From model simulations towards vegetation properties mapping:

automating, optimizing & simplifying

J. Verrelst, J.P. Rivera & J. Moreno

ISSI Workshop – 21-25 Nov2016
AUTOMATING RTMs

- RTMs
- ARTMO/forward
- Retrieval toolboxes

OPTIMIZING exploiting spectroscopy data

- Retrieval
- Band selection
- Dimensionality & sample reduction

SIMPLIFYING RTMs

- Global sensitivity analysis
- Emulation
- Retrieval
AUTOMATING

- RTMs
- ARTMO/forward
- Retrieval toolboxes
A diversity of RTMs exist with different complexity.
RTMs are important tools in EO research, but not always easy to use. Only very few of them offer user-friendly interfaces.

- No interface exists that brings multiple RTMs together in one GUI.
- None of existing (publicly available) GUIs provide post-processing tools.
Toolbox for EO applications:

**ARTMO**

**Automated**

**Radiative**

**Transfer**

**Models**

**Operator**

http://ipl.uv.es/artmo/
ARTMO

Atmospheric models
MODTRAN
Vegetation indices

Time series analysis
Ray tracing model
RPV model

Neural nets
Support vectors
Gaussian Processes

BRDF apps
Spectral unmixing
Classifiers
ARTMO v. 3: modular design

[Diagram showing various options and modules for ARTMO v. 3, including File, Models, Forward, Retrieval, Tools, and Help categories with sub-options like Load Project, New Project, DB administration, Settings, Model inputs, Leaf, Canopy, Combined, Spectral Indices, MLRA, LUT-based Inversion, Sensor, Graphics, Spectral resample, GSA, Emulator, GSA configuration, GSA results, and others.]
Conceptual architecture ARTMO
Forward
RTM outputs only a few clicks away...

Data flow
ARTMO’s leaf models

PROSPECT-4

PROSPECT-5

DLM

LIBERTY

Fluspect-B

New!
ARTMO’s canopy models

SAIL

INFORM

FLIGHT
ARTMO’s combined models: SCOPE
Sensor

Simulations can be generated according to band settings of a selected sensor.

- New sensor settings can be imported by clicking on the ‘Import’ button in the top bar.
- Existing band settings can be modified or new ones can be added by clicking on the ‘Edit’ button.
- Also a spectral filter of a sensor can be imported or viewed by clicking on the ‘Spectral Filter’ button.

Default sensors:
- Landsat 7 TM
- Landsat 7 ETM+
- SPOT-4 VMI
- SPOT-4 HRVIR
- CHRIS Mode-3
- MODIS
- MERIS
- Sentinel-2
- Sentinel-3 OLCI
- Sentinel-3 SLSTR
- Landsat 8
- Pleiades-1A
- Quickbird
Graphics

Visualization options:

Export:
- Spectral data
- Associated metadata
Retrieval
### Retrieval families

<table>
<thead>
<tr>
<th>Parametric regression</th>
<th>Non-parametric regression</th>
<th>RTM inversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral relationships that are sensitive to specific vegetation properties</td>
<td>Advanced techniques that search for relationships between spectral data and biophysical variables</td>
<td>Models that simulate interactions between vegetation and radiation</td>
</tr>
</tbody>
</table>

**Normalized Difference Vegetation Index**

\[
NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}
\]

Methods of these different families can be combined: *hybrid methods*
ARTMO’s retrieval toolboxes:

Spectral indices toolbox

[Diagram of spectral indices toolbox]

Machine learning regression algorithm toolbox

[Diagram of MLRA toolbox]

LUT-based inversion toolbox

[Diagram of LUT-based inversion toolbox]

Optimizing and generating maps of vegetation properties only a few clicks away...
General structure:

1. **Input**
2. **Settings**
3. **Validation**
4. **Retrieval**

- Database [spectral+variables]
- Spectral data
- Biophysical variables
- T/V
- % Training data/ % validation data

- **Training**
  - Spectral data
  - Biophysical variables

- **Noise**
- Retrieval algorithm
- Developed model
- Estimated biophysical variable
- Goodness-of-fit statistics

- **Validation**
  - Spectral data
  - Biophysical variables

- **Optimized model**
- Spectral data
- Biophysical variable map

To account for natural variability
OPTIMIZING

- Retrieval: parametric/non-parametric/inversion
- Band selection
- Dimensionality & sample reduction
Spectral indices toolbox:

Properties:
- Calculates all possible band combinations.
- For index formulations with up to 10-band indices ($#b^{10}$, for a 10-band sensor that would be 10 billion combinations)
- Includes multiple fitting functions (linear, exponential, logarithmic, power, polynomial)
- Noise & Cross-validation options
- Results stored in MySQL
- Top-performing indices per formulation and fitting function are given.
- Can process both image or individual spectra.

Best-performing index can be applied to an image.
SPARC dataset (Barrax Spain); HyMap data

SR and ND for different fitting functions

Best performing model.

- Which VI method is most correct?
- Why restricting to a few bands only?
Machine learning regression algorithm toolbox

Properties:
- About 15 MLRAs implemented
- Single-output & multi-output
- Noise & Cross-validation options
- Dimensionality reduction options
- Results stored in MySQL
- GPR properties: band relevance & uncertainties
- Can process both images or individual spectra.
- Active learning, GPR-BAT, dim. reduction

Simpler to execute than SI: no band selection needed.

Non-parametric models:
- SimpleR [Camps-Valls et al., 2013]
  - http://www.uv.es/gcamps/code/simpleR.html

Also:
- Bagging trees (BAGTREE)
- Boosting trees (BOOST)
- Neural networks (NN)
- Extreme Learning Machines (ELM)
- Support Vector Regression (SVR)
- Relevance Vector Machine (RVM)
- Variational Heteroscedastic Gaussian Process Regression (VHGPR)

GPR in Bayesian framework also provides:
- Band relevance
- Uncertainty estimates
CAS1500
- pixel size: 1.4m
- 288 bands

Same GP model was applied
• **Uncertainty maps** provide additional info which may be hidden on the images.
• Implausible estimations are detected.
• In turn, **despite low uncertainties also good estimations.** No impact on recently irrigated areas (other methods have difficulties with wet soils).

• For operational applications, of interest to flag/mask regions with high uncertainties.
GPR to S2

GPR Retrievals with uncertainties <40% masked out (removes directly non-vegetated surfaces).
The challenge of machine learning applied to imaging spectroscopy

- **Machine learning methods are adaptive and can be very powerful. However that goes a computational cost. This can be problematic when large datasets are involved, either in the sampling or in the spectral domain (e.g. for hybrid methods).**

- **Moreover, when many bands are involved: multicollinearity leads to statistical problems (suboptimal performance).**

**Solutions to deal with large datasets:**

1. Reducing spectral data:
   - I. band selection (GPR-BAT),
   - II. dimensionality reduction

2. Samples reducing: Active learning
I) Band selection: GPR-BAT

Gaussian processes regression – Band analysis Tool (GPR-BAT).

Sequential Backward Band Removal: remove band with highest sigma (least informative)

The lower the sigma, the more important the band is! 😊

Best-performing method can be applied to an image.
GPR-BAT example with simulated data (PROSAIL)

Experimental setup:
- PROSAIL: LHS 100@ 10 nm, 220b; Cab, LAI
- 4k cross-var sampling

<table>
<thead>
<tr>
<th># band</th>
<th>R2</th>
<th>wavelengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.9997</td>
<td>815, 1145, 1205, 122, 1245</td>
</tr>
<tr>
<td>4</td>
<td>0.9997</td>
<td>815, 1145, 1205, 1245</td>
</tr>
<tr>
<td>3</td>
<td>0.9213</td>
<td>815, 1145, 1205</td>
</tr>
<tr>
<td>2</td>
<td>0.8104</td>
<td>815, 1145</td>
</tr>
<tr>
<td>1</td>
<td>0.8104</td>
<td>815</td>
</tr>
</tbody>
</table>

Best performances achieved between 70 and 4 bands
(using all bands or <3 bands not recommended)
GPR-BAT example with A. Gitelson field data (maize/soybean, OO, 300#b)

LCC (best with 9 bands)

- All bands:
  - 482, 500, 564, 566, 710, 712, 714, 878, 966, 980
  - 406, 746, 770, 790, 792, 794, 798, 808, 858, 878

gLAI (best with 7 bands)

- All bands:
  - 482, 500, 564, 710, 712, 714, 878, 966
  - 406, 746, 790, 792, 794, 798, 808, 858, 878

- 9 bands:
  - R^2_{CV}: 0.79
  - RMSE_{CV}: 7.26
  - NRMSE_{CV}: 12.90%

- 7 bands:
  - R^2_{CV}: 0.94
  - RMSE_{CV}: 0.40
  - NRMSE_{CV}: 7.19%
GPR-BAT example with field data (SPARC, Barrax, Spain; Hymap, 125#b)

**LAI (best with 4 bands)**

**CWC (best with 6 bands)**

---

**All bands**

462, 478, 708, 723, 1215, 1243, 1272, 1327, 1635, 2483
462, 478, 708, 723, 1215, 1243, 1272, 1327, 2483
462, 478, 708, 723, 1215, 1243, 1272, 1327
462, 478, 708, 723, 1215, 1327
462, 478, 708, 723, 1327
462, 708, 723, 1327
462, 708, 1327
462, 1327
462

**4 bands**

- $R^2_{CV} = 0.95$
- $\text{RMSE}_{CV} = 0.37$
- $\text{NRMSE}_{CV} = 6.50\%$

---

**All bands**

462, 723, 1128, 1157, 1272, 1286, 1299, 1327, 1419, 2483
723, 1128, 1157, 1272, 1286, 1299, 1327, 1419, 2483
723, 1128, 1157, 1272, 1286, 1327, 1419, 2483
723, 1128, 1157, 1272, 1286, 1327, 1419
723, 1157, 1272, 1286, 1157, 1272, 1286
1157, 1272, 1286
1157, 1286
1286

**6 bands**

- $R^2_{CV} = 0.99$
- $\text{RMSE}_{CV} = 14.51$
- $\text{NRMSE}_{CV} = 7.24\%$
**BIOHYPE & SCOPE: SIF (200#b)**

**SLA (best with 3 bands)**

<table>
<thead>
<tr>
<th>#</th>
<th>R2</th>
<th>SD</th>
<th>MIN</th>
<th>MAX</th>
<th>Wavelengths (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.704</td>
<td>0.057</td>
<td>0.652</td>
<td>0.796</td>
<td>675,676,694,761,762</td>
</tr>
<tr>
<td>4</td>
<td>0.704</td>
<td>0.058</td>
<td>0.652</td>
<td>0.797</td>
<td>675,694,761,762</td>
</tr>
<tr>
<td>3</td>
<td>0.704</td>
<td>0.057</td>
<td>0.652</td>
<td>0.797</td>
<td>675,694,761</td>
</tr>
<tr>
<td>2</td>
<td>0.299</td>
<td>0.084</td>
<td>0.194</td>
<td>0.396</td>
<td>694,761</td>
</tr>
<tr>
<td>1</td>
<td>0.169</td>
<td>0.051</td>
<td>0.105</td>
<td>0.243</td>
<td>694</td>
</tr>
</tbody>
</table>

**SCOPE: LAI (best at 44 bands)**

**LWC (best with 116 bands, stable results from 4 bands onwards)**

<table>
<thead>
<tr>
<th>#</th>
<th>R2</th>
<th>SD</th>
<th>MIN</th>
<th>MAX</th>
<th>Wavelengths (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.70</td>
<td>0.07</td>
<td>0.64</td>
<td>0.82</td>
<td>669,700,701,760,761</td>
</tr>
<tr>
<td>4</td>
<td>0.70</td>
<td>0.07</td>
<td>0.64</td>
<td>0.82</td>
<td>669,701,760,761</td>
</tr>
<tr>
<td>3</td>
<td>0.35</td>
<td>0.11</td>
<td>0.18</td>
<td>0.46</td>
<td>701,760,761</td>
</tr>
<tr>
<td>2</td>
<td>0.35</td>
<td>0.11</td>
<td>0.18</td>
<td>0.46</td>
<td>701,761</td>
</tr>
<tr>
<td>1</td>
<td>0.18</td>
<td>0.12</td>
<td>0.07</td>
<td>0.38</td>
<td>701</td>
</tr>
</tbody>
</table>

**LCC (best with 10 bands)**

<table>
<thead>
<tr>
<th>#</th>
<th>R2</th>
<th>SD</th>
<th>MIN</th>
<th>MAX</th>
<th>Wavelengths (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.77</td>
<td>0.07</td>
<td>0.68</td>
<td>0.85</td>
<td>668,669,695,719,720,726,749,750,794,832</td>
</tr>
<tr>
<td>9</td>
<td>0.75</td>
<td>0.08</td>
<td>0.64</td>
<td>0.85</td>
<td>668,669,695,719,720,726,749,750,794</td>
</tr>
<tr>
<td>8</td>
<td>0.75</td>
<td>0.08</td>
<td>0.64</td>
<td>0.85</td>
<td>668,669,695,719,720,726,749,750,794</td>
</tr>
<tr>
<td>7</td>
<td>0.75</td>
<td>0.08</td>
<td>0.64</td>
<td>0.85</td>
<td>668,669,695,719,720,726,749,750,794</td>
</tr>
<tr>
<td>6</td>
<td>0.72</td>
<td>0.09</td>
<td>0.62</td>
<td>0.83</td>
<td>669,695,719,726,749,750</td>
</tr>
<tr>
<td>5</td>
<td>0.72</td>
<td>0.09</td>
<td>0.62</td>
<td>0.83</td>
<td>669,695,719,748,750</td>
</tr>
<tr>
<td>4</td>
<td>0.72</td>
<td>0.09</td>
<td>0.62</td>
<td>0.83</td>
<td>669,695,719,748,750</td>
</tr>
<tr>
<td>3</td>
<td>0.68</td>
<td>0.09</td>
<td>0.55</td>
<td>0.86</td>
<td>669,719,749</td>
</tr>
<tr>
<td>2</td>
<td>0.54</td>
<td>0.13</td>
<td>0.32</td>
<td>0.68</td>
<td>719,749</td>
</tr>
<tr>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.20</td>
<td>749</td>
</tr>
</tbody>
</table>
II) Dimensionality reduction: SIMFEAT

13 dimensionality reduction methods implemented.

Experimental setup:

PROSAIL: 500 random samples

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1.3</td>
<td>2.5</td>
</tr>
<tr>
<td>Cab</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>Cw</td>
<td>0.002</td>
<td>0.05</td>
</tr>
<tr>
<td>Cm</td>
<td>0.002</td>
<td>0.05</td>
</tr>
<tr>
<td>LAD</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>LAI</td>
<td>0.01</td>
<td>7</td>
</tr>
</tbody>
</table>

LAI input samples

directional reflectance (2101 bands)
Impact of DR methods on PROSAIL data (2101#b) for LAI retrieval

<table>
<thead>
<tr>
<th>No DR (2101#)</th>
<th>PCA (5#)</th>
<th>Best DR method (5#)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LR</strong></td>
<td></td>
<td>CCA</td>
</tr>
<tr>
<td>R2: 0.1</td>
<td>R2: 0.48</td>
<td>R2: 0.92</td>
</tr>
<tr>
<td>NRMSE: 179.7%</td>
<td>NRMSE: 21.3%</td>
<td>NRMSE: 8.2%</td>
</tr>
</tbody>
</table>

By combining advanced DR methods with (advanced) regression methods, hyperspectral data can be exploited to the fullest. Or, full spectral dataset into regression require (advanced) DR methods.
2) Sample reduction: Active learning (AL)

- Active learning (AL) searches for new samples from a data pool based on **uncertainty** (PAL, EQB, RSAL) and **diversity** (ABD, CBD, EBD).

- AL method search more efficiently for relevant samples than random sampling or when using all data.

**Best-performing method can be applied to an image.**
Background LUT-based inversion

Best match is obtained through a ‘Cost function’, or ‘Minimum distance function’.

<table>
<thead>
<tr>
<th>Cost Function</th>
<th>Definition</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{LS}$</td>
<td>$\sum(Y - \bar{X})^2$</td>
<td>0</td>
<td>$\infty$</td>
</tr>
<tr>
<td>$C_{NC}$</td>
<td>$\frac{\sum(XY)}{\sqrt{\sum X^2} \sqrt{\sum Y^2}}$</td>
<td>$-1$</td>
<td>1</td>
</tr>
<tr>
<td>$C^W$</td>
<td>$\sum k \frac{\text{Var}(Y_k)}{N}$</td>
<td>0</td>
<td>$\infty$</td>
</tr>
<tr>
<td>$C_{GR}$</td>
<td>$\frac{1}{\text{Var}(Y)} \sum k \frac{\text{Var}(Y_k)}{N}$</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Other important factors:
- Adding **noise** (to account for natural variability)
- Selecting mean/median of multiple solutions
LUT-based inversion toolbox:

Properties:
- LUT ARTMO RTMs or external LUT
- Over 60 different cost functions
- Noise & multiple solutions
- Results stored in MySQL
- Top-performing inversion strategies are given.
- Can apply inversion to both image or individual spectra.

Matching a pixel against a part of the LUT.

Best-performing method can be applied to an image.
SPARC dataset, CHRIS resampled to S2, S3

Optimized cost function for each variable.

One optimized cost function for simultaneous retrieval of multiple variables.

<table>
<thead>
<tr>
<th>Sentinel</th>
<th>S2-10m</th>
<th>S2-20m</th>
<th>S3-300m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial resolution [m]</td>
<td>10</td>
<td>20</td>
<td>300</td>
</tr>
<tr>
<td># bands</td>
<td>4</td>
<td>8 (4 + 4 at &lt; 20 m)</td>
<td>19</td>
</tr>
<tr>
<td>Band position</td>
<td>B2, B3, B4 and B8</td>
<td>B2 to B8a</td>
<td>O2 to O20</td>
</tr>
<tr>
<td>Wavelengths [nm]</td>
<td>490-665 and 842</td>
<td>490-865</td>
<td>413-940</td>
</tr>
</tbody>
</table>

Estimates

- Chl [µg/cm²]
- LAI [m²/m²]
- Chl × LAI [µg/cm²]

Uncertainties (CV)

- Chl
- LAI
- Chl × LAI
SIMPLIFYING

- Global sensitivity analysis
- Emulation
- Retrieval
Global sensitivity analysis

Global sensitivity analysis: explores the full input parameter space, i.e. all input parameters are changed together.

Variance-based methods: the output variance is decomposed to the sum of contributions of each individual input parameter and the interactions (coupling terms) between different parameters.

Based on the work of Sobol’, variance-based sensitivity measures are represented as follows:

\[ 1 = \sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \cdots + S_{12}, \ldots, k, \]

in this equation, \( S_i, S_{ij}, \ldots, S_{12}, \ldots, k \) are Sobol’s global sensitivity indices.

The first order sensitivity index \( S_i \) measures and quantifies the sensitivity of model output \( Y \) to the input parameter \( X_i \) (without interaction terms), whereas, \( S_{ij}, \ldots, S_{12}, \ldots, k \) are the sensitivity measures for the higher order terms (interaction terms).

The total effect sensitivity index \( S_{T_i} \) measures the whole effect of the variable \( X_i \), i.e. the first order effect as well as its coupling terms with the other input variables:

\[ S_{T_1} = S_1 + S_{12} + S_{13} + S_{123} \]
**GSA toolbox**

**PROSPECT-4 Reflectance (#1000)**

**PROSAIL Directional Reflectance (#1000)**

**Properties:**
- ARTMO RTMs
- *Saltelli 2010* GSA method
- Various sample distributions
- Results stored in MySQL
- First order or total order Sobol Sensitivity indices
- Can process multiple RTM outputs.
Any difference?

Which one would you choose?

 SCOPE

Metamodel (emulator)
Emulators are regression models that are able to approximate the processing of an RTM, at a fraction of the computational cost: 

*making a statistical model of a physical model*

**Emulators applied to RTMs:**

- In principle any nonlinear, adaptive machine learning regression algorithms (MLRAs) can serve as emulators.

- To emulate RTMs, the emulator should have the capability to reconstruct multiple outputs, i.e. the complete spectrum: resolved with dimensionality reduction techniques (e.g. PCA).

**Processing steps:**

1. Input (variables + spectra)
2. Splitting into training/validation
3. PCA on spectra
4. MLRA training looping over components
5. Prediction of components
6. Reconstruction of full spectrum
7. Validation
8. Emulator
**Emulating SCOPE fluorescence outputs**

Because of the smooth profiles, SIF outputs are easy to emulate.

In Emulation, physical models go hand in hand with machine learning.
Emulating a complex 3D RTM: DART

Experimental setup:
- DART: LUT1000# @ 1 nm; 7 variables
- 3 MLRAs tested: KRR, NN, GPR
- Various # PCA components tested (5, 10, 20, 30)

Emulator
- KRR
- NN
- GPR

Experimental setup:
- DART: LUT1000# @ 1 nm; 7 variables
- 3 MLRAs tested: KRR, NN, GPR
- Various # PCA components tested (5, 10, 20, 30)

Emulator
- KRR
- NN
- GPR
Emulators applied into GSA: PROSPECT-4

PROSPECT reference (122 s)

KRR (55 s)

NN (55 s)

GPR (58 s)

1000#/variable
PROSAIL emulators

PROSAIL reference (279 s)

KRR (73 s)

NN (80 s)

GPR (79 s)

Total SI [%]

Wavelength [nm]

1000#/variable

N  Cw  Cab  Cm  LAI  LAD  skyl  soil coeff  SZA  VZA  RAA
MODTRAN atmospheric transfer functions:

\[ L_{TOA} = L_0 + \frac{(E_{\text{dir}} \mu_s + E_{\text{dif}})(T_{\text{dif}} + T_{\text{dir}}) \rho}{\pi(1 - S\rho)} \]

- \(E_{\text{dif}}\) (GPR: 121 s)
- \(T_{\text{dif}}\) (NN: 157 s)
- \(L_0\) (GPR: 121 s)
- \(E_{\text{dir}}\) (KRR: 101 s)
- \(T_{\text{dir}}\) (NN: 151 s)
- \(S\) (GPR: 166 s)

With emulators, processing speed is boosted in the orders of hundred thousand. Using the original MODTRAN simulations would take more than a month.

1000#/variable

VZA  SZA  RAA  ELEV  AOT  AMS  G  CWV
Emulators into numerical inversion

Minimization algorithm: lsqnonlin

Output maps of RTM variables

Per-pixel RTM iterations: very slow method, inapplicable to computationally expensive RTMs.

Emulation of an RTM
DART KRR emulator applied to HyPlant DUAL (450-2500 nm)

SCOPE KRR emulator applied HyPlant DUAL (bare soil spectra added)

Retrieval quality depends on: (1) emulator, (2) number and type of included variables.
Conclusions

Tools
- Leaf RTMs
  - COSINE
  - PROSPECT-5
  - DLM
  - Fluspect-B
  - LIBERTY
- Canopy RTMs
  - SAIL
  - INFORM
  - FLIGHT
- Combined RTMs
  - SCOPE
- Retrieval Toolboxes
  - MODTRAN
  - 6Sv
  - libRadtran
  - Numerical inversion
  - Spectral indices
  - LUT-based inversion
  - MLRA retrieval
  - Spatial resampling
  - Global sensitivity analysis (GSA)
  - Emulator
  - Graphics
  - Sensor
  - Scene Generator

Thanks

http://ipl.uv.es/artmo/