General comments:

This project aims to assess the effectiveness of several ongoing bias-correction approaches funded by NOAA and to identify a potential replacement for the existing bias-correction approach used in NOAA's operational air quality models (AQMs). A bias-correction technique, the Kalman Filter Analog Ensemble (KFAN), developed by OAR/PSL, has been implemented by AQMs since 2016 to improve air quality forecasts for PM2.5 and O3. Overall, the KFAN approach has successfully mitigated most symmetric biases across different regions and seasons. However, it struggles to accurately predict high concentrations of PM2.5 and O3 during air-quality exceedance episodes such as wildfires and windblown dust events. This limitation stems from operational timing constraints, which allow for only limited training data utilization. Additionally, the number of analogs is another critical parameter in analog-ensemble-related bias-correction approaches that current operational models use. Having too many analog members may hurt peak value predictions due to constraints associated with the ensemble mean used in current bias-correction products. Therefore, it is imperative to seek alternative bias-correction approaches to further improve air quality forecast products, particularly for high-impact air quality events.

Specific comments:

1. It is crucial to provide background information on the bias-correction approaches evaluated in this project. Specifically, how many bias-correction approaches will be assessed, and what are their features? How do these candidates address the bottleneck issues encountered by the current bias-correction method (i.e., analog ensemble)?
2. Please make sure that the ongoing bias-correction development projects are able to provide bias-correction results for inter-comparison when this project starts. This is perhaps my most concern as compared to the evaluation tool(s).
3. Regarding the evaluation tool(s), if EVS or Metplus is difficult for the project to pick up, MELODIES-MONET could be an alternative option for completing inter-comparison tasks. MELODIES-MONET has been widely utilized to evaluate AQMv7 performance during model development.
4. Consistency of data files and formats across different bias-correction projects is another concern. EVS requires grib2-format input data, while it's unclear if Metplus has the same requirements. Original bias-correction methods typically provide hourly data in netcdf format. This requires further development of post-processing code to convert bias-correction data from netcdf to grib2 format. Again, MELODIES-MONET does not require grib2 data files.
5. Further details are needed regarding the inter-comparison tasks proposed by the project.

5.1) Evaluation metrics: Time series comparison including diurnal variation patterns, traditional verification metrics such as RMSE, correlation coefficients, mean biases, etc., and forecast skill scores such as CSI, POD, FAR, etc.

5.2) Verification time periods (winter, summer) and regions (EUS, WUS, rural, urban, VOC-limited, and NOx-limited formation regimes for O3) should be specified.

5.3) How do these new bias correction methods outpace the currently operational BC method in addressing high-impact air quality events such as intense wildfires and dust storms?