Constraining Future Summer Austral Jet Stream Positions in the CMIP5 Ensemble by Process-Oriented Multiple Diagnostic Regression*

Sabrina Wenzel and Veronika Eyring

Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

Edwin P. Gerber

Courant Institute of Mathematical Sciences, New York University, New York, New York

Alexey Yu. Karpechko

Arctic Research, Finnish Meteorological Institute, Helsinki, Finland

ABSTRACT

Stratospheric ozone recovery and increasing greenhouse gases are anticipated to have a large impact on the Southern Hemisphere extratropical circulation, shifting the jet stream and associated storm tracks. Models participating in phase 5 of the Coupled Model Intercomparison Project poorly simulate the austral jet, with a mean equatorward bias and 10° latitude spread in their historical climatologies, and project a wide range of future trends in response to anthropogenic forcing in the representative concentration pathway (RCP) scenarios. Here, the question is addressed whether the unweighted multimodel mean (uMMM) austral jet projection of the RCP4.5 scenario can be improved by applying a process-oriented multiple diagnostic ensemble regression (MDER). MDER links future projections of the jet position to processes relevant to its simulation under present-day conditions. MDER is first targeted to constrain near-term (2015–34) projections of the austral jet position and selects the historical jet position as the most important of 20 diagnostics. The method essentially recognizes the equatorward bias in the past jet position and provides a bias correction of about 1.5° latitude southward to future projections. When the target horizon is extended to midcentury (2040–59), the method also recognizes that lower-stratospheric temperature trends over Antarctica, a proxy for the intensity of ozone depletion, provide additional information that can be used to reduce uncertainty in the ensemble mean projection. MDER does not substantially alter the uMMM long-term position in jet position but reduces the uncertainty in the ensemble mean projection. This result suggests that accurate observational constraints on upper-tropospheric and lower-stratospheric temperature trends are needed to constrain projections of the austral jet position.

1. Introduction

Uncertainty in the circulation response to anthropogenic forcing remains a pressing problem in climate projections (Shepherd 2014). The models participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5) simulate a wide spread in the austral jet position trends in both the historical and future scenarios, particularly in austral summer (Eyring et al. 2013; Gerber and Son 2014). Shifts in the jet and the associated storm track in this season have had significant impacts on regional temperatures and precipitation across the Southern Hemisphere (SH) in recent decades (e.g., Kang et al. 2011; Thompson et al. 2011) and have also impacted the meridional overturning of the ocean, with implications for carbon and heat uptake (e.g., Waugh et al. 2013). It is therefore important to provide reliable projections of future summer austral jet position trends.

Historical trends in the austral jet stream have been largest in austral summer (Marshall 2003), as the

circulation has been impacted by two anthropogenic forcings in this season: stratospheric ozone loss and greenhouse gas (GHG) increase (Arblaster and Meehl 2006). Ozone depletion led to radiative cooling of the lower stratosphere over Antarctica in the late twentieth century and strongly impacted the SH extratropical circulation, shifting the jet stream poleward (Gillett and Thompson 2003; Son et al. 2010). The recovery of ozone is expected to have the opposite effect as ozone depletion, thus tending to shift the jet equatorward (Perlwitz et al. 2008; Son et al. 2008). Increasing GHGs appear to drive a poleward expansion of the jet streams in both hemispheres (Yin 2005), and controlled CO_2-doubling experiments suggest that the response of the jet in the SH is strongest in austral summer (Kushner et al. 2001).

The balance between ozone recovery and increasing GHGs will influence future austral jet position (Son et al. 2008; Arblaster et al. 2011). While ozone appears to have dominated the response in the past (Polvani et al. 2011), the balance in the future depends in part on the speed of ozone recovery and the strength of future greenhouse gas emissions (Son et al. 2010; Simpkins and Karpechko 2012; Barnes and Polvani 2013; Eyring et al. 2013). Even for a given forcing scenario, however, there is still considerable spread. Among the CMIP5 models, Gerber and Son (2014) found that in a moderate carbon future, as characterized by the representative concentration pathway 4.5 (RCP4.5), differences in ozone changes contributed most significantly to the spread in future climate projections. There was also considerable spread associated with processes independent of the thermodynamic trends, however, suggesting that uncertainty in the dynamical response to temperature trends also plays a role in model spread.

CMIP5 models differ substantially in their ability to simulate the basic climatology and trends of the twentieth century (Eyring et al. 2013). The austral circulation has long presented a particular challenge to climate models, with substantial biases in the basic position and variability of the jet stream (e.g., Kidston and Gerber 2010; Swart and Fyfe 2012). These biases have significant implications; for example, Bracegirdle et al. (2015) emphasize that a model's ability to represent the austral circulation is one of the most important factors influencing future projections of the Antarctic climate.

In this study, we diagnose relationships between models' ability to simulate the historical climate and their ability to simulate the future, with an ultimate goal of better discriminating among their projections of the future. This relates to the question whether the ordinary arithmetic ensemble mean (i.e., the "one model–one vote" approach; Knutti et al. 2010) gives the best estimate of future austral jet position. We use the multiple diagnostic ensemble regression (MDER) methodology of Karpechko et al. (2013) to relate future projections to process-oriented diagnostics based on the twentieth century in order to see if one can improve on the unweighted multimodel mean (uMMM) projection of future climate.

We first explain the MDER method and detail the process-oriented diagnostics that are used to evaluate the models' ability to simulate the austral climate in section 2. We include the main diagnostics that have been linked to the austral jet position in the recent literature. Section 3 then outlines the observational and reanalysis constraints on these diagnostics and lists the CMIP5 models used in this study. In section 4, we use MDER to improve projections of the position of the jet stream in the near term (2015–34) and midterm (2040–59). We conclude our study in section 5 with a discussion of the results.

2. Method and diagnostics

a. MDER

Karpechko et al. (2013) developed the MDER method to show how Antarctic total column ozone projections in October are related to observable process-oriented present-day diagnostics in chemistry–climate models. The method identified key biases in model transport processes and used them to establish future ozone projections with higher precision compared to the uMMM projection.

The method is based on statistical relationships between models' simulation of the historical climatology and their future projections, which are often referred to as "emergent constraints" (e.g., Bracegirdle and Stephenson 2012). If there is a robust linear relationship between future projections of a target variable (e.g., the position of the austral jet) and a diagnostic of the past climate, one can use observations to make an improved forecast, as illustrated schematically in Fig. 1. The key idea is to use the models to establish a relationship between the historical climatology and future projections (i.e., the linear regression illustrated by the red line) and use this relationship to estimate the future projection based on historical observations. The method thus depends on 1) the existence of robust correlations between key processes and the future variable to be projected and 2) the ability to constrain the relationships with available observations.

As emphasized by Bracegirdle and Stephenson (2012), one must be wary of spurious relationships
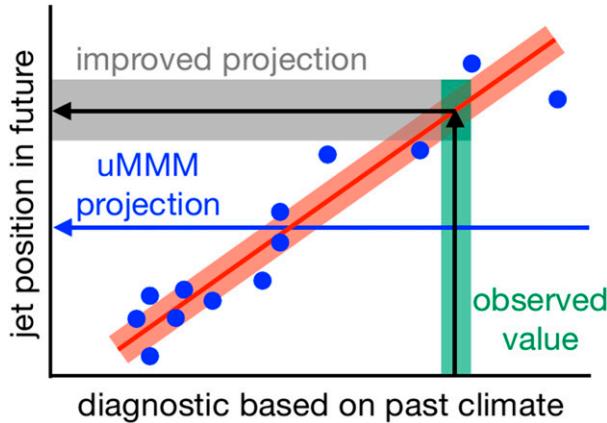
FIG. 1. A schematic diagram illustrating the linear regression model for constraining future projections of the jet position. Each blue dot represents (hypothetical) output from different climate models, comparing a model's performance on a diagnostic based on the historical scenario integration (x axis) with its projection of the jet position in the future (y axis). The uMMM projection is the average of all blue dots in the y direction and marked by the horizontal blue arrow. The linear relationship between the past diagnostic and future projection illustrates an emergent constraint, which is quantified by linear regression (red line). The linear relationship can be used to estimate the future projection based on the observations of the past diagnostic, as marked by the black arrows. Uncertainty in the new projection (gray shading) arises from two sources: uncertainty in the observational constraint (green shading) and uncertainty in the linear regression (red shading).

between the past climatology and future projections. This danger of overfitting grows larger when considering multiple diagnostics at once, and the main difficulty of the MDER method stems from the need to systematically reject spurious relationships and avoid using redundant information (i.e., cases where the same effective emergent constraint is captured by two different diagnostics). Cross validation is used to help filter out spurious relationships, and redundancy is avoided by a stepwise regression procedure, as detailed below.

More formally, the method exploits relationships between a climate response variable y and a set of m diagnostics of the present climate x_j, where $j = 1, 2, \ldots, m$. For a set of n climate models, the multiple linear regression of the relation can be written in matrix form:

$$\mathbf{Y} = \mathbf{1}\beta_0 + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{Y} = (\begin{matrix} y_1 & y_2 & \ldots & y_n \end{matrix})^{\mathrm{T}}$ is the vector of the climate response variables in the model projection (a superscript T denotes the transpose), $\mathbf{1} = (\begin{matrix} 1 & 1 & \cdots & 1 \end{matrix})^{\mathrm{T}}$ is a column vector of size n,

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,1} & \cdots & x_{n,m} \end{pmatrix}$$

is the matrix of diagnostics, and $\boldsymbol{\varepsilon}$ is the vector of independent random variables of size n representing the uncertainty in the projections. The parameters β_0 and $\boldsymbol{\beta}$ of the multiple regression represented in Eq. (1), where $\boldsymbol{\beta}$ is a column vector of size m, are estimated by a least squares fit. A key additional assumption for MDER is that the relationship defined by Eq. (1) and parameters estimated from the model ensemble simulations also hold for the true climate—and not just for the climate models. Under this assumption, Eq. (1) can be used to estimate the climate response y_0, given the vector of observed diagnostics $\mathbf{X}_0$:

$$\hat{y}_0 = \hat{\beta}_0 + \mathbf{X}_0^{\mathrm{T}}\hat{\boldsymbol{\beta}}, \tag{2}$$

where the hatted quantities indicate that a variable is the best fit determined from the regression analysis.

The selection of the diagnostics x_j in MDER is done in a two-step process. First, physical processes that are expected to influence the climate response y must be identified. A set of diagnostics representing these processes are selected based on expert judgement, as discussed in section 2b. This step is necessarily subjective, and Eyring et al. (2005) and Bracegirdle et al. (2015) provide practical examples of diagnostic selections. Second, a stepwise regression procedure (von Storch and Zwiers 1999) is applied in order to choose only a subset of diagnostics for the multiple linear regression that contribute significantly to intermodel variation in the climate response y. In the stepwise regression, diagnostics are iteratively added to and removed from the regression model given by Eq. (1). This will continue until the regression sum of squares is not further increased by adding more diagnostics according to an F test, with the level of significance chosen in this study being $p = 0.05$. A more detailed description of the stepwise regression can be found in von Storch and Zwiers (1999).

An example of a model-weighting strategy that uses only the first (subjective) step for diagnostic selection is given by Waugh and Eyring (2008). However, as discussed in Räisänen et al. (2010), Bracegirdle and Stephenson (2012), and Karpechko et al. (2013), it is not necessary that all the subjectively selected diagnostics play a discernible role in climate response or contribute significantly to intermodel spread in the response. As a result, the statistical model in Eq. (1) may become

overfitted and not necessarily provide the best estimate of the climate response.

For example, Karpechko et al. (2013) initially selected 19 diagnostics known to be relevant to stratospheric ozone under present-day conditions, but only 1–4 diagnostics, depending on the forecast period, were selected by the stepwise algorithm during the second step (i.e., $m \leq 4$ in their study). Similarly, Räisänen et al. (2010) found that up to four diagnostics could be added to the regression model before overfitting problems started to emerge. Räisänen et al. (2010) applied a multiple regression model, as in Eq. (1), to diagnose the climate response in surface air temperature but used ad hoc diagnostics that were not necessarily directly related to physically relevant processes.

To assess whether projections following from the MDER algorithm may be susceptible to overfitting, we perform a cross-validation strategy (Michaelsen 1987). In the field of weather forecasting, one can test a predictive model against subsequent observations, but clearly we cannot wait to verify climate model projections. Thus, we perform cross validation in a "pseudo reality," where one model at a time is chosen to represent reality (hence the term pseudo reality) and withdrawn from the model ensemble. As a measure of prediction error, a squared difference between the projected future jet position and the jet change in this pseudo reality is calculated for both MDER and uMMM approaches. The process is repeated n times, once using each model as the pseudo reality, and the resulting root-mean-square error (RMSE) quantifies the accuracy of the prediction.

Diagnostics that have been known to impact on the austral jet stream are discussed in the following subsection and listed in Table 1. The MDER method and the calculation of the key process-oriented diagnostics for austral jet position were implemented in the earth system model evaluation tool (ESMValTool; Eyring et al. 2015), and individual results of the diagnostics calculated from models and observations or reanalyses are shown in the supporting information. The austral jet position is calculated as the December–February (DJF) latitude of maximum zonal mean zonal wind at 850 hPa between 30° and 80°S, following Son et al. (2009). To diagnose the exact latitude of the maximum zonal mean zonal wind, a parabolic fit around the three points of maximum wind speed was calculated for each time step.

b. Key process-oriented diagnostics for austral jet position

Several processes have been linked to the austral jet position in the literature. For most diagnostics, we include both the climatological value (appended by "_c") and the linear trend (appended by "_t") over the observation period, which we defined to be 1979–2005. An exception is the ASR-SH diagnostic, which was defined only for a shorter period (2000–05), resulting from a lack of observations before 2000. The choice of 1979–2005 restricts us to the satellite era, where we have some confidence in the reanalyses, and ends with the historical scenario in the CMIP5. The precise definition of each diagnostic, its value in the reanalysis or observational dataset, and its multimodel mean value from the CMIP5 ensemble are listed in Table 1. The values from each individual model and the observational or reanalysis datasets are presented in the supplemental material (Figs. S1–S11).

In the list below we briefly justify the inclusion of each diagnostic in our analysis. Note, however, that the vast majority of the diagnostics will not ultimately be utilized by MDER to predict future jet position. This is largely because many diagnostics are correlated with each other (e.g., biases in the climatological position of the jet stream are highly correlated with biases in the natural variability; Kidston and Gerber 2010). The acronyms defined in the list below are used in the figures and are also specified in Table 1.

- Stratospheric ozone at 50 hPa averaged over Antarctica (O3-SP) directly captures differences in the strength of the ozone hole and recovery (Eyring et al. 2013). Many models used the Cionni et al. (2011) dataset generated by Stratosphere–Troposphere Processes and their Role in Climate (SPARC), a few models interactively simulated ozone, and others used datasets generated by related chemistry–climate models.
- The near-global mean ozone at 50 hPa (O3-NGlob) provides a complementary measure of ozone loss and recovery and impacts near-global lower-stratospheric temperatures trends in particular (Eyring et al. 2013).
- Antarctic stratospheric temperature at 100 hPa (T-SP) is another indicator of ozone change (depletion or recovery). Because of differences in models' radiation schemes and dynamical feedbacks, models with the same ozone can simulate different thermal trends despite having the same underlying ozone. The radiative cooling in the lower stratosphere due to ozone depletion results in an enhanced temperature gradient in the upper troposphere–lower stratosphere (UTLS) and therefore accelerates the austral jet (Wilcox et al. 2012). Gerber and Son (2014) found variance in T-SP to be a significant source of spread in CMIP5 models in both the historical and future scenario integrations.
- The near-global mean temperature at 100 hPa (T-NGlob) is again a complementary measure of

TABLE 1. Description of the diagnostics, reanalysis, or observational data used to constrain the models, their mean value and uncertainty for the diagnostic, the corresponding value in the CMIP5 ensemble, and a reference. The trend and climatological mean of the diagnostics are calculated for the period 1979–2005, except for the ASR-SH diagnostic, which was calculated for the period 2000–05.

Acronym	Diagnostic	Reanalysis or observations	Reanalysis or observational value	CMIP5 mean ± std dev	References
Impact of Antarctic ozone depletion on the position of the jet stream					
T-SP_t	Trend in October–January (ONDJ) polar stratospheric temperatures at 100 hPa over Antarctica (60°–90°S)	ERA-Interim (Dee et al. 2011)	-1.17 ± 0.63 K decade^{-1}	-1.40 ± 0.72 K decade^{-1}	Fig. 10d of Eyring et al. (2013); Fig. 10 of Gerber and Son (2014)
T-SP_c	As above, but for climatological mean	As above	219.47 ± 0.51 K	218.17 ± 2.75 K	
O3-SP_t	Trend and in September–December ozone at 50 hPa over Antarctica (60°–90°S)	Binary Database of Profiles (BDBP; Hassler et al. 2009)	-0.42 ± 0.05 ppmv decade^{-1}	-0.32 ± 0.10 ppmv decade^{-1}	Fig. 10c of Eyring et al. (2013)
O3-SP_c	As above, but for climatological mean	As above	2.02 ± 0.077 ppmv	2.49 ± 0.23 ppmv	
Impact of GHG warming and climate sensitivity on the position of the jet stream					
T-NGlob_t	Trend in annual mean near-global (82.5°S–82.5°N) temperature at 100 hPa	ERA-Interim (Dee et al. 2011)	-0.14 ± 0.062 K decade^{-1}	-0.09 ± 0.09 K decade^{-1}	Fig. 10b of Eyring et al. (2013)
T-NGlob_c	As above, but for climatological mean	As above	204.47 ± 0.05 K	205.88 ± 1.43 K	
O3-NGlob_t	Trend in annual-mean near-global (82.5°S–82.5°N) ozone at 50 hPa	BDBP (Hassler et al. 2009)	-0.13 ± 0.017 ppmv decade^{-1}	-0.05 ± 0.02 ppmv decade^{-1}	Fig. 10a of Eyring et al. (2013)
O3-NGlob_c	As above, but for climatological mean	As above	2.08 ± 0.02 ppmv	2.20 ± 0.13 ppmv	
T-Trop_t	Trend in DJF upper-tropospheric tropical (30°S–30°N) temperatures at 250 hPa	ERA-Interim (Dee et al. 2011)	0.29 ± 0.09 K decade^{-1}	0.38 ± 0.13 K decade^{-1}	Fig. 10f of Eyring et al. (2013); Fig. 10 of Gerber and Son (2014)
T-Trop_c	As above, but for climatological mean	As above	230.67 ± 0.08 K	229.23 ± 1.68 K	
U-Jet_t	Trend in DJF SH jet position at 850 hPa	ERA-Interim (Dee et al. 2011)	$-0.79° \pm 0.32°$ lat decade^{-1}	$-0.45° \pm 0.48°$ lat decade^{-1}	Fig. 10e of Eyring et al. (2013)
U-Jet_c	As above, but for climatological mean	As above	$-50.02° \pm 0.27°$ lat	$-48.49° \pm 2.32°$ lat	
H-SH_t	Trend of the location of the SH Hadley cell boundary defined by $\Psi = 0$ at 500 hPa	ERA-Interim (Dee et al. 2011)	$-0.65° \pm 0.18°$ lat decade^{-1}	$-0.26° \pm 0.22°$ lat decade^{-1}	Fig. 5e of Son et al. (2010)
H-SH_c	As above, but for climatological mean	As above	$-36.27° \pm 0.17°$ lat	$-35.59° \pm 1.59°$ lat	
P-SH_t	Trend of the extratropical zonal mean tropopause pressure integrated south of 50°S	ERA-Interim (Dee et al. 2011)	-0.16 ± 0.07 hPa decade^{-1}	-0.32 ± 0.17 hPa decade^{-1}	Fig. 5c of Son et al. (2010)
P-SH_c	As above, but for climatological mean	As above	280.11 ± 0.62 hPa	252.18 ± 13.77 hPa	
SAM_efold_c	Climatological mean of the e-folding time scale of SAM in the troposphere	ERA-Interim (Dee et al. 2011)	12 ± 0.84 days	24.19 ± 10.21 days	Fig. 1c of Kidston and Gerber (2010)
ASR-SH_c	Climatological index of the meridional gradient in absorbed solar radiation (ASR) throughout the atmosphere	CERES EBAF (Doelling et al. 2013)	136.38 ± 16.83	130.78 ± 6.57	Ceppi et al. (2014)
Impact of Antarctic sea ice on SH winds and the position of the jet stream					
SIE-SP_t	Trend of annual mean Antarctic sea ice extent	NSIDC (Cavalieri et al. 1996)	$0.068 \pm 0.109 \times 10^6$ km^2 decade^{-1}	$-0.04 \pm 0.05 \times 10^6$ km^2 decade^{-1}	Fig. 3b of Stroeve et al. (2012)
SIE-SP_c	As above, but for climatological mean	As above	$12.17 \pm 0.06 \times 10^6$ km^2	$11.15 \pm 4.38 \times 10^6$ km^2	

stratospheric trends, seeking to identify differences between the models that are not confined to the polar cap.

- Changes in upper-tropospheric temperatures in the tropics at 250 hPa (T-Trop) influence temperature gradients in the UTLS (Wilcox et al. 2012) and were also a key driver of model spread in the analysis of Gerber and Son (2014). Upper-tropospheric temperatures in the tropics are influenced by both changes in surface temperatures and changes in the atmospheric stability.

- The historical DJF SH jet position at 850 hPa (U-Jet) has been found to correlate with a model's response (Kidston and Gerber 2010). This could reflect geometric constraints on the circulation (Barnes and Polvani 2013) and/or differences in the dynamics of the jet with latitude (Garfinkel et al. 2013). Recent trends in the jet also provide a measure of how sensitive the jets are to forcings and may also reflect natural variability, as discussed in section 5.

- Along with U-Jet, the latitude of the SH Hadley cell boundary defined by $\Psi = 0$ at 500 hPa (H-SH) gives us information about circulation biases and trends associated with ozone depletion over the past period (Son et al. 2010), where Ψ denotes the meridional streamfunction.

- A decrease in extratropical zonal mean tropopause pressure integrated south of 50°S (P-SH) is associated with warming of the troposphere and cooling of the lower stratosphere (two signatures of global warming) and has been strongly linked to the position of the extratropical jet streams (Lorenz and DeWeaver 2007).

- The e-folding time scale of a model's southern annular mode (SAM) in the troposphere (SAM-efold) characterizes the strength of interactions between baroclinic eddies and the extratropical jet stream (Lorenz and Hartmann 2001; Gerber et al. 2008a). Fluctuation dissipation theory suggests that the time scales of natural variability may be related to the response to external forcing (Gerber et al. 2008b; Ring and Plumb 2008), and there is evidence for this in comprehensive climate models (Kidston and Gerber 2010; Son et al. 2010; Barnes and Polvani 2013).

- Ceppi et al. (2014) link changes in the jet stream to changes in the meridional gradient of SH ASR throughout the atmosphere (ASR-SH). Changes in the ASR gradient can force changes in the equator-to-pole temperature gradient, directly impacting the baroclinicity of the atmosphere.

- Changes and biases in the climatological mean sea ice extent in the Southern Ocean (SIE-SP) impact the local energy budget and could influence the equator-to-pole temperature gradient (Stroeve et al. 2012; Ceppi et al. 2014; Bracegirdle et al. 2015).

3. Models used and observational and reanalysis constraints

The MDER method was applied to 28 models of the CMIP5 ensemble, as listed in Table 2, created and run by 18 different modeling centers. Many centers provided multiple ensemble member integrations of the same model and scenario. We use all the available ensemble members, which helps reduce the impact of natural variability. In order not to bias the MDER method toward models that ran more ensemble integrations, we first average all ensemble members for each individual model together prior to the calculations. Hence, MDER sees only one historical and future (RCP4.5) time series for each model. Only models that provided output for all process-oriented present-day diagnostics are included in the analysis because the method does not allow for missing values (Karpechko et al. 2013).

The future trends in the austral jet position were calculated from monthly means from the RCP4.5 scenario integrations, which are forced by changing GHGs concentrations, but also include aerosol, ozone, and land-use changes and natural forcings (Taylor et al. 2012). The present-day diagnostics were calculated from the monthly mean CMIP5 historical simulations, in general for the period 1979–2005 (see details in Table 1), and results are shown in the supplemental material. Each of the present-day diagnostics is compared with monthly mean reanalysis data or observations as listed in Table 1.

Direct measurements are used in the diagnostics where available, but for many diagnostics we had to rely on meteorological reanalysis. For the evaluation, monthly means for the period 1979–2005 are used except for the zonal means of net balanced climatology top-of-atmosphere (TOA) fluxes, which are only available for the period 2000–14. A list of the reanalysis data or observations used in this study is given in Table 1.

4. Application of MDER to projections of the summertime austral jet position

To highlight how the most important factors constraining the jet stream evolve over time, we apply MDER to two time horizons. We first focus on the jet position in the near term from 2015–34. A 20-yr period was selected to reduce the influence of natural variability in the jet stream. Over this short time horizon, no significant changes in anthropogenic forcings occur in the RCP4.5 scenario, so we expect the method to focus on correcting biases in the historical climatologies. We then focus the method on a midcentury projection

TABLE 2. Overview of CMIP5 models that are used in this study, including the number of ensembles for which concentration scenarios were simulated by each model. (Expansions of acronyms are available at http://www.ametsoc.org/PubsAcronymList.)

No.	Model	Modeling center	RCP4.5	Main reference
01	ACCESS1.0	Centre for Australian Weather and Climate Research,	1	Dix et al. (2013)
02	ACCESS1.3	Australia	1	
03	BCC_CSM1.1	Beijing Climate Center, China Meteorological Administration,	1	Wu (2012)
04	BCC_CSM1.1(m)	China	1	
05	BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University, China	1	—
06	CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	5	Arora et al. (2011)
07	CCSM4	National Center for Atmospheric Research, United States	5	Meehl et al. (2012)
08	CESM1(BGC)	Community Earth System Model contributors	1	Gent et al. (2011)
09	CESM1(CAM5)		3	
10	CMCC-CMS	Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy	1	Vichi et al. (2011)
11	CNRM-CM5	Centre National de Recherches Météorologiques, France	1	Voldoire et al. (2013)
12	CSIRO Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation in collaboration with Queensland Climate Change Centre of Excellence, Australia	10	Rotstayn et al. (2012)
13	FGOALS-g2	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, and Center for Earth System Science, China	1	Li et al. (2013)
14	GFDL CM3	NOAA/Geophysical Fluid Dynamics Laboratory, United	1	Donner et al. (2011)
15	GFDL-ESM2G	States	1	Dunne et al. (2013)
16	GFDL-ESM2M		1	
17	HadGEM2-AO	National Institute of Meteorological Research, Korea Meteorological Administration, South Korea	1	Martin et al. (2011)
18	INM-CM4.0	Institute of Numerical Mathematics, Russia	1	Volodin et al. (2010)
19	IPSL-CM5A-LR	L'Institut Pierre-Simon Laplace, France	4	Dufresne et al. (2013)
20	IPSL-CM5A-MR		1	
21	IPSL-CM5B-LR		1	
22	MIROC5	Japan Agency for Marine-Earth Science and Technology,	3	Watanabe et al. (2011)
23	MIROC-ESM	Atmosphere and Ocean Research Institute (The University	1	Watanabe et al. (2011)
24	MIROC-ESM-CHEM	of Tokyo), and National Institute for Environmental Studies, Japan	1	Watanabe et al. (2011)
25	MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	3	Giorgetta et al. (2013)
26	MPI-ESM-MR		3	
27	MRI-CGCM3	Meteorological Research Institute, Japan	1	Yukimoto et al. (2012)
28	NorESM1-M	Norwegian Climate Centre, Norway	1	Iversen et al. (2012)

(2040–59), a time when the stratospheric ozone and greenhouse gas concentrations have changed.

a. Near-term projections of the austral jet position

Figure 2a shows the absolute value of the correlation coefficients between the short-term projection of the austral jet position and our 20 process-oriented present-day diagnostics. The coefficients reveal a strong correlation between U-Jet_c and the near-term projection of the austral jet position. The correlation coefficient is near unity with a tight uncertainty envelope, as quantified by the 95% prediction interval. Models simulating the jet equatorward too far in the historical simulations (see Fig. S6) also do so for the near term, and vice versa. The high correlation between the historical and the projected austral jet position will cause the MDER algorithm to recognize and correct for this well-known equatorial bias in the CMIP5 model ensemble.

The climatological mean of the Hadley cell boundary (Fig. S7) position ($r = 0.90$) and trend ($r = 0.58$) are also highly correlated with the jet position from 2015 to 2034, although the relationship is of opposite sign for the trend. Biases in the position of the SH Hadley cell mirror biases in the extratropical jet stream (Son et al. 2010; Arblaster and Meehl 2006), such that the first relationship is strongly linked to the connection with U-Jet_c discussed above. At face value, the negative correlation between the near-term jet position and H-SH_t suggests that models that saw more expansion of the tropics in the late twentieth century tend to have a more equatorward jet in coming decades. Given that the near-term jet is so highly correlated with the jet in the past, this could reflect the fact that models with an equatorward bias in their climatology are more sensitive to external forcing (and so exhibited larger trends in the twentieth century), as found by Kidston and Gerber (2010) for
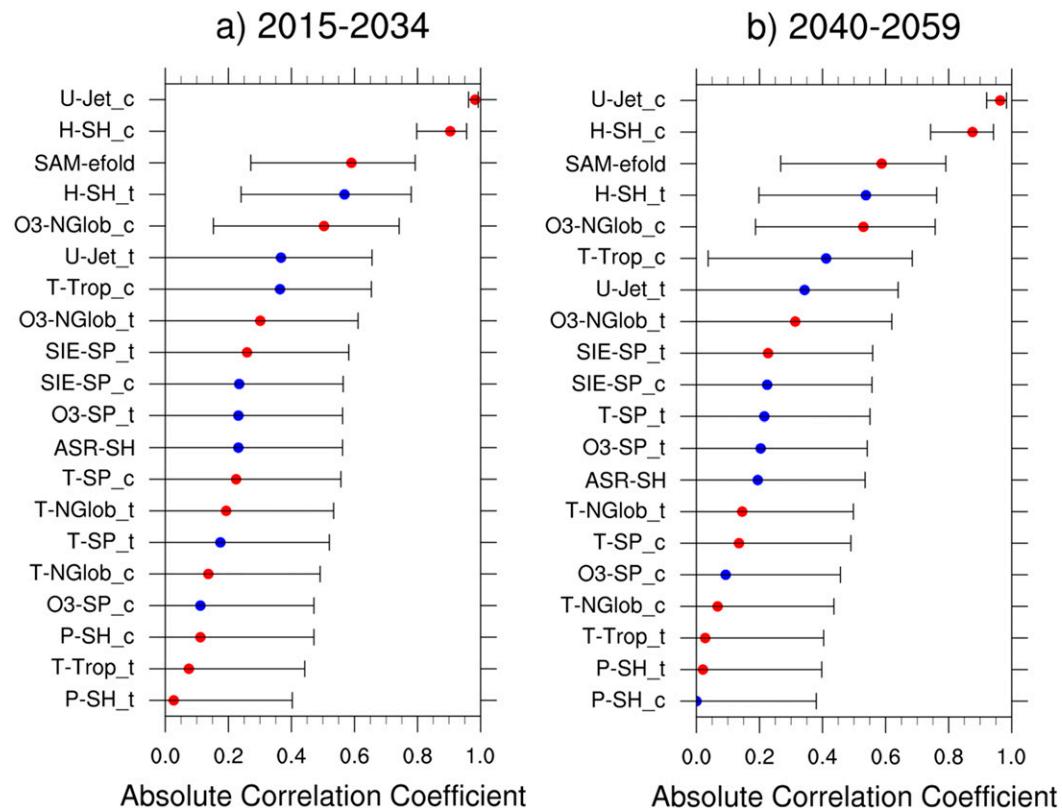
FIG. 2. Absolute values of the correlation coefficient between future austral jet position and present-day diagnostics as listed in Table 1 across the CMIP5 model ensemble (see Table 2), for (a) the near-term austral jet position climatological mean (2015–34) and (b) the midterm austral jet position climatological mean (2040–59). Error bars show the 95% prediction intervals for the correlation coefficients. Colored markers indicate positive (red) and negative (blue) correlations.

future jet shifts. The late twentieth-century U-Jet_t is also negatively correlated with the 2015–34 jet position, albeit more weakly. It is unclear to us why the trend in the Hadley cell is more strongly associated with jet position than the trend in the jet itself.

The e-folding time scale of SAM (Fig. S10) also exhibits a statistically significant positive correlation ($r =$ 0.59) with the near-term projection of the austral jet. As in the case of the Hadley cell, the SAM e-folding time scale is linked to U-Jet_c (e.g., Kidston and Gerber 2010) and so again may be a manifestation of the same relationship. Since the H-SH and SAM-efold diagnostics ultimately provide somewhat redundant information compared to the diagnostic U-Jet_c, the MDER algorithm rejects them from the regression model.

The diagnostic of near-global climatological mean ozone (Fig. S1) shows the fifth-highest correlation, and the link is statistically significant ($r =$ 0.50) at the 95% confidence level. The correlation could reflect the fact that models experiencing larger ozone loss over the historical period (and so exhibit a climatology with less

ozone) also experienced a stronger ozone hole and thus a poleward shift in the jet stream (Eyring et al. 2013).

The remaining correlations in Fig. 2a are not statistically significant at the 95% level of the linear regression. In general, however, diagnostics indicating biases in the SH circulation climatology show a stronger correlation to the near-term austral jet stream position than diagnostics that characterize trends over the historical period.

From all the diagnostics included, the MDER algorithm creates a parsimonious regression model to predict the near-term austral jet position, focusing exclusively on the diagnostic U-Jet_c, as shown in Fig. 3a. The model is simply $-1.36 + 0.98 \times$ U-Jet_c. In essence, the algorithm detects the equatorward bias of the CMIP5 models in the jet stream in the past and provides a correction to the future projection. As the result depends on a single parameter, Fig. 3a can be compared quite easily with our schematic diagram in Fig. 1. MDER focuses on the nearly perfect correlation between U-Jet_c and jet location in 2015–34. The
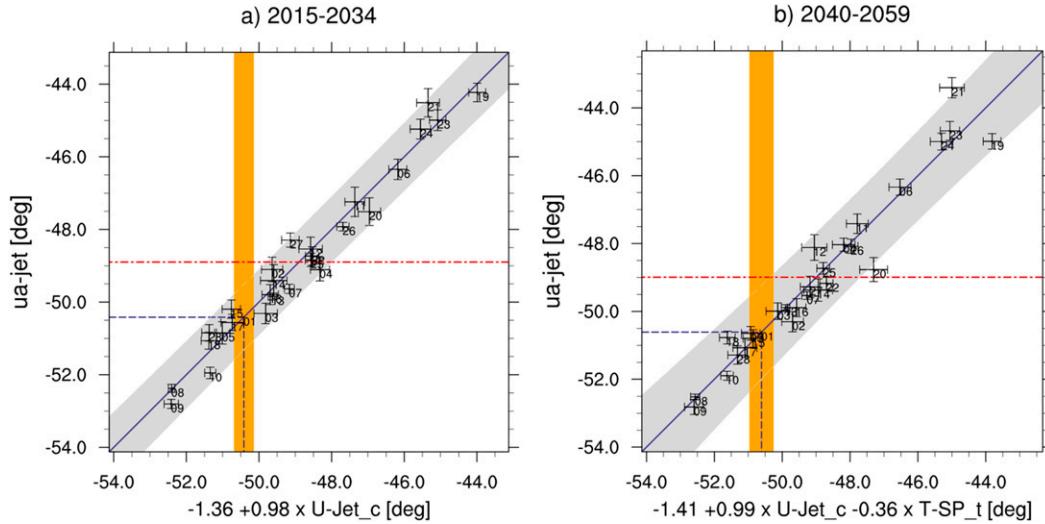
FIG. 3. Scatterplot showing the correlation between the future austral jet position and (a) the quantity ($-1.36 + 0.98 \times$ U-Jet_c) for the near-term climatological mean (2015–34) and (b) the quantity ($-1.41 + 0.99 \times$ U-Jet_c $- 0.36 \times$ T-SP_t) for the midterm climatological mean (2040–59). Numbers indicate estimates of simulated climatological mean values of each CMIP5 model and the error bars show one standard deviation of the means, calculated from seasonal means. The solid blue line shows the least squares linear fit to the CMIP5 model ensemble and the gray shading marks the 95% prediction interval for the least squares linear regression. The orange shading indicates one standard deviation of the observed climatological mean values calculated using historical values. The red dotted line shows the uMMM and the blue dashed line the MDER prediction.

uMMM projection puts the jet at 48.9°S (red horizontal line), but knowing that the historical jet was biased in the CMIP5 models (located on average at 48.5°S instead of 50.0°S), MDER suggests that it should also be 1.5° poleward of the uMMM in 2015–34, at 50.4°S, as indicated by the blue dashed lines.

While the result is almost trivial, this is the first time, to our knowledge, that projections of the future multimodel jet position have been bias corrected. Taking the uMMM would place the jet at 48.9°S over the period 2015–34, substantially equatorward of its current position in reanalysis. MDER suggests that it should be at 50.4°S, just a bit poleward of its current location.

Cross validation of the results indicates that MDER can reduce uncertainty in the jet projection. This is realized by comparing the results of future austral jet position estimates with the MDER method against the uMMM in pseudo reality, following Karpechko et al. (2013). The root-mean-square error of the projection of the near-term austral jet positions is nearly an order of magnitude lower using the MDER method compared to uMMM (Fig. 4; RMSE$_{MDER}$ = 0.42° latitude and RMSE$_{uMMM}$ = 2.37° latitude). This dramatic drop in uncertainty in the cross validation can be understood more easily by viewing the time series of the jet position, shown in Fig. 5. In the cross-validation test with an uMMM methodology, one is effectively seeking to predict one model's jet position (i.e., the pseudo reality)

using the positions projected by all the other models. The RMSE$_{uMMM}$ thus reflects the spread in the mean jet position from 2015 to 2034, a spread on the order of degrees. The errors are large because the uMMM cannot successfully predict cases when the pseudo reality is
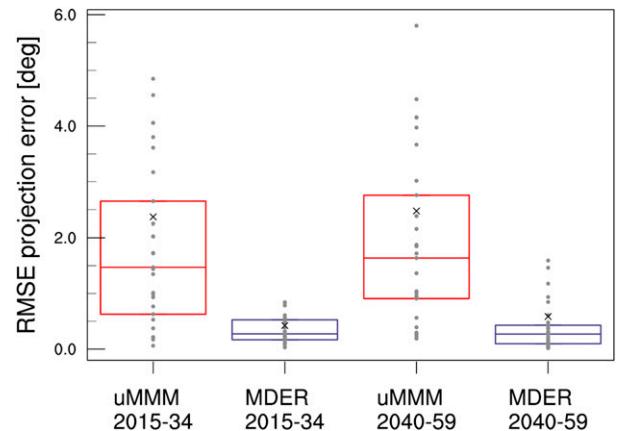

FIG. 4. RMSE differences between the ensemble mean future climatological mean (2015–34 and 2040–59) austral jet position and the future climatological mean austral jet position in pseudo reality for each pseudo reality considered (gray dots) under the RCP4.5. The ensemble mean is calculated for each scenario from the uMMM (red boxes) and the MDER method (blue boxes). The cross indicates the RMSE for each case and the boxes show the 25th–75th percentiles across the error ensemble. The lines inside the box indicate the median of the ensemble.
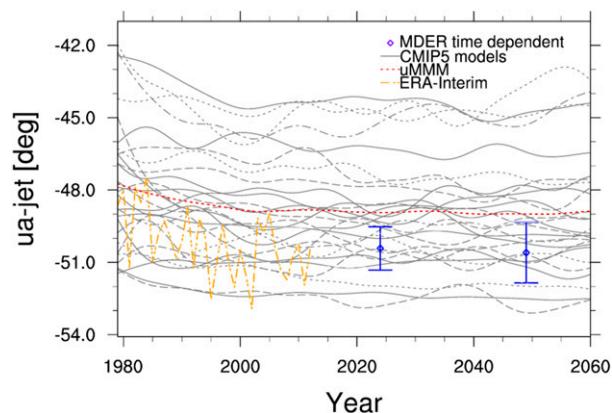
FIG. 5. Time series of the austral jet position for RCP4.5 scenario between 1980 and 2100. Gray lines show the individual models (iteratively smoothed with a 1–2–1 filter, repeated 30 times, to reduce the noise) and the red dotted line the uMMM across all CMIP5 models in Table 2. Diamonds show the predicted mean estimate resulting from the MDER analysis, for the near-term (2015–34) and midterm (2040–59) climatological mean austral jet position. Error bars indicate the 95% prediction interval of the regression analysis. The orange dashed line shows the reanalysis data from ERA-Interim.

an outlier model. With MDER, however, we explicitly take into account information on the historical jet position in the model chosen as the pseudo reality and only use the other models to estimate the jet shift between 1979–2005 and 2015–34. For this short time horizon, the forced signal is small, on the order of 0.1°.

We should emphasize that the RMSE bounds obtained in the cross-validation exercise provide a nice illustration of the actual prediction errors associated with uMMM and MDER. Formal estimates of the prediction errors from the full model ensemble further demonstrate how the prediction uncertainty is reduced by MDER in comparison to uMMM. Based on 28 realizations of climate change under the RCP 4.5 scenario, the 95% prediction intervals for MDER and uMMM methods are 0.8° and 4.8°, respectively. Here, the MDER error is calculated in a standard way as a prediction interval for the response variable of regression [e.g., Karpechko et al. 2013; their Eq. (6)]. For uMMM the corresponding prediction interval is given by $t_{(1+\tilde{p})/2}s$, where s is the standard deviation across individual model projections, $t_{(1+\tilde{p})/2}$ is the $(1 + \tilde{p})/2$ quantile of Student's t distribution, and $\tilde{p} = 0.95$. The MDER uncertainty is calculated assuming perfect knowledge of the observed diagnostics.

A more realistic uncertainty bound should reflect both uncertainty in the multimodel estimate of the climate signal (in the case of MDER, uncertainty in the change between 1979–2005 and 2015–34) and uncertainty associated with calculation of the diagnostics. The latter is

affected by reanalysis errors and internal variability. While reanalysis errors can only be estimated qualitatively (see discussion in section 5), the influence of the internal variability can be directly incorporated into the prediction uncertainty. In 27 years of reanalysis, the mean jet can only be bounded to the range 50.0° ± 0.5° with 95% confidence. When uncertainty associated with internal variability is taken into account (by the law of error propagation), the uncertainty of MDER prediction becomes 1° latitude, still considerably less than the uncertainty of uMMM method.

b. Midterm projections of the austral jet position

A key finding from our application of MDER to the near-term jet position is that the climatological biases in CMIP5 historical integrations are larger than any of the shifts predicted in the next two decades. We next apply the MDER to midterm (2040–59) jet position where the forcing signal is larger. As we will show, however, the mean trends in the jet remain small, likely due to the fact that stratospheric ozone loss and greenhouse gas increases tend to oppose each other in coming decades (e.g., Perlwitz et al. 2008; Son et al. 2008). Nonetheless, MDER suggests that we can glean more information than a simple bias correction when focusing on longer-term projections.

Figure 2b illustrates correlations between the process-oriented diagnostics and the midterm austral jet projections. Even at midcentury, SH circulation biases in the historical integrations are still the most important. The top five diagnostics with the strongest correlations to midterm austral jet positions are the same as those for the near term. The importance of the remaining 15 process-oriented diagnostics has changed, although those correlation coefficients are generally not statistically significant.

Despite the similarities in the correlation structure, MDER obtains a more complex result for the midterm projection. The method initially constructs a regression model ($-1.66 + 1.02 \times$ U-Jet_c $- 0.40 \times$ T-SP_t $- 0.10 \times$ T-SP_c) involving three diagnostics: U-Jet_c, T-SP_t, and T-SP_c. While the U-Jet_c term can again be interpreted as a bias correction of the austral position in the CMIP5 models, the T-SP terms indicate that the diagnostics associated with the formation of the ozone in the historical period can be used to improve future projections of the jet position.

The negative sign of the T-SP_t term reflects the fact that models that experienced larger stratospheric cooling over the historical period tend to exhibit a more equatorward shift of the jet in the future. Wilcox et al. (2012) and Gerber and Son (2014) found that models with more cooling over the polar cap tend to

experience a more poleward shift in the jet, suggesting that the jet is responding to the equator-to-pole temperature gradient in the upper troposphere–lower stratosphere. Here, the relationship has changed sign because we are comparing cooling over the historical period to an equatorward shift in the future. Models that experienced a strong thermodynamic response to ozone loss in the past are likely to have an equal and opposite response to ozone recovery in the future (i.e., more warming) and so a more equatorward jet shift. T-SP_t can thus be acting as a proxy for the strength of ozone loss and recovery, a key driver of austral jet shifts. We emphasize, however, that it is the temperature response to ozone loss that appears to be crucial. The regression model picks T-SP_t over the actual historical trend in ozone, O3-SP_t, even though both statistics are nearly equally correlated with future jet position. Many models used similar ozone data (Cionni et al. 2011) but do not exhibit a uniform thermal response because of differences in their radiation schemes.

We were concerned that the negative sign of the correlation with T-SP_c could reflect a similar connection to the ozone hole, as ozone depletion already occurred over the entire historical period (1979–2005); a colder historical climatology is indicative of a larger ozone hole. It is thus unclear how the climatology would contain information independent from the polar cap temperature trend, which raises the danger that MDER could be overfitting the diagnostics. To avoid inclusion of redundant information with unclear physical interpretation, we recalculated the regression model, intentionally removing the T-SP_c diagnostic, and obtained the following result: $-1.41 + 0.99 \times$ U-Jet_c $- 0.36 \times$ T-SP_t. The difference between the projections made by these two models is 0.2° latitude, much smaller than the uncertainty of either statistical model (see below). Based on further cross-validation tests (not shown), we believe the simple model is more robust and apply it in Fig. 3b. It incorporates two physically justified constraints: a correction for biases in the climatological jet position and a correction based on the intensity of thermodynamic response to stratospheric ozone loss.

Figure 4 shows also the cross-validation tests for the midrange jet projection. As one might expect, the RMSE$_{MDER}$ prediction error (0.59°) is larger for the mid-twenty-first-century case than for the near-term analysis (where it was 0.42° latitude) but still more than 4 times less than the uMMM prediction error (RMSE$_{uMMM}$ = 2.47° latitude). Again, the key is that the shifts in the jet stream, even 50 years away, are small relative to the biases in the models' historical climatology. As noted in the discussion of section 4a, the RMSEs reflect our uncertainty in light of 28 realizations of the

future and do not account for uncertainty in a jet associated with a single realization, as will be the case with our one earth.

From the regression model in Fig. 3b, the MDER analysis predicts an austral jet stream position for the midterm climatological mean of 50.6°S, implying a mean shift of 0.2° southward compared to the 2015–34 position of the austral jet (or 0.6° latitude southward from its historical climatology). The uMMM projection, 50.0°S, suggests a small southward shift from the 2015–34 mean as well, but only by 0.1° latitude. Note that this is still northward of the jet location in historical reanalysis; naïvely comparing the future projection with historical reanalysis would give one the opposite trend.

In our near-term application, MDER took the shift in the uMMM projection and bias corrected for the mean jet location. With inclusion of information on stratospheric polar cap temperature trends, MDER modifies the jet trend as well. We emphasize, however, that this modification (and the total trends themselves) is very small relative to the 1.5° latitude bias in the models' historical jet position climatology. The trends are also small relative to uncertainty in the jet position associated with natural variability; given 1979–2005 reanalysis data, we can only say that the mean jet position was within $50.0° \pm 0.5°$S with 95% confidence.

5. Summary and discussion

We have used a multiple diagnostic ensemble regression (MDER) algorithm to analyze the austral jet position in projections of the twenty-first century under the RCP4.5 scenario, a moderate carbon future. MDER allowed us to incorporate 20 process-oriented constraints from observations and reanalysis to improve upon the unweighted multimodel mean (uMMM) projection. The method can be interpreted as a reweighting of models based on biases in their historical climatologies (Karpechko et al. 2013).

We first applied the MDER method to the near-term climatological mean (2015–34) of the austral jet position. The method removed the equatorward bias in the jet stream, suggesting that the best estimate of its future position should be 1.5° latitude southward of that found in the uMMM projection (48.9°S). We next focused on a midcentury austral jet stream projection, a target period of 2040–59. In addition to the same need to correct for the climatological jet position bias, MDER found that lower-stratospheric polar cap temperature trends over the historical period could be used to effectively discriminate future trends. From a physical standpoint, historical temperature trends are an indicator of the intensity of the ozone hole. It is likely that models with

more intense cooling over the historical period of ozone loss will experience more intense warming as ozone recovery and hence a more equatorward shift in the jet stream as it responds to changes in the upper troposphere–lower stratosphere temperature gradient.

Expected shifts in the jet stream in coming decades are generally small, on the order of 0.1°, owing in part to cancellation between the impacts of stratospheric ozone recovery and increased greenhouse gas loading (e.g., Perlwitz et al. 2008). Biases in some models' climatological jet position, on the other hand, are on the order of degrees, and the multimodel mean position is 1.5° latitude poleward of that found in ERA-Interim. Thus, a naïve use of the uMMM to project the mean jet position in the near or midterm places the future jet equatorward of its current position, even though most models project that it should shift slightly poleward over this period. While this bias correction is a fairly straightforward result, it is, to our knowledge, the first effort to account for this bias in future projections.

Getting the jet in the right place has significant implications. First, it is collocated with the storm track and thus tightly linked with the boundary between the subtropical dry zone and extratropical precipitation maximum. Shifts in the jet have significant impacts on regional precipitation (e.g., Kang et al. 2011; Thompson et al. 2011), and it is critical that regional modeling efforts to downscale climate information from global models account for this bias. Second, the surface wind stress associated with the jet stream plays a key role in the overturning circulation of the ocean (Waugh et al. 2013). Biases in the austral jet position limit our ability to accurately model the heat and carbon uptake of the deep ocean (Swart and Fyfe 2012).

Given these large model biases, an alternative approach would be to first compute the jet shift from the historical period to the future using the models and then to simply add this to the historical climatology based on reanalyses (e.g., Räisänen 2007). MDER effectively led to this result for the near-term projection. This change-based approach, however, relies on the explicit assumption that biases in simulated present-day and future climates remain constant (i.e., that the jet shift only depends on the applied forcing and is independent of present jet positions). MDER does not make this assumption, and it did make a difference (albeit a small one) for the midterm projection.

Our regression model for the midrange jet projection suggests that we can use a historical trend in polar stratospheric temperatures to better estimate the future jet position. Constraining this trend with reanalysis, however, is problematic, as changes in the observational network can lead to spurious trends. Calvo et al. (2012)
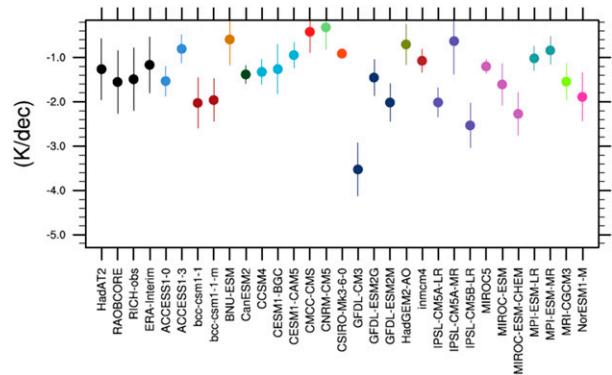

FIG. 6. Trends in ONDJ temperature anomalies at 100 hPa over Antarctica for radiosonde data (HadAT2, RAOBCORE, and RICH-obs), ERA-Interim, and the individual models of the CMIP5 ensemble. Vertical lines indicate the sample standard deviation of the mean value.

suggest that Antarctic lower-stratospheric cooling due to ozone depletion (T-SP_t) may be underestimated by ERA-Interim by as much as a factor of 2 compared to radiosonde observations. On the other hand, the interannual variability of the temperatures is so large that the discrepancy between trend estimates based on ERA-Interim and radiosondes is within statistical uncertainty (Calvo et al. 2012).

To test this for our study, Fig. 6 compares the T-SP diagnostics derived from the CMIP5 models with ERA-Interim data and the radiosonde observations that were analyzed by Young et al. (2013): Hadley Centre Atmospheric Temperatures, version 2 (HadAT2; Thorne et al. 2005); Iterative Universal Kriging (IUK; Sherwood et al. 2008); Radiosonde Observation Correction Using Reanalyses (RAOBCORE, version 1.5; Haimberger et al. 2008); and Radiosonde Innovation Composite Homogenization (RICH) comparing the observations of a tested time series with observations of neighboring radiosonde time series (RICH-obs, version 1.5; Haimberger et al. 2012). For DJF 1979–2005 considered in our study, the mean trend in ERA-Interim is approximately $-1.4\,\mathrm{K\,decade^{-1}}$, and so slightly smaller than that in the radiosonde datasets, where the trends vary between -1.6 and $-2.2\,\mathrm{K\,decade^{-1}}$. The ERA-Interim trend, however, is still mostly within the given observational uncertainty. We also found that the ERA-Interim climatology (Fig. S4, bottom) is very similar to the radiosonde climatology.

The focus of MDER on different time periods provides additional insight into which physical processes are important for projections at the midterm horizon. In the near term, diagnostics focused on biases in the climatology are most important. At midcentury, uncertainty associated with stratospheric ozone trends also becomes

important. Toward the end of the century, when the ozone hole is mostly recovered, uncertainty in tropical warming trends begins to appear in the MDER results (not shown). The tropical warming trends over the historical period give an indication of how sensitive a model is to greenhouse gas warming; models that warm more over the historical period tend to warm more in the future and so project greater circulation trends. We did not present these results here, however, because of the lack of reliable direct measurements of upper-tropospheric temperature trends. Our study thus emphasizes the need for reliable long-term climate records, which may prove critical for constraining future model projections.

REFERENCES

Arblaster, J. M., and G. A. Meehl, 2006: Contributions of external forcings to southern annular mode trends. *J. Climate*, **19**, 2896–2905, doi:10.1175/JCLI3774.1.

——, ——, and D. J. Karoly, 2011: Future climate change in the Southern Hemisphere: Competing effects of ozone and greenhouse gases. *Geophys. Res. Lett.*, **38**, L02701, doi:10.1029/2010GL045384.

Arora, V. K., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, doi:10.1029/2010GL046270.

Barnes, E. A., and L. Polvani, 2013: Response of the midlatitude jets, and of their variability, to increased greenhouse gases in the CMIP5 models. *J. Climate*, **26**, 7117–7135, doi:10.1175/JCLI-D-12-00536.1.

Bracegirdle, T. J., and D. B. Stephenson, 2012: Higher precision estimates of regional polar warming by ensemble regression of climate model projections. *Climate Dyn.*, **39**, 2805–2821, doi:10.1007/s00382-012-1330-3.

——, and Coauthors, 2015: A multi-disciplinary perspective on climate model evaluation for Antarctica. *Bull. Amer. Meteor. Soc.*, doi:10.1175/BAMS-D-15-00108.1, in press.

Calvo, N., R. R. Garcia, D. R. Marsh, M. J. Mills, D. E. Kinnison, and P. J. Young, 2012: Reconciling modeled and observed temperature trends over Antarctica. *Geophys. Res. Lett.*, **39**, L16803, doi:10.1029/2012GL052526.

Cavalieri, D. J., C. L. Parkinson, P. Gloersen, and H. Zwally, 1996: Sea ice concentrations from *Nimbus-7* SMMR and DMSP SSM/I-SSMIS passive microwave data, version 1. NASA National Snow and Ice Data Center, accessed May 2015, doi:10.5067/8GQ8LZQVL0VL.

Ceppi, P., M. D. Zelinka, and D. L. Hartmann, 2014: The response of the Southern Hemispheric eddy-driven jet to future changes in shortwave radiation in CMIP5. *Geophys. Res. Lett.*, **41**, 3244–3250, doi:10.1002/2014GL060043.

Cionni, I., and Coauthors, 2011: Ozone database in support of CMIP5 simulations: Results and corresponding radiative forcing. *Atmos. Chem. Phys.*, **11**, 11 267–11 292, doi:10.5194/acp-11-11267-2011.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

Dix, M., and Coauthors, 2013: The ACCESS coupled model: Documentation of core CMIP5 simulations and initial results. *Aust. Meteor. Oceanogr. J.*, **63**, 83–99.

Doelling, D. R., and Coauthors, 2013: Geostationary enhanced temporal interpolation for CERES flux products. *J. Atmos. Oceanic Technol.*, **30**, 1072–1090, doi:10.1175/JTECH-D-12-00136.1.

Donner, L. J., and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Climate*, **24**, 3484–3519, doi:10.1175/2011JCLI3955.1.

Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth system model: From CMIP3 to CMIP5. *Climate Dyn.*, **40**, 2123–2165, doi:10.1007/s00382-012-1636-1.

Dunne, J. P., and Coauthors, 2013: GFDL's ESM2 global coupled climate–carbon Earth system models. Part II: Carbon system formulation and baseline simulation characteristics. *J. Climate*, **26**, 2247–2267, doi:10.1175/JCLI-D-12-00150.1.

Eyring, V., and Coauthors, 2005: A strategy for process-oriented validation of coupled chemistry–climate models. *Bull. Amer. Meteor. Soc.*, **86**, 1117–1133, doi:10.1175/BAMS-86-8-1117.

——, and Coauthors, 2013: Long-term ozone changes and associated climate impacts in CMIP5 simulations. *J. Geophys. Res. Atmos.*, **118**, 5029–5060, doi:10.1002/jgrd.50316.

——, and Coauthors, 2015: ESMValTool (v1.0)—A community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. *Geosci. Model Dev. Discuss.*, **8**, 7541–7661, doi:10.5194/gmdd-8-7541-2015.

Garfinkel, C. I., D. W. Waugh, and E. P. Gerber, 2013: The effect of tropospheric jet latitude on coupling between the stratospheric polar vortex and the troposphere. *J. Climate*, **26**, 2077–2095, doi:10.1175/JCLI-D-12-00301.1.

Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, doi:10.1175/2011JCLI4083.1.

Gerber, E. P., and S. W. Son, 2014: Quantifying the summertime response of the austral jet stream and Hadley cell to stratospheric ozone and greenhouse gases. *J. Climate*, **27**, 5538–5559, doi:10.1175/JCLI-D-13-00539.1.

——, L. M. Polvani, and D. Ancukiewicz, 2008a: Annular mode time scales in the intergovernmental panel on climate change fourth assessment report models. *Geophys. Res. Lett.*, **35**, L22707, doi:10.1029/2008GL035712.

——, S. Voronin, and L. M. Polvani, 2008b: Testing the annular mode autocorrelation time scale in simple atmospheric general circulation models. *Mon. Wea. Rev.*, **136**, 1523–1536, doi:10.1175/2007MWR2211.1.

Gillett, N., and D. Thompson, 2003: Simulation of recent Southern Hemisphere climate change. *Science*, **302**, 273–275, doi:10.1126/science.1087440.

Giorgetta, M. A., and Coauthors, 2013: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model intercomparison project phase 5. *J. Adv. Model. Earth Syst.*, **5**, 572–597, doi:10.1002/jame.20038.

Haimberger, L., C. Tavolato, and S. Sperka, 2008: Toward elimination of the warm bias in historic radiosonde temperature records—some new results from a comprehensive intercomparison of upper-air data. *J. Climate*, **21**, 4587–4606, doi:10.1175/2008JCLI1929.1.

——, ——, and ——, 2012: Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *J. Climate*, **25**, 8108–8131, doi:10.1175/JCLI-D-11-00668.1.

Hassler, B., G. E. Bodeker, I. Cionni, and M. Dameris, 2009: A vertically resolved, monthly mean, ozone database from 1979 to 2100 for constraining global climate model simulations. *Int. J. Remote Sens.*, **30**, 4009–4018, doi:10.1080/01431160902821874.

Iversen, T., and Coauthors, 2012: The Norwegian Earth System Model, NorESM1-M—Part 2: Climate response and scenario projections. *Geosci. Model Dev. Discuss.*, **5**, 2933–2998, doi:10.5194/gmdd-5-2933-2012.

Kang, S. M., L. M. Polvani, J. C. Fyfe, and M. Sigmond, 2011: Impact of polar ozone depletion on subtropical precipitation. *Science*, **332**, 951–954, doi:10.1126/science.1202131.

Karpechko, A. Yu., D. Maraun, and V. Eyring, 2013: Improving Antarctic total ozone projections by a process-oriented multiple diagnostic ensemble regression. *J. Atmos. Sci.*, **70**, 3959–3976, doi:10.1175/JAS-D-13-071.1.

Kidston, J., and E. P. Gerber, 2010: Intermodel variability of the poleward shift of the austral jet stream in the CMIP3 integrations linked to biases in 20th century climatology. *Geophys. Res. Lett.*, **37**, L09708, doi:10.1029/2010GL042873.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2010: Challenges in combining projections from multiple climate models. *J. Climate*, **23**, 2739–2758, doi:10.1175/2009JCLI3361.1.

Kushner, P. J., I. M. Held, and T. L. Delworth, 2001: Southern Hemisphere atmospheric circulation response to global warming. *J. Climate*, **14**, 2238–2249, doi:10.1175/1520-0442(2001)014<0001:SHACRT>2.0.CO;2.

Li, L., and Coauthors, 2013: The flexible global ocean–atmosphere–land system model, grid-point version 2: FGOALS-g2. *Adv. Atmos. Sci.*, **30**, 543–560, doi:10.1007/s00376-012-2140-6.

Lorenz, D. J., and D. L. Hartmann, 2001: Eddy–zonal flow feedback in the Southern Hemisphere. *J. Atmos. Sci.*, **58**, 3312–3327, doi:10.1175/1520-0469(2001)058<3312:EZFFIT>2.0.CO;2.

——, and E. T. DeWeaver, 2007: Tropopause height and zonal wind response to global warming in the IPCC scenario integrations. *J. Geophys. Res.*, **112**, D10119, doi:10.1029/2006JD008087.

Marshall, G. J., 2003: Trends in the southern annular mode from observations and reanalyses. *J. Climate*, **16**, 4134–4143, doi:10.1175/1520-0442(2003)016<4134:TITSAM>2.0.CO;2.

Martin, G. M., and Coauthors, 2011: The HadGEM2 family of Met Office Unified Model climate configurations. *Geosci. Model Dev.*, **4**, 723–757, doi:10.5194/gmd-4-723-2011.

Meehl, G. A., and Coauthors, 2012: Climate system response to external forcings and climate change projections in CCSM4. *J. Climate*, **25**, 3661–3683, doi:10.1175/JCLI-D-11-00240.1.

Michaelsen, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.*, **26**, 1589–1600, doi:10.1175/1520-0450(1987)026<1589:CVISCF>2.0.CO;2.

Perlwitz, J., S. Pawson, R. L. Fogt, J. E. Nielsen, and W. D. Neff, 2008: Impact of stratospheric ozone hole recovery on Antarctic climate. *Geophys. Res. Lett.*, **35**, L08714, doi:10.1029/2008GL033317.

Polvani, L. M., D. W. Waugh, G. J. P. Correa, and S.-W. Son, 2011: Stratospheric ozone depletion: The main driver of twentieth-century atmospheric circulation changes in the Southern Hemisphere. *J. Climate*, **24**, 795–812, doi:10.1175/2010JCLI3772.1.

Räisänen, J., 2007: How reliable are climate models? *Tellus*, **59A**, 2–29, doi:10.1111/j.1600-0870.2006.00211.x.

——, L. Ruokolainen, and J. Ylhäisi, 2010: Weighting of model results for improving best estimates of climate change. *Climate Dyn.*, **35**, 407–422, doi:10.1007/s00382-009-0659-8.

Ring, M. J., and R. A. Plumb, 2008: The response of a simplified GCM to axisymmetric forcings: Applicability of the fluctuation–dissipation theorem. *J. Atmos. Sci.*, **65**, 3880–3898, doi:10.1175/2008JAS2773.1.

Rotstayn, L. D., S. J. Jeffrey, M. A. Collier, S. M. Dravitzki, A. C. Hirst, J. I. Syktus, and K. K. Wong, 2012: Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: A study using single-forcing climate simulations. *Atmos. Chem. Phys.*, **12**, 6377–6404, doi:10.5194/acp-12-6377-2012.

Shepherd, T. G., 2014: Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.*, **7**, 703–708, doi:10.1038/ngeo2253.

Sherwood, S. C., C. L. Meyer, R. J. Allen, and H. A. Titchner, 2008: Robust tropospheric warming revealed by iteratively homogenized radiosonde data. *J. Climate*, **21**, 5336–5350, doi:10.1175/2008JCLI2320.1.

Simpkins, G. R., and A. Yu. Karpechko, 2012: Sensitivity of the southern annular mode to greenhouse gas emission scenarios. *Climate Dyn.*, **38**, 563–572, doi:10.1007/s00382-011-1121-2.

Son, S.-W., and Coauthors, 2008: The impact of stratospheric ozone recovery on the Southern Hemisphere westerly jet. *Science*, **320**, 1486–1489, doi:10.1126/science.1155939.

——, N. F. Tandon, L. M. Polvani, and D. W. Waugh, 2009: Ozone hole and Southern Hemisphere climate change. *Geophys. Res. Lett.*, **36**, L15705, doi:10.1029/2009GL038671.

——, and Coauthors, 2010: Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel

assessment. *J. Geophys. Res.*, **115**, D00M07, doi:10.1029/2010JD014271.

Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier, 2012: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophys. Res. Lett.*, **39**, L16502, doi:10.1029/2012GL052676.

Swart, N. C., and J. C. Fyfe, 2012: Observed and simulated changes in the Southern Hemisphere surface westerly wind-stress. *Geophys. Res. Lett.*, **39**, L16711, doi:10.1029/2012GL052810.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Thompson, D. W. J., S. Solomon, P. J. Kushner, M. H. England, K. M. Grise, and D. J. Karoly, 2011: Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nat. Geosci.*, **4**, 741–749, doi:10.1038/ngeo1296.

Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper air temperatures from 1958 to 2002. *J. Geophys. Res.*, **110**, D18105, doi:10.1029/2004JD005753.

Vichi, M., E. Manzini, P. G. Fogli, A. Alessandri, L. Patara, E. Scoccimarro, S. Masina, and A. Navarra, 2011: Global and regional ocean carbon uptake and climate change: Sensitivity to a substantial mitigation scenario. *Climate Dyn.*, **37**, 1929–1947, doi:10.1007/s00382-011-1079-0.

Voldoire, A., and Coauthors, 2013: The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dyn.*, **40**, 2091–2121, doi:10.1007/s00382-011-1259-y.

Volodin, E. M., N. A. Dianskii, and A. V. Gusev, 2010: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izv. Atmos. Oceanic Phys.*, **46**, 414–431, doi:10.1134/S000143381004002X.

von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research.* Cambridge University Press, 494 pp.

Watanabe, S., and Coauthors, 2011: MIROC-ESM: Model description and basic results of CMIP5-20c3m experiments. *Geosci. Model Dev. Discuss.*, **4**, 1063–1128, doi:10.5194/gmdd-4-1063-2011.

Waugh, D. W., and V. Eyring, 2008: Quantitative performance metrics for stratospheric-resolving chemistry–climate models. *Atmos. Chem. Phys. Discuss.*, **8**, 10 873–10 911, doi:10.5194/acpd-8-10873-2008.

——, F. Primeau, T. Devries, and M. Holzer, 2013: Recent changes in the ventilation of the southern oceans. *Science*, **339**, 568–570, doi:10.1126/science.1225411.

Wilcox, L. J., A. J. Charlton-Perez, and L. J. Gray, 2012: Trends in austral jet position in ensembles of high- and low-top CMIP5 models. *J. Geophys. Res.*, **117**, D13115, doi:10.1029/2012JD017597.

Wu, T. W., 2012: A mass-flux cumulus parameterization scheme for large-scale models: Description and test with observations. *Climate Dyn.*, **38**, 725–744, doi:10.1007/s00382-011-0995-3.

Yin, J. H., 2005: A consistent poleward shift of the storm tracks in simulations of 21st century climate. *Geophys. Res. Lett.*, **32**, L18701, doi:10.1029/2005GL023684.

Young, P. J., A. H. Butler, N. Calvo, L. Haimberger, P. J. Kushner, D. R. Marsh, W. J. Randel, and K. H. Rosenlof, 2013: Agreement in late twentieth century Southern Hemisphere stratospheric temperature trends in observations and CCMVal-2, CMIP3, and CMIP5 models. *J. Geophys. Res. Atmos.*, **118**, 605–613, doi:10.1002/jgrd.50126.

Yukimoto, S., and Coauthors, 2012: A new global climate model of the Meteorological Research Institute: MRI-CGCM3—model description and basic performance. *J. Meteor. Soc. Japan*, **90A**, 23–64, doi:10.2151/jmsj.2012-A02.