# Towards Operational Products Development from Earth Observation: Exploration of SimSphere Land Surface Process Model Sensitivity using a GSA Approach

George P. Petropoulos[1], Hywel Griffiths[1], Stefano Tarantola[2]

(1) Institute of Geography and Earth Sciences, University of Aberystwyth, Wales, United Kingdom

(2) Institute for the Protection and Security of the Citizen, Joint Research Centre, European Commission, Ispra (VA) - Italy

## 1. INTRODUCTION

The use of **simulation process models** has played a key role in extending our abilities to study Earth system processes and enhancing our understanding on how different components of it interplay. The use of such models **combined with Earth Observation (EO) data** provides a promising direction towards deriving more accurately spatio-temporal estimates of key parameters characterising land surface interactions, as it combines the horizontal coverage and spectral resolution of EO data with the vertical coverage and fine temporal continuity of models.

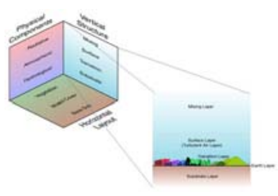
**This study aimed to perform a Global Sensitivity Analysis (GSA) on the SimSphere Surface Vegetation–Atmospheric Transfer (SVAT) model** to further extend our understanding of the model structure and to establish its coherence. For consistency and comparability to previous studies, the GSA implemented herein has also been based on a cutting edge, sophisticated, yet simple to apply method based on Bayesian Analysis of Computer Code Outputs (BACCO; Kennedy and O'Hagan, 2001). Whereas previous SA studies on SimSphere using BACCO assumed normal probability distribution functions (PDFs) for the model input parameters, in our study we assume uniform PDFs. It also uses PDFs of the most sensitive model inputs derived directly from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) imagine radiometer.

## 2. BACKGROUND / METHODOLOGY

### 1. SIMSPHERE :

**SimSphere** was originally developed by Carlson & Boland (1978) and significantly improved by Gillies et al. (1995) & Petropoulos et al. (2013a).

It is a deterministic mathematical model that provides representations of the physical mechanisms controlling land surface interactions in a vertical profile. It implicitly refers to a horizontal area that can be composed of a mixture of bare soil & vegetation (*Fig. 1*).



It simulates various parameters over a 24-hour cycle and at a chosen time step, starting from a set of initial conditions in the early morning. The model inputs requirements and the number of outputs produced related to characterising land surface processes are shown in *Fig. 2*.

**Fig. 2:** *List of SimSphere inputs (left) and variables predicted (right)*

**SimSphere verification and applications:**

- As **a stand alone tool**, in studies of:
  - **Carbon [ ] in the atmosphere & plant responses** *(e.g. Gottschalck et al., 2001)*
  - **Plant transpiration, plant stress, stomatal resistance & GPP** under different scenarios *(e.g. Olioso et al., 1996)*
  - **Sensitivity testing** of both the vegetation/soil components *(e.g. Petropoulos et al., 2009; 2010)*
  - **Comparative studies** *(e.g. Petropoulos et al., 2010b)*
- Studies **linked with RS data** *(e.g. Gillies et al., 1997; Petropoulos and Carlson, 2011)*
- In 3 **Ph.D theses** *(Gillies, 1990; Brunsell, 2003; Petropoulos, 2008)*
- Is currently used as an **educational tool** in 5 Universities in the world.
- Distributed globally from Aberystwyth University (http://www.aber.ac.uk/simsphere)

### 2. BACCO GEM SA:

❑ GSA is achieved from an *emulator*, derived from a relatively small number of model runs covering a multidimensional input space, which deducts all the SA measures related to the original deterministic model code without the need to execute further runs of the original code.

❑ Self-measure of *emulator* performance in matching the original model code is embedded, providing an accurate and reliable indication of the trustworthiness of its analysis.

$$V(Y) = \sum_{i=1}^{s} D_i + \sum_{i \lhd j} D_{ij} + ... + D_{1...s}$$

$$D_{ij} = V(E(Y|X_i, X_j)) - V(E(Y|X_i)) - V(E(Y|X_j))$$

$$S_i = \frac{D_i}{V(Y)} \quad S_{ij} = \frac{D_{ij}}{V(Y)} \quad ST_i = \frac{D_i + D_{i,\sim i}}{V(Y)}$$

$$\sum_i S_i + \sum_{i \lhd j} S_{ij} + \sum_{i \lhd j \lhd m} S_{ijm} + ... + S_{12...k} = 1$$

***where:***

- $V(Y)$ is the total variance of the output variable Y
- $D_i$ is the importance measure for input $X_i$,
- $D_{ij}$ is the importance measure for the interaction between inputs $X_i$ and $X_j$
- $D_{1...s}$ denotes analogous formulae for the higher order terms
- *s* denotes the number of inputs (so-called 'factors')
- $E(Y|X_i)$ stands for the expectation of Y conditional on fixed value of *Xi* and the variance is taken over all factors which are fixed in the conditional expectations.
- $S_i$ is a sensitivity index
- $ST_i$ is the total sensitivity index
- $D_{i,\sim i}$ indicates all interactions involving the factor $X_i$.

The basic SA output from GEM SA includes the computation of the ***main and joint effects*** (pairwise interactions only) of the input parameters, as well as of the ***total effects***. In addition, GEM SA provides a set of ***main effects plots***, which are calculated for each of the model inputs.

### 3. BACCO GEM SA on SimSphere:

The sensitivity of the following parameters was examined due to their importance in surface energy balance studies:

- Daily Average Net Radiation ($\overline{Rn_{daily}}$),
- Daily Average Latent Heat flux ($\overline{LE_{daily}}$)
- Daily Average Sensible Heat flux ($\overline{H_{daily}}$)
- Daily Average Air Temperature ($\overline{Tair_{daily}}$)
- Daily Average Surface Moisture Availability ($\overline{Mo_{daily}}$)

A design space of *400* simulations was created. All model inputs were allowed to vary except the geographical location and atmospheric profile. For these, *a priori* values were used from *Borgo Cioffi* CarboEurope site in central Italy (40 31' 25.5'' N, 14 57' 26.8'' E) on 17 Nov 2004. All inputs were initially assumed to be uniform, then for the most sensitive model inputs (i.e. slope, aspect, Fr) PDFs from ASTER imagery were taken.

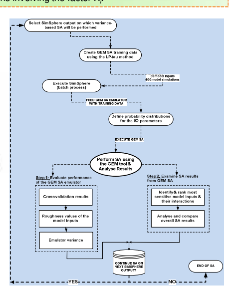



**Fig. 3:** *Overview of the overall SA methodology followed.*

## 3. RESULTS: *Emulator Accuracy*

**1. Emulator performance :**

| | Model Input | Roughness $\overline{Rn_{daily}}$ | $\overline{LE_{daily}}$ | $\overline{H_{daily}}$ | $\overline{Tair_{daily}}$ | $\overline{Mo_{daily}}$ |
|---|---|---|---|---|---|---|
| X1 | *Slope* | 3.602 | 0.038 | 0.038 | 2.664 | 0.059 |
| X2 | *Aspect* | 18.501 | 0.467 | 0.467 | 7.928 | 0.268 |
| X3 | *Station Height* | 0.004 | 0.843 | 0.843 | 0.000 | 0.002 |
| X4 | *Fractional Vegetation Cover* | 0.507 | 2.006 | 2.006 | 1.675 | 0.006 |
| X5 | *LAI* | 0.472 | 0.096 | 0.096 | 0.190 | 0.000 |
| X6 | *Foliage emissivity* | 0.010 | 0.168 | 0.168 | 0.018 | 0.000 |
| X7 | *[Ca]* | 0.024 | 0.148 | 0.148 | 0.022 | 0.004 |
| X8 | *[Ci]* | 0.000 | 0.016 | 0.016 | 0.021 | 0.005 |
| X9 | *[03] in the air* | 0.000 | 0.000 | 0.000 | 0.000 | 0.105 |
| X10 | *Vegetation height* | 0.036 | 0.630 | 0.630 | 1.103 | 0.000 |
| X11 | *Leaf width* | 0.011 | 0.024 | 0.024 | 0.000 | 0.064 |
| X12 | *Minimum Stomatal Resistance* | 0.028 | 0.058 | 0.058 | 0.000 | 0.000 |
| X13 | *Cuticle Resistance* | 0.017 | 0.178 | 0.178 | 0.074 | 0.000 |
| X14 | *Critical leaf water potential* | 0.018 | 0.026 | 0.026 | 0.493 | 0.000 |
| X15 | *Critical solar parameter* | 0.011 | 0.049 | 0.049 | 0.014 | 0.067 |
| X16 | *Stem resistance* | 0.001 | 0.019 | 0.019 | 0.002 | 0.047 |
| X17 | *Surface Moisture Availability* | 0.035 | 3.535 | 3.534 | 0.244 | 8.258 |
| X18 | *Root Zone Moisture Availability* | 0.000 | 0.101 | 0.101 | 0.000 | 0.148 |
| X19 | *Substr. Max. Volum. Water Cont.* | 0.000 | 0.040 | 0.040 | 0.000 | 1.960 |
| X20 | *Substrate clim. mean temper.* | 0.003 | 0.006 | 0.006 | 0.028 | 0.000 |
| X21 | *Thermal inertia* | 0.111 | 0.017 | 0.017 | 0.225 | 0.012 |
| X22 | *Ground emissivity* | 0.005 | 0.027 | 0.027 | 0.000 | 0.000 |
| X23 | *Atmospheric Precipitable water* | 0.007 | 0.003 | 0.003 | 0.069 | 0.009 |
| X24 | *Surface roughness* | 0.105 | 0.774 | 0.774 | 0.396 | 0.000 |
| X25 | *Obstacle height* | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 |
| X26 | *Fractional Cloud Cover* | 0.001 | 0.146 | 0.145 | 0.004 | 0.000 |
| X27 | *RKS* | 0.000 | 0.030 | 0.031 | 0.000 | 0.944 |
| X28 | *CosbyB* | 0.009 | 0.255 | 0.255 | 0.110 | 0.614 |
| X29 | *THM* | 1.671 | 0.366 | 0.366 | 0.033 | 0.206 |
| X30 | *PSI* | 0.124 | 0.000 | 0.000 | 0.000 | 3.000 |

| | $\overline{Rn_{daily}}$ | $\overline{LE_{daily}}$ | $\overline{H_{daily}}$ | $\overline{Tair_{daily}}$ | $\overline{Mo_{daily}}$ |
|---|---|---|---|---|---|
| **FITTED MODEL PARAMETERS** | | | | | |
| sigma-squared: | 0.3691 | 1.042 | 1.042 | 1.314 | 0.244 |
| **CROSS-VALIDATION RESULTS:** | | | | | |
| Cross- validation root mean squared-error ($Wm^{-2}$): | 27.850 | 0.077 | 0.077 | 0.911 | 0.086 |
| Cross-validation root mean squared relative error (%): | 8.930 | 20.321 | 18.725 | 6.688 | 24.705 |
| Cross-validation root mean squared standardised error: | 1.389 | 1.287 | 1.287 | 1.489 | 3.073 |

**Table 1:** *Summary of statistics concerning the emulator performance when uniform PDF for inputs was assumed. Shading highlights the roughness values of the model inputs with values greater than 1.0.*

| | Model Input | Roughness $\overline{Rn_{daily}}$ | $\overline{LE_{daily}}$ | $\overline{H_{daily}}$ | $\overline{Tair_{daily}}$ | $\overline{Mo_{daily}}$ |
|---|---|---|---|---|---|---|
| X1 | *Slope* | 2.6660 | 0.2491 | 1.0044 | 1.1963 | 0.4268 |
| X2 | *Aspect* | 16.9225 | 8.1534 | 4.9068 | 6.5718 | 0.0786 |
| X3 | *Station Height* | 0.0078 | 0.9689 | 0.2072 | 0.0010 | 0.3230 |
| X4 | *Fractional Vegetation Cover* | 0.6988 | 1.9564 | 1.0975 | 1.8938 | 0.0116 |
| X5 | *LAI* | 0.9337 | 0.0111 | 0.2063 | 0.0828 | 0.0010 |
| X6 | *Foliage emissivity* | 0.0502 | 0.0078 | 0.1355 | 0.0301 | 0.0026 |
| X7 | *[Ca]* | 0.0083 | 0.0710 | 0.0048 | 0.1500 | 0.0366 |
| X8 | *[Ci]* | 0.0722 | 0.0843 | 0.0035 | 0.0172 | 0.0081 |
| X9 | *[03] in the air* | 0.1042 | 0.1284 | 0.0338 | 0.0010 | 0.0001 |
| X10 | *Vegetation height* | 0.1933 | 0.0268 | 0.7700 | 0.7651 | 0.0001 |
| X11 | *Leaf width* | 0.0001 | 0.0555 | 0.0007 | 0.0031 | 0.0001 |
| X12 | *Minimum Stomatal Resistance* | 0.0001 | 0.0102 | 0.0004 | 0.0127 | 0.0074 |
| X13 | *Cuticle Resistance* | 0.0452 | 0.0604 | 0.1886 | 0.0569 | 0.0072 |
| X14 | *Critical leaf water potential* | 0.0289 | 0.0023 | 0.0388 | 0.2143 | 0.0083 |
| X15 | *Critical solar parameter* | 0.0001 | 0.0001 | 0.0062 | 0.0001 | 0.0036 |
| X16 | *Stem resistance* | 0.0002 | 0.1568 | 0.0022 | 0.0198 | 0.1686 |
| X17 | *Surface Moisture Availability* | 0.4368 | 1.4709 | 0.4533 | 0.1154 | 3.1706 |
| X18 | *Root Zone Moisture Availability* | 0.0293 | 0.0387 | 0.0317 | 0.0759 | 0.0543 |
| X19 | *Substr. Max. Volum. Water Cont.* | 0.0001 | 0.0118 | 0.0001 | 0.0923 | 0.0129 |
| X20 | *Substrate clim. mean temper.* | 0.0333 | 0.3054 | 0.2938 | 0.1741 | 0.0011 |
| X21 | *Thermal inertia* | 0.2169 | 0.2840 | 0.2754 | 0.2514 | 0.0001 |
| X22 | *Ground emissivity* | 0.0001 | 0.0001 | 0.0682 | 0.0722 | 0.0333 |
| X23 | *Atmospheric Precipitable water* | 0.0218 | 0.0106 | 0.0140 | 0.0066 | 0.0001 |
| X24 | *Surface roughness* | 0.2668 | 0.0077 | 0.8880 | 0.7983 | 0.0001 |
| X25 | *Obstacle height* | 0.0001 | 0.0001 | 0.0064 | 0.0001 | 0.0001 |
| X26 | *Fractional Cloud Cover* | 0.0224 | 0.0037 | 0.0001 | 0.0001 | 0.0031 |
| X27 | *RKS* | 0.0069 | 0.0003 | 0.0001 | 0.0001 | 3.7857 |
| X28 | *CosbyB* | 0.0766 | 0.0049 | 0.0156 | 0.0001 | 1.1833 |
| X29 | *THM* | 0.0301 | 0.0001 | 0.0005 | 0.0100 | 0.5326 |
| X30 | *PSI* | 0.0517 | 0.0968 | 0.0213 | 0.0001 | 2.9997 |

| | $\overline{Rn_{daily}}$ | $\overline{LE_{daily}}$ | $\overline{H_{daily}}$ | $\overline{Tair_{daily}}$ | $\overline{Mo_{daily}}$ |
|---|---|---|---|---|---|
| **FITTED MODEL PARAMETERS** | | | | | |
| sigma-squared: | 0.5229 | 0.8752 | 1.4494 | 1.6321 | 0.1272 |
| **CROSS-VALDIATION RESULTS:** | | | | | |
| Cross- validation root mean squared-error ($Wm^{-2}$): | 23.0515 | 20.2858 | 18.1837 | 0.8349 | 0.0452 |
| Cross-validation root mean squared relative error (%): | 25.2553 | 24.4602 | 36.0452 | 6.4749 | 19.9241 |
| Cross-validation root mean squared standardised error: | 1.3367 | 1.1569 | 1.4070 | 1.5411 | 1.5587 |

**Table 2:** *Summary of statistics concerning the emulator performance assuming normal PDF for inputs. Shading highlights the roughness values of the model inputs with values greater than 1.0.*

## 4. RESULTS: *SA assuming uniform PDFs*

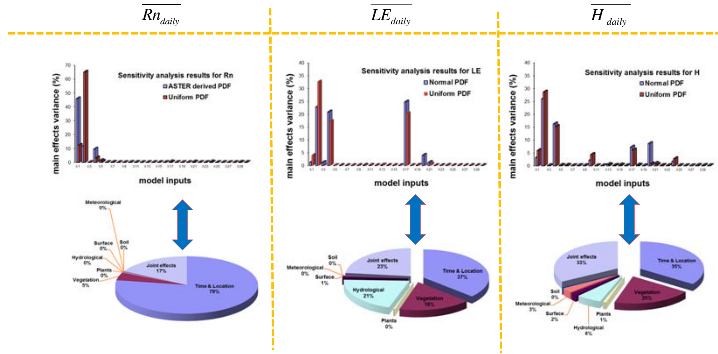



**Fig. 4:** *Sensitivity analysis results using GEM SA software comparing uniform PDF and normal PDFs (upper) and the importance of various variables for uniform PDF (lower) here shown for daily Rn, daily LE and daily H.*

## 5. RESULTS: *SA using PDFs from ASTER satellite*

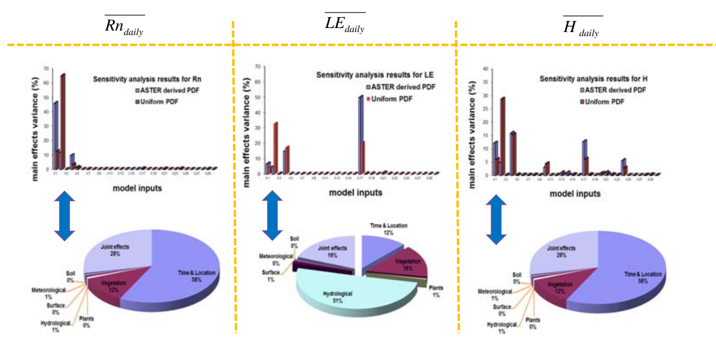



**Fig. 5:** *Sensitivity analysis results using GEM SA comparing uniform PDFs (derived in Obj.1) and ASTER-derived PDF (upper) and the SA inputs for examples of variables of which SA was performed for ASTER-derived data (lower) for daily Rn, daily LE & daily H.*

## 4. CONCLUSIONS

❑ Our study showed comparable results to previous studies in terms of identifying the most sensitive model inputs. Yet, the PDFs assumption can influence the absolute SA measures of the model input parameters in respect to the target quantity considered each time. Some of the most sensitive model inputs for all outputs were parameters relatively easily estimated from EO data, which has important implications for the integration of the model with such data and also its use in general in future research.

❑ This work is significant to the community of model users and is also very timely given the efforts currently examining its use in an EO-based method for deriving operationally regional estimates of energy fluxes and soil moisture from EO data (e.g *Chauhan et al., 2003*).

## References

**Carlson, T.N.; Boland, F.E.** Analysis of urban-rural canopy using a surface heat flux/temperature model. Journal of Applied Meteorology, 1978, 17, pp: 998-1013. **Chauhan, N. S., Miller, S. & Ardanuy, P.,** Spaceborne soil moisture estimation at high resolution: a microwave-optical/IR synergistic approach. International Journal of Remote Sensing, 22, 4599-46, 2003. **Gottschalck, J. C.; Gillies, R. R.; Carlson, T. N.** The simulation of canopy transpiration under doubled CO2: The evidence and impact of feedbacks on transpiration in two 1-D Soil-Vegetation-Atmosphere-Tranfer models. Agriculture and Forest Meteorology, 2001, 106, pp: 1-21. **Kennedy,M.C., O'Hagan, A.,** 2001. Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (3), 425–464. **Olioso, A.; Carlson, T. N.; Brisson, N.** Simulation of diurnal transpiration and photosynthesis of a water stressed soybean crop. Agriculture and Forest Meteorology, 1996, 81, 41-59. **Petropoulos G.P. and T.N., Carlson,** 2011. Retrievals of turbulent heat fluxes and soil moisture content by Remote Sensing, pages 469-501, in "Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications", Ed. Taylor and Francis, pp. 556, ISBN: 978-1-4200-9175-5. **Petropoulos, G., M. Ratto and S. Tarantola** (2010a): A comparative analysis of emulators for the sensitivity analysis of a land surface process model. 6th International Conference on Sensitivity Analysis of Model Output, July 19-22nd, 2010, Milan, Italy, on Procedia - Social and Behavioral Sciences, vol. 2 (6), 7716-17. **Petropoulos, G., Wooster, M. J., M. Kennedy, Carlson, T.N., and M. Scholze**, 2009. A global sensitivity analysis study of the 1d SimSphere SVAT model using the GEM SA software. Ecological Modelling, 220 (19), 2427-2440. **Petropoulos, G., Wooster, M. J., T.N. Carlson, and N. Drake,** 2010b. Verification Of The SimSphere SVAT Model Performance In Simulating Key Land Surface Parameters At Selected CarboEurope IP Sites. European Geosciences Union 2010, May 2-7th, 2010, Vienna, Austria. **Petropoulos, G.P., I. Konstas and T.N. Carlson,** 2013a. Automation of SimSphere Land Surface Model Use as a Standalone Application and Integration with EO Data for Deriving Key Land Surface Parameters. European Geosciences Union, April 7-12th, 2013, Vienna, Austria.

## Acknowledgments

Preparation of this work was partially supported by the European Space Agency (ESA) Support to Science Element (STSE) under contract STSE-TEBM-EOPG-TN-08-0005 as well as the Marie Curie Career Re-Integration Grant TANSFORM-EO project. Dr. Petropoulos gratefully acknowledges the support provided. **Contact: gep9@aber.ac.uk**