The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)

3	Ronald Gelaro ¹ , Will McCarty ¹ , Max J. Suárez ² , Ricardo Todling ¹ , Andrea Molod ¹ ,
4	Lawrence Takacs ³ , Cynthia Randles ^{4†} , Anton Darmenov ¹ , Michael G. Bosilovich ¹ , Rolf
5	Reichle ¹ , Krzysztof Wargan ³ , Lawrence Coy ³ , Richard Cullather ⁵ , Clara Draper ² ,
6	Santha Akella ³ , Virginie Buchard ² , Austin Conaty ³ , Arlindo da Silva ¹ , Wei Gu ³ ,
7	Gi-Kong Kim ¹ , Randal Koster ¹ , Robert Lucchesi ³ , Dagmar Merkova ^{3‡} , Jon Eric
8	Nielsen ³ , Gary Partyka ³ , Steven Pawson ¹ , William Putman ¹ , Michele Rienecker ¹ ,
9	Siegfried D. Schubert ¹ , Meta Sienkiewicz ³ , and Bin Zhao ⁶
10	¹ Global Modeling and Assimilation Office, NASA Goddard Space Flight Center,
11	Greenbelt, MD
12	² Universities Space Research Association, Columbia, MD
13	³ Science Systems and Applications, Inc., Lanham, MD
14	⁴ Morgan State University, Baltimore, MD
15	⁵ Earth System Science Interdisciplinary Center, College Park, MD
16	⁶ Science Applications International Corporation, Beltsville, MD
17	[†] Current affiliation: ExxonMobil Corporate Strategic Research, Annandale, NJ
18	[‡] Current affiliation: I.M. System Group, Inc., Rockville, MD
19	Revised 23 March 2017
20	Accepted for publication in <i>Journal of Climate</i> on 29 March 2017
21	MERRA-2 Special Collection

22 Corresponding author: Ronald Gelaro, Global Modeling and Assimilation Office, NASA

²³ Goddard Space Flight Center, Greenbelt, MD 20771, USA. Email: ron.gelaro@nasa.gov

24 Abstract

The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-25 2) is the latest atmospheric reanalysis of the modern satellite era produced by NASA's 26 Global Modeling and Assimilation Office (GMAO). MERRA-2 assimilates observation 27 types not available to its predecessor, MERRA, and includes updates to the Goddard 28 Earth Observing System (GEOS) model and analysis scheme so as to provide a viable 29 ongoing climate analysis beyond MERRA's terminus. While addressing known limita-30 tions of MERRA, MERRA-2 is also intended to be a development milestone for a future 31 integrated Earth system analysis (IESA) currently under development at GMAO. This 32 paper provides an overview of the MERRA-2 system and various performance metrics. 33 Among the advances in MERRA-2 relevant to IESA are the assimilation of aerosol 34 observations, several improvements to the representation of the stratosphere including 35 ozone, and improved representations of cryospheric processes. Other improvements in 36 the quality of MERRA-2 compared with MERRA include the reduction of some spuri-37 ous trends and jumps related to changes in the observing system, and reduced biases 38 and imbalances in aspects of the water cycle. Remaining deficiencies are also identified. 39 Production of MERRA-2 began in June 2014 in four processing streams, and converged 40 to a single near-real time stream in mid 2015. MERRA-2 products are accessible online 41 through the NASA Goddard Earth Sciences Data Information Services Center (GES 42 DISC). 43

44 1. Introduction

Reanalysis is the process whereby an unchanging data assimilation system is used to 45 provide a consistent reprocessing of meteorological observations, typically spanning an 46 extended segment of the historical data record. The process relies on an underlying 47 forecast model to combine disparate observations in a physically consistent manner, en-48 abling production of gridded data sets for a broad range of variables including ones that 49 are sparsely or not directly observed. As such, and with appropriate consideration of the 50 inherent uncertainties, reanalysis products have not only become a staple of the atmo-51 spheric research community, but are used increasingly for climate monitoring as well as 52 for business applications in, for example, energy and agriculture. Recent reanalyses from 53 the National Oceanic and Atmospheric Administration/National Centers for Environ-54 mental Prediction (NOAA/NCEP), the European Centre for Medium-Range Weather 55 Forecasts (ECMWF), the National Aeronautics and Space Administration/Global Mod-56 eling and Assimilation Office (NASA/GMAO), and the Japan Meteorological Agency 57 (JMA) provide a rich ensemble of climate data products beginning more or less with 58 the period of regular conventional and satellite observations in the mid to late twentieth 59 century (Saha et al. 2010; Dee et al. 2011; Rienecker et al. 2011; Kobavashi et al. 2015). 60 However, there have also been successful efforts to extend atmospheric reanalyses back 61 to the late nineteenth and early twentieth centuries using only surface pressure obser-62 vations (Compo et al. 2011) or surface and mean sea level pressure observations plus 63 surface marine winds (Poli et al. 2013). As noted by Dee et al. (2011), these century-64 long reanalyses have also sparked remarkable data recovery and digitization efforts by 65 various groups around the world. 66

The GMAO's reanalysis development effort began (under its predecessor organization,
the Data Assimilation Office) with the production of the Goddard Earth Observing

System, version 1 (GEOS-1) reanalysis (Schubert et al. 1993), but advanced significantly 69 with the more recent production of the Modern-Era Retrospective Analysis for Research 70 and Applications (MERRA, Rienecker et al. 2011). MERRA encompassed the period 71 1979–2016 and was undertaken with two primary objectives: to place NASA's Earth 72 Observing System (EOS) satellite observations in a climate context and to improve 73 the representation of the atmospheric branch of the hydrological cycle compared with 74 previous reanalyses. MERRA succeeded in meeting these objectives overall and was 75 found to be of comparable quality to contemporaneous reanalyses produced by NCEP 76 and ECMWF (e.g., Decker et al. 2011). However, it also suffered from a number 77 of known, but not necessarily unique, deficiencies. These include unphysical jumps 78 and trends in precipitation in response to changes in the observing system, biases and 79 imbalances in certain atmospheric and land surface hydrological quantities, and a poor 80 representation of the upper stratosphere (e.g., Bosilovich et al. 2011; Robertson et al. 81 2011; Reichle et al. 2011; Rienecker et al. 2011). In addition, the long-term viability 82 of MERRA was limited by system constraints that precluded the incorporation of new 83 satellite data sources beyond NOAA-18, which launched in 2005. At the time of its 84 termination in March 2016, MERRA was at risk of suffering a significant degradation 85 in quality were certain observing platforms to fail, including, for example, EOS Aqua, 86 which was already well beyond its designed lifetime and provided MERRA with its only 87 sources of hyperspectral infrared and afternoon-orbit microwave radiances. 88

The Modern Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) was undertaken to provide a timely replacement for MERRA and to sustain GMAO's commitment to having an ongoing near-real-time climate analysis. MERRA-2 is intended as an intermediate reanalysis; one that leverages recent developments at GMAO in modeling and data assimilation to address some of the known limitations of MERRA, but also provides a stepping stone to GMAO's longer term goal of developing an in-

tegrated Earth system analysis (IESA) capability that couples assimilation systems for 95 the atmosphere, ocean, land and chemistry. Toward the latter goal MERRA-2 includes 96 aerosol data assimilation, thereby providing a multi-decadal reanalysis in which aerosol 97 and meteorological observations are jointly assimilated within a global data assimilation 98 system. Other new developments in MERRA-2 relevant to IESA focus on aspects of 99 the cryosphere and stratosphere, including the representation of ozone, and on the use 100 of precipitation observations to force the land surface. At the same time, basic aspects 101 of the MERRA-2 system, such as the variational analysis algorithm and observation 102 handling, are largely unchanged since MERRA. Also unchanged is the preparation of 103 most conventional data sources used originally in MERRA. 104

This paper presents an overview of MERRA-2, including a description of the data as-105 similation system and various measures of performance. Some of these measures focus 106 on difficulties encountered in MERRA while others highlight new capabilities such as 107 the assimilation of aerosol observations. This paper also serves as an introduction to a 108 series of companion papers that provide more detailed analyses of the topics covered in 109 this overview as well as others. For example, a detailed description of the MERRA-2 110 aerosol analysis system and its validation are presented in Randles et al. (2017) and 111 Buchard et al. (2017). Reichle et al. (2017a,b) assess the land surface precipitation 112 and land surface hydrology, while Draper et al. (2017) examine the land surface en-113 ergy budget. Bosilovich et al. (2017) evaluate the global water balance and water cycle 114 variability in MERRA-2. Collow et al. (2016) examine MERRA-2's representation of 115 US summertime extreme precipitation events, and Lim et al. (2017) investigate aspects 116 of major El Niño events. Collow and Miller (2016) examine the radiation budget and 117 cloud radiative effect over the Amazon. Segal-Rosenhemier et al. (2017) examine surface 118 radiative fluxes in polar marginal ice zones. Several papers investigate aspects of the 119 stratosphere in MERRA-2: Wargan et al. (2017) examine the representation of lower 120

stratospheric ozone and the effect of assimilating ozone observations from NASA's Aura
satellite; Coy et al. (2016) examine the representation of the quasi-biennial oscillation
in MERRA-2; and Wargan and Coy (2016) present a case study of the 2009 sudden
stratospheric warming.

Section 2 provides an overview of the MERRA-2 data assimilation system, focusing pri-125 marily on developments since MERRA, including new observation sources. Basic met-126 rics of the assimilation system performance are presented in section 3. The MERRA-2 127 aerosol analysis is described in section 4, along with sample results and validation statis-128 tics. Section 5 examines global and regional aspects of the representation of precipitation 129 in MERRA-2, focusing on areas of difficulty in MERRA. Stratospheric processes and 130 the representation of ozone are discussed in section 6. Section 7 addresses the represen-131 tation of the cryosphere in MERRA-2, with focus on glaciated land surface processes. 132 Section 8 provides information about MERRA-2 products and how they can be accessed. 133 It is noted here that each MERRA-2 data collection has its own digital object identifier 134 (DOI) number, so data used in scientific publications can be cited exactly. Most of the 135 results shown for MERRA-2 in this paper are derived from these collections, which are 136 individually cited in the corresponding figure captions. Finally, a brief summary and 137 perspective on future work are presented in section 9. A list of acronyms is given in the 138 Appendix. 139

¹⁴⁰ 2. MERRA-2 system description

MERRA-2 is produced with version 5.12.4 of the Goddard Earth Observing System (GEOS-5.12.4) atmospheric data assimilation system. The key components of the system are the GEOS atmospheric model (Rienecker et al. 2008; Molod et al. 2015) and the Gridpoint Statistical Interpolation (GSI) analysis scheme (Wu et al. 2002; Kleist et

al. 2009b). The model includes the finite-volume dynamical core of Putman and Lin 145 (2007), which uses a cubed sphere horizontal discretization at an approximate resolution 146 of $0.5^{\circ} \times 0.625^{\circ}$ and 72 hybrid-eta levels from the surface to 0.01 hPa. The analysis is 147 computed on a latitude-longitude grid at the same spatial resolution as the atmospheric 148 model using a three-dimensional variational (3DVAR) algorithm based on the GSI with a 149 6-h update cycle and the so-called first-guess-at-appropriate-time (FGAT) procedure for 150 computing temporally accurate observation-minus-background departures. The analysis 151 is applied as a correction to the background state using an incremental analysis update 152 (IAU) procedure (Bloom et al. 1996). 153

The MERRA-2 system has many of the same basic features as the MERRA system 154 (GEOS-5.2.0) described in Rienecker et al. (2011) but includes a number of important 155 updates. An overview of these updates is provided here, with additional details provided 156 in companion publications as cited. Unless otherwise stated, other aspects of the system 157 configuration and preparation of the input data are as described in Rienecker et al. 158 (2011). The updates discussed here include changes to the forecast model (section 2a), 159 the analysis algorithm (section 2b), the observing system (section 2c), the radiance 160 assimilation (section 2d), the bias correction of aircraft observations (section 2e), the 161 mass conservation and water balance (section 2f), the precipitation used to force the land 162 surface and drive wet aerosol deposition (section 2g), the boundary conditions for sea 163 surface temperature and sea ice concentration (section 2h), and reanalysis production 164 (section 2i). 165

166 a. Forecast model

¹⁶⁷ Since MERRA, the GEOS model has undergone changes to both its dynamical core ¹⁶⁸ and its physical parameterizations. Whereas in MERRA the horizontal discretization of

the model was computed on a latitude-longitude grid, MERRA-2 uses a cubed sphere 169 grid. This allows relatively uniform grid spacing at all latitudes and mitigates the more 170 severe grid spacing singularities that occur on a latitude-longitude grid. Upgrades to 171 the physical parameterization schemes include increased re-evaporation of frozen pre-172 cipitation and cloud condensate, changes to the background gravity wave drag, and 173 an improved relationship between the ocean surface roughness and ocean surface stress 174 (Molod et al. 2015). The MERRA-2 model also includes a Tokioka-type trigger on deep 175 convection as part of the Relaxed Arakawa-Schubert (RAS, Moorthi and Suárez 1992) 176 convective parameterization scheme, which governs the lower limit on the allowable en-177 trainment plumes (Bacmeister and Stephens 2011). A new glaciated land representation 178 and seasonally-varying sea ice albedo have been implemented, leading to improved air 179 temperatures and reduced biases in the net energy flux over these surfaces (Cullather et 180 al. 2014). 181

182 b. Analysis algorithm

The control variable for moisture used in recent versions of GSI and MERRA-2 differs 183 from the one used in MERRA. Whereas MERRA used the so-called pseudo-relative 184 humidity (Dee and da Silva, 2003) defined by the water vapor mixing ratio scaled by 185 its saturation value, MERRA-2 uses the normalized pseudo-relative humidity (Holm 186 2003) defined by the pseudo-relative humidity scaled by its background error standard 187 deviation. The latter has a near Gaussian error distribution, making it more suitable for 188 the minimization procedure employed in the assimilation scheme. Also within the GSI, 189 a tangent linear normal mode constraint (TLNMC, Kleist et al. 2009a) is applied during 190 the minimization procedure to control noise and improve the overall use of observations. 191 The background error statistics used in the GSI have been updated as well in MERRA-192

2. As in MERRA, the statistics are estimated using the 'NMC' method (Parrish and Derber, 1992) by calculating variances and covariances from the differences between 24-h and 48-h forecasts, but from a more recent version of GEOS. Compared with the MERRA system, the background error statistics for the MERRA-2 system exhibit generally smaller standard deviations for most variables, but both larger and smaller correlation length scales depending on the variable, latitude and vertical level.

199 c. Observing system

MERRA included no new satellite observation sources after the introduction of NOAA-18 in 2005. MERRA-2, in contrast, includes numerous additional satellite observations both before and after this time. The complete set of input observations assimilated in MERRA-2 is summarized in Table 1, while a detailed description of their use is provided in McCarty et al. (2016). Additions to the MERRA-2 observing system compared with MERRA include:

- Atmospheric motion vectors from the Advanced Very High Resolution Radiometer (AVHRR);
- Surface wind speeds from the Special Sensor Microwave Imager/Sounder (SSMIS);
- Surface wind vectors from the Meteorological Operational Satellite-A (Metop-A)
 Advanced Scatterometer (ASCAT) and WindSat;
- Temperature and ozone profiles from the EOS Aura Microwave Limb Sounder (MLS);
- Total column ozone from the EOS Aura Ozone Monitoring Instrument (OMI);
- Bending angle from Global Positioning System radio occultations (GPSRO);
 - 7

- Microwave and infrared sounding radiances from the Advanced TIROS Operational Vertical Sounder (ATOVS) on NOAA-19, Metop-A and -B;
 Microwave sounding radiances from the Advanced Technology Microwave Sounder (ATMS) on the Suomi National Polar-orbiting Partnership (SNPP);
 Hyperspectral infrared radiances from the Infrared Atmospheric Sounding Interferometer (IASI) on Metop-A and -B, and from the Cross-track Infrared Sounder (CrIS) on SNPP;
- Geostationary radiances from the Meteosat Second Generation (MSG) Spinning
 Enhanced Visible Infrared Imager (SEVIRI) and Geostationary Operational En vironmental Satellites (GOES-11, -13 and -15).

Time series of the various types of observations assimilated in MERRA and MERRA-2 225 are shown in Figure 1. The number of assimilated observations in MERRA-2 grows 226 from approximately two million per 6-h cycle in 2002 to almost five million in 2015, 227 while MERRA assimilates approximately 1.5 million observations per 6-h cycle from 228 2002 onward. The GSI in MERRA-2 is also capable of assimilating microwave and hy-229 perspectral infrared radiances from planned future satellites including Metop-C and the 230 Joint Polar Satellite System (JPSS). The temporary spike in the number of QuikSCAT 231 data assimilated in MERRA-2 in late 2000 is due to an error in preprocessing which 232 led to observations beyond the mid-swath "sweet spot" being used in the analysis. This 233 has no discernible impact on the quality of the analyzed fields or on the use of other 234 observations in the assimilation system. 235

MERRA-2 also assimilates reprocessed versions of some of the same satellite observation
 types used in MERRA. In MERRA-2, Remote Sensing Systems version 7 (RSS v7)
 recalibrated radiances and retrieved surface wind speeds from the Defense Meteorological

Satellite Program (DMSP) Special Sensor Microwave Imager (SSM/I) are used, whereas
MERRA used RSS v6. The use of retrieved ozone from the Solar Backscatter Ultraviolet
Radiometer (SBUV) also differs, with MERRA-2 assimilating version 8.6 on 21 layers
from 1980 thru 2004 before switching to OMI and MLS in October 2004. In contrast,
MERRA used SBUV version 8 throughout, in a form degraded from its original 21 layers
to 12.

245 d. Radiance assimilation

Radiative transfer calculations necessary for the assimilation of satellite radiances in 246 MERRA-2 are performed using the Community Radiative Transfer Model (CRTM, Han 247 et al. 2006, Chen et al. 2008). MERRA-2 uses version 2.1.3 of the CRTM for assimilation 248 of all satellite radiances, whereas MERRA used a prototype version of the CRTM for 249 all radiances except those from the Stratospheric Sounding Unit (SSU), for which the 250 Goddard Laboratory for Atmospheres TOVS forward model (GLATOVS, Susskind et 251 al. 1983) was used. Differences between the prototype and version 2.1.3 of the CRTM 252 are too numerous to mention here, but a detailed description of the latter can be found 253 in Liu and Boukabara (2014). 254

The actively assimilated channels for each satellite sensor type in MERRA-2 are sum-255 marized in Table 2. Microwave temperature sounding channels with strong surface 256 sensitivity—so-called window channels—are not assimilated in MERRA-2, in part be-257 cause of the strong sensitivity of global precipitation and humidity to these data found in 258 MERRA (Robertson et al. 2011). These include channels 1–3 and 15 on the Advanced 259 Microwave Sounding Unit-A (AMSU-A), channels 1–4 and 16 on ATMS, and channel 260 1 on the Microwave Sounding Unit (MSU). For microwave humidity sounders includ-261 ing the the Advanced Microwave Sounding Unit-B (AMSU-B) and Microwave Humid-262

ity Sounder (MHS), window channels are actively assimilated along with the sounding
channels. For heritage infrared sounders, channels 13–15 on the High-resolution Infrared
Radiation Sounder (HIRS) were assimilated in MERRA but are excluded in MERRA-2.
The channel selections for hyper-spectral infrared sounders and performance assessments
for selected instruments are provided in McCarty et al. (2016).

Like MERRA, MERRA-2 uses an automated bias correction scheme for the assimilation 268 of most satellite radiance observations. Bias estimates for individual sensor channels 269 are represented by a small number of predictors which can depend on the atmospheric 270 state, the radiative transfer model, and the sensor characteristics. Air-mass- and viewing 271 angle-dependent biases are estimated using a variational scheme in which the predictor 272 coefficients are updated as part of the control vector used to minimize the analysis cost 273 function (Derber and Wu, 1998). Satellite scan-position-dependent bias is estimated 274 directly as an exponential moving average filter of the observation-minus-background 275 departures for brightness temperature. For both the variational and scan-position predic-276 tors, initial values of the coefficients for MERRA-2 were derived from GEOS operations 277 and other long production runs using system versions similar to that used for MERRA-278 2. In the few cases where no recent coefficient information was available, initial values 279 were derived from MERRA. Note that no bias correction is applied to a small number of 280 sensor channels that peak in the upper stratosphere, including channel 14 on AMSU-A, 281 channel 15 on ATMS, and channel 3 on SSU. This is done to prevent the variational 282 bias correction scheme—which is formulated to remove systematic discrepancies between 283 the observations and the background state irrespective of the source—from making er-284 roneous adjustments to the observations at levels where model biases are known to be 285 large. 286

287 e. Bias correction of aircraft temperature observations

A bias correction scheme for aircraft temperature observations has been implemented in 288 MERRA-2, motivated by the known warm bias of these measurements compared with 289 other data sources (Cardinali et al. 2003, Ballish and Kumar 2008; Rienecker et al. 290 2011). The scheme uses the mean observed-minus-background departures to estimate 291 the bias for temperature reports from individual aircraft, identified by their tail number. 292 The bias estimates are updated after each analysis. The scheme is used to correct 293 Aircraft Meteorological Data Relay (AMDAR) and Aircraft Communications Addressing 294 and Reporting (ACARS) reports only, since other sources of aircraft observations in 295 MERRA-2 do not have unique identifiers by which they can be tracked. As of 2015, bias 296 corrections for approximately 3700 separate aircraft are tracked in MERRA-2. 297

The performance of the scheme is discussed in McCarty et al (2016). As expected, 298 the scheme is shown to reduce the bias between the corrected aircraft observations 299 and the background forecast, as well as reduce the variance of the corrected background 300 departures, allowing more aircraft observations to be used in the analysis. Unfortunately, 301 the MERRA-2 background state was found to have a larger than expected positive bias 302 in the mid- to upper troposphere, which feeds back to the bias estimates. The result 303 is that the bias correction actually increases the aircraft temperatures in some cases, 304 and the fit to other unbiased observation types such as radiosondes is degraded. This is 305 discussed further in section 3. 306

307 f. Mass conservation and water balance

³⁰⁸ Studies have documented the difficulty of maintaining realistic balances between varia-³⁰⁹ tions in total mass and total water content in previous reanalyses (e.g., Trenberth and

Smith 2005; Bosilovich et al. 2011; Berrisford et al. 2011). These studies show that 310 analysis adjustments to moisture are often large (when, ideally, they should be small), 311 highly sensitive to changes in the observing system, and mostly balanced by unphysical 312 changes in precipitation. Takacs et al. (2016) argue that, in attempting to analyze the 313 total mass of the atmosphere from surface pressure observations, reanalyses may violate 314 the simple physical constraint that, to an excellent approximation, the total dry mass of 315 the atmosphere is invariant, and so changes in total mass must be essentially equivalent 316 to changes in total water mass. At the same time, Berrisford et al. (2011) argue that, 317 while the observing system may not provide the data to determine exactly the total mass 318 of the atmosphere, the degree to which dry mass is preserved in a reanalysis provides a 319 useful diagnostic of reanalysis quality. 320

Reconsideration of these issues during the development of MERRA-2 prompted modifications to GEOS to conserve atmospheric dry mass and to guarantee that the net source of water from precipitation and surface evaporation equals the change in total atmospheric water. As described by Takacs et al. (2016), this has been achieved by making the following changes to the forecast model and assimilation procedure:

- Sources and sinks of atmospheric water have been added to the model continuity equation so that changes in total mass are driven purely by changes in total water.
- A constraint that penalizes analysis increments of dry air has been added to the GSI.
- 330

• Tendencies in the IAU are rescaled so that the global mean is removed from the analysis increment of water.

³³² The global impact of these modifications is illustrated in Figures 2 and 3, which compare

different components of atmospheric mass in MERRA and MERRA-2. Figure 2 shows 333 monthly mean anomalies from the mean seasonal cycle for total mass, total water, and 334 dry-air mass in the two reanalyses. In MERRA, there is an increase in total water over 335 the period, with significant inter-annual variations, but these features do not necessarily 336 match the changes in total mass. There also are spurious anomalies in dry-air mass 337 throughout, some of which track closely with the changes in total mass. In MERRA-2, 338 changes in total mass and total water track each other almost perfectly, by design, and 339 the dry-air mass remains a constant whose value must be specified. For the latter, the 340 value 983.24 hPa is chosen based on MERRA. This value falls within 0.1% of the values 341 derived from other recent reanalyses (Takacs et al. 2016). 342

Figure 3 shows monthly mean values of evaporation minus precipitation (E-P), or water 343 source term), the vertically integrated analysis increment of water, and the atmospheric 344 water storage. Note that the atmospheric water storage has similar magnitude in both 345 reanalyses and is dominated by the seasonal cycle. In MERRA, however, the storage 346 is determined by a near balance between the large and highly variable contributions 347 from the analysis increment on the one hand, and unphysical variations in E - P of 348 the opposite sign on the other hand. This includes an abrupt change in the sign of 349 these quantities after the introduction of AMSU-A in 1998 (Robertson et al. 2011). In 350 MERRA-2, the globally integrated analysis increment is zero, by design, and the water 351 storage is determined as in nature by small seasonal differences in E and P. It should be 352 noted that removing the global mean analysis increments of total mass and water mass 353 does not imply that the analysis increments of water vapor or surface pressure vanish 354 locally, as shown in Section 3 of this paper and discussed in further detail by Bosilovich 355 et al. (2017). 356

³⁵⁷ g. Observation-corrected precipitation forcing

The precipitation generated by the atmospheric model during the IAU segment of the 358 assimilation procedure is subject to considerable errors that can propagate into land 359 surface hydrological fields and beyond (Reichle et al. 2011). To mitigate these effects 360 in MERRA-2, the model-generated precipitation is corrected with observations before 361 being used to force the land surface or affect the wet deposition of aerosols over land 362 and ocean. Both the model-generated precipitation and the precipitation seen by the 363 land surface and the aerosols are available in the MERRA-2 output. MERRA-2 is 364 one of several recent applications of GEOS that uses observation-corrected precipitation 365 estimates. Others include the GMAO seasonal forecasting system (Ham et al. 2014), the 366 MERRA-Land data product (Reichle et al. 2011), and the MERRA ero aerosol reanalysis 367 (Buchard et al. 2015). Precipitation observations have also been used in reanalyses 368 produced by NOAA, including the North American Regional Reanalyis (Mesinger et al. 369 2006) and in the Climate Forecast System Reanalysis (CFSR, Saha et al. 2010; Meng 370 et al. 2012), although in both cases the approaches differ from that used in MERRA-2. 371 Some discussion of the differences between the approaches used in MERRA-2 and CFSR 372 can be found in Reichle et al. (2017a). 373

The corrected precipitation in MERRA-2 is derived from publicly available, observa-374 tionally based global precipitation products disaggregated from daily or pentad totals 375 to hourly accumulations using precipitation estimates from MERRA (Reichle and Liu 376 2014; Reichle et al. 2017a). The land surface in MERRA-2 sees a combination of cor-377 rected and model-generated precipitation depending on latitude, with the land surface 378 forced primarily by the corrected estimates at low to mid-latitudes, by the MERRA-2 379 model-generated precipitation at high latitudes, and by a weighted mixture in between 380 to prevent spatial discontinuities in climatological means. This is illustrated in Figure 4, 381

which shows the annual average adjustment made to the model-generated precipitation in MERRA-2 for the period 1980–2015 using this technique. The greatest adjustments are made in the tropics, where precipitation is greatest and the corrected estimates are given most weight, while no adjustments are made poleward of 62.5° in either hemisphere.

Based on the evaluation of several metrics, Reichle et al. (2017a) found the observation-387 corrected precipitation to be more realistic overall than that generated by the model 388 within the cycling MERRA-2 system, or that of the MERRA and MERRA-Land data 389 products. Exceptions include discontinuities in the MERRA-2 corrected precipitation 390 that result from errors in the underlying gauge products, for example, in Myanmar and 391 South America. Another issue is the high bias in MERRA-2 summer precipitation in the 392 high latitudes (where precipitation observations are not used). Moreover, the diurnal 393 cycle of the MERRA-2 corrected precipitation has reasonable amplitudes compared to 394 independent observations, but the time-of-day of maximum precipitation is inherited 395 from MERRA and is unrealistic. 396

The improvements in the precipitation forcing are also reflected in the MERRA-2 land 397 surface estimates. Reichle et al. (2017b) show that soil moisture, snow, terrestrial 398 water storage, and runoff in MERRA-2 agree better with independent observations than 399 estimates from MERRA. Draper et al. (2017) further demonstrate that the temporal 400 behavior and long term mean values of the land-atmosphere turbulent fluxes in MERRA-401 2 are improved. Moreover, by applying the precipitation corrections within the coupled 402 atmosphere-land modeling system, MERRA-2 can provide more self-consistent surface 403 meteorological data than were used for MERRA-Land (Reichle et al. 2017a). This self-404 consistency is important for applications such as forcing land-only model simulations. 405

Finally, it should be noted that the atmospheric water and energy prognostic variables associated with the creation of precipitation in MERRA-2 are not directly modified by the corrected estimates, although they can be indirectly modified through subsequent feedback with the land surface.

410 h. Sea surface temperature and sea ice concentration

The boundary conditions for sea surface temperature (SST) and sea ice concentration 411 (SIC) in MERRA were based on the 1° weekly (or monthly) product of Reynolds et al. 412 (2002). In MERRA-2, SST and SIC boundary conditions are instead based on currently 413 available high-resolution (finer than 1°) daily products. However, as there exists no 414 continuous source of daily global high-resolution SST and SIC for the entire period 415 of MERRA-2—and no source of daily data whatsoever prior to 1982—the following 416 products were used in combination (Table 3): monthly 1° data from the Coupled Model 417 Intercomparison Project (CMIP) as in Taylor et al. (2000) for the period prior to 418 1982; daily 1/4° data from the NOAA Optimum Interpolation Sea Surface Temperature 419 (OISST) as in Reynolds et al. (2007) from 1982 thru March 2006; and daily $1/20^{\circ}$ 420 data from the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) as 421 in Donlon et al. (2012) from April 2006 onwards. Note that different versions of the 422 NOAA OISST product are used prior to and after January 2003, the latter including 423 satellite data from both AVHRR and the Advanced Microwave Scanning Radiometer-424 EOS (AMSR-E) on NASA's Aqua satellite, and the former including satellite data from 425 AVHRR only. The processing of these products into a unified gridded set of daily SST 426 and SIC boundary conditions for MERRA-2 is described in Bosilovich et al. (2015). 427 Care was taken to use both SST and SIC from the same data source to avoid potential 428 inconsistencies, especially in marginal ice zones. 429

Figure 5 shows 12-month running averaged values of SST between $60^{\circ}N$ and $60^{\circ}S$ for 430 MERRA-2 and several other reanalyses, including MERRA. In all cases, there is a 431 positive trend in SST throughout the period. The running means for all the reanalyses 432 are within 1 K for the 30 years spanning 1980–2010, and the anomalies (not shown) are 433 separated by less than 0.2 K. At the same time, there are clear systematic differences 434 between reanalyses, with the MERRA-2 SST's on the one hand being cooler than those 435 used in the other reanalyses shown except CFSR (which used similar input data sets), 436 especially before the transition to OSTIA in 2006. The values for JRA-55, on the other 437 hand, are on the order of 0.1 K higher than other reanalyses throughout the 35-year 438 period. It can also be seen that the MERRA-2 SSTs increase slightly with the change 439 in NOAA OISST versions after 2003. The reader is referred to Bosilovich et al. (2015) 440 for a more detailed list of known issues with the SST and SIC boundary conditions for 441 MERRA-2. 442

443 *i.* Production

MERRA-2 was produced in four separate streams, each of which was spun up for a year 444 at full resolution beginning on 1 January 1979 (stream 1), 1 January 1991 (stream 2), 445 1 January 2000 (stream 3) and 1 January 2010 (stream 4). The land surface restart 446 files for each MERRA-2 stream were themselves spun up for at least 20 years using the 447 off-line MERRA-2 land model forced by MERRA surface meteorological fields, and with 448 the precipitation replaced by the observation-corrected estimates described in section 449 2g. The final MERRA-2 product distribution is from stream 1 for 1 January 1980–31 450 December 1991, followed by stream 2 for 1 January 1992–31 December 2000, then stream 451 3 for 1 January 2001–31 December 2010, and finally stream 4 for 1 January 2011–present. 452 With streams 1–3 complete, MERRA-2 production continues as a near-real time climate 453

analysis from stream 4 alone. The decision to begin stream 1 in January 1979 and
distribute products beginning in January 1980—a year later than the schedule followed
in MERRA—was based on the fact that the products used to create the observationcorrected precipitation estimates for MERRA-2 only start on 1 January 1979, leaving no
viable way to initialize the land surface properly before this time (which requires several
months of spin-up, after initialization from climatological conditions).

The overlap periods between successive streams were examined to determine the ad-460 equacy of the spin-up procedure and to quantify the uncertainty in individual fields. 461 Differences between overlapping MERRA-2 streams were found to be minimal for most 462 fields after one year, with the exception of certain land surface variables including the 463 deep-level soil temperature and land surface soil moisture storage at high latitudes. 464 The spin-up of the land surface is addressed separately in Reichle et al. (2017a); sec-465 tion 3d and Figure 13 of that paper discuss specific examples of the aforementioned 466 discontinuities across consecutive MERRA-2 streams. Users should be aware of these 467 discontinuities when the data are used for specific applications. 468

469 3. Data assimilation diagnostics

By-products of the data assimilation procedure in the form of differences between fore-470 casts and observations, analysis increments, and estimates of bias can be used effectively 471 to monitor the quality of both the input and output of the assimilation. In this section, 472 examples of such diagnostics are presented for MERRA-2, focused mainly on feedbacks 473 with respect to in-situ conventional observations and on the net correction, or incre-474 ment, brought by the entirety of the assimilated observations. The reader is referred 475 to McCarty et al. (2016) for examples of feedbacks related to the treatment of satellite 476 radiance observations. 477

478 a. Background departure statistics

Differences between the assimilated observations and the background forecast, referred 479 to as innovations or background departures, provide important information about the 480 quality of the assimilation. In particular, it is important that the assimilation system be 481 able to predict high-quality observations, especially for conventional data types which 482 provide direct measurements of the analyzed variables. In addition to affecting the anal-483 ysis directly, many conventional data play an important role in anchoring the variational 484 bias estimates used in the assimilation of satellite radiances. Generally speaking, smaller 485 background departures indicate a higher quality assimilation. The results shown here 486 are selected to highlight both strengths and weaknesses of MERRA-2 in this regard. 487 As in MERRA, for convenience, gridded versions of the observations and corresponding 488 departures used in MERRA-2 will be made available to users. 489

Figure 6 shows time series of monthly mean and root mean square (RMS) background 490 departure statistics for all assimilated surface pressure observations in MERRA and 491 MERRA-2 for both the Northern and Southern Hemisphere. Also shown are the monthly 492 mean numbers of surface pressure observations assimilated in each 6-h assimilation cycle 493 in MERRA-2. The RMS values decrease with time in both reanalyses, especially in the 494 Southern Hemisphere after the early to mid 1990's when the number of observations 495 begins to increase significantly. The RMS values in the Northern Hemisphere (Fig-496 ure 6a) are smaller than in the Southern Hemisphere initially and decrease more slowly 497 with time, reflecting the greater number of conventional observations available over land 498 throughout the period. This decrease is slightly more pronounced in MERRA-2 after 499 the mid 1990's when the number of surfaces pressure observations from land stations 500 increases significantly. In the Southern Hemisphere (Figure 6b), the RMS values are 501 larger in MERRA-2 than in MERRA before the mid 1990's but smaller by the end of 502

the period. The larger values early on are due to the use of larger observation errors 503 for surface ship observations (and some other conventional data types) in MERRA-2. 504 allowing more "outliers" with larger departure values to pass the quality control pro-505 cedure in the analysis.¹ The impact diminishes by the mid 1990's as other observation 506 types, including from satellites, become more abundant. There is no similar effect in the 507 Northern Hemisphere where surface pressure observations from land stations are domi-508 nant early in the period; the observation errors specified for these data are the same in 509 MERRA and MERRA-2. Finally, the jump in RMS values in the Southern Hemisphere 510 evident in both reanalyses at the beginning of 1985 coincides with the introduction of 511 regularly spaced synthetic surface pressure observations over southern ocean areas. 512

The mean background departures for surface pressure in the Northern Hemisphere are 513 consistently less biased in MERRA-2 than in MERRA, especially after the mid 1990's. 514 In the Southern Hemisphere, however, the departures for MERRA-2 show a negative 515 bias throughout the period; this is discussed further in section 3b. The mean departures 516 in MERRA-2 also show a more pronounced annual cycle in this hemisphere. As a 517 point of reference, the background departure statistics for other reanalyses including, 518 for example, ERA-Interim (Dee at al. 2011) exhibit a clear annual cycle, but with 519 somewhat smaller amplitude than in MERRA-2. 520

Figure 7 shows global background departure statistics for radiosonde temperatures for MERRA and MERRA-2 at selected pressure levels in the troposphere (300 hPa and 700 hPa) and stratosphere (10 hPa and 50 hPa). Also shown for each level are the monthly mean numbers of radiosonde temperature observations assimilated in each 6-h assimilation cycle in MERRA-2. In the troposphere (Figures 7c and d), the performance of MERRA-2 is degraded compared to that of MERRA, especially at 300 hPa. The

¹The observation errors for conventional data types have been adjusted since MERRA-2.

RMS values for MERRA-2 decrease with time but remain 10–20% larger than those for MERRA during much of the period. Again, this is due at least partially to the use of larger observation errors for radiosonde temperatures and other conventional data types in MERRA-2. Noticeable improvements occur first in the mid 1990's when satellite observations become more abundant, and again in 2006 when the number of GPSRO observations increases significantly.

The mean departure values at 300 hPa for both MERRA and MERRA-2 exhibit a clear 533 negative bias. The bias is generally larger in MERRA-2, reaching a maximum amplitude 534 of greater than 0.5 K during the early 2000's. This is due to a warm model bias in the 535 upper troposphere which worsened during the course of development between MERRA 536 and MERRA-2 (see also Figure 10). However, aspects of the assimilation process may 537 exacerbate the problem. It can be seen for example that the bias in the background 538 departures at 300 hPa increases noticeably after the mid 1990's, especially in MERRA-2, 539 when the numbers of both aircraft temperature observations and satellite radiances begin 540 to increase significantly (Figure 1). The design of the bias correction procedures for both 541 observation types is such that they result in an adjustment of the observations regardless 542 of the source of the bias. In the presence of a strong model bias this can reinforce the 543 actual observational bias and cause the assimilation system to drift further toward the 544 model state, as noted in the case of the aircraft bias corrections described in section 2e. 545 A similar, though less direct, effect may occur through the observational bias corrections 546 used to assimilate satellite radiances, although other aspects of the variational scheme 547 used to adjust these data act to reduce this risk (Dee and Uppala 2009). At 700 hPa, the 548 mean departures for both reanalyses are generally more comparable and considerably 549 less biased. 550

⁵⁵¹ In the stratosphere (Figures 7a and b), there are fewer significant differences between

the results for MERRA and MERRA-2 although the departures at 10 hPa for MERRA-2 552 show a larger negative bias of -0.2 K to -0.3 K prior to the early 2000's. After 2002, 553 when assimilation of AIRS radiances begins, the biases at 10 hPa in both reanalyses 554 exhibit an upward trend and eventually become positive, first in MERRA around 2003 555 and then in MERRA-2 in 2005. There is a discernible jump in the mean departures 556 at this level for MERRA-2 in 2005. This is around the time when assimilation of both 557 MLS temperature retrievals (above 5 hPa) and GPSRO bending angle observations (up 558 to approximately 10 hPa) begins in MERRA-2, but this does not appear to improve 559 the fit to radiosondes at 10 hPa compared with MERRA. After 2006, the biases in both 560 reanalyses have average values of 0.2 K to 0.3 K. Finally, at 50 hPa, the departure values 561 for both reanalyses are very similar and exhibit only a small positive bias throughout. 562

Figure 8 shows statistics for radiosonde specific humidity background departures at 500 and 850 hPa in the tropics. The performance of MERRA-2 is slightly worse than that of MERRA in the middle troposphere in terms of both RMS and bias, but similar or slightly better in the lower troposphere. Again, the mean departure values are consistent with known biases in the GEOS model.

568 b. Analysis increments

The analysis increments represent the net adjustment to the background state by the assimilation scheme in response to all the observations. As this adjustment depends in a complex way on assumed or crudely estimated errors in the observations and background state, and on the forward operator that transforms the model variables to observation space, the increments do not necessarily represent errors in the background state. Nonetheless, their spatial and temporal variations provide an important diagnostic of system performance, including how changes in the observing system may affect the consistency of the analysis. Systematic increments often indicate the presence of biases in
the model or observations which may complicate the use of reanalyses for estimating
budgets and identifying trends (Dee et al. 2011).

As described in section 2, the GEOS assimilation system uses an IAU procedure which, instead of correcting the initial condition, applies the analysis increment to the model as a constant tendency term during the 6-h assimilation window. It is this contribution to the time tendency from the analysis that is provided as a standard output quantity in MERRA-2, examples of which are presented here. For convenience, these are referred to as simply the analysis increments in the discussion that follows.

Figure 9 shows the mean and standard deviation in time of the monthly mean analysis 585 increment of surface pressure in MERRA-2 for the period January 1980 through De-586 cember 2015. The monthly means themselves have been computed from sub-daily data, 587 eight times per day. The pattern of the mean increments indicates that the analysis 588 tends to move mass from the oceans to the continents, as noted also by Takacs et al. 589 (2016), although this pattern is arguably most robust in the Southern Hemisphere. (The 590 mostly negative surface pressure increments over Canada provide an obvious counter ex-591 ample.) These results are consistent with those in Figure 6 showing a negative bias in 592 the Southern Hemisphere background departures in MERRA-2. The standard devia-593 tion of the increments shows that the largest variations in surface pressure occur in the 594 middle and high latitudes, and especially over coastal Antarctica and the mountainous 595 regions of southern and eastern Asia, as well as southern Alaska. 596

Time series of the global monthly mean and standard deviation of the analysis increments of temperature from the surface to 70 hPa in MERRA-2 are shown in Figure 10. The most striking feature in the mean increments is the persistent cooling by the analysis

in the layer between 250 and 400 hPa. This is consistent with the negative bias in the 600 background departures at 300 hPa shown in Figure 7 and provides further evidence of 601 the warm model bias at these levels. Except for seasonal variations, the magnitude 602 of the cooling remains relatively constant throughout much of the period, although 603 noticeable changes occur, for example, beginning in the mid to late 1990's as the number 604 of aircraft and satellite observations increase, and again in 2006, possibly in response 605 to the introduction of data from IASI and GPSRO. Warming by the analysis is evident 606 above 200 hPa and below 700 hPa. In this global view, the mean increments close to the 607 surface exhibit a negative trend with strong warming before the early 1990's turning to 608 slight cooling after 2010, but this is in fact the net effect of distinct regional differences 609 in the increments (not shown). In particular, near-surface warming by the analysis in 610 response to a cold model bias over northern midlatitude land masses is offset by cooling 611 over southern oceans that generally increases with time beginning with the assimilation 612 of data from the first microwave humidity sensors in the late 1980's. These differences 613 also contribute to the large variability of the increments below 700 hPa (Figure 10b). 614 The variability in the mid troposphere is noticeable but small compared with that at 615 low levels, again highlighting the consistency of the cooling by the observations between 616 250 and 400 hPa. 617

The increments of specific humidity in the tropics are shown in Figure 11 for levels 618 between the surface and 250 hPa (the values become exceedingly small above this level). 619 The mean increments indicate distinct biases in the middle and lower troposphere, with 620 systematic drying between 600 and 300 hPa, and mostly moistening below 700 hPa. The 621 corrections are generally larger during the second half of the period and especially after 622 the late 1990's as more satellite observations of humidity become available. There is an 623 abrupt increase in the variability of the increments corresponding to the introduction of 624 the first SSM/I instrument in mid 1987, with additional increases corresponding to the 625

use of a second and third SSM/I instrument in late 1990 and mid-1995, respectively. The
use of multiple SSM/I instruments from the early 1990's to late 2000's also corresponds
to a strong drying and a marked increase in variability at levels very close to the surface.
The introduction of AMSU-B data in 1998 corresponds to marked increases in the mean
and variability of the increments, the latter being most pronounced in the layer between
800 and 900 hPa. The sensitivity of the precipitation to these observing system changes
is discussed in section 5.

633 4. Aerosol data assimilation

In addition to a standard meteorological analysis, MERRA-2 includes an aerosol analysis 634 as described in Randles et al. (2016, 2017) and Buchard et al. (2017). The multi-decadal 635 coverage and the coupling between aerosols and the circulation is a step forward com-636 pared to previous EOS-era reanalyses such as MERRAero, the Navy Aerosol Analysis 637 and Prediction System (NAAPS) reanalysis (Lynch et al. 2016), the Monitoring At-638 mospheric Composition and Climate (MACC) reanalysis (Inness et al. 2013), and the 639 more recent Copernicus Atmosphere Monitoring Service (CAMS) reanalysis (Flemming 640 et al. 2017). The MERRA-2 system produces 3-hourly analyses and gridded output of 641 both observable parameters and aerosol diagnostics not easily observed, especially on a 642 global scale, with potential applications ranging from air quality forecasting to studies 643 of aerosol-climate and aerosol-weather interactions (e.g., Bocquet et al. 2015). 644

An analysis splitting technique (Randles et al. 2017) is used to assimilate aerosol optical depth (AOD) at 550 nm, in which a two-dimensional analysis is performed first using error covariances derived from innovation data and then the horizontal increments are projected vertically and across species using an ensemble method. AOD observations are derived from several sources, including

- Reflectances from AVHRR (1979–2002, ocean-only, Heidinger et al. 2002);
 Reflectances from the Moderate Resolution Imaging Spectroradiometer (MODIS) on Terra (2000–present) and Aqua (2002–present) (Remer et al. 2005; Levy et. al. 2007);
 AOD retrievals from the Multi-angle SpectroRadiometer (MISR) (2000–2014, bright,
- AOD retrievals from the Multi-angle SpectroRadiometer (MISR) (2000–2014, bright desert regions only, Kahn et al. 2005);
- Direct AOD measurements from the ground-based Aerosol Robotics Network (AERONET)
 (1999–2014, Holben et. al. 1998).

MODIS provides the vast majority of AOD observations assimilated in MERRA-2, es-658 pecially after 2002 when data from both the Terra and Aqua satellites become available. 659 Prior to 2000, only AVHRR reflectances over ocean are used in MERRA-2. AOD for 660 both MODIS and AVHRR are derived from cloud-cleared reflectances using a neural net 661 procedure trained on AERONET measurements (Randles et al. 2017). By construction, 662 these AOD retrievals are unbiased with respect to AERONET observations. AOD from 663 MISR and AERONET observations are used without bias correction. Additional details 664 about the aerosol observing system in MERRA-2 can be found in Randles et al. (2016, 665 2017). 666

The Goddard Chemistry, Aerosol, Radiation and Transport model (GOCART; Chin et al. 2002; Colarco et al. 2010) is coupled with the GEOS atmospheric model to simulate the life cycles of five externally-mixed aerosol species, including dust, sea salt, black carbon, organic carbon, and sulfate. The model carries three-dimensional mass mixing ratios of these five aerosol species as prognostic aerosol tracers. The AOD at 550 nm is a column- and species-integrated optical quantity, which is calculated as the summed product of each species mass and its extinction coefficient based on aerosol

optical properties derived largely from the Optical Properties of Aerosols and Clouds 674 (OPAC) dataset (see Randles et al. 2017 and references within.) Emissions of both dust 675 and sea salt are wind-driven for each of five size bins, parameterized following Martio-676 corena and Bergametti (1995) and Gong (2003), respectively. Sulfate and carbonaceous 677 aerosol emissions derive from both natural and anthropogenic sources as described in 678 Randles et al. (2017). In particular, MERRA-2 includes volcanic sources (Diehl et al., 679 2012) and biomass burning emissions that utilize satellite observations, and are based on 680 the Reanalysis of the Tropospheric chemical composition, version 2 (RETRO-2, Schultz 681 et al. 2008), the Global Fire Emissions Database, version 3.1 (GFED-3.1, van der Werf et 682 al. 2006), and the Quick Fire Emission Dataset, version 2.4r6 (QFED-2.4.r6, Darmenov 683 and da Silva, 2015). 684

It should be noted that AOD observations can only directly constrain the total, species-685 integrated and vertically-integrated aerosol extinction — a quantity that can be related 686 to column aerosol mass by assuming a set of optical properties. Non-analyzed aerosol 687 properties such as the vertical distribution, aerosol speciation, and absorption are not 688 fully constrained by the observations and are chiefly determined by the underlying model 689 physics and error covariance assumptions. Despite this fact, Buchard et al. (2017) 690 show that the MERRA-2 aerosol reanalysis has considerable skill in simulating numer-691 ous observable aerosol properties. Randles et al. (2017) show that the AOD fields in 692 MERRA-2 generally have both high correlation and low bias relative to independent 693 (non-assimilated) sun-photometer and aircraft observations. 694

As in the case of the meteorological analysis discussed in section 3, statistics of background and analysis departures provide a basic metric of the quality of the aerosol assimilation. Figure 12 shows probability distribution functions of collocated observationminus-forecast and observation-minus-analysis departures from MERRA-2 for each sen⁶⁹⁹ sor in the aerosol observing system. Statistics are shown in terms of the log-transform ⁷⁰⁰ AOD analysis variable (i.e., $\ln[AOD + 0.01]$) which is approximately normally dis-⁷⁰¹ tributed (Randles et al., 2017). Note that AOD is a dimensionless quantity and log-⁷⁰² transformed AOD is typically in the range (-4, 2). As expected, compared to the forecast ⁷⁰³ departures, the analysis departures show reduced bias with respect to the observations. ⁷⁰⁴ Note also that the innovation variances are much larger over land than ocean, a direct ⁷⁰⁵ consequence of the signal-to-noise limitation of aerosol retrievals over land.

Regional aspects of the global distribution of aerosols are illustrated in Figure 13, which 706 shows time series of analyzed AOD from MERRA-2 area-averaged over several major 707 aerosol source regions. The contribution of each aerosol species to the total AOD is 708 indicated by the colored shading. The seasonal cycles of dust and biomass burning 709 (carbonaceous) AOD are apparent in all regions. Large increases in sulfate aerosol occur 710 in all regions after the El Chichon (1982) and Pinatubo (1991) volcanic eruptions. Over 711 the Asian region (Figure 13a), the analysis captures high carbonaceous aerosol associated 712 with the 2003 Siberian fires and the increasing trend in AOD between the late 1990s 713 and present (commensurate with increasing anthropogenic aerosol emissions reported by 714 Diehl et al. 2012). The AOD over northern Africa (Figure 13b) is dominated by dust, 715 and major dust transport events such as in 2010 are captured (see Buchard et al. 2017 716 for details). Carbonaceous aerosol from biomass burning in major source regions such 717 as the Amazon Basin are also well captured (Figure 13c), especially after 2000 when 718 emissions inventories derive from MODIS observations (Darmenov and da Silva, 2015). 719

Figure 14 compares values of AOD from several recent aerosol reanalyses for the period 2003–2010. Where such information is available, the results are partitioned by species and identified as either fine or coarse mode (see caption for details). Also shown are multi-model average results from Phase I of the Aerosol Comparison (AeroCom)

inter-comparison project (Kinne et al. 2006), as well as both model and observational 724 estimates from Yu et al. (2006). The latter study includes an attempt to account for 725 satellite clear-sky biases by combining MODIS and MISR observations with the GO-726 CART model. Compared to MERRAero, for example, MERRA-2 has slightly higher 727 global average AOD due to increased contributions from dust (related to the assimi-728 lation of MISR AOD over bright surfaces) and sea salt (related to changes in model 729 physics). MERRA-2 and NAAPS show similar global average AOD, both for fine and 730 coarse mode aerosol. Models without assimilation (AeroCom and Yu_Model) underesti-731 mate global average AOD compared to both observational estimates (Yu_Obs) and the 732 aerosol reanalyses. The MACC aerosol reanalysis has the highest global mean AOD 733 (Bellouin et al. 2013), which is close to the MODIS-only value of 0.188 for the period 734 2003–2010 (Yu et al. 2006). MACC also has more dust and sea salt aerosol compared 735 to the other reanalyses, particularly over the ocean (not shown). 736

The direct aerosol impact on the radiative energy balance of Earth is dependent on the 737 vertical distribution of aerosol scattering and absorption, which is not fully constrained 738 by the vertically integrated AOD measurements that MERRA-2 assimilates. An as-739 sessment of the aerosol vertical structure and absorption is presented in a companion 740 paper (Buchard et al. 2017). Long-term aerosol reanalyses can potentially reduce un-741 certainty in how aerosol direct effects have changed over time, particularly once better 742 observational constraints on aerosol absorption become available. The direct radiative 743 effect (DRE) of all aerosols is defined as the flux difference in W m^{-2} between clear-744 sky and clear clean-sky conditions (no aerosols or clouds). In the absence of clouds, 745 this quantity is less sensitive to the vertical distribution of aerosol absorption, although 746 it remains sensitive to absorbing aerosols over surfaces with high albedo (Chýlek and 747 Coakley, 1974). 748

Table 4 compares the DRE from MERRA-2, MERRAero, MACC, model inter-comparisons, 749 and the observationally constrained estimate of Yu et al. (2006). Listed are the top-of-750 the atmosphere (TOA), surface (SFC), and atmospheric (ATM) estimates of DRE for the 751 period 2003–2010, averaged over land and ocean separately. Note that the atmospheric 752 contribution to the DRE is defined as the difference between top-of-the-atmosphere 753 and surface values, ATM = TOA - SFC. Over land, the DRE estimate from MACC 754 best agrees with the observationally-constrained estimate. TOA and SFC forcing in 755 MERRA-2 and MERRA are lower than in MACC due to their lower AOD, although 756 the atmospheric forcing is similar. Over ocean, the DRE estimates from MERRA-2 and 757 MACC are lower and higher, respectively, than the observational estimate, and both re-758 analyses have lower estimates of atmospheric absorption. Much of the uncertainty in the 759 DRE reported by the Intergovernmental Panel on Climate Change (IPCC) arises from 760 differences between estimates from global models and satellite-based estimates (Myhre 761 2009). However, as aerosol reanalyses such as MERRA-2 continue to mature and incor-762 porate additional observations (e.g., from lidars and multi-spectral sensors), we expect 763 a narrowing of the gap between simulated and satellite-based estimates of the DRE. 764

765 5. Precipitation

The representation of precipitation in a reanalysis is key to applications in weather and climate as it ties together aspects of both the water and energy cycles. It also presents a significant challenge, however, as estimates of precipitation are only indirectly constrained by observations and are strongly dependent on model physics whose parameterizations have known errors and can be highly sensitive to even small changes in large-scale temperature and humidity fields. The observations themselves can sometimes introduce additional uncertainty in these estimates as a result of heterogeneous sampling, changes in instrumentation, and time-varying calibration (Bosilovich et al.2017).

While improved representation of the hydrological cycle was a primary focus of MERRA, 775 the character of its global precipitation in particular was found to be highly sensitive to 776 the assimilated observations and thus to changes in the observing system (e.g., Robert-777 son et al. 2011). Among the development aspects of MERRA-2 intended to address 778 this issue are modifications to GEOS to conserve atmospheric dry mass and ensure that 779 changes in global atmospheric total mass are equivalent to changes in total water (sec-780 tion 2f), exclusion from the analysis of microwave temperature sounding channels with 781 strong surface sensitivity (section 2d) and, less directly, forcing of the land surface by 782 observation-corrected precipitation estimates (section 2g). 783

784 a. Global aspects

Bosilovich et al. (2015, 2017) have investigated the global water cycle variability in 785 MERRA-2 using comparisons with observational data sets and other recent reanalyses. 786 Those studies present a broad range of metrics on this topic, a small subset of which 787 are summarized here. Figure 15 shows time series of global mean precipitation for sev-788 eral recent reanalyses and the observation-based estimates from the Global Precipitation 789 Climatology Project (GPCP, Adler et al. 2003). MERRA-2 exhibits larger temporal 790 variability than GPCP but similar temporal variability as other recent reanalyses, and 791 noticeably less spurious temporal variability than MERRA. The largest improvements 792 compared with MERRA in this regard relate to the decreased sensitivity of MERRA-2 793 to the introduction of AMSU-A radiances on NOAA-15 and -16 in the late 1990s, and 794 to the loss of SSM/I radiances in the late 2000s. There is still an obvious sensitivity 795 in MERRA-2 to the introduction of SSM/I in 1987, but the response to these data 706

is comparable in magnitude to those of the other reanalyses shown. The response in 797 MERRA-2 appears accentuated due to the decrease and subsequent recovery of precip-798 itation through the mid 1980s. This behavior is not reflected in the GPCP time series, 799 but is evident to lesser degrees in CFSR and ERA-Interim, especially after 1983. For 800 MERRA-2 and CFSR, this may be related to the fact that the SST boundary condi-801 tions used in these reanalyses reach their global minimum value for the entire reanalysis 802 period after 1985 (Figure 5), but further investigation is required to confirm this. The 803 increasing trend in global precipitation in MERRA-2 from approximately 2.9 mm day⁻¹ 804 in 1988 to approximately 3.0 mm day⁻¹ in 1998 is likely due to increasing evaporation 805 over oceans driven by the assimilation of additional SSM/I wind speed observations and 806 the tight coupling of evaporation and precipitation in MERRA-2 through the global 807 mass constraint (Bosilovich et al. 2017). Overall, the global mean precipitation values 808 are higher than those of GPCP but well within the envelope of other recent reanalyses. 809

Spatial comparisons provide additional insight into the strengths and weaknesses of the 810 representation of precipitation globally in MERRA-2. Figure 16 shows maps of time-811 averaged differences in precipitation during boreal summer for MERRA and MERRA-2 812 compared with GPCP. MERRA-2 shows general improvement compared to MERRA 813 over oceanic regions in both the tropics and extratropics, but an increase in positive 814 bias over northern high latitudes. A notable deficiency in MERRA-2 is the excessive 815 precipitation in the vicinity of high topography in the tropics, especially along the Andes 816 and over the maritime continent. This is related to the partitioning between resolved 817 (large scale) and parameterized (convective) precipitation in the MERRA-2 model which, 818 being more heavily skewed toward the former, results in large-scale precipitation over 819 high topography that is difficult to control. In comparing these features with available 820 gauge data, Bosilovich et al. (2015) point out that the maximum precipitation values 821 in MERRA-2 do not always coincide with the maximum terrain height, so that other 822

effects also may play a role locally. Despite this deficiency over tropical land areas, the positive bias over the warm pool present in MERRA is slightly improved in MERRA-2. Additionally, the high precipitation bias over the Central America Sea in MERRA has been reduced significantly in MERRA-2 and precipitation over the Bay of Bengal and Arabian Sea is slightly improved. Results for other seasons (not shown) are qualitatively similar to those in Figure 16.

⁸²⁹ b. US summertime precipitation variability

⁸³⁰ Deficiencies in the ability of MERRA to reproduce certain aspects of the summer-⁸³¹ time seasonal precipitation over the United States (US) have been well documented ⁸³² (Bosilovich 2013). In particular, MERRA was unable to produce seasonal highs and ⁸³³ lows in regional precipitation that were similar to observations. For example, droughts ⁸³⁴ and floods were only weakly reproduced.

Figure 17 shows the time series of summertime seasonal precipitation anomalies over 835 the midwestern US as derived from the NOAA Climate Prediction Center (CPC) gauge 836 observations and from MERRA and MERRA-2 model-generated precipitation. (The 837 correlation values between various reanalyses and the gauge data for this and other 838 regions of the US are shown in Figure 18.) The limitations of MERRA are apparent, 839 especially when comparing values for 1988 (regional drought) and 1993 (large-scale flood-840 ing) with the observed values. In contrast, MERRA-2 is able to reproduce the 1988 and 841 1993 anomalies and is generally much better at tracking the overall variability of the 842 observed anomalies. The poor performance of MERRA-2 in 1980 is a notable exception. 843 A significant drought occurred in the southern Great Plains that year, but its location 844 in MERRA-2 extended too far northeastward into the midwestern US. 845

Figure 18 presents regional summary statistics for US summer seasonal precipitation 846 anomalies for selected reanalyses. The regions are defined as in Bosilovich (2013). For 847 each region, the temporal mean, standard deviation, and anomaly correlation with re-848 spect to the CPC data are derived from time series like those shown in Figure 17. In 849 general, precipitation mean values across the US are improved in MERRA-2 compared 850 with MERRA (Figure 18a), and in many regions the values for MERRA-2 improve over 851 those of other reanalyses as well. There is also a marked increase in the standard de-852 viation of the MERRA-2 time series relative to MERRA (Figure 18b). As discussed 853 above, for example, MERRA-2 more realistically reproduces the seasonal extremes in 854 midwestern US precipitation. Note, however, that MERRA-2 overestimates the stan-855 dard deviation with respect to the CPC estimates in some regions. Ancillary results 856 indicate that this is due to an excess in the number of days with rain in MERRA-2. 857 Improvements in MERRA-2 are most evident in the anomaly correlation of the seasonal 858 time series (Figure 18c). In this measure, the two most recent reanalyses, JRA-55 and 859 MERRA-2, generally outperform the others. MERRA-2 produces the highest values of 860 the reanalyses shown in most regions, with substantially higher values in a few of these 861 regions. 862

The detection and analysis of extreme weather, including extreme precipitation events, 863 is a topic of societal interest and another potential application of reanalyses. At least 864 some of this interest is related to assessing changes in the risk of such events in the con-865 text of climate change. For example, observation-based studies cite strong evidence of 866 an upward trend in the frequency and intensity of extreme precipitation events averaged 867 over the US during the last 50 years (Kunkel et al. 2013), although the causes of the 868 observed trends are less certain. Figure 19 shows the accumulated precipitation amounts 869 for the largest precipitation events (at the 99th percentile) as derived from gauge ob-870 servations, MERRA, and MERRA-2. Compared with the observations, MERRA shows 871
very low values, and very little structure across the continental US. MERRA-2, on the 872 other hand, exhibits a spatial pattern more similar to the observations, and the magni-873 tude of the extreme rainfall is also more similar to the observations. MERRA-2 does, 874 however, overestimate the precipitation values over the Midwestern US. While the re-875 sults in Figure 19 provide an indicator of how the representation of extreme events has 876 improved in MERRA-2 compared with MERRA, the relatively coarse resolution of both 877 reanalyses limits their utility for studying such events in detail. Presumably, the trend 878 toward increasing resolution, among other improvements, will reduce these limitations 879 in future global reanalyses. 880

⁸⁸¹ 6. The stratosphere

In MERRA-2 the stratospheric meteorology and ozone have benefited from improve-882 ments to the GEOS atmospheric model and GSI analysis scheme, as well as from the 883 addition of observations that were not incorporated into MERRA. The model changes 884 most relevant to the stratosphere are the use of the cubed sphere grid and the re-tuning 885 of the gravity wave drag (GWD) parameterization. The amplitude of the non-orographic 886 GWD was increased in the tropics, enabling a model-generated Quasi-Biennial Oscilla-887 tion (QBO) that was not found in the model version used for MERRA (Molod et al. 888 2015). Having a model-generated QBO, in turn, results in smaller lower-stratospheric 889 analysis wind increments in MERRA-2 than in MERRA (Cov et al. 2016). The strength 890 of the orographic GWD was also increased in the Southern Hemisphere to better model 891 the strong, late-winter westerlies found there (Molod et al. 2015). 892

The main GSI change relevant to the stratosphere is the use in MERRA-2 of the CRTM for the assimilation of SSU radiances while in MERRA the SSU assimilation was based on GLATOVS (section 2d). These SSU radiance channels are a major source of stratospheric information during the 1980's and 1990's, although the SSU instruments during
these decades span several satellite platforms, each with different bias characteristics
(Kobayashi et al. 2009). The CRTM has been enhanced for SSU data assimilation since
MERRA and now accounts for these biasing factors.

The main additional observations relevant to the stratosphere for MERRA-2 are GPSRO 900 bending angle observations from the suite of platforms beginning in July 2004, and 901 temperature and ozone measurements of the middle atmosphere from MLS and OMI on 902 the EOS Aura satellite beginning later the same year (Froidevaux et al. 2006; Schwartz 903 et al. 2008; McPeters et al. 2008). MERRA-2 assimilates GPSRO bending angle 904 observations up to 30 km. Details of the GPSRO platforms assimilated by MERRA-2 905 can be found in McCarty et al. (2016). The GPSRO observations aid lower stratospheric 906 bias correction by providing a stable source of temperature and moisture measurements. 907 The MLS-retrieved temperature profiles are assimilated in MERRA-2 at altitudes above 908 5 hPa, providing a strong constraint on the dynamics of the stratopause and lower 909 mesosphere. As shown below in section 6a, this improves the quality of the synoptic 910 meteorological fields at these altitudes but may complicate the study of trends. The 911 MLS and OMI contributions to ozone assimilation are discussed in section 6b. 912

913 a. Meteorology

The cubed sphere discretization of the MERRA-2 model eliminates computational instabilities near the poles, a characteristic of latitude-longitude grids. This is especially important for stratospheric analysis where strong cross-polar flow events occur frequently, especially during major sudden warming events, as planetary-scale Rossby waves disturb the polar vortex. Ertel's Potential Vorticity (EPV), a scalar based on the horizontal vorticity, is often used to characterize the stratospheric circulation (Andrews et al. 1987),

where stronger EPV gradients imply stronger flow. Figure 20 illustrates a case where 920 the analyzed wind speeds in MERRA-2 reached nearly 170 m s⁻¹ close to the polar 921 stratopause on 2 January 1995 at 12 UTC. On a global scale (Figure 20a and b), the 922 MERRA and MERRA-2 EPV fields appear similar, with the polar vortex (indicated by 923 green and orange colors) displaced well off the pole. In both cases, strong winds cross 924 the North Pole as they circle around the region of high EPV. However, a closer look 925 reveals that the EPV in MERRA (Figure 20c) has anomalous radial perturbations near 926 the pole, while the EPV in MERRA-2 (Figure 20d) shows a smooth and strong EPV 927 gradient in this region. Note also that while the largest discontinuities in the MERRA 928 EPV field occur close to the pole itself, their effects can extend well beyond this location. 929

Figure 21 provides an example of how the assimilation of MLS temperature measure-930 ments in MERRA-2 improves the representation of the dynamics near the stratopause. 931 The figure shows the time-height evolution of polar temperatures during the 2005–2006 932 Northern Hemisphere winter in which a major stratospheric sudden warming occurred. 933 In a comprehensive study of this winter based on MLS observations, Manney et al. 934 (2008) documented the disappearance of the warm polar stratopause during the warm-935 ing and its later high-altitude reformation and subsequent descent. This breakdown 936 and high-altitude reformation in early February 2006 is now well captured in MERRA-2 937 (Figure 21b), in contrast to MERRA (Figure 21a). 938

The characteristics of the assimilation on longer time scales is illustrated in Figure 22, which shows the time-height evolution of global monthly averaged temperature anomalies in MERRA-2. The 35-year mean and annual cycle for the period 1980–2015 have been subtracted from each pressure level. The global temperatures in the lower stratosphere (100–10 hPa) show no obvious discontinuities as different instruments become available. There is a slight cooling with time over the 35 years, which is generally consistent with

recent analyses of the satellite-based stratospheric climate data record (see Seidel et al., 945 2016 and references therein). There are also episodic temperature increases associated 946 with the two large volcanic eruptions, El Chichon in 1982 and Pinatubo in 1991. In 047 the upper stratosphere, several discontinuities can be seen. There is a marked decrease 948 in temperature near 1 hPa in 1995 when the transition from assimilating NOAA-11 949 to NOAA-14 SSU channel 3 radiances occurs. The latter are demonstrably cooler (see 950 Figure 16 of McCarty et al. 2016) and are assimilated without bias correction because of 951 the relatively large model errors at this level. There is an overall increase in temperature 952 when AMSU-A data are first assimilated in 1998, which was not as apparent in MERRA 953 (Rienecker et al. 2011) due to the overlapping use of SSU channel 3 and AMSU-A 954 channel 14 radiances in that reanalysis. The overall effect of assimilating the MLS 955 temperature profiles beginning in 2004 is to sharpen the stratopause with warming at 956 approximately 1 hPa and cooling above and below this level. 957

958 b. Ozone

The most notable aspects of the MERRA-2 ozone analysis, and those that constitute the main differences with MERRA, are the use of the improved version of SBUV data prior to October 2004 and subsequent assimilation of OMI and MLS observations. The latter provides high vertical resolution (~ 2.5 km) measurements of stratospheric ozone profiles during both night and day. The specification of background errors for ozone has also been upgraded to account for flow dependent error standard deviations as described in Wargan et al. (2015).

Many ozone data sets exist for various periods between 1980 and present. The decision to use only SBUV, MLS and OMI observations in MERRA-2 was motivated by the desire to avoid introducing multiple discontinuities into the ozone observing system while taking advantage of high-quality data offered by SBUV and EOS Aura retrievals. This approach leads to a relatively homogeneous MERRA-2 ozone record with only one major discontinuity in 2004 when MLS and OMI data replace SBUV observations. The price is a degraded quality of the analyzed ozone during the short periods when the selected data are not available, most notably in the Southern Hemisphere in late 1994, as discussed below.

An initial evaluation of the representation of ozone in MERRA-2 was presented in 975 Bosilovich et al (2015). A more comprehensive validation against independent satel-976 lite and ozonesonde data, including evaluation of the vertical structure and variability, 977 is given in Wargan et al. (2017). In particular, it is shown there that the assimilation of 978 MLS observations in MERRA-2 leads to significant improvements in the representation 979 of lower stratospheric ozone when compared with MERRA or compared with the period 980 of SBUV assimilation in MERRA-2. The QBO signal in ozone is discussed in Coy et al. 981 (2016), who demonstrate an improvement in the vertical structure of the ozone QBO 982 signature from 2004 onward, when MLS data are assimilated in MERRA-2. The focus 983 here is on the Antarctic total column ozone in order to illustrate that MERRA-2 has 984 realistic climatic ozone in a poorly observed region, while also highlighting some of its 985 uncertainties. Two examples are presented: a comparative evaluation of the South Pole 986 ozone in MERRA and MERRA-2 and the representation of Antarctic ozone holes in the 987 present reanalysis. The former follows Wargan et al. (2017). 988

Figure 23a shows the time series of total ozone derived from ozonesonde measurements at the South Pole, along with MERRA and MERRA-2 output sampled at the ozonesonde times and location between 1986 and 2015. The ozonesonde data, including the integrated column values were obtained from the Earth System Research Laboratory website (http://www.esrl.noaa.gov/gmd/ozwv/ozsondes/spo.html). Note that the vertical range

of balloon-borne measurements typically does not extend to pressure levels above 10 hPa 994 and so the upper-stratospheric portion of the column is obtained by extrapolating the 995 mixing ratios from 7 hPa or from the highest observed altitude, whichever is lower. For 996 completeness, Figure 23a also shows the reanalysis data between 1980 and 1985. In 997 the absence of ozonesondes, the reanalyses are sampled four times monthly in one-week 998 intervals for that period. The differences between each reanalysis and the ozonesonde 999 values are plotted in Figure 23b. Overall, both reanalyses capture the annual cycle and 1000 much of the interannual variability observed in the ozonesonde data, although there 1001 are large discrepancies (greater than 50%) during austral summer months in MERRA-2 1002 prior to 2005 and in MERRA throughout the period of comparison. This is consis-1003 tent with the fact that the reanalyses are not constrained by SBUV data during polar 1004 night. In addition, in late 1994, the SBUV coverage was limited to latitudes north of 1005 approximately 30° S owing to an orbital drift of the NOAA-11 satellite, which left the 1006 middle and high southern latitudes unobserved in both reanalyses. Nonetheless, these 1007 differences are reduced in MERRA-2 compared to MERRA. MERRA-2 performs sig-1008 nificantly better than MERRA relative to the South Pole ozonesondes from October 1009 2004 onward, when EOS Aura ozone data are assimilated. In particular, the standard 1010 deviation of the differences between MERRA-2 and the ozonesonde values drops from 1011 12.5% between 1991 and 2004 to 5% between 2005 and 2014. At the same time, the 1012 correlation between MERRA-2 and the ozonesonde measurements increases from 0.88 1013 to 0.98. The large excursions seen in Figure 23b in MERRA between 2008 and 2012 are 1014 due to degraded coverage of the NOAA-17 SBUV instrument. In contrast, the behavior 1015 of the MERRA-2 South Pole ozone is remarkably steady relative to the ozonesondes in 1016 the period when MLS and OMI data are assimilated. Only small seasonal variations 1017 are seen during that period. The MERRA-2 South Pole total ozone exhibits a small 1018 negative bias of approximately 6.7 Dobson units (DU), or roughly 2%, throughout the 1019

period of comparison. This bias does not vary significantly between the periods when
either SBUV or EOS Aura ozone data are assimilated.

As discovered by Molina and Rowland (1974), anthropogenic emissions of chlorofluo-1022 rocarbons provide the main contribution to the chlorine loading in the stratosphere. 1023 leading to destruction of the ozone layer. One prominent feature of the ozone loss in 1024 recent decades is the occurrence of springtime ozone holes over Antarctica since the 1025 early 1980's (Farman et al. 1985). Ozone holes are regions of extremely low values of 1026 total ozone forming inside the polar vortex due to a series of chlorine-catalyzed reactions 1027 (WMO 2014). The climatological importance of this phenomenon warrants its accurate 1028 representation in long-term reanalyses. The discussion here focuses on only one simple 1029 diagnostic, the ozone hole area, defined as the region with total ozone values less than 1030 220 DU. 1031

Figure 24 shows the time series of the ozone hole area calculated from the MERRA-2 1032 total ozone averaged between 20 September and 10 October in each year between 1980 1033 and 2015. Also plotted in Figure 24 are the ozone hole area values derived from the Total 1034 Ozone Mapping Spectrometer (TOMS) instruments on Nimbus-7 (1980–1992), Meteor-3 1035 (1992–1994) and Earth Probe (1996–2005), and from OMI (2004–2015). Note that OMI 1036 data are assimilated in MERRA-2 but TOMS observations are not. With the excep-1037 tion of 1994 there is remarkable agreement between MERRA-2 and these observations. 1038 In particular, MERRA-2 realistically captures the ozone hole interannual variability 1039 throughout the period of the reanalysis. There is an upward trend between 1980 and 1040 the mid-1990s followed by a plateau with the area oscillating around 22×10^6 km². 1041 This is consistent with the late twentieth century increase of anthropogenic chlorine and 1042 bromine loadings and the subsequent slow recovery after the gradual implementation of 1043 the Montreal Protocol of 1986 (WMO 2014). The Protocol, which went into effect in the 1044

late 1990s, banned the release of the main ozone depleting substances. Because the rate of the springtime polar ozone depletion depends on temperature and the strength of the Antarctic polar vortex in a given year, the size of the ozone hole exhibits a dynamically driven interannual variability superimposed on decadal-scale trends. This dynamical modulation is also evident in Figure 24. The extremely small (less than 3×10^6 km²) ozone hole in 2002 occurred in conjunction with the only major sudden stratospheric warming in the Southern Hemisphere on record (Newman and Nash 2005).

It should be noted that the southern high-latitude ozone for 1994 in MERRA-2 is not 1052 recommended for scientific use. The degraded result for that year is due to limited 1053 SBUV data coverage, as explained above, and the decision not to use data sources other 1054 than SBUV, OMI and MLS throughout the reanalysis. This particular deficiency is 1055 not shared with other major reanalyses (except MERRA), which replaced the missing 1056 data with other available observations such as from the short-lived Meteor-3 TOMS 1057 instrument (ERA-Interim) or NOAA-9 SBUV (CFSR and JRA-55). The latter were not 1058 considered in MERRA-2 because of the poorer quality of its partial columns compared 1059 to other SBUV instruments. 1060

Realistic ozone hole interannual variability is also present in MERRA (Sean M. Davis, personal communication 2016) with the exception of 1993, 1994 (as in MERRA-2), and the period between 2010 and 2012 when poor coverage from NOAA-17 SBUV resulted in degraded quality of the Antarctic ozone. The inferior performance of MERRA in 1993 compared to MERRA-2 is a consequence of applying more stringent data quality criteria to the older version of the SBUV data, resulting in limited data coverage near the terminator.

1068 7. Representation of the cryosphere

Reanalyses provide a global context for assessing recent, pronounced high latitude cli-1069 mate variability and provide seamless information on linkages to lower latitudes. As 1070 compared to midlatitudes, reanalyses in polar regions are particularly challenged by the 1071 paucity of the *in-situ* observational network, by the difficulty of satellite microwave and 1072 infrared sensors to profile the lower atmosphere over snow and ice surfaces, and by an 1073 inadequate representation of physical processes in models that are specific to these areas. 1074 Of these three challenges, improvement of model representations of physical processes— 1075 particularly as they relate to ice and snow surfaces—was seen as the most tractable in 1076 the development of MERRA-2. 1077

Several changes in the representation of physical processes between MERRA and MERRA-2 are directly relevant to polar regions. These include the use in MERRA-2 of the cubed-sphere computational grid (e.g., Putman and Lin 2007), which removes the need for gravity wave filtering at high latitudes, as well as daily sea ice concentration and sea surface temperature boundary conditions (Donlon et al. 2012; Reynolds et al. 2007; Taylor et al. 2000), as compared with the weekly fields used in MERRA.

In MERRA, a fixed surface albedo of 0.6 was used with sea-ice cover. This resulted in 1084 erroneously warm surface temperatures in the Arctic spring, when the observed albedo 1085 is typically much higher (Cullather and Bosilovich 2012). In MERRA-2, Northern Hemi-1086 sphere sea-ice albedo varies seasonally based on flux tower observations from the Surface 1087 Heat Budget of the Arctic Ocean (SHEBA) field experiment (Duynkerke and de Roode 1088 2001). Monthly values are computed and then linearly interpolated in time to produce 1089 instantaneous values. Sea-ice albedo in the Southern Hemisphere remains fixed as in 1090 MERRA, as there are few reliable albedo observations there. Sea ice in the Southern 1091

Hemisphere also does not endure an extended period of surface melting and a resulting
decreased albedo as in the Northern Hemisphere. Comparisons with SHEBA observations indicate a substantial reduction in 2-m air temperature biases during boreal spring
in MERRA-2.

These comparisons also find a warm bias in winter months over sea ice in MERRA-2 1096 of approximately 1.2°C in comparison to SHEBA. Larger air temperature differences 1097 of greater than 3°C are found in comparison to Soviet ice drifting station observations 1098 made during the 1980's (Colony et al. 1992). Simmons et al. (2016) showed that 1099 MERRA-2 is an outlier in near-surface temperature trends in polar regions as compared 1100 to ERA-Interim, JRA-55, and several conventional data sets. For the period 1980-1101 2009, annual 2-m air temperatures for the north polar cap bounded by 60° N increased 1102 by 0.35 ± 0.08 °C per decade in MERRA-2. This is the trend determined from linear 1103 regression; the uncertainty denotes the standard error of the trend. By comparison, 1104 north polar cap temperatures increased by $0.46 \pm 0.09^{\circ}$ C per decade in NOAA CFSR, 1105 by $0.55 \pm 0.10^{\circ}$ C per decade in ERA-Interim, and by $0.56 \pm 0.09^{\circ}$ C per decade in JRA-1106 55. The behavior in MERRA-2 may be attributable to spurious changes in the SST and 1107 SIC boundary conditions and the response of the model to changes in surface forcing. 1108 Investigation of these issues is ongoing. 1109

A particular focus during the development of MERRA-2 was on the representation of glaciated land surfaces (Cullather et al. 2014). In MERRA, ice sheets had an unrealistic design, with a fixed surface albedo and no representation of surface hydrology. Surface energy fluxes were computed using a fixed sub-surface temperature of 230 K (-43° C). In MERRA-2, energy conduction properties of the upper 15 meters of ice are represented, as well as the energy and hydrologic properties of an overlying, variable snow cover. Snow hydrology follows a modified version of the Stieglitz model that is also used over terrestrial land surfaces (Lynch-Stieglitz 1994; Stieglitz et al. 2001). This provides an explicit representation of snow densification, meltwater runoff, percolation, refreezing, and a prognostic surface albedo based on Greuell and Konzelmann (1994).

Figure 25 shows the effects of the different surface configurations in MERRA and 1120 MERRA-2 on near-surface air temperatures over ice sheets. In MERRA, biases are 1121 found when the observed surface temperature differs markedly from the fixed sub-surface 1122 temperature of -43° C. This includes South Pole station in winter (Figure 25a), where 1123 MERRA values are more than 5 K too warm; over the central Ross Ice Shelf in summer 1124 (Figure 25b), where MERRA is 8 K too cold; and over central Greenland in summer 1125 (Figure 25c), where MERRA is 4 K too cold. It may be seen from Figure 25 that these 1126 seasonal air temperature differences between MERRA and the station values are signifi-1127 cant over interannual time periods. In contrast, 2-m air temperatures for these locations 1128 in MERRA-2 more closely agree with the observed values. 1129

The surface representation in MERRA-2 also allows for the computation of surface 1130 mass balance over ice sheets, which may be defined as the net of precipitation minus 1131 evaporation minus runoff. The MERRA system does not provide runoff over land ice 1132 and, as seen in Figure 26, lacks ablation areas (in which the annual surface mass balance 1133 is negative) along the periphery of the ice sheet. For Greenland these occur mostly as a 1134 result of runoff from surface melt. The corresponding fields in MERRA-2, on the other 1135 hand, compare well with those from the widely-used Modèle Atmosphérique Régional 1136 regional climate model (MAR; Fettweis 2007), particularly in terms of the accumulation 1137 distribution in southeastern and western Greenland and the location of the zero-contour 1138 line along the western coast. However, some differences are also evident. For example, 1139 the regional climate model indicates average annual mean ablation values of up to 4 m 1140 yr^{-1} in southwestern Greenland, as compared with values of approximately 1 m yr^{-1} in 1141

MERRA-2. In addition to differing surface representations, differences in grid spacing between MAR (25 km) and MERRA-2 (roughly 50 km) may also play a role. A final point of comparison in Figure 26 is with regard to topography. The MERRA system used a dated topography which contained large errors of up to 600 m over the Greenland Ice Sheet (Box and Rinke 2003). These differences are apparent in the topography contours shown for MERRA and MERRA-2 in Figure 26.

1148 8. MERRA-2 products and access

The complete list of analyzed and diagnosed fields produced by MERRA-2 is given in 1149 the product file specification document available at the GMAO's MERRA-2 web site 1150 (https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf). The GEOS IAU procedure 1151 allows for higher-frequency products than just the 6-hourly ones generated directly from 1152 the analysis. There are three time-invariant and 39 time-varying product collections, 1153 all produced on a $0.625^{\circ} \times 0.5^{\circ}$ horizontal grid. Variables are provided on either the 1154 native vertical grid (at 72 model layers or the 73 edges), or interpolated to 42 standard 1155 pressure levels. Detailed information and a description of each variable are available in 1156 the MERRA-2 file specification document. As in MERRA, MERRA-2 provides closed 1157 atmospheric budgets, including the analysis increment terms. The observational forcing 1158 from the assimilation increments during the IAU segment is summed in the output 1159 budgets of the model. Bosilovich et al. (2015) show the magnitudes of these terms in 1160 water and energy budgets. 1161

The NASA Goddard Earth Sciences Data Information Services Center (GES DISC) provides access to MERRA-2 products through a new unified user interface connected to three different search engines. Many of the tools will be familiar to MERRA users, such as the popular Giovanni visualization and analysis tool, web based FTP servers

and OpenDAP web services. The subsetting capability has been updated to include 1166 grid transformation options, while retaining the essential functionality of selecting lev-1167 els, variables, time and domain. Citations for the individual MERRA-2 data collections 1168 are included in the GES DISC MERRA-2 data access pages. As noted in section 1, 1169 these citations are included in the figure captions of this paper (except where results for 1170 MERRA-2 are derived from other sources such as diagnostic output from the data assim-1171 ilation scheme). Results shown for MERRA are from similarly named data collections, 1172 as described by Rienecker et al. (2011). 1173

1174 9. Summary and outlook

The Modern Era Retrospective Analysis for Research Applications Version 2 (MERRA-1175 2) was developed with two primary objectives: to provide an ongoing near-real time cli-1176 mate analysis of the satellite era that addresses known limitations of the now-completed 1177 MERRA reanalysis (January 1979–February 2016), and to demonstrate progress toward 1178 development of a future integrated Earth system analysis (IESA) capability. MERRA-2 1179 has achieved those objectives in several respects. These include the assimilation of satel-1180 lite observations not available to MERRA—which assimilated no new satellite observa-1181 tions after NOAA-18 (launched in 2005)—the reduction of certain biases and imbalances 1182 in the water cycle, and the reduction of spurious trends and jumps in precipitation related 1183 to changes in the observing system. As a step toward a future IESA, MERRA-2 includes 1184 aerosol data assimilation and improved representations of aspects of the cryosphere and 1185 stratosphere, including ozone, as compared with MERRA. 1186

At the same time, because of the fairly rapid development schedule required to produce a timely replacement for MERRA, other aspects of the MERRA-2 development received less attention. For example, there was little focus on the preparation and improvement of

input conventional data types and minimal tuning of the model physics for the current 1190 application. Notable shortcomings of MERRA-2 compared with MERRA include an 1191 increased warm bias in the upper troposphere—as revealed by the background forecast fit 1192 to radiosonde temperature observations and mean analysis increments of temperature— 1193 as well as excessive precipitation over high topography in the tropics and, to a lesser 1194 extent, over northern high latitudes. Subsequent experimentation indicates that these 1195 behaviors are most affected by the model parameterizations of deep convection and 1196 gravity wave drag in GEOS, as well as the representation of topography. They are being 1197 addressed in more recent model versions. 1198

Ongoing development in other aspects of modeling and data assimilation are likely to pro-1199 vide benefit for reanalyses in the near future. For example, while MERRA-2 assimilates 1200 only clear-sky satellite radiances, the use of cloud- and rain-affected radiances—referred 1201 to as all-sky assimilation (Bauer et al. 2010)—has matured or become operational at 1202 several centers including GMAO. This should improve the assimilation of moisture-1203 sensitive data types which, as shown here and by Bosilovich et al. (2017), can still 1204 induce unexpected changes in global precipitation and moisture fields. Direct assimi-1205 lation of land surface observations, including remotely sensed soil moisture and snow 1206 cover fraction, is another area of improving capability that is likely to provide bene-1207 fit to reanalysis, especially for capturing extreme events like droughts and heat waves. 1208 Implementation of an improved land model that includes dynamic phenology and pho-1209 tosynthesis is a key component of the GMAO's land surface modeling and assimilation 1210 efforts (Koster et al. 2014). To improve the specification of ocean surface boundary 1211 conditions, many centers are developing some form of coupled ocean-atmosphere anal-1212 vsis system. The GMAO has recently implemented a coupled data assimilation scheme 1213 for analyzing ocean skin temperature within the existing atmospheric analysis (Akella 1214 et al. 2016). It uses background fields from a near-surface ocean diurnal layer model 1215

to assimilate surface-sensitive radiances plus in-situ observations along with all other observations in the atmospheric assimilation system. The scheme may be described as being weakly coupled in the sense that the atmospheric observations do not affect the ocean fields directly, but only through the increment of ocean skin temperature during the next analysis cycle.

Improving the representation of aerosol effects on climate is another important area of 1221 development for reanalysis. As the aerosol observing system continues to evolve and 1222 provide additional global information on aerosol absorption, size and vertical distribu-1223 tion, the discrepancy among reanalyses and satellite-only estimates of aerosol radiative-1224 climate effects should decrease. For example, the GMAO is working to incorporate 1225 aerosol vertical distribution information from space-based lidars, as well as implicit spe-1226 ciation and size information from multi-channel radiometers on low-orbiting and geosta-1227 tionary satellites. Unlike satellite estimates alone, reanalyses like MERRA-2 can provide 1228 detailed information on how the anthropogenic component of aerosols, and thus radia-1229 tive forcing, has changed during the modern satellite era, as well as its interaction with 1230 the circulation and the climate at large. This should lead to reduced uncertainty in 1231 assessing, for example, the human impact on climate. 1232

More extensive analysis coupling between the atmosphere, ocean, land and chemistry as 1233 envisioned for IESA, while progressing, still presents significant challenges (e.g., Brass-1234 ington et al. 2015). These include model biases that can be exacerbated when coupled, 1235 component systems with different physical characteristics and different spatial and tem-1236 poral scales, and component observations in different media with different spatial and 1237 temporal frequencies and different latencies. These challenges may be offset at least 1238 partially by the fact that, in practice, where the time scales and observation laten-1239 cies between components differ greatly—as between the deep ocean and atmosphere for 1240

example—a weak coupling approach may suffice. Prospects for success are also bolstered by the fact that the numerical weather prediction community is placing increasing focus on the need to analyze currently uncoupled components of the Earth system in a more consistent manner. The GMAO strategy is to progress incrementally toward an IESA through an evolving combination of coupled systems and offline component reanalyses driven by, for example, MERRA-2 atmospheric forcing.

Quantifying uncertainty in reanalyses remains important for expanding their utility, 1247 especially as a potential tool for climate change assessment. Dee et al. (2011) argued 1248 that advances in observational bias correction and other aspects of data assimilation 1249 have reduced uncertainty in the representation of low-frequency variability to the point 1250 where ERA-Interim can be used to estimate certain atmospheric temperature trends. 1251 More recently, Simmons et al. (2014) compared multi-annual variability and trends 1252 in atmospheric temperature from ERA-Interim, JRA-55 and MERRA and found them 1253 to be in generally good agreement in the upper troposphere and lower stratosphere 1254 but more uncertain in the middle stratosphere. Nonetheless, for less well constrained 1255 quantities such as precipitation and surface fluxes, there still appear to be substantial 1256 differences between recent reanalyses. For example, the 12-month running mean values 1257 of global precipitation in ERA-Interim, MERRA-2, and JRA-55 can at times differ by 1258 almost 20%. Uncertainty in sea surface temperature, as illustrated by the surprising 1259 differences between the prescribed values used in different reanalyses (Figure 5) is likely 1260 to be a contributing factor. Impacts from observing system changes also appear to play 1261 a significant role in explaining these precipitation differences, pointing to the need for 1262 new sources of high-quality observations of these or closely related variables not only 1263 for assimilation but for improving our understanding and modeling of the underlying 1264 physical processes. Ongoing efforts to improve the quality of existing historical data 1265 sets are also critical in this regard. 1266

The increasing use of ensemble and hybrid ensemble-variational methods in Earth sys-1267 tem data assimilation has the potential to make at least some measures of uncertainty a 1268 standard component of reanalysis data sets (e.g., Compo et al. 2011; Poli et al. 2013). 1269 The GMAO has recently implemented a hybrid four-dimensional ensemble-variational 1270 (4D-ENVAR) assimilation scheme with similar capability. Finally, ECMWF, JMA and 1271 GMAO are conducting multi-decadal atmospheric model integrations (without data as-1272 similation) for comparison with reanalyses as a means of assessing internal variability 1273 and distinguishing boundary-forced climate signals from those imposed by changes in 1274 the observing system. All these efforts will benefit from the continued assessment of ex-1275 isting reanalysis products by the research community, and from the sharing of key data 1276 assimilation diagnostic quantities (e.g., background departures, analysis increments, bias 1277 estimates) between both reanalysis developers and data providers. 1278

1279 Acknowledgments

Development of the GEOS data assimilation system and the MERRA-2 project were 1280 funded by NASA's Modeling Analysis and Prediction program. Computational resources 1281 and support for the execution of MERRA-2 were provided by the NASA High-End 1282 Computing Capability Project and NASA Center for Climate Simulation. GEOS and 1283 MERRA-2 are the result of years of dedicated research, development and analysis by 1284 many individuals at GMAO whose efforts are greatly appreciated. We gratefully ac-1285 knowledge the GMAO operations group for monitoring the production of MERRA-2 1286 and the GMAO software integration group who helped improve the performance and 1287 flexibility of GEOS. We thank Julio Bacmeister for his contribution to the development 1288 of the model physics, Qing Liu for processing the precipitation input data, and Allison 1289 Collow and Edmond Brent Smith for providing some of the figures for this paper. We 1290

- ¹²⁹¹ also thank the GES DISC for providing on-line access to MERRA-2 products. Finally,
- ¹²⁹² we thank the three reviewers, whose comments and suggestions helped improve the paper
- 1293 substantially.

1294 Appendix: Acronyms

3DVAR	Three-dimensional variational data assimilation		
4DENVAR	Four-dimensional ensemble-variational data assimilation		
AAOD	Aerosol absorption optical depth		
ACARS	Aircraft Communications Addressing and Reporting		
AeroCom	Aerosol Comparison Project		
AERONET	Aerosol Robotics Network		
AIREP	Aircraft report		
AIRS	Advanced Infrared Sounder		
AMDAR	Aircraft Meteorological Data Relay		
AMSR-E	Advanced Microwave Scanning Radiometer-EOS		
AMSU-A	Advanced Microwave Sounding Unit-A		
AMSU-B	Advanced Microwave Sounding Unit-B		
AOD	Aerosol optical depth		
ASCAT	Advanced Scatterometer		
ASDAR	Aircraft to Satellite Data Relay		
ATMS	Advanced Technology Microwave Sounder		
ATOVS	Advanced TIROS Operational Vertical Sounder		
AVHRR	Advanced Very High Resolution Radiometer		
CFSR	Climate Forecast System Reanalysis		
CAMS	Copernicus Atmosphere Monitoring Service		
CMIP	Coupled Model Intercomparison Project		
CPC	Climate Prediction Center		
CrIS	Cross-track Infrared Sounder		
CRTM	Community Radiative Transfer Model		

DMSP	Defense Meteorological Satellite Program	
DRE	Direct radiative effect	
ECMWF	European Centre for Medium-Range Weather Forecasts	
EOS	Earth Observing System	
ERA-20C	ECMWF Reanalysis from 1900–2010	
ERA-Interim	ECMWF Reanalysis from 1979–present	
ERS	Environmental Research Satellite	
FGAT	First guess at appropriate time	
GEOS	Goddard Earth Observing System	
GES DISC	SC Goddard Earth Sciences Data Information Services Center	
GLATOVS Goddard Laboratory for Atmospheres TOVS forward model		
GMAO	Global Modeling and Assimilation Office	
GMS	Geostationary Meteorological Satellite	
GOCART	GOCART Goddard Chemistry, Aerosol, Radiation and Transport mode	
GOES	DES Geostationary Operational Environmental Satellites	
GPCP	Global Precipitation Climatology Project	
GPSRO	Global Positioning System radio occultation	
GSI	Gridpoint Statistical Interpolation	
GWD	Gravity wave drag	
HIRS	IIRS High-resolution Infrared Radiation Sounder	
IASI	I Infrared Atmospheric Sounding Interferometer	
IAU	Incremental analysis update	
IESA	Integrated Earth system analysis	
IPCC	Intergovernmental Panel on Climate Change	
JMA	Japan Meteorological Agency	

JPSS	Joint Polar Satellite System	
JRA-55	Japanese 55-year Reanalysis	
MACC	Monitoring Atmospheric Composition and Climate project	
MAR	Modèle Atmosphérique Régional regional climate model	
MDCRS	Meteorological Data Collection and Reporting System	
MERRA	Modern-Era Retrospective Analysis for Research and Applications	
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications, Version 2	
Metop	Meteorological Operational Satellite	
MHS	Microwave Humidity Sounder	
MISR	Multi-angle SpectroRadiometer	
MLS	Microwave Limb Sounder	
MODIS	Moderate Resolution Imaging Spectroradiometer	
MSG	Meteosat Second Generation satellite	
MSU	Microwave Sounding Unit	
MTSAT	Multifunctional Transport Satellite	
NAAPS	Navy Aerosol Analysis and Prediction System	
NASA	National Aeronautics and Space Administration	
NCEP	National Centers for Environmental Prediction	
NEXRAD	Next-Generation Radar	
NOAA	National Oceanic and Atmospheric Administration	
NRL	Naval Research Laboratory	
OISST	Optimum Interpolation Sea Surface Temperature	
OMI	Ozone Monitoring Instrument	
OPAC	Optical Properties of Aerosols and Clouds	
OSTIA	Operational Sea Surface Temperature and Sea Ice Analysis	
PAOB	Synthetic surface pressure observation	

Pibal	Pilot balloon	
PIREP	Pilot report	
QBO	Quasi-Biennial Oscillation	
QFED	Quick Fire Emission Dataset	
Raob	Radiosonde observation	
RAS	Relaxed Arakawa-Schubert convection scheme	
RMS	Root mean square	
RSS	Remote Sensing Systems	
SBUV	Solar Backscatter Ultraviolet Radiometer	
SEVIRI	Spinning Enhanced Visible Infrared Imager	
SHEBA	IEBA Surface Heat Budget of the Arctic Ocean	
SIC	Sea ice concentration	
SMAP	Soil Moisture Active Passive satellite	
SMOS	OS Soil Moisture and Ocean Salinity satellite	
SNPP	Suomi National Polar-orbiting Partnership	
SSM/I	Special Sensor Microwave Imager	
SSMIS	Special Sensor Microwave Imager/Sounder	
SST	Sea surface temperature	
SSU	Stratospheric Sounding Unit	
TIROS	Television Infrared Observation Satellite	
TLNMC	Tangent linear normal mode constraint	
TMI	Tropical Rainfall Measuring Mission Microwave Image	
TOA	Top of the atmosphere	
TOMS	Total Ozone Mapping Spectrometer	
VAD	Velocity Azimuth Display	
WMO	World Meteorological Organization	

1299 References

- Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Cimatology1301Project (GPCP) monthly precipitation analysis (1979–present). J. Hydrometeor.,13024, 1147–1167.
- Akella, S., R. Todling, M., and M. Suárez, 2016: Assimilation for skin SST in the NASA
 GEOS atmospheric data assimilation system. *Quart. J. Roy. Meteor. Soc.*,
 doi:10.1002/qj.2988.
- Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: Middle Atmosphere Dynamics.
 Academic Press, 489 pages.
- Bacmeister, J. T. and Stephens, G., 2011: Spatial statistics of likely convective clouds
 in CloudSat data. J. Geophys. Res., 116, D04104, doi:10.1029/2010JD014444.
- Ballish, B. A., and V. K. Kumar, 2008: Systematic differences in aircraft and radiosonde temperatures. *Bull. Amer. Meteor. Soc.*, **89**, 1689–1707.
- Bauer, P., A. J. Geer, P. Lopez, and D. Salmond, 2010: Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Quart. J. Roy. Meteor. Soc.*, 136, 1868–1885. doi:10.1002/qj.659
- Bellouin, N., J. Quaas, J.-J. Morcrette and O. Boucher, 2013: Estimates of aerosol radiative forcing from the MACC re-analysis. Atmos. Chem. Phys., 13, 2045–2062, doi:10.5194/acp-13-2045-2013.
- Berrisford, P., P. Kallberg, S. Kobayashi, D. Dee, S. Uppala, A. J. Simmons, P. Poli, and H. Sato, 2011: Atmospheric conservation properties in ERA-Interim. *Quart. J. Roy. Meteor. Soc.*, **137**, 1381–1399.
- ¹³²¹ Bloom, S., L. Takacs, A. DaSilva, and D. Ledvina, 1996: Data assimilation using incre-¹³²² mental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.
- Bocquet M., and Coauthors, 2015:, Data assimilation in atmospheric chemistry models:
 Current and future prospects for coupled chemistry meteorology models. Atmos.
 Chem. Phys., 15 (10), 5325–5358, doi:10.5194/acp-15-5325-2015.
- Bosilovich, M. G., 2013: Regional climate and variability in NASA MERRA and recent reanalyses: US summertime precipitation and temperature, J. Appl. Meteor. *Climatol.*, 52, 1939–1951, doi: http://dx.doi.org/10.1175/JAMC-D-12-0291.1.
- ¹³²⁹ Bosilovich, M.G., F. R. Robertson, and J. Chen, 2011: Global energy and water budgets

- Bosilovich, M.G., and Coauthors, 2015: MERRA-2: Initial Evaluation of the Climate. NASA/TM2015104606, Vol. 43, 139 pp.
- https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf.
- Bosilovich, M., F. Robertson, L. Takacs, A. Molod, and D. Mocko, 2017: Atmospheric
 water balance and variability in the MERRA-2 reanalysis. J. Climate, 30, 1177–
 1196, doi: 10.1175/JCLI-D-16-0338.1.
- Box, J. E., and A. Rinke, 2003: Evaluation of Greenland Ice Sheet surface climate in the HIRHAM regional climate model using automatic weather station data. J. *Climate*, 16, 1302-1319, doi:10.1175/1520-0442-16.9.1302.
- Brassington, G. B., M. J. Martin, H. L. Tolman, S. Akella, M. Balmeseda, C. R. S. Chambers, J. A. Cummings, Y. Drillet, P. A. E. M. Jansen, P. Laloyaux, D. Lea, A.
 Mehra, I. Mirouze, H. Ritchie, G. Samson, P. A. Sandery, G. C. Smith, M. Suárez,
 and R. Todling, 2015: Progress and challenges in short- to medium- range coupled
 prediction, J. Op. Oceanogr., 8, 239–258, doi:10.1080/1755876X.2015.1049875.
- Buchard, V., and Coauthors, 2015: Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA Aerosol Reanalysis. Atmos. Chem.
 Phys., 15 (10), 5743–5760, 10.5194/acp-15-5743-2015.
- Buchard, V., C. A. Randles, A. M. da Silva, A. Darmenov, P. R. Colarco, R. Govindaraju, R. Ferrare, J. Hair, A. J. Beyersdorf, L. D. Ziemka, and H. Yu, 2017: The MERRA-2 Aerosol Reanalysis, 1980–onward, Part 2: Evaluation and case studies. J. Climate, in review.
- Cardinali, C., L. Isaksen, and E. Anderson, 2003: Use and impact of automated aircraft
 data in a global 4DVAR data assimilation system. *Mon. Wea. Rev.*, 131, 1865–
 1877.
- ¹³⁵⁵ Chen Y., F. Weng, Y. Han, and Q. Liu, 2008: Validation of the community radiative transfer model (CRTM) by using CloudSat Data. J. Geophys. Res., 113 (D8), 2156–2202.
- Chin, M., P, Ginoux, S. Kinne, O. Torres, B. N. Holben, B. N. Duncan, R. V. Martin,
 J. A. Logan, A. Higurashi, and T. Nakajima, 2002: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and
 sun photometer measurements. J. Atmos. Sci., 59, 461–483, 10.1175/15200469(2002)059 <0461:TAOTFT>2.0.CO;2,
- http://dx.doi.org10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2.

- 1364 Chýlek, P., and J. A. Coakley, 1974: Aerosol and climate. Science, 183, 75–77.
- Colarco, P., A. da Silva, M. Chin, and T. Diehl, 2010: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. J. Geophys. Res., **115** (D14207), 10.1029/2009JD012820, http://dx.doi.org/10.1029/2009JD012820.
- Collow, A. B. M., M. G. Bosilovich, and R. D. Koster, 2016: Large scale influences on
 summertime extreme precipitation in the northeastern United States. To appear
 in J. Hydromet., doi: 10.1175/JHM-D-16-0091.1.
- ¹³⁷² Collow, A. B. M., and M. A. Miller, 2016: The seasonal cycle of the radiation budget
 ¹³⁷³ and cloud radiative effect in the Amazon rainforest of Brazil. J. Climate, doi:
 ¹³⁷⁴ 10.1175/JCLI-D-16-0089.1.
- ¹³⁷⁵ Colony, R., I. Appel, and I. Rigor, 1992: Surface air temperature observations in the
 ¹³⁷⁶ Arctic Basin. Tech. Memo. TM 1-92, 120 pp. Available from Applied Physics
 ¹³⁷⁷ Laboratory, University of Washington, Seattle, WA 98195.
- ¹³⁷⁸ Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. Quart.
 ¹³⁷⁹ J. Roy. Meteor. Soc., 137, 1–28, doi:10.1002/qj.776.
- Coy, L., K. Wargan, A. M. Molod, W. R. McCarty, and S. Pawson, 2016: Structure and dynamics of the quasi-biennial oscillation in MERRA-2. J. Climate, 29, 5339–5354, doi:10.1175/JCLI-D-15-0809.1.
- ¹³⁸³ Cullather, R. I., and M. G. Bosilovich, 2012: The energy budget of the polar atmosphere ¹³⁸⁴ in MERRA. J. Climate, **25**, 5–24, doi:10.1175/2011JCLI4138.1.
- ¹³⁸⁵ Cullather, R.I., S.M.J. Nowicki, B. Zhao, and M. J. Suárez, 2014: Evaluation of the
 ¹³⁸⁶ surface representation of the Greenland Ice Sheet in a general circulation model.
 ¹³⁸⁷ J. Climate, 27, 4835–4856, doi: 10.1175/JCLI-D-13-00635.1.
- Darmenov, Anton, and Arlindo da Silva, 2015. The Quick Fire Emissions Dataset
 (QFED): Documentation of versions 2.1, 2.2 and 2.4. NASA/TM2015104606,
 Vol. 38, 201 pp.
- Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. Zeng, and M. G. Bosilovich, 2011:
 Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux
 tower observations. J. Climate, 24, 221–249, doi:10.1175/JCLI-D-11-00004.1.
- ¹³⁹⁴ Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and per-¹³⁹⁵ formance of the data assimilation system. Q. J. R. Meteorol. Soc., **137**, 553–597,

doi:10.1002/qj.828.

1396

- ¹³⁹⁷ Dee D. P., and A. M. da Silva, 2003: The choice of variable for atmospheric moisture ¹³⁹⁸ analysis. *Mon. Weather Rev.*, **131** 155–171.
- ¹³⁹⁹ Dee, D, and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.*, **135**, 1830–1841.
- ¹⁴⁰¹ Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the
 ¹⁴⁰² NCEP SSI analysis system. Mon. Wea. Rev., **126**, 2287–2299.
- Diehl, T., A. Heil, M. Chin, X. Pan, D. Streets, M. Schultz, and S. Kinne, 2012: Anthropogenic, biomass burning, and volcanic emissions of black carbon, organic
 carbon, and SO2 from 1980 to 2010 for hindcast model experiments. Atmos. *Chem. Phys. Discuss.*, 12 (9), 24 895–24 954, 10.5194/acpd-12-24895-2012,
 http://www.atmos-chem-phys-discuss.net/12/ 24895/2012/.
- Donlon, C.J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012:
 The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system.
 Remote Sens. Environ., 116, 140–158, doi:10.1016/j.rse.2010.10.017.
- ¹⁴¹¹ Draper, C., R. Reichle, and R. Koster, 2017: Assessment of the MERRA-2 land surface
 ¹⁴¹² energy flux estimates. J. Climate, in review.
- ¹⁴¹³ Duynkerke, P., and S. de Roode, 2001: Surface energy balance and turbulence characteristics observed at the SHEBA Ice Camp during FIRE III. J. Geophys. Res.,
 ¹⁴¹⁵ 106, 15313–15322, doi:10.1029/2000JD900537.
- 1416Farman, J., B. Gardiner, B., and J. Shanklin, 1985: Large losses of total ozone in1417Antarctica reveal seasonal ClOx/NOx interaction Nature, **315** (6016), 207–210,1418doi: 10.1038/315207a0.
- Fettweis, X., 2007: Reconstruction of the 1979–2006 Greenland ice sheet surface mass balance using the regional climate model MAR. *The Cryosphere*, **1**, 21–40, doi:10.5194/tc-1-21-2007.
- Flemming, J., A. Benedetti, A. Inness, R. Engelen, L. Jones, V. Huijnen, S. Remy, M.
 Parrington, M. Suttie, A. Bozzo, V.-H. Peuch, D. Akritidis, and E. Katragkou,
 2017: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for
 2003–2015, Atmos. Chem. Phys., 17, 1945–1983, doi:10.5194/acp-17-1945-2017.
- Froidevaux, L., and Coauthors, 2006: Early validation analyses of atmospheric profiles
 from EOS MLS on the Aura satellite. IEEE Transactions on Geoscience and

- ¹⁴²⁸ Remote Sensing 44, no. 5, doi:10.1109/TGRS.2006.864366.
- Global Modeling and Assimilation Office (GMAO), 2015a: MERRA-2 inst1_2d_asm_Nx:
 2d, 3-Hourly, Instantaneous, Single-Level, Assimilation, Single-Level Diagnostics
 V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information
 Services Center (GES DISC), accessed June 2016, doi:10.5067/3Z173KIE2TPD.
- Global Modeling and Assimilation Office (GMAO), 2015b: MERRA-2 inst1_2d_int_Nx:
 2d,1-Hourly, Instantaneous, Single-Level, Assimilation, Vertically Integrated Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed June 2016,
 doi:10.5067/G0U6NGQ3BLE0.
- Global Modeling and Assimilation Office (GMAO), 2015c: MERRA-2 inst3_3d_asm_Np:
 3d, 3-Hourly, Instantaneous, Pressure-Level, Assimilation, Assimilated Meteorological Fields, V5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center
 Distributed Active Archive Center (GSFC DAAC), accessed June 2016,
 doi:10.5067/QBZ6MG944HW0.
- Global Modeling and Assimilation Office (GMAO), 2015d: MERRA-2 tavg1_2d_flx_Nx:
 2d, 1-Hourly, Time-Averaged, Single-Level, Assimilation, Surface Flux Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information
 Services Center (GES DISC), accessed June 2016, doi:10.5067/7MCPBJ41Y0K6.
- Global Modeling and Assimilation Office (GMAO), 2015e: MERRA-2 tavg1_2d_int_Nx:
 2d, 1-Hourly, Time-Averaged, Single-Level, Assimilation, Vertically Integrated
 Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and
 Information Services Center (GES DISC), accessed June 2016,
 doi:10.5067/Q5GVUVUIVGO7.
- Global Modeling and Assimilation Office (GMAO), 2015f: MERRA-2 tavg1_2d_slv_Nx:
 2d, 1-Hourly, Time-Averaged, Single-Level, Assimilation, Single-Level Diagnos tics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information
 Services Center (GES DISC), accessed June 2016, doi:10.5067/VJAFPLI1CSIV.
- Global Modeling and Assimilation Office (GMAO), 2015g: MERRA-2 tavgM_2d_aer_Nx:
 2d, Monthly mean, Time-averaged, Single-Level, Assimilation, Aerosol Diagnostics V5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center Distributed
 Active Archive Center (GSFC DAAC), accessed June 2016,
 doi:10.5067/FH9A0MLJPC7N.
- Global Modeling and Assimilation Office (GMAO), 2015h: MERRA-2 tavgM_2d_flx_Nx:
 2d, Monthly mean, Time-Averaged, Single-Level, Assimilation, Surface Flux Di-

1463 1464 1465	agnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed June 2016, doi:10.5067/0JRLVL8YV2Y4.
1466 1467 1468 1469	Global Modeling and Assimilation Office (GMAO), 2015i: MERRA-2 tavgM_2d_glc_Nx: 2d, Monthly mean, Land Ice Surface Diagnostics, V5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), accessed June 2015, doi:10.5067/5W8Q3I9WUFGX.
1470 1471 1472 1473	Global Modeling and Assimilation Office (GMAO), 2015j: MERRA-2 tavgM_2d_lnd_Nx: 2d, Monthly mean, Land Surface Diagnostics, V5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), accessed June 2015, doi:10.5067/8S35XF81C28F.
1474 1475 1476 1477 1478	Global Modeling and Assimilation Office (GMAO), 2015k: MERRA-2 tavgM_2d_int_Nx: 2d, Monthly mean, Time-Averaged, Single-Level, Assimilation, Vertically Integrated Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed June 2015, doi:10.5067/FQPTQ4OJ22TL.
1479	Global Modeling and Assimilation Office (GMAO), 2015l: MERRA-2 tavgM_3d_qdt_Np:
1480	3d, Monthly mean, Time-Averaged, Pressure-Level, Assimilation, Moist Tenden-
1481	cies V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Informa-
1482	tion Services Center (GES DISC), accessed June 2016,
1483	doi:10.5067/2ZTU87V69ATP.
1484	Global Modeling and Assimilation Office (GMAO), 2015m: MERRA-2 tavgM_2d_slv_Nx:
1485	2d, Monthly mean, Single-Level Diagnostics, V5.12.4, Greenbelt, MD, USA: God-
1486	dard Space Flight Center Distributed Active Archive Center (GSFC DAAC),
1487	accessed April 2015, doi:10.5067/AP1B0BA5PD2K.
1488	Global Modeling and Assimilation Office (GMAO), 2015n: MERRA-2 tavgM_3d_tdt_Np:
1489	3d, Monthly mean, Time-Averaged, Pressure-Level, Assimilation, Temperature
1490	Tendencies V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and
1491	Information Services Center (GES DISC), accessed June 2016,
1492	doi:10.5067/VILT59HI2MOY.
1493	Gong, S. L., 2003: A parameterization of sea-salt aerosol source function for sub- and
1494	super-micron particles. <i>Global Biogeochemical Cycles</i> , 17 (4),
1495	doi:10.1029/2003GB002079.
1496	Greuell, W., and T. Konzelmann, 1994: Numerical modelling of the energy balance
1497	and englacial temperature of the Greenland Ice Sheet. Calculations for the ETH-

- ¹⁴⁹⁸ Camp location (West Greenland, 1155m a.s.l.). *Global Planet. Change*, **9**, 91–114, doi:10.1016/0921-8181(94)90010-8.
- Ham, Y.-G., S. Schubert, Y. Vikhliaev, and M. J. Suárez, 2014: An assessment of the
 ENSO forecast skill of GEOS-5 system. *Clim. Dynam.*, doi:10.1007/s00382-0142063-2.
- Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006:
 JCSDA Community Radiative TransferModel (CRTM)–Version 1. NOAA Tech.
 Rep. 122, 33 pp.
- Heidinger, A. K., C. Cao, and J. T. Sullivan, 2002: Using Moderate Resolution Imaging
 Spectrometer (MODIS) to calibrate Advanced Very High Resolution Radiometer
 reflectance channels. J. Geophys. Res. Atmos., 107 (D23), 10.1029/2001JD002035,
 http://dx.doi.org/10.1029/2001JD002035.
- ¹⁵¹⁰ Hoch, S.W., 2005: Radiative flux divergence in the surface boundary layer. A study
 ¹⁵¹¹ based on observations at Summit, Greenland. Ph.D. dissertation, Swiss Federal
 ¹⁵¹² Institute of Technology (ETH), Zurich, 164 pp.
- Holben, B., and Coauthors, 1998: AERONET A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.*, 66 (1), 1–16, http://dx.doi.org/10.1016/S0034-4257(98)00031-5,
- http://www.sciencedirect.com/science/article/ pii/S0034425798000315.
- Holm, E. V., 2003: Revision of the ECMWF humidity analysis: Construction of a Gaussian control variable. In *Proceedings of Workshop on Humidity Analysis*, 1–3. ECMWF/GEWEX: Reading, UK.
- Inness, A., F. Baier, A. Benedetti, I. Bouarar, S. Chabrillat, H. Clark, H., and Coauthors,
 2013: The MACC reanalysis: An 8 yr data set of atmospheric composition. *Atmos. Chem. Phys.*, **13**, 4073–4109, doi:10.5194/acp-13-4073-2013.
- Kahn, R. A., B. J. Gaitley, J. V. Martonchik, D. J. Diner, K. A. Crean, and B. Holben,
 2005: Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth
 validation based on 2 years of coincident Aerosol Robotic Network (AERONET)
 observations. J. Geophys. Res. Atmos., 110 (D10), 10.1029/2004JD004706,
 URL http://dx.doi.org/10.1029/ 2004JD004706.
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009a:
 Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, 137, 1046–1060.

- ¹⁵³¹ Kleist D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009b:
 ¹⁵³² Introduction of the GSI into the NCEPs Global Data Assimilation System. Wea.
 ¹⁵³³ Forecasting, 24, 1691–1705.
- Kinne, S., and Coauthors, 2006:, An AeroCom initial assessment optical properties in
 aerosol component modules of global models. Atmos. Chem. Phys., 6, 1815–
 1834, doi:10.5194/acp-6-1815-2006.
- Kobayashi, S., M. Matricardi, D. Dee, and S. Uppala, 2009: Toward a consistent reanalysis of the upper stratosphere based on radiance measurements from SSU and
 AMSU-A. *Quart. J. Roy. Meteor. Soc.*, **135**, 2086–2099, doi:10.1002/qj.514.
- Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. J. Meteorol. Soc. Japan, 93 (1), 5–48. doi:10.2151/jmsj.2015-001.
- Koster, R. D., G. Walker, G. J. Collatz, and P. E. Thornton, 2014: Hydroclimatic controls on the means and variability of vegetation phenology and carbon uptake. *J. Climate*, 27, 5632–5652. doi: 10.1175/JCLI-D-13-00477.
- Kunkel, K. E., and Coauthors, 2013: Monitoring and understanding trends in extreme storms: State of knowledge. Bull. Amer. Meteor. Soc., 94, 499–514, doi:10.1175/BAMS-D-11-00262.1.
- Levy, R. C., L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman, 2007: Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Res- olution Imaging Spectroradiometer spectral reflectance. J. Geophys. Res. Atmos., **112** (**D13**), 10.1029/2006JD007811, URL http://dx.doi.org/10.1029/2006JD007811.
- Lim, Y.-K., R. Kovach, S. Pawson, and G. Vernieres, 2017: The 2015/2016 El Niño event in context of the MERRA-2 reanalysis: A comparison of the tropical Pacific with 1982/1983 and 1997/1998. J. Climate, doi:10.1175/JCLI-D-16-0800.1, in press.
- Liu, Q. and S. Boukabara, 2014: Community Radiative Transfer Model (CRTM) applications in supporting the Suomi National Polar-orbiting Partnership (SNPP) mission validation and verification. *Remote Sens. Environ.*, **140**, 744–754.
- Lynch, P., and Coauthors, 2016:, An 11-year global gridded aerosol optical thickness
 reanalysis (v1.0) for atmospheric and climate sciences. *Geosci. Model Dev.*, 9, 1489–1522, doi:10.5194/gmd-9-1489-2016.

Lynch-Stieglitz, M., 1994: The development and validation of a simple snow model for the GISS GCM. J. Climate, 7, 1842–1855, doi:10.1175/1520.0442(1994)007.1842:TDAVOA.2.0.CO:2

- doi:10.1175/1520-0442(1994)007,1842:TDAVOA.2.0.CO;2.
- Manney, G. L., K. Kruger, S. Pawson, K. Minschwaner, M. J. Schwartz, W. H. Daffer,
 N. J. Livesey, M. G. Mlynczak, E. E. Remsberg, J. M. Russell, J.W. Waters,
 2008: The evolution of the stratopause during the 2006 major warming: Satellite
 data and assimilated meteorological analyses. J. Geophys. Res., 113, D11115,
 doi:10.1029/2007JD00909.
- Marticorena, B., and G. Bergametti, 1995: Modeling the atmospheric dust cycle: 1.
 Design of a soil-derived dust emission scheme. J. Geophys. Res. Atmos., 100
 (D8), 16415–16430, doi:10.1029/95JD00690.
- McCarty, W., L. Coy, R. Gelaro, A, Huang, D. Merkova, E. B. Smith, M. Sienkiewicz, and K. Wargan, 2016: MERRA-2 input observations: Summary and initial assessment. NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2016-104606, Vol. 46, 61 pp.
- McPeters, R., M. Kroon, G. Labow, E. Brinksma, D. Balis, I. Petropavlovskikh, J. P.
 Veefkind, P. K. Bhartia, and P. F. Levelt, 2008: Validation of the Aura Ozone
 Monitoring Instrument total column ozone product. J. Geophys. Res., 113,
 D15S14, doi:10.1029/2007JD008802.
- Meng, J., R. Yang, H. Wei, M. Ek, G. Gayno, P. Xie, and K. Mitchell, 2012: The
 Land Surface Analysis in the NCEP Climate Forecast System Reanalysis. J. *Hydrometeor.*, 13, 1621–1630, doi: 10.1175/JHM-D-11-090.1.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. Bull. Amer.
 Meteor. Soc., 87, 343–360, doi:10.1175/BAMS-87-3-343.
- ¹⁵⁸⁸ Molina, M. J., and F. S. Rowland, 1974: Stratospheric sink for chlorouoromethanes: ¹⁵⁸⁹ Chlorine atom-catalysed destruction of ozone. *Nature*, **249** (**28**), 810–812.
- Molod, A., L. Takacs, M. Suárez, and J. Bacmeister, 2015: Development of the GEOS-5
 atmospheric general circulation model: evolution from MERRA to MERRA2,
 Geosci. Model Dev., 8, 1339–1356, doi:10.5194/gmd-8-1339-2015.
- ¹⁵⁹³ Moorthi, S., and M. J. Suárez, 1992: Relaxed Arakawa-Schubert: A parameterization of ¹⁵⁹⁴ moist convection for general circulation models. *Mon. Wea. Rev.*, **120**, 978–1002.
- Newman, P. A., and E. R. Nash, 2005: The unusual southern hemisphere stratosphere
 winter of 2002. J. Atmos. Sci., 62, 614–628, doi: 10.1175/JAS-3323.1.

- ¹⁵⁹⁷ Myhre, G., 2009:, Consistency between satellite-derived and modeled estimates of the ¹⁵⁹⁸ Direct Aerosol Effect. *Science*, **325**, 187–190.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral
 statistical-interpolation analysis system. Mon. Wea. Rev., 120, 1747–1763.
 http://dx.doi.org/10.1175/1520-0493(1992)120;1747:TNMCSS;2.0.CO;2
- Poli, P., and Coauthors, 2013: The data assimilation system and initial performance
 evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface
 observations only (ERA-20C). ERA Report Series 14
- Putman, W. and S.-J. Lin, 2007: Finite Volume Transport on Various Cubed Sphere
 Grids. J. Comput. Phys., 227, 55–78. doi:10.1016/j.jcp.2007.07.022.
- Randles, C. A., A. M. da Silva, V. Buchard, A. Darmenov, P. R. Colarco, V. Aquila, H.
 Bian, E. P. Nowottnick, X. Pan, A. Smirnov, H. Yu, and R. Govindaraju, 2016:
 The MERRA-2 aerosol assimilation. NASA Technical Report Series on Global
 Modeling and Data Assimilation, NASA/TM-2016-104606, Vol. 45, 143 pp.
- Randles, C. A., A. da Silva, V. Buchard, P.R. Colarco, A. Darmenov, R. Govindaraju,
 A. Smirnov, B. Holben, R. Ferrare, J. Hair, Y. Shinozuka, and C.J. Flynn, 2017:
 The MERRA-2 Aerosol Reanalysis, 1980–onward, Part 1: System description and
 data assimilation evaluation. J. Climate, in review.
- Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. Mahanama, and A. Toure, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. J. Climate, 24, 6322–6338. doi:10.1175/JCLI-D-10-05033.1.
- Reichle, R. H., and Q. Liu, 2014: Observation-Corrected Precipitation Estimates in
 GEOS-5. NASA/TM2014-104606, Vol. 35.
- Reichle, R. H., C. S. Draper, Q. Liu, M. Girotto, S. P. P. Mahanama, R. D. Koster, and
 G. J. M. De Lannoy, 2017b: Assessment of MERRA-2 land surface hydrology
 estimates, J. Climate, doi:10.1175/JCLI-D-16-0720.1.
- Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. P. Mahanama, and G. S.
 Partyka, 2017a: Land surface precipitation in MERRA-2, *J. Climate*, **30**, 1643–
 1664, doi:10.1175/JCLI-D-16-0570.1.
- Remer, L. A., and Coauthors, 2005: The MODIS aerosol algorithm, products, and
 validation. J. Atmos. Sci., 62, 947–973, 10.1175/JAS3385.1, http://dx.doi.org/
 10.1175/JAS3385.1.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609–1625.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax,
 2007: Daily high-resolution-blended analyses for sea surface temperature. J.
 Climate, 20, 5473–5496, doi:10.1175/2007JCLI1824.1.

- Rienecker, M. M., and Coauthors, 2008: The GEOS-5 Data Assimilation System–
 Documentation of versions 5.0.1 and 5.1.0, and 5.2.0. NASA Technical Report
 Series on Global Modeling and Data Assimilation, NASA/TM-2008-104606, Vol.
 27, 118 pp.
- https://gmao.gsfc.nasa.gov/pubs/docs/Rienecker369.pdf.
- Rienecker and Coauthors, 2011: MERRA NASA's Modern-Era Retrospective Analysis
 for Research and Applications. J. Climate, 24, 3624–3648, doi:10.1175/JCLI-D11-00015.1.
- Robertson, F. R., M. G. Bosilovich, J. Chen, and T. L. Miller, 2011: The effect of satellite
 observing system changes on MERRA water and energy fluxes. J. Climate, 24,
 5197–5217. doi: http://dx.doi.org/10.1175/2011JCLI4227.1.
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. Bull.
 Amer. Meteor. Soc., 91, 1010–1057, doi:10.1175/2010BAMS3001.1.
- Schubert, S. D., R. Rood, and J. Pfaendtner, 1993: An assimilated dataset for earth sci ence applications. *Bull. Amer. Meteor. Soc.*, 74, 2331–2342, doi:10.1175/1520 0477(1993)0742.0.CO;2.
- Schultz, M. G., A. Heil, J. J. Hoelzemann, A. Spessa, K. Thonicke, J. Goldammer,
 A. C. Held, J. M. Pereira, and M. van het Bolscher, 2008: Global Wildland
 Fire Emissions from 1960 to 2000. *Global Biogeochem. Cyc.*, 22, GB2002,
 doi:10.1029/2007GB003031.
- Schwartz, M.J., and Coauthors, 2008: Validation of the Aura Microwave Limb Sounder
 temperature and geopotential height measurements. J. Geophys. Res. 113,
 D15S11, doi:10.1029/2007JD008783.
- Segal-Rosenhemier, M., N. Barton, J. Redmann, S. Schmidt, S. LeBlanc, B. Anderson,
 E. Winstead, C. Corr, R. Moore, L.K. Thornhill, and R.I. Cullather., 2017: Improving our understanding of surface radiative flux bias in Arctic reanalysis over
 the marginal ice zone: Observational based sensitivity analysis during ARISE. J.
 Clim., in review.

Seidel, D. J., J. Li, C. Mears, I. Moradi, J. Nash, W. J. Randel, R. Saunders, D. W. J.
 Thompson, and C.-Z. Zou, 2016: Stratospheric temperature changes during the satellite era. J. Geophys. Res. Atmos., 121, doi:10.1002/2015JD024039.

Simmons, A. J., P. Berrisford, D. P. Dee, H. Hersbach, S. Hirahara, and J.-N. Thpaut, 2016: A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets. *Quart. J. Roy. Meteor. Soc.*, doi10.1002/qj.2949.

- Simmons, A. J., P. Poli, D. P. Dee, P. Berrisford, H. Hersbach, S. Kobayashi, and
 C. Peubey, 2014: Estimating low-frequency variability and trends in atmospheric
 temperature using ERA-Interim. Q. J. R. Meteorol. Soc., 140, 329–353, doi:10.1002/qj.2317.
- Stieglitz, M., A. Ducharne, R. D. Koster, and M. J. Suárez, 2001: The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. J. Hydrometeor., 2, 228–242, doi:10.1175/1525-7541(2001)002,0228:TIODSP.2.0.CO;2.
- ¹⁶⁷⁶ Susskind, J., J. Rosenfield, and D. Reuter, 1983: An accurate radiative transfer model
 ¹⁶⁷⁷ for use in the direct physical inversion of HIRS and MSU temperature sounding
 ¹⁶⁷⁸ data. J. Geophys. Res., 88, 8550–8568.
- Takacs, L. L., M. J. Suárez, and R. Todling, 2016: Maintaining atmospheric mass and
 water balance in reanalyses. *Quart. J. Roy. Meteor. Soc.*, 142 1565–1573. doi:
 10.1002/qj.2763.
- Taylor, K. E., D. Williamson, and F. Zwiers, 2000: The sea surface temperature and
 sea ice concentration boundary conditions for AMIP II simulations. Program for
 Climate Model Diagnosis and Intercomparison (PCMDI). Report 60, Lawrence
 Livermore National Laboratory.
- ¹⁶⁸⁶ Trenberth, K.E., and L. Smith, 2005: The mass of the atmosphere: A constraint on ¹⁶⁸⁷ global analysis. J. Climate, **18**, 864–875.
- Turner, J., S. Colwell, G. Marshall, T. Lachlan-Cope, A. Carleton, P. Jones, V. Lagun,
 P. Reid, and S. Iagovkina, 2004: The SCAR READER Project: Toward a High Quality Database of Mean Antarctic Meteorological Observations. J. Climate,
 1691 17, 2890–2898, doi:10.1175/1520-0442(2004)017<2890:TSRPTA>2.0.CO;2.
- van der Werf, G. R., J. Y. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A.
 F. Arellano Jr., 2006: Interannual variability in global biomass burning emissions
 from 1997 to 2004. Atmos. Chem. Phys., 6, 3423–3441, doi:10.5194/acp-6-34232006.

- Wargan, K., and L. Coy, 2016: Strengthening of the Tropopause Inversion Layer during
 the 2009 Sudden Stratospheric Warming: A MERRA-2 Study. J. Atmos. Sci., **73**, 1871–1887, doi: 10.1175/JAS-D-15-0333.1
- Wargan, K., G. Labow, S. Frith, S. Pawson, and G. Partyka, 2017: Evaluation of the
 ozone fields in NASAs MERRA-2 reanalysis. J. Climate, doi: 10.1175/JCLI-D16-0699.1.
- Wargan, K., S. Pawson, M. A. Olsen, J. C. Witte, A. R. Douglass, J. R. Ziemke, S. E.
 Strahan, and J. E. Nielsen, 2015: The global structure of upper troposphere-lower
 stratosphere ozone in GEOS-5: A multiyear assimilation of EOS Aura data, J. *Geophys. Res. Atmos.*, **120**, 2013–2036, doi:10.1002/2014JD022493.
- World Meteorological Organization (WMO 2014): Scientific Assessment of Ozone Depletion: 2014. Global Ozone Research and Monitoring Project - Report no. 55.
- Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.
- Xie, P., A. Yatagai, M. Chen, T. Hayasaka, Y. Fukushima, C. Liu, and S. Yang, 2007:
 A gauge-based analysis of daily precipitation over East Asia. J. Hydrometeorol., 8, 607–626.
- Yu, H., and Coauthors, 2006: A review of measurement-based assessments of the aerosol direct radiative effect and forcing. *Atmos. Chem. Phys.*, **6**, 613–666, doi:10.5194/acp-6-613-2006.

Table 1: Observation types assimilated in MERRA-2, including their usage dates and sources. Bold fonts indicate observation types not assimilated in MERRA. Acronyms are defined in the Appendix.

Data Type	MERRA-2 Dates	Source
	Conventional	
Raob, Pibal, Dropsonde	1 Jan 1980–present	See Rienecker et al. (2011)
AIREP, PIREP, ASDAR, MDCRS aircraft	1 Jan 1980–present	NCEP, ECMWF, JMA
PAOB	1 Jan 1980–17 Aug 2010	BOM
Surface land	1 Jan 1980–present	NCEP
Surface ship and buoy	1 Jan 1980–present	ICOADS
Ground-E	Based Remotely Sensed	
Wind profiler	14 May 1992–present	UCAR, NCEP
NEXRAD VAD wind	16 June 1997–present	NCEP
Satel	lite-Derived Wind	
GMS, MTSAT, Himawari atmos. motion vector	1 Jan 1980–present	NCEP, JMA
Meteosat atmos. motion vector	1 Jan 1980–present	NCEP, EUMETSAT
GOES atmos. motion vector	1 Jan 1980–present	NCEP
AVHRR atmos. motion vector	1 Oct 1982–present	CIMSS
SSM/I surface wind speed	9 Jul 1987–4 Nov 2009	RSS
ERS-1 surface wind vector	5 Aug 1991–21 May 1996	ESA
ERS-2 surface wind vector	19 Mar 1996–29 Mar 2011	ESA
QuikSCAT surface wind vector	19 Jul 1999–22 Nov 2009	JPL
MODIS atmos. motion vector	2 Jul 2002–present	CIMSS, NCEP
SSMIS surface wind speed	23 Oct 2003–29 Oct 2013	RSS
WindSat surface wind vector	13 Aug 2007–4 Aug 2012	NCEP
ASCAT surface wind vector	$15 \operatorname{Sep} 2008 - \operatorname{present}$	NCEP
Sa	tellite- $Retrieved$	
SBUV, SBUV/2 ozone	1 Jan 1980–31 Sep 2004	NASA/GES DISC
SSM/I rain rate	9 Jul 1987–16 Sep 2009	NASA/GES DISC
TMI rain rate	1 Jan 1998–8 Apr 2015	NASA/GES DISC
MLS temperature	13 Aug 2004–present	NASA/GES DISC
MLS ozone	$1 { m Oct} 2004 { m - present}$	NASA/GES DISC
OMI total column ozone	$1 { m Oct} 2004 { m - present}$	NASA/GES DISC
Ra	dio Occultation	
GPSRO bending angle	14 July 2004–present	NCAR, NCEP
Sa	tellite Radiance	
TOVS	1 Jan 1980–10 Oct 2006	NCAR, NESDIS
SSM/I	9 Jul 1987–4 Nov 2009	RSS
ATOVS (NOAA-15, -16, -17, -18)	21 Jul 1998–present	NESDIS
GOES (G08, G10, G11, G12 Low Res.)	24 April 2001–31 March 2007	NCEP, NESDIS
AMSU-A (Aqua)	1 Sep 2002–present	NASA/GES DISC
AIRS	1 Sep 2002-present	NASA/GES DISC
GOES (G11, G12, G13, G15 Full Res.)	1 April 2007–present	NESDIS
ATOVS (NOAA-19, Metop-A, -B)	21 May 2007–present	NESDIS
IASI	$17 \operatorname{Sep} 2008 - \operatorname{present}$	NESDIS
ATMS	16 Nov 2011–present	NESDIS
SEVIRI	$15 { m Feb} 2012 - { m present}$	NESDIS
CrIS	7 Apr 2012–present	NESDIS
Table 2: Nominal channel selections for satellite radiances assimilated in MERRA-2. Usage can vary for individual satellite platforms as a result of sensor failure or quality control decisions.

Sensor	Assimilated Channels
MSU	2-4
AMSU-A	4 - 14
ATMS	$5\!-\!15, 17\!-\!22$
AMSU-B	1-5
MHS	1-5
SSM/I	$1\!-\!7$
SSU	1–3
HIRS	2-8, 10-12
AIRS	See McCarty et al. 2016
IASI	See McCarty et al. 2016
CrIS	See McCarty et al. 2016
GOES Sounder	1-8, 10-12
SEVIRI	2, 3

Table 3: Sea surface temperature and sea ice concentration data products used in MERRA-2.

MERRA-2 dates	SST and SIC product
1 January 1980 – 31 December 1981	CMIP mid-monthly 1°
1 January 1982 – 31 December 2002	NOAA OISST daily $1/4^{\circ}$ (AVHRR)
1 January 2003 – 31 March 2006	NOAA OISST daily $1/4^{\circ}$ (AVHRR, AMSR-E)
1 April 2006 – present	OSTIA daily $1/20^{\circ}$

	Yu et al. (2006) Obs. ^a	Yu et al. (2006) Models ^b	MERRA-2 ^c	MERRAero ^c	$\mathrm{MACC}^{\mathrm{d}}$
		Land-are	ea Average		
AOD AAOD TOA DRE SFC DRE ATM DRE	$\begin{array}{c} 0.225 \pm 0.038 \\ - \\ -4.85 \pm 0.45 \\ -11.70 \pm 1.20 \\ 6.85 \pm 0.75 \end{array}$	$\begin{array}{c} 0.178 \pm 0.029 \\ - \\ -2.80 \pm 1.19 \\ -7.20 \pm 1.86 \\ 4.90 \pm 0.81 \end{array}$	$\begin{array}{l} \textbf{0.180} \pm 0.027 \\ \textbf{0.012} \pm 0.002 \\ \textbf{-3.09} \pm 0.62 \\ \textbf{-8.35} \pm 1.82 \\ \textbf{5.26} \pm 1.23 \end{array}$	$\begin{array}{c} 0.171 \pm 0.030 \\ 0.016 \pm 0.003 \\ -3.11 \pm 0.70 \\ -8.64 \pm 2.04 \\ 5.53 \pm 1.37 \end{array}$	$\begin{array}{c} 0.203 \pm 0.030 \\ 0.010 \pm 0.003 \\ -6.40 \pm 1.00 \\ -11.50 \pm 1.90 \\ 5.10 \end{array}$
Ocean-area Average					
AOD AAOD TOA DRE SFC DRE ATM DRE	$\begin{array}{c} 0.138 \pm 0.024 \\ - \\ -5.45 \pm 0.70 \\ -8.80 \pm 1.65 \\ 3.60 \pm 1.30 \end{array}$	$\begin{array}{c} 0.100 \pm 0.042 \\ - \\ - 3.50 \pm 1.28 \\ - 4.80 \pm 1.60 \\ 1.30 \pm 0.72 \end{array}$	$\begin{array}{c} \textbf{0.123} \pm 0.008 \\ \textbf{0.005} \pm 0.001 \\ \textbf{-3.65} \pm 0.21 \\ \textbf{-5.74} \pm 0.41 \\ \textbf{2.09} \pm 0.27 \end{array}$	$\begin{array}{c} 0.111 \pm 0.010 \\ 0.005 \pm 0.001 \\ -3.44 \pm 0.24 \\ -5.58 \pm 0.47 \\ 2.14 \pm 0.29 \end{array}$	$\begin{array}{c} 0.170 \pm 0.030 \\ 0.007 \pm 0.001 \\ -7.70 \pm 1.50 \\ -10.60 \pm 1.90 \\ 2.90 \end{array}$

Table 4: Clear-sky Direct Radiative Effect (DRE) from Reanalyses and Observations

^aMedian and standard deviation from satellite-derived estimates in Yu et al. (2006).

^bMedian and standard deviation from 4 global models in Yu et al. (2006).

^cClimatological global area-weighted average (\pm monthly standard deviation) for Y2003–Y2010.

^dFor MACC, the Y2003–Y2010 global mean and uncertainty is given following Bellouin et al. (2013).

1716 List of Figures

1717 1718 1719 1720	1	Observations assimilated per 6-hr cycle in (a) MERRA and (b) MERRA- 2. The temporary spike in the number of surface wind observations assim- ilated in MERRA-2 in late 2000 is due to an error in the pre-processing of QuikSCAT data.	78
1721 1722 1723 1724 1725 1726	2	Globally integrated monthly-mean mass anomalies from the mean sea- sonal cycle for (a) MERRA and (b) MERRA-2. Shown are the anoma- lies of total mass (black dotted), and their decomposition into atmo- spheric water (blue) and dry air (orange). The units are hPa. Results for MERRA-2 are derived from the data collection described in GMAO (2015b).	79
1727 1728 1729 1730 1731	3	Globally integrated monthly-mean total water budget terms for (a) MERRA and (b) MERRA-2. Shown are the water source term $(E-P, \text{blue})$, vertically integrated analysis increment of water (green), and atmospheric water storage (black dotted). The units are mm day ⁻¹ . Results for MERRA-2 are derived from the data collections described in GMAO (2015b, d, e).	80
1732 1733 1734 1735	4	Mean difference (1980–2015) between the (corrected) MERRA-2 precip- itation seen by the land surface and the model-generated precipitation within the MERRA-2 system. The units are mm d^{-1} . Results are de- rived from the data collections described in GMAO (2015h, j)	81
1736 1737 1738 1739	5	Time series of 12-month running mean prescribed sea surface temperature for various reanalyses, averaged between 60°N and 60°S. The units are K. Results for MERRA-2 are derived from the data collection described in GMAO (2015f).	82
1740 1741 1742 1743 1744 1745	6	Monthly mean (thick lines) and RMS (thin lines) background depar- tures for surface pressure observations assimilated in MERRA (blue) and MERRA-2 (red). Results are shown for the (a) Northern Hemisphere and (b) Southern Hemisphere. The units are hPa. Also shown are the corresponding monthly mean counts of surface pressure observations as- similated in MERRA-2 (gray shaded).	83

1746 1747 1748 1749 1750 1751	7	Global monthly mean (thick lines) and RMS (thin lines) background de- partures for radiosonde temperature observations assimilated in MERRA (blue) and MERRA-2 (red). Results are shown for the pressure levels (a) 10 hPa, (b) 50 hPa, (c) 300 hPa and (d) 700 hPa. The units are K. Also shown are the corresponding monthly mean counts of radiosonde temperature observations assimilated in MERRA-2 (gray shaded)	84
1752 1753 1754	8	As in Figure 7, except for radiosonde specific humidity observations in the tropics $(20^{\circ}N-20^{\circ}S)$ at (a) 500 hPa and (b) 850 hPa. The units are g kg ⁻¹ .	85
1755 1756 1757 1758 1759	9	(a) Mean and (b) standard deviation of the monthly mean analysis ten- dency of surface pressure for the period January 1980 through December 2015. Monthly mean values are based on four synoptic times daily. The units are hPa day ⁻¹ . Results are derived from the data collection de- scribed in GMAO (2015k).	86
1760 1761 1762 1763	10	Global (a) mean and (b) standard deviation of the monthly mean analysis tendency of temperature from 1000 to 70 hPa. Monthly means values are based on four synoptic times daily. The units are K day ⁻¹ . Results are derived from the data collection described in GMAO (2015n)	87
1764 1765 1766	11	As in Figure 10, except for specific humidity in the tropics $(20^{\circ}N-20^{\circ}S)$ from 1000 to 250 hPa. The units are g kg ⁻¹ day ⁻¹ . Results are derived from the data collection described in GMAO (20151).	88
1767 1768 1769 1770 1771 1772 1773 1774	12	Probability distribution functions (PDFs) of observation minus forecast (O-F, dashed) and observation minus analysis (O-A, solid) differences in observation space, collocated in space and time for each sensor in the MERRA-2 aerosol observing system. The PDFs are calculated from innovation data in log-transformed space (ln(AOD+0.01)) to ensure distributions are positive and Gaussian. The time periods considered include AVHRR (1993–1999), MODIS Terra (2001–2014), MODIS Aqua (2003–2014), MISR (2001–2012), and AERONET (ANET 2000–2013).	89
1775 1776 1777 1778 1779 1780 1781	13	Time series of area-weighted aerosol optical depth (AOD) from the MERRA-2 aerosol reanalysis averaged over major aerosol source regions: (a) South and East Asia $[5^{\circ}N-55^{\circ}N, 65^{\circ}W-160^{\circ}W]$, (b) northern Africa $[2.5^{\circ}S-30^{\circ}N, 45^{\circ}W-15^{\circ}E]$, and (c) the Amazon Basin in South America $[20^{\circ}S-7.5^{\circ}N, 80^{\circ}W-30^{\circ}W]$. The total AOD (thick black line) is the sum of contributions from sea salt (blue), dust (yellow), carbonaceous (black and organic carbon, green), and sulfate (grey) AOD. Results are derived from the data	
1782		collection described in GMAO (2015g).	90

1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793	14	Aerosol optical depth (AOD) from aerosol reanalyses (MERRA-2, MER- RAero, NAAPS, MACC), inter-model comparisons (AeroCom Phase I, Yu_Model), and observations (Yu_Obs) for the period 2003–2010. Where available, total AOD is broken down by component species (left bar) and by fine and coarse mode (right bar). For MERRA-2 and MERRAero, the error bar represents the standard deviation of the monthly-mean AOD for the period 2003–2010. For MACC, the error bar is the uncertainty in the total AOD from Bellouin et al. (2013). AeroCom (Kinne et al., 2006) and Yu et al. (2006) uncertainty are the inter-model or inter-observational standard deviations. Coarse mode is defined as the sum of dust plus sea salt AOD, with the remainder of the AOD assigned to the fine mode.	
1794 1795		Results for MERRA-2 are derived from the data collection described in GMAO (2015g).	91
1796 1797 1798 1799	15	Time series of 12-month running mean globally averaged precipitation for several reanalyses and the GPCP merged gauge satellite data product. The units are mm day ^{-1} . Results for MERRA-2 are derived from the data collection described in GMAO (2015h).	92
1800 1801 1802 1803	16	Time-averaged precipitation differences during June-July-August for (a) MERRA minus GPCP and (b) MERRA-2 minus GPCP for the period 1980–2015. The units are mm day ⁻¹ . Results for MERRA-2 are derived from the data collection described in GMAO (2015h).	93
1804 1805 1806 1807 1808 1809	17	Time series of midwestern US summer seasonal precipitation anomalies, following Bosilovich (2013). The anomalies are computed from the June-July-August mean for the period 1980–2011. The gauge data are from NOAA/CPC gridded daily data for the US (Xie et al. 2007). The units are mm day ⁻¹ . Results for MERRA-2 are derived from the data collection described in GMAO (2015h).	94
1810 1811 1812 1813 1814 1815 1816 1817 1818 1819	18	Regional summary statistics for the US summer seasonal anomaly time series of precipitation: (a) mean (mm day ⁻¹), (b) standard deviation (mm day ⁻¹), and (c) anomaly correlation to CPC gauge observations. The anomalies are computed from the June-July-August mean for the period 1980–2011. The regions lie within the continental US and are defined as in Bosilovich (2013): Northeast (NE), Southeast (SE), Midwest (MW), Great Plains (GP), Southern Great Plains (SGP), Northern Great Plains (NGP), Northwest (NW), Southwest (SW), and the accumulation of all area in these regions (US). Results for MERRA-2 are derived from the data collection described in GMAO (2015h).	95

1820 1821 1822 1823 1824 1825 1825 1826 1827 1828 1829	19	Average amount of precipitation that exceeds the 99th percentile during June-July-August for the period 1980–2013 for (a) MERRA, (b) MERRA-2, and (c) CPC gauge observations. Panel (d) shows the closeness of each reanalysis to the CPC observations for the same period, defined as $ MERRA-2 - CPC - MERRA - CPC $, where the vertical bars indicate absolute differences and the names indicate the set of time-averaged gridpoint values for each data type. In (d), blue (red) shades indicate that MERRA-2 (MERRA) is closer to the CPC observations. The units in all panels are mm day ⁻¹ . Results for MERRA-2 are derived from the data collection described in GMAO (2015d).	96
1830 1831 1832 1833 1834 1835 1836	20	Ertel's potential vorticiity (EPV, $\times 10^3$ potential vorticity units, PVU; 1 PVU = 10^{-6} m ⁻² s ⁻¹ K kg ⁻¹) at 0.7 hPa on 2 January 1995 12 UTC for (a) MERRA and (b) MERRA-2 for the Northern Hemisphere. Polar cap detail (80° -90°N) for (c) MERRA and (d) MERRA-2. Color shading interval is 2.5×10^3 PVU. Black contour interval is 10×10^3 PVU in (a) and (b) and 5×10^3 PVU in (c) and (d). Cyan circle denotes 80° N latitude. Results are derived from the data collection described in GMAO (2015c).	97
1837 1838 1839	21	Time-altitude section of zonally averaged temperature at 70°N for (a) MERRA and (b) MERRA-2. The time resolution is twice daily (00 and 12 UTC) for December 2005–March 2006. The contour interval is 5 K. $$.	98
1840 1841 1842 1843 1844	22	Monthly and globally averaged temperature anomaly for MERRA-2 as a function of time. The annual cycle and mean for 1980–2015 have been removed. The MLS temperatures were introduced at levels above 5 hPa beginning in August 2004. Results are derived from the data collection described in GMAO (2015c).	99
1845 1846 1847 1848 1849 1850 1851 1852 1853	23	Time series of (a) total ozone (Dobson units, DU) at the South Pole derived from individual ozonesonde measurements (gray) and from collocated values in MERRA (blue) and MERRA-2 (red). Note that ozonesonde measurements are unavailable prior to 1986; see text for details. The reanalysis-minus-ozonesonde differences divided by sonde total ozone are shown in (b) for MERRA (blue) and MERRA-2 (red). The black vertical line in (b) separates the SBUV and Aura periods. (Figure from Wargan et al. 2016.) Results for MERRA-2 are derived from the data collection described in GMAO (2015a).	100

1854 1855 1856 1857 1858 1859 1860	24	Time series of the Antarctic ozone hole area calculated from MERRA-2 ozone fields averaged between 20 September and 10 October for the years 1980–2015 (red curve with circles). Also shown are values derived from TOMS (gray squares) and OMI (black triangles) observations. The units are 10 ⁶ km ² . Results for MERRA-2 in 1994 are excluded due to insuf- ficient SBUV data coverage in the Southern Hemisphere, which signifi- cantly degraded the analysis; see text for details. Results for MERRA-2
1861		are derived from the data collection described in GMAO (2015a) 101
1862 1863 1864 1865 1866 1867 1868	25	Average annual cycle of 2-m air temperature in MERRA and MERRA-2 at (a) South Pole station (90°S; 1980–2014; Turner et al., 2004), (b) Gill automatic weather station (80°S, 179°W; 1985–2014; Turner et al., 2004), and (c) Summit, Greenland (73°N, 38°W; 2000–2002; Hoch, 2005). The units are °C. Vertical bars denote ± 1 standard deviation of the multi-year time series for each month. Results for MERRA-2 are derived from the data collections described in GMAO (2015i, j, m) 102
1869 1870 1871 1872 1873 1873	26	Surface mass balance for the Greenland Ice Sheet for the period 1980–2012 in (a) MERRA, (b) MERRA-2, and (c) MAR regional climate model (Fet- tweis 2007). The units are mm yr ⁻¹ water-equivalent. Surface topography (including ice sheet) is contoured with dashed lines every 200 m. Results for MERRA-2 are derived from the data collections described in GMAO (2015i, j, m)



Figure 1: Observations assimilated per 6-hr cycle in (a) MERRA and (b) MERRA-2. The temporary spike in the number of surface wind observations assimilated in MERRA-2 in late 2000 is due to an error in the pre-processing of QuikSCAT data.



Figure 2: Globally integrated monthly-mean mass anomalies from the mean seasonal cycle for (a) MERRA and (b) MERRA-2. Shown are the anomalies of total mass (black dotted), and their decomposition into atmospheric water (blue) and dry air (orange). The units are hPa. Results for MERRA-2 are derived from the data collection described in GMAO (2015b).



Figure 3: Globally integrated monthly-mean total water budget terms for (a) MERRA and (b) MERRA-2. Shown are the water source term (E-P, blue), vertically integrated analysis increment of water (green), and atmospheric water storage (black dotted). The units are mm day⁻¹. Results for MERRA-2 are derived from the data collections described in GMAO (2015b, d, e).



Figure 4: Mean difference (1980–2015) between the (corrected) MERRA-2 precipitation seen by the land surface and the model-generated precipitation within the MERRA-2 system. The units are mm d^{-1} . Results are derived from the data collections described in GMAO (2015h, j).



Figure 5: Time series of 12-month running mean prescribed sea surface temperature for various reanalyses, averaged between 60°N and 60°S. The units are K. Results for MERRA-2 are derived from the data collection described in GMAO (2015f).



Figure 6: Monthly mean (thick lines) and RMS (thin lines) background departures for surface pressure observations assimilated in MERRA (blue) and MERRA-2 (red). Results are shown for the (a) Northern Hemisphere and (b) Southern Hemisphere. The units are hPa. Also shown are the corresponding monthly mean counts of surface pressure observations assimilated in MERRA-2 (gray shaded).



Figure 7: Global monthly mean (thick lines) and RMS (thin lines) background departures for radiosonde temperature observations assimilated in MERRA (blue) and MERRA-2 (red). Results are shown for the pressure levels (a) 10 hPa, (b) 50 hPa, (c) 300 hPa and (d) 700 hPa. The units are K. Also shown are the corresponding monthly mean counts of radiosonde temperature observations assimilated in MERRA-2 (gray shaded).



Figure 8: As in Figure 7, except for radiosonde specific humidity observations in the tropics $(20^{\circ}N-20^{\circ}S)$ at (a) 500 hPa and (b) 850 hPa. The units are g kg⁻¹.



Figure 9: (a) Mean and (b) standard deviation of the monthly mean analysis tendency of surface pressure for the period January 1980 through December 2015. Monthly mean values are based on four synoptic times daily. The units are hPa day⁻¹. Results are derived from the data collection described in GMAO (2015k).



Figure 10: Global (a) mean and (b) standard deviation of the monthly mean analysis tendency of temperature from 1000 to 70 hPa. Monthly means values are based on four synoptic times daily. The units are K day⁻¹. Results are derived from the data collection described in GMAO (2015n).



Figure 11: As in Figure 10, except for specific humidity in the tropics $(20^{\circ}N-20^{\circ}S)$ from 1000 to 250 hPa. The units are g kg⁻¹ day⁻¹. Results are derived from the data collection described in GMAO (20151).



Figure 12: Probability distribution functions (PDFs) of observation minus forecast (O-F, dashed) and observation minus analysis (O-A, solid) differences in observation space, collocated in space and time for each sensor in the MERRA-2 aerosol observing system. The PDFs are calculated from innovation data in log-transformed space (ln(AOD+0.01)) to ensure distributions are positive and Gaussian. The time periods considered include AVHRR (1993–1999), MODIS Terra (2001–2014), MODIS Aqua (2003–2014), MISR (2001–2012), and AERONET (ANET 2000–2013).



Figure 13: Time series of area-weighted aerosol optical depth (AOD) from the MERRA-2 aerosol reanalysis averaged over major aerosol source regions: (a) South and East Asia $[5^{\circ}N-55^{\circ}N, 65^{\circ}W-160^{\circ}W]$, (b) northern Africa $[2.5^{\circ}S-30^{\circ}N, 45^{\circ}W-15^{\circ}E]$, and (c) the Amazon Basin in South America $[20^{\circ}S-7.5^{\circ}N, 80^{\circ}W-30^{\circ}W]$. The total AOD (thick black line) is the sum of contributions from sea salt (blue), dust (yellow), carbonaceous (black and organic carbon, green), and sulfate (grey) AOD. Results are derived from the data collection described in GMAO (2015g).



Figure 14: Aerosol optical depth (AOD) from aerosol reanalyses (MERRA-2, MER-RAero, NAAPS, MACC), inter-model comparisons (AeroCom Phase I, Yu_Model), and observations (Yu_Obs) for the period 2003–2010. Where available, total AOD is broken down by component species (left bar) and by fine and coarse mode (right bar). For MERRA-2 and MERRAero, the error bar represents the standard deviation of the monthly-mean AOD for the period 2003–2010. For MACC, the error bar is the uncertainty in the total AOD from Bellouin et al. (2013). AeroCom (Kinne et al., 2006) and Yu et al. (2006) uncertainty are the inter-model or inter-observational standard deviations. Coarse mode is defined as the sum of dust plus sea salt AOD, with the remainder of the AOD assigned to the fine mode. Results for MERRA-2 are derived from the data collection described in GMAO (2015g).



Figure 15: Time series of 12-month running mean globally averaged precipitation for several reanalyses and the GPCP merged gauge satellite data product. The units are mm day⁻¹. Results for MERRA-2 are derived from the data collection described in GMAO (2015h).



Figure 16: Time-averaged precipitation differences during June-July-August for (a) MERRA minus GPCP and (b) MERRA-2 minus GPCP for the period 1980–2015. The units are mm day⁻¹. Results for MERRA-2 are derived from the data collection described in GMAO (2015h).



Figure 17: Time series of midwestern US summer seasonal precipitation anomalies, following Bosilovich (2013). The anomalies are computed from the June-July-August mean for the period 1980–2011. The gauge data are from NOAA/CPC gridded daily data for the US (Xie et al. 2007). The units are mm day⁻¹. Results for MERRA-2 are derived from the data collection described in GMAO (2015h).



Figure 18: Regional summary statistics for the US summer seasonal anomaly time series of precipitation: (a) mean (mm day⁻¹), (b) standard deviation (mm day⁻¹), and (c) anomaly correlation to CPC gauge observations. The anomalies are computed from the June-July-August mean for the period 1980–2011. The regions lie within the continental US and are defined as in Bosilovich (2013): Northeast (NE), Southeast (SE), Midwest (MW), Great Plains (GP), Southern Great Plains (SGP), Northern Great Plains (NGP), Northwest (NW), Southwest (SW), and the accumulation of all area in these regions (US). Results for MERRA-2 are derived from the data collection described in GMAO (2015h).



Figure 19: Average amount of precipitation that exceeds the 99th percentile during June-July-August for the period 1980–2013 for (a) MERRA, (b) MERRA-2, and (c) CPC gauge observations. Panel (d) shows the closeness of each reanalysis to the CPC observations for the same period, defined as |MERRA-2 - CPC| - |MERRA - CPC|, where the vertical bars indicate absolute differences and the names indicate the set of time-averaged grid-point values for each data type. In (d), blue (red) shades indicate that MERRA-2 (MERRA) is closer to the CPC observations. The units in all panels are mm day⁻¹. Results for MERRA-2 are derived from the data collection described in GMAO (2015d).



Figure 20: Ertel's potential vorticiity (EPV, $\times 10^3$ potential vorticity units, PVU; 1 PVU = $10^{-6}m^{-2}s^{-1}K \text{ kg}^{-1}$) at 0.7 hPa on 2 January 1995 12 UTC for (a) MERRA and (b) MERRA-2 for the Northern Hemisphere. Polar cap detail ($80^{\circ}-90^{\circ}N$) for (c) MERRA and (d) MERRA-2. Color shading interval is 2.5×10^3 PVU. Black contour interval is 10×10^3 PVU in (a) and (b) and 5×10^3 PVU in (c) and (d). Cyan circle denotes $80^{\circ}N$ latitude. Results are derived from the data collection described in GMAO (2015c).



Figure 21: Time-altitude section of zonally averaged temperature at 70°N for (a) MERRA and (b) MERRA-2. The time resolution is twice daily (00 and 12 UTC) for December 2005–March 2006. The contour interval is 5 K.



Figure 22: Monthly and globally averaged temperature anomaly for MERRA-2 as a function of time. The annual cycle and mean for 1980–2015 have been removed. The MLS temperatures were introduced at levels above 5 hPa beginning in August 2004. Results are derived from the data collection described in GMAO (2015c).



Figure 23: Time series of (a) total ozone (Dobson units, DU) at the South Pole derived from individual ozonesonde measurements (gray) and from collocated values in MERRA (blue) and MERRA-2 (red). Note that ozonesonde measurements are unavailable prior to 1986; see text for details. The reanalysis-minus-ozonesonde differences divided by sonde total ozone are shown in (b) for MERRA (blue) and MERRA-2 (red). The black vertical line in (b) separates the SBUV and Aura periods. (Figure from Wargan et al. 2016.) Results for MERRA-2 are derived from the data collection described in GMAO (2015a).



Figure 24: Time series of the Antarctic ozone hole area calculated from MERRA-2 ozone fields averaged between 20 September and 10 October for the years 1980–2015 (red curve with circles). Also shown are values derived from TOMS (gray squares) and OMI (black triangles) observations. The units are 10⁶ km². Results for MERRA-2 in 1994 are excluded due to insufficient SBUV data coverage in the Southern Hemisphere, which significantly degraded the analysis; see text for details. Results for MERRA-2 are derived from the data collection described in GMAO (2015a).



Figure 25: Average annual cycle of 2-m air temperature in MERRA and MERRA-2 at (a) South Pole station (90°S; 1980–2014; Turner et al., 2004), (b) Gill automatic weather station (80°S, 179°W; 1985–2014; Turner et al., 2004), and (c) Summit, Greenland (73°N, 38°W; 2000–2002; Hoch, 2005). The units are °C. Vertical bars denote ± 1 standard deviation of the multi-year time series for each month. Results for MERRA-2 are derived from the data collections described in GMAO (2015i, j, m).



Figure 26: Surface mass balance for the Greenland Ice Sheet for the period 1980–2012 in (a) MERRA, (b) MERRA-2, and (c) MAR regional climate model (Fettweis 2007). The units are mm yr^{-1} water-equivalent. Surface topography (including ice sheet) is contoured with dashed lines every 200 m. Results for MERRA-2 are derived from the data collections described in GMAO (2015i, j, m).