Reconstructing Aerosol Vertical Profiles with Aggregate Output Learning

Sofija Stefanović, Shahine Bouabid, Philip Stier, Athanasios Nenes, Dino Sejdinović
Motivation

GEOS-5 10km resolution
Red: Dust  Blue: Sea Salt  Green: Smoke  White: Sulfate

William Putman, NASA/Goddard
Motivation

IPCC 2013
Motivation

2D proxies (vertically aggregated data) often insufficient to understand aerosol distribution
e.g. aerosol optical depth from satellites \[ \text{AOD} = \int_0^H b_{\text{ext}}(h) \, dh \]
Problem Statement

General setup:

- Collection of bagged observations: \( \{ \{ x^{(i)}_j \} \}_{i=1}^n; y_j; z_j \}_{j=1}^n \)
- Function to disaggregate: \( f : \mathbb{R}^d_x \rightarrow \mathbb{R} \)
- Aggregation operator over column height: \( \text{Agg}_j : f \mapsto \int_{\text{column}} f(x) \, dh(x) \)
- Aggregate observation model:

\[
    z_j = \text{Agg}_j(f) + \varepsilon_j
\]

Specific problem to develop a proof of concept for the methodology:

Reconstruct vertical profiles of sulfate concentrations from aggregated column mass density + chemical and meteorological covariates

\[
    \sigma_{SO_4} = \int_0^H [SO_4](h) \, dh
\]
Dataset

**NASA’s GEOS-5 Nature Run** output used as dataset basis.

|            | Name                  | Notation | Units      |
|------------|-----------------------|----------|------------|
| 2D         | SO$_4$ column density  | $\sigma_{SO_4}$ | kg·m$^{-2}$ |
|            | Liquid water path     | LWP      | kg·m$^{-2}$ |
| 3D         | SO$_4$ mass mixing ratio | $r_{SO_4}$ | kg·kg$^{-1}$ |
|            | SO$_2$ mass mixing ratio | $r_{SO_2}$ | kg·kg$^{-1}$ |
|            | Relative Humidity     | RH       | 1          |
|            | Air temperature       | $T$      | K          |
|            | Vertical velocity     | $w$      | m·s$^{-1}$ |
|            | Cloud liquid water    | $q$      | kg·kg$^{-1}$ |
|            | Moist air density     | $\rho$   | kg·m$^{-3}$ |

*Table 1. Dataset variables, “2D” corresponds to variables indexed by time, latitude and longitude while “3D” corresponds to variables that also have a height dimension.*
Initial Solutions - Baseline 1

Input 3D covariates: \[ x = (\text{latitude, longitude, altitude, } r_{SO_2}, RH, T, w, q) \]

Objective: 
\[ \min_{f} \sum_{i=1}^{n} \left( \sigma_{SO_4} - \int_{1^{th \text{ column}}} f(x) dh(x) \right)^2 \]

Hypothesis: 
\[ f(x) = \beta^\top x \]

Solution: Closed form ridge regressor of column-aggregate inputs against AOD
Initial Solutions - Baseline 2

Input 3D covariates: $x = (\text{latitude, longitude, altitude, } r_{\text{SO}_2}, \text{RH}, T, w, q)$

Input 2D covariates: $y = (\text{latitude, longitude, } \sigma_{\text{SO}_4}, \text{LWP})$

Step 1: Fit $g : y_i \mapsto \int_{\text{ith column}} f(x) dh(x)$  
Step 2: $\min_f \sum_{i=1}^{n} (\sigma_{\text{SO}_4} i - g(y_i))^2$

Hypothesis: $f(x) = \beta^T x$  
$g(y) = \gamma^T y$

Solution: Closed form two-stage ridge regressor
Experiments

|          | RIDGE  | TWO-STAGE |
|----------|--------|-----------|
| RMSE (10^{-6}) | 3.47   | 3.52      |
| 2D MAE (10^{-6}) | 3.39   | 3.39      |
| Corr. (%) | 93.5   | 87.5      |
| RMSE (10^{-10}) | 2.71   | 2.50      |
| 3D MAE (10^{-10}) | 1.07   | 1.10      |
| Corr. (%) | 62.5   | 63.9      |

*Table 2.* Evaluation scores on vertical profile reconstruction; “2D” refers to evaluation against aggregate $\sigma_{SO_2}$ targets used for training; “3D” refers to evaluation against vertical groundtruth.
Discussion & future work

- Two-stage regression shows a slight increase in performance over simple kernel ridge regression with metrics used
- Unclear why the influence of SO2 profiles (important sulfate precursor gas) on predictions varies in experiments
- Metrics more suited to the problem should be developed
- Next step: use specialised aerosol models with lidar simulator and develop kernel-based model to tackle the AOD disaggregation problem:

Reconstruct vertical profiles $b_{\text{ext}}(h)$ from aggregated observations of the AOD and chemistry + meteorological covariates

Code and data available at:
https://github.com/shahineb/aerosols-vertical-profiles