Detection and Representation of Atmospheric Blocking: An Examination of Block Characteristics

By

MARIELLE C. PINHEIRO
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Atmospheric Science

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Dr. Paul A. Ullrich, Chair

Dr. Shu-Hua Chen

Dr. Michael Wehner

Committee in Charge

2020

This dissertation is dedicated to every teacher and TA who nudged me a little bit further towards the completion of this degree; and to my family, who have always provided unwavering support, even when the homework assignments stopped making sense.

CONTENTS

LIST OF TABLES

<div align="center">ABSTRACT</div>

<div align="center">**Detection and Representation of Atmospheric Blocking:**</div>
<div align="center">**An Examination of Block Characteristics**</div>

The availability of global-scale climate data, both observational and modeled, has provided opportunities to address countless research questions as pertains to the atmosphere, both in terms of understanding the physical processes and exploring model scenarios. The research community has been aided by increasingly powerful computational resources and higher quality data. However, a number of challenges still remain, in terms of both the representation and detection of weather features within gridded data.

This dissertation explores these issues in the context of atmospheric blocking events, which are large-scale, quasi-stationary, high pressure features that divert the jet stream and are subsequently correlated with extreme temperatures and precipitation. Chapter 1 provides an overview of the current literature pertaining to the questions that will be addressed in Chapters 2-4. There are three main topics addressed in this document: the detection of atmospheric blocks and the conditions that lead to block identification; the representation of blocks in a suite of CMIP5 climate models compared to observations, and correlation of model bias and resulting blocking feature bias; and finally, the relative change of blocking representation in the new CMIP6 generation of models as compared to their CMIP5 counterparts. Throughout this research, a number of new metrics and analytical techniques were developed in order to assess the features of detected blocks and compare the blocking results for different objective detection algorithms and models. Using the StitchBlobs package, blocks are detected as individual, 3-dimensional (latitude, longitude, and time) features that contain per-timestep information on each block's extent, persistence, and intensity. Differences are assessed in terms of the resultant distributions of block characteristics, as well as the spatial patterns of the blocks and the underlying characteristics of the model fields.

The results show that there is still much work to be done on all of these topics. A comparison of three objective detection algorithms in Chapter 2 concludes that the choice

of algorithm and the study region will produce significantly different results in blocking statistics. In Chapter 3, the final results show that while biases in blocking frequency can be explained fairly well by regressing the normalized bias values against normalized bias values in block and model characteristics, there are some outlier models, such as HadCM3 and HadGEM2-CC, which cannot be explained by the utilized set of explanatory variable biases. Finally, in Chapter 4, it is shown that, while the CMIP6 models themselves demonstrate some notable improvement in the model characteristics over their CMIP5 counterparts, the relationship between model biases established in Chapter 3 remains largely unchanged. Furthermore, replacing some of the CMIP5 models with their CMIP6 counterparts serves to slightly decrease model accuracy with respect to the explanation of the variance. The analytical methods that have been utilized in this dissertation could be further improved with the inclusion of additional, more complex metrics that better incorporate spatial trends and a better overall definition of blocking.

Chapter 1

Introduction and Background

Atmospheric blocking is a synoptic-scale weather phenomenon with important social and ecological impacts that arise due to its correlation with many kinds of extreme weather, such as heat waves (Pfahl and Wernli, 2012; Grotjahn, 2011; Lee and Grotjahn, 2015), cold spells (Sillmann et al., 2011; Grotjahn and Faure, 2008), and floods (Houze et al., 2011; Hong et al., 2011). Yet it is a phenomenon that is not fully understood in terms of the underlying physics. From initial subjective studies conducted in the mid-1900s to the more powerful objective methods later developed with the aid of computers, atmospheric blocks remain an area of active research because there is no clear consensus on the optimal method for detecting and classifying blocks; this has led to the development of multiple objective blocking detection methods, each with its own biases. While some of the literature has started to display a convergence towards a common blocking climatology, there is not consistent agreement across all detection methods. Studies tend to be heavily focused on a couple of regions, especially the North Atlantic; this is problematic when attempting to apply blocking methods to other regions, such as the North Pacific, which also experience a high frequency of blocking but tends towards blocks with characteristics that are not as easily captured by certain detection algorithms. Furthermore, due to the continuously evolving understanding of the mechanisms that lead to block formation, maintenance, and dissipation, blocking climatologies also vary across individual climate models. Generally, climate models do not yet produce a blocking climatology that is consistent with observations; North Atlantic blocking is notoriously underrepresented in climate models,

1

and blocking in other sectors is also lower than observed, although frequency biases are seasonally and regionally dependent (Schiemann et al., 2016). Possible solutions include increased spatial resolution(Schiemann et al., 2016), correcting underlying biases in model mean state (Scaife et al., 2011), and improved representation of orography (Berckmans et al., 2013) or physical parameterizations (Martinez et al.). However, the aforementioned regional dependency of block characteristics means that corrections to one source of model blocking error will not necessarily fix overall blocking representation.

The following sections will provide context for subsequent chapters. Section 1.1 defines atmospheric blocking and describes some of the observed regional and seasonal trends in blocking climatology, as well as discussing some of the potential causes for block formation. Section 1.2 discusses the representation of blocks in climate data, both in terms of algorithmic definition and blocking in climate models.

1.1   A brief overview of atmospheric blocking

The American Meteorological Society (AMS) definition of blocking (Glickman, 2000) uses three criteria for classifying a flow pattern as blocked:

1. persistent obstruction of the normal west-to-east flow pattern,

2. pronounced meridional flow in the upper levels, and

3. anticyclonic circulation at high latitudes accompanying cyclonic circulation at low latitudes.

The onset of a blocking feature results in temporary redirection of the jet stream, which is in turn responsible for the aforementioned anomalous weather conditions. Blocks take several forms, including the well-known omega block, a poleward high co-located between two equatorward lows; high/low dipoles; and persistent ridges, and often take on more than one form during their life cycles. All of these varieties nonetheless satisfy the criteria established by the AMS definition.

1.1.1 Formation and maintenance of atmospheric blocks

In their examination of geopotential height anomalies, Dole and Gordon (1983) showed that the probability of anomaly persistence increases with time up until five days, after which the probability of persistence is nearly constant; this, along with later papers, have found that five days is an acceptable minimum threshold for blocking persistence. However, despite the body of literature that examines the existence and effects of blocking, the causes are less clear.

A common theme of blocking formation theories involves interference of large-scale waves. The phase speed for westward propagation of barotropic Rossby waves has latitudinal dependence via the β parameter:

$$c = -\frac{\beta}{k^2} \tag{1.1.1}$$

where $\beta = \frac{2\Omega \cos \phi_0}{a}$ decreases in the poleward direction. Thus, Rossby waves with the same wavenumber that propagate at lower latitudes will have a larger phase speed. At small scales, latitudinal variations can be neglected; however, they are significant when one considers weather phenomena of the scale on which blocking occurs. For example, Grose and Hoskins (1979) modeled the impact of orography on zonal flow, and showed that the interaction of the resultant Rossby wavetrains, produced northwards and southwards of the mountain feature, led to downstream blocking where the waves were out of phase due to latitudinal variations. Hoskins and Karoly (1981) found that tropical sea surface temperature (SST) anomalies can produce similar interactions; upper-level divergence and vorticity anomalies that arise from SST anomalies form stationary Rossby wavetrains as in the topography case. Here, the smaller wavenumbers propagate strongly poleward while the larger wavenumbers remain confined to lower latitudes, resulting in the split wavetrains. However, the influence of SST anomalies displays regional variations. Mullen (1989) demonstrated that tropical warm SST anomalies enhance blocking over the Northern Pacific but suppress it over the central Pacific, and extratropical cool SSTs similarly suppress central Pacific blocking; while SST anomalies, both tropical and extratropical, had little to no effect on the Atlantic basin at all.

Another important component of blocking involves forcing by synoptic-scale transient eddies, which are one of the main drivers of extratropical weather. Eddy vorticity helps maintain the anticyclonic structure of a block against the mean westerly flow, and the eddy straining mechanism proposed by Shutts (1983) states that the deformation of eddy vorticity leads to an energy transfer which maintains and strengthens the split flow in the vicinity of the block.

1.1.2   Seasonality and interannual variability

Generally, the winter and spring seasons of the respective hemispheres display a higher frequency of blocking, although there are regional variations (Lupo and Smith, 1995); and the blocks that occur during these seasons are also more intense (Dole and Gordon, 1983). Recurring seasonal and multiannual patterns are shown to have influences on the various factors which affect the dynamics surrounding blocking. For example, Barriopedro et al. (2006) found that some of the less frequent climatological oscillations, such as the NAO, explained as much as 45% of the variance in blocking trends in the Northern Hemisphere; when the NAO is in the negative phase, blocking events are more likely to be longer and more frequent. Another major influence for the cold seasons is the positive phase of the Scandanavian (SCAN) pattern, which accounts for 30% of blocking variance further eastward over Europe and results in increased frequency and a northeastern shift of the frequency maxima center. In the Pacific, the Western Pacific (WP) pattern accounts for 25% of spring blocking variability and the and Eastern Pacific (EP) accounts for 35% of winter variability, with the negative phase of each pattern correlated to increased blocking frequency. Additionally, La Niña leads to a westward shift to blocking patterns in the Pacific compared to El Niño.

1.1.3   Geographical trends

There is extensive literature on favored locations for block formation and persistence (Treidl et al., 1981; Tibaldi et al., 1994; Wiedenmann et al., 2002; Barriopedro et al., 2006), with patterns of blocking maxima displaying a few distinctive trends. The Northern Hemisphere is more prone to block formation than the Southern Hemisphere, which is

likely due to a greater proportion of land mass, leading to increased drag and subsequently a jet stream with more curvature. Also notable is the fact that blocks are prevalent on the east side of ocean basins and the west coasts of continents; blocks over land masses, while not uncommon, are less frequent and generally not as intense as those located over the ocean (Dole and Gordon, 1983).

In the Northern Hemisphere, the Atlantic (centered about 10W-30E) and Pacific (centered about 180) basins are generally the most favored locations for blocking due to storm track activity downstream of North America and Asia, respectively. A third, continental region located over Eurasia shows enhanced activity during JJA that is linked to the Mediterranean Sea (Whittaker and Horn, 1981). The circumstances for blocking formation have been shown to differ between the Atlantic and the Pacific. For example, Nakamura et al. (1997) proposed that Pacific blocking is caused mainly by eddy forcing, as opposed to the Atlantic basin, where existing quasi-stationary Rossby waves are found to be an important ingredient in block formation. Similarly, Altenhoff et al. (2008) performed an analysis utilizing PV-streamers, and proposed that the Atlantic basin experiences weaker surface baroclinicity which leads to stronger anticyclone formation, as opposed to the more strongly baroclinic Pacific basin. Dunn-Sigouin et al. (2012) finds that Euro-Atlantic blocking is correlated with geopotential height anomaly dipoles over the Atlantic, which is similar to the negative phase of the NAO; Pacific blocking events are associated with a wave train pattern that resembles the negative phase of the PNA.

Southern Hemisphere blocks tend to be shorter in duration than their Northern Hemisphere counterparts, with the Pacific blocks (maximum blocking frequency at about 120W) and near Australia (from 140E-180) showing the greatest persistence. Wiedenmann et al. (2002) suggests that the relatively lower frequency of Southern Hemisphere blocking is possibly due to a lower frequency of favorable interactions between synoptic- and planetary-scale waves, as well as less frequent cyclone development which enhances or maintains blocks.

1.2 Quantification and representation of blocking

1.2.1 Objective detection of atmospheric blocks

Blocking has been studied for decades, but the first attempts at finding blocks required visual inspection of flow patterns and were thus limited in scope (Rex, 1950). Such subjective assessment leads different individuals to potentially different conclusions about a flow pattern. The development of an objective procedure for identifying blocks is thus important for a number of reasons:

1. **Consistency:** While the development of objective methods requires some preliminary human judgment in the choice of parameters, an algorithmic definition of blocking removes the human subjectivity from the rest of the procedure and thus produces results that are internally consistent for a given dataset.

2. **Efficiency:** These methods can be automated, making computation and comparison of blocking climatologies possible across very large volumes of high-resolution, multi-decadal data.

3. **Improved scientific understanding:** Algorithms are based on current concepts of block formation and maintenance. Objective detection methods can allow for these concepts to be rigorously tested and improved as more information is gathered.

Objective methods based on a variety of fields and techniques have been developed over the years (see Figure 1 of Barriopedro et al. (2010) for an overview), and intercomparison studies such as Barnes et al. (2012) have explored whether these methods produce consistent blocking climatologies. Barnes et al. (2012) compares three longitudinally-varying (1D) methods (i.e. calculated about a single time-varying central latitude): Pelly and Hoskins 2003 (potential temperature (θ) on a constant potential vorticity surface), Tibaldi and Molteni 1990 (500 hPa geopotential height ($Z500$) gradient over a latitudinal range), and Scaife et al. 2010 (zonal wind over a latitudinal range). The analysis, which was performed on 43 years of Northern Hemisphere data, concludes that these methods yield similar results in terms of calculated blocking frequency and duration across the

time and longitude axes. However, these are only two possible metrics under which objective methods can be examined, and other papers have noted inherent differences in the methods due to both the data and the chosen method. For example, Davini et al. (2012) notes that there are distinct regional differences in both the geopotential height fields and the resultant characteristics of detected blocks. Over Greenland, blocks principally correspond to cyclonic Rossby wave breaking with a dipole structure, and split-flow blocking generally happen in the midlatitudes over central Europe. The structure of a block impacts the effectiveness of the detection method; Scherrer et al. (2006) compared detection of an omega block versus a persistent ridge, using the aforementioned $Z500$ gradient method of Tibaldi and Molteni (1990) as well as two potential vorticity (PV)-based metrics. All three detection methods produced similar results for the omega block, but displayed notable difference in both the size and center locations of the blocked areas for the ridge. The authors attribute the differences to both the choice of the variable ($Z500$ versus PV) and the use of an anomaly versus total field.

In Chapter 2, the importance of a robust blocking detection methodology is highlighted by comparing the blocking climatologies and block characteristics of features detected by three different objective detection algorithms.

1.2.2 Representation of blocking in climate models

Beyond differences introduced by the methodology, blocking results are also altered by the characteristics of the input dataset. An assessment of Northern Hemisphere blocking in CMIP5 data (Masato et al., 2013), where the historical and RCP8.5 experiments were compared to the 40-year ECMWF reanalysis (ERA-40) using both 1D and 2D blocking detection methods, notes that the mean of historical models greatly underestimates wintertime blocking compared to the reanalysis, particularly over Europe. Such underprediction has been noted elsewhere, both in terms of short-term forecasting (Tibaldi and Molteni, 1990) and other uses of CMIP data (Dunn-Sigouin and Son, 2013; Davini and D'Andrea, 2016). Models tend to underestimate both the number and persistence of blocking events, which both contribute to the reduced blocking frequency noted in the models.

The difference in blocking within climate models has been associated with many underlying model biases. For example, Scaife et al. (2010) finds that blocking representation in models is improved by correcting biases in the winter-mean geopotential height. Coupled ocean-atmosphere models tend to exhibit higher biases, due to the additional influence of sea surface temperature biases Scaife et al. (2011). Increasing the model resolution can reduce the bias in SST and therefore reduce the resultant bias in blocking frequency, although improvements are largely seen in the Atlantic. In fact, improvements in one region might worsen blocking representation elsewhere. For example, while Schiemann et al. (2016) noted a marked improvement in Euro-Atlantic blocking with increased spatial resolution, there was actually a detrimental effect to the representation of Pacific blocking in the CAM5 model. Berckmans et al. (2013) found similar trends in the representation of blocking with respect to the Atlantic versus the Pacific in the Northern Hemisphere. They observed that increased orographic resolution led to higher topography in the Rocky Mountains, resulting in increased suppression of storm tracks and the jet, a more intense SW-NE jet tilt in the Atlantic, and increased blocking activity downstream. Pacific blocking, however, is influenced by the topography of the Tibetan Plateau, which does not change as dramatically between the low and high resolution models and thus does not lead to the same magnitude of improvement; other factors, such as the representation of the Pacific jet, exhibited a stronger influence on the resultant blocking climatology in the Pacific basin. Dunn-Sigouin et al. (2012) observes that improvements in the representation of Pacific blocking appear to be associated with improved model physics rather than higher spatial resolution.

1.2.3 Methods of quantifying blocking results

Many blocking studies assess blocking on a per-gridpoint basis rather than tracking blocks as unique features, which limits the metrics that can be used to assess model performance. The most commonly measured quantities in blocking studies are blocking frequency and block duration, due to their ease of assessment using per-gridpoint tracking. In gridded fields, blocking frequency is defined as a per-gridpoint average of blocked days (blocked versus not blocked) over the extent of the study period, which does not require knowledge

of individual block extents; duration, as defined by studies such as Barnes et al. (2012), involves searching for consecutively blocked longitudes and then identifying regions of persistent blocking over time. Other papers have created additional metrics to assess blocking characteristics, such block intensity (Lupo and Smith, 1995; Wiedenmann et al., 2002), speed (Sinclair, 1996), and size (Croci-Maspoli et al., 2007).

1.3 Outline of Thesis

This thesis addresses three main topics, as presented in the previous sections: block detection, block representation, and assessment of block characteristics. Chapter 2 presents an intercomparison of three objective detection algorithms: Dole and Gordon 1983, a geopotential height anomaly-based algorithm; Tibaldi and Molteni 1990, a geopotential height gradient-based algorithm; and Schwierz et al 2004, a potential vorticity anomaly-based alorithm. This chapter examines the differences in resultant blocking climatologies for each of the three algorithms in terms of block characteristics and the underlying flow patterns that lead to these differences.

Chapter 3 examines compares blocking between reanalysis datasets and the CMIP5 climate model ensemble, both in terms of the averaged blocking climatologies and the characteristics of the blocks produced within the models. These differences are connected to differences in the underlying physical fields within the models. Chapter 4 builds upon this analysis by comparing a subset of the CMIP5 models with corresponding CMIP6 models.

Chapter 2

Intercomparison of Objective Blocking Algorithms

This chapter expands upon previous intercomparison efforts; blocking is assessed in terms of distinct blocking features rather than per gridpoint, and each algorithm is applied across the full latitude-longitude (2D) range of the study regions, which include both the Northern (NH) and Southern (SH) Hemisphere midlatitudes. Assessing algorithmic differences via individual blocking events allows determination of block characteristics beyond blocking frequency: here, we consider the size, duration, distance traveled, and zonal speed of each block as determined by each algorithm. We choose to utilize 2D rather than 1D blocking indices in order to more fully examine regional variations in blocking; most notably, low-latitude blocking is often missed by 1D methods. Furthermore, it demonstrates how these algorithms perform in regional climatologies outside of those for which they were developed. In particular, attention is paid to the underlying flow patterns that lead to differences in objective blocking climatologies. Our analysis shows that each of the assessed algorithms only capture a subset of meteorological patterns defined by the AMS definition of blocking, and the level of agreement between algorithms is highly dependent on region and block type. This an important point to consider when attempting to assess current and future blocking trends and the impacts of corresponding extreme weather. A further benefit of this study is that the metrics and algorithms developed through this work may be leveraged for evaluation of global climate datasets, either from

individual model runs or from coordinate intercomparison efforts.

Section 2.1 outlines the three objective detection algorithms and the analysis framework, which was developed with the goal of standardizing the detection methodology as much as possible across the algorithms. Section 2.2 compares results between the three algorithms in terms of both the averaged and instantaneous blocking patterns, as well as some of the characteristics of the detected blocks. Section 2.3 explains some of the meteorological factors which influence the algorithms' results, and Section 2.4 summarizes and discusses the implications of differences between algorithm results.

2.1   Data and Methodology

2.1.1   Data

Our dataset is the ERA-Interim reanalysis from the European Center for Medium-Range Weather Forecasts (Dee et al., 2011). Temperature, meridional and zonal wind, and geopotential variables are 6-hourly at 1-degree spatial resolution in the time period of March 1, 1979- February 28, 2018 (39 years). The latitude range employed spans 25-75 degrees in each hemisphere, and the longitudinal extents of each region are outlined in Table 2.1; the abbreviations in the table will be used for each region hereafter. These regions are based on the suggested ranges in Wiedenmann et al. (2002); each region is roughly centered over a local maximum of blocking frequency.

	Continent (NC)	Pacific (NP)	Atlantic (NA)
NH	40E, 140E	140E, 100W	100W, 40E
	Indian Ocean (SI)	**Pacific (SP)**	**Atlantic (SA)**
SH	30E, 130E	130E, 60W	60W, 30E

Table 2.1: Longitudinal extents of study regions; each region has a latitudinal extent of 25-75 degrees in their respective hemispheres. The regions can be seen outlined on the maps in Figures 2.2-2.6. The two-letter abbreviations will be used to refer to these regions throughout the paper.

Two of the methods (Tibaldi and Molteni (1990), hereafter referred to as TM90; and Dole and Gordon (1983), hereafter referred to as DG83) are based on the $Z500$ variable,

11

Initial Calculations

PV^* and Z^*: Look for $\widetilde{V^*} \geq 1$

$\overline{V}$: LTDM V $\rightarrow$ $V^* = V - \overline{V}$

$\alpha = 1.5 \times \sigma_{V^*}$ or α_{min} $\rightarrow$ $\widetilde{V^*} = \frac{V^*}{\alpha}$

AGP: Satisfy $GHGN$ and $GHGS$ criteria

$GHGN = \frac{Z(\phi_n) - Z(\phi_0)}{\phi_n - \phi_0}$

$GHGS = \frac{Z(\phi_0) - Z(\phi_s)}{\phi_0 - \phi_s}$

NH: $AGP = 1$ if $GHGN < -10$ m/deg lat $GHGS > 0$ m/deg lat

SH: $AGP = 1$ if $GHGN < 0$ m/deg lat $GHGS > 10$ m/deg lat

StitchBlobs

Inputs

file list | variable | region constraint | time/size constraints

Per time step (instantaneous blocking):

Identify points where variable ≥ 1 $\rightarrow$ Determine clusters (nearest neighbor) $\rightarrow$ Eliminate clusters smaller than min size

Create 3D (t, x, y) blobs: Over time axis, join clusters that overlap by at least 1 grid box

Per blob

Eliminate blobs persisting less than min duration $\rightarrow$ Assign a unique identifier $\rightarrow$ **Per time step:** find min/max/center lat/lon values, area

Figure 2.1: Schematic of workflow for blocking calculations using StitchBlobs

while the third method (Schwierz et al. (2004), hereafter referred to as S04) is based on vertically averaged PV (VPV). The $Z500$ fields are derived from the geopotential variable ($Z = \Phi/g$). From temperature and the horizontal wind components, we calculated Ertel PV (EPV, described in Appendix A.1) and then averaged this over the 150-500 hPa layer to produce VPV.

2.1.2 Blocking Detection Methods

The blocking climatology varies with the choice of detection scheme, as shown below. In order to explore some of the points raised by Davini et al. (2012) and Scherrer et al. (2006)—particularly the differences due to variable choice and region— we utilize schemes that are based on two different variables ($Z500$ and VPV) and field types (anomaly-based versus total field).

A standardized analysis framework, StitchBlobs, was developed for the intercompar-

ison of global blocking detection schemes. The blocking detection workflow is outlined in Figure 2.1, and further details on StitchBlobs, which is part of the TempestExtremes package (Ullrich and Zarzycki, 2017), are provided in Appendix A.2. Instantaneous blocks that meet the spatial constraint (minimum area of 10^6 km^2) are stitched together across time into distinct blocking events. StitchBlobs identifies events that fit the minimum time constraint (5 days, the characteristic length of persistent height anomalies as per DG83), then tags each event with a unique identifier and provides per-time-step information on each block's location (in terms of block center latitude/longitude coordinates) and size (in terms of either maximum latitude/longitude extent or the area of the cluster). This allows users to follow individual blocking events from formation to dissipation, and examine seasonal and regional trends in block characteristics such as size and zonal distance traveled on a per-block basis.

The methods are briefly summarized in Table 2.2, and explained in greater detail in the following sections. In order to differentiate between the methods from the original papers and the ones presented here, we use different abbreviations: our version of TM90 is AGP, DG83 is Z^*, and S04 is PV^*.

2.1.2.1 Geopotential Height Gradient (AGP)

The most frequently cited blocking detection method in the literature is TM90, which is itself based on Lejenäs and økland (1983). Two gradients are calculated about a central latitude as follows:

$$GHGN = \frac{Z500(\phi_n) - Z500(\phi_0)}{\phi_n - \phi_0} \tag{2.1.1}$$

$$GHGS = \frac{Z500(\phi_0) - Z500(\phi_s)}{\phi_0 - \phi_s} \tag{2.1.2}$$

where ϕ_0, ϕ_n, and ϕ_s are the reference latitude and the latitudes 20° above and below ϕ_0, respectively, and $GHGN$ and $GHGS$ are the height gradients. For $GHGN < -10$ m/deg lat and $GHGS > 0$ m/deg lat, the point is considered instantaneously blocked; the negative $GHGN$ and positive $GHGS$ values imply a large-scale high in the 500 hPa geopotential height field.

TM90 performed these calculations about a single reference latitude band (60°N), and

Method	Tibaldi and Molteni 1990 (AGP)	Dole and Gordon 1983 (Z^*)	Schwierz et al. 2004 (PV^*)
Variable	$Z500$ latitudinal gradient ($GHGN$, $GHGS$)	Z500 anomaly ($Z500^*$) with latitudinal scaling factor	Vertically averaged PV anomaly (VPV^*)
Detected Feature	Change in $Z500$ 15° above/below point, implying presence of high	Positive $Z500^*$ with respect to climatological mean	Reversal of flow with respect to climatological mean
Original blocking criteria	GHGN<-10 m/deg lat and GHGS>0 m/deg lat over 4 days, per gridpoint	$Z500^* \geq 100$ m over 10 days, per gridpoint	$VPV^* \leq$ -1.2 PVU, with at least 70% overlap between contours over 5 days
Change to original method	Extend analysis to all latitudes within 25-75°, 5 days' persistence with contour overlap	Varying anomaly threshold, 5 days' persistence with contour overlap	Varying anomaly threshold with positive sign for SH

Table 2.2: Summary of original methods and modifications.

Barnes et al. (2012) performed these calculations about a varying central latitude. The TM90 algorithm was modified by Scherrer et al. (2006) to extend analysis to latitudes 35-75°N and define ϕ_n and ϕ_s as 15° away from ϕ_0; we further extend the analysis to 25° in both hemispheres. Furthermore, to apply this method in the SH, it is necessary to switch the criteria and signs for the two gradients, since the orientation of ridges is flipped and the SH latitudes are negative. Therefore, in AGP, $GHGN < 0$ m/deg lat and $GHGS > 10$ m/deg lat in the SH.

In TM90, a blocking episode is defined as a region of blocked flow that extends over at least 12 degrees longitude for a minimum of 4 days. This satisfies the second and third points of the AMS definition (meridional flow and anticyclonic circulation); however, the fact that this method is based on total fields means that it does not necessarily satisfy the first point (obstruction of normal flow), since there is no reference to the mean climatology.

2.1.2.2 Geopotential Height Anomaly (Z^*)

DG83 utilizes $Z500$ anomaly ($Z500^*$), which is first calculated as the height departure from the long term seasonal average, h', then normalized by a latitudinal coefficient:

$$Z500^* = \left(\frac{\sin 45°}{\sin \phi}\right) h'. \tag{2.1.3}$$

This modification is necessitated by the latitudinal change in planetary vorticity; the conservation of absolute vorticity means that there must be an increase in relative vorticity due to the decrease in planetary vorticity in the poleward direction. At the higher latitudes, the convergence of latitudes leads to a bias in the representation of meridional energy propagation.

A single grid point is defined as blocked if $Z500^*$ exceeds 100 m for 10 days, although subsequent papers have used different combinations of heights and durations (for a 5-day minimum duration, Sausen et al. (1995) used 250 m). As with TM90, this detection method works by searching for high geopotential heights, although in this case the high is defined with respect to the long term average. DG83 is theoretically capable of satisfying the AMS criteria for blocking because anomalously high $Z500$ will modify the flow pattern in a manner consistent with all three requirements. With that said, the relationship

between the climatological mean and the instantaneous field can lead to overprediction of blocking (particularly in the SH), as discussed in Sections 2.3.1 and 2.3.3.

2.1.2.3 Potential Vorticity Anomaly (PV^*)

S04 proposes a blocking detection method which entails searching for regions of persistent column-averaged (150-500 hPa) negative PV anomalies (VPV^*) in the NH (in PV^*, the relevant anomalies are positive in the SH). As with DG83, anomalies are calculated as instantaneous departures from a long term daily mean (defined as the 15-year monthly mean in S04). S04 favors the use of VPV over $Z500$ for anomaly-based detection, as VPV^* more closely follows the shape of the dynamic tropopause, compared to a similarly situated $Z500$ anomaly (Figures 1b and 2b in S04). While S04 does not explicitly account for easterly flow as in TM90, the negative sign indicates anticyclonic circulation at and below the layer with easterlies countering the mean flow, thus signifying underlying higher pressure as in point 3 of the AMS definition. The use of VPV^* also accounts somewhat for parts 1 and 2 of the AMS definition, but strongly negative (positive) values of vorticity in the NH (SH), discussed in Section 2.3.2, or relatively low values of VPV with respect to the climatological mean, discussed in Section 2.3.3, can cause this method to mistakenly identify unobstructed flow as blocked.

2.1.2.4 Modifications to anomaly methods

For the purposes of global intercomparison, a few minor modifications are needed for the anomaly-based detection methods. Since DG83 and S04 use different definitions of climatological means and thresholds, it is necessary to redefine these quantities using a consistent methodology rather than those of the original algorithms. The choice of threshold is an important consideration when applying a method to a new dataset; as DG83 and S04 were initially developed using data from NA DJF, their hardcoded thresholds are not directly applicable in other sectors. The applicability of threshold values to new datasets is also an important consideration for multi-model intercomparisons; Woollings et al. (2018) found that the timing and spatial distribution of projected changes in blocking using CMIP5 model output were dependent on region and methodology. Therefore, a constant threshold that was determined using one model may not be representative of
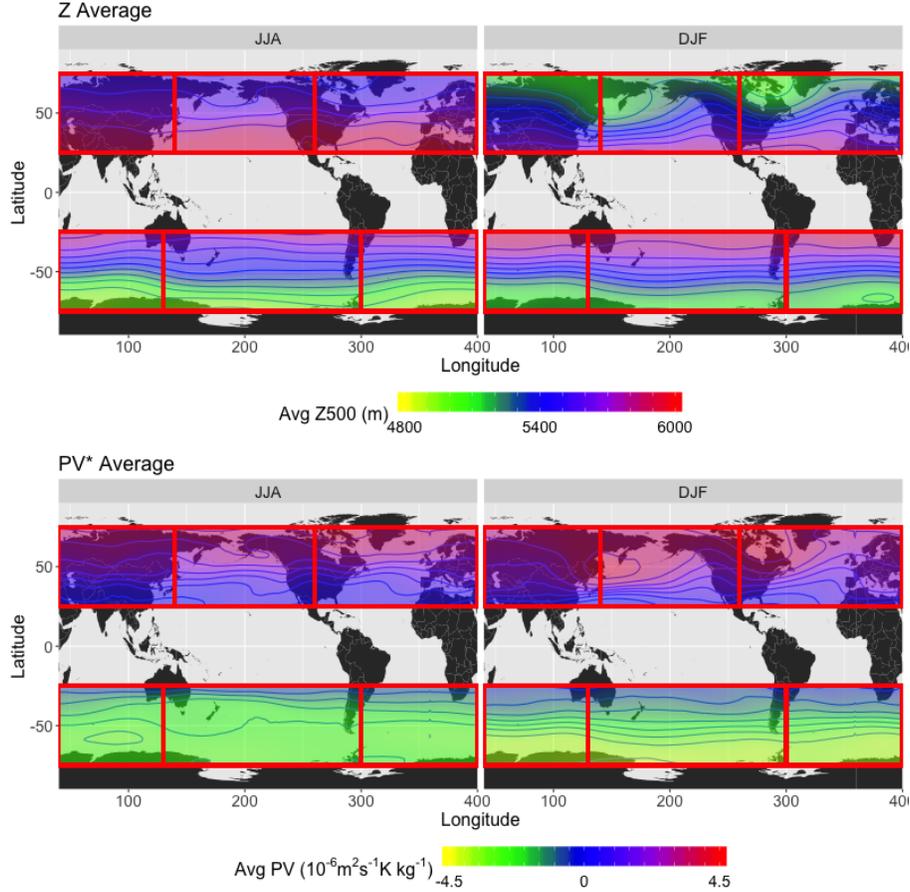
Figure 2.2: Seasonal averages of long term daily mean Z500 (top) and vertically averaged PV (bottom) values for (left) JJA, and (right) DJF, ERA-Interim 1979-2018. Each seasonal average contains 39 years' worth of data. Red rectangles denote the study regions as outlined in Table 2.1; from left to right, the NH regions are NC, NP, and NA, and the SH regions are SI, SP, and SA. The contours are in intervals of 100 m for Z500 and 0.5 PVU for PV.

the climatology of another model. For a unified global study such as this one, a constant threshold definition will lead to either under- or over-detection of blocks in other regions because the anomaly thresholds are calibrated to the climatology of that region and time period. Here we follow Barriopedro et al. (2010) and Dunn-Sigouin et al. (2012), who address this problem by replacing the constant threshold definition with one derived from the standard deviation of anomaly values.

Long term daily mean and anomalies: For this study, mean fields are computed for each of the 365 days in a year (excluding leap days) across 26 years. The values on these
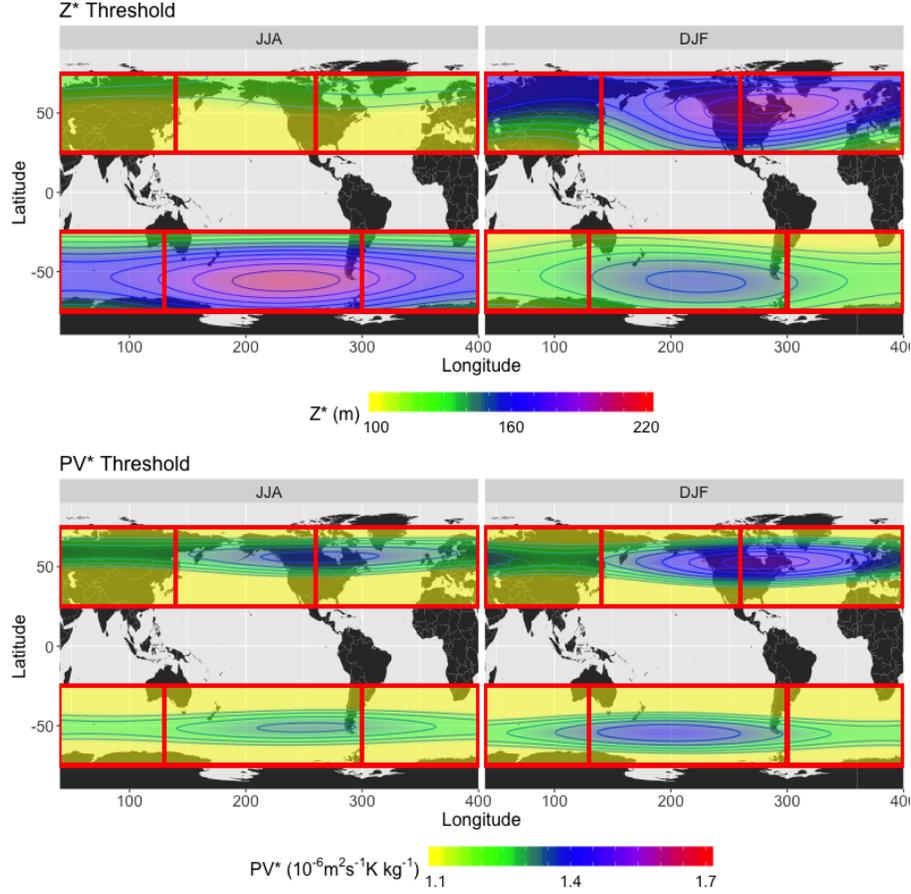
Figure 2.3: Seasonal averages of (top) Z500 anomaly ($Z500^*$) and (bottom) vertically averaged PV anomaly (VPV^*) threshold values, as described in Section 2.1.2.4, for (left) JJA and (right) DJF. Red rectangles denote same regions described in Figure 2.2 caption. The contours are in intervals of 10 m for $Z500^*$ and 0.05 PVU for VPV^*.

365 days are Fourier transformed and only the first 6 harmonics (0-5, with 0 corresponding to the mean and 5 corresponding to a 73 day span, or slightly less than a season) are used to back transform and obtain the long term daily mean (LTDM, denoted here as $\overline{V}$, the average of instantaneous variable V). This procedure is further explained in Grotjahn and Zhang (2017); it serves to smooth out the mean, which can otherwise have excessive day-to-day variations. Figure 2.2 shows seasonal averages of the LTDM for the $Z500$ and VPV fields. Anomalies from the LTDM, V^*, can thus be defined as

$$V^* = V - \overline{V} \tag{2.1.4}$$

18

Threshold and normalized anomalies: Whereas DG83 and S04 use constant minimum anomaly value threshold values in their blocking definitions (thresholds of magnitude 100 m and 1.2 PVU, respectively), we calculate a threshold field that varies in both space and time. At each grid point in the relevant field, we calculate the standard deviation of that grid point's time series; the threshold is defined as 1.5 times the standard deviation. These values are then smoothed in both the zonal and meridional directions in a manner similar to the LTDM computation. This procedure is then repeated across time with the first 6 harmonics as with the LTDM calculation. However, an additional minimum threshold criterion, α_{min} (100 m for Z^* and 1.1 PVU for PV^*), is imposed. This is necessary for regions with little to no blocking activity, where the anomaly values (and thus the standard deviation and threshold values) are very low. Using the standard deviation of the anomalies for each day over the course of the 26-year time period, σ_{V^*}, threshold α is defined as

$$\alpha = \begin{cases} 1.5 \times \sigma_{V^*}, & \text{if } \alpha \geq \alpha_{min} \\ \alpha_{min}, & \text{otherwise} \end{cases} \qquad (2.1.5)$$

The range of threshold values as defined by the standard deviation is highly dependent on region and season, as apparent in Figure 2.3. DG83 noted that the distribution of persistent $Z500^*$ values varied from region to region in the wintertime, with standard deviations of 170-180 m for the Atlantic and Pacific regions, as opposed to 45 m for Eastern Asia; it is evident that using a constant threshold on a global analysis is not advisable for Z^*. The distribution of VPV^* values, while less drastically variable with respect to the magnitude of values compared to $Z500^*$, also displays some regional and seasonal differences with respect to anomaly magnitudes, particularly when comparing summer to winter. The spatiotemporally varying, standard deviation-based definition ensures that only anomaly values that are at the tails of the local distribution at that particular time step will be classified as blocked, irrespective of region or season.

The normalized anomalies, $\widetilde{V^*}$, are calculated as

$$\widetilde{V^*} = \frac{V^*}{\alpha} \qquad (2.1.6)$$

19

and areas with $\widetilde{V^*} \geq 1$ are considered blocked.

2.1.3 Quantification of agreement between methods

Three metrics are used for quantifying agreement between methods: pattern correlation between the seasonally-averaged blocking patterns, probability of co-occurrence, and block spatial similarity. We utilized multiple methods in order to separate out seasonally averaged versus instantaneous agreement; pattern correlation is useful for assessing the blocking climatology, but similarity and probability metrics provide further insight into the agreement between methods with regards to individual blocks. The quantification of probability and similarity serve similar purposes, but inevitably are different metrics for comparing detection algorithms. Two methods could have a high probability of co-occurrence if they consistently detect the same features, but a lesser value of similarity if the resultant clusters produced by the two methods are very different in size and shape.

Pearson pattern correlation: This metric, denoted by $C(M1, M2)$, measures the strength of linear relationships between frequency values at corresponding coordinates for methods $M1$ and $M2$. Higher correlation is seen when patterns are more similar, regardless of the relative magnitudes of the two data points. Pattern correlation has a range of possible values from -1 to 1, in which magnitudes of 0.3 and below are considered weak, and strong at 0.7 and above. Negative values indicate an inverse relationship between $M1$ and $M2$. Centered pattern correlation is computed using the NCL `pattern_cor` function with cosine latitude weighting.

Probability of co-occurrence: This metric, denoted by $P(M1|M2)$, quantifies the likelihood that a block will be detected by method $M1$, given that method $M2$ also detects it. The methodology for calculating probability of co-occurrence can be found in Appendix A.3.

Spatial similarity: This metric, denoted by $S(M1, M2)$, quantifies the match between areas designated as part of a block by methods $M1$ and $M2$. S is the intersection divided by the union of the two areas. Since S varies for each block, an interquartile range will be presented in the results. The methodology for calculating spatial similarity can be found

in Appendix A.4.

2.2 Evaluation and intercomparison of atmospheric block metrics

We begin with a single case study, to highlight the importance of selecting a method, then present results for the entire ERA-Interim period in subsequent sections. The full ERA-Interim results for JJA and DJF are presented as follows. First, each algorithm is assessed in terms of its individual statistics (averaged blocking climatology, block duration, size, distance traveled, and zonal speed measurements). Agreement between detection methods (in terms of both averaged and instantaneous detection of blocks) is also assessed. Results are considered statistically significant if there is a meaningful relationship in the data that could not have occurred from random chance or sampling error alone. The methodology for establishing statistical significance for all of the metrics in Section 2.2 is explained in Appendix A.5. These metrics demonstrate that agreement in blocking frequency does not imply consistency in the character of individual blocks. Namely, one needs to be careful when drawing conclusions based on the blocking climatology alone, since there may be significant differences in the meteorological character of individual blocks. Further, these results are indicative that conclusions drawn with one blocking detection scheme may not hold when another algorithm is used.

2.2.1 Case study: The Ridiculously Resilient Ridge

We present a case study of a persistent and pronounced ridge pattern that repeatedly appeared off the western coast of North America in late 2013, then reoccurred during the winters of 2014-2015 and 2015-2016. This feature, dubbed the "Ridiculously Resilient Ridge" (or RRR for short) by Swain et al. (2014), was responsible for redirecting moisture-heavy air northwards during the winter months. Because California receives the bulk of its precipitation from December to March, the RRR was a key player in the drought that devastated the state for almost 6 years.

Figure 2.4 shows the frequency of detected blocking in NP DJF from December 2012 to February 2016, and Table 2.3 provides the corresponding agreement metrics. Out of

Method pair	Correlation	Probability	Similarity
PV^* and AGP	-0.01	$P(PV^*|AGP)$: 0.32** $P(AGP|PV^*)$: 0.40**	*0.11* to 0.48**
Z^* and AGP	0.06*	$P(Z^*|AGP)$: **0.78** $P(AGP|Z^*)$: 0.35**	0.23 to 0.47**
PV^* and Z^*	0.52**	$P(PV^*|Z^*)$: 0.32** $P(Z^*|PV^*)$: **0.90**	**0.30** to 0.51**

Table 2.3: Pattern correlation, co-occurrence, and spatial similarity for RRR data (DJF, December 2012-February 2016). Probability values above 0.7 are bolded. Similarity values are formatted based on relative distributions of values—25th percentile values above (below) 0.29 (0.15) and 75th percentile values above (below) 0.54 (0.39) are bolded (italicized) to denote particularly high or low quantities relative to other values. Statistical significance is denoted by a "*" for $0.01 < p < 0.05$ and "**" for $p < 0.01$.
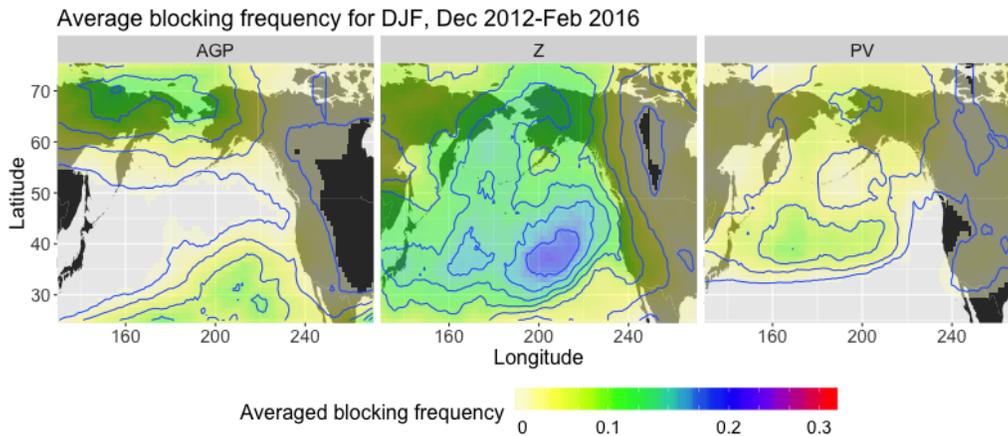
Figure 2.4: Blocking frequency, averaged over winters from 2012-2016, for the AGP method (left), Z^* method (center), and PV^* method (right)

22
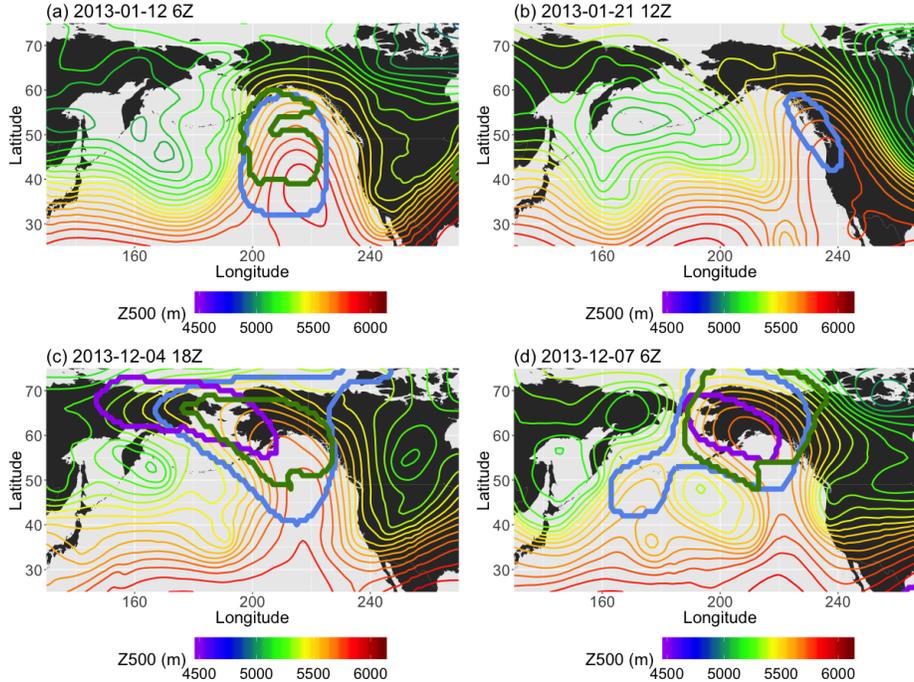
Figure 2.5: (a-b) Examples of ridges that were detected by only the anomaly methods (Z^*, blue, and PV^*, green) in January 2013. (c-d) Examples of ridges that were detected by all three methods (including AGP, purple) in December of 2013.

the three methods, the one that most closely coincides with the location of the RRR—in terms of a blocking maximum centered off of the West Coast—is the Z^* method. The PV^* method displays an inclination towards contours that co-locate with Z^* contours ($P(Z^*|PV^*) = 0.90$), but the maximum is positioned further northwards and there is only a modest correlation of $C(PV^*, Z^*) = 0.52$. Due to the PV^* method's propensity to focus on the ridge peak, PV^* also picks up about half as many instances of ridging (averaged blocking frequency of 9%, $\sim$8.28 days per season at its maximum over Alaska) as Z^* (averaged blocking 18%, $\sim$16.56 days per season in the location of the PV^* blocking frequency maximum). The AGP method has an averaged blocking frequency that about equals PV^* over Alaska, and a high probability of co-occurrence with Z^* ($P(Z^*|AGP) = 0.78$), but its maximum during this time is positioned over the Western Pacific, indicating that it is less likely to identify the particular feature that we are interested in identifying here.

The reason for this difference in the average blocking patterns with respect to the

RRR becomes apparent in Figure 2.5. In each of these examples, the blocks detected by the AGP, Z^* and PV^* methods are outlined in purple, blue, and green solid lines, respectively, and the thin contours depict 500 hPa geopotential height in 50 m increments. Many times, the ridge that appears off the North American west coast has a north/south oriented ridge axis with little to no horizontal tilt; therefore, the $GHGS$ criterion of the AGP method is not fulfilled, as will be discussed in Section 2.3.1. The blocking pattern seen on December 7th, 2013 (Figure 2.5d) is one of the exceptions, due to both the slight westward tilt (with increasing latitude) of the ridge axis and the local Z500 maximum at 220E, both of which satisfy $GHGS > 0$. However, even in this example, where all three methods detected the feature, they defined the extent of the block differently; comparing the three methods, $S(PV^*, Z^*) = 0.56$, $S(PV^*, AGP) = 0.38$, and $S(Z^*, AGP) = 0.25$. If these detection algorithms were to be used in some sort of predictive capacity, the Z^* method would pick up RRR-like features every time they occurred, but detect many additional features as well. The PV^* method would detect some of the same features as Z^*, but it would also miss some instances, as in Figure 2.5b, and would not define the extent of the block in the same way. Like PV^*, the AGP method would also miss some of these features, as well as identifying others that are not relevant here.

If the goal is to detect a ridge configuration similar to that of the RRR, then the results suggest that of the three algorithms discussed in this paper, the Z^* algorithm is the most reliable method to find the ridge, with PV^* acting as a more conservative substitute method. The algorithm design of AGP does not deal well with the particular block shape that appears frequently during this time; therefore, the use of AGP is not ideal for performing an analysis on future trends in blocking specific to the western United States.

2.2.2   Blocking climatology by algorithm

Figure 2.6 compares seasonal blocking frequency obtained from applying the three detection methods to the full ERA-Interim reanalysis dataset from 1979-2018. Section 2.2.4 quantifies the amount of agreement between each of the methods, but it is obvious just from viewing the frequency plots that, even using the same time and size thresholds,
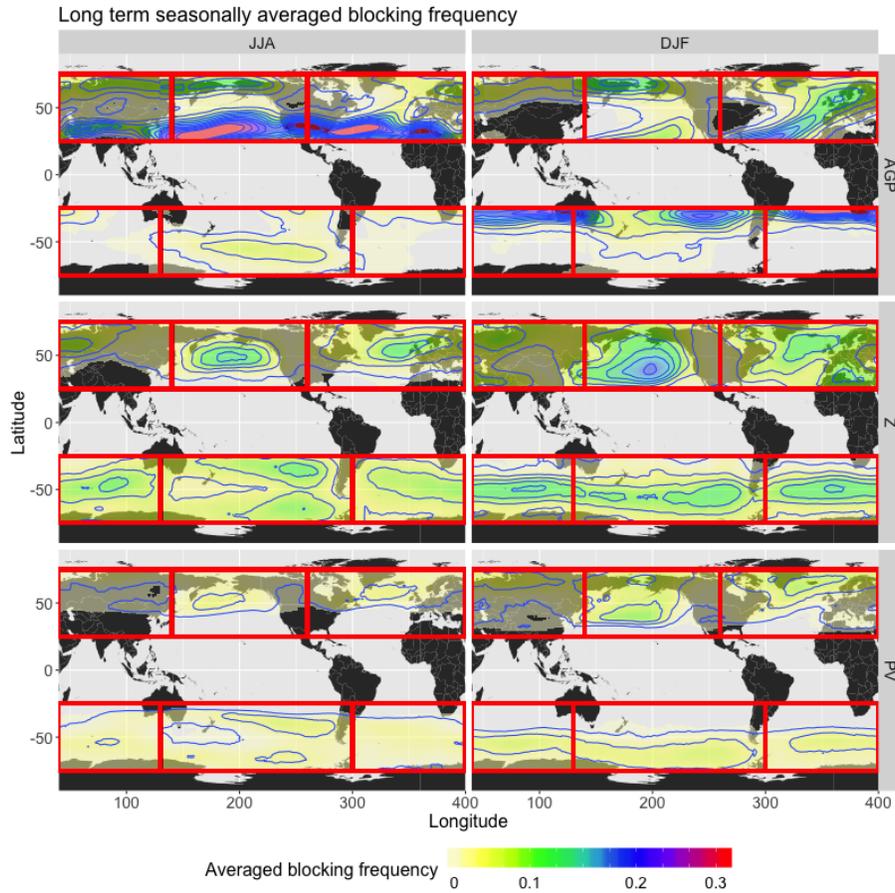
Figure 2.6: Long term seasonally averaged blocking frequency for (left) JJA and (right) DJF, (top row) *AGP* method, (center row), Z^* method, (bottom row) PV^* method. Frequency values represent the fraction of blocked days per season as averaged over the 39 years of the study, with frequencies here ranging from 0.01 (less than one day per season) to 0.40 (about 37 days per season). Contour lines have intervals of 0.03.

these algorithms do not agree on the definition of blocking. Each of the three algorithms produce distinct regional and seasonal differences in their overall global blocking climatologies. Z^* is the least discriminating, detecting blocks over the entire extent of each study region (particularly in the winter hemispheres), with maxima of about 19% (18 days per season); in contrast, AGP appears more optimized for the NH than the SH, as it detects scarcely any blocks in the SH midlatitudes with the exception of SP, but has distinct maxima of almost 40% (37 days) at the lower latitudes in the summer hemispheres. PV^* has maxima that are similar in location to Z^* maxima, but the lowest overall magnitudes of blocking frequency (maximum values of $\approx 9\%$, or 8 days per season).

These differences are analogous to results reported in previous blocking studies. Certain maxima seen in Figure 2.6 are also present in other studies, i.e. the maxima that are centered over Scandinavia and the Aleutian Islands in the NH (Croci-Maspoli et al. 2007, using S04; Dunn-Sigouin et al. 2012, using a hybrid of TM90 and DG83; Cheung et al. 2013, using TM90); and in the SH, the maximum centered in SP (Wiedenmann et al. 2002, using TM90; Parsons et al. 2016, using DG83). The results between papers diverge in terms of the blocking frequency magnitude and the overall extent of blocking, but one could reasonably assume that this was in part due to differences in methodology (choice of variable, anomaly threshold, 1D vs 2D, extent of study region). Here, despite attempts to standardize the detection methodology as much as possible, the three objective detection methods still display differences in magnitude and extent.

One of the most notable differences between methods is seen in the lower latitudes. The 2D blocking indices utilized here select certain blocking patterns that are typically missed by the 1D indices of Tibaldi and Molteni (1990) and others, because 1D methods only identify blocks in the vicinity of a central latitude. Low-latitude blocks are the most obvious example of features that are missed by the 1D methods, and the largest blocking frequencies are all at the lower latitudes. These features will be discussed in further detail in Section 2.3.3; while they tend to be less stationary than other blocking features, they still impair the zonal flow as per the AMS definition and are therefore worth including in the results. AGP detects a high percentage of low-latitude blocks

in the summer hemispheres (maximum frequency of $\approx 40\%$, or 37 days per season), as well as in the winter hemispheres to a lesser extent. The summer low-latitude blocks are rarely detected by the anomaly methods because deviation from the mean accounts for seasonally high values of their respective variables, but AGP has no reference to the mean climatology. However, the anomaly methods do detect low-latitude blocks in winter (particularly in NP DJF), due to seasonally low values of the LTDM.

2.2.3 Blocking duration, zonal distance traveled, speed, size

This section addresses differences in the block duration, zonal distance traveled, zonal block speed, and block size that emerge from each detection scheme (calculation methods for each of these items is briefly explained in Appendix A.2). It should be noted that all of these characteristics are somewhat intertwined with one another. Smaller detected regions often correlate with shorter durations due to the size constraint, and block speed, being a function of distance and duration, will be skewed by relatively high or low values of either.

Results for each of the methods are presented in Figures 2.7-2.10, with the summer season values of the respective hemisphere on the top row and the winter season values on the bottom row. Statistical significance was established by using permutation, as explained in Appendix A.5.

Block duration: These medians of these results fall within the range of previously presented values (about 6-8 days), such as results from Wiedenmann et al. 2002. The 5-day minimum threshold provides a lower bound to the range of possible 25th percentile values, but the upper bound is quite variable per method, as seen in Figure 2.7. Z^*-detected blocks have the overall largest durations (as well as most of the largest outlier values), although low-latitude blocks with longer durations values contribute to the upper tail of the AGP in the summer hemispheres. PV^* has the smallest overall duration values, but the difference is not always significant, particularly when compared to AGP. Smaller duration values are partly an artifact of the size threshold, and PV^*-detected regions tend to be smaller and therefore meet the minimum size requirement less often.
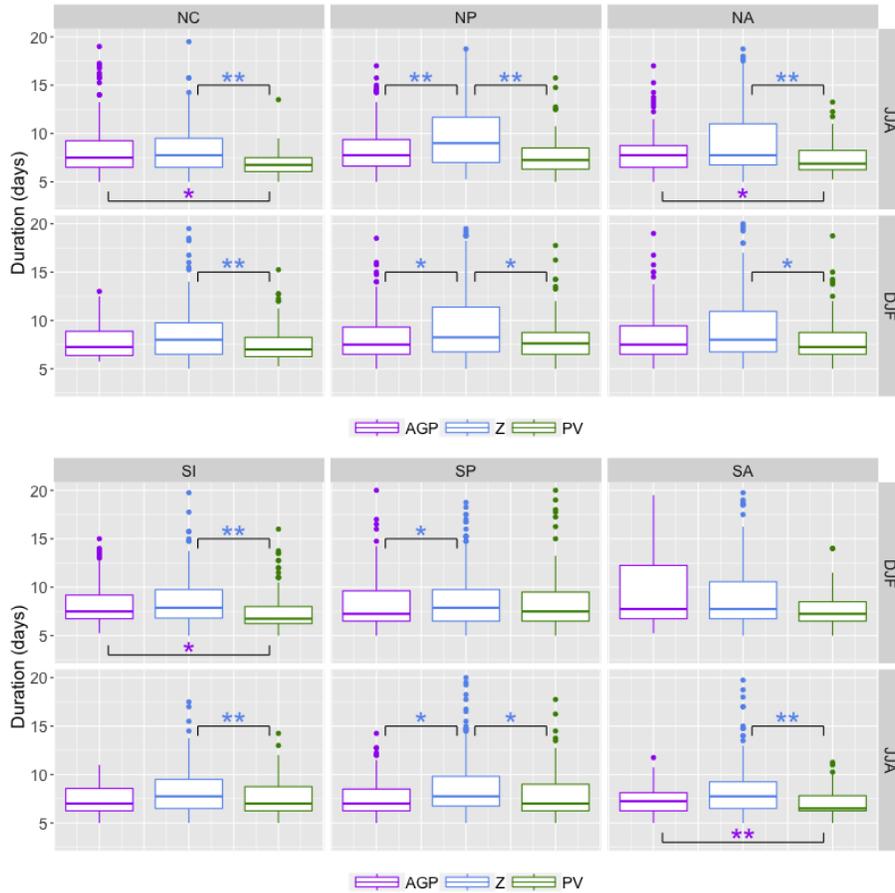
Figure 2.7: Boxplots of block duration values for NH (top) and SH(bottom), in days. The upper and lower bounds of the box correspond to the 25th and 75th percentile values; the ends of the whiskers correspond to 1.5 times the 25th and 75th percentiles. Dots signify outliers beyond the whiskers. The brackets indicate pairs with statistically significant differences in the median values, with a "*" denoting $0.01 < p < 0.05$ and a "**" denoting $p < 0.01$. The colors of the asterisks indicate which method's median value is larger (i.e. a purple asterisk indicates that the median value for AGP is larger).
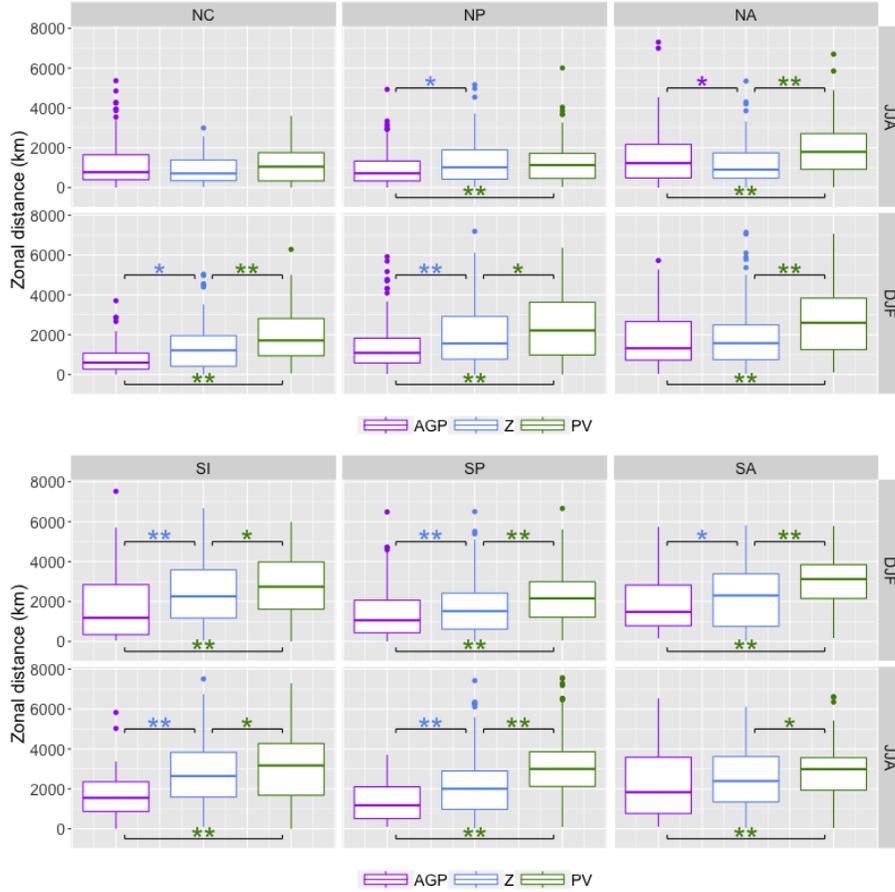
Figure 2.8: Similar to Figure 2.7 except with distance values in km.

Block zonal distance traveled: As most previous blocking studies determine blocking on a per-gridpoint basis, it is difficult to draw direct comparisons to results here. Figure 2.8 shows that regions detected by PV^* tend to travel longer distances in the zonal direction (particularly in winter), despite having shorter duration values; statistically significant ($p < 0.01$) differences in median values between PV^* and other methods are in the 500-1500 km range. There is a fairly consistent trend in which $PV^* > Z^* > AGP$ in terms of the distribution means, although the differences in the overall distributions are not always significant. It is worth nothing that for the AGP method, a larger proportion of cases with distance values greater than 2000 km (≈ 24 degrees longitude) are associated with low-latitude blocking than the other methods; 68% (232 out of 341 cases) of AGP blocks with distance values greater than 2000 km were detected equatorward of 40 degrees, compared to 4% (29 out of 696 cases) for PV^* and 28% (153 out of 758 cases) for Z^*. In
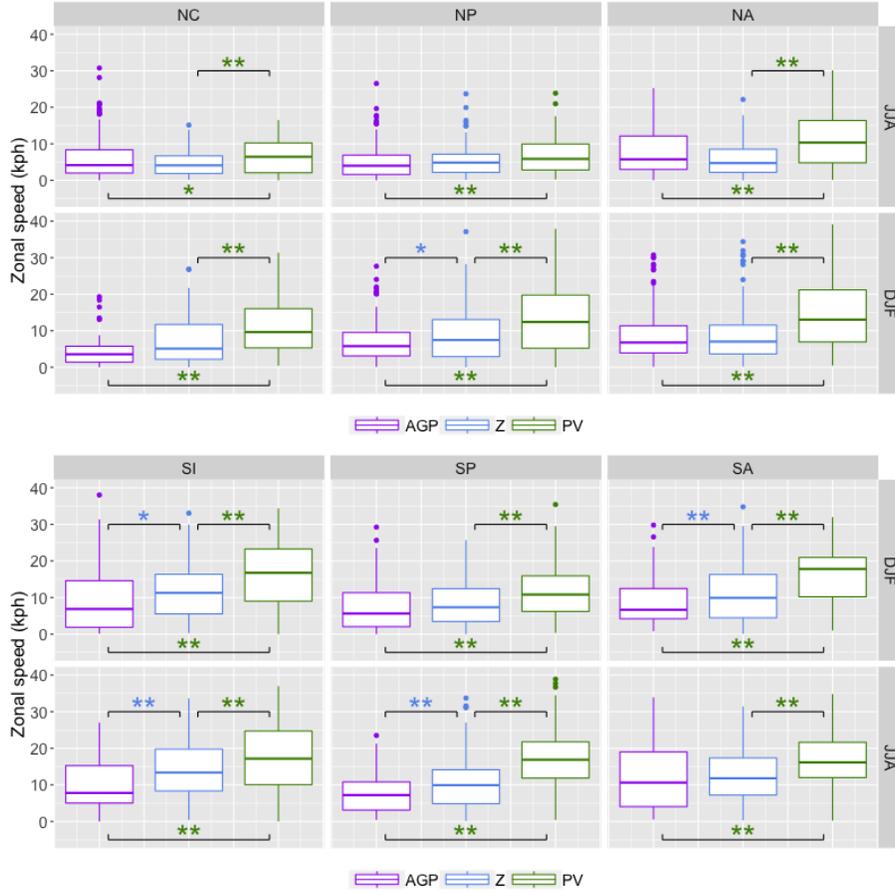
Figure 2.9: Similar to Figure 2.7 except with zonal speed in km/hr.

general, interquartile ranges of SH distance values are greater than interquartile ranges of NH distance values, and winter is greater than summer.

Block zonal speed: As with zonal distance, it is difficult to compare these results to previous studies, but Sinclair (1996), which tracks anticyclones in the SH, used a criterion of 3000 km in 5 days (or 25 km/hr) as a threshold for limiting tracking to slower-moving features. This threshold was in terms of 2D distance, rather than zonal distance as presented here, but assuming that the trajectory is mainly zonal with slight latitudinal variation, an estimate of 25 km/hr is a reasonable estimate for an upper bound. Speed values, as shown in Figure 2.9, tend to follow a similar pattern to distance values, PV^* displaying the largest values and faster speeds in the SH than the NH. PV^*-detected regions have the largest zonal speed values due to the combination of shorter duration and longer zonal distance, particularly in the SH; the 25 km/hr benchmark from Sinclair is
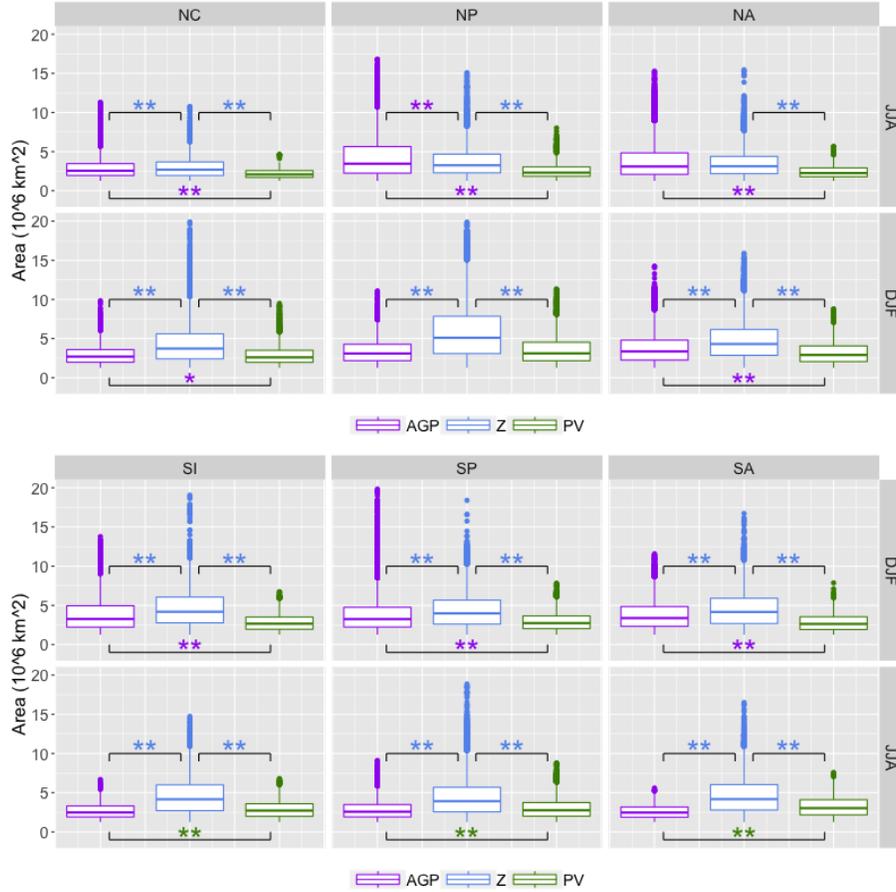
Figure 2.10: Similar to Figure 2.7 except with area in 10^6 km^2

exceeded more often by PV^*-detected regions than other methods, and the 75th percentile value is regularly up to 50% larger than that of AGP or Z^*. The relationship between AGP and Z^* consistently shows $AGP < Z^*$ in the SH, but the pattern is less clear in the NH, particularly in NH summer where AGP detects a substantial number of low-latitude blocks with large distance and duration values, as noted in the previous sections.

Block size: It is worth noting that this is the one metric where almost every possible combination of methods, seasons, and regions displayed significant differences between distributions of values. Croci-Maspoli et al. (2007), which uses S04 (but without the size threshold as defined here), presents results in which detected regions in the NH range from 0.5 to 4×10^6 km^2; Figure 2.10 displays the range of block sizes for each method here. The interquartile range for PV^* is largely in agreement with this finding (unsurprising as the algorithm originated in S04), while those of AGP and particularly Z^* tend to have

larger 75th percentile values and a substantial number of outliers that reach a magnitude of up to 20×10^6 km^2, which is 5 times larger than the S04 value. Size is an important consideration here because it can determine whether or not a detected region is rejected due to the threshold constraint. When a blocked feature is detected by both the PV^* and Z^* methods, the larger Z^* contour will frequently both appear earlier and persist longer than the smaller PV^* contour (this is seen in Figure 2.16). This is not an artifact of the threshold magnitude, as altering the magnitude of the threshold did not significantly impact the distribution of block size; PV^* blocks on average will be smaller than their Z^* counterparts.

2.2.4 Intercomparison of blocking algorithms

Tables 2.4-2.6 provide a summary of the various co-occurrence metrics between the three methods. Each section displays results for the NH and SH as separate subtables, with the summer seasons on top. Correlation and probability tables contain single numbers, as they were calculated as single quantities, but as spatial similarity was calculated for instantaneous fields at each time step containing blocking, there are a range of similarity values, presented here as interquartile ranges similar to the tables in the previous section; there are instances of fields where $S(M1, M2) > 0.9$, but these are rare.

Pearson pattern correlation: The results in Table 2.4 highlight the clear differences in blocking frequency that will arise given the choice of an anomaly versus a total field (discussed further in Section 2.3.1). Correlation between AGP and the other methods is consistently negative in the summer hemispheres due to AGP detecting blocks mainly at the lower latitudes, and NC DJF is the only instance of correlation values that are both non-negative and greater than 0.2. In comparison, $C(PV^*, Z^*)$ is almost always the strongest positive correlation value for a particular region or season (with the exception of $C(PV^*, AGP)$ in NC DJF), with particularly strong correlation values in SH summer, NA DJF, SI (both seasons) and NP (both seasons).

Probability of co-occurrence: The interpretation of the values in Table 2.5 must take the relative quantities of each method's detected features into account. For example, it was previously noted that PV^* and Z^* had similar averaged patterns, with some of the

(a) Northern Hemisphere

Season	Method pair	NC	NP	NA
JJA	PV^* and AGP	*-0.21***	-0.56**	-0.46**
	Z^* and AGP	*-0.19***	-0.38**	-0.49**
	PV^* and Z^*	0.51**	**0.76****	**0.73****
DJF	PV^* and AGP	0.62**	*0.01*	*-0.06**
	Z^* and AGP	*0.26***	*-0.04*	*-0.01*
	PV^* and Z^*	0.49**	**0.71****	0.45**

(b) Southern Hemisphere

Season	Method pair	SI	SP	SA
DJF	PV^* and AGP	-0.55**	-0.54**	-0.46**
	Z^* and AGP	-0.68**	-0.64**	-0.58**
	PV^* and Z^*	**0.76****	**0.71****	**0.79****
JJA	PV^* and AGP	-0.48**	*-0.06**	-0.59**
	Z^* and AGP	-0.58**	*0.09***	-0.53**
	PV^* and Z^*	0.70**	0.44**	0.67**

Table 2.4: Pearson correlation values between blocking frequencies of each region, as shown in Figure 2.6. Magnitudes above (below) 0.7 (0.3) are bolded (italicized) to emphasize patterns of consistently high or low values. Negative values imply an inverse relationship between corresponding gridpoint values. Statistical significance is denoted by a "*" for $0.01 < p < 0.05$ and "**" for $p < 0.01$.

(a) Northern Hemisphere

Season	Method pair	Probability	NC	NP	NA
JJA	PV^* and AGP	$P(PV^*\|AGP)$	*0.04*	*0.05*	*0.04*
		$P(AGP\|PV^*)$	0.45	0.46	0.34
	Z^* and AGP	$P(Z^*\|AGP)$	*0.11*	*0.18*	*0.13*
		$P(AGP\|Z^*)$	0.43	0.66	0.44
	PV^* and Z^*	$P(PV^*\|Z^*)$	*0.14*	*0.29*	*0.24*
		$P(Z^*\|PV^*)$	0.48	**0.73**	0.60
DJF	PV^* and AGP	$P(PV^*\|AGP)$	0.38	0.34	*0.21*
		$P(AGP\|PV^*)$	0.30	*0.38*	0.48
	Z^* and AGP	$P(Z^*\|AGP)$	0.59	0.55	0.52
		$P(AGP\|Z^*)$	*0.18*	0.41	0.53
	PV^* and Z^*	$P(PV^*\|Z^*)$	*0.24*	0.48	0.31
		$P(Z^*\|PV^*)$	0.61	**0.73**	0.70

(b) Southern Hemisphere

Season	Method pair	Probability	SI	SP	SA
DJF	PV^* and AGP	$P(PV^*\|AGP)$	*0.01*	*0.04*	*0.00*
		$P(AGP\|PV^*)$	*0.03*	*0.11*	*0.01*
	Z^* and AGP	$P(Z^*\|AGP)$	*0.09*	*0.24*	*0.09*
		$P(AGP\|Z^*)$	*0.16*	0.31	*0.14*
	PV^* and Z^*	$P(PV^*\|Z^*)$	0.40	0.34	0.34
		$P(Z^*\|PV^*)$	**0.78**	**0.73**	**0.78**
JJA	PV^* and AGP	$P(PV^*\|AGP)$	*0.13*	*0.24*	*0.18*
		$P(AGP\|PV^*)$	*0.04*	*0.20*	*0.11*
	Z^* and AGP	$P(Z^*\|AGP)$	0.49	0.66	0.67
		$P(AGP\|Z^*)$	*0.06*	*0.18*	*0.14*
	PV^* and Z^*	$P(PV^*\|Z^*)$	*0.26*	0.31	*0.28*
		$P(Z^*\|PV^*)$	**0.72**	**0.71**	**0.79**

Table 2.5: Probability of co-occurrence between instantaneously blocked fields. Probability values above (below) 0.7 (0.3) are bolded (italicized) to emphasize high or low values. All probability values are significant at the $p < 0.01$ level.

highest correlation values in Table 2.4; however, this does not mean that they are equally likely to predict one another. While $P(Z^*|PV^*)$ is consistently the highest value in a given region, the reverse is not true, since Z^* also detects many other regions that do not coincide with PV^*-detected regions. Also, averaged agreement does not necessarily also mean instantaneous agreement; for example, $C(PV^*, AGP)$ is 0.62 in NC DJP but the probabilities of co-occurrence for these two methods are only about 30-38%, while $C(PV^*, Z^*)$ is 0.49 in NC DJF but $P(Z^*|PV^*) = 0.61$. In this instance, PV^* and AGP detect blocks in similar areas, but the methods are not simultaneously detecting the same features on a per-timestep basis.

Spatial similarity: The third measure of agreement between methods, spatial similarity, is highly influenced by the size and location of detected regions relative to one another. Note that the largest 75th percentile value is 0.62, which is unsurprising given the differences in detected block size between methods even if the same feature is detected, as was seen in Figure 2.5d. The results in Table 2.6 provide additional insight to the results from the previous two sections. While PV^* and Z^* again have some of the highest values for similarity relative to other combinations of methods, $S(PV^*, Z^*)$ rarely exceeds 0.6, and more often in the respective summer hemispheres when the Z^* blocks are smaller. Additionally, $S(PV^*, AGP) < S(Z^*, AGP)$ in most instances; this is sometimes due to the relative sizes of the detected contours, but the PV^*-detected contour is also sometimes shifted with respect to the $Z500$-based methods (example in Section 2.3.2).

2.3 Meteorological drivers of differences between blocking algorithms

The various metrics displayed in Section 2.2 show that all three of the methods have widely varying definitions of blocks, from the block's physical characteristics to whether or not a block is actually present at a particular point in time. The RRR case study, in particular, highlights the importance of considering the nature of the region's flow field and prevailing block type when selecting the appropriate detection method. A few meteorological factors that influence the differences between block detection methods are

(a) Northern Hemisphere

	Method pair	NC	NP	NA
JJA	PV^* and AGP	0.21 to 0.46	*0.08* to *0.38*	0.16 to 0.43
	Z^* and AGP	**0.35** to **0.58**	*0.13* to *0.37*	0.19 to 0.50
	PV^* and Z^*	**0.33** to **0.57**	**0.32** to **0.54**	**0.31** to **0.54**
DJF	PV^* and AGP	0.27 to **0.53**	0.15 to 0.43	*0.13* to 0.43
	Z^* and AGP	**0.35** to **0.58**	0.19 to 0.42	0.25 to 0.50
	PV^* and Z^*	**0.29** to **0.57**	0.25 to 0.53	0.28 to **0.54**

(b) Southern Hemisphere

	Method pair	SI	SP	SA
DJF	PV^* and AGP	*0.02* to *0.21*	*0.08* to *0.37*	*0.02* to *0.08*
	Z^* and AGP	*0.07* to *0.23*	0.18 to 0.39	*0.06* to *0.24*
	PV^* and Z^*	**0.36** to **0.59**	**0.34** to **0.62**	**0.37** to **0.58**
JJA	PV^* and AGP	*0.05* to *0.26*	*0.08* to *0.31*	*0.05* to *0.31*
	Z^* and AGP	0.24 to 0.46	0.21 to 0.43	0.28 to 0.47
	PV^* and Z^*	0.26 to 0.52	0.23 to 0.48	0.27 to **0.54**

Table 2.6: Interquartile ranges of spatial similarity between instantaneously blocked fields. The 25th and 75th percentile values are formatted based on the relative distributions of these values— 25th percentile values above (below) 0.29 (0.15) are bolded (italicized), and 75th percentile values above (below) 0.54 (0.39) are bolded (italicized) to denote particularly high or low quantities relative to other values. All ranges of similarity values were significantly different from the generated null distribution.
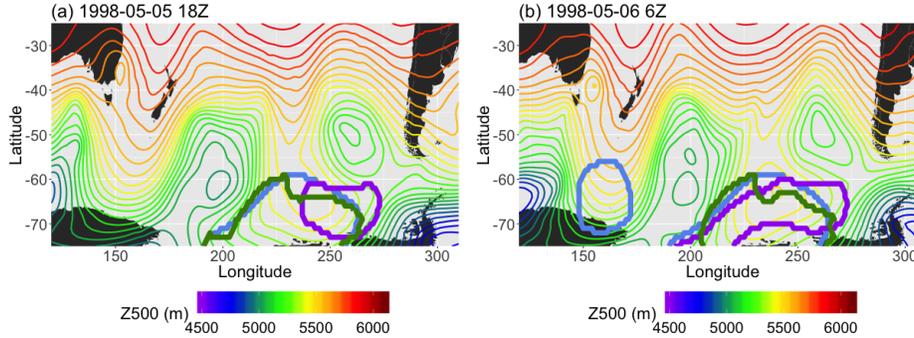
Figure 2.11: Example, 12 hours apart in 1995 MAM, of instances in which there is (a) less and (b) more agreement between the AGP method (purple) and the two anomaly methods (blue and green) in SP.

discussed next.

2.3.1 Anomaly versus total field Z500-based methods

In Section 2.2.4, the magnitudes of the agreement metrics show that there is a clear distinction between circumstances when blocks are detected by AGP, a method based on the total $Z500$ field, as opposed to the other two methods that are based on anomalies with respect to the LTDM. We limit the discussion here to the two methods that are based on $Z500$ in order to reduce other sources of variability.

Both Z^* and AGP were created with the purpose of detecting high values of $Z500$, but AGP has no reference to the mean climatology; instead, it requires a significant change in $Z500$ over a latitude range (30 degrees) that equates to more than half of each study region's latitude range (50 degrees). The difference in block detection is most evident in the SH, where there is a much stronger zonal flow component than in the NH. This stronger zonal flow implies that the requirements of the AGP method are fulfilled far less often in the SH midlatitudes, since there is not sufficient distortion of the flow field; AGP-detected blocks in the SH midlatitudes are often co-located with lows (dipole or omega blocks) or have a tilt in their north-south axis, since this guarantees that there is a sufficiently large $Z500$ gradient. In contrast, Z^* is much less discriminating; it detects a high number of blocks in the SH because the $\overline{Z500}$ field has a strong meridional gradient in addition to its mainly zonal flow, meaning that even fairly shallow ridges in the $Z500$ field are more likely to satisfy Z^* anomaly thresholds.

These points are illustrated in Figure 2.11, which shows a snapshot of an omega block in SP on May 5th 18Z 1998 (left) and 12 hours later (right). At the earlier time, the two anomaly methods produce clusters that are centered over the ridge ($S(PV^*, Z^*) = 0.78$), but the AGP method only picks out an area centered polewards of the lower geopotential heights to the east of the ridge (before the size constraint was applied, there was also a detected region in the vicinity of 190E,65S); $S(PV^*, AGP) = 0.24$ and $S(Z^*, AGP) = 0.26$. Later, the intensification of the high in the 60-70S latitudes produces the necessary height gradient to satisfy the criteria for AGP; $S(PV^*, AGP) = 0.46$ and $S(Z^*, AGP) = 0.58$.

2.3.2 PV^* links to shear and vorticity

The PV^* method picks out regions with PV that are highly anomalous with respect to the climatological mean in the upper troposphere. Most often, these regions are areas with particularly pronounced anticyclonic circulation in the $Z500$ field, such as dipoles or omega blocks. Since anomalous vorticity and anomalous highs are linked, the two anomaly methods are often very similar in terms of the location of the detected block, even if the size of the detected cluster of grid points is not always the same. However, the EPV field can be influenced by phenomena such as vertical shear (first two terms in Equation A.1.2) or locally strong winds; it is a factor to consider in flow that does not have the necessary meridional component to satisfy points 1 and 2 of the AMS definition of blocking. For example, a jet streak embedded in otherwise zonal flow will lead to mistaken identification of a region as blocked even though there is neither persistent obstruction nor pronounced meridional flow. The PV^* method tends to detect gridpoints at higher latitudes than the other two methods during each respective hemisphere's summer, and over a wider range of latitudes in the winter; this corresponds roughly to the location of the 500 hPa jet, which shifts from polewards of 45 degrees latitude in summer to 40 degrees latitude and below in the wintertime. Therefore, the PV^* method is somewhat linked to the presence of the jet stream and embedded jet streaks.

Figure 2.12 demonstrates a case in which a PV^*-detected cluster is shifted relative to clusters detected by the other two methods, despite all of the methods identifying
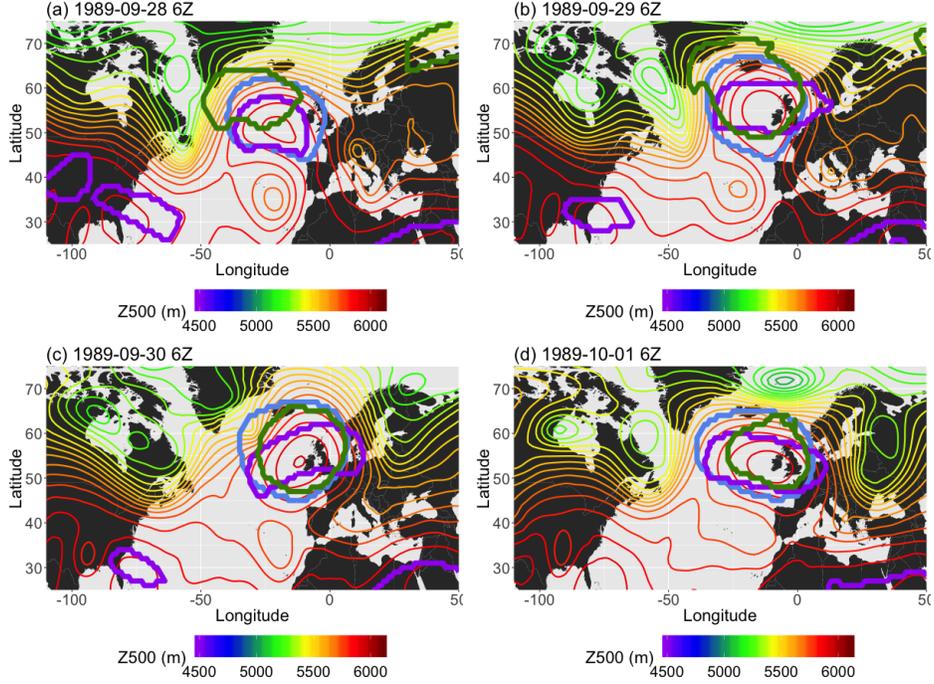
Figure 2.12: Example, in 24-hour increments, of omega block detection in 1989 NA SON. The PV^* method is denoted by the green contour, the Z^* method is blue, and the AGP method is purple.

the same block. The 4-panel figure shows an omega block that is detected by all three methods in 1989 NA SON; however, in Figure 2.12a, the Z^* and AGP contours both center on the high that is the top half of the block ($S(Z^*, AGP) = 0.56$), while the PV^* contour is shifted westwards ($S(PV^*, Z^*) = 0.29$, $S(PV^*, AGP) = 0.18$). Daily composites of the corresponding total wind (Figure 2.13) show that, on September 28th 6Z, 1989 (Figure 2.13a), there was a jet streak on the upstream side of the high pressure feature which is being detected by the three algorithms. The vorticity ($-g\zeta\frac{\partial\theta}{\partial p}$ in Equation A.1.2) associated with the anticyclonic curvature of the high is strongly negative at this time, and further enhanced by vertical shear in the jet streak region. The combination of these factors leads to the westward extension of the PV^*-detected cluster relative to the other two clusters. As time progresses (Figure 2.13b), a portion of the PV^*-detected contour continues to track that jet streak, which coincides with strongly negative vorticity, until it reaches the downstream side of the omega block (Figure 2.13c). In Figure 2.13d, the next jet streak has reached a higher latitude, where $\overline{PV}$ is larger and there is positive

39
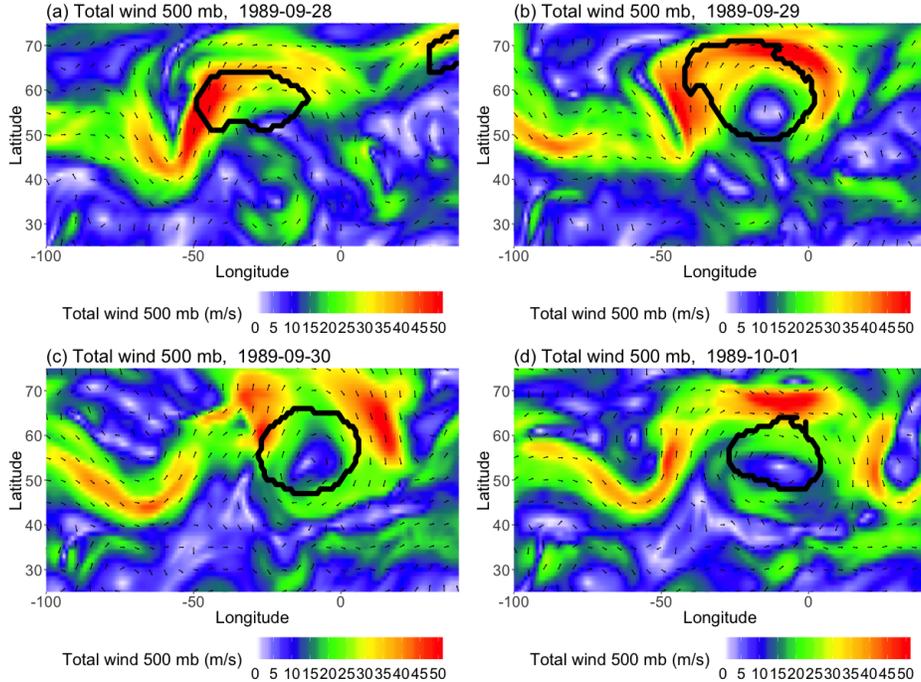
Figure 2.13: 500 hPa vector wind field corresponding to previous figure (September 28th-October 1st), showing location of jet streaks. Wind speeds upwards of 45 m/s are visualized as the red areas. The thick black contour corresponds to the blocked region detected by PV^*.

vorticity associated with the low. The anomaly is no longer exceeding the threshold; therefore, the PV^* cluster separates from the jet streak and is in better agreement with the other two methods $(S(PV^*, Z^*) = 0.52, S(PV^*, AGP) = 0.54)$.

2.3.3 Low-latitude blocking and flow impairment

Low-latitude blocking with respect to NH summer and the AGP method is discussed in Davini et al. (2012);their paper questions whether these features are correctly characterized as blocks, since they are linked to poleward displacement of subtropical easterlies and are less intense and persistent than those at higher latitudes. Low-latitude blocking detection is present in all three methods: the AGP method has relative maxima in the averaged blocking patterns in the respective summer hemispheres, and all three methods find low-latitude blocks in the respective winter hemispheres (PV^* and Z^* have distinct maxima in the lower latitudes of NP DJF). Many of these low-latitude features are nearly stationary (block speed averages 2-12 km/hr in both hemispheres) and persistent (block
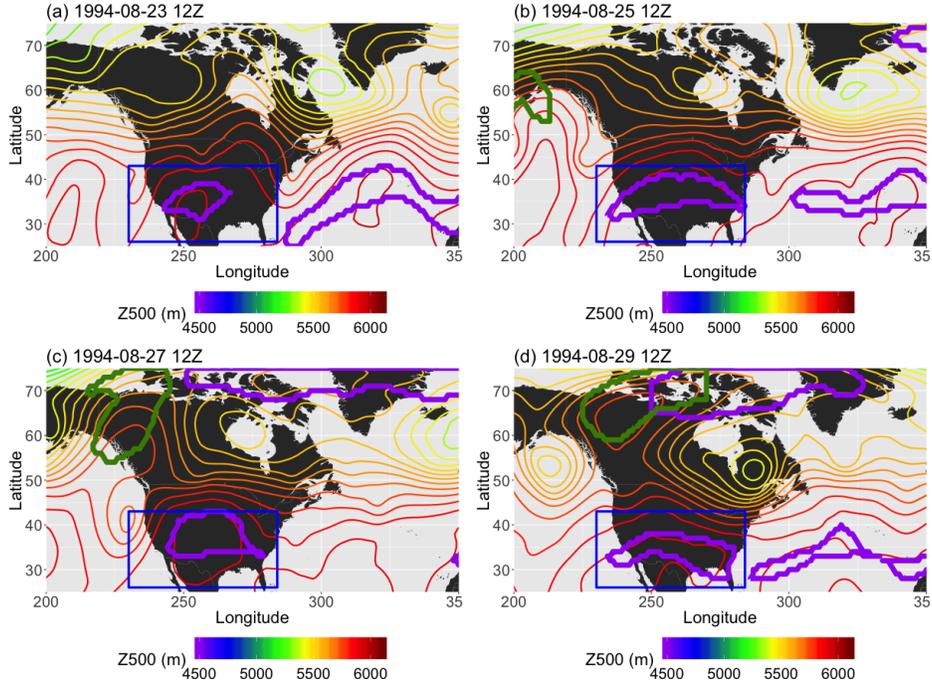
Figure 2.14: Example, in 48-hours increments, from NP JJA 2014, of a low-latitude block detected by the *AGP* (the other two methods do not detect a block here). Thin contours are Z500 in 50m intervals, and the thick purple contour denotes the detected feature. The blue box spans [230E-234W, 25N-43N] and outlines the extent of the detected block.
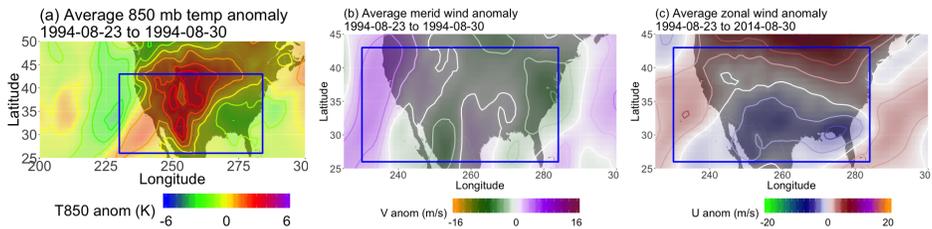
Figure 2.15: Averaged (a) 850 hPa temperature, (b) 500 hPa meridional, and (c) 500 hPa zonal wind anomalies for June 8th-17th 1984. The temperature contour intervals are 1K and the wind contour intervals are 2 m/s. The blue box corresponds to the one seen in Figure 2.14.

duration averages 6-12 days in both hemispheres). We present two cases of low-latitude blocking here: one summer case, and one winter case.

Figure 2.14 shows an example of one of the more stationary low-latitude *AGP*-detected regions in JJA, a persistent ridge over the central United States in 1994 that lasted from August 23rd 12Z to August 30th 0Z and had an average zonal speed of 0.74 km/hr over 7.75 days. The other two methods did not detect this feature, although the *PV**

41
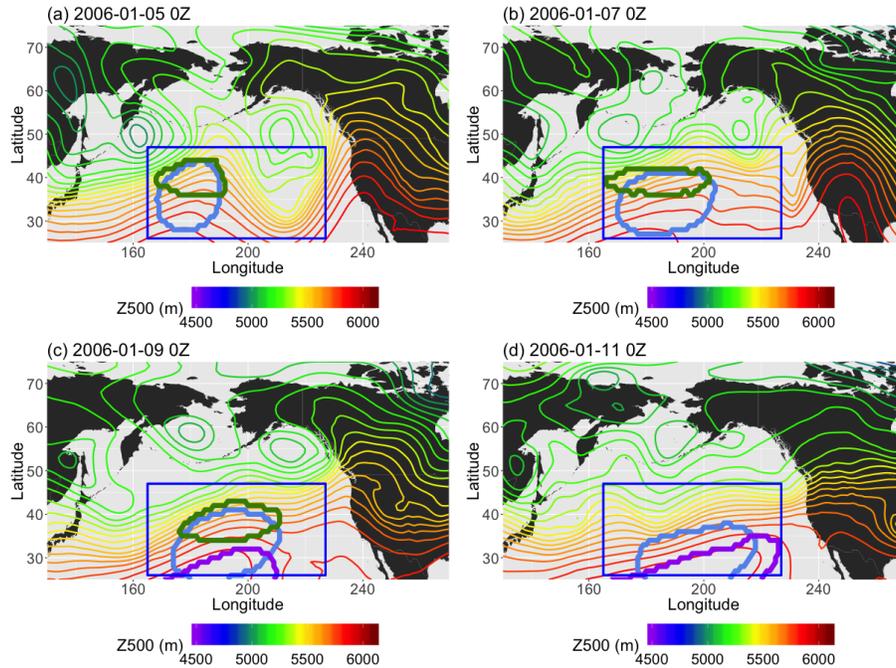
Figure 2.16: Example, in 24-hours increments, from NP DJF 2006, of a low-latitude block detected by the AGP (purple), Z^* (blue), and PV^* (green) methods. The blue box spans [165E-227E, 26N-47N] and outlines the extent of the detected block.

method detected a second feature at the higher latitudes of Figures 2.14b-d. Averaged over the detected block's lifespan, the 850 hPa temperature anomaly (T_a) in the vicinity of the detected block is approximately +4K, the 500 hPa meridional wind anomaly (v_a) is approximately ± 5 m/s on either side of the block, and the 500 hPa zonal wind anomaly (u_a) is up to -5 m/s in the blocked flow region (see Figure 2.15). The size and shape of the AGP-detected region fluctuates from panel to panel as the height gradient changes, a trait that is more common among AGP-detected regions (particularly at the lower latitudes) than their Z^* or PV^* counterparts. The resultant averaged wind anomalies are somewhat weaker than higher latitude cases, but the flow is indeed being consistently diverted over the course of the ridge's existence, and therefore satisfies the first point of the AMS definition (as well as somewhat satisfying the second point of meridional flow).

In contrast, Figure 2.16 shows a low-latitude case in NP DJF in which all of the methods identify a block in the midst of flow that is not sufficiently diverted or slowed. This example shows January 5th 0Z to January 11th 0Z, 2006, where all three methods
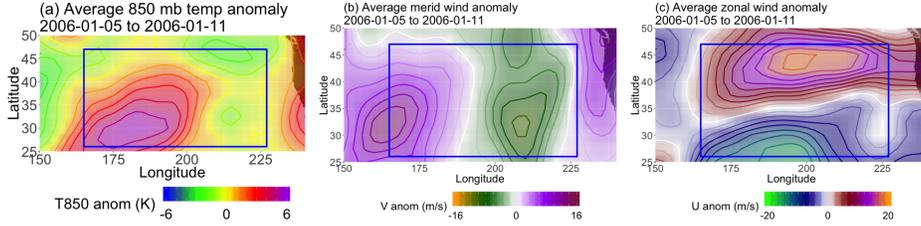
Figure 2.17: Averaged (a) 850 hPa temperature, (b) 500 hPa meridional, and (c) 500 hPa zonal wind anomalies for January 6th-11th 2006. The temperature contour intervals are 1K and the wind contour intervals are 2 m/s. The blue box corresponds to the one seen in Figure 2.16

detected a relatively shallow ridge that evolves into mostly zonal flow by the end of the detected block's lifespan, although the AGP method only detects the ridging at the lower edge of the study region during the last few days (Figures 2.16c and d) and the PV^* method disappears in Figure 2.16d. The PV^*-detected region has an average zonal speed of 12 km/hr over 8 days and the Z^*-detected region has an average zonal speed of 10 km/hr over 12.5 days (the difference is mainly due to the longer lifespan of the Z^*-detected region). The AGP-detected region has an average zonal speed of 7.86 km/hr over 5.25 days (and it mainly tracks the ridge after the other two methods have stopped following it, outside of the time window examined here). u_a is -16 m/s in bottom half of the blocked flow region outlined in Figure 2.17, but northwards of 35N, u_a has a maximum magnitude of 20 m/s. v_a is ±10 m/s on either side of the block and T_a is +5K within the blocked region, both of which are larger anomalies than the JJA case. However, one must always approach anomalies with caution; the Z500 field in Figure 2.16 does not indicate that the zonal flow has been reduced at later times. Indeed, composites of the total wind fields in Figure 2.18 show that while the JJA case has reduced wind speeds (maximum 5 m/s in the blocked region, compared to 20 m/s outside the impaired region), the DJF case has wind speeds of up to 35 m/s in the northern portion of the "blocked" region, which also coincides with the positive u_a value in the northern half of the outlined region in Figure 2.17. In the case of Z^*, this mistaken identification can be attributed to a fairly shallow ridge in an area with a strong meridional gradient (as in Section 2.3.1). In the case of PV^*, vorticity values were only slightly negative or close to 0, but when considering
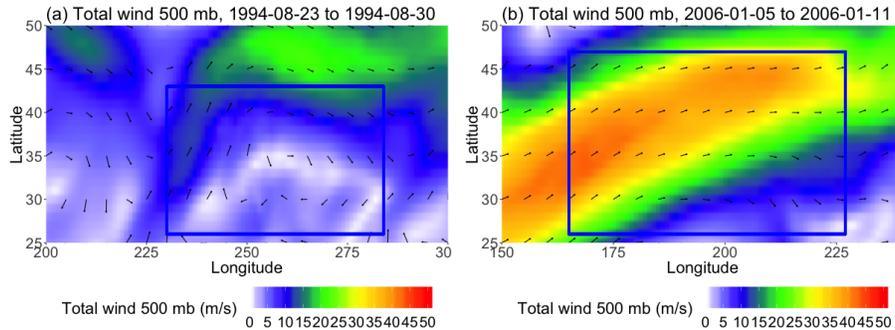
Figure 2.18: 500 hPa total wind fields and vectors for (a) the JJA blocking case in Figure 2.14 and (b) the DJF blocking case in Figure 2.16. The vectors indicate the wind direction, and the colors indicate the wind magnitude.

VPV compared to $\overline{VPV}$, sufficiently anomalous for VPV^* to surpass the local threshold value of 1.2 PVU. In both instances of AGP-detected blocking, the presence of a sustained ridge was enough to qualify as blocked, regardless of other criteria.

These examples suggest that not all low-latitude blocking should be discounted, at least according to certain metrics; thus, applying blocking algorithms over a range of latitudes rather than a single central latitude will produce a more complete climatology of blocking hotspots. A number of detected low-latitude blocks will partially meet, or fail to meet the AMS flow diversion criteria; the first case detected a region of impaired flow, albeit a fairly weak one; and the second case did not produce demonstrably impaired flow at all. However, algorithm "failures" are not unique to the lower latitudes; therefore, rather than restricting analysis to a limited range of latitudes, future research should consider algorithm limitations, which have been mentioned in previous sections, and make use of additional diagnostic metrics to filter out non-blocked flow.

2.4 Conclusions

This study examines the blocking climatologies of six regions over summer and winter seasons, and highlights some of the block characteristics from three detection methods, which can be seen in Table 2.7. Maximum blocking frequency ranged from 9% (PV^*) to 40% (AGP), with the locations of maximum blocking differing between AGP and the anomaly methods. Some methods, such as the PV^* method, are quite conservative,

	AGP	Z^*	PV^*
Flow field sensitivity	Equatorward low Strong poleward $Z500$ gradient Subtropical high pressure	Strong high Strong $Z500$ meridional gradient	Strong curvature Strong winds, vertical shear
Favored block types	Omega block or ridge with meridional axis tilt Cutoff low	All kinds (with sufficiently large $Z500^*$)	All kinds (with sufficiently large VPV^*)
Spatial distribution of frequency	NH: Higher latitudes, Atlantic Ocean basin (and some East Pacific), summer low latitudes SH: summer low latitudes, SP midlatitudes	Blocks detected over full range of study regions	Higher latitudes in summer seasons, wider range of latitudes in winter seasons
Location (frequency) of maximum	NH JJA equatorwards of 45 degrees (40%)	NP DJF equatorwards of 45 degrees (19%)	NP DJF equatorwards of 45 degrees (9%)
Notable features	Less similar to other two methods, highest frequency of low-latitude blocking	Highest frequency of SH blocking	Lowest overall frequency magnitude between methods
Duration	$\approx 6.5 - 10$ days	$\approx 6.5 - 11$ days	Shortest overall between methods ($\approx 6.25 - 9$ days)
Zonal distance traveled	$\approx 600 - 2500$ km; shortest between methods in NH winter and SH (both seasons)	$\approx 800 - 3100$ km	Longest between methods ($\approx 1200 - 3700$ km) with exception of NH summer ($\approx 600 - 2100$ km)
Zonal speed	$\approx 3 - 12$ km/hr; slowest between methods in NH winter and SH (both seasons)	$\approx 4 - 13$ km/hr	Fastest overall betweem methods ($\approx 6.5 - 18$ km/hr)
Size	Smallest between methods in SH winter ($\approx 1.5 - 2.8 \times 10^6$ km^2); $\approx 1.8 - 4.2 \times 10^6$ km^2 otherwise	Largest overall ($2.2 - 5.2 \times 10^6$ km^2)	Smallest between methods ($1.6 - 3.4 \times 10^6$ km^2) except SH winter

Table 2.7: Summary of notable blocking frequency distribution and block characteristics.

	PV^* and AGP	Z^* and AGP	PV^* and Z^*
Correlation	Weak to moderate negative correlation	Weak to moderate negative correlation	Moderate to strong positive correlation
Probability of co-occurrence	Low to moderate for both $P(PV^*|AGP)$ and $P(AGP|PV^*)$; particularly low in SH	Moderate $P(Z^*|AGP)$ in winter, $P(AGP|Z^*)$ in NH summer; low otherwise	Moderate to high $P(Z^*|PV^*)$ in all regions
Spatial similarity	Lowest ranges of similarity values compared to other method pairs	$S(Z^*, AGP) > S(PV^*, AGP)$	Highest range of similarity values compared to other method pairs in almost all regions and seasons

Table 2.8: Summary of notable observations for intercomparison of objective detection methods.

showing comparatively smaller frequencies and detecting smaller clusters; while others, such as the Z^* method, are less discriminating with respect to regions that are defined as blocked, displaying both higher frequencies and larger clusters.

The intercomparison of the results, as summarized in Table 2.8, raises some points that should be considered in future blocking studies that use objective detection algorithms. Previous blocking studies often present averaged results, rather than examining them on a per-block basis, but high averaged pattern correlation does not imply that two methods will simultaneously detect the same features; furthermore, even when all three methods detect the same feature, they do not necessarily define the block using the same cluster of gridpoints. The AGP and anomaly methods had a much lower degree of agreement than that between PV^* and Z^*, in both the averaged and instantaneous sense. However, even the agreement between PV^* and Z^* is affected by the relative sizes and placement of the detected clusters, and the RRR case study demostrates that the block configuration is an important element to "successful" detection of a block. This has implications to studies

which attempt to link blocking to extreme weather, because attempts to correlate the location of the block with the location of extremes will produce different results based on the chosen algorithm.

The ideal blocking detection algorithm (one that detects all features that satisfy the AMS definition) likely requires elements from multiple detection algorithms (such as Dunn-Sigouin et al. (2012) or Barriopedro et al. (2010), which combine elements of TM90 and DG83), as well as measurements of metrics such as block intensity (developed by Wiedenmann et al. 2002) and flow diversion. Additionally, such an algorithm would need to take seasonal and regional differences of the flow field into account. The results show that there remains a significant discrepancy between published methods with regards to how the AMS definition is interpreted, from the calculated blocking frequencies to the average size and speed of the detected features. Many of the detected regions are persistent in the sense that they are relatively nonstationary for at least 5 days, but the resultant changes to wind speed and temperature are inconsistent. Blocks are a continuum of forms, rather than clearly delineated idealized shapes; and each method is optimized for detecting certain kinds of features under certain kinds of climatological conditions.

This chapter has summarized the factors that influence block detection by three different algorithms, and noted the relative strengths and weaknesses of these methods. Additionally, it has outlined a number of metrics for both assessing individual methods and comparing the results between methods. Objective algorithms show promise for analyzing current and future trends, given their applicability to extremely large volumes of high resolution data. Using results from historical models can provide a confidence bound on anticipated changes in blocking characteristics in future climate simulations, but careful consideration of algorithm biases should be factored into the analysis.

Chapter 3

Assessment of atmospheric blocking in reanalysis and model data

The results of Chapter 2, which show that the choice of algorithm impacts the block characteristics as well as the blocking frequency, prompts a related question as it pertains to the representation of blocking in model data. We now turn our attention to a related question: how does the choice of model impact the block characteristics, and what are some of the features of the model that cause these differences? Section 1.2.2 outlined some of the known issues facing the accurate representation of blocking in climate models versus the reanalysis datasets. Understanding the influence of particular model biases with respect to blocking is useful for modeling groups in terms of both assessing and improving climate models throughout the development process. By quantifying the relationships between model biases and blocking biases, researchers are then able to select a subset of models that will more accurately reflect realistic block formation, as well as working to reduce the existing model biases that impact accurate block representation. Additionally, understanding the correlations between model bias and resultant errors in block location and formation will help to further inform understanding the dynamics of these features.

As in Chapter 2, the analysis is expanded beyond comparison of blocking climatologies; the blocks are again assessed in terms of individual events, and we utilize a similar list of block characteristics in this analysis, with the addition of further metrics that were not calculated during the course of the Chapter 2 research. Analysis is restricted to one

objective detection algorithm; we utilize the geopotential height anomaly method of Dole and Gordon (1983) (e.g. DG83) that was modified as detailed in Section 2.1.2.4 to use a spatiotemporally varying threshold. While many blocking studies utilize the geopotential height gradient method (e.g. TM90), Chapter 2 demonstrated that this method fails to capture certain block configurations, such as those with an insufficient tilt in the ridge axis. The height anomaly method, being the least discriminating of the three detection methods examined in Chapter 2, provides the widest possible variety of blocks for examination in this chapter, which will enable us to determine whether there is a particular pattern in the types of blocks that are produced by a model compared to reanalysis results.

This chapter will focus on the Pacific Ocean basins of both the Northern (NP) and Southern (SP) Hemispheres; the boundaries of these regions remain unchanged from those noted in Chapter 2. As previously noted in Chapter 1, increased spatial resolution does not significantly improve representation of blocking in the NP basin, and using this study area eliminates one potential source of variation introduced by the spatial resolution, as might be seen in the Atlantic basin. All blocking calculations are performed on the provided grid resolution, and only the blocking frequency results are regridded to a common 1-degree spatial resolution in order to assess differences between reanalysis and model results for the gridded variables.

We begin by presenting the blocking statistics for the average of four reanalysis datasets in Section 3.2.1; these results form the benchmark against which the various models are compared. Section 3.2.2 then puts forth the model results as deviations from the reanalysis mean; we assess differences in the distributions of block characteristics and jet latitude positions, as well as the horizontal wind and Z500 fields. The results are shown as values normalized by the reanalysis interquartile range (explained in Section 3.1.3), with additional figures for the gridded variables in order to further discuss significant spatial differences. The normalized bias values are then used as input explanatory variables for a linear regression analysis in Section 3.2.3; we begin with a linear regression fit to all models, then explore the change in the model fit when outlier models are dropped from the calculation.

3.1 Data and Methodology

3.1.1 Data

Blocking is assessed on four reanalysis datasets (Table 3.1), each with 6-hourly temporal resolution. The fields were averaged on a daily basis in order to provide similar temporal resolution to the model data. The blocking results for each reanalysis did not display significant differences from the multi-reanalysis average; therefore, Section 3.2.1, which presents the baseline reanalysis blocking statistics, is in terms of the multi-reanalysis mean. The 14 CMIP5 models listed in Table 3.2, which were obtained from the Lawrence Livermore National Laboratory Earth System Grid Federation (LLNL ESGF) portal, are comprised of the first ensemble member (r1i1p1) output from historical simulations and have a daily temporal resolution with the timestep centered on 12Z. The overlap period between all reanalysis and model data is 1979-2005 (27 years), which is less than the 30-year convention that is considered robust enough to account for natural variability, but an assessment of 25-year versus 30-year blocking climatologies did not produce any significant differences, implying that the 27-year time period is sufficiently robust for the purposes of this chapter.

Institute	Dataset	horizontal resolution
JMA	JRA-55	1.25x1.25
ECMWF	ERA-Interim	0.75x0.75
NASA	MERRA-2	0.5x0.625
NCEP	CFSR	0.5x0.5

Table 3.1: Reanalysis models and horizontal resolution (lat/lon)

The results for the HadCM3 and HadGEM2-CC models (Table 3.3), which are also CMIP5 models, are shown for comparison purposes, but were not included in the CMIP5 model ensemble average. Therefore, when referring to the CMIP5 models, we refer to the models listed in Table 3.2; we refer to the models in Table 3.3 as the Hadley models. The reason for this exclusion is that the seasonally averaged blocking frequency in the Hadley models is approximately 5-10% higher than that of the CMIP5 models in both DJF

and JJA, a result that has also been noted in Masato et al. (2013) (although a different detection algorithm was used to produce the blocking climatology). We use these two models to explore whether biases in the wind or height fields are sufficient to explain the significant deviations of these two models from the rest of the CMIP5 ensemble.

Institute	Dataset	Horizontal Resolution
BCC	BCC-CMS1.1	T106 (1.12°)
CCCMA	CanESM2	T42 (2.79°)
CMCC	CMCC-CESM	3.4x3.75
	CMCC-CM	T159 (0.75°)
CNRM-CERFACS	CNRM-CM5	T85 (1.4°)
NOAA-GFDL	GFDL-CM3	2x2.5
	GFDL-ESM2M	1.5x2.5
IPSL	IPSL-CM5A-LR	1.9 x 3.75
MIROC	MIROC5	T85 (1.4°)
	MIROC-ESM	T42 (2.8°)
MPI-M	MPI-ESM-MR	T63 (1.88°)
MRI	MRI-CGCM3	T106 (1.12°)
	MRI-ESM1	T106 (1.12°)
NCC	NorESM1-M	1.9 x 2.5

Table 3.2: CMIP5 models and horizontal resolution (lat/lon)

Institute	Dataset	Horizontal Resolution
MOHC	HadCM3	2.5x3.75
	HadGEM2-CC	1.25x1.875

Table 3.3: Hadley Center models and horizontal resolution (lat/lon)

3.1.2 Metrics for assessing agreement between datasets

We are assessing model agreement both in terms of model characteristics (Table 3.4) as well as the resultant blocking metrics (Table 3.5) once the detection algorithm has

been run on all of the models. The methodology for calculating block duration, distance traveled, size, and speed are all briefly summarized in Table 3.5, and were previously described in Chapter 2.

Blocking intensity (BI) was developed by Lupo and Smith (1995) and modified in Wiedenmann et al. (2002); we further modify it here to accommodate our 2D results (see Appendix B.1). BI is a measure of the Z500 gradient relative to overall field's climatology. In the reference papers, blocks are classified as weak ($BI < 2$), moderate ($2 < BI < 4.3$), and strong ($BI > 4.3$). We also present an anomaly intensity (AI) metric (Appendix B.2), which is an area-weighted sum of anomaly values over the detected block extent. The AI index complements the BI index, presenting the block's intensity while also accounting for the block's size and the underlying atmospheric mean state.

We approximate both upper and lower level daily mean jet positions for each model in the 27-year time period, as well as reanalysis and CMIP5 (excluding the Hadley models) means. Li and Wettstein (2012) note that jets in NP are both thermally- and eddy-driven, and that the dominant processes are seasonally dependent. NP jets are considered to be more subtropical in nature, due the influence of tropical Pacific heating patterns such as the El Niño Southern Oscillation (ENSO), as well as variations in geopotential height patterns such as the Pacific/North American teleconnection pattern (PNA). However, eddy forcing is still a driving force in development and maintenance of jet position as well; and in summertime, when the equator-pole temperature gradient weakens relative to wintertime, the maximum wind speed is observed to be in the vicinity of the eddy-driven winds.

The jet position calculations are performed using a methodology inspired by Woollings et al. (2010); the jet position is defined as the latitude of the maximum bandpass-filtered, longitudinally averaged wind speed over 15-65 degrees in the respective hemisphere. The multi-year averaged latitude values are further smoothed retaining only the first five harmonics of the multi-year averaged time series, similar to the procedure used in Chapter 2 for determining the LTDM.

The blocking frequency, U, V, and Z500 fields are all gridded fields, and we assess

Metric	Abbrev	Units	Description
Geopotential height	Z500	m	Geopotential height at 500 mb
zonal wind	U	m/s	Zonal wind velocity at 500 mb
meridional wind	V	m/s	Meridional wind velocity at 500 mb
Eddy-driven jet position	J850	°N/°S	Latitude of maximum zonally averaged wind at 850 mb
Subtropical jet position	J250	°N/°S	Latitude of maximum zonally averaged wind at 250 mb

Table 3.4: Model variables and abbreviations used in this paper

the model differences both in terms of the field's mean value and its variability over time (as measured by the standard deviation of per-gridpoint time series).While we utilize the median field bias values (see Appendix **??** for the calculation procedure) for both the seasonal mean and variability of the fields in the linear regression calculations, we also present the full multi-year seasonally averaged fields for JJA and DJF, with statistically significant differences indicated by stippling. These significant differences were determined by taking per-year seasonal averages (standard deviations) of the field for both the reanalysis average and the model of interest, then performing the t-test at each gridpoint, using the respective distributions of per-year values to assess whether there is a significant difference in the mean (standard deviation).

By evaluating both the model characteristics (Table 3.4) and the resultant block characteristics (Table 3.5), we gain a more comprehensive understanding of both the blocks produced by a given dataset, as well as the connection between the block characteristics and the inherent model physics. Blocks are formally defined as synoptic-scale, quasi-stationary features, but the thresholds for quantities such as duration or size vary by methodology, and consequently change the resultant blocking climatology. While some thresholds (5 day duration, 10^6 km^2 minimum size) have basis in the literature, other metrics, such as zonal speed, have no prior constraint (although we attempt to eliminate faster moving features by requiring a minimum of 50% overlap from one contour to the

Metric	Abbrev	Units	Description
Blocks per season	NB	count	Number of blocks per season
Mean centroid latitude	LA	°N/°S	Mean absolute latitude of block's centroid coordinate
Mean centroid longitude	LO	°E	Mean longitude of blocks centroid coordinate
Duration	DD	days	Number of time steps (in days) for which the block is identified (min 5 days).
Zonal Distance	ZD	km	Difference between block center start and end longitudes
Zonal Speed	ZS	km/hr	Zonal distance divided by duration
Block Area	BA	$km^2 \times 10^6$	Area of the contour which captures the cluster of detected points (min 10^6 km ^{2})
Anomaly Intensity	AI	none	Area-weighted sum of block anomaly values
Blocking Intensity	BI	none	Maximum geopotential height normalized by the local conditions

Table 3.5: Blocking metrics and abbreviations used in this paper

next over time).

It is worth noting that some of the block characteristics listed in Table 3.5 are not fully independent of one another. For example, speed is the result of duration divided by distance; also, larger blocks tend to have longer durations due to the minimum size threshold constraint.

3.1.3   Linear regression calculations

In order to assess the extent to which various model biases impact the resultant blocking frequency, we regress median difference values in the variables from both Table 3.5 and

3.4 against blocking frequency differences for the respective models. Additionally, we regress each variable bias in Table 3.4 variable against Table 3.5 variable biases. A similar analysis was previously conducted in Davini and D'Andrea (2016) to assess the impact of SST biases on blocking frequency. However, their analysis used root mean square error (RMSE), which, while quantifying absolute model error, does not retain the sign of the bias. The intent here is to investigate the relationship between model fields and the characteristics of the blocks that are produced by these models; therefore, RMSE does not work for our purposes. We use the block bootstrapping procedure described in Appendix ?? in an attempt to determine a meaningful bias magnitude in the context of the overall spatial pattern while retaining the sign of the bias.

For each of these analyses, the regression equation can be generally expressed as

$$m_i = c_0 + c_1 f_1 + c_2 f_2 + ... \tag{3.1.1}$$

where m_i is the normalized difference of the reanalysis average and the model for the variable in question, c_0 is a constant, and c_n are the slope coefficients for biases f_n.

The regression models are calculated using the ordinary least squares (OLS) regression calculation from the Python `statsmodels` package. Each variable bias V^* is normalized prior to the regression calculations by the reanalysis average's corresponding interquartile range, IQR_{V_R} (the difference between the 75th and 25th percentiles) for that region and season:

$$V_N^* = \frac{V_M - V_R}{IQR_{V_R}} \tag{3.1.2}$$

where V_M and V_R are the seasonally averaged values of the model and reanalyses, respectively and V_N^* is the normalized variable value. Normalization is necessary to ensure that biases with larger absolute magnitudes (such as the Z500 biases) do not factor disproportionately into the regression calculations, as well as taking regional and seasonal climatology differences into consideration.

Before performing the regression calculation, we determine whether any of the variables exhibit high collinearity, which will disproportionately weight the resulting model fit. The variance inflation factor (VIF) flags variables that have significant multicollinearity with other variables, which could skew the fit of the linear regression. The `statsmodels` package calculates VIF for each possible input value relative to all others, and an explanation of the calculation steps is found in Appendix B.4. Variables with a VIF greater than 5 (this is a commonly accepted threshold in the literature) are dropped from the list of input variables. We then regress blocking frequency against the remaining list of explanatory variables and examine the model fit in terms of the parameters in the subsequent section.

Interpretation of regression summary output: The results in Section 3.2.3 are displayed in terms of the variable coefficients c_n, the coefficient of determination, R^2, the t-statistic, and the p-value. Each of these values relays a different piece of information about the fit of the regression equation:

- c_n: The relative weight of the corresponding explanatory variable in the overall linear equation. Since all of the explanatory variables are standardized, coefficients with higher magnitudes can be considered to explain a larger proportion of variance in the outcome. Negative values imply a negative correlation between the dependent and independent variables.

- R^2: The amount of variance in the independent variable that can be described by the explanatory variables. A high R^2 value implies that much of the variance in the independent variable is correlated with variance in the explanatory variables.

- t-statistic: The coefficient divided by the standard error of the coefficient. A larger value of the t-statistic implies that the coefficient is large, the standard error of the coefficient is small, or both.

- AIC and BIC: metrics for quantifying relative model quality, accounting for both overfitting (too many parameters) and underfitting (poor predictive value). Lower

values of AIC and BIC imply a better model. (See Appendix B.4 for the formulas for AIC and BIC).

- p-value: The probability that the coefficient c_n is 0 (implying that the variable has little to no influence in the resultant regression equation). At the 95% confidence interval, $p < 0.05$ implies that the coefficient is nonzero and the corresponding explanatory variable should be included in the linear regression equation.

High influence outliers: The `statsmodels` package contains a function for detecting models that contain significantly anomalous input or output values. The outlier models are described in terms of the following quantities:

- Leverage: A measure of how different the individual model's explanatory variables are from the entire set of explanatory variables as a whole.

- Studentized residual: Residuals are the difference between the observed and predicted values of the regression equation. A studentized residual is an observation's residual divided by the standard deviation of all other residuals (i.e. excluding the residual that is the dividend).

- Cook's distance: A measure of how much the values in the regression model change when a particular observation is removed from the calculations.

See Appendix B.4 for the equations related to these measures, as well as the supplemental figures for plots of the models. The function outputs labels for the models that surpass either the leverage or residual thresholds, regardless of the resultant influence value (measured by Cook's distance). Note that a high influence value is not necessarily a detrimental impact to the model fit; it merely signifies that removing that observation from the calculation will result in a significantly different model fit.

3.2 Results

We begin in Section 3.2.1 with the baseline values from the average of ERA-Interim, JRA, MERRA2, and CFSR for JJA and DJF, with some brief remarks on seasonal and regional

differences. These values are the basis for the model comparisons in Section 3.2.2, and the interquartile range, or the difference between the 75th and 25th percentiles, is the quantity that is used to normalize the model-reanalysis differences. Section 3.2.3 presents a linear regression analysis between blocking frequency model bias and biases in other block and model characteristics. The regression is first performed on the full set of models, and then a second time with significant outlier models removed in order to determine if variable significance changes without the outlier models present. The individual block variables are also regressed against the model variables in order to explore possible explanations for variations in block characteristics between models.

3.2.1 Reanalysis average of blocking metrics

Tables 3.6 and 3.7 present the per-season median and interquartile range of blocking metrics for NP and SP, with the corresponding summer season values on the left and the corresponding winter season values on the right of the table. For the blocking frequency, U, V, and Z500, the discussion also references Figures 3.1-3.4, which show the reanalysis multi-year seasonal average and variability plots for the specified region and season.

Blocking Frequency (BF): Blocking frequency ranges from the smallest blocking frequency in SP DJF, with a maximum of 6%, to the largest in NP DJF, with a maximum blocking frequency of 14%. Figures 3.1 to 3.4 display BF fields in the top left corner of each figure. Blocks are more concentrated towards the poleward latitudes in the respective summer hemispheres compared to the winter counterparts, which contain blocking spread out across the extent of the study region. Note that the BF field patterns for the average and temporal variability are very similar, although the magnitude of the variability field is on average more than double that of the mean field. This is unsurprising, as areas with higher blocking frequency will have more instances of switching from a non-blocked state (0) to a blocked state (1) in the instantaneous field.

Number of blocks (NB): The number of blocks per regions and season ranges from 3 to 8 blocks per season, with NP JJA having the lowest range of values (3 to 7) and SP JJA having the highest range of values (4.75 to 8).
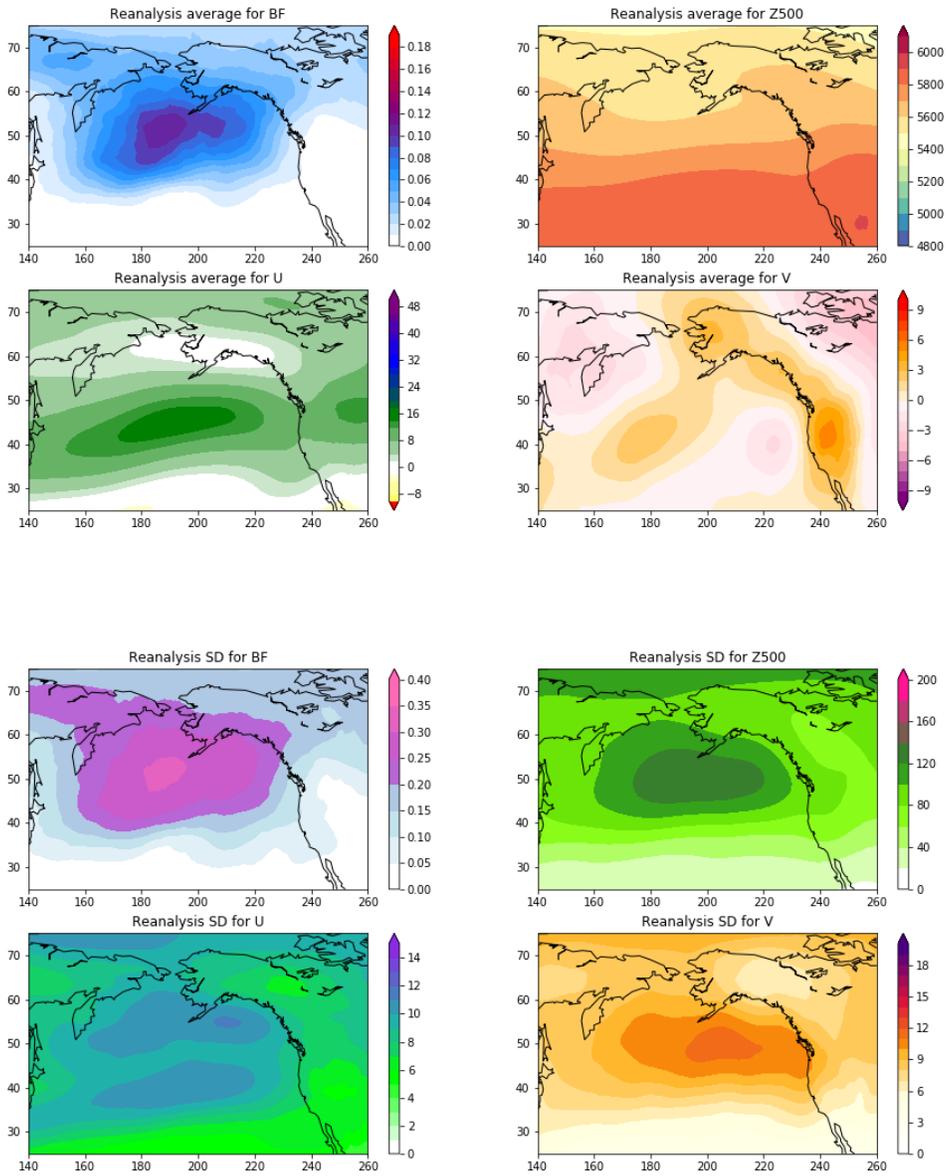
Figure 3.1: (Top subfigure) NP JJA mean fields for (top left) blocking frequency in 0.02 interval spacing starting at 0.02, (top right) Z500 with 100 m interval spacing, (bottom left) U with 5 m/s interval spacing, and (bottom right) V with 2 m/s interval spacing. (Bottom subfigure) NP JJA standard deviation fields for (top left) blocking frequency in 0.05 interval spacing starting at 0.05, (top right) Z500 with 20 m interval spacing, (bottom left) U with 1 m/s interval spacing, and (bottom right) V with 1 m/s interval spacing.
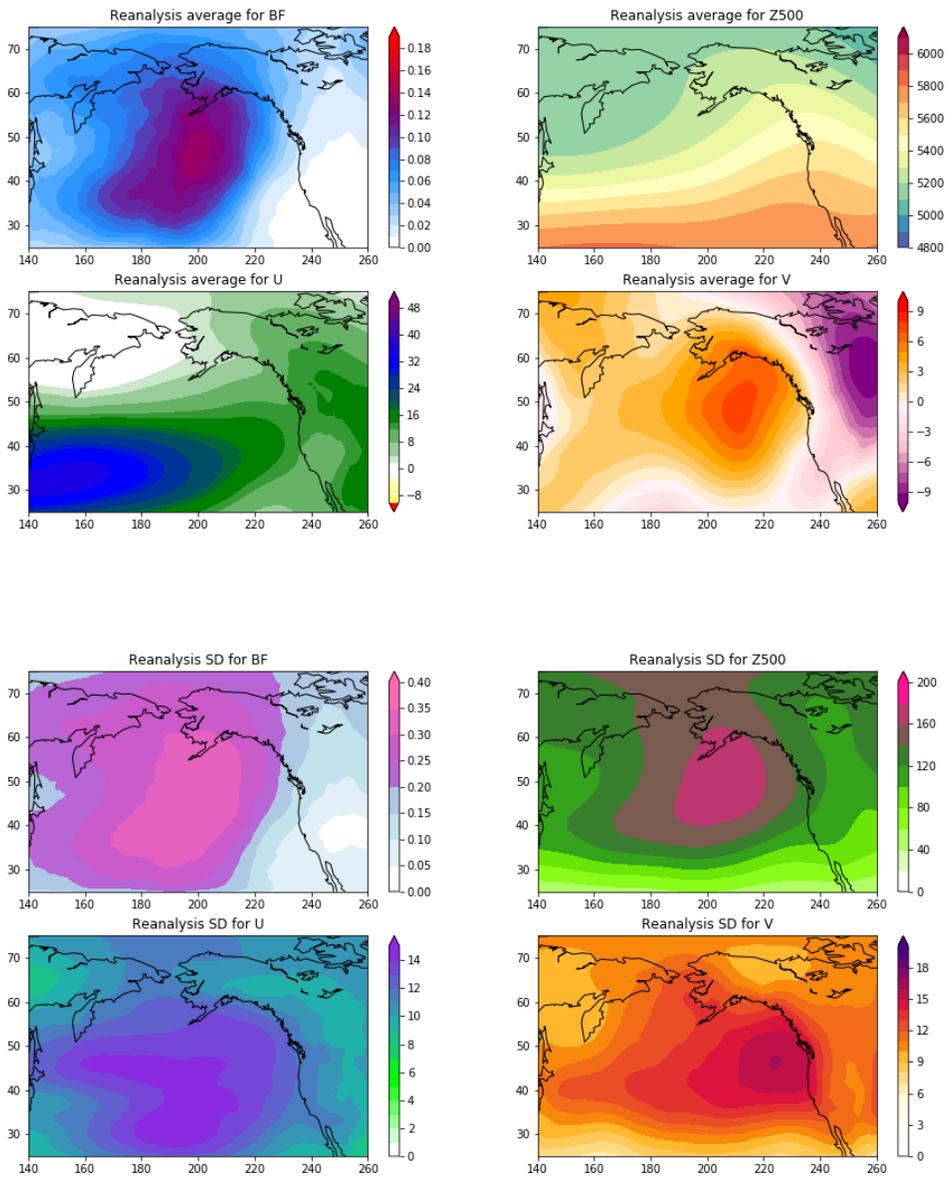
Figure 3.2: As for 3.1, with NP DJF.
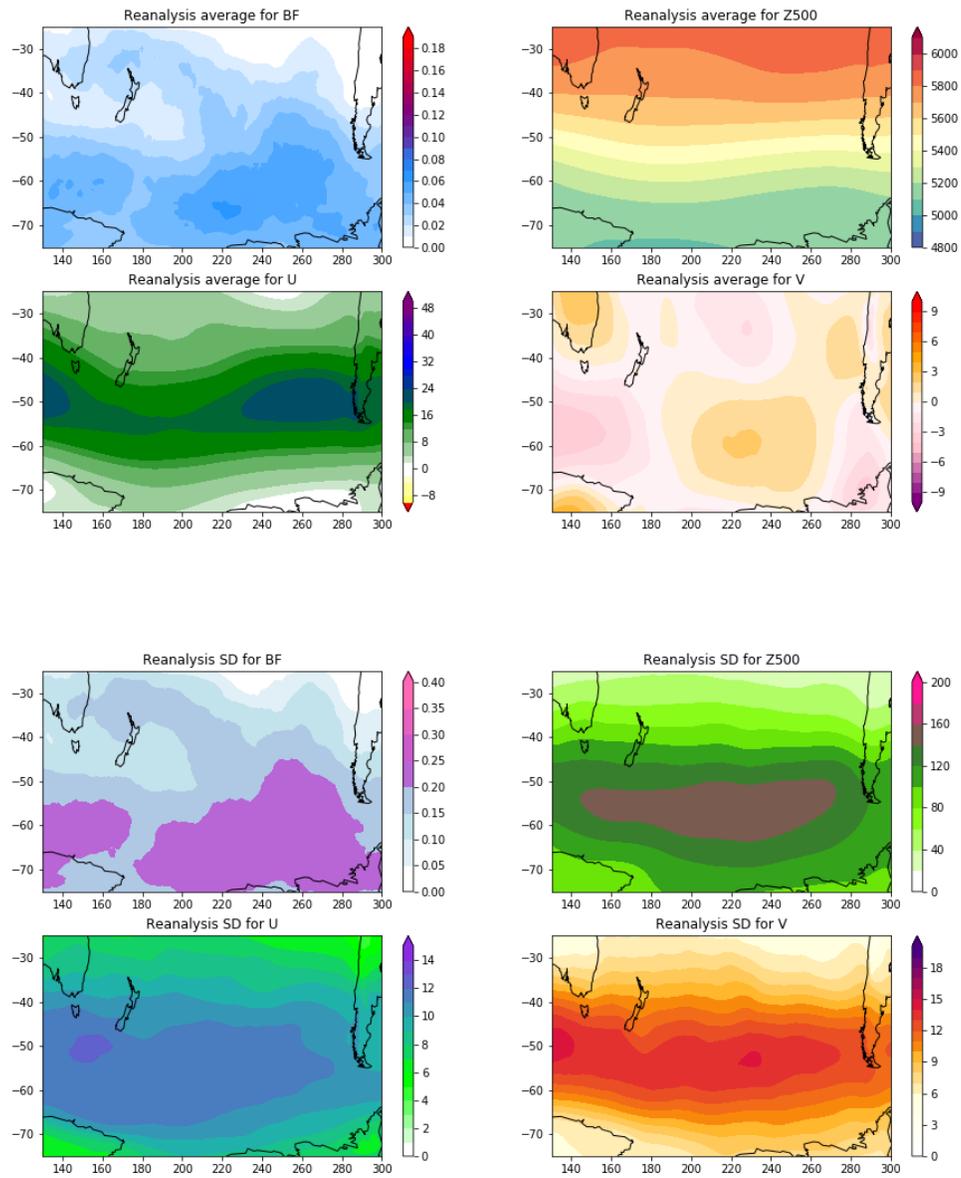
Figure 3.3: As for 3.1, with SP DJF.
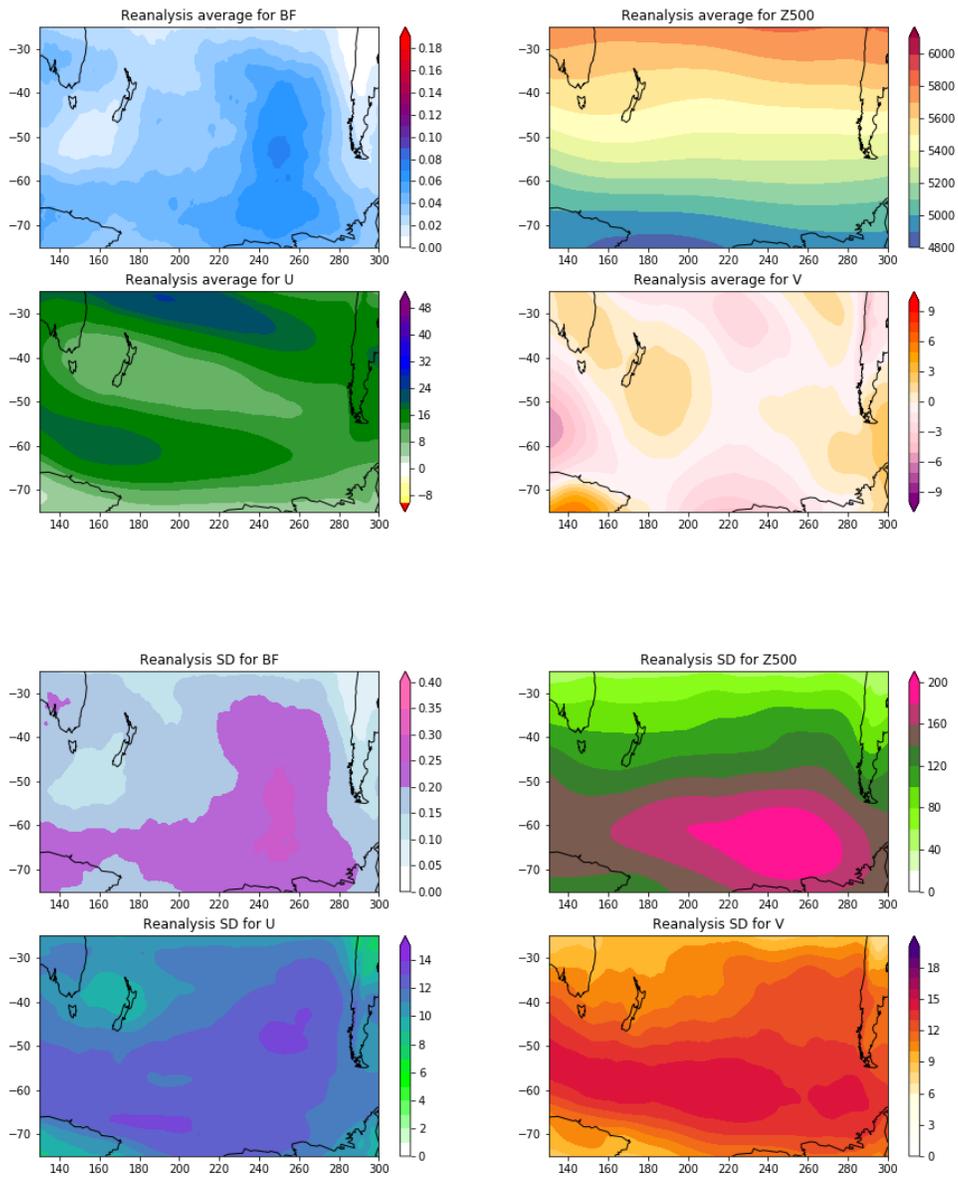
Figure 3.4: As for 3.1, with SP JJA.

	JJA			DJF		
	Q25	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**
BF (%)	1.349	3.660	5.610	3.537	6.703	9.160
BF SD (%)	11.665	18.951	23.069	20.320	25.618	29.003
NB (#)	3.500	4.250	7.000	4.500	5.750	6.750
LA (N)	48.320	51.690	59.900	35.660	44.680	56.310
LO (E)	174.970	195.000	217.990	174.050	192.390	219.310
DD (days)	5.000	7.000	9.000	6.000	7.000	10.000
ZD (km)	431.420	837.210	1418.190	429.510	1022.250	1834.100
ZS (km/hr)	3.960	6.570	9.560	4.280	7.420	11.530
BA (km2)	2.280	3.070	3.750	3.390	4.420	5.950
AI	4.260	6.440	8.100	9.280	13.190	20.540
BI	1.510	1.830	2.180	2.410	3.540	4.120
Z500 (m)	5599.895	5683.674	5814.641	5203.299	5330.329	5533.028
Z500 SD (m)	75.685	91.664	102.626	119.685	134.379	149.527
U (m/s)	3.792	6.251	9.475	5.180	11.813	16.206
U SD (m/s)	8.094	9.135	9.895	10.960	12.194	13.453
V (m/s)	-0.730	0.591	1.598	0.490	2.373	3.642
V SD (m/s)	7.454	8.461	9.393	10.913	12.050	13.251
J850 (N)	44.160	46.530	48.820	38.150	38.820	40.370
J250 (N)	44.820	48.350	51.310	33.510	33.910	35.430

Table 3.6: NP values for the blocking frequency (BF), the various blocking metrics listed in Table 3.5, and the field variables listed in Table 3.4. The values are listed for the 25th (Q25), 50th (Q50), and 75th (Q75) percentiles.

Mean centroid latitude (LA): The summer (winter) seasons tend towards more poleward (equatorward) latitude values, with a 50th percentile value of 52N (45N) in NP and 50S (47S) in SP. However, there is less of a difference between the winter and summer seasons in SP than in NP; the overall flow pattern in SP is much more zonal due to the comparative lack of land mass (with the exception of SE Australia on the western edge

	DJF			JJA		
	Q25	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**
BF (%)	2.196	3.301	4.535	2.999	3.981	4.892
BF SD (%)	14.664	17.672	20.800	17.217	19.712	21.857
NB (#)	3.250	5.500	7.000	4.750	6.750	8.000
LA (S)	39.040	50.290	59.110	35.920	47.480	59.010
LO (E)	172.880	206.790	252.130	184.780	225.590	255.740
DD (days)	5.000	6.000	8.000	5.000	6.000	9.000
ZD (km)	379.870	978.780	1662.690	850.130	1514.100	2232.320
ZS (km/hr)	4.460	8.800	12.420	6.280	10.460	14.050
BA (km2)	2.730	3.540	4.270	2.880	3.640	4.570
AI	6.470	9.470	12.680	8.150	11.510	15.590
BI	1.770	3.370	4.250	2.030	2.860	4.170
Z500 (m)	5258.736	5494.509	5703.509	5176.352	5378.461	5534.993
Z500 SD (m)	91.907	118.692	135.107	110.503	142.010	165.860
U (m/s)	9.335	13.560	17.250	12.186	13.777	16.062
U SD (m/s)	9.676	10.660	11.304	11.417	12.123	12.460
V (m/s)	-0.936	-0.039	0.815	-0.944	-0.153	0.601
V SD (m/s)	9.252	11.122	12.633	11.715	12.876	13.795
J850 (S)	52.850	53.260	53.540	53.470	54.280	56.040
J250 (S)	47.820	50.240	50.920	32.130	32.530	32.760

Table 3.7: As for Table 3.6, with SP values

and the tip of South America on the eastern edge).

Mean centroid longitude (LO): In contrast to LA, there is almost no difference between the winter and summer seasons in NP, again likely due to the land mass, which is responsible for reinforcing a standing wave pattern in NP regardless of season. In SP, a greater number of blocks tend more eastwards within the region in winter compared to summer, with a median value of 207E in DJF and 226E in JJA.

Duration (DD): In general, the summer months tend towards shorter blocks than the winter months, with SP DJF having the shortest durations (5-8 days) and NP DJF having the longest durations (6-10 days). The longer-lasting blocks in NP DJF explain why this study region has the largest blocking frequency, despite SP DJF experiencing more blocking events (but shorter ones) overall.

Zonal distance traveled (ZD): Here, we see that in general SP displays larger distance values for a given season than the northern counterpart, although SP DJF has a larger proportion of short-distance blocks than NP JJA (likely connected to shorter durations).

Zonal speed (ZS): SP has faster speeds compared to the NP speeds for corresponding seasons, which is due to both longer distances and shorter durations than NP.

Block area (BA): NP DJF has the largest blocks and NP JJA has the smallest, with SP showing similar distributions between winter and summer seasons.

Anomaly Intensity (AI): Summer blocks have lower AI values than winter blocks in both regions, with the highest overall AI values in NP DJF. The 75th percentile AI value in NP DJF is 30% higher than the 75th percentile value for SP winter.

Block Intensity (BI): NP JJA produces mainly weak ($BI < 2$) blocks; NP DJF and SP in both seasons have blocks with stronger Z500 gradient values. The 75th percentile BI value for SP JJA is slightly higher than the 75th percentile value for NP DJF, which is seemingly at odds with the previously noted comparison of AI values. However, NP DJF blocks are on average much larger; the BI index does not factor this into the calculation.

500 mb geopotential height (Z500): NP JJA has the highest heights on average (about 5700 m) with a maximum height of almost 6000 m in the equatorward parts of the region. The winter seasons of both regions average about 5400 m with maximum heights of about 5800 m, but the patterns are quite different. NP DJF has a tighter horizontal Z500 gradient over the lower western half of the region, which is mainly ocean, compared to the weaker horizontal gradient in the eastern part of the region (over the North Atlantic land mass and the Pacific Coast Ranges). SP in both seasons displays a mostly zonal average height pattern with the strongest Z500 gradient across the middle latitudes in

the region, with lower heights in JJA (winter). The largest amount of variability is seen in the upper mid-Pacific in NP (approximately co-located with the BF maximum), and slightly polewards of the center of SP.

500 mb zonal wind (U):   The bottom left panels of Figures 3.1 and 3.2 show that in NP, there is a concentrated region of stronger zonal wind in the lower half of the region that corresponds to the jet core, with a maximum of about 15 m/s in JJA and 36 m/s in DJF and wind speeds averaging about 6 m/s in JJA and 11 m/s in DJF. In SP (Figures 3.3 and 3.4, bottom left), the summer (DJF) has a more concentrated band of stronger wind around 50S (poleward of the Australian continent), with average winds of about 11 m/s; the winter wind speed displays a maximum of 22 m/s at the equatormost (25S) latitudes, with a lesser local maximum of 18 m/s around 60S and overall average wind speeds around 14 m/s. As with Z500, variability tends to be higher in winter seasons compared to the summer. The location of variability maxima are not co-located with the location of the mean field maxima; rather, the largest variability is associated with shifting of the maximum zonal wind position, in both the zonal and meridional directions.

500 mb meridional wind (V):   There is an alternating pattern of positive/negative regions of V, with the largest opposing magnitudes in NP DJF (the central part of NP DJF, off the coast of North America, is about -10 m/s). The mean meridional wind speeds are an order of magnitude smaller than zonal wind, ranging from about -2 to 4 m/s among the various study regions. The positive maxima in the summer seasons of the respective hemispheres are shifted westwards relative to winter, which corresponds to the westward shift of the zonal wind maximum. This indicates that the large-scale wave pattern is different between the two seasons, with the winter wave pattern displaying a larger amplitude, higher wind speeds, and a trough centered over the ocean basin, although the difference is much more notable in NP than in SP. However, unlike the largely positive zonal wind values observed in the seasonal averages of U, V has regular occurrences of both positive and negative values depending on the position of the large scale wave pattern; thus, the V variability field has a larger maximum magnitude than the U variability field.
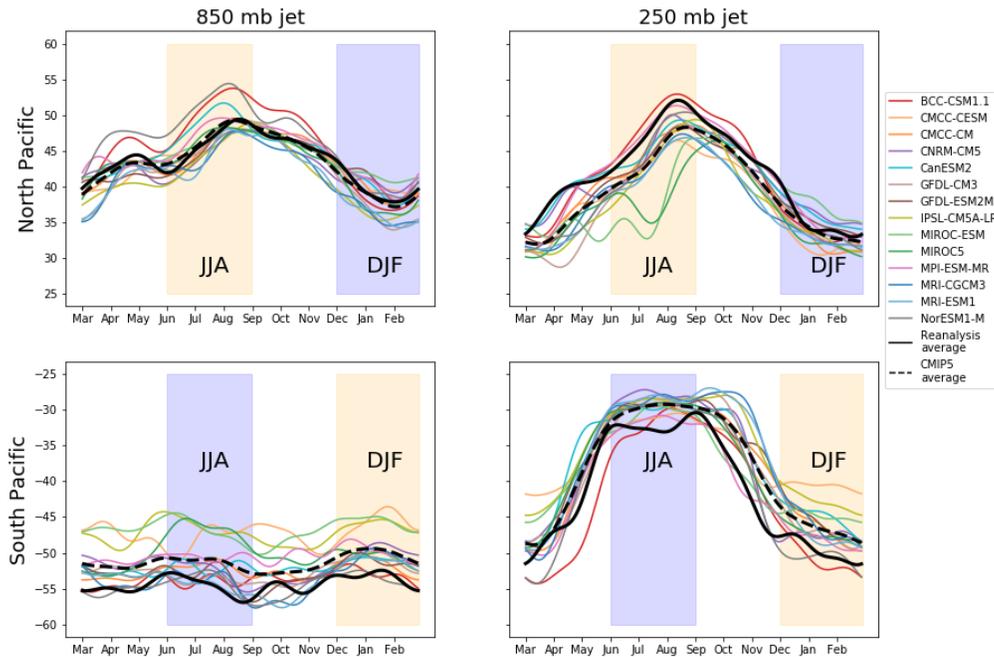
Figure 3.5: Position of the low-level (850 mb, left column) and upper level (250 mb, right column) jets for NP (top row) and SP (bottom row). The reanalyis average is denoted by the black line and the individual models are the colored lines. Summer months are highlighted in orange and winter months are highlighted in blue for the respective regions.

850 and 250 mb jet position (J850, J250): Woollings et al. (2010) notes a weak bimodality in jet position in both NP and SP, and that observation is reflected here. The relative positions of upper and lower jet positions display distinct seasonal trends for both hemispheres. In NP, there is a 2-degree difference between the median maximum wind positions, and in SP there is a 3-degree difference; this vertical coherence is also evidenced by the latitudinal position of the zonal wind maxima (see supplemental material, Figures 3.1 and 3.3). During the summertime, the meridional temperature gradient and, subsequently, the upper level jet, is weaker; thus, the maximum wind position will be in the vicinity of the eddy-driven jet in both the upper and lower levels.

In contrast, there is more separation between the positions of the upper-level and lower-level winds in wintertime, with approximately a 5 degree positional difference between

the upper and lower winds in NP and a 22 degree positional difference between the upper and lower winds in SP. During the winter, the eddy-driven jet is present at both the upper and lower vertical levels, but the upper level jet traps eddies, which subsequently reduces meridional propagation (Nakamura and Sampe, 2002). In SP, the 850 mb jet position is fairly constant throughout the year, without the equatorward migration in the summertime as observed in NP; this is likely due to the strength of the meridional temperature gradient (influenced by the Antarctic land mass) and the unimpaired flow in this region.

3.2.2 Model biases with respect to the reanalysis average

Figure 3.6 visualizes the strength of the relationship between pairs of variable model bias values, using the Pearson correlation (similar to Chapter 2) on the set of model bias values (CMIP5 and the outlier Hadley models both) for both hemispheres and seasons. Figure 3.7 presents the sign and significance of these biases in all measured quantities— blocking frequency, block characteristics, model fields, and the calculated jet positions. More detailed per-region figures with additional detail on difference magnitudes are available in the supplemental materials (Figures C.1 to C.4), as well as individual plots of the model fields in terms of the mean and variability, and Taylor plots that summarize the pattern and magnitude differences for these fields (Figures C.5 to C.52). We also present the interquartile range of all model bias values in Tables 3.8 and 3.9, where differences that meet the level of significance are noted with bolding. Significant differences are denoted by dark red (positive) and dark blue (negative), while non-significant differences are light red and blue for the respective signs of the biases. The differences for the field variables (blocking frequency (BF), geopotential height (Z500), and the two wind variables (U and V)) come with the caveat that the difference values somewhat oversimplify the more complex spatial differences in the full field. To maintain consistency with NP, the biases in the latitude values (LA, J850, and J250) are positive for a poleward bias and negative for an equatorward bias (this is not reflected in Figure 3.5, where the latitude values maintain the sign).

As discussing individual model differences would be exceedingly long-winded, we

choose to focus on broader trends in the the significant differences, and reserve discussion of individual models for the linear regression analysis.

3.2.2.1 Correlation between model biases

Figure 3.6 shows the Pearson correlation values that signify the strength and direction of the relationship between the denoted row and column variables. This relationship is calculated using the 64 bias values (4 bias values for each of the 16 models, as this includes the NH and SH winter and summer biases) for each of the two variables being compared. The correlation is calculated using the per-model variable bias pairs.

BF (leftmost column in Figure 3.6) is highly correlated with BA (0.89), and moderately correlated with DD (0.55), NB (0.49), AI (0.56), and BI (0.66). The almost perfect correlation between BF bias and BF variability bias (0.99) echoes the previous observation that the patterns between the mean and variability fields were very similar in the reanalysis. In contrast to previous literature, there is a very low correlation (-0.11) between BF and the Z500 bias value, and even less (-0.06) with the Z500 variability bias field. This is interesting, considering the literature that has explored correcting biases in blocking frequency by correcting the mean bias, as in Scaife et al. (2011). This may simply be a limitation of the mean bias value without its spatial component, but it is also important to consider that we are examining the accurate depiction of extreme events, and the representation of the mean state does not necessarily correlate with the representation of outlier cases.

Some other notable correlation values lead to those variables being excluded from the final list of input variables that are used in the linear regression. This is due to the fact that strong correlations between input variables will introduce redundancy into the regression model and increase the standard error of the coefficients. The following variables have a VIF of more than 5 and are excluded from the input list:

- BA: BA bias is strongly correlated with BF bias (0.89), but its strong relationships with AI (0.73), BI (0.84) and moderately strong relationship with DD (0.66)

- U SD: U SD is highly correlated with V SD (0.89) and Z500 SD (0.7) and moderately correlated with J250 (0.5) and LO (0.43)
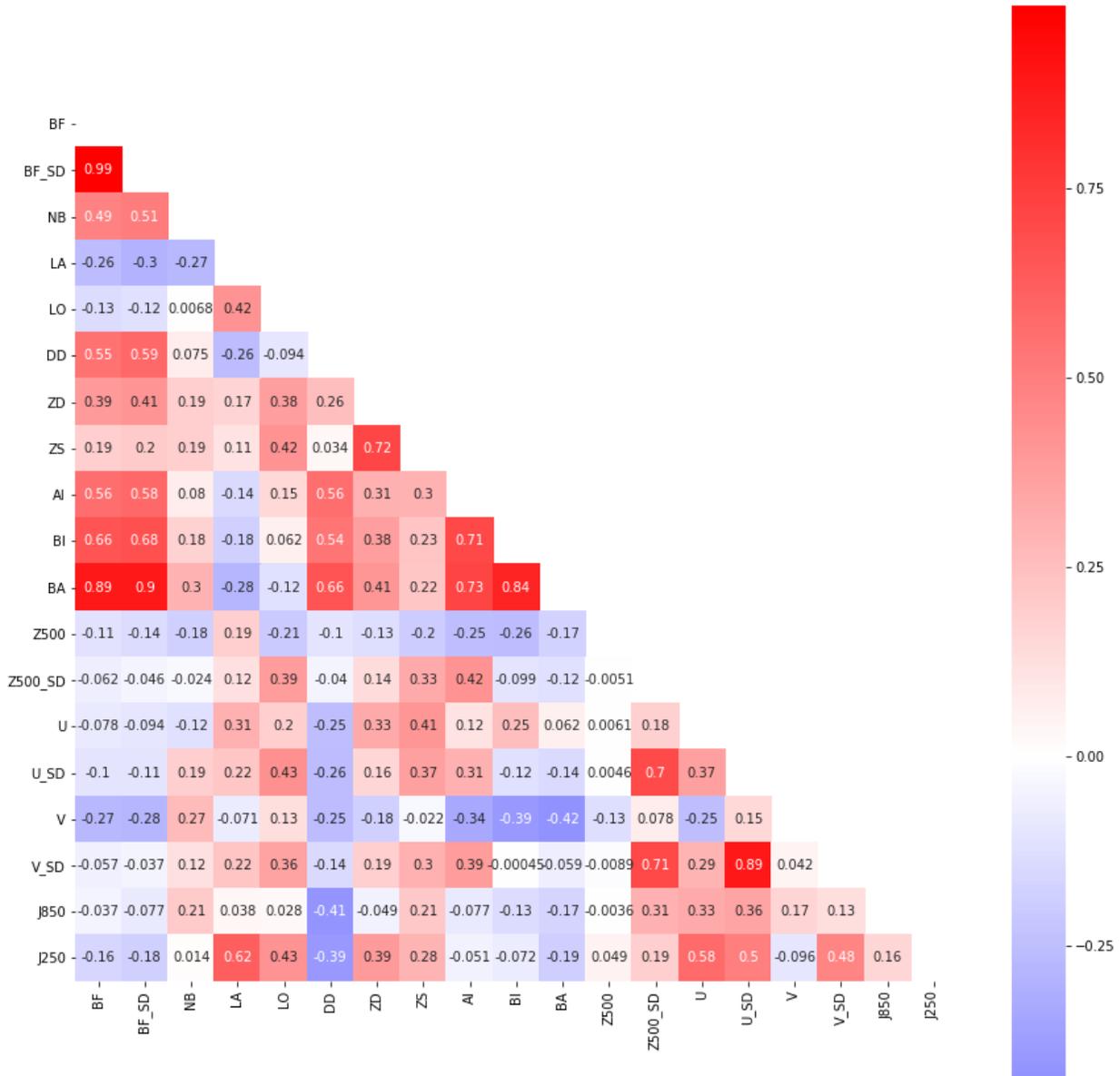
Figure 3.6: Pearson correlation values between model biases of the corresponding row and columns variables. Each relationship is determined between the set of all CMIP5 (and Hadley) model biases (one pair per model, region and season for a total of 64 bias pairs) for the specified variables.

	JJA			DJF		
	Q25	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**
BF	**-0.497**	**-0.121**	**0.309**	0.158	**0.533**	**0.917**
BF SD	**-2.151**	**-0.316**	**1.183**	**0.296**	**0.907**	**1.800**
NB	0.500	0.750	1.000	-0.750	0.250	0.250
LA	-1.093	0.401	0.751	1.091	2.409	3.476
LO	-1.869	0.654	3.604	6.511	**11.359**	**16.463**
DD	-1.000	-0.250	0.000	0.000	0.000	0.250
ZD	-82.958	-24.221	26.760	63.779	**198.322**	**227.634**
ZS	-0.693	-0.375	-0.031	0.516	1.205	1.826
BA	**-0.374**	-0.165	-0.063	-0.317	-0.175	-0.028
AI	**-1.577**	**-1.110**	-0.441	-0.042	0.565	1.775
BI	**-0.372**	-0.194	-0.032	**-0.456**	-0.283	0.022
Z500	**-39.806**	**-19.686**	**11.509**	**-55.265**	**-36.010**	-3.319
Z500 SD	**-6.510**	**-3.347**	-0.384	-0.797	**6.985**	**15.819**
U	**-1.120**	**-0.283**	**0.452**	**-1.072**	**-0.786**	**0.441**
U SD	**-1.746**	**-1.263**	**-0.833**	**-1.299**	**-1.053**	**-0.708**
V	**-0.109**	**0.128**	**0.177**	**-0.452**	0.004	**0.296**
V SD	**-1.645**	**-1.290**	**-0.918**	**-1.427**	**-1.188**	**-0.745**
J850	-1.269	-0.235	**1.836**	**-2.332**	0.051	**1.029**
J250	**-5.407**	**-4.061**	**-2.319**	**-1.681**	**-0.873**	0.119

Table 3.8: NP model bias value quantiles, as seen in Figure 3.7 (excluding the CMIP5 mean but including the Hadley models). The values are listed for the 25th (Q25), 50th (Q50), and 75th (Q75) percentiles. Significant difference values are bolded.

- J250: J250 is moderately correlated with LA (0.62), LO (0.43), U (0.58), U SD (0.5) and V SD (0.48)

- AI: AI is highly correlated with BA (0.73) and BI (0.71). It is also moderately correlated with DD (0.56), Z500 SD (0.42)
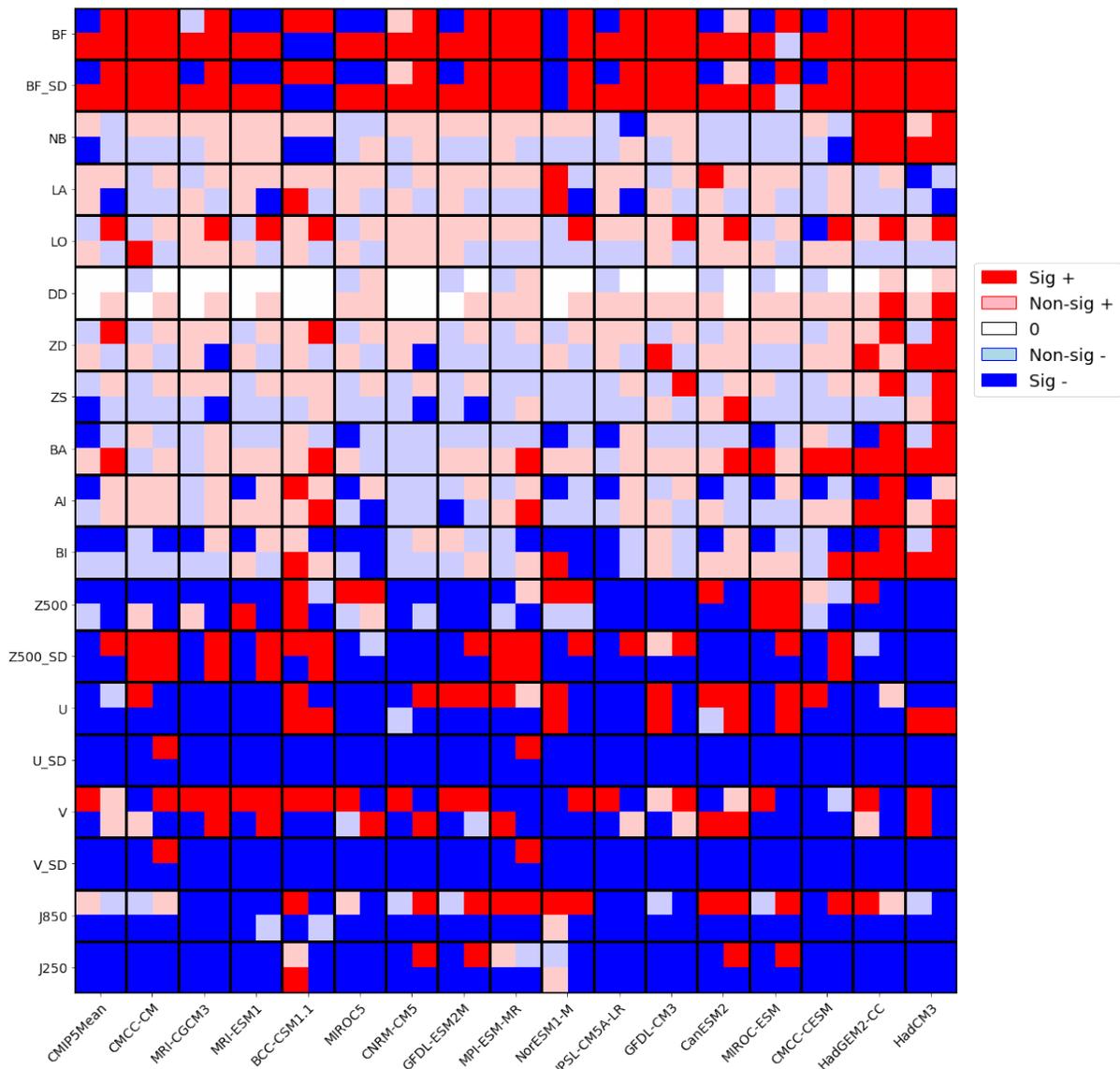
Figure 3.7: Significance and sign of biases for metrics in Tables 3.4 and 3.5. The top half of each box represents NP biases and the bottom half of each box represents SP biases. The left side of each box represents the summer months and the right side of each box represents the winter months. Significant positive (negative) biases are denoted by dark red (blue); non-significant bias values are indicated by lighter shades. White represents instances in which the median values for both the reanalysis and model values are the same. The gridded variables (BF, Z500, U, and V) have rows for both the mean and variability bias values; the variability component is denoted with the suffix "_SD".

	DJF			JJA		
	Q25	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**
BF	**0.214**	**0.570**	**0.720**	**0.360**	**0.605**	**0.831**
BF SD	**0.497**	**1.359**	**1.790**	**0.736**	**1.374**	**1.750**
NB	**-1.500**	-0.500	-0.250	-1.000	-0.250	0.500
LA	-0.937	1.268	2.486	**-4.782**	**-3.919**	-2.848
LO	-0.452	7.747	12.287	-5.732	-3.903	-0.875
DD	0.000	0.750	1.000	0.875	1.000	1.000
ZD	17.644	93.462	273.108	-199.203	-128.822	-44.176
ZS	**-1.145**	-0.848	-0.351	-1.410	-0.726	0.074
BA	-0.001	0.178	0.424	0.153	**0.272**	**0.497**
AI	-0.519	0.033	0.379	-0.883	0.372	**2.047**
BI	-0.090	0.011	0.396	-0.235	-0.016	0.196
Z500	**-26.022**	-7.397	1.593	**-33.978**	**-15.903**	-8.781
Z500 SD	**-13.801**	**-6.677**	**-3.978**	**-7.544**	**-3.284**	**4.200**
U	**-1.454**	**-0.805**	-0.030	**-1.433**	**-0.858**	-0.137
U SD	**-2.205**	**-1.749**	**-1.386**	**-1.605**	**-1.237**	**-0.766**
V	**-0.209**	-0.100	0.028	**-0.189**	-0.069	**0.076**
V SD	**-2.657**	**-2.381**	**-1.739**	**-2.121**	**-1.782**	**-1.465**
J850	**-3.462**	**-2.438**	**-2.171**	**-4.782**	**-2.137**	**-0.800**
J250	**-4.444**	**-2.986**	**-2.300**	**-3.774**	**-3.087**	**-2.138**

Table 3.9: SP model bias value quantiles, as seen in Figure 3.7 (excluding the CMIP5 mean but including the Hadley models). The values are listed for the 25th (Q25), 50th (Q50), and 75th (Q75) percentiles. Significant difference values are bolded.

3.2.2.2 Trends in bias values

Figure 3.7 shows an overall summary of the sign and significance of the biases for the various metrics in Tables 3.4 and 3.5. Some trends are particularly notable. First, few of the block characteristic metrics display significant differences compared to the reanalysis; the main exceptions are the two Hadley models (the two rightmost columns), which

have significant positive biases in NB, ZD, BA, and BI for almost all regions and seasons. Therefore, the Hadley models are consistently producing more blocks which are also larger and more intense, which subsequently means that they are able to persist longer and travel further. All of the models (including the Hadley models) overall display consistently negative bias values in many of the model variables, particularly the jet position variables and the wind variability variables. However, these model variables are not significantly correlated with BF bias on an individual basis, nor with many of the block characteristic bias values. There are strong correlations between the variability biases for Z500 and the wind fields, but this is unsurprising given the relationship between pressure gradient and wind speed.

Another observation of interest is the fact that BF bias, with the exception of NP JJA, is mainly positive, in both the mean and variability bias values. Previous blocking literature has maintained that models tend to underpredict the frequency of blocking, but many of these studies were focused on a different region (NH Atlantic) and used a different method (TM90, from Chapter 2). Here, we are using an anomaly-based objective detection algorithm rather than one based on an absolute field. Removing the mean bias from the absolute field will result in a better representation of the meridional Z500 gradient that is necessary for the TM90 algorithm. However, that is not the case here; we previously noted the apparent disconnect between the BF and Z500 field biases in the discussion of bias correlations. An anomaly-based method will search for deviations from the mean, regardless of the magnitude of the mean state, and the spatiotemporally-varying threshold ensures that extremes are detected in the context of the local climatology. Thus, it is possible that a mean model bias will not be reflected in the resultant blocking frequency bias, as long as the extreme events are properly represented.

3.2.3   Linear regression of bias values

3.2.3.1   Linear regression of blocking frequency against individual variables

Table 3.10 summarizes the resulting OLS fits for normalized BF bias against the normalized bias of each individual variable in the table. LO, ZS, Z500 (mean and variability), U (mean and variability), V variability, J250, and J850 did not have significant R^2 values

Variable	R^2	R^2 (adj)	intercept	coef
BA	0.79	0.79	0.1636	3.3613
BI	0.43	0.42	0.675	2.6872
AI	0.32	0.31	0.5595	3.1736
DD	0.30	0.29	-	3.9603
NB	0.24	0.23	0.4166	1.6664
ZD	0.15	0.14	0.3426	3.2181
LA	0.07	0.05	0.5002	-2.1561
V	0.07	0.06	0.406	-1.9492

Table 3.10: R^2, adjusted R^2, and coefficient values for linear regression fits of BF versus individual explanatory variables. LO, ZS, Z500, Z500 SD, U, U SD, V SD, J250, and J850 are not shown.

and are not included in the table. Of the individual fits, BA bias has the largest R² value at 0.79 (almost double that of BI, with the second largest R² values), which echoes the correlation results from earlier. While the Hadley models produce notably outsized BF biases compared to the rest of the models, eliminating them does not actually increase the R^2 values of most of the individual fits, as seen in Figure 3.8. It is clear from the left column in the figure that the Hadley models are exerting an outsized influence on the resulting fit; however, it is also clear from the overall R^2 values of the individual OLS results that one-to-one relationships are not sufficient to explain the variability in BF bias. While BA has the largest R^2 value out of all of the possible explanatory variables, it is also moderately to highly correlated with a number of other explanatory variables (hence its exclusion from the input list for the BF OLS calculations), which implies that BF bias could also be explained by a combination of the variables that are highly correlated with BA (i.e. the intensity variables), as well as other variables that are moderately correlated with BF, such as NB.
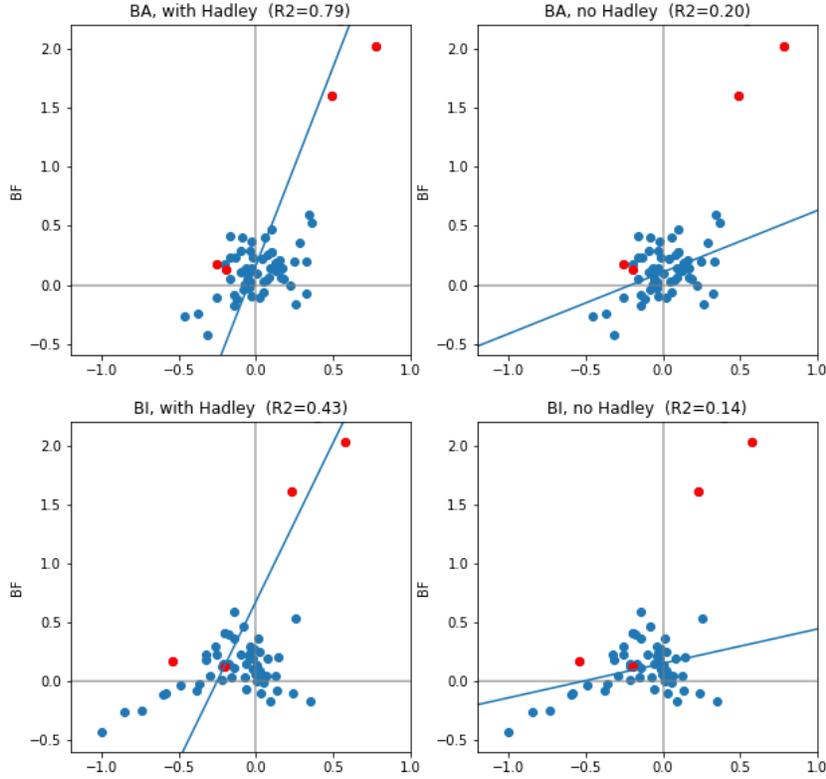
Figure 3.8: Individual scatter plots of BF versus (top) BA and (bottom) AI, including (left) versus excluding (right) the Hadley models (marked as red scatter points).

3.2.3.2 Linear regression of block characteristics against model characteristics

Table 3.11 shows the results of a series of OLS fits, in which each block characteristic from Table 3.5 was regressed against the set of model characteristics from Table 3.4, i.e.

$$C_i = a_0 + a_1 * Z500 + a_2 * U + a_3 * V + a_4 * J850 + a_5 * J250 + a_6 * Z500_{SD}$$

$$+ a_7 * U_{SD} + a_8 * V_{SD} \quad (3.2.1)$$

From these OLS fits, we can determine a number of relationships between block biases and the underlying model biases, although the largest R^2 value is 0.5 and therefore only explains a moderate amount of variability in the dependent variable. While correlation

variable	R^2	R^2 (adj)	intercept	Z500	U	V	J850	J250	Z500 SD	U SD	V SD
AI	0.50	0.43	-	-0.4592	-	-0.5362	-	-0.0424	-	-	-
LA	0.43	0.35	-	-	-	-	-	0.0588	-	-	-
LO	0.42	0.34	-	-0.2389	-	-	-0.0174	0.0351	0.2295	0.1791	-0.3163
BI	0.39	0.30	-	-0.6297	0.7022	-0.5282	-	-0.0568	-	-	-
DD	0.37	0.28	-	-	-	-0.2549	-0.0247	-0.0349	-	-	-
BA	0.32	0.22	-	-	-	-0.8265	-	-0.0681	-	-	-
ZS	0.28	0.18	-	-	0.2479	-	-	-	-	-	-
ZD	0.24	0.13	-	-	-	-	-	0.0286	-	-	-

Table 3.11: R^2, adjusted R^2, and coefficient values for linear regression fits of each block characteristic variable against the model characteristic variables. A dash denotes a non-significant coefficient value; BF and NB did not have a significant regression fit and the values are not shown.

and R^2 are not quantifying the same parameters, the sign of the individual OLS coefficients match that of the correlation values between the block and model variables, and significant relationships are mostly seen between variables that have correlation magnitudes greater than 0.3.

This section of the analysis reinforces the idea that the relationship between models and the resulting block characteristics is complex and not easily defined by pairwise relationships. For example, the AI bias OLS equation has Z500, V, and J250 biases as the significant explanatory variables. OLS fits calculated for AI bias against each of the other individual biases yield R^2 values on the order of 0.1; however, using multiple variables increases the R^2 value to 0.5. When this analysis was performed using only model biases against BF bias, the resulting OLS fit resulted in a non-significant regression (i.e. no relationship between BF bias and the model biases in combination). Therefore, BF bias cannot be explained by model biases alone. However, by including all possible variables in the BF bias regression calculation (excluding those that introduce multicollinearity), we are able to capture a greater amount of the variance in the BF bias and create a better model.

3.2.3.3 Linear regression of BF bias against block and model biases

These OLS calculations examine the relationship between per-model blocking frequency bias and all other possible explanatory variables, excluding the variables with a high VIF

(as noted in Section 3.2.2.1). The equation can generally be expressed as

$$BF[m] = B_0 + B_1 * NB[m] + B_2 * LA[m] + B_3 * LO[m] + B_4 * DD[m] + B_5 * ZD[m]$$
$$+ B_6 * ZS[m] + B_7 * BI[m] + B_8 * U[m] + B_9 * V[m] + B_{10} * V_{SD}[m] + B_{11} * Z500[m]$$
$$+ B_{12} * Z500_{SD} + B_{13} * J850[m] \quad (3.2.2)$$

where B_0 is a constant and each B_n is the coefficient of the corresponding explanatory variable. In the first model fit (which will be referenced as OLS 3.1), all 64 data points (16 models with the winter and summer biases for NP and SP) are included in the regression calculation. The model is assessed in terms of the R^2 value (which determines how much of the variance is explained) and the variables that have significant coefficients in the regression equation. The Cook's distance metric is used to determine data points representing influential outlier models that are impacting the model fit. These points are excluded from the second calculation (which will be referenced as OLS 3.2), and the analysis procedure is repeated a second time.

Linear regression with all variables (OLS 3.1): Table 3.12 shows the output of OLS 3.1, which results in $R^2 = 0.74$ (adjusted $R^2 = 0.67$). While a total of 13 possible variables were used in the calculation, the majority of the variance in BF bias can be explained by five variables: number of blocks (NB), blocking intensity (BI), zonal wind (U), zonal distance (ZD), and centroid longitude (LO). LO has a negative coefficient value, meaning that a westward bias in block location (i.e. blocks forming further upstream) is associated with a positive blocking frequency bias, and vice versa. A negative U bias has been demonstrated in the literature to correlate with more blocking (He et al., 2014). BI (and to a certain extent, ZD) is associated with more intense, longer-lasting blocking events, and NB is associated with more overall events, so it stands to reason that positive biases in these variables would result in corresponding positive blocking frequency biases. From Figure 3.6, recall that BF bias is moderately correlated with NB(0.49) and BI (0.66), and less well correlated with ZD (0.39). U (-0.08) and LO (-0.13) have the lowest correlation values, which is interesting considering that both are somewhat associated

variable	coef	se	t	pval
NB	1.3291	0.318	4.178	0.000
BI	2.1322	0.467	4.568	0.000
U	-1.7406	0.737	-2.362	0.022
ZD	2.2493	0.997	2.256	0.028
LO	-1.9138	0.865	-2.212	0.032
V	-1.0191	0.650	-1.568	0.123
Z500 SD	0.7887	0.586	1.346	0.184
V SD	-0.5167	0.517	-1.000	0.322
DD	0.7790	0.813	0.958	0.343
J850	0.0474	0.051	0.935	0.354
LA	0.7110	0.800	0.888	0.379
Z500	0.5073	0.742	0.684	0.497
ZS	-0.4902	1.047	-0.468	0.642
B0	0.1131	0.362	0.312	0.756

$R^2 = 0.74$, Adjusted $R^2 = 0.67$, AIC=160.9, BIC=191.2

Table 3.12: Linear regression of blocking frequency biases as outlined in Equation 3.2.2. Variables are sorted from smallest to largest p-value, with the five significant variables above the double line.

with zonal flow (or lack thereof).

The variables that fall below the significance threshold are also a matter of interest, even if they are not relevant in the context of the regression equation. The J850 bias has a coefficient value that is close to 0. In hindsight, this is unsurprising, as the BF bias as quantified here is somewhat divorced from the spatial component of the original variable, and the correlation value between the BF and J850 biases is low. The large biases in jet position are notable, but do not help to explain BF bias in this instance. The model Z500 mean and variability biases also have low correlation values, despite many of the individual models exhibiting significant biases; this further reinforces the observation that the biases between BF and Z500 are largely unrelated.

Model	Lev	Resid	CD	BF	NB (+)	BI (+)	U (-)	ZD (+)	LO (-)
MIROC-ESM (NP JJA)	**0.631**	0.753	**0.070**	-0.27	-0.07	-0.85	-0.22	0.04	-0.19
HadCM3 (SH JJA)	0.414	**2.803**	**0.349**	6.86	0.69	1.05	0.13	0.51	-0.08
HadGEM2-CC (SH JJA)	0.301	**7.265**	**0.799**	7.55	0.69	0.77	-0.39	0.02	-0.06
MIROC5 (SH JJA)	0.391	**-2.278**	**0.219**	0.41	0.38	-0.20	-0.87	-0.16	-0.08

Table 3.13: Significant outliers in OLS 3.1 that exceed the threshold for the leverage (0.438), the studentized residual (2.008) and the Cook's distance (0.063), as well as the normalized bias values for BF and variables with $p < 0.1$. The signs next to the variable names signify the sign of the coefficient.

Using the Cook's distance metric described in the methodology section, four points were flagged as significantly influential outliers (Table 3.13); these points will be excluded in OLS 3.2. MIROC-ESM (NP JJA) has a large leverage value (i.e. significantly anomalous input values). The other three models have large studentized residual values (i.e. significantly anomalous predicted BF bias values compared to the actual BF bias).

Linear regression with influential outliers removed (OLS 3.2): In Section 3.2.3.1, removing the Hadley outlier points had a significant impact on the R^2 values of the OLS fits between BF bias and individual variables, often decreasing the amount of explained BF bias. Here, when the outlier points from Table 3.13 are removed from the dataset, the R^2 value increases slightly, from 0.74 in OLS 3.1 to 0.76 in OLS 3.2. While the amount of explained variance in BF bias has not greatly increased compared to OLS 3.1, it was

variable	coef	se	t	pval
NB	0.8607	0.143	6.016	0.000
ZD	1.6814	0.436	3.855	0.000
BI	0.7429	0.240	3.095	0.003
U	-0.7363	0.409	-1.799	0.079
V	-0.5197	0.296	-1.757	0.086
DD	0.5791	0.365	1.585	0.120
LO	-0.5085	0.405	-1.254	0.216
Z500	0.3669	0.474	0.773	0.443
B0	0.1235	0.165	0.748	0.458
ZS	-0.3332	0.489	-0.681	0.499
LA	-0.2020	0.371	-0.545	0.589
Z500 SD	0.1438	0.267	0.538	0.593
V SD	-0.1158	0.237	-0.488	0.628
J850	-0.0033	0.022	-0.148	0.883

$R^2 = 0.76$, Adjusted $R^2 = 0.70$ AIC=50.26, BIC=79.58

Table 3.14: Linear regression of blocking frequency biases against all possible explanatory variables as outlined in Equation 3.2.2, with high influence outliers from Table 3.13 removed

already high to begin with, and values of the model fit criteria, AIC and BIC, are much lower (i.e. better) in OLS 3.2 (granted, some of this is due to fewer data points, but also a better fit).

The magnitudes of NB, BI, and ZD have decreased (but so have the standard errors), and LO is now below the significance threshold. U has dropped to slightly below the $p < 0.05$ significance threshold, and V is now at almost the same significance level as U. The fact that 3 out of 5 of the original significant explanatory variables remain significant in OLS 3.2 indicates that overall, the relationship between NB, BI, ZD, and BF bias holds even when there are outliers present.

Model	Lev	Resid	CD	BF	NB (+)	BI (+)	ZD (+)	U (-)	V (-)
CanESM2 (NP JJA)	0.235	**2.136**	**0.093**	-0.08	-0.07	-0.38	-0.20	0.03	-0.04
HadGEM2-CC (SP DJF)	0.268	**6.335**	**0.568**	3.28	0.93	0.21	0.26	-0.18	0.01
CMCC-CESM (SP JJA)	**0.479**	-0.531	0.019	0.53	-0.54	0.25	0.02	-0.09	-0.87

Table 3.15: Significant outliers in OLS 3.2 that exceed the threshold for the leverage (0.467), the studentized residual (2.013) and the Cook's distance (0.067), as well as the normalized bias values for BF and variables with $p < 0.1$. The signs next to the variable names signify the sign of the coefficient.

The change in the model fit also means that there is a new set of outlier points in OLS 3.2. Table 3.15 shows the various parameters associated with the significant influential outliers in this new fit. CMCC-CESM (SP JJA) has anomalous input values, but a non-significant Cook's distance value. Another Hadley model (HadGEM2-CC SP DJF) has emerged as a significant outlier; in the previous model fit, it had a studentized residual value of 0.347, but with the removal of the outlier points in the previous iteration, its new residual value is 6.335 (the previously higher coefficients for NB and BI, as well as the influence of the extreme outlier Hadley models, had contributed to a better predicted fit value for HadGEM2-CC in the previous model). CanESM2 (NP JJA) also has a large positive residual value, although in this case the BF bias is slightly negative.

Discussion: Of the two models, the second model (with the dropped model outliers) has a better fit by all criteria; the R^2 value is higher and the AIC and BIC values are lower (Table 3.16) . There are still outliers in the second model fit as well, but considering

82

Regression	R^2	AIC	BIC	Significant Variables
OLS 3.1	0.74	160.9	191.2	NB, BI, ZD, U, LO
OLS 3.2	0.76	50.26	79.58	NB, BI, ZD, (U, V)

Table 3.16: Summary table of linear regression results from OLS 3.1 (all data points) and OLS 3.2 (outliers excluded). In the significant variables column, variables with $0.05 < p < 0.1$ are included in parentheses.

that the final regression equation has an R^2 value of 0.76 (which implies that 76% of the variation in blocking frequency bias can be explained by the explanatory variable bias), it is reasonable to draw conclusions using this model fit.

For each of the models, NB, ZD, and BI were consistently significant explanatory variables, with LO, U, and V as secondary explanatory variables. Influential data outliers like the SP JJA Hadley models had a significant impact on OLS 3.1; once those outliers were removed, the coefficients for some of the significant explanatory variables decreased by as much as half of the previous value. NB is related to the frequency of blocking occurrence, ZD and LO are related to the position of the blocking frequency, and BI is related to the relative persistence of blocks. U and V also affect position and persistence, but the relative uncertainty surrounding the averaged field bias values creates a higher standard error relative to the model coefficient, which can result in a non-significant coefficient value.

ZS, LA, DD, Z500, Z500 SD, V SD, and J850 were not significant explanatory variables in either of the two model fits. For all of these variables, the standard error equaled or exceeded the variable coefficient; there was too much uncertainty in the inputs to create a good fit with the BF bias.

Outlier data points in the models contained either anomalous inputs (leverage) or predicted BF bias values relative to the actual bias (residual). One example of a high leverage model is MIROC-ESM in NP JJA, which has a positive Z500 bias (Figure 3.9, top) that is an order of magnitude larger than that of other models (see Figure C.17 in the supplemental figures for reference.) While the mean field pattern closely matched that of the reanalysis (correlation of 0.95, see Figure C.19 for a Taylor plot), the variability of

Type of Outlier	Model	OLS	Observations
High Leverage (Inputs)	MIROC-ESM (NP JJA)	3.1	Strong + Z500 bias
			Strong - BI bias
	CMCC-CESM (SP JJA) *	3.2	low pattern correlation in V
High Residual (Prediction)	Hadley models (SP JJA)	3.1	Systematic positive BF bias
			Strong + NB bias
			Strong + BI bias
	MIROC5 (SP JJA)	3.1	Strong - U bias
	HadGEM2-CC (SP DJF)	3.2	Systematic positive BF bias
			Strong + NB bias
	CanESM2 (NP JJA)	3.2	Poleward BF concentration

Table 3.17: Outlier data points from OLS 3.1 and OLS 3.2, in terms of high leverage and high residual values. Note that CMCC-CESM was a significant outlier, but not an influential outlier.

the Z500 field over time (Figure 3.9, middle) was much lower compared to the reanalysis, while also producing blocks that had much weaker BI values on average (strong negative BI bias), and less of them (Figure 3.9, bottom). However, this was accounted for within the model; and while the predicted BF bias was higher than the actual BF bias value, it was not significantly so. Neither Z500 nor Z500 SD are considered to be significant explanatory variables within OLS 3.1, but this was due to high standard errors in the coefficient estimates rather than low coefficient magnitudes. Removing MIROC-ESM from the input data lowered the coefficient values in OLS 3.2 while not significantly changing the resulting model fit.

The high residual models consist of three Hadley models, MIROC5 in SP JJA, and CanESM2 in NP JJA. MIROC5 is the only data point with a negative residual value; the negative bias in the poleward half of SP is a major contributor to the model's predicted BF bias value, but the actual bias value ended up being lower. The Hadley models all display strongly positive biases in NB, and the SP JJA Hadley data points also have strongly positive BI biases as well. While the signs of the biases generally follow a pattern that contributes to positive BF bias values (positive NB and BI, negative U) these variables are
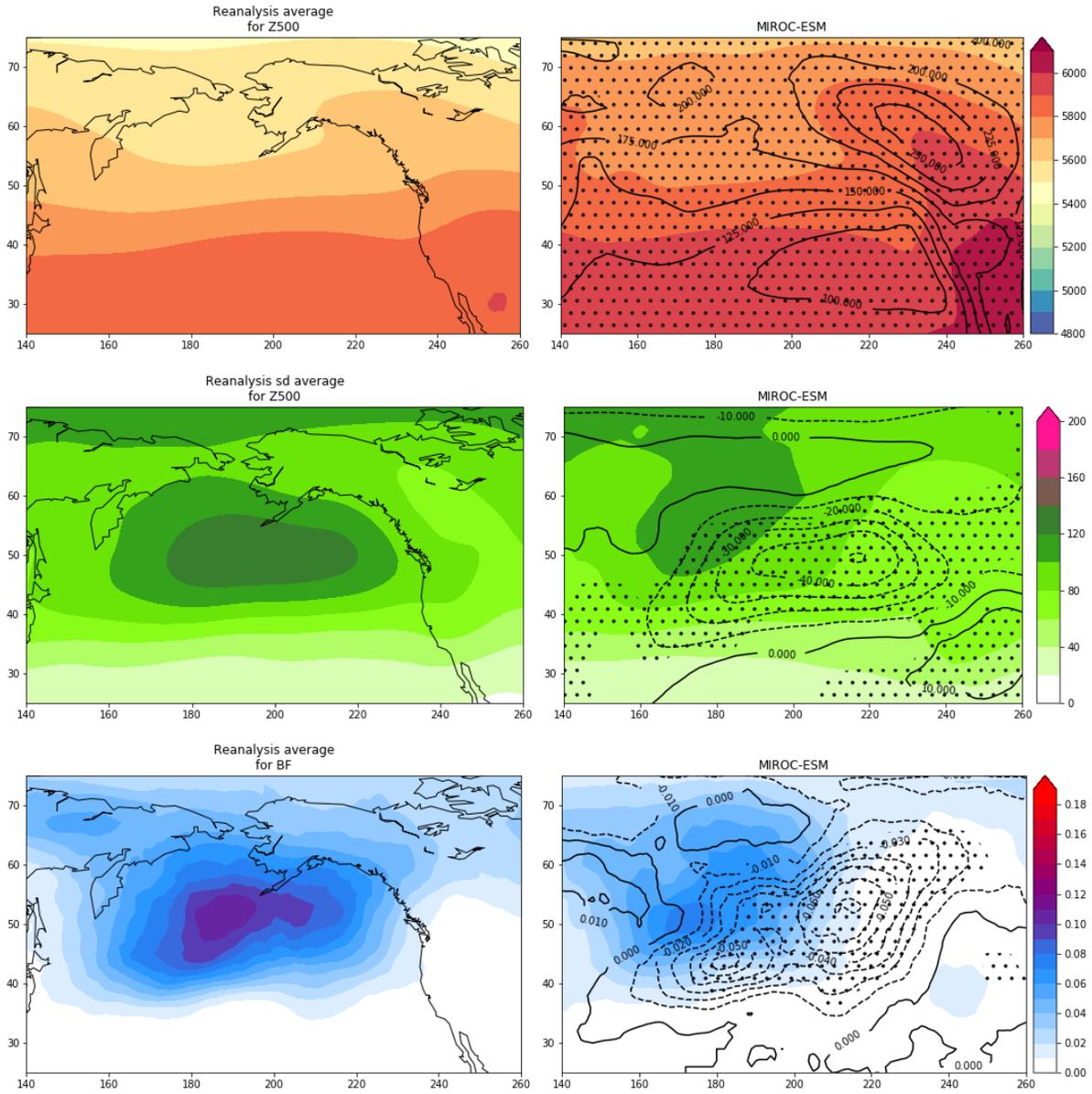
Figure 3.9: (Top) Reanalysis and MIROC-ESM NP JJA mean Z500 fields, (middle) corresponding Z500 variability fields, and (bottom) corresponding mean BF fields.

not sufficient to explain why the Hadley models have such outsized bias values compared to the other CMIP5 models. CanESM2 also has a positive residual value, but a slightly negative BF bias. However, in this case, the magnitude of the bias does not capture what is happening in the BF field pattern. Like MIROC-ESM in NP JJA, CanESM2 has a region of negative BF bias in the mid-Pacific (Figure 3.10; however, unlike MIROC-ESM, this is somewhat offset by a positive bias in the poleward half of the region, leading to a
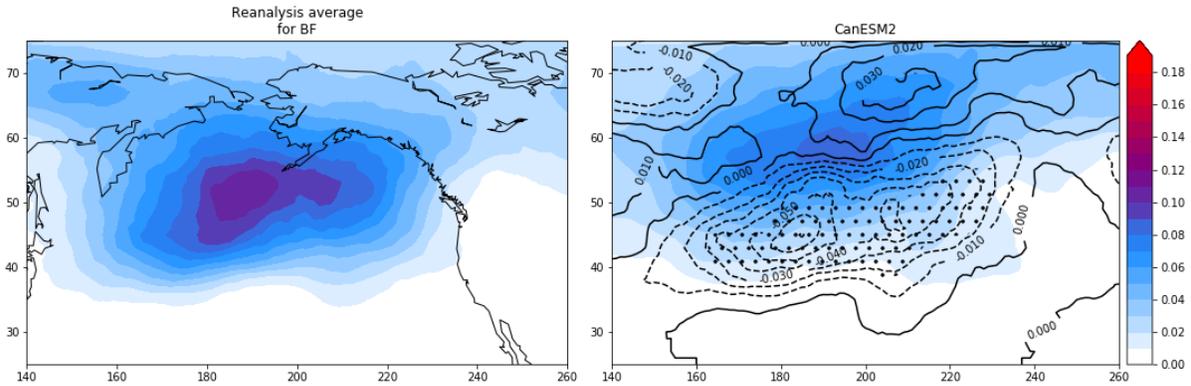
Figure 3.10: Reanalysis and CanESM2 NP JJA mean BF fields

net bias value that is slightly negative. This is indicative of a pattern shift rather than an overall magnitude bias; this model also has significantly positive LA and J850 bias values, which suggests that the overall flow pattern is biased in the poleward direction.

3.3 Conclusion

We have explored a number of potential explanations for blocking frequency bias in CMIP5 models, as well as possible connections between model characteristics and resultant block characteristics. The analysis was focused on the Pacific basins of the Northern and Southern Hemisphere for the respective winter and summer seasons. The reanalysis results show little to no difference between individual reanalyses in either the model characteristics or the suite of metrics used to assess various facets of blocking. The CMIP5 models, meanwhile, portray a wide range of both model characteristics and resultant blocking characteristics. In contrast to past studies on Northern Atlantic blocking, the CMIP5 models produce a higher median blocking frequency relative to the reanalysis mean climatology in the Northern Pacific winter and Southern Hemisphere winter and summer. Some of this could be due to the choice of detection algorithm; Chapter 2 demonstrated that the AGP algorithm commonly used in many blocking studies is sensitive to the block configuration and does not detect ridges with a N/S oriented axis. We use the anomaly-based Z^* algorithm instead, which could account for the observed positive BF bias in many models, since there is a negative Z500 bias in some of those same models.

There are moderate correlations between the model variable biases and the resultant

block characteristic biases, although none of the R^2 values for the linear regression equations exceeded 0.5; these relationships were also reflected in the pairwise correlation values between model and block characteristics. While a linear regression between BF and BA initially yielded $R^2 = 0.79$, this fit was heavily influenced by outlier Hadley models, and once these points were removed, the R^2 value dropped to 0.2. In contrast, both OLS 3.1 and 3.2, which utilize multiple input variables, explain over 74% of the variance in model blocking frequency bias, and the R^2 value does not greatly change when outlier points are dropped from the dataset. Biases in number of blocks, zonal distance, and block intensity are significant explanatory variables in both regression fits, with the wind field biases and centroid longitude serving as secondary explanatory variables. However, the relationship between BF bias and the possible explanatory variables is complicated, and in certain circumstances, the bootstrap averaging does not adequately capture the spatial pattern of the gridded field bias.

We will reserve further conclusions for the end of the next chapter. We now proceed to examine the changes to model results in the next generation CMIP6 models in Chapter 4.

Chapter 4

Assessment of model progress in between CMIP5 and CMIP6 models

This chapter is a continuation of Chapter 3; we continue to examine the connection between blocking frequency and model bias, as well as any connections between model characteristics and resultant block characteristics. Chapter 3 concluded that the number of blocks, the intensity of blocking events, and the distance traveled by blocks were significant variables in explaining biases in blocking frequency; there were also a number of modest correlations between model biases and corresponding block characteristic biases. Here, we follow a similar procedure to that outlined in Chapter 3, but with a new suite of climate models. The reanalysis average of ERA, JRA, MERRA2, and CFSR again serve as the reference from which to calculate relative model biases, and a subset of the CMIP5 models serve as the basis to compare whether or not there has been improvement in the next generation of models.

The CMIP6 generation of models (Eyring et al., 2016) builds upon past efforts to create a standardized model format which enable researchers to understand both historical and future scenarios under a number of real-world and idealized conditions. As with previous CMIP generations, the CMIP6 *historical* experiment is largely forced by observation-based datasets, although the overview paper notes that there have been efforts to better quantify aerosols in CMIP6 compared to CMIP5. Since both the CMIP5 and CMIP6 datasets are theoretically representing the same atmospheric conditions, this assessment will quantify

the extent to which a CMIP6 model has changed relative to the CMIP5 counterpart. Blocks are assessed as individual blocking events that are detected using the geopotential height anomaly-based objective detection algorithm outlined in Chapter 2 and previously utilized in Chapter 3. Analysis is limited to the Northern and Southern Pacific basins, for the time period 1979-2005, and blocking calculations are performed on the native grid. We also use the same list of block and model characteristic metrics that were used in the Chapter 3 assessment. We refer the reader to the Chapter 3 methods and appendices for reference to the various techniques and equations, as well as the abbreviations and definitions for the various metrics.

4.1 Data and Methodology

4.1.1 Data

We attempt to match as many model characteristics in between the CMIP5 and CMIP6 models as availability allowed. The CMIP6 models were obtained from the LLNL ESGF portal, which was also the source of the CMIP5 data; all but CNRM-CM6-1 are variant r1i1p1f1 (this was not available for CNRM-CM6-1; we chose to go with variant r1i1p1f2). Almost all of the model pairs have the same resolution between the CMIP5 and CMIP6 generations; only IPSL-CM6A-LR increased horizontal resolution to 1.3x2.5 degrees, from 1.9x3.75 in IPSL-CM5A-LR.

In Section 4.2.1, the variable pairs are compared against the reanalysis mean, as in Chapter 3; the CMIP6 model is further compared against the CMIP5 counterpart in order to assess where there has been significant improvement or deterioration of model accuracy with respect to the various model and block characteristics. Significant differences for both the CMIP5 and CMIP5 models are assessed relative to the reanalysis, and the mean improvement or deterioration of model accuracy is judge by comparing the absolute differences of CMIP5 and CMIP6 relative to the reanalysis.

Comparative model error change is calculated by finding the difference between the absolute bias values for CMIP5, $C5_e$, and CMIP6, $C6_e$:

$$M_{ec} = |C5_e| - |C6_e| \tag{4.1.1}$$

Institute	CMIP5	Resolution	CMIP6	Resolution
BCC	BCC-CSM1.1	T106 (1.12°)	BCC-CSM2-MR	T106 (1.12°)
CCCMA	CanESM2	T42 (2.79°	CanESM5	T42 (2.79°)
CNRM-CERFACS	CNRM-CM5	T85 (1.4°)	CNRM-CM6-1*	T85 (1.4°)
NOAA-GFDL	GFDL-CM3	2x2.5	GFDL-CM4	2x2.5
IPSL	IPSL-CM5A-LR	**1.9x3.75**	IPSL-CM6A-LR	**1.3x2.5**
MIROC	MIROC5	T85 (1.4°)	MIROC6	T85 (1.4°)
MRI	MRI-ESM1	T106 (1.2°)	MRI-ESM2	T106 (1.2°)
NCC	NorESM1-M	1.9x2.5	NorESM2-LM	1.9x2.5

Table 4.1: CMIP5 and CMIP6 counterparts used in this chapter, and horizontal resolution (lat/lon). CNRM-CM6-1 is the only CMIP6 model that is not the r1p1i1f1 variant.

If the CMIP6 bias is smaller in magnitude than the CMIP5 bias, M_{ec} will be positive and CMIP6 has improved over CMIP5; otherwise, the CMIP6 model has a stronger absolute bias value than before.

We will focus the majority of the analysis on the variables that exhibited significant influence on blocking frequency bias as calculated in Chapter 3, but will comment on any other significant changes that occur between CMIP5 and CMIP6.

4.2 Results

As the reanalysis baseline has already been addressed in Chapter 3, we begin with a discussion of the CMIP6 model and block biases. Section 4.2.1 presents the CMIP6 model biases and change relative to CMIP5 model bias, with additional figures for the model fields. Section 4.2.2 substitutes CMIP5 models in the linear regression fits for their CMIP6 counterparts, and the calculations from Section 3.2.3 are repeated. Finally, we compare the regression results from Chapter 3 to those from this chapter, and discuss the relative changes and comparative goodness of fit between the models.

Variable	Slope	R^2	Corr
J250	**1.16**	0.84	0.91
U_SD	0.78	0.76	0.87
V_SD	0.67	0.75	0.87
U	0.94	0.73	0.86
AI	0.75	0.69	0.83
BI	0.74	0.64	0.80
Z500_SD	0.68	0.63	0.79
LA	0.90	0.60	0.78
BF	0.85	0.58	0.76
BF_SD	0.87	0.56	0.75
J850	0.88	0.56	0.75
BA	0.65	0.49	0.70
LO	**1.06**	0.45	0.67
ZS	0.65	0.44	0.66
Z500	0.91	0.43	0.65
ZD	0.59	0.32	0.57
V	0.47	0.24	0.49
NB	0.54	0.20	0.45
DD	0.42	0.16	0.40

Table 4.2: Slope of fit line between CMIP5 and corresponding CMIP6 biases, R^2 value of the fit line, and Pearson correlation between bias values. The table is sorted by the R^2 values in descending order. A slope of 1 implies no change between CMIP5 and CMIP6 bias values. A slope of less than 1 implies that the CMIP6 bias values have decreased relative to their CMIP5 counterparts. A slope of more than 1 implies that CMIP6 biases have increased (these are bolded in the table). Lower R^2 values signify that there is more scatter in the bias values and that the change was not consistent across all of the model pairs.
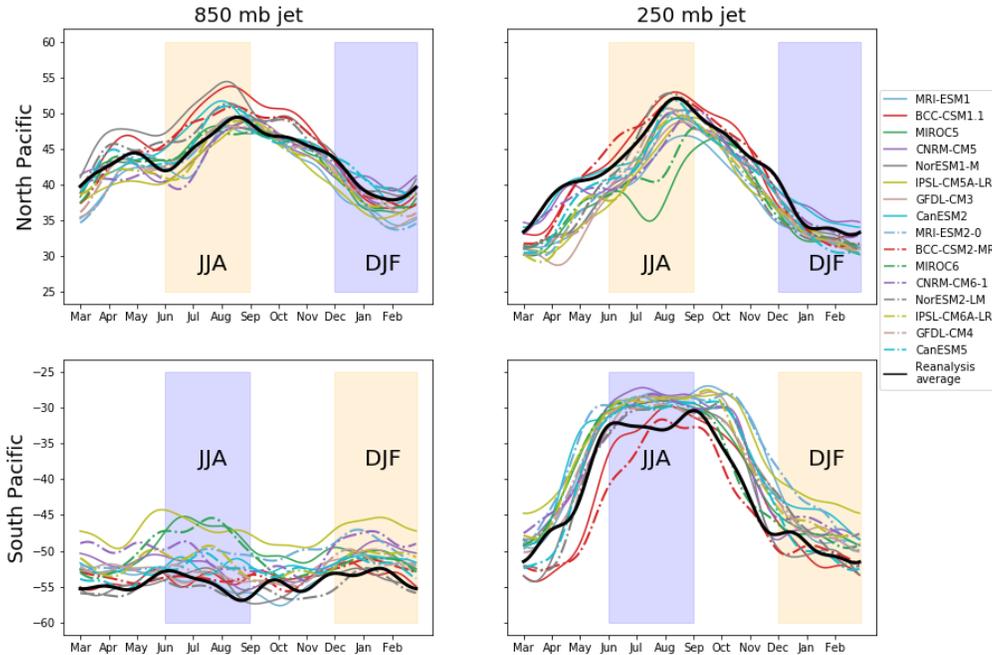
Figure 4.1: CMIP5 and CMIP6 jet positions, with the reference reanalysis line.

4.2.1 CMIP5 and CMIP6 comparison of metric biases

Table 4.2 shows the relationship between CMIP5 biases and their CMIP6 counterparts, and quantifies the mean direction of change in the bias values in between CMIP5 and CMIP6. A slope parameter of less than 1, such as that for the U bias (0.94, with an R^2 value of 0.73) implies that overall, bias in the U field has decreased in CMIP6 relative to the CMIP5 equivalent. However, some of the relationships outlined in this table have low R^2 and correlation values, which means that there is not a systematic relationship between changes in the bias values.

Two variables (J250 and LO) have slope values that are slightly greater than 1, implying that CMIP6 biases have actually increased, albeit very slightly. The right panels of Figure 4.1 show that while some particularly large outliers, such as MIROC5/MIROC6 in NH JJA, have been reduced in CMIP6, others, such as BCC-CSM1.1/BCC-CSM2-MR in SH JJA, show increased bias in CMIP6. Others, such as the wind variability biases,

exhibit a consistent improvement in model bias; Figure 4.2 shows one such example, where all CMIP6 models but CanESM5 display net improvements over CMIP5 for U variability in SP DJF. NB is an example of a variable where the improvement is more mixed. The slope of the NB fit line is 0.54, which suggests a trend in decreased CMIP6 bias compared to CMIP5. However, the R^2 value is only 0.2, meaning that there is a lot of variability in the relationship between the CMIP5 and CMIP6 data points. This is not necessarily a bad thing, as it might be due to a significant reduction in bias in a few outlier cases.

Some general observations are worth mentioning. Both the distribution of bias values (i.e. Tables 4.3 and 4.4) and the individual biases (Figure 4.3) should be taken into consideration; some of the variable quantiles values show that, on average, the CMIP6 models have reduced bias values compared to the CMIP5 models (although it should also be noted that the CMIP6 quantiles are only comprised of the values from 8 models as opposed to the 16 models in the CMIP5 quantiles). However, on a model-by-model basis, the picture is much more complicated, particularly when factoring in the spatial patterns of the gridded variables, as noted in Chapter 3.

Some of the trends observed in Chapter 3 continue to hold true with the CMIP6 models. The Pearson correlation values between bias pairs, as previously assessed in Chapter 3, have not significantly changed for the most part, which signifies that the per-model relationships between these variable biases also remains unchanged even when some of the data points are replaced. There are a few exceptions, which changes the variables that are dropped from the list of possible explanatory variables. As before, BA, U SD, and AI are dropped due to previously mentioned high correlation values with other input bias values. However, in this analysis, J250 is retained; its correlation with LO dropped from 0.43 to 0.28, and previously noted moderate correlations with U, U SD, and V SD dropped in value by 0.05. In its place, V SD is dropped from the list of input variables; it is highly correlated with Z SD (0.81) and U SD (0.91).

There is still a positive BF bias in all but NP JJA for the majority of the CMIP6 models (Figure 4.3, left subplot, top row), although the bias magnitude for many of those models has also decreased (Figure 4.3, right subplot, top row). The model variables also
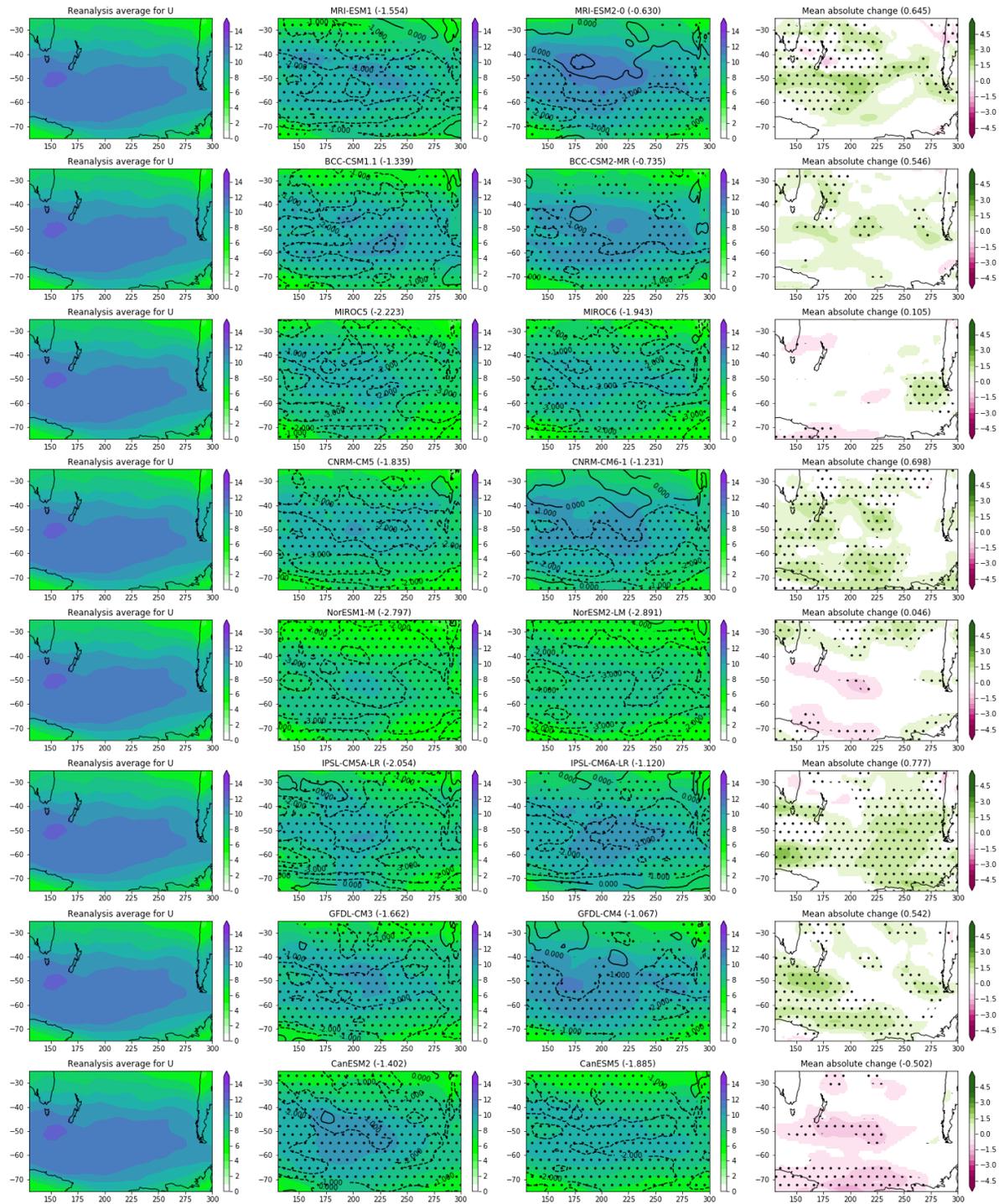
Figure 4.2: (Left column) Reanalysis average, (middle columns) CMIP5 and CMIP6 biases, and (right column) relative change in model bias magnitude in between CMIP5 and CMIP6 for U variability in SP DJF. In the bias change plot, green implies that the bias magnitude has decreased and pink implies that the model bias has increased. Significant differences are denoted by stippling.

94
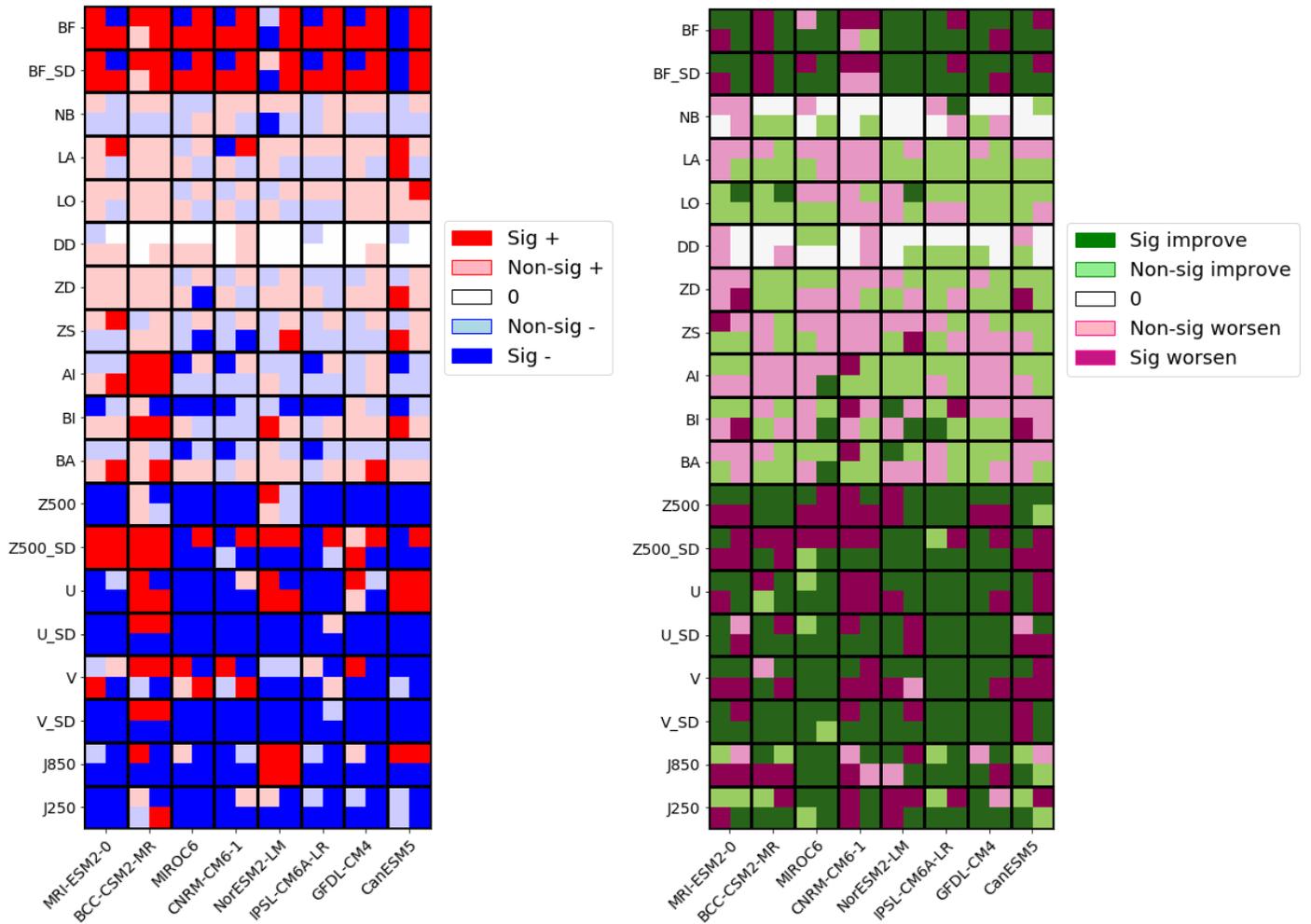
Figure 4.3: (left) CMIP6 differences with respect to the reanalysis average, where (light) dark blue signifies (non-)significant negative differences and (light) dark red signifies (non-)significant postive differences (right) CMIP6 improvement (green) or deterioration (pink) relative to CMIP5 bias

	JJA			DJF		
	Q25	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**
BF	**-0.922**	**-0.342**	0.033	**0.396**	**0.538**	**0.603**
BF SD	**-2.021**	**-1.002**	0.224	**0.679**	**0.974**	**1.190**
NB	-0.500	0.750	1.000	0.000	0.250	0.250
LA	-0.118	1.443	2.623	2.008	3.480	4.685
LO	-5.584	0.259	2.141	5.637	6.527	8.555
DD	-1.000	0.000	0.000	0.000	0.000	0.000
ZD	-116.211	-54.144	4.136	45.045	93.139	159.739
ZS	-0.798	-0.639	-0.485	0.338	1.108	1.363
BA	**-0.492**	-0.263	-0.020	-0.260	-0.159	-0.093
AI	**-1.486**	-0.899	-0.573	-0.356	0.864	1.814
BI	**-0.455**	-0.314	-0.046	**-0.576**	-0.447	-0.347
Z500	**-36.560**	**-15.904**	**-4.400**	**-40.476**	**-33.369**	-12.367
Z500 SD	**-5.570**	-1.411	**2.494**	**6.530**	**9.695**	**14.628**
U	**-1.302**	-0.014	**0.432**	-0.684	-0.283	-0.008
U SD	**-1.416**	**-1.265**	**-0.558**	**-0.906**	**-0.749**	-0.275
V	-0.026	0.059	**0.136**	**-0.284**	**-0.239**	-0.040
V SD	**-1.530**	**-1.248**	**-0.576**	**-1.267**	**-0.691**	-0.257
J850	-0.081	0.562	**2.595**	**-1.032**	**-0.707**	0.043
J250	**-4.101**	-2.059	0.028	**-1.518**	**-1.431**	**-1.344**

Table 4.3: NP CMIP6 model bias value quantiles. The values are listed for the 25th (Q25), 50th (Q50), and 75th (Q75) percentiles. Significant difference values are bolded.

continue to display negative biases in wind variability and jet position, although Table 4.2 and Figure 4.3 show that the many of the models have reduced biases relative to CMIP5. As these variables are not highly correlated with BF (and wind variability is excluded from the linear regression analysis), it is doubtful that this will result in any meaningful change in the linear regression fit.

	DJF			JJA		
	Q25	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**
BF	0.026	**0.203**	**0.390**	**0.220**	**0.610**	**0.730**
BF SD	-0.080	**0.652**	**1.102**	**0.420**	**1.257**	**1.569**
NB	-0.750	-0.500	-0.500	-0.750	-0.750	-0.500
LA	0.832	2.184	3.432	-3.818	-2.965	-2.184
LO	2.144	4.493	9.377	-3.270	-1.764	2.378
DD	0.000	0.000	0.250	0.000	1.000	1.000
ZD	45.567	90.390	238.607	-253.909	6.015	126.752
ZS	-0.652	-0.195	0.102	-1.070	-0.450	0.372
BA	0.018	0.126	0.342	0.137	**0.272**	**0.331**
AI	-0.532	-0.211	0.574	-0.061	-0.025	1.441
BI	0.025	0.118	**0.626**	-0.007	0.122	0.382
Z500	**-36.713**	**-30.388**	-14.469	**-54.245**	**-38.957**	-25.108
Z500 SD	**-11.039**	**-3.107**	1.591	**-5.761**	**-2.330**	1.120
U	**-0.985**	-0.109	**0.562**	**-1.621**	**-0.554**	**0.532**
U SD	**-1.899**	**-1.176**	**-0.984**	**-1.478**	**-1.333**	**-0.767**
V	-0.090	-0.027	0.005	**-0.438**	-0.239	**0.086**
V SD	**-2.351**	**-1.781**	**-1.468**	**-2.001**	**-1.719**	**-1.200**
J850	**-3.388**	**-1.861**	**-1.539**	**-3.554**	**-2.499**	**-1.418**
J250	**-3.779**	**-2.341**	-1.455	**-3.333**	**-2.680**	**-2.395**

Table 4.4: SP CMIP6 model bias value quantiles. The values are listed for the 25th (Q25), 50th (Q50), and 75th (Q75) percentiles. Significant difference values are bolded and marked with an asterisk.

4.2.2   Linear regression fit with CMIP6 variables

In order to compare the effect of the CMIP6 models on the previous regression fits, we substitute the CMIP6 models for the corresponding CMIP5 models but retain the rest of the CMIP5 and Hadley models in the regression calculation. The list of input variables remains unchanged with the exception of J250, which replaces V SD.

variable	R^2	R^2 (adj)	intercept	Z500	U	V	J850	J250	Z500 SD	U SD	V SD
AI	0.49	0.41	-	-0.4291	-	-0.4813	-	-	0.3619	-	-
LA	0.47	0.39	-	-	-	-	-	0.0482	-	-	-
BI	0.37	0.28	-	-0.5373	0.7696	-	-	-0.0669	-	-	-
DD	0.29	0.19	-	-	-	-	-0.0226	-	-	-	-
ZS	0.29	0.19	-	-	0.3774	-	-	-	-	-	-
BA	0.26	0.16	-	-	-	-0.6958	-	-	-	-	-
NB	0.26	0.15	0.5161	-	-	-	-	-	-0.6863	-	-
LO	0.25	0.14	-	-0.2199	-	-	-	-	-	-	-

Table 4.5: R^2, adjusted R^2, and coefficient values for linear regression fits of each block characteristic variable against the model characteristic variables. A dash denotes a non-significant coefficient value; BF and ZD did not have a significant regression fit and the values are not shown.

The single BF-explanatory variable regression calculations that were performed in the first part of Section 3.2.3 remain largely unchanged in terms of the regression fits to the scatter plots; therefore, we skip discussion of this question and advance to examining the change in the relationship between the block variables and model variables.

4.2.2.1 Linear regression of block characteristics against model characteristics

Table 4.5 again shows the results of a series of regression fits, with each block characteristics regressed against all of the model characteristics. Previously, BF and NB did not have significant fit equations; this time, BF and ZD do not display any significant relationships. As the model variables did not exhibit a strong relationship with the block characteristics in the first place, this is unsurprising.

The wind variability biases, which only had a significant relationship with LO in the previous version of this analysis, do not exhibit any significant relationships with any of the block variables this time. However, the Z500 variability bias now has a meaningful relationship with both AI and NB bias. The linear regression equation for AI in Chapter 3 previously had Z500, V, and J250 as explanatory variables; the coefficient values for Z500 and V remain largely unchanged, and the addition of the Z500 variability bias slightly drops the R^2 values from 0.5 to 0.49. However, J250 did not have a particularly large coefficient value in the previous analysis, and some of the largest biases in these

variable	coef	se	t	pval
NB	1.3568	0.363	3.739	0.000
BI	1.6000	0.536	2.986	0.004
U	-2.3202	0.951	-2.440	0.018
V	-1.3474	0.666	-2.025	0.048
DD	1.5375	0.844	1.821	0.075
LO	-1.8162	1.008	-1.801	0.078
ZD	1.9293	1.126	1.713	0.093
J850	0.0712	0.051	1.400	0.168
B0	0.2586	0.241	1.071	0.289
Z500	0.5684	0.825	0.689	0.494
Z500 SD	0.2064	0.396	0.521	0.605
LA	0.3986	1.089	0.366	0.716
ZS	-0.2829	1.162	-0.243	0.809
J250	0.0171	0.103	0.167	0.868

$R^2 = 0.71$, Adjusted $R^2 = 0.64$, AIC=168.6, BIC=198.9

Table 4.6: Linear regression of blocking frequency biases against model and block biases, but with CMIP6 models substituted for CMIP5 when available. Variables are sorted from smallest to largest p-value, with the four significant variables above the double line.

values (like MIROC5 in NP JJA) were corrected in the CMIP6 equivalent; the previously established relationship in Chapter 3 no longer holds now that these biases have been corrected, which is likely also why J250 is now included in the list of input variables for the BF linear regression analysis.

4.2.2.2 Linear regression of BF biases against block and model biases

Linear regression with all variables (OLS 4.1) The first model fit (Table 4.6) has an R^2 value of 0.71, which is lower than the previous R^2 values of 0.74 for the regression

model with all variables. The significant variables from the first model fit are still relevant here, although ZD and LO are slightly below the $p < 0.05$ significance level. V is also included in OLS 4.1; while the standard error of the coefficient is the same as before, the value of the coefficient itself is high enough to yield a significant t statistic.

It should be noted that the Hadley models, which were the most extreme BF outliers in Chapter 3, have not been replaced by CMIP6 equivalent data in this analysis; neither has MIROC-ESM, which was one of the outliers that was flagged in OLS 3.1. These data points are still significant outliers in the new model fit, although MIROC-ESM does not have a sufficiently high Cook's threshold value to designate it as a significant outlier; this new regression fit is able to predict a BF bias value that somewhat matches the resulting MIROC-ESM BF bias, even though the some of the inputs for this model are anomalously large. Meanwhile, the fourth outlier model in OLS 3.1, MIROC5, was replaced with MIROC6. While the changes in the gridded model field biases was not particularly large in MIROC6 (Table 4.7; also see Figure D.34 and Figure D.46 in the supplemental figures), the magnitudes of the NB and BI biases did significantly decrease.

Model	Lev	Resid	CD	BF	NB (+)	BI (+)	ZD (+)	U (-)	V (-)
MIROC5 (SP JJA)	0.391	**-2.278**	**0.219**	0.41	0.38	-0.2	-0.16	-0.87	0.31
MIROC6 (SP JJA)	0.365	-1.701	**0.114**	0.37	0.08	-0.07	-0.23	-0.78	0.30

Table 4.7: MIROC5 and MIROC6 (NP JJA) leverage, residual, and Cook's distance values from OLS 3.1 and 4.1, as well as normalized bias values for the relevant variables.

As MIROC-ESM NP JJA does not significantly impact the model fit, we retain it in the input data (but not the two SP JJA Hadley models) and proceed to OLS 4.2.

Linear regression with SP JJA Hadley models removed (OLS 4.2) Once the two outlier Hadley models are removed, the R^2 value increases slightly but the AIC and BIC values greatly improve, which is similar to what happened in Chapter 3 when the outliers were removed in OLS 3.2. The coefficients for U and BI greatly decrease, and their p-values drop below significance as a result; while the coefficient of NB also drops, its standard error also decreased, and so it remains significant. ZD, which had dropped below significance in OLS 4.1, is significant once more, and DD also joins the list. Interestingly,

variable	coef	se	t	pval
NB	0.8783	0.155	5.653	0.000
ZD	1.6818	0.478	3.522	0.001
DD	0.9571	0.356	2.688	0.010
V	-0.5914	0.286	-2.065	0.044
U	-0.6274	0.420	-1.495	0.141
BI	0.3314	0.240	1.380	0.174
ZS	-0.5830	0.488	-1.194	0.239
B0	0.0712	0.102	0.696	0.490
Z500	0.2405	0.349	0.690	0.493
LA	-0.2356	0.459	-0.514	0.610
Z500 SD	0.0175	0.167	0.105	0.917
J250	-0.0041	0.043	-0.095	0.925
LO	-0.0262	0.442	-0.059	0.953
J850	0.0006	0.022	0.025	0.980

$R^2 = 0.73$, Adjusted $R^2 = 0.66$, AIC=55.85, BIC=85.63

Table 4.8: Linear regression of blocking frequency biases against model and block biases with SP JJA Hadley model removed. Variables are sorted from smallest to largest p-value, with the four significant variables above the double line.
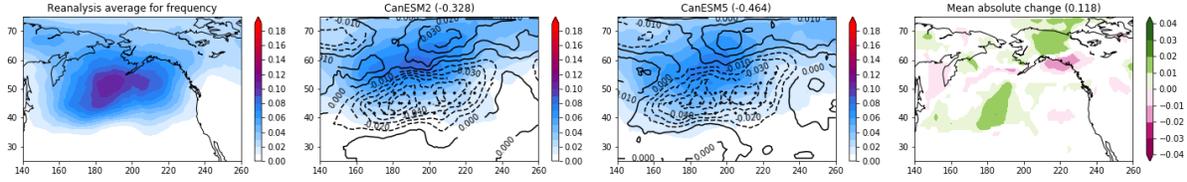
Figure 4.4: (Left) Reanalysis mean BF field, (left center), CanESM2, (right center) CanESM5, and (right) the relative change in between CanESM2 and CanESM5 for NP JJA.

removing MIROC-ESM NP JJA as well (not shown) does not change which variables are significant explanatory variables, nor does it greatly change their coefficient or standard error values. Removing it does impact the variables for which it exhibits particularly large biases, such as Z500. But it is no longer an influential outlier as in OLS 3.1; therefore, the model fit is not overall greatly impacted.

OLS 4.2 has 3 influential outlier data points: HadGEM2-CC (SP DJF), which was also an outlier in OLS 3.2; HadCM3 SP DJF; and CanESM5. Previously, in OLS 4.1 we observed that the improved bias values in MIROC6 as opposed to MIROC5 meant that MIROC6 was not a significant outlier. CanESM2 was a significant outlier in OLS 3.2, and the fact that CanESM5 has again been identified as an outlier point suggests that existing biases in CanESM2 have not been reduced in CanESM5. Figure 3.10 shows that the blocking patterns have not significantly changed in between CanESM2 and CanESM5. Additionally, a comparison of relevant bias values for CanESM2 vs CanESM5 (see Table 4.9) shows that the bias magnitudes for BI, V, and DD actually increased relative to CanESM2, while those for BF, NB, ZD, and U were similar.

Model	Lev	Resid	CD	BF	NB (+)	BI (+)	ZD (+)	U (-)	V (-)	DD (+)
CanESM2 (NP JJA)	0.235	**2.136**	**0.093**	-0.08	-0.07	-0.38	-0.2	0.03	-0.04	-0.12
CanESM5 (NP JJA)	0.281	**2.164**	**0.121**	-0.11	-0.07	-0.55	-0.19	0.02	-0.10	-0.25

Table 4.9: CanESM2 and CanESM5 (NP JJA) leverage, residual, and Cook's distance values from OLS 3.2 and 4.2, as well as normalized bias values for the relevant variables.

Discussion   Table 4.10 summarizes the model results from the Chapter 3 and 4 regression models. Exchanging the CMIP6 model for the CMIP5 models actually served to slightly decrease the model quality, although dropping outliers let to slight increases in

Regression	R^2	AIC	BIC	Significant Variables
OLS 3.2	0.76	50.26	79.58	NB, BI, ZD, (U, V)
OLS 4.2	0.73	55.85	85.63	NB, ZD, DD, V
OLS 3.1	0.74	160.9	191.2	NB, BI, ZD, U, LO
OLS 4.1	0.71	168.6	198.9	NB, BI, U, V (DD, LO, ZD)

Table 4.10: Comparison of all model fits from Chapters 3 and 4, sorted with the best (low AIC and BIC) on top. Variable biases with $0.05 < p < 0.1$ are listed in parentheses.

Type of Outlier	Model	OLS	Observations
High Leverage (Inputs)	MIROC-ESM (NP JJA)	3.1, 4.1 *	Strong + Z500 bias
			Strong - BI bias
	CMCC-CESM (SP JJA) *	3.2	low pattern correlation in V
High Residual (Prediction)	Hadley models (SP JJA)	3.1, 4.1	Systematic positive BF bias
			Strong + NB bias
			Strong + BI bias
	MIROC5 (SP JJA)	3.1	Strong - U bias
	HadGEM2-CC (SP DJF)	3.2, 4.2	Systematic positive BF bias
			Strong + NB bias
	HadCM3 (SP DJF)	4.2	Systematic positive BF bias
			Strong + NB bias
	CanESM2 (NP JJA)	3.2	Poleward BF concentration
	CanESM5 (NP JJA)	4.2	Poleward BF concentration

Table 4.11: Outlier data points from Chapters 3 and 4, in terms of high leverage and high residual values. Note that CMCC-CESM was a significant outlier, but not an influential outlier in OLS 3.2; the same applies for MIROC-ESM in OLS 4.2 but not OLS 3.2.

the R^2 value and much smaller AIC and BIC values. If there were CMIP6 counterparts for all of the CMIP5 models (particularly the outlier Hadley models) then there would likely be a better fit in the Chapter 4 models, as all of the models in the regression would be initialized with the same set of parameters and exhibit more consistent bias trends.

Across all models, number of blocks and the relative block intensity are consistently associated with biases in blocking frequency. Biases in wind, zonal distance, and centroid

longitude, which are associated with block position and the mean flow, are also significant explanatory variables for a majority of the models. The regression fits were highly impacted by outlier data points from the Hadley models, as well as a few others; nonetheless, all of the model fits outlined here have R^2 values that exceed 0.7, which means that a majority of the BF bias have a significant association with the variable biases noted above. The outlier Hadley models were CMIP5 models with no CMIP6 counterpart, which meant that they remained in the dataset throughout both Chapters 3 and 4 and as a result skewed the fits of each. MIROC6 did not exhibit significant improvements in the model fields, but improvements in some of the resultant block characteristics, which meant that it was not flagged as a significant outlier in OLS 4.1. CanESM2 was an outlier in OLS 3.2, but CanESM5 had similar or larger bias magnitudes for the relevant variables in OLS 4.2, and thus remained a significant outlier.

4.3   Conclusion

We have examined the change in the new generation of CMIP6 blocking with respect to previously calculated CMIP5 blocking which was examined in Chapter 3. Some improvements were seen in the model characteristic bias values, although these improvements were not reflected either in the resultant block characteristics or the new regression fits. When CMIP6 models were substituted for their CMIP5 counterparts into linear regression fit calculations, the model quality did not significantly deteriorate relative to the previous fit, but using the CMIP5 observations yields a slightly better regression fit. Overall, all of the models had R^2 values that exceed 0.7; these models explain a large proportion of BF bias and a few variables remain consistently significant throughout most or all of the models. NB was a consistently significant variable, which is unsurprising. ZD and BI were significant in 3 out of 4 models (and ZD was slightly below significance in 4.1). U was significant in 3.1 and 4.1 and slightly below significance in 3.2. V was significant in 4.1 and 4.2 and slightly below significance in 3.2. OLS 4.2 was the only model to include DD as one of the significant components, although it was also slightly below significance in 4.1. As the per-model DD bias is comparatively low, with no significant biases in

either CMIP5 or CMIP6, is is not entirely clear why this variable emerged as a potential explanatory variable in the Chapter 4 models.

The outlier models, which are summarized in Table 4.11, show that the majority of error arises from models in the CMIP5 dataset, even when CMIP6 models are substituted into the new regression calculations. Unfortunately, not all CMIP5 models had an available CMIP6 counterpart, which means that most of the models that were the problem in the first place were not substituted with a more recent model version. The Hadley models, which produce extremely high blocking frequencies that are not adequately captured by the existing explanatory variables, are the most extreme example, although not all of the Hadley models had bias values that were extreme enough to be flagged as outliers. MIROC5 was the one case in which its CMIP6 counterpart corrected explanatory variable biases, resulting in a predicted BF bias that better matched the observed BF bias. Meanwhile, CanESM5 did not exhibit the same level of improvement over CanESM2.

Moving forward, there are other variables that would be worthwhile including in a similar analysis. For example, SST is an important consideration in the Pacific basins, as well as some of the interannual climate modes such as PNA. We have indices such as mean centroid latitude and longitude as an approximation of block position; however, a block trajectory metric, such as those used in tropical cyclone studies, would be useful for assessing the mean path of blocking events throughout the study basins. Additionally, characterizing the energetics leading up to and during the lifespan of blocks in the various models will be useful for better assessing the biases in block characteristics such as block intensity. Finally, the wave breaking index variable presented in Davini et al. (2012) would be useful in determining whether the blocking was associated with cyclonic or anticylonic wave breaking.

Chapter 5

Summary, Conclusions, and Future Directions

With improvements in both model approximation of atmospheric processes and the computational power required to run these models, the research community continues to advance towards more accurate simulations of weather and climate. This has important societal and scientific implications, as models are essentially a sandbox in which the researcher can explore a wide range of scenarios, from aquaplanet simulations to the effect of increased $C0_2$ levels in the atmosphere. However, there is still much work to be done with the existing tools available.

We have focused on atmospheric blocking, an extreme weather phenomenon that is still an area of active research in terms of the governing physics. The mere definition of a block, as described in terms that can be put into a computer program and used to process a dataset, is still nebulous enough to warrant further study. We have demonstrated that the definition of blocking in the context of different algorithms will yield very different results in terms of both the block location and the block characteristics, even when the same feature is detected in the same dataset by three different algorithms. We then explored a different avenue in block representation by attempting to connect blocking frequency bias to model biases in a number of explanatory variables, both in terms of the inherent model characteristics and the characteristics of the blocks that were produced. While the regression calculations focused on blocking frequency, a number of connections

were drawn between other combinations of variables as well. Throughout the course of this work, a number of new metrics and statistical methods were created to assess results, including

- Modified versions of existing anomaly-based objective detection algorithms in which the constant threshold is replaced with a spatiotemporally varying threshold

- A probability of co-occurrence metric that quantifies how likely a feature will be detected by an algorithm, given another algorithm's results

- A spatial similarity metric that quantifies agreement between algorithms when identifying the same blocking event

- A non-parametric, generalized method of quantifying the difference between two distributions of values (particularly relevant when working with extreme value data)

- A non-parametric way to quantify a signed median field bias, using blocking bootstrap sampling to construct a distribution of difference values between two gridded fields and the permutation test to determine the significance of the median difference value

- A methodology for normalizing biases by the distribution of the reference data in order to compare values across different regions, seasons, and scales while also retaining the sign of the bias

Finally, we explored the effect of including some of the most recent CMIP6 models in place of their CMIP5 counterparts and examining the change in the OLS fit. Despite containing significant outliers which impacted the overall regression model fits, all of the models explained a significant proportion of the variance.

Some limitations remain that would be worth addressing in future work. First, a more sophisticated definition of a block will be required to accurately capture the desired features while filtering out others that are similar in some ways (i.e. persistent anomalous geopotential height) but different in others (i.e. subtropical highs). Accomplishing this

task requires a deeper examination of blocking characteristics as they relate to the relative success of an individual algorithm; the RRR case study highlights a clear limitation of the AGP algorithm with regards to its requirement of a block axis tilt, and there are steps that could be taken to limit the results of the Z^* algorithm.

Second, calculating bias values in weather data, with its tendency toward skewed distributions and spatiotemporal dependence, is no trivial task, and in reducing bias down to a single number, potentially relevant information is lost. There are some cases in which the bias was a result of a pattern shift rather than a difference in magnitude, which suggests that when it comes to spatial data, two quantities are needed to assess bias: one for the magnitude component, and one for the spatial component. BF was the independent variable of choice in our linear regression calculations, as it is the most commonly used method for assessing model accuracy in blocking; however, there were some notable outlier models that could not be explained by the multivariate relationship that was established in the linear regression equation.

Finally, the comparison of the results between CMIP5 and CMIP6 highlights an opportunity to assess model improvement between generations, not just in terms of blocking frequency but also in terms of individual block characteristics. Future blocking research should make use of an expanded set of metrics beyond blocking frequency in model comparisons and assessments of model improvement. These additional metrics can aid researchers in selecting a subset of models that suits their specific research need, such as isolating the cause of a particular blocking characteristic bias as it relates to the models and what can be done to reduce the bias. We hope that this work will inspire new questions while also addressing the research questions set forth at the beginning of this dissertation, and inform future work with regards to both blocking and the underlying methodologies to assess this phenomenon.

Appendices

Appendix A

Chapter 2 Appendices

A.1 Vertically averaged potential vorticity

In its simplest form, Ertel's potential vorticity (EPV) is written as

$$EPV = \frac{\zeta_a \cdot \nabla \theta}{\rho} \tag{A.1.1}$$

Assuming hydrostatic balance in order to eliminate density, (A.1.1) can be written as

$$EPV = g \left(\frac{1}{r \cos \phi} \frac{\partial v}{\partial p} \frac{\partial \theta}{\partial \lambda} - \frac{1}{r} \frac{\partial u}{\partial p} \frac{\partial \theta}{\partial \phi} - (\zeta + f) \frac{\partial \theta}{\partial p} \right) \tag{A.1.2}$$

in which g is the gravitational constant, r is the Earth's radius, u and v are the zonal and meridional wind components respectively, θ is potential temperature, and $\zeta + f$ is absolute vorticity.

The form presented in Equation A.1.2 is in spherical coordinates (λ, ϕ) in the horizontal direction and pressure level coordinates (p) in the vertical.

For S04, the EPV is vertically averaged over the 150-500 hPa layer using the integral method,

$$VPV = \frac{1}{(p_1 - p_2)} \int_{p_1}^{p_2} EPV(p) dp \tag{A.1.3}$$

where $p_1 = 1.5 \times 10^4$ Pa and $p_2 = 5 \times 10^4$ Pa, respectively.

A.2   Blocking Events and the StitchBlobs Software

All of code used to produce these results are included in the TempestExtremes software package (Ullrich and Zarzycki, 2017). The algorithm for the AGP method is contained to a single binary, `BlockingGHG`, while the calculations for PV^* and Z^* involve a multi-step process which is outlined in Section 2.1.2.4 and Figure 2.1.

The `StitchBlobs` binary takes the outputs from the desired algorithm and applies the spatiotemporal constraints, and the corresponding `BlobStats` binary produced summary text files which include per-feature information about:

- Minimum/maximum latitude and longitude coordinates

- coordinates of block centroid

- block area (in terms of fractional area)

This information is mainly used in Section 2.2.3, and the quantities are defined as follows:

- *Block duration:* the number of time steps for which block is present, multiplied by the time resolution.

- *Block zonal distance:* The difference between the start and end longitudes, converted to km.

- *Block zonal speed:* Distance divided by duration.

- *Block size:* Fractional area multiplied by the Earth's surface area, $SA = 4\pi(6371)^2$ km^2

A previous attempt to average speeds calculated over each 6 hour period was discarded because the center coordinate would occasionally "jump" in between subsequent time steps due to a change in the detected block's size—usually caused by merging of distinct features— and a subsequent shift in the centroid. This "jump" led to large zonal distance values and, consequently, artificially high block speed values. The method used here

(distance from start to finish, divided by duration) has its own drawbacks, as two methods tracking the same block might yield different block speed measurements if the duration and distance measurements differ between the two methods.

A.3 Probability of co-occurrence calculation

Co-occurrence is defined as the overlap between blocked regions detected by two or more different methods, which corresponds to a region of $C_k = 2$ in similarity calculations. We quantified co-occurrence by counting the time steps in which there was overlap ($n[M1 \cap M2]$) during the similarity calculations as well as the number of blocks for each of the methods ($n[M1]$, $n[M2]$, etc), then finding $\min(n[M1], n[M2], n[M1 \cap M2])$ per time step.

The probability of co-occurrence is defined as

$$P(M1|M2) = \frac{P(M1 \cap M2)}{P(M2)} \tag{A.3.1}$$

which translates to the number of times that blocks detected by methods M1 and M2 overlap over the number of times that blocks detected by method M2 occur.

A.4 Spatial similarity calculation

This method is similar in concept to the Jaccard index (Jaccard, 1908), also known as intersection over union:

$$J(A, B) = \frac{A \cap B}{A \cup B} = \frac{A \cap B}{A + B - A \cap B} \tag{A.4.1}$$

where $A \cap B$ are the common points between A and B and $A \cup B$ are the sum total of points in A and B, minus the number of common points. In this instance, spatial similarity calculations are performed on two fields $M1$ and $M2$, consisting of regularly spaced grid points whose values are either 1 (block) or 0 (not a block) according to the corresponding objective detection method.

However, there are two points to consider in these calculations:

1. We only wish to count common presence, not common absence (in order to quantify the amount of agreement when two methods detect a block).

2. A simple count of commonly blocked gridpoints between two methods will overemphasize spatial agreement at higher latitudes, since the meridians converge at the poles and the actual distance between gridpoints is shorter.

To address point 1, the two fields are summed together ($F = M1 + M2$), where gridpoints have values of 0 (no detection), 1 (one method detects blocking), or 2 (both methods detect blocking); only nonzero gridpoints in F are considered.

For each cluster C_k of contigous nonzero gridpoints on F

$$S(C_k) = \frac{\sum_{n=1}^{P_k} (v_n - 1)\cos(\phi)_n}{\sum_{n=1}^{P_k} cos(\phi)_n} \tag{A.4.2}$$

where P_k is the number of points in C_k, v_n is the value of the gridpoint (either 1 or 2), and $cos(\phi)_n$ is the cosine of that gridpoint's latitude value.

Summing the cosine latitude values of gridpoints, rather than the number of points in C_k, addresses point 2, as $\cos(\phi)_n$ is smaller at higher latitudes (thus approximating the smaller area between gridpoints). The numerator is essentially the intersection of blocked points ($M1 \cap M2$); the sum will only include $\cos(\phi)_n$ where $v_n = 2$ (since $v_n = 1$ means that $\cos(\phi)_n$ is multiplied by 0). The denominator is the union ($M1 \cup M2$), where all $\cos(\phi)_n$ are included in the sum.

A.5   Significance testing

The quantities in Section 2.2 are considered to be statistically significant if $p < 0.05$ (and highly significant if $p < 0.01$), which implies that the quantity does not satisfy the null hypothesis. The following sections explain the methods used to establish significance.

A.5.1   Permutation test

The permutation test is a nonparametric method of hypothesis testing similar to bootstrapping (but without replacement). A sampling distribution is constructed by resampling and/or shuffling data, thus avoiding assumption of a known sampling distribution (for example, a normal distribution).

The general methodology is as follows:

1. Calculate the reference statistic R (correlation, difference of medians, etc.) between the two datasets

2. Construct the sample distribution D by repeating the following 10000 times:

    (a) Generate samples $S1$ and $S2$ (specifics in subsequent subsections)

    (b) Repeat the calculation of R using $S1$ and $S2$ to generate sample statistic $\widetilde{R}$

3. Count the number of times, n_D, in D where $\widetilde{R}$ is more extreme than R:

$$
n_D = \begin{cases} \text{number of } \widetilde{R} \geq R, & \text{if } R > D_{50} \\ \text{number of } \widetilde{R} \leq R & \text{otherwise} \end{cases} \tag{A.5.1}
$$

where D_{50} representing the 50th percentile of D and n_D representing the number of instances in D which satisfy the appropriate case.

The p-value is calculated as

$$
p = \frac{n_D}{10000} \tag{A.5.2}
$$

A.5.1.1   Blocking duration, zonal distance traveled, size, speed (Section 2.2.3)

Null hypothesis: Vectors of values $V1$ and $V2$ are drawn from the same distribution. Therefore, if $V1$ and $V2$ are pooled and two new samples $S1$ and $S2$ are drawn from that common pool, the difference in medians, $M(S1, S2)$, of the samples' distributions should be similar to $M(V1, V2)$.

Alternate hypothesis: The distributions of $V1$ and $V2$ are different enough that $M(V1, V2)$ will differ significantly from $M(S1, S2)$

Sampling method:

1. Combine all values from $V1$ and $V2$ into common pool P

2. Draw new samples $S1$ (same length as $V1$) and $S2$ (same length as $V2$) from P without replacement and calculate $M(S1, S2)$

A.5.1.2 Pattern correlation (Section 2.2.4)

Null hypothesis: For fields $F1$ and $F2$, correlation $C(F1, F2)$ will not change significantly if one of the grids is shuffled (implying that the correlation value would be similar if the pattern was entirely random).

Alternate hypothesis: $C(F1, F2)$ is a result of this particular arrangement of frequency values in each of the two grids being compared.

Sampling method:

1. Generate a list of 1000 random array indices i

2. Create two array subsamples, $S1 = F1[i]$ and $S2 = F2[i]$

3. Shuffle the order of $S2$ to create $\widehat{S2}$ and calculate $C(S1, \widehat{S2})$

A.5.2 Probability of co-occurrence (Section 2.2.4)

In order to verify that the probabilities would not significantly change with the addition of more data, we took samples of the entire dataset for a particular region and season, with the samples consisting of the equivalent of $n = \{1, 2, ..., length(data)\}$ years that were randomly selected from the set of possible years in the ERA-Interim data; probability of co-occurrence was re-calculated with those samples. This procedure was repeated 1000 times and stored as pairs of years and probability (n, P).

We calculated a linear regression between n and each of the probability values ($lm(n \sim P(M1|M2))$ and $lm(n \sim P(M2|M1))$) using the linear model (`lm`) function in R. Since the slope of the line was approximately zero, the estimated equation was effectively the coefficient, which equaled the initially calculated probability value. The `lm` function returns a p-value for each coefficient, therefore giving us a p-value for each calculated probability.

A.5.3 Spatial Similarity (Section 2.2.4)

This methodology was similar to the permutation test, although the end calculation for the p-value was slightly different.

Null hypothesis: The algorithms detect features equivalently. Therefore, for fields $F1$ and $F2$, $S(F1, F2)$ will not change significantly if the values of gridpoints on $F1$ and $F2$ are pooled and randomly reassigned.

Alternate hypothesis: $S(F1, F2)$ is dependent upon the results of the two algorithms being compared.

For each pair of overlapping blocks, for the set of all nonzero gridpoint coordinates G in both $F1$ and $F2$, the following is repeated 100 times:

1. Determine the number of blocked grid points in $F1$ and $F2$, $nF1$ and $nF2$.

2. Set all $F1[G]$ and $F2[G]$ to 0.

3. Randomly reassign $nF1$ grid points in $F1[G]$ a value of 1 to create $S1$; randomly reassign $nF2$ grid points in $F2[G]$ a value of 1 to create $S2$.

4. Calculate $S(S1, S2)$ and store the new value.

The result is two distributions; the distribution of calculated similarity values, S, and the distribution of the sampled similarity values, $\widetilde{S}$. The difference between the two distributions is established using the Wilcoxon test (`wilcox.test` in R), which returns the p-value.

Appendix B

Chapter 3 Appendices

B.1   Blocking intensity calculation

Block intensity (BI) was first proposed in Lupo and Smith 1995:

$$BI = 100 \left(\frac{MZ}{RC} - 1 \right) \tag{B.1.1}$$

where MZ is the maximum Z500 value and RC is the value of the Z500 contour that outlines the wave containing the blocking event. It measures the block's geopotential height gradient with respect to a subjectively determined height contour that represents the block. This measure provides a way of comparing block intensity across different regions and season, as the maximum geopotential height value of the blocked region is normalized to values on a common scale of 1-10, with blocks classified as weak ($BI < 2$), moderate ($2 < BI < 4.3$), and strong ($BI > 4.3$).

This equation was modified in Wiedenmann et al. (2002) to accommodate longitudinally-varying detection methods, where RC was calculated as the average between MZ and selected height values that are upstream (Z_u) and downstream (Z_d) from the coordinate of MZ:

$$RC = \frac{(Z_u + MZ) \times 0.5 + (Z_d + MZ) \times 0.5}{2} \tag{B.1.2}$$

We further modify Equation B.1.2 to accommodate the results from our blocking detection algorithm, which produces 2D blocking results as opposed to the 1D results

from (LS95) and (W02). MZ is the average of the highest 5% and Z_{min} is the average of the lowest 5% of Z500 values in the detected region. Thus,

$$RC = (MZ + Z_{min}) \times 0.5 \tag{B.1.3}$$

Thus, plugging MZ and the RC value from Equation B.1.3 into B.1.1, we obtain a blocking intensity index.

B.2 Anomaly intensity calculation

This metric was inspired by the total power dissipation measure used in hurricane studies, which accounts for the hurricane size in the quantification of hurricane energy. Here, we account for detected block size when assessing the height anomaly magnitude of a particular block.

The anomaly index, AI, is defined as

$$AI = \sum_{i,j=1,1}^{m,n} Z_{ij}^* A_{ij} \tag{B.2.1}$$

where A_{ij} is the area of a single grid box, in km^2, and Z^* is the average of the four anomaly values at the corners of the grid boxes. For ease of comparison, the values in Chapter 3 have been divided by 10^8. The units are effectively m^3, but as this does not make sense in the context of the variable, we omit the units from the discussion.

B.3 Difference charts methodology

The difference plot for the block characteristics (Figure 3.7) displays a set of variables with a wide range of units and values on a single normalized scale. The calculations for the metrics and fields differences are similar in concept but contain some slight differences in methodology. The calculation steps are outlined in the following sections.

B.3.1 Field differences

The differences in two spatial fields is difficult to collapse down into a single number, particularly when there are competing positive and negative differences that can cancel each other out. Root mean square error is a useful metric in that it quantifies the average

absolute deviation from the expected value, but does not work for our purposes because it does not include the overall sign of the difference between the fields.

We utilize a block bootstrap resampling method to generate distributions for model and reanalysis fields. This method involves generating a 10 (latitude) x20 (longitude) degree boundary box (this size captures significant features on the averaged maps fairly well while avoiding averaging out significant features) and taking subregion samples from both the model and reanalysis time-averaged fields.

The following steps are repeated over 5000 iterations (sampling with replacement) to generate the distributions:

1. Generate a sample coordinate pair, (ϕ_s, λ_s) from the array of possible latitude values $[\phi_{min}, (\phi_{max} - 10)]$ and longitude values $[\lambda_{min}, (\lambda_{max} - 20)]$

2. Take subsample boxes R_s and M_s with boundaries $[\phi_s, (\phi_s + 10)]$, $[\lambda_s, (\lambda_s + 20)]$ from the time-averaged fields $\overline{R}$ and $\overline{M}$

3. Find the box averages R_b and M_b

4. append R_b and M_b to corresponding lists of average values R_L and M_L

5. Find the difference between the two lists, $D_L = M_L - R_L$. The median bias value of the field is defined as the median of D_L

This method is necessary to maintain some semblance of spatial information in the median difference; the difference between the two lists is determined on a per-index basis, so the most common difference value will theoretically be the median value (note: this method becomes problematic in cases where there is significant bimodality with positive and negative differences in the field, but doing the block bootstrapping is an attempt to remedy this issue).

B.4 Linear regression and influence plots: a brief explanation of the statistical terms and equations

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC):
AIC and BIC are often used in conjunction to measure relative model goodness-of-fit. The formula for AIC is

$$AIC = 2p - 2\ln(L) \tag{B.4.1}$$

where p is the number of model parameters and L is the maximum of the likelihood function, which is an estimate of goodness-of-fit for possible models with a varying number of parameters.

BIC has a similar purpose as AIC, but it imposes a larger penalty for more model parameters; the first terms is also scaled by the number of observations n:

$$BIC = \ln(n)p - 2\ln(L) \tag{B.4.2}$$

For both criteria, models with lower values are considered to have a better fit.

Leverage: In the influence plots, the x-axis denotes the models' leverage values. Points with high leverage have unusually divergent input values compared to the rest of the inputs.

In the influence plot function, the threshold for significant influence values is determined by

$$T = 2 \times \frac{DF_M + 1}{N} \tag{B.4.3}$$

where threshold T is a function of the model degrees of freedom (DF_M) and the number of observations (N).

Studentized Residuals: In the influence plots, the y-axis denotes the models' studentized residuals. Points with high studentized residuals have significantly large deviations from the other predicted outcome values.

In the influence plot function, the threshold for significant residuals is determined by

$$T = T.PPF\left(1 - \frac{\alpha}{2}, DF_R\right) \tag{B.4.4}$$

where threshold T is determined by the inverse of the t-distribution, $T.PPF$, which here is a function of parameter α (set to 0.05 by default in the function) and the residual degrees of freedom DF_R. The percent point function is the inverse of the cumulative distribution function, and returns the 97.5th percentile value of the distribution.

Cook's Distance: The size of the points in the influence plots is scaled by Cook's distance:

$$D_i = \frac{\sum_{j=1}^{n}(\hat{Y}_j - \hat{Y_{j(i)}})^2}{(p+1)\hat{\sigma}^2} \tag{B.4.5}$$

The distance D_i is a measure of how much the values in the regression model change when the ith observation is removed from the calculations; $\hat{Y}_j$ is the fit, and $\hat{Y_{j(i)}}$ is the fit excluding i. Large points denote models that significantly impact the resulting fit of the regression equation.

The rule of thumb for the Cook's distance threshold T is determined by the number of observations,

$$T = \frac{4}{N} \tag{B.4.6}$$

Variance inflation factor (VIF):   The procedure can be generally described as follows:

1. Calculate the regression between the chosen input and all other inputs:

$$X_i = a_0 + a_1 X_1 + a_2 X_2... \tag{B.4.7}$$

   and obtain the R^2 value for each X_i

2. Calculate VIF:

$$VIF_i = \frac{1}{1 - R_i^2} \tag{B.4.8}$$

3. If VIF_i is above 5, drop X_i from the list of all possible X

4. Repeat the above steps until there are no VIF_i greater than 5

121

Appendix C

Chapter 3 Supplemental Figures

Significant differences in NH JJA

variable	CMIP5Mean	CMCC-CM	MRI-CGCM3	MRI-ESM1	BCC-CSM1.1	MIROC5	CNRM-CM5	GFDL-ESM2M	MPI-ESM-MR	NorESM1-M	IPSL-CM5A-LR	GFDL-CM3	CanESM2	MIROC-ESM	CMCC-CESM	HadGEM2-CC	HadCM3
BF	-0.151	0.602		0.149	0.287	-1.027		0.727	0.237	-0.420	-1.753	0.372	0.328	-1.147	-0.256	0.718	0.516
BF_SD	-0.616	1.603	-0.346	0.286	1.330	-2.625		-2.066	1.134	-2.405	-4.652	1.089	0.777	-3.445	-1.173	2.182	1.423
NB																2.750	
LA									5.466				4.260				-3.305
LO														-11.812			
DD																	
ZD																	
ZS																	
BA	-0.180				-0.548			-0.372	-0.466				-0.677		-0.380		
AI	-0.810			-1.130	0.859	-1.880			-1.090	-1.916		-1.741	-2.741	-1.175	-1.522		-1.142
BI	-0.210		-0.244	0.329		-0.495			-0.392	-0.673		-0.256	-0.570		-0.365		
Z500	-3.781	29.024	-3.604	-9.181	80.915	10.282	20.944	1.05	-18.374	0.05	-47.644	-4.724	4.69	60.461		7.551	-34.936
Z500_SD	-1.179	4.828	-1.586	3.125	6.763	10.024	-3.769	-6.339	5.015	-3.568	6.171		2.506	-7.020	12.383		-9.044
U	-0.228	0.541	-1.234	-1.145	0.422	-1.100	-0.888	0.267	0.631	0.192	-1.566	0.594	0.162	-1.257	0.582	-1.112	-0.728
U_SD	-1.321	-0.141	-0.921	-1.042	-0.359	-1.814	1.023	-1.650	-0.414	-2.041	-2.376	-1.134	-1.392	-1.595	-2.044	-0.568	-1.724
V	-0.103	-0.216	0.141	0.159	0.406	0.430	0.264	0.139	-0.237	-0.206	0.144		0.081	0.099	-0.193	0.233	0.118
V_SD	-1.140	-0.246	-0.937	-0.996	-0.655	-1.597	0.989	-1.495	-0.158	-2.127	-2.232	-1.108	-1.473	-1.566	-1.927	-0.858	-1.788
J850			-2.490	-1.861	5.413				2.233	5.637	-1.528		2.946		-1.365	1.704	
J250	-4.818	-2.192	5.349	5.271		-9.771	2.362	3.711		-4.410	-3.539	-2.592	12.947	-4.750	6.763		-5.581

Figure C.1: More detailed version of Figure 3.7, for NP JJA. Boxes are shaded according to the normalized bias values, while numbers indicate the actual bias values. Boxes with numbers indicate significant differences.
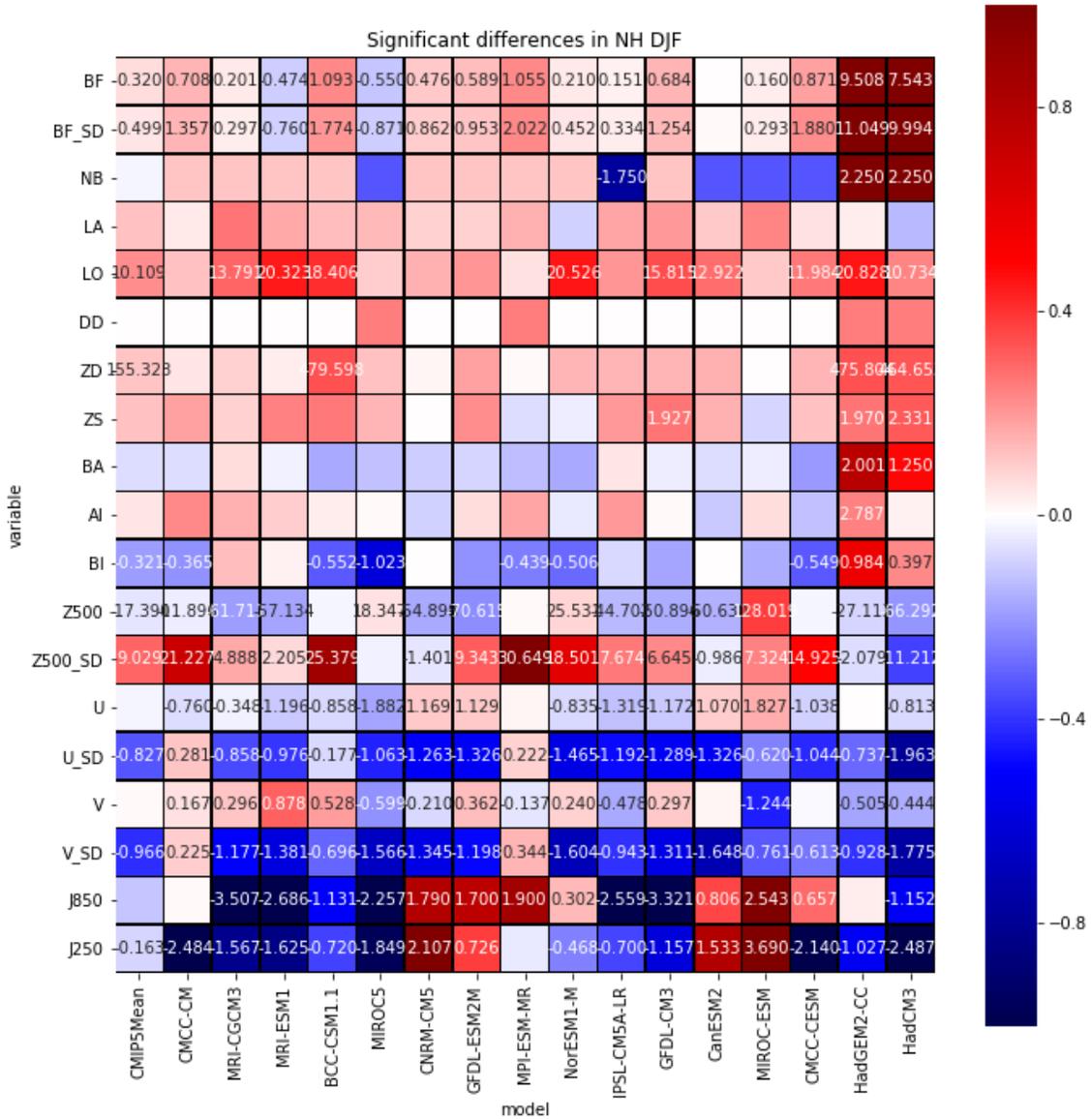
Figure C.2: As for Figure C.1, with NP DJF.
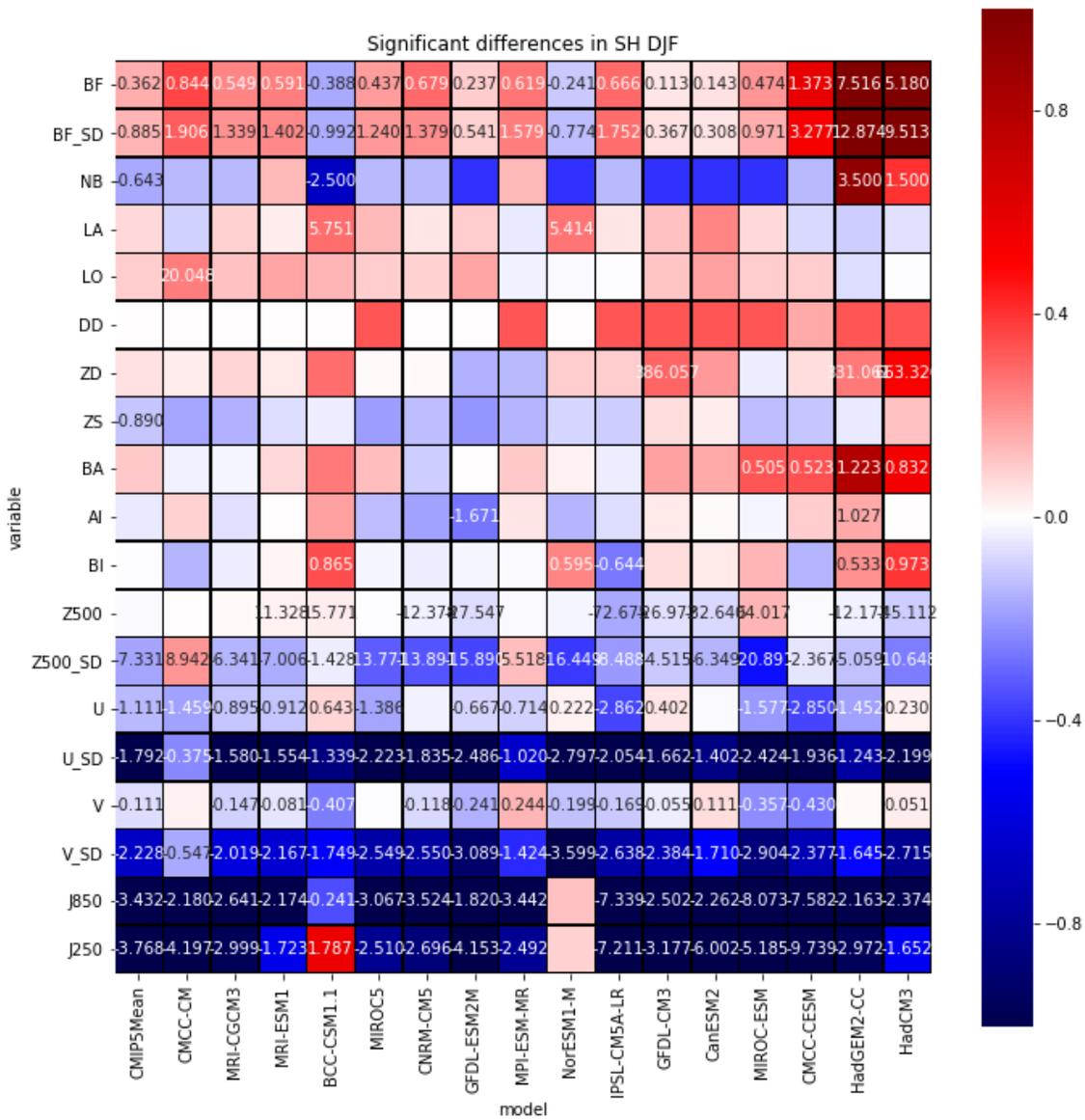
Figure C.3: As for Figure C.1, with SP DJF.

Significant differences in SH JJA

variable	CMIP5Mean	CMCC-CM	MRI-CGCM3	MRI-ESM1	BCC-CSM1.1	MIROC5	CNRM-CM5	GFDL-ESM2M	MPI-ESM-MR	NorESM1-M	IPSL-CM5A-LR	GFDL-CM3	CanESM2	MIROC-ESM	CMCC-CESM	HadGEM2-CC	HadCM3
BF	-0.471	0.424	0.428	0.216	-0.145	0.799	0.776	0.268	0.692	0.780	0.925	0.519	0.391		0.988	14.641	13.366
BF_SD	-0.829	0.870	0.808	0.520	-0.374	1.644	1.554	0.382	1.546	1.646	2.064	1.203	0.850		2.209	18.962	18.183
NB					-1.750										-1.750	2.250	2.250
LA	-3.412			-5.169					-6.649	-5.000							3.910
LO																	
DD																2.000	2.000
ZD			-332.480				-416.766										702.02
ZS			-1.796				-2.004	2.212					1.699				2.620
BA	-0.224				0.544				0.481				0.420		0.617	2.716	2.641
AI					2.062	-2.043			2.501							5.564	3.656
BI						-0.425			-0.426						0.542	1.635	2.246
Z500	-18.508	9.770	16.171	-18.492	20.175			-31.739	25.576		76.174	-41.356	62.139	66.591	11.866	15.424	19.843
Z500_SD	-1.668	7.901	4.135	3.405	4.395	16.929	8.783	11.733	7.830	-6.561	-2.598	7.131	3.970	11.313	5.463	2.087	4.821
U	-0.944	-1.382	0.934	-0.783	1.005	-3.461	-1.731	-1.409	-0.348	-0.942	-2.119	-0.568	0.766	0.818	-0.362	-1.505	0.498
U_SD	-1.217	-0.456	0.557	-0.578	-0.783	-1.988	-1.572	-1.949	-0.714	-2.163	-1.566	-1.702	-1.174	-1.299	-1.091	-0.801	-1.520
V		-0.167	0.154	0.337	-0.251	0.419	0.162		-0.168	-0.291			0.050	-0.128	-1.244	-0.150	-0.700
V_SD	-1.819	-0.764	-1.579	-1.700	-1.378	-2.653	-2.063	-2.750	-1.215	-2.929	-1.560	-2.115	-1.864	-2.142	-1.486	-1.403	-1.995
J850	-3.297	-2.405	-0.749			-7.756	2.693	-1.383	-3.995	-0.818	-7.191	-0.360	-2.582	-7.147	-7.141	-1.868	-1.536
J250	-2.974	-3.084	-3.742	-3.895	-0.982	-3.268	-4.414	-3.120	-0.873	-3.089	-3.869	-4.048	-2.794	-2.351	-1.258	2.191	-1.977

Figure C.4: As for Figure C.1, with SP JJA.
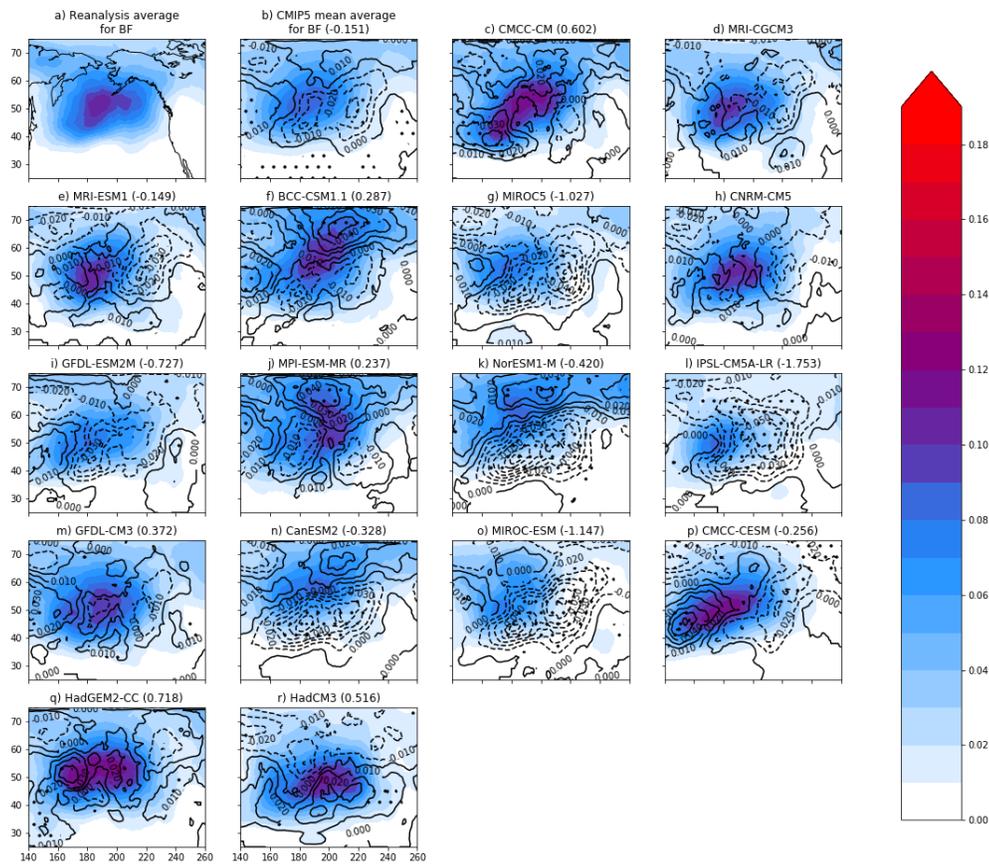
Figure C.5: Blocking frequency for NP JJA, with the reanalysis mean (top left), the CMIP5 model mean frequency, and the individual model frequency patterns as filled contours in 0.01 interval spacing. For each model, the differences from the reanalysis average are denoted by the black contours in 0.02 interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
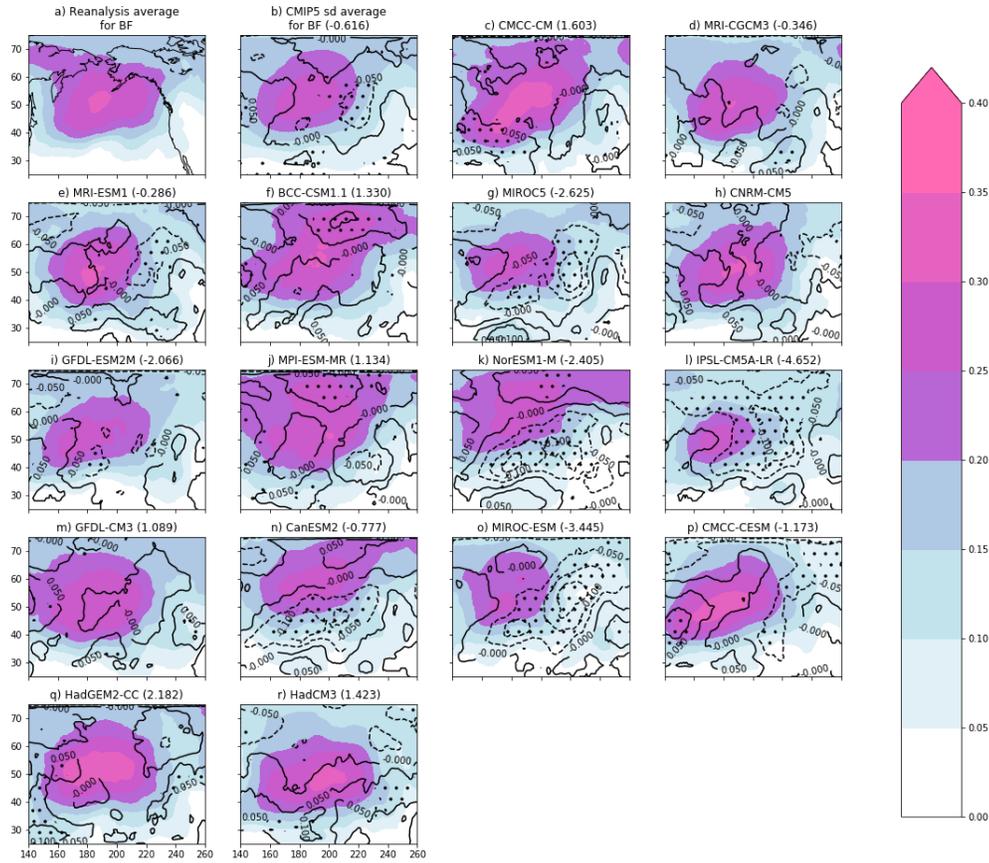
Figure C.6: Blocking frequency variability for NP JJA, with the reanalysis mean standard deviation (top left), the CMIP5 model mean standard deviation frequency, and the individual model frequency standard deviation patterns as filled contours in 0.05 interval spacing. For each model, the differences from the reanalysis average are denoted by the black contours in 0.05 interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
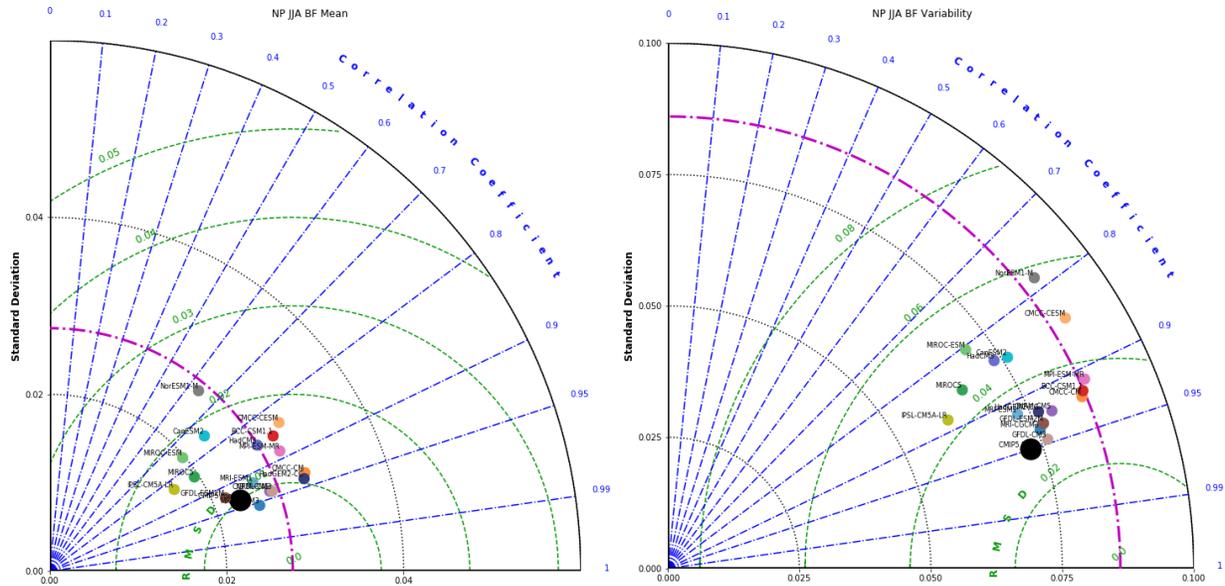
Figure C.7: Taylor plot of (left) the seasonal mean BF field and (right) the temporal variability in the seasonal BF field (measured as the standard deviation of the per-gridpoint time series). The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. The large black dot signifies the CMIP5 (excluding the Hadley models) average. Note that the maximum extent of the standard deviation axis is 0.06 for the mean plot while it is 0.10 for the variability plot.

Figure C.8: As for Figure C.5, with NP DJF
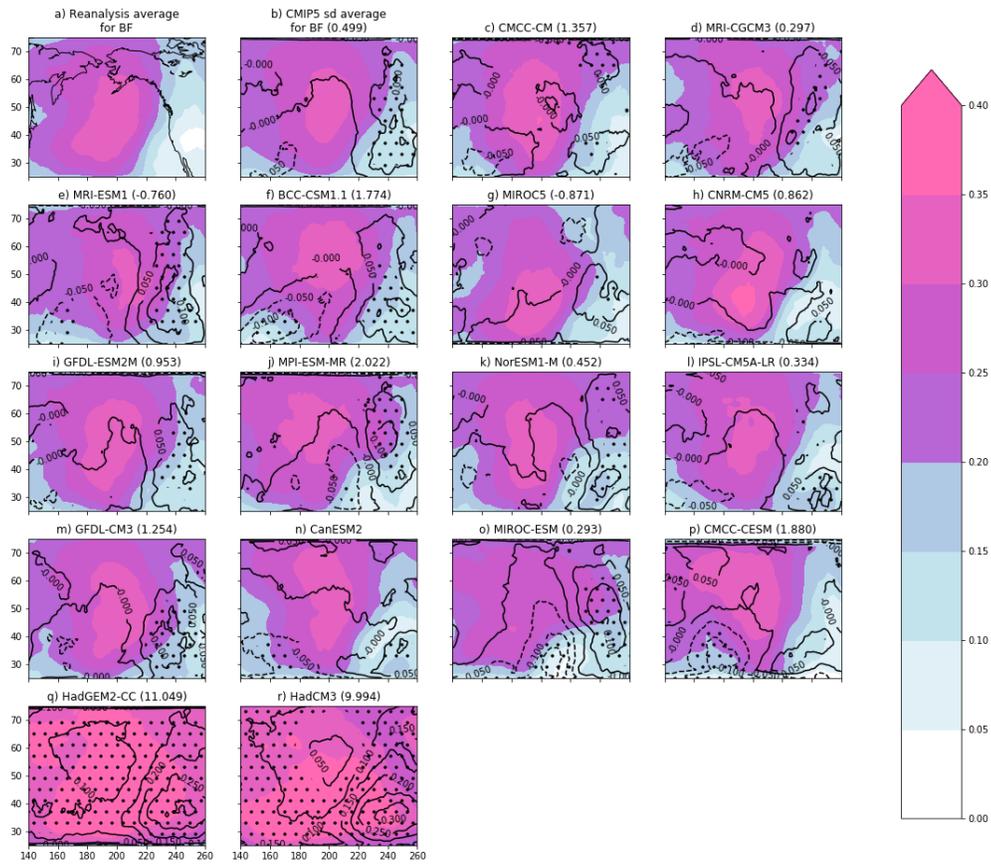
Figure C.9: As for Figure C.6, with NP DJF

131
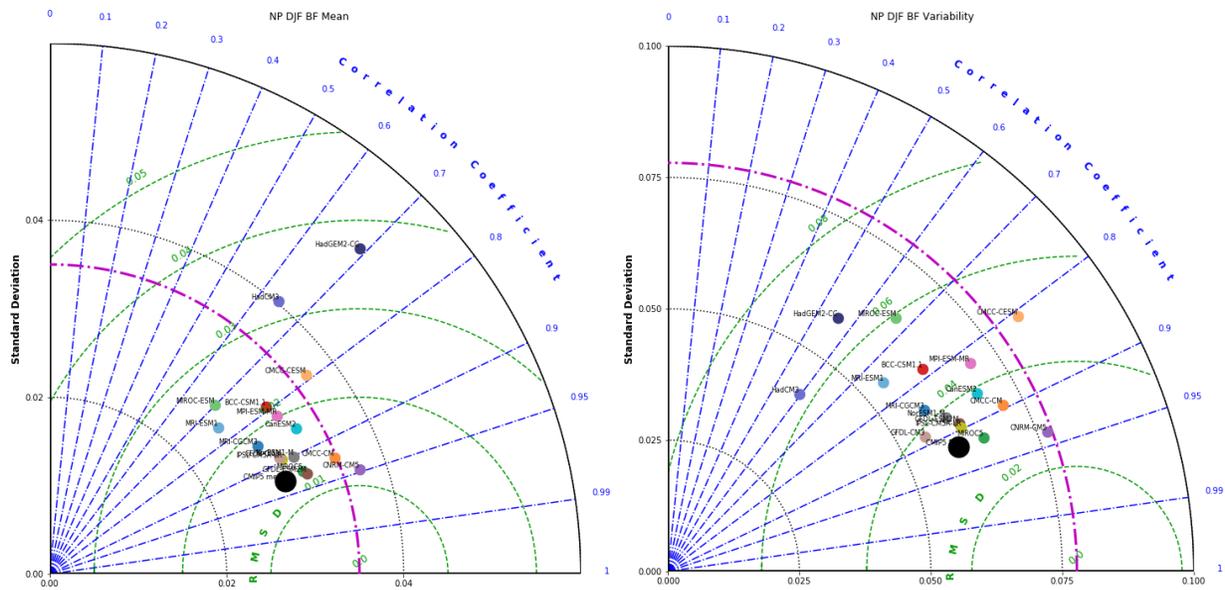
Figure C.10: As for Figure C.7, with NP DJF.
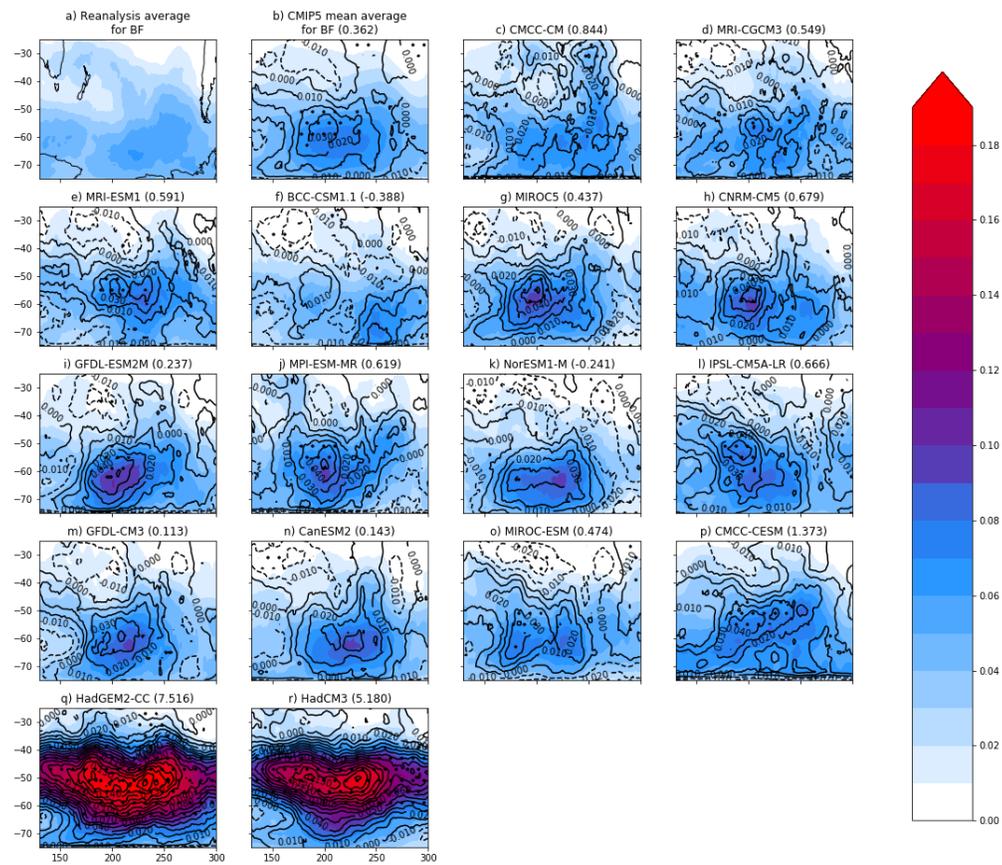
Figure C.11: As for Figure C.5, with SP DJF
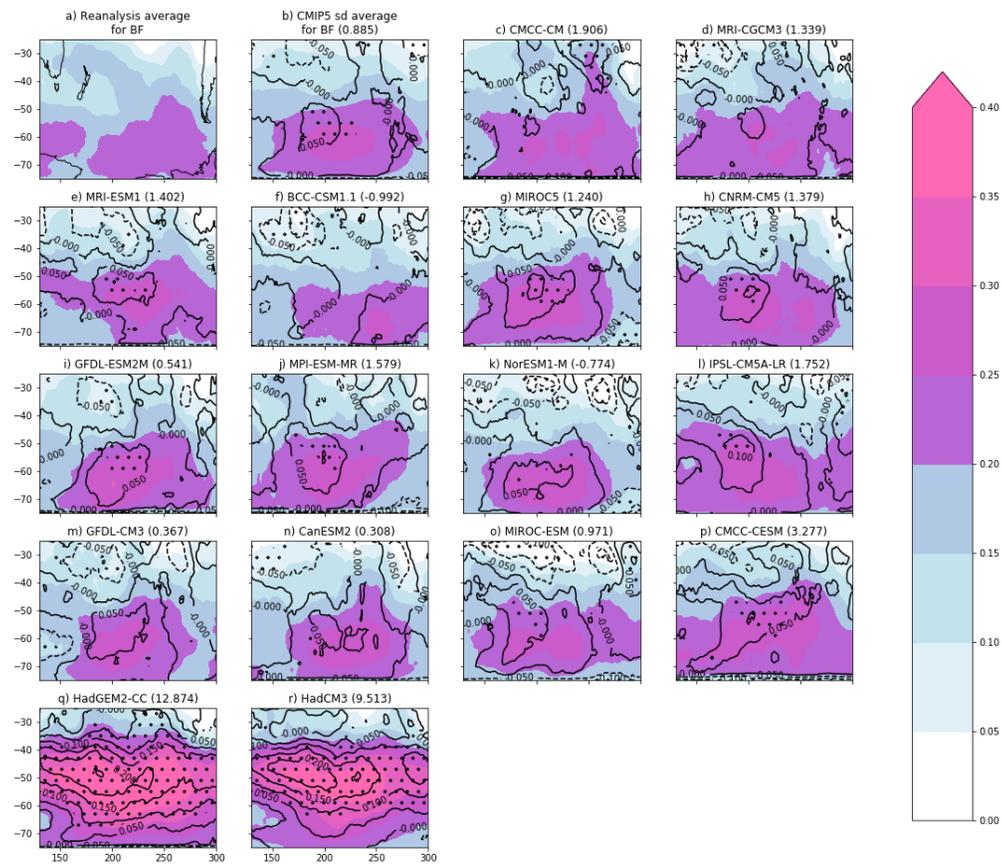
Figure C.12: As for Figure C.6, with SP DJF
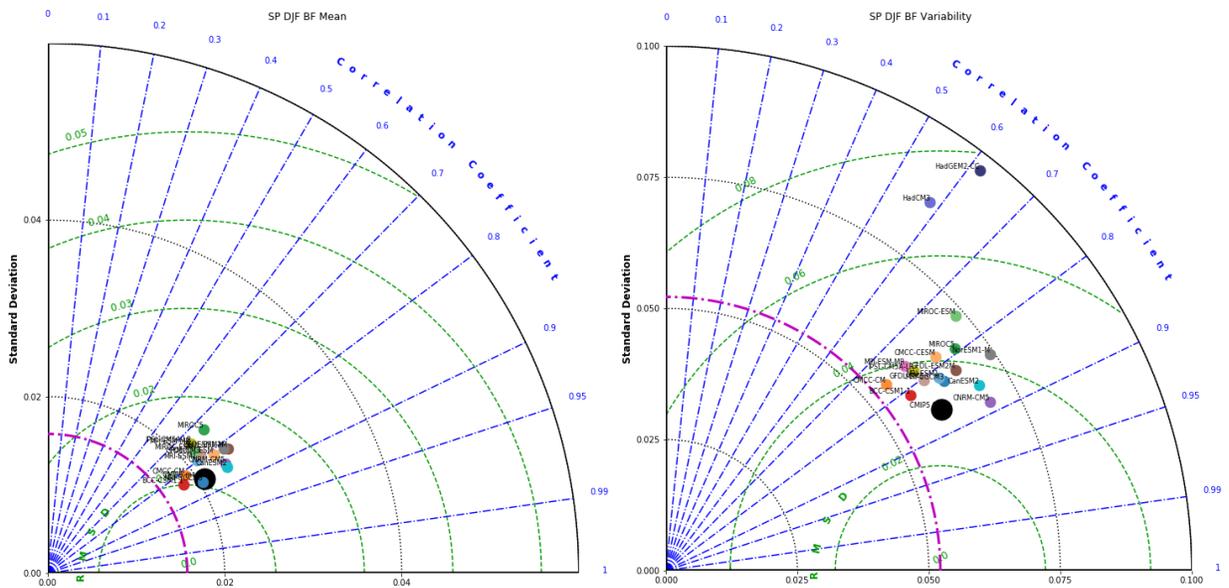
Figure C.13: As for Figure C.7, with SP DJF. Note that the Hadley models are missing from the mean plot, as their standard deviation values exceeded the left plot axis limits.
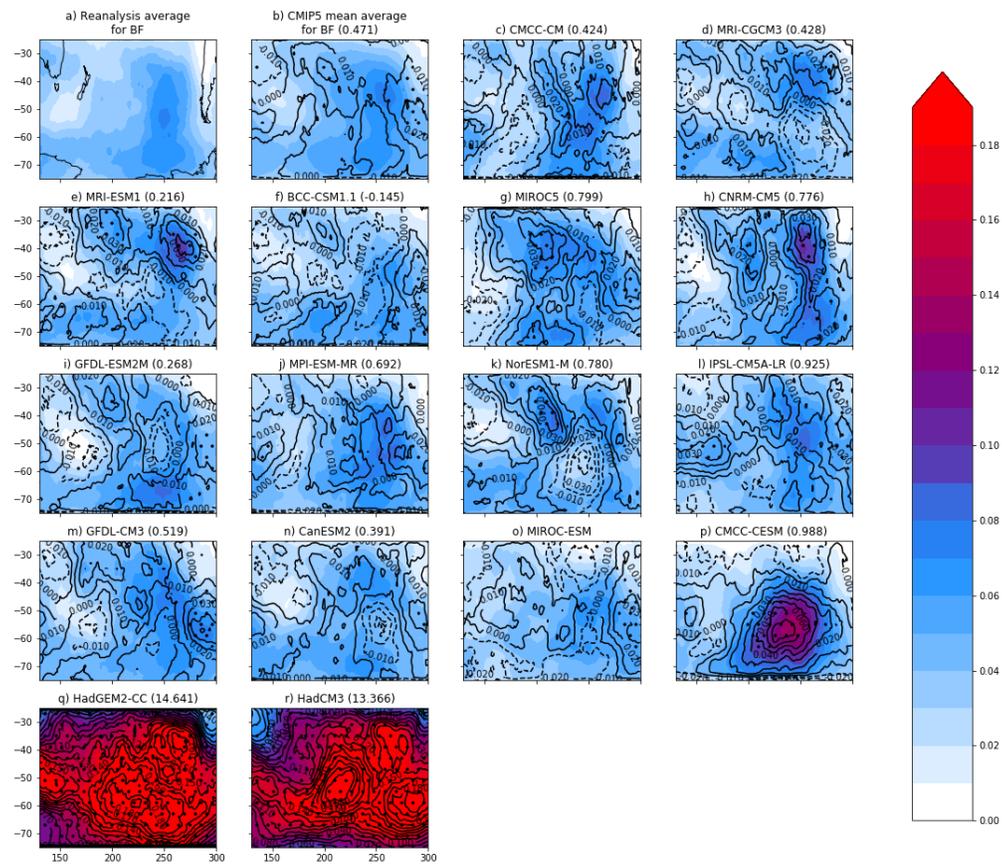
Figure C.14: As for Figure C.5, with SP JJA
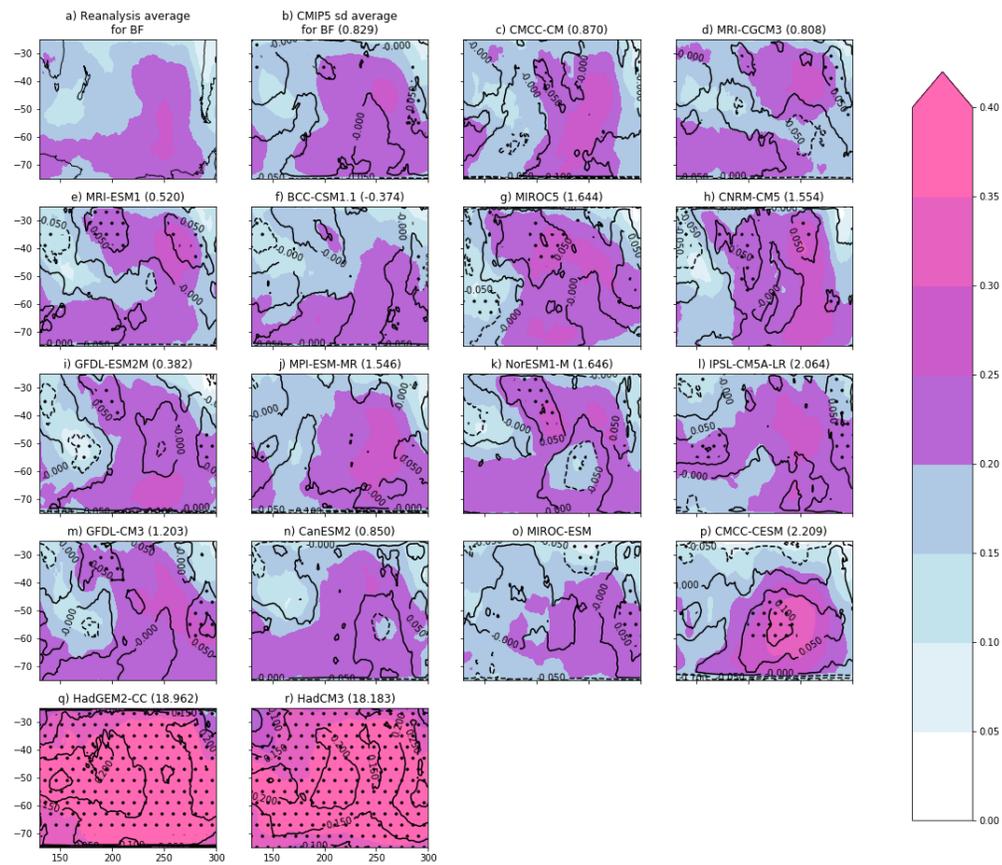
Figure C.15: As for Figure C.6, with SP JJA
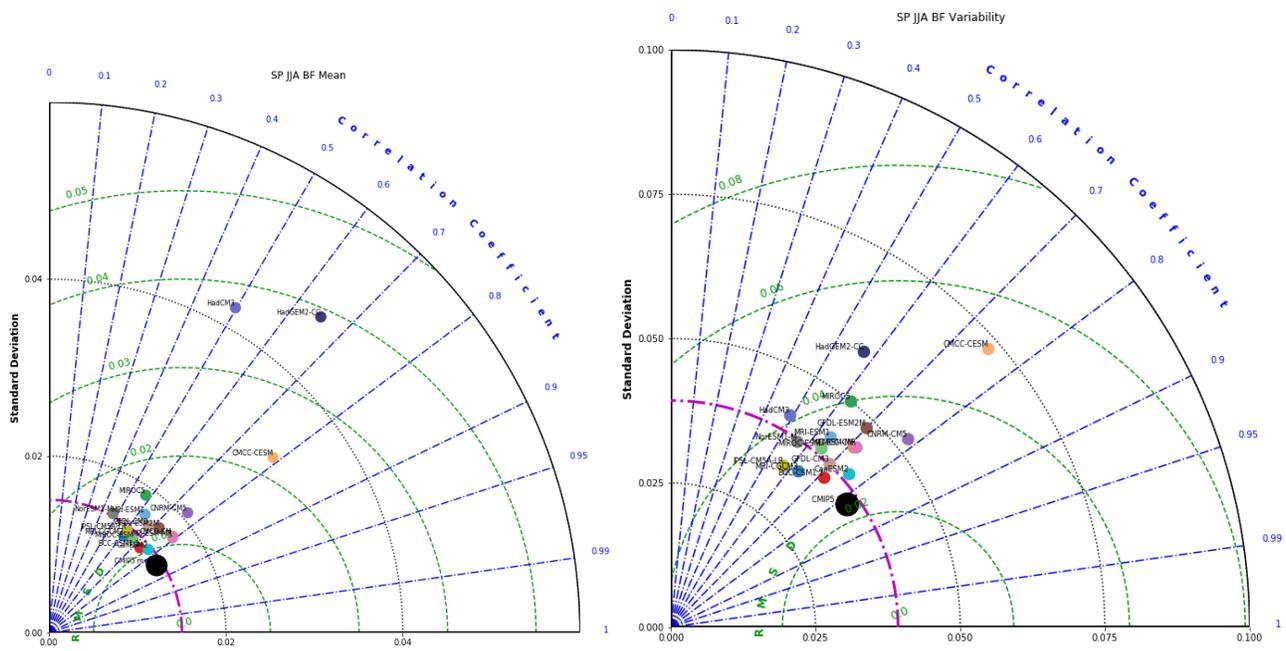
Figure C.16: As for Figure C.7, with SP JJA.
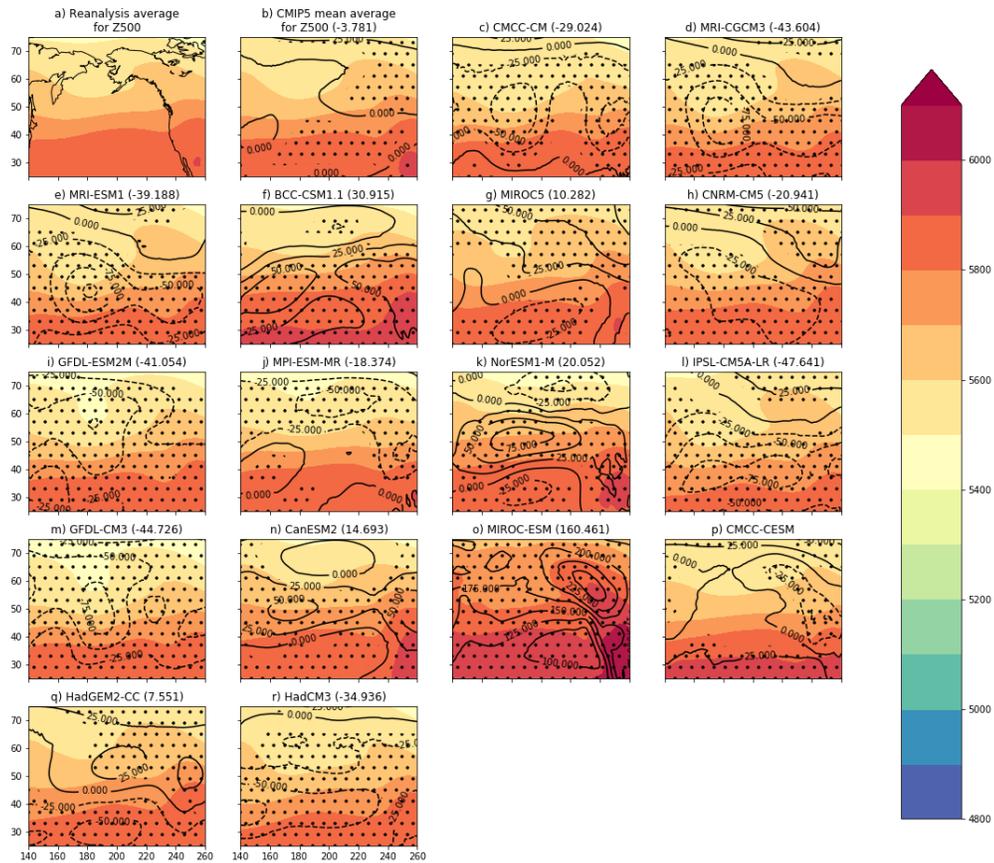
Figure C.17: 500 mb geopotential height for NP JJA, with the reanalysis mean (top left), the CMIP5 model mean geopotential height, and the individual model geopotential height averages as filled contours with 100 mb interval spacing. For each model, the differences from the reanalysis average are denoted by black contours in 25 mb interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
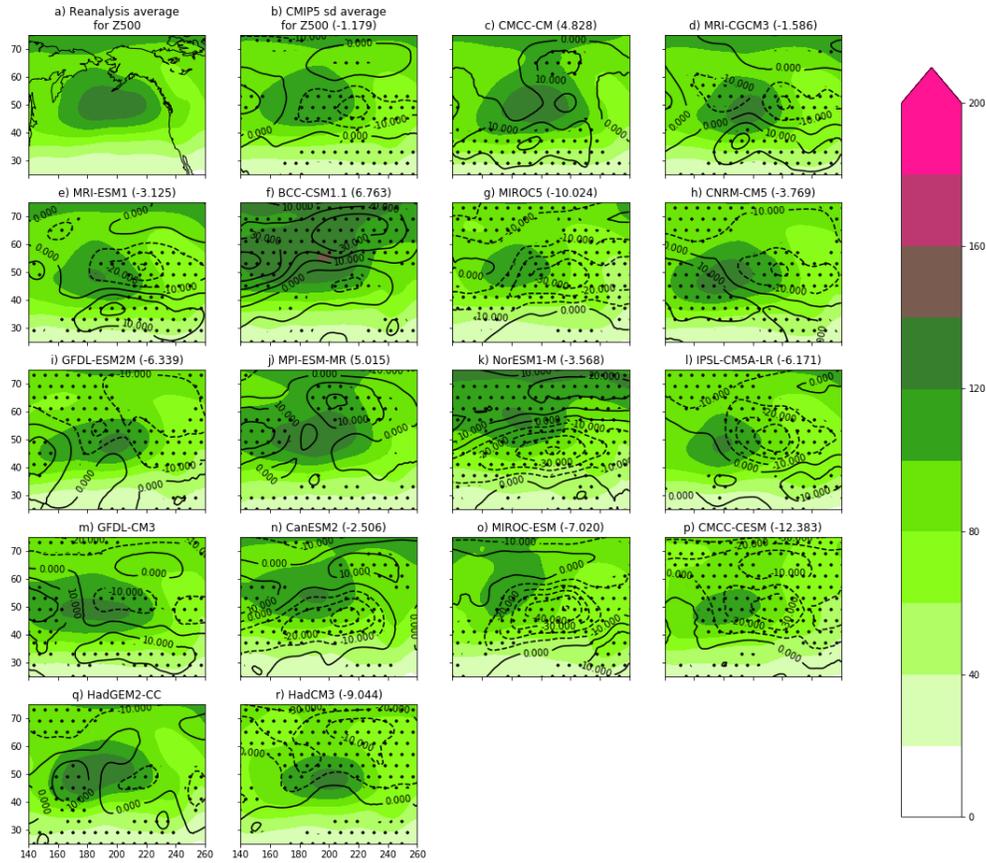
Figure C.18: 500 mb geopotential height variability for NP JJA, with the reanalysis standard deviation (top left), the CMIP5 model mean standard deviation of geopotential height, and the individual model geopotential height averages as filled contours with 20 mb interval spacing. For each model, the differences from the reanalysis average are denoted by black contours in 10 mb interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
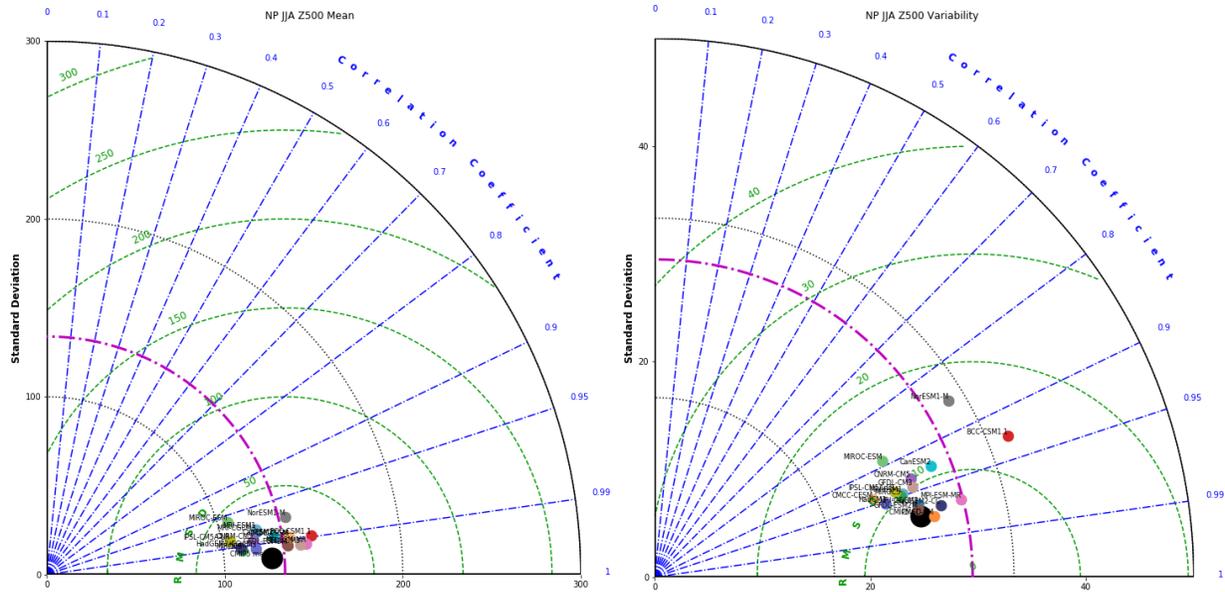
Figure C.19: Taylor plot of (left) the seasonal mean Z500 field and (right) the temporal variability in the seasonal Z500 field (measured as the standard deviation of the per-gridpoint time series). The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. The large black dot signifies the CMIP5 (excluding the Hadley models) average. Note that the maximum extent of the standard deviation axis is 300 for the mean plot while it is 50 for the variability plot.

Figure C.20: As for Figure C.17, with NP DJF
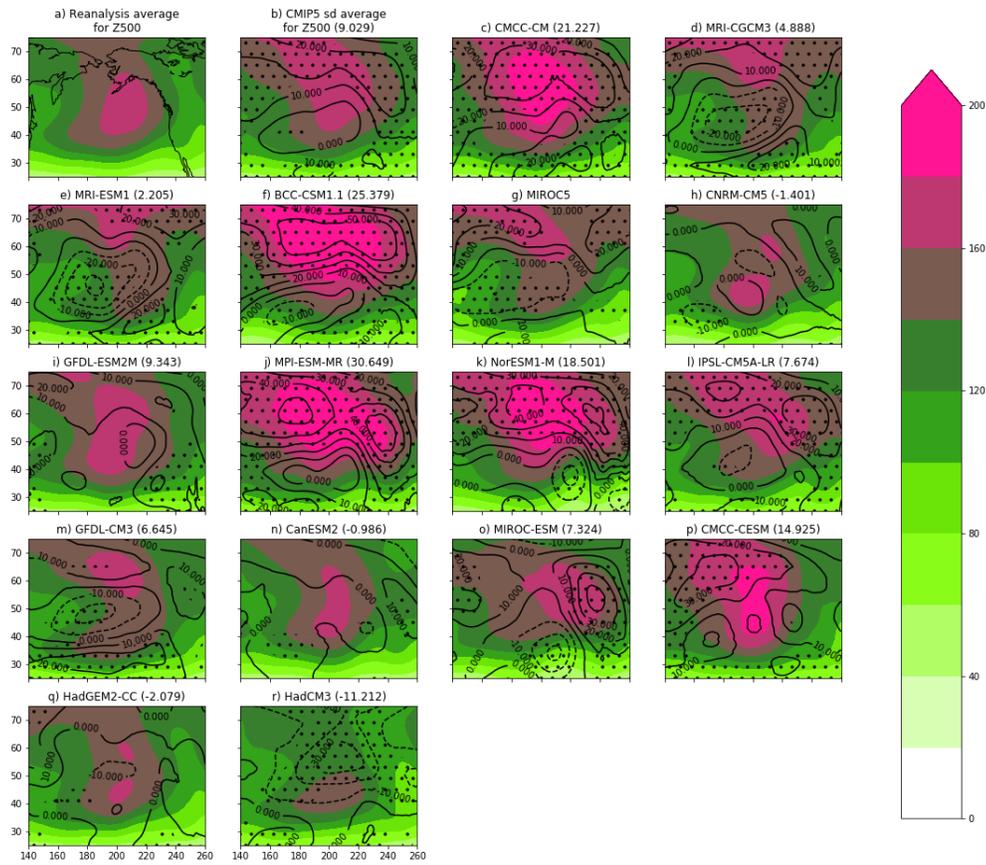
Figure C.21: As for Figure C.18, with NP DJF
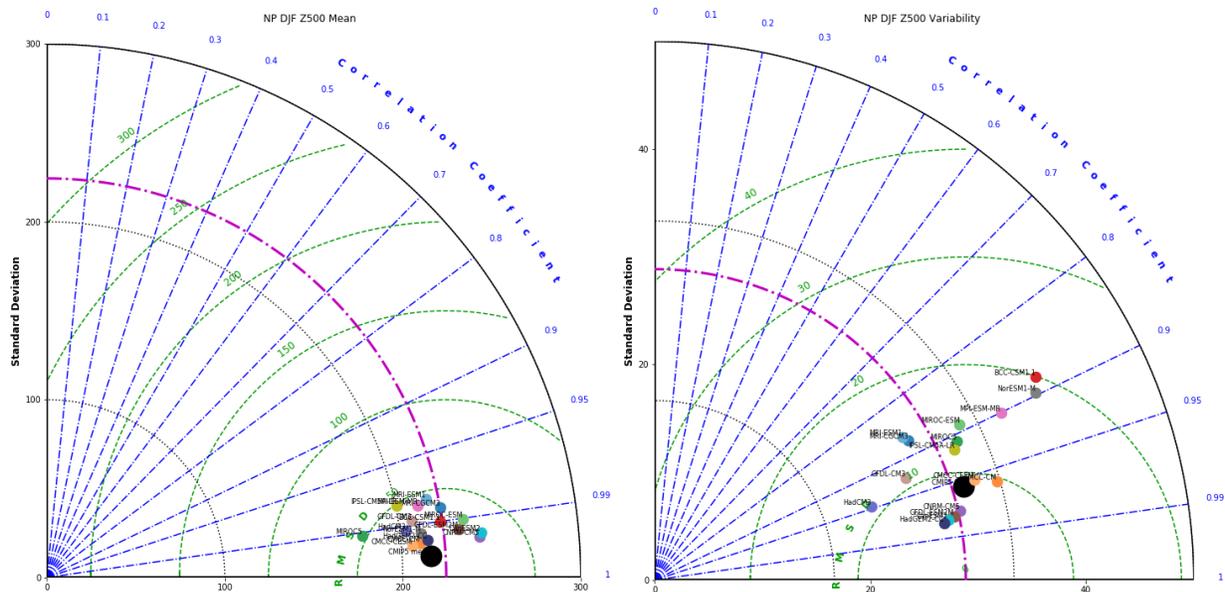
Figure C.22: As for Figure C.19, with NP DJF.
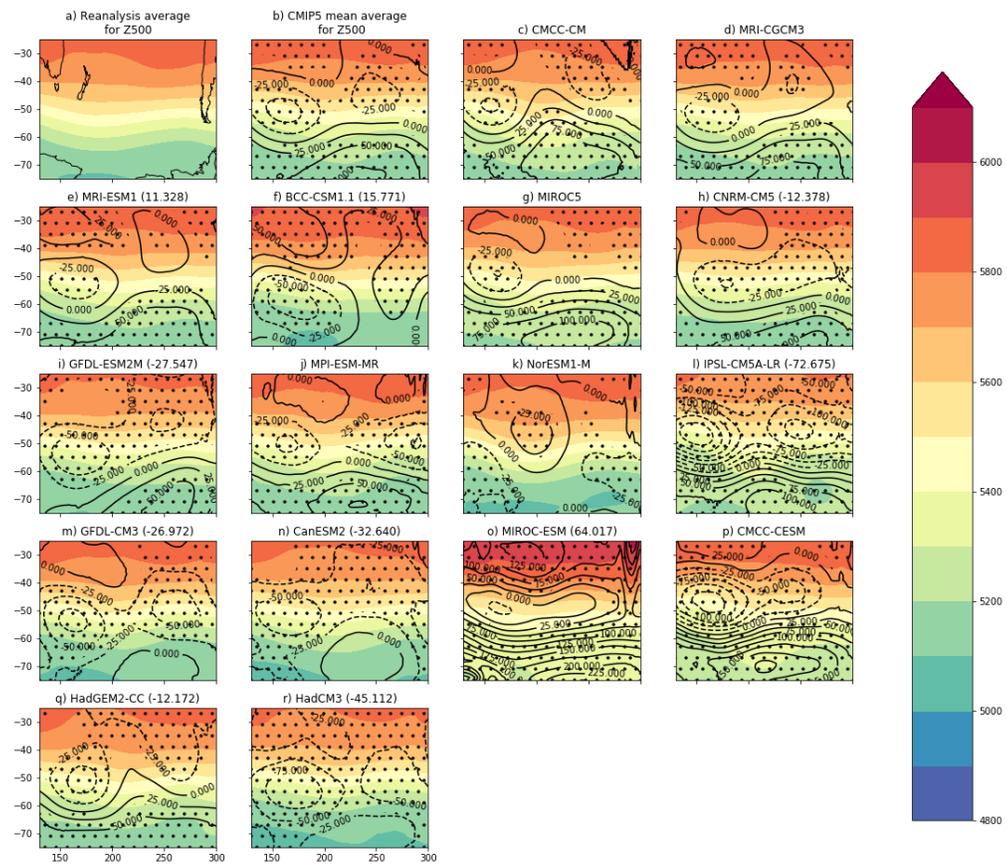
Figure C.23: As for Figure C.17, with SP DJF
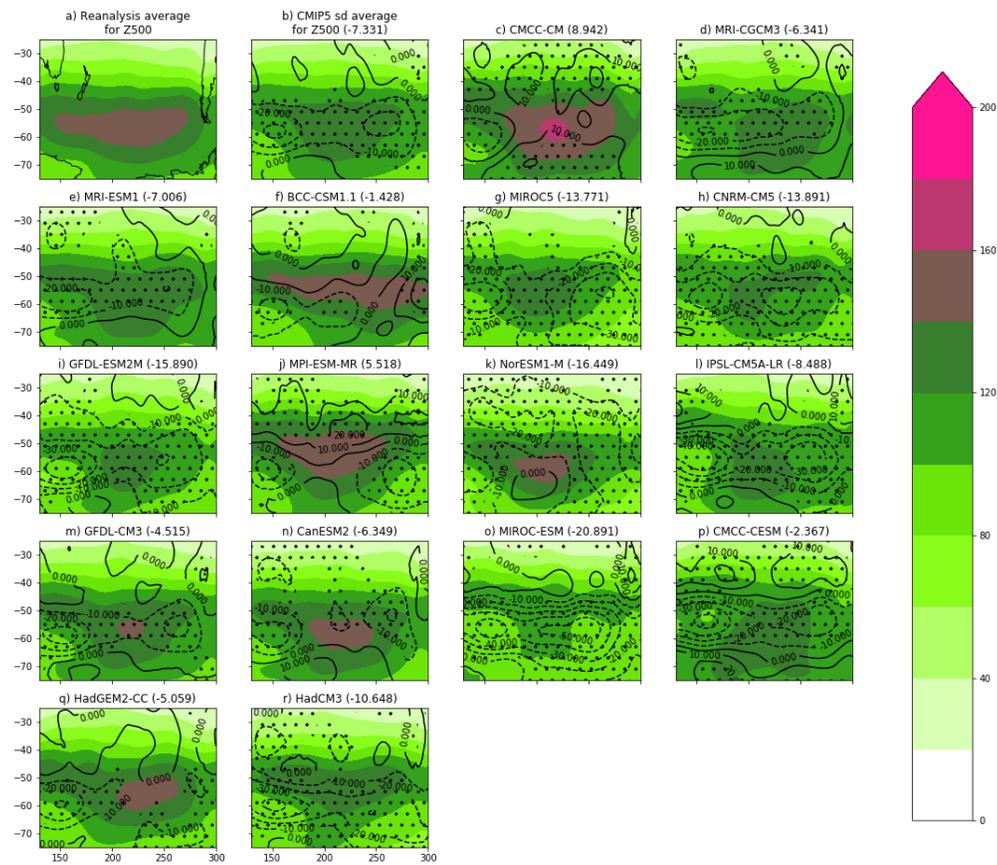
Figure C.24: As for Figure C.18, with SP DJF
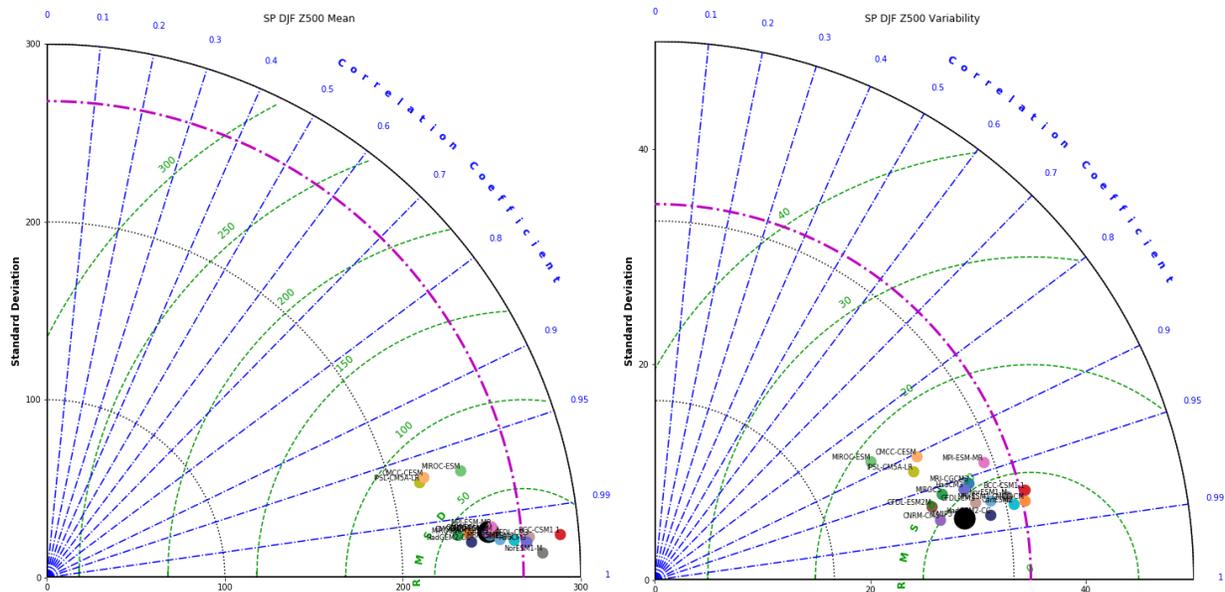
Figure C.25: As for Figure C.19, with SP DJF.
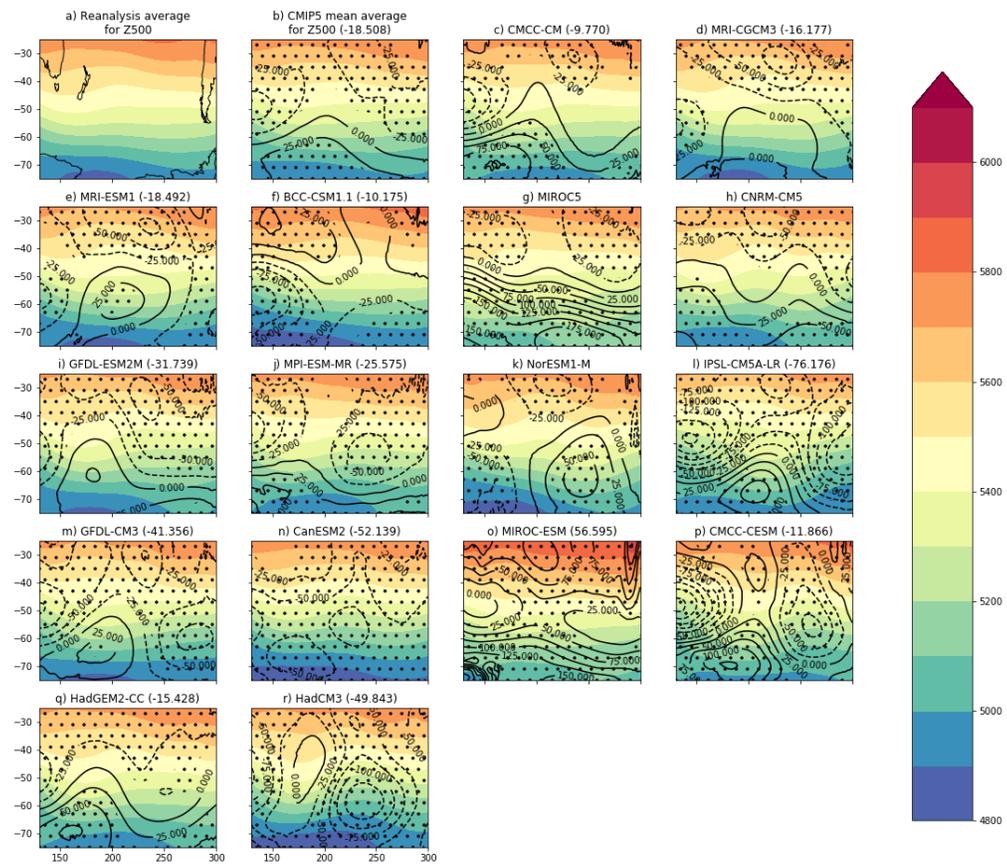
Figure C.26: As for Figure C.17, with SP JJA
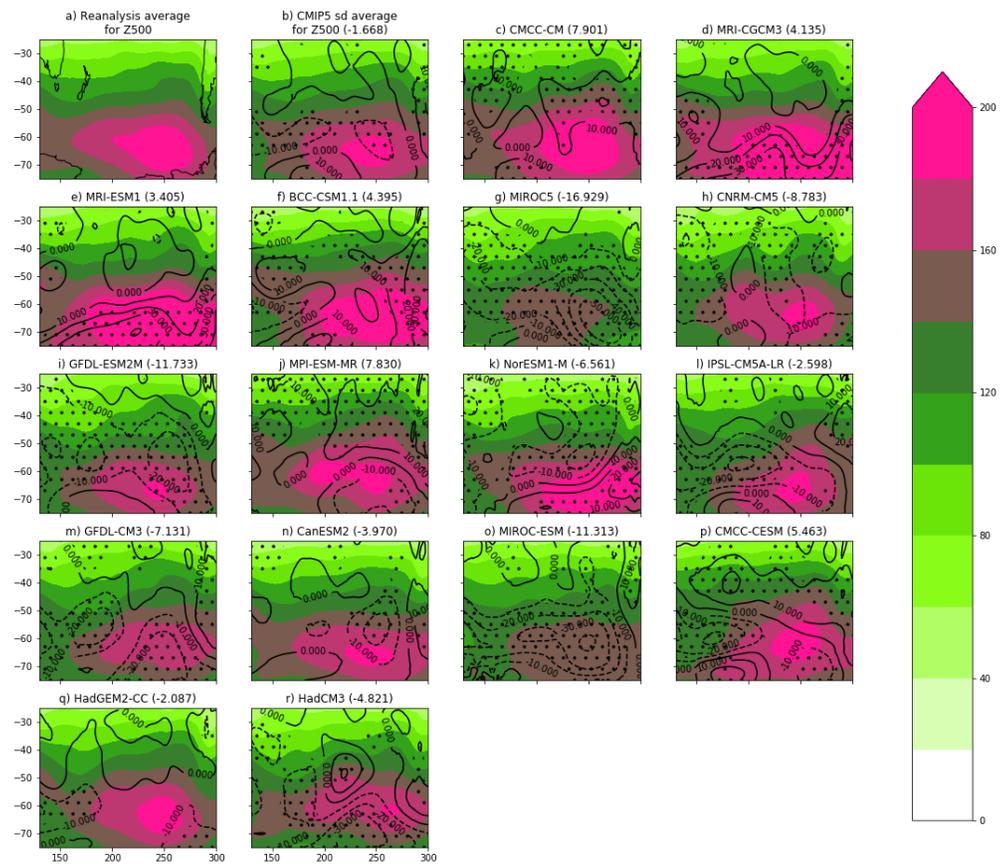
Figure C.27: As for Figure C.18, with SP JJA
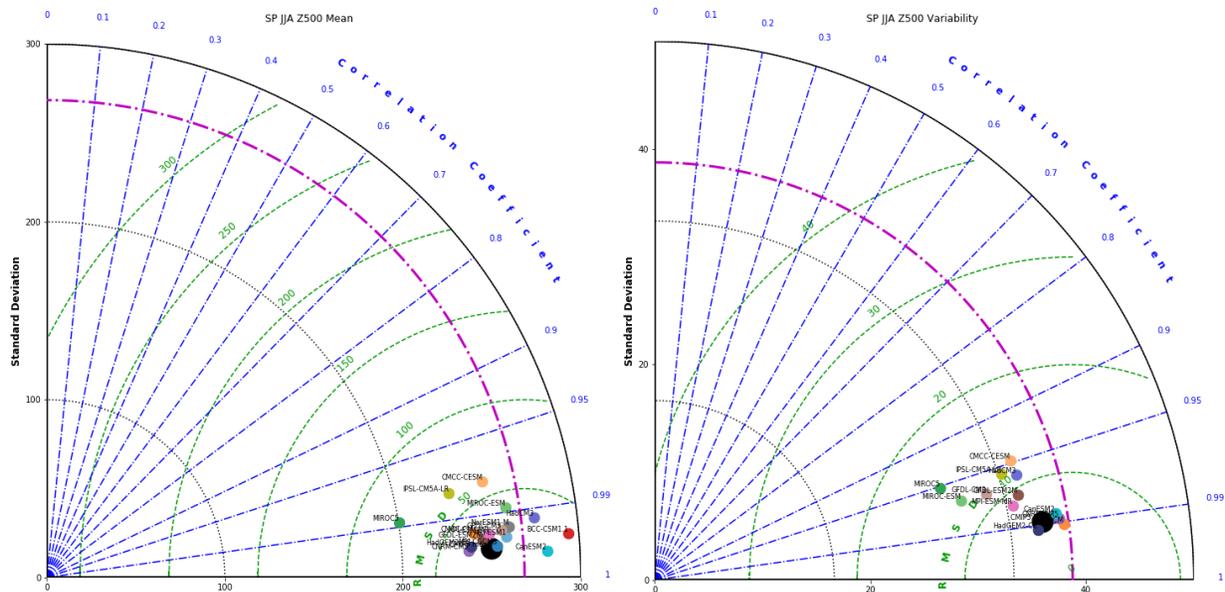
Figure C.28: As for Figure C.19, with SP JJA.
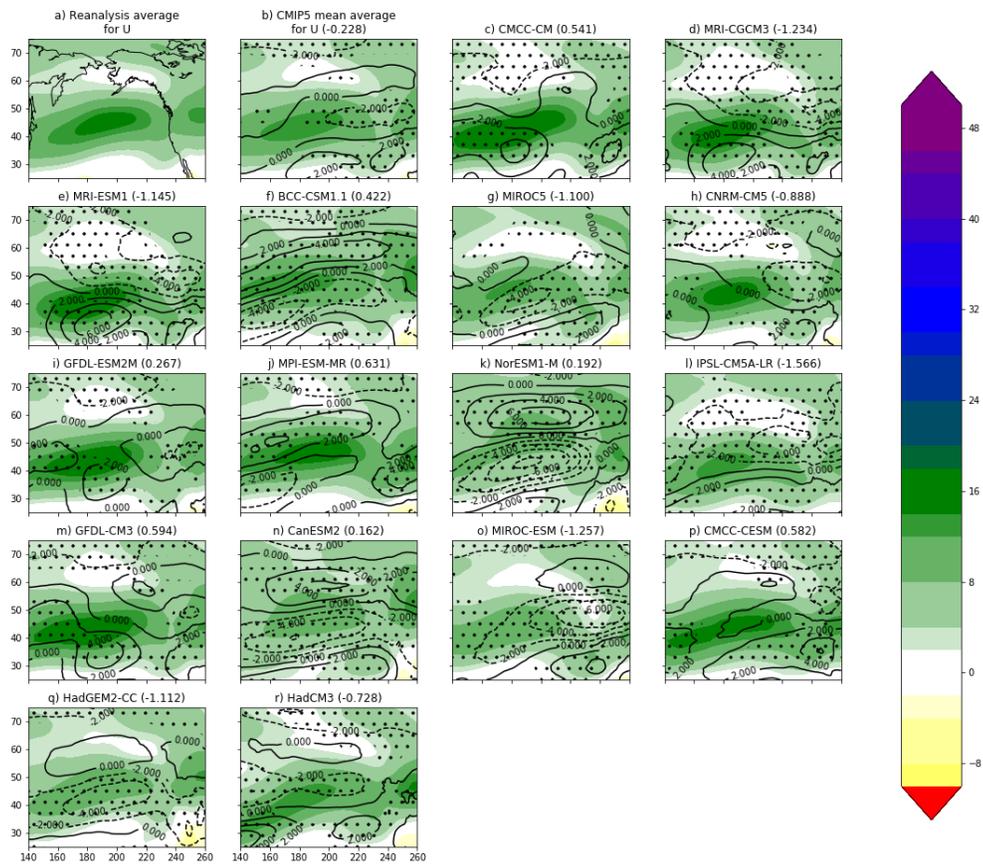
Figure C.29: 500 mb zonal wind for NP JJA, with the reanalysis mean (top left), the CMIP5 model mean zonal wind, and the individual model zonal wind averages as filled contours with 2 m/s interval spacing. For each model, the differences from the reanalysis average are denoted by black contours in 3 m/s interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
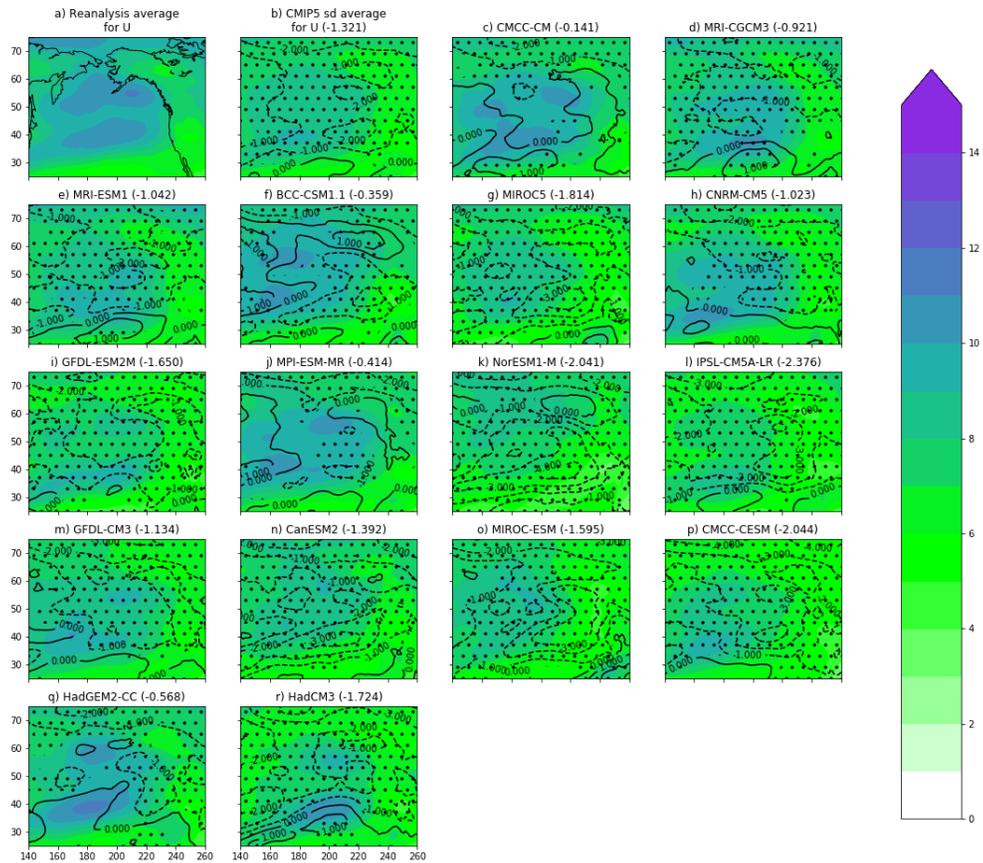
Figure C.30: 500 mb zonal wind variability for NP JJA, with the reanalysis mean standard deviation (top left), the CMIP5 model mean standard deviation of zonal wind, and the individual model zonal wind averages as filled contours with 1 m/s interval spacing. For each model, the differences from the reanalysis average are denoted by black contours in 1 m/s interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
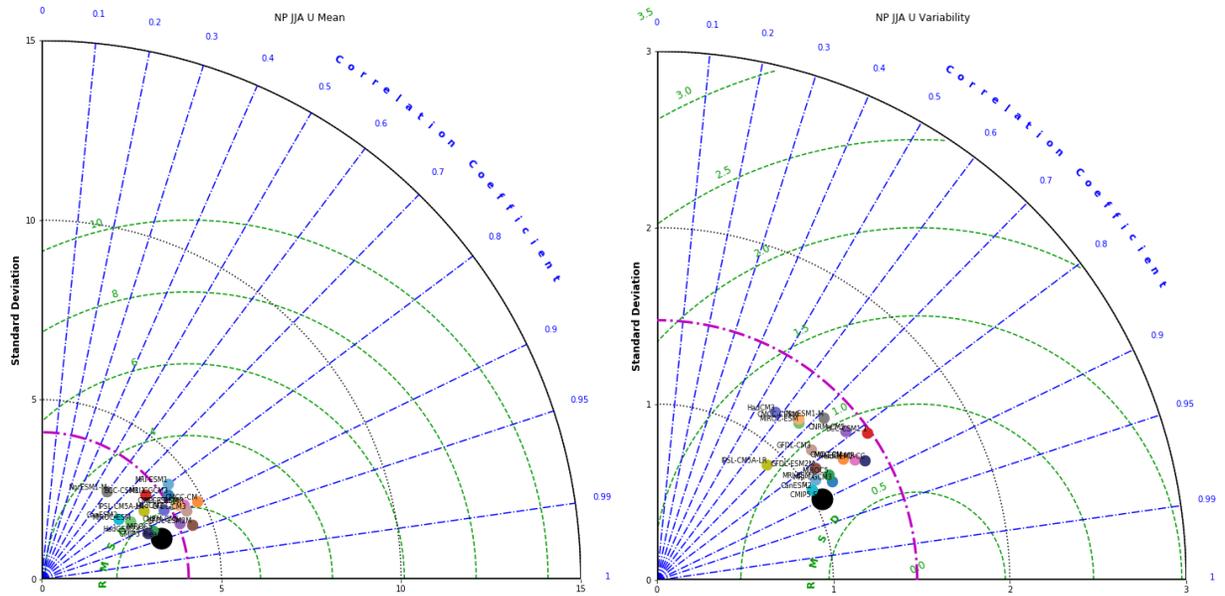
Figure C.31: Taylor plot of (left) the seasonal mean U field and (right) the temporal variability in the seasonal U field (measured as the standard deviation of the per-gridpoint time series). The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. The large black dot signifies the CMIP5 (excluding the Hadley models) average. Note that the maximum extent of the standard deviation axis is 15 for the mean plot while it is 3 for the variability plot.

Figure C.32: As for Figure C.29, with NP DJF
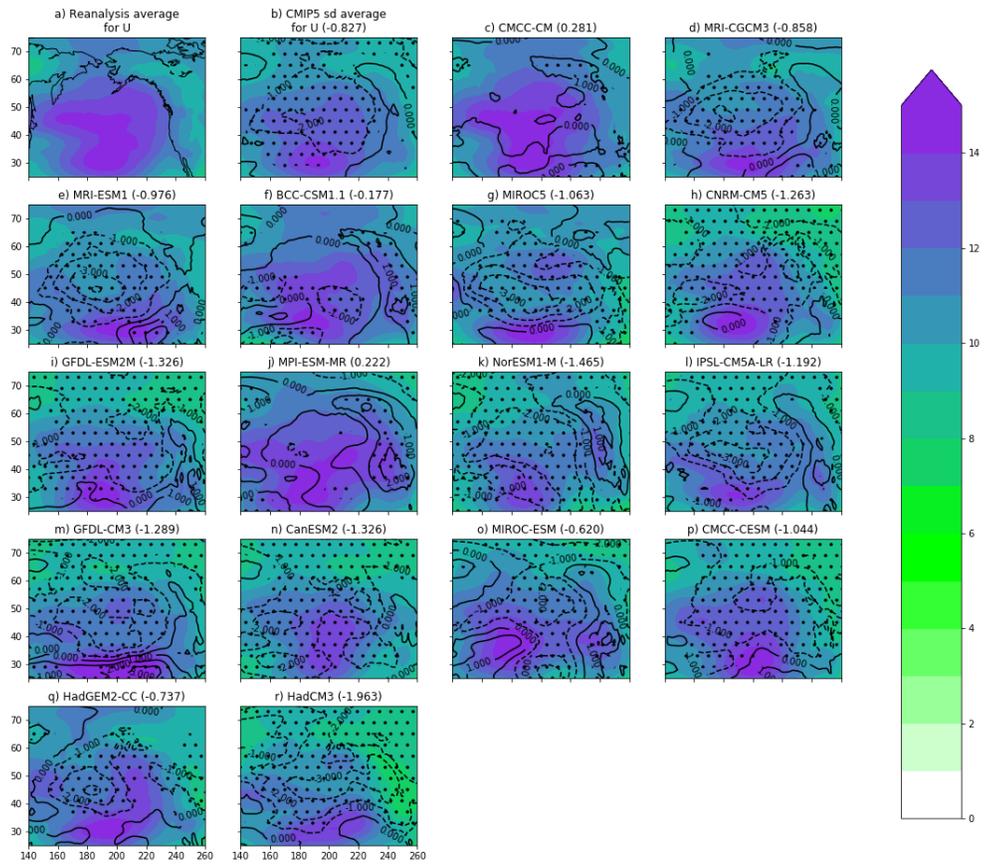
Figure C.33: As for Figure C.30, with NP DJF
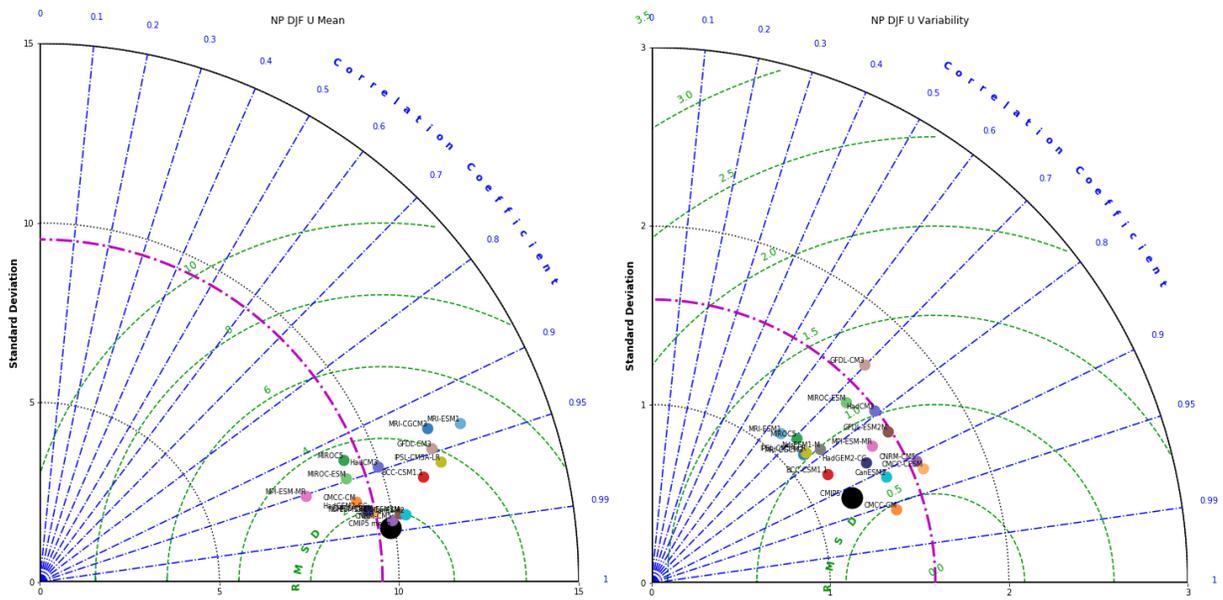
Figure C.34: As for Figure C.31, with NP DJF.
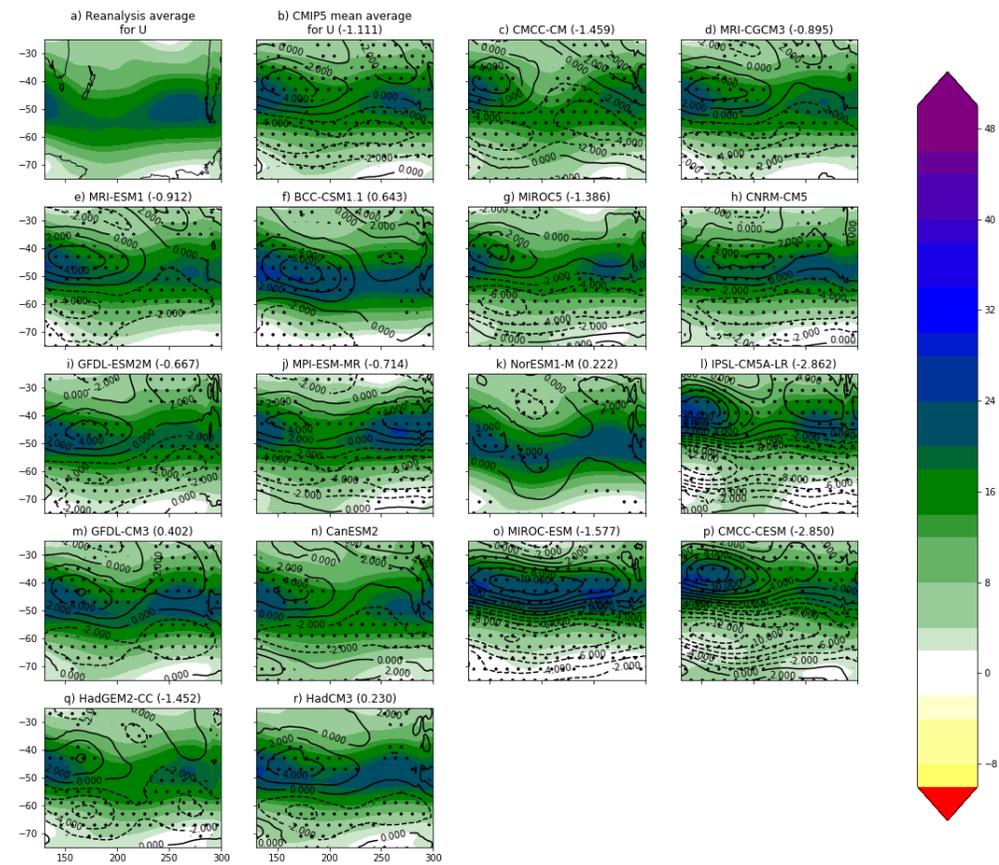
Figure C.35: As for Figure C.29, with SP DJF
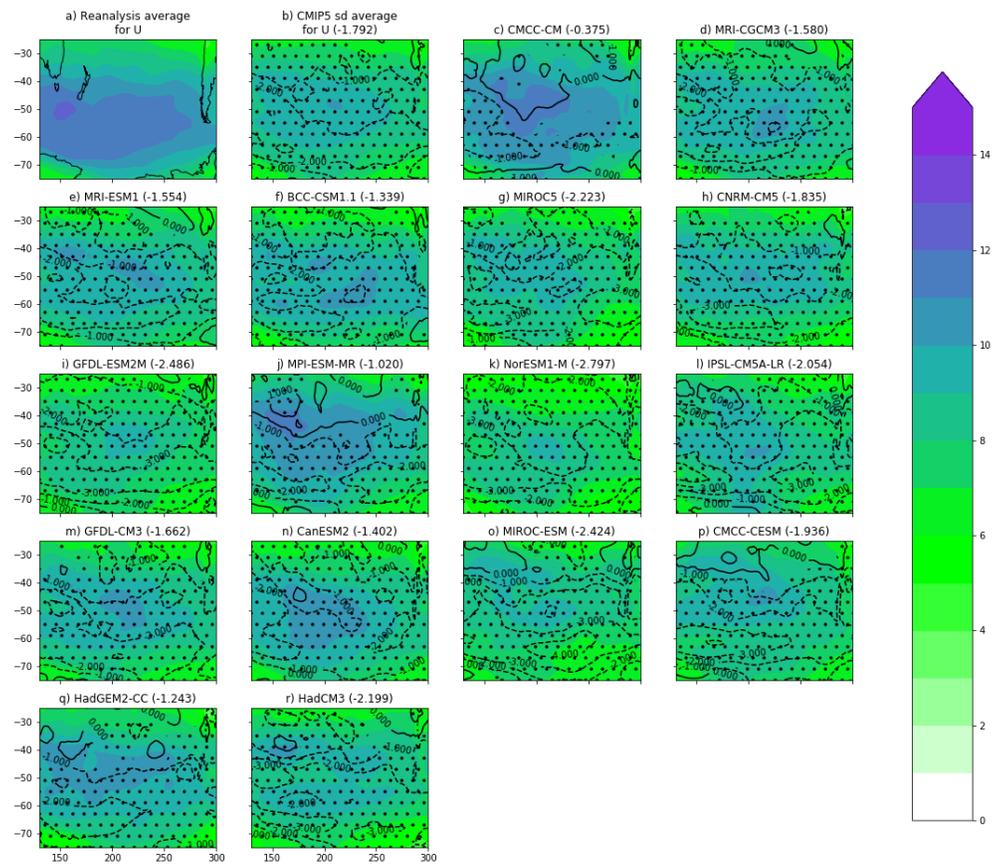
Figure C.36: As for Figure C.30, with SP DJF
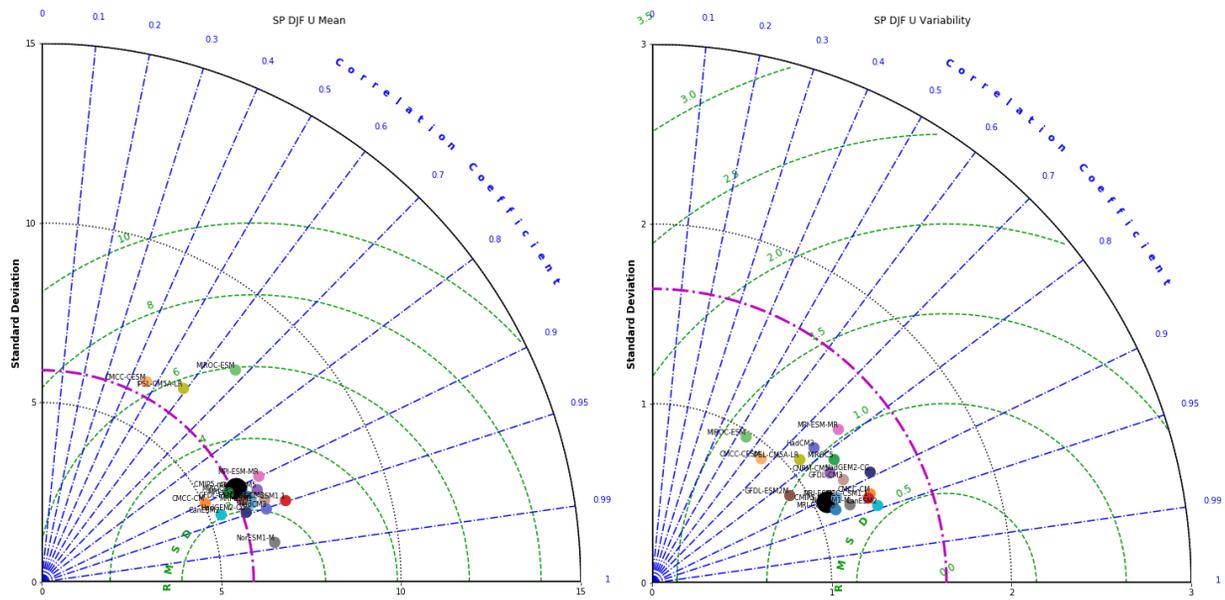
Figure C.37: As for Figure C.31, with SP DJF.
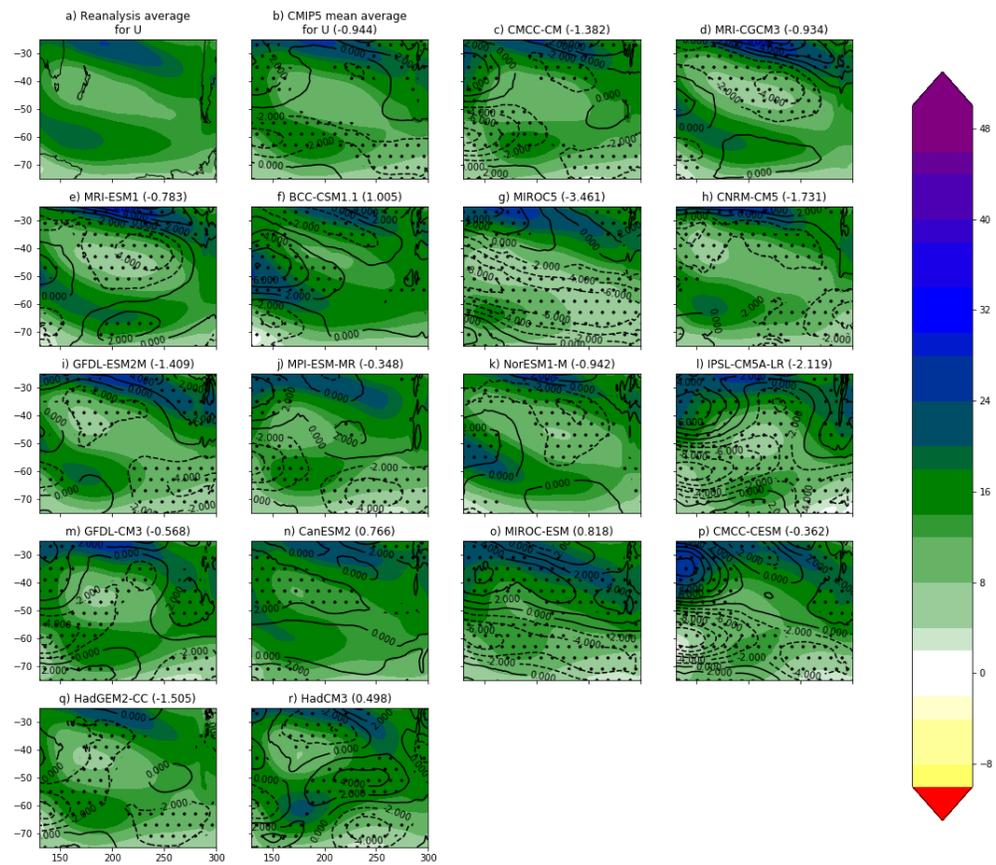
Figure C.38: As for Figure C.29, with SP JJA
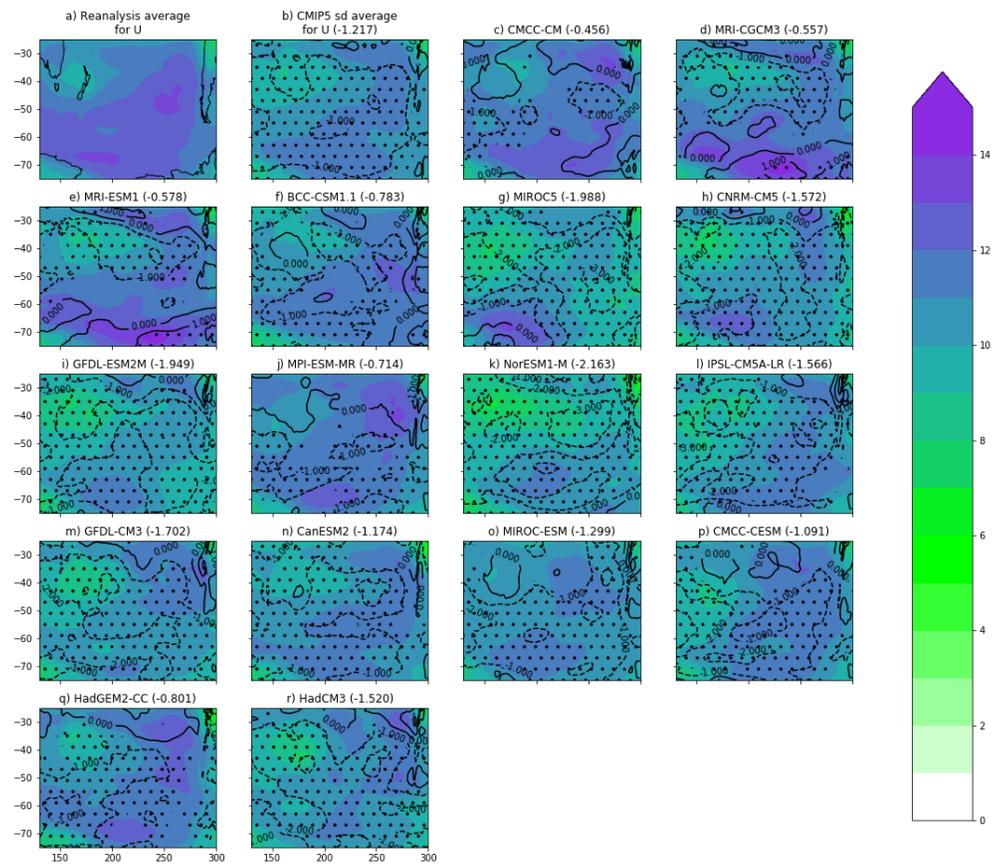
Figure C.39: As for Figure C.30, with SP JJA
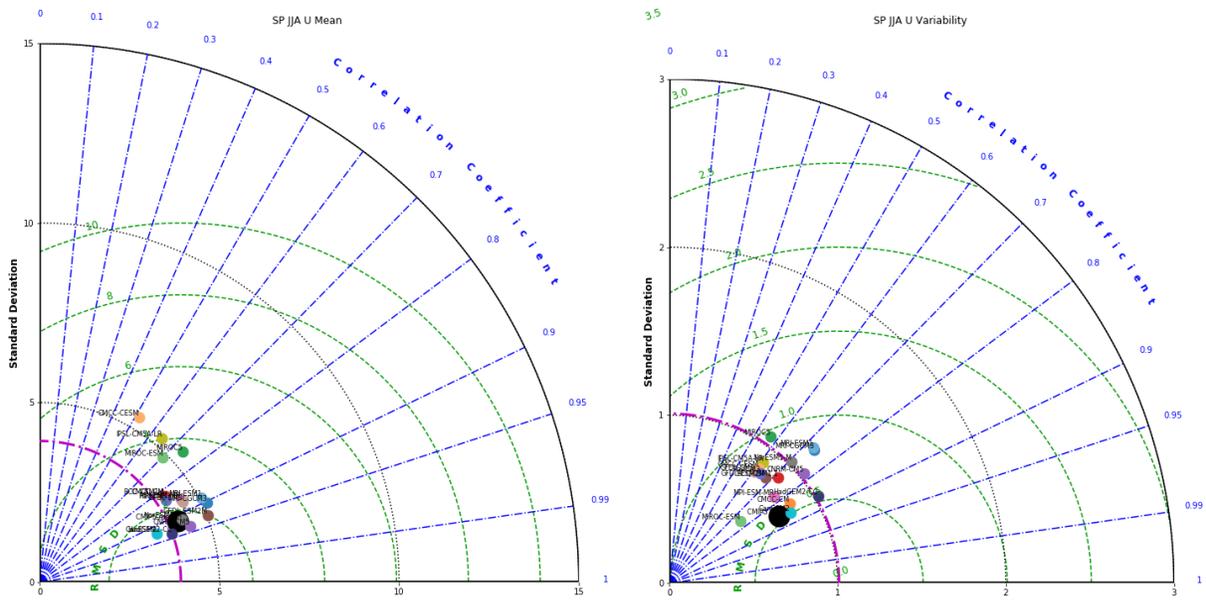
Figure C.40: As for Figure C.31, with SP JJA.
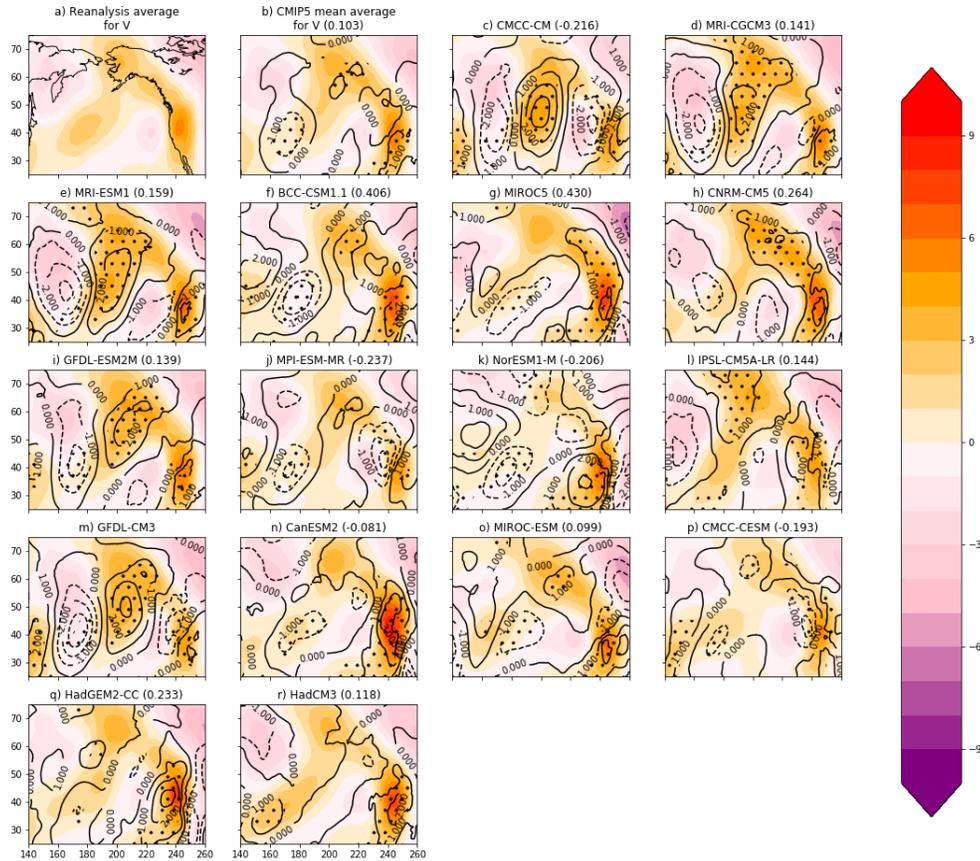
Figure C.41: 500 mb meridional wind for NP JJA, with the reanalysis mean (top left), the CMIP5 model mean meridional wind, and the individual model meridional wind averages as filled contours with 2 m/s interval spacing. For each model, the differences from the reanalysis average are denoted by black contours in 1 m/s interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
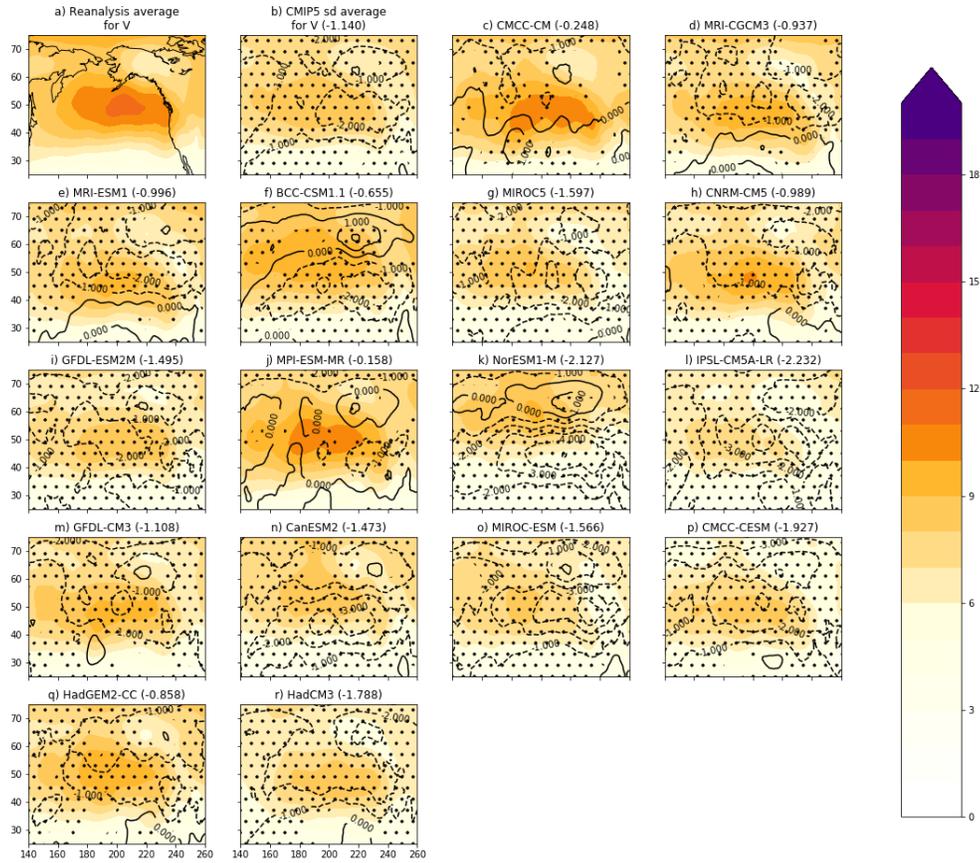
Figure C.42: 500 mb meridional wind variability for NP JJA, with the reanalysis mean standard deviation (top left), the CMIP5 model mean standard deviation of meridional wind, and the individual model meridional wind standard deviations as filled contours with 1 m/s interval spacing. For each model, the differences from the reanalysis average are denoted by black contours in 1 m/s interval spacing, with negative differences represented by dashed lines. Regions with significant differences are denoted by stippling.
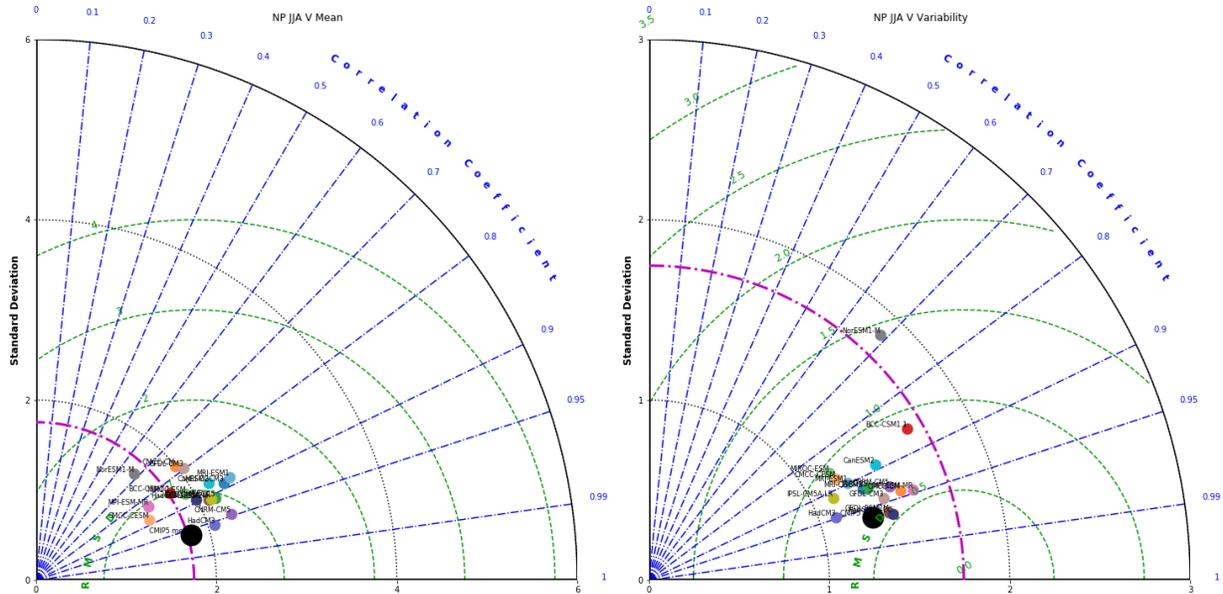
Figure C.43: Taylor plot of (left) the seasonal mean V field and (right) the temporal variability in the seasonal V field (measured as the standard deviation of the per-gridpoint time series). The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. The large black dot signifies the CMIP5 (excluding the Hadley models) average. Note that the maximum extent of the standard deviation axis is 6 for the mean plot while it is 3 for the variability plot.

Figure C.44: As for Figure C.41, with NP DJF
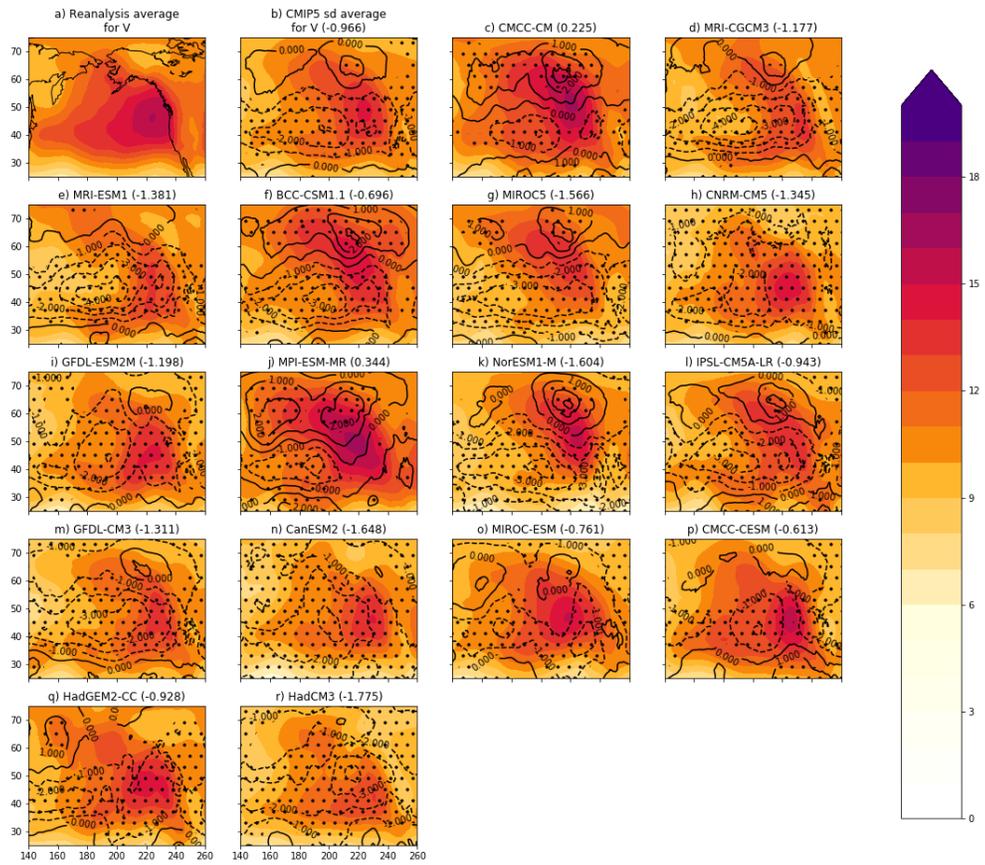
Figure C.45: As for Figure C.42, with NP DJF
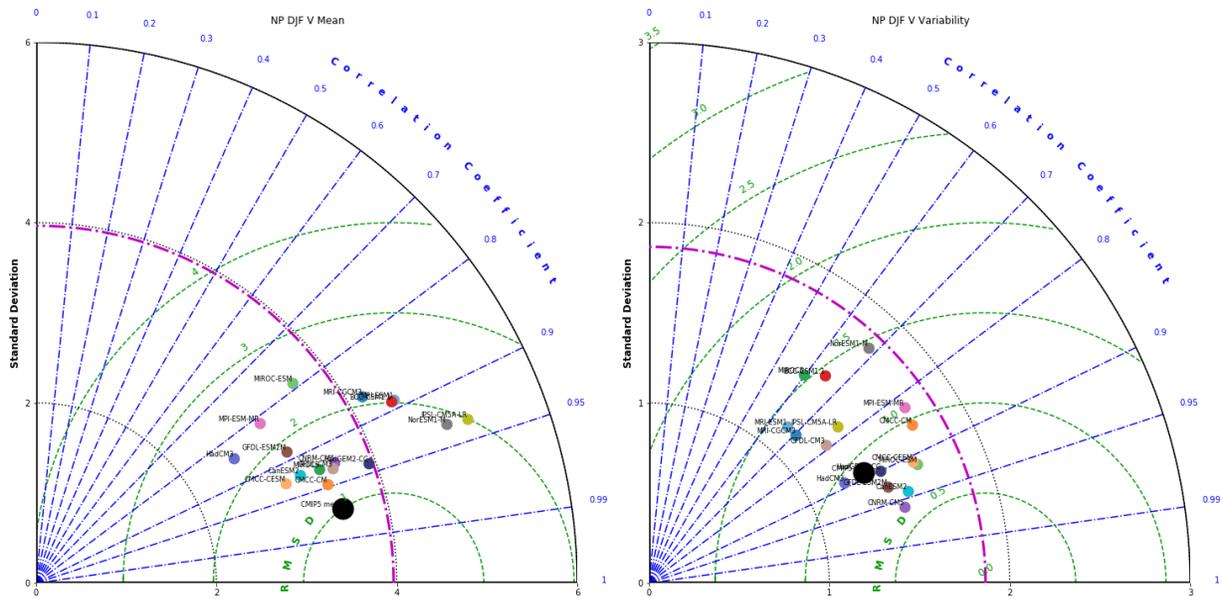
Figure C.46: As for Figure C.43, with NP DJF.

Figure C.47: As for Figure C.41, with SP DJF
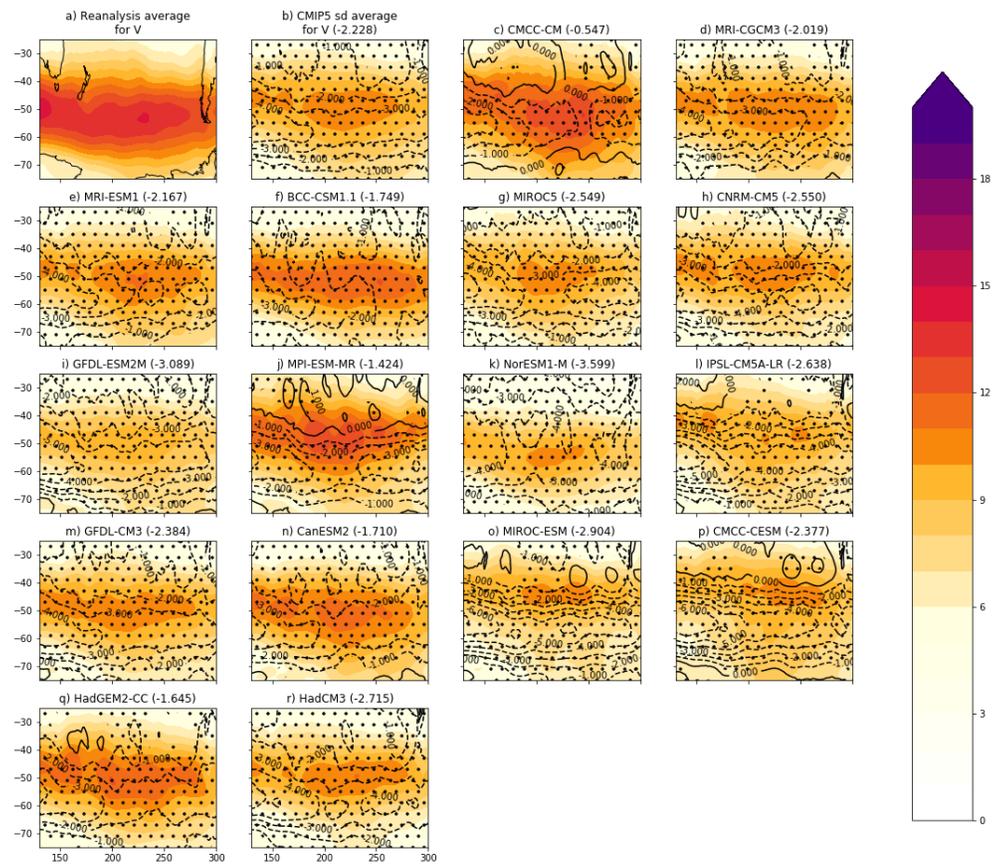
Figure C.48: As for Figure C.42, with SP DJF
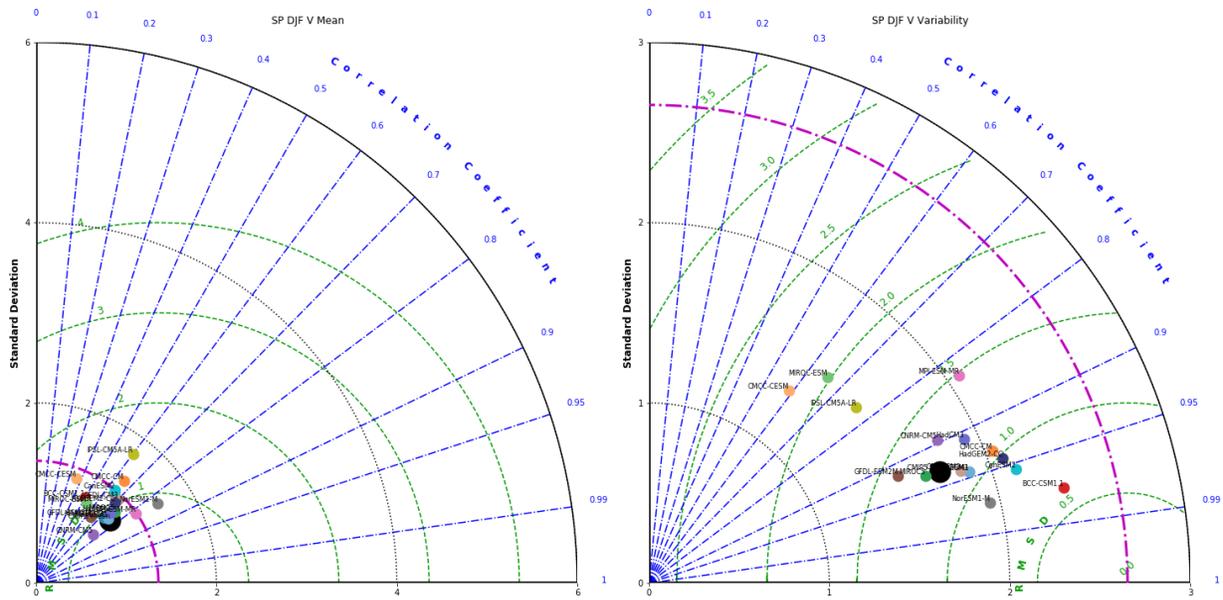
Figure C.49: As for Figure C.43, with SP DJF.

Figure C.50: As for Figure C.41, with SP JJA
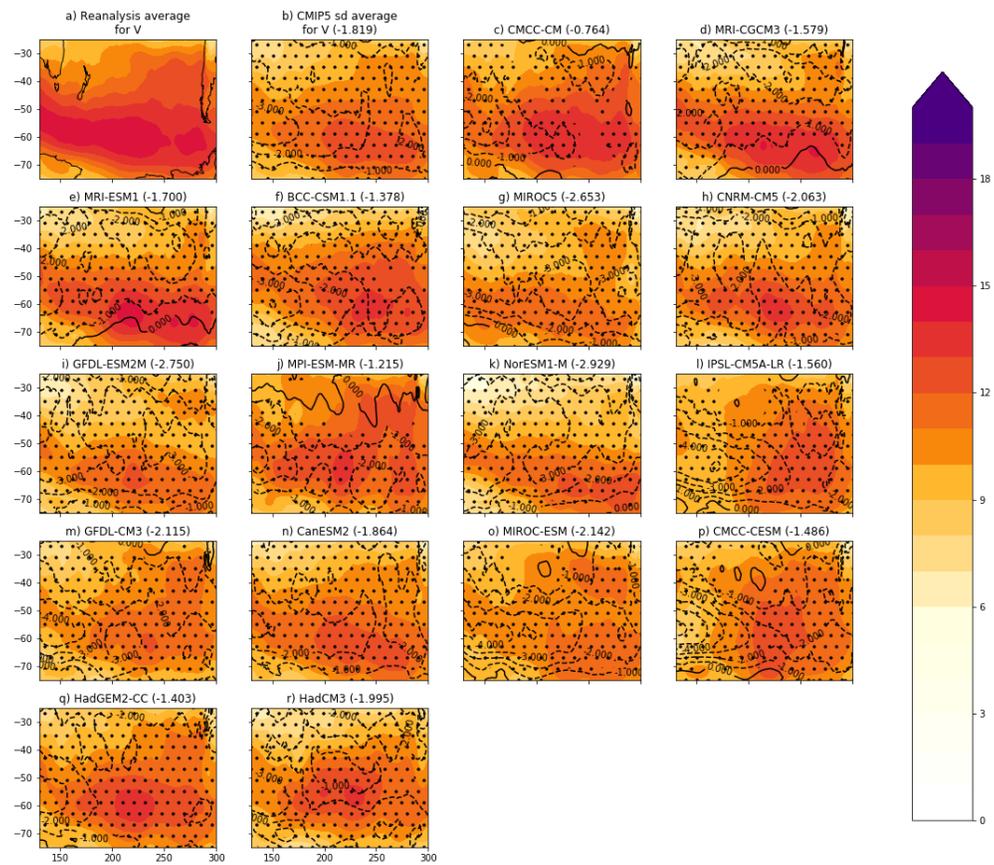
Figure C.51: As for Figure C.42, with SP JJA

Figure C.52: As for Figure C.43, with SP JJA.

Appendix D

Chapter 4 Supplemental Figures
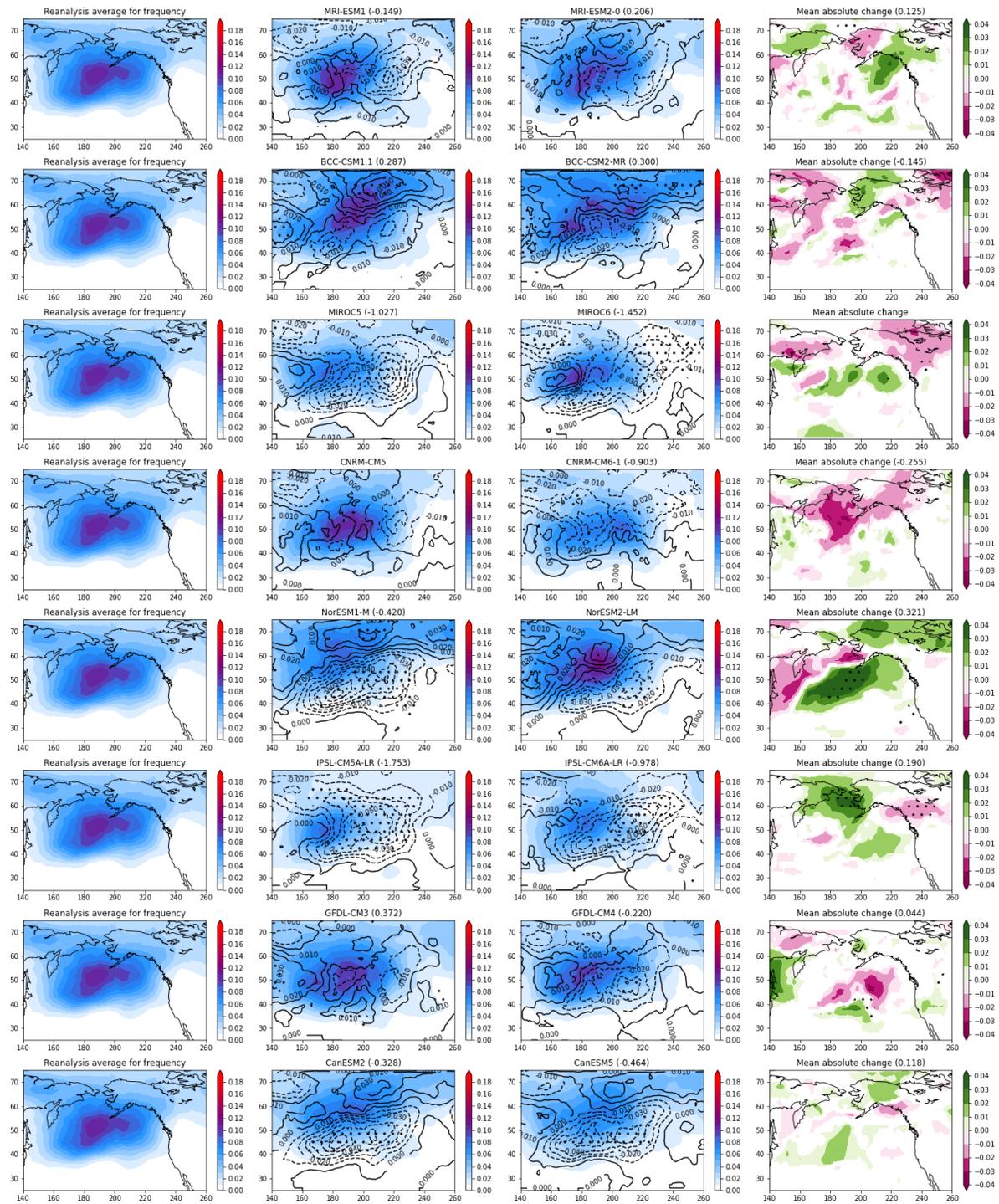
Figure D.1: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for blocking frequency in NP JJA.
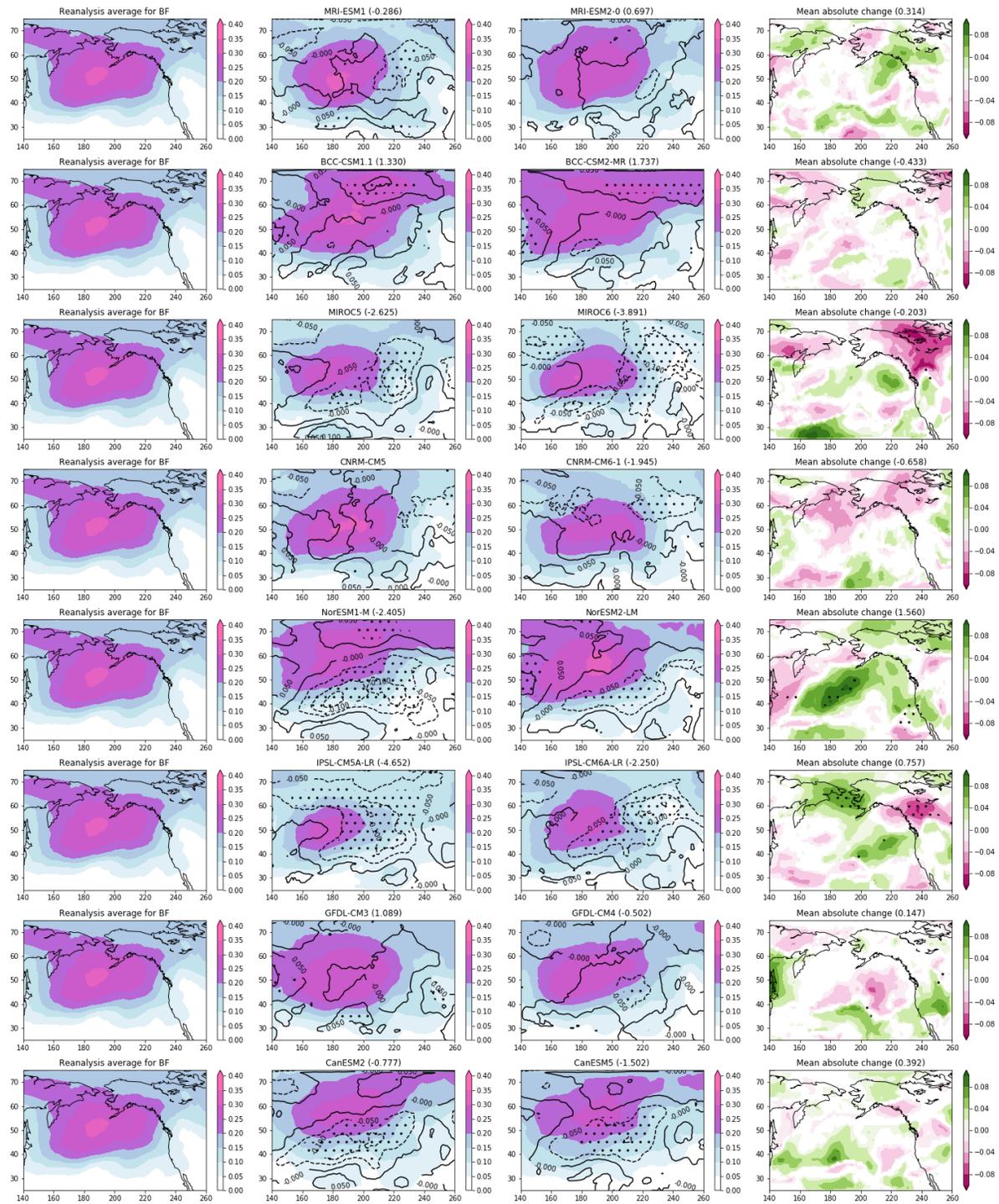
Figure D.2: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for blocking frequency variability in NP JJA.
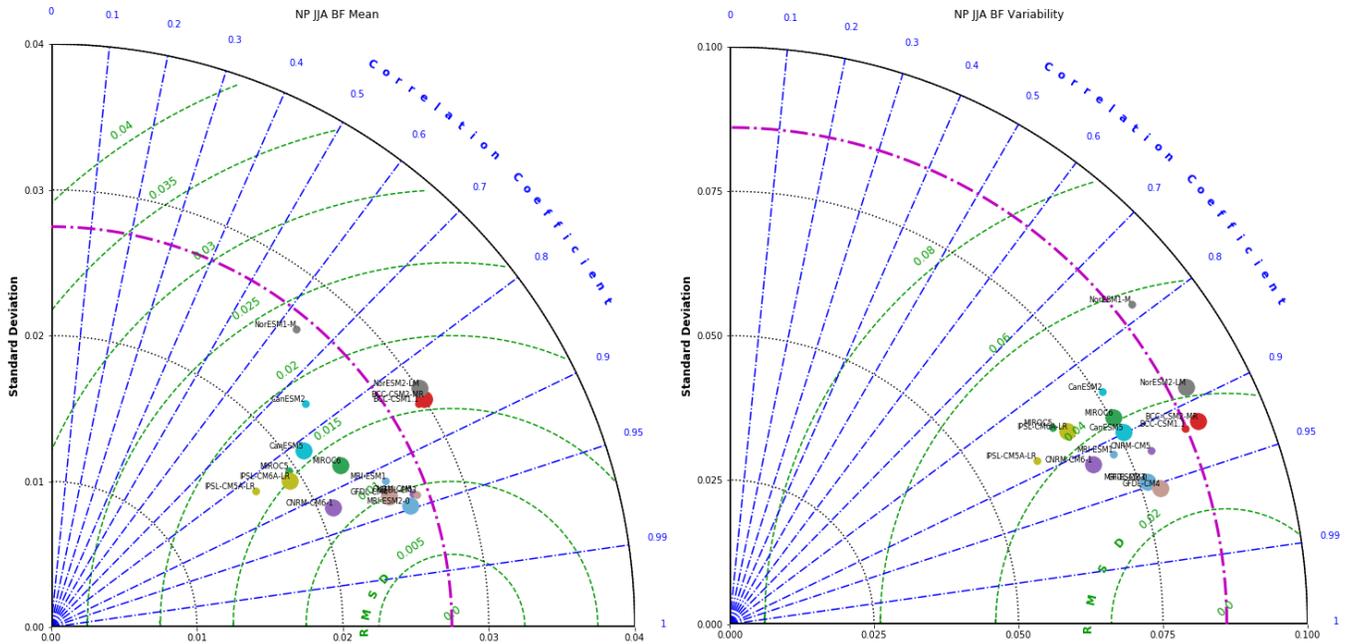
Figure D.3: Taylor plot of (left) the seasonal mean BF field and (right) the temporal variability in the seasonal BF field (measured as the standard deviation of the per-gridpoint time series) for CMIP5 (smaller points) and CMIP6 (larger points) models. The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. Note that the maximum extent of the standard deviation axis is 0.06 for the mean plot while it is 0.10 for the variability plot.

178

Figure D.4: As for Figure D.1, with NP DJF
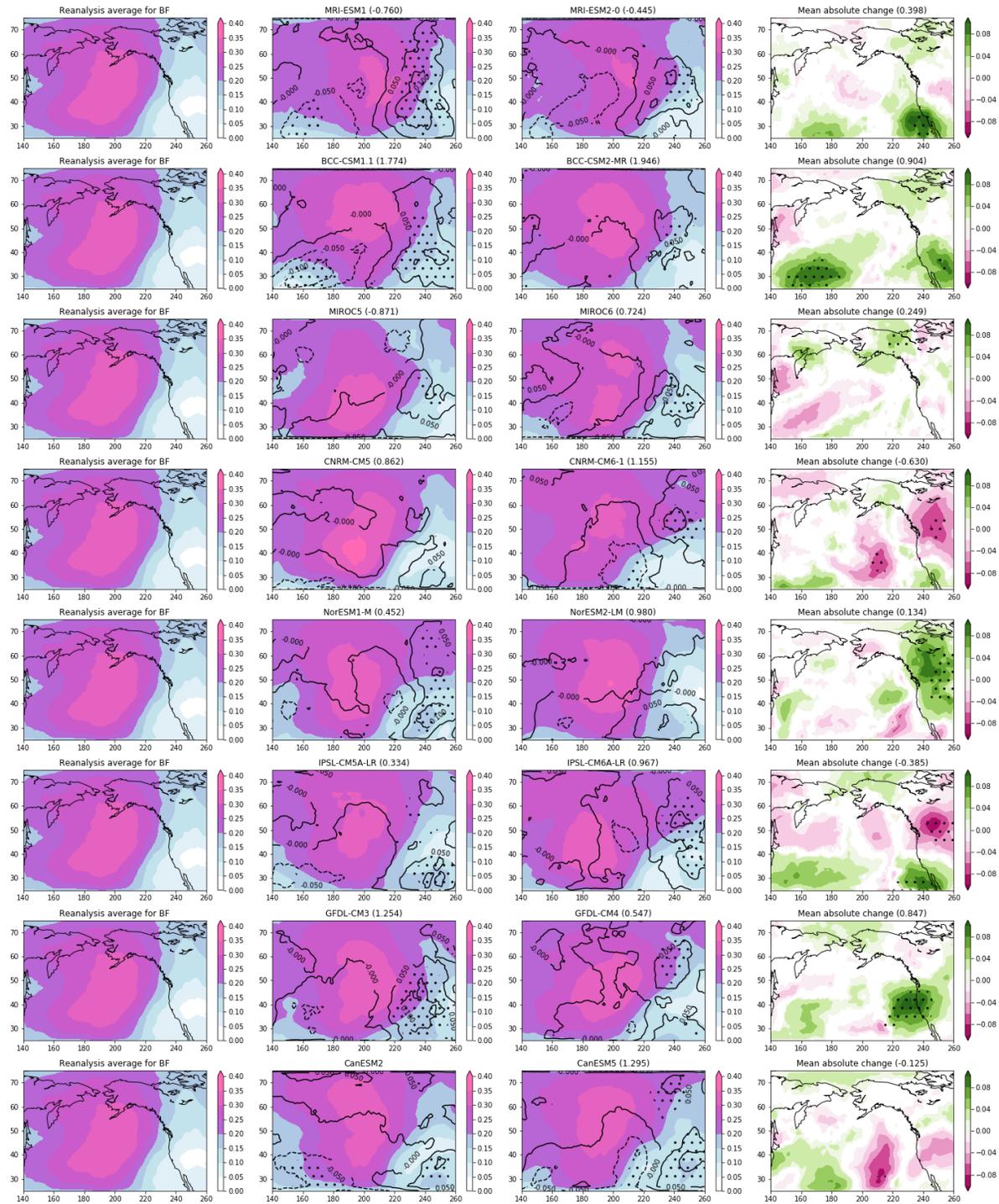
Figure D.5: As for Figure D.2, with NP DJF

Figure D.6: As for Figure D.3, with NP DJF.
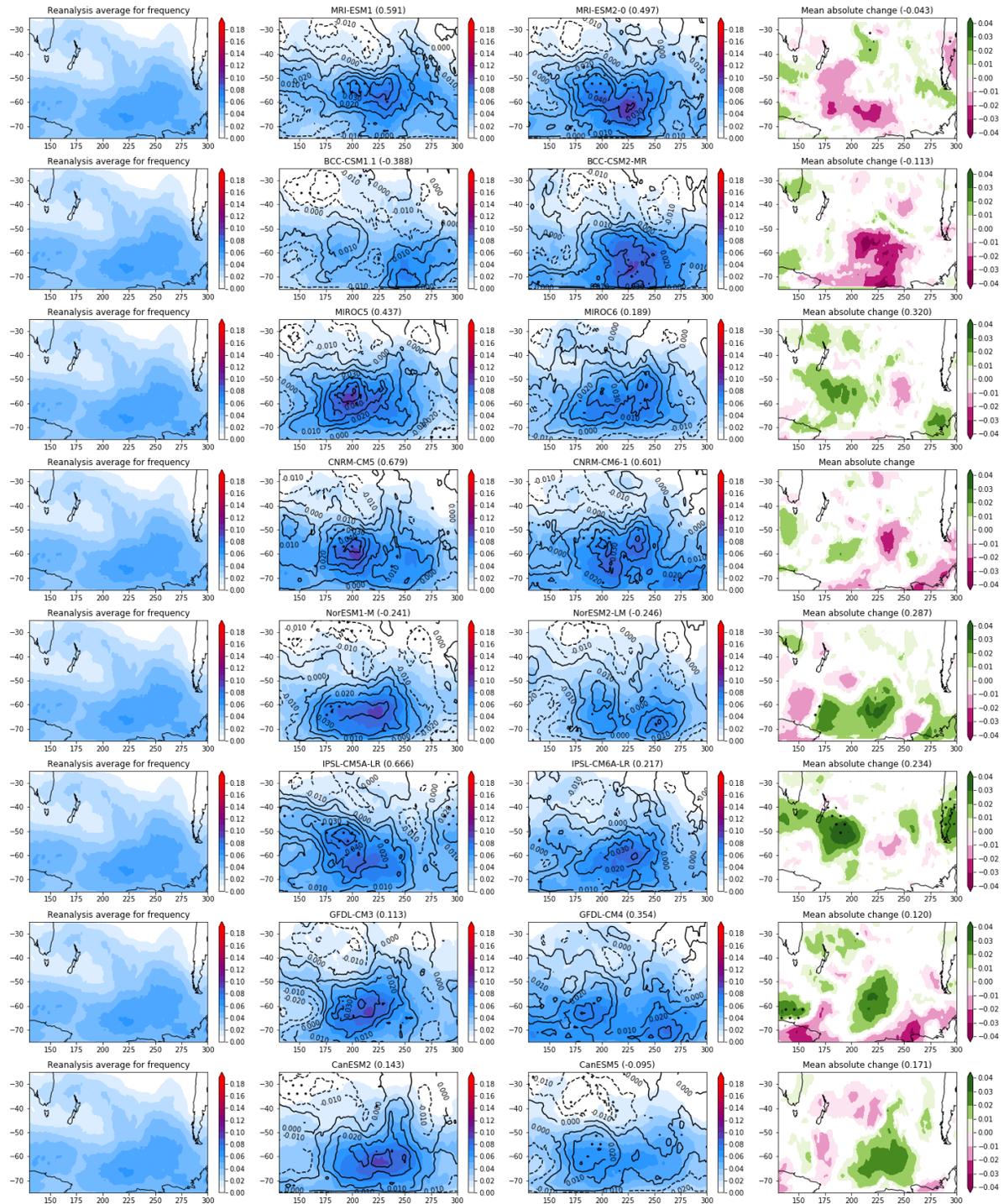
Figure D.7: As for Figure D.1, with SP DJF
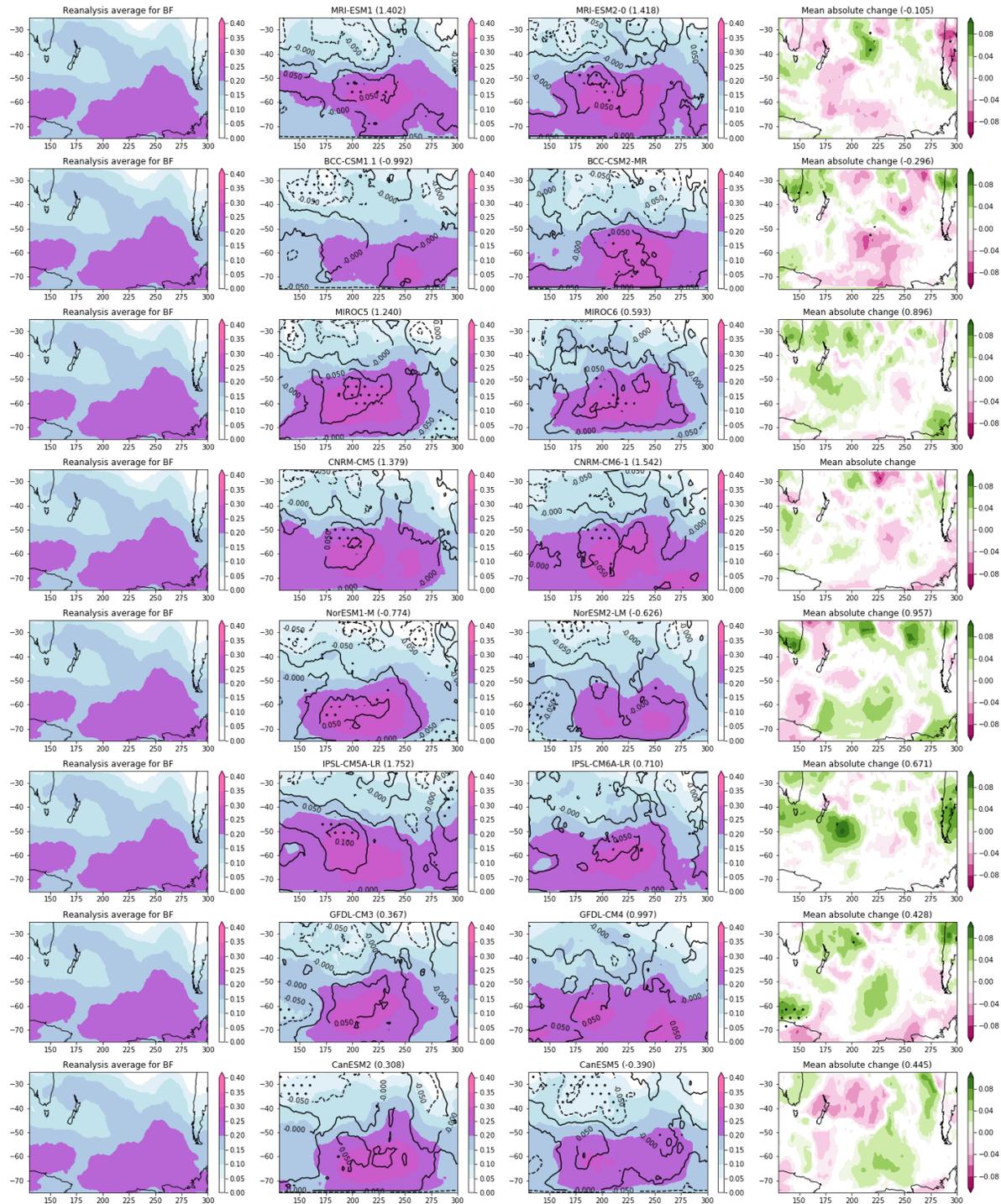
Figure D.8: As for Figure D.2, with SP DJF

Figure D.9: As for Figure D.3, with SP DJF.
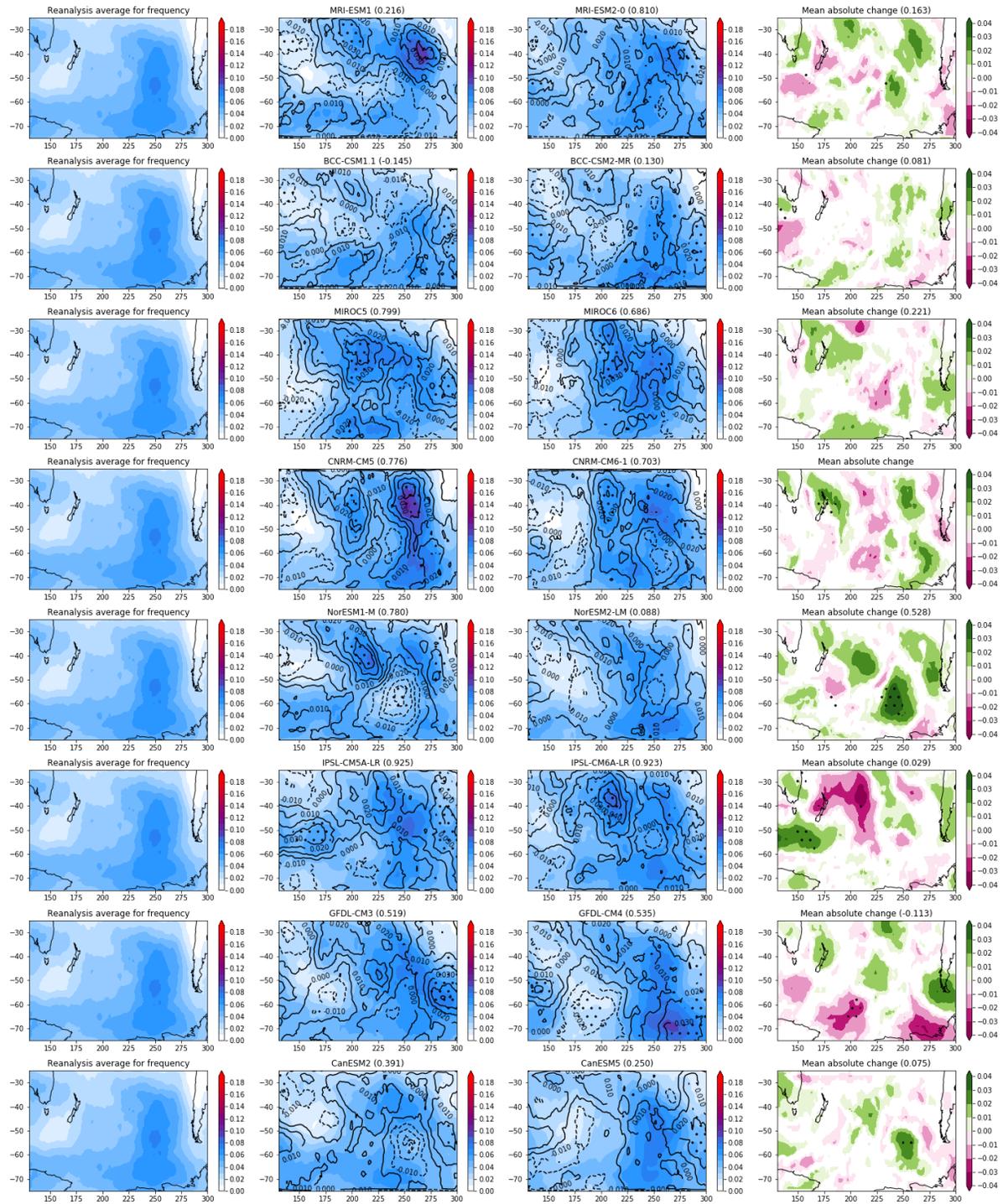
Figure D.10: As for Figure D.1, with SP JJA
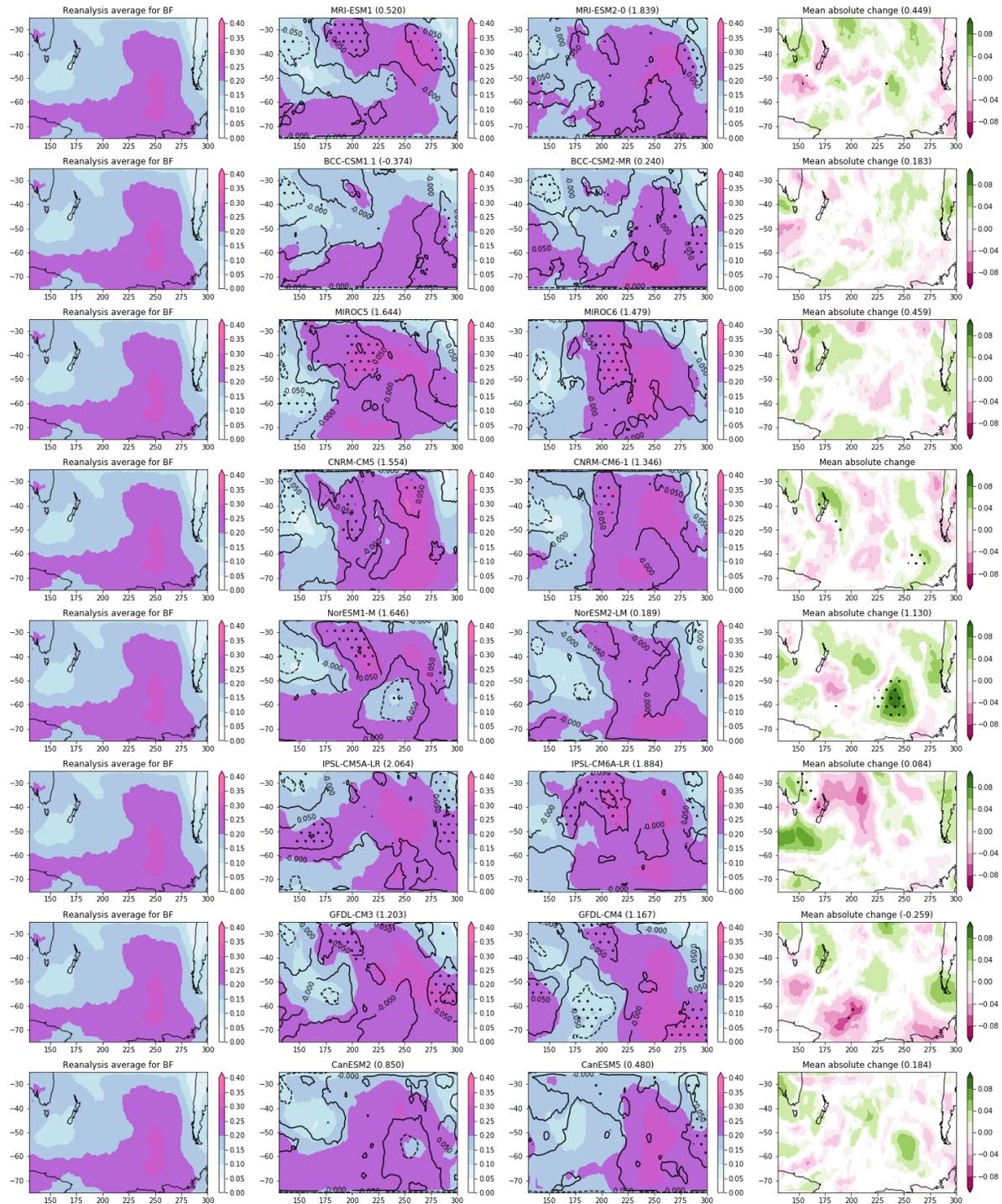
Figure D.11: As for Figure D.2, with SP JJA
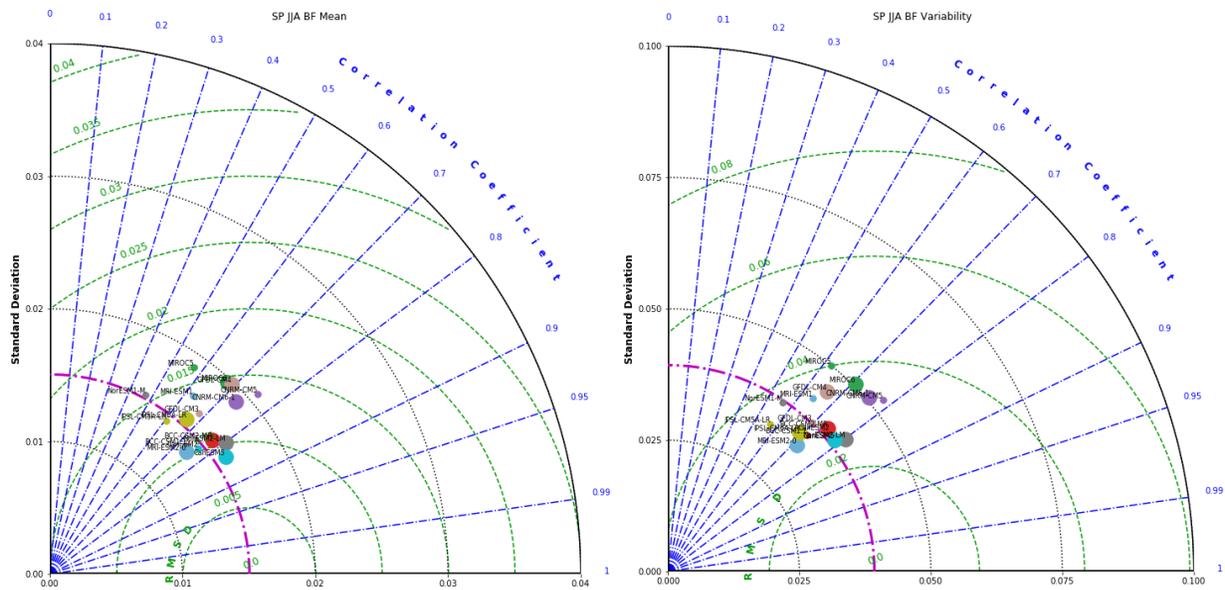
Figure D.12: As for Figure D.3, with SP JJA.
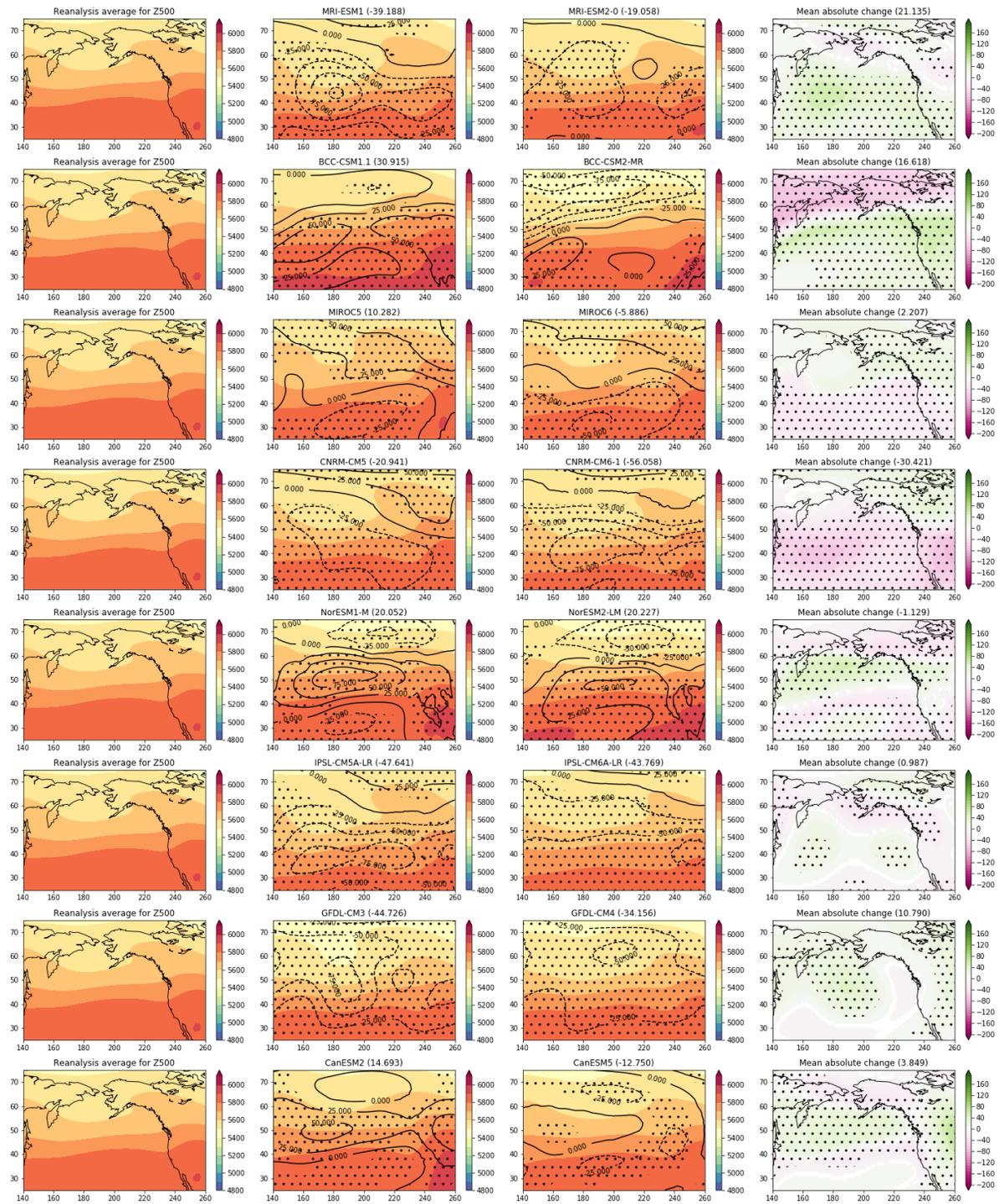
Figure D.13: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for Z500 in NP JJA.
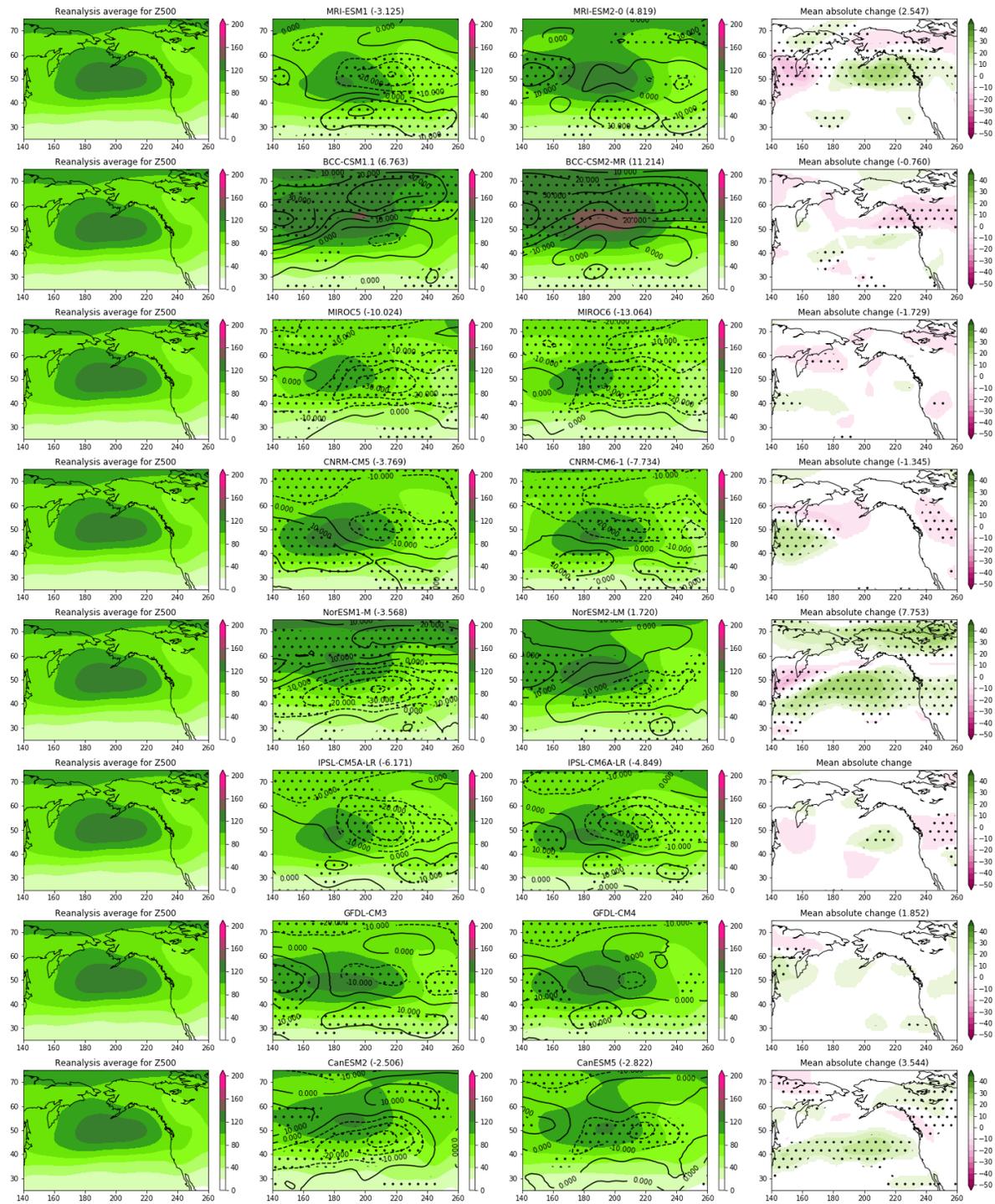
Figure D.14: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for Z500 variability in NP JJA.
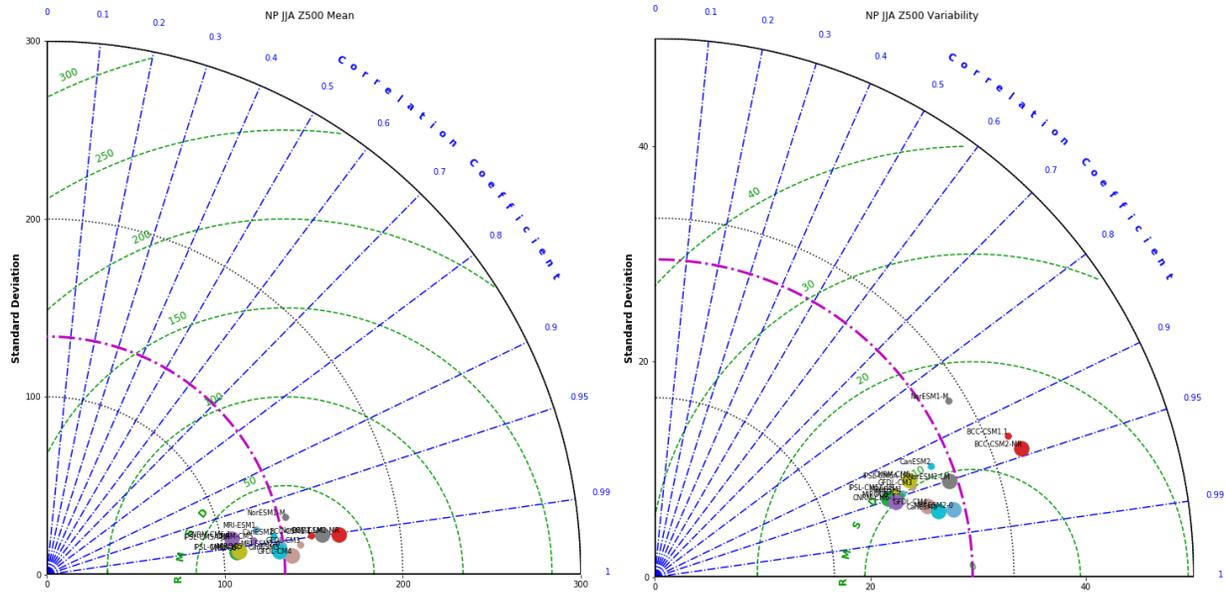
Figure D.15: Taylor plot of (left) the seasonal mean BF field and (right) the temporal variability in the seasonal Z500 field (measured as the standard deviation of the per-gridpoint time series) for CMIP5 (smaller points) and CMIP6 (larger points) models. The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. Note that the maximum extent of the standard deviation axis is 300 for the mean plot while it is 50 for the variability plot.

Figure D.16: As for Figure D.13, with NP DJF
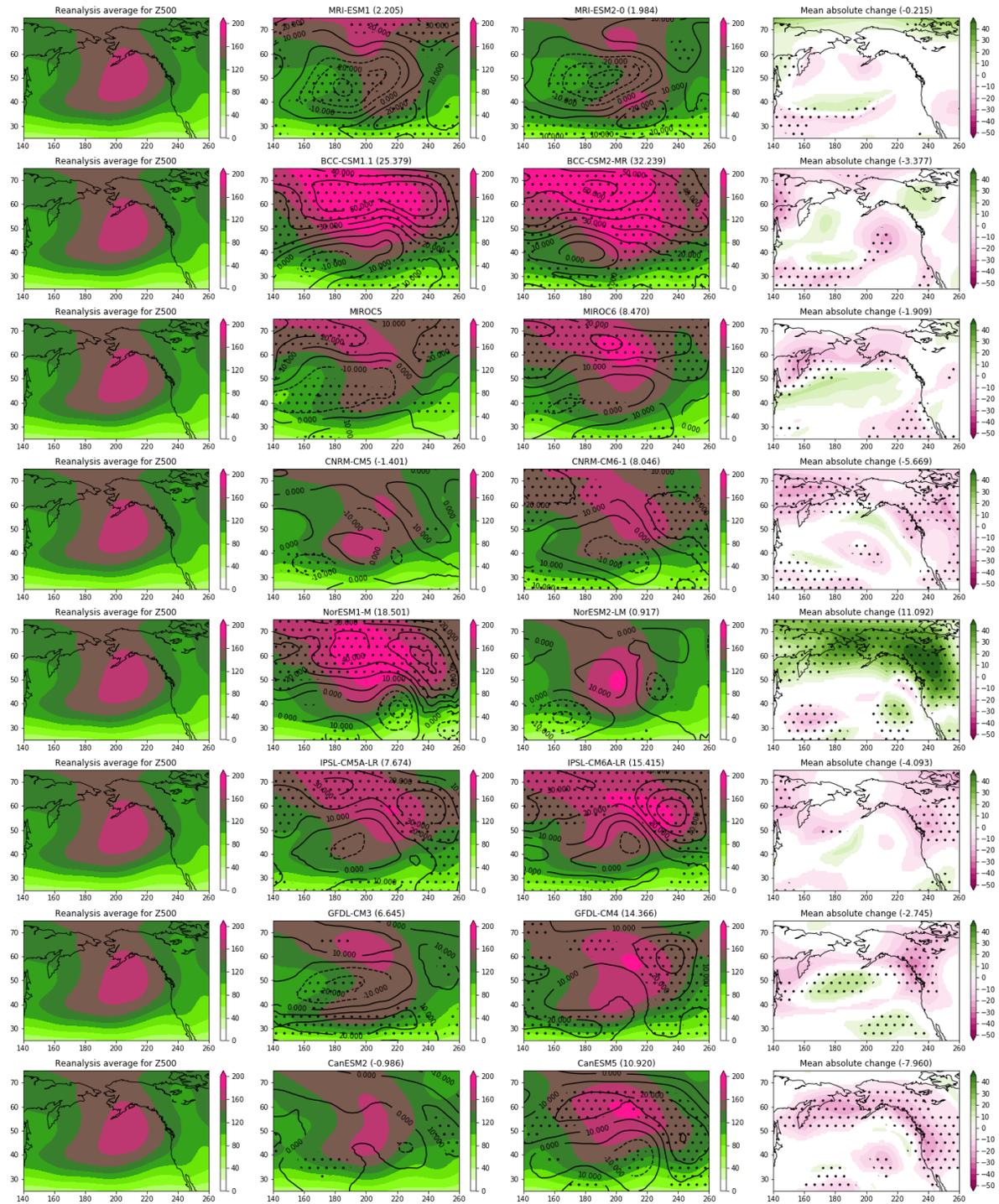
Figure D.17: As for Figure D.14, with NP DJF

Figure D.18: As for Figure D.15, with NP DJF.
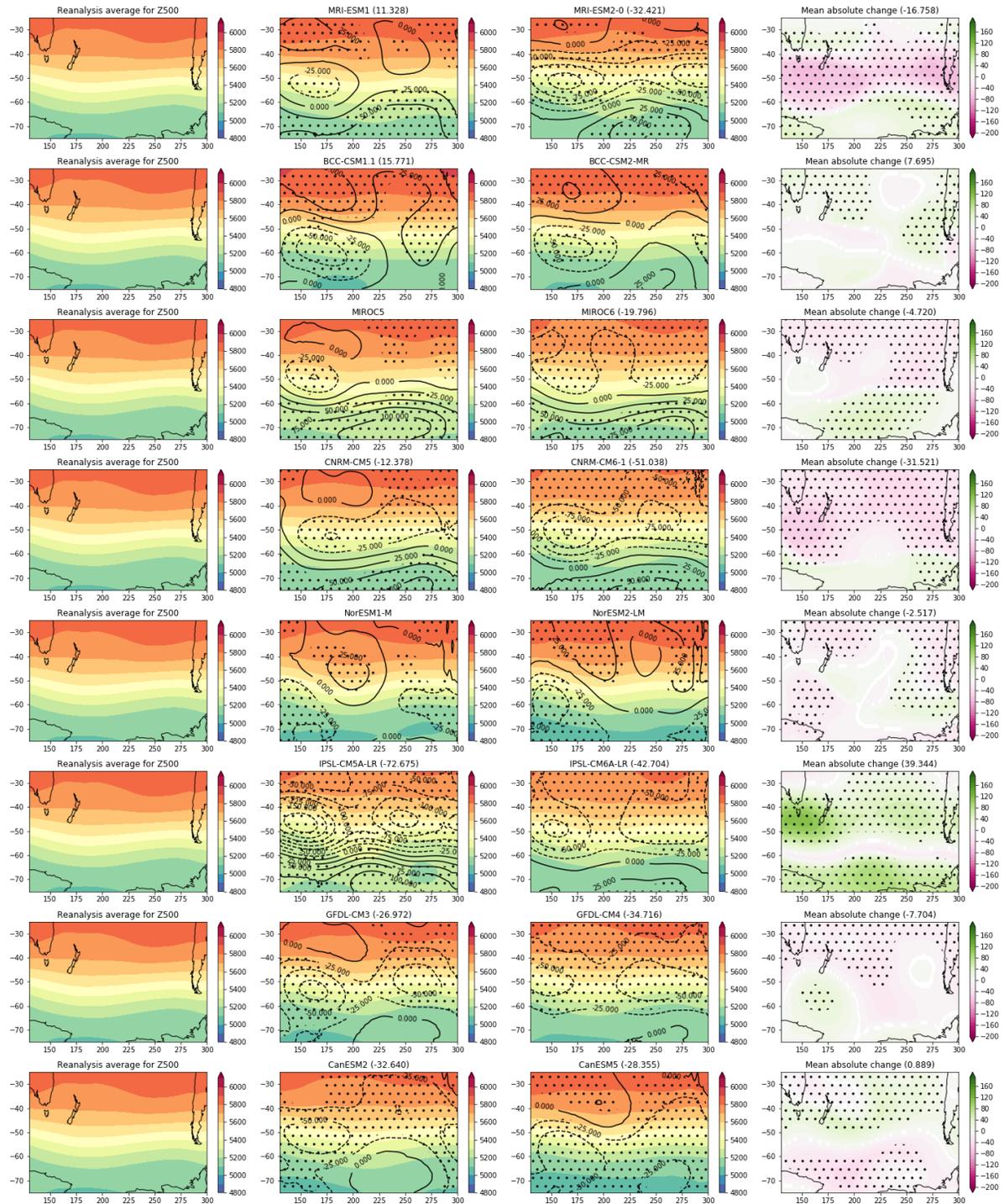
Figure D.19: As for Figure D.13, with SP DJF
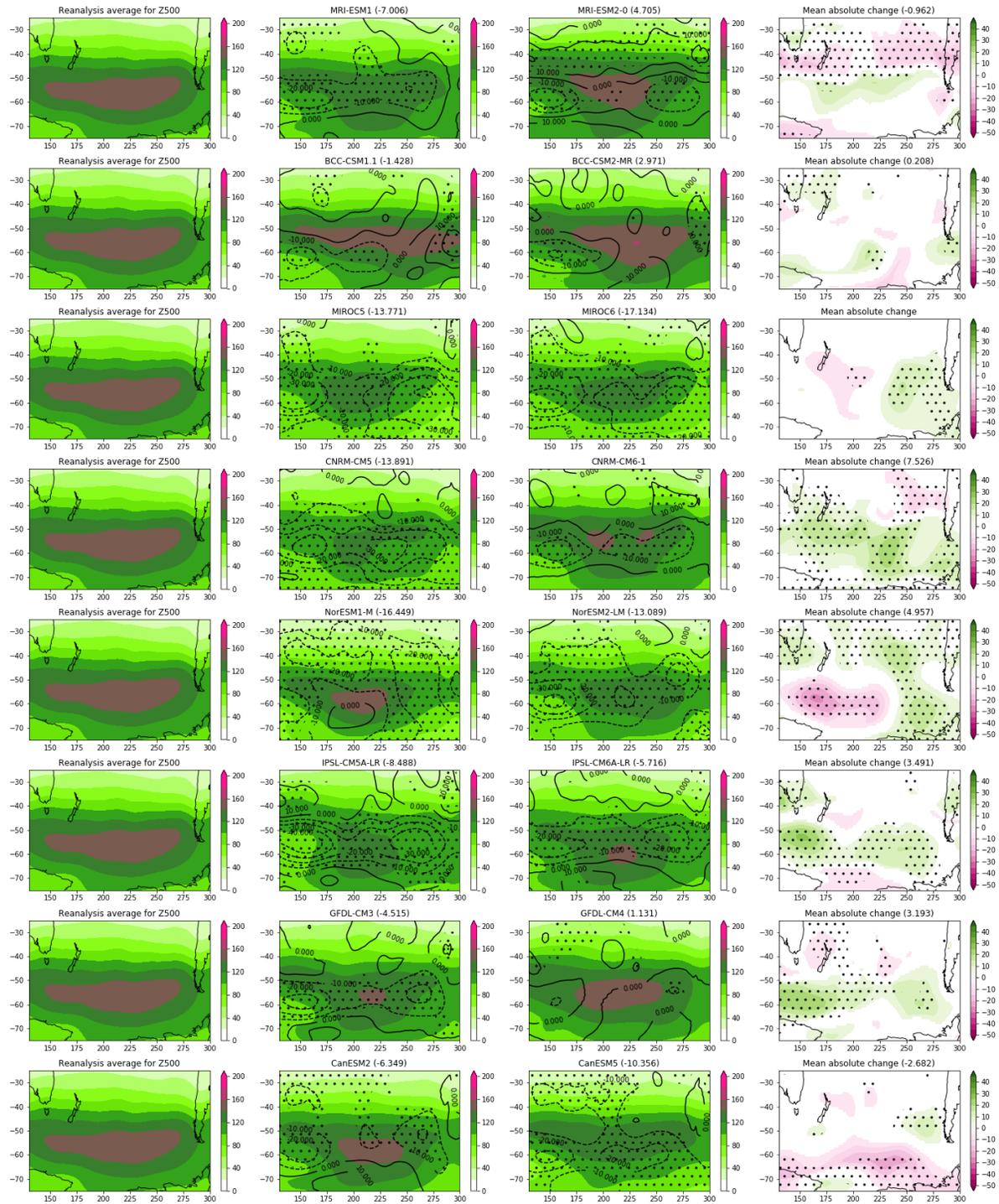
Figure D.20: As for Figure D.14, with SP DJF
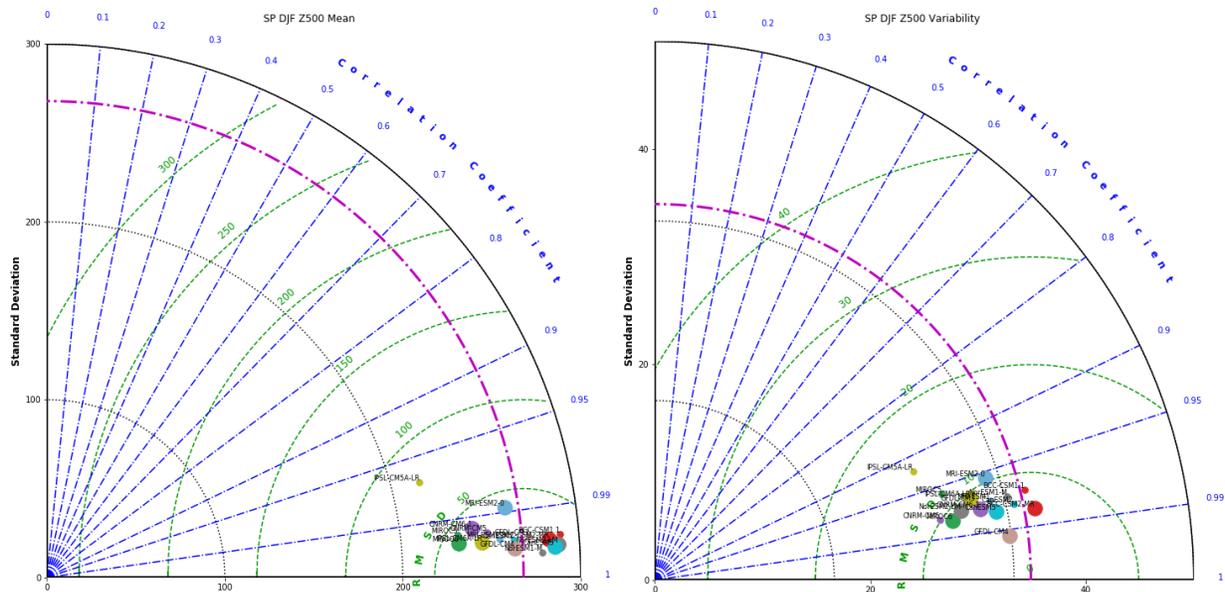
Figure D.21: As for Figure D.15, with SP DJF.
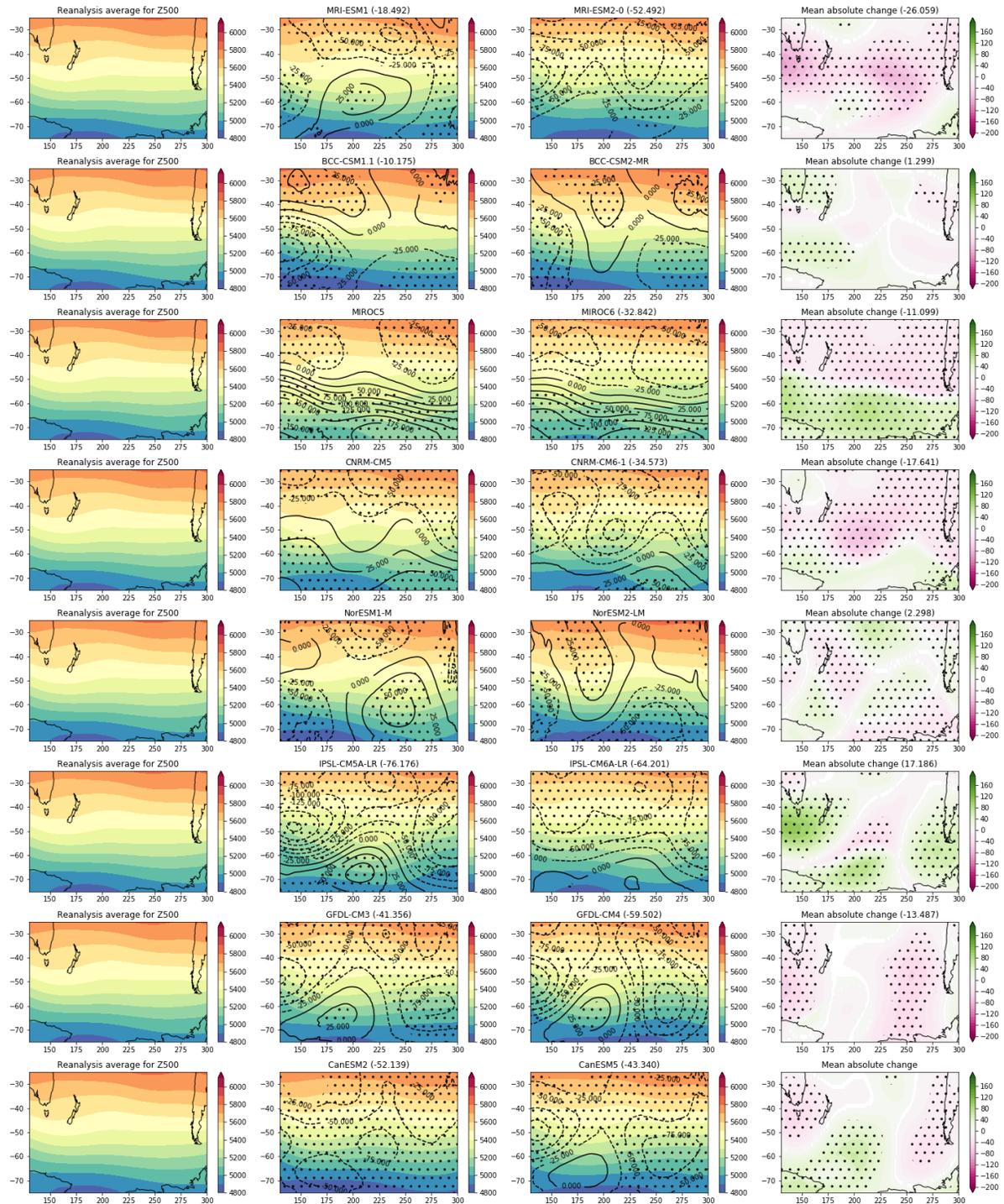
Figure D.22: As for Figure D.13, with SP JJA
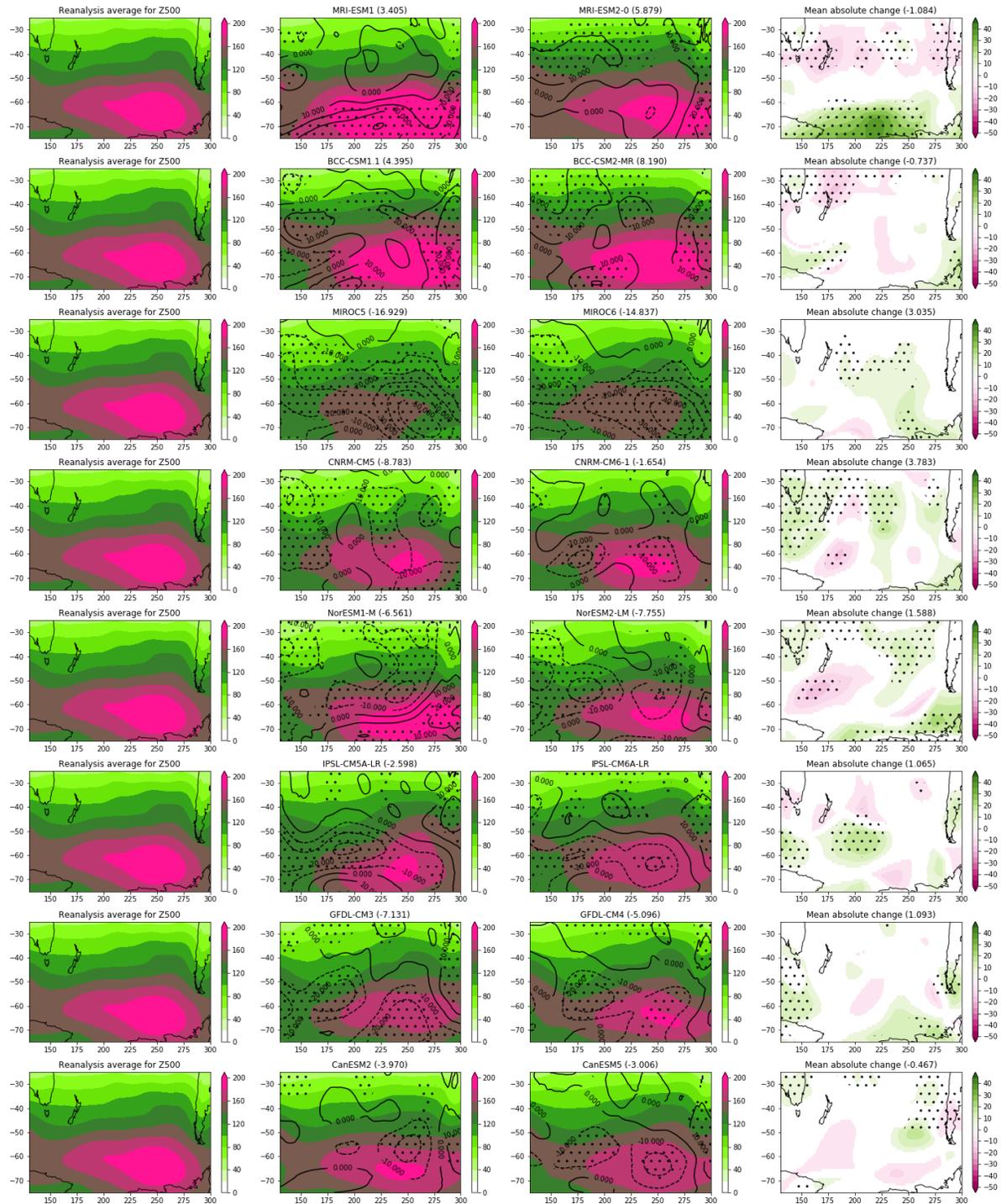
Figure D.23: As for Figure D.14, with SP JJA
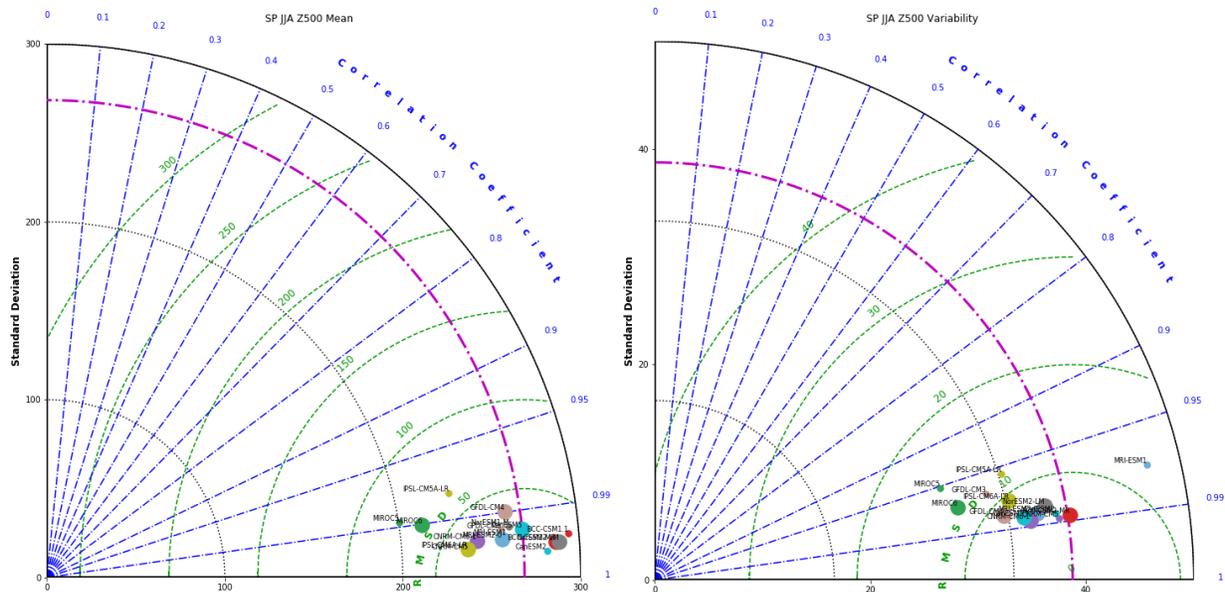
Figure D.24: As for Figure D.15, with SP JJA.
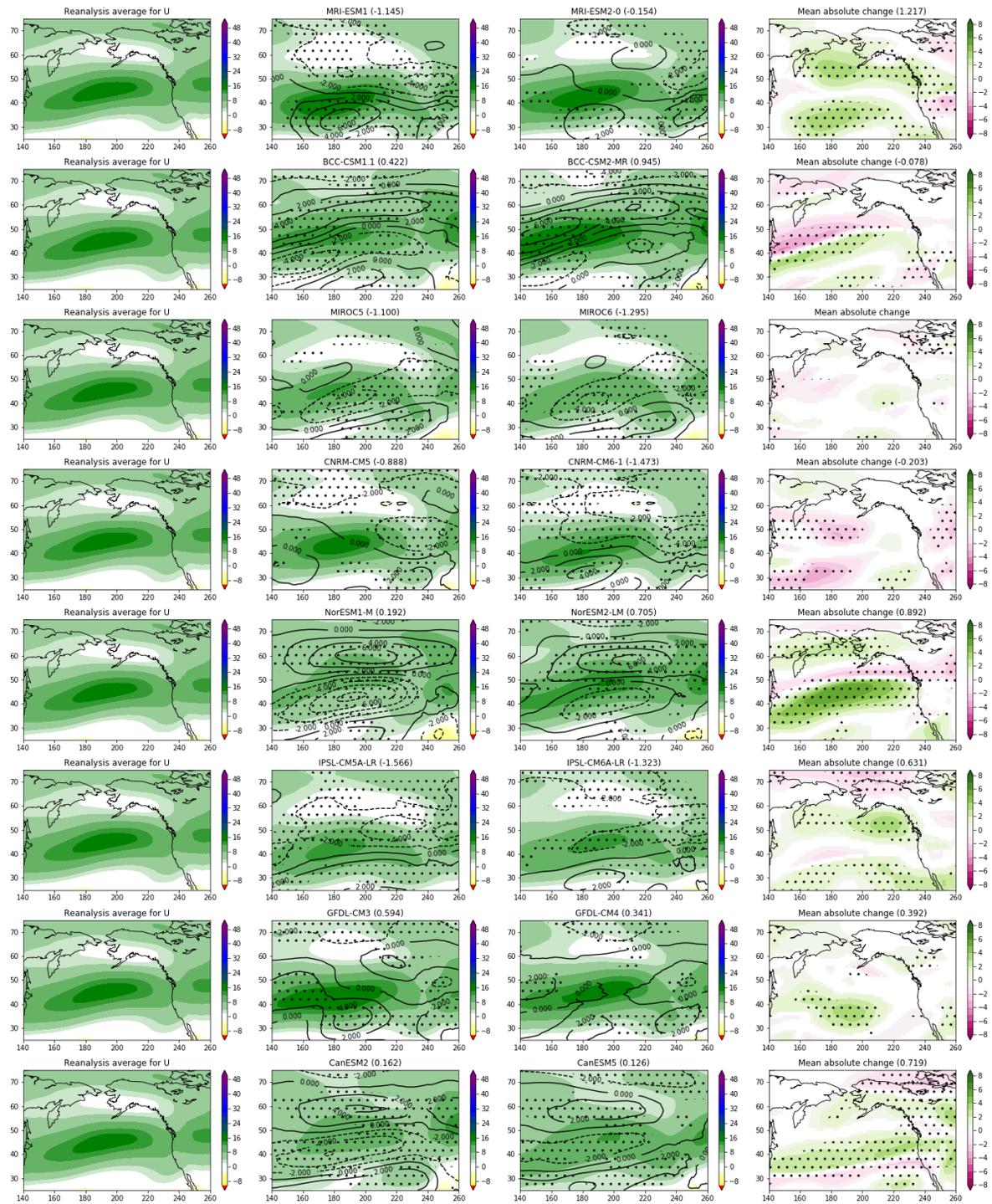
Figure D.25: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for U in NP JJA.
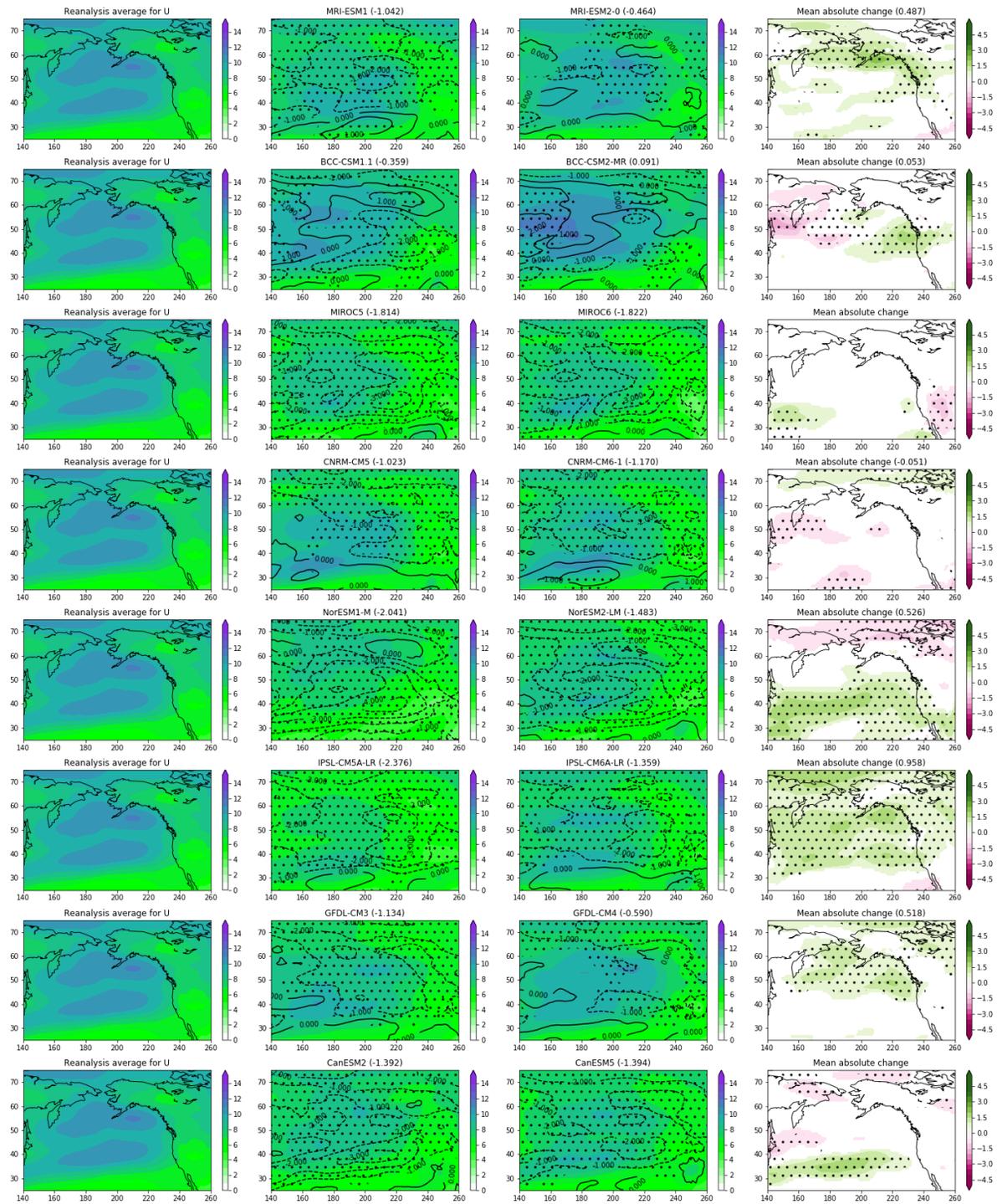
Figure D.26: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for U variability in NP JJA.
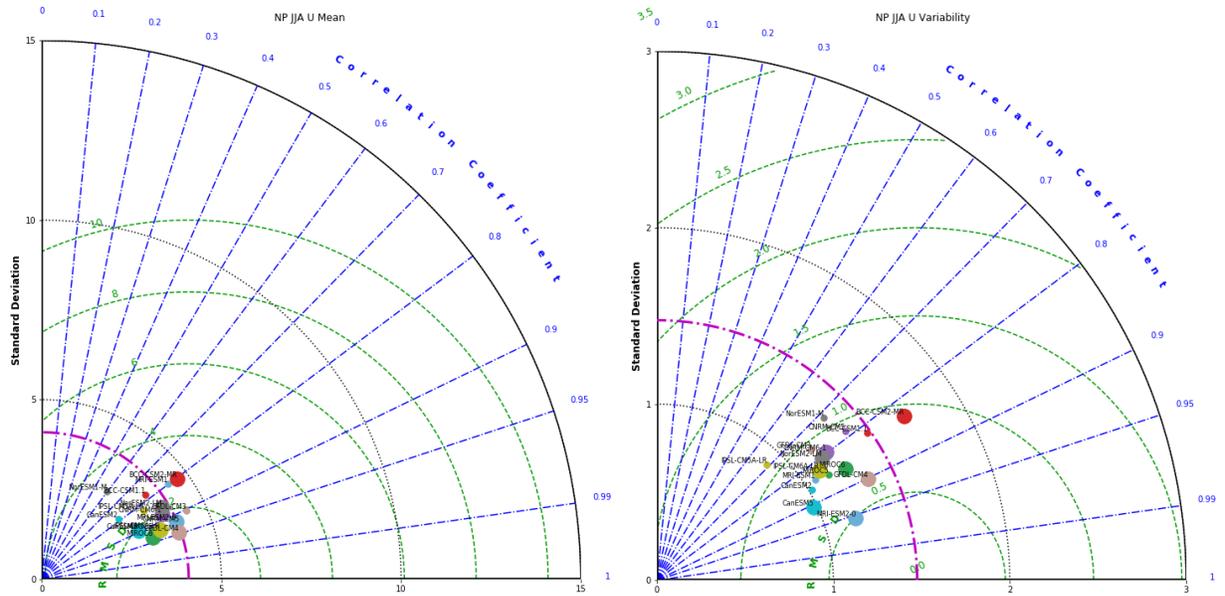
Figure D.27: Taylor plot of (left) the seasonal mean BF field and (right) the temporal variability in the seasonal U field (measured as the standard deviation of the per-gridpoint time series) for CMIP5 (smaller points) and CMIP6 (larger points) models. The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. Note that the maximum extent of the standard deviation axis is 15 for the mean plot while it is 3 for the variability plot.

Figure D.28: As for Figure D.25, with NP DJF
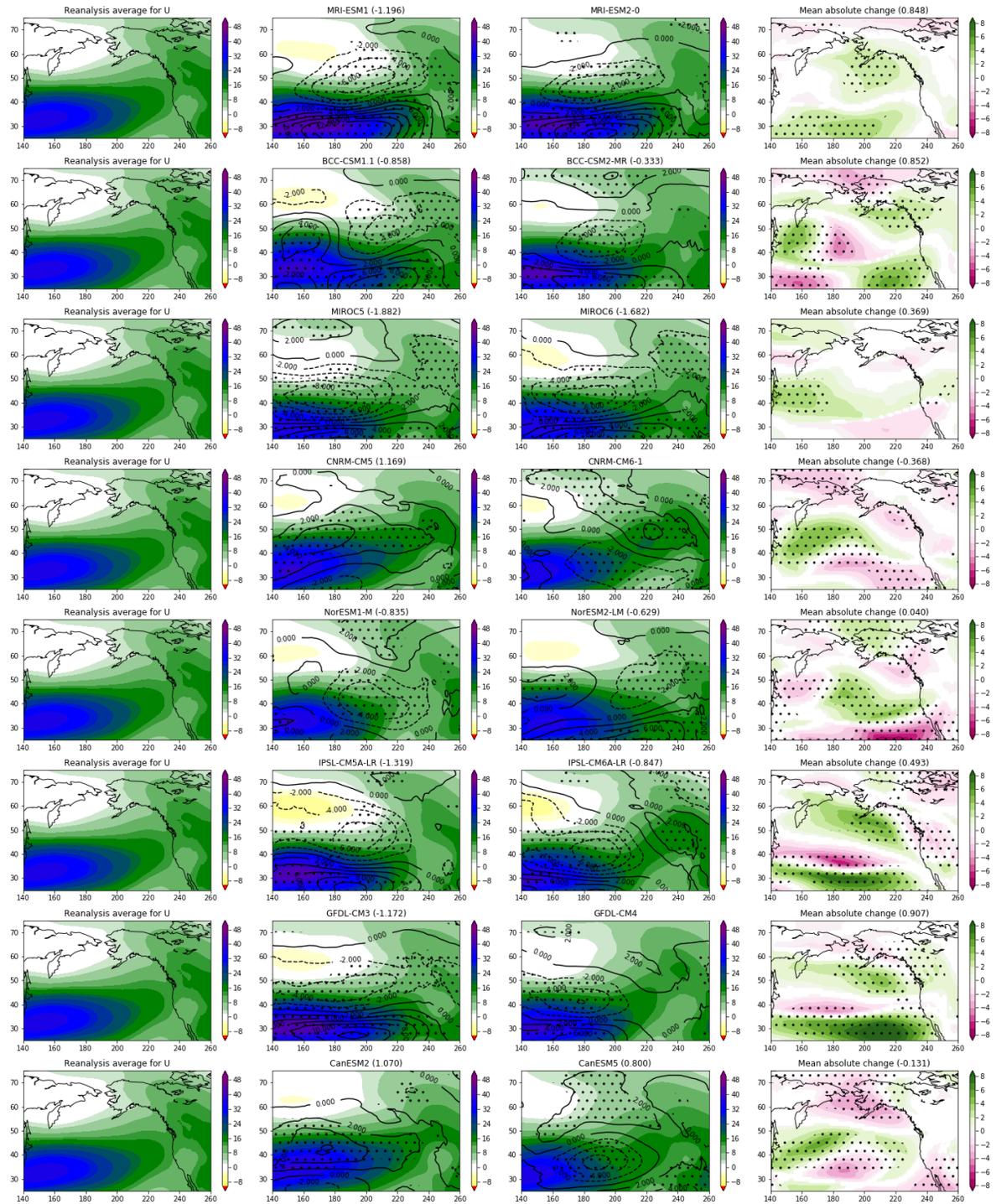
Figure D.29: As for Figure D.26, with NP DJF
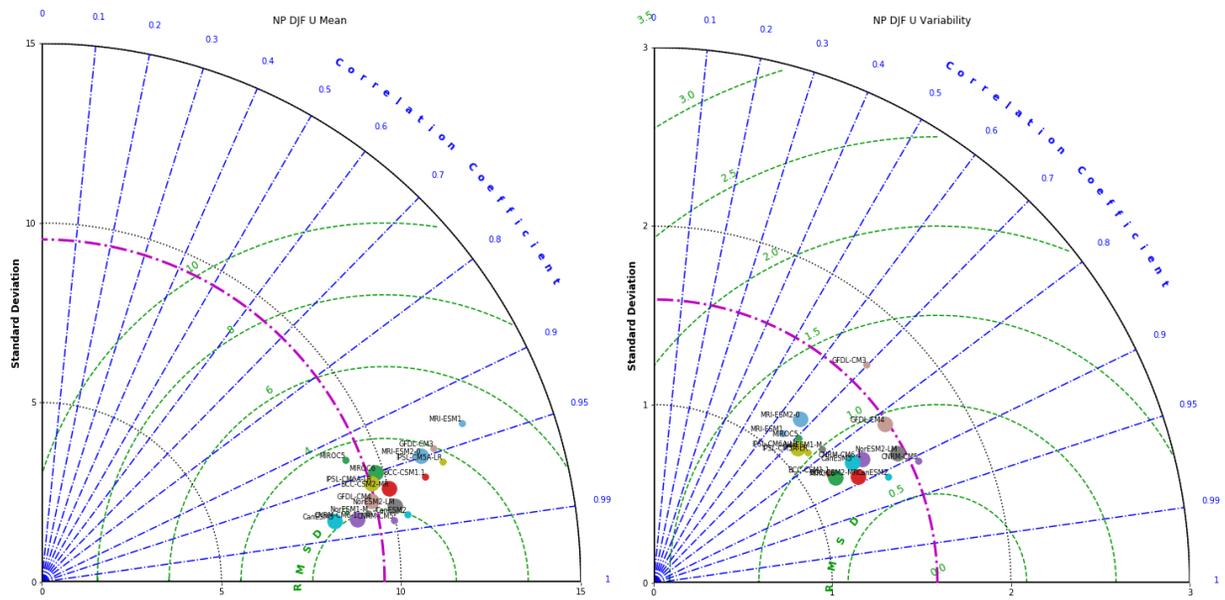
Figure D.30: As for Figure D.27, with NP DJF.
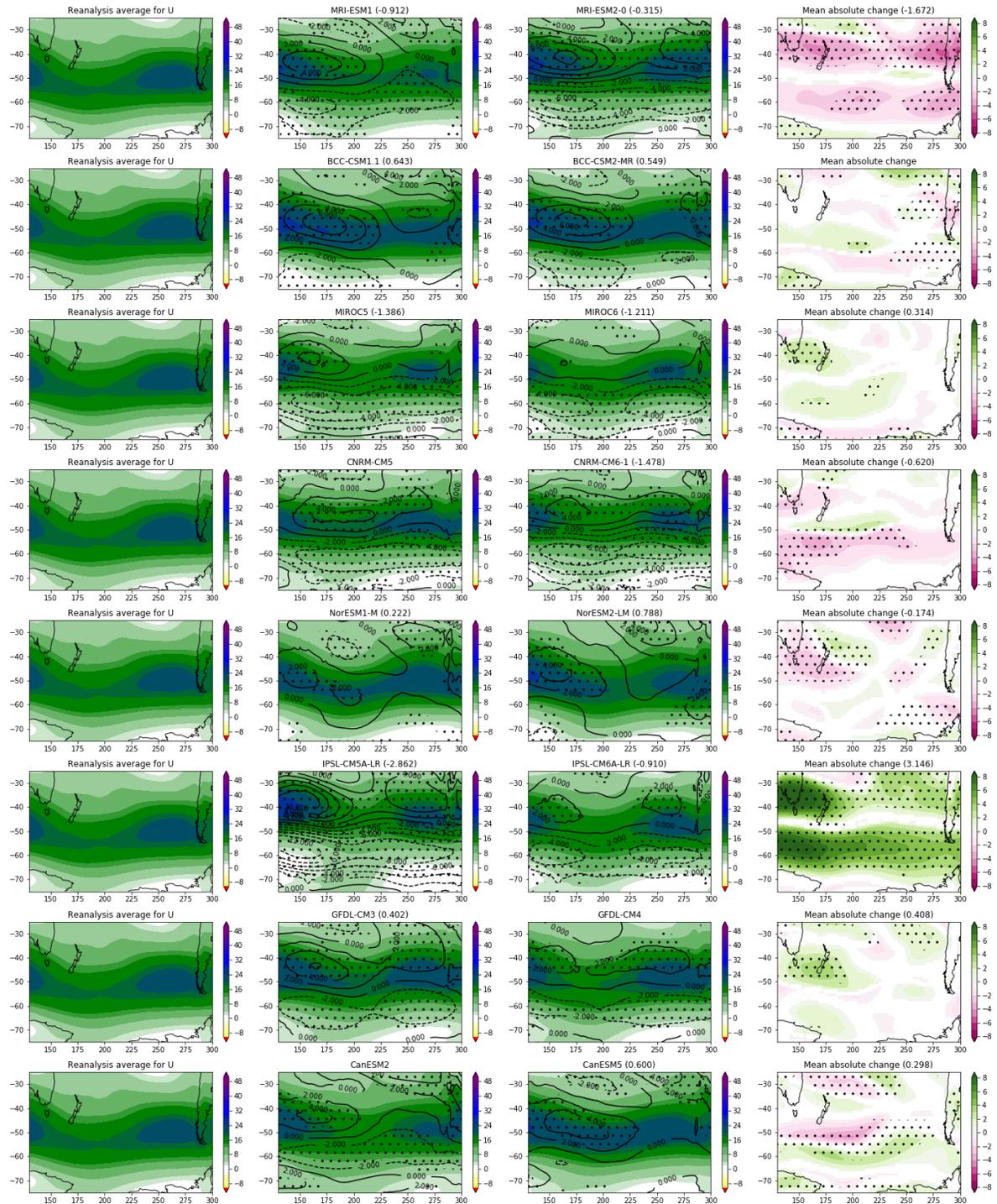
Figure D.31: As for Figure D.25, with SP DJF
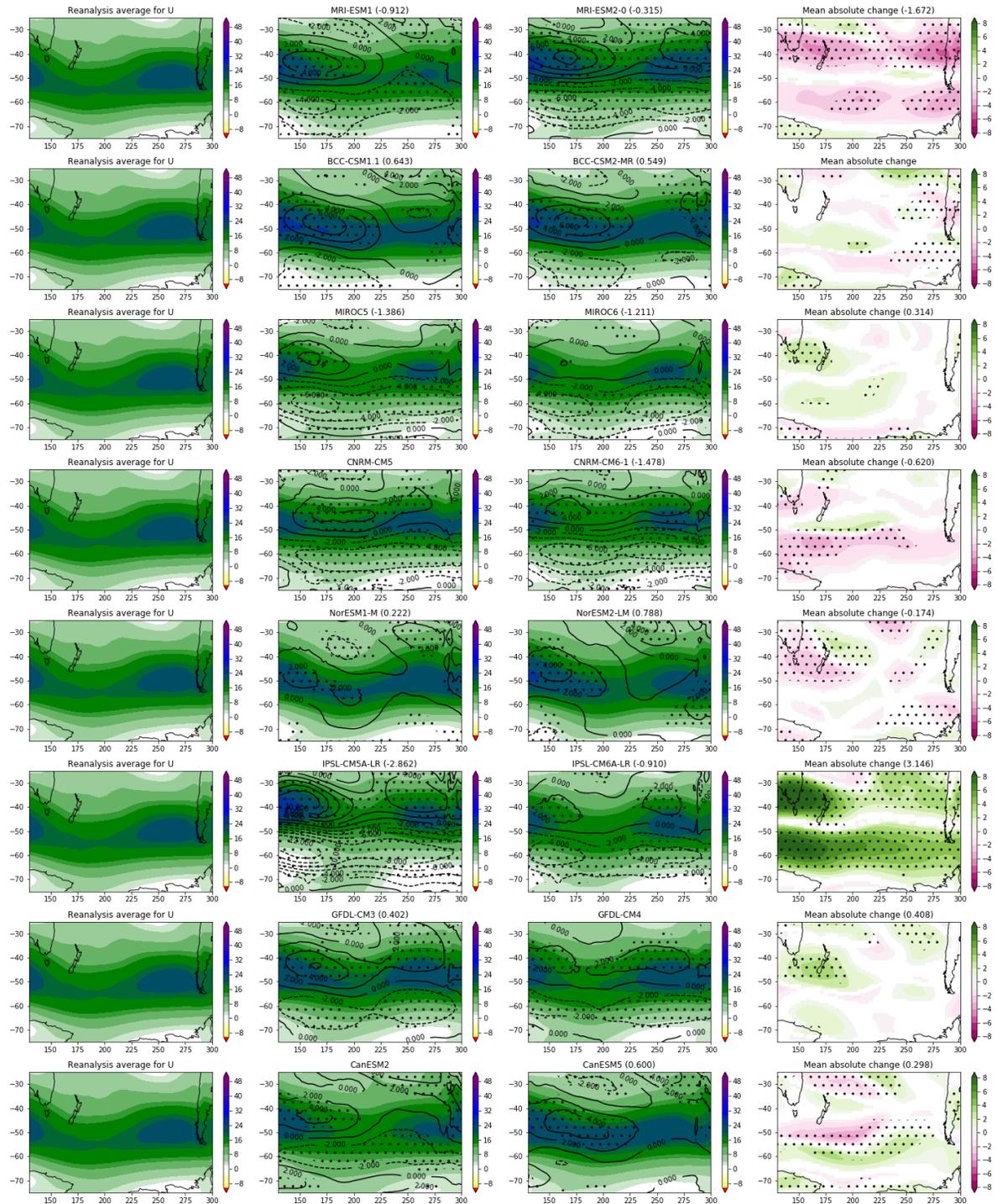
Figure D.32: As for Figure D.26, with SP DJF

Figure D.33: As for Figure D.36, with SP DJF.
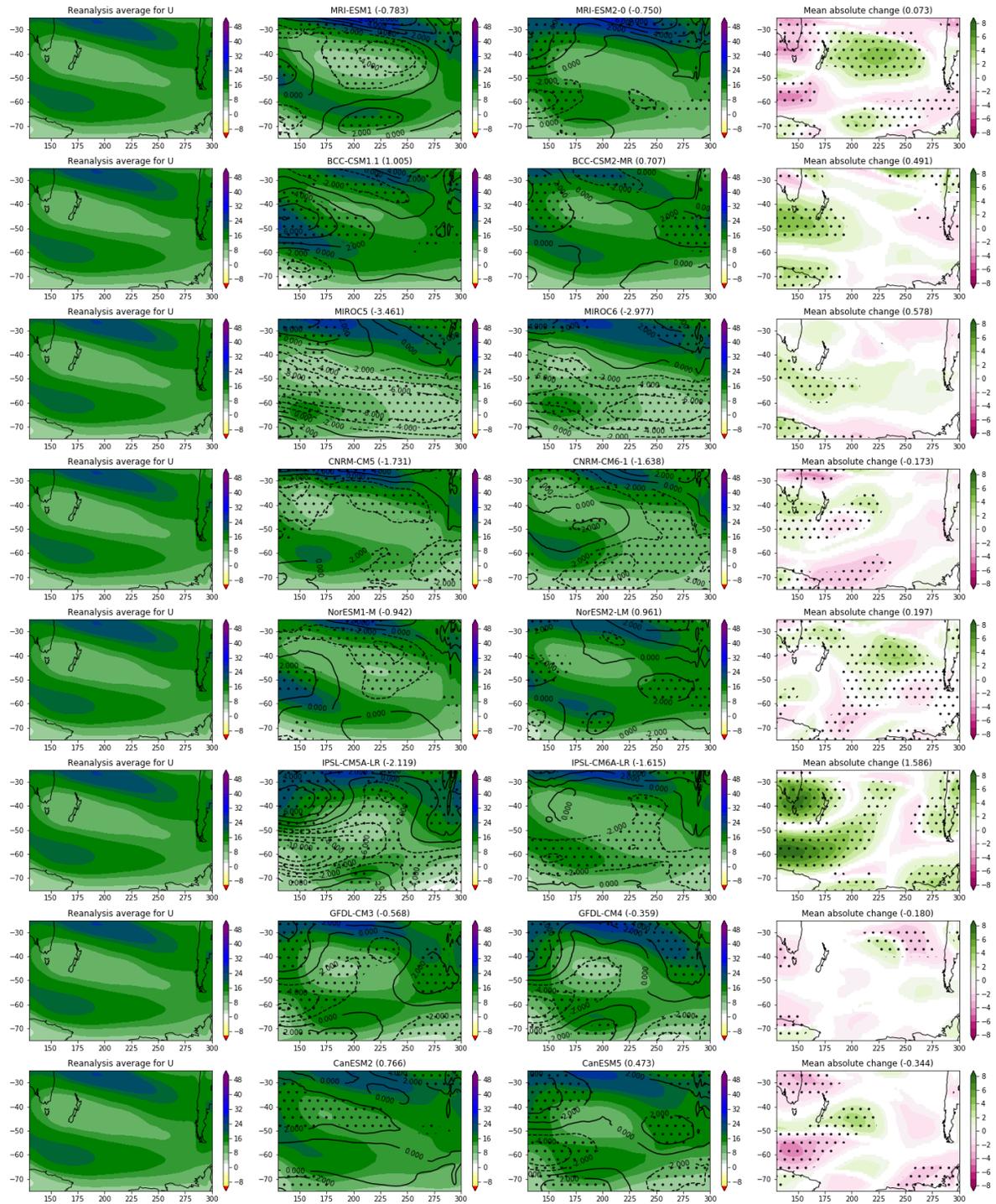
Figure D.34: As for Figure D.25, with SP JJA
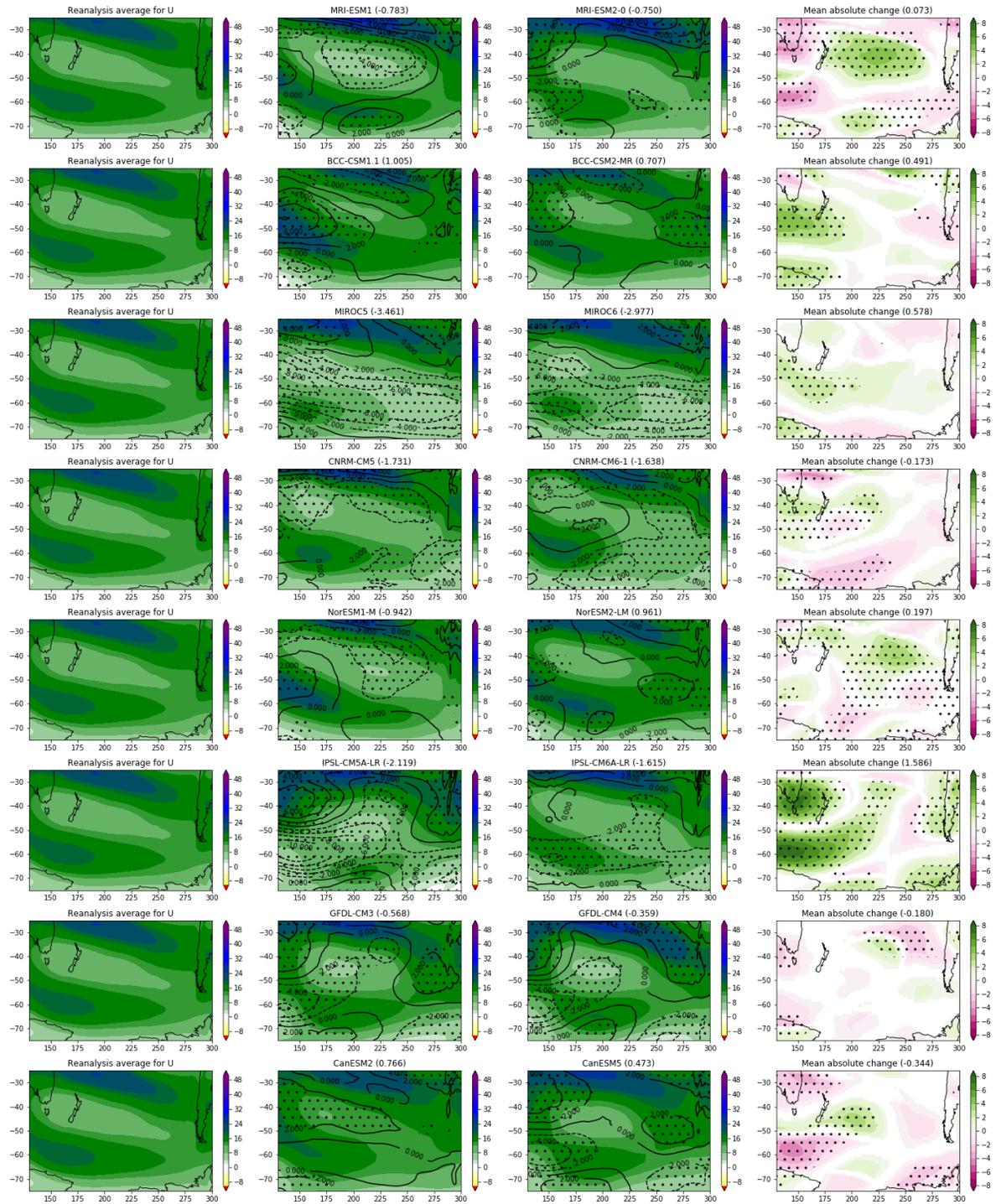
Figure D.35: As for Figure D.26, with SP JJA
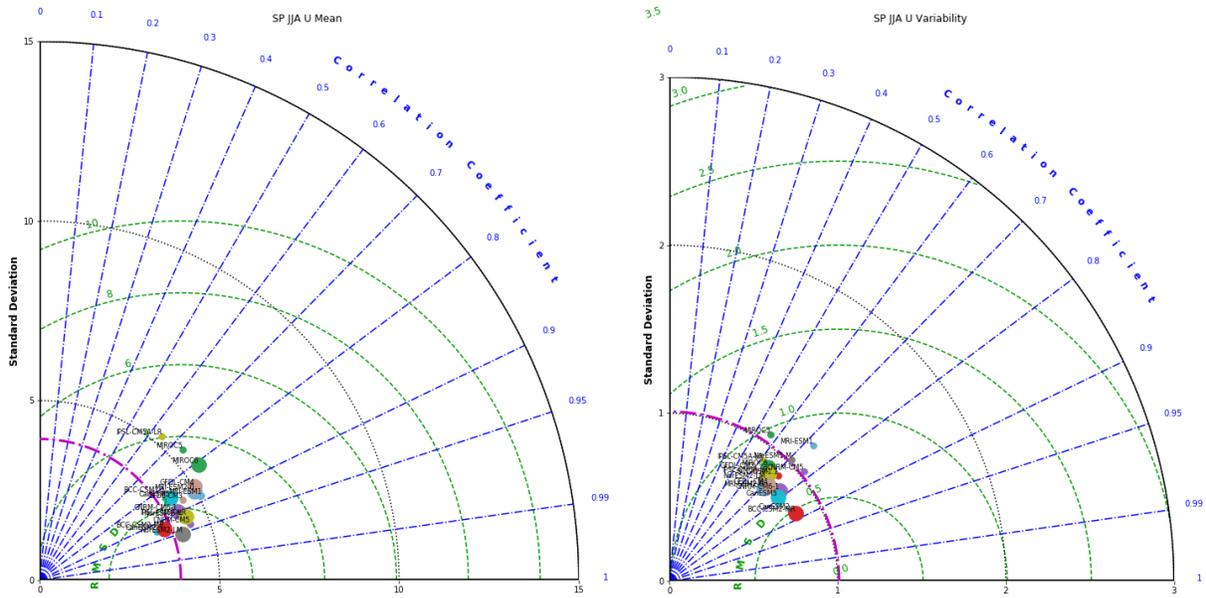
Figure D.36: As for Figure D.27, with SP JJA.
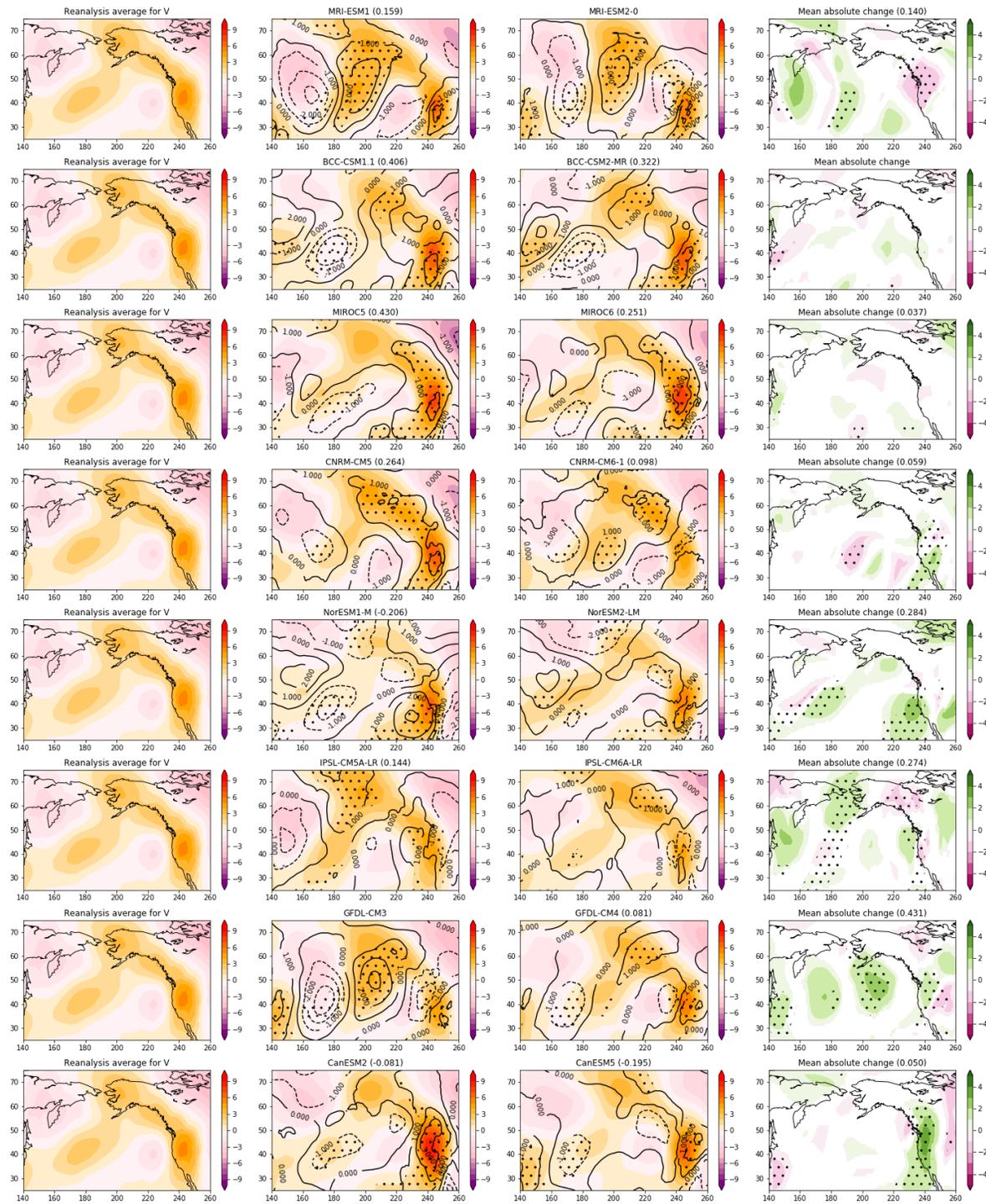
Figure D.37: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for V in NP JJA.
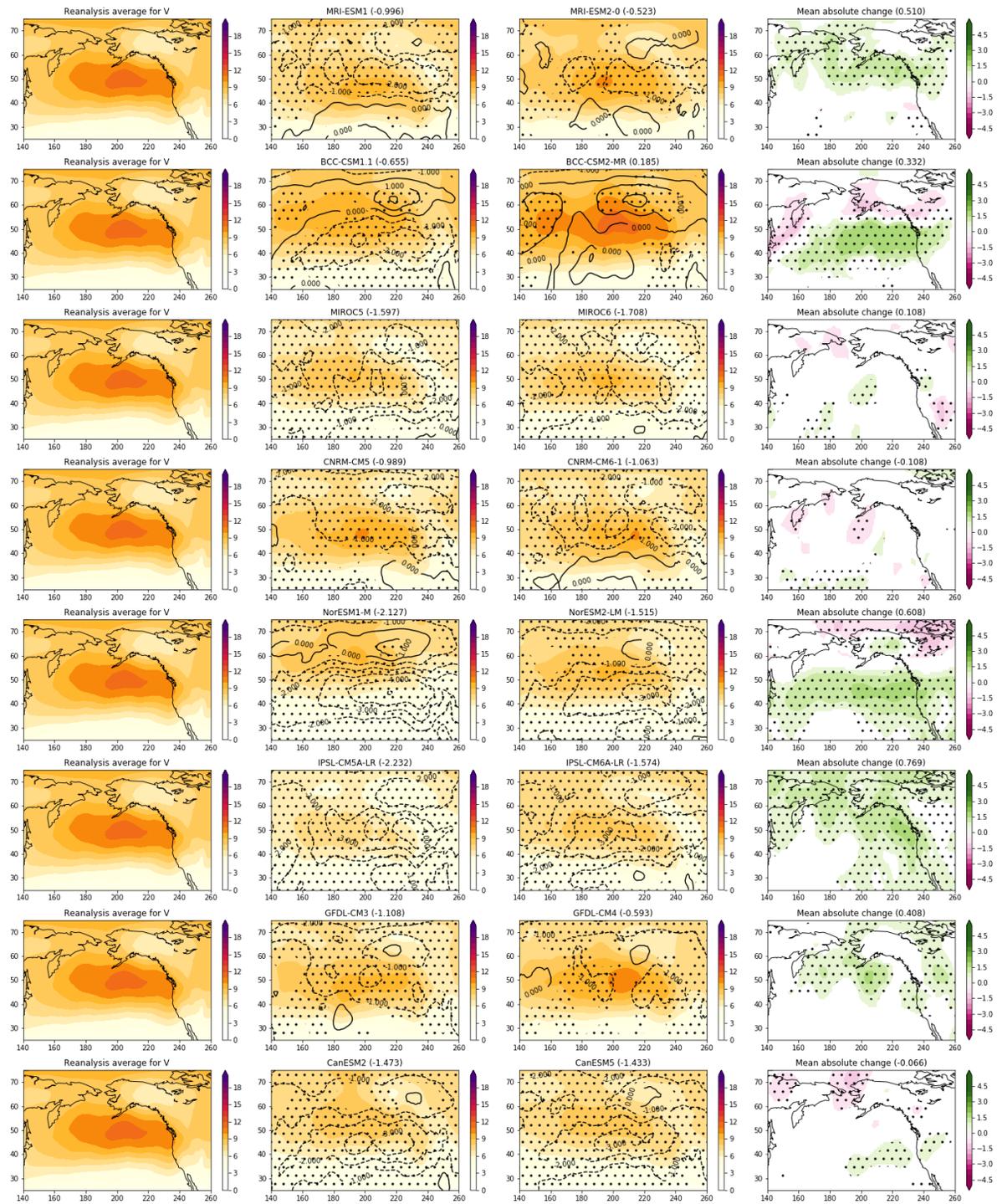
Figure D.38: (left) Reanalysis average, (middle left) CMIP5 average, (middle right) CMIP6 average, and (right) CMIP5/CMIP6 comparison of relative error change for V variability in NP JJA.
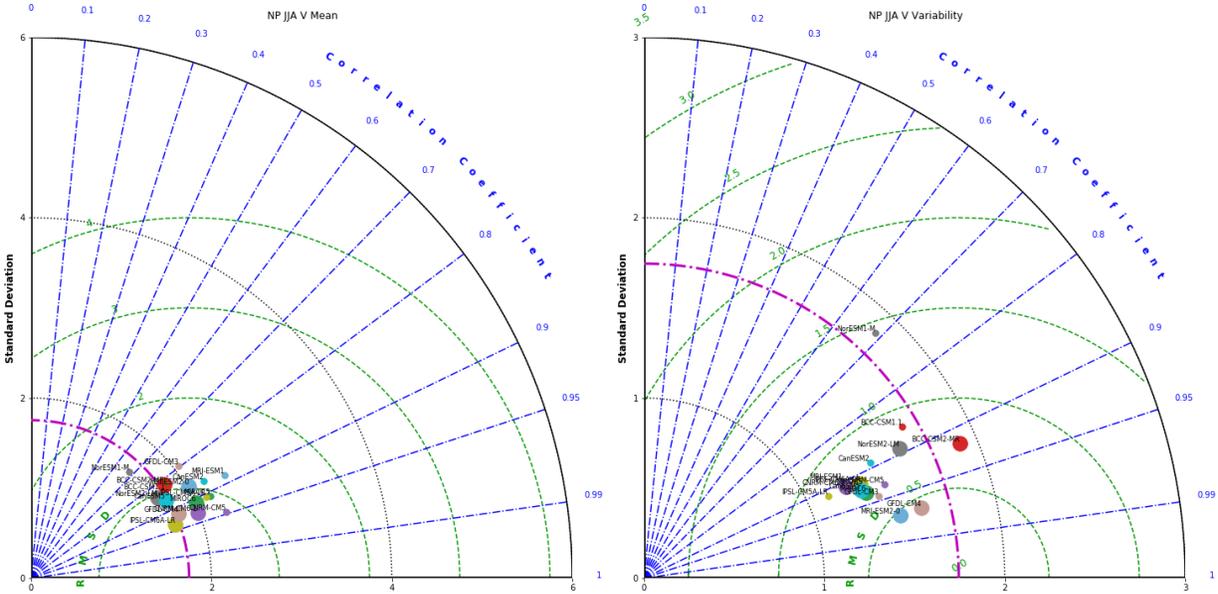
Figure D.39: Taylor plot of (left) the seasonal mean BF field and (right) the temporal variability in the seasonal V field (measured as the standard deviation of the per-gridpoint time series) for CMIP5 (smaller points) and CMIP6 (larger points) models. The plot axes denote the standard deviation of the field magnitudes (and black dotted lines denote the axis ticks); the reference reanalysis standard deviation for the NP JJA field is denoted as a magenta broken line. The blue broken lines denote the level of pattern correlation between the models and the reanalysis, and the green dashed lines denote the centered root mean square deviation (RMSD) values relative to the reanalysis. The positions of the individual model points are plotted according to their values along these axes. Note that the maximum extent of the standard deviation axis is 6 for the mean plot while it is 3 for the variability plot.

214

Figure D.40: As for Figure D.37, with NP DJF
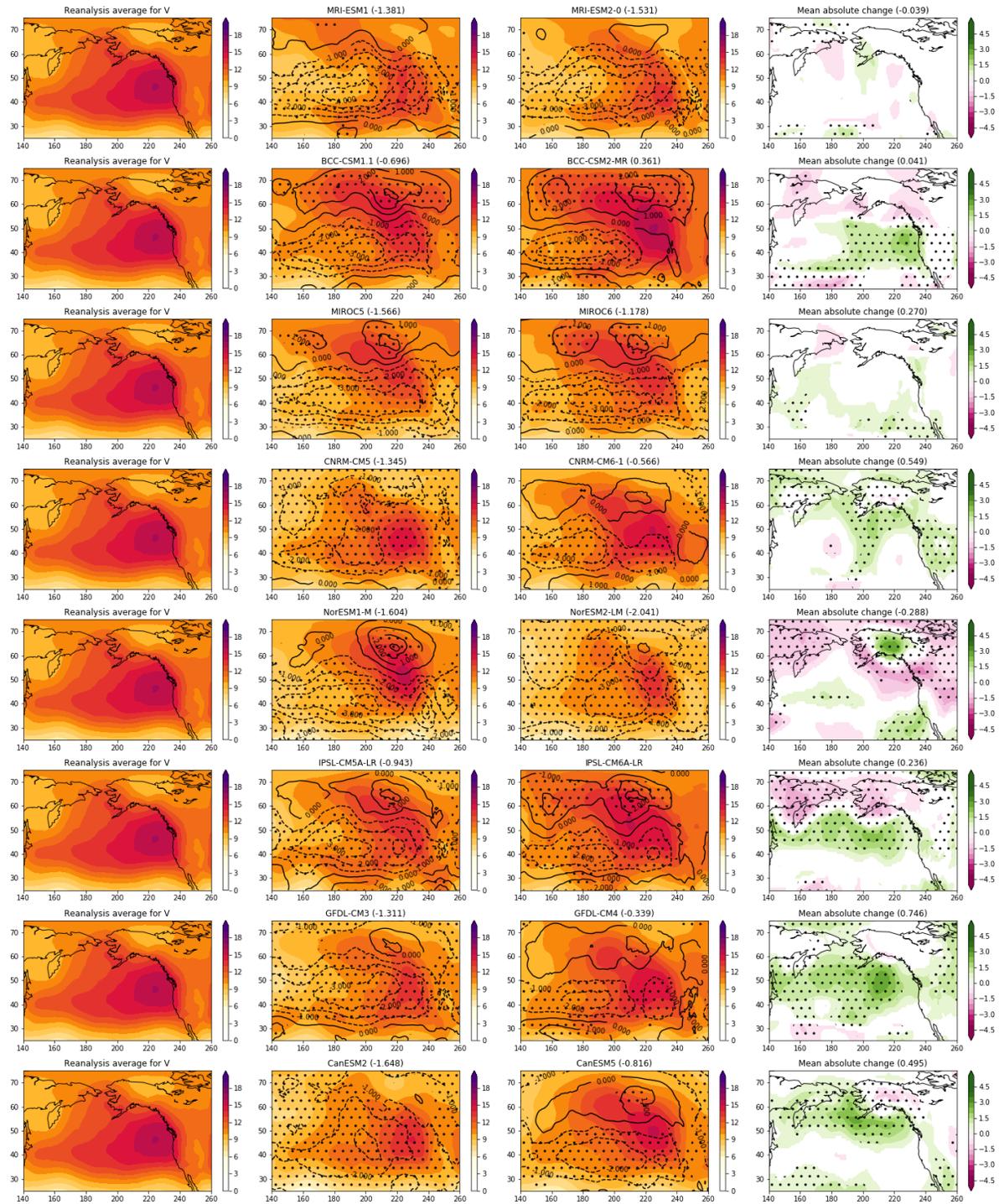
Figure D.41: As for Figure D.38, with NP DJF

216
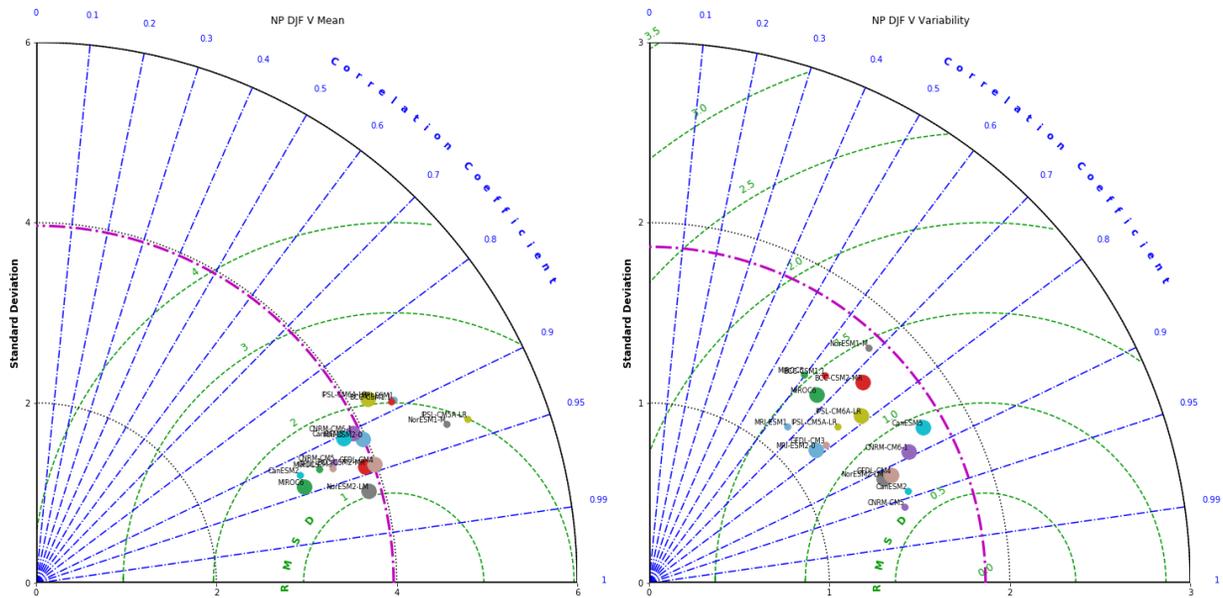
Figure D.42: As for Figure D.39, with NP DJF.
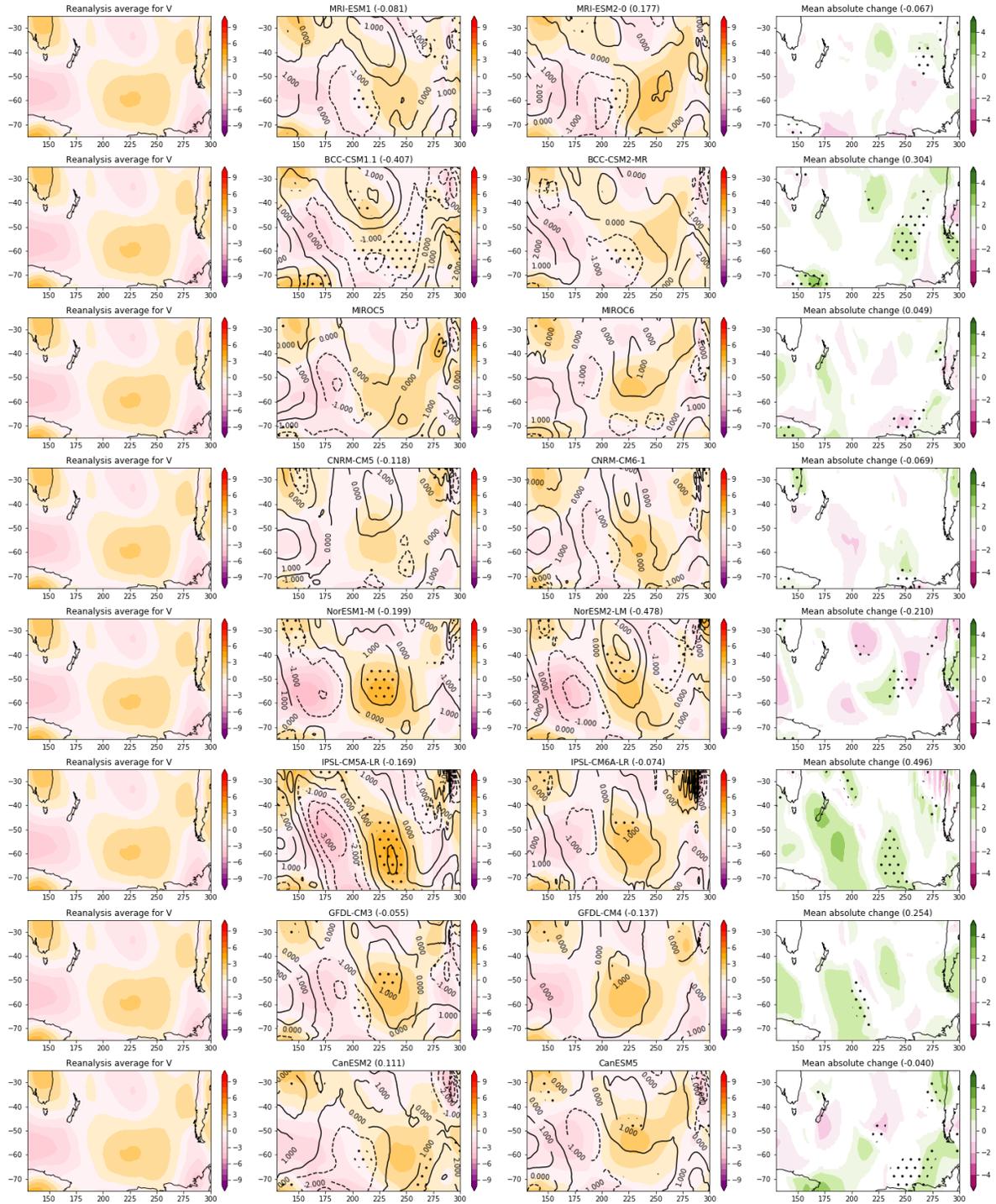
Figure D.43: As for Figure D.37, with SP DJF
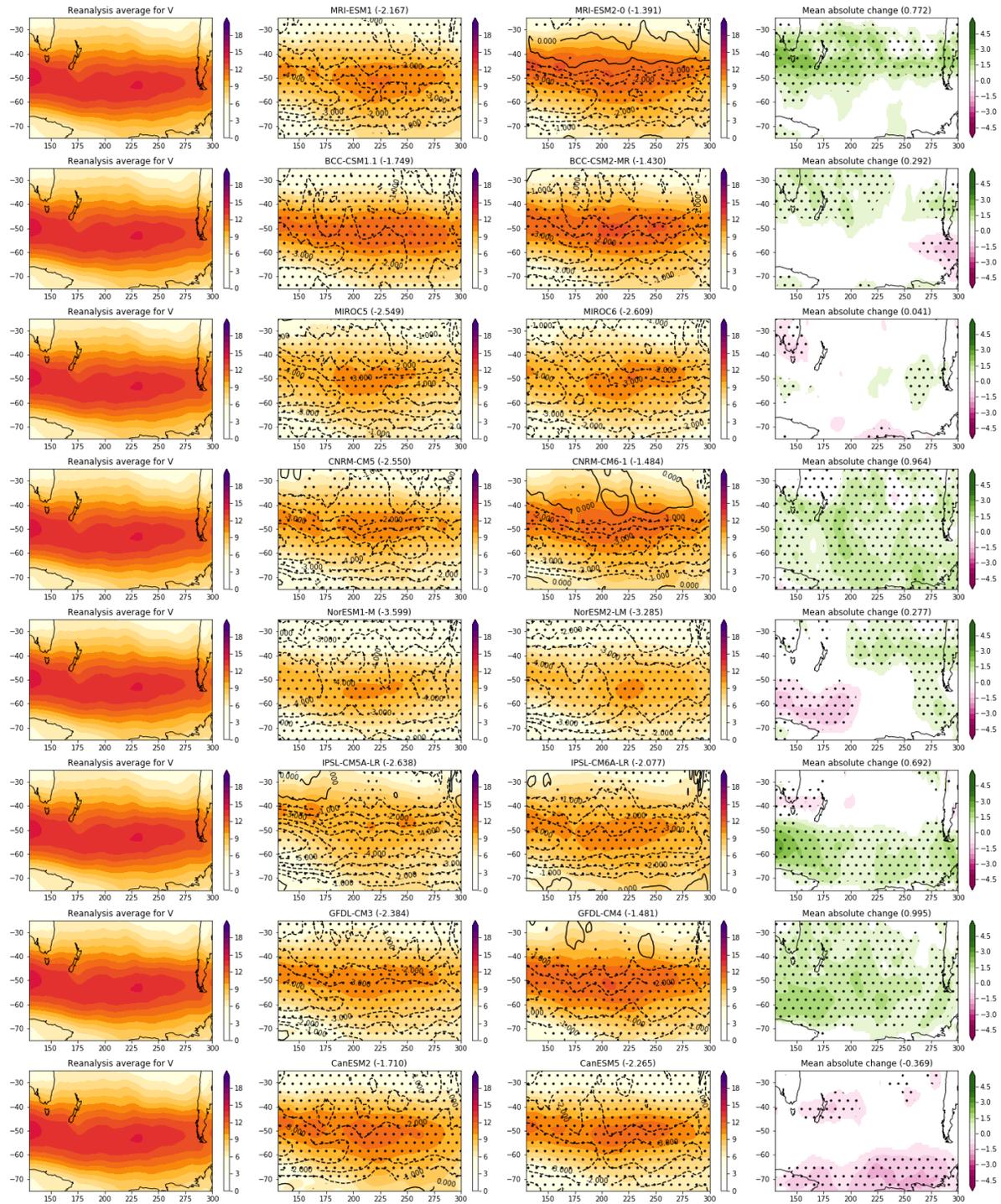
Figure D.44: As for Figure D.38, with SP DJF
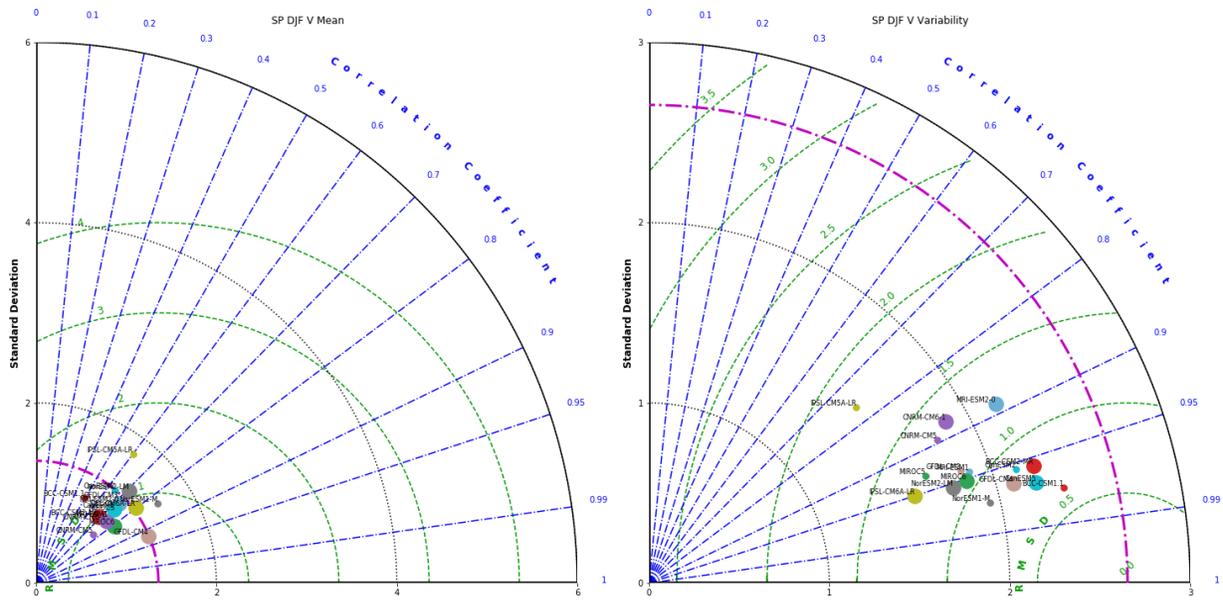
Figure D.45: As for Figure D.39, with SP DJF.
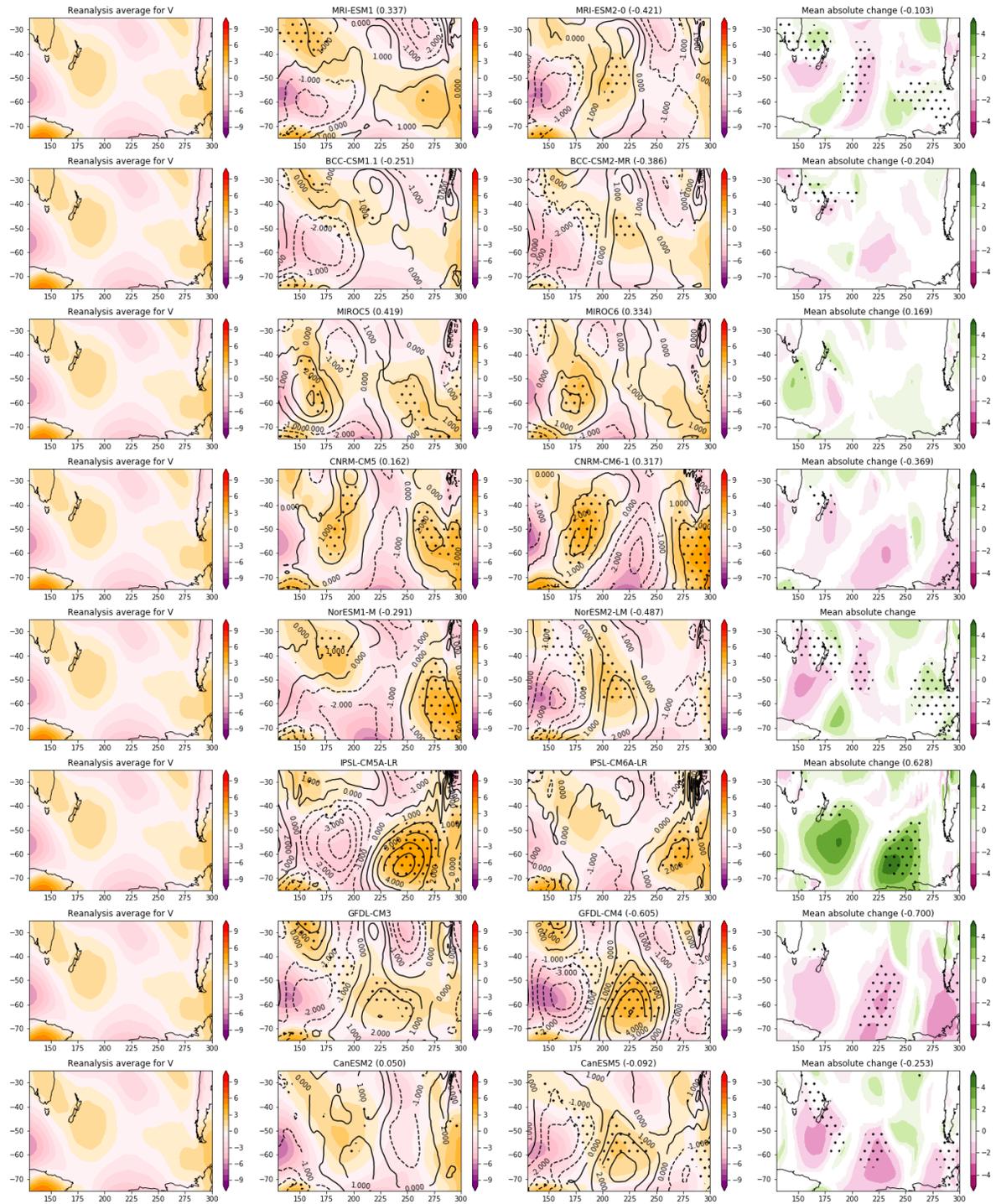
Figure D.46: As for Figure D.37, with SP JJA

221
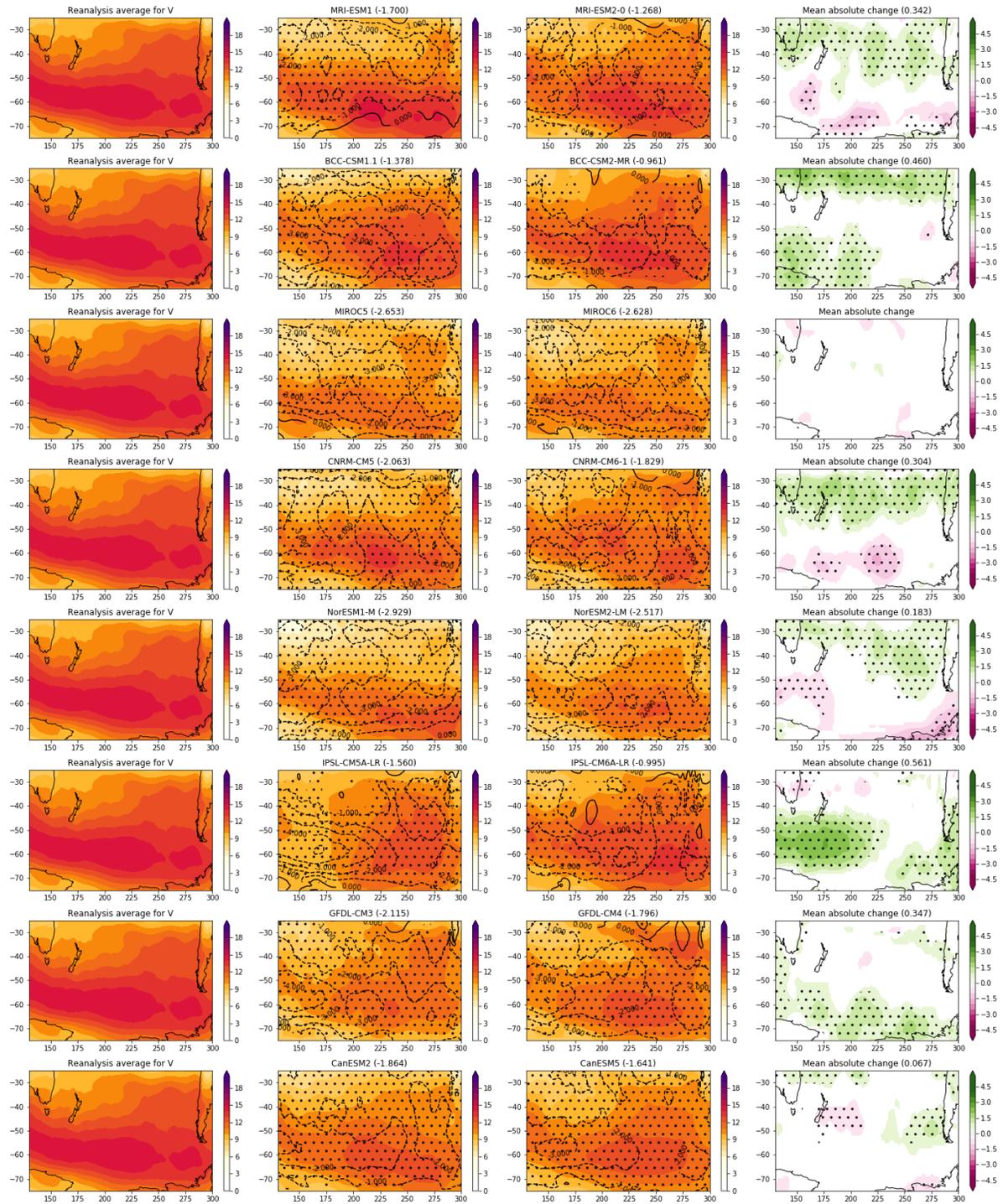
Figure D.47: As for Figure D.38, with SP JJA
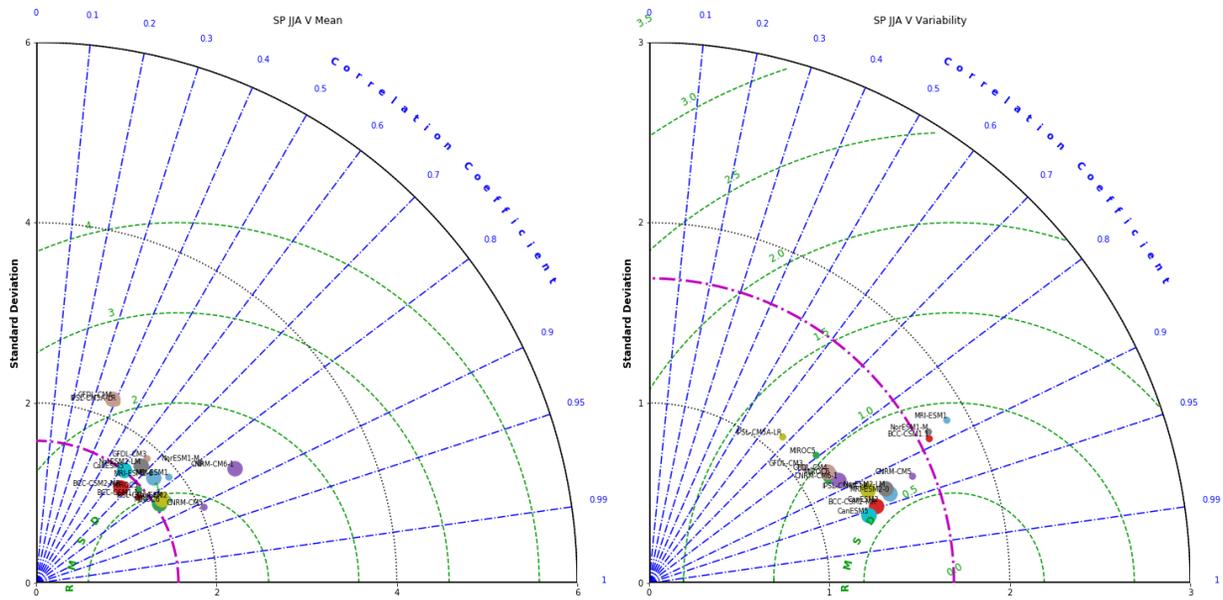
Figure D.48: As for Figure D.39, with SP JJA.

References

Adrian M. Altenhoff, Olivia Martius, Mischa Croci-Maspoli, Cornelia Schwierz, and Huw C. Davies. Linkage of atmospheric blocks and synoptic-scale Rossby waves: a climatological analysis. Tellus A, 60(5):1053–1063, October 2008. ISSN 1600-0870. doi: 10.1111/j.1600-0870.2008.00354.x. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0870.2008.00354.x/abstract`.

Elizabeth A. Barnes, Julia Slingo, and Tim Woollings. A methodology for the comparison of blocking climatologies across indices, models and climate scenarios. Climate Dynamics, 38(11-12):2467–2481, June 2012. ISSN 0930-7575, 1432-0894. doi: 10.1007/s00382-011-1243-6. URL `http://link.springer.com/10.1007/s00382-011-1243-6`.

D. Barriopedro, R. García-Herrera, and R. M. Trigo. Application of blocking diagnosis methods to General Circulation Models. Part I: a novel detection scheme. Climate Dynamics, 35(7-8):1373–1391, December 2010. ISSN 0930-7575, 1432-0894. doi: 10.1007/s00382-010-0767-5. URL `http://link.springer.com/10.1007/s00382-010-0767-5`.

David Barriopedro, Ricardo García-Herrera, Anthony R. Lupo, and Emiliano Hernández. A Climatology of Northern Hemisphere Blocking. J. Climate, 19(6):1042–1063, March 2006. ISSN 0894-8755. doi: 10.1175/JCLI3678.1. URL `https://journals.ametsoc.org/doi/abs/10.1175/JCLI3678.1`.

Julie Berckmans, Tim Woollings, Marie-Estelle Demory, Pier-Luigi Vidale, and Malcolm Roberts. Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. Atmospheric Science Letters, 14(1):34–40, 2013. ISSN 1530-261X. doi: 10.1002/asl2.412. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/asl2.412`.

Ho Nam Cheung, Wen Zhou, Hing Yim Mok, Man Chi Wu, and Yaping Shao. Revisiting the climatology of atmospheric blocking in the Northern Hemisphere.

Adv. Atmos. Sci., 30(2):397–410, March 2013. ISSN 0256-1530, 1861-9533. doi: 10.1007/s00376-012-2006-y. URL `https://link.springer.com/article/10.1007/s00376-012-2006-y`.

M. Croci-Maspoli, C. Schwierz, and H. C. Davies. A Multifaceted Climatology of Atmospheric Blocking and Its Recent Linear Trend. J. Climate, 20(4):633–649, February 2007. ISSN 0894-8755. doi: 10.1175/JCLI4029.1. URL `https://journals.ametsoc.org/doi/full/10.1175/JCLI4029.1`.

Paolo Davini and Fabio D'Andrea. Northern Hemisphere Atmospheric Blocking Representation in Global Climate Models: Twenty Years of Improvements? Journal of Climate, 29(24):8823–8840, December 2016. ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-16-0242.1. URL `http://journals.ametsoc.org/doi/10.1175/JCLI-D-16-0242.1`.

Paolo Davini, Chiara Cagnazzo, Silvio Gualdi, and Antonio Navarra. Bidimensional Diagnostics, Variability, and Trends of Northern Hemisphere Blocking. Journal of Climate, 25(19):6496–6509, October 2012. ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-12-00032.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00032.1`.

D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656):553–597, April 2011. ISSN 00359009. doi: 10.1002/qj.828. URL `http://doi.wiley.com/10.1002/qj.828`.

Randall M. Dole and Neil D. Gordon. Persistent Anomalies of the Extratropical Northern

Hemisphere Wintertime Circulation: Geographical Distribution and Regional Persistence Characteristics. <u>Monthly Weather Review</u>, 111(8):1567–1586, August 1983. ISSN 0027-0644, 1520-0493. doi: 10.1175/1520-0493(1983)111⟨1567:PAOTEN⟩2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281983%29111%3C1567%3APAOTEN%3E2.0.CO%3B2`.

Etienne Dunn-Sigouin and Seok-Woo Son. Northern Hemisphere blocking frequency and duration in the CMIP5 models: CMIP5 BLOCKING. <u>Journal of Geophysical Research: Atmospheres</u>, 118(3):1179–1188, February 2013. ISSN 2169897X. doi: 10.1002/jgrd.50143. URL `http://doi.wiley.com/10.1002/jgrd.50143`.

Etienne Dunn-Sigouin, Seok-Woo Son, and Hai Lin. Evaluation of Northern Hemisphere Blocking Climatology in the Global Environment Multiscale Model. <u>Mon. Wea. Rev.</u>, 141(2):707–727, August 2012. ISSN 0027-0644. doi: 10.1175/MWR-D-12-00134.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-12-00134.1`.

Veronika Eyring, Sandrine Bony, Gerald A. Meehl, Catherine A. Senior, Bjorn Stevens, Ronald J. Stouffer, and Karl E. Taylor. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. <u>Geosci. Model Dev.</u>, 9(5):1937–1958, May 2016. ISSN 1991-9603. doi: 10.5194/gmd-9-1937-2016. URL `https://www.geosci-model-dev.net/9/1937/2016/`.

Todd S. Glickman. <u>American Meteorological Society Glossary of Meteorology</u>. American Meteorological Society, Boston, Mass, 2nd ed edition, 2000. ISBN 978-1-878220-34-9.

William L. Grose and Brian J. Hoskins. On the Influence of Orography on Large-Scale Atmospheric Flow. <u>J. Atmos. Sci.</u>, 36(2):223–234, February 1979. ISSN 0022-4928. doi: 10.1175/1520-0469(1979)036⟨0223:OTIOOO⟩2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1979)036%3C0223:OTIOOO%3E2.0.CO;2`.

Richard Grotjahn. Identifying extreme hottest days from large scale upper air data: a pilot scheme to find California Central Valley summertime maximum surface temperatures. <u>Clim Dyn</u>, 37(3-4):587–604, August 2011. ISSN 0930-7575, 1432-0894. doi:

10.1007/s00382-011-0999-z. URL `https://link.springer.com/article/10.1007/s00382-011-0999-z`.

Richard Grotjahn and Ghislain Faure. Composite Predictor Maps of Extraordinary Weather Events in the Sacramento, California, Region. Wea. Forecasting, 23(3): 313–335, June 2008. ISSN 0882-8156. doi: 10.1175/2007WAF2006055.1. URL `https://journals.ametsoc.org/doi/abs/10.1175/2007WAF2006055.1`.

Richard Grotjahn and Rui Zhang. Synoptic Analysis of Cold Air Outbreaks over the California Central Valley. J. Climate, 30(23):9417–9433, September 2017. ISSN 0894-8755. doi: 10.1175/JCLI-D-17-0167.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-17-0167.1`.

Yongli He, Jianping Huang, and Mingxia Ji. Impact of land–sea thermal contrast on interdecadal variation in circulation and blocking. Clim Dyn, 43(12):3267–3279, December 2014. ISSN 1432-0894. doi: 10.1007/s00382-014-2103-y. URL `https://doi.org/10.1007/s00382-014-2103-y`.

Chi-Cherng Hong, Huang-Hsiung Hsu, Nai-Hsin Lin, and Hsun Chiu. Roles of European blocking and tropical-extratropical interaction in the 2010 Pakistan flooding. Geophysical Research Letters, 38(13), July 2011. ISSN 1944-8007. doi: 10. 1029/2011GL047583. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL047583`.

Brian J. Hoskins and David J. Karoly. The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing. J. Atmos. Sci., 38(6):1179–1196, June 1981. ISSN 0022-4928. doi: 10.1175/1520-0469(1981)038⟨1179:TSLROA⟩2.0.CO; 2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1981)038%3C1179%3ATSLROA%3E2.0.CO%3B2`.

R. A. Houze, K. L. Rasmussen, S. Medina, S. R. Brodzik, and U. Romatschke. Anomalous Atmospheric Events Leading to the Summer 2010 Floods in Pakistan. Bull. Amer. Meteor. Soc., 92(3):291–298, February 2011. ISSN 0003-0007. doi:

10.1175/2010BAMS3173.1. URL `https://journals.ametsoc.org/doi/10.1175/2010BAMS3173.1`.

Paul Jaccard. Nouvelles Recherches Sur La Distribution Florale. Bulletin de la Société vaudoise des Sciences Naturelles, 44:223–270, 1908.

Yun-Young Lee and Richard Grotjahn. California Central Valley Summer Heat Waves Form Two Ways. J. Climate, 29(3):1201–1217, December 2015. ISSN 0894-8755. doi: 10.1175/JCLI-D-15-0270.1. URL `https://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-15-0270.1`.

Harald Lejenäs and Hans økland. Characteristics of northern hemisphere blocking as determined from a long time series of observational data. Tellus A, 35A(5):350–362, October 1983. ISSN 02806495, 16000870. doi: 10.1111/j.1600-0870.1983.tb00210.x. URL `http://tellusa.net/index.php/tellusa/article/view/11446`.

Camille Li and Justin J. Wettstein. Thermally Driven and Eddy-Driven Jet Variability in Reanalysis. Journal of Climate, 25(5):1587–1596, March 2012. ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-11-00145.1. URL `http://journals.ametsoc.org/doi/10.1175/JCLI-D-11-00145.1`.

Anthony R. Lupo and Phillip J. Smith. Climatological features of blocking anticyclones in the Northern Hemisphere. Tellus A, 47(4):439–456, August 1995. ISSN 1600-0870. doi: 10.1034/j.1600-0870.1995.t01-3-00004.x. URL `http://onlinelibrary.wiley.com/doi/10.1034/j.1600-0870.1995.t01-3-00004.x/abstract`.

Carlos Martinez, Young-Oh Kwon, Hyodae Seo, and Justin Small. North Atlantic Amospheric Blocking and Atlantic Multidecadal Oscillation in CESM1 Large Ensemble Simulations. URL `http://nldr.library.ucar.edu/repository/assets/soars/SOARS-000-000-000-440.pdf`.

Giacomo Masato, Brian J. Hoskins, and Tim Woollings. Winter and Summer Northern Hemisphere Blocking in CMIP5 Models. J. Climate, 26(18):7044–7059, March 2013.

ISSN 0894-8755. doi: 10.1175/JCLI-D-12-00466.1. URL `https://journals.ametsoc.org/doi/full/10.1175/JCLI-D-12-00466.1`.

Steven L. Mullen. Model Experiments on the Impact of Pacific Sea Surface Temperature Anomalies on Blocking Frequency. J. Climate, 2(9):997–1013, September 1989. ISSN 0894-8755. doi: 10.1175/1520-0442(1989)002⟨0997:MEOTIO⟩2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1989)002%3C0997%3AMEOTIO%3E2.0.CO%3B2`.

Hisashi Nakamura and Takeaki Sampe. Trapping of synoptic-scale disturbances into the North-Pacific subtropical jet core in midwinter. Geophysical Research Letters, 29(16): 8–1–8–4, 2002. ISSN 1944-8007. doi: 10.1029/2002GL015535. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002GL015535`.

Hisashi Nakamura, Mototaka Nakamura, and Jeffrey L. Anderson. The Role of High- and Low-Frequency Dynamics in Blocking Formation. Mon. Wea. Rev., 125(9):2074–2093, September 1997. ISSN 0027-0644. doi: 10.1175/1520-0493(1997)125⟨2074:TROHAL⟩2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1997)125%3C2074:TROHAL%3E2.0.CO%3B2`.

Simon Parsons, James A. Renwick, and Adrian J. McDonald. An Assessment of Future Southern Hemisphere Blocking Using CMIP5 Projections from Four GCMs. J. Climate, 29(21):7599–7611, July 2016. ISSN 0894-8755. doi: 10.1175/JCLI-D-15-0754.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-15-0754.1`.

J. L. Pelly and B. J. Hoskins. A New Perspective on Blocking. Journal of the Atmospheric Sciences, 60(5):743–755, March 2003. ISSN 0022-4928, 1520-0469. doi: 10.1175/1520-0469(2003)060⟨0743:ANPOB⟩2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0469%282003%29060%3C0743%3AANPOB%3E2.0.CO%3B2`.

S. Pfahl and H. Wernli. Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the Northern Hemisphere on (sub-)daily time scales. Geophys.

Res. Lett., 39(12):L12807, June 2012. ISSN 1944-8007. doi: 10.1029/2012GL052261. URL http://onlinelibrary.wiley.com/doi/10.1029/2012GL052261/abstract.

Daniel F. Rex. Blocking Action in the Middle Troposphere and its Effect upon Regional Climate. Tellus, 2(3):196–211, August 1950. ISSN 00402826, 21533490. doi: 10.1111/j.2153-3490.1950.tb00331.x. URL http://tellusa.net/index.php/tellusa/article/view/8546.

R. Sausen, W. König, and F. Sielmann. Analysis of blocking events from observations and ECHAM model simulations. Tellus A, 47(4):421–438, August 1995. ISSN 1600-0870. doi: 10.1034/j.1600-0870.1995.t01-3-00003.x. URL http://onlinelibrary.wiley.com/doi/10.1034/j.1600-0870.1995.t01-3-00003.x/abstract.

Adam A. Scaife, Tim Woollings, Jeff Knight, Gill Martin, and Tim Hinton. Atmospheric Blocking and Mean Biases in Climate Models. Journal of Climate, 23(23):6143–6152, December 2010. ISSN 0894-8755, 1520-0442. doi: 10.1175/2010JCLI3728.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3728.1.

Adam A. Scaife, Dan Copsey, Chris Gordon, Chris Harris, Tim Hinton, Sarah Keeley, Alan O'Neill, Malcolm Roberts, and Keith Williams. Improved Atlantic winter blocking in a climate model. Geophysical Research Letters, 38(23), 2011. ISSN 1944-8007. doi: 10.1029/2011GL049573. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL049573.

Simon C. Scherrer, Mischa Croci-Maspoli, Cornelia Schwierz, and Christof Appenzeller. Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. Int. J. Climatol., 26(2):233–249, February 2006. ISSN 1097-0088. doi: 10.1002/joc.1250. URL http://onlinelibrary.wiley.com/doi/10.1002/joc.1250/abstract.

Reinhard Schiemann, Marie-Estelle Demory, Len C. Shaffrey, Jane Strachan, Pier Luigi Vidale, Matthew S. Mizielinski, Malcolm J. Roberts, Mio Matsueda, Michael F. Wehner, and Thomas Jung. The Resolution Sensitivity of Northern Hemisphere Blocking in Four

25-km Atmospheric Global Circulation Models. J. Climate, 30(1):337–358, September 2016. ISSN 0894-8755. doi: 10.1175/JCLI-D-16-0100.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-16-0100.1`.

C. Schwierz, Mischa Croci-Maspoli, and H. C. Davies. Perspicacious indicators of atmospheric blocking. Geophysical Research Letters, 31(6), 2004. ISSN 0094-8276. doi: 10.1029/2003GL019341. URL `http://doi.wiley.com/10.1029/2003GL019341`.

G. J. Shutts. The propagation of eddies in diffluent jetstreams: Eddy vorticity forcing of 'blocking' flow fields. Q.J.R. Meteorol. Soc., 109(462):737–761, October 1983. ISSN 1477-870X. doi: 10.1002/qj.49710946204. URL `http://onlinelibrary.wiley.com/doi/10.1002/qj.49710946204/abstract`.

Jana Sillmann, Mischa Croci-Maspoli, Malaak Kallache, and Richard W. Katz. Extreme Cold Winter Temperatures in Europe under the Influence of North Atlantic Atmospheric Blocking. Journal of Climate, 24(22):5899–5913, November 2011. ISSN 0894-8755, 1520-0442. doi: 10.1175/2011JCLI4075.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2011JCLI4075.1`.

Mark R. Sinclair. A Climatology of Anticyclones and Blocking for the Southern Hemisphere. Mon. Wea. Rev., 124(2):245–264, February 1996. ISSN 0027-0644. doi: 10.1175/1520-0493(1996)124⟨0245:ACOAAB⟩2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1996)124%3C0245:ACOAAB%3E2.0.CO;2`.

Daniel L. Swain, Michael Tsiang, Matz Haugen, Deepti Singh, Allison Charland, Bala Rajaratnam, and Noah S. Diffenbaugh. The Extraordinary California Drought of 2013/2014: Character, Context, and the Role of Climate Change. Bulletin of the American Meteorological Society; Boston, 95(9):S3–S7, September 2014. ISSN 00030007. URL `https://search.proquest.com/docview/1623231385/abstract/D199604F9FF14513PQ/1`.

S. Tibaldi, E. Tosi, A. Navarra, and L. Pedulli. Northern and Southern Hemisphere Seasonal Variability of Blocking Frequency and Predictability. Monthly Weather

Review, 122(9):1971–2003, September 1994. ISSN 0027-0644, 1520-0493. doi: 10.1175/
1520-0493(1994)122⟨1971:NASHSV⟩2.0.CO;2. URL `http://journals.ametsoc.org/`
`doi/abs/10.1175/1520-0493%281994%29122%3C1971%3ANASHSV%3E2.0.CO%3B2`.

Stefano Tibaldi and Franco Molteni. On the operational predictability of blocking. Tellus
A, 42(3):343–365, May 1990. ISSN 0280-6495, 1600-0870. doi: 10.1034/j.1600-0870.
1990.t01-2-00003.x. URL `http://tellusa.net/index.php/tellusa/article/view/`
`11882`.

R.A. Treidl, E.C. Birch, and P. Sajecki. Blocking action in the northern hemisphere: A Cli-
matological study. Atmosphere-Ocean, 19(1):1–23, March 1981. ISSN 0705-5900, 1480-
9214. doi: 10.1080/07055900.1981.9649096. URL `http://www.tandfonline.com/doi/`
`abs/10.1080/07055900.1981.9649096`.

P. A. Ullrich and C. M. Zarzycki. TempestExtremes: a framework for scale-insensitive
pointwise feature tracking on unstructured grids. Geosci. Model Dev., 10(3):1069–
1090, March 2017. ISSN 1991-9603. doi: 10.5194/gmd-10-1069-2017. URL `https:`
`//www.geosci-model-dev.net/10/1069/2017/`.

Linda M. Whittaker and Lyle H. Horn. Geographical and Seasonal Distribution of
North American Cyclogenesis, 1958–1977. Mon. Wea. Rev., 109(11):2312–2322, Novem-
ber 1981. ISSN 0027-0644. doi: 10.1175/1520-0493(1981)109⟨2312:GASDON⟩2.0.CO;
2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1981)109%`
`3C2312:GASDON%3E2.0.CO%3B2`.

Jason M. Wiedenmann, Anthony R. Lupo, Igor I. Mokhov, and Elena A. Tikhonova.
The Climatology of Blocking Anticyclones for the Northern and Southern Hemi-
spheres: Block Intensity as a Diagnostic. J. Climate, 15(23):3459–3473, Decem-
ber 2002. ISSN 0894-8755. doi: 10.1175/1520-0442(2002)015⟨3459:TCOBAF⟩2.0.CO;
2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(2002)015%`
`3C3459:TCOBAF%3E2.0.CO%3B2`.

Tim Woollings, Abdel Hannachi, and Brian Hoskins. Variability of the North Atlantic

eddy-driven jet stream. Quarterly Journal of the Royal Meteorological Society, 136 (649):856–868, 2010. ISSN 1477-870X. doi: 10.1002/qj.625. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.625`.

Tim Woollings, David Barriopedro, John Methven, Seok-Woo Son, Olivia Martius, Ben Harvey, Jana Sillmann, Anthony R. Lupo, and Sonia Seneviratne. Blocking and its Response to Climate Change. Curr Clim Change Rep, 4(3):287–300, September 2018. ISSN 2198-6061. doi: 10.1007/s40641-018-0108-z. URL `https://doi.org/10.1007/s40641-018-0108-z`.