Natural Resources Canada | Ressources naturelles Canada

GEOMATICS CANADA
OPEN FILE 6

Moderate Resolution Time Series Data Management and Analysis: Automated Large Area Mosaicking and Quality Control

R. Latifovic, D. Pouliot, L. Sun, J. Schwarz, and W. Parkinson

2015

Canada

GEOMATICS CANADA
OPEN FILE 6

Moderate Resolution Time Series Data Management and Analysis: Automated Large Area Mosaicking and Quality Control

R. Latifovic, D. Pouliot, L. Sun, J. Schwarz, and W. Parkinson

2015

©Her Majesty the Queen in Right of Canada, as represented by the Minister of Natural Resources Canada, 2015

doi:10.4095/296204

This publication is available for free download through GEOSCAN (http://geoscan.nrcan.gc.ca/).

Recommended citation
Latifovic, R., Pouliot, D., Sun, L., Schwarz, J., and Parkinson, W., 2015. Moderate Resolution Time Series Data Management and Analysis: Automated Large Area Mosaicking and Quality Control; Geomatics Canada, Open File 6, 25 p. doi:10.4095/296204

Publications in this series have not been edited; they are released as submitted by the author.

Disclaimer

Her Majesty the Queen in right of Canada, as represented by the Minister of Natural Resources ("Canada"), does not warrant or guarantee the accuracy or completeness of the information ("Data") on these maps and does not assume any responsibility or liability with respect to any damage or loss arising from the use or interpretation of the Data. The Data on these maps are intended to convey regional trends and should be used as a guide only.

ABSTRACT

The Canada Centre for Remote Sensing[1] (CCRS) maintains national-scale Long Term Satellite Data Records (LTSDRs) as an essential component of Earth Observation (EO) based land surface monitoring. The CCRS LTSDR framework provides long-term capability to generate, archive and provide access to value-added satellite data and thematic products addressing various land surface monitoring needs of the Government of Canada. For many years, coarse-resolution LTSDRs supported basic land-cover and land-use information needs over large areas. While these LTSDRs, with pixel sizes between 250m and 1000m, are important for ongoing long-term time series analysis, increasingly there is an opportunity to use greater spatial resolution data to more effectively address monitoring and assessment of both anthropogenic and natural land surface changes.

Until recently, cost and availability limited the usefulness of medium resolution (~30m pixel size) EO data for such analyses. Then, in 2009, the United States Geological Survey made Landsat data freely available. The potential for medium-resolution time series monitoring has been further strengthened by Landsat-8 and the pending launch of ESA's Sentinels. For this potential to be realized, new methods and algorithms are required to extract and analyse information from the medium resolution data, such as Landsat Time Series data, to monitor aspects of land surface dynamics. CCRS' new medium resolution LTSDR framework, the Time Series Data Management and Analysis System (TSDMAS), generates and manages value-added LTSDRs based on the TM, ETM and OLI sensors on board the Landsat 5, 7 and 8 missions, respectively. An overview of the TSDAMS system, and the algorithms implemented therein, will be presented.

A new value-added data product is presented: a Top of Atmosphere Reflectance Coverage of Canada, at 30 m spatial resolution. This circa 2010 product has been generated by the TSDAMS from data acquired by the TM and ETM sensors. Product generation, quality control, and characteristics of the underlying dataset are described. By providing readily available, national-scale Landsat data products, of "research quality", CCRS' TSDMAS harnesses the potential of medium resolution EO data, and can exponentially increase the downstream generation and use of medium resolution land surface information products. The new system and example data product which are described were designed to assist government agencies, the scientific community, natural resources managers and non-governmental groups engaged in land cover mapping, and the generation of geophysical and biophysical products for the assessment of surface dynamics at national and regional scales.

[1] A division of Natural Resources Canada's Canada Centre for Mapping and Earth Observation (CCMEO)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The Canada Centre for Remote Sensing (CCRS) maintains national-scale Long Term Satellite Data Records (LTSDR) as an essential component of Earth Observation (EO) based land surface monitoring. The CCRS LTSDR framework provides long-term capability to generate, archive and provide access to value-added satellite data and thematic products addressing various land surface monitoring needs of the Government of Canada. LTSDRs have been used by various federal departments in support of a broad range of endeavours, such as sustainably developing natural resources, assessing climate change impacts and adaptation, monitoring ecological integrity of protected lands, estimating carbon emissions, and mapping changes which affect land cover and water availability. While coarse-resolution EO data, such as those underpinning the existing LTSDR, have been used for many years to derive basic land-cover and land-use information over large areas, EO data with greater spatial resolution is now readily available. Jensen and Cowen (1999) and Franklin et al. (2003) illustrate that data and products with pixel sizes between 250 and 1000m do not satisfy all requirements for adequate monitoring and reconstruction of both anthropogenic and natural land surface changes.

Until recently the primary factors limiting use of medium resolution remote sensing data (~30m pixel size) over large areas were cost and availability. In 2009, the United States Geological Survey (USGS) made Landsat data freely available (Woodcock et al 2008), facilitating the use of higher-resolution time series for monitoring land cover dynamics. Gutman et al., 2008; and Wulder et. al., 2008 consider medium-spatial resolution sensors, with 30m pixel size, appropriate for today's land surface motoring requirements. In fact, there has been an increase in the number of published approaches for generating large area Landsat datasets for national monitoring and reporting needs across land-use and resource sectors. Numerous datasets, methods and algorithms for extracting information from the Landsat Time Series have been developed to monitor aspects of land change dynamics including spatial and temporal patterns of forest disturbance in Cohen et al., 2010; Huang et al., 2010; Kennedy et al., 2010; Zhu et al., 2012, forest structure in Pflugmacher et al., 2012, forest biomass in Powell et al., 2010 and gradual change in Vogelmann et al., 2012.

Most of the recently established national land-surface monitoring programs are based on medium resolution satellite data such as USA National Land Cover Database Jin et al., 2011, and the Australian Collaborative Land Use and Management Program (ACLUMP) etc. The USGS' new "open data" policy also enables improved land surface monitoring in Canada, allowing CCRS to address limitations of the coarse resolution LTSDRs.

In this study, CCRS' new medium resolution LTSDR framework, the Time Series Data Management and Analysis System (TSDMAS), is presented. It generates and manages value-added LTSDRs based on the TM, ETM and OLI sensors on board the Landsat 5, 7 and 8 missions, respectively. An overview of the TSDAMS system, and the algorithms implemented therein, is presented.

A new value-added data product is also presented: a Top of Atmosphere Reflectance Coverage of Canada, at 30 m spatial resolution. This circa 2010 product has been generated by the TSDAMS from data acquired by the TM and ETM sensors. Product generation, quality control, and characteristics of the underlying dataset are described. By providing readily available, national-scale Landsat data products, of "research quality", CCRS' TSDMAS harnesses the potential of medium resolution EO data, and can exponentially increase the downstream generation and use of medium resolution land surface information products.

The new system and example data product which are described have been designed to assist government agencies, the scientific community, natural resources managers and non-governmental groups engaged in land cover mapping, and the generation of geophysical and biophysical products for the assessment of surface dynamics at national and regional scales.

2. METHOD

2.1 Landsat time series database

In total, 13,350 Landsat TM/ETM+ scenes were acquired from the U.S. Geological Survey to produce a complete coverage of Canada. Scenes spanned the 2009-2011 period and were all acquired in July and August. Of the scenes used, 48% were acquired by TM and 52% by ETM. The temporal distribution (month and year) of all available scenes, by sensor, is shown in Fig.1.

The acquired scenes were level-one-terrain-corrected products (L1T) and were available in GeoTIFF format in the Universal Transverse Mercator (UTM) map projection based on the World Geodetic System 84 (WGS84) datum. The processing of the Level one product includes radiometric calibration, systematic geometric correction, precision correction using ground control points (GCP) and orthorectification. Average sub-pixel accuracy of the L1T product is generally adequate and additional geometric processing in most cases was not required. However, certain acquisitions with significant cloud cover did not have a sufficient number of GCPs or good quality elevation data necessary for precise geolocation. For scenes with georeferencing errors, an additional geometric rectification was applied to the L1T product using additional manually collected GCPs.

The UTM map projection, used for the acquired L1T Landsat data, is not an appropriate projection for large areas that cover several UTM zones. The Lambert Conformal Conic (LCC) projection is typically used in Canada for large area spatial datasets as it does not require separate zones, yet keeps distortion to an acceptable level. The LCC projection was therefore used, with the added advantage of consistency with CCRS' other LTSDR and national vector products. The LCC projection is specified with two standard parallels at 49° N and 77° N and central meridian at -95° (Table 1).

Table 1. The parameters of the Lambert Conformal Conic (LCC) projection and Earth ellipsoid model used for output imagery over Canada

Parameter	Value
Earth ellipsoid	GRS 1980
Major semi-axis, a	6378137 [m]
First eccentricity	0.00669438002290
Ellipsoid flattening, f	0.00335281068118
Projection	LCC
1st parallel	49.00 [degree]
2nd parallel	77.00 [degree]
Central meridian	–95.00 [degree]
Upper left corner	(–2600000.0 E [m]; 10500000.0 N [m])
Lower right corner	(3100000.0 E [m]; 5700000.0 N [m])
Easting	0
Northing	0
Gridbox size, x	30 [m]
Gridbox size, y	30 [m]

A Landsat time series database covering all of Canada was created, based on an ingestion procedure which included i) setting up database structure, ii) defining scene boundaries to coincide with a common national spatial framework, iii) re-projecting all scenes into LCC using the nearest neighbor resampling algorithm, iv) saving images in binary raw format and v) standardizing metadata. Subsequent processing steps included cloud screening and cloud shadow detection. The compositing procedure used to generate the Landsat-based mosaic of Canada, circa 2010, is one of several applications developed to compliment TSDMAS.

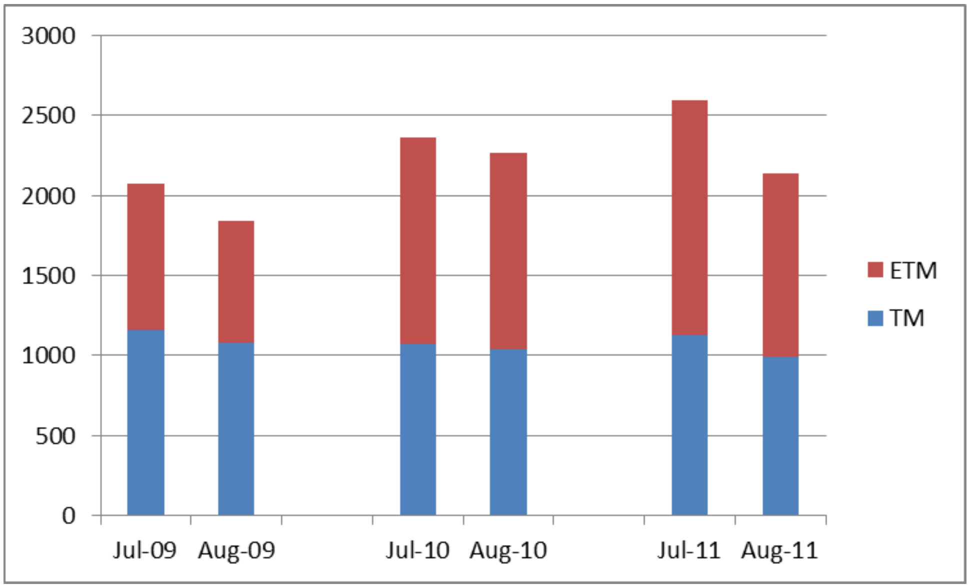

Figure 1. Number of available Landsat scenes 2009-2011 during July-August period.

2.2 Processing approach

The Landsat coverage of Canada was generated using a new image processing environment referred to as the Time Series Data Management and Analysis System (TSDMA, Latifovic et al., 2014) The TSDMA manages data records acquired by TM Landsat 5, ETM Landsat 7 and OLI Landsat 8 sensors processed to L1T. The TSDMA system is comprised of the following functional modules:

(1) Data Base Management: includes data structure and import, metadata, re-projection and data format standardization;
(2) Geometry: includes resampling, geolocation quality control and refinement;
(3) Scene Identification: includes cloud, cloud shadow, water, snow and land identification;
(4) Radiometry: includes atmospheric correction and scene normalization; and
(5) Time Series Analysis Toolbox: includes compositing, trend analysis, gap filling / interpolation and machine learning and computer vision algorithms.

A detailed description of the algorithms implemented in the TSDMA system will be provided in a separate paper. This article provides only a short description of the algorithms used to generate the product presented here.

2.3 Top of Atmosphere Reflectance Computation

The spectral radiance measured by the TM and ETM+ sensors is stored as an 8-bit digital number (Markham et al. 2006). Digital numbers are converted back into spectral radiance (W m^{-2} sr^{-1} μm^{-1}) using sensor calibration gain and bias coefficients provided in the L1T metadata (Chander et al., 2009). Spectral radiances are then converted to reflectance at the top of atmosphere using the standard formula

$$\rho_\lambda = \frac{\pi\, L_\lambda\, d^2}{ESUN_\lambda\, \theta_s}$$

where ρ_λ is the top of atmosphere (TOA) reflectance (unitless), L_λ is the TOA spectral radiance (W m^{-2} sr^{-1} μm^{-1}), d is the Earth-Sun distance (astronomical units) , $ESUN_\lambda$ is the mean TOA solar spectral irradiance provided by Chander at al., 2009, and θ_s is solar zenith angle (radians).

The radiance sensed in the Landsat low and high gain thermal bands were converted to TOA brightness temperature (i.e., assuming unit surface emissivity) using the standard formula

$$T = \frac{K_2}{log\left(\frac{K_1}{L_\lambda + 1}\right)}$$

where T is the 10.40–12.50 μm TOA brightness temperature (Kelvin), K1 and K2 are thermal calibration constants set as 666.09 (W m−2 sr−1 μm−1) and 1282.71 (Kelvin), respectively, and Lλ is the TOA spectral radiance (Chander et al., 2009).

2.4 Scene Identification

Clouds and cloud shadows are a significant source of noise in Landsat data. A number of methods have been developed for cloud identification with variable success. FMask, one of the latest algorithms (Zhu and Woodcock, 2012), builds on the results of previous approaches and is currently widely used by the remote sensing community. It detects both clouds and cloud shadows for Landsat TM and ETM images. In the first pass, the FMask algorithm uses basic spectral tests such as Normalized Difference Vegetation Index (NDVI), Normalized Difference Snow Index (NDSI) and temperature, Whiteness, Haze Optimized Transformation (HOT), and B4/B5 to identify likely cloud pixels. After separating likely cloud pixels in the first pass, the remaining pixels are considered clear-sky pixels and are used to compute cloud probability for all pixels in the image. FMask computes cloud probability separately for water and land. To classify pixels as land or water, the Water Test is applied in pass one (Zhu and Woodcock, 2012).

In the current study, the FMask algorithm for cloud detection was implemented in TSDMA system and applied to the Landsat scenes over Canada. This algorithm works well in high latitudes for snow free conditions, separating clouds from snow and detecting thin clouds. A separate cloud shadow identification algorithm (Fig. 2) was developed and applied after FMask cloud detection.

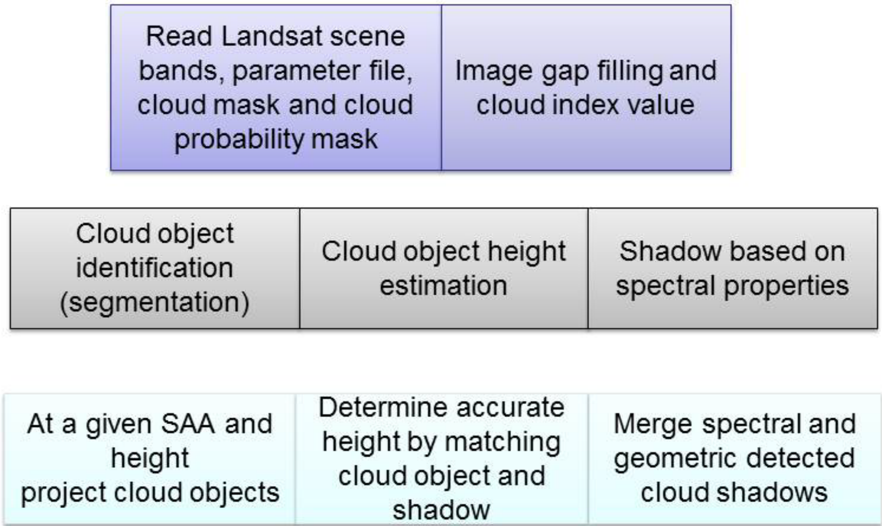

Figure 2. Cloud shadow identification algorithm

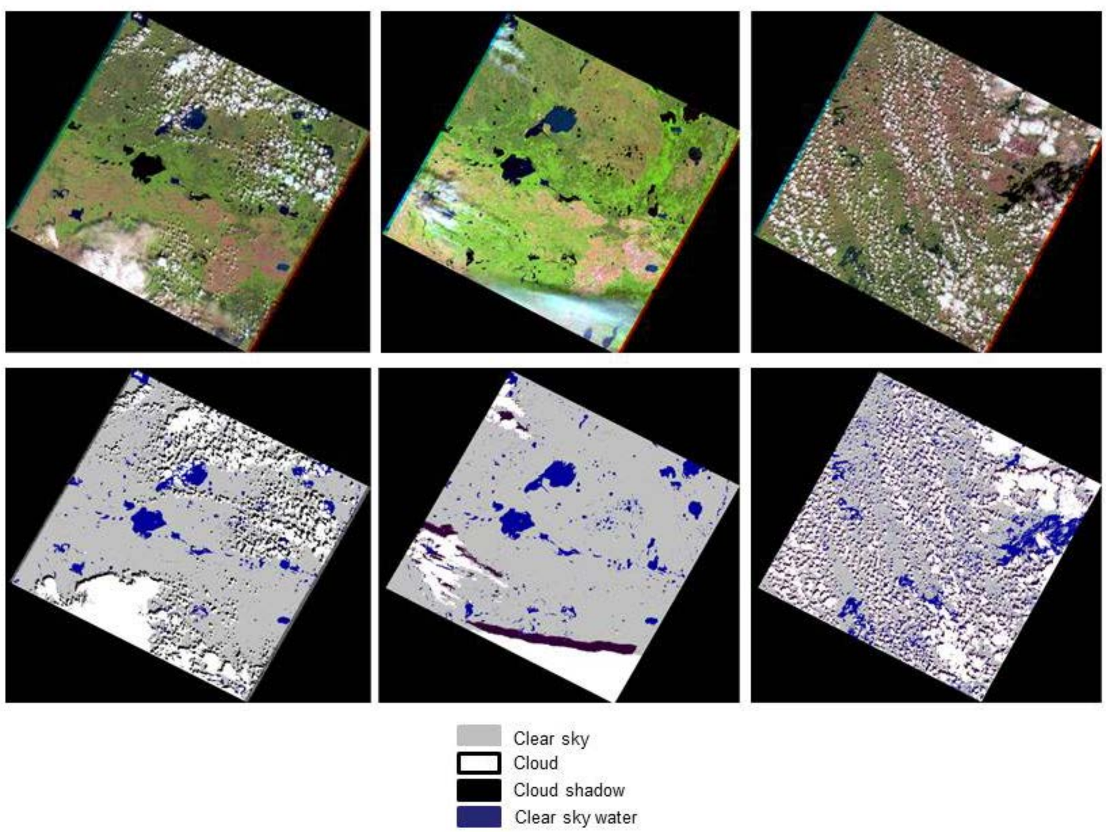

Figure 3. Cloud and cloud shadow examples

2.5 Compositing

Compositing Landsat scenes over particular time periods to produce a single image with the best selection of clear-sky pixels, is a standard procedure. It allows for the production of datasets with reduced contamination from atmosphere, clouds, and cloud shadows. Several compositing methods have been used historically (Cihlar et al., 1994). They are usually variants of so-called “maximum (or minimum) value compositing (MVC)” techniques. However, single criteria selection algorithms may perform poorly given certain land cover types or data acquisition conditions. For example Max NDVI does not perform well over land with no, low or sparse vegetation (e.g. some arid regions, burned land, bare rock, water bodies). Likewise, Min Red’s performance is compromised by selecting shadow over clear-sky observations. To overcome these problems, a new compositing technique was developed and used for generating the current data set. It relies, firstly, on the cloud and cloud shadow masks. Then, identification of the “best” pixel among available observations not flagged as cloud or shadow is based on a multi-criteria selection algorithm. The algorithm combines five criteria (Table 2) designed to detect remaining cloud shadow and haze. All available observations for a given pixel are ranked by each criterion and the final score for each observation is computed as the weighed sum of all rankings.

In order to define a set of weight coefficients, composite images produced with different weights were analyzed for variability of spectral reflectance within spectrally stable land cover types. The expectation was that the best set of coefficients would result in the smallest variability in spectral reflectance for areas with homogenous land cover. However, different sets of weights were found to produce the best result for different land cover types. For example, compositing with greater weight placed on minimal red criteria produced better results over water, while compositing with greater weight on the vegetation indices would produce the most stable composite image for forested land. Therefore, based on scene identification and vegetation indexes, different weight coefficients were applied for water, snow and land pixels.

To provide a clear sky observation for each location, data from years 2009, 2010, and 2011 were used, but 2010 was preferentially selected where possible. Due to differences in ground target conditions (moisture and phenology) and atmosphere, the variability of composite results over homogenous areas decreased if adjacent pixels were selected from the same scene. To increase the probability of selecting from 2010, if clear sky observations for that year were available, selection criteria computed as time difference between the middle of the compositing period (e.g. July 31, 2010) and the actual time of the acquisition were included in the composite scoring scheme.

Table 2. Multitemporal selection criteria

Compositing criteria	Description and rationale
Max NDVI	appearance of clouds, poor atmospheric conditions, and off-nadir viewing geometries would depress the NDVI values disadvantage: poor performance over water and non-vegetated land
Min Red	appearance of clouds, poor atmospheric conditions would increase Red reflectance disadvantage: preferable selection of cloud shadows over truly clear-sky pixels and discriminating snow from clouds
Max (NIR/Blue, SWIR1/Blue)	a certain equivalent of cloud–haze optical depth
(Blue-Red)/(SWIR1+1)	a certain equivalent of cloud–haze optical depth
Min (Delta Time)	prefer clear pixels closer to the middle of the compositing period

Compositing was performed on 25 tiles, as opposed to one large image, to produce images that could be handled with standard desktop computers. Fig. 4 shows the tile system, size and naming. Processing by tiles was necessary due to the size of the data set, which included 13,350 Landsat scenes covering 10 million square kilometers. Compositing by tiles facilitates parallel processing, offers manageable data size, and enables easier loading into viewing tools for performing of visual quality control.

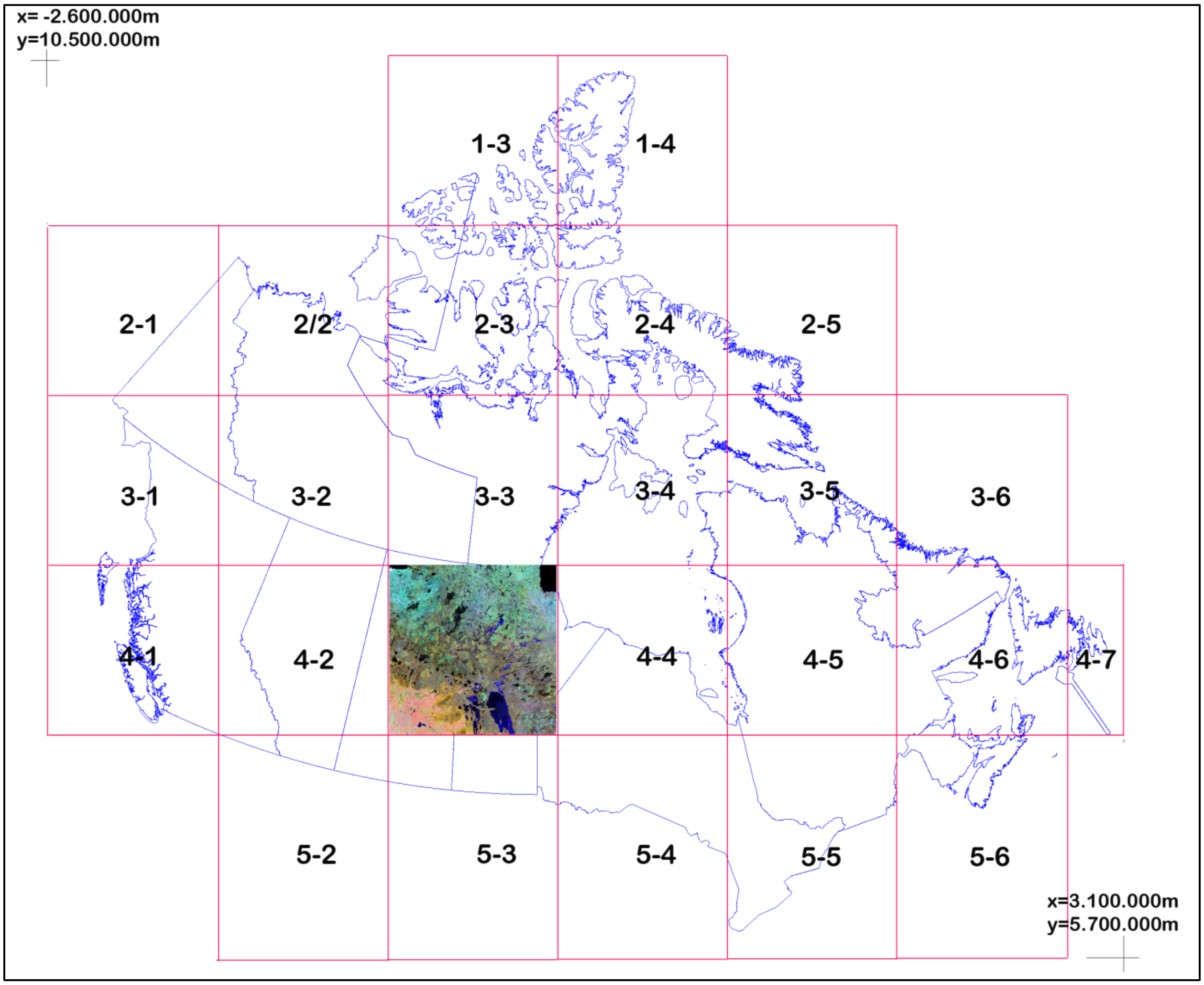

Figure 4 Tile arrangement, size and naming convention used in processing

2.6 Scan-Line Striping Correction in Composite Landsat 5 and 7 Data Products

The Scan Line Corrector (SLC) which should compensate for the forward motion of Landsat 7 failed on May 31, 2003, causing missing scan lines in acquired data. In composite products, these missing data are in-filled with data from other dates. However, differences in atmosphere and ground target conditions (phenology and moisture) at acquisition pose problems for this approach. Atmosphere correction is often ineffective, due to unavailability of detailed spatial-temporal information on atmosphere conditions, and no correction is possible for ground target variability. This leads to strong systematic spatial striping in composites including Landsat 7 data, and errors in the 5-10% range. Such variability is sufficient to introduce errors in land cover classification and change detection applications. Figure 5 a) shows an example of the striping problem in a Landsat 5-7 composite product. To correct this effect, a frequency space filter approach was used. An example result of the correction is shown in Figure 5 b).

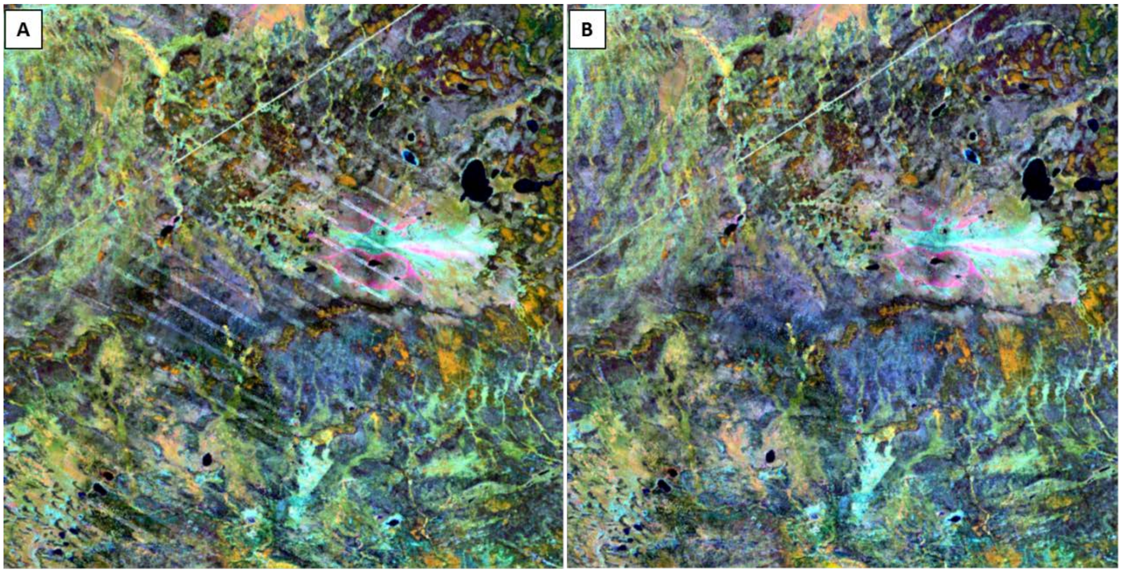

Figure 5. a) Example of SLC-Off scan line noise in composite results. b) Results of the FFT-correction.

The correction is based on a Fast Fourier Transformation (FFT) of the image to frequency space. The systematic orientation of the scan-line errors causes predictable anomalies in frequency space. Figure 6 a) shows an example of frequency space for a composite product affected by scan-lines errors. The anomalies are clearly seen in the figure as local maxima in the frequency surface. While the orientation is predictable, the frequency is only partially predictable since very low and very high frequencies are not affected. Based on the orientation angle and frequency range, a mask is defined in frequency space to search for local maxima. Local maxima are determined as the 99th percentile of the frequency data within the mask. Values above this threshold are identified as local maxima and their value set to zero. Thus these frequencies are ignored in the inverse-FFT, eliminating the striping effect. However, to avoid ringing effects in the inverse-FFT, which can occur with hard frequency space boundaries, a Gaussian filter is applied to soften the edges around identified local maxima. An example of the detected local maxima is shown in Figure 6 b).

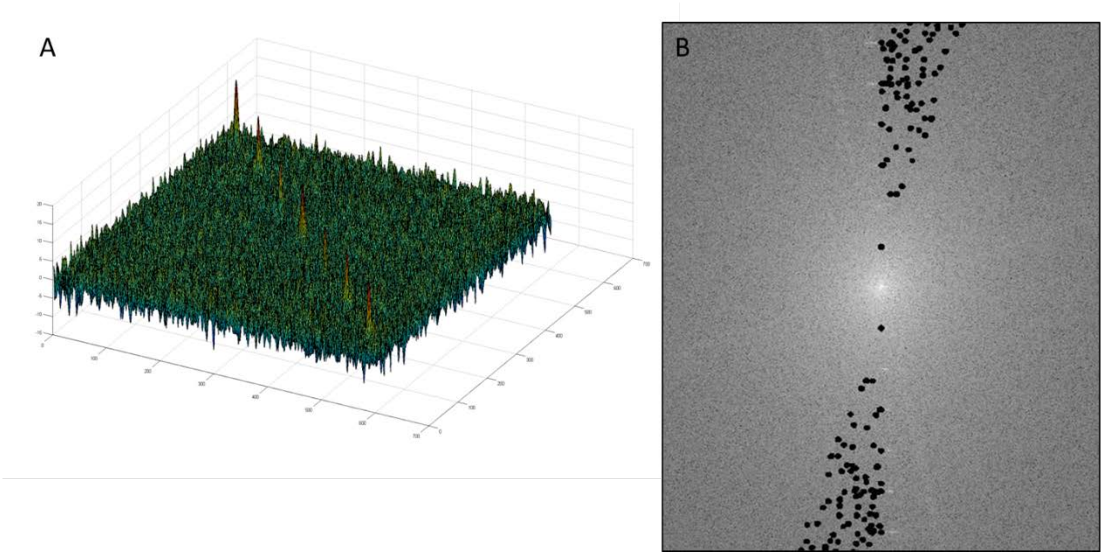

Figure 6. a) Frequency space anomalies associated with a scan-line error from composite infill. b) Results of anomaly detection and masking in frequency space.

To predict the scan-line error orientation angle in the composite imagery, a model was developed using the Easting and Northing coordinates for samples of orientation angle collected across Canada. In total, 100 locations were sampled. Fig. 7 a) shows the relation between the orientation angle and Easting, while Fig. 7 b) shows the relation with Northing. The Northing was the stronger factor in the prediction. Exploratory analysis revealed that a multivariate regression with Easting and Northern terms to the fourth order was the most effective for prediction. Further enhancement was obtained by developing a piece-wise model for four ranges of Easting. The performance of the final model is shown in Figure 7 c). On average, model error was less than 4%. In implementation, the frequency space mask was set to ±10% of the predicted orientation angle to account for model error.

In regions where striping artifacts were present, the FFT-correction was effective. However, in areas where no striping existed, the correction itself sometimes created some small-magnitude striping. To avoid this, a method was developed to identify regions in image composites where FFT-correction was required. This was implemented by examining the difference between edge detection, before and after correction, using a Sobel edge operator. In regions where the correction was effective, the total edge magnitude was reduced, whereas in areas where the correction was not required, edge magnitude was increased. A threshold was determined, on an image by image basis, to ensure the FFT-correction caused minimal image distortion.

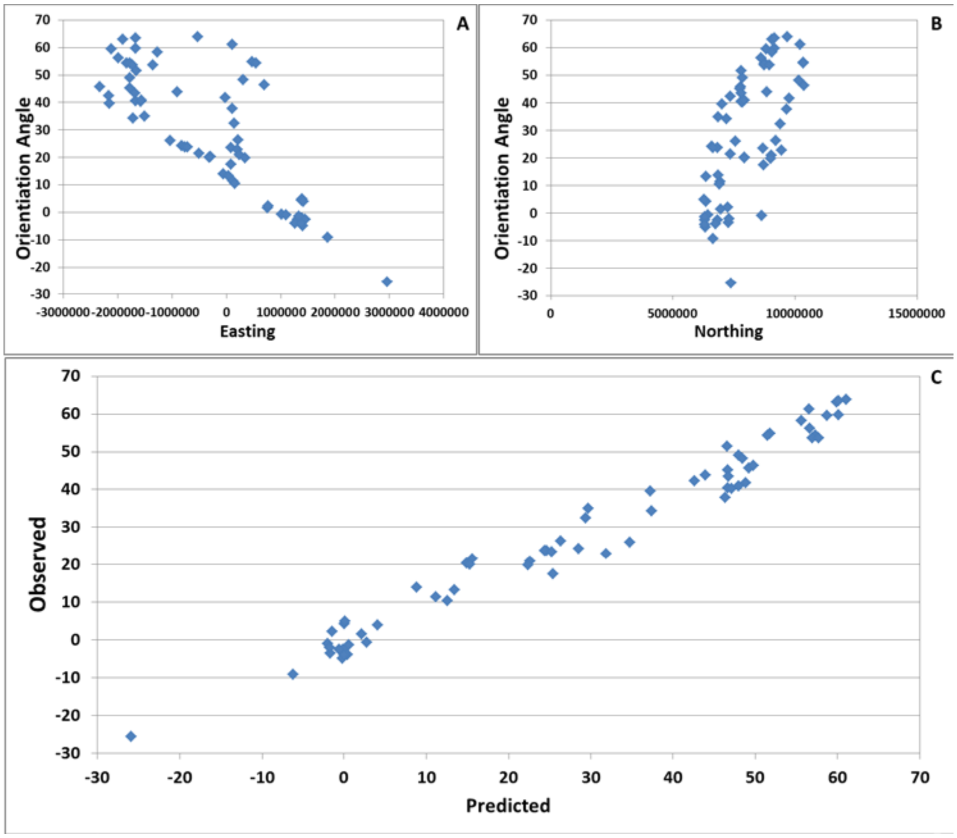

Figure 7. Model used to predict SLC-OFF orientation angle for input to FFT correction. a) Relation of samples with Easting, b) relation of samples with Northing, c) final model results between observed and predicted samples.

2.7 Quality control

A quality control (Q/C) procedure was implemented as an integral part of the process of generating large area Landsat composite products. A systematic Q/C approach is required to ensure quality requirements for both the process and product are met. Following each processing step, Q/C determined the appropriate action: i) proceed with the next processing step, or ii) reiterate previous steps with adjusted processing parameters. The Q/C procedure is semi-automate: it relies on visual inspection to detect problems while refinement is, in most cases, automated. The Q/C procedure workflow (Fig. 8) visually assesses all tiles by undertaking the following steps: 1) overlaying a grid to ensure an entire tile will be systematically assessed, 2) inspecting georeferencing quality for each grid cell, 3) determining if there are areas with missing data, 4) detecting the presence of cloud and cloud shadow, and 5) detecting the presence of scan line striping.

If a tile georeferencing error is detected, the correction procedure follows two sub-steps: i) determination of the scenes with georeferencing error, and ii) removal of the scenes with error from the list of available scenes if there are sufficient alternative scenes to cover the area in question. In instances where there are no alternative scenes in the original dataset i.e. a scene provides observations critical to the output composite, geometric correction is performed based on manual collection of GCPs to develop a 1st to 3rd polynomial model. The correction procedure in cases of missing data involves acquiring additional scenes from the USGS archive with extended time window. For the south, the acceptable temporal range for additional scenes was extended to include the last two weeks of June and the first two weeks of September. In the north, where phenology is more pronounced and land surface changes are less common, scenes from either 2008 or 2012 were sought.

Should the Q/C procedure detect residual cloud contamination (quality control step 4), despite all scenes having passed through cloud and cloud shadow masking, problematic tiles were reprocessed to rectify the problem. In case 5) when scan-line stripping is detected FFT correction and masking procedure described in section 2.2 was applied.

Additional procedures available for further improving the visual aspect of the data includes despike correction, radiometric adjustment and custom correction applied on composite. For this Landsat Mosaic of Canada, despike correction was not applied. Custom correction was developed and applied to reduce SLC-Off scan line noise over water bodies.

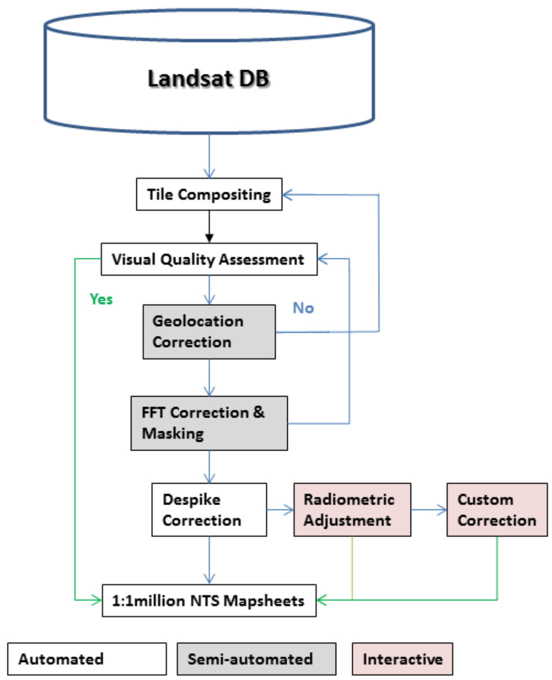

Figure 8. Tile quality control flow chart.

3. RESULTS AND DISCUSSION

Fig. 9 shows an image of the new July-August, circa 2010, Landsat Mosaic of Canada, created from the composite tiles. This data product has been developed for land cover mapping applications and biophysical parameters retrieval across Canada.

Many other applications of these data are also possible, but the user should consider the temporal and spectral limitations of the product. In order to reduce the significant phenological differences which result from Canada's strong gradient in vegetation growing conditions, particularly south-north, a core acquisition time window of just two months (July-August, 2010) was selected. Since such a short period cannot provide clear-sky coverage over Canada, July-August periods from adjacent years (2009 and 2011) were used to compliment the 2010 coverage. The distribution of selected observations, by year of acquisition and by tile, is provided in Fig 10. Overall, 10% of observations were selected from 2009, 70% from 2010 and 30% from 2011.

The use of pixels from multiple years could create inconsistencies in areas where land surface changes occurred in the 2009 – 2011 period, especially if adjacent pixels were selected from years before and after a significant change. To minimize this possibility, the weight for the compositing criteria was adjusted for selecting observations from 2010 over no change areas and from 2011 over changed areas. Guindon et al. (2014) suggest that approximate 4 million hectares of change occurred due to fires and harvesting between 2009 and 2010. If the compositing procedure was completely biased to select 2009 instead of 2010 pixels for areas with change, then 10% of the total 2009-2010 would occur. Relative to the forested area in Canada this is a very small around (e.g. 0.1%). Since the composite is, in fact, largely based on 2010 pixels, including in areas with change, this worse-case scenario cannot occur.

For each pixel, a scene ID with time of acquisition is provided as a separate layer. Thus, if data are used in studies which involve temporal aspects, the time-of-acquisition information provided can and should be considered. For change detection studies, a more appropriate approach would be to consider all available observations in a time series analysis instead of just the one provided in this composite. The same is suggested for water quality studies because this composite includes data from different times, and water quality conditions would understandably vary with time. For those interested in mapping water body extent, the new composite is perhaps sufficient, assuming variability during the three year period is not too significant an issue.

The quality control procedure described in section 2.7 was time consuming but ultimately ensured the highest possible product quality. After the initial processing, almost all tiles had to be improved in one or more quality aspects through iterative reprocessing. The Scan-Line Striping Correction had to be applied on each tile. One third of tiles had one or more scenes where georeferencing correction was required. In a number of scenes covering Canadian north, the problem was inadequate cloud shadow masking and confusion between water bodies and cloud shadow. In some cases, additional criteria were used to separate cloud shadow and water bodies, while in other cases, scenes with high cloud coverage were removed from the list of available scenes. Removal of scenes was not problematic due to the high number of observations over Canada's north, where considerable scene overlap occurs (see Fig.11).

Bands 3 (0.63-0.69 µm), 4 (0.76-0.90 µm), 5 (1.55-1.75 µm), and 7 (2.08-2.35 µm) are provided in this version of the composite. As significant atmosphere effects strongly limit the quality of the blue (0.45-0.52 µm) and green (0.52-0.60 µm) bands, they have not been provided. An atmospherically corrected version is being developed for future release. Research to reduce the temporal range used to capture clear sky observations is also in development as a future enhancement.

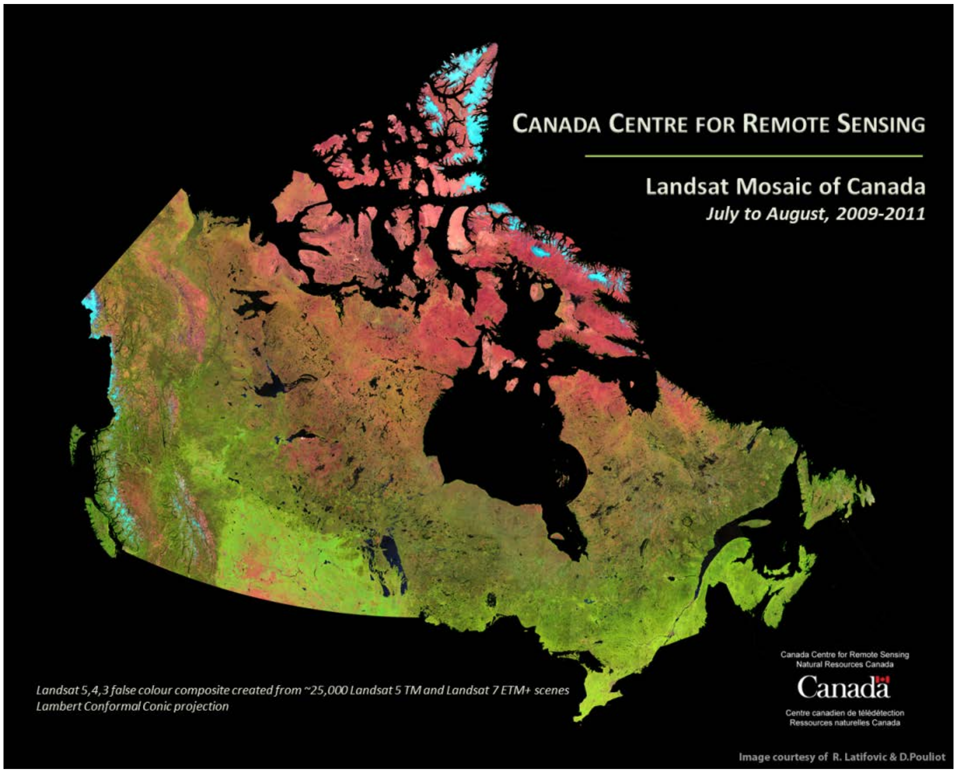

Figure 9. Landsat Mosaic of Canada false color composite TOA reflectance. **R**ed: band-5 (1.55-1.75 µm), **G**reen: band-4 (0.76-0.90 µm) **B**lue: band-3 3 (0.63-0.69 µm)

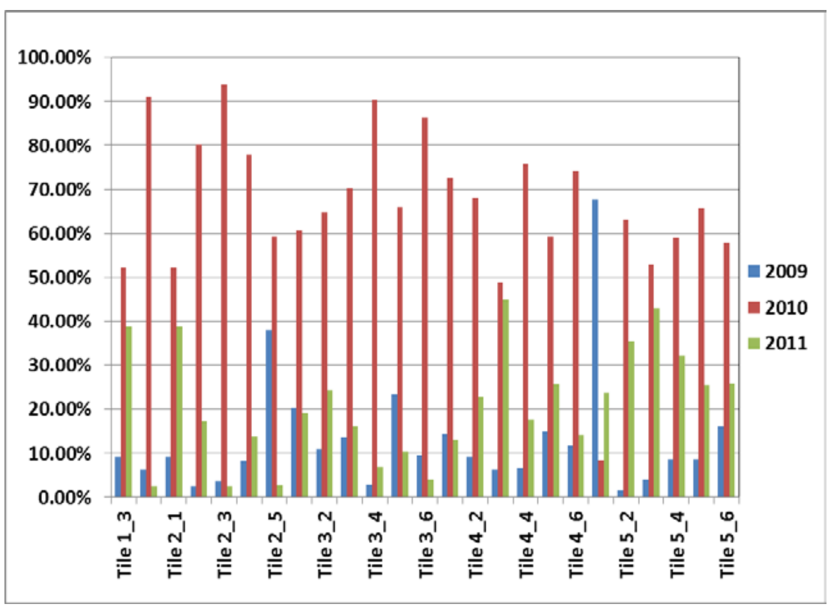

Figure 10. Pixels used in the new composite, by tile and acquisition year.

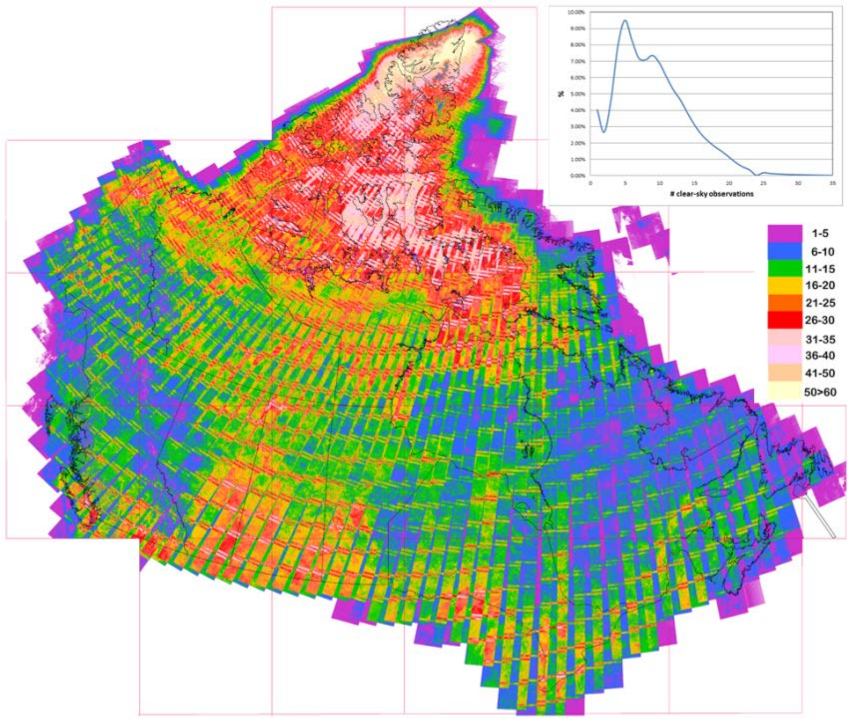

Figure 11. Available number of clear-sky observations per pixel

3.1 Product data layers

The tiles have been cutout into 1:1000000 scale National Topographic System Maps sheets for each band and saved in tiff file format Fig 12. At this scale, there are 100 map sheets covering Canada. The data are provided as top of atmosphere reflectance for the bands listed in Table 3. All bands are in unsigned 16-bit format.

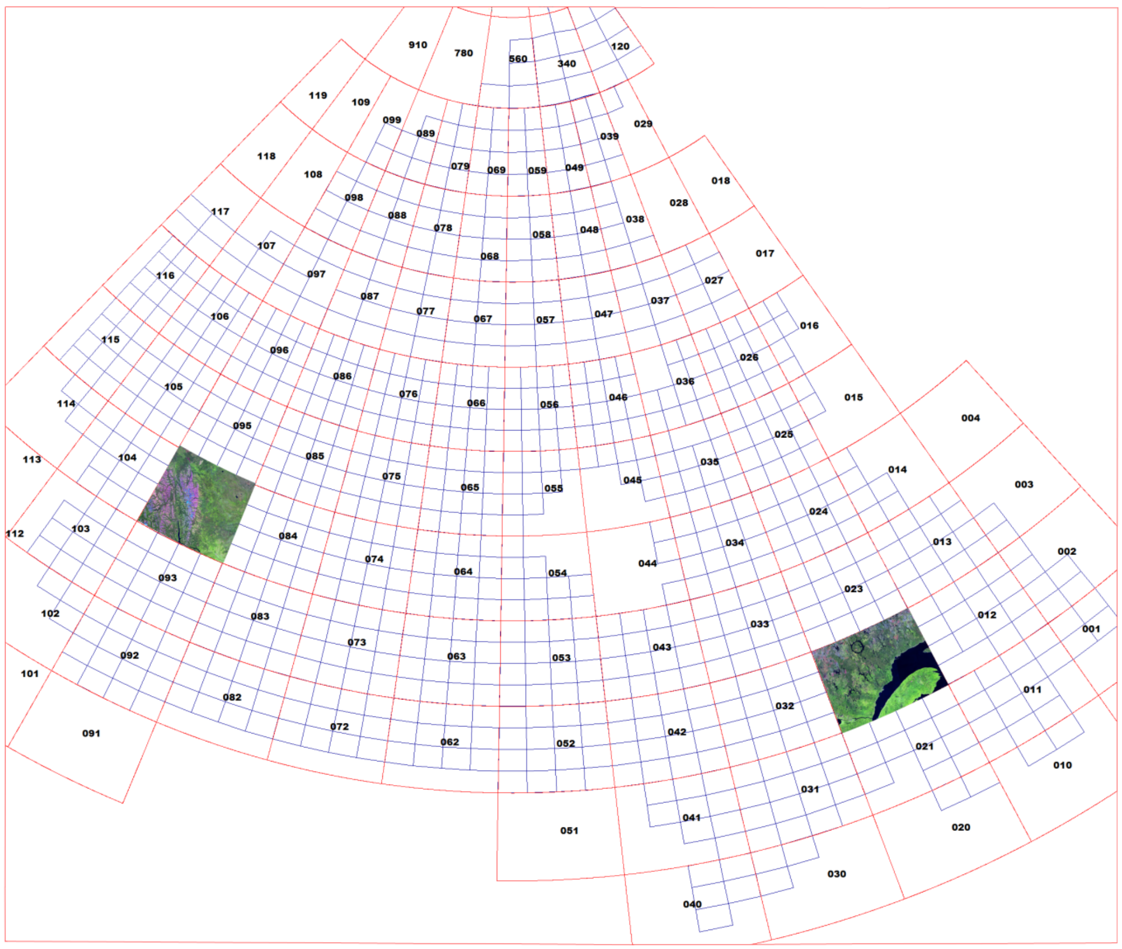

Figure 12. 1:1000000 scale National Topographic System Maps. Sheets 022 and 094 are provided as examples of this data product.

Table 3. Image Layer Descriptions

IMAGE LAYER	DESCRIPTION	DATA TYPE	FILL VALUE	VALID RANGE	SCALE FACTOR	OFF-SET	Unit
Mapsheet#_band 3	Band 3 (0.63-0.69 µm) reflectance at TOA	16-bit unsigned	65535	1~10000	0.0001	0	None
Mapsheet#_band 4	Band 4 (0.76-0.90 µm) reflectance at TOA	16-bit unsigned	65535	1~10000	0.0001	0	None
Mapsheet#_band 5	Band 5 (1.55-1.75 µm) reflectance at TOA	16-bit unsigned	65535	1~10000	0.0001	0	None
Mapsheet#_band 7	Band 7 (2.08-2.35 µm) reflectance at TOA	16-bit unsigned	65535	1~10000	0.0001	0	None
Mapsheet#_ind	Numeric code to identity scene in Logfile.	16-bit unsigned	65535	1~65535	1	0	None
Logfile.txt	Scene path\row and date information corresponding to the scene ID code in the scene ID layer.	text					

4. CONCLUSION

The work outlined here has contributed new methods and algorithms for extracting and analysing information from medium resolution optical data for improved land surface monitoring. The algorithms outlined here have been implemented in the Moderate Resolution Time Series Data Management and Analysis System – a system which enables CCRS to establish and manage its value-added Landsat-based LTSDR framework. The system and related applications facilitate processing, analysis and extraction of large volumes of data to generate information required for monitor different aspects of land surface dynamics.

The new Top of Atmosphere Reflectance Mosaic of Canada, circa 2010, at 30 m spatial resolution, which is described here, represents one of the outputs of a comprehensive body of research which has the goal of developing new methods, algorithms, value-added data sets and products to enable users to more effectively address monitoring and assessment of both anthropogenic and natural land surface changes.

To produce this particular LTSDR, 13,350 Landsat senses acquired during July-August 2009-2011 were ingested and processed following standards that preserve the sensor measured values of each pixel. The new Mosaic of Canada includes only a fraction of this Landsat data, selected to ensure full coverage with clear-sky observations. Enhancement of this data includes correction of the regions where scan-line striping artifacts were present. Quality control implemented as an integral part of the production process ensures quality of the product including geometric fidelity, cloud and cloud shadow screening, radiometric consistency and completeness of coverage.

By providing this national scale Landsat-based value-added data product with, “science quality”, readily available, CCRS is contributing to realizing the potential for new medium-resolution land surface information products. The new system and example data product which were described here are designed to assist government agencies, the scientific community, natural resources managers and non-governmental groups engaged in land cover mapping, and the generation of geophysical and biophysical products for the assessment of surface dynamics at national and regional scales. The new Mosaic of Canada will be made available via the National Earth Observation Data Framework (NEODF). Please contact neodf.mailbox@nrcan.gc.ca or Ltsdr@nrcan.gc.ca.

5. REFERENCES

C. E. Woodcock, R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. B. Cohen, et al., "Free access to Landsat imagery", *Science*, **2008**, vol. 320,1101.

C. Huang, S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann, "An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks," *Remote Sensing of Environment*, **2010**, 114 (1), pp 183-198.

D. Pflugmacher, W. B. Cohen, and R. E. Kennedy, Using Landsat-derived disturbance history (1972–2010) to predict current forest structure, *Remote Sensing of Environment*, **2012**, vol. 122, pp. 146-165.
G. Chander, B. L. Markham, and D. L. Helder, "Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors," *Remote Sensing of Environment*,**2009**, vol. 113, pp. 893-903.

G. Gutman, R. Byrnes, J. Masek, S. Covington, C. Justice, S. Franks, et al. Towards monitoring land-cover and land-use changes at a global scale: The Global Land Survey 2005. *Photogrammetric Engineering and Remote Sensing*, **2008**, 74, 6−10.

IEEE Transactions on Geoscience and Remote Sensing, **1994**, 32, pp. 427–437
J. Cihlar, D. Manak, M. D'Iorio. Evaluation of compositing algorithms for AVHRR data over land
J. Franklin, S.R. Phinn, C.E. Woodcock, J. Rogan. Rationale and conceptual framework for classification approaches to assess forest resources and properties. In: Wulder, M., Franklin, S.E. (Eds.), Methods and Applications for Remote Sensing of Forests: Concepts and Case Studies, Kluwer Academic Publishers, Dordrecht, **2003**, pp. 279-300

J.R. Jensen, D.J. Cowen. Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing*, **1999**, 65, pp.611–622.
R. E. Kennedy, Z. Yang, and W. B. Cohen. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms, *Remote Sensing of Environment*, **2010**,vol. 114, pp. 2897-2910.

R. Latifovic and D. Pouliot. Monitoring Cumulative Long-Term Vegetation Changes over the Athabasca Oil Sands Region. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, **2014**, vol.7, Number 7, pp. 3380-3392

S. Jin, L. Yang, P. Danielson, C. Homer, J. Fry, and G. Xian. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, **2011**,132, pp.159 – 175.

S. L. Powell, W. B. Cohen, S. P. Healey, R. E. Kennedy, G. G. Moisen, K. B. Pierce, et al., Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, **2010**, vol. 114, pp. 1053-1068.

W. B. Cohen, Z. Yang, and R. Kennedy, "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. "*Remote Sensing of Environment*, **2010**, vol. 114, pp. 2911-2924.

Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sensing of Environment*, **2012**,vol. 118, pp. 83-94.

Z. Zhu, C. E. Woodcock, and P. Olofsson. Continuous monitoring of forest disturbance using all available Landsat imagery, *Remote Sensing of Environment*, **2012**, vol. 122, pp.75-91.