30-year rolling windows of regime occurrence for the multireanalysis, CMIP5 and CMIP6 ensembles, with the x-axis marking the central year of the window. Thick lines show the ensemble mean, while shading indicates the full ensemble spread. Dashed red and black lines show the 1900-2010 average occurrence for reanalysis and CMIP6 respectively.

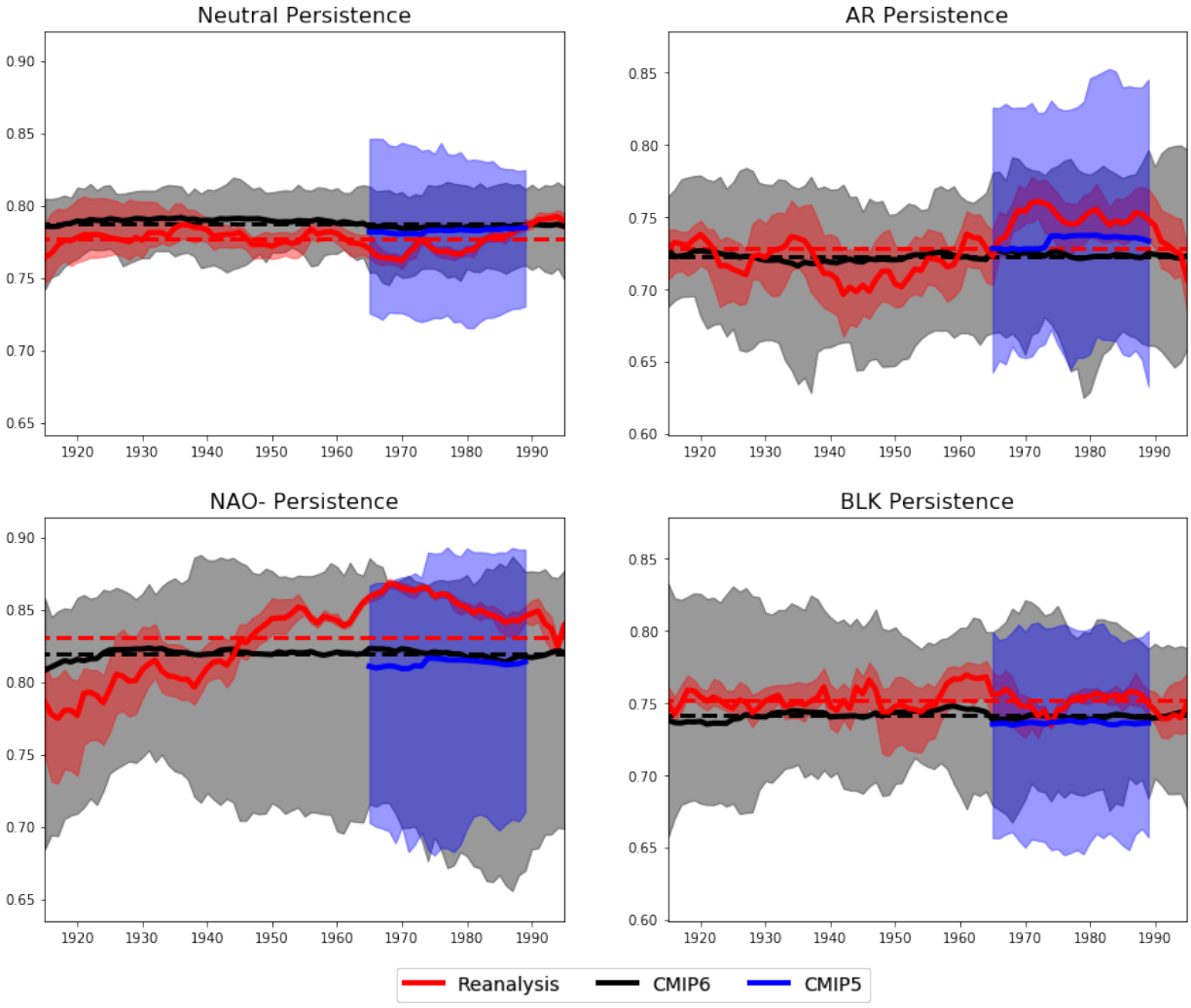


Figure 3.7: As in figure 3.6 but showing regime persistence.

seen in reanalysis, and the ensemble mean occurrence is almost flat across the century, consistent with the perspective that slow SST variability – which will not be coherent across an average of coupled models – is a major driver of regime variability. Analysing AMIP regime statistics could validate this claim, and is left for future work. This emphasises the importance of evaluating model regimes against an appropriately long historical benchmark. For example, looking only at the years 1980-2010, CMIP6 would seem to have a positive NAO- occurrence 'bias', while between 1955 and 1985 it would seem to have a large negative bias of 5%. Comparing centennial averages however, there is no NAO- occurrence bias at all.

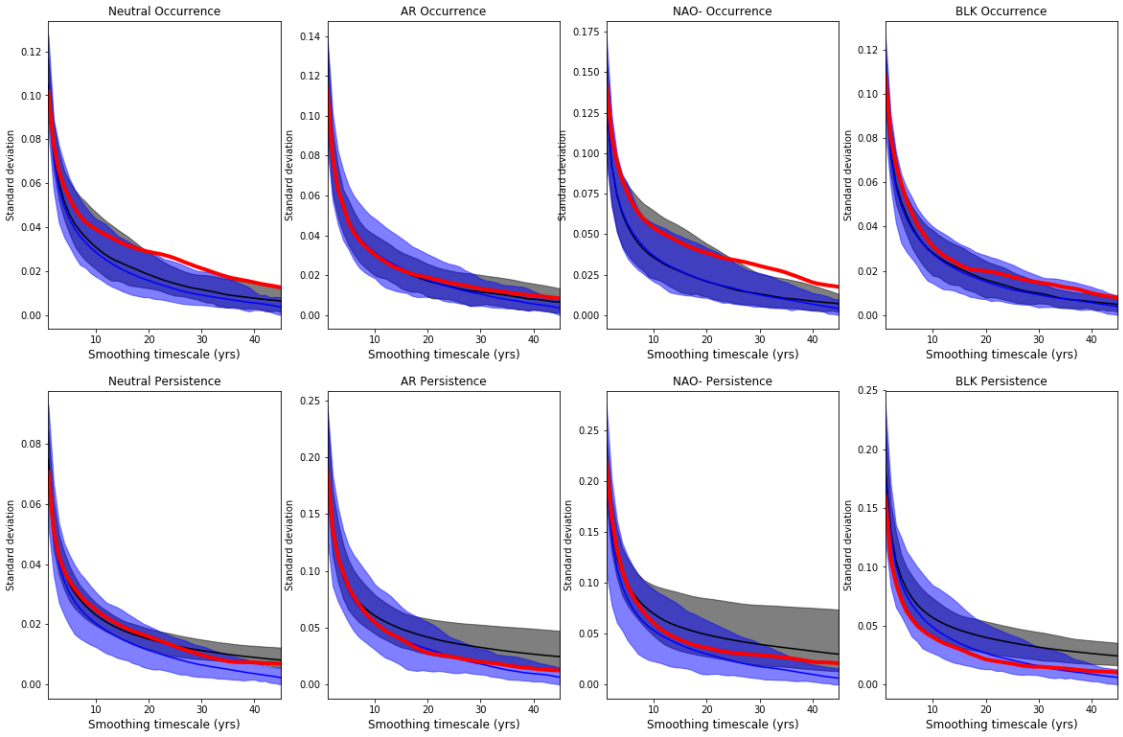


Figure 3.8: Standard deviation of seasonal regime occurrence and persistence as a function of smoothing timescale, showing the variability of the regime statistics on a range of timescales, with CMIP5 shown in blue, CMIP6 in black, and reanalysis in red. Thick lines indicate ensemble means of each models' standard deviation, while shading shows the model spread.

Regime persistence and its variability is shown in figure 3.7. Here the disagreements between reanalysis estimates of historical regime persistence are even larger than for occurrence, and worryingly divergent spread in NAO- persistence is seen between reanalyses in the first half of the 20th century. It is quite possible that this spread is spurious due to poor historical data constraints, especially as low-persistence is consistent with the CMIP6 model bias, which shows BLK and NAO- persistence probabilities 2% lower than in the reanalysis ensemble. The AR regime shows no persistence bias, while there is a tendency for CMIP6 to produce too persistent neutral events, in line with the known biases towards overly zonal flows. The underpersistence of the NAO- and BLK regimes is not improved in CMIP6 over CMIP5, again in line with the comparison in blocking lifetimes between CMIP generations in [Davini and D'Andrea, 2016].

The variability of the CMIP ensemble means is of course far smaller than of the reanal-

ysis mean, and it is also valuable to consider whether the individual models are able to represent the correct variability structure of occurrence and persistence, as well as their mean values. Figure 3.8 shows the standard deviation of the occurrence and persistence metrics for a range of smoothing timescales. e.g. the 1-year smoothing timescale captures the interannual variability while the 30-year smoothing timescale captures multidecadal variability. In the presence of non-stationarity in some reanalyses metrics and model time series of different lengths, such an approach was found more robust than a more-conventional inspection of the Fourier power spectra.

The interannual variability of occurrence and persistence is well captured in models for all regimes, but biases appear on longer timescales of variability. In particular, models underestimate the inter- and multi-decadal variability of the Neutral and NAO- regime occurrence, and to a lesser extent for BLK occurrence as well. Low-frequency variability in regime persistence however is too high in CMIP6, representing a notable departure from CMIP5, which has little bias in persistence variability.

3.5 Impact of climate change on regimes

In the previous section we established that the three geopotential jet regimes introduced in chapter 2 are well reproduced in CMIP5 and CMIP6 models, and developed a detailed understanding of their historical variability and the biases in the occurrence and persistence of those regimes. We found variability to be mostly in line with reanalysis especially on shorter timescales, and relatively small biases. In short, although there is progress still to be made until Euro-Atlantic regimes are perfectly captured by climate models, the geopotential jet regime structure in CMIP6 is close enough to reality that there is clear value in considering its forced response under an anthropogenic climate change scenario.

In both reanalyses and CMIP simulations, the historical regime pattern variability has been shown to be small, as captured by values of the regime stability very close to 1.

However, the substantial forcing of the climate system due to anthropogenic warming may induce circulation changes that dominate over internal variability in the 21st century, and could therefore produce qualitatively different regime dynamics. Our motivation for considering how the forced climate response will impact regimes is twofold. Firstly, we wish to test the hypothesis of [Palmer, 1999] (P99) which predicts forcing should manifest primarily in a shift in regime occurrence, rather than changes in the regime patterns themselves. This hypothesis has never been explicitly tested in complex Earth system models. Secondly, the details of any forced regime response is inherently interesting due to the link between persistent regime events and extreme weather, and the general uncertainties in future European climate.

To this end we consider the representation of geopotential-jet regimes in the CMIP6 SSP5-8.5 scenario. This scenario has been characterised as relatively unlikely and represents an extreme future rather than a baseline 'best guess' emissions scenario [Burgess et al., 2020]. However [Fabiano et al., 2021] showed that for 4 classical circulation regimes, the magnitude of the regime response was approximately linear with increasing warming, and so we consider only this most extreme scenario here in order to obtain the clearest dynamical signal possible.

3.5.1 Future regime pattern changes

Here we work with the 20 CMIP6 models with available SSP5-8.5 experiments as detailed in table 3.1. For each model we compute regimes for the end-of-century period 2070-2100 and for the period 1950-2010, which we shorthand as 'future' and 'historical' respectively, and compute spatial Z500 composites. These composites are then averaged across all models to produce a representative set of regime patterns for the recent historical period and another for the end of century climate. The anomaly between these two sets of regime patterns provides the ensemble mean forced response of the spatial regime structure. In order to assess whether pattern anomalies are statistically significant, a

bootstrap approach was used. For each model, 100 30-season samples of data, randomly selected non-consecutively from the period 1950-2010, were clustered and their Z500 composites computed. These were then pooled across the models to create 2000 cluster composites, capturing the distribution of pattern sampling variability in the recent historical period.

Figure 3.9 shows the resulting cluster composites, with stippling indicating gridpoints at which anomalies between the future and historical period are significant at the 5% level. The regime composites for the historical and future periods are visually similar, with no qualitative changes in the fundamental character of the circulation patterns. There are regions of statistically significant differences however, particularly clearly in the NAO- regime, which in the future climate has a weakened meridional dipole, and has the low shifted further east. The AR regime also features a slightly weakened dipole and a very minor eastward shift of the ridge. The BLK regime is largely unchanged but features a slight strengthening of its zonally oriented dipole. The sizes of the anomalies between future and historical regime patterns are at all gridpoints less than 25% of the strength of the circulation anomalies themselves.

We can follow the same procedure for the 4 classical circulation regimes which are shown in figure 3.10. Here much larger differences between the historical and future patterns can be seen but, with the exception of a small number of gridpoints in the NAO- regime, none of the changes are statistically significant; a consequence of the large historical sampling variability making it difficult to distinguish any signal from the noise. Nevertheless, the predicted changes in the NAO- pattern in the four regime case is consistent with what we see for the 3 geopotential-jet regimes, while the BLK and AR pattern changes are partially consistent, featuring the same increased positive blocking anomaly over Europe, and the same eastward displacement of the ridge.

By comparing figures 3.9 and 3.10 we see that the geopotential-jet regime framework delivers on its promise of allowing for more confident estimates of future regime change,via

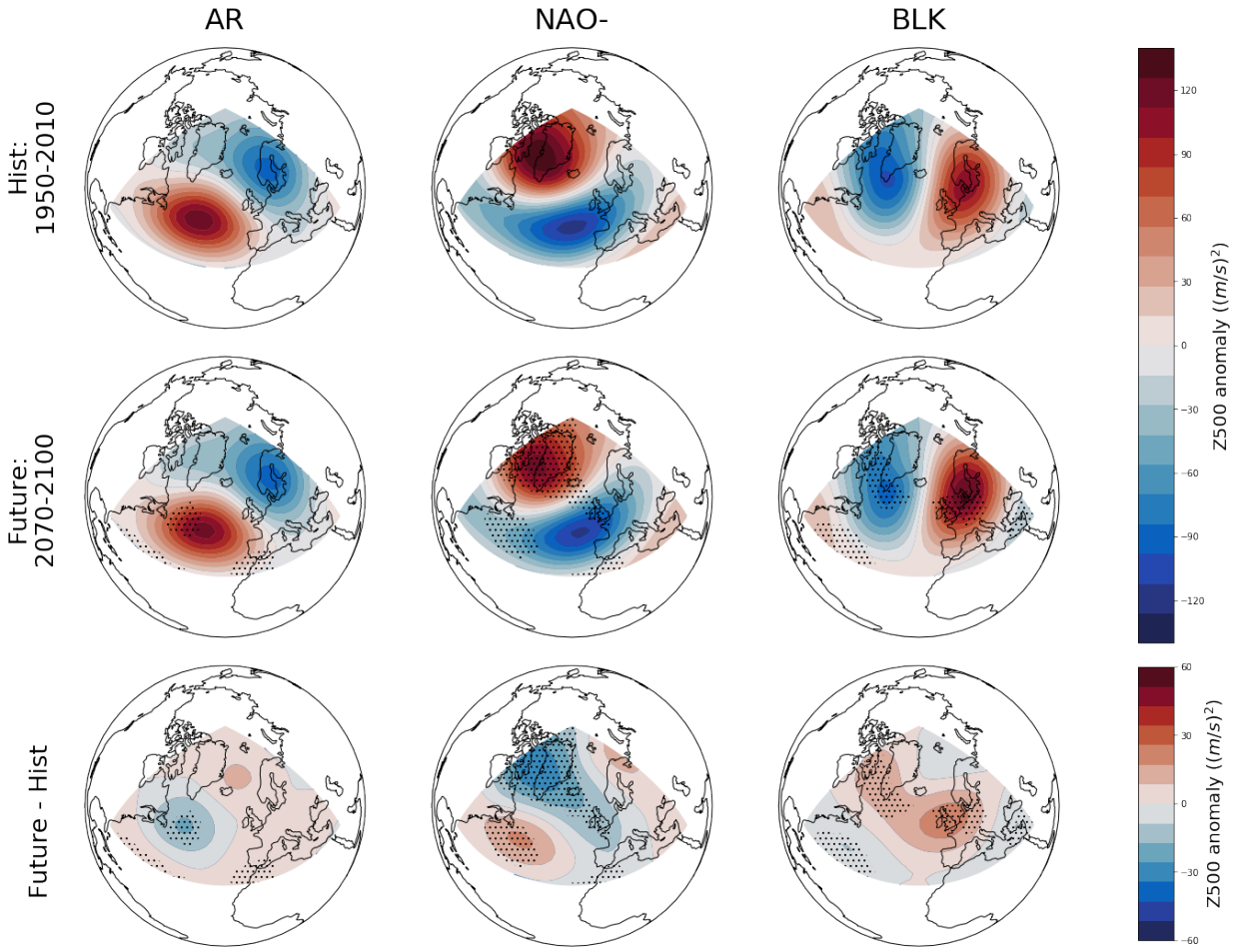


Figure 3.9: Top: Ensemble mean CMIP6 historical Z500 regime composites computed over the shared 1950-2010 period. Middle: Ensemble mean composites under the SSP5-8.5 forcing scenario between 2070-2100. Bottom: Anomalies between the future and historical regime composites. Stippling in the middle and bottom rows indicate grid-points at which the anomaly exceeds the 2.5th or 97.5th percentile of the bootstrapped sampling distribution (see main text for details). Note that the bottom row uses a different colourbar, in order to better show small changes.

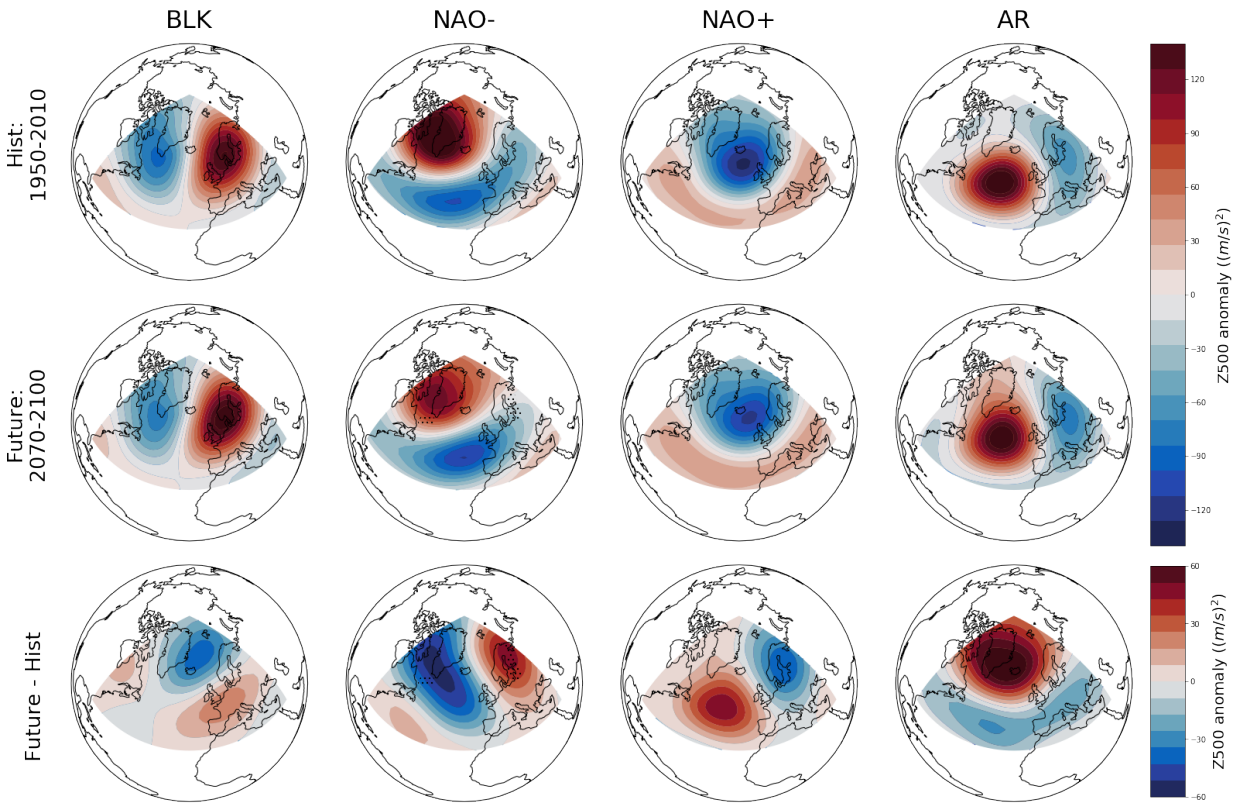


Figure 3.10: As in figure 3.9 but for 4 classical circulation regimes.

detection of statistically significant changes in circulation patterns.

These changes are small, and hard to detect by visual inspection of the regime anomalies. We can make a quantitative assessment of exactly how meaningful, if at all, these small pattern changes are to the overall dynamics by comparing how well the variability of the end-of-century Z500 field is explained by both the historical and future regime patterns. For each model, the pattern correlation between the daily detrended DJF Z500 anomaly field from 2070-2100 and the future regime pattern assigned to each day was calculated and then averaged across the 30 years, providing an assessment of how well the average flow pattern is explained by the three regime framework. This was then repeated, using the historical regime patterns. If the regime patterns really are fundamentally the same in past and future climates then the historical patterns should do just as well at explaining the future Z500 fields as the future patterns, whereas if the character of the regime structure is altered we should anticipate a commensurate drop in the ability of historical regimes to

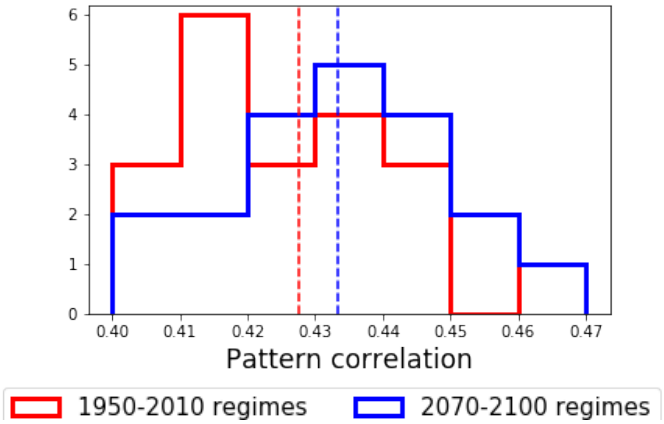


Figure 3.11: A histogram across CMIP6 models of the average pattern correlation between the daily detrended Z500 anomaly over the period 2070-2100 and the regime pattern assigned to each day, computed using both the historical and future regime patterns of each model. Dashed lines show the multimodel ensemble mean.

explain the variability of future circulation anomalies. A histogram of these correlations across the CMIP6 models is shown in figure 3.11. While the historical regime patterns show slightly lower correlation with future daily weather than the future patterns do, the difference is fairly small, with the ensemble mean correlation differing by only $\sim$0.05 (equivalently, a change in variance explained from 18.8% to 18.3%).

The ability of a regime framework to characterise seasonal anomalies in the circulation is also important, especially on climate timescales. We assess this as a second check, computing DJF mean Z500 anomalies for each model, and constructing regime proxies of the seasonal anomaly by averaging regime patterns together weighted by their seasonal regime occurrence. This was done both with and without the inclusion of the neutral state in the definition of regime occurrence and persistence, to assess the sensitivity of the results. Just as for the daily anomalies, seasonal proxies were constructed using both the historical and future regime patterns, and the correlation between each true seasonal anomaly and the corresponding regime proxy anomaly were computed and averaged across the 30 years as shown in figure 3.12. In this case the differences in the ensemble mean correlations are completely negligible with both sets of regime patterns equally suited to capturing the seasonal regime signal, with the historical patterns even

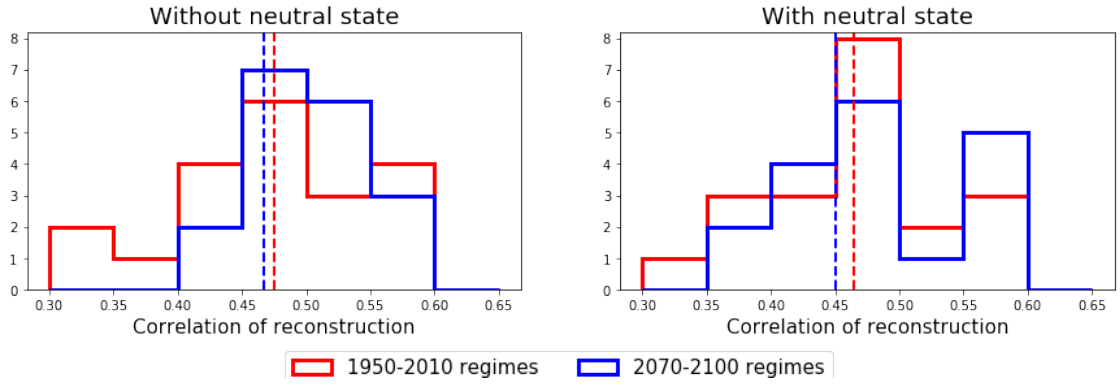


Figure 3.12: Left: A histogram across CMIP6 models of the average pattern correlation between seasonal mean Z500 anomalies, and regime composites constructed using seasonal regime occurrence, computed using both the historical and future regime patterns of each model. Dashed lines show the multimodel ensemble mean. Right: The same, but using regime occurrences computed while including a neutral state.

doing fractionally better.

Therefore we have validated the P99 conjecture that regime patterns are left essentially unaltered by climate change; there is a negligible spatial component to the forced regime response. The other half of this conjecture is that the forcing should manifest via a temporal regime response – altered regime occurrence – and we will now consider whether this also holds in the CMIP6 ensembles.

3.5.2 Anthropogenic trends in regime statistics

Regime occurrence and persistence are computed for the SSP5-8.5 CMIP6 experiments exactly as they were for the historical experiments except that while a 30-year rolling window was used to analyse historical regime variability, here we use a 60-year window in order to better capture the centennial timescale trend. In order to best highlight the changes in models' regime statistics we consider anomalies of regime occurrence and persistence, relative to each model's 1950-2010 mean value, rather than analyse the absolute values directly. Figure 3.13 shows the CMIP6 ensemble mean occurrence and persistence anomalies, with corresponding best fit linear trends and error-bars estimated using a leave-one-out bootstrap approach.

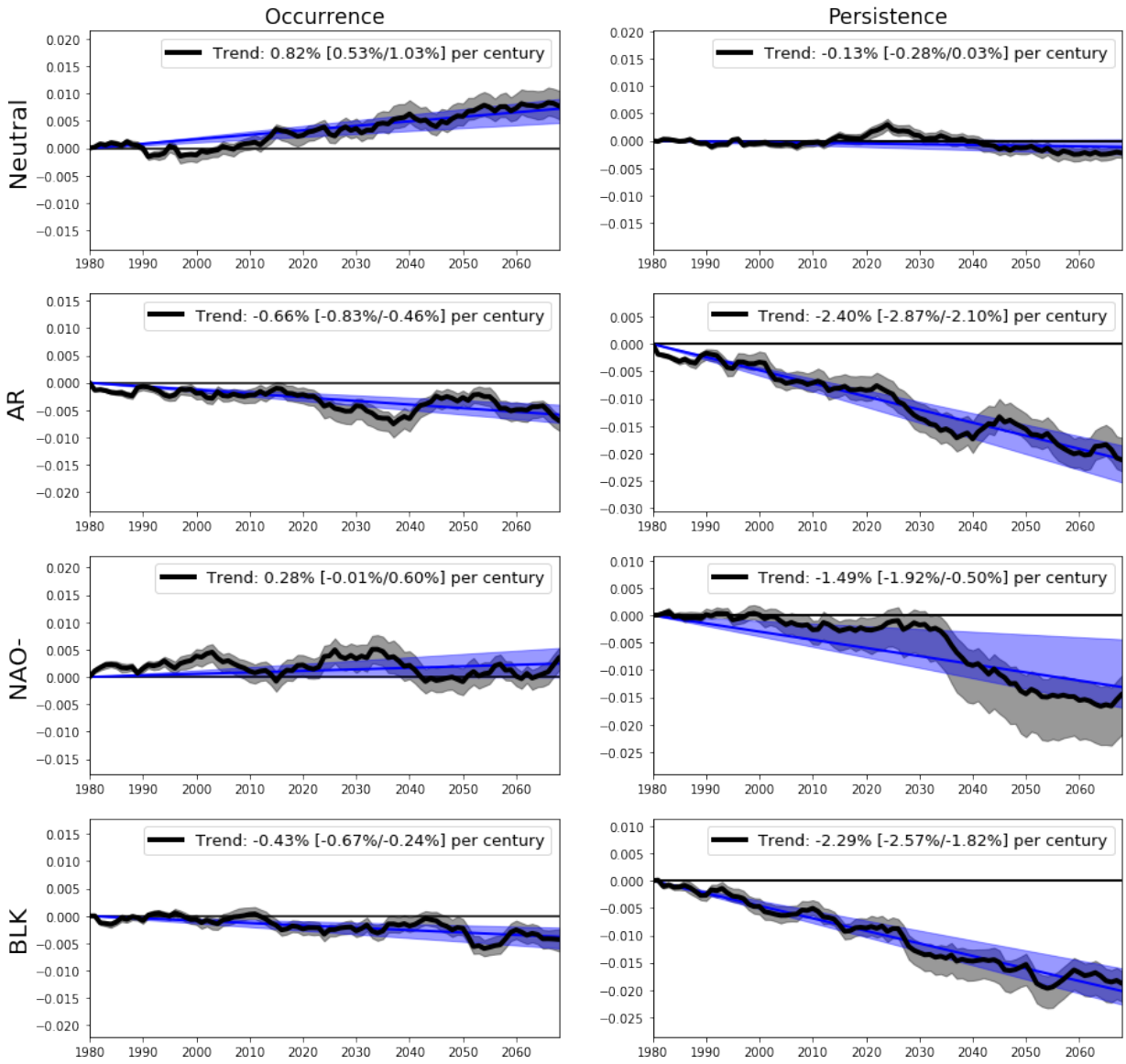


Figure 3.13: Black lines show the CMIP6 ensemble mean of 60-year average regime occurrence and persistence anomalies relative to the 1950-2010 reference period. Dates are the central year of the averaging window. Shading shows the standard error of the ensemble mean. The blue line shows a linear trend of occurrence/persistence against time, with shading showing the upper and lower bounds of trend errors calculated using a drop-one-out bootstrap.

Trends in regime occurrence are quite weak, in all cases less than 1% shifts over a 100-year period. The Neutral regime shows the largest response, becoming increasingly frequent at the expense of reduced AR and BLK frequency, while the NAO- regime shows no trend. The active regimes show a much stronger response in their regime persistences, universally showing a trend towards reduced regime lifetimes, while the persistence of the Neutral regime does not alter. Again the signal is strongest for the AR and BLK regimes, which show reductions in the probability of persistence of 2.4% and 2.3% respectively. The NAO- regime also shows a robust decrease in regime persistence, although not as strongly, with a 1.5% decrease in persistence probability over the century. These trends are weak compared to the interannual and even interdecadal regime variability seen in the historical record, but still represent significant shifts, equivalent to the magnitude of model bias seen in the historical period for some regimes. The ensemble mean picture is one of weakening regime dynamics in a warming climate, with regime events occurring less often, and lasting for less time when they do.

The ensemble mean picture however obscures considerable inter-model variability which adds complexity to the interpretation. Ensemble plumes of occurrence and persistence anomalies are shown in figures 3.14 and 3.15 respectively, with the ensemble mean trends reproduced in red for reference. For completeness, the variability prior to the 1950-2010 window is also included for those historical model experiments initialised at the beginning of the 20th century. In all cases the ensemble spread includes the origin, and there is no absolute consensus between the models on which regimes become more or less frequent in the future. The future plumes of regime occurrence anomaly tend to exceed the magnitude of the historical variability seen for the sub-set of full 20th century model runs, for the Neutral and BLK regimes in particular. The overall mean trends in occurrence come from a near cancellation of strongly disagreeing ensemble members, with a 60/40% split in the sign of the AR and Neutral trends, and a 55/45% split for the BLK occurrence trend.

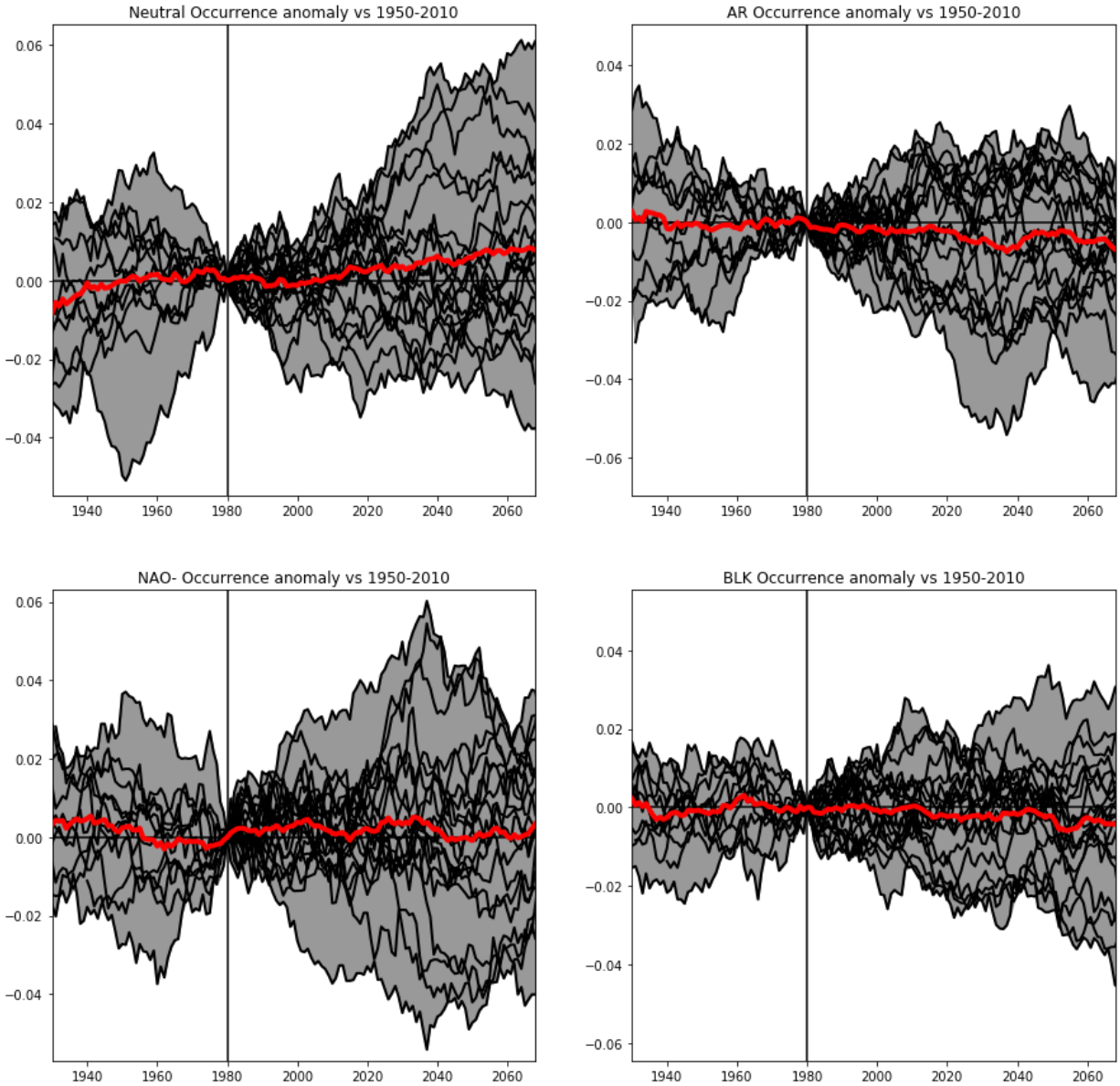


Figure 3.14: Red lines show 60-year ensemble mean regime occurrence anomalies, as shown in the left panels of figure 3.13. Dates are the central year of the averaging window. Black lines show each individual ensemble member, with the total model spread enclosed in the grey shading. The vertical line marks the reference period of 1950-2010.

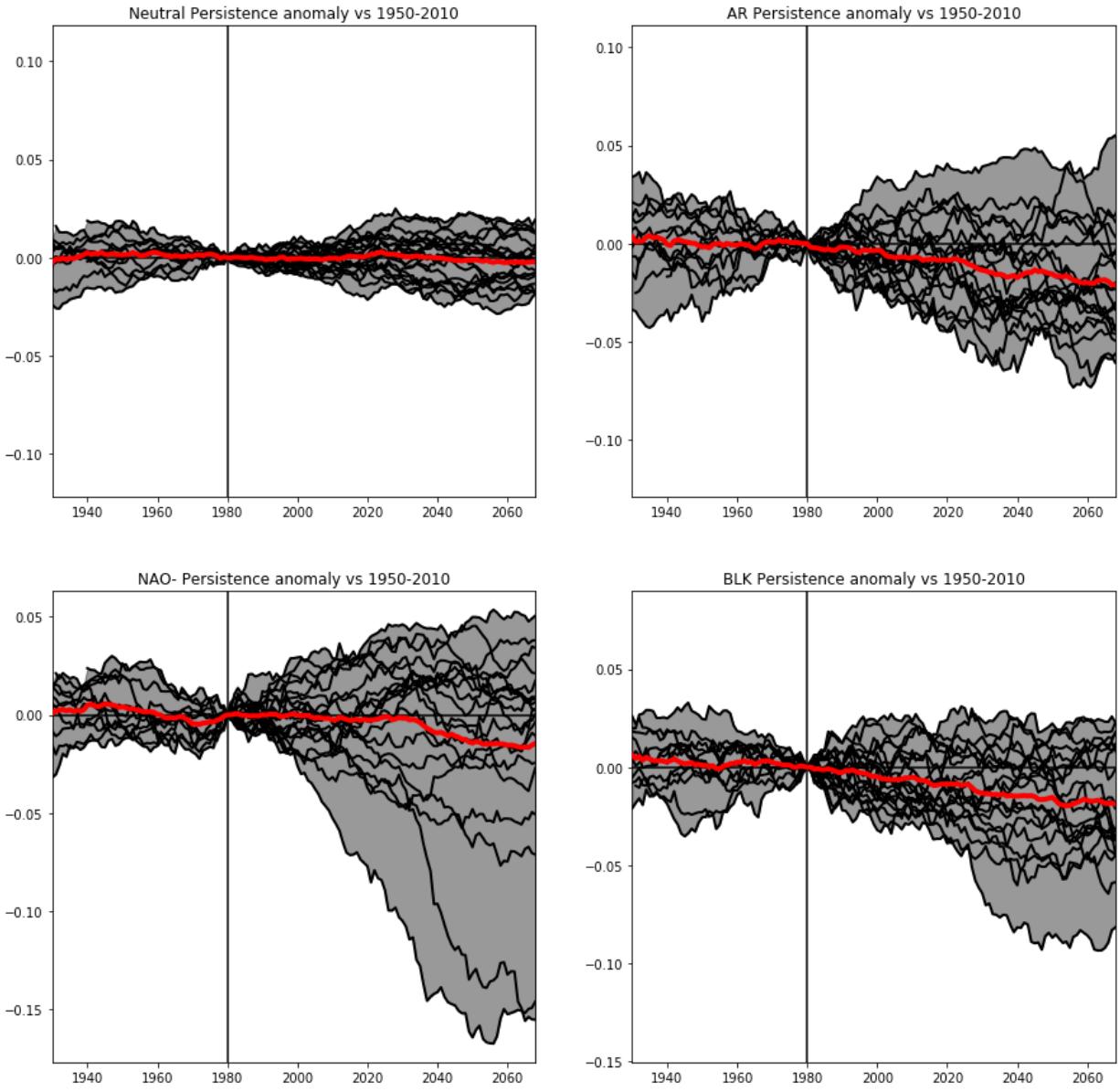


Figure 3.15: As in figure 3.14 but for regime persistence.

For regime persistence, there is slightly higher coherence, especially for the reduced AR persistence on which 75% of models agree (an outcome that would only occur 4% of the time by chance). The trend in NAO- regime persistence is particularly uncertain. The ensemble mean NAO- persistence trend does not feature an approximately linear decrease as in the other regimes, but seems to instead drop suddenly after the 2005-2065 period. In the plumes, this can be seen to result from the severe persistence decrease of a small number of models, which experience persistence drops exceeding 10%. In fact, only 7 of the 20 models predict an overall decrease in NAO- regime persistence in the 21st century, but the mean value is heavily skewed by these outlying experiments. It is worth noting that the two most extreme outliers in NAO- persistence are from the same centre, the Met Office UKESM1-0-LL and HadGEM3-GC31-LL models, and so are not completely independent models. The same effect can be seen to a much lesser degree in the plume of BLK persistence trends, with a few models projecting particularly strong decreases in persistence. The BLK and NAO- persistence trends are linked, as models which project decreased BLK persistence also tend to project decreased NAO- persistence (not shown).

Overall then, while there are large disagreements between climate models on the exact details of future changes in regime statistics, many models do show large alterations in the frequency of occurrence and persistence of their regimes, indicating a temporal regime response to climate change that dominates over any small spatial changes. CMIP6 projects an overall weakening of the Euro-Atlantic regime structure, with weakened meridional dipoles and with rarer, less long-lived regime events. While the overall signal of this change is small, some models predict very large decreases in regime persistence for the NAO-, and to a lesser extent the BLK, regime. We should also bear in mind the well-known problem in seasonal-forecasting of the signal-to-noise problem, whereby ensemble mean signals in regime systems can actually be far lower in amplitude than they should be given the predictive skill of the forecast [Scaife and Smith, 2018]. If such a signal-to-noise issue is also relevant on climactic timescales – as hinted at by its presence in decadal forecasts [Smith et al., 2019] – then the true response of the Euro-Atlantic

regime structure might be far larger than the modest trends predicted here suggest. Assuming that the trends here are representative of the true climate response however, then outside the most extreme cases discussed, interdecadal variability is likely to dominate, or at least be equal in magnitude, to the forced trend in the coming century – take for example the 6% variation in NAO- regime persistence between 1960-1990 and 1980-2010. This emphasises the importance of using initialised multidecadal forecasts to understand the actual regime shifts likely to be seen in the 21st century. Conjecturally, it may be that part of the reason climate projections over Europe are so uncertain is that they are not properly taking account of the importance of initial conditions. While even recently, skilful decadal forecasts of the Euro-Atlantic circulation were purely an ambition [Smith et al., 2016], the latest generation of models are showing some success in predicting multidecadal variations in blocking and the NAO [Athanasiadis et al., 2020], and so their useful application to future climate projection is becoming a real possibility.

3.6 Regime response of the Lorenz '63 system under multiplicative forcing

Just as was hypothesised in [Palmer, 1999] (P99), CMIP6 models experience only very small changes in their regime patterns under climate change, with historical regime patterns explaining future Z500 variability just as effectively as future regime patterns. However the details of the temporal response are only partially in agreement with P99, which predicted forcing to project primarily onto occurrence changes. Instead we saw only modest changes in regime occurrence in CMIP6, while the bulk of the response was in a decrease in regime persistence.

In P99, where additive forcing terms are applied to the Lorenz 63 system, changes in regime occurrence are directly correlated to persistence changes – the more frequently occurring regime also shows increased persistence – whereas in CMIP6 we see all active

regimes becoming less persistent, without a clearly linked decrease in occurrence in all cases.

We will now show that a forced persistence decrease can be reproduced in the Lorenz '63 system, if we generalise the applied forcing to also include modifications to the ρ and β coefficients as well as the additive forcing term. Specifically we consider the system:

$$\dot{x} = \sigma\left(y - x\right) + F \tag{3.1}$$

$$\dot{y} = x\left(\rho_F - z\right) - y + F \tag{3.2}$$

$$\dot{z} = xy - \beta_F z \tag{3.3}$$

where $\rho_F = \rho + \gamma F$ and $\beta_F = \beta - \frac{\beta\gamma F}{\gamma F + \rho - 1}$, and γ is a coefficient setting the relative amplitude of the multplicative and additive forcings. The purely additive model of P99 corresponds to the case $\gamma = 0$, whereas here we will use $\gamma = 10$.

The particular form of the multiplicative forcing is chosen such that $\beta_F(\rho_F - 1)$ remains constant for all F, which ensures the x- and y- coordinates of the unstable fixed points that define the dynamics of the Lorenz attractor also remain constant. In general, the z-values of the equilibria, which are equal to $\rho - 1$, will change with F when γ is nonzero.

Integrations of the forced system were performed for F varying between 0 and 2.8 for 200,000 MTU with samples taken every 0.1 MTU after a 5000 MTU spinup period. Two regimes were defined simply by evaluating the sign of the x variable.

Figure 3.16 shows the Lorenz attractor both for $F = 0$ and for the strongest forcing applied, $F = 2.8$, with the mean value of the points in each regime marked for comparison. In the $x - y$ plane the size of the attractor expands slightly but the mean positions of points in each regime – the centroids – remain approximately stationary. The whole attractor is translated however towards higher z values, and again expands slightly.

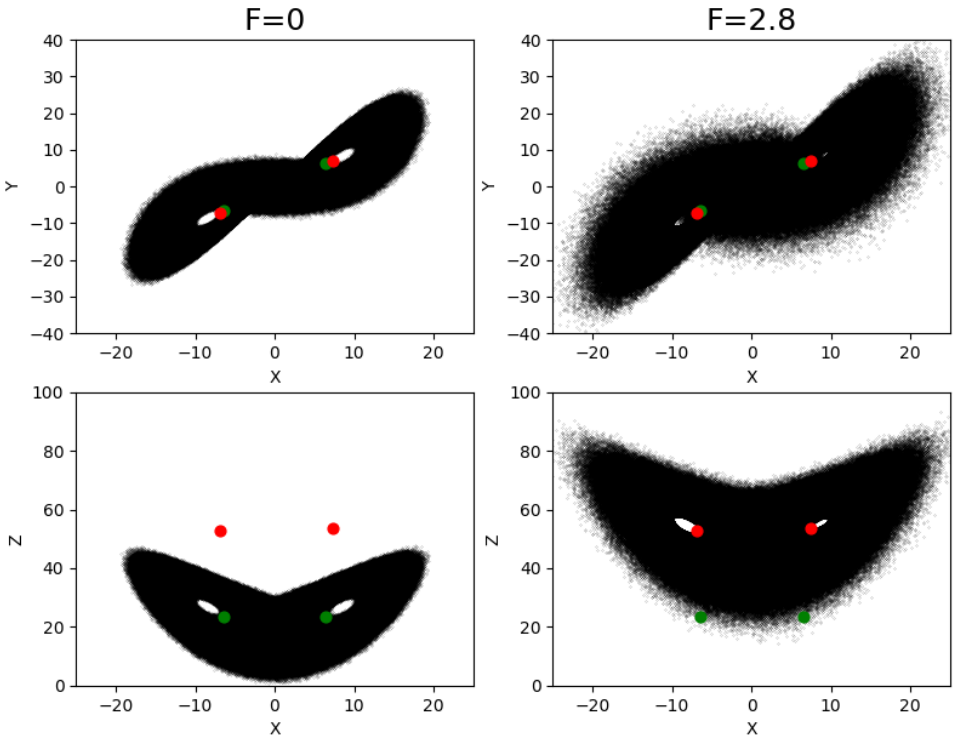


Figure 3.16: The x-y (top) and x-z (bottom) crossections of the forced Lorenz attractor, for $F = 0$ and $F = 2.8$. Green dots mark the mean position of points in the $F = 0$ regimes, while red dots mark the same for points in the $F = 2.8$ regimes. Only very small changes in the average position in the $x - y$ plane can be seen as forcing increases, whereas the z value increases by 28.

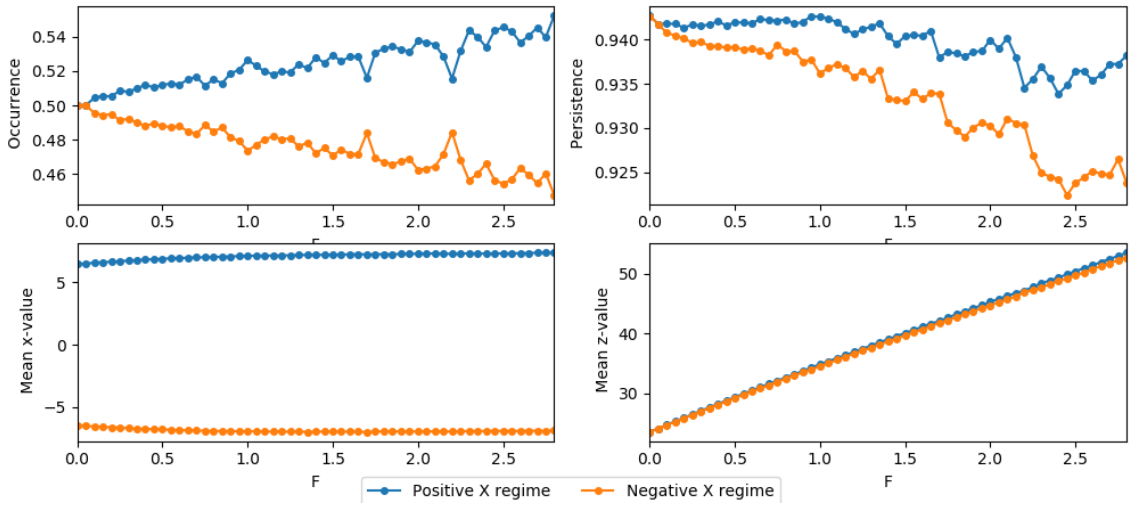


Figure 3.17: Changes in summary regime statistics as the strength of the joint additive and multiplicative forcing applied to the Lorenz attractor is increased. Due to the symmetry of the system, changes in the mean y state are exactly equivalent to changes in the mean x state and so are not shown.

The resulting changes in regime occurrence and persistence, as well as the shift of the regime centroids is shown for the full range of forcings in figure 3.17. The temporal aspect of the regime response captures well that seen in the CMIP6 SSP5-8.5 experiments: the occurrence frequencies change, with a more frequent positive x regime, while the persistence of both regimes decreases. The decrease is asymmetrical, with lower persistence observed for the negative x regime. Shifts in the x (and the not-shown y) component of the centroid are very minor, again in line with the minor regime pattern changes seen in CMIP6. The linear increase of the average z value with F does not appear to match the CMIP6 results however, nor strictly the hypothesised stability of regime centroids to forcing supposed in P99.

However, in our analysis of regime structure in reanalysis and complex models we have focused on the horizontal spatial structure of regimes, on the basis that blocking anomalies and the dynamics of the NAO can be understood as close to equivalent barotropic. Therefore, the vertical structure of pressure anomalies is considered of only secondary importance. Within the original approximation of Rayleigh-Bénard convection that [Lorenz,

1963] used to derive the L63 model, the x and y variables represent the intensity of convection and the horizontal temperature gradient, resulting from ascending and descending air, respectively. On the other hand, the z variable represents the deviation of the vertical temperature profile from linearity, capturing the degree of vertical instability. Therefore, that the regimes remain approximately constant in spatial structure within the $x-y$ plane can be seen as maintaining a constant horizontal regime structure, while the vertical profile does respond to forcing.

Viewed in this light, the translation of the attractor to increasing z values with forcing is quite consistent with the geopotential-jet regime picture, as the vertical expansion of the atmosphere under global warming introduces a trend in geopotential height, which was removed before any principal components or regimes were computed. Therefore we have already implicitly allowed for a translation of the phase space under forcing, without considering it to have altered the regime dynamics themselves, which fundamentally are defined by anomalies to the mean state.

In summary, by extending the L63 system to contain additive and multiplicative forcing terms, we can qualitatively reproduce the simultaneous occurrence and persistence trends seen in CMIP6 without significantly altering the fundamental structure of the regimes, as long as we understand translations of the system to represent in some sense trivial changes, just as is routinely done for more complex regime systems. This validates the usefulness of the Lorenz system for qualitatively understanding the forced response of regime systems, as part of the hierarchy of model complexity.

3.7 Discussion and conclusions

In this chapter we have showed the usefulness of the geopotential-jet regime framework for studying both the historical variability and forced response of the Euro-Atlantic circulation in climate models, and demonstrated that three highly stable anticyclonic regimes characterised by increased blocking frequency can be found in almost all models consid-

ered.

We have shown that CMIP6 models have a generally improved regime structure over the previous generation of models, in terms of robust improvements in their spatial patterns. This has been accompanied by lower biases in both the jet speed and jet latitude indices, and we have found that a smaller error in the distribution of jet latitudes is a robust indicator of the fidelity of a model's regime patterns, highlighting the interconnection between the jet and circulation regime perspectives as recently emphasised in [Dorrington and Strommen, 2020, Madonna et al., 2017, Madonna et al., 2021]. While we have not examined the physical origins of regime and jet biases here, it is quite possible that systematic improvements in resolution between CMIP5 and CMIP6 are responsible for much of the improved regime structure, as was found in [Dawson et al., 2012, Strommen et al., 2019, Fabiano et al., 2020]. Investigating the impact of model biases on regime structure is an important extension of the work in this chapter which is currently ongoing.

We have found that CMIP6 models have a systematic bias towards a too infrequent BLK regime and that both the NAO- and BLK regime are underpersistent when comparing against a centennial reference period. The underestimated NAO- persistence is also seen in the classical circulation regime study of Primavera done by [Fabiano et al., 2020]. Occurrence biases seen in CMIP5 are heavily reduced, but no clear improvement in persistence is seen between the two model ensembles. We have also shown that while the historical interannual variability in regime occurrence and persistence is well captured by models, the multidecadal variability in regime occurrence is too low, while multidecadal persistence variability is too high in CMIP6. It is not entirely clear what might drive such a variability bias.

It is natural to ask how climate model biases towards underpersistent regimes can be ultimately eliminated. Given that the resolution upgrade associated with the transition from CMIP5 to CMIP6 does not suffice to eliminate the persistence bias, it is likely that simply continuing to improve model resolution is not enough. Studies investigating the impact of

moderate horizontal resolution upgrades (i.e. investigating resolution variations between 100 km and 10 km) often report only minor improvements in blocking representation [Schiemann et al., 2017, Davini et al., 2017a] or even no improvement [Schiemann et al., 2020].[Scaife et al., 2019] reported a substantial improvement in the strength of synoptic eddy forcing as horizontal grid spacing dropped below 0.4°, but such high resolutions are currently unfeasible for practical applications. If correctly capturing real world eddy forcing is the main barrier to modelling persistent regimes then we may expect to be haunted by this model persistence bias for the foreseeable future. However, [Pithan et al., 2016] suggests that much of the CMIP5 biases in the climatology of blocking and the jet can be explained as a result of poorly represented orographic drag. This is in line with [Berckmans et al., 2013] which showed that it was the improved orography associated with a resolution upgrade that most improved the Euro-Atlantic circulation by heightening the Rocky Mountains and changing the climatological standing wave pattern. If it is in fact poor representation of orography that is the main barrier to persistent regimes, then we may be able to substantially reduce persistence bias by making use of more sophisticated – possibly stochastic – orographic drag schemes. It is probable that these factors are not independent of each other; most likely both a moderately high resolution *and* improved orographic parameterisations are needed in tandem. However, to move such claims beyond mere conjecture, there is need for additional theoretical work exploring the interaction of improved parameterisations with improved horizontal, vertical and ocean resolution on the Euro-Atlantic circulation.

The impact of climate change, as represented by the SSP5-8.5 scenario, only weakly projects onto the regime patterns themselves, with a slight weakening of the NAO- and AR dipoles consistent with [Masato et al., 2014] but with little quantitative degradation in the ability of historically established regime patterns to explain future Z500 variability. Changes in regime occurrence and persistence paint a picture of a less-regime dominated circulation, with both less persistent and less frequent anticyclonic regimes, which is consistent with previous jet, classical circulation regime, and blocking based approaches to

understanding Euro-Atlantic circulation trends. These trends however are relatively small compared to the internal variability in regime occurrence and persistence, to the extent that the 21st century European circulation might still be defined more by initial conditions than by the changing boundary conditions of anthropogenic warming. This suggests an important potential role for multi-decadal forecasts with initialised ocean states in making useful predictions of future European climate, which can also be validated against a shorter historical baseline.

We also found that the mean CMIP6 trends were heavily skewed in some cases by a minority of models with very strong forced responses, such that the ensemble mean trend does not actually represent a valid ensemble member. Since we have no a priori grounds for discounting the behaviour of such outliers, and their difficulties with interpreting a multi-model ensemble of opportunity in a probabilistic sense [Knutti et al., 2013], it is difficult to quantitatively synthesise the model trends. The value of using storylines – different physically coherent scenarios of future change without precise probabilities attached – for presenting circulation changes has been emphasised in e.g. [Zappa and Shepherd, 2017], and developing a story line framing for the climate responses documented here is also a basis of ongoing work. This would also allow for an exploration of non-linear tipping points which may be behind the sudden mid-century declines in regime persistence seen in some models. By comparison to [Fabiano et al., 2021] we can suggest that the occurrence and persistence trends would be even smaller still in more optimistic emissions scenarios than SSP5-8.5, which would suggest a minimal role for circulation changes under climate change. However given the known signal-to-noise issue in the North Atlantic, we cannot discount the possibility that the strength of the regime trends in CMIP6 is underestimated.

Through our separation and individual consideration of changes in regime patterns and regime statistics we have provided the first test of the Palmer '99 hypothesis that regime response in a forced system should manifest through roughly stationary regime patterns,

and shifts in regime occurrence. We have shown that this is broadly accurate although the simple additive forcing used there can not explain substantial persistence decreases in regimes without corresponding reductions in occurrence. By extending the Lorenz model to include multiplicative forcing we can indeed capture a qualitatively similar forced regime response to that seen in CMIP6, providing a qualified validation of the Palmer '99 hypothesis, and once again demonstrating the capacity of simple toy systems to capture the complex non-linear dynamical phenomena seen in our most advanced Earth system models.

4 | Conclusions

In this thesis we have aimed to deepen our understanding of circulation regimes in the Euro-Atlantic wintertime atmosphere by addressing three distinct questions:

1. How does stochastic forcing influence regime systems?
2. Why has the circulation regime structure of the Euro-Atlantic been so hard to statistically pin down in a robust way?
3. How will anthropogenic global warming impact the Euro-Atlantic regime structure?

Our strategy for tackling these questions has involved analysis of regimes in real-world observations, state-of-the-art Earth system models and low-dimensional dynamical systems, as we attempted to use simplicity and comprehensibility where possible, and realism and complexity where necessary.

How does stochastic forcing influence regime systems?

Since the response of sophisticated models to stochastic forcing strongly depends on the details of the parameterised processes that are randomly perturbed, and results in global mean state changes, it can be challenging to understand why a given model's regime structure responds in a particular way to stochasticity. We therefore looked to understand in detail the phenomena of stochastically induced regime persistence in the Charney deVore model, to look for lessons in a reduced-complexity setting that would be useful in interpreting the regime response in SPHINX simulations.

We obtained a complete understanding of the CdV model's response to stochastic forcing and succeeded in framing the dynamical mechanisms in physically interpretable terms, thereby providing a better understanding of the results of [Crommelin et al., 2004] and [Kwasniok, 2014]. The blocking regime features a reduced dimension of the unstable manifold, partly caused by a phase locking between wave modes, confining the the sys-

tem to a single one-dimensional trajectory via which thermal relaxation can destabilise the blocked flow state. We showed that symmetric perturbations to the zonally symmetric modes, which can be understood as westerly or easterly wind gusts, would on average lengthen the amount of time taken for a blocking event to decay, while perturbations to the wave modes are strongly damped by the phase locking. This was compounded by the tendency of the system to linger near an unstable fixed point representing a strongly blocked flow. We found that eddy forcing played no role in the increased persistence; instead the changes in the model mean state were key, supporting the view that the main benefit of stochasticity is that it allows a model to explore a larger part of phase space. Of course, the lack of eddy forcing in this heavily truncated model does not preclude its relevance in more complex models with true multi-scale dynamics. Building on this work, a natural next step is to ask how stochasticity impacts regimes in quasigeostrophic settings, and to explore how the mechanisms uncovered here evolve (or indeed, are rendered extinct) as the complexity of the flow increases. Order-to-disorder transitions played an important role in the stochastic persistence of the CdV system, and much of the direct relevance of these results therefore depends on whether such transitions have a significant presence in the true atmospheric flow – a currently unanswered question. In order to better link simple model behaviour to that of complex models we must continue to develop tools that can be used to study both kinds of systems. The theory of unstable periodic orbits may be useful here, and it has recently been suggested by [Maiocchi, C. C. and Lucarini, 2021] that regime dynamics may be thought of as movements between bundles of UPOs. How well does this conception of regimes match our traditional intuition? Can the influence of stochasticity on UPO transitions be formalised? Tackling these questions may ultimately provide an avenue for a more rigorous understanding of the regime dynamics explored in this thesis.

We found that the introduction of stochastic parameterisation into the SPHINX EC-Earth model had the effect of significantly improving spatial regime structure, and eliminated biases in the AR regime, results we have shown to be robust to sub-sampling. Notably,

this improvement was only detectable due to the large decrease in regime sampling variability that came from moving to the geopotential-jet regime framework. Our attempts to explain the stochastically forced response of the SPHINX EC-Earth simulations, however, were more limited, and it was not easy to leverage insights from the CdV model. A major difference between the regimes of CdV and Euro-Atlantic cirulation regimes is that while the former has periodic zonal boundaries, and so represents a global hemispheric wave, the latter is rooted in a sectoral understanding of regimes. How similar are the dynamics of resonant stationary waves and local blocking anomalies? While the term âblockingâ is often applied to both, it may be important to disentangle these different concepts to better understand their origin and to identify any distinctions in their behaviour. Another challenge of comparison between the models was the different types of stochasticity introduced: the additive CdV forcing is quite different to the multiplicative forcing that dominates in the EC-Earth simulations. Again, it may be that different kinds of stochastic parameterisation have very different dynamical impacts, and these differences should be considered more closely in the future.

While in SPHINX we were ultimately able to link an increase in NAO- occurrence with a decrease in zonal wind speed south of Greenland, the NAO- and BLK regimes both showed persistence decreases which we could not explain. Although the AR regime experienced increased persistence, it was unlikely to be a result of a CdV-like mechanism, due to the lack of orography in the middle of the Atlantic. We conjecture that well-documented improvements in tropical climate with stochastic physics drive the improved AR regime. However, we still do not know exactly what the causal impacts of stochasticity are in the SPHINX model. Going forward, more targeted experiments may be needed to isolate the cause and effect of stochastically induced circulation changes, such as applying stochasticity only in specific regions, or with regional relaxation experiments. This would address the important question of where the impact of stochasticity is strongest in models, and whether its impacts are primarily local, or rooted in modified teleconnection patterns. Intermediate complexity models might be well suited to such a

task if their regime structure can be shown to be adequate.

Why has the circulation regime structure of the Euro-Atlantic been so hard to statistically pin down in a robust way?

The Atlantic jet stream is now well understood to feature both latitudinal variability between three preferred states, and 'pulsing' variations in its speed that are linearly distributed, orthogonal to the latitudinal shifting, and subject to multidecadal variability. By recognising the blocking dynamics of circulation regimes as being inherently linked to the behaviour of the jet stream, we posited a similar orthogonality between the jet speed and the regime structure captured in circulation anomalies. In CMIP6 data we did in fact see that models with large jet latitude biases had worse regime structure, whereas the same could not be said for jet speed biases, providing evidence that such an orthogonality exists. We suggested that this fundamentally linear variability might be obscuring the non-linear structure of the Z500 phase space and so make it difficult for clustering algorithms to find the underlying regimes. We regressed out jet speed prior to clustering, and found that the resulting geopotential jet regime framework produced three statistically robust regimes. When considering stability of regime patterns to sub-sampling as a way of choosing regime number, the choice of three regimes naturally emerges.

While in recent years statistical sharpness tests have proved popular for choosing regime number, our approach is very similar to the subsampling recommended in [Michelangeli et al., 1995]. These regime patterns map closely to the three anticyclonic patterns found when using four classical circulation regimes, but with the cyclonic NAO+ pattern excluded. This is consistent with the view of [Woollings et al., 2008] which suggested that although blocking exhibited non-linear persistence and recurrence, the zonal flow associated with the NAO+ is simply a default state. Similarly, [Faranda et al., 2016] argues that the North Atlantic is a single basin attractor that irregularly visits the vicinity of a Greenland-blocking-like fixed point. We remember that regimes are not weather types; it is not required that all qualitatively distinct flow configurations correspond to some

distinct phase space structure.

We isolated these geopotential-jet regime patterns in five reanalyses, in the SPHINX EC-Earth experiments and in a large number of CMIP5/6 models. In all cases they proved highly robust – often more stable to subsampling than the NAO dipole – providing strong evidence for their physical significance. We did observe some disagreement between the temporal statistics of regimes in different reanalyses, especially in the pre-satellite era, and we noted that the stability properties of ERA5 were quite different to the other reanalyses, even when accounting for their differing time periods. We have not delved deeply into these inter-reanalysis differences in this thesis, but it is certainly important to try and understand which analyses are closest to the true history of the atmosphere, and to explore how the different data assimilation strategies of different reanalyses impact the resulting regime time series obtained.

Despite these disagreements, the general robustness of our results in many different datasets suggests there is considerable evidence that jet speed is indeed obscuring classical circulation regime structure, and that adopting the geopotential-jet regime approach put forward here is one way to overcome that, although other approaches may also exist. Regardless of approach, reducing sampling variability in regime identification considerably extends the usefulness and power of regime analysis, making it easier to compare models, characterise historical variability and project forced circulation changes.

We have not addressed the question of *why* different climate models do better or worse at representing the Euro-Atlantic circulation in this thesis, which of course is a vital step in ultimately eliminating model bias. Ongoing work based on this thesis is tackling this question and revealing strong correlations between CMIP model biases and geopotential-jet regime statistics, promising to shed some light on the physical causes of regime dynamics. Another avenue for future work is the exploration of regime dynamics in AMIP runs and in initialised models, so that we can see whether the historical variability patterns in occurrence and persistence seen in reanalysis can be replicated. By comparing

the free-running models used in this thesis, with models with a prescribed ocean and models with an initialised ocean and atmosphere, it should be possible to isolate the main drivers of long-term regime variability.

While we have explored the dynamics of regimes in both historical and future contexts in considerable detail, we have not directly addressed the reason for their practical importance – namely, how do regime patterns influence surface weather? Given that we introduce a new set of regime patterns in this thesis, there are important questions to answer in future work concerning the teleconnections between these regimes and surface temperature, precipitation and wind anomalies over Western Europe. Similarly, the flow-dependent predictability of these states may be different from their corresponding classical regimes, and this must also be explored. Such work will help to reveal whether the more detailed seven regime picture we briefly discuss in Chapter 2 is necessary for impact related work.

There is also more work to do to understand the covariability of circulation and jet regimes, and a clearer connection between the variation of blocking indices and regime statistics is needed before we can truly describe the full extent of Euro-Atlantic circulation variability. Importantly, despite the value of the geopotential-jet regime approach we have demonstrated here, we do not have a clear theoretical justification for why we can consider blocking dynamics as independent of jet speed. While we can appeal to the orthogonality of the jet latitude and jet speed, we also do not have a clear model of the trimodal jet dynamics or of why the "wobbling" and "pulsing" modes of the jet vary independently. Indeed, more fundamentally, we do not know why there are any regime dynamics at all in the atmosphere! In parallel to these conceptual issues, the implications for applied meteorology in terms of the synoptic impacts of jet and circulation regimes need further exploration, especially in terms of their covariable impacts; that is how jet latitude and circulation regimes interact to impact surface weather.

How will anthropogenic global warming impact the Euro-Atlantic regime structure?

For the first time, we tested the hypothesis of [Palmer, 1999] and found that the influence of anthropogenic global warming on the Euro-Atlantic circulation does indeed primarily manifest via changes in temporal regime statistics rather than shifts in the spatial regime patterns. The overall trend is towards a less pronounced regime structure in a warmer climate, with robust decreases in the persistence of the AR regime, and several models predicting substantial decreases in NAO- and BLK persistence. Neutral conditions are also projected to become more prevalent in the future, and the slight spatial changes seen were towards weaker meridional dipoles in the NAO- and AR regimes. These results are consistent with work suggesting a more zonal and narrower jet in the future, and validates the conclusions of previous studies such as [Fabiano et al., 2021], extending them by considering spatial changes and by considering the inter-model variability. Although the overall trends we have identified are fairly small, on the order of occurrence and persistence changes of 1-2% per century even in the worst case SSP5-8.5 scenario we consider, the mean value masks substantial model spread. Thus, if the low signal-to-noise issue seen in seasonal and decadal forecasts is relevant on climactic timescales, the true impact may be larger than predicted here, even in the absence of non-linear tipping points. If this is the case, then the societal implications are severe – we may be drastically underestimating the potential for wintertime extremes over Western Europe. However despite the importance of the question, it is not clear how to test this possibility. One approach may lie in leveraging seamless prediction, using models that can be run in both forecast and climate modes, and characterising how the signal-to-noise ratio evolves on longer averaging timescales in order to project forward the reliability of the climate model ensemble. Even if the CMIP6 ensemble spread is reliable, it is still uncomfortably large, and future work needs to explore why different models predict such radically different changes in regime dynamics. Can we find emergent constraints that will tell us which models we should trust?

By exploring future regime change in CMIP6, we explicitly tested the hypothesis of [Palmer, 1999], which introduced additive forcing into the Lorenz '63 system. While

they proposed that climate change would manifest mainly in regime occurrence changes, rather than in the persistence trends that dominated here, we were able to extend the forcing profile of Lorenz '63 to qualitatively match the CMIP6 results. In the case of climate change, where we are considering the impact of large scale climactic changes on large scale circulation patterns, low-dimensional regime systems are perhaps better suited for aiding our understanding here than they are in the more multi-scale stochastic forcing problem.

A valuable extension to the climate change analysis done here would be to frame the results in a story line approach, which would allow qualitatively different model scenarios to be considered on an equal footing. Additionally, constant regime patterns may experience climate dependent impacts, and exploring the distribution of extreme weather events associated with persistent regime events in both the current and possible future climate would provide an explicit connection to applied impacts.

In Summary

One of the key ideas that guided the approach taken in this thesis was a belief in the value of working across a hierarchy of models. In simple models we can easily explore new ideas in a clean, controllable setting, while in complex models we can always connect ourselves back to realistic dynamics, and consider the full impacts of the many interacting components of the climate system. How well has this attempt to work across the model hierarchy gone? We have certainly managed to achieve some important successes, particularly in Chapter 3, where we have used the conceptual framework of regime forcing in Lorenz '63 to understand the drastically more complex issue of anthropogenically forced regime change in CMIP6. This provides a firm footing for future regime analysis and has highlighted additional subtleties – the role of persistence trends – that have helped us refine the original, simple model. We have also managed to explain the interactions of stochasticity with the regimes of CdV, and so have highlighted the underlying geometric mechanisms involved and related them to physically meaningful variables. However

we have also faced challenges along the way that we could not always overcome. The stochastic regimes of SPHINX could not be closely connected to those of CdV – the gulf in the complexity of the systems was too large to span. This issue was compounded by the difficulties in extending the CdV model to more modes.

As we have developed our understanding of the CdV model we have also highlighted some of the unrealistic characteristics it possesses: it lacks the number of modes required to produce a true separation between eddies and the basic state, its blocking dynamics are closer to resonant waves than localised blocks, and its bimodal streamfunction behaviour is closely linked to a corresponding bimodality in zonal wind speed (again a consequence of the small number of modes). The important role of weak versus strong westerlies in the dynamics of the CdV model, has been noted and questioned in [Benzi et al., 1986], and is particularly in tension with the geopotential jet regime framework developed in Chapter 2, where we argue for neglecting the role of zonal jet speed in regime dynamics. Of course, the insights that guided Chapters 2 and 3 were not available at the time the work in Chapter 1 was developed, and so the potential importance of multiple regimes existing for a single value of zonal wind speed was not fully appreciated.

This is not to say that the work of Chapter 1 has been invalidated by our later approach, or that the CdV model is too simplified to be of use. The realism of the system is greater than, say, for the Lorenz '63 model, and the vorticity dynamics at play have a clear connection to the true atmosphere. Ultimately, however, rather than managing to build a bridge between the theories of simple and complex stochastic regimes, we have instead only isolated important dynamical building blocks. For future work on the theoretical basis of regimes, especially when trying to consider the implications for both jet latitude shifts and blocking anomalies, we recommend the use of more fully converged models – perhaps with several dozen modes – with multiple regimes emerging for a single zonal wind speed. Such a regime system can emerge via non-linear resonance bending in a channel model, as discussed in [Benzi et al., 1986], or in a spherical model with reason-

ably realistic orography which permits a meandering jet to form. Such a model would ideally still be simple enough to be understood in terms of the dynamical building blocks we have isolated in CdV, but complex enough to eliminate its worst flaws.

At the other end of the model hierarchy, we believe the introduction of the geopotential-jet framework to provide a valuable addition to the study of blocking and jet regimes in complex models. Regime analysis has certainly already been used very successfully in climate contexts, such as in [Fabiano et al., 2020, Fabiano et al., 2021] but the driving motivation behind regimes has often come from the context of numerical weather prediction. That is still the domain in which it finds most use, in hybrid statistical-dynamical modelling of surface weather, analyses of flow-dependent predictability, and analyses of flow-dependent forecast model bias.

Arguably, one reason climate science as a whole does not make wider use of regime perspectives is due to the various technical issues involved. The challenge of data-driven methods, such as clustering, is the tendency for results to be very dataset dependent. This is a particular issue when using many different models, and very long integrations as is common in a climate context. By showing a way to identify reproducible, stable, and well-captured regime patterns, regime approaches become more feasible to use as a routine analysis tool. Of course, not every application requires a regime approach, which adds complexity and reduces sample sizes when conditioning statistics on regime state, but for the analysis of non-linear flows there are many benefits to a methodology which intrinsically captures flow-dependence.

In summary, this thesis has explored many aspects of regimes and has obtained a range of interesting results – although not always the ones we were originally seeking. Just as when a regime system is forced along one axis it might shift along another, we have found answers to questions we had not originally intended to ask. The work that dominates Chapters 2 and 3 of this thesis, for example, began out of simple frustration with the difficulty of obtaining stable results, before developing into a new approach to under-

standing regime dynamics. Of course, we have raised many more questions than we have provided answers, and the possibilities for future extensions, generalisations and deeper investigations of the results presented here are extensive. However, in this author's opinion, the most substantial open question, is not one of those raised in this thesis for the first time, but is the one that has remained stubbornly unanswered for at least the last 40 years:

Why does the atmosphere possess regimes at all?

Bibliography

[Araújo and Viana, 2009] Araújo, V. and Viana, M. (2009). Hyperbolic Dynamical Systems. In *Encyclopedia of Complexity and Systems Science*, pages 4723–4737. Springer New York.

[Athanasiadis et al., 2020] Athanasiadis, P. J., Yeager, S., Kwon, Y.-O., Bellucci, A., Smith, D. W., and Tibaldi, S. (2020). Decadal predictability of North Atlantic blocking and the NAO. *npj Climate and Atmospheric Science 2020 3:1*, 3(1):1–10.

[Bacer et al., 2021] Bacer, S., Jomaa, F., Beaumet, J., Gallée, H., Le Bouëdec, E., Ménégoz, M., and Staquet, C. (2021). Impact of climate change on wintertime European atmospheric blocking.

[Barnes and Polvani, 2013] Barnes, E. A. and Polvani, L. (2013). Response of the Midlatitude Jets, and of Their Variability, to Increased Greenhouse Gases in the CMIP5 Models. *Journal of Climate*, 26(18):7117–7135.

[Baur, 1951] Baur, F. (1951). *Compendium of meteorology*.

[Baur et al., 1944] Baur, F., Hess, P., and Nagel, H. (1944). Kalendar der Grosswetterlagen Europas 1881-1939. *Bad Homburg (DWD)*.

[Beerli and Grams, 2019] Beerli, R. and Grams, C. M. (2019). Stratospheric modulation of the large-scale circulation in the Atlantic-European region and its implications for surface weather events. *Quarterly Journal of the Royal Meteorological Society*, 145(725):3732–3750.

[Benzi et al., 1986] Benzi, R., Malguzzi, P., Speranza, A., and Sutera, A. (1986). The statistical properties of general atmospheric circulation: Observational evidence and a minimal theory of bimodality. *Quarterly Journal of the Royal Meteorological Society*, 112(473):661–674.

[Berckmans et al., 2013] Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L., and Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, 14(1):34–40.

[Berner et al., 2017] Berner, J., Achatz, U., Batté, L., Bengtsson, L., De La Cámara, A., Christensen, H. M., Colangeli, M., Coleman, D. R., Crommelin, D., Dolaptchiev, S. I., Franzke, C. L., Friederichs, P., Imkeller, P., Järvinen, H., Juricke, S., Kitsios, V., Lott, F., Lucarini, V., Mahajaajaajan, S., Palmer, T. N., Penland, C., Sakradzijaja, M., Von Storch, J. S., Weisheimer, A., Weniger, M., Williams, P. D., and Yano, J. I. (2017). Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, 98(3):565–587.

[Berner et al., 2012] Berner, J., Jung, T., and Palmer, T. N. (2012). Systematic Model Error: The Impact of Increased Horizontal Resolution versus Improved Stochastic and Deterministic Parameterizations. *Journal of Climate*, 25(14):4946–4962.

[Bishop, 2006] Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Springer.

[Blackport and Screen, 2019] Blackport, R. and Screen, J. A. (2019). Influence of Arctic Sea Ice Loss in Autumn Compared to That in Winter on the Atmospheric Circulation. *Geophysical Research Letters*, 46(4):2213–2221.

[Bowie and Weightman, 1914] Bowie, E. H. and Weightman, R. H. (1914). Types of storms of the United States and their average movements. *US Government Printing Office*.

[Branstator, 2002] Branstator, G. (2002). Circumglobal Teleconnections, the Jet Stream Waveguide, and the North Atlantic Oscillation. Technical Report 14.

[Brayshaw et al., 2009] Brayshaw, D. J., Hoskins, B., and Blackburn, M. (2009). The Basic Ingredients of the North Atlantic Storm Track. Part I: Land-Sea Contrast and Orography. *Journal of the Atmospheric Sciences*, 66(9):2539–2558.

[Brunner et al., 2018] Brunner, L., Schaller, N., Anstey, J., Sillmann, J., and Steiner, A. K. (2018). Dependence of Present and Future European Temperature Extremes on the Location of Atmospheric Blocking. *Geophysical Research Letters*, 45(12):6311–6320.

[Buizza et al., 1999] Buizza, R., Miller, M., and Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125(560):2887–2908.

[Burgess et al., 2020] Burgess, M. G., Ritchie, J., Shapland, J., and Pielke, R. (2020). IPCC baseline scenarios have over-projected CO2 emissions and economic growth. *Environmental Research Letters*, 16(1):014016.

[Butler et al., 2017] Butler, A. H., Sjoberg, J. P., Seidel, D. J., and Rosenlof, K. H. (2017). A sudden stratospheric warming compendium. *Earth System Science Data*, 9(1):63–76.

[Cassou, 2008] Cassou, C. (2008). Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. *Nature*, 455(7212):523–527.

[Cehelsky and Tung, 1987] Cehelsky, P. and Tung, K. K. (1987). Theories of multiple equilibria and weather regimes - a critical reexamination. Part II: baroclinic two-layer models. *Journal of the Atmospheric Sciences*.

[Champneys and Kirk, 2004] Champneys, A. R. and Kirk, V. (2004). The entwined wiggling of homoclinic curves emerging from saddle-node/Hopf instabilities. *Physica D: Nonlinear Phenomena*, 195(1-2):77–105.

[Charlton-Perez et al., 2018] Charlton-Perez, A. J., Ferranti, L., and Lee, R. W. (2018). The influence of the stratospheric state on North Atlantic weather regimes. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1140–1151.

[Charney and DeVore, 1979] Charney, J. G. and DeVore, J. G. (1979). Multiple Flow Equilibria in the Atmosphere and Blocking. *Journal of the Atmospheric Sciences*,

36(7):1205–1216.

[Charney and Straus, 1980] Charney, J. G. and Straus, D. M. (1980). Form-drag instability, multiple equilibria and propagating planetary waves in baroclinic, orographically forced, planetary wave systems. *Journal of the Atmospheric Sciences*.

[Christensen et al., 2015a] Christensen, H. M., Moroz, I. M., and Palmer, T. N. (2015a). Simulating weather regimes: impact of stochastic and perturbed parameter schemes in a simple atmospheric model. *Climate Dynamics*, 44(7-8):2195–2214.

[Christensen et al., 2015b] Christensen, H. M., Moroz, I. M., and Palmer, T. N. (2015b). Stochastic and perturbed parameter representations of model uncertainty in convection parameterization. *Journal of the Atmospheric Sciences*, 72(6):2525–2544.

[Christiansen, 2007] Christiansen, B. (2007). Atmospheric circulation regimes: Can cluster analysis provide the number? *Journal of Climate*, 20(10):2229–2250.

[Compo et al., 2011] Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, 137(654):1–28.

[Coppola et al., 2021] Coppola, E., Nogherotto, R., Ciarlo', J. M., Giorgi, F., van Meijgaard, E., Kadygrov, N., Iles, C., Corre, L., Sandstad, M., Somot, S., Nabat, P., Vautard, R., Levavasseur, G., Schwingshackl, C., Sillmann, J., Kjellström, E., Nikulin, G., Aalbers, E., Lenderink, G., Christensen, O. B., Boberg, F., Sørland, S. L., Demory, M.-E., Bülow, K., Teichmann, C., Warrach-Sagi, K., and Wulfmeyer, V. (2021). Assessment of the European Climate Projections as Simulated by the Large EURO-

CORDEX Regional and Global Climate Model Ensemble. *Journal of Geophysical Research: Atmospheres*, 126(4).

[Crommelin, 2002] Crommelin, D. T. (2002). Homoclinic Dynamics: A Scenario for Atmospheric Ultralow-Frequency Variability. *Journal of the Atmospheric Sciences*, 59(9):1533–1549.

[Crommelin et al., 2004] Crommelin, D. T., Opsteegh, J. D., and Verhulst, F. (2004). A Mechanism for Atmospheric Regime Behavior. *Journal of the Atmospheric Sciences*.

[D'Andrea, 2002] D'Andrea, F. (2002). Extratropical low-frequency variability as a low-dimensional problem. II: Stationarity and stability of large-scale equilibria. *Quarterly Journal of the Royal Meteorological Society*, 128(582):1059–1073.

[Davini et al., 2012] Davini, P., Cagnazzo, C., Gualdi, S., and Navarra, A. (2012). Bidimensional diagnostics, variability, and trends of northern hemisphere blocking. *Journal of Climate*, 25(19):6496–6509.

[Davini et al., 2017a] Davini, P., Corti, S., D'Andrea, F., Rivière, G., and von Hardenberg, J. (2017a). Improved Winter European Atmospheric Blocking Frequencies in High-Resolution Global Climate Simulations. *Journal of Advances in Modeling Earth Systems*, 9(7):2615–2634.

[Davini and D'Andrea, 2016] Davini, P. and D'Andrea, F. (2016). Northern Hemisphere atmospheric blocking representation in global climate models: Twenty years of improvements? *Journal of Climate*, 29(24):8823–8840.

[Davini and D'Andrea, 2020] Davini, P. and D'Andrea, F. (2020). From CMIP3 to CMIP6: Northern hemisphere atmospheric blocking simulation in present and future climate. *Journal of Climate*, 33(23):10021–10038.

[Davini et al., 2017b] Davini, P., Von Hardenberg, J., Corti, S., Christensen, H. M., Juricke, S., Subramanian, A., Watson, P. A. G., Weisheimer, A., and Palmer, T. N. (2017b). Climate SPHINX: Evaluating the impact of resolution and stochastic physics

parameterisations in the EC-Earth global climate model. *Geoscientific Model Development*, 10(3):1383–1402.

[Davini et al., 2021] Davini, P., Weisheimer, A., Balmaseda, M., Johnson, S. J., Molteni, F., Roberts, C. D., Senan, R., and Stockdale, T. N. (2021). The representation of winter Northern Hemisphere atmospheric blocking in ECMWF seasonal prediction systems. *Quarterly Journal of the Royal Meteorological Society*, 147(735):1344–1363.

[Dawson and Palmer, 2015] Dawson, A. and Palmer, T. N. (2015). Simulating weather regimes: impact of model resolution and stochastic parameterization. *Climate Dynamics*, 44(7-8):2177–2193.

[Dawson et al., 2012] Dawson, A., Palmer, T. N., and Corti, S. (2012). Simulating regime structures in weather and climate prediction models. *Geophysical Research Letters*, 39(21):1–6.

[Dawson et al., 1994] Dawson, S., Grebogi, C., Sauer, T., and Yorke, J. A. (1994). Obstructions to shadowing when a lyapunov exponent fluctuates about Zero. *Physical Review Letters*, 73(14):1927–1930.

[De Swart, 1988] De Swart, H. E. (1988). Low-order spectral models of the atmospheric circulation: A survey. *Acta Applicandae Mathematicae*, 11(1):49–96.

[Delworth and Zeng, 2016] Delworth, T. L. and Zeng, F. (2016). The impact of the North Atlantic Oscillation on climate through its influence on the Atlantic meridional overturning circulation. *Journal of Climate*, 29(3):941–962.

[Delworth et al., 2017] Delworth, T. L., Zeng, F., Zhang, L., Zhang, R., Vecchia, G. A., and Yang, X. (2017). The central role of ocean dynamics in connecting the North Atlantic oscillation to the extratropical component of the Atlantic multidecadal oscillation. *Journal of Climate*, 30(10):3789–3805.

[Dole and Gordon, 1983] Dole, R. M. and Gordon, N. D. (1983). Persistent anomalies of the extratropical Northern Hemisphere wintertime circulation: geographical distri-

bution and regional persistence characteristics. *Monthly Weather Review*.

[Domeisen et al., 2020] Domeisen, D. I. V., Grams, C. M., and Papritz, L. (2020). The role of North Atlantic-European weather regimes in the surface impact of sudden stratospheric warming events. *Weather and Climate Dynamics*, 1(2):373–388.

[Dorrington et al., 2021] Dorrington, J., Strommen, K., and Fabiano, F. (2021). Explaining the representation of historical Euro-Atlantic variability in CMIP models (In preparation).

[Dorrington and Strommen, 2020] Dorrington, J. and Strommen, K. J. (2020). Jet Speed Variability Obscures Euro-Atlantic Regime Structure. *Geophysical Research Letters*, 47(15).

[Drouard et al., 2021] Drouard, M., Woollings, T., Sexton, D. M., and McSweeney, C. F. (2021). Dynamical Differences Between Short and Long Blocks in the Northern Hemisphere. *Journal of Geophysical Research: Atmospheres*, 126(10):e2020JD034082.

[Düben et al., 2014] Düben, P. D., McNamara, H., and Palmer, T. N. (2014). The use of imprecise processing to improve accuracy in weather & climate prediction. *Journal of Computational Physics*, 271:2–18.

[Dunstone et al., 2016] Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N., Andrews, M., and Knight, J. (2016). Skilful predictions of the winter North Atlantic Oscillation one year ahead. *Nature Geoscience 2016 9:11*, 9(11):809–814.

[Eade et al., 2014] Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L., and Robinson, N. (2014). Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophysical Research Letters*, 41(15):5620–5628.

[ECMWF, 2010] ECMWF (2010). *IFS Documentation CY36R1 - Part IV: Physical Processes*. IFS Documentation. ECMWF.

[Eichelberger and Hartmann, 2007] Eichelberger, S. J. and Hartmann, D. L. (2007). Zonal Jet Structure and the Leading Mode of Variability. *Journal of Climate*, 20(20):5149–5163.

[Eyring et al., 2016] Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958.

[Fabiano et al., 2020] Fabiano, F., Christensen, H. M., Strommen, K., Athanasiadis, P., Baker, A., Schiemann, R., and Corti, S. (2020). Euro-Atlantic weather Regimes in the PRIMAVERA coupled climate simulations: impact of resolution and mean state biases on model performance. 54:5031–5048.

[Fabiano et al., 2021] Fabiano, F., Meccia, V. L., Davini, P., Ghinassi, P., and Corti, S. (2021). A regime view of future atmospheric circulation changes in northern mid-latitudes. *Weather and Climate Dynamics*, 2(1):163–180.

[Falkena et al., 2020] Falkena, S. K., de Wiljes, J., Weisheimer, A., and Shepherd, T. G. (2020). Revisiting the identification of wintertime atmospheric circulation regimes in the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, 146(731):2801–2814.

[Faranda et al., 2016] Faranda, D., Masato, G., Moloney, N., Sato, Y., Daviaud, F., Dubrulle, B., and Yiou, P. (2016). The switching between zonal and blocked mid-latitude atmospheric circulation: a dynamical system perspective. *Climate Dynamics*, 47(5-6):1587–1599.

[Fereday, 2017] Fereday, D. (2017). How Persistent Are North Atlantic-European Sector Weather Regimes? *Journal of Climate*, 30(7):2381–2394.

[Fereday et al., 2008] Fereday, D. R., Knight, J. R., Scaife, A. A., Folland, C. K., and Philipp, A. (2008). Cluster Analysis of North Atlantic-European Circulation Types and Links with Tropical Pacific Sea Surface Temperatures. *Journal of Climate*, 21(15):3687–3703.

[Ferranti et al., 2015a] Ferranti, L., Corti, S., and Janousek, M. (2015a). Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, 141(688):916–924.

[Ferranti et al., 2015b] Ferranti, L., Corti, S., and Janousek, M. (2015b). Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, 141(688):916–924.

[Forzieri et al., 2016] Forzieri, G., Feyen, L., Russo, S., Vousdoukas, M., Alfieri, L., Outten, S., Migliavacca, M., Bianchi, A., Rojas, R., and Cid, A. (2016). Multi-hazard assessment in Europe under climate change. *Climatic Change*, 137(1-2):105–119.

[Frame et al., 2013] Frame, T. H. A., Methven, J., Gray, S. L., and Ambaum, M. H. P. (2013). Flow-dependent predictability of the North Atlantic jet. *Geophysical Research Letters*, 40(10):2411–2416.

[Franzke et al., 2011] Franzke, C., Woollings, T., and Martius, O. (2011). Persistent Circulation Regimes and Preferred Regime Transitions in the North Atlantic. *Journal of the Atmospheric Sciences*, 68(12):2809–2825.

[Garrido-Perez et al., 2020] Garrido-Perez, J. M., Ordóñez, C., Barriopedro, D., García-Herrera, R., and Paredes, D. (2020). Impact of weather regimes on wind power variability in western Europe. *Applied Energy*.

[Gerber and Vallis, 2009] Gerber, E. P. and Vallis, G. K. (2009). On the Zonal Structure of the North Atlantic Oscillation and Annular Modes. *Journal of the Atmospheric Sciences*, 66(2):332–352.

[Ghil, 2015] Ghil, M. (2015). *Climate Change: Multidecadal and Beyond*, volume 6. World Scientific Publ. Co./Imperial College Press.

[Ghil and Childress, 1987] Ghil, M. and Childress, S. (1987). Persistent Anomalies, Blocking and Predictability. pages 125–201.

[Gold, 1920] Gold, E. (1920). *Aids to forecasting: types of pressure distribution*. Gt. Brit. Meteorological office. Geophysical memoirs, no. 16. Meteorological office [Printed by Neill & co., Edinburgh], London.

[Gottwald and Gugole, 2020] Gottwald, G. A. and Gugole, F. (2020). Detecting Regime Transitions in Time Series Using Dynamic Mode Decomposition. *Journal of Statistical Physics*, 179:1028–1045.

[Grams et al., 2017a] Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H. (2017a). Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change*, 7(8):557–562.

[Grams et al., 2017b] Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H. (2017b). Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change*, 7(8):557–562.

[Haines and Marshall, 1987] Haines, K. and Marshall, J. (1987). EddyâForced Coherent Structures As A Prototype of Atmospheric Blocking. *Quarterly Journal of the Royal Meteorological Society*, 113(476):681–704.

[Hanna et al., 2016] Hanna, E., Cropper, T. E., Hall, R. J., and Cappelen, J. (2016). Greenland Blocking Index 1851-2015: a regional climate change signal. *International Journal of Climatology*, 36(15):4847–4861.

[Hannachi and Hannachi, 2007] Hannachi, A. and Hannachi, A. (2007). Tropospheric Planetary Wave Dynamics and Mixture Modeling: Two Preferred Regimes and a Regime Shift. *Journal of the Atmospheric Sciences*, 64(10):3521–3541.

[Hannachi et al., 2012] Hannachi, A., Woollings, T., and Fraedrich, K. (2012). The North Atlantic jet stream: A look at preferred positions, paths and transitions. *Quarterly Journal of the Royal Meteorological Society*, 138(665):862–877.

[Harvey et al., 2020] Harvey, B. J., Cook, P., Shaffrey, L. C., and Schiemann, R. (2020). The Response of the Northern Hemisphere Storm Tracks and Jet Streams to Climate Change in the CMIP3, CMIP5, and CMIP6 Climate Models. *Journal of Geophysical Research: Atmospheres*, 125(23).

[Hazeleger et al., 2012] Hazeleger, W., Wang, X., Severijns, C., Stefanescu, S., Bintanja, R., Sterl, A., Wyser, K., Semmler, T., Yang, S., van den Hurk, B., van Noije, T., van der Linden, E., and van der Wiel, K. (2012). EC-Earth V2.2: Description and validation of a new seamless earth system prediction model. *Climate Dynamics*, 39(11):2611–2629.

[Hersbach et al., 2020] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

[Hewitson and Crane, 2002] Hewitson, B. C. and Crane, R. G. (2002). Self-organizing maps, Application to synoptic climatology. *Climate Research*, 22:13–26.

[Hoskins and Ambrizzi, 1993] Hoskins, B. J. and Ambrizzi, T. (1993). Rossby Wave Propagation on a Realistic Longitudinally Varying Flow. *Journal of the Atmospheric Sciences*, 50(12):1661–1671.

[Itoh and Kimoto, 1996] Itoh, H. and Kimoto, M. (1996). Multiple Attractors and

Chaotic Itinerancy in a Quasigeostrophic Model with Realistic Topography: Implications for Weather Regimes and Low-Frequency Variability. *Journal of the Atmospheric Sciences*, 53(15):2217–2231.

[Itoh and Kimoto, 1997] Itoh, H. and Kimoto, M. (1997). Chaotic itinerancy with preferred transition routes appearing in an atmospheric model. *Physica D: Nonlinear Phenomena*, 109(3-4):274–292.

[Izhikevich, 2006] Izhikevich, E. M. (2006). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. MIT Press.

[Jiménez-Esteve and Domeisen, 2018] Jiménez-Esteve, B. and Domeisen, D. I. (2018). The tropospheric pathway of the ENSO-North Atlantic teleconnection. *Journal of Climate*, 31(11):4563–4584.

[Kallen, 1981] Kallen, E. (1981). The Nonlinear Effects of Orographic and Momentum Forcing in a Low-Order, Barotropic Model. *Journal of the Atmospheric Sciences*, pages 2150–2163.

[Kallenberg, 2017] Kallenberg, O. (2017). *Random Measures, Theory and Applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, Cham.

[Kim et al., 2011] Kim, W., Yeh, S. W., Kim, J. H., Kug, J. S., and Kwon, M. (2011). The unique 2009-2010 El Niño event: A fast phase transition of warm pool El Niño to la Niña. *Geophysical Research Letters*, 38(15).

[Kimoto and Ghil, 1993] Kimoto, M. and Ghil, M. (1993). Multiple Flow Regimes in the Northern Hemisphere Winter. Part I: Methodology and Hemispheric Regimes. *Journal of the Atmospheric Sciences*, 50(16):2625–2644.

[Kimoto et al., 1993] Kimoto, M., Ghil, M., Kimoto, M., and Ghil, M. (1993). Multiple Flow Regimes in the Northern Hemisphere Winter. Part II: Sectorial Regimes and Preferred Transitions. *Journal of the Atmospheric Sciences*, 50(16):2645–2673.

[King et al., 2018] King, M. P., Herceg-Bulić, I., Bladé, I., García-Serrano, J., Keenlyside, N., Kucharski, F., Li, C., and Sobolowski, S. (2018). Importance of late fall ENSO teleconnection in the euro-atlantic sector. *Bulletin of the American Meteorological Society*, 99(7):1337–1343.

[Knutti et al., 2013] Knutti, R., Masson, D., and Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, 40(6):1194–1199.

[Kondrashov et al., 2004] Kondrashov, D., Ide, K., and Ghil, M. (2004). Weather Regimes and Preferred Transition Paths in a Three-Level Quasigeostrophic Model. *Journal of the Atmospheric Sciences*, 61(5):568–587.

[Kornhuber et al., 2019] Kornhuber, K., Coumou, D., Vogel, E., Lesk, C., Donges, J. F., Lehmann, J., and Horton, R. M. (2019). Amplified Rossby waves enhance risk of concurrent heatwaves in major breadbasket regions. *Nature Climate Change 2019 10:1*, 10(1):48–53.

[Kornhuber et al., 2017] Kornhuber, K., Petoukhov, V., Karoly, D., Petri, S., Rahmstorf, S., and Coumou, D. (2017). Summertime planetary wave resonance in the Northern and Southern hemispheres. *Journal of Climate*, 30(16):6133–6150.

[Kwasniok, 2007] Kwasniok, F. (2007). Reduced atmospheric models using dynamically motivated basis functions. *Journal of the Atmospheric Sciences*, 64(10):3452–3474.

[Kwasniok, 2014] Kwasniok, F. (2014). Enhanced regime predictability in atmospheric low-order models due to stochastic forcing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2018):20130286–20130286.

[Laloyaux et al., 2018] Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J. R., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D., Haimberger, L., Hersbach, H.,

Kosaka, Y., Martin, M., Poli, P., Rayner, N., Rustemeier, E., and Schepers, D. (2018). CERA-20C: A Coupled Reanalysis of the Twentieth Century. *Journal of Advances in Modeling Earth Systems*, 10(5):1172–1195.

[Lee et al., 2020] Lee, J. C. K., Lee, R. W., Woolnough, S. J., and Boxall, L. J. (2020). The links between the Madden-Julian Oscillation and European weather regimes. *Theoretical and Applied Climatology*, 141(1-2):567–586.

[Lee and Kim, 2003] Lee, S. and Kim, H. K. (2003). The dynamical relationship between subtropical and eddy-driven jets. *Journal of the Atmospheric Sciences*.

[Legras and Ghil, 1985] Legras, B. and Ghil, M. (1985). Persistent anomalies, blocking and variations in atmospheric predictability. *Journal of the Atmospheric Sciences*.

[Li and Wettstein, 2012] Li, C. and Wettstein, J. J. (2012). Thermally driven and eddy-driven jet variability in reanalysis. *Journal of Climate*, 25(5):1587–1596.

[Lorenz, 1963] Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Universality in Chaos, Second Edition*, 20(2):367–378.

[Lucarini and Gritsun, 2020] Lucarini, V. and Gritsun, A. (2020). A new mathematical framework for atmospheric blocking events. *Climate Dynamics*, 54(1-2):575–598.

[Lund, 1963] Lund, I. (1963). Map pattern classification by statistical methods. *J Appl Meteor*, 2:56–65.

[Madonna et al., 2021] Madonna, E., Battisti, D. S., Li, C., and White, R. H. (2021). Reconstructing winter climate anomalies in the Euro-Atlantic sector using circulation patterns. *Weather and Climate Dynamics*.

[Madonna et al., 2017] Madonna, E., Li, C., Grams, C. M., and Woollings, T. (2017). The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Quarterly Journal of the Royal Meteorological Society*, 143(708):2960–2972.

[Maiocchi, C. C. and Lucarini, 2021] Maiocchi, C. C. and Lucarini, V. (2021). Unstable Periodic Orbits Sampling and Its Applications to Climate Models. In *EGU General Assembly 2021*.

[Malguzzi and Speranza, 1981] Malguzzi, P. and Speranza, A. (1981). Local multiple equilibria and regional atmospheric blocking. *Journal of the Atmospheric Sciences*, 38(9):1939–1948.

[Masato et al., 2013] Masato, G., Hoskins, B. J., and Woollings, T. (2013). Winter and Summer Northern Hemisphere Blocking in CMIP5 Models. *Journal of Climate*, 26(18):7044–7059.

[Masato et al., 2009] Masato, G., Hoskins, B. J., and Woollings, T. J. (2009). Can the Frequency of Blocking Be Described by a Red Noise Process? *Journal of the Atmospheric Sciences*, 66(7):2143–2149.

[Masato et al., 2012] Masato, G., Hoskins, B. J., and Woollings, T. J. (2012). Wave-breaking characteristics of midlatitude blocking. *Quarterly Journal of the Royal Meteorological Society Q. J. R. Meteorol. Soc*, 138:1285–1296.

[Masato et al., 2014] Masato, G., Woollings, T., and Hoskins, B. J. (2014). Structure and impact of atmospheric blocking over the Euro-Atlantic region in present-day and future simulations. *Geophysical Research Letters*, 41(3):1051–1058.

[Matsueda and Palmer, 2018] Matsueda, M. and Palmer, T. N. (2018). Estimates of flow-dependent predictability of wintertime Euro-Atlantic weather regimes in medium-range forecasts. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1012–1027.

[Meccia et al., 2020] Meccia, V. L., Fabiano, F., Davini, P., and Corti, S. (2020). Stochastic Parameterizations and the Climate Response to External Forcing: An Experiment With EC-Earth. *Geophysical Research Letters*, 47(3):e2019GL085951.

[Meurer et al., 2017] Meurer, A., Smith, C. P., Paprocki, M., Čertík, O., Kirpichev, S. B., Rocklin, M., Kumar, A. T., Ivanov, S., Moore, J. K., Singh, S., Rathnayake, T., Vig, S., Granger, B. E., Muller, R. P., Bonazzi, F., Gupta, H., Vats, S., Johansson, F., Pedregosa, F., Curry, M. J., Terrel, A. R., Roučka, Š., Saboo, A., Fernando, I., Kulal, S., Cimrman, R., and Scopatz, A. (2017). SymPy: Symbolic computing in python. *PeerJ Computer Science*, 2017(1):e103.

[Mezzina et al., 2020] Mezzina, B., García-Serrano, J., Bladé, I., and Kucharski, F. (2020). Dynamics of the ENSO teleconnection and NAO variability in the North Atlantic-European late winter. *Journal of Climate*, 33(3):907–923.

[Michelangeli et al., 1995] Michelangeli, P.-A., Vautard, R., Legras, B., Michelangeli, P.-A., Vautard, R., and Legras, B. (1995). Weather Regimes: Recurrence and Quasi Stationarity. *Journal of the Atmospheric Sciences*, 52(8):1237–1256.

[Mo, 1986] Mo, K. C. (1986). Quasi-stationary states in the Southern Hemisphere. *Monthly Weather Review*.

[Mo and Ghil, 1988] Mo, K. C. and Ghil, M. (1988). Statistics and Dynamics of Persistent Anomalies. *Journal of the Atmospheric Sciences*, 44(5):877–902.

[Molteni and Corti, 1998] Molteni, F. and Corti, S. (1998). Long-term fluctuations in the statistical properties of low-frequency variability: Dynamical origin and predictability. *Quarterly Journal of the Royal Meteorological Society*, 124(546):495–526.

[Molteni et al., 1990] Molteni, F., Tibaldi, S., and Palmer, T. N. (1990). Regimes in the wintertime circulation over northern extratropics. I: Observational evidence. *Quarterly Journal of the Royal Meteorological Society*, 116(491):31–67.

[Monahan et al., 2001] Monahan, A. H., Pandolfo, L., and Fyfe, J. C. (2001). The preferred structure of variability of the Northern Hemisphere atmospheric circulation. *Geophysical Research Letters*, 28(6):1019–1022.

[Mullen, 1987] Mullen, S. (1987). Transient eddy forcing of blocking flows. *Journal of the Atmospheric Sciences*.

[Nakamura and Wallace, 1993] Nakamura, H. and Wallace, J. M. (1993). Synoptic Behavior of Baroclinic Eddies during the Blocking Onset. *Monthly Weather Review*.

[O'Reilly et al., 2016] O'Reilly, C. H., Minobe, S., and Kuwano-Yoshida, A. (2016). The influence of the Gulf Stream on wintertime European blocking. *Climate Dynamics*, 47(5-6):1545–1567.

[O'Reilly et al., 2019] O'Reilly, C. H., Woollings, T., Zanna, L., and Weisheimer, A. (2019). An Interdecadal Shift of the Extratropical Teleconnection From the Tropical Pacific During Boreal Summer. *Geophysical Research Letters*, 46(22):13379–13388.

[Oudar et al., 2020] Oudar, T., Cattiaux, J., and Douville, H. (2020). Drivers of the Northern Extratropical Eddy-Driven Jet Change in CMIP5 and CMIP6 Models. *Geophysical Research Letters*, 47(8).

[P Cvitanovic, 1991] P Cvitanovic, B. E. (1991). Periodic orbit expansions for classical smooth flows. *J Phys A*, 24:L237.

[Palmer, 1993] Palmer, T. N. (1993). A nonlinear dynamical perspective on climate change. *Weather*, 48(10):314–326.

[Palmer, 1999] Palmer, T. N. (1999). A nonlinear dynamical perspective on climate prediction. *Journal of Climate*, 12(2):575–591.

[Palmer and Weisheimer, 2011] Palmer, T. N. and Weisheimer, A. (2011). Diagnosing the causes of bias in climate models - why is it so hard? *Geophysical and Astrophysical Fluid Dynamics*, 105(2-3):351–365.

[Parker et al., 2019] Parker, T., Woollings, T., Weisheimer, A., O'Reilly, C., Baker, L., and Shaffrey, L. (2019). Seasonal Predictability of the Winter North Atlantic Oscil-

lation From a Jet Stream Perspective. *Geophysical Research Letters*, 46(16):10159–10167.

[Peings et al., 2017] Peings, Y., Cattiaux, J., Vavrus, S., and Magnusdottir, G. (2017). Late Twenty-First-Century Changes in the Midlatitude Atmospheric Circulation in the CESM Large Ensemble. *Journal of Climate*, 30(15):5943–5960.

[Pithan et al., 2016] Pithan, F., Shepherd, T. G., Zappa, G., and Sandu, I. (2016). Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag. *Geophysical Research Letters*, 43(13):7231–7240.

[Poli et al., 2016] Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., Fisher, M., Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M. (2016). ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. *Journal of Climate*, 29(11):4083–4097.

[Pusuluri and Shilnikov, 2018] Pusuluri, K. and Shilnikov, A. (2018). Homoclinic chaos and its organization in a nonlinear optics model. *Physical Review E*, 98(4):040202.

[Reinhold and Pierrehumbert, 1982] Reinhold, B. B. and Pierrehumbert, R. T. (1982). Dynamics of Weather Regimes: Quasi-Stationary Waves and Blocking. *Monthly Weather Review*, 110(9):1105–1145.

[Riddle et al., 2013] Riddle, E. E., Stoner, M. B., Johnson, N. C., L'Heureux, M. L., Collins, D. C., and Feldstein, S. B. (2013). The impact of the MJO on clusters of wintertime circulation anomalies over the North American region. *Climate Dynamics*, 40(7-8):1749–1766.

[Robin et al., 2017] Robin, Y., Yiou, P., and Naveau, P. (2017). Detecting changes in forced climate attractors with Wasserstein distance. *Nonlinear Processes in Geo-*

physics, 24(3):393–405.

[Rodríguez-Fonseca et al., 2016] Rodríguez-Fonseca, B., Suárez-Moreno, R., Ayarzagüena, B., López-Parages, J., Gómara, I., Villamayor, J., Mohino, E., Losada, T., and Castaño-Tierno, A. (2016). A Review of ENSO Influence on the North Atlantic. A Non-Stationary Signal. *Atmosphere*, 7(7):87.

[Rossby, 1940] Rossby, C. G. (1940). Planetary flow patterns in the atmosphere. *Quart. J. Roy. Meteor. Soc. (suppl),*, (66):68–87.

[Rousi et al., 2021] Rousi, E., Selten, F., Rahmstorf, S., and Coumou, D. (2021). Changes in North Atlantic Atmospheric Circulation in a Warmer Climate Favor Winter Flooding and Summer Drought over Europe. *Journal of Climate*, 34(6):2277–2295.

[Scaife et al., 2019] Scaife, A. A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., MacLachlan, C., Martin, N., Nie, Y., Ren, H.-L., Roberts, M., Robinson, W., Smith, D., and Vidale, P. L. (2019). Does increased atmospheric resolution improve seasonal climate predictions? *Atmospheric Science Letters*, 20(8):e922.

[Scaife and Smith, 2018] Scaife, A. A. and Smith, D. (2018). A signal-to-noise paradox in climate science. *npj Climate and Atmospheric Science*, 1(1):28.

[Schiemann et al., 2020] Schiemann, R., Athanasiadis, P., Barriopedro, D., Doblas-Reyes, F., Lohmann, K., Roberts, M. J., Sein, D. V., Roberts, C. D., Terray, L., and Vidale, P. L. (2020). Northern Hemisphere blocking simulation in current climate models: evaluating progress from the Climate Model Intercomparison Project Phase 5 to 6 and sensitivity to resolution. *Weather and Climate Dynamics*, 1(1):277–292.

[Schiemann et al., 2017] Schiemann, R., Demory, M.-E., Shaffrey, L. C., Strachan, J., Vidale, P. L., Mizielinski, M. S., Roberts, M. J., Matsueda, M., Wehner, M. F., and Jung, T. (2017). The Resolution Sensitivity of Northern Hemisphere Blocking in Four 25-km Atmospheric Global Circulation Models. *Journal of Climate*, 30(1):337–358.

[Selten and Branstator, 2004] Selten, F. M. and Branstator, G. (2004). Preferred Regime Transition Routes and Evidence for an Unstable Periodic Orbit in a Baroclinic Model. *Journal of the Atmospheric Sciences*, 61(18):2267–2282.

[Shen et al., 2021] Shen, B. W., Pielke, R. A., Zeng, X., Baik, J. J., Faghih-Naini, S., Cui, J., and Atlas, R. (2021). Is weather chaotic? Coexistence of chaos and order within a generalized lorenz model. *Bulletin of the American Meteorological Society*, 102(1):E148–E158.

[Shepherd, 2014] Shepherd, T. G. (2014). Atmospheric circulation as a source of uncertainty in climate change projections. *Nature Geoscience*, 7(10):703–708.

[Shilnikov et al., 1995] Shilnikov, A., Nicolis, G., and Nicolis, C. (1995). Bifurcation and predictability analysis of a low-order atmopsheric circulation model. *International Journal of Bifurcation and Chaos*, 05(06):1701–1711.

[Shukla and Mo, 1983] Shukla, J. and Mo, K. C. (1983). Seasonal and geographical variation of blocking. *Monthly Weather Review*.

[Shutts, 1983] Shutts, G. J. (1983). The propagation of eddies in diffluent jetstreams: Eddy vorticity forcing of 'blocking' flow fields. *Quarterly Journal of the Royal Meteorological Society*, 109(462):737–761.

[Slivinski et al., 2021] Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan, P., Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster, D. C., Kennedy, J. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H., Lee, T. C., Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C., Slonosky, V. C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P. (2021). An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3. *Journal of Climate*, 34(4):1417–1438.

[Smith et al., 2019] Smith, D. M., Eade, R., Scaife, A. A., Caron, L. P., Danabasoglu, G., DelSole, T. M., Delworth, T., Doblas-Reyes, F. J., Dunstone, N. J., Hermanson, L., Kharin, V., Kimoto, M., Merryfield, W. J., Mochizuki, T., Müller, W. A., Pohlmann, H., Yeager, S., and Yang, X. (2019). Robust skill of decadal climate predictions. *NPJ Climate and Atmospheric Science*, 2(1).

[Smith et al., 2016] Smith, D. M., Scaife, A. A., Eade, R., and Knight, J. R. (2016). Seasonal to decadal prediction of the winter North Atlantic Oscillation: emerging capability and future prospects. *Quarterly Journal of the Royal Meteorological Society*, 142(695):611–617.

[Stephenson et al., 2004] Stephenson, D. B., Hannachi, A., and O'Neill, A. (2004). On the existence of multiple climate regimes. *Quarterly Journal of the Royal Meteorological Society*, 130(597):583–605.

[Straus et al., 2007] Straus, D. M., Corti, S., and Molteni, F. (2007). Circulation regimes: Chaotic variability versus SST-forced predictability. *Journal of Climate*, 20(10):2251–2272.

[Straus and Molteni, 2004] Straus, D. M. and Molteni, F. (2004). Circulation regimes and SST forcing: Results from large GCM ensembles. *Journal of Climate*.

[Strommen et al., 2021] Strommen, K., Chantry, M., Dorrington, J., and Otter, N. (2021). A topological perspective on regimes in dynamical systems.

[Strommen et al., 2019] Strommen, K., Mavilia, I., Corti, S., Matsueda, M., Davini, P., von Hardenberg, J., Vidale, P. L., and Mizuta, R. (2019). The Sensitivity of Euro-Atlantic Regimes to Model Horizontal Resolution. *Geophysical Research Letters*, 46(13):7810–7818.

[Strommen and Palmer, 2018] Strommen, K. and Palmer, T. N. (2018). Signal and noise in regime systems: A hypothesis on the predictability of the North Atlantic Oscillation. *Quarterly Journal of the Royal Meteorological Society*.

[Strommen and Palmer, 2019] Strommen, K. and Palmer, T. N. (2019). Signal and noise in regime systems: A hypothesis on the predictability of the North Atlantic Oscillation. *Quarterly Journal of the Royal Meteorological Society*, 145(718):147–163.

[Sura et al., 2005] Sura, P., Newman, M., Penland, C., and Sardeshmukh, P. (2005). Multiplicative Noise and Non-Gaussianity: A Paradigm for Atmospheric Regimes? *Journal of the Atmospheric Sciences*, 62(5):1391–1409.

[Taylor et al., 2012] Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93(4):485–498.

[Toth, 1993] Toth, Z. (1993). Preferred and unpreferred circulation types in the Northern Hemisphere wintertime phase space. *Journal of the Atmospheric Sciences*.

[Trevisan and Buzzi, 1980] Trevisan, A. and Buzzi, A. (1980). Stationary Response of Barotropic Weakly Nonlinear Rossby Waves to Quasi-Resonant Orographic Forcing. *Journal of the Atmospheric Sciences*.

[Van Der Wiel et al., 2019] Van Der Wiel, K., Bloomfield, H. C., Lee, R. W., Stoop, L. P., Blackport, R., Screen, J. A., and Selten, F. M. (2019). Environmental Research Letters The influence of weather regimes on European renewable energy production and demand Recent citations The influence of weather regimes on European renewable energy production and demand. *Environ. Res. Lett*, 14(9).

[Vautard and Legras, 1988] Vautard, R. and Legras, B. (1988). On the source of mid-latitude low-frequency variability. Part II: Nonlinear equilibration of weather regimes. *Journal of the Atmospheric Sciences*.

[Vissio et al., 2020] Vissio, G., Lembo, V., Lucarini, V., and Ghil, M. (2020). Evaluating the Performance of Climate Models Based on Wasserstein Distance. *Geophysical Research Letters*, 47(21).

[Waliser et al., 2012] Waliser, D. E., Moncrieff, M. W., Burridge, D., Fink, A. H., Gochis, D., Goswami, B. N., Guan, B., Harr, P., Heming, J., Hsu, H. H., Jakob, C., Janiga, M., Johnson, R., Jones, S., Knippertz, P., Marengo, J., Nguyen, H., Pope, M., Serra, Y., Thorncroft, C., Wheeler, M., Wood, R., and Yuter, S. (2012). The "year" of tropical convection (May 2008-April 2010): Climate variability and weather highlights.

[Watson et al., 2017] Watson, P. A., Berner, J., Corti, S., Davini, P., von Hardenberg, J., Sanchez, C., Weisheimer, A., and Palmer, T. N. (2017). The impact of stochastic physics on tropical rainfall variability in global climate models on daily to weekly time scales. *Journal of Geophysical Research*, 122(11):5738–5762.

[Weeks et al., 1997] Weeks, E. R., Tian, Y., Urbach, J. S., Ide, K., Swinney, H. L., and Ghil, M. (1997). Transitions Between Blocked and Zonal Flows in a Rotating Annulus with Topography. *Science*, 278(5343):1598–1601.

[Weisheimer et al., 2017] Weisheimer, A., Schaller, N., O'Reilly, C., MacLeod, D. A., and Palmer, T. (2017). Atmospheric seasonal forecasts of the twentieth century: multi-decadal variability in predictive skill of the winter North Atlantic Oscillation (NAO) and their potential value for extreme event attribution. *Quarterly Journal of the Royal Meteorological Society*, 143(703):917–926.

[White et al., 2019] White, R. H., Hilgenbrink, C., and Sheshadri, A. (2019). The Importance of Greenland in Setting the Northern Preferred Position of the North Atlantic Eddy-Driven Jet. *Geophysical Research Letters*, 46(23):14126–14134.

[Woollings et al., 2018] Woollings, T., Barnes, E., Hoskins, B., Kwon, Y. O., Lee, R. W., Li, C., Madonna, E., McGraw, M., Parker, T., Rodrigues, R., Spensberger, C., and Williams, K. (2018). Daily to decadal modulation of jet variability. *Journal of Climate*, 31(4):1297–1314.

[Woollings et al., 2014] Woollings, T., Czuchnicki, C., and Franzke, C. (2014). Twenti-

eth century North Atlantic jet variability. *Quarterly Journal of the Royal Meteorological Society*, 140(680):783–791.

[Woollings et al., 2010] Woollings, T., Hannachi, A., and Hoskins, B. (2010). Variability of the North Atlantic eddy-driven jet stream. *Quarterly Journal of the Royal Meteorological Society*, 136(649):856–868.

[Woollings et al., 2008] Woollings, T., Hoskins, B., Blackburn, M., and Berrisford, P. (2008). A New Rossby Wave-Breaking Interpretation of the North Atlantic Oscillation. *Journal of the Atmospheric Sciences*.

[Xinhua Cheng and Wallace, 1993] Xinhua Cheng and Wallace, J. M. (1993). Cluster analysis of the Northern Hemisphere wintertime 500-hPa height field: spatial patterns. *Journal of the Atmospheric Sciences*.

[Yamazaki and Itoh, 2013] Yamazaki, A. and Itoh, H. (2013). Vortex-vortex interactions for the maintenance of blocking. part I: The selective absorption mechanism and a case study. *Journal of the Atmospheric Sciences*, 70(3):725–742.

[Yang et al., 2019] Yang, C., Corti, S., Davini, P., von Hardenberg, J., and Christensen, H. M. (2019). The impact of stochastic physics on the El Niño Southern Oscillation in the EC-Earth coupled model. *Climate Dynamics*, pages 1–17.

[Yang et al., 1997] Yang, S., Reinhold, B., and Källé, E. (1997). Multiple Weather Regimes and Baroclinically Forced Spherical Resonance. *Journal of the Atmospheric Sciences*, 54(11):1397–1409.

[Zappa and Shepherd, 2017] Zappa, G. and Shepherd, T. G. (2017). Storylines of Atmospheric Circulation Change for European Regional Climate Impact Assessment. *Journal of Climate*, 30(16):6561–6577.

[Zhang et al., 2019] Zhang, W., Wang, Z., Stuecker, M. F., Turner, A. G., Jin, F. F., and Geng, X. (2019). Impact of ENSO longitudinal position on teleconnections to the NAO. *Climate Dynamics*, 52(1-2):257–274.

A | Appendix

A.1 Dynamical systems glossary

In this glossary we informally define some of the dynamical systems concepts used in Chapter 1 that less frequently appear in the atmospheric science literature. For more in-depth explanations the peer-reviewed digital reference Scholarpedia is recommended.

Bifurcation A qualitative change in the behaviour of a dynamical system due to a change in the system's parameters.

Chaotic Intermittency The movement of a dynamical system between qualitatively distinct types of behaviour. This can either be a crisis-induced intermittency, where the system moves between distinct chaotic flows, as in the Lorenz '63 system, or a Pomeau-Manneville intermittency where the system moves between chaotic and quasi-predictable dynamics, as in the CdV system.

Heteroclinic orbit A trajectory in phase space which connects one fixed point to another. Heteroclinic orbits often bifurcate into the more general and less symmetric homoclinic orbit.

Homoclinic orbit A trajectory in phase space which starts from and returns to the same fixed point.

Hopf Bifurcation A bifurcation in which a periodic orbit, which might be stable or unstable, is created or destroyed.

Hyperbolicity Hyperbolic systems have no zero Lyapunov exponents at any point in phase space, which prevents the number of unstable and stable dimensions from changing. Nonhyperbolic systems in contrast can show such variation in stability.

Invariant measure A set of points in the phase space that evolve to themselves under the system's dynamics. A system lying on an invariant measure will not leave it at any

future time. Fixed points, periodic orbits, and strange attractors are all examples of invariant measures.

Manifold A subspace of the system's phase space, obeying some technical conditions including smoothness which make it well-behaved.

Saddle-node Bifurcation A bifurcation in which two fixed point solutions are created or destroyed. Often related to the onset of hysteresis dynamics in a system.

Shadowing A chaotic trajectory is said to shadow an exact solution of the system if it stays close to the exact solution for some period of time.

Stable manifold A manifold along which trajectories converge. By contrast, along an unstable manifold trajectories will diverge.

A.2 The hidden Markov model

Hidden markov models are an approach to clustering time series that uses both spatial and temporal data. One of the main drawbacks of the commonly used K means algorithm is that it ignores any temporal coherence in the data; it considers only density in state space, not persistence or consistency of dynamics in time. The hidden Markov model however makes use of the concept of Markov chains - a series of datapoints x_t which are conditional only on the previous value x_{t-1}.

We must choose the number of clusters *a priori* and assume that each datapoint x_t is assigned to a single cluster k, an assignment we indicate by the state vector $S = [s_1, ...s_T]$ where $s_t \in [1, ..., K]$. We consider the evolution of S to act as a Markov chain, depending only on the previous state. This allows us to define a transition matrix $\underline{\underline{A}}$, where $A_{ij} = P(s_t = i|s_{t-1} = j)$. How do we link this underlying series of state transitions S to the observations X? We consider each cluster k to be associated with a Gaussian distribution Γ_k, and assume that x_t is drawn from Γ_{s_t}

x_1 is special as it is unconstrained by a previous datapoint, so we have to also consider a probability distribution over the starting states, π. Between π, the transition matrix $\underline{\underline{A}}$ and the means and covariances of the $\{\Gamma_k\}$ there are a total of $K^2 + KD(D+1)$ free parameters, collectively referred to as θ, that must be learnt in order to fit the hidden Markov model to data.

How can this be done? In summary, we want to maximise the probability,$P(X|S,\theta)$, of obtaining our dataset from the hidden Markov model, given some model parameters and some state sequence. This state sequence is simply the joint probability of obtaining each datapoint $x_t \in X$:

$$P(X|S,\theta) = \prod_{t=1}^{T} P(x_t|S,\theta) \tag{A.1}$$

By taking a logarithm we can define the log-probability given some parameters, and turns the product into a more workable summation:

$$\mathcal{L}(\theta) = \log P(X|S,\theta) = \sum_{t=1}^{T} \log \left[\sum_{S} P(x_t, S|\theta) \right] \tag{A.2}$$

By taking the state sequence as a variable, we have introduced the sum over all possible state sequences S which has K^T terms and so is incalculable for even modestly long time series. We proceed as follows:

1. We make an initial random guess at our parameters, θ=θ^{old}
2. We define a posterior distribution over the hidden state at time t:

$$\gamma(s_t) = P(s_t|X,\theta^{\text{old}}) \tag{A.3}$$

and a posterior joint distribution over the state at two sequential times t and $t-1$:

$$\xi(s_{t-1}, s_t) = P(s_{t-1}, s_t|X,\theta^{\text{old}}) \tag{A.4}$$

3. We evaluate the values of γ and ξ for each point in the data series using the forward backward algorithm
4. We find a function $Q(\theta, \theta^{\text{old}})$ that provides a lower bound on $\mathcal{L}(\theta)$ and that can be written in terms of γ and ξ
5. We maximise $Q(\theta, \theta^{\text{old}})$ with respect to θ, and use that to update θ^{old}
6. We check if Q has converged, and if not, we return to step 2.

A.2.1 Full derivation

For the interested reader, we now provide a more detailed discussion of the hidden Markov algorithm.

We start by explicitly deriving Q. By basic algorithmic manipulation we can write $\mathcal{L}(\theta)$ as :

$$\mathcal{L}(\theta) = \log \left[\sum_S P(S|X, \theta^{\text{old}}) \frac{P(X, S|\theta)}{P(S|X, \theta^{\text{old}})} \right] \tag{A.5}$$

Where we have just introduced the log probability of our current best guess at the parameters θ. Because the logarithm is a concave function, we can say that:

$$\log \left[\sum_S P(S|X, \theta^{\text{old}}) \frac{P(X, S|\theta)}{P(S|X, \theta^{\text{old}})} \right] \geq \sum_S P(S|X, \theta^{\text{old}}) \log \left[\frac{P(X, S|\theta)}{P(S|X, \theta^{\text{old}})} \right] \tag{A.6}$$

Which is of course entirely equivalent to:

$$\mathcal{L}(\theta) \geq \sum_S \left[P(S|X, \theta^{\text{old}}) \log P(X, S|\theta) - P(S|X, \theta^{\text{old}}) \log P(S|X, \theta^{\text{old}}) \right] \tag{A.7}$$

Therefore if we maximise the RHS of the above expression, we will at the very least not decrease $\mathcal{L}(\theta)$ (and practically speaking will always increase it). Since the second term has no θ dependence, we can ignore it when optimising with respect to θ, and therefore

define:

$$Q(\theta, \theta^{\text{old}}) = \sum_{S} \left[P(S|X, \theta^{\text{old}}) \log P(X, S|\theta) \right] \tag{A.8}$$

Q will prove a much easier quantity to maximise than $\mathcal{L}(\theta)$, but first we must find out how to explicitly evaluate it. The joint distribution over the observations and the hidden states, $P(X, S|\theta)$, is given by:

$$P(X, S|\theta) = \left[P(s_1|\pi) \prod_{t=2}^{T} P(s_t|s_t - 1, A) \right] \prod_{s=1}^{T} P(x_s|s_s, \Gamma_{s_s}) \tag{A.9}$$

The term in square brackets expresses that the state at time t is determined by its previous state, as governed by the transition matrix A. The term outside of brackets expresses that the observations are derived by the current state and the corresponding Gaussian distribution.

As mentioned in the summary, we now introduce some new notation, $\gamma(s_t)$ for the marginal posterior distribution of the state at time t (represented by a length K vector for each timestep), and $\xi(s_{t-1}, s_t)$ for the joint posterior distribution of the state at times t and $t-1$ (represented by a KxK matrix for each timestep):

$$\gamma(s_t) = P(s_t|X, \theta^{\text{old}}) \tag{A.10}$$

$$\xi(s_{t-1}, s_t) = P(s_{t-1}, s_t|X, \theta^{\text{old}}) \tag{A.11}$$

Equations A.9, A.10 and A.11 let us rewrite Q explicitly as:

$$Q(\theta, \theta^{\text{old}}) = \sum_{k=1}^{K} \gamma(s_1)_k \log \pi_k + \sum_{t=2}^{T} \sum_{j=1,k=1}^{K} \xi(s_{t-1,j}, s_{t,k}) \log A_{jk} + \sum_{t=2}^{T} \sum_{k=1}^{K} \gamma(s_t)_k \log P(x_t|\Gamma_k) \tag{A.12}$$

If we can evaluate γ and ξ, then we can maximise Q.

This is done with the use of the forward-backward algorithm, sometimes called the

Baum-Welch algorithm. We define two additonal quantities; $\alpha(s_t)$, which represents the probability of observing all the data up until time t as well as the hidden state at t:

$$\alpha(s_t) := P(x_1, ..., x_t, s_t|\theta^{\text{old}}) \tag{A.13}$$

and $\beta(s_t)$, which represents the probability of observing all the data after time t, given the hidden state at time t:

$$\beta(s_t) = P(x_{t+1}, ...x_T|s_t, \theta^{\text{old}}) \tag{A.14}$$

These quantities are useful to us, because we can exploit the Markov structure of our system to calculate them recursively. We start with $\alpha(s_{1,k}) = P(x_1, s_1 = k) = \pi_k P(x_1|\Gamma_k)$. Using this starting point, we can say:

$$\alpha(s_{t,k}) = P(x_t|s_t) \sum_j \alpha(s_{t-1,j}) A_{j,k} \tag{A.15}$$

Which just says the probability of the current hidden state and observation, is the probability of observing each of the K possible previous states, multiplied by the corresponding transition and emission probabilities. Each of these iterations involves a $K \times K$ matrix multiplication, and so the time taken to compute the whole chain of α values scales as $O(K^2T)$. To compute the β values, we follow a similar procedure, but work backwards in time:

$$\beta(s_{t,k}) = \sum_j \beta(s_{t+1,j}) P(x_{t+1}|s_{t+1,j}) A_{k,j} \tag{A.16}$$

This says the probability of observing all measurements after time t, is the probability of observing all measurements after time $t + 1$, summed over the K possible ways to go from s_t to s_{t+1}. Here our initial condition is $\beta(s_T) = 1$, and so we are able to compute the whole chain of β values, again in $O(K^2T)$ time.

We now have values of $\alpha(s_t)$ and $\beta(s_t)$ for all t, and so all that remains is to rewrite γ and ξ in these terms. $\gamma(s_t)$ can be written as:

$$\gamma(s_t) = P(s_t|X) = \frac{P(X|s_t)P(s_t)}{P(X)} = \frac{P(x_1, ...x_t, s_t)P(x_{t+1}, ..., x_T|s_t)}{P(X)} = \frac{\alpha(s_t)\beta(s_t)}{P(X)} \tag{A.17}$$

While $\xi(s_t, s_{t-1})$ can be written as:

$$\xi(s_t, s_{t-1}) = P(S_{t-1}, s_t|X) = \frac{P(X|s_{t-1}, s_t)P(s_{t-1}, s_t)}{P(X)} = \frac{\alpha(s_{t-1}\beta(s_t)P(x_t|s_t)P(s_t|s_{t-1})}{P(X)} \tag{A.18}$$

We now know how to calculate all terms in these equations, except for the $P(X)$ which in fact will turn out to always cancel. Thus we can calculate Q exactly. Now all that remains is to maximise the parameters $\theta = \{A, \pi, \Gamma\}$ with respect to Q. This can be done very simply by setting the derivative of Q with respect to each parameter to zero, and including relevant Lagrange multipliers and non-negativity constraints where needed. This gives us that:

$$\pi_k = \frac{\gamma(s_{1,k})}{\sum_j^K \gamma(s_{1,j})} \tag{A.19}$$

$$A_{jk} = \frac{\sum_{t=2}^T \xi(s_{t-1,j}, s_{t,k})}{\sum_l^K \sum_{t=2}^T \xi(s_{t-1,j}, s_{t,l})} \tag{A.20}$$

And that the mean μ and covariance Σ for each gaussian distribution Γ is:

$$\mu_k = \frac{1}{T_k} \sum_{t=1}^T \gamma(s_{t,k}) x_t \tag{A.21}$$

$$\Sigma_k = \frac{1}{T_k} \sum_{t=1}^T \gamma(s_{t,k})(x_t - \mu_k)(x_t - \mu_k)^T \tag{A.22}$$

By substituting in these new, optimised parameters, we can get new updates for γ and ξ, and so repeat the optimisation procedure until convergence. Once we have a fully

trained model, we can use that to find the most probable sequence of state transitions that produced the observed data X, and we define that as our state sequence.

A.3 K means algorithm

The K-means algorithm is commonly used in atmospheric sciences in order to pick out distinct clusters from a dataset. No features of the clusters are specified in advance (although the number of clusters, K, must be prescribed) and so K-means falls into the class of unsupervised machine learning algorithms; the learnt structure emerges directly from the data. We consider a dataset X of dimension D and series length T. We will index over the T separate datapoints in X with $t \in [1, 2, ..., T]$, and over the K individual clusters with $k \in [1, ..K]$. Our aim is to assign each data point x_t in X to one of the K clusters, so for each x_t we define K indicator variables $r_{tk} \in \{0, 1\}$. If $r_{sj} = 1$ that means the s^{th} datapoint in X is assigned to cluster j. We assume that the assignment is hard, so that $r_{sj} = 1$ implies $r_{sk} = 0$ for $k \neq j$; a single datapoint can only lie in a single cluster. Relaxing this restriction, and letting r_{tk} take any real value between 0 and 1, produces the probabilistic or 'fuzzy' version of the K means algorithm, but we shall focus on the non-probabilistic case.

If different datapoints belong to the same cluster that should imply some similarity between them not shared by datapoints in different clusters. We can formalise this by saying the Euclidean distance (measured in the D-dimensional state space) between points in the same cluster should be smaller than that between points in different clusters. For each cluster we introduce a D-dimensional vector, μ_k, which defines the cluster's centre (called the centroid). We then aim to find values of the $\{\mu_k\}$, and the $\{r_{tk}\}$ that minimises:

$$J := \sum_{t=1}^{T} \sum_{k=1}^{K} r_{tk} \|x_t - \mu_k\|^2 \tag{A.23}$$

Here J, which is simply the total sum of distances between datapoints and cluster cen-

troids, is called the distortion measure.

To minimise J, we follow a two-step iterative procedure. We make an arbitrary initial guess at the $\{\mu_k\}$, typically by setting them equal to randomly selected members of X.

Our first step is then to find r_{tk} that minimise J, holding the μ_k constant. J is linear in r_{tk}, which means we can minimise for each t independently. Given our hard assignments for r_{tk}, this naturally implies the minimisation:

$$r_{tk} = \begin{cases} 1 & \text{if } k = \arg\min_j \|x_t - \mu_j\|^2 \\ 0 & \text{otherwise} \end{cases} \tag{A.24}$$

This just says that each datapoint should be assigned to the cluster centroid it is closest to. Our second step is to find values of μ_k that minimise J, holding the r_{tk} constant.This is done by setting the first derivative of J to 0. Again, we can simply minimise for each value of k independently:

$$\frac{\partial J}{\partial \mu_k} = 2 \sum_{t=1}^{T} r_{tk}(x_t - \mu_k) = 0 \tag{A.25}$$

That easily rearranges to:

$$\mu_k = \frac{\sum_t r_{tk} x_t}{\sum_t r_{tk}} \tag{A.26}$$

which tells us that the k^{th} centroid should be set to the mean position of every datapoint assigned to cluster k (which perhaps could have been guessed from the algorithm's name).

These two steps:

1. Minimise J with respect to the $\{\mu_k\}$ by assigning datapoints to the nearest centroid.
2. Minimise J with respect to the $\{r_{tk}\}$ by setting the centroid to the mean position of the datapoints it contains.

are then repeated until convergence. Each step strictly decreases the size of J, so convergence to a local minimum is guaranteed. To try and find the global minimum, the K means algorithm should be rerun many times, especially in high dimensional datasets, a step that is often not commented on in the atmospheric science literature. Anecdotally, in typical atmospheric fields, many local minima seem to exist, meaning quite large number of repetitions (50-100) are required to get consistent results. In all work in this thesis, 100 repetitions are used for all applications of the K means algorithm.