



Can current reanalyses accurately portray changes in Southern Annular Mode structure prior to 1979?

Gareth J. Marshall¹ · Ryan L. Fogt² · John Turner¹ · Kyle R. Clem³

Received: 25 November 2021 / Accepted: 6 April 2022
© Crown 2022

Abstract

Early reanalyses are less than optimal for investigating the regional effects of ozone depletion on Southern Hemisphere (SH) high-latitude climate because the availability of satellite sounder data from 1979 significantly improved their accuracy in data sparse regions, leading to a coincident inhomogeneity. To determine whether current reanalyses are better at SH high-latitudes in the pre-satellite era, here we examine the capabilities of the European Centre for Medium-range Weather Forecasts (ECMWF) fifth generation reanalysis (ERA5), the Twentieth Century Reanalysis version 3 (20CRv3), and the Japanese Meteorological Agency (JMA) 55-year reanalysis (JRA-55) to reproduce and help explain the pronounced change in the relationship between the Southern Annular Mode (SAM) and Antarctic near-surface air temperatures (SAT) between 1950 and 1979 (EARLY period) and 1980–2020 (LATE period). We find that ERA5 best reproduces Antarctic SAT in the EARLY period and is also the most homogeneous reanalysis across the EARLY and LATE periods. ERA5 and 20CRv3 provide a good representation of SAM in both periods with JRA-55 only similarly skilful in the LATE period. Nevertheless, all three reanalyses show the marked change in Antarctic SAM-SAT relationships between the two periods. In particular, ERA5 and 20CRv3 demonstrate the observed switch in the sign of the SAM-SAT relationship in the Antarctic Peninsula: analysis of changes in SAM structure and associated meridional wind anomalies reveal that in these reanalyses positive SAM is linked to cold southerly winds during the EARLY period and warm northerly winds in the LATE period, thus providing a simple explanation for the regional SAM-SAT relationship reversal.

Keywords Southern Annular Mode (SAM) · Climate change · Antarctica · Temperature · Reanalysis · Validation

1 Introduction

Global gridded reanalysis data are, in theory, particularly appropriate for studying climate variability and change in observation-sparse regions of the globe such as Antarctica. However, although the forecast models are kept constant to avoid temporal inhomogeneities related to changes in the model physics, it has long been known that changes in the availability of the type and coverage of assimilated data in remote regions can in itself produce spurious climate

changes within a reanalysis (e.g., Hines et al. 2000; Marshall and Harangozo 2000; Bromwich and Fogt 2004). In particular, reanalyses have always struggled in the data-sparse Southern Hemisphere (SH) high-latitudes before 1979, prior to the availability of satellite sounder data from the Television InfraRed Observation Satellite (TIROS) (e.g., Marshall 2003). Although earlier sounder data exist, the TIROS Operational Vertical Sounder (TOVS) instrument provided the first meteorological data over the Southern Ocean to be widely assimilated into reanalyses. Utilising reanalyses to better understand Antarctic climate variability before this time is difficult because of the sometimes considerable temporal changes in their accuracy. Significantly, 1979 coincides approximately with the advent of the Antarctic ozone hole, the biggest driver of recent climate change in the SH high-latitudes (e.g., Thompson and Solomon 2002; Thompson et al. 2011; Polvani et al. 2011) and thus reanalyses have been less than optimal for comparing SH high-latitude climate pre- and post-ozone depletion.

✉ Gareth J. Marshall
gjma@bas.ac.uk

¹ British Antarctic Survey, UK Research and Innovation, Cambridge, UK

² Department of Geography, Ohio University, Athens, OH, USA

³ School of Geography, Environment and Earth Sciences, Victoria University of Wellington, Wellington, New Zealand

One of the principal climatic impacts of ozone depletion has been through driving changes in the Southern Annular Mode (SAM), which have, in turn, contributed to changes in SH high-latitude surface climate. The SAM is the leading mode of extra-tropical SH climate variability (for a recent review, see Fogt and Marshall 2020). The 'characteristic circumpolar SAM structure' (Wachter et al. 2020), which refers to the first empirical orthogonal function (EOF) of sea level pressure (SLP) or geopotential height (or the pattern of regression or correlation between the SAM and SLP or geopotential height), broadly approximates to a zonally symmetric or annular structure, with synchronous pressure anomalies of opposite sign in the SH mid- and high-latitudes. However, it is important to note that within the climatological SAM structure there is a weak zonal wave-number 3 pattern (e.g., Raphael 2004; Goyal et al. 2021) but with one of the nodes comprising an often significant asymmetric component, centred in the South Pacific (Fan 2007; Fogt et al. 2012), and associated with teleconnections from the tropical Pacific (e.g., Ding et al. 2012; Schneider et al. 2012; Clem et al. 2016). Figure 1a shows the correlation between the SAM and SLP based on the 35-year period of European Centre for Medium-range Weather Forecasts (ECMWF) fifth generation reanalysis (ERA5) data from 1980 to 2014 (ERA5 is described in Sect. 2.2.1). Essentially, it shows the expected annular structure, with statistically significant positive (negative) correlations north of $\sim 50^{\circ}\text{S}$ (south of $\sim 60^{\circ}\text{S}$). The primary departure from this pattern is the northward extension of the region with a negative correlation to 45°S at $\sim 105^{\circ}\text{W}$, in the region of the climatological Amundsen Sea Low (ASL) (Raphael et al. 2016). The positive polarity of the SAM (hereinafter SAM+) is defined as being when relatively high (low) pressures occur over mid-latitudes (Antarctica) and is thus associated with stronger circumpolar westerlies around Antarctica. Recent positive trends in the SAM have occurred during austral summer and autumn, with the former driven primarily by ozone depletion (e.g., Thompson et al. 2011; Fogt and Marshall 2020).

Variability in the SAM polarity has a marked influence on Antarctic near-surface temperatures (SAT). Typically, SAM+ is associated with positive SAT anomalies over the Antarctic Peninsula and negative SAT anomalies over the majority of the rest of the continent, a spatial pattern first described by Thompson and Solomon (2002) and examined further by others (e.g., Gillett et al. 2006; Marshall 2007; Marshall and Thompson 2016). It results from the interaction of the orography of the Antarctic continent with the regional circulation. The Antarctic Peninsula extends northwards into the latitudes of the circumpolar westerlies and the climate of its western side is often influenced by northerly maritime winds along the eastern edge of the ASL to the west (c.f., Fig. 1a). With SAM+, the enhanced westerly winds impinging on the Peninsula are more likely to push

air masses over the mountainous spine, whereafter they warm adiabatically as they descend on the lee side, leading to downslope föhn winds that cause rapid increases in temperature on the eastern Peninsula (e.g., Orr et al. 2008). The positive trend in the SAM in austral summer has led to a greater frequency of these warming events and the resultant surface melting has likely contributed to the collapse of parts of the Larsen B Ice Shelf (e.g., Marshall et al. 2006; Cape et al. 2015). Elsewhere across Antarctica, the stronger circumpolar westerlies accompanying SAM+ reduce the meridional flux of heat and moisture into the continent at most longitudes, isolating it from the warmer mid-latitude maritime air so that SAT anomalies are predominantly negative (e.g., Marshall and Thompson 2016). In addition, SAM+ is associated with weaker katabatic drainage over East Antarctica, reducing turbulent heat flux towards the surface, resulting in a stronger surface temperature inversion and thus lower SAT (van den Broeke and van Lipzig 2003).

While this SAM-SAT spatial pattern may be the current climatological situation, several studies have described regional temporal variations across periods of several years post 1979 (Marshall et al. 2011; 2013; Wachter et al. 2020). These tend to be associated with changes in the phase and/or magnitude of the atmospheric wave number 3 pattern, which may in turn be linked to changes in tropical convection (Goyal et al. 2021) or associated with large-scale patterns of ocean sea surface temperature (SST) variability such as the Interdecadal Pacific Oscillation (IPO) and Atlantic Multidecadal Variability (AMV; also known as the Atlantic Multidecadal Oscillation) (Wachter et al. 2020). However, observations also suggest an earlier, notable shift in the Antarctic SAM-SAT relationship in the 1970s, as portrayed recently in Fig. 10 of Turner et al. (2020). Previously, Silvestri and Vera (2009) described a poleward migration of the mid-latitude pressure anomaly centres within the SAM structure from the 1960s/70s to the 1980s/90s, especially over South America and Australia. Fogt et al. (2012) also noted marked changes in the asymmetric component of the SAM around 1980.

An apparent consequence of this change is depicted in Fig. 1b, which shows running decadal (10-year) correlations between annual SAT observations at 14 Antarctic meteorological stations (locations given in Fig. 1a) and the SAM. Both the SAT and SAM datasets utilised here are described in Sect. 2 and can be considered temporally homogeneous. The correlations are provided for the 55 running decades from 1957 to 66 to 2011–2020. It is apparent there is a major change in the SAM-SAT relationship across Antarctica between the decades up to 1970–79 and from 1971 to 80 onwards (hereinafter, termed the EARLY (1950–1979) and LATE (1980–2020) periods, respectively). A similar figure, based on 5-year running means (not shown) indicates a clear change between 1975

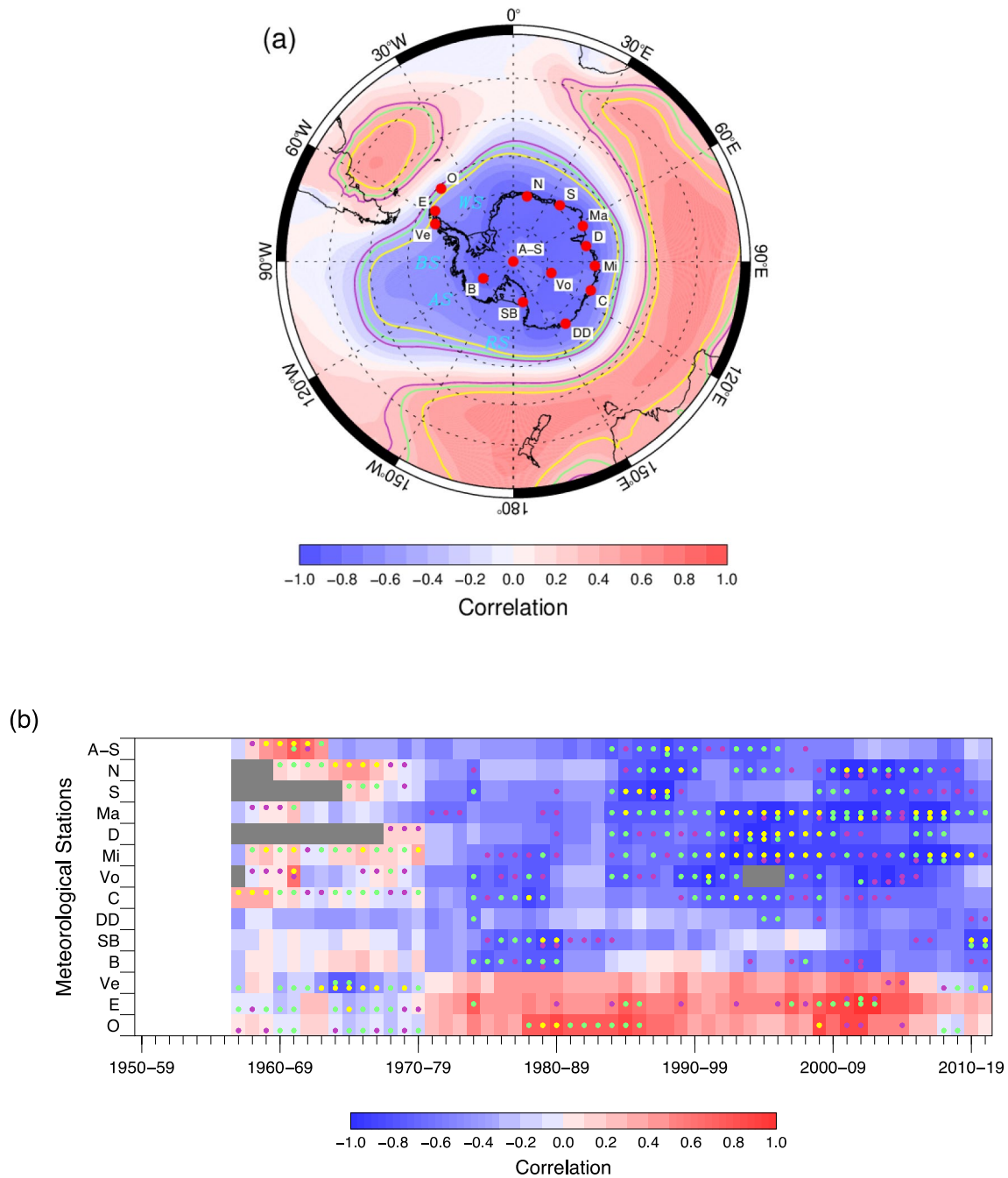


Fig. 1 **a** Mean correlation between the annual SAM index and SLP for 1980–2014 from ERA5. **b** Correlations between the decadal running observation-based annual SAM index and annual mean SAT at 14 Antarctic meteorological stations. Dots in the centre of the cell show statistical significance; dots below (above) centre indicate if the correlation is significantly lower (higher) than the climatological value (1981–2015). Statistical significance values are $p < 0.10$ (pur-

ple), $p < 0.05$ (green) and $p < 0.01$ (yellow). Meteorological station acronyms: Amundsen-Scott (A-S) at the South Pole and then east from the Greenwich Meridian (0°), Novolazarevskaya (N), Syowa (S), Mawson (Ma), Davis (D), Mirny (Mi), Vostok (V), Casey (C), Dumont d'Urville (DD), Scott Base (SB), Byrd (B), Vernadsky (Ve), Esperanza (E) and Orcadas (O). Oceanic names: Amundsen Sea (AS), Bellingshausen Sea (BS), Ross Sea (RS) and Weddell Sea (WS)

and 79 and 1976–80, signifying that this shift in annual Antarctic SAM-SAT relationships is essentially independent of the length of the correlation period.

In the LATE period the well-documented positive (negative) correlation between the SAM and SAT in the Antarctic Peninsula (West and East Antarctica) is dominant with very

few exceptions to this. Most notably, there are some negative correlations in recent decades at both Vernadsky and Orcadas on the Peninsula that, while not statistically significant, lie outside the expected range of decadal correlation values at $p < 0.05$, as determined using synthetic data based on the 1981–2015 period (c.f., Sect. 2.4). Turner et al. (2016) provide a detailed explanation for these episodes.

However, in the EARLY period there are several decades when all the stations apart from Dumont d'Urville had SAM-SAT correlations of the opposite sign to the LATE period, both in the Peninsula and East Antarctica. Some of these 'reversed' decadal correlations are statistically significant: for example, positive correlations at Amundsen-Scott and Vostok and negative correlations at Vernadsky. Furthermore, many of the other decadal correlations in the EARLY period lie outside their expected ranges as defined above. Finally, it is also worth noting that while the majority of the Antarctic stations do show clear switches in the magnitude and/or sign of the SAM-SAT correlation between 1970–79 and 1971–80, some, such as Amundsen-Scott and Mawson, do not: in these cases, the switch occurs several years earlier (c.f., Fig. 1b).

In this paper, we examine the ability of three current reanalyses that extend backwards in time beyond 1979 to (i) reproduce this marked change in the spatial pattern of Antarctic SAM-SAT relationships between the EARLY and LATE periods and (ii) help explain the reasons behind it. The three reanalyses are ERA5, the Twentieth Century Reanalysis version 3 (20CRv3), and the Japanese Meteorological Agency (JMA) 55-year reanalysis (JRA-55). Following a description of the data and statistical methods employed (Sect. 2), we begin by validating the reanalyses against the SAT observations from the 14 Antarctic meteorological stations and the observation-based SAM Index (Sect. 3). In particular, we focus on the homogeneity of the reanalyses between the EARLY and LATE periods to ascertain whether they may be suitable for undertaking comparisons of pre- and post-ozone SH high-latitude climate. In Sect. 4 we compare the three reanalyses in terms of changes in the circumpolar SAM structure between the EARLY and LATE periods. Finally, in Sect. 5, we summarise our conclusions and briefly consider whether the spatial changes in SAM structure can be linked to changes in tropical SST variability.

2 Materials and methods

2.1 SAT observations

We utilise monthly SAT observations from 14 Antarctic stations to produce the annual SAT values (simply the mean of the 12 calendar months, January–December) (Fig. 1). These time-series all started at least one decade prior to the change from the EARLY to LATE period, with most commencing in the International Geophysical Year (IGY) of 1957/58. The majority of these data were obtained from the quality-controlled READER (REference Antarctic Data for Environmental Research) dataset (Turner et al. 2004; Turner et al. 2020). In addition, we use the reconstructed Byrd station SAT record (Bromwich et al. 2013, 2014) because of the dearth of other station data from West Antarctica. Note that we do not include the SAT record from Halley station due to inhomogeneities within it that are associated with station moves (King et al. 2021). Regarding missing data, annual SAT values are only computed if all 12 monthly values are available and subsequent decadal SAM-SAT correlations calculated if eight of the ten annual SAT values exist.

2.2 Reanalysis data

Here we provide a brief description of the three different reanalyses used in this study. Key characteristics and the online archives from where the reanalysis data were downloaded are given in Table 1.

2.2.1 European centre for medium-range weather forecasts (ECMWF) fifth generation reanalysis (ERA5)

The latest reanalysis from ECMWF, called ERA5, is described in Hersbach et al. (2020). Its predecessor, ERA-Interim, was generally considered as one of the best to use in the Antarctic region (e.g., Bracegirdle and Marshall 2012) although, as with many reanalyses, there were significant errors in the model orography that led to marked biases in SAT. Compared to ERA-Interim, ERA5 employs an improved version of the IFS Earth System Model (cycle 41r2) and associated 4D-Var assimilation scheme, as used in

Table 1 Details of the three reanalyses used in the study

Reanalysis	Reference	Horizontal grid of downloaded data	Assimilation method	Data link
ERA5	Hersbach et al. (2020)	1440×721	4D VAR	https://cds.climate.copernicus.eu
20CRv3	Silvinski et al. (2019)	360×181	4D IAU	https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html
JRA-55	Kobayashi et al. (2015)	640×320	4D VAR	https://rda.ucar.edu/datasets/ds628.1

the ECMWF operational product in 2016. ERA5 also has a considerably higher spatial resolution of 31 km, 137 vertical levels to 1 hPa, and a temporal resolution of 1 h.

Zhu et al. (2021) demonstrated that ERA5 performs marginally better than ERA-Interim when validated against meteorological station observations across Antarctica as a whole for the 1979–2018 period. Bozkurt et al. (2020) also found that the bias in ERA5 is broadly similar to the older reanalysis in the Antarctic Peninsula region, although some of the seasonal SAT trends are different. However, Tetzner et al. (2019) suggested that ERA5 showed significant improvements over ERA-Interim in reproducing SAT in the southern Peninsula.

ERA5 SAT and SLP data were acquired on a 721×1440 lat/lon grid. The EARLY period ERA5 data are called the ERA5 back extension (BE) data, and this paper represents the first time they have been analysed for the Antarctic region. These data have been termed the preliminary version of ERA5 BE, primarily because tropical cyclones are too intense, which is unlikely to have any significant impact on Antarctic SAT. ERA5 BE benefits from lessons learnt from ECMWF centennial reanalyses, in particular the optimal use of ocean boundary datasets, radiation forcing and coupling of the atmosphere with other components of the Earth system (Herzbach et al. 2020). Nonetheless, Gusti (2020) noted a warm bias of up to 1 °C in SAT in ERA5 BE over Australia prior to 1970.

2.2.2 Twentieth century reanalysis version 3 (20CRv3)

The 20CRv3 reanalysis is the latest version of 20CR, originally described by Compo et al. (2011), with the improvements incorporated into version 3 outlined by Silviski et al. (2019). 20CRv3 is different to the other two reanalyses examined here, in that it has been specifically designed to span more than a century, termed a historical reanalysis, and only assimilates surface pressure data. Therefore, the change in satellite sounder availability in 1979 should not have any direct effect and thus the EARLY and LATE periods might be expected to be more homogeneous in this reanalysis. 20CRv3 uses the National Centers for Environmental Prediction (NCEP) Global Forecast System v14.0.1 model, operational in 2017. Key updates in version 3 include an upgraded assimilation scheme, which incorporates a 4D incremental update analysis and uses an adaptive rather than fixed localisation length for the quality control, a higher-resolution forecast model and a larger set of pressure observations made available through various data-mining projects. 20CRv3 also has 80 ensemble members, the spread of which provides an internal measure of reanalysis uncertainty. In this study we primarily focus on the ensemble mean but do use all ensemble members to qualitatively examine the

temporal change in uncertainty of the derived SAM Index. The 20CRv3 data were acquired on a 360×181 lat/lon grid.

Silviski et al. (2019) described a problem with spurious SLP trends at SH high latitudes in 20CRv3, whereby modern values are approximately 5 hPa lower than in the early twentieth century. Moreover, SH confidence fields demonstrate that uncertainty in SLP over parts of Antarctica remains higher than climatology in the early 20th Century. It is also worth noting that the original 20CR had several major issues in reproducing Antarctic SAT, such as artificial statistically significant negative SAT trends since the late 1970s and a distinct jump in temperature around 1950, likely associated with changes in the assimilated observation counts (Zhang et al. 2018).

2.2.3 Japanese meteorological agency (JMA) 55-year reanalysis (JRA-55)

The JRA-55 has been produced with the 2009 version of JMA's operational assimilation scheme. Key improvements in JRA-55 compared to earlier JMA products include a revised longwave radiation scheme and 4D VAR assimilation scheme, including variational bias correction for satellite radiances (Kobayashi et al. 2015). Figure 2e in their paper is a clear demonstration of the seasonal cycle in the number of surface observations at SH high latitudes, with a marked peak when meteorological data from summer-only Antarctic bases become available. The spatial resolution of JRA-55 is ~55 km with 60 vertical levels and the data were acquired on a 640×320 lat/lon grid.

There have been a number of studies comparing JRA-55 with other, older reanalyses. Jones et al. (2016) stated that JRA-55 showed a marked improvement in reproducing summer temperatures compared to winter for the Amundsen Bay region of Antarctica: the authors surmised that this reanalysis has more skill at capturing the weakly stable summer boundary layer than others with a similar bias across all seasons. Wang et al. (2016) showed that JRA-55, along with some other reanalyses, produced spurious warming trends from 1979 to 2014 in parts of East Antarctica, where observations indicate cooling has occurred. Importantly, Huai et al. (2019) noted an abrupt jump in annual mean Antarctic SAT in JRA-55 around 1979, suggesting that the inhomogeneity is a direct result of the assimilation of the satellite sounder data.

2.3 Validation

There are two principal sections to the validation exercise. The first is to compare the capability of the reanalyses to reproduce Antarctic SAT at the 14 station locations (bilinearly interpolated from the different model grids), as established using three error statistics. These are the root mean

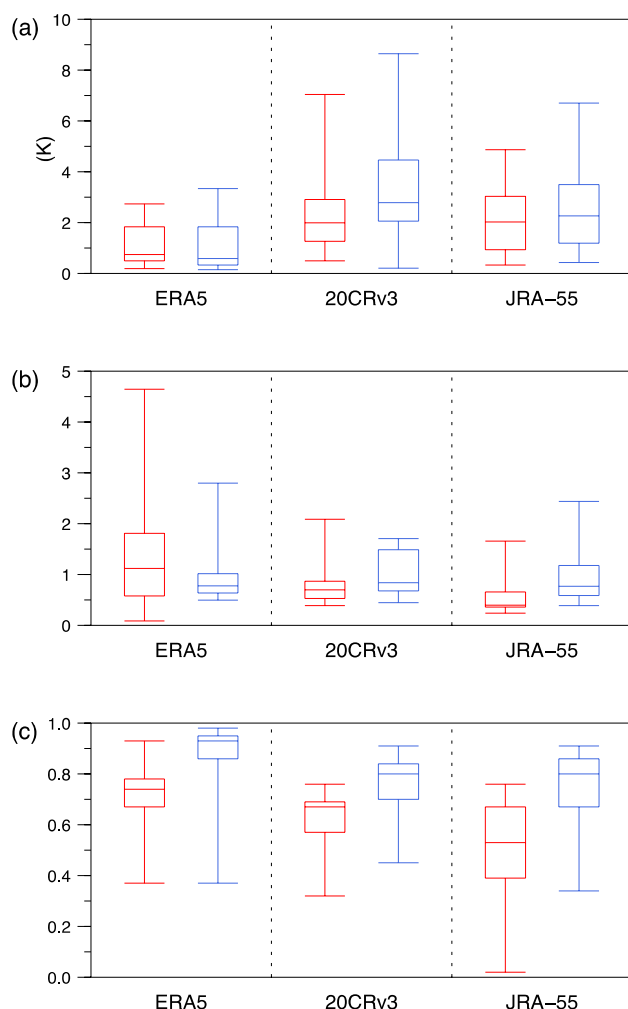


Fig. 2 Box-whisker plots of **a** RMSE error, **b** variance ratio and **c** correlation of SAT at 14 Antarctic meteorological stations in three reanalyses for 1958–1978 (red) and 1979–2015 (blue). The median, upper and lower quartiles (boxes) and range (whiskers) of the 14 values are shown

square error (RMSE), the ratio of the variance (reanalysis to observations), and the correlation coefficient r . Note that the RMSE is partially dependent on the accuracy of the model orography at the individual station locations, which can be markedly different from reality, leading to large biases in the reanalysis SATs: for example, in this analysis the largest mean annual SAT bias exceeds 11 K at Vostok in 20CRv3. We have not made any adjustment to account for the inaccurate station heights in the model orography because lapse rates are likely to be highly variable, particularly for coastal stations (Jones et al. 2016), so any adjustment would be highly approximate, and its accuracy would vary temporally. Moreover, here we are primarily interested in the temporal homogeneity of the model data rather than their absolute accuracy. Using annual SAT data, we determine how homogenous each reanalysis is between (i) the EARLY

and LATE periods, and (ii) between each other, again considering both periods separately. To enable a direct comparison of the three reanalyses with their slightly differing time periods, in this validation the EARLY and LATE periods are restricted to 1958–79 and 1980–2015, respectively. In addition, as previous studies of reanalyses in Antarctica have revealed marked seasonal variation in their accuracy (e.g., Marshall 2003; Bromwich and Fogt 2004), we also examine the accuracy of ERA5 across the four standard SH meteorological seasons: autumn (March–April–May), winter (June–July–August), spring (September–October–November) and summer (December–January–February).

The second validation is to compare an annual SAM Index derived from each reanalysis with the observation-based index of Marshall (2003), which was created to be independent of the availability of satellite sounder data and thus should be homogeneous across the EARLY and LATE periods. It does not begin until 1957 because the majority of Antarctic stations did not commence observations until the IGY of 1957/58 (Turner et al. 2004). To enable a direct comparison, for validation purposes only the reanalysis SAM indices are similarly defined: that is the normalised SLP difference between the mean of six meteorological station locations at $\sim 40^\circ\text{S}$ and six at $\sim 65^\circ\text{S}$. All the SAM Indices were normalised using the 1981–2010 period, rather than the original 1971–2000 period of the online observation-based SAM, to clearly separate any differences between the EARLY and LATE periods. In addition, we include a statistically-derived station-based reconstructed SAM Index in the validation to compare with the reanalysis-derived indices. Using a principal component regression methodology, the 'Fogt reconstruction' makes use of the availability of mid-latitude SLP observations throughout the twentieth century and earlier to derive a SAM index (calibrated to the Marshall (2003) index) prior to the advent of widespread high-latitude SH SLP data in the late 1950s (Fogt et al. 2009; Jones et al. 2009). Note that the Fogt reconstruction comprises the mean of four standard seasonal values so we might expect small differences due to the two-month shift in calculating the annual data used.

2.4 Calculation of the SAM Index

We use a slightly different definition of the SAM for our analysis outside of the validation exercise, based on that employed by Gong and Wang (1999). The annual mean zonal SLP at 40°S and 65°S in the three reanalyses is calculated as the average of 5° longitudinally spaced grid-point values around each latitude circle to provide a true hemispheric signal. The SAM Index is then computed as the normalised difference between these two values using the standard methodology.

2.5 Statistical methods

In the figures showing the spatial correlation between the SAM and SLP the significance is calculated using the false discovery rate (FDR) method outlined by Wilks (2016), which accounts for spatial autocorrelation within the data. Following the suggestion within that paper, $\alpha_{FDR} = 2\alpha_{global}$ where α is the significance level, as the SAM–SLP correlation data exhibit moderate to strong spatial correlation. Using running decadal correlations means that there is also significant temporal autocorrelation. For statistical analyses comparing the EARLY and LATE periods, the likely degrees of freedom in the data are estimated to be 1.5 times the number of non-overlapping segments (see Allen and Smith 1996).

To determine whether the reanalyses show statistically significant different levels of skill in reproducing observations, either between the EARLY and LATE periods or between each other, the Wilcoxon signed-rank test is employed (Wilcoxon 1945). This is a non-parametric test as the three validation statistics cannot be assumed to be normally distributed.

To determine the periods of non-stationarity in the decadal relationship between the SAM and Antarctic SAT, we calculate synthetic running correlations, using a similar method to Gallant et al. (2013). Non-stationarity is defined when the running decadal correlation coefficient lies outside a given confidence interval generated from 10,000 sets of synthetic stochastic data for each location. Each set was 35 years in length (or 26 running decades) and produced by adding 'local climate noise' (c.f., Gallant et al. 2013) to the regression relationship between SAM and SAT for 1981–2015, as calculated from either observations or reanalysis data. Thus, there are a total of 260,000 values with which to form a probability distribution function (PDF) of correlations. These are then transformed into Fisher z scores so that the data are normally distributed and the six confidence intervals of the z score PDF—at 90%, 95% and 99%, above and below the mean—can be computed. Finally, these z scores are transformed inversely back to the correlation values defining non-stationarity at the different confidence intervals.

In order to quantitatively analyse changes in SAM structure between the EARLY and LATE periods, the zonal SAM–SLP correlation values at 55°S are decomposed into the first four waves using standard Fourier analysis techniques, similar to Marshall and Bracegirdle (2015). The statistical significance of differences between the populations of the amplitude, phase and variance explained of these waves between the EARLY and LATE periods are calculated using a Wilcoxon test, as discussed above.

3 Results

3.1 SAT validation

3.1.1 Comparison of the three reanalyses

Results for the individual stations and the mean of the 14 stations are given in Tables S1, S2 and S3 for ERA5, 20CRv3 and JRA-55, respectively.

Analysis of the RMSE data reveals that there is no statistically significant difference between values in the EARLY and LATE periods for any of the three reanalyses. However, while the two RMSE values for ERA5 are indeed very similar (Fig. 2a), those in 20CRv3 and JRA-55 are actually larger in the LATE period: e.g. 20CRv3 has median RMSE values of 2.12 °C and 2.91 °C in the EARLY and LATE periods, respectively. It is apparent that 20CRv3 and JRA-55 have some RMSE values that are particularly high: these are predominantly for stations located on the Antarctic Plateau, such as Amundsen–Scott and Vostok, indicative of errors in the model orography. Interestingly, the RMSE values for Byrd in West Antarctica are smaller, and of similar magnitude to those of coastal stations in ERA5 and JRA-55 (Tables S1–S3). It is clear from Fig. 2a that ERA5 is the best reanalysis in terms of RMSE, with values being significantly lower than both 20CRv3 and JRA-55 at $p < 0.05$ and $p < 0.01$ in the EARLY and LATE periods, respectively. There is no significant difference between the RMSE values of the other two reanalyses in either period (Table 2). Moreover, the similarity of the two ERA5 distributions in Fig. 2a indicates it is also the most homogeneous reanalysis in terms of this validation statistic.

The distribution of the ratio of variance values is shown in Fig. 2b. Although ERA5 has the greatest range of values, it is the other two reanalyses that demonstrate significant improvements (values closer to 1) from the EARLY to LATE periods (populations of ratio of variance values are statistically different at $p < 0.10$ for 20CRv3 and $p < 0.01$ for JRA-55). In the EARLY period the vast majority of the ratio values for these two reanalyses are less than one, indicating that they exhibit too little inter-annual Antarctic SAT variability at this time. In ERA5, the ratio of variance varies

Table 2 Statistical significance of differences in the RMSE populations from the three reanalyses

	EARLY Period (1958–1978)		LATE Period (1979–2015)	
	20CRv3	JRA-55	20CRv3	JRA-55
ERA5	$p < 0.05$	$p < 0.05$	$p < 0.01$	$p < 0.01$
20CRv3	–		–	

Table 3 Statistical significance of differences in the ratio of variance populations from the three reanalyses

	EARLY Period (1958–1978)		LATE Period (1979–2015)	
	20CRv3	JRA-55	20CRv3	JRA-55
ERA5	$p < 0.10$	$p < 0.01$		
20CRv3	–	$p < 0.05$	–	

Table 4 Statistical significance of differences in the correlation coefficient populations from the three reanalyses

	EARLY Period (1958–1978)		LATE Period (1979–2015)	
	20CRv3	JRA-55	20CRv3	JRA-55
ERA5	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.01$
20CRv3	–	$p < 0.01$	–	

from 0.15 at Vernadsky to 4.71 at Novolazarevskaya in the EARLY period. Although there is no significant change in the distribution of ratio values in the LATE period, it is clear that the range of values has decreased considerably. Table 3 reveals that in the EARLY period ERA5 has the best ratio of variance values, being significantly higher than 20CRv3 ($p < 0.10$) and JRA-55 ($p < 0.01$), while the former is also significantly better than the latter ($p < 0.05$). However, in the LATE period there are no significant differences in the ratio of variance values between any combination of the three reanalyses. Similar to the RMSE values, ERA5 is the most homogeneous of the three reanalyses in portraying the magnitude of interannual SAT variability across both periods.

The final validation statistic is the correlation coefficient r , which is shown in Fig. 2c, with the significance of any differences between the reanalyses given in Table 4. In contrast to the previous two validation statistics, all three reanalyses demonstrate a clear improvement ($p < 0.01$) in the correlation of annual SAT from the EARLY to LATE periods. In terms of the median values of r , ERA5 changes from 0.74 to 0.93, 20CRv3 from 0.67 to 0.80 and JRA-55 from 0.53 to 0.80. Figure 2c indicates that all reanalyses/periods have at least one correlation coefficient less than 0.40. These

values are generally for either Amundsen-Scott or Novolazarevskaya. Marshall (2003) demonstrated that an earlier ECMWF reanalysis, ERA-40, also had problems reproducing the climate at Novolazarevskaya both before and after 1979: in ERA5 the RMSE and correlation values are actually worse in the LATE period (c.f., Table S1). The exception is for JRA-55 during the EARLY period, when there are four stations where $r < 0.40$ (Mawson, Mirny, Vostok and Vernadsky), encompassing the Antarctic Plateau, coastal East Antarctica and the western Peninsula. Unsurprisingly, many of these stations also have high RMSE and low variance ratio values (c.f., Table S3). The lowest r of 0.02 is at Vostok, which improves to 0.75 in the LATE period despite a coincident increase in an already high RMSE. Regarding differences between the three reanalyses, ERA5 has significantly better correlation coefficients than the two other reanalyses during both periods while 20CRv3 is superior to JRA-55 in the EARLY period only (Table 5). We note again that, in contrast to ERA5 and JRA-55, SAT observations are not assimilated into 20CRv3.

Based on the locations of the 14 meteorological stations it is apparent that ERA5 has the most skill of the three reanalyses in reproducing Antarctic annual-mean SAT variability during both periods. All three validation statistics are statistically significantly better than for 20CRv3 and JRA-55 in the EARLY period while in the LATE period both the RMSE and correlation coefficient values are superior. Moreover, the populations of RMSE and ratio of variance values from the EARLY and LATE periods in ERA5 are not distinct, which is not the case for the latter statistic in either 20CRv3 or JRA-55. Therefore, ERA5 also appears to demonstrate the greatest homogeneity in SAT before and after the advent of the satellite sounder data.

3.1.2 Analysis of seasonal data in ERA5

Summary validation statistics for each season are given in Tables S4–S7 and presented as box-whisker plots in Fig. 3. There are no significant differences in the RMSE and variance ratio values between the four seasons or between the EARLY and LATE periods. Interestingly, the highest RMSE values were larger (worse) in the LATE period, although the median was lower in the LATE period in all seasons

Table 5 SAM validation statistics from the three reanalyses and the Fogt reconstruction for the EARLY period (left) and LATE period (right)

Reanalysis/ reconstruction	EARLY Period (1958–1978)					LATE Period (1979–2015)				
	n	bias	RMSE	var ratio	r	n	bias	RMSE	var ratio	r
ERA5	21	– 0.08	0.37	1.36	0.94	37	0.00	0.35	0.91	0.97
20CRv3	21	– 0.18	0.43	0.95	0.89	37	– 0.08	0.51	0.99	0.94
JRA-55	21	– 2.22	1.61	1.82	0.67	37	– 0.07	0.31	0.91	0.98
Fogt	21	1.65	1.19	0.89	0.87	27	0.41	0.81	0.59	0.84

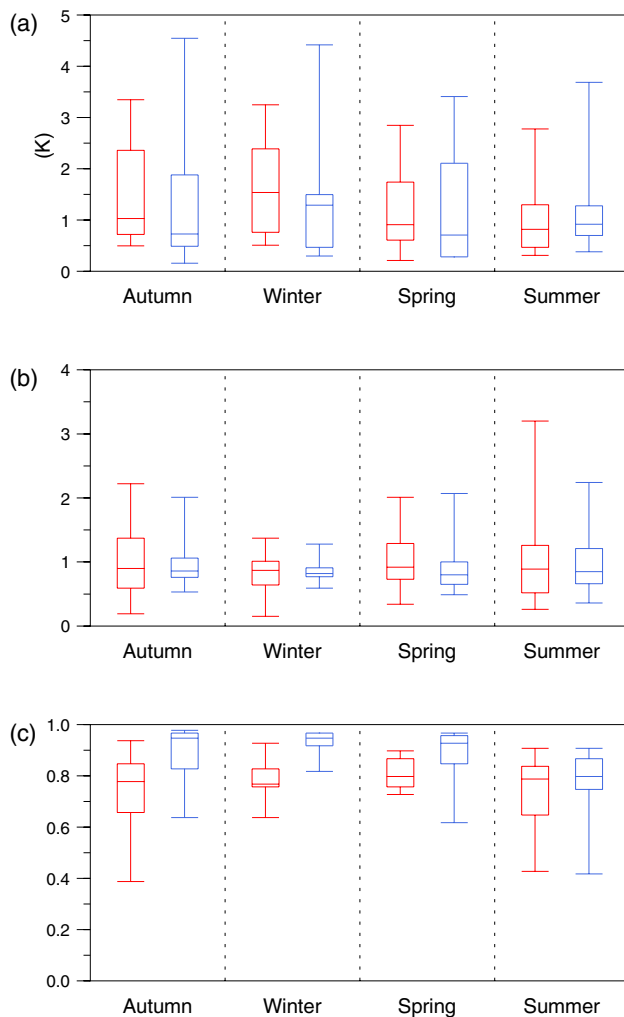


Fig. 3 As Fig. 2 but for the four seasons in ERA5

except summer. While the relatively high RMSE values for Novolazarevskaya are similar across all four seasons, at Amundsen-Scott there is a major improvement in summer: 0.93 °C compared to 4.61 °C in autumn (c.f., Tables S4 and S7). This suggests that the RMSE value here is not simply a function of incorrect model orography but also a response to an inability to reproduce local processes. The improvement in summer is indicative of the model being unable to correctly represent surface temperature inversions, which are less frequent and weaker in this season (e.g., Hudson and Brandt 2005).

Figure 3c shows that the ERA5 correlation coefficient statistics for autumn, winter and spring are similar to the annual data in that there is a statistically significant improvement between the EARLY and LATE periods. However, in austral summer there is only a slight improvement. Consequently, in the LATE period the values of r in the other three seasons are all significantly better than in summer ($p < 0.01$). Thus, based on the three validation statistics, the ERA5 SAT data

can be considered more homogeneous across the EARLY and LATE periods in summer than at other times of year but at a cost of being less accurate in the LATE period. We note that these seasonal differences in temporal improvements and homogeneity match those from validation studies of older reanalyses (e.g., Bromwich and Fogt 2004).

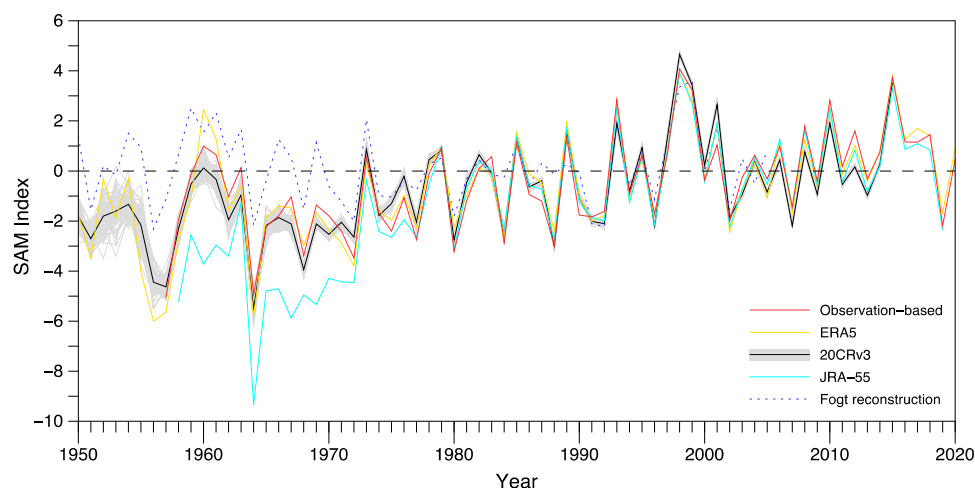
3.2 Analysis of SAM indices

The validation statistics for the different SAM indices versus the observation-based index of Marshall (2003) are given in Table 5 and the indices plotted in Fig. 4. It is apparent that the SAM indices for ERA5 and 20CRv3 ensemble mean both produce very good facsimiles of the observed data and are relatively homogeneous across both the EARLY and LATE periods. The bias, RMSE and correlation of SAM values are marginally better in ERA5 while the interannual variability is more closely aligned to observations in 20CRv3. In contrast, the JRA-55 SAM Index shows a marked improvement in all the validation statistics from the EARLY to LATE periods such that it is as skilled as the other two reanalyses in the latter. The Fogt reconstruction demonstrates an improvement in bias and RMSE from the EARLY to LATE periods whereas the variance ratio is poorer. We note that the reconstruction data end in 2005, which will impact a direct comparison with the other reanalyses in the LATE period, in addition to the different methodology used to produce it.

The bias values and Fig. 4 reveal marked negative and positive biases in SAM values in the EARLY period in JRA-55 and the Fogt reconstruction, respectively. Comparing the separate normalised SLP anomalies at 40°S and 65°S indicates that errors in the latter are responsible for the negative bias in JRA-55. For example, the strongly negative polarity of the observation-based SAM in 1964 (− 4.92) comprises normalised anomalies of − 3.18 and 1.75 at 40°S and 65°S, respectively, in the observations. For JRA-55, the equivalent values are − 9.34, − 3.50 and 5.84, establishing that SLP at SH high-latitudes is much too high. This positive SLP bias is similar to that observed in older reanalyses (e.g., Marshall 2003; Bromwich and Fogt 2004).

The spread in SAM values between the 80 20CRv3 ensemble members is expressed by the width of the grey region around the black line of the ensemble mean in Fig. 4. There is clearly more variability in the EARLY period, with a maximum standard deviation between members of 0.65 occurring in 1955. As expected, the spread declines following the IGY of 1957/58 when many new pressure observations from SH high-latitudes became available, (standard deviation of 0.29 in 1957). Subsequently, there is a further gradual reduction, with 1983 the last year when the standard deviation is greater than 0.10. Values in the LATE period are as low as 0.05.

Fig. 4 Comparison of annual SAM indices from observations, three reanalyses and a statistical reconstruction. The indices are normalised over 1981–2010. For 20CRv3 each of the 80 ensemble members is shown as a grey line with the ensemble mean in black



3.3 Changes in SAM-SAT relationship

In Fig. 5 we reproduce Fig. 1b for ERA5, 20CRv3 and JRA-55 as Figs. 5a, b and c, respectively. There is evidence of a switch in the decadal SAM-SAT relationship between 1970–79 and 1971–80 in all three reanalyses, although it is less distinct than in the observations, particularly for station locations in East Antarctica. The SAM-SAT relationships in the reanalyses during the LATE period generally match the observations well, not unexpectedly given the accuracy of both Antarctic SAT and the SAM indices at this time, as described in Sects. 3.1 and 3.2. As an illustration, the significantly less positive (and sometimes negative) SAM-SAT relationship at some Peninsula stations in recent decades (c.f., Fig. 1b) is present in ERA5 and JRA-55 (Figs. 5a, c) (20CRv3 ends too early to show this).

ERA5 (Fig. 5a) is the reanalysis that most closely reproduces the spatial and temporal decadal Antarctic SAM-SAT variability in the observations during the EARLY period (c.f., Fig. 1b). For example, the periods of positive correlation at Amundsen-Scott and Novolazarevskaya are clearly apparent. However, ERA5 has fewer decades with a significantly more positive SAM-SAT correlation from Syowa to Davis, while having a greater number of positive decadal correlations further east, from Dumont d'Urville to Byrd in West Antarctica. Moreover, the majority of these SAM-SAT correlations with opposite sign to those in the LATE period occur several decades before 1970–79. Another difference from the observations is that the shift to strong positive decadal correlations at Vernadsky and Esperanza on the Peninsula begins a year earlier (1970–79). Furthermore, in ERA5 all three Peninsula stations also have a positive SAM-SAT relationship for the decades from 1955–64 to 1962–71, similar to the LATE period. However, the reanalysis suggests that for 1950–59 to 1956–65, which are the decades encompassing the period before the observation-based SAM

begins, there is generally a negative SAM-SAT relationship across much of Antarctica.

In the 20CRv3 reanalysis (Fig. 5b) the 1970–79 to 1971–80 change in the SAM-SAT relationship in East Antarctica is more distinct than in ERA5, being especially marked at Scott Base. Decades with positive SAM-SAT correlations at Amundsen-Scott are present at the beginning of the analysis period (1950–59) in 20CRv3 but, strangely, switch to being predominantly negative during the decades when the observations (and ERA5) indicate a positive relationship. Moreover, at Syowa station there is a positive SAM-SAT relationship throughout the EARLY period. Further east, the SAM-SAT relationship at Dumont d'Urville in the EARLY period in 20CRv3 matches observations in that there are no decades with a positive correlation. While the switch between the EARLY and LATE periods is distinct at Orcadas, this is not the case for the other two Peninsula stations: for example, at Vernadsky there is a significant positive SAM-SAT correlation in 1970–79 (c.f., Fig. 5b). In the decades before the observation-based SAM, the spatial and temporal distributions of Antarctic SAM-SAT correlation values are quite different in 20CRv3 and ERA5, with the significantly low correlation values for the Peninsula stations being the principal feature common to the two reanalyses (cf. Fig. 5a, b).

JRA-55 is the reanalysis least able to reproduce the observed SAM-SAT relationships (Fig. 5c). The greatest difference is that there is no change in the sign of the SAM-SAT relationship of the three Peninsula stations between the EARLY and LATE periods. There are weaker positive correlations before 1971–80 but all three stations have an unbroken positive SAM-SAT relationship back to 1958–65, which is not seen in the other reanalyses. In addition, most of East Antarctica has continuous negative SAM-SAT correlations, including several decades with statistically significant negative correlations at Amundsen-Scott and Mirny stations, some of which are actually significantly positive

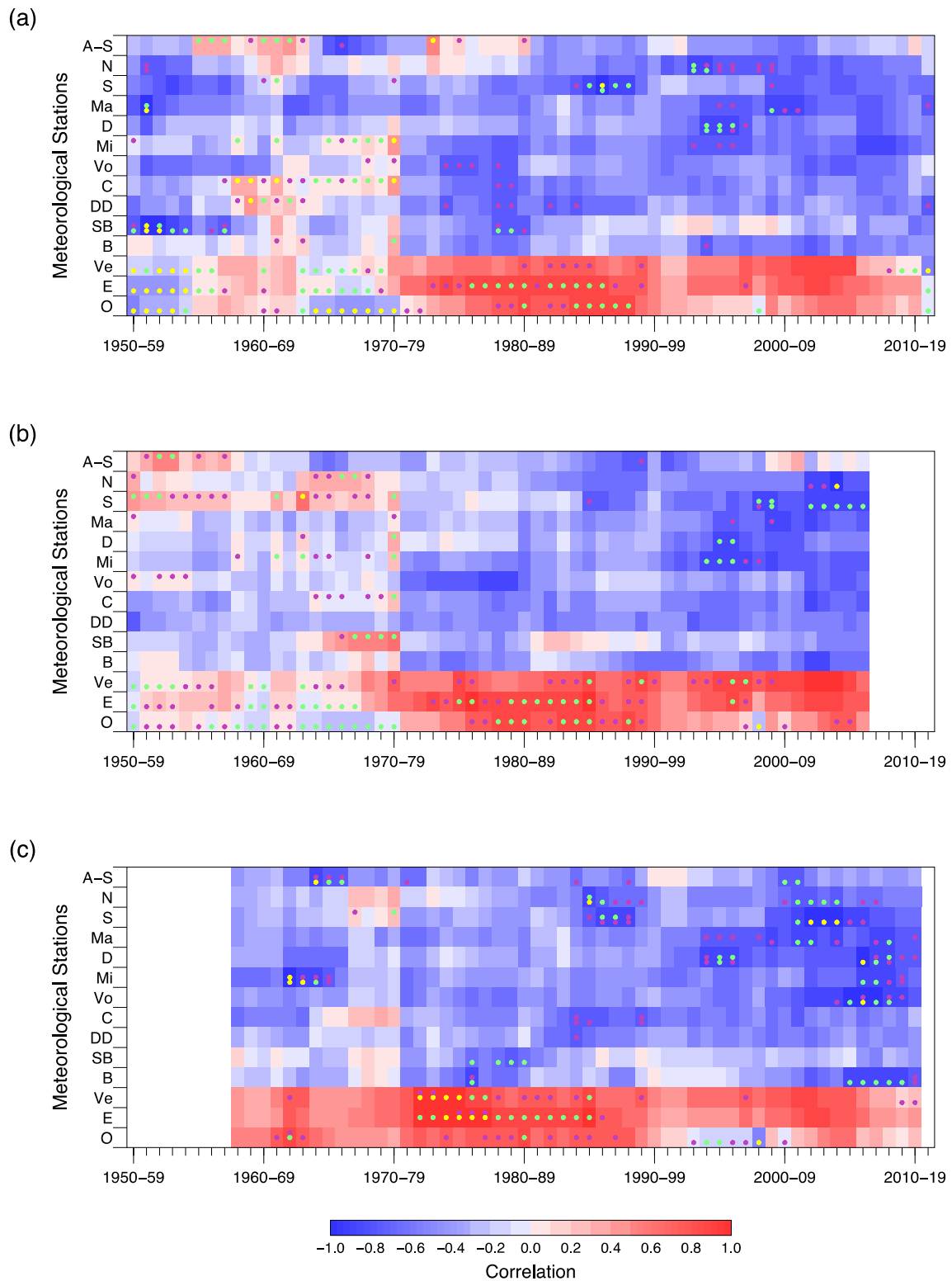


Fig. 5 As Fig. 1b for **a** ERA5, **b** 20CRv3, **c** JRA-55. The SAM indices were computed from the individual reanalyses using an identical methodology to the observation-based SAM index

in the observations, ERA5 and 20CRv3. Similar to the latter, there is a period of a few decades at the end of the EARLY period with positive correlations in the longitudes from Novolazarevskaya to Syowa and Casey east to Byrd (excluding Dumont d'Urville). The change to the negative SAM-SAT relationships of the LATE period in these regions of East Antarctica most clearly defines the switch between the two periods in JRA-55.

4 Changes in the SAM structure between the EARLY and LATE periods

The validation exercise demonstrates that the reanalyses have some skill in reproducing the marked change in decadal Antarctic SAM-SAT relationships between 1970–79 and 1971–80 seen in the observations, particularly ERA5 and 20CRv3. Previous analyses of variations in the SAM-SAT relationships within the LATE period, have pointed to changes in the SAM structure as being responsible (e.g., Marshall et al. 2011; 2013; Wachter et al. 2020) and next we investigate how this varies between the EARLY and LATE periods and whether it can explain some of the differences between the reanalyses. First, we undertake a qualitative analysis comparing the changes in the Antarctic SAM-SAT relationship observed in Fig. 5 with those in the SAM structure, as defined using the spatial variability in the SAM-SLP correlation field. The SAM-SLP correlation for each available decade for the three reanalyses across the EARLY and LATE periods are provided in Figs. 6, 7, 8.

In the first few decades of ERA5 and 20CRv3, from 1950–59 to 1954–63, there is a northward projection of negative correlations (hereinafter NPNC) from the Weddell Sea to the western South Atlantic ($\sim 45^\circ\text{W}$), extending north beyond 30°S in 20CRv3 (Figs. 6 and 7). There will be associated anomalous northward advection of heat and moisture on the western side of a region of NPNC and, similarly, anomalous southward flow to the east. The opposite circulation structure will occur for an area of southward projecting negative correlations (hereinafter SPNC). Previous work has demonstrated that the sign of the SAM-SAT relationship in Antarctica is often highly dependent on the local meridional wind anomaly associated with the SAM (e.g., Marshall and Thompson 2016; Wachter et al. 2020). A northerly wind will advect warm maritime air towards Antarctica while a southerly wind will draw cold air from the continent. There is also a smaller NPNC at $\sim 150^\circ\text{W}$ extending to about 50°S , located slightly further west than in the LATE period. However, around East Antarctica the structure is broadly annular and so the SAM-SAT relationships in ERA5 are similar to the LATE period (Fig. 5a).

In ERA5 the decadal SAM structure changes to having a primary NPNC at about 30°E from 1955 to 64, which

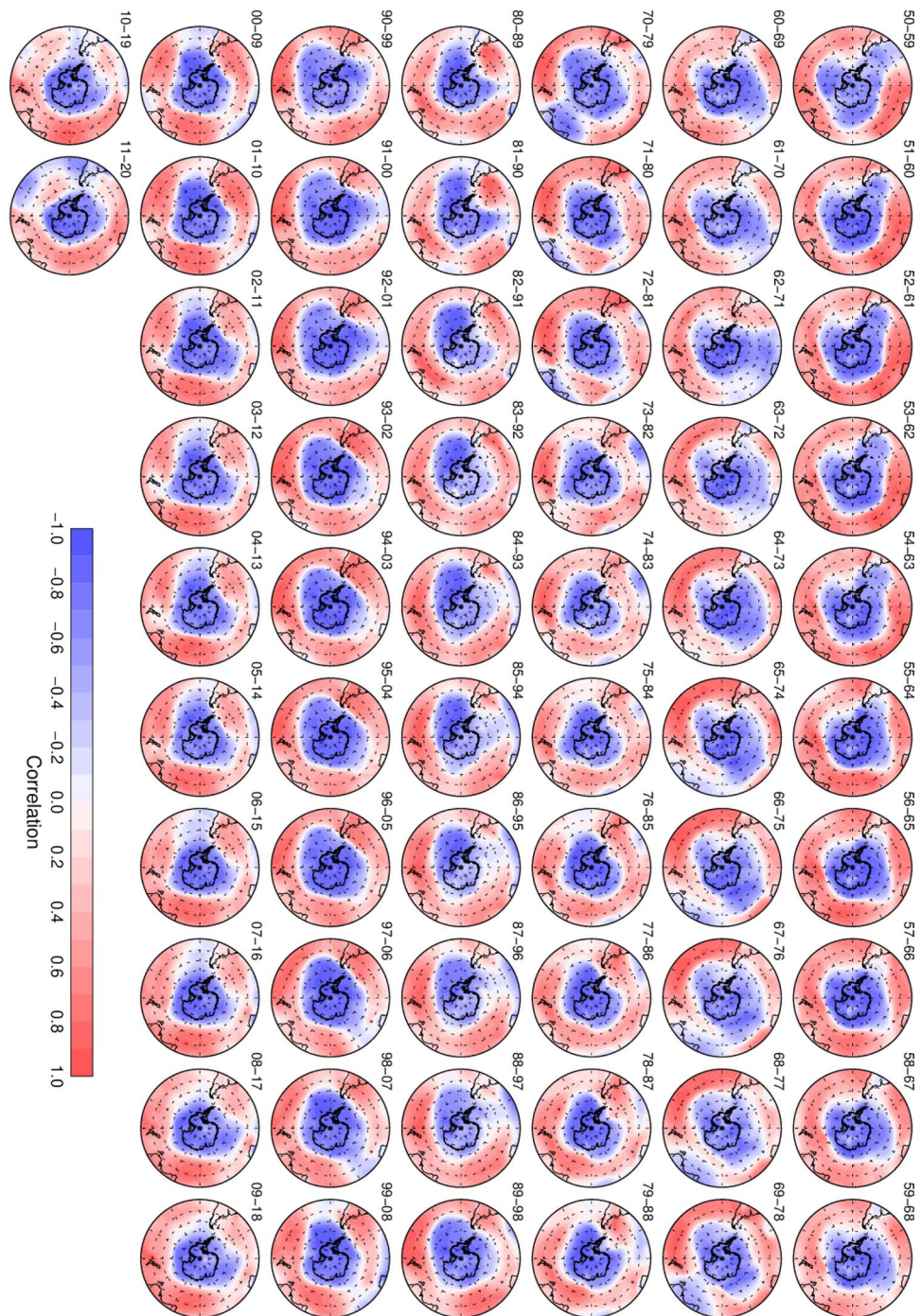
subsequently switches to being close to the Greenwich Meridian, particularly in 1962–71 and 1963–72 (Fig. 6). During these two modes of SAM structure variability there is a corresponding positive SAM-SAT relationship at Amundsen-Scott (Fig. 5a). In 20CRv3 there is also a switch to the primary NPNC near to the Greenwich Meridian from 1958–67 to 1963–72 (Fig. 7), although here it marks the end of the earlier positive decadal correlations between the SAM and Amundsen-Scott SAT. Throughout these decades the SAM structure in JRA-55 is similar to the LATE period (Fig. 8) and, consequently, the SAM-SAT relationships in these decades in Fig. 5c also resemble those in the LATE period.

During the remainder of the EARLY period, from 1964–73 to 1970–79, in ERA5 there is a broad NPNC to north of 50°S across the South Atlantic and eastward into the Indian Ocean (Fig. 6). In 20CRv3 a similar SAM structure occurs but extends further east with negative correlation values of smaller magnitude (Fig. 7). In JRA-55, there is a contemporaneous NPNC towards Australia at $\sim 120^\circ\text{E}$ (Figs. 8), not observed in the two other reanalyses, which explains the positive or weaker negative SAM-SAT relationships at Casey and Dumont d'Urville from 1964–73 to 1970–79 (Fig. 5c). Throughout the EARLY period, ERA5 and 20CRv3 have a stronger and narrower NPNC at $\sim 150^\circ\text{W}$ than JRA-55 (Figs. 6, 7, 8).

The transition from EARLY to LATE periods in the observations is readily apparent from 1970–79 to 1971–80 but, as previously mentioned, is less distinct in the reanalyses. In ERA5, the largest change in the SAM-SAT relationships at two Peninsula stations (Vernadsky and Esperanza) is a year earlier, between 1969–78 and 1970–79 (Fig. 5a) and occurs even earlier in 20CRv3. In ERA5, this appears to be due to the formation of the NPNC in the Amundsen Sea in 1969–78, similar to the LATE period climatology (Fig. 6). In 20CRv3 the timing of the switch from a negative to positive SAM-SAT relationship on the Peninsula is less apparent from changes in the SAM structure (Fig. 7). In JRA-55 there is no marked change in the SAM structure in the South Pacific between the EARLY and LATE periods (Fig. 8). Orcadas, situated to the north-east of the Peninsula (c.f. Figure 1a), appears to be less influenced by the changes in the SAM structure over the Amundsen-Sea and the temporal decadal SAM-SAT relationship there in ERA5 and 20CRv3 better resembles that in the observations.

While there are various decades when the SAM-SAT relationship is positive at some East Antarctic stations in the EARLY period in all three reanalyses, 1970–79 is the decade when this relationship is most consistent, especially in ERA5 (Fig. 5a). Figure 6a reveals that this is due to an NPNC at $\sim 135^\circ\text{E}$, similar to that seen in earlier decades in JRA-55. In both ERA5 and 20CRv3 this feature is most evident in this decade (Figs. 6 and 7), while in JRA-55,

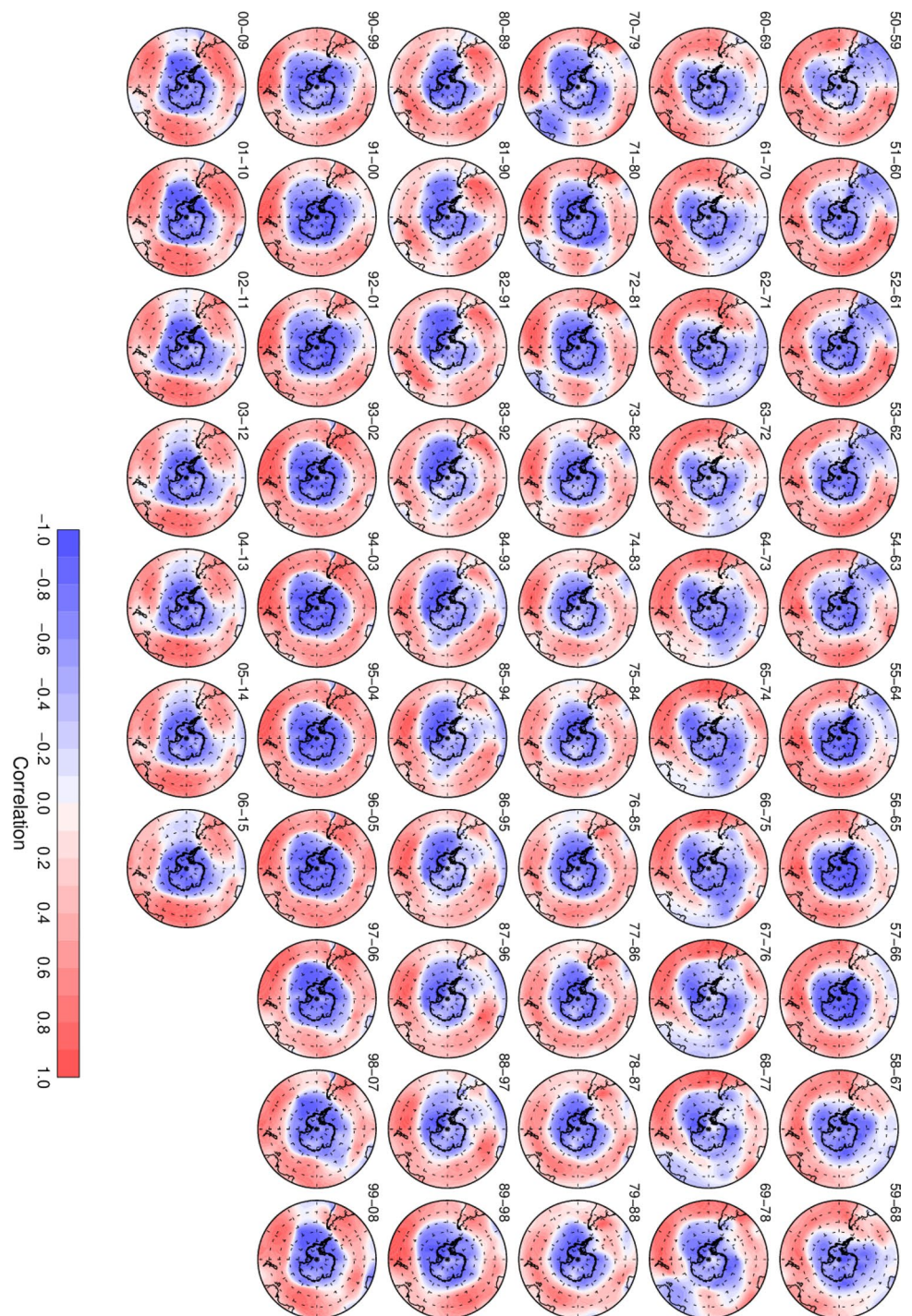
Fig. 6 Time series of decadal correlation between the annual SAM index and SLP for ERA5



1970–79 represents the last decade when an NPNC at this longitude is especially marked (Fig. 8).

In the LATE period all three reanalyses generally do well at reproducing the broadly consistent SAM-SAT relationships at the Antarctic stations. Nevertheless, Fig. 6 does reveal that there are marked changes in the annual SAM-structure within the LATE period, consistent with the findings of Wachter et al. (2020). In particular, all the

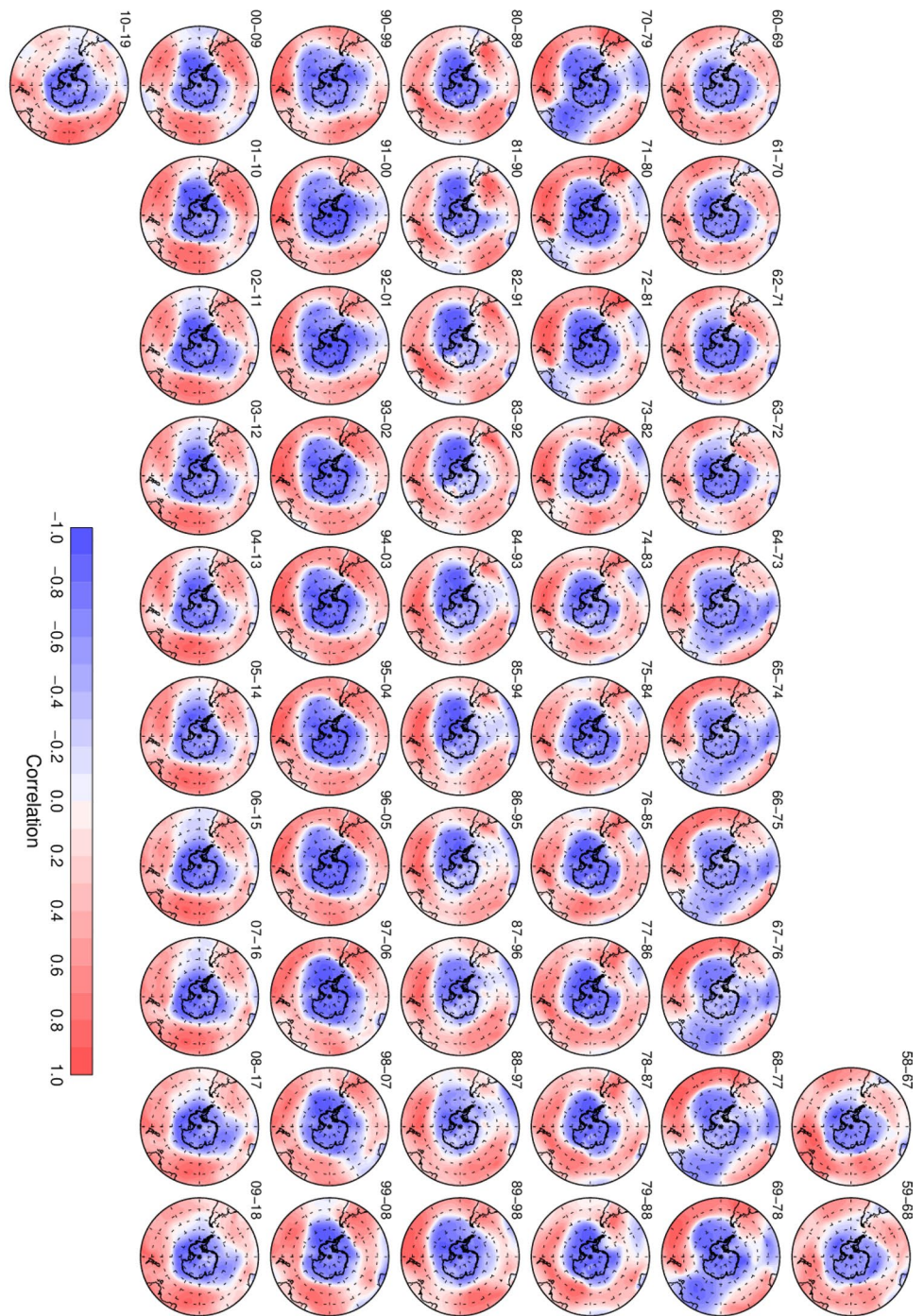
reanalyses have decades when the NPNC in the Amundsen Sea extends further north than the climatology: for example, in ERA5 this is 1980–89 to 1988–97 and 2000–09 to 2007–16. There are also decades when the NPNC near the Greenwich Meridian stretches further north into the South Atlantic, some of which occur contemporaneously with the extended negative correlations in the South Pacific. Such periods, from 1984–93 to 1988–97 for example, are

Fig. 7 As Fig. 6a for 20CRv3

broadly consistent among the three reanalyses (Figs. 6, 7, 8). However, the associated changes in SAM structure have relatively little impact on the decadal annual SAM-SAT relationships across Antarctica (Fig. 5). The only clear differences in SAM-SAT relationships between the reanalyses and the observations in the LATE period are three decades of positive correlation at Amundsen-Scott (1999–2008 to 2001–2010) in 20CRv3 (Fig. 5b) and the decades of

negative correlation at Orcadas that predominate between 1993–2002 and 2000–2009 in JRA-55 (Fig. 5c).

To quantify the variability in the SAM structure shown in Fig. 6, 7, 8, for each decade we extract the zonal SAM-SLP correlation anomalies at 55°S per degree longitude to use as a summary diagnostic. This latitude typically lies between positive (negative) SAM-SLP correlations to the north (south) (cf. Fig. 1a). Therefore, any anomalous changes in

Fig. 8 As Fig. 6a for JRA-55

SAM structure, as represented by changes in the location of NPNCs and SPNCs, will be manifested at this latitude. The inter-quartile ranges of this diagnostic for the EARLY and LATE periods are shown in Fig. 9, together with the longitudes where the population of SAM-SLP correlations is significantly different between the two periods.

For ERA5, the mean decadal SAM structure in the EARLY period reveals an extensive region of NPNCs from

the South Atlantic to the Indian Ocean (70°W–120°E), that is when the upper quartile contour is located north of 55°S in Fig. 9a. Conversely, SPNCs, when the contour defining the lower quartile is south of 55°S, are apparent in the remainder of the hemisphere, apart from a local NPNC centred at ~140°W (Fig. 9a). In the LATE period ERA5 has the characteristic SAM structure, as already illustrated in Fig. 1a: the principal NPNC is located over the Amundsen Sea that,

together with regions of smaller magnitude NPNCs at $\sim 40^\circ\text{E}$ and 150°E , gives a much stronger zonal wave 3 pattern to the SAM-SLP correlations than in the EARLY period. The greatest variability in SAM structure in the LATE period in ERA5 is observed in the Amundsen-Bellinghousen Seas (hereinafter ABS), and the other two regions of local NPNC maxima, which corresponds to the findings of Wachter et al. (2020: their Fig. 4c). Figure 9a indicates statistically significant differences in the SAM structure between the two periods in four regions: most notably in the ABS, but also the western Weddell Sea and two regions of the southern Indian Ocean.

The SAM structure in 20CRv3 data in the EARLY period is broadly similar to ERA5. The principal differences are that (i) the broad region of NPNCs is reduced in size, extending to only 90°E and (ii) the NPNC at 150°W has a greater magnitude (Fig. 9b). As the SAM structure in the LATE period is essentially identical to ERA5, the differences in the EARLY period mean that there are fewer and slightly different regions where there is a significant difference between the two periods in 20CRv3. As mentioned previously with regard to the Antarctic SAM-SAT relationships, JRA-55 is markedly different to the other two reanalyses in the EARLY period and, unsurprisingly, this is

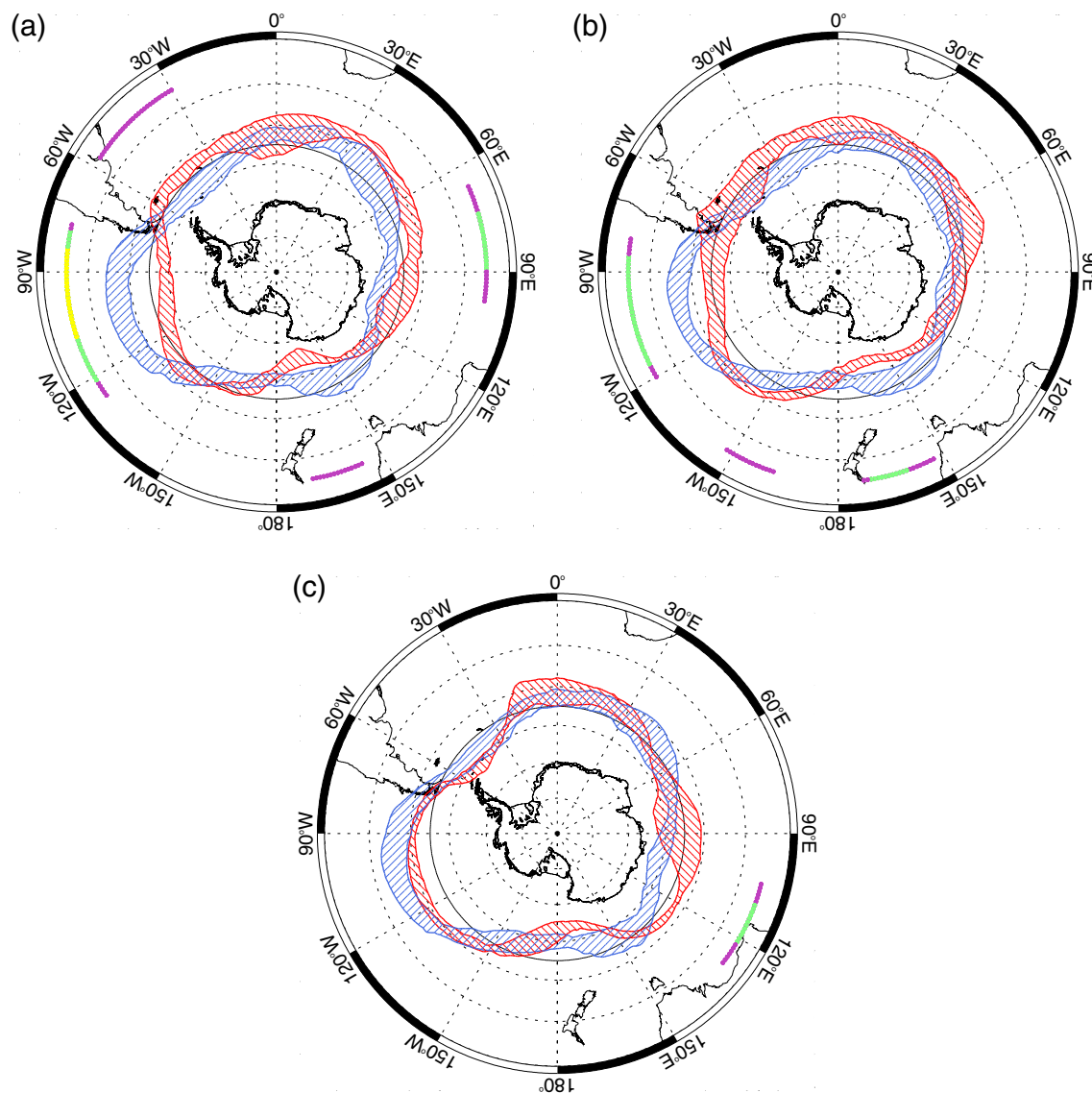


Fig. 9 Mean decadal SAM structure based on SAM-SLP correlations at 55°S for the EARLY period (red) and LATE period (blue) for **a** ERA5, **b** 20CRv3 and **c** JRA-55. The shading represents the range in SAM-SLP correlations between the upper and lower quartiles. Positive (negative) values are plotted proportionately south (north) of

55°S . The dots at 35°S indicate the presence of a statistically significant difference in the SAM-SLP correlations between the two periods at that longitude: purple, green and yellow dots represent $p < 0.10$, $p < 0.05$ and $p < 0.01$, respectively

also true of the SAM structure. Figure 9c indicates that there are three well-defined NPNCs during this period, located at $\sim 90^\circ\text{E}$, 120°W and 20°W , giving a distinct wave-number 3 structure. The second of these is relatively close to the major LATE-period NPNC in the ABS, although of smaller magnitude and centred slightly further west. This, together with the smaller sample size, means there are no significant differences in SAM structure in the ABS or Weddell Sea, in contrast to the other two reanalyses. The only region where there is a significant difference between the EARLY and LATE periods in JRA-55 is in the south-east Indian Ocean, centred at $\sim 110^\circ\text{E}$, which is distinct from the regions of significant difference in either ERA5 or JRA-55.

By differentiating the SAM-SLP structure with respect to longitude—that is, calculating the local change in the SAM-SLP correlation at 55°S per degree of longitude—we approximate the mean meridional wind direction associated with SAM+ in the EARLY and LATE periods (Fig. 10). As mentioned previously, the sign of the SAM-SAT relationship in some parts of Antarctica is primarily determined by the regional meridional wind anomaly associated with the SAM (e.g., Marshall and Thompson 2016; Wachter et al. 2020).

In ERA5 there are several sectors where the meridional wind direction associated with one polarity of the SAM reversed between the EARLY and LATE periods (Fig. 10a). In the Peninsula region SAM+ was linked to southerly winds in the EARLY period and northerlies in the LATE period, providing a simple explanation for the reversal in the SAM-SAT relationship seen in Fig. 5a, which closely mirrors the observations (Fig. 1b). In the EARLY period, the combination of southerlies (weak northerlies) in the Peninsula (eastern Weddell Sea) regions in ERA5 is indicative of anomalously cyclonic flow in the Weddell Sea associated with SAM+. Clem et al. (2020) established that this circulation pattern, which they linked to positive SSTs in the western tropical Pacific combined with SAM+, is responsible for warmer SAT at Amundsen-Scott and hence it may explain the positive SAM-SAT correlations at the Pole during substantial parts of the EARLY period (Fig. 5a). From 40 to 100°E , there is little change in mean meridional wind direction between the EARLY and LATE periods and thus there is only minor variability in the SAM-SAT correlation at stations within this sector (Syowa east to Davis). Further east, there are additional reversals in the meridional wind direction, such as the change from northerlies to southerlies at 100 – 160°W , which is likely a contributing factor to the decades with a positive SAM-SAT relationship at Byrd in West Antarctica in the EARLY period (Fig. 5a).

The changes in meridional wind direction in 20CRv3 closely match those in ERA5 (c.f., Figs. 10a, b). The greatest differences occur in the Weddell Sea region. For example, in 20CRv3 there is a switch from northerlies to weak southerlies at $\sim 40^\circ\text{W}$ between the EARLY and LATE periods

that is not seen in ERA5. This explains the more frequent positive SAM-SAT correlations at Novolazarevskaya and Syowa in the early period in 20CRv3 (c.f., Figs. 5a, b). In JRA-55 there are fewer sectors where there is a clear difference in the meridional wind direction between the two periods examined than in the other two reanalyses and thus the EARLY and LATE period SAM-SAT relationships in JRA-55 also demonstrate greater temporal homogeneity. The most prominent difference is centred at $\sim 90^\circ\text{E}$ (Fig. 10c), with the northerlies in the EARLY period likely responsible for the decades of positive SAM-SAT correlations at Casey in Fig. 5c. The situation in the Weddell Sea sector is more complex during the EARLY period. We note that the especially strong northerly wind component at $\sim 20^\circ\text{W}$ associated with SAM+ in the EARLY period does not result in a positive SAM-SAT relationship at Amundsen-Scott in JRA-55.

To quantify longitudinal variations in the SAM structure, we decompose the zonal SAM-SLP correlation anomalies into the first four zonal wave-numbers for each decade in the EARLY and LATE periods using standard Fourier analysis techniques. The mean amplitude, phase and variance explained for each wave-number from the two periods are provided in Tables 6, 7, 8 and illustrated in Fig. 11.

Statistically significant differences between the EARLY and LATE periods in the wave structure of the SAM-SLP correlation anomalies in ERA5 occur in the amplitude ($p < 0.10$) and variance explained ($p < 0.05$) of Wave 1 and the phase ($p < 0.01$) of Wave 3 (Table 6). In the EARLY period Wave 1 is dominant, explaining more than half of the variance, indicating that the SAM structure is much more annular at that time. The mean phase of Wave 3 changes from 32° in the EARLY period to 77° in the LATE period: the difference of 45° is relatively close to being a complete phase reversal (60°). Although not significant, we also note the increases in the mean amplitude and variance explained by Wave 3 from the EARLY to LATE period, making it the dominant wave-number during the latter. Figure 11a indicates that all zonal wave-numbers are contributing to the LATE period NPNC in the ABS ($\sim 110^\circ\text{W}$), as they are all negative at this longitude. Similarly, Fig. 11a demonstrates that the position of the adjacent SPNC, which is located over the Weddell Sea in the LATE period ($\sim 45^\circ\text{W}$), corresponds closely with positive nodes in Waves 2–3. However, given the significance in the change in the phase of Wave 3 between the EARLY and LATE periods and its amplitude in the latter, temporal variability in this wave-number is the primary driver behind the switch from southerlies to northerlies associated with SAM+ in the Antarctic Peninsula and thus the reversal of the regional SAM-SAT relationship.

The differences in the zonal wave-numbers between the EARLY and LATE periods in 20CRv3 have some similarities with ERA5. In particular, the difference in the phase of Wave 3 is also significant at $p < 0.01$ (Table 7). The variance

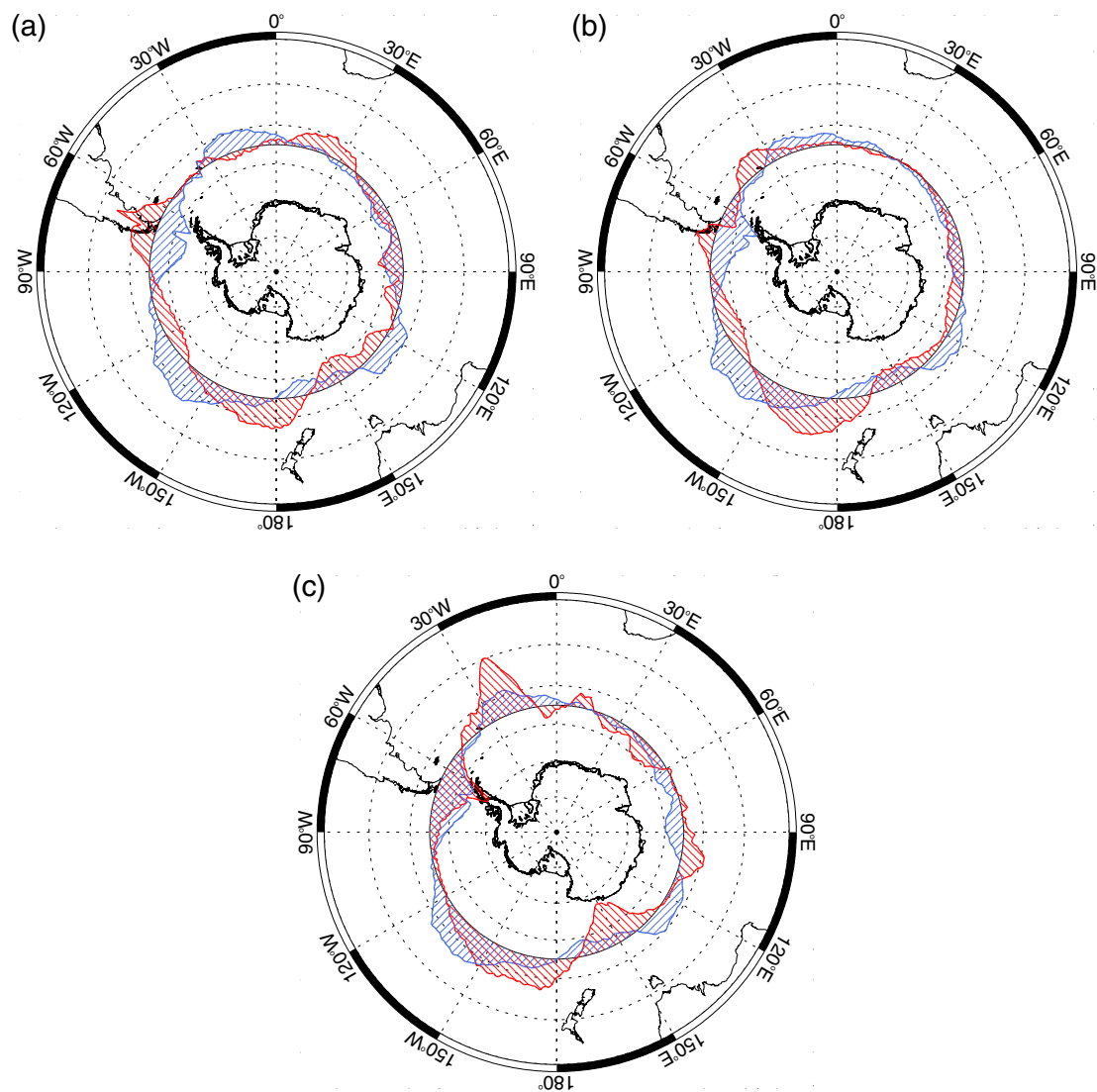


Fig. 10 Longitudinal variability in the strength of the meridional wind component associated with SAM+ during the EARLY period (red) and LATE period (blue) for **a** ERA5, **b** 20CRv3 and **c** JRA-55.

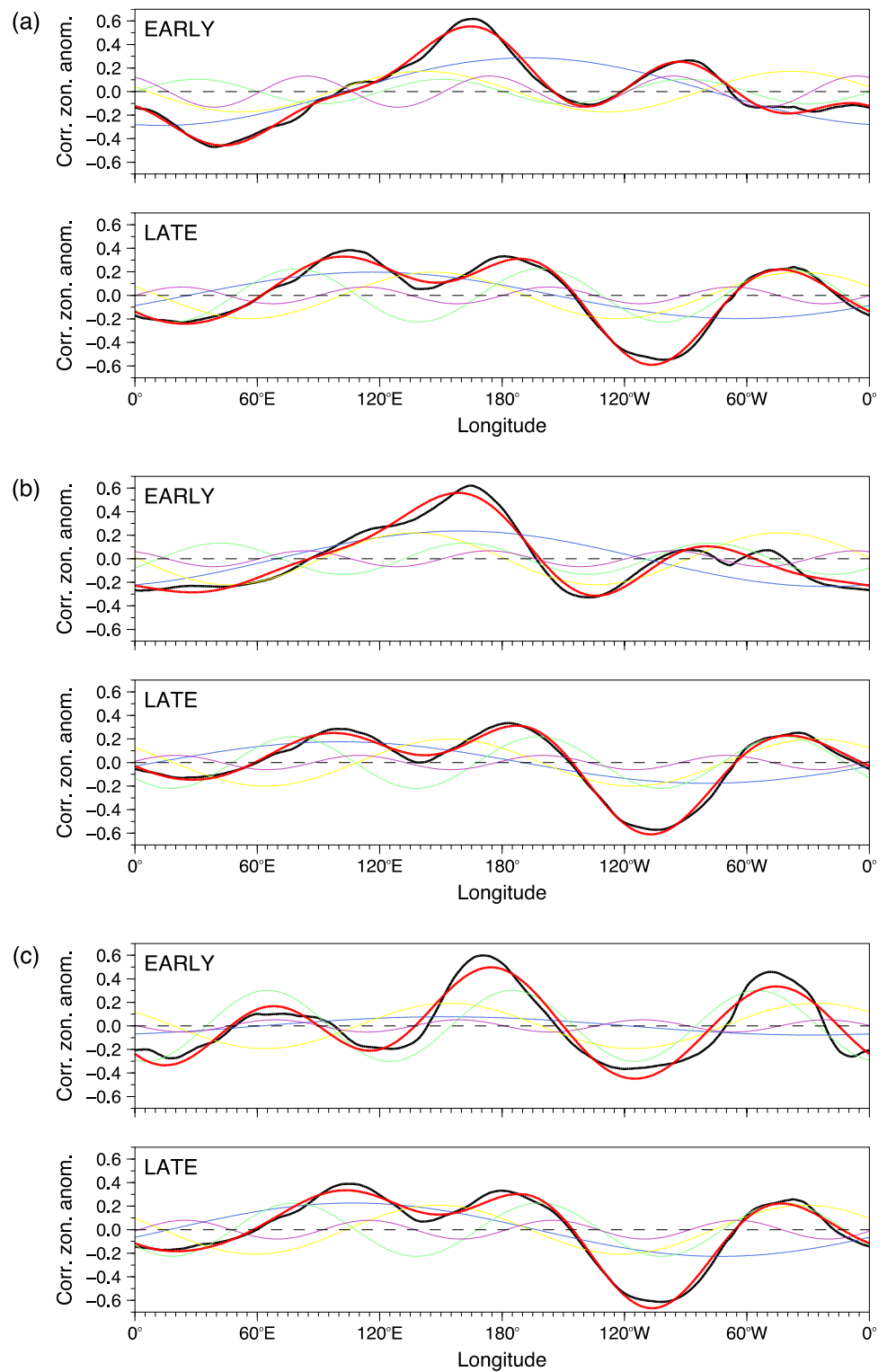
The relative magnitude of the northerly (southerly) winds are plotted proportionately south (north) of 55°S

explained by Wave 1 is significantly different between the two periods but the amplitude is not, with the decrease from the EARLY to LATE period being less than observed in ERA5. Given the relatively small amplitudes of the other wave-numbers, the divergence in amplitude and phase of Wave 1 between ERA5 and 20CRv3 in the EARLY period appears primarily responsible for the different SAM structure from 0°–90°E (Fig. 11a, b), with zonal SAM–SLP correlation anomalies less negative in 20CRv3.

Despite the marked differences between JRA-55 and the other two reanalyses in the EARLY period, there is a still a significant change between the phase of Wave 3 in the EARLY and LATE periods in this reanalysis ($p < 0.10$)

(Table 8). However, in contrast to ERA5 and 20CRv3, Wave 3 contributes the most variability in the EARLY period in JRA-55 and the variance explained diminishes in the LATE period, although it remains higher than the other wave-numbers. This, in combination with the reduced annular structure in the EARLY period compared to the two other reanalyses (Tables 6, 7, 8), explains why, uniquely, there is no distinction in the Peninsula SAM–SAT relationship between the two periods in JRA-55. Also dissimilar to the two other reanalyses is the increase in the variance explained by Wave 1. Thus, in contrast to ERA5 and 20CRv3 the annularity of the SAM structure increases from the EARLY to LATE period in JRA-55 (Fig. 11c).

Fig. 11 The phase and magnitude of the first four planetary waves computed from the zonal SAM-SLP correlation anomalies at 55°S for the EARLY and LATE periods. **a** ERA5, **b** 20CRv3 and **c** JRA-55. Wave-numbers 1–4 are shown in blue, yellow, green and purple, respectively. The thick black and red lines represent the actual correlation anomaly and the sum of the first four wave-numbers, respectively



5 Discussion and conclusions

The advent of the ozone hole in the late 1970s has had a major impact on surface SH high-latitude climate change through increasing the frequency of SAM+ events (Polvani

et al. 2011; Fogt and Marshall 2020). However, older reanalyses have been found to be less than optimal for investigating the effects of ozone depletion because the coincident availability of satellite sounder data from 1979 significantly improves their accuracy, meaning that they are not

Table 6 Zonal wave-number parameters within the SAM structure in ERA5 in the EARLY and LATE periods

	Amplitude	Phase	Variance explained
EARLY Period			
Wave 1	0.38*	194°	53.9%**
Wave 2	0.17	91°	13.9%
Wave 3	0.18	32°***	15.4%
Wave 4	0.16	49°	12.3%
LATE Period			
Wave 1	0.22	152°	23.3%
Wave 2	0.22	128°	23.6%
Wave 3	0.30	77°	39.1%
Wave 4	0.14	33°	10.1%

Asterisks in the EARLY period data indicate a significant difference between the two periods: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 7 As Table 6 for 20CRv3

	Amplitude	Phase	Variance explained
EARLY Period			
Wave 1	0.28	149°	39.4%*
Wave 2	0.14	104°	10.5%
Wave 3	0.27	33°***	37.0%
Wave 4	0.11	50°	8.9%
LATE Period			
Wave 1	0.18	141°	18.1%
Wave 2	0.22	124°	24.3%
Wave 3	0.30	77°	40.9%
Wave 4	0.15	33°	12.3%

homogeneous before and after this time. To determine whether more recent reanalyses are better than their predecessors in the pre-satellite era, we examined the capabilities of three current reanalyses that begin before 1979 to reproduce and help explain the pronounced observed change in SAM structure and its effect on Antarctic SAT between 1950 and 1979 (EARLY period) and 1980–2020 (LATE period).

In the validation exercise we find that ERA5 is the best reanalysis at reproducing Antarctic SAT in the pre-satellite era and is also the most homogeneous between the EARLY and LATE periods. In the EARLY period its RMSE and variance ratio statistics are significantly better than 20CRv3 and JRA-55 (Tables 2 and 3, Figs. 2a, b). All three reanalyses show significant improvement in their correlation with observed SAT from the EARLY to LATE period, with ERA5 having significantly higher correlations than the other two reanalyses in both periods (Table 5, Fig. 2c). Interestingly, despite not assimilating surface SAT measurements, the 20CRv3 correlations are significantly better than JRA-55

Table 8 As Table 6 for JRA-55

	Amplitude	Phase	Variance explained
EARLY Period			
Wave 1	0.12	197°	13.7%
Wave 2	0.14	63°	17.6%
Wave 3	0.27	61°*	56.9%
Wave 4	0.08	62°	5.0%
LATE Period			
Wave 1	0.24	132°	28.1%
Wave 2	0.21	123°	22.1%
Wave 3	0.29	76°	37.6%
Wave 4	0.13	32°	8.7%

in the EARLY period. Seasonally, in ERA5 there are no significant differences in RMSE and variance ratio between the EARLY and LATE periods (Fig. 3a, b). However, correlation shows significant improvement in all seasons except summer, mirroring the annual data (Fig. 3c). Thus, ERA5 is more homogeneous in summer but at the cost of reduced skill in the LATE period relative to the other seasons.

Regarding the SAM indices, ERA5 and the 20CRv3 ensemble mean produce a good representation of SAM in both periods: JRA-55 is only as good as the two other reanalyses in the LATE period. Marked negative SAM values in the EARLY period in JRA-55 (Fig. 4) are principally due to it having too high SLP at SH high-latitudes, typical of earlier reanalyses and corresponding with the findings of Huai et al. (2019). By comparing the 80 different ensemble members of 20CRv3, we show a reduced spread of annual SAM values with time, especially after the IGY in 1957/58, when many new Antarctic stations began observing SLP.

All three reanalyses indicate a marked change in decadal Antarctic SAM-SAT relationships between the EARLY and LATE periods. ERA5 most closely matches observations while JRA-55 is the poorest at reproducing them. In particular, there is no change in the sign of the SAM-SAT relationship at the three Peninsula stations in this reanalysis because a positive correlation already exists in the EARLY period (Fig. 5c). The switch in sign in the Peninsula is well-represented in ERA5 and 20CRv3, even though the change occurs a year earlier than observed—1970–79 rather than 1971–80—at Vernadsky and Esperanza (Figs. 5a, b). The variable sign of the SAM-SAT relationship at East Antarctic stations in the EARLY period is generally apparent in ERA5 and 20CRv3 although the sign associated with a particular decade doesn't always match observations. JRA-55 is also the least accurate of the three reanalyses in this region, with continuous negative decadal SAM-SAT relationships across both periods at the majority of East Antarctic stations (Fig. 5c).

We utilised the zonal SAM-SLP correlation anomalies at 55°S as a summary diagnostic to investigate changes in SAM structure between the EARLY and LATE periods. During the EARLY period there is an extensive region of northward projecting negative correlations (NPNCs) from the South Atlantic to the Indian Ocean in ERA5, with southward projecting negative correlations (SPNCs) in the ABS, contrasting with the principal NPNC located there in the LATE period. ERA5 demonstrates the greatest zonal extent of statistically significant differences in SAM structure between the two periods, most notably the ABS, but also the western Weddell Sea and parts of the southern Indian Ocean (Fig. 9a). 20CRv3 reveals a SAM structure broadly similar to ERA5 during the EARLY period although the region of NPNCs is reduced in size (Fig. 9b). However, the SAM structure in JRA-55 is markedly different from the other two reanalyses during the EARLY period: it is less annular, having a strong wave-number 3 component. One of the NPNCs is close to that in the ABS in the LATE period and, therefore, in contrast to ERA5 and 20CRv3, there are no significant differences in SAM structure in this region in JRA-55 (Fig. 9c).

By estimating the meridional wind component associated with SAM structure, we show that in the Antarctic Peninsula region both ERA5 and 20CRv3 demonstrated a switch from SAM+ being linked with cold southerly winds in the EARLY period to warm northerly winds in the LATE period, thus providing a simple explanation for the SAM-SAT relationship reversal (Fig. 10a, b). In JRA-55 both periods have SAM+ associated with northerlies and thus no reversal occurs in this region (Fig. 10c). Elsewhere in Antarctica, differences in the sign of the meridional wind associated with SAM+ between ERA5 and 20CRv3 are likely responsible for the differences in the SAM-SAT relationship between these two reanalyses. One such example is in the Weddell Sea sector (e.g., at Novolazarevskaya and Syowa; c.f., Fig. 5a, b). We also examined evidence for changes in the annual SAM-SAT structure in SH mid-latitude meteorological stations (not shown): while present at some individual locations, there were no clear regional signals similar to Antarctica.

To quantify longitudinal variations in the SAM structure we decomposed zonal SAM-SLP correlation anomalies into the first four zonal wave numbers. The primary significant difference across all three reanalyses is an eastward change in the phase of wave number 3 (Table 6, 7, 8). However, in ERA5 and 20CRv3 this is associated with an increase in its contribution to SAM structure in the LATE period whereas for JRA-55 it is reversed. Thus, in the first two reanalyses SAM structure is more zonal during the EARLY period whereas for JRA-55 the structure is actually less zonal.

Wachter et al. (2020) demonstrated that some of the SAM structure variability within the LATE period could

be attributed to changes in broad-scale oceanic SST variability in the Pacific and Atlantic Oceans, as represented by the Pacific Decadal Oscillation (PDO) and AMV teleconnection patterns, respectively, although we note the existence of the latter has recently been disputed (Mann et al. 2020). Here, we briefly consider the possible impact of the IPO (Henley et al. 2015) and AMV (Enfield et al. 2001) on the identified changes in SAM structure between the EARLY and LATE periods. There have been many previous studies linking tropical Pacific SST changes with SAM variability (e.g., Ding et al. 2012; Schneider et al. 2012; Clem et al. 2016): they find that the teleconnection is made via the Pacific South American pattern Rossby wave train and this feature is primarily responsible for the asymmetric component of the SAM structure in the LATE period. A specific example of the influence of tropical Pacific SSTs on Antarctic SAM-SAT relationships is described by Clem et al. (2020), who established that the changing relationship between the polarity of the IPO (residual after removing ENSO) and the SAM has led to the recent warming at the South Pole (Amundsen-Scott). In Fig. 1b this is apparent as a less strongly negative SAM-SAT relationship (actually turning positive in some decades in ERA5 and 20CRv3: c.f., Fig. 5a, b). Other authors have described a teleconnection between the AMV and Antarctic climate, with positive SSTs in the north and tropical Atlantic leading to extratropical SLP changes resembling SAM+ (Li et al. 2014).

The switch from the EARLY to LATE periods approximates to the time when both the decadal-smoothed IPO and AMO are at their most negative (Fig. S1) but there is no clear qualitative relationship between the teleconnection patterns and the marked change in SAM structure. Indeed, both the decadal-smoothed IPO and AMO change sign within the LATE period while the AMV also does so in the EARLY period. A quantitative analysis of the relationship between the IPO and AMO and the characteristics of the first four zonal wave-numbers (described in the Supplementary Information) reveals no statistically significant correlations across the full 62 decades or either of the two periods.

Of course, a major issue with this simple analysis is the small number of independent decadal-length samples available from the 62-year time series. One potential methodology to determine whether statistically significant relationships between tropical SST variability and decadal changes in SAM structure might become apparent in longer time series would be to examine climate model control runs of several hundred years in length. However, Marshall and Bracegirdle (2015) demonstrated that climate models from the fifth Climate Model Intercomparison Project (CMIP5) (Taylor et al. 2012) were generally poor at reproducing observed SAM-SAT relationships and thus the changes in SAM structure driving them. As an example, in Fig. S2 we reproduce Fig. 1b

using data from the historical run of HadGEM2-ES, the GCM able to replicate the SAM-SAT relationship at the most Antarctic stations in Marshall and Bracegirdle (2015). Fig. S5 reveals very little similarity with the observations: the greatest difference is the marked spatial and temporal variability in the sign of the decadal SAM-SAT correlations at East Antarctic stations. There is perhaps some evidence of a longer-term switch from a negative to positive correlation in the Antarctic Peninsula region but this occurs five years after observations. Given that many climate models also struggle to reproduce correct facsimiles of the IPO and AMV (Han et al. 2016; Henley et al. 2017), it seems unlikely that they will be able to accurately replicate any relationship between SST variability and Antarctic SAM-SAT relationships.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00382-022-06292-3>.

Acknowledgements We thank the reviewers for their constructive comments, which helped improve the clarity of the manuscript. GJM and JT were supported by the UK Natural Environment Research Council (NERC) through the British Antarctic Survey research programme "Polar Science for Planet Earth." RFL acknowledges partial support from NSF Office of Polar Programs grant 1744998. KRC acknowledges support from the Royal Society of New Zealand Marsden Fund grant MFP-VUW2010.

Funding RFL acknowledges partial support from NSF Office of Polar Programs grant 1744998. KRC acknowledges support from the Royal Society of New Zealand Marsden Fund grant MFP-VUW2010.

Data availability The READER data were obtained from <https://legacy.bas.ac.uk/met/READER/data.html> (doi: GB/NERC/BAS/PDC/00794); the SAT data for Byrd station from https://polarmet.osu.edu/datasets/Byrd_recon/byrd_tdemp_recon_monthly_revised.txt; the observation-based SAM Index from <https://legacy.bas.ac.uk/met/gjma/sam.html>; the 'Fogt' SAM reconstruction from https://polarmet.osu.edu/ACD/sam/sam_recon.html; the ERA5 reanalysis data were obtained from <https://cds.climate.copernicus.eu>; the 20CRv3 reanalysis data were obtained from https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html; the JRA-55 reanalysis data were obtained from <https://rda.ucar.edu/datasets/ds628.1>; the IPO data from <https://psl.noaa.gov/data/timeseries/IPOTP/tpi.timeseries.hasisst11.data> and the AMV data from <https://psl.noaa.gov/data/correlation/amon.us.data>.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allen MR, Smith LA (1996) Monte Carlo SSA: detecting irregular oscillations in the presence of colored noise. *J Clim* 9:3373–3404. [https://doi.org/10.1175/1520-0442\(1996\)009%3c3373:MCS-DIO%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009%3c3373:MCS-DIO%3e2.0.CO;2)
- Bozkurt D, Bromwich DH, Carrasco J, Hines KM, Maureira JC, Rondanelli R (2020) Recent near-surface temperature trends in the Antarctic Peninsula from observed, reanalysis and regional climate data. *Adv Atmos Sci* 37:477–493. <https://doi.org/10.1007/s00376-020-9183-x>
- Bracegirdle TJ, Marshall GJ (2012) The reliability of Antarctic tropospheric pressure and temperature in the latest global reanalyses. *J Clim* 25:7138–7146. <https://doi.org/10.1175/JCLI-D-11-00685.1>
- Bromwich DH, Fogt RL (2004) Strong trends in the skill of ERA-40 and NCEP-NCAR reanalyses in the high and midlatitudes of the Southern Hemisphere, 1958–2001. *J Clim* 17:4603–4619. <https://doi.org/10.1175/3241.1>
- Bromwich DH, Nicolas JP, Monaghan AJ, Lazzara MA, Keller LM, Weidner GA, Wilson AB (2013) Central West Antarctica among the most rapidly warming regions on Earth. *Nature Geosci* 6:139–145. <https://doi.org/10.1038/NGEO1671>
- Bromwich DH, Nicolas JP, Monaghan AJ, Lazzara MA, Keller LM, Weidner GA, Wilson AB (2014) Correction: corrigendum: central west Antarctica among the most rapidly warming regions on Earth. *Nature Geosci* 7:76. <https://doi.org/10.1038/NGEO2016>
- Cape MR, Vernet M, Skvarca P, Marinsek S, Scambos T, Domack E (2015) Foehn winds link climate-driven warming to ice shelf evolution in Antarctica. *J Geophys Res: Atmos* 120:11037–11057. <https://doi.org/10.1002/2015JD023465>
- Clem KR, Renwick JA, McGregor J, Fogt RL (2016) The relative influence of ENSO and SAM on Antarctic Peninsula climate. *J Geophys Res: Atmos* 121:9324–9341. <https://doi.org/10.1002/2016JD025305>
- Clem KR, Fogt RL, Turner J, Lintner BR, Marshall GJ, Miller JR, Renwick JA (2020) Record warming at the South Pole during the past three decades. *Nat Clim Change* 10:762–770. <https://doi.org/10.1038/s41558-020-0815-z>
- Compo GP et al (2011) The twentieth century reanalysis project. *Q J Roy Met Soc* 137:1–28. <https://doi.org/10.1002/qj.776>
- Ding Q, Steig EJ, Battisti DS, Wallace JM (2012) Influence of the tropics on the Southern Annular Mode. *J Clim* 25:6330–6348. <https://doi.org/10.1175/JCLI-D-11-00523.1>
- Enfield DB, Mestas-Núñez AM, Trimble PJ (2001) The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys Res Lett* 28:2077–2080. <https://doi.org/10.1029/2000GL012745>
- Fan K (2007) Zonal asymmetry of the Antarctic Oscillation. *Geophys Res Lett* 34. <https://doi.org/10.1029/2006GL028045>
- Fogt RL, Marshall GJ (2020) The Southern Annular Mode: variability, trends, and climate impacts across the Southern Hemisphere. *Wires Clim Change* 11:e652. <https://doi.org/10.1002/wcc.652>
- Fogt RL, Perlwitz J, Monaghan AJ, Bromwich DH, Jones JM, Marshall GJ (2009) Historical SAM variability. Part II: Twentieth-century variability and trends from reconstructions, observations and the IPCC AR4 models. *J Clim* 22:5346–5365. <https://doi.org/10.1175/2009JCLI2786.1>

- Fogt RL, Jones JM, Renwick J (2012) Seasonal zonal asymmetries in the Southern Annular Mode and their impact on regional temperature anomalies. *J Clim* 25:6253–6270. <https://doi.org/10.1175/JCLI-D-11-00474.1>
- Gallant AJE, Phipps SJ, Karoly DJ, Mullan AB, Lorrey AM (2013) Nonstationary Australasian teleconnections and implications for paleoclimate reconstructions. *J Clim* 26:8827–8849. <https://doi.org/10.1175/JCLI-D-12-00338.1>
- Gillett NP, Kell TD, Jones PD (2006) Regional climate impacts of the Southern Annular Mode. *Geophys Res Lett* 33:L23704. <https://doi.org/10.1029/2006GL027721>
- Gong D, Wang S (1999) Definition of antarctic oscillation index. *Geophys Res Lett* 26:459–462. <https://doi.org/10.1029/1999GL900003>
- Goyal R, Jucker M, Sen Gupta A, Hendon HH, England MH (2021) Zonal wave 3 pattern in the Southern Hemisphere generated by tropical convection. *Nat Geosci* 14:732–738. <https://doi.org/10.1038/s41561-021-00811-3>
- Gusti M (2020) ERA5 back extension 1950–1978 (preliminary version): large bias in surface analysis over Australia prior to 1970. ECMWF. <http://confluence.ecmwf.int/display/CKB/ERA5+back+extension+1950-1978+%28preliminary+version%29%3A+large+bias+in+surface+analysis+over+Australia+prior+to+1970>. Accessed 1 May 2021
- Han Z, Luo F, Li S, Gao Y, Furevik T, Svendsen L (2016) Simulation by CMIP5 models of the Atlantic multidecadal oscillation and its climate impacts. *Adv Atmos Sci* 33:1329–1342. <https://doi.org/10.1007/s00376-016-5270-4>
- Henley BJ, Gergis J, Karoly DJ, Power S, Kennedy J, Folland CK (2015) A tripole index for the Interdecadal Pacific Oscillation. *Clim Dyn* 45:3077–3090. <https://doi.org/10.1007/s00382-015-2525-1>
- Henley BJ et al (2017) Spatial and temporal agreement in climate model simulations of the Interdecadal Pacific Oscillation. *Environ Res Lett* 12:044011. <https://doi.org/10.1088/1748-9326/aa5cc8>
- Hersbach H et al (2020) The ERA5 global reanalysis. *Q J Royal Met Soc* 146:1999–2019. <https://doi.org/10.1002/qj.3803>
- Hines KM, Bromwich DH, Marshall GJ (2000) Artificial surface pressure trends in the NCEP-NCAR reanalysis over the Southern Ocean and Antarctica. *J Clim* 13:3940–3952. [https://doi.org/10.1175/1520-0442\(2000\)013%3c3940:ASPTIT%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013%3c3940:ASPTIT%3e2.0.CO;2)
- Huai B, Wang Y, Ding M, Zhang J, Dong X (2019) An assessment of recent global atmospheric reanalyses for Antarctic near surface air temperature. *Atmos Res* 226:181–191. <https://doi.org/10.1016/j.atmosres.2019.04.029>
- Hudson SR, Brandt RE (2005) A look at the surface-based temperature inversion on the Antarctic Plateau. *J Clim* 18:1673–1696. <https://doi.org/10.1175/JCLI3360.1>
- Jones JM, Fogt RL, Widmann M, Marshall GJ, Jones PD, Visbeck M (2009) Historical SAM variability Part i: century-Length Reconstructions. *J Clim* 22:5319–5345. <https://doi.org/10.1175/2009JCLI2785.1>
- Jones RW, Renfrew IA, Orr A, Webber BGM, Holland DM, Lazzara MA (2016) Evaluation of four global reanalysis products using in situ observations in the Amundsen Sea Embayment, Antarctica. *J Geophys Res: Atmos* 121. <https://doi.org/10.1002/2015JD02468>
- King JC et al (2021) Inhomogeneity of the surface air temperature record from Halley, Antarctica. *J Clim* 34:4771–4783. <https://doi.org/10.1175/JCLI-D-20-0748.1>
- Kobayashi S et al (2015) The JRA-55 reanalysis: general specification and basic characteristics. *J Met Soc Japan* 93:5–48. <https://doi.org/10.2151/jmsj.2015-001>
- Li X, Holland DM, Gerber EP, Yoo C (2014) Impacts of the north and tropical Atlantic Ocean on the Antarctic Peninsula and sea ice. *Nature* 505:538–542. <https://doi.org/10.1038/nature12945>
- Mann ME, Steinman BA, Miller SK (2020) Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nature Comm* 11:49. <https://doi.org/10.1038/s41467-019-13823-w>
- Marshall GJ (2003) Trends in the Southern Annular Mode from observations and reanalyses. *J Clim* 16:4134–4143. [https://doi.org/10.1175/1520-0442\(2003\)016%3c4134:TITSAM%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016%3c4134:TITSAM%3e2.0.CO;2)
- Marshall GJ (2007) Half-century seasonal relationships between the Southern Annular Mode and antarctic temperatures. *Int J Climatol* 27:373–383. <https://doi.org/10.1102/joc.1407>
- Marshall GJ, Bracegirdle TJ (2015) An examination of the relationship between the Southern Annular Mode and Antarctic surface air temperatures in the CMIP5 historical runs. *Clim Dyn* 45:1513–1535. <https://doi.org/10.1007/s00382-014-2406-z>
- Marshall GJ, Harangozo SA (2000) An appraisal of NCEP/NCAR reanalysis MSLP data viability for climate studies in the South Pacific. *Geophys Res Lett* 27:3057–3060. <https://doi.org/10.1029/2000GL011363>
- Marshall GJ, Thompson DWJ (2016) The signatures of large-scale patterns of atmospheric variability in Antarctic surface temperatures. *J Geophys Res: Atmos* 121:3276–3289. <https://doi.org/10.1002/2015JD024665>
- Marshall GJ, Orr A, van Lipzig NPM, King JC (2006) The impact of a changing Southern Hemisphere Annular Mode on Antarctic Peninsula summer temperature. *J Clim* 19:5388–5404. <https://doi.org/10.1175/JCLI3844.1>
- Marshall GJ, di Battista S, Naik SS, Thamban M (2011) Analysis of a regional change in the sign of the SAM-temperature relationship in Antarctica. *Clim Dyn* 36:277–287. <https://doi.org/10.1007/s00382-009-0682-9>
- Marshall GJ, Orr A, Turner J (2013) A predominant reversal in the relationship between the SAM and East Antarctic temperatures during the twenty-first century. *J Clim* 26:5196–5204. <https://doi.org/10.1175/JCLI-D-12-00671.1>
- Orr A et al (2008) Characteristics of summer airflow over the Antarctic Peninsula in response to recent strengthening of westerly circumpolar winds. *J Atmos Sci* 65:1396–1413. <https://doi.org/10.1175/2007JAS2498.1>
- Polvani LM, Waugh DW, Correa GJP, Son S-W (2011) Stratospheric ozone depletion: the main driver of twentieth-century atmospheric circulation changes in the Southern Hemisphere. *J Clim* 24:795–812. <https://doi.org/10.1175/2010JCLI3772.1>
- Raphael MN (2004) A zonal wave 3 index for the Southern Hemisphere. *Geophys Res Lett* 31:L23212. <https://doi.org/10.1029/2004GL020365>
- Raphael MN et al (2016) The Amundsen Sea Low: variability, change and impact on Antarctic climate. *Bull Amer Met Soc* 97:111–121. <https://doi.org/10.1175/BAMS-D-14-00018.1>
- Schneider DP, Okumura Y, Deser C (2012) Observed Antarctic interannual climate variability and tropical linkages. *J Clim* 25:4048–4066. <https://doi.org/10.1175/JCLI-D-11-00273.1>
- Silvestri G, Vera C (2009) Nonstationary impacts of the Southern Annular Mode on Southern Hemisphere climate. *J Clim* 22:6142–6148. <https://doi.org/10.1175/2009JCLI3036.1>
- Silvinski LC et al (2019) Towards a more reliable historical reanalysis: improvements for version 3 of the Twentieth Century Reanalysis system. *Q J Roy Met Soc* 145:2876–2908. <https://doi.org/10.1002/qj.3598>
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 experiment design. *Bull Am Meteorol Soc* 93:485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- Tetzner D, Thomas E, Allen C (2019) A validation of ERA5 reanalysis data in the southern Antarctic Peninsula-Ellsworth Land region, and its implications for ice core studies. *Geosci* 9:289. <https://doi.org/10.3390/geosciences9070289>

- Thompson DWJ, Solomon S (2002) Interpretation of recent Southern Hemisphere climate change. *Science* 296:895–899. <https://doi.org/10.1126/science.1069270>
- Thompson DWJ, Solomon S, Kushner PJ, England MH, Grise KM, Karoly DJ (2011) Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nat Geosci* 4:741–749. <https://doi.org/10.1038/NCEO1296>
- Turner J et al (2004) The SCAR READER project: toward a high-quality database of mean Antarctic meteorological observations. *J Clim* 17:2890–2898. [https://doi.org/10.1175/1520-0442\(2004\)017%3c2890.TSRPTA%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017%3c2890.TSRPTA%3e2.0.CO;2)
- Turner J et al (2016) Absence of 21st century warming on Antarctic Peninsula consistent with natural variability. *Nature* 505:538–542. <https://doi.org/10.1038/nature18645>
- Turner J, Marshall GJ, Clem K, Colwell S, Phillips T, Lu H (2020) Antarctic temperature variability and change from station data. *Int J Climatol* 40:2986–3007. [10.02/joc.6378](https://doi.org/10.1002/joc.6378)
- van den Broeke MR, van Lipzig NPM (2003) Response of winter-time Antarctic temperatures to the Antarctic Oscillation: results of a regional climate model. In: Domack E et al. (eds) *Antarctic Peninsula climate variability: historical and palaeoenvironmental perspectives*, Antarctic Research Series 79 Amer Geophys Union, pp 43–58. <https://doi.org/10.1029/AR079p0043>
- Wachter P, Beck C, Philipp A, Höppner K, Jacobeit J (2020) Spatiotemporal variability of the Southern Annular Mode and its influence on Antarctic surface temperatures. *J Geophys Res: Atmos* 125:e2020JD033818. <https://doi.org/10.1029/2020JD033818>
- Wang Y, Zhou D, Bunde A, Havlin S (2016) Testing reanalysis data sets in Antarctica: trends, persistence properties, and trend significance. *J Geophys Res: Atmos* 121:12839–12855. <https://doi.org/10.1002/2016JD024864>
- Wilcoxon F (1945) Individual comparisons by ranking methods. *Biometrics Bull* 1:80–83. <https://doi.org/10.2307/3001968>
- Wilks D (2016) “The stippling shows statistically significant grid points”: How research results are routinely overstated, overinterpreted, and what to do about it. *Bull Amer Met Soc* 97:2263–2273. <https://doi.org/10.1175/BAMS-D-15-00267.1>
- Zhang Y, Wang Y, Huai B, Ding M, Sun W (2018) Skill of the two 20th century reanalyses in representing Antarctic near-surface temperature. *Int J Climatol* 38:4225–4238. <https://doi.org/10.1002/joc.5563>
- Zhu J, Xie A, Qin X, Wang Y, Xu B, Wang Y (2021) An assessment of ERA5 reanalysis for Antarctic near-surface air temperature. *Atmos* 12:217. <https://doi.org/10.3390/atmos12020217>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.