Remote Sensing of Environment 212 (2018) 103–113

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Determination of uncertainty characteristics for the satellite data-based estimation of fractional snow cover

Miia Salminen[a,*], Jouni Pulliainen[a], Sari Metsämäki[b], Jaakko Ikonen[a], Kirsikka Heinilä[b], Kari Luojus[a]

[a] Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland
[b] Finnish Environment Institute, P.O. Box 140, FI-00251 Helsinki, Finland

ARTICLE INFO

Keywords:
Fractional snow cover
Reflectance
Seasonal snow
Optical
Remote sensing
Snow mapping
Accuracy

ABSTRACT

We developed a methodology to evaluate quantitative uncertainty characteristics of satellite data retrievals including the contribution of systematic error and statistical error. This is introduced by assessing the total product error of optical, satellite data-based, Fractional Snow Cover (*FSC*) estimates. Here the *FSC* estimation is based on an algorithm allowing the consideration of the effect of different error sources; a semi-empirical reflectance model describing the relationship of the observed reflectance and *FSC* through several variables and parameters. We assume that after the statistical error analysis, the remaining portion of total product error arises due to systematic factors. Hence, first we define a statistical error component through the theory of error propagation, and then estimate the total product error (PE) by using *in situ* observations on *FSC*, and finally derive the systematic error from these two error components. The experimental approach for estimating PE is conducted through an analysis of the observed estimation errors (*i.e.* residuals) in the GlobSnow Snow Extent (SE) v2.1 products on *FSC*. In practice, independent *in situ* snow course observations from Finland on *FSC* are compared to corresponding satellite *FSC* estimates to quantify the residuals. The approach is then illustrated for an extended region of corresponding European boreal forest. Our results show that the total *PE* in the GlobSnow *FSC* product is significantly higher than the originally provided statistical error. This is due to deficiencies in the parameterization of the applied forward modelling approach, in particular in the consideration of the forest canopy effects.

1. Introduction

Estimation of Snow Extent (SE) through optical Earth Observation (EO) data is particularly advantageous during the end of the melting season. Passive microwave-based SE detection methods typically fail to detect patchy wet snow areas, whereas optical observations have been shown to provide far more reliable SE estimates (Dietz et al., 2012a; Frei et al., 2012). However, cloud cover and lack of sun light (or high sun zenith angles) during winter hamper the use of optical instruments. Together the two complementary satellite remote sensing methods enable both temporally frequent as well as high resolution spatial retrievals of snow cover necessary for global and regional scale hydrological and climate models (Nolin, 2010; Dietz et al., 2012b). In addition to annual regional and global snow monitoring, acquiring long decadal scale time-series of snow observations has high relevance for climate research (*e.g.* for building consistent climate data records), provided that the accuracy of the different snow products is known and the products exhibit sufficient accuracy (Ramsay, 1998; Hall and Riggs, 2007; Helfrich et al., 2007; Nolin, 2010; Brown and Robinson, 2011; Dietz et al., 2012a; Frei et al., 2012; Rittger et al., 2013; Metsämäki et al., 2015; Hori et al., 2017). Uncertainties and errors contained in different snow products can be also inherited by operational hydrological and meteorological models when snow information is assimilated or used as input to model forecasts (Dozier et al., 2009; Rittger et al., 2013). Therefore, the provision of quantitative uncertainty estimates for each remote sensing product, such as Fractional Snow Cover (*FSC*) within pixel is highly important. Note that the real value of *FSC* is denoted here by italics letters, whereas its estimate is FSC. Often, information on uncertainties is in fact a requirement for data assimilation schemes, for both hydrological and meteorological models. Previously *e.g.* Hall and Riggs (2007), Rittger et al. (2013) and Metsämäki et al. (2005, 2012, 2015) have discussed and quantified the uncertainty in measurements of FSC.

There are few extensively used global snow extent data sets, such as

* Corresponding author.
E-mail address: miia.salminen@fmi.fi (M. Salminen).

https://doi.org/10.1016/j.rse.2018.04.038
Received 4 May 2017; Received in revised form 20 April 2018; Accepted 21 April 2018

NOAA Climate Data Record (CDR) of Northern Hemisphere Snow Cover Extent (Robinson et al., 2012), NASA MODIS (Moderate Resolution Imaging Spectroradiometer) and VIIRS (Visible Infrared Imaging Radiometer Suite) snow products (https://modis-snow-ice.gsfc.nasa.gov/) as well as ESA's GlobSnow fractional snow cover products (www.globsnow.info). MODIS/VIIRS and GlobSnow databases include validated daily FSC products, of which the GlobSnow database also provides pixel-wise uncertainty for each FSC estimate (so called statistical error) (Metsämäki et al., 2015; Hall and Riggs, 2007). Further characterization and quantification of uncertainty is the topic of this research, focusing on the analysis of all components of FSC retrieval error. This is carried out for the GlobSnow FSC product that applies the -*SCAmod* method based on the inversion of a semi-empirical forward model of landscape reflectance (Metsämäki et al., 2015).

The FSC defines the patchiness of snow cover by a value of FSC ranging from 0 to 1 (or from 0 to 100%-units). Metsämäki et al. (2012) showed that *SCAmod* algorithm is feasible for continental-scale snow mapping and is relatively accurate also in forested regions. For the current optical FSC-methods, the presence of forest cover in particular is typically problematic and usually causing an underestimation of the area covered by snow (Salomonson and Appel, 2004, 2006; Nolin, 2010; Metsämäki et al., 2012; Rittger et al., 2013). The success of the *SCAmod* algorithm in forested areas is related to the utilization of the *apparent forest transmissivity* (derived from space borne reflectance observations under full snow cover) that enables the consideration of the masking effect of forest canopy. Since the *SCAmod* is in practice the only widely used FSC method that is based on the forward modelling of (Top-of-Atmosphere (TOA)/scene) reflectance, here we examine its uncertainty characteristics. Currently the *SCAmod*-based GlobSnow products include a statistical error layer (Metsämäki et al., 2015), but they do not consider systematic errors, and thus, the provided error does not describe the observed residuals (*RMSE*) very well.

The statistical accuracy of FSC information can be investigated by analysing the statistics of factors that affect the space-borne observed TOA reflectance and thus, the FSC retrieval performance of *SCAmod*. These factors include the reflectance signatures of the various constituents of snow covered and snow-free landscapes. With *SCAmod*, these constituents comprise of the effects of forest (canopy closure and transmissivity), as well as reflectance of snow-free and melting snow covered ground. The statistical (random) error (which is unbiased by nature and corresponds to *Precision*) is obtained by applying an error propagation analysis to the inverted reflectance model (Metsämäki et al., 2005, 2012, 2015; Salminen et al., 2013). However, the total error in FSC estimation is composed of statistical error and systematic error (caused by *Bias*), and the goal of this study is to define both of them. PE is determined using *in situ* validation to calculate the residuals; PE is then assumed to be a sum of (squared) systematic error and statistical error. This implies that the systematic error can be determined by subtracting statistical error component from PE. This approach is considered a novelty in this study. Systematic error is demonstrated here by analysing satellite-data (European Space Agency's Envisat AATSR (Advanced Along-Track Scanning Radiometer) from 2003 until 2012) derived GlobSnow FSC products and their related FSC retrieval uncertainty estimates (*i.e.* statistical errors) for each studied pixel together with the *in situ* analysis. The factors typically affecting the systematic error are inaccuracies or excluded factors in the parametrization of the applied inverse model, *e.g.* the solar zenith angle, sensor zenith angle and topography. In addition, the levels and variations in atmospheric transmittance and reflectivity have an effect on TOA reflectance. These issues are only indirectly considered here, since part of the analyses are carried out using TOA reflectance observations (additionally, surface reflectance values are converted to TOA reflectances using constant atmospheric parameters when appropriate). As the different systematic error components (as well as their proportions and magnitudes) are intangible, we start off with the statistical analysis and assume that the remaining component to the observed residuals stems from the systematic factors.

To summarize, we analyse the behaviour of the FSC-product error using first the theory of error propagation for defining the statistical component, and then utilize *in situ* regional data from Finland for calculating the remaining systematic component (such distributed *in situ* reference data are not available elsewhere). Finally, the error considerations are demonstrated for a larger region of boreal forest. Even though the analyses are performed for a single approach, the *SCAmod* method, the technique is relevant for other quantitative snow mapping approaches given that they are based on the modelling of scene reflectance or provide FSC as a function of observed reflectances (Kaufman et al., 2002; Salomonson and Appel, 2004). Also if a retrieval approach is able to provide a quantitative value of statistical accuracy, the developed method can be used to assess the systematic error component and thereby to yield an estimate for *PE* (total error).

2. Materials

2.1. Data for analysing the systematic error component

Quantitative analyses are carried out for the region of Finland using the GlobSnow Snow Extent (SE) v2.1 products on FSC, generated by using the *SCAmod* algorithm (Metsämäki et al., 2015). As a study area, Finland represents southern, central and northern boreal forests. In Finland, the total forest area of 22 million ha is mostly dominated by conifers (Scots pine, Norway spruce); even though broad-leaved trees (mainly Birch) can be locally dominant (ACIA, 2005). As the same coniferous species are typical for the boreal forest belt as a whole, data from the Finnish study area can be extrapolated to larger regions, basically, for all boreal forests in the Northern Hemisphere with similar landscapes of evergreen conifer dominance (except larch dominated regions in Siberia). Therefore, this investigation can also be considered relevant for the evergreen regions in Russia, Alaska, Canada and Scandinavia between latitudes of roughly 45° and 70° N, where the largest part of the world's boreal forests are located (ACIA, 2005). However, as both the parameterization used in the statistical error analysis and the *in situ* data applied in the analysis of residuals (estimation errors) are mostly conducted for moderately dense forests, the results are not likely to be representative for the most dense forests existing *e.g.* in Siberia. Fig. 1 shows the current GlobSnow FSC product extracted for the limited region of boreal forests of this investigation. The product shows the estimated FSC and the corresponding map of statistical error for the period of April 02 – April 15, 2010.

The applied ground truth data consists of the Finnish Environment Institute's (SYKE) snow course observations that include spatially distributed information on FSC for the deep winter and spring melt period. This unique data set enables the quantitative investigation of residuals in FSC retrieval in case of southern, central and northern boreal forests. The Finnish snow courses are typically 2–4 km long paths where observations are made at constant intervals through various landscapes including forests. FSC is assessed through visual estimation for areas within a radius of 25 m of observers' location from 80 samples along the course.

The employed GlobSnow FSC product is based on Envisat AATSR data from the period of 2003–2011. The product pixel size is $0.01° \times 0.01°$, *i.e.* about 1 km × 1 km. We extracted (non-cloudy) FSC estimates from daily products corresponding to each snow course observation within a time window of three days (only those estimates where FSC product showed values $0\% < FSC < 100\%$ to represent the melting conditions). The mean FSC values of each snow course visit were applied as reference ground truth. The GlobSnow FSC estimate was extracted from the pixel hitting the central point of the snow course, and from eight surrounding pixels. This cluster of 3 × 3 pixels approximately corresponds to the areal extent of snow courses, and thus, FSC values averaged over the cluster are used for the analysis (in case of product provided statistical error of FSC estimate the average is