THE RAMI ON-LINE MODEL CHECKER (ROMC)

J.-L. Widlowski [a], *, M. Robustelli [a], M. Disney [b], J.-P. Gastellu-Etchegorry [c], T. Lavergne [e], P. Lewis [b], E. Martin [c], P. R. J. North [d], B. Pinty [a], M. Taberner [a], R. Thompson [f], M. M. Verstraete [a]

[a] DG Joint Research Center, European Commission, Via E. Fermi, 1, 21020 Ispra (VA), Italy
(Jean-Luc.Widlowski@jrc.it, Bernard.Pinty@jrc.it, Monica.Robustelli@ext.jrc.it, Malcolm.Taberner@jrc.it)
[b] Department of Geography, University College London, 26 Bedford way, London, WC1H 0AP, UK and NERC Centre
for Terrestrial Carbon Dynamics (mdisney@geog.ucl.ac.uk, plewis@geog.ucl.ac.uk)
[c] Centre d'Etudes Spatiales de la Biosphère (CESBIO), 18 avenue Edouard Belin, bpi 2801, 31401 Toulouse cedex 9,
France (jean-philippe.gastellu@cesbio.cnes.fr, emmanuel.martin@cesbio.cnes.fr)
[d] Climate and Land-Surface Systems Interaction Centre, Department of Geography, University of Wales Swansea,
Singleton Park, Swansea, SA2 8PP, UK (p.r.j.north@swansea.ac.uk)
[e] Section for Remote Sensing, Research and Development Department, Norwegian Meteorological Institut, P.O.Box
43, Blindern, N-0313 Oslo, Norway (thomasl@met.no)
[f] Alachua Research Institute, Alachua, Florida, USA (richardlthompson@cox.net).

KEY WORDS: benchmarking, radiation transfer model intercomparison, RAMI, quality control, BRDF, vegetation canopy.

ABSTRACT:

The RAMI On-line Model Checker (ROMC) is a web-based tool (http://romc.jrc.it/)[**] allowing for the automated evaluation of canopy radiation transfer (RT) models. This is achieved by comparing simulation results submitted by registered ROMC users to a reference dataset established from credible 3-D Monte Carlo models during the third phase of the RAdiation transfer Model Intercomparison (RAMI) exercise. Access to the ROMC is free and allows users to assess their model performance either in "debug mode" (if users want to evaluate the impact of changes in their model's software code/parameter input with respect to some user-selected test cases identical to those in RAMI), or, in "validate mode" (if users want to evaluate the performance of their model against a random selection of test cases not included within RAMI). Upon successful submission, registered ROMC users will be presented with various results statistics and graphical displays that document the performance of their model with respect to the reference dataset. In addition to providing an indication of the overall skill of a model to correctly match the reference data, the ROMC also allows for interactive comparison/evaluation of the differences existing between various versions of the same model. Furthermore, ROMC users can download encapsulated postscript versions of these graphical results files, and – in debug mode only – may also receive ASCII files containing the ROMC reference data. ROMC graphs come with a reference number and may be used in presentations and – in validate mode – also in refereed scientific publications to underline the quality of a given model. As such the ROMC is a good example, on how a concerted, voluntary community effort (i.e., RAMI) has led to tangible results that benefit both model developers and their customers.

1. INTRODUCTION

Indicators as to the quality of physically-based radiation transfer (RT) models are of relevance to their developers and users, as well as the scientists, space agencies and policy makers that use or support the products and information derived on the basis of such model simulations/inversions. The radiation transfer model intercomparison (RAMI) initiative was launched in an attempt to shed light on the reliability and accuracy of existing canopy radiation transfer models. RAMI is a triennial community exercise that encourages the systematic evaluation of canopy reflectance models under well-controlled experimental conditions and on a voluntary basis (Pinty et al., 2001, 2004, Widlowski et al., 2007). Significant efforts were being made during the first (1999) and second (2002) phases of RAMI in order to document and reduce the dispersion between increasing numbers of participating models, but it was not until the completion of the third phase (2005) that a sufficiently

strong consensus emerged among RT models capable of simulating the entire palette of RAMI test cases: from simple plane-parallel turbid medium scenarios to complex heterogeneous vegetation canopies with and without underlying topography. During RAMI-3 the mutual agreement between the reflectance simulations of six state-of-the-art 3-D Monte Carlo models was estimated to lie at ~1%. This is significantly lower than anything previously achieved and well below the absolute accuracy of current space borne measurements, which prompted the usage of these model simulations in order to establish a reference dataset against which other/future radiative transfer models may now be compared.

2. THE ROMC REFERENCE DATA SET

The self-consistency (e.g., energy conservation) together with the absolute and relative performance of models were evaluated in great detail during RAMI-3. Substantial improvements in

---

* Corresponding author; ** Due to a renaming of all European Commission websites this URL is likely to change in the near future to http://romc.jrc.ec.europa.eu/ .

model agreement were observed, in particular, for the 3-D Monte Carlo (MC) models that participated. These models allow for explicit 3-D representations of complex canopy architectures and avoid unnecessary assumptions and approximations in the solving of the radiation transfer equation due to their stochastic sampling of the surface leaving radiation field (Disney et al., 2000). The average model-to-ensemble dispersion between the total BRF simulations of six 3-D MC models (e.g., dart (Gastellu-Etchegorry et al., 2004), drat (Lewis 1999), flight (North, 1996), rayspread (Widlowski et al., 2006), raytran (Govaerts and Verstraete 1998) and sprint3 (Goel and Thompson 2000)) was found to have almost halved from RAMI-2 (1.37 %) to RAMI-3 (0.72 %) for canopies with finite-sized foliage, whereas in turbid medium cases it improved by a factor of ~7 from RAMI-2 (6.36 %) to RAMI-3 (0.91 %) – see Table 1 for exact values of the dispersion in RAMI-3 (computed according to equations given in section 3.3.2 of Widlowski et al., 2007). The overall performance of these six 3D MC models in RAMI-3 thus supported their usage in the generation of a "surrogate truth" against which other RT models could then be evaluated.

model name	foliage description	
	discrete	turbid
dart	-	1.46
drat	0.55	-
flight	0.97	1.06
rayspread	0.55	0.64
raytran	0.60	0.69
sprint3	1.01	0.69

Table 1. Model-to-ensemble dispersion (in percent) for the total BRF simulations over both homogeneous and heterogeneous canopy scenes for six "credible" 3-D Monte Carlo models identified during the third phase of RAMI.

In order to obtain a "surrogate truth" estimate it was decided to simply average the BRFs that were obtained from the above set of credible MC models. Due to variable model participation (and performance) the precise number and names of the available 3-D MC models could, however, change from one experiment and measurement to the next. As a consequence, the number and names of the credible models that featured in the computation of the "surrogate truth" had to be chosen separately for every RAMI experiment and measurement. The following list of criteria was finally applied to derive the "surrogate truth" estimates from RAMI-3 simulations:

1. For every RAMI BRF (flux) measurement, identify at least two (one) 3-D Monte Carlo models that do not belong to the same RT modelling school/family,
2. If two (or more) models from the same RT modelling school/family are available, e.g., rayspread and raytran, choose the one with the least amount of apparent MC noise,
3. Remove all those 3-D Monte Carlo models from the reference set that are noticeably different from the main cluster of 3-D MC simulations,
4. If sufficient models are contained in the main cluster of 3-D MC simulations then remove those models that would introduce noticeable levels of "MC noise" into the reference set,

5. If there are two distinct clusters of 3-D Monte Carlo models, or, no obvious cluster at all, then use all available 3-D RT models to define a reference solution.

A drawback of this selection procedure lies in the fact that the ROMC reference dataset may not be fully compliant with energy conservation. This is a direct consequence of 1) not all credible models performing all test cases and measurements, and 2) the variable performance of credible models for different measurements types, as well as, structural, spectral and illumination scenarios. Nevertheless, the average absolute deviation of the ROMC reference dataset from energy conservation was found to be 0.00025 in the red, 0.00007 in the near-infrared, and $9.5*10^{-7}$ for conservative scattering conditions (purist corner). A synoptic table featuring the names and performances of the various 3-D Monte Carlo models that contributed toward the computation of the "surrogate truth" for all the RAMI-3 experiments and measurement types individually, can be found on the following internet page: http://romc.jrc.it/WWW/PAGES/ROMC_Home/RAMIREF.php
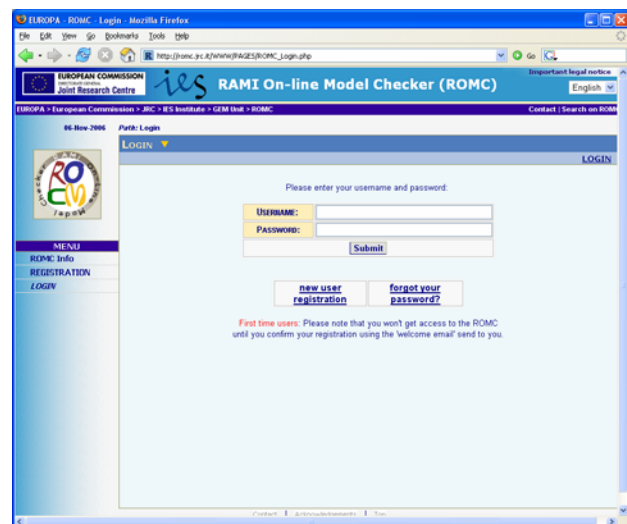

Figure 1. The RAMI On-line model checker (ROMC) website.

3. THE RAMI ON-LINE MODEL CHECKER

With this valuable reference dataset at hand, it becomes possible to allow canopy RT model owners, developers and customers to evaluate the performance of a given model even outside the frame of RAMI. To facilitate such an undertaking the RAMI On-line Model Checker (ROMC) was developed at the Joint Research Centre of the European Commission in Ispra, Italy. The ROMC is a web-based interface allowing for the on-line evaluation of RT models using as reference the "surrogate truth" derived from among the six credible 3-D Monte Carlo models identified during RAMI-3. Access to the ROMC can be obtained either via the RAMI website or directly using the URL[***] http://romc.jrc.it/. After providing a username and valid email address, the ROMC can be utilised in two different ways: 1) in "debug mode", which allows to repeatedly compare the output of a RT model to that of one or more experiments and/or

---

[***] Due to a renaming of all European Commission websites this URL is likely to change to http://romc.jrc.ec.europa.eu/.

measurements from RAMI, i.e., the simulation results are known since they are available on the RAMI website, and 2) in "validate mode", which enables the once-only testing of the RT model against a randomly selected set of test cases that are similar but not quite equivalent to those from RAMI, i.e., the solutions are not known *a priori* and the model runs cannot be repeated.

- In "debug mode" users may choose to execute one particular experiment and/or measurement from the set of RAMI-3 test cases *ad infinitum*, or, at least until they are satisfied with the performance of their model. Detailed descriptions of the structural, spectral, illumination and measurement conditions are available. Once the ROMC user has generated his/her model simulations, they can be up-loaded via the web-interface, and - provided that they adhere to the RAMI filenaming and formatting conventions - this process will result in a series of graphical results files being made available for all test cases. In debug mode users may not only download their ROMC results but also an ASCII file containing the actual "surrogate truth" data.

- In "validate mode" ROMC users may choose between structurally homogeneous leaf canopies and/or heterogeneous "floating spheres" scenes in order to verify the performance of their model. The actual set of test cases will, however, be drawn randomly from a large list of possible ones, such that it is unlikely to obtain the same test case twice, i.e., in all likelihood one will not "know" the solution *a priori*. Again, the "surrogate truth" was derived from simulations generated by models belonging to the same set of 3-D MC models as was the case for the debug mode. In validate mode the reference data will, however, not be available for downloading. The procedure for data submission, on the other hand, is identical to that of the debug mode, and - provided that all RAMI formatting and filenaming requirements are applied - leads to a results page featuring a variety of intercomparison graphics.

Registered ROMC users may download their model performance results either as jpeg formatted images directly from the ROMC website, or else, opt for receiving them via email in postscript form. Both the debug and validate mode ROMC results files feature a unique reference number. Graphs are available for individual experiments (see Figure 2) as well as for the ensemble of submitted experiments (see Figure 3). Currently the set of downloadable graphs include 1) plots of Bidirectional Reflectance Factors (BRF) from both the user's model and the ROMC reference dataset along either the principal or the orthogonal plane, 2) 1 to 1 plots of the model and reference BRFs (or fluxes), 3) histograms of the deviations between the model and reference BRFs, 4) χ^2 graphs depicting model equivalency with the reference data set for the various submitted measurements, 5) graphs depicting the deviation of the model and reference fluxes using barcharts, and 6) Taylor diagrams (Taylor, 2001) expressing model performance in terms of correlation, standard deviation and root mean square error (see Figure 4). Statistics like the root mean square error (RMS), the signal to noise ratio (S/N) and the number of experiments included are also provided. In addition some these graphs can be received in black and white only form.
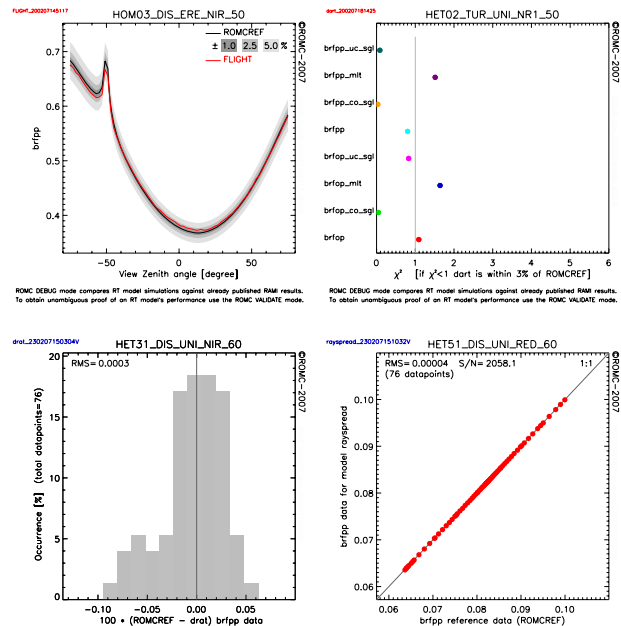

Figure 2. Examples of downloadable ROMC graphics showing model performance for individual experiments in debug mode (top) and validate mode (bottom). Shown are plots of user and reference BRFs (top left), a χ^2 graph (top right), a histogram of BRF differences (bottom left), and a 1:1 plot (bottom right).
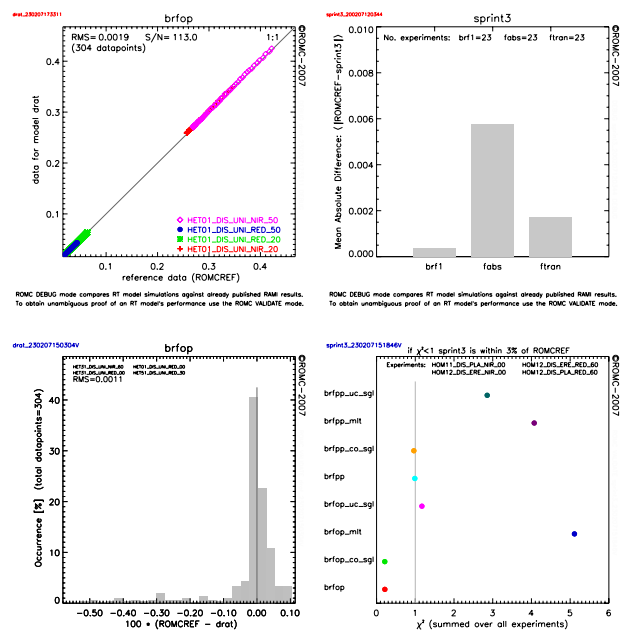

Figure 3. Examples of ROMC graphics summarising model performance for ensembles of submitted experiments carried out in debug mode (top row) and validate mode (bottom row). Shown are a 1:1 BRF plot (top left), a barchart of mean absolute flux differences (top left), a histogram of BRF differences (bottom left), and a χ^2 graph (bottom right).
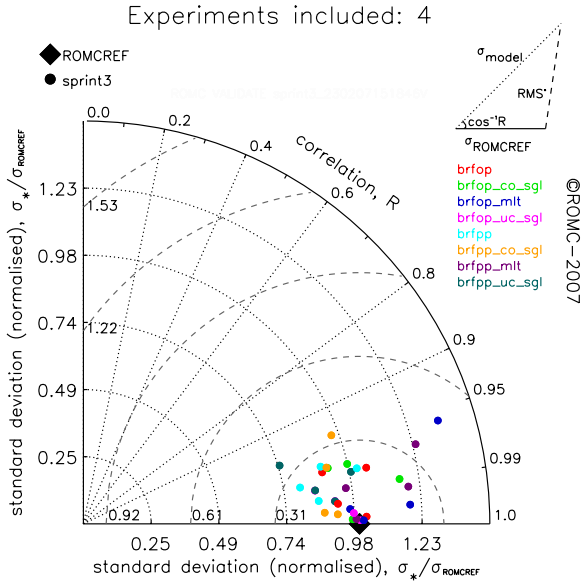
Experiments included: 4

Figure 4. Examples of a Taylor graph in validate mode depicting model performance in terms of the ability to match the pattern (correlation R) and variability (normalised standard deviation) of the reference solution. The RMS difference error is also indicated (dashed isolines).

In order to provide an overall indication of a model's capability to match the reference data set a skill metric is defined and applied to each measurement type. The ability of a skill metric should be such that 1) for any given variance the score increases monotonically with increasing correlation, and 2) for any given correlation the score increases as the modelled variance approaches the variance of the reference data set. For visualisation purposes the ROMC skill scores were defined from zero (least skillful) to 100 (most skillful). The metric that was chosen for the ROMC skill score evaluates the correlation, R between the model simulations and the reference data set, the standard deviation, σ of the model and reference dataset, and the mean, $\overline{x}$ of both datasets:

$$Skill = 100 \frac{(1+R)^4}{\left(\dfrac{\overline{x}_{mod}}{\overline{x}_{ref}} + \dfrac{\overline{x}_{ref}}{\overline{x}_{mod}}\right)^2 \left(\dfrac{\sigma_{mod}}{\sigma_{ref}} + \dfrac{\sigma_{ref}}{\sigma_{mod}}\right)^2} \quad (1)$$

Since different models may perform different numbers of experiments in debug mode the representability of the skill metric has to be related to the number of test cases on which it is based (see Figure 5).

The ROMC possesses also a series of interactive features allowing users with more than one registered model (maximum is 3), or, users that have submitted simulation results of different versions of the same model to compare these data either against the ROMC reference data or against one of their own ROMC submissions. Similar graphics as in the standard submission to the ROMC are provided both in encapsulated postscript format and as jpeg images (see Figure 6).
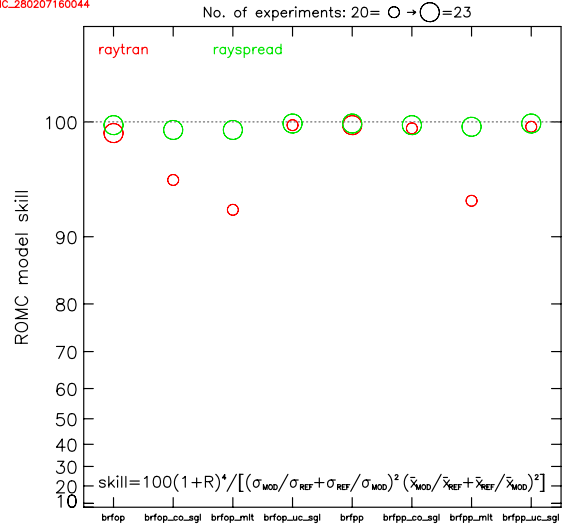
Figure 5. The ROMC skill describing the performance of two models – belonging to the same user – over their respective sets of available simulations for the homogeneous scenes in RAMI.

Users of ROMC are encouraged to utilise only ROMC results that were obtained in validate mode for publication purposes. Those obtained in debug mode, obviously, can not qualify as unambiguous proof of model performance since all simulation results may readily be viewed on the RAMI website. It should also be noted that it is not permissible to modify, change or edit the results provided by the ROMC. This applies to ROMC graphs, statistics, as well as reference data which must all be used 'as provided' and duly referenced using either Widlowski et al., (2007) or this publication.
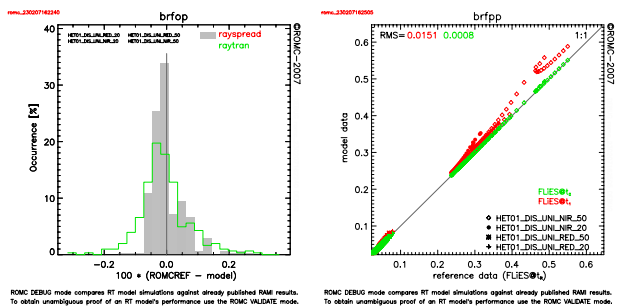

Figure 6. Examples of interactive ROMC graphics in debug mode, showing (left panel) a histogram of model-to-reference-differences for two models belonging to the same ROMC user, and (right panel) a 1:1 plot of BRFs for two different versions of the same model of a ROMC user.

Last but not least a large ensemble of FAQs are available on the ROMC website, that explain the conceptual functioning of the ROMC, the statistics and graphics provided, as well as the submission procedure and guidelines to address eventual troubleshooting. It is hoped that the ROMC will prove useful for the RT modelling community as a whole, not only by providing a convenient means to evaluate RT models outside the triennial phases of RAMI (something that was rather tedious

in the past if authors wished to rely on the experiences gained from RAMI, e.g., Gastellu-Etchegorry et al., (2004)) but also to attract participation in future RAMI activities.

In this context it is important to realize that the ROMC is primarily intended to provide rapid, indicative evaluations of the performance of RT model, in particular in validate mode. RAMI, on the other hand, approaches model evaluation in a more comprehensive manner covering many different aspects of model performance under a large variety of structural, spectral and illumination conditions. The usage of the ROMC should thus be seen as a first step toward the full participation in a next phase of RAMI. Finally it should also be noted, that it is foreseen to regularly update the reference data set of the ROMC after every new phase of RAMI, thus reflecting the improvements that are being made by the RT model community.

4. CONCLUSION

The RAMI On-line Model Checker (ROMC), a web-based tool for the autonomous benchmarking of canopy radiation transfer models was presented (http://romc.jrc.it/). The ROMC was made possible due to substantial improvements in model agreement observed during the third phase of the RAdiation transfer Model Intercomparison (RAMI) exercise. The ROMC utilizes a reference data set established after several years of efforts of the international modeling community to identify a series of "credible" 3-D Monte Carlo models. It is expected that the ROMC will evolve into a yardstick for scientists, funding agencies, and policy makers alike when it comes to appreciating the quality of a given model and ultimately also the information that it helps retrieving from Earth Observation data.

5. ACKNOWLEDGEMENTS

6. REFERENCES

Disney, M. I., P. Lewis, and P. R. J. North (2000). Monte Carlo ray-tracing in optical canopy reflectance modeling. Remote Sensing Reviews, 18, 163-196.

Gastellu-Etchegorry, J-P, E. Martin, and F. Gascon (2004). DART: a 3D model for simulating satellite images and studying surface radiation budget. *International Journal of Remote Sensing*, 25, 73-96.

Goel, N. S. and Thompson, R. L. (2000). A snapshot of Canopy Reflectance Models and a Universal Model for the Radiation Regime. *Remote Sensing Reviews*, 18, 197-206.

Govaerts, Y., and M. M. Verstraete (1998). Raytran: A Monte Carlo ray tracing model to compute light scattering in threedimensional heterogeneous media. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 493.505, 1998.

Lewis, P. (1999). Three-dimensional plant modelling for remote sensing simulation studies using the botanical plant modelling system. *Agronomie - Agriculture and Environment*, 19, 185-210.

North, P. R. J. (1996). Three-dimensional forest light interaction model using a Monte Carlo method. *IEEE Transactions on Geoscience and Remote Sensing*, 34, 946.956, 1996.

Pinty, B., N. Gobron, J.-L. Widlowski , S. A. W. Gerstl, M. M. Verstraete, M. Antunes, C. Bacour, F. Gascon, J.-P. Gastellu, N. Goel, S. Jacquemoud, P. North, W. Qin, and R. Thompson (2001). Radiation Transfer Model Intercomparison (RAMI) Exercise. *Journal of Geophysical Research*, 106, 11,937-11,956.

Pinty, B., J-L. Widlowski, M. Taberner, N. Gobron, M. M. Verstraete, M. Disney, F. Gascon, J.-P. Gastellu, L. Jiang, A. Kuusk, P. Lewis, X. Li, W. Ni-Meister, T. Nilson, P. North, W. Qin, L. Su, S. Tang, R. Thompson, W. Verhoef, H. Wang, J. Wang, G. Yan, and H. Zang (2004). RAdiation transfer Model Intercomparison (RAMI) exercise: Results from the second phase. *Journal of Geophysical Research*, 109, D06210 10.1029/2003JD004252.

Taylor, K. E. (2006). Summarizing multiple aspects of model performance in a single diagram. J*ournal of Geophysical Research*, 106, D7, 7183-7192.

Widlowski, J-L., M. Taberner, B. Pinty, V. Bruniquel-Pinel, M. Disney, R. Fernandes, J.-P. Gastellu-Etchegorry, N. Gobron, A. Kuusk, T. Lavergne, S. Leblanc, P. Lewis, E. Martin, M. Mottus, P. J. R. North, W. Qin, M.Robustelli, N. Rochdi, R.Ruiloba, C.Soler, R.Thompson, W. Verhoef, M. M.Verstraete, and D. Xie (2007). The third RAdiation transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance modelling. *Journal of Geophysical Research*, 2007, accepted.

Widlowski, J.-L., T. Lavergne, B. Pinty, M. M. Verstraete, and N. Gobron (2007). Rayspread: A virtual laboratory for rapid BRF simulations over 3-D plant canopies, in *Computational Methods in Transport*, edited by Graziani, F., pp. 211.231, ISBN.10 3.540.28,122.3, Lecture Notes in Computational Science and Engineering Series, 48, Springer Verlag, Berlin.