Historical SAM Variability. Part I: Century-Length Seasonal Reconstructions*

JULIE M. JONES,[+] RYAN L. FOGT,[#,@] MARTIN WIDMANN,[&] GARETH J. MARSHALL,[**]
PHIL D. JONES,[++] AND MARTIN VISBECK[##]

[+] *Department of Geography, University of Sheffield, Sheffield, United Kingdom, and GKSS Research Centre, Geesthacht, Germany*
[#] *Polar Meteorology Group, Byrd Polar Research Center, The Ohio State University, Columbus, Ohio*
[&] *School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom, and GKSS Research Centre, Geesthacht, Germany*
[**] *British Antarctic Survey, Cambridge, United Kingdom*
[++] *Climatic Research Unit, University of East Anglia, Norwich, United Kingdom*
[##] *Leibniz Institute for Marine Sciences (IFM-Geomar), Kiel, Germany*

ABSTRACT

Seasonal reconstructions of the Southern Hemisphere annular mode (SAM) index are derived to extend the record before the reanalysis period, using station sea level pressure (SLP) data as predictors. Two reconstructions using different predictands are obtained: one [Jones and Widmann (JW)] based on the first principal component (PC) of extratropical SLP and the other (Fogt) on the index of Marshall. A regional-based SAM index (Visbeck) is also considered.

These predictands agree well post-1979; correlations decline in all seasons except austral summer for the full series starting in 1958. Predictand agreement is strongest in spring and summer; hence agreement between the reconstructions is highest in these seasons. The less zonally symmetric SAM structure in winter and spring influences the strength of the SAM signal over land areas, hence the number of stations included in the reconstructions. Reconstructions from 1865 were, therefore, derived in summer and autumn and from 1905 in winter and spring.

This paper examines the skill of each reconstruction by comparison with observations and reanalysis data. Some of the individual peaks in the reconstructions, such as the most recent in austral summer, represent a full hemispheric SAM pattern, while others are caused by regional SLP anomalies over the locations of the predictors. The JW and Fogt reconstructions are of similar quality in summer and autumn, while in winter and spring the Marshall index is better reconstructed by Fogt than the PC index is by JW. In spring and autumn the SAM shows considerable variability prior to recent decades.

1. Introduction

The Southern Hemisphere annular mode (SAM) (or Antarctic Oscillation) is the dominant mode of extratropical atmospheric circulation in the SH (Kidson 1988; Limpasuvan and Hartmann 2000; Thompson and Wallace 2000). It is approximately zonally symmetric

---

Corresponding author address: Julie M. Jones, Department of Geography, University of Sheffield, Winter Street, Sheffield S10 2TN, United Kingdom.
E-mail: julie.jones@sheffield.ac.uk

and characterizes fluctuations in the strength and location of the eddy-driven jet (Codron 2005). A positive (negative) index represents negative (positive) high-latitude and positive (negative) midlatitude pressure anomalies and hence stronger (weaker) westerly circumpolar flow. Because it is a hemispheric climate mode, the SAM influences diverse aspects of Southern Hemisphere climate, including temperatures over the Antarctic (e.g., Kwok and Comiso 2002; Marshall 2007; Schneider et al. 2004), temperature and precipitation in the Southern Hemisphere midlatitudes (Jones and Widmann 2003, hereafter JW03; Gillett et al. 2006), and sea ice and ocean circulation (e.g., Sen Gupta and England 2006; Hall and Visbeck 2002; Lefebvre et al. 2004).

Two definitions of the SAM index have been used. One is the first principal component (PC) of Southern
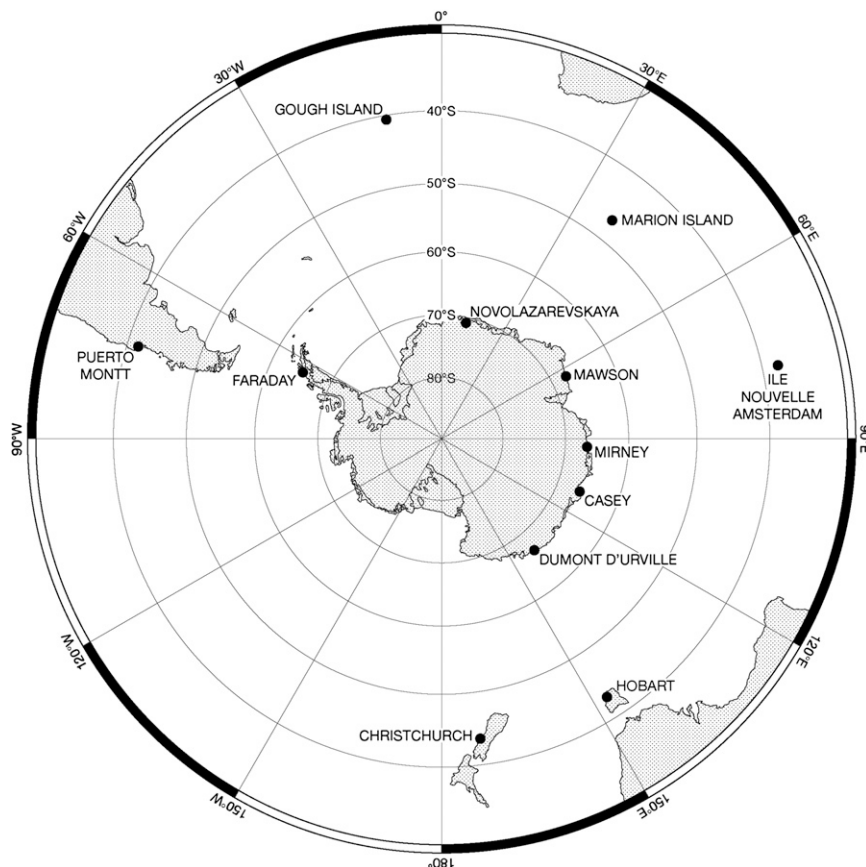
FIG. 1. The stations used by Marshall (2003) to estimate the SAM index.

Hemisphere extratropical sea level pressure (SLP), geopotential height, or zonal winds (Thompson and Wallace 2000). The other is the difference in normalized zonal mean pressure between 40° and 65°S (Gong and Wang 1999). As there is no "correct" definition of the SAM index, it is important to know to what extent these different definitions lead to differences in the indices. A complicating factor is that the main hemispheric data available for the last 50 yr come from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) (Kalnay et al. 1996) reanalysis and from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) (Uppala et al. 2005). These datasets are known, particularly in the nonsummer seasons, to have problems in data-sparse regions such as the high-latitude SH before 1979 when extensive satellite data were first assimilated (Bromwich et al. 2007; Bromwich and Fogt 2004; Hines et al. 2000; Marshall 2002). Thus uncertainties exist in the SAM indices and hence in any trends calculated using reanalysis data prior to 1979. However, Bromwich et al. (2007) also find considerable differences in the SAM

trends from various datasets after 1979, reflecting the differences between reanalyses even after this period. To provide a more reliable estimate of the SAM index for recent decades, Marshall (2003, hereafter M03) derived a SAM index that approximates the Gong and Wang (1999) definition by calculating the difference in mid- and high-latitude proxy zonal mean SLP based on SLP observations around 40° and 65°S, starting in 1957 and updated to the present (see online at http://www.antarctica.ac.uk/met/gjma/sam.html); this is hereafter referred to as the Marshall index. The stations used are shown in Fig. 1.

The SAM has attracted much interest because of statistically significant positive trends in recent decades during austral summer and autumn (e.g., Marshall 2007). Studies investigating these trends have mostly concentrated on anthropogenic causes, specifically stratospheric ozone depletion (e.g., Thompson and Solomon 2002; Gillett and Thompson 2003; Miller et al. 2006; Perlwitz et al. 2008) and greenhouse gas emissions (e.g., Stone et al. 2001; Kushner et al. 2001; Miller et al. 2006). To determine the significance of these recent trends it is necessary to establish the magnitude of SAM changes

during previous decades. JW03 and Jones and Widmann (2004, hereafter JW04) reconstructed the austral summer [November–January (NDJ) and December–January (DJ)] SAM index using station SLP for the periods 1878–2000 and 1905–2000, respectively. JW03 also presented a reconstruction for the NDJ SAM index based on tree-ring-width chronologies that extends from 1743 to 2001. The DJ SAM index reconstructions indicated that the trends in recent decades are not unprecedented, and thus natural climate forcings and internal variability can also strongly influence the state of the SAM (JW04).

In this paper we derive and compare reconstructions for the four standard seasons: austral summer [December–February (DJF)] and autumn [March–May (MAM)] back to 1865 and winter [June–August (JJA)] and spring [September–November (SON)] back to 1905 [shorter because of reconstruction quality issues (section 3b)], derived using principal component regression (PCR). One set of reconstructions [hereafter the Jones and Widmann (JW) reconstructions] use the first PC of extratropical SLP as predictand, while the other (hereafter the Fogt reconstruction) uses the Marshall index. The index definitions are based on weighted SLP anomalies. These defining weight patterns differ between the indices and between seasons. For the JW and Fogt reconstructions SLP anomalies are not weighted according to the defining patterns but by using reconstruction weights that follow from the regression-based reconstruction technique and can be seen as a result of the defining weight patterns and the spatial correlation structure of the SLP field. These reconstruction weights are also closely related to the correlation patterns of the local SLP with the different indices, which for the purpose of this paper we call the SAM signal of the index and which also depend on the defining weights and the spatial correlation structure of the SLP field. We also compare these reconstructions to those of Visbeck (2009), which use fixed and predefined weights.

The aims of this paper are as follows:

(i) to analyze the temporal variability of the PC-based, Gong and Wang, and Marshall indices separately for each season, and to analyze the link between differences in the variability and the defining weight patterns;

(ii) to determine how the quality of the reconstructions is related to the correlations between local SLP and the SAM indices and to the spatial distribution of the predictors;

(iii) to determine whether there is a "best" reconstruction (based on fitting statistics and validation on independent data); and

(iv) to determine the SAM behavior over the past 150 yr based on a joint analysis of the Fogt, JW, and Visbeck reconstructions, in light of their uncertainties and robust features.

These objectives are valuable not only with respect to observed SAM behavior but also for model-based studies of SAM variability. In this regard, the reconstructions described herein are used in a companion paper (Fogt et al. 2009, hereafter Part II) to evaluate the SAM as simulated by 17 simulations from the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (Meehl et al. 2007).

2. Data and methods

a. Data

The hemispheric SLP data were taken from the ERA-40 reanalysis (Uppala et al. 2005) and cover the period 1958–2001. For computational ease they were regridded from the original N80 Gaussian grid to a $5° \times 5°$ grid. Gridded SLP data were also taken from the Second Hadley Centre Sea Level Pressure dataset (HadSLP2) (Allan and Ansell 2006), which is based on an interpolation of global land and marine pressure observations.

The primary source of station SLP data is stations collected and digitized by Jones (1987) and Jones et al. (1999). Data were also obtained from the NCAR Data Support Section (DSS) 570.0 dataset and the National Climatic Data Center (NCDC) Global Historical Climatology Network (GHCN). Additional stations were obtained from Rob Allan and Tara Ansell (Allan and Ansell 2006). Antarctic station data are from the Reference Antarctic Data for Environmental Research (READER) database (Turner et al. 2004).

b. SAM indices

A PC-based SAM index (hereafter, ERA40PC) was defined as the first PC of ERA-40 seasonal mean SLP for the domain 20°–80°S and the SAM pattern as the first EOF of these data. Because of uncertainties in the early ERA-40 data (Bromwich and Fogt 2004; Marshall 2003), the EOFs were calculated based on the 1979–2001 period (detrended), and the PCs calculated by projecting the SLP anomalies (nondetrended) for the full period on the seasonal EOFs. The ERA-40 Gong and Wang index (ERA40GW) was calculated as the difference of area-weighted normalized zonal mean ERA-40 SLP at 40° and 65°S. The Marshall index approximates the Gong and Wang (1999) index using zonal means from six station pressure records located at approximately 40°S and six stations at 65°S (Fig. 1). This reconstruction uses

a similar number of stations to the Fogt and longer JW reconstructions.

The SAM index calculated from HadSLP2 is considered in Part II, where it is shown to be markedly different from the reconstructed indices prior to 1957, when most Antarctic data first became available. The differences are potentially due to uncertainties in the dataset away from the input station data (Jones and Lister 2007), particularly in the southeastern Pacific Ocean (Allan and Ansell 2006). However, these data are used for deriving spatial anomaly patterns (as well as ERA-40 after 1958) to provide an estimate of the hemispheric anomalies at points away from observations.

c. Methods

Both JW and Fogt employ PCR, as used previously for reconstructions of the austral summer SAM (JW03 and JW04). The main difference is in choice of predictand: the JW reconstructions use the ERA40PC SAM index while the Fogt reconstructions use the Marshall index.

1) JW RECONSTRUCTIONS

The JW reconstructions extend the austral summer reconstructions of JW03 and JW04 to include all seasons. The full ERA-40 period of 1958–2001 was used for model fitting, as 1979–2001 was deemed too short. Four different station networks were constructed (extending to 2005) with start dates of 1866, 1905, 1951, and 1958, hereafter termed JW1866, JW05, JW51, and JW58, respectively. Stations that were significantly correlated (at either the $p < 0.05$ or $p < 0.01$ level) with the ERA40PC SAM index were considered as predictors. The stations retained are listed in the appendix (Tables A1–A3) and displayed for each season in Figs. 3–6. In all cases we present the reconstruction that had the best fitting and validation statistics.

2) FOGT RECONSTRUCTION

The Fogt reconstruction uses only the long-term 1865 and 1905 station networks. Only stations significantly correlated with the Marshall index ($p < 0.05$) were used in the reconstruction. The stations retained are listed in the appendix (Table A4) and displayed in Figs. 3–6. The period 1957–2005 was used for model fitting. The Marshall index was further adjusted to have zero mean over the fitting period.

3) RECONSTRUCTION METHODOLOGY

PCR was used to derive both the Fogt and the JW reconstructions. The methodology follows that of JW03 and JW04, where further detail can be found. The model was fitted with both datasets detrended, while the model was applied to undetrended predictors. PCR was performed on the data from each of the networks, whereby the seasonal SAM indices were regressed onto the leading PCs of normalized station SLP. The optimal number of PCs to be retained was determined based on the correlation between detrended reconstructed and detrended true SAM index in an independent validation period.

To obtain the reconstructions, the station records were normalized by dividing by the standard deviation from the fitting period, and the PCs were calculated by projecting the normalized values on the EOFs defined in the fitting period. Multiplication of these PCs with the PCR weights derived from the fitting data yields the reconstructions. As noted in other studies (Cook et al. 1999; JW03; JW04), the reconstruction methodology can equivalently be expressed as a weighted sum of the station records.

To test the skill of the reconstruction methodology, a leave-one-out cross-validation procedure was employed. We performed the PCR methodology 44 (JW) or 49 (Fogt) times for each season, each time estimating a year not included in the fitting data. The individual years were then concatenated to obtain a verification SAM time series. To ensure independence of the individual years within this series, 2 yr on either side of the estimated year were left out during the PCR process.

4) VISBECK RECONSTRUCTIONS

The Visbeck SAM index from 1880 to 2002 was calculated using a two-step selection algorithm (Visbeck 2009). In a first pass all station on Antarctica and on the subpolar islands and southern ends of the continents that had more than 75% of valid monthly data between 1970 and 2000 and with a height below 950 hPa were selected from the SLP dataset of Jones (1991) and Jones et al. (1999). The remaining stations were sorted into four regions: a polar region between 90° and 60°S (AN) and three subtropical ring segments between 60° and 20°S from 10°W to 80°E [South Africa (AF)], 80°E to 120°W [Australia/New Zealand (AU)], and 120° to 10°W [South America (SA)]. For each region normalizing the mean SLP anomaly with the standard deviation generated a SLP-based index. From that a base SAM index was constructed by computing the difference between the mean subtropical indices (SA + AF + AU)/3 and Antarctica (AA).

In a second step only stations with correlations >0.3 with the preliminary SAM index for the subtropical regions and >0.7 for the Antarctic zone were retained and new regional averaged pressure anomalies and indices constructed. From the remaining 11 stations for Antarctica, 7 for South Africa, 12 for Australia–New Zealand, and 10 for South America a final station-based

SAM index was derived covering the time span from 1954 to 2005.

To arrive at an earlier SAM index during the time of very few Antarctic stations, Visbeck (2009) proposed a method to reconstruct the Antarctic SLP variability using the concept of atmospheric mass conservation between Antarctica and the subtropical latitudes. This allowed extending the index back in time to 1884. (A detailed description and the data are provided online at http://www.ifm-geomar.de/~sam).

To enable estimation of the quality of the Visbeck reconstruction using a varying number of stations (hereafter Vvar) earlier in the series (when less stations are included), reconstructions using a fixed number of stations (hereafter Vfixed) were also undertaken.

5) STANDARDIZATION OF RECONSTRUCTIONS

To compare the reconstructions, a common standardization was required as the original Fogt and JW reconstructions were calibrated against predictands with different standard deviations, and the Visbeck reconstructions were not calibrated against a predictand index. The JW and Fogt SAM indices describe SLP anomalies. The reference period, which defines the mean pressure field with respect to which the anomalies are calculated and which is also often used for scaling the variance of the index, also differs for the original indices. To have a consistent definition of the ERA40PC and GW indices, we will use here the same reference period 1979–2001 for all indices. For the JW and Fogt reconstructions and the Marshall index (assumed to be approximating ERA40GW) the scaling was chosen such that the predictand indices have unit variance and zero mean for 1979–2001. We did not choose an earlier start of the reference period because of uncertainty in the presatellite ERA-40 SLP. As the Visbeck reconstruction is not based on estimating a predictand, this normalization approach cannot be applied. The closest approximation to this idea can be obtained by scaling the Visbeck reconstructions such that their variance is the correlation of the Visbeck index with ERA40GW for the reference period. Note that the standardization means that the variance of the Visbeck reconstructions and the Marshall index are lower than in the figures of Visbeck (2009) and Marshall (2003), respectively.

3. Results

a. Comparison of the ERA-40 SAM indices and the Marshall and Visbeck indices

We first analyze the similarity of the observational and reanalysis SAM indices (i.e., ERA40PC, ERA40GW, and the Marshall index), as well as Vvar and Vfixed,
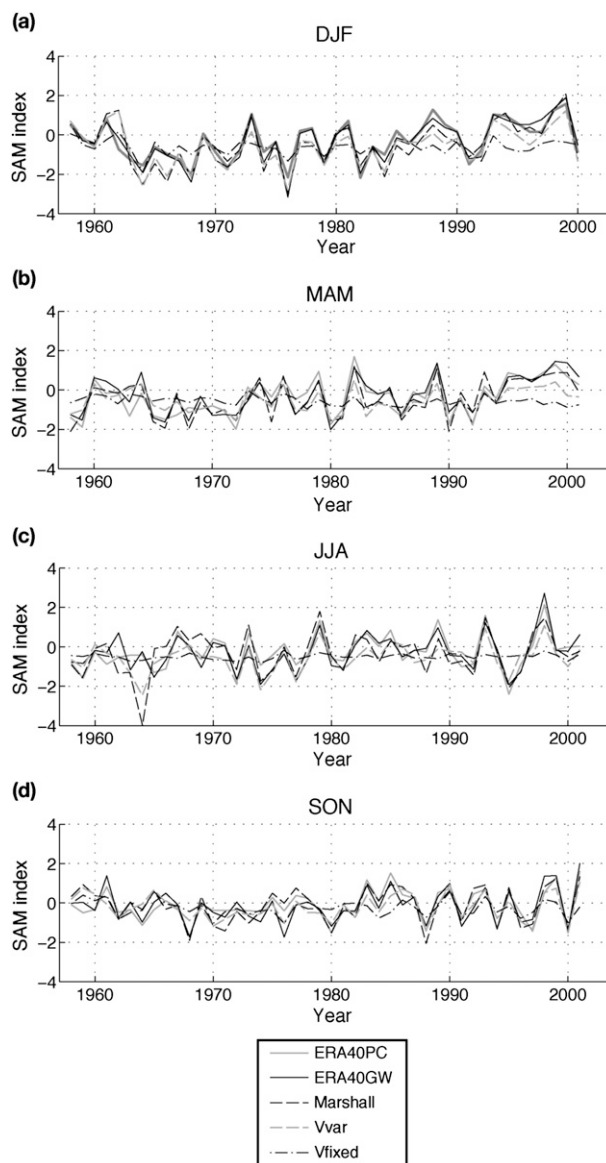

FIG. 2. The ERA-40, Marshall, and Visbeck SAM indices 1958–2001.

shown in Fig. 2. Table 1 presents the cross correlations during 1979–2001 and 1958–2001. The comparison for these different periods assesses whether the relationships change when the ERA-40 SAM indices are calculated from potentially less reliable reanalysis data. Spatial comparisons are undertaken in section 3b. Data were detrended over the calculation period for all analyses.

Table 1 indicates that the Marshall index approximates the GW SAM index excellently in DJF and SON for 1979–2001 (correlations of 0.91 in both seasons) and well in MAM and JJA (0.85 and 0.86, respectively). Correlations drop for the period 1958–2001, particularly

TABLE 1. Correlations between different indices of the SAM index (detrended). All values are statistically significant at the 1% level.

	ERA40PC/ERA40GW		Marshall/ERA40GW		Marshall/ERA40PC		Vvar/ERA40PC		Vfixed/ERA40PC		Vvar/ERA40GW		Vfixed/ERA40GW	
	79–01	58–01	79–01	58–01	79–01	58–01	79–01	58–01	79–01	58–01	79–01	58–01	79–01	58–01
DJF	0.96	0.95	0.91	0.89	0.88	0.84	0.91	0.85	0.52	0.61	0.92	0.91	0.59	0.69
MAM	0.95	0.87	0.85	0.87	0.92	0.81	0.93	0.84	0.61	0.52	0.83	0.82	0.47	0.45
JJA	0.95	0.87	0.86	0.60	0.87	0.65	0.82	0.67	0.28	0.26	0.81	0.73	0.34	0.40
SON	0.94	0.89	0.91	0.80	0.95	0.85	0.90	0.76	0.54	0.46	0.89	0.80	0.62	0.65

in JJA where values fall to 0.60 and SON with a fall from 0.91 to 0.80 (partly because of a strong negative Marshall index in JJA, which is not present in any of the ERA-40 indices; Fig. 2c). We note that this 1964 value is a result of anomalous SLP across both the SH mid- and high latitudes rather than some erroneous observation at one or more of the stations used to derive the Marshall index: it is believed to be a consequence of the Agung eruption of the previous year, which had a significant impact on SH climate (e.g., Angell 1988). Correlations are higher between ERA40PC and the Marshall index than between ERA40GW and the latter in JJA and SON, potentially because the PC is a weighted mean of all points between 20° and 80°S (with the weights given by the EOF loadings) and is therefore proportionally less dependent on reanalysis SLP in the high latitudes than ERA40GW (particularly as area-weighting reduces the influence of the high-latitude grid boxes). Thus, although the Marshall index is formulated as an approximation to the GW index, it is in some seasons actually closer to the PC-based index.

Comparison of Vvar and Vfixed with ERA40PC and ERA40GW shows how well the Visbeck method reconstructs the SAM index (equivalent to the validation correlations for Fogt and JW of Table 2). Vvar shows good agreement with both ERA40PC and ERA40GW in all seasons (Table 1), with lowest correlations in JJA. Agreement between Vfixed and the ERA-40 indices is considerably lower, particularly in MAM and JJA, with correlations with ERA40GW of 0.47 and 0.34, respectively. This suggests that the lower number of stations may not be sufficient to fully capture SAM variability. As the Visbeck reconstructions used here are standardized according to correlations with ERA40GW for the period 1979–2001 [section 2c(5)] where the correlations are higher because of the denser station network, the reconstructed variability may be overestimated during the earlier parts of Vvar (where this reconstruction is based on fewer stations).

The maximum agreement that can be expected between the JW and Fogt reconstructions is indicated by correlations between ERA40PC and the Marshall index for 1958–2001, as the JW reconstructions use ERA40PC as the predictand, and the Fogt reconstructions use the Marshall index. The maximum correlation is 0.84 (DJF and SON), with again JJA lowest (0.65). Thus one could expect, and indeed one sees, lower agreement in JJA between JW and Fogt (see sections 3c and 3d).

b. Model fitting and validation statistics and stations included

The fitting and validation statistics for the JW and Fogt reconstructions are shown in Table 2. The locations

TABLE 2. Fitting and validation statistics for the Jones and Widmann and Fogt reconstructions. The fitting–calibration period is 1958–2001 for the Jones and Widmann reconstructions and 1957–2005 for the Fogt reconstruction.

	Number of stations	Fitting correlation	Validation correlation	Reduction of error
DJF				
JW1866	16	0.84	0.80	0.63
JW1905	29	0.87	0.84	0.70
JW1951	40*	0.89	0.84	0.71
JW1958	49*	0.91	0.86	0.75
Fogt	9	0.84	0.81	0.65
MAM				
JW1866	12	0.75	0.74	0.55
JW1905	14*	0.82	0.80	0.63
JW1951	27*	0.85	0.83	0.70
JW1958	33*	0.90	0.89	0.78
Fogt	8	0.74	0.72	0.50
JJA				
JW1905	11	0.73	0.70	0.49
JW1951	16	0.81	0.77	0.60
JW1958	22	0.85	0.78	0.61
Fogt	16	0.84	0.80	0.62
SON				
JW1905	8	0.74	0.72	0.52
JW1951	15	0.78	0.73	0.54
JW1958	23	0.81	0.81	0.65
Fogt	10	0.86	0.83	0.67

* Reconstructions that use stations significantly correlated with predictand at the 1% level.

of the stations and their regression weights are shown in Figs. 3–6. These figures also include the SAM signals for comparison. Although the standard definition of the linear SLP signal would be the regression maps (e.g., Widmann 2005), it is more relevant for the purpose of this paper to consider correlation maps, as the correlations between local SLP and the SAM indices influence the selection of predictor stations as well as the reconstruction weights. The correlation maps are identical to the regression maps divided by the standard deviations of the local SLP. For the PC-based index, the regression map is proportional to SLP EOF1 (e.g., Bretherton et al. 1992), which is the defining weight pattern, whereas for the Marshall index there is no direct mathematical relationship between the defining weight pattern and the correlation maps. The correlations are based on the respective SAM index used for model calibration (detrended) and ERA-40 SLP 1979–2001 (detrended). The weights of the stations included in the Visbeck reconstructions are fixed and hence not discussed here.

As well as the fitting and validation correlations, the reduction of error (RE) is used to assess reconstruction quality. The RE compares the residuals from the reconstruction (validation series) with the residuals relative to an estimate based on no knowledge, which here is taken as the calibration period mean of the predictand (Fritts et al. 1990; Cook et al. 1999). The RE can have values from $-\infty$ to $+1$, with positive values indicating

skill relative to climatology, and a value of $+1$ indicating perfect reconstruction for the validation period.

The reconstruction quality that can be achieved differs between seasons, related to where the centers of high correlations between the SAM index and local SLP are located in relation to land areas containing measurement stations. In DJF (Fig. 3) and MAM (Fig. 4) the SAM has more regions of high correlations in midlatitudes located over the continents and New Zealand than in JJA and SON. In DJF (Fig. 3) the SAM is most annular, with zonally oriented bands of positive correlations at midlatitudes, a zero line at around 50°S, and negative correlations at high latitudes. In MAM the pattern is slightly less annular (Fig. 4), with low correlations extending northward in the eastern Pacific to north of 45°S. The SAM is least annular in JJA and SON (Figs. 5 and 6), and low correlations again extend northward into the southeast Pacific.

The seasonal differences in the SAM structure have been identified previously. Szeredi and Karoly (1987) found their EOF of station pressure representing "out of phase" structure between mid- and high latitudes to be less zonally symmetric in winter. Rogers and van Loon (1982) also found that midlatitude anomalies are greater over Chatham Island in winter (JJA) rather than over the central Indian Ocean as in summer (DJF) for their first EOF of daily SLP, as can be seen in our Figs. 3 and 5. The stronger annularity in summer is related to the
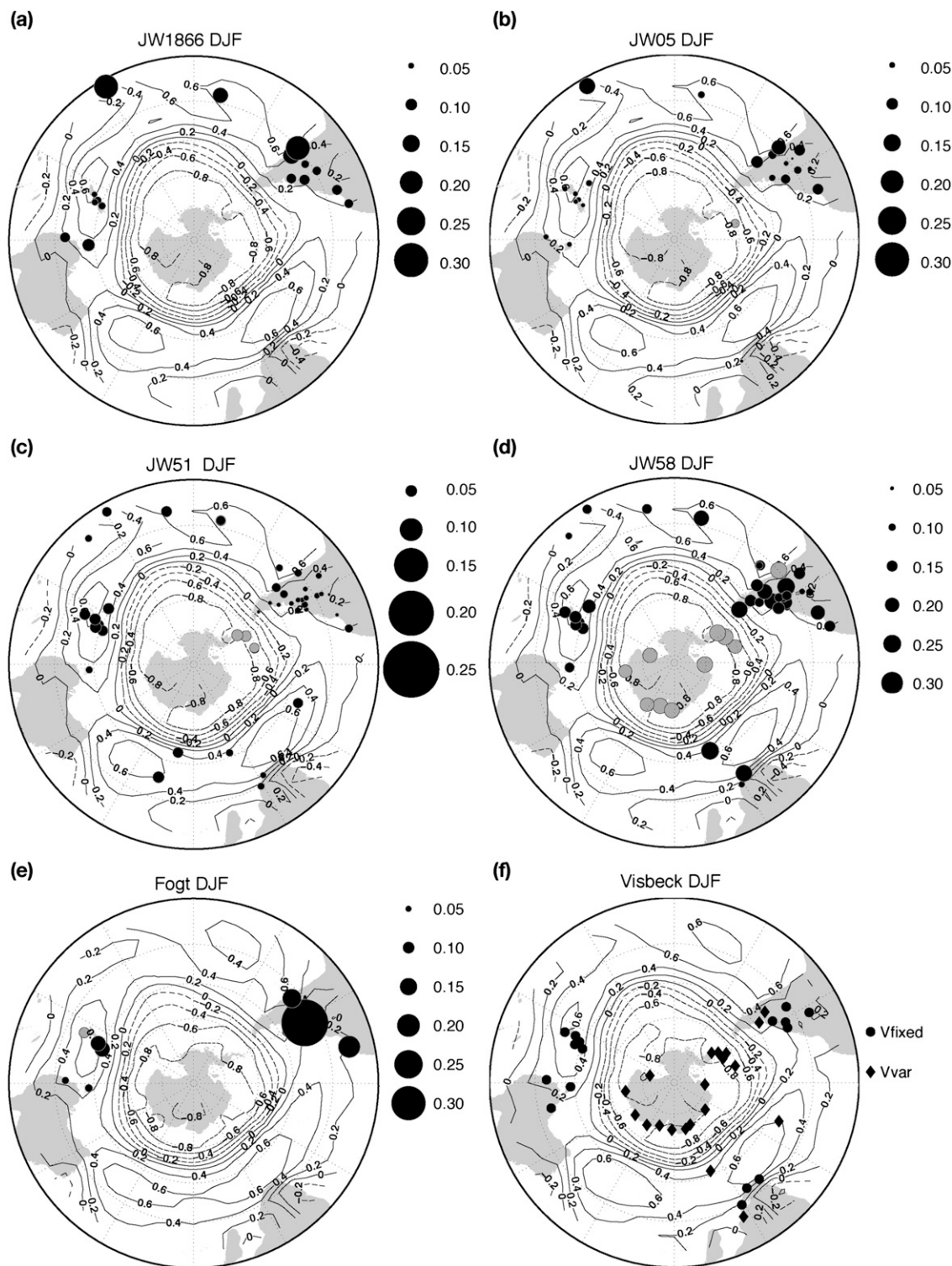
FIG. 3. The regression weights used in the DJF reconstructions for (a) JW1866, (b) JW05, (c) JW51, (d) JW58, and (e) Fogt. Contours are the correlation map between the respective calibration indexes detrended for 1979–2001 and detrended ERA-40 SLP. Black (gray) circles show positive (negative) weights, circle size proportional to regression weight. Note the different scales in (c) and (d). (f) The stations used for the Visbeck reconstructions, circles are stations used for Vfixed, all stations used for Vvar, and contours are the correlation map between detrended ERA40GW and detrended ERA-40 SLP 1979–2001.
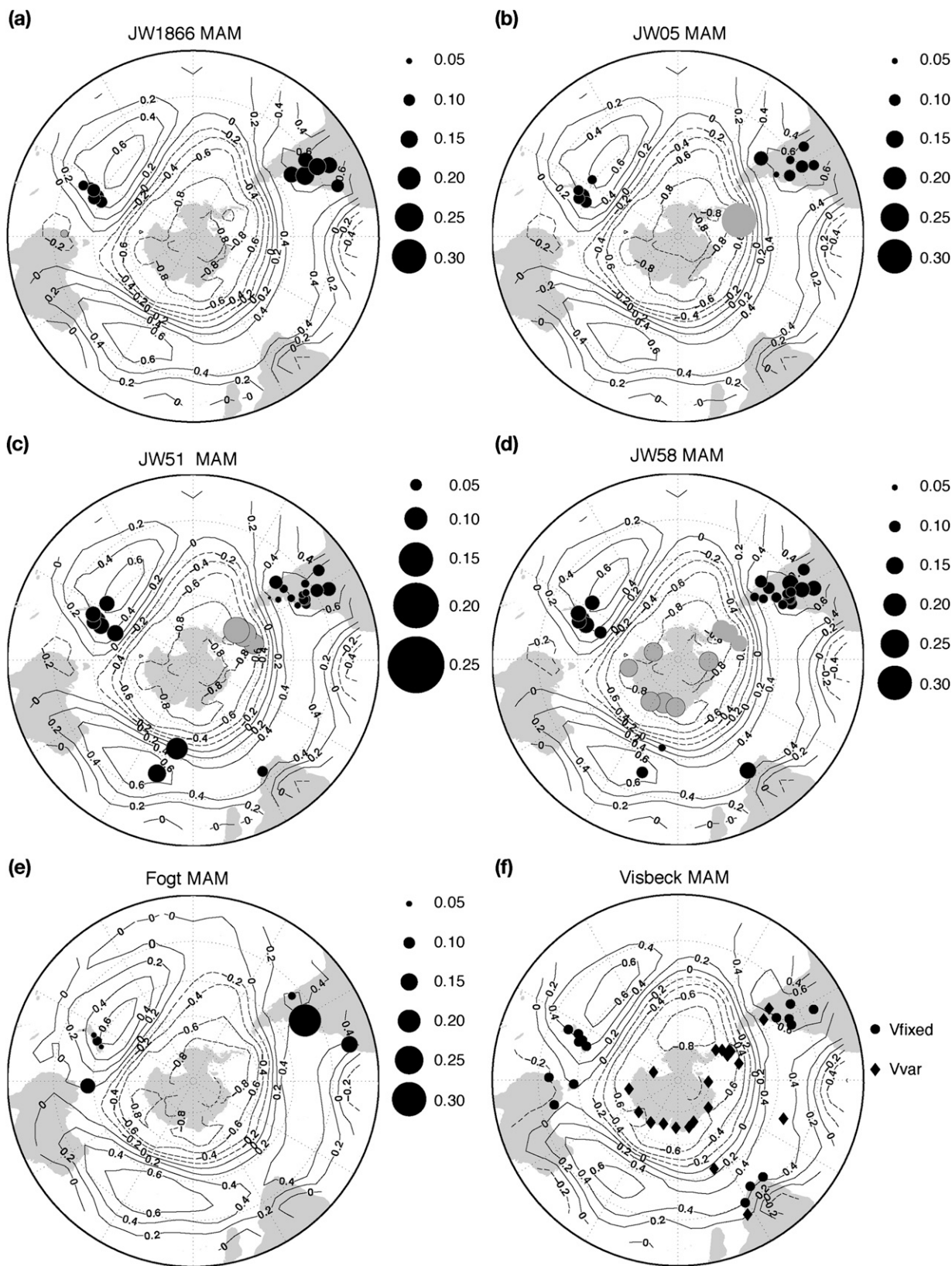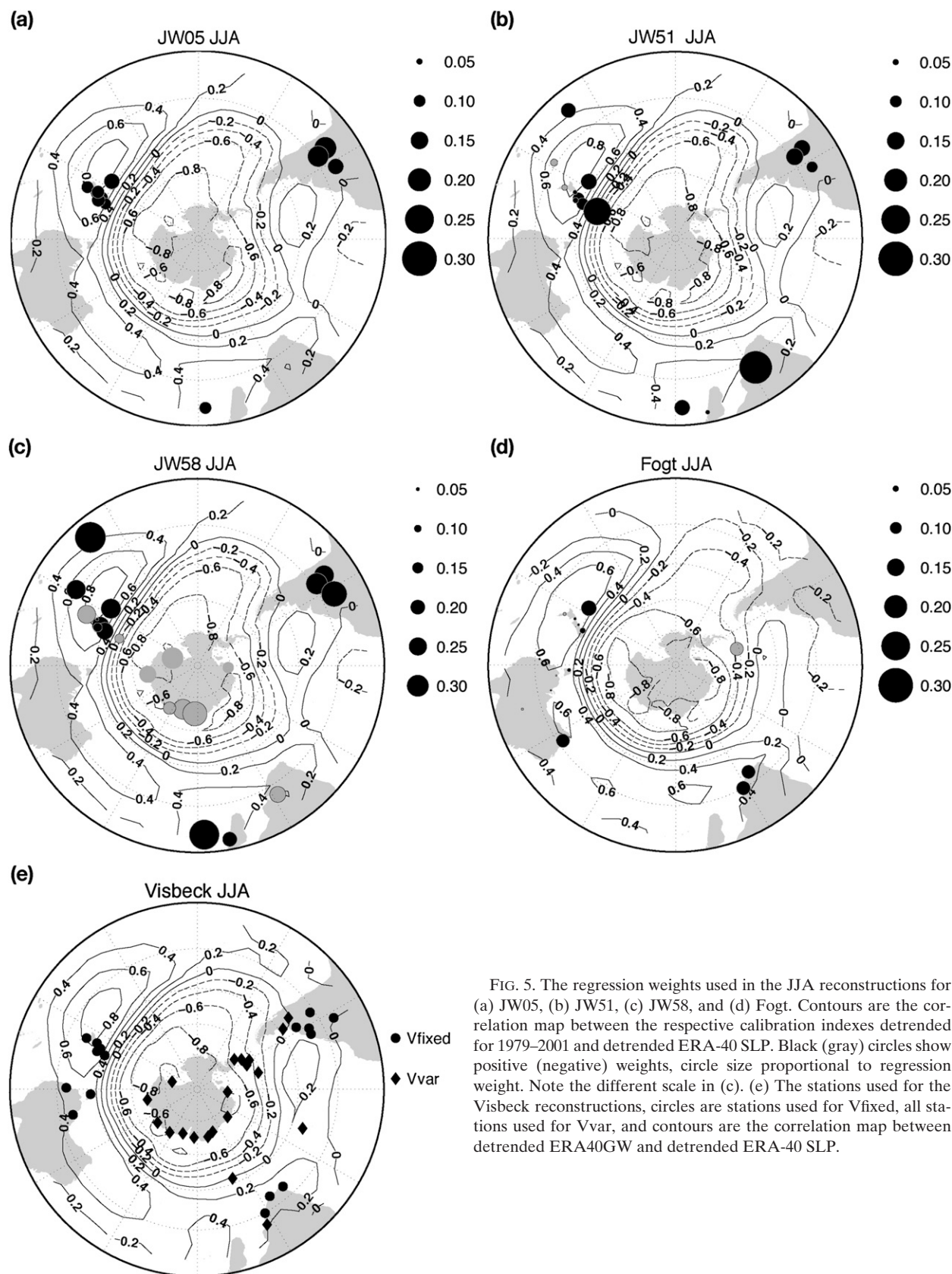
FIG. 4. As in Fig. 3, but for MAM. Note the different scale in (c).

(a) JW05 JJA

(b) JW51 JJA

(c) JW58 JJA

(d) Fogt JJA

(e) Visbeck JJA

FIG. 5. The regression weights used in the JJA reconstructions for (a) JW05, (b) JW51, (c) JW58, and (d) Fogt. Contours are the correlation map between the respective calibration indexes detrended for 1979–2001 and detrended ERA-40 SLP. Black (gray) circles show positive (negative) weights, circle size proportional to regression weight. Note the different scale in (c). (e) The stations used for the Visbeck reconstructions, circles are stations used for Vfixed, all stations used for Vvar, and contours are the correlation map between detrended ERA40GW and detrended ERA-40 SLP.

(a)

JW05 SON

(b)

JW51 SON

(c)

JW58 SON

(d)

Fogt SON

(e)
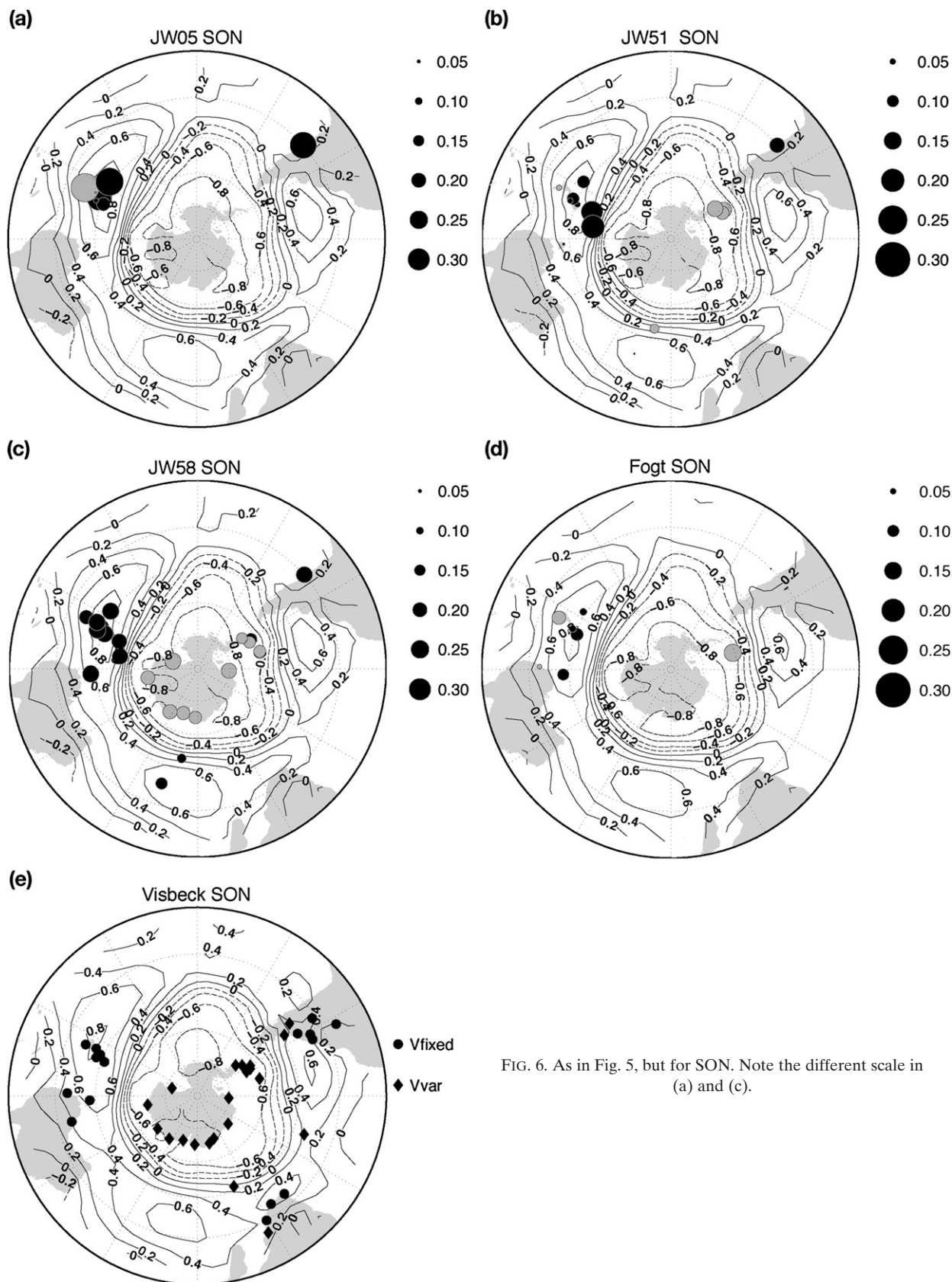
Visbeck SON

● Vfixed

◆ Vvar

FIG. 6. As in Fig. 5, but for SON. Note the different scale in (a) and (c).

greater zonal symmetry of the eddy-driven jet (Codron 2005; 2007) and hence storm tracks (e.g., Trenberth 1991) in this season.

Because of this strong zonal symmetry, in DJF areas of moderate to high correlations cover most midlatitude areas (with the exception of northern and central Australia; Fig. 3a). Stations from similar regions (South America and New Zealand) are predominantly used in the Fogt and JW05 and JW1866 reconstructions (Figs. 3a,b; the appendix), as the SLP pattern associated with the ERA40PC and that associated with the Marshall index are very similar. The addition of Antarctic and extra midlatitude stations in JW51 and JW58 (Figs. 3c,d; the appendix) brings some improvement in reconstruction quality; however, good validation statistics for JW1866 and JW05 indicate that the SAM is well captured in this season. The supplementary information in JW04 shows that the larger number of New Zealand and Australian stations compared to those in other centers of action does not greatly alter their DJ SAM reconstruction compared to one with fewer input stations in this region.

Areas with moderate correlations between the SAM index and local SLP cover South Africa and southern South America in MAM (Fig. 4). The surface pressure signal over South America in this season is stronger over many areas for ERA40PC (0.60) than the Marshall index (0.40); hence more stations are included from this area in the JW reconstructions than Fogt (Fig. 4; see appendix). The inclusion of more stations over the Antarctic and in the southern Indian Ocean leads to greater improvements in model quality than in DJF, the RE increasing from 0.63 (JW05) to 0.70 and 0.78 in JW51 and JW58, respectively. Orcadas, in the southern center of action, was very nearly significantly correlated in this season (-0.29, $p < 0.10$) with ERA40PC, thus was included to give information from this center in the JW reconstructions, and is indeed relatively strongly weighted.

The less zonally symmetric SAM in JJA (Fig. 5) results in lower signal strength over the continents and a lower correlation between the two predictands. As mentioned above, the JW reconstructions have their lowest quality in this season. Stations from South America are included in all JW reconstructions but not in Fogt (the appendix; Fig. 5): correlations with SLP are weakly positive with ERA40PC and negative with the Marshall index in this region. The SAM signal over New Zealand and Australia is stronger for the Marshall index (correlations of 0.4–0.6 over much of Australia) compared to for ERA40PC (0.0–0.2), perhaps because of the greater number of stations included from this region in the former. Thus the Fogt reconstruction in this season is

based primarily on Australian and New Zealand stations as well as Orcadas. Although reconstruction quality improves considerably for JW51 and JW58, the Fogt reconstruction quality is higher despite the fewer stations included (Table 2). Therefore it is easier to reconstruct the Marshall index rather than the PC index in this season.

The reconstruction quality for Fogt in SON is similar to JW58 (Table 2), despite being based on 10 stations compared to 23. All stations except one are located over New Zealand and Australia for both Fogt and JW05, as the spatial structure for both indices gives weak correlations between the SAM index and local SLP over large areas of the other midlatitude landmasses (Fig. 6). Additionally, the correlations between the detrended predictor stations and the detrended SAM index in this season are lower than in DJF and MAM (the appendix). There is little improvement in the JW51 reconstruction despite additional stations over the Antarctic Peninsula, but it is greater with the addition of stations around the Antarctic and in the Indian Ocean in JW58. The aforementioned greater uncertainty in the pre-1979 ERA-40 data at high latitudes in JJA and SON may contribute toward the poorer reconstruction quality in JJA and SON for the JW reconstructions.

In summary:

1) The quality of the JW reconstructions is best in DJF, followed by MAM, and is poorer in SON and particularly in JJA. This results from the smaller number of predictor stations in SON and JJA (because of a less annular structure, giving a weaker SAM signal over the midlatitude continents), from the lower average correlation of these selected stations, and from the greater uncertainty in the pre-1979 ERA-40 data.

2) The Fogt reconstructions perform equally well in all seasons except in MAM, where the lower number of stations captures less well the correlation pattern (especially over New Zealand). The Fogt reconstructions include more stations in JJA than JW05 because of the higher correlations of the Marshall index over Australia compared to ERA40PC.

c. Reconstruction comparison and characteristics 1958–2005

As SAM behavior in past decades has been of strong interest (e.g., Thompson and Solomon 2002; Gillett and Thompson 2003; Marshall 2007) we first compare the reconstructions in the recent period 1958–2001 (Fig. 7). The trends during this period are analyzed in more detail in Part II. The JW and Fogt reconstructions
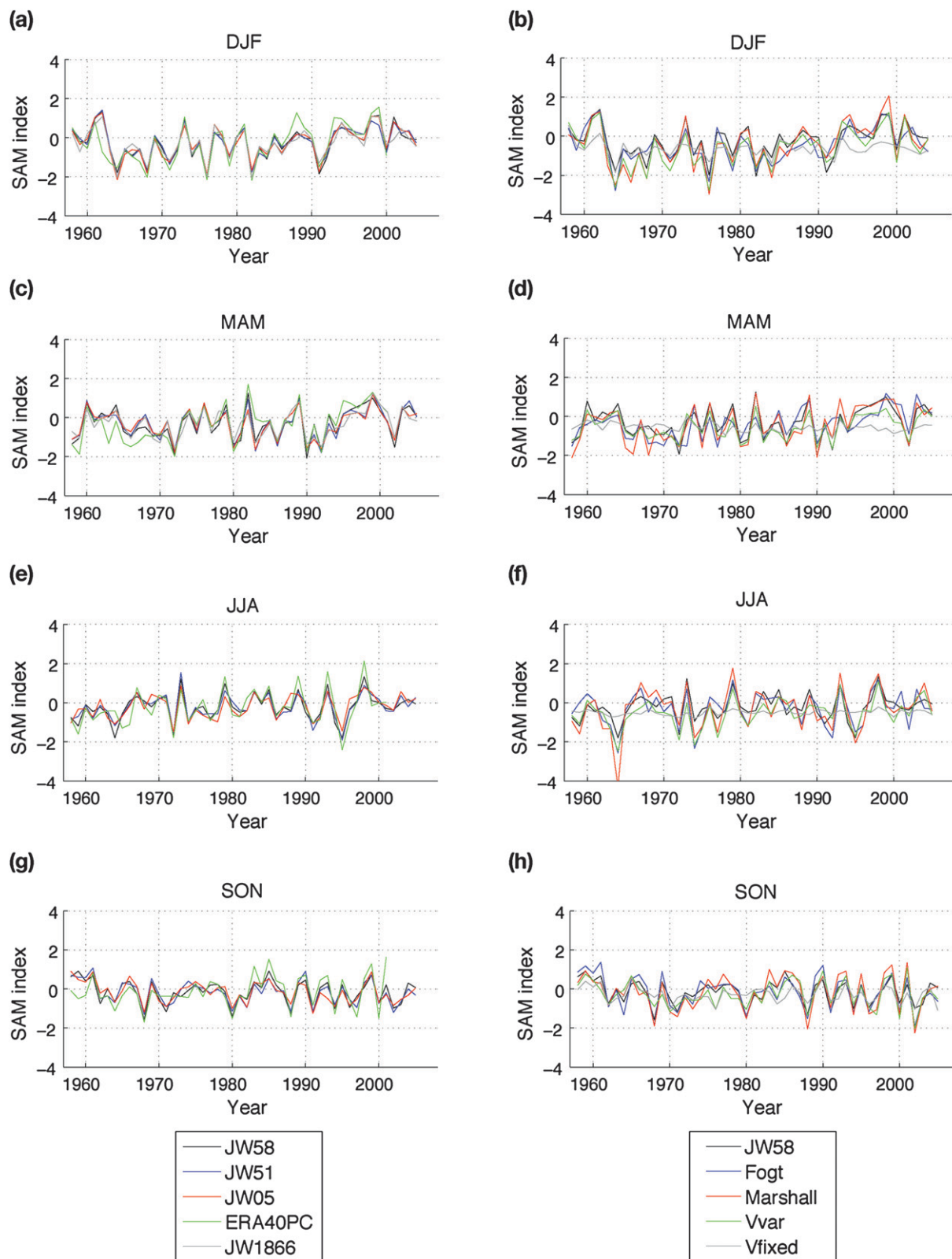
FIG. 7. (left) The JW reconstructions and ERA40PC for the period 1958–2005. (right) The JW58, Fogt, and Visbeck reconstructions and the Marshall index for the period 1958–2005.
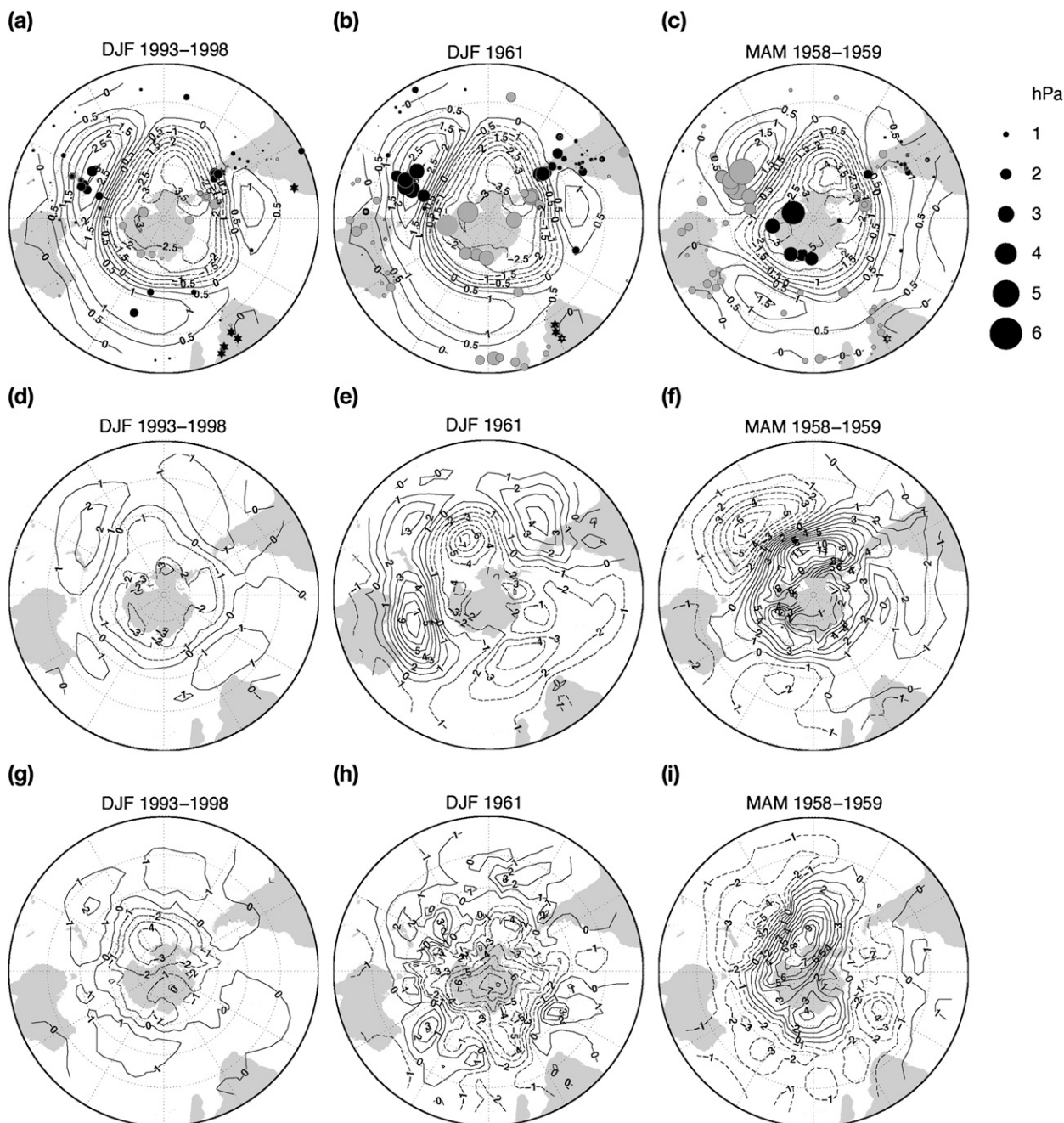
FIG. 8. (top) Station SLP anomalies for (a) DJF 1993–98, (b) DJF 1961, and (c) MAM 1958–59. Anomaly magnitude is proportional to circle size, positive anomalies are black, and negative are gray. Stars signify stations with missing data over the composite period. Contour lines show regression map between detrended ERA40PC DJF SAM index and ERA-40 mean SLP (MSLP), 1979–2001. (middle) Mean ERA-40 SLP anomalies for (d) DJF 1993–98, (e) DJF 1961, and (f) MAM 1958–59. (bottom) Mean HadSLP2 anomalies for (g) DJF 1993–98, (h) DJF 1961, and (i) MAM 1958–59. All anomalies are relative to the 1979–2001 mean.

inherently have lower variability than the Marshall index, as the regression models explain less than 100% of the variance. Vfixed has very low variability because of the low correlations with ERA40GW, hence only little variance of the SAM index is explained, thus discussion

will focus on Vvar. To investigate the reasons for the differences in the reconstructions, SLP anomalies from the 1979–2001 mean from all available stations for selected periods of interest were calculated. To compare these to the SAM centers of action, they are plotted over

TABLE 3. Correlations between the detrended reconstructions for the period 1958–2001. Values in parentheses are for correlations between series filtered with a 9-yr Hamming window; values not in parentheses are the interannual correlations.

	JW58	JW51	JW05	JW1866	Marshall	Fogt
DJF						
Marshall	0.90 (0.90)	0.88 (0.88)	0.87 (0.84)	0.84 (0.89)		
Fogt	0.82 (0.77)	0.84 (0.84)	0.86 (0.75)	0.87 (0.83)	0.84 (0.91)	
Vvar	0.92 (0.87)	0.92 (0.90)	0.88 (0.81)	0.80 (0.84)	0.94 (0.90)	0.81 (0.80)
Vfixed	0.69 (0.76)	0.71 (0.73)	0.65 (0.66)	0.61 (0.67)	0.67 (0.72)	0.63 (0.62)
MAM						
Marshall	0.89 (0.91)	0.79 (0.80)	0.77 (0.75)	0.67 (0.69)		
Fogt	0.70 (0.72)	0.70 (0.74)	0.74 (0.71)	0.78 (0.76)	0.73 (0.67)	
Vvar	0.95 (0.93)	0.87 (0.86)	0.84 (0.85)	0.71 (0.76)	0.87 (0.94)	0.58 (0.68)
Vfixed	0.59 (0.45)	0.56 (0.48)	0.54 (0.53)	0.28 (0.32)	0.56 (0.55)	0.19 (0.15)
JJA						
Marshall	0.80 (0.89)	0.65 (0.82)	0.59 (0.65)			
Fogt	0.72 (0.86)	0.67 (0.66)	0.65 (0.60)		0.83 (0.83)	
Vvar	0.78 (0.55)	0.62 (0.47)	0.62 (0.58)		0.82 (0.63)	0.75 (0.73)
Vfixed	0.33 (0.11)	0.19 (−0.03)	0.21 (−0.14)		0.44 (0.21)	0.50 (0.49)
SON						
Marshall	0.89 (0.95)	0.77 (0.67)	0.73 (0.79)			
Fogt	0.92 (0.81)	0.86 (0.73)	0.81 (0.74)		0.85 (0.82)	
Vvar	0.75 (0.66)	0.72 (0.59)	0.58 (0.47)		0.89 (0.74)	0.75 (0.66)
Vfixed	0.54 (0.13)	0.54 (0.07)	0.35 (−0.01)		0.59 (0.34)	0.66 (0.37)

the regression map between detrended ERA40PC and ERA-40 SLP for 1979–2001 (Fig. 8). ERA-40 and HadSLP2 anomalies for the same years relative to 1979–2001 were also plotted to provide another estimate of the spatial structure (Fig. 8). Where ERA-40 and HadSLP2 agree, greater confidence exists in the full spatial SLP anomaly pattern.

1) DJF 1958–2005

Agreement between reconstructions is strongest in this season. Interannual correlations between JW58 and the Marshall index are 0.90; the lowest correlation of 0.84 is with JW1866 (Table 3). Correlations between JW05, JW1866, and Fogt are higher than for the shorter JW reconstructions. This may be because the long reconstructions are based on very similar predictor networks, with stations predominantly in Australasia and South America, whereas JW51 and JW58 include stations in the Antarctic and the southern Indian and Atlantic Oceans (Fig. 3; the appendix). Correlations with all reconstructions and Vvar are highest in this season, with interannual correlations of over 0.90 with JW58, JW51, and Marshall. This may be because the Visbeck (2009) method (i.e., the assumption of an equal contribution of each ocean basin) is most valid in this season because of the aforementioned higher zonal symmetry of the SAM. Vfixed correlations are also highest in this season (above 0.60 with all reconstructions).

The well-documented positive trend in austral summer (DJF) from the mid-1960s to present (Thompson and Solomon 2002; Gillett and Thompson 2003; Marshall 2007) is evident in all indices (Figs. 7a,b). It is strongest in the Marshall index because of stronger negative values in the mid-1970s and to more positive recent values. The station anomalies during this recent peak (1993–98; Fig. 8a) show a very SAM-like pattern, with negative anomalies over the Antarctic and positive anomalies in all midlatitude centers of action, all of which are approximately proportional to the regression coefficients. The ERA-40 and HadSLP2 anomalies corroborate the positive SAM structure (Figs. 8d,g). The more positive Marshall index results therefore from the inclusion of stations in all centers of action, whereas the greater weighting toward South American stations in the Fogt and JW reconstructions (Fig. 3) results in lower values.

The 1960s positive peak noted by JW04 is present and of similar magnitude in all series except Vfixed. Anomalies for 1961 (Fig. 8) show a less zonally symmetric structure, although the strongest anomalies do occur in areas of strongest SAM signal. The station anomalies (Fig. 8b) indicate that the midlatitude zonal symmetry breaks down east of Africa, with negative anomalies at all latitudes. The ERA-40 and HadSLP2 spatial plots (Figs. 8e,h) generally agree with the observations, although there are subtle differences (e.g., negative ERA-40 anomalies in the Atlantic sector and the HadSLP2 plot is much noisier). The lower zonal symmetry in ERA-40 compared to indications from the observations helps to explain why the 1960s peak is stronger and lasts longer in the reconstructions than in ERA-40; however, the fact that all reconstructions have similar values suggests this peak is a prominent SAM episode.

2) MAM 1958–2005

Agreement in this season is generally weaker than in DJF between the JW reconstructions and both the Marshall index and Fogt (Table 3), except for Marshall/ JW58. This is due to the aforementioned slightly less good fitting and validation statistics in this season for all reconstructions, especially Fogt (section 3b), and the lower agreement between predictands for 1958–2001 (section 3a). As in DJF the strongest correlations between the JW reconstructions and the Fogt reconstruction are with the JW05 and JW1866 reconstructions, reflecting the greater similarity of their predictor networks (Fig. 4).

There is a positive trend in the MAM reconstructions both from 1958 and the early 1970s to present (Figs. 7c,d), although the JW reconstructions and the Marshall index have less negative values in the early 1960s than Fogt, ERA40PC, and Vvar. The period of positive SAM index in the 1990s is higher in the Marshall index and JW58 than in Fogt and Vvar. The MAM station SLP anomalies for 1996–99 (not shown) show a pronounced positive hemispheric SAM anomaly. Hence, the more positive Marshall index and JW58 reconstruction during this period may result from the inclusion of Antarctic stations with negative anomalies.

Station anomalies during the period of negative reconstructed and observed SAM index values during 1958–59 (Fig. 8c) are consistent with the negative SAM pattern in all regions except South America where anomalies are predominantly weakly negative. Both ERA-40 (Fig. 8f) and HadSLP2 (Fig. 8i) indicate a negative SAM structure. Both spatial plots also have the strong blocking in the southeast Pacific, a region of preferential blocking in the SH (Renwick and Revell 1999; Renwick 2005), which leads to a greater frequency of cyclones deflected farther north over New Zealand, creating the low pressure anomalies there. The most negative SAM index is in Marshall, JW51, and JW58 (Figs. 7c,d), reflecting the fact that these reconstructions are based on stations in more centers of strong anomalies in Figs. 8f,i (the Antarctic and New Zealand), whereas the longer reconstructions have a larger proportion of South American stations as predictors, and Vvar equally weights all areas.

3) JJA 1958–2005

Agreement between the JW reconstructions with both the Fogt and Marshall indices is the lowest of all seasons (Table 3); this is likely related to the lower correlation between the two predictands in this season (section 3a) and the differing spatial structure of the SLP anomalies

related to these indices (cf. contours in Figs. 5a,d). Correlations for the JW58 reconstruction are higher than for the longer JW reconstructions, with correlations of 0.80 and 0.72 with the Marshall index and Fogt reconstructions, respectively, reflecting that added value is gained by inclusion of Antarctic stations. The correlation between Fogt and Vvar of 0.75 is stronger than that between Fogt and all JW reconstructions; however, correlations with Vfixed drop to 0.50, reflecting the lower zonal symmetry in this season, which is assumed in construction of the Visbeck reconstructions. The very weak agreement between the JW reconstructions and Vfixed reflect the uncertainty in both sets of reconstructions.

All reconstructions show weak positive trends from the mid-1960s to the present, although, as discussed further in Part II, most of this trend results from a strong negative index in 1964, which is most pronounced in the Marshall index. This year is characterized by a full hemispheric SAM pattern, which is evident in observations, ERA-40 and HadSLP (not shown). M03 identified that, in this year, all Antarctic stations except Novolazarevskaya have a stronger positive anomaly than in ERA-40. It is thus not surprising that the anomaly is most marked in the Marshall index (Fig. 7f). As this may be related to the Agung eruption of 1963, ERA-40 potentially does not capture this so strongly because of a lack of radiosonde data in this region to constrain the reanalysis.

4) SON 1958–2005

The JW reconstructions (except for JW05) experience their strongest interannual correlations with both Fogt and the Marshall index in this season (Table 3), despite the lower fitting and validation statistics than in DJF and MAM. This is probably because of the strong correlation between predictands (section 3a; Table 1). Agreement between Fogt and Marshall is also strongest in this season (0.85). Interannual agreement between Vvar and both Marshall and Fogt is good in this season, with interannual correlations of 0.89 and 0.75, respectively, and it is reasonable with Vfixed. However, agreement between the JW reconstructions and Vfixed is poor, with low-frequency correlations as low as −0.01. All reconstructions agree that there is no trend throughout this period (Figs. 7g,h; see also Part II).

In summary:

1) There is a positive trend in the SAM index in recent decades in all seasons except SON, although the JJA trend is strongly influenced by a pronounced negative index in 1964.

2) The Vfixed reconstructions, based on 14–17 stations, agree poorly with all other reconstructions (with

strongest agreement in DJF when the assumption of zonality is better met), indicating potential uncertainties early in Vvar when it is also based on a similar number of stations.

3) Agreement between all reconstructions is strongest in DJF; the JW and Fogt reconstructions also agree strongly in SON.

4) Analysis of station and ERA-40 SLP anomalies during positive and negative reconstructed SAM phases shows that some periods, such as the recent positive DJF SAM index, are very SAM-like. During others, however, such as the 1960s peak in DJF, anomalies do not occur in all SAM centers of action.

d. Comparison and characteristics of full reconstructions

1) CALCULATION AND COMPARISON OF ERROR BARS

The JW and Fogt reconstructions are fitted over different periods (1958–2001 and 1957–2005, respectively). For a consistent skill intercomparison the period 1958–2001 was used to calculate the error bars. These are defined as ±1.96 standard deviations of the residuals of the reconstructed seasonal index from the predictand. The JW error bars were calculated relative to ERA40PC, Fogt relative to Marshall, and Marshall and Vvar relative to ERA40GW.

The magnitudes of these error bars are in accordance with the fitting and validation statistics (Table 2). The interannual error bars are shown (Figs. 9b,e,h,k) added to the reconstructions (both smoothed with a 9-yr Hamming filter). Low-frequency errors, which can be expected to be smaller, were not calculated because of the short period of data available for calculation (44 yr). The errors for JW are smallest in DJF and MAM and reduce in size with the shorter reconstructions. The Fogt error bars are bigger in DJF than JW1866 (the closest equivalent network) because of the low Marshall index values in the mid–late 1960s that are not fully captured in the reconstruction. In JJA and SON, the error bars for the Fogt reconstructions are smaller or equivalent to those for JW1905, and even are of similar magnitude to those of JW1958, as the Fogt reconstruction has better fitting and validation statistics. The Marshall error bars reflect how well this index captures the true zonal mean pressure differences, being largest in JJA (although this also is influenced by the aforementioned strong negative index in 1964, which is not captured by the reanalysis data).

The full period reconstructions for all seasons are shown in Fig. 9, together with the Marshall index. The JW reconstructions have been concatenated to produce JWconcat, to provide a best estimate of the SAM index from 1865. The Vfixed reconstructions are not shown because of their uncertainty; however, their running correlations with JWconcat are shown. The Vfixed error bars are considerably larger than for Vvar because of the low variability and low correlation with ERA40GW. Hence, prior to about 1960 (when then number of stations included reduces), these Vvar error bars in Fig. 9 can be thought of as lower estimates.

2) RECONSTRUCTION COMPARISON AND CHARACTERISTICS DJF 1865–2005

The positive trend from the mid-1960s to 2000 is the clearest feature of the reconstructions (Figs. 9a,b) and is strongest in the Marshall index (section 3c; see also Part II). As first noted in JW04 in their DJ reconstruction, there is also a positive trend from negative values in 1939 to a peak of higher values in 1962/63, but this is weaker than the recent trend. Station anomalies are more similar to a SAM structure during the late 1990s than during this peak [section 3c(1)]. Note that the reconstructions are standardized to have zero mean for 1979–2001 (a period of positive SAM index), whereas those in JW04 have zero mean for 1958–2001, hence the values in Fig. 9 are more negative than in JW03 and JW04. The reconstructions trace each other well throughout the full period, although it is evident that the variability is higher in Fogt, which also has a lower mean prior to 1960. There is a single positive value in 1927/28 in all reconstructions (Fig. 9a) that is of similar value to the 1960s and recent peaks, but this is not visible in the low-frequency values as it is in a period of predominantly lower values. All reconstructions also have a negative peak in 1911/12. Negative SAM anomalies are associated with warmer temperatures over the Antarctic (Marshall 2007; Kwok and Comiso 2002). This summer indeed experienced high temperatures, as recorded on the Scott expedition to the South Pole (Schwerdtfeger 1984; Villalba et al. 1997).

The degree of agreement between the reconstructions is shown in Fig. 9c, through correlations over running 20-yr windows, and in Table 4 (full period correlations). Correspondence between JW1866 and Fogt is good throughout the reconstruction, with correlations generally above 0.80, and correlations over the full period of 0.88 (Table 4). Although still high (20 yr > 0.75, 1866–2005 = 0.81), correlation between JWconcat and Fogt is lower because of the less similar station networks [section 3c(1)]. Correlations between JWconcat and Vvar are also above 0.80 for windows after 1905, reflecting the good agreement, also evident with Fogt (Table 4).
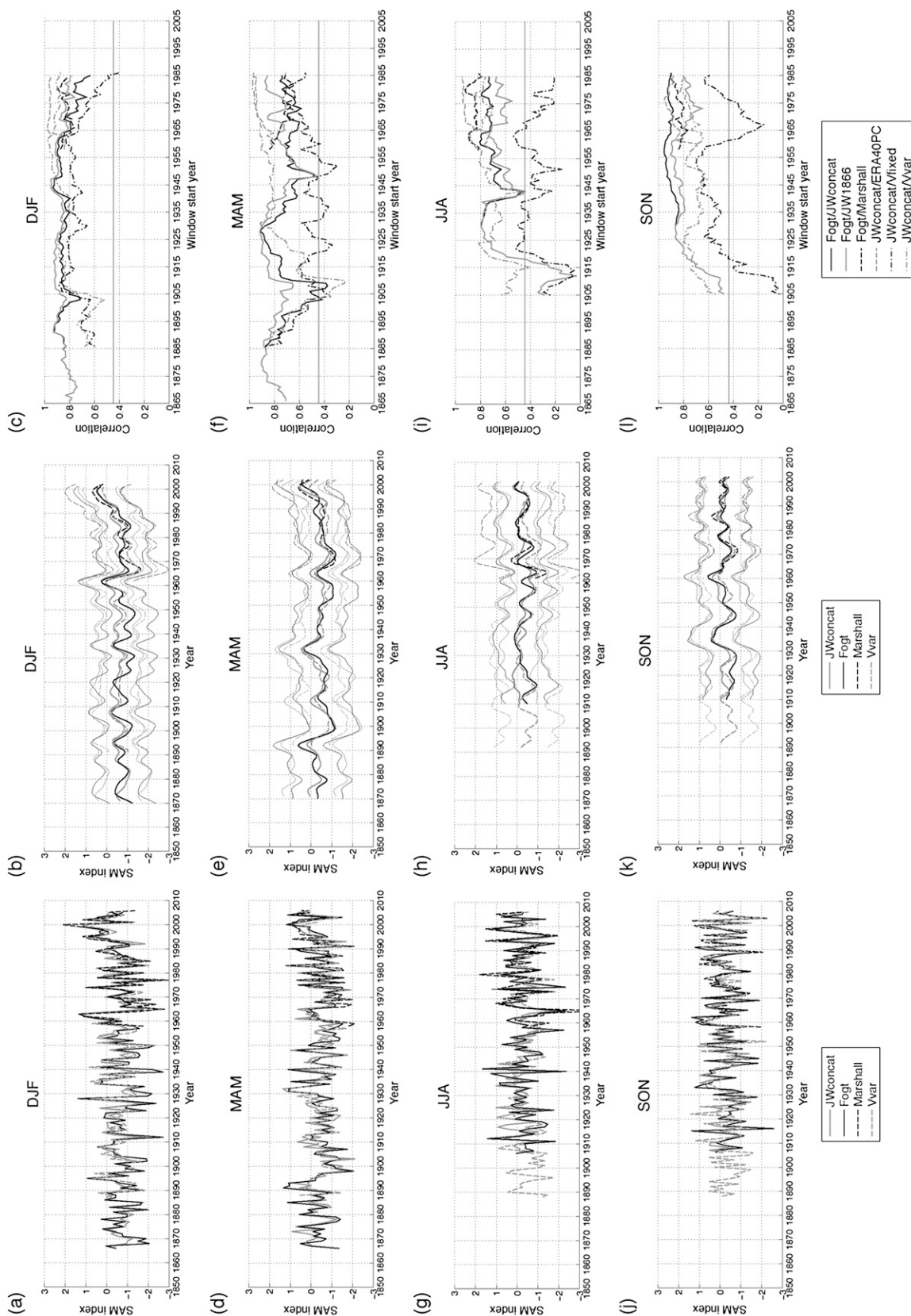
FIG. 9. The Fogt, JWconcat, and Vvar reconstructions and the Marshall index: (top to bottom) DJF to SON. (left) Interannual values and (center) interannual values filtered with a 9-yr Hamming window. Thin lines in (center) are the interannual 95% confidence intervals (with 9-yr Hamming filter applied). (right) 20-yr moving window correlations of annual values of the JWconcat and JW1866 reconstructions with the Fogt reconstructions, of the Vfixed and Vvar reconstructions with JWconcat, and of the Fogt and Marshall reconstructions and JWconcat and ERA-40. Horizontal black line is the 5% significance level for 18 degrees of freedom.

TABLE 4. Correlations between the detrended JW, Fogt, and Visbeck reconstructions over their full period. Values in parentheses are for correlations between series filtered with a 9-yr Hamming window; values not in parentheses are the interannual correlations.

	JW05 1905–2005	JWconcat 1905–2005	JW1866 1905–2005	JWconcat 1866–2005	JW1866 1866–2005	Vvar 1905–2005	Vfixed 1905–2005
DJF							
Fogt	0.84 (0.80)	0.82 (0.79)	0.88 (0.85)	0.81 (0.79)	0.86 (0.84)	0.78 (0.71)	0.65 (0.48)
JWconcat	0.95 (0.96)						
Vvar	0.84 (0.83)	0.87 (0.83)	0.78 (0.80)	0.84 (0.83)	0.77 (0.79)		
Vfixed	0.68 (0.68)	0.70 (0.68)	0.64 (0.62)	0.69 (0.65)	0.64 (0.59)		
MAM							
Fogt	0.73 (0.72)	0.70 (0.78)	0.79 (0.82)	0.72 (0.78)	0.80 (0.82)	0.66 (0.77)	0.13 (0.13)
JWconcat	0.89 (0.94)						
Vvar	0.76 (0.83)	0.81 (0.89)	0.75 (0.84)	0.81 (0.89)	0.74 (0.84)		
Vfixed	0.63 (0.51)	0.49 (0.47)	0.19 (0.52)	0.54 (0.61)	0.29 (0.52)		
JJA							
Fogt	0.58 (0.32)	0.62 (0.40)				0.67 (0.75)	0.61 (0.63)
JWconcat	0.91 (0.88)						
Vvar	0.62 (0.45)	0.70 (0.38)					
Vfixed	0.27 (0.02)	0.32 (0.01)					
SON							
Fogt	0.76 (0.80)	0.82 (0.83)				0.72 (0.79)	0.75 (0.63)
JWconcat	0.93 (0.96)						
Vvar	0.60 (0.57)	0.68 (0.71)					
Vfixed	0.38 (0.44)	0.47 (0.50)					

Running window correlations between JW58 and ERA40PC also shown in Fig. 9c are higher than between Fogt and Marshall, reflecting the slightly better validation statistics of the former. The Fogt means are, however, generally lower than JW, a feature not represented by the correlations.

3) RECONSTRUCTION COMPARISON AND CHARACTERISTICS MAM 1866–2005

As in DJF, the reconstructions have considerable decadal-scale variability, both the interannual (Fig. 9d) and 9-yr filtered data (Fig. 9e) show that the peaks in recent decades are not unprecedented. A strong positive SAM in the early 1890s is followed by a sharp decline, which is followed by a positive trend leading to the early 1930s. An early 1960s peak is also evident: this is higher in JWconcat than in Fogt. All reconstructions agree on the main characteristics of the MAM SAM, as the reconstruction means are similar throughout. There are periods of lower correlation between the reconstructions (Fig. 9f), namely 1905–30 and 1950–60. Agreement is stronger between Vvar and the JW reconstructions than with Fogt, with correlation between Vfixed and the latter very low (0.13).

To investigate whether the peak around 1930 is associated with a regional SLP anomaly captured by the predictors or indeed reflects a hemispheric SAM pattern, station pressure anomalies for 1930–32 were plotted (Fig. 10a). Anomalies are positive at most available midlatitude stations (with strongest anomalies over New Zealand), and there is a negative anomaly at Orcadas. The positive anomalies in southern Africa and South America indicate positive anomalies in this center of action, however, and together with the negative Orcadas anomaly these suggest a hemispheric SAM-like pattern. However, the HadSLP2 anomalies (Fig. 10c) suggest a negative SAM index, with anomalies being positive over the Antarctic and negative at midlatitudes. This tendency for a more negative SAM index in HadSLP2 than in the JW03 reconstructions during the early twentieth century was noted by Allan and Ansell (2006), and it is also more negative than the Fogt reconstructions during this period (Part II). This indicates potential uncertainties in the HadSLP2 SAM estimates during this period because of the aforementioned sparse data network (although uncertainties in the JW and Fogt reconstruction methods cannot be ruled out).

4) RECONSTRUCTION COMPARISON AND CHARACTERISTICS JJA 1905–2005

Before the positive trend from the mid-1960s to present, due predominantly to a very strong negative SAM index in 1964 [section 3c(3)], there is little decadal-scale variability in JJA (Fig. 9h). Agreement between the reconstructions is low in the early parts with correlations between JWconcat and Fogt of less than 0.10 for the window beginning in 1912, becoming significant for 20-yr windows with start dates after 1917. A number of
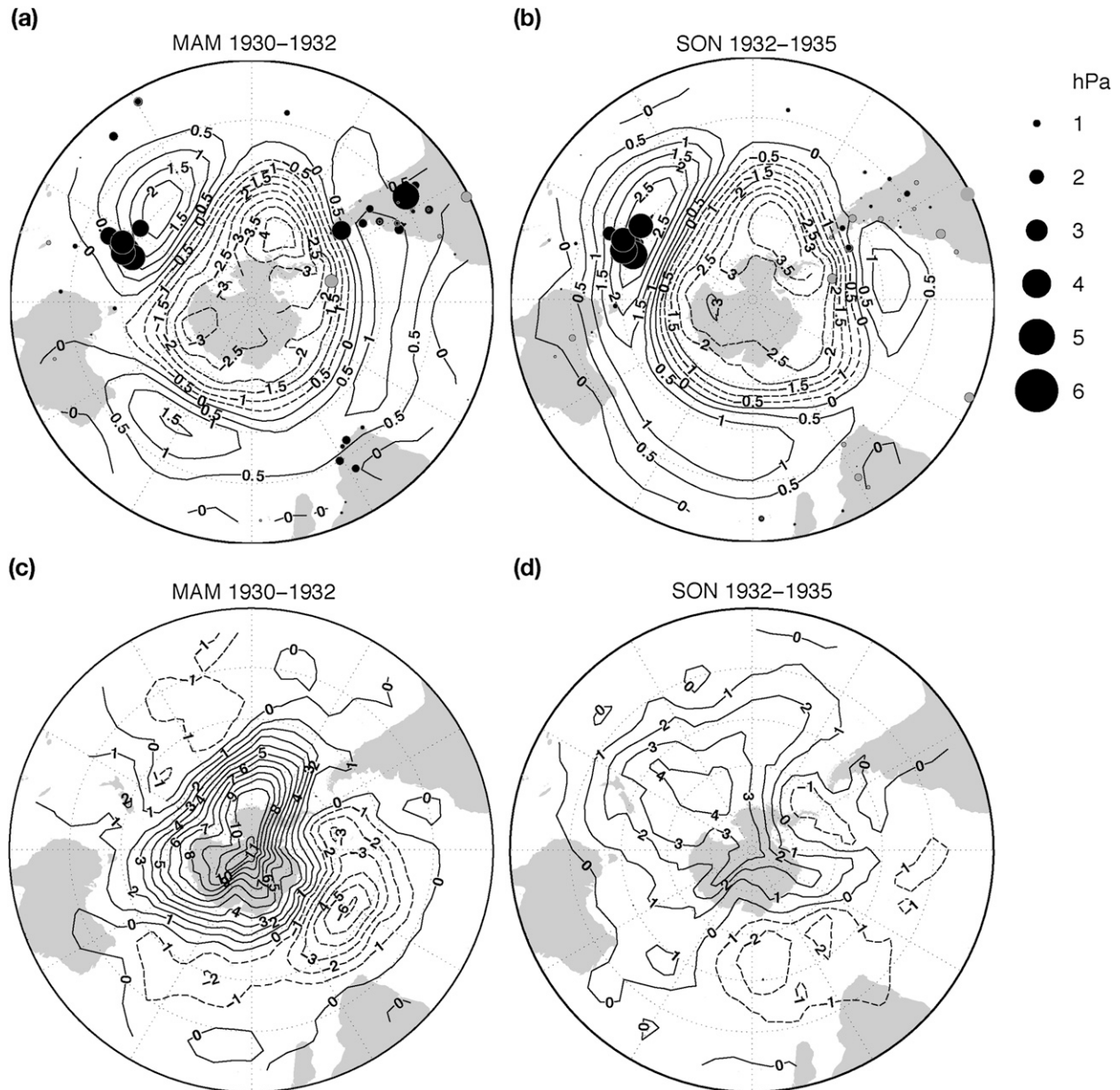
FIG. 10. (top) Station SLP anomalies for (a) MAM 1930–32 and (b) SON 1932–35 relative to the 1979–2001 mean. The anomaly magnitude is proportional to circle size, positive anomalies are black, and negative are gray. Contour lines show the regression map between detrended ERA40PC DJF SAM index and ERA-40 MSLP, 1979–2001. Stars signify stations with missing data over the composite period. (bottom) HadSLP2 anomalies relative to the 1979–2001 mean for (c) MAM 1930–32 and (d) SON 1932–35.

years with SAM values of opposite sign are evident in the 1910s and 1920s (Fig. 9g); Vvar agrees with Fogt on more negative values in this period. Full period inter-annual and low-frequency correlations between Fogt and both JWconcat and JW05 are the lowest of any season, with low-frequency correlations of 0.40 and 0.32, respectively (Table 4). Low-frequency agreement appears poor until around 1975 (Fig. 9h). As in SON, and in contrast to DJF and MAM, correlations for windows

past 1965 are higher between Fogt and JWconcat than with JW1905 (around 0.75 compared to 0.60). This suggests that the JJA JW05 reconstruction, as indicated by the poorest fitting and validation statistics of all reconstructions in all seasons (section 3b), is less reliable than Fogt, but confirms that the inclusion of additional stations for the JW58 reconstruction improves its quality [section 3c(3)]. Running window correlations between JWconcat and Vvar drop as the number of

stations included decreases. However, agreement is stronger between Fogt for both Vvar and Vfixed than with JW.

5) RECONSTRUCTION COMPARISON AND CHARACTERISTICS SON 1905–2005

Low-frequency variability in all reconstructions is greater during the first part of the reconstruction than in the latter (Figs. 9j,k); the standard deviation of the low-frequency JWconcat (Fogt) reconstruction is 0.28 (0.38) for 1905–65 compared to 0.17 (0.17) for 1965–2005 ($p < 0.05$). This variability is due to a period of positive values between 1930 and 1940, with negative values before and after (Fig. 9j). Although significant ($p < 0.05$) throughout, correlations between JW and Fogt are weaker (0.50) at the beginning of the reconstructions. This can be seen in both the interannual and low-frequency values (Figs. 9j,k), where JWconcat is more positive than Fogt. After the late 1920s agreement improves, with correlations above 0.80 (Fig. 9l). The additional stations in the JW58 reconstruction improve its quality, as agreement with Fogt is higher for JWconcat than JW05; indeed full period correlations of 0.82 are as high as in DJF. The 1960s peak is not present in Vvar; the 1930s peak is present although lower than in both JW and Fogt.

During 1930–40, the reconstructed SAM index is positive in all years except 1936. The SON anomalies for 1932–35 (Fig. 10b) show that, in contrast to the 1930–32 peak in MAM, station anomalies are only strong in New Zealand but weak elsewhere. Orcadas shows a negative anomaly, but the anomalies north of this in South America do not show a consistent sign and likely reflect regional SLP anomalies rather than a hemispheric SAM structure. The HadSLP2 anomaly (Fig. 10d) also does not resemble a SAM structure. Rather, the station and HadSLP2 anomalies show some similarity with those associated with the 1960s peak (not shown), where anomalies are also strongest over New Zealand, with weaker negative anomalies cantered over the Antarctic Peninsula and South America, that is, more of a zonal wavenumber-1 pattern than a SAM pattern.

In summary:

1) The DJF and MAM SAM reconstructions show considerable decadal-scale variability: in DJF this is greatest for the recent trends, whereas in MAM it is strong throughout. The recent JJA positive trend is the main low-frequency feature, and in SON reconstructed SAM variability is greatest in the first half of the twentieth century.
2) Agreement on these main features is strongest between all reconstructions in DJF and lowest in JJA, when Vvar agrees more strongly with Fogt than does JW.

4. Discussion and conclusions

In the introduction, four aims to be addressed in the paper were outlined. We organize our discussion around these four points.

a. To analyze the temporal variability of the PC-based, Gong and Wang, and Marshall indices separately for each season, and to analyze the link between differences in the variability and the defining weight patterns

For the post-1979 period, the two SAM indices derived from the two definitions, zonal pressure differences at mid- and high latitudes from M03 and PC-based (Thompson and Wallace 2000), are highly similar, with correlations >0.85 in all seasons. When data prior to 1979 are included, agreement drops in all seasons. The smallest drop is for DJF. The differences in JJA are particularly large: the correlations between the two indices are only 0.65. It thus seems that the difference in definition, and thus in defining weight patterns, is small enough to not result in large differences in indices when one is not considering data with uncertainties in some regions. The higher differences between the indices calculated from ERA-40 data prior to 1979 are likely to be caused by an unrealistic relation between SLP anomalies in areas with good data coverage and those in areas with poor coverage, where ERA-40 is more uncertain. ERA40PC is less dependent on high-latitude SLP than ERA40GW because of the area weighting and hence may be more reliable. The degree to which the two SAM indices differ if they are calculated from standard GCM simulations (without assimilation) may thus depend on how well the dynamical relationships between different regions are simulated and on the simulated position of the mean jet and associated features.

b. To determine how the quality of the reconstructions is related to the correlations between local SLP and the SAM indices and to the spatial distribution of the predictors

Correlation maps of the SAM indices with ERA-40 SLP (1979–2001) (contours in Figs. 3a,e, 4a,e, 5a,d, and 6a,d) show that the spatial signals of the ERA40PC and Marshall SAM indices are very similar, but with important differences relevant to reconstruction quality, in all seasons but JJA. In this season, the differences are largest in the eastern Pacific, a region with no station data available. Nonetheless, these correlation maps confirm previous findings that the SAM is least annular in austral winter (e.g., Fan 2007). This results from the differing jet and storm-track structure in this season (Codron 2007),

with a subtropical jet in the Indian and Pacific Oceans. Kidston et al. (2009) suggest the high baroclinicity over the southern flanks of the midlatitude continents results in the greater zonal symmetry of the SAM in summer.

While the high latitudes have high correlations between the SAM index and local SLP (>0.8) in all seasons, the differences in annularity change the SAM signal primarily in the midlatitudes. This strongly influences the strength of correlations between individual stations and the SAM index and hence the number of stations from the midlatitudes included in the reconstructions in each season. Station correlations are strongest in DJF and MAM when midlatitude zonal symmetry is highest (especially for ERA40PC), contributing to the better reconstruction quality of particularly JW in these seasons (section 3b). This results in best agreement between JW and Fogt reconstructions in these seasons (section 3c).

The structural difference between the ERA40PC and the Marshall index correlation maps in JJA results in higher correlations over New Zealand and Australia for Fogt, contributing to its better reconstruction quality in this season compared to JW. The uncertainties in the predictand pre-1979 in this season (and SON) also contribute to the lower reconstruction quality of JW.

SLP from observations, ERA-40, and HadSLP2 were used to investigate the spatial structure of the SLP anomalies in periods of strong positive and negative SAM index, to determine how well the few predictors do in capturing the SLP anomaly in each season. In many cases—for example, the recent period with positive SAM index in DJF and MAM—SLP anomalies are very SAM-like. Although the early 1960s are also associated with a strong positive SAM index, the hemispheric anomaly pattern is not SAM-like in all sectors, with regional asymmetries, for example, in the midlatitude Indian Ocean. Similarly, a number of other periods have positive SLP anomalies in regions known to be preferential to blocking but not a zonally symmetric SAM signature. For example, SON 1961 and JJA 1973 (not shown), show a wavenumber 1 pattern at high latitudes. The latter was identified by Trenberth and Mo (1985) as a season with strong blocking, and the station, ERA-40, and HadSLP2 anomalies correspond well with their Figs. 12 and 13. MAM 1958–59 shows strong positive anomalies in the southeast Pacific, a region where blocking is linked to ENSO related Rossby wave propagation (Renwick 2005; Renwick and Revell 1999; Kiladis and Mo 1998). Frequent winter blocking in the eastern Pacific may be related to the split-jet structure (and hence to the zonal asymmetry), with a weaker high-latitude jet (Codron 2007) providing less support for cyclogenesis and hence lower cyclone activity in this region (Simmonds et al. 2003).

These results corroborate the findings of Part II, where it is shown that the spatial patterns of the trends in the seasons that most probably result from greenhouse or ozone forcing are very SAM-like (DJF and MAM). The lower reconstructed trends in SON indicate low ozone and greenhouse forcing on the SAM or that natural variability overrides any trends. Some SLP anomaly patterns in SON project on the SAM but are not canonical SAM events. The pressure trends as calculated from HadSLP2 (Part II) in this season also do not project strongly onto the SAM, showing a wavenumber 3 pattern in the extratropics (Fig. 5f in Part II).

c. To determine whether there is a ''best'' reconstruction (based on fitting statistics and validation on independent data)

This question really addresses whether the Marshall index or the PC index can be better reconstructed. In DJF and MAM the JW1866 and Fogt reconstructions are of similar quality (from consideration of validation correlations, reduction of error, and error bars). Agreement is stronger with the JW1866–JW05 and Fogt than with JW51–JW58 and Fogt in DJF and MAM because of their more similar station networks. This indicates that, in these two seasons, JW51 and JW58 *may* be slightly more reliable, but JW and Fogt reconstructions are of similar quality pre-1958 (where there are the biggest unknowns). Hence both indices can be equally well reconstructed. In the recent period those reconstructions that use data from all centers of action (Marshall, JW58) may be less prone to potential errors because of the above-mentioned non-SAM-like SLP anomalies in some periods.

In JJA and SON the Fogt reconstructions are more reliable than JW05; hence the Marshall index can be better reconstructed. The similar validation statistics for JW51 and JW58 to Fogt indicate that the latter achieves a similar level of accuracy to these shorter reconstructions with a lower number of stations. This is corroborated by the higher correlations between the short JW reconstructions and Fogt than with JW05. Thus for the more recent period one can place a similar level of confidence in Fogt and JW58, but prior to this the Fogt reconstruction is more reliable (although in SON, particularly post-1920, agreement is strong between JW and Fogt). Use of corrected ERA-40 surface pressure data (Trenberth et al. 2005) may address these uncertainties.

The Visbeck reconstructions with the fixed network of 14–17 stations show much lower correlations with all other reconstructions, although this drop is lowest in DJF when zonal symmetry is strongest. This suggests that the early sections of the Vvar reconstructions that are constructed using fewer stations may also contain considerable uncertainty outside of austral summer. Nevertheless a number of features are present in all three reconstructions in all seasons.

d. To determine the SAM behavior over the past 150 yr based on a joint analysis of the Fogt, JW, and Visbeck reconstructions, in light of their uncertainties and robust features

To give a best estimate of the SAM index from the JW reconstructions, the concatenated index was considered. Reconstruction quality is high in both JW and Fogt in DJF (point c) and agreement in variability between the JW, Fogt, and Visbeck reconstructions is strongest in this season, thus we can place strong confidence in the features of reconstructions in this season. However, the JW mean is higher than the Fogt mean for much of the twentieth century. When put into a century-scale context, the low-frequency trend from mid-1960s to present is the strongest in the series in all reconstructions (see Part II for a detailed analysis of trends). Individual years in reconstructions and trends to individual points (e.g., 1910–62) are of similar magnitude in the Fogt and JW reconstructions.

In MAM agreement between long reconstructions is generally high (with short periods of exception around 1900 and 1950, and except between Fogt and Visbeck). Periods of positive SAM index and low-frequency variability of similar magnitude to the recent period occur throughout the record in all reconstructions. The greater uncertainty and hence wider error bars than in DJF should be noted. Station SLP anomalies associated with the peak in the 1930s project strongly onto the SAM pattern, and thus this peak is likely to represent a hemispheric SAM signal.

The most prominent feature in JJA is the low decadal-scale variability, particularly prior to 1960. There is disagreement between reconstructions between 1910 and 1930, with Fogt more negative than JW. Greater confidence can be placed in the former as it has been found to be more reliable in this season, and Vvar corroborates this. Marshall index trends in the latter part of the reconstruction are stronger because of negative values in the mid-1960s not present in JW or Fogt, but all trends are relatively weak and insignificant.

As found by other authors (e.g., M03), there is no recent positive trend in SON. The longer reconstructions reveal greater variability, both interannual and decadal, prior to compared to post-1965, which is particularly marked in the more reliable Fogt reconstruction. As this reconstruction is based on the same number of stations throughout, this variability is not a result of greater uncertainties during the earlier period. Station SLP anomalies for strong positive SAM during 1930–40 appear to result from a zonal wavenumber 1 pattern, but this is consistent with the peak in the 1960s as well. Therefore the SAM may have more of a zonal wavenumber 1 structure in this season than a purely zonally symmetric signature.

The analysis presented here makes it clear that differences in SAM structure between the seasons are important for reconstructions based on station data, as the precise location of SAM centers of action and the associated strength of the correlations between the SAM and local SLP over regions containing stations influences the strength of the relationship of the stations with the predictand. This is important because of the large areas of oceans in SH midlatitudes. This geographical limitation on station availability means that, in the early parts of the twentieth century, there is the possibility that the SAM index during some periods is either over- or underestimated in the reconstructions, depending on the regions in which the pressure anomalies are located. The magnitude of this effect can be partly estimated from correlations of the earlier reduced network with the fuller network reconstructions (Table 3) and from comparison of their validation statistics. The dropoff in correlations is greater for JW in JJA and SON. For the former this may reflect the influence of tropical forcing on the SAM (Fogt and Bromwich 2006). These structural differences also influence the validity of the assumption of zonal symmetry in the Visbeck reconstructions.

If evenly distributed stations were available around the hemisphere throughout the period of the reconstruction, this would not be a problem; however, as such a network of stations does not exist, one can only bear this in mind when considering the reconstructions. These factors add to the challenge of interpreting past climate variability in the SH. Nevertheless, this study also demonstrates that, for the period from 1957 onward, when high-latitude (Antarctic) data are available, the station-based Marshall index provides a very good representation of the SAM, which captures most of the variability described by a PC-based index.

As well as aiding understanding of past climate changes in the SH, SAM index estimates are needed to constrain, evaluate, and understand past and recent SAM variability in climate simulations. Part II demonstrates how such comparisons not only give insight in which models may provide better estimates of future SAM development but also aid process understanding.

APPENDIX

Tables

TABLE A1. Stations with data available since at least 1958 that are significantly correlated with the detrended ERA40PC 1958–2001 and hence used for the JW reconstructions, and their correlations with this index. Correlations in italic are significant at the 5% level, those in roman at the 1% level.

Station	Lat	Lon	Correlation with DJF ERA40PC	Correlation with MAM ERA40PC	Correlation with JJA ERA40PC	Correlation with SON ERA40PC
Davis	−68.6	78.0	−0.83	−0.80	−0.68	−0.45
Dumont D'Urville	−66.7	140.0	−0.74	−0.60	−0.67	−0.54
Halley	−75.5	−26.6	−0.89	−0.76	−0.49	−0.69
Mawson	−67.6	62.9	−0.85	−0.75	−0.68	−0.47
Mirny	−66.5	93.0	−0.79	−0.73	−0.56	−0.53
Scott	−77.9	166.8	−0.86	−0.74	−0.73	−0.73
Signy	−60.7	−45.6				*−0.31*

TABLE A2. Stations with data available since at least 1951 that are significantly correlated with the detrended ERA40PC 1958–2001, and hence used for the JW reconstructions, and their correlations with this index. Correlations in italic are significant at the 5% level, those in roman at the 1% level.

Station	Lat	Lon	Correlation with DJF ERA40PC	Correlation with MAM ERA40PC	Correlation with JJA ERA40PC	Correlation with SON ERA40PC
Aituti	−18.8	−159.8	0.40			
Bellingshausen	−62.2	−58.9	−0.63			−0.42
Campbell	−52.6	169.2	*0.35*	0.44	*0.39*	0.70
Carnarvon	−24.5	113.4	*−0.36*			
East London	−33.0	27.8	0.42	0.41		
Eziaza	−34.8	−58.5		0.56		
Esperenza	−63.4	−57.0	−0.68	−0.41		−0.52
Faraday/Vernadsky	−65.4	−64.4	−0.71	−0.47		−0.45
Funafiti	−8.5	179.2	*0.37*			
Gough Island	−40.4	−9.9	0.72			
Ile Nouvea	−37.8	77.5	0.62	0.39		*0.32*
Il Pen	−9.0	−158.1	0.55			
Juan Fernandez	−33.6	−78.8	0.45			
Johannesburg	−26.2	28.1			0.45	
Junin	−34.6	−61.0	0.44	0.59		
Kerguelen	−49.3	70.2	0.64	0.44		0.36
Mar del Plata	−37.9	−57.6	0.44	0.53		
Marion Island	−46.9	37.9	0.50			
MacQuarie Island	−54.6	158.9				0.58
Neuquen	−39.0	−68.1	0.59	0.47		
Pitcairn	−24.1	−130.1	0.56			
Pahuajo	−35.9	−65.9	0.44			
Puka puka	−10.9	−165.8	0.59			
Raoul	−29.2	−177.9			*0.34*	
Rarotonga	−21.2	−159.8	*0.38*		*0.32*	
Resistencia	−27.5	−59.1	0.47			
Rio Gallegos	−51.6	−69.3	0.44			
Rosario	−32.9	−60.8	0.43	0.46		
Sarmiento	−45.6	−69.1	0.42	0.51		
St. Denis	−20.5	55.3			0.43	
Tamatave	−18.1	49.4			0.39	
Trelew	−43.2	−65.3	0.42	0.60		

TABLE A3. Stations with data available since at least 1905 that are significantly correlated with the detrended ERA40PC 1958–2001, and hence used for the JW reconstructions, and their correlations with this index. Those with data available since 1865 that have been used in the JW1866 reconstructions are highlighted in bold. Correlations in italic are significant at the 5% level, those in roman at the 1% level.

Station	Lat	Lon	Correlation with DJF ERA40PC	Correlation with MAM ERA40PC	Correlation with JJA ERA40PC	Correlation with SON ERA40PC
Apia	−13.8	−171.8	*0.37*	*0.35*	*0.32*	
Ascuncion	−25.2	−57.5	0.39	**0.62**	***0.30***	
Auckland	−36.9	174.8	**0.48**	***0.34***	0.49	0.40
Bahia Blanca	−38.7	−62.2	**0.50**	**0.48**		
Buenos Aires	−34.6	−58.6	**0.53**	**0.64**		
Catamarca	−28.6	−64.5	*0.34*	*0.36*		
Chatham	−44.0	−176.6	0.61	0.55	0.69	0.65
Christchurch	−43.5	172.6	**0.55**	**0.46**	0.54	0.67
Cordoba	−31.4	−64.2	**0.47**	**0.45**		
Curitiba	−15.6	−56.1	**0.40**	***0.35***		
Dunedin	−45.9	170.5	**0.46**	**0.40**	0.53	0.56
Durban	−30.0	31.0	0.44			
Easter Island	−27.2	−109.4	**0.42**			
Goya	−29.2	−59.7	**0.53**	**0.54**		
Hobart	−42.9	147.3	**0.42**			0.55
Hokitika	−42.7	171.0	**0.56**	**0.41**	0.54	0.60
Majunga	−15.7	46.4			*0.37*	
Mauritius	−20.4	57.7			*0.37*	
Melbourne	−37.8	145.0				*0.30*
Montevideo	−34.9	−56.2		0.49		
Orcadas*	−60.7	−44.7	−0.46	**−0.25**		
Port Elizabeth	−34.0	25.6		*0.34*		
Punta Arenas	−53.0	−70.9	*0.31*			
Punta Tortuga	−29.9	−71.4	**0.64**			
Salta	−24.9	−65.5	0.50	0.59	*0.33*	
Santiago	−33.4	−70.8	***0.34***			
Stanley	−51.7	−57.9		*0.37*		
Sydney	−33.9	151.2	***0.36***	***−0.33***		
Tahiti	−17.6	−149.6	**0.58**			
Valdivia	−39.8	−73.2	0.57	0.57		
Wellington	−41.3	174.8	**0.52**	**0.44**	0.56	0.52

* Orcadas included when nearly significant in MAM ($p < 0.10$).

TABLE A4. Stations that are significantly correlated at the 5% level with the detrended Marshall SAM index 1957–2005 and hence included in the Fogt reconstructions.

Station	Lat	Lon	Correlation with DJF Marshall index	Correlation with MAM Marshall index	Correlation with JJA Marshall index	Correlation with SON Marshall index
Adelaide	−34.9	138.5			0.44	
Alice Springs	−23.8	133.9			0.33	
Auckland	−36.9	174.8	0.53	0.50	0.57	0.46
Brisbane	−27.4	153.1			0.47	
Buenos Aires	−34.6	−58.6	0.48	0.51		
Chatham	−44.0	−176.6			0.66	0.69
Christchurch	−43.5	172.6	0.66	0.61	0.66	0.71
Dunedin	−45.9	170.5	0.66	0.57	0.68	0.74
Durban	−30.0	31.0			0.40	
Hobart	−42.9	147.3	0.51	0.29	0.52	0.65
Hokitika	−42.7	171.0			0.64	0.62
Melbourne	−37.8	145.0			0.49	0.44
Orcadas	−60.7	−44.7			−0.19	−0.56
Perth	−31.9	116			0.54	
Port Elizabeth	−34.0	25.6			0.35	
Rio de Janeiro	−22.9	−43.2	0.41	0.31		
Santiago/Pudah	−33.4	−70.8	0.29	0.26		
Sydney	−33.9	151.2	0.39		0.50	0.34
Wellington	−41.3	174.8	0.62	0.59	0.66	0.66

REFERENCES

Allan, R. J., and T. Ansell, 2006: A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *J. Climate,* **19,** 5816–5842.

Angell, J. K., 1988: Variations and trends in tropospheric and stratospheric global temperatures, 1958–87. *J. Climate,* **1,** 1296–1313.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate,* **5,** 541–560.

Bromwich, D. H., and R. L. Fogt, 2004: Strong trends in the skill of the ERA-40 and NCEP–NCAR reanalyses in the high and middle latitudes of the Southern Hemisphere, 1958–2001. *J. Climate,* **17,** 4603–4619.

——, ——, K. I. Hodges, and J. E. Walsh, 2007: A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. *J. Geophys. Res.,* **112,** D10111, doi:10.1029/2006JD007859.

Codron, F., 2005: Relation between annular modes and the mean state: Southern Hemisphere summer. *J. Climate,* **18,** 320–330.

——, 2007: Relations between annular modes and the mean state: Southern Hemisphere winter. *J. Atmos. Sci.,* **64,** 3328–3339.

Cook, E. R., D. M. Meko, D. W. Stahle, and M. K. Cleveland, 1999: Drought reconstructions for the continental United States. *J. Climate,* **12,** 1145–1163.

Fan, K., 2007: Zonal asymmetry of the Antarctic Oscillation. *Geophys. Res. Lett.,* **34,** L02706, doi:10.1029/2006GL028045.

Fogt, R. L., and D. H. Bromwich, 2006: Decadal variability of the ENSO teleconnection to the high-latitude South Pacific governed by coupling with the southern annular mode. *J. Climate,* **19,** 979–997.

——, J. Perlwitz, A. J. Monaghan, D. H. Bromwich, J. M. Jones, and G. J. Marshall, 2009: Historical SAM variability. Part II: Twentieth-century variability and trends from reconstructions, observations, and the IPCC AR4 models. *J. Climate,* **22,** 5346–5365.

Fritts, H. C., J. Guiot, G. A. Gordon, and F. Schweingruber, 1990: Methods of calibration, verification, and reconstructions. *Methods of Dendrochronology,* E. R. Cook and C. A. Kairiukstis, Eds., Kluwer, 163–217.

Gillett, N. P., and D. W. J. Thompson, 2003: Simulation of recent Southern Hemisphere climate change. *Science,* **302,** 273–275.

——, T. D. Kell, and P. D. Jones, 2006: Regional climate impacts of the Southern Annular Mode. *Geophys. Res. Lett.,* **33,** L23704, doi:10.1029/2006GL027721.

Gong, D. Y., and S. W. Wang, 1999: Definition of Antarctic Oscillation index. *Geophys. Res. Lett.,* **26,** 459–462.

Hall, A., and M. Visbeck, 2002: Synchronous variability in the Southern Hemisphere atmosphere, sea ice, and ocean resulting from the annular mode. *J. Climate,* **15,** 3043–3057.

Hines, K. M., D. H. Bromwich, and G. J. Marshall, 2000: Artificial surface pressure trends in the NCEP–NCAR reanalysis over the Southern Ocean. *J. Climate,* **13,** 3940–3952.

Jones, J. M., and M. Widmann, 2003: Instrument- and tree-ring-based estimates of the Antarctic Oscillation. *J. Climate,* **16,** 3511–3524.

——, and ——, 2004: Early peak in Antarctic oscillation index. *Nature,* **432,** 290–291.

Jones, P. D., 1987: The early twentieth century Arctic high—Fact or fiction? *Climate Dyn.,* **1,** 63–75.

——, and D. H. Lister, 2007: Intercomparison of four different Southern Hemisphere sea level pressure datasets. *Geophys. Res. Lett.,* **34,** L10704, doi:10.1029/2007GL029251.

——, M. J. Salinger, and A. B. Mullan, 1999: Extratropical circulation indices in the Southern Hemisphere based on station data. *Int. J. Climatol.,* **19,** 1301–1317.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kidson, J. W., 1988: Interannual variations in the Southern Hemisphere circulation. *J. Climate,* **1,** 1177–1198.

Kidston, J., J. A. Renwick, and J. McGregor, 2009: Hemispheric-scale seasonality of the southern annular mode and impacts on the climate of New Zealand. *J. Climate,* **22,** 4759–4770.

Kiladis, G. N., and K. C. Mo, 1998: Interannual and intraseasonal variability in the Southern Hemisphere. *Meteorology of the Southern Hemisphere, Meteor. Monogr.,* No 49, Amer. Meteor. Soc., 307–336.

Kushner, P. J., I. M. Held, and T. L. Delworth, 2001: Tropospheric response to stratospheric ozone loss. *Geophys. Res. Lett.,* **28,** 1547–1550.

Kwok, R., and J. C. Comiso, 2002: Spatial patterns of variability in Antarctic surface temperature: Connections to the Southern Hemisphere Annular Mode and the Southern Oscillation. *Geophys. Res. Lett.,* **29,** 1705, doi:10.1029/2002GL015415.

Lefebvre, W., H. Goosse, R. Timmermann, and T. Fichefet, 2004: Influence of the Southern Annular Mode on the sea ice–ocean system. *J. Geophys. Res.,* **109,** C09005, doi:10.1029/2004JC002403.

Limpasuvan, V., and D. L. Hartmann, 2000: Wave-maintained annular modes of climate variability. *J. Climate,* **13,** 4414–4429.

Marshall, G. J., 2002: Trends in Antarctic geopotential height and temperature: A comparison between radiosonde and NCEP–NCAR reanalysis data. *J. Climate,* **15,** 659–674.

——, 2003: Trends in the southern annular mode from observations and reanalyses. *J. Climate,* **16,** 4134–4143.

——, 2007: Half-century seasonal relationships between the Southern Annular Mode and Antarctic temperature. *Int. J. Climatol.,* **27,** 373–383.

Meehl, G. A., and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 747–846.

Miller, R. L., G. A. Schmidt, and D. T. Shindell, 2006: Forced annular variations in the 20th century Intergovernmental Panel on Climate Change Fourth Assessment Report models. *J. Geophys. Res.,* **111,** D18101, doi:10.1029/2005JD006323.

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.

Renwick, J. A., 2005: Persistent positive anomalies in the Southern Hemisphere circulation. *Mon. Wea. Rev.,* **133,** 977–988.

——, and M. J. Revell, 1999: Blocking over the South Pacific and Rossby wave propagation. *Mon. Wea. Rev.,* **127,** 2233–2247.

Rogers, J. C., and H. van Loon, 1982: Spatial variability of sea level pressure and 500-mb height anomalies over the Southern Hemisphere. *Mon. Wea. Rev.,* **110,** 1375–1392.

Schneider, D. P., E. J. Steig, and J. C. Comiso, 2004: Recent climate variability in Antarctica from satellite-derived temperature data. *J. Climate,* **17,** 1569–1583.

Schwerdtfeger, W., 1984: *Weather and Climate of the Antarctic.* Developments in Atmospheric Science, Vol. 15, Elsevier, 261 pp.

Sen Gupta, A. S., and M. H. England, 2006: Coupled ocean–atmosphere–ice response to variations in the southern annular mode. *J. Climate,* **19,** 4457–4486.

Simmonds, E., K. Keay, and E. P. Lim, 2003: Synoptic activity in the seas around Antarctica. *Mon. Wea. Rev.,* **131,** 272–288.

Stone, D. A., A. J. Weaver, and R. J. Stouffer, 2001: Projection of climate change onto modes of atmospheric variability. *J. Climate,* **14,** 3551–3565.

Szeredi, I., and D. Karoly, 1987: The vertical structure of monthly fluctuations of the Southern Hemisphere troposphere. *Aust. Meteor. Mag.,* **35,** 19–30.

Thompson, D. W. J., and J. M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate,* **13,** 1000–1016.

——, and S. Solomon, 2002: Interpretation of recent Southern Hemisphere climate change. *Science,* **296,** 895–899.

Trenberth, K. E., 1991: Storm tracks in the Southern Hemisphere. *J. Atmos. Sci.,* **48,** 2159–2178.

——, and K. C. Mo, 1985: Blocking in the Southern Hemisphere. *Mon. Wea. Rev.,* **113,** 3–21.

——, D. P. Stepaniak, and L. Smith, 2005: Interannual variability of patterns of atmospheric mass distribution. *J. Climate,* **18,** 2812–2825.

Turner, J., and Coauthors, 2004: The SCAR READER project: Toward a high-quality database of mean Antarctic meteorological observations. *J. Climate,* **17,** 2890–2898.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.

Villalba, R., E. R. Cook, R. D. D'Arrigo, G. C. Jacoby, P. D. Jones, M. J. Salinger, and J. Palmer, 1997: Sea-level pressure variability around Antarctica since A.D. 1750 inferred from subantarctic tree-ring records. *Climate Dyn.,* **37,** 375–390.

Visbeck, M., 2009: A station-based southern annular mode index from 1884–2005. *J. Climate,* **22,** 940–950.

Widmann, M., 2005: One-dimensional CCA and SVD, and their relationship to regression maps. *J. Climate,* **18,** 2785–2792.