CHANGES IN INTERNAL CLIMATE VARIABILITY DUE TO GLOBAL WARMING.

by
Emerson Nicole LaJoie
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Climate Dynamics

Committee:

_____ Dr. Timothy M. DelSole, Dissertation Director

_____ Dr. Barry A. Klinger, Committee Member

_____ Dr. Edwin K. Schneider, Committee Member

_____ Dr. Edwin K. Schneider, Department Chair

_____ Dr. Donna M. Fox, Associate
Dean for Academic and Student
Affairs, College of Science

_____ Dr. Peggy Agouris, Dean,
College of Science

Date: _____ Summer Semester 2016
George Mason University
Fairfax, VA

Changes in Internal Climate Variability Due to Global Warming.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Emerson Nicole LaJoie
Master of Science
San Jose State University, 2009
Bachelor of Science
Indiana University, 2005

Director: Dr. Timothy M. DelSole, Professor
Department of Atmospheric, Oceanic, and Earth Sciences

Summer Semester 2016
George Mason University
Fairfax, VA

Dedication

To Matt Weber.
Without his dedication and encouragement, this endeavor would have been a lot less fun.

Acknowledgments

I would like to thank my advisor, Dr. Timothy DelSole, for his invaluable teaching style and exceptional standards. His ever encouraging counsel and criticism on how to be a better scientist were exactly what I had hoped for when I chose him as my mentor. His infinite patience and guidance allowed me to find my own path as a researcher. I also thank my committee members Dr. Barry Klinger and Dr. Ed Schneider for their advisement, support, and time. My research was also helped by the efforts of Dr. Xiaoqin Yan and Dr. Liwei Jia, both of whom made data available for my analysis. I also thank fellow colleagues who went before me, and with me, on this journey. I am indebted to you for not only your invaluable help, but also your friendship. I also thank the remaining faculty of the Climate Dynamics department for helping me along the way.

Table of Contents

List of Tables

List of Figures

Abstract

CHANGES IN INTERNAL CLIMATE VARIABILITY DUE TO GLOBAL WARMING.

Emerson Nicole LaJoie, PhD

George Mason University, 2016

Dissertation Director: Dr. Timothy M. DelSole

Extreme weather events can have serious impacts on society, infrastructure, and human life. Evidence is growing that the frequency and intensity of extreme weather events will increase in response to rising greenhouse gas concentrations. However, a consensus has yet to be reached as to whether these changes can be explained by a simple shift in the underlying probability distribution, or by a change in shape of the distribution (namely variance) as well. Previous studies have investigated this question by aggregating data across space, but aggregation requires normalizing data in some way to allow data from different geographic regions to be combined into a single distribution. Unfortunately, subsequent studies showed that the normalization procedure introduces biases. This dissertation proposes a new methodology for quantifying changes in variance that is rigorous, multivariate, and invariant to linear transformation (and thus independent of normalization). The new methodology is applied to simulations from state-of-the-art climate models and reveals significant changes in seasonal- and annual-mean 2m temperature and precipitation in response to anthropogenic forcing. The models consistently predict decreases in temperature variance in regions of seasonal sea-ice formation and across the Southern Ocean by the end of the twenty-first century. While more than half the models also predict significant changes

in variance over ENSO regions and the North Atlantic Ocean, the direction of this change is model dependent. Models also consistently predict widespread increases to precipitation variability, particularly in the tropics, extratropics, and polar latitudes. Some models predict more than a doubling in variance, raising questions about the adequacy of doubling uncertainty estimates to test robustness in detection and attribution studies.

Chapter 1: Introduction

Prolonged heatwaves and persistent droughts are inherent to our climate system and can have a serious impact on society, infrastructure, and human life (IPCC, 2012). Observational records show that as greenhouse gas concentrations increase, global mean temperature also rises, increasing the frequency of extreme weather events. Since the middle of the twentieth century, most land areas for which there are sufficient observational records have experienced increases in the frequency and intensity of warm extremes, and decreases in the frequency and intensity of cold extremes (Collins et al., 2013a; IPCC, 2012). These changes are consistent with the hypothesis that anthropogenic global warming acts to shift the distribution of temperature toward a warmer climate. An illustration of this shift is given by the top panel of fig. 1.1, which was published in the Special Report on Extreme Events (IPCC, 2012). However, changes in extreme weather could also result from changes in the variance of climate variability. The purpose of this dissertation is to investigate how internal climate variability will change in a future climate.

There is a general consensus that as global mean temperature rises, the global hydrological cycle intensifies and thus, the amount of rainfall increases (Trenberth et al., 2007). The moisture-holding capacity of the atmosphere increases at about 7% per degree C. However, the spatial distribution of that rainfall, as well as the frequency, intensity, duration, and type of rainfall can be geographically-dependent. For example in the tropics, where atmospheric convection leads to frequent thunderstorms, increased availability of water vapor leads to more intense precipitation. In the subtropics, or regions of subsidence, increased temperature leads to greater evaporation and thus enhanced surface drying. This precipitation pattern has been coined the 'rich get richer and poor get poorer' hypothesis in response to global warming (Held and Soden, 2006). On the one hand, these changes could be argued as due to simple shifts in the distribution of precipitation for regional climate zones.

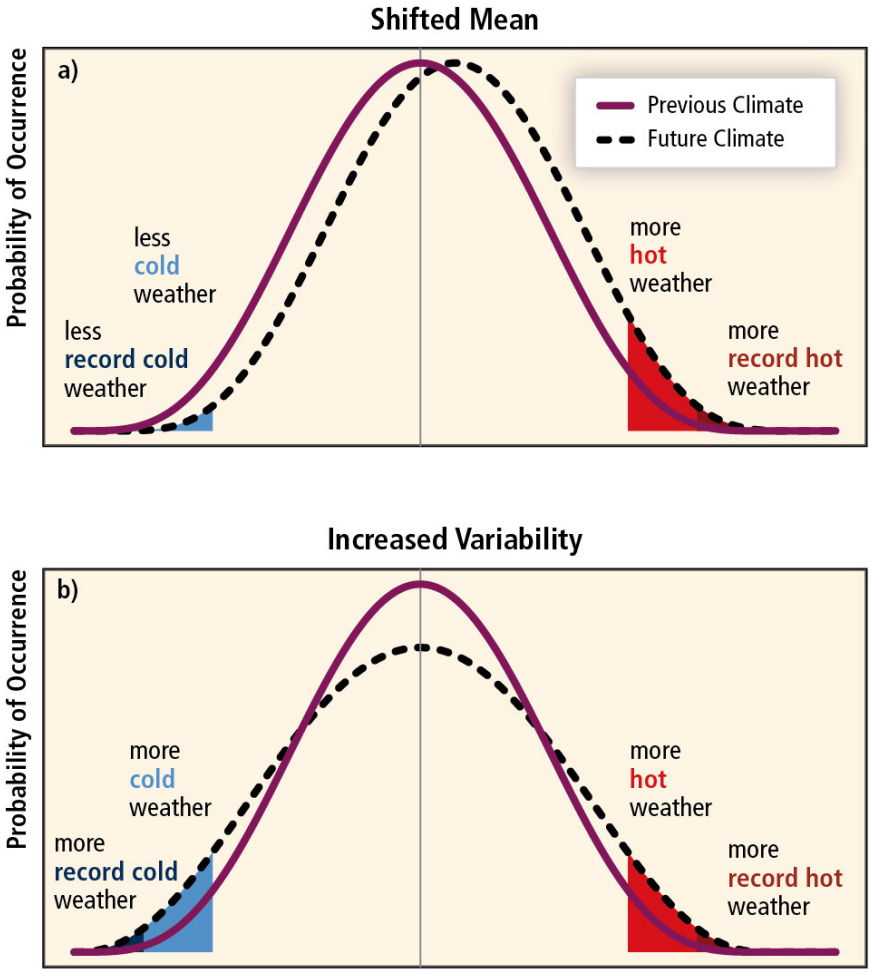

Figure 1.1: Reproduced from the IPCC Special Report on Extreme Events Chapter 1: Illustration of the effect on temperature distribution, specifically on extremes, of a warming climate; (a) effects of a simple shift of the entire distribution toward a warmer climate; (b) effects of an increase in temperature variability with no shift in the mean.

For instance, the subtropics receiving less rain could be explained by a simple shift toward drier weather, but it is also possible that drier conditions could be explained by changes in variance as well.

The middle panel in fig. 1.1 illustrates the affect of increased variability; notice that a wider distribution would mean more extreme weather events. The IPCC Special Report on Extreme Events states that the frequency of heavy precipitation (or the proportion of total rainfall from a heavy rain event) will likely increase for high latitudes and the tropics (IPCC, 2012). While at a medium confidence level, droughts are projected to intensify in the 21st century due to reduced precipitation and/or increased evapotranspiration in marginal convective zones. However, at other latitudes, projected changes in drought intensity are at a lower confidence (based on model inconsistencies and definitional issues).

While the observational record supports a shift in the mean distribution of temperature, the variability of temperature might also be changing, however, methods for quantifying global-scale changes in variance have been criticized. For instance, Hansen et al. (2012) claimed that the distribution of globally aggregated summer temperatures has both shifted toward a higher mean and broadened. Subsequent studies have supported Hansen et al. (2012) with respect to a shifting mean, but disagree that changes in variance have contributed to observed summer mean hot extremes (Coumou and Robinson, 2013; Huntingford et al., 2013; Rhines and Huybers, 2013). One source of disagreement is the procedure for normalizing temperatures at different geographic locations before aggregating them to obtain a distribution. In particular, removing the mean temperature in one period based on temperature statistics of an earlier period, as done in Hansen et al. (2012), imparts a positive bias to the variance because the sample mean does not vanish in the study period (Rhines and Huybers, 2013; Tingley, 2011). Another complicating factor is that the number of surface stations in a geographic region has changed over time. In particular, a decline in station density implies fewer stations for averaging, which in turn leads to larger variance (Rhines and Huybers, 2013). Finally, differences in trends between different geographic regions also contributes to differences in variance. After accounting for issues

related to normalization, trends, and data density, Rhines and Huybers (2013) find that changes to the variance of summer mean temperature cannot be detected. Consistent with this conclusion, Huntingford et al. (2013) find that if trends are removed by computing temperature anomalies relative to an 11-year local running mean, then changes in variability of seasonal/annual mean temperature also cannot be detected in observations. Looking at a high emissions scenario, Coumou and Robinson (2013) show that the exceedences of extreme monthly mean temperatures during summer are projected to increase and that those changes can be explained by a shift in the local mean temperature. Their results imply that local variability does not change.

Cold-season variability also has been investigated. Francis and Vavrus (2012) hypothesize that the decline of sea ice extent, due to Arctic warming, causes the jet stream to grow more wavy, resulting in more frequent cold extremes. If the frequency of cold extremes were increasing, and one accepts that the frequency of warm extremes also are increasing (IPCC, 2012), then these two changes cannot be explained solely by a shift in the mean because if cold *and* warm extremes are both increasing, this indicates a distribution is broadening and thus overall variance increases. However, Barnes (2013) showed that trends in planetary-scale waviness are sensitive to methodology and thus inconclusive. Similarly, Screen and Simmonds (2013) find that trends in wave amplitude are sensitive to whether they are measured in the zonal or meridional direction.

Screen (2014) examined a different quantity, namely zonal means of the local variance of daily temperature over Northern Hemispheric land, and concluded that temperature variance had *decreased* since 1979 for fall, winter, and spring. Screen (2014) argues that this decrease in variance is caused by Arctic amplification. Specifically, cold extremes in the Northern Hemisphere are invariably associated with winds that blow from the north. Arctic amplification, however, increases temperatures more in the Arctic than at low-latitudes, thereby reducing cold advection by northerly winds. Consequently, Arctic amplification causes the coldest days to warm faster than the warmest days, thus reducing variance. Consistent with this mechanism, climate models project less variable land temperatures

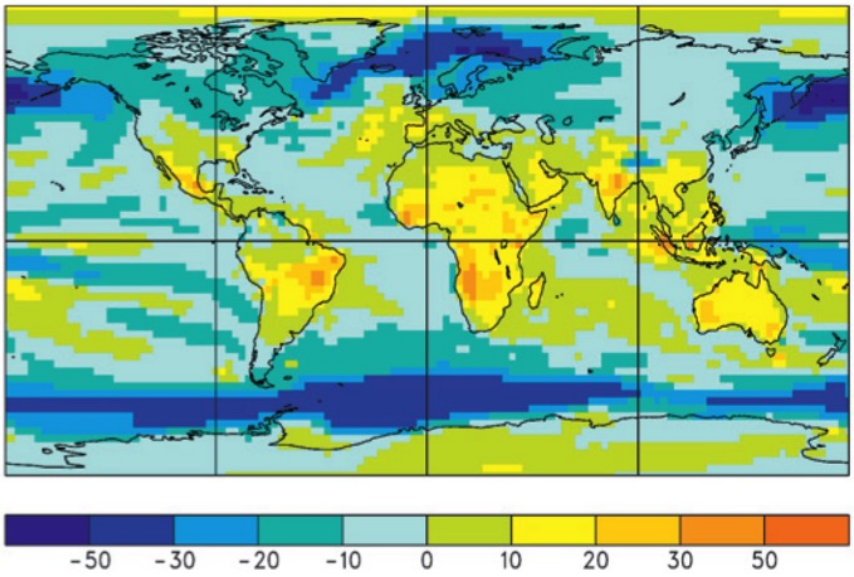

Figure 1.2: Reproduced from Boer (2009) fig. 4. The percentage change in the multimodel standard deviations for the last 150 years of the A1B stabilization simulation, relative to multimodel preindustrial control simulations, after removal of the forced component. Values are statistically different from zero at the 1% level with the exception of a narrow band about the zero lines (see colorbar).

in northern latitudinal bands during fall, winter, and spring, and models with stronger Arctic amplification tend to exhibit stronger decreases in variance (Screen, 2014). Other independent studies support this conclusion. For instance, Huntingford et al. (2013) showed that climate models predict, on average, a decrease in total variability of annual mean temperature in high emissions scenarios relative to an 11-year running mean, with some of this decrease associated with reductions in sea-ice cover. Also, Boer (2009) showed that climate models predict, on average, that the variance of temperature anomalies (relative to a low-order polynomial in time) will decrease in mid-latitudes and increase slightly in the tropics (see fig. 1.2).

As is the case with the studies discussed above, temperature anomalies at different geographic locations are often combined using spatial aggregation or averaging methods in order to draw a single conclusion about an *overall* change in variance. In addition to the loss of information about local changes in variance, these methods must also standardize

temperature anomalies to remove local differences in the means, variances, and trends prior to aggregating or averaging. Unfortunately, there is no unique normalization procedure and so criticisms can be raised about any chosen normalization procedure. Furthermore, there appears to be no rigorous significance test to evaluate the changes seen from these methods.

An alternate approach to combining data is to compute local changes in variance and then display maps of those changes. While this approach preserves information about local changes, it leads to a field significance problem in which the likelihood of the computed field of changes needs to be quantified relative to the null hypothesis of no local change in variance. Standard field significance techniques (e.g. Livezey and Chen (1983)) are designed for correlation maps and it is not clear how to apply them to variance ratio maps to test for field significance.

This dissertation was motivated by these challenges and sets forth two new methodologies for assessing spatially distributed changes in internal variability in response to increasing greenhouse gases. One approach is well-suited for identifying changes in a single mode of climate variability, while the other approach can determine if multiple components of variability are changing variance. While changes in variance have been quantified previously in a univariate or aggregate sense, the tests presented here are distinct, yet complimentary, approaches that account for spatial and temporal relationships within the chosen domain. In addition, the tests are invariant to normalization procedures or any other affine transformation of the data. While these tests do avoid certain problems that arise in spatial aggregation and have well-defined significance measures, they require severely restricting the dimension of the state space. Since no single test can capture all possible changes in variance, it is argued that applying these two tests is a reasonable and comprehensive approach to advancing our understanding of how the climate system responds to increasing greenhouse gases.

Chapter 2: Methodology

In this chapter, the data and the methodologies used to investigate changes in internal variability in a future climate are presented. Also presented, the rationale and techniques applied to check assumptions made about the data (e.g., that the data is distributed as a Gaussian or normal). In addition, the techniques applied to derive significance thresholds are also detailed.

2.1 Data

Global climate model simulations from phase 5 of the Coupled Model Intercomparison Project (CMIP5) are evaluated for changes in variance. Two types of simulations are analyzed: preindustrial control runs, in which the forcings do not change from year to year, and projections based on Representative Concentration Pathway 8.5 (RCP8.5), in which concentrations and emissions increase such that the radiative forcing peaks at 8.5 W m^{-2} in 2100 (Collins et al., 2013b). Seasonal- and annual- mean 2m temperature and precipitation are analyzed. Only the models with at least a 500-year long preindustrial control simulation and at least three ensemble members for their RCP8.5 simulations covering the period 2006 to 2095 are used in this study. These criteria resulted in a selection of seven global climate models (see Table 2.1 for details). All data were interpolated onto a common 5 x 5 degree grid, yielding 2592 total grid points for each model. The control simulations were detrended to remove the effects of model drift. To demonstrate robustness, the simulations were partitioned and the analyses were performed in each half separately. For instance, the 500 year control simulations were partitioned such that each half contains 250 years. Similarly, each member of the RCP8.5 90-yr simulations were divided into a first half (45

years) and second half (also 45 years). This yields 135 total years (3 * 45) for each half of an RCP8.5 simulation.

Table 2.1: List of climate models used in this study. Included in the table is the modeling center, the long-form model name, and the short-form model name generated for this investigation and referenced herein.

Coupled Global Cimate Models		
Model Center:	Model name:	Short name:
Canadian Centre for Climate Modelling and Analysis	CCCMA CanESM2	CCCma
Centre National de Recherches Meteorogiques Centre Europeen de Recherche et Formation Avancee en Cacul Scientifique	CNRM-CERFACS CM5	CNRM
Institut Pierre-Simon Laplace	IPSL-CM5A-LR	IPSL
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC5	MIROC5
Met Office Hadley Centre	MOHC HadGEM2-ES	HadGEM2
Max-Planck-Institut for Meteorology	MPI-M ESM-LR	MPI
National Center for Atmospheric Research	NCAR CCSM4	NCAR

2.2 Statistical Model

To quantify changes in internal variability due to anthropogenic forcing, assume that a climate variable, such as 2m temperature, $\mathbf{t}$, can be modeled as

$$\mathbf{t} = \mathbf{F} + \mathbf{u}, \tag{2.1}$$

where $\mathbf{F}$ is the forced response and $\mathbf{u}$ is a random term representing internal (or unforced) variability. The statistical model (2.1) is commonly used in climate change detection and attribution studies and it assumes that the statistics of internal variability are independent of time, have a known distribution, and are additive relative to the variability of the forced

component (e.g., Allen and Tett (1999), Jones et al. (2013), and Imbers et al. (2013)).

To the extent that this statistical model is correct, changes in internal variability due to anthropogenic forcing can be evaluated using ensemble techniques and then compared with estimates of internal variability simulated from preindustrial control runs. To see this, consider an ensemble of simulations initialized from different states but driven by the same forcings. For such an ensemble, $\mathbf{F}$ in (2.1) is the same for different ensemble members, hence the difference between ensemble member and ensemble mean leads to cancellation of the forced response (i.e., cancellation of $\mathbf{F}$), in which case, the residual can be considered internal variability. The residual has slightly less variance than the true internal variability because the ensemble mean that is removed also contains some internal variability due to the finite ensemble size. In this work, the ensemble mean from a 3-member RCP8.5 simulation is evaluated and subtracted from each member. The future emissions scenario is shorthanded as "21C." Let $t^{21C}_{s,y,e}$ be a climate variable from the e^{th} ensemble member at the s^{th} spatial grid point and the y^{th} year. Then, a (slightly damped) realization of internal variability in the twenty-first century is

$$U^{21C}_{s,y,e} = t^{21C}_{s,y,e} - [t]^{21C}_{s,y}, \tag{2.2}$$

where the ensemble mean of the twenty-first century simulations is

$$[t]^{21C}_{s,y} = \frac{1}{E}\sum_{e=1}^{E} t^{21C}_{s,y,e}, \tag{2.3}$$

and E is the total ensemble size. For the preindustrial control runs, let $t^{ctr}_{s,y}$ be a climate variable from the preindustrial control simulation ("ctr") at the s^{th} spatial grid point and the y^{th} year. Since the forcing does not change from year to year in the preindustrial control run, $\mathbf{F}$ in (2.1) is constant. Therefore, internal variability in the control run can be estimated from the residual about the time mean. Again, the residual has slightly less variance than

the true internal variability because the time mean contains internal variability due to finite sample size. A (slightly damped) realization of internal variability in the absence of anthropogenic forcing is

$$U_{s,y}^{ctr} = t_{s,y}^{ctr} - \bar{t}_s^{ctr}, \tag{2.4}$$

where the climatological mean is

$$\bar{t}_s^{ctr} = \frac{1}{Y_{ctr}} \sum_{y=1}^{Y_{ctr}} t_{s,y}^{ctr}. \tag{2.5}$$

2.3 Univariate Test for Changes to Internal Variability

At each grid point, s, changes in variability caused by anthropogenic forcing can be assessed by testing the null hypothesis that (2.2) and (2.4) were drawn from populations with equal variances. Standard Analysis of Variance (ANOVA) techniques show that an *unbiased* estimate of the variance from a realization of internal variability in a twenty-first century simulation can be determined from

$$\sigma_{s,21C}^2 = \frac{1}{Y_{21C}(E-1)} \sum_{y=1}^{Y_{21C}} \sum_{e=1}^{E} (U_{s,y,e}^{21C})^2, \tag{2.6}$$

and an unbiased estimate of variance in the control simulations is

$$\sigma_{s,ctr}^2 = \frac{1}{Y_{ctr}-1} \sum_{y=1}^{Y_{ctr}} (U_{s,y}^{ctr})^2. \tag{2.7}$$

If the samples are independent and identically distributed (*iid*) as a Gaussian (or normal distribution), then standard statistical theory states that the statistic

$$F_s = \frac{\sigma^2_{s,21C}}{\sigma^2_{s,ctr}}, \tag{2.8}$$

has an F-distribution with $Y_{ctr} - 1$ and $Y_{21C}(E-1)$ degrees of freedom. The above statistic will be called the *21C-noise to control ratio*. A priori, it is unknown which direction the internal variability may change, so two-tailed test is used to determine the significance of the ratio in (2.8). If the null hypothesis is true, this ratio should be close to one, whereas values far from one indicate that the variances differ and suggest that anthropogenic forcing changes internal variability. The statistic (2.8) is *univariate* because it compares variances at a single grid point. The spatial distribution of the variance ratios, F_s, can be visualized as a *field* of individual ratios, referred to as a variance ratio map.

2.4 Null Hypothesis for Field Significance

The F-test defined above determines the significance of changes in internal variability at individual grid points. The next step is to quantify the likelihood that a collection of variance ratios, F_s, could have occurred by random chance under a null hypothesis of no change in local variance and no change in the covariance between grid points. The central issue in testing field significance is accounting for dependencies between grid points. These dependencies can be quantified by a covariance matrix. It is proposed that the appropriate null hypothesis for testing differences between fields of variances is that the respective distributions have the same covariance matrix. Thus, if $\mathbf{\Sigma}_{21C}$ and $\mathbf{\Sigma}_{CTR}$ are the covariance matrices of internal variability in the 21C and preindustrial control simulations, respectively, then the null hypothesis is

$$H_0 : \mathbf{\Sigma}_{21C} = \mathbf{\Sigma}_{CTR}. \tag{2.9}$$

In each covariance matrix, the diagonal elements give the variances at different spatial locations and the off-diagonal elements quantify the degree of dependence (covariance) between locations. The off-diagonal elements are included in the null hypothesis because these elements define the dependencies between grid points that are essential to determining field significance.

Testing hypotheses about covariance matrices requires estimating the covariance matrix itself. Unfortunately, sample covariance matrices estimated from gridded data will be singular because the number of grid points far exceeds the number of samples. The standard approach to this problem is to project the data onto a smaller dimensional subspace and then apply the test in the reduced space. Let the basis vectors for this subspace be denoted $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_T$. Borrowing terminology from machine learning, the basis vectors will be called *feature vectors*, and the space spanned by these vectors will be called the *feature space*. These vectors can be collected as column vectors of the matrix

$$\mathbf{E} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_T \end{pmatrix}. \tag{2.10}$$

The standard choice for feature vectors are the leading empirical orthogonal functions (EOFs, also known as principal components) of the data. Unfortunately, the EOFs are model dependent and comparison between models is not straightforward. To avoid these problems, basis vectors are chosen based on a truncated set of eigenvectors of the Laplace operator. Laplacian eigenvectors are orthogonal in space and can be ordered by a measure of length scale. Thus, representing data with a truncated set of Laplacian eigenvectors is equivalent to filtering out variability on small spatial scales. In addition, Laplacian eigenvectors depend only on the geometry of the domain and therefore will be the same for each model. Over a global domain, Laplacian eigenvectors are equivalent to spherical harmonics. Over limited domains on a sphere, the Laplacian eigenvectors are determined numerically using the method of DelSole and Tippett (2015).

Time series for the Laplacian eigenfunctions are derived from the *pseudo-inverse*, $\mathbf{E}^i$,

which has the property

$$\mathbf{E}^T\mathbf{E}^i = \mathbf{I}. \tag{2.11}$$

It is convenient to use the matrix notation

$$\left(\mathbf{U}^{21C}\right)_{s,y+Y(e-1)} = U^{21C}_{s,y,e}, \tag{2.12}$$

where the ensemble members of a 21C simulation have been "stacked" into a single time series. Similarly, the matrix notation for a control run is,

$$\left(\mathbf{U}^{ctr}\right)_{s,y} = U^{ctr}_{s,y}. \tag{2.13}$$

Then, projecting the pseudo-inverse on a twenty-first century simulation yields a matrix with Y_{21C} years for each T time series, referred to as *feature variables*,

$$\left(\mathbf{U}^{21C}\right)^T\mathbf{E}^i = \mathbf{F}_{21C}. \tag{2.14}$$

Also, projecting the pseudo-inverse on a control simulation yields a matrix with Y_{ctr} years for each T time series,

$$\left(\mathbf{U}^{ctr}\right)^T\mathbf{E}^i = \mathbf{F}_{ctr}. \tag{2.15}$$

Recalling that estimates of a realization of internal variability have zero sample mean, an unbiased estimate of the covariance matrix of internal variability in a twenty-first century simulation is

$$\tilde{\mathbf{\Sigma}}_{21C} = \frac{1}{Y_{21C}(E-1)}\mathbf{F}^T_{21C}\mathbf{F}_{21C}, \tag{2.16}$$

where the tilde indicates a sample covariance matrix in feature space. Similarly, an unbiased

estimate of the covariance matrix of a control simulation is

$$\tilde{\mathbf{\Sigma}}_{CTR} = \frac{1}{Y_{ctr} - 1}\mathbf{F}_{ctr}^T\mathbf{F}_{ctr}. \tag{2.17}$$

2.5 Discriminant Analysis

Discriminant analysis is an optimization technique that finds a linear combination of variables that maximize a variance ratio. Let the weights for the linear combination be represented by the vector $\mathbf{q}$, such that the variates for the 21C and CTR simulations are

$$\mathbf{r}_{21C} = \mathbf{F}_{21c}\mathbf{q} \quad \text{and} \quad \mathbf{r}_{ctr} = \mathbf{F}_{ctr}\mathbf{q}. \tag{2.18}$$

Since both variates are centered, the sample variance is simply the sum square of a variate divided by the number of degrees of freedom:

$$\sigma_{q,21C}^2 = \frac{1}{Y_{21C}(E-1)}\mathbf{r}_{21C}^T\mathbf{r}_{21C} = \frac{1}{Y_{21C}(E-1)}\mathbf{q}^T\mathbf{F}_{21C}^T\mathbf{F}_{21C}\mathbf{q} = \mathbf{q}^T\tilde{\mathbf{\Sigma}}_{21C}\mathbf{q} \tag{2.19}$$

$$\sigma_{q,ctr}^2 = \frac{1}{Y_{ctr} - 1}\mathbf{r}_{ctr}^T\mathbf{r}_{ctr} = \frac{1}{Y_{ctr} - 1}\mathbf{q}^T\mathbf{F}_{ctr}^T\mathbf{F}_{ctr}\mathbf{q} = \mathbf{q}^T\tilde{\mathbf{\Sigma}}_{ctr}\mathbf{q}. \tag{2.20}$$

By using (2.18), it follows that the ratio of variances and covariances between the 21C and CTR can be written as,

$$\lambda = \frac{\sigma_{q,21C}^2}{\sigma_{q,ctr}^2} = \frac{\mathbf{q}^T\tilde{\mathbf{\Sigma}}_{21C}\mathbf{q}}{\mathbf{q}^T\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q}}. \tag{2.21}$$

If the null hypothesis of equal covariances is true, then $\lambda = 1$ for all possible $\mathbf{q}$. Conversely, if the null is *not* true, then $\lambda \neq 1$ for at least one $\mathbf{q}$. The goal, then, is to find the weighting

coefficients that makes λ an *extremum*. The extremum can be found by solving $\partial\lambda/\partial\mathbf{q} = \mathbf{0}$:

$$\begin{aligned}\frac{\partial\lambda}{\partial\mathbf{q}} &= \frac{2\tilde{\mathbf{\Sigma}}_{21C}\mathbf{q}}{\mathbf{q}^T\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q}} - 2\frac{\mathbf{q}^T\tilde{\mathbf{\Sigma}}_{21C}\mathbf{q}}{(\mathbf{q}^T\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q})^2}\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q} \\ &= \frac{2}{\mathbf{q}^T\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q}}\left(\tilde{\mathbf{\Sigma}}_{21C}\mathbf{q} - \lambda\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q}\right) = 0. \end{aligned} \tag{2.22}$$

Since $\tilde{\mathbf{\Sigma}}_{CTR}$ is positive definite, the derivative vanishes if

$$\tilde{\mathbf{\Sigma}}_{21C}\mathbf{q} = \lambda\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q}. \tag{2.23}$$

Equation (2.23) is a generalized eigenvalue problem. Solving this problem yields T distinct eigenvector solutions, or *weighting vectors*, denoted as $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_T$. The corresponding eigenvalues, or *discriminant ratios*, characterize differences between the covariance matrices and, by convention, are ordered largest to smallest as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_T$. The largest ratio, λ_1, is called the *leading* discriminant ratio and gives the maximum variance ratio out of all the possible weighting vectors $\mathbf{q}$. The last ratio, λ_T, is called the *trailing* discriminant ratio and gives the minimum variance ratio out of all possible weighting vectors.

Discriminant analysis is an attractive technique because if the covariance matrices differ, then the differences can be diagnosed with a set of variates and spatial patterns. In particular, the variates can be plotted as time series to facilitate a physical interpretation of the differences. To see this, let the eigenvectors $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_T$, obtained from the optimal weight vectors, be collected into the matrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \ldots & \mathbf{q}_T \end{pmatrix}. \tag{2.24}$$

The variates, which are time series in this application, can be derived from (2.18). The

entire collection of variates for the 21C simulation are given by the columns of

$$\mathbf{R}_{21C} = \mathbf{F}_{21C}\mathbf{Q}. \tag{2.25}$$

Similarly, the variates for the control simulation are

$$\mathbf{R}_{CTR} = \mathbf{F}_{CTR}\mathbf{Q}. \tag{2.26}$$

For each variate, there is a corresponding pattern, or *loading vector*. To find the loading vector, let $\mathbf{P}$ be an $S \times T$ matrix that minimizes

$$\|\mathbf{U}^{21C} - \mathbf{R}_{21C}\mathbf{P}^T\|^2 \tag{2.27}$$

where $\|\mathbf{A}\|^2$ denotes the Frobenius norm of matrix $\mathbf{A}$. Equation (2.27) presents a standard least squares problem with the following solution,

$$\mathbf{P} = (\mathbf{U}^{21C})^T\mathbf{R}_{21C}\left(\mathbf{R}_{21C}^T\mathbf{R}_{21C}\right)^{-1} = \frac{1}{Y_{21C}(E-1)}(\mathbf{U}^{21C})^T\mathbf{R}_{21C}. \tag{2.28}$$

The j'th loading vector is merely the j'th column of $\mathbf{P}$, which is

$$\mathbf{p}_j = \frac{1}{Y_{21C}(E-1)}\mathbf{U}^{21C}\mathbf{r}_{21C,j}. \tag{2.29}$$

This expression reveals that each spatial point of the loading vector is simply the sample regression coefficient (i.e., the slope) between the data at that location and the j'th variate.

It is convenient to normalize the variates to have unit variance in the control simulation. According to (2.20), this normalization implies

$$\mathbf{q}_k^T\tilde{\mathbf{\Sigma}}_{CTR}\mathbf{q}_k = 1. \tag{2.30}$$

Because the covariance matrices in the generalized eigenvalue problem (2.23) are symmetric, the eigenvectors satisfy the orthogonality constraints

$$\mathbf{Q}^T \tilde{\mathbf{\Sigma}}_{21C} \mathbf{Q} = \mathbf{\Lambda} \quad \text{and} \quad \mathbf{Q}^T \tilde{\mathbf{\Sigma}}_{CTR} \mathbf{Q} = \mathbf{I}. \tag{2.31}$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements equal the eigenvalues. These orthogonality constraints imply that the variates satisfy the constraints

$$\mathbf{R}_{21C}^T \mathbf{R}_{21C} = \mathbf{\Lambda} \quad \text{and} \quad \mathbf{R}_{ctr}^T \mathbf{R}_{ctr} = \mathbf{I}, \tag{2.32}$$

which in turn imply that the variates $\mathbf{r}_{21C,k}$ and $\mathbf{r}_{21C,j}$ are uncorrelated for $k \neq j$, and similarly, the variates from the control simulation $\mathbf{r}_{ctr,k}$ and $\mathbf{r}_{ctr,j}$ are uncorrelated for $k \neq j$.

Another property of the eigenvalues is that they are invariant to affine transformations of the data. For instance, centering and normalizing by a standard deviation are special cases of affine transformations. This invariance property means that any function of the eigenvalues is also invariant to affine transformations. Additionally, the invariance property means that the sampling distribution of the eigenvalues are independent of the mean and covariance matrix of the population. As such, significance thresholds can be estimated by straightforward Monte Carlo methods, as discussed below. However, significance is also tested under non-normal assumptions. These procedures are discussed in detail in sec. 2.6.

Finally, as with all statistical optimization procedures, *overfitting* is a concern when the number of parameters being estimated (e.g., the eigenvectors) is not a small fraction of the sample size. To guard against overfitting, conclusions are checked for robustness by comparing results across independent data sets.

2.6 Deriving Significance Thresholds

In order to evaluate the significance of the eigenvalues themselves, or functions of the eigenvalues, it is necessary to estimate the sampling distribution of the eigenvalues. Recall

that eigenvalues are invariant to non-singular linear transformations of the data. Since any positive definite matrix can be transformed into any other positive definite matrix (of the same dimension) by a suitable linear transformation, it follows that the sampling distribution of the eigenvalues is independent of the covariance matrix of the population. Therefore, one can choose the most convenient covariance matrix, which is the identity matrix. Accordingly, a random number generator is used to produce normally distributed and independent and identically distributed (*iid*) numbers. These numbers are used to fill up data matrices of the same size as those in (2.14) and (2.15), then the covariance matrices (2.16) and (2.17) are computed and discriminant analysis is performed. These steps are repeated 1000s of times to build an empirical distribution of the eigenvalues or functions of the eigenvalues, from which significance thresholds can be derived.

The above Monte Carlo method assumes a normal distribution. An alternate way to evaluate significance, with fewer assumptions, is based on permutation methods, in which the data is repeatedly sampled to build the sampling distribution. For instance, a control simulation is randomly sampled (without replacement) to create two matrices with dimensions that match (2.14) and (2.15). Then the covariance matrices (2.16) and (2.17) are computed and discriminant analysis is performed. These steps are repeated 1000s of times to build an empirical distribution of the eigenvalues, or functions thereof, from which significance thresholds can be derived. Comparing significance thresholds derived from the permutation method and from the Monte Carlo procedure then allows us to check if the normal and iid assumptions made about the data via the Monte Carlo method, are a good fit for the data.

The permutation method can also account for auto-correlations. After performing an auto-correlation test on the preindustrial control simulations (not shown), serial correlations ranging from 2 - 5 years were discovered that violate the assumption that the data are independent in time. To account for such autocorrelations, the permutation procedure samples consecutive blocks of data in order to preserve the autocorrelation structure. By adjusting the block size from 1- to 5-year blocks, serial correlations could be accounted for

when deriving the significance thresholds.

2.7 Union-Intersection Test

A standard test for differences in covariance matrices is the *union-intersection test* (Flury, 1985). This test is based on evaluating the significance of the leading and trailing discriminant ratios, λ_1 and λ_T. In essence, H_0 in (2.9) is rejected if λ_1 is too large or λ_T is too small (i.e., if the ratio is far from 1 in either direction). In this application, a significant λ_1 implies anthropogenic forcing *increases* internal variability, while a significant λ_T implies anthropogenic forcing *decreases* internal variability. If a change in variance is detected, the associated eigenvector can be used to derive a spatial pattern and time series that explains the difference, thereby facilitating visualization and physical interpretation of those changes in internal variability.

The union-intersection test is well suited for identifying changes in variance caused by a single component of internal variability. For example, if anthropogenic forcing causes a global-scale ENSO teleconnection pattern to change variance, then the union-intersection test has the potential to detect this change. Shown in the following chapter, the results from applying the union-intersection test to CMIP5 simulations led to conclusions that were, for the most part, sensitive to truncation and thus difficult to interpret. In addition, projecting the discriminants onto an independent data set led to variance ratios that were mostly marginally significant. The fact that the leading or trailing discriminant ratios tended to be insignificant, or only marginally significant, implies that changes to internal variability in a single component are weak or non-existent. This result, however, does not imply that internal variability does not change. For instance, numerous independent modes might change their variances, but the change in any *individual* mode might be too small to satisfy statistical significance using the union-intersection test. Therefore, a test that can detect relatively small changes in variance that might be "spread" across many independent modes is needed.

2.8 Divergence

Another measure of the difference between two covariance matrices is the following:

$$D_T = \frac{1}{2}\,\mathrm{tr}\left[(\tilde{\mathbf{\Sigma}}_{21C}^{-1} + \tilde{\mathbf{\Sigma}}_{CTR}^{-1})(\tilde{\mathbf{\Sigma}}_{CTR} - \tilde{\mathbf{\Sigma}}_{21C})\right]. \tag{2.33}$$

This measure will be called *divergence*. For Gaussian distributions, this measure can be derived from the Kullback-Leilber Divergence, which itself is fundamental to a wide range of applications, including information theory, finance, coding theory, and quantum entanglement (Cover and Thomas, 1991; Jaeger, 2007; Kullback, 1968). An equivalent expression for divergence can be written in terms of a sum of discriminant ratios,

$$D_T = \frac{1}{2}\sum_{i=1}^{T}\left(\lambda_i + \frac{1}{\lambda_i} - 2\right), \tag{2.34}$$

where λ_i is the i^{th} eigenvalue from (2.21) (see Kullback, 1968, ch 9 Eq. 6.7, and note the means are zero). Notice that if the covariance matrices are equal, then all the eigenvalues equal 1 and $D_T = 0$. More specifically, the function $\lambda + 1/\lambda$ is a minimum when $\lambda = 1$ and becomes large when λ is either very large or close to zero because the function involves both the eigenvalue and its inverse. In contrast to the union-intersection test, D_T depends on the whole spectrum of eigenvalues up to the cutoff T. Thus, changes in variance that are "spread" across many independent feature variables will inflate individual eigenvalues and thereby accumulate in the sum to produce a large value of D_T.

Chapter 3: Results

In this chapter, the methodologies presented in chapter 2 are applied to model simulations of temperature and precipitation, for annual and seasonal means, to assess changes in internal variability in response to anthropogenic forcing. Changes in the absence of anthropogenic forcing are also investigated and discussed. In addition to global-scale changes, changes on sub-global scales, such as land and ocean only domains, are also presented.

3.1 Annual Mean 2m Temperature

3.1.1 Changes in Internal Variability Due to Anthropogenic Forcing: Global

The local changes in internal variability of annual mean 2m temperature due to anthropogenic forcing in a future climate are shown in fig. 3.1 for the 7 models investigated in this study (see Table 2.1 for model long-form names). The changes are quantified by the ratio of variance of twenty-first century residuals (relative to an ensemble mean) from an RCP8.5 emissions scenario over the variance of a preindustrial control run (for the respective model). Insignificant values at the 10% significance level (according to a standard F-test) are not colored. Along with the model name, in the title of each F-map is the percentage of total grid points that are deemed significant for that model.

All the models project significant local changes in internal variability for their RCP8.5 emissions scenario (relative to the model's control variability). Statistically speaking, one could expect to find 10% of the area (or about 25 grid points, in this study) to be significant just by random chance, but clearly fig. 3.1 shows more than that. In particular, the models show between 39% (NCAR) and 66% (MIROC5) of their respective total area are significantly changing variance. Grid points where the ratios are greater than 1 (warm

colors) indicate 21st century variability increases in response to anthropogenic forcing. For instance, a value of 2 indicates that variance is projected to double relative to control variability. Conversely, ratios less than 1 (cool colors) indicate anthropogenic forcing decreases variability. Consistently, each model projects decreases in variability in regions of seasonal sea-ice formation such as the Southern Ocean, the seas around Greenland, and the Bering Sea. This decrease could be explained by the loss of sea-ice due to a warmer future climate; sea-ice loss exposes the underlying sea surface, which has a larger effective heat capacity relative to sea-ice, thereby reducing temperature differences with the overlying atmosphere. Huntingford et al. (2013) and DelSole et al. (2013) also noted variance decreases associated with areas of sea-ice formation.

Other local changes in variance are model dependent. For instance, most models project significant changes in variability in the tropical oceans and in regions of the El Niño Southern Oscillation (ENSO), but the direction of these changes is model dependent. The North Atlantic Ocean exhibits significant changes in variance, but again, the direction of that change is also model dependent. At the North Pole, a majority of the models indicate increases in variability. There are several interesting, smaller-scale changes, like those in the Amazon Basin, however, the scale and model-dependent direction of these changes makes them difficult to interpret.

Climate variability is often expressed in terms of "modes" of variability. Accordingly, the above changes in variance are investigated to determine if they can be explained by a few patterns of internal variability. Instead of guessing a pattern and investigating its variability as a function of forcing, the pattern that maximizes the variance ratio is identified first, and then the union-intersection test is applied to assess the statistical significance. It turns out, however, that the union-intersection test led to results that were difficult to interpret because the significance of the leading (or trailing) ratios were marginally significant for most models. For instance, fig. 3.2 shows a representative example for one model, CCCma. The maximized ratio as a function of the accumulated number of Laplacian eigenvectors is shown by the blue circle-dashed curves, and the corresponding 5% and 95% significance

thresholds are shown as solid blue curves (derived from Monte Carlo techniques). It can be seen that the maximized ratio is sometimes inside and sometimes outside the significance band, depending on the number of Laplacian eigenvectors. This behavior was deemed to characterize marginally significant, sensitive, and/or not robust results. The minimized ratio is shown in red and exhibits similar, marginally significant behavior. Note that the solid curves are monotonic functions of the number of Laplacians. This monotonic behavior reflects the impact of overfitting, in other words, the variance ratio moves farther from 1 as a mathematical necessity, even under a no-change hypothesis, because each additional Laplacian provides extra freedom to fit differences in variances.

To confirm the lack of robustness, the variance ratio computed from the first half of the 21C and the first half of the control simulation was maximized (via discriminant analysis), then the resulting weights were projected onto the second halves of a 21C and control simulation. The resulting variance ratios computed from the independent data are shown as the square-dash curves in fig. 3.2. The blue and red square-dash curves show the independent maximized and minimized ratios, respectively. The corresponding significance lines (black straight dashed lines) were derived from an F-distribution at the 10% significance level. It is clear that the independent ratios also move in and out of the significance thresholds as a function of the number of Laplacian eigenvectors. Thus the component that maximized the variance ratio in the first half of the data set exhibits no significant change in variance in the second half of the data, consistent with the conclusion that the discriminants are not robust. Similar (marginally significant) results occurred for all the models in this study. These results indicate that the differences in internal variability of annual mean 2m temperature that were displayed in fig. 3.1, cannot be explained by changes in any *single* component of climate variability. This result suggests that large scale components of variability, such as ENSO, are not significantly changing variance.

However, this result does not imply that internal variability is not changing in response to anthropogenic forcing. For instance, changes in individual components could be small and insignificant on an individual basis, but large and significant in an overall aggregate sense.

As explained in sec. 2.8, an alternate multivariate approach for measuring variance changes is based on divergence. More precisely, divergence measures the difference in covariance matrices and vanishes if and only if all covariances are identical.

The divergence of annual mean 2m temperature in the model simulations is shown in fig. 3.3. Again, to check robustness, the divergence procedure was repeated on two separate halves of the 21C and preindustrial control simulations. The divergence for the first half is given by the blue curve and the divergence results for the independent half is given by the red curve. Also plotted in each panel is a yellow-brown curve, the relevance of which will be explained in the next section. The conclusions hold for truncations beyond 40, but for clarity a cutoff was chosen. Both divergence curves lie above the solid significance curves (at a 10% significance level), except for one model, IPSL, which indicates that the first half is marginally significant for this model. In LaJoie and DelSole (2016 accepted), the same divergence test was applied, however in that study, the discriminants were calculated from empirical orthogonal functions (EOFs) instead of Laplacians. With EOFs, both divergence results were significant for IPSL (fig. 3 LaJoie and DelSole (2016)). This difference reflects that changes in variability project more strongly on EOFs compared to Laplacian eigenvectors and, with regard to IPSL, this difference highlights that changes in variability in that model are more small scale relative to other models.

The impact of overfitting also can be seen in the monotonic increase of the significance curves as a function of the number of Laplacians used in the computation. However notice that for each model, the actual divergence increases faster than that of the significance curves, indicating that the change in variance is larger than expected by random chance. Note also that changes are larger in the second half of the century for all but one model (MIROC5), indicating that simulated changes in internal variability are generally larger in the second half of the 21st century as compared to the first half (as one would expect).

The significance thresholds (at 10%) shown in fig. 3.3 were computed three different ways with the intent to test the sensitivity of certain assumptions about the population. First, the significance thresholds were estimated by Monte Carlo methods in which independent and

identically distributed (*iid*) random numbers were drawn from a normal distribution (solid brown curve). The upper and lower 5% significance thresholds are hard to see because they fit nearly perfectly beneath the significance threshold derived from a permutation method using a 1-year block size (solid black curve). Recall that the permutation method randomly draws years in a control run to construct sample covariance matrices and therefore assumes only that the data are independent in time and identically distributed; it does not make the assumption that the data are normally distributed. The similarity between the brown and black significance thresholds implies that the assumption that the control data is normally distributed is reasonable. On the other hand, differences between the thresholds derived from the permutation method with 1-year blocks and 5-year blocks (solid gray curves) indicates that internal variability in the leading Laplacians are autocorrelated to some extent. However, it is important to note that even after accounting for auto correlations, changes in internal variability in response to anthropogenic forcing are much greater than those expected under a no-change null hypothesis, even when serial correlations are present: this is indicated by noting that the blue and red curves are above the gray curves.

A final remark: the results in fig. 3.1 show that two models (MIROC5 and MPI) exhibit widespread areas in which the variance of internal variability more than doubles by the end of the twenty-first century. Some detection and attribution studies artificially inflate a model's internal variability by a factor of two to assess the robustness of uncertainty in the estimates of internal variability (Hegerl et al., 2007). Such studies also assume that the statistical properties of internal variability do not change in response to climate forcing. Thus, for these models, not only is the assumption of constant internal variability incorrect, but doubling the internal variability may not be sufficient to account for changes in variability due to anthropogenic forcing.

3.1.2 Changes in Control Variability

Previous studies find that internal variability can change on multicentennial timescales even in the absence of anthropogenic forcing (e.g., Wittenberg, 2009). It is reasonable, then, to

question if the projected changes in variance for the 21C might occur "naturally". This question is investigated by using the methods discussed in chapter 2 to quantify changes in variance solely within preindustrial control simulations. In particular, a 500-yr control simulation is partitioned such that there are two segments, each with 250-years. The local differences in variance between the segments can be quantified with an F-test. The resulting F-ratios are shown in fig. 3.4 and insignificant values (according to an F-test distribution at a 10% significance level) are masked out. The color scale and tick marks are the same as those in fig. 3.1 and can be interpreted as the percent change in variance between the two halves of a given control run. The percent of total grid points deemed significant is given in the title of each F-map. There are numerous (locally) significant changes for each model. In general, one might expect to find 10% of any given field of F-ratios to be significant just by random chance and for two models (CCCma and NCAR) the percentage of significant grid points falls into this random category. However, for the remaining 5 models, fig. 3.4 shows more changes than would be expected by random chance.

Recall that an F-test assumes the data are normally distributed and *iid*. To determine if the differences in variance could be explained by non-normal behavior or serial correlations, significance thresholds were derived for each grid point individually using a permutation method with a 1-yr and a 5-yr block size. Recall that the permutation method with a 1-yr block only assumes *iid*, but not normal. Permuting with a 5-yr block does not make either the normal or *iid* assumption and can account for autocorrelations. After applying the 1-yr permutation method, the significance thresholds at each grid point were close to those derived from an F-distribution (not shown). Repeating permutation with a 5-year block, again revealed little difference relative to an F-distribution. As such, the assumptions made by the F-test were deemed appropriate for these control simulations, even after accounting for possible autocorrelations by resampling with 5-yr blocks.

The union-intersection test was applied to determine if the changes displayed in fig. 3.4 could be explained by changes in a single component of variability. The resulting leading and trailing discriminant ratios are given in fig. 3.5 (log plot). Also shown are the

significance thresholds (at 10%) derived from Monte Carlo techniques. A majority of the models indicate that the significance of the leading discriminant is sensitive to truncation, and with the exception of one model, the trailing ratios are not significant (MIROC5 dark green curve). These results indicate that the local changes in variance displayed in fig. 3.5 cannot be explained by changes in a single component of variability (for a majority of models).

Next, the divergence was computed for two segments of each control run. The resulting values are shown for each model in fig. 3.6. Also shown are the significance thresholds via permutation methods (at 10%). For most of the models, the divergence curves lie just above the 1-yr significance curves (black), indicating that there are some significant changes in variance present in these control runs. However, for all but two models (IPSL and HadGEM2), those curves become insignificant when compared to the 5-yr significance curves (gray). This finding contrasts with conclusions from LaJoie and DelSole (2016), in which the same analysis was performed, however as already noted, LaJoie and DelSole (2016) used EOFs instead of Laplacians. In that study, the divergence computed for some control runs were significant (CCCma, CNRM, IPSL, HadGEM2, and MPI), even after applying the thresholds derived from the 5-yr permutation method. Nonetheless, as noted by the authors, changes in control variability were considerably less than changes in the presence of anthropogenic forcing.

Having examined changes in the absence of anthropogenic forcing, and finding that there are some locally significant changes, it is worth comparing comparing figs. 3.1 and 3.4 and noting that changes in the presence of anthropogenic forcing are double or triple the changes that occur in the absence of anthropogenic forcing. However, this comparison is not perfect because the sample sizes differ. To explore this fully, the control simulations were sampled to mock the dimensions of (2.16) and (2.17) and the divergence of these mock covariance matrices was evaluated for each model. The resulting divergence as a function of accumulated Laplacians is plotted as the yellow-brown curve in each panel in fig. 3.3. Comparing this curve with the other divergence curves reveals that changes in the presence

of anthropogenic forcing (red and blue curves) are much greater than changes that occur naturally in the unforced climate system (yellow-brown curve), even when the sample sizes are the same. Also note that most of the yellow-brown curves lie below the gray curves in each panel, indicating that the unforced changes estimated from the smaller sample size in that model would still not be significant after accounting for autocorrelations.

3.1.3 Changes in Internal Variability: Sub-global Domains

The result of applying the union-intersection test to different geographic domains is shown in fig. 3.7. Each panel in the figure displays the leading discriminants for each model as a function of the number of Laplacians for the specified domain, where the domain is identified in the title of the panel. Most domain names are self explanatory, others names, such as NH, NALAND, and NASS, are shorthand for the Northern Hemisphere, the land area of the North American continent, and the sea surface area defined by the North Atlantic Ocean, respectively. For details on each domain, please refer to Table 3.1. Different domains are investigated to gain further insight into the source of changes in variance and to identify changes that may occur on sub-global scales, but not project on the global domain. Note that the global domain has been included and although these results have already been discussed (and shown), the panel is provided again for convenience. In short, the figure shows that there are no robustly significant leading discriminants in any domain for annual mean 2m temperature. There are also no significant trailing ratios (not shown). It is worth noting that one model shows significant ratios for nearly every truncation in each domain (MIROC5 dark green). These results suggest that while there are many locally significant changes in variance for each model, as computed by an F-test and displayed in fig. 3.1, significant changes to any single component of variability cannot be detected within the global or sub-global domains considered in this study.

Divergence was computed from the same domain-specific data and the results from the second half of the 21C are provided in fig. 3.8. The conclusions for the first halves are

generally consistent (not shown), with exceptions discussed shortly. The same domain-name conventions apply and the global domain results are repeated for convenience. While significant changes in variance for a global domain have already been discussed, this figure shows that there are also robust significant changes in variance detected in the ocean, NH, and NASS domains. For the remaining domains: land, NHLAND, and NALAND, the changes are significant for a majority of models. These results suggest that global-scale changes in the variability of annual mean 2m temperature are predominantly influenced by changes over ocean-surfaces. Changes in variance over land-only surfaces are more robust in the second half of the 21C, indicating that land-based changes get stronger in time as greenhouse gas concentrations increase.

3.2 Seasonal Mean 2m Temperature

The above sections considered changes in variability of annual means. Now changes in individual seasons is considered. The local changes in internal variability of seasonal mean 2m temperature are plotted in the following figures: JFM in fig. 3.9, AMJ in fig. 3.10, JAS in fig. 3.11, and OND in fig. 3.12. The seasonal mean changes are quantified by an F-test between twenty-first century residuals (relative to an ensemble mean) from an RCP8.5 emissions scenario and a preindustrial control. Insignificant values (at the 10% significance level) are not colored. Each season generally shows the same spatial pattern of change as was quantified by the annual mean (and displayed in fig. 3.1). For example, the robust annual mean decreases in variance found in regions of sea-ice formation appear to be largely explained by the changes in OND, JFM, and AMJ. Other regional and tropical ocean changes remain model-dependent, with the possible exception of summer (fig. 3.11), in which a majority of models indicate increases in either the tropical or extratropical latitudes.

Changes in the polar latitudes are model-dependent. HadGEM2, for instance, shows few to no significant changes in variance in its annual mean at polar latitudes, but the

seasonal-mean analysis reveals that there are large decreases in variance at polar latitudes in OND, followed by large increases in variance in JAS (see figs. 3.12 and 3.11, respectively, for the HadGEM2 panels). The lack of annual mean changes at polar latitudes for this model might reflect a cancellation between the opposing changes detected in OND and JAS. CNRM shows strong increases in annual mean variability for the northern-most latitudes and the seasonal mean analysis indicates that those changes are likely influenced by the strong changes in JAS; in fact, in JAS the variance doubles in some of those latitudes. Again, raising questions about the adequacy of doubling uncertainty estimates. CCCma shows many more widespread changes in each seasonal mean, however, the direction of these changes is seasonally-dependent as such, the annual mean F-map shows comparably fewer local changes in variance. The widespread increases in variance in the annual mean F-map for IPSL can be linked to the increases in local variance in JFM. Both MIROC5 and MPI project increases in variance in tropical and subtropical latitudes. This dominant structure appears in each season for both models. Differences in the direction of change occur in the northern-most latitudes for the transitional seasons. For NCAR, the annual mean structure is consistent with each season, however, there are significant decreases in variance which appear in AMJ and JAS that are not evident in this model's annual mean variance ratio map.

The union-intersection test was applied to these seasonal-mean data. There are no robustly significant leading or trailing discriminants for any season or domain for 2m temperature. Figure 3.13 is provided as a representative example of the results that led to this conclusion. One model is again consistently significant (MIROC5 dark green). Overall, however, these results indicate that significant 21C changes in variance in any single component of variability cannot be identified in the seasonal mean 2m temperature data.

Divergence was computed for the seasonal mean data and the results for each domain are given in separate, season-specific figures: JFM in fig. 3.14, AMJ in fig. 3.15, JAS in fig. 3.16, and OND in fig.3.17. In each seasonal figure, the models robustly agree on significant changes in variance in the global and ocean domains. The models also robustly agree on

significant changes in variance in the NH in every season, except JAS, in which case only the second half of the century is robustly significant. The other domains appear to exhibit seasonally-dependent changes. For example, in the land-only related domains, the JAS and OND seasons show more robust changes when compared to JFM and AMJ for the same domains.

3.3 Annual Mean Precipitation

The previous sections examined changes in temperature variability due to anthropogenic forcing. This section examines changes in precipitation variability.

3.3.1 Changes in Internal Variability Due to Anthropogenic Forcing

The local changes in the internal variability of annual mean precipitation between twenty-first century residuals (relative to an ensemble mean) from an RCP8.5 emissions scenario and a preindustrial control run are shown in fig. 3.18. The changes are quantified by an F-test and insignificant values (at the 10% significance level) are not colored. Along with the model name, in the title of each F-map is the percentage of significant grid points. Again, grid points where the ratios are greater than 1 (warm colors) indicate increases in response to anthropogenic forcing and ratios less than 1 (cool colors) indicate anthropogenic forcing decreases variability.

All the models project locally significant changes in variance. A general comment about the changes displayed in fig. 3.18: there appears to be robust agreement that the changes in variance are in the positive direction for most of the globe, in other words, the internal variability of annual mean precipitation is projected to increase in the 21C. Consistently at the polar latitudes, the models project significant increases in variance. A majority of the models project significant increases in the tropical Pacific Ocean, as well. For the most part, the models show that middle-latitudes are also projected to see increases in the variability of annual mean precipitation. While the results primarily suggest that significant increases in variance are widely projected for the 21C, there are some equatorial latitudes in the

North Atlantic Ocean in which a majority of models project decreases in the variability of annual mean precipitation. The direction of other changes, like those in the Sahel of Africa, are model dependent.

To determine if these changes could be explained by changes in a single (or a few) components of climate variability, the union-intersection test was applied to the annual mean precipitation data. The leading discriminants for each model as a function of the number of Laplacians, are displayed for specified domains in fig. 3.19. For the global domain, the leading discriminants are consistently significant after $T = 35$ (marked by vertical red line), for 5 out of 7 models. These results indicate that a single component of variability significantly changes variance. The trailing ratios were not significant (not shown). Recall that weighting vectors $\mathbf{q}$ can be used to construct a time series and a spatial map that help visualize the changes detected in that component. When the leading ratios are deemed not sensitive to truncation, it turns out that the leading *variates* are also not sensitive to truncation.

To demonstrate this finding, fig. 3.20 shows a representative set of leading variates from one ensemble member for one model at $T = 35$, however, the individual variates are difficult to see because they are virtually identical. This finding was true for the other ensemble members (not shown) and for all the ensemble members in the other models as well (also not shown). This figure is provided to illustrate that the variates are not sensitive to choosing truncation $T = 35$. Shown in fig. 3.21 is a segment of the corresponding time series for both the significant control variate (blue curve) and the significant 21C variate (one ensemble member; red curve) at $T = 35$. This figure is provided to demonstrate that variance appears to increase with time, i.e., the variance of the 21C (red curve) is greater than the variance of the control (blue curve). This finding was also true for the same significant component identified in the other models (not shown).

Recall that discriminant analysis finds a linear combination of time series that maximize (or minimize) a variance ratio. The resulting time series, called the variates (eq. (2.25)), can then be used to construct regression patterns to describe the spatial structure of the

component. As shown above, the variates are not sensitive to the number of Laplacian eigenvectors, hence the regression pattern derived from the variates is also insensitive to truncation. The regression pattern for the component that increases in variance over the global domain is displayed for each model in fig. 3.22. Note that the phrase 'Not Significant' appears on the map of two panels, identifying that the conclusions for that model were sensitive to truncation (see CNRM and NCAR). In general, the spatial patterns are confined to the tropical Pacific Ocean, with some expansion to the subtropical and middle latitudes. These results suggest that on a global-scale, a significant change in the variability of annual mean precipitation may occur in the tropical Pacific Ocean. Based on these findings, it may not be surprising that the leading discriminants are also significant in the ocean-only domain (see OCEAN panel in fig. 3.19). The leading pattern that explains the significant component of change in the ocean domain (at $T = 30$) is displayed for each model in fig. 3.23. Note that conclusions were sensitive to truncation for the same two models (CNRM and NCAR). The models generally agree on a dipole of changes in the western portion of the equatorial and subtropical Pacific Ocean. A few models show changes extend toward the middle latitudes. Other changes, like those in the Indian Ocean, are model-dependent.

Divergence was computed for the annual mean precipitation data and those results are shown in fig. 3.24. Of particular notice, the robustly significant results for the NASS domain, indicating that several components are significantly changing their variance in that area. The divergence results also indicate that significant changes in variance can be detected in the NALAND domain. In the land, NH, and NHLAND domains, all but one or two models indicate significant changes in variance. In all the domains, the changes in the second half of the 21C are robust among models.

3.3.2 Changes in Control Variability

As previously discussed, studies have shown that significant changes in the variability of annual mean temperature can occur in the absence of anthropogenic forcing. To determine if this may also be true in the annual mean precipitation data, the 500-yr control simulations

were partitioned into equal halves and the local differences in variance were quantified with an F-test. The resulting F-ratios are displayed for each model's precipitation control simulation in fig. 3.25. The percentage of significant grid points is given in the title of each F-map and, speaking broadly, those percentages are close to what one might expect due to random chance (about 10% of the total area). Also, the spatial distribution of the significant grid points lacks structure, i.e., they appear to be randomly scattered throughout the domain. However, there are two models (MIROC5 and MPI) which show that about 20% of their total area is deemed to be significantly changing variance and the changes appear to be concentrated in the tropical Pacific Ocean.

To determine if these changes could be explained by a single component of climate variability the union-intersection test was performed on these data. Figure 3.26 shows the leading and trailing discriminant ratios for the global domain. It is interesting to note that MIROC5 (dark green) exhibits significant trailing ratios for all truncations, indicating that the tropical-Pacific "zone" of decreasing variance displayed in it's F-map can likely be explained by changes in a single component of climate variability in this model. Similarly, MPI (dark blue) exhibits significant leading ratios for all truncations beyond 1. These results indicate that the tropical-Pacific "zone" of increasing variance displayed in MPI's F-map, could likely be explained by changes in a single component of climate variability present in that model. The union-intersection results imply that for these two models, the local changes displayed in the F-map are likely explained by changes in a single component of variability, however, while these results are noteworthy and demonstrate a straight-forward result from applying the union-intersection test, these findings are not robust among the models, as such, the corresponding spatial patterns that explain these significant ratios are not shown. Divergence was also applied to the control discriminants and those results are shown in 3.27. With the exception of MIROC5 and MPI, the other models indicate variance in the control simulations is not significantly changing in an aggregate sense (although single isolated components may experience significant changes).

3.4 Seasonal Mean Precipitation

The local changes in the internal variability of seasonal mean precipitation between twenty-first century residuals (relative to an ensemble mean) from an RCP8.5 emissions scenario and a preindustrial control run are shown in the following figures: JFM in fig. 3.28, AMJ in fig.3.29, JAS in fig. 3.30, and OND in fig. 3.31. The local changes are quantified by an F-test and insignificant values (at the 10% significance level) are not colored. Each season, respectively, shows generally the same spatial pattern of changes as their annual mean counterparts (displayed in fig. 3.18). For example, the robust increases in variance for middle and polar latitudes are evident in every season. Some differences include that in summer, fig. 3.30, HadGEM2 displays tropical Pacific changes that are in a different direction than its annual mean counterpart. Also, generally speaking across models, the changes in AMJ are weaker compared to the annual mean (but still significant). In OND, fig. 3.31, the general pattern of increased variance at polar latitudes is most pronounced. Also, in contrast to the annual mean, in OND most of the models *agree* about the direction of change along the tropical Pacific; suggesting increases in variance are projected for precipitation in that region for autumn in a warmer climate.

Other differences between the annual and seasonal means, can be found in the tropics and extratropics. For instance in JFM, fig. 3.28, there are comparably more extratropical latitudes showing decreases in variability, especially in the lower section of the North Atlantic Ocean, in which case, the decreases are robust across models. Similar decreases are also found in AMJ for 5 out of 7 models (fig. 3.29).

To determine if these changes could be explained by changes in a single (or a few) component(s) of climate variability, the union-intersection test was performed on these data. In two cases, the leading discriminants were significant. Figure 3.32 shows the leading discriminants as a function of the number of Laplacians for JAS in the global domain and JFM in the land-only domain. The trailing ratios were not significant (not shown). In the global domain for JAS, the ratios were significant for 5 out of 7 models after $T = 15$ (marked

by a vertical red line). In the land-only domain for JFM, the ratios were significant for 6 out of 7 models after $T = 50$ (marked by a vertical red line).

The time series associated with the leading variates for the significant components indicate that variance increases with time (not shown). The leading pattern that explains the significant component of change in the global domain for JAS is displayed for each model in fig. 3.33. Note that the phrase 'Not Significant' appears on the map of two panels, identifying that the conclusions for that model were sensitive to truncation (see CNRM and NCAR). In general, the spatial patterns are confined to the tropical Pacific Ocean, with some expansion to the subtropical and middle latitudes. These results suggest that significant changes in the variability of JAS mean precipitation may occur in the tropical Pacific Ocean in the 21C, which supports conclusions from the annual mean study that were displayed in fig. 3.22.

The leading pattern that explains the significant component of change in the land-only domain for JFM is displayed for each model in fig. 3.34. Note that conclusions were sensitive to truncation for only one model, HadGEM2. In general, the spatial patterns are confined to the subtropical latitudes in both hemispheres, suggesting that significant changes in the variability of JFM mean precipitation over subtropical land-areas are projected for the 21C in these models.

Divergence was also computed for the seasonal mean precipitation data. Some results were consistent not only across models, but also across season and domain. For instance, the NASS domain was significant for every season; suggesting the models robustly agree on 21C changes in the variability of seasonal mean precipitation over the North Atlantic Ocean. For practical reasons, these results are demonstrated for only two seasons, JFM and OND; the results are similar for AMJ and JAS, but not shown. For JFM, JAS, and OND, the changes in the NHLAND domain were also robustly significant. In addition, in the JFM season, every domain was significant for all 2nd-halves, and significant for at least 5 out of 7 models in all the 1st-halves (see fig. 3.35). For OND, fig. 3.36, changes were robust in the following domains: ocean, land, NHLAND, and NASS (as previously noted).

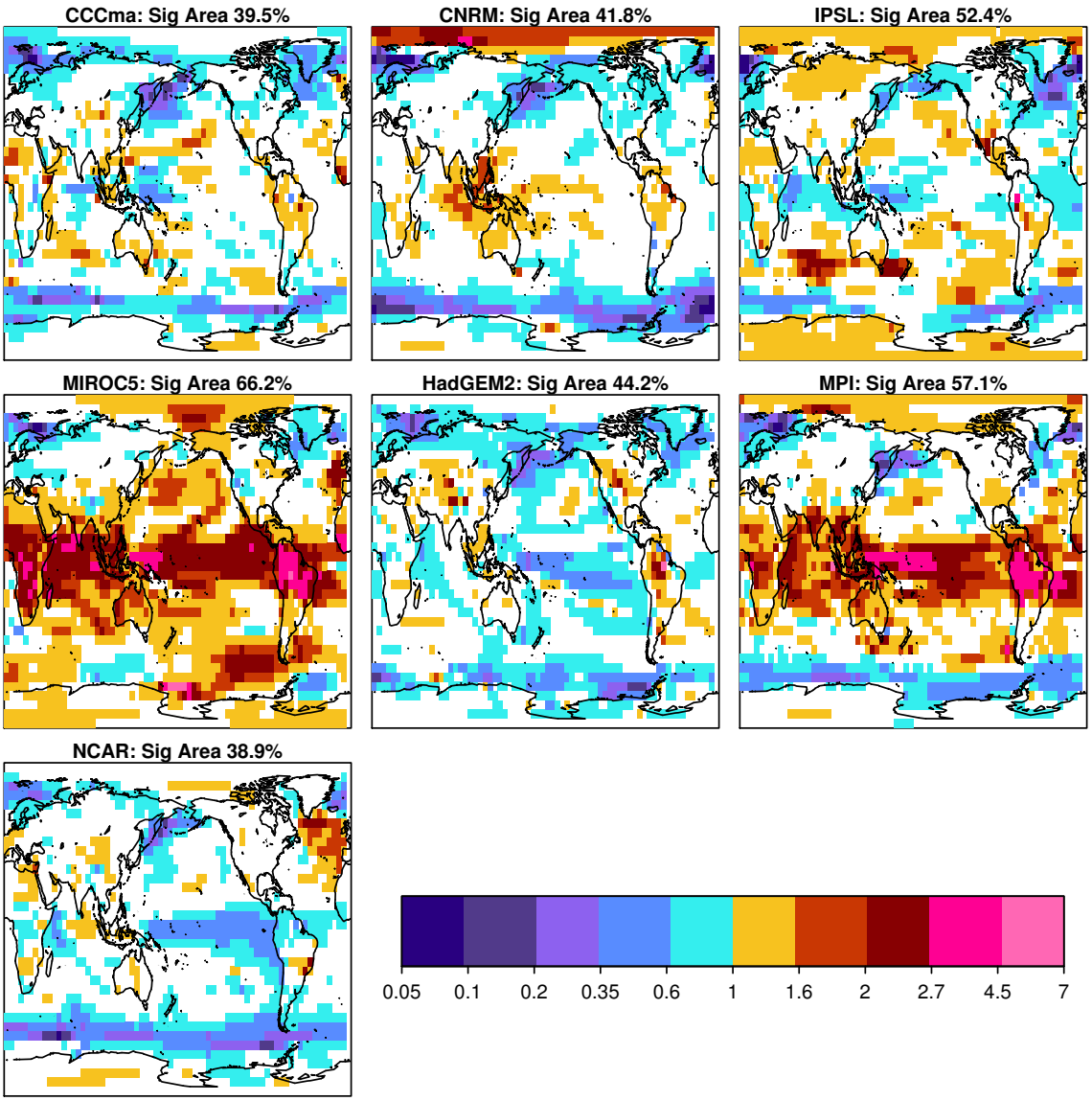

Figure 3.1: Change in internal variability of annual mean 2m temperature due to anthropogenic forcing, as quantified by the local ratio of variance in the 21st century (21C) to preindustrial control internal variability in each model (via eq. (2.8)). The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. The variance of internal variability for preindustrial forcing is computed from the same model's 500-year control simulation. A ratio larger than one indicates internal variability increases in the 21C. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out (i.e., not colored). Total percent of significant grid points is listed in the title along with the model name.

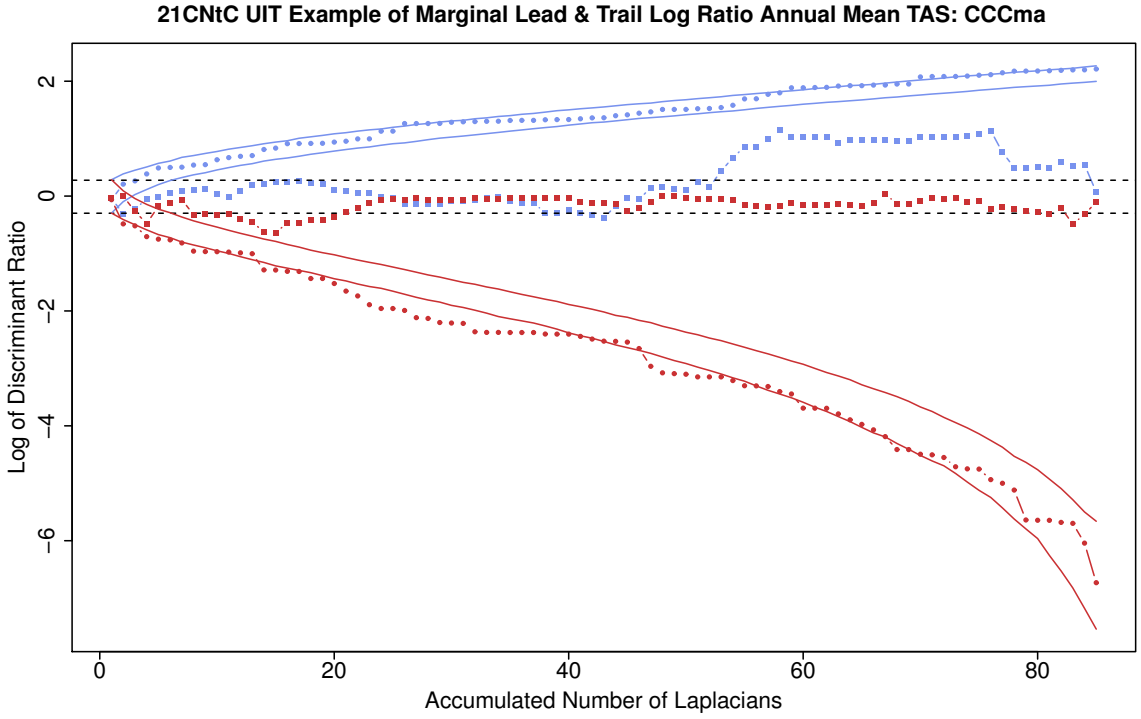

Figure 3.2: An example of results from applying the union-intersection test to the discriminant 21C-noise to control ratios for the annual mean 2m temperature (TAS) data set is shown for one model: CCCma. The log of the leading discriminant ratios, as a function of accumulated Laplacians, are given by the blue circle-dash curve. The analysis was repeated in an independent data set and the results for the independent leading ratios are given by the blue square-dash curve. The log of the trailing discriminant ratios are given by the red circle-dash curve and the trailing ratios from the independent analysis are given by the red square-dash curve. Also shown are the upper and lower 5% significance thresholds computed from Monte Carlo techniques (blue and red, respectively) and the significance thresholds derived from an F-distribution (straight dashed black lines). The results appear to be sensitive to truncation.

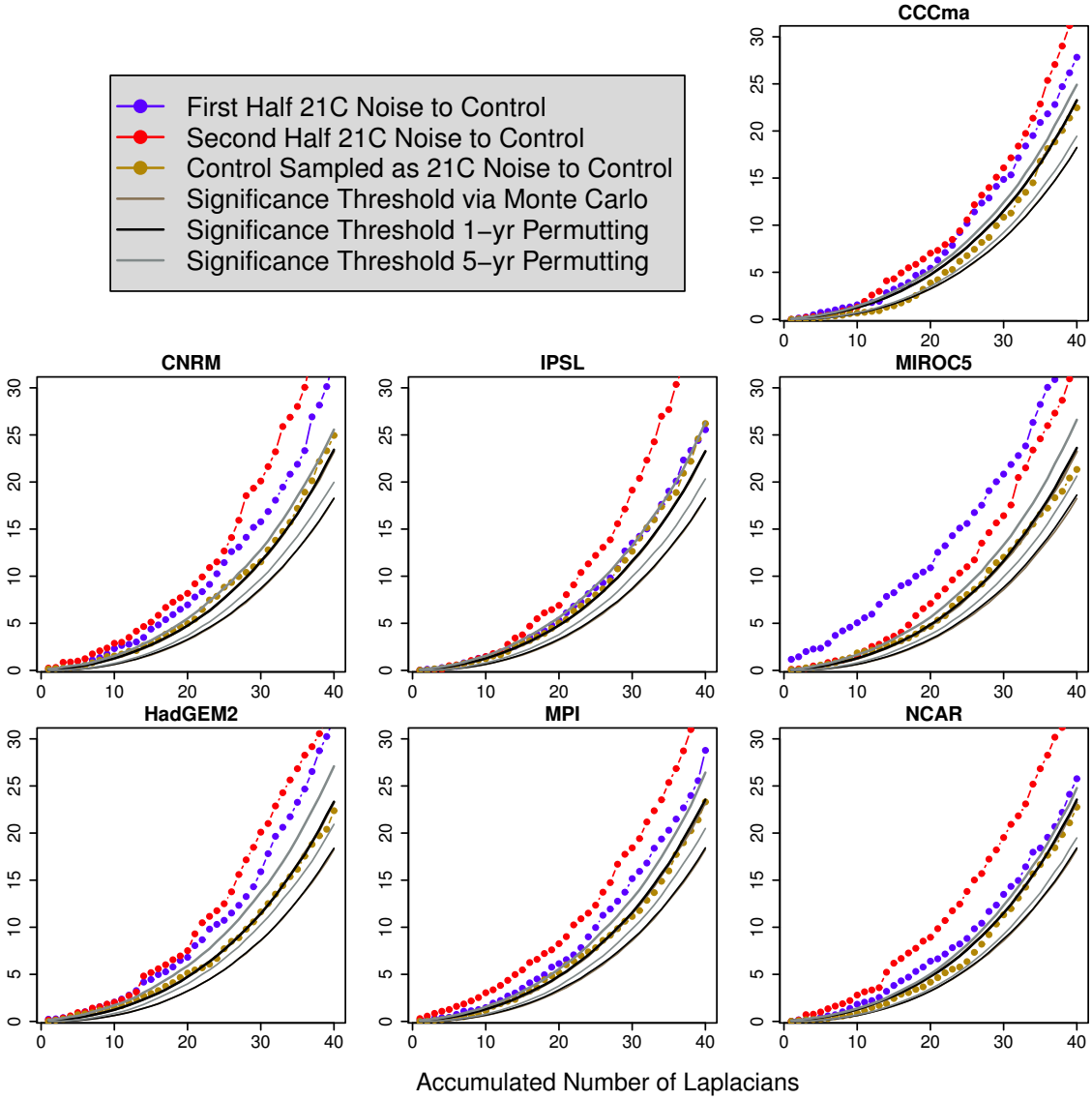

Figure 3.3: The divergence D_T of internal variability of annual mean 2m temperature between 21st century and preindustrial control simulations, as a function of the number of Laplacians included in the computation. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by blue curves and the remaining halves are given by red curves. The yellow-brown curves show the divergence within a climate model's preindustrial control simulation, in particular, with dimensions that match those of the other divergences. Also shown are the upper and lower 5% significance threshold computed from Monte Carlo techniques (black curve) and permutation techniques; 1-yr (gray curve) and 5-yr (brown curve). Significant results lie outside the solid curves.

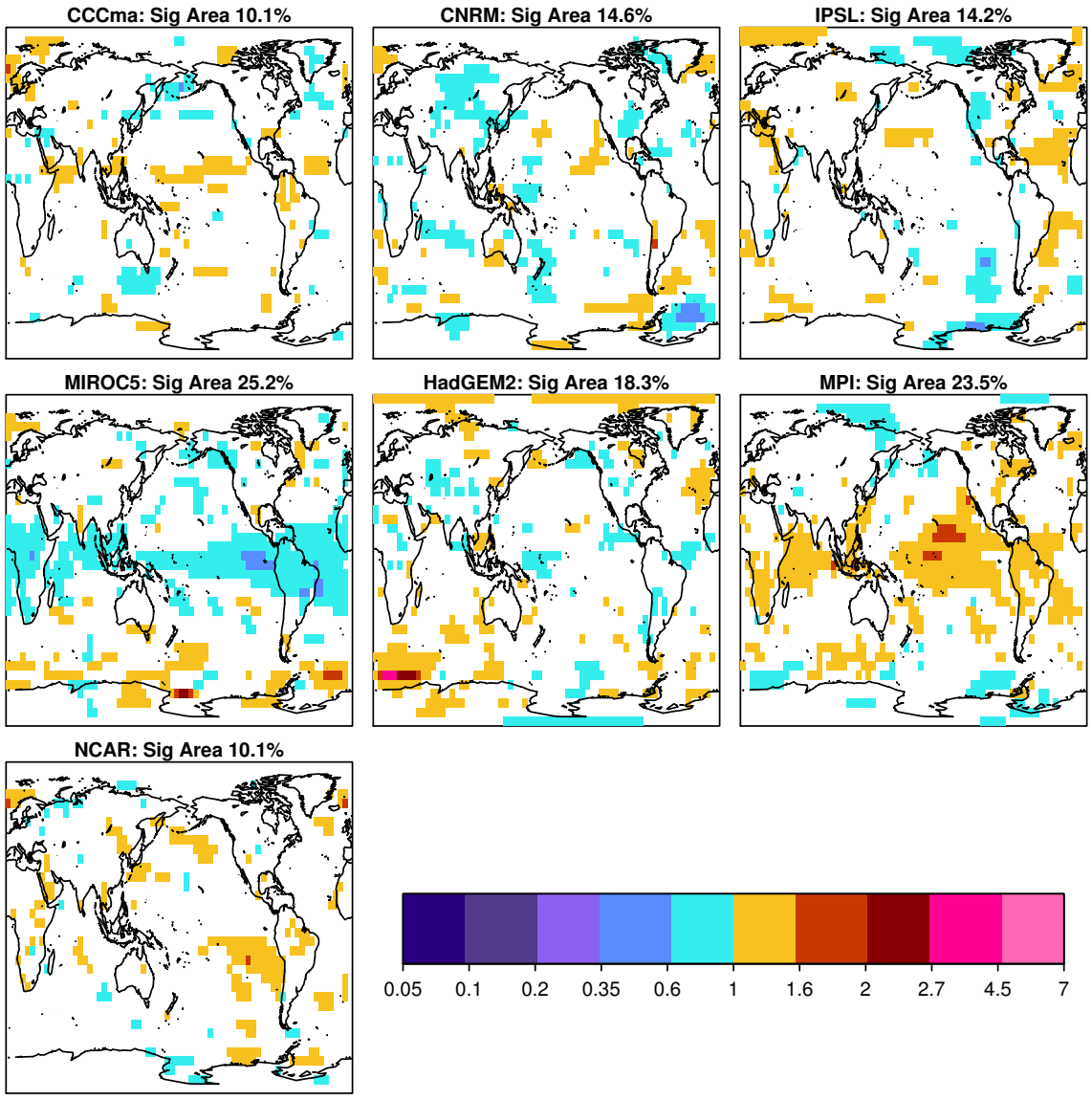

Figure 3.4: Local changes in internal variability of annual mean 2m temperature between two non-overlapping, 250-year segments of each model's preindustrial control simulation. Insignificant values (according to an F-test distribution at a 10% significance level) are masked out (i.e., not colored). Total percent of significant grid points is listed in the title along with the model name. Changes in variance occur mostly in the tropical oceans.

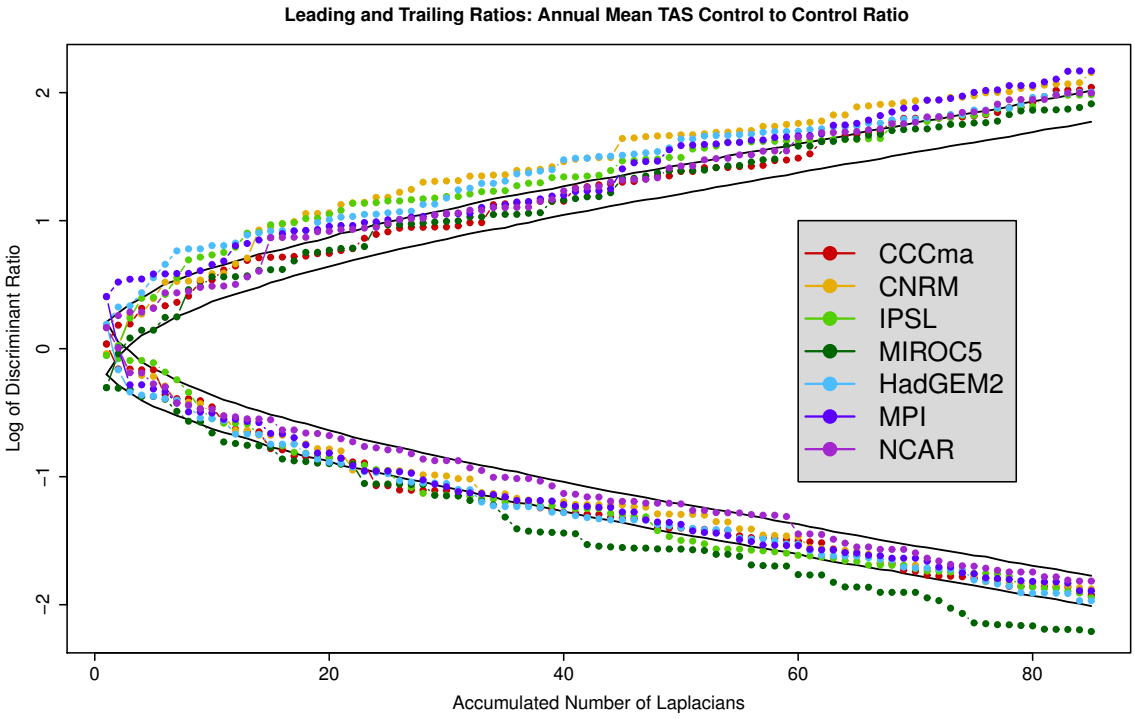

Figure 3.5: Log of the leading and trailing discriminant ratios from the union-intersection test on control to control ratios, as a function of the accumulated number of Laplacians. Note the curves are color-coded by model (see legend). Also shown are the upper and lower 5% significance thresholds computed from Monte Carlo techniques (solid black curve). The significance of the leading discriminants is sensitive to truncation and the trailing ratios are generally not significant, with the exception of MIROC5 (dark green). These results indicate that, for the most part, changes in any single component of variability is not significant.

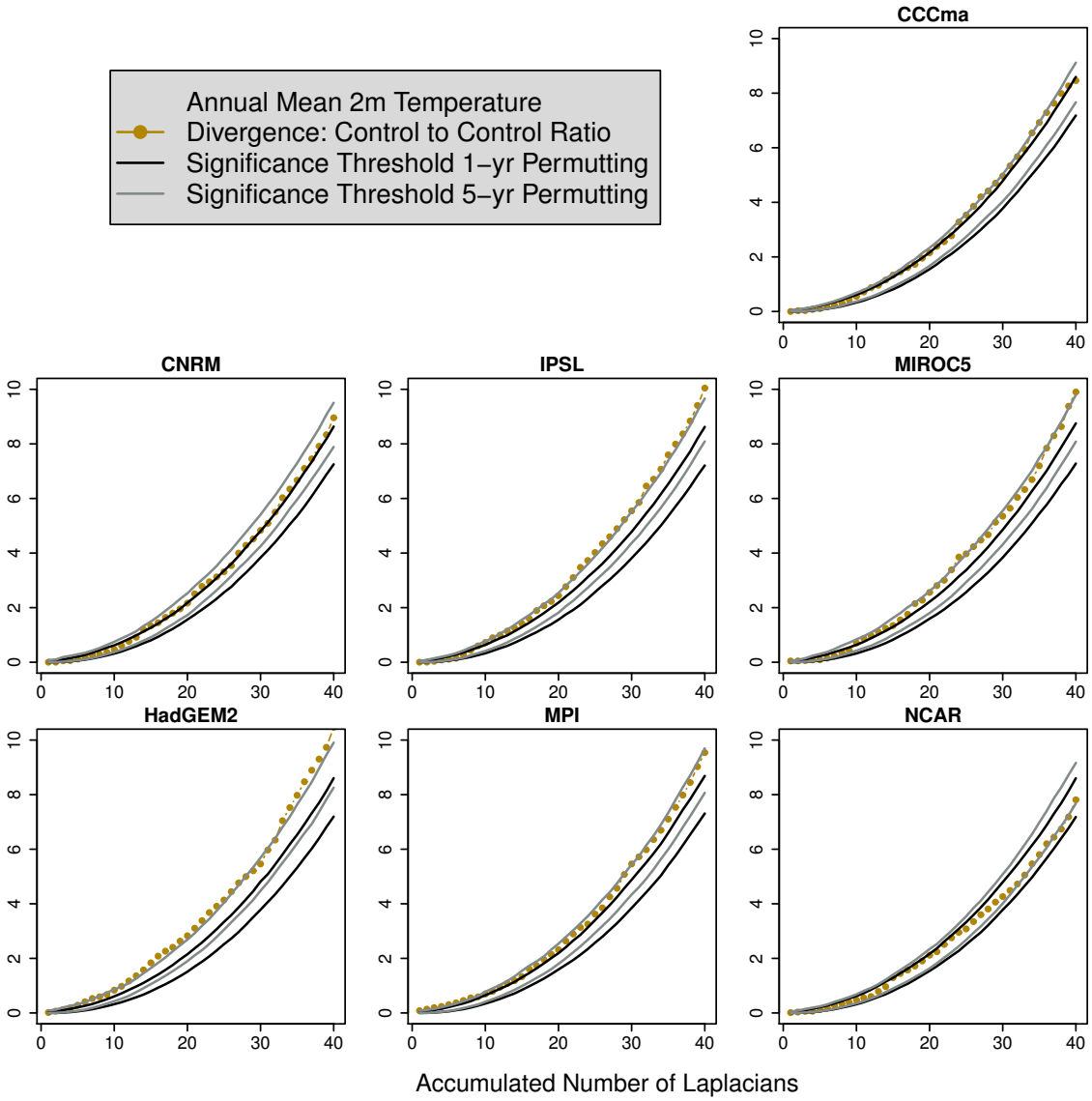

Figure 3.6: The divergence D_T of internal variability within each model's preindustrial control simulation as a function of the number of Laplacians included in the computation. The divergence is given by the yellow-brown curves. Also shown are the upper and lower 5% significance threshold computed from permutation techniques; 1-yr (black curve) and 5-yr (gray curve). Notice that after accounting for 5-yr autocorrelations (gray curves) the results are generally not significant. Also notice the comparably smaller range of the y-axis, indicating that changes in variance in the absence of anthropogenic forcing are considerably less than in the presence of anthropogenic forcing.

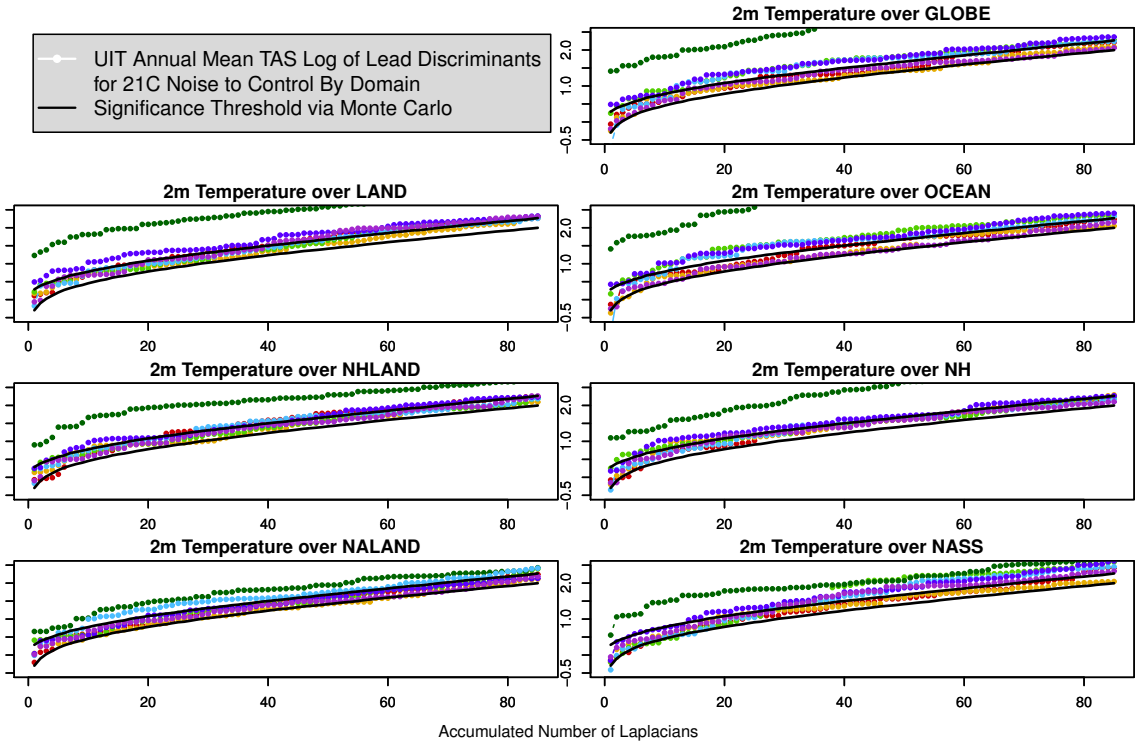

Figure 3.7: Shown by domain, indicated in the title of the individual panels, is the log of the leading discriminant ratios for internal variability of annual mean 2m temperature between the first half of the 21st century and the first half of the preindustrial control simulations, as a function of accumulated number of Laplacians. Note the curves are color-coded for each model (see legend inset fig. 3.5). Also shown are the upper and lower 5% significance thresholds computed from Monte Carlo techniques (black curves). For the domain specified in the title, only the grid points in that domain were considered in the computation. For example, in the panel marked 'LAND', only the land-based grid points were used to compute the covariance matrices. The results indicate that in each domain, the significance of the leading discriminants is sensitive to truncation, with the exception of one model, MIROC5 (dark green), which appears to be significant in all domains and for every truncation. The trailing discriminants are not significant (not shown). These results indicate that changes in variability in any single component of variability are not significant.

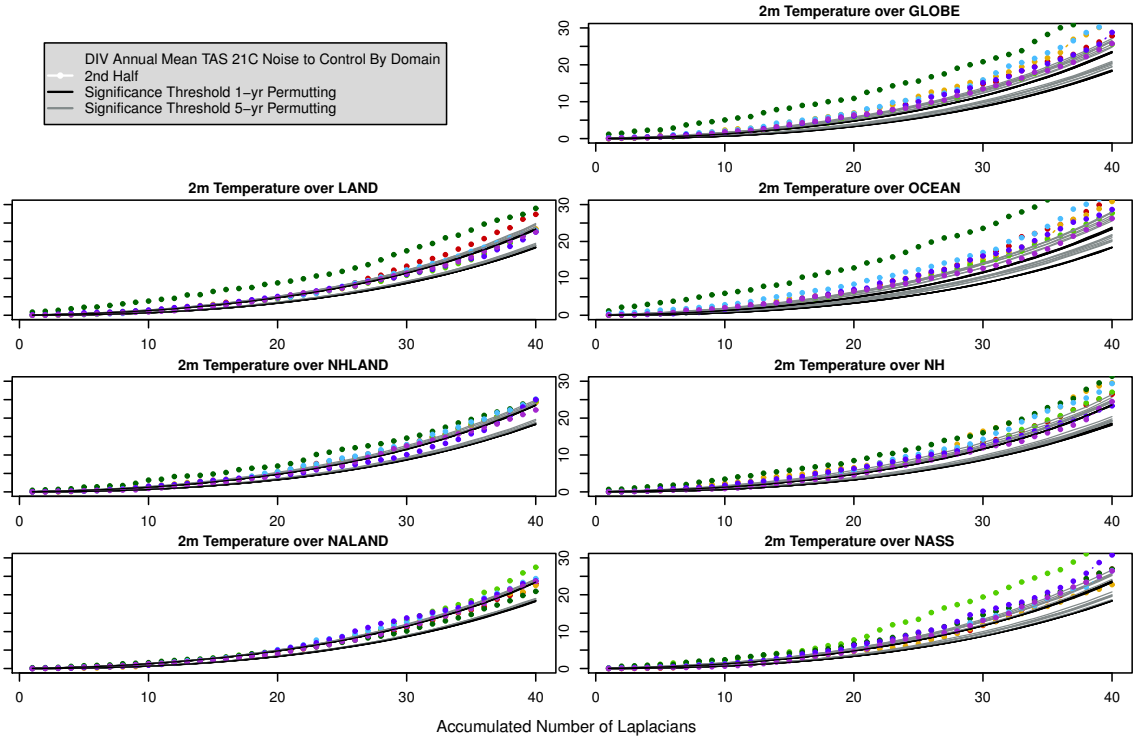

Figure 3.8: Shown by domain, indicated in the title of the individual panels, is the divergence D_T of internal variability of annual mean 2m temperature between 21st century residuals and preindustrial control simulations, as a function of the number of accumulated Laplacians. The divergence for the second half of the 21st century and second half of the preindustrial control run are given by the colored square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.5. Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). There are significant changes in variance in each domain, indicating that variability significantly changes in an aggregate sense.

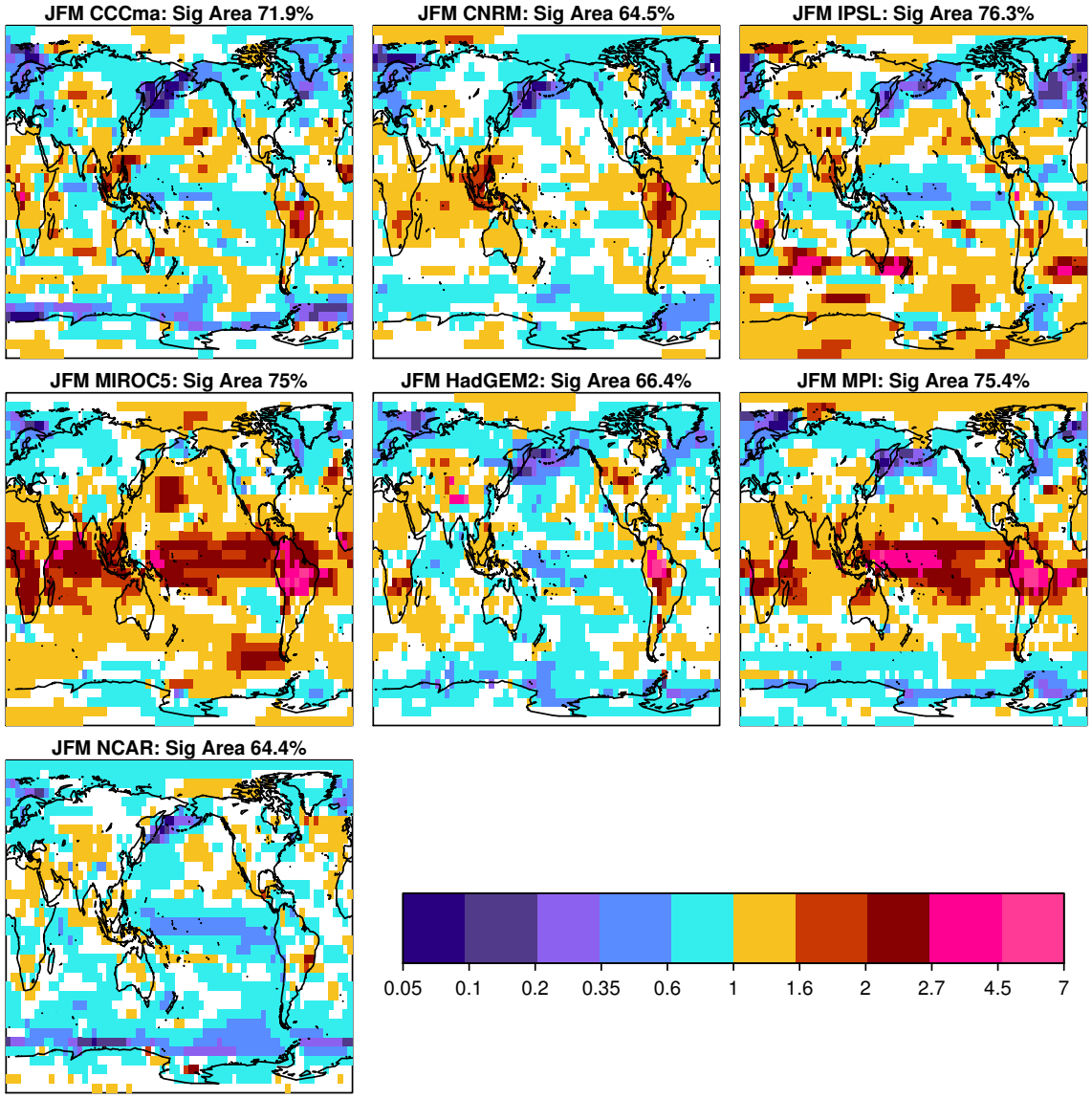

Figure 3.9: The local changes in internal variability of January-February-March (JFM) mean 2m temperature due to anthropogenic forcing and as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel.

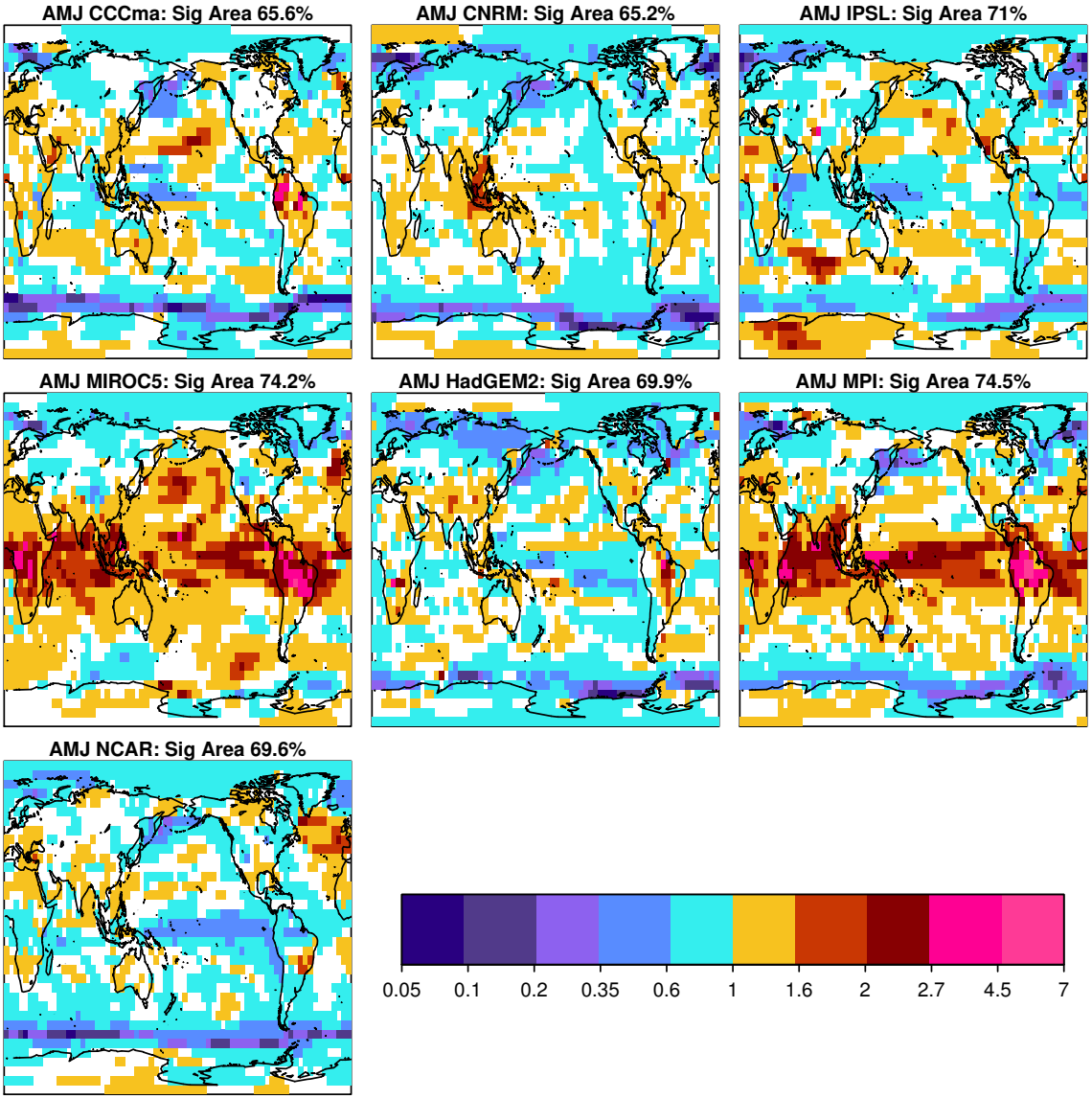

Figure 3.10: The local changes in internal variability of April-May-June (AMJ) mean 2m temperature due to anthropogenic forcing and as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel.

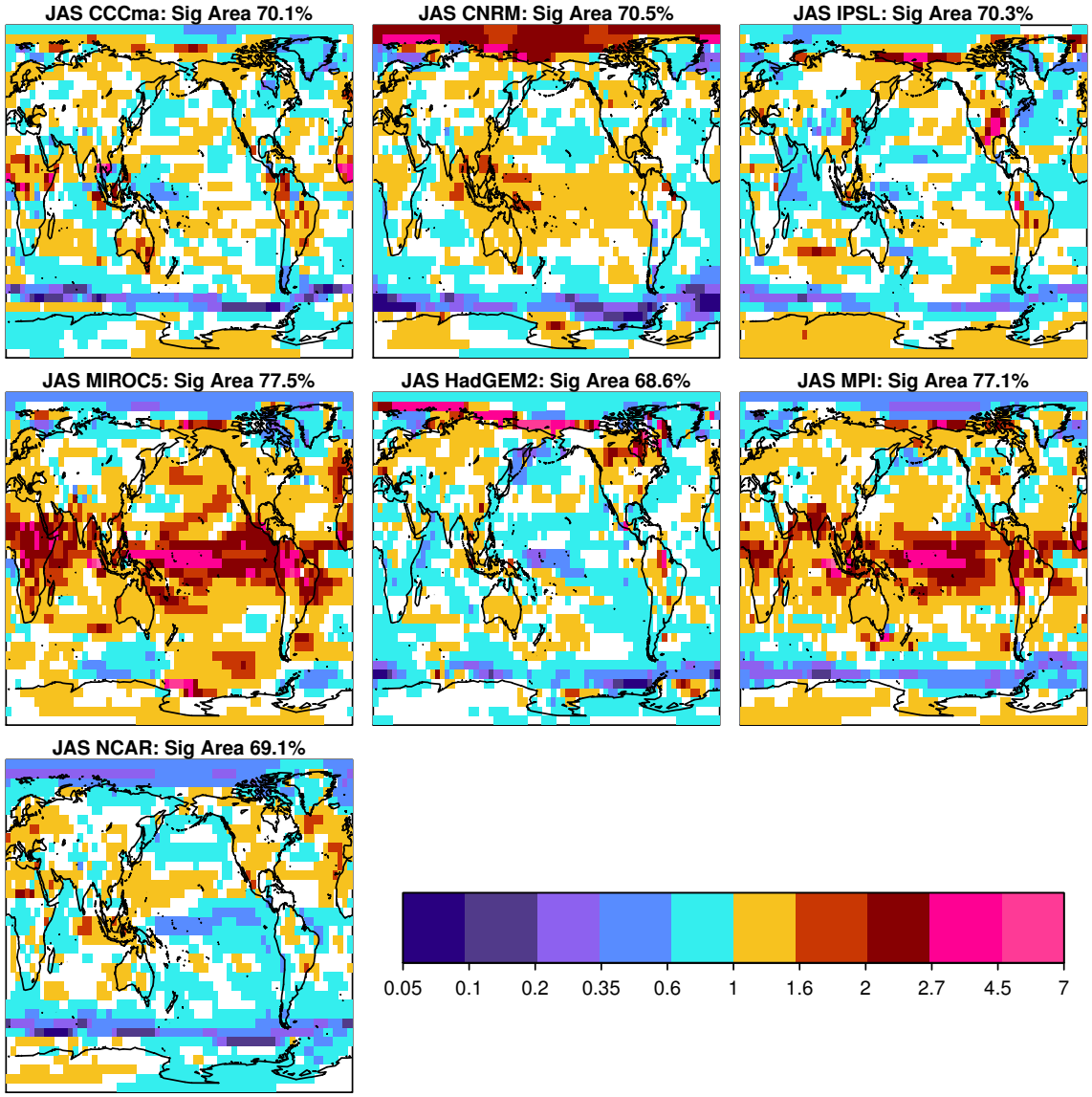

Figure 3.11: The local changes in internal variability of July-August-September (JAS) mean 2m temperature due to anthropogenic forcing and as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel.

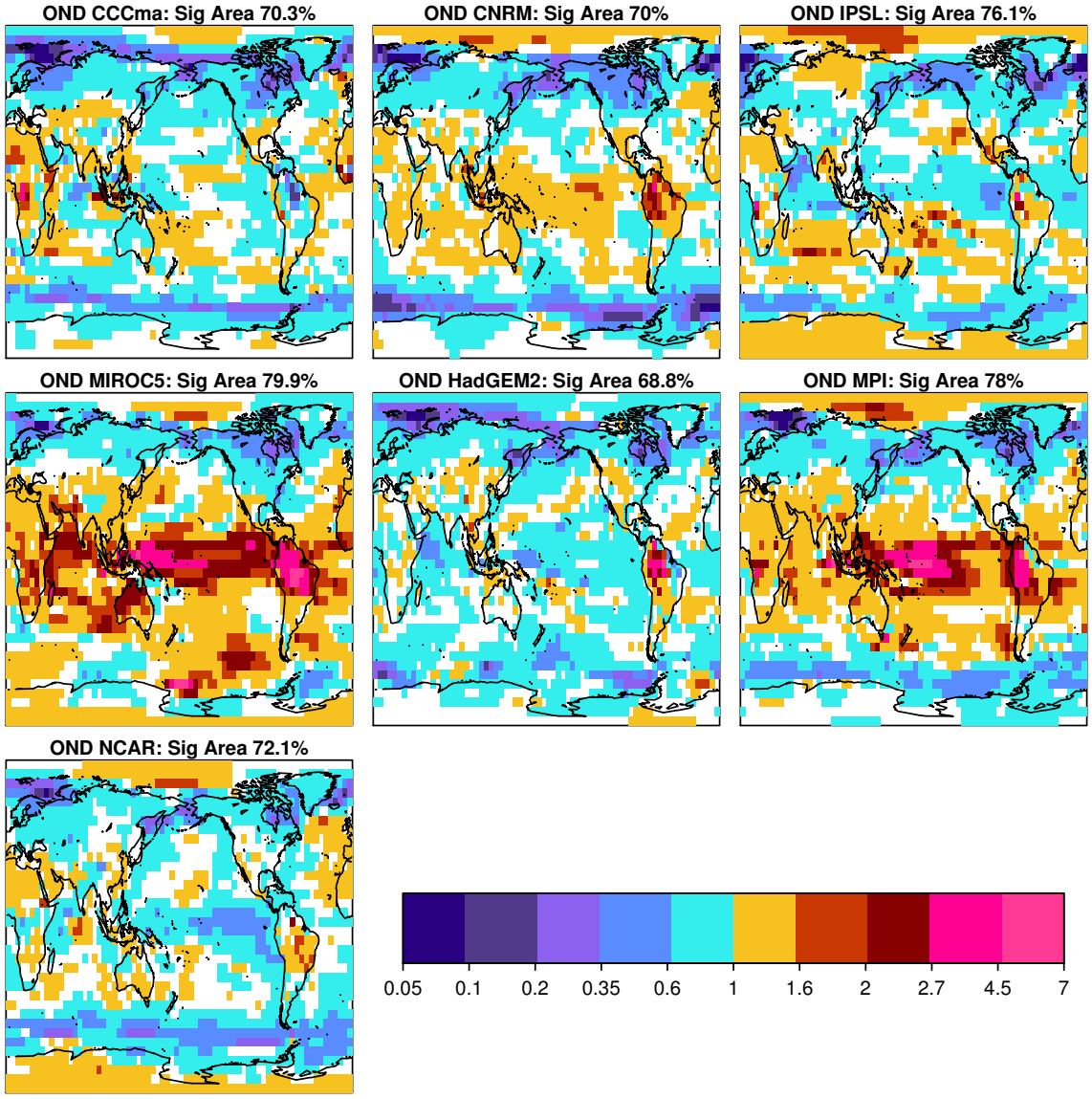

Figure 3.12: The local changes in internal variability of October-November-December (OND) mean 2m temperature due to anthropogenic forcing and as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel..

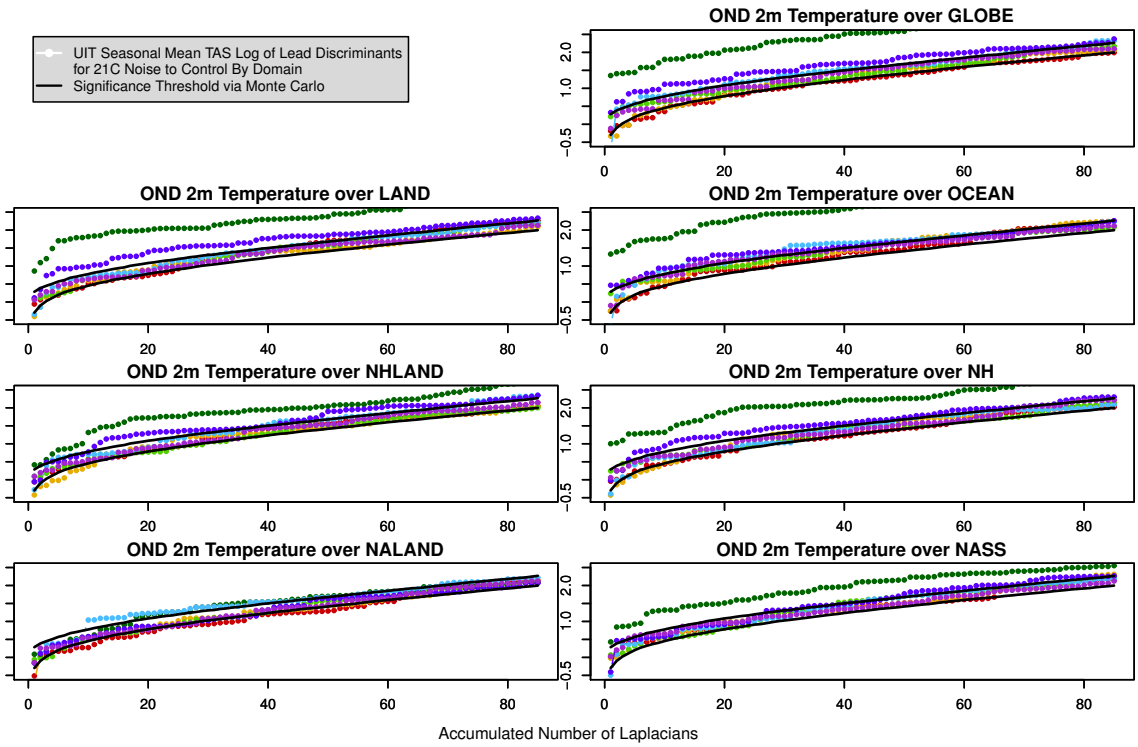

Figure 3.13: Shown by domain, indicated in the title of the individual panels, is the log of the leading discriminant ratios for internal variability of annual mean 2m temperature between 21st century residuals and preindustrial control simulations as a function of the accumulated number of Laplacians for the October-November-December mean (OND). Note the curves are color-coded by model (see legend inset fig. 3.5). Also shown are the upper and lower 5% significance thresholds (for a 10% significance level) computed from Monte Carlo techniques (solid black curves). The significance of the leading discriminants is sensitive to truncation for all the models except MIROC5 (dark green), which appears to be significant in all domains except NALAND. The trailing discriminants are not significant (not shown). These results are shown in order to provide a representative sample of the results found in each season and domain after applying the union-intersection test to the discriminants derived from the seasonal mean 2m temperature data. This figure exemplifies the conclusion that the variability of any single component is not significantly changing variance in any season.

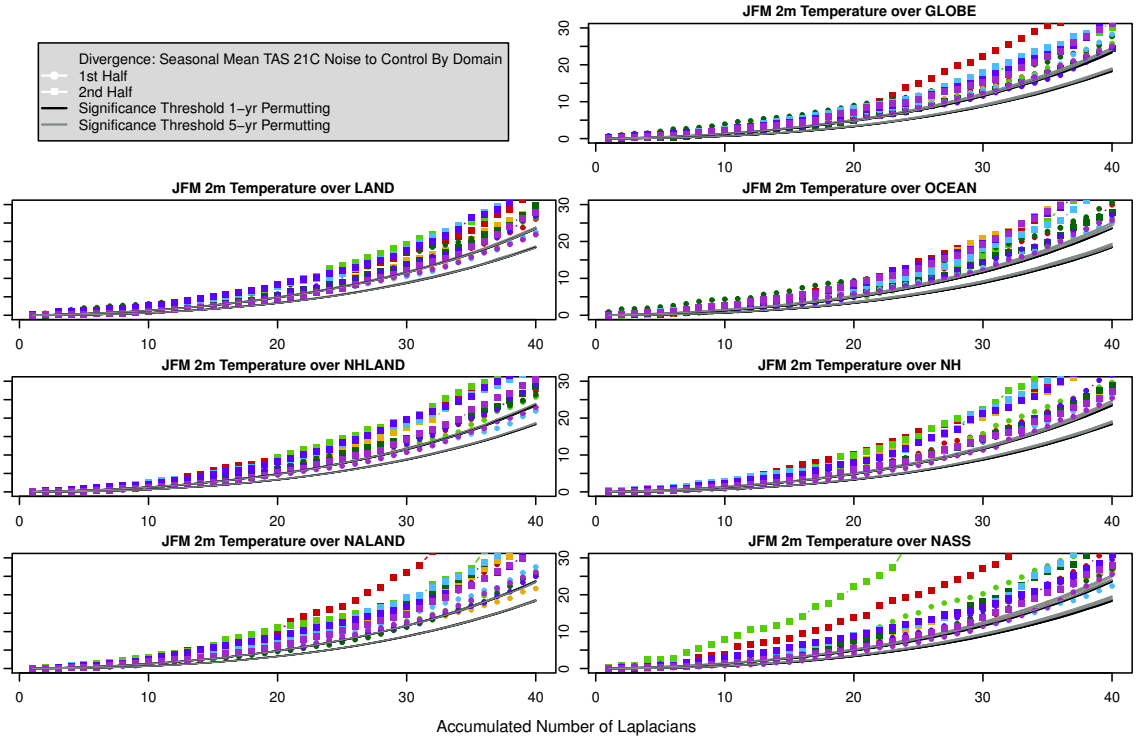

Figure 3.14: Shown for domain, indicated in the title of each panel, is the divergence D_T of internal variability of JFM mean 2m temperature between 21st century residuals and preindustrial control simulations as a function of accumulated Laplacians. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by circle-dash curves and the remaining halves are given by square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.5). Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). There are significant changes in variance in each domain, indicating that several components significantly change variance in each domain for the JFM season.

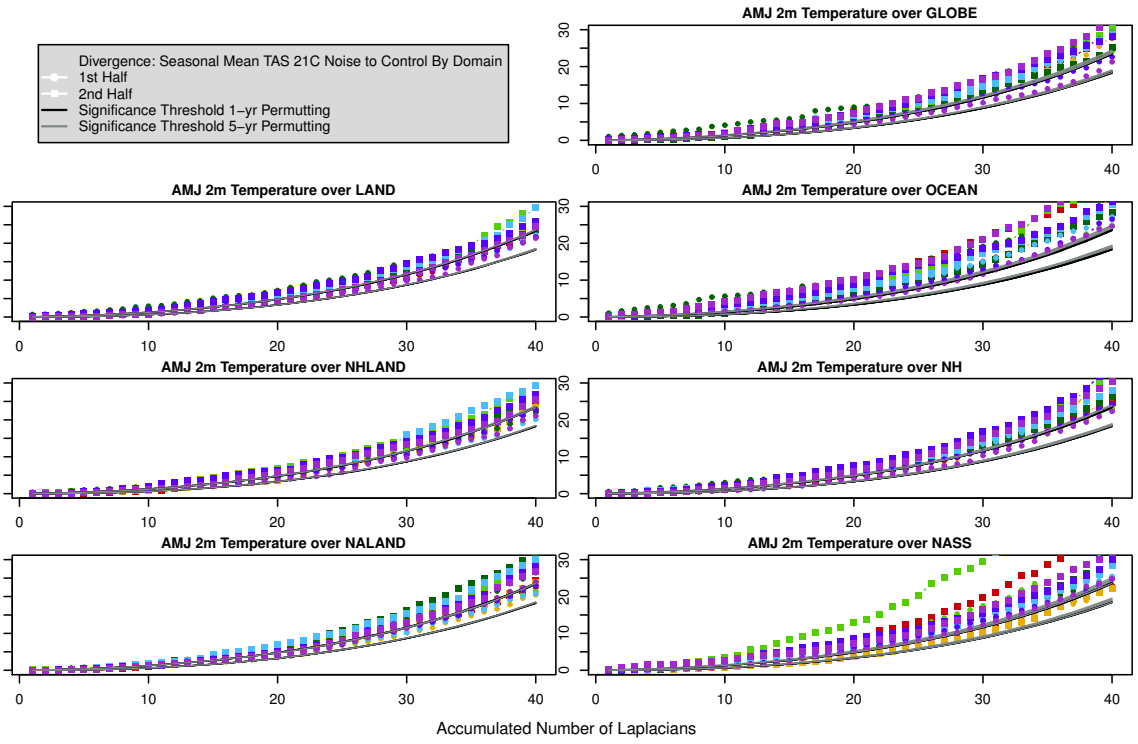

Figure 3.15: Shown for domain, indicated in the title of each panel, is the divergence D_T of internal variability of AMJ mean 2m temperature between 21st century residuals and preindustrial control simulations, as a function of accumulated Laplacians. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by circle-dash curve and the remaining halves are given by square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.5). Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). There are significant changes in variance in the global, ocean, NH, and NASS domains, indicating that several components significantly change variance in those domains for the AMJ season.

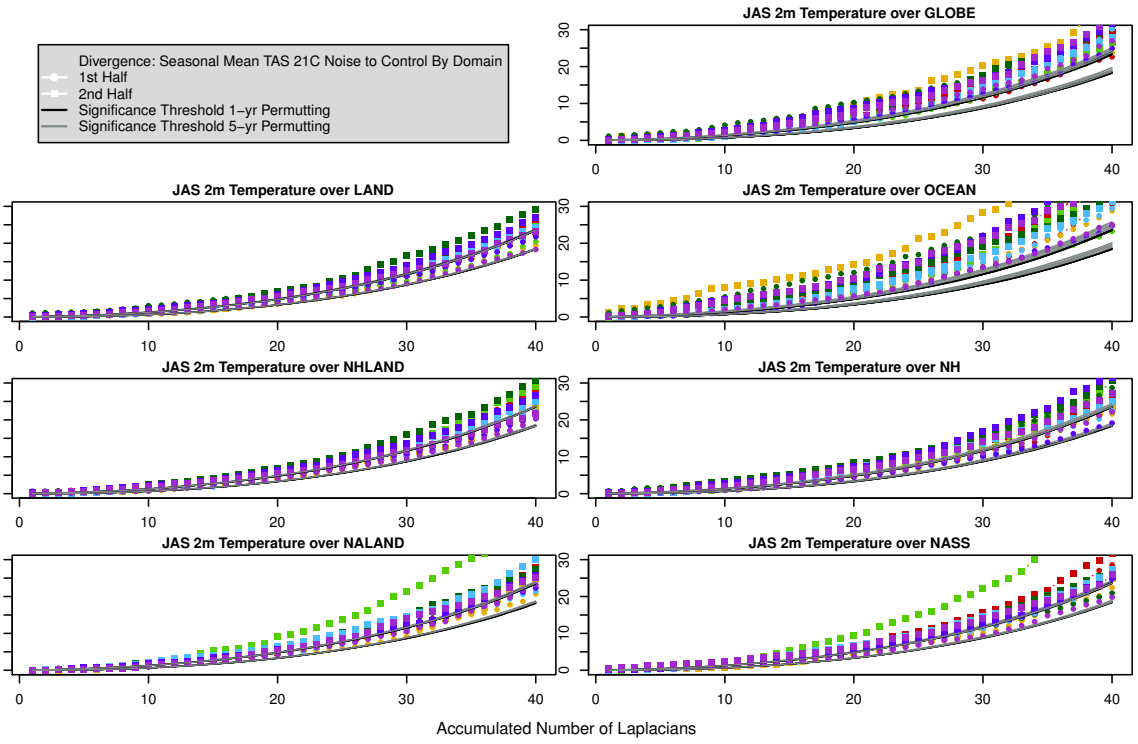

Figure 3.16: Shown for domain, indicated in the title of each panel, is the divergence D_T of internal variability of JAS mean 2m temperature between 21st century residuals and preindustrial control simulations, as a function of accumulated Laplacians. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by circle-dash curves and the remaining halves are given by square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.5). Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). There are significant changes in variance in the global, ocean, NH, and NASS domains, indicating that several components significantly change variance in those domains for the JAS season.

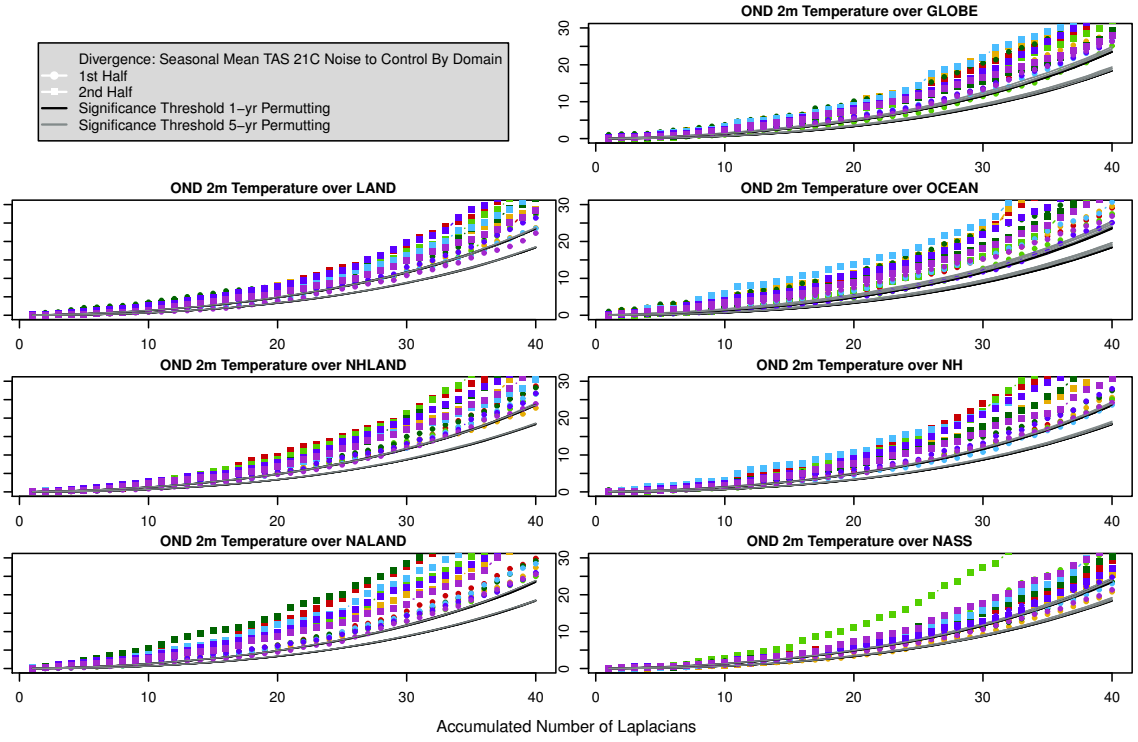

Figure 3.17: Shown for domain, indicated in the title of each panel, is the divergence D_T of internal variability of OND mean 2m temperature between 21st century residuals and preindustrial control simulations, as a function of accumulated Laplacians. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by circle-dash curves and the remaining halves are given by square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.5). Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). There are significant changes in variance in each panel, indicating that several components significantly change variance in each domain for the OND season.

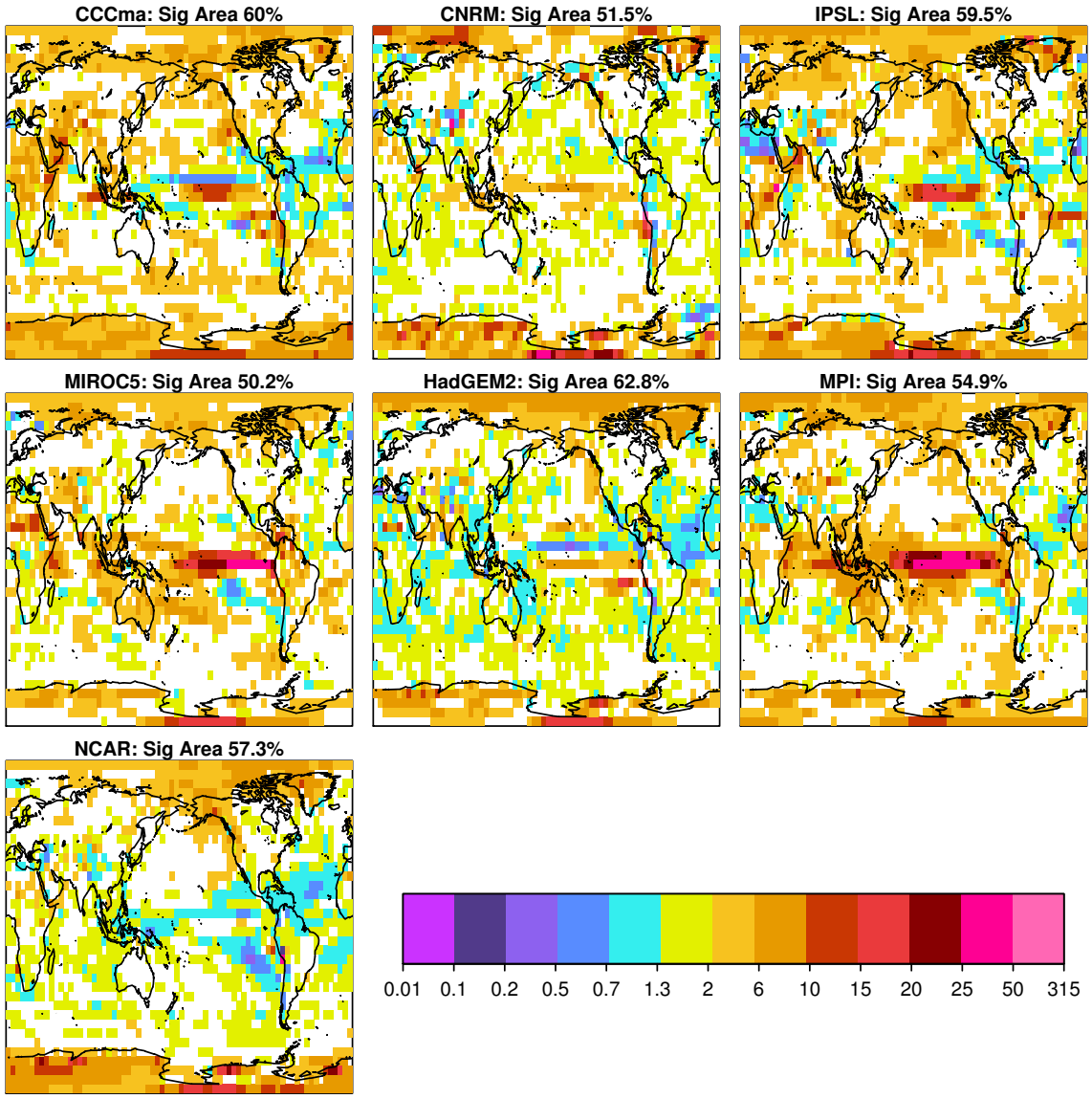

Figure 3.18: The local changes in internal variability of annual mean precipitation due to anthropogenic forcing as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. The variance of internal variability for preindustrial forcing is computed from the same model's 500-year control simulation. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel along with the model name.

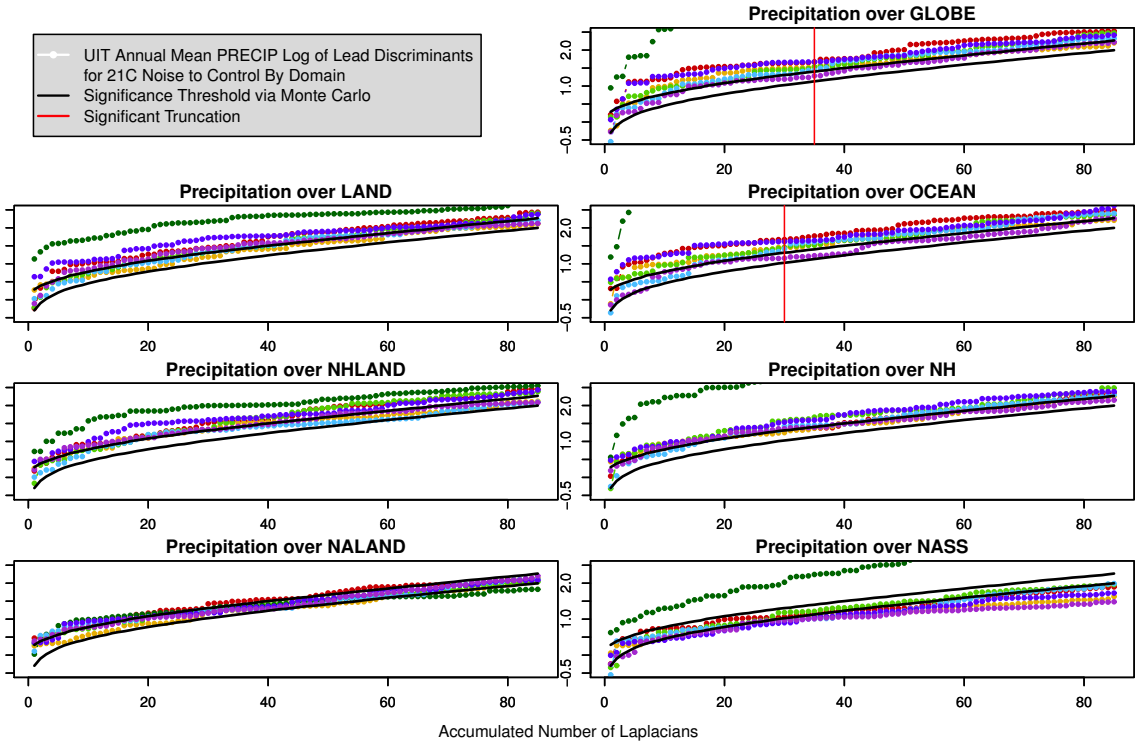

Figure 3.19: Shown for domain, specified in the title of each panel, is the log of the leading discriminant ratios for internal variability of annual mean precipitation between the 21st century residuals and preindustrial control simulations as a function of accumulated Laplacians. Note the curves are color-coded for each model (see legend inset fig. 3.5). Also shown are the upper and lower 5% significance thresholds computed from Monte Carlo techniques (black curves). In the global domain, 5 of the 7 models indicate significant ratios after truncation $T = 35$ (marked by a vertical red line). For the ocean domain, the leading discriminants are significant for 5 out of 7 models after $T = 30$ (marked by a vertical red line). These results are consistent in the 2nd-halves (not shown) and indicate that a single component of variability significantly changes variance. The trailing discriminants are not significant (not shown).

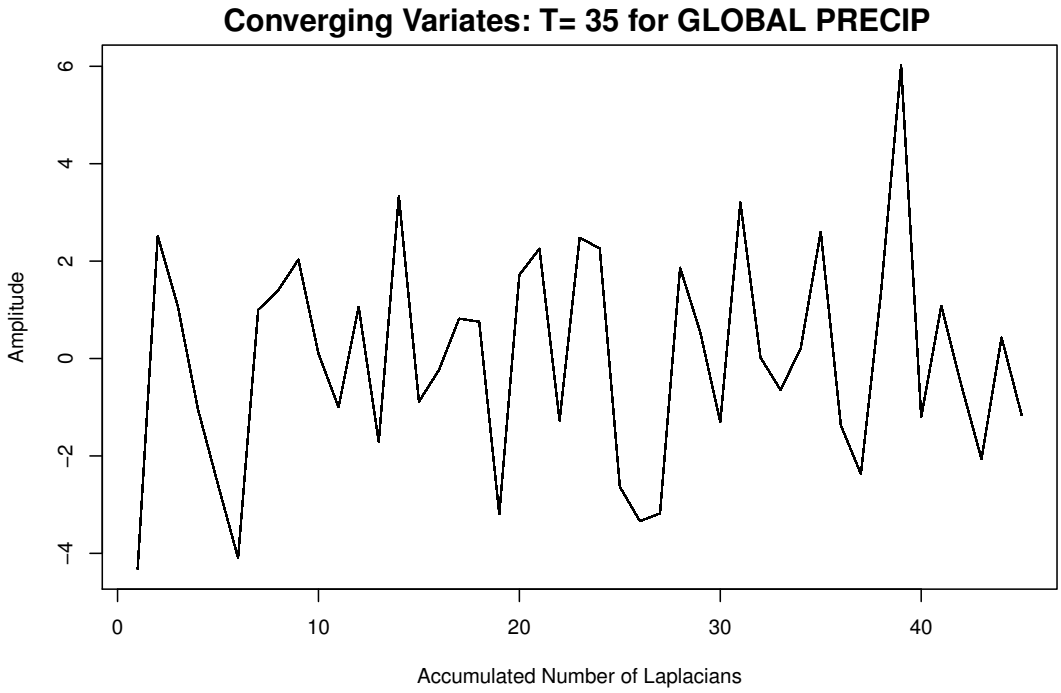

Figure 3.20: Example of converging variates plotted as centered time series from one representative model. The significant variate from one ensemble member for the first half of the 21C at $T = 35$ is plotted in black, and the leading variate for every following truncation is plotted on top. The individual variates are difficult to see because they are virtually identical. As such, it has been demonstrated that the significant ratio is not sensitive to truncation after $T = 35$. Results were similar for all models at $T = 35$.

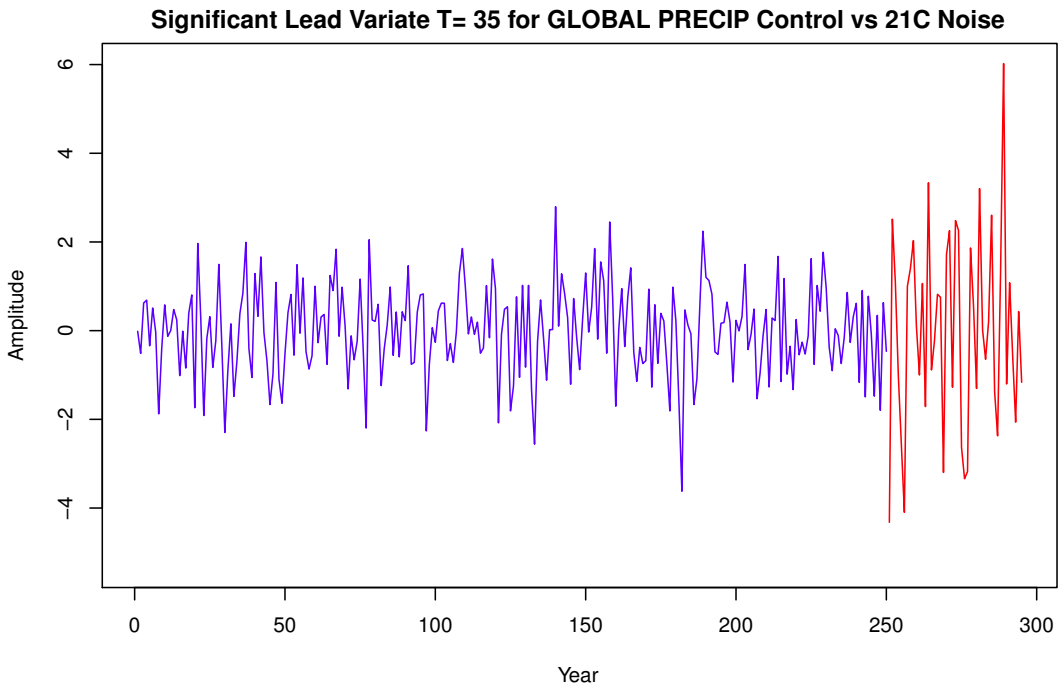

Figure 3.21: An example of significant leading variates, plotted as centered time series, from one representative model; the blue time series indicates a 250-yr segment from the model's preindustrial control simulation and the red time series indicates a 45-yr segment from a 21C simulation (only one ensemble member is shown for clarity). Notice the variance appears larger in the red time series as compared to the blue time series, indicating that variance increases with time. Results were similar for all models at $T = 35$.

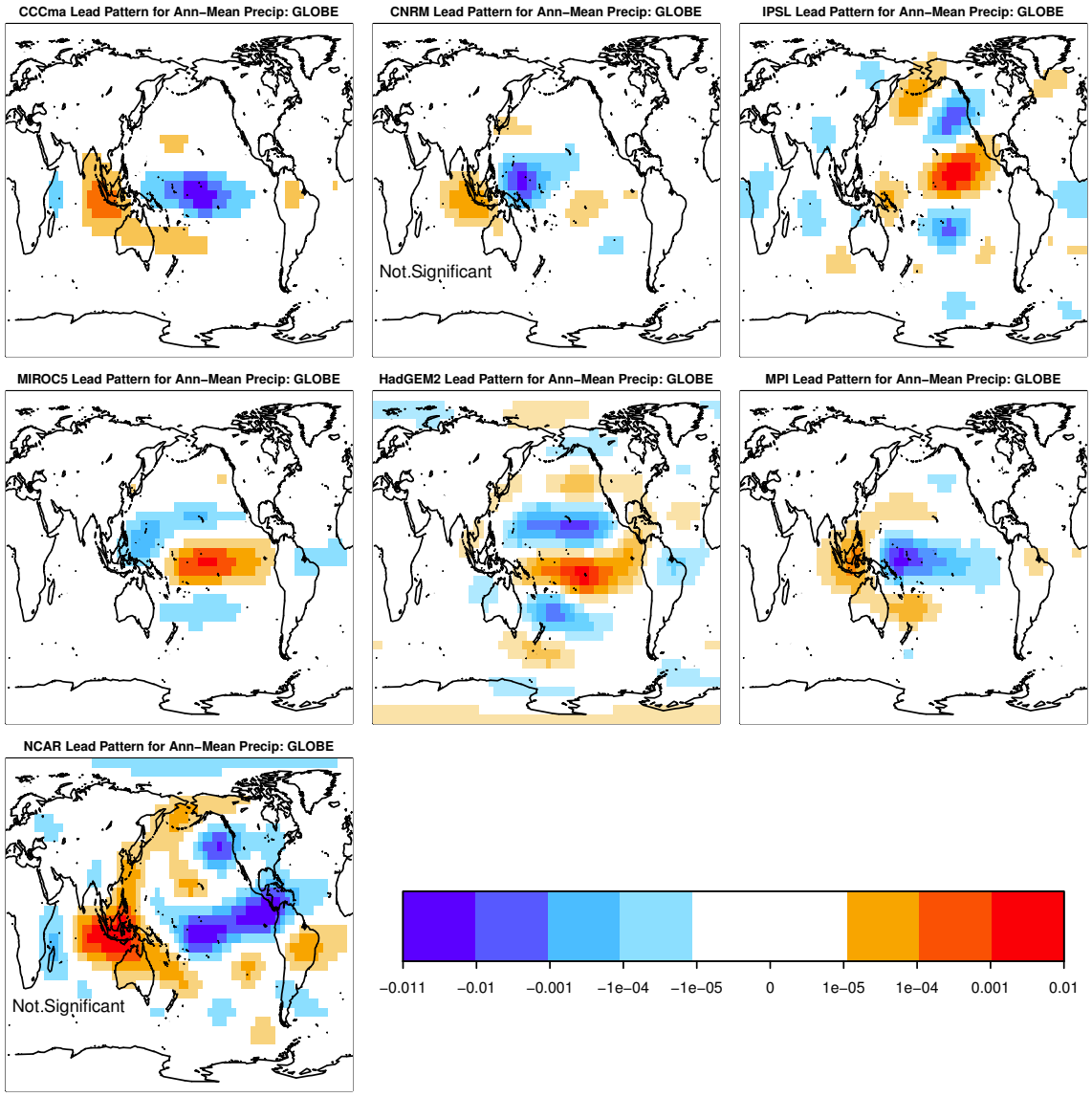

Figure 3.22: Spatial pattern that explains the significantly changing component of climate variability in the annual mean precipitation data in the global domain at $T = 35$. For two models, the phrase 'Not Significant' appears on the map, indicating that the leading component was sensitive to truncation (see CNRM and NCAR). Changes are across the tropical Pacific Ocean in each model. Some models indicate that the tropical changes expand to include subtropical and middle latitudes.

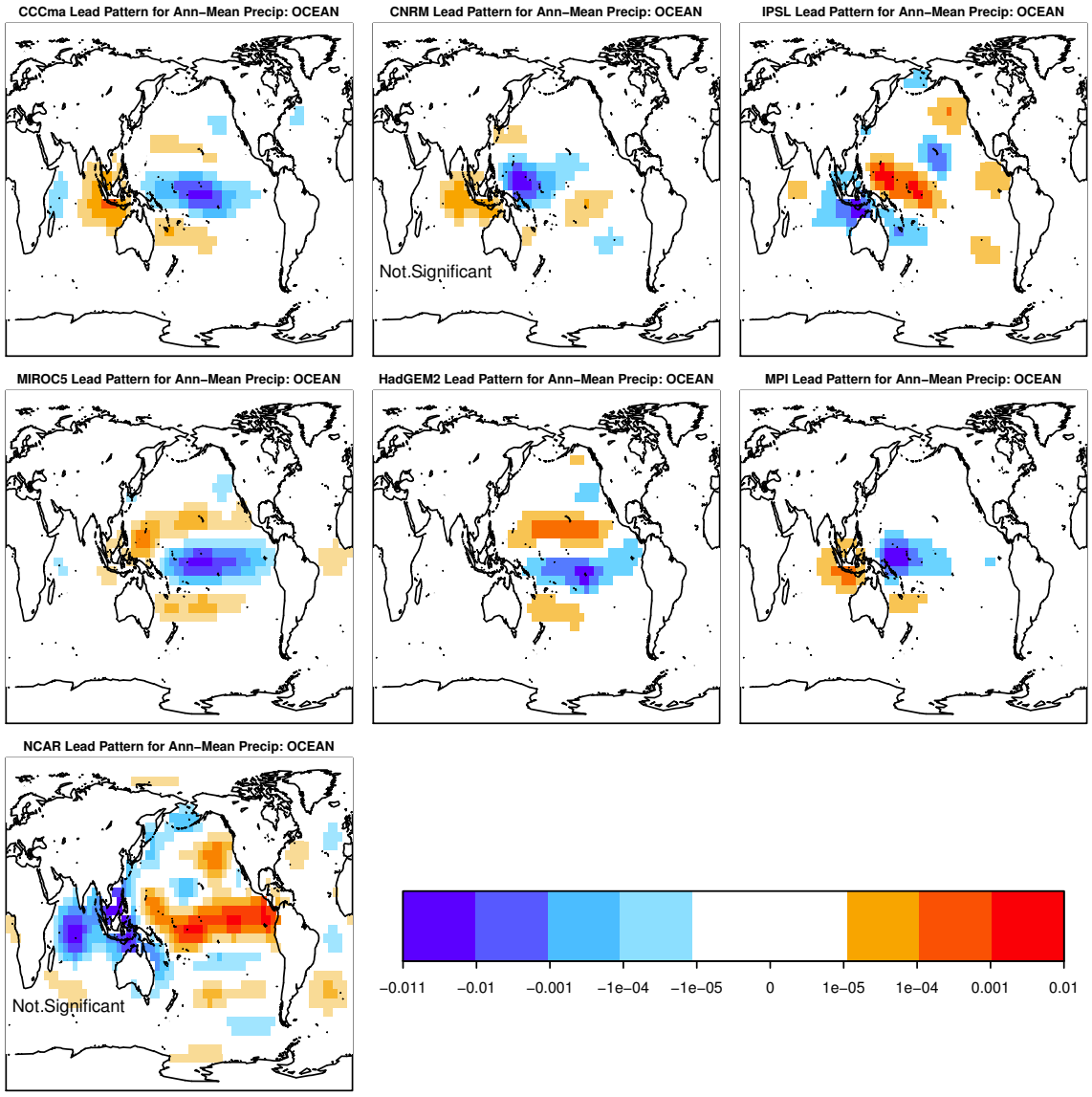

Figure 3.23: Spatial pattern that explains the significantly changing component of climate variability in the annual mean precipitation data in the ocean domain at $T = 30$. For the same two models in the preceding figure, the phrase 'Not Significant' appears on the map, indicating that the leading component was sensitive to truncation (see CNRM and NCAR). Changes are across the tropical Pacific Ocean in each model. Some models indicate that the changes expand to include the subtropical ocean latitudes and the north-eastern portions of the Pacific Ocean.

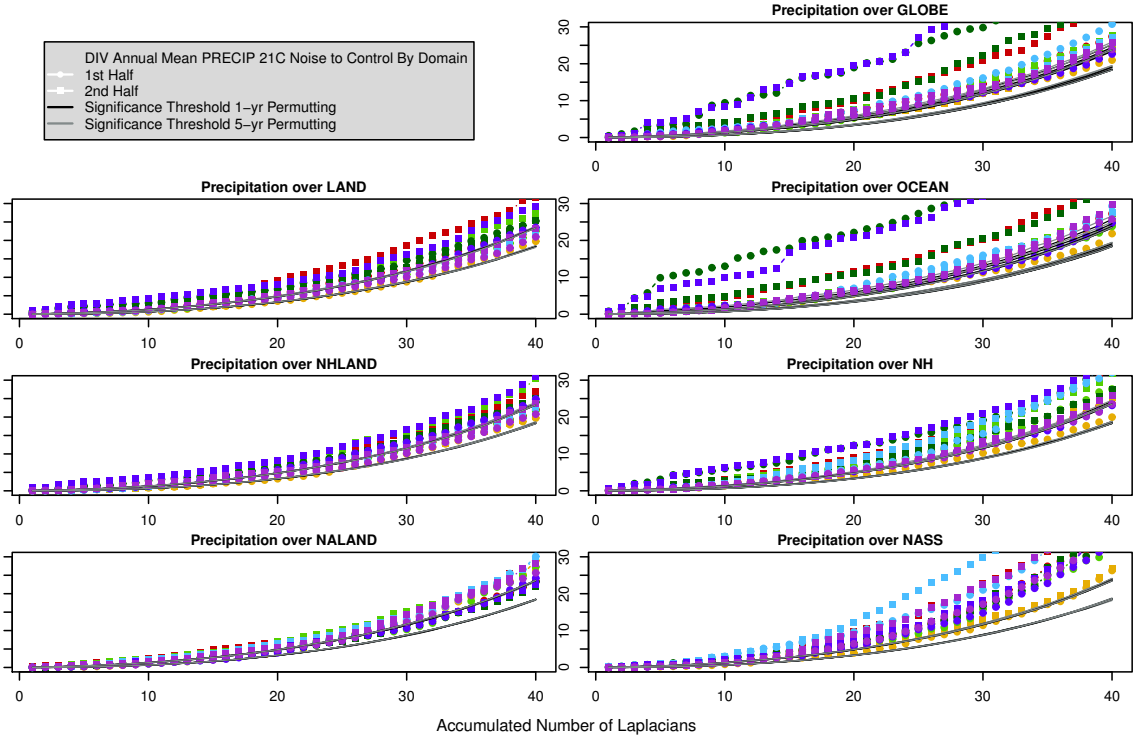

Figure 3.24: Shown for domain, specified in the title of each panel, is the divergence D_T of internal variability of annual mean precipitation between 21st century residuals and preindustrial control simulations as a function of accumulated Laplacians. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by circle-dash curves and the remaining halves are given by square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.26). Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). There are robust significant changes in variance detected in the global, ocean, NH, NASS, and NALAND domains. In the land and NHLAND domains, the results are slightly less robust in the first half of the century.

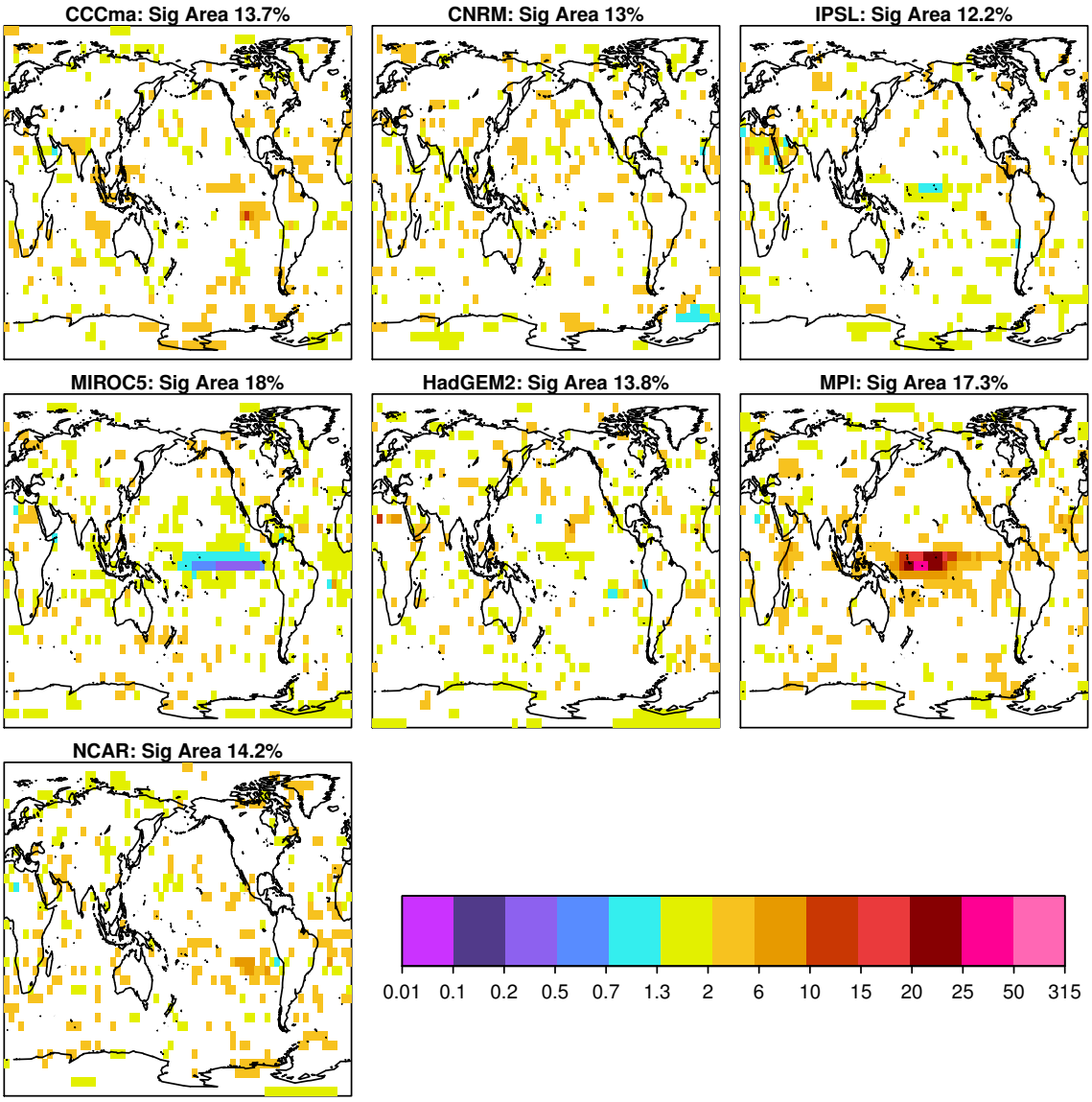

Figure 3.25: Local changes in the variance of annual mean precipitation in the absence of anthropogenic forcing. The changes are quantified by an F-test between non-overlapping (250 year) segments within each model's preindustrial control simulation. Insignificant values (according to an F-test distribution at a 10% significance level) are masked out (i.e., not colored). The percentage of grid points deemed significant in each F-map is provided in the title along with the model name. The changes in variability are close to what one would expect by random chance (about 10% of the area). In addition, the changes appear scattered and lack spatial structure. There are notable exceptions; MIROC5 shows significant decreases along the equatorial Pacific Ocean, while in the same area, MPI shows increases.

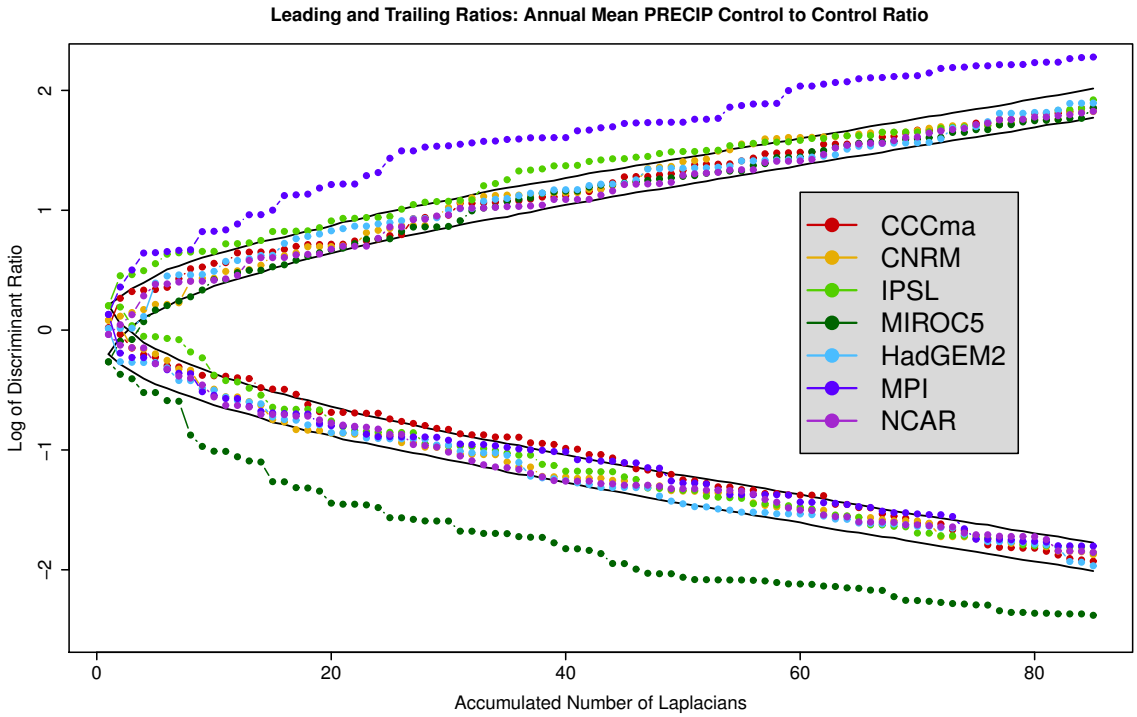

Figure 3.26: Shown for only the global domain, the log of the leading and trailing discriminant ratios computed from the internal variability of annual mean precipitation in the preindustrial control runs as a function accumulated Laplacians. Note the curves are color-coded for each model (see legend inset). Also shown are the upper and lower 5% significance thresholds computed from Monte Carlo techniques (black curves). In the leading curves, MPI (dark blue) is significant for all but the first ratio, indicating that the changes in its F-map can likely be explained by changes in a single, large-scale component of climate variability present in this model. Similarly for the trailing ratios, MIROC5 (dark green) is significant for all truncations, again indicating that the changes in its F-map can also likely be explained by changes in a single, large-scale component of climate variability present in that model. While these results are noteworthy and demonstrate a clear case in which changes in variance can be captured by one or two, large-scale structures of climate variability using the UIT, overall, the results are not robust across models.

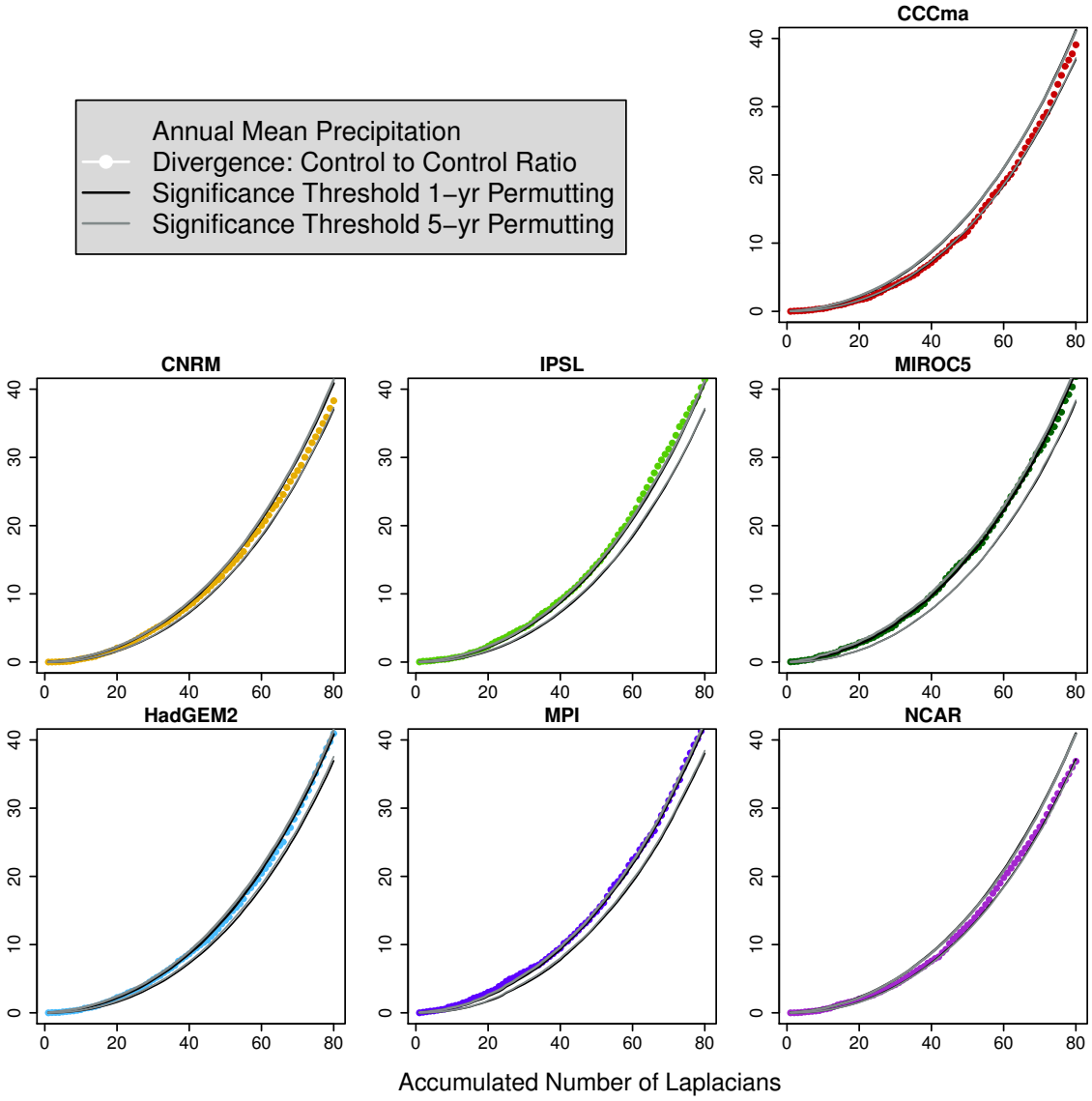

Figure 3.27: Shown for only the global domain, the divergence D_T of internal variability of annual mean precipitation in the preindustrial control simulations as a function of accumulated Laplacians. Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). This result demonstrates that changes in variance across several components is not significant within the global domain and thus, changes in the associated F-maps are not significant in an aggregate sense within the global domain.

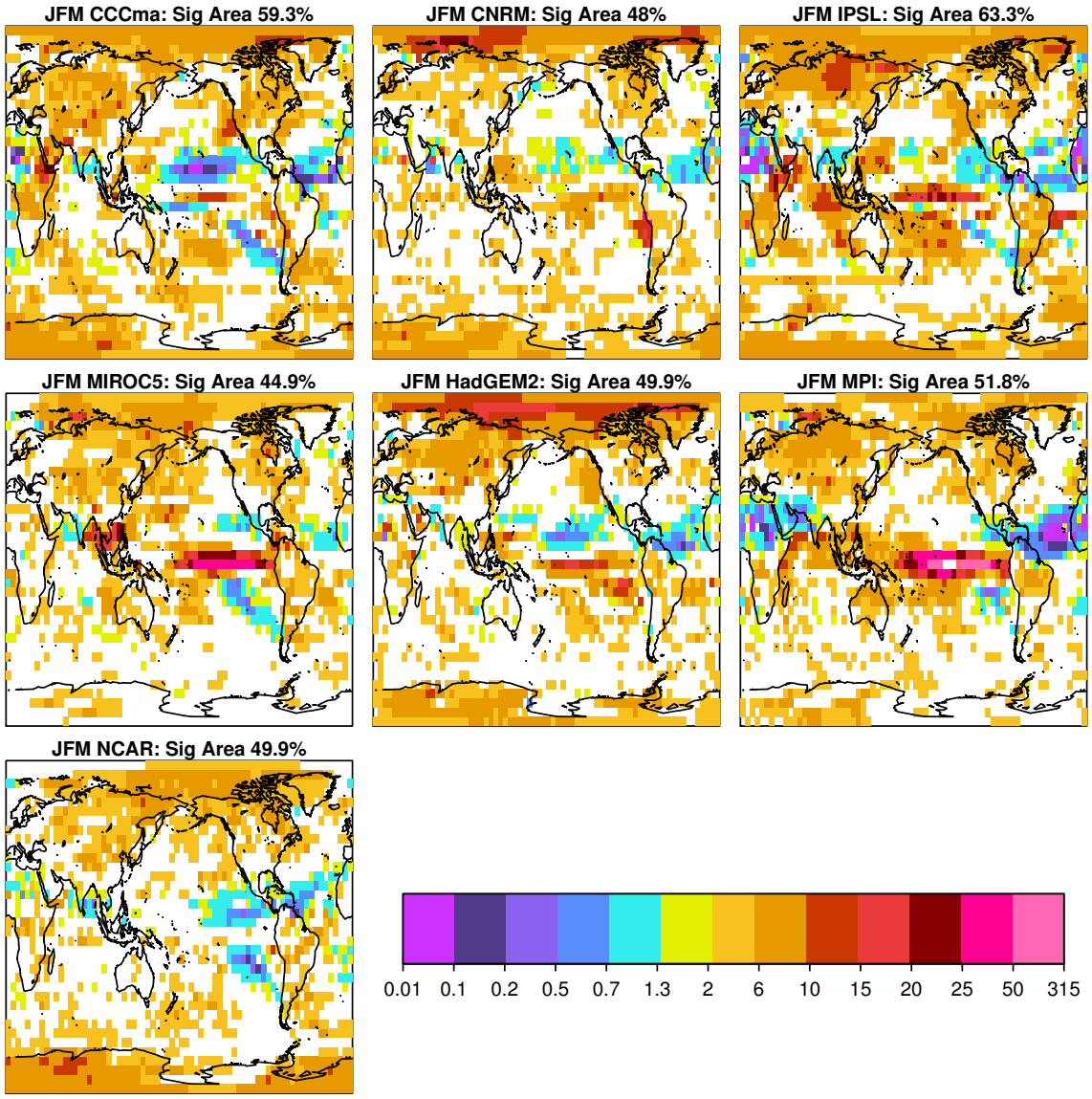

Figure 3.28: The local changes in internal variability of January-February-March (JFM) mean precipitation due to anthropogenic forcing as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel. The models generally agree about the pattern and direction of change in the polar latitudes. The direction of change in the tropics is model dependent.

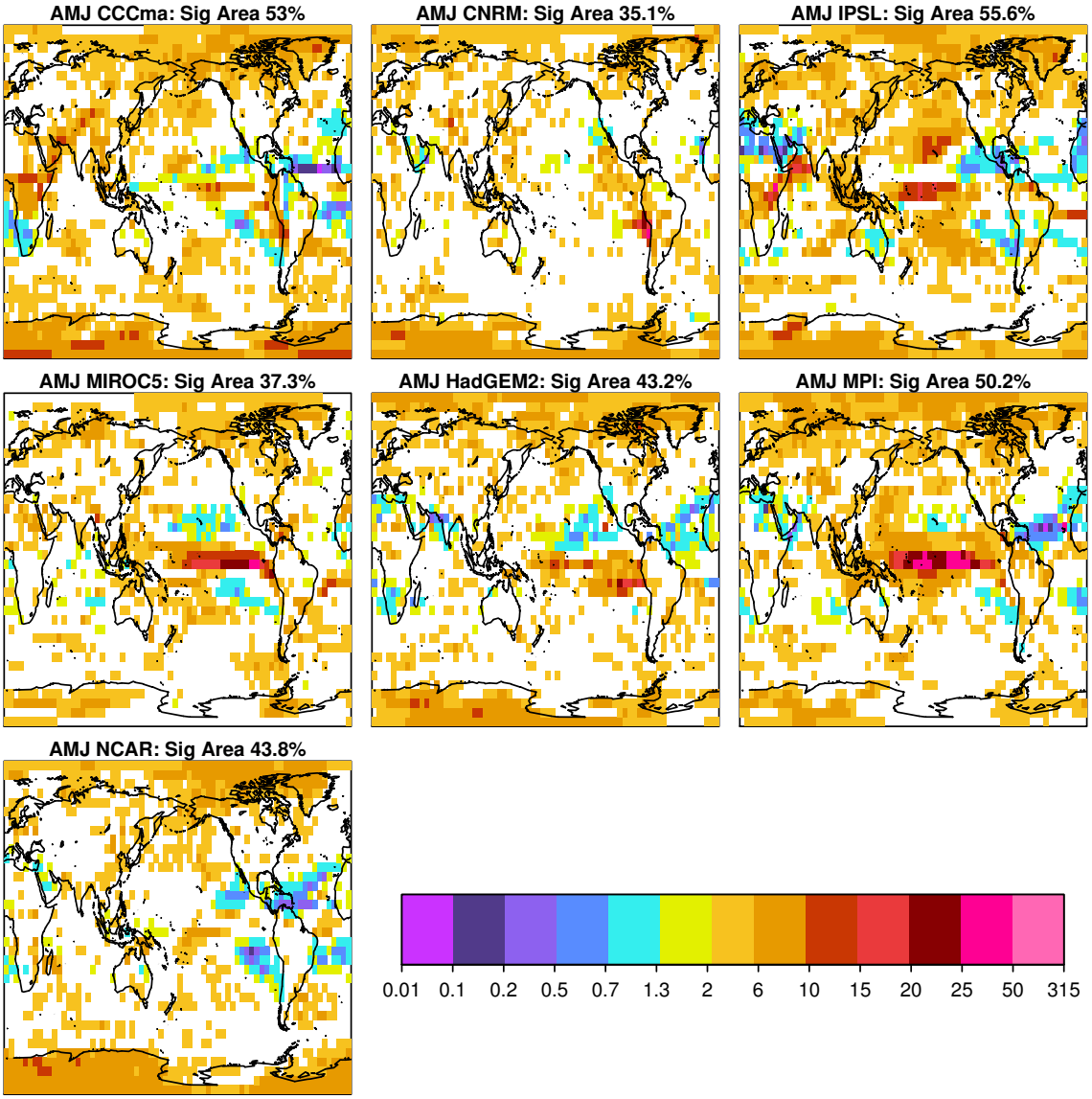

Figure 3.29: The local changes in internal variability of April-May-June (AMJ) mean precipitation due to anthropogenic forcing as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel. The models generally agree about the pattern and direction of change in the polar latitudes.

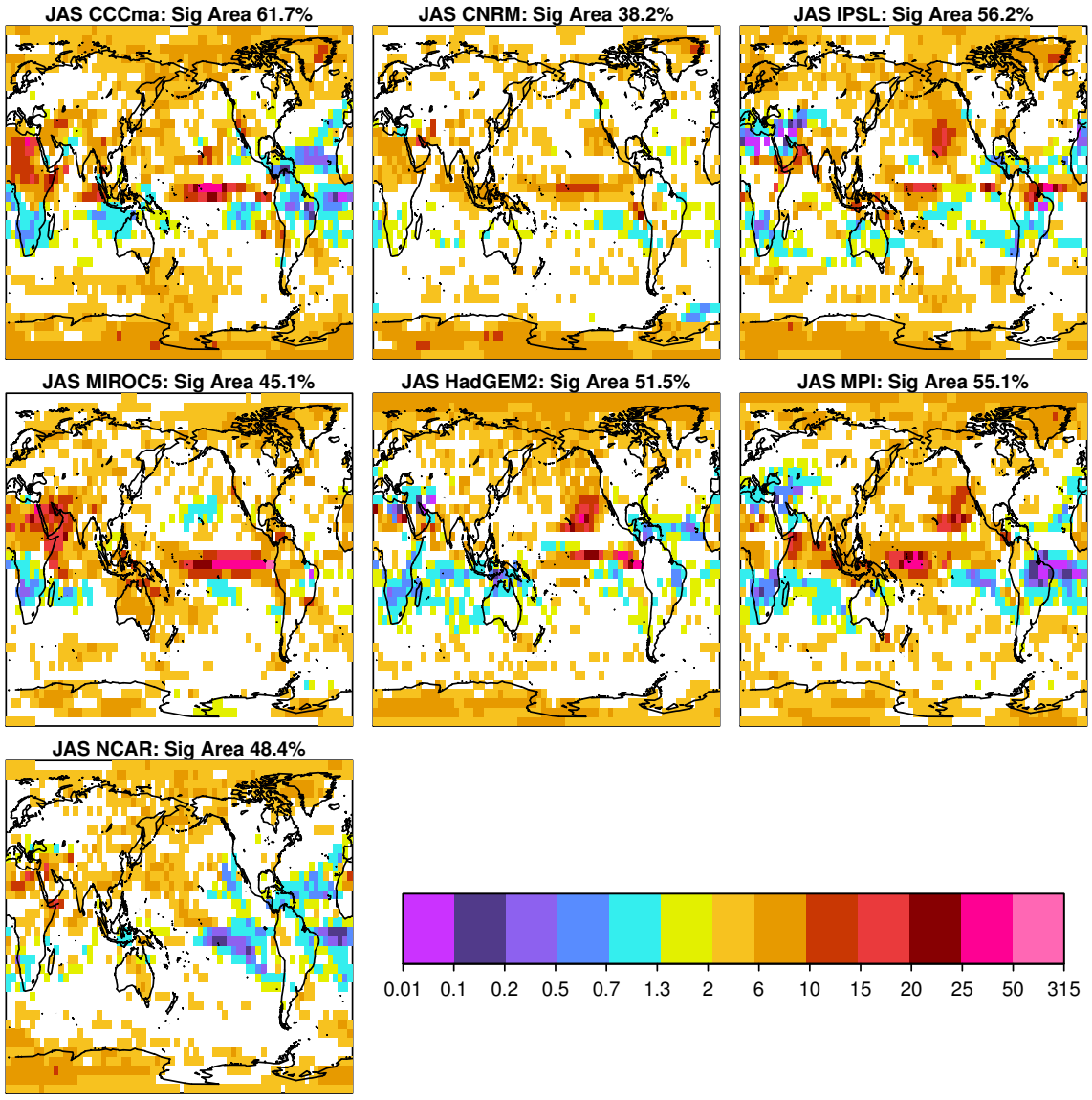

Figure 3.30: The local changes in internal variability of July-August-September (JAS) mean precipitation due to anthropogenic forcing as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel. The models generally agree about the pattern and direction of change in the polar latitudes. The increase in variance for the tropics is consistent across models for all but one model; NCAR.

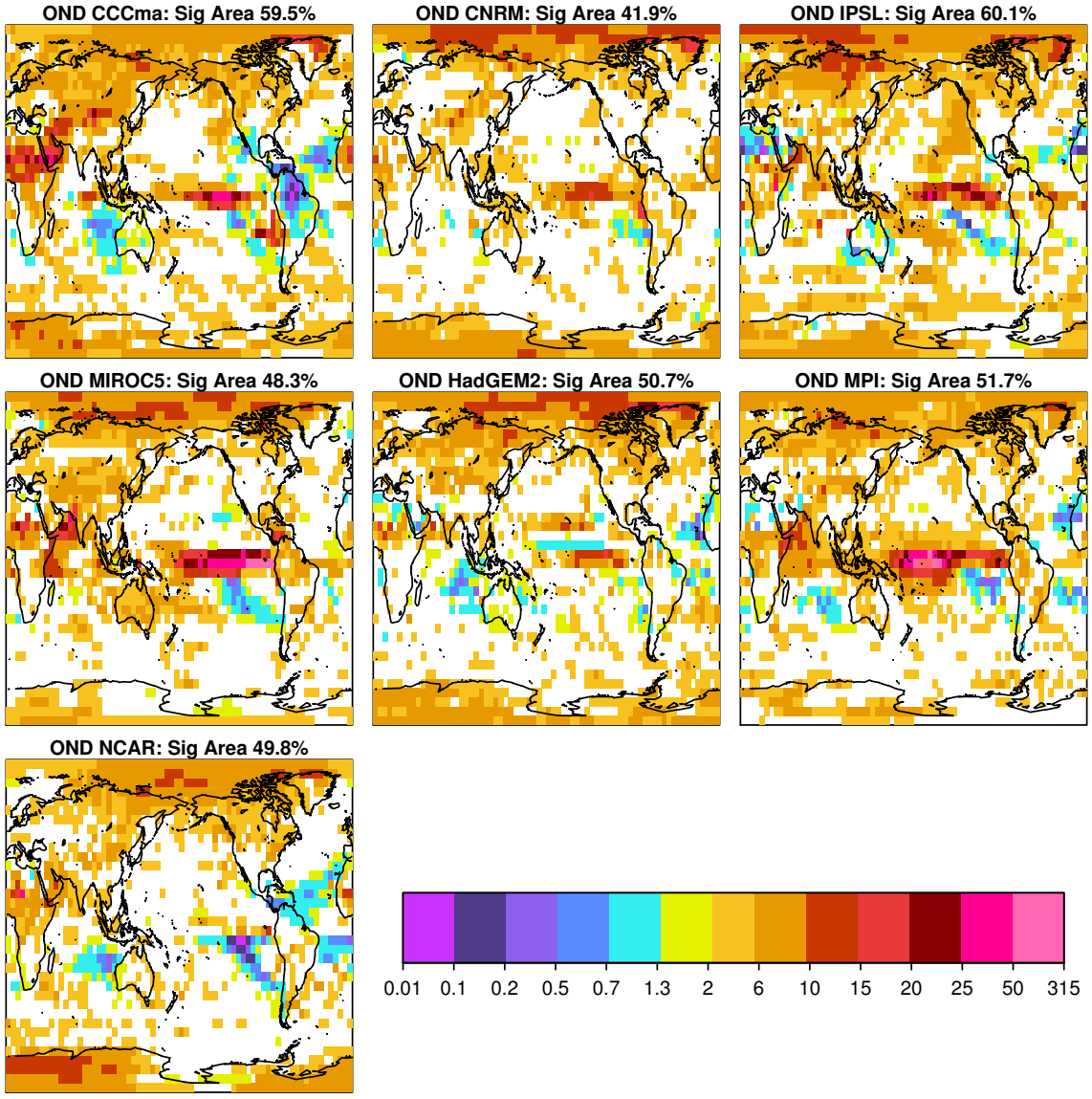

Figure 3.31: The local changes in internal variability of October-November-December (OND) mean precipitation due to anthropogenic forcing as quantified by an F-test between the 21C and preindustrial control simulations for each model. The variance of internal variability during the 21C is computed from residuals about the ensemble mean of a three member ensemble using a high emissions scenario (RCP8.5) for the 90-year period from 2006 to 2095. A ratio larger than one indicates internal variability increases in the 21st century. Insignificant values (according to the F-test distribution at a 10% significance level) are masked out. The percent area of significant grid points is indicated in the title of each panel. The models generally agree about the pattern and direction of change in the polar latitudes. The increase in variance for the tropics is consistent across models for all but one model; NCAR.

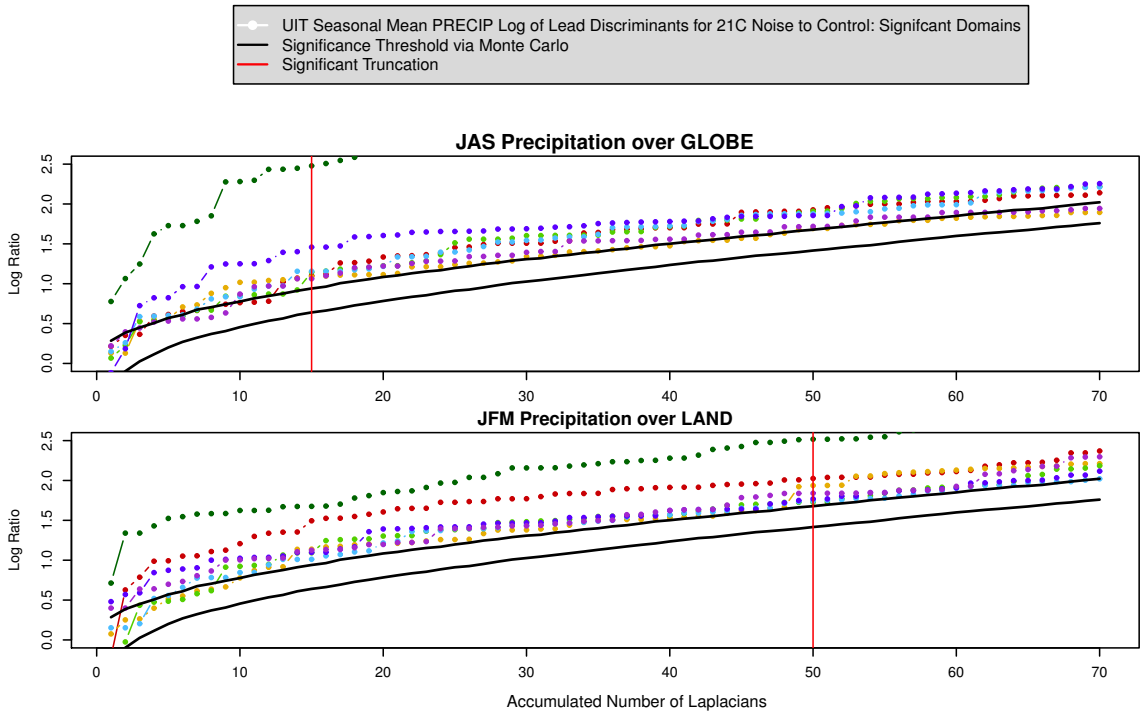

Figure 3.32: Shown are the log of the leading discriminant ratios for select seasons for precipitation as a function of accumulated Laplacians. The top panel displays the leading ratios for July-August-September (JAS) for the global domain and the bottom panel displays the leading ratios for January-February-March (JFM) for the land-only domain. Trailing ratios were not significant (not shown). Note the curves are color-coded by model (see legend inset fig. 3.26). Also shown are the upper and lower 5% significance thresholds computed from Monte Carlo techniques (black curves). These conclusions were verified in the independent data (not shown). A vertical red line in each panel marks the significant truncation. For JAS the significant truncation is $T = 15$ for 5 out of 7 models; note that NCAR (purple curve) and CNRM (yellow curve) are marginally significant for all truncations. For JFM in the land domain, the significant truncation is $T = 50$ for 6 out of 7 models; note that HadGEM2 (light blue curve) is the only model that becomes marginally significant beyond this truncation.

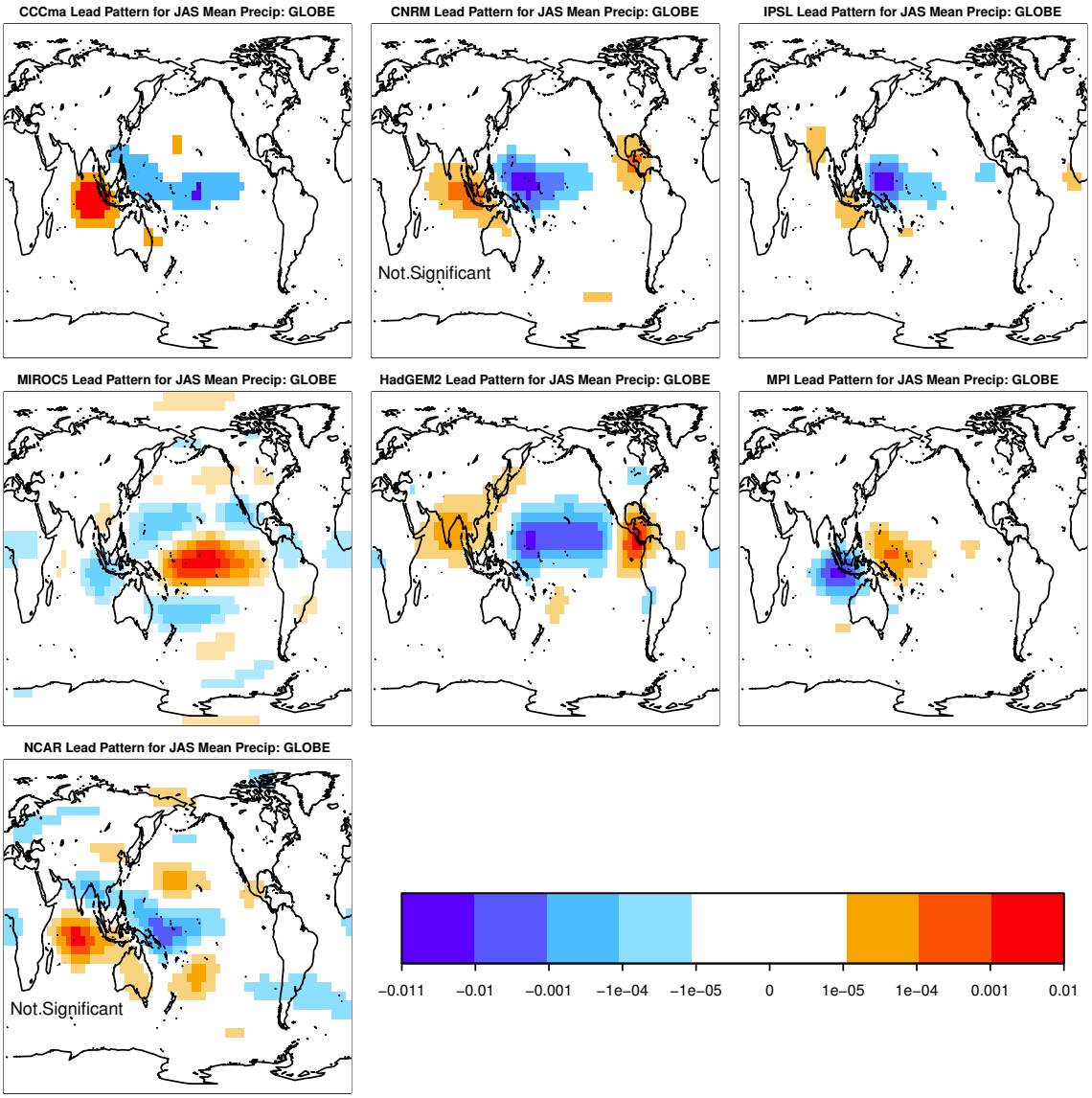

Figure 3.33: Spatial pattern that explains the significantly changing component of climate variability in the JAS mean precipitation data in the global domain at $T = 15$. For two models, the phrase 'Not Significant' appears on the map, indicating that the leading components were sensitive to truncation (see CNRM and NCAR). The changes are generally confined to the tropical Pacific Ocean, with some expansion to the subtropical and middle latitudes displayed by some models. These results suggest that significant changes in the variability of summer precipitation may occur in the tropical Pacific Ocean in a warmer climate, which supports conclusions from the annual mean study that were displayed in fig. 3.22

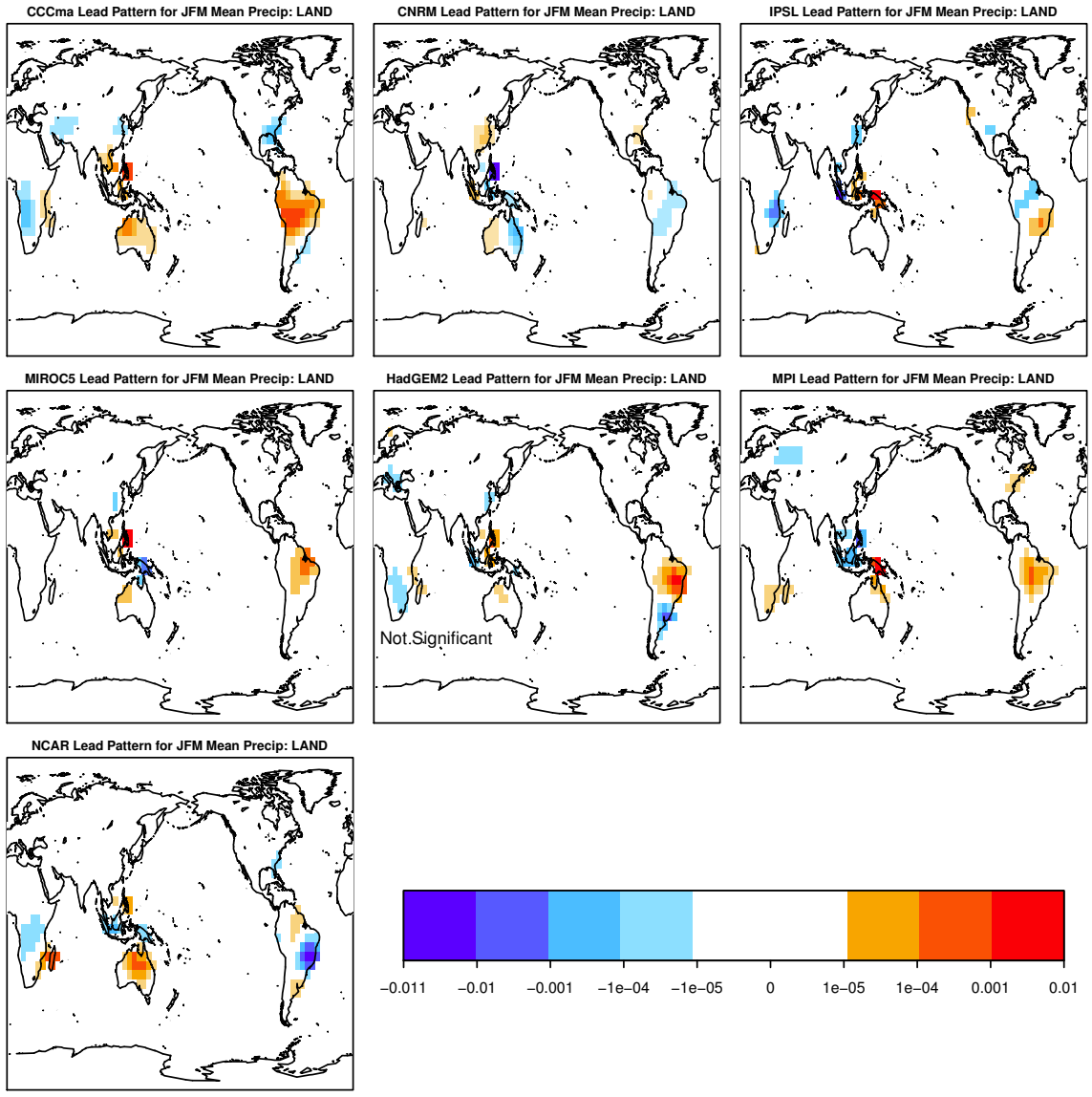

Figure 3.34: Spatial pattern that explains the significantly changing component of climate variability in the JFM mean precipitation data in the land domain at $T = 50$. For one model, the phrase 'Not Significant' appears on the map, indicating that the leading component was sensitive to truncation (see HadGEM2). The models indicate that winter changes in variability are projected for the land-areas between subtropical zones.

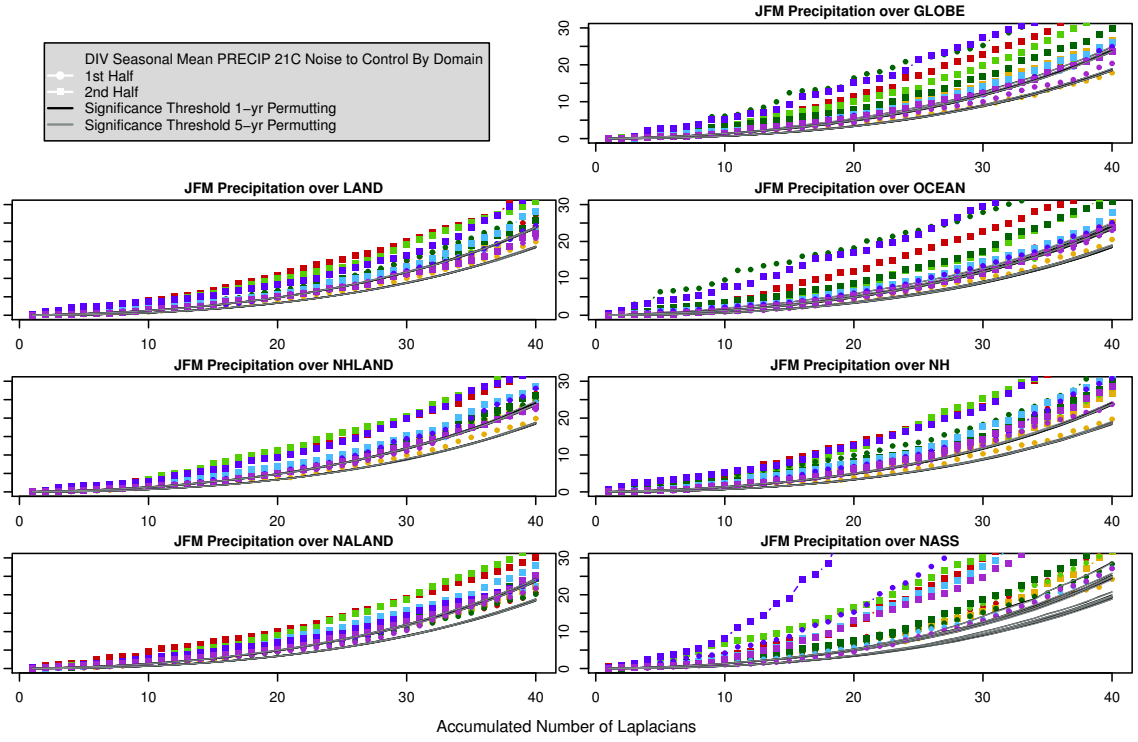

Figure 3.35: Shown for domain, specified in the title of each panel, is the divergence D_T of internal variability of JFM mean precipitation between 21st century residuals and preindustrial control simulations, as a function of accumulated Laplacians. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by circle-dash curves and the remaining halves are given by square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.26). Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). In every domain the 2nd-halves are robustly significant, and for the 1st-halves, at least 5 out of 7 models indicate significant changes in each domain.

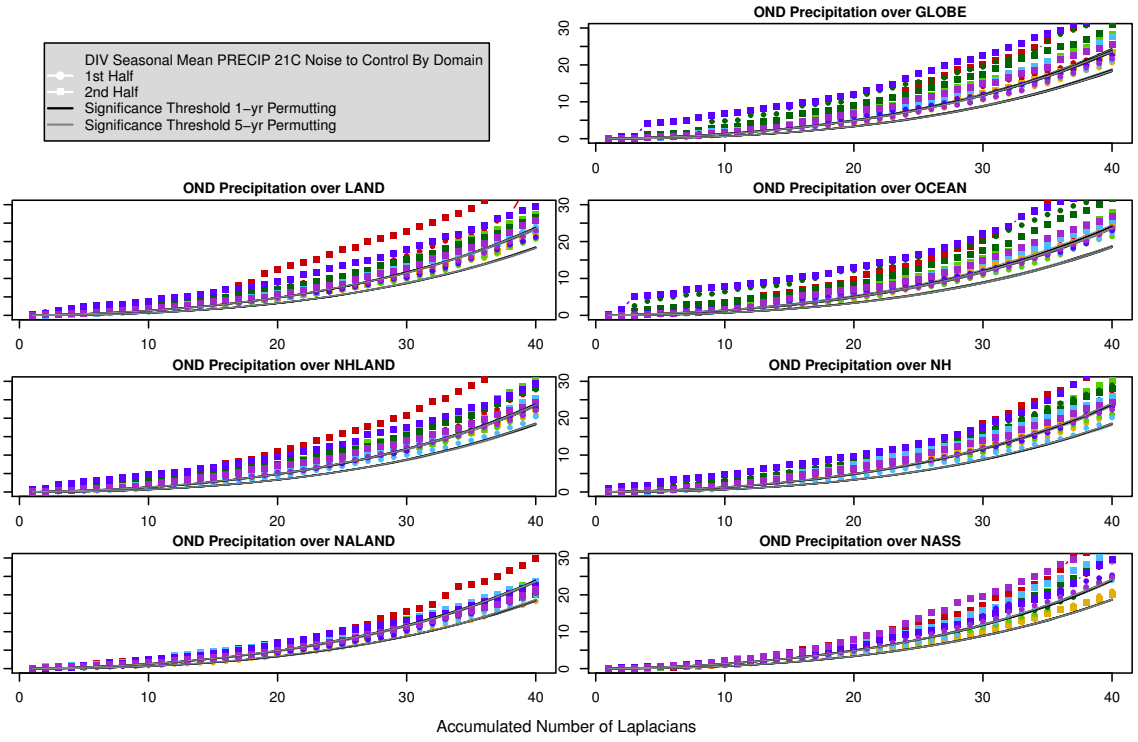

Figure 3.36: Shown for domain, specified in the title of each panel, is the divergence D_T of internal variability of OND mean precipitation between 21st century residuals and preindustrial control simulations, as a function of Laplacians. The divergence for the first half of the 21st century and first half of the preindustrial control run are given by circle-dash curves and the remaining halves are given by square-dash curves. Note the curves are color-coded by model (see legend inset fig. 3.26). Also shown are the upper and lower 5% significance thresholds computed from permutation techniques; 1-yr (black curves) and 5-yr (gray curves). Robustly significant changes are found in the ocean, land, NHLAND, and NASS domains. In the remaining domains, changes are robust in the 2nd-half of the 21st century.

Table 3.1: Seven domains are used in this study. This table provides a description of each domain, the short-form domain name, and the boundaries for each domain, given as longitude range and latitude range. The terms 'seas' refers to the Mediterranean, the Baltic, and the Caspian seas in particular.

Domain Information		
Description:	Short-form:	Lon and Lat:
Grid points for whole globe	GLOBE	0 to 360; -90 to 90
Global ocean-only grid points; no land or seas	OCEAN	0 to 360; -60 to 90
Global land-only grid points; no oceans or seas	LAND	0 to 360; -60 to 90
Ocean and land grid points in the Northern Hemisphere	NH	0 to 360; 0 to 90
Land-only grid points in the Northern Hemisphere	NHLAND	0 to 360; 0 to 90
Land-only grid points in the North American continent	NALAND	190 to 310; 15 to 70
Ocean grid points in the North Atlantic Ocean	NASS	285 to 352; 0 to 60

Chapter 4: Conclusions

4.1 Summary of Work and Discussion

A basic paradigm in climate science is that the variability of the climate system can be split into two kinds: external variability that is driven by external forcings, such as human-caused increases to greenhouse gases, and internal variability that arises naturally from coupled atmosphere-ocean-land-ice interactions. This dissertation examined whether internal variability would change in response to changes in external forcing. This topic is both an important and challenging area of research. For instance, a change in temperature variability would mean a change in the frequency and intensity of extreme hot and cold weather events, which would have significant impacts on society.

In part, this dissertation is a response to studies like that of Hansen et al. (2012), which claimed that the distribution of observed seasonal mean temperature anomalies has shifted toward higher temperatures and that the range of anomalies also has increased. This conclusion was based on a technique in which local temperatures were individually normalized and then aggregated over space to obtain a distribution. Unfortunately, the resulting distribution turned out to be sensitive to the chosen method of normalization.

This dissertation proposes a methodology for testing changes in variability that overcomes the problems with spatial aggregation. This methodology requires using ensemble climate simulations. The more difficult problem of estimating such changes in observational data is left for future work. Simulations known as RCP8.5 are examined. In these simulations the carbon dioxide concentrations increase over time such that they reach 1250 ppm by 2100. Climate models suggest that global mean temperature will rise in response to such increases by at least 2.6 degrees C, and as much as 4.8 degrees C, by the end of the century. The internal variability in these simulations was estimated by subtracting a

3-member ensemble mean from each ensemble member. The residual yields an estimate of projected internal variability in the 21st century. The projected internal variability can be compared to estimates of internal variability from preindustrial control simulations, in which the forcing remains constant.

A straightforward approach to testing local changes in internal variability is to perform a standard F-test. In this dissertation, an F-test was applied to a ratio of estimates of internal variability, as projected by CMIP5 model simulations of the RCP8.5 emissions scenario, and respective preindustrial control simulations in order to quantify future changes in the internal variability of temperature and precipitation. Results show that there are widespread, significant local changes in the variance of annual mean 2m temperature in response to anthropogenic forcing (see fig. 3.1). The number of changes exceeds what one would expect to occur by random chance (about 10%). In some cases, the number of changes exceeded 60% of the total area (see MIROC5 in fig. 3.1). All the models predict significant decreases in variance in regions of sea-ice formation and across the Southern Ocean (see fig. 3.1). This decrease is a plausible consequence of disappearing sea-ice due to global warming. In a warmer climate, melting sea-ice exposes the underlying sea surface which has a much larger effective heat capacity than sea-ice due to its coupling with the oceanic mixed layer. This interpretation also is consistent with previous research (e.g., DelSole et al., 2013; Huntingford et al., 2013; Screen, 2014; Screen et al., 2014).

The analysis also reveals that more than half the models predict significant future changes in the variance of temperature over the tropical oceans, land-areas, and the North Atlantic Ocean, but the sign of these changes is model dependent. ENSO's global influence on temperature and precipitation extremes makes it an important player in climate variability. Unfortunately, this study shows that the response of annual mean 2m temperature in the ENSO region to increasing greenhouse gases is highly model-dependent; a result that is in accord with other studies (e.g., Collins et al., 2010; Vecchi and Wittenberg, 2010).

Since the F-test is univariate, it cannot assess the field significance of the estimated changes in variability. To be field significant, it should be demonstrated that the resulting

F-map is *unlikely* given the spatial and temporal relationships between the grid points. Thus, a proper field significance test requires estimating the covarying properties of the data (for example a covariance matrix).

Discriminant analysis provides an attractive approach to quantifying multivariate differences in variance. Discriminant analysis is a method that identifies an uncorrelated set of components and orders them such that the first maximizes the ratio of variances, the second maximizes the ratio of variance subject to being uncorrelated with the first, and so on. The method depends on the entire covariance matrix (i.e, the variances and the covariances between grid points) and therefore takes into account the full spatio-temporal relations between grid points (multivariate). Also, the components and variance ratios derived from discriminant analysis are invariant to invertible affine transformations, which implies that the results from discriminant analysis are independent of how the data are normalized.

Unfortunately, discriminant analysis cannot be applied when the number of grid points exceeds the number of samples, because then the covariance matrices are singular and variance ratios like 0/0 can be obtained. To avoid this problem, the data were projected into a reduced subspace spanned by a small number of eigenvectors of the Laplace operator. These eigenvectors were chosen as a basis set because they quantify changes in variability of large-scale spatial structures, which is the primary interest of this study. Laplacian eigenvectors provide a natural, orthogonal set of vectors that can be ordered by length scale and are independent of model data (unlike EOFs), which facilitates comparison between models. Although projecting data into a reduced dimensional subspace avoids singularities, it introduces a new concern, namely a model selection problem in which the dimension of the subspace must be specified. The standard approach is to choose the dimension based on some objective criterion. After a variety of criteria were explored, they were ultimately rejected because the results were sensitive to the truncation parameter, as explained shortly.

A standard test for equality of covariance matrices is the Union-Intersection Test (UIT). This test is based on the extreme value of the variance ratio, which as discussed above can be obtained from discriminant analysis. If the covariance matrices of the two populations

are equal, then any linear combination of variables yields the same variance in the two populations. In such a case, the variance ratio would be one, or statistically indistinguishable from one. Conversely, if the two covariance matrices differ, then some linear combination can be found in which the ratio differs from one. The UIT is based on the most extreme variance ratio obtained from all possible linear combinations. The UIT was applied to climate model simulations of internal variability of annual mean 2m temperature. Notably, the extreme variance ratio was found to be only marginally significant, if significant at all, for most models. In other words, the significance of the variance ratio was sensitive to the number of Laplacian eigenvectors used to represent the data. This result suggests that changes in the variance of *any single component* of climate variability (e.g., ENSO) cannot be identified in these model simulations.

Despite the above conclusion, the variance ratio maps (F-maps) clearly showed changes in variance in response to anthropogenic forcing. These apparently contradictory results can be reconciled if one considers that changes in variance can occur over a wide spectrum of components, but that the change in variance in any single component may be too small to be significant by itself. In other words, the change in variance could be significant in an aggregate sense and not in an individual sense. For this reason, the UIT fails to detect a change in variance because it relies on detecting such changes as would occur in a single, maximized component. Accordingly a new test was proposed, namely, one that measures the difference between covariance matrices in an aggregate sense. The chosen measure is divergence, which can be derived from an information theoretic measure called relative entropy for the case of Gaussian distributions with equal means. Divergence depends only on the optimized variance ratios derived from discriminant analysis, implying that divergence also is invariant to affine transformation. In addition, divergence can be written as a sum of variance ratios plus a sum of inverse variance ratios, clearly demonstrating that it is an aggregate measure of differences in variance. Divergence was also applied to climate model simulations of internal variability of annual mean 2m temperature and it was found that differences in variance could be easily detected.

To learn more about the source of these changes, the analyses were repeated in ocean-only and land-only domains (as well as smaller domains). Since the F-test is performed on individual grid points, the local data remains the same and does not change with domain selection. In contrast, the UIT and divergence tests can depend on domain. Nevertheless, the results were largely the same as in the global domain: changes in variance could be detected over both ocean and land using the divergence measure. Recall that divergence measures changes in variance without distinguishing the direction of the changes. Perhaps local decreases in variance in regions of sea-ice are responsible for the significant changes in variance detected by the divergence measure. Unfortunately, due to the nature of the divergence measure and the limitations of the F-test, it is not possible to go beyond speculation when comparing the results from these two techniques. Also, while it is clear that decreased variability over regions of sea-ice formation are robust, and likely explained by melting sea ice, other changes in variance are model-dependent thus it is difficult to infer mechanisms for those changes.

Some studies have argued that in a warmer climate, temperature variability can be expected to decrease. In addition to melting sea-ice, another proposed mechanism for this decrease is fewer, or less severe, cold days due to changes in cold-wind advection (e.g., Schneider et al., 2015; Screen, 2014). Screen (2014) argues that Arctic amplification (the more rapid warming of the Arctic as compared to middle latitudes) leads to decreases in Northern Hemisphere temperature variability. In particular, a warm Arctic leads to a decreased temperature gradient in autumn and winter and thus weaker cold-air advection from northern latitudes to the middle latitudes. This dissertation generally supports these findings, especially with regard to projected decreases in temperature variability for the North American continent and Greenland, however, not over Eurasia.

Seasonal mean changes were also investigated in this study and in general, the changes were similar to the annual mean changes. For instance, the decreased variability of temperature that was found over regions of sea-ice formation in the annual-mean study could also be found in the OND, JFM, and AMJ seasonal mean F-tests. Most of the differences between

the annual and seasonal mean studies were found in the polar latitudes of the Northern Hemisphere and/or were model dependent. For instance HadGEM2 showed very few locally significant changes at polar latitudes in the annual mean F-test, but in the seasonal mean study, those same areas showed significant decreases in OND followed by significant increases in JAS. Opposing changes, such as these, would likely cancel in an annual mean.

Changes in the variability of annual mean precipitation were also investigated. An F-test demonstrates that there are widespread, significant local changes in the variance of annual mean precipitation in response to anthropogenic forcing (see fig. 3.18). The models robustly project increases in variance for northern-most and southern-most latitudes. In general, changes in the tropics and subtropics are model dependent, however, a majority of models indicate increased variability occurs there in a warmer climate. In particular, for some models there are very strong increases in variability for the equatorial Pacific Ocean, a result that is in accord with Power et al. (2013).

The union intersection test on this data indicated that a component of climate variability was significantly changing variance in the global domain. In particular, a pattern of increasing variability was identified for the western portions of the tropical Pacific Ocean. Adjacent areas such as the Bay of Bengal, the South and East China Seas, as well as some middle latitudes were also significantly changing variance. A similar component of change was deemed significant in the ocean-only domain. Divergence revealed robust changes in the annual mean variability of precipitation for the North Atlantic Ocean.

The seasonal mean changes in precipitation variability were similar to the annual mean changes, particularly with respect to the robust polar increases in variability. The seasonal mean results indicated that the annual mean changes were most influenced by the strong increases in OND. Additionally, in OND, a majority of models agreed about the direction of change along the tropical Pacific Ocean, indicating increased variability of autumn precipitation is projected in a warmer climate. In the equatorial latitudes of the North Atlantic Ocean, the models robustly agreed on decreased variability of winter precipitation.

In the IPCC's Special Report on Extreme Events (IPCC, 2012), it was reported at high

confidence that heavy rainfall is expected to increase at higher latitudes while at those same latitudes, there is low confidence that droughts will increase IPCC (2012). If both changes occur together, i.e., heavy rainfall and increased drought-frequency, then this would be characteristic of changes in variance for precipitation at those latitudes, namely, increasing variance. Other research indicates that in both observations and model simulations, there is a link between a warmer climate and an increase in extreme precipitation events in the tropics (Allan and Soden, 2008). Lau et al. (2013) found a global increase in heavy precipitation, decrease in moderate precipitation, and increase in light precipitation. If one thinks of these results as describing changes in a globally aggregated PDF of precipitation, then, they would indicate a *broadening and flattening* of the distribution of precipitation. In this case, speaking broadly about global changes, the variability of precipitation could be characterized as 'globally' increasing in response to a warmer climate. Regionally speaking, Lau et al. (2013) also found more heavy precipitation over climatologically wet zones and more dry periods in the subtropics and marginal convective zones (e.g., the Sahel region in Africa). The authors associate their results with the hypothesis of the "wet get wetter and the dry get dryer" in a warmer climate (Held and Soden, 2006). The results presented in this dissertation support the finding that precipitation variability is projected to increase with global warming, especially for convergence zones, mid-latitude storm track zones, and polar latitudes. In addition, for some models, there is also evidence of decreased variability in the subtropics and marginal convective zones.

The possibility that variance changes in the absence of anthropogenic forcing was also investigated. An F-test performed on only the preindustrial control simulations reveals that most models exhibit significant changes in the variability of annual mean temperature and precipitation in the absence of anthropogenic forcing. However, these changes are not as large as the changes that occur in the presence of anthropogenic forcing. In some models, the largest unforced, centennial-scale changes in variance occur along the equatorial Pacific Ocean, suggesting a connection to ENSO. This finding is consistent with previous studies that have shown significant changes to ENSO variability on centennial timescales in the

absence of anthropogenic forcing (Wittenberg, 2009).

A final area of work also discussed in this dissertation is the validity of common assumptions made about the internal variability of temperature and precipitation. On the one hand, consistency between permutation and Monte Carlo techniques suggests that Gaussian, *iid* assumptions are reasonable for both the seasonal- and annual-mean temperature and precipitation data from these climate models. However, our results also suggest that the practice of doubling uncertainty estimates, as is often done in detection and attribution studies (Hegerl et al. (2007)), may not be sufficient for some models in capturing the amplitude of their variability changes in response to anthropogenic forcing.

The methods and analyses described in this dissertation can be used to investigate changes in the internal variability of other important climate variables such as, sea-level pressure, wind fields, and ocean salinity. Also, these tools can be adapted for various time-means. For instance, changes in decadal-mean variance might reveal additional information about the response of internal climate variability in response to anthropogenic forcing, while 3- to 5-yr means could help to further understand the future of ENSO variability.

Bibliography

Allan, R. P., and B. J. Soden, 2008: Atmospheric warming and the amplification of precipitation extremes. *Science*, **321 (5895)**, 1481–1484, doi:10.1126/science.1160787, URL `http://science.sciencemag.org/content/321/5895/1481`, `http://science.sciencemag.org/content/321/5895/1481.full.pdf`.

Allen, M. R., and S. F. B. Tett, 1999: Checking for model consistency in optimal fingerprinting. *Clim. Dyn.*, **15**, 419–434.

Barnes, E. A., 2013: Revisiting the evidence linking arctic amplification to extreme weather in midlatitudes. *Geophysical Research Letters*, **40 (17)**, 4734–4739.

Boer, G., 2009: Changes in interannual variability and decadal potential predictability under global warming. *Journal of Climate*, **22 (11)**, 3098–3109.

Collins, M., and Coauthors, 2010: The impact of global warming on the tropical Pacific Ocean and el niño. *Nature Geosci.*, **3**, 391–397.

Collins, M., and Coauthors, 2013a: Long-term climate change: projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 12, 1029–1136.

Collins, M., and Coauthors, 2013b: Long-term climate change: Projectsion, commitments, and irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of*

Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 12, 1029–1136.

Coumou, D., and A. Robinson, 2013: Historic and future increase in the global land area affected by monthly heat extremes. *Environmental Research Letters*, **8 (3)**, 034 018, URL `http://stacks.iop.org/1748-9326/8/i=3/a=034018`.

Cover, T. M., and J. A. Thomas, 1991: *Elements of information theory*. Wiley-Interscience, 576 pp.

DelSole, T., and M. K. Tippett, 2015: Laplacian eigenfunctions for climate analysis. *J. Climate*, **submitted**.

DelSole, T., X. Yan, P. A. Dirmeyer, M. Fennessy, and E. Altshuler, 2013: Changes in seasonal predictability due to global warming. *J. Climate*, **27**, 300–311, doi:10.1175/JCLI-D-13-00026.1.

Flury, B. N., 1985: Analysis of linear combinations with extreme ratios of variance. *J. Amer. Stat. Assoc.*, **80**, 915–922.

Francis, J. A., and S. J. Vavrus, 2012: Evidence linking arctic amplification to extreme weather in mid-latitudes. *Geophysical Research Letters*, **39 (6)**.

Hansen, J., M. Sato, and R. Ruedy, 2012: Perception of climate change. *Proceedings of the National Academy of Sciences*, **109 (37)**, E2415–E2423, doi:10.1073/pnas.1205276109, URL `http://www.pnas.org/content/109/37/E2415.abstract`, `http://www.pnas.org/content/109/37/E2415.full.pdf`.

Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon,

D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, 663–745.

Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *Journal of Climate*, **19 (21)**, 5686–5699, doi:10.1175/JCLI3990.1, URL http://dx.doi.org/10.1175/JCLI3990.1, http://dx.doi.org/10.1175/JCLI3990.1.

Huntingford, C., P. D. Jones, V. N. Livina, T. M. Lenton, and P. M. Cox, 2013: No increase in global temperature variability despite changing regional patterns. *Nature*, **500 (7462)**, 327–330, URL http://dx.doi.org/10.1038/nature12310.

Imbers, J., A. Lopez, C. Huntingford, and M. Allen, 2013: Sensitivity of climate change detection and attribution to the characterization of internal climate variability. *Journal of Climate*, **27 (10)**, 3477–3491, doi:10.1175/JCLI-D-12-00622.1, URL http://dx.doi.org/10.1175/JCLI-D-12-00622.1.

IPCC, 2012: Managing the risks of extreme events and disasters to advance climate change adaptation. *A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*, C. Field, V. Barros, T. Stocker, D. Qin, D. Dokken, K. Ebi, M. Mastrandrea, K. Mach, G.-K. Plattner, S. Allen, M. Tignor, and P. Midgley, Eds., Cambridge University Press, 582.

Jaeger, G., 2007: *Quantum information*. Springer.

Jones, G. S., P. A. Stott, and N. Christidis, 2013: Attribution of observed historical nearsurface temperature variations to anthropogenic and natural causes using cmip5 simulations. *Journal of Geophysical Research: Atmospheres*, **118 (10)**, 4001–4024, doi:10.1002/jgrd.50239, URL http://dx.doi.org/10.1002/jgrd.50239.

Kullback, S., 1968: *Information Theory and Statistics*. Dover, 399 pp.

Lau, W. K.-M., H.-T. Wu, and K.-M. Kim, 2013: A canonical response of precipitation characteristics to global warming from cmip5 models. *Geophysical Research Letters*, **40 (12)**, 3163–3169, doi:10.1002/grl.50420, URL http://dx.doi.org/10.1002/grl.50420.

Livezey, R. E., and W. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon Wea. Rev.*, **111**, 46–59.

Power, S., F. Delage, C. Chung, G. Kociuba, and K. Keay, 2013: Robust twenty-first-century projections of el[thinsp]nino and related precipitation variability. *Nature*, **502 (7472)**, 541–545, URL http://dx.doi.org/10.1038/nature12580.

Rhines, A., and P. Huybers, 2013: Frequent summer temperature extremes reflect changes in the mean, not the variance. *Proceedings of the National Academy of Sciences*, **110 (7)**, E546–E546.

Schneider, T., T. Bischoff, and H. Płotka, 2015: Physics of changes in synoptic midlatitude temperature variability. *Journal of Climate*, **28 (6)**, 2312–2331, doi: 10.1175/JCLI-D-14-00632.1, URL http://dx.doi.org/10.1175/JCLI-D-14-00632.1, http://dx.doi.org/10.1175/JCLI-D-14-00632.1.

Screen, J. A., 2014: Arctic amplification decreases temperature variance in northern mid- to high-latitudes. *Nature Clim. Change*, **4 (7)**, 577–582, URL http://dx.doi.org/10.1038/nclimate2268.

Screen, J. A., C. Deser, and L. Sun, 2014: Reduced risk of north american cold extremes due to continued arctic sea ice loss. *Bulletin of the American Meteorological Society*.

Screen, J. A., and I. Simmonds, 2013: Exploring links between arctic amplification and mid-latitude weather. *Geophysical Research Letters*, **40 (5)**, 959–964, doi:10.1002/grl.50174, URL http://dx.doi.org/10.1002/grl.50174.

Tingley, M. P., 2011: A bayesian anova scheme for calculating climate anomalies, with applications to the instrumental temperature record. *Journal of Climate*, **25 (2)**, 777–791, doi: 10.1175/JCLI-D-11-00008.1, URL `http://dx.doi.org/10.1175/JCLI-D-11-00008.1`.

Trenberth, K. E., and Coauthors, 2007: Observations: Surface and atmospheric climate change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Eds., Cambridge University Press, chap. 3, 235–335.

Vecchi, G. A., and A. T. Wittenberg, 2010: El niño and our future climate: where do we stand? *Wiley Interdisciplinary Reviews: Climate Change*, **1 (2)**, 260–270, doi: 10.1002/wcc.33, URL `http://dx.doi.org/10.1002/wcc.33`.

Wittenberg, A. T., 2009: Are historical records sufficient to constraint ENSO simulations? *Geophys. Res. Lett.*, **36**, doi:10.1029/2009GL038 710.

BIOGRAPHY

Emerson Nicole LaJoie received her GED and attended community college before completing her Bachelor of Science from Indiana University in 2005. She received her Master of Science from San Jose State University in 2009, while teaching an introductory weather and climate course. She pursued her Doctor of Philosophy degree with the Climate Dynamics program because this program offered a unique opportunity to work with highly regarded researchers on a range on climate change topics as they emerge from a multiple disciplines.