RESEARCH ARTICLE
10.1029/2018JD029693

Key Points:
• The GEOS-Chem model does not capture the observed sulfate response to hydroclimate variations during summertime.
• The model deficiency is traced to too large a decrease in clouds and hence aqueous phase sulfate production from wet to dry conditions.
• Simulations with updated cloud fields better reproduce both sulfate concentration and wet deposition flux across different levels of dryness.

Supporting Information:
• Supporting Information SI

Correspondence to:
Y. Xie and Y. Wang, ywang246@central.uh.edu; xieyy12@mails.tsinghua.edu.cn

Abstract
Understanding the response of sulfate to climate change is crucial given tight couplings between sulfate and the hydrological cycle. As the sources and sinks of sulfate are sensitive to cloud and precipitation processes, the accuracy of model simulations depends on the accuracy of these meteorological inputs. In this study, we evaluate the GEOS-Chem model in simulating summertime surface sulfate concentrations in the continental United States across different levels of dryness and compare the model performance based on two sets of meteorological fields: Modern Era Retrospective Analysis for Research and Applications (MERRA) and MERRA-2. Both simulations fail to reproduce observed increases in sulfate during drought, as indicated by negative correlation slopes between surface sulfate concentrations and the standardized precipitation evapotranspiration index (SPEI). This deficiency can be largely attributed to too large a decrease in clouds and hence aqueous phase sulfate production as conditions shift from wet to dry. MERRA-2-driven GEOS-Chem (M2GC) shows improvements in cloud and precipitation fields relative to the MERRA-driven GEOS-Chem, hence eliminating approximately half of the bias in the simulated sulfate-SPEI slope. However, M2GC still underestimates boundary layer cloud fraction, overestimates liquid water content, and overestimates the rates of the decrease in both quantities as conditions become drier. Explicitly correcting these cloud biases in M2GC results in a 60–80% reduction of the bias in the simulated sulfate-SPEI slope. The strong sensitivity of simulated sulfate to prescribed cloud fields suggests the need for more comprehensive assessment of cloud inputs for sulfate simulations under current and future climate change scenarios.

1. Introduction
Sulfate is the dominant inorganic species in aerosols and accounts for a large portion of the total aerosol mass. It determines the hygroscopicity of aerosols, thereby affecting microphysical and optical properties as well as their climate impacts. Sulfate is also tightly coupled to the hydrological cycle (Ervens, 2015; Ramanathan et al., 2001). Clouds and precipitation affect the source, sink, and spatial distribution of sulfate. More than half of the atmospheric sulfate burden forms within clouds (Barth et al., 2000; Seinfeld et al., 2018). Variations in clouds and precipitation affect gas phase sulfate production by modulating radiation and temperature (Koch et al., 2003; Mueller et al., 2006), while cloud condensate may also act as a reservoir of sulfate, releasing it upon evaporation (Berg et al., 2015; Pruppacher & Jaenicke, 1995). Sulfate is most effectively removed by wet scavenging (Barth et al., 2000; Rasch et al., 2000). The atmospheric residence time of sulfate thus varies greatly between dry and wet conditions.

Sulfate scatters solar radiation, acts as cloud condensation nuclei, and affects cloud development and subsequent precipitation processes (Albrecht, 1989; Feingold & Kreidenweis, 2000; Rosenfeld et al., 2008; Twomey, 1977; H. Yu et al., 2005). Changes in sulfate concentrations thus have significant impacts on climate and the hydrological cycle at a wide range of scales (Charlson et al., 1992; Haywood & Boucher, 2000; Lohmann & Feichter, 2005). The hydrological cycle over land regions is projected to intensify as climate warms, with heavier precipitation events distributed among longer dry spells (Dai et al., 2017; Giorgi et al., 2018; Trenberth et al., 2014). The response and feedbacks of sulfate to changing
hydroclimate remain highly uncertain owing to the complexity and potential compensations among these processes (Ervens, 2015; Koch et al., 2003).

Given the strong links between sulfate and the hydrological cycle, the credibility of sulfate simulations by atmospheric models depends intrinsically on the accuracy of the meteorological inputs. Previous studies indicate that climate models tend to underestimate the fraction of low clouds (Kay et al., 2012; M. H. Zhang, 2005). This cloud bias would affect simulations of both the mean and variability of sulfate concentrations. Mueller et al. (2006) found that model underestimation of sulfate is greatly reduced after correcting underestimates in cloud cover. In a prior analysis, we revealed the deficiency of coupled chemistry-climate models in simulating sulfate variability between drought and wet conditions and related this problem with model overestimation of the rate of decrease in clouds from wet to dry conditions (Y. Wang et al., 2017). Shen et al. (2017) also suggested that biases in cloud fields in a global chemical transport model were responsible for the model deficiency in the relationship between sulfate and temperature. In addition to this pervasive model bias in cloud fraction, Komurcu et al. (2014) found that total column liquid water significantly exceeded observations in four out of six climate model simulations. J. Zhang et al. (2007) likewise reported a 24–99% overestimate of cloud liquid water content (LWC) in the Canadian Global Environmental Multiscale model (GEM) as compared to an aircraft campaign conducted in 2004. Such a bias poses a potential problem for air quality simulations based on these models. Biases in model precipitation and associated wet deposition processes can also affect sulfate simulations (Gong et al., 2006; Luo et al., 2011; von Salzen et al., 2000). A better understanding of the sensitivity of sulfate simulations to hydroclimatic variables such as clouds and precipitation is needed to improve our ability to predict the response of sulfate to climate change.

In this study, we evaluate the performance of a global chemical transport model driven by assimilated meteorological fields, GEOS-Chem, where GEOS stands for Goddard Earth Observing System, in simulating summertime sulfate. We focus on the reliability of simulations across different levels of dryness ranging from extremely wet conditions to extremely dry conditions and aim to characterize and attribute model biases in the response of sulfate to hydroclimatic variations. By analyzing GEOS-Chem sulfate simulations driven by two different sets of reanalysis fields, we quantify the extent to which the model sulfate bias may originate in the prescribed meteorological fields. Through sensitivity simulations in which biases in the prescribed cloud fields are corrected, we further assess whether the model can simulate sulfate changes under different levels of dryness. We focus on summertime (June to August) in the continental United States, where observations are relatively abundant, sulfate concentrations are high, and surface sulfate is thought to be more influenced by cloud processing due to stronger mixing within the boundary layer during summer (Rasch et al., 2000; Walker et al., 2012). Drought frequency is also expected to increase during summer (Cook et al., 2018; Swain & Hayhoe, 2015; G. Wang, 2005; Wetherald & Manabe, 2002).

2. Methods

2.1. GEOS-Chem Model

The GEOS-Chem model (version 11-1) is driven using assimilated meteorological fields provided by NASA's Global Modeling and Assimilation Office. The model includes fully coupled O3-NOx-volatile organic compound-aerosol chemistry (Bey et al., 2001; Park et al., 2004). Gas-aerosol partitioning is calculated using the thermodynamic module ISORROPIA II (Fountoukis & Nenes, 2007). In the model, sulfate is produced through gas phase oxidation of SO2 by OH and aqueous phase oxidation of SO2 by H2O2 and O3 in clouds (Alexander et al., 2012; Park et al., 2004). Heterogeneous production of sulfate on the surface of deliquescent aerosols was suggested to be important during winter in polluted urban environments in China (Y. X. Wang et al., 2014), but this mechanism is not considered in this study. Anthropogenic emissions are from the Environmental Protection Agency (EPA) National Emission Inventory for year 2011, with annual scaling factors applied to the study period (2000 to 2014) based on estimated trends in air pollutant emissions (U.S. EPA, 2016). Biomass burning emissions are based on the Global Fire Emission Database (Akagi et al., 2011; Giglio et al., 2013; Randerson et al., 2012; van der Werf et al., 2010). Aerosol wet scavenging due to large-scale and convective precipitation is parameterized following the scheme described by Liu et al. (2001) with modifications by Q. Wang, et al. (2014). Sulfate is fully scavenged in warm (T > 258 K) and ice (T < 237 K) clouds but retained in rimed ice crystals in large-scale mixed-phase clouds (half retained...
in convective mixed-phase clouds). SO$_2$ scavenging is treated like sulfate in the model with the soluble fraction constrained by the concentration of H$_2$O$_2$ (Chin et al., 1996; Mari et al., 2000).

To diagnose the potential impact of prescribed meteorological fields on the model sulfate bias, two sets of long-term simulations are performed in GEOS-Chem, one driven by the Modern Era Retrospective Analysis for Research and Applications (MERRA) and the other by the more recent Modern Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Both reanalysis data sets are produced using the Goddard Earth Observing System atmospheric model and data assimilation system. MERRA was specifically designed to improve representations of precipitation and the atmospheric water cycle for climate research (Rienecker et al., 2011) and is thus relevant for simulating sulfate burdens under a variety of hydroclimate conditions. The newer MERRA-2 contains significant updates to the forecast model, analysis algorithm, and assimilated data relative to MERRA (Gelaro et al., 2017). For example, the stratiform cloud parameterization in MERRA-2 uses new observationally based constraints on total water (Molod, 2012), while the convection scheme has been modified to eliminate clouds that detrain within the atmospheric boundary layer (Molod et al., 2015). MERRA-2 also uses observations to correct the model-generated precipitation field over tropical and midlatitude land areas (60°S–60°N), which improves the simulation of land surface fluxes and subsequent feedback to the atmospheric water cycle (Reichle et al., 2017). Both simulations are performed for the summer months (June–July–August) of 2000 to 2014 at a horizontal resolution of 2° × 2.5° and on 47 vertical levels. The model spin-up time is 3 months (March to May) for each year. In the following, we refer to the GEOS-Chem simulation driven by MERRA as MGC and to the GEOS-Chem simulation driven by MERRA-2 as M2GC. Differences between MGC and M2GC reflect the impacts of different prescribed meteorological fields on sulfate simulation.

2.2. Observations

Surface SO$_2$ measurements are obtained from the U.S. EPA Air Quality System (https://www.epa.gov/outdoor-air-quality-data). Surface sulfate concentrations are collected from the EPA Speciation Trend Networks and Interagency Monitoring of Protected Visual Environments (https://views.cira.colostate.edu/fed/QueryWizard/Default.aspx) monitoring networks (Malin et al., 1994). Wet deposition data are from the National Atmospheric Deposition Program (https://nadp.slh.wisc.edu/). Measured sulfur deposition includes both wet-scavenged sulfate and sulfate produced via the oxidation of dissolved SO$_2$ in rainwater (Appel et al., 2011; Y. Zhang et al., 2018). To match these data, the model sulfur deposition flux for comparison with National Atmospheric Deposition Program flux measurements is calculated as the sum of the sulfate wet deposition flux and 1.5 times the SO$_2$ wet deposition flux, where 1.5 is the ratio of molecular mass between sulfate and SO$_2$. To ensure sufficient coverage of both dry and wet conditions, only sites with more than 10 years of sulfate observations are selected. The screening process yields a total of 652 sites for observations of surface SO$_2$, 408 sites for sulfate, and 220 sites for sulfur wet deposition.

We obtain satellite-observed cloud fraction and liquid water path in the lower troposphere (1,000–680 hPa) during June–July–August 2000 to 2014 from the Clouds and the Earth’s Radiant Energy System ISCCP-D2like data set (https://ceres-tool.larc.nasa.gov/ord-tool/jsp/CldTypHistSelection.jsp). This data set combines observations from Moderate Resolution Imaging Spectroradiometer and geostationary (GEO) satellite instruments to produce diurnally consistent cloud properties at a 1° × 1° spatial resolution (Minnis, Sun-Mack, Chen, et al., 2011; Minnis, Sun-Mack, Young, et al., 2011). Additional information on cloud vertical profiles is obtained from the combined CloudSat and CALIPSO spaceborne lidar cloud fraction data set during the period for which this data set is available (2006 to 2011, https://climatedataguide.ucar.edu/climatedata/combined-cloudsat-spaceborne-radar-and-calipso-spaceborne-lidar-cloud-fraction-dataset; Kay & Gettelman, 2009). These data are provided at a spatial resolution of 2° × 2° and a vertical resolution of 480 m. We use data from the Global Precipitation Climatology Project at 2.5° × 2.5° resolution to evaluate the model precipitation (Adler et al., 2003).

Local dryness is represented by the 1-month standardized precipitation evapotranspiration index (SPEI v2.5, http://spei.csic.es/) at a 0.5° × 0.5° resolution (Vicente-Serrano, Begueria, & López-Moreno, 2010). The SPEI is calculated based on standardized anomalies in precipitation and potential evapotranspiration estimated based on the Climate Research Unit Timeseries version 3.23 gridded analysis of precipitation and other meteorological variables. By including both precipitation and potential evapotranspiration, this index
considers the influences of both precipitation and temperature changes on the local water balance (Vicente-Serrano, Beguería, López-Moreno, Angulo, & El Kenawy, 2010). Negative SPEI indicates dry conditions relative to the local norm, while positive SPEI represents wet conditions. The index is normalized to have a standard deviation equal to 1.

Site-level observations and SPEI are regridded onto 2° × 2.5° grids to match the model resolution. To test whether surface dryness in the model is consistent with that indicated by the standard SPEI data set, we calculate SPEI based on model surface temperature and total precipitation inputs using the R package (https://cran.r-project.org/web/packages/SPEI/index.html). Although the normalization period differs, the model SPEI is well correlated with the standard SPEI data set over most of the continental United States (r > 0.7 on average; supporting information Figure S1). SPEI derived from MERRA-2 yields a higher correlation coefficient than that derived from MERRA, especially over the northeastern United States. This suggests that updates in MERRA-2 relative to MERRA improve the representation of local dryness, as expected since the precipitation flux prescribed to the land surface model is replaced by observations in MERRA-2 (Reichle et al., 2017). Overall, these comparisons suggest that the observationally based SPEI data set is an appropriate metric for distinguishing different levels of surface dryness in the model simulations.

2.3. Model Evaluation

Figure 1 compares the observed and simulated (MGC and M2GC) mean summertime SO2 and sulfate concentrations over the continental United States during the 2000–2014 study period. Both MGC and M2GC reproduce well the spatial distribution of SO2 (r = 0.64) and sulfate (r > 0.95). The mean of the observations is 2.0 (±1.2) ppbv for SO2 and 2.3 (±1.7) μg/m³ for sulfate (Figure 1). The normalized mean bias of the simulated SO2 is −12% in MGC and +3% in M2GC, indicating that the model represents SO2 emissions well. However, the normalized mean bias for sulfate is +40% in MGC. Limiting the comparison to rural sites (Interagency Monitoring of Protected Visual Environments) has little impact on this high bias, ruling out emissions as the main driver of the sulfate bias. M2GC has a smaller normalized mean bias in sulfate...
(+29%) compared to MGC. Our simulation results are consistent with previous work based on GEOS-Chem and other regional models (e.g., CMAQ), which consistently produce a 10–30% high bias in surface sulfate across the United States (McKeen et al., 2007; S. Yu et al., 2008; Walker et al., 2012). We discuss the sulfate high bias and its causes in detail across a range of dryness levels in section 3.

Despite the mean bias, both MGC and M2GC capture the long-term decreasing trend of SO2 and sulfate (Figures 2a and 2b), further validating the model emissions. Correlations between the observed and the simulated sulfate, after removing the long-term trends, are shown in Figures 2c and 2d for every model grid cell where observations are available. MGC sulfate shows a significant but moderate correlation with observations over the eastern United States (average $r = 0.46$). Over the western United States, the correlation is typically weak and not statistically significant. M2GC shows a higher correlation with observations over the eastern United States (average $r = 0.63$). Therefore, compared to MGC, M2GC not only reduces the mean sulfate high bias but also better reproduces temporal variations in the summertime sulfate burden. Since both simulations perform better over the eastern United States where sulfate concentrations are on average higher, we focus on this region in the following analysis. The eastern United States is further divided into southeastern and northeastern sectors as shown in Figure 2d.

### 3. Sulfate Simulation at Different Dryness Conditions

Figure 3 compares the observed and simulated sulfate composited at different levels of dryness for the southeastern and northeastern United States separately. For each grid cell, we divide the SPEI into five bins covering the local range of SPEI, defined so that each bin has approximately the same number of samples. We then average the SPEI values and sulfate concentrations within each bin. Long-term trends in sulfate were removed from both the observations and the model simulations to reduce the impacts of changes in anthropogenic emissions. The mean values of SPEI and sulfate concentrations within each SPEI bin are then averaged over all grid cells in the southeastern United States and northeastern United States, respectively, to derive the regional mean values shown in Figure 3. Observed summertime mean sulfate concentrations during the study period were 3.7 $\mu$g/m³ over the southeastern United States and 3.3 $\mu$g/m³ over the northeastern United States.
United States. Sulfate concentrations were negatively correlated with SPEI over both the southeastern United States \((p < 0.05)\) and the northeastern United States \((p = 0.15)\), estimated via linear regression of mean sulfate across the five SPEI bins. A negative correlation indicates that sulfate concentrations tend to increase with increasing dryness, as negative SPEI indicates drought. The sulfate concentration under extremely dry condition \((SPEI < -1)\) is 4–12% higher than that under extremely wet conditions \((SPEI > 1)\). The regression slope (in units of micrograms per cubic meter per unit change of SPEI) is more negative over the southeastern United States \((-0.18)\) than over the northeastern United States \((-0.08)\). This suggests a higher sensitivity of sulfate to hydroclimate variations over the southeastern United States.

As shown in Figure 3, neither MGC nor M2GC captures the observed negative correlations between sulfate and SPEI. Our previous work (Y. Wang et al., 2017) identified a similar problem in coupled chemistry-climate models. This suggests that properly representing the sulfate response to hydroclimate variations is a nontrivial problem for chemical transport models and climate models alike. Consulting Figure 3, we can infer that the model’s inability to represent the sulfate-SPEI relationship arises mainly from a larger high bias of sulfate under wet conditions. For example, MGC has a mean sulfate bias of 38% during extremely wet months \((SPEI > 1.3)\) over the southeastern United States, while the bias is only 10% during extremely dry months \((SPEI < -1.3)\). Comparing the two simulations, the high bias in sulfate is smaller in M2GC \((12–40\%)\) than MGC \((23–55\%)\), and the sulfate-SPEI slope in M2GC is likewise much less positive \((0.03–0.11 \mu g/m^3 per unit change of SPEI)\), a notable improvement relative to MGC \((0.21–0.24 \mu g/m^3 per unit change of SPEI)\).

4. Attributing the Improvement From MGC to M2GC

As discussed in the previous section, simulated sulfate concentrations in M2GC are on average 10% smaller than those in MGC. M2GC thus has a 50% lower sulfate high bias compared to observations, which leads to a better simulation of the sulfate-SPEI relationship. In this section, we provide a detailed comparison between the two simulations to better understand the causes behind the improvements in M2GC relative to MGC. As M2GC still overestimates both mean sulfate concentrations and the sulfate-SPEI slope relative to observations, we follow this discussion by diagnosing some remaining problems in M2GC in section 5.

The sulfate high bias can be explained in part by the model simulations overestimating surface \(SO_2\) concentrations, particularly over the northeastern United States (Figure 1). In addition, the models cannot reproduce observed changes in \(SO_2\) across different levels of surface dryness (Figure S2). The observations show negative correlations between surface \(SO_2\) and SPEI over the southeastern United States (slope = -0.11,
and the northeastern United States (slope = −0.12, p < 0.01), which partially explains the observed negative relationship between sulfate and SPEI. In comparison, the models predict a positive or flat change of SO2 as SPEI changes from wet to dry. SO2 emissions in the model are dominated by anthropogenic sources which vary by less than 2% across different levels of dryness, and the models adequately capture the significant increase in SO2 emissions due to enhanced biomass burning during dry conditions (Figure S3).

Potential anthropogenic impacts on SO2 emissions and sulfate production across different levels of dryness are not currently represented in the model. However, differences in SO2 between the two simulations are unlikely to be a major cause of the improvement in M2GC.

4.1. Wet Deposition and Precipitation

Since wet scavenging is one of the most important drivers of sulfate variations under wet conditions, we continue by comparing the model wet deposition flux with observations (Figures 4a and 4b). Like our calculation of the sulfate-SPEI relationship described above, we calculate the mean wet deposition flux for five equally sized SPEI bins in each grid cell and then average these gridded values to construct the regional mean deposition flux in each SPEI bin. The observed mean sulfate wet deposition is 1.0–1.1 kg·ha⁻¹·month⁻¹ across the eastern United States and shows a significant positive correlation with SPEI. Wet removal is

![Figure 4](image_url)

Figure 4. Comparisons of total sulfur wet deposition (a, b) and precipitation rate (c, d) from observations and models under different levels of dryness over the southeastern and northeastern United States. Error bars indicate the standard error of the mean in each bin. Slopes in (a)-(d) indicate linear trends based on each data set, with p the confidence level for the corresponding slope. MGC = MERRA-driven GEOS-Chem; M2GC = MERRA-2-driven GEOS-Chem; SPEI = standardized precipitation evapotranspiration index.
41–47% less during dry months than during wet months, a consequence of a ~60% reduction in precipitation intensity (Figures 4c and 4d). This reduction in scavenging by precipitation partially explains the observed enhancement of sulfate and SO2 loading during dry conditions.

MGC overestimates mean wet deposition fluxes of sulfate by 37–55%, with a mean value of 1.59 (1.49) kg·ha\(^{-1}\)·month\(^{-1}\) over the southeastern (northeastern) United States. MGC predicts a 30% decrease in wet deposition from wet to dry conditions, which underestimates the observed change. As a result, MGC underestimates wet deposition-SPEI slopes by 22% over the southeast United States and by 37% over the northeast United States relative to observations. By comparison, the M2GC-simulated average wet deposition flux is within 10% and the wet deposition-SPEI slope is within 5% of the observed values over both regions. Mean sulfate wet deposition is about 50% lower during dry months than during wet months, close to the observed enhancement of sulfate and SO2 loading during dry conditions.

Figure 5. Comparisons between simulated gas and aqueous phase sulfate production in MGC and M2GC over the southeastern United States (a) and the northeastern United States (b) at the surface level. Slopes in (a) and (b) indicate linear trends based on each data set, with \(p\) the confidence level for the corresponding slope. MGC = MERRA-driven GEOS-Chem; M2GC = MERRA-2-driven GEOS-Chem; SPEI = standardized precipitation evapotranspiration index. 10.1029/2018JD029693

MGC overestimates mean wet deposition fluxes of sulfate by 37–55%, with a mean value of 1.59 (1.49) kg·ha\(^{-1}\)·month\(^{-1}\) over the southeastern (northeastern) United States. MGC predicts a 30% decrease in wet deposition from wet to dry conditions, which underestimates the observed change. As a result, MGC underestimates wet deposition-SPEI slopes by 22% over the southeast United States and by 37% over the northeast United States relative to observations. By comparison, the M2GC-simulated average wet deposition flux is within 10% and the wet deposition-SPEI slope is within 5% of the observed values over both regions. Mean sulfate wet deposition is about 50% lower during dry months than during wet months, close to the observed change. Therefore, the improved simulation of wet deposition in M2GC contributes to the better performance of this model compared to MGC in simulating both sulfate and its variations with SPEI. As M2GC has effectively achieved an unbiased simulation of sulfate wet deposition fluxes, this rules out wet deposition as the process most likely responsible for the remaining M2GC biases in both sulfate and sulfate-SPEI relationships noted above. This finding is consistent with our previous work that shows the chemistry-climate models can well reproduce the magnitude of decrease in sulfate wet deposition during dry conditions, suggesting processes other than wet deposition play a larger role in causing model deficiencies with respect to the sulfate-SPEI relationship (Y. Wang et al., 2017).

The improvement in precipitation fields in MERRA-2 can explain the better simulation of sulfur wet deposition fluxes in M2GC as compared to MGC. Figures 4c and 4d show that MERRA-2 reproduces the precipitation-SPEI slope to within 6% (1%) of the observed value over the southeastern (northeastern) United States, suggesting that the updates in MERRA-2 ensure a better representation of precipitation across different levels of dryness. By comparison, the precipitation-SPEI slope in MERRA is 35–48% lower than that indicated by observations. This inability to represent precipitation variability is likely the primary reason that MGC underestimates the wet deposition-SPEI slope. Nevertheless, precipitation in MERRA-2 still has low biases of 20% and 6% over the southeastern and northeastern United States, respectively. This suggests that the overestimated wet deposition flux in M2GC is not caused by biases in model precipitation input but stems instead from the high bias in sulfate concentrations.

4.2. Sulfate Production

Figure 5 provides a comparison of surface gas and aqueous phase sulfate production as simulated by MGC and M2GC. On average, relative to MGC, the total production simulated by M2GC is 67% smaller in the
southeastern United States and 33% smaller in the northeastern United States. These differences are mainly caused by significantly smaller aqueous phase production, as gas phase production is similar between the two simulations. The 59–87% reduction in aqueous phase sulfate production in M2GC relative to MGC accounts for the ~10% lower sulfate concentrations at the surface level as well as the 17–33% smaller column total wet deposition fluxes and produces a better match to observations. The implication is that aqueous phase production at the surface level is substantially overestimated in MGC. The relationship between total production and SPEI is dominated by the relationship between aqueous phase production and SPEI, with both relationships characterized by positive correlations in the model. Gas phase production of sulfate presents little variations across different levels of dryness in both simulations. This may be explained by the lack of variability in SO2 concentrations across different levels of dryness (Figure S2) as well as relatively small changes in low cloud fractions (discussed further in section 5). The significant increase in aqueous phase production with increasing SPEI is also the main driver of the positive relationship between SPEI and sulfate simulated by MGC. The total production-SPEI slope is 16–67% smaller in M2GC than in MGC, primarily owing to weaker increases in aqueous phase production as SPEI increases. This results in a shallower sulfate-SPEI slope that is more consistent with observations. Y. Wang et al. (2017) also reported a significant decrease in aqueous phase sulfate production from normal to dry conditions, which drives an overall decrease in sulfate concentrations, opposite to the relationship inferred from observations.

The spatial patterns of differences in surface sulfate concentrations and total wet deposition fluxes between the two simulations match that of differences in surface aqueous phase production (Figure 6). Aqueous phase production of sulfate is reduced over the entire eastern United States in M2GC relative to MGC, with especially pronounced differences over the southeastern United States. Consequently, simulated surface sulfate concentrations and total wet deposition fluxes are also much smaller over the eastern United States in M2GC than in MGC. This further emphasizes the major role played by more realistic aqueous phase production in underpinning the improvements in M2GC relative to MGC. Comparisons of vertical profiles within the boundary layer confirm that the largest differences in sulfate concentrations between the two simulations can be attributed to differences in aqueous phase production, especially close to the surface (Figure S4). Aqueous phase production within the boundary layer shows a smaller difference (~10%) between the two simulations than does aqueous phase production at the model levels closest to the surface. By contrast, despite little difference close to the surface, average gas phase production within the boundary layer is 10–20% higher in M2GC than in MGC. The moderate decrease in aqueous phase production within the boundary layer, reduced wet deposition, and enhanced gas phase production all play a role in explaining why a difference of more than 60% in surface aqueous phase sulfate production is associated with only a 10% decrease in sulfate concentrations at the same level. Berg et al. (2015) also reported a substantial compensation between enhanced aqueous phase production and enhanced wet deposition over the southeastern United States, leading to only a small net decrease in sulfate concentrations.

### 4.3. Model Cloud Comparison

What causes the smaller aqueous phase production of sulfate in M2GC relative to MGC? Since the two simulations are conducted with the same chemical transport model, the changes in aqueous phase sulfate production should result from differences in the meteorological fields. In the GEOS-Chem model, aqueous phase production depends mainly on the in-cloud LWC (g/m³) and cloud fraction (CF) inputs. LWC describes the volume density of cloud water, while CF denotes the percentage of a grid cell containing clouds in which aqueous phase production takes place. All other factors being equal, a larger LWC accelerates SO2 oxidation in clouds by enhancing oxidant uptake and reducing droplet acidity, while a larger CF increases the fraction of SO2 in the grid cell oxidized through the aqueous pathway. In the model, the in-cloud LWC is calculated by dividing the grid-scale LWC by the corresponding CF.

Figure 7 shows distributions of LWC and CF inputs from MERRA and MERRA-2 at the lowest model level. Relative to MERRA, surface CF of MERRA-2 is 96% lower over the southeastern United States and 76% lower over the northeastern United States. LWC is about 50% smaller in MERRA-2 over the southeastern United States but differs little between the two simulations over the northeastern United States. The spatial pattern of the differences in CF matches well with the spatial pattern of differences in aqueous phase sulfate production and surface sulfate concentrations (Figure 6). Therefore, the reduction in aqueous phase sulfate production in M2GC relative to MGC appears to be caused primarily by the decrease in surface CF in
MERRA-2. Vertical profile comparisons within the boundary layer also show that the difference in CF between the two simulations peaks at the lowest model level (Figure S5), while differences in LWC are largest at approximately 2,000 m above ground level.

To examine whether the much smaller CF at the surface level in MERRA-2 is realistic, we compare CF inputs for the two simulations with CFs derived from combined CloudSat and CALIOP lidar observations (Kay & Gettelman, 2009) at the lowest layer, averaged over the period 2006–2010 for which these observations are available (Figure 7). A direct quantitative comparison is not possible because the observations report surface CF at 240 m (0–480 m) above sea level while the lowest model level is centered approximately 60 m (0–120 m) above ground level. However, as the average topography over the eastern United States is less than 100 m, these observations can provide a rough estimate of the surface CF. The surface CF from

---

**Figure 6.** Differences between M2GC and MGC with respect to surface sulfate concentration (a), aqueous-phase production at the lowest model level (b), and total sulfur wet deposition (c) during summertime from 2000 to 2014. MGC = MERRA-driven GEOS-Chem; M2GC = MERRA-2-driven GEOS-Chem.
the satellite observations is 2.1% over the southeastern United States and 1.3% over the northeastern United States. MERRA significantly overestimates surface CF with average values greater than 5% in both regions. By contrast, MERRA-2 has surface CFs ranging from 0.2% to 1.1%, closer to those based on satellite observations but consistently smaller. This comparison indicates that surface CFs in MERRA are unrealistically large and that MERRA-2 largely corrects this bias.

The improvement of cloud fields in MERRA-2 upon MERRA can be partly traced to differences in how the cloud base is defined in the moist convection parameterization used in the GEOS-5 model, although the core of the convective parameterization (Moorthi & Suarez, 1992) is the same in both versions (Fujiwara et al., 2017). Cloud base for convective plumes is diagnosed in MERRA-2 as the boundary layer height, while that in MERRA is defined as the average of the lowest two model levels (Molod et al., 2015). The lower cloud base in MERRA introduces a larger possibility for convective clouds to form and detrain cloud condensate near the surface. MERRA-2 also uses a larger critical relative humidity threshold than MERRA for non-convective cloud formation within the boundary layer, which helps to reduce stratiform cloud fractions near the surface (Molod et al., 2015). More realistic land-atmosphere enthalpy exchange due to the corrected precipitation flux may also play a role (Reichle et al., 2017). However, it should be noted that both MERRA and MERRA-2 have significant low biases in CF within the lower atmosphere (1,000–3,000 m) relative to the combined CloudSat-CALIPSO product. The largest underestimates are located at around 2,000-m height, where the observations show ~20% CF but both reanalyses produce cloud fractions of 5–10%. Although

Figure 7. Liquid water content from MERRA (a) and the difference between MERRA-2 and MERRA (b); cloud fraction from MERRA (c), the difference between MERRA-2 and MERRA (d), from satellite observation (e) and comparison of the vertical profile of cloud fraction within boundary layer (f). Error bars in panel (f) show standard deviations in CF. MERRA = Modern Era Retrospective Analysis for Research and Applications; LWC = liquid water content; CF = cloud fraction; CALIOP = Cloud-Aerosol Lidar with Orthogonal Polarization.

XIE ET AL.
differences in LWC are also large between MERRA-2 and MERRA (Figure S5), we do not currently have observations to constrain the profile of LWC within the boundary layer. We provide a more detailed assessment of boundary layer clouds in MERRA-2 in section 5.

To summarize, our results demonstrate that a large portion of the surface sulfate high bias in MGC is associated with a positive bias in cloud cover at the lowest level, which in turn causes the model to overestimate aqueous phase sulfate production. A stronger decrease in aqueous phase production with decreasing SPEI then drives the positive relationship between sulfate and SPEI in MGC. MERRA-2 largely corrects the bias in surface CF and consequently provides a better simulation of mean sulfate concentrations at the surface in M2GC. The reduced slope of the relationship between aqueous phase production and SPEI is key to the improved relationship between sulfate concentrations and SPEI in M2GC.

5. Remaining Cloud-Related Sulfate Biases in M2GC

5.1. Biases in Boundary Layer Clouds

Despite the substantial improvement relative to MGC, M2GC still does not fully reproduce the observed relationship between mean sulfate concentrations and SPEI. In section 4 we presented improvements in the mean surface cloud fields in MERRA-2 as compared to MERRA. However, comparison with observations suggests that MERRA-2 still significantly underestimates cloud cover in the boundary layer. Moreover, it remains unclear whether large values of LWC near 2,000 remains unclear whether large values of LWC near 2,000 still significantly underestimates cloud cover in the boundary layer. Moreover, it remains unclear whether large values of LWC near 2,000 are realistic. In this section, we first evaluate boundary cloud properties from MERRA-2 within the 1,000- to 700-hPa layer under different levels of dryness against estimates from the Clouds and the Earth’s Radiant Energy System ISCCP-D2like data set for the 1,000- to 680-hPa layer (Figure 8). We then quantify the potential impacts of cloud biases on model sulfate simulation. To facilitate comparisons, satellite-based estimates of liquid water path are divided by the approximate height of the layer (~3 km) to obtain estimates of LWC.

Relative to satellite retrievals, MERRA-2 CFs are approximately a factor 2 (40–53%) smaller than the observed values, while LWCs are a factor 3–4 larger than observed (Figure 8). MERRA-2 uses a two-moment probability distribution function of total water content to estimate stratiform cloud properties in every grid cell (Bacmeister et al., 2006; Molod, 2012). The parameterization also considers “anvil” cloud condensate detrained by the convection scheme (Moorthi & Suarez, 1992), which may detrain at any level above the diagnosed boundary layer height (Molod et al., 2015). Biases in low cloud cover in MERRA-2 may therefore arise from biases in simulated total water content, biases in the critical relative humidity threshold (used to represent subgrid-scale variations in total water), or biases in the amount of cloud condensate detrained just above the boundary layer by shallow convective plumes (Molod et al., 2015; Naud et al., 2010). Part of the overestimation of LWC shown in Figure 8 is also related to the underestimation of CF, as in-cloud LWC is calculated by dividing the grid-scale LWC by CF. Moreover, this combination is consistent with previous studies based on other models, which have also found significant underestimates of low cloud cover associated with overestimates of LWC (Kay et al., 2012; Komurcu et al., 2014; M. H. Zhang, 2005; J. Zhang et al., 2007).

The CF-SPEI and LWC-SPEI slopes in M2GC are also biased high relative to slopes based on the observations by a factor of 4–9 (Figure 8). These significant positive biases in the slopes suggest that CF and LWC decrease too quickly in MERRA-2 as conditions become drier. This is consistent with the results of Y. Wang et al. (2017), who found a factor of 8 overestimation of the CF-SPEI slope in coupled climate-chemistry models. The large decreases in CF and LWC from dry to wet conditions may in turn result in sharper decreases in aqueous phase production in M2GC, and may partially explain why the model simulates lower sulfate concentrations during dry months than during wet months. Accordingly, the more positive sulfate-SPEI relationship simulated in M2GC relative to that observed may be linked to excessive reductions in CF and LWC as SPEI decreases.

5.2. Cloud Correction Sensitivity Tests

To further evaluate the M2GC simulation, we perform a series of sensitivity tests in which we perturb the model cloud fields. We first select June 2002 to represent dry conditions and June 2003 to represent wet conditions (Figure S6). We then scale the model CF and LWC to better match the observed values. Rather than applying scaling factors specific to the months selected, we apply values derived from long-term mean...
comparisons (Figure 8) to better represent the impacts of the mean cloud bias. The scaling factors, derived as the average ratios between observations and M2GC under different conditions, are 2.5 for dry months and 1.5 for wet months for CF and 0.33 for dry months and 0.25 for wet months for LWC. We only apply these adjustments to the cloud fields in the aqueous phase production module, leaving other aspects of the model alone. Changes in the cloud properties do not affect radiative transfer in the model. As such, gas phase production rates are not affected by changes in available solar radiation. However, cloud-induced changes in aqueous phase production do alter the mixing ratios of SO2 and other oxidants (e.g., OH and H2O2), which could in turn impact the rate of gas phase oxidation. Since cloud biases are similar over the southeastern and northeastern United States, we perturb cloud fields using the same scaling factors over the entire eastern United States. We first discuss the impacts of cloud biases under wet and dry conditions separately and then discuss how these changes in combination might influence the relationship between sulfate and SPEI in the model. The sensitivity experiments are compared against the standard simulation for the corresponding month over the entire eastern United States.

For the dry month (June 2002), increasing CF by a factor of 2.5 leads to a 40.0 kt S (50%) increase in vertically integrated aqueous phase production in the troposphere over the eastern United States, while decreasing

![Figure 8. Comparison of cloud fraction (a, b) and liquid water content (c, d) between MERRA-2 (blue) and satellite observations from the CERES ISCCP-D2like analysis (black) across different levels of dryness for the southeastern United States (a, c) and the northeastern United States (b, d). Slopes in (a)-(d) indicate linear trends based on each data set, with r the confidence level for the corresponding slope. Error bars indicate the standard error of the mean in each bin. MERRA = Modern Era Retrospective Analysis for Research and Applications; LWC = liquid water content; SPEI = standardized precipitation evapotranspiration index.](image-url)
LWC by two thirds reduces aqueous phase production by 14.9 kt S (19%) (Figure S6). When the two perturbations are applied simultaneously, the tropospheric total aqueous phase production increases by 22.8 kt S (29%). The average surface sulfate concentration increases by 5% (0.25 μg/m³), much less than the fractional increase in sulfate production (Table 1). The nonlinearity of the response can be explained by further changes in gas phase production and wet deposition. The reduction in SO₂ due to increased oxidation in clouds leads to a 5.4 kt S (4%) decrease in gas phase production, while the increase in total sulfate production (17.4 kt S) results in a 7.5 kt S (7%) increase in the wet deposition flux. Reduced gas phase production and increased wet deposition together offset about 57% of the increase in aqueous phase sulfate production. In summary, the perturbations to the cloud fields, despite their large relative magnitudes (+150% for CF and −67% for LWC), lead to a change in surface sulfate concentrations of merely 5% during the dry month due to (1) the compensating effects of the CF increase and the LWC decrease and (2) changes in gas phase production and wet deposition that offset the increase in aqueous phase production.

During the wet month (June 2003), increasing CF by a factor of 1.5 leads to 13.7 kt S (14%) increase in aqueous phase production over the eastern United States, while decreasing LWC by 75% reduces aqueous phase production by 26.1 kt S (27%) (Figure S6). In combination, the two perturbations lead to a 9.6 kt S (10%) reduction in aqueous phase sulfate production. Surface sulfate concentrations decrease by 2% (0.08 μg/m³), with the reduced magnitude of this change again due to the compensating effects of changes in gas phase production and wet scavenging. The response of wet deposition to the net change in production is 43% for the dry month and 47% for the wet month. Such a similar response between the dry and wet months is unexpected since area-mean precipitation is 31% smaller during the dry month than during the wet month. Detailed comparisons show that precipitation frequency was larger during the dry month than during the wet month over parts of the eastern United States, even though the precipitation intensity was reduced (Figure S7). The area of enhanced precipitation frequency corresponds well to the area where the simulated wet deposition flux is larger, indicating that precipitation frequency is also an important factor to consider in terms of the response of wet deposition to changes in sulfate production. The simulated surface sulfate concentration response (in units of micrograms per cubic meter per kiloton Sulfate) is 0.025 for the dry month and 0.023 for the wet month. The slightly stronger response under dry conditions may be attributed to stronger boundary layer mixing, which brings a larger portion of the sulfate produced in the cloud layer down to the surface layer.

Overall, cloud correction experiments for both the dry and wet month produce relatively small responses (2–5%) in surface sulfate concentrations, not sufficient to correct the 22–45% sulfate high bias in M2GC. However, the cloud corrections do alter the sulfate–SPEI relationship in the model from positive to negative, owing to the 5% increase in sulfate concentrations for the dry month and the 2% decrease for the wet month. Over the southeastern United States, the corrected slope is −0.11 (cf. the original value of 0.03), more consistent with the slope based on observations (−0.18). We find a similar correction for the northeastern United States, where the original slope of 0.11 reduces to −0.04, as compared to a slope of −0.08 based on observations. These results suggest that biases in model cloud fields account for around 60% of the bias in the simulated relationship between sulfate and SPEI, and supports our previous argument that the model...
sulfate-SPEI bias is partially caused by the model overestimating the rate of decrease in aqueous phase sulfate production as conditions become drier.

The sensitivity experiments described above are based on a period when anthropogenic SO2 emissions in the United States were relatively high (June 2002 and 2003). To test whether our results might be sensitive to changes in emissions, we apply the same cloud corrections for June 2012 (dry month) and June 2013 (wet month), when anthropogenic SO2 emissions were about 50% lower than in 2002 and 2003. The resulting response in surface sulfate concentrations is slightly larger, with a 4% increase for the dry month and a 5% decrease for the wet month (Table 1). The associated sulfate-SPEI slopes are $-0.15$ over the southeastern United States and $-0.08$ over the northeastern United States, within 20% of those based on observations. This suggests that the sensitivity of sulfate to cloud fields may increase with decreasing SO2 emissions, and that the impact of the cloud bias on the sulfate-SPEI relationship may therefore be underestimated in our prior calculation. It also suggests that the impact of cloud biases on simulations of surface sulfate and sulfate-SPEI relationships may increase in the future if SO2 emissions continue to decline.

6. Discussion and Conclusion

Our cloud correction sensitivity experiments suggest that model clouds play a critical role in causing biases in the simulated relationship between sulfate and SPEI, consistent with our prior work based on evaluation of coupled climate-chemistry models (Y. Wang et al., 2017). Although we derive correction factors for model cloud parameters from observations and provide quantitative assessment of the cloud impact, these results are subject to several sources of uncertainty. First, the corrections are based on satellite observations, which are themselves subject to large retrieval uncertainties (Lin & Rossow et al., 1994; Stephens & Kummerow, 2007; M. H. Zhang, 2005). For example, the Moderate Resolution Imaging Spectroradiometer-derived liquid water path has been reported to have an uncertainty range of ±30% when validated against surface observations (Dong et al., 2008). Second, we neglected potentially large spatial variability in biases in the cloud fields, as we perturbed cloud parameters homogeneously across the entire eastern United States. Third, we perturbed boundary layer clouds in aggregate, without considering biases in the vertical profiles of cloud fraction or cloud mass. However, Figure 7 indicates that the largest differences are centered at around 2000 m altitude, while surface cloud fractions are more consistent.

The cloud correction experiments focus on bulk cloud properties (CF and LWC) at grid scales. However, biases in subgrid cloud variability can also have significant impacts. In the GEOS-Chem model, cloud droplets are generally assumed to be homogeneous with a bulk pH value. Heterogeneity of cloud droplets is only considered under certain conditions over ocean (Fahey, 2003). For pH-dependent sulfate production pathways, the assumption of a homogeneous cloud droplet pH may result in biases in the sulfate production rate (McVay & Ervens, 2017). Besides, Berg et al. (2015) showed a 40% increase in sulfate due to enhanced aqueous phase production after considering interactions between cloud and aerosol in subgrid shallow and deep convection. In addition, the model time step may be too long to properly represent clouds that reside for only a few minutes and dissipate. These issues pose challenges to the assumption of thermodynamic equilibrium for gas dissolution, and neglect portions of the sulfate cycle that are present in the actual atmosphere.

Although cloud biases in the model can account for most of the problems in the simulated sulfate-SPEI relationship, the tendency of the model to overestimate surface sulfate concentrations cannot be explained by these biases. The persistent sulfate high bias in M2GC may be related in part to the coarse resolution of the model grid, which cannot fully resolve the boundary layer or the processes that take place within it. K. Yu et al. (2018) have shown that weaker transient vertical motion at coarser model resolution can cause overestimates of up to 40% in certain components at the surface level. Overestimation of oxidant levels could also result in excessive sulfate formation in the model. Indeed, previous studies suggest that the GEOS-Chem model tends to overestimate OH and H2O2 compared to surface-based proxy measurements and satellite observations (Allen et al., 2013; Hu et al., 2018; Prinn et al., 2005).

To summarize, we present comprehensive evaluations of the GEOS-Chem simulation of sulfate under different levels of dryness as driven by MERRA (MGC) and MERRA-2 (M2GC) reanalysis products. While both simulations fail to reproduce the observed negative relationship between sulfate and SPEI, M2GC
represents a significant improvement over MGC owing to the more reliable representations of clouds and precipitation in MERRA-2 relative to MERRA. Significant overestimates of surface cloud cover in MGC are largely eliminated in M2GC, reducing overestimates of sulfate under wet conditions. This reduces biases in the simulated sulfate-SPEI relationship in M2GC by up to 50% relative to MGC. M2GC also better reproduces differences in precipitation and wet deposition across different SPEI bins, further alleviating problems in reproducing differences between mean sulfate concentrations under dry and wet conditions relative to MGC. However, MERRA-2 still contains significant biases in boundary layer cloud properties. Although cloud correction experiments suggest that these biases in cloud fields account for most of the bias in the simulated relationship between sulfate and SPEI under M2GC, correcting these biases does little to reduce the positive bias in sulfate concentrations. Our results emphasize the critical role of meteorological inputs in simulating sulfate, and provide a new perspective on a theme discussed by previous studies (Choi et al., 2017; Gong et al., 2006; Liu et al., 2009; Luo et al., 2011; Mueller et al., 2006; von Salzen et al., 2000; Yu et al., 2017; Zhang et al., 2007). Future research will be needed to evaluate model clouds at higher spatial, temporal, and vertical scales, considering the large sensitivity of sulfate simulations to clouds.

Acknowledgments

We gratefully acknowledge the individuals and groups involved in providing observations at IMPROVE (https://views.cira.colostate.edu/fed/QueryWizard/Default.aspx), U.S. EPA (https://www.epa.gov/outdoor-air-quality-data), and NADP (https://nadp.silwisc.edu/) networks for the surface quality XIE ET AL. 16

References


