Environmental Modelling & Software 65 (2015) 94–107

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Addressing the ability of a land biosphere model to predict key biophysical vegetation characterisation parameters with Global Sensitivity Analysis

Gareth Ireland [a], George P. Petropoulos [a,*], Toby N. Carlson [b], Sarah Purdy [c]

[a] Department of Geography and Earth Sciences, University of Aberystwyth, SY23 2EJ, UK
[b] Department of Meteorology, Pennsylvania State University, University Park, PA 16802, USA
[c] Institute of Biological Environmental and Rural Sciences, Aberystwyth University, SY23 3EE, UK

ARTICLE INFO

Article history:
Received 26 August 2014
Received in revised form
13 November 2014
Accepted 14 November 2014
Available online 2 January 2015

Keywords:
Sensitivity analysis
CO_2 flux
Ambient CO_2
O_3 flux
SimSphere
Gaussian process emulators
BACCO GEM-SA

ABSTRACT

Sensitivity Analysis (SA) of the SimSphere Soil Vegetation Atmosphere Transfer (SVAT) model has been performed in this study using a cutting edge and robust Global Sensitivity Analysis (GSA) approach, based on the use of the Gaussian Emulation Machine for Sensitivity Analysis (GEM-SA) tool. The sensitivity of the following model outputs was evaluated: the ambient CO_2 concentration, the rate of CO_2 uptake by the plant, the ambient O_3 concentration, the flux of O_3 from the air to the plant/soil boundary and the flux of O_3 taken up by the plant alone. The most sensitive model inputs for the majority of outputs were: The Leaf Area Index (LAI), Fractional Vegetation Cover (Fr), Cuticle Resistance (CR) and Vegetation Height (VH). The influence of the external CO_2 on the leaf and O_3 concentration in the air as input parameters was also significant. Our study provides an important step forward in the global efforts towards SimSphere verification given the increasing interest in its use as an independent modelling or educational tool. Results of this study are also timely given the ongoing global efforts focused on deriving, at an operational level, spatio-temporal estimates of energy fluxes and soil moisture content using SimSphere synergistically with Earth Observation (EO) data.

1. Introduction

The complex interplay between the different facets of land–surface interactions play a fundamental role in the spatio-temporal variations of carbon dioxide (CO_2) and ozone (O_3) fluxes within the Earth system. Within the atmospheric boundary layer, the exchange of CO_2 at the surface is primarily the result of complex vegetation processes (Boussetta et al., 2013). CO_2 is assimilated in vegetation by photosynthesis (expressed as gross primary productivity) and is returned to the atmosphere by a variety of above- and below-ground metabolic processes (Bloemen et al., 2013). Stomatal behaviour provides the main short-term control of both transpiration and CO_2 assimilation. Crops grown under CO_2 enrichment usually exhibit increased mass, which is attributed to greater photosynthetic capacity and enhanced water use efficiency (Olvera et al., 2013). Carbon assimilation and water dynamics are inherently linked to crop yield and so an understanding of these relationships is fundamental to our ability to understand, or predict, plant productivity. The Intergovernmental Panel on Climate Change (IPCC) predicts concentrations of atmospheric CO_2 will continue rising from their current concentrations of ~395 ppmv–>420 ppmv by 2050 (IPCC, 2001). Therefore, understanding the interactions between plants and environment is a central requirement when forecasting the effects of future climate change and variability (Williams et al., 2012).

Tropospheric ozone is a phytotoxic air pollutant responsible for crop and forest damage worldwide (Panek, 2004). The impact of O_3 exposure usually manifests as necrotic lesions, decreased photosynthesis and accelerated senescence (Pell et al., 1997; Wiese and Pell, 1997). The exact mechanism by which O_3 stress is imposed is unclear but probably occurs through the formation of reactive oxygen species such as superoxide and hydrogen peroxide (Pell et al., 1997). Furthermore, stomatal control may be reduced following exposure to O_3 which causes greater susceptibility to drought stress (McLaughlin et al., 2007). This has major implications for tree and crop performance in the future climate as temperatures are predicted to increase (thus increasing transpiration) along with a greater area affected by more frequent drought (IPCC, 2014). The

* Corresponding author. Tel.: +44 0 1970 621861.
E-mail address: george.petropoulos@aber.ac.uk (G.P. Petropoulos).

http://dx.doi.org/10.1016/j.envsoft.2014.11.010
1364-8152/

effect of O_3 stress on stomatal regulation and implications for drought susceptibility has been found to offset much of the projected benefit of rising CO_2 for plant growth in some biogeochemical models (Ollinger et al., 2002). The direct reduction in forest productivity currently caused by O_3 stress is estimated to be 1–10% for forests in Europe and North America (Broadmeadow, 1998). Thus, O_3 is clearly an important economic factor in biomass production and there is a need to understand its influence on plant processes (Ainsworth et al., 2012; Konovalov et al., 2012). To this end, an improved quantification of the effect of land surface-atmosphere interactions on the spatio-temporal distribution of CO_2 and O_3 fluxes will enable the development of coherent plans to manage ecosystems for future climate mitigation and agricultural production (Pitman et al., 2012; Williams et al., 2012).

In this context, a representative description of land surface–atmosphere interactions requires mathematical models able to accurately describe interdependent physical and biological processes in vegetation and soil, as well as physical processes within the atmospheric boundary layer (Marras et al., 2011). Several modelling approaches have been developed to represent the terrestrial carbon cycle depending on the main goals of the modelling effort. Aggregated 'big leaf' models for instance, act to simulate selected mass, water and energy transfers from a representative leaf surface which is scaled up to the whole canopy, either based on simple linear scaling or on a non-linear scaling by partitioning between Sunlit and shaded leaves (Ganzeveld et al., 2012). Generally, these models are widely regarded as the simplest group of models, but behold vast application value such as data gap filling and tracing gas fluxes (Baldocchi, 2010). Soil Vegetation Atmosphere Transfer (SVAT) models are more complex 'distributed multi-layer models' which differ in their approach to estimate surface exchanges. These embedded modelling efforts are numerical representations of the multifarious interactions of energy and mass transfers through the soil/vegetation/atmospheric 1-dimensional vertical column (Marras et al., 2011). They require an application context constrained by input variables (atmospheric forcing and vegetation variables) and input parameters (soil and vegetation properties, initialisation) to simulate the water and energy budget at the surface. The number of parameters is generally related to the complexity of the model and their calibration requires the development of optimisation methodologies (Coudert and Ottlé, 2007).

A SVAT model, termed SimSphere, was originally developed by Carlson and Boland (1978) within the Department of Meteorology of Pennsylvania State University, USA, and has continued to be developed over a period of more than two decades. Notably, during this period the model has undergone significant modifications by a number of contributors, most recently by Petropoulos et al. (2013a). Briefly, it is a one-dimensional boundary layer model with a plant component implicitly referring to a horizontal area of undefined size that can be composed of a mixture of bare soil and vegetation. In addition to its use as a stand-alone modelling tool, SimSphere is also integrated synergistically with Earth Observation (EO) data via a method termed the "triangle" method (Carlson, 2007; Petropoulos and Carlson, 2011). This method interprets the relationship between a Vegetation Index (VI) and surface radiative temperature (Ts) derived from a satellite-derived scatter plot, linked with SimSphere to deduce evaporative fraction (EF) over large areas (Long and Singh, 2013). Variants of this method are at present being considered for the development of operational products from EO data, some anticipated to be delivered on a global scale (Chauhan et al., 2003; ESA STSE, 2012). However, being a mathematical representation of natural processes, such modelling approaches require a considerable number of assumptions on model structure, model parameter values and model input variables. These input parameters can lead to output uncertainty and inaccuracy (Cosenza et al., 2014; Vanuytrecht et al., 2014). Thus, Sensitivity Analysis (SA) is an essential and well-established tool that has been used in evaluating robustness of model based results (Feyissa et al., 2012; Ratto et al., 2012). In particular, SA quantifies the influence of each uncertain factor (parameter or driving variable) on the model's output variability (Gan et al., 2014). It can help to determine the relationship between independent and dependent variables to get a better understanding of the model performance. Reasons for performing SA are diverse; it allows for Factor Fixing (FF), where factors that are non-influential can be set to a fixed value anywhere in their uncertainty range and it would not affect model output variance. Factor Prioritisation (FP) on the other hand, is where the modeller focuses on the parameters that have the potential to maximally reduce model output variance if determined. For the case of FP, SA allows for better estimation of the actual factor value and distribution (Nossent et al., 2011; Gamerith et al., 2013).

SA methods are generally classified as either Local Sensitivity Analysis (LSA) or Global Sensitivity Analysis (GSA). In LSA methods, each factor is perturbed in turn from randomly generated reference parameter sets, whilst holding all others to their central value and computing the difference in the outputs (Wainwright et al., 2013; Baroni and Tarantola, 2014). Although a computationally frugal method, LSA can be criticised for being inadequate for analysing complex biophysical process models which may have many parameters, and may be high-dimensional and/or non-linear (Song et al., 2012; Wainwright et al., 2013). Compared with LSA, GSA provides quantitative importance measures that relate the variance of the output with each input dependent variable on different sources of variation over the entire parameter space (Wei et al., 2013). Furthermore, GSA approaches are not limited by model complexity and provide robust sensitivity measures in the presence of non-linearity and interactions amongst parameters. However, the model complexity and high number of parameters can be computationally intensive and inefficient (Gatelli et al., 2009; Wainwright et al., 2013; Gan et al., 2014). Given the intricacy of the physical interconnections involved in modelling land surface–atmosphere interactions, GSA has become popular in the environmental modelling field in recent years. The complexity of such models and their ability to incorporate parameter interactions can be a significant advantage when deriving simulation outputs that are fully analogous of the real world system in terms of accuracy, generality and realism (Anderson et al., 2008). A number of studies have thus performed advanced GSA on SimSphere based on a Gaussian process emulator (Petropoulos et al., 2009b, 2010, 2013b–d). These allowed, for the first time, an insight into the model architecture and the mapping of the sensitivity between the model inputs and outputs. However, SA studies on SimSphere have been limited to only a small number of output parameters, and the effect of different simulation times on model sensitivity has yet to be explored.

In this context, the objective of this study is two-fold: i) To perform a GSA to explore the sensitivity of target quantities simulated by SimSphere for the model inputs/outputs, which have not been previously investigated. These parameters are namely; the ambient CO_2 concentration [ppmv], the rate of uptake of CO_2 by the plant [μmol m^{-2} s^{-1}], the ambient O_3 concentration [ppmv $\times 10^{-3}$], the flux of O_3 from the air to the plant/soil boundary [mg m^{-2} s^{-1}], the flux of O_3 taken up by the plant alone [mg m^{-2} s^{-1}]. ii) To extend the GSA on SimSphere and to explore the sensitivity of the same target quantities at different simulation times (9:30/11:30/13:30), which coincide closely to overpass times of different satellite sensors. These objectives will allow us to foster an understanding of the model structure and further establish its coherence.