Geosci. Model Dev., 8, 3257–3284, 2015
www.geosci-model-dev.net/8/3257/2015/
doi:10.5194/gmd-8-3257-2015

Geoscientific Model Development
Open Access

Quantifying the prediction accuracy of a 1-D SVAT model at a range of ecosystems in the USA and Australia: evidence towards its use as a tool to study Earth's system interactions

G. P. Petropoulos[1], **M. R. North**[1], **G. Ireland**[1], **P. K. Srivastava**[2,3,4], **and D. V. Rendall**[1]

[1]Department of Geography and Earth Sciences, Aberystwyth University, Aberystwyth, SY23 3DB, UK
[2]Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[3]Earth System Science Interdisciplinary Center, University of Maryland, Maryland, USA
[4]Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, India

Correspondence to: G. P. Petropoulos (george.petropoulos@aber.ac.uk)

Received: 9 January 2015 – Published in Geosci. Model Dev. Discuss.: 5 March 2015
Revised: 25 June 2015 – Accepted: 28 July 2015 – Published: 16 October 2015

Abstract. This paper describes the validation of the SimSphere SVAT (Soil–Vegetation–Atmosphere Transfer) model conducted at a range of US and Australian ecosystem types. Specific focus was given to examining the models' ability in predicting shortwave incoming solar radiation (R_g), net radiation (R_{net}), latent heat (LE), sensible heat (H), air temperature at 1.3 m ($T_{air\ 1.3\ m}$) and air temperature at 50 m ($T_{air\ 50\ m}$). Model predictions were compared against corresponding in situ measurements acquired for a total of 72 selected days of the year 2011 obtained from eight sites belonging to the AmeriFlux (USA) and OzFlux (Australia) monitoring networks. Selected sites were representative of a variety of environmental, biome and climatic conditions, to allow for the inclusion of contrasting conditions in the model evaluation.

Overall, results showed a good agreement between the model predictions and the in situ measurements, particularly so for the R_g, R_{net}, T_{air} 1.3 m and T_{air} 50 m parameters. The simulated R_g parameter exhibited a root mean square deviation (RMSD) within 25 % of the observed fluxes for 58 of the 72 selected days, whereas an RMSD within ∼ 24 % of the observed fluxes was reported for the R_{net} parameter for all days of study (RMSD = 58.69 W m^{-2}). A systematic underestimation of R_g and R_{net} (mean bias error (MBE) = −19.48 and −16.46 W m^{-2}) was also found. Simulations for the T_{air} 1.3 m and T_{air} 50 m showed good agreement with the in situ observations, exhibiting RMSDs of 3.23 and 3.77 °C (within ∼ 15 and ∼ 18 % of the observed) for all days of analysis, respectively. Comparable, yet slightly less satisfactory simulation accuracies were exhibited for the H and LE parameters (RMSDs = 38.47 and 55.06 W m^{-2}, ∼ 34 and ∼ 28 % of the observed). Highest simulation accuracies were obtained for the open woodland savannah and mulga woodland sites for most of the compared parameters. The Nash–Sutcliffe efficiency index for all parameters ranges from 0.720 to 0.998, suggesting a very good model representation of the observations.

To our knowledge, this study presents the most detailed evaluation of SimSphere done so far, and the first validation of it conducted in Australian ecosystem types. Findings are important and timely, given the expanding use of the model both as an educational and research tool today. This includes ongoing research by different space agencies examining its synergistic use with Earth observation data towards the development of global operational products.

1 Introduction

The importance of studying land surface–atmosphere interactions to develop a better understanding of Earth's physical processes and feedbacks is evident from several investigations. Today, particularly so in the face of climate change, it has been recognised by the global scientific community as a topic requiring further attention and investigation (Battrick and Herland, 2006; Petropoulos et al., 2014). This is docu-

Published by Copernicus Publications on behalf of the European Geosciences Union.

mented by the fact that it is of crucial importance to help address directives such as the European Parliament Directive 2000/60/EC aimed at establishing a framework for community action in the field of water policy, namely the EU Water Framework Directive. On this basis, the need to develop a holistic understanding of how land surface parameters characterising the planet's energy and water budget in different ecosystems has never been more important (ESA, 2012).

Land surface parameterisation schemes (LSPs, also known as land surface models (LSMs)) are one of the preferred scientific tools to quantify at fine spatial and temporal resolutions Earth system interactions. LSPs simulate a number of parameters characterising land surface interactions within the lower atmospheric boundary from a predefined set of surface characteristics (i.e. properties of soil, vegetation and water). Often LSPs are utilised, amongst other schemes, to assess water resources, to evaluate the hydrological impacts of changes in climate and land use, and to model land–atmosphere exchanges and emissions of aerosols (Prentice et al., 2015). Recent developments in mathematical modelling have been driven primarily by the progress in computer technology, the expansion of modelling into new fields and disciplines and the need for increased accuracy in model predictions (Bellocchi et al., 2010). As a result, LSPs have advanced considerably to include detailed parameterisations of momentum, energy, mass and biogeochemistry (Rosolem et al., 2013).

One group of LSPs include the Soil–Vegetation–Atmosphere Transfer (SVAT) models. These are mathematical representations of vertical "views" of the physical mechanisms controlling energy and mass transfers in the soil–vegetation–atmosphere continuum. These deterministic models are able to provide estimates of the time course of soil and vegetation state variables at time steps compatible with the dynamics of atmospheric processes. During the last number of decades SVAT models have evolved from simple energy balance parameterisations, e.g. from the bucket schemes adopted by Manabe (1969), through the schemes of Deardorff (1978), to the biosphere–atmosphere transfer scheme (BATS) of Dickinson and Henderson-Sellers (1986) and the simple biosphere (SiB) model of Sellers et al. (1986). At present, SVATs are able to describe the multifarious transfer processes through varying degrees of complexity, including the energy, water and carbon dioxide (CO_2) fluxes between the ground surface covered by different vegetation types and the atmosphere over different temporal and spatial scales (Olchev et al., 2008). These require an application context constrained by input variables (atmospheric forcing and vegetation) and input parameters (soil and vegetation properties, initialisation) to simulate the water and energy budget at the surface (Coudert et al., 2008; Ridler et al., 2012).

However, before applying a computer simulation model to perform any kind of analysis or operation, a variety of validatory tests need to be executed. The process of validating a mathematical model's performance, coherence and representation of the natural environment is regarded as an essential step in its development. This allows for an evaluation of its ability to systematically reproduce the system being simulated (model reliability) and the level of accuracy in which the model reproduces the natural environment (model usefulness) (Huth and Holzworth, 2005; Wallach, 2006). Numerous model validation techniques exist; for a comprehensive overview read for example Bellocchi et al. (2010). The procedures to perform the task of validation appear in several forms, depending on data availability, system characteristics and researchers' opinions (Hsu et al., 1999). A common strategy is to examine the model's simulation versus actual observations acquired from the real world using common statistical metrics, and several validation studies of this type have been undertaken globally (Henderson-Sellers et al., 1995; Viterbo and Beljaars, 1995; Liang et al., 1998; Wang et al., 2007; Abramowitz et al., 2008; Slevin et al., 2015). In addition, Kramer et al. (2002), in an attempt to holistically assess the capability of a model for portraying a real world system, have proposed a set of model assessment criteria, namely accuracy, generality and realism. Accuracy is described by the authors as the "goodness of fit" to in situ measurements. Generality is described as the applicability of the model in numerous ecosystems. Realism is described as the ability of the model to address relationships between modelled phenomena.

The SimSphere land biosphere model is one example of a SVAT model. Formerly known as the Penn State University Biosphere-Atmosphere Modeling Scheme (PSUBAMS; Carlson and Boland, 1978; Carlson et al., 1981; Lynn and Carlson, 1990), this 1-D model was considerably modified to its current state by Gillies et al. (1997) and Petropoulos et al. (2013a). Since its early development, the model has become highly variable in its applicational use (for a recent overview of the model use and its applications see Petropoulos et al., 2009a). Amongst other uses, it has been involved in studies concerning the study of land surface interactions (Todhunter and Terjung, 1987; Ross and Oke, 1988) and the examination of hypothetical scenarios examining feedback processes (Wilson et al., 1999; Grantz et al., 1999). Furthermore, its synergistic use with Earth observation (EO) data is being considered at present for the development of operational products of energy fluxes and/or soil moisture on a global scale (Chauhan et al., 2003; ESA, 2012). These investigations have been based around the implementation of a technique commonly termed in the literature as the "triangle" (Carlson, 2007; Petropoulos and Carlson, 2011). A variant of this method, although it does not use SimSphere, is already deployed over Spain to operationally deliver surface soil moisture at 1 km spatial resolution from ESA's own SMOS satellite (Soil Moisture Ocean Salinity; Piles et al., 2011).

As SimSphere's use is rapidly expanding worldwide as both a research and educational tool, its validation and establishment of its coherence and correspondence to what it has been built to simulate is of paramount importance. In this respect, a series of SA (sensitivity analysis) experiments have already been conducted on the model (Olioso et al., 1996; Petropoulos et al., 2009b, 2013a–c). Such studies have allowed quantifying the relative influence of each model input to the simulation of key parameters by the model, ranking them in order of importance to understand how different parts of the model interplay. Yet, to our knowledge, validation studies involving direct comparisons of SimSphere predictions against in situ observations have as yet been scarce and incomprehensive. Such validation exercises have so far only been performed over a very small range of land use/cover types and on earlier versions of the model, when it was still under development (e.g. Todhunter and Terjung, 1987; Ross and Oke, 1988). Furthermore, to our knowledge, very few studies, if any, have validated SimSphere to numerous global ecosystems, for example, over Australian ecosystems. In this context, and given SimSphere use is currently expanding globally, a fully inclusive and comprehensive validation of the model is now of fundamental importance.

In preview of the above, the main objective of this study was to evaluate SimSphere's ability to model key parameters characterising land surface interactions. In this context, the main focus of this study has been to understand specifically the model's ability in predicting shortwave incoming radiation (R_g), net radiation (R_net), latent heat (LE), sensible heat (H), and air temperature (T_air) at a height of 1.3 and 50 m. Model validation is assessed through a comparison of the model results with corresponding observations from actual in situ measurements acquired at local scale from eight experimental sites (72 days in total) belonging to the OzFlux (Australia) and AmeriFlux (USA) global monitoring networks. This allowed including contrasting conditions in the model evaluation.

2 SimSphere model description

This work deals with the SimSphere 1-D boundary layer model devoted to the study of energy and mass interactions of the Earth system. Formerly known as PSUBAMS (Carlson and Boland, 1978; Carlson et al., 1981), this model was considerably modified to its current state by Gillies et al. (1997) and Petropoulos et al. (2013a). It is currently maintained and freely distributed from Aberystwyth University, United Kingdom (http://www.aber.ac.uk/simsphere). Further details about the model architecture can be found in Gillies (1993). In brief, the *physical* components ultimately determine the microclimate conditions in the model and are grouped into three categories, radiative, atmospheric and hydrological. The primary forcing of this component is the available clear sky radiant energy reaching the surface or the plant canopy, calculated as a function of sun and Earth geometry, atmospheric transmission factors for scattering and absorption, the atmospheric and surface emissivities and surface (including soil and plant) albedos.

The *vertical structure* effectively corresponds to the components of the planetary boundary layer (PBL) that are divided into four layers – *a surface mixing layer*, a *surface of constant flux layer*, a *surface of vegetation layer* and a *bare soil layer*. The depths of all these layers are somewhat variable with time. The top of the mixing layer is identified by the presence of a temperature inversion that caps the air in convective contact with the surface layer. At night, the situation is reversed as the Earth cools down more rapidly than the atmosphere. The surface "constant flux" layer evolves in the model as a series of equilibrium states between the transition layer below and the mixing layer above. Heat and moisture are assumed to be instantaneously conveyed between the surface and the top of the surface layer, which is chosen to be at a height of 50 m. In reality this height varies between approximately 20 and 50 m. The transition layer applies to a layer in which the vertical exchanges are dominated by molecular and radiative effects as well as by vertical wind changes. In the case of vegetation, the transition layer is represented by the microclimate within and at the top of the vegetation canopy. The substrate layer refers to the depth of the soil over which heat and water is conducted. It consists of two layers, a surface layer and a root zone. Water flows from the surface and the root zone to the atmosphere, respectively, by direct evaporation or through the plants as well as between the two layers. Soil water content is specified by assigning a fractional volume of field capacity, which essentially is the "soil moisture availability". Five layers are used to compute the flow of heat in the substrate. An initial soil temperature profile is assigned on the basis of the initial surface temperature (furnished from a meteorological sounding), as well as a climatological substrate temperature which one obtains from mean data. A governing parameter for heat conduction is the "thermal inertia" that contains both soil conductivity and soil diffusivity (or alternately, the volumetric heat content). This parameter is the one that also governs the rate of H flux to or from the atmosphere through the soil surface.

The *horizontal* component of the model is composed of four parts: (i) PBL, (ii) surface layer, (iii) transition layer and (iv) substrate layer. Due to SimSphere simulating parameters in a one-dimensional vertical column, the model is restricted horizontally only to areas representative of its initialised conditions; therefore, the model has an undefined spatial coverage. The vegetation component is dormant at night, that is, after radiation sunset. The night-time dynamics for the surface fluxes differ from those during the daytime. Heat and moisture fluxes are exchanged between both the ground and foliage, between plant and interplant airspaces through stomatal and cuticular resistances in the leaf (for water vapour) and the air, between soil and the interplant air spaces and between the entire vegetation canopy and the air. A separate component exists for the bare soil fluxes between the surface and the air. Vegetation and soil fluxes meld at the top of the vegetation canopy, their relative weights depending on the fractional vegetation cover, which is specified as an input to the model. As such, SimSphere is thus referred to as a form of two-stream or two-source model. The soil hydraulic parameters are prescribed from the Clapp and Hornberger (1978) classification. The soil surface turbulent fluxes are determined following the Monin and Obukov (1954) similarity theory which takes into account atmospheric stability.

SimSphere represents various physical processes taking place in a column that extends from the root zone below the soil surface up to a level well above the surface canopy, the top of the surface mixing layer. The processes and interactions simulated by the model are allowed to develop over a 24 h cycle at a chosen time step (typically 30 min), starting from a set of initial conditions given in the early morning. For its parameterisation, input parameters are categorised into seven defined groups; time and location, vegetation, surface, hydrological, meteorological, soil and atmospheric (Table 1). From initialisation, over a 24 h cycle SimSphere assesses the evolution of more than 30 variables associated with the radiative, hydrological and atmospheric physical domains.

3 Experimental set up

A total of five AmeriFlux and three OzNet experimental sites were used, providing a comprehensive data set of measured micrometeorological parameters together with general meteorological observations. The potential use of several FLUXNET sites was evaluated before deciding on the final eight experimental sites used in the study. Sites were excluded form analysis based on the requirement to fulfil specific criteria, namely (a) sites needed to incorporate different land cover types for the evaluation of the model's ability to simulate fluxes over different land cover/land use types; (b) sites were required to show homogeneous land cover, invariable topography and limited anthropogenic intervention; and (c) site data needed to include measurements of the six parameters validated in the study simultaneously for the same day. Any sites which did not successfully meet this criteria were excluded. Experimental days were further excluded following the pre-processing steps outlined in Sect. 4.1. Table 2 provides an overview of the characteristics of the experimental sites used in this study. At each site, micrometeorological measurements of various parameters are acquired including the turbulent fluxes of heat and moisture, R_g, R_{net} (at the surface) and T_{air} (often at different heights). Flux measurement methods and calculations performed within the FLUXNET sites are designed with the same specifications at all sites. All collected data are quality-controlled and standard procedures for error corrections are prescribed. Details on the FLUXNET measurements and the raw data processing can be found in Aubinet et al. (2000).

The sites were representative of a range of ecosystem types with markedly different site characteristics to include contrasting conditions in the model evaluation. All in situ data acquired from each site were collected covering the year 2011, allowing for a sufficient database for model parameterisation and validation to be developed. All data were obtained from the FLUXNET database (http://fluxnet.ornl.gov/obtain-data) at Level 2 processing, to allow for consistency and interoperability. This processing level includes the originally acquired in situ data from which any erroneous data caused by obvious instrumentation error have been removed. Additionally, atmospheric in situ data were collected from the freely distributed University of Wyoming's weather balloon data archive (http://weather.uwyo.edu/upperair/sounding.html). Local profiles of temperature, dew point temperature, wind direction, wind speed and atmospheric pressure were taken from the nearest possible experimental sites which were also used in model parameterisation.

4 SimSphere parameterisation and validation

This section provides a synopsis of the methodology followed in parameterising and subsequently evaluating SimSphere's ability to simulate key parameters characterising land surface interactions. An overview of the main steps included is furnished in Fig. 1.

4.1 Data set pre-processing

Following data acquisition, further analysis was implemented aimed at identifying the specific days for which SimSphere would be parameterised and validated for each experimental site. Initially, for each site, cloudy days were identified and eliminated from any further analysis. Judgment on which days (or time periods) were cloud free was based on the diurnal observations of R_g. In particular as cloud-free days were flagged as those having smoothly symmetrical R_g curves, a property signifying clear-sky conditions (e.g. Carlson et al., 1991).

Subsequently, for the subset of days which included only the cloud-free days, the energy balance closure (EBC) was computed. EBC evaluation has been accepted as a valid method for accuracy assessment of turbulent fluxes derived from eddy covariance measurements (Wilson et al., 2002; Barr et al., 2006). Energy imbalance provides important information on how they should be compared with model simulations (e.g. Twine et al., 2000; Culf et al., 2002). In this study, EBC was principally evaluated by performing a regression analysis (e.g. see Wilson and Baldocchi, 2000; Wilson et al., 2002; Castellvi et al., 2006). The linear regression coefficients (slope and intercept) as well as the coefficient of determination (R^2) were calculated from the ordinary least squares (OLS) relationship between the 30 min estimates of the dependent flux variables ($\mathrm{LE}+H$) and the independently derived available energy ($R_{net}-G-S$). In addition to this, the

Table 1. Summary of the main SimSphere inputs. The units of each of the model inputs are given in parentheses where applicable.

Name of the model input	Process in which parameter is involved	Min value	Max value
Slope (°)	time & location	0	45
Aspect (°)	time & location	0	360
Station height (m)	time & location	0	4.92
Fractional vegetation cover (%)	vegetation	0	100
LAI ($m^2\,m^{-2}$)	vegetation	0	10
Foliage emissivity (unitless)	vegetation	0.951	0.990
[Ca] (external [CO_2] in the leaf) (ppmv)	vegetation	250	710
[Ci] (internal [CO_2] in the leaf) (ppmv)	vegetation	110	400
[O_3] (ozone concentration in the air) (ppmv)	vegetation	0.0	0.25
Vegetation height (m)	vegetation	0.021	20.0
Leaf width (m)	vegetation	0.012	1.0
Minimum stomatal resistance ($s\,m^{-1}$)	plant	10	500
Cuticle resistance ($s\,m^{-1}$)	plant	200	2000
Critical leaf water potential (bar)	plant	−30	−5
Critical solar parameter ($W\,m^{-2}$)	plant	25	300
Stem resistance ($s\,m^{-1}$)	plant	0.011	0.150
Surface moisture availability (vol/vol)	hydrological	0	1
Root zone moisture availability (vol/vol)	hydrological	0	1
Substrate max volume water content (vol/vol)	hydrological	0.01	1
Substrate climatol. mean temperature (°C)	surface	20	30
Thermal inertia ($W\,m^{-2}\,K^{-1}$)	surface	3.5	30
Ground emissivity (unitless)	surface	0.951	0.980
Atmospheric precipitable water (cm)	meteorological	0.05	5
Surface roughness (m)	meteorological	0.02	2.0
Obstacle height (m)	meteorological	0.02	2.0
Fractional cloud cover (%)	meteorological	1	10
RKS (satur. thermal conduct.; $W\,m^{-1}\,K^{-1}$) (Farouki, 1981)	soil	0	10
Cosby B (unitless parameter) (see Cosby et al., 1984)	soil	2.0	12.0
THM (satur. vol. water cont.) (vol/vol) (Cosby et al., 1984)	soil	0.3	0.5
PSI (satur. water/matric potential) (kPa) (Wilson et al., 1994)	soil	1	7
Wind direction (°)	wind sounding profile	0	360
Wind speed (kn)	wind sounding profile	–	–
Altitude (1000 ft)	wind sounding profile	–	–
Pressure (mbar)	moisture sounding profile	–	–
Temperature (°C)	moisture sounding profile	–	–
Temperature – dew point temperature (°C)	moisture sounding profile	–	–

energy balance ratio (EBR) parameter was computed by cumulatively summing $R_{net}-G-S$ and LE$+H$ from the 30 min mean average surface energy flux components and then rationing each of the cumulative sums as follows (e.g. Wilson et al., 2002; Liu et al., 2006):

$$\mathrm{EBR}=\frac{\sum(\mathrm{LE}+H)}{\sum(\mathrm{Rn}-G-S)}. \quad (1)$$

In the above equation, G refers to the soil surface heat flux and S refers to the above-ground heat storage in the vegetation. This index ranges generally from 0 to 1, with values closer to 1 highlighting a satisfactory diurnal energy closure, indicating a good quality of in situ measurements. All days with poor EBC (EBR <0.750, slope <0.85, $R^2<0.930$) were excluded from further analysis.

Further conditions were subsequently employed to ensure that selected days were of the highest possible class in terms of in situ data quality. Firstly, all days selected were within the same year to eliminate effects ascribed from interannual variability in vegetation phenology or climatic conditions. Secondly, selected simulation days were assessed for atmospherically stable conditions, namely low wind speeds and low available energy (Maayar et al., 2001). Such conditions were identified by the evaluation of the in situ conditions, where direct measurements of wind speed and energy flux amplitude and diurnal trend were used as indicators of atmospherically stable conditions. As a result, a final set of a total of 72 non-consecutive days from the selected experimental

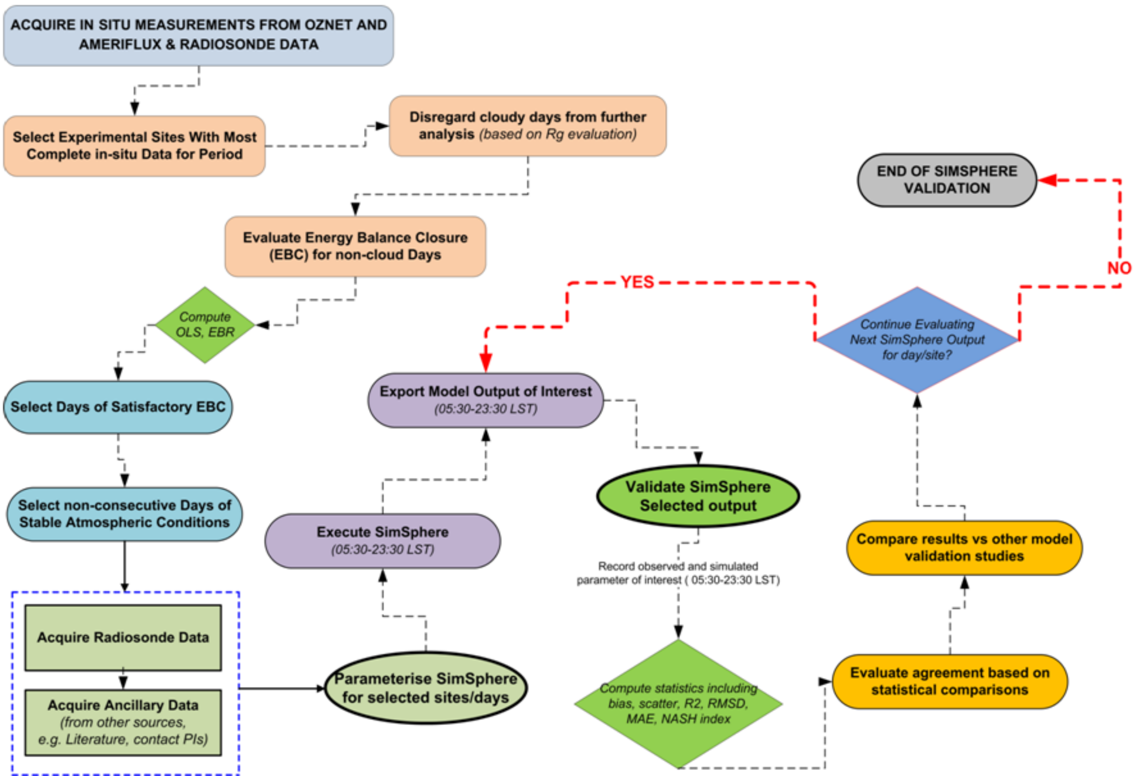

Figure 1. Flowchart of the overall methodology followed in evaluating SimSphere's outputs.

sites were identified as being suitable to be included in this study.

4.2 Model parameterisation

SimSphere was parameterised to the daily conditions existent at the flux tower for each of the selected days. In situ data sets provided measurements of soil water content, temperature, wind speed, wind direction and atmospheric pressure at the corresponding time of initialisation, 06:00 LT (local time). Ancillary parameters, critical for the model's initialisation, were largely acquired through either the site's respective principle investigator (for the case of OzFlux), or the FLUXNET database (for the case of AmeriFlux). Such measurements included detailed information on the vegetation (LAI, FVC, vegetation height, cuticle resistance), pedological (soil morphology and soil classification) and topographical (slope, aspect, surface roughness) characteristics of each site. If no further ancillary information was available, specific parameters were acquired through the analysis of standard literature sources (e.g. Mascart et al., 1991; Carlson et al., 1991). The soil type parameters were obtained using the soil texture data provided at each FLUXNET test site and information supplied in some instances by the experimental site managers themselves. This was also the case for the topographical information required in model initialisation. Wind and water vapour sounding profiles which were attained at 06:00 GMT from the University of Wyoming database to correspond to the model's initialisation were also used in model parameterisation. Upon completion of its initialisation, the model was executed for each site/day, forced by the observations acquired from the site on which it was parameterised. The 30 min average value of each of the targeted model outputs per site for the period 05:30–23:30 LT was subsequently exported in Statistical Package for the Social Sciences (SPSS) to validate the model predictions.

4.3 Model performance assessment

A series of statistical terms included to evaluate the agreement between the in situ and the SimSphere predictions, including the mean bias error (MBE or bias; Eq. 2) and mean standard deviation (MSD or scatter; Eq. 3) of the observed and modelled values, the RMSD (Eq. 4), the mean absolute difference (MAD; Eq. 5), the linear regression fit model coefficient of determination (R^2; Eq. 6) and the Nash and Sutcliffe (1970; denoted as Nash) index (Eq. 7):

$$\text{bias} = \text{MBE} = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i), \quad (2)$$

$$\text{scatter} = \text{MSD} = \frac{1}{(N-1)}\sum_{i=1}^{N}(P_i - O_i - \overline{(P_i - O_i)})^2, \quad (3)$$

Table 2. Description of selected experimental sites used for validating SimSphere.

Site name	Site abbreviation	Country	Geographic location (lat, long)	PFT	Ecosystem type	Dominant species	Elevation	Climate
Alice Springs	–	Australia	−22.283, 133.249	MWO	mulga woodland	*Acacia aneura*	606 m	desert: hot and dry summers and cold winters
Calperum	–	Australia	−34.003, 140.588	PAS	grazing pasture	*Eucalyptus stricta*	200 m	subtropical dry summer
Howard Springs	–	Australia	−12.495, 131.15	WSV	woody savannah	*Eucalyptus miniata* and *Eucalyptus tetrodonta*	64 m	tropical wet and dry: hot and humid summers
Vaira Ranch	US_VAR	USA	38.406, −120.950	GRA	grassland	*Brachypodium distachyon, Hypochaeris glabra, Trifolium dubium*	129 m	Mediterranean: hot and dry summers, wet and cold winters
Missouri Ozark	US_MOZ	USA	38.7441, −92.200	DBL	deciduous broadleaf	*Quercus alba, Quercus velutina, Carya ovata*	219 m	temperate continental
Fermi Agricultural	US_IB1	USA	41.8593, −88.2227	CRO	cropland	soybean (C3)	225 m	wet and hot summers and mild winters
Tonzi Ranch	US_TON	USA	38.4316, −120.9660	WSV	woody savannah	*Quercus douglasii, Pinus sabiniana, Brachypodium distachyon*	169 m	Mediterranean: hot and dry summers, wet and cold winters
Lucky Hills Shrubland	US_WHS	USA	31.7438, −110.0522	SHR	shrubland	*Larrea tridentata, Acacia constricta, Flourensia cernua*	1372 m	semi-arid

$$\mathrm{RMSD} = \sqrt{\mathrm{bias}^2 + \mathrm{scatter}^2}, \tag{4}$$

$$\mathrm{MAD} = N^{-1}\sum_{i=1}^{N} |P_i - O_i|, \tag{5}$$

$$R^2 = \left[\sum_{i=1}^{N}\left(P_i - \bar{P}\right)\left(O_i - \bar{O}\right) / \left[\sum_{i=1}^{N}(O_i - \bar{O})^2 \sum_{i=1}^{N} P_i - \bar{O})^2\right]^{0.5}\right]^2, \tag{6}$$

$$\mathrm{Nash} = 1 - \left[\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2}\right]. \tag{7}$$

P denotes the "predicted" values obtained from SimSphere and O denotes the "observed" values from the selected OzFlux- and AmeriFlux-site days.

The utilisation of these statistics has been widely demonstrated in a number of previous studies comparing model outputs to observational networks (e.g. Alexandris and Kerkides, 2003; Marshall et al., 2013). All statistical metrics were computed from comparisons performed at identical 30 min intervals between the two data sets for each day of comparison. In addition, these statistical parameters, where appropriate, were also computed for each site, providing a summary of the model predictions per experimental site on which the model was validated.

5 Results

The main results from the comparisons between the SimSphere predictions and the corresponding in situ data for the different parameters evaluated in this study are summarised in Tables 3–8. In addition, Fig. 2 provides a graphical illustration of the agreement between the simulated values and in situ measurements per parameter for all sites together and Fig. 3 illustrates the diurnal agreement between the modelled outputs and in situ-observed fluxes for a selected site and day. The detailed findings from the comparisons performed are made available next.

5.1 Incoming shortwave radiation (R_g) at the surface

Simulation accuracy of R_g was largely accurate, exhibited by low RMSD (within ∼ 19 % of the observed fluxes) and MAE values (RMSD = 67.83 W m^{-2}, MAE = 46.43 W m^{-2}; Table 3, Fig. 2). A moderate underestimation of the observed fluxes was also evident (MBE = −19.48 W m^{-2}). Notably, R_g yielded the highest correlated results of all parameters assessed ($R^2 = 0.971$, Nash = 0.963), further illustrated in Fig. 2, where the distribution of points within the feature space were predominantly centred on the 1 : 1 line, showing a strong relationship between both variables.

Table 3. Daily simulation accuracy and average site simulation accuracy for R_g fluxes. Bias, scatter, RMSD and MAE are expressed in watts per square metre ($W m^{-2}$). Nash index is unitless.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash
Alice Springs	23/03/2011	−5.53	33.38	33.83	8.25	24.74	0.998
	15/04/2011	13.56	28.84	31.87	8.90	19.10	0.956
	23/04/2011	3.96	29.62	29.88	8.40	19.37	0.974
	10/05/2011	1.82	20.40	20.48	6.37	13.41	0.979
	24/05/2011	−16.47	25.45	30.32	10.29	20.29	0.924
	31/05/2011	−13.52	21.89	25.73	8.73	17.08	0.996
	18/06/2011	−26.93	32.75	42.40	15.37	28.03	0.949
	25/06/2011	−35.78	39.47	53.27	19.01	35.84	0.993
	18/07/2011	−34.00	33.93	48.04	16.73	34.00	1.000
	20/08/2011	−48.38	40.44	63.06	17.87	48.38	0.975
	Average	−19.48	62.36	67.825	21.20	46.29	0.974
Calperum	24/02/2011	9.68	23.06	25.01	5.85	19.077	0.994
	02/03/2011	8.41	22.63	24.14	5.71	18.314	0.979
	31/03/2011	30.48	28.25	41.56	12.30	30.482	0.986
	24/04/2011	41.93	33.67	53.78	20.58	41.932	0.975
	22/07/2011	−58.28	61.06	84.41	40.79	60.624	0.978
	28/07/2011	−67.87	71.01	98.22	46.28	70.950	0.974
	28/08/2011	−108.13	102.92	149.29	52.81	110.484	0.889
	01/12/2011	−110.33	75.49	133.69	26.40	112.586	0.899
	23/12/2011	−76.00	62.66	98.50	19.34	78.332	0.978
	29/12/2011	−74.10	62.08	96.67	18.56	76.348	0.991
	Average	−40.42	80.91	90.45	24.52	61.91	0.964
Howard Springs	18/04/2011	18.24	20.76	27.64	7.64	18.78	0.975
	23/04/2011	7.81	15.15	17.04	4.67	11.64	0.978
	13/05/2011	−0.93	20.24	20.26	5.91	15.11	0.989
	27/05/2011	24.47	29.62	38.42	12.84	25.10	0.978
	03/06/2011	−8.37	34.64	35.64	10.94	27.60	0.935
	14/06/2011	−20.95	43.62	48.39	14.86	35.50	0.974
	22/06/2011	−15.48	42.38	45.12	14.31	33.86	0.976
	22/07/2011	−37.30	56.85	67.99	21.94	48.96	0.982
	28/07/2011	−63.83	69.49	94.36	28.24	67.30	0.989
	27/09/2011	−52.80	51.87	74.01	19.51	54.04	0.979
	Average	−14.913	50.367	52.528	15.64	33.789	0.976
US_MOZ	28/06/2011	−48.13	51.40	70.42	15.12	59.86	0.976
	01/08/2011	−5.55	34.91	35.35	8.87	24.81	0.976
	18/08/2011	−2.57	35.53	35.63	8.84	27.93	0.991
	31/08/2011	42.46	42.04	59.76	17.57	42.46	0.974
	01/09/2011	34.48	30.62	46.11	13.23	34.48	0.978
	07/09/2011	4.83	41.09	41.38	10.62	30.60	0.987
	12/09/2011	16.18	33.51	37.21	10.52	24.67	0.969
	30/09/2011	29.14	34.46	45.10	14.38	29.22	0.988
	29/09/2011	42.10	34.04	54.14	23.88	42.10	0.978
	11/11/2011	48.52	44.14	65.59	33.89	48.52	0.972
	Average	16.50	47.58	50.36	14.67	36.57	0.979

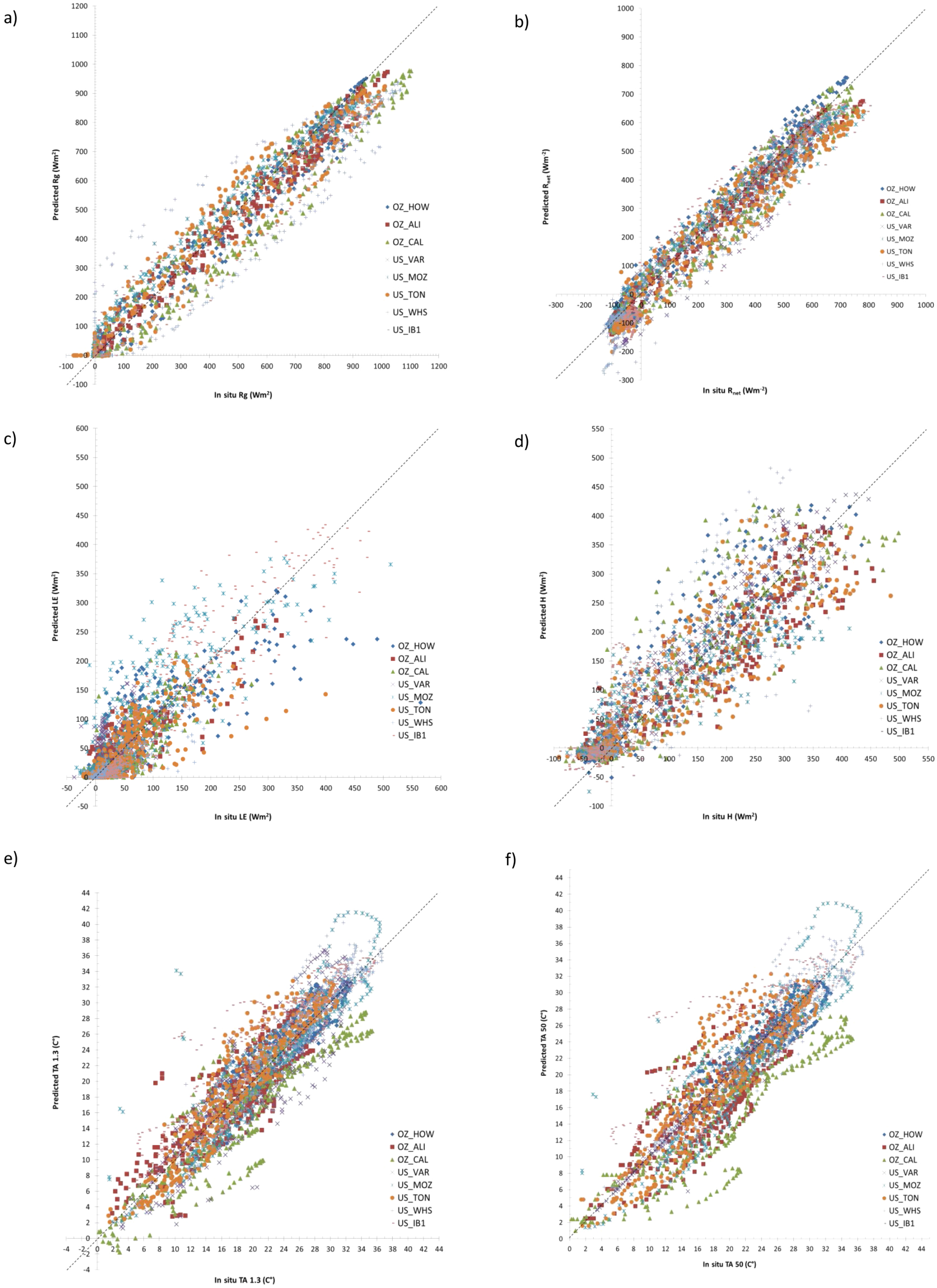

Figure 2. Scatterplot comparison of SimSphere-predicted and in situ **(a)** R_g flux, **(b)** R_{net} flux, **(c)** LE flux, **(d)** H flux, **(e)** $T_{air\ 1.3\ m}$, and **(f)** $T_{air\ 50\ m}$.

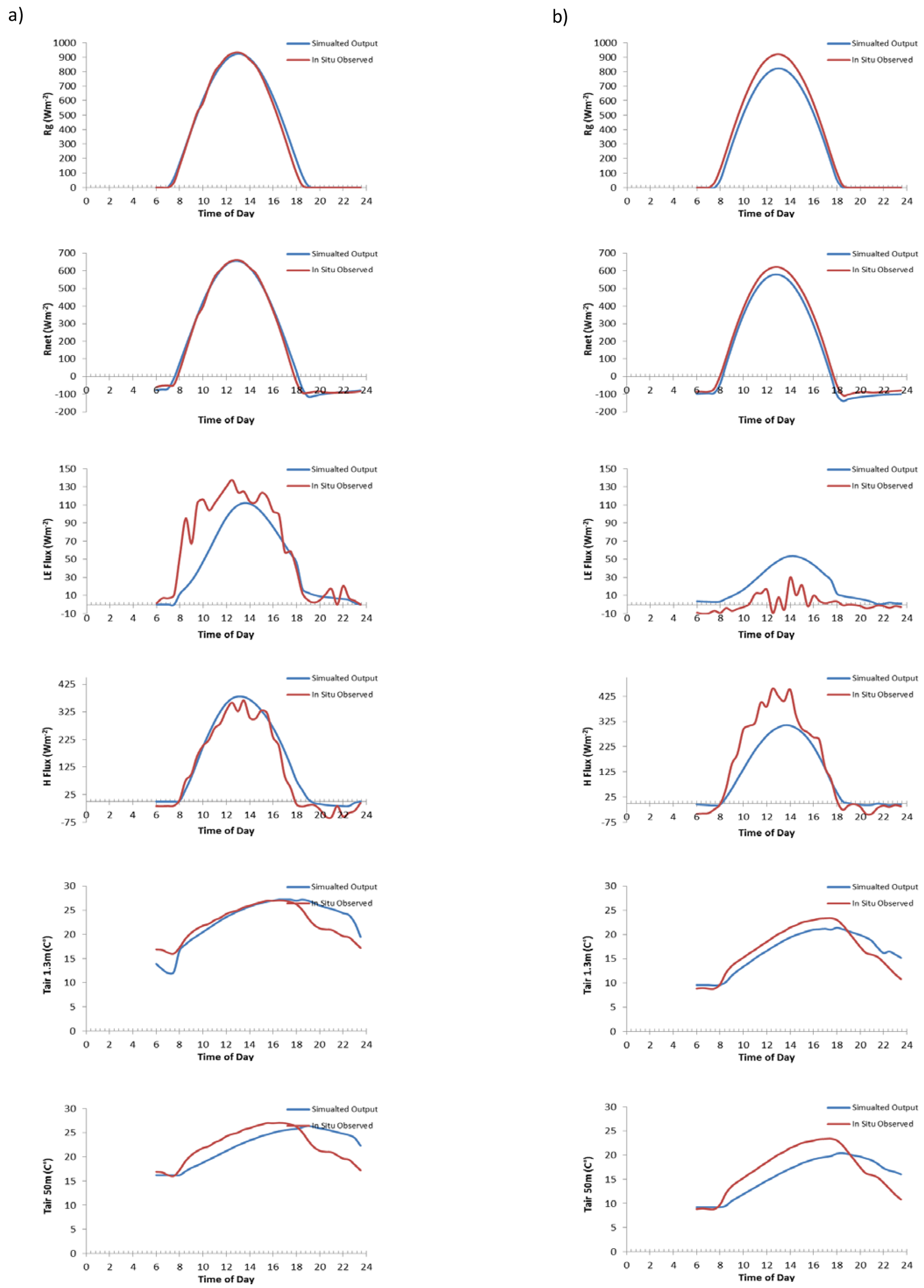

Figure 3. Simulated and observed fluxes for the Alice Springs site (shrubland); **(a)** illustrates the diurnal trend of the simulated fluxes from SimSphere against the observed in situ fluxes for 15 April 2011 (spring); **(b)** illustrates the diurnal trend of the simulated fluxes from SimSphere against the observed in situ fluxes for 20 August 2011 (summer).

Table 3. Continued.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash
US_IB1	30/05/2011	−70.94	67.44	97.88	22.57	70.94	0.939
	07/06/2011	−64.46	68.10	93.76	21.27	65.04	0.898
	28/06/2011	−69.64	69.19	98.17	20.03	72.25	0.899
	08/07/2011	−55.80	74.50	93.08	19.71	67.98	0.937
	24/08/2011	7.96	56.42	56.98	15.31	38.42	0.986
	13/09/2011	12.64	43.93	45.71	12.96	31.17	0.978
	15/09/2011	−2.54	43.42	43.50	12.71	29.90	0.940
	01/10/2011	13.80	42.18	44.38	12.00	27.31	0.977
	15/10/2011	12.39	47.00	48.61	17.53	29.42	0.949
	24/10/2011	15.15	45.93	48.365	19.38	28.51	0.997
	Average	−20.15	68.20	71.114	18.71	46.09	0.950
US_TON	27/02/2011	39.37	24.89	46.58	37.72	39.68	0.961
	17/03/2011	−88.37	74.91	115.85	37.22	88.37	0.899
	24/05/2011	−77.28	51.05	92.61	20.19	77.28	0.961
	24/06/2011	−62.15	40.59	74.23	15.30	62.15	0.965
	30/07/2011	−10.44	17.10	20.04	4.62	15.34	0.973
	07/08/2011	−19.86	27.43	33.87	7.76	24.87	0.984
	28/08/2011	−1.79	19.71	19.79	4.83	14.83	0.991
	15/09/2011	46.82	36.15	59.15	17.80	46.82	0.974
	01/11/2011	66.77	55.13	86.59	40.25	66.77	0.925
	16/11/2011	58.47	50.65	77.36	43.03	58.47	0.941
	Average	−4.85	69.54	69.71	20.59	49.46	0.957
US_WHS	08/02/2011	−119.41	122.29	170.92	35.60	119.474	0.899
	16/02/2011	−124.62	114.72	169.39	55.35	124.624	0.845
	25/03/2011	−141.67	114.86	182.38	44.63	141.666	0.880
	22/06/2011	−73.15	48.54	87.79	17.72	73.152	0.937
	13/07/2011	−77.12	63.05	99.61	20.11	78.604	0.913
	02/08/2011	−42.92	63.54	76.68	17.01	59.743	0.986
	28/08/2011	−21.54	47.97	52.59	12.80	41.999	0.983
	03/08/2011	−11.92	36.71	38.59	9.59	29.599	0.997
	05/10/2011	−1.32	35.02	35.04	10.04	24.874	0.985
	20/10/2011	11.97	27.15	29.67	9.50	18.541	0.991
	Average	−56.40	83.36	100.65	24.48	67.45	0.942
	Average of all sites	−19.48	62.36	67.83	19.19	46.42	0.963

On a per site basis, the highest simulation accuracies were attained within the US_MOZ deciduous broadleaf site, in comparison to all other sites ($RMSD = 50.36\,W\,m^{-2}$, within $\sim 15\,\%$ of the observed fluxes, $MAE = 36.57\,W\,m^{-2}$). The Howard Springs woody savannah site also attained comparably high simulation accuracies ($RMSD = 52.53\,W\,m^{-2}$, within $\sim 16\,\%$ of the observed fluxes, $MAE = 33.79\,W\,m^{-2}$). Contrarily, model predictions of R_g for the Australian Calperum grazing pasture site were significantly lower, indicating a weaker model performance ($RMSD = 100.65\,W\,m^{-2}$, within $\sim 25\,\%$ of the observed fluxes, $MAE = 61.91\,W\,m^{-2}$), closely followed by the US_WHS shrubland site ($RMSD = 90.45\,W\,m^{-2}$, within $\sim 25\,\%$ of the observed fluxes, $MAE = 46.09\,W\,m^{-2}$). Within the majority of sites, model simulations consistently underestimated the in situ measurements ($MBE = -4.85$ to $-56.40\,W\,m^{-2)}$, with the US_MOZ deciduous forest site being the only exception ($MBE = 16.47\,W\,m^{-2}$). That is, the true change (in situ observations), for six of the seven sites tends to be larger than the model-based estimates. Intersite variability was minimal for the simulation of this parameter, with only a difference of $\sim 9\,\%$ between the minimum and maximum RMSD as a percentage of the observed fluxes on a per site basis.

Evidently, agreement over the Australian sites generally increased for the period between February and June, with a significant decrease in accuracy from August to early February. For example, over the Calperum grazing pasture site,

Table 4. Daily simulation accuracy and average site simulation accuracy for R_{net} fluxes. Bias, scatter, RMSD and MAE are expressed in $W m^{-2}$. Nash index is unitless.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
Alice Springs	23/03/2011	−47.84	39.66	62.14	18.48	49.88	0.989
	15/04/2011	5.37	20.58	21.27	8.37	15.35	0.978
	23/04/2011	5.82	20.03	20.86	8.93	15.03	0.982
	10/05/2011	0.24	19.92	19.92	10.08	16.86	0.981
	24/05/2011	15.02	14.52	20.89	11.86	17.07	0.968
	31/05/2011	−16.37	18.30	24.55	14.32	20.45	0.991
	18/06/2011	−32.89	21.07	39.06	22.95	34.37	0.974
	25/06/2011	−40.45	18.12	44.32	27.28	40.62	0.979
	18/07/2011	−17.88	11.17	21.08	11.86	18.28	0.998
	20/08/2011	−34.57	13.29	37.04	16.38	34.57	0.964
	Average	−16.35	29.69	33.90	16.23	26.25	0.980
Calperum	24/02/2011	28.31	33.37	43.76	14.33	38.93	0.979
	02/03/2011	2.23	22.55	22.66	7.88	17.92	0.998
	31/03/2011	10.28	26.72	28.63	13.03	24.49	0.982
	24/04/2011	36.99	44.56	57.91	36.50	49.76	0.981
	22/07/2011	−62.63	39.68	74.14	69.82	62.63	0.968
	28/07/2011	−42.48	38.93	57.62	53.47	42.56	0.964
	28/08/2011	−76.72	58.52	96.49	55.19	76.72	0.945
	01/12/2011	−70.84	52.79	88.34	23.33	74.16	0.911
	23/12/2011	−18.27	33.56	38.21	10.26	26.07	0.965
	29/12/2011	−40.99	41.01	57.98	15.64	42.62	0.971
	Average	−23.41	56.46	61.12	24.63	45.59	0.966
Howard Springs	18/04/2011	22.80	32.62	39.79	13.93	32.82	0.963
	23/04/2011	17.03	30.42	34.86	11.58	28.66	0.944
	13/05/2011	40.73	28.01	49.44	21.98	40.77	0.956
	27/05/2011	54.63	44.72	70.60	38.42	56.14	0.939
	03/06/2011	20.03	27.17	33.75	17.42	25.21	0.985
	14/06/2011	16.26	33.68	37.39	19.99	29.82	0.985
	22/06/2011	10.77	39.44	40.89	22.60	29.58	0.989
	22/07/2011	−0.61	34.49	34.50	17.89	26.80	0.967
	28/07/2011	−51.75	47.36	70.15	32.05	57.36	0.995
	27/09/2011	−26.45	29.78	39.82	14.85	30.20	0.997
	Average	10.35	45.89	47.05	21.03	35.74	0.972
US_VAR	10/05/2011	−32.46	19.86	38.05	12.24	32.46	0.974
	23/06/2011	−36.76	33.67	49.85	14.69	44.40	0.987
	19/07/2011	−10.81	34.63	36.28	11.48	31.93	0.989
	30/07/2011	−2.93	49.87	49.95	17.07	43.81	0.974
	07/08/2011	4.39	40.18	40.42	14.71	32.47	0.911
	27/08/2011	40.92	61.81	74.13	32.86	68.51	0.978
	22/09/2011	43.98	65.16	78.61	49.50	72.56	0.946
	07/10/2011	−2.19	85.26	85.29	52.10	78.18	0.998
	26/11/2011	3.42	61.11	61.21	74.33	54.67	0.996
	19/12/2011	−8.42	47.35	48.09	102.45	43.57	0.996
	Average	−0.09	58.64	58.64	26.52	50.26	0.975

Table 4. Continued.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
US_MOZ	28/06/2011	−88.46	58.74	106.19	26.25	91.19	0.957
	01/08/2011	−8.96	31.83	33.07	9.28	23.32	0.984
	18/08/2011	−29.16	31.88	43.20	13.01	38.60	0.989
	31/08/2011	−7.51	36.16	36.93	12.47	31.74	0.969
	01/09/2011	5.45	26.09	26.65	9.00	20.74	0.968
	07/09/2011	−26.40	51.75	58.09	20.09	43.98	0.964
	12/09/2011	−2.30	29.74	29.83	10.55	23.89	0.981
	30/09/2011	−17.85	46.09	49.42	22.39	37.06	0.991
	29/09/2011	33.28	35.39	48.58	34.83	33.77	0.905
	11/11/2011	54.81	64.20	84.28	87.69	56.09	0.886
	Average	−13.25	49.83	51.56	19.00	38.46	0.959
US_IB1	30/05/2011	−86.39	70.85	111.73	26.35	86.39	0.842
	07/06/2011	−35.43	40.05	53.47	14.38	37.86	0.986
	28/06/2011	−38.58	33.74	51.25	13.39	40.59	0.972
	08/07/2011	−52.02	19.96	55.72	15.01	52.02	0.976
	24/08/2011	19.23	54.20	57.51	18.55	41.64	0.946
	13/09/2011	15.26	54.05	56.16	18.53	48.64	0.977
	15/09/2011	−1.69	70.25	70.27	27.59	59.80	0.899
	01/10/2011	15.91	58.94	61.05	23.83	45.12	0.985
	15/10/2011	24.75	73.02	77.10	43.41	68.48	0.978
	24/10/2011	−28.90	73.82	79.27	51.27	71.18	0.996
	Average	−16.79	67.54	69.59	23.15	55.17	0.956
US_TON	27/02/2011	−101.40	51.67	113.80	73.72	101.40	0.911
	17/03/2011	−88.31	35.39	95.13	46.41	88.31	0.913
	24/05/2011	−70.18	38.19	79.89	21.08	70.18	0.952
	24/06/2011	−83.36	42.99	93.79	24.11	83.36	0.962
	30/07/2011	−65.26	42.12	77.67	21.57	66.65	0.986
	07/08/2011	−53.89	54.31	76.51	22.52	58.28	0.965
	28/08/2011	−39.97	57.08	69.69	22.73	58.79	0.971
	15/09/2011	2.42	38.27	38.35	16.01	30.94	0.966
	01/11/2011	26.56	47.53	54.45	51.09	46.09	0.984
	16/11/2011	12.42	48.78	50.34	52.70	48.18	0.963
	Average	−46.10	62.96	78.03	30.30	65.22	0.957
US_WHS	08/02/2011	−56.66	73.69	92.95	36.12	66.57	0.912
	16/02/2011	−71.45	65.15	96.69	61.91	75.32	0.872
	25/03/2011	−70.67	57.33	91.00	39.10	75.11	0.874
	22/06/2011	−55.39	72.62	91.33	34.65	59.76	0.929
	13/07/2011	−10.84	27.38	29.45	10.03	23.78	0.985
	02/08/2011	−15.37	36.24	39.37	11.95	30.58	0.964
	28/08/2011	5.33	26.54	27.07	10.11	18.49	0.996
	03/08/2011	−24.34	51.80	57.24	22.11	41.30	0.996
	05/10/2011	48.88	27.23	55.95	29.01	48.88	0.968
	20/10/2011	8.07	52.60	53.22	34.52	50.05	0.978
	Average	−26.24	64.52	69.653	28.94	50.271	0.947
	Average of all sites	−16.49	54.44	58.69	23.81	45.90	0.964

Table 5. Daily simulation accuracy and average site simulation accuracy for LE fluxes. Bias, scatter, RMSD and MAE are expressed in $W m^{-2}$. Nash index is unitless.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
Alice Springs	23/03/2011	−23.75	45.45	51.28	18.34	36.85	0.997
	15/04/2011	−17.30	23.04	28.81	23.21	19.96	0.992
	23/04/2011	2.76	23.88	24.04	30.85	14.13	0.989
	10/05/2011	20.87	19.88	28.82	87.51	21.32	0.935
	24/05/2011	4.59	4.68	6.56	36.44	5.44	0.969
	31/05/2011	5.12	8.63	10.04	51.10	6.65	0.968
	18/06/2011	−0.34	8.61	8.61	26.74	6.70	0.979
	25/06/2011	3.25	9.22	9.77	44.77	7.45	0.950
	18/07/2011	12.90	13.33	18.55	124.66	13.42	0.914
	20/08/2011	19.44	14.83	24.45	145.53	19.44	0.758
	Average	2.75	24.59	24.75	36.03	15.16	0.945
Calperum	24/02/2011	−9.77	31.40	32.89	20.08	23.06	0.995
	02/03/2011	−13.83	25.93	29.39	25.35	21.17	0.992
	31/03/2011	−8.48	18.35	20.21	22.22	13.19	0.994
	24/04/2011	−8.26	17.96	19.76	32.63	13.20	0.990
	22/07/2011	−7.97	15.53	17.45	54.76	10.97	0.979
	28/07/2011	−9.24	13.33	16.22	35.06	11.54	0.983
	28/08/2011	−17.69	24.64	30.33	63.45	19.45	0.979
	01/12/2011	−5.22	20.11	20.78	21.99	15.76	0.988
	23/12/2011	24.57	39.14	46.21	31.35	31.75	0.993
	29/12/2011	−11.57	30.29	32.43	21.10	24.78	0.993
	Average	−6.75	27.20	28.02	29.41	18.49	0.989
Howard Springs	18/04/2011	−31.86	46.21	56.13	22.11	40.76	0.997
	23/04/2011	−17.90	77.00	79.06	24.84	46.29	0.998
	13/05/2011	−5.36	23.19	23.80	14.63	17.17	0.997
	27/05/2011	35.70	44.91	57.37	71.24	39.41	0.970
	03/06/2011	26.12	37.60	45.78	74.56	29.79	0.976
	14/06/2011	7.11	16.14	17.64	30.44	12.01	0.984
	22/06/2011	31.51	35.67	47.60	52.70	36.33	0.982
	22/07/2011	13.30	29.13	32.02	30.11	20.23	0.993
	28/07/2011	−10.94	20.67	23.39	15.82	17.39	0.996
	27/09/2011	−25.35	70.48	74.90	32.73	39.03	0.965
	Average	2.23	50.06	50.11	22.23	29.84	0.986
US_VAR	10/05/2011	−9.01	13.06	15.87	12.82	12.66	0.968
	23/06/2011	29.67	38.13	48.31	76.27	31.90	0.978
	19/07/2011	23.91	29.52	37.99	193.52	25.48	0.928
	30/07/2011	27.99	31.61	42.22	357.06	29.02	0.292
	07/08/2011	22.12	25.56	33.80	354.25	22.98	0.654
	27/08/2011	24.33	29.46	38.21	532.37	24.56	0.665
	22/09/2011	17.85	21.54	27.97	403.04	17.85	0.414
	07/10/2011	6.59	27.20	27.98	43.26	19.53	0.979
	26/11/2011	−2.67	13.20	13.47	27.84	8.58	0.992
	19/12/2011	−2.61	10.60	10.91	34.99	7.21	0.985
	Average	13.817	28.93	32.06	92.96	19.98	0.786

Table 5. Continued.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
US_MOZ	28/06/2011	−11.80	56.09	57.32	16.41	43.455	0.912
	01/08/2011	66.84	84.61	107.83	42.92	73.193	0.912
	18/08/2011	25.06	59.74	64.79	22.93	45.616	0.937
	31/08/2011	37.95	49.68	62.51	39.44	41.24	0.912
	01/09/2011	46.76	62.26	77.87	49.50	53.78	0.927
	07/09/2011	21.02	48.81	53.14	38.55	38.27	0.869
	12/09/2011	40.56	50.34	64.65	49.34	45.22	0.945
	30/09/2011	15.96	38.19	41.39	37.46	28.55	0.974
	29/09/2011	16.38	35.63	39.22	119.95	35.57	0.945
	11/11/2011	28.35	32.97	43.48	115.51	32.72	0.841
	Average	25.65	55.92	61.52	37.32	42.02	0.917
US_IB1	30/05/2011	−28.88	61.84	68.26	16.15	54.17	0.899
	07/06/2011	40.29	71.27	81.87	28.77	65.32	0.927
	28/06/2011	32.16	51.86	61.02	31.59	49.59	0.982
	08/07/2011	−35.32	28.67	45.49	17.58	35.36	0.947
	24/08/2011	1.74	37.11	37.15	9.69	31.07	0.972
	13/09/2011	−1.04	50.50	50.51	15.28	43.88	0.821
	15/09/2011	−6.30	15.45	16.68	6.14	13.25	0.998
	01/10/2011	0.80	37.23	37.24	16.76	28.78	0.964
	15/10/2011	38.31	53.74	66.00.	43.70	52.64	0.979
	24/10/2011	−14.13	17.31	22.35	14.22	18.56	0.978
	Average	2.76	52.47	52.54	19.64	39.26	0.947
US_TON	27/02/2011	−5.85	22.86	23.60	31.85	17.43	0.981
	17/03/2011	−16.50	43.06	46.11	33.43	32.99	0.969
	24/05/2011	−56.28	73.75	92.78	39.70	62.52	0.899
	24/06/2011	−3.14	35.44	35.58	21.81	27.23	0.948
	30/07/2011	6.05	29.06	29.68	41.56	20.93	0.969
	07/08/2011	2.09	20.96	21.06	24.63	16.99	0.990
	28/08/2011	0.90	16.51	16.54	31.22	11.71	0.985
	15/09/2011	7.75	22.49	23.79	63.47	14.02	0.983
	01/11/2011	−2.22	14.10	14.28	20.75	11.12	0.991
	16/11/2011	4.30	10.10	10.98	30.59	7.15	0.987
	Average	−6.29	38.27	38.79	40.36	22.21	0.970
US_WHS	08/02/2011	9.61	12.40	15.69	217.20	10.35	0.886
	16/02/2011	1.03	7.80	7.87	102.72	4.61	0.946
	25/03/2011	−0.038	5.98	5.98	103.62	4.22	0.925
	22/06/2011	−2.64	6.02	6.57	63.29	4.47	0.913
	13/07/2011	−5.69	21.22	21.97	42.26	16.75	0.956
	02/08/2011	−43.53	36.74	56.96	27.02	44.83	0.975
	28/08/2011	−39.80	37.57	54.73	46.11	41.24	0.979
	03/08/2011	−12.72	15.97	20.41	25.42	15.11	0.986
	05/10/2011	−13.01	17.25	21.61	51.87	13.88	0.973
	20/10/2011	0.18	7.57	7.57	40.99	4.81	0.966
	Average	−11.49	25.52	27.99	50.61	15.36	0.951
	Average of all sites	2.836	37.870	39.472	33.70	25.591	0.936

Table 6. Daily simulation accuracy and average site simulation accuracy for H fluxes. Bias, scatter, RMSD and MAE are expressed in $W\,m^{-2}$. Nash index is unitless.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
Alice Springs	23/03/2011	−24.28	61.35	65.98	36.42	56.40	0.996
	15/04/2011	25.00	28.48	37.90	16.51	29.89	0.963
	23/04/2011	2.38	42.43	42.50	17.16	32.46	0.965
	10/05/2011	−24.02	64.04	68.40	28.09	53.23	0.975
	24/05/2011	9.20	27.77	29.25	12.40	24.61	0.921
	31/05/2011	−17.74	44.73	48.12	20.25	34.45	0.932
	18/06/2011	−16.03	37.98	41.22	19.41	28.27	0.983
	25/06/2011	−11.18	39.11	40.68	21.86	26.44	0.998
	18/07/2011	−7.95	28.68	29.76	12.63	22.79	0.999
	20/08/2011	−37.00	65.84	75.52	26.10	54.33	0.973
	Average	−10.16	49.35	50.39	22.57	36.29	0.970
Calperum	24/02/2011	58.73	62.79	85.97	41.06	69.62	0.981
	02/03/2011	4.58	46.74	46.96	16.77	35.21	0.963
	31/03/2011	8.70	42.43	43.31	20.97	30.60	0.899
	24/04/2011	67.41	72.42	98.93	70.00	74.96	0.991
	22/07/2011	−19.03	34.44	39.34	29.72	25.54	0.997
	28/07/2011	−1.21	32.85	32.88	30.46	25.32	0.998
	28/08/2011	−14.37	31.47	34.60	19.36	22.87	0.998
	01/12/2011	−20.74	38.84	44.02	11.19	36.18	0.986
	23/12/2011	−15.69	33.46	36.96	11.17	30.30	0.951
	29/12/2011	−12.29	38.80	40.70	12.26	32.77	0.932
	Average	5.61	54.53	54.81	23.70	38.34	0.970
Howard Springs	18/04/2011	56.78	50.31	75.86	51.67	58.88	0.995
	23/04/2011	24.08	34.73	42.26	32.73	29.46	0.996
	13/05/2011	69.81	67.25	96.93	65.29	70.17	0.995
	27/05/2011	12.17	32.14	34.36	16.66	24.12	0.973
	03/06/2011	12.11	42.25	43.95	21.14	30.03	0.963
	14/06/2011	19.13	46.53	50.31	21.14	34.01	0.932
	22/06/2011	−18.82	44.08	47.93	26.97	34.39	0.998
	22/07/2011	−9.05	26.81	28.29	15.32	19.52	0.937
	28/07/2011	−14.96	43.91	46.39	25.46	31.70	0.974
	27/09/2011	3.94	39.00	39.20	20.99	29.47	0.912
	Average	15.52	51.92	54.19	29.97	36.18	0.967
US_VAR	10/05/2011	37.64	40.41	55.22	23.14	41.20	0.889
	23/06/2011	−5.64	26.33	26.93	8.81	19.04	0.987
	19/07/2011	10.05	25.86	27.74	8.07	22.16	0.931
	30/07/2011	−7.48	31.14	32.03	9.83	23.88	0.847
	07/08/2011	11.30	24.19	26.70	8.75	21.24	0.869
	27/08/2011	29.36	37.65	47.74	19.25	37.53	0.899
	22/09/2011	34.80	28.53	45.00	24.92	38.05	0.899
	07/10/2011	29.17	25.74	38.90	25.23	30.29	0.997
	26/11/2011	28.17	32.33	42.88	67.81	30.92	0.984
	19/12/2011	13.81	18.96	23.46	40.82	19.18	0.994
	Average	13.82	33.48	38.07	17.13	28.35	0.930

Table 6. Continued.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
US_MOZ	28/06/2011	−9.39	35.77	36.98	34.11	26.10	0.943
	01/08/2011	−34.10	58.25	67.49	50.95	44.07	0.926
	18/08/2011	19.00	35.01	39.83	23.82	29.07	0.911
	31/08/2011	−5.01	61.27	61.48	36.50	45.51	0.954
	01/09/2011	−14.39	60.86	62.54	36.40	47.65	0.938
	07/09/2011	−20.00	83.89	86.24	38.73	70.20	0.847
	12/09/2011	−1.37	45.67	45.69	25.26	36.45	0.970
	30/09/2011	−16.75	79.20	80.95	44.61	62.64	0.899
	29/09/2011	31.91	47.11	56.91	41.40	40.83	0.964
	11/11/2011	12.38	39.64	41.52	45.15	35.47	0.745
	Average	1.24	57.63	57.64	42.44	42.44	0.910
US_IB1	30/05/2011	43.82	42.74	61.21	96.12	55.53	0.912
	07/06/2011	−26.18	35.35	43.99	32.53	35.86	0.938
	28/06/2011	−21.76	24.51	32.77	13.97	26.23	0.981
	08/07/2011	27.47	13.96	30.82	26.27	27.47	0.987
	24/08/2011	66.89	39.50	77.69	74.73	67.52	0.949
	13/09/2011	40.24	33.83	52.57	86.18	43.64	0.945
	15/09/2011	44.11	35.65	56.71	99.42	44.87	0.974
	01/10/2011	70.61	49.18	86.05	60.18	70.61	0.960
	15/10/2011	20.11	36.1	41.37	37.97	31.27	0.958
	24/10/2011	36.48	24.821	44.12	120.21	36.85	0.987
	Average	30.18	46.56	55.48	68.45	43.99	0.959
US_TON	27/02/2011	−31.49	54.12	62.62	47.89	48.24	0.974
	17/03/2011	−32.30	53.99	62.91	42.14	41.69	0.949
	24/05/2011	20.70	66.34	69.49	25.01	50.30	0.891
	24/06/2011	−29.63	48.44	56.79	18.84	38.08	0.963
	30/07/2011	−26.67	65.91	71.10	21.16	49.32	0.964
	07/08/2011	−33.82	59.47	68.42	20.66	51.35	0.985
	28/08/2011	1.24	58.79	58.80	19.55	44.20	0.961
	15/09/2011	18.72	47.12	50.70	21.14	36.56	0.979
	01/11/2011	43.03	29.34	52.08	68.88	45.21	0.894
	16/11/2011	26.49	28.39	38.82	43.20	28.90	0.979
	Average	−4.37	59.77	59.93	26.84	43.39	0.954
US_WHS	08/02/2011	−18.24	59.82	62.54	21.88	47.84	0.896
	16/02/2011	−32.83	49.03	59.01	30.47	46.02	0.921
	25/03/2011	−27.28	38.85	47.47	16.42	38.03	0.973
	22/06/2011	−43.74	88.41	98.64	34.20	62.97	0.954
	13/07/2011	11.17	38.21	39.81	13.40	26.23	0.970
	02/08/2011	66.41	49.29	82.71	53.07	66.83	0.931
	28/08/2011	68.22	63.93	93.49	50.47	70.74	0.929
	03/08/2011	18.90	36.66	41.24	17.56	30.47	0.974
	05/10/2011	77.51	66.79	102.31	48.15	77.81	0.969
	20/10/2011	36.28	40.16	54.12	29.51	41.09	0.997
	Average	17.47	67.73	69.94	30.07	48.97	0.951
	Average of all sites	8.66	52.62	55.06	28.40	40.14	0.951

Table 7. Daily simulation accuracy and average site simulation accuracy for $T_{\text{air 1.3 m}}$. Bias, scatter, RMSD and MAE are expressed in degrees Celsius. Nash index is unitless.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
Alice Springs	23/03/2011	−1.19	1.81	2.16	9.42	1.87	0.822
	15/04/2011	0.56	2.60	2.66	11.95	1.99	0.842
	23/04/2011	3.70	1.87	4.14	21.71	3.72	0.839
	10/05/2011	−0.09	2.75	2.75	17.22	2.52	0.871
	24/05/2011	2.97	3.48	4.58	30.80	3.06	0.850
	31/05/2011	−1.66	2.20	2.76	21.86	2.37	0.927
	18/06/2011	−0.07	2.41	2.41	17.78	2.15	0.911
	25/06/2011	−2.97	2.68	3.99	26.59	3.34	0.915
	18/07/2011	−1.25	1.92	2.29	14.21	2.08	0.911
	20/08/2011	−0.33	2.10	2.13	12.55	1.93	0.917
	Average	−0.03	3.11	3.11	18.34	2.50	0.881
Calperum	24/02/2011	−3.28	2.68	4.24	15.08	3.69	0.874
	02/03/2011	0.82	2.26	2.40	12.84	1.68	0.914
	31/03/2011	1.01	3.31	3.46	21.74	2.65	0.886
	24/04/2011	−0.45	3.47	3.50	21.99	3.21	0.903
	22/07/2011	−2.56	1.58	3.01	38.32	2.61	0.904
	28/07/2011	−3.21	2.76	4.24	30.76	3.51	0.867
	28/08/2011	−7.92	3.43	8.63	61.07	7.98	0.791
	01/12/2011	−3.30	1.50	3.63	18.09	3.30	0.785
	23/12/2011	−5.55	2.91	6.26	22.00	5.64	0.833
	29/12/2011	−4.45	1.77	4.79	18.18	4.45	0.835
	Average	−2.89	3.76	4.74	25.05	3.87	0.859
Howard Springs	18/04/2011	1.80	0.88	2.01	7.62	1.86	0.743
	23/04/2011	−0.03	0.78	0.78	2.71	0.68	0.915
	13/05/2011	0.39	1.59	1.64	7.20	1.26	0.923
	27/05/2011	2.14	2.01	2.93	12.70	2.60	0.813
	03/06/2011	2.11	1.98	2.89	12.40	2.70	0.826
	14/06/2011	1.27	2.41	2.72	14.25	2.47	0.794
	22/06/2011	−0.98	1.90	2.13	9.04	2.01	0.871
	22/07/2011	0.17	2.14	2.15	8.85	1.82	0.888
	28/07/2011	−1.38	1.74	2.22	8.61	2.08	0.851
	27/09/2011	0.07	1.10	1.10	3.88	0.95	0.910
	Average	0.56	2.10	2.17	8.83	1.84	0.853
US_VAR	10/05/2011	−3.70	2.79	4.64	25.05	3.91	0.862
	23/06/2011	1.37	2.61	2.94	11.44	1.94	0.939
	19/07/2011	−0.69	2.34	2.44	9.69	2.16	0.927
	30/07/2011	2.53	3.34	4.19	17.02	3.21	0.915
	07/08/2011	0.55	2.85	2.90	12.09	2.27	0.933
	27/08/2011	−0.79	2.80	2.90	10.38	2.63	0.926
	22/09/2011	−3.78	2.99	4.82	16.48	4.14	0.884
	07/10/2011	0.08	2.95	2.95	19.95	2.73	0.846
	26/11/2011	1.93	1.49	2.44	23.92	1.99	0.863
	19/12/2011	1.42	1.28	1.92	27.01	1.56	0.890
	Average	−0.11	3.34	3.35	16.14	2.66	0.898

Table 7. Continued.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
US_MOZ	28/06/2011	−0.70	0.75	1.03	4.28	0.97	0.821
	01/08/2011	1.67	1.04	1.97	7.04	1.68	0.909
	18/08/2011	−0.49	1.09	1.19	4.73	1.03	0.898
	31/08/2011	−0.97	1.21	1.55	5.05	1.23	0.903
	01/09/2011	3.87	2.58	4.65	14.55	3.87	0.631
	07/09/2011	1.14	1.67	2.02	10.73	1.45	0.890
	12/09/2011	1.73	0.91	1.96	7.72	1.73	0.883
	30/09/2011	0.70	2.03	2.14	12.43	1.79	0.830
	29/09/2011	−2.59	1.31	2.90	23.41	2.65	0.844
	11/11/2011	−1.70	2.12	2.72	21.14	2.45	0.924
	Average	0.23	2.37	2.38	10.52	1.84	0.853
US_IB1	30/05/2011	1.81	1.82	2.57	9.36	1.81	0.753
	07/06/2011	0.49	1.19	1.29	4.18	1.01	0.923
	28/06/2011	3.82	2.17	4.39	19.44	3.82	0.585
	08/07/2011	0.88	3.72	3.82	14.94	3.04	0.782
	24/08/2011	4.18	1.67	4.50	17.50	4.18	0.752
	13/09/2011	8.40	4.44	9.50	32.96	8.40	0.625
	15/09/2011	2.83	2.96	4.09	23.65	2.96	0.768
	01/10/2011	2.18	0.93	2.37	24.16	2.19	0.710
	15/10/2011	4.08	1.41	4.31	34.43	4.08	0.272
	24/10/2011	0.98	2.67	2.84	25.42	2.49	0.850
	Average	3.01	3.44	4.57	21.57	3.44	0.702
US_TON	27/02/2011	−1.68	0.94	1.93	25.67	1.71	0.833
	17/03/2011	−1.68	2.13	2.71	26.02	2.33	0.837
	24/05/2011	−0.69	1.34	1.51	9.06	1.18	0.922
	24/06/2011	1.51	1.36	2.03	8.59	1.79	0.906
	30/07/2011	1.47	2.03	2.51	10.34	1.86	0.923
	07/08/2011	3.11	2.78	4.18	17.63	3.11	0.875
	28/08/2011	2.08	2.42	3.20	14.78	2.12	0.919
	15/09/2011	4.26	3.15	5.30	24.52	4.29	0.788
	01/11/2011	1.27	2.14	2.49	14.90	2.27	0.873
	16/11/2011	0.39	0.96	1.03	7.08	0.82	0.919
	Average	1.00	2.77	2.94	16.30	2.15	0.880
US_WHS	08/02/2011	−1.32	1.92	2.33	7.79	2.05	0.901
	16/02/2011	0.79	1.89	2.05	11.97	1.79	0.869
	25/03/2011	−1.21	1.45	1.89	13.17	1.50	0.924
	22/06/2011	−0.56	2.59	2.66	8.27	2.07	0.880
	13/07/2011	2.26	2.24	3.18	11.24	2.98	0.745
	02/08/2011	0.55	1.37	1.48	4.98	1.17	0.907
	28/08/2011	0.65	1.35	1.50	5.11	1.20	0.940
	03/08/2011	2.76	4.31	5.12	17.83	4.27	0.739
	05/10/2011	0.56	1.23	1.35	6.61	1.10	0.934
	20/10/2011	−0.91	2.34	2.51	11.18	2.02	0.909
	Average	0.49	2.56	2.61	10.34	1.99	0.875
	Average of all sites	0.28	2.93	3.23	15.37	2.54	0.850

Table 8. Daily simulation accuracy and average site simulation accuracy for $T_{air\ 50\ m}$. Bias, scatter, RMSD and MAE are expressed in degrees Celsius. Nash index is unitless.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
Alice Springs	23/03/2011	−2.14	2.23	3.09	13.45	2.55	0.758
	15/04/2011	−0.05	3.10	3.10	13.91	2.71	0.785
	23/04/2011	3.49	2.91	4.54	23.82	3.49	0.849
	10/05/2011	−1.02	3.49	3.64	22.77	3.34	0.829
	24/05/2011	1.89	4.15	4.56	30.73	3.37	0.835
	31/05/2011	−2.59	3.05	4.00	31.72	3.32	0.898
	18/06/2011	−0.87	3.14	3.26	24.08	2.92	0.880
	25/06/2011	−3.61	3.41	4.97	33.05	3.96	0.899
	18/07/2011	−2.28	2.49	3.38	20.98	2.87	0.877
	20/08/2011	−1.28	3.01	3.27	19.24	2.95	0.872
	Average	−0.84	3.74	3.84	22.65	3.15	0.848
Calperum	24/02/2011	−4.35	3.88	5.83	20.74	4.91	0.833
	02/03/2011	0.15	3.03	3.03	16.23	2.58	0.868
	31/03/2011	0.78	4.36	4.43	27.79	3.77	0.837
	24/04/2011	−1.19	4.67	4.82	30.33	4.56	0.862
	22/07/2011	−2.09	2.81	3.50	44.57	2.73	0.900
	28/07/2011	−3.91	3.27	5.10	37.00	4.14	0.843
	28/08/2011	−8.46	4.52	9.59	67.82	8.76	0.771
	01/12/2011	−4.36	2.73	5.14	25.63	4.36	0.717
	23/12/2011	−6.68	3.54	7.56	26.56	6.78	0.800
	29/12/2011	−5.29	2.57	5.88	22.32	5.31	0.803
	Average	−3.54	4.57	5.78	30.56	4.79	0.823
Howard Springs	18/04/2011	0.85	1.20	1.47	5.58	1.07	0.852
	23/04/2011	−0.70	1.46	1.62	5.61	1.37	0.828
	13/05/2011	−0.52	1.57	1.66	7.29	1.47	0.910
	27/05/2011	2.14	1.19	2.44	10.57	2.15	0.845
	03/06/2011	1.92	1.07	2.19	9.40	1.92	0.876
	14/06/2011	0.82	1.07	1.35	7.05	1.20	0.900
	22/06/2011	−1.38	1.97	2.40	10.18	2.18	0.860
	22/07/2011	−0.39	2.24	2.27	9.38	1.93	0.881
	28/07/2011	−1.90	2.01	2.76	10.69	2.33	0.833
	27/09/2011	−0.30	1.65	1.68	5.93	1.44	0.863
	Average	0.05	2.04	2.04	8.30	1.71	0.865
US_VAR	10/05/2011	−4.69	3.78	6.02	32.55	5.17	0.818
	23/06/2011	0.64	3.98	4.03	15.67	3.19	0.899
	19/07/2011	−1.89	3.44	3.93	15.60	3.46	0.884
	30/07/2011	1.58	4.43	4.70	19.12	3.55	0.906
	07/08/2011	−0.43	4.00	4.03	16.78	3.42	0.898
	27/08/2011	−1.79	4.01	4.39	15.70	4.00	0.888
	22/09/2011	−4.33	4.06	5.94	20.32	4.89	0.863
	07/10/2011	−0.80	3.62	3.71	25.06	3.45	0.805
	26/11/2011	1.66	2.41	2.92	28.69	2.45	0.831
	19/12/2011	1.16	1.89	2.22	31.27	1.88	0.867
	Average	−0.89	4.24	4.34	20.92	3.55	0.866

Table 8. Continued.

Location	Date	Bias	Scatter	RMSD	RMSD (%) of observed	MAE	Nash index
US_MOZ	28/06/2011	−1.44	1.26	1.91	7.97	1.772	0.674
	01/08/2011	1.38	1.69	2.18	7.79	1.677	0.910
	18/08/2011	−1.44	1.70	2.22	8.81	1.83	0.819
	31/08/2011	−1.78	1.86	2.58	8.40	2.02	0.842
	01/09/2011	3.49	3.43	4.89	15.29	3.62	0.655
	07/09/2011	0.23	2.35	2.37	12.54	2.07	0.843
	12/09/2011	1.09	1.81	2.11	8.33	1.59	0.893
	30/09/2011	0.12	2.82	2.82	16.35	2.50	0.762
	29/09/2011	−3.44	1.58	3.79	30.60	3.44	0.798
	11/11/2011	−1.96	1.75	2.63	20.50	2.14	0.934
	Average	−0.46	2.81	2.85	12.56	2.22	0.813
US_IB1	30/05/2011	1.23	2.41	2.71	9.87	1.83	0.750
	07/06/2011	0.43	2.35	2.39	7.75	2.09	0.840
	28/06/2011	3.08	3.14	4.40	19.47	3.12	0.661
	08/07/2011	−0.19	4.09	4.10	16.03	3.61	0.741
	24/08/2011	4.36	3.29	5.46	21.23	4.36	0.741
	13/09/2011	8.20	5.50	9.88	34.27	8.20	0.491
	15/09/2011	1.86	3.84	4.26	23.98	3.32	0.740
	01/10/2011	1.76	1.50	2.31	23.63	1.76	0.767
	15/10/2011	4.10	2.34	4.73	37.73	4.10	0.267
	24/10/2011	0.33	3.17	3.19	27.71	2.84	0.829
	Average	2.52	4.11	4.82	22.69	3.52	0.683
US_TON	27/02/2011	−2.08	1.44	2.53	33.73	2.08	0.797
	17/03/2011	−1.98	2.84	3.46	33.20	2.93	0.795
	24/05/2011	−1.41	2.13	2.55	15.30	2.37	0.844
	24/06/2011	0.81	2.51	2.64	11.17	1.96	0.897
	30/07/2011	0.60	3.14	3.19	13.17	2.52	0.895
	07/08/2011	2.45	4.01	4.70	19.85	3.04	0.878
	28/08/2011	1.17	3.62	3.80	17.59	2.92	0.889
	15/09/2011	3.41	4.21	5.42	25.07	3.63	0.821
	01/11/2011	0.53	2.69	2.74	16.40	2.51	0.859
	16/11/2011	−0.13	1.57	1.58	10.84	1.49	0.853
	Average	0.34	3.42	3.43	19.02	2.55	0.853
US_WHS	08/02/2011	−1.43	2.64	3.00	10.03	2.65	0.872
	16/02/2011	1.15	2.02	2.32	13.55	1.79	0.870
	25/03/2011	−1.61	2.54	3.01	21.01	2.52	0.873
	22/06/2011	−1.00	3.04	3.20	9.97	2.81	0.838
	13/07/2011	1.25	2.59	2.88	10.18	2.21	0.811
	02/08/2011	−0.37	2.15	2.18	7.34	2.01	0.841
	28/08/2011	−0.32	2.10	2.13	7.26	1.94	0.903
	03/08/2011	1.84	4.70	5.05	17.59	4.16	0.746
	05/10/2011	−0.67	2.04	2.15	10.54	1.93	0.884
	20/10/2011	−1.43	3.13	3.44	15.34	3.02	0.864
	Average	−0.19	3.03	3.04	12.03	2.51	0.850
	Average of all sites	−0.376	3.496	3.766	17.90	3.00	0.825

RMSD ranged from 24.14 to 53.78 W m^{-2} (or within ∼ 6 and ∼ 21 % of the observed fluxes) for all the test days located within the period from 24 February to 24 April 2011. In contrast, for the same site, RMSD varied from 84.41 to 149.29 W m^{-2} (or within ∼ 41 and ∼ 53 % of the observed fluxes) for all the test days for the period between 22 July and 29 December 2011. Similar trends were observed for all other Australian sites, although some anomalies were present. In relation to the US sites, the adverse was found: the highest simulation accuracies were predominantly derived for the test days during the period between October and late April. Clearly, the periods of highest simulation accuracy for both the Australian and US sites correspond to their respective summer season, and are thus consistent between the two countries. Generally, the results for the US sites suggested that the conditions prevalent within the wet season (October–May) may have had an influence on model accuracy.

5.2 Net radiation (R_{net}) at the surface

Table 4 and Fig. 2 indicate a high overall performance in the model's ability to accurately predict R_{net}, confirmed by the high simulation accuracy (RMSD = 58.69 W m^{-2}, within ∼ 24 % of the observed fluxes, MAE = 46.42 W m^{-2}) reported for all sites. Furthermore, comparisons of R_{net} for all days of simulation showed a low average MSD of 54.44 W m^{-2}, indicating the model's capability to precisely represent the amplitude of the R_{net} flux, with low dispersion of variance from the in situ trends, as evidenced in Fig. 2. MBE results indicated a moderate underestimation of the in situ measurements by the model (−16.49 W m^{-2}), with seven of the eight site averages showing an underestimation of the in situ trends (negative MBE values in a range of −0.09 to −46.10 W m^{-2}). A much larger intersite variability was reported for the model simulation accuracies of the R_{net} parameter, where RMSD ranged between 33.90 and 78.03 W m^{-2} (also reflected in the RMSD as a percentage of observed fluxes ranging between ∼ 16 and ∼ 30 % on a per site average basis) showing to some extent a deficiency in the capability of the model to capture the land surface process over varying land cover types. The R_{net} results exhibited largely similar statistical agreement with those observed for the R_g parameter.

Most noticeably, in correspondence with the R_g parameter results, SimSphere showed superior simulation accuracy within the Alice Springs mulga woodland site in comparison to the other land cover types, with the reported accuracies significantly above the overall average (RMSD = 33.90 W m^{-2}, within ∼ 16 % of the observed fluxes, MAE = 26.25 W m^{-2}). Moreover, the woody savannah site of Howard Springs again exhibited high simulation accuracies (RMSD = 47.05 W m^{-2}, within ∼ 21 % of the observed fluxes, MAE = 35.74 W m^{-2}), with comparable accuracies to the simulation of the R_g parameter. Conversely, the model showed an inferior performance when simulating R_{net} within the US_TON wooded savannah site where a systematic and more pronounced underestimation of R_{net} was evident (MBE = −46.10 W m^{-2}). This constant underestimation by the model led to a poorer agreement between the model predictions and in situ observations for the US_TON site, as reflected in the statistical analysis (RMSD = 78.03 W m^{-2}, within ∼ 30 % of the observed fluxes, MAE = 65.22 W m^{-2}). It should be noted that the accuracy of the model estimations on a per site basis did not correlate between both the R_g and R_{net} parameter estimations, with only the US_WHS shrubland site exhibiting weaker simulation accuracies for both parameters and, as indicated above, a relatively high simulation accuracy for the Howard Springs woody savannah site.

Evidently, as indicated in Table 4, trends in simulation accuracy dependent on test day were apparent. Although comparable; the trends were not as prominent as those exhibited for the R_{net} parameter. Within the Australian sites, low RMSD was exhibited predominantly for the test days within the period from March to July, although some discrepancies were present during specific days. For example, the 27 May simulation date for the Howard Springs site reported an RMSD of 70.60 W m^{-2} (within ∼ 38 % of the observed fluxes), indicating a day of unusually high error for this period. However, such anomalies were limited. Generally, for the US sites, highest RMSD was exhibited for the period concurrent to the wet season (October–April), with the highest error rates exhibited during the dry period, for example, during the 27 February simulation date for the US_TON site (RMSD = 113.80 W m^{-2}, within ∼ 73 % of the observed fluxes), although, again, the anomalies in such trends were notable yet uncommon.

5.3 Latent heat (LE)

As presented in Table 5, the highest RMSD in relation to the observed fluxes was reported for the LE parameter in comparison to all other parameters evaluated (RMSD = 39.47 W m^{-2}), where SimSphere showed some deficiencies when reproducing LE fluxes in varying land cover, both in terms of its seasonal and diurnal evolution. An average R^2 value of 0.700 is also indicative of a poorer correlation between the predictions and observations of LE (Fig. 2). When averaged over all days and sites, the model-based estimates tended towards a conservative overestimation of the observed fluxes, indicated by an average MBE of 2.84 W m^{-2}.

On a site by site basis, the US_IB1 cropland site consistently yielded the highest statistical agreement between model-predicted and observed values, with low error and high correlation results (RMSD = 52.54 W m^{-2}, within 20 % of the observed fluxes, MAE = 15.16 W m^{-2}, $R^2 = 0.827$, Nash = 0.945). Notably, all other sites exhibited poorer agreement, with RMSD values in relation to the observed fluxes above 30 % for six of the eight sites (RMSD varying within ∼ 34 and 83 % of the observed fluxes). Generally, each site exhibited a significant range of MBE, from $-11.49\,\mathrm{W\,m^{-2}}$ (US_WHS) to $25.65\,\mathrm{W\,m^{-2}}$ (US_MOZ), suggesting high variability between the partitioning of LE in each ecosystem. Peak LE flux values exhibited high intersite variability, with both the US_IB1 (cropland) and US_MOZ (deciduous broadleaf forest) sites containing the highest LE flux peaks of 458.5 and $376\,\mathrm{W\,m^{-2}}$, respectively. In comparison, a maximum LE flux peak of just $143.7\,\mathrm{W\,m^{-2}}$ was reported for the US_WHS (shrubland) site, suggesting a substantial range of $314.8\,\mathrm{W\,m^{-2}}$ between the lowest daily peak LE and maximum daily peak LE. Noticeably, trends in simulation accuracy dependent on test day were comparable to both the R_g and R_{net} parameter results, however, with significantly higher intersite variability in RMSD ranges.

5.4 Sensible heat (H)

SimSphere showed a satisfactory ability to accurately simulate H fluxes in numerous ecosystems for the 72 days included in this study, with average RMSD and R^2 values of $55.06\,\mathrm{W\,m^{-2}}$, within ∼ 28 % of the observed fluxes, and 0.829, respectively. Results were largely similar to those of the LE flux simulation accuracies, although the model performance for the LE parameter underperformed compared that of the H flux for the majority of statistical metrics computed herein.

Average RMSD values ranged from 38.07 to $69.94\,\mathrm{W\,m^2}$ (US_VAR and US_WHS) and within ∼ 17 and ∼ 68 % of the observed fluxes (US_VAR and US_IB1) when analysed on a site by site basis, underlining the greatest intersite variability was reported for this parameter. In addition, R^2 values ranged from 0.73 (US_IB1) to 0.94 (US_VAR). The latter was suggestive that model predictions were generally in good agreement with the in situ measurements, showing a strong relationship between both variables. The grassland site (US_VAR) consistently showed superior model performance in comparison to all other sites, with values indicating an excellent agreement with the observed diurnal evolution (RMSD $= 38.07\,\mathrm{W\,m^{-2}}$, within ∼ 17 % of the observed fluxes, MAE $= 28.35\,\mathrm{W\,m^{-2}}$). MSD values reported for US_VAR were $19.41\,\mathrm{W\,m^{-2}}$ lower than the all site average, suggesting a systematically accurate representation of H fluxes at this site. MSDs for H flux were directly comparable to the overall average MSD values reported for R_g and R_{net}, yet significantly higher than the LE fluxes. Simulation accuracy was comparably high for the simulated H fluxes for five of the eight sites, with RMSD values in relation to the observed fluxes above 30 % (RMSD within percentage of the observed fluxes varying between ∼ 17 and 30 %). Notably, results for the US_IB1 site exhibited significant error, with RMSD and MSD values of $69.94\,\mathrm{W\,m^{-2}}$, within ∼ 68 % of the observed fluxes, and $67.73\,\mathrm{W\,m^{-2}}$, respectively.

For the Australian sites no significant trends were evident dependent on simulation day, with generally comparable accuracy ranges for the specific test days including anomalistic days which exhibited significantly higher error ranges. For example, the Howard Springs woody savannah site indicated an RMSD for the majority of simulation days ranging between 28.29 and $50.31\,\mathrm{W\,m^{-2}}$ (within ∼ 15 and ∼ 21 % of the observed fluxes) on a per test day basis, with the 18 April and 13 May experimental days exhibiting RMSDs of 75.86 and $96.93\,\mathrm{W\,m^{-2}}$ (within ∼ 52 and ∼ 65 %), respectively. Similar intra-site variability was notable for the US sites.

5.5 Air temperature 1.3 m ($T_{air,\,1.3\,m}$)

SimSphere showed a high capability for simulating $T_{air,\,1.3\,m}$ with an average RMSD as low as 3.23 °C (within ∼ 15 % of the observed) and relatively high R^2 value of 0.843, see Table 7. Furthermore, $T_{air,\,1.3\,m}$ exhibited neither a consistent over- or underestimation, with an overall average MBE of 0.28 °C. Simulation accuracy for $T_{air,\,1.3\,m}$ was relatively stable, with a low range of RMSD values reported over all sites. RMSD values ranged from 2.17 °C (within ∼ 9 % of the observed) in the woodland savannah site of Howard Springs to 4.74 °C (within ∼ 25 % of the observed) in the grazing pasture site of Calperum. Overall, agreement between the predictions and observations was greatest for the Howard Springs site, with results confirming a high overall correlation to the observed diurnal evolution of $T_{air,\,1.3\,m}$. The deciduous broadleaf site of US_MOZ also exhibited a comparably high simulation accuracy (RMSD = 2.38 °C, within ∼ 11 % of the observed, MAE = 1.84 °C, Nash = 0.853). The Calperum site exhibited the weakest agreement of $T_{air,\,1.3\,m}$ with an average RMSD 1.51 °C higher than the all-site average. The R^2 analysis further appraised the model's ability to accurately simulate air temperature, with a range of values indicating high correlation between model-predicted and observed $T_{air,\,1.3\,m}$ (0.74–0.93). MSD displayed a high range of values (2.1–3.76 °C), showing to some extent the inability of the model to consistently predict $T_{air,\,1.3\,m}$ with a high level of precision. The trends in simulation accuracy dependent on test day were again insignificant for the $T_{air\,1.3\,m}$ parameter, exhibiting similar patterns to those indicated for the H flux parameter.

5.6 Air Temperature 50 m ($T_{air\,50\,m}$)

The model showed a slightly inferior performance in predicting $T_{air\,50\,m}$ (RMSD = 3.77 °C, within ∼ 18 % of the observed) when compared to $T_{air\,1.3\,m}$, with an average RMSD difference of 0.54 °C (∼ 3 % percentage difference in relation to the observed) (Table 8, Fig. 2). A lower average R^2 value of 0.775 is reported compared to that of $T_{air,\,1.3\,m}$ (R^2 = 0.843), indicating a weaker, yet close, agreement between both variables. However, the values reported still showed a highly acceptable correlation between the modelled estimates and the in situ measurements, as indicated by an average Nash value of 0.825. Once averaged, $T_{air\ 50\ m}$ exhibited a minor underestimation of −0.38 °C; however, the range of MBEs reported between sites was significantly less (2.1 °C), suggesting a more consistent simulation of T_{air} at 50 m compared to at 1.3 m by SimSphere. In contrast, agreement between the simulated $T_{air\ 50\ m}$ and in situ measurements resulted in a higher MSD than that reported for the $T_{air,\ 1.3\ m}$ parameter, with the exception of the Howard Springs site. When analysed on a per site basis, notably, in correspondence with the $T_{air,\ 1.3\ m}$ parameter, agreement between the estimated and measured values over both the Howard Springs and US_MOZ sites exhibited the highest simulation accuracy (RMSD = 2.04 and 2.85 °C, within ∼ 8 and ∼ 13 % of the observed, respectively). Moreover, the weakest agreement was reported over the Calperum site, again, in correspondence with the results of the $T_{air,\ 1.3\ m}$ parameter. No systematic trends were apparent in the intersite variability of simulation accuracy dependent on test day.

6 Discussion

The present study evaluated the ability of the SimSphere SVAT model to accurately represent key parameters characterising land surface interactions within eight ecosystems in two continents. A total of 72 days (10 days per site of the eight sites selected) from year 2011 were selected from Australia and the USA to validate the model's ability to predict R_g, R_{net}, LE, H, and T_{air} at a height of 1.3 and 50 m.

Variable model performance was clearly evident when simulating both the LE and H fluxes within contrasting land cover types. For example, as discussed, the highest simulation accuracy was attained within the grassland study sites. In contrast, simulation accuracy within forested ecosystems was less satisfactory. The deciduous forest stand (US_MOZ), with an average canopy height of 24.2 m attained significantly low simulation accuracy and was also outperformed by the mulga forested ecosystem (Alice Springs), characterised by a sparse canopy at a height of 6.5 m. Such results suggest that the increased complexity and heterogeneity of forested environments, particularly those with understory vegetation, can have profound effects on the overall exchange of mass and energy which cannot be represented within the model's parameterisation and hence can influence LE and H outputs. The partitioning of LE and H fluxes are also highly susceptible to a number of other factors. Small changes in the moisture availability, particularly from the deep layer soil water content (SWC), can have a strong influence (Carlson and Lynn, 1991), on the representativeness of the radiosonde data to the existent local conditions (Taconet et al., 1986). As reported by Taconet et al. (1986), an error of just ∼ 2 °C in the sounding profile temperature can cause a variation of ∼ 45 W m^{-2} in the corresponding fluxes, particularly for H flux. SimSphere was forced with surface and root zone moisture availability taken directly from the in situ data sets. These highly influential parameters were consistently misrepresented within the model's parameterisation, providing a possible reason, in part, for the lower simulation accuracies attained.

R_g was estimated by the model to a high level of accuracy (error within ∼ 19 % of the observed fluxes), where an R^2 value of 0.971 and a Nash value of 0.960 reported for all days of analysis suggest that model predictions had excellent correlation to the observed data set. This indicates that SimSphere was able to simulate the trend of R_g well. A possible reason for the underestimation of R_g by the model is perhaps linked to the solar transmission model and/or the surface albedo calculation in the model, as has also been pointed out previously by Todhunter and Terjung (1978). Furthermore, previous sensitivity analysis studies undertaken upon the model confirm that R_g is significantly influenced by a site's aspect (Petropoulos et al., 2014). Therefore, simulation accuracy may partly be related to the model's representation of a site's topographical characteristics.

In the majority of the experimental sites a general underestimation of R_{net} was attained by the model, which led to mean RMSD and R^2 values of 58.69 W m^{-2} and 0.960, respectively. These results are also comparable to those reported in other analogous validation studies (Carlson and Boland, 1978; Todhunter and Terjung, 1987; Ross and Oke, 1988). Todhunter and Terjung (1987) compared predicted R_{net} from the model versus corresponding R_{net} values obtained from the literature for Los Angeles, USA, and showed both daytime and night time simulations to be in agreement within the range reported in the literature. Ross and Oke (1988) also confirmed the capability of the model for simulating the day-to-day variation of R_{net} for comparisons using 18 cloud-free days over an urban area of Vancouver, B.C., Canada. Ross and Oke (1988) reported an overall average RMSD error of 43 W m^{-2} for comparisons for all cloud-free days, a minor improvement on the RMSD of 58.69 W m^{-2} presented herein. Disparity in the results between this study and those studies could be the result of utilising model simulations over dissimilar land cover types, where it is largely accepted that R_{net} partitioning into LE and H fluxes is highly dependable on the vegetation and surface characteristics of the site (Olioso et al., 2000). Previous sensitivity analysis studies undertaken on the SimSphere further confirm this observation (Petropoulos et al., 2014). Similarly to R_g, simulation accuracy of R_{net} was described by Ross and Oke (1988) to be a factor of long-wave radiation, mainly the values of atmospheric and surface emissivities (which affect the surface temperature estimation). Increased representation of the surface optical properties and long-wave radiation estimation of the model could greatly enhance simulation accuracy.

Overall, simulation accuracies were lower for estimates of $T_{air\ 50\ m}$ compared to estimates of $T_{air,\ 1.3\ m}$ in all but one site, Howard Springs. One possible explanation for this may be the fundamental problem that model estimates of $T_{air\ 50\ m}$

could only be validated against ancillary air temperature data obtained directly from the site's flux tower; thus, direct comparison specifically at 50 m could not be achieved. Similarly to the LE and H fluxes, variable simulation accuracies dependent on land cover types were also evident. Three sites – Calperum, US_VAR and US_IB1 – exhibit noticeably weaker simulation accuracies in comparison to the remaining sites. Upon further investigation, all three sites showed ecosystems which are characterised by high interannual variability of vegetation phenology, such as vegetation height, leaf width, FVC etc. Modelled T_{air} peaked between 10:30 and 14:30 LT. For instances where a time lag between the predicted and observed T_{air} comparisons is observed, such effects may be linked with the energy storage in the vegetation and the air, as it is not taken into account in the SimSphere simulations. This may partly explain some of the inaccuracies reported for T_{air} estimation in Alice Springs and US_MOZ, as this effect is more important for forested sites. Carlson and Boland (1978) and Carlson et al. (1991) also described a similar hysteresis effect in comparisons which they performed for different vegetation canopies and environmental conditions (urban and rural environments). Carlson and Boland (1978) suggested thermal inertia to be related proportionally to an increase in the time lag between solar noon and the time of maximum H flux and T_{s}, whereas Carlson et al. (1991) admitted that they were unable to practically explain this "hysteresis" trend. Through comprehensive sensitivity analysis studies (Petropoulos et al., 2009b, 2013a, 2014), parameters closely associated with vegetation phenology have been previously outlined to have a highly influential control on air temperature magnitude and extent. Conversely, sites which show relatively stable vegetation phenology such as US_TON (wooded savannah) exhibited more accurate temperature estimates. Furthermore, the air temperature of the site covered by the dead forest had greater daily fluctuation compared to the stands covered by mature forest which generally had the smallest daily fluctuations. However, more studies are required in this direction to categorise the dead forest from mature forest, currently which is not possible in the given land cover database. Improved land cover information can provide more insights into the performances during the validation. As SimSphere assumes a homogenous canopy layer, some discrepancies may occur in the air temperature simulation, which is also the case over here. Furthermore, a very important point to also consider in the overall interpretation of the results is that the model does not account for advective conditions which might be important for instance when strong winds exist. Yet, generally, air temperature at 1.3 and 50 m were well represented by the model with the results obtained showing a significant improvement on values reported in previous validation attempts (Carlson and Boland, 1978; Carlson et al., 1991).

All in all, SimSphere demonstrated a high capability for simulating parameters associated with Earth's energy balance. It is also apparent that the model fulfils three of the Kramer et al. (2002) model assessment criteria, namely *accuracy, generality and realism* (see also Sect. 1) In regards to accuracy, no significant systematic prediction errors occurred within all of the fluxes analysed, with the exception of a consistent underestimation of R_{g} and R_{net}. Additionally, simulated peak heat and water flux values were in high accordance with the in situ data, typically at 12:30–13:30 LT, with a slight lag for LE and H fluxes (13:00–14:00 LT). In terms of *generality*, the model has shown high levels, with acceptable simulation accuracies attained in the majority of sites validated. In order to improve the model's generality, the inclusion of more forested environments would comprehensively assess the model's applicability to different land cover types, particularly heterogeneous forest stands where simulation accuracy tends to be lower. Finally, *realism* in the model has been most notable in the simulation of LE, H and T_{air} fluxes, where a slight change in the vegetation phenology or SWC was accountable for characterising the diurnal evolution of fluxes in all sites validated.

This study can advance our understanding on SimSphere's capability to simulate the interactions between different components of our Earth system and related land surface processes. As no model is perfect, some discrepancies between predictions and measurements will always appear. Identification of these discrepancies are most interesting, because they can teach us more about causes of model uncertainties in the prediction of hydro-meteorological variables and help us to improve the model structure and performance. Some large discrepancies between the simulated and observed data sets could be due to model parameterisation. Apart from environmental factors, some instrumentation errors in the tower flux measurements, indicated by the presence of many spikes (too large or too small values), can also affect the accuracy, even if model-simulated results are in agreement with actual conditions. The other possible reasons is the presence of spikes in the fluxes, observed particularly on the days of low agreement, which could occur because of horizontal advection, footprint changes and non-stationarity of turbulent regimes (Papale et al., 2006). Unfortunately, such conditions cannot be captured and replicated by SimSphere.

Overall, it is important to recognise that uncertainty is inevitable in any model and that a model will never be as complex as the reality it portrays. In this way the model fulfills its objective as a tool that identifies the patterns of change expected, if not always the magnitudes, indicating its usefulness in practical applications either as a stand-alone tool or in combination with remote sensing data as done for instance through the implementation of the "triangle" technique.

7 Conclusions

This study evaluated the ability of the SimSphere land biosphere model in predicting a number of parameters characterising land surface interactions for eight sites from the global terrestrial monitoring network, FLUXNET. A rigorous comparison was performed for 72 selected days from the year 2011. The main findings of this study are concluded as follows.

Overall, SimSphere estimates of instantaneous energy fluxes and air temperature showed good agreement in all ecosystems evaluated, apart from a minor underestimation of R_g and R_{net} (MBE $= -19.48$ and $-16.49\,W\,m^{-2}$, respectively). Some ecosystems exhibited poorer simulation accuracies than others, most noticeably cropland (US_IB1) and grazing pasture (Calperum); whilst the woodland savannah (Howard Springs) and mulga woodland (Alice Springs) ecosystems both attained the highest overall simulation accuracies. Comparisons showed a good agreement between modelled and measured fluxes, especially for the days with smoothed daily flux trends. Very high values of the Nash–Sutcliffe efficiency index were also reported for all parameters, ranging from 0.720 to 0.998, suggesting, overall, a very good model representation of the observations. The highest simulation accuracies were obtained for the open woodland savannah and mulga woodland sites for most of the compared parameters.

The process of validating any physical model is imperative to understanding its representation of real-world scenarios. It helps in identifying any deficiencies in the models' predictive ability and to identify any possible sources of error and uncertainty associated with a model. To our knowledge, very few studies, if any, have focused specifically on validating SimSphere to numerous ecosystems in the USA and Australia. On this basis, with the use of this model as either a stand alone research or educational tool, or for its synergy with EO data, its validation is not only timely but essential. SimSphere, despite its inherent architectural limitations, can be applied in future for solving various theoretical and applied tasks. There is certainly room for further improvements to the model to develop it further in terms of its representation of the various physical processes characterising land surface interactions. This is a promising research direction on which model development efforts should be focused in future.

Acknowledgements. This research was conducted with the support of the European Commission under the Marie Curie Reintegration Grant TRANSFORM-EO project. the Authors gratefully acknowledge the funding body for their financial support and also extend a thank you to the PI's from the AmeriFlux and OzFlux networks for the data provision which made this study possible.

Edited by: M.-H. Lo

References

Abramowitz, G., Leuning, R., Clark, M., and Pitman, A.: Evaluating the performance of land surface models, J. Climate, 21, 5468–5481, 2008.

Alexandris, S. and Kerkides, P.: New empirical formula for hourly estimations of reference evapotranspiration, Agr. Water Manag., 60, 157–180, 2003.

Aubinet, M., Grelle A., Ibrom, A., Rannik, Ü., Moncrieff, J., Foken, T., Kowalski, A. S., Martin, P. H., Berbigier, P., Bernhofer, Ch., Clement, R., Elbers, J., Granier, A., Grünwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the Annual Net Carbon and Water Exchange of Forests: The EUROFLUX Methodology, Adv. Ecol. Res., 30, 113–175, 2000.

Barr, A. G., Morgenstern, K., Black, T. A., McCaughey, J. H., and Nesic, Z.: Surface energy balance closure by the eddy covariance method above three boreal forest stands and implications for the measurement of the CO_2 flux, Agr. Forest Meteorol., 140, 322–337, 2006.

Battrick, B. and Herland, E. A.: The changing Earth. New scientific challenges for ESA's Living Planet Programme, ESA SP-1304, ESA, Publications Division, ESTC, The Netherlands, 2006.

Bellocchi, G., Rivington, M., Donatelli, M., and Matthews, K.: Validation of biophysical models: issues and methodologies, A review, Agron. Sustain. Dev., 30, 109–130, 2010.

Carlson, T. N.: An overview of the "triangle method" for estimating surface evapotranspiration and soil moisture from satellite imagery, Sensors, 7, 1612–1629, 2007.

Carlson, T. N. and Boland, F. E.: Analysis of urban-rural canopy using a surface heat flux/temperature model, J. Appl. Meteorol., 17, 998–1013, 1978.

Carlson, T. N. and Lynn, B.: The effects of plant water storage on transpiration and radiometric surface temperature, Agr. Forest Meteorol., 57, 171–186, 1991.

Carlson, T. N., Dodd, J. K., Benjamin, S. G., and Cooper, J. N.: Satellite estimation of the surface energy balance, moisture availability and thermal inertia, J. Appl. Meteorol., 20, 6–87, 1981.

Castellvi, F., Martinez-Cob, A., and Perez-Coveta, O.: Estimating sensible and latent heat fluxes over rice using surface renewal, Agr. Forest Meteorol., 139, 164–169, 2006.

Chauhan, N. S., Miller, S., and Ardanuy, P.: Spaceborne soil moisture estimation at high resolution: a microwave-optical/IR synergistic approach, Int. J. Remote Sens., 22, 4599–4646, 2003.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic-properties, Water Resour. Res., 14, 601–604, 1978.

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T.: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, Water Resour. Res, 20, 682–690, 1984.

Culf, A. D., Folken, T., and Gash, J. H. C.: The energy balance closure problem, In: Vegetation, Water, Humans and the Climate, Berlin, Springer-Verlag, 2002.

Coudert, B., Ottlé, C., and Briottet, X.: Monitoring land surface processes with thermal infrared data: Calibration of SVAT parameters based on the optimisation of diurnal surface temperature cycling features, Remote Sens. Environ., 112, 872–887, 2008.

Deardorff, J. W.: Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation, J. Geophys. Res.-Ocean., 83, 1889–1903, 1978.

Dickinson, R. E. and Henderson-Sellers, A.: Modelling tropical deforestation: A study of GCM land-surface parametrizations, Q. J. Roy. Meteor. Soc., 114, 439–462, 1988.

European Space Agency: Support to Science Element: A pathfinder for innovation in Earth Observation, ESA, available at: http://due.esrin.esa.int/stse/files/document/STSE_report_121016.pdf (last access: 10 July 2013), 2012.

Farouki, O. T.: The thermal properties of soils in cold regions, Cold Regions Sci. Tech., 5, 67–75, 1981.

Gillies, R. R.: A physically-based land sue classification scheme using remote solar and thermal infrared measurements suitable for describing urbanisation, PhD Thesis, University of Newcastle, UK, 121 pp., 1993.

Gillies, R. R., Kustas, W. P., and Humes, K. S.: A verification of the 'triangle' method for obtaining surface soil water content and energy fluxes from remote measurements of the Normalized Difference Vegetation Index (NDVI) and surface, Int. J. Remote Sens., 18, 3145–3166, 1997.

Granz, D., Zhang, X., and Carlson, T. N.: Observations and model simulations link stomatal inhabitation to impaired hydraulic conductance following ozone exposure in cotton, Plant Cell Environ., 22, 1201–1210, 1999.

Henderson-Sellers, A., Pitman, A. J., Love, P. K., Irannejad, P., and Chen, T. H.: The project for intercomparison of land surface parameterization schemes (PILPS): Phases 2 and 3, Bull. Am. Meteorol. Soc., 76, 489–503, 1995.

Hsu, M. H., Kuo, A. Y., Kuo, J. T., and Liu, W. C.: Procedure to calibrate and verify numerical models of estuarine hydrodynamics, J. Hydraul. Eng., 125, 166–182, 1999.

Huth, N. and Holzworth, D.: Common sense in model testing, in: Proc. MODSIM 2005 International Congress on Modelling and Simulation: Advances and applications for management and decision making, edited by: Zerger, A. and Argent, R. M., 12–15 December, Melbourne, Australia, 2804–2809, 2005.

Kramer, K., leinonen, I., Bartelink, H., Berbigier, P., Borgnetti, M., Bernhofer, C., Cienciala, E., Dolman, A. J., Froer, O., Gracia, A., Granier, A., Grunwald, T., Hari, P., Jans, W., Kellomaki, S., Loustau,, D., Magnani, F., Markkanen, T., Matteucci, G., Mohren, G. M., Moors, E., Nissenen, A., Peltola, H., Sabate, S., Sanchez, A., Sontag, M. Valentini, R., and Vesala, T.: Evaluation of six-process-based forest growth models using eddy-covariance measurements of CO2 and H2O fluxes at six forest sites in Europe, Global Change Biol., 8, 213–230, 2002.

Liang, X., Wood, E. F., Lettenmaier, D. P., Lohmann, D., Boone, A., Chang, S., and Zeng, Q. C.: The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) phase 2 (c) Red-Arkansas River basin experiment: 2. Spatial and temporal analysis of energy fluxes, Global Planet. Chang., 19, 137–159, 1998.

Liu, Y., Hiyama, T., and Yamaguchi, Y.: Scaling of land surface temperature using satellite data: A case examination on ASTER and MODIS products over a heterogeneous terrain area, Remote Sens. Environ., 105, 115–128, 2006.

Lynn, B. H. and Carlson, T. N.: A stomatal resistance model illustrating plant vs. external control of transpiration, Agr. Forest Meteorol., 52, 5–43, 1990.

Maayar, M., Price, D. T., Delire, C., Foley, J. A., Black, T. A., and Bessemoulin, P.: Validation of the Integrated Biosphere Simulator over Canadian deciduous and coniferous boreal forest stands, J. Geophys. Res-Atmos., 106, 14339–14355, 2001.

Manabe, S.: Climate and the ocean circulation 1. The atmospheric circulation and the hydrology of the earth's surface, Mon. Weather. Rev., 97, 739–774, 1969.

Marshall, M., Tu, K., Funk, C., Michaelsen, J., Williams, P., Williams, C., Ardö, J., Boucher, M., Cappelaere, B., de Grandcourt, A., Nickless, A., Nouvellon, Y., Scholes, R., and Kutsch, W.: Improving operational land surface model canopy evapotranspiration in Africa using a direct remote sensing approach, Hydrol. Earth Syst. Sci., 17, 1079–1091, doi:10.5194/hess-17-1079-2013, 2013.

Mascart, P., Taconet, O., Pinty, J. P., and Mehrez, M. B.: Canopy resistance formulation and its effect in mesoscale models: a HAPEX perspective, Agr. Forest Meteorol., 54, 319–351, 1991.

Monin, A. S. and Obukhov, A.: Basic laws of turbulent mixing in the surface layer of the atmosphere, Contrib. Geophys. Inst. Acad. Sci. USSR, 151, 163–187, 1954.

Nash, J. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I - A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Olchev, A., Ibrom, A., Ross, T., Falk, U., Rakkibu, G., Radler, K., Grotea, S., Kreileina, H., and Gravenhorst, G.: A modelling approach for simulation of water and carbon dioxide exchange between multi-species tropical rain forest and the atmosphere, Ecol. Model, 212, 122–130, 2008.

Olioso, A., Carlson, T. N., and Brisson, N.: Simulation of diurnal transpiration and photosynthesis of a water stressed soybean crop, Agr. Forest Meteorol., 81, 41–59, 1996.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, Biogeosciences, 3, 571–583, doi:10.5194/bg-3-571-2006, 2006.

Petropoulos, G., Carlson, T., and Wooster, M. J.: An Overview of the Use of the SimSphere Soil vegetation Atmospheric Transfer (SVAT) Model for the Study of Land Atmosphere Interactions, Sensors, 9, 4286–4308, 2009a.

Petropoulos, G., Wooster, M. J., Kennedy, K., Carlson, T. N., and Scholze, M.: A global sensitivity analysis study of the 1d SimSphere SVAT model using the GEM SA software, Ecol. Model., 220, 2427–2440, 2009b.

Petropoulos, G. and Carlson, T. N.: Retrievals of turbulent heat fluxes and soil moisture content by remote sensing, in: Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications, Taylor and Francis, 556, 667–502, 2011.

Petropoulos, G. P., Griffiths, H. M., and Tarantola, S.: Towards Operational Products Development from Earth Observation: Exploration of SimSphere Land Surface Process Model Sensitivity using a GSA approach, 7th International Conference on Sensitivity Analysis of 25 Model Output, 1–4 July 2013, Nice, France, 2013a.

Petropoulos G., Griffiths, H. M., Ioannou-Katidis, P., and Srivastava P. K.: Sensitivity exploration of SimSphere land surface model towards its use for operational products development from Earth observation data, in: Remote Sensing Applications in Environmental Research, edited by: Srivastava, P. K., Mukherjee, S., Gupta, M., and Islam, T., Springer, Switzerland, 2013b.

Petropoulos, G. P., Griffiths, H., and Tarantola, S.: Sensitivity analysis of the SimSphere SVAT model in the context of EO-based operational products development, Environ. Modell. Softw., 49, 166–179, 2013c.

Petropoulos, G. P., Griffiths, H. M., Carlson, T. N., Ioannou-Katidis, P., and Holt, T.: SimSphere model sensitivity analysis towards establishing its use for deriving key parameters characterising land surface interactions, Geosci. Model Dev., 7, 1873–1887, doi:10.5194/gmd-7-1873-2014, 2014.

Piles, M., Camps, A., Vall-Llossera, M., Corbella, I., Panciera, R., Rudiger, C., Kerr, Y. H., and Walker, J.: Downscaling SMOS-derived soil moisture using MODIS visible/infrared data, IEEE T. Geosci. Remote., 49, 3156–3166, 2011.

Prentice, I. C., Liang, X., Medlyn, B. E., and Wang, Y.-P.: Reliable, robust and realistic: the three R's of next-generation land-surface modelling, Atmos. Chem. Phys., 15, 5987–6005, doi:10.5194/acp-15-5987-2015, 2015.

Ridler, M. E., Sandholt, I., Butts, M., Lerer, S., Mougin, E., Timouk, F., Kergoat, L., and Madsen, H.: Calibrating a soil–vegetation–atmosphere transfer model with remote sensing estimates of surface temperature and soil surface moisture in a semi-arid environment, J. Hydrol., 436, 1–12, 2012.

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Gonçalves, L. G. G., and Zeng, X.: Towards a comprehensive approach to parameter estimation in land surface parameterization schemes, Hydrol. Process., 27, 2075–2097, 2013.

Ross, S. L. and Oke, T. R.: Tests of three urban energy balance models, Bound-Lay. Meteorol., 44, 73–96, 1988.

Sellers, P. J., Mintz, Y. C. S. Y., Sud, Y. E. A., and Dalcher, A.: A simple biosphere model (SiB) for use within general circulation models, J. Atmos. Sci., 43, 505–531, 1986.

Slevin, D., Tett, S. F. B., and Williams, M.: Multi-site evaluation of the JULES land surface model using global and local data, Geosci. Model Dev., 8, 295–316, doi:10.5194/gmd-8-295-2015, 2015.

Taconet, O., Carlson, T., Bernard, R., and Vidal-Madjar, D.: Evaluation of a surface/vegetation parameterisation using satellite measurements of surface temperature, J. Clim. Appl. Meteorol., 25, 1752–1767, 1986.

Todhunter, P. E. and Terjung, W. H.: Intercomparison of three urban climate models, Bound-Lay. Meteorol., 42, 181–205, 1987.

Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P., Meyers, T. P., Prueger, J. H., Tarks, P. J., and Wesley, M. L.: Correcting eddy-covariance flux underestimates over a grassland, Agr. Forest. Meteorol., 103, 279–300, 2000.

Viterbo, P. and Beljaars, A. C.: An improved land surface parameterization scheme in the ECMWF model and its validation, J. Climate, 8, 2716–2748, 1995.

Wallach, D.: Evaluating crop models, in: Working with dynamic crop models, edited by: Wallach D., Makowski D., and Jones J. W., Elsevier, Amsterdam, The Netherlands, 11–53, 2006.

Wang, Y. P., Baldocchi, D., Leuning, R. A. Y., Falge, E. V. A., and Vesala, T.: Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight Fluxnet sites, Global Change Bio., 13, 652–670, 2007.

Wilson, K., Carlson, T., and Bunce, J. A.: Feedback significantly influences the simulated effect of CO2 on seasonal evapotranspiration from two agricultural species, Global Change Biol., 5, 903–917, 1999.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P. and Verma, S.: Energy balance closure at FLUXNET sites, Agr. Forest Meteorol., 113, 223–243, 2002.

Wilson, K. B. and Baldocchi, D. D.: Seasonal and inter-annual variability of energy fluxes over a broadleaved temperate deciduous forest in North America, Agr. Forest Meteorol., 100, 1–18, 2000.

Wilson, L. G., Everett, L. G., and Cullen, S. J.: Handbook of vadose zone characterization & monitoring, CRC Press, 1994.