Remote Sensing of Environment 219 (2018) 62–71

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

New neural network cloud mask algorithm based on radiative transfer simulations

Nan Chen[a,*], Wei Li[a], Charles Gatebe[b], Tomonori Tanikawa[c], Masahiro Hori[d], Rigen Shimada[d], Teruo Aoki[e], Knut Stamnes[a]

[a] Department of Physics, Stevens Institute of Technology, Hoboken, USA
[b] NASA GSFC, Greenbelt, MD, USA
[c] Meteorological Research Institute, Tsukuba, Japan
[d] Japan Aerospace Exploration Agency, Tsukuba, Japan
[e] Okayama University, Okayama, Japan

ARTICLE INFO

Keywords:
Cloud mask
Machine learning
Radiative transfer

2010 MSC:
00-01
99-00

ABSTRACT

Cloud detection and screening constitute critically important first steps required to derive many satellite data products. Traditional threshold-based cloud mask algorithms require a complicated design process and fine tuning for each sensor, and they have difficulties over areas partially covered with snow/ice. Exploiting advances in machine learning techniques and radiative transfer modeling of coupled environmental systems, we have developed a new, threshold-free cloud mask algorithm based on a neural network classifier driven by extensive radiative transfer simulations. Statistical validation results obtained by using collocated CALIOP and MODIS data show that its performance is consistent over different ecosystems and significantly better than the MODIS Cloud Mask (MOD35 C6) during the winter seasons over snow-covered areas in the mid-latitudes. Simulations using a reduced number of satellite channels also show satisfactory results, indicating its flexibility to be configured for different sensors. Compared to threshold-based methods and previous machine-learning approaches, this new cloud mask (i) does not rely on thresholds, (ii) needs fewer satellite channels, (iii) has superior performance during winter seasons in mid-latitude areas, and (iv) can easily be applied to different sensors.

1. Introduction

1.1. Background

A reliable cloud mask is essential for satellite remote sensing of land, ocean, or cryospheric properties. Due to the significant impact of clouds on shortwave and longwave radiation, mis-identification of cloudy pixels as surface or vice versa can significantly affect the quality of any satellite remote sensing product. Traditionally, threshold-based tests have been employed in many cloud mask algorithms. Such algorithms include the Automated Cloud Cover Assessment (ACCA) algorithm (Irish et al., 2006) applied to the Landsat ETM+ sensor, the cloud tests applied in the MOD35 algorithm (Ackerman et al., 2010) for the moderate-resolution imaging spectroradiometer (MODIS) sensor and the Clouds from AVHRR (CLAVR) (Stowe et al., 1999) as well as its extension CLAVR-x algorithm. These algorithms typically use a combination of threshold tests, which employ a number of satellite channels located in the visible (VIS), near infrared (NIR), shortwave infrared (SWIR), and thermal infrared (TIR) wavelength ranges (e.g. MOD35 uses 19 bands — 10 reflectance bands and 9 thermal infrared bands) to detect clouds and snow/ice. The thresholds used in these tests are generally from 1) model simulations, 2) statistics of cloud/clear-sky scenes, and 3) expert experience. New algorithms, such as fmask (Zhu and Woodcock, 2012; Zhu et al., 2015), employ dynamic thresholds derived from object-based cloud and cloud shadow statistics. In our previous work (Chen et al., 2014), a model based dynamic threshold method was developed, tested, and shown to have superior performance compared to the MODIS MOD35 algorithm over the snow-covered Greenland Plateau.

Because of the similarity of cloud and snow/ice optical properties in VIS and near NIR channels, snow detection has always been essential in cloud mask algorithm designs. Indices for mapping snow cover using VIS and SWIR data were developed in the mid-1970s. The Normalized Difference Snow Index (NDSI) was introduced by Hall et al. (1995) to

* Corresponding author.
E-mail address: nchen@stevens.edu (N. Chen).

https://doi.org/10.1016/j.rse.2018.09.029
Received 10 February 2018; Received in revised form 25 September 2018; Accepted 30 September 2018

map snow using MODIS data. Prior to that, Dozier (1987, 1989) used a VIS/SWIR index algorithm to map snow based on Landsat data. Most threshold-based cloud mask algorithms will use NDSI in their processing chain (Ackerman et al., 1998,2010; Irish et al., 2006; Zhu and Woodcock, 2012) for cloud screening, which highlights the importance of snow detection since its accuracy will also affect that of cloud detection.

Enhanced computational power and improvements in machine learning techniques have allowed machine learning algorithms, such as decision trees, logistic regressions, support vector machines, and artificial neural networks, to be used for cloud masking and snow/ice detection. Taravat et al. (2015) used a multi-layer perceptron neural network model to detect clouds in Landsat images. Hollstein et al. (2016) compared several methods, including decision tree, classical Bayesian, random forest, support vector machine, and stochastic gradient descent, applied to Sentinel-2 MultiSpectral Instrument (MSI) images. Hughes and Hayes (2014) used a neural network based method trained with a subset of the United States Geological Survey Landsat Data Continuity Mission (USGS LDCM) Cloud Cover Assessment Data (Scaramuzza et al., 2012) and a comparison with fmask (Zhu and Woodcock, 2012) showed favorable results.

Bayesian methods have shown significant improvements over threshold based methods. Notably, model based Bayesian statistical methods have shown that simulated datasets can be used as a predictor to improve the cloud detection accuracy. Merchant et al. (2005) first applied this method for cloud screening over ocean areas in order to retrieve sea surface temperature. Bulgin et al. (2014), and Bulgin et al. (2018) extended this method to be applied over land areas. In these studies, manually classified datasets were used for validation. An automatic Bayesian classifier, derived using collocated AVHRR and CALIOP data by Heidinger et al. (2012), showed improvements over threshold-based methods and the ability to derive uncertainties in the cloud masking process. The dependence on CALIOP data to derive posterior cloud probability was also introduced in this paper.

Recently, a support vector machine (SVM) approach has been used in the latest CLAUDIA3 algorithm (Ishida et al., 2018). High quality training datasets are essential to machine-learning-based methods and manually-generated datasets such as the ACCA reference dataset (Irish et al., 2006) and the Sentinel-2 MSI dataset constructed by Hollstein et al. (2016) are often used by current machine-learning-based cloud detection schemes. In Ishida et al. (2018), the training dataset for the SVM classification is also selected subjectively from actual satellite measurements by carefully examining the typical surface type and eliminating irregular data.

1.2. Limitations of traditional methods

Traditional threshold-based cloud mask methods still face serious challenges over snow- and ice-covered areas, especially in Arctic and sub-Arctic regions where there are frequent temperature inversions (affecting TIR-based tests) and over mid-latitude regions where the reflected signal is often from pixels with mixed snow and vegetation/soil cover. In order to handle such complicated surface conditions, the threshold-based logic becomes increasingly complex (as can be seen in plates 1–5 of Irish et al., 2006) and a large number of satellite channels is often required. Sometimes these tests will produce conflicting results and additional "clear restoral tests" are needed (Ackerman et al., 2010) to avoid mis-classification. The need to detect possible snow-covered areas also adds uncertainty to the results. As reported by Wang et al. (2008), mis-classifications of snow-covered areas as "cloud" or vice versa are still a serious problem in results produced by traditional threshold-based methods such as the MODIS cloud mask as will be shown in Section 3.

Machine learning methods, on the other hand, generally have no dependence on thresholds and do not rely on detecting snow before cloud screening. However, the dependence on manually-generated datasets has limited the development and operational use of machine learning based algorithms. It is difficult to generate a reliable training dataset due to the large amount of human resources needed to classify hundreds of images with millions of pixels. The limited amount of manually-classified images also makes it hard to cover all possible solar/viewing geometries, which limits the operational use of trained algorithms. Most importantly, manually-classified images are usually available only post-launch. This circumstance impedes pre-launch evaluation of algorithm performance and makes its application to a different sensor difficult.

2. New approach

In this paper, we present a new machine-learning based approach to cloud and snow detection and discrimination to overcome the limits of previous methods. Instead of using manually-generated datasets, we simulate the training dataset needed by machine learning algorithms. Compared to manually-generated training data based on actual measurements, simulated training data have the following advantages:

- There is no need for humans to identify hundreds of images with millions of pixels, which greatly saves human effort.
- The number of training samples can be as large as desired/needed, which can help avoid overfitting problems and be used to fully explore the potential of machine learning techniques.
- The training dataset can cover the full range of possible solar/viewing geometries.
- The algorithm can easily be modified for application to different sensors; only new training datasets are needed.

In order to create such a training dataset, it is necessary to take into account the interaction of incident solar radiation with different types of surfaces, aerosols and clouds. This requirement implies that it is crucially important to have access to a comprehensive radiative transfer model. In order to simulate the reflectance from complex land surfaces, we constructed such a model; the details are provided in the following section.

2.1. Radiative transfer simulations

In order to simulate the light signal received by a satellite instrument, we need to solve the radiative transfer equation (RTE) pertinent for light propagation in the coupled atmosphere-surface system. The diffuse radiance $I(\tau,\theta,\phi)$ at wavelength λ is found by solving the following RTE:

$$\mu\frac{dI(\tau,\theta,\phi)}{d\tau} = I(\tau,\theta,\phi) - \frac{\varpi(\tau)F_0e^{-\tau/\mu_0}}{4\pi}p(\tau,\theta',\phi';\theta_0,\phi_0) - \frac{\varpi(\tau)}{4\pi}\int_0^{2\pi}d\phi'\int_{-1}^{1}d\mu' p(\tau,\theta',\phi';\theta,\phi)I(\tau,\theta',\phi'). \tag{1}$$

Here, F_0 is the incident top-of-the-atmosphere (TOA) solar irradiance (normal to the beam), while the differential optical depth $d\tau = -(\alpha+\beta)dz$, the single scattering albedo $\varpi = \beta/(\alpha+\beta) = \beta/\gamma$, and the scattering phase function $p(\tau,\theta',\phi';\theta,\phi)$ are the inherent optical properties (IOPs) of the scattering/absorbing medium. Note that we have used the Greek letters α, β, and $\gamma = \alpha+\beta$ to denote the absorption, scattering, and extinction coefficients, respectively. θ_0 and ϕ_0 represent solar zenith and azimuth angles, $\mu_0 = \cos\theta_0$; θ' and ϕ' are sensor zenith and azimuth angles prior to a scattering event, and θ and ϕ the corresponding angles after the scattering event, $\mu = \cos\theta$. In our training dataset, the TOA bidirectional reflectance factor (hereafter simply referred to as the reflectance), defined as $R(\tau,\theta,\phi) = \pi I(\tau,\theta,\phi)/F_0 \cos\theta_0$, is simulated using the latest version of the DISORT radiative transfer model (RTM) (DISORT 4.0, Lin et al., 2015; Stamnes et al.,