Creating TROPOMI super observations for use in TM5-4dvar

Johann Rasmus Nüß
Fabian Piwowaczyk and Mihalis Vrekoussis

LAMOS group, IUP, University Bremen

Outline

Objective and Motivation

Objective

- Reduce data volume and computational cost by gridding, i.e. aggregating many observation in an area into a single superobservation
- Based on work by Miyazaki et. al. 2012 [1]
- Extended / modified to meet requirements of TM5-4dvar inverse model

[1] doi.org/10.5194/acp-12-2263-2012

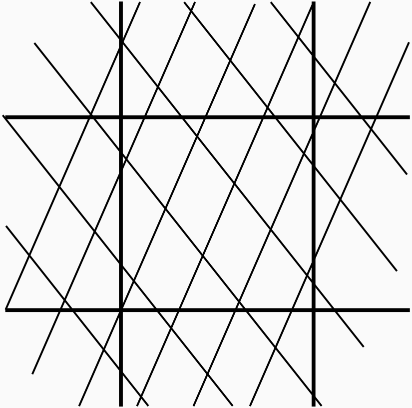
TROPOMI observations

- **TROPO**spheric **M**onitoring **I**nstrument onboard of **S**entinel-**5** **P**recursor
- Daily global coverage
- Local overpass time 13:30
- High resolution (up to $7{\times}7\,\text{km}^2$)
- Especially sensitive to troposphere/boundary layer

Image: ESA

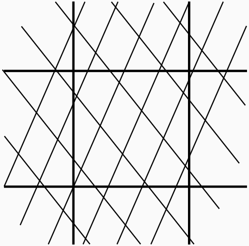
Gridding

The easy approach



- Set up global, regular grid
- Aggregate any observations in a grid cell into a single superobservation based on the location of their center
- How to weigh the observations?

Area-weighting



$$\hat{y}_o = \frac{\sum_{i=1}^{m} w_i \hat{y}_i^o}{\sum_{i=1}^{m} w_i}, \tag{1}$$

- Calculate intersection areas w_i of footprints $\hat{y}_i^o$ with each grid cell
- Get area-weighted mean
- Can also be applied to averaging kernel, pressure levels, time, and a-priori profile...
- ... but not to the observational error

Observational error and overlapping footprints

- Observations can contribute to multiple grid cells
- Inflate their error by $\sqrt{\frac{A_i}{w_i}}$ to keep their weight in the cost function constant, A_i is the total footprint area

$$\sigma = \frac{\sum_{i=1}^{m} \sqrt{\frac{A_i}{w_i}} w_i \sigma_i^o}{\sum_{i=1}^{m} w_i} = \frac{\sum_{i=1}^{m} \sqrt{A_i w_i} \sigma_i^o}{\sum_{i=1}^{m} w_i} \tag{2}$$

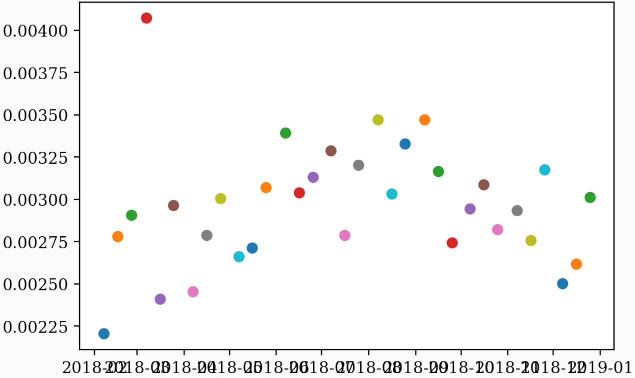
Error deflation and systematic errors

- Many independent observations reduce error by $\sqrt{n}$
- Adjacent satellite observations not independent
- Correlations in errors form assumptions about albedo etc.
- Eskes et. al. 2003 suggest $\sigma_o = \sigma\sqrt{\frac{1-c}{n} + c}$
- Miyazaki et. al. 2012 set $c = 15\%$

Representativeness error

- Handle grid cells on partly covered by observations
- introduce factor $f_{\mathrm{rep}}(\alpha)$ based on the relative coverage $0 \leq \alpha \leq 1$
- Estimate f_{rep} by artificially reducing coverage and comparing the resulting superobservations
- We aggregate f_{rep} into bins of 1 % coverage each
- Only use well covered cells, Miyazaki et. al. 2012 used $\alpha > 90\,\%$, we use $\alpha > 50\,\%$, to accommodate coarser grids

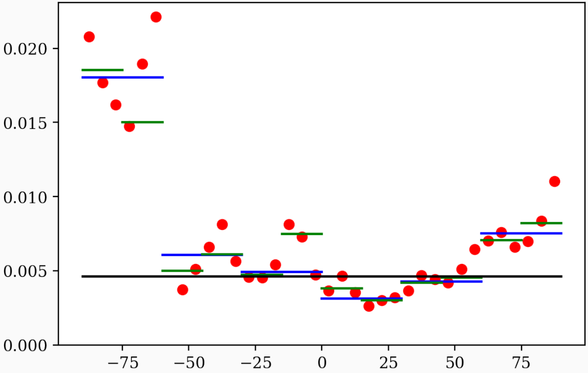
Representativeness error - intra-annual variations



- f_{rep} seems to weakly depend on season, likely due to differences in land mass between NH and SH
- daily variation has similar magnitude $\rightarrow$ use one consistent f_{rep} for the whole year

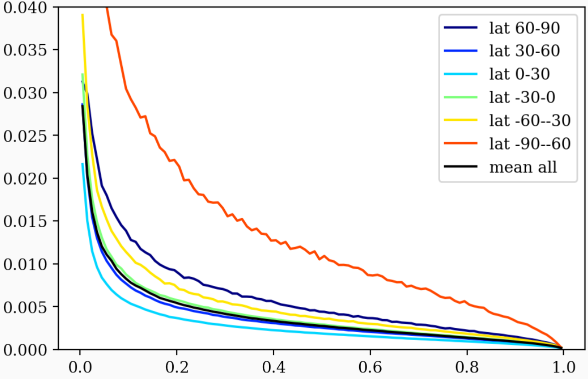
Representativeness error - latitudinal variations

- Magnitude of f_{rep} seems to strongly depend on latitude, likely linked to distribution of oceans and grid cell size



Representativeness error - latitudinal variations

- Magnitude of f_{rep} seems to strongly depend on latitude, likely linked to distribution of oceans and grid cell size
- Shape of f_{rep} appears to be unaffected by latitude



Representativeness error - latitudinal variations

- Magnitude of f_{rep} seems to strongly depend on latitude, likely linked to distribution of oceans and grid cell size
- Shape of f_{rep} appears to be unaffected by latitude

→ Calculate global $f_{\mathrm{rep}}(\alpha)$

→ Recalculated f_{rep} in 12° latitude bands to get scaling factor $\bar{f}_{rep}(\phi)$ to the global curve

$$f_{rep}(\alpha, \phi) = \bar{f}_{rep}(\phi) \cdot f_{\mathrm{rep}}(\alpha) \quad (3)$$

Total superobservation error

- Representativeness error for a specific coverage and latitude, based on the area-weighted observation:

$$\sigma_r = f_{\mathrm{rep}}(\alpha, \phi) \cdot \hat{y}_o. \tag{4}$$

- Total superobservation error by combing inflation through representativeness error and deflation through number of observations:

$$\sigma_s = \sqrt{\sigma_o^2 + \sigma_r^2}. \tag{5}$$

Summary & Outlook

Summary & Outlook

Two papers in the making:

- Technical paper, showcasing TROPOMI (super)observations in TM5-4dvar

Summary & Outlook

Two papers in the making:

- Technical paper, showcasing TROPOMI (super)observations in TM5-4dvar

$\rightarrow$ $0.5^\circ \times 0.5^\circ$ superobservation as input for inversion with NH in $3^\circ \times 2^\circ$

Summary & Outlook

Two papers in the making:

- Technical paper, showcasing TROPOMI (super)observations in TM5-4dvar

$\rightarrow$ $0.5^\circ \times 0.5^\circ$ superobservation as input for inversion with NH in $3^\circ \times 2^\circ$

$\rightarrow$ Investigate large scale emission increments over Eastern-Europe and decrements over China

Summary & Outlook

Two papers in the making:

- Technical paper, showcasing TROPOMI (super)observations in TM5-4dvar

$\rightarrow$ $0.5^\circ \times 0.5^\circ$ superobservation as input for inversion with NH in $3^\circ \times 2^\circ$

$\rightarrow$ Investigate large scale emission increments over Eastern-Europe and decrements over China

- Scientific paper for Californian fires after that

Acknowledgments

- The computations were performed on the HPC cluster Aether at the University of Bremen, financed by DFG in the scope of the Excellence Initiative.
- The PhD position is paid for by the University Bremen.
- Special thanks to the TM5 community, especially Maarten Krol and Sourish Basu for provision of and help with the TM5-4DVAR model.

- ... and of course thank You for your attention

Inverse Modeling - Mathematical description

- Forward model **F**
  - takes parameters $\vec{p}$ (meteorology, chemistry, ...)
  - and state $\vec{x}$ (emissions)
  - yields observation $\vec{y}$ (satellite measurements)

$$\vec{y} = \mathbf{F}(\vec{x}, \vec{p})$$

Inverse Modeling - Mathematical description

- Forward model $\mathbf{F}$
  - takes parameters $\vec{p}$ (meteorology, chemistry, ...)
  - and state $\vec{x}$ (emissions)
  - yields observation $\vec{y}$ (satellite measurements)

$$\vec{y} = \mathbf{F}(\vec{x}, \vec{p}) + \vec{\varepsilon}_O$$

with observational error $\vec{\varepsilon}_O$ (error of measurements, model, and parameters)

Inverse Modeling - Cost function

- Least squares approach
- Assume a priori state $\vec{x}_A$
- Error covariance matrices $\mathbf{S_O}$ and $\mathbf{S_A}$

Inverse Modeling - Cost function

- Least squares approach
- Assume a priori state $\vec{x}_A$
- Error covariance matrices $\mathbf{S_O}$ and $\mathbf{S_A}$

$$\text{Cost} = \qquad\qquad + $$

Inverse Modeling - Cost function

- Least squares approach
- Assume a priori state $\vec{x_A}$
- Error covariance matrices $\mathbf{S_O}$ and $\mathbf{S_A}$

$$\text{Cost} = \qquad + \frac{(\text{obs} - \text{model}(\text{state}))^2}{\text{error}_{\text{obs}}{}^2}$$

Inverse Modeling - Cost function

- Least squares approach
- Assume a priori state $\vec{x_A}$
- Error covariance matrices $\mathbf{S_O}$ and $\mathbf{S_A}$

$$\text{Cost} = \frac{(\text{state} - \text{a priori})^2}{{\text{error}_{\text{apri}}}^2} + \frac{(\text{obs} - \text{model}(\text{state}))^2}{{\text{error}_{\text{obs}}}^2}$$

Inverse Modeling - Cost function

- Least squares approach
- Assume a priori state $\vec{x_A}$
- Error covariance matrices $\mathbf{S_O}$ and $\mathbf{S_A}$

$$\text{Cost} = \frac{(\text{state} - \text{a priori})^2}{{\text{error}_{\text{apri}}}^2} \qquad + \frac{(\text{obs} - \text{model}(\text{state}))^2}{{\text{error}_{\text{obs}}}^2}$$

$$J(\vec{x}) = (\vec{x} - \vec{x_A})^T \mathbf{S_A^{-1}} (\vec{x} - \vec{x_A}) \quad + (\vec{y} - \mathbf{F}(\vec{x}))^T \mathbf{S_O^{-1}} (\vec{y} - \mathbf{F}(\vec{x}))$$

Inverse Modeling - Cost function

- Least squares approach
- Assume a priori state $\vec{x_A}$
- Error covariance matrices $\mathbf{S_O}$ and $\mathbf{S_A}$

$$\mathrm{Cost} = \frac{(\mathrm{state} - \mathrm{a\ priori})^2}{\mathrm{error_{apri}}^2} + \frac{(\mathrm{obs} - \mathrm{model}(\mathrm{state}))^2}{\mathrm{error_{obs}}^2}$$

$$J(\vec{x}) = (\vec{x} - \vec{x_A})^T \mathbf{S_A^{-1}} (\vec{x} - \vec{x_A}) + (\vec{y} - \mathbf{F}(\vec{x}))^T \mathbf{S_O^{-1}} (\vec{y} - \mathbf{F}(\vec{x}))$$

- Repeat cycle of forward $\rightarrow$ correcting $\rightarrow$ inverse until mismatch is sufficiently small