ARTICLE IN PRESS
RSE-09354; No of Pages 16
Remote Sensing of Environment xxx (2015) xxx–xxx
Contents lists available at ScienceDirect
Remote Sensing of Environment
journal homepage: www.elsevier.com/locate/rse

Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction

Souhail Boussetta, Gianpaolo Balsamo, Emanuel Dutra, Anton Beljaars, Clement Albergel

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG29AX, UK

ARTICLE INFO

Article history:
Received 12 November 2014
Received in revised form 3 March 2015
Accepted 13 March 2015
Available online xxxx

Keywords:
Near Real Time Leaf Area Index
Near real time albedo
Satellite data assimilation
Surface fluxes
Surface-atmosphere interaction

ABSTRACT

The vegetation state can have a prominent influence on the global energy, water and carbon cycles. This has been particularly evident during extreme conditions in recent years (e.g. Europe 2003 and Russia 2010 heat waves, Horn of Africa 2010 drought, and Australia 2010 drought recovery). Weather parameters are sensitive to the vegetation state and particularly to albedo and Leaf Area Index (LAI) that controls the partitioning of the surface energy fluxes into latent and sensible fluxes, and the development of planetary boundary conditions and clouds. An optimal interpolation analysis of a satellite-based surface albedo and LAI is performed through the combination of satellite observations and derived climatologies, depending on their associated errors. The final analysis products have smoother temporal evolution than the direct observations, which makes them more appropriate for environmental and numerical weather prediction.

The impact of assimilating these near-real-time (NRT) products within the land surface scheme of the European Centre of Medium-Range Weather Forecasts (ECMWF) is evaluated for anomalous years. It is shown that: (i) the assimilation of these products enables detecting/monitoring extreme climate conditions where the LAI anomaly could reach more than 50% and in wet years albedo anomaly could reach 10% , (ii) extreme NRT LAI anomalies have a strong impact on the surface fluxes, while for the albedo, which has a smaller inter-annual variability, the impact on surface fluxes is small, (iii) neutral to slightly better agreement with in-situ surface soil moisture observations and surface energy and CO_2 fluxes from eddy-covariance towers is obtained, and (iv) in forecast using a land–atmosphere coupled system, the assimilation of NRT LAI reduces the near-surface air temperature and humidity errors both in wet and dry cases, while NRT albedo has a small impact, mainly in wet cases (when albedo anomalies are more noticeable).

1. Introduction

Owing to their influence on the partitioning of energy, mass and momentum fluxes between the land surface and the atmosphere, land surface processes have been shown to substantially impact weather forecasting at short and medium rages (Beljaars, Viterbo, Miller, & Betts, 1996; Koster et al., 2010; Rowell & Blondin, 1990). The impact of land surface processes also extend over different spatial and time scales, and can affect long-term climate projections (Avissar & Liu, 1996; Balsamo et al., 2009; Betts, Ball, Beljaars, Miller, & Viterbo, 1996; Boussetta, Koike, Kun, Graf, & Pathmathevan, 2008; Sellers et al., 1996; Xue, Fennessy, & Sellers, 1996).

As a fundamental component of a land surface model (LSM), the vegetation layer plays a crucial role in the land–atmosphere exchanges. The vegetation contributes to the evaporation through the plant transpiration and direct evaporation of the plant-intercepted precipitation. In addition, it affects the available surface energy through the radiative transfer within the canopy by modifying the surface albedo (Deardorff, 1978).

In most LSMs, the Leaf Area Index (LAI), is used as an indicator of the vegetation state (e.g. greening, mature, senescent, dormant). Traditionally, the LAI was represented through look-up tables dependant on the vegetation type (Viterbo & Beljaars, 1995). Although its spatial variation was commonly specified according to the biome types, the temporal variation of the LAI was often neglected and sometimes climatological seasonality was introduced together with other major changes. However, the impact of the LAI seasonality could not be assessed (Dorman & Sellers, 1989; Giard & Bazile, 2000). With the advent of LAI that is based on satellite observations, the impact of these products within LSMs has been tested at different temporal and spatial scales. By using a satellite-based climatology of LAI within a mesoscale numerical weather prediction (NWP) model, Knote, Bonafe, and DI Guiseppe (2009) showed that more realistic LAI information is able to improve the short-range forecast scores of lower-level variables but not their biases. Other studies focused on global circulation models (GCM) and the implications of introducing observed seasonally-varying LAI on the simulations. These studies generally discussed the impact of LAI seasonal (Lawrence & Slingo, 2004; Van den Hurk, Viterbo, & Los, 2003)

http://dx.doi.org/10.1016/j.rse.2015.03.009
0034-4257/

Please cite this article as: Boussetta, S., et al., Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction, *Remote Sensing of Environment* (2015), http://dx.doi.org/10.1016/j.rse.2015.03.009

and inter-annual (Guillevic et al., 2002) variability in terms of annual cycle of the hydrological fluxes. They showed that seasonal LAI can have a non-negligible impact on the seasonality of the surface evaporation and precipitation over land. Recently, Boussetta, Balsamo, Beljaars, Kral, and Jarlan (2013a) introduced a monthly climatology for LAI based on a MODIS satellite product (Myneni et al., 2002) within the European Centre for Medium-Range Weather Forecast (ECMWF) Integrated Forecasting System (IFS) to replace the fixed maximum LAI previously used. They showed that this results in a reduction of near-surface temperature errors in the tropical and mid-latitude areas, especially during spring and summer seasons. However, the direct impact of the use of near-real-time (NRT) LAI observations on surface fluxes and screen-level variables was not investigated.

In addition, the surface albedo was shown to be one of the important parameters that controls the land surface energy balance and subsequently affects the atmospheric boundary layer through the surface radiative balance (Pielke & Avissar, 1990). Initially, NWP systems relied on surface albedo derived from soil type and vegetation type maps, neglecting its seasonal and inter-annual variabilities. Preuss and Geleyn (1980) showed that introducing a seasonally-varying satellite-based climatology of surface albedo into the ECMWF IFS has a small impact on medium-range forecasts and a potentially larger impact at longer ranges. While at a regional scale, Berbet and Costa (2003) showed that over Amazonia, most of the spatial and seasonal variability in the simulated climate after a tropical deforestation can be explained by the variability in surface albedo. These previous studies trigger the question whether the use of NRT surface albedo has a significant impact on the near-surface atmosphere in comparison to climatological data.

Recently, the Copernicus Global Land products (GEOV1) of surface albedo and LAI based on observations from the VEGETATION sensor on board SPOT satellite (Baret et al., 2007; Verger, Baret, & Weiss, 2011) have become available in NRT with a foreseen operationally-maintained chain. However, the direct use of these products within a NWP system is not possible without quality checks given the spatial and temporal discontinuities they may contain.

In this study, we explore the assimilation of NRT LAI and surface albedo products in a NWP system using optimal interpolation analysis. We evaluate the impact of the NRT analysis products on surface fluxes and surface soil moisture derived from land-only offline simulations of the ECMWF land surface model CHTESSEL (Balsamo et al., 2009, 2011; Boussetta et al., 2013a,2013b; Dutra et al., 2010; Van den Hurk & Viterbo, 2003; Viterbo & Beljaars, 1995; Viterbo, Beljaars, Mahouf, & Teixeira, 1999). The evaluation is then extended to the near-surface air temperature, humidity and precipitation derived from coupled land–atmosphere simulations using the fully-coupled IFS.

2. The analysis procedure

The purpose of the analysis is to provide an optimal and complete initial condition to the forecasting system based on available information. Optimal is used here in the sense that different pieces of LAI and Albedo information (climatology, and NRT information) are combined and weighted on the basis of their estimated errors. To obtain such an analysis, the observations are first processed and quality checked, a 10-day climatology is created by averaging over 14 years, and finally the NRT time data is combined with the climatology, to obtain a complete time series. The data products are described in subsection 2.1, the procedure to build the climate is documented in subsection 2.2 and the final analysis procedure is outlined in subsection 2.3.

2.1. The Copernicus Global Land products (GEOV1)

The GEOV1 LAI and albedo products are based on observations from the VEGETATION sensor on board SPOT satellite. They are produced every 10 days using a composite observation from a 30 days moving window at 1/112^{o} spatial resolution (about 1 km at the Equator) with a global coverage. Each GEOV1 product is provided with its associated error measure σ_o. To take advantage of previous algorithmic experience and existing LAI products, the "best-performing" LAI data (Garrigues et al., 2008) were combined and then used to train a neural network system (Verger et al., 2011). The combined LAI data are the CYCLOPES-V3.1 (Baret et al., 2007) and the collection 5 of MODIS LAI (Myneni et al., 2002). This fusion benefits from the good performance of MODIS LAI for high values and CYCLOPES-V3.1 LAI at low values. After being trained with the fused data, the neural network system is then fed with the atmospherically-corrected reflectances in red, near-infrared, and shortwave-infrared bands from VEGETATION as well as the solar zenith angles and the satellite overpass timing which results into the GEOV1 LAI (Baret et al., 2013).

The GEOV1 surface albedo is also based on observations from the VEGETATION sensor and its derivation follows Geiger and Samain (2004). The method includes cloud screening (Hagolle et al., 2004), atmospheric correction (Rahman & Dedieu, 1994), directional reflectance normalization (Roujean, Leroy, & Deschamps, 1992), and albedo determination for the different integration angles (direct and diffuse) and spectral intervals (visible, near-infrared, and broadband).

Under the Copernicus Global Land framework, the GEOV1 products were validated by analysing their spatial and temporal continuity and consistency as well as their accuracy at the global and regional scales against other global products and the BELMANIP2 sites network (Benchmark Land Multi-site Analysis and Inter-comparison of Products). The conclusion of this validation was that the GEOV1 products are of good quality, show consistent temporal and spatial distributions, and have reasonable accuracy which can meet the requirements for use within LSMs (Camacho, Cernicharo, Lacaze, Baret, & Weiss, 2013).

2.2. Derivation of climatological series

Before processing the climatological product, unreliable retrievals have been discarded from the analysis using the quality flag (QA). In particular: dead detectors, significant clouds and/or snow contaminated pixels, and failure of the radiative transfer model due to problems other than geometry have been filtered out during this pre-processing. Afterwards, the 1-km products were aggregated to 10-km resolution through a nine-by-nine-point spatial smoothing. The 10-km value is computed when more than 30% of the 1-km products at the grid point scale have not been flagged and a further snow-free screening is performed on the data. Then a first version of the climatological time series is obtained by averaging data from 1999 to 2012 (ALB_cv1/LAI_cv1).

The first version of the climatological time-series still contains gaps, especially in snow-covered high latitude regions. To overcome this deficiency, a second version of the climatological time series (ALB_cv2/LAI_cv2) is generated by spatially filling the data gaps with values from 36 10-daily "self-derived" look-up tables of LAI and albedos for each vegetation type. The look-up tables are derived through stratification of ALB_cv1/LAI_cv1 by vegetation type based on a 90% vegetation cover threshold for each type. In CHTESSEL, the land use classification follows from the Global Land Cover Characteristics (GLCC) data (Loveland et al., 2000) and use is made of the Biosphere-Atmosphere Transfer Scheme (BATS) classification to assign the vegetation types.

Finally, a three-point temporal smoothing is applied to this second version climatological data to obtain a final version (ALB_c/LAI_c). These data are then re-projected and interpolated to a target model simulation grid, together with their associated error σ_c in order to be used in the ECMWF model. In this study the target model is run at T511 reduced Gaussian grid horizontal resolution which is about 40x40 km.

2.3. The optimal interpolation method

The main objective of data assimilation is to optimize the use of observational data to get the best estimate from all available information. It usually attempts to combine data from different sources in an optimal

Please cite this article as: Boussetta, S., et al., Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction, *Remote Sensing of Environment* (2015), http://dx.doi.org/10.1016/j.rse.2015.03.009