MODIS Vegetation Index Product Series
Collection 5 Change Summary

Kamel Didan, Alfredo Huete
TBRS Lab.,
The University of Arizona
June 29th, 2006

INTRODUCTION

This document details the MODIS VI product series changes implemented in collection 5.0 reprocessing. There are four categories of change:

- Science changes
- Structural changes
- Processing rules changes
- Metadata changes

In designing the refinements for collection 5.0, the emphasis was placed on the following items:

- User requested additions
  - Addition of new parameters
  - Combined some parameters
- Improve the spatial and temporal characteristics
  - Improved the CV-MVC compositing method
  - Dealing with sub-pixel clouds and mislabeled clouds
  - Improve the Aerosol filtering
  - Inland water bodies (this tested approach was canceled)
  - Phased production for improved temporal frequency
- Adoption of internal compression to reduce disk space usage
- New scheme for data reliability classification
- Elimination of some redundancies

SUMMARY OF CHANGES

1. Replaced the NDVI_QA and EVI_QA with one VI_QA layer
2. Restructured the VI_QA data layer to eliminate redundant information
3. Added the composite day of the year output parameter
4. Added pixel reliability output parameter
5. Modified the compositing approach :
   a. When all input days are cloudy we simply use the MVC
   b. The CV-MVC approach was modified to favor smaller view angle
6. Adopted a stricter aerosol quality filter
7. Adopted the adjacent cloud filter
8. Adopted a threshold technique to identify disproportionate surface reflectance over inland water bodies and suspended VI computation
   a. This was later dropped from the algorithm. The information is left in this document to alert the user community of this spatial issue

9. Adopted a phased production approach for Terra and Aqua data streams in order to keep them separate while increasing the temporal frequency of the VI product
10. Added new Metadata to identify Collection 5.0 QA structure
11. Documented all these changes to help end users work with C5 more effectively

COLLECTION 5.0 CHANGES

Science changes

Clouds and Aerosols

An in-house analysis of the 5+ years of MODIS C4 data record revealed residual issues with cloud and aerosol in the final product. This was attributed mainly to mislabeled clouds (where cloudy pixels are labeled clear, and vice versa, clear pixels are labeled cloudy). Moreover, and in collaboration with the MODIS surface reflectance team, we concluded that aerosol correction is not very effective over heavy and average aerosol loads and very poor when climatology data is used. To address the above, we modified our input data filtering procedure in order to reduce the potential of picking problematic data. The new daily input filtering scheme is carried in two separate steps:

1. At the filtering level:

- Data with cloud shadow is separated from data with no shadow
- New internal input data ranking is employed to separate between the different levels of aerosol loads.
- Data adjacent to clouds is now separated from the rest and ranked low

2. At the compositing level

If all the input data was cloudy, or mixed cloud, or adjacent to cloud, we no longer use the CV-MVC, but use the simpler MVC instead. This measure enabled us to identify mislabeled clear data. However, the final output QA is not adjusted and is still labeled based on the input (although we're fairly certain the pixel was mislabeled).

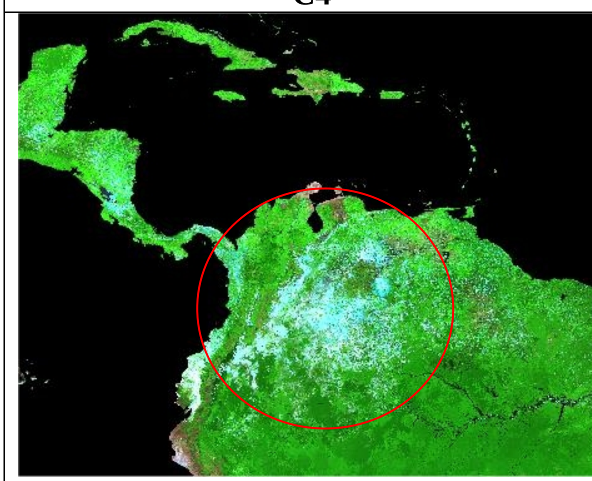
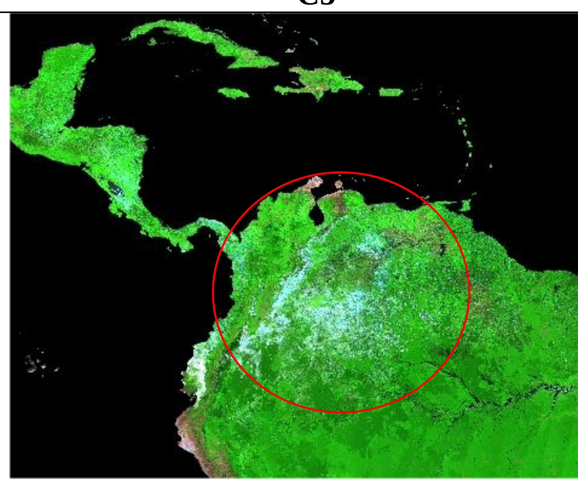
C4	**C5**
 Clouds in C4	 Less clouds in C5

Figure 1: By modifying the compositing rules, we were able to improve the overall data quality and reduce the impact of clouds on the product (clouds are less in C5). The output data still carries the input QA information.

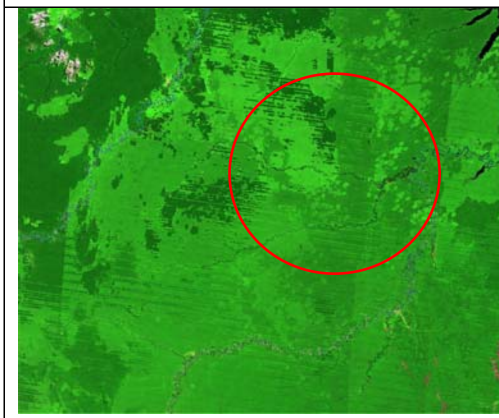
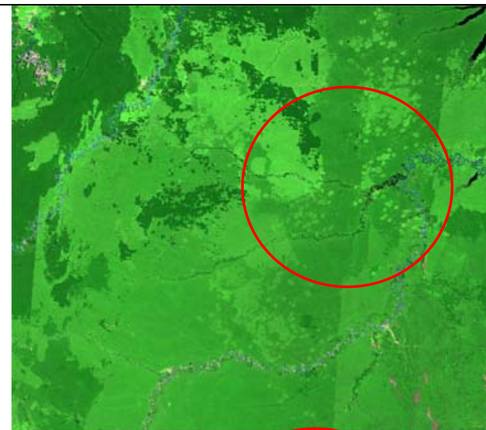
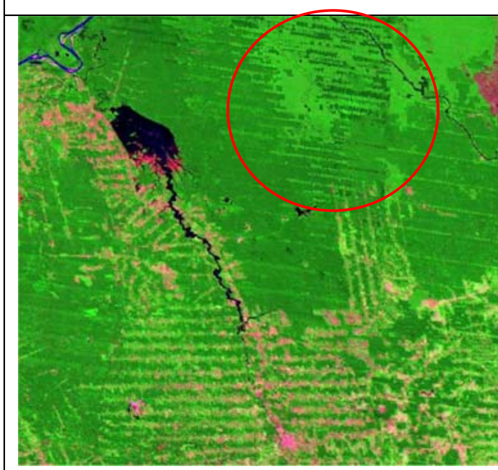
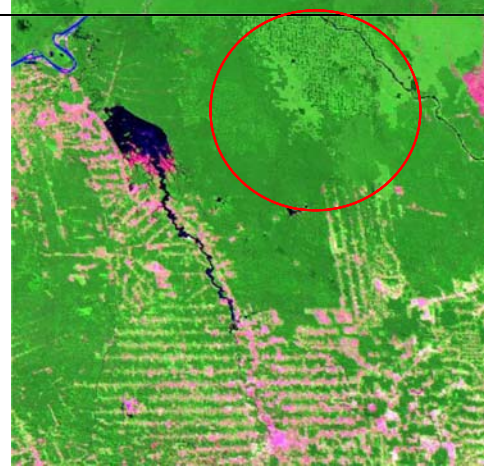
C4	C5
	
 C4 spatial problems	 C5 improved spatial integrity

Figure 2: Aerosol related problems (blockyness, stripes, etc…) were minimized. This resulted from an improved aerosol correction at the surface Reflectance level (MOD09) and a new filtering scheme in the VI algorithm.

Inland water exclusion

In an attempt to partially address the inland water bodies VI problem (high VI values Fig. 3) we evaluated the suspension of VI production over what is labeled as **Shallow and Deep inland water** in the MODIS Land Water mask. An earlier test was restricted to the **Caspian Sea and Lake Victoria** (Fig. 3), and subsequently applied globally. And although, this solution seems to have worked by eliminating unrealistic VI, the global evaluation revealed that this approach was futile because of the poor geolocation accuracy of the MODIS LW mask resulting in large land/wetland areas being excluded (Fig-4). Short of a more accurate LW mask or better Land Surface Reflectance product over water bodies this problem will persists. We decided to lift the exclusion from the algorithm and allow normal processing.

We plan to provide an explanation (through the product's User guide, our SCF website, and EROS/DAAC) and warn the user community of these issues.

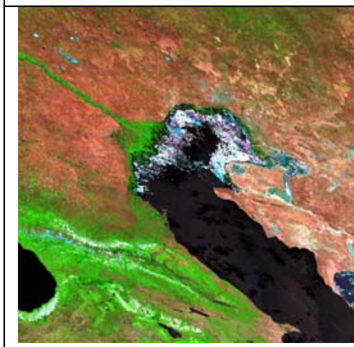
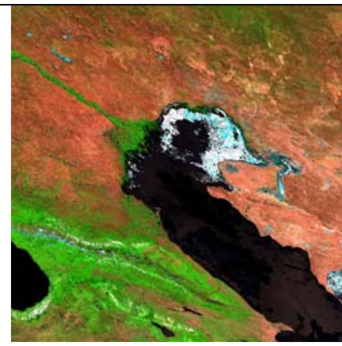
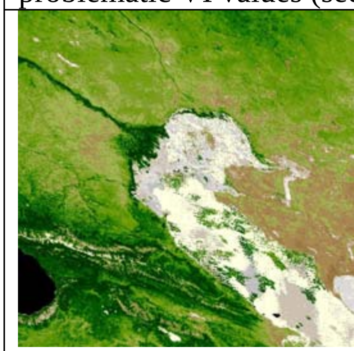
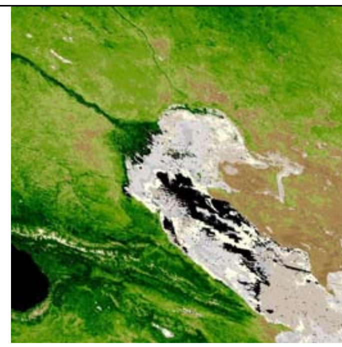
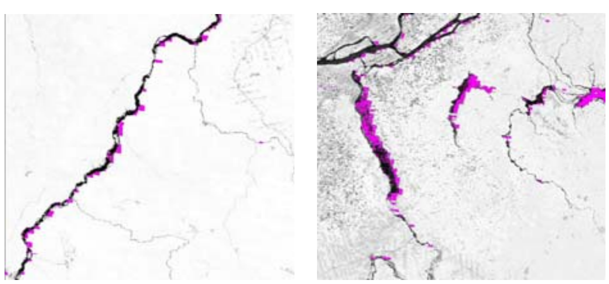
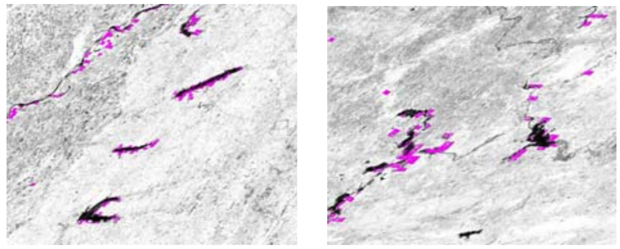
C4	**C5**
 C4 MNR image of the Caspian sea. The image looks fine visually, but the disproportionate red/NIR leads to problematic VI values (see below)	 C5 MNR image of the Caspian seas
 C4 NDVI image of the Caspian sea. The image reveals the problem (Green color over water)	 C5 NDVI image of the Caspian sea. The VI is no longer generated if the final VI data seems out of the normal range (black color)

Figure 3: Earlier implementation of a solution to the inland water bodies problems. Inland water bodies VI issues in C4 were addressed using a series of threshold tests in C5. The VI is no longer produced over such areas if the input data does not satisfy these thresholds.



Amazon, Tile h12v09 (C5 test)



US southeast, Tile h10v05 (C5 test)

***Figure 4:** Issues with the Inland Water bodies' exclusion approach. Valid land and wetlands are being excluded (magenta color) due to the poor geolocation accuracy of the MODIS LW mask*

EVI backup algorithm

In previous versions of the MODIS VI compositing algorithm we used the SAVI index as a backup algorithm for EVI. The SAVI index is used to replace the EVI over cloudy pixel, snow/ice covered pixels, or if the blue band is problematic (out of range). In order to provide for a better continuity with the standard EVI index, and starting collection 5.0 a new index called EVI2 (equation 1) is used instead.

$$EVI2 = 2.5\frac{\rho_{NIR} - \rho_{red}}{1 + \rho_{NIR} + \rho_{red}} \tag{1}$$

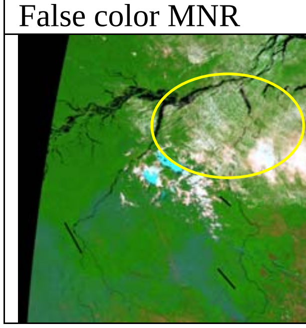
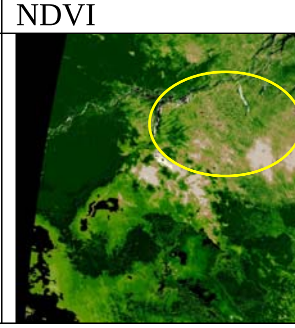
False color MNR	EVI2	Standard EVI	NDVI
			

Figure 5: EVI backup algorithm. Under cloudy conditions (and snow/ice) the EVI index is replaced by the EVI2 index to reduce the affects of the blue band (over bright targets, the blue band as well as the Red&NIR bands exhibit proportionality problems that are aggravated by the atmosphere correction). These problems severely affect the EVI index values (abnormally high).

Spatial anomalies in the VI product (stripes and speckles)

The CV-MVC compositing method, used in the MODIS VI algorithm, was designed to minimize the view angle effects and consequently BRDF related issues. Special care had to be placed on the input data filtering and an NDVI threshold of 30% used to constrain the deviation from the absolute maximum value (MVC). Only the top 2 observations are kept and passed to the CV-MVC algorithm as a safeguard against outliers and contaminated observations. From these top 2 observations, the one with the smaller view angle is selected for output. In doing so, we expected to consistently select observations with the highest NDVI value (characteristics of good atmosphere conditions).

Depending on the daily input data this method has an inherent problem, where adjacent pixels are derived from different days (mosaic of composite days) even when equally good observations (from the same day) are present with smaller view angle but slightly lower NDVI value. This is a direct result of minimizing the view angle while maximizing the NDVI value. Table 1, 2, and 3 illustrate such instances (for adjacent pixels).

DOY	Rank	NDVI	EVI	Red	NIR	Blue	View_Zenith
Selected_194	0	**0.9133**	0.604	0.0159	0.3511	0.0079	4.97
-1	-1	-0.3	-0.3	-0.1	-0.1	-0.1	-100
New_Method_194	0	**0.9133**	0.604	0.0159	0.3511	0.0079	4.97
195	1	0.905	0.4748	0.0129	0.2587	0.0056	64.13
196	3	0.7837	0.8338	0.0427	0.3523	0.0907	32.11
197	1	0.9138	0.4798	0.0116	0.2577	0.006	52.91
198	1	0.8979	0.6295	0.0204	0.3795	0.0101	50.67
199	0	**0.9167**	0.511	0.0121	0.2785	0.0064	35.5
200	1	0.9004	0.644	0.0204	0.3896	0.0105	62.63
201	0	**0.9129**	0.5485	0.014	0.3075	0.0072	9.49
-1	-1	-0.3	-0.3	-0.1	-0.1	-0.1	-100
203	0	**0.9032**	0.6156	0.0185	0.364	0.0096	20.02
204	1	0.9218	0.4811	0.0105	0.2581	0.0046	58.93
205	0	0.8998	0.6316	0.0201	0.3811	0.0097	42.68
206	1	0.9162	0.4865	0.0115	0.2632	0.0052	45.04
207	1	0.8905	0.6388	0.0226	0.3903	0.0116	57.37
208	0	**0.9115**	0.5145	0.0131	0.2832	0.0066	23.47

Table 1: *In this case, DOY 194 (olive highlight) was selected since it provided the highest NDVI and the smallest View angle compared to DOY 199 (light green) which has the absolute highest NDVI but a larger view angle. Notice how close the NDVI values from other days are.*

DOY	Rank	NDVI	EVI	Red	NIR	Blue	View_Zenith
Selected_199	0	**0.9158**	0.5169	0.0124	0.2824	0.0068	35.5
-1	-1	-0.3	-0.3	-0.1	-0.1	-0.1	-100
New_Method_194	0	**0.9146**	0.6039	0.0156	0.35	0.0079	4.97
195	1	0.905	0.4748	0.0129	0.2587	0.0056	64.13
196	3	0.7837	0.8338	0.0427	0.3523	0.0907	32.11
197	1	0.9038	0.4755	0.013	0.2573	0.0068	52.91
198	1	0.8979	0.6292	0.0204	0.3793	0.0101	50.66
199	0	**0.9158**	0.5169	0.0124	0.2824	0.0068	35.5
200	1	0.9004	0.6436	0.0204	0.3896	0.0104	62.63
201	0	**0.9146**	0.5637	0.0142	0.3186	0.0072	9.49
-1	-1	-0.3	-0.3	-0.1	-0.1	-0.1	-100
203	0	**0.9025**	0.6186	0.0188	0.3672	0.0096	20.02

204	1	0.9196	0.4761	0.0107	0.2555	0.0046	58.87
205	0	**0.899**	0.6378	0.0206	0.3876	0.0097	42.67
206	0	**0.9182**	0.5077	0.0118	0.277	0.0056	44.94
207	1	0.8905	0.6388	0.0226	0.3903	0.0116	57.37
208	0	**0.9149**	0.5288	0.013	0.2926	0.0065	23.42

***Table 2**: In this case, DOY 199 was selected since it provided the highest NDVI, the absolute highest NDVI (DOY 206) wasn't selected because it had a larger view angle.*

DOY	Rank	NDVI	EVI	Red	NIR	Blue	View_Zenith
Selected_201	0	0.9171	0.5606	0.0136	0.3146	0.0072	9.49
-1	-1	-0.3	-0.3	-0.1	-0.1	-0.1	-100
New_Method_194	0	**0.9093**	0.5918	0.0163	0.3435	0.0079	4.97
195	1	0.9092	0.4641	0.0119	0.2503	0.005	64.13
196	3	0.6168	0.5897	0.0915	0.3861	0.0915	32.11
197	1	0.9138	0.4858	0.0118	0.2621	0.006	52.91
198	1	0.9012	0.6327	0.0197	0.3794	0.0102	50.63
199	0	**0.9167**	0.511	0.0121	0.2785	0.0064	35.5
200	1	0.8971	0.636	0.0209	0.3857	0.0103	62.63
201	0	0.9171	0.5606	0.0136	0.3146	0.0072	9.49
-1	-1	-0.3	-0.3	-0.1	-0.1	-0.1	-100
203	0	**0.9029**	0.6137	0.0185	0.3628	0.0095	20.02
204	1	0.9218	0.4811	0.0105	0.2581	0.0046	58.93
205	0	**0.9004**	0.6425	0.0204	0.3893	0.0102	42.68
206	1	0.9163	0.4986	0.0117	0.2682	0.007	45.04
207	1	0.8909	0.6392	0.0225	0.3903	0.0116	57.37
208	0	**0.9133**	0.5214	0.013	0.2869	0.0069	23.51

***Table 3:** Similarly, DOY 201 was chosen as opposed to DOY 199. The above three pixels were adjacent, and the corresponding VI product was derived from a different day.*

It is important to realize that the ratio-ing nature of the VI makes the final product look very consistent spatially, even though the pixels came from different days. However, inspecting the resulting surface reflectances and view angle reveals the spatial consistency issue (Look at the Red/NIR and Blue of the selected days, which tends to change considerably from day to day for the same pixel and under what seems to be ideal atmosphere conditions in response to view angle mostly-Fig.6).

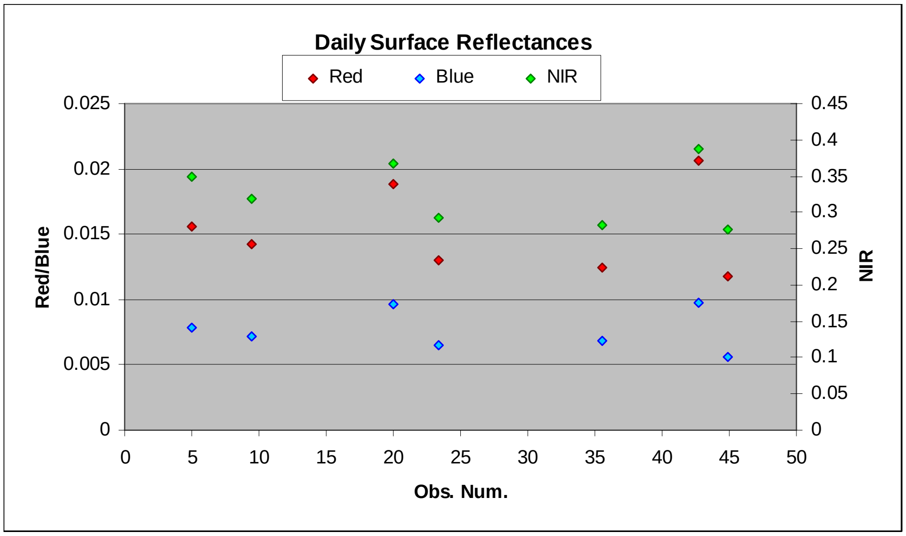


Figure 6: *Change in daily surface reflectance for the same pixel (only what is considered ideal data is plotted).*

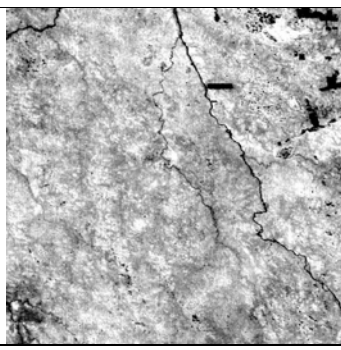
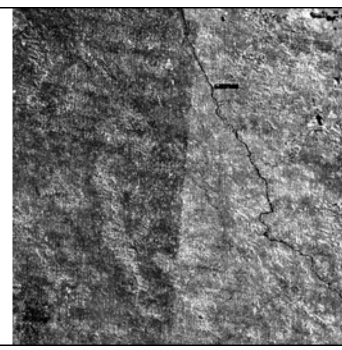
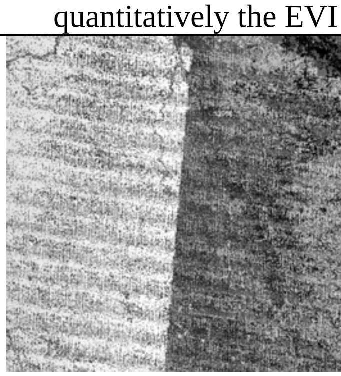
<table>
<tr><td></td><td></td></tr>
<tr><td colspan="2">NDVI EVI
The spatial problems resulting from the different days selected is hardly noticeable in the NDVI. In the case of EVI, the use of the blue band aggravates the problems. However, quantitatively the EVI values are not as bad as the image conveys.</td></tr>
<tr><td>
The view angle is highly variable and suffered the most.</td><td>
The selected composite day (mosaic) resulted from the algorithm sensitivity to the NDVI values combined with the view angle differences from day to day.</td></tr>
</table>

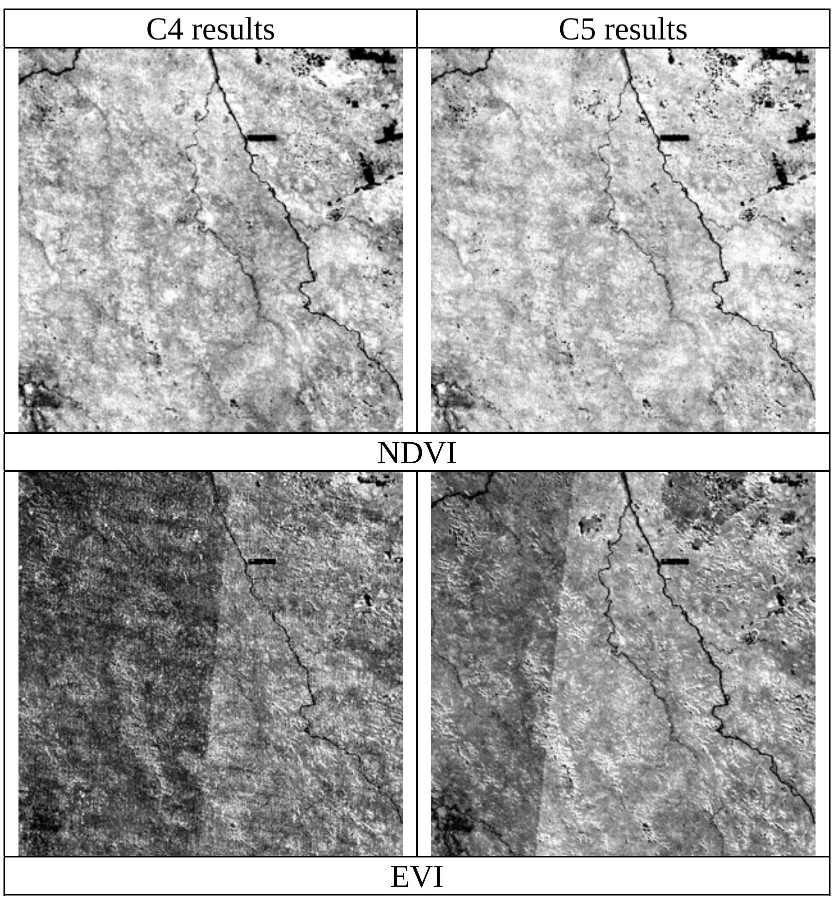
Figure 6: Shows the spatial problems related to the selection of different DOY for adjacent pixels.

Although the above issue is well understood, its manifestation during production is highly erratic owing to the behavior of the daily land surface reflectance and how it affects the VI product.

To partially address this problem we adopted the following changes:

- Modify the CV-MVC to use a contextual approach where the number of observations retained for CV-MVC compositing will vary with the pixel temporal behavior (during the composite cycle) and the number of ideal observations retained for compositing.
- Adopt a <u>**10%**</u> NDVI threshold to constrain the deviation from the absolute highest NDVI value (MVC). This approach permitted the selection of observations with smaller view angle when their corresponding NDVI is not in the top-2 values.

The above changes were implemented and evaluated during C5 testing. This new method resulted in a slight VI decrease (less than 1%) while constraining the view angle to smaller values (Fig. 7).



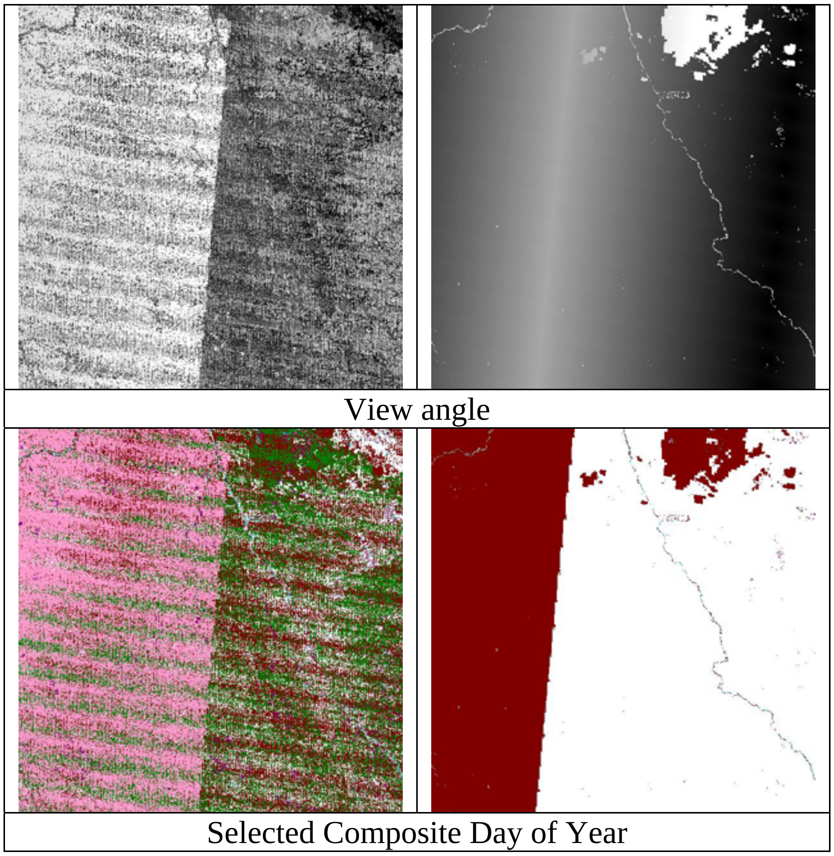


Figure 7: Results of the new CV-MVC compositing method. The improvements are most obvious in the View angle and Day of composite images.

Structural changes

Addition of the Day of composite and per pixel reliability

Two new parameters (Fig. 8) were added to all MODIS VI products:

- Day of composite: Useful in post processing analysis. This parameter is not present in some MODIS VI product due to the way they are generated (Monthly VI, 16 day and monthly CMG).
- Pixel reliability: Useful in post processing analysis and recaps the QA status of the product. This parameter is a simple decimal number that ranks the product into five categories [**Good-Marginal-Snow/Ice-Cloudy-NoData**]. Users can consult this layer instead of working with the VI_QA layer. The VI_QA layer is still present and can be used for detailed analysis.

	1km Product	500m Product	250m Product

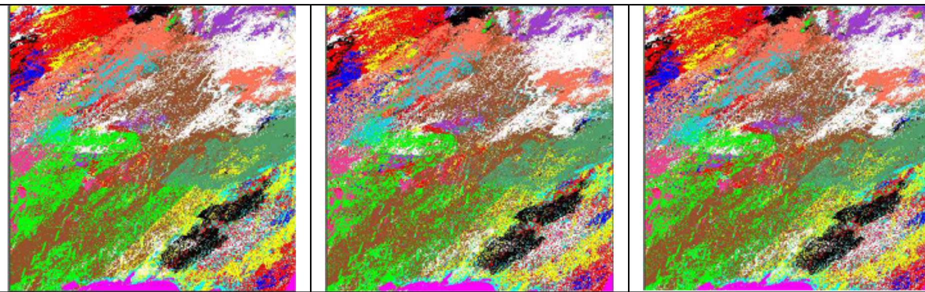
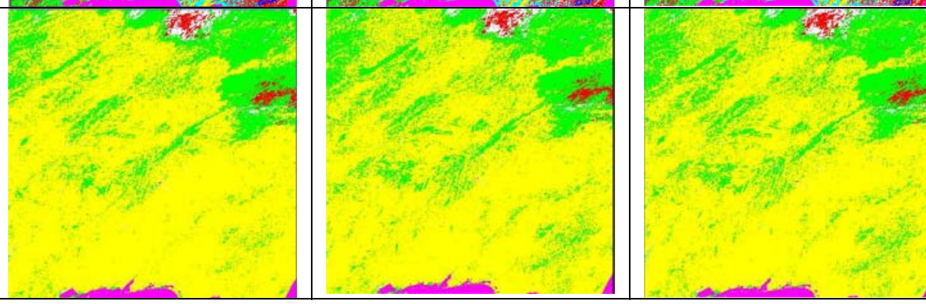
Day of composite	
Pixel Rank Yellow= Good Green = Marginal Red= Snow/Ice White = Cloud	

Figure 8: Example images of the new parameters (DOY and Rank) stored in the VI products.

Modified QA layer

An In-House evaluation of both NDVI and EVI QA information revealed a near identical content on a global basis for the last 5 years and for all resolutions. In order to eliminate this redundancy and reduce the disk space usage in the process, we adopted a single QA layer instead of two separate (NDVI and EVI based layers). This new VI QA layer applies to both NDVI and EVI and will save about 8.33 % disk volume.

In structuring this new VI_QA layer, we also decided to use the full Land water mask (7 categories requiring 3 bits) field (in previous versions the LW mask was resampled to 4 categories in the algorithm using 2 bits). The extra bit needed to code the LW mask resulted from the elimination of the compositing field bit (BIT 15 in previous version). Since the algorithm only uses the CV-MVC/MVC technique, there is no longer a need to indicate the compositing method in the QA layer. Overall, the VI QA layer changed to the following structure:

Bit	Old QA map (C4 and earlier)	New QA map (C5 and later)
0	MODLAND QA	MODLAND QA
1		
2	VI usefulness	VI usefulness
3		
4		
5		
6	Aerosol	Aerosol
7		
8	Adjacent cloud	Adjacent cloud
9	Atm. BRDF correction	Atm. BRDF correction

10	Mixed cloud	Mixed cloud
11	Land/water flag 00: Ocean 01: Coast 10: Wetland 11: Land	Land/water flag 000: Shallow ocean 001: Land (Nothing else but land) 010: Ocean coastlines and lake shorelines 011: Shallow inland water 100: Ephemeral water 101: Deep inland water 110: Moderate or continental ocean 111: Deep ocean
12		
13	Snow/Ice	
14	Shadow	Snow/Ice
15	Composite method	Shadow

Table1: The new VI Quality layer structure. Note the elimination of the compositing method filed (BIT 15) and the expansion of the LandWater mask field BITS 11-13.

Phased production

Currently all multi-day products from Terra and Aqua are generated on the same date, which compels users to either use one data stream or adopt strategies to combining the two streams. In most cases users will simply use one sensor and discard the other due to their identical nature. Moreover, most multi-day algorithms discard large mounts of data already (compositing scheme), especially the selection operator based algorithm (like the VI compositing algorithm). In-House evaluation indicates that the two datasets are highly similar in terms of daily surface reflectance, and higher order products. Further evaluation of the daily cloud cover and other quality aspects of the products also show high similarity. However, we also realized that combining the two data streams (Terra and Aqua) into one might create some serious problems:

- The logistics could be unnecessarily complicated (should one sensor exhibit hardware problems, etc…)
- There still exist slight differences between the two data sets
- Keeping the two data streams separate may be more desirable
- Combining the two data sets to improve quality might not be effective

For the above reasons we investigated the option of optimizing the two data streams without actually combining them. The idea here is to simply increase the temporal frequency of the products with minimum change to the current production. To do so, we decided that phasing Terra and Aqua production seems to be the most reasonable. Under phased production (implemented in collection 5.0 reprocessing) the algorithm will use the same compositing period, but starting at different days (one starts in the middle of the compositing period of the other), as detailed in the following figure. This will effectively create an 8-day product that improves temporal change detection.

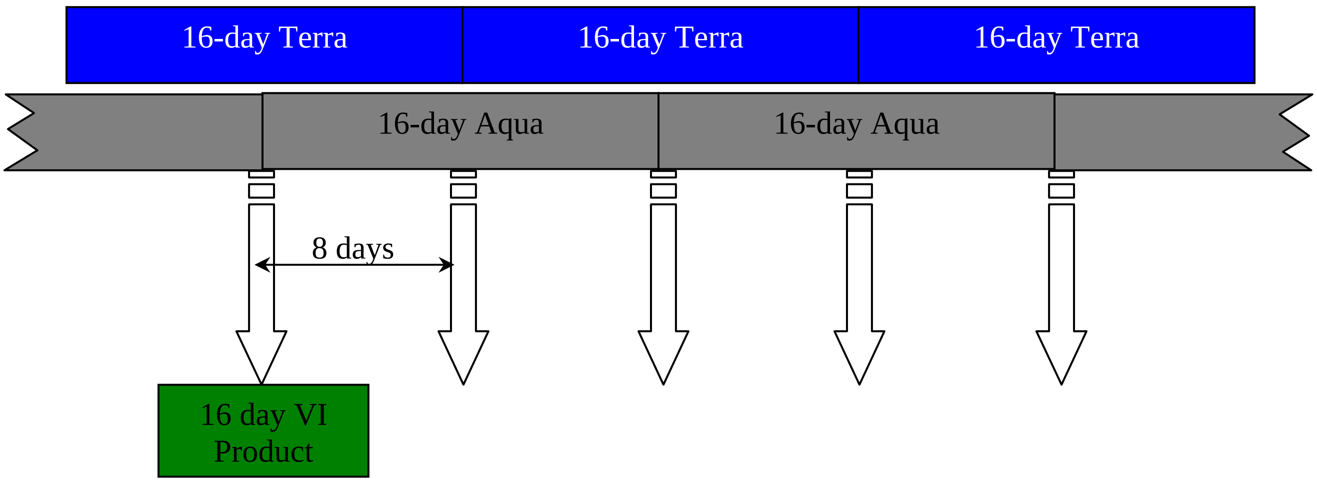


Figure 9: Terra and Aqua 8 day phased production

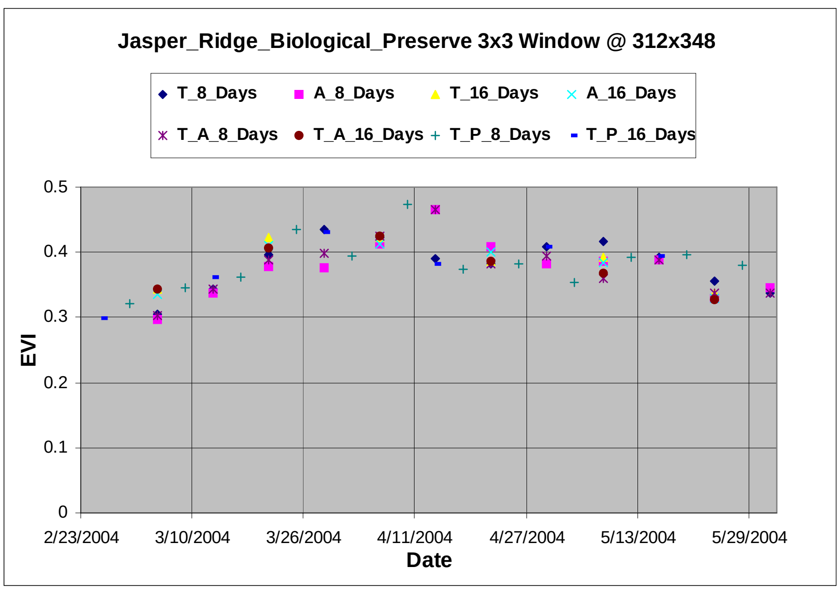


Figure 10: Phased production impact on time series analysis and the resulting improved temporal frequency. Shown are the options evaluated for Collection 5.0 reprocessing (separate 8 day Terra and Aqua streams, Combined 8 day T&A, Current same production day 16-day T&A, 8 day phased T&A, or the adopted 16 day phased T&A).

Miscellaneous changes

Internal compression

Internal compression was implemented in all MODIS Land products starting collection 5.0. The internal compression will permit disk space savings of up to 75% and should be transparent to end users. The product will contain the same SDS list as the regular (non compressed product), in addition another list of Vdata (vertex data) is also stored in the file. This list of Vdata is for internal use by the HDF library, in order to properly decompress the files, and should be ignored by users.

Metadata changes

Due to the change in the structure of the VI quality layer and to reduce the impact on downstream VI production (Monthly 1km and CMG VI product) a new metadata parameters called "**QA_STRUCTURE_STYLE**" was added to the product. These parameters will indicate whether the QA structure is the old style (2 BITS LW mask) or the new style (3 BITS LW mask).
Under automatic post processing, users can consult this parameter in order to properly decompose the quality data bit fields.

Another metadata parameter (**PRODUCTIONTYPE**)was also added, to indicate whether the product was generated under regular production or phased production

- Regular production [1-16, 17-32, ……,353-2/3]
- Phased production [9-24, 25-40,…..,361-365+1-10/11]

Conclusions

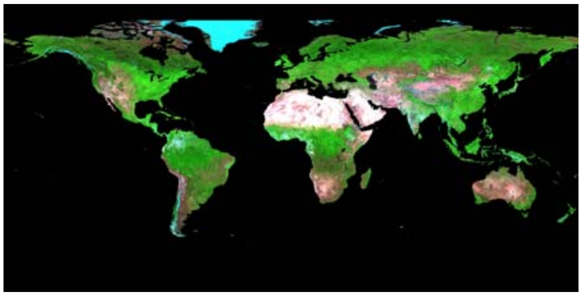
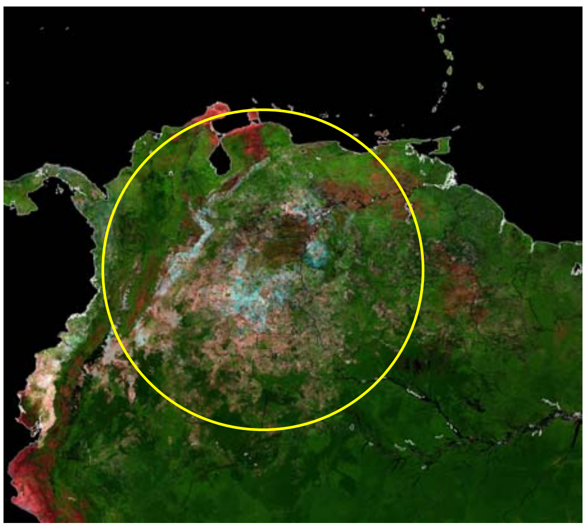
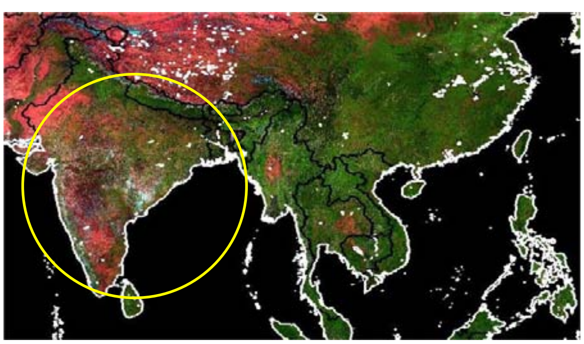
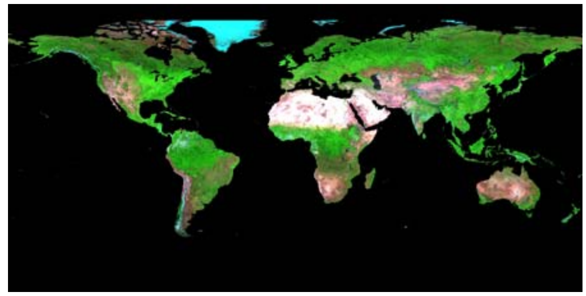
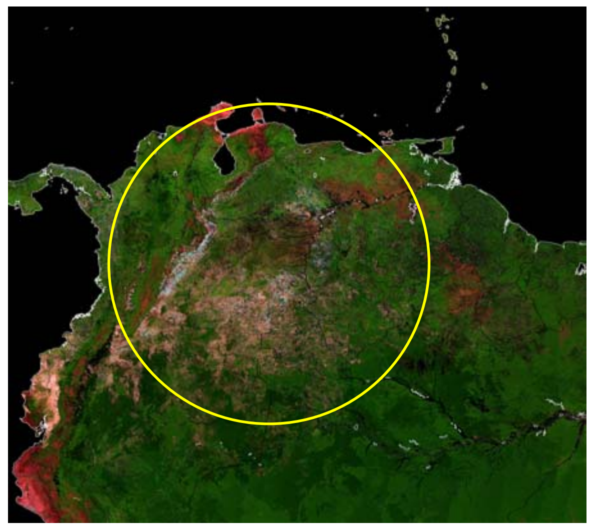
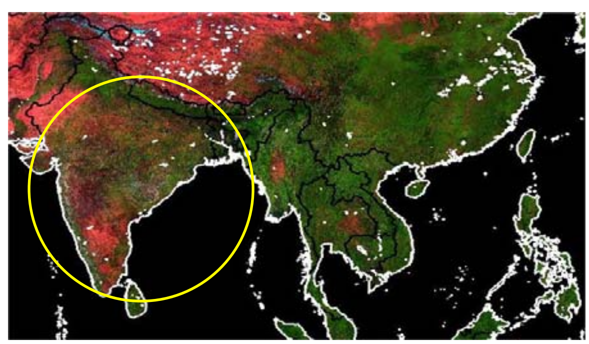
The most important improvements in C5 were related to the new quality based filtering scheme and a modified compositing method to deal with residual and mislabeled clouds. These changes positively impacted all the VI products, since we're now able to identify the least cloudy observation from the daily inputs.

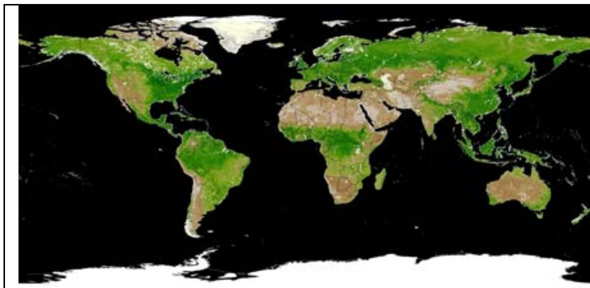
We've also eliminated some of the product redundancies, by adopting a one quality layer, instead of the two separate NDVI and EVI QA layers that were found to be almost identical.

In order to benefit from the presence of two identical data streams (Terra and Aqua) we made modification to the production rules. Terra and Aqua data streams are now processed out of phase (8 days apart) providing a quasi-8-day frequency and improving the change detection capabilities of the products.

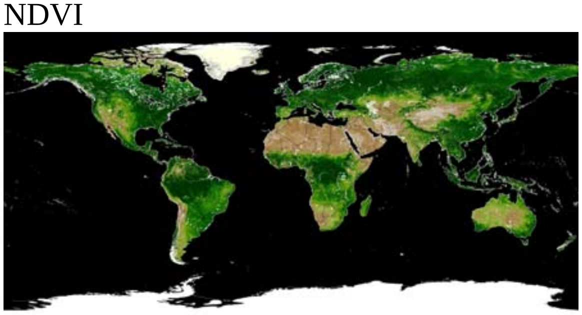
Two new output parameters were also added to the MODIS VI products (Day of composite and Data reliability) in response to users request and learned lesson from previous product version. These two parameters are expected to have a tremendously positive affect on post-processing and addressed end users needs.

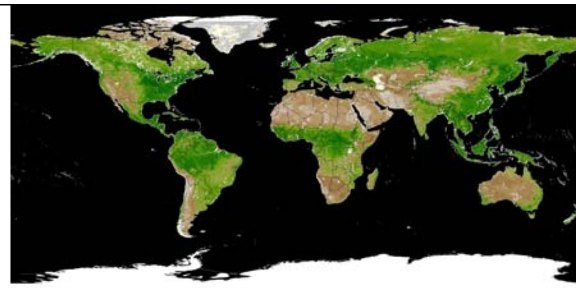
Collection 5 changes example images

C4 (July, 2003)	C5 (July, 2003)
MNR false color	MNR false color
   EVI	   EVI

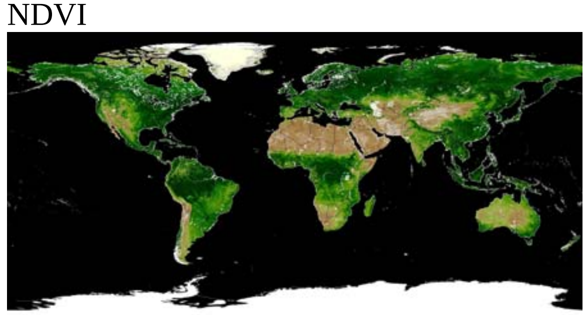


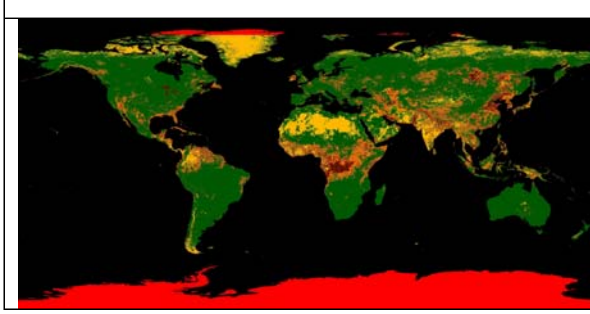
NDVI

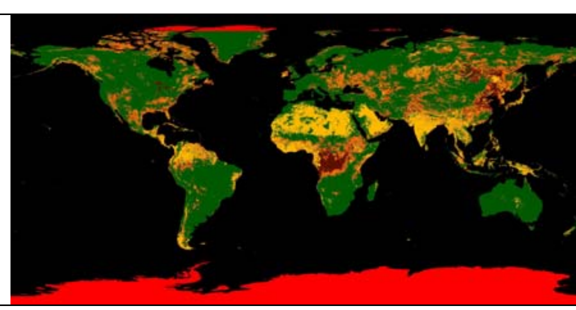


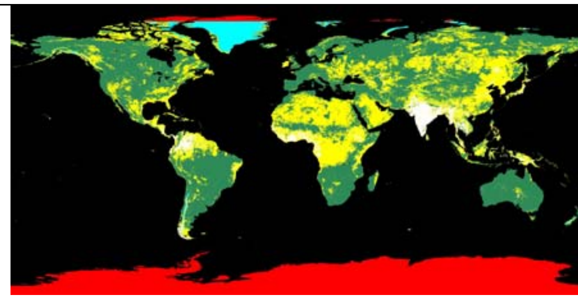
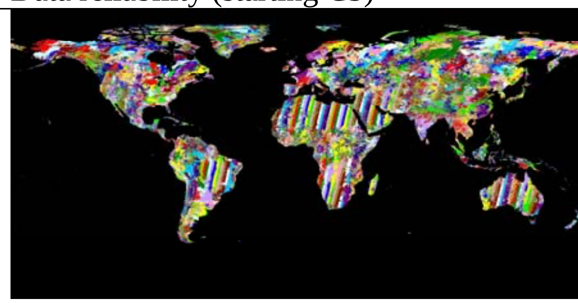


NDVI







Aerosol quality	Aerosol quality
No Data reliability in C4	 Data reliability (starting C5)
No composite DOY in C4	 Composite DOY (starting C5)