<u>Scientific report on the</u>

<u>STSM of Dr. Sarah Asam at the Laboratory of Image Processing (IPL, University of Valencia)</u>

09.03. – 27.03.2015

Purpose of the STSM

The main purpose of the STSM of Dr. Sarah Asam at the Laboratory of Image Processing (IPL) at the University of Valencia was to get to know in depth the Automated Radiative Transfer Models Operator (ARTMO) software, and to use it for biophysical parameter derivation from remote and proximal sensing data. Since Ms. Asam has already worked on the derivation of grassland Leaf Area Index (LAI) from high spatial resolution optical remote sensing data using a radiation transfer model (RTM) during her PhD, her current task within the MONALISA project (www.monalisa-project.eu) comprehends the derivation of biophysical parameters for different biomes on different scales, and hence from different earth observation data. Therefore, different RTMs and a range of inversion optimization approaches need to be tested. The ARTMO software is a valuable tool for this kind of analysis, as different models, inversion algorithms, and sensor characteristics can be quickly set up.

Work carried out during the STSM

After installing the latest ARTMO version and creating the necessary training data files and satellite imagery formats, Ms. Asam used the ARTMO toolbox in a first step on multispectral RapidEye data to carry out a range of analysis:

- The results achieved with different leaf models, namely the PROSPECT-4 and PROSPECT-5 models, were compared first.
- The parameterization of the RTM adapted to the grassland biome was set up based on field measurements carried out in the Ammer catchment in 2011/2012 and refined based on the spectral match between the RTM output and the remotely sensed reflectances. Since the ARTMO toolbox does not allow for pixel-wise selection of Look-up tables (LUT) entries based on specific illumination and viewing angles, the specification of the sensor and sun zenith angles as well as the relative azimuth angle included all occurring local angles of the scene during the forward modeling as well as during model inversion.
- Different sampling distributions of the input parameters in the range of predefined boundary values, namely uniform, normal, exponential, and 'fullRandom' distributions, were tested.
- The influence of the LUT size on the LAI retrieval accuracy was systematically tested within the range of 5000 to 180000 LUT entries.
- For the inversion of the differently generated LUTs, a range of optimization techniques was systemically tested each time:
    - the use of multiple solutions
    - the addition of different levels of synthetic noise
    - the use of 15 different cost function out of 3 mathematical groups
    - the usage of normalization/non-normalization within the cost function

ARTMO iteratively uses all cost functions, noise levels and multiple solution samples selected by the user and automatically identifies the best performing setting.

After identifying the preferable settings, an LAI map as well as inversion statistic layers (inversion residuals and the standard deviation and coefficient of variation of LAI estimates in a multiple solution) were derived based on the RapidEye image. Then, the same optimization was performed for the use of an airborne hyperspectral data set (HySpex). Firstly, a LUT covering the 336 spectral bands of the HySpex sensor was generated based on the optimal LUT size identified before and based on a similar parameterization (only difference: a varying leaf water content parameter). To optimize the model parameterization for the HySpex data, soil spectra sampled from the HySpex image were additionally used as input parameter. An LAI map was then derived from the HySpex data based on the best performing inversion settings. Additional to the LAI information, the inversion statistics as well as the local geometry angles were derived. Also, the influence of using only a subset of all HySpex features for inversion has been tested.

In a last step, the results generated with the RTM and LUT-based inversion approach were compared to those generated with a range of 13 non-parametric machine learning regression algorithms (MLRA). In the special case of using Gaussian processes, associated uncertainty estimates, expressed as the standard deviation or the coefficient of variation, can be mapped and relative band importance can be derived, which is useful for further feature reduction. As for the LUT approaches, ARTMO enables a systematic trial of different noise settings and training/validation partitioning for the training of the MLRAs. Due to the low number of field measurements, a 5-fold cross-validation was chosen. After MLRA-inversion was applied to the complete data set, a range of automatic feature selection procedures (standard principle component analysis as well as transformations used in the yet unpublished SIMFEAT module) as well as the simulation of sensor-specific settings that include a lower number of bands (RapidEye, Landsat, Sentinel-2) was tested.

Main results

- LAI maps were be derived based on multispectral as well as hyperspectral data with a lowest normalized RMSE of 14 – 18 % given different inversion settings, using both physical as well as statistical machine learning techniques. For most maps, an uncertainty estimate layer could be generated in addition.
- The comparison of PROSPECT4 and PROSPECT5 revealed no significant differences for LAI derivation accuracy (lowest achieved LAI derivation RMSEs of 14.66 and 14.91, respectively). Therefore, the PROSPECT 4 was selected for further use since it has fewer parameters.
- The ‘fullRandom’ sampling of model input parameters implemented in the ARTMO toolbox, that generates parameter combinations based on continuous sampling and random matching of the different parameters, achieved best results (see Figure 1). In addition, with this option continuous LAI values are stored in the LUT, which enables the derivation of continuous LAI values during inversion of the LUT.
- Overall, using a non-normalized distance measure achieves higher accuracies in most inversion settings (see Figure 1 and Figure 2).

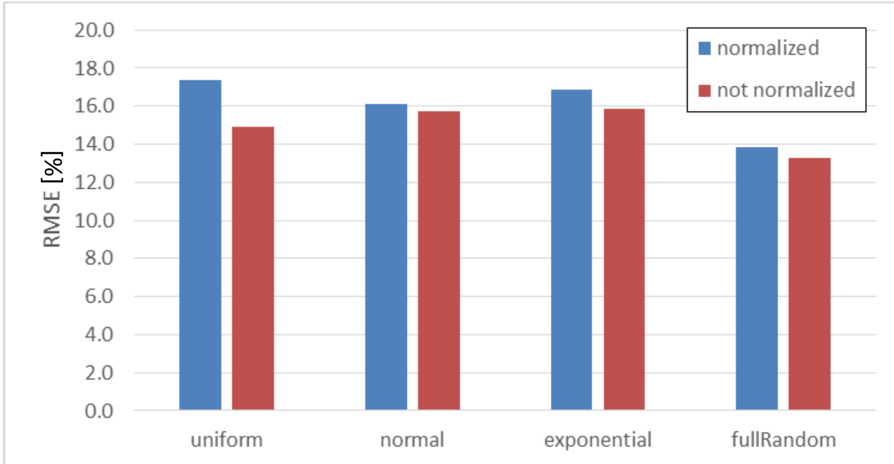

Fig. 1: Normalized RMSE values of the respective best performing inversion setting using different RTM parameter distributions as well as cost function normalization.

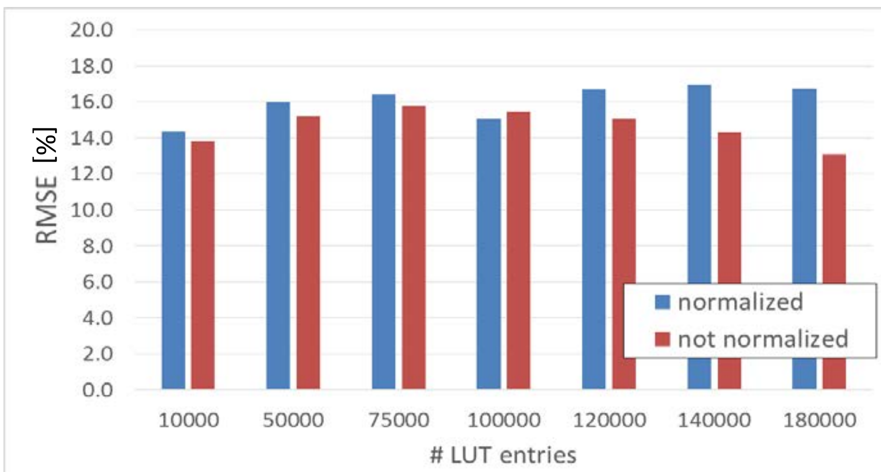

Fig. 2: Influence of the number of LUT entries on the LAI estimation accuracy. In all LUTs a full random sampling over the same parameter ranges has been performed. In addition, for each LUT a normalized and a non-normalized inversion was tested.

- The analysis of the influence of the LUT size on the retrieval accuracy revealed that a reduction of the LUT size down to 10000 entries does not afflict the mapping accuracy significantly (see Figure 2 and Figure 3). This is a relevant finding since LAI mapping based on a LUT inversion is speed up a lot when using smaller LUTs.
- The selection of the cost function has a bigger influence on the inversion accuracy than the number of LUT entries (see Figure 3). A range of approximately 15 % normalized RMSE is spanned by the respective best and worst performing cost function. However, no single best-performing cost function could be identified, but the chosen cost function varies with other settings such as the sampling distribution or normalization. In addition, no group of cost functions (i.e. of 'divergence information measures', 'Maximum-likelihood type measures', and 'Minimum constraint estimation') performed generally better than the others. Although

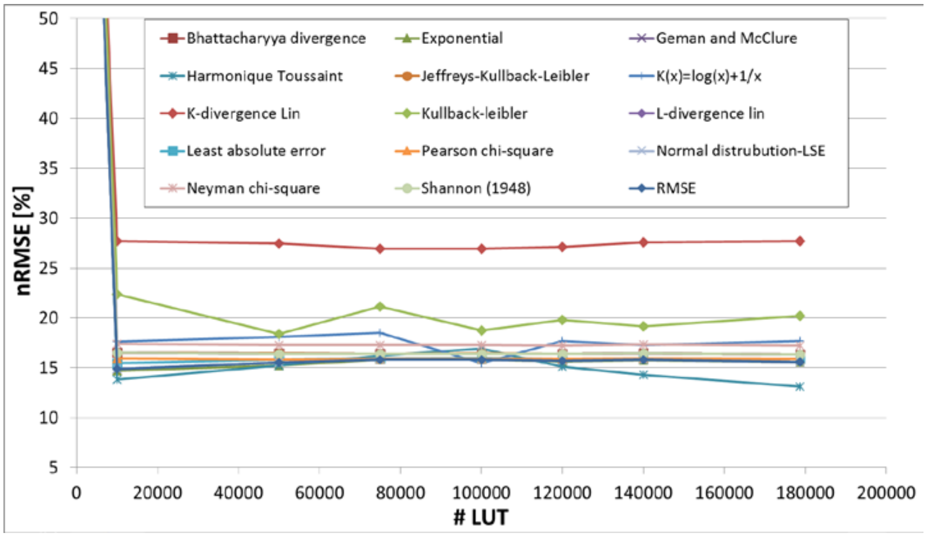

Fig. 3: Normalized RMSE values of the LAI estimation depending on the number of LUT entries and on the selected cost function.

the usually used root mean square error (RMSE) almost never performed best, it was however under the best five performing cost functions in about two third of all trial runs. It achieved in average only 1-2% higher errors than the respective best performing algorithm.

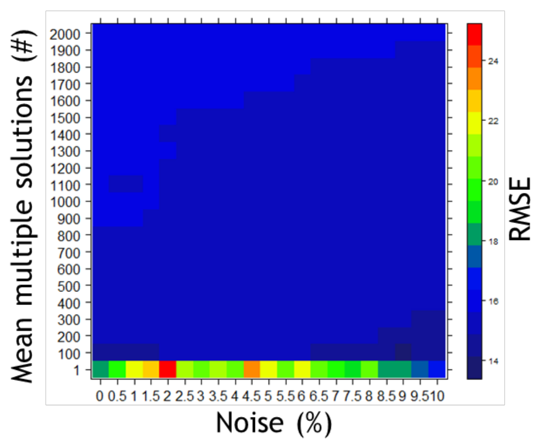

Fig. 4: Heat map of the RMSE of LAI derivation based on the added synthetic noise and on the use of multiple solutions.

- Not using the single best solution but the mean of a range of best estimation accuracy solutions within the LUT increases the estimation accuracy in most cases (see Figure 4 as an example). The use of synthetic noise levels does not show such a clear pattern. While in Figure 4 the estimation accuracy increases with increasing noise levels, this pattern varies for other data sets or inversion settings (i.e. parameter sampling distribution, LUT size, or cost function). This optimization should therefore ideally be conducted separately for each new data set.
- Inverting the LUT over each pixel without fixing the local viewing and illumination angles for each pixel seems not to affect the LAI derivation accuracy. However, when comparing the true

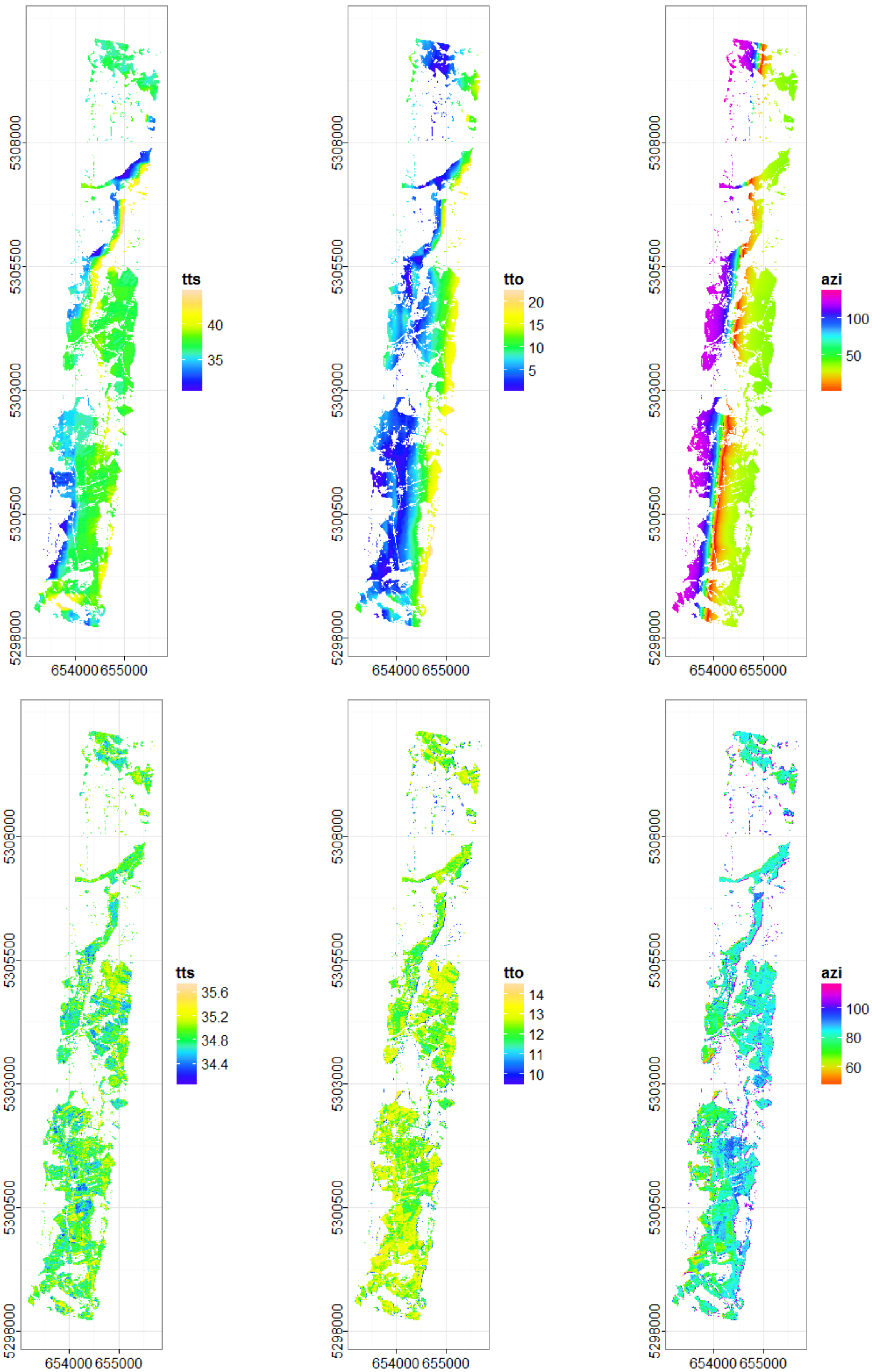

Fig. 5: Comparison of the true sun zenith (tts), sensor zenith (tto) and relative azimuth angles (azi) of HySpex flight strip A02 in the trop row to the respective angle maps derived through LUT inversion in the bottom row.

- sun and sensor zenith angles as well as the relative azimuth maps (all corrected for local topography) with the respective angle maps derived through inversion (see Figure 5), it becomes clear that the correct system geometry is not reproduced. Obviously only a small range of simulated angles is selected, and the patter of the maps is different. While the true angle maps reflect the flight direction, sensor characteristics und underlying topographic conditions, the modeled angles seem rather to change with individual fields, i.e. probably compensating for spectral changes due to vegetation. This underlines the necessity to account for the local geometry during inversion, although a sensitivity analysis on the influence of the three angles is needed to quantify this error.
- The MLRA techniques and feature selection approaches used achieve slightly better results than the LUT-based inversion and are much faster. However, due to the small number of field sampling data available, the full power of these methods could probably not been exploited. In addition, the MLRAs trained on the field data are sometimes unstable, in the sense that no or no reasonable values are derived when mapping LAI.

The conducted analysis overall improved the LAI derivation accuracy by about 15 % in comparison to previous works based on the RapidEye data, and provided very useful findings for the optimized setup of LUT based biophysical parameter derivation. Further, the ARTMO software simplified working with hyperspectral data. This work laid the foundation for further research on biophysical parameter estimation also of biomes that are more complex and the estimation of ecosystem processes using the SCOPE model.

For the host institution, the work carried out during the STSM revealed some needs for clarification within the software handbook and a couple of minor bugs. Updated ARTMO versions were published on March 9th, 21st, and 26th 2015.

Future collaboration

Since the STSM provided the possibility for Ms. Asam to gain a comprehensive knowledge on the structure and functioning of the ARTMO toolbox, some advancements on ARTMO were discussed during the stay of Ms. Asam. One option would be to implement an application that enables the use of variable scene geometries, i.e. an algorithm that identifies all occurring scene geometries first and adapts the identified angles or angle rages for the spectral modeling. During inversion, this application would enable to subset the LUT based on the pixel specific geometries and hence reducing the dimensionality problem. This extension would enable the correct analysis of airborne data, of composite images, time series, as well as of images acquired over complex terrain. The development of such an application by Ms. Asam is foreseen.

Foreseen publications from the STSM

An output of the STSM is a contribution to the 9th EARSeL 'SIG Imaging Spectroscopy Workshop' which was presented on April 15th, 2015, on the use of hyperspectral data for grassland LAI derivation in comparison to the use of multispectral data. On the same subject, a peer-reviewed publication in the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS) special issue is planned.

Confirmation by the host institution of the successful execution of the STSM

VNIVERSITAT
DE VALÈNCIA

Image Processing Laboratory (IPL)
Unidad de Administración y Servicios
P.O. Box 22085
E-46071 Valencia, Spain

01/04/2015

To whom it may concern,

This is to confirm that Dr. Sarah Asam stayed for three weeks for a lab visit at the Laboratory of Earth Observation (LEO) at the Image Processing Laboratory (IPL) at the University of Valencia.

She started joining our group on March 9th and she stayed until March 27th 2015.

During those three weeks Sarah Asam learned various techniques regarding the retrieval of biophysical parameters by using the ARTMO toolbox. She was supervised by Dr. Jochem Verrelst.

Kind regards,

Jochem Verrelst