$PM_{2.5}$ Assimilation within JEDI for NOAA's Regional Air Quality Model (AQMv7): Application to the September 2020 Western U.S. Wildfires

Hongli Wang[1,2], Cory Martin[3], Jérôme Barré[4,5], Ruifang Li[1,2], Steve Weygandt[2], Jianping Huang[3], Youhua Tang[6,7], Hyundeok Choi[8,3], Andrew Tangborn[8,3], Kai Wang[9,3], Haixia Liu[9,3], Jeffrey Lee[10]

1. Cooperative Institute for Research In Environmental Sciences, University of Colorado, Boulder, CO 80305
2. NOAA Global Systems Laboratory, Boulder, CO 80305
3. NOAA/NWS/NCEP/EMC, College Park, MD 20740
4. NASA Global Modeling and Assimilation Office, Greenbelt, MD, USA
5. Morgan State University, Baltimore, MD, USA
6. Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA 22030
7. NOAA Air Resources Laboratory (ARL), College Park, MD 20740
8. SAIC@NOAA/NWS/NCEP/EMC, College Park, MD 20740
9. LINKER@NOAA/NWS/NCEP/EMC, College Park, MD 20740
10. School of Meteorology, University of Oklahoma, Norman, OK 73072

Correspondence to: Hongli Wang (hongli.wang@noaa.gov)

Abstract. This paper describes efforts to establish aerosol data assimilation capabilities for NOAA's National Air Quality Forecasting Capability (NAQFC), a regional online air quality modeling (AQM) system under NOAA's Unified Forecast System (UFS), by assimilating measurements of fine particulate matter ($PM_{2.5}$, particles with aerodynamic diameters less than 2.5 μm). $PM_{2.5}$ assimilation is developed within the Joint Effort for Data assimilation Integration (JEDI) framework and tested using its 3D-Var data assimilation (DA) component. The $PM_{2.5}$ observation operator is constructed by combining newly developed $PM_{2.5}$ transformation recipes in the JEDI Variable Derivation Repository (VADER) with a general spatial interpolation operator in the Unified Forward Operator (UFO).

Cycled DA and forecast experiments were conducted from 1 to 21 September 2020, during a period of Western U.S. wildfires, to assess the impact of assimilating $PM_{2.5}$ observations from the AirNow and PurpleAir networks. The control and analysis variables include individual aerosol species, with background error standard deviations generated by scaling their respective background values. Prognostic variables such as aerosol particle number and total particulate surface area are updated accordingly following each analysis update. All DA experiments use a 3-hourly cycling interval, with $PM_{2.5}$ observations assimilated every 3 hours. The control experiment uses the same configuration but without any data assimilation. Results show that assimilating either AirNow or PurpleAir $PM_{2.5}$

data reduces 1–24 h forecast errors in terms of mean absolute error (MAE) and root mean square error (RMSE) compared to the control run over Continental United States(CONUS. Substantial improvements are in regions where fire events took place and largely affected by transported smoke. AirNow observations have a greater impact overall, while PurpleAir shows its strongest impact over northern Utah, Colorado, and southwestern New Mexico. Overall, the assimilation of PurpleAir observations in addition to AirNow data leads to a slight reduction in 3–24 h MAE.

1 Introduction

Particulate matter with an aerodynamic diameter of 2.5 micrometers or smaller ($PM_{2.5}$) is a key contributor to poor air quality in the United States, posing significant risks to public health and the environment, and contributing to substantial loss of life (Cohen et al. 2017; Colmer et al. 2020; Huang et al., 2025). Over the past few decades, poor air quality in the U.S. has contributed to over 100,000 premature deaths annually, far exceeding fatalities from all other weather-related causes combined, which average around 500 per year (Huang et al., 2025). Given its public health significance, $PM_{2.5}$ is one of the important pollutants used in calculating the Air Quality Index (AQI)—a standardized system designed to communicate daily air pollution levels to the public at the U.S. Environmental Protection Agency (EPA). Elevated $PM_{2.5}$ concentrations frequently result in "unhealthy" AQI ratings, triggering health advisories and public warnings.

$PM_{2.5}$ in the United States originates from a range of both anthropogenic and natural sources. Anthropogenic sources include agricultural activities and combustion processes, such as emissions from motor vehicles, power plants, industrial facilities, and residential heating systems. Among natural sources, wildfires are a particularly significant contributor, especially in the western United States, where their frequency and intensity have escalated dramatically over the past two decades (Wen and Burke, 2021). According to the U.S. Environmental Protection Agency (EPA), wildfires account for approximately 15% to 30% of total $PM_{2.5}$ emissions nationwide (EPA, 2017). While national seasonal averages of $PM_{2.5}$ have generally declined, summer $PM_{2.5}$ concentrations in the western U.S. have remained persistently high, primarily due to wildfire smoke (O'Dell et al., 2019). In addition to degrading air quality, wildfires have caused widespread property loss. Since 2005, more than 129,000 homes, businesses, and other structures have been destroyed by wildfire-related events (https://headwaterseconomics.org/natural-hazards/structures-destroyed-by-wildfire, last access on February 18, 2026), underscoring the urgent need for more effective strategies in air quality monitoring, forecasting, and wildfire management.

The National Oceanic and Atmospheric Administration (NOAA) has developed an advanced regional Air Quality Modeling (AQM) prediction system within the Unified Forecast System (UFS) framework to enhance the accuracy of air quality forecasts across the United States, particularly during wildfire events (Huang et al. 2025). The National Air Quality Forecast Capability (NAQFC), operated by NOAA's National Weather Service (NWS), has been providing operational air quality forecast guidance for over 20 years, with continuous inclusion of new capabilities.

Under NAQFC, the AQM version 7 was implemented and became operational on May 14, 2024. The system features online coupling of atmospheric and chemical models, allowing dynamic interactions between meteorology and atmospheric chemistry. This integration improves the representation of emissions and ensures real-time feedback of meteorological fields that influence chemical transformations and the transport of pollutants in the atmosphere. The UFS-AQM online system has consistently shown improved performance in simulating major wildfire events, including the significant wildfires in the northwestern coastal regions of the U.S. in September 2020, and widespread smoke transport from Canadian wildfires in the summer of 2023. This system was officially implemented on May 14, 2024 as NOAA's operational air quality prediction system (AQMv7), replacing the previous offline-coupled Global Forecast System using the Finite Volume Cube-Sphere dynamical core (GFS-FV3) version 15 with the Community Multiscale Air Quality modeling system (CMAQv5.0.2)(Chen et al. 2021).

$PM_{2.5}$ data assimilation (DA) has proven effective in reducing errors in air quality forecasts (e.g., Pagowski et al. 2010, 2012; Schwartz et al. 2012; Wu et al. 2015; Robichaud 2017; Lee et al. 2021; Chen et al. 2022, Ha 2022; Vogel et al. 2025, among others). Pagowski et al. (2010) demonstrated that fine aerosol forecasts benefit from AirNow $PM_{2.5}$ DA, showing improved verification scores for a period of at least 24 hours. Schwartz et al. (2012) found that assimilating AirNow $PM_{2.5}$ observations significantly improved surface $PM_{2.5}$ forecasts over the CONUS compared to forecasts without DA. Wu et al. (2015) reported that incorporating ground-based $PM_{2.5}$ observations notably enhanced 24-hour forecasts during a severe pollution episode in Shanghai. Similarly, Chen et al. (2022) showed that assimilating multi-source $PM_{2.5}$ data significantly improved WRF-Chem $PM_{2.5}$ forecasts with benefits lasting up to 48 hours. Lee et al. (2021) highlighted the effectiveness of assimilating ground in-situ surface $PM_{2.5}$ observations in improving the short-term $PM_{2.5}$ predictions in Northeast Asia.

Many operational regional air quality prediction systems around the world use some form of data assimilation to initialize the forecasts. These approaches vary in complexity, ranging from simple optimal interpolation to full variational or ensemble Kalman filter methods (e.g. Robichaud et al. 2016; Wei et al. 2024; Colette et al. 2024). In NOAA's current regional air quality model (AQM) operations, aerosol and chemical initial conditions are "warm-started" using 6-hour forecasts from the previous model cycle. The implementation of an aerosol data assimilation system can further enhance short-term air quality forecasts by providing more accurate spatial analyses of initial aerosol distributions.

To establish aerosol data assimilation capabilities for NOAA's regional operational AQM system, we employ the Joint Effort for Data assimilation Integration (JEDI) (Trémolet and Auligné, 2020). JEDI is a flexible, agnostic, and modern data assimilation system applicable to a wide range of forecasting systems (e.g. Liu et al. 2023; Huang et al. 2023; Sluka, 2024). JEDI offers a platform that supports efficient scientific development and facilitates the transition from research to operations. As part of a broader strategic shift, NOAA and partner agencies are transitioning their data assimilation systems to JEDI, opening the door for rapid integration of new scientific advancements, greater consistency across modeling systems, and enhanced collaboration across research communities and operational centers.

This study aims to develop and evaluate an initial aerosol analysis capability for the NOAA's regional AQM system by assimilating $PM_{2.5}$ observations using the JEDI three-dimensional variational (3D-Var) data assimilation framework. Compared to previous $PM_{2.5}$ data assimilation studies, this research adopts the NOAA's regional operational AQMv7 system and incorporates a new $PM_{2.5}$ transform in JEDI for assimilating $PM_{2.5}$ observations. In addition to evaluating the impact of assimilating AirNow $PM_{2.5}$ measurements on air quality prediction, this study also examines the impact of assimilating low-cost PurpleAir observations. Although PurpleAir data are valuable for real-time air quality monitoring, their impact on numerical air quality prediction has not been thoroughly investigated. To the authors' best knowledge, this is the first study to demonstrate the value of PurpleAir observations for air quality prediction during a major fire event using the AQMv7 system.

The paper is organized as follows: section 2 provides a description of Methodology including the NOAA's AQM system, 3D-Var approach, and JEDI $PM_{2.5}$ assimilation. Experimental setup is presented in section 3 including case description, AQM configuration, AirNow and PurpleAir $PM_{2.5}$ observations and background errors setup. Results are described in section 4. A summary and discussion are presented in the final section.

2 Methodology

2.1 AQMv7 overview

The NOAA's regional operational AQMv7 system was developed through the online coupling of the Finite-Volume version 3 (FV3) dynamical core-based atmospheric model (Black et al 2021) with the EPA's Community Multiscale Air Quality (CMAQ) model v5.2.0 within the UFS framework (Huang et al., 2025). In this UFS-AQM online system, CMAQ is treated as an atmospheric chemistry column model to simulate atmospheric chemistry reactions that govern concentrations of chemical species including gas- and aerosol-phase species. The transport terms of chemical species are handled by the FV3 dynamical core in the same way as other physics tracers (Huang et al., 2025). Aerosol module version 6 (AERO6) (Zhang et al. 2018) is utilized by CMAQ to simulate aerosol processes.

2.2 $PM_{2.5}$ assimilation within JEDI 3D-Var

In the JEDI framework, a series of components are provided to create a flexible, comprehensive data assimilation system (Trémolet. and Auligné, 2020). The JEDI 3D-Var component is used to assimilate $PM_{2.5}$ for AQMv7. The 3D-Var method is chosen for its operational feasibility, primarily due to its low computational cost and the fact that it does not require an ensemble prediction system, as is needed in (hybrid) ensemble–variational data assimilation.

In practice, a 3D-Var data assimilation system typically uses an incremental approach to minimize a quadratic cost function which is defined in terms of the analysis increment δx relative to the guess state x_g:

$$J(\delta x) = \frac{1}{2}(\delta x - \delta x_g)B^{-1}(\delta x - \delta x_g)^T + \frac{1}{2}(\mathbf{H}[\delta x] - d))R^{-1}(\mathbf{H}[\delta x] - d)^T \quad (1)$$

Where:

- $\delta x_g = x_b - x_g$ is the guess state departure from background state x_b, which is usually taken from a previous short-term forecast.
- **H** is the linearized observation operator of nonlinear observation operator ***H***.
- **B** and **R** are the background and observation error covariance matrices, respectively.
- d is the innovation vector, defined as:

$$d = y - H(x_g) \quad (2)$$

with y representing the observation vector.

Once the increment δx is obtained, the analysis state x^a is reconstructed as:

$$x^a = x_g + \delta x \quad (3)$$

2.2.1 $PM_{2.5}$ observation operator

To assimilate $PM_{2.5}$ data, a $PM_{2.5}$ transform that builds relationships between the model aerosol variables and the observed $PM_{2.5}$ needs to be developed. In AQMv7, the modal approach taken in the CMAQ model represents aerosol particle size distributions as the superposition of three lognormal modes: Aitken (I), accumulation (J), and coarse (K). It predicts only three integral properties of the size distribution for each mode: the total particle number concentration, the total surface area concentration, and the total mass concentration of the individual chemical components.

The total $PM_{2.5}$ concentration is calculated as a weighted sum of the individual aerosol concentration across these three modes:

$$PM_{2.5}=ATOTI\cdot PM25AT+ATOTJ\cdot PM25AC+ATOTK\cdot PM25CO \quad (4)$$

Here, ATOTI, ATOTJ, and ATOTK represent the total aerosol mass concentrations in the Aitken, accumulation, and coarse modes, respectively. For example, ATOTI is the combined mass of 14 prognostic aerosol variables in the Aitken mode from the AERO6 aerosol module. Similarly, ATOTJ and ATOTK are the aggregated mass concentrations of 49 and 7 aerosol variables in the accumulation and coarse modes, respectively. PM25AT, PM25AC, and PM25CO are mass scaling factors for the three modes that vary by location and time. The aerosol variables within the same mode share the same mass scaling factor.

The $PM_{2.5}$ observation operator is constructed by combining the newly developed $PM_{2.5}$ transformation recipes in the JEDI Variable Derivation Repository (VADER) with an existing general spatial interpolation operator in the Unified Forward Operator (UFO). VADER is responsible for transforming model variables using user-defined "recipes" to generate new variables in model space. For $PM_{2.5}$ assimilation, VADER computes $PM_{2.5}$ from individual aerosol species using model-specific transformation, specifically using the equation 4 for this application. Since $PM_{2.5}$ composition varies by model, these transforms are implemented within VADER to match the specific structure of the regional air quality model AQMv7. Once $PM_{2.5}$ is derived in model space, UFO applies a generic spatial interpolation operator to map the model-simulated values to the observation locations, enabling computation of the observed minus forecast values.

The input for the $PM_{2.5}$ transformation are mixing ratio of the 70 aerosol variables wrt dry air in unit ug/kg, the three mass scaling factors in the three modes, and dry air density for unit conversion. The output product is the $PM_{2.5}$ in unit ug/m^3. It is noted that a new recipe to derive dry air density from temperature, surface pressure, and delta pressure has been added to VADER for cases where the dry air density is not provided.

The new JEDI/VADER $PM_{2.5}$ recipe provides nonlinear (NL), tangent linear (TL), and adjoint (AD) transforms of $PM_{2.5}$ that keeps the output products in the same grid space as the input variables. Hence, the generic interpolation operator in UFO is used to connect the model-derived $PM_{2.5}$ fields with observed surface $PM_{2.5}$ measurements. This respects the JEDI paradigm of keeping the UFO part of the JEDI model independent.

2.2.2 Background error covariance modeling

In a 3D-Var system, the background error covariance (BEC) determines both the spatial spreading of information from observations and the magnitude of the analysis increments along with the observation error variance.

The background error covariance matrix B can be decomposed into a standard deviation matrix (Σ) and a correlation matrix (C), as follows:

$$B=\Sigma C\Sigma \quad (5)$$

The correlation matrix C is generally non-diagonal. Σ is a diagonal matrix, with the standard deviations of the background errors for each variable on the diagonal.

The error modeling of the correlation matrix and standard deviations usually apply to control variables. In the first implementation of aerosol data assimilation in JEDI for AQMv7, the control variables are defined as individual forecast aerosol variables, resulting in 70 control variables for AQMv7 with the AERO6 aerosol mechanism. The setup of background error standard deviation and correlation modeling will be described in Section 3: Experimental setup.

2.2.3 Minimization Algorithm (DRIPCG)

JEDI provides several minimization algorithm options. In this paper, we use the Derber–Rosati Inexact Preconditioned Conjugate Gradient (DRIPCG) algorithm (Derber and Rosati, 1989), as implemented in the JEDI's OOPS (Object-Oriented Prediction System) framework. DRIPCG has been extensively tested and is chosen here for stability and convergence efficiency.

3 Experimental setup

3.1 The September 2020 fire event and AQMv7 system setup

The wildfires of September 2020 ranked among the most intense in the U.S. in recent years. These fires produced dense smoke that initially moved westward over the Willamette Valley in western Oregon and eventually blanketed the broader region. As a result, air quality rapidly across Oregon, Washington, and Idaho deteriorated to hazardous levels, marking one of the worst air quality periods in recent decades (Mass et al., 2021). Wildfire smoke originating from California, Oregon, and Washington was injected into the free troposphere and transported across the country by prevailing winds, leading to hazy conditions in several states. According to Li et al. (2021), from August to October 2020, wildfires in the western U.S. contributed 23% of surface $PM_{2.5}$ across CONUS, with higher contributions observed along the Pacific Coast (43%) and in mountain region (42%). This study focuses on the peak fire activity occurring between September 1 and 21.

In this research, the model configuration is almost the same as the operational AQMv7 setup except for running over the CONUS domain with a 3 hourly cycling interval. The AQMv7 system is configured over the CONUS domain with a grid-spacing of 13 km and 65 vertical levels, extending up to 0.2 hPa. The system uses the Global Forecast System version 16 (GFSv16) physics package within the Common Community Physics Package (CCPP) framework to generate the meteorological fields driving air quality predictions. Meteorological initial conditions and lateral boundary conditions are generated using GFS forecast outputs with lead times up to 30 hours at 3-hour intervals from the previous GFS cycle. Fire-related emissions are represented using real-time Regional hourly Advanced Baseline Imager (ABI) and Visible Infrared Imaging Radiometer Suite (VIIRS) Emissions (RAVE) data at 0.03° spatial resolution. Anthropogenic emissions are based on the 2016 U.S. EPA NEI Collaborative (NEIC2016v1) modeling platform. Gas-phase chemistry is simulated using the Carbon Bond Mechanism version 6 (CB6r3) with updated isoprene chemistry and revised photolysis rates. More detailed information including physics, chemistry options, anthropogenic emissions, and fire emissions about the model configuration can be found in Huang et al. (2025).

3.2 $PM_{2.5}$ observations

In this study, surface $PM_{2.5}$ observations were obtained from two sources: AirNow and PurpleAir observing networks. These datasets differ in sensor type, spatial coverage, and quality control (QC) requirements. AirNow provides regulatory-grade measurements from federal, state, and local monitoring stations, while PurpleAir is a low-cost, community-based network of air quality sensors. PurpleAir sensors are widely deployed by individuals

and communities, providing real-time data on $PM_{2.5}$ concentrations as well as meteorological variables such as temperature, pressure, and relative humidity. Only the data reported from outdoor $PM_{2.5}$ sensors are used in this study. The PurpleAir data were available for registered users through the PurpleAir API. (https://community.purpleair.com/t/api-use-guidelines/1589: last access on February 18 2026)

3.2.1 PurpleAir $PM_{2.5}$ quality control and correction

Quality control and correction of PurpleAir data followed the methodology described in Barkjohn et al. (2021). Readers are referred to that paper for further details. A correction is required because the PurpleAir raw data usually overestimate $PM_{2.5}$ concentrations under typical ambient and smoke-impacted conditions. The following quality control (QC) filters were applied to the raw PurpleAir $PM_{2.5}$ measurements:

- Reported $PM_{2.5}$ values from two Plantower sensors within the PurpleAir sensor (channels A and B) must be nonnegative.
- The PurpleAir sensor channel A and B consistency:
    - Absolute difference < 5 μg/m³, *or*
    - Relative difference within 61%.
- $PM_{2.5}$ values must not exceed PM10 values.
- $PM_{2.5}$ values must be less than 3000 μg/m³ (upper threshold).
- Gross check of relative humidity with range 0-100%.

Only PurpleAir $PM_{2.5}$ measurements that passed all the above QC criteria were retained for subsequent correction.

3.2.2 PurpleAir $PM_{2.5}$ correction

Correction of PurpleAir $PM_{2.5}$ measurements was performed using a multiple linear regression model based on sensor-reported $PM_{2.5}$ (PA) and relative humidity (RH), following the correction formula proposed by Barkjohn et al. (2021):

$$PM_{2.5}=0.524\times PA-0.0862\times RH+5.75 \quad (6)$$

We adopt the above equation because it was United States-wide valid by fitting data from September 2017 until January 2020. In addition, although the correction was originally derived for 24-hour averaged $PM_{2.5}$, it is consistent with a regression equation obtained from the September 2020 dataset based on 1-hour averaged $PM_{2.5}$.

3.2.3 Observation error assignment

Observation error standard deviations were assigned to each network:

- AirNow $PM_{2.5}$: 5% of observed value

- PurpleAir $PM_{2.5}$: 10% of observed value

The values above are based on the EPA's definition of acceptable measurement uncertainty, which specifies a 10% coefficient of variation for total precision. The AirNow $PM_{2.5}$ observation errors were set to 5% of the observed values. For PurpleAir $PM_{2.5}$ data, the observation errors were set to 10%, reflecting the higher likelihood of greater uncertainties associated with lower-cost sensors.

Figure 1a–b shows the spatial distribution of AirNow and PurpleAir $PM_{2.5}$ monitoring stations at 1200 UTC on September 16, 2020. PurpleAir sensors are especially concentrated in densely populated areas, leading to notable spatial variability in observation coverage during the September 2020 wildfire events. Coverage is particularly dense in urban regions of the western United States (e.g. California, Oregon, Washington, Utah, Arizona and Colorado), while rural and remote areas have significantly fewer sensors, for example, Nevada and North Dakota. Figure 1c displays the time series of domain averaged $PM_{2.5}$ values and station counts from the AirNow and PurpleAir networks, including matched stations between the two. The number of AirNow stations ranges from approximately 800 to 900, while PurpleAir stations number between 1,160 and 1,300. Dropouts in the AirNow network lead to sudden decreases in station count and corresponding drops in the $PM_{2.5}$ time series. In contrast, the PurpleAir network shows a general upward trend in station count, with no major data dropouts observed.

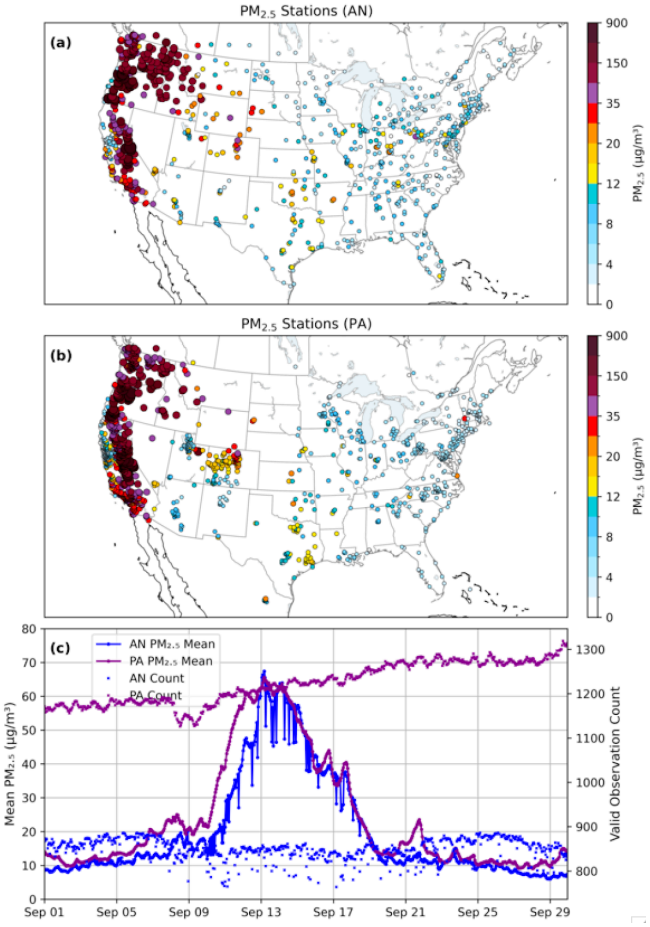


Figure 1. (a-b). Spatial distribution of AirNow(AN) and PurpleAir(PA) $PM_{2.5}$ monitoring stations on 1200 UTC 16 September 2020. (c). Time series of domain averaged $PM_{2.5}$ values and numbers from AirNow and PurpleAir observing networks.

3.3 Background error covariance

In this study, the background error standard deviation (Σ) for each control variable is constructed based on the background forecast; specifically, the error standard deviations of an aerosol variable are prescribed as proportional to its background values.

The proportional scaling factor *s* is approximately estimated by building a linear relationship between the $PM_{2.5}$ standard error (Σ) and the background forecast $PM_{2.5}^{bkg}$ $PM_{2.5}$ concentrations:

$$\Sigma = s.PM_{2.5}^{bkg} \quad (7)$$

The scaling factor *s* is subsequently applied to all $PM_{2.5}$ components, i.e., the 70 prognostic aerosol variables, to construct their error standard deviations.

This proportionality-based approach has also been adopted in the MOCAGE operational system (Colette et al., 2024), where background error standard deviations are similarly prescribed relative to background concentrations as a first-order approximation.

Tang et al. (2023) tested a similar method, in which the background $PM_{2.5}$ error variance is first estimated using the Hollingsworth–Lönnberg method (Hollingsworth and Lönnberg, 1986). A linear relationship is then established between the estimated $PM_{2.5}$ standard error and the background forecast $PM_{2.5}^{bkg}$.

Here we take the same idea but using an alternative approach to roughly estimate the background $PM_{2.5}$ forecast error variance. The background $PM_{2.5}$ error variance ($\mathbf{\Sigma^2}$) is estimated using $PM_{2.5}$ innovation information d defined in the subsection 2.2, and observation error information R specified in the subsection 3.2.3**,** specifically,

$$\Sigma^2 = E(d\,d^T) - R \quad (8)$$

In Eq. 8, $E(\cdot)$ denotes the mathematical expectation operator. The superscript T denotes the transpose of a vector. Equation 9 is valid under the assumption that observation and background errors are uncorrelated. This assumption is reasonable when the innovation vector d is calculated using forecasts from a free-running model without any aerosol data assimilation.

In this study, short-term (e.g., 3-hour) $PM_{2.5}$ forecasts from a free run conducted during 1–21 September 2020 were used to compute the innovation vector d . This free run, referred to as the *control run*, is described in detail in the

following section. Using the innovations and observation errors from subsection 2.2 and 3.2.3 as inputs to Eq. 8, the background error variance of $PM_{2.5}$ was first estimated. This error, along with the background values, was then used in Equation 7 to estimate the scaling factor, *s*. This scaling factor was subsequently applied in all assimilation experiments presented in this study.

This proportionality-based approach implicitly assumes that displacement errors in background don't dominate, focusing the assimilation process on correcting amplitude. It offers several benefits:

- It helps constrain analysis increments to physically meaningful regions. For example, it prevents the generation of sea salt aerosol increments over inland areas where no sea salt is present in the background. This is a problem that can occur when using GSI's height-dependent or latitude–height-dependent background error variance formulations, particularly when individual aerosol species are used as control variables.
- It introduces location- and time-dependent background error variance information, improving the realism of background error specification. Moreover, the aerosol variables that dominate background errors vary by location and assimilation cycle, rather than being consistently dominated by the same species when using constant static background error statistics. For example, organic and black carbon typically exhibit the largest errors in wildfire regions and downwind areas affected by smoke, whereas other regions may be dominated by non-organic aerosols.

An example of domain averaged background $PM_{2.5}$ error standard deviation from a data assimilation run that assimilated both AirNow and PurpleAir $PM_{2.5}$ is shown in Figure 2. This figure is intended to illustrate the main difference to static constant background errors, though the actual errors used in the data assimilation experiments are the errors of the individual aerosol control variables. It is obvious that this approach produces dynamically location- and time-dependent varying error estimates that yields particularly large error variances during the peak fire events from 10 to 20 September 2020.

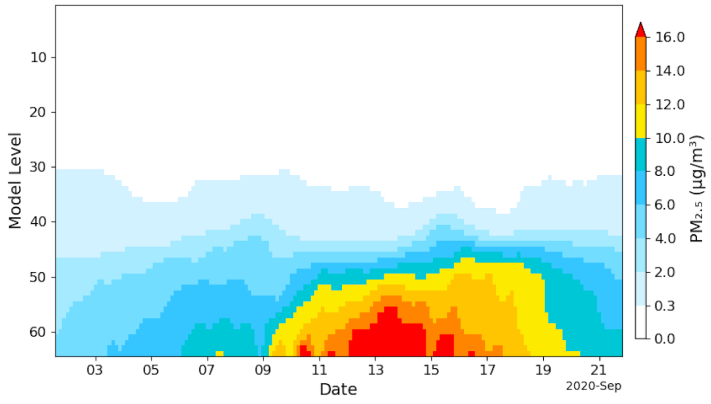


Figure 2. Domain averaged $PM_{2.5}$ error standard deviations for the data assimilation run that assimilated both AirNow and PurpleAir $PM_{2.5}$.

The background error correlation matrix C is modeled using a generic diffusion correlation operator designed for short length scales, as implemented in the System-Agnostic Background Error Representation (SABER) repository (Sluka, 2024). A horizontal cutoff scale of 100 km is applied, consistent with estimates derived from NMC statistics in previous GSI applications (Wang et al., 2021). For vertical correlations, this study uses a cutoff scale of 12 model levels, which helps confine the influence of surface $PM_{2.5}$ observations within the average daytime planetary boundary layer (PBL) height (~1450 m) and has demonstrated improved surface $PM_{2.5}$ prediction as will be discussed in Section 4.

3.4 Update of total particle number and surface area concentrations

After the aerosol mass concentration has been analyzed, total particle number concentration, total surface area concentration can be updated accordingly. For simplicity, it is assumed that the ratio of the particle number concentration to total particulate volume within each mode (I, J, K) remains the same as in the background. Total particulate volume is used instead of mass mixing ratio because it is proportional to the particle number concentration (see Eq. 3 in Binkowski and Roselle, 2003). A similar assumption was adopted by Li (2013) to update number concentrations for the WRF-Chem model.

The number of particles is updated using the following relation:

$$N_a = N_b / V_b \times V_a \quad (9)$$

Where:

- N_a and N_b are the number of particles in the analysis and background, respectively, within each mode.
- V_a and V_b are the total particulate volumes in the analysis and background, respectively, within the same mode.

The total particulate volume (V_a or V_b) within each mode is calculated by dividing the mass concentration of each aerosol variable by its corresponding density in that mode, and then summing the results. This updating approach implicitly assumes that changes in volume across the three modes are driven solely by variations in particle number, rather than shifts in the aerosol size distribution. The total particulate surface area within each mode is then updated using the same volume ratio, i.e., V_a / V_b (Eq. 9) multiplied by the background surface area.

In preparatory work for this study, six-hourly cycling experiments (Wang et al., 2025) have shown that updating these variables is crucial for improving AQMv7 performance. In contrast, previous work using GSI with earlier developmental versions of AQM did not update these variables, primarily because those model versions were less advanced than the current operational AQMv7. As a result, there was still significant room for improving prediction

skills.

3.5 Experiments

In the operational AQMv7 configuration, aerosol initial conditions (IC) are warm-started from a previous forecast cycle. In contrast, in cycling data assimilation and forecasting experiments, the data assimilation system provides aerosol initial conditions for the subsequent forecast, while the short-term (3-hour) forecast serves as the background for the next data assimilation cycle. A schematic of the data assimilation and forecasting cycles is shown in Fig. 3.

In the assimilation cycle, JEDI updates the aerosol analysis by combining $PM_{2.5}$ observations with the background information. The updated aerosol analysis, together with meteorological initial conditions, emissions, and other inputs, is then used to initialize the subsequent forecasts in the forecast cycle. Note that the meteorological initial conditions are not updated by JEDI but are generated from GFS forecast outputs of the previous GFS cycle.

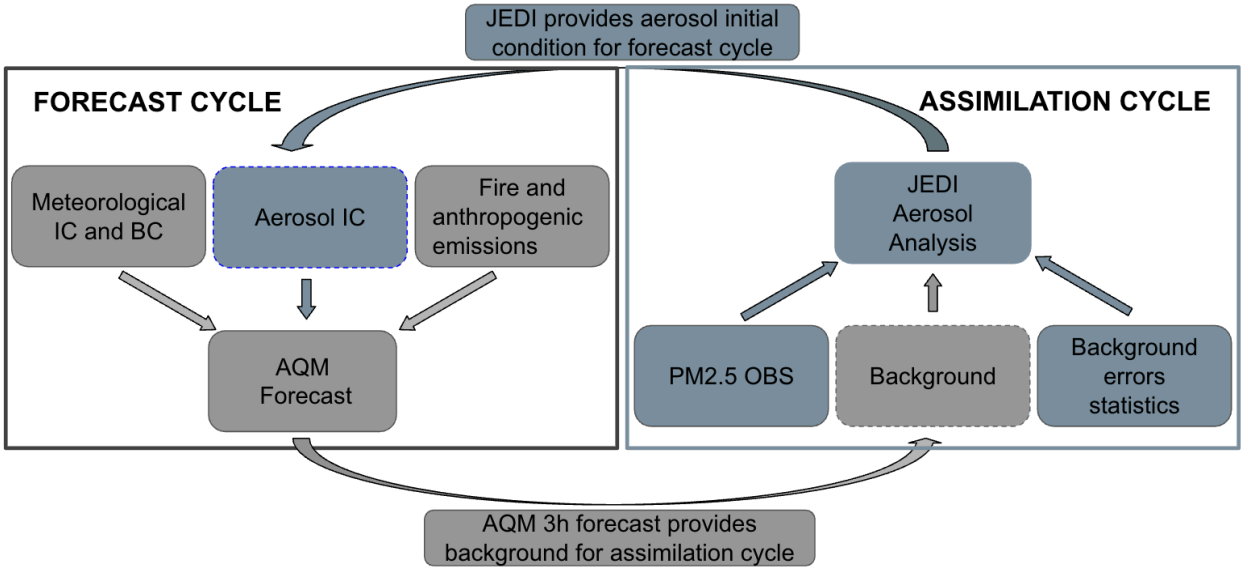


Figure 3. A schematic of the data assimilation and forecasting cycles.

Four experiments were conducted to evaluate the performance of JEDI/AQM $PM_{2.5}$ DA. Table 1 provides a description of the experiments.The first experiment is a control run (CTR), in which meteorological initial and boundary conditions are updated every 3 hours, while chemical and aerosol fields are carried over from the 3-hour forecast of the previous cycle. The other three experiments incorporate data assimilation: DA_AN, DA_PA, and DA_ANPA, which assimilate AirNow $PM_{2.5}$ only, PurpleAir $PM_{2.5}$ only, and both AirNow and PurpleAir $PM_{2.5}$ observations, respectively.

Table 1. Descriptions of the experiments.

Experiment	Data Assimilation	$PM_{2.5}$ Observations Assimilated	Aerosol Fields

CTR	No	None	Carried over from previous cycle's 3-hour forecast
DA_AN	Yes	AirNow $PM_{2.5}$ only	Updated by Assimilation
DA_PA	Yes	PurpleAir $PM_{2.5}$ only	Updated by Assimilation
DA_ANPA	Yes	AirNow + PurpleAir $PM_{2.5}$	Updated by Assimilation

Like the CTR experiment, all DA experiments are conducted as 3-hourly cycling runs. Data assimilation is performed every 3 hours, and a 3-hour forecast is launched at each cycle. This 3-hour forecast serves as the background for the subsequent data assimilation and forecasting cycle. In addition, forecasts initialized at 0000, 0600, 1200, and 1800 UTC are extended to 24 hours for evaluation purposes. The experimental period spans from 1200 UTC on September 1 to 1800 UTC on September 21, 2020. It is noted that to reduce random sensor noise and improve comparability with the model resolution (~13 km), the corrected PurpleAir $PM_{2.5}$ data were spatially averaged onto a 0.1° × 0.1° latitude–longitude grid.

4 Results

This section provides an overview of the impact of DA on $PM_{2.5}$ forecasts. A total of 80 forecasts—initialized four times daily from 0000 UTC on September 2 to 1800 UTC on September 21, 2020—are used to evaluate model performance. AirNow PM2.5 observations are used to verify the forecast. Forecast errors are assessed using bias, mean absolute error (MAE), and root mean square error (RMSE). Forecast performance is evaluated using box plots, which illustrate the distribution, spread, and central tendency of forecast errors. Time series of $PM_{2.5}$ at various forecast hours are presented to examine the temporal evolution of forecast performance. Additionally, spatial distributions of $PM_{2.5}$ including observations, forecasts, forecast errors, and forecast differences are analyzed to evaluate the spatial impact of data assimilation on $PM_{2.5}$ predictions.

Figure 4 presents the bias, mean absolute error (MAE), and root mean square error (RMSE) for the 1–24 h forecast of domain-averaged $PM_{2.5}$. Domain averages are computed over all states in the mainland United States. From the bias statistics (Fig. 4a), it is seen that the bias in the control run follows an upward trend initially, then reverses into a downward trend. The data assimilation runs show a similar trend, as data assimilation primarily corrects the model state and does not fully resolve inherent model bias. The control run underpredicted surface $PM_{2.5}$ throughout the 24-hour forecast period by about 1 μg m^{-3}. This underprediction was improved in the data assimilation experiments. The two assimilated experiments, DA_AN and DA_ANPA, reduced the 1-24h mean bias to −0.2 and −0.3 μg m^{-3},

respectively. The PurpleAir $PM_{2.5}$ assimilation experiment (DA_PA) also slightly improved the 1-24h mean bias. It is seen that DA_PA produces a near zero bias at 1h forecast, however this is due to large positive and negative biases canceling out. Therefore, it should be used together with norms like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), which quantify the actual magnitude of forecast errors.

In terms of MAE and RMSE, all the three data assimilation experiments reduced the surface $PM_{2.5}$ forecast error throughout the 24-hour forecast period. The added value of assimilating PurpleAir $PM_{2.5}$ data alongside AirNow observations is evident in the consistent MAE reduction (Fig. 4b). Its impact on RMSE (Fig. 4c) is also positive, though relatively small. Overall, all the data assimilation experiments show improved forecast skill compared to the control run.

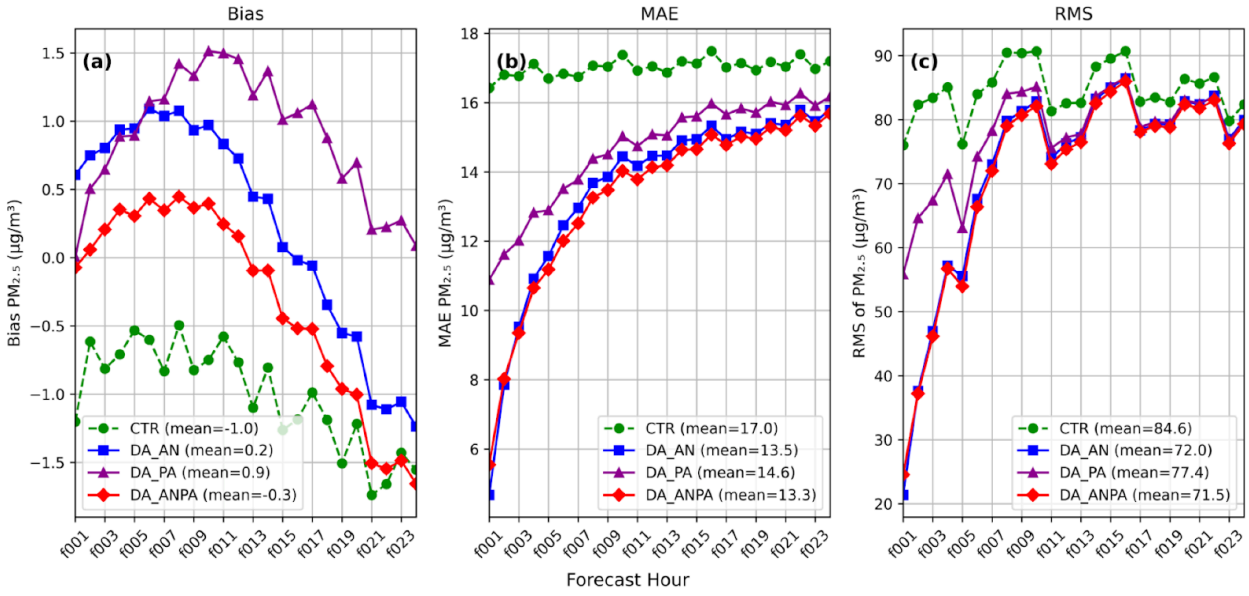


Figure 4. $PM_{2.5}$ forecast errors for 1–24 h lead times based on 80 forecasts initialized four times daily during September 2–21, 2020. Domain-averaged over CONUS. The x-axis represents forecast lead times from 1 to 24 hours.

(a) Bias, (b) Mean Absolute Error (MAE), (c) Root Mean Square Error (RMSE).

Figure 5 shows box-and-whisker plots of $PM_{2.5}$ forecast bias. Across all forecast hours, the interquartile range (IQR)—represented by the height of the boxes—is consistently smaller for the DA experiments compared to the control run. This indicates reduced forecast error spread between the 25th and 75th percentiles and suggests more consistent forecasts in the DA experiments. Although the median forecast bias in the control run is sometimes closer to zero, the DA_ANPA experiment performs comparably in terms of central tendency while showing clear improvements in reducing the mean forecast bias, as also reflected in Fig. 4a. Among the DA experiments, DA_AN and DA_ANPA show the most consistent improvement at 24-hour lead times, with DA_ANPA slightly outperforming others during the early forecast hours (e.g., hour 1 to 12). This suggests that assimilating PurpleAir observations in addition to AirNow helps reduce bias and brings the forecasts closer to observed $PM_{2.5}$ values in the short term.

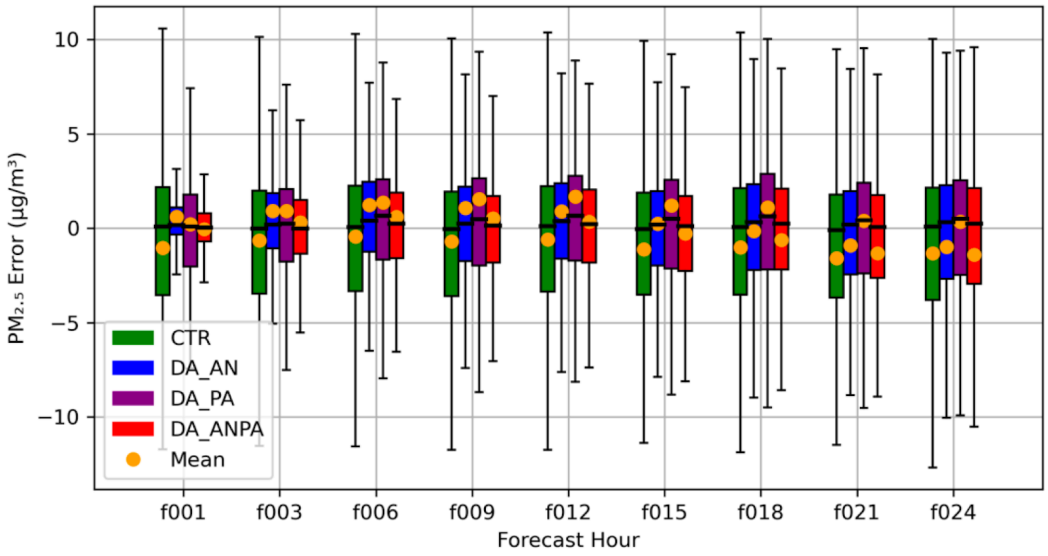
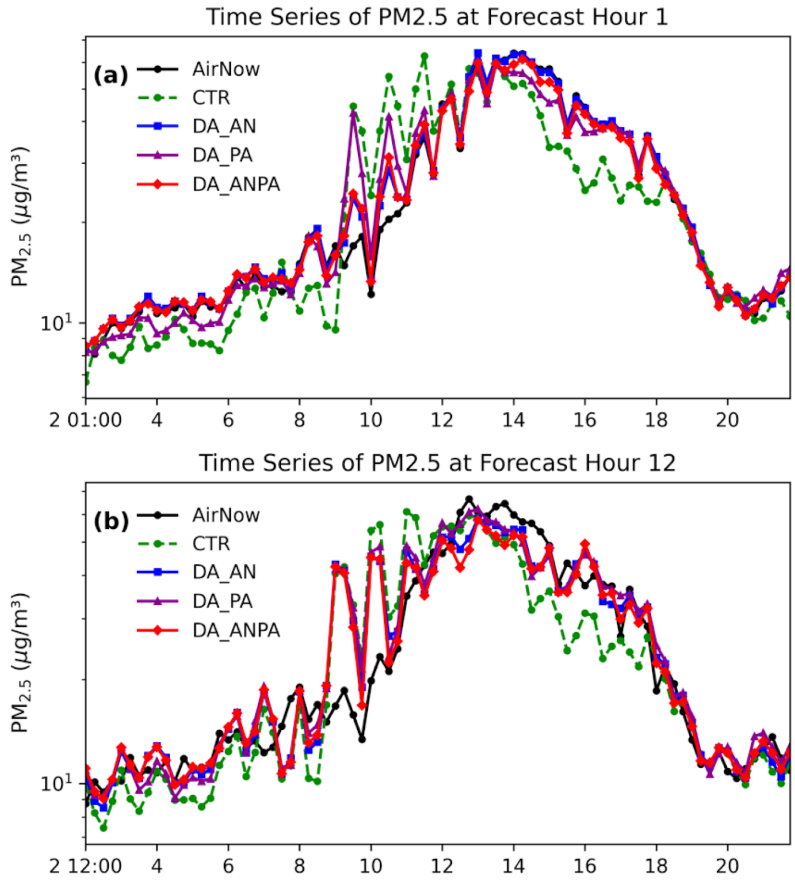
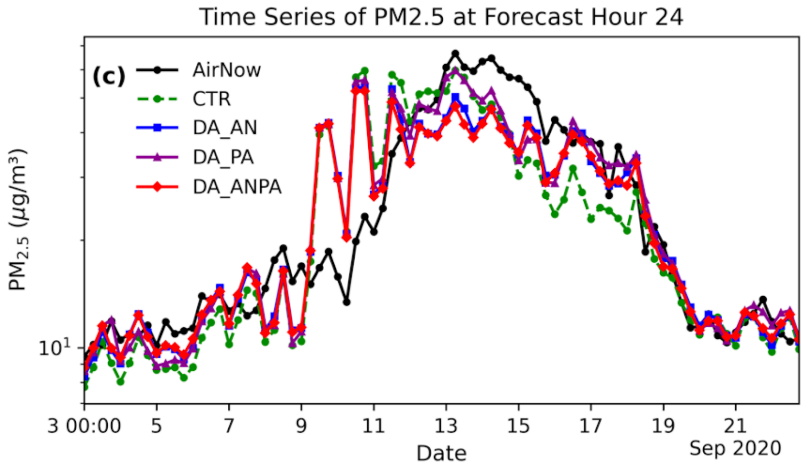


Figure 5. Box-and-whisker plot of $PM_{2.5}$ forecast bias. Orange dot: domain averaged mean; Bottom edge = Q1 (25th percentile); Top edge = Q3 (75th percentile); Height = Interquartile Range (IQR = Q3 − Q1); Horizontal line inside box: The median (50th percentile); Whiskers: Extend to the min and max values within 1.5 × IQR from Q1 and Q3.





426

427 **Figure 6.** Time series of $PM_{2.5}$ averaged over CONUS for (a) forecast hour 1, (b) forecast hour 12, and (c) forecast
428 hour 24. The y-axis is shown on a logarithmic scale.

429 Figure 6 shows time series of $PM_{2.5}$ averaged over CONUS at forecast hours 1, 12, and 24, respectively. Consistent
430 with the evaluations in Figures 4 and 5, all DA experiments generally improve $PM_{2.5}$ forecasts. Notably, all DA
431 experiments help correct underpredictions during September 2–9 and 14–17. In addition, the substantial
432 overprediction during September 9-13 observed in the control run is partially mitigated by the DA experiments.
433 Among the DA experiments, DA_AN and DA_ANPA show comparable performance and both outperform DA_PA.

434 While we have investigated the impact of DA on $PM_{2.5}$ forecasts in terms of temporal evolution, it is also important
435 to examine the spatial distribution of forecast fields, associated errors, and the impact of $PM_{2.5}$ data assimilation.

436 Figure 7 presents the spatial distribution of temporally averaged $PM_{2.5}$ forecasts at forecast hour 1, based on 80
437 forecasts initialized four times daily (0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC) from September 2 to 21.
438 The effects of wildfire events are clearly seen in Fig 7a, c and e across California, Oregon, and Washington—where
439 the fires occurred—as well as in downstream regions such as Montana, Wyoming, Utah, and Colorado impacted by
440 smoke advection and transport.

441 1-h $PM_{2.5}$ forecast errors in the control run are evident in Fig. 6a but are more clearly highlighted in Fig. 7b, which
442 shows the difference between the control run and AirNow observations. Significant overpredictions appear along the
443 California coast, as well as in parts of the Midwest and Northeast U.S., including Tennessee, Kentucky, West
444 Virginia, and Virginia. Conversely, notable underpredictions are found over Colorado, New Mexico, much of Texas
445 and Oklahoma, and several Gulf Coast states.

446 Both DA_AN (Fig. 7c-d) and DA_PA (Fig. 7e-f) show similar spatial correction patterns across California, Oregon,
447 and Washington, particularly in reducing overpredictions along the California coast. They also produce comparable
448 large-scale adjustments across the Northeast, Midwest, and Southern U.S., with error patterns (Fig. 7d and 7f)
449 largely opposite in sign to those in the CTR–AirNow difference (Fig. 7b). This indicates that both DA experiments

effectively mitigate the control run's over- and underpredictions.

However, the magnitude of the corrections is generally smaller in DA_PA than in DA_AN. DA_PA shows its strongest impact over Nevada, northern Utah, Colorado, and southwestern New Mexico, where it helps alleviate regional underpredictions (Fig. 7e and 7f). These improvements are also observed in DA_ANPA, whose spatial pattern closely resembles DA_AN except over these few states (figures not shown).

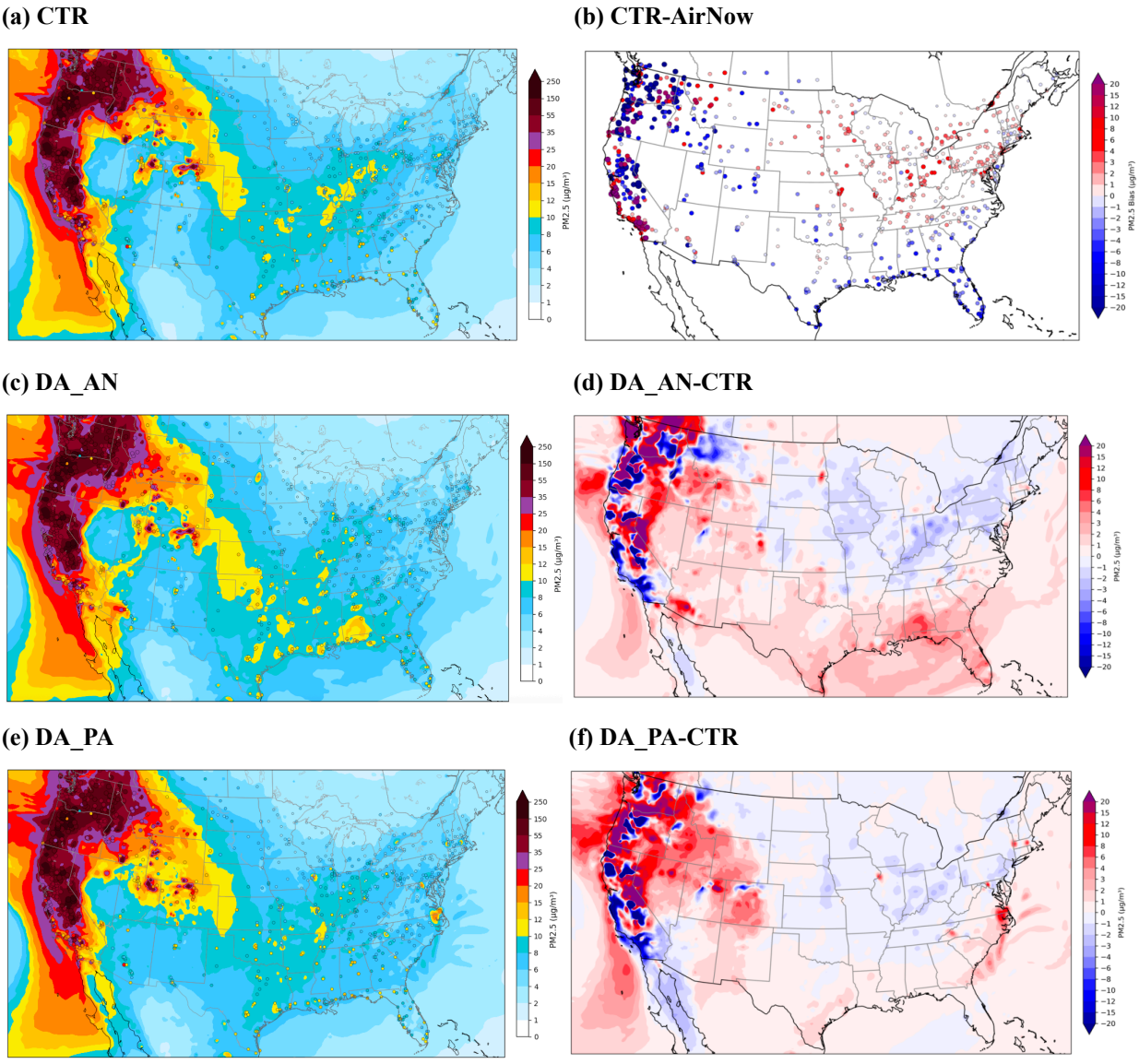


Figure 7. Spatial distribution of average $PM_{2.5}$ at forecast hour 1, based on 80 forecasts initialized four times daily (0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC) during 2–21 September.
(a) $PM_{2.5}$ in experiment CTR (shaded) overlaid with AirNow $PM_{2.5}$ observations (filled dots).
(b) $PM_{2.5}$ bias in experiment CTR.
(c) $PM_{2.5}$ in experiment DA_AN (shaded) overlaid with AirNow $PM_{2.5}$ observations.
(d) $PM_{2.5}$ difference between experiments DA_AN and CTR.

(e) $PM_{2.5}$ in experiment DA_PA (shaded) overlaid with AirNow $PM_{2.5}$ observations.
(f) $PM_{2.5}$ difference between experiments DA_PA and CTR.

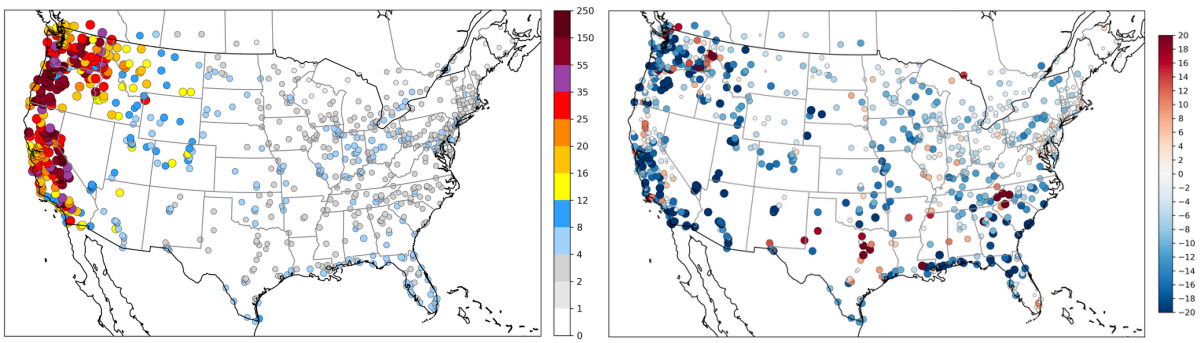
The above analysis shows that data assimilation effectively corrects the 1-h forecast errors in the control run, particularly in the DA_AN and DA_ANPA experiments. It is therefore of interest to examine how data assimilation affects longer forecast lead times. Figure 8 presents the MAE from the CTR experiment and the percentage change in MAE (%) between the data assimilation (DA) experiments and the CTR experiment at the 24-h forecast lead time. Negative values in Figs. 8b–d indicate a reduction in MAE.

It is seen from Fig. 8a that the largest MAE values occur in California, Oregon, Washington, Idaho, Montana, Wyoming, Utah, Colorado, and Arizona. The $PM_{2.5}$ MAE in these states is generally greater than 10 μg m^{-3}, with maximum values exceeding 150 μg m^{-3} at certain stations in California, Oregon, and Washington. In contrast, regions less affected by wildfire smoke exhibit MAE values below 10 μg m^{-3}.

All DA experiments (Figs. 8b–d) show overall reductions in MAE at most stations. MAE reduction varies by location, with substantial improvements observed in regions with large forecast errors, such as California, Oregon, Washington, and Idaho, where MAE is reduced by approximately 20%. On average, assimilation of AirNow $PM_{2.5}$ observations alone significantly improves the 24-h forecast skill, with MAE reductions of 6.6% based on MAE percentage changes averaged over all stations. In contrast, assimilation of PurpleAir data alone reduces the MAE by only 1.7%. Assimilating both AirNow and PurpleAir $PM_{2.5}$ observations reduces the MAE by 7.6%, indicating that PurpleAir data provides complementary value to AirNow in improving forecast skill.

In the DA_PA experiment, a few eastern coastal states from Georgia to Virginia exhibit large percentage increases in MAE. These increases are associated with relatively small absolute errors, as MAE values in the CTR experiment over these states are typically only 2–4 μg m^{-3}(Fig. 8a). Increased MAE is also present at several stations at the 1-h forecast lead time in the DA_PA experiment in these areas (figure not shown). In contrast, the MAE increase at the 24-h forecast lead time is much less pronounced in the experiment assimilating only AirNow $PM_{2.5}$ observations. This behavior suggests potential quality issues in the assimilated PurpleAir data in those aereas that require further investigation, although a contribution from model errors cannot be ruled out.

(a) **(b)**



(c) (d)

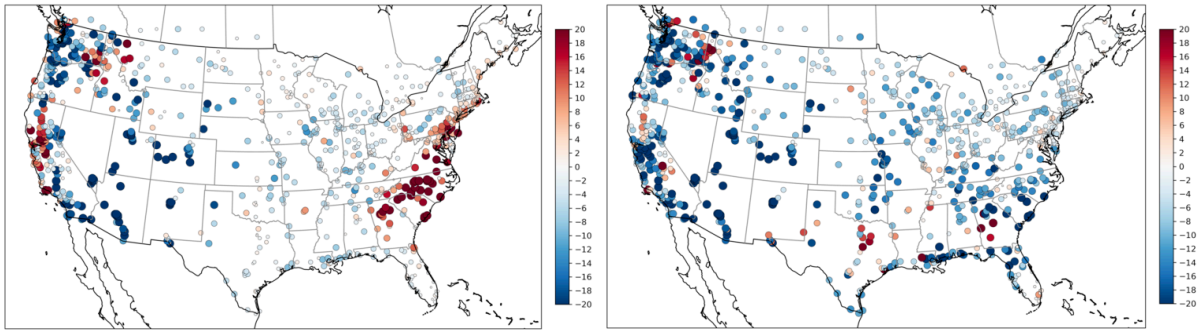


Figure 8. (a) MAE from the control (CTR) experiment. Percentage change in MAE (%) relative to CTR for (b) DA_AN, (c) DA_PA, and (d) DA_ANPA. The percentage change is calculated as (MAE(DA)-MAE(CTR))/MAE(CTR)×100.

To further examine model performance across different regions, error statistics (bias, MAE, and RMSE) for the averaged 1–24 h $PM_{2.5}$ forecasts from the control and data assimilation experiments over the 10 EPA regions are analyzed. Figure 9 presents the EPA Regions 1-10 and Averaged 1–24 h $PM_{2.5}$ forecasts errors statistics in each EPA region.

In terms of bias (Fig. 9b), the DA_AN and DA_ANPA forecasts show improved performance relative to the control run. The DA_PA experiment also improves bias over most EPA regions, except for increased overprediction in EPA regions 8 and 9. For MAE and RMSE, the DA_AN and DA_ANPA experiments generally show improved performance across all EPA regions. The DA_PA experiment also exhibits improved, or at least comparable, MAE relative to the control run. However, DA_PA shows slightly increased RMSE over EPA regions 1–4, while reduced RMSE is found over other EPA regions.

Substantial reductions in MAE and RMSE are observed over EPA regions 8–10, where fire events occurred and/or where regions were most influenced by transported smoke. This is consistent with the large 1-h MAE reduction shown in Fig. 8. Notably, the DA_PA experiment, which assimilates PA observations alone, reduces both MAE and RMSE over EPA regions 8–10. Assimilating PurpleAir data together with AirNow data results in slightly smaller MAE and RMSE over these regions, suggesting that PurpleAir observations nonetheless provide complementary information to AirNow in these regions.

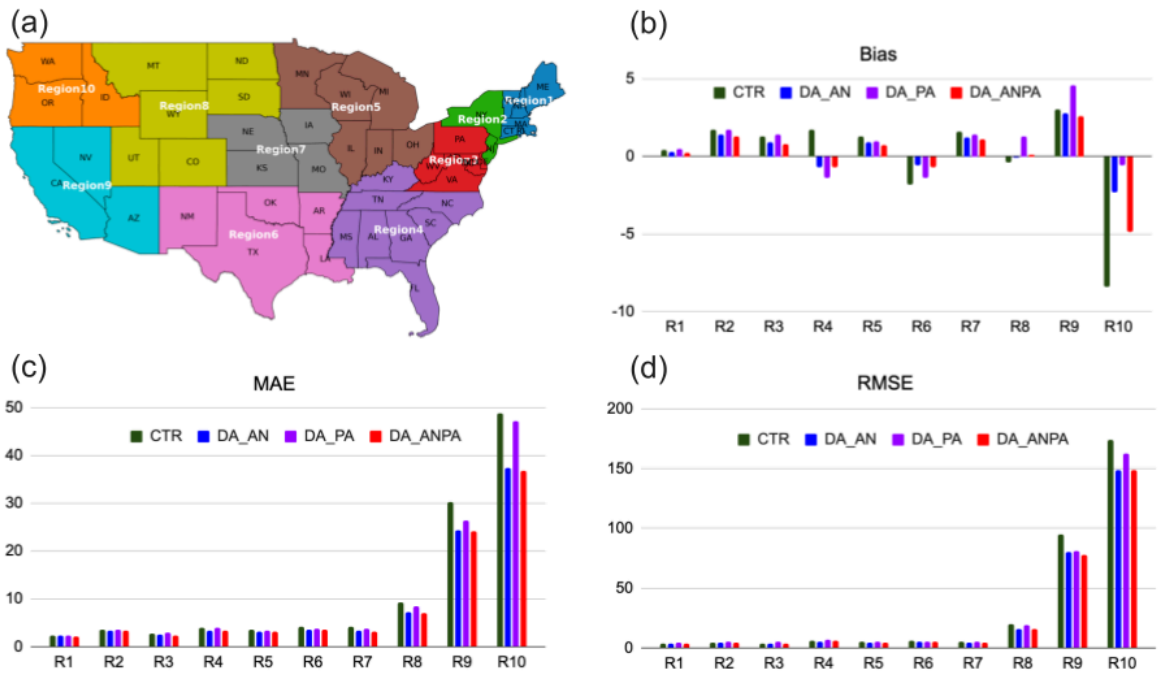


Figure 9. (a) EPA Regions 1-10. Averaged 1–24 h $PM_{2.5}$ forecasts errors: (b) Bias, (c) MAE and (d) RMSE.

5 Summary and discussion

The latest version of NOAA's regional AQM system became operational on May 14, 2024. This system is based on the online coupling of the Finite Volume Cubed Sphere (FV3) atmospheric model with the Environmental Protection Agency (EPA)'s Community Multiscale Air Quality (CMAQ) model within the Unified Forecast System (UFS) framework.

To improve initial conditions for AQM and enhance predictions of wildfire impacts on air quality, the capability to assimilate $PM_{2.5}$ observations into AQMv7 was developed within JEDI and evaluated using its 3D-Var assimilation component in this study. The developed assimilation scheme can also be used to generate analysis (reanalysis) dataset for other applications, for example, providing data for training artificial intelligent models used in air quality prediction.

Data assimilation experiments were conducted for the September 2020 Western U.S. wildfire episode, using 3-hourly cycling with observations from the AirNow and PurpleAir networks. In the data assimilation experiments, the location- and time-dependent background error standard deviations of an aerosol variable are specified as proportional to its background values, using a diagnosed scaling factor. In addition to updating the analyzed aerosol variables in each mode, the particle number concentration and surface area of each mode are also updated. The results show that assimilating AirNow $PM_{2.5}$ observations improves 1–24 h forecast skill. Assimilating PurpleAir data alone yields modest improvements in MAE. Combining PurpleAir with AirNow observations provides additional benefit by slightly further reducing MAE relative to AirNow-only assimilation, indicating that PurpleAir

observations nonetheless provide complementary information to AirNow. The AirNow data assimilation alone or with PurpleAir data generally show reduced MAE and RMSE across all EPA regions, whereas the largest reductions in MAE and RMSE are observed over regions affected by fire events and/or strongly influenced by transported smoke. The positive impact of PurpleAir data assimilation on smoke prediction during the September 2020 wildfires has also been demonstrated in an experimental Rapid Refresh Forecast System coupled with the Smoke and Dust Model (Wang et al., 2023), where it significantly reduced the model's 24-h underprediction of surface $PM_{2.5}$. Given that PurpleAir data coverage has improved since September 2020, the results of this study further highlight its potential to complement AirNow observations by filling spatial gaps and improving $PM_{2.5}$ forecast skill.

In this initial development and evaluation of aerosol data assimilation in JEDI for AQMv7, the control variables are defined as individual forecast aerosol variables. In previous work on aerosol data assimilation for an earlier version of AQM using the GSI system (Wang et al., 2021), one option for the control variables was to define them as the total aerosol mass in each of the three modes, resulting in just three control variables. A control variable transform (CVT) was then applied to partition the analysis increments across these modes to individual aerosol species, based on the ratio of each species' mass to the total mass within the corresponding mode. The use of total aerosol mass in the three modes as control variables—thereby reducing the number of control variables from 70 to 3—is planned for a future phase of development. The use of total masses as control variables also reduces the cost of the background error statistics calculation and iterative minimization (Kumar et al. 2019).

This study focused on surface-level $PM_{2.5}$ and did not incorporate vertical profile constraints with satellite-based aerosol optical depth (AOD) retrievals, which could further enhance forecast skill. A key challenge is the need for a robust forward operator in the CRTM AOD module—specifically, the creation and validation of lookup tables (LUTs) for AOD calculations with AQM. As an intermediate solution, existing LUTs in CRTM, such as the GEOS-5 LUTs, have been tested by grouping and mapping AQM aerosol species to those used in GEOS-5 (Wang et al. 2025). AOD assimilation also depends on an accurate vertical distribution of aerosols in the background field so that the CRTM AOD operator can provide meaningful gradient information at the correct vertical levels to constrain the analysis update. However, AQM models have shown deficiencies for the September 2020 fire events in representing smoke concentrations at and above plume rise levels, largely due to how fire emissions are injected into the model. This will be improved in the next update of the operational AQM.

Code and data availability

The AQMV7 model, JEDI software and $PM_{2.5}$ and fire emission data we used in this research are publicly available on on Zenodo (https://doi.org/10.5281/zenodo.17049857; Wang et al., 2025b).

Users are referred to the guidance on compiling and running the model:
https://ufs-srweather-app.readthedocs.io/en/develop/UsersGuide/index.html (Last accessed on August 26, 2025).

Global Forecast System analysis data were downloaded from the NCAR Research Data Archive:
https://doi.org/10.5065/D65D8PWK (last access: Aug 26 2025)

Author contribution

HW designed and developed the $PM_{2.5}$ DA capability within JEDI for the AQM model, conducted experiments, and evaluated performance; CM and JB contributed to $PM_{2.5}$ DA methodology, advised on code implementation, and assisted in performance analysis; SW contributed to $PM_{2.5}$ DA methodology and experimental design; RL conducted control experiments and contributed to workflow development; JL and KW contribute to model configuration and control run setup; YT contributed to background error modeling and observational error specification; HC, AT and HL contributed to workflow development; JL performed quality control and correction of PurpleAir observations.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors sincerely thank Dr. Ming Hu and the three anonymous reviewers for their constructive comments and insightful suggestions, which significantly improved the quality and clarity of this manuscript. The authors also thank Dr. Mohmmed Farooqui at Texas A&M University-Kingsville for assisting with Python scripts to download the PurpleAir observations.

This research was supported by the Fire Weather and Precipitation Research and Development in Support of the Disaster Relief Supplemental Appropriations Act (DRSA) project (NA23OAR4050200D), and in part by a NOAA Cooperative Agreement NA22OAR4320151 with the University of Colorado. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect those of NOAA or the Department of Commerce.

References

Abatzoglou, J. T., Rupp, D. E., O'Neill, L. W., & Sadegh, M. (2021). Compound extremes drive the western Oregon wildfires of September 2020. *Geophysical Research Letters*, 48(8), e2021GL092520. https://doi.org/10.1029/2021gl092520

Albores, I. S., Buchholz, R. R., Ortega, I., Emmons, L. K., Hannigan, J. W., Lacey, F., et al. (2023). Continental-scale atmospheric impacts of the 2020 western US wildfires. *Atmospheric Environment*, 294, 119436. https://doi.org/10.1016/j.atmosenv.2022.119436

Barkjohn, K. K., Gantt, B., and Clements, A. L.: Development and application of a United States-wide

correction for $PM_{2.5}$ data collected with the PurpleAir sensor, Atmos. Meas. Tech., 14, 4617–4637, https://doi.org/10.5194/amt-14-4617-2021, 2021.DOI: 10.5194/amt-2020-413

Black, T. L., Abeles, J. A., Blake, B. T., Jovic, D., Rogers, E., Zhang, X., et al. (2021). A limited area modeling capability for the finite-volume cubed-sphere (FV3) dynamical core and comparison with a global two-way nest. *Journal of Advances in Modeling Earth Systems*, 13(6), e2021MS002483. https://doi.org/10.1029/2021MS002483

Binkowski, F. S., and S. J. Roselle, Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component, 1, Model description, J. Geophys. Res., 108(D6), 4183, doi:10.1029/2001JD001409, 2003.

Chen L, Mao F, Hong J, Zang L, Chen J, Zhang Y, Gan Y, Gong W, Xu H. Improving $PM_{2.5}$ predictions during COVID-19 lockdown by assimilating multi-source observations and adjusting emissions. Environ Pollut. 2022 Mar 15;297:118783. doi: 10.1016/j.envpol.2021.118783. Epub 2021 Dec 30. PMID: 34974086; PMCID: PMC8717716.

Chen, X., Zhang, Y., Wang, K., Tong, D., Lee, P., Tang, Y., Huang, J., Campbell, P. C., Mcqueen, J., Pye, H. O. T., Murphy, B. N., and Kang, D.: Evaluation of the offline-coupled GFSv15–FV3–CMAQv5.0.2 in support of the next-generation National Air Quality Forecast Capability over the contiguous United States, Geosci. Model Dev., 14, 3969–3993, https://doi.org/10.5194/gmd-14-3969-2021, 2021.

Colette, A., Collin, G., Besson, F., and coauthors.: Copernicus Atmosphere Monitoring Service – Regional Air Quality Production System v1.0, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-3744, 2024.

Derber, J. C., and A. Rosati, 1989*:* A global ocean data assimilation system. *J. Phys. Oceanogr.*, ***19***, 1333–1347, https://doi.org/10.1175/1520-0485(1989)019<1333:AGODAS>2.0.CO;2.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope III, C. A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C. J. L., and Forouzanfar, M. H.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, *Lancet*, **389**, 1907–1918, 2017.

Colmer, J., Hardman, I., Shimshack, J., and Voorheis, J.: Disparities in $PM_{2.5}$ air pollution in the United States, *Science*, 369(6503), 575–578, https://doi.org/10.1126/science.aaz9353, 2020.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli. 2005. Diagnostic of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc*. 131:3385–96. doi:10.1256/qj.05.108.doi:10.1256/qj.05.108

Ha, S.: Implementation of aerosol data assimilation in WRFDA (v4.0.3) for WRF-Chem (v3.9.1) using the RACM/MADE-VBS scheme, Geosci. Model Dev., 15, 1769–1788, https://doi.org/10.5194/gmd-15-1769-2022, 2022.

631 Hollingsworth, A., and P. Lönnberg. 1986. The statistical structure of short-range forecast errors as determined
632 from radiosonde data. Part I: The wind field. Tellus A38:111–36. doi:10.1111/j.1600-0870.1986.tb00460.x
633 Huang, J., I. Stajner, R. Montuoro, F. Yang, K. Wang, H.-C. Huang, C.-H. Jeon, B. Curtis, J. McQueen, H. Liu,
634 B. Baker, D. Tong, Y. Tang, P. Campbell, G. Grell, G. Frost, R. Schwantes, S. Wang, S. Kondragunta, F. Li, Y.
635 Jung. Development of the next-generation air quality prediction system in the unified forecast system
636 framework: enhancing predictability of wildfire air quality impacts. Bull. Am. Meteorol. Soc. (2025),
637 10.1175/BAMS-D-23-0053.1
638
639 Environmental Protection Agency. Technical Note on Reporting $PM_{2.5}$ Continuous Monitoring and Speciation
640 Data to the Air Quality System (AQS). November 8, 2006.
641 https://www.epa.gov/aqs/aqs-memos-technical-note-reporting-pm25-continuous-monitoring-and-speciation-data
642 -air-quality.
643 EPA (2017). Wildfire and Air Quality. U.S. Environmental Protection Agency.
644 Huang, B., Pagowski, M., Trahan, S., Martin, C. R., Tangborn, A., Kondragunta, S., & Kleist, D. T. (2023).
645 JEDI-based three-dimensional Ensemble-Variational Data Assimilation System for global aerosol forecasting at
646 NCEP. *Journal of Advances in Modeling Earth Systems*, 15(4), e2022MS003232.
647 https://doi.org/10.1029/2022MS003232

648 Lee Sojin, Chul Han Song, Kyung Man Han, Daven K. Henze, Kyunghwa Lee, Jinhyeok Yu, Jung-Hun Woo,
649 Jia Jung, Yunsoo Choi, Pablo E. Saide, Gregory R. Carmichael, Impacts of uncertainties in emissions on aerosol
650 data assimilation and short-term PM2.5 predictions over Northeast Asia, Atmospheric Environment, Volume
651 271, 2022, 118921, ISSN 1352-2310, https://doi.org/10.1016/j.atmosenv.2021.118921.

652 Li, Y., Tong, D., Ma, S., Zhang, X., Kondragunta, S., Li, F., & Saylor, R. (2021). Dominance of wildfires impact
653 on air quality exceedances during the 2020 record-breaking wildfire season in the United States. *Geophysical*
654 *Research Letters*, 48(21), e2021GL094908. https://doi.org/10.1029/2021GL094908

655 Li, Z., Zang, Z., Li, Q. B., Chao, Y., Chen, D., Ye, Z., Liu, Y., and Liou, K. N.: A three-dimensional variational
656 data assimilation system for multiple aerosol species with WRF/Chem and an application to PM2.5 prediction,
657 Atmos. Chem. Phys., 13, 4265–4278, https://doi.org/10.5194/acp-13-4265-2013, 2013.

658 Liu, Z., Snyder, C., Guerrette, J. J., Jung, B.-J., Ban, J., Vahl, S., Wu, Y., Trémolet, Y., Auligné, T., Ménétrier,
659 B., Shlyaeva, A., Herbener, S., Liu, E., Holdaway, D., and Johnson, B. T.: Data assimilation for the Model for
660 Prediction Across Scales – Atmosphere with the Joint Effort for Data assimilation Integration (JEDI-MPAS
661 1.0.0): EnVar implementation and evaluation, Geosci. Model Dev., 15, 7859–7878,
662 https://doi.org/10.5194/gmd-15-7859-2022, 2022.

663 Kumar, R., Monache, L. D., Bresch, J., Saide, P. E., Tang, Y., Liu, Z., Silva, A. M. da, Alessandrini, S., Pfister,
664 G., Edwards, D., Lee, P., and Djalalova, I.: Toward Improving Short-Term Predictions of Fine Particulate Matter

Over the United States Via Assimilation of Satellite Aerosol Optical Depth Retrievals, J. Geophys. Res.-Atmos., 124, 2753–2773, https://doi.org/10.1029/2018JD029009, 2019.

Mass, C.F.; Ovens, D.; Conrick, R.; Saltenberger, J. The September 2020 wildfires over the Pacific Northwest. *Weather Forecast.* **2022**, *36*, 1843–1865.

Ménard, R., Deshaies-Jacques, M., & Gasset, N. (2016). A comparison of correlation-length estimation methods for the objective analysis of surface pollutants at Environment and Climate Change Canada. *Journal of the Air & Waste Management Association*, *66*(9), 874–895. https://doi.org/10.1080/10962247.2016.1177620

O'Dell, K. et al. (2019). Open-access study on wildfire smoke and $PM_{2.5}$ trends in the western U.S.

Robichaud, A. (2017). Surface data assimilation of chemical compounds over North America and its impact on air quality and Air Quality Health Index (AQHI) forecasts. Air Quality, Atmosphere & Health, 10(8), 955–970. https://doi.org/10.1007/s11869-017-0485-9

National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce: NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory [data set], https://doi.org/10.5065/D65D8PWK (last access: 26 August 2025).

Robichaud A, Ménard R, Zaïtseva Y, Anselmo D. Multipollutant surface objective analyses and mapping of air quality health index over North America. Air Qual Atmos Health. 2016;9(7):743-59. doi: 10.1007/s11869-015-0385-9.

Pagowski, M., G. A. Grell, S. A. McKeen, S. E. Peckham, and D. Devenyi (2010), Three-dimensional variational data assimilation of ozone and fine particulate matter observations: Some results using the Weather Research and Forecasting–Chemistry model and grid-point statistical interpolation, *Q. J. R. Meteorol. Soc.*, 136, 2013–2024, doi:10.1002/qj.700.

Pagowski, M., Liu, Z., Grell, G. A., Hu, M., Lin, H.-C., and Schwartz, C. S.: Implementation of aerosol assimilation in Gridpoint Statistical Interpolation (v. 3.2) and WRF-Chem (v. 3.4.1), Geosci. Model Dev., 7, 1621–1627, https://doi.org/10.5194/gmd-7-1621-2014, 2014.

Park, S.-Y., Dash, U. K., Yu, J., Yumimoto, K., Uno, I., and Song, C. H.: Implementation of an ensemble Kalman filter in the Community Multiscale Air Quality model (CMAQ model v5.1) for data assimilation of ground-level PM2.5, Geosci. Model Dev., 15, 2773–2790, https://doi.org/10.5194/gmd-15-2773-2022, 2022.

Safford, H. D., Paulson, A. K., Steel, Z. L., Young, D. J. N., & Wayman, R. B. (2022). The 2020 California fire season: A year like no other, a return to the past or a harbinger of the future? *Global Ecology and Biogeography*, 31(10), 2005–2025. https://doi.org/10.1111/geb.13498

Sluka Travis. Generic Explicit Diffusion Operator Added to JEDI. JCSDA News Letter. No. 74, Fall 2024.

Schwartz C.S., Z. Liu, H.C. Lin, S.A. McKeen. Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth. J. Geophys. Res. Atmos., 117 (2012), 10.1029/2011jd017383

Sun W. , Z. Liu, D. Chen, P. Zhao, M. Chen. Development and application of the WRFDA-Chem three-dimensional variational (3DVAR) system: aiming to improve air quality forecasting and diagnose model deficiencies. Atmos. Chem. Phys., 20 (2020), pp. 9311-9329, 10.5194/acp-20-9311-2020

Tang Y., and coauthors, 2023-1-11: *Develop and Evaluate JEDI-Based Regional Aerosol Data Assimilation for NOAA UFS-AQM System.* The 103rd AMS Annual Meeting, Denver, Colorado.

Trémolet, Y. and Auligné, T.: The Joint Effort for Data Assimilation Integration (JEDI), JCSDA Quarterly Newsletter, 66, 1–5, https://doi.org/10.25923/RB19-0Q26, 2020.

Vogel, A., Ménard, R., Abu, J., and Chen, J.: Potential of error-evolving tracer forecasts for operational assimilation of PM2.5 during wildfire smoke episodes, EGU General Assembly 2025, Vienna, Austria, 27 Apr–2 May 2025, EGU25-9428, https://doi.org/10.5194/egusphere-egu25-9428, 2025.

Wang. H. and coauthors. 2021-9-13. Assimilation of Aerosol Optical Depth (AOD) retrievals and $PM_{2.5}$ in NCEP's Next-Generation Regional Air Quality Forecasting System. WCRP-WWRP Symposium on DA and reanalysis. 2021.

Wang, H, and coauthors, 2023-5-16: Assimilation of Surface Particulate Matter Observations in the experimental Rapid Refresh Forecast System coupled with Smoke and Dust Model. CU/CIRES Rendezvous 2023, Boulder, Colorado. (https://insidecires.colorado.edu/rendezvous/uploads/Rendezvous_2023_7732_1683821981.pdf, last access, July 19 2025)

Wang, H, and coauthors, 2025-01-15: Aerosol Data Assimilation within JEDI for the NOAA's Regional Air Quality Model(AQM). The 105th Annual Meeting of the American Meteorological Society, New Orleans, LA.

Wei Y., X. Zhao, Z. Zhang, J. Xu, S. Cheng, Z. Liu, W. Sun, X. Chen, Z. Wang, X. Hao, J. Li, D. Chen. Impact of model resolution and its representativeness consistency with observations on operational prediction of PM2.5 with 3D-VAR data assimilation. Atmos. Pollut. Res., 15 (2024), Article 102141, 10.1016/j.apr.2024.102141

Wang, H., Martin, C., Barré, J., Li, R., Weygandt, S., Huang, J., Tang, Y., Choi, H., Wang, K., Liu, H., & Lee, J. (2025). $PM_{2.5}$ Assimilation within JEDI for NOAA's Regional Air Quality Model (AQMv7): Application to the September 2020 Western U.S. Wildfires [Data set]. Zenodo. https://doi.org/10.5281/zenodo.17049857

Wen, J., M. Burke. Wildfire smoke plume segmentation using geostationary satellite imagery. arXiv. https://doi.org/10.48550/arXiv.2109.01637. 2021.

Zhang, H., Yee, L. D., Lee, B. H., Curtis, M. P., Worton, D. R., Isaacman-VanWertz, G., Offenberg, J. H., Lewandowski, M., Kleindienst, T. E., Beaver, M. R., Holder, A. L., Lonneman, W. A., Docherty, K. S., Jaoui, M., Pye, H. O. T., Hu, W., Day, D. A., Campuzano-Jost, P., Jimenez, J. L., Guo, H., Weber, R. J., de Gouw, J., Koss, A. R., Edgerton, E. S., Brune, W., Mohr, C., Lopez-Hilfiker, F. D., Lutz, A., Kreisberg, N. M., Spielman, S. R., Hering, S. V., Wilson, K. R., Thornton, J. A. and Goldstein, A. H.: Monoterpenes are the largest source of summertime organic aerosol in the southeastern United States. Proc. Natl. Acad. Sci., 115, 2038-2043, doi.org/10.1073/pnas.1717513115, 2018.