The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation

C. A. RANDLES,a,j A. M. DA SILVA,a V. BUCHARD,a,b P. R. COLARCO,c A. DARMENOV,a R. GOVINDARAJU,a,d A. SMIRNOV,d,e B. HOLBEN,e R. FERRARE,f J. HAIR,f Y. SHINOZUKA,g,h AND C. J. FLYNNi

a Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland
b Goddard Earth Sciences Technology and Research/Universities Space Research Association, Columbia, Maryland
c Atmospheric Chemistry and Dynamics Lab, NASA Goddard Space Flight Center, Greenbelt, Maryland
d Science Systems and Applications, Inc., Lanham, Maryland
e NASA Biospheric Sciences Laboratory, Greenbelt, Maryland
f NASA Langley Research Center, Hampton, Virginia
g Bay Area Environmental Research Institute, Petaluma, California
h NASA Ames Research Center Cooperative for Research in Earth Science and Technology, Moffett Field, California
i Pacific Northwest National Laboratory, Richland, Washington

(Manuscript received 18 August 2016, in final form 10 March 2017)

ABSTRACT

The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), updates NASA’s previous satellite-era (1980 onward) reanalysis system to include additional observations and improvements to the Goddard Earth Observing System, version 5 (GEOS-5), Earth system model. As a major step toward a full Integrated Earth Systems Analysis (IESA), in addition to meteorological observations, MERRA-2 now includes assimilation of aerosol optical depth (AOD) from various ground- and space-based remote sensing platforms. Here, in the first of a pair of studies, the MERRA-2 aerosol assimilation is documented, including a description of the prognostic model (GEOS-5 coupled to the GOCART aerosol module), aerosol emissions, and the quality control of ingested observations. Initial validation and evaluation of the analyzed AOD fields are provided using independent observations from ground, aircraft, and shipborne instruments. The positive impact of the AOD assimilation on simulated aerosols is demonstrated by comparing MERRA-2 aerosol fields to an identical control simulation that does not include AOD assimilation. After showing the AOD evaluation, this paper takes a first look at aerosol–climate interactions by examining the shortwave, clear-sky aerosol direct radiative effect. The companion paper (Part II) evaluates and validates available MERRA-2 aerosol properties not directly impacted by the AOD assimilation (e.g., aerosol vertical distribution and absorption). Importantly, while highlighting the skill of the MERRA-2 aerosol assimilation products, both studies point out caveats that must be considered when using this new reanalysis product for future studies of aerosols and their interactions with weather and climate.

1. Introduction

the MERRA-2 system incorporates new observations not available for MERRA and reduces spurious trends and jumps related to changes in the meteorological observing system (McCarty et al. 2016). Numerous improvements have been made to the MERRA-2 Goddard Earth Observing System, version 5 (GEOS-5), modeling system (Molod et al. 2012, 2015), including improvements in the representation of the hydrologic cycle (Takacs et al. 2015; Reichle and Liu 2014), the stratosphere, ozone, and cryospheric processes (Bosilovich et al. 2016).

In a significant step toward an Integrated Earth Systems Analysis (IESA), MERRA-2 for the first time includes analyzed aerosol fields that are radiatively coupled to the atmosphere. To our knowledge, this is the first multidecadal reanalysis within which meteorological and aerosol observations are jointly assimilated into a global assimilation system, although other operational forecasting centers are actively developing similar capabilities (e.g., Benedetti et al. 2009; Sekiyama et al. 2010; Lynch et al. 2016). Previously, the GMAO had performed an offline aerosol reanalysis (the MERRA Aerosol Reanalysis or MERRAero) in which bias-corrected Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth (AOD) from Terra and Aqua was assimilated into an earlier version of the NASA GEOS-5 model driven by meteorology from MERRA-1 (Buchard et al. 2015, 2016). In addition to assimilating bias-corrected Collection 5 MODIS AOD as in MERRAero, MERRA-2 now also includes assimilation of bias-corrected AOD from the Advanced Very High Resolution Radiometer (AVHRR) instruments (Heidinger et al. 2014), AOD retrievals from the Multiangle Imaging SpectroRadiometer (MISR) over bright surfaces (Kahn et al. 2005), and ground-based Aerosol Robotic Network (AERONET) direct measurements of AOD (Level 2; Holben et al. 1998).

Despite their rapid increase in complexity, aerosol models remain uncertain due to poorly constrained emissions and physical process parameterizations such as hygroscopic growth, mixing, and aerosol–cloud interactions (Textor et al. 2006; Kinne et al. 2006; Benedetti et al. 2009; Schutgens et al. 2010). Similarly, many aerosol observations such as those from remote sensing platforms, both satellite- and ground-based, suffer from limited coverage (e.g., due to their orbit and/or cloud contamination), contextual biases such as “clear-sky” bias, and biases due to assumptions made in retrieval algorithms (Zhang and Reid 2009; Shi et al. 2011; Colarco et al. 2014). Similarly, many aerosol observations such as those from remote sensing platforms, both satellite- and ground-based, suffer from limited coverage (e.g., due to their orbit and/or cloud contamination), contextual biases such as “clear-sky” bias, and biases due to assumptions made in retrieval algorithms (Zhang and Reid 2009; Shi et al. 2011; Colarco et al. 2014). Although there are several important caveats, these fields can serve as initial conditions for regional modeling and air quality forecasting (Zhang et al. 2012; Giordano et al. 2015; Buchard et al. 2016), as a tool to investigate aerosol–climate or aerosol–weather interactions (Bellouin et al. 2013; Reale et al. 2014), and for optimal network/satellite sensor design in the context of Observing System Simulation Experiments (OSSEs; Bocquet et al. 2015).

Because standard minimum variance data assimilation algorithms are designed to minimize random errors under the assumption of no biases, systematic errors in model background and observations must be carefully accounted for, or, to the extent possible, removed prior to the assimilation process (Dee and da Silva 1999; Zhang and Reid 2009; Benedetti et al. 2009). On the observational side, this can be accomplished by imposing strict quality control on the observing system used by the reanalysis (Zhang and Reid 2006). Additionally, it is essential to have a well-performing and well-tuned prognostic model to minimize the corrections imposed by the assimilation since large corrections can cause spurious trends and jumps in the analyzed fields when the observing system changes over time (Lynch et al. 2016).

Several important caveats must be understood when using and evaluating the MERRA-2 aerosol reanalysis products. First, the relative paucity of (nonassimilated) aerosol observations makes independent validation of the analyzed AOD a challenge. Second, despite the best efforts at harmonizing the observing system through quality control, differences in data coverage can and do impact the analyzed AOD, particularly between the pre- and post-NASA Earth Observing System (EOS) periods (1980–99 and 2000 onward, respectively). Finally, nonanalyzed aerosol properties (e.g., vertical distribution, aerosol speciation, absorption) are not fully constrained by the assimilation and strongly resemble the assimilating model in most cases. Nevertheless, despite some deficiencies, previous studies (e.g., Buchard et al. 2015, 2016), the current study, and a companion evaluation paper (Buchard et al. 2017, hereinafter Part II) demonstrate that the aerosol assimilation system does indeed show considerable skill in simulating numerous observable aerosol properties.

The purpose of this study is to describe the MERRA-2 aerosol data assimilation system, provide initial validation of the analyzed AOD fields, and suggest applications of the aerosol products while highlighting both the model’s skill and deficiencies. In section 2, we provide details about the GEOS-5/GOCART model and aerosol emissions. Section 3 discusses the aerosol assimilation system and the AOD observing system. We then...
evaluate the performance of the analyzed AOD fields with respect to the observing system and demonstrate the stability of the assimilation system (section 4a). The impact of the AOD assimilation is shown by comparisons to a control simulation driven by MERRA-2 meteorology with identical aerosol emissions, but without AOD assimilation (section 4b). In section 4c, we present a validation of MERRA-2 AOD with available nonassimilated observations. Finally, we examine the MERRA-2 clear-sky shortwave aerosol direct radiative effect (DRE; section 4d) and compare to results from other AOD reanalyses. A summary is presented in section 5.

Part II of this study presents an evaluation of the MERRA-2 system with respect to nonassimilated aerosol properties (e.g., absorption, vertical profile, PM$_{2.5}$) and case studies during major aerosol events (e.g., the 1991 Pinatubo eruption). In both this work and Part II, we demonstrate important aspects of the aerosol products that users must consider. We note that the full MERRA-2 reanalysis dataset including aerosol fields is currently publicly available online through the Goddard Earth Sciences Data and Information Services Center (GES DISC; http://disc.sci.gsfc.nasa.gov/mdisc/) with technical documentation (Randles et al. 2016) and file specifications available at http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/. Individual MERRA-2 data collections are accessible via their own digital object identifier (doi) codes. For this study and Part II we use aerosol (GMAO 2015a,b,c), meteorological (GMAO 2015d,e), and radiation (GMAO 2015f) collections.

2. The MERRA-2 modeling system

An overview of the MERRA-2 modeling system is found in Gelaro et al. (2017). Briefly, MERRA-2 is produced using the GEOS-5 atmospheric model and data assimilation system version 5.12.4 (Rienecker et al. 2008; Molod et al. 2015) and the three-dimensional variational data analysis (3DVAR) Gridpoint Statistical Interpolation (GSI) meteorological analysis scheme (Wu et al. 2002; Kleist et al. 2009). The GSI uses an incremental analysis update procedure every 6 h (Bloom et al. 1996). The discretization of the dynamical core (Lin 2004) is computed on a cubed-sphere grid that mitigates grid-spacing singularities (Putman and Lin 2007). The GEOS-5 model resolution on the native cubed-sphere grid is roughly 50 km with 72 hybrid-eta layers from the surface to 0.01 hPa, while most products are saved on a standard 0.5° × 0.625° latitude by longitude grid. As noted previously, the MERRA-2 meteorological observing system includes numerous additions that are detailed in McCarty et al. (2016), and Bosilovich et al. (2016) presents the validation of the MERRA-2 meteorological, radiation, ozone, and cryospheric fields. The remainder of this section focuses on aspects of the MERRA-2 modeling system relevant for the aerosol reanalysis.

a. GOCART aerosol module

Aerosols in MERRA-2 are simulated with a radiation-coupled version of the Goddard Chemistry, Aerosol, Radiation, and Transport model (GOCART; Chin et al. 2002; Colarco et al. 2010). GOCART treats the sources, sinks, and chemistry of 15 externally mixed aerosol mass mixing ratio tracers: dust (five noninteracting size bins), sea salt (five noninteracting size bins), hydrophobic and hydrophilic black and organic carbon (BC and OC, respectively; four tracers), and sulfate (SO$_4$). Both dust and sea salt have wind speed–dependent emissions. Primary sulfate and carbonaceous aerosol species have emissions principally from fossil fuel combustion, biomass burning, and biofuel consumption, with additional biogenic sources of organic carbon. Secondary sources of sulfate include chemical oxidation of sulfur dioxide gas (SO$_2$) and dimethyl sulfide (DMS), and we include a database of volcanic SO$_2$ emissions and injection heights. Note that we use a monthly mean climatology of oxidant fields in GOCART. Loss processes for all aerosols include dry deposition (including gravitational settling), large-scale wet removal, and convective scavenging. We note that precipitation-induced aerosol deposition does not depend on model-generated precipitation fields. Rather, we use the MERRA-2 corrected precipitation product of Reichle et al. (2017) that better represents diurnal precipitation changes compared to observations. Aerosol hygroscopic growth depends on simulated relative humidity and is considered in computations of particle fall velocity, deposition velocity, and optical parameters. Numerous studies have demonstrated the skill of the GOCART aerosol module in simulating AOD and other observable aerosol properties (e.g., Colarco et al. 2010; Nowottnick et al. 2010, 2011; Bian et al. 2013).

b. Emissions

Figure 1 shows the EOS-period climatological distribution of aerosol emissions in MERRA-2 from GOCART. Dust emissions (Fig. 1a) use a map of potential dust source locations based on the observed correlation of dust emitting regions with large-scale topographic depressions (Ginoux et al. 2001, updated to 0.25° resolution). Emissions of both dust and sea salt (Fig. 1b) are wind driven for each size bin, parameterized following Marticorena and Bergametti (1995) and Gong (2003), respectively. Sea salt emission strength is modulated
FIG. 1. Annual-mean aerosol emissions climatology (2000–14) from MERRA-2. Emissions are shown for (a) dust, (b) sea salt, (c) particulate organic matter (POM), (d) black carbon (BC), (e) primary sulfate (SO$_4$) and sulfate from oxidation of sulfur dioxide gas (SO$_2$), and (f) SO$_4$ from aqueous production. We define POM = 1.4 × OC, where OC is organic carbon (Textor et al. 2006). Emissions are from all sectors (fossil fuel, biofuel, biomass burning, and biogenic, if applicable). The global, annual mean emissions are given on each panel.
Table 1. Aerosol and precursor emissions in MERRA-2. Precursor gases include SO2, DMS, and MSA for sulfate aerosol and biogenic terpene for particulate organic matter (POM). The native resolution of emission inventories is given in °lat × °lon. For time-varying emissions, the final year of emissions is persisted; however, for volcanoes, only degassing volcanic emissions are persisted after 2010. Note that POM = 1.4 × OC as in Textor et al. (2006).

<table>
<thead>
<tr>
<th>Aerosol type</th>
<th>Source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dust</td>
<td>Wind-driven</td>
<td>Static topographic depression source map</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.3125° × 0.25°; Ginoux et al. 2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>See section 2b</td>
</tr>
<tr>
<td>Sea salt</td>
<td>Wind-driven</td>
<td>Daily degassing and eruptive volcanoes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1980–2010)</td>
</tr>
<tr>
<td>Volcanic SO2</td>
<td>AeroCom Phase II (HCA0 v2; Diehl et al. 2012)</td>
<td>Monthly mean climatology (2° × 2.5°)</td>
</tr>
<tr>
<td>Biogenic terpene</td>
<td>Guenther et al. (1995)</td>
<td>Monthly mean varying (1980–96; 0.3125° × 0.25°)</td>
</tr>
<tr>
<td>Dimethyl sulfide (DMS) and</td>
<td>Lana et al. (2011)</td>
<td>Monthly mean varying (1997–2009; 0.3125° × 0.25°)</td>
</tr>
<tr>
<td>methanesulfonic acid (MSA)</td>
<td></td>
<td>Daily (2010 onward; 0.3125° × 0.25°)</td>
</tr>
<tr>
<td>Biomass burning</td>
<td>Scaled RETROv2 (Duncan et al. 2003)</td>
<td>Monthly mean varying (1980–2008; 0.1° × 0.1°)</td>
</tr>
<tr>
<td>Anthropogenic SO2</td>
<td>QFED 2.4-r6 (energy + non-energy) (European Commission 2011)</td>
<td>Annually varying (1980–2006; 1° × 1°)</td>
</tr>
<tr>
<td>Anthropogenic SO4, BC, and POM</td>
<td>AeroCom Phase II (HCA0 v1; Diehl et al. 2012)</td>
<td>Annually varying (1980–2005; 1° × 1°)</td>
</tr>
<tr>
<td>International ships (SO4, POM, BC)</td>
<td>AeroCom Phase II (HCA0 v1; Diehl et al. 2012) and Eyring et al. (2005)</td>
<td>Monthly varying (1980–2006; 1° × 1.25° × 72 levels)</td>
</tr>
<tr>
<td>Aircraft (SO2)</td>
<td>AeroCom Phase II (HCA0 v1; Diehl et al. 2012)</td>
<td>Monthly varying (1980–2006; 1° × 1.25° × 72 levels)</td>
</tr>
</tbody>
</table>

With a sea surface temperature (SST)-derived correction similar to the work of Jaeglé et al. (2011). The wind-dependent term in the Gong (2003) sea salt emission parameterization is updated to depend on friction velocity rather than 10-m velocity (i.e., $u_\kappa^{2.41}$, where $u_\kappa$ is the friction velocity). We note that there are errant emissions of sea salt over the Great Lakes and Caspian and Black Seas due to an issue in the lake-masking algorithm during the simulation. While there is minimal impact of lake emissions on the assimilated AOD fields, these inland sea salt emissions have consequences for other aerosol products such as speciated PM2.5 in some locations (Part II; Randles et al. 2016).

Sulfate and carbonaceous aerosol emissions (Figs. 1c–f) derive from both natural and anthropogenic sources. Table 1 outlines the emission inventories with additional details given in Randles et al. (2016). With the exception of volcanic SO2 emissions, when a given emissions inventory ends the final emission year is persisted in the model. Natural emissions of SO2 from volcanoes derive from the AeroCom Phase II project (Diehl et al. 2012; http://aerocom.met.no/) and cover eruptive and degassing volcanoes on all days from 1 January 1979 to 31 December 2010. Only a repeating annual cycle of degassing volcanoes is included in MERRA-2 after 2010. Eruptive volcanoes emit in the upper third of the column defined by the volcano elevation and the plume-top height; degassing volcanoes emit at the volcano elevation. With the exception of aircraft and energy-sector sulfur dioxide, anthropogenic aerosol sources emit into the lowest model layer. Energy-sector emissions of SO2 (EDGARv4.2; European Commission 2011) are emitted between 100 and 500 m above the surface as in Buchard et al. (2014).

Owing to the intense vertical mixing associated with fires, we distribute biomass burning emissions uniformly throughout the planetary boundary layer (PBL) in the grid box where the fire emission occurs. A diurnal cycle, which is more prominent in the tropics and gradually weakens in the higher-latitude extratropical temperate zones, is imposed online for biomass burning emissions. Importantly, biomass burning emissions of carbonaceous and sulfate aerosols in MERRA-2 derive from a variety of inventories over the course of the reanalysis. Figure 2 shows global and regional time series of the carbonaceous aerosol biomass burning emissions.

From 2010, daily emissions of biomass burning OC, BC, and SO2 derive from the Quick Fire Emissions Dataset (QFED) version 2.4-r6 (Darmenov and da Silva 2015). QFED is based on the top-down fire radiative power (FRP) approach. QFED draws on the cloud correction method used in the Global Fire Assimilation System (GFAS; Kaiser et al. 2012), but in addition employs a more sophisticated treatment of emissions from nonobserved land areas (Darmenov and da Silva 2015). FRP and locations of fires are obtained from
MODIS level-2 fire and geolocation products. Level-2 fire products are gridded at 0.3125° lon × 0.25° lat horizontal resolution and combined to create daily mean emissions at the same resolution. QFED emissions have been independently evaluated in the WRF-Chem model in the context of the recent NASA SEAC4RS field campaign (Saide et al. 2015) and in the offline version of the GOCART aerosol model (Petrenko et al. 2012).

Between 1997 and 2009 we employ a scaled version of the MODIS burned-area based Global Fire Emission Dataset (GFED) version 3.1 monthly biomass burning emissions (Randerson et al. 2006; van der Werf et al. 2006). We apply biome-dependent correction factors based on the period 2003–11 where both GFED v3.1 and QFEDv2.4-r6 were available. Spatially varying fractional contributions of emissions from the tropical sources reached the atmosphere.

**Fig. 2.** Time series of carbonaceous aerosol emissions (BC + POM; Tg month⁻¹) from biomass burning sources averaged globally (gray line; same in all panels) and over several major source regions. For the regional emissions, red and green shading indicate the relative contribution of BC and POM, respectively, to the total carbonaceous aerosol emissions; the total carbonaceous aerosol emissions are the sum of the POM and BC contributions (i.e., stacked area plots are shown, but POM emissions far exceed BC emissions). Comparing the regional and global time series, we show regional emissions events of global importance, such as the 1997 Indonesian fires.
forest, extratropical forest, savannah, and grassland biomes were calculated using the QFED climatology. These fractional contributions were then used to stratify the monthly GFED emissions by biome. Next, for each biome, the estimated GFED emissions were used to determine a scaling factor by means of a linear regression between the globally integrated QFED and GFED emissions. The final product, the QFED-scaled GFED emission, was computed as the sum over the biomes of the scaled biome-stratified emissions.

We apply this same type of correction to the RETROv2 monthly mean emissions (Duncan et al. 2003) that cover 1980–96 and are based on AVHRR, the Along Track Scanning Radiometer (ATSR), and the Total Ozone Mapping Spectrometer (TOMS) Aerosol Index (AI). We use the same aforementioned biome-specific scaling factors, justified by the similarities between global emissions from RETROv2 and GFEDv3.1. As can be seen in Fig. 2, on monthly mean time scales, the scaling process results in emissions that are globally consistent over time.

3. Goddard Aerosol Assimilation System

The MERRA-2 aerosol analysis uses the Goddard Aerosol Assimilation System (GAAS) that was first detailed in Buchard et al. (2015, 2016). Every 3 h, this system assimilates quality-controlled AOD at 550 nm into the GEOS-5/GOCART modeling system. The prognostic variable in GOCART is the three-dimensional forecasted column-integrated aerosol optical depth $\tau_f$ for each species $i$. The forecasted column-integrated aerosol optical depth $\tau_f$ is expressed as

$$\tau_f = \sum_{z_d} x_i \times b_{\text{ext},i}(\text{RH}, \lambda) \times \delta z,$$

(1)

where $\delta z$ is the atmospheric layer thickness and $b_{\text{ext},i}(\text{RH}, \lambda)$ is the species-specific extinction coefficient at wavelength $\lambda$ derived from Mie theory for spherical particles (Wiscombe 1980) or the T-matrix approach using the updated optics for nonspherical dust as described in Meng et al. (2010). The aerosol species $i$ include the 15 tracers previously described. The extinction coefficients $b_{\text{ext},i}(\text{RH}, \lambda)$ for sulfate and hydrophilic carbonaceous aerosol are a function of relative humidity (RH) following Chin et al. (2002); hygroscopic growth of sea salt aerosol follows Gerber (1985). Assumed optical properties are primarily from the Optical Properties of Aerosols and Clouds (OPAC) dataset (Hess et al. 1998) with updated dust optical properties that incorporate nonsphericity (Meng et al. 2010; Colarco et al. 2014). Table S1 (available in the online supplemental material) provides aerosol optical properties (extinction coefficient, single scattering albedo, and asymmetry parameter) by species at 550 nm as a function of relative humidity; assumed dry sizes for each aerosol species are also given.

The AOD analysis in MERRA-2 is performed by means of the so-called analysis splitting method. First, a 2D analysis of AOD is performed using error covariances derived from innovation data using the maximum-likelihood method of Dee and da Silva (1999). The AOD analysis equation is

$$\tau^a = Hx^a = H(x^f + \delta x^a)$$

$$= \tau^f + \delta \tau^a$$

$$= \tau^f + \text{HP}^fH^T(\text{HP}^fH^T + R)^{-1}(\tau^o - Hx^f),$$

(2)

where the superscripts $a, f,$ and $o$ indicate the analysis, forecast (background), and observation, respectively, and $H$ is the linear observation operator that converts aerosol mass to AOD. The operators $P^f$ and $R$ are the background and observation error covariance matrices, respectively. The AOD analysis increments $\delta \tau^a$ are computed using a 2D version of the Physical-Space Statistical Analysis System (PSAS; Cohn et al. 1998). For algorithm consistency, this analysis is performed using a natural log-transformed control variable $[\eta = \ln(\text{AOD} + \epsilon)$, where $\epsilon = 0.01]$ as detailed in section 3a.

Once the AOD analysis increments are obtained, the next step is to derive 3D analysis increments for the mixing ratio of each aerosol species, $\delta x^a_i$. Previous studies (e.g., Zhang et al. 2008, and references therein) have opted to simply scale the mixing ratio increments to promote consistency with the analyzed AOD at each grid point, a procedure that does not make any use of error covariance information. Here we seek a relationship relating analysis increments of AOD to analysis increments of aerosol species mixing ratio that involve the corresponding error covariance operators. From the mixing ratio analysis equation implicit in Eq. (2), the vertical structure of $\delta x^a$ is determined by the operator $P^fH^T$. Therefore, we seek an operator $Q$ such that

$$\delta x^a = P^fH^TQ\delta \tau^a.$$

(3)

Imposing the condition $H\delta x^a = \delta \tau^a$ leads to

$$Q = (\text{HP}^fH^T)^{-1}.$$

(4)

Substituting Eq. (4) into Eq. (3) we arrive at

$$\delta x^a = P^fH^T(\text{HP}^fH^T)^{-1}\delta \tau^a.$$

(5)

Notice that the observation error covariance matrix $R$ is not involved in Eq. (5), and that this expression is
invariant to any scaling of the background error covariance operator $P^f$.

In principle, solving Eq. (5) requires the 3D error covariance operator, including vertical and horizontal correlations. However, for computational reasons, we solve this equation for each vertical column separately, as the main purpose of this step is to project the horizontal AOD increments into the vertical and across species. As described in Buchard et al. (2015) and section 3b, we employ a local displacement ensemble (LDE) formulation to solve Eq. (5).

a. Choice of control variable for the 2D AOD analysis

Since AOD is not a normally distributed variable (e.g., O’Neill et al. 2000), the 2D analysis in Eq. (2) is performed using the natural log-transformed AOD,

$$\eta = \ln(\tau + \varepsilon),$$

as the control variable. The parameter $\varepsilon = 0.01$ is chosen to render the distribution closest to a Gaussian using a probability plot technique (Chambers et al. 1983). Notice that for small values of $\tau$ the log-transformed variable $\eta$ is linear in $\tau$ and approaches $\ln\tau$ for large AOD. This choice of control variable avoids the classical problem of log-normal distributions for small values of $\tau$ and allows for multiplicative rather than additive corrections for large $\tau$ (Henze et al. 2009; Saide et al. 2013). Notice that AOD errors $\tau'$ are related to $\eta'$ errors by

$$\tau' = (\tau + \varepsilon)\eta'.$$

Therefore, even when $\eta$ errors are Gaussian and flow independent, AOD errors are a sum of flow-dependent ($\eta'\tau'$) and flow-independent ($\eta'\varepsilon$) components. To preserve linearity, analysis increments $\delta\eta'$ are converted back to $\delta\tau'$ before use in Eq. (10).

b. Local displacement ensembles

To evaluate Eq. (5) we employ an ensemble formulation. Let

$$X = (x_1 \ x_2 \ \ldots \ x_E),$$

where $X$ is a $n_q \times n_E$ matrix ($n_q$ is the number of aerosol concentration tracers times the number of vertical levels and $n_E$ is the number of ensemble perturbations), for a particular column. From Eq. (2) it follows that

$$Y = HX = (Hx_1 \ Hx_2 \ \ldots \ Hx_E) = (\tau_1 \ \tau_2 \ \ldots \ \tau_E).$$

Approximating the background error covariance matrix by $P^f \sim XDX^T$, where $D$ is a diagonal matrix allowing for the weight of the ensemble perturbations, Eq. (5) can be written as the unbiased linear regression equation:

$$\delta x^a = XDY^T(YD^TY^T)^{-1}\delta r^a.$$  

At this point we have made no assumption about the nature of the ensemble perturbations. MERRA-2 did not include an ensemble of aerosol forecasts, and this practical approach was developed to produce ensemble perturbations capable of producing realistic speciation and vertical structures for the mixing ratio analysis increments. The underlying assumption of our error covariance modeling exercise is that aerosol forecast errors are due primarily to misplacements of aerosol plumes. Implicit in this assumption is that the AOD analysis [Eq. (2)] removes any systematic biases. For each grid point, ensemble perturbations are formed by computing the difference of background aerosol mixing ratios from this central grid point and adjacent grid points within a given radius (taken as 1000 km in MERRA-2). Ensemble perturbations are weighted according to $\exp[-4(\tau' - \tau^a)]$ so that nearby grid points that better match the AOD analysis receive higher weights.

c. AOD background correction and approximate analysis averaging kernel

The MERRA-2 meteorological analysis is performed every 6 h, while the AOD analysis occurs every 3 h. For efficiency reasons, the overall analysis cycle in MERRA-2 is controlled by the meteorological assimilation, with two independent AOD analyses performed within each 6-h cycle. Consider the 1200 UTC analysis cycle. The AOD analysis is performed at 0900 and 1200 UTC, using backgrounds that are forecasts from 0600 UTC. The corresponding update of the GEOS-5 aerosol state occurs at 0900 and 1200 UTC. At 1200 UTC the proper background state should be a forecast from 0900 not from 0600 UTC as in the (off-line) AOD analysis. Therefore a background correction is in order to account for this mismatch in background states. Denoting the previous AOD analysis and background by $\tau^a$ and $\tau^f$, respectively, it can be shown that

$$\tau^a = \tau^f + (I - A)(\tau^f - \tau^a),$$

where $\tau^a$ and $\tau^f$ are the proper analysis and forecast at 1200 UTC, and $A = KH$ is the analysis averaging kernel with $K$ being the usual Kalman gain. For typical satellite swaths, the operator $A$ evaluates to approximately zero outside the swath leading to a simple replacement of the background in those regions; elsewhere, Eq. (11)
provides a background correction that depends on the details of the analysis.

In practice, a diagonal approximation for the analysis averaging kernel $A$ is utilized. In such an approximation, an additional AOD analysis is performed with all innovations set to 1 while preserving the actual observational coverage. This averaging kernel field is computed as the second step of each AOD analysis and is provided as an additional diagnostic for MERRA-2. Such an algorithm can be derived as a limiting case of a banded approximation for the Kalman gain $K$ (derivation not shown).

d. AOD observing system and bias correction

Table 2 summarizes the AOD observing system used in MERRA-2 while Fig. 3 shows the total number of monthly observations by sensor for the entire reanalysis. Assimilation of satellite aerosol products requires careful data quality control and bias removal (Zhang and Reid 2006; Lary et al. 2009). MERRA-2 includes assimilation of bias-corrected AOD derived from AVHRR and MODIS radiances. Our approach involves cloud screening and homogenization of the observing system by means of a neural net retrieval (NNR) that translates cloud-cleared observed radiances into AERONET-calibrated AOD. Additional details on the NNR algorithm are given below, and supplemental Table S2 provides observational error standard deviations and Kalman gains (the ratio of the background error variances to the innovation error variances) for each sensor used in the PSAS-based analysis of the log-transform AOD $\eta$. For reference, the background error decorrelation length scale is 140 km, and the background error standard deviation is 0.45. Homogeneous and isotropic covariance models are assumed, with spatially constant variances. Error covariance model parameters were estimated using the maximum-likelihood method of Dee and da Silva (1999).

For the pre-EOS period and until 2002, we assimilate bias-corrected AOD derived from the 25-yr record of AVHRR radiances (Heidinger et al. 2014). We discontinue AVHRR assimilation after 2002 when MODIS Aqua becomes available; both instruments have afternoo

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Temporal coverage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVHRR NNR</td>
<td>1980–August 2002</td>
<td>PATMOS-x radiances over ocean only (PM orbit)</td>
</tr>
<tr>
<td>AERONET</td>
<td>Station dependent (1999–October 2014)</td>
<td>AOD from land station network</td>
</tr>
<tr>
<td>MISR</td>
<td>February 2000–June 2014</td>
<td>AOD over bright land surfaces only (albedo &gt; 0.15)</td>
</tr>
<tr>
<td>MODIS Terra NNR</td>
<td>March 2000 onward (NRT)</td>
<td>Collection 5 “Dark Target” land and ocean radiances (AM orbit)</td>
</tr>
<tr>
<td>MODIS Aqua NNR</td>
<td>August 2002 onward (NRT)</td>
<td>Collection 5 “Dark Target” land and ocean radiances (PM orbit)</td>
</tr>
</tbody>
</table>
adaptive buddy check of Dee et al. (2001) is used before assimilating AOD data into GEOS-5.

To highlight the consistency imposed by the NNR algorithm on the MERRA-2 aerosol observing system, Fig. 4 compares the AVHRR and MODIS *Aqua* NNR AOD during a month when both instruments were available. Saide et al. (2013) independently evaluated the MODIS NNR AOD against other observational datasets using the WRF-Chem data assimilation system. Postprocessing techniques such as the NNR were shown to reduce biases in the assimilation relative to independent AERONET AOD observations. Additionally, compared to other bias-corrected MODIS AOD products (e.g., Zhang and Reid 2006), the NASA NNR retrieval produced a more positive impact on assimilated AOD because its less restrictive cloud fraction requirement increases data availability.

e. Control simulation without AOD assimilation

GEOS-5 can be run in a replay mode whereby a previous meteorological analysis, generated with an identical model version, is used to adjust the model’s meteorological state (winds, temperature, specific humidity) much like a chemical transport model (CTM). However, unlike a CTM, in a replay simulation the aerosol transport dynamics are entirely consistent with the model thermodynamic state at every time step between analysis updates. Using the same version of GEOS-5 as MERRA-2, we perform an EOS-period control simulation (M2REPLAY) driven by the exact MERRA-2 meteorology but without the aerosol optical depth assimilation. The aerosol fields from this control simulation are compared to those from MERRA-2 to demonstrate the impact of the AOD assimilation on aerosol fields (section 4 herein; see also Part II).

4. Results

Figure 5a shows the time series of global, area-weighted monthly mean analyzed AOD from MERRA-2. The contribution of different aerosol species to the total AOD is also shown. For comparison, we show the AOD from the previous MERRAero reanalysis (Buchard et al. 2015, 2016). Major features of the
FIG. 5. (a) Global, area-weighted average monthly mean 550-nm AOD from MERRA-2 (black line) and MERRAero (magenta) and contributions from various aerosol species [color shading MERRA-2 only]. Carbonaceous aerosol AOD (green) is the sum of organic plus black carbon AOD. Note that only the total AOD (thick black line) is directly constrained by the aerosol assimilation; aerosol speciation depends strongly on emission and loss processes in the model. (b) Monthly mean, collocated in time and space (pairwise in observation space), comparison of the 3-h model forecast $F$ and observations $O$ for AVHRR NNR, MODIS Aqua land NNR (MYDL), and MODIS Aqua ocean NNR (MYDO). Here the lines represent the observed AOD, and shading indicates the difference between $F$ and $O$. (c) Monthly mean statistics computed in log-transformed observation space. Statistics comparing $O$ and $F$ are shown as dashed lines, and relationships between the observations and analysis $A$ are shown as solid lines for the same sensors as in (b). Supplemental figures show the statistics for the other sensors in the MERRA-2 aerosol observing system (Table 2). The log-space errors (standard deviation) in AOD are related to linear errors according to Eq. (7). Note that we expect the monthly mean statistics presented here to outperform statistics sampled at higher time frequencies.
reanalysis include large increases in global AOD following major volcanic eruptions in the pre-EOS period and the seasonal cycle of AOD associated primarily with dust and biomass burning (carbonaceous) aerosols. During the EOS period, MERRA-2 AOD is slightly higher than the AOD from MERRAero, due to higher dust and sea salt AOD that result from changes in model physics, meteorology, and the inclusion of MISR AOD assimilation over bright desert regions in MERRA-2 (Randles et al. 2016).

**a. AOD innovation statistics**

Figure 5b shows collocated observed O and forecasted F AOD from AVHRR NNR and MODIS Aqua NNR (separately over land and ocean). Figure 5c shows the statistical relationship between the collocated assimilated observations of AOD, the forecasted AOD, and the analyzed AOD A for these sensors. Figures S1–S3 show similar plots for other sensors in the MERRA-2 aerosol observing system (Table 2). Additionally, Fig. S4 shows the probability distribution functions of $O - F$ and $O - A$ for each sensor in the AOD observing system. These comparisons serve as a sanity check, as it is expected that the analysis should have higher correlation and lower bias and root-mean-square error (RMSE) with respect to the assimilated observations compared to the forecast. Note that while the analysis is statistically closer to the observations than the forecast, for any given sensor the bias is not completely eliminated due in part to the influence of other sensors on the analyzed AOD. During the EOS period, analysis statistics are also slightly better over ocean compared to land, due to a combination of observability (fewer observations over land versus ocean) and a prevalence of aerosol source regions over land (greater aerosol type variability) (Gelaro et al. 2017; Randles et al. 2016).

Generally, forecasted AOD departs only slightly from collocated assimilated observations (Fig. 5b). The exception to this is after major volcanic eruptions (El Chichón in 1982 and Pinatubo in 1991) when the MERRA-2 forecasted AOD is significantly higher than the assimilated AVHRR observations. Part II describes the reasons for this in more detail for the Pinatubo eruption. Briefly, MERRA-2 first overestimates the volcanic plume height from Pinatubo, injecting sulfur dioxide gas higher into the stratosphere than more recently available data suggest. This higher injection height has implications for transport and lifetime of the subsequently formed stratospheric sulfate aerosol,
particularly since the assimilation does not constrain the precursor sulfur dioxide gas. Next, GOCART assumes all sulfate aerosols share the same (dry) size distribution when calculating aerosol optical properties. The sulfate aerosol size in GOCART is more appropriate for tropospheric aerosols rather than the particles that formed in the stratosphere after the Pinatubo eruption (Aquila et al. 2014). Smaller aerosol particles are more efficient at scattering light, and the underestimate in stratospheric aerosol particle size contributes to an overestimate of stratospheric and total AOD in the forecast.

b. Impact of the AOD analysis

Figure 6 compares, for each of the satellite sensors during the EOS period, the 3-h forecasted and analyzed climatological AOD. We also show the climatological AOD from the replay control simulation where we have sampled M2REPLAY using the daytime Terra and Aqua orbits. In most regions, the analyzed MERRA-2 AOD resembles the forecasted AOD. Exceptions appear, for example, over China and northern Africa. Both the MERRA-2 forecasted AOD (which retains the influence of the previous assimilation step) and the analyzed AOD fields show differences relative to the control experiment. Without AOD assimilation, M2REPLAY generally has higher overall AOD, particularly over the Southern Ocean and North Atlantic, over and downwind of China, and over the Sahara.

Figure 7 shows observed, forecasted, and analyzed AOD from AVHRR for all of MERRA-2 in which we assimilate observations from this sensor and for the two years during and subsequent to the Pinatubo eruption. Here we see, as in Figs. 5b and 5c, that the 3-h forecast overestimates the observed AOD after Pinatubo. However, the analysis resembles the observations after Pinatubo. This implies a negative aerosol increment. Importantly, when a negative AOD increment is applied, it impacts all aerosol species, proportionally reducing their masses to optimally reproduce observed AOD. This can be seen in the time series of AOD in Fig. 5a, where dust, carbonaceous, and sea salt AOD is reduced immediately after the Pinatubo eruption. This speciated AOD reduction is an artifact of applying relatively large, negative AOD increments in MERRA-2. We note that both two-dimensional total AOD increments and three-dimensional speciated aerosol mass mixing ratio increments are available as part of the MERRA-2 dataset (GMAO 2015g,h). Ongoing
work is focused on improved forward modeling of large stratospheric volcanic eruptions to mitigate the need for large, negative AOD analysis increments [i.e., δτ_a in Eq. (2)].

c. Independent verification

Here we assess the performance of the analyzed AOD fields by comparing them with other, independent (nonassimilated) observations from both surface-based and aircraft sensors. To examine the impact of the AOD analysis, we perform these comparisons for the analyzed fields and, where available, AOD fields from the control experiment. We reiterate that it is challenging to independently validate the MERRA-2 aerosol products because most of the global, readily available spaceborne and ground-based observations are included in the assimilation. A more in-depth validation of the aerosol analysis, including other aerosol properties (vertical distribution, absorption, surface concentration) and case studies of aerosol events, is presented in Part II. In this work, we focus on validation of the total column AOD only, since this is the aerosol property directly constrained in the assimilation.

1) COMPARISONS WITH AERONET

The Aerosol Robotic Network (AERONET; http://aeronet.gsfc.nasa.gov/new_web/) is a federated global network of ground-based, automatic sun photometers that measure direct sun and sky radiances at several wavelengths (Holben et al. 1998). AOD is obtained from direct sun measurements with an accuracy to within ±0.015. In the assimilation, we use cloud-screened level 2.0 data (quality assured; Smirnov et al. 2000). Although AERONET is assimilated in MERRA-2 after 1999 (Table 2) and cannot be considered for independent verification, a subset of sites (~40) have unassimilated level 2.0 data going back to 1993. It is instructive to examine the performance of the MERRA-2 analyzed AOD fields relative to these observations.


<table>
<thead>
<tr>
<th>Station name</th>
<th>Dominant aerosol type</th>
<th>Geographic region</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Number of collocations</th>
<th>Natural log-transformed AOD (r)</th>
<th>Natural log-transformed AOD (RMSE)</th>
<th>MERRA-2</th>
<th>M2REPLAY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jabiru</td>
<td>Smoke</td>
<td>Australia</td>
<td>−12.661</td>
<td>123.893</td>
<td>21 917</td>
<td>0.82 (0.44)</td>
<td>0.68 (0.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lake Argyle</td>
<td>Smoke</td>
<td>Australia</td>
<td>−16.108</td>
<td>128.749</td>
<td>27 894</td>
<td>0.81 (0.71)</td>
<td>0.80 (0.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alta Floresta</td>
<td>Smoke</td>
<td>Amazon basin</td>
<td>−9.871</td>
<td>−56.104</td>
<td>16 942 (1650)</td>
<td>0.92 (0.77)</td>
<td>0.77 (0.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bratts Lake</td>
<td>Smoke</td>
<td>Canada</td>
<td>50.280</td>
<td>−104.700</td>
<td>16 825 (919)</td>
<td>0.85 (0.85)</td>
<td>0.68 (0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rimrock</td>
<td>Smoke</td>
<td>U.S. Northwest</td>
<td>46.487</td>
<td>−116.992</td>
<td>15 771</td>
<td>0.75 (0.51)</td>
<td>0.59 (0.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSRN BAO</td>
<td>Smoke, dust</td>
<td>U.S. Great Plains</td>
<td>40.045</td>
<td>−105.006</td>
<td>19 545</td>
<td>0.82 (0.44)</td>
<td>0.70 (0.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boulder</td>
<td>Smoke, dust</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mongu</td>
<td>Smoke, dust</td>
<td>Africa</td>
<td>−15.255</td>
<td>23.151</td>
<td>19 387 (4831)</td>
<td>0.93 (0.71)</td>
<td>0.81 (0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skukuza</td>
<td>Smoke, dust</td>
<td>Southeast Africa</td>
<td>−24.992</td>
<td>31.587</td>
<td>17 149 (1803)</td>
<td>0.89 (0.84)</td>
<td>0.75 (0.45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ilorin</td>
<td>Smoke, dust</td>
<td>Sahel</td>
<td>8.320</td>
<td>4.340</td>
<td>15 081 (2255)</td>
<td>0.92 (0.76)</td>
<td>0.71 (0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Railroad Valley</td>
<td>Smoke, dust</td>
<td>U.S. Southwest</td>
<td>38.504</td>
<td>−115.962</td>
<td>23 984</td>
<td>0.81 (0.60)</td>
<td>0.70 (0.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sevilleta</td>
<td>Dust</td>
<td>U.S. Southwest</td>
<td>34.355</td>
<td>−106.885</td>
<td>22 144 (8222)</td>
<td>0.84 (0.68)</td>
<td>0.72 (0.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arica</td>
<td>Dust</td>
<td>Northern Chile</td>
<td>−18.472</td>
<td>−70.313</td>
<td>15 906 (3446)</td>
<td>0.67 (0.47)</td>
<td>0.18 (0.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saada</td>
<td>Dust</td>
<td>Northwest Africa</td>
<td>31.626</td>
<td>−8.156</td>
<td>21 290</td>
<td>0.84 (0.43)</td>
<td>0.82 (0.58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capo Verde</td>
<td>Dust</td>
<td>Northwest Africa</td>
<td>16.733</td>
<td>−22.935</td>
<td>24 324 (4445)</td>
<td>0.94 (0.84)</td>
<td>0.81 (0.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dakar</td>
<td>Dust</td>
<td>Sahel</td>
<td>14.394</td>
<td>2.665</td>
<td>35 365 (3958)</td>
<td>0.80 (0.64)</td>
<td>0.65 (0.58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Banizoumbou</td>
<td>Dust</td>
<td>Sahel</td>
<td>13.541</td>
<td>23.594</td>
<td>28 947</td>
<td>0.82 (0.40)</td>
<td>0.68 (0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IER Cinzana</td>
<td>Dust</td>
<td>Sahel</td>
<td>13.278</td>
<td>−7.594</td>
<td>28 947</td>
<td>0.82 (0.40)</td>
<td>0.68 (0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dalandzadag</td>
<td>Dust</td>
<td>Inner Mongolia</td>
<td>43.577</td>
<td>104.419</td>
<td>18 701 (2692)</td>
<td>0.82 (0.75)</td>
<td>0.70 (0.60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solar Village</td>
<td>Dust</td>
<td>Saudi Arabia</td>
<td>24.907</td>
<td>46.397</td>
<td>33 837 (2704)</td>
<td>0.94 (0.92)</td>
<td>0.73 (0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sede Boker</td>
<td>Dust</td>
<td>Israel</td>
<td>30.855</td>
<td>24.782</td>
<td>39 027 (3225)</td>
<td>0.89 (0.80)</td>
<td>0.73 (0.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMS-METU-ERDEMLI</td>
<td>Dust</td>
<td>Turkey</td>
<td>36.565</td>
<td>24.355</td>
<td>16 444</td>
<td>0.83 (0.32)</td>
<td>0.67 (0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nes Ziona</td>
<td>Dust</td>
<td>Israel</td>
<td>31.922</td>
<td>24.782</td>
<td>23 640</td>
<td>0.89 (0.31)</td>
<td>0.70 (0.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSFC</td>
<td>Polluted</td>
<td>U.S. Northeast</td>
<td>38.992</td>
<td>−76.840</td>
<td>24 331 (6479)</td>
<td>0.94 (0.84)</td>
<td>0.79 (0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD Science Center</td>
<td>Polluted</td>
<td>U.S. Northeast</td>
<td>39.283</td>
<td>−76.617</td>
<td>19 285 (528)</td>
<td>0.89 (0.86)</td>
<td>0.72 (0.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wallops</td>
<td>Polluted</td>
<td>U.S. Northeast</td>
<td>37.942</td>
<td>−75.475</td>
<td>10 652 (1924)</td>
<td>0.93 (0.90)</td>
<td>0.79 (0.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexico City</td>
<td>Polluted</td>
<td>Central Mexico</td>
<td>19.334</td>
<td>−99.182</td>
<td>15 830 (1045)</td>
<td>0.58 (0.54)</td>
<td>0.54 (0.97)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
to observations at these sites. Additionally, comparing the performance of MERRA-2 AOD to M2REPLAY shows the impact of the analysis at these stations.

Table 3 compares simulated AOD to observations from long-term AERONET stations (>7 yr of data), some of which also had some portion of their observations occurring prior to 1999. Statistics are computed based on hourly observations and collocated model output, and are provided both for each station’s full EOS-period record (2000–14), and, in parentheses, just that part of the data record available prior to 1999. For each station, the dominant aerosol source type is indicated. For most stations, comparing MERRA-2 to M2REPLAY reveals that the assimilation increases correlation while reducing the RMSE of the differences with the observed hourly AOD. Notably, even without assimilation, M2REPLAY shows considerable skill in capturing the monthly mean AOD without aerosol assimilation (due to the use of observationally based biomass burning emissions inventories described in section 2b).

In all cases, the variability in observed AOD is better captured by the assimilation (MERRA-2) compared to the control run (M2REPLAY).

Prior to 1999, MERRA-2 assimilates AOD from AVHRR over ocean only, and land-based observations from AERONET are assimilated starting in 1999. EOS-era satellite observations are assimilated after 2000, providing satellite observations of AOD over both land and ocean. We cannot disentangle the influence of AERONET observations alone on the AOD assimilation. However, we do see that the EOS-era observations generally better simulate AOD variability (standard deviation) compared to the impact of AVHRR ocean observations only prior to 1999. For example, prior to 1999, MERRA-2 does not generally have a standard deviation comparable to the observations in summer at GSFC (Fig. 8a), a location with AOD dominated by anthropogenic sources that have only annually varying emissions (Table 1).
The time series of observed and simulated AOD at Mongu (Fig. 8b) illustrates the combined impact of changing biomass burning emissions and the observing system on the analyzed AOD. Recall that prior to 1997, MERRA-2 uses emissions based on a scaled version RETROv2 (Duncan et al. 2003) and does not include assimilated observations over land. After 1997, emissions are based on a scaled version of GFED and QFED, both of which derive from MODIS data products. There is a clear improvement both in the monthly mean AOD and standard deviation of AOD after 1997 when emissions change. Further improvement is seen during the EOS period when many more observations over land become available. We note that the change from monthly mean to daily biomass burning emissions (before and after 2010, respectively) has only a minor impact on the multiyear statistical agreement between MERRA-2 and AERONET (not shown); however, the biomass burning emissions frequency is likely more important in the context of individual events. We conclude that, importantly, time series analysis of MERRA-2 AOD is sensitive both to a changing observing system and changes and/or trends imposed by emission inventories.

2) COMPARISONS WITH MARITIME NETWORKS

We use historical shipborne sun photometer data to validate MERRA-2 AOD prior to the EOS period. These data, summarized in Smirnov et al. (2002) and Sakerin and Kabanov (2002) and partially used by Liu et al. (2004) for AVHRR validation, span the time frame 1982–96 and cover the cruises shown in Fig. 9a. We compare the cosampled MERRA-2 AOD at the observed wavelength closest to 550 nm (range of 500–570 nm depending on the cruise). Observations are reported as morning or afternoon averages with an observed accuracy of ±0.02. The model is sampled every 3 h for a given observation date and then averaged from 0800 to 1200 or from 1200 to 1600 local time for purposes of comparison. Figure 9b shows a scatterplot of the AOD comparison, with statistics reported by year. The correlation between MERRA-2 and the observations is best near dust-influenced regions (e.g., the Red Sea, the Mediterranean, and near North Africa). In the remote Pacific, MERRA-2 AOD is not well correlated with the observations, but the bias is low. Observations from a single cruise indicate that MERRA-2 AOD is biased high after the Pinatubo eruption.

During the EOS period, the Maritime Aerosol Network (MAN; http://aeronet.gsfc.nasa.gov/new_web/man_data.html) employs Microtops II sun photometers aboard ships of opportunity to measure AOD. The sun photometers used in MAN are calibrated to have an estimated uncertainty in AOD of ±0.02 (Smirnov et al. 2009). MAN cruises cover the period from 2004 to the present (Fig. 10a). These observations have not been assimilated in the MERRA-2 GAAS and therefore serve as independent validation of the assimilated AOD product. Figure 10b compares all available MAN observations with collocated MERRA-2 AOD. A high degree of correlation is found between the MERRA-2 and MAN observations, and the bias is generally low, although MERRA-2 does tend to slightly overestimate the lowest observed AOD. The impact of AOD assimilation is apparent in Fig. 10c, where the M2REPLAY control simulation has lower correlation and higher bias relative to the MAN observations than MERRA-2.

3) COMPARISONS WITH AIRCRAFT OBSERVATIONS

The NASA Langley Research Center (LaRC) Differential Absorption Lidar (DIAL) system implements the High Spectral Resolution Lidar (HSRL) technique to retrieve aerosol extinction and AOD at 532 nm (Hair et al. 2008). The instrument also retrieves aerosol backscatter coefficients and is sensitive to polarization at three wavelengths (355, 532, and 1064 nm), measuring both above and below the aircraft (i.e., zenith and nadir). AOD values are derived from nadir aerosol extinction profiles when the aircraft is above 6 km; a 1-km region below the aircraft is omitted from the AOD column as the laser and telescope are not fully aligned in this region.

During the NASA Studies of Emissions and Atmospheric Composition, Clouds, and Climate Coupling by Regional Surveys (SEAC4RS) campaign from August to September 2013 (http://www.nasa.gov/mission_pages/seac4rs), the DIAL/HSRL system aboard the NASA DC8 aircraft measured AOD over a large portion of the southeastern and western continental United States, including several smoke plumes from large fires (e.g., the Rim Fire). We sample MERRA-2 along the aircraft trajectories for the entire SEAC4RS campaign and calculate the AOD over the same portion of the column as the observations. Figure 11a shows the performance of the MERRA-2 AOD relative to this independent observation. There is a high degree of correlation and low bias between MERRA-2 and AOD measured by the DIAL/HSRL system. As with the MAN comparison (Fig. 10), there is a high and low bias for low and high AOD, respectively. Figure 11b shows the same comparison for the control run simulation (M2REPLAY) and again demonstrates the positive impact of AOD assimilation. While we have shown the AOD comparison for the entire SEAC4RS campaign, Part II uses DIAL/HSRL observations to...
validate vertical profiles of aerosol optical properties for several aircraft campaigns and during the Yosemite Rim Fire.

The Spectrometer for Sky-Scanning, Sun-Tracking Atmospheric Research (4STAR) measured the AOD above the DC8 during the SEAC4RS campaign. 4STAR combines airborne sun tracking and sky scanning with spectroscopy to cover the full 350–1700-nm spectral range (Dunagan et al. 2013). The 4STAR AOD is computed from the observed intensity of the direct solar beam based on the Beer–Lambert law. Shinozuka et al. (2013) give details on AOD data acquisition, screening, reduction, calibration, and uncertainty analysis. During SEAC4RS, where ambient temperature varied more widely than in 4STAR’s previous deployments, the measurements exhibited a systematic high bias in AOD. An empirical correction determined from field data and laboratory tests has been applied to the AOD such that, where the magnitude of the correction is comparable to the AOD itself or where the 4STAR temperature is below $215^\circ C$, AOD values are flagged and not reported. This quality screening improves the accuracy of reported AODs, but also disproportionally eliminates AODs under conditions of low aerosol burden or high altitude. Comparisons with ground sun photometers before and after each flight allowed for characterization of 4STAR’s optical throughput, enabling an estimate of measurement uncertainty of $\pm 0.02$ for SEAC4RS.

Figure 12 compares probability distribution functions (PDFs) of collocated simulated AOD to 4STAR for both
Fig. 9. Comparison of MERRA-2 and historical shipborne AOD observations. (a) Map showing the location of the ship cruises spanning the period 1982–96. (b) Scatterplot of AOD spanning various years (indicated by different colors) and different cruises (indicated by marker shape). The x-axis error bars represent the assumed observed AOD error of 0.02, and y-axis error bars are the standard deviation of the MERRA-2 AOD used to make the morning or afternoon averages (see text for sampling strategy).
FIG. 10. (a) MAN cruises 2004–15, color coded by year (see the MAN website for information on specific cruises). (b) Joint PDF comparison of collocated MAN-observed and MERRA-2 AOD for the same period. (c) As in (b), but for the M2REPLAY control simulation without AOD assimilation. Note that statistics are calculated in natural log-transformed AOD space, the color bar represents probability density, and the x and y axes have been relabeled in linear AOD space for clarity.
MERRA-2 and M2REPLAY for the entire SEAC4RS campaign. Here, we have sampled the simulations along the aircraft track, and AOD is calculated for only the portion of the atmospheric column above the aircraft. Note that differences between observed and simulated AOD at the lowest tail of the PDFs, which are representative of cleaner background air remote from major aerosol sources, are generally within the reported instrumental uncertainty. Because AOD here is not representative of the full atmospheric column, simulated AOD is sensitive to the vertical distribution of aerosol, which is not directly constrained in the assimilation. The assimilation best characterizes the midrange of AODs (0.13–0.21) compared to the control run. However, it underestimates the highest observed AODs (associated with biomass burning events; not shown); without assimilation (M2REPLAY), the model overestimates these high-AOD occurrences.

d. Clear-sky aerosol direct radiative effects

Atmospheric aerosols, both natural and anthropogenic, impact climate through scattering and absorption of radiation [direct radiative effect (DRE)], modification of cloud microphysics (indirect effects), and thermodynamic effects (semidirect effect of aerosol absorption). Estimating the direct radiative effect requires knowledge of the three-dimensional distribution of aerosols and their optical properties. While satellites can measure AOD, a key aerosol property for determination of the DRE, cloud contamination and satellite viewing geometries can combine to produce spatial and temporal sampling biases. Furthermore, global observations of aerosol absorption optical depth (AAOD) and vertical distributions are currently even more sparsely available. Although they simulate the full aerosol life cycle without data gaps, global aerosol models are complex and produce varying estimates of DRE (Kinne et al. 2006; Schulz et al. 2006; Yu et al. 2006). In the MERRA-2 aerosol reanalysis, however, AOD is continuously available for over two decades and is optimally constrained by quality controlled satellite- and ground-based observations.

Before examining the DRE from MERRA-2, we first compare the global average AOD and AAOD to other models and reanalyses. Several forecasting centers are currently producing global aerosol reanalyses. For example, the Naval Research Laboratory (NRL) has produced an 11-yr offline aerosol reanalysis at 1° resolution using the Navy Aerosol Analysis and Prediction System (NAAPS) to assimilate quality-assured and controlled MODIS Collection 5 and MISR AOD (Lynch et al. 2016). As part of the Monitoring Atmospheric Composition and Climate (MACC) project, the European Centre for Medium-Range Weather Forecasts (ECMWF) has assimilated MODIS AOD from 2003 to the present (Morcrette et al. 2009; Benedetti et al. 2009).

Table 4 and Gelaro et al. (2017, their Fig. 14) compare the global-average AOD and AAOD from MERRA-2 to recent reanalysis estimates, including our own MERRAero offline aerosol reanalysis. Where such information is available, results are also partitioned by
species and identified as either fine or coarse mode (see table footnotes for details). Also shown are the multimodel average results from phase I of the AeroCom intercomparison project (Kinne et al. 2006), as well as both model and observational estimates from Yu et al. (2006). Yu et al. (2006) attempt to account for satellite clear-sky biases by combining MODIS and MISR observations with the GOCART aerosol model as part of their observational estimate. Compared to MERRAero, MERRA-2 has slightly higher global average AOD due to increased contributions from dust (related to the assimilation of MISR AOD over bright surfaces) and sea salt (related to changes in model physics). MERRA-2 and NAAPS have similar global average AOD both for fine and coarse model aerosol. Global aerosol model estimates of AOD (AeroCom and the Yu model) are lower than MERRA-2, and the observational estimate (Yu Obs) is higher. The MACC reanalysis (Bellouin et al. 2013) AOD is close to the MODIS-only AOD of 0.188 (Yu et al. 2006). MACC also has more dust and sea salt AOD compared to MERRA-2, especially over the ocean (see Table S3).

Both AOD and AAOD influence the direct impact of aerosols on the radiative energy balance of the planet. Recall that AAOD is only indirectly constrained by assimilating AOD. MERRA-2 AAOD is slightly lower than our previous aerosol reanalysis but agrees well with the AAOD from Bellouin et al. (2013) (note that the AAOD is not the MACC-native AAOD but is derived in that study). While the column-integrated AAOD from MERRA-2 generally agrees well with OMI observations in the near-UV (especially near source regions; Buchard et al. 2016), MERRA-2 overestimates absorption in regions remote from aerosol sources (e.g., the free troposphere) due to excessive amounts of black carbon aerosol in these regions [see Randles et al. (2016) for comparisons of MERRA-2 to black carbon vertical distributions from aircraft observations].

Recall that the DRE is defined as the shortwave flux difference (in W m$^{-2}$) between clear-sky (i.e., no clouds) and clear, clean-sky conditions (i.e., no aerosols or clouds). In the absence of clouds, the radiative effect of aerosols is less sensitive to the vertical distribution of aerosol absorption, although it remains sensitive to absorbing aerosols over bright surfaces (e.g., snow and deserts; Chylek and Coakley 1974). As long-term aerosol reanalyses such as MERRA-2 continue to evolve and improve, they can potentially reduce uncertainty in the DRE, particularly once better observational constraints on aerosol absorption and vertical distribution become available and are included in the assimilation. For comparison (Fig. 13), we consider DRE estimates derived from the MACC project (Bellouin et al. 2013). DRE estimates based on observations and calculated based on the results of four global aerosol models are also considered (Yu et al. 2006).

Much of the uncertainty in the DRE reported by the Intergovernmental Panel on Climate Change (IPCC) arises from differences between estimates from global models and satellite-based estimates (Myhre 2009). Figure 13 shows the time series of clear-sky shortwave DRE from MERRA-2 and MERRAero separately over land and ocean. We also indicate DRE estimates derived from the MACC reanalysis (Bellouin et al. 2013, yellow shading) and Yu et al. (2006; both model and observational estimates, red and gray shading, respectively). Top-of-the-atmosphere (TOA), surface (SFC), and atmospheric (ATM) DRE is shown, where TOA = SFC + ATM. Table S3 presents the results of Fig. 13 in tabular form and also includes a comparison of
AOD, AAOD, and DRE efficiency (DRE/AOD) separately over land and ocean. Differences in the DRE efficiency highlight where model aerosol assumptions (e.g., absorption, size distribution) and environmental properties (e.g., surface albedo) contribute to discrepancies between DRE estimates from reanalyses. The DRE efficiency for MERRA-2 is within about 20% of the observational estimate (Yu et al. 2006).

As the global aerosol observing system continues to grow and provide additional information on aerosol absorption, size, type, and vertical distribution that can be assimilated, reanalysis estimates of aerosol radiative-climate effects should become more consistent with satellite-derived effects and thus reduce DRE uncertainty. For example, the GMAO is working to incorporate aerosol vertical distribution information from the Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) into a future aerosol reanalysis. Unlike satellite estimates alone, reanalyses like MERRA-2 can provide detailed information on how the anthropogenic component of aerosols, and thus radiative forcing, has changed during the modern satellite era. This should lead to reduced uncertainty in assessing the human impact on climate.

### 5. Summary and conclusions

This paper describes the NASA GMAO’s MERRA-2 aerosol reanalysis, the first satellite era (1980 onward) reanalysis in which both meteorological and aerosol observations are jointly assimilated. We use the GOCART aerosol module coupled to the GEOS-5 Data Assimilation System (DAS) to include radiatively active prognostic aerosol mass tracers and the Goddard Aerosol Assimilation System (GAAS) to assimilate bias-corrected AOD from AVHRR and MODIS (Terra and Aqua), MISR AOD over desert regions, and ground-based AERONET AOD. Publicly available gridded output is available at hourly, 3-hourly, and monthly time scales for both column-integrated and three-dimensionally resolved aerosol mass, optical properties, and other aerosol diagnostics. These data are available through the GES DISC at http://disc.sci.gsfc.nasa.gov/mdisc/.

In this paper, we have provided information about the MERRA-2 Data Assimilation System relevant for the aerosol assimilation, including a description of the aerosol module and emissions inventories. We describe both the process of AOD assimilation and the necessary prior data quality control. In MERRA-2 the only property directly constrained by the assimilation is the AOD. Other aerosol properties such as aerosol speciation and vertical distribution, all of which are available, are indirectly constrained by the assimilation. Here, we focus our evaluation and validation efforts on the AOD only. Our companion study (Part II) presents an evaluation and validation of aerosol properties that are indirectly impacted by the MERRA-2 AOD assimilation.
In both studies, to examine the overall impact of the AOD assimilation, we perform an EOS-period control simulation in replay mode where the same version of the model is driven by MERRA-2 analyzed meteorology, but without AOD assimilation.

We begin by assessing innovation statistics as a sanity check, as we expect analyzed AOD fields to better match assimilated AOD than the forecast AOD. As shown by Lynch et al. (2016), a well-performing forward model is as equally important as the AOD assimilation process itself. During the EOS period, a period for which the GEOS-5/GOCART system has been extensively evaluated and refined, we show that in most regions there is little difference between the 3-h forecast and analyzed AOD fields. Thus, on average, the assimilation process needs to apply relatively small AOD analysis increments \( \Delta a \) to agree with the assimilated observations. This is an indication that the model is relatively unbiased with respect to the AOD measurements it assimilates.

The forward modeling system in MERRA-2, however, has not been optimized to deal with the larger, stratospheric aerosol produced from SO\(_2\) oxidation after major volcanic eruptions. We are currently working on improving the representation of stratospheric SO\(_2\) injection by past volcanoes, in terms of plume height and injection magnitude by incorporating data from Carn et al. (2016). This dataset combines UV measurements from TOMS, OMI, and OMPS with infrared data from TOVS, AIRS, and IASI to produce a database of volcanic SO\(_2\) injection amount and plume top height from 1979 to present. Future inclusion of a separate volcanic aerosol tracer or aerosol microphysics will improve the representation of stratospheric aerosol size. Improvements in the forecasted AOD after major eruptions will mitigate the need for large, negative AOD increments, and we expect less of an artificial impact on the time series of AOD (and DRE) for individual aerosol species.

Next we focus on evaluating and validating the MERRA-2 analyzed AOD fields. First, we compare hourly AOD to AERONET observations, which were not assimilated until 1999. We find that at all stations, the assimilated AOD better matches the observations than the control simulation, not only in terms of correlation but also in terms of more realistic variability and
reduced RMSE. The assimilation produces the least impact for biomass burning regions, where emissions during the EOS period have been well tuned to MODIS data products. Comparisons to AERONET also reveal that both changes in the aerosol observing system between the pre-EOS and EOS periods and changes in prescribed aerosol emissions inventories (especially for biomass burning) impact MERRA-2 AOD.

Shipborne sun photometer data are used for independent validation of MERRA-2 AOD both for the pre-EOS and EOS periods. Compared to historic AOD cruise measurements, we find that MERRA-2 correlates reasonably well with the observations reported as morning or afternoon averages. The bias between analyzed and observed AOD is generally within the instrumental uncertainty. In part because they are recorded with time precision that allows for unambiguous sampling of the model, present-day Maritime Aerosol Network observations show even greater agreement with analyzed AOD fields. Furthermore, there is a clear, positive impact of the assimilation on the AOD as compared to the control simulation.

We conclude our model AOD validation using aircraft observations taken during the recent NASA SEAC4RS campaign over the southeastern and western United States in August–September 2013. First, we compare to AOD retrievals from the DIAL/HSRL instrument. Considering that the model is spatially coarse compared to the high-resolution observations, both the MERRA-2 AOD and AOD from the control simulation simulate observed AOD well, with improvements seen for the analyzed AOD fields. Similarly, comparisons between these two simulations and the 4STAR instrument, which measured AOD above the aircraft during SEAC4RS, reveal that the AOD assimilation tends to increase the lowest-simulated and midrange AOD while decreasing the highest-simulated AOD during the campaign. Overall, the resulting campaign-wide AOD distribution from MERRA-2 better matches the observed 4STAR distribution. Notably, this improved agreement occurs despite the fact that the above-aircraft AOD is a function of the aerosol vertical distribution, an aerosol property only indirectly impacted by the assimilation (Part II).

Finally, we compare MERRA-2 AOD, AAOD, and the clear-sky shortwave direct radiative effect (DRE) among recent aerosol reanalyses, models, and observations. MERRA-2 global average AOD is higher than the model-simulated AOD, but AAOD from MERRA-2 is lower than observation studies suggest. Estimates of clear-sky DRE from MERRA-2 are closer to observationally based estimates than models, but examination of the DRE efficiency reveals a potential need to revise aerosol optical property assumptions (e.g., absorption and size distribution). We caution that in all-sky (cloudy) conditions, the DRE of aerosols from MERRA-2 will be highly sensitive to the vertical distribution of aerosol absorption (Chylek and Coakley 1974). While the overall MERRA-2 vertical profile of aerosol extinction is reasonable compared to observations (Buchard et al. 2016), black carbon aerosol and its associated absorption are overestimated at higher altitudes, particularly away from source regions [e.g., compared to High-Performance Instrumented Airborne Platform for Environmental Research Pole-to-Pole Observations (HIPPO) data from Schwarz et al. (2013) as shown in Randles et al. (2016)]. Furthermore, anthropogenic aerosol preindustrial to present-day radiative forcing (RF), which we do not consider in this study, depends both on aerosol distributions and aerosol specification, the latter of which is strongly controlled by lower boundary conditions (i.e., emissions inventories) rather than the assimilation of AOD. Both the AOD and DRE are impacted by a changing observing system and, in some cases, changing emissions inventories. Nevertheless, long-term aerosol reanalyses like MERRA-2 have the potential to reduce uncertainty in estimates of aerosol-radiation effects, particularly as they are further constrained by additional aerosol observations (e.g., vertical distribution, absorption, multiwavelength information to distinguish aerosol types).

As demonstrated in this paper, in Randles et al. (2016), and in Part II, in many cases MERRA-2 shows considerable skill in simulating numerous observable aerosol properties. However, we emphasize that only the 550-nm AOD is constrained in the reanalysis, and it is only constrained when and where data are available. Prior to the EOS period, observations were primarily from AVHRR over ocean only. While data volume, especially over land, increased markedly after 2000, AOD observations are only available for the sunlit portion of the globe, depend strongly on satellite viewing geometry, and are subject to meteorological conditions (e.g., cloudiness). Without available data to assimilate, the assimilation naturally draws toward the forecast from the GEOS-5/GOCART model. Care must always be taken when considering aerosol products indirectly constrained by the assimilation, and trends in reanalysis aerosol properties including AOD must be considered within the context of a changing observing system and a forecast influenced by prescribed aerosol emissions inventories.

The MERRA-2 aerosol reanalysis is a major step toward an Integrated Earth Systems Analysis that will one day incorporate atmospheric, constituent, oceanic, and land observations to provide a scientific, internally
consistent gridded description of the state of the Earth system and how it is evolving over the satellite era. Future work on the aerosol analysis will focus on improving the aerosol forecast in order to minimize the needed analysis increments. We are also actively working on incorporating observations from additional space- and ground-based sensors from an ever-expanding aerosol observation network, particularly observations that provide additional information content (e.g., multispectral measurements and vertical structure from space-based lidar).

Acknowledgments. MERRA-2 is an official product of the Global Modeling and Assimilation Office at NASA GSFC, supported by NASA’s Modeling, Analysis, and Prediction (MAP) program. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at the Goddard Space Flight Center. The authors would like to acknowledge our colleagues at the GMAO who produced MERRA-2. We thank the AERONET and MAN PIs for their effort in establishing and maintaining these essential ground- and ship-based networks. We also thank all SEAC4RS and DISCOVER-AQ PIs, participants, and funding organizations that made these campaigns possible. Finally, we thank the three reviewers who provided useful feedback that helped to improve this manuscript.

REFERENCES


