The AntSMB dataset: a comprehensive compilation of surface mass balance field observations over the Antarctic Ice Sheet

Yetang Wang[1], Minghu Ding[2], Carleen H. Reijmer[3], Paul C. J. P. Smeets[3], Shugui Hou[4], Cunde Xiao[5]

[1]College of Geography and Environment, Shandong Normal University, Jinan 250014, China
[2]Institute of Tibetan Plateau and Polar Meteorology, Chinese Academy of Meteorological Sciences, Beijing 100081, China
[3]Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, Netherlands
[4]School of Oceanography, Shanghai Jiao Tong University, Shanghai 200240, China
[5]State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

Correspondence to: Yetang Wang(yetangwang@sdnu.edu.cn), and Cunde Xiao (cdxiao@bnu.edu.cn)

Abstract. A comprehensive compilation of observed records is needed for accurate quantification of surface mass balance (SMB) over Antarctica, which is a key challenge for calculation of Antarctic contribution to global sea level change. Here, we present the AntSMB dataset: a new quality-controlled dataset of a variety of published field measurements of the Antarctic Ice Sheet SMB by means of stakes, snow pits, ice cores, ultrasonic sounders and ground-penetrating radars. The dataset collects 268913 individual multi-year averaged observations, 687 annual resolved time series from 675 sites extending back the past 1000 years, and 78968 records at daily resolution from 32 sites across the whole ice sheet. These records are derived from ice core, snow pits, stakes/stake farms, ultrasonic sounders and ground-penetrating radar measurements. This is the first ice-sheet-scale compilation of SMB records at different temporal (daily, annual and multi-year) resolutions from multiple types of measurements, which is available at: https://doi.org/10.11888/Glacio.tpdc.271148 (Wang et al., 2021). The database has potentially wide applications such as the investigation of temporal and spatial variability in SMB, model validation, assessment of remote sensing retrievals and data assimilation.

1 Introduction

Under the background of rapid global warming, wide international concerns have been arouse on changes in the Antarctic Ice Sheet (AIS) mass balance, which positively contributed 14.0±2.0 mm to global sea level rise over 1979-2017 (Rignot et al., 2019). Antarctic mass balance is dependent on the partitioning between ice discharge into the ocean and net snow accumulation at the surface, i.e., surface mass balance (SMB). Recent negative mass balance of the ice sheet reflects larger ice dynamical loss than mass gain from SMB (e.g., Shepherd et al., 2012; Shepherd et al., 2018). Despite the responsibility of ice discharge for Antarctic mass balance on the decadal or longer time scales, considerable inter-annual variability is largely determined by fluctuations in SMB (Wouters et al., 2013). Because annual net mass input into the entire ice sheet through snowfall is equivalent to about 6 mm global sea level decline (Church et al., 2001), any small fluctuations in the Antarctic SMB can even result in large variability and trends of global sea level.

SMB is defined as the sum of precipitation, surface and drifting snow sublimation, erosion/deposition caused by drifting snow, and surface meltwater run-off. Since the first international polar year 1957/58 (IPY), a number of scientific Antarctic traverses/expeditions have been performed with the goals of SMB measurements by means of stakes, ice cores/snow pits,

35 ultrasonic sounders, or ground-penetrating radar (GPR). Due to logistical constraints in the harsh environment, gaps in the spatial coverage of SMB measurements are still large, and long-term samplings are also scarce (Favier et al., 2013). As a result, substantial caveats have been encountered when quantifying SMB at the ice sheet scale by using simple interpolation of these observations (Magand et al., 2008; Genthon et al., 2009). Climate models and various atmospheric reanalysis products provide an important choice to assess SMB for large areas. The outputs of regional climate models have been used

40 to calculate ice sheet SMB in recent decades by a wealth of Antarctic mass change estimate studies (e.g., Rignot et al., 2011; Shepherd et al., 2012; Rignot et al., 2019). However, these simulations depend on ground-based observations to improve their accuracy and resolution. Before application, the model's performance need to be carefully assessed based on in situ observations, as done by some previous studies (Medley et al., 2013; Wang et al., 2015; Van Wessem 2018; Agosta et al., 2019; Wang et al., 2020). To improve ice sheet SMB estimates, field measurements have been used by cross comparison

45 with remotely sensed data (Arthern et al., 2006), or outputs of the climate models (e.g., Monaghan et al., 2006; Van de Berg et al., 2006; Medley et al., 2019; Wang et al., 2019). Thus, it is still pivotal to compile all available observations from the past to present to better estimate spatial and temporal variability in SMB, and to constraint climate models and remote sensing algorithm.

50 Vaughan and Russell (1997) performed the pioneering work to compile all multi-year averaged SMB field measurement data over the AIS, and this compilation was detailly introduced by Vaughan et al. (1999). However, according to Magand et al. (2007), this dataset includes a lot of unreliable data, and should be used with caution. To improve this, Favier et al. (2013) updated the database using the new field measurements carried out during 1999-2012, through a quality control proposed by Magand et al. (2007). Recently, several compilations of SMB measurements at annual resolution have been published (e.g.,

55 Mayewski et al., 2013; Altnau et al., 2015; Thomas et al., 2017, Montgomery et al., 2018). In spite of numerous field measurements in these datasets, most cover only limited area of the AIS. In particular, these datasets missed a large amount of annually resolved stake/stake farm observations, such as data from the Japanese Antarctic Research Expedition (JARE), South Pole and Vostok, and so on. In addition, available SMB measurements derived from GPR are not or at least not fully collected into these datasets. Furthermore, all available ultrasonic sounder data from automatic weather stations (AWSs) at

60 daily or higher resolution have not been compiled until now.

In this study, our objective is to generate a comprehensive SMB database for Antarctica, using all available measurements by means of stake or stake network, snow pit or ice core, GPR, and ultrasonic sounder, with the control of data quality. This dataset includes SMB measurements at daily, annual, and multi-year resolutions, which can be applied for validation and

65   calibration of climate models and remote sensing, developments of remotely sensed algorithm, examination of spatial and temporal patterns in Antarctic SMB and estimate of the drivers of SMB changes across multiple scales. As a case of model validation, we make a comparison of the dataset with ERA5 reanalysis.

2 Description of the AntSMB dataset

2.1 Data collections and sources

70   We compile the dataset of SMB measurements over the AIS by searching the literature and public data portal platforms (e.g., the National Snow and Ice Data Center, NSIDC, PANGAEA and World Data Service for Paleoclimatology, NOAA), by collecting the supplements of publications, and by asking individual data generators to contribute their field measurements by email. If two or more request emails were not replied, we consider the data to be unavailable for the public, and thus they are not included in this dataset.

75

The new data resources of the records in the database include 175373 individual GPR measurements over West Antarctic coastal zones during 2010-2017 (Dattler et al., 2019), 6818 over the Thwaites Glacier in 2009 (Medely et al., 2013), and 1038 along between Dome C and Vostok in 2012 (Le Meur et al., 2018), respectively (Fig.1). A large amount of new stake measurements were acquired by revisiting the traverses from Zhongshan Station to Dome Argus (Ding et al., 2015), from

80   Syowa Station to Dome Fuji (Motoyama et al., 2015), and between Progress Station and Vostok Station (Khodzher et al., 2014). In addition to a new long-term ice core SMB records at the South Pole (Winski et al., 2019), this dataset includes previously published but unreleased time series of SMB records from ice cores drilled over the Lambert Glacier Basin (Xiao et al., 2001; Li et al., 2009; Ding et al., 2017). Furthermore, an important update of annual resolved SMB data results from the continuous stake network measurements performed at the South Pole, Vostok, and six sites of the transverse between

85   Syowa Station and Dome Fuji. In addition, this is the first public release of the published high-resolution ultrasonic sounder observations on Berkner Island (Reijmer et al., 1999; Reijmer and Van den Broeke, 2003), Dronning Maud Land (Van den Broeke et al., 2004), East Antarctic Plateau (Reijmer and Van den Broeke, 2003), and Chinese transvers from Zhongshan Station to Dome Argus (Liu et al., 2019), which are very useful for the investigation of intra-annual and seasonal cycles of SMB. The remainder of the database are obtained from existing SMB data compilations, including the multi-year averaged

90   SMB measurements by Favier et al. (2013) and Wang et al. (2016), time series of ice core records at annual resolution by Mayewski et al. (2013), Altnau et al. (2015) and Thomas et al. (2017), and SMB component measurements over the Antarctic Ice Sheet and Greenland Ice Sheet (SUMup dataset) by Montgomery et al. (2018)..

2.2 Selection criteria

In order to establish a comprehensive, complete and quality-controlled AIS SMB product for a variety of scientific

95   application, quantitative criteria are designed for record inclusion in the database to center on the high-resolution and well-dated records, and to optimize data spatial coverage. The criteria are as follows.

Firstly, the records must be published through peer-review or public available. The duration and temporal resolution of the records vary by the measurement types. We select the ultrasonic sounder records with the minimum duration of one year.

100   For annually resolved archives (ice core, stake and stake network measurements), the duration of records included in this dataset should be at least 10 years, but smaller than 1000 years. For the multi-year averaged observations, the included records for average span more than 3 years, which are the minimum number of years for an accurate estimation of the mean local SMB with the uncertainty of < 10% (Magand et al., 2007).

105   Secondly, the essential parameters for each SMB data are provided, including location, measurement methodology, data time coverage, and references to the primary data sources.

Thirdly, to ensure the multi-year averaged SMB data reliable at each site, we select the data determined by the anthropogenic radionuclides and volcanic horizons with errors of <10%, or stake measurements for more than three years, as suggested by

110   Magand et al. (2007). The observations based on both stable isotopes and chemical markers, and natural radionuclide are reliable (Magand et al., 2007), and thus included in the dataset. The available GPR-based snow accumulation rate data are included, because their uncertainties are <5% at a firn depth of 10 m, and decrease with the increase of the depth (Spikes et al., 2004; Eisen et al., 2008). SMB records of annual resolved ice cores should be either cross-dated or layer-counted. Their chronology should include at least two age control points, with one near the youngest part and another near the oldest part of

115   the time series. Also, they must be confirmed by the data generator. Furthermore, ice core SMB records are corrected for the impact of firn density and the vertical strain rate profile. The preliminary quality control for AWS snow accumulation data has been performed by data owner by means of removing the null measurements and physically anomalous snow accumulation data (i.e., data outside of the initial and final accumulation values). Some high-frequency noises still occur in the AWS snow accumulation data. To reduce the noises, we discard the data points outside of one standard deviation of a

120   running daily value as done by Fountain et al. (2010), and Cohen and Dean (2013)..

2.3 Types of data measurements collected in the AntSMB dataset

2.3.1 Stakes

Stakes are the easiest and most traditional way to measure SMB. After placing a stake vertically in the snow or ice, relative variations of snow surface heights over a certain period can be determined by repeated measurements of the distance

125 between the top of the stake and the surface. Changes in snow heights are multiplied by snow density to yield the corresponding SMB. This simple method has been widely applied over Antarctica by almost all national glacier surveys. However, in most cases, spatial representative of a single stake records is very limited due to small-scale disturbance from post-depositional effects such as the interactions between the stake and local wind. To reduce the related uncertainties, stake lines along a transect or stake farms are often used (e.g., Frezzotti et al., 2005; Kameda et al., 2008; Ding et al., 2011). In

130 particular, these measurements are useful for the investigation of the spatial distribution of SMB at the scale of < kilometer.

Given the repeated measurements, stake observations are only performed over the easily accessible regions. Due to logistic constraint in the extreme environment of Antarctica, the time span for the measurements usually ranges from 1 year to several years or even more.

2.3.2 Snow pits/ice cores

Snow pits and ice cores are used to construct SMB changes in time by determining the age and density of different layers. The dating is dependent on the different time markers preserved in the column of snow pit and ice core. Annual layer is dated through counting of seasonal changes in various parameters including the visual stratigraphy, the oxygen and hydrogen isotopic composition, major chemical ion content, hydrogen peroxide, electric conductivity, and so on. When intergraded

140 with the prominent horizons of known age from volcanic or radioactive markers, accuracy of dating is largely improved and results in time series of annual snow accumulation. Furthermore, the valuable reference horizons can be also used for the estimation of the SMB between horizons.

In Antarctica, counting annual layer based on the seasonal variations of multiparameter records combining with reference

145 horizons can calculate annual SMB on the high accumulation zones. However, seasonal cycles can hardly be identified at regions with low accumulation of <100 kg m-2 yr-1, especially for the East Antarctic Plateau. Thus, reference horizon may be the most reliable dating method at the low accumulation area, and only yields a mean SMB between two reference horizons.

2.3.3 GPR

150 GPR maps firn stratigraphy along a profile from the surface, and the radar identified firn layer with equal age along the continuous profiles can allow to gain a detailed insight into SMB patterns. To calculate SMB, the isochronous layers must be well dated, which is usually dependent on complementary depth-age of highly resolved ice core records along the radar profile. During the past few decades, grounded GPR has been widely used for the estimation of spatial variation of recent and historical SMB over Antarctica (e.g., Frezzotti et al., 2007; Anschutz et al., 2008; Müllerr et al., 2010). Most recently,

155 the newly developed airborne radar systems provide the revolutionized SMB measurements over the Antarctic Ice Sheet (Kanagaratnam et al., 2004, 2007). It can robustly resolve the stratigraphy at the shallow (10 m) to intermediate (100 m)

depths and hence to measure annual and multi-year accumulation rates at the width of hundreds of kilometers along aircraft flight tracks. The systems were firstly developed by the Center for Remote Sensing of Ice Sheets, and flown on the National Aeronautics and Space Administration Operation IceBridge (OIB) campaigns (Leuschen, 2010; Rodriguez-Morales et al.,

160   2013). The AntSMB database contains records of grounded and airborne GPR observations for the 2009-2019 OIB campaigns.

Relative to point measurements such as stakes, snow pits/ice cores, the advantage of GPR observations is to yield a more accurate representation of spatial variations of SMB. Furthermore, the radar images of deep internal horizons allow us to

165   quantify long-term variability in SMB. The errors of GPR-based SMB observations are associated with the depth and age of the reflector, and extrapolation of density along the radar profile. The resulting uncertainties were estimated to be about 4% of the calculated SMB at a firn depth of 10 m, and about 0.5% at the depth of 60 m (Spikes et al., 2004).

2.3.4 AWS

In Antarctica, some AWSs equipped with ultrasonic sensors measure snow surface height changes by detecting the vertical

170   distance to the surface. Combining with density observation, snow height changes can be converted to SMB. Despite the poor quality occasionally when the blowing snow or fog happens, this method can continuously yield a high (typically hourly) temporal resolution records of SMB (van den Broeke et al., 2004; Gorodetskaya et al., 2013), which can be utilized to identify individual accumulation/ablation events, to quantify seasonal cycle of snow accumulation, and also to calculate the surface energy balance coupled with other AWS observations.

175

Same as single ice core or stake observation, AWS measurements represent a single location, and spatial representativity is possibly limited. In addition, after collection of raw snow height data, the temperature-dependent speed of sound correction must be performed. The uncertainty of AWS height measurements is estimated to be ±1 cm or 0.4% of the distance to the surface. This means that the measurements are not sufficient to examine the smaller snow accumulation events usually

180   occurring on the interior of East Antarctic Plateau.

2.4 Structure and metadata

The AntSMB dataset includes three subsets which are composed of multi-year averaged SMB observations from stakes, ice cores and GPR measurements, annual resolved SMB measurements by ice cores, stakes and stake networks, and AWS daily

185   snow height measurements. To facilitate data reuse, subsampling and re-analysis for scientific research efforts, each record in the three sub-datasets include some essential information, i.e., the name of measurement sites, site locations, measurement method, time coverage of the measurements, and citations. Site locations include latitude, longitude and surface elevation. Each location of the measurements is in units of decimal degrees relative to the WGS84 ellipsoid. As listed in the dataset's

metadata, measurement techniques include firn/ice core, snow pits, stake or stake network, ultrasonic sounder and GPR.
190 Table 1 summarize the essential information for each measurement. Uncertainties of any measurement methods have been detailly discussed by Eisen et al. (2008).

Among the three subdatsets, the number of the multi-year mean SMB subdataset is largest, including 265334 unique measurements by radar isochrones, 2276 stake measurements, and 1303 ice core and snow pit observations (Figure 2). The
195 majority of these observations (~ 69%) are derived from airborne snow radar measurements in the coastal zone of West Antarctica and Antarctic Peninsula, Ronne Ice Shelf, South Pole, Pine Island and Thwaites glaciers (Medley et al., 2013; Medley et al., 2014; Dattler et al., 2019). GPR data from two transects in East Antarctica account for 30% of all measurements in this subdatset. In most cases, SMB values come from original measurements. However, for the Japanese traverse route from Syowa Station to Dome F, we updated multi-year averaged SMB by combining new stake surface height
200 measurements during 2007-2013 with the improved snow density data from Wang et al. (2015).

Annual resolved SMB subdataset contains 687 time series of records, of which 79 records come from the compilation of ice core snow accumulation by Thomas et al. (2017), and 26 from the shallow firn core records in Dronning Maud Land (DML) collected by Altnau et al. (2015). Continuous stake surface height measurements at sub-annual resolution are available for
205 the transverse route from Syowa Station to Dome F since the 1970s (Motoyama et al., 2015). We converted the measurements to SMB for the subdataset by multiplying the height changes by the improved snow density estimation by Wang et al. (2015).

AWS snow accumulation data are measured by the determination of the variations of the vertical heights between the sensor
210 and snow using surface ultrasonic height rangers. The measurements cover a wide range of areas, including coastal zone of East Antarctica (McMorrow et al., 2001) and West Antarctica (van Lipzig et al., 2004b) with high snow accumulation, the dry East Antarctic Plateau (Reijmer and Broeke, 2003; van den Broeke et al., 2004), Ross Ice Shelf (Cohen and Dean, 2013), Berkner Island, Lambert Glacier drain, McMurdo Dry Valleys (Doran et al., 2002). To minimize the impact of changes in the speed of sound, these data have been corrected using the simultaneous air temperature measurements from the AWSs.

215 **3. Spatial and temporal analysis of the AntSMB dataset**

3.1 Spatial coverage of SMB records

The comprehensive observed SMB database collects SMB field data at the daily, annual and multi-year scales from the whole AIS. Spatial distribution of the records is uneven within Antarctica. AWS snow accumulation measurements were obtained at 32 sites, of which ten are located at Dronning Maud Land, seven at the Ross Ice Shelf, and four along Chinese
220 transverse route from Zhongshan Station to Dome A (Figure 1a). Availability of time series from the annual resolution SMB

subdataset is rich for West Antarctica, Dronning Maud Land, Berkner Island and traverse from Syowa Station to Dome F (Figure 1b). However, large parts of Antarctic interior with low snow accumulation remain undocumented, which is easily understood because the seasonal stratigraphy in ice cores is almost unavailable at the regions with the accumulation of < 70 kg m^{-2} yr^{-1} (Frezzotti et al., 2007; Frezzotti et al., 2013). Compared with SMB compilation by Favier et al. (2013), spatial

225   coverage of the multi-year SMB subdataset has greatly improved, especially for West Antarctica and the Antarctic Peninsula. Despite the improvement of spatial distribution, SMB records are still poor for the region from the Filchner-Ronne ice shelf via the South Pole to Dome C, and for the coastal zone of East Antarctica.

3.2 Temporal variability in the SMB records

The records in the comprehensive SMB dataset cover different time spans, ranging from a minimum of 3 years to a
230   maximum of 1000 years. The covered time periods are closely associated with the measurement method. AWS provides very high-resolution measurements of snow height changes, but the records generally span only a few years (1-18 years). Although a significant advantage of ice cores is to record SMB changes over the long timespans, it is difficult to perform these observations at a high spatial density. Stake farms are the easiest method to observe SMB, but continuous measurements are available between several years and tens of years, largely due to the logistic constraints in the extreme
235   Antarctic environment. GPR can detect the local SMB from the last tens of years to about 1000 years along continuous profiles of the snowpack. The temporal resolution of GPR measurements is dependent on the age estimates of reflection horizons, and the resulting records in our dataset ranges from decadal to centennial.

For annual resolved SMB subdataset, of 183 time series from ice core and stake network measurements, 47 span the last 200
240   years (Fig. 3a). The number of time series peaks during the early 2000s when ice cores were retrieved in Dronning Maud Land (Altnau et al., 2015) and West AIS (Mayewski et al., 2013). Prior to 1800, the number of time series decreases greatly, with only ten with the duration beyond the past 500 years, and five beyond the past 1000 years (Fig. 3a). The sharp decline since the mid-2000s results from a lack of coring efforts. Annual resolved stake measurements cover the past 40 years, peaking from the mid-1990s to the early 2000s (Fig. 3b).
245

For the multi-year averaged SMB subdataset, 83% of the records with the exception of radar measurements cover <20 years, and 43% span <5 years. Figure 3c presents the distribution of years when these records were measured from 1950 onwards. The distribution of the measurements is relatively even, until the 1990s when the number of samples increase. The temporal coverage of radar observations ranges from 25 years to 185 years.

250   ### 4. Inter-comparison of the different types of SMB measurements

The dataset compiles the different types of SMB measurements including ice cores/snow pits, stakes, ultrasonic sensors and GPR approaches. It is critical to investigate if the resulting data have systematic discrepancy due to the distinct measurement

methods. In particular, the measurements by ice core, stake, and ultrasonic sensor are performed at the centimeter scale, whereas GPR samples at the meter scale. Despite the scale difference, near 100-year averaged GPR measurements agree well

255 with 5-year averaged single stake at the corresponding locations along the transect near Talos Dome, with the differences of around 10% (Frezzotti et al., 2007). Given that no existing observed SMB dataset can be used as an independent reference to the different types of Antarctic SMB observations, the inter-comparison of SMB determined by different methods at the same or near locations are made, as presented in Figure 4. It is clear that despite the different averaged time coverage, they provide a reasonable match with each other, with the largest discrepancy of <20%, which are consistent with the previous

260 similar inter-comparison (e.g., Vaughan et al., 2004; Frezzotti et al., 2005; Anschütz et al., 2007). In addition, no systematic errors between the different methods are found.

5. Comparison with the previous AIS SMB observation datasets

Here, we present an unprecedentedly comprehensive compilation of SMB observations at the daily, annual and multi-year scales. For the compilation of the multi-year averaged SMB ground based observations including stake/stake farm, snow

265 pit/ice core, and GPR, we apply the same quality control criteria as used in the compilation by Favier et al. (2013) updated by Wang et al. (2016). Compared with the dataset, our compilation greatly improved the data spatial coverage by updating records using more recent published observations along the margins of West Antarctica, across Marine Byrd Land and Antarctic Peninsula, around the South Pole, between Dome C and Vostok, and along the transects of Progress station–Vostok station, Dumont d'Urville–Dome C, and Talos Dome, etc. We also updated SMB records along the transects of

270 Zhongshan Station-Kunlun Station and Syowa Station-Dome F based on the recent revisiting measurements. In particular, our dataset provides the first comprehensive compilation of grounded and airborne GPR measurements.

In terms of the collections of annual resolution SMB measurements from ice core, the SUMup dataset focused on the limited ice core records at West Antarctica from the US International Trans-Antarctic Scientific Expedition during the early 2000s

275 (US ITASE, Mayewski et al., 2013) and several cores drilled in 2010/2011 (Medley et al., 2013), and over the DML and Berkner Island from the European Project for Ice Coring in Antarctica (EPICA, Oerter, 2008a-l). The Antarctica 2k database constructed by Thomas et al. (2017) included 80 ice core records spanning at least 30 years, and the shorter and other ground-based measurement records are omitted. However, our Ant-SMB dataset focuses on the collection of annually resolved snow accumulation records from different kinds of measurements covering the whole ice sheet. As a result, this

280 dataset contains 175 annually resolved ice core snow accumulation records, 8 stake network measurements covering at least 10 years, and 512 time series of continuous stake measurements spanning more than 18 years.

Previous SMB compilations centered on glaciological observations on annual and longer timescales (e.g., Vaughan et al.,1999; Frezzotti et al., 2013; Thomas et al., 2017), which are useful for the examination of trends and large-scale

285 variability of AIS snow accumulation. Nevertheless, they do not shed insight on SMB changes at much shorter timescales,

such as synoptic scale and accumulation events. AWSs provide high resolution (typically hour) snow accumulation measurements, which is an advantage to quantify seasonal cycle of SMB, and to examine the synoptic sources of individual accumulation events, relative to the other methods such as snow pits, ice cores, and stakes. Snow accumulation data from individual or several AWSs at the different sectors of Antarctica have been published by some previous studies (e.g.,

290 Reijmer and Van den Broeke, 2003; Thiery et al., 2012; Cohen and Dean, 2013; Thomas et al., 2015). However, these data have been not well compiled until now. Our dataset is the first attempt to collect all AWS snow accumulation measurements in Antarctica.

6. Comparison with ERA5

6.1 ERA5 output

295 Reanalysis utilizes a large amount of observations assimilated into a numerical model to generate a spatially and temporally complete state of the atmosphere. Because the main assimilated data are atmospheric and oceanic measurements, reanalysis outputs are not entirely subjective to the density of surface observations, and thus have the potential to provide important information over the regions with few or even no surface observation. Recent studies have revealed that European Centre for Medium Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim) is likely to be the best or among the best

300 reanalysis dataset for the representation of Antarctic precipitation (e.g., Bromwich et al., 2011; Wang et al., 2016).

ERA5 is the fifth generation ECMWF reanalysis product produced by the Integrated Forecasting System (IFS) Cy41r2 operational in 2016 (Hersbach et al., 2020). Compared with ERA-Interim, a major advantage of ERA5 is much higher horizontal vertical resolutions (~ 31km and 137 pressure levels, respectively), and more enhanced outputs (hourly). Furthermore, IFS Cy41r2 includes a more advanced 4DVar assimilation scheme together with an uncertainty estimation, and

305 much more observations are assimilated. Detailed improvements can be found in Hersbach et al. (2020). This reanalysis dataset has replaced ERA-Interim, of which updates were stopped on August, 2019. Here, our main objective is to determine if the AntSMB dataset is also capable of representing main features of SMB in space and time, compared to ERA5. Despite the recent release of ERA5 data extending back to 1950, we only use the outputs for the 1979-2018 period, due to the spurious shift of reanalysis outputs in 1979 largely caused by changes in the amount of assimilated observations (e.g., Zhang

310 et al., 2018; Huai et al., 2019; Wang et al., 2020). Although the problem is likely solved by ERA5, a careful assessment of discontinuity in ERA5 time series is still required before application. However, this is beyond the scope of this study.

6.2 A subset of data used for the comparison with ERA5

6.2.1 Multi-year averaged SMB observations

Given that the output of climate models center on climate information since 1979 in Antarctica, it is necessary to define a

315 special dataset for the model comparison. To match with the coverage period of the models, we only retain observations starting from 1950 onwards in the multi-year averaged SMB subdataset. In particular, we discard observations starting for

the 1950-1978 period with the time coverage of no more than 10 years. Because blowing snow processes are not schemed by ERA5, 190 measurements in blue ice regions (SMB values <0) are excluded. Finally, 191816 multi-year averaged observations are left for model-observation comparison.

6.2.2 Annual resolved SMB observations

To estimate the temporal performance of ERA5 for snow accumulation, we use the records from annually resolved SMB subdatabase covering at least 10 years starting from 1979. This results in 159 time series of annual resolved SMB. The representativeness of SMB measurements at a simple site for a region is influenced by local noises from the interaction between wind and local snow surface, especially in the regions with accumulation rate of <120 kg m-2 a-1 (e.g., Frezzotti et al., 2005, 2007; Ding et al., 2011). This can be confirmed by that ERA5 simulated individual records highly correlate with each other (r>0.70), but exhibit a variety of relationships with their corresponding ice core SMB time series at 35 cores on the DML plateau, including significantly negative, positive and insignificant correlations (Fig. 5a). Various linear relationships between the simulated and observed time series are also found over the Berkner Island and Ronne Ice Shelf with high density of cores, with r values ranging from -0.35 to 0.67. At the South Pole, ERA5 shows a significant correlation (r=0.68, p<0.05) with stake farm measurements, but fails to do so with the individual ice core records. A main possibility is that SMB derived from stake networks is less noisy by removing small-scale spatial variability based on the average of a lot of stakes together. To reduce local noise and better assess the performance of ERA5, we first average the individual observation records in the same grid cell, and then stack the averaged time series at the same geographic region. If there are ice core records and stake farm observations in the same location, the measurements of stake farm are utilized. Despite only one core at the top of four ice domes where the local noises are minor (Monaghan et al., 2006a), the records are not discarded in the estimate. Following Frezzotti et al. (2007) and (2013), a single ice core site with accumulations of >700 kg m-2 yr-1 allow the determination of annual SMB at ±10% accuracy, which corresponds to the accuracy derived from the instrumental measurement, and hence the corresponding ice core records are retained. If the records from a single ice core are confirmed to be less local noisy by data owners, we also don't omit them. After the composite and filtering, 48 locations or regions with annual resolved SMB are left to compare with ERA5 simulations.

6.3 Spatial performance of ERA5 output

A comparison of the density distribution of ERA5 precipitation minus evaporation (P-E) with the filtered multi-year averaged SMB observations reveals that the multi-year averaged dataset is representative of the entire P-E spectrum of the model at the continental scale (Fig.6a). As shown in Fig.6b, the dataset also represents well the samples elevation distribution of SMB in relation to the whole ice sheet, especially between 200 and 1000m elevations where it was not correctly sampled by the SMB observation dataset compiled by Favier (2013).

ERA5 reveals large spatial gradients of snow accumulation over the AIS (Fig. 7a), with values higher than 1000 kg m^{-2} yr^{-1} at the margins, and lower values (<30 kg m^{-2} yr^{-1}) on the hinterland of East Antarctic Plateau. There is a very high

350   correlation between ERA5 output and the observed SMB (R^2= 0.93, p < 0.01, which is calculated based on the logarithm of SMB values, due to the lognormal SMB distributions). This suggests a good representation of the major spatial patterns as presented by observations, such as coast-to-plateau SMB gradients. No systematic spatial bias is observed on the West AIS, whereas dry biases occur in most sites of inland East Antarctic Plateau, and wet biases in the section of between 30°W and 150°E of the East AIS margins (Fig. 7b). The mean bias accounts for 6.6% of the average of observed SMB, which is

355   slightly higher than regional climate models (MAR and RACMO2.3p2) (Agosta et al., 2019; Van Wessem et al., 2018). It is obvious that ERA5 robustly capture the sharp decrease in SMB with elevation (Fig.7c). Compared with observation in each 200 m elevation bin, ERA5 is slightly wet below 1600 m elevation, whereas dry biases occur in inland Antarctica with the elevations above 3000 m.

6.4 Temporal performance of ERA5 output

360   The correlation coefficients (r) between ERA5 simulations and SMB observations at 48 locations are shown in Fig.5b. Significant and high correlations are observed at two out of five sites over the Antarctic Peninsula, with r values of >0.7 (p<0.05). Over the WAIS, ERA5 simulations are correlated significantly with observations (r>0.45, p<0.05) at 14 out of 18 sites, and correlation coefficients exceed 0.8 at five sites, suggesting relatively good skills of simulated records for capturing the observed variability in accumulation rates. Significant and positive correlations are present over the plateau, and western

365   and eastern coastal areas of the DML. Correlation coefficients show that a large fraction of inter-annual changes (>70%) in SMB observations at Law Dome of the Wilkes Land. The performances are relatively good for South Pole, Vostok, and Talos Dome. However, no significant or negative correlations are observed at the sites of the Lambert Basin, the Princess Elizabeth Land, middle DML coasts and Adélie Land.

370   To further assess the temporal performance of ERA5, we use the continuous time series of stake measurements along the JARE traverse route from Syowa station to Dome F. These stake measurements are divided into four subgroups, as done for this traverse route by Wang et al. (2015). Stake measurements in each subgroup are stacked, and then compared with the composites of ERA5 simulations at the respective subgroup (Fig.8). ERA5 overestimates the observed SMB at the coastal and katabatic regions, but underestimates those at the inland plateau region. The modeled records match particularly well

375   with observations at the coastal, higher katabatic and inland plateau regions, with higher r2 values of >0.5. Observed SMB at the lower katabatic region is simulated well by the reanalysis dataset.

Overall, ERA5 fits interannual variability in observed SMB acceptely at most sites over the AIS, and this reveals much of atmospheric circulation is represented by this reanalysis product. Nevertheless, its performance is limited at some sites of

380   Lambert Basin, inland West Antarctica, and parts of East Antarctic coasts. These may result from the limited performance of

ERA5 for the storm frequency related to synoptic-scale circulations, and sublimation because of circulation variations. Detailed interpretation of uncertainty of ERA5 is beyond the scope of this study.

7. Data availability

The comprehensive SMB observation dataset is available through a Big Earth Data Platform for Three Poles. The dataset can

385 be downloaded from https://doi.org/10.11888/Glacio.tpdc.271148 (Wang et al., 2021). In this repository, the three subdatasets included in the entire dataset are provided in Excel spreadsheet format together with metadata files.

8. Discussion and conclusions

The dataset provides an unprecedentedly comprehensive compilation of SMB observations, with better spatial coverage than previous studies. In particular, our compilation greatly improves spatial density of measurements in the 200-1000 m

390 elevations where are not correctly sampled by the dataset from Favier et al. (2013). However, there is a clear need to increase the spatial density of annual resolved SMB measurements over the inland East AIS, and daily SMB observations over West Antarctica, and 90°-170°E sector of East Antarctica.

This dataset can be used to estimate the temporal and spatial changes in the AIS SMB. A temporal homogeneous

395 climatology of SMB for the second half of the 20th century may be obtained by temporal rescaling of the multi-year averaged SMB subdataset against ERA5 outputs. The available syntheses of time series of records from annual resolved SMB subdataset will allow to investigate regional snow accumulation changes during the past several decades or centuries (Kaspari, et al., 2004; Frezzotti et al., 2013; Altnau et al., 2015; Thomas et al., 2017). The combination of annual SMB subdataset with reanalysis products or the outputs of regional climate models can generate gridded datasets to better

400 constraint the temporal and spatial variability AIS SMB at the different scales (Monaghan et al., 2006b; Medley et al., 2019; Wang et al., 2019). The availability of AWS snow height measurements will allow insights into synoptic and seasonal patterns of SMB, which are vital for ice core dating studies.

In the current study, we have made a comparison between observation data and ERA5 output. As a result, in spite of

405 discrepancy in magnitude, ERA5 represents spatial variations of SMB observations well, and captures a large proportion of the inter-annual variability. Similarly, this dataset can be used to evaluate the quality of other atmospheric reanalyses, and regional or global climate models such as JRA-55, MERRA-2, RACMO2.3, MAR and CESM. Moreover, a high spatial density of stake and GPR measurements along several transections from coasts to inland are included in the dataset, which correctly sample the actual distribution of SMB, and thus allow to provide stringent constraints on the models in these

410 specific regions. Annual resolved SMB observations in the database are also likely to be used as an important input of data assimilation for paleoclimate reconstructions (Dalaiden et al., 2020). The dataset is of vital importance for improvement of

remote sensing algorithm for Antarctic snow accumulation/snowfall rate, such as CloudSat 2C-SNOW-PROFILE product (Palerme et al., 2014; Behrangi et al., 2016).

415 The scientific community are expected to apply this dataset for Antarctic hydrological studies, model-data inter-comparison and remotely sensed algorithm developments. The cryospheric community are also encouraged to further share their SMB observation data to update this dataset in the future.

References

425 Altnau, S., Schlosser, E., Isaksson, E., and Divine, D.: Climatic signals from 76 shallow firn cores in Dronning Maud Land, East Antarctica, The Cryosphere., 9, 925–944, https://doi.org/10.5194/tc-9-925-2015, 2015.

Agosta, C., Amory, C., Kittel, C., Orsi, A., Favier, V., Gallée, H., van den Broeke, M. R., Lenaerts, J. T. M., van Wessem, J. M., van de Berg, W. J., and Fettweis, X.: Estimation of the Antarctic surface mass balance using the regional climate model
430 MAR (1979– 2015) and identification of dominant processes, The Cryosphere., 13, 281–296, https://doi.org/10.5194/tc-13-281-2019, 2019.

Anschütz, H., Müller, K., Isaksson, E., McConnell, J. R., Fischer, H., Miller, H., Albert, M., and Winther, J.-G.: Revisiting sites of the South Pole Queen Maud Land Traverses in East Antarctica: Accumulation data from shallow firn cores, J.
435 Geophys. Res., 114, D24106, https://doi.org/10.1029/2009JD012204, 2009.

Arthern, R. J., Winebrenner, D. P., and Vaughan, D. G.: Antarctic snow accumulation mapped using polarization of 4.3-cm wavelength microwave emission, J. Geophys. Res., 111, D06107, https://doi.org/10.1029/2004JD005667, 2006.

440 Behrangi, A., Christensen, M., Richardson, M., Lebsock, M., Stephens, G., Huffman, G. J., Bolvin, D., Adler, R. F., Gardner, A., Lambrigtsen, B., and Fetzer, E.: Status of high-latitude precipitation estimates from observations and reanalyses, J. Geophys. Res. Atmos., 121, 4468–4486, https://doi.org/10.1002/2015JD024546, 2016.

Bromwich, D. H., Nicolas, J. P., and Monaghan, A. J.: An assessment of precipitation changes over Antarctica and the

445   Southern Ocean since 1989 in contemporary global reanalyses. J. Clim., 24, 4189–4209. https://doi.org/10.1175/2011JCLI4074.1, 2011.

Church, J. A., J. M. Gregory, P. Huybrechts, Kuhn, M., Lambeck, K., Nhuan, M. T., Qin, D., and Woodworth, P. L.: Changes in sea level. Climate Change 2001: The Scientific Basis, Cambridge University Press, Cambridge, 639-693 pp.,

450   2001.

Cohen, L., Dean, S., and Renwick, J.: Synoptic weather types for the Ross Sea region, Antarctica, J. Clim., 26, 636–649, https://doi.org/10.1175/JCLI-D-11-00690.1, 2013.

455   Dalaiden, Q., Goosse, H., Klein, F., Lenaerts, J. T. M., Holloway, M., Sime, L., and Thomas, E. R.: How Useful is snow accumulation in reconstructing surface air temperature in Antarctica? A study combining ice core records and climate models, The Cryosphere., 14, 1187–1207, https://doi.org/10.5194/tc-14-1187-2020, 2020.

Dattler, M. E., Lenaerts, J. T. M., and Medley, B.: Significant spatial variability in radar-derived West Antarctic

460   accumulation linked to surface winds and topography, Geophys. Res. Lett., 46, 126–134, https://doi.org/10.1029/2019GL085363, 2019.

Ding, M., Xiao, C., Li, Y., Ren, J., Hou, S., Jin, B., and Sun, B.: Spatial variability of surface mass balance along a traverse route from Zhongshan station to Dome A, Antarctica, J. Glaciol., 57, 204, 658–666,

465   https://doi.org/10.3189/002214311797409820, 2011.

Ding, M., Xiao, C., Li, C., Qin, D., Jin, B., Shi, G., Xie, A., and Cui, X.: Surface mass balance and its climate significance from the coast to Dome A, East Antarctica, Sci. China: Earth Sci., 58, 1787–1797, https://doi.org/10.1007/s11430-015-5083-9, 2015.

470

Favier, V., Agosta, C., Parouty, S., Durand, G., Delaygue, G., Gallée, H., Drouet, A.-S., Trouvilliez, A., and Krinner, G.: An updated and quality controlled surface mass balance dataset for Antarctica, The Cryosphere., 7, 583–597, https://doi.org/10.5194/tc-7-583-2013, 2013.

475 Frezzotti, M., Pourchet, M., Flora, O., Gandolfi, S., Gay, M., Urbini, S., Vincent, C., Becagli, S., Gragnani, R., Proposito, M., Severi, M., Traversi, R., Udisti, R., and Fily, M.: Spatial and temporal variability of snow accumulation in East Antarctica from traverse data, J. Glaciol., 51, 113–124, https://doi.org/10.3189/172756505781829502, 2005.

Frezzotti, M., Urbini, S., Proposito, M., Scarchilli, C., and Gandolfi, S.: Spatial and temporal variability of surface mass
480 balance near Talos Dome, East Antarctica. J. Geophys. Res., 112, F02032, https://doi.org/10.1029/2006JF000638, 2007.

Frezzotti, M., Scarchilli, C., Becagli, S., Proposito, M., and Urbini, S.: A synthesis of the Antarctic surface mass balance during the last 800 yr. The Cryosphere., 7, 303-319. https://doi.org/10.5194/tc-7-303-2013, 2013.

485 Genthon, C., Magand, O., Krinner, G., and Fily, M.: Do climate models underestimate snow accumulation on the Antarctic plateau? A re-evaluation of/from in situ observations in East Wilkes and Victoria Lands, Ann. Glaciol., 50, 50, 61–65, https://doi.org/10.3189/172756409787769735, 2009.

Gorodetskaya, I. V., van Lipzig, N. P. M., van den Broeke, M. R., Mangold, A., Boot, W., and Reijmer, C. H.:
490 Meteorological regimes and accumulation patterns at Utsteinen, Dronning Maud Land, East Antarctica: Analysis of two contrasting years. J. Geophys. Res.: Atmospheres, 118, 1700–1715, https://doi.org/10.1002/jgrd.50177, 2013.

Le Meur, E., Magand, O., Arnaud, L., Fily, M., Frezzotti, M., Cavitte, M., Mulvaney, R., Urbini, S.: Spatial and temporal distributions of surface mass balance between Concordia and Vostok stations, Antarctica, from combined radar and ice core
495 data: first results and detailed error analysis, The Cryosphere., 12, 5, 1831-1850, https://doi.org/10.5194/tc-12-1831-2018, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita,
500 M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. R. Meteorol. Soc., 146, 730, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

505 Huai, B., Wang, Y., Ding, M., Zhang, J., and Dong, X.: 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, 2019.

Kaspari, S., Mayewski, P. A., Dixon, D. A., Spikes, V. B., Sneed, S. B., Handley, M. J., and Hamilton, G. S.: Climate variability in West Antarctica derived from annual accumulation-rate records from ITASE firn/ice cores, Ann. Glaciol., 39, 510   585–594, https://doi.org/10.3189/172756404781814447, 2004.

Li, Y., Cole-Dai, J., and Zhou, L.: Glaciochemical evidence in an East Antarctica ice core of a recent (AD 1450–1850) neoglacial episode, J. Geophys. Res., 114, D08117, https://doi.org/10.1029/2008JD011091, 2009.

515   Magand, O., Genthon, C., Fily, M., Krinner, G., Picard, G., Frezzotti, M., and Ekaykin, A. A.: An up-to-date quality-controlled surface mass balance data set for the 90–180°E Antarctica sector and 1950–2005 period, J. Geophys. Res., 112, D12106, https://doi.org/10.1029/2006JD007691, 2007.

Mayewski, P. and Dixon, D. A.: US International Trans-Antarctic Scientific Expedition (US ITASE) Glaciochemical Data, 520   Version 2, US_ITASE_Core Info-SWE-Density_2013.xlsx, National Snow and Ice Data Center, Boulder, Colorado, USA, 2013.

Medley B., Joughin I., Das S. B., Steig E. J., Conway, H., Gogineni S., Criscitiello A. S., McConnell J. R., Smith B. E., van den Broeke M. R., Lenaerts J. T. M., Bromwich, D. H., and Nicolas J. P.: Airborne-radar and ice-core observations of annual 525   snow accumulation over Thwaites Glacier, West Antarctica confirm the spatiotemporal variability of global and regional atmospheric models, Geophys. Res. Lett., 40, 3649–3654, https://doi.org/doi:10.1002/grl.50706, 2013.

Medley, B., Joughin, I., Smith, B. E., Das, S. B., Steig, E. J., Conway, H., Gogineni, S., Lewis, C., Criscitiello, A. S., Mc-Connell, J. R., van den Broeke, M. R., Lenaerts, J. T. M., Bromwich, D. H., Nicolas, J. P., and Leuschen, C.: Constraining 530   the recent mass balance of Pine Island and Thwaites glaciers, West Antarctica, with airborne observations of snow accumulation, The Cryosphere., 8, 1375–1392, https://doi.org/10.5194/tc-8-1375-2014, 2014.

Medley, B., and Thomas, E. R.: Increased snowfall over the Antarctic Ice Sheet mitigated 20th century sea-level rise, Nat. Clim. Change, 9, 34–39. https://doi.org/10.1038/s41558-018-0356-x, 2019.

535

Montgomery, L., Koenig, L., and Alexander, P. The SUMup dataset: compiled measurements of surface mass balance components over ice sheets and sea ice with analysis over Greenland, Earth Syst. Sci. Data, 10, 1959–1985, https://doi.org/10.5194/essd-10-1959-2018, 2018.

540

Monaghan, A. J., Bromwich, D. H., Fogt, R. L., Wang, S.-H., Mayewski, P. A., Dixon, D. A., Ekaykin, A., Frezzotti, M., Goodwin, I., Isaksson, E., Kaspari, S. D., Morgan, V. I., Oerter, H., Van Ommen, T. D., Van der Veen, C. J., and Wen, J.: Insignificant Change in Antarctic Snowfall Since the International Geophysical Year, Science, 313, 827–831, https://doi.org/10.1126/science.1128243, 2006a.

545

Monaghan, A. J., Bromwich, D. H., and Wang, S-H.: Recent trends in Antarctic snow accumulation from Polar MM5 simulations, Phil. Trans. R. Soc. A., 364, 1683–1708, https://doi.org/10.1098/rsta.2006.1795, 2006b.

Motoyma, H., Furukawa, T., Fujita, S., Shinbori, K., Tanaka, Y., Li, Y., Chung, J-W., Nakazawa, F., Fukui, K., Enomoto, H., Sugiyama, S., Asano, H., Takeda, Y., Hirabayashi, M., Nishimura, D., Masunaga, T., Kuramoto, T., Kobashi, T., Kusaka, R., 550 Kinase, T., Ikeda, C., Suzuki, T., Ohno, H., Hoshina, Y., Hayakawa, Y., and Kameda T.: Glaciological Data Collected by the 48th–54th Japanese Antarctic Research Expeditions during 2007–2013, JARE Data Rep., 341, Glaciology 35, 2015.

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML76S05_11, https://doi.org/10.1594/PANGAEA.708113, 2008a.

555

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML77S05_12, https://doi.org/10.1594/PANGAEA.708114, 2008b.

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML78S05_13, 560 https://doi.org/10.1594/PANGAEA.708115,2008c.

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML79S05_14, https://doi.org/10.1594/PANGAEA.708116, 2008d.

565 Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML80S05_15, https://doi.org/10.1594/PANGAEA.708117, 2008e.

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML81S05_16, https://doi.org/10.1594/PANGAEA.708118, 2008f.

570

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML82S05_17, https://doi.org/10.1594/PANGAEA.708119, 2008g.

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML83S05_18,
575   https://doi.org/10.1594/PANGAEA.708120, 2008h.

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML84S05_19,
https://doi.org/10.1594/PANGAEA.708121, 2008i.

580   Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML85S05_20,
https://doi.org/10.1594/PANGAEA.708122, 2008j.

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML86S05_21,
https://doi.org/10.1594/PANGAEA.708123, 2008k.
585

Oerter, H.: Annual means of d18O and accumulation rates of snow pit DML87S05_22,
https://doi.org/10.1594/PANGAEA.708124, 2008l.

Palerme, C., Kay, J. E., Genthon, C., L'Ecuyer, T., Wood, N. B., and Claud, C.: How much snow falls on the Antarctic ice
590   sheet? The Cryosphere., 8, 1577–1587, https://doi.org/ 10.5194/tc-8-1577-2014, 2014.

Rignot, E., Mouginot, J., Scheuchl, B., van den Broeke, M., van Wessem, M. J., and Morlighem, M.: Four decades of
Antarctic Ice Sheet mass balance from 1979–2017. PNAS, 116:1095–1103, https://doi.org/10.1073/pnas.1812883116, 2019.
Rignot, E., Velicogna, I., van den Broeke, M. R., Monaghan, A. J., and Lenaerts, J. T.: Acceleration of the contribution of
595   the Greenland and Antarctic ice sheets to sea level rise, Geophys. Res. Lett., 38, L05503,
https://doi.org/10.1029/2011GL046583, 2011.

Reijmer, C. H., Greuell, W., and Oerlemans, J.: The annual cycle of meteorological variables and the surface energy balance
on Berkner Island, Antarctica, Ann. Glaciol., 29, 49-54, 1999.
600

Reijmer, C. H. and Van den Broeke, M. R.: Temporal and spatial variability of the surface mass balance in Dronning Maud
Land, Antarctica, as derived from automatic weather stations, J. Glaciol., 49, 167, 512–520,
https://doi.org/10.3189/172756503781830494, 2003.

605   Shepherd, A., Ivins, E., Rignot, E., Smith, B., van den Broeke, M., Velicogna, I., Whitehouse, P., Briggs, K., Joughin, I.,
Krinner, G., Nowicki, S., Payne, T., Scambos, T., Schlegel, N., A, G., Agosta, C., Ahlstrøm, A., Babonis, G., Barletta, V.,
Blazquez, A., Bonin, J., Csatho, B., Cullather, R., Felikson, D., Fettweis, X., Forsberg, R., Gallee, H., Gardner, A., Gilbert,

L., Groh, A., Gunter, B., Hanna, E., Harig, C., Helm, V., Horvath, A., Horwath, M., Khan, S., Kjeldsen, K. K., Konrad, H., Langen, P., Lecavalier, B., Loomis, B., Luthcke, S., McMillan, M., Melini, D., Mernild, S., Mohajerani, Y., Moore, P.,
610 Mouginot, J., Muir, A., Nagler, T., Nield, G., Nilsson, J., Noel, B., Otosaka, I., Peltier, R., Pie, N., Rietbroek, R., Rott, H., Sandberg-Sørensen, L., Sasgen, I., Save, H., Schrama, E., Schroder, L., Seo, K.-W., Simonsen, S., Slater, T., Spada, G., Sutterley, T., Talpe, M., Tarasov, L., van de Berg, W. J., van der Wal, W., van Wessem, M., Vishwakarma, B. D., Wiese, D., Wouters, B., Wu, X., and Zwally, J.: Mass balance of the Antarctic Ice Sheet from 1992 to 2017, Nature, 558, 219–222, https://doi.org/10.1038/s41586-018-0179-y, 2018.

615

Sheperd, A., Ivins, E. R., A. G., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., Bromwich, D. H., Forsberg, R., Galin, N., Horwath, M., Jacobs, S., Joughin, I., King, M. A., Lenaerts, J. T. M., Li, J., Ligtenberg, S. R. M., Luckman, A., Luthcke, S. B., McMillan, M., Meister, R., Milne, G., Mouginot, J., Muir, A., Nicolas, J. P., Paden, J., Payne, A. J., Pritchard, H., Rignot, E., Rott, H., Sørensen, L. S., Scambos, T. A., Scheuchl, B., Schrama, E. J. O., Smith, B., Sundal, A. V., van
620 Angelen, J. H., van de Berg, W. J., van den Broeke, M. R., Vaughan, D. G., Velicogna, I., Wahr, J., Whitehouse, P. L., Wingham, D. J., Yi, D., Young, D., and Zwally, H. J.: A reconciled estimate of ice-sheet mass balance, Science, 338, 1183–1189, https://doi.org/10.1126/science.1228102, 2012.

Thomas, E. R. and Bracegirdle, T. J.: Precipitation pathways for five new ice core sites in Ellsworth Land, West Antarctica,
625 Clim. Dyn., 44, 2067–2078, https://doi.org/10.1007/s00382-014-2213-6, 2015.

Thiery, W., Gorodetskaya, I. V., Bintanja, R., Van Lipzig, N. P. M., Van den Broeke, M. R., Reijmer, C. H., and Kuipers Munneke, P.: Surface and snowdrift sublimation at Princess Elisabeth station, East Antarctica, The Cryosphere., 6, 841–857, https://doi.org/10.5194/tc-6-841-2012, 2012.

630

Thomas, E. R., van Wessem, J. M., Roberts, J., Isaksson, E., Schlosser, E., Fudge, T. J., Vallelonga, P., Medley, B., Lenaerts, J., Bertler, N., Van den Broeke, M. R., Dixon, D. A., Frezzotti, M., Stenni, B., Curran, M., and Ekaykin, A. A.: Regional Antarctic snow accumulation over the past 1000 years, Clim. Past., 13, 1491–1513, https://doi.org/10.5194/cp-13-1491-2017, 2017.

635

Van den Broeke, M. R., Reijmer, C. H., and van de Wal, R. S. W.: A study of the surface mass balance in Dronning Maud Land, Antarctica, using automatic weather station. J. Glaciol., 50, 565–582. https://doi.org/10.3189/172756504781829756, 2004.

640   Van de Berg, W. J., Van den Broeke, M. R., Reijmer, C. H., and Van Meijgaard, E.: Reassessment of the Antarctic surface
mass balance using calibrated output of a regional atmospheric climate model, J. Geophys. Res., 111, D11104,
doi:10.1029/2005JD006495, 2006.

Vaughan, D. G. and Russell, J.: Compilation of surface mass balance measurements in Antarctica, Internal Rep., ES4 , 56,
645   1997.

Vaughan, D. G., Bamber, J. L., Giovinetto, M., Russell, J., and Cooper, A. P. R.: Reassessment of net surface mass balance
in Antarctica, J. Climate, 12, 933–946, https://doi.org/10.1175/1520-0442(1999)012<0933:RONSMB>2.0.CO;2, 1999.

650   Vaughan, D. G., Anderson, P. S., King, J. C., Mann, G. W., Mobbs, S. D., and Ladkin R. S.: Imaging of firn isochrones
across an Antarctic ice rise and implications for patterns of snow accumulation rate, J. Glaciol., 50, 413–418, 2004.

Van Wessem, J. M., van de Berg, W. J., Noël, B. P. Y., van Meijgaard, E., Amory, C., Birnbaum, G., Jakobs, C. L., Krüger,
K., Lenaerts, J. T. M., Lhermitte, S., Ligtenberg, S. R. M., Medley, B., Reijmer, C. H., van Tricht, K., Trusel, L. D., van Ulft,
655   L. H., Wouters, B., Wuite, J., and van den Broeke, M. R.: Modelling the climate and surface mass balance of polar ice sheets
using RACMO2 –Part 2: Antarctica (1979–2016), The Cryosphere., 12, 1479–1498, https://doi.org/10.5194/tc-12-1479-2018,
2018.

Wang, Y., Ding, M., van Wessem, J., Schlosser, E., Altnau, S., van den Broeke, M. R., Lenaerts, J. T. M., Thomas, E. R.,
660   Isaksson, E., Wang, J., and Sun, W.: A comparison of Antarctic Ice Sheet surface mass balance from atmospheric climate
models and in situ observations, J. Climate., 29, 5317-5337, 2016

Wang, Y., Ding, M., Reijmer C., Smeets P., Hou, S., and Xiao, C.: A comprehensive dataset of surface mass balance field
observations over the Antarctic Ice Sheet version 1.0. A Big Earth Data Platform for Three Poles, 2021.
665
Wang, Y., Huai, B., Thomas, E. R., van den Broeke, M. R., van Wessem, J. M., and Schlosser, E.: A new 200-year spatial
reconstruction of West Antarctic surface mass balance, J. Geophys. Res.: Atmos., 124, 5282-5295,
https://doi.org/10.1029/2018JD029601, 2019.

670   Wang, Y., Hou, S., Sun, W., Lenaerts, J. T. M., van den Broeke, M. R., and van Wessem, J. M.: Recent surface mass balance
from Syowa Station to Dome F, East Antarctica: comparison of field observations, atmospheric reanalyses, and a regional
atmospheric climate model, Clim. Dyn., 45, 2885-2899. https://doi.org/ 10.1007/s00382-015-2512-6, 2015.

Wang, Y., Hou, S., Ding, M., and Sun W.: On the performance of twentieth century reanalysis products for Antarctic snow
675   accumulation, Clim. Dyn., 54, 435-455, https://doi.org/10.1007/s00382-019-05008-4, 2020.

Winski, D. A., Fudge, T. J., Ferris, D. G., Osterberg, E. C., Fegyveresi, J. M., Cole-Dai, J., Thundercloud, Z., Cox, T. S.,
Kreutz, K. J., Ortman, N., Buizert, C., Epifanio, J., Brook, E. J., Beaudette, R., Severinghaus, J., Sowers, T., Steig, E. J.,
Kahle, E. C., Jones, T. R., Morris, V., Aydin, M., Nicewonger, M. R., Casey, K. A., Alley, R. B., Waddington, E. D., Iverson,
680   N. A., Dunbar, N. W., Bay, R. C., Souney, J. M., Sigl, M., and McConnell, J. R.: The SP19 chronology for the South Pole
Ice Core – Part 1: volcanic matching and annual layer counting, Clim. Past., 15, 1793–1808, https://doi.org/10.5194/cp-15-
1793-2019, 2019.

Wouters, B., Martin-Español, A., Helm, V., Flament, T., van Wessem, J. M., Ligtenberg, S. R. M., van de Broeke, M. R.,
685   and Bamber, J. L.: Dynamic thinning of glaciers on the Southern Antarctic Peninsula, Science, 348, 899–903.
https://doi.org/10.1126/science.aaa5727, 2015.

Xiao, C., Ren, J., Qin, D. H., Li, Z. Q., Sun, W. Z., and Allison, I.: Complexity of the climatic regime over the Lambert
Glacier basin of the East Antarctic Ice Sheet: Firn core evidences, J. Glaciol., 47, 160–163,
690   https://doi.org/10.3189/172756501781832539, 2001.

Zhang Y., Wang Y., Huai B., Ding M., and Sun W.: Skill of the two 20th century reanalyses in representing Antarctic near-
surface air temperature, Int J Climatol., 38, 11, 4225–4238. https://doi.org/10.1002/joc.5563, 2018
.
695

700

705

Table 1 Brief description of metadata fields used in the Antarctic SMB observations database

Column	Name of field in database	Description	Format	Unit
Site name	Geo_siteName	Name of the site	Number Code	Number Code
Dataset ID	DatasetName	Specific identifier assigned to all SMB records from a given site and publication	Number Code	Number Code
Latitude	Geo_latidute	Latitude of the site	WGS84	Decimal degrees (-90° to 90°
Longitude	Geo_longitude	Longitude of the site	WGS84	Decimal degrees (-180° to 180°)
Elevation	Geo_elevation	Elevation of the site	Height above the EGM geoid	m above sea level
Variable name	SMB	SMB in millimetre of water equivalent per year	Mass loss is defined as negative	$kg\ m^{-2}\ yr^{-1}$
Method	Method	How each measurement was collected	_____	_____
Starting date	MinYear	Starting date of measurement, i.e., minimum (oldest) year of each SMB record	Number	Year
Ending date	MaxYear	Ending date of measurement, i.e., maximum (more recent) year of each SMB record	Number	Year
Citation	Citation	Citation for the first publication presenting the SMB record.	_____	_____

710

715
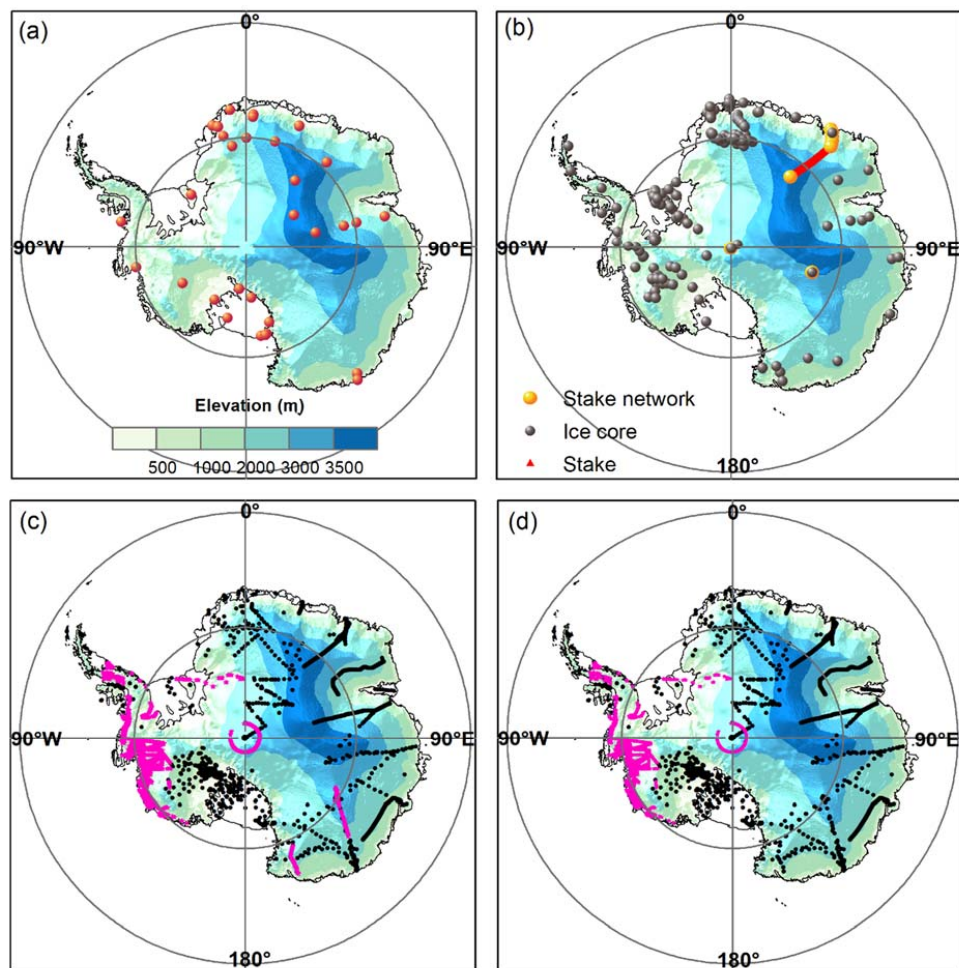
720 **Figure 1: The comprehensive dataset of Antarctic SMB observations. (a) Spatial distribution of available AWS observations; (b) Locations of available annual resolved SMB observations. (c) Locations of available multi-year averaged SMB field data. Purple five-pointed stars standard for GPR measurements. Black points represent reliable SMB determined by stake/stake farms, ice cores/snow pits. (d) Location of multi-year averaged SMB data only during the second half 20th century selected for model validation.**

725
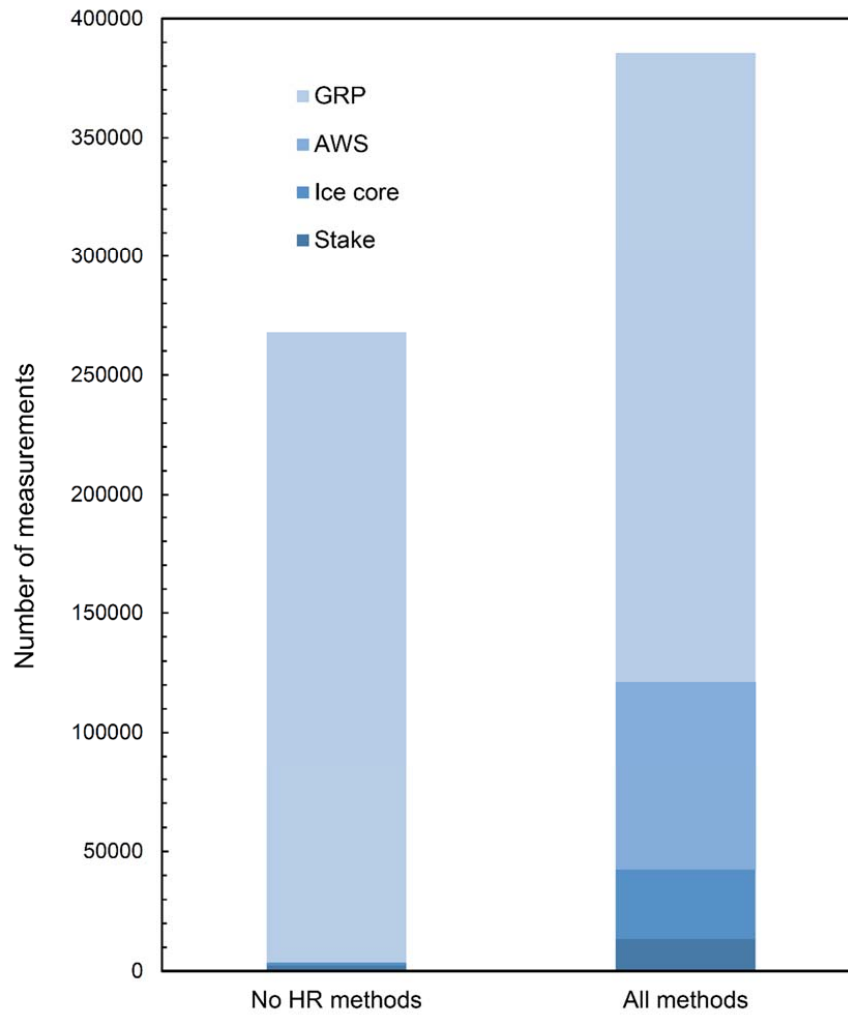
Figure 2: Bar charts indicating the number of the different types of measurement techniques in the SMB observation dataset. The left bar demonstrates the distribution of approaches with the exclusion of high-resolution snow accumulation measurements, and the distribution of all measurement techniques is shown in the right bar.

730
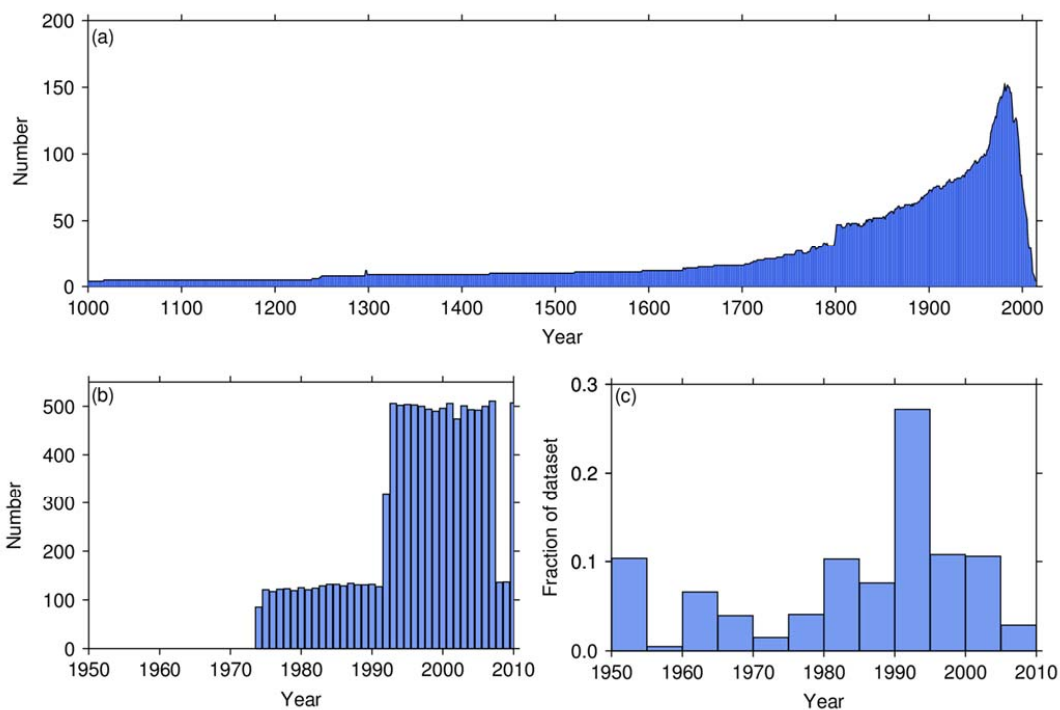
Figure 3: (a) Availability of records excluding the stake measurements along the traverse from Syowa Station to Dome F in the annual resolved SMB sub-database over time during the past 1000 years. (b)Time coverage of the stake measurements along the traverse from Syowa Station to Dome F. (c) Histograms indicating the date taken of the multi-year averaged SMB subdaset only including ice core and stake measurements.
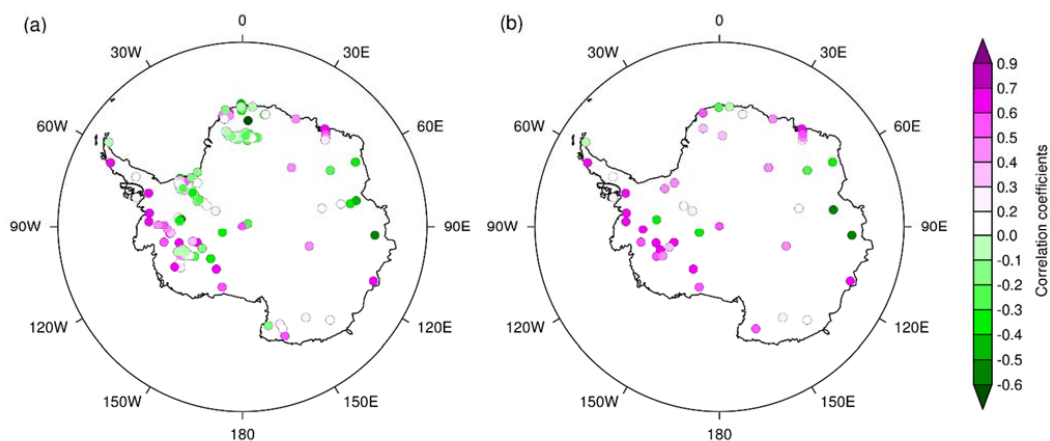
735

740

745
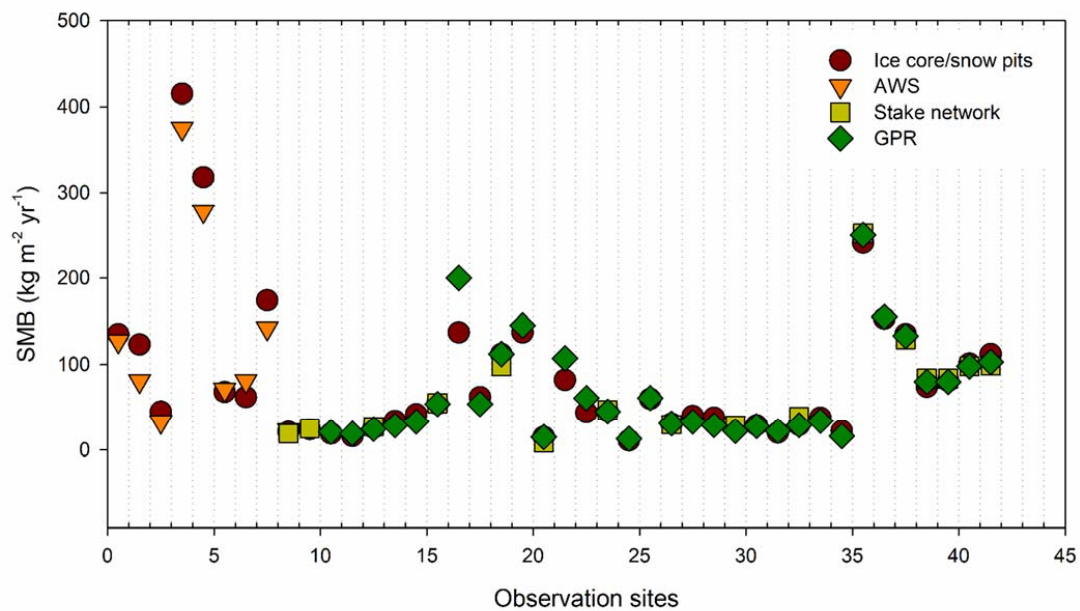
Figure 4: Inter-comparison between different types of SMB measurements including AWS, snow pit/ice core and GPR at 42 locations.

750

755

760

Figure 5: Spatial distribution of the correlation coefficients (a) between annual resolved SMB observation and ERA5 simulations for their overlapping period, and (b) between averaged observed time series in the same location/region and the corresponding simulations from ERA5.
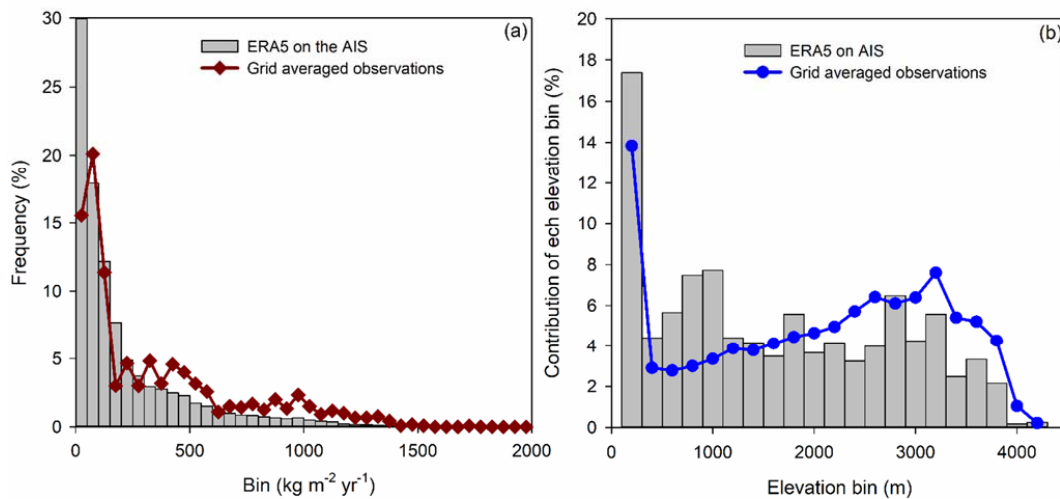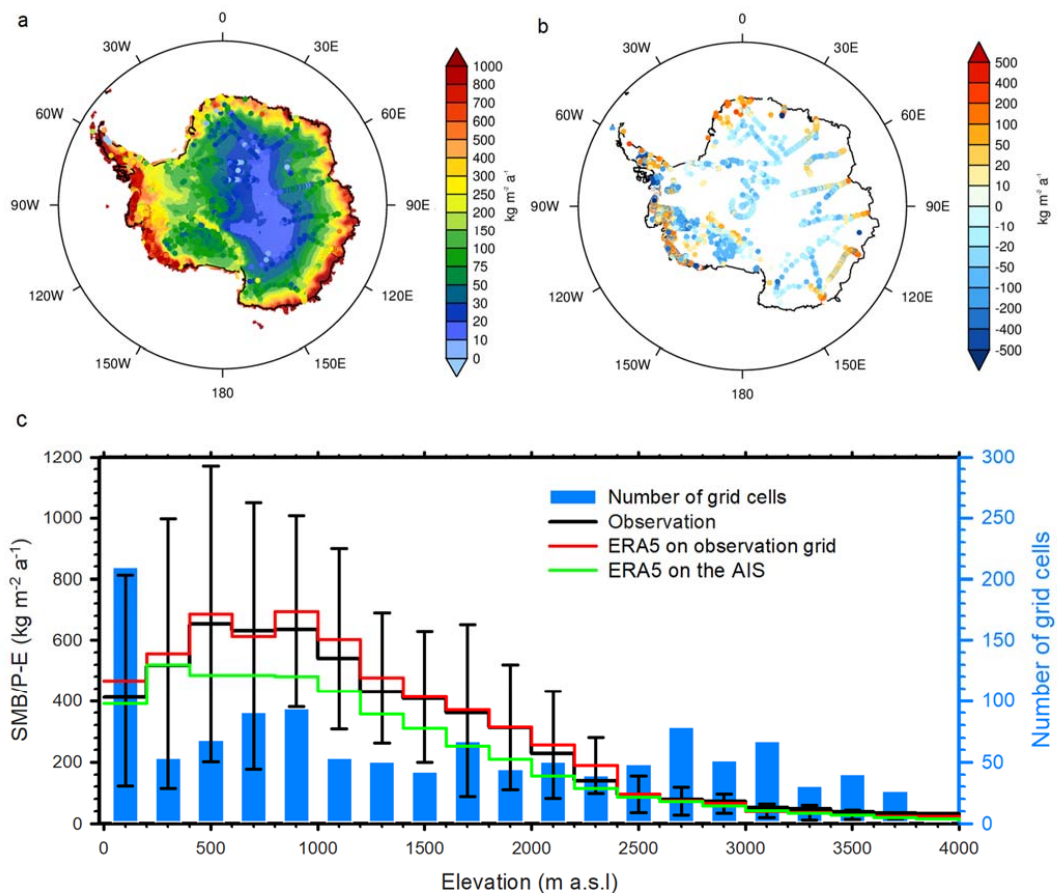
785

Figure 6 (a) Relative frequency of ERA5 P-E field over the AIS and gridded averaged records from the multi-year averaged SMB

790 **subdatabase, with a bin range of 50kg m^{-2} yr^{-1}. (b) The contribution of elevation bin for ERA5 grid cells containing measurements,**

**and entire elevation range to the AIS. The 200m elevation bins are used.**

795

800

805

Figure 7: (a) Spatial distribution of ERA5 mean precipitation minus evaporation (approximated as SMB) for the period 1979–2018, and multi-year averaged SMB measurements. (b) ERA5 minus observed SMB on the ERA5 grid cells, (c) Multi-year averaged observations and ERA5 simulations, binned in 200 m elevation intervals. The number of ERA5 grid cells with in situ measurements in each elevation bin is shown by the blue line (right axis).

810
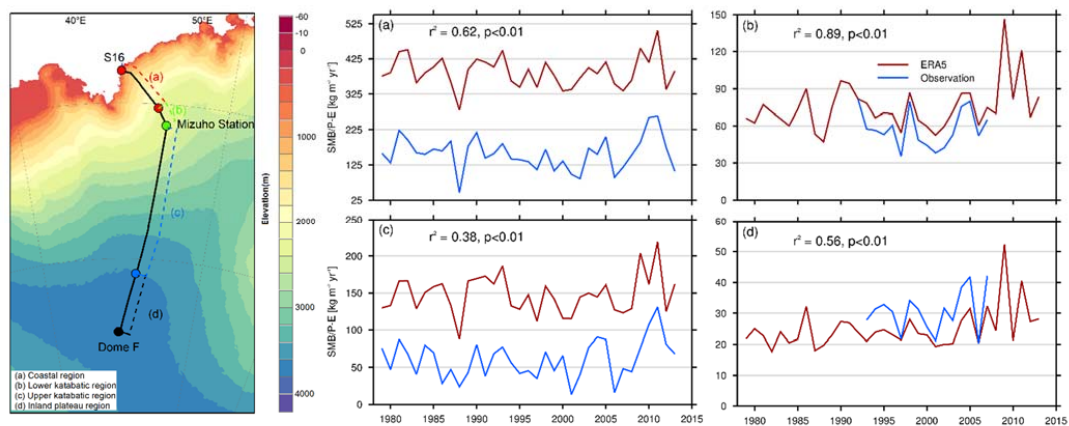
Figure 8: The left map showing the locations of stake measurements along the traverse between Syowa Station and Dome F, and the regional boundaries. The right four charts showing the comparison of the inter-annual variability in spatially-averaged stake measurements and snow accumulation simulated by ERA5 for (a) the coastal region, (b) lower katabatic region, (c) upper katabatic region, and (d) inland plateau region.