

Integration of satellite remote sensing data in ecosystem modelling at local scales: Practices and trends

Damiano Pasetto¹  | Salvador Arenas-Castro²  | Javier Bustamante³  | Renato Casagrandi⁴  | Nektarios Chrysoulakis⁵  | Anna F. Cord⁶  | Andreas Dittrich⁶ | Cristina Domingo-Marimon⁷  | Ghada El Serafy^{8,9} | Arnon Karnieli¹⁰  | Georgios A. Kordelas¹¹ | Ioannis Manakos¹¹  | Lorenzo Mari⁴  | Antonio Monteiro²  | Elisa Palazzi¹²  | Dimitris Poursanidis⁵  | Andrea Rinaldo^{1,13} | Silvia Terzago¹² | Alex Ziembra^{8,9} | Guy Ziv¹⁴ 

¹Laboratory of Ecohydrology, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland; ²CIBIO/InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, Vairão, Portugal; ³Estación Biológica de Doñana, CSIC, Sevilla, Spain; ⁴Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy; ⁵Institute of Applied and Computational Mathematics, Foundation for Research and Technology Hellas, Heraklion, Greece; ⁶Department of Computational Landscape Ecology, UFZ – Helmholtz Centre for Environmental Research, Leipzig, Germany; ⁷Grumets Research Group, CREA, Universitat Autònoma de Barcelona, Bellaterra, Spain; ⁸Deltares, Delft, The Netherlands; ⁹Department of Applied Mathematics, Delft University of Technology, Delft, The Netherlands; ¹⁰Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Beersheba, Israel; ¹¹Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece; ¹²Institute of Atmospheric Sciences and Climate, National Research Council, Turin, Italy; ¹³Department of Civil Environmental and Architectural Engineering, University of Padova, Padova, Italy and ¹⁴School of Geography, Faculty of Environment, University of Leeds, Leeds, UK

Correspondence

Damiano Pasetto

Email: damiano.pasetto@epfl.ch

Funding information

Horizon 2020 Framework Programme, Grant/Award Number: 641762

Handling Editor: Nathalie Pettorelli

Abstract

1. Spatiotemporal ecological modelling of terrestrial ecosystems relies on climatological and biophysical Earth observations. Due to their increasing availability, global coverage, frequent acquisition and high spatial resolution, satellite remote sensing (SRS) products are frequently integrated to in situ data in the development of ecosystem models (EMs) quantifying the interaction among the vegetation component and the hydrological, energy and nutrient cycles. This review highlights the main advances achieved in the last decade in combining SRS data with EMs, with particular attention to the challenges modellers face for applications at local scales (e.g. small watersheds).
2. We critically review the literature on progress made towards integration of SRS data into terrestrial EMs: (1) as input to define model drivers; (2) as reference to validate model results; and (3) as a tool to sequentially update the state variables, and to quantify and reduce model uncertainty.
3. The number of applications provided in the literature shows that EMs may profit greatly from the inclusion of spatial parameters and forcings provided by vegetation and climatic-related SRS products. Limiting factors for the application of such models to local scales are: (1) mismatch between the resolution of SRS products and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

model grid; (2) unavailability of specific products in free and public online repositories; (3) temporal gaps in SRS data; and (4) quantification of model and measurement uncertainties. This review provides examples of possible solutions adopted in recent literature, with particular reference to the spatiotemporal scales of analysis and data accuracy. We propose that analysis methods such as stochastic downscaling techniques and multi-sensor/multi-platform fusion approaches are necessary to improve the quality of SRS data for local applications. Moreover, we suggest coupling models with data assimilation techniques to improve their forecast abilities.

4. This review encourages the use of SRS data in EMs for local applications, and underlines the necessity for a closer collaboration among EM developers and remote sensing scientists. With more upcoming satellite missions, especially the Sentinel platforms, concerted efforts to further integrate SRS into modelling are in great demand and these types of applications will certainly proliferate.

KEYWORDS

data assimilation, ecohydrological models, satellite remote sensing, stochastic downscaling

1 | INTRODUCTION

Anthropogenic and climate change pressures constitute serious threats to the integrity of the delicate ecosystems of several protected areas, such as National Parks, UNESCO World Heritage sites and Natura 2000 sites (Marris, 2011). Ecosystem models (EMs) help researchers to understand the dynamics of these terrestrial environments, and improve monitoring capabilities by filling spatiotemporal data gaps and predicting short- and long-term impacts of different management strategies. Mechanistic ecohydrological models that couple hydrological and vegetation processes (Chen, Wang, Ma, & Liu, 2015) can estimate features like forest productivity and growth (Huber et al., 2013), or evaluate the system water stress under different climate scenarios (Bhattarai, Wagle, Gowda, & Kakani, 2017). These EMs mainly differ in the complexity of the vegetation component and its interaction with the carbon, nutrient and water cycles. Research in this field has mainly focused on improving the physical description of physiological processes (Chen et al., 2015; Fatichi, Ivanov, & Caporali, 2012) to accurately quantify vegetation photosynthesis and growth. However, this modelling effort requires the measurement or estimation of several biophysical parameters (e.g. LAI, canopy height) and input fluxes (e.g. precipitation, land surface temperature, irradiation) that are heterogeneous at the landscape scales and evolve in time (Fatichi, Pappas, & Ivanov, 2016; Pappas, Fatichi, & Burlando, 2016).

Monitoring systems based on in situ, airborne and unmanned aerial vehicles (UAVs) measurements may not always be sufficient to provide the large amount of data required by EMs. Despite the very high spatial resolution (few centimetres), UAV data suffer from low spectral resolution, limited flight endurance of drones, and, for several countries, strict laws regulating UAV use for research purposes (Paneque-Gálvez, McCall, Napoletano, Wich, & Koh, 2014), while in situ and airborne measurements are time-, cost-

and labour-intensive. The constantly expanding list of ready-to-use satellite remote sensing (SRS) products is useful to integrate and, if necessary, replace these measurements (Lawley, Lewis, Clarke, & Ostendorf, 2016) with the main advantages of being freely available for the past two or three decades (e.g. since 1972 for Landsat or 2000 for MODIS), with almost-regular repetition (depending on atmospheric conditions), and at high spectral resolution.

In a seminal work, Plummer (2000) highlighted the crucial role that SRS was already playing two decades ago for the improvement of terrestrial models, defining four strategies for linking SRS data and EMs:

1. using SRS data for the estimation of EMs forcings;
2. using SRS data for the calibration and validation of EMs;
3. using SRS data for updating the state variables of EMs;
4. using EMs to interpret SRS data.

As foreseen by Plummer's conclusions, research efforts in the past decade have been promoted by a closer collaboration among ecological modellers and the RS community (e.g. the ESA Climate Change Initiative), which led to improved SRS products together with the estimation of SRS uncertainties (Merchant et al., 2017). Nowadays, a number of online repositories allow direct access to images from different satellite missions and sensors in near real-time, and to higher level global SRS products certified through quality assurance (QA) tests and standards. While SRS data have been extensively used for updating the state of global terrestrial models, e.g. for hydrological (De Lannoy & Reichle, 2016) and carbon cycle (Scholze, Buchwitz, Dorigo, Guanter, & Quagan, 2017) models, the potential of combining SRS data and EMs for near real-time monitoring at local scales, such as those of small watersheds and protected areas (area up to few hundreds of km²), is yet to be fully exploited. This step is fundamental to achieve the objectives of projects such as the ECOPOTENTIAL H2020 project (

ecopotential-project.eu), which brought together ecologists, park managers and SRS experts with the goal of monitoring several European protected areas through integration of SRS products into EMs.

Given the marked advances since Plummer (2000), we present updates on the state-of-the-art approaches and discuss some remaining difficulties that modellers may face when integrating SRS data into EMs. As high spatial and temporal resolutions of model drivers and parameters are particularly important to describe vegetation dynamics at local scales, Section 2 highlights the strategies adopted in the literature to properly downscale high-level SRS products (i.e. level 2 and level 3 products, Section 2.1) or to directly obtain the quantities of interest through processing of the observations at the sensor level (level 1 products, Sections 2.2 and 2.3). Best practices to compare EM outputs to the associated SRS products are described in Section 3. Data assimilation (DA) techniques that are proving successful to enhance EMs forecast abilities through SRS data are described in Section 4, together with techniques for the evaluation of model and data uncertainties (Sections 4.1 and 4.2). Finally, Section 5 concludes the paper and suggests strategies for further advancements.

2 | PREPARING SRS DATA FOR USE IN LOCAL TERRESTRIAL EMS

Satellite remote sensing products are useful to characterize spatial parameters and forcings related to the hydrological component (precipitation, evapotranspiration [ET], soil moisture, snow cover, etc.), the energy balance (land surface temperature [LST], irradiation, albedo) and the vegetation properties (land use, vegetation cover, vegetation height/biomass, fraction of absorbed photosynthetically active radiation [FAPAR], LAI; Amitrano et al., 2014). Vegetation indices (VIs) have been frequently used as indicators of vegetation cover type and vegetation photosynthetic activity, and for the estimation of phenological parameters. van der Kwast et al. (2009), for example, used the normalized difference vegetation index (NDVI) derived from ASTER to estimate the input parameters of the Surface Energy Balance System, while Turner et al. (2006) adopted land cover and seasonal LAI derived from Landsat Enhanced Thematic Mapper+ (ETM+) as input to the Biome-BGC carbon cycle model for the estimation of landscape gross primary production (GPP) and net primary production (NPP; see Table 1 for further examples).

The spatial resolution, temporal frequency and accuracy of “off-the-shelf” high-level SRS products (Table S1) frequently do not match modelling requirements for local applications, where grid cells from tens to hundreds of metres in size are used. Moreover, the assumptions and ancillary data (which are often “hidden” in the technical documentation, cascades of scientific articles or provider’s websites for more recent changes) used for the computation of these high-level products might not be consistent with the assumptions or other inputs used in local scale EMs. For example, LAI estimated by MODIS are based on specific land cover map characterized by eight biomes (Yan et al., 2016), which may differ from those used in the

EM. We present three typical strategies that modellers are currently adopting and combining to address the aforementioned problems:

1. downscaling low-resolution SRS products, which is particularly useful for climatic products having resolutions of several kilometres (Table S1);
2. deriving high-resolution products from SRS level 1 data, which is relevant to obtain accurate vegetation-related parameters for the domain under study (Table S2);
3. applying multi-sensor/multi-platform fusion techniques, which can fill temporal gaps among the estimated vegetation parameters.

2.1 | Downscaling methods for climatic products

Downscaling helps modellers overcome the scale mismatch between high-level (levels 2, 3) SRS products and the desired model resolution. While applicable to many types of SRS, downscaling methods are frequently used for climatic variables (e.g. precipitation and LST), which are among the main driving factors of terrestrial EMs describing ecosystem seasonality and long-term trends.

2.1.1 | Precipitation

Satellite remote sensing-based precipitation measurements (Table S1) constitute a valid alternative to datasets obtained through spatial interpolation of in situ measurements. These datasets achieve regional-(DAYMET, PRISM over North America) or global-(WorldClim) scale coverage with resolutions up to 1 km, but are known to have poor reliability in areas where the density of ground sensors is low and uneven (Mourtzinis, Rattalino Edreira, Conley, & Grassini, 2017). Moreover, from the temporal perspective, the monthly resolution of the provided climatologies is too coarse to effectively drive many EMs relying on daily or sub-daily forcing data (Table 1). Main problems arising from SRS-based rainfall products concern spatial coverage, because different satellites cover different ranges of latitudes (i.e. GPCP, CMAP or GPM), and their generally coarse spatial resolution, from 2.5 deg (about 280 km at the equator) up to 0.1 deg (about 11 km). A recent comparison of precipitation datasets derived from gauges, models and SRS data also highlighted the variability of different products, in particular in SRS-derived seasonal precipitation and distribution of extreme events (Sun et al., 2018).

Statistical downscaling has been used to obtain rainfall data at about 1 km resolution from e.g. TRMM using classic geostatistical analysis (Chen, Liu, Liu, & Li, 2014; Shi & Song, 2015). Covariates available at higher spatial resolution (e.g. VIs, elevation and other topographic parameters, and in situ weather data) are used to explain part of the large spatiotemporal variability of the precipitation field. Stochastic methods are particularly suitable to generate synthetic precipitation patterns whose statistical properties are consistent with those of observed precipitation. One example is represented by the Rainfall Filtered Autoregressive Model (RainFARM) based on

the extrapolation to the small scales of the Fourier power spectrum of coarse-scale precipitation (D’Onofrio, Palazzi, von Hardenberg, Provenzale, & Calmanti, 2014; Terzago, Palazzi, & von Hardenberg, 2018).

2.1.2 | Land surface temperature

Satellite remote sensing measurements provide daily global spatial coverages of LST at resolutions that vary from 1 km (Sentinel 3) to 56 km (AMSRE-E). Higher resolution data (30 m) are provided at 16 days interval from Landsat 8 (Table S1). Spatial downscaling of LST (also known as sharpening or disaggregating) relies on information about the soil type, emissivity and vegetation cover, frequently using NDVI or LAI as proxy for the latter. Downscaling has been applied to estimate LST at higher spatial resolution (e.g. 250 m) from

MODIS and AVHRR data while maintaining their original temporal resolution (Liu & Pu, 2008; Metz, Rocchini, & Neteler, 2014). Fusion approaches have been tested in keeping the temporal resolution of MODIS (daily) while downscaling it down to Landsat resolution (30 m) or ASTER (90 m; Weng, Fu, & Gao, 2014; Yang et al., 2016). It is worth stressing that LST may differ by several degrees from near-surface air temperature measured by surface stations (Good, 2016). Approaches to estimate near-surface air temperature from satellite observations are still under development, and daily global mapping will be accomplished in the EUSTACE H2020 project (Brugnara, Auchmann, & Brönnimann, 2017).

It is important to remember that SRS climate products might have low accuracy in topographically complex areas mainly due to their low resolution. To improve the accuracy for local applications, downscaling can be applied in conjunction to bias correction techniques

TABLE 1 A non-exhaustive list of examples of ecosystem models (EMs) applications at local scales using satellite remote sensing (SRS) products to describe model parameters and/or forcings. Detailed model descriptions and references are available in Supplementary Information 3

Study	EM model	EM description	Modelled area (km ²)	Cell size/ resolution (m)	Time step	SRS products used
Turner et al. (2006)	Biome-BGC	Process-based model derived from Forest-BCG considering the physical and biological processes governing energy, water, carbon and nitrogen fluxes among the vegetation and soil layers	25	25	1 day	Land cover, LAI (estimated from Landsat ETM+) NPP/GPP (from MODIS for validation)
van der Kwast et al. (2009)	SEBS	Single-source model for the estimation of turbulent heat fluxes and ET	5	90	—	Albedo, Vegetation cover (estimated from ASTER)
Govind et al. (2009)	BEPS-TerrainLab	Couples the hydrological model TerrainLab with the biophysical simulator BEPS Estimates mass and energy fluxes among soil, vegetation and atmosphere, together with plant growth and C-cycle	~40	25	1 day	Land cover (from ESOD) LAI (estimated from NDVI of Landsat TM)
Strauch and Volk (2013)	SWAT	Semi-distributed watershed model for simulations of daily discharge, nutrient, pesticide and sediment loads. Spatial heterogeneities are considered by dividing the domain into sub-basins	242	—	1 day	ET and LAI (from MODIS)
Hwang et al. (2008)	RHESSys	Ecohydrological modelling framework designed to simulate carbon, water and nutrient fluxes	15.8	30	1 day	LAI (estimated from NDVI of Landsat ETM+)
Lopes, Aranha, Walford, O’Brien, and Lucas (2014)	Forest-BCG	Process-based ecophysiological model for NPP estimations based on canopy interception, evaporation, photosynthesis, growth, carbon allocation, litter fall, nitrogen mineralization and mortality	60 and 24	30	1 day	LAI (estimated from EVI and NDVI of Landsat ETM+)

and considering the ancillary data used to obtain the original SRS products (e.g. Maggioni, Meyers, & Robinson, 2016).

2.2 | Deriving high-resolution SRS products

Assessing the key parameters for describing photosynthesis, ET, and NPP is crucial to compute energy, water and carbon fluxes. Several global SRS products related to vegetation biogeochemistry are freely available (Table S1). Although these products have the clear advantage of passing external QA, they differ in terms of algorithms, ancillary data and product uncertainty, and might not be consistent with the particular requirements and assumptions of EMs, especially for local applications. Exploitation of satellite images using empirical, semi-empirical or physically based approaches may be required to obtain consistent products at high spatiotemporal resolutions. We underline, however, that a deep expertise on remote sensing is needed to exploit correctly the available dataset depending on the desired accuracy and application, in particular for assessing the uncertainty of the retrieved variables. The indirect nature of SRS measurements makes them potentially hard to interpret and relate to physically measurable quantities (Disney, 2016).

2.2.1 | Empirical approaches

Empirical approaches establish mathematical relationships between SRS data and the biophysical variables of interest via calibration on in situ data (Chuvieco & Huete, 2009). Various empirical approaches to identify leaf and canopy properties from SRS data have been proposed in literature (Tables S2 and S3), including estimations of biophysical parameters, such as water content, nitrogen, chlorophyll content from VIs or, more recently, from solar-induced chlorophyll fluorescence. Empirical models require extensive field measurements for calibration and the modelling results are dependent of the site condition, time period and sensor, thus limiting their applicability to other sites (Croft, Chen, & Zhang, 2014). However, their reduced computational cost makes them an appreciated asset. For instance, recent applications to Sentinel-2 data have provided LAI and chlorophyll content at high resolution (10 m) and 5 days revisit period (Clevers, Kooistra, & van den Brande, 2017). Empirical approaches are also implemented within cloud computing platforms such as Google Earth Engine, which contains both raw imagery archives (including Landsat and Sentinel-2) and machine-learning algorithms (e.g. robust linear regression, random forest, support-vector machines).

2.2.2 | Physically based models

Physically based models attempt to describe the surface reflectance through physical laws of the radiation transfer inside the canopy and its interaction with the soil surface, and offer an explicit connection between the biophysical variables of vegetation and soil and canopy reflectance (Banskota et al., 2015; Houborg, McCabe, Cescatti, et al., 2015). Physically based models have strong advantages over

empirical approaches: they permit to infer causality and perform predictions, can be adapted to a wide range of land cover situations, time periods and sensor configurations, while at the same time not requiring the simultaneous acquisition of in situ and SRS data. Physically based models are fine-tuned using inversion techniques (Chuvieco & Huete, 2009), including quasi-Newton algorithm, look-up tables and artificial neural networks (e.g. Sehgal, Chakraborty, & Sahoo, 2016). On top of the effort to adequately describe the light interaction processes, one of the main challenges of radiative transfer models (RTMs) remains the development of correction factors to take into account the uncertainty in the radiative response associated to 3D-heterogeneous vegetation. The RAMI4PILPS comparison experiments (Widlowski et al., 2011), which evaluate the consistency of several simple RTMs used within EMs, show that as the structure of the plant canopies becomes more complex, model-to-model agreements generally deteriorate and model-to-reference deviation increases as well.

2.2.3 | Semi-empirical models

Semi-empirical models rely on the theoretical formulations used in physical models, while adjusting some parameters through empirical relationships based on SRS data. Such models reduce the complexity of physical models by reducing the number of parameters requiring calibration. A possibility is to approximate the outputs of RTMs through surrogate functions, e.g. Gaussian process emulators, whose evaluation is faster. Calibration procedures such as Markov Chain Monte Carlo (MC) or DA schemes are necessary for the inversion also in this case (Gómez-Dans, Lewis, & Disney, 2016). Semi-empirical models have been applied to estimate, among others, canopy height and structure and time series of LAI (Kumar, Kumari, & Saha, 2013).

2.3 | Multi-sensor/multi-platform approaches

Expert SRS users are frequently adopting multi-sensor approaches to overcome the fixed temporal and spatial resolutions of single data sources. Data fusion techniques, which require knowledge of sensor limitations and uncertainties, offer advantages for temporally continuous mapping of vegetation parameters (Dusseux, Corpetti, Hubert-Moy, & Corgne, 2014) and for the spatial extrapolation of in situ observations. Limitations of VIs derived from optical data, such as saturation effects and dependence on weather conditions, can be reduced by integration with radar data, taking into account the differences between the responses of optical and radar sensors. In the fields of long-term land use mapping and monitoring, a wide range of studies employ multi-sensor SRS data, mostly combining Landsat with ALOS/PALSAR, Radarsat, or ERS datasets (reviewed in Joshi et al., 2016).

LAI products are highly relevant for EM applications, because the derivation through physically based RTMs makes LAI directly connected to EM parameters and inputs like plant productivity, transpiration and energy fluxes (Asner, Scurlock, & Hicke, 2003).

Recent studies showed that the low temporal frequency (16 days) of LAI products derived from Landsat can be improved by combining MODIS reflectance and LAI data with higher temporal resolution (Myneni, Knyazikhin, & Park, 2015), while keeping Landsat spatial resolution (Houborg, McCabe, & Gao, 2015).

Mountain areas pose particular challenges for SRS applications, making indispensable the application of thorough topographic and atmospheric correction (for optical data), as well as of methods to account for foreshortening and layover effects and variations in surface water content affecting dielectric properties (for radar data, Gupta, 2018). In particular, multi-sensor/multi-platform approaches require radiometrically homogeneous and consistent reflectance among images from different sources and different time periods. For instance, Attarchi and Gloaguen (2014) were able to develop a statistical model for above-ground biomass in a mountain forest site in Iran by combining corrected and co-registered Landsat ETM+ and ALOS/PALSAR data, which quality significantly improved compared to the single use of (uncorrected) optical and radar data. In (semi-)arid regions, synergistic optical and radar data were used to model daily ET (Hu & Jia, 2015), a key parameter in this kind of ecosystem. Multi-sensor approaches can also be effective in estuarine environments: Chakraborty, Ferrazzoli, and Rahmoune (2014) analysed MODIS Enhanced Vegetation Index and AMSR-E radar signatures to estimate the amount of vegetation biomass that is submerged during monsoon flooding, while Rangoonwala, Enwright, Ramsey, and Spruce (2016) related persistence of marshland flooding (radar data) to changes in vegetation (optical data).

While the use of multi-sensor approaches offers new opportunities with respect to time-series analysis, cross-sensor calibration is a challenging task. The long-term availability of data from Landsat and SPOT missions makes these sensors still the most suitable for multi-temporal analysis. As Sentinel-2, with the first platform launched in 2015, has a similar spectral coverage but better spatial and temporal resolution, current efforts focus on combining it with Landsat data to provide near daily global coverage at 30 m resolution (NASA Harmonized Landsat-Sentinel-2 project, Claverie & Masek, 2017). Although some research has been carried out to develop methods to obtain radiometric homogenization between different sensor time series (Padró et al., 2017; Pons, Pesquer, Cristóbal, & González-Guerrero, 2014), there are still challenges to be met for the use of multi-temporal, multi-source SRS data in ecology.

3 | MODEL CALIBRATION AND VALIDATION USING SRS DATA

Satellite remote sensing is currently used not only as an input to EMs, but also to assess model reliability through validation techniques (Bennett et al., 2013). A large number of SRS products have served to assess EM results at the ecosystem level such as LAI, FAPAR, soil moisture, GPP and NPP (see examples in Table S4).

The operations to process the model outputs to obtain a variable consistent with the measured data constitute the so-called

“observation operator” (Kaminski & Mathieu, 2017), which is necessary for the implementation of calibration, validation and assimilation schemes. Two strategies can be used to compare the EM outputs to SRS data (Plummer, 2000), namely indirect or direct comparison, which is related to two different observation operators.

3.1 | Indirect comparison

Indirect comparison considers high-level SRS products (e.g. LAI, FAPAR, etc.). In this case, the observation operator adapts model outputs to the measurements, typically through downscaling/upscaling procedures. The correction of biases due to different assumptions between a particular EM and products (e.g. over-simplification of the vegetation layer in the EM) is a particularly important step that, if not considered, might lead to large discrepancies in the results (Liu et al., 2018).

3.2 | Direct comparison

Direct comparison couples ecological and reflectance models, where the EM outputs become input of RTMs, so as to directly compare the measured and modelled radiances. Direct comparison is particularly appealing since it avoids possible discrepancies between SRS products and model outputs, and the inversion of a RTM. Examples are the coupling between ecological and reflectance models in hydrological applications, e.g. using temperature brightness from SMOS microwave sensor (De Lannoy & Reichle, 2016), or the computation of modelled canopy reflectance to be compared to MODIS data (Quaife et al., 2008). The observation operator required for direct comparison is not easily implemented in most EMs, because it considers the adaptation of the EM outputs to the RTMs inputs, and the operations in the RTM including backscatter from soil, vegetation and atmosphere. These processes are rarely considered explicitly due to the computational burden of complex dedicated RTMs, and simplified RT schemes are employed (e.g. 1D, effective LAI, etc.). Thus, direct comparison requires the calibration of several additional parameters. Due to these difficulties and the free availability of many SRS products, indirect comparison is still frequently adopted, especially when validation is simply performed by qualitative approaches (Table S4).

Note that indirect comparison requires the evaluation of error metrics between raster maps (Stow et al., 2009), which are typically characterized by strong spatial autocorrelation. Due to the errors introduced by SRS downscaling procedures, the accuracy of the sensor and the complex nature of environmental systems, classical validation metrics based on a per-pixel comparison (such as the root mean squared error, the Pearson's correlation coefficient or Nash-Sutcliffe efficiency) might not be able to evince common spatial patterns, thus limiting the comparison to mere qualitative considerations. In these cases, residual-based metrics should be replaced by the analysis of statistical moments, spectra and other quantitative measures of spatial structure that have been developed to objectively reveal common patterns among maps (Koch, Jensen, & Stisen, 2015). The quantified uncertainties associated with the model results and SRS observations are frequently neglected during the validation of model

results, but should be included to weigh their relative contribution to the error metric (see Section 4.2).

4 | ASSIMILATION OF SRS DATA FOR THE UPDATE OF MODEL STATE VARIABLES

Ecosystem models suffer from the presence of many sources of errors that may propagate in time amplifying the uncertainty on the model outputs. Model-Data Fusion techniques (MDFs) reduce and control this model uncertainty by consistently combining EMs and data (Peng, Guiot, Wu, Jiang, & Luo, 2011). DA schemes are a particular family of MDFs mainly developed by the meteorological community to improve model forecast by updating the state variables using measured data. DA has been applied in several fields of ecology (Luo et al., 2011), e.g. for modelling the spread of infectious diseases, fire, fisheries (Niu et al., 2014), to improve carbon cycle models (Benavides Pinjosovsky et al., 2017) and hydrological predictions (Lahoz & De Lannoy, 2014).

Open access repositories such as Open DA, the Parallel DA Framework (Kurtz et al., 2016) and the DA Research Testbed (Mizzi, Arellano, Edwards, Anderson, & Pfister, 2016) have been developed to help the coupling between EMs and state-of-the-art DA methods. A particularly useful software for EM applications at local scales is ESA's project EO-Land Data Assimilation System (EO-LDAS, Lewis et al., 2012), which provides a Python library for the retrieval of geophysical parameters by assimilation of the optical medium-resolution data of Sentinel-2.

The core of DA schemes is the update (or analysis) step that corrects the model state variables towards the observations, reducing the forecast uncertainty. Different implementations of the update step characterize three main approaches (Montzka, Pauwels, Hendricks Franssen, Han, & Vereecken, 2012):

1. Variational methods, which minimize a cost function associated with the residual between forecast and observations (Benavides Pinjosovsky et al., 2017).
2. Kalman-based methods, which extend the well-known Kalman filter to nonlinear/non-Gaussian models, e.g. using MC simulations as in the Ensemble Kalman filter (EnKF; Quaife et al., 2008).
3. Particle filters, which directly apply MC sampling using a Bayesian approach (De Bernardis, Vicente-Guijalba, Martinez-Marin, & Lopez-Sanchez, 2016).

The analysis step of DA updates the model state variables based on a balance between model and observation uncertainties, described by estimates of their probability distribution. This is a particularly difficult task for both EM outputs and SRS data.

4.1 | SRS measurement uncertainty

The assimilation of SRS data requires the computation of error cross-covariances at the pixel level, taking into account the cumulative

effect of different sources of uncertainty: sensor errors, errors in the RTM, errors of representativity (due to upscaling or downscaling steps) and errors introduced when processing the data, e.g. atmospheric and radiometric corrections as well as filtering or masking optical encumbrances such as clouds and haze (Pfeifer, Disney, Quaife, & Marchant, 2012; Waller, Dance, Lawless, & Nichols, 2014). However, the propagation of the instrument and parameter errors is frequently neglected during the inversion of RTMs, and the product accuracy is assessed a posteriori, by means of costly validations against in situ measurements, or comparison with the output of calibrated process-based models (Wanders, Karssenber, De Roo, De Jong, & Bierkens, 2012). QA of online products usually provides qualitative information on pixel values (e.g. QA band for Landsat VIs specifies the pixel condition about cloud, snow, water, etc.) and only rarely quantitative information on the accuracy of the data (e.g. SD at the pixel level for MODIS LAI/FAPAR) which is what needed for DA. The lack of information on the errors of SRS products can be relieved by using the statistics associated with the residuals between EMs outputs and SRS measurements evaluated during the DA updates (Crow & Reichle, 2008). Direct coupling represents a valid alternative, allowing the assimilation of the optical signal at the sensor, thus describing the observation error as the accuracy of the sensor (e.g. De Lannoy & Reichle, 2016; Zhang, Shi, & Dou, 2012). In this latter case, the evaluation of model uncertainty (see Section 4.2) has to be propagated through the observation operator (described in Section 3). Moreover, model uncertainties might be amplified by the unknown parameters of the RTM. As an example, EnKF assimilation of canopy reflectance from MODIS has been shown to improve EM estimates of GPP and reduce model uncertainty (Quaife et al., 2008).

4.2 | EM uncertainty

Correct estimation of model error is fundamental for DA. Underestimating model uncertainty reduces the relevance of the data (the assimilation would marginally correct the system state variables), with increased risk of divergence from the actual state of the system. Overestimating it, instead, would result in poor forecast capabilities, with the model forecast that spans a large range of possible solutions. The latter is more favourable to DA analysis, since the true system state is more likely to fall within the prediction interval.

Ecosystem Models uncertainties are mainly evaluated during the forecast step, which drives forward the state variables until the following observation time. The main sources of uncertainty are input variables (initial conditions, external forcing), unknown parameters, structural uncertainties due to the physical simplification of the governing processes, and numerical approximations for the discretization of continuous processes (Refsgaard, van der Sluijs, Hojberg, & Vanrolleghem, 2007). MC-based DA schemes, such as EnKF and particle filters, estimate forecast uncertainties by running an ensemble of EM realizations, each associated with different samples of the inputs, forcing term and parameter probability density functions (PDFs). The definition of these PDFs is typically straightforward for the forcing,

based on direct error statistics (e.g. De Lannoy & Reichle, 2016). Assessment of the parameter PDFs is more difficult and computationally expensive. For many systems, literature provides insights into the distribution limits, type of density function and modal value, supplemented through local knowledge and expertise. PDF tuning can be done through sensitivity analysis and Bayesian parameter calibration techniques (Harrison, Kumar, Peters-Lidard, & Santanello, 2012). Finally, the characterization of model structural uncertainty is non-trivial and constitutes one of the main issues in model identification, with important implications for forecasting. Structural uncertainty represents the fundamental inability of a model to represent the processes it is designed to replicate. Numerical errors, which arise from the spatiotemporal discretization of model equations, are frequently considered together with the structural uncertainty (El Serafy et al., 2011). The simplest approach to account for structural uncertainties is to consider the model error as a Gaussian distribution, where the estimation of the covariance matrix follows a procedure developed ad hoc for each model and application (Reichle, 2008). As an example of EM uncertainty estimation, De Lannoy and Reichle (2016) assigned static perturbation statistics for the error of the land surface model GEOS-5 CLSM. In this case, two model state variables are perturbed with an additive noise characterized by fix temporal and spatial correlations. In a more sophisticated fashion, El Serafy et al. (2011) proposed an iterative procedure based on a MC sampling method to estimate the error covariance of a model for suspended particulate matter concentration (Delft3D-WAQ). In general, the determination of structural uncertainties through sensitivity analyses, parameterization, and conceptual models (Matott, Babendreier, & Purucker, 2009; Refsgaard, van der Sluijs, Brown, & van der Keur, 2006) is heavily reliant on a sufficiently large number of in situ data (Uusitalo, Lehikoinen, Helle, & Myrberg, 2015).

Many studies neglect these different sources of uncertainty, thus possibly underestimating the model output variances (Matott et al., 2009). Further research and inquiry are required to provide standard methods for the estimation of model structural uncertainty.

5 | CONCLUSIONS

In the two decades since Plummer (2000), the ecological community has acquired more confidence in the use of SRS data, mainly owing to the increased availability of global products subject to rigorous QA tests. Global SRS datasets are nowadays used also by ecologists and conservation managers who may not have experience or knowledge of how SRS datasets are generated. For modellers, major risks of integrating SRS products into EMs are that the specific assumptions underlying their production (e.g. the land cover considered in the RTM) might be incompatible with the assumptions of the EMs, or that the practices in downscaling the SRS data for local-scale modelling (e.g. nearest neighbour resampling) are inappropriate. Better information about the conditions, assumptions and ancillary data

behind SRS data production needs to be provided by remote sensing experts with off-the-shelf products, perhaps via community-driven adoption of detailed metadata standards. On the other hand, modellers need to be educated to properly use SRS and their QA metadata, starting from improved university programmes in ecology and physical geography, which currently lack sufficient depth on SRS (Bernd et al., 2017).

Global SRS products are increasingly available to the ecological modelling community, but their application to local scale EMs is not straightforward, requiring downscaling, or the adaptation of the developed algorithms to higher resolution satellite imagery. So far, the lack of shared codes, the difficulties of adapting them to different SRS products and the need for larger computational resources have hindered the use of such algorithms and products. Two emerging trends may be key to changing this. First, the surge of cloud computing (e.g. Google Earth Engine, ESA RUS) allows SRS experts to upscale their algorithms and produce large-scale coverage (e.g. Pekel, Cottam, Gorelick, & Belward, 2016). This will be further pushed by the six Copernicus Data and Information Access Services (DIAS) platforms to be launched in 2018. Second, the adoption of open research practice (and not just open data) for new projects, such as those funded by the European Framework Programmes (e.g. Horizon 2020), is pushing the SRS community to release the codes developed for local applications. New initiatives, such as the "Model Web" within the GEO System of Systems, aim to make such models and algorithms accessible and executable online.

To fully explore the potentiality of SRS data, we suggest that the development of spatially explicit EMs should adopt the following strategies:

1. Direct coupling of EMs with RTMs, a strategy that is rarely adopted in ecological applications at present. Although direct coupling would require the calibration of a larger number of model parameters, the possibility of directly simulating the system reflectance clearly gives considerable advantages for (1) the validation of model outputs, (2) the assimilation of SRS data in near real time (without the necessity to wait the production of higher level products) and (3) the estimation of measurement uncertainties, which is required in DA techniques and allows for prediction.
2. structuring the simulation codes, so that they can be easily connected to the available DA platforms. In fact, the relatively small investment required in terms of updating model implementation would be greatly compensated by the potential to access a number of state-of-the-art, well-tested algorithms for the assimilation of SRS data and the assessment of model uncertainty.

Further collaborations between the remote sensing and the ecosystem modelling scientific communities will help to overcome the outlined difficulties and develop standardized techniques to include SRS data into EMs. There are promising prospects for new SRS missions, and for the development of new algorithms, standards and platforms.

Realizing these will require collaborative projects demonstrating the use of SRS in EMs, similar to the Horizon 2020 ECOPOTENTIAL project whose focus was on improving protected areas management using remote sensing.

The coming years promise great potential for SRS and EMs. In the next decade, several new missions are planned to be launched, including P-band SAR (BIOMASS mission) and ISS-mounted LiDAR (GEDI) for estimating biomass, hyperspectral imaging spectrometers for more accurate RTM inversion (EnMAP, HyspIRI, Hisui), high-resolution sun-induced fluorescence for estimating plant photosynthetic activity (FLEX), and a thermal radiometer to monitor water stress (ECOSTRESS). In addition, constellations of small and less-expensive satellites taking high-resolution imagery of the entire Earth every day, like PlantScope (optical) and ICEYE (SAR), are becoming a reality. In combination with longer time-series data from Sentinel missions and the computational power of cloud platforms such as Google Earth Engine and the upcoming DIAS platforms—the sky is the limit for ecological modellers.

ACKNOWLEDGEMENTS

This work has been carried out within the H2020 project “ECOPOTENTIAL: Improving Future Ecosystem Benefits Through Earth Observations”, coordinated by CNR-IGG (<http://www.ecopotential-project.eu>). The project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement no. 641762.

AUTHORS' CONTRIBUTION

D.P., I.M., A.F.C. and G.Z. conceived the ideas and the scope of the review; D.P. led the writing of the manuscript. All authors critically contributed to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

The manuscript does not include any data.

ORCID

Damiano Pasetto  <http://orcid.org/0000-0001-6892-9826>
 Salvador Arenas-Castro  <http://orcid.org/0000-0001-8671-4968>
 Javier Bustamante  <http://orcid.org/0000-0001-7515-0677>
 Renato Casagrandi  <http://orcid.org/0000-0001-5177-803X>
 Nektarios Chrysoulakis  <http://orcid.org/0000-0002-5208-626X>
 Anna F. Cord  <http://orcid.org/0000-0003-3183-8482>
 Cristina Domingo-Marimon  <http://orcid.org/0000-0001-6822-8704>
 Arnon Karnieli  <http://orcid.org/0000-0001-8065-9793>
 Ioannis Manakos  <http://orcid.org/0000-0001-6833-294X>
 Lorenzo Mari  <http://orcid.org/0000-0003-1326-9992>

Antonio Monteiro  <http://orcid.org/0000-0002-9462-5938>

Elisa Palazzi  <http://orcid.org/0000-0003-1683-5267>

Dimitris Poursanidis  <http://orcid.org/0000-0003-3228-280X>

Guy Ziv  <http://orcid.org/0000-0002-6776-0763>

REFERENCES

- Amitrano, D., di Martino, G., Iodice, A., Riccio, D., Ruello, G., Ciervo, F., ... Koussoubé, Y. (2014). Effectiveness of high-resolution SAR for water resource management in low-income semi-arid countries. *International Journal of Remote Sensing*, 35, 70–88. <https://doi.org/10.1080/01431161.2013.862605>
- Asner, G. P., Scurlock, J. M. O., & Hicke, J. A. (2003). Global synthesis of leaf area index observations: Implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, 12, 191–205. <https://doi.org/10.1046/j.1466-822X.2003.00026.x>
- Attarchi, S., & Gloaguen, R. (2014). Improving the estimation of above ground biomass using dual polarimetric PALSAR and ETM+ data in the Hyrcanian mountain forest (Iran). *Remote Sensing*, 6, 3693–3715. <https://doi.org/10.3390/rs6053693>
- Banskota, A., Serbin, S. P., Wynne, R. H., Thomas, V. A., Falkowski, M. J., Kayastha, N., ... Townsend, P. A. (2015). An LUT-based Inversion of DART model to estimate forest LAI from hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 3147–3160. <https://doi.org/10.1109/JSTARS.2015.2401515>
- Benavides Pinjosovsky, H. S., Thiria, S., Ottlé, C., Brajard, J., Badran, F., & Maugis, P. (2017). Variational assimilation of land surface temperature within the ORCHIDEE Land Surface Model Version 1.2.6. *Geoscientific Model Development*, 10, 85–104. <https://doi.org/10.5194/gmd-10-85-2017>
- Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., ... Andreassian, V. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, 40, 1–20. <https://doi.org/10.1016/j.envsoft.2012.09.011>
- Bernd, A., Braun, D., Ortmann, A., Ulloa-Torrealba, Y. Z., Wohlfart, C., & Bell, A. (2017). More than counting pixels – Perspectives on the importance of remote sensing training in ecology and conservation. *Remote Sensing in Ecology and Conservation*, 3, 38–47. <https://doi.org/10.1002/rse2.27>
- Bhattacharai, N., Wagle, P., Gowda, P. H., & Kakani, V. G. (2017). Utility of remote sensing-based surface energy balance models to track water stress in rain-fed switchgrass under dry and wet conditions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 133, 128–141. <https://doi.org/10.1016/j.isprsjprs.2017.10.010>
- Brugnara, Y., Auchmann, R., & Brönnimann, S. (2017). Global LSAT dataset with discontinuities identified where possible. EU Surface Temperature for All Corners of Earth Deliverable (EUSTACE). Available from https://www.eustaceproject.eu/eustace/static/media/uploads/Deliverables/eustace_d1-7.pdf
- Chakraborty, R., Ferrazzoli, P., & Rahmoune, R. (2014). Application of passive microwave and optical signatures to monitor submerging of vegetation due to floods. *International Journal of Remote Sensing*, 35, 6310–6328. <https://doi.org/10.1080/01431161.2014.950764>
- Chen, F., Liu, Y., Liu, Q., & Li, X. (2014). Spatial downscaling of TRMM 3B43 precipitation considering spatial heterogeneity. *International Journal of Remote Sensing*, 35, 3074–3093. <https://doi.org/10.1080/01431161.2014.902550>
- Chen, L., Wang, L., Ma, Y., & Liu, P. (2015). Overview of ecohydrological models and systems at the watershed scale. *IEEE Systems Journal*, 9, 1091–1099. <https://doi.org/10.1109/JSYST.2013.2296979>

- Chuvieco, E., & Huete, A. (2009). *Fundamentals of satellite remote sensing*. Boca Raton, FL: CRC Press/Taylor & Francis.
- Claverie, M., & Masek, J. (2017). Harmonized Landsat-8 Sentinel-2 (HLS) Product's Guide. v.1.3. Available from <https://hls.gsfc.nasa.gov/documents/>
- Clevers, J. G. P. W., Kooistra, L., & van den Brande, M. M. M. (2017). Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sensing*, 9, 405. <https://doi.org/10.3390/rs9050405>
- Croft, H., Chen, J. M., & Zhang, Y. (2014). The applicability of empirical vegetation indices for determining leaf chlorophyll content over different leaf and canopy structures. *Ecological Complexity*, 17, 119–130. <https://doi.org/10.1016/j.ecocom.2013.11.005>
- Crow, W. T., & Reichle, R. H. (2008). Comparison of adaptive filtering techniques for land surface data assimilation. *Water Resources Research*, 44, W08423. <https://doi.org/10.1029/2008WR006883>
- de Bernardis, C., Vicente-Guijalba, F., Martínez-Marin, T., & Lopez-Sanchez, J. (2016). Particle filter approach for real-time estimation of crop phenological states using time series of NDVI images. *Remote Sensing*, 8, 610. <https://doi.org/10.3390/rs8070610>
- de Lannoy, G. J. M., & Reichle, R. H. (2016). Global assimilation of multi-angle and multipolarization SMOS brightness temperature observations into the GEOS-5 catchment land surface model for soil moisture estimation. *Journal of Hydrometeorology*, 17, 669–691. <https://doi.org/10.1175/JHM-D-15-0037.1>
- Disney, M. (2016). Remote sensing of vegetation: Potentials, limitations, developments and applications. In K. Hikosaka, Ü. Niinemets, & N. Anten (Eds.), *Canopy photosynthesis: From basics to applications* (pp. 289–331). Dordrecht: Springer Netherlands. <https://doi.org/10.1007/978-94-017-7291-4>
- D'Onofrio, D., Palazzi, E., von Hardenberg, J., Provenzale, A., & Calmanti, S. (2014). Stochastic rainfall downscaling of climate models. *Journal of Hydrometeorology*, 15, 830–843. <https://doi.org/10.1175/JHM-D-13-096.1>
- Dusseux, P., Corpetti, T., Hubert-Moy, L., & Corgne, S. (2014). Combined use of multi-temporal optical and radar satellite images for grassland monitoring. *Remote Sensing*, 6, 6163–6182. <https://doi.org/10.3390/rs6076163>
- El Serafy, G. Y., Eleveld, M. A., Blaas, M., van Kessel, T., Aguilar, S. G., & van der Woerd, H. J. (2011). Improving the description of the suspended particulate matter concentrations in the southern North Sea through assimilating remotely sensed data. *Ocean Science*, 46, 179–204. <https://doi.org/10.1007/s12601-011-0015-x>
- Fatichi, S., Ivanov, V. Y., & Caporali, El. (2012). A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments: 1. Theoretical framework and plot-scale analysis. *Journal of Advances in Modelling Earth Systems*, 4, M05002. <https://doi.org/10.1029/2011MS000086>
- Fatichi, S., Pappas, C., & Ivanov, V. Y. (2016). Modeling plant-water interactions: An ecohydrological overview from the cell to the global scale. *WIREs Water*, 3, 327–368. <https://doi.org/10.1002/wat2.1125>
- Gómez-Dans, J. L., Lewis, P. E., & Disney, M. (2016). Efficient emulation of radiative transfer codes using gaussian processes and application to land surface parameter inferences. *Remote Sensing*, 8, 119. <https://doi.org/10.3390/rs8020119>
- Good, E. J. (2016). An *in situ*-based analysis of the relationship between land surface “skin” and screen-level air temperatures. *Journal of Geophysical Research: Atmospheres*, 121, 8801–8819. <https://doi.org/10.1002/2016JD025318>
- Govind, A., Chen, J. M., Margolis, H., Ju, W., Sonnentag, O., & Giasson, M.-A. (2009). A spatially explicit hydro-ecological modeling framework (BEPS-TerrainLab V2.0): Model description and test in a boreal ecosystem in Eastern North America. *Journal of Hydrology*, 367, 200–216. <https://doi.org/10.1016/j.jhydrol.2009.01.006>
- Gupta, R. P. (2018). Interpretation of SAR Imagery. In *Remote sensing geology*. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-662-55876-8>
- Harrison, K. W., Kumar, S. V., Peters-Lidard, C. D., & Santanello, J. A. (2012). Quantifying the change in soil moisture modeling uncertainty from remote sensing observations using Bayesian inference techniques. *Water Resources Research*, 48, W11514. <https://doi.org/10.1029/2012WR012337>
- Houborg, R., McCabe, M., Cescatti, A., Gao, F., Schull, M., & Gitelson, A. (2015). Joint leaf chlorophyll content and leaf area index retrieval from Landsat data using a regularized model inversion system (REGFLEC). *Remote Sensing of Environment*, 159, 203–221. <https://doi.org/10.1016/j.rse.2014.12.008>
- Houborg, R., McCabe, M., & Gao, F. (2015). Downscaling of coarse resolution LAI products to achieve both high spatial and temporal resolution for regions of interest. Geoscience and Remote Sensing Symposium (IGARSS), IEEE International, Milan.
- Hu, G., & Jia, L. (2015). Monitoring of evapotranspiration in a semi-arid inland river basin by combining microwave and optical remote sensing observations. *Remote Sensing*, 7, 3056–3087. <https://doi.org/10.3390/rs70303056>
- Huber, O. M., Eastaugh, C. S., Gschwantner, T., Hasenauer, H., Kindermann, G., Ledermann, T., ... Sterba, H. (2013). Comparing simulations of three conceptually different forest models with National Forest Inventory data. *Environmental Modelling & Software*, 40, 88–97. <https://doi.org/10.1016/j.envsoft.2012.08.003>
- Hwang, T., Kang, S., Kim, J., Kim, Y., Lee, D., & Bend, L. (2008). Evaluating drought effect on MODIS Gross Primary Production (GPP) with an eco-hydrological model in the mountainous forest, East Asia. *Global Change Biology*, 14, 1037–1056. <https://doi.org/10.1111/j.1365-2486.2008.01556.x>
- Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., & Waske, B. (2016). A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing*, 8, 70. <https://doi.org/10.3390/rs8010070>
- Kaminski, T., & Mathieu, P.-P. (2017). Reviews and syntheses: Flying the satellite into your model: On the role of observation operators in constraining models of the Earth system and the carbon cycle. *Biogeosciences*, 14, 2343–2357. <https://doi.org/10.5194/bg-14-2343-2017>
- Koch, J., Jensen, K. H., & Stisen, S. (2015). Toward a true spatial model evaluation in distributed hydrological modeling: Kappa statistics, Fuzzy theory, and EOF-analysis benchmarked by the human perception and evaluated against a modeling case study. *Water Resources Research*, 51, 1225–1246. <https://doi.org/10.1002/2014WR016607>
- Kumar, V., Kumari, M., & Saha, S. K. (2013). Leaf area index estimation of lowland rice using semi-empirical backscattering model. *Journal of Applied Remote Sensing*, 7, 73474. <https://doi.org/10.1117/1.JRS.7.073474>
- Kurtz, W., He, G., Kollet, S. J., Maxwell, R. M., Vereecken, H., & Hendricks Franssen, H.-J. (2016). TerrSysMP-PDAF (version 1.0): A modular high-performance data assimilation framework for an integrated land surface-subsurface model. *Geoscientific Model Development*, 9, 1341–1360. <https://doi.org/10.5194/gmd-9-1341-2016>
- Lahoz, W. A., & de Lannoy, G. J. M. (2014). Closing the gaps in our knowledge of the hydrological cycle over land: Conceptual problems. *Surveys in Geophysics*, 35, 623–660. <https://doi.org/10.1007/s10712-013-9221-7>
- Lawley, V., Lewis, M., Clarke, K., & Ostendorf, B. (2016). Site-based and remote sensing methods for monitoring indicators of vegetation condition: An Australian review. *Ecological Indicators*, 60, 1273–1283. <https://doi.org/10.1016/j.ecolind.2015.03.021>
- Lewis, P., Gómez-Dans, J., Kaminski, T., Settle, J., Quaipe, T., Gobron, N., ... Berger, M. (2012). An earth observation land data assimilation

- system (EO-LDAS). *Remote Sensing of Environment*, 120, 219–235. <https://doi.org/10.1016/j.rse.2011.12.027>
- Liu, D., & Pu, R. (2008). Downscaling thermal infrared radiance for sub-pixel land surface temperature retrieval. *Sensors*, 8, 2695–2706. <https://doi.org/10.3390/s8042695>
- Liu, Y., Xiao, J., Ju, W., Zhu, G., Wu, X., Fan, W., ... Zhou, Y. (2018). Satellite-derived LAI products exhibit large discrepancies and can lead to substantial uncertainty in simulated carbon and water fluxes. *Remote Sensing of Environment*, 206, 174–188. <https://doi.org/10.1016/j.rse.2017.12.024>
- Lopes, D. M., Aranha, J. T., Walford, N., O'Brien, J., & Lucas, N. (2014). Accuracy of remote sensing data versus other sources of information for estimating net primary production in *Eucalyptus globulus* Labill. and *Pinus pinaster* Ait. ecosystems in Portugal. *Canadian Journal of Remote Sensing*, 35(1), 37–53. <https://doi.org/10.5589/m08-078>
- Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., ... Schimel, D. S. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, 21, 1429–1442. <https://doi.org/10.1890/09-1275.1>
- Maggioni, V., Meyers, P. C., & Robinson, M. D. (2016). A review of merged high-resolution satellite precipitation product accuracy during the tropical rainfall measuring mission (TRMM) era. *Journal of Hydrometeorology*, 17, 1101–1117. <https://doi.org/10.1175/JHM-D-15-0190.1>
- Marris, E. (2011). The end of the wild. *Nature*, 469, 150–152. <https://doi.org/10.1038/469150a>
- Matott, L. S., Babendreier, J. E., & Purucker, S. T. (2009). Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resources Research*, 45, W06421. <https://doi.org/10.1029/2008WR007301>
- Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., ... Wagner, W. (2017). Uncertainty information in climate data records from Earth observation. *Earth System Science Data*, 9, 511–527. <https://doi.org/10.5194/essd-9-511-2017>
- Metz, M., Rocchini, D., & Neteler, M. (2014). Surface temperatures at the continental scale: Tracking changes with remote sensing at unprecedented detail. *Remote Sensing*, 6, 3822–3840. <https://doi.org/10.3390/rs6053822>
- Mizzi, A. P., Arellano, A. F., Edwards, D. P., Anderson, J. L., & Pfister, G. G. (2016). Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: A regional chemical transport/ensemble Kalman filter data assimilation system. *Geoscientific Model Development*, 9, 965–978. <https://doi.org/10.5194/gmd-9-965-2016>
- Montzka, C., Pauwels, V., Hendricks Franssen, H.-J., Han, X., & Vereecken, H. (2012). Multivariate and multiscale data assimilation in terrestrial systems: A review. *Sensors*, 12, 16291–16333. <https://doi.org/10.3390/s121216291>
- Mourtzinis, S., Rattalino Edreira, J. I., Conley, S. P., & Grassini, P. (2017). From grid to field: Assessing quality of gridded weather data for agricultural applications. *European Journal of Agronomy*, 82, 163–172. <https://doi.org/10.1016/j.eja.2016.10.013>
- Myneni, R., Knyazikhin, Y., & Park, T. Boston University and MODAPS SIPS - NASA. (2015). MOD15A3H MODIS/Combined Terra+Aqua Leaf Area Index/FPAR Daily L4 Global 500m SIN Grid. NASA LP DAAC.
- Niu, S., Luo, Y., Dietze, M. C., Keenan, T. F., Shi, Z., Li, J., & Chapin, F. S. III (2014). The role of data assimilation in predictive ecology. *Ecosphere*, 5, 1–16. <https://doi.org/10.1890/ES13-00273.1>
- Padr , J.-C., Pons, X., Aragon s, D., D az-Delgado, R., Garc a, D., Bustamante, J., ... Lange, M. (2017). Radiometric correction of simultaneously acquired Landsat-7/Landsat-8 and Sentinel-2A imagery using pseudoinvariant areas (PIA): Contributing to the Landsat time series legacy. *Remote Sensing*, 9, 1319. <https://doi.org/10.3390/rs9121319>
- Paneque-G lvez, J., McCall, M. K., Napoletano, B. M., Wich, S. A., & Koh, L. P. (2014). Small drones for community-based forest monitoring: An assessment of their feasibility and potential in tropical areas. *Forests*, 5, 1481–1507. <https://doi.org/10.3390/f5061481>
- Pappas, C., Fatichi, S., & Burlando, P. (2016). Modeling terrestrial carbon and water dynamics across climatic gradients: Does plant trait diversity matter? *New Phytologist*, 209, 137–151. <https://doi.org/10.1111/nph.13590>
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540, 418–422. <https://doi.org/10.1038/nature20584>
- Peng, C., Guiot, J., Wu, H., Jiang, H., & Luo, Y. (2011). Integrating models with data in ecology and palaeoecology: Advances towards a model-data fusion approach. *Ecology Letters*, 14, 522–536. <https://doi.org/10.1111/j.1461-0248.2011.01603.x>
- Pfeifer, M., Disney, M., Quaife, T., & Marchant, R. (2012). Terrestrial ecosystems from space: A review of earth observation products for macroecology applications. *Global Ecology and Biogeography*, 21, 603–624. <https://doi.org/10.1111/j.1466-8238.2011.00712.x>
- Plummer, S. E. (2000). Perspectives on combining ecological process models and remotely sensed data. *Ecological Modelling*, 129, 169–186. [https://doi.org/10.1016/S0304-3800\(00\)00233-7](https://doi.org/10.1016/S0304-3800(00)00233-7)
- Pons, X., Pesquer, L., Crist bal, J., & Gonz lez-Guerrero, O. (2014). Automatic and improved radiometric correction of Landsat imagery using reference values from MODIS surface reflectance images. *International Journal of Applied Earth Observation and Geoinformation*, 33, 243–254. <https://doi.org/10.1016/j.jag.2014.06.002>
- Quaife, T., Lewis, P., de Kauwe, M., Williams, M., Law, B. E., Disney, M., & Bowyer, P. (2008). Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter. *Remote Sensing of Environment*, 112, 1347–1364. <https://doi.org/10.1016/j.rse.2007.05.020>
- Rangoonwala, A., Enwright, N. M., Ramsey, E. W. III, & Spruce, J. P. (2016). Radar and optical mapping of surge persistence and marsh dieback along the New Jersey Mid-Atlantic coast after Hurricane Sandy. *International Journal of Remote Sensing*, 37, 1692–1713. <https://doi.org/10.1080/01431161.2016.1163748>
- Refsgaard, J. C., van der Sluijs, J. P., Brown, J., & van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, 29, 1586–1597. <https://doi.org/10.1016/j.advwatres.2005.11.013>
- Refsgaard, J. C., van der Sluijs, J. P., Hojberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software*, 22, 1543–1556. <https://doi.org/10.1016/j.envsoft.2007.02.004>
- Reichle, R. H. (2008). Data assimilation methods in the Earth sciences. *Advances in Water Resources*, 31, 1411–1418. <https://doi.org/10.1016/j.advwatres.2008.01.001>
- Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., & Quagan, S. (2017). Reviews and synthesis: Systematic earth observations for use in terrestrial carbon cycle data assimilation systems. *Biogeosciences*, 14, 3401–3429. <https://doi.org/10.5194/bg-14-3401-2017>
- Sehgal, V. K., Chakraborty, D., & Sahoo, R. N. (2016). Inversion of radiative transfer model for retrieval of wheat biophysical parameters from broadband reflectance measurements. *Information Processing in Agriculture*, 3, 107–118. <https://doi.org/10.1016/j.inpa.2016.04.001>
- Shi, Y., & Song, L. (2015). Spatial downscaling of monthly TRMM precipitation based on EVI and other geospatial variables over the Tibetan plateau from 2001 to 2012. *Mountain Research and Development*, 35, 180–194. <https://doi.org/10.1659/MRD-JOURNAL-D-14-00119.1>
- Stow, C. A., Jolliff, J., Doney, S. C., Allen, J. I., Friedrichs, M. A. M., Rose, K. A., & Wallhead, P. (2009). Skill assessment for coupled biological/physical models of marine systems. *Journal of Marine Systems*, 76, 4–15. <https://doi.org/10.1016/j.jmarsys.2008.03.011>
- Strauch, M., & Volk, M. (2013). SWAT plant growth modification for improved modeling of perennial vegetation in the

- tropics. *Ecological Modelling*, 269, 98–112. <https://doi.org/10.1016/j.ecolmodel.2013.08.013>
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and inter-comparisons. *Reviews of Geophysics*, 56, 79–107. <https://doi.org/10.1002/2017RG000574>
- Terzago, S., Palazzi, E., & von Hardenberg, J. (2018). Stochastic downscaling of precipitation in complex orography: A simple method to reproduce a realistic fine-scale climatology. *Natural Hazards Earth System Sciences Discussions*, <https://doi.org/10.5194/nhess-2017-452>
- Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Running, S. W., ... Ahl, D. E. (2006). Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sensing of Environment*, 102, 282–292. <https://doi.org/10.1016/j.rse.2006.02.017>
- Uusitalo, L., Lehtikoinen, A., Helle, I., & Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*, 63, 24–31. <https://doi.org/10.1016/j.envsoft.2014.09.017>
- van der Kwast, J., Timmermans, W., Gieske, A., Su, Z., Oliso, A., Jia, L. & de Jong, S. (2009). Evaluation of the Surface Energy Balance System (SEBS) applied to ASTER imagery with flux-measurements at the SPARC 2004 site (Barrax, Spain). *Hydrology and Earth System Sciences*, 13, 1337–1347. <https://doi.org/10.5194/hess-13-1337-2009>
- Waller, J. A., Dance, S. L., Lawless, A. S., & Nichols, N. K. (2014). Estimating correlated observation error statistics using an ensemble transform Kalman filter. *Tellus A: Dynamic Meteorology and Oceanography*, 66, 23294. <https://doi.org/10.3402/tellusa.v66.23294>
- Wanders, N., Karssenberg, D., de Roo, A., de Jong, S. M., & Bierkens, M. F. P. (2012). Observation uncertainty of satellite soil moisture products determined with physically-based modeling. *Remote Sensing of Environment*, 127, 341–356. <https://doi.org/10.1016/j.rse.2012.09.004>
- Weng, Q., Fu, P., & Gao, W. (2014). Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sensing of Environment*, 145, 55–67. <https://doi.org/10.1016/j.rse.2014.02.003>
- Widłowski, J. L., Pinty, B., Clerici, M., Dai, Y., de Kauwe, M., Kallel, A., ... Olchev, A. (2011). RAMI4PILPS: An intercomparison of formulations for the partitioning of solar radiation in land surface models. *Journal of Geophysical Research*, 116, G02019.
- Yan, K., Park, T., Yan, G., Chen, C., Yang, B., & Myneni, R. B. (2016). Evaluation of MODIS LAI/FPAR product collection 6. Part 1: Consistency and improvements. *Remote Sensing*, 8, 359. <https://doi.org/10.3390/rs8050359>
- Yang, G., Weng, Q., Pu, R., Gao, F., Sun, C., Li, H., & Zhao, C. (2016). Evaluation of ASTER-like daily land surface temperature by fusing ASTER and MODIS data during the HiWATER-MUSOEXE. *Remote Sensing*, 8, 75. <https://doi.org/10.3390/rs8010075>
- Zhang, S., Shi, J., & Dou, Y. (2012). A soil moisture assimilation scheme based on the microwave Land Emissivity Model and the Community Land Model. *International Journal of Remote Sensing*, 33, 2770–2797. <https://doi.org/10.1080/01431161.2011.620032>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Pasetto D, Arenas-Castro S, Bustamante J, et al. Integration of satellite remote sensing data in ecosystem modelling at local scales: Practices and trends. *Methods Ecol Evol*. 2018;9:1810–1821. <https://doi.org/10.1111/2041-210X.13018>