Synthetic spectra generation for TEMPO

Chris Chan Miller, Xiong Liu, Gonzalo Gonzalez Abad, Peter Zoogman, Rob Spurr

cmiller@cfa.harvard.edu

5th TEMPO Science Team Meeting
Cambridge, MA, May 31, 2016
Why do we need synthetic spectra?

(1) To test the Science Data Processing Center Pipeline

(2) To validate TEMPO retrieval algorithms

Why do we need synthetic spectra?

(1) To test the Science Data Processing Center Pipeline

(2) To validate TEMPO retrieval algorithms

Truth Atmosphere

Radiative Transfer Model

Synthetic TEMPO Spectra

Operational Algorithms

Validation

Retrieved Atmosphere

SDPC

2.6 million spectra / hour

TEMPO
Synthetic spectra generation for TEMPO

GEOS-5 Truth Atmosphere

Surface Ozone - 20130715 - 13:00 EST

“G5NR” - GEOS-Chem coupled to the GEOS-5 GCM run globally at ~12x12 km² resolution

Courtesy of Christoph Keller
Construction of truth atmosphere

Python Sampling Tool

G5NR

Truth Atmosphere

Synthetic spectra generation for TEMPO
Construction of truth atmosphere

G5NR

Sampled Fields:
- Trace Gases
- Aerosols
- Pressure
- Temperature
- Relative Humidity
- Winds (u,v)
- Cloud fraction & Optical depth
- Snow fraction & depth

Example: G5NR HCHO

10^{15} \text{ molecules cm}^{-2}
Sample G5NR fields based on simulated geolocation
Sample G5NR fields based on simulated geolocation
Construction of truth atmosphere

Example: G5NR HCHO

Sampled HCHO Field

Sample G5NR fields based on simulated geolocation
Construction of truth atmosphere

- G5NR
- Spacecraft Orbit (Carr Astronautics)

“Proxydynamics” simulation of orbit + spacecraft pitch/yaw/roll

Python Sampling Tool

Example: G5NR HCHO

Truth Atmosphere

Sampled HCHO Field

10^15 molecules cm^-2

Latitude (degrees)
Pressure (hPa)

Spacecraft Orbit (Carr Astronautics)
Profiles adjusted for subgrid altitude variation using GMTED2010 7.5 arc second digital elevation database
Tropospheric NO2 Column - 15/7/13 1pm EST

G5NR appears adequate for resolving range of trace gas concentrations expected for TEMPO
Tropospheric NO2 Column - 15/7/13 1pm EST

G5NR can resolve filamented plumes from urban emissions
Comparison with previous truth atmosphere

... but the G5NR spatial resolution is still lower than the TEMPO pixels
Hour-to-hour pointing variation

Tropospheric NO2 Column - 15/7/13

13:00 ET  14:00 ET  15:00 ET

Expected hour-to-hour jitter >> pixel dimension
Implications for oversampling?
Simulation of TEMPO Spectra

G5NR

Scan Geolocation

GMTED DEM Database

Python Sampling Tool

Truth Atmosphere

Synthetic spectra generation for TEMPO
Simulation of TEMPO Spectra

**GEOCAPE-Tool - Fortran interface to the VLIDORT RT Model**

- **Trace gas absorption cross sections**
- **Aerosol/cloud scattering (Mie theory)**
- **Surface reflectance (MODIS-BRDF)**

**Python Sampling Tool**

- **G5NR**
- **Scan Geolocation**
- **GMTED DEM Database**
Simulation of TEMPO Spectra

- G5NR
- Scan Geolocation
- GMTED DEM Database

Python Sampling Tool

- Truth Atmosphere

GEOCAPE Tool

- Trace gas absorption cross sections
- Aerosol/cloud scattering (Mie theory)
- Surface reflectance (MODIS-BRDF)

VLIDORT v2.7

I,Q,U

5/31/17

Synthetic spectra generation for TEMPO
Simulation of TEMPO Spectra

- G5NR
- Scan Geolocation
- GMTED DEM Database
- Python Sampling Tool
- Truth Atmosphere
- GEOCAPE Tool
- Trace gas absorption cross sections
- Aerosol/cloud scattering (Mie theory)
- Surface reflectance (MODIS-BRDF)
- TEMPO Instrument Model
- Synthetic L0
- SDPC
- VLIDORT v2.7

Synthetic spectra generation for TEMPO
Typical optical depths for simulated species

- $\text{O}_3 (T)$
- $\text{SO}_2$
- $\text{O}_2-\text{O}_2 (T)$
- $\text{NO}_2 (T)$
- $\text{BrO}$
- $\text{H}_2\text{CO}$
- $\text{CHOCHO}$
- $\text{H}_2\text{O}(p,T)$
- $\text{O}_2 (p,T)$
- Aerosol cross sections/phase functions calculated from Mie theory
- Aerosol refractive indices and size distribution from GADs
- The RH dependence of sea salt, sulfate and organic carbon optical properties is accounted for
Clouds modeled using the independent pixel approx.

Cloud scattering is explicitly simulated (i.e. no lambertian surface approximation)

\[ I(\lambda) = \sum_{j=1}^{n} f_j I_j(\lambda) \]
How do we infer cloud subpixels from an aggregate cloud profile?
The simulated radiance will be sensitive to cloud overlap assumptions.

Clouds are archived on a 0.5x0.625 grid. This is an order of magnitude larger than a TEMPO pixel.

Prather (2015)
(1) Divide clouds into low, medium, and high altitude type layers
(2) Assume clouds randomly overlap between each coarse layer
(3) Assume clouds are maximally overlapped within each layer
Synthetic spectra generation for TEMPO
Stochastic column generation

Cloud Optical Depth CDF with MRAN Assumption

Cloud Column Optical Depth

Sample CDF onto 0.1x0.1 degree sub grid
Stochastic column generation

Cloud Optical Depth CDF with MRAN Assumption

Impose spatial correlation using Gaussian Copula (Wind et al. 2013)

Not accounting for spatial correlation between pixels will underestimate the number of cloud free pixels
**Stochastic column generation**

Cloud Optical Depth CDF with MRAN Assumption

- **Cloud Column Optical Depth**
- **Pressure (hPa)**
- **Cloud Column O. Depth**

**Impose spatial correlation using Gaussian Copula (Wind et al. 2013)**

Sample clouds to TEMPO grid
Surface reflection

- Snow Free Land
  - MODIS BRDF with EOF-Interpolation

- Snow Covered Land/Sea Ice
  - GISS GCM Snow parameterization

- Ocean
  - Cox-Munk ocean BRDF

Synthetic spectra generation for TEMPO
Diurnal variation in surface reflectivity

Forested Scene, NY State
(Viewing Zenith Angle = 54°)

Large variations over the course of a day

Zoogman et al. (2016)
MODIS uses a 3 Kernel BRDF approximation:

\[
f_r(\lambda, \vec{\omega}_i, \vec{\omega}_f) = f_{iso}(\lambda) + f_{vol}(\lambda)K_{vol}(\vec{\omega}_i, \vec{\omega}_f) + f_{geo}(\lambda)K_{geo}(\vec{\omega}_i, \vec{\omega}_f)
\]
Consider reflectance for one geometry \((w_i, w_f)\):

\[
f_r(\lambda, \vec{w}_i, \vec{w}_f) = f_{iso}(\lambda) + f_{vol}(\lambda)K_{vol}(\vec{w}_i, \vec{w}_f) + f_{geo}(\lambda)K_{geo}(\vec{w}_i, \vec{w}_f)
\]
EOFs derived from USGS/ASTER databases

Factor loadings derived from snow free SCIAMACHY LER observations

Both loading vectors can explain >99% of variability in their respective datasets
Consider reflectance for one geometry \((w_i, w_f)\):

![Graph showing BRDF vs Wavelength (nm).]

- BRDF values range from 0.00 to 0.25.
- Wavelength (nm) ranges from 300 to 900.
Consider reflectance for one geometry \((w_i, w_f)\):

Fit database EOFs (Zoogman et al.) over 400-900 nm
Consider reflectance for one geometry \((w_i, w_f)\):

Perform least squares fit of SCIAMACHY factor loadings between 400 - 900 nm to predict reflectances < 400 nm
Consider reflectance for one geometry \((w_i, w_f)\): 

Interpolate values below 400nm using a smoothed spline to reduce potential discontinuity at 400nm.
Interpolate values below 400nm using a smoothed spline to reduce potential discontinuity at 400nm
The synthetic view from TEMPO

Courtesy of J. Carr
Future improvements
- RRS approximation using VLIDORT jacobians
- Chlorophyll fluorescence
- Liquid Water + Chlorophyll Absorption in ocean
- More trace gases - HONO, OCIO, IO, Methylglyoxal, ....

Requests?
EXTRA SLIDES
(1) Sample 3 numbers from a uniform distribution corresponding to the cloud layers

\[ r_l, r_m, r_h \sim U(0, 1) \]
MRAN Stochastic column generator

(1) Sample 3 numbers from a uniform distribution corresponding to the cloud layers

$$r_l, r_m, r_h \sim \mathcal{U}(0, 1)$$
(2) Layer is cloudy if cloud fraction > r
(3) Sample cloud optical depths for layers classified as cloudy
(1) Generate cloud subcolumns using stochastic column algorithm with MRAN assumption

(2) Distribute on 0.1x0.1 sub grid
Stochastic column generation

(3) Sample columns for each TEMPO pixel

NB - This implicitly assumes no spatial correlation between pixels on the downscaled grid - and likely underestimates the fraction of clear TEMPO pixels
Stochastic column generation

(1) Sort columns by optical depth to create marginal CDF

Wind et al. (2013)

(2) Use gaussian copula to specify relationship between cloud fields

\[ r_i = \frac{1}{2} \left( 1 + \text{erf} \left( \frac{x_i}{\sqrt{2}} \right) \right) \quad r_i \in [0, 1] \]

(3) Compute inverse transform for each \( x_i \) and sample marginal CDF

\[ \bar{x} \sim \mathcal{N}(\bar{0}, S) \]

\[ S_{i,j} = \exp \left( -\frac{d_{i,j}}{d_c} \right) \]

Distance between \( x_i \) and \( x_j \)

Correlation length scale

Downscaled grid

Wind et al. (2013)
Stochastic column generation

(1) Sort columns by optical depth to create marginal CDF

\[ x_i \sim \mathcal{N}(\bar{0}, S) \]

\[ S_{i,j} = \exp \left( -\frac{d_{i,j}}{d_c} \right) \]

(2) Use gaussian copula to specify relationship between cloud fields

\[ r_i = \frac{1}{2} \left( 1 + \text{erf} \left( \frac{x_i}{\sqrt{2}} \right) \right) \quad r_i \in [0, 1] \]

(3) Compute inverse transform for each \( x_i \) and sample marginal CDF
(1) Sort columns by optical depth to create marginal CDF

Wind et al. (2013)
Effect of correlation length scale

Synthetic spectra generation for TEMPO

$\mathbf{d=0}$

$\mathbf{d=\infty}$
Effect of correlation length scale

\[ d = 0 \]

\[ d = \infty \]

Correlation length scale (Degrees)

Clear Pixel Fraction