Data assimilation for Numerical Smoke Prediction

Edward J. Hyer[1]

Christopher P. Camacho[1]

David A. Peterson[1]

Elizabeth A. Satterfield[1]

Pablo E. Saide[2]

[1]Naval Research Laboratory, Marine Meteorology Division, Monterey CA

[2] Department of Atmospheric & Oceanic Sciences (AOS) and Institute of the Environment and Sustainability (IoES), University of California Los Angeles, Los Angeles CA

1 Abstract

Skillful forecasts of weather phenomena in numerical models begin with the most accurate set of initial conditions achievable from observational datasets. The process of combining observations with numerical model predictions is called data assimilation. This chapter describes the types of observations available for data assimilation in models that predict the transport, fate, and impacts of smoke pollution. Observation properties needed for effective data assimilation are identified based on experiences with a variety of observation types in data assimilation experiments, compiled from the published literature. The second half of the chapter surveys the data assimilation methodologies that have been applied to smoke aerosols, and describes specific problems associated with the smoke observations that require innovative techniques in data assimilation. The chapter concludes by providing an outlook for future research and development in data assimilation for smoke prediction models. Data assimilation for prediction of smoke is an emerging area of development that promises to greatly improve forecast skill as new datasets and techniques are applied.

2 Introduction

A cascading sequence of numerical models are used to predict downwind impacts of smoke: *fire behavior and smoke emission models* translate observations of fires into model-relevant quantities, *plume rise models* describe fine-scale vertical motion of smoke at the fire source, and atmospheric *transport models* simulate the evolution and transport of smoke and the effects of local, mesoscale, and synoptic meteorology on the distribution of smoke. In this chapter, we refer to an atmospheric transport model used to predict the movement and distribution of smoke plumes as a *smoke transport model*.

For use in operational forecasting, these models require very timely information about fires, smoke, and meteorological factors. Observations of smoke are used to constrain the location, thickness, and composition of atmospheric smoke plumes. Under the right circumstances, these observations can be combined with model predictions of plume extent

and direction to improve predictions of downwind smoke. This process of combining numerical model predictions with observations is called *data assimilation*. Data assimilation is an essential component in skillful atmospheric models of all kinds, but it demands accurate numerical estimates of the uncertainty in both the forecast model and the relevant observations. Initializing models based on sparse observations also requires understanding how these uncertainties vary in space and time (spatial and temporal correlations).

Previous chapters have discussed the observations of fires that are relevant to this problem. The condition and behavior of the fire itself are *boundary conditions* in the terminology of an atmospheric model. Data on fire characteristics (see Chapters 3 & 4) is used to inform sub-models of fire behavior and smoke emissions (see Chapter 5) that translate events outside the model domain (fires on the ground) into changes to the smoke transport model state (the three-dimensional, time-variable numerical description of the model atmosphere).

As discussed in Chapter 8, models of fire behavior and plume rise may be *tightly coupled* to the smoke transport model, meaning that temperature, wind and humidity predictions from the atmospheric model are fed into the sub-model of fire behavior, and sensible and latent heat flux from the fire are introduced as boundary conditions to influence the smoke transport model.

The application of fire observations into models of fire behavior can be treated as a data assimilation problem. However, the formal methods of data assimilation are more fully developed for assimilating observations of the atmosphere that can be used to directly perturb the smoke transport model state to obtain a more accurate starting point for a model prediction, and thus a more accurate forecast. A different category of data assimilation problem is when observations are used to constrain emissions or other non-meteorological parameters, this is generally referred to as *inverse modeling*. Many operational aerosol prediction models use data assimilation to improve predictions of plumes of pollution, dust, and smoke. This chapter discusses the technical considerations for improving model predictions of downwind smoke impacts through data assimilation.

The first half of this chapter discusses the types of observations available and the characteristics of those observations that make them more or less useful in the context of data assimilation for smoke prediction. Many observations that are clearly useful in managing health impacts of smoke are difficult to apply in numerical forecasts, for a variety of reasons we will discuss. We will attempt to outline some key characteristics that apply to existing and future observations affecting their potential for data assimilation.

The second half of the chapter discusses some of the mechanics of data assimilation and how they relate to the problem of smoke from fires. Many technical details in the implementation of data assimilation have very substantial consequences for the specific problem of smoke plumes.

At the end of the chapter, we briefly describe some directions of research that could lead to improved communication between observations and models, and ultimately improved forecasts of downwind smoke impacts.

3 Matching observations to forecast problems

This section describes the principal types of observations relevant to the smoke forecasting problem and the domains in which each provides information to improve the forecast. The template for a smoke prediction has two main components, "how much smoke?" and "for how long?", or simply the magnitude and duration of smoke impacts. For most purposes, smoke impacts at the surface are the target of prediction. When smoke transport models are used to aid in these predictions, they serve two purposes. First, the model gives an estimate of conditions that is spatially continuous across the model domain. Second, the model computes integration in time including the effects of the atmosphere on movement, dilution, and removal of smoke. While there are many applications that require models in conjunction with contemporary observations (such as estimation of exposure for epidemiological studies, e.g. (Hammer et al., 2020)), this section specifically considers the condition where the model must make a prediction at a location and time where observations are unavailable.

If we break this problem into components, we can examine how different types of observations contribute to this smoke prediction problem. The model provides a coarse representation of a smoke plume: model resolution is often insufficient to capture much of the variation in smoke, especially in the vertical, and the model representation of smoke composition is inevitably vastly oversimplified. In spite of these simplifications, we find that all of the available observations provide limited information, leaving the model to fill in the rest. Therefore, benefit to the forecast can be increased by integration of multiple observation types, utilizing each observation's strengths and supporting weak points of observations with strong model treatments.

3.1 Overview of observations relevant to smoke plumes

The objective of data assimilation in a smoke transport model is to constrain key aspects of the smoke plume by observations, and thus estimate the most accurate initial condition and/or emissions for a numerical prediction of the smoke plume's evolution. The *model state* of the smoke prediction model is a 3D gridded representation of the atmosphere with various properties of the atmosphere, including aerosols and trace gases, represented at each grid cell. For inverse modeling, the state may include spatial and temporal patterns of emissions magnitude and composition. At a more abstract level, we can identify the properties of a smoke plume that will impact the ability of the model to correctly predict its evolution. These properties are, roughly: position of the plume, horizontal extent, vertical extent, quantity of smoke, and smoke composition. A wide range of observations can be used to reduce uncertainties in these properties, but no single observation provides the comprehensive information needed by the model.

Table 1 overviews the main types of observations that have been used to constrain atmospheric models used to predict smoke. This list is not exhaustive but captures the types of data that are most likely to be practically useful in a forecast context (see Section 3.2.2). Products from research satellites may provide detailed information on smoke (for instance, MISR (K. T. J. Noyes et al., 2020)), but are often not suited to near-real-time application. Each of these types of observations has been examined in detail in the references cited in this chapter and other chapters of this book.

Table 1.Observation types and some of their relevant properties for use in smoke forecasting. Acronyms are defined at the end of this chapter.

Observation	Representative Data Source	Coverage	Reduction of Model State Uncertainty	Key limitation for smoke prediction
Surface PM	AirNow	Sparse in CONUS; very sparse globally	Moderate-High	2m height; no information on elevated smoke
Sun / Moon Photometer	AERONET	Sparse in CONUS and globally	Moderate	Column-integrated retrieval, no vertical specificity
Satellite AOD retrieval (polar)	MODIS MxD14; MAIAC; NOAA Enterprise Aerosol Products	Global ~1x/day per satellite; daytime only	Low-Moderate	Column-integrated; resolution/noise tradeoffs; retrieval failures in thick smoke
Satellite AOD (geostationary)	GOES-R ABI NOAA Enterprise Aerosol Product	Every 15 minutes or better within satellite domain	Moderate	Column-integrated; Resolution/noise tradeoffs; retrieval failures in thick smoke
Ground-based Lidar	NASA MPLNET	Very sparse	Moderate	<200 sites globally
Space-based Lidar	CALIOP	Global w/16-day repeat	Moderate-High	Sparse; cannot see under thick clouds
Total column Space-based trace gas retrievals	TROPOMI CO, NO_2, and HCHO	Global ~1x/day; daytime only	Moderate	Column-integrated
Space-based trace gas retrievals from sounders or profilers	MOPITT, CrIS, and IASI CO, OMI/OMPS	Global coverage every ~1-3 days, day	Low-Moderate	Coarse resolution; limited vertical signal; limited

	CO, HCHO and NO2	and night data		sensitivity near surface

The general outlook is that the greater the level of accuracy, precision, and detail provided by an observation, the more sparse those observations are. Surface PM and trace gas monitors provide the most direct observation of the smoke quantity, but are too sparsely located to constrain plume location and extent (see Chapter 6). PM measurements may also be biased relative to the intended predictand. For instance, evaporation of organic aerosol at the surface (Selimovic, Yokelson, McMeeking, & Coefield, 2019) can act as a confounding factor. Ground-based lidar observations, which observe the backward scattering of laser light, capture the vertical mixing that is a critical determinant of surface smoke, but the locations of these few instruments rarely facilitate application to prediction of a local or regional smoke event. Some lidar observations, especially low-power instruments such as ceilometers, may attenuate for thick smoke plumes, limiting their vertical coverage. Satellite instruments provide far greater numbers of observations over far greater areas, but the resolution of these observations is, compared to the size of the active burning area, coarse (1km or more). The weak signal of smoke in satellite observations of reflected sunlight often demands additional spatial aggregation to reduce noise. Aerosol optical depth (AOD) is an optical quantity, which not only depends on aerosol mass but also on other variables like relative humidity and the aerosol composition and size distribution. Conversions between aerosol mass and optical properties have been shown to vary substantially with smoke age (Kleinman et al., 2020). Additionally, multiple algorithms filter out AOD retrievals on thick smoke (see Section 3.2.3), and the ones that do provide them can be uncertain due to assumptions on smoke absorption. Ongoing work seeks to advance the capability of space-based sensors to constrain aerosol composition and other properties (e.g. (Rogozovsky et al., 2021)).

Despite the limitations, assimilation of many of these types of observations has been shown to greatly increase the predictive skill of smoke transport models (Jianglong, Reid, Westphal, Baker, & Hyer, 2008; Pablo E. Saide et al., 2014; Pablo E. Saide et al., 2015; Yumimoto et al., 2018). Much of the published literature has focused on global models, which are recognized for accurately representing long-range transport of heavy aerosol plumes. However, for many practical smoke prediction scenarios, observations that may show significant benefit in a global model may provide very little skill at the scale where smoke impacts are most significant. The next section discusses the characteristics of observations that make them impactful for real-world forecast scenarios.

3.2 Applicability of observations to real-world smoke forecasting

3.2.1 Types of forecast uncertainties

The goal of data assimilation is to apply observations to correct errors in a short forecast and produce the most accurate initial conditions for subsequent predictions. Therefore, a good place to start to is to consider common types of forecast errors and how the inclusion of observational information could act to mitigate such issues. This section will outline these forecast errors in detail. Subsequent sections will examine specific topics of timeliness and spatial matching.

3.2.1.1 Missing fires

Without a smoke source, there is very little a numerical model can achieve for smoke prediction. Numerical smoke models use the fire observations discussed in the earlier chapters of this book to estimate smoke production (see Chapter 5). The availability of those observations to the model is a precondition for effectively utilizing observations of smoke. This is mostly a matter of timing, which we discuss in more detail below.

Errors in fire timing can arise from the temporal resolution of the fire observations. Polar-orbiting weather satellites that can detect relatively smaller fires require parameterization of fire behavior over the diurnal cycle; coarser, less-sensitive geostationary fire observations have typically been used for this purpose (e.g. (E. J. Hyer et al., 2013; Mingquan Mu et al., 2011; Reid et al., 2004b)). The errors associated with infrequent fire observations can be mitigated by assimilation of observations with better temporal resolution, but satellite observations have been shown less effective than *in situ* observations due to lack of satellite aerosol retrievals near the fire source (Pablo E. Saide et al., 2015) (see Section 3.2.3).

The spatial resolution of satellite fire observations is a significant source of error for wildfire responders on the ground, and for matching fire observations to fuel maps (Edward J. Hyer & Reid, 2009), but represents a small error at the scale of significant smoke plumes. Atmospheric retrievals of smoke properties such as aerosol optical depth (AOD) are almost universally less precise spatially than fire observations, so assimilation of these observations will do little to mitigate these spatial errors.

3.2.1.2 Too much / Too little smoke

Smoke production magnitude is a recognized weakness of satellite-based fire inventories (e.g. (Al-Saadi et al., 2008; French et al., 2011; F. Zhang et al., 2014)). Error in smoke source magnitude is the error most readily addressed by assimilation of smoke observations.

If a smoke source is present in the model at the approximately correct time and location, observations of the smoke downwind can be used to correct the smoke magnitude. As noted in Section 4, the data assimilation system is sometimes used to correct the smoke loading in the model state, and sometimes used to apply a correction to the smoke source magnitude (i.e., an emission inversion). The higher the spatial resolution, the more precise the correction can be, but the improved description of smoke magnitude from high-resolution observations can only be achieved if errors in simulated wind speed and direction are small.

3.2.1.3 Vertical placement and transport

Ventilation of smoke is a key consideration for downwind impacts, and a key component of any prediction of smoke air quality. Complex interactions between the fire and the atmosphere are explicitly simulated in some high-resolution models, as discussed in Chapter 8 of this volume, but in most regional and global models, these interactions are represented by simple parameterizations. The energy release from the fire is relevant to this question, but studies using existing observations of fire energetics have found only low correlation with smoke vertical placement (Paugam, Wooster, Freitas, & Martin, 2016; D. Peterson, Hyer, & Wang, 2014; Val Martin et al., 2012), due to large uncertainties in fire energy (Louis Giglio & Schroeder, 2014), the importance of model representation of stable

layers (Val Martin et al., 2012), and the complexities of 3D plume behavior (Y. Q. Liu et al., 2019; Nelson, Butler, & Weise, 2012; Potter, 2012).

Observations at the surface are obviously relevant to correct errors in smoke vertical placement, but assimilation of these observations requires careful consideration of the vertical and horizontal spread of the information, represented in data assimilation as error correlations (Z. Q. Liu et al., 2011; J. L. Zhang et al., 2011). Profile information such as that provided by a lidar can be very impactful to improve placement of smoke (Cheng et al., 2019). Several recent papers have demonstrated the potential of passive satellite observations to retrieve some information on smoke plume height (Lyapustin, Wang, Korkin, Kahn, & Winker, 2020) (Choi et al., 2021; Xu et al., 2019; Xu et al., 2017) (Carr et al., 2020) or smoke over cloud (Alfaro-Contreras, Zhang, Campbell, Holz, & Reid, 2014; Jethva et al., 2016; Jethva, Torres, Waquet, Chand, & Hu, 2014). This information could improve predictions of long-range transport.

Downwind of a fire, a plume will remain aloft or mix into the surface layer controlled by a variety of processes at scales from synoptic subsidence to turbulent mixing (e.g. (Colarco et al., 2004)). When the smoke transport model erroneously mixes or fails to mix smoke between the surface and layers aloft, only observations resolving smoke at the surface can be used to place the smoke correctly. The surface pollution may have long spatial correlation lengths in temporal averages (Kaku et al., 2018; Y. Z. Wang et al., 2020), but spatial correlation lengths associated with smoke events are often short, with the implication that corrections from surface observations cannot be applied to large areas. Networks of low-cost sensors which provide much finer spatial coverage could be a solution to this problem (Delp & Singer, 2020; Robinson, 2020).

3.2.1.4 Horizontal movement of smoke

If errors in wind speed and direction are large, assimilation of smoke observations without vertical information will likely do little to improve the forecast. Even if the plume position, shape, and size are fully mapped with satellite data, the lack of any model prior for the vertical placement of the smoke will result in significant error in placement of aerosols in the model, even in aerosol species not corresponding to smoke.

In practical terms, wind errors will generally be small in the short model forecasts. However, in the case of small plumes and observations very near the smoke source, small errors in wind direction can be just as costly as large errors downwind. Assimilation of smoke observations onto a background with a misplaced plume will result in unrealistic depiction of the smoke in the analysis, as well as a mismatch between the smoke and other meteorological fields in the simulation (e.g. (J. L. Zhang, Reid, Christensen, & Benedetti, 2016)). Adding or subtracting mass from the downwind tail of a plume is one of the primary means by which smoke predictions are improved by data assimilation, so correcting small errors in wind speed is a clear use for data assimilation in smoke prediction.

3.2.1.5 Persistence in emission forecasts

The current state of air quality forecast models is to continue the fire emissions estimated from the latest satellite fire detections into the next forecast cycle, which we define here as "persistence". However, fires can often change substantially from day to day, which can generate large uncertainties in smoke forecasts (Ye et al., 2021). This results in systematic underprediction of smoke after periods of extreme fire weather and overprediction when

fires extinguish rapidly (e.g. due to precipitation). Limited studies have been performed to predict changes in fire behavior and smoke emissions during the weather forecast period (e.g., (D. Peterson, Hyer, & Wang, 2013; D. A. Peterson et al., 2015)), and this issue remains an active area of research in the smoke forecasting community.

3.2.1.6 Prescribed diurnal cycles

Many smoke emission models rely on fire observations collected only once a day. In order to distribute emissions at the hourly resolution, emission models typically assume that fires follow a prescribed diurnal pattern based on observations with better temporal resolution (L. Giglio, 2007) (M. Mu et al., 2011), with a peak in fire activity in the afternoon and low emissions during evening and morning hours. Fires in practice often deviate from the climatological pattern (Pablo E. Saide et al., 2015; Ye et al., 2021), especially for extreme fire behavior which often burns throughout the night (D. A. Peterson et al., 2015). Recent studies have found strong correlation between geostationary FRP and smoke concentrations (Wiggins et al., 2020) which is motivating multiple groups to move to using these observations to distribute emissions to an hourly time-scale (e.g. (Reid et al., 2004a; Roberts et al., 2015; Wang et al., 2006)). However, there are two caveats to this. First, in order for these observations to make a difference there needs to be a post run (see next section) of the forecast or assimilation system so that corrected emissions are introduced at the time they occurred. Second, similarly to the previous subsection, persistence of diurnal cycles is likely not to occur, and thus methods need to be develop to forecast emission diurnal cycles as well.

All of these different types of model errors can be corrected to some degree by assimilation of smoke observations. In each case, we see that correction for one type of model error is most effective when other errors are small. Operational prediction systems thus achieve accumulating benefits from repeated cycles of short forecasts followed by data assimilation. But for the specific case of forecasting impacts of significant smoke events, important limitation on the practical results are imposed by the logistics of timing and the placement of observations. These dependencies are discussed in the following sections.

3.2.2 Timeliness of observations

The single most important characteristic for utilization of observations to support model predictions is timeliness. Models at the global scale typically run four forecast cycles each day, and at each cycle an initialization is performed using observations in a six-hour window around the nominal forecast start time. We describe this using a notation where *T* is the valid time of the forecast initialization, so the observation window for initialization at *T* is from *T-3* to *T+3*. The delay between the time of an observation and its availability for use in the forecast is referred to as *latency*. If the forecast cycle is started four hours after *T* (*T+4*), and the latency is 1 hour, then all observations up to *T+3* will be used in the initialization. If the cycle is started at *T+4* and the latency is 3 hours, then each cycle will use observations only from *T-3* to *T+1*. Many global systems therefore include a "post" run at ~*T+8* (also referred to as reanalysis or final analysis), where all observations can be used to prepare for the following forecast cycle.

In a system configured this way, fires and smoke observed at (for example) 20 UTC may be reflected in forecasts initialized at 18UTC which may be available as early as 22UTC (two hours after observation time). Observations at 21UTC will only be included in the "post" run, if latency is too high, as it often is for derived satellite products or research products. These observations will only be included in the forecast products initialized at 00UTC which become available no earlier than 04UTC the following day (seven hours after observation time). This scenario leaves out several other potential sources of delay. A forecaster compiling a forecast at 12UTC for the following day may have a numerical model output incorporating observations as late as 9UTC, but only observations from 3UTC or earlier are guaranteed to be incorporated in model runs available for preparing a forecast at 12UTC.

When the observations come from a satellite in a sun-synchronous polar orbit such as MODIS or VIIRS, the interactions between the observation times and forecast cycles produce aliasing effects where forecast cycles have widely varying utility for specific locations. A further aliasing effect will be noted in the verification if the model is verified using observations available at specific times (such as daytime-only Sun photometer observations). This effect can be seen in (J. Zhang et al., 2014) where 18-hour forecasts in many cases verified better than 12-hour forecasts; the satellite data assimilated were all collected near local noon, and the benefits of data assimilation manifested in a limited zonal range such that impacts would be missing from the daytime-only verification of the 12-hour forecast. This effect is illustrated in Figure 1 and Table 2, and will be highly noticeable in a regional prediction system.

High-resolution regional-scale models such as the NOAA RAP start new prediction cycles as often as each hour. From Lin et al. (Lin, Weygandt, Benjamin, & Hu, 2017): "RAP and HRRR use a short observation data cutoff time of about +35 min after the analysis time… For RAP, this window extends back no more than 1.5 h from the model initial time." Some types of observations, such as geostationary satellite radiances, are available for assimilation within 10 minutes of acquisition, but observations with more than 30 minutes latency will frequently miss the cutoff, and observations with >2 hours latency have no path to ever be included in a system of this type.

The latency of satellite based AOD observations is the difference between the observation time and the time that the observation is available for integration into the modeling system. Thus, downlink from the satellite, generation of data products in the satellite ground station network, and delivery of products to end users are all included in the overall latency. This question of latency, together with the common satellite infrastructure used by both NWP observations as well as observations of fire and smoke, explains why weather satellite data are the most commonly assimilated type in global models. Many research satellites (e.g. CALIOP, MISR) deliver data very applicable to problems of smoke forecasting, but the space and ground systems were not originally designed to deliver data within the tight latency requirements of NWP models.

At the Naval Research Laboratory Marine Meteorology Division, the MODIS Aqua and Terra AOD products (MYD04_L2 and MOD04_L2, respectively) are downloaded from NASA LANCE's near real-time servers every 30 minutes. The VIIRS (SNPP/NOAA-20) and ABI (GOES-16/17, full disk) AOD products (NOAA Enterprise algorithms) are

downloaded from NOAA every 10 minutes. Maintaining these efficient download workflows helps to achieve the minimal latency for satellite observations (Figure 2). The GOES AOD products exhibit the lowest latencies, with modes around 13 minutes and 22 minutes. The MODIS and VIIRS latencies are similar in their distributions, but there is an observed platform dependence for granule latency. AQUA and SNPP show a bimodal distribution, with peaks around 63-66 minutes and 120-135 minutes. TERRA and NOAA-20 latency is unimodal with means around 77 minutes (σ=18) and 57 minutes (σ=10), respectively. This discrepancy is related to the satellite downlink; Terra and NOAA-20 utilize the Tracking And Data Relay Satellite System (https://www.nasa.gov/directorates/heo/scan/services/networks/tdrs_main) to achieve more frequent downlink leading to reduced latency.

Ground-based observation networks have made great advances in latency over the past decade due to the improvements in electronics and communication networks. The AirNow air quality observation network is partnership between multiple U.S. Federal agencies (including the Environmental Protection Agency (EPA), National Oceanic and Atmospheric Administration (NOAA), and the U.S. Forest Service among others) and local agencies. Figure 3 shows a snapshot of the timeliness of available hourly PM2.5 and PM10 site observations within a three day window. A total of 1131 sites reported PM2.5 or PM10 during this window, with the bulk of the sites (80%) reporting within a two-hour window. The AirNow network has far lower latencies compared to the Aerosol Robotic Network (AERONET, (Holben et al., 1998)). A small subset of AERONET sites (20%) are capable of reporting Level 1.5 data within three hours, but it may take up to 20 hours for roughly half of the data to become available. AERONET Level 2 products include additional quality control, but are not available in near real-time (Giles et al., 2019).

With a rapid update model and a low-latency system to deliver observations, many types of observations can contribute to smoke prediction. However, latency of observations remains the primary bounding condition for utility of numerical smoke predictions. In the best-case scenario, the first forecast incorporating observations of a fire or smoke plume will likely be 2-4 hours after the observation. The inference is that the numerical forecast is far less useful for predictions of smoke less than 2-4 hours downwind, corresponding to 30-60km for moderate wind speeds of 4 m s-1. For a global prediction system, this "radius of inutility" can be much larger, exceeding 9 hours in many current operational configurations. For prediction of smoke impacts in the near field within this radius, a different type of modeling such as trajectory modeling may be more appropriate.

Geostationary observations mitigate this effect because observations are available throughout the day. Several methods of retrieving aerosol optical depth at nighttime have been prototyped (Fu et al., 2018; Johnson, Zhang, Hyer, Miller, & Reid, 2013; McHardy et al., 2015; Wang, Aegerter, Xu, & Szykman, 2016; J. Wang, M. Zhou, et al., 2020; J. L. Zhang et al., 2019; Zhao, Shi, Yu, & Yang, 2016), but so far there is no operational nighttime aerosol product available. Nighttime observations also provide enhanced information on fire behavior by combining the visible signal of flame with the infrared heat signature of fires; this can provide important insight into overnight fire behavior that is very important to predict fire activity the following day (Polivka, Wang, Ellison, Hyer, & Ichoku, 2016; J. Wang, S. Roudini, et al., 2020). From a practical standpoint, it is easy to see how nighttime

observations could be very valuable, in terms of providing information about overnight progress of a fire and early insight into smoke conditions for the following day. Assimilation of CO observations from sounders shows promise for smoke prediction due to their nighttime availability. While not directly constraining aerosol, increments can be applied through covariation of smoke and CO as they are co-emitted in biomass burning plumes (Arellano, Hess, Edwards, & Baumgardner, 2010). But currently available geostationary observations of smoke are daytime only, and observations with solar elevation lower than 10° have large uncertainty, so these observations will still not be available near sunrise and sunset.

Timing of observations interacts with the practicalities of forecasting to impose important limits on the ability of observations to improve a real-time forecast. Numerical forecasts available during the morning hours may have overnight fire observations applied in the smoke source, but the smoke plume characteristics will likely reflect only smoke observations from the previous day, even with geostationary observations. Only around midday will model outputs incorporating same-day smoke observations become available.

3.2.3 Range from smoke source

The consequence of the timing considerations from the previous section is that only predictions of smoke at least two hours or multiple tens of kilometers from the source will see improvement from assimilation of smoke observations. The proximity of the observation to the smoke source is a key factor determining how effectively that observation can mitigate different types of errors in the forecast model.

Figure 4 shows an array of different satellite aerosol products for two days of extreme fire activity in northern California during September 2020. The top row is true color imagery giving the context of land, ocean, and smoke in these scenes. Below that, rows depict satellite AOD retrievals, ordered from fine to coarse spatial resolution, starting with nominal 750m VIIRS data using the NOAA Enterprise aerosol retrieval (H. Zhang et al., 2016), and proceeding through the nominal 1km MAIAC (Lyapustin et al., 2011), the nominal 2km GOES-17 ABI AOD using the NOAA Enterprise retrieval (H. Zhang, Kondragunta, Laszlo, & Zhou, 2020), the nominal 10km MODIS combined "Dark Target/ Deep Blue" retrieval (Levy et al., 2013), and the 0.5 degree Level 3 NRL/UND DA-quality AOD product (Edward J. Hyer, Reid, & Zhang, 2011; Y. Shi et al., 2011). For data and visualization sources, see Section 6.

Several aspects of satellite aerosol retrievals relevant to smoke modeling are apparent in Figure 4. Most obviously, in all of the products there is some part of the thickest smoke that is excluded from retrieval. At a given wavelength, there are hard physical limitations on the maximum retrievable AOD because a complete retrieval requires at least some light reflected from the surface to penetrate through the plume to reach the sensor. Many sensors based on retrieval at 550nm adopt an AOD of 5 as a retrieval limit for this reason. Longer wavelengths are more transparent to smoke, and some proposed methods exploit this to obtain retrievals even in smoke which is opaque in visible wavelengths (Eck et al., 2019).

Heavy smoke conditions can also trigger textural filters intended to ensure exclusion of clouds. The MODIS Collection 6.1 "Dark Target" aerosol retrievals are the basis for the bottom two panels in Figure 4 (the bottom row is the NRL/UND DA-quality product with

additional stringent QC filtering), and the conservative approach to clouds taken by these retrievals results in the gaps in these datasets over the heaviest smoke. This is in spite of significant effort to implement a "call-back" method in MODIS Collection 6 aerosol retrievals to "rescue" retrievals in heavy smoke initially flagged as cloud (Levy et al., 2013). An experimental method based on MODIS achieved significantly improved retrieval of heavy smoke, at the cost of weakening the cloud filtering (Y. X. R. Shi et al., 2019). The higher-resolution retrievals shown in Figure 4 have smaller excluded areas, and a greater ability to retrieve heavy smoke and separate smoke and cloud in the NOAA Enterprise algorithm (Laszlo & Liu, 2016) and the MAIAC (Lyapustin, Korkin, Wang, Quayle, & Laszlo, 2012). All retrievals will exclude some part of thick smoke near the fire source. This will results in significant underestimation of heavy aerosol events (Y. X. R. Shi et al., 2019) and is an ongoing area of improvement for aerosol retrievals.

The filtering methods applied to exclude flawed and cloud-contaminated retrievals from aerosol retrievals are necessary to assimilate retrievals on a systematic basis (J. L. Zhang, Reid, & Holben, 2005), but these impose a low bias on the assimilated observations, which will be transferred to the model analysis if not mitigated. Some mitigation can be achieved by using an ensemble approach to obtain flow-dependent background errors that can help to preserve gradients (Rubin et al., 2017) (also see Section 4.2); it may also be profitable to explore automated means for identifying contexts where cloud filtering can be relaxed to permit improved retrieval of thick smoke.

3.3 Discussion: how and when are aerosol observations helpful in smoke prediction?

The discussion above identifies critical properties of observations that make them useful for data assimilation in smoke prediction systems:

1) Observations must be available in a timely fashion with fully automated data flows;
2) Observations must project onto the model state (e.g. some observations may contain information on smoke composition, but this information can only be assimilated into a model that simulates composition information);
3) Observations should not exhibit systematic bias in the presence of thick smoke (although many commonly used observations do this);
4) The smoke model must adequately constrain the characteristics of the smoke plume that are not observed (e.g. plume height for column-integrated retrievals, plume shape for point retrievals, accurate conversions between aerosol mass and optical properties)

Based on these considerations, the trajectory of development of aerosol data assimilation systems in aerosol prediction models can be understood. The wide-area, low-latency data from meteorological satellites is the most straightforward fit into model systems. However, numerous other observation types may satisfy these criteria (e.g. aircraft observations (Pablo E. Saide et al., 2015) or surface measurements (Ma et al., 2019)). Many potentially beneficial data sources have not yet been explored.

4 Considerations for assimilation of smoke observations

Equally important to what observations are included in data assimilation is the method applied to combine the observations and the model to obtain an optimal state. This section briefly surveys the published literature on aerosol data assimilation, and discusses some of the unique interactions between observations of smoke aerosols and data assimilation systems that complicate the problem and demand advanced techniques. The final section outlines some areas with future work in the near and medium term.

4.1 Development and scope of aerosol data assimilation efforts

The effort to assimilate aerosol observations to improve model simulations of the atmosphere goes back 20 years (Yu et al., 2003) (Wang, Nair, & Christopher, 2004). The Navy Aerosol Analysis and Prediction System (NAAPS) began to use assimilation of MODIS AOD observations for forecast initialization in 2009 (J. Zhang, J. S. Reid, D. L. Westphal, N. L. Baker, & E. J. Hyer, 2008), making it the first operational aerosol model to include data assimilation. Since that time, many different groups have worked on aerosol data assimilation systems for operational and research models.

Many different datasets have been used in aerosol data assimilation including EPA surface monitor PM2.5 data (Schwartz, Liu, Lin, & McKeen, 2012) and surface aerosol optical measurements (Chang, Zhang, Li, Chen, & Li, 2021), visibility networks (Clark et al., 2008), AERONET Sun photometer data (Rubin et al., 2017), and ground-based lidars (Liang et al., 2020) (Yumimoto et al., 2008). Many kinds of satellite data have also been assimilated into aerosol models: visible/near-infrared satellite data from MODIS (Edward J. Hyer et al., 2011; Z. Q. Liu et al., 2011), MISR (Lynch et al., 2016), UV observations from OMI (Lee, Zupanski, Zupanski, & Park, 2017; J. L. Zhang et al., 2021), multi-angle, multi-polarization observations from POLDER (Generoso et al., 2007), lidar backscatter from CALIOP (J. L. Zhang et al., 2014), and geostationary observations (Pablo E. Saide et al., 2014; Yumimoto et al., 2016) and even cloud retrievals (P. E. Saide, Carmichael, Spak, Minnis, & Ayers, 2012). Assimilation methods tested include a wide variety of both variational (Benedetti et al., 2009; Yumimoto & Takemura, 2013), optimal interpolation (Tang et al., 2017), and ensemble (Dai, Schutgens, & Nakajima, 2013; Rubin et al., 2016; Yumimoto & Takemura, 2011) methods.

While most aerosol data assimilation efforts have been aimed at forecast improvement (Xian et al., 2019), important work has also been done applying data assimilation to retrieval of aerosol sources (Dai, Cheng, Goto, et al., 2019; Dubovik et al., 2008), evaluation of aerosol property assumptions used in models and remote sensing (Goto, Schutgens, Nakajima, & Takemura, 2011), long-term reanalysis of atmospheric aerosol (Gelaro et al., 2017; Lynch et al., 2016; Yumimoto, Tanaka, Oshima, & Maki, 2017), and to study aerosol interactions with weather (P. E. Saide et al., 2015). A review of assimilation methods and applications for trace gases and aerosols can be found in (Bocquet et al., 2015).

4.2 Problems particular to aerosol data assimilation

While the techniques of data assimilation are generally applicable to aerosol, obtaining the most accurate forecast results requires a data assimilation scheme that accommodates the realities of the aerosol observations as well as the realities of the atmospheric aerosol, both as it exists in nature and in the simplified form in which it is represented by the model. This

section discusses a few of the specifics of aerosols and aerosol observations and how they affect data assimilation. The distinguishing characteristics of smoke plumes interact with the behavior of both aerosol measurements and models to pose specific challenges for data assimilation to improve smoke prediction.

4.2.1 Non-Gaussian distribution of aerosol-related quantities

Most operational data assimilation systems, whether variational, ensemble or hybrid, employ the convenient assumptions that the errors obey Gaussian statistics and that the observation operator, which maps state variables to observation space, is linear ((Amezcua & Van Leeuwen, 2014) (Bannister, Chipilski, & Martinez-Alvarado, 2020), among others). The assumption of Gaussianity allows for PDFs to be defined by a mean and covariance. For variables such as AOD or atmospheric moisture, their probability distributions (PDFs) can deviate significantly from Gaussian. Nonlinear observation operators are often required for non-local observations. Atmospheric aerosols, measured in terms of mass concentration or optical extinction, cannot have negative values, and for this reason alone violate the assumption of Gaussian distribution. The true distribution of atmospheric aerosol quantity is generally recognized as lognormal (O'Neill, Ignatov, Holben, & Eck, 2000). Detailed analysis has shown that variation in aerosol within narrow spatio-temporal bounds influenced primarily by fine-scale meteorology may appear more Gaussian but at scales where variability is driven by aerosol sources and sinks and synoptic meteorology, a lognormal distribution best matches the observations (Alexandrov et al., 2016; Sayer & Knobelspiesse, 2019). When considering specifically the errors in aerosol observations, these errors will include errors of spatial and temporal sampling which can be modeled as draws from the lognormal regime, as well as instrumental errors and errors associated with underdetermined retrievals, which may or may not have a Gaussian distribution (Sayer et al., 2020).

The consequences of assuming Gaussianity for a non-Gaussian parameter can be significant. As discussed in Bannister et al. (2020), a Gaussian distribution with the same mean and variance will not represent the true PDF (which is especially detrimental for a multi-modal case), negative values may be allowed for bounded variables, and the mode may be a physically improbably state. There are several approaches to dealing with non-Gaussianity in data assimilation. The most computationally efficient is a variable transformation. As discussed in Amezcua and Van Leeuwen (Amezcua & Van Leeuwen, 2014), using a variable transformation will only provide an approximate solution when the requirement of Gaussianity is not met in the original space; however, it is clear that some transformations are better than others. Further, Bocquet et al. (Bocquet, Pires, & Wu, 2010) discussed the potential drawback of introducing an observation operator in a new space following a typically nonlinear transform. A more difficult approach is to use a more appropriate filter that does not require a Gaussian assumption. Although such methods can be computationally expensive, this is an area of active research. Particle filters (van Leeuwen, 2009; van Leeuwen, Kunsch, Nerger, Potthast, & Reich, 2019) generalize the Bayesian data assimilation problem by approximating PDFs through weighted samples. Bishop (Bishop, 2016) introduced an Ensemble Kalman Filter (EnKF) that extends the Gaussian assumption to also account for Gamma or Inverse Gamma PDFs. Fletcher and

Zupanski (Fletcher & Zupanski, 2006) show how a cost function can be modified to account for lognormal distributions.

Power transformation (Following (Yeo & Johnson, 2000)) can be applied, such that a variable u is mapped to "log" space through the transform:

$$u \rightarrow \log(1+u). \quad (1.1)$$

This transformation effectively pulls the tails of the distribution inward and results in a distribution that is closer to a Gaussian distribution. A number of centers use, or have previously used, log humidity as a control variable ((Bannister et al., 2020; Dee & Da Silva, 2003),among others).

Examination of modeled aerosol optical depth from NAAPS and AOD observations used in data assimilation shows how the untransformed AOD data deviate very significantly from the assumed Gaussian distribution (Figure 5, left column). The inset plots show the quantile-to-quantile comparison—normally distributed data will fall on the 1:1 line (red line in the figure). Thus, a variable transform could significantly improve the performance of an aerosol data assimilation system.

Several scaling based transformations have been applied for atmospheric moisture variables, most notably Dee and DeSilva (2003)and Hólm (Hólm et al., 2002). Dee and DeSilva (2003) proposed the use of pseudo relative humidity (PRH). PRH is defined by scaling the mixing ratio by the background saturation mixing. Hólm (2002) defined a transformation based on normalized forecast differences, which gives a more symmetric and closer to Gaussian PDF. Both PRH and the Hólm transform have been used for atmospheric moisture analyses. is worth noting that the approach of Hólm (2002) is more generalizable and could extend to other variables of interest.

Some AOD data assimilation systems do employ a transform of the form

$$AOD' \rightarrow \log(c + AOD) \quad (1.2)$$

with c taking the value of 0.01(Randles et al., 2017) or 0 (Pablo E. Saide et al., 2013), while others do not. Benedetti et al. (2009) explicitly tested this transform in part of their aerosol data assimilation system, and described their results this way:

> *As a side note, the implementation of the logarithmic variable did not dramatically improve the analysis performance. On the contrary, the RMS is higher and the correlation lower in the analysis with the logarithmic control variable than in the analysis with total mass mixing ratio. The reason for this could lie in the fact that the logarithmic control variable is only used at the level of the minimization, whereas the rest of the model is formulated in terms of mass mixing ratio. A more effective way to handle tracers could be to formulate the whole forward model in terms of logarithmic (hence positive definite) variables. This however would involve an extensive effort in modifying and rewriting the model, and it is not a viable option at this point. The use of alternative normalized control variables with a more Gaussian error distribution can still be investigated for future developments, following existing examples (Hólm et al., 2002).*

The right-hand column of Figure 5 shows that application of this transform with c=0.01 yields a more Gaussian distribution of background, observations, and innovations. However, significant distortions remain in the distribution because of the need to truncate negative AOD observations in order to perform the log transform (see next section).

Gaussian anamorphasis (GA) is another approach that has gained recent popularity in atmospheric data assimilation (e.g. (Brankart et al., 2012; Simon & Bertino, 2012; Zhou, Gomez-Hernandez, Franssen, & Li, 2011)) GA employs a transform based on the cumulative distribution (CDF) such that the moments of the transformed Gaussian variable are set to those of the original prior distribution (Bertino, Evensen, & Wackernagel, 2003). The simplest case of GA would be to transform the state variables and observations independently. Lien et al. (2013) used GA applied to precipitation, based on its model climatology, under the assumption that a forecast variable with more Gaussian climatological distribution would result in a more Gaussian error distribution. In their study, zero precipitation was handled as a delta function located at the median of the zero-precipitation part of the normal distribution. Lien et al. (2016) further showed that this approach also showed benefit in correcting the amplitude-dependent biases. Amezcua and Van Leeuwen (Amezcua & Van Leeuwen, 2014) explored GA in joint state variable and observation space.

4.2.2 Satellite retrievals permit negative values of AOD

AOD values of zero or higher are physically realistic, but the retrieval of AOD from reflected radiance is constructed as a problem of solving for the aerosol contribution as a deviation from the expected clear-sky radiance (Kaufman et al., 1997). Errors in the estimation of the clear-sky radiance can be positive or negative, so this formulation of the retrieval problem will always result in some negative values of retrieved AOD. Exclusion of these observations is not appropriate, because they constrain the true AOD to a narrow range and have a high probability of being cloud-free. However, truncation of these values to 0 or to a low nominal AOD value distorts the distribution of the observations.

4.2.3 Observation quality control systematically removes heavy smoke

The most important error in AOD retrievals obtained by inversion of observed radiances is contamination by cloud. Because the scattering efficiency of particles smaller than the wavelength of incident light (Rayleigh scattering) scales with the sixth power of the particle radius (r^6), the presence of even a very low number of cloud droplets (radius ~10μm) can overwhelm the signal of scattering aerosols (radius < 1μm) (J. L. Zhang et al., 2005). Separation of the signals of cloud droplets and aerosol particles has been attempted but the only practical results have concerned regimes where a discrete aerosol layer sits above a layer of optically opaque cloud (Jethva et al., 2016); separation of cloud and aerosol particle scattering using multispectral data is very challenging. Remotely sensed data is arbitrarily discretized in space such that condensed water particles can be present in any quantity; thus cloud clearing mechanisms have employed textural and contextual tests to exclude cloud droplets (Remer et al., 2005). These tests, to the extent that they rely on textural information in the visible scattering regime, also systematically remove heavy aerosol and smaller aerosol plumes (Levy et al., 2013), as can be seen in Figure 4.

This creates an asymmetric problem when assimilating strong aerosol plumes: if the model background is too high, data assimilation of observations adjacent to the core of the plume will act to reduce this error. If the model background is low, the results of data assimilation will necessarily also be low because of the exclusion of the plume core (Dai, Cheng, Suzuki, et al., 2019). This problem can be mitigated with the use of an ensemble that

includes uncertainty in the source magnitude (Rubin et al., 2016), and better results will be obtained with the use of retrievals that are able to retain more of the thick plumes (MAIAC and VIIRS NOAA Enterprise on Figure 4).

4.2.4 Observations do not constrain aerosol type or vertical structure

The atmospheric aerosol is a wildly heterogeneous mixture of particles of varying size, shape, and composition, in varying equilibria between water and chemical components in gas, liquid, and solid phases. Numerical models of aerosol reduce this complexity enormously, attempting to represent the aerosol properties that most impact targeted atmospheric outcomes such as climate forcing (Stier et al., 2013).

Optical retrieval of aerosol is underdetermined in different ways. Information can be added to the optical retrieval by incorporation of multiple wavelengths, viewing geometries, and polarization angles. Advanced methods have shown promise to obtain detailed aerosol composition information (K. J. Noyes et al., 2020) as well information on plume height (Lyapustin et al., 2020) and combustion phase (J. Wang, S. Roudini, et al., 2020) for certain fires. However, observing in all of these dimensions together still captures only a small part of the variability in the real atmosphere (Choi et al., 2021). Operational retrievals available for smoke forecasting provide a moderately uncertain estimate of column-integrated aerosol optical depth, and do not effectively constrain any other properties (Reid et al., 2013).

Some data assimilation systems simply place the speciation and vertical profile outside of the data assimilation solution itself. By transforming the model background to AOD and solving for an increment in AOD space, the critical step is applying the AOD increment to obtain a new 3D initial state for each simulated aerosol type. Simply conserving the prior was recognized early on to be inadequate because the model deficiencies corrected by assimilation of AOD project onto the speciation and vertical distribution—heavy aerosol events are unlikely to be realistically represented by linear scaling of a background aerosol. In systems that solve for a single AOD increment, a climatology is employed to more realistically distribute added aerosol in cases where the model background is strongly biased low (J. L. Zhang, J. S. Reid, D. L. Westphal, N. L. Baker, & E. J. Hyer, 2008). This problem manifests differently in systems where an increment is calculated for each aerosol type, but the underlying problem remains that large increments of aerosol loading require accounting for patterns of speciation and vertical placement at high and low aerosol loadings. For models resolving aerosol size distribution an intermediate solution involves keeping speciation fixed within each aerosol size bin. This approach reduces the degrees of freedom but still allows changes on the overall speciation (Pablo E. Saide et al., 2013).

5 Future Research Directions

The content of this chapter covered the observations available to improve smoke predictions via data assimilation, the characteristics that determine usefulness of those observations, the scenarios in which smoke predictions can benefit from data assimilation, and the interactions between the particular problems of smoke and smoke observations in data assimilation. The scientific territory of this topic is only beginning to be explored, and significant practical gains in smoke forecast skill are likely achievable over current practice. Smoke observational datasets are continuously being added and improved, with significant

and rapid innovation. Advanced data assimilation methods have been developed that are relevant to the smoke prediction problem, but have not yet been applied. Techniques that maximize the effectiveness of combining multiple observation types (e.g. (Pablo E. Saide et al., 2013; Schwartz et al., 2012) (Pablo E. Saide et al., 2015)) need to be advanced into operational smoke prediction systems. In all of these areas, careful attention to both the scientific details of the observations, as well as the practical realities of the forecast problem is necessary to obtain significant progress.

6 Sources for Data Products

The data in Figure 4 was obtained from the following public website sources:

- MODIS RGB from NASA Worldview (http://worldview.earthdata.nasa.gov/)
- NOAA Enterprise VIIRS AOD can be viewed at NOAA JPSS STAR Mapper: https://www.star.nesdis.noaa.gov/jpss/mapper
- MODIS MAIAC, Dark Target / Deep Blue combined, and NRL/UND L3 AOD can be viewed at NASA Worldview (NRL/UND L3 is labeled as "MODIS Combined Value-Added Aerosol Optical Depth")
- NOAA Enterprise GOES ABI AOD can be viewed at NOAA Aerosol Watch: https://www.star.nesdis.noaa.gov/smcd/spb/aq/AerosolWatch/

Model predictions of dust, smoke, and other aerosols from the Navy Aerosol Analysis and Prediction System can be viewed at https://www.nrlmry.navy.mil/aerosol/

Data download of NASA products is available through the Earth Science Data and Information Service: http://earthdata.nasa.gov/

Download of NOAA products is available through the NOAA Comprehensive Large Array-data Stewardship System (CLASS) system: http://www.class.noaa.gov/

AirNow data are available through AirNowTech http://airnowtech.org/

AERONET data are available from NASA: http://aeronet.gsfc.nasa.gov/

7 Acknowledgements

This chapter was written with support from the Office of Naval Research code 322 as well as the NASA FIREX-AQ program.

8 References

9 Acronyms

ABI	Advanced Baseline Imager
AERONET	Aerosol Robotic Network
AOD	Aerosol Optical Depth
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CDF	Cumulative Distribution Function
CONUS	Continental United States
CrIS	Cross-Track Infrared Spectrometer
FRP	Fire Radiative Power
GA	Gaussian Anamorphosis
IASI	Infrared Atmospheric Sounding Interferometer
MAIAC	Multi-Angle Implementation of Atmospheric Correction
MISR	Multi-Angle Imaging Spectroradiometer
MODIS	Moderate-Resolution Imaging Spectroradiometer
MOPITT	Measurement of Pollution in the Troposphere
MPLNet	Micropulse Lidar Network
NAAPS	Navy Aerosol Analysis and Prediction System
NRL/UND	Naval Research Laboratory / University of North Dakota
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapping Profiler Suite
PBL	Planetary Boundary Layer
PDF	Probability Distribution Function
PM	Particulate matter
PM10	Particulate matter with a diameter of 10 microns or less
PM2.5	Particulate matter with a diameter of 2.5 microns or less
POLDER	Polarization and Directionality of the Earth's Reflectances
PRH	Pseudo Relative Humidity
TROPOMI	Tropospheric Monitoring Instrument
VIIRS	Visible Infrared Imaging Radiometer Suite

10 Figures and Table

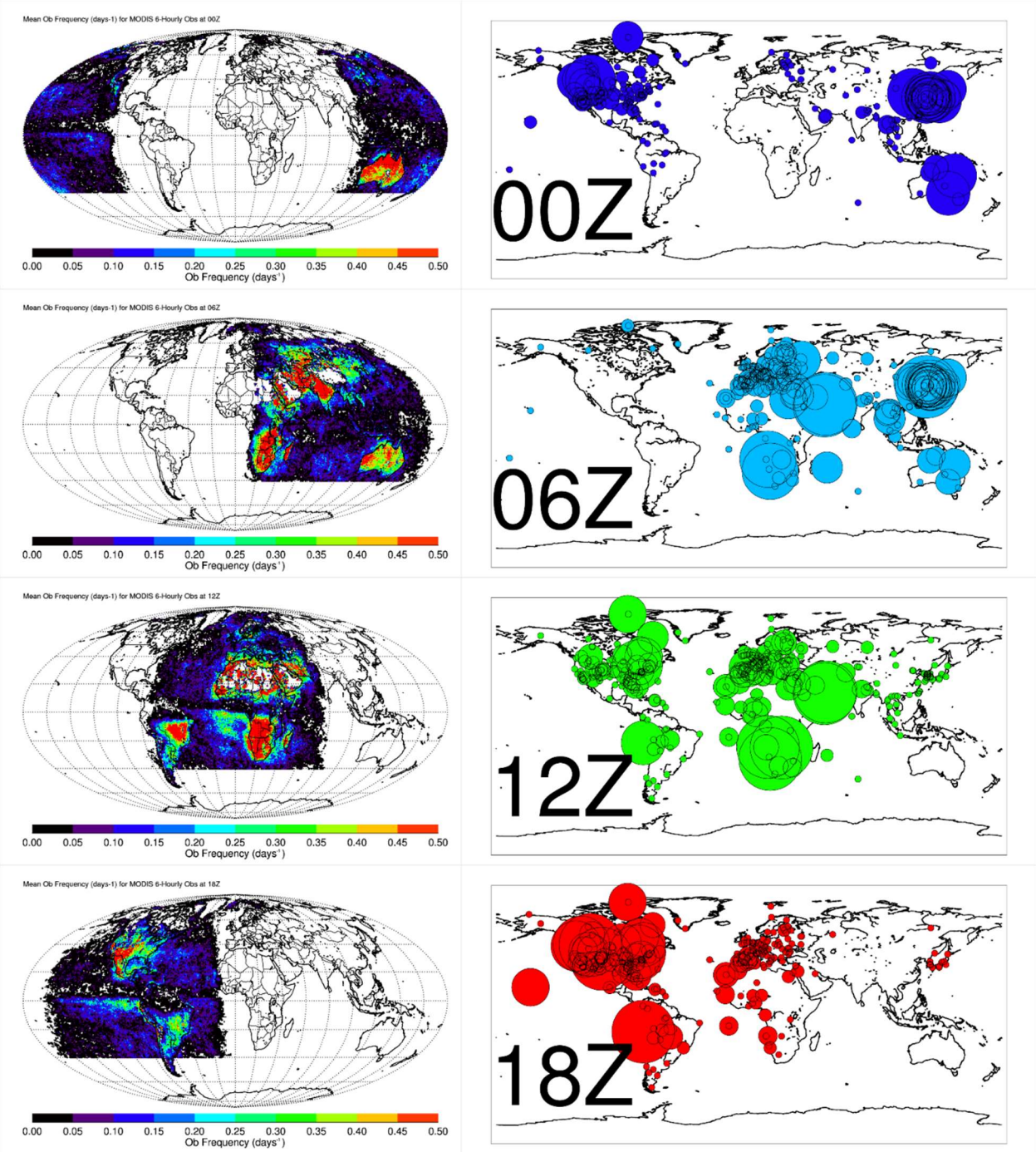

Figure 1. ***Density of AOD observations at synoptic times 00Z, 06Z, 12Z, 18Z. (Left): MODIS NRL/UND Level 3 observations assimilated into the NAAPS model; Colored grid cells indicate density of observations. (Right) AERONET AOD retrievals used for model verification; Symbols indicate locations of AERONET sites, size of circles is proportional to data volume over the study period of 1 May 2016 to 18 June 2016.***

Table 2. ***Availability of MODIS AOD for assimilation vs times/locations of AERONET forecast verification observations. Blue color indicates overlap between MODIS at analysis time and AERONET observations at 24-hour, 48-hour, 72-hour, etc. forecast lead times. Red color indicates overlap between MODIS at analysis time and AERONET at 12-hour, 36-hour, 60-hour, etc. forecast lead times. Green colors indicate availability of MODIS AOD observations at the times and locations of AERONET observations used to verify the NAAPS analysis.***

Analysis Time	AERONET observation time 21Z-3Z	3Z-9Z	9Z-15Z	15Z-18Z
00Z	16%	10%	1%	2%
06Z	18%	20%	9%	1%
12Z	0%	20%	25%	14%
18Z	12%	0%	9%	16%
PAIRED	26%	28%	28%	25%

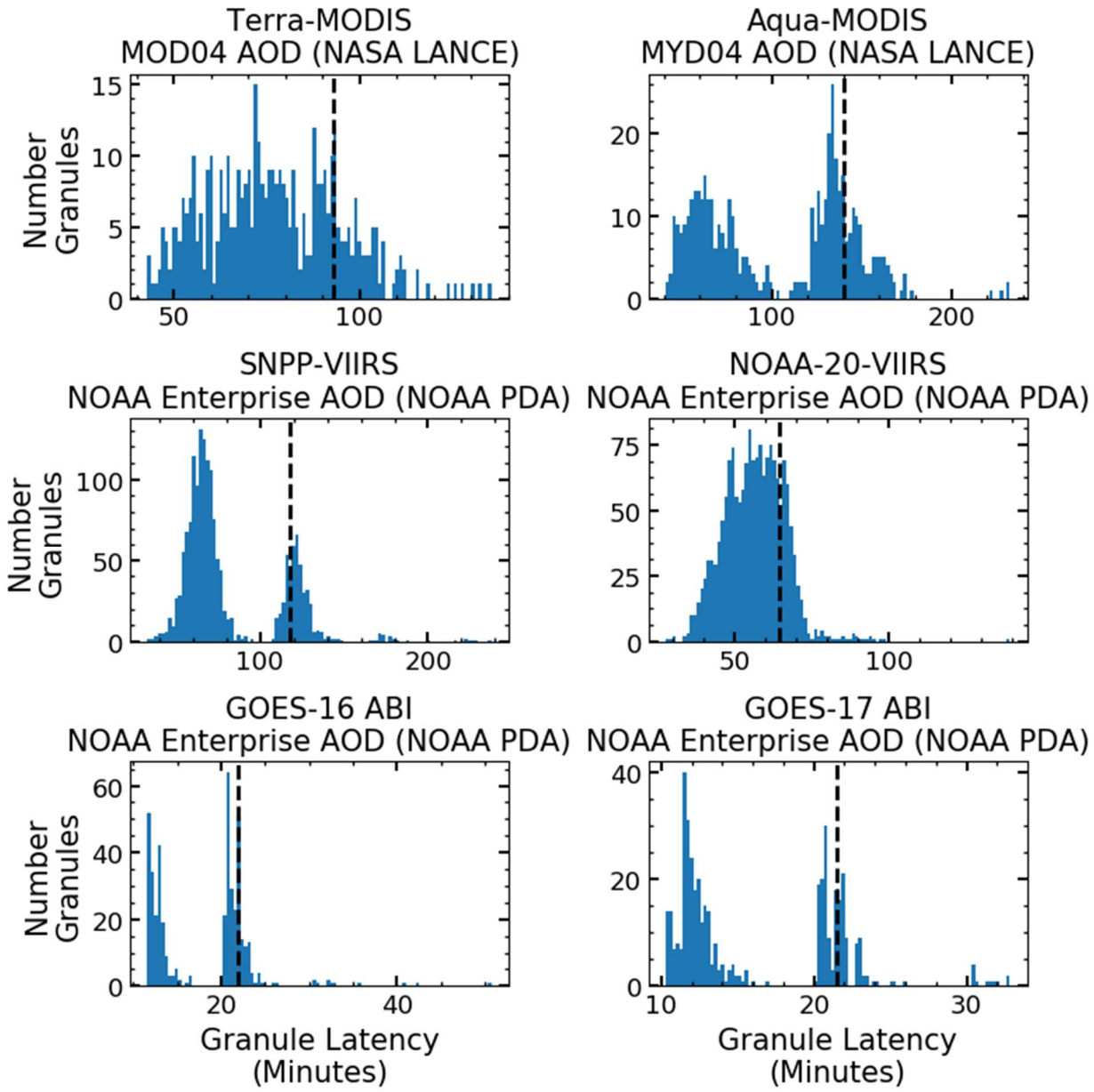

Figure 2. ***Latency of satellite aerosol products. Each plot shows the time gap between the observation time and the time the observation files are available on the NRL filesystem for use by the model. The dashed vertical line on each plot represents the 80th percentile of data latency. Details of the observations are in the text, plots are based on May 2021 data.***

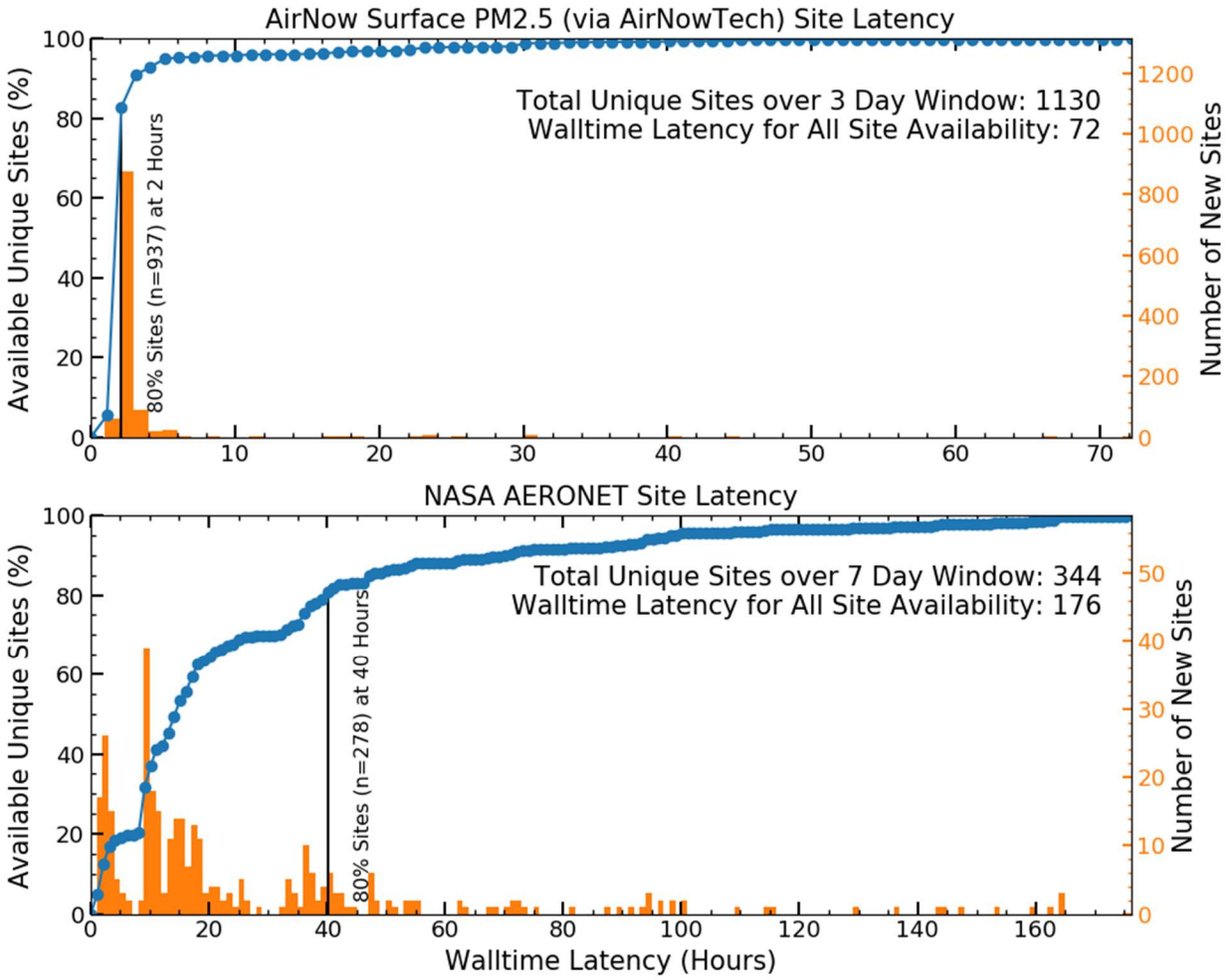

Figure 3. ***Latency of ground-based datasets. These plots show the cumulative availability of data from different stations as a function of delay. Vertical lines indicate the delay associated with 80% of the total sites reporting. Details of the observations are in the text, plots are based on May 2021 data.***

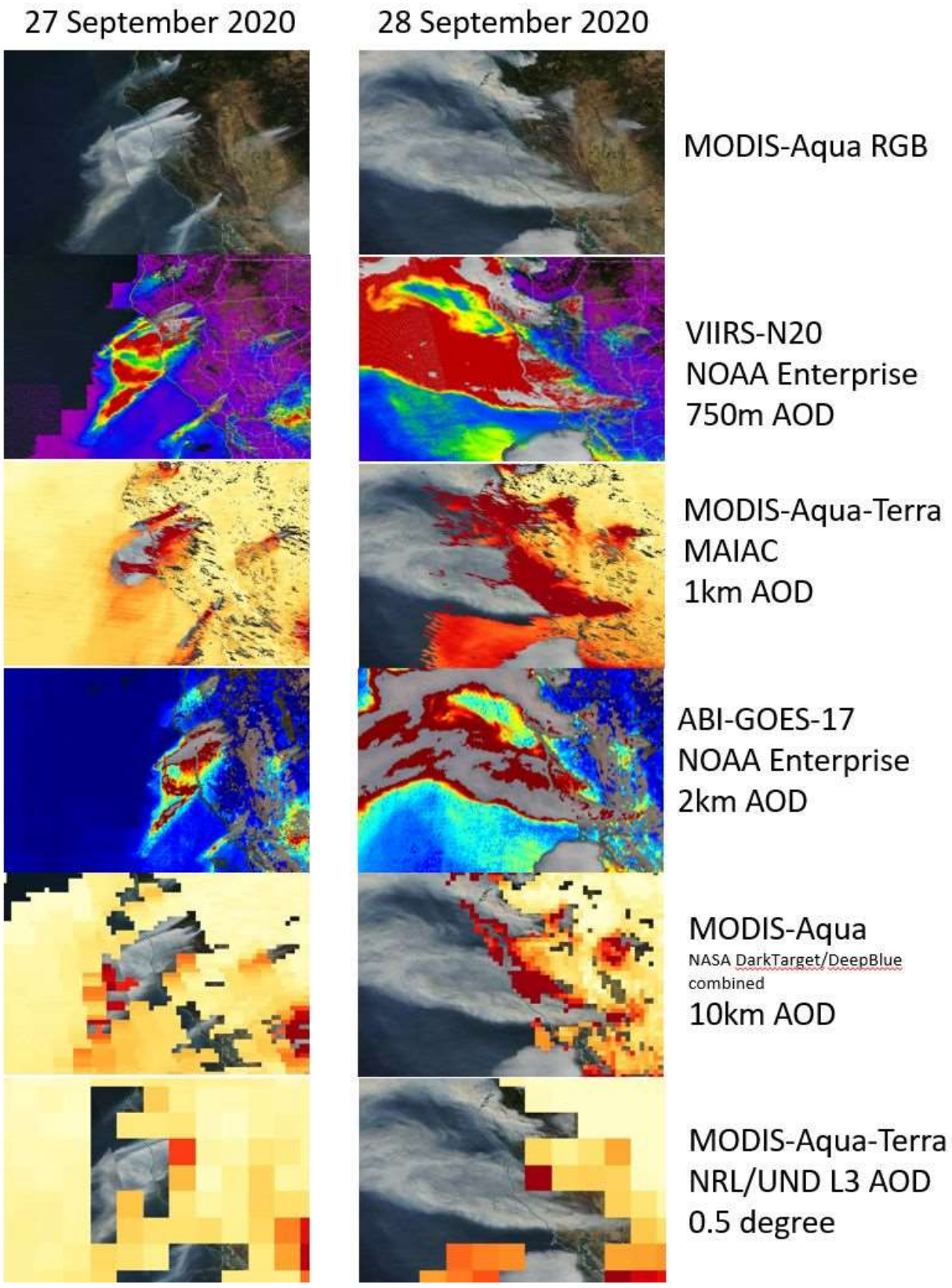

Figure 4. ***Retrieval of smoke by different AOD products for two days in September 2020. Datasets are described in the text. Dark red colors indicate the highest AOD; lower AODs are indicated by blue or beige colors on the maps.***

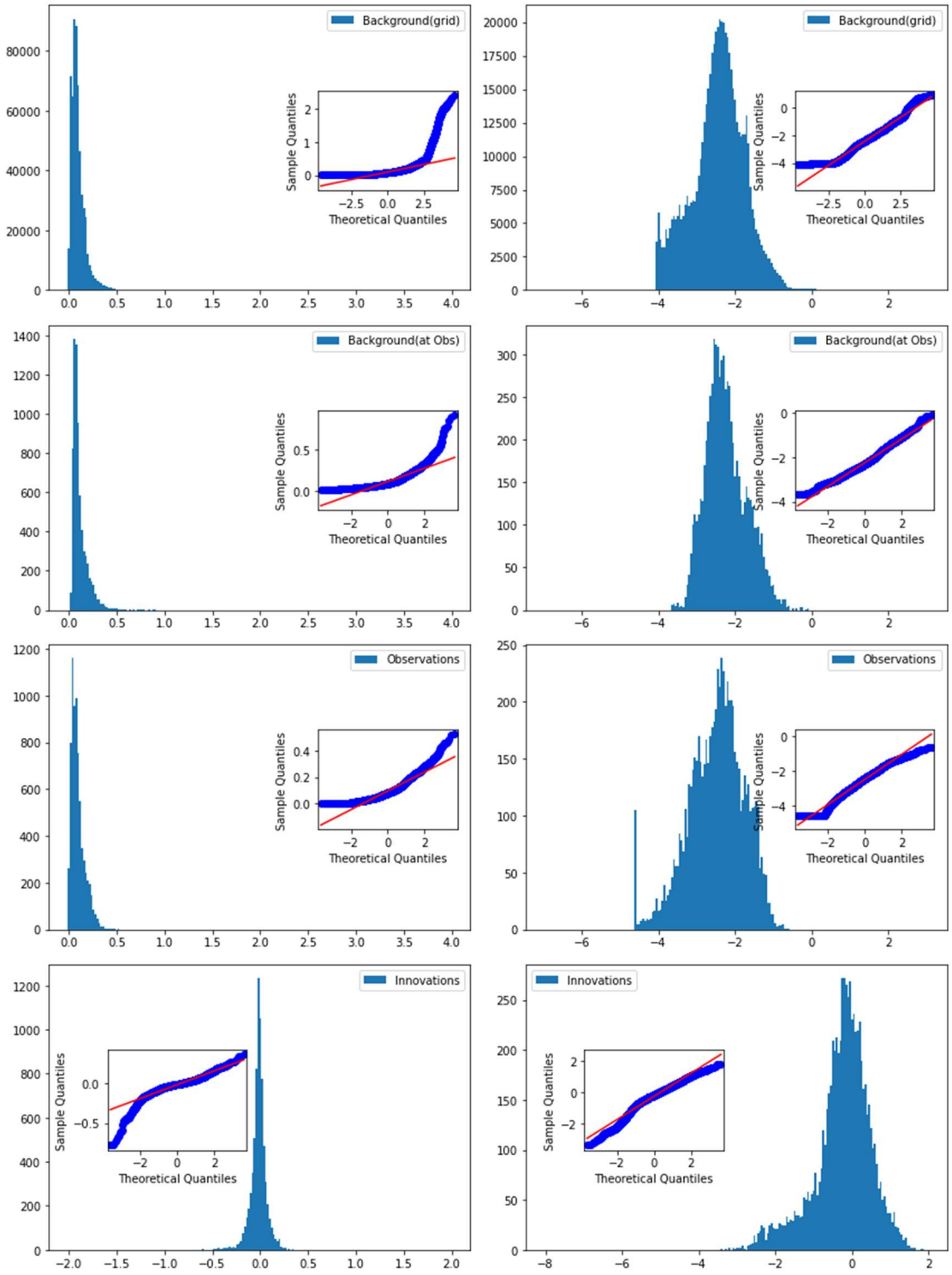

Figure 5. ***Distribution of modeled and observed AOD values, and effects of a logarithmic transformation. The left-hand side shows the distribution of AOD from (top to bottom) the NAAPS model grid, NAAPS sampled to observation locations, MODIS DA-quality observations, and lastly the innovations, computed as the differences Observation – Background. The inset plots on each histogram are the quantile-quantile plots indicating how well each distribution conforms to a Gaussian shape. The right-hand column shows the same values with the AOD'=ln(AOD+0.01) transform applied (see details in the text).***

Al-Saadi, J., Soja, A., Pierce, R. B., Szykman, J., Wiedinmyer, C., Emmons, L., . . . Bowman, K. (2008). Intercomparison of near-real-time biomass burning emissions estimates constrained by satellite fire data. *Journal Of Applied Remote Sensing, 2*.

Alexandrov, M. D., Geogdzhayev, I. V., Tsigaridis, K., Marshak, A., Levy, R., & Cairns, B. (2016). New Statistical Model for Variability of Aerosol Optical Thickness: Theory and Application to MODIS Data over Ocean. *Journal Of The Atmospheric Sciences, 73*(2), 821-837. doi:10.1175/jas-d-15-0130.1

Alfaro-Contreras, R., Zhang, J. L., Campbell, J. R., Holz, R. E., & Reid, J. S. (2014). Evaluating the impact of aerosol particles above cloud on cloud optical depth retrievals from MODIS. *Journal Of Geophysical Research-Atmospheres, 119*(9), 5410-5423. doi:10.1002/2013jd021270

Amezcua, J., & Van Leeuwen, P. J. (2014). Gaussian anamorphosis in the analysis step of the EnKF: a joint state-variable/observation approach. *Tellus Series a-Dynamic Meteorology and Oceanography, 66*, 1-18. doi:10.3402/tellusa.v66.23493

Arellano, A. F., Hess, P. G., Edwards, D. P., & Baumgardner, D. (2010). Constraints on black carbon aerosol distribution from Measurement of Pollution in the Troposphere (MOPITT) CO. *Geophysical Research Letters, 37*, 5. doi:10.1029/2010gl044416

Bannister, R. N., Chipilski, H. G., & Martinez-Alvarado, O. (2020). Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales. *Quarterly Journal Of The Royal Meteorological Society, 146*(726), 1-48. doi:10.1002/qj.3652

Benedetti, A., Morcrette, J. J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., . . . Suttie, M. (2009). Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation. *Journal Of Geophysical Research-Atmospheres, 114*. doi:D13205

10.1029/2008jd011115

Bertino, L., Evensen, G., & Wackernagel, H. (2003). Sequential data assimilation techniques in oceanography. *International Statistical Review, 71*(2), 223-241. doi:10.1111/j.1751-5823.2003.tb00194.x

Bishop, C. H. (2016). The GIGG-EnKF: ensemble Kalman filtering for highly skewed non-negative uncertainty distributions. *Quarterly Journal Of The Royal Meteorological Society, 142*(696), 1395-1412. doi:10.1002/qj.2742

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Zabkar, R., Carmichael, G. R., . . . Seigneur, C. (2015). Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models. *Atmospheric Chemistry And Physics, 15*(10), 5325-5358. doi:10.5194/acp-15-5325-2015

Bocquet, M., Pires, C. A., & Wu, L. (2010). Beyond Gaussian Statistical Modeling in Geophysical Data Assimilation. *Monthly Weather Review, 138*(8), 2997-3023. doi:10.1175/2010mwr3164.1

Brankart, J. M., Testut, C. E., Beal, D., Doron, M., Fontana, C., Meinvielle, M., . . . Verron, J. (2012). Towards an improved description of ocean uncertainties: effect of local anamorphic transformations on spatial correlations. *Ocean Science, 8*(2), 121-142. doi:10.5194/os-8-121-2012

Carr, J. L., Wu, D. L., Daniels, J., Friberg, M. D., Bresky, W., & Madani, H. (2020). GEO-GEO Stereo-Tracking of Atmospheric Motion Vectors (AMVs) from the Geostationary Ring. *Remote Sensing, 12*(22), 47. doi:10.3390/rs12223779

Chang, W. Y., Zhang, Y., Li, Z. Q., Chen, J., & Li, K. T. (2021). Improving the sectional Model for Simulating Aerosol Interactions and Chemistry (MOSAIC) aerosols of the Weather Research and Forecasting-Chemistry (WRF-Chem) model with the revised Gridpoint Statistical Interpolation system and multi-wavelength aerosol optical measurements: the dust aerosol observation campaign at Kashi, near the Taklimakan Desert, northwestern China. *Atmospheric Chemistry And Physics, 21*(6), 4403-4430. doi:10.5194/acp-21-4403-2021

Cheng, Y. M., Dai, T., Goto, D., Schutgens, N. A. J., Shi, G. Y., & Nakajima, T. (2019). Investigating the assimilation of CALIPSO global aerosol vertical observations using a four-dimensional ensemble Kalman filter. *Atmospheric Chemistry And Physics, 19*(21), 13445-13467. doi:10.5194/acp-19-13445-2019

Choi, M., Sander, S. P., Spurr, R. J. D., Pongetti, T. J., van Harten, G., Drouin, B. J., . . . Fu, D. J. (2021). Aerosol profiling using radiometric and polarimetric spectral measurements in the O-2 near infrared bands: Estimation of information content and measurement uncertainties. *Remote Sensing Of Environment, 253*, 20. doi:10.1016/j.rse.2020.112179

Clark, P. A., Harcourt, S. A., Macpherson, B., Mathison, C. T., Cusack, S., & Naylor, M. (2008). Prediction of visibility and aerosol within the operational Met Office Unified Model. I: Model formulation and variational assimilation. *Quarterly Journal Of The Royal Meteorological Society, 134*(636), 1801-1816. doi:10.1002/qj.318

Colarco, P. R., Schoeberl, M. R., Doddridge, B. G., Marufu, L. T., Torres, O., & Welton, E. J. (2004). Transport of smoke from Canadian forest fires to the surface near Washington, D. C.: Injection height, entrainment, and optical properties. *Journal Of Geophysical Research-Atmospheres, 109*(D6), D06203, doi:06210.01029/02003JD004248.

Dai, T., Cheng, Y. M., Goto, D., Schutgens, N. A. J., Kikuchi, M., Yoshida, M., . . . Nakajima, T. (2019). Inverting the East Asian Dust Emission Fluxes Using the Ensemble Kalman Smoother and Himawari-8 AODs: A Case Study with WRF-Chem v3.5.1. *Atmosphere, 10*(9). doi:10.3390/atmos10090543

Dai, T., Cheng, Y. M., Suzuki, K., Goto, D., Kikuchi, M., Schutgens, N. A. J., . . . Nakajima, T. (2019). Hourly Aerosol Assimilation of Himawari-8 AOT Using the Four-Dimensional Local Ensemble Transform Kalman Filter. *Journal of Advances in Modeling Earth Systems, 11*(3), 680-711. doi:10.1029/2018ms001475

Dai, T., Schutgens, N. A. J., & Nakajima, T. (2013). Applying a Local Ensemble Transform Kalman Filter Assimilation System to the NICAM-SPRINTARS Model. In R. F. Cahalan & J. Fischer (Eds.), *Radiation Processes in the Atmosphere and Ocean* (Vol. 1531, pp. 744-747).

Dee, D. P., & Da Silva, A. M. (2003). The choice of variable for atmospheric moisture analysis. *Monthly Weather Review, 131*(1), 155-171. doi:10.1175/1520-0493(2003)131<0155:tcovfa>2.0.co;2

Delp, W. W., & Singer, B. C. (2020). Wildfire Smoke Adjustment Factors for Low-Cost and Professional PM(2.5)Monitors with Optical Sensors. *Sensors, 20*(13), 21. doi:10.3390/s20133683

Dubovik, O., Lapyonok, T., Kaufman, Y. J., Chin, M., Ginoux, P., Kahn, R. A., & Sinyuk, A. (2008). Retrieving global aerosol sources from satellites using inverse modeling. *Atmospheric Chemistry And Physics, 8*(2), 209-250.

Eck, T. F., Holben, B. N., Giles, D. M., Slutsker, I., Sinyuk, A., Schafer, J. S., . . . Aldrian, E. (2019). AERONET Remotely Sensed Measurements and Retrievals of Biomass Burning Aerosol Optical Properties During the 2015 Indonesian Burning Season. *Journal Of Geophysical Research-Atmospheres, 124*(8), 4722-4740. doi:10.1029/2018jd030182

Fletcher, S. J., & Zupanski, M. (2006). A data assimilation method for log-normally distributed observational errors. *Quarterly Journal Of The Royal Meteorological Society, 132*(621), 2505-2519. doi:10.1256/qj.05.222

French, N. H. F., de Groot, W. J., Jenkins, L. K., Rogers, B. M., Alvarado, E., Amiro, B., . . . Turetsky, M. (2011). Model comparisons for estimating carbon emissions from North American wildland fire. *Journal Of Geophysical Research-Biogeosciences, 116*. doi:G00k05 10.1029/2010jg001469

Fu, D., Xia, X., Duan, M., Zhang, X., Li, X., Wang, J., & Liu, J. (2018). Mapping nighttime PM2.5 from VIIRS DNB using a linear mixed-effect model. *Atmospheric Environment, 178*, 214-222. doi:10.1016/j.atmosenv.2018.02.001

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., . . . Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal Of Climate, 30*(14), 5419-5454. doi:10.1175/jcli-d-16-0758.1

Generoso, S., Breon, F. M., Chevallier, F., Balkanski, Y., Schulz, M., & Bey, I. (2007). Assimilation of POLDER aerosol optical thickness into the LMDz-INCA model: Implications for the Arctic aerosol burden. *Journal Of Geophysical Research-Atmospheres, 112*(D2), 15. doi:10.1029/2005jd006954

Giglio, L. (2007). Characterization of the tropical diurnal fire cycle using VIRS and MODIS observations. *Remote Sensing Of Environment, 108*(4), 407-421.

Giglio, L., & Schroeder, W. (2014). A global feasibility assessment of the bi-spectral fire temperature and area retrieval using MODIS data. *Remote Sensing Of Environment, 152*(0), 166-173. doi:http://dx.doi.org/10.1016/j.rse.2014.06.010

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., . . . Lyapustin, A. I. (2019). Advancements in the Aerosol Robotic Network (AERONET) Version 3 database - automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. *Atmospheric Measurement Techniques, 12*(1), 169-209. doi:10.5194/amt-12-169-2019

Goto, D., Schutgens, N. A. J., Nakajima, T., & Takemura, T. (2011). Sensitivity of aerosol to assumed optical properties over Asia using a global aerosol model and AERONET. *Geophysical Research Letters, 38*. doi:10.1029/2011gl048675

Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., . . . Martin, R. V. (2020). Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998-2018). *Environmental Science & Technology, 54*(13), 7879-7890. doi:10.1021/acs.est.0c01764

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., . . . Smirnov, A. (1998). AERONET - A federated instrument network and data archive for aerosol characterization. *Remote Sensing Of Environment, 66*(1), 1-16.

Hólm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F., Simmons, A., & Thépaut, J.-N. (2002). Assimilation and modelling of the hydrological cycle: ECMWF's status and plans. *Technical Memoranda, 383*.

Hyer, E. J., & Reid, J. S. (2009). Baseline uncertainties in biomass burning emission models resulting from spatial error in satellite active fire location data. *Geophysical Research Letters, 36*, doi:10.1029/2008GL037145.

Hyer, E. J., Reid, J. S., Prins, E. M., Hoffman, J. P., Schmidt, C. C., Miettinen, J. I., & Giglio, L. (2013). Patterns of fire activity over Indonesia and Malaysia from polar and geostationary satellite observations. *Atmospheric Research, 122*, 504-519. doi:10.1016/j.atmosres.2012.06.011

Hyer, E. J., Reid, J. S., & Zhang, J. (2011). An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals. *Atmospheric Measurement Techniques, 4*, 379-408. Retrieved from doi:10.5194/amt-4-379-2011

Jethva, H., Torres, O., Remer, L., Redemann, J., Livingston, J., Dunagan, S., . . . Spurr, R. (2016). Validating MODIS above-cloud aerosol optical depth retrieved from "color ratio" algorithm using direct measurements made by NASA's airborne AATS and 4STAR sensors. *Atmospheric Measurement Techniques, 9*(10), 5053-5062. doi:10.5194/amt-9-5053-2016

Jethva, H., Torres, O., Waquet, F., Chand, D., & Hu, Y. X. (2014). How doA- train sensors intercompare in the retrieval of above- cloud aerosol optical depth? A case study- based assessment. *Geophysical Research Letters, 41*(1), 186-192. doi:10.1002/2013gl058405

Jianglong, Z., Reid, J. S., Westphal, D. L., Baker, N. L., & Hyer, E. J. (2008). A system for operational aerosol optical depth data assimilation over global oceans. *Journal of Geophysical Research - Part D - Atmospheres, 113*(D10). doi:10.1029/2007jd009065

Johnson, R. S., Zhang, J., Hyer, E. J., Miller, S. D., & Reid, J. S. (2013). Preliminary investigations toward nighttime aerosol optical depth retrievals from the VIIRS day/night band. *Atmospheric Measurement Techniques, 6*, 1245-1255. doi:doi:10.5194/amt-6-1245-2013

Kaku, K. C., Reid, J. S., Hand, J. L., Edgerton, E. S., Holben, B. N., Zhang, J., & Holz, R. E. (2018). Assessing the Challenges of Surface-Level Aerosol Mass Estimates From Remote Sensing During the SEAC(4)RS and SEARCH Campaigns: Baseline Surface Observations and Remote Sensing in the Southeastern United States. *Journal Of Geophysical Research-Atmospheres, 123*(14), 7530-7562. doi:10.1029/2017jd028074

Kaufman, Y. J., Tanre, D., Remer, L. A., Vermote, E. F., Chu, A., & Holben, B. N. (1997). Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. *Journal Of Geophysical Research-Atmospheres, 102*(D14), 17051-17067.

Kleinman, L. I., Sedlacek, A. J., Adachi, K., Buseck, P. R., Collier, S., Dubey, M. K., . . . Yokelson, R. J. (2020). Rapid evolution of aerosol particles and their optical properties downwind of wildfires in the western US. *Atmospheric Chemistry And Physics, 20*(21), 13319-13341. doi:10.5194/acp-20-13319-2020

Laszlo, I., & Liu, H. (2016). *NOAA EPS Aerosol Optical Depth (AOD) Algorithm Theoretical Basis Document, version 3.0.1*. Retrieved from https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_EPS_Aerosol_AOD_v3.0.1.pdf

Lee, E., Zupanski, M., Zupanski, D., & Park, S. K. (2017). Impact of the OMI aerosol optical depth on analysis increments through coupled meteorology-aerosol data assimilation for an Asian dust storm. *Remote Sensing Of Environment, 193*, 38-53. doi:10.1016/j.rse.2017.02.013

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., & Hsu, N. C. (2013). The Collection 6 MODIS aerosol products over land and ocean. *Atmospheric Measurement Techniques, 6*(11), 2989-3034. doi:10.5194/amt-6-2989-2013

Liang, Y. F., Zang, Z. L., Liu, D., Yan, P., Hu, Y. W., Zhou, Y., & You, W. (2020). Development of a three-dimensional variational assimilation system for lidar profile data based on a size-resolved aerosol model in WRF-Chem model v3.9.1 and its application in PM2.5 forecasts across China. *Geoscientific Model Development, 13*(12), 6285-6301. doi:10.5194/gmd-13-6285-2020

Lien, G. Y., Kalnay, E., & Miyoshi, T. (2013). Effective assimilation of global precipitation: simulation experiments. *Tellus Series a-Dynamic Meteorology and Oceanography, 65*. doi:10.3402/tellusa.v65i0.19915

Lien, G. Y., Kalnay, E., Miyoshi, T., & Huffman, G. J. (2016). Statistical Properties of Global Precipitation in the NCEP GFS Model and TMPA Observations for Data Assimilation. *Monthly Weather Review, 144*(2), 663-679. doi:10.1175/mwr-d-15-0150.1

Lin, H. D., Weygandt, S. S., Benjamin, S. G., & Hu, M. (2017). Satellite Radiance Data Assimilation within the Hourly Updated Rapid Refresh. *Weather And Forecasting, 32*(4), 1273-1287. doi:10.1175/waf-d-16-0215.1

Liu, Y. Q., Kochanski, A., Baker, K. R., Mell, W., Linn, R., Paugam, R., . . . McNamara, D. (2019). Fire behaviour and smoke modelling: model improvement and measurement needs for next-generation smoke research and forecasting systems. *International Journal Of Wildland Fire, 28*(8), 570-588. doi:10.1071/wf18204

Liu, Z. Q., Liu, Q. H., Lin, H. C., Schwartz, C. S., Lee, Y. H., & Wang, T. J. (2011). Three-dimensional variational assimilation of MODIS aerosol optical depth: Implementation and application to a dust storm over East Asia. *Journal Of Geophysical Research-Atmospheres, 116*. doi:D23206

10.1029/2011jd016159

Lyapustin, A., Korkin, S., Wang, Y., Quayle, B., & Laszlo, I. (2012). Discrimination of biomass burning smoke and clouds in MAIAC algorithm. *Atmospheric Chemistry And Physics, 12*(20), 9679-9686. doi:10.5194/acp-12-9679-2012

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., . . . Reid, J. S. (2011). Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *Journal Of Geophysical Research-Atmospheres, 116*, 15. doi:10.1029/2010jd014986

Lyapustin, A., Wang, Y. J., Korkin, S., Kahn, R., & Winker, D. (2020). MAIAC Thermal Technique for Smoke Injection Height From MODIS. *Ieee Geoscience And Remote Sensing Letters, 17*(5), 730-734. doi:10.1109/lgrs.2019.2936332

Lynch, P., Reid, J. S., Westphal, D. L., Zhang, J. L., Hogan, T. F., Hyer, E. J., . . . Walker, A. L. (2016). An 11-year global gridded aerosol optical thickness reanalysis (v1.0) for atmospheric and climate sciences. *Geoscientific Model Development, 9*(4), 1489-1522. doi:10.5194/gmd-9-1489-2016

Ma, C. Q., Wang, T. J., Mizzi, A. P., Anderson, J. L., Zhuang, B. L., Xie, M., & Wu, R. S. (2019). Multiconstituent Data Assimilation With WRF-Chem/DART: Potential for Adjusting Anthropogenic Emissions and Improving Air Quality Forecasts Over Eastern China. *Journal Of Geophysical Research-Atmospheres, 124*(13), 7393-7412. doi:10.1029/2019jd030421

McHardy, T., Zhang, J., Reid, J. S., Miller, S. D., Hyer, E. J., & Kuehn, R. E. (2015). An improved method for retrieving nighttime aerosol optical thickness from the VIIRS Day/Night Band. *Atmospheric Measurement Techniques, 8*, 4773-4783.

Mu, M., Randerson, J., van der Werf, G., Giglio, L., Kasibhatla, P., Morton, D., . . . Wennberg, P. O. (2011). Daily and hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide. *Journal of Geophysical Research*. Retrieved from doi:10.1029/2011JD016245

Mu, M., Randerson, J. T., van der Werf, G. R., Giglio, L., Kasibhatla, P., Morton, D., . . . Wennberg, P. O. (2011). Daily and 3-hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide. *Journal Of Geophysical Research-Atmospheres, 116*. doi:D24303

10.1029/2011jd016245

Nelson, R. M., Butler, B. W., & Weise, D. R. (2012). Entrainment regimes and flame characteristics of wildland fires. *International Journal Of Wildland Fire, 21*(2), 127-140. doi:10.1071/wf10034

Noyes, K. J., Kahn, R., Sedlacek, A., Kleinman, L., Limbacher, J., & Li, Z. Q. (2020). Wildfire Smoke Particle Properties and Evolution, from Space-Based Multi-Angle Imaging. *Remote Sensing, 12*(5), 23. doi:10.3390/rs12050769

Noyes, K. T. J., Kahn, R. A., Limbacher, J. A., Li, Z. Q., Fenn, M. A., Giles, D. M., . . . Winstead, E. L. (2020). Wildfire Smoke Particle Properties and Evolution, From Space-Based Multi-Angle Imaging II: The Williams Flats Fire during the FIREX-AQ Campaign. *Remote Sensing, 12*(22), 27. doi:10.3390/rs12223823

O'Neill, N. T., Ignatov, A., Holben, B. N., & Eck, T. F. (2000). The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET sunphotometer data. *Geophysical Research Letters, 27*(20), 3333-3336.

Paugam, R., Wooster, M., Freitas, S., & Martin, M. V. (2016). A review of approaches to estimate wildfire plume injection height within large-scale atmospheric chemical transport models. *Atmospheric Chemistry And Physics, 16*(2), 907-925. doi:10.5194/acp-16-907-2016

Peterson, D., Hyer, E., & Wang, J. (2013). A short-term predictor of satellite-observed fire activity in the North American boreal forest: Toward improving the prediction of smoke emissions. *Atmospheric Environment, 71*, 304-310. doi:10.1016/j.atmosenv.2013.01.052

Peterson, D., Hyer, E., & Wang, J. (2014). Quantifying the potential for high-altitude smoke injection in the North American boreal forest using the standard MODIS fire products and subpixel-based methods. *Journal Of Geophysical Research-Atmospheres, 119*(6), 3401-3419. doi:10.1002/2013jd021067

Peterson, D. A., Hyer, E. J., Campbell, J. R., Fromm, M. D., Hair, J. W., Butler, C. F., & Fenn, M. A. (2015). THE 2013 RIM FIRE Implications for Predicting Extreme Fire Spread, Pyroconvection, and Smoke Emissions. *Bulletin Of The American Meteorological Society, 96*(2), 229-247. doi:10.1175/bams-d-14-00060.1

Polivka, T. N., Wang, J., Ellison, L. T., Hyer, E. J., & Ichoku, C. M. (2016). Improving Nocturnal Fire Detection With the VIIRS Day-Night Band. *Ieee Transactions On Geoscience And Remote Sensing, 54*(9), 5503-5519. doi:10.1109/tgrs.2016.2566665

Potter, B. E. (2012). Atmospheric interactions with wildland fire behaviour - I. Basic surface interactions, vertical profiles and synoptic structures. *International Journal Of Wildland Fire, 21*(7), 779-801. doi:10.1071/wf11128

Randles, C. A., da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., . . . Flynn, C. J. (2017). The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation. *Journal Of Climate, 30*(17), 6823-6850. doi:10.1175/jcli-d-16-0609.1

Reid, J. S., Hyer, E. J., Johnson, R. S., Holben, B. N., Yokelson, R. J., Zhang, J. L., . . . Liew, S. C. (2013). Observing and understanding the Southeast Asian aerosol system by remote sensing: An initial review and analysis for the Seven Southeast Asian Studies (7SEAS) program. *Atmospheric Research, 122*, 403-468. doi:10.1016/j.atmosres.2012.06.005

Reid, J. S., Prins, E. M., Westphal, D. L., Schmidt, C. C., Richardson, K. A., Christopher, S. A., . . . Hoffman, J. P. (2004a). Real-time monitoring of South American smoke particle emissions and transport using a coupled remote sensing/box-model approach. *Geophysical Research Letters, 31*(6), 5. doi:10.1029/2003gl018845

Reid, J. S., Prins, E. M., Westphal, D. L., Schmidt, C. C., Richardson, K. A., Christopher, S. A., . . . Hoffman, J. P. (2004b). Real-time monitoring of South American smoke particle emissions and transport using a coupled remote sensing/box-model approach. *Geophysical Research Letters, 31*(6), doi:10.1029/2003GL018845.

Remer, L. A., Kaufman, Y. J., Tanre, D., Mattoo, S., Chu, D. A., Martins, J. V., . . . Holben, B. N. (2005). The MODIS aerosol algorithm, products, and validation. *Journal Of The Atmospheric Sciences, 62*(4), 947-973.

Roberts, G., Wooster, M. J., Xu, W., Freeborn, P. H., Morcrette, J. J., Jones, L., . . . Kaiser, J. W. (2015). LSA SAF Meteosat FRP products - Part 2: Evaluation and demonstration for use in the Copernicus Atmosphere Monitoring Service (CAMS). *Atmospheric Chemistry And Physics, 15*(22), 13241-13267. doi:10.5194/acp-15-13241-2015

Robinson, D. L. (2020). Accurate, Low Cost PM(2.5)Measurements Demonstrate the Large Spatial Variation in Wood Smoke Pollution in Regional Australia and Improve Modeling and Estimates of Health Costs. *Atmosphere, 11*(8), 21. doi:10.3390/atmos11080856

Rogozovsky, I., Ansmann, A., Althausen, D., Heese, B., Engelmann, R., Hofer, J., . . . Chudnovsky, A. (2021). Impact of aerosol layering, complex aerosol mixing, and cloud coverage on high-resolution MAIAC aerosol optical depth measurements: Fusion of lidar, AERONET, satellite, and ground-based measurements. *Atmospheric Environment, 247*, 12. doi:10.1016/j.atmosenv.2020.118163

Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Collins, N., Hoar, T. J., . . . Zhang, J. L. (2016). Development of the Ensemble Navy Aerosol Analysis Prediction System (ENAAPS) and its application of the Data Assimilation Research Testbed (DART) in support of aerosol forecasting. *Atmospheric Chemistry And Physics, 16*(6), 3927-3951.

Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Holben, B. N., Xian, P., . . . Zhang, J. L. (2017). Assimilation of AERONET and MODIS AOT observations using variational and ensemble data assimilation methods and its impact on aerosol forecasting skill. *Journal Of Geophysical Research-Atmospheres, 122*(9), 4967-4992. doi:10.1002/2016jd026067

Saide, P. E., Carmichael, G. R., Liu, Z., Schwartz, C. S., Lin, H. C., Da Silva, A. M., & Hyer, E. J. (2013). Aerosol optical depth assimilation for a size-resolved sectional model: impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale forecasts. *Atmospheric Chemistry And Physics, 13*, 10425-10444. doi:doi:10.5194/acp-13-10425-2013

Saide, P. E., Carmichael, G. R., Spak, S. N., Minnis, P., & Ayers, J. K. (2012). Improving aerosol distributions below clouds by assimilating satellite-retrieved cloud droplet number. *Proceedings Of The National Academy Of Sciences Of The United States Of America, 109*(30), 11939-11943. doi:10.1073/pnas.1205877109

Saide, P. E., Kim, J., Song, C. H., Choi, M., Cheng, Y., & Carmichael, G. R. (2014). Assimilation of next generation geostationary aerosol optical depth retrievals to improve air quality simulations. *Geophysical Research Letters, 41*(24), 9188-9196. doi:10.1002/2014gl062089

Saide, P. E., Peterson, D. A., da Silva, A., Anderson, B., Ziemba, L. D., Diskin, G., . . . Carmichael, G. R. (2015). Revealing important nocturnal and day-to-day variations in fire smoke emissions through a multiplatform inversion. *Geophysical Research Letters, 42*(9), 3609-3618. doi:10.1002/2015gl063737

Saide, P. E., Spak, S. N., Pierce, R. B., Otkin, J. A., Schaack, T. K., Heidinger, A. K., . . . Carmichael, G. R. (2015). Central American biomass burning smoke can increase tornado severity in the US. *Geophysical Research Letters, 42*(3), 956-965. doi:10.1002/2014gl062826

Sayer, A. M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., . . . Witek, M. L. (2020). A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing. *Atmospheric Measurement Techniques, 13*(2), 373-404. doi:10.5194/amt-13-373-2020

Sayer, A. M., & Knobelspiesse, K. D. (2019). How should we aggregate data? Methods accounting for the numerical distributions, with an assessment of aerosol optical depth. *Atmospheric Chemistry And Physics, 19*(23), 15007-15032. doi:10.5194/acp-19-15023-2019

Schwartz, C. S., Liu, Z. Q., Lin, H. C., & McKeen, S. A. (2012). Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth. *Journal Of Geophysical Research-Atmospheres, 117*. doi:D13202

10.1029/2011jd017383
Selimovic, V., Yokelson, R. J., McMeeking, G. R., & Coefield, S. (2019). In situ measurements of trace gases, PM, and aerosol optical properties during the 2017 NW US wildfire smoke event. *Atmospheric Chemistry And Physics, 19*(6), 3905-3926. doi:10.5194/acp-19-3905-2019
Shi, Y., Zhang, J., Reid, J. S., Holben, B., Hyer, E. J., & Curtis, C. (2011). An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation. *Atmospheric Chemistry And Physics, 11*(2), 557-565. doi:10.5194/acp-11-557-2011
Shi, Y. X. R., Levy, R. C., Eck, T. F., Fisher, B., Mattoo, S., Remer, L. A., . . . Zhang, J. L. (2019). Characterizing the 2015 Indonesia fire event using modified MODIS aerosol retrievals. *Atmospheric Chemistry And Physics, 19*(1), 259-274. doi:10.5194/acp-19-259-2019
Simon, E., & Bertino, L. (2012). Gaussian anamorphosis extension of the DEnKF for combined state parameter estimation: Application to a 1D ocean ecosystem model. *Journal of Marine Systems, 89*(1), 1-18. doi:10.1016/j.jmarsys.2011.07.007
Stier, P., Schutgens, N. A. J., Bellouin, N., Bian, H., Boucher, O., Chin, M., . . . Zhou, C. (2013). Host model uncertainties in aerosol radiative forcing estimates: results from the AeroCom Prescribed intercomparison study. *Atmospheric Chemistry And Physics, 13*(6), 3245-3270. doi:10.5194/acp-13-3245-2013
Tang, Y. H., Pagowski, M., Chai, T. F., Pan, L., Lee, P., Baker, B., . . . Kim, H. C. (2017). A case study of aerosol data assimilation with the Community Multi-scale Air Quality Model over the contiguous United States using 3D-Var and optimal interpolation methods. *Geoscientific Model Development, 10*(12), 4743-4758. doi:10.5194/gmd-10-4743-2017
Val Martin, M., Kahn, R. A., Logan, J. A., Paugam, R., Wooster, M., & Ichoku, C. (2012). Space-based observational constraints for 1-D fire smoke plume-rise models. *Journal of Geophysical Research, 117*(D22), 27. doi:DOI: 10.1029/2012JD018370
van Leeuwen, P. J. (2009). Particle Filtering in Geophysical Systems. *Monthly Weather Review, 137*(12), 4089-4114. doi:10.1175/2009mwr2835.1
van Leeuwen, P. J., Kunsch, H. R., Nerger, L., Potthast, R., & Reich, S. (2019). Particle filters for high-dimensional geoscience applications: A review. *Quarterly Journal Of The Royal Meteorological Society, 145*(723), 2335-2365. doi:10.1002/qj.3551
Wang, J., Aegerter, C., Xu, X. G., & Szykman, J. J. (2016). Potential application of VIIRS Day/Night Band for monitoring nighttime surface PM2.5 air quality from space. *Atmospheric Environment, 124*, 55-63. doi:10.1016/j.atmosenv.2015.11.013
Wang, J., Christopher, S. A., Nair, U. S., Reid, J. S., Prins, E. M., Szykman, J., & Hand, J. L. (2006). Mesoscale modeling of Central American smoke transport to the United States: 1. "Top-down" assessment of emission strength and diurnal variation impacts. *Journal Of Geophysical Research-Atmospheres, 111*(D5).
Wang, J., Nair, U. S., & Christopher, S. A. (2004). GOES 8 aerosol optical thickness assimilation in a mesoscale model: Online integration of aerosol radiative effects. *Journal Of Geophysical Research-Atmospheres, 109*(D23), D23203, doi:23210.21029/22004JD004827.
Wang, J., Roudini, S., Hyer, E. J., Xu, X. G., Zhou, M., Garcia, L. C., . . . da Silva, A. M. (2020). Detecting nighttime fire combustion phase by hybrid application of visible and infrared radiation from Suomi NPP VIIRS. *Remote Sensing Of Environment, 237*, 14. doi:10.1016/j.rse.2019.111466
Wang, J., Zhou, M., Xu, X. G., Roudini, S., Sander, S. P., Pongetti, T. J., . . . Spurr, R. (2020). Development of a nighttime shortwave radiative transfer model for remote sensing of nocturnal aerosols and fires from VIIRS. *Remote Sensing Of Environment, 241*, 16. doi:10.1016/j.rse.2020.111727
Wang, Y. Z., Bechle, M. J., Kim, S. Y., Adams, P. J., Pandis, S. N., Pope, C. A., . . . Marshall, J. D. (2020). Spatial decomposition analysis of NO2 and PM2.5 air pollution in the United States. *Atmospheric Environment, 241*, 8. doi:10.1016/j.atmosenv.2020.117470
Wiggins, E. B., Soja, A. J., Gargulinski, E., Halliday, H. S., Pierce, R. B., Schmidt, C. C., . . . Moore, R. H. (2020). High Temporal Resolution Satellite Observations of Fire Radiative Power Reveal Link Between Fire Behavior and Aerosol and Gas Emissions. *Geophysical Research Letters, 47*(23), 12. doi:10.1029/2020gl090707

Xian, P., Reid, J. S., Hyer, E. J., Sampson, C. R., Rubin, J. I., Ades, M., . . . Zhang, J. L. (2019). Current state of the global operational aerosol multi-model ensemble: An update from the International Cooperative for Aerosol Prediction (ICAP). *Quarterly Journal Of The Royal Meteorological Society, 145*, 176-209. doi:10.1002/qj.3497

Xu, X. G., Wang, J., Wang, Y., Zeng, J., Torres, O., Reid, J. S., . . . Remer, L. A. (2019). Detecting layer height of smoke aerosols over vegetated land and water surfaces via oxygen absorption bands: hourly results from EPIC/DSCOVR in deep space. *Atmospheric Measurement Techniques, 12*(6), 3269-3288. doi:10.5194/amt-12-3269-2019

Xu, X. G., Wang, J., Wang, Y., Zeng, J., Torres, O., Yang, Y. K., . . . Miller, S. (2017). Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: First results from EPIC/DSCOVR at Lagrange-1 point. *Geophysical Research Letters, 44*(14), 7544-7554. doi:10.1002/2017gl073939

Ye, X. X., Arab, P., Ahmadov, R., James, E., Grell, G. A., Pierce, B., . . . Saide, P. E. (2021). Evaluation and intercomparison of wildfire smoke forecasts from multiple modeling systems for the 2019 Williams Flats fire. *Atmospheric Chemistry And Physics, 21*(18), 14427-14469. doi:10.5194/acp-21-14427-2021

Yeo, I. K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika, 87*(4), 954-959. doi:10.1093/biomet/87.4.954

Yu, H. B., Dickinson, R. E., Chin, M., Kaufman, Y. J., Holben, B. N., Geogdzhayev, I. V., & Mishchenko, M. I. (2003). Annual cycle of global distributions of aerosol optical depth from integration of MODIS retrievals and GOCART model simulations. *Journal Of Geophysical Research-Atmospheres, 108*(D3).

Yumimoto, K., Nagao, T. M., Kikuchi, M., Sekiyama, T. T., Murakami, H., Tanaka, T. Y., . . . Maki, T. (2016). Aerosol data assimilation using data from Himawari-8, a next-generation geostationary meteorological satellite. *Geophysical Research Letters, 43*(11), 5886-5894. doi:10.1002/2016gl069298

Yumimoto, K., & Takemura, T. (2011). Direct radiative effect of aerosols estimated using ensemble-based data assimilation in a global aerosol climate model. *Geophysical Research Letters, 38*, 6. doi:10.1029/2011gl049258

Yumimoto, K., & Takemura, T. (2013). The SPRINTARS version 3.80/4D-Var data assimilation system: development and inversion experiments based on the observing system simulation experiment framework. *Geoscientific Model Development, 6*(6), 2005-2022. doi:10.5194/gmd-6-2005-2013

Yumimoto, K., Tanaka, T. Y., Oshima, N., & Maki, T. (2017). JRAero: the Japanese Reanalysis for Aerosol v1.0. *Geoscientific Model Development, 10*(9), 3225-3253. doi:10.5194/gmd-10-3225-2017

Yumimoto, K., Tanaka, T. Y., Yoshida, M., Kikuchi, M., Nagao, T. M., Murakami, H., & Maki, T. (2018). Assimilation and Forecasting Experiment for Heavy Siberian Wildfire Smoke in May 2016 with Himawari-8 Aerosol Optical Thickness. *Journal of the Meteorological Society of Japan, 96B*, 133-149. doi:10.2151/jmsj.2018-035

Yumimoto, K., Uno, I., Sugimoto, N., Shimizu, A., Liu, Z., & Winker, D. M. (2008). Adjoint inversion modeling of Asian dust emission using lidar observations. *Atmospheric Chemistry And Physics, 8*(11), 2869-2884. doi:10.5194/acp-8-2869-2008

Zhang, F., Wang, J., Ichoku, C., Hyer, E. J., Yang, Z., Ge, C., . . . da Silva, A. (2014). Sensitivity of mesoscale modeling of smoke direct radiative effect to the emission inventory: a case study in northern sub-Saharan African region. *Environmental Research Letters, 9*(7), 075002 (075014 pp.)-075002 (075014 pp.). doi:10.1088/1748-9326/9/7/075002

Zhang, H., Kondragunta, S., Laszlo, I., Liu, H. Q., Remer, L. A., Huang, J. F., . . . Ciren, P. (2016). An enhanced VIIRS aerosol optical thickness (AOT) retrieval algorithm over land using a global surface reflectance ratio database. *Journal Of Geophysical Research-Atmospheres, 121*(18), 10717-10738. doi:10.1002/2016jd024859

Zhang, H., Kondragunta, S., Laszlo, I., & Zhou, M. (2020). Improving GOES Advanced Baseline Imager (ABI) aerosol optical depth (AOD) retrievals using an empirical bias correction algorithm. *Atmospheric Measurement Techniques, 13*(11), 5955-5975. doi:10.5194/amt-13-5955-2020

Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Westphal, D. L., & Johnson, R. S. (2014). Evaluating the impact of multisensor data assimilation on a global aerosol particle transport model. *Journal Of Geophysical Research-Atmospheres, 119*(8), 4674-4689. doi:10.1002/2013jd020975

Zhang, J., Reid, J. S., Westphal, D. L., Baker, N. L., & Hyer, E. J. (2008). A system for operational aerosol optical depth data assimilation over global oceans. *Journal Of Geophysical Research-Atmospheres, 113*(D10). doi:D10208 10.1029/2007jd009065

Zhang, J. L., Campbell, J. R., Hyer, E. J., Reid, J. S., Westphal, D. L., & Johnson, R. S. (2014). Evaluating the impact of multisensor data assimilation on a global aerosol particle transport model. *Journal Of Geophysical Research-Atmospheres, 119*(8), 4674-4689. doi:10.1002/2013jd020975

Zhang, J. L., Campbell, J. R., Reid, J. S., Westphal, D. L., Baker, N. L., Campbell, W. F., & Hyer, E. J. (2011). Evaluating the impact of assimilating CALIOP-derived aerosol extinction profiles on a global mass transport model. *Geophysical Research Letters, 38*. doi:L14801

10.1029/2011gl047737

Zhang, J. L., Jaker, S. L., Reid, J. S., Miller, S. D., Solbrig, J., & Toth, T. D. (2019). Characterization and application of artificial light sources for nighttime aerosol optical depth retrievals using the Visible Infrared Imager Radiometer Suite Day/Night Band. *Atmospheric Measurement Techniques, 12*(6), 3209-3222. doi:10.5194/amt-12-3209-2019

Zhang, J. L., Reid, J. S., Christensen, M., & Benedetti, A. (2016). An evaluation of the impact of aerosol particles on weather forecasts from a biomass burning aerosol event over the Midwestern United States: observational-based analysis of surface temperature. *Atmospheric Chemistry And Physics, 16*(10), 6475-6494. doi:10.5194/acp-16-6475-2016

Zhang, J. L., Reid, J. S., & Holben, B. N. (2005). An analysis of potential cloud artifacts in MODIS over ocean aerosol optical thickness products. *Geophysical Research Letters, 32*(15), 4. doi:10.1029/2005gl023254

Zhang, J. L., Reid, J. S., Westphal, D. L., Baker, N. L., & Hyer, E. J. (2008). A system for operational aerosol optical depth data assimilation over global oceans. *Journal Of Geophysical Research-Atmospheres, 113*(D10), 13. doi:10.1029/2007jd009065

Zhang, J. L., Spurr, R. J. D., Reid, J. S., Xian, P., Colarco, P. R., Campbell, J. R., . . . Baker, N. L. (2021). Development of an Ozone Monitoring Instrument (OMI) aerosol index (AI) data assimilation scheme for aerosol modeling over bright surfaces - a step toward direct radiance assimilation in the UV spectrum. *Geoscientific Model Development, 14*(1), 27-42. doi:10.5194/gmd-14-27-2021

Zhao, X. R., Shi, H. Q., Yu, H., & Yang, P. L. (2016). Inversion of Nighttime PM2.5 Mass Concentration in Beijing Based on the VIIRS Day-Night Band. *Atmosphere, 7*(10). doi:10.3390/atmos7100136

Zhou, H. Y., Gomez-Hernandez, J. J., Franssen, H. J. H., & Li, L. P. (2011). An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering. *Advances in Water Resources, 34*(7), 844-864. doi:10.1016/j.advwatres.2011.04.014