Recent trends in the atmospheric circulation



Rhidian Thomas

Jesus College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2025

Abstract

To build confidence in projections of future change, models should be able to reproduce observed trends in the current climate. However, despite successes in modelling some aspects of climate change associated with the rise in global temperature, models struggle to capture trends in various regional and circulation-driven features. The broad aim of this thesis is to develop new tools to study recent large-scale trends in the climate system, with a particular focus on the contribution of the atmospheric circulation.

The main approach is to explore the range of trends in the initialised hindcasts of a near-term prediction model, DePreSys3, and use these to provide context for observed trends. In the tropics, upper-level temperature trends in the hindcasts are cooler than in the equivalent free-running model, bringing them into agreement with reanalyses, and the improvement is explained through linear surface temperature scaling theory. Much of the improvement is shown to occur through capturing the magnitude of SST warming in the tropical Pacific. The hindcasts also support a robust weakening of the NH Hadley Cell, in line with pressure observations.

A key conclusion is that the hindcasts are consistent with a remarkably wide range of trends over much of the climate system, most notably in the Arctic. Sea ice trends are constrained by the initialisation and explain very little of this spread; instead, the spread in hindcast Arctic warming, as well as in Eurasian cooling and even global surface temperature trends, is mainly a function of trends in annular mode-like structures. While not always a dominant contributor to the observed trend, this highlights the role that seasonal variability, likely atmospheric in origin, can play in modulating multidecadal trends.

Despite considerable ensemble spread, the hindcasts also support the emergence of poleward jet shifts on the global scale. The effect of the tropical Pacific on the jets is more subtle than on tropical temperature trends, but it does appear to contribute to the improvement in the hindcast wind trends, as well as regional temperature and circulation trends.

Finally, a novel budget-based method is developed to quantify the dynamical contribution to atmospheric temperature trends. The dynamical and diabatic contributions largely balance each other, and trends are identified in the mean and eddy circulations. Trends in the ERA5 budget terms largely lack significance after accounting for temporal autocorrelation, suggesting that a longer observational record may be necessary to achieve robust trends.

Acknowledgments

I am deeply indebted to my supervisor, Tim Woollings, both for his help with the scientific questions of this thesis, and for guiding me through my time at Oxford. I am also very grateful to my CASE supervisor at the Met Office, Nick Dunstone, for his guidance during the DPhil and his advice for life after it. I could not have asked for a better supervisory team.

Throughout my studies, I have benefited from conversations with colleagues at Oxford and abroad. I would like to thank my collaborators at the University of Chicago, Joonsuk Kang and Tiffany Shaw, for our many stimulating discussions of all things Pacific. I am also grateful to those who have read and commented on my work during Transfer and Confirmation reviews: Myles Allen, Lesley Gray, Tim Palmer, David Marshall, and Antje Weisheimer. I would like to thank the members of the Met Office Monthly to Decadal Prediction team, not only for their helpful comments on this work but also for making me feel welcome in Exeter. Thanks also to everyone in the Climate Dynamics group at AOPP, which has been a fun and engaging home at Oxford.

When things have gone wrong, I have been fortunate to have had support from all corners. From AOPP and the DTP, I am indebted to Victoria Forth and David Marshall for their help in the final months of this project; without them this thesis would have suffered enormously. I have also enjoyed meeting old friends and new in Oxford – Henry, Joe, Rosie, Pete, Sol, and the AOPP RC – and am grateful for their support throughout. Most of all, I would like to thank my parents, to whom I owe everything, and Hannah, for her endless encouragement and love - and for her saint-like patience during the write-up.

Contents

1 Introduction **1**

1.1 Overview . . . 1
1.2 What is a trend? . . . 2
1.3 Thesis structure . . . 4
1.4 Relation to published work . . . 4

2 Background **5**

2.1 Shifting jets . . . 5
2.1.1 Response of the jets to forcing . . . 6
2.1.1.1 Tug-of-war . . . 6
2.1.2 Observed trends . . . 8
2.1.2.1 Zonal-mean trends . . . 8
2.1.2.2 Regional trends . . . 10
2.1.3 Are models failing to capture observed trends? . . . 10
2.1.4 Summary and outstanding questions . . . 11
2.2 Tropical tropospheric warming . . . 12
2.2.1 Response to forcing . . . 12
2.2.2 Observed trends . . . 13
2.2.3 Explanations of the trend discrepancy . . . 13
2.2.3.1 Internal variability . . . 13
2.2.3.2 Observational uncertainty . . . 14
2.2.3.3 Model physics . . . 14
2.2.4 Summary and outstanding questions . . . 15
2.3 Arctic amplification . . . 15
2.3.1 Mechanisms of amplification . . . 15
2.3.2 Observed trends . . . 15
2.3.3 Internal variability . . . 17
2.3.4 Impacts of AA on the midlatitudes . . . 18

2.3.4.1 Role of coupling . . . 19
2.3.4.2 The role of deep Arctic warming . . . 19
2.3.5 Summary and outstanding questions . . . 21
2.4 Pacific SST trends . . . 22
2.4.1 Mechanisms for a weakening SST gradient . . . 22
2.4.2 Observed trends . . . 23
2.4.2.1 Internal variability . . . 24
2.4.2.2 Forced response . . . 24
2.4.3 Implications of the tropical Pacific SST trend . . . 25
2.4.4 Summary and outstanding questions . . . 26
2.5 Hadley Circulation . . . 26
2.5.1 Strength of the circulation . . . 26
2.5.1.1 Weakening response . . . 26
2.5.1.2 Observed trends . . . 27
2.5.2 Summary and outstanding questions . . . 28
2.6 Seasonal-to-decadal prediction . . . 28
2.6.1 What is near-term prediction? . . . 29
2.6.2 Terminology for initialised and free-running models . . . 29
2.6.3 New methods using S2D models . . . 30
2.7 Research questions . . . 31

3 Zonal-mean trends in the hindcasts 34
3.1 Introduction . . . 34
3.2 Methods . . . 37
3.2.1 Hindcast data . . . 37
3.2.2 Reanalysis data . . . 38
3.2.3 Free-running model data . . . 39
3.2.4 Trend analysis . . . 39
3.3 Tropical tropospheric temperatures . . . 40
3.3.1 Trends in the reanalyses and the free-running model . . . 41
3.3.2 Trends in the hindcasts . . . 42
3.3.3 Variability and unseen trends in the hindcasts . . . 44
3.4 Arctic amplification . . . 44
3.4.1 Lower tropospheric AA . . . 44
3.4.2 Surface amplification . . . 46
3.4.3 Sea ice scaling . . . 49

3.4.4 Variability and unseen trends in the hindcasts 49
3.4.5 Summary of Arctic results 52
3.5 Shifting jets 52
3.5.1 Trends in the reanalyses and the free-running model 52
3.5.2 Trends in the hindcasts 53
3.5.2.1 Winter jet trends 54
3.5.2.2 Summer jet trends 55
3.5.2.3 Jet biases 55
3.5.3 Variability and unseen trends in the hindcasts 57
3.5.4 Relation to meridional temperature gradients 58
3.5.5 Summary of jet trends 60
3.6 Hadley circulation 61
3.7 Conclusions 65

4 Spatial structure of trends in the hindcasts 67
4.1 Surface-air temperature 67
4.1.1 Summary of trend improvements 67
4.1.2 Internal variability in the hindcasts 71
4.1.3 Eurasian cooling 74
4.2 Sea-level pressure 79
4.2.1 Trends in the North Atlantic Oscillation 82
4.3 Discussion and conclusions 83

5 The impact of tropical Pacific SST trends on global and regional circulation 86
5.1 Introduction 86
5.2 Methods 88
5.2.1 Hindcast data 88
5.2.2 Pacemaker experiments 89
5.2.3 Observational and reanalysis data 90
5.2.4 Trend analysis 90
5.2.5 ENSO composites 90
5.2.6 Performance of the conditioned hindcasts 91
5.2.7 Nomenclature 92
5.3 Trend response 92
5.3.1 Zonal-mean response 92
5.3.2 Regional response 96

5.3.2.1 Tropics and subtropical Pacific 97
5.3.2.2 North Pacific and Americas 101
5.3.2.3 North Atlantic and European sector 101
5.4 Comparison with reanalysis 103
5.4.1 Tropical temperatures 103
5.4.2 Shifting jets 104
5.4.3 Regional circulation 105
5.4.4 Regional temperature and precipitation 108
5.5 Boreal summer results 110
5.5.1 Methods 110
5.5.2 South Pacific circulation 111
5.6 Discussion and conclusions 113

6 Diagnosing the circulation contribution to ERA5 temperature trends 118
6.1 Introduction 118
6.2 Methods 120
6.2.1 Reanalysis data 120
6.2.2 Tendency budget 120
6.3 Accumulated heating trends 123
6.3.1 Method 123
6.3.2 Results 124
6.4 Drivers of tendency trends 125
6.5 Seasonal tendency trends 128
6.6 Significance testing 131
6.7 Discussion and conclusions 133

7 Conclusions 137
7.1 Summary of results 137
7.2 Themes and open questions 140
7.3 Limitations and future work 142

A Extended Figures for Chapters 3 and 4 145

Bibliography 148

List of Figures

2.1 Zonal-mean temperature and zonal wind response under RCP8.5 forcing. 8
2.2 Trends in zonal wind indices at 500hPa plotted against trends in the 250hPa meridional temperature gradient. 9
2.3 AA distributions in CMIP models. 16
2.4 Eurasian cooling and deep Arctic warming. 20
2.5 SST and SLP trend discrepancies between models and observations. . 32
2.6 CMIP5 and reanalysis trends in DJF Hadley circulation strength. . . 33

3.1 SAT bias and coupled variability in the hindcasts. 35
3.2 DJF zonal-mean temperature trends in the hindcasts. 40
3.3 JJA zonal-mean temperature trends in the hindcasts. 41
3.4 Vertical warming structure in the tropics. 43
3.5 Lower tropospheric AA in the hindcasts. 45
3.6 Surface AA in the hindcasts. 47
3.7 Climatological winter inversion strength. 48
3.8 Zonal wind trends associated with AA in the hindcasts. 50
3.9 SAT and SLP trends associated with AA in the hindcasts. 51
3.10 Zonal-mean zonal wind trends in the hindcasts. 53
3.11 Jet shifts as measured by trends in zonal wind indices. 54
3.12 Mean model zonal wind biases. 56
3.13 Mean model biases in the zonal wind indices. 57
3.14 Trends in zonal wind indices plotted against trends in the 250hPa meridional temperature gradient. 59
3.15 Distribution of trends in Hadley Circulation strength. 62
3.16 Reanalysis trends in the DJF meridional mass streamfunction. 63
3.17 Model trends in the DJF meridional mass streamfunction. 64

4.1 Area-averaged SAT trend distributions for DJF. 68
4.2 Maps of ensemble-mean hindcast SAT trends in DJF. 70

4.3 Spatial patterns associated with warmest and coolest hindcast GMST trends. 72
4.4 GMST trend as a function of the trend in the zonal SAT gradient in the tropical Pacific. 73
4.5 Eurasian temperature trends plotted against Ural pressure trends in the 1m hindcasts. 75
4.6 Mean temperature and circulation trends associated with the Eurasian cooling. 76
4.7 Trends in randomly selected members of the EC subset. 77
4.8 Zonal-mean temperature trend in the EC subset, relative to the hindcast ensemble-mean. 78
4.9 Maps of ensemble-mean hindcast SLP trends. 81
4.10 Hindcast trends in the winter and summer NAO. 82

5.1 Definition of ENSO events used for composites. 90
5.2 Ensemble-mean tropical Pacific SST trends in the conditioned and unconditioned models. 91
5.3 Ensemble-mean, zonal-mean temperature and zonal wind trends in the unconditioned models 93
5.4 The Pacific-forced zonal-mean trend response. 94
5.5 Interannual zonal-mean DJF La Nina minus El Nino response. 95
5.6 Ensemble-mean spatial trends in the unconditioned models. 97
5.7 The Pacific-forced spatial trend response. 99
5.8 Interannual DJF La Nina minus El Nino spatial response. 100
5.9 Tropical tropospheric temperature trends. 103
5.10 Trends in zonal wind indices plotted against trends in meridional temperature gradient. 105
5.11 Trends in regional circulation indices for the conditioned and unconditioned models. 107
5.12 Regional SAT trend improvement in the conditioned models. 108
5.13 Regional precipitation trend improvement in the conditioned models. 109
5.14 Conditioned model SLP trend responses in JJA. 111
5.15 Conditioned hindcast SAT and precipitation trend responses in JJA . 112
5.16 Regional trend indices in the conditioned and unconditioned hindcasts for JJA. 113

6.1 Trend in zonal-mean, annual-mean temperature in ERA5 over 1989-2020.119

6.2 Climatology of terms in the local temperature tendency budget. . . . 122
6.3 Trends in annual-mean accumulated heating anomalies. 125
6.4 Drivers of the tendency trend due to the MMC. 127
6.5 Drivers of the tendency trend due to the eddies. 128
6.6 Seasonal trends in temperature tendency due to the net dynamics. . . 129
6.7 Seasonal trends in temperature tendency due to the MMC. 130
6.8 Seasonal trends in temperature tendency due to the eddies. 131
6.9 Significance testing of the dynamical tendency trends. 132

A.1 Area-averaged SAT trend distributions for JJA. 146
A.2 Maps of ensemble-mean hindcast SAT trends in JJA. 147

List of abbreviations

AA Arctic amplification

AL Aleutian Low

AMIP Atmospheric Model Intercomparison Project

AMO/AMV Atlantic Multidecadal Oscillation / Atlantic Multidecadal Variability

AMOC Atlantic Meridional Overturning Circulation

AR Atmospheric river

AW Arctic warming

B-K Barents-Kara seas

BL Boundary layer

CMIP Coupled Model Intercomparison Project

DCPP Decadal Climate Prediction Project

EC Eurasian cooling

ENSO El Niño–Southern Oscillation

EP East Pacific

GCM General circulation model

GC2 HadGEM3-GC2

GHG Greenhouse gases

GMST Global mean surface temperature

HC Hadley cell

ITCZ Intertropical convergence zone

LE Large ensemble

LT Lower troposphere

MALR Moist adiabatic lapse rate

MCS Mesoscale convective system

MMC Mean meridional circulation

NAM Northern Annular Mode

NAO North Atlantic Oscillation

NH Northern hemisphere

NonCond Non-conditioned hindcast subset (Ch. 5)

NonPace Non-Pacemaker version of CESM2 (i.e. CESM2-LE, Ch. 5)

PAMIP Polar Amplification Model Intercomparison Project

PDO Pacific Decadal Oscillation

PWC Pacific Walker Cell

PacCond Conditioned hindcast subset (Ch. 5)

PacPace Pacific Pacemaker version of CESM2 (Ch. 5)

RCP Representative Concentration Pathway

RMS Root mean square

S2D Seasonal-to-decadal

SAM Southern Annular Mode

SH Southern Hemisphere

SLP Sea-level pressure

SNAO Summer North Atlantic Oscillation

SP dipole South Pacific dipole

SSP Shared Socioeconomic Pathway

SST Sea surface temperature

T300, T500 Temperature at 300hPa, 500hPa

TAS Surface-air temperature

TEM Transformed Eulerian Mean

TUT Tropical upper troposphere

UT Upper troposphere

UNSEEN UNprecedented Simulated Extremes using ENsembles

WP West Pacific

WTG Weak temperature gradient approximation

Chapter 1

Introduction

1.1 Overview

The effects of climate change are emerging globally, and anthropogenic forcing is expected to lead to further disruptions in years to come (Collins et al. 2024). To build confidence in projections of future change, an important benchmark for models is to be able to reproduce observed trends in the current climate. A key achievement of climate science has been to accurately model the global-mean surface warming of nearly 1.5C since preindustrial times, and to confidently attribute it to anthropogenic influences (Intergovernmental Panel on Climate Change (IPCC) 2023). Much of our successes in both understanding and modelling of human influence on the climate, such as changes in the hydrologoical cycle (Held and Soden 2006), stems from the thermodynamic changes associated with this rise in global temperature. This is in stark contrast with changes to the atmospheric dynamics, which are more poorly understood and represent considerable uncertainty in model projections (Shepherd 2014). Consequently, there remain doubts over the ability of models to represent observed change in other parts of the climate system; this is particularly true of changes in the circulation itself, or in regional features, which are particularly sensitve to dynamical changes (Shaw et al. 2024).

The broad aim of this thesis is to develop new tools to study trends in the climate system. The main approach is to explore trends in the initialised hindcasts of a near-term prediction model, which forms the basis of chapters 3 through 5. In chapter 6, a budget method is developed to quantify the dynamical contribution to atmospheric temperature trends.

A holistic survey of trends is given throughout the atmosphere, with a particular focus on five features of the climate system where models are in tension with observed trends:

1. **Shifting jets**: A range of trends in jet latitude are seen in models, partly reflecting uncertainty in changes to meridional temperature gradients in the midlatitudes. However, observed shifts over the satellite era appear to be large, especially relative to changes in temperature gradients.

2. **Tropical tropospheric warming**: Coupled models consistently warm more in the tropical mid-upper troposphere than most observational and reanalysis datasets. By altering the equator-to-pole temperature gradient, the tropical warming response is expected to be a leading ingredient in the jet response to climate change.

3. **Arctic amplification**: In the NH, the effect of tropical warming on the jet may be opposed by amplified Arctic warming. Warming appears to be occurring towards the upper edge of model simulations, although the magnitude of warming is sensitive to the choice of time period and observational dataset.

4. **Trends in the tropical Pacific**: Observed trends suggest a strengthening of the east-west sea surface temperature (SST) gradient and an intensified Pacific Walker Circulation. Models often struggle to replicate these trends, raising potential implications outside the Pacific.

5. **Hadley circulation trends**: The strength of the Hadley circulation (HC) depends on several factors. Despite the complexity of the system, model simulations robustly show a weakening of the NH winter HC over the satellite era and in future projections. However, successive generations of reanalyses have instead shown a strengthening of the circulation.

1.2 What is a trend?

Before discussing emerging trends in the climate system, it is necessary to first define the term. The word "trend" can be interpreted in different ways, for example:

- The trend in a quantity X is the change in X over a specified time interval, often taken as the linear regression of X against time.

- Informally, "trend" may also suggest monotonicity of change, or permanence of change.

- In the context of climate science, "trend" is sometimes taken to imply a forced response to climate change.

This thesis will exclusively use the first definition. A linear trend in some quantity can always be defined, regardless of whether the change in that quantity is understood to be a result of external forcing or purely internally generated, and regardless of whether the trend is considered significant according to some statistical test. Trends as considered here therefore consist of a forced and unforced component, and both components may contribute to trend discrepancies between models and observations (Jain et al. 2023).

The focus of this thesis is on trends calculated over periods broadly aligned with the era of satellite observations, beginning around 1980. As well as the greater observational coverage afforded by space-borne instruments, this period coincides with rapid observed changes in various features of the climate system that have been attributed to anthropogenic forcing, such as amplified warming at high latitudes (Gillett et al. 2008). Note that the precise periods used in the trend calculations differ slightly between chapters due to the differing availability of model data, as explained in each chapter: Chapters 3 and 4 use 1981/82-2022/23, Chapter 5 uses 1981/82-2018/19, and Chapter 6 uses 1979/80-2020/21.

Linear trends are calculated using ordinary least squares regression throughout. While this method can be sensitive to outliers, its widespread use in the scientific literature means that it can be easily implemented at scale using standard software libraries, aiding reproducibility and comparison to other studies. Despite acknowledging that variability and forced changes in the climate need not be linear, this thesis will also focus only on linear trends. Much of the relevant literature discusses emerging linear trends in the climate system; for the results of this thesis to be directly comparable, a linear approach is adopted here. More broadly, a key motivation for investigating changes in the climate system stems from the observed increase in global-mean surface temperature, which is approximately linear over the satellite era (Intergovernmental Panel on Climate Change (IPCC) 2023). The central importance of this fact to other aspects of the climate further justifies a linear approach as a starting point for trend analysis.

Full details of significance calculations are provided in each chapter. In general, where only a single timeseries exists (such as for trends in observations or reanalyses, such as Figure 3.2a), significance is calculated using a two-tailed Student's t-test with a null hypothesis of no trend. For trends in model ensembles consisting of multiple realisations, significance is defined relative to the model spread: for example, stippling on Figures 3.2b-d show where at least 80% of members agree on the sign of the trend.

1.3 Thesis structure

The next chapter reviews the literature concerning the emerging trends outlined above, introduces the field of near-term prediction, and outlines the main research questions addressed in this thesis. Chapters 3 and 4 present overviews of zonal-mean and spatial trends in the hindcasts, respectively. Chapter 5 focuses on the impact of trends in the tropical Pacific on global climate trends. The temperature budget results are in Chapter 6. The results are discussed and synthesised in Chapter 7, and outstanding questions are identified for future research.

1.4 Relation to published work

Some of the hindcast results in Chapters 3 and 4 are contained in a paper under review at Journal of Climate, co-authored by my supervisors Tim Woollings and Nick Dunstone. Those chapters also contains additional material that will not feature in that article. All writing, analysis, and figure creation in both the submitted manuscript and in these chapters is my own; however, I would also like to acknowledge improvements made to the manuscript following comments from my co-authors, as well as Mika Rantanen and two anonymous reviewers.

Chapter 5 contains work done in collaboration with colleagues at the University of Chicago, Joonsuk Kang and Tiffany Shaw, as well as Tim Woollings and Nick Dunstone. Work in this chapter forms the basis of a paper currently in preparation, with equal contribution and first authorship from JK and me; the remaining authors will be ND, TS, and TW. Analysis of the hindcast data was performed by me, and the Pacemaker data by JK. While performing the analysis, we each worked in parallel to create preliminary figures using our own datasets, before sharing data in order to plot both models on the same axes. The manuscript versions of some figures were created by JK, and are included in this thesis with his permission; they are Figures 5.2 to 5.13 in Chapter 5. All other figures in Chapter 5 were my own creation. Interpretation of both the hindcast and Pacemaker results is my own, as is all of the written text. I am grateful to my collaborators for our many discussions of this work over the last year.

Chapter 2

Background

Emerging trends

The following sections review a selection of trends that are emerging in observations but which are poorly captured by coupled climate models. Each section begins by discussing how these features respond to anthropogenic forcing in models, using arguments from theory where appropriate. Then, evidence of observed trends is summarised and compared to trends in models; where these are in tension, possible explanations of the discrepancy are evaluated. For trends in the Arctic and the tropical Pacific, I additionally review the implications of trends in these regions for the rest of the climate system. Finally, I summarise the literature for each emerging trend, highlighting outstanding questions.

2.1 Shifting jets

Jets are a general feature of rotating planetary bodies (Galperin and Read 2019) and arise through distinct mechanisms. In the subtropics, poleward moving air in the upper branch of the Hadley Circulation conserves angular momentum and is accelerated eastwards with respect to the Earth. This leads to a region of intense wind shear, giving rise to the subtropical jet, which is strongest in the upper troposphere. A distinct process occurs in the eddy-dominated midlatitudes. Baroclinic waves, such as those forced by horizontal contrasts in heating in the storm tracks, propagate upwards and equatorwards from their source in the midlatitudes. The associated horizontal flux of momentum into the midlatitudes accelerates and deepens the upper-level westerlies, leading to a deeper, eddy-driven jet (e.g. Vallis 2017). In practice, the subtropical and eddy-driven jets are not always well separated, but it is sometimes useful to distinguish between them due to their distinct origins.

2.1.1 Response of the jets to forcing

An early target of climate models was the investigation of changes in the jets in response to anthropogenic forcing (Manabe and Wetherald 1980). Various mechanisms have been proposed to explain forced changes in the midlatitude circulation, reflecting the complex processes which affect the jets. These include processes local to the midlatitude troposphere; for example, the warming response of high clouds in the midlatitudes has been shown to lead to a poleward jet shift, independent of the underlying SSTs (Voigt and Shaw 2016). The jets are also influenced by remote factors originating from both higher and lower latitudes, as well as the stratosphere (e.g. Butler et al. 2010).

Reflecting these complex and often competing influences, models generally show a modest poleward shift of the zonal-mean jets in response to CO2 forcing (Barnes and Polvani 2015). However, a large spread is seen between models, suggesting structural uncertainty in the relative importance of the different processes outlined above. In addition, the spread in jet responses remains large within single-model ensembles, pointing to the importance of internal variability, which is in general larger than the forced response.

2.1.1.1 Tug-of-war

Considering the zonal-mean circulation, Shaw (2019) classified jet shift mechanisms according to their thermodynamic starting points (independent of the midlatitude circulation). Implicated in many of these mechanisms are changes in the meridional temperature gradient $\frac{\partial T}{\partial y}$ across the jet. On timescales longer than a few days, the extratropical atmosphere obeys thermal wind balance:

$$f\frac{\partial u_g}{\partial p} = \frac{R}{p}\frac{\partial T}{\partial y}, \tag{2.1}$$

where u_g is the geostrophic zonal wind, p is the pressure, R is the ideal gas constant, and f is the Coriolis parameter (Vallis 2017). Changes in the latitude of maximum $\frac{\partial T}{\partial y}$ are thus associated with shifts in the latitude of the jet. Dynamically, $\frac{\partial T}{\partial y}$ affects baroclinicity and hence eddy growth, with subsequent changes in the associated momentum fluxes feeding back onto the mean flow. The temperature gradient can also alter the refractive properties of the atmosphere, and hence affect momentum fluxes by changing the propagation of Rossby waves.

The temperature response in a strongly forced CMIP5 model is shown in Figure 2.1. A robust feature of coupled model projections in response to anthropogenic forcing is amplified warming of the tropical upper troposphere, shown for a CMIP5 model in Figure 2.1 and discussed further in Section 2.2. By itself, this will increase $\frac{\partial T}{\partial y}$ at upper levels, consistent with a poleward shift of the jet (Butler et al. 2010; Rivière 2011). An additional contribution is made by the stratosphere, which cools in response to anthropogenic forcing through radiative adjustment (Manabe and Wetherald 1967; Manabe and Wetherald 1975). Cooling of the polar lower stratosphere strengthens $\frac{\partial T}{\partial y}$ in the upper troposphere/lower stratosphere, independently of the tropical warming, and thus leads to a poleward jet shift in models (Butler et al. 2010). Over the historical period, stratospheric cooling exerts an additional effect through ozone depletion, leading to cooling of the SH polar stratosphere and a strengthened $\frac{\partial T}{\partial y}$ (Thompson and Solomon 2002; Gillett and Thompson 2003). An important hallmark of the ozone-forced response in the upper troposphere/lower stratosphere is its distinct seasonality, peaking in austral spring and early summer shortly after photochemical ozone depletion is maximised (Gillett and Thompson 2003; Son et al. 2010; Seviour et al. 2017).

Although $\frac{\partial T}{\partial y}$ strengthens at upper levels in response to anthropogenic forcing, an oppositely signed change is seen in the NH near the surface (Figure 2.1). Models suggest that the Arctic warming response to sea ice loss is strongest near the surface, weakening the meridional temperature gradient and potentially leading to an equatorward shift of the jet (Smith et al. 2022). These competing effects can be thought of as exerting a tug of war on the jet (Screen et al. 2018a; Screen et al. 2022). Confidence in the ultimate fate of the jet is therefore conditional on our understanding of the forced thermal response in these regions (Shaw 2019), with doubts remaining in both cases. Trends in the tropical upper troposphere and in the Arctic are discussed in Sections 2.2 and 2.3 respectively.

Finally, it is important to note that thermal wind balance is not a prognostic equation; by itself, it does not imply that changes in the temperature field cause changes in the jet. Changes in the midlatitudes are likely to themselves influence the climate of both lower latitudes, e.g. through modulating the strength of the tropical Hadley Cells (see Section 2.5), and higher latitudes (e.g. through decreased poleward energy transport by the circulation (Graversen 2006; Graversen et al. 2008; Screen et al. 2012). Separating causality in observations or in a single, fully coupled model run is therefore difficult, with progress largely made through targeted model experiments (Barnes and Screen 2015). However, though modulated by changes in

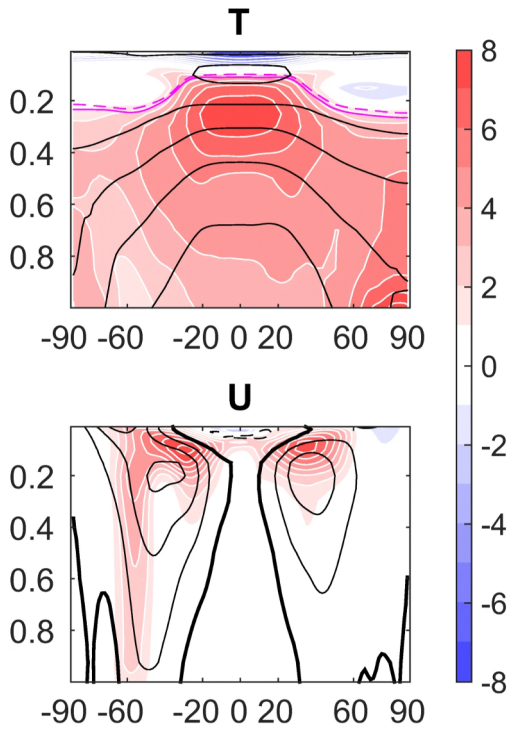


Figure 2.1: **Zonal-mean temperature and zonal wind response under RCP8.5 forcing.** Change in zonal-mean temperature (top, shading with contour interval 1K) and zonal wind (bottom, shading with contour interval 1m/s), between RCP8.5 (2080–2099) minus historical (1986–2005) simulations using the CMIP5 ECHAM6 model (Stevens et al. 2013). Solid contours show the model climatology, with spacing 20K to 300K for temperature and 10m/s for the winds. Pink lines in the top panel denote the tropopause in climatological (solid) and increased CO2 (dashed) runs. From Shaw (2019).

the midlatitudes, both the tropical and Arctic responses are thought to arise from independent processes, justifying their use as thermodynamic starting points (Shaw 2019).

2.1.2 Observed trends

2.1.2.1 Zonal-mean trends

Changes over the satellite era are also poorly constrained in models, and vary seasonally and by hemisphere. The vertical axis in Figure 2.2 shows trends in zonal indices, which measure the meridional jet shift in each hemisphere (see the caption). Wind trends in the CMIP6 ensemble are varied, with trends of both sign possible in each season and hemisphere.

By contrast, poleward trends in the observed jets are increasingly clear and ro-

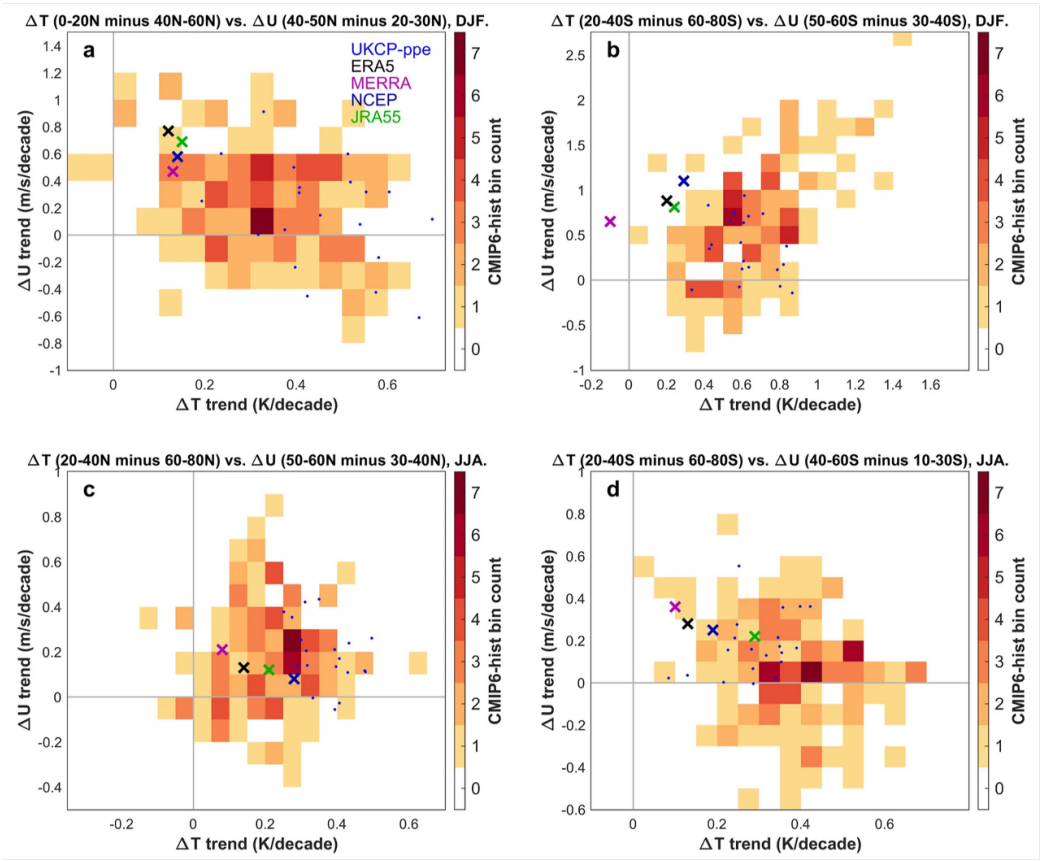


Figure 2.2: **Trends in zonal wind indices at 500hPa plotted against trends in the 250hPa meridional temperature gradient.** Crosses indicate reanalysis trends (index in top left panel). Blue dots indicate members of the UKCP-PPE20 perturbed parameter ensemble (Yamazaki et al. 2021). The 2D shaded histogram shows the frequency of trends in the CMIP6 ensemble. Trends are calculated over 1979-2014 for CMIP6, and over 1979-2019 otherwise. From Woollings et al. (2023).

bust (coloured crosses in Figure 2.2, and other figures in Woollings et al. (2023)). Though the reanalysis wind trends are consistent with those in the CMIP6 ensemble, in many cases they are towards the upper edge of the range simulated by models. The exception is NH summer (panel c), where the observed shift is weak and well within the model envelope. On the other hand, a much stronger poleward shift is seen in austral summer, consistent with the ozone-forced reduction in upper-level $\frac{\partial T}{\partial y}$ over the satellite era. Some evidence has emerged of a reversal of the poleward forced trend as stratospheric ozone recovers (Banerjee et al. 2020), though the trend over the whole period still appears to be dominated by depletion.

2.1.2.2 Regional trends

Averaging zonally risks obscuring regional features of the dynamical trends (Manney and Hegglin 2018). However, evidence is also emerging of strong regional trends, including a robust poleward shift of the winter North Pacific jet (Keel et al. 2024). In the Atlantic sector, changes in jet strength and location are closely related to the North Atlantic Oscillation (NAO) pressure dipole (Woollings et al. 2010). The North Atlantic is a region of substantial variability on seasonal through to decadal timescales (Woollings et al. 2015). Several studies have suggested that models underestimate the true decadal variability of the North Atlantic (Bracegirdle et al. 2018; Bracegirdle 2022 and hence underestimate the range of possible multidecadal trends (Eade et al. 2022). For example, the positive trend in the winter NAO between the 1960-1990s is larger than trends seen in models (Eade et al. 2022). More recently, Blackport and Fyfe (2022) found positive trends in the NAO and North Atlantic winter jet speed over a longer period, from 1951-2020. Over this 70 year period, they showed that the observed trends in jet speed and the NAO lie outside the range of trends simulated by the CMIP6 ensemble. It is worth noting that this time period includes the 1960-90s strengthening trend, and the early decades appear to be important for the discrepancy; for trends over the satellite era, the jet speed trend is consistent with the CMIP6 ensemble (Blackport and Fyfe 2022, their Fig. S5).

2.1.3 Are models failing to capture observed trends?

Figure 2.2 shows that the observed poleward jet shifts in the zonal mean are consistent with the range of trends seen in CMIP6 models (Woollings et al. 2023). However, compared to the coupled models, reanalyses show relatively large wind trends given their modest changes in upper tropospheric $\frac{\partial T}{\partial y}$. In Figure 2.2, this corresponds to being on the top left of the model envelope. Woollings et al. (2023) suggest that this is consistent with a greater sensitivity of the observed circulation to tropical warming than in models, noting reanalysis trends towards increased poleward eddy heat and momentum fluxes in the subtropical upper troposphere. On interannual timescales, La Nina events are similarly characterised by weak upper level $\frac{\partial T}{\partial y}$, and by poleward fluxes of heat and momentum by atmospheric eddies triggered by zonal contrasts in tropical heating (Lee 2014). This raises the possibility that the stronger poleward jet shifts in reanalyses may be related to the long-term strengthening of the zonal SST gradient in the tropical Pacific, in contrast to the more uniform warming typical of coupled models (Seager et al. 2019; see Section 2.4).

While observed trends appear consistent with a dominant role for the tropics, evidence is emerging that the Arctic influence on midlatitude circulation may be underestimated in models. Idealised modelling studies suggest that Arctic warming can lead to an equatorward jet shift, but that this response is weak compared to internal variability (Barnes and Screen 2015; Smith et al. 2022). However, Smith et al. (2022) also find that models with stronger eddy feedback show a greater sensitivity to Arctic forcing. Since the observed eddy feedback strength is underestimated by the PAMIP models in Smith et al. (2022), then the modelled circulation response to Arctic warming may also be too weak. Screen et al. (2022) develop an emergent constraint based on the strength of the observed eddy feedback. Applying this to model projections, they show that, in models with more realistic eddy feedback, the Arctic response overcomes the effect of future ocean warming (with the latter being closely related to tropical warming). However, Saffin et al. (2024) argue that the eddy feedback parameter used in these studies is not independent of the circulation, and that estimates of the feedback strength using an alternative metric show that models and reanalyses are consistent. This tension represents considerable uncertainty in the fidelity of modelled trends in the midlatitude circulation: if models do underestimate the response to Arctic warming, then the observed poleward shift would be even further towards the edge of the model envelope. Possible impacts of Arctic warming on the midlatitudes are discussed further in Section 2.3.4.

2.1.4 Summary and outstanding questions

Observed trends in the jets are on the extreme edge of model simulations, which show a wide range of jet shifts in the present and future climate. Although many processes affect the jets, an important factor in their future is likely to be the tug-of-war between warming of the tropics and the poles, manifest as changes in $\frac{\partial T}{\partial y}$. With only a single observed realisation, it is difficult to know whether observed trends are simply unusual or indicate a shared deficiency or insensitivity in models. In addition to process-based studies of the individual mechanisms, there is a need to quantify uncertainty in jet trends in our most realistic models.

2.2 Tropical tropospheric warming

2.2.1 Response to forcing

A robust feature of coupled models since the experiments of Manabe and Wetherald (1980) is amplified warming of the tropical upper troposphere (TUT) relative to lower levels. This behaviour in models is generally understood through moist adiabatic adjustment theory. In tropical regions of intense convective updrafts, the temperature profile approximately follows the moist adiabatic lapse rate (MALR). Though the updrafts occupy only a small fraction of the total area, they are responsible for setting the lapse rate of the entire tropics. At low latitudes, the atmosphere is unable to sustain large horizontal temperature gradients, as these are quickly damped by gravity waves (Bretherton and Smolarkiewicz 1989). This is encapsulated in the weak temperature gradient approximation (WTG) (Sobel and Bretherton 2000; Sobel et al. 2001). Through the WTG, the entire tropical troposphere lies along a moist adiabat determined by the regions of strongest convection.

The MALR is a function of the saturation vapour pressure, itself a strong function of temperature T, and weakens in a warming atmosphere. This can be understood through the Clausius-Clapeyron relation for the saturation vapour pressure, e_s, written here in ideal gas form:

$$\frac{de_s}{dT} = \frac{Le_s}{R^v T^2} \tag{2.2}$$

where L is the specific latent heat of vaporisation in Jkg^{-1} and R^v is the specific gas constant of water vapour (Vallis 2017; Schneider et al. 2010). For small changes in temperature, the fractional change in saturation vapour pressure is thus:

$$\frac{\delta e_s}{e_s} \approx \frac{L}{R^v T^2}\, \delta T. \tag{2.3}$$

Using values characteristic of the present-day lower troposphere, (2.3) yields an increase in saturation vapour pressure of approximately 7% per degree of warming (Held and Soden 2006). This means that, under warmer conditions, a saturated air parcel contains more water vapour; under ascent, the same parcel would therefore release more latent heat at upper levels, partially offsetting the adiabatic cooling associated with convective ascent. Assuming the tropical troposphere stays approximately moist adiabatic, then the warming response is amplified with height.

2.2.2 Observed trends

Despite the robust signal in coupled models, there has long been tension between trends in models and observations in the TUT (Thorne et al. 2011). The discrepancy first arose in the early 1990s, when the first decade of satellite temperature records failed to show a tropospheric warming trend (Spencer and Christy 1990). Observed trends remained cooler than in models; for example, only a single radiosonde dataset (Santer et al. 2005), and no satellite datasets (Fu et al. 2011), were consistent with the amplified upper tropospheric warming in CMIP3 models. Since then, the discrepancy has persisted through subsequent CMIP generations (Mitchell et al. 2013; Mitchell et al. 2020). As a result, IPCC AR5 has *low confidence* in TUT trends prior to 2001 (Gulev et al. 2021).

2.2.3 Explanations of the trend discrepancy

2.2.3.1 Internal variability

Several explanations have been proposed of the model-observation trend discrepancy. One possibility is differences in the phasing of internal decadal variability between coupled models and the observed system, as was shown to explain the slowdown in global surface warming during the early 21st century (Kosaka and Xie 2013.) Similar arguments have been made for the TUT trend, using both uncoupled and coupled models. Firstly, models forced with observed evolution of SST have been shown to have improved TUT trends compared to their coupled counterparts (Kamae et al. 2015; Po-Chedley and Fu 2012; Mitchell et al. 2013; Mitchell et al. 2020). In coupled models, members with cooler TUT trends (i.e. closer to the observed trend) are associated with a PDO-negative like pattern, with cooling of the East Pacific as in Kosaka and Xie (2013) (Kamae et al. 2015; Po-Chedley et al. 2021).

Other studies, while noting the importance of internal variability, suggest that variability cannot explain the discrepancy entirely (Suárez-Gutiérrez et al. 2017; Mitchell et al. 2020; Keil et al. 2021; Chung et al. 2024). Both Suárez-Gutiérrez et al. (2017) and Mitchell et al. (2020) (hereafter SG17 and M20) show that models struggle to reproduce observed radiosonde trends over the final decades of the 20th century (1979-1997 in M20, and 1986-2005 in SG17), even when forced by observed SSTs in M20. CMIP6 AMIP runs also warm more than satellite observations over 1979-2014 (Chung et al. 2024). However, radiosonde trends since 1997 are well captured by AMIP runs (M20), potentially pointing to the existence of "period-dependent

mechanisms or biases rather than systematic model flaws" for the latter decades of the century (SG17).

2.2.3.2 Observational uncertainty

Observational uncertainty is considerable in the TUT. Structural uncertainty arises from the decisions taken by each research team in processing the raw radiosonde or satellite data: in bias correction (Fu and Johanson 2005), the homogenization of sparse data (Zhou et al. 2021), and in the merging of records from different instruments (Santer et al. 2021). The result is substantial spread in temperature trends between different observational datasets throughout the depth of the troposphere (Thorne 2008; Steiner et al. 2020.

As well as uncertainty in tropospheric measurements, SST trends can also vary significantly between datasets. In an important study, Flannaghan et al. (2014) (hereafter F14) showed that differing TUT trends in two different AMIP ensembles were explained by differences in the trends of their SST boundary conditions (each being forced by a different SST dataset). The tropical surface is coupled to the free troposphere through deep convection; motivated by this, F14 defined a precipitation-weighted SST index $pSST$:

$$pSST = \frac{\langle T_s P \rangle}{\langle P \rangle}, \tag{2.4}$$

where P is the precipitation, T_s the SST, and $\langle \cdot \rangle$ denotes an average over tropical oceans. F14 showed that the TUT trend was most sensitive to the SST trend in strongly convecting regions, as highlighted by $pSST$. In light of the differences between SST datasets, F14 argued that "one cannot conclude that atmospheric general circulation models have systematic biases in the tropical temperature trend profile" (Flannaghan et al. 2014).

2.2.3.3 Model physics

A final possibility is that deficiencies in model physics explain part of the discrepancy. Keil et al. (2021) show that TUT warming is overestimated in all CMIP6 AMIP models relative to most radiosonde datasets, as in Mitchell et al. (2020). They further show that varying the entrainment parameter in a global atmospheric model can reproduce the spread in CMIP6 TUT trends. They therefore argue that at least part of the TUT discrepancy is due to choices made in the parametrisation of convection in climate models.

2.2.4 Summary and outstanding questions

Uncertainty in the observed UT warming is substantial, particularly in the first half of the satellite era (Gulev et al. 2021). Nevertheless, given the observations available, most models appear to overpredict the amplification of UT warming compared to most observational datasets (Po-Chedley et al. 2021; Mitchell et al. 2020; Keil et al. 2021). The discrepancy is particularly large in coupled models, but is not entirely removed even in models forced with observed SSTs (Mitchell et al. 2020; Keil et al. 2021). This suggests that, although SST variability clearly contributes to the discrepancy (Po-Chedley et al. 2022), other factors (such as deficiencies in model physics) may also contribute (Keil et al. 2021; Keil 2023). However, even the observed SSTs are uncertain, and differences between SST forcing datasets can lead to non-negligible differences in model TUT trends (Flannaghan et al. 2014; Fueglistaler et al. 2015). Studies of TUT trends should therefore account for the underlying SST trend directly, ideally accounting for differences in the pattern of atmosphere-ocean coupling using the $pSST$ method (Flannaghan et al. 2014).

2.3 Arctic amplification

2.3.1 Mechanisms of amplification

A robust feature of climate models is amplified warming of the Arctic relative to the rest of the globe (Manabe and Wetherald 1975). Many mechanisms have been suggested to contribute to Arctic amplification (AA). These include processes local to the Arctic, such as sea ice loss (and associated ocean-atmosphere heat fluxes; Screen and Simmonds 2010a; Screen and Simmonds 2010b) and Planck and lapse rate feedbacks (Pithan and Mauritsen 2014). Remote poleward energy transport also contributes, both in the atmosphere (Graversen 2006; Graversen et al. 2008; Lee 2014; Woods and Caballero 2016) and the ocean (Xu et al. 2024), and is tightly coupled with local processes (Previdi et al. 2021). A recent review of dynamical AA mechanisms is provided by Previdi et al. (2021). Finally, a variety of AA metrics have been used (Davy et al. 2018), with most metrics using some ratio of linear trends in the Arctic relative to global; AA corresponds to a ratio greater than 1.

2.3.2 Observed trends

As well as being a feature of models, strong AA has been observed over the satellite era. The IPCC concluded that the Arctic has *likely* warmed more than twice as much

as the global average over the last two decades(i.e. AA > 2; Intergovernmental Panel on Climate Change (IPCC) 2022), while AA factors of between 2 to 4 or more have been reported over parts of the satellite era (Rantanen et al. 2022; Chylek et al. 2022; Hay et al. 2024). These differing values reflect the sensitivity of AA to the choice of dataset, amplification metric, and time period (Davy et al. 2018; Hay et al. 2024). The rate of AA also varies seasonally, peaking in winter in part due to the strong ocean-atmosphere thermal gradient and associated heat fluxes (Screen and Simmonds 2010a). In light of these subtleties, differing conclusions can be reached regarding the ability of models to simulate observed AA.

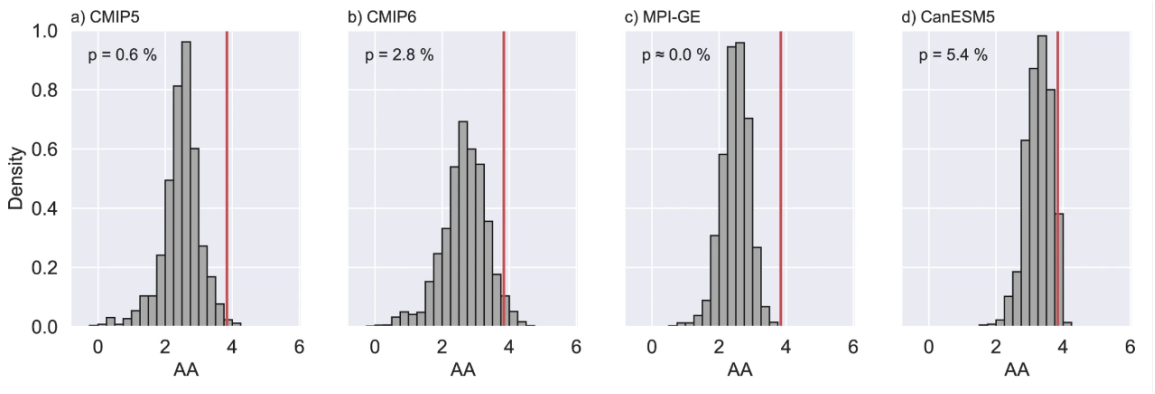


Figure 2.3: **AA distributions in CMIP models.** All possible 43-year AA ratios between 1970-2040 in a range of model ensembles: (a) CMIP5, (b) CMIP6, (c) the 100-member MPI-GE ensemble (Maher et al. 2019), and (d) the 50-member CanESM5 (Swart et al. 2019). In (a), one member is used from each CMIP5 model. In (b), all available members are used from each CMIP6 model, with the exception of CanESM5, which is excluded. CanESM5 is plotted separately in (d) to highlight the role of internal variability in a single-model large ensemble. Model runs follow historical forcing until 2005 (CMIP5 and MPI-GE) or 2014 (CMIP6 and CanESM5), then RCP4.5 (CMIP5 and MPI-GE) or SSP2-4.5 (CMIP6 and CanESM5) thereafter. Observed AA over 1979-2021 (red line) is calculated using an average of trends in 3 surface temperature datasets and the ERA5 reanalysis, with the Arctic defined as all longitudes poleward of 66.5°N. From Rantanen et al. (2022).

In the ensemble mean, CMIP6 models can generally reproduce observed Arctic warming (the numerator of the AA ratio), but tend to overestimate the global temperature trend over the satellite era – leading to an underestimate of AA (Hahn et al. 2021; Sweeney et al. 2023; Chylek et al. 2024). When the full range of model trends is considered, observed AA is generally within the CMIP6 range, although is still towards the upper edge of the model envelope (Ye and Messori 2021; Rantanen et al. 2022; Hay et al. 2024). Shared model biases could explain part of this under-

estimation – for example, Xu et al. (2024) find a moderate increase in oceanic heat transport into the Arctic in a coupled GCM with a very high-resolution ocean model (0.1° grid). Shared biases would not explain the wide AA spread in models, however, which we turn to next.

2.3.3 Internal variability

The wide spread in model AA suggests that, while there is a strong forced component to AA, internal variability is also an important factor which has received less attention in the literature. One important consideration is the short time period over which observed AA is typically calculated. Hay et al. (2024) partition AA uncertainty in CMIP6 projections into model, scenario, and irreducible internal uncertainty following Hawkins and Sutton (2009). They find that the largest uncertainty in 30-year AA ratios throughout the 21st century is due to internal variability. This supports studies which have found a wide range of satellite-era Arctic trends in single-model large ensembles (e.g. Rantanen et al. (2022)), suggesting a possible role for internal variability in modulating recent AA.

How could internal variability have influenced AA? Coupled modes of variability can affect Arctic temperatures directly – for example, previous work has connected multidecadal variations in the rate of Arctic warming to oceanic variability in the Atlantic (Chylek et al. 2009; Mahajan et al. 2011). Variability that affects the global temperature trend also affects AA, as the denominator of the AA ratio. The well-established slowdown of global warming associated with Pacific decadal variability (Kosaka and Xie 2013; Trenberth and Fasullo 2013) thus contributed to a larger AA factor in recent years (Sweeney et al. 2023; Zhou et al. 2024).

In addition to large-scale coupled variability, evidence is emerging of the role of atmospheric variability in the observed AA. Recently, Zhou et al. (2024) identified a mode of atmospheric variability internal to the Arctic which has contributed to the enhanced Arctic warming. The satellite era coincides with a transition from the negative to positive phase of this mode, corresponding to a low-level anticyclonic anomaly over northern Eurasia. This circulation anomaly is associated with warming and moistening of the Barents-Kara seas region and cooling of the Eurasian continental interior, discussed further in the next section. The Arctic may be particularly affected by atmospheric variability due to its small area, with moisture transport by a small number of synoptic eddies able to exert a strong influence on Arctic climate variability (Woods and Caballero 2016).

In summary, AA is a robust feature of model simulations, although the observed amplification is on the edge of the range simulated by models. Coupled variability has contributed to both the numerator and denominator of AA. Additionally, intrinsic atmospheric variability may also have contributed to the observed amplification, through an anticyclonic anomaly over northern Eurasia.

2.3.4 Impacts of AA on the midlatitudes

Some of the interest in AA has concerned its potential impacts on circulation at lower latitudes. Francis and Vavrus (2012) drew attention to an increase in cold weather extremes in the NH midlatitudes. They proposed that Arctic warming was causing a weaker, wavier jet stream, leading to more persistent winter weather over the northern continents. This suggestion provoked immediate and sustained debate (Barnes and Polvani (2013) and Blackport and Screen (2020)).

A useful framework is provided by Barnes and Screen (2015), who distinguish three questions regarding an Arctic influence on the midlatitude circulation: "Can it? Has it? Will it?" Regarding "Has it?", proponents of an observed Arctic influence have stated that "observational studies overwhelmingly support" their position (Cohen et al. 2020). However, Blackport and Screen (2020) convincingly demonstrate that many of the short-term trends identified in these studies (from approximately 1988-2012) have not proved robust over a longer time period (1980-2020), highlighting the difficulty in identifying robust changes amid midlatitude variability. The case for a decisive Arctic influence on observed midlatitude weather has therefore weakened in recent years (Screen et al. 2018a; Blackport and Screen 2020).

On the other hand, evidence has grown that the Arctic *can* influence the midlatitudes, though with varying responses between models (e.g. Screen et al. 2018b). For example, both positive and negative NAO responses to sea ice loss have been reported (Screen et al. 2014; Deser et al. 2015). An effort was made to systematise experimental procedure through the Polar Amplification Model Intercomparison Project (PAMIP, Smith et al. 2019b). This found robust evidence of a weak equatorward shift of the NH zonal winds in response to sea ice loss (Smith et al. 2022), although the response may be underestimated due to weak eddy feedback in models (Screen et al. 2022, and Section 2.1.3). However, even after correcting for the weak eddy feedback using an emergent constraints approach, the adjusted response to AA is small relative to internal variability (Smith et al. 2022), as shown for the non-adjusted case by Ye et al. (2024).

2.3.4.1 Role of coupling

An interesting caveat to the PAMIP results is the models' lack of ocean-atmosphere coupling. Several studies have compared the response to Arctic sea-ice loss in coupled and uncoupled models (Deser et al. 2015; Deser et al. 2016; Tomas et al. 2016; Smith et al. 2017; England et al. 2020). Coupling both (i) strengthens the local response and (ii) induces a response at more remote latitudes from the Arctic. Both effects are a consequence of a warmer global SST response when coupling is enabled: the high latitude heating is redistributed by the atmospheric circulation (and ocean circulation in models with fully dynamical oceans), warming the SSTs at lower latitudes. The SST response in turn warms and moistens the atmosphere above, leading to a stronger circulation response than in uncoupled models (Deser et al. 2015; Deser et al. 2016; England et al. 2020). Although strongest in the NH, the atmospheric response in coupled simulations is global in scale, with amplified warming of the tropical UT partially offsetting the equatorward shift of the eddy-driven jet in both hemispheres. This response is strongest in models with fully dynamical oceans, but notably, a stronger response at low latitudes is also seen in slab ocean models compared to those with prescribed SSTs. This may be important for trends in seasonal hindcasts, introduced in the following chapter: these allow the spreading of Arctic warming to SSTs in adjacent regions, though slower changes in ocean circulation will not affect the trends.

Smith et al. (2017) also find a strong local effect of coupling on the response to sea ice loss. However, they show that this response is strongly conditioned by the model mean state, which differs substantially between their coupled and uncoupled runs; differences in climatological refractive index between middle and high latitudes affect the propagation of planetary waves and hence the midlatitude dynamical response (Smith et al. 2017). Together, these results show that the modelled response to AA depends on both (i) air-sea coupling, and (ii) a realistic climatological background state in the midlatitudes. These requirements support the use of initialised hindcasts for studying the response to AA, as discussed later in this thesis.

2.3.4.2 The role of deep Arctic warming

A final comment on AA linkages to the midlatitudes concerns the vertical structure of Arctic warming. In models, the winter warming response to sea ice loss is mostly confined to the lower troposphere (Ogawa et al. 2018), particularly in the absence of coupling (e.g. Deser et al. 2015). Dynamically, this is because the stable winter

boundary layer inhibits vertical motion, meaning that low-level heating is not primarily balanced by cooling through ascent (Siew et al. 2024). Instead, warming of the mid-upper troposphere appears to be mainly associated with influences from lower latitudes (Graversen et al. 2008; Screen et al. 2012; Ogawa et al. 2018), although sea ice may contribute to a lesser degree via a stratospheric pathway (Xu et al. 2023).

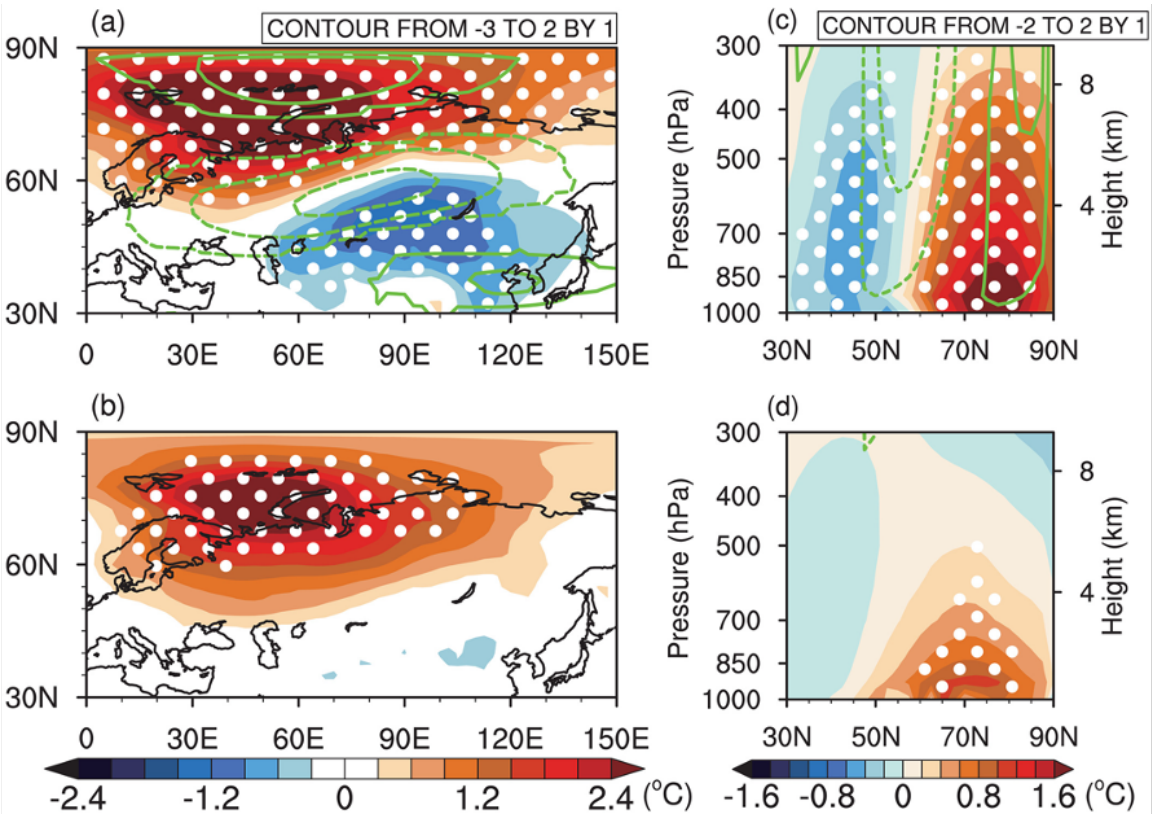


Figure 2.4: **Winter Eurasian cooling and deep Arctic warming.** DJF surface-air temperature anomalies (shading) and 300-hPa zonal wind anomalies (green contours), composited for winters with deep (a) and shallow Arctic warming (b). Data is from an ensemble-mean of 36 CMIP5 historical simulations between 1900–2004. Only the first ensemble member is used from each model. Panels (c) and (d) show the corresponding composites of winter zonal-mean temperature and zonal wind anomalies in the $0° - 150°$E region. Stippling indicates points where at least 90% of the models agree on the sign of the ensemble mean temperature anomaly, and green contours show the same but for the winds. From He et al. (2020).

Deep warming also seems to have a greater impact on the midlatitude circulation, though again the direction of causality is difficult to establish (Barnes and Polvani 2015). For example, He et al. (2020) showed that winters with deep Arctic warming are associated with anomalous Eurasian cooling. Although their study focused on interannual timescales, it supports the prospect of a link between recent Arctic

warming and Eurasian cooling trends – but a link that is not primarily mediated by sea ice, contrary to previous suggestions (e.g. Honda et al. 2009). Two important questions remain from their work: (i) checking whether such a link exists for trends, rather than on interannual timescales; and (ii) quantifying how unusual such a trend might be due to internal variability alone.

2.3.5 Summary and outstanding questions

In summary, amplified Arctic warming appears to be a feature of both climate models and the observed climate system, with a variety of mechanisms potentially contributing (Previdi et al. 2021). The observed amplification is large but not inconsistent with model ranges, although shared model biases may still exist (Xu et al. 2024). A wide range of amplification trends is possible in models through internal variability alone, both over the recent past and in projections. This is consistent with previous links between AA and large-scale coupled variability such as the AMO (Chylek et al. 2009). However, studies are increasingly showing that intrinsic atmospheric variability may also contribute to the AA trend; in particular, atmospheric variability may explain an anticyclonic trend centred over the Urals that also contributed to cooling of the Eurasian interior (Zhou et al. 2024).

Establishing the direction of causality is difficult when considering Arctic-midlatitude linkages. Could the Ural high-pressure trend itself have been caused by Arctic processes? Modelling studies do support a weak influence of the Arctic on midlatitude circulation (Smith et al. 2022), though there is also evidence that models may underestimate or misrepresent the response due to weak eddy feedback, lack of atmosphere-ocean coupling, and unrealistic climatological basic states.

There remains a need to further explore the dynamical response to AA in models that account for some of these shortcomings. It seems unlikely that the Arctic played a leading role in the Eurasian cooling identified by Cohen et al. (2012) and others, given that the observed trend appears possible (though unlikely) in a large uncoupled model ensemble (McCusker et al. 2016). However, contributions from internal variability and Arctic forcing are not mutually exclusive (Outten et al. 2023). It remains to be quantified exactly how unusual this trend is in a realistic coupled model, given the current climate.

2.4 Pacific SST trends

This section discusses trends in the coupled atmosphere-ocean system of the tropical Pacific, a region of high variability on seasonal and interannual timescales through the El Nino-Southern Oscillation (ENSO). Coupled models forced with GHG tend to show a tropical Pacific that warms more in the East Pacific cold tongue (EP) than the West Pacific warm pool (WP) (Collins et al. 2010), sometimes referred to as an "El Nino-like" warming pattern. In contrast, recent observed trends have appeared more "La Nina-like", with a strengthening of the climatological SST gradient. This discrepancy has implications far beyond the tropical Pacific due to its role as a "pacemaker" of global climate (Kosaka and Xie 2016). Linked to changes in the SSTs are changes in the Pacific Walker Circulation (PWC), the zonal overturning cell over the Pacific. While in most models the PWC weakens in response to GHG forcing, recent trends have instead shown a strengthening.

2.4.1 Mechanisms for a weakening SST gradient

Two main physical arguments have been put forward to explain the changes seen in models. The first concerns the different rates of evaporative damping in the WP and EP. Evaporative energy loss of the ocean surface depends on the absolute surface temperature; for a unit increase in SST, warmer regions cool more through evaporation than colder regions. As the WP is climatologically warmer than the EP, it can therefore cool more efficiently in response to a uniform perturbation, weakening the zonal SST gradient (Knutson and Manabe 1995).

The second argument concerns the atmospheric moisture budget. Held and Soden (2006) considered the global moisture budget, where in the time average, precipitation P is balanced by moisture entering the free troposphere through updrafts:

$$P = Mq \tag{2.5}$$

Where q is the specific humidity and M is the global upward convective mass flux at the top of the boundary layer. Separately, the global mean tropospheric heating due to precipitation must be balanced by radiative cooling to space. Linearising (2.5), changes in each term can be constrained as follows:

$$\frac{dM}{M} \approx \frac{dP}{P} - \frac{dq}{q} \tag{2.6}$$

- $\frac{dq}{q} \sim 7\%/\text{K}$ follows from Clausius-Clapeyron (2.2), assuming that changes in relative humidity are small (Held and Soden 2006);

- $\frac{dP}{P}$ is constrained by the change in tropospheric heating, and is likely 2-3%/K (e.g. Allen and Ingram 2002).

Hence $\frac{dM}{M} < 0$, meaning a reduction in global convective mass flux. Although strictly a global constraint, this argument has also been applied to the tropics in isolation due to its dominance of global moist convection. For example, Vecchi et al. (2006) and Vecchi and Soden (2007) argued that (2.6) implied a weakening of the Pacific Walker circulation in response to climate change.

However, a weakening of the global convective mass flux may not entail a weakening of the PWC specifically (Li et al. 2015). Shrestha and Soden (2023) compared satellite-era trends in M in coupled and uncoupled versions of the same model (CESM2/CAM6). They found that M decreased in both models; however, while the PWC weakened in CESM2, it strengthened in the runs forced by observed SSTs. A similar conclusion was reached by Watanabe et al. (2023), who showed that the weakening of the PWC according to (2.6) in prescribed SST runs was outweighed by the strengthening SST gradient. This suggests that the primary factor influencing changes in the Walker circulation is the underlying SSTs (Sandeep et al. 2014; Watanabe et al. 2023; Kosovelj and Zaplotnik 2023).

2.4.2 Observed trends

In contrast to the weakening SST gradient in models, observed trends show a strengthening gradient. Karnauskas et al. (2009) showed a strengthening of the SST gradient in boreal autumn over 1880-2005 in three different SST datasets. Since then, strengthening trends have been shown over the more recent periods of 1950-2018 (Seager et al. 2022) and the satellite era from 1979-2020 (shown in Figure 2.5 from Wills et al. 2022).

As Figure 2.5 shows, SST trends in the tropical and subtropical South Pacific appear to be on the edge of the range simulated in coupled models (Seager et al. 2022; Wills et al. 2022). Similarly, the PWC has strengthened over the satellite era in observations and reanalyses (Zhao and Allen 2019). Next, I review whether the observed trends are consistent with internal variability, or indicate a shared error in the forced response of models.

2.4.2.1 Internal variability

The difference in trend over the Pacific basin (panel (e) of Figure 2.5) resembles the negative phase of the Pacific Decadal Oscillation (PDO) (Zhang et al. 1997). Since the PDO phase has been mostly negative since the 1990s, an aliasing of the PDO could contribute to the strengthening trend over the satellite era (Chung et al. 2019; Watanabe et al. 2021).

However, it seems less plausible that internal variability could explain the longer term discrepancy. Firstly, as noted above, strengthening trends are seen over multi-decadal and centennial timescales, incorporating many PDO cycles (Karnauskas et al. 2009). Over longer timescales, the observed trend is also more distinct from the PDO pattern, with the cooling trend confined more narrowly to the equator (Seager et al. 2022). Secondly, the observed trend is on the edge of the range of trends simulated by climate models (Seager et al. 2022; Wills et al. 2022), suggesting either model error or an underestimation of true variability in the models. Finally, evidence has emerged that trends in some seasonal forecast models also show a weakening of the SST gradient (Shin and Huang 2019; L'Heureux et al. 2022; Beverley et al. 2024b). Since these models are initialised from observed conditions, they inherit the phasing of coupled variability from the observed system. L'Heureux et al. (2022) and Beverley et al. (2024b) therefore argue that internal variability cannot explain the trend difference between the hindcasts and the observed system. However, these studies considered only the ensemble-mean hindcast trends. An outstanding question is the extent to which short-timescale internal variability can lead to a spread of trends within the hindcast ensemble.

2.4.2.2 Forced response

Another possibility is that the observed SST trend pattern arises as a forced response to GHG. For example, warming in the EP is determined by the temperature of upwelling subsurface waters; as these are not in contact with the ocean surface, their warming may be delayed, leading to lagged warming of the EP (Clement et al. 1996; Seager et al. 2019). This so-called ocean thermostat mechanism is present but weak in some CMIP6 models (Heede and Fedorov 2023). However, in those models that do simulate such a response, it appears to be transient, with the long-term equilibrium response to CO2 forcing being a weakening of the gradient (Heede et al. 2020; Heede and Fedorov 2021). Biases in the tropical mean state, including widespread cold biases in the equatorial cold tongue, have also been suggested to contribute to an erroneous

forced response in models (Seager et al. 2019). Intriguingly, coupled simulations that resolve the ocean mesoscale appear to show weaker warming of the EP and Southern Ocean than when run at lower resolution (Yeager et al. 2023), suggesting that model deficiencies may contribute to the discrepancy. Another interesting possibility is that the weakening PWC in models is driven by the enhanced tropical UT warming in models, and hence that deficiencies in convective parametrisations could lead to incorrect responses of the PWC and Pacific SSTs. For further discussion of potential forced mechanisms, see a recent synthesis by Watanabe et al. (2024).

2.4.3 Implications of the tropical Pacific SST trend

The previous sections reviewed evidence for the SST trend discrepancy and two possible origins: either as internal variability, or as a forced response to GHG. In this section, possible implications of the trend discrepancy are discussed for circulation outside the tropical Pacific.

On interannual timescales, variability in the tropical Pacific influences global climate through teleconnections (e.g. Yeh et al. 2018). Decadal variability in the Pacific also modulates global temperature, such as during the global warming "hiatus" in the early 21st century (Kosaka and Xie 2013; England et al. 2014). Consequently, the strengthening of the tropical Pacific SST gradient may be expected to have implications for circulation globally.

Some studies have considered the Pacific influence on remote trends in variability. For example, the observed SST trend has been linked to a poleward shift of atmospheric rivers (AR) in both hemispheres (Li and Ding 2024; Zhao and Knutson 2024) and a northward shift of mesoscale convective system (MCS) activity in the NH (Zhao and Knutson 2024). The Pacific SST trend has also been linked to a weakening of the austral winter storm tracks (Kang et al. 2024) and enhanced NH jet waviness in boreal summer (Sun et al. 2022). It may also have contributed to an increase in Atlantic hurricane activity over the satellite era (Sobel et al. 2023), and models forced with an SST pattern similar to the observed trend project future increases in Atlantic hurricanes and tropical storms (Karnauskas et al. 2023; Zhao and Knutson 2024).

Regionally, the differing Pacific SST trends in models and observations could lead to radically different climate impacts. The Americas are particularly affected due to their proximity to the Pacific and in being downstream in terms of wave propagation. Over California, Zhao and Knutson (2024) find a sharp decrease in the number of AR and MCS days in a model forced with the observed Pacific SST trend, in contrast to the weak trends seen when the model's own SST response is used.

Qiu et al. (2024) used interannual ENSO teleoconnections to develop a statistical adjustment for coupled model precipitation trends over the Americas, in response to the tropical Pacific SST trend. They showed that the linear adjustment brought the model into agreement with observed trends over Mexico and over tropical South America, pointing to the role of the Pacific SST trend in driving the observed changes.

2.4.4 Summary and outstanding questions

In summary, coupled models struggle to capture the observed strengthening of the zonal SST gradient in the tropical Pacific. The satellite era trend includes a large contribution from decadal variability. However, a discrepancy exists even for trends which contain multiple cycles of the PDO, although trends over longer time periods could be sensitive to recently identified biases in early 20th century SST datasets (Sippel et al. 2024). This may point to deficiencies in the transient forced response of the current generation of models, possibly linked to cold biases in the eastern Pacific or teleconnections from the Southern Ocean. The failure of models to reproduce the observed SST trend could contribute to discrepancies with observed trends outside the Pacific, including in the large scale atmospheric circulation.

2.5 Hadley Circulation

This section discusses changes in the Hadley circulation (HC), the zonal-mean meridional overturning circulation in the tropics. Traditionally the HC has been understood through axisymmetric theories, with angular momentum conserved as air moves poleward in the upper branch of the circulation (Held and Hou 1980). Increasingly, the tropical circulation is also seen as coupled with the storm tracks, with the strength and extent of the cells influenced by the extraction of momentum by midlatitude eddies (e.g. Bordoni and Schneider 2010). A recent review by Lionello et al. (2024) provides an accessible introduction to changes in the HC framed through these two perspectives.

2.5.1 Strength of the circulation

2.5.1.1 Weakening response

The response of the HC strength to global warming is complex and differs by hemisphere. In future projections, the NH winter HC weakens in both CMIP5 and CMIP6 under RCP8.5 and SSP585 respectively (Vallis et al. 2015; Chemke and Polvani 2021;

Xia et al. 2020; Wu et al. 2024). The SH response is less clear, with only a weak annual-mean response in CMIP5 (Chemke and Polvani 2021) and CMIP6 (Xia et al. 2020). Given uncertainty in the long-term response in the SH, this sections will focus on changes in the NH HC.

Several arguments have been proposed to explain the future weakening response seen in the NH cell. For example, in Section 2.4.1, a theory is outlined for the weakening of the global convective mass flux in a warmer atmosphere due to moisture budget constraints ((2.5), following Held and Soden (2006) - hereafter HS06). The same argument may be expected to imply a weakening of the HC in response to GHG forcing. As noted previously, Shrestha and Soden (2023) found a weakening of the overturning circulation and reduction of the global convective mass flux in CMIP6 historical simulations. However, the weakening occurred preferentially in the zonally asymmetric circulation, with little change in the zonally symmetric circulation (as already suggested by HS06). Chemke and Polvani (2021) further investigated the HS06 mechanism, as well as five others suggested in the literature to support a weakening of the circulation (Knutson and Manabe 1995; Bony et al. 2013; Merlis 2015; Vallis et al. 2015; Seo et al. 2014). While five of the six mechanisms were consistent with a weakened NH HC, none captured the time evolution of the HC strength by themselves, suggesting a complex interplay of the different mechanisms. The dynamics of future changes in HC strength are discussed by Schneider et al. (2010) and Vallis et al. (2015), as well as in the review of Lionello et al. (2024) mentioned above.

2.5.1.2 Observed trends

There is also uncertainty over whether a weakening of the NH HC has been observed over the historical period. On the one hand, historical model simulations generally show a weakening trend over the satellite era (Held and Soden 2006; Chemke and Polvani 2019; Xia et al. 2020). Chemke and Yuval (2023) further showed that sea level pressure (SLP) datasets also support a weakening trend.

On the other hand, reanalyses have consistently shown a strengthening of the NH HC over the satellite era (Quan et al. 2004; Mitas and Clement 2005; Mitas and Clement 2006; Nguyen et al. 2013; Chemke and Polvani 2019; Zaplotnik et al. 2022). Mitas and Clement (2005) note that their motivation for exploring the HC in reanalyses was the satellite findings of Chen et al. (2002), which also implied an intensification of the HC. The discrepancy is summarised in Figure 2.6, where the reanalysis trends are clearly inconsistent with those in CMIP5 models.

The difference in HC trends has been linked to other trend discrepancies between models and reanalyses. For example, the amplified upper-level warming seen in climate models (Section 2.2) would imply a greater reduction in tropical static stability, which in isolation would weaken the HC (Held and Soden 2006; Mitas and Clement 2006. In future projections, increased static stability is an important driver of the HC weakening (Knutson and Manabe 1995; Vallis et al. 2015; Chemke and Polvani 2021).

To test this, Chemke and Polvani (2019) directly analysed the drivers of satellite era HC trends in CMIP5 models and reanalyses using the Kuo-Eliassen equation. Rather than static stability, they traced the discrepancy to much larger reanalysis trends in tropical precipitation than in models, with the associated latent heating causing a spurious intensification of the HC. Zaplotnik et al. (2022) confirmed the latent heating discrepancy using the ERA5 reanalysis (Hersbach et al. 2020). However, they argued that the HC strengthening could also be partly physical, with the intensity of the circulation strongly correlated with decadal variability of the Atlantic in ERA5.

2.5.2 Summary and outstanding questions

The NH winter HC weakens in model simulations, both in future projections and over the satellite era. Trends in SLP records support a weakening, but reanalyses have consistently shown a strengthening of the circulation over the satellite era. Deficiencies in reanalysis diabatic heating trends appear to contribute to the discrepancy. However, Atlantic coupled variability has also been suggested to have promoted a strengthening of the HC (Zaplotnik et al. 2022).

2.6 Seasonal-to-decadal prediction

The previous sections of this chapter have discussed five emerging trends that will be investigated in this thesis. One of the primary tools that will be used is an ensemble of initialised hindcasts, helping to bridge the divide between the communities of weather and climate science (Randall and Emanuel 2024). This section briefly introduces the porperties of these near-term prediction models that make them an appropriate tool for this purpose.

2.6.1 What is near-term prediction?

The development of dynamical modelling over the 20th century has allowed weather to be skillfully forecast in the medium range. Over the last 2-3 decades, predictive skill in dynamical models has been extended to seasonal and even decadal lead times (e.g. Kushnir et al. 2019). On these timescales model skill does not consist of predicting individual weather systems, but rather the statistics of weather. Seasonal-to-decdal prediction (S2D) is used here to encompass prediction of the climate on timescales ranging from the seasonal to the decadal, often referred to as the near-term.

Weather prediction is primarily an initial value problem, with the boundary conditions – such as external forcing, or the ocean and cryospheric state – changing little over the course of the prediction window. On the other hand, climate projections on multidecadal and centennial timescales are dominated by boundary forcings, with knowledge of the initial conditions being quickly lost. S2D exists between these two regimes, with both boundary and initial conditions being important (Smith et al. 2019a). Earlier generations of S2D models used fixed external forcings, leading to bias development over time (e.g. a cold bias at longer leads, due to the lack of increasing GHG forcing (Doblas-Reyes et al. 2006; Cai et al. 2009). Modern S2D models therefore typically include time-varying external forcings, as in climate models, but differ from climate models in having initial conditions derived from the observed system.

Skill in the near-term is derived from both the response to external forcing (Patterson et al. 2022) and from initial knowledge of the slowly evolving components of the climate (e.g. Collins 2002). On interannual timescales, the greatest source of predictive skill arises through ENSO and its predictable teleconnections to the extratropics (e.g. Doblas-Reyes et al. 2013; Yeh et al. 2018). Aside from ENSO, the stratosphere and sea ice also offer predictive skill from seasonal to interannual lead times (Boer et al. 2016; Tietsche et al. 2014). Decadal modes of ocean variability, such as the Pacific Decadal Oscillation and Atlantic Multidecadal Variability, are potential sources of predictability on longer timescales.

2.6.2 Terminology for initialised and free-running models

Initialisation refers to the process of providing initial conditions to a S2D model, and an initialised model is one which has received this information. Different initialisation strategies exist, with the main distinction being in whether the full model field or only its anomaly is initialised (Smith et al. 2013). In the DePreSys3 model used in this thesis, the full model fields are initialised.

Later chapters will compare trends in an initialised hindcast ensemble to trends in an ensemble of continuous model integrations. Such a continuous integration is referred to in this thesis as a free-running simulation. This definition applies to any typical climate model simulation (e.g. CMIP Eyring et al. 2016). These runs will sometimes be referred to as being "uninitialised", to distinguish from trends in the initialised hindcasts.

2.6.3 New methods using S2D models

S2D models differ positively from climate models in several ways. Firstly, their shorter simulation time means that S2D models are often run at higher resolution and with larger ensemble sizes than is typical of most climate models. In models that use full-field initialisation, mean state biases may also be reduced for a period following the initialisation. Finally, confidence can be built in the model's performance by validating its skill and reliability in a hindcast setting (Weisheimer and Palmer 2014), which may be relevant for some of the physical processes involved in its response to external forcing (Palmer et al. 2008; Palmer and Weisheimer 2018). Given these advantages, and the considerable computational expense in generating such ensembles, increasing attention has been given to developing novel methods to exploit existing S2D datasets.

An example of this is the UNSEEN methodology developed at the Met Office Hadley Centre ("UNprecedented Simulated Extremes using ENsembles", Thompson et al. 2017). The observed state is only a single realisation of a chaotic climate system; UNSEEN leverages the advantages of S2D models outlined above to provide plausible alternative realisations of the system. On seasonal timescales, large ensembles of S2D models have been used to estimate the risk of seasonal extremes (Thompson et al. 2017; Thompson et al. 2018; Kay et al. 2020; Kay et al. 2022). Kelder et al. (2020) and Berghald et al. (2024) also adapted the UNSEEN approach to study longer-term trends in extreme precipitation and heatwaves. In common with previous UNSEEN studies, their approaches were regionally targeted and focused on impact-relevant metrics. As described in this chapter, however, there remains considerable uncertainty concerning many trends in the atmospheric mean state. The following chapters will test whether a S2D ensemble can also provide useful information for placing these larger scale trends in context.

2.7 Research questions

This chapter has reviewed the literature concerning trends in five key features of the climate system, aiming to highlight areas where outstanding questions remain. Not all of these questions can be tackled with the tools in this thesis, with progress in some likely to require targeted experiments. Nevertheless, the hindcasts offer a largely untapped resource for studying trends in the atmospheric mean state, and chapters 3 to 5 will aim to answer the following broad questions:

1. **Are the hindcasts better able to capture observed trends than their free-running counterpart?** For example, do the hindcasts have tropospheric warming trends similar to the free-running model, to observations, or neither?

2. **What range of trends is possible in the hindcasts, and can they be used to provide context for observed trends?** For example, how unusual are observed jet trends relative to trends in the hindcast ensemble?

3. **How are trends in a given feature related to trends elsewhere in the hindcasts?** For example, what impact does the tropical Pacific SST trend have on trends outside the Pacific region?

The above grouping is necessarily broad in order to cover each emerging trend, although interesting questions particular to individual features will also be investigated in the following chapters. Although Chapter 6 does not use the hindcasts, in exploring the connection between circulation and temperature trends, it follows on from Q3 above.

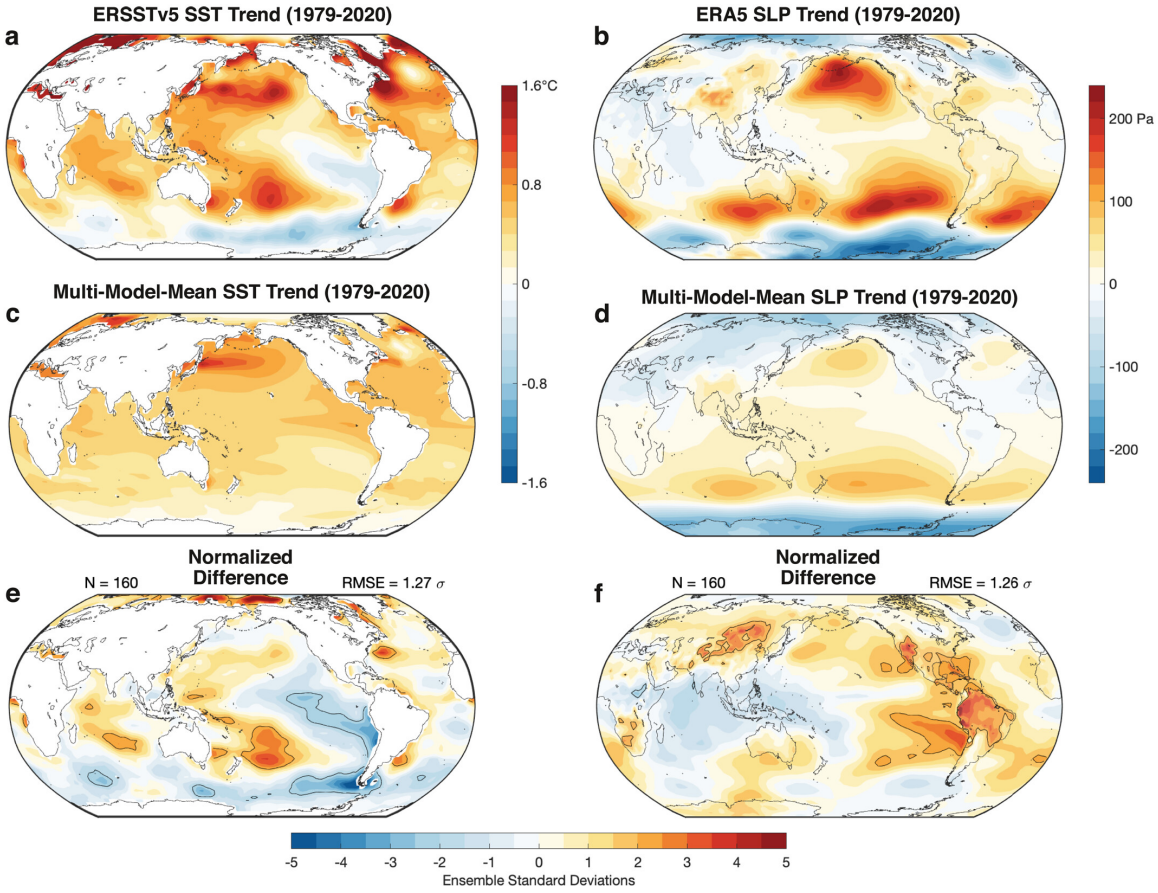


Figure 2.5: **SST and SLP trend discrepancies between models and observations.** Observed 1979-2020 trends (per 41 yr) in annually-averaged ERSSTv5 SST (a, Huang et al. 2017) and ERA5 SLP (b, Hersbach et al. 2020). (c, d): Multi-model mean trends in SST and SLP, respectively, from 16 CMIP5 and CMIP6 large ensembles (LEs). (e, f): Observed SST and SLP trends, respectively, written in units of ensemble standard deviations away from the multi-model means in panels (c, d). Only the first 10 members of each LE are used to ensure equal weighting of models. The black contour demarcates regions of ±2 standard deviation. In panels (e) and (f), the RMS error averaged over ±60° is shown on the top right of the map. From Wills et al. (2022).

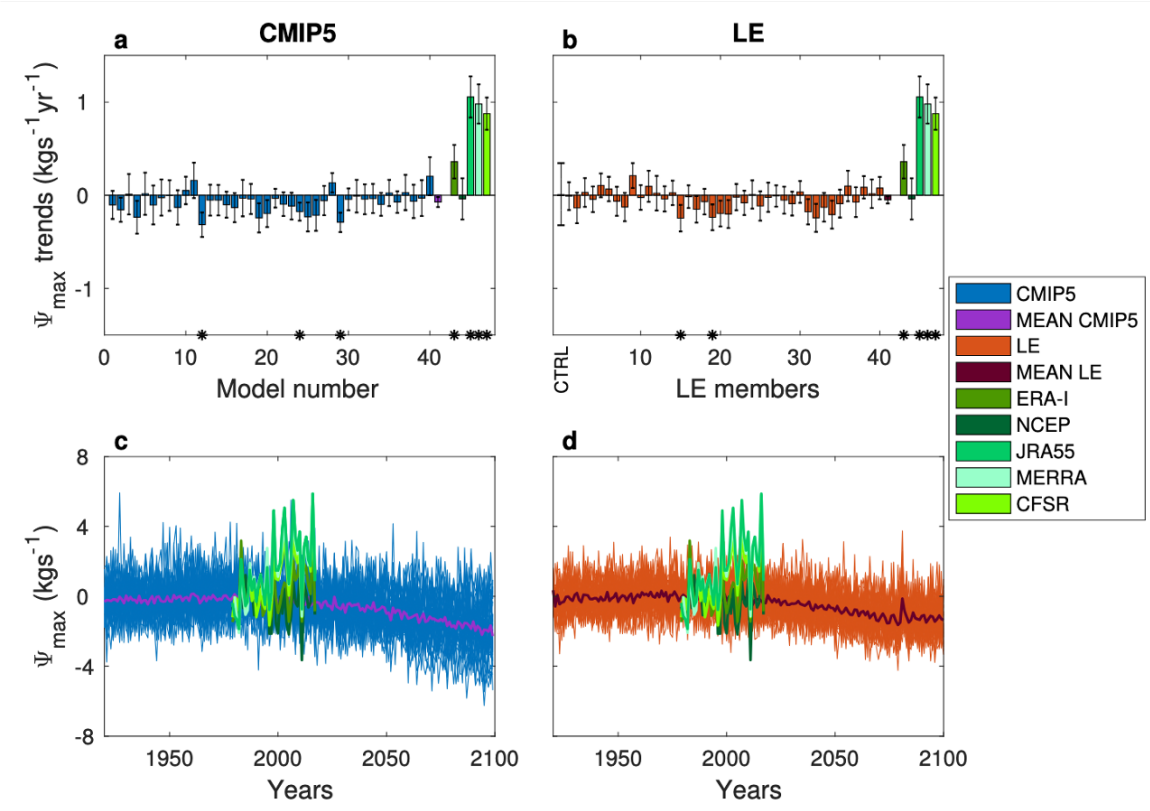


Figure 2.6: **CMIP5 and reanalysis trends in DJF Hadley circulation strength.** (a, b): 1979-2017 trends in DJF NH Hadley Circulation strength, ψ_{max}, in CMIP5 models (blue) and their multimodel mean (purple), and CESM1-LE members (red) and their mean (maroon). Reanalysis trends are in green. In panel (b), the whiskers on the far left show the spread of 39-year trends calculated from the piControl CESM1 simulation (denoted CTRL). (c, d): ψ_{max} anomalies relative to 1979–1989, with the colours as in the panels above. From Chemke and Polvani (2019). Note that this picture is essentially unchanged for the NH cell in the annual-mean (their Fig. 1).

Chapter 3

Zonal-mean trends in the hindcasts

3.1 Introduction

In recent years, some notable discrepancies have emerged between observed trends in the atmospheric mean state and trends in coupled model simulations. As discussed in the previous chapter, these include uncertainty over the rate of warming in the tropical upper troposphere (Thorne et al. 2011; Po-Chedley et al. 2021) and changes in the tropical overturning circulation (Chemke and Polvani 2019; Zaplotnik et al. 2022). In the midlatitudes, reanalysis jet shifts appear to be on the edge of coupled model simulations (Woollings et al. 2023), and models also struggle to capture the observed strengthening of the winter North Atlantic jet speed (Blackport and Fyfe 2022). Finally, while models robustly show amplified Arctic warming over the satellite era and in projections, the observed amplification remains larger than in most simulations (Rantanen et al. 2022; Sweeney et al. 2023).

In light of these trend discrepancies, this chapter explores trends in near-term hindcasts, to bridge the gap between trends in coupled models and reanalyses. In the modelling community, attention has often focused on model trends as a source of skill or error in initialised forecasts (Shin and Huang 2019; L'Heureux et al. 2022; Becker et al. 2022). However, the spread of trends within hindcast ensembles can also provide useful context for trends emerging in observations and reanalyses.

Near-term hindcasts offer a unique tool for studying multidecadal trends in the atmospheric mean state. Figure 3.1 shows two advantages of initialised hindcasts over the equivalent uninitialised model. The hindcast and free-running models in this figure - DePreSys3 and HadGEM3-GC2 - are introduced fully in Section 3.2. Initialisation means that, at short lead times, the phasing of coupled modes of variability in the hindcasts will be similar to the observed system, in contrast to the random variability in uninitialised runs. Figures 3.1c and d show this for the AMV and ENSO,

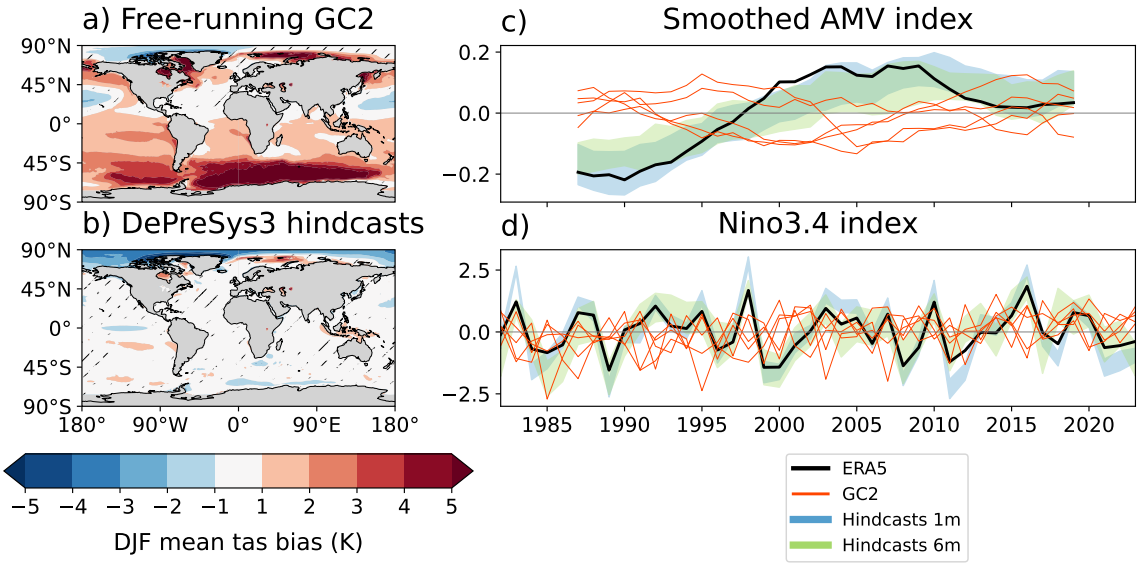


Figure 3.1: **SAT bias and coupled variability in the hindcasts**. Left column: DJF surface-air temperature (SAT) bias in (a) the free-running model (GC2), and (b) the hindcasts at 1 month lead (DePreSys3). The bias is calculated as the time-mean difference between the model and ERA5 (model minus reanalysis). The time-mean is taken over the same period as the trend calculations (DJF 1981/82-2022/23). Hatching indicates where greater than 20% of the ensemble members disagree on the sign of the bias. Right column: timeseries plots of (c) the AMV index (Trenberth and Shea 2006), and (d) the Nino3.4 index (Trenberth 1997), each for DJF. ERA5 is in black and the five free-running members are in orange. Green shading indicates the 5-95% range of values in the hindcast ensemble. A 10-year centred rolling mean is applied to the AMV index.

respectively: the hindcast variability (shading) is in phase with the reanalysis (black line) due to the frequent reinitialisation, whereas the free-running simulations are not (orange lines). In the case of full-field initialisation, as in DePreSys3, coupled biases may also be reduced following initialisation; Figures 3.1 a and b show that mean surface temperature biases remain substantially reduced at lead times of 1-3 months compared to the free-running HadGEM3-GC2 runs.

Similar considerations could motivate the use of uncoupled, prescribed-SST ensembles. However, in contrast to uncoupled models, the DePreSys3 hindcasts retain full coupling during the forecast phase. This is particularly important in the extratropics, where the oceanic influence on the atmosphere is not always directly mediated by SST anomalies (Sutton and Mathieu 2002) and where a lack of direct coupling can lead to underestimates of low frequency variability (Barsugli and Battisti 1998; Mori

et al. 2024). The hindcasts offer a unique tool for studying trends in a coupled model, which is nonetheless constrained to follow large-scale modes of coupled variability.

The ability to validate coupled model performance on shorter timescales is another unique advantage of initialised hindcasts. In recent years, skill in patterns of extratropical circulation has been realised on seasonal and interannual timescales (Dunstone et al. 2016). Validation on shorter timescales builds confidence in the model's representation of some of the processes important for longer term trends (Palmer et al. 2008), further bolstered by the high model resolution and relatively large ensemble size.

Finally, additional motivation for the present study is given by recent work identifying mean temperature and circulation trend errors in a range of seasonal hindcast models (Beverley et al. 2024b). In general, trend discrepancies between models and observations can arise due to errors in the model forced response, the aliasing of internal variability, or both. Since individual members of an initialised hindcast ensemble share the same phasing of coupled variability as the observed system, the potential for internal variability to contribute to trend discrepancies is reduced. However, the extent to which spread in trends can arise even in these tightly constrained models has not yet been quantified. At short lead times, any spread in trends is presumably largely due to intrinsic, seasonal and subseasonal atmospheric variability, although short-timescale coupled processes may also contribute. Following the UNSEEN approach applied on weather and seasonal timescales (Thompson et al. 2017; Thompson et al. 2018; Kay et al. 2020; Kay et al. 2022) and later adapted for trends (Kelder et al. 2020; Berghald et al. 2024), this chapter asks: given the large-scale modes of variability, how different could these trends have been?

This chapter focuses on four of the five emerging trends identified in the previous chapter: jet shifts, tropical tropospheric warming, Arctic amplification (AA), and trends in the Hadley Circulation. The range of model trends in these metrics – for both the DePreSys3 hindcasts and equivalent free-running simulations – is compared to trends in three modern reanalysis products. Which of the reanalysis trends are better captured by the hindcasts than the free-running model? What spread of trends is possible in the hindcasts? And in some cases, what does this mean for trends in other parts of the system?

Details of the models used and the generation of the trend ensemble is in Section 3.2. Tropical warming is discussed in Section 3.3, AA in Section 3.4, jet shifts in Section 3.5, and Hadley circulation trends in Section 3.6. Results are briefly summarised in Section 3.7. The primary focus of this chapter is on results in the zonal-mean;

the spatial pattern of trends is discussed in the following chapter, along with a fuller discussion of the hindcast trends approach (including some reference to the results of this chapter).

3.2 Methods

Given substantial structural uncertainty, previous studies of hindcast trends have pooled prediction models as is common in operational systems (Shin and Huang 2019; L'Heureux et al. 2022; Beverley et al. 2024b). However, this can make it difficult to cleanly compare the hindcast trends with those in uninitialised models due to differences in model resolution and forcings. Instead, we take advantage of the Met Office's seamless prediction approach, by comparing trends in a recent version of the Met Office Decadal Prediction System (Dunstone et al. 2016) to those in a free-running historical simulation using an otherwise identical model configuration. This allows a novel isolation of the impact of initialisation on the range of possible model trends.

3.2.1 Hindcast data

Monthly hindcast data is used from DePreSys3, the third iteration of the Met Office decadal prediction system (Dunstone et al. 2016). DePreSys3 uses the HadGEM3-GC2 coupled model (Williams et al. 2015; Senior et al. 2016), incorporating the Unified Model (UM) atmospheric dynamical core, JULES land model (Best et al. 2011), CICE sea ice model (Rae et al. 2015), and NEMO ocean model (Megann et al. 2014).

Both the UM and JULES use an N216 grid, corresponding to a horizontal resolution of 0.83°x0.56° or approximately 60km in the midlatitudes. The atmosphere has 85 levels in the vertical and an 85km model top with a well-resolved stratosphere. NEMO and CICE both use a tripolar grid with an eddy-permitting horizontal resolution of 0.25° and 75 vertical levels (1m top level).

Initial conditions are derived from an assimilation run in which the model fields are nudged towards full-field analyses (relaxation timescale in brackets): 6-hourly ERA-Interim (Dee et al. 2011) temperature and zonal and meridional winds (6hr); monthly MOSORA (Smith et al. 2015) salinity and temperature in the ocean (10 day), and monthly HadISST (Rayner 2003) sea-ice concentration (1 day). During the assimilation run, non-nudged fields (e.g. snow cover and soil moisture) are free to evolve in response to the nudged parameters. This nudging strategy aims to reduce

the shock that could arise if initialised using instantaneous fields at the launch time. Simulations are launched biannually on the 1st of November and 1st of May for every year since 1960, although the work presented here focuses on the data-rich period since the introduction of satellite observations in 1980.

An ensemble of 40 members for each start date is generated by applying stochastic perturbations to the model physics (Bowler et al. 2009). Simulations last between 13 and 66 months. External forcing in DePreSys3 is time-varying, following CMIP5 historical forcing until 2005 and RCP4.5 thereafter.

3.2.2 Reanalysis data

Trends in the models are compared to those in three modern atmospheric reanalyses: ERA5 (Hersbach et al. 2020), JRA-55 (Kobayashi et al. 2015), and JRA-3Q (Kosaka et al. 2024). Monthly mean data is used for all datasets between June 1981-February 2023. ERA5 is retrieved at 1° horizontal grid resolution and the two JRA reanalyses at 1.25°. Although reanalysis trends can be susceptible to biases or discontinuities (see below), they are used here for dynamical consistency between the thermodynamic and dynamical trends (e.g. in the precipitation-weighted surface temperature described in Section 2.2, (2.4)).

Caution is necessary when interpreting trends in reanalyses due to changes in the assimilated observing systems over time (Bengtsson et al. 2004). In JRA-3Q, different SST forcing datasets are used before and after June 1985: COBE-SST2 before, and MGDSST after (Kosaka et al. 2024). The change in dataset could affect trends that span the changeover date, particularly for variables that are especially sensitive to the lower boundary conditions, such as surface-air temperature or tropospheric temperatures in the tropics. No evidence of a sharp discontinuity in 1985 is found for the variables considered here (not shown). In addition, the JRA-55 dataset is included for comparison; in JRA-55, the COBE-SST2 forcing dataset is used throughout (Kobayashi et al. 2015). The overall similarity between trends in JRA-3Q and JRA-55 builds confidence in the use of JRA-3Q. JRA-3Q sea ice data is not used due to previously documented quality issues (Japan Meteorological Agency 2022).

Kosaka et al. (2024) note that surface-air temperature (TAS) observations assimilated in JRA-55 and JRA-3Q over ocean may be influenced by biased ship observations (Simmons et al. 2004); when calculating global mean TAS in the JRA reanalyses, they therefore use analysis fields over land and the background forecast field over ocean. This paper adopts the same approach for the JRA reanalyses. Without this

important step, the analysis global TAS trends are approximately 20% weaker than in ERA5 (Simmons 2022; not shown).

3.2.3 Free-running model data

The hindcasts are compared to five free-running members of the same HadGEM3-GC2 model with historical forcing (i.e. CMIP5 historical forcing until 2005 and RCP4.5 thereafter). These simulations (denoted "GC2") have identical spatial resolution and external forcing as described above for DePreSys3, differing only in being uninitialised. Members are initialised from different points along the pre-industrial control run in 1850. Though the ensemble size is limited, including different models would dilute the like-for-like comparison which is used to isolate the effect of initialisation on the modelled trends.

3.2.4 Trend analysis

Seasonal mean trends are presented for DJF and JJA. DJF trends are calculated from DJF 1981/82 to DJF 2022/23, and JJA trends are from 1981-2022. Both May and November start dates are exploited to study trends at months 2-4 ("1m lead") and months 8-10 ("6m lead").

Multidecadal trends are constructed by randomly selecting a single member for each start date (i.e. one member out of 40 is chosen each year, and these are stitched together to form a timeseries). Hindcast labelling is arbitrary; no special relationship exists between the i^{th} member for one start date and the i^{th} member for another. This allows a free choice of all 40 members for each start date between 1981-2023. The selection process is repeated 10,000 times to generate a distribution of trends consistent with the hindcast predictions (Kay et al. 2022). In the following sections, "ensemble" is used to refer to this distribution of 10,000 42-year sequences ("members"), rather than to the 40-member ensemble for each forecast run.

Linear trends are calculated using least-squares regression. Significance in the reanalyses is calculated using a two-tailed t-test with null hypothesis of zero trend. The test uses an effective sample size that accounts for lag-1 autocorrelation (Santer et al. 2000). As the trends are calculated here using seasonal mean data, this effect is generally small. To account for multiple testing in the reanalyses, only those regions where the local test statistic passes the false discovery rate criterion (Wilks 2016), with a global control level of $\alpha = 0.05$, are stippled in later figures. DePreSys3 and GC2 output is conservatively regridded to the coarser 1° resolution for comparison

with ERA5. The JRA reanalyses are not regridded as no point-by-point comparison is made with the other datasets. Only data above the surface is used for all pressure-level variables.

3.3 Tropical tropospheric temperatures

The top rows of Figures 3.2 and 3.3 present the zonal-mean temperature trends in ERA5, JRA-55 and the model ensemble-means, for DJF and JJA respectively. The reanalyses show a warming of the troposphere and cooling of the stratosphere in both seasons. Two notable exceptions are seen – a warming trend in the winter Arctic stratosphere, and tropospheric cooling in the southern high latitudes - though the trends in these regions are not statistically significant.

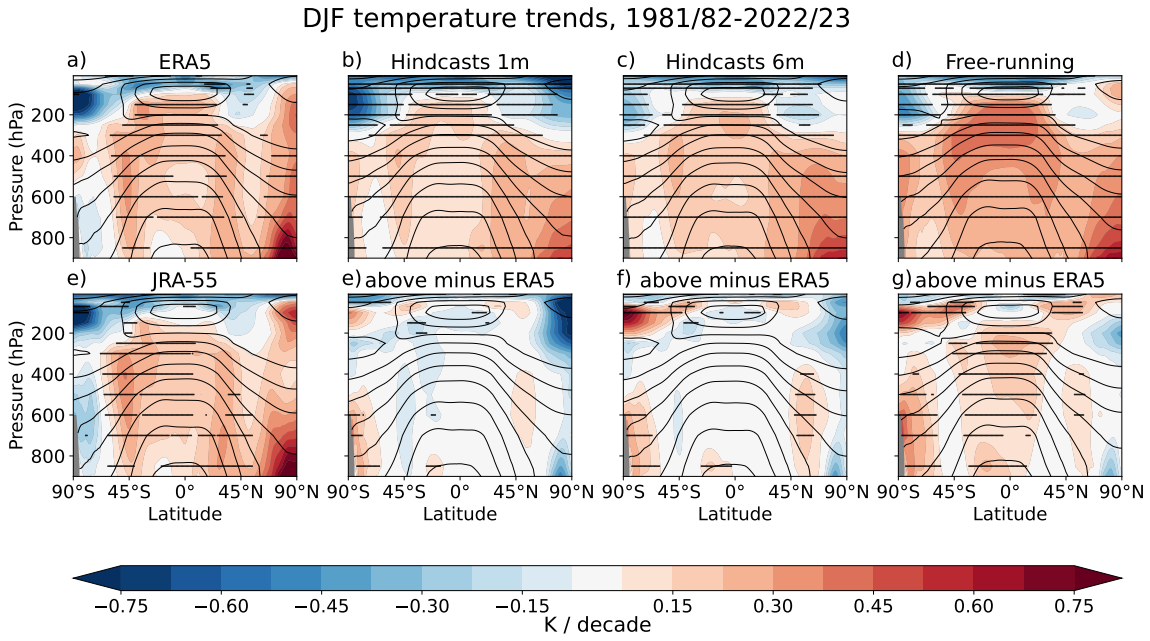


Figure 3.2: **DJF zonal-mean temperature trends.** (Top row) DJF zonal mean temperature trends (1981/82-2022/23) for (a) ERA5; (b)-(c) the hindcast ensemble-mean for 1m and 6m leads, respectively; and (d) the free-running GC2 ensemble-mean. Solid contours show the corresponding climatological values with spacing of 10K. Stippling on (a) indicates a trend significantly different from zero at the 5% level after accounting for multiple testing (see Section 3.2). Stippling on (b)-(d) indicates where at least 80% of the model members agree on the sign of the trend. (Bottom row) differences of the plots above from the ERA5 trend in (a). Stippling on (e)-(g) indicates where the hindcast mean is more than $2S$ away from the ERA5 mean, where S is the standard error on the ERA5 regression.

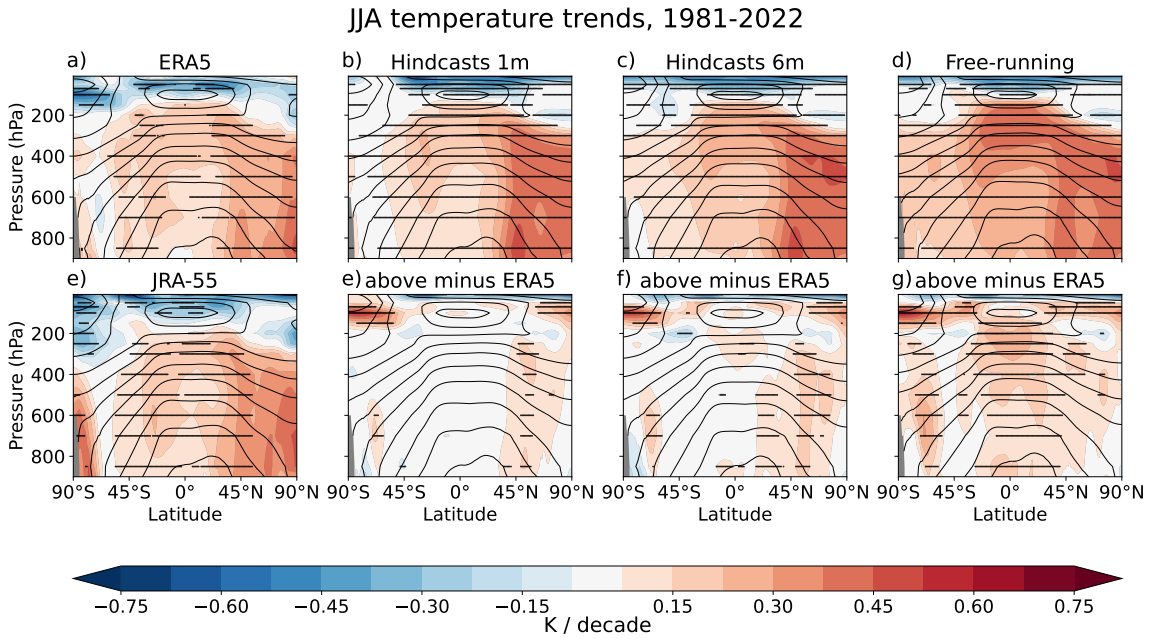


Figure 3.3: **JJA zonal-mean temperature trends.** As in Figure 3.2, but for JJA.

The bottom row of the figures shows the difference in trend between ERA5 and the model ensemble-means. Overall, trends in the free-running GC2 tend to be warmer than in ERA5 throughout most of the atmosphere, but this difference is substantially reduced in the hindcasts at both lead times. The following sections focus on two key features of the trends, as discussed in the previous chapter: tropical tropospheric warming, and Arctic amplification.

3.3.1 Trends in the reanalyses and the free-running model

Figures 3.4a and 3.4b show vertical profiles of the tropical average trend. The entire tropical atmosphere warms faster in GC2 than in the reanalyses. In common with other free-running coupled models, the tropical warming in GC2 strongly peaks in the mid-upper troposphere in both DJF and JJA. This amplified upper-level warming is commonly understood to be a result of adjustment to a moist adiabat in the models (e.g. Santer et al. 2005). The tropical tropospheric warming in the reanalyses is also greatest above 400hPa, although the upper-level trend is approximately half as large as in GC2. The vertical structure of the trends shows little seasonality in either the reanalyses or GC2, and the profiles of each GC2 member are smoother than the reanalyses. The three reanalysis trends differ somewhat, particularly above 300hPa

and to a lesser extent between 400-700hPa, but are much more similar in magnitude to each other than they are to the GC2 trends.

3.3.2 Trends in the hindcasts

Tropical warming is reduced in the hindcasts at both lead times compared to the free-running GC2 (Figures 3.4a and 3.4b). The reanalysis trends in the mid to upper troposphere are broadly within the 5-95% hindcast range at either 1m or 6m lead. Trends in the hindcasts are generally smallest at 1m lead. Even by leads of 6m, however, they remain substantially smaller than the free-running members, indicating that the effect of the initialisation persists beyond seasonal timescales.

Why do the hindcasts better capture tropical temperature trends than GC2? To first order, the diminished tropospheric warming in the hindcasts can be understood as a response to cooler trends at the tropical surface than in the free-running model. The tropical surface is coupled to the free troposphere by deep convection. Following Flannaghan et al. (2014), Fueglistaler et al. (2015) and others, a precipitation-weighted surface temperature is defined:

$$pTAS = \langle T_s \cdot P \rangle / \langle P \rangle \tag{3.1}$$

where T_s is the surface air temperature (tas), P is the precipitation, and $\langle \cdot \rangle$ denotes an average over tropical oceans (previously introduced in Section 2.2, (2.4)).

Figures 3.4c and 3.4d shows the temperature trend at 300hPa, T_{300}, as a function of the $pTAS$ trend. The $pTAS$ trends are smaller in the hindcasts than in GC2, and the spread of T_{300} trends is well explained by the $pTAS$ trend in the hindcasts and in GC2. The model behaves approximately linearly, and the trends at 6m lead neatly span the trends at 1m lead and in GC2. This important result suggests that much of the upper-level difference between the models is due to smaller ocean surface temperature trends in the hindcasts, particularly in convecting regions.

Some spread is seen between the different reanalysis products, particularly in DJF. Each of the reanalyses used here is forced by a unique combination of SST datasets over this time period, as discussed above: Section 3.2). The MOSORA ocean reanalysis used to initalise DePreSys3 is informed by different SST data again. The spread in reanalysis $pTAS$ trends therefore reflects uncertainty in both the underlying SST trends and the simulated patterns of precipitation (Flannaghan et al. 2014; Po-Chedley et al. 2021, their figure S1). Although only 2 of the 6 reanalysis points lies

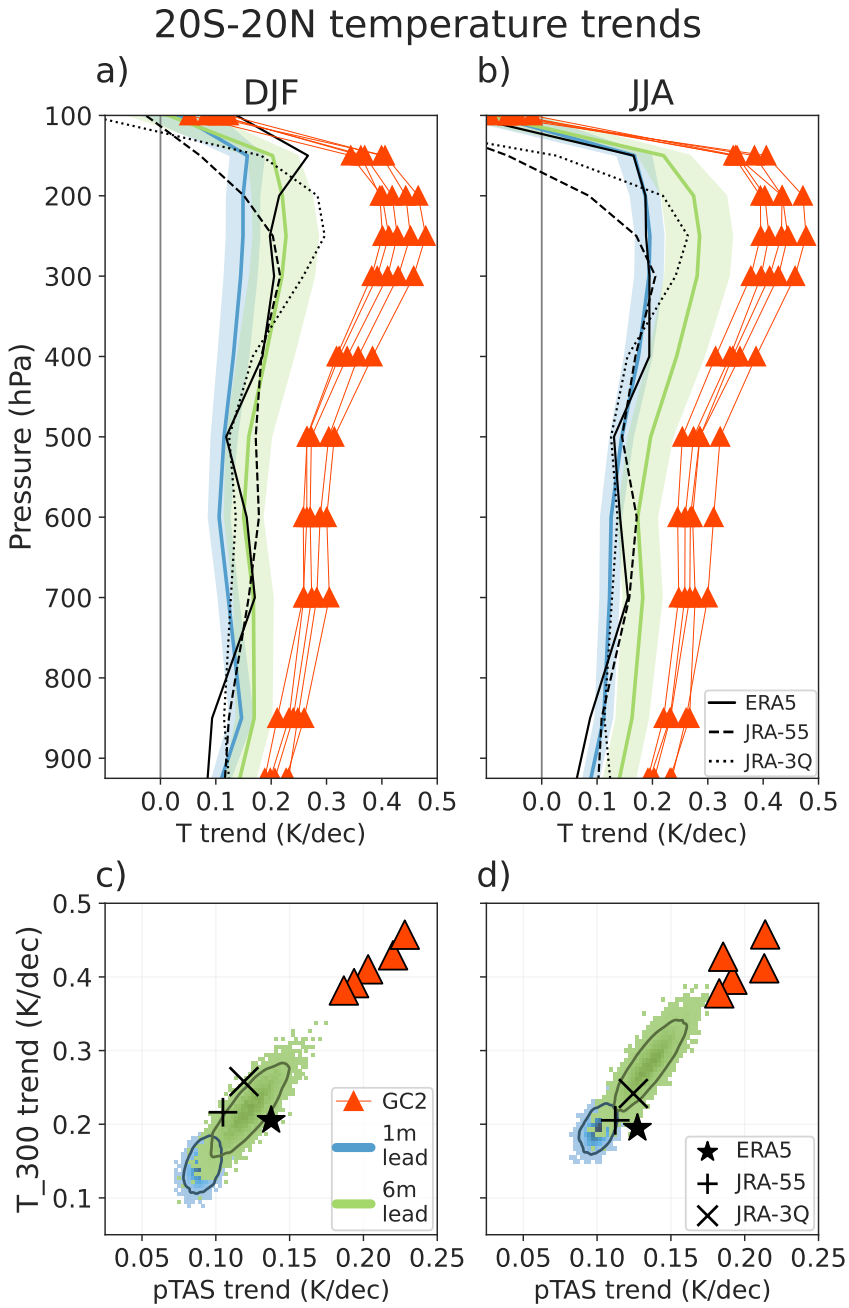


Figure 3.4: **Vertical warming structure in the tropics.** (Top row) Mean temperature trend averaged between ±20° for DJF (a) and JJA (b). The 5-95% ranges of hindcast trends at 1m and 6m lead are shown by the blue and green shading respectively, and the mean hindcast trends are shown by the solid lines of the same colours. Reanalyses are in black and the five free-running GC2 members are in orange. (Bottom row) Trend at 300hPa as a function of the pTAS trend for (c) DJF and (d) JJA (see text). Solid contours enclose 95% of the hindcast trends at each lead time. DJF trends are calculated over 1981/82-2022/23, and JJA trends are calculated over 1981-2022.

within the 5-95% ranges of hindcast trends, this is due to the large spread between the reanalysis products.

In summary, the hindcasts broadly capture the relatively cool reanalysis temperature trends in the tropical mid-upper troposphere. Trends in the 6m lead hindcasts are warmer than in the 1m lead runs. Trends in the free-running GC2 simulations are warmer still, and inconsistent with the reanalyses, although uncertainty in the reanalysis trends remains considerable. Within the model runs, linear scaling theory using $pTAS$ trends explains the spread of T_{300} trends very well, suggesting that the model-reanalysis discrepancy can be attributed to differing SST trends in strongly convecting regions.

3.3.3 Variability and unseen trends in the hindcasts

The range of tropical tropospheric temperature trends is small in the hindcasts, reflecting the primary importance of the underlying surface temperature for tropical temperatures further aloft. The spread in tropical tropospheric trends grows with lead time as the range of surface temperature trends predicted by the hindcasts diverges. However, even at 1m lead we see non-negligible spread in the hindcasts which is not entirely explained by differences in the $pTAS$ trend between members.

Previous work has suggested an important role for internal variability in modelled tropical tropospheric warming. Mitchell et al. (2020) show that the range of tropospheric temperature trends in CMIP6 large ensembles is approximately half as large as the intermodel spread, and that the spread is reduced (but not eliminated) in simulations with prescribed SSTs. Taking the ratio of warming at two levels, then for T_{300}/T_{500} the 5-95% range is 1.26-1.43 for the 1m hindcasts in JJA (not shown), or slightly less than half the annual-mean spread among uninitialised CMIP5 members (O'Gorman and Singh 2013). By removing the confounding factors of differences in model architecture or SST forcing datasets (Fueglistaler et al. 2015), the hindcasts highlight the scope for internal variability on seasonal timescales to affect tropical temperature trends in a coupled model.

3.4 Arctic amplification

3.4.1 Lower tropospheric AA

In the reanalyses and in the GC2 mean, the strongest zonal-mean warming is found at low levels in the wintertime Arctic (Figure 3.2). Although the trend peaks below

700hPa, the entire Arctic troposphere warms faster than the midlatitudes to the south. Arctic amplification (AA) is first defined using a lower tropospheric metric following Barnes and Polvani (2015) (see caption of Figure 3.5), which may be of more significance for circulation at lower latitudes (Barnes and Polvani 2015). Throughout this section, only winter data is used (DJF).

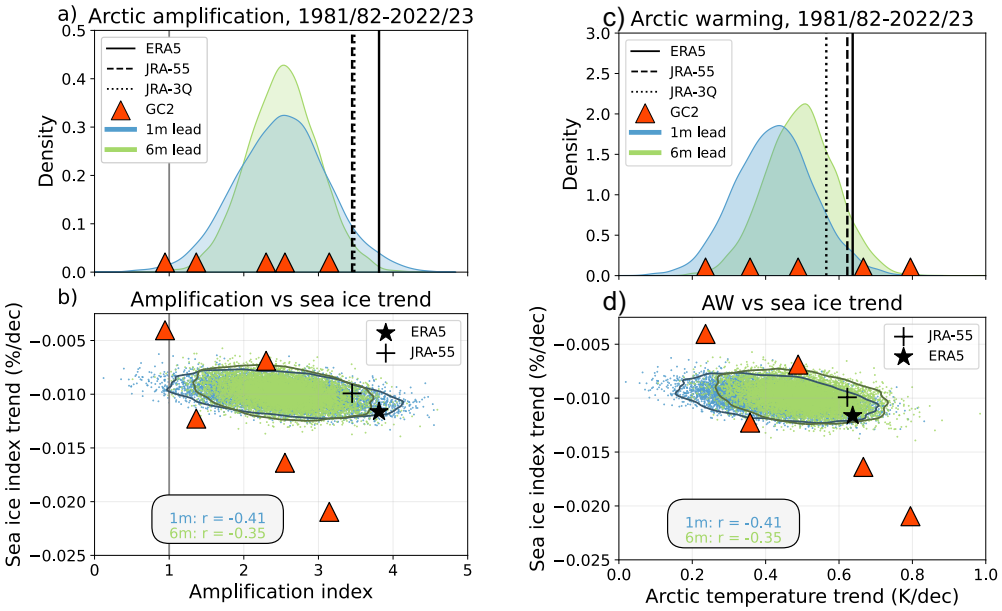


Figure 3.5: **Lower tropospheric AA and Arctic warming.** (a) Arctic amplification (AA) in the lower troposphere, and (b) AA as a function of the sea ice trend. The amplification index is defined as the ratio of the area-weighted temperature trend between 70-90°N and 925-700hPa, divided by the average trend at all latitudes for these same pressure levels (Barnes and Polvani 2015). Values greater than 1 (grey line) correspond to amplified warming of the Arctic relative to the global mean. The sea ice trend is defined as the area-weighted mean trend in sea ice fraction over all latitudes north of 70°N; results are similar using all latitudes north of 60°N (not shown). Sea ice data from JRA-3Q is not used due to quality issues (Japan Meteorological Agency 2022). Colours and symbols are as in Figure 3.4, and as before, solid contours in panel (b) enclose 95% of the hindcast trends at each lead time. (c, d): As in (a) and (b), but for the raw Arctic warming instead of AA (i.e. the numerator of the AA ratio).

Figure 3.5a shows the range of Arctic amplification ratios in the models and reanalyses; the corresponding plot using the raw Arctic warming (i.e. the numerator

of the AA ratio) is shown in panel (c). The Arctic lower troposphere warms between 3-4 times as much as the global average in the reanalyses (Figure 3.5a), consistent with recent work (Rantanen et al. 2022; Zhou et al. 2024). Amplification in the five GC2 members ranges from values greater than 3, to less than 1 (i.e. less warming of the Arctic lower troposphere than the global average). Although the total Arctic warming tends to be slightly weaker in the hindcasts than in the free-running model (Figures 3.2 and 3.5c), the mean amplification is similar due to the greater warming at low latitudes in GC2 (Figure 3.5a). The reanalysis trends fall within the hindcast range, but are extreme: the ERA5 amplification is greater than 98% of hindcast amplification ratios at 1m lead and 99% at 6m lead; for the JRA reanalyses, the equivalent values are 94% and 97% in turn.

3.4.2 Surface amplification

Lower tropospheric (LT) AA in the reanalyses is on the upper end of the hindcast distributions. Are trends at the surface better captured? Figure 3.6 (a) plots surface AA, defined using the surface-air temperature (SAT), against the previous LT definition. JRA-3Q is omitted due to documented issues identified with sea ice trends, which strongly influences the Arctic SAT trend (Japan Meteorological Agency 2022). Surface AA is better captured by the hindcasts than LT AA: both the JRA-55 and ERA5 values are now within the central 10-90th percentile range of hindcast AA values (88th and 48th percentiles respectively). Surface and LT AA are similarly closely coupled in both the hindcasts and GC2, although the relationship is not 1:1; rather, surface AA in the models is consistently smaller than in the LT. Although this is also true of the two reanalyses, the ERA5 value is close to the 1:1 line and falls outside the range of model trends. Why is the amplification of warming confined more to the surface in the models than in ERA5?

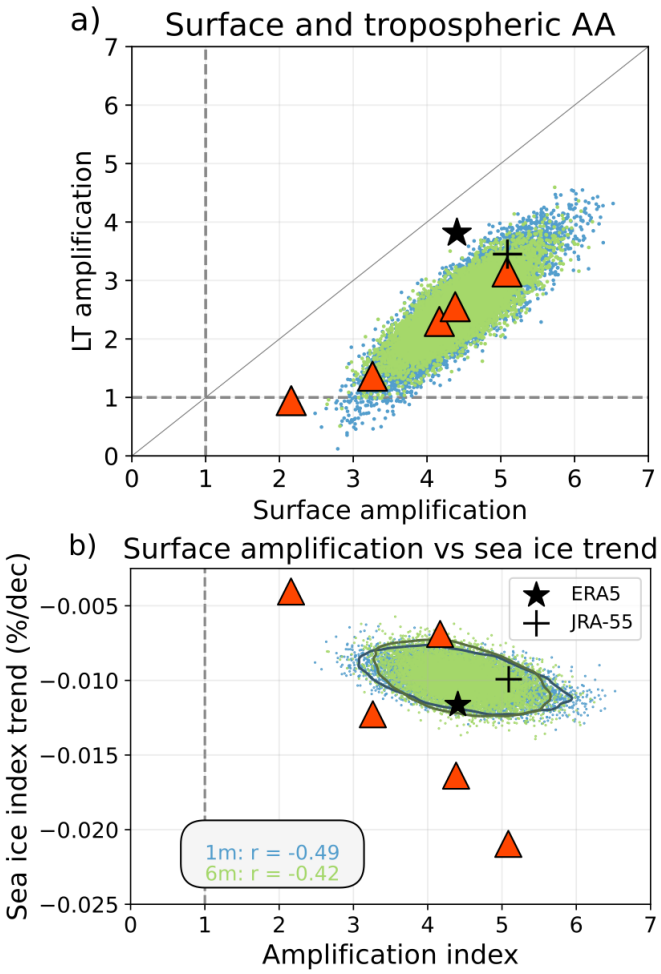


Figure 3.6: **Surface AA.** (a) AA at the surface and in the lower troposphere (LT), and (b) surface AA as a function of the sea ice trend. Surface AA is defined as the same ratio and for the same latitudes (70°−90°N) as LT AA, but using the surface-air temperature (SAT). Colours and symbols are as in Figure 3.5, and as before, solid contours in panel (b) enclose 95% of the hindcast trends at each lead time. The diagonal grey line in (a) is the 1:1 line. As surface warming is especially sensitive to sea ice trends, JRA-3Q is not shown due to quality concerns regarding sea ice data (Japan Meteorological Agency 2022).

In winter, the Arctic boundary layer (BL) is typically stably stratified, corresponding to low-level temperature inversions (e.g. Persson 2002; Zhang et al. 2011). Figure 3.7 shows the difference in temperature between the surface and the upper BL, a simple measure of the inversion strength. A climatological inversion is seen in ERA5 and in both models over the central Arctic. However, the inversion is stronger in the models than in ERA5, and Medeiros et al. (2011) find that climate models in general tend to overestimate Arctic stability compared to reanalyses. A more stably stratified BL inhibits vertical mixing, which may lead to difficulties in propagating surface warming upwards in the hindcasts and GC2 (Bintanja et al. 2012).

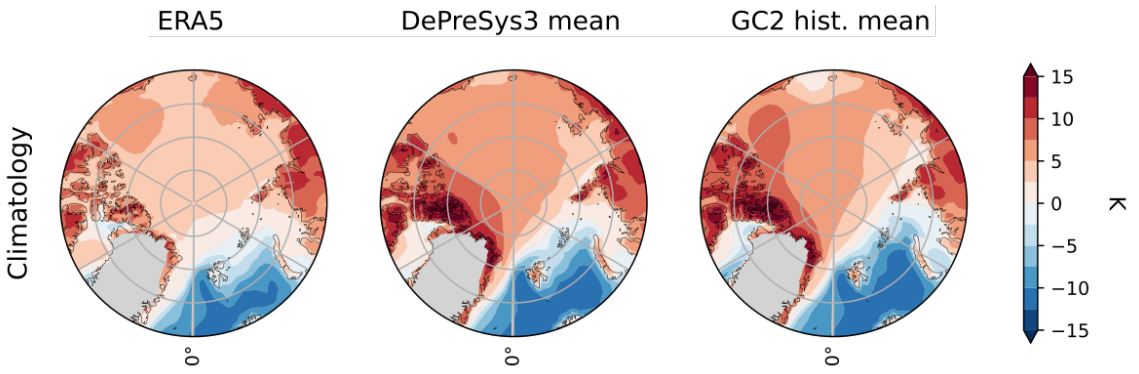


Figure 3.7: **Climatological winter inversion strength.** Inversion strength is estimated by $\Delta T \equiv T_{850} - T_s$; this corresponds to the temperature difference between the surface and a pressure level above the inversion layer (850hPa), following Medeiros et al. (2011). DePreSys3 hindcasts at 1m are shown in the middle panel. Positive values indicate a stable temperature inversion.

An important caveat here is that the ΔT metric in Figure 3.7 is a coarse measure of inversion strength that is appropriate for use with reanalyses and climate model output. It is also sensitive to surface temperature biases recently identified in the ERA5 reanalysis, which, over regions of thick sea ice in the central Arctic, has climatological SAT values up to 3K warmer than in satellite-derived data (Tian et al. 2024). Rather than the models overestimating the inversion strength, it may instead be that the inversion in ERA5 is weaker than in actuality. A deeper investigation is beyond the scope of this work, but in either case, the stronger inversion in the models could explain why their AA is confined more to the surface and the LT AA is underestimated.

3.4.3 Sea ice scaling

Figures 3.5 b, d and 3.6 b show that AA and Arctic warming scale with sea ice loss in both the hindcasts and the free-running model. However, since the sea ice concentration is initialised in the hindcasts, the range of sea ice trends is much smaller than in GC2, and the reanalysis trends are well captured by the hindcasts.

Why is the scaling stronger in GC2 than in the hindcasts? One possibility is the presence of slow, interseasonal feedbacks in the free-running model that are neglected in the hindcast timeseries. For example, sea ice is primarily melted in the summer months, allowing greater storage of heat in the ocean mixed layer. In the free-running model, realisations that see greater ice loss in a given summer may see increased heat release from the ocean in the subsequent winter.

By construction, the timeseries of 1m-lead hindcasts severs this link, as each member is initialised from the same observed sea ice conditions during the growing season in November. This could explain the weak sea ice-AA scaling for the 1m-lead hindcasts. However, the same argument does not apply to the 6m-lead hindcasts, which are initialised during the melt season in May. The scaling for the 6m-lead hindcasts is similar to the 1m-lead scaling, even though the slow feedback outlined above could act in the former but not the latter.

Regardless, Figure 3.5 shows that a wide range of AA trends is possible in the hindcasts even when sea-ice trends are well constrained, highlighting the potential for seasonal variability to contribute to AA.

3.4.4 Variability and unseen trends in the hindcasts

The GC2 ensemble-mean supports a strong forced component of Arctic amplification in the lower troposphere. However, even with just five members, a wide range of amplification ratios is seen, suggesting an important role for internal variability in modulating the forced response (Figure 3.5a).

Remarkably, not only is the mean amplification similar in the hindcasts to GC2, but the hindcasts also support a wide range of Arctic warming and amplification, both in the LT and at the surface. At 1m lead, the 5-95% LT amplification range is 1.5-3.5; approximately as many members show an amplification of 4 or more, as have values less than 1 – indicating no amplification at all (86 vs 89 of the 10,000 members). Moreover, this spread cannot be explained by differences in sea ice trends between hindcast members, which explain less than a quarter of the variance in AA at the surface (Figure 3.6b) and less than a sixth in the LT (Figure 3.5b, d). Coupled

internal variability, such as the AMV, has previously been linked to Arctic amplification (e.g. Chylek et al. 2009); however, at the short lead times considered here, hindcast members will have the same phases of ENSO, AMV and other modes of coupled interannual and decadal variability (Figures 3.1 c, d). This suggests that internal variability on short timescales, which is likely to be largely atmospheric in origin, is also sufficient to lead to a wide range of Arctic temperature trends (Zhou et al. 2024).

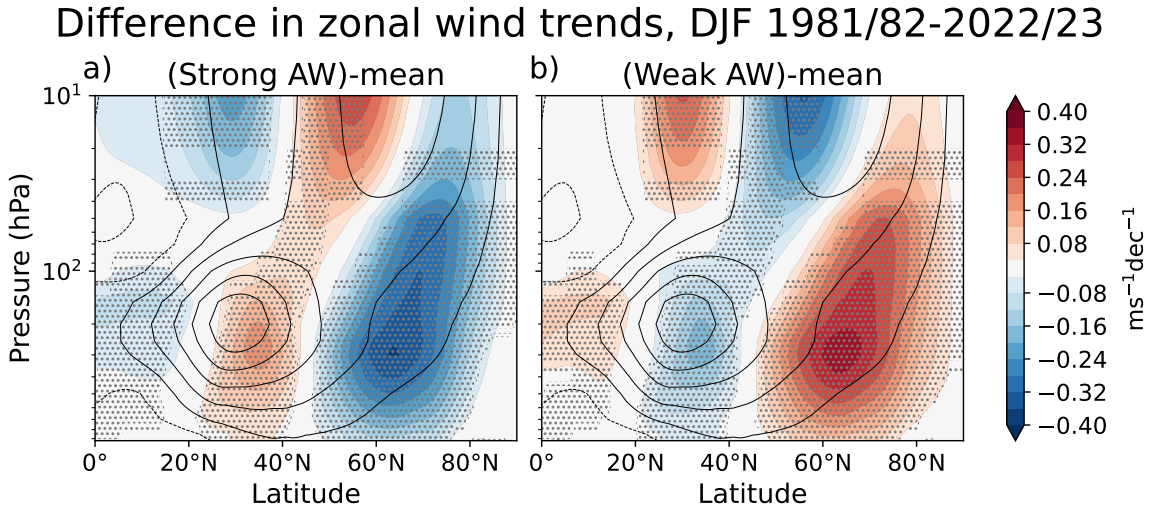


Figure 3.8: **Zonal wind trends associated with AA in the hindcasts.** Difference in DJF zonal-mean zonal-wind trends in the $N = 200$ 1m hindcast members with the strongest (a) and weakest (b) Arctic warming, relative to the ensemble mean (i.e. $trend_{200} - trend_{ens-mean}$). Arctic warming is defined in the lower troposphere, as shown in Figure 3.5c. Solid contours show the climatology. Stippling indicates a significant difference between trends in the 200-member subset and the remaining 9800 members, using a two-tailed t-test at the 5% level.

How are trends in the large-scale dynamics related to the spread in Arctic warming? Figure 3.8 shows the difference in zonal wind trends in the 1m hindcast members with the strongest and weakest Arctic warming (AW), relative to the ensemble-mean. Strong Arctic warming is associated with an expanded stratospheric polar vortex and more equatorward jet trends in the NH compared to the ensemble-mean, and vice versa for weak warming. The difference in trends extends to the surface, indicating that transient eddies have contributed to deepening the structure. This picture is largely unchanged if the subset size N is varied or if AA is used instead of the raw Arctic warming; the only difference when using AA occurs in the SH, which appears to reflect a tropical influence on global temperature rather than being associated with

the Arctic warming directly. If the raw Arctic warming is used, no trend response is seen in the SH (not shown).

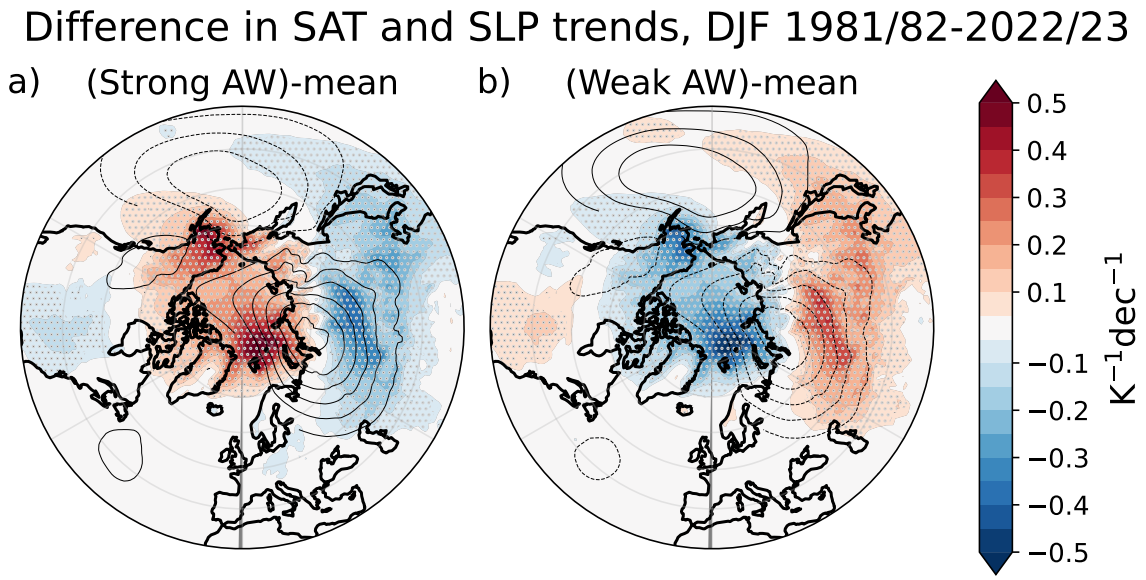


Figure 3.9: **SAT and SLP trends associated with AA in the hindcasts.** As in Figure 3.8, but for differences in SAT (shading) and SLP (contours) trends. Stippling indicates a significant difference in the SAT trends using the same method as in Figure 3.8, and only significantly different SLP trends are drawn in the contours. Solid and dashed contours indicate positive and negative trend differences, respectively.

Differences in the spatial trends for the two subsets are shown in Figure 3.9. The largest differences in Arctic warming occur over the marginal ice zones in the Barents-Kara (B-K) and Bering seas. Relative to the ensemble-mean, the strong AW subset has an anomalous high pressure trend over northern Eurasia and low pressure over the Aleutian region, leading to the distinctive "Warm Arctic, Cold Continents" pattern. Trends in the weak AW case are similar in pattern and opposite in sign, consistent with the action of a mode that can be positive or negative. Little difference is seen outside the NH extratropics, with no substantial role for seasonal-mean processes at lower latitudes in driving the spread of Arctic warming in the hindcasts.

Strong winter Arctic warming in the hindcasts is therefore associated with anomalously cool SAT trends over Eurasia. This midlatitude cooling occurs in members with strong Arctic warming, not just strong AA – and hence is not simply a byproduct of selecting for members with small trends in globally-averaged temperature (the denominator of AA). It is important to note that subsetting in this way does not imply causality: linkages in both directions have been proposed between Arctic warming

and the midlatitude circulation (e.g. Cohen et al. 2014 and Lee 2014, and discussion in Chapter 2.3.4). The "Warm Arctic, Cold Eurasia" trend is discussed further in Section 4.1.3, and the hindcasts' ability to capture the observed trend is quantified.

3.4.5 Summary of Arctic results

Two definitions of Arctic warming and amplification have been presented: the first in the lower troposphere, and the second at the surface. The hindcasts are able to capture reanalysis trends in each, but may be biased low in the lower troposphere due to a climatologically strong inversion in the model. Regardless of which definition is used, the Arctic is characterised by a wide spread of hindcast trends, even at 1m lead. AA scales with sea ice loss, but only weakly in the hindcasts, possibly due to the lack of interseasonal feedbacks by construction. No more than a quarter of the variance in hindcast AA trends is explained by differences in sea ice trends, and less in the lower troposphere. The spread therefore appears linked to seasonal variability, likely mostly atmospheric in origin. Warm and cool Arctic members, relative to the ensemble mean, appear to arise from the same dynamical mode: with an equatorward shift of the eddy-driven jet and anticyclonic trends over Eurasia for the warm Arctic subset, and vice versa in the cool members. This mode is associated with cooling of the NH continents, discussed further in Section 4.1.3.

3.5 Shifting jets

3.5.1 Trends in the reanalyses and the free-running model

Figure 3.10 shows the zonal wind trends in ERA5 and the model ensemble-means. To show the spread between model members, Figure 3.11 summarises trends toward meridional jet shifts using the zonal indices of Woollings et al. (2023). In ERA5, the trend over the satellite era is towards a poleward shift of the jets in both hemispheres in DJF, and in the SH in JJA (Figures 3.10 and 3.11). The exception is boreal summer, where wind trends in the midlatitudes are weak (Figure 3.10e). Despite the strong trends emerging in ERA5, none of the tests meet significance at the 5% level after accounting for multiple testing (Figures 3.10a and 3.10e). Good agreement is seen between the zonal index jet shifts in the three different reanalysis products (Figure 3.11).

By comparison to the reanalyses, the trends in GC2 are weak; in DJF, none of the five members show a poleward shift as large as the reanalysis in either hemisphere

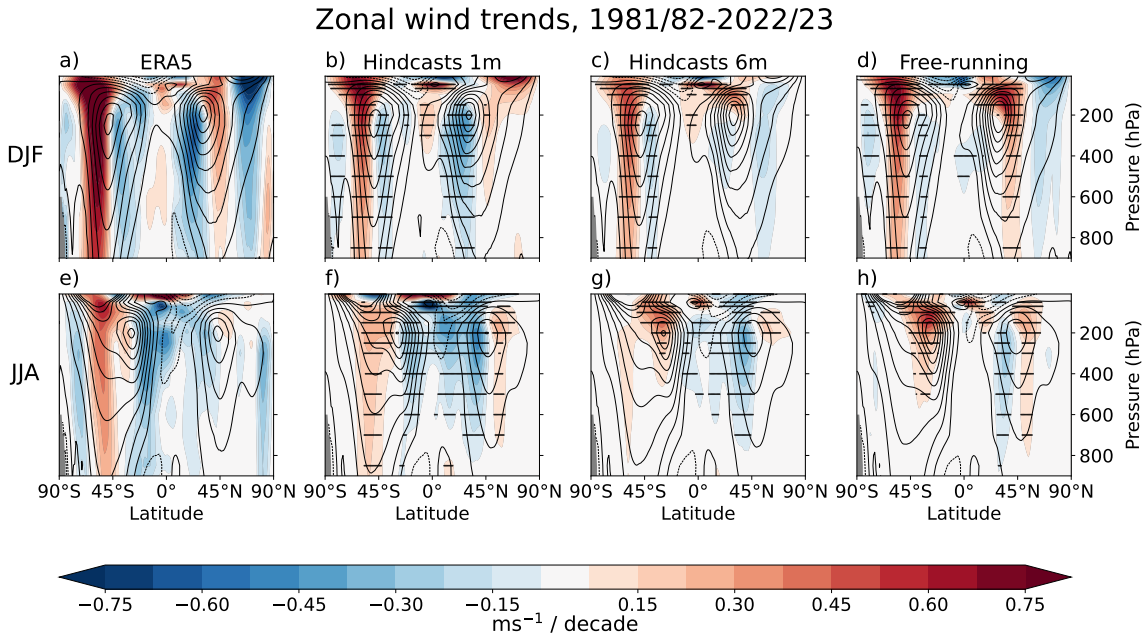


Figure 3.10: **Zonal-mean zonal wind trends.** Seasonal-mean trends shown for (a), (e) ERA5; (b), (f) and (c), (g) the hindcast ensemble-mean at 1m and 6m leads respectively; and (d), (h) the free-running ensemble mean. DJF trends are shown on the top row and JJA on the bottom. Solid contours show the corresponding climatologies with dashed lines indicating negative values and a spacing of 5m/s. Stippling follows the corresponding panels in the top row of Figure 3.2. DJF trends are calculated over 1981/82-2022/23, and JJA trends are calculated over 1981-2022.

(Figures 3.11 a, b), while in austral winter only two show any poleward shift at all (Figure 3.11 d). Combining hemispheres, on average a poleward shift is seen in ERA5 for both seasons (Figures 3.11 c and 3.11 f). The GC2 members show a weaker average poleward shift in DJF and shifts of both signs in JJA.

3.5.2 Trends in the hindcasts

Averaged over both hemispheres, the hindcasts at 1m lead robustly show a poleward shift of the jets in both DJF and JJA (Figures 3.11c and 3.11f); only 2% of members fail to show an average poleward shift in DJF, and less than 1% in JJA. The hindcasts hence provide additional evidence that poleward jet shifts are emerging on the global scale. In each season and hemisphere, the reanalysis trends lie within the hindcast distribution, although often towards the edge.

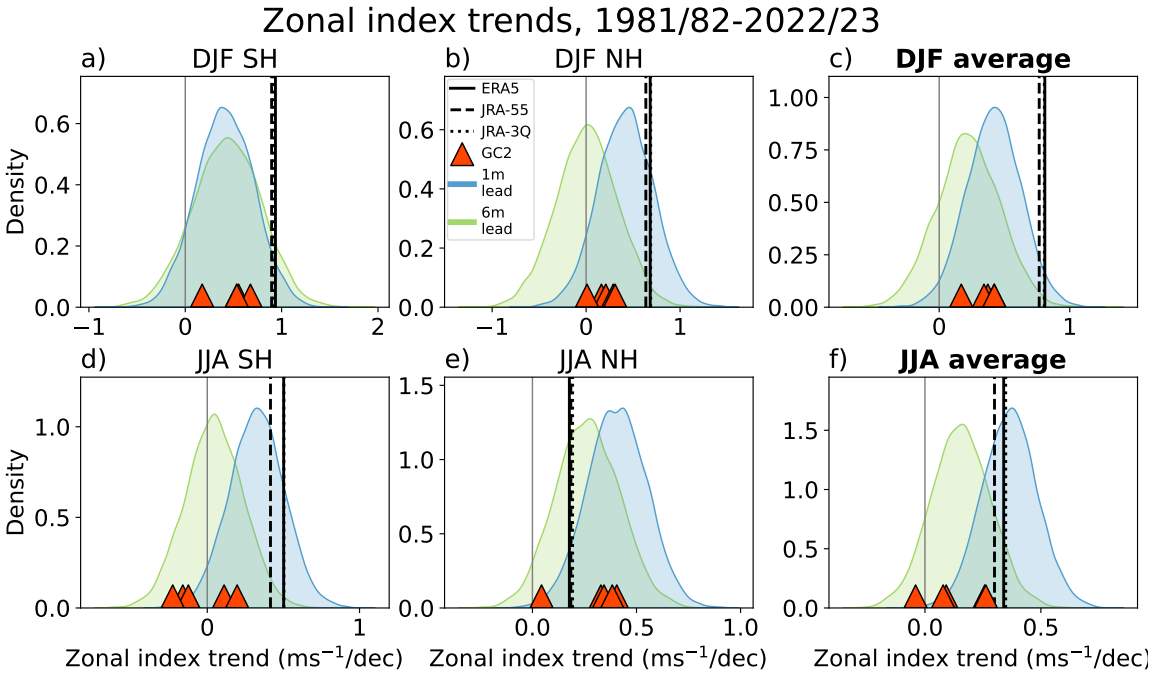


Figure 3.11: **Jet shifts as measured by trends in zonal wind indices.** Each zonal index is the difference in zonal wind trend between the poleward and equatorward sides of the climatological jet, averaged over 7 levels between 200-700hPa. The latitude bands used to define each index vary by season and hemisphere (see Woollings et al. (2023)). Poleward shifts correspond to positive trends in the index in each hemisphere. (Top row) Zonal index trends in the SH (a), NH (b) and the average over both hemispheres (c) for DJF; (bottom row) as above but for JJA. Shading and symbols are as in previous figures. Note the different x-axes in each panel.

3.5.2.1 Winter jet trends

The effect of the initialisation is greatest in the winter hemispheres. At 1m lead following initialisation, most hindcast members show poleward shifts in both boreal and austral winter, in contrast to the weak trends in GC2 (Figure 3.11). However, after just 6 months of integration, the hindcast trend distribution is symmetric about zero.

As well as having more realistic trends in the zonal indices, the zonal-mean structure of the trends at 1m lead closely resemble ERA5 (Figure 3.10). Compared to the free-running model, the hindcasts have stronger easterly trends in the subtropics, particularly in winter. In addition, the deep acceleration on the poleward side of the austral winter jet is well captured by the hindcasts at 1m lead, in contrast to GC2. Here again, the ensemble-mean trend at 6m lead more closely resembles the

free-running model.

3.5.2.2 Summer jet trends

Initialisation does not lead to similar improvements in summertime jet trends. The hindcast distribution in austral summer is nearly identical at both lead times (Figure 3.11a), and the ensemble-mean trend patterns broadly match the uninitialised model (Figures 3.10b, c and d). Although trends in the reanalyses are within the range of hindcast trends, only 4% of the hindcast members at 1m lead have as strong a poleward shift as in ERA5. Circulation trends in this season are strongly influenced by changes in stratospheric ozone over the satellite era (Gillett and Thompson 2003; Son et al. 2010; Seviour et al. 2017). Changes in ozone concentration are nonlinear in time (Banerjee et al. 2020), although the linear trend here is consistent with depletion dominating in the satellite era overall. The hindcasts share the same external forcings (including ozone) as the free-running GC2; the insensitivity of the hindcast trends to lead time could therefore point to the importance of ozone forcing in austral summer. Even here, however, trends of both sign are possible within the hindcasts, which indicates a role for internal variability on seasonal timescales (Seviour et al. 2017).

While the reanalysis trends in austral summer are on the stronger edge of the hindcast distribution, the opposite is true of the boreal summer jet. A modest poleward shift of the NH jet is seen in the reanalyses (Figure 3.11e), but the pattern of wind trends is broadly weaker than in other seasons (e.g. ERA5 in Figure 3.10e). By contrast, the hindcasts robustly estimate a northward shift of the zonal-mean circulation at 1m lead (Figure 3.11e), with only 25 of the 10,000 hindcast trends failing to show an overall poleward shift. The ensemble-mean trends in the hindcasts largely resemble the GC2 ensemble mean, although the weakening of the subtropical winds is again greater in the hindcasts than the free-running model (Figures 3.10f, g and h).

3.5.2.3 Jet biases

The initialisation seems to improve the winter jet trends in the hindcasts, but not in summer. Interestingly, a similar pattern is seen in the improvement in the mean zonal wind bias. Figure 3.12 shows the zonal wind bias in the models relative to ERA5. Biases in the free-running model are large in both hemispheres and seasons, with a pronounced equatorward bias in winter. In general, the biases are greatly reduced in the 1m lead hindcasts, and to a lesser degree in the 6m hindcasts.

Figure 3.13 summarises this picture for the zonal indices used above, demonstrating the substantial reduction in the equatorward winter jet bias following initialisation. In the NH, the winter bias is reduced primarily over the Pacific and Atlantic sectors; in the SH, the reduction is greatest in the South Pacific sector (not shown), reflecting the strong zonal asymmetry of the austral winter jet (Breul et al. 2023). By 6m lead, the biases have begun to redevelop, though remain smaller than in the free-running GC2.

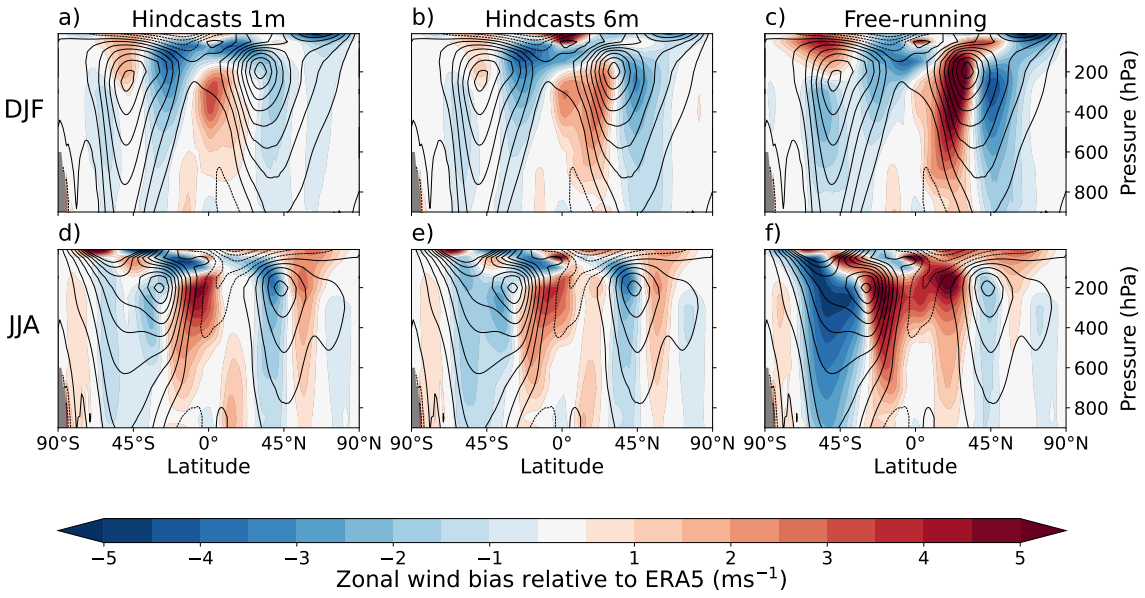


Figure 3.12: **Mean model zonal wind biases.** Shown for DJF (top row) and JJA (bottom row). The mean bias is calculated as the time-mean difference between the model and ERA5 fields, as in Figures 3.1 a and b. The time-mean is taken over the same period as the trend calculations: DJF 1981/82-2022/23, and JJA 1981-2022.

The free-running model biases are much smaller in the summer jets than in winter. Although zonal wind biases are somewhat reduced in austral summer (Figure 3.12), the reduction does not project strongly onto the zonal index (Figure 3.13) and the effect on the trends remains modest. On the other hand, the poleward jet bias in boreal summer is larger following initialisation, which may be reflected in the stronger hindcast trends. In light of previous work highlighting generally low model skill in the European and Atlantic sectors in boreal summer (Patterson et al. 2022), the larger bias limits confidence in the hindcast jet trends in northern summer. Trends in the summer NAO are presented in Section 4.2.1.

Overall, compared to the free-running model, the hindcasts better capture reanalysis trends in the winter zonal-mean jets. This coincides with a substantial reduction

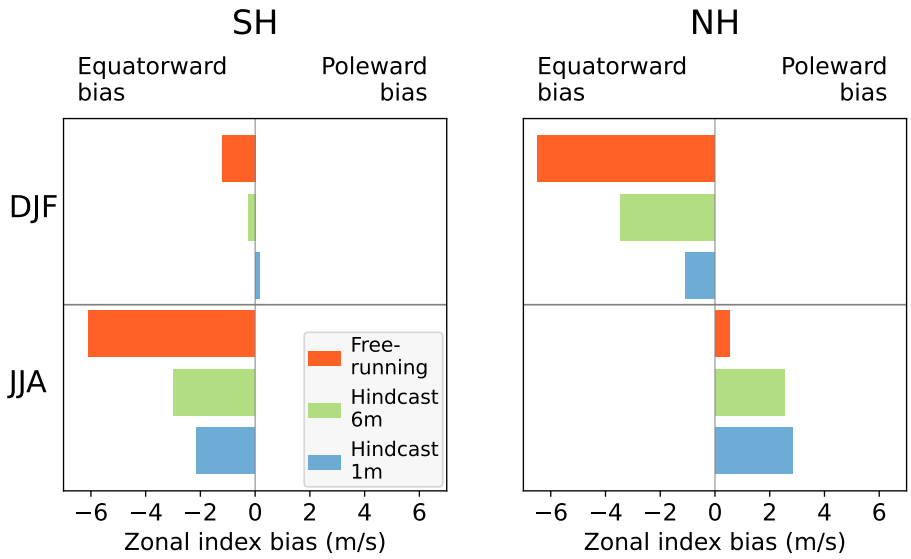


Figure 3.13: **Mean model biases in the zonal indices of Figure 3.11**. Shown for DJF (top row) and JJA (bottom row). The bias is calculated as in Figure 3.12.

in the model's mean equatorward bias following initialisation, although the bias begins to redevelop by 6m. The initial free-running bias is smaller in the summer jets than in winter, which may explain why initialisation from the observed conditions fails to improve agreement with the reanalysis summer jet trends. In boreal summer, the more poleward jet trends in the 1m hindcasts may reflect the stronger poleward bias following the initialisation.

3.5.3 Variability and unseen trends in the hindcasts

In each season and hemisphere, the hindcasts support a wide spread of trends in the midlatitude jets. Both poleward and equatorward shifts are possible in the hindcasts in each case, even at 1m lead. Confidence in the hindcast trends is limited in boreal summer, where the mean bias grows after initialisation. For the other seasons, however, the reduced mean biases build trust in the plausibility of the hindcast trend distributions. In these cases, as in the wintertime Arctic (previous section), it seems that intrinsic seasonal and subseasonal variability can lead to a broad array of possible hindcast trends.

For example, the mean poleward shift of the austral summer jet is similar in the hindcasts to the free-running model, likely reflecting the strong forced trend in this season and the models' shared ozone scheme. Even so, 8% of the hindcast trends at 1m lead (and 11% at 6m) show an equatorward shift – opposite to the model's forced

response. This highlights the scope for internal variability to modulate trends in the jets in this season, as noted by Seviour et al. (2017). The hindcasts are somewhat more confident in poleward shifts in the winter jets, though 8% and 4% of the hindcast trends show equatorward shifts in boreal and austral winter, respectively. Even where the hindcasts are relatively confident in the sign of the trend, clearly its magnitude can vary substantially between members.

Finally, Figures 3.11c and 3.11f show that, averaged over both hemispheres, the hindcasts support poleward jet shifts on a global scale. However, in some cases, an average poleward shift globally may hide a weak or equatorward shift in one hemisphere. Accounting for this, 85% of 1m-lead members show concurrent poleward shifts in both hemispheres for DJF, and 96% for JJA. In other words, given the same initial conditions, 15% of the hindcast members had an equatorward jet shift in at least one hemisphere in DJF.

These results highlight the difficulty in identifying forced trends in the jets over the short historical period: taken at face value, they suggest that seasonal variability – likely largely atmospheric in origin – can lead to a wide range of trends, even when members share forcing and large-scale coupled modes of variability. In a non-negligible number of members, this short-timescale variability is sufficient to generate trends opposite in sign to the forced response, as in austral summer.

3.5.4 Relation to meridional temperature gradients

A key source of uncertainty in understanding changes in the jets is their relation to changing meridional temperature gradients. Figure 3.14 shows the zonal index trends plotted against trends in upper-level meridional temperature gradients (ΔT); the original figure of Woollings et al. (2023) using the same indices is shown in Figure 2.2.

The overall picture in the CMIP6 ensemble is of a wide scatter between models, with the spread of dynamical trends across the ensemble poorly explained by changes in ΔT (Woollings et al. 2023). By comparison, trends in the hindcasts are more structured, particularly in the winter hemispheres. The similar behaviour at both 1m and 6m lead further builds confidence in the robustness of these relationships. In austral winter (panel c), a moderate positive correlation is seen between ΔT and ΔU. This is consistent with the typical response to zonal-mean tropical warming in model projections, with stronger temperature gradients associated with more poleward jet trends (Shaw 2019).

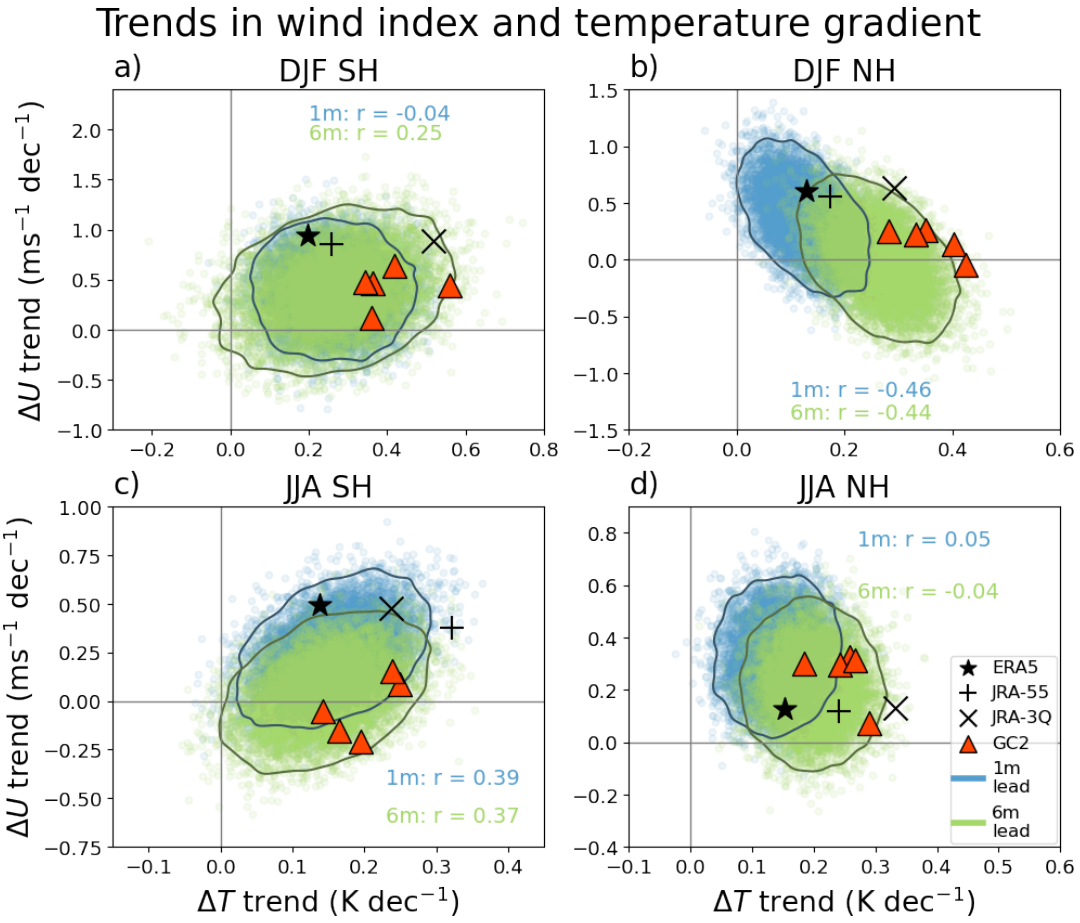


Figure 3.14: **Trends in the zonal wind indices plotted against trends in the 250hPa meridional temperature gradient**. Colours and symbols are as in Figure 3.5, and solid contours again enclose 95% of the hindcast trends at each lead. All correlations noted on the subplots are significant at $p < 10^{-6}$. Note that the ordering of the panels matches Figure 3.11 but not the original figure of Woollings et al. (2023), which is reproduced in Figure 2.2.

Contrary to the picture in the SH, however, the strongest relationship occurs in boreal winter (panel b), where more poleward jet trends are instead associated with *weaker* ΔT trends. Such a relationship seems surprising on multidecadal timescales, but on interannual timescales, a steeper meridional temperature gradient is not always associated with a poleward shift of the circulation. La Nina events are characterised by a narrow zonal-mean cooling of the tropical troposphere, but poleward shifted jets through altering the refraction of transient eddies (Seager et al. 2003). It is possible that the anticorrelation in boreal winter reflects a common association of weaker ΔT and more positive ΔU trends, with trends towards a more La Nina-like state or towards increased La Nina variability. This appears to be consistent with

the downgradient heat flux trends of Woollings et al. (2023), and with the $\Delta T - \Delta U$ anticorrelation seen in the UKCP runs in that paper. The influence of SST trends in the tropical Pacific on the global circulation is discussed further in Chapter 5.

It should be remembered that trends in the jets are complex and their behaviour in any season will depend on a combination of factors - including both the expected response to zonal-mean tropical warming and any potential influence from trends towards zonally asymmetric tropical heating, in addition to extratropical factors. Compensation between these opposing influences could explain the weaker relationships seen in the other seasons. In austral summer (panel a), in addition to any tropical influence, stratospheric ozone is likely to strengthen ΔT over this period and is associated with a poleward shift of the SH jet (e.g. Gillett and Thompson 2003). This additional forcing could be sufficient to blunt the potential La Nina-associated anticorrelation seen in the NH for DJF. Curiously, austral summer is the only season where a clear correlation between ΔT and ΔU is seen in the CMIP6 ensemble. One possible explanation of this is that the positive correlation in Woollings et al. (2023) reflects differences in the ozone-forced response among different CMIP6 models, which could vary considerably across the ensemble. It is notable that within the single-model UKCP ensemble, no robust relationship is apparent, consistent with the hindcasts picture.

Table 3.1: Latitude bands used to calculate the zonal wind (ΔU) and meridional temperature gradient (ΔT) indices in each season and hemisphere. ΔU is calculated using a vertical average over 200-700hPa, as in the previous figures, and ΔT is calculated at the 250hPa level, as in Woollings et al. (2023). The latitude bands for the temperature gradients correspond to gradients across the time-mean jet latitude.

	DJF	JJA
ΔU_{SH}	(50-60S) minus (30-40S)	(40-60S) minus (10-30S)
ΔU_{NH}	(40-50N) minus (20-30N)	(50-60N) minus (30-40N)
ΔT_{SH}	(20-40S) minus (60-80S)	(20-40S) minus (60-80S)
ΔT_{NH}	(0-20N) minus (40-60N)	(20-40N) minus (60-80N)

3.5.5 Summary of jet trends

Trends in the zonal-mean jets have been presented, and summarised using the zonal wind indices of Woollings et al. (2023). The hindcasts offer support to the emergence of poleward jet shifts on a global scale; averaged over both hemispheres, 98% of 1m

hindcasts show a global poleward jet shift in DJF, and 85% show a concurrent poleward shift in each hemisphere. Compared to the free-running GC2, hindcast trends in the winter jets are improved, although the effect of the initialisation appears to be substantially diminished by 6m lead. Trends in the summer jets are not improved, however. This appears to coincide with a reduction in the equatorward jet bias in winter but not in summer, although it is unclear if improvements in the biases and trends are linked. Regardless, in both seasons and hemispheres, the hindcasts support a wide spread of trends in the jets, even in cases with a strong forced component such as austral summer. Neglecting the role of internal variability in this way could lead to incorrect conclusions regarding the identification of circulation trend errors in initialised models. Finally, poleward jet shifts are associated with strengthened meridional temperature gradients in austral winter, but weakened gradients in boreal winter. The latter relationship is surprising, but could be consistent with more La Nina-like trends in the tropical Pacific leading to stronger poleward jet shifts in NH winter.

3.6 Hadley circulation

Section 2.5 introduced the discrepancy in recent Hadley Circulation trends, between coupled models, which tend to show a weakening of the boreal winter cell, and the strengthening seen in reanalyses. Do the hindcasts lend support to either of these positions?

The meridional overturning circulation in the latitude-pressure plane ($\phi - p$) is described by the mass streamfunction ψ,

$$\psi(\phi, p) = \frac{2\pi R_E \cos\phi}{g} \int_0^p [v(\phi, p')] \, dp, \tag{3.2}$$

where v is the meridional wind, R_E is the Earth radius, g is the acceleration due to gravity, and square brackets represent a zonal average (Peixoto and Oort 1992). The strength of the Hadley circulation (HC) is commonly taken as the maximum value of the streamfunction, $|\psi|_{max}$, in the midtroposphere. The midtroposphere is typically taken either as the 500hPa level (e.g. Chemke and Polvani 2019), or using a midtropospheric average, as is used here (700-300hPa); in practice, this choice does not alter the results that follow.

Figure 3.15 shows timeseries of DJF and JJA $|\psi|_{max}$ anomalies. The hindcasts capture much of the reanalysis variability in $|\psi|_{max}$, with significant detrended correlation skill of 0.53-0.67 (Table 3.2). Much of the interannual variability is related

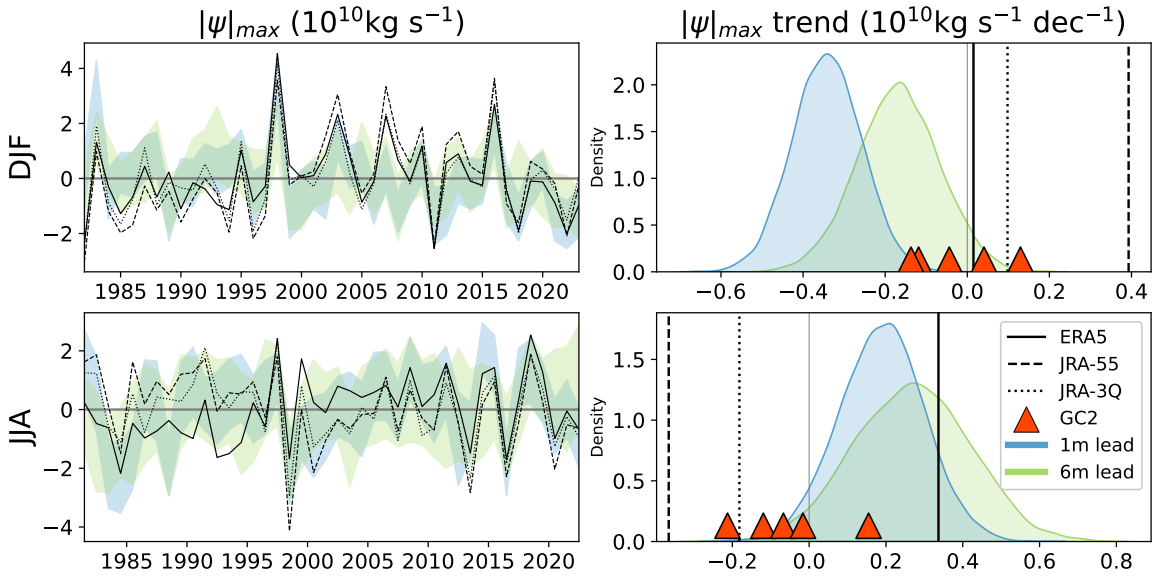


Figure 3.15: **Distribution of trends in Hadley Circulation strength.** $|\psi|_{max}$ anomalies (left) and trends (right), for DJF on the top row and JJA on the bottom row. Shading in the anomaly plots corresponds to the 10-90% range of hindcast anomalies. Colours and symbols are as in this chapter's previous plots. Note the different x-axes in the trend plots.

to ENSO, with the HC strengthening and weakening in response to El Nino and La Nina respectively (Nguyen et al. 2013; D'Agostino and Lionello 2017).

The range of $|\psi|_{max}$ trends is shown in the right-hand column of Figure 3.15. Trends of both sign are seen in JJA, in the reanalyses and in the hindcasts. Given the lack of robustness of the reanalysis trends, this section will instead focus on DJF, where each of the reanalysis trends is positive. On the one hand, the DJF HC strengthens in the reanalyses, as found by previous studies cited in Section 2.5; here, it is shown that this is also true of the new JRA-3Q product, which has not yet been included in the published literature. The strengthening trends in the reanalyses are contrary to the weakening trends in free-running historical simulations. Do the initialised hindcasts offer support for the reanalysis trends?

Remarkably, despite being initialised from observed conditions, none of the 10,000 1m DJF hindcast trends are positive. Interestingly, though, 5% of the 6m lead hindcasts show a strengthening trend, and 2 of the 5 free-running members. Rather than supporting the reanalysis trends, it appears that the more strongly constrained the model is by the observed conditions, the more negative its trends in the DJF HC strength.

Table 3.2: (Top three rows): $|\psi|_{max}$ trend p-value, and detrended correlation skill of the hindcast ensemble-means with the reanalysis, for DJF. (Bottom three rows): as above, but for JJA.

	ERA5	JRA-55	JRA-3Q
p_{trend}	0.93	0.07	0.59
$corr_{1m}$	0.62	0.59	0.66
$corr_{6m}$	0.54	0.56	0.56
p_{trend}	0.03	0.04	0.22
$corr_{1m}$	0.53	0.67	0.66
$corr_{6m}$	0.54	0.56	0.56

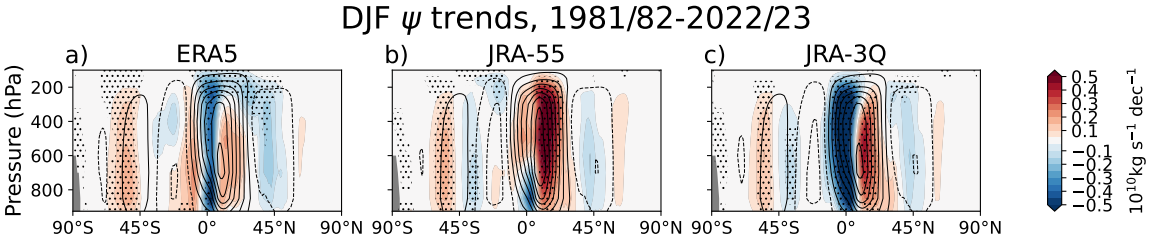


Figure 3.16: **Reanalysis trends in the DJF meridional mass streamfunction.** Solid contours indicate positive climatological values (clockwise circulation) and dashed contours the opposite. Stippling indicates significant trends at the 5% level using a two-tailed t-test.

Confidence in the reanalysis trends is further undermined when the full spatial structures are compared (Figure 3.16). Some features are similar in all reanalyses: for example, the HC strengthens between approximately $5° - 20°N$ in each reanalysis product. However, the magnitude of the strengthening differs considerably between the reanalyses, and the overall structure of the tropical trends is not consistent. A clear difference is seen compared to the 1m hindcast ensemble mean (Figure 3.17 a), which almost entirely lacks the region of positive trend seen in the reanalyses. Instead, the most striking feature of the hindcast trend is the equatorial weakening, which is present in JRA-3Q but much weaker in the other reanalyses. The weak ensemble-mean trend in GC2 suggests that the forced component of the trend is not a dominant factor over this period, implying that the observed evolution of coupled variability likely contributes to the weakening trend in the 1m hindcasts.

Early suggestions linked the DJF HC strengthening in reanalyses to weak or negative temperature trends in the tropical mid-upper troposphere; it was argued that

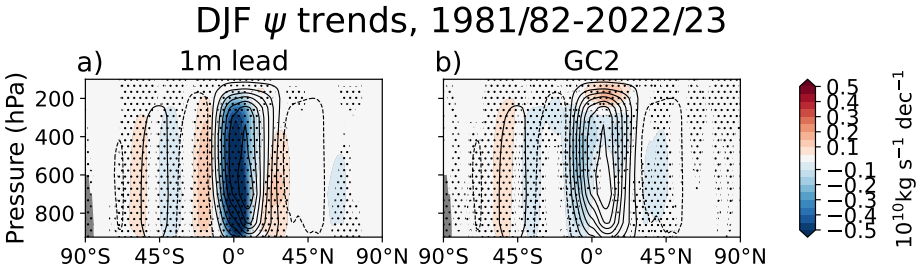


Figure 3.17: **Model trends in the DJF meridional mass streamfunction.** Solid contours indicate positive climatological values (clockwise circulation) and dashed contours the opposite. Stippling indicates points where $> 75\%$ of members agree on the sign of the trend.

nudging the forecast model towards a more unstable temperature profile may induce a spurious strengthening of the HC (Mitas and Clement 2006). However, as shown above in Section 3.3, upper-level warming in the hindcasts is comparable to, or at 1m lead slightly weaker than, the reanalysis trends (Figure 3.4 c). It therefore seems unlikely that the NH HC strengthening in reanalyses is due to their tropospheric temperature trends.

Another explanation was proposed by Zaplotnik et al. (2022), who suggest that Atlantic Multidecadal Variability (AMV) contributes to the strengthening HC in ERA5 (although they note that the relationship is weaker in other reanalyses). The hindcasts offer a direct test of this hypothesis, since the phasing of AMV is similar in the hindcasts and ERA5 (Figure 3.1). Despite this, the 1m hindcasts still fail to support a strengthening of the DJF HC. On the contrary, comparing the robust trends in the initialised runs with the weak ensemble-mean trend in the uninitialised model (Figure 3.17), it seems more likely that coupled variability has instead contributed to a *weakening* of the DJF HC.

Overall, it is suggested that these results support the emerging evidence of errors in the reanalyses, such as the erroneous trends in tropical latent heating identified by Chemke and Polvani (2019) and Zaplotnik et al. (2022). Errors in streamfunction-based trends could also arise due to the meridional winds in reanalyses being less well constrained by observations than zonal winds, and not being constrained dynamically by other fields (e.g. through thermal wind balance; Baldassare et al. 2023). A true weakening of the DJF HC, as the hindcast results would suggest, is also consistent with pressure observations (Chemke and Yuval 2023) and with the expected forced response to climate change as seen in models (Vallis et al. 2015; Chemke and Polvani 2021).

3.7 Conclusions

The main results of this chapter are summarised below. A fuller discussion of the hindcast trends approach, including some discussion of this chapter's results, is provided in the conclusions of Chapter 4.

1. Tropical tropospheric warming

    - Upper-level warming is substantially reduced following initialisation, reflecting the strong control exerted by the underlying SSTs on warming in the tropical atmosphere.
    - This improvement is well explained by differences in SST trends in the regions of deepest convection, scaling linearly with $pTAS$. Given the uncertainty between different products, the hindcasts appear to behave consistently with the reanalysis trends.

2. Arctic amplification

    - Reanalysis AA trends at the surface and to some extent in the lower troposphere are captured by the hindcast ensemble, although model trends in the latter may be biased low due to a climatologically strong inversion.
    - The Arctic displays a remarkably wide spread of hindcast trends at both lead times, and at both the surface and in the troposphere.
    - No more than a quarter of the variance in hindcast AA trends is explained by differences in sea ice trends, and less in the lower troposphere. The spread therefore appears linked to seasonal variability, likely mostly atmospheric in origin. Arctic warming is associated with an equatorward shift of the eddy-driven jet and anticyclonic trends over Eurasia.

3. Jet trends

    - The hindcasts offer support to the emergence of poleward jet shifts on a global scale (Woollings et al. 2023), with 85% of the 1m hindcasts showing a concurrent poleward jet shift in each hemisphere.
    - Initialisation appears to improve winter jet trends, though not in summer. This coincides with reduced equatorward winter jet biases following initialisation, although improvements in the trends and biases may not be linked.

- The hindcasts support a wide spread of trends in the jets, even in cases with a strong forced component such as austral summer.
- Poleward jet shifts are associated with strengthened meridional temperature gradients in austral winter, but weakened gradients in boreal winter; the latter relationship may reflect a relationship between more La Nina-like trends and stronger poleward jet shifts.

4. Hadley circulation

    - The DJF Hadley Cell weakens in all 1m hindcasts and in 95% of members at 6m. This provides strong evidence against the strengthening seen in reanalyses, and is consistent with the weakening inferred from SLP observations (Chemke and Yuval 2023).

Chapter 4

Spatial structure of trends in the hindcasts

The previous chapter focused mainly on trends in the zonal mean. In this chapter, some of the spatial structure of the hindcast trends is explored. Even when zonally averaged, we have seen that the hindcasts support a wide range of trends; we can expect that variability between members will be even more important for trends on regional scales.

The structure of this chapter is as follows. The first section below deals with trends in surface-air temperature (SAT), beginning with trends in the global and hemispheric means, before exploring the dynamical structures that influence trends on these larger scales. A particular aspect of the SAT trend is then discussed in depth, namely the regional Eurasian cooling that arose earlier in the zonal-mean Arctic amplification section (Section 3.4). Finally, the broader trends in sea-level pressure (SLP) are discussed, including a regional focus on the North Atlantic sector where 20th Century trends are poorly captured by models.

4.1 Surface-air temperature

4.1.1 Summary of trend improvements

Figures 4.1 shows the distribution of surface air temperature (SAT) trends for DJF and JJA respectively. Trends are averaged globally, by hemisphere, and by land or ocean. For brevity, this section mostly focuses on DJF, but the broad conclusions also apply to JJA (shown in Figure A.1)

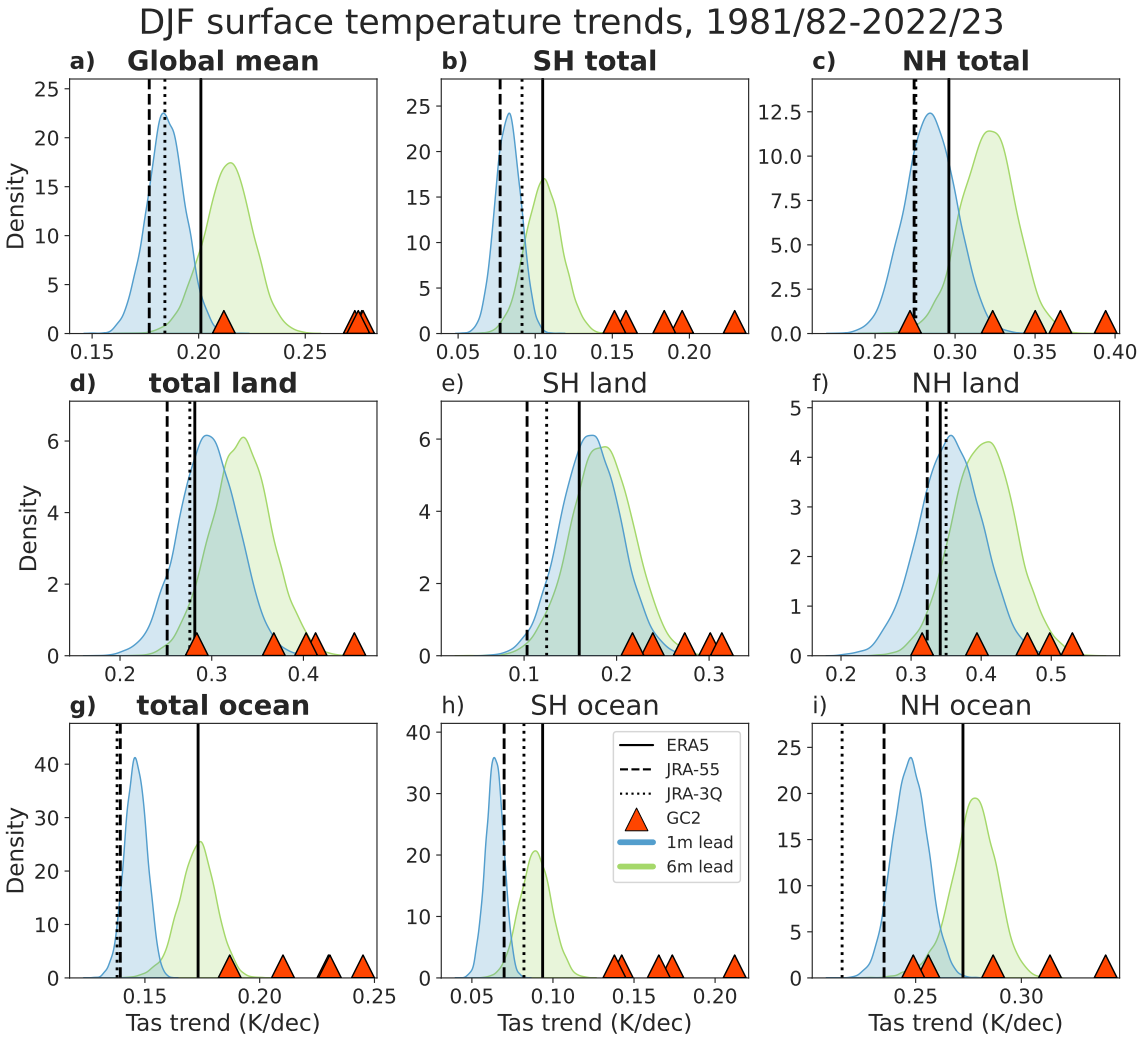


Figure 4.1: **Area-averaged SAT trend distributions for DJF.** DJF area-weighted surface air temperature trends (1981/82-2022/23), averaged (a) globally; (b) over the SH; and (c) over the NH; and over all land (d) and ocean (h) grid boxes. Plots (e) and (f) show the averages over land in the SH and NH respectively; plots (h) and (i) show the same but averaged over the ocean. Colours and symbols are as in Figure 3.5. Note the different x-axes in each panel.

Globally, GC2 warms more than the reanalyses in both DJF and JJA (panel (a) on Figures 4.1 and A.1). The different reanalyses show global warming trends of between 0.17-0.20K/decade in DJF, compared to trends above 0.27K/decade in four of the five GC2 members. Compared to GC2, the 1m lead hindcasts much better capture the reanalysis trends; even by 6m lead, the hindcast trends remain smaller than all but one of the GC2 members, showing that the effect of the initialisation on SAT trends persists beyond seasonal timescales.

Most of the global difference between the hindcasts and GC2 is attributable to the SH (c.f. panels (b) and (c) of each figure). This likely reflects the greater proportion of the SH that is covered by ocean, with temperatures over ocean better constrained by the initialisation. Over SH ocean, GC2 warms much more than the reanalyses; the hindcasts better capture the reanalysis trends here, although it is notable that the 1m-lead trends are surprisingly somewhat cooler than in the reanalyses (panel (h) of each figure).

In the NH, the difference between the hindcasts and GC2 is smaller (panels (c) in each season); although a systematic difference is still seen between the 1m and 6m distributions, showing that the initialisation does affect the SAT trends, it appears that the uninitialised GC2 is still able to produce trends similar to the reanalysis. Two factors likely contribute to the NH picture: firstly, the greater proportion of land as noted above, and secondly, the relative insensitivity of Arctic warming to the initialisation (Section 3.4).

The ensemble-mean model trends are shown for DJF in Figure 4.2 and for JJA in Figure A.2, along with the ERA5 and JRA-55 trends. Panels (d), (f), and (h) show the model "trend error" relative to ERA5.

Over ocean, the ensemble-mean trend in the hindcasts closely resembles the reanalyses; the spatial correlation between ERA5 and the 1m hindcast trend over ocean is 0.89 for DJF and 0.63 for JJA ($p < 0.01$), and the difference plots over ocean show only minor discrepancies at 1m lead. Regionally, improvements in the 1m hindcasts are seen in the PDO negative-like trend in the Pacific, and in the pattern of weak or negative trends over the Southern Ocean. In both cases, however, by 6m lead the ensemble-mean is beginning to resemble the more uniform Pacific and Southern Ocean warming seen in GC2.

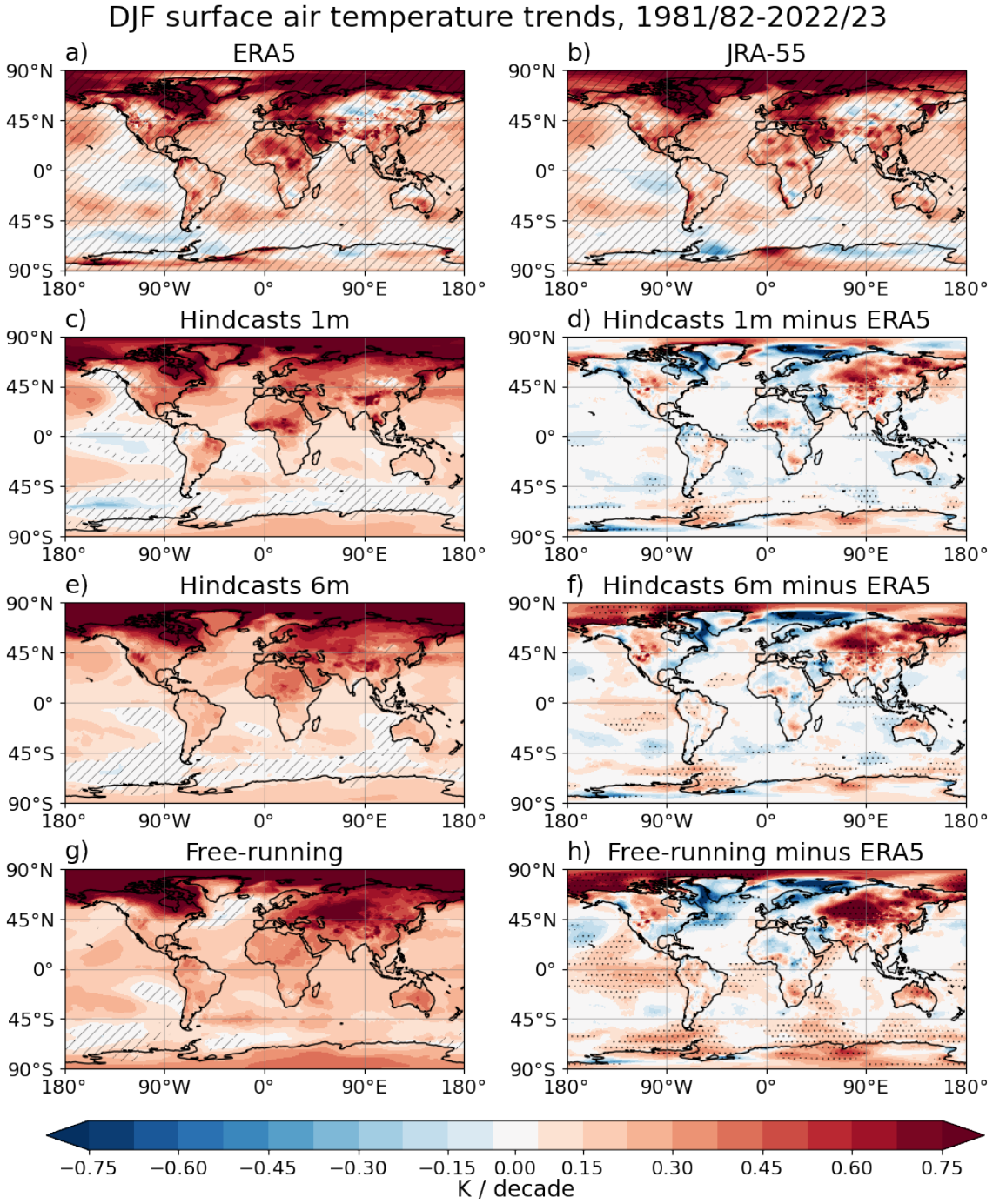


Figure 4.2: **Maps of ensemble-mean hindcast SAT trends in DJF.** Trends are calculated over 1981/82-2022/23, for (a) ERA5, (b) JRA-55; (c), (e) the hindcast ensemble-means at 1m and 6m lead respectively; and (g) the free-running ensemble-mean. Hatching on these plots is the inverse of the stippling in Figure 3.2 (here, hatching *hides* non-significant points, defined as before). (d, f, h): Differences of the plots to the left from the ERA5 trend in (a). As in Figure 3.2, stippling on (d), (f) and (h) indicates where the hindcast mean is more than $2S$ away from the ERA5 mean, where S is the standard error on the ERA5 regression.

A somewhat different picture is seen over land. Some improvement is seen in the hindcast ensemble-means, particularly where GC2 warms much more than the reanalyses - most notably over central Eurasia. However, the hindcast trends remain much smoother over land than in the reanalyses, similar to other initialised forecast systems (Choi et al. 2016; Becker et al. 2022). Consequently, the pattern of trend discrepancies over land is broadly similar in the hindcasts and GC2 (panels (d), (f) and (h)). One interpretation of these figures is as trend errors in the hindcasts, since they inherit the same phasing of coupled variability as the observed system. However, as the following section will show, short-timescale variability can lead to a wide range of SAT trends over land, even when the ocean is initialised.

4.1.2 Internal variability in the hindcasts

Figures 4.1 and A.1 show that a range of SAT trends are possible in the hindcasts. Even at 1m lead, a non-trivial spread is seen in global mean surface-air temperature (GMST), with a 5-95% range of 0.17-0.20K/dec in DJF and 0.15-0.18K/dec in JJA. Over the satellite era, this equates to a difference of 0.12K of warming in DJF and 0.11K in JJA across the 1m hindcast ensemble, equivalent to roughly a tenth of the observed warming since preindustrial times (Intergovernmental Panel on Climate Change (IPCC) 2023).

Two other notable features can be seen in these figures. Firstly, the spread in SAT trends is wider over land than ocean; secondly, the spread over land is wider in the winter hemisphere than in the summer hemisphere. The first point follows from the initialisation strategy: as the ocean is directly initialised, the range of oceanic SAT trends is constrained, whereas the the land surface – though free to adjust during the assimilation run – is not directly constrained. The second point, regarding seasonality, hints at the mechanisms responsible for the spread in hindcast GMST trends.

To probe this, the 200 hindcast members with the warmest and coolest GMST trends in each season are selected, and the differences in their SAT trends relative to the ensemble mean are plotted in Figure 4.3 (i.e. $SAT_{200} - SAT_{ens-mean}$). Overlaid in contours are the corresponding differences in SLP trends. This figure tell us the following:

1. Very little of the spread in GMST trends occurs at low latitudes; instead, the largest differences are found over the mid to high latitude continents.

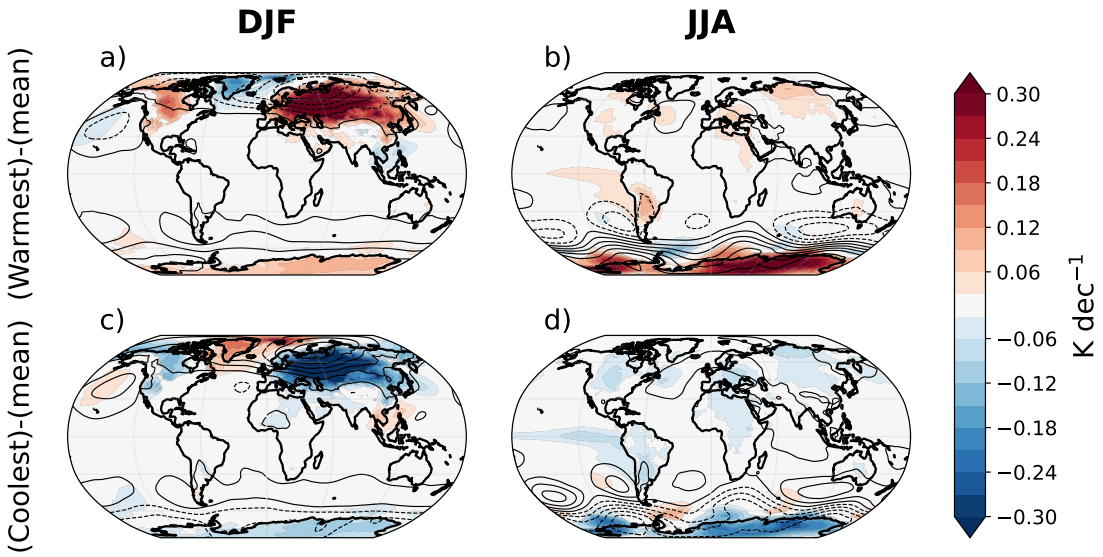


Figure 4.3: **Spatial patterns associated with warmest and coolest hindcast GMST trends.** Difference in SAT and SLP trends for the N=200 1m hindcast members with the warmest (top row) and coolest (bottom row) GMST trends in each season. Differences are shown relative to the 1m hindcast ensemble-mean, as in Figure 3.9. Only significantly different trends are shaded, using a two-tailed t-test as in Figure 3.9. Differences in SLP trend are shown in contours, with solid contours indicating a positive difference.

2. The warmest members are not simply warmer everywhere - for example, in DJF, trends over Greenland and the Arctic are actually *cooler* in the warm subset than in the ensemble-mean (panel a).

3. This reflects the distinctive extratropical circulation patterns which drive the differences in SAT trends and which partly resemble the annular modes: the Northern Annular Mode (NAM, or sometimes referred to as the Arctic Oscillation) and Southern Annular Mode (SAM). Month-to-month, the positive phase of the NAM is associated with warming over Eurasia and North America and cooling over Greenland and the Arctic (panel a), and the negative phase of the SAM is associated with warming of Antarctica (panels a and b). The same temperature and circulation relations are seen here for the hindcast trends, although interestingly the pressure anomaly is stronger over Eurasia than the North Atlantic; consequently, the anomalous SAT trends over Eurasia are associated mainly with the strength of the westerly flow, rather than meridional advection. The coolest members have annular mode trends of the opposite polarity, i.e. negative NAM (panel c) and positive SAM (panels c and d).

4. The spread in warming over land (panel d in Figures 4.1 and A.1) is large for the same reason that monthly variability in hemispheric temperatures is dominated by land: the heat capacity of the land surface is much smaller than the sea surface, and hence its temperature (and the overlying SAT) can adjust much more rapidly in response to changes in the large scale circulation (Wallace et al. 1996). Wallace et al. (1995) demonstrated that this "cold ocean, warm land" phenomenon, associated with the NAM, explains half of the month-to-month variance in NH SAT anomalies over the 20th century.

5. Variability in the annular modes is stronger in the winter than in the summer (e.g. Thompson and Wallace 1998), which explains why the spread in SAT trends over land is largest in the winter hemispheres.

Variability in the 1m hindcast GMST trends therefore appears to be largely related to trends in the annular modes. On interannual timescales, variability in GMST is strongly correlated with ENSO activity. How much do trends in the tropical Pacific contribute to the spread in hindcast GMST trends?

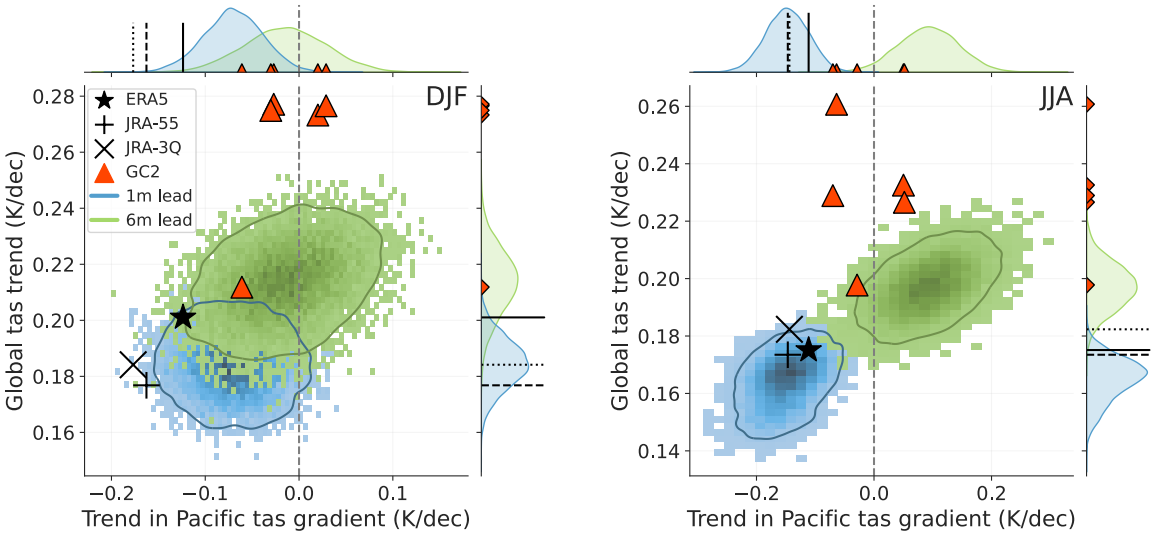


Figure 4.4: **GMST trend as a function of the trend in the zonal SAT gradient in the tropical Pacific,** $\Delta_x TAS$. Shown for DJF (left) and JJA (right). As in Figures 3.4 c-d, solid contours enclose 95% of the hindcast trends at each lead time. $\Delta_x TAS$ is defined as the difference in SAT between boxes in the West Pacific and East pacific as in Karnauskas et al. (2009). DJF trends are calculated over 1981/82-2022/23 and JJA trends are calculated over 1981-2022 as in previous figures.

Table 4.1: Correlation between trends in GMST and $\Delta_x TAS$, as shown in Figure 4.4. All correlations are significant with $p < 10^{-8}$.

corr(GMST, $\Delta_x TAS$)	DJF	JJA
1m lead	−0.06	0.39
6m lead	0.34	0.46

Figure 4.4 shows trends in GMST as a function of trends in the zonal SST gradient $\Delta_x TAS$ in the tropical Pacific; the corresponding trend correlation scores are presented in Table 4.1. In the maps of Figure 4.3, the tropical Pacific appears to contribute to the GMST spread in JJA but not DJF, at least at 1m lead. This is borne out in Figure 4.4, with the $\Delta_x TAS$ trend explaining around 10% of the variance in GMST trends in JJA but virtually none in DJF. The relationships are stronger at 6m lead, as the SSTs evolve away from their initialised state and more El Nino-like trends are possible. Nevertheless, there remains a large scatter between members; while the tropical Pacific contributes to the spread in GMST trends, particularly in JJA, it appears that unrelated trends in the annular modes also contribute. Impacts of the pacific on circulation trends are discussed in greater depth in Chapter 5.

How do the results of this section relate to trends in free-running climate models, or indeed the real atmosphere? The long-term GMST trend is anthropogenically driven, and the modelled response to anthropogenic changes in CO2 and ozone tends towards positive phases of the annular modes in each hemisphere. The externally forced GMST trend is modulated by internally generated variability, with the largest contributions on multidecadal timescales being from decadal variability in the ocean (e.g. in the Pacific; Kosaka and Xie 2013). Superposed on the contributions from external forcing and unforced coupled variability are contributions from seasonal variability of the kind seen in the hindcasts, which is likely mostly atmospheric in origin. The hindcast results are a useful reminder that, while not a dominant factor over longer timescales, variability in the annular modes can noticeably affect GMST trends, particularly over shorter time periods such as the satellite era.

4.1.3 Eurasian cooling

Figure 4.2 shows that winter warming over central Eurasia is reduced in the hindcasts compared to the free-running GC2, although the ensemble-mean trends are still smoother and more positive than the regional cooling in the reanalyses. Several figures in this thesis have qualitatively shown that particular subsets of the hindcast

members - those with strong Arctic warming (Figure 3.9), or those with cool hemispheric GMST trends (Figure 4.3) - have cooler SAT trends over Eurasia. Are any hindcast members able to properly capture the full magnitude of the observed trend? If so, how many – and what do they share in common dynamically?

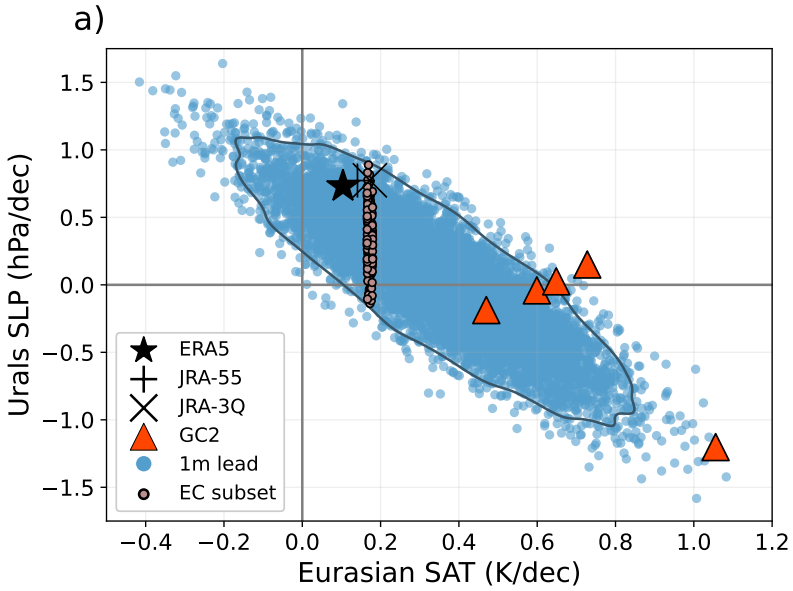


Figure 4.5: **Eurasian temperature and circulation trends in the 1m hindcasts.** Trends in SLP over the Urals, SLP_{Ur}, plotted against Eurasian SAT trends T_{Eur}. The regions used to define each index are shown in Figure 4.6, as defined by Outten et al. (2023). Pink circles indicate the "EC subset", comprising the 200 hindcast members with the closest T_{Eur} trend to the JRA-3Q reanalysis. Other colours and symbols are as in previous figures, although only the 1m hindcasts are plotted here for simplicity.

As the recent Eurasian wintertime cooling is associated with anomalous high-pressure over the Urals, Figure 4.5 plots trends in sea-level pressure (SLP) over the Urals (SLP_{Ur}) against Eurasian SAT trends (T_{Eur}). The two indices are indeed robustly anticorrelated in the 1m hindcasts. The ensemble-mean circulation trend is near zero in the hindcasts, and hence the mean SAT trend is warmer than in the reanalyses. Nevertheless, 11-20% of the hindcasts show a T_{Eur} trend comparable to or cooler than in the reanalyses (the precise fraction depending on the reanalysis product). It is notable that none of the free-running members are close to capturing the reanalysis trends, although confidence is again limited by the small ensemble size. This could be consistent with McCusker et al. (2016), who found that, in a 50 member free-running coupled ensemble of historical simulations, only a single member was able to capture the observed magnitude of Eurasian cooling. Why exactly the

free-running models seemingly struggle more than the hindcasts here is unclear, and could merit further investigation.

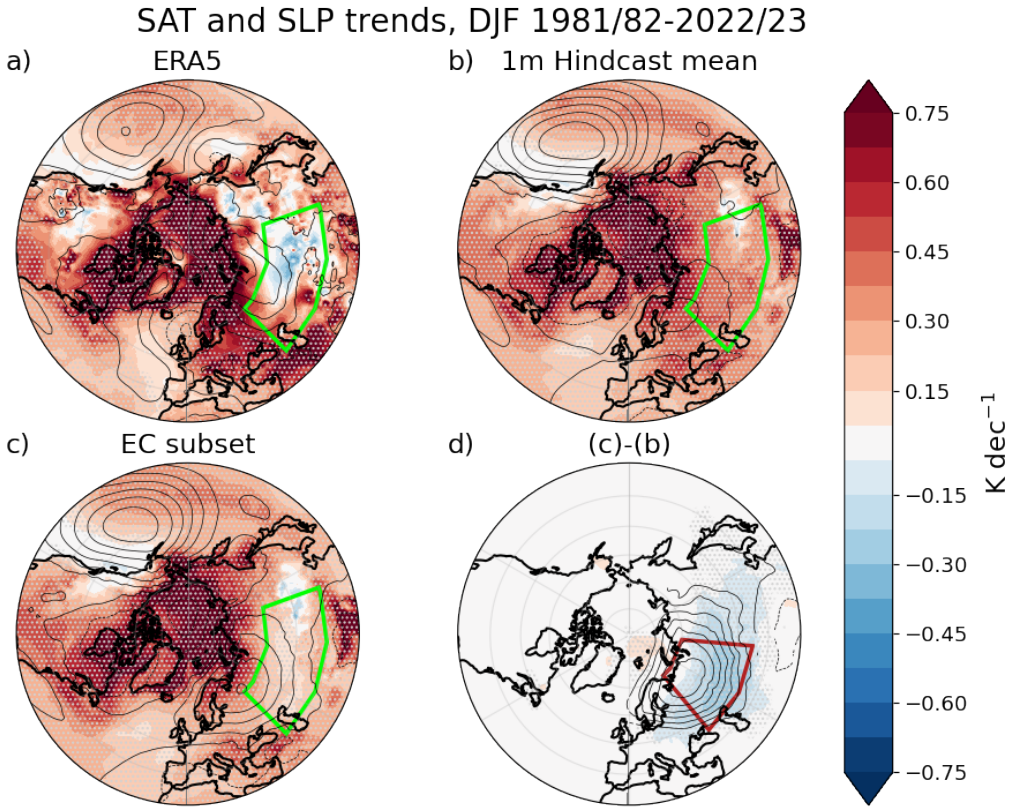


Figure 4.6: **Mean temperature and circulation trends associated with the Eurasian cooling.** DJF SAT (shading) and SLP (contour) trends in (a) ERA5, (b) the hindcast ensemble-mean, and (c) the 200-member EC subset ensemble-mean, defined in Figure 4.5. The difference in trends between (c) and (b) is shown in panel (d). Stippling on (a) indicates significant ERA5 SAT trends at the 5% level; on (b) and (c), stippling indicates where at least 75% of members agree on the sign of SAT trend; and in (d) indicates a significantly different trend at the 5% level, as in Figure 3.9. Only significantly different SLP trends are plotted in (d). The regions used for calculating T_{Eur} and SLP_{Ur} are shown in green and red boxes, respectively.

The pink dots in Figure 4.5 indicate the 200 hindcast members with the closest T_{Eur} trend to the JRA-3Q reanalysis (the "Eurasian Cooling (EC) subset"). The ensemble-mean trends for the EC subset are shown in Figure 4.6, and the difference from the hindcast ensemble-mean in panel (d). As expected, the ensemble-mean trend in the EC subset is cooler over central Eurasia than the hindcast ensemble-mean (green boxes), but it also lacks the localised pockets of strong cooling seen in ERA5 (c.f. panels (c) and (a)). Similarly, although the subset ensemble-mean has a robustly stronger SLP_{Ur} trend (red box), its location is slightly to the west of the

anticyclonic trend in ERA5.

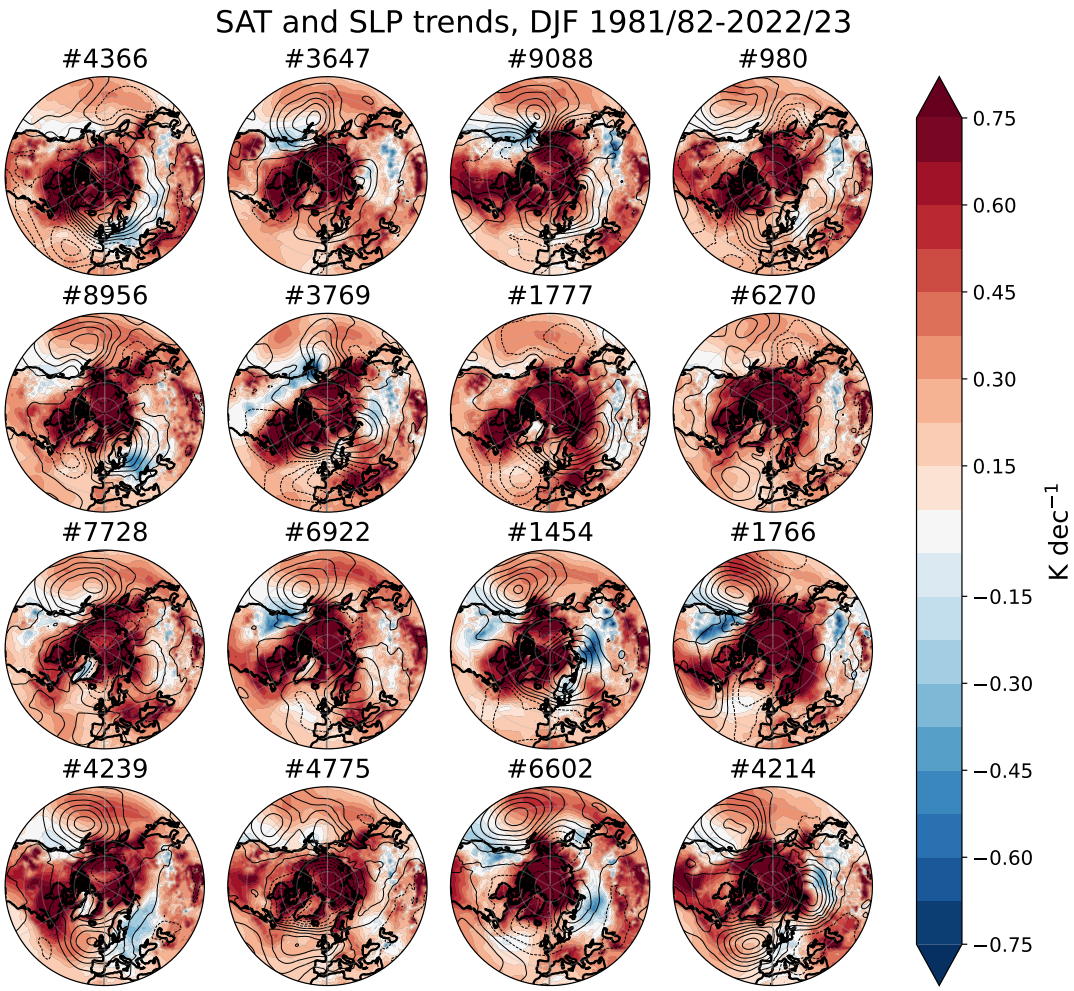


Figure 4.7: **Trends in randomly selected members of the EC subset.** DJF SAT (shading) and SLP (contour) trends in 16 randomly selected members of the 200-member EC subset. Plot titles refer to the member's arbitrary label in the original 10,000-member ensemble.

It is clear from Figure 4.5 that, even within the EC subset, the magnitude of the SLP_{Ur} trend can vary substantially, and the ensemble averaging in Figure 4.6 may be obscuring regional cooling trends that occur at different longitudes in individual members. This is seen in Figure 4.7, which shows trends in 16 randomly chosen members of the 200-member EC subset. Most of these members have regional cooling trends that are comparable in magnitude and spatial extent to the reanalyses, and in some individual members the pressure and temperature trends are approximately collocated with the ERA5 trends (e.g. #3769 on the second row, and #1454 on the third). In other words, an appreciable proportion of the 1m hindcasts are able to

reproduce the overall observed T_{Eur} trend, and trends in individual members feature regions of cooling that resemble the observed trend.

This is a useful demonstration of one of the central points of this thesis: that even in a conditioned subset of hindcast members, not to mention across the whole 10,000 member ensemble, trends in individual hindcast members can vary substantially, and "trend errors" of the hindcast ensemble-mean relative to the reanalysis do not necessarily indicate model error. This is particularly salient for regional trends such as the Eurasian cooling, which are highly sensitive to differences in circulation trends.

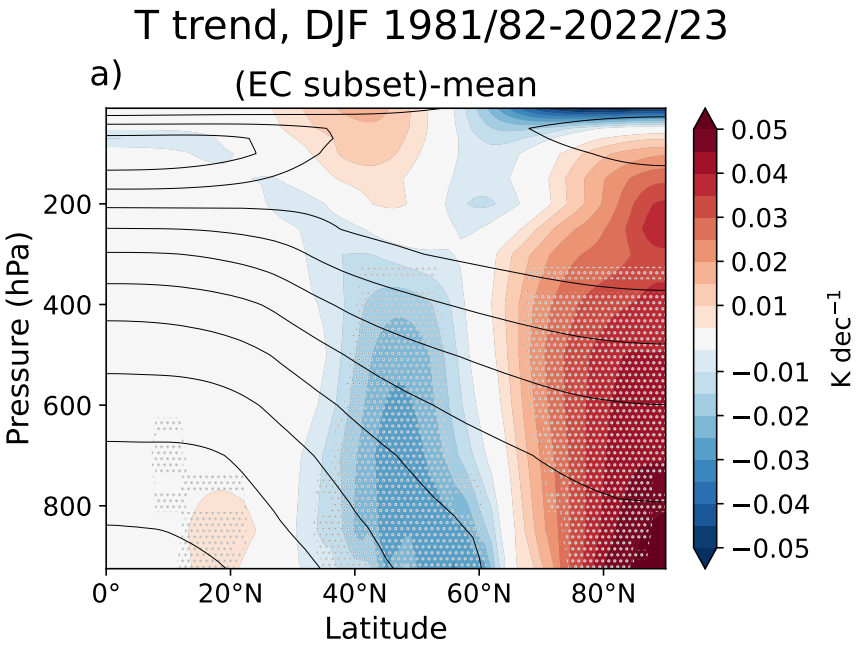


Figure 4.8: **Zonal-mean temperature trend in the EC subset, relative to the hindcast ensemble-mean.** Difference in DJF zonal-mean temperature trends for the N=200 "EC subset", shown in pink in Figure 4.5. Differences are shown relative to the 1m hindcast ensemble-mean. Stippling indicates significantly different trends using a two-tailed t-test as in Figure 3.9.

Finally, while stronger Arctic surface warming is seen in the EC subset members, the vertical extent of the warming is also of interest. As noted in Section 2.3.4.2, He et al. (2020) found that winters with cold SAT anomalies over Eurasia tended to feature deeper Arctic warming extending into the mid-troposphere. Figure 4.8 shows the difference in temperature trends in the EC subset relative to the hindcast ensemble-mean. Trends in these members are warmer not only near the surface, but throughout the depth of the Arctic troposphere; moreover, the significance of the trend difference is greater in the mid-troposphere than directly over the polar

surface. This suggests that the relationship found in variability (He et al. 2020) and in uncoupled model simulations (Ogawa et al. 2018) is also present in the coupled hindcast trend ensemble.

In summary, although the mean T_{Eur} trend in the hindcasts is larger than in reanalyses, between 11-20% of 1m hindcast members have trends of a similar magnitude, and individual ensemble members show regional cooling trends comparable to the reanalyses. Dynamically, the hindcasts reproduce the link between Arctic warming and Eurasian cooling via an anticyclonic trend over the Urals, similar to the annular mode pattern which was shown in the previous section to be responsible for most of the spread in hindcast GMST trends. Finally, members with comparable T_{Eur} trends to the reanalyses show much deeper warming of the Arctic troposphere, confirming the link suggested by Ogawa et al. (2018) and He et al. (2020) in a fully coupled model.

4.2 Sea-level pressure

Previous sections have shown SLP trends in various hindcast subsets, such as those with strong AA (Figure 3.9), extreme GMST trends (Figure 4.3), or anomalously cool Eurasian SAT trends (Figure 4.6). The overall picture has been one of variability: even when a subset exhibits a robust difference in circulation trends relative to the hindcast ensemble-mean, we have also seen that SLP trends vary a lot between members of the subset. This prompts the question: are there *any* regional SLP trends that are robust across the hindcast ensemble?

Figure 4.9 shows ensemble-mean SLP trends in the hindcasts and GC2, as well as in ERA5; in these figures, as in the SAT maps (Figure 4.2), unhatched regions of the model plots indicate where $> 80\%$ of ensemble members agree on the sign of the trend. The NH in ERA5 is characterised by a rich spatial structure of sea-level pressure (SLP) trends (Figure 4.9), although with substantial noise and no gridcells reaching statistical significance after accounting for multiple testing. In some cases, the hindcasts offers support for regional trends in the reanalysis that lack significance: for example, the Aleutian Low robustly weakens at 1m lead in most hindcast members (panels (c) and (d)), and the summer low pressure trend over North America is also broadly consistent with the hindcasts and GC2 (c.f. panel (b) with panels (d), (f) and (h)). In other cases, the ensemble-mean trends differ to the reanalyses, notably over the North Atlantic and Eurasia. As shown in the previous section, however, this need not imply that individual hindcasts members fail to capture the reanalysis

trend, particularly if it owes a lot to atmospheric variability. Trends in the North Atlantic sector are discussed further in the following section.

In the SH, much of the SH tropics and subtropics show a general high pressure trend in the reanalyses in both seasons (Figures 4.9 a and b). The 1m hindcasts support this broad pattern (Figures 4.9 c and d), with most members showing high pressure trends over the Pacific, south Atlantic, and southern Indian Ocean. A striking feature of the reanalysis trend in SH winter is the strengthened SLP gradient in the south Pacific (Figure 4.9 b; Garreaud et al. 2021), though discrepancies have been noted between reanalyses and observations in this region (Swart et al. 2015). At 1m lead, the hindcasts confidently show similar trends to ERA5 in both the subtropical and midlatitude South Pacific, although the ensemble-mean trend is weaker than in ERA5 (Figure 4.9 c). The strengthened SLP gradient here is consistent with a poleward shifted jet, which is responsible for much of the deep zonal wind trend seen in the zonal-mean in Figures 3.10 and 3.11 (not shown).

The reanalysis trend in the SH is more zonally uniform in DJF than JJA, and the hindcast ensemble means at 1m lead reflect this difference (Figures 4.9 a-d). Across longitudes in the SH midlatitudes, the meridional SLP gradient strengthens in DJF, consistent with the poleward jet shift in Figure 3.11 a. Both the hindcasts and free-running model show a similar trend pattern here, with the similarity between the model ensemble-means potentially reflecting their shared ozone forcing.

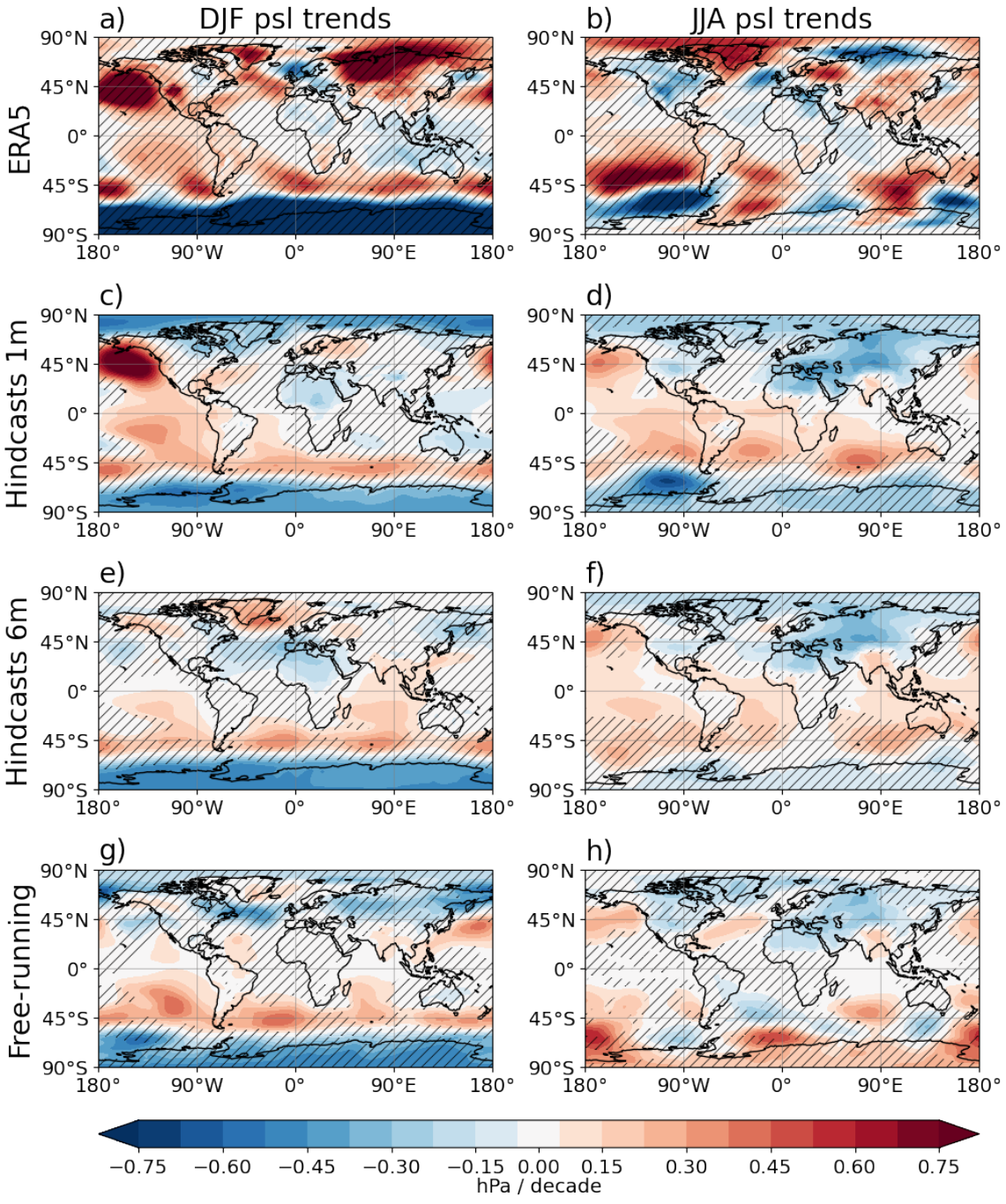


Figure 4.9: **Maps of ensemble-mean SLP trends.** Trends shown for DJF (left column) and JJA (right column). (a)-(b) show the ERA5 trend; (c)-(d) and (e)-(f) show the hindcast ensemble mean trends at 1m and 6m lead respectively; and (g)-(h) show the free-running ensemble mean trends. Hatching hides non-significant trends, as in the left column of Figure 4.2. DJF trends are calculated over 1981/82-2022/23 and JJA trends are calculated over 1981-2022.

4.2.1 Trends in the North Atlantic Oscillation

As discussed in Section 2.1.2.2, models have struggled to capture multidecadal trends in the North Atlantic Oscillation (NAO), particularly for trends that include the mid-20th Century (Eade et al. 2022; Blackport and Fyfe 2022). Are the hindcasts able to capture the satellite era trend?

Hindcast SLP trends in the North Atlantic are not robust across the ensemble, as shown by the hatching in Figure 4.9 (panels (c)-(f)). The winter SLP trend in ERA5 shows more structure, but does not project strongly onto the NAO (Figure 4.9a). Figure 4.10 (a) shows trends in the winter NAO anomaly index. The ERA5 winter NAO index has a weak positive trend between 1981-2022 (p=0.76), falling within the hindcast range at both lead times and even the 5-member uninitialised ensemble. The ensemble-mean trends are weak at both hindcast lead times, and the spread is large such that trends of both sign are clearly possible within the hindcasts. Over the satellite era, both the uninitialised model and the hindcasts are capable of simulating trends in the winter NAO comparable to the reanalyses.

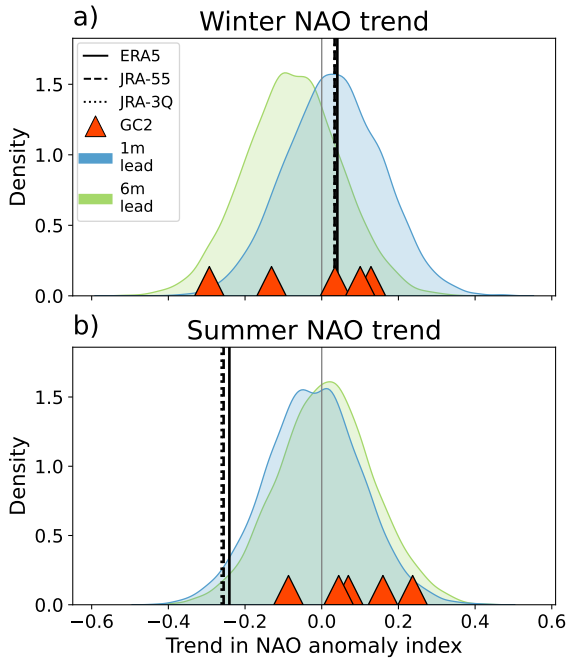


Figure 4.10: **Hindcast trends in the winter and summer NAO**. The winter and summer indices are calculated following Dunstone et al. (2016) and Dunstone et al. (2023) respectively. Colours and symbols are as in previous figures.

Unlike in winter, the reanalysis trend projects strongly onto negative summer NAO (SNAO) (Figure 4.9 (b); Hanna et al. 2015). Ensemble-mean trends in the

hindcasts are weak over the North Atlantic, and are not robust across the ensemble. Neither the 1m nor 6m hindcasts show a strong tendency towards trends of either sign in the SNAO (Figure 4.10 (b)). While ERA5 is within the range of hindcast trends, fewer than 4% of members at 1m lead, and 3% at 6m lead, have as strong a negative SNAO trend. Interestingly, the hindcast SNAO trend – and the wider European/North African summer circulation – appears to be largely insensitive to lead time. The similarity of the model trends here could reflect the poor seasonal predictability of European summer climate (Patterson et al. 2022), and could also point to the importance of shared external forcings such as aerosols (Dong and Sutton 2021).

4.3 Discussion and conclusions

The main results of this chapter are summarised as follows:

1. Surface-air temperature

    - SAT trends improved more over ocean than land due to direct ocean initialisation.
    - See a non-trivial spread of GMST trends in the hindcasts. At short leads, this seems to primarily reflect atmospheric variability in the annular modes, and the greater range of SAT trends over land than ocean. At longer leads, trends in the tropical Pacific also contribute.
    - The hindcasts are able to capture the regional winter cooling trends over Eurasia, which is associated with an anticyclonic trend over the Urals and deep warming throughout the Arctic troposphere.

2. Sea-level pressure

    - A robust weakening of the Aleutian Low is seen in the hindcasts, as well as high pressure trends over much of the tropics and subtropics.
    - Satellite era trends in the NAO are well captured by the hindcasts, although the strongly negative SNAO trend is on the extreme edge of the distribution.

The last two chapters have presented an overview of satellite era trends in the DePreSys3 near-term prediction system, and compared them to five free-running simulations with identical forcings and resolution to the hindcasts. Though confidence

is limited by the small free-running ensemble size, this clean comparison allows a novel isolation of the impact of the initialisation on trends. Initialisation generally improves agreement with reanalysis temperature trends at the surface and in the tropical upper troposphere (Figures 4.1 and A.1, and 3.4 respectively). Differences in upper tropospheric tropical warming can be well explained by the differences in SST warming in convective regions (Figures 3.4 c and d). In light of continuing uncertainty over tropospheric warming rates at low latitudes, the hindcast trends may be a useful complement to existing coupled and uncoupled models.

Trends in the zonal-mean jets are also better captured in the hindcasts, particularly for the winter hemispheres (Figures 3.10 and 3.11). Initialisation minimises development of the model's equatorward winter jet biases at short lead times, coinciding with the improvement in trends (Figure 3.13). The hindcasts also lend support to emerging poleward jet shifts on a global scale (Woollings et al. 2023), and highlight an intriguing anticorrelation between meridional temperature gradients and jet shifts in boreal winter.

The second key aspect of this work is its focus on the spread of trends that are possible within the hindcast ensemble across the climate system (corresponding to the second research question in Chapter 2.7). A wide range of trends is consistent with the hindcasts over much of the climate system. A striking example is Arctic amplification: after just a month of integration, the hindcasts support a remarkably wide range of trends in Arctic amplification over the satellite era (Figure 3.5; 5-95% range of 1.5-3.5) even when coupled patterns such as ENSO and AMV are constrained (Figure 3.1). In line with recent work, this highlights the importance of internal variability for the observed Arctic warming (Ye and Messori 2021; Sweeney et al. 2023; Zhou et al. 2024).

Beyond the Arctic, a wide spread of hindcast trends is seen in jet latitudes, North Atlantic variability, and even global mean surface temperature. These can be interpreted as "unseen trends": trends that could plausibly have occurred given the observed climate state, but which remained unseen due to seasonal and subseasonal variability (Kelder et al. 2020; Berghald et al. 2024). For example, although the hindcast ensemble-means show a poleward jet shift in winter and summer for both hemispheres, in 3 of the 4 cases between 4%-8% of the members show an equatorward shift at 1m lead. This includes austral summer, where 8% of members show an equatorward trend at 1m lead, despite the forced ozone signal common to all members.

Finally, a non-negligible spread is seen in surface temperature trends, even when averaged globally. For the 1m lead hindcasts, the 5-95% trend ranges for DJF and JJA are 0.17K/dec-0.20K/dec and 0.15-0.18K/dec respectively. Though small compared to the ensemble mean warming trends, this still equates to 0.12K of spread over the satellite era, equivalent to approximately a tenth of the observed warming since preindustrial times (Intergovernmental Panel on Climate Change (IPCC) 2023). This surprising result underlines the importance of seasonal variability for trends over the relatively short satellite era, even for those with strong and well-understood forced components. Unlike in a prescribed SST ensemble, this variability can include coupled processes, though is likely largely atmospherically generated at the lead times considered here.

Near-term hindcasts are a largely untapped resource for studying longer term climate. Substantial computation is routinely spent in generating operational systems, often with large ensemble sizes and at high resolution. This chapter suggests that this existing data can be exploited to provide useful context for trends seen in the climate system. We echo calls for CMIP modelling centres to perform seasonal and decadal forecasts, particularly where this may assist in identifying shared nonstationary biases or trend errors (Beverley et al. 2024b). However, caution is necessary when using model ensemble-means in isolation for diagnosing trend errors. A clear message of these chapters is that, for a variety of trends across the climate system, seasonal and subseasonal variability leads to a wide range of plausible trends and cannot be neglected.

Chapter 5

The impact of tropical Pacific SST trends on global and regional circulation

5.1 Introduction

In recent years, evidence has emerged of a trend discrepancy in the tropical Pacific, between coupled models – in which the zonal SST gradient generally weakens – and the observed strengthening of the SST gradient over the satellite era (e.g. Wills et al. 2022). In light of the importance of the tropical Pacific for global climate, considerable attention has been devoted to quantifying the scale and nature of the discrepancy (Karnauskas et al. 2009; Seager et al. 2022; Wills et al. 2022) and to investigations of its potential causes (Seager et al. 2019; Watanabe et al. 2024). As discussed in Chapter 2, explanations of the origin of the discrepancy have broadly fallen into two categories (Lee et al. 2022): an underestimation of internal Pacific variability in coupled models (Chung et al. 2019; Watanabe et al. 2021), or deficiencies in their forced response, such as a too-weak transient ocean thermostat response (Seager et al. 2019; Heede and Fedorov 2023).

Fewer studies have quantified the implications of the Pacific trend discrepancy for global circulation. In the tropics and subtropics, trend biases in the Pacific SSTs have been linked to underestimates of Atlantic tropical cyclone activity in coupled models (Sobel et al. 2023) and errors in model precipitation trends (Qiu et al. 2024). Further poleward, recent work has connected the SST trend discrepancy to an intensification of the austral winter storm tracks (Kang et al. 2024). Deser et al. (2017a) also showed the contribution of tropical Pacific variability to decadal circulation trends over the northern extratropics, focusing on the brief global warming hiatus period during the

early 21st century.

Given the stark discrepancy between models and observations in this key climate "pacemaker" region (Kosaka and Xie 2016), an assessment of its broader implications for the circulation is needed. This chapter quantifies the impact of the observed tropical Pacific SST trend on circulation trends over the satellite era. How has the strengthened SST gradient influenced circulation trends regionally over the Pacific, and in the remote extratropics? Moreover, in simulations with realistic Pacific SST trends, are circulation trends in the extratropics in better agreement with observed changes?

Further motivation is given by projections of the tropical Pacific response to climate change. Coupled models project a future weakening of the SST gradient over the coming century, in contrast to the recent observed strengthening (Lee et al. 2022). If the recent trend is a manifestation of internal variability, then it demonstrates the possibility of multidecadal trends that are counter to the long-term forced response. On the other hand, it could point to deficiencies in the models' transient forced response in the tropical Pacific, or even cast doubt on their future projections. In any case, this underlines the importance of understanding the remote impacts of the Pacific SST trend on the atmospheric circulation.

As coupled models generally struggle to reproduce the observed Pacific SST trend, two complementary methods are used here. The first tool is a large ensemble of plausible trends from the DePreSys3 hindcasts introduced in the previous chapter (Dunstone et al. 2016). In this ensemble, the influence of the Pacific is constrained by forming a composite subset of members with SST trends comparable to the observed trend – the same approach that was used to study the dynamical trends associated with observed winter Eurasian cooling (Section 4.1.3). This subensemble is referred to as the conditioned hindcast subset. The second method is to use Pacemaker simulations, in which SST anomalies in a coupled model are relaxed to observed values over the tropical Pacific.

These two approaches complement each other. In near-term hindcasts, mean state biases are reduced, which could improve the representation of teleconnections (Li et al. 2020). The hindcasts also retain full coupling between model components, which may be important for teleconnection strength from the tropical Pacific (Molteni et al. 2020). However, causality can be difficult to infer from the conditioned hindcast subset – the extratropical trends are consistent with a strengthened Pacific SST gradient, but may not necessarily be caused by it. By directly perturbing the SSTs in the tropical Pacific, the Pacemaker experiments can isolate the causal impact of the Pacific

SST trend on the wider circulation. The Pacemaker approach has been successfully employed to investigate the Pacific influence on global temperature and regional circulation (Kosaka and Xie 2013; Deser et al. 2017a). Broad agreement between these two very different methods builds confidence in our conclusions.

This chapter primarily focuses on the DJF season. Section 5.2 details the models used. The DJF model response is discussed in Section 5.3.1, and comparison is made to reanalyses in Section 5.4. A selection of results for the JJA season are presented briefly in Section 5.5. Discussion and conclusions follow in Section 5.6.

5.2 Methods

5.2.1 Hindcast data

Hindcast data is used from the DePreSys3 decadal hindcasts (Dunstone et al. 2016). As in previous chapters, an ensemble of 10,000 plausible trends is generated from the hindcasts – see Chapter 3 for a description of the model and the ensemble generation procedure. As before, the "hindcast ensemble" refers here to the distribution of 10,000 trends, rather than the 40 independent members initialised per start date. However, the ensemble differs from the previous chapters in that longer lead times are used here (months 13-15), as discussed at the end of this subsection.

To investigate the impact of the tropical Pacific SST trend, a conditioned subset of hindcast members is formed, whose trends in the tropical Pacific are closest to the observed trend. The target region is defined as the tropical East Pacific, using the same region as the Pacemaker nudging (the wedge region in Figure 5.2). This region is relatively meridionally broad, extending up to 20 degrees either side of the equator. Seager et al. (2022) note that the region of diminished warming in observations is confined to the equator, in contrast to the broader SST patterns associated with Pacific variability (e.g. ENSO and the PDO). The broad target region used here is therefore necessary to condition the hindcasts on the distinctive meridional structure of the observed trend.

In order to select members for the conditioned set, a metric is required that represents how similar a given member's SST trend is to the observed trend. A root-mean-squared (RMS) error metric is calculated by taking the RMS difference between the hindcast and observed trend at each gridpoint in the target domain. For consistency, the Pacemaker SSTs are used as the observed trend in the target region, as these are nudged to the observed ERSSTv5 anomaly field (see the following section). An overall score for each hindcast member is calculated by averaging the

RMS differences over the target region, weighting by gridcell area. Gridcells in the Pacemaker buffer regions are additionally weighted by the same forcing value as in the Pacemaker experiments before the average is calculated. The N=500 members (i.e. 5% of the 10,000 member ensemble) with the lowest RMS difference relative to the observed trend are used to form the Pacific-conditioned ensemble.

The choice of ensemble size for the conditioned hindcast ensemble involves a balance between competing considerations. On the one hand, averaging the response over larger samples minimises noise and allows stronger statistics; on the other, a larger sample will also include members with a broader range of Pacific trends, potentially diluting the Pacific-forced signal (Section 5.2.6). Sensitivity tests were performed with a conditioned ensemble size of N=50. These confirmed that the pattern of the significant responses was substantively the same as with the larger ensemble; though the response was somewhat stronger, the greater extratropical noise complicated the interpretation of the midlatitude trend response. Results are thus shown here for the larger ensemble size.

Previous chapters have used hindcast data from months 2-4 and 7-9 of the simulations. At these lead times, the hindcasts have high skill and little ensemble spread in the tropical Pacific. In order to define a meaningfully different subset, it was found that a broader ensemble spread was needed, which is achieved by exploiting data at longer lead times. This chapter therefore uses data from months 14-16 of the runs initialised in November, corresponding to the second DJF of each simulation.

5.2.2 Pacemaker experiments

Output data from the 10-member CESM2 Pacific Pacemaker ensemble is analysed (Rosenbloom et al. 2022). In these runs, SST anomalies are relaxed to observed ERSSTv5 anomalies (Huang et al. 2017) over the tropical central and eastern Pacific; trends in this wedge-shaped region are shown in Figure 5.2. Outside this region, the coupled model is free to evolve. Historically observed external forcing is prescribed in the Pacemaker runs until 2014, and they follow SSP3-7.0 until 2019. The experiments use the CAM6 configuration of CESM2 (Danabasoglu et al. 2020), run at a nominal 1° horizontal resolution in the atmosphere and 32 vertical levels, with a top height of approximately 40km (a low top model).

The Pacemaker simulations are compared with the equivalent freely evolving (non-relaxed) CESM2 large ensemble (CESM2-LE). An ensemble of 50 members is analysed at the same grid resolution and subject to the same external forcings as the Pacemaker runs, in order to isolate the impact of the Pacific SSTs.

5.2.3 Observational and reanalysis data

Model trends are compared to those from three reanalysis datasets: ERA5 (Hersbach et al. 2020), JRA-3Q (Kosaka et al. 2024), and MERRA2 (Gelaro et al. 2017). Precipitation data comes from GPCP, which blends satellite retrievals over ocean and gauge-based observations over land (Huffman et al. 2023).

5.2.4 Trend analysis

All model and reanalysis data is retrieved at monthly temporal resolution. As in the previous chapter, the monthly data is averaged to give DJF seasonal means; JJA is discussed in Section 5.5. Trends are calculated over DJF 1981/82-2018/19, as this is the common time period for which both Pacemaker and hindcast data are available. Note the slightly shorter time period than in the previous chapter, which considered DJF trends over 1981/82-2022/23.

5.2.5 ENSO composites

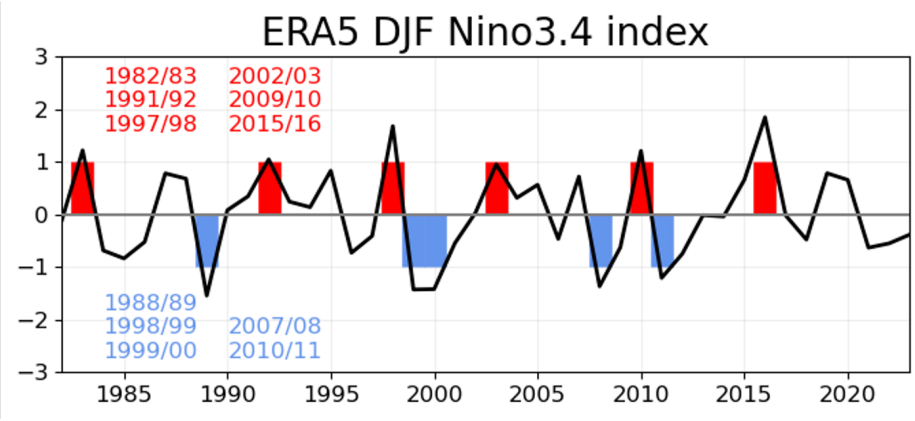


Figure 5.1: **Definition of ENSO events used for composites.** Detrended DJF-mean Nino3.4 index in ERA5, following Deser et al. (2017b). Red bars indicate El Nino events and blue bars La Nina, following the criteria defined in the text. The 6 El Nino and 5 La Nina events in this period are shown in red and blue text respectively.

The Pacific-forced trend response is compared to the interannual response to ENSO. El Nino and La Nina events are defined in ERA5 following Deser et al. (2017b). Briefly, the monthly Nino3.4 timeseries is calculated (SST anomalies are averaged over ±5°N, 120°–170°W, Bamston et al. 1997) and averaged over DJF, before being detrended. El Nino and La Nina events occur when the detrended seasonal index exceeds one standard deviation above or below zero, respectively. Figure 5.1

shows the 6 El Nino and 5 La Nina seasons as defined using this metric. Composites in the hindcasts and Pacemaker runs use these same years for consistency. The composite responses are of La Nina minus El Nino (LN-EN).

5.2.6 Performance of the conditioned hindcasts

Figure 5.2 show the ensemble mean SST trends over the tropical Pacific in each of the models. The unconditioned models (Fig 5.2 b, d) both show enhanced warming of the cold tongue region, while the East Pacific trends are weak or negative in the conditioned models (Figs 5.2 a, c). The conditioned hindcasts capture much of the structure of the Pacemaker trend, particularly the region of cooling south of the equator; in the NH, the zonal gradients are much weaker in the conditioned hindcasts. The mean trend in the conditioned hindcasts is a strengthening of the east-west zonal SST gradient, in contrast to the mean weakening of the gradient in the unconditioned hindcasts (Fig 5.2e).

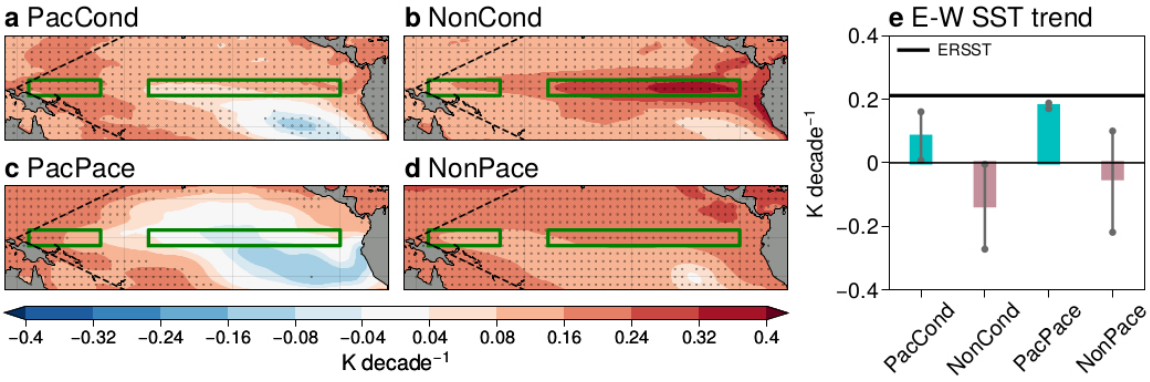


Figure 5.2: **Ensemble-mean tropical Pacific SST trends in the conditioned and unconditioned models.** SST trends in the conditioned (a) and unconditioned hindcasts (b), and in Pacemaker (c) and CESM2-LE (d). The Pacific wedge region that is relaxed in the Pacemaker runs is enclosed by the black dashed lines. Stippling indicates statistically significant trends at the 5% level. (e) Trends in the zonal SST gradient in the conditioned and unconditioned models, and as observed in ERSSTv5 (black line). The SST gradient is defined as the difference in SST between boxes in the East Pacific ($\pm 3°$N, 140°–170°E) minus a west Pacific box $\pm 3°$N, 170°–90°W following Seager et al. (2022).

However, it is important to stress that the conditioning is imperfect. Two important conclusions can be drawn from Figure 5.2. Firstly, the magnitude of the Pacific trend is somewhat weaker in the conditioned hindcasts than in the Pacemaker runs, particularly north of the equator. Secondly, the *spread* in trends is much wider

in the conditioned hindcasts than in the Pacemaker ensemble, with trends of both sign seen in the SST gradient. We can therefore expect a weaker and more varied response to the Pacific SST trend in the conditioned hindcasts than in the Pacemaker runs. These features reflect both the less stringent conditioning method compared to the Pacemaker SST relaxation, but also a strength of the hindcast method: namely, the much larger ensemble size. Qualitative agreement between the two conditioned models despite their differences will therefore build confidence in the results.

5.2.7 Nomenclature

The Pacific-conditioned DePreSys3 hindcast subset is denoted "PacCond" in the figures that follow, as distinct from the remaining non-conditioned hindcasts ("NonCond"). The Pacific Pacemaker simulations are labelled "PacPace", whereas the CESM2-LE historical experiment (no pacemaking) is "NonPace". Where possible, the text will use the long form names (e.g. conditioned hindcasts, or simply Pacemaker) for the sake of clarity. When Pacemaker and the conditioned hindcasts are referred to together, they may be termed "conditioned models".

5.3 Trend response

5.3.1 Zonal-mean response

Figure 5.3 shows the ensemble-mean trends in the unconditioned models: the unconditioned 12m hindcasts on the left and CESM2-LE on the right. Tropical warming in CESM2-LE (NonPace) is strongly amplified with height, as in most coupled models (e.g. Mitchell et al. 2020, and Chapter 2.2). Figure 5.3(a) shows that, by 12m lead, the same is true of the hindcasts, despite the improvement at shorter lead times (the 1m and 6m hindcasts of Chapter 3.3). The ensemble-mean wind trends differ more between the models. In CESM2-LE, a poleward jet shift is seen in both hemispheres, whereas trends on the equatorward sides of the jets are substantially weaker in the hindcasts. The ensemble-mean wind trends at 1m lead look broadly similar to the trends at 6m lead and lack the strong poleward shifts seen at 1m lead (Chapter 3.5).

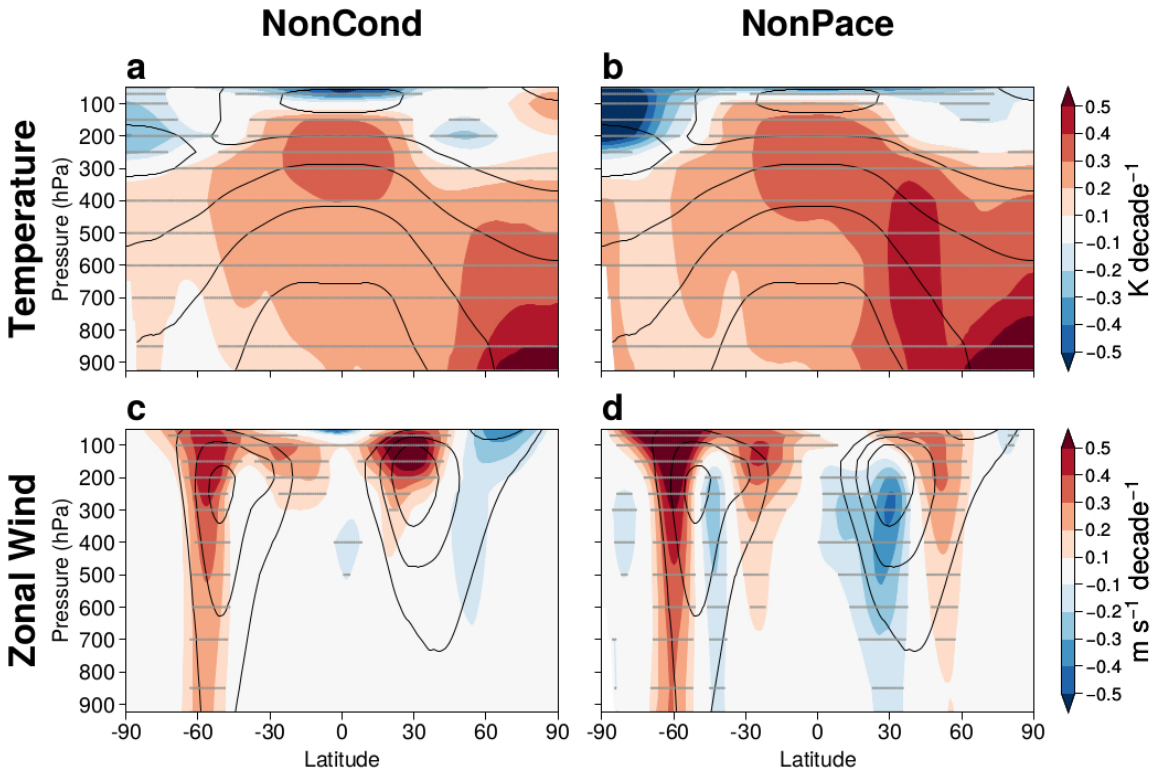


Figure 5.3: **Ensemble-mean, zonal-mean temperature and zonal wind trends in the unconditioned models**. The unconditioned hindcasts are in the left column (a, c) and CESM2-LE the right column (b, d). Trends are in the DJF seasonal means, calculated over 1981/82-2018/19 inclusive. Solid contours indicate the DJF climatology over the same period. Stippling indicates a statistically significant trend at the 5% level, using a two-tailed t-test.

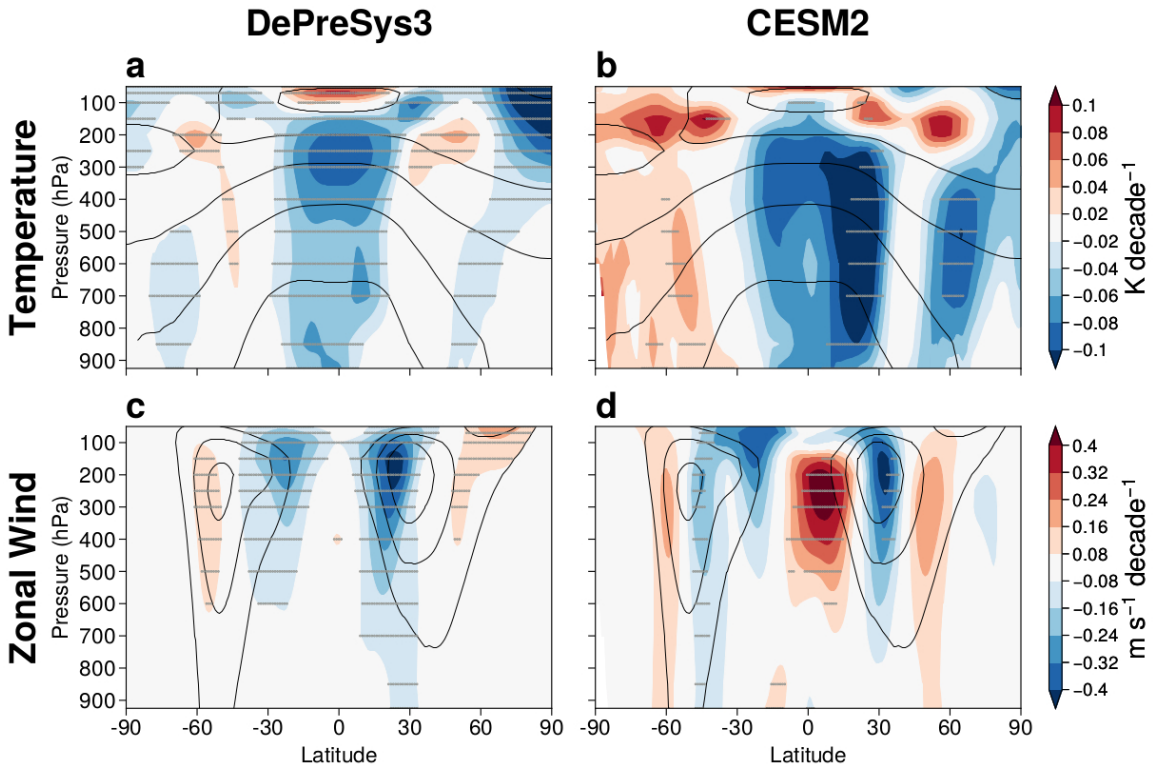


Figure 5.4: **The Pacific-forced zonal-mean trend response.** The trend response is defined as the ensemble-mean trend in the conditioned models minus the ensemble-mean trend in the corresponding unconditioned model (Figure 5.3): Conditioned minus unconditioned hindcasts (left column) and Pacemaker minus CESM2-LE (right column). As before, trends are in the DJF seasonal means, calculated over 1981/82-2018/19 inclusive. Stippling indicates a statistically significant trend at the 5% level, using a two-tailed t-test.

Figure 5.4 shows the trend response in the zonal-mean temperature and zonal-winds, defined as the difference in ensemble-mean trends between the conditioned and non-conditioned models. Many aspects of the Pacific-forced trend response are consistent between the models. A meridionally narrow cooling response is seen in the tropics (Figure 5.4 c, d), in contrast to the broad warming extending into the subtropics characteristic of the response to GHG forcing (Lu et al. 2008). In the conditioned hindcasts and in the NH of the Pacemaker response, vertical bands of warming and cooling are seen in the extratropics, steepening the meridional temperature gradient in the midlatitudes. Correspondingly, the zonal winds weaken on the subtropical side of the jets through thermal wind balance (Seager et al. 2003).

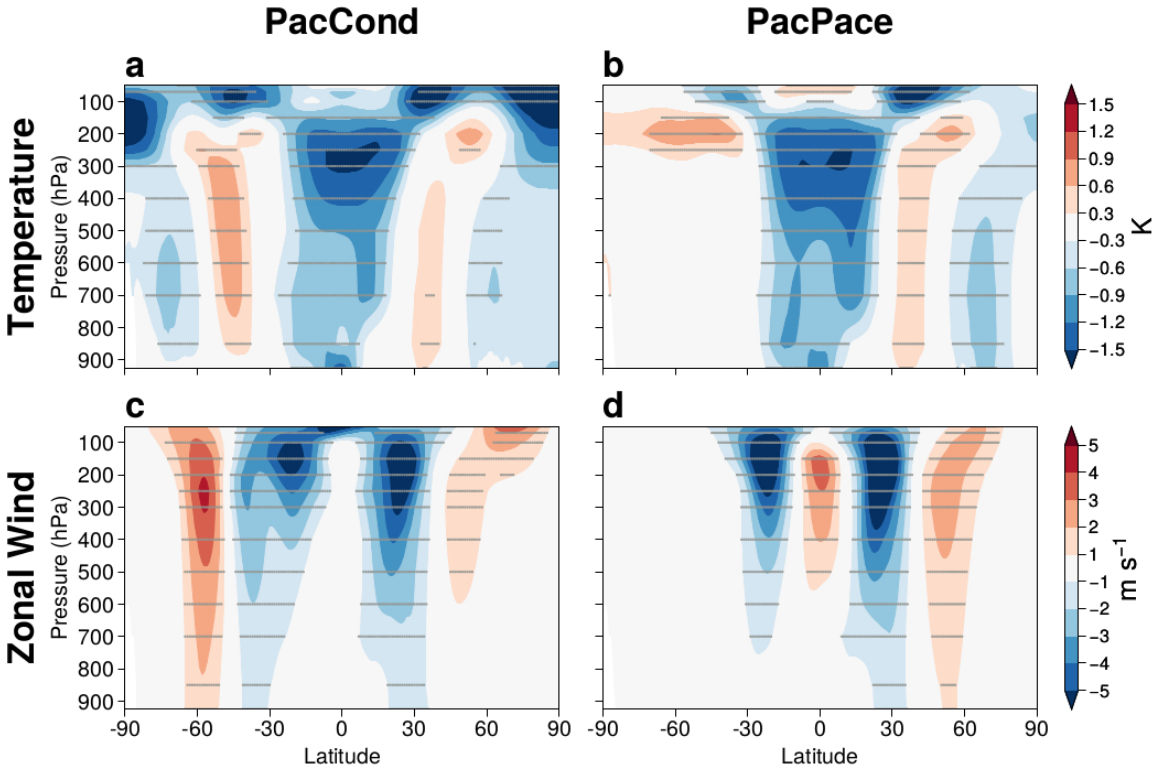


Figure 5.5: **Interannual zonal-mean DJF La Nina minus El Nino response.** Shown in the hindcasts (left) and Pacemaker (right), for temperature and zonal winds (top and bottom rows respectively). LN and EN years are defined in Figure 5.1. Stippling indicates a significant response at the 5% level using a two-tailed t-test.

These trend responses are immediately recognisable as resembling the interannual response to La Nina (Figure 5.5), which suggests a likely mechanism for the extratropical response (Seager et al. 2003). On interannual timescales, the weakening and poleward shift of the subtropical jets leads to a poleward shift in the region of strongest baroclinicity and eddy generation. The propagation of eddies out of this

region converges momentum in the midlatitudes, leading to westerly anomalies in the mid-high latitudes of each hemisphere (Seager et al. 2003; Schneider and Sobel 2021). Although lacking the high frequency data to diagnose it directly, such a mechanism would appear to be consistent with Figure 5.4; the deep midlatitude westerly response, particularly visible in Figure 5.4 (b), indicates a transient eddy-mediated response. In addition, the trend response is largely symmetric about the equator, which is also characteristic of the interannual response to ENSO (Seager et al. 2003).

The conditioned hindcast trend response is weaker overall than in Pacemaker, reflecting the weaker tropical Pacific forcing (Figure 5.2). Some differences are also seen between the response patterns of the two models. In the SH, a weak warming is seen in the mid-high latitudes in Pacemaker, in contrast to the cooling response in the conditioned hindcasts. In addition, the Arctic polar vortex strengthens in the conditioned hindcasts while little response is seen in Pacemaker. Modelling evidence supports a weakening of the vortex and increased occurrence of sudden stratospheric warmings during El Nino (Ineson and Scaife 2009; Domeisen et al. 2019), which may be consistent with the strengthened vortex seen in the conditioned hindcasts. The difference in response between the models likely reflects the higher model top and better resolved stratosphere in DePreSys3 than in CESM2-LE.

The Pacemaker response also shows a substantial acceleration of the zonal-mean zonal winds at upper levels in the tropics (500-200hPa). This zonal-mean trend is dominated by strengthened westerly flow in the upper branch of the Pacific Walker cell. The zonal-mean response in the hindcasts is much weaker due to compensating easterly trends outside the tropical Pacific (shown in the following section).

5.3.2 Regional response

This section details the spatial structure of the Pacific-forced trend response. Trends in the unconditioned models are shown in Figure 5.6 for SLP, (a, b), SAT (c, d), and precipitation (e, f). The unconditioned models both show increased precipitation over the Central and Eastern Pacific. In the extratropics, the ensemble-mean trends differ, with oppositely signed SLP trends in the North Pacific and North Atlantic.

Figure 5.7 shows the trend response for the same variables plus the 200hPa zonal winds. The trend response is defined as the ensemble-mean trend in the conditioned model minus the ensemble-mean trend in the unconditioned model.

Although the impact is strongest in the Pacific region, Figure 5.7 shows that the Pacific SST trend influences trends on the global scale. In the following sections, the circulation response is discussed by region, and loosely by the robustness of the

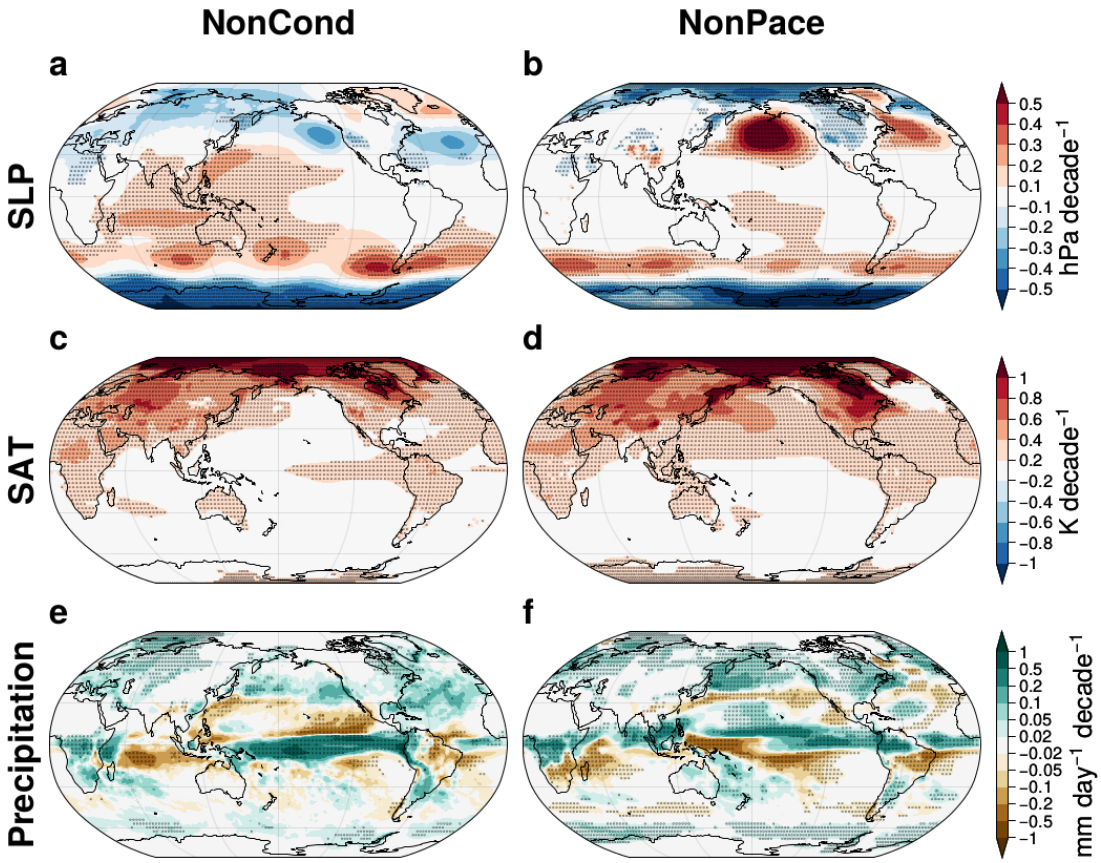


Figure 5.6: **Ensemble-mean spatial trends in the unconditioned models.** Trends in sea-level pressure (top), surface-air temperature (middle), and precipitation (bottom) for the unconditioned models: the unconditioned hindcasts (left column: a, c, e) and CESM2-LE (right column: b, d, f). Trends are in the DJF seasonal means, calculated over 1981/82-2018/19 inclusive. Stippling indicates a statistically significant trend at the 5% level as in Figure 5.3.

signal: trends over the tropical Pacific itself (very good agreement between methods); the extratropical Pacific and Americas (moderate agreement between methods); and finally, the North Atlantic-European sector (mixed agreement between methods).

5.3.2.1 Tropics and subtropical Pacific

The tight coupling between the SST trend and the tropical circulation over the low-latitude Pacific is apparent in both sets of conditioned runs. Over the central and Eastern tropical Pacific, the cooler SST trends are associated with decreased precipitation (Figures 5.7 e, f) and higher pressure (Figures 5.7 a, b). A deepening of the low pressure over the west Pacific warm pool is seen in the conditioned hindcasts, though the Pacemaker response here is weaker. At upper levels, a westerly acceleration is

seen over the tropical Pacific (Figures 5.7g, h), associated with a strengthening of the Pacific Walker Circulation (discussed further in Section 5.4.3).

Outside the Pacific, the tropical response is generally similar in both models. Good agreement is seen in precipitation trends over much of the tropics, including positive trends over tropical South America, Southern Africa, and much of the Indian Ocean, and drying trends over equatorial Africa and the horn of Africa (Figure 5.7 e, f). On the poleward edge of the tropics, the subtropical jets weaken over the Pacific (Figures 5.7g, h), and at all longitudes in the conditioned hindcasts.

The trend response described above is essentially the inverse of the ensemble-mean trend in the unconditioned runs, characterised by warming and wetting of the Central-Eastern Pacific, an upper-level easterly trend, and a weakening of the Pacific Walker circulation (Figure 5.6). In the subtropics, the unconditioned models show either asymmetric subtropical jet trends (CESM2-LE, which weakens in the NH but strengthens in the SH) or trends that are opposite to the conditioned runs (hindcast mean, with strengthened subtropical jets in both hemispheres). Unsurprisingly, the Pacific SST trend is shown to be a leading order control on circulation trends in the tropical and subtropical Pacific.

As in the zonal mean, the trend response largely resembles the interannual La Nina response in most respects (Figure 5.8). One exception is the U200 winds in the Pacemaker experiments (Figure 5.8g, h). Here, we see that the trend response is tilted northwest-southeast, as in the hindcasts, while the Pacemaker La Nina response is more zonally oriented. Previous work has noted a similar tilting bias in the North Pacific ENSO response of five different seasonal forecast models (Williams et al. 2023).

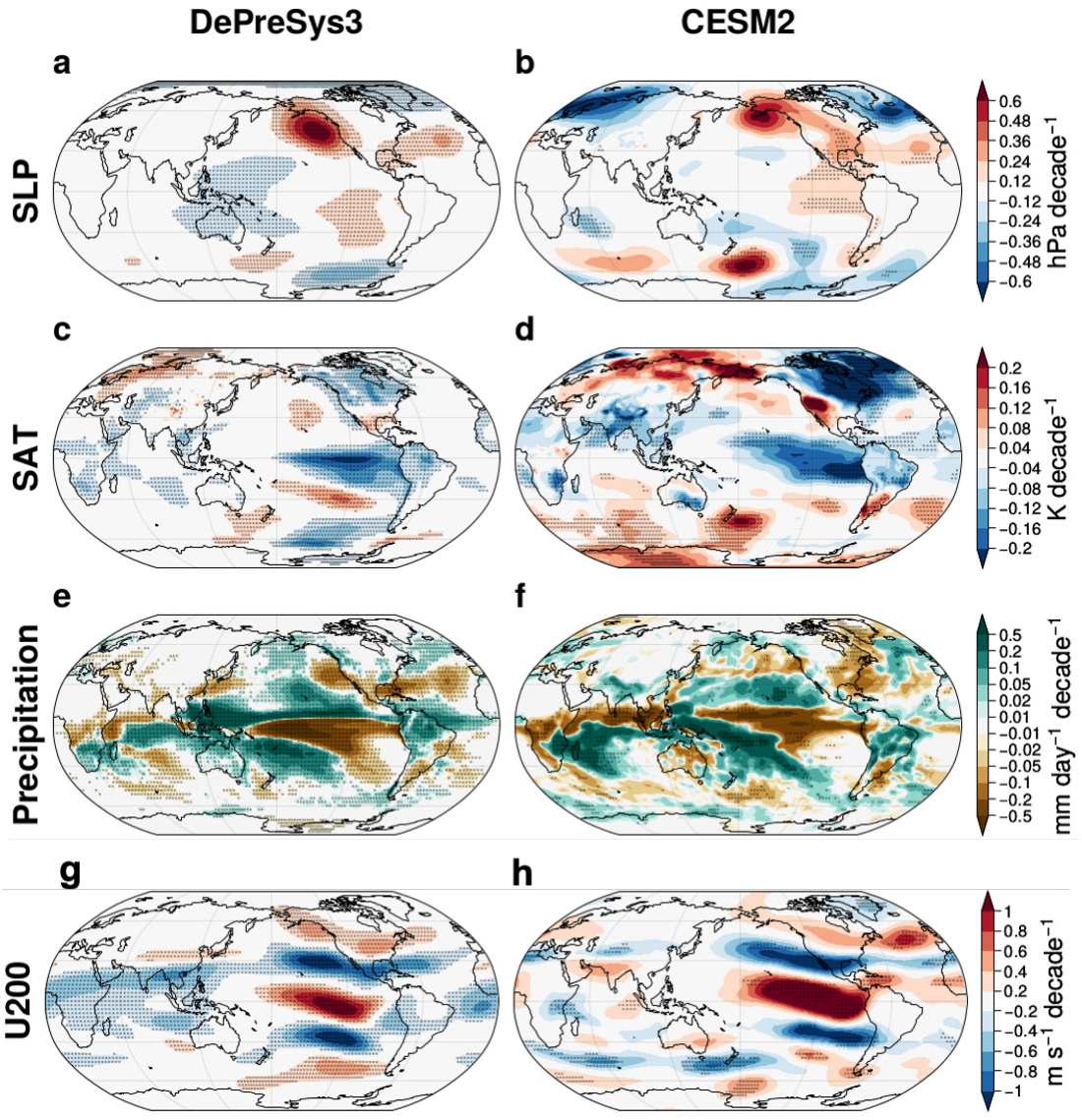


Figure 5.7: **The Pacific-forced spatial trend response.** As before, the trend response is defined as the ensemble-mean trend in the conditioned models minus the ensemble-mean trend in the corresponding unconditioned model: Conditioned minus unconditioned hindcasts (left column) and Pacemaker minus CESM2-LE (right column). As before, trends are in the DJF seasonal means, calculated over 1981/82-2018/19 inclusive. Stippling indicates a statistically significant trend response at the 5% level as in Figure 5.4.

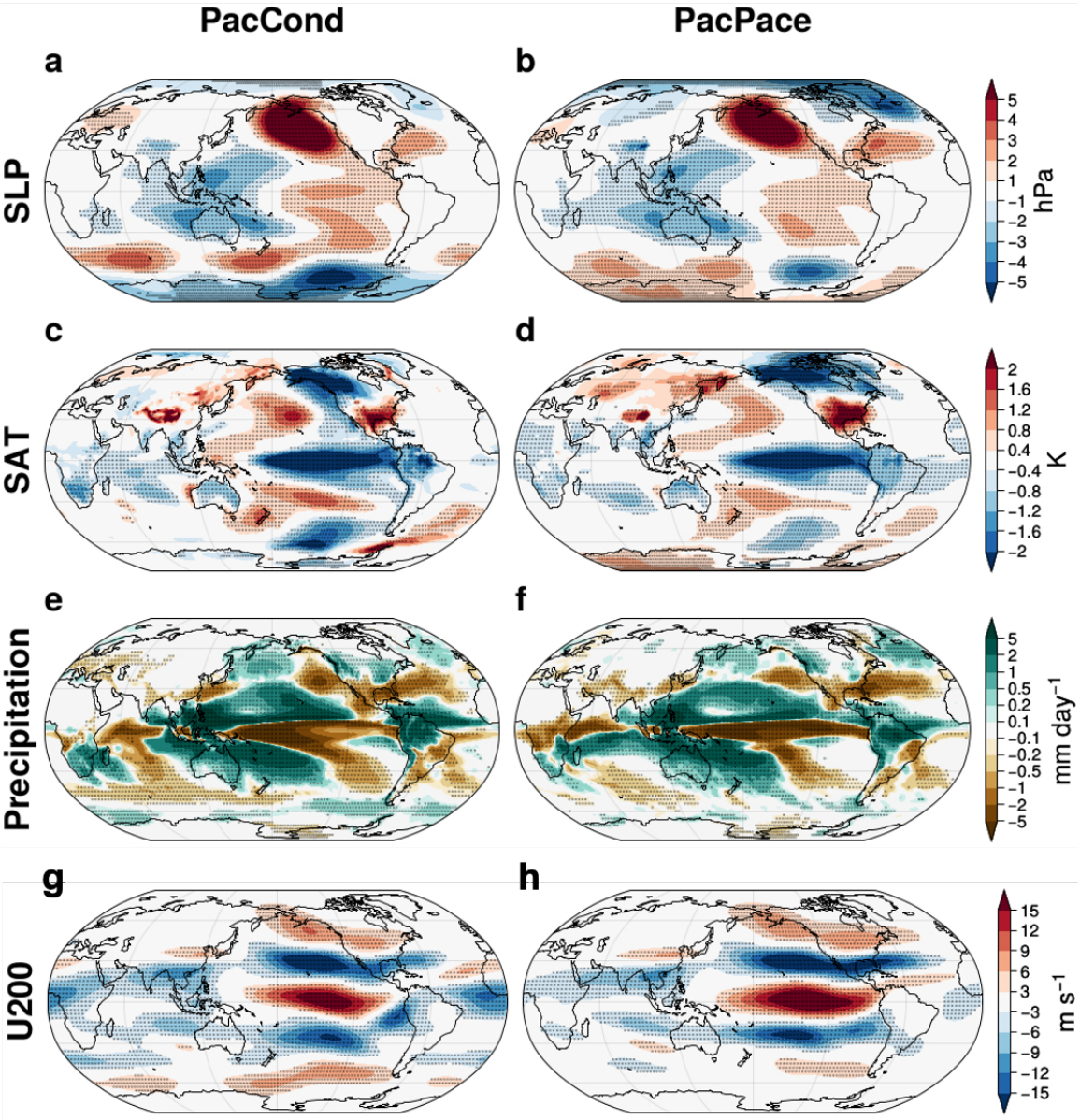


Figure 5.8: **Interannual DJF La Nina minus El Nino spatial response.** Interannual DJF La Nina minus El Nino response in the hindcasts (left) and Pacemaker (right), for the same variables as Figure 5.7. LN and EN years are defined in Figure 5.1. Stippling indicates a statistically significant response at the 5% level as in Figure 5.5.

5.3.2.2 North Pacific and Americas

Trends in the unconditioned models differ over the North Pacific and North America. The Aleutian Low (AL) weakens in CESM2-LE but deepens in the unconditioned hindcast mean (though the hindcast trend is not robust across the ensemble, as indicated by the lack of stippling; Figures 5.6 a, b). Over North America, a low-pressure trend is seen in CESM2-LE, while the hindcast trend is weak.

A robust trend response is also seen over the North Pacific and North America. Once again, much of the trend response is similar between the models, and largely resembles the response to La Nina (Figure 5.8). A weakening of the AL is seen in both models, and cooling over the northern half of the continent is separated from warming and drying trends over the south/southeastern US and Mexico (Figures 5.7 a, d). These divergent trends are dynamically driven by a poleward shift of the jet (Figures 5.6 g, h). The trend responses in pressure, winds, and precipitation, are plotted on the same scale as the corresponding unconditioned trends; that is, the Pacific-forced trend response has a leading order effect on circulation trends over this region (c.f. colour scales on Figures 5.6 and 5.7).

Differences between the Pacific models are subtle. While a broad separation of southern warming and northern cooling responses is seen in each, the warming trend response in Pacemaker is concentrated further to the west. This reflects the high-pressure trend which extends further across North America in Pacemaker. As in the zonal mean, the magnitude of both responses is generally larger in Pacemaker than in the conditioned hindcasts. One possibility for this is the stronger Pacific SST trend in Pacemaker, as was suggested to explain the differing zonal-mean responses. Another is differences in the zonal structure of the tropical forcing; while the overall forcing appears weaker in DePreSys, the response in the west Pacific is larger than in Pacemaker (e.g. compare Figures 5.7 a and b, or 5.7 e and f). The Rossby wave forcing may therefore be expected to differ between models, which could explain differences in the zonal structure of the extratropical responses.

5.3.2.3 North Atlantic and European sector

Similar ensemble-mean trend responses are also seen in the North Atlantic circulation. In both models, the ensemble-mean North Atlantic trend response is broadly consistent with the positive phase of the NAO (Figures 5.7 a, b), although the Pacemaker response is stronger and extends further east. Again, the trend response follows the

DJF-mean La Nina response in each model (Figures 5.7 and 5.8), which itself resembles the late winter La Nina teleconnection (consistent with the strong Aleutian response; Ineson and Scaife 2009; Molteni and Brookshaw 2023). The agreement between Pacemaker and the conditioned hindcasts in the North Atlantic is perhaps surprising given the weak stratospheric response in the Pacemaker runs, pointing to the importance of tropospheric teleconnection pathways (Jiménez-Esteve and Domeisen 2020).

Downstream of the Atlantic sector, the trend response is less robust between models. Some high latitude warming is seen in the Eurasian response of both models, but the warming is confined further to the west in the conditioned hindcasts. The dynamical drivers appear to be different too; the Pacemaker warming is associated with the eastward extension of the North Atlantic low-pressure trend, whereas no robust trend response is seen in the hindcast circulation. Notably, neither model shows an anticyclonic trend over the Urals, of the kind that is linked to the recent Arctic warming and Eurasian cooling trends (Outten et al. 2023, and Section 4.1.3).

Links between the tropical Pacific and Arctic climate have been explored by Lee (2014), who suggests that zonally asymmetric tropical heating – as in La Nina states – can trigger poleward propagating Rossby waves which heat the high latitudes. Although some Arctic warming is seen in the ensemble-mean responses, it is weak in the conditioned hindcasts and lacks significance in Pacemaker. Over the Barents Sea, the ensemble-mean response in the conditioned hindcasts is around 0.05K/dec, whereas the strongest response in Pacemaker is around 0.16-0.20K/dec, just north of Siberia (Figures 5.7 c and d respectively). Given the total Arctic SAT trend is of order 1K/dec(Figure 5.6 c, d), and in light of their divergent circulation responses over Siberia, neither model supports a dominant role for the Pacific in recent wintertime Arctic warming and Eurasian cooling.

Summary of trend response

This section has shown that the Pacific-forced trend response is similar in the conditioned hindcasts and Pacemaker, and that the response can be well understood using interannual La Nina teleconnection theory. The response is generally somewhat stronger in Pacemaker than the conditioned hindcasts, potentially reflecting the weaker tropical forcing in the latter. The magnitude of the response is greatest over the Pacific and Americas, but a robust NAO+ response is also seen over the North Atlantic. Further downstream, the response is less consistent over Eurasia,

and the models do not support a large role for the Pacific in recent Arctic warming and Eurasian cooling trends.

5.4 Comparison with reanalysis

The previous section outlined the trend response within the models. In this section, it is shown that accounting for the Pacific SST trend improves agreement with reanalyses for a range of trends that the unconditioned models otherwise struggle to capture.

5.4.1 Tropical temperatures

Figure 5.4 showed that the Pacific SST trend is associated with a cooling of the tropical troposphere. The trend profile averaged over the tropics is shown in Figure 5.9. Coupled models generally overestimate observed and reanalysis tropical warming, as discussed in Section 2.2. This is the case for the unconditioned models here, with all three reanalysis trends lying almost entirely outside the 10-90% range of unconditioned model trends, throughout the troposphere (the black lines lie outside the pink shading).

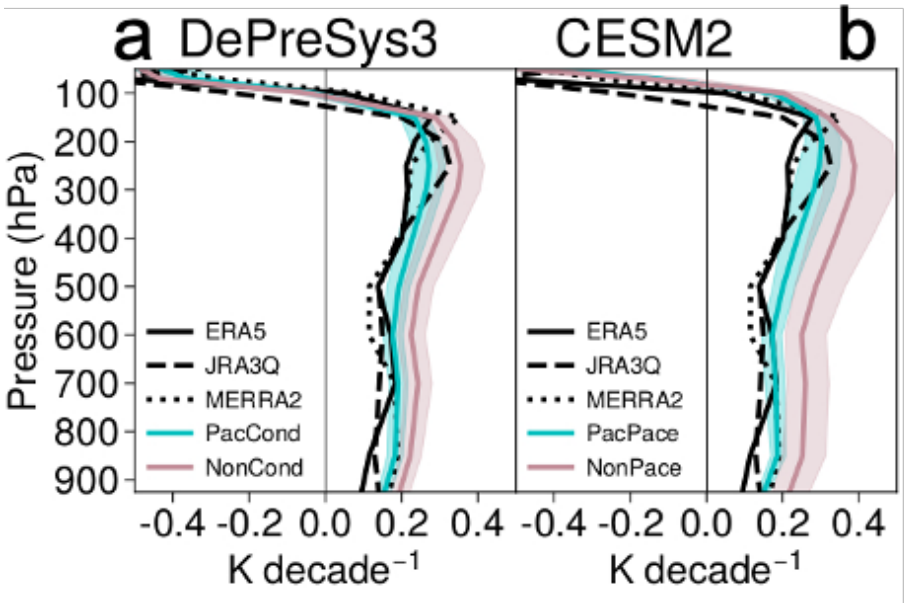


Figure 5.9: **Tropical tropospheric temperature trends.** Trends are averaged over the tropics (±20°) for the conditioned (cyan) and unconditioned models (pink): (a) DePreSys3 hindcasts, and (b) Pacemaker/CESM2-LE. Trends in three reanalyses are in black. Shading shows the 10-90% range of model trends.

When conditioned on the Pacific SST trend, the same models are in much better agreement with the reanalysis trends (cyan shading). As seen previously (Section 3.3), it is important to note the spread in trends between the different reanalyses, which gives an indication of the uncertainty in the true trend. Only at one vertical level, 500hPa, are all three reanalyses outside the range of conditioned model trends, despite the much narrower spread of trends in the conditioned models than in their unconditioned counterparts. In summary, the discrepancy between reanalysis and model trends is considerably reduced by capturing trends in the tropical Pacific - consistent with the conclusions of Po-Chedley et al. (2021) and others. This suggests that much of the improvement in tropospheric warming trends seen in Section 3.3 is due to improved SST trends in the tropical Pacific.

5.4.2 Shifting jets

Woollings et al. (2023) find that poleward jet shifts in reanalyses are large given their relatively small increases in upper-level temperature gradient $\frac{\partial T}{\partial y}$ (discussed in Chapter 2.1.3). The previous section showed that the conditioned runs have reduced tropical warming, consistent with a smaller increase in $\frac{\partial T}{\partial y}$. Does the Pacific conditioning also lead to improved jet trends?

Figure 5.10 tests this by recreating Figure 3 of Woollings et al. (2023) for boreal winter. Here, the temperature gradient index is denoted ΔT. The reanalysis points, in blue, are on the edge of the unconditioned CESM2-LE envelope, and are outside the 95th percentile of the unconditioned hindcast trends (black)[1]. On the other hand, the conditioned models generally show much smaller increases in ΔT and somewhat more poleward jet trends (red), and are consequently in good agreement with the reanalyses. This provides strong evidence that capturing the tropical Pacific SST trend helps to resolve discrepancies in jet shifts and temperature gradient trends. It could also be consistent with the anomalous poleward heat flux trends of Woollings et al. (2023), which they note may be associated with the observed SST trend.

It is nevertheless worth noting that the ensemble-mean jet responses (triangles) are modest compared to the spread of trends within the conditioned ensembles. The mean changes in temperature gradient are larger, although a wide range of ΔT trends is still seen even in the conditioned runs. Figure 5.9 shows that the tropical warming is relatively well constrained, so the spread is mostly due to the extratropics. While capturing the Pacific SST trend does improve jet trends *on average*, trends between

[1]This is because the 12m hindcasts are used in this chapter. As shown in the previous chapter, the shorter lead hindcasts are better able to capture tropical temperature and jet trends.

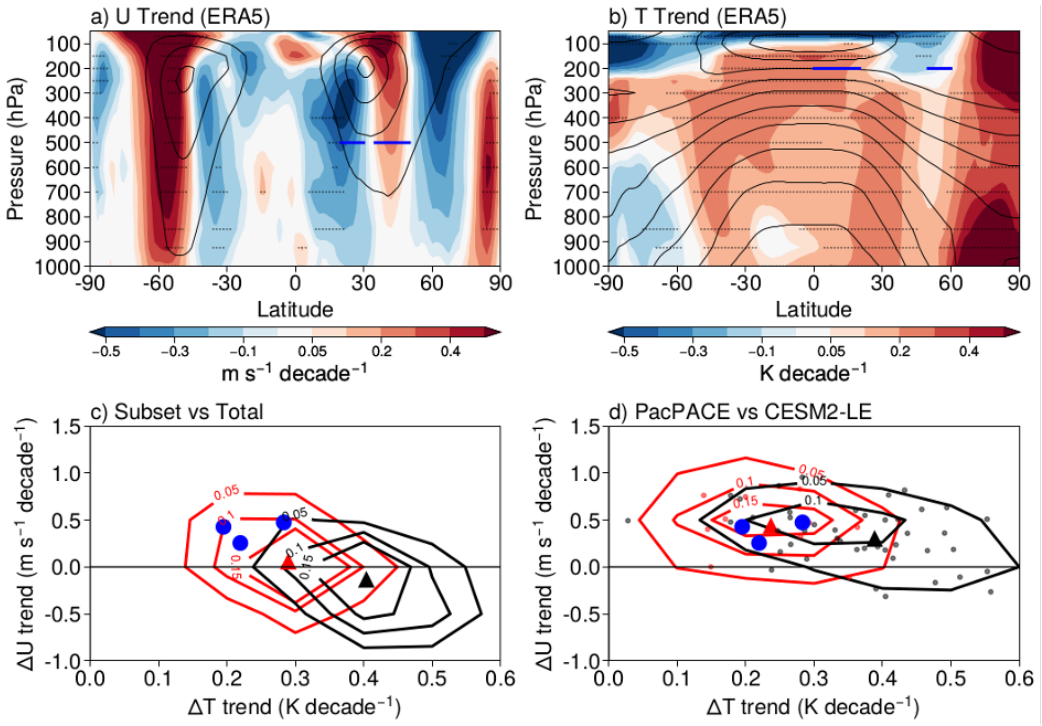


Figure 5.10: **Trends in zonal wind indices plotted against trends in meridional temperature gradient.** Joint density distribution of DJF trends in zonal indices and meridional temperature gradients in the NH, using the indices of Woollings et al. (2023) (their original figure is Figure 2.2). The latitude bands used to calculate the wind index ΔU (at 500hPa) and the temperature index ΔT (250hPa) are shown by the blue horizontal lines in panels (a) and (b), respectively, and the latitudes ranges are listed in a previous chapter (Table 3.1). Positive values of the wind and temperature indices correspond to a poleward shift of the jet and a strengthening of the meridional temperature gradient, respectively. Reanalyses are in blue, the unconditioned models are in black, and the conditioned models in red. Triangles mark the model ensemble means.

any pair of members will in general differ more due to factors external to the tropical Pacific (including intrinsic atmospheric variability).

5.4.3 Regional circulation

The previous section showed that both models show similar ensemble-mean responses in regional circulation metrics: a strengthening of the Pacific Walker Circulation (PWC), a weakening of the Aleutian Low (AL), and a trend towards positive NAO. Trends in these metrics are summarised in Figure 5.11. Panel (a) shows that the PWC response is fairly robust between ensemble members, particularly in the tightly-constrained Pacemaker runs; most of the conditioned hindcasts also show a positive

trend, but few capture the magnitude of the reanalysis trends. Nevertheless, compared to the unconditioned runs, conditioning on the SST trend clearly improves agreement with the reanalysis PWC trends, as may be expected from their tight coupling (Watanabe et al. 2023).

On the other hand, a much larger spread is seen in the extratropical circulation responses. The interquartile range of trends spans zero for both the AL and NAO, except for the Pacemaker NAO (though one member still shows a negative trend here). As for the zonal-mean jets, the ensemble-mean response in these features is small compared to the ensemble spread. The reanalysis trends are consistent with the unconditioned models in both regions; although the Pacific conditioning does generally move the model mean closer towards the reanalyses, comparable trends are still possible in the models without it. This highlights the contribution of internal variability beyond the tropical Pacific to trends in the AL, as also seen in the seasonal response to ENSO (Deser et al. 2017b).

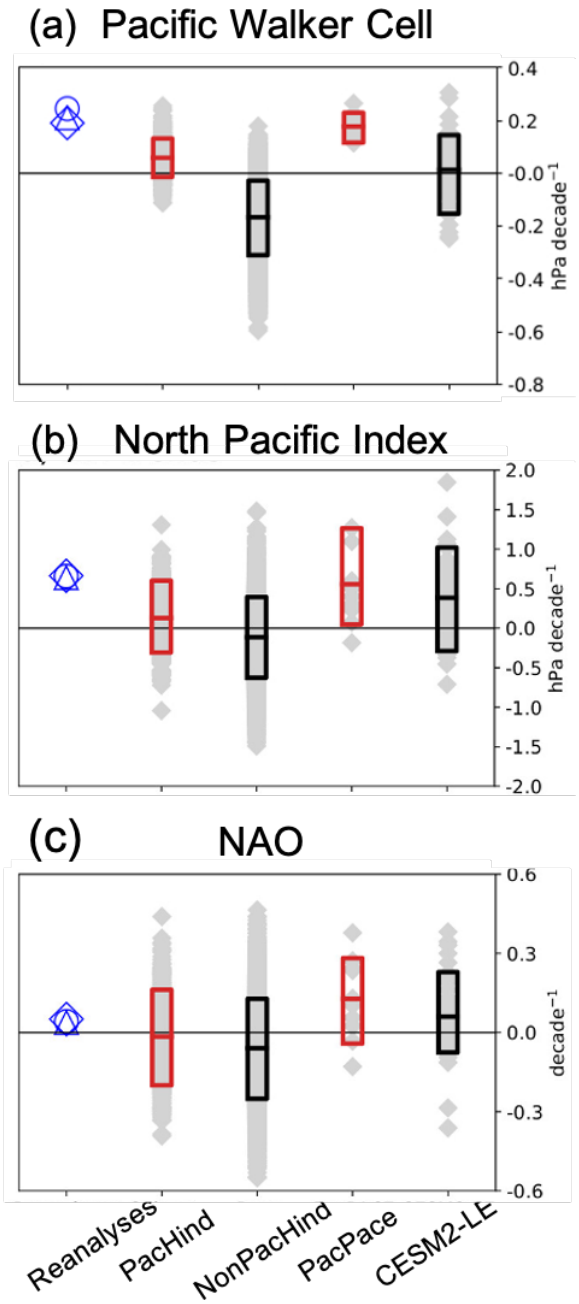


Figure 5.11: **Trends in regional circulation indices for the conditioned and unconditioned models.** Trends in the (a) Pacific Walker circulation (Vecchi and Soden 2007), (b) North Pacific Index (Trenberth and Hurrell 1994), and (c) NAO (Dunstone et al. 2016). Reanalyses are in blue, unconditioned models in black, and the conditioned models in red. Labelling of the models differs somewhat from previous figures: here, the conditioned and unconditioned hindcasts are denoted PacHind and NonPacHind respectively. Boxes show the interquartile range and median of the models, and grey dots are individual ensemble members.

5.4.4 Regional temperature and precipitation

Finally, this section shows that accounting for the Pacific SST trend improves model trends in two key impact variables – surface-air temperature (SAT) and precipitation – in regions which are affected interannually by ENSO.

Figure 5.12 shows the ensemble-mean SAT trends in ERA5 and the conditioned models (panels a-c), and the spread of trends in four boxed regions (panels d-g). The ERA5 trends in these regions are generally weak and lacking significance. In the ensemble mean, the unconditioned models overestimate the warming in each region; the reanalysis trends are either outside or on the very edge of the ensemble spread in each case, with the exception of central North America where the unconditioned hindcasts are largely consistent with the reanalyses.

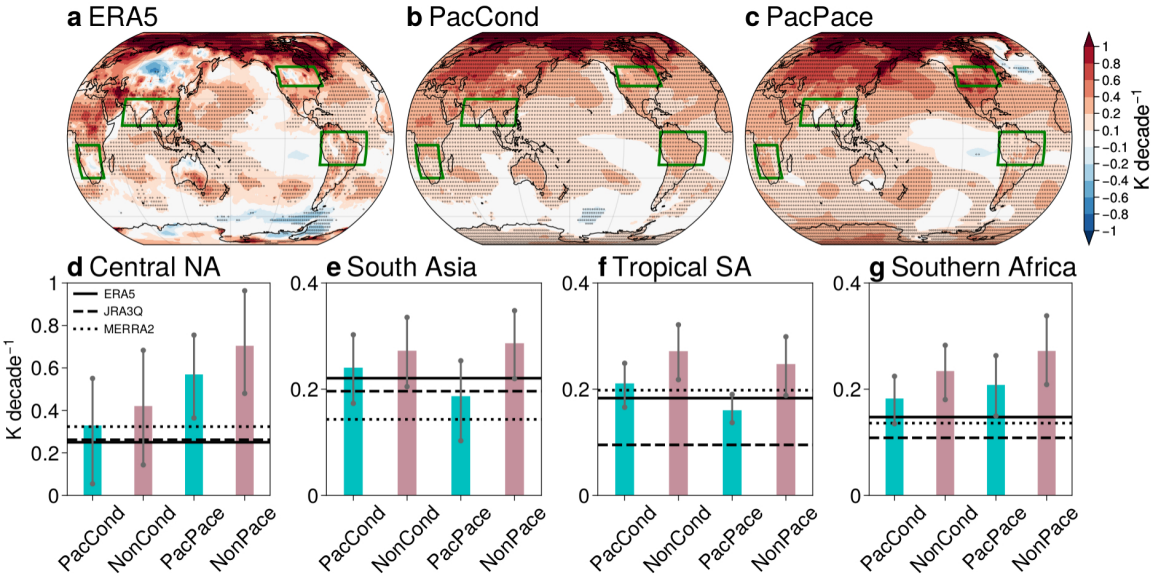


Figure 5.12: **Regional SAT trend improvement in the conditioned models.** Surface-air temperature (SAT) trends in ERA5 (a), the conditioned hindcast ensemble-mean (b), and Pacemaker ensemble-mean (c). Stippling denotes significant trends at the 5% level, as in Figure 5.3. (d-g) SAT trends averaged over the green boxed regions. Conditioned models are shown in cyan, and unconditioned in pink. Black horizontal lines denote reanalysis trends. Error bars show the 10-90% range of model trends.

SAT trends in each region are consistently reduced in the conditioned models, bringing the mean warming trends closer to the reanalyses. Considerable spread is seen in the reanalysis trends, so not all reanalysis products are within the conditioned model ensemble spread in each region. However, unlike the unconditioned runs, at least one reanalysis product falls within the conditioned ensemble spread of each

model in every region, barring the Pacemaker trends over Central North America. Capturing the Pacific SST trend therefore helps the models to capture SAT trends over the four regions identified here.

The Pacific similarly improves regional precipitation trends in the models. Panels (a-c) of Figure 5.13 shows precipitation trends in GPCP and the conditioned model ensemble means. The observed trend is generally well captured throughout much of the tropics and subtropics; as before, the ensemble spread is shown for four notable regions in panels (d-g).

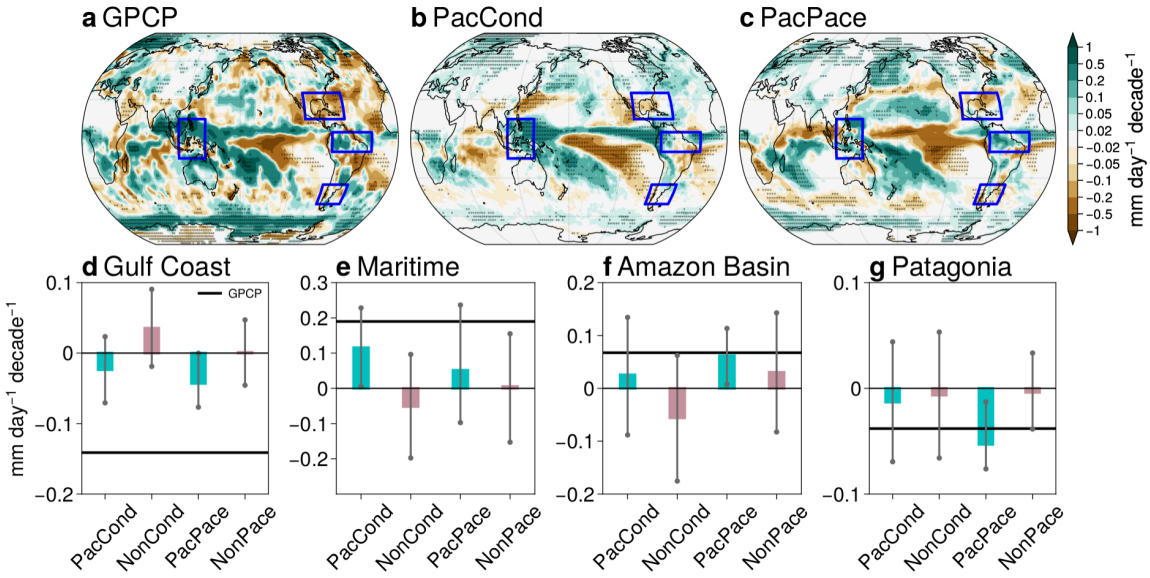


Figure 5.13: **Regional precipitation trend improvement in the conditioned models.** Precipitation trends in GPCP (a), the conditioned hindcast ensemble-mean (b), and Pacemaker ensemble-mean (c). Stippling is as in Figure 5.12. (d-g) Precipitation trends averaged over the blue boxed regions. Conditioned models are shown in cyan and unconditioned in pink. The black horizontal line denotes the GPCP trend. Error bars show the 10-90% range of model trends.

Increases in precipitation are seen over the Maritime Continent and Amazon basin (panels e-f). The ensemble-mean trends in the unconditioned models are weak or negative, and only in one case is the observed trend within the model range (CESM2-LE in the Amazon). However, in both regions, the ensemble-mean trend in the conditioned models is positive, and the observed trend is well captured.

In contrast to the positive trends in the tropics, drying trends are observed in the extratropical American regions identified in panels (d) and (g). The trends are broadly hemispherically symmetric, characteristic of the regions of anomalous descent associated with tropical expansion in response to La Nina (Seager et al. 2005).

As above, the ensemble-mean trend is weak in the unconditioned models, and only in Patagonia is the observed trend comfortably within ensemble range (for the unconditioned hindcasts). Again, the ensemble-mean trend in the conditioned models captures the sign of the observed trend, and its magnitude in Patagonia.

Even with the Pacific conditioning, however, the observed trend in the Gulf Coast region remains much larger than in the models. Qiu et al. (2024) showed that a linear adjustment, based on ENSO teleconnections, can bring coupled model precipitation trends over Mexico and the Southern US into agreement with the observed trend. Why do the conditioned models fail to do the same? Firstly, the Gulf Coast region used here does not match theirs exactly, being further east and including ocean tiles; this could contribute to the difference. However, if the discrepancy persists after matching the regions used, then it could point to deficiencies in the model trends besides the tropical Pacific. This highlights an avenue for future research which is beyond the scope of the present study.

5.5 Boreal summer results

The main focus of this chapter is the Pacific-forced impact on trends in DJF. In this section, a brief summary is given of the impacts of the Pacific SST trend in JJA, as found using the same methods as above. In general, these impacts are weaker than those in DJF, for two main reasons: a somewhat weaker tropical teleconnection source (not shown), and seasonal changes in teleconnection pathways. In particular, stronger jets in the winter hemisphere serve as better waveguides for tropically generated Rossby waves. Consequently, we may expect the greater impacts of the Pacific SST trend in JJA to be found in the SH. This section identifies two key impacts of the Pacific on trends in JJA: a strengthening of the South Asian monsoon, and a pressure dipole trend in the South Pacific.

5.5.1 Methods

The Pacemaker experiments are the same as in the previous section. The hindcast trend ensemble is also the same as above, but here months 8-10 are used rather than months 14-16. Despite the shorter lead time, we find that the trend ensemble is still sufficiently dispersed in the tropical Pacific. The trend bias metric is calculated again for JJA, so the 500 members with the lowest trend bias in DJF are not required to be the same as those in JJA. As before, trends are calculated from 1981-2018 inclusive.

5.5.2 South Pacific circulation

Panels (a), (b) and (c) of Figure 5.14 show the ERA5 SLP trend, and the Pacific-forced trend responses in the conditioned hindcasts and in Pacemaker, respectively. The NH response is not robust between the models; while the Aleutian Low weakens in Pacemaker, little response is seen in the conditioned hindcasts.

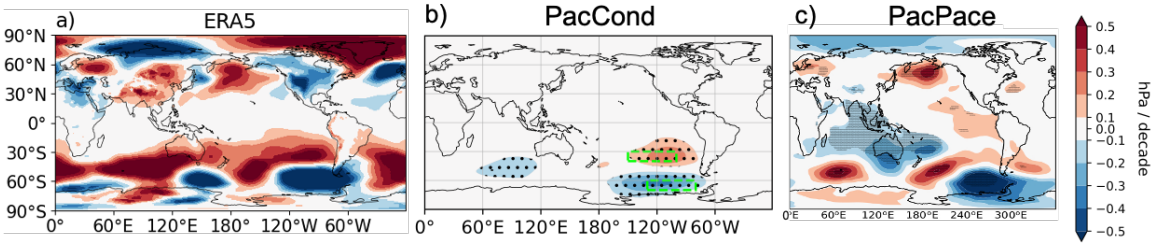


Figure 5.14: **Conditioned model SLP trend responses in JJA.** (a): SLP trend in ERA5. (b, c): Pacific-forced trend response, defined as before (i.e. the ensemble-mean trend in the conditioned models minus the ensemble-mean trend in the corresponding unconditioned model), for the conditioned hindcasts (b) and Pacemaker (c). Trends are in the JJA seasonal means, calculated over 1981-2018 inclusive. Stippling indicates a statistically significant trend/trend response at the 5% level. The green boxes shows the regions used to calculate the South Pacific SLP dipole in Figure 5.16.

As expected, the stronger response occurs in the winter hemisphere. The most striking feature of the circulation response is the pressure dipole trend in the South Pacific, with high pressure extending between New Zealand and Chile and low pressure to the south. Figure 5.16 (a) shows dipole trends in the hindcasts, using the definition of Garreaud et al. (2021) (green boxed regions in Figure 5.14 (b)). The difference between the conditioned and unconditioned distributions is small, despite the robust signal in Figure 5.14 (b). However, this subtle difference is important: the interquartile range of trends in the conditioned hindcasts is positive, and the ERA5 trend is within it. On the other hand, the interquartile range in the unconditioned hindcasts spans zero, and fails to capture the ERA5 trend.

The dipole response is present to some degree in both models, but differs slightly in its structure: the strongest response in Pacemaker is southeast of New Zealand, while the response in the conditioned hindcasts is more zonal and is strongest downstream (Figure 5.14, panels (c) and (b) respectively). Both the upstream and downstream aspects are present in the ERA5 trend (Figure 5.14 (a)). Why does the response differ between the models?

Multiple factors may contribute. An important ingredient could be the SST trend in the subtropical Southwest Pacific, east of New Zealand – termed by Garreaud et al.

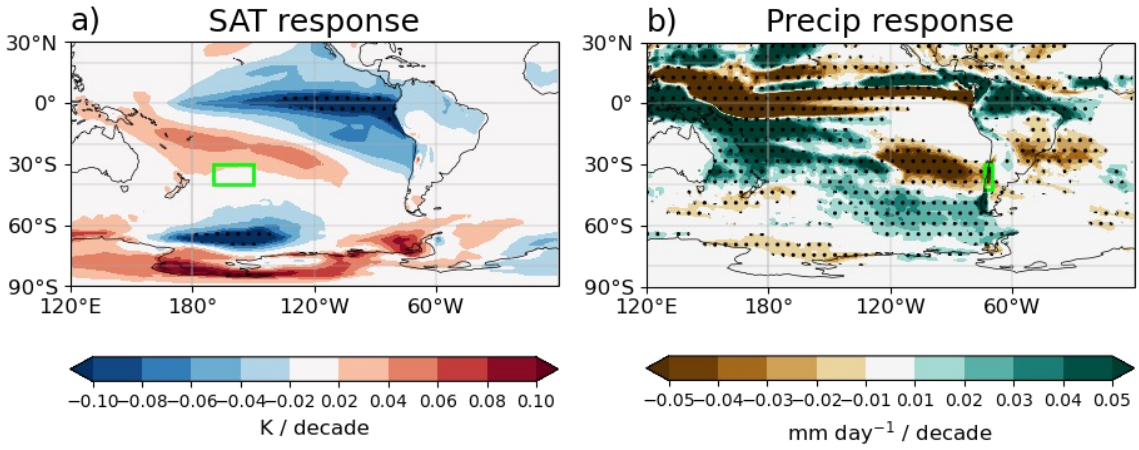


Figure 5.15: **Conditioned hindcast SAT and precipitation trend responses in JJA.** As in Figure 5.14 (b), but for (a) SAT and (b) precipitation. The green boxes in (a) and (b) denote the regions used to calculate the Southern Blob and Chile precipitation indices, respectively (Figure 5.16).

(2021) the "Southern Blob". Using atmosphere-only simulations, they showed that circulation trends in the downstream extension are related to warming in the blob region. Figure 5.15 a shows the SH SAT response in the conditioned hindcasts, with the blob region of Garreaud et al. (2021) shown in green. A warming response is seen in the subtropical Southwest Pacific, although the response is strongest north of the blob region. SAT trends in this region are broadly similar in the conditioned and unconditioned hindcasts (Figure 5.16 b). The magnitude of warming in the blob region is large relative to the hindcast trends: the conditioned runs fail to capture the ERA5 SAT trend, and neither hindcast ensemble captures the magnitude of the ERSSTv5 SST trend. This result is not sensitive to the precise latitude and longitude ranges chosen for the blob (not shown). The lack of response in the blob region is surprising given that the conditioned hindcast circulation response is strongest downstream.

Understanding the downstream response is of particular importance for impacts in South America, not least the ongoing Chilean "megadrought" (Garreaud et al. 2017; Garreaud et al. 2020). Figure 5.15 b shows the precipitation response in the conditioned hindcasts. A zonal precipitation dipole is approximately collocated with the South Pacific pressure dipole, with both corresponding to a poleward shift of the jet. Precipitation trends over central Chile, the green boxed region in Figure 5.15, are shown in Figure 5.16. Again, the effect of the conditioning is small, although the conditioned hindcasts are somewhat better able to capture the GPCP drying trend.

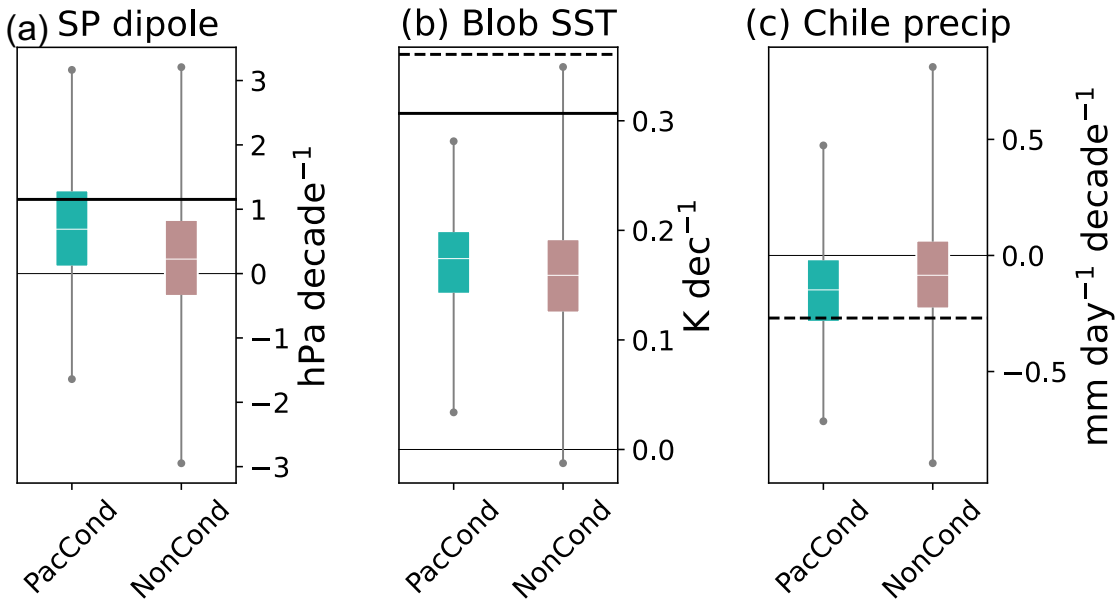


Figure 5.16: **Regional trend indices in the conditioned and unconditioned hindcasts for JJA.** (a) The South Pacific pressure dipole trend, defined as the difference in SLP trend between boxes over the subtropical Pacific (40°–30°S, 210°–260°E) and the Amundsen–Bellingshausen Sea (70°–60°S, 230°–280°E) following Garreaud et al. (2021). (b) SAT trends in the Southern Blob region: 40°–30°S, 190°–210°E (Garreaud et al. 2021). (c) Precipitation trends over central Chile (74°-70°W, 42°-30°S). All regions are shown in green in the previous figures. Solid black lines are the ERA5 trends, and dashed black lines in (b) and (c) are ERSSTv5 and GPCP trends respectively.

In summary, the tropical Pacific SST trend is associated with a pressure dipole trend in the South Pacific, and a slight tendency towards drying trends over central Chile. However, the magnitude of the effect in the conditioned hindcasts is small in each case, and observed SST trends in the southern blob region of the Southwestern Pacific remain outside the hindcast range.

5.6 Discussion and conclusions

This chapter has shown that the tropical Pacific SST trend has large and robust impacts on trends in the global climate. The main results can be summarised as follows:

1. The boreal winter circulation response to the tropical Pacific SST trend is largely consistent with the model's interannual La Nina response.

- In the zonal mean, the tropics cool and the subtropical jets weaken and shift poleward.
- The Pacific SST trend corresponds to a strengthened PWC, a weakened Aleutian Low and a poleward shift of the jet over North America.
- The Eurasian circulation response is not robust between models. Though some Arctic warming and Eurasian cooling is seen, in neither model does the magnitude of the response indicate a leading role for the tropical Pacific in the observed trends.

2. Conditioning on the tropical Pacific trend also improves the models' agreement with observed and reanalysis trends.

    - Reanalysis tropical tropospheric temperature trends are mostly within the 10-90% range of trends in the conditioned models.
    - The Pacific trend leads to much weaker upper-level temperature gradients and slightly more poleward jet shifts. This largely resolves the discrepancies found by Woollings et al. (2023), though the impact on the temperature gradient is much larger than on the jet. Non-Pacific variability remains a leading order consideration for trends in the jets over this period.
    - Very few (CESM2-LE) or none (hindcasts) of the unconditioned runs capture the PWC strengthening, whereas the conditioned runs do. However, the responses in the North Pacific and North Atlantic are modest and the unconditioned runs already capture observed trends here.
    - The Pacific SST trend is necessary for capturing observed temperature and precipitation trends over certain regions in Southern Africa, the Maritime Continent, and the Americas. However, the magnitude of the Gulf Coast drying trend is not reproduced in the conditioned runs – an example of discrepancies in the models beyond the tropical Pacific.

3. The trend response in NH summer is generally weak. In the SH winter, a pressure dipole response is seen over the South Pacific in both models, although the response size in the conditioned hindcasts is small and does not meaningfully affect downstream precipitation trends in central Chile.

The results presented above have been interpreted through teleconnections from the tropical Pacific that are well established on seasonal timescales, and which appear

to also hold in projections of future climate (Collins and The CMIP Modelling Groups 2005). However, with only a small pool of ENSO events in the observed record, even these relationships are somewhat uncertain, particularly outside the Pacific and North American sectors (Deser et al. 2017b). Unlike ENSO composites, which isolate the influence of the Pacific, the single observed trend also contains contributions from outside the Pacific, as well as the externally forced GHG trend.

Although the trend responses above are robust and are well supported by ENSO theory, some limitations of the results remain. An important consideration for the hindcast subsetting approach is the extent to which even the "best" hindcast trends resemble the observed trend over the tropical Pacific: just how close is close enough? Figure 5.2 shows that while the conditioned hindcast trends are clearly more La Nina-like than the ensemble mean, they still differ from the observed trend. The weaker Pacific trend in the conditioned hindcasts, compared to Pacemaker, likely explains the generally weaker trend response outside the tropical Pacific. However, although the trend response is somewhat weaker, the pattern of the global response is broadly similar in both Pacemaker and the conditioned hindcasts, building confidence in the robustness of our findings despite small differences in the tropical Pacific.

Another feature of the hindcast trends is that they are built of discrete snapshots in time, rather than showing changes in a single continuous model integration. Slow feedbacks, which couple the climate state in successive years, are therefore neglected by construction. Most such processes involve subsurface ocean processes. For example, on seasonal timescales a NAO+ causes anomalous cooling of the North Atlantic ocean surface. In a continuous integration, the cooler surface waters are denser and can alter the subsurface ocean circulation; the effect of the altered circulation will then persist beyond the seasonal timescale (Khatri et al. 2022). In the hindcast timeseries, such coupling across timescales is absent: an NAO+ in year Y has no bearing on the climate state in year Y+1, since all members are reinitialised from observed conditions in the intervening period. The North Atlantic response may indeed be consistent with this possibility, with the continuous Pacemaker runs showing a much stronger cooling response than the discrete hindcasts (despite both showing trend responses towards NAO+). Another potential case concerns the response of the stratospheric polar vortex to Pacific SSTs, which has been shown to weaken in response to El Nino but strengthen during the positive phase of the PDO. The sign change of the response over longer time periods has been traced to the slow build-up of stratospheric water vapour on decadal timescales (Seabrook et al. 2023). To the extent that the Pacific SST trend resembles the PDO, this process could also contribute to the trend

response in the polar stratosphere, but would not be present in the hindcast trends. This is an unavoidable feature of the hindcast trends approach, but comparison with the Pacemaker ensemble suggest that it is unlikely to be a leading order effect.

As noted in the introduction, the hindcast and Pacemaker approaches are complementary. The DePreSys3 hindcasts are run at high horizontal and vertical resolution, and have demonstrated good skill in simulating ENSO teleconnections (Dunstone et al. 2016). The trend ensemble is large, allowing robust statistics and the minimisation of noise in the Pacific-forced response. A limitation of the hindcast subsetting approach is that, since the SSTs are not directly perturbed, a direction of causality is not implied; changes in the tropical Pacific may arise as a consequence of extratropical processes, rather than being their immediate cause, as suggested recently by Toda et al. (2024). On the other hand, the Pacemaker approach directly tests the causal impacts of the Pacific SST trend, although at the expense of preserving full dynamical coupling between the ocean and atmosphere, which may lead to weaker teleconnections (Molteni et al. 2020). The good agreement between these very different methods builds confidence in our conclusions.

This chapter has explored the modelled impact of the Pacific SST trend on remote circulation trends. The relevance of the modelled response to the real climate depends on the extent to which these models faithfully represent teleconnections from the tropical Pacific. There remain doubts over teleconnection strength in models, linked to the well-known signal-to-noise issues prevalent in models (Williams et al. 2023; Molteni and Brookshaw 2023). Considering an ensemble of 5 different seasonal forecast models, Williams et al. (2023) showed that the magnitude of the extratropical ENSO response was too weak in each model, despite them skilfully capturing some of the spatial structure of the response. Shared model deficiencies in the Pacific-forced response cannot be ruled out and are an unavoidable concern for any model-based study addressing this topic.

Our focus has been primarily on trends in the mean state due to the limited availability of higher frequency hindcast data. However, as discussed in Section 5.1 and Chapter 2, many potential impacts of the Pacific SST trend concern changes in atmospheric variability and extremes, such as atmospheric rivers (Li and Ding 2024), tropical cyclones (Sobel et al. 2023) and midlatitude waviness (Sun et al. 2022). If sub-monthly hindcast data were to be available, future work could adapt our approach to explore the Pacific impact on trends in variability. Additionally, although our focus has been on the standard summer and winter seasons, the Pacific has also been linked to observed trends in other seasons, such as the central equatorial African drought

in April-May-June (Hua et al. 2016). A simple extension of this method would be to explore the Pacific impact on trends in seasons besides DJF and JJA. Finally, it bears repeating that this work has quantified the Pacific-forced trend response over the recent historical period. Coupled models robustly predict a reversal of the strengthening zonal SST gradient over the coming decades. However, even if the current trend continues, changes in the basic state that are independent of the Pacific could lead to qualitatively different remote impacts in future (Drouard and Cassou 2019; Beverley et al. 2024a).

Chapter 6

Diagnosing the circulation contribution to ERA5 temperature trends

6.1 Introduction

Describing trends in the large-scale temperature structure of the atmosphere is a primary goal in climate science. Figure 6.1 shows the trend in zonal-mean temperature over the satellite era in the ERA5 reanalysis. Previous chapters have discussed various features of this figure, from tropical tropospheric warming to amplified low-level warming over the Arctic.

Studies of the circulation response to climate change commonly take as their starting point a particular zonal-mean temperature profile, such as in Figure 6.1, and seek to understand the altered circulation as a response to this (e.g. Shaw 2019). This approach has yielded numerous advances in understanding the circulation; for example, relating the differing shifts of the extratropical jets seen in El-Nino events versus under greenhouse forcing to upper-level meridional temperature gradients (Lu et al. 2008). The dynamical mechanisms often depend sensitively on the spatial temperature distribution.

Implicit in this approach is the treatment of the atmospheric dynamics as separate from, and a consequence of, the temperature structure. Fundamentally, however, the dynamics and thermodynamics of the atmosphere are coupled in the equations of motion: the advection of warm or cool air, and the adiabatic expansion or compression of air parcels undergoing vertical motion, both contribute to the temperature change at a given location (Peixoto and Oort 1992). Even when it is thermodynamic processes that are driving temperature changes, the circulation will tend to adjust along with

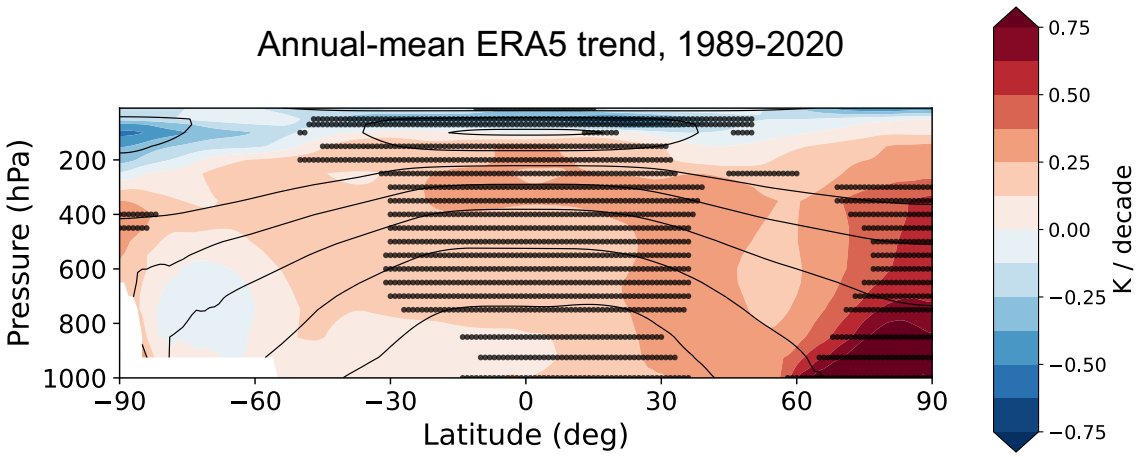


Figure 6.1: **Trend in zonal-mean, annual-mean temperature in ERA5 over 1989-2020.** Colours show the trend in K / decade. Black contour lines show the time-mean and solid lines indicate positive values. Stippling denotes a significant trend at the 5% level using a two-tailed t-test.

the thermal structure to bring the system back into balance.

Previous chapters have given support to significant emerging trends in the circulation. This chapter asks: how has the atmospheric circulation itself contributed to the trends in Figure 6.1? The approach taken is to use a heat budget, which decomposes the zonal-mean temperature tendency into contributions from diabatic and dynamical processes. To make explicit the link with temperature *trends*, we look at accumulations in the tendency terms, and thereby express the annual- and zonal-mean temperature trend in ERA5 as the sum of a dynamical and a diabatic contribution. We then assess how changes in the wind and temperature fields each contribute to driving the overall temperature tendency trend. Finally, we explore the seasonality of the temperature tendency trends.

Section 6.2 outlines the data used and defines the temperature tendency budget, before the accumulated budget is introduced in Section 6.3. In Section 6.4 the dynamical and thermodynamical drivers of the tendency terms are assessed, and section 6.5 presents the seasonal trends in the tendencies. Significance testing is performed in Section 6.6. The discussion and concluding remarks are in Section 6.7.

6.2 Methods

6.2.1 Reanalysis data

The ECMWF's ERA5 reanalysis dataset is used in this study (Hersbach et al. 2020). As described in previous chapters, ERA5 uses a fixed data assimilation method to generate an observationally-constrained dataset intended to best represent the atmospheric state at each output timestep. Previous chapters used monthly-mean reanalysis data. Here, 6-hourly data is averaged to daily resolution on a 1°x1° grid, for 20 pressure levels between 1000hPa and 10hPa. Only data above the surface is used. As before, analysis is restricted to the satellite era where global instrument coverage is greatest (1979-2020 here; previous chapters have used slightly different start and end dates).

All reanalyses, including ERA5, suffer from time-varying observational coverage as new instruments become available, and hence their relation to trends in the real atmosphere is not straightforward (Bengtsson et al. 2004). However, ERA5 is one of the newest and highest quality reanalyses, and has been shown to have fewer discontinuities in its temperature timeseries than another recent dataset (Kozubek et al. 2020). Prior concerns regarding trends in tropical diabatic heating are repeated here and are discussed in the relevant sections below, noting that this bias appears common to most modern reanalysis products (Chemke and Polvani 2019).

6.2.2 Tendency budget

The Eulerian-mean budget for the zonal-mean temperature tendency is

$$\begin{aligned}\frac{\partial\langle T\rangle}{\partial t} = &- \underbrace{\left[\frac{\langle v\rangle}{a}\frac{\partial\langle T\rangle}{\partial\phi} + \langle\omega\rangle\left(\frac{\partial\langle T\rangle}{\partial p} - \frac{R}{c_p p}\langle T\rangle\right)\right]}_{\text{MMC}} \\ &\underbrace{- \frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\langle v^*T^*\rangle\cos\phi\right) - \frac{\partial}{\partial p}\langle\omega^*T^*\rangle + \frac{R}{c_p p}\langle\omega^*T^*\rangle}_{\text{Eddies}} \\ &+ \frac{\langle Q\rangle}{c_p},\end{aligned} \tag{6.1}$$

where $\langle\cdot\rangle$ denotes a zonal mean and * a deviation from the zonal mean, a is the planetary radius, and c_p is the specific heat capacity of air at constant pressure (Peixoto and Oort 1992).

Following Seager et al. (2003), terms on the right-hand side can be grouped into tendency contributions from the mean meridional circulation (MMC), eddies, and diabatic processes $\langle \mathrm{Q} \rangle$. The MMC is the zonal-mean circulation in the latitude-height plane, including the Hadley cells in the tropics and the Ferrel and polar cells at mid-high latitudes. The eddy terms include stationary and transient eddies. The first two terms in the MMC and eddy groups relate to the advection of warm/cool air, while the third is the temperature change due to the adiabatic work associated with local ascent/descent.

As a shorthand, the dynamical terms can be combined into a total dynamics D:

$$\frac{\partial \langle T \rangle}{\partial t} = \partial \langle D(t) \rangle + \frac{\langle Q \rangle}{c_p}. \tag{6.2}$$

The net temperature tendency $\frac{\partial \langle T \rangle}{\partial t}$ is evaluated directly as the centred difference of the temperature field at each daily timestep. In ERA5, the diabatic tendency due to parametrisations is provided on model levels as a mean over the last time increment, whereas the winds and temperature are output as instantaneous variables on pressure levels. Hence, it was decided to calculate Q as a budget residual, rather than using the model diabatic tendencies. The thermodynamic budget residual at each daily timestep is therefore taken as the implied diabatic heating, following previous work (Yanai et al. 1973; Nigam et al. 2000; Seager et al. 2003). A shortcoming of this approach is that it offers no way to check that the budget closes, and contributions due to numerical diffusion will be present in the residual. However, the discrepancy between heating calculated as parametrised output or as the budget residual is likely to be small compared to systematic biases in the diabatic heating between different reanalysis products, or between reanalyses and soundings (Ling and Zhang 2013).

In the time-mean, the net temperature tendency reduces to the temperature difference between the first and last days in the data period, divided by the length of the data period. On daily timescales, large temperature tendencies are possible, of similar magnitude to the heating terms in (6.1); averaged over several decades, however, the time-mean tendency is smaller by at least 3 orders of magnitude than the average dynamical and diabatic heating terms shown in Figure 6.2. Consequently, the sum of the adiabatic dynamical heating (Figure 6.2a) nearly perfectly balances the diabatic term (Figure 6.2b) on timescales longer than a week, as noted previously (Peixoto and Oort 1992; Lachmy and Kaspi 2020). This points to the efficiency of the circulation in redistributing heat through the atmosphere on timescales longer than a few days.

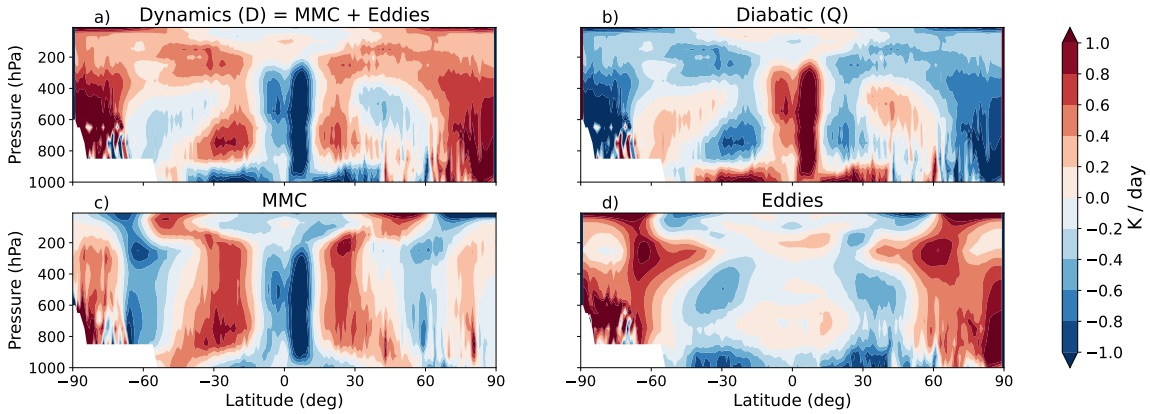


Figure 6.2: **Climatology of terms in the local temperature tendency budget**. The tendency due to (a) the total dynamics, (b) diabatic processes, (c) the mean meridional circulation (MMC) and (d) eddies. Units are in K/day of heating, and the time mean is calculated over 1979-2020.

The relative contribution of the dynamical terms is shown in Figure 6.2c and Figure 6.2d. In the tropics, the MMC dominates, with deep ascent in the ITCZ cooling the mid-upper troposphere and the corresponding latent heat release heating it diabatically (Figure 6.2b). The region of strongest cooling is slightly north of the equator, reflecting the climatological position of the ITCZ over the Atlantic and Pacific oceans (Gruber 1972; Philander et al. 1996). Over land, the strongest convective activity occurs over a wider range of latitudes, following the seasonal cycle. In austral summer, this deep convection over land is strongest south of the equator, leading to the broad tropical ascent and cooling in the annual-mean MMC (Figure 6.2c). The winter hemisphere subtropics are heated dynamically by mean descent in the Hadley cell, but cool radiatively. Meridional heat flux convergence by the MMC is an order of magnitude smaller than in the vertical, and the primary balance is between vertical heat flux convergence and adiabatic cooling due to expansion associated with large-scale ascent (not shown).

The diabatic term, diagnosed here as a residual, agrees well with other studies on large scales (Ling and Zhang 2013), although regionally the diabatic heating can vary substantially between reanalyses (e.g. the tropical upper troposphere/lower stratosphere, Wright and Fueglistaler 2013). The heating pattern in Figure 6.2b can be considered the superposition of two regimes: at all latitudes, the troposphere cools diabatically to space (and to the surface at high latitudes), with localised regions of diabatic heating mapping onto well-understood processes. In the tropics, latent heating is associated with deep convection as described above. Latent heating in the

midlatitude storm tracks (Figure 6.2b) is balanced by poleward and upward transient eddy heat transport in both hemispheres (Figure 6.2d). The boundary layer in the tropics and midlatitudes is heated diabatically by turbulent surface heat fluxes.

6.3 Accumulated heating trends

6.3.1 Method

The budget equation in (6.1) is for the temperature tendency, in units of K/day. To make explicit the link with temperature trends, this section introduces a budget for the accumulated tendencies between some reference time t_0 and day t:

$$\big\langle T(t)\big\rangle - \big\langle T_0\big\rangle = \int_{t_0}^{t} \big\langle D(t')\big\rangle \; dt' + \int_{t_0}^{t} \frac{\big\langle Q(t')\big\rangle}{c_p} dt', \tag{6.3}$$

where $T_0 = T(t_0) = const.$ In this way, trends in $\big\langle T(t)\big\rangle$ can be directly related to trends in accumulated heating due to the circulation, and due to diabatic processes.

The total accumulated heating for each tendency term $\big\langle X(t)\big\rangle$ is dominated by a steady climatological accumulation. Each term $\big\langle X(t)\big\rangle$ is therefore written as the sum of a climatological mean $\big\langle X_C\big\rangle$ and anomaly $\big\langle X_a(t)\big\rangle$ components before integrating, with the analysis focusing on anomalous trends beyond simply the climatological accumulation of heating.

For the dynamical tendencies, $\big\langle X_C\big\rangle$ is calculated as the all-year mean over the first 10 years of data (1979-88), with $t = \tau$ denoting the final date in this reference period. Trends will be calculated since the end of the reference period, which is therefore chosen to be as short as possible while remaining reasonably representative of the climatological base state. By definition, the anomalous accumulation at the end of the reference period is zero, and so the accumulated heating up to each timestep t is calculated with respect to $t = \tau$.

Motivated by the close balance between the diabatic and dynamical tendencies, we define $\dfrac{\big\langle Q_c\big\rangle}{c_p} = -\big\langle D_c\big\rangle$ such that the sum of the accumulations in the mean tendencies cancel, leaving:

$$\big\langle T(t)\big\rangle - \big\langle T_0\big\rangle = \int_{\tau}^{t} \big\langle D_a(t')\big\rangle \; dt' + \int_{\tau}^{t} \frac{\big\langle Q_a(t')\big\rangle}{c_p} dt'. \tag{6.4}$$

Now, trends in $\big\langle T(t)\big\rangle$ can be decomposed into a component due to trends in the accumulation of anomalous dynamical tendencies, and anomalous diabatic tendencies. τ is identified as the first data point following the end of the reference period (i.e.

01/01/1989). In the plots to follow, annual-mean trends are calculated over the period 1989-2020 using least-squares regression. Significance testing is discussed separately in Section 6.

6.3.2 Results

Figure 6.3 shows trends in the accumulated tendencies due to the dynamic and diabatic anomalies. In Section 6.2, it is noted that the time-mean tendency is small and so, climatologically, the diabatic heating tendency in (6.1) is approximately balanced by the dynamical tendency. In Figure 6.3, we see that the same is true for decadal trends in the accumulated tendency budget: the net temperature trend in Figure 6.3a is the result of a minor imbalance between the much larger trends in accumulated heating due to diabatic processes and due to dynamical processes.

Several trends in the dynamical heating are noteworthy. In the tropics, the ITCZ narrows and intensifies, particularly north of the equator, with strengthened adiabatic cooling (Figures 6.3a and 6.3c) due to increased ascent balanced by increased latent heating (Figure 6.3b). Interestingly, a strong dynamical heating trend is seen between 0-5°S and below 650hPa. Dynamical heating due to descent is reinforced in the subtropics in both hemispheres, with a concurrent increase in the diabatic cooling here. Together with positive trends in the ITCZ heating, this is consistent with an overall intensification of the tropical overturning circulation in ERA5, as noted in other reanalysis studies (Mitas and Clement 2005; Chemke and Polvani 2019). As discussed in previous chapters, the NH strengthening has been suggested to be an artefact of reanalyses related to spurious trends in diabatic heating (Section 3.6).

The MMC dominates the accumulated dynamical heating in the tropics and subtropics. The deep dynamical cooling trend at 60°S is associated with strengthened ascent in the poleward branch of the SH Ferrel cell, consistent with previous reanalysis studies of the satellite era (Chen et al. 2011) and with modelled responses to greenhouse gas forcing (Rojas 2013). Trends in the MMC contribution are generally weaker in the NH extratropics.

In the midlatitudes, the mid-tropospheric eddy cooling in the storm track strengthens moderately in both hemispheres. The eddy heating in the midlatitude upper troposphere strengthens on the equatorward side of its annual-mean climatological peak in the NH (45°N) and to a lesser extent in the SH (60°S). The peak in eddy heating moves modestly upward in both hemispheres, consistent with a lifting of the extratropical general circulation in response to warming (Singh and O'Gorman 2012; Thompson et al. 2019). Trends in the midlatitude accumulated eddy heating are

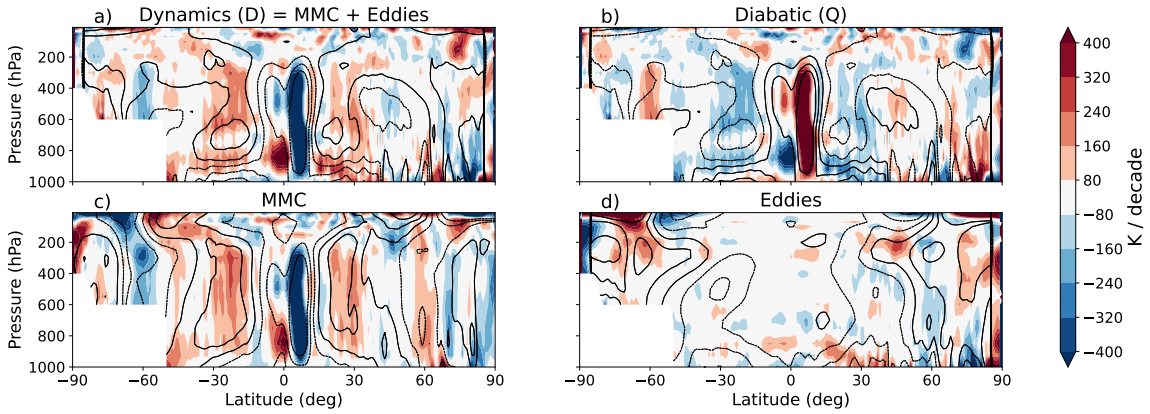


Figure 6.3: **Trends in annual-mean accumulated heating anomalies**. Heating trends due to a) the total dynamics, (b) diabatic processes, (c) the mean meridional circulation (MMC) and (d) eddies. Colours show the trend in K / decade. Black contour lines show the time-mean of the climatological accumulated heating with spacing of 2000K and solid lines indicating positive values.

likely obscured in the annual mean (see Section 6.5 for trends in the tendency terms by season).

6.4 Drivers of tendency trends

Section 6.3 shows the trends in the heating tendencies and their accumulations. Here, the unintegrated tendency terms are temporally decomposed in order to quantify the factors driving their trends, analogously to the moist static energy transport decomposition of Clark et al. (2022). Each variable is written

$$\begin{aligned}\langle T \rangle &= \langle \overline{T} \rangle + \langle T' \rangle \\ \langle v \rangle &= \langle \overline{v} \rangle + \langle v' \rangle \\ \langle \omega \rangle &= \langle \overline{\omega} \rangle + \langle \omega' \rangle,\end{aligned}$$

where an overbar denotes the time-mean over the whole data period and primes the temporal anomalies. The MMC thus becomes

$$
\begin{aligned}
\text{MMC} = &- \underbrace{\left[\frac{\langle\overline{v}\rangle}{a}\frac{\partial\langle\overline{T}\rangle}{\partial\phi} + \langle\overline{\omega}\rangle\left(\frac{\partial\langle\overline{T}\rangle}{\partial p} - \frac{R}{c_p p}\langle\overline{T}\rangle\right)\right]}_{\text{(1) MEAN}} \\
&- \underbrace{\left[\frac{\langle\overline{v}\rangle}{a}\frac{\partial\langle T'\rangle}{\partial\phi} + \langle\overline{\omega}\rangle\left(\frac{\partial\langle T'\rangle}{\partial p} - \frac{R}{c_p p}\langle T'\rangle\right)\right]}_{\text{(2) THERM}} \\
&- \underbrace{\left[\frac{\langle v'\rangle}{a}\frac{\partial\langle\overline{T}\rangle}{\partial\phi} + \langle\omega'\rangle\left(\frac{\partial\langle\overline{T}\rangle}{\partial p} - \frac{R}{c_p p}\langle\overline{T}\rangle\right)\right]}_{\text{(3) DYN}} \\
&- \underbrace{\left[\frac{\langle v'\rangle}{a}\frac{\partial\langle T'\rangle}{\partial\phi} + \langle\omega'\rangle\left(\frac{\partial\langle T'\rangle}{\partial p} - \frac{R}{c_p p}\langle T'\rangle\right)\right]}_{\text{(4) NONLINEAR}}
\end{aligned}
\tag{6.5}
$$

Group 1 (MEAN) has no time dependence and is neglected for the trend analysis. The second group denotes the advection of the anomalous temperature field by the steady MMC; group 3 is the advection of the climatological temperature field by the anomalous MMC; and group 4 is the nonlinear advection of the anomalous temperature field by the anomalous circulation. As in Clark et al. (2022), groups 2, 3, and 4 are identified as being thermodynamically driven (THERM), dynamically driven (DYN), or arising from nonlinear interactions (NONLINEAR) respectively. Applying the same decomposition to the eddies yields

$$
\begin{aligned}
\text{Eddies} = &- \underbrace{\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\langle\overline{v}^*\overline{T}^*\rangle\cos\phi\right) - \frac{\partial}{\partial p}\langle\overline{\omega}^*\overline{T}^*\rangle + \frac{R}{c_p p}\langle\overline{\omega}^*\overline{T}^*\rangle}_{\text{(1) MEAN}} \\
&- \underbrace{\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\langle\overline{v}^*T'^*\rangle\cos\phi\right) - \frac{\partial}{\partial p}\langle\overline{\omega}^*T'^*\rangle + \frac{R}{c_p p}\langle\overline{\omega}^*T'^*\rangle}_{\text{(2) THERM}} \\
&- \underbrace{\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\langle v'^*\overline{T}^*\rangle\cos\phi\right) - \frac{\partial}{\partial p}\langle\omega'^*\overline{T}^*\rangle + \frac{R}{c_p p}\langle\omega'^*\overline{T}^*\rangle}_{\text{(3) DYN}} \\
&- \underbrace{\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\langle v'^*T'^*\rangle\cos\phi\right) - \frac{\partial}{\partial p}\langle\omega'^*T'^*\rangle + \frac{R}{c_p p}\langle\omega'^*T'^*\rangle}_{\text{(3) NONLINEAR}},
\end{aligned}
\tag{6.6}
$$

where the terms are grouped similarly to above.

Figure 6.4 shows the trends in the MMC decomposition terms in (6.5). Note that the whole period 1979-2020 is used in this chapter's remaining figures since there is

no need to define an initial climatological period as in the accumulated budget. The net trend is dominated by the trend in the dynamic term containing the anomalous winds; both the thermodynamic and nonlinear terms are negligible, particularly in the extratropics.

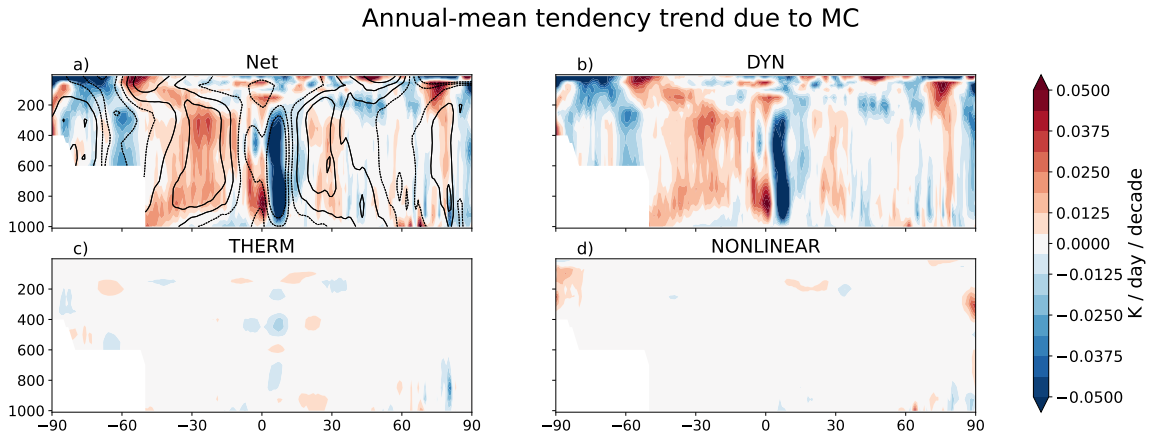


Figure 6.4: **Drivers of the tendency trend due to the MMC**. (a) Net annual-mean trend in temperature tendency due to MMC; (b)-(d) contributions to the net trend in (a) from dynamic, thermodynamic, and nonlinear anomalies respectively. Colours show the trend in K / day / decade. Black contour lines in (a) show the time-mean with spacing of 0.35K / day and solid lines indicating positive values. Unlike in Figure 6.3, the whole 1979-2020 period is used here.

The three terms in the eddy decomposition (6.6) are all of a similar magnitude. Closer inspection reveals some cancellation between the dynamic and thermodynamic terms, and so the net eddy tendency trend is similar to the nonlinear interaction term between the anomalous circulation and the anomalous temperature field. The residual between trends in the left- and right-hand sides of (6.5) and (6.6) is generally smaller by at least an order of magnitude in each case, indicating that the decomposition is approximately closed (not shown).

Clark et al. (2022) find the same terms identified above are important in driving trends in moist static energy transport by the MMC and eddies (i.e. primarily DYN for the MMC, and a combination of all three for the eddies). This indicates that the tendency trends cannot be explained by an unchanged circulation passively acting on changing temperature gradients; rather, for both the MMC and the eddies, the circulation is actively changing and modifying the temperature field.

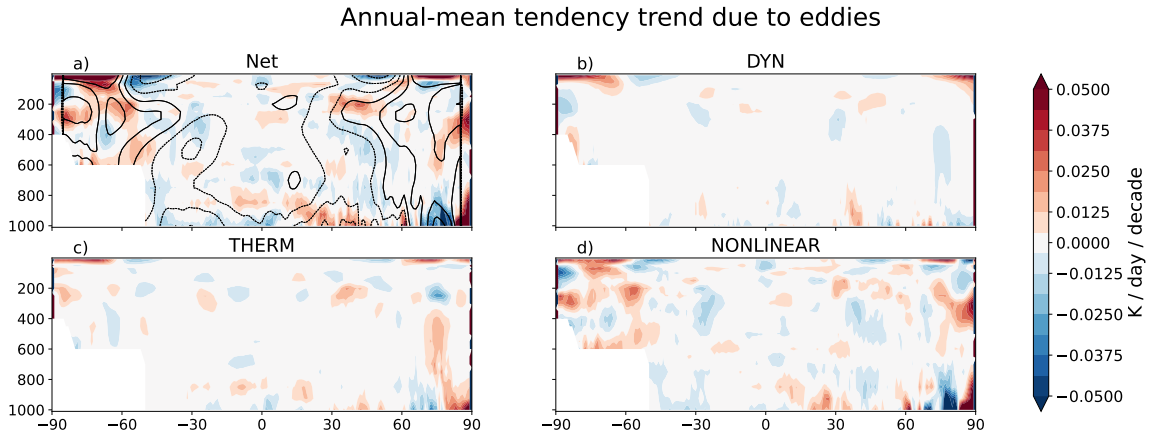


Figure 6.5: **Drivers of the tendency trend due to the eddies**. (a) Net annual-mean trend in temperature tendency due to eddies; (b)-(d) contributions to the net trend in (a) from dynamic, thermodynamic, and nonlinear anomalies respectively. Colours show the trend in K / day / decade. Black contour lines in (a) show the time-mean with spacing of 0.35K / day and solid lines indicating positive values. Unlike in Figure 6.3, the whole 1979-2020 period is used here.

6.5 Seasonal tendency trends

The accumulated tendency approach allows the direct estimation of the dynamical contribution to the annual mean temperature trend. However, it also conflates the contribution in different seasons, since the heating accumulated over the course of one season remains at the beginning of the following season. To study seasonal trends independently, trends in the unintegrated tendencies are plotted here for each season. As shown in the previous sections, trends in the diabatic term are essentially equal in size and opposite in sign to the net dynamics and are omitted for brevity.

Substantial seasonal variation is seen even at low latitudes, where the MMC is the largest contributor to the total dynamical heating. In MAM and SON, the cooling due to ascent in the zonal-mean ITCZ narrows and intensifies, as previously reported in the Pacific region (Wodzicki and Rapp 2016, Figures 6.6b, 6.6d and 6.7b, 6.7d). The most striking trend in the ITCZ is seen in DJF (Figures 6.6a and 6.7a), with a deep dynamical cooling trend due to ascent in the zonal mean.

Notable trends in the tropical and subtropical dynamical heating are seen in the transition seasons, SON and MAM (Figures 6.6b and 6.6d). In the climatology, the tropical dynamical heating in both seasons is broadly hemispherically symmetric. The trends in SON are asymmetric however, with a strengthening of the subtropical warming in the SH resembling the JJA climatology. Similarly, although MAM trends

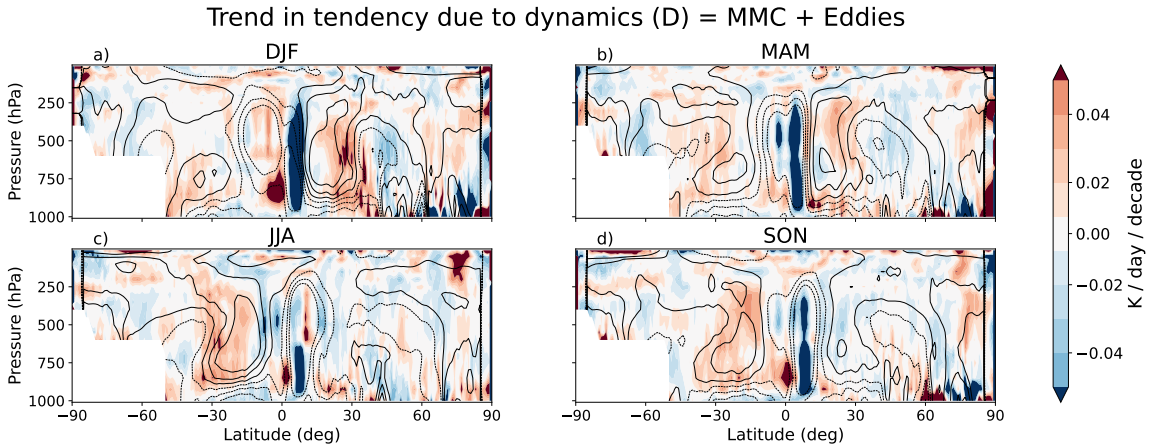


Figure 6.6: **Seasonal trends in temperature tendency due to the net dynamics**. Trends in (a) DJF, (b) MAM, (c) JJA, and (d) SON. Colours show the trend in K / day / decade. Black contour lines show the time-mean with spacing of 0.35K / day and solid lines indicating positive values.

are more symmetric about the equator than SON, the subtropical descent strengthens more in the NH, pushing MAM towards the DJF mean. Overall, the SH Hadley cell strengthens to varying degrees in every season except DJF, where trends are weak.

Further hemispheric asymmetry is seen in the midlatitude storm tracks. In the climatology, the eddies transport heat poleward and upward in the midlatitudes, cooling the mid-troposphere and warming the upper troposphere (Figure 6.2). This pattern strengthens somewhat in winter in the SH (Figure 6.8c) and substantially in the NH (Figure 6.8a). By contrast, the summertime storm track heat transport strengthens in the SH (Figure 6.8a) but is weakened in the NH (Figure 6.8c), consistent with studies showing a weakening of the NH summer storm track over recent decades (Coumou et al. 2015; Chang et al. 2016; Gertler and O'Gorman 2019).

While the MMC heating trends are largest in the tropics, some non-negligible trends are found even in the extratropics (Figure 6.7). In all seasons, a robust strengthening of the climatological MMC cooling is seen near 60°S, at the ascending branch of the Ferrel cell. This is consistent with a poleward strengthening of the SH Ferrel cell since 1979 (Chen et al. 2011) and its modelled response to CO2 forcing (Rojas 2013). The cooling appears well connected to the stratosphere in DJF (Figure 6.7a), which may indicate a stratospheric influence in this season. As discussed in previous chapters, stratospheric ozone depletion has partly driven the positive summer Southern Annular Mode trend (Thompson and Solomon 2002; Gillett and Thompson 2003). However, the consistency of the cooling trend at 60°S across all seasons suggests that mechanisms other than ozone depletion are also important here,

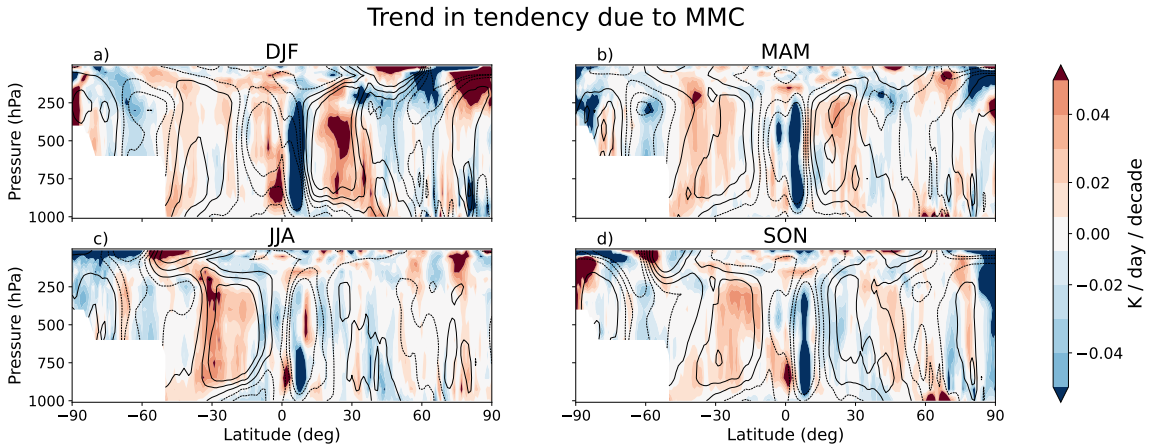


Figure 6.7: **Seasonal trends in temperature tendency due to the MMC**. Trends in (a) DJF, (b) MAM, (c) JJA, and (d) SON. Colours show the trend in K / day / decade. Black contour lines show the time-mean with spacing of 0.35K / day and solid lines indicating positive values.

as suggested by the role for internal variability in the hindcast SH jet trends (Section 3.5).

Over the Arctic a distinct seasonal pattern is seen. In SON, DJF, and MAM, a dipolar trend emerges consisting of subpolar dynamical cooling, and heating over the polar cap (Figures 6.6a, 6.6b, and 6.6d). By contrast, the subpolar region is warmed by the dynamics in JJA. This may reflect the seasonality of heat fluxes associated with the trend in sea ice loss. Ice is primarily melted during the warm summer months, resulting in a net downward heat flux out of the atmosphere (Serreze et al. 2009). As the Arctic lower troposphere warms with climate change, summer sea ice loss is enhanced and the trend is towards increased diabatic heat loss from the atmospheric column over the ice edge. We find that the trend in heat lost from the atmosphere to the ice and ocean in JJA is largely compensated by dynamical heating (Figure 6.6b).

As sea ice recedes year-on-year, more ocean surface is exposed over the cold period from early autumn until the following summer. A larger area of the cool winter atmosphere can thus be heated from below by the comparatively warm subpolar ocean. Figures 6.6a, 6.6c, and 6.6d suggest the atmosphere redistributes this heat poleward, largely through eddy heat flux convergence (compare the subpolar surface regions in Figures 6.6a, 6.6c, 6.6d with 6.8a, 6.8c, 6.8d). In a model with imposed sea ice loss, Deser et al. (2010) also find that horizontal eddy mixing is the dominant process in redistributing heat released from the Arctic Ocean over the high latitude continents.

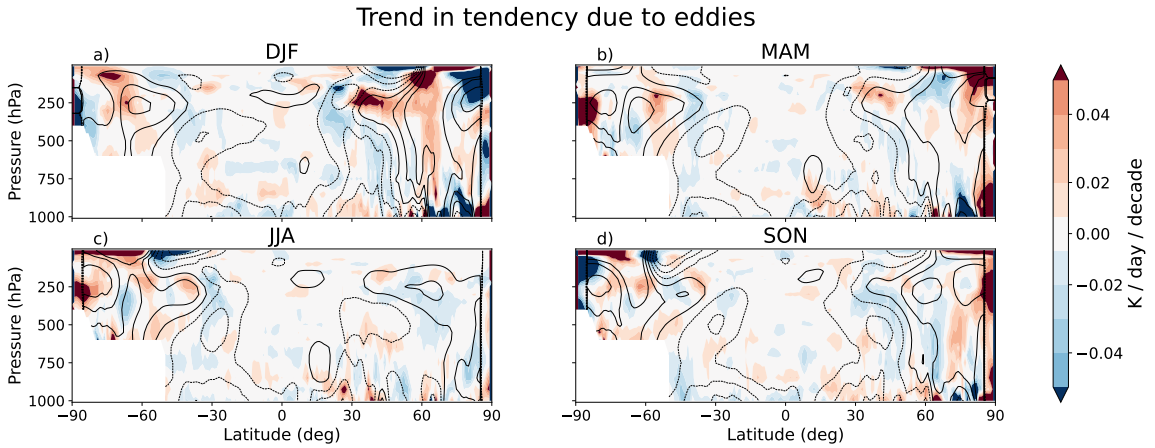


Figure 6.8: **Seasonal trends in temperature tendency due to the eddies**. Trends in (a) DJF, (b) MAM, (c) JJA, and (d) SON. Colours show the trend in K / day / decade. Black contour lines show the time-mean with spacing of 0.35K / day and solid lines indicating positive values.

The strongest dynamical cooling in the subpolar region occurs in DJF (Figure 6.6a), when the air-sea thermal contrast is greatest and the upward surface heat flux into the atmosphere peaks (Screen and Simmonds 2010b; Deser et al. 2015). In SON and DJF, the trend in eddy transport is relatively shallow, possibly due to the low-level winter inversion (Deser et al. 2010); in MAM, when inversions are less frequent and intense (Zhang et al. 2011), the eddy heating deepens and heat is mixed vertically to a greater extent.

6.6 Significance testing

In this section, significance testing is performed for a subset of the fields above. Prior chapters have used monthly data, averaged to give seasonal means. Consequently, temporal autocorrelation between successive data points was found to have a negligible impact on the significance of trends. However, the budget method uses daily data, and the budget term timeseries are highly autocorrelated in both the tendency and accumulated forms. Initial tests which did not account for this showed highly significant trends across most of the field (not shown). In this section, a statistical test is used which explicitly accounts for the reduced degrees of freedom due to the temporal autocorrelation, in order to avoid spuriously high estimates of significance.

In the plots below, a two-tailed Student's t-test is performed to detect trends significantly different from 0 at the 5% level using the `scipy.stats.t` class. In an

uncorrelated timeseries of length N, the degrees of freedom $df = N$. An autocorrelation timescale of 10 days is used to estimate the reduced degrees of freedom here $df = N/10$ (Lorenz 1982).

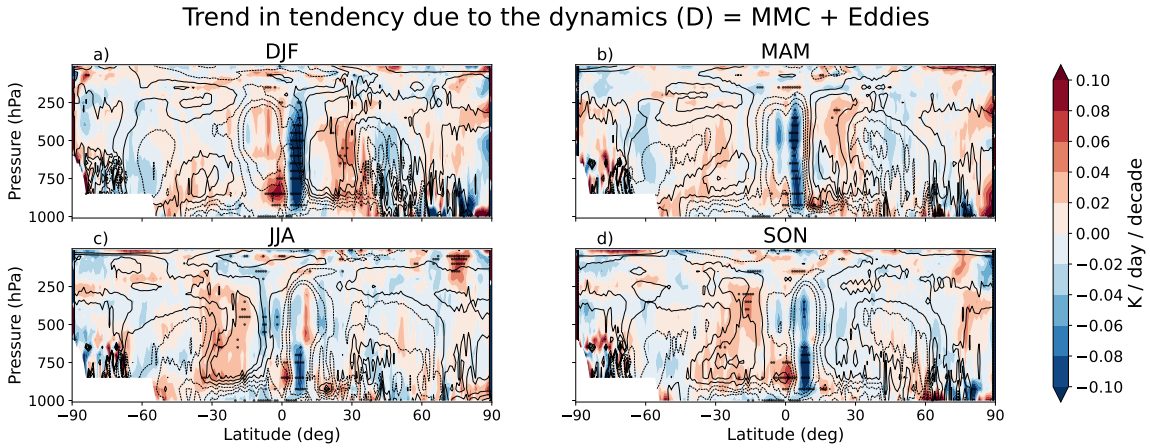


Figure 6.9: **Significance testing of the dynamical tendency trends.** As in Figure 6.6, but with stippling indicating significant trends at the 5% level using a two-tailed Student's t-test, accounting for autocorrelation in the tendency timeseries. Colours show the tendency trend due to the dynamics in K / day / decade. Black contour lines show the time-mean with spacing of 0.35K / day and solid lines indicating positive values.

Figure 6.9 shows seasonal trends in the dynamical component of the tendency budget, with significant trends indicated by the stippling. Very few regions achieve statistical significance using this test. Some residual significance is seen in the tropics and subtropics, particularly in the ITCZ and the descending branches of the Hadley Cells. However, as noted above, concerns have been raised over the robustness of reanalysis trends in tropical diabatic heating and overturning circulations, which casts doubt on even these trends. For brevity, only the total dynamical tendency trends are shown; the lack of significance is the same for the individual MMC and eddy components, and for terms in the accumulated budget.

The budget terms exhibit high variability on a range of timescales, with the time period used here potentially too short to demonstrate statistically robust trends. Reasons for the low statistical significance, and potential solutions, are discussed in the following section.

6.7 Discussion and conclusions

The structure of the trend in Figure 6.1 is well understood as being driven by GHG. However, at any given location this will be realised in the balance between several different processes, with the dynamics acting to spread out the heating effect from the diabatics. This chapter aimed to quantify the contribution of the atmospheric circulation to the zonal-mean temperature trend in ERA5, using a novel budget-based method. Starting with the heat budget for the local temperature tendency, and by considering trends in the accumulated heating over time, the ERA5 temperature trend is decomposed into a component due to the circulation and a component due to diabatic processes (all in units of K/decade). The main results are as follows:

1. **Well-balanced budget:** In the tendency budget, the time-mean net tendency is small compared to the individual heating components (Figure 6.2); likewise, it is shown that the net temperature trend is much smaller than trends in the accumulated diabatic and dynamical heating (c.f. Figure 6.1 and Figure 6.3). For decadal trends in temperature, the accumulated diabatic heating is balanced by the circulation to a high degree of precision (Section 6.3, Figure 6.3). Without any circulation at all, the temperature trends would reflect the highly localised diabatic heating; by redistributing this heat through the atmosphere, the dynamics act to smooth the resulting temperature field and its trend.

2. **Features of the budget terms:** Trends are identified in the dynamical and diabatic heating contributions in ERA5.

   (a) In the annual mean, the accumulated heating in the ITCZ intensifies and narrows, and the dynamical heating in the regions of subtropical descent is reinforced in each hemisphere (Figure 6.3c).

   (b) In the SH, a deep accumulated cooling trend is seen in the annual-mean MMC at 60°S (Figure 6.3c), and in each season in the tendency budget (Figure 6.7), consistent with previous work that noted a poleward strengthening of the SH Ferrel cell over recent decades (Chen et al. 2011).

   (c) In the NH, the accumulated dynamical cooling associated with heat transport in the storm track strengthens in the annual-mean (Figure 6.3d). This cooling trend is reflected in the tendencies in winter and the transition seasons, but is reversed in NH summer, in line with observational and modelling studies (Figure 6.8; Coumou et al. 2015; Chang et al. 2016; O'Gorman 2010).

(d) Over the Arctic, the results are consistent with the circulation transporting heat associated with sea ice loss poleward from September to March. When heat is lost from the warm summertime atmosphere to melt sea ice, the dynamics compensates (Figures 6.6 and 6.8).

3. **Dynamical decomposition of budget terms:** It is shown that trends in the heating contributions cannot be understood by an unchanged circulation passively advecting the altered temperature field; rather, the results indicate that the circulation in ERA5 has actively adjusted in conjunction with the temperature field (Figures 6.4 and 6.5, Section 6.4).

It is also important to note some limitations of this work. An important consideration concerns how to interpret the tendency budget (6.1) in reanalysis products. In a model, the total temperature tendency is given exactly by the sum of the tendencies due to the dynamics and diabatic processes. In a reanalysis product, the total tendency on the LHS of (6.1) is calculated directly from analysis temperature fields, and so will include the analysis increments used to update the background forecast when assimilating observational data. Since the increments are not explicitly calculated, they will be included in the budget residual. This is a potentially important caveat concerning interpretation of the diabatic term. Future work should instead calculate the diabatic term explicitly by summing the reanalysis diabatic tendencies (e.g. from radiation and model physics); comparing this to the residual above would indicate whether errors introduced from the analysis increments are important. Nevertheless, this chapter has primarily focused on trends in the dynamical terms, which are calculated directly and are unaffected by the residual approach. It is also notable that a tendency budget similar to (6.1) has previously been applied to reanalysis data, with the diabatic term calculated as a residual and no consideration of analysis increments (e.g. Seager et al. 2003).

As discussed in previous chapters, reanalyses can also suffer from discontinuities associated with changes in assimilated observing systems, leading to spurious trends (Bengtsson et al. 2004). A particular concern here regards trends in tropical diabatic heating, which show much larger positive trends in reanalyses than are seen in model simulations of the satellite era, as discussed in previous chapters. As the dynamical and diabatic heating terms are tightly coupled in the temperature budget, errors in diabatic heating trends thus lead to errors in the dynamical trends too – providing a direct example of the mechanism suggested by Chemke and Polvani (2019). In the present generation of reanalyses, the utility of the budget approach is therefore limited

in the tropics; however, anticipating future improvements in reanalysis products, this technique may prove useful in future.

The third key limitation of this work concerns the lack of statistical significance for most of the budget term trends (Section 6.6). It is worth underlining that, although trends in the individual budget terms largely lack significance, trends in their sum – the total temperature – are highly significant throughout much of the atmosphere (Figure 6.1). The lack of significance refers to the difficulty in robustly estimating the relative contributions of the different budget terms to the total trend. Below, some of the contributory factors are identified, and potential solutions are discussed:

1. Trends are calculated over a relatively short time period (the satellite era). For highly variable quantities, such as dynamical heat flux convergences, this may simply be too short a record to identify robust trends.

    - **Solution 1:** Use extended reanalyses which include the pre-satellite era (e.g. the NOAA 20th Century Reanalysis, Slivinski et al. 2019). However, trends in these datasets are likely more poorly constrained than in the observation-rich satellite era, particularly for the high frequency wind data required for this budget. Also, trends are likely strongest over the recent period anyway, so using pre-satellite data may not help.
    - **Solution 2:** Revisit this approach in future with a longer observational record and improved reanalyses.

2. The observed climate system is only a single realisation, and ERA5 is only a single, imperfect representation of it.

    - **Solution 1:** Future work could compare trends in different reanalysis products to see if they are consistent with those in ERA5 (e.g. the JRA-55 or JRA-3Q reanalyses used in Chapter 2). This would test the sensitivity of the results above to differences in the reanalysis forecast models; for example, vertical velocity ω is outputted as a forecast field of the reanalysis, and therefore may be less well constrained between different products. However, these are not independent datasets and assimilate many of the same observations, so this is not a full solution.
    - **Solution 2:** Exploit large model ensembles, as in previous chapters. A seemingly logical extension to this work would be to implement the budget method in the DePreSys3 hindcasts. However, they lack the submonthly

data needed here. Future work could instead look in a high frequency hindcast dataset (e.g. DePreSys4) with multiple members, allowing robust statistics to be applied to the budget trends.

3. The observed forcing over the satellite era is too weak to drive robust trends in the budget terms.

    - **Solution 1:** Adapt the method to study projected trends under more strongly forced future scenarios. This method may find more statistically robust trends under stronger forcing scenarios.
    - **Solution 2:** The period of strong anthropogenic forcing appears set to continue, and so revisiting this method in future may allow more strongly forced trends to emerge.

A final point of consideration concerns the budget itself. For example, Clark et al. (2022) found significant trends in moist static energy (MSE) transport using a similar budget approach. However, they considered vertically integrated trends, which may have increased the significance. Alternatively, the budget could be reformulated using the transformed Eulerian-mean (TEM) (e.g. Palmer 1981). Although the net dynamical heating would be unchanged using the TEM, it may offer greater insight in the midlatitudes, where our Eulerian framework separates the baroclinic eddies from the meridional circulation they drive.

This chapter has quantified the local contribution of the dynamics as a whole to the zonal-mean temperature trend. Anomalies in circulation and diabatic heating are clearly tightly coupled, with the impact of anomalous diabatic heating being smoothed out by the circulation. Conversely, it is also true that anomalous circulation can result in diabatic heating. The latter is clearly important, since the diabatic heating terms do not look like the simple radiative change that may be expectd from increased GHG foring. Instead, strong changes are seen in the distribution of diabatic heating, for example in latent heating associated with anomalous ascent.

Ultimately, forced changes in the climate system, such as greenhouse warming due to anthropogenic climate change, manifest as a diabatic imbalance in the Earth system's heat budget. While the driving of these changes ultimately comes from diabatic processes, the characteristics of the circulation can shape the resulting warming structure, for example by influencing the regions and length scales over which heat is redistributed.

Chapter 7

Conclusions

7.1 Summary of results

This thesis has sought to develop tools to study satellite-era trends. The main results are summarised below.

Shifting jets: improved winter jet trends, but internal variability remains large.

The hindcasts support the emergence of poleward jet shifts on the global scale and in each season (Woollings et al. 2023), with 85% and 96% of 1m-lead hindcasts showing a simultaneous poleward shift in both hemispheres for DJF and JJA, respectively. Common to all seasons are stronger subtropical easterly trends in the hindcast ensemble-mean than in the free-running GC2, which may result from the La Nina-like SST trends in the hindcasts. The effect of the initialisation on jet trends varies seasonally and by hemisphere. Broadly, winter jet trends are improved by the initialisation whereas the summer jets are not. However, the improvement in winter appears to be largely lost by 6m lead, possibly due to the redevelopment of the model's equatorward bias which is reduced through initialisation.

A wide spread is seen in the hindcast jet shifts, and the spread does not change substantially with lead time. In fact, seasonal variability can lead to shifts opposite in sign to the externally forced response. An example is the 8% of members that show an equatorward jet shift in austral summer despite the ozone signal, supporting the conclusions of Seviour et al. (2017). Intriguingly, in the hindcasts, poleward jet shifts are associated with strengthened meridional temperature gradients in austral winter, but weakened gradients in boreal winter.

Tropical tropospheric warming: hindcasts capture observed trends after accounting for surface warming, with the tropical Pacific contributing.

Discrepancies in tropical tropospheric warming are substantially reduced in the hindcasts compared to GC2. Upper-level temperature trends in the models are found to scale linearly with trends in pTAS, a precipitation-weighted index of mean surface temperature over tropical ocean. The scaling not only explains the hindcast trends at both 1m and 6m lead, but also trends in the free-running GC2. We therefore conclude that the improvement in tropospheric warming trends following initialisation is due to the hindcasts having cooler tropical SST trends, particularly in regions of deep convection. The hindcasts appear to be consistent with upper-level trends in the reanalyses after accounting for differences in their pTAS trends, particularly given uncertainty in the reanalysis trends. Given the prominence of deep convection over the tropical Pacific, it seems plausible that tropical-mean pTAS trends will be dominated by the Pacific sector. In Chapter 5, we find that in models that better capture observed SST trends in the tropical Pacific (Pacemaker and the conditioned hindcasts), reanalysis tropospheric warming trends are consistent with the range of model trends.

Arctic amplification: a wide spread of trends, largely due to seasonal variability.

Arctic amplification of warming is seen in at least 99% of the hindcasts, both at surface level and in the lower troposphere. However, while surface AA in the reanalyses is well captured by the hindcasts, at most 6% of the hindcasts capture the reanalysis LT AA, likely due to the model's stable bias in the Arctic boundary layer.

Irrespective of definition, the Arctic is characterised by a wide spread of AA and warming trends. Sea ice trends are constrained by the initialisation and explain very little of this spread ($r \lesssim -0.4$); seasonal variability, likely atmospheric in origin, appears to contribute substantially. Hindcast members with the strongest Arctic warming show anticyclonic trends over northern Eurasia, delivering warm air to the Arctic and cooling the continental interior, and an equatorward shift of the zonal-mean eddy-driven jet. These results do not imply a direction of causality between the Arctic and midlatitude circulation, but are consistent with an important role for variability in the warm arctic, cold Eurasia trend pattern, as suggested by McCusker et al. (2016), Ogawa et al. (2018) and others. Between $11-20$% of the hindcasts have Eurasian winter cooling trends as large as the observed trend, with the magnitude

and location of trends in individual members varying, consistent with an internally generated pattern of variability.

Tropical Pacific trends: robust global impacts and improved agreement with reanalyses.

Most 1m hindcasts show a strengthening of the tropical Pacific SST gradient, but by 6m trends are substantially worse, as the model drifts towards weaker trends in the SST gradient. It is then shown that the Pacific has a large and robust impact on global circulation trends, including the weakened tropical tropospheric warming discussed above. The Pacific also contributes on average to poleward jet shifts in each hemisphere, although the spread in hindcast jet trends remains large even in the conditioned models; factors outside the tropical Pacific are therefore also important. Regionally, capturing the Pacific SST trend improves agreement with temperature and precipitation trends in regions traditionally affected by La Nina on interannual timescales, supporting recent work by Qiu et al. (2024). However, not all discrepancies are resolved; for example, Gulf Coast drying is still underestimated in the conditioned runs, pointing to potential deficiencies in models beyond the Pacific.

Hadley Circulation: a weakening circulation in NH winter.

None of the 10,000 1m-lead hindcasts show a strengthening of the NH winter Hadley Cell (HC), and only 5% at 6m lead do. The hindcasts therefore lend support to a weakening of the circulation, in line with trends in sea-level pressure and contrary to the strengthening trends in reanalyses (Chemke and Polvani 2019; Chemke and Yuval 2023). The strengthening in ERA5 has been argued to be partly physical and related to the evolution of Atlantic decadal variability (Zaplotnik et al. 2022). However, we find that the weakening is more robust in the hindcasts, where decadal variability is constrained, than in the free-running GC2, where the evolution of coupled variability differs more between members. Rather than strengthening the DJF HC, this suggests that internal variability has weakened it.

Summary of other important results

GMST trends: Hindcasts support positive trends in GMST in both DJF and JJA. The spread in SAT trends is small over ocean but much larger over land, where it is associated with trends in the annular modes. Although interannual-decadal variability contributes much more to GMST (e.g. Kosaka and Xie 2013), these results are a useful reminder that the influence of seasonal and subseasonal variability is not negligible.

NAO: The satellite-era NAO trend is consistent with the hindcasts, and even with the 5 free-running GC2 members. On the other hand, the negative summer NAO trend is on the edge of the range of hindcast trends at both lead times, potentially reflecting the shared aerosol forcing. For both summer and winter, a wide spread of NAO trends is seen, consistent with our understanding of a substantial fraction of NAO variability arising through internal atmospheric dynamics.

Integrated temperature budget: A novel integrated budget method is developed in Chapter 6 for directly quantifying the dynamical contribution to temperature trends. The dynamical and diabatic contributions to temperature trends largely balance each other, as they do in the tendency budget. Trends in the ERA5 budget terms are not significant after accounting for temporal autocorrelation. However, the total warming trend due to GHG forcing is significant; the difficulty is in identifying significant trends in the relative contributions at each location. Nevertheless, some trends do appear to align with physical expectations, such as a weakening of heating in the NH summer storm tracks, and an eddy-mediated poleward transport of heat from the subpolar region during autumn-winter.

7.2 Themes and open questions

This section discusses some common themes that emerged in this thesis, as well as some questions that remain open.

Novel contributions

The overarching theme of this work has been to develop methods to study climate trends. The integrated budget approach in Chapter 6 is a novel technique. Although the results lack significance in ERA5, it can be considered successful in having diagnosed the dynamical contribution to the well-known zonal-mean temperature trend.

Building ensembles of trends from near-term hindcasts is not an entirely novel approach, having been previously used to study "unseen" trends in regional extremes (Kelder et al. 2020; Berghald et al. 2024). However, it has not been applied to the type and breadth of features in this thesis. Most work on trends in hindcasts focuses on the ensemble-mean trend (e.g. L'Heureux et al. 2022; Beverley et al. 2024b), typically to understand its impact on hindcast skill. Using the *spread* of hindcasts trends to provide contextual information for trends in the large-scale circulation is a

novel contribution, as is the clean comparison with an otherwise-identical uninitialised model (GC2).

Role of internal variability

A recurring theme of this work is the wide range of trends that are possible through internally generated seasonal variability. This is strikingly clear in the Arctic, but seems to be the case to some degree for nearly every feature investigated. The importance of internal variability for climate trends is not a novel observation (e.g. Deser et al. 2012), but the very large ensemble size and constrained nature of the hindcast ensemble affords this thesis a unique perspective on its role. It is worth noting that, despite the wide spread of trends, it is still possible to make robust statements regarding the hindcast trends: for example, the hindcast ensemble does not permit negative GMST trends, surface AA< 1, or, notably, a strengthening of the NH winter Hadley Cell. The hindcasts also allow probabilistic estimates of the likelihood of trends in highly variable regions, such as the 85% of hindcast members that show concurrent poleward jet shifts in DJF.

Implications for studies of hindcast trends

Hindcasts inherit the observed phasing of coupled variability from the initialisation (Figure 3.1). Discrepancies between observed trends and the hindcast ensemble-mean have therefore been interpreted in the literature as "trend errors", since they are argued to be unrelated to differences in variability between the model and the real world. However, a key message of this thesis is that there is not a single trend in a given hindcast model (the ensemble mean), but rather a range of possible trends, with the spread arising due to variability on short timescales. This complicates statements about the model's trend error. For example, the mean boreal winter jet shift is weaker in the 1m hindcasts than in the reanalyses – but a quarter of the hindcasts actually show a stronger poleward shift than in the reanalyses. Should an equatorward trend error be diagnosed in this case? The observed trend contains contributions from seasonal and subseasonal variability in addition to slower modes; unless the former contributions are expected to be negligible, then the hindcast ensemble-mean gives only a partial picture. This likely explains why the largest ensemble-mean trend errors in seasonal hindcasts are found in dynamical variables (Beverley et al. 2024b), where atmospheric variability has presumably contributed meaningfully to the observed reference trend.

7.3 Limitations and future work

There also remain several limitations to this work. The first two below are unavoidable features of the hindcast approach, whereas statistical and data-related limitations are solvable in principle. Possible extensions to this work are also discussed.

Lack of slow feedbacks

Arguably the main limitation of the hindcast approach is that the resulting trends are statistical constructions, rather than trends in a continuous model integration. Firstly, this means that slow feedbacks which couple successive seasons are absent. This is seen in the Arctic, where the lack of interseasonal ice-ocean-atmosphere feedbacks leads to a weaker AA-sea ice scaling in the hindcasts than in the free-running GC2. Whether this leads to an over- or underestimation of the range of possible trends will depend on whether the feedback amplifies or dampens the initial signal, which will vary by region. In the Arctic, the AA-sea ice feedback is positive, and a larger GC2 ensemble may ultimately show a wider range of AA trends than the hindcasts because of this. On the other hand, it seems possible that in other regions, the spread is overestimated; a rule that applies across all regions is unlikely.

Secondly, it could be objected that the statistical timeseries may be unrealistic, particularly at the extremes of the distribution. One way of quantifying this would be to systematically apply the more rigorous quality control steps introduced by Kelder et al. (2020), which verify that the statistical moments of the observed timeseries are consistent with those of the synthetic hindcast timeseries. The DePreSys3 hindcasts passed these benchmarks for the handful of individual features that were tested (e.g. DJF HC strength and AA; not shown), building confidence in their suitability for this study. In summary, with the caveat that they are not interchangeable with continuous integrations, the hindcasts can still offer a useful point of comparison for observed trends.

Signal-to-noise issues

Recent years have seen the identification of a low ratio of predictable signal to noise in seasonal-to-decadal (S2D) models (Scaife et al. 2014; Eade et al. 2014; Smith et al. 2020), at least since the latter part of the 20th century (Weisheimer et al. 2020). Multiple potential contributing factors have been identified (Weisheimer et al. 2024), such as an underestimation of tropical-extratropical teleconnection strength (Williams et al. 2023). If the response to internal drivers such as ENSO is too weak

in DePreSys3, then the range of hindcast trends in the affected regions may be biased. This could affect conclusions regarding, for example, the weak influence of tropical Pacific trends on Arctic warming (Lee 2012). As evidence of the signal-to-noise issue has been found in various S2D models, including DePreSys3 (Dunstone et al. 2016), it is at present an unavoidable limitation of this approach. However, a mitigating factor is that it primarily affects the Atlantic (Hodson et al. 2023); while regional hindcast trends may be impacted, its influence on trends on the global scale is likely more limited.

Statistical limitations

Linear trends in this thesis have been calculated using ordinary least squares regression, which is sensitive to outliers. Multidecadal trends can also be sensitive to end point values. Future work could therefore consider using regression methods that are robust to outliers, such as the Theil-Sen estimator (Theil 1950). Additionally, future researchers may wish to test the sensitivity of the trends to different endpoints, such as by varying the start and end year of the trends. Timeseries plots such as in Fig. 3.1 could also help highlight the effect of outliers and endpoint biases.

Data limitations

Single hindcast model: The hindcast results may be sensitive to the particular biases of DePreSys3 and may not generalise to other models. Many other near-term prediction models exist (e.g. DCPP, Boer et al. 2016) which could be used to verify the results in this thesis; in particular, it may be interesting to compare the effect of different initialisation protocols, as anomaly initialisation preserves the mean model bias.

Small free-running ensemble: Only 5 Historical simulations were available at the same N216 resolution as the hindcasts. This makes it difficult to be fully confident that an observed trend is inconsistent with the model, except when the difference is egregiously large (e.g. tropospheric temperature trends).

Monthly data: The choice of features to investigate was constrained by the lack of submonthly data in the hindcasts. Limited higher frequency data is available for DePreSys4, which would allow the investigation of trends in variability (e.g. the heat fluxes of Woollings et al. (2023), or weakening storm tracks (Chemke and Coumou

2024)). Higher frequency data could also be necessary for identifying links between trends in different regions that are mediated by transient eddies, such as potential tropical-AA links outlined by Lee (2014).

Satellite era trends: Future work could consider extending the window back to include trends earlier in the 20th century, either using DePreSys3 itself (which begins in 1960), or other 20th century hindcasts (Weisheimer et al. 2020). For example, free-running models struggle to capture the positive NAO trend from the 1960-1990s. Do initialised models fare better? This may also lead to higher significance for trends in the budget terms of Chapter 6.

Regional focus: Although much of this thesis has focused on trends that are hemispheric or global in scale, perhaps the greatest benefit of the hindcasts could be in providing context for emerging regional trends (Shaw et al. 2024), as was done for the Eurasian winter cooling trend. This could be particularly useful for impact-related trends that are poorly captured by smaller ensembles of free-running models, such as the period of declining East African rainfall in the last century (Rowell et al. 2015).

Appendix A

Extended Figures for Chapters 3 and 4

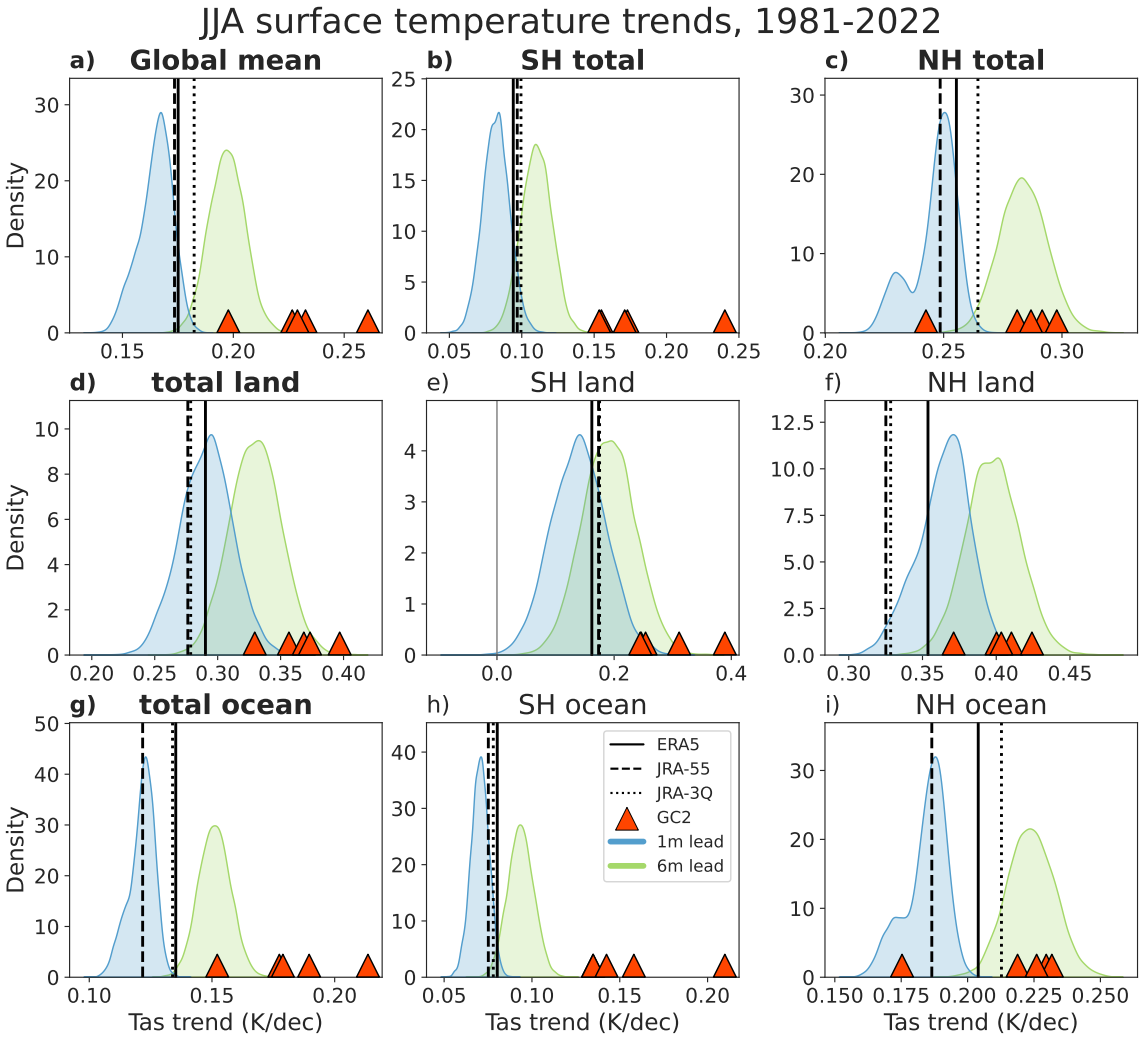


Figure A.1: As in Figure 4.1, but for JJA.

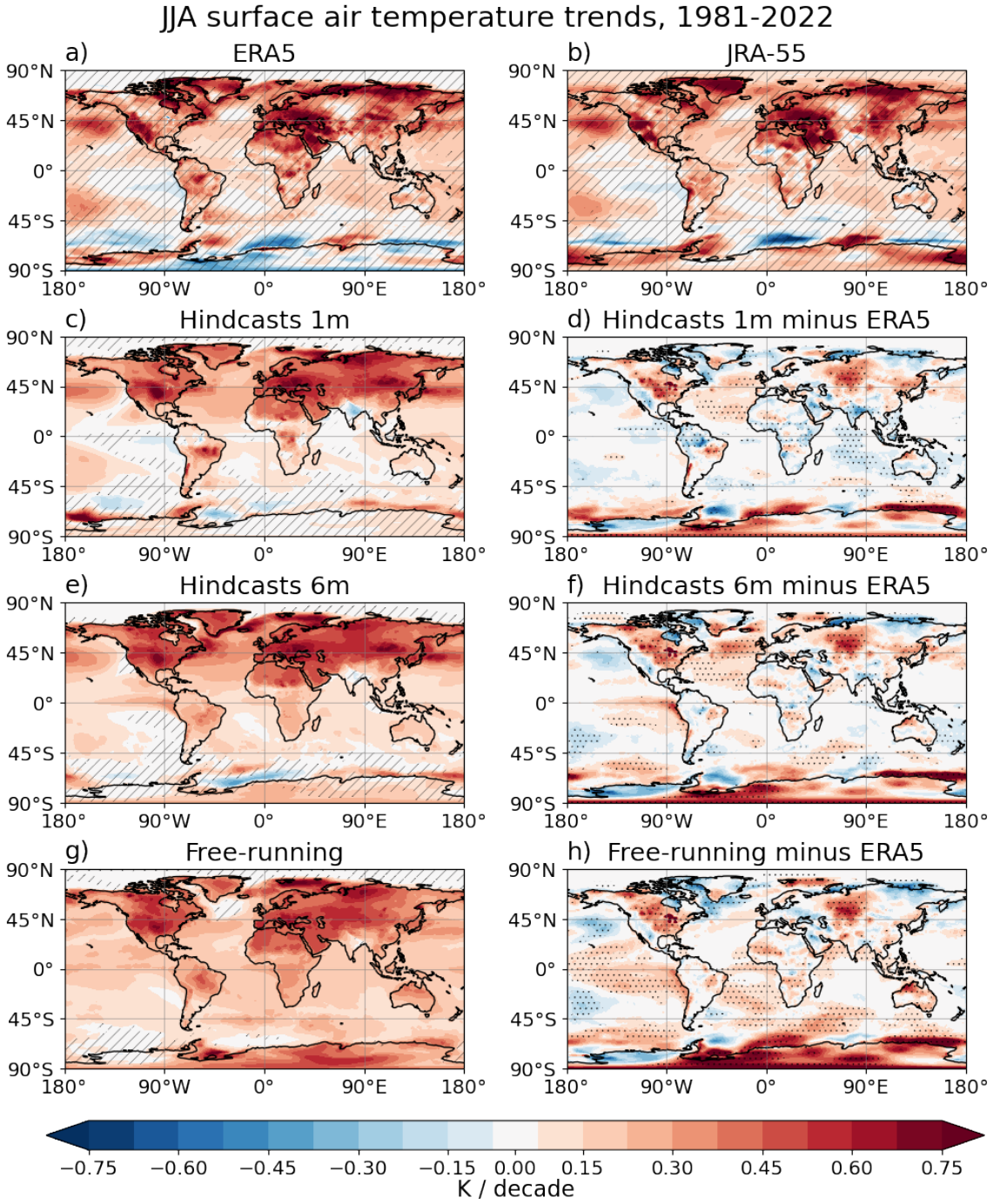


Figure A.2: **Maps of ensemble-mean hindcast SAT trends in JJA.** Trends (1981-2022) for (a) ERA5, (b) JRA-55; (c), (e) the hindcast ensemble-means at 1m and 6m lead respectively; and (g) the free-running ensemble-mean. Hatching on these plots is the inverse of the stippling in Figure 3.2 (here, hatching *hides* non-significant points, defined as before). (d, f, h): Differences of the plots to the left from the ERA5 trend in (a). As in Figure 3.2, stippling on (d), (f) and (h) indicates where the hindcast mean is more than $2S$ away from the ERA5 mean, where S is the standard error on the ERA5 regression.

Bibliography

Allen, Myles R. and William J. Ingram (2002). "Constraints on future changes in climate and the hydrologic cycle". *Nature* 419.6903, pp. 224–232. ISSN: 1476-4687. DOI: 10.1038/nature01092.

Baldassare, Daniel et al. (2023). "Large uncertainty in observed estimates of tropical width from the meridional stream function". *Weather and Climate Dynamics* 4.2, pp. 531–541. DOI: 10.5194/wcd-4-531-2023.

Bamston, Anthony G., Muthuvel Chelliah, and Stanley B. Goldenberg (1997). "Documentation of a highly ENSO-related SST region in the Equatorial Pacific". *Atmosphere-Ocean* 35.3, pp. 367–383. ISSN: 0705-5900. DOI: 10.1080/07055900.1997.9649597.

Banerjee, Antara et al. (2020). "A pause in Southern Hemisphere circulation trends due to the Montreal Protocol". *Nature* 579.7800, pp. 544–548. ISSN: 0028-0836. DOI: 10.1038/s41586-020-2120-4.

Barnes, Elizabeth A and James A Screen (2015). "The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?" *WIREs Climate Change* 6.3, pp. 277–286. ISSN: 1757-7799. DOI: 10.1002/wcc.337.

Barnes, Elizabeth A. and Lorenzo Polvani (2013). "Response of the midlatitude jets, and of their variability, to increased greenhouse gases in the CMIP5 models". *Journal of Climate* 26.18, pp. 7117–7135. ISSN: 08948755. DOI: 10.1175/JCLI-D-12-00536.1.

Barnes, Elizabeth A. and Lorenzo M. Polvani (2015). "CMIP5 Projections of Arctic Amplification, of the North American/North Atlantic Circulation, and of Their Relationship". *Journal of Climate* 28.13, pp. 5254–5271. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-14-00589.1.

Barsugli, Joseph J. and David S. Battisti (1998). "The Basic Effects of Atmosphere–Ocean Thermal Coupling on Midlatitude Variability". *Journal of the Atmospheric Sciences* 55.4, pp. 477–493. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1998)055<0477:TBEOAO>2.0.CO;2.

Becker, Emily J. et al. (2022). "A Decade of the North American Multimodel Ensemble (NMME): Research, Application, and Future Directions". *Bulletin of the American Meteorological Society* 103.3, E973–E995. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-20-0327.1.

Bengtsson, Lennart, Stefan Hagemann, and Kevin I. Hodges (2004). "Can climate trends be calculated from reanalysis data?" *Journal of Geophysical Research D: Atmospheres* 109.11. ISSN: 01480227. DOI: 10.1029/2004JD004536.

Berghald, S., S. Mayer, and P. Bohlinger (2024). "Revealing trends in extreme heatwave intensity: applying the UNSEEN approach to Nordic countries". *Environmental Research Letters* 19.3, p. 034026. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ad2893.

Best, M. J. et al. (2011). "The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes". *Geoscientific Model Development* 4.3, pp. 677–699. ISSN: 1991-9603. DOI: 10.5194/gmd-4-677-2011.

Beverley, J. D. et al. (2024a). "Drivers of Changes to the ENSO–Europe Teleconnection Under Future Warming". *Geophysical Research Letters* 51.10, e2023GL107957. ISSN: 1944-8007. DOI: 10.1029/2023GL107957.

Beverley, Jonathan D., Matthew Newman, and Andrew Hoell (2024b). "Climate model trend errors are evident in seasonal forecasts at short leads". *npj Climate and Atmospheric Science* 7.1, pp. 1–13. ISSN: 2397-3722. DOI: 10.1038/s41612-024-00832-w.

Bintanja, R., E. C. van der Linden, and W. Hazeleger (2012). "Boundary layer stability and Arctic climate change: a feedback study using EC-Earth". *Climate Dynamics* 39.11, pp. 2659–2673. ISSN: 0930-7575. DOI: 10.1007/s00382-011-1272-1.

Blackport, Russell and John C. Fyfe (2022). "Climate models fail to capture strengthening wintertime North Atlantic jet and impacts on Europe". *Science Advances* 8.45, eabn3112. DOI: 10.1126/sciadv.abn3112.

Blackport, Russell and James A. Screen (2020). "Weakened evidence for mid-latitude impacts of Arctic warming". *Nature Climate Change* 10.12, pp. 1065–1066. ISSN: 17586798. DOI: 10.1038/s41558-020-00954-y.

Boer, George J. et al. (2016). "The Decadal Climate Prediction Project (DCPP) contribution to CMIP6". *Geoscientific Model Development* 9.10, pp. 3751–3777. ISSN: 1991-959X. DOI: 10.5194/gmd-9-3751-2016.

Bony, Sandrine et al. (2013). "Robust direct effect of carbon dioxide on tropical circulation and regional precipitation". *Nature Geoscience* 6.6, pp. 447–451. ISSN: 1752-0908. DOI: 10.1038/ngeo1799.

Bordoni, Simona and Tapio Schneider (2010). "Regime Transitions of Steady and Time-Dependent Hadley Circulations: Comparison of Axisymmetric and Eddy-Permitting Simulations". *Journal of the Atmospheric Sciences* 67.5, pp. 1643–1654. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/2009JAS3294.1.

Bowler, Neill E. et al. (2009). "The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system". *Quarterly Journal of the Royal Meteorological Society* 135.640, pp. 767–776. ISSN: 00359009. DOI: 10.1002/qj.394.

Bracegirdle, Thomas J. (2022). "Early-to-Late Winter 20th Century North Atlantic Multidecadal Atmospheric Variability in Observations, CMIP5 and CMIP6". *Geophysical Research Letters* 49.11, e2022GL098212. ISSN: 1944-8007. DOI: 10.1029/2022GL098212.

Bracegirdle, Thomas J. et al. (2018). "Do CMIP5 Models Reproduce Observed Low-Frequency North Atlantic Jet Variability?" *Geophysical Research Letters* 45.14, pp. 7204–7212. ISSN: 1944-8007. DOI: 10.1029/2018GL078965.

Bretherton, Christopher S. and Piotr K. Smolarkiewicz (1989). "Gravity Waves, Compensating Subsidence and Detrainment around Cumulus Clouds". *Journal of the Atmospheric Sciences* 46.6, pp. 740–759. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1989)046<0740:GWCSAD>2.0.CO;2.

Breul, Philipp, Paulo Ceppi, and Theodore G. Shepherd (2023). "Revisiting the wintertime emergent constraint of the southern hemispheric midlatitude jet response to global warming". *Weather and Climate Dynamics* 4.1, pp. 39–47. ISSN: 2698-4016. DOI: 10.5194/wcd-4-39-2023.

Butler, Amy H., David W. J. Thompson, and Ross Heikes (2010). "The Steady-State Atmospheric Circulation Response to Climate Change–like Thermal Forcings in a Simple General Circulation Model". *Journal of Climate* 23.13, pp. 3474–3496. ISSN: 0894-8755. DOI: 10.1175/2010JCLI3228.1.

Cai, Ming et al. (2009). "The Role of Long-Term Trends in Seasonal Predictions: Implication of Global Warming in the NCEP CFS". *Weather and Forecasting* 24.4, pp. 965–973. ISSN: 1520-0434, 0882-8156. DOI: 10.1175/2009WAF2222231.1.

Chang, Edmund K. M. et al. (2016). "Observed and projected decrease in Northern Hemisphere extratropical cyclone activity in summer and its impacts on maximum temperature". *Geophysical Research Letters* 43.5. ISSN: 00948276. DOI: 10.1002/2016GL068172.

Po-Chedley, Stephen and Qiang Fu (2012). "Discrepancies in tropical upper tropospheric warming between atmospheric circulation models and satellites". *Environmental Research Letters* 7.4, p. 044018. ISSN: 1748-9326. DOI: 10.1088/1748-9326/7/4/044018.

Po-Chedley, Stephen et al. (2021). "Natural variability contributes to model–satellite differences in tropical tropospheric warming". *Proceedings of the National Academy of Sciences of the United States of America* 118.13. ISSN: 10916490. DOI: 10.1073/PNAS.2020962118/-/DCSUPPLEMENTAL.

Po-Chedley, Stephen et al. (2022). "Internal variability and forcing influence model–satellite differences in the rate of tropical tropospheric warming". *Proceedings of the National Academy of Sciences* 119.47, e2209431119. DOI: 10.1073/pnas.2209431119.

Chemke, R. and L. M. Polvani (2021). "Elucidating the Mechanisms Responsible for Hadley Cell Weakening Under 4 × CO2 Forcing". *Geophysical Research Letters* 48.3, e2020GL090348. ISSN: 1944-8007. DOI: 10.1029/2020GL090348.

Chemke, Rei and Dim Coumou (2024). "Human influence on the recent weakening of storm tracks in boreal summer". *npj Climate and Atmospheric Science* 7.1, pp. 1–8. ISSN: 2397-3722. DOI: 10.1038/s41612-024-00640-2.

Chemke, Rei and Lorenzo M. Polvani (2019). "Opposite tropical circulation trends in climate models and in reanalyses". *Nature Geoscience* 12.7. ISSN: 17520908. DOI: 10.1038/s41561-019-0383-x.

Chemke, Rei and Janni Yuval (2023). "Human-induced weakening of the Northern Hemisphere tropical circulation". *Nature* 617.7961, pp. 529–532. ISSN: 1476-4687. DOI: 10.1038/s41586-023-05903-1.

Chen, Junye, Barbara E. Carlson, and Anthony D. Del Genio (2002). "Evidence for Strengthening of the Tropical General Circulation in the 1990s". *Science* 295.5556, pp. 838–841. DOI: 10.1126/science.1065835.

Chen, Liqi, Mirong Song, and Jiping Liu (2011). "Ferrel Circulation variability in the Southern Hemisphere and its linkages with tropical and subtropical sea surface temperature". *Journal of Geophysical Research: Atmospheres* 116 (D12). ISSN: 2156-2202. DOI: 10.1029/2010JD015409.

Choi, Jung et al. (2016). "Seasonal-to-Interannual Prediction Skills of Near-Surface Air Temperature in the CMIP5 Decadal Hindcast Experiments". *Journal of Climate* 29.4, pp. 1511–1527. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-15-0182.1.

Chung, Eui Seok et al. (2019). "Reconciling opposing Walker circulation trends in observations and model projections". *Nature Climate Change* 9.5, pp. 405–412. ISSN: 17586798. DOI: 10.1038/s41558-019-0446-4.

Chung, Eui-Seok et al. (2024). "Multi-decadal climate variability and satellite biases have amplified model-observation discrepancies in tropical troposphere warming estimates". *Communications Earth & Environment* 5.1, pp. 1–8. ISSN: 2662-4435. DOI: 10.1038/s43247-024-01510-8.

Chylek, Petr et al. (2009). "Arctic air temperature change amplification and the Atlantic Multidecadal Oscillation". *Geophysical Research Letters* 36.14. ISSN: 1944-8007. DOI: 10.1029/2009GL038777.

Chylek, Petr et al. (2022). "Annual Mean Arctic Amplification 1970–2020: Observed and Simulated by CMIP6 Climate Models". *Geophysical Research Letters* 49.13, e2022GL099371. ISSN: 1944-8007. DOI: 10.1029/2022GL099371.

Chylek, Petr et al. (2024). "Why Does the Ensemble Mean of CMIP6 Models Simulate Arctic Temperature More Accurately Than Global Temperature?" *Atmosphere* 15.5, p. 567. ISSN: 2073-4433. DOI: 10.3390/atmos15050567.

Clark, Joseph P., Steven B. Feldstein, and Sukyoung Lee (2022). "Moist Static Energy Transport Trends in Four Global Reanalyses: Are They Downgradient?" *Geophysical Research Letters* 49.20, e2022GL098822. ISSN: 1944-8007. DOI: 10.1029/2022GL098822.

Clement, Amy C. et al. (1996). "An Ocean Dynamical Thermostat". *Journal of Climate* 9.9, pp. 2190–2196. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/1520-0442(1996)009<2190:AODT>2.0.CO;2.

Cohen, Judah et al. (2012). "Arctic warming, increasing snow cover and widespread boreal winter cooling". *Environmental Research Letters* 7.1, p. 014007. ISSN: 1748-9326. DOI: 10.1088/1748-9326/7/1/014007.

Cohen, Judah et al. (2014). "Recent Arctic amplification and extreme mid-latitude weather". *Nature Geoscience* 7.9, pp. 627–637. ISSN: 1752-0894. DOI: 10.1038/ngeo2234.

Cohen, Judah et al. (2020). "Divergent consensuses on Arctic amplification influence on midlatitude severe winter weather". *Nature Climate Change* 10.1, pp. 20–29. ISSN: 1758-6798. DOI: 10.1038/s41558-019-0662-y.

Collins, Matthew (2002). "Climate predictability on interannual to decadal time scales: the initial value problem". *Climate Dynamics* 19.8, pp. 671–692. ISSN: 1432-0894. DOI: 10.1007/s00382-002-0254-8.

Collins, Matthew and The CMIP Modelling Groups (2005). "El Niño- or La Niña-like climate change?" *Climate Dynamics* 24.1, pp. 89–104. ISSN: 0930-7575, 1432-0894. DOI: 10.1007/s00382-004-0478-x.

Collins, Matthew et al. (2010). "The impact of global warming on the tropical Pacific Ocean and El Niño". *Nature Geoscience* 3.6, pp. 391–397. ISSN: 1752-0908. DOI: 10.1038/ngeo868.

Collins, Matthew et al. (2024). "Emerging signals of climate change from the equator to the poles: new insights into a warming world". *Frontiers in Science* 2. ISSN: 2813-6330. DOI: 10.3389/fsci.2024.1340323.

Coumou, Dim, Jascha Lehmann, and Johanna Beckmann (2015). "The weakening summer circulation in the Northern Hemisphere mid-latitudes". *Science* 348.6232, pp. 324–327. ISSN: 10959203. DOI: 10.1126/SCIENCE.1261768/SUPPL_FILE/PAPV2.PDF.

D'Agostino, Roberta and Piero Lionello (2017). "Evidence of global warming impact on the evolution of the Hadley Circulation in ECMWF centennial reanalyses". *Climate Dynamics* 48.9, pp. 3047–3060. ISSN: 1432-0894. DOI: 10.1007/s00382-016-3250-0.

Danabasoglu, G. et al. (2020). "The Community Earth System Model Version 2 (CESM2)". *Journal of Advances in Modeling Earth Systems* 12.2, e2019MS001916. ISSN: 1942-2466. DOI: 10.1029/2019MS001916.

Davy, Richard, Linling Chen, and Edward Hanna (2018). "Arctic amplification metrics". *International Journal of Climatology* 38.12, pp. 4384–4394. ISSN: 1097-0088. DOI: 10.1002/joc.5675.

Dee, D. P. et al. (2011). "The ERA-Interim reanalysis: configuration and performance of the data assimilation system". *Quarterly Journal of the Royal Meteorological Society* 137.656, pp. 553–597. ISSN: 00359009. DOI: 10.1002/qj.828.

Deser, Clara, Ruixia Guo, and Flavio Lehner (2017a). "The relative contributions of tropical Pacific sea surface temperatures and atmospheric internal variability to the recent global warming hiatus". *Geophysical Research Letters* 44.15, pp. 7945–7954. ISSN: 1944-8007. DOI: 10.1002/2017GL074273.

Deser, Clara, Robert A. Tomas, and Lantao Sun (2015). "The Role of Ocean–Atmosphere Coupling in the Zonal-Mean Atmospheric Response to Arctic Sea Ice Loss". *Journal of Climate* 28.6, pp. 2168–2186. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-14-00325.1.

Deser, Clara et al. (2010). "The Seasonal Atmospheric Response to Projected Arctic Sea Ice Loss in the Late Twenty-First Century". *Journal of Climate* 23.2, pp. 333–351. ISSN: 0894-8755. DOI: 10.1175/2009JCLI3053.1.

Deser, Clara et al. (2012). "Uncertainty in climate change projections: the role of internal variability". *Climate Dynamics* 38.3, pp. 527–546. ISSN: 1432-0894. DOI: 10.1007/s00382-010-0977-x.

Deser, Clara et al. (2016). "Does ocean coupling matter for the northern extratropical response to projected Arctic sea ice loss?" *Geophysical Research Letters* 43.5, pp. 2149–2157. ISSN: 1944-8007. DOI: 10.1002/2016GL067792.

Deser, Clara et al. (2017b). "The Northern Hemisphere Extratropical Atmospheric Circulation Response to ENSO: How Well Do We Know It and How Do We

Evaluate Models Accordingly?" *Journal of Climate* 30.13, pp. 5059–5082. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0844.1.

Doblas-Reyes, F. J. et al. (2006). "Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts". *Geophysical Research Letters* 33.7. ISSN: 1944-8007. DOI: 10.1029/2005GL025061.

Doblas-Reyes, Francisco J. et al. (2013). "Seasonal climate predictability and forecasting: status and prospects". *WIREs Climate Change* 4.4, pp. 245–268. ISSN: 1757-7799. DOI: 10.1002/wcc.217.

Domeisen, Daniela I.V., Chaim I. Garfinkel, and Amy H. Butler (2019). "The Teleconnection of El Niño Southern Oscillation to the Stratosphere". *Reviews of Geophysics* 57.1, pp. 5–47. ISSN: 1944-9208. DOI: 10.1029/2018RG000596.

Dong, Buwen and Rowan T. Sutton (2021). "Recent Trends in Summer Atmospheric Circulation in the North Atlantic/European Region: Is There a Role for Anthropogenic Aerosols?" *Journal of Climate* 34.16, pp. 6777–6795. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-20-0665.1.

Drouard, Marie and Christophe Cassou (2019). "A Modeling- and Process-Oriented Study to Investigate the Projected Change of ENSO-Forced Wintertime Teleconnectivity in a Warmer World". *Journal of Climate* 32.23, pp. 8047–8068. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-18-0803.1.

Dunstone, Nick et al. (2016). "Skilful predictions of the winter North Atlantic Oscillation one year ahead". *Nature Geoscience* 9.11, pp. 809–814. ISSN: 1752-0908. DOI: 10.1038/ngeo2824.

Dunstone, Nick et al. (2023). "Skilful predictions of the Summer North Atlantic Oscillation". *Communications Earth & Environment* 4.1, pp. 1–11. ISSN: 2662-4435. DOI: 10.1038/s43247-023-01063-2.

Eade, R. et al. (2022). "Quantifying the rarity of extreme multi-decadal trends: how unusual was the late twentieth century trend in the North Atlantic Oscillation?" *Climate Dynamics* 58.5, pp. 1555–1568. ISSN: 14320894. DOI: 10.1007/S00382-021-05978-4/FIGURES/3.

Eade, Rosie et al. (2014). "Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?" *Geophysical Research Letters* 41.15, pp. 5620–5628. ISSN: 1944-8007. DOI: 10.1002/2014GL061146.

England, Mark R. et al. (2020). "Tropical climate responses to projected Arctic and Antarctic sea-ice loss". *Nature Geoscience* 13.4, pp. 275–281. ISSN: 1752-0908. DOI: 10.1038/s41561-020-0546-9.

England, Matthew H. et al. (2014). "Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus". *Nature Climate Change* 4.3, pp. 222–227. ISSN: 1758-6798. DOI: 10.1038/nclimate2106.

Eyring, Veronika et al. (2016). "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization". *Geoscientific Model Development* 9.5, pp. 1937–1958. ISSN: 1991-959X. DOI: 10.5194/gmd-9-1937-2016.

Flannaghan, T. J. et al. (2014). "Tropical temperature trends in Atmospheric General Circulation Model simulations and the impact of uncertainties in observed SSTs".

Journal of Geophysical Research: Atmospheres 119.23, pp. 13, 327–13, 337. ISSN: 2169-8996. DOI: 10.1002/2014JD022365.

Francis, Jennifer A. and Stephen J. Vavrus (2012). "Evidence linking Arctic amplification to extreme weather in mid-latitudes". *Geophysical Research Letters* 39.6, n/a–n/a. ISSN: 00948276. DOI: 10.1029/2012GL051000.

Fu, Qiang and Celeste M. Johanson (2005). "Satellite-derived vertical dependence of tropical tropospheric temperature trends". *Geophysical Research Letters* 32.10. ISSN: 1944-8007. DOI: 10.1029/2004GL022266.

Fu, Qiang, Syukuro Manabe, and Celeste M. Johanson (2011). "On the warming in the tropical upper troposphere: Models versus observations". *Geophysical Research Letters* 38.15. ISSN: 1944-8007. DOI: 10.1029/2011GL048101.

Fueglistaler, S., C. Radley, and I. M. Held (2015). "The distribution of precipitation and the spread in tropical upper tropospheric temperature trends in CMIP5/AMIP simulations". *Geophysical Research Letters* 42.14, pp. 6000–6007. ISSN: 1944-8007. DOI: 10.1002/2015GL064966.

Galperin, Boris and Peter L. Read, eds. (2019). *Zonal Jets: Phenomenology, Genesis, and Physics*. Cambridge: Cambridge University Press. ISBN: 978-1-107-04388-6. DOI: 10.1017/9781107358225.

Garreaud, René D., Kyle Clem, and José Vicencio Veloso (2021). "The South Pacific Pressure Trend Dipole and the Southern Blob". *Journal of Climate* 34.18, pp. 7661–7676. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-20-0886.1.

Garreaud, René D. et al. (2017). "The 2010–2015 megadrought in central Chile: impacts on regional hydroclimate and vegetation". *Hydrology and Earth System Sciences* 21.12, pp. 6307–6327. ISSN: 1027-5606. DOI: 10.5194/hess-21-6307-2017.

Garreaud, René D. et al. (2020). "The Central Chile Mega Drought (2010–2018): A climate dynamics perspective". *International Journal of Climatology* 40.1, pp. 421–439. ISSN: 1097-0088. DOI: 10.1002/JOC.6219.

Gelaro, Ronald et al. (2017). "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)". *Journal of Climate* 30.14, pp. 5419–5454. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0758.1.

Gertler, Charles G. and Paul A. O'Gorman (2019). "Changing available energy for extratropical cyclones and associated convection in Northern Hemisphere summer". *Proceedings of the National Academy of Sciences of the United States of America* 116.10, pp. 4105–4110. ISSN: 10916490. DOI: 10.1073/pnas.1812312116.

Gillett, Nathan P. and David W. J. Thompson (2003). "Simulation of Recent Southern Hemisphere Climate Change". *Science* 302.5643, pp. 273–275. ISSN: 0036-8075. DOI: 10.1126/science.1087440.

Gillett, Nathan P. et al. (2008). "Attribution of polar warming to human influence". *Nature Geoscience* 1.11, pp. 750–754. ISSN: 1752-0908. DOI: 10.1038/ngeo338.

Graversen, R. G. (2006). "Do Changes in the Midlatitude Circulation Have Any Impact on the Arctic Surface Air Temperature Trend?" *Journal of Climate* 19.20, pp. 5422–5438. ISSN: 1520-0442. DOI: 10.1175/JCLI3906.1.

Graversen, Rune G. et al. (2008). "Vertical structure of recent Arctic warming". *Nature* 451.7174, pp. 53–56. ISSN: 0028-0836. DOI: 10.1038/nature06502.

Gruber, A. (1972). "Fluctuations in the Position of the ITCZ in the Atlantic and Pacific Oceans". *Journal of the Atmospheric Sciences* 29.1, pp. 193–197. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1972)029<0193:FITPOT>2.0.CO;2.

Gulev, S.K. et al. (2021). "Changing State of the Climate System". *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson- Delmotte et al. Cambridge University Press.

Hahn, L. C. et al. (2021). "Contributions to Polar Amplification in CMIP5 and CMIP6 Models". *Frontiers in Earth Science* 9. ISSN: 2296-6463. DOI: 10.3389/feart.2021.710036.

Hanna, Edward et al. (2015). "Recent seasonal asymmetric changes in the NAO (a marked summer decline and increased winter variability) and associated changes in the AO and Greenland Blocking Index: RECENT SEASONAL ASYMMETRIC CHANGES IN THE NAO". *International Journal of Climatology* 35.9, pp. 2540–2554. ISSN: 08998418. DOI: 10.1002/joc.4157.

Hawkins, Ed and Rowan Sutton (2009). "The Potential to Narrow Uncertainty in Regional Climate Predictions". *Bulletin of the American Meteorological Society* 90.8, pp. 1095–1108. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/2009BAMS2607.1.

Hay, Stephanie, James A. Screen, and Jennifer L. Catto (2024). "Steady but model dependent Arctic amplification of the forced temperature response in 21st century CMIP6 projections". *Environmental Research: Climate* 3.3, p. 031003. ISSN: 2752-5295. DOI: 10.1088/2752-5295/ad4201.

He, Shengping et al. (2020). "Eurasian Cooling Linked to the Vertical Distribution of Arctic Warming". *Geophysical Research Letters* 47.10, e2020GL087212. ISSN: 1944-8007. DOI: 10.1029/2020GL087212.

Heede, Ulla K. and Alexey V. Fedorov (2021). "Eastern equatorial Pacific warming delayed by aerosols and thermostat response to CO2 increase". *Nature Climate Change* 11.8, pp. 696–703. ISSN: 1758-6798. DOI: 10.1038/s41558-021-01101-x.

— (2023). "Colder Eastern Equatorial Pacific and Stronger Walker Circulation in the Early 21st Century: Separating the Forced Response to Global Warming From Natural Variability". *Geophysical Research Letters* 50.3, e2022GL101020. ISSN: 1944-8007. DOI: 10.1029/2022GL101020.

Heede, Ulla K., Alexey V. Fedorov, and Natalie J. Burls (2020). "Time Scales and Mechanisms for the Tropical Pacific Response to Global Warming: A Tug of War between the Ocean Thermostat and Weaker Walker". *Journal of Climate* 33.14, pp. 6101–6118. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-19-0690.1.

Held, Isaac M. and Arthur Y. Hou (1980). "Nonlinear Axially Symmetric Circulations in a Nearly Inviscid Atmosphere". *Journal of the Atmospheric Sciences* 37.3, pp. 515–533. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1980)037<0515:NASCIA>2.0.CO;2.

Held, Isaac M. and Brian J. Soden (2006). "Robust responses of the hydrological cycle to global warming". *Journal of Climate* 19.21. ISSN: 08948755. DOI: 10.1175/JCLI3990.1.

Hersbach, Hans et al. (2020). "The ERA5 global reanalysis". *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. ISSN: 1477-870X. DOI: 10.1002/QJ.3803.

Hodson, Daniel L.R., Rowan T. Sutton, and Adam A. Scaife (2023). "Signal-to-noise and predictable modes of variability in winter seasonal forecasts". *Quarterly Journal of the Royal Meteorological Society* 149.755, pp. 2598–2616. ISSN: 1477-870X. DOI: 10.1002/qj.4522.

Honda, Meiji, Jun Inoue, and Shozo Yamane (2009). "Influence of low Arctic sea-ice minima on anomalously cold Eurasian winters". *Geophysical Research Letters* 36.8. ISSN: 1944-8007. DOI: 10.1029/2008GL037079.

Hua, Wenjian et al. (2016). "Possible causes of the Central Equatorial African long-term drought". *Environmental Research Letters* 11.12, p. 124002. ISSN: 1748-9326. DOI: 10.1088/1748-9326/11/12/124002.

Huang, Boyin et al. (2017). "Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons". *Journal of Climate* 30.20, pp. 8179–8205. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0836.1.

Huffman, George J. et al. (2023). "The New Version 3.2 Global Precipitation Climatology Project (GPCP) Monthly and Daily Precipitation Products". *Journal of Climate* 36.21, pp. 7635–7655. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-23-0123.1.

Ineson, S. and A. A. Scaife (2009). "The role of the stratosphere in the European climate response to El Niño". *Nature Geoscience* 2.1, pp. 32–36. ISSN: 1752-0908. DOI: 10.1038/ngeo381.

"Polar Regions" (2022). *The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change*. Ed. by Intergovernmental Panel on Climate Change (IPCC). Cambridge: Cambridge University Press, pp. 203–320. ISBN: 978-1-00-915796-4. DOI: 10.1017/9781009157964.005.

"Changing State of the Climate System" (2023). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by Intergovernmental Panel on Climate Change (IPCC). Cambridge: Cambridge University Press, pp. 287–422. ISBN: 978-1-00-915788-9. DOI: 10.1017/9781009157896.004.

Jain, Shipra et al. (2023). "Importance of internal variability for climate model assessment". *npj Climate and Atmospheric Science* 6.1, pp. 1–7. ISSN: 2397-3722. DOI: 10.1038/s41612-023-00389-0.

Japan Meteorological Agency (2022). *JRA-3Q Usage: Local issues with sea ice parameters*. Japan Meteorological Agency.

Jiménez-Esteve, Bernat and Daniela I. V. Domeisen (2020). "Nonlinearity in the tropospheric pathway of ENSO to the North Atlantic". *Weather and Climate Dynamics* 1.1, pp. 225–245. DOI: 10.5194/wcd-1-225-2020.

Kamae, Youichi et al. (2015). "Recent slowdown of tropical upper tropospheric warming associated with Pacific climate variability". *Geophysical Research Letters* 42.8, pp. 2995–3003. ISSN: 1944-8007. DOI: 10.1002/2015GL063608.

Kang, Joonsuk M. et al. (2024). "Revisiting the reanalysis-model discrepancy in Southern Hemisphere winter storm track trends". *npj Climate and Atmospheric Science* 7.1, pp. 1–10. ISSN: 2397-3722. DOI: 10.1038/s41612-024-00801-3.

Karnauskas, Kristopher B., Ulla K. Heede, and Lei Zhang (2023). "The Impact of Eastern Pacific Warming on Future North Atlantic Tropical Cyclogenesis". *Geophysical Research Letters* 50.17, e2023GL105551. ISSN: 1944-8007. DOI: 10.1029/2023GL105551.

Karnauskas, Kristopher B. et al. (2009). "Observed Strengthening of the Zonal Sea Surface Temperature Gradient across the Equatorial Pacific Ocean". *Journal of Climate* 22.16, pp. 4316–4321. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/2009JCLI2936.1.

Kay, G et al. (2022). "Assessing the chance of unprecedented dry conditions over North Brazil during El Niño events". *Environmental Research Letters* 17.6, p. 064016. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac6df9.

Kay, Gillian et al. (2020). "Current likelihood and dynamics of hot summers in the UK". *Environmental Research Letters* 15.9, p. 094099. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ABAB32.

Keel, Tom et al. (2024). "Exploring Uncertainty of Trends in the North Pacific Jet Position". *Geophysical Research Letters* 51.16, e2024GL109500. ISSN: 1944-8007. DOI: 10.1029/2024GL109500.

Keil, P. et al. (2021). "Variations of Tropical Lapse Rates in Climate Models and Their Implications for Upper-Tropospheric Warming". *Journal of Climate* 34.24, pp. 9747–9761. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-21-0196.1.

Keil, Paul (2023). "Tropical tropospheric temperatures, warming and involved mechanisms", p. 16405344. DOI: 10.17617/2.3509472.

Kelder, T. et al. (2020). "Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes". *npj Climate and Atmospheric Science* 3.1, pp. 1–13. ISSN: 23973722. DOI: 10.1038/s41612-020-00149-4.

Khatri, Hemant et al. (2022). "Fast and Slow Subpolar Ocean Responses to the North Atlantic Oscillation: Thermal and Dynamical Changes". *Geophysical Research Letters* 49.24, e2022GL101480. ISSN: 1944-8007. DOI: 10.1029/2022GL101480.

Knutson, Thomas R. and Syukuro Manabe (1995). "Time-Mean Response over the Tropical Pacific to Increased C02 in a Coupled Ocean-Atmosphere Model". *Journal of Climate* 8.9, pp. 2181–2199. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/1520-0442(1995)008<2181:TMROTT>2.0.CO;2.

Kobayashi, Shinya et al. (2015). "The JRA-55 Reanalysis: General Specifications and Basic Characteristics". *Journal of the Meteorological Society of Japan. Ser. II* 93.1, pp. 5–48. ISSN: 0026-1165, 2186-9057. DOI: 10.2151/jmsj.2015-001.

Kosaka, Yu and Shang-Ping Xie (2013). "Recent global-warming hiatus tied to equatorial Pacific surface cooling". *Nature* 501.7467, pp. 403–407. ISSN: 1476-4687. DOI: 10.1038/nature12534.

— (2016). "The tropical Pacific as a key pacemaker of the variable rates of global warming". *Nature Geoscience* 9.9, pp. 669–673. ISSN: 1752-0908. DOI: 10.1038/ngeo2770.

Kosaka, Yuki et al. (2024). "The JRA-3Q Reanalysis". *Journal of the Meteorological Society of Japan. Ser. II* 102.1, pp. 49–109. DOI: 10.2151/jmsj.2024-004.

Kosovelj, Katarina and Žiga Zaplotnik (2023). "Indices of Pacific Walker Circulation Strength". *Atmosphere* 14.2, p. 397. ISSN: 2073-4433. DOI: 10.3390/atmos14020397.

Kozubek, Michal, Peter Krizan, and Jan Lastovicka (2020). "Homogeneity of the Temperature Data Series from ERA5 and MERRA2 and Temperature Trends". *Atmosphere 2020, Vol. 11, Page 235* 11.3, p. 235. ISSN: 2073-4433. DOI: 10.3390/ATMOS11030235.

Kushnir, Yochanan et al. (2019). "Towards operational predictions of the near-term climate". *Nature Climate Change* 9.2, pp. 94–101. ISSN: 1758-678X. DOI: 10.1038/s41558-018-0359-7.

L'Heureux, Michelle L., Michael K. Tippett, and Wanqiu Wang (2022). "Prediction Challenges From Errors in Tropical Pacific Sea Surface Temperature Trends". *Frontiers in Climate* 4. ISSN: 2624-9553. DOI: 10.3389/fclim.2022.837483.

Lachmy, Orli and Yohai Kaspi (2020). "The Role of Diabatic Heating in Ferrel Cell Dynamics". *Geophysical Research Letters* 47.23, e2020GL090619. ISSN: 1944-8007. DOI: 10.1029/2020GL090619.

Lee, Sukyoung (2012). "Testing of the Tropically Excited Arctic Warming Mechanism (TEAM) with Traditional El Niño and La Niña". *Journal of Climate* 25.12, pp. 4015–4022. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-12-00055.1.

— (2014). "A theory for polar amplification from a general circulation perspective". *Asia-Pacific Journal of Atmospheric Sciences* 50.1, pp. 31–43. ISSN: 1976-7633. DOI: 10.1007/s13143-014-0024-7.

Lee, Sukyoung et al. (2022). "On the future zonal contrasts of equatorial Pacific climate: Perspectives from Observations, Simulations, and Theories". *npj Climate and Atmospheric Science* 5.1, pp. 1–15. ISSN: 2397-3722. DOI: 10.1038/s41612-022-00301-2.

Li, Ronald K. K. et al. (2020). "Effect of the North Pacific Tropospheric Waveguide on the Fidelity of Model El Niño Teleconnections". *Journal of Climate* 33.12, pp. 5223–5237. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-19-0156.1.

Li, Tim, Lei Zhang, and Hiroyuki Murakami (2015). "Strengthening of the Walker circulation under globalwarming in an aqua-planet general circulation model simulation". *Advances in Atmospheric Sciences* 32.11, pp. 1473–1480. ISSN: 1861-9533. DOI: 10.1007/s00376-015-5033-7.

Li, Zhe and Qinghua Ding (2024). "A global poleward shift of atmospheric rivers". *Science Advances* 10.41, eadq0604. DOI: 10.1126/sciadv.adq0604.

Ling, Jian and Chidong Zhang (2013). "Diabatic Heating Profiles in Recent Global Reanalyses". *Journal of Climate* 26.10, pp. 3307–3325. ISSN: 0894-8755. DOI: 10.1175/JCLI-D-12-00384.1.

Lionello, Piero et al. (2024). "The Hadley circulation in a changing climate". *Annals of the New York Academy of Sciences* 1534.1, pp. 69–93. ISSN: 1749-6632. DOI: 10.1111/nyas.15114.

Lorenz, E. N. (1982). "Atmospheric predictability experiments with a large numerical model". *Tellus A: Dynamic Meteorology and Oceanography* 34.6. ISSN: 1600-0870. DOI: 10.3402/tellusa.v34i6.10836.

Lu, Jian, Gang Chen, and Dargan M. W. Frierson (2008). "Response of the Zonal Mean Atmospheric Circulation to El Niño versus Global Warming". *Journal of Climate* 21.22, pp. 5835–5851. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/2008JCLI2200.1.

Mahajan, Salil, Rong Zhang, and Thomas L. Delworth (2011). "Impact of the Atlantic Meridional Overturning Circulation (AMOC) on Arctic Surface Air Temperature and Sea Ice Variability". *Journal of Climate* 24.24, pp. 6573–6581. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/2011JCLI4002.1.

Maher, Nicola et al. (2019). "The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability". *Journal of Advances in Modeling Earth Systems* 11.7, pp. 2050–2069. ISSN: 1942-2466. DOI: 10.1029/2019MS001639.

Manabe, Syukuro and Richard T. Wetherald (1967). "Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity". *Journal of the Atmospheric Sciences* 24.3, pp. 241–259. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.

— (1975). "The Effects of Doubling the CO $_2$ Concentration on the climate of a General Circulation Model". *Journal of the Atmospheric Sciences* 32.1, pp. 3–15. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2.

— (1980). "On the Distribution of Climate Change Resulting from an Increase in CO2 Content of the Atmosphere". *Journal of the Atmospheric Sciences* 37.1, pp. 99–118. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1980)037<0099:OTDOCC>2.0.CO;2.

Manney, Gloria L. and Michaela I. Hegglin (2018). "Seasonal and Regional Variations of Long-Term Changes in Upper-Tropospheric Jets from Reanalyses". *Journal of Climate* 31.1, pp. 423–448. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-17-0303.1.

McCusker, Kelly E., John C. Fyfe, and Michael Sigmond (2016). "Twenty-five winters of unexpected Eurasian cooling unlikely due to Arctic sea-ice loss". *Nature Geoscience* 9.11, pp. 838–842. ISSN: 1752-0894. DOI: 10.1038/ngeo2820.

Medeiros, Brian et al. (2011). "Arctic Inversion Strength in Climate Models". *Journal of Climate* 24.17, pp. 4733–4740. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/2011JCLI3968.1.

Megann, A. et al. (2014). "GO5.0: the joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications". *Geoscientific Model Development* 7.3, pp. 1069–1092. ISSN: 1991-959X. DOI: 10.5194/gmd-7-1069-2014.

Merlis, Timothy M. (2015). "Direct weakening of tropical circulations from masked CO2 radiative forcing". *Proceedings of the National Academy of Sciences* 112.43, pp. 13167–13171. DOI: 10.1073/pnas.1508268112.

Mitas, Christos M. and Amy Clement (2005). "Has the Hadley cell been strengthening in recent decades?" *Geophysical Research Letters* 32.3. ISSN: 1944-8007. DOI: 10.1029/2004GL021765.

— (2006). "Recent behavior of the Hadley cell and tropical thermodynamics in climate models and reanalyses". *Geophysical Research Letters* 33.1. ISSN: 1944-8007. DOI: 10.1029/2005GL024406.

Mitchell, D. M. et al. (2013). "Revisiting the controversial issue of tropical tropospheric temperature trends". *Geophysical Research Letters* 40.11, pp. 2801–2806. ISSN: 1944-8007. DOI: 10.1002/grl.50465.

Mitchell, Dann M. et al. (2020). "The vertical profile of recent tropical temperature trends: Persistent model biases in the context of internal variability". *Environmental Research Letters* 15.10, 1040b4. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab9af7.

Molteni, Franco and Anca Brookshaw (2023). "Early- and late-winter ENSO teleconnections to the Euro-Atlantic region in state-of-the-art seasonal forecasting systems". *Climate Dynamics* 61.5, pp. 2673–2692. ISSN: 1432-0894. DOI: 10.1007/s00382-023-06698-7.

Molteni, Franco et al. (2020). "Boreal-winter teleconnections with tropical Indo-Pacific rainfall in HighResMIP historical simulations from the PRIMAVERA project". *Climate Dynamics* 55.7, pp. 1843–1873. ISSN: 1432-0894. DOI: 10.1007/s00382-020-05358-4.

Mori, Masato et al. (2024). "Northern Hemisphere winter atmospheric teleconnections are intensified by extratropical ocean-atmosphere coupling". *Communications Earth & Environment* 5.1, pp. 1–16. ISSN: 2662-4435. DOI: 10.1038/s43247-024-01282-1.

Nguyen, H. et al. (2013). "The Hadley Circulation in Reanalyses: Climatology, Variability, and Change". *Journal of Climate* 26.10, pp. 3357–3376. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-12-00224.1.

Nigam, Sumant, Chul Chung, and Eric DeWeaver (2000). "ENSO Diabatic Heating in ECMWF and NCEP–NCAR Reanalyses, and NCAR CCM3 Simulation". *Journal of Climate* 13.17, pp. 3152–3171. ISSN: 0894-8755. DOI: 10.1175/1520-0442(2000)013.

O'Gorman, Paul A. (2010). "Understanding the varied response of the extratropical storm tracks to climate change". *Proceedings of the National Academy of Sciences of the United States of America* 107.45, pp. 19176–19180. ISSN: 10916490. DOI: 10.1073/PNAS.1011547107/SUPPL_FILE/PNAS.1011547107_SI.PDF.

O'Gorman, Paul A. and Martin S. Singh (2013). "Vertical structure of warming consistent with an upward shift in the middle and upper troposphere". *Geophysical Research Letters* 40.9, pp. 1838–1842. ISSN: 1944-8007. DOI: 10.1002/grl.50328.

Ogawa, Fumiaki et al. (2018). "Evaluating Impacts of Recent Arctic Sea Ice Loss on the Northern Hemisphere Winter Climate Change". *Geophysical Research Letters* 45.7, pp. 3255–3263. ISSN: 1944-8007. DOI: 10.1002/2017GL076502.

Outten, Stephen et al. (2023). "Reconciling conflicting evidence for the cause of the observed early 21st century Eurasian cooling". *Weather and Climate Dynamics* 4.1, pp. 95–114. DOI: 10.5194/wcd-4-95-2023.

Palmer, T. N. (1981). "Diagnostic Study of a Wavenumber-2 Stratospheric Sudden Warming in a Transformed Eulerian-Mean Formalism". *Journal of the Atmospheric Sciences* 38.4, pp. 844–855. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1981)038<0844:DSOAWS>2.0.CO;2.

Palmer, T. N. and A. Weisheimer (2018). "A Simple Pedagogical Model Linking Initial-Value Reliability with Trustworthiness in the Forced Climate Response".

Bulletin of the American Meteorological Society 99.3, pp. 605–614. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-16-0240.1.

Palmer, T. N. et al. (2008). "Toward Seamless Prediction: Calibration of Climate Change Projections Using Seasonal Forecasts". *Bulletin of the American Meteorological Society* 89.4, pp. 459–470. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-89-4-459.

Patterson, Matthew et al. (2022). "The strong role of external forcing in seasonal forecasts of European summer temperature". *Environmental Research Letters* 17.10, p. 104033. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac9243.

Peixoto, Jose P. and Abraham H. Oort (1992). *Physics of Climate*. Melville, NY: American Institute of Physics.

Persson, P. Ola G. (2002). "Measurements near the Atmospheric Surface Flux Group tower at SHEBA: Near-surface conditions and surface energy budget". *Journal of Geophysical Research* 107 (C10), p. 8045. ISSN: 0148-0227. DOI: 10.1029/2000JC000705.

Philander, S. G. H. et al. (1996). "Why the ITCZ Is Mostly North of the Equator". *Journal of Climate* 9.12, pp. 2958–2972. ISSN: 0894-8755. DOI: 10.1175/1520-0442(1996)009<2958:WTIIMN>2.0.CO;2.

Pithan, Felix and Thorsten Mauritsen (2014). "Arctic amplification dominated by temperature feedbacks in contemporary climate models". *Nature Geoscience* 7.3, pp. 181–184. ISSN: 1752-0894. DOI: 10.1038/ngeo2071.

Previdi, Michael, Karen L. Smith, and Lorenzo M. Polvani (2021). "Arctic amplification of climate change: a review of underlying mechanisms". *Environmental Research Letters* 16.9, p. 093003. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac1c29.

Qiu, Weiteng et al. (2024). "Tropical Pacific trends explain the discrepancy between observed and modelled rainfall change over the Americas". *npj Climate and Atmospheric Science* 7.1, pp. 1–7. ISSN: 2397-3722. DOI: 10.1038/s41612-024-00750-x.

Quan, Xiao-Wei, Henry F. Diaz, and Martin P. Hoerling (2004). "Change in the Tropical Hadley Cell Since 1950". *The Hadley Circulation: Present, Past and Future*. Ed. by Henry F. Diaz and Raymond S. Bradley. Dordrecht: Springer Netherlands, pp. 85–120. ISBN: 978-1-4020-2944-8. DOI: 10.1007/978-1-4020-2944-8_4.

Rae, J. G. L. et al. (2015). "Development of the Global Sea Ice 6.0 CICE configuration for the Met Office Global Coupled model". *Geoscientific Model Development* 8.7, pp. 2221–2230. ISSN: 1991-959X. DOI: 10.5194/gmd-8-2221-2015.

Randall, David A. and Kerry Emanuel (2024). "The Weather–Climate Schism". *Bulletin of the American Meteorological Society* 105.1, E300–E305. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-23-0124.1.

Rantanen, Mika et al. (2022). "The Arctic has warmed nearly four times faster than the globe since 1979". *Communications Earth & Environment* 3.1, pp. 1–10. ISSN: 2662-4435. DOI: 10.1038/s43247-022-00498-3.

Rayner, N. A. (2003). "Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century". *Journal of Geophysical Research* 108 (D14), p. 4407. ISSN: 0148-0227. DOI: 10.1029/2002JD002670.

Rivière, Gwendal (2011). "A Dynamical Interpretation of the Poleward Shift of the Jet Streams in Global Warming Scenarios". *Journal of the Atmospheric Sciences* 68.6, pp. 1253–1272. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/2011JAS3641.1.

Rojas, M. (2013). "Sensitivity of Southern Hemisphere circulation to LGM and 4 × CO2 climates". *Geophysical Research Letters* 40.5, pp. 965–970. ISSN: 1944-8007. DOI: 10.1002/GRL.50195.

Rosenbloom, Nan, Isla Simpson, and Adam Phillips (2022). *CESM2 Pacific Pacemaker Ensemble*. DOI: 10.26024/GTRS-TF57.

Rowell, David P. et al. (2015). "Reconciling Past and Future Rainfall Trends over East Africa". *Journal of Climate* 28.24, pp. 9768–9788. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-15-0140.1.

Saffin, Leo et al. (2024). "Large Uncertainties When Diagnosing the "Eddy Feedback Parameter" and Its Role in the Signal-To-Noise Paradox". *Geophysical Research Letters* 51.11, e2024GL108861. ISSN: 1944-8007. DOI: 10.1029/2024GL108861.

Sandeep, S. et al. (2014). "Pacific Walker Circulation variability in coupled and uncoupled climate models". *Climate Dynamics* 43.1, pp. 103–117. ISSN: 1432-0894. DOI: 10.1007/s00382-014-2135-3.

Santer, B. D. et al. (2000). "Statistical significance of trends and trend differences in layer-average atmospheric temperature time series". *Journal of Geophysical Research: Atmospheres* 105 (D6), pp. 7337–7356. ISSN: 2156-2202. DOI: 10.1029/1999JD901105.

Santer, B. D. et al. (2005). "Amplification of Surface Temperature Trends and Variability in the Tropical Atmosphere". *Science* 309.5740, pp. 1551–1556. DOI: 10.1126/science.1114867.

Santer, Benjamin D. et al. (2021). "Using Climate Model Simulations to Constrain Observations". *Journal of Climate* 34.15, pp. 6281–6301. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-20-0768.1.

Scaife, A. A. et al. (2014). "Skillful long-range prediction of European and North American winters". *Geophysical Research Letters* 41.7, pp. 2514–2519. ISSN: 0094-8276. DOI: 10.1002/2014GL059637.

Schneider, Tapio, Paul A. O'Gorman, and Xavier J. Levine (2010). "Water Vapor and the Dynamics of Climate Changes". *Reviews of Geophysics* 48.3. ISSN: 1944-9208. DOI: 10.1029/2009RG000302.

Schneider, Tapio and Adam H. Sobel, eds. (2021). *The Global Circulation of the Atmosphere*. Princeton University Press. ISBN: 978-0-691-23691-9. DOI: 10.2307/j.ctv1t1kg52.

Screen, J. A., T. J. Bracegirdle, and I. Simmonds (2018a). "Polar Climate Change as Manifest in Atmospheric Circulation". *Current Climate Change Reports* 4.4, pp. 383–395. ISSN: 2198-6061. DOI: 10.1007/s40641-018-0111-4.

Screen, J. A., C. Deser, and I. Simmonds (2012). "Local and remote controls on observed Arctic warming". *Geophysical Research Letters* 39.10. ISSN: 1944-8007. DOI: 10.1029/2012GL051598.

Screen, James A. and I. Simmonds (2010a). "The central role of diminishing sea ice in recent Arctic temperature amplification". *Nature* 464.7293. ISSN: 00280836. DOI: 10.1038/nature09051.

Screen, James A. and Ian Simmonds (2010b). "Increasing fall-winter energy loss from the Arctic Ocean and its role in Arctic temperature amplification". *Geophysical Research Letters* 37.16, p. 16707. ISSN: 1944-8007. DOI: 10.1029/2010GL044136.

Screen, James A. et al. (2014). "Atmospheric impacts of Arctic sea-ice loss, 1979–2009: separating forced change from atmospheric internal variability". *Climate Dynamics* 43.1, pp. 333–344. ISSN: 1432-0894. DOI: 10.1007/s00382-013-1830-9.

Screen, James A. et al. (2018b). "Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models". *Nature Geoscience* 11.3, pp. 155–163. ISSN: 1752-0908. DOI: 10.1038/s41561-018-0059-y.

Screen, James A. et al. (2022). "Net Equatorward Shift of the Jet Streams When the Contribution From Sea-Ice Loss Is Constrained by Observed Eddy Feedback". *Geophysical Research Letters* 49.23, e2022GL100523. ISSN: 1944-8007. DOI: 10.1029/2022GL100523.

Seabrook, M. et al. (2023). "Opposite Impacts of Interannual and Decadal Pacific Variability in the Extratropics". *Geophysical Research Letters* 50.2, e2022GL101226. ISSN: 1944-8007. DOI: 10.1029/2022GL101226.

Seager, R. et al. (2005). "Mechanisms of ENSO-forcing of hemispherically symmetric precipitation variability". *Quarterly Journal of the Royal Meteorological Society* 131.608, pp. 1501–1527. ISSN: 1477-870X. DOI: 10.1256/qj.04.96.

Seager, Richard, Naomi Henderson, and Mark Cane (2022). "Persistent Discrepancies between Observed and Modeled Trends in the Tropical Pacific Ocean". *Journal of Climate* 35.14, pp. 4571–4584. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-21-0648.1.

Seager, Richard et al. (2003). "Mechanisms of hemispherically symmetric climate variability". *Journal of Climate* 16.18. ISSN: 08948755. DOI: 10.1175/1520-0442(2003)016<2960:MOHSCV>2.0.CO;2.

Seager, Richard et al. (2019). "Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases". *Nature Climate Change* 9.7, pp. 517–522. ISSN: 1758-6798. DOI: 10.1038/s41558-019-0505-x.

Senior, Catherine A. et al. (2016). "Idealized climate change simulations with a high-resolution physical model: HadGEM3-GC2". *Journal of Advances in Modeling Earth Systems* 8.2, pp. 813–830. ISSN: 19422466. DOI: 10.1002/2015MS000614.

Seo, Kyong-Hwan, Dargan M. W. Frierson, and Jun-Hyeok Son (2014). "A mechanism for future changes in Hadley circulation strength in CMIP5 climate change simulations". *Geophysical Research Letters* 41.14, pp. 5251–5258. ISSN: 1944-8007. DOI: 10.1002/2014GL060868.

Serreze, M. C. et al. (2009). "The emergence of surface-based Arctic amplification". *Cryosphere* 3.1, pp. 11–19. ISSN: 19940424. DOI: 10.5194/TC-3-11-2009.

Seviour, William J. M. et al. (2017). "Robustness of the Simulated Tropospheric Response to Ozone Depletion". *Journal of Climate* 30.7, pp. 2577–2585. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0817.1.

Shaw, Tiffany A. (2019). "Mechanisms of Future Predicted Changes in the Zonal Mean Mid-Latitude Circulation". *Current Climate Change Reports* 5.4, pp. 345–357. ISSN: 2198-6061. DOI: 10.1007/s40641-019-00145-8.

Shaw, Tiffany A. et al. (2024). "Regional climate change: consensus, discrepancies, and ways forward". *Frontiers in Climate* 6. ISSN: 2624-9553. DOI: 10.3389/fclim.2024.1391634.

Shepherd, Theodore G. (2014). "Atmospheric circulation as a source of uncertainty in climate change projections". *Nature Geoscience* 7.10, pp. 703–708. ISSN: 1752-0908. DOI: 10.1038/ngeo2253.

Shin, Chul-Su and Bohua Huang (2019). "A spurious warming trend in the NMME equatorial Pacific SST hindcasts". *Climate Dynamics* 53.12, pp. 7287–7303. ISSN: 1432-0894. DOI: 10.1007/s00382-017-3777-8.

Shrestha, Sisam and Brian J. Soden (2023). "Anthropogenic Weakening of the Atmospheric Circulation During the Satellite Era". *Geophysical Research Letters* 50.22, e2023GL104784. ISSN: 1944-8007. DOI: 10.1029/2023GL104784.

Siew, Peter Yu Feng et al. (2024). "Circulation responses to surface heating and implications for polar amplification". *Weather and Climate Dynamics* 5.3, pp. 985–996. DOI: 10.5194/wcd-5-985-2024.

Simmons, A. J. et al. (2004). "Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature". *Journal of Geophysical Research: Atmospheres* 109 (D24). ISSN: 2156-2202. DOI: 10.1029/2004JD005306.

Simmons, Adrian J. (2022). "Trends in the tropospheric general circulation from 1979 to 2022". *Weather and Climate Dynamics* 3.3, pp. 777–809. DOI: 10.5194/wcd-3-777-2022.

Singh, Martin S. and Paul A. O'Gorman (2012). "Upward Shift of the Atmospheric General Circulation under Global Warming: Theory and Simulations". *Journal of Climate* 25.23, pp. 8259–8276. ISSN: 0894-8755. DOI: 10.1175/JCLI-D-11-00699.1.

Sippel, Sebastian et al. (2024). "Early-twentieth-century cold bias in ocean surface temperature observations". *Nature* 635.8039, pp. 618–624. ISSN: 1476-4687. DOI: 10.1038/s41586-024-08230-1.

Slivinski, Laura C. et al. (2019). "Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system". *Quarterly Journal of the Royal Meteorological Society* 145.724, pp. 2876–2908. ISSN: 1477-870X. DOI: 10.1002/qj.3598.

Smith, D. M. et al. (2019a). "Robust skill of decadal climate predictions". *npj Climate and Atmospheric Science 2019 2:1* 2.1, pp. 1–10. ISSN: 2397-3722. DOI: 10.1038/s41612-019-0071-y.

Smith, D. M. et al. (2020). "North Atlantic climate far more predictable than models imply". *Nature* 583.7818, pp. 796–800. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2525-0.

Smith, D. M. et al. (2022). "Robust but weak winter atmospheric circulation response to future Arctic sea ice loss". *Nature Communications* 13.1, p. 727. ISSN: 2041-1723. DOI: 10.1038/s41467-022-28283-y.

Smith, Doug M., Rosie Eade, and Holger Pohlmann (2013). "A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction". *Climate*

Dynamics 41.11, pp. 3325–3338. ISSN: 1432-0894. DOI: 10.1007/s00382-013-1683-2.

Smith, Doug M. et al. (2015). "Earth's energy imbalance since 1960 in observations and CMIP5 models". *Geophysical Research Letters* 42.4, pp. 1205–1213. ISSN: 0094-8276. DOI: 10.1002/2014GL062669.

Smith, Doug M. et al. (2017). "Atmospheric Response to Arctic and Antarctic Sea Ice: The Importance of Ocean–Atmosphere Coupling and the Background State". *Journal of Climate* 30.12, pp. 4547–4565. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0564.1.

Smith, Doug M. et al. (2019b). "The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification". *Geoscientific Model Development* 12.3, pp. 1139–1164. ISSN: 1991-9603. DOI: 10.5194/gmd-12-1139-2019.

Sobel, Adam H. and Christopher S. Bretherton (2000). "Modeling Tropical Precipitation in a Single Column". *Journal of Climate* 13.24, pp. 4378–4392. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/1520-0442(2000)013<4378:MTPIAS>2.0.CO;2.

Sobel, Adam H., Johan Nilsson, and Lorenzo M. Polvani (2001). "The Weak Temperature Gradient Approximation and Balanced Tropical Moisture Waves". *Journal of the Atmospheric Sciences* 58.23, pp. 3650–3665. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(2001)058<3650:TWTGAA>2.0.CO;2.

Sobel, Adam H. et al. (2023). "Near-term tropical cyclone risk and coupled Earth system model biases". *Proceedings of the National Academy of Sciences* 120.33, e2209631120. DOI: 10.1073/pnas.2209631120.

Son, S.-W. et al. (2010). "Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment". *Journal of Geophysical Research: Atmospheres* 115 (D3). ISSN: 2156-2202. DOI: 10.1029/2010JD014271.

Spencer, Roy W. and John R. Christy (1990). "Precise Monitoring of Global Temperature Trends from Satellites". *Science* 247.4950, pp. 1558–1562. ISSN: 00368075. DOI: 10.1126/SCIENCE.247.4950.1558.

Steiner, A. K. et al. (2020). "Observed Temperature Changes in the Troposphere and Stratosphere from 1979 to 2018". *Journal of Climate* 33.19, pp. 8165–8194. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-19-0998.1.

Stevens, Bjorn et al. (2013). "Atmospheric component of the MPI-M Earth System Model: ECHAM6". *Journal of Advances in Modeling Earth Systems* 5.2, pp. 146–172. ISSN: 1942-2466. DOI: 10.1002/jame.20015.

Suárez-Gutiérrez, Laura et al. (2017). "Internal variability in simulated and observed tropical tropospheric temperature trends". *Geophysical Research Letters* 44.11, pp. 5709–5719. ISSN: 1944-8007. DOI: 10.1002/2017GL073798.

Sun, Xiaoting et al. (2022). "Enhanced jet stream waviness induced by suppressed tropical Pacific convection during boreal summer". *Nature Communications* 13.1, p. 1288. ISSN: 2041-1723. DOI: 10.1038/s41467-022-28911-7.

Sutton, R. and P.-P. Mathieu (2002). "Response of the atmosphere–ocean mixed-layer system to anomalous ocean heat-flux convergence". *Quarterly Journal of the Royal Meteorological Society* 128.582, pp. 1259–1275. ISSN: 1477-870X. DOI: 10.1256/003590002320373283.

Swart, Neil C. et al. (2015). "Comparing Trends in the Southern Annular Mode and Surface Westerly Jet". *Journal of Climate* 28.22, pp. 8840–8859. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-15-0334.1.

Swart, Neil C. et al. (2019). "The Canadian Earth System Model version 5 (CanESM5.0.3)". *Geoscientific Model Development* 12.11, pp. 4823–4873. ISSN: 1991-959X. DOI: 10.5194/gmd-12-4823-2019.

Sweeney, Aodhan J. et al. (2023). "Internal Variability Increased Arctic Amplification During 1980–2022". *Geophysical Research Letters* 50.24, e2023GL106060. ISSN: 1944-8007. DOI: 10.1029/2023GL106060.

Theil, H. (1950). "A rank-invariant method of linear and polynomial regression analysis". *Indagationes Mathematicae* 1.2. ISSN: 1872-6100.

Thompson, David W. J. and John M. Wallace (1998). "The Arctic oscillation signature in the wintertime geopotential height and temperature fields". *Geophysical Research Letters* 25.9, pp. 1297–1300. ISSN: 1944-8007. DOI: 10.1029/98GL00950.

Thompson, David W.J., Paulo Ceppi, and Ying Li (2019). "A Robust Constraint on the Temperature and Height of the Extratropical Tropopause". *Journal of Climate* 32.2, pp. 273–287. ISSN: 0894-8755. DOI: 10.1175/JCLI-D-18-0339.1.

Thompson, David W.J. and Susan Solomon (2002). "Interpretation of recent Southern Hemisphere climate change". *Science* 296.5569, pp. 895–899. ISSN: 00368075. DOI: 10.1126/SCIENCE.1069270/ASSET/7B2127B4-90B7-4C9F-82FC-41123EAA32BE/ASSETS/GRAPHIC/SE1620429004.JPEG.

Thompson, Vikki et al. (2017). "High risk of unprecedented UK rainfall in the current climate". *Nature Communications* 8.1, pp. 1–6. ISSN: 20411723. DOI: 10.1038/s41467-017-00275-3.

Thompson, Vikki et al. (2018). "Risk and dynamics of unprecedented hot months in South East China". *Climate Dynamics 2018 52:5* 52.5, pp. 2585–2596. ISSN: 1432-0894. DOI: 10.1007/S00382-018-4281-5.

Thorne, Peter W. (2008). "Arctic tropospheric warming amplification?" *Nature* 455.7210, E1–E2. ISSN: 0028-0836. DOI: 10.1038/nature07256.

Thorne, Peter W. et al. (2011). "Tropospheric temperature trends: history of an ongoing controversy". *Wiley Interdisciplinary Reviews: Climate Change* 2.1, pp. 66–88. ISSN: 1757-7799. DOI: 10.1002/WCC.80.

Tian, Tian et al. (2024). "Cooler Arctic surface temperatures simulated by climate models are closer to satellite-based data than the ERA5 reanalysis". *Communications Earth & Environment* 5.1, pp. 1–6. ISSN: 2662-4435. DOI: 10.1038/s43247-024-01276-z.

Tietsche, S. et al. (2014). "Seasonal to interannual Arctic sea ice predictability in current global climate models". *Geophysical Research Letters* 41.3, pp. 1035–1043. ISSN: 1944-8007. DOI: 10.1002/2013GL058755.

Toda, Masaki et al. (2024). "Walker circulation strengthening driven by sea surface temperature changes outside the tropics". *Nature Geoscience* 17.9, pp. 858–865. ISSN: 1752-0908. DOI: 10.1038/s41561-024-01510-5.

Tomas, Robert A., Clara Deser, and Lantao Sun (2016). "The Role of Ocean Heat Transport in the Global Climate Response to Projected Arctic Sea Ice Loss".

Journal of Climate 29.19, pp. 6841–6859. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-15-0651.1.

Trenberth, Kevin E. (1997). "The Definition of El Niño". *Bulletin of the American Meteorological Society* 78.12, pp. 2771–2778. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2.

Trenberth, Kevin E. and John T. Fasullo (2013). "An apparent hiatus in global warming?" *Earth's Future* 1.1, pp. 19–32. ISSN: 2328-4277. DOI: 10.1002/2013ef000165.

Trenberth, Kevin E. and James W. Hurrell (1994). "Decadal atmosphere-ocean variations in the Pacific". *Climate Dynamics* 9.6, pp. 303–319. ISSN: 1432-0894. DOI: 10.1007/BF00204745.

Trenberth, Kevin E. and Dennis J. Shea (2006). "Atlantic hurricanes and natural variability in 2005". *Geophysical Research Letters* 33.12. ISSN: 1944-8007. DOI: 10.1029/2006GL026894.

Vallis, Geoffrey K. (2017). *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*. 2nd ed. Cambridge: Cambridge University Press. ISBN: 978-1-107-06550-5. DOI: 10.1017/9781107588417.

Vallis, Geoffrey K. et al. (2015). "Response of the large-scale structure of the atmosphere to global warming". *Quarterly Journal of the Royal Meteorological Society* 141.690, pp. 1479–1501. ISSN: 1477-870X. DOI: 10.1002/qj.2456.

Vecchi, Gabriel A. and Brian J. Soden (2007). "Global warming and the weakening of the tropical circulation". *Journal of Climate* 20.17, pp. 4316–4340. ISSN: 08948755. DOI: 10.1175/JCLI4258.1.

Vecchi, Gabriel A. et al. (2006). "Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing". *Nature* 441.1, pp. 73–76. ISSN: 14764687. DOI: 10.1038/nature04744.

Voigt, Aiko and Tiffany A. Shaw (2016). "Impact of Regional Atmospheric Cloud Radiative Changes on Shifts of the Extratropical Jet Stream in Response to Global Warming". *Journal of Climate* 29.23, pp. 8399–8421. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0140.1.

Wallace, John M., Yuan Zhang, and Louis Bajuk (1996). "Interpretation of Interdecadal Trends in Northern Hemisphere Surface Air Temperature". *Journal of Climate* 9.2, pp. 249–259. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/1520-0442(1996)009<0249:IOITIN>2.0.CO;2.

Wallace, John M., Yuan Zhang, and James A. Renwick (1995). "Dynamic Contribution to Hemispheric Mean Temperature Trends". *Science* 270.5237, pp. 780–783. DOI: 10.1126/science.270.5237.780.

Watanabe, Masahiro et al. (2021). "Enhanced warming constrained by past trends in equatorial Pacific sea surface temperature gradient". *Nature Climate Change* 11.1, pp. 33–37. ISSN: 1758-6798. DOI: 10.1038/s41558-020-00933-3.

Watanabe, Masahiro et al. (2023). "Two Competing Drivers of the Recent Walker Circulation Trend". *Geophysical Research Letters* 50.23, e2023GL105332. ISSN: 1944-8007. DOI: 10.1029/2023GL105332.

Watanabe, Masahiro et al. (2024). "Possible shift in controls of the tropical Pacific surface warming pattern". *Nature* 630.8016, pp. 315–324. ISSN: 1476-4687. DOI: 10.1038/s41586-024-07452-7.

Weisheimer, A. and T. N. Palmer (2014). "On the reliability of seasonal climate forecasts". *Journal of The Royal Society Interface* 11.96, p. 20131162. DOI: 10.1098/rsif.2013.1162.

Weisheimer, Antje et al. (2020). "Seasonal Forecasts of the Twentieth Century". *Bulletin of the American Meteorological Society* 101.8, E1413–E1426. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-19-0019.1.

Weisheimer, Antje et al. (2024). "The Signal-to-Noise Paradox in Climate Forecasts: Revisiting Our Understanding and Identifying Future Priorities". *Bulletin of the American Meteorological Society* 105.3, E651–E659. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-24-0019.1.

Wilks, D. S. (2016). ""The Stippling Shows Statistically Significant Grid Points": How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It". *Bulletin of the American Meteorological Society* 97.12, pp. 2263–2273. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-15-00267.1.

Williams, Keith et al. (2015). "The Met Office Global Coupled model 2.0 (GC2) configuration". *Geoscientific Model Development* 8, pp. 1509–1524.

Williams, N. C., A. A. Scaife, and J. A. Screen (2023). "Underpredicted ENSO Teleconnections in Seasonal Forecasts". *Geophysical Research Letters* 50.5, e2022GL101689. ISSN: 1944-8007. DOI: 10.1029/2022GL101689.

Wills, Robert C. J. et al. (2022). "Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change". *Geophysical Research Letters* 49.17. ISSN: 0094-8276. DOI: 10.1029/2022GL100011.

Wodzicki, K. R. and A. D. Rapp (2016). "Long-term characterization of the Pacific ITCZ using TRMM, GPCP, and ERA-Interim". *Journal of Geophysical Research: Atmospheres* 121.7, pp. 3153–3170. ISSN: 2169-8996. DOI: 10.1002/2015JD024458.

Woods, Cian and Rodrigo Caballero (2016). "The Role of Moist Intrusions in Winter Arctic Warming and Sea Ice Decline". *Journal of Climate* 29.12, pp. 4473–4485. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-15-0773.1.

Woollings, T. et al. (2015). "Contrasting interannual and multidecadal NAO variability". *Climate Dynamics* 45.1, pp. 539–556. ISSN: 1432-0894. DOI: 10.1007/s00382-014-2237-y.

Woollings, Tim, Abdel Hannachi, and Brian Hoskins (2010). "Variability of the North Atlantic eddy-driven jet stream". *Quarterly Journal of the Royal Meteorological Society* 136.649, pp. 856–868. ISSN: 1477-870X. DOI: 10.1002/qj.625.

Woollings, Tim et al. (2023). "Trends in the atmospheric jet streams are emerging in observations and could be linked to tropical warming". *Communications Earth & Environment* 4.1, p. 125. ISSN: 2662-4435. DOI: 10.1038/s43247-023-00792-8.

Wright, J. S. and S. Fueglistaler (2013). "Large differences in reanalyses of diabatic heating in the tropical upper troposphere and lower stratosphere". *Atmospheric Chemistry and Physics* 13.18, pp. 9565–9576. DOI: 10.5194/ACP-13-9565-2013.

Wu, Mingna, Chao Li, and Zhongshi Zhang (2024). "Recalibrated projections of the Hadley circulation under global warming". *Environmental Research Letters* 19.10, p. 104041. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ad751f.

Xia, Yan, Yongyun Hu, and Jiping Liu (2020). "Comparison of trends in the Hadley circulation between CMIP6 and CMIP5". *Science Bulletin* 65.19, pp. 1667–1674. ISSN: 2095-9273. DOI: 10.1016/j.scib.2020.06.011.

Xu, Gaopeng et al. (2024). "High-resolution modelling identifies the Bering Strait's role in amplified Arctic warming". *Nature Climate Change* 14.6, pp. 615–622. ISSN: 1758-6798. DOI: 10.1038/s41558-024-02008-z.

Xu, Mian et al. (2023). "Important role of stratosphere-troposphere coupling in the Arctic mid-to-upper tropospheric warming in response to sea-ice loss". *npj Climate and Atmospheric Science* 6.1, pp. 1–9. ISSN: 2397-3722. DOI: 10.1038/s41612-023-00333-2.

Yamazaki, Kuniko et al. (2021). "A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: part 2: global performance and future changes". *Climate Dynamics* 56.11, pp. 3437–3471. ISSN: 1432-0894. DOI: 10.1007/s00382-020-05608-5.

Yanai, Michio, Steven Esbensen, and Jan-Hwa Chu (1973). "Determination of Bulk Properties of Tropical Cloud Clusters from Large-Scale Heat and Moisture Budgets". *Journal of the Atmospheric Sciences* 30.4, pp. 611–627. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1973)030.

Ye, Kunhui and Gabriele Messori (2021). "Inter-model spread in the wintertime Arctic amplification in the CMIP6 models and the important role of internal climate variability". *Global and Planetary Change* 204, p. 103543. ISSN: 0921-8181. DOI: 10.1016/j.gloplacha.2021.103543.

Ye, Kunhui et al. (2024). "Response of winter climate and extreme weather to projected Arctic sea-ice loss in very large-ensemble climate model simulations". *npj Climate and Atmospheric Science* 7.1, pp. 1–16. ISSN: 2397-3722. DOI: 10.1038/s41612-023-00562-5.

Yeager, Stephen G. et al. (2023). "Reduced Southern Ocean warming enhances global skill and signal-to-noise in an eddy-resolving decadal prediction system". *npj Climate and Atmospheric Science* 6.1, pp. 1–13. ISSN: 2397-3722. DOI: 10.1038/s41612-023-00434-y.

Yeh, Sang-Wook et al. (2018). "ENSO Atmospheric Teleconnections and Their Response to Greenhouse Gas Forcing". *Reviews of Geophysics* 56.1, pp. 185–206. ISSN: 1944-9208. DOI: 10.1002/2017RG000568.

Zaplotnik, Žiga, Matic Pikovnik, and Lina Boljka (2022). "Recent Hadley Circulation Strengthening: A Trend or Multidecadal Variability?" *Journal of Climate* 35.13, pp. 4157–4176. ISSN: 0894-8755. DOI: 10.1175/JCLI-D-21-0204.1.

Zhang, Yehui et al. (2011). "Climatological Characteristics of Arctic and Antarctic Surface-Based Inversions". *Journal of Climate* 24.19, pp. 5167–5186. ISSN: 0894-8755. DOI: 10.1175/2011JCLI4004.1.

Zhang, Yuan, John M. Wallace, and David S. Battisti (1997). "ENSO-like Interdecadal Variability: 1900–93". *Journal of Climate* 10.5, pp. 1004–1020. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2.

Zhao, Ming and Thomas Knutson (2024). "Crucial role of sea surface temperature warming patterns in near-term high-impact weather and climate projection". *npj*

Climate and Atmospheric Science 7.1, pp. 1–13. ISSN: 2397-3722. DOI: 10.1038/s41612-024-00681-7.

Zhao, Xueying and Robert J. Allen (2019). "Strengthening of the Walker Circulation in recent decades and the role of natural sea surface temperature variability". *Environmental Research Communications* 1.2, p. 021003. ISSN: 2515-7620. DOI: 10.1088/2515-7620/ab0dab.

Zhou, Chunlüe et al. (2021). "A New Approach to Homogenize Global Subdaily Radiosonde Temperature Data from 1958 to 2018". *Journal of Climate* 34.3, pp. 1163–1183. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-20-0352.1.

Zhou, Wenyu, L. Ruby Leung, and Jian Lu (2024). "Steady threefold Arctic amplification of externally forced warming masked by natural variability". *Nature Geoscience*, pp. 1–8. ISSN: 1752-0908. DOI: 10.1038/s41561-024-01441-1.