Fast Forward Modeling in Scattering Atmospheres with Optimum Spectral Sampling

Jean-Luc Moncet, Gennady Uymin, Alan Lipton, Ned Snell

Atmospheric and Environmental Research, Inc.
Topics

- Optimal Spectral Sampling brief overview
- Global training
  - Application to AIRS and IASI
- Use of principal components of radiances
- Treatment of multiple scattering
  - Current training method
  - Minimizing the number of scattering RT operations
  - Application to AIRS and MODIS
Review of the Basic OSS Method

- OSS channel radiances modeled as (Moncet et al. 2003, 2001, 2008)

\[
\overline{R} = \int_{\Delta \nu} \phi(\nu) R(\nu) d\nu \approx \sum_{i=1}^{N} w_i R_i(\nu_i); \quad \nu_i \in \Delta \nu
\]

- Channel-average radiance is modeled as weighted average of monochromatic radiances
- Wavenumbers \( \nu_i \) (nodes) and weights \( w_i \) are optimally selected to fit calculations from a reference line-by-line model for a globally representative set of profiles (training set)
- Monochromatic absorption coefficients from look-up tables
- Demonstrated to be generally faster and more accurate than methods that use regression to fit the space-to-level band transmittance (here called total path transmittance regression = TPTR methods)
  - For molecular absorption, non-scattering computations
  - Scattering case is discussed in this presentation
Training Approaches: 1) Local Training

- Operates on individual channels, one at a time
- Nodes for each channel required to be within spectral range of channel response
  - Nodes may be shared between channels with overlapping responses

AIRS (2378 channels):
- Average: 11 nodes weighted per channel
- Average: 1.3 nodes/channel overall (accounts for sharing)
Training Approaches: 2) Global Training

- Operates on groups of channels (up to the full channel set) simultaneously
- Uses clustering of nodes to efficiently account for spectral correlations
  - Condenses the information of the full channel set into a minimal number of nodes
- Monochromatic RT at a relatively few nodes determines radiances for full channel set
- Optionally, can be fit to channel subset, or first X principal components of channel set, or radiances filtered by PC transformation
  - Reduces information relative to full channel set
AIRS Example

<table>
<thead>
<tr>
<th></th>
<th>AIRS - full channel set</th>
<th>AIRS - 281 channel subset</th>
</tr>
</thead>
<tbody>
<tr>
<td># channels</td>
<td>2378</td>
<td>281</td>
</tr>
<tr>
<td># nodes</td>
<td>5340</td>
<td>1809</td>
</tr>
<tr>
<td># nodes / # channels</td>
<td>2.25</td>
<td>6.44</td>
</tr>
<tr>
<td>N'</td>
<td>9.84</td>
<td>11.75</td>
</tr>
</tbody>
</table>

N' = number of nodes contributing to radiance computation in 1 channel (on average)

Training conditions:
- 0.05 K accuracy requirement
- Extra wide range of incidence angles (0°–70.5°)*
- Variable gases H₂O, O₃

Global20

- Surface emissivity assumed linear over contiguous 20 cm⁻¹ intervals (needs to be verified with real data)
- Global training applied independently to each interval
- Provides extra robustness by avoiding reliance on correlations from distant spectral points
  - Extra robustness for surfaces whose emissivity spectra are outliers

*difference from data in chart 4
Application to IASI

**Global and local training results**

<table>
<thead>
<tr>
<th>IASI band</th>
<th>Spectral range (cm⁻¹)</th>
<th>Number of channels</th>
<th>Local</th>
<th>Global</th>
<th>Global nodes/ channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>645-1210</td>
<td>2261</td>
<td>1855</td>
<td>220</td>
<td>0.097</td>
</tr>
<tr>
<td>2</td>
<td>1210-2000</td>
<td>3160</td>
<td>2927</td>
<td>281</td>
<td>0.089</td>
</tr>
<tr>
<td>3</td>
<td>2000-2760</td>
<td>3040</td>
<td>2639</td>
<td>321</td>
<td>0.11</td>
</tr>
<tr>
<td>Total</td>
<td>2000-2760</td>
<td>8461</td>
<td>7421</td>
<td>822</td>
<td></td>
</tr>
</tbody>
</table>

Training conditions:
- 0.05 K accuracy requirement
- 13 variable gases: H₂O, O₃, CO₂, CO, CH₄, N₂O, F11, F12, CCl₄, HNO₃, SO₂, OCS, CF₄
- 5 fixed gases: O₂, NO, NO₂, NH₃, N₂
- Sources: ECMWF for H₂O, O₃; Global Modeling Initiative chem model for CO₂, CO, CH₄, N₂O, F11, M. Matricardi for F12, CCl₄, HNO₃
- 2002-2012 secular trends added for CO₂ and CH₄
- Randomization was applied to all species for robust training
- Emissivity spectra for global training is random walk, with 20-cm⁻¹ steps
IASI Validation

- Validation with 48 independent UMBC profiles
- Each profile assigned 3 CO$_2$ profiles:
  - Minimum (●), mean (●), maximum (●)
- Validates robustness of training over CO$_2$ trends

From local training
OSS with Principal Components

- Option may be useful when some information loss is accepted as trade-off for speed
  - When eigenvector truncation goes beyond eliminating redundancy
- Can be done without significant revision to OSS training
  - Filter training-profile radiances with truncated eigenvectors
    - Convert to PCs, then use reverse transformation to recover physical-space radiances
- OSS training achieves required accuracy for every channel (PC filtered)
- OSS coefficients project only on retained PCs (within training accuracy)

Forward model output in terms of PCs efficiently done by combining eigenvectors with OSS coefficients in advance

\[ R_{\text{chan}} = W R_{\text{node}} \quad \text{PC} = U_m R_{\text{chan}} \quad \text{with } m \text{ retained PCs} \]

\[ \text{PC} = U_m W R_{\text{node}} = W_m R_{\text{node}} \quad \text{where} \quad U_m W = W_m \]
OSSSCAT is single-wavelength version of CHARTS adding-doubling RTM

- Uses same molecular absorption and weighted monochromatic radiances as non-scattering RTM
- Cloud module converts from physical properties (e.g., IWP, LWP, $D_{\text{eff}}$, top, thickness, $T(p)$) to optical properties (absorption and scattering optical depths, asymmetry parameter)
  - Look-up table
    - Size distributions based on in-situ aircraft measurements
    - Mie for liquid
    - MADA for ice - with temperature-dependent shape recipes
- Optical properties linearly interpolated from hinge points to OSS nodes
**Cloudy Training**

- **Must include cloud/aerosol optical properties in training**
  - Over wide bands: training can be done by using a database of cloud/aerosol optical properties
  - More general training obtained by breaking spectrum in intervals of the order of 10 cm\(^{-1}\) in width (impact of variations in cloud/aerosol properties on radiances is quasi-linear) and by performing independent training for each interval
    - Lower computational gain but increased robustness

- **Direct cloudy radiance training approach**
  - Clouds tend to mask molecular structure, which makes training less demanding
  - If trained for mixture of clear and cloudy atmospheres in direct training, clear-sky performance degrades
  - Train with clear-sky and several clouds simultaneously, requiring all to meet the accuracy criterion
**Cloudy and Clear Fit**

- **OSS selection requires accuracy threshold be met for each training set individually and simultaneously**

  *AIRS Channel Set*
  *Scattering included*
  *Fit error -- all meet 0.05-K rms requirement*
  *Nadir view shown*
**Multiple Scattering Acceleration**

- With scattering, execution time is dominated by radiative transfer integration
  - Contrasts with non-scattering, where band transmittance calculation may be a bigger factor
  - OSS RT timing ~proportional to number of nodes
  - TPTR RT timing ~proportional to number of channels
  - OSS is faster than TPTR methods only when the number of nodes / number of channels <~1

- Scattering calculations do not have to be performed for each node
  - Scattering correction may be predicted based on a few nodes only:
    \[
    \bar{R} \approx \sum_{i=1}^{N} w_i R_{ns}^i(v_i) + \sum_{k \in S} C_k [R(v_k) - R_{ns}^k(v_k)]
    \]
    - \( R \) is radiance from scattering model
    - \( R_{ns} \) is radiance from non-scattering model
    - \( w \) are the ordinary OSS weights
    - \( k \) are a subset of the set of the OSS nodes \((S)\) for the channel
    - \( C \) are regression coefficients

- Number of predictors can be tuned to control balance between cloudy radiance accuracy and computation speed
  - Some relaxation of accuracy may be tolerable in clouds with high optical depth, in proportion to uncertainties in optical properties

In thermal regime
### Scattering Prediction Performance for MODIS

<table>
<thead>
<tr>
<th>MODIS Channel #</th>
<th>Bandpass (µm)</th>
<th>Number of nodes*</th>
<th>Number of predictor nodes†</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>3.660 - 3.840</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>21</td>
<td>3.929 - 3.989</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>3.929 - 3.989</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>4.433 - 4.498</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>4.482 - 4.549</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>6.535 - 6.895</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>7.175 - 7.475</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>29</td>
<td>8.400 - 8.700</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>31</td>
<td>10.780 - 11.280</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>11.770 - 12.270</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>13.185 - 13.485</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>34</td>
<td>13.485 - 13.785</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>13.785 - 14.085</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>36</td>
<td>14.085 - 14.385</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td><strong>13.6</strong></td>
<td><strong>1.7</strong></td>
</tr>
</tbody>
</table>

* for error threshold 0.05 K, clear and cloudy training
† for scattering prediction error threshold 0.2 K

Selected IR channels

Localized training used

Generalized may require fewer predictors
Simulated TPTR method

- Ideal case of no error in transmittance regression
- Effectively approximates reflected component as product of band-averages

\[ \bar{R}_{\text{refl}} \approx \bar{r} \bar{T} \bar{R} \]

effects in-band correlations

biggest impact is with low emissivity = high reflectivity \( r \)
OSS Performance for MODIS Accelerated with Scattering Selection

Scattering: 2 predictor nodes 3 predictor nodes all nodes

Channel 27

Channel 28

Channel 29 Desert (high reflectivity surface)

Ice cloud (nadir) Ice cloud (60°)
Liquid cloud (nadir) Liquid cloud (60°)

Generally 1-2 predictors are sufficient to exceed TPTR accuracy
Summary

- Global training with correlated clustering minimizes number of nodes for channel set as a whole
- Flexibility - same monochromatic (physical, general) framework provides options to meet user requirements
  - Produce radiances for full channel set
  - Retrieve/assimilate at OSS nodes
    - Avoids computational cost of mapping from nodes to channels
    - Involves channel $\rightarrow$ node transformation of measurement error covariance
      - Treatment of scene-dependent noise depends on application
- Scattering version maintains accuracy in clear areas
- Scattering can be accelerated with process to select subset of nodes to do scattering
  - Requires testing with global training
Backup
Variational retrieval methods:
- Average channel uses ~150 nodes
- Mapping Jacobians from node to channel space partially offsets speed gain

Alternatives:
(a) PC (reduces first dimension of matrix A)
(b) Operate directly in node space

\[ y^m = A \hat{y}_0 \rightarrow \hat{y}_0 = H y^m \]

Avoids Jacobian transformation altogether and reduce K-matrix size (inversion speed up)

for AIRS: 2378 channels → 250 nodes

\[ \delta x_{n+1} = \left( K_n^T S^{-1}_n K_n + S^{-1}_x \right)^{-1} K_n^T S^{-1}_n \left[ (y_n - y^m) + K_n \delta x_n \right] \]

where,
\[ y = A \hat{y}_0 \text{ and} \]
\[ K = A K_0^\theta \]

\[ \tilde{y}^m = \left( A^T S^{-1}_x A \right)^{-1} A^T S^{-1}_x y^m \]
\[ \tilde{S}_x^{-1} = A^T S^{-1}_x A \]
\[ \delta x_{n+1} = \left( \tilde{K}_n^T \tilde{S}_x^{-1} \tilde{K}_n + S^{-1}_x \right)^{-1} \tilde{K}_n^T \tilde{S}_x^{-1} \left[ (\tilde{y}_n - \tilde{y}^m) + \tilde{K}_n \delta x_n \right] \]

**Equivalent to**
\[ \delta x_{n+1} = \left( K_n^T S^{-1}_x K_n + S^{-1}_x \right)^{-1} K_n^T S^{-1}_x \left[ (y_n - HA y^m) + K_n \delta x_n \right] \]
Inversion (continued)

Need strategy for handling input-dependent noise

- Scene temperature dependence (clear/cloudy)
- worse in SW band
- Cloud clearing noise amplification
- H-transformation not overly sensitive to noise

For clear retrievals: sufficient to update noise covariance regionally

Retrieval performance - constant noise

Channel space retrieval
Node space retrieval