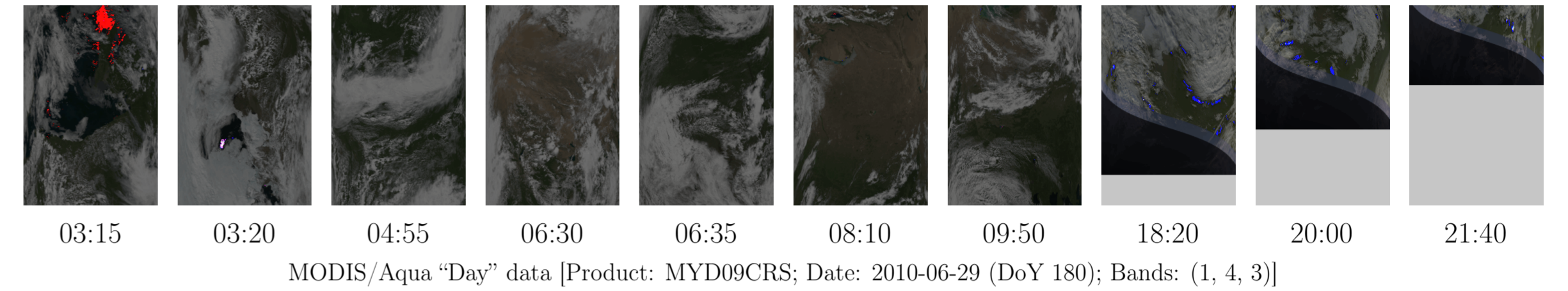
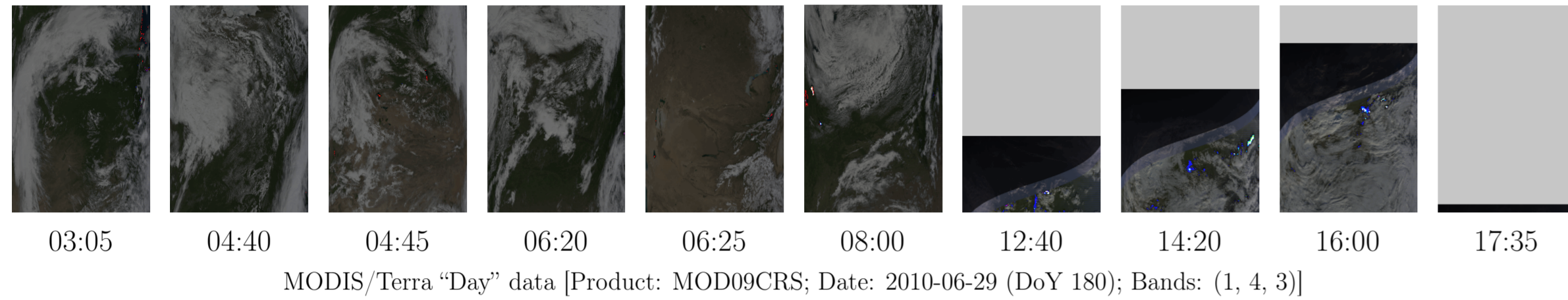


Daily maps of the Bi-directional Reflectance Distribution Function (BRDF) over the Siberian region based on the MODIS data*

Ivan Shmakov, Anatoly Lagutin

mailto:ivan@theory.asu.ru, mailto:lagutin@theory.asu.ru



Abstract

This talk summarizes the approaches and codes used at Altai State University for Bi-directional Reflectance Distribution Function (BRDF) and Albedo retrieval using the MODIS data. The issues related to the deployment of the NASA PGE Collection 5-based data processing system on the Rackspace CloudServers platform are discussed.

1 Introduction

Spacecraft-based sensors are virtually the sole mean of assessing land surface change at *near-daily* time intervals at the *regional* and *global* scales. Apart for the obvious need of such an assessment for the land surface studies and economics, it could also be of a value to the Earth sciences as a whole.

E. g., land surface albedo is a key climatologic property, since it determines the lower boundary condition for the (Earth, Atmosphere) system.

There're a range of effects that influence the land surface reflectances as "seen" by a spacecraft-based sensor, including, but not limited to:

- poor geolocation;
- atmospheric changes;
- (illumination, view) geometry changes.

The MODIS instrument onboard the Terra satellite was the first space-based spectroradiometer which allowed these issues to be addressed. Namely:

- the geolocation's accuracy is verified to be within 50 m (typically), which is just 20% of the MODIS finest resolution (250 m);
- the careful choice of the MODIS' 36 bands have for the first time allowed for accurate atmospheric correction by means of physically-sound retrieval of the key atmospheric variables;
- the availability of the data on the near-daily basis allows for the time series to be accumulated within a 16-day interval and the Bi-directional Reflectance Distribution Function (BRDF) be retrieved.

The currently available MODIS/(Terra + Aqua) BRDF/Albedo products (of the NASA "Level 3" (L3) 500 m MCD43An and 1 km MCD43Bn series), are computed over 16-day windows with such windows starting every 8 days.

However, there're certain uses (such as fire affected area determination, as per Roy et al, 2005) that may presumably benefit from different selection of the time window width or the shift between the windows.

Earlier, we've already explored the possibility of retrieving "daily" BRDF/Albedo maps with PGE Collection 4 software. Within this study we've attempted to deploy the newer, Collection 5 version of it on a Rackspace CloudServers platform.

2 BRDF Model

The albedo retrieval algorithm relies on the approximation of the land surface bidirectional reflectance distribution function (BRDF) by the following half-empirical model:

$$R_s(\Lambda, \theta, \vartheta, \varphi) = \sum_{k=1}^{n=3} f_k(\Lambda) K_k(\theta, \vartheta, \varphi) = f_1(\Lambda) K_{\text{iso}} + f_2(\Lambda) K_{\text{tt}}(\theta, \vartheta, \varphi) + f_3(\Lambda) K_{\text{geo}}(\theta, \vartheta, \varphi), \quad (1)$$

where $K_{\text{iso}} = 1$ is the kernel corresponding to isotropic scattering, and the other kernels are as follows:

$$K_{\text{tt}} = \frac{(\pi/2 - \xi) \cos \xi + \sin \xi}{\cos \theta + \cos \vartheta} - \frac{\pi}{4},$$

$$K_{\text{geo}} = O(\theta, \vartheta, \varphi) - \sec \theta - \sec \vartheta + \frac{1}{2}(1 + \cos \xi) \sec \theta \sec \vartheta, \quad (2)$$

where:

$$O = \max \left(0, \frac{1}{\pi} (t - \sin t \cos t) (\sec \theta + \sec \vartheta) \right),$$

$$\cos t = \max \left(0, \min \left(\frac{2\sqrt{D^2 + (\tan \theta \tan \vartheta \sin \phi)^2}}{\sec \theta + \sec \vartheta}, 1 \right) \right),$$

$$D = \max \left(0, \sqrt{\tan^2 \theta + \tan^2 \vartheta - 2 \tan \theta \tan \vartheta \cos \phi} \right),$$

$$\cos \xi = \cos \theta \cos \vartheta + \sin \theta \sin \vartheta \cos \phi.$$

Minimizing the least-squares error function:

$$e_{\Lambda}^2 = \frac{1}{d} \sum_{l=1}^m \frac{(\rho_s(\Lambda, \theta_l, \vartheta_l, \varphi_l) - R_s(\Lambda, \theta_l, \vartheta_l, \varphi_l))^2}{w_{\Lambda,l}}, \quad (3)$$

the coefficients $f_k(\Lambda)$ could be determined:

$$\sum_{i=1}^n \left(\sum_{l=1}^m \frac{K_{k,l}}{w_{\Lambda,l}} \right) f_i = \sum_{l=1}^m \rho_s(\Lambda, \theta_l, \vartheta_l, \varphi_l) \frac{K_{k,l}}{w_{\Lambda,l}}, \quad (4)$$

$$f_k = \sum_{i=1}^n \left\{ \left(\sum_{l=1}^m \frac{K_{k,l}}{w_{\Lambda,l}} \right)^{-1} \sum_{j=1}^m \frac{K_{i,j}}{w_{\Lambda,j}} \right\}, \quad (5)$$

where: $\left(\sum_{l=1}^m \frac{K_{k,l}}{w_{\Lambda,l}} K_{i,l} \right)^{-1}$ is the inverse of the matrix (4).

Given BRDF, black- $a_{\Lambda}^b(\theta)$ and whitesky a_{Λ}^w albedo could be calculated as follows:

$$a_{\Lambda}^b(\theta) = \int_0^{2\pi} d\varphi \int_0^{\pi/2} d\vartheta \rho_s(\Lambda, \theta, \vartheta, \varphi) \sin \vartheta \cos \vartheta,$$

$$a_{\Lambda}^w = 2 \int_0^{\pi/2} d\theta a_{\Lambda}^b(\theta) \sin \theta \cos \theta.$$

Assuming diffuse skylight to be an isotropic fraction $S(\theta, \tau_{\lambda})$ of total illumination, the actual albedo could be calculated as follows:

$$a_{\Lambda}(\theta) = [1 - S(\theta, \tau_{\lambda})] a_{\Lambda}^b(\theta) + S(\theta, \tau_{\lambda}) a_{\Lambda}^w.$$

3 Data processing

Early on it was decided to focus on the NASA archives' data processing, which implied the use of an Internet connection with a decent (a few Mbit/s or more) bandwidth.

Of the platforms that provides such a connectivity, Rackspace CloudServers was chosen, mainly because of the availability of the per-hour billing option. (Another one that was considered was Amazon EC2. For now, we don't have a compelling reason to prefer one over the other.)

3.1 Preparations

The instances For the preparations and testing, it was decided to order a single 1 GiB RAM, 40 GiB FS space instance (henceforth, *hesby*.) Later on, an additional instance (*utmyg*) was ordered under the same specification.

The OS Debian GNU/Linux 5.0 "Lenny" was chosen as the base OS distribution, mainly because of our years-long experience with it.

Beyond the packages included into the base OS image, the following were also installed (from "Lenny"):

```
busybox, bzip2, debootstrap, ed, less, locales-all,
ncurses-term, ntpdate, rsync, schroot, screen, time,
uuid-runtime and zip.
```

Also, the Squid caching proxy (as of 3.1.3-2~bpo50+1) was installed from the "backports", as well as *debian-backports-keyring*.

The OS: reloaded Of the packages above, *debootstrap* and *schroot* deserve additional explanation. Actually, the "Squeeze" (testing) OS distribution offers additional packages (namely, GCTP, HDF-EOS and HDF-EOS5 libraries) and fixes some known bugs, so it makes sense to use this distribution instead. To get the best of the two worlds, a *child* system was installed on the instance with the following *debootstrap* invocation:

```
# debootstrap --verbose --variant=minbase \
--keep-debootstrap-dir --exclude=login \
--include=autoconf,automake,bind9-host,binutils,
busybox,bzip2,cadaver,dnsutils,ed,emacs23-nox,file,\
gawk,gcc,gforth,gfortran,git-core,hdf4-tools,\
hdf5-tools,iamerican,ispell,less,libc6-i386,\
libdanga-socket-perl,libdigest-sha-perl,libgctp-dev,\
libhdf4-dev,libhdf5-serial-dev,libhdf5-dev,\
libhe5-hdf5-dev,libio-socket-inet6-perl,\
libtest-unit-perl,liburi-perl,libwww-perl,libxml2-dev,\
locales-all,make,man-db,metalink,mg,mr,ncurses-term,\
nd,patch,prolog-el,psmisc,rasqal-utils,redland-utils,\
strace,swi-prolog-nox,tcl8.5,time,tk8.5,tree,zip,\
uuid-runtime,vim-nox,wget,xmllstarlet,xutils-dev,zile,\
squeeze /srv/chroot/$(date -uI) / \
http://http.us.debian.org/debian/
```

(The set of packages to install was mostly modelled after a typical *child* system's installation and includes packages not directly relevant to the task.)

The *schroot.conf* section could be like:

```
[chroot-2010-07-06]
description=Debian Squeeze (testing)
directory=/srv/chroot/2010-07-06
priority=999
groups=users
aliases=squeeze-amd64,squeeze
```

Upon the foundation described above, the payload software is built.

3.2 The processing

The processing chain leading to *MCD43cn* passes through the *MxD09GST* and *MxD09GHK* products (MODIS reflectance, L2G), which are, to the best of our knowledge, not available through the NASA archives.

It was therefore chosen to perform the following steps:

- "de-stripping" (PGE 03; *MxD021KM* and *MxD02HKM* to *MxD1KMDS* and *MxDHKMDS*, respectively).

- atmospheric correction (PGE 11; *MxDsKMDS*, *MxD02QKM*, *MxD03* and *MxD35_L2* to *MxD09*);
- reprojection (PGE 12; *MxD03* to *MxDPT...*; then using the result, *MxD03* and *MxD09* to *MxDMGGAD* and *MxD09G...*, respectively);
- BRDF retrieval (PGE 23; *MxDPT...*, *MxDMGGAD* and *MxD09G...* for both the platforms to *MCD43cn*).

3.3 Performance and resource usage

Example CPU and "wall clock" times for the processing performed are as follows.

3.3.1 2010-06-30Z (hesby)

process	real	user, s	system, s	ratio, %
destriping	6 : 43.45	292.54	35.05	81
MODIS/Terra (10 day + 1 night)				
MOD09...	4 : 35 : 02.	11666.89	183.98	71
File ID: 51022003				
MODPT...	7 : 35.57	402.27	6.40	89
MODMGGAD	0 : 10.22	7.11	1.00	79
MOD09GST	0 : 13.41	4.25	0.67	36
MOD09GHK	1 : 12.24	50.45	5.34	77
MOD09GQK	3 : 14.21	86.78	8.92	49
File ID: 51023003				
MODPT...	7 : 44.20	415.13	6.17	90
MODMGGAD	0 : 11.01	7.76	1.19	81
MOD09GST	0 : 11.01	5.46	3.74	83
MOD09GHK	1 : 20.91	53.07	5.88	72
MOD09GQK	2 : 01.82	91.87	10.64	84
MODIS/Aqua (10 day + 3 night)				
MYD09...	4 : 47 : 03.	13258.66	196.25	78
File ID: 51022003				
MYDPT...	7 : 16.61	408.39	6.72	95
MYDMGGAD	0 : 10.55	7.14	1.44	81
MYD09GST	0 : 13.37	4.38	0.62	37
MYD09GHK	1 : 19.32	58.50	5.28	80
MYD09GQK	2 : 10.76	89.29	8.41	74
File ID: 51023003				
MYDPT...	7 : 21.48	398.76	6.49	91
MYDMGGAD	0 : 11.73	8.52	2.19	91
MYD09GST	0 : 12.86	6.12	3.66	76
MYD09GHK	1 : 32.23	64.48	6.42	76
MYD09GQK	2 : 16.74	100.18	11.36	81