Associations between optical, physical and chemical properties of aerosols measured at ground-based networks

Meredith Franklin[1], Meytar Sorek-Hamer[2], Olga Kalashnikova[3], Michael Garay[3], and David Diner[3]

[1]University of Southern California, Los Angeles, CA, United States
[2]NASA Ames Research Center, Moffett Field, CA, USA
[3]Jet Propulsion Laboratory, Pasadena, CA, USA

Correspondence: Meredith Franklin (meredith.franklin@usc.edu)

Abstract. The abstract goes here. It can also be on *multiple lines*.

1 Introduction

A large body of literature has shown that satellite-derived aerosol optical depth (AOD), typically retrieved at 550 nm wavelength, reliably correlates with mass-volume concentrations of fine mode particulate matter with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$). Studies that have used satellite observations to generate $PM_{2.5}$ have been instrumental for air pollution - health effects research. The associations between AOD and different chemical components of $PM_{2.5}$ are lesser known. A handful of studies using the Multiangle Imaging SpectroRadiometer (MISR), an instrument onboard the NASA Terra satellite that provides observations of optical properties by particle type (size, shape, absorption), have provided evidence that different optical properties relate to different physical and chemical properties of particulate matter. Results have been somewhat inconsistent, showing differences depending on geographic area of analysis, optical components used, and statistical tools applied.

A large body of literature has shown that satellite-derived aerosol optical depth (AOD), typically retrieved at 550 nm wavelength, reliably correlates with mass-volume concentrations of fine mode particulate matter with aerodynamic diameter less than 2.5 um ($PM_{2.5}$). Studies that have used satellite observations to generate PM2.5 have been instrumental for air pollution - health effects research. The associations between AOD and different chemical components of PM2.5 are lesser known. A handful of studies using the Multiangle Imaging SpectroRadiometer (MISR), an instrument onboard the NASA Terra satellite that provides observations of optical properties by particle type (size, shape, absorption), have provided evidence that different optical properties relate to different physical and chemical properties of particulate matter. Results have been somewhat inconsistent, showing differences depending on geographic area of analysis, optical components used, and statistical tools applied.

The purpose of this analysis is to make a detailed examination of the statistical relationships between ground-level PM2.5 and PM2.5 chemical components (nitrate, sulfate, elemental carbon, organic carbon, dust) and optical measures of aerosols (e.g. aerosol optical depth, angstrom exponent).

AERONET Holben et al. (1998); **?**); Shin et al. (2019a, b)

– Specific sections of the manuscript:

  – `running` with `title` and `author`
  – `competinginterests`
  – `copyrightstatement` (optional)
  – `availability` (strongly recommended if any used), one of `code`, `data`, or `codedata`
  – `authorcontribution`
  – `disclaimer`
  – `acknowledgements`

See the defaults and examples from the skeleton and the official Copernicus documentation for details.

Important: Always double-check with the official manuscript preparation guidelines at https://publications.copernicus.org/for_authors/manuscript_preparation.html, especially the sections "Technical instructions for LaTeX" and "Manuscript composition". Please contact Daniel Nüst, `daniel.nuest@uni-muenster.de`, with any problems.

2 Methods

The study encompasses the San Joaquin Valley region of central California (Figure 1). We leverage datasets available at four sites - Bakersfield, Fresno, Modesto, and Visalia. At these sites there are co-located instruments from EPA's chemical speciation network (CSN), EPA's air quality system (AQS) and NASA's AERONET network.

2.1 Data

The CSN monitors are on a 1-in-3 or 1-in-6 day sampling schedule, providing $PM_{2.5}$mass and component $PM_{2.5}$ concentrations of metals (e.g. Aluminium Al, Silicon Si, Calcium Ca, Titanium Ti, Iron Fe) obtained from X-ray fluorescence (XRF), ions (nitrate NO_3^- and sulfate SO_4^{2-}) from ion chromatography, and carbons (organic OC and elemental EC) from thermal/optical analysis. To quantify dust we use the following equation **?**: dust = 2.2Al + 2.49Si + 1.63Ca + 1.94Ti + 2.42Fe

The AQS monitors provide daily concentrations of $PM_{2.5}$ mass by the EPA's Federal Reference Method, which is the highest quality gravimetric measurement method used for regulatory purposes.

AERONET sites are sunphotometers providing a "ground-up" measurement of aerosol optical properties at multiple wavelengths and have been used extensively to validate "top-down" satellite observations of related properties. Wavelength-specific

AOD and angstrom exponents are the primary sunphotometer variables. Using quadratic log-log interpolation we calculated AOD 550 nm from AOD 440, 500, 675, 870 nm in log-log space. AOD at 550 nm is the most common wavelength retrieved from satellite instruments. A retrieval-based AERONET product, called the inversion product, provides an additional suite of aerosol properties that help distinguish size (fine, coarse effective radius), shape (asymmetry), and absorption. We excluded sunphotometer and inversion variables that had a significant proportion of missing data (~90% missing). A list of the variables included in the analysis are shown in the Appendix. In a separate test we examine data from the SPARTAN site in Rehovot, Israel. The SPARTAN network provides data for $PM_{2.5}$ mass and speciation concentrations on an integrated 1 in 9 day sampling schedule, and is colocated with an AERONET site (We don't have the speciation data for this site so we could only look at $PM_{2.5}$ for now).

The CSN monitors are on a 1 in 3 or 1 in 6 day sampling schedule, providing PM2.5 mass and component PM2.5 concentrations of metals (e.g. Aluminium Al, Silicon Si, Calcium Ca, Titanium Ti, Iron Fe) obtained from X-ray fluorescence (XRF), ions (nitrate NO_3^- and sulfate SO_4^{2-}) from ion chromatography, and carbons (organic OC and elemental EC) from thermal/optical analysis. To quantify dust we use the following equation: dust = 2.2Al + 2.49Si + 1.63Ca + 1.94Ti + 2.42Fe

The AQS monitors provide daily concentrations of PM2.5 mass by the EPA's Federal Reference Method, which is the highest quality gravimetric measurement method used for regulatory purposes.

AERONET sites are sunphotometers providing a "ground-up" measurement of aerosol optical properties at multiple wavelengths and have been used extensively to validate "top-down" satellite observations of related properties. Wavelength-specific AOD and angstrom exponents are the primary sunphotometer variables. Using quadratic log-log interpolation we calculated AOD 550 nm from AOD 440, 500, 675, 870 nm in log-log space. AOD at 550 nm is the most common wavelength retrieved from satellite instruments. A retrieval-based AERONET product, called the inversion product, provides an additional suite of aerosol properties that help distinguish size (fine, coarse effective radius), shape (asymmetry), and absorption. We excluded sunphotometer and inversion variables that had a significant proportion of missing data (~90% missing). A list of the variables included in the analysis are shown in the Appendix. In a separate test we examine data from the SPARTAN site in Rehovot, Israel. The SPARTAN network provides data for PM2.5 mass and speciation concentrations on an integrated 1 in 9 day sampling schedule, and is colocated with an AERONET site (We don't have the speciation data for this site so we could only look at PM2.5 for now).

2.2 Statistical methods

Prior to model building we examined a cluster-based correlation heat map (Figure 2), which provides the Pearson correlations between all pairs of AERONET variables grouped by a decision tree. To avoid collinearity in the regression models, we kept the most relevant of a group of variables that had a correlation coefficient > |0.9|. We then examined and picked a subset of variables connected at the mid-tier level of the tree to construct interactions. We fit simple linear regression models separately for PM2.5 mass, sulfate, nitrate, EC, OC, and dust with AOD 550 nm as the sole predictor variable. Multiple linear regression models were again fit separately for PM2.5 mass, sulfate, nitrate, EC, OC, and dust, but with the combined AERONET sunphotometer and inversion product as predictor variables and model selection was conducted using the "all possible subset method". This method constructs models based on all combinations from 1 to k variable models. We select the best model from the combinations based on highest R2, lowest RMSE, and Mallow's Cp statistic that is close to k+1. Model selection for the Fresno and Bakersfield sites were examined separately and in combination in a "total CA" analysis, which combined data from Fresno, Bakersfield, Modesto, Visalia, and a special DRAGON campaign in late 2012-early 2013 over the region (8 co-located sites with PM2.5 mass).

All models were cross validated (CV) with 10-fold CV, and we report the CV R2 and RMSE. Models were fit in R using the leaps() library.

3 Results

3.1 Statistical methods

Prior to model building we examined a cluster-based correlation heat map (Figure 2), which provides the Pearson correlations between all pairs of AERONET variables grouped by a decision tree. To avoid collinearity in the regression models, we kept the most relevant of a group of variables that had a correlation coefficient $> |0.9|$. We then examined and picked a subset of variables connected at the mid-tier level of the tree to construct interactions. We fit simple linear regression models separately for PM2.5 mass, sulfate, nitrate, EC, OC, and dust with AOD 550 nm as the sole predictor variable. Multiple linear regression models were again fit separately for PM2.5 mass, sulfate, nitrate, EC, OC, and dust, but with the combined AERONET sunphotometer and inversion product as predictor variables and model selection was conducted using the "all possible subset method". This method constructs models based on all combinations from 1 to k variable models. We select the best model from the combinations based on highest R2, lowest RMSE, and Mallow's Cp statistic that is close to k+1. Model selection for the Fresno and Bakersfield sites were examined separately and in combination in a "total CA" analysis, which combined data from Fresno, Bakersfield, Modesto, Visalia, and a special DRAGON campaign in late 2012-early 2013 over the region (8 co-located sites with PM2.5 mass).

All models were cross validated (CV) with 10-fold CV, and we report the CV R2 and RMSE. Models were fit in R using the leaps() library.

4 Results

5 Content section with R code chunks

You should always use `echo = FALSE` on R Markdown code blocks as they add formatting and styling not desired by Copernicus. The hidden workflow results in 42.

You can add verbatim code snippets without extra styles by using `` ``` `` without additional instructions.

```
sum <- 1 + 41
```

6 Content section with list

If you want to insert a list, you must

- leave
- empty lines
- between each list item

because the `\tightlist` format used by R Markdown is not supported in the Copernicus template. Example:

```
- leave

- empty lines

- between each list item
```

7 Examples from the official template

7.1 FIGURES

When figures and tables are placed at the end of the MS (article in one-column style), please add

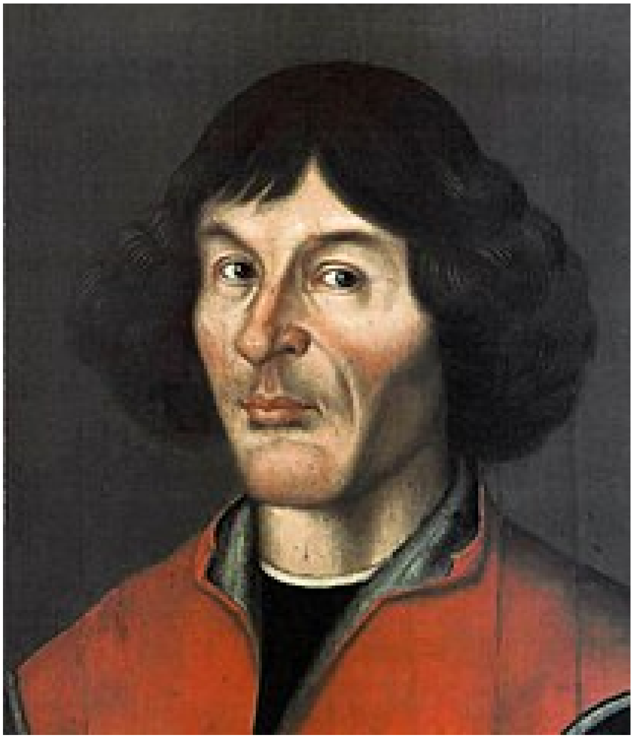

Figure 1. one column figure

between bibliography and first table and/or figure as well as between each table and/or figure.

7.1.1 ONE-COLUMN FIGURES

Include a 12cm width figure of Nikolaus Copernicus from Wikipedia with caption using R Markdown.

7.1.2 TWO-COLUMN FIGURES

You can also include a larger figure.

7.2 TABLES

You can ad LaTeXtable in an R Markdown document to meet the template requirements.

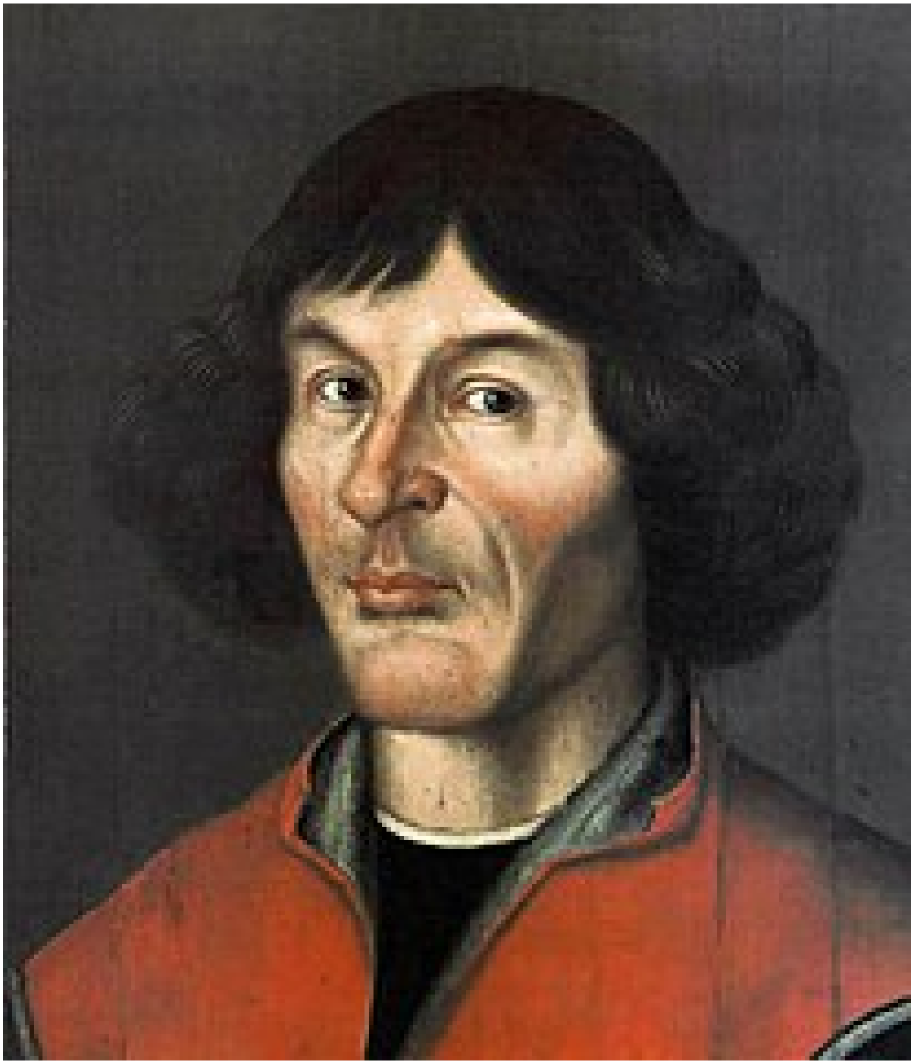

Figure 2. two column figure

Table 1. TEXT

a	b	c
1	2	3

Table Footnotes

Table 2. TEXT

a	b	c
1	2	3

Table footnotes

7.2.1 ONE-COLUMN TABLE

7.2.2 TWO-COLUMN TABLE

7.3 MATHEMATICAL EXPRESSIONS

All papers typeset by Copernicus Publications follow the math typesetting regulations given by the IUPAC Green Book (IUPAC: Quantities, Units and Symbols in Physical Chemistry, 2nd Edn., Blackwell Science, available at: http://old.iupac.org/publications/boo 1993).

Physical quantities/variables are typeset in italic font (t for time, T for Temperature)

Indices which are not defined are typeset in italic font (x, y, z, a, b, c)

Items/objects which are defined are typeset in roman font (Car A, Car B)

Descriptions/specifications which are defined by itself are typeset in roman font (abs, rel, ref, tot, net, ice)

Abbreviations from 2 letters are typeset in roman font (RH, LAI)

Vectors are identified in bold italic font using $\boldsymbol{x}$

Matrices are identified in bold roman font

Multiplication signs are typeset using the LaTeX commands `\times` (for vector products, grids, and exponential notations) or `\cdot`

The character * should not be applied as mutliplication sign

7.4 EQUATIONS

7.4.1 Single-row equation

Unnumbered equations (i.e. using `$$` and getting inline preview in RStudio) are not supported by Copernicus.

$$1 \times 1 \cdot 1 = 42 \tag{1}$$

$$A = \pi r^2 \tag{2}$$

$$x = \frac{2b \pm \sqrt{b^2 - 4ac}}{2c}. \tag{3}$$

7.4.2 Multiline equation

$$3 + 5 = 8 \tag{4}$$

$$3 + 5 = 8 \tag{5}$$

$$3 + 5 = 8 \tag{6}$$

7.5 MATRICES

$$\begin{matrix} x & y & z \\ x & y & z \\ x & y & z \end{matrix}$$

7.6 ALGORITHM

If you want to use algorithms, you can either enable the required packages in the header (the default, see `algorithms: true`), or make sure yourself that the LaTeXpackages `algorithms` and `algorithmicx` are installed so that `algorithm.sty` respectively `algorithmic.sty` can be loaded by the Copernicus template. Copernicus staff will remove all undesirable packages from your LaTeX source code, so please stick to using the header option, which only adds the two acceptable packages.

7.7 CHEMICAL FORMULAS AND REACTIONS

For formulas embedded in the text, please use `\chem{}`, e.g. $\mathrm{A \rightarrow B}$.

The reaction environment creates labels including the letter R, i.e. (R1), (R2), etc.

Algorithm 1 Algorithm Caption

$i \leftarrow 10$
if $i \geq 5$ **then**
  $i \leftarrow i-1$
else
  if $i \leq 3$ **then**
    $i \leftarrow i+2$
  end if
end if

- `\rightarrow` should be used for normal (one-way) chemical reactions
- `\rightleftharpoons` should be used for equilibria
- `\leftrightarrow` should be used for resonance structures

$$A \rightarrow B \tag{R1}$$

$$Coper \rightleftharpoons nicus \tag{R2}$$

$$Publi \leftrightarrow cations \tag{R3}$$

7.8 PHYSICAL UNITS

Please use `\unit{}` (allows to save the math/`$` environment) and apply the exponential notation, for example $3.14\,\mathrm{km\,h^{-1}}$ (using LaTeX mode: `\( 3.14\,\unit{...} \)`) or $0.872\,\mathrm{m\,s^{-1}}$ (using only `\unit{0.872\,m\,s^{-1}}`).

8 Conclusions

The conclusion goes here. You can modify the section name with `\conclusions[modified heading if necessary]`.

Code and data availability. use this to add a statement when having data sets and software code available

Sample availability. use this section when having geoscientific samples available

Appendix A: Figures and tables in appendices

Regarding figures and tables in appendices, the following two options are possible depending on your general handling of figures and tables in the manuscript environment:

A1 Option 1

If you sorted all figures and tables into the sections of the text, please also sort the appendix figures and appendix tables into the respective appendix sections. They will be correctly named automatically.

A2 Option 2

If you put all figures after the reference list, please insert appendix tables and figures after the normal tables and figures.

To rename them correctly to A1, A2, etc., please add the following commands in front of them: `\appendixfigures` needs to be added in front of appendix figures `\appendixtables` needs to be added in front of appendix tables

Please add `\clearpage` between each table and/or figure. Further guidelines on figures and tables can be found below.

Author contributions. M. Franklin conducted analyses and wrote the manuscript. M. Sorek-Hamer conducted analyses and reviewed the manuscript. O. Kalashnikova and D. Diner conceptualized the study. D. Diner edited the manuscript.

Competing interests. The authors declare no competing interests.

Disclaimer. We like Copernicus.

Acknowledgements. This work was supported by NASA Grant 80NSSC19K0225

References

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET - A federated instrument network and data archive for aerosol characterization, Remote Sensing of Environment, 66, 1–16, https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Shin, S. K., Tesche, M., Müller, D., and Noh, Y.: Technical note: Absorption aerosol optical depth components from AERONET observations of mixed dust plumes, Atmospheric Measurement Techniques, 12, 607–618, https://doi.org/10.5194/amt-12-607-2019, 2019a.

Shin, S. K., Tesche, M., Noh, Y., and Müller, D.: Aerosol-type classification based on AERONET version 3 inversion products, Atmospheric Measurement Techniques, 12, 3789–3803, https://doi.org/10.5194/amt-12-3789-2019, 2019b.