Available online at www.sciencedirect.com

ScienceDirect

Remote Sensing of Environment 107 (2007) 138–148

Remote Sensing of Environment

www.elsevier.com/locate/rse

A data-mining approach to associating MISR smoke plume heights with MODIS fire measurements

Dominic Mazzoni [a,c,*], Jennifer A. Logan [b,c], David Diner [a,c], Ralph Kahn [a,c], Lingling Tong [b,c], Qinbin Li [a,c]

[a] Jet Propulsion Laboratory, California Institute of Technology, United States
[b] Division of Engineering and Applied Sciences, Harvard University, United States
[c] Department of Computer Sciences, University of Texas, Austin, United States

Received 4 March 2006; received in revised form 11 August 2006; accepted 15 August 2006

Abstract

Satellites provide unique perspectives on aerosol global and regional spatial and temporal distributions, and offer compelling evidence that visibility and air quality are affected by particulate matter transported over long distances. The heights at which emissions are injected into the atmosphere are major factors governing downwind dispersal. In order to better understand the environmental factors determining injection heights of smoke plumes from wildfires, we have developed a prototype system for automatically searching through several years of MISR and MODIS data to locate fires and the associated smoke plumes and to retrieve injection heights and other relevant measurements from them. We are refining this system and assembling a statistical database, aimed at understanding how injection height relates to the fire severity and local weather conditions. In this paper we focus on our working proof-of-concept system that demonstrates how machine-learning and data mining methods aid in processing of massive volumes of satellite data. Automated algorithms for distinguishing smoke from clouds and other aerosols, identifying plumes, and extracting height data are described. Preliminary results are presented from application to MISR and MODIS data collected over North America during the summer of 2004.

Keywords: MISR; MODIS; Smoke plumes; Boreal fires; Injection height; Aerosol transport; Data mining; Support vector machines; Neural networks

1. Introduction

The injection height of smoke plumes from forest fires is a large source of uncertainty in transport models used to predict the effect of emissions from fires on air quality and climate. It is well known that crown fires generate sufficient energy to loft smoke plumes above the boundary layer (Cofer et al., 1996; Lavoue et al., 2000), facilitating long-range transport of gases and particulate matter (e.g., Bertschi et al., 2004; Colarco et al., 2004; Kahn et al., in press). A large fraction of smoke aerosols remain in the near-surface boundary layer, and do not form discrete "plumes" that are the focus of this paper. Emissions that rapidly escape the boundary layer are more likely to contribute to long distance transport.

Case studies have shown that smoke from large boreal fires can be injected to the lower stratosphere by supercell convection (Fromm & Servranckx, 2003). The frequency of high-altitude (and thus long-lifetime) smoke injection has not been quantified systematically (Fromm et al., 2004). It is possible that boreal and mid-latitude fires may become more common in the future as a result of global warming (e.g., Brown et al., 2004; Flannigan et al., 2000). Understanding the impacts of fires on air quality and climate requires, in part, the use of transport models to relate particle and gas emissions to their downstream dispersal. Observations of aerosol injection are necessary to initialize and validate the models and to develop relationships between injection height and local surface and meteorological conditions. To facilitate progress in this area, we are using data

* Corresponding author. Google, Inc., 604 Arizona Ave., Santa Monica, CA, 90401.
E-mail address: dominic@minorninth.com (D. Mazzoni).

0034-4257/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2006.08.014

from the Terra satellite to obtain statistics on the geographic distribution, extent, orientation, and injection height of plumes (Averill et al., 2005). Terra data acquisition began in February 2000, and NASA recently approved an extension of the mission through 2009. This paper describes the specific process we used to automatically find and extract measurements from smoke plumes, using machine learning techniques and custom image analysis algorithms. We are in the process of refining these algorithms, but our initial results demonstrate the utility of the automated approach.

The development work discussed here uses data collected from June to September 2004 over North America. We selected this period for initial study because of record setting fires in Alaska and the adjacent Yukon Territory of Canada. More than 2.6×10^6 ha burned in Alaska and 1.7×10^6 ha in the Yukon Territory (CFS, 2004; NIFC, 2004). Records were also set for the number of days with reduced visibility caused by wildfire smoke, 42 days in Fairbanks, compared to the previous record of 19 days in 1977, as noted by Averill et al. (2005).

Smoke plumes from these fires were intercepted by aircraft on the INTEX-NA field campaign over the United States which took place from July 1 to August 15 (Singh et al., 2002, in press), and enhanced CO was observed by MOPITT (Measurements Of Pollution In The Troposphere) as a continental scale plume over North America (e.g., Pfister et al., 2005). A case of pyroconvection was documented in June over Alaska, with aerosol enhancements observed near the tropopause (Damoah et al., 2006). Knowledge of the injection heights of the emissions from severe fires such as those in 2004 is required for a quantitative assessment of their effects on atmospheric composition. The work described below is a first step towards providing this information, for plumes immediately downwind of fires.

The Terra Multi-angle Imaging SpectroRadiometer (MISR) instrument observes the Earth in reflected sunlight with a 10:30 AM local time equator crossing, and its typical data collection mode is to observe the Earth globally at nine different view zenith angles in four spectral bands (446, 558, 672, and 866 nm) (Diner et al., 1998). The fore–aft cameras are paired in a symmetrical arrangement and acquire images with nominal view angles, relative to the Earth's surface, at 0°, 26.1°, 45.6°, and 70.5°. In its global observing mode, the nadir camera data in all bands, and the red band data of all of the off-nadir cameras are downlinked at the full spatial resolution of the instrument, 275 m. All other channels are averaged on-board to 1.1-km resolution. The swath width observed in common among all nine cameras is about 380 km. Complete coverage between ±82° latitude is obtained every 9 days. Absolute geolocation uncertainty for the nadir camera is about ±45 m, and relative co-registration errors among the nine cameras are typically less than 275 m.

MISR data make possible unique smoke plume identification and characterization approaches. The use of oblique-angle imagery from MISR enhances plume sensitivity because of the longer optical path through the atmosphere, and the combination of multiangle and multispectral information assists in distinguishing smoke from clouds or other types of aerosols (Mazzoni et al., 2006-this issue). Furthermore, automated pattern matching algorithms (Moroney et al., 2002; Muller et al., 2002; Zong et al., 2002) take advantage of the stereoscopic nature of MISR data, and as part of MISR operational data processing determine the geometric parallax (horizontal displacement) that occurs for a given plume due to its altitude above the surface. Pattern matching is aided by the moderately high spatial resolution of MISR imagery and the 14-bit radiometric depth. Photogrammetric algorithms using accurate camera geometric models transform the derived parallaxes into cloud-top or plume-top heights. Using the nadir and near-nadir cameras, as is done for the standard MISR product, the quantized precision of the resulting height field is ±560 m. Height accuracies for low clouds have been validated to a few hundred meters (Naud et al., 2004); since the technique is purely geometric, comparable accuracy is expected for smoke plumes. Altitudes for clouds as well as smoke, dust, and volcanic plumes are routinely retrieved, and reported on a 1.1-km resolution geolocated grid. However, the MISR standard stereo product does not provide a scene classifier along with the heights which identifies whether the observed target is cloud or aerosol. In this paper, we describe the use of a Support Vector Machine approach for providing this classification in conjunction with height extraction from the MISR stereo product.

Data captured by the MODerate-resolution Imaging Spectroradiometer (MODIS) instrument (Barnes et al., 1998), co-located with MISR on the Terra spacecraft, provides invaluable information about fires, including fire occurrence maps and mean radiative power. Burned area products are being developed with the algorithms of Roy et al. (2005) for the entire MODIS data set. MODIS observes the Earth in 36 spectral bands from 0.4 to 14.4 μm. Its scan pattern sees a 2330-km swath, providing near global daily coverage. MODIS's thermal-infrared sensing capabilities give it the ability to detect active fires with high temporal resolution. By combining MISR and MODIS data over hundreds of fires, and supplementing the Terra data with meteorological information, we are developing a statistical database that will make it possible to empirically relate a fire's power and local atmospheric conditions to the resulting smoke plume injection height. While case studies have shown that this technique is feasible (Kahn et al., in press), performing this computation for hundreds of smoke plumes is daunting, not only due to the calculations required, but also due to the relative rarity of smoke plumes in terabytes of satellite images and the challenge of finding them.

2. Method

Our method was developed based on the assumption that it is unreasonably time-consuming to identify and extract data from every smoke plume manually. Using machine learning and data mining techniques, we have developed an approach for extracting smoke plume and fire data automatically from MISR and MODIS imagery and higher-level data products and retrieving several properties including the plume direction and injection height. Since this automated system cannot identify plumes with perfect accuracy, all individual plumes found are