A National Crop Monitoring System prototype (NCMS-P) using MODIS data: Near-Real-Time Agricultural Assessment from Space

A. Davidson¹, A. Howard¹,², K. Sun¹, M. Pregitzer¹, P. Rollin¹ and Z. Aly¹

¹ National Land and Water Information Service, AAFC.
² National Agriclimate Information Information Service, AAFC.

Document Revision #: 
Date of Issue: May 29 2009
# TABLE OF CONTENTS

1. **BACKGROUND** ......................................................................................................................... 3

2. **CROP CONDITION ASSESSMENT USING NDVI AT AAFC** .................................................. 6

3. **A NATIONAL CROP MONITORING SYSTEM PROTOTYPE (NCMS-P)** .................................... 7
   3.1. **THE NCMS-P TOOLSET** ..................................................................................................... 7
   3.2. **REQUIRED MODIS DATASETS** .......................................................................................... 8
   3.3. **ORDERING DATA** ............................................................................................................. 14
   3.4. **DOWNLOADING DATA (FTP PULL)** .................................................................................. 15
   3.5. **OPERATIONALLY RECEIVING DATA FROM LP DAAC (FTP PUSH)** .................................. 15
   3.6. **SCIENCE DATASET EXTRACTION** .................................................................................... 16
   3.7. **CHECKING FOR MISSING DATA** ....................................................................................... 17
   3.8. **WEEKLY (N-DAY) MAX-NDVI COMPOSITING TOOL** ..................................................... 18
   3.9. **HISTORICAL BASELINE CALCULATION TOOL** ............................................................. 21
   3.10. **WEEKLY ANOMALY CALCULATION TOOL** ................................................................. 23

4. **PROCESSING PERFORMANCE: A NEAR-REAL-TIME SYSTEM?** ........................................... 25

5. **COMPARISON OF L-2G AND L-1B WEEKLY MAX-NDVI PRODUCTS** .................................... 26

6. **BIBLIOGRAPHY** ......................................................................................................................... 28

APPENDIX A: **MODIS SIN-PROJECTED TILES (GRANULES)** .......................................................... 31
APPENDIX B: **MODIS SEARCH QUERY PARAMETERS** .................................................................... 32
APPENDIX C: **ORDERING DATA THROUGH WIST** ....................................................................... 33
APPENDIX D: **BATCH FILE FOR DOWNLOADING MODIS HDF TILES** ......................................... 37
APPENDIX E: **IMPLEMENTING NDVICOMP, NDVIBASE AND NDVIANOM IN ARCGIS V9.3 AND PYTHON V9.2** ................................................................. 38
APPENDIX F: **MAX-NDVI COMPOSITING TUTORIAL (DVD DEMO)** ............................................. 38
National Crop Monitoring System Prototype (NCMS-P)  
Near-Real-Time Agricultural Assessment from Space

Development Team:  
A. Davidson, A. Howard,  
K. Sun, M. Pregitzer,  
C. Dillabaugh,  
P. Rollin

First Near-Real-Time  
Max-NDVI  
Composite for Canada  
from MODIS-Terra  
6-12 April 2009

First Near-Real-Time  
Max-NDVI Anomaly  
Composite for Canada  
from MODIS-Terra  
6-12 April 2009
1. BACKGROUND

Canada’s agriculturally important lands include lands that are currently used for agricultural purposes (e.g. croplands, rangeland, pastures) as well as those that show potential for future agricultural development (e.g. natural grasslands). However, because these lands generally cover large geographical extents, their assessment by conventional ground survey techniques are often time-consuming and costly (Asrar et al. 1986; Bakhtiari and Zoughi 1991; Tucker 1980; Weiser et al. 1986). Thus, other monitoring approaches must be utilized. One possible approach is the use of satellite remote sensing systems.

Remote sensing, through the unique combination of extensive spatial, spectral and frequent temporal data collection, can provide scientists and managers with a powerful monitoring tool at a variety of landscape scales. The mounting of sensors on earth-orbiting vehicles has increased not only our "spatial" reach (i.e. the distance from which we are able to monitor the earth) but also our "spectral" reach (i.e. our ability to gather information from non-visible wavelengths of the electromagnetic spectrum) (Tucker 1980). Remotely-sensed data collection has the potential to provide quantitative information on the amount, condition, and type of vegetation, provided that the effects of physical and physiological processes on the spectral characteristics of canopies are fully understood.

One of the greatest challenges in the remote sensing of agricultural systems has been the reliable estimation of biophysical variables – such as aboveground biomass, net primary productivity and yield – from satellite platforms. This is largely a consequence of the "mixed pixel" problem, where factors other than the presence and amount of green vegetation (e.g. senescent vegetation, soil, shadow) combine to form composite spectra (see Asner 1998; Asner et al. 1998; Fourtý et al. 1996; Goel 1988; Myeni et al. 1989; Ross 1981). Spectral mixing often makes the discrimination of green vegetation difficult and has prompted the development of numerous spectral vegetation indices (VIs). VIs are dimensionless radiometric measures that combine two or more spectral bands to enhance the vegetative signal, while simultaneously minimizing background effects. Vegetation indices are one of the most widely used remote sensing measurements, and thus, many exist. The most common VIs utilize red and near-infrared canopy reflectances in the form of ratios (e.g. NDVI) or linear combination (e.g. PVI), while others are more complex and also require the derivation of soil correction factors (e.g. MSAVI). Although many indices are well correlated with various plant biophysical parameters, some - such as the NDVI - have received more attention than others.

The Normalized Difference Vegetation Index (NDVI) is a computationally simple index that can be calculated from the red and near infrared data acquired by many satellite systems. The NDVI is calculated as NDVI = (\(\rho_{nir} - \rho_{red}\))/(\(\rho_{nir} + \rho_{red}\)), where \(\rho_{red}\) and \(\rho_{nir}\) are the reflected radiant fluxes in the red and near-infrared wavelengths, respectively (Rouse et al. 1973). The principle behind the NDVI is based on the relationship between the physiological properties of healthy vegetation and the type and amount of radiation it can absorb and reflect (Gitelson and Kaufman 1998). More specifically, plant chlorophyll strongly absorbs solar radiation in the red portion of the electromagnetic spectrum, while plant spongy mesophyll strongly reflects solar radiation in the near-infrared region of the spectrum (Jackson and Ezra 1985; Tucker 1979; Tucker et al. 1991). As a result, vigorously growing healthy vegetation has low red-light reflectance and high near-infrared reflectance, and hence, high NDVI values. The NDVI produces output values in the range of -1.0 to 1.0. Increasing positive NDVI values indicate increasing amounts of green vegetation, while NDVI values near zero and decreasing negative values are characteristic of non-vegetated surfaces such as barren surfaces (rock and soil) and water, snow, ice, and clouds (Jensen 2007). It is important to note, however, that because the NDVI becomes less sensitive to plant chlorophyll at high chlorophyll contents, the NDVI approaches saturation
asymptotically under moderate-to-high biomass conditions (Baret and Guyot 1991; Gitelson and Kaufman 1998; Huete et al. 2002; Myneni et al. 2002; Sellers 1985). As a result, although the NDVI has been shown to correlate well with many canopy biophysical properties – including vegetation abundance (Hurcom and Harrison 1998; Purevdorj et al. 1998), aboveground biomass (Boutton et al. 1980; Davidson and Csillag 2001; Weiser et al. 1986), green leaf area (Asrar et al. 1986; Baret and Guyot 1991; Weiser et al. 1986), photosynthetically active radiation (PAR) (Asrar et al. 1986; Baret and Guyot 1991; Hatfield et al. 1984; Tucker et al. 1986; Weiser et al. 1986), and productivity (Box et al. 1989; Prince 1991; Running et al. 1989) – it generally does so in a non-linear fashion across low-to-high productivity gradients (Figure 1).

The NDVI has emerged as one of the most robust tools for monitoring natural vegetation and crop conditions. This is largely due to its use in the various NDVI datasets produced from daily reflectance observations collected by the Advanced Very High Resolution Radiometer (AVHRR) instruments flown onboard 14 of NOAA’s Polar Orbiting Satellites since 1978. The AVHRR, originally designed for meteorological applications, is a four-channel (AVHRR-1), five-channel (AVHRR-2) or six-channel (AVHRR-3) scanner that senses in the visible, near-infrared, and thermal infrared portions of the electromagnetic spectrum at a spatial resolution of 1.1km (at nadir) (Table 1). The data collected by this series of sensors comprise the longest-lived and most influential series of Earth observing satellites ever launched (Hastings and Emery 1992).

The most commonly-used products derived from the AVHRR are the n-day maximum-value NDVI composites produced by several U.S and Canadian Government agencies (e.g. National Oceanic and Atmospheric Administration (NOAA); National Aeronautics and Space Administration (NASA), Canada Centre for Remote Sensing (CCRS), and Manitoba Remote Sensing Centre (MRSC)) (Cracknell 2001). While the detailed methodologies for creating these datasets vary, maximum-value compositing usually involves (a) examining each NDVI value pixel by pixel for each observation date during the n-day compositing period, (b) determining the maximum-value NDVI for each pixel during the n-day period, and (c) creating a single output image that contains only the maximum NDVI value for each pixel for the n-day period. Maximum-value NDVI compositing has become a popular resource management tool because it captures the dynamics of green-vegetation and minimizes problems common to single-date AVHRR NDVI data, such as those associated with cloud contamination, atmospheric attenuation, surface directional reflectance, and view and illumination geometry (Holben 1986). At present, two Canadian Government Agencies produce maximum-value NDVI composite datasets focusing on a Canadian coverage (National Coverage / 10-day compositing period (CCRS); Prairie Region Coverage / 7-day compositing period (MRSC)).
However, because the AVHRR sensor was not originally designed for monitoring vegetation, it suffers from limitations regarding the design of its red and near infrared channels when formulating NDVI (Fensholt and Sandholt 2005). Two particularly important limitations of the AVHRR are (a) the overlap of the near infrared channel (0.725 – 1.100µm) with a region of considerable atmospheric water vapour absorption (0.9 to 0.98 µm) that can introduce noise to the remotely sensed signal (Huete et al. 2002; Justice et al. 1991); and (b) the relatively “quick” saturation of the red channel, and hence NDVI, over medium-to-dense vegetation (Gitelson and Kaufman 1998; Huete 1988; Jensen 2007; Myneni et al. 1997). These limitations were directly addressed with the development of a new generation of EO platforms, including the Moderate resolution Imaging Spectroradiometer (MODIS) launched onboard NASA’s Terra satellite in December 1999. MODIS, which has been acquiring data in 36 narrow spectral bands since February 2000, was designed to provide data for vegetation and land cover mapping applications. The MODIS sensor offers a number of improvements over the AVHRR for NDVI calculation (Fensholt and Sandholt 2005; Huete et al. 2002; Trishchenko et al. 2002). These include improved (a) spectral resolution; (b) radiometric resolution (c) spatial resolution; (d) geolocation accuracy; and (e) on-board radiometric calibration for producing scaled reflectances (Jensen 2007). The MODIS red and near-infrared channels were selected to avoid the spectral regions of water absorption that constitute a major limitation of the AVHRR (Justice et al. 1991; Vermote and Saleous 2006). Furthermore, the unprecedented radiometric resolution of MODIS Terra makes its red and near-infrared channels more sensitive to small variations in chlorophyll content, thereby lessening how quickly its NDVI saturates over denser vegetation. As a result of these improvements, MODIS Terra holds promise for environmental monitoring in general, and the estimation of vegetation indices in particular (Fensholt and Sandholt 2005).
2. CROP CONDITION ASSESSMENT USING NDVI AT AAFC

Severe droughts, increasing competition among grain exporters, and the instability of grain markets have highlighted the importance of having accurate and timely information on crop conditions and potential yield. There is thus a need to produce operational information and applications to help address both the risk of a weather-related disaster occurring and the potential impact of a current weather-related event.

Under normal operational conditions regular monitoring and forecasting will provide a risk assessment of weather–related impact to the agriculture industry and an estimate of the severity of the impact. Providing advanced warning of such impacts will help in planning anticipated mitigation costs and assist the agricultural industry with preparedness. When low- or moderate-level impacts occur, timely information is available to assist the industry in mitigation efforts.

Agriculture and Agri-Food Canada (AAFC) requires regular information on crop conditions across Canada in near-real-time for crop condition assessment purposes. To this end, AAFC has funded Statistics Canada from 2004-2008 to distribute/deliver AVHRR-derived 1km-resolution NDVI composites of the Canadian agriculture extent and the northern half of the United States. These data, available as 7-day national composites are available through the Crop Condition Assessment Program (CCAP), developed and maintained by Statistics Canada since 1987. CCAP is an interactive Web-based application that monitors changing cropland and pasture conditions during the growing season. These changes in vegetation health are monitored using 1km-resolution digital satellite (AVHRR) data, thematic maps, vegetation index graphs, and tabular data of current and historical cropland and pasture conditions.

However, there are a number of limitations regarding the use of CCAP AVHRR data in AAFC applications. These are (a) the relatively coarse radiometric, spectral and spatial resolutions of the AVHRR’s red and near-infrared channels; (b) the non-availability of daily (or n-day) data products; (c) the restriction of data access to web viewing and web mapping, and (d) the cost of $90K per year for AAFC to fund CCAP. Limitation (c) is especially important because there is a strong need to have access to the digital products through an Application Programming Interface (API) or web services that are not provided through the Statistics Canada application.

Limitations relating to AVHRR data quality (i.e. radiometric, spectral and spatial resolution) can be directly addressed by using the new generation earth observation platforms, including the Moderate Resolution Imaging Spectroradiometer (MODIS) launched onboard NASA’s Terra satellite. MODIS data is available at no cost to the public.

The MODIS Terra sensor offers a number of improvements over the AVHRR for NDVI calculation. These include improved

(a) spectral resolution;
(b) radiometric resolution (12-bit vs 10-bit for the AVHRR sensor);
(c) spatial resolution (250m for NDVI, compared to 1 km for the AVHRR sensor);
(d) geolocation accuracy;
(e) on-board radiometric calibration for producing scaled reflectances.

In response to the above limitations, the NLWIS Climate Data Team has developed a prototype NDVI compositing sub-system (the National Crop Monitoring System Prototype, NCMS-P) that uses data acquired from MODIS to produce weekly / 7-day (or n-day) Max-NDVI composites for Canada south of 60°N. NLWIS plans not only to allow users to access these data through web mapping, but also allow access to the digital products themselves.
3. A NATIONAL CROP MONITORING SYSTEM PROTOTYPE (NCMS-P)

3.1. THE NCMS-P TOOLSET

The NLWIS Climate Data Team has developed a set of individual tools for generating weekly Max-NDVI composites and their associated weekly Max-NDVI anomalies (differences from the 10-year average Max-NDVI conditions for that week). These tools use MODIS Level-2 Gridded (L-2G) surface reflectance data (collection V005).

The toolset is designed to run from batch (.bat) files (although a prototype GUI will be completed by September 2009). Six tools are used to create weekly Max-NDVI composites, weekly average Max-NDVI baselines and weekly Max-NDVI anomalies in near-real-time from these MODIS data. The tools carry out the following tasks:

(a) Download of MODIS HDF data tiles (granules) from USGS Data Archive;
(b) Extraction of required science datasets (as TIF format) from HDF tiles;
(c) Check for missing TIF files and gap filling if tiles are unavailable on USGS archive;
(d) Generation of weekly Max-NDVI and associated Day-of-Week composites;
(e) Generation of weekly historical (10-year) NDVI average baselines (2000-2009); and
(f) Generation of weekly Max-NDVI anomalies.

Successfully running these scripts requires various software programs. The data download script (a) requires installation of the NcFTP FTP client. The science dataset extraction script (b) requires installation of the MODIS Reprojection Tool (MRT). Both the NcFTP client at the MRT are also available on the Climate Data team’s data directories on ||Onotta105a|SWAPS||.

The scripts to check for and replace missing TIF files (c), the Max-NDVI compositing scripts (d-e) and the Max-NDVI anomaly script (f) require access to Python 2.4. These scripts use the ArcGIS Geoprocessor Object, an object that that provides a single access point and environment for the execution of any geoprocessing tool in ArcGIS, including extensions. Instructions for using NCMS-P with ArcGIS V9.3 and Python 2.5 are provided later in this report.

Note that although the primary output products of our near-real-time processing are generated for weekly (7-day) time periods, NCMS-P can be used to generated Max-NDVI-related composites for any n-day period.

Figure 2, illustrates the steps involved in generating weekly Max-NDVI composites in near-real-time. These steps are discussed in further detail in the following sections.
3.2. REQUIRED MODIS DATASETS

NCMS-P uses two MODIS/Terra daily Level-2G MODIS/Terra products (V005):

(a) Surface Reflectance Daily L2G Global 250m SIN Grid V005 [MOD09GQ].
(b) Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V005 [MOD09GA].

Data for each of the above products are stored in NASA’s Land Processes Distributed Active Archive Center (LP DAAC) as tiles (granules) in HDF-EOS format (Hierarchical Data Format; *.hdf). Each tile covers an area of approximately 10º by 10º and is in a sinusoidal (SIN) projection using the WGS84 datum. Twelve tiles are required to completely cover Canada’s landmass south of 60ºN latitude for any single day. A weekly Max-NDVI composite thus uses 7 days *12 tiles = 84 tiles for each of the MOD09GQ and MOD09GA products (168 tiles in total). The SIN-projected tiles used to cover Canada south of 60ºN latitude are shown in Appendix A.

Each tile contains >1 science dataset (SD). A single tile for the MOD09GQ product contains 5 SDs. A single tile for the MOD09GA product contains 21 SDs. NCMS-P uses 7 SDs in total for each day of the compositing period (three SDs from MOD09GQ and four SDs from MOD09QA). A weekly Max-NDVI composite thus uses 7 days *12 tiles * 7 SDs = 588 SDs. SDs are extracted from the HDF-EOS source files using the MODIS Reprojection Tool. The extraction process converts SDs from the HDF-EOS file format to the GeoTIFF file format (note that the SIN projection of the *.hdf tiles are retained).

The SDs contained in the MOD09GQ and MOD09GA products are outlined in Figures 3a and b and 4a, b, and c, below.

<table>
<thead>
<tr>
<th>Science Data Sets (HDF Layers) (5)</th>
<th>UNITS</th>
<th>BIT TYPE</th>
<th>FILL</th>
<th>VALID RANGE</th>
<th>MULTIPLY BY SCALE FACTOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>num_observations: number of observations within the pixel</td>
<td>None</td>
<td>8-bit signed integer</td>
<td>-1</td>
<td>0–127</td>
<td>na</td>
</tr>
<tr>
<td>250m Surface Reflectance Band 1 (020–070 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-26672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m Surface Reflectance Band 2 (041–0876 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-26672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td>250m Reflectance Band Quality</td>
<td>Bit Field</td>
<td>16-bit unsigned integer</td>
<td>2995</td>
<td>0–4095</td>
<td>na</td>
</tr>
<tr>
<td>obs_cov: percentage of the grid cell area is covered by the observation</td>
<td>Percent</td>
<td>8-bit signed integer</td>
<td>-1</td>
<td>0–100</td>
<td>(0.001) 0.00999999976482562</td>
</tr>
</tbody>
</table>

Figure 3a: Science Data Sets for MODIS Terra Surface Reflectance Daily L2G Global 250m SIN Grid V005 (MOD09GQ). Red arrows indicate the science datasets used in the NCMS-P process. Note that the 250m reflectance band quality SD is a bit field, where the values of sequences of bits (rather than the actual pixel value) describe the quality of the product (see Figure 3b).
<table>
<thead>
<tr>
<th>Bit No.</th>
<th>Parameter Name</th>
<th>Bit Comb.</th>
<th>Sur_refl_qc_250m</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1</td>
<td>MCDLAND QA bits</td>
<td>00</td>
<td>corrected product produced at ideal quality all bands</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01</td>
<td>corrected product produced at less than ideal quality some or all bands</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>corrected product not produced due to cloud effects all bands</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>corrected product not produced due to other reasons some or all bands may be fill value [Note that a value of (11) overrides a value of (01)].</td>
</tr>
<tr>
<td>2–3</td>
<td>cloud state</td>
<td>00</td>
<td>clear</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01</td>
<td>cloudy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>mixed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>not set, assumed clear</td>
</tr>
<tr>
<td>4–7</td>
<td>band 1 data quality four bit range</td>
<td>0000</td>
<td>highest quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>dead detector; data interpolated in L1B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1001</td>
<td>solar zenith &gt;= 88 degrees</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1010</td>
<td>solar zenith &gt;= 85 and &lt; 88 degrees</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1011</td>
<td>missing input</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1100</td>
<td>internal constant used in place of climatological data for at least one atmospheric constant</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1101</td>
<td>correction out of bounds, pixel constrained to extreme allowable value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1110</td>
<td>L1B data faulty</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1111</td>
<td>not processed due to deep ocean or clouds</td>
</tr>
<tr>
<td>8–11</td>
<td>band 2 data quality four bit range</td>
<td>SAME AS BAND ABOVE</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>atmospheric correction performed</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no</td>
</tr>
<tr>
<td>13</td>
<td>adjacency correction performed</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>14–15</td>
<td>spare (unused)</td>
<td>-</td>
<td>---</td>
</tr>
</tbody>
</table>

*Figure 3b: MOD09GQ.005 250-meter Surface Reflectance Data QA Descriptions (16-bit). Green arrows indicate the bit combinations used in the NCMS-P quality control process.*
Figure 4a: Science Data sets for MODIS Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V005 (MOD09GA). Red arrows indicate the science datasets used in the NCMS-P process. Note that the 1km reflectance state QA SD is a bit field, where the values of sequences of bits (rather than the actual pixel value) describe the quality of the product.

Table continued on next page…
<table>
<thead>
<tr>
<th>DATA GROUP</th>
<th>Science Data Sets (HDF Layers) (12)</th>
<th>UNITS</th>
<th>BIT TYPE</th>
<th>FILL</th>
<th>VALID RANGE</th>
<th>MULTIPLY BY SCALE FACTOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>500m (2)</td>
<td>run_observations_500m</td>
<td>none</td>
<td>8-bit signed integer</td>
<td>-1</td>
<td>0–127</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>sur_refl_b01: 500m Surface Reflectance Band 1 (620-670 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-28672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>sur_refl_b02: 500m Surface Reflectance Band 2 (841-876 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-28672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>sur_refl_b03: 500m Surface Reflectance Band 3 (459-479 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-28672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>sur_refl_b04: 500m Surface Reflectance Band 4 (545-565 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-28672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>sur_refl_b05: 500m Surface Reflectance Band 5 (1230-1250 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-28672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>sur_refl_b06: 500m Surface Reflectance Band 6 (1628-1652 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-28672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>sur_refl_b07: 500m Surface Reflectance Band 7 (2105-2155 nm)</td>
<td>Reflectance</td>
<td>16-bit signed integer</td>
<td>-28672</td>
<td>-100–16000</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>GC_500m: 500m Reflectance Band Quality</td>
<td>Bit Field</td>
<td>32-bit unsigned integer</td>
<td>787410671</td>
<td>0–4294967295</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Obs_cov_500m: Observation coverage</td>
<td>Percent</td>
<td>8-bit signed integer</td>
<td>-1</td>
<td>0–100</td>
<td>(0.01) 0.00999999977648258</td>
</tr>
<tr>
<td></td>
<td>lols_res: Observation number</td>
<td>none</td>
<td>8-bit unsigned integer</td>
<td>255</td>
<td>0–254</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>q_scan: 250m scan value information</td>
<td>Bit Field</td>
<td>8-bit unsigned integer</td>
<td>255</td>
<td>0–255</td>
<td>na</td>
</tr>
</tbody>
</table>

*Figure 4a (Continued): Science Data sets for MODIS Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V005 (MOD09GA). Red arrows indicate the science datasets used in the NCMS–P process.*
<table>
<thead>
<tr>
<th>Bit No.</th>
<th>Parameter Name</th>
<th>Bit Comb.</th>
<th>state_1km</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1</td>
<td>cloud state</td>
<td>00</td>
<td>clear</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01</td>
<td>cloudy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>mixed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>not set, assumed clear</td>
</tr>
<tr>
<td>2</td>
<td>cloud shadow</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no</td>
</tr>
<tr>
<td>3-5</td>
<td>land/water flag</td>
<td>0000</td>
<td>shallow ocean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>001</td>
<td>land</td>
</tr>
<tr>
<td></td>
<td></td>
<td>010</td>
<td>ocean coastlines and lake shorelines</td>
</tr>
<tr>
<td></td>
<td></td>
<td>011</td>
<td>shallow inland water</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>ephemeral water</td>
</tr>
<tr>
<td></td>
<td></td>
<td>101</td>
<td>deep inland water</td>
</tr>
<tr>
<td></td>
<td></td>
<td>110</td>
<td>continental/moderate ocean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>111</td>
<td>deep ocean</td>
</tr>
<tr>
<td>6-7</td>
<td>aerosol quantity</td>
<td>00</td>
<td>climatology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01</td>
<td>low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>average</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>high</td>
</tr>
<tr>
<td>8-9</td>
<td>cirrus detected</td>
<td>00</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01</td>
<td>small</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>average</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>high</td>
</tr>
<tr>
<td>10</td>
<td>internal cloud algorithm flag</td>
<td>1</td>
<td>cloud</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no cloud</td>
</tr>
<tr>
<td>11</td>
<td>internal fire algorithm flag</td>
<td>1</td>
<td>fire</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no fire</td>
</tr>
<tr>
<td>12</td>
<td>MOD35 snow/ice flag</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no</td>
</tr>
<tr>
<td>13</td>
<td>Pixel is adjacent to cloud</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no</td>
</tr>
<tr>
<td>14</td>
<td>BRDF correction performed</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no</td>
</tr>
<tr>
<td>15</td>
<td>internal snow mask</td>
<td>1</td>
<td>snow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>no snow</td>
</tr>
</tbody>
</table>

Figure 4b: MOD09GA 1-kilometer State QA Descriptions (16-bit). Green arrows indicate the bit combinations used in the NCMS-P quality control process.
The necessary SDs are extracted from the downloaded *.hdf tiles using the MODIS Reprojection Tool software (see following sections).

Once SD extraction is complete, the Max-NDVI compositing algorithm is implemented (see following sections). This algorithm uses Band 1 and Band 2 surface reflectance at 250m resolution, along with the QA flags, to produce an n-day Max-NDVI composite from only the “best quality” reflectance retrievals. The general concept is to retain pixels in the Max-NDVI procedure that are (i) not influence by atmospheric effects, and (ii) fall within the range of acceptable sensor and sun illumination angles.

Specifically, screening involves the elimination of pixels that are contaminated by clouds, cloud shadow, high amounts of aerosols, and high amounts of cirrus. In addition, pixels collected at unacceptable viewing and sun illumination angles are also eliminated. The screening rules for the selection of acceptable retrievals are summarized in Table 3, Section 3.8.

While output products are focused on weekly (7-day) periods, it is important to note that n-day “custom” products can also be generated.
3.3. ORDERING DATA [VIDEO]

MODIS L-2G reflectance data is ordered from NASA's Land Processes Distributed Active Archive Center (LP DAAC) using the Warehouse Inventory Search Tool (WIST).

When generating weekly Max-NDVI composites for Canada, data is acquired over a 23 week growing season from the 2nd week of April (Week 15) until the 2nd week of September (Week 37). Weeks of the year are defined according to the ISO 8601 week-numbering standard. For each week, data is ordered from Monday at 00:00:00hrs (day 1 of week) to the following Sunday at 23:59:59hrs (day 7 of week).

To access the data archives, one must first create a user account (user name and password). On the first login, a user must set up the search criteria (WIST remembers previous searches, so the search criteria needs to be set up only once; in subsequent searches the user need only change the start and end dates in the query).

An example of a search query is shown in Figure 5 and in Appendix B. Detailed instructions on query building in WIST are provided in Appendices B and C.

![Example search query](image.png)

Figure 5: Example search criteria for 27 April – 3 May 2009. This queries two datasets (MOD09GQ and MOD09QA for 12 tiles (ranging from h09 to h14 to v03 - v04 (i.e. h09v03, h09v04, h10v03 etc)) (Also see Appendix B).
3.4. **DOWNLOADING DATA (FTP PULL)** [VIDEO]

Once the query has been submitted to WIST, a series of emails (four in total) are sent to the user's account to inform the user of the status of their data search (see Appendix C for examples of these emails). The first three emails confirm the data being ordered, its format and its cost (free for MODIS data). The arrival of the fourth email indicates that tiles are available for download. This email contains an ftp location of the data and download instructions.

The time taken for data requests to be processed by WIST varies, and depends on (a) the number of requests the DAAC is processing at that given time, and (b) the volume of data in an order. For example, small data orders (e.g. a few tiles) are often available for download within an hour of the order being placed. In comparison, large orders (e.g. one week of data comprising a total of 168 tiles) are usually available 24 hours after the order has been placed. In addition, queries submitted at “low traffic” times (late evening, early morning) are generally processed more quickly than queries submitted during peak times (9 to 5pm).

The requested MODIS tiles are downloaded using a batch file that downloads tiles from the ftp site location (note: To run this file, drop the .txt extension from the filename and double click the file in windows explorer). Batch files require installation of the NcFTP FTP client.

Batch files take the form:

```bash
@echo off
:: FTPdatadownload.bat
::
:: This batch file automatically downloads MODIS data from their ftp site
:: using NcFTP FTP client software. The -z flag makes sure that files
:: already downloaded are not overwritten if the download is restarted for
:: any reason.
::
:: **Note**: If a download fails and needs to be restarted, the file
:: being downloaded at time of failure will only be partially downloaded.
:: This file should be deleted before restarting the download.
::
mkdir 2009117to123
cd 2009117to123
ncftpget ftp://e4ftl01u.ecs.nasa.gov/PullDir/0301617268QdcAPB/* -z
```

where the ftp location corresponds to the Ftp Pull Download Links specified in the email described above. It is important to note that (a) the ftp location will change from order to order; (b) the ftp location is usually only valid for 5 days from when the order is made available; and (c) directory locations specified in the batch file can be relative or absolute (also see Appendix D).

Two download options are provided by the LPDAAC (many tiles vs a single ZIP packaged order). We favour the former because interrupted downloads can be resumed more easily (the \-z flag ensure downloaded files are not overwritten when a download is restarted. However, as noted, partial file downloads must be deleted before the batch file is rerun to avoid problems later in the processing chain). Both methods take the same amount of time to download files.

3.5. **OPERATIONALLY RECEIVING DATA FROM LP DAAC (FTP PUSH)**

Data delivery as FTP push from the LP DAAC would automate the above process. However, additional AAFC FTP support (not currently available) is required for this to occur.
3.6. SCIENCE DATASET EXTRACTION [VIDEO]

The MODIS Science Datasets (SDS) described earlier are extracted on a tile-by-tile basis using the MODIS Reprojection Tool (MRT). The MRT executable program is run from a batch file to extract the required SDSs and convert them to a GeoTIFF file with a standardized name (ProductCode_Date_TileID_SDS). The tiles are kept in their original sinusoidal projection.

The MRT is software designed to help individuals work with MODIS data by reprojecting MODIS land data image products into more standard map projections. The user may reproject selected portions of the image (spatial subsetting) and selected image bands (spectral subsetting).

The MRT outputs MODIS data in file formats that are supported by existing software packages (raw binary and GeoTIFF) as well as HDF-EOS. The MRT runs on several platforms, including Sun Solaris workstations, SGI IRIX workstations, Linux, 64-bit Linux, Microsoft Windows, and Mac OS X.

The heart of the MRT is the resample and mrtmosaic executable programs that may be run either from the command-line or from the MRT Graphical User Interface (GUI). The GUI is an easy, user-friendly way to run the MRT. However, individuals with intensive data processing requirements are advised to investigate the more powerful command-line interface, which allows background processing of large MODIS data files when the system load is light.

We use the resample executable in a batch file to extract the seven science datasets described previously (four from MOD09GA and three from MOD09GQ) that are used in the NDVI compositing procedure (note: To run this file, drop the .txt extension from the filename and double click the file in windows explorer).

The batch script (a) extracts the SDSs required from the MOD09GA and MOD09GQ hdf files to GeoTiff files, (b) renames the SDSs to a standardized name, and (c) moves the GeoTiffs to an output directory.

An example batch file is shown in Figure 6.

![Figure 6: Example batch file for extracting science datasets from downloaded *.hdf tiles.](image)
3.7. CHECKING FOR MISSING DATA [VIDEO]

In some cases, not all *.hdf tiles are available for a given week. This is problematic because the current version of NCMS-P does not handle missing data files. To address this problem, data gaps are “filled”. Filling gaps involves (a) identifying the GeoTIFF files expected for an n-day, n-tile composite (usually, 7-day, 12-tiled), (b) identifying if any of the expected tiles are missing, and (c) using the available tiles to fill gaps in the data.

The python script Check4MissingDataTIF.py is used to carry out the gap-filling. The script is run from a batch file that takes five input arguments:

```
Check4MissingDataTIF.py 2009117 7 C:/MOD/117to123/ C:/MOD/tile_list.txt 0
```

These input arguments are (i) the start date of the compositing period, (ii) the number of days in the compositing period, (iii) the input directory containing the TIF files, (iv) a file containing a list of tiles to be used, and (v) a flag that specifies whether the script should be run in report mode (0) or report/replace mode (1). Note that directory and file names should be specified as absolute paths.

When run in “report mode”, the script identifies the GeoTIFF files expected for the n-day period for n-tiles, and identifies if any tiles are missing. The script generates a report file that contains the results of the search. No gap-filling is done in this mode.

When run in “report and replace mode”, the script identifies the GeoTIFF files expected for the n-day period for n-tiles, and identifies if any tiles are missing. The script then fills in the gaps in the missing data. It does this for each missing TIF file by identifying the corresponding tile for the previous day, then copies and renames the file to the name of the missing file, thus “filling the data gap”. If the TIF for the previous day is not available (i.e. it is also missing), the TIF for the following day is instead used to fill the gap. The report file generated in this mode logs the missing tile names, and which tiles they were “filled” with.

```
Expected 588 Files

Missing files:
MOD09GQ.A2009123.h14v03.hdf.sur_refl_b01_1.tif

587 files counted | 1 file missing
Gap Filling Files:
Filling: MOD09GQ.A2009123.h14v03.hdf.sur_refl_b01_1.tif | With:
MOD09GQ.A2009122.h14v03.hdf.sur_refl_b01_1.tif
```

Note that the pixel values of Max-NDVI composites generated by the compositing tool are not affected by this process.

Fortunately, missing data is not a common problem. For the 2000-2008 growing seasons, gap-filling needed to be used the following number of times: 0 (2003, 2004, 2006), 1 (2005, 2007), 2 (2001, 2002), 3 (2000), 4 (2008). [Note: This does not include the 3 weeks in 2000 and 2 weeks in 2001 where no data was available and composites were not created].

The Data Archive shows the weeks that contain gap-filled data, and the tiles and dates that were unavailable for download. Note: the current version of this document is located at.

\http://Onotta105a\SWAPS\AEData-Climate\Climate\MODIS\b.MODIS.7DAY.MAX.NDVI\Data\Data Archive.xls
3.8. WEEKLY (N-DAY) MAX-NDVI COMPOSITING TOOL [VIDEO]

The weekly (n-day) Max-NDVI compositing tool is written in python 2.4 and runs using ArcGIS 9.2 with the spatial analysis extension. Instructions on how to run this tool in ArcGIS 9.3 and Python 2.5 are provided in Appendix E. The use of this tool is demonstrated using data from week 18 2009 (Day-of-year 2009117 to 2009123).

3.8.1 Implementing the Max-NDVI compositing tool [NDVICOMP.py]

Max-NDVIcomposing is carried out using the python script NDVICOMP.py (that calls another python script mapping.py). NDVICOMP.py is run from a batch file that takes five input arguments.

```
NDVICOMP.py F:/hdf/117to123/ 2009117 7 F:/2009/2009117 F:/ tile_list.txt
```

These input arguments are (i) the directory containing the GeoTIFF files to be used as input to the compositing procedure, (ii) the start date of the compositing period, (iii) the number of days in the compositing period, (iv) the output directory in which composites are placed, and (v) a file containing the list of tiles to be used as input to the compositing algorithm (e.g. tile_list.txt). Note that directory and file names should be specified as absolute paths.

3.8.2 Screening for the best quality surface reflectance retrievals

The NDVI compositing algorithm uses Band 1 and Band 2 surface reflectance at 250m resolution along with the QA flags to decide which pixels to retain for the compositing. The general concept is to choose the pixels that are the most cloud free that fall within the range of acceptable sensor and sun illumination angles. Quality flags in the L2G data are used to eliminate pixels that are contaminated by clouds, high aerosols, cloud shadow, and high cirrus, and pixels that were collected at unacceptable viewing and sun illumination angles. The rules for the compositing algorithm are summarized in Table 3, below.

<table>
<thead>
<tr>
<th>Data Product</th>
<th>Science Dataset</th>
<th>Bits (if bit field)</th>
<th>Quality Criteria</th>
<th>Bitmap Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOD09GQ</td>
<td>250m Reflectance Band Quality</td>
<td>0-1</td>
<td>MODLAND QA = Corrected product at ideal quality all bands</td>
<td>== 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-3</td>
<td>Cloud state = clear</td>
<td>== 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4-7</td>
<td>Band 1 data quality = highest quality</td>
<td>== 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8-11</td>
<td>Band 2 data quality = highest quality</td>
<td>= 0</td>
</tr>
<tr>
<td></td>
<td>State_1km: Reflectance Data State QA</td>
<td>2</td>
<td>Cloud shadow = no</td>
<td>= 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6-7</td>
<td>Aerosol quantity = not high</td>
<td>&lt;&gt; 3</td>
</tr>
<tr>
<td>MOD09GA</td>
<td>Sensor Zenith</td>
<td>8-9</td>
<td>Cirrus quantity = not high</td>
<td>&lt;&gt; 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sensor zenith between 0 and 45 degrees</td>
<td>&gt;= 0 and &lt;= 4500</td>
</tr>
</tbody>
</table>

Table 3: Pixel screening criteria used to generate MODIS Max-NDVI composites of high quality.

3.8.3 Output Products: Max-NDVI and Day-of-Week Composites
The compositing tool generates four outputs. These are: (i) a Max-NDVI composite for the 7-day 12-tile compositing period, where NDVI ranges from -1 to 1 (nmcom*); (ii) a rescaled version of this NDVI composite, where negative NDVI values (NDVI < 0) are set to 0 (rncom*); (iii) a day-of-week (DOW) composite where pixel values correspond to the day of the compositing period from which their Max-NDVI value is taken (dow*); and (iv) a logfile that contains the information on the successful or unsuccessful implementation of the script.

Figures 7 and 8 illustrate output rasters generated by the compositing tool. The example shown here uses 12 tiles of MODIS data for a 7-day period (DOY 117-123, 2009). NDVI composites are generated as 32-bit float Geotiff rasters. DOW composites are generated as 8-bit Geotiff rasters. All GeoTiff rasters remain in their native sinusoidal projection (Fig. 8).

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>dow2009117x123.e0x</td>
<td>11.666</td>
<td>2009-05-15 1:33:44 AM</td>
</tr>
<tr>
<td>dow2009117x123.e1d</td>
<td>10.460,388</td>
<td>2009-05-15 1:33:44 AM</td>
</tr>
<tr>
<td>dow2009117x123.e2d</td>
<td>277,561,736</td>
<td>2009-05-15 1:33:18 AM</td>
</tr>
<tr>
<td>dow2009117x123.e3d</td>
<td>231</td>
<td>2009-05-15 1:33:32 AM</td>
</tr>
<tr>
<td>dow2009117x123.e4d</td>
<td>1,624</td>
<td>2009-05-15 1:33:34 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e0x</td>
<td>534,761</td>
<td>2009-05-15 1:26:23 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e1d</td>
<td>537,074,170</td>
<td>2009-05-15 1:26:23 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e2d</td>
<td>1,157,201,736</td>
<td>2009-05-15 1:25:20 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e3d</td>
<td>1,592</td>
<td>2009-05-15 1:25:24 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e4d</td>
<td>534,590</td>
<td>2009-05-15 1:25:24 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e5d</td>
<td>86,908,152</td>
<td>2009-05-15 1:30:35 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e6d</td>
<td>1,127,541,862</td>
<td>2009-05-15 1:30:35 AM</td>
</tr>
<tr>
<td>rnmcom2009117x123.e7d</td>
<td>553</td>
<td>2009-05-15 1:30:37 AM</td>
</tr>
</tbody>
</table>

Figure 7: Directory listing showing GeoTiff rasters generated by the n-day Max-NDVI compositing tool (example shown is a 7-day period from day-of-year 117 to 123, 2009, using 12 tiles).

Figure 8: (a) Re-scaled Maximum NDVI composite (where NDVI ranges from 0-1), and (b) Day-of-week (DOW) composite, where pixel values are assigned the value corresponding to the day of the compositing period from which their Max-NDVI value is taken. Note that composites remain in their native sinusoidal projection.
Figure 9: When NDVICOMP.py is implemented, messages are echoed to the screen and to a logfile. These messages reflect successful or unsuccessful implementation of the script. If script failure occurs, the logfile should be assessed to identify where problems in implementation occurred.

3.8.4 Using Max-NDVI composites as input to other NCMS tools

The generation of weekly Max-NDVI and DOW is not the end of the processing chain. We are often interested in how the current week’s NDVI compares to the “historical” average (baseline) Max-NDVI conditions for that week. To do this, we must calculate the mean NDVI conditions for the week in question across the entire MODIS historical record, then calculate the difference (anomaly) between the current week’s Max-NDVI (current conditions) and the baseline. These processes are described in detail in the following sections.
3.9. **HISTORICAL BASELINE CALCULATION TOOL [VIDEO]**

The baseline (mean) calculation tool is written in python 2.4 and runs using ArcGIS 9.2 with the spatial analysis extension. No modification of this tool should be needed to run with python 2.5 and ArcGIS 9.3. The use of this tool is demonstrated using data from week 18 2000-2009.

3.9.1 **Implementing the baseline calculation tool [NDVIBASE.py]**

The calculation of weekly baselines (averages) is carried out using the python script `NDVIBASE.py`, which is run from a **batch file** that takes five input arguments.

```
```

These input arguments are (i) the week number to be processed; (ii) the start year in the baseline calculation period; (iii) the end year in the baseline calculation period; (iv) the directory containing the Max-NDVI GeoTiff files to be used as input to the baseline calculations; and (v) the directory to which generated baselines are written. Note that directory and file names should be specified as absolute paths.

We use the rescaled Max-NDVI composites (rncom*; NDVI ranges from 0-1) in the baseline calculations. Note that all GeoTiffs in the input directory are used in the calculations.

3.9.2 **Output Products: Baseline and Standard Deviation composites**

The baseline calculation tool generates three outputs. These are: (i) a baseline Max-NDVI composite for each week in the growing season (used to describe the long term “baseline” Max-NDVI conditions for a given week) (Av.MaxNDVI.*); (ii) a standard-deviation Max-NDVI composite for each week in the growing season (used to describe the historical variability in Max-NDVI values for a given week) (St.MaxNDVI.*); and (iii) a **logfile** that contains information on the files used as input to the calculations (Logfile.AvStdv.MaxNDVI.*).

Figures 10 and 11 illustrate output rasters generated by the baseline calculation tool. The example shown here uses 12 tiles of MODIS data for week 18 2000-2009. Baseline and standard deviation composites are generated as 32-bit float Geotiff rasters. All GeoTiff rasters remain in their native sinusoidal projection (Fig. 11).

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Modified</th>
</tr>
</thead>
</table>

**Figure 10:** Directory listing showing GeoTiff rasters generated by the baseline calculation tool: Averages (Av.MaxNDVI.*) and standard deviation (St.MaxNDVI.*) (example shown for week 18, 2000-2009). Note also the logfile that contains information on the files used as input to the calculations.
Figure 11: (a) The baseline (mean) of Max-NDVI for week 18 (2000-2009) and (b) the standard deviation of Max-NDVI for week 18 (2000-2009). We use the rescaled Max-NDVI composites (rncom*; NDVI ranges from 0-1) in the baseline and standard deviation calculations. Note that composites remain in their native sinusoidal projection.

Figure 12: When NDVIBASE.py is implemented, messages are echoed to the screen and to a logfile. These messages reflect successful or unsuccessful implementation of the script. In the above, example, baselines are calculated using the rescaled Max-NDVI for week 18 (2000-2009).
At the time of writing (June 2009), mean baseline NDVI conditions exist for all growing season weeks (weeks 15-37) using data from the period 2000-2008 (n=9), and for weeks 15-21 using data from the period 2000-2009 (n=10). The availability of these data and their storage locations are described in the Data Archive.

3.10. WEEKLY ANOMALY CALCULATION TOOL [VIDEO]

The weekly anomaly calculation tool is written in python 2.4 and runs using ArcGIS 9.2 with the spatial analysis extension. No modification of this tool should be needed to run with python 2.5 and ArcGIS 9.3. The use of this tool is demonstrated using data from week 18 2009 (Day-of-year 2009117 to 2009123) and the long-term (10-year) average Max-NDVI for week 18 generated by the baseline calculation tool.

3.10.1 Implementing the weekly anomaly calculation algorithm [NDVIANOM.py]

The calculation of weekly baselines (means) is carried out using the python script NDVIANOM.py, which is run from a batch file that takes three input arguments.

```
NDVIANOM.py C:/W18/rncom2009117to123.tif
```

These input arguments are (i) the weekly Max-NDVI composite that the anomaly will be generated for; (ii) the average Max-NDVI composite (baseline) for the week in question; and (iii) the name of the output anomaly composite. Note that directory and file names should be specified as absolute paths. We use the weekly rescaled Max-NDVI composites (rncom*; NDVI ranges from 0-1) in the anomaly calculations.

3.10.2 Output Product: Anomaly composites

The compositing tool generates one output: the anomaly composite for the weekly Max-NDVI composite specified.

Figures 13 and 14 illustrate output rasters generated by the anomaly calculation tool. The example shown uses the Max-NDVI composite for week 18 2009 (DOY 117-123, 2009) and the baseline created for week 18 using the ten weekly composites for week 18 from 2000-2009. Anomaly composites are generated as 32-bit float Geotiff rasters. All Geotiff rasters remain in their native sinusoidal projection (Fig. 14).

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anom2009117to123.Week.18.aux</td>
<td>534,966</td>
<td>2009-05-15 8:14:00 AM</td>
</tr>
<tr>
<td>Anom2009117to123.Week.18.ind</td>
<td>90,754,074</td>
<td>2009-05-15 8:13:59 AM</td>
</tr>
<tr>
<td>Anom2009117to123.Week.18.tif</td>
<td>1,107,541,652</td>
<td>2009-05-15 8:13:59 AM</td>
</tr>
<tr>
<td>Anom2009117to123.Week.18.tif.xml</td>
<td>943</td>
<td>2009-05-15 8:14:00 AM</td>
</tr>
</tbody>
</table>

Figure 13: Directory listing showing GeoTiff rasters generated by the calculation tool (example shown for difference between week 18 2009 and the baseline calculated by week 18 Max-NDVI composites for the years 2000-2009).
Figure 14: The NDVI anomaly for week 18, 2009. This anomaly is calculated as the difference in NDVI between the Max-NDVI composite for this week and the ten-year average (baseline) NDVI for this week. Note that composites remain in their native sinusoidal projection.

Figure 15: When NDVIANOM.py is implemented, messages are echoed to the screen. These messages reflect successful or unsuccessful implementation of the script.

At the time of writing, weekly anomalies are available for weeks 15-21 for 2009. The availability of these data and their storage locations are described in the Data Archive.
4. PROCESSING PERFORMANCE: A NEAR-REAL-TIME SYSTEM?

Daily MODIS L-2G reflectance and QA/QC tiles are usually added to the LPDAAC archive 3 days after data acquisition. This means that the 184 tiles required to produce a weekly (Monday–Sunday) Max-NDVI composite are usually all available for download some time during the Wednesday of the following week (= +2.5 days after reflectance retrieval).

However, sometimes this is not the case. There are weeks when data are available for download more quickly (e.g. Wednesday morning). There are also weeks when data availability on the LPDAAC is delayed by a few days and/or when data requests through WIST have taken much longer to be processed than usual. The LPDAAC Support Team can be emailed for advice if you do not receive an order notification within 1-3 days (quote your order number).

To produce weekly Max-NDVI composites as close to near-real-time as possible, data tiles should be ordered, downloaded and extracted in two- or three-day blocks during the week of interest, leaving as few tiles as possible to be ordered, downloaded and extracted the following Wednesday. By doing this, Max-NDVI compositing can usually commence 3-4 hours after placing the final data order (+ 1hr to receive final order confirmation and data ftp location; +1hr to download *.hdf tiles; +1 hr to extract SDs from *.hdf tiles).

The Max-NDVI compositing process takes approximately 8-10 hours to complete. Thus, weekly Max-NDVI and DOW composites are usually available <12 hours of placing the final data order. Note that this is not the end of the processing chain: The weekly Max-NDVI composite is then used to update the historical baseline (average) NDVI for the week (+6 hours), from which the week’s NDVI anomalies are calculated (+ 0.5 hours). Thus, all weekly products – a Max-NDVI composite, updated historical weekly baseline and the Max-NDVI anomaly – can usually be generated within 10-22 hour after placing the final data order.

Table 4 shows an “ideal” processing schedule for week 18. Here, the NDVI products are generated by early Thursday afternoon (+3.5 days from the end of the previous week).

<table>
<thead>
<tr>
<th>Week</th>
<th>Day of Week (DOW)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>1 Monday</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 Tuesday</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 Wednesday</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 Thursday</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 Friday</td>
<td>- Order data for Week 18 DOW 1-2, Download, Extract SDs</td>
</tr>
<tr>
<td></td>
<td>6 Saturday</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7 Sunday</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1 Monday</td>
<td>- Order data for Week 18 DOW 3-5, Download, Extract SDs</td>
</tr>
<tr>
<td></td>
<td>2 Tuesday</td>
<td>- Order data for Week 18 DOW 6, Download, Extract SDs</td>
</tr>
<tr>
<td></td>
<td>3 Wednesday</td>
<td>- Order data for Week 18 DOW 7, Download, Extract SDs</td>
</tr>
<tr>
<td></td>
<td>4 Thursday</td>
<td>- Max-NDVI and DOW compositing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Output Max-NDVI and DOW products available.</td>
</tr>
<tr>
<td></td>
<td>5 Friday</td>
<td>- Order data for Week 19 DOW 1-2, Download, Extract SDs... etc</td>
</tr>
<tr>
<td></td>
<td>6 Saturday</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: An “ideal” processing schedule for generating weekly Max-NDVI and DOW composites, baselines and weekly anomalies. Output products can usually be generated by Thursday evening (< +3.5 days from end of week) assuming all MODIS data products are available for download on the Wednesday. For the 2009 growing season, we have generated weekly composites as quickly as +3 days from the end of week (Wednesday) and as late as > +6 days from the end of week (Saturday).
5. COMPARISON OF L-2G AND L-1B WEEKLY MAX-NDVI PRODUCTS

As part of the quality control process, we compared weekly Max-NDVI composites generated by NCMS-P to weekly Max-NDVI composites generated by the CCRS processing system that uses MODIS L-1B (Swath) data.

We found high correlations between products for average within-unit Max-NDVI and total within-unit Max-NDVI at various geographic scales (Census Agriculture Regions; Census Subdivisions; Dissemination Areas; Townships) using four weekly composites for growing season weeks in 2006. These results suggest that during the growing season, weekly L-2G Max-NDVI composites from NCMS-P provide similar outputs to those of the CCRS L-1B system.

These results are illustrated in Figures 16(a)-(d). In all cases, high correlations were found between the Max-NDVI generated by the L-2G MODIS data (NCMS-P) and the L-1B MODIS Data (CCRS). Correlations between the data were $r > 0.98$ (Census Agriculture Regions for Alberta, Saskatchewan and Manitoba), $r > 0.92$ (Census Subdivisions and Dissemination Areas, for Alberta, Saskatchewan and Manitoba) and $r > 0.97$ (Townships within Canada’s Prairie Agricultural extent).

We do not utilize the CCRS L1-B processing software because its data and processing requirements cannot be presently met by AAFC to generate weekly composites in near-real-time.

Figure 16(a): Correlation between NCMS-P L-2G and CCRS L-1B weekly Max-NDVI (total and average) for Census Agriculture Regions (CAR) for various growing season weeks in 2006.
Figure 16(b): Correlation between NCMS-P L-2G and CCRS L-1B weekly Max-NDVI (total and average) for Census Subdivisions (CSD) for various growing season weeks in 2006.

Figure 16(c): Correlation between NCMS-P L-2G and CCRS L-1B weekly Max-NDVI (total and average) for Dissemination Areas (DA) for various growing season weeks in 2006.
Figure 16(d): Correlation between NCMS-P L-2G and CCRS L-1B weekly Max-NDVI (total and average) for Townships (Twp) for various growing season weeks in 2006.

6. BIBLIOGRAPHY


Cracknell, A.P. (2001). The exciting and titally unanticipated asuccess of the AVHRR in applications for which it was never intended. Advances in Space Research, 28, 233-240


APPENDIX A: MODIS SIN-PROJECTED TILES (GRANULES)
APPENDIX B: MODIS SEARCH QUERY PARAMETERS

Parameters used in MODIS search query `MODIS_SEARCH.QUERY.TXT`:

```python
GROUP = INVENTORY_SEARCH
MESSAGE_ID = "0"
DATASET_ID = {
    "MODIS/Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V005",
    "MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid V005"
}
GROUP = RANGE_LOC
NORTH_LATITUDE = 90.0000
SOUTH_LATITUDE = -90.0000
WEST_LONGITUDE = -180.0000
EAST_LONGITUDE = 180.0000
END_GROUP = RANGE_LOC
START_DATE = 2009-04-27T00:00:00Z
STOP_DATE = 2009-05-03T23:59:59Z
GRANULE_LIMIT = 1000
GUIDE_LIMIT = 100
TIME_LIMIT = 90
GROUP = EXTENDED_SEARCH
GROUP = SPECIALIZED_CRITERIA
CATEGORY_NAME = "Spatial"
VARIANT = "HORIZONTALTILENUMBER"
CRITERIA_NAME = "HORIZONTALTILENUMBER"
CRITERIA_TYPE = "INTEGER"
CRITERIA_MIN = 9
CRITERIA_MAX = 14
END_GROUP = SPECIALIZED_CRITERIA
GROUP = SPECIALIZED_CRITERIA
CATEGORY_NAME = "Spatial"
VARIANT = "VERTICALTILENUMBER"
CRITERIA_NAME = "VERTICALTILENUMBER"
CRITERIA_TYPE = "INTEGER"
CRITERIA_MIN = 3
CRITERIA_MAX = 4
END_GROUP = SPECIALIZED_CRITERIA
END_GROUP = EXTENDED_SEARCH
SEARCH_NAME = "Modis_Search"
WG_MISSIONS = {
    "edg_terra_modis_land"
}
WG_SRCHTYPE = "INVENTORY_SEARCH"
WG_OPTSTYPE = "All Options"
WG_GEOTYPE = "Type in Lat/Lon Range"
WG_TIMETYPE = "Standard Date Range"
WG_DEPVALIDTYPE = "Valids"
END_GROUP = INVENTORY_SEARCH
END
```
APPENDIX C: ORDERING DATA THROUGH WIST

1. To set up the first search criteria:
   (a) Log in to WIST;
   (b) Click Search. You will be taken to the Primary Data Search screen.
   (c) Click Save/Restore Search.
   (d) Choose a pre-defined query and restore it [Browse | Modis_Search.query.txt | Restore Search Criteria].

2. To order data:
   (a) Log in to WIST;
   (b) Change the start and end dates of the query (if necessary). Dates can be specified as a standard date range (YYYY-MM-DD) or as Julian date (YYYY-DDD):
   (c) Click Start Search
   (d) Query results are returned to the screen (84 granules are returned for a full week).
   (e) Note: In some cases, 84 granules are not available for a 7-day compositing period. If this is the case, the user must decide if there is enough data to warrant creating the 7-day composite. If only a few tiles are missing, then the user can proceed with the order since missing data gaps are “filled” later in the process. However, if there are significant data gaps, the user must consider not creating a composite at all for the week in question. In this case, the process ends here, and data are not ordered.
(f) To order data returned by the query, select all returned tiles (granules) [Select | All | List Data Granules].

Results: Data Set

1 Browse files hosted on http://browse.echo.nasa.gov are currently unavailable. All other browse or online access links will continue to work.

(f) To order data returned by the query, select all returned tiles (granules) [Select | All | List Data Granules].

(g) The data tiles will be listed. Select All then Add selections to cart:

Granule List

1 Browse files hosted on http://browse.echo.nasa.gov are currently unavailable. All other browse or online access links will continue to work.

Select All

(h) If delivery options have already been selected and saved to profile, select Go to step 2: Order Form. If delivery options have not yet been selected and saved to profile, select Change to select and save the preferred delivery options.

Shopping Cart: Step 1: Choose Ordering Options

1 Browse files hosted on http://browse.echo.nasa.gov are currently unavailable. All other browse or online access links will continue to work.

Order ID

Empty Entire Shopping Cart

Order Options | Data Granule ID (Local Granule ID) | Size (MB) | Start Date | Stop Date | Ordered
--- | --- | --- | --- | --- | ---
Delivery Option | SC:MOD09GA.005:2056874515 (MOD09GA.A2000251.H164012340003289891802929112313291123) | 1.092G | 09 Sep 2003, 08:20:00:00.000 | 09 Sep 2003, 23:30:00:00.000 | ☐
Delivery Option | SC:MOD09GA.005:2056874515 (MOD09GA.A2000251.H164012340003289891802929112313291123) | 1.092G | 09 Sep 2003, 08:20:00:00.000 | 09 Sep 2003, 23:30:00:00.000 | ☐
Delivery Option | SC:MOD09GA.005:2056874515 (MOD09GA.A2000251.H164012340003289891802929112313291123) | 1.092G | 09 Sep 2003, 08:20:00:00.000 | 09 Sep 2003, 23:30:00:00.000 | ☐
Delivery Option | SC:MOD09GA.005:2056874515 (MOD09GA.A2000251.H164012340003289891802929112313291123) | 1.092G | 09 Sep 2003, 08:20:00:00.000 | 09 Sep 2003, 23:30:00:00.000 | ☐
(i) If changing delivery options, select Media Type = FTP Pull; Ftp Pull media Format = File. Then select Save these order options as my order preference & apply them (if applicable) to all granules currently in my shopping cart for data set.

Choose Ordering Options

- Media Type: Ftp Pull (Required)
- Ftp Pull Media Format: File (Required)

Select Save these order options as my order preference & apply them (if applicable) to all granules currently in my shopping cart for data set.

(j) Fill out contact details on order form and select Submit Order Now.

(k) A web page will open summarizing your order.
3. Order Confirmation:

WIST will send you four emails concerning your order:

(a) An email from echo@acho.nada.gov will be sent immediately, indicating that you have placed an order:

(b) An email from wist_support@echo.nasa.gov will summarize the order in more detail:

(c) An email from custserv@usgs.gov will also detail the order, the delivery format and the cost (free):
(d) An email from lpdaac@eos.nasa.gov will arrive containing the ftp download location:

(e) Data can be downloaded from the ftp site given using the NcFTP FTP client described earlier in this document.

4. Notes:

The time taken for data requests to be processed by WIST varies, and depends on (a) the number of requests the DAAC is processing at that given time, and (b) the volume of data in an order. For example, small data orders (e.g. a few tiles) are often available for download within an hour of the order being placed. In comparison, large orders (e.g. one week of data comprising a total of 168 tiles) are usually available 24 hours after the order has been placed. In addition, queries submitted at “low traffic” times (late evening, early morning) are generally processed more quickly than queries submitted during peak times (9 to 5pm).

APPENDIX D: BATCH FILE FOR DOWNLOADING MODIS HDF TILES

@echo off
:: FTPdatadownload.bat
::
:: This batch file automatically downloads MODIS data from their ftp site
:: using NcFTP FTP client software. The -z flag makes sure that files
:: already downloaded are not overwritten if the download is restarted for
:: any reason.
::
:: **Note**: If a download fails and needs to be restarted, the file being
:: downloaded at time of failure will only be partially downloaded. This
:: file should be deleted before restarting the download.

mkdir 2009117to123
cd 2009117to123
ncftpget ftp://e4ftl01u.ecs.nasa.gov/PullDir/0301617268QdcAPB/* -z
APPENDIX E: IMPLEMENTING NDVICOMP, NDVIBASE AND NDVIANOM IN ARCGIS V9.3 AND PYTHON V9.2 [VIDEO]

All python scripts should be able to be implemented with ArcGIS9.3 and Python 2.5. The following, however, should be done:

(a) Go to the C:\Program Files\ArcGIS\Bin directory and run the file EsriPython251.exe to install python v2.5 to a local directory on your computer (usually, C:/python25/).

(b) Use windows explorer to match the .py extension with the python v2.5 executable (Tools | Folder Options | File Types | Select .py | Choose the yellow and blue icon that matches the python2.5 executable).

(c) Open the Max-NDVI compositing tool (NDVICOMP.py) and go to line 67. Change the python path to: python_path = 'c:/Python25/python.exe'

(d) Restart your PC.

APPENDIX F: MAX-NDVI COMPOSITING TUTORIAL (DVD DEMO)

This demonstration uses MODIS data for 2 dates (days of year 2008189 and 2008190) and 2 tiles (h10v03; h10v04) to create a 2-day 2-tile Max-NDVI composite. Before processing, you must download the contents of the demonstration DVD to your local drive directory C:\MODPROC\.

1. DATA ACQUISITION [VIDEO] and DOWNLOAD [VIDEO]

1.1 Use data provided on DVD (Skip order and download)

If you are using the data provided on DVD, skip to step 2.

1.2. Order and download data from WIST

Login to WIST and order MOD09GQ and MOD09GA data for the above two tiles and dates following the methods described on Appendix C. To do this, use the query file

C:\MODPROC\1.Query_and_download_from_USGS\Modis_Search.query.

You will receive an email from the NASA LP DAAC containing the FTP location of the data and data downloading instructions. Copy and paste the FTP location into the following batch file:

C:\MODPROC\1.Query_and_download_from_USGS\FTPdatadownload.bat.

Make sure that the NCTFP ftp client is installed correctly on your PC and double click the batch file. The batch file will download the .hdf tiles to the directory:

C:\MODPROC\2.Extract_SDs_From_HDF_Download\demo_HDF\
2. DATA EXTRACTION [VIDEO]

2.1. Installing the MODIS Reprojection Tool (MRT)

Install the MRT as directed in the MRT installation guide. The MRT Tool is located on C:\MODPROC\9.Software\.

(a) Create a new folder on the C drive (C:\MRTv4),
(b) Extract the contents of the zipped file to the new folder.
(c) Double click / run file: C:\MRTv4\install.bat
(d) When prompted for bin directory: Enter C:/MRTv4/
(e) When prompted for windows version: Enter number from list (1=XP or later).
(f) When prompted for Java** bin directory: Enter C:/Program Files/Java/jre1.6.0_03/bin (or whatever version you have installed).
(g) Press <Enter> to finish MRT installation.
(h) Check Environment variables (Settings | Control Panel | System | Advanced | Environment Variables) MRT_HOME and MRTDATADIR are correctly set.
(i) Restart computer.

* Note that DOS paths use forward slashes to separate directories ("/").

** Java v 1.50 or higher needed [www.java.com/downloads]

2.2. Extracting Scientific Datasets using the MRT

(a) Open C:\MODPROC\2.Extract_SDs_From_HDF_Download\Run_MRT_Demo.bat in a text editor.
(b) Edit the directory location of various files:
   • Edit the input file directory to its current HDF data directory (the default specified is demo_HDF).
   • Edit the location of the attribute.prm file to its current location (the default is the same directory as Run_MRT_Demo.bat).
   • Check that the location of resample.exe is C:\mrtv4\bin\resample.
(c) Double-click MRT.bat to execute.
(d) Output files are created in the demo_HDF\TIFF directory.
(e) Check the file resample.log for any errors after the script has been executed.
3. CHECK FOR MISSING DATA [VIDEO]

Check for missing Geotiff files using the python checking tool.

(a) To start the checking, open the batch file:

C:\MODPROC\D\3.Check_for_missing_data\runC4MDTIF_demo.bat

(b) Replace parameters if need be with the current program and data directories

(c) Move or delete one of the Geotiff files and rerun Check4MissingDataTIF.py in report and replace mode (argument #5 = 1). Note how the file is replaced with a copy from the existing dataset.

4. NDVI COMPOSITING [VIDEO]

Run the NDVI Compositing Tool

(a) To start NDVI compositing, open the batch file:

C:\MODPROC\D\4.NDay_NDVI_Compositing\run_NDVI_demo.bat

(b) Replace parameters with the current program and data directories

(c) Replace parameters with the current program and data directories