Natural Resources Canada | Ressources naturelles Canada

GEOMATICS CANADA
OPEN FILE 52

Automated surface water extraction from RapidEye imagery including cloud and cloud shadow detection

I. Olthof

2019

Canada

GEOMATICS CANADA
OPEN FILE 52

Automated surface water extraction from RapidEye imagery including cloud and cloud shadow detection

I. Olthof

Canada Centre for Mapping and Earth Observation, 560 Rochester Street, Ottawa, Ontario K1S 5K2

2019

Permanent link: https://doi.org/10.4095/315176

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

Recommended citation
Olthof, I., 2019. Automated surface water extraction from RapidEye imagery including cloud and cloud shadow detection; Geomatics Canada, Open File 52, 20 p. https://doi.org/10.4095/315176

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

ABSTRACT

Mapped surface water extents represent fundamental geospatial information required by a large number of public and private sector stakeholders in Canada. Currently available surface water maps from Canada's National Hydrographic Network (NHN) are out-of-date in many locations in part due to the vintage of the maps themselves, and also because water extents are dynamic and changing. Previously, a significant amount of Canada's NHN base geospatial data was manually interpreted from airphotos and other sources of high-resolution imagery, which enabled detailed mapping but required significant human and financial resources. Satellite imagery can provide a cheaper alternative due to the free availability of certain data and potential for automated feature extraction. Freely available medium resolution data from sensors such as Landsat can provide information at a scale of 1:50 k while more detail is needed to map smaller water courses that can only be resolved from high resolution sensors (< 5 m) such as RapidEye. This Open File describes a robust, fully automated procedure to extract surface water from RapidEye imagery including cloud and cloud shadow screening, as a potential method to enhance future NHN updating. The method is applied to RapidEye imagery residing in the Government of Canada's image archive over the St-John, Red and Richelieu Rivers. Qualitative assessment of the Richelieu product shows high overall agreement with water features in Google Earth imagery and NHN, with errors stemming from spectral confusion with dark soil and urban shadow. Canada's RapidEye archive currently covers approximately 30 % of provinces while its 5 m spatial resolution provides a good compromise between detail and data volume that can be processed on modern workstations. Similar spectral bands available in other higher resolution satellite sensors (< 2 m) such as WorldView suggest more detailed surface water extents may become available using methods developed here and adapted to these sensors.

TABLE OF CONTENTS

Page

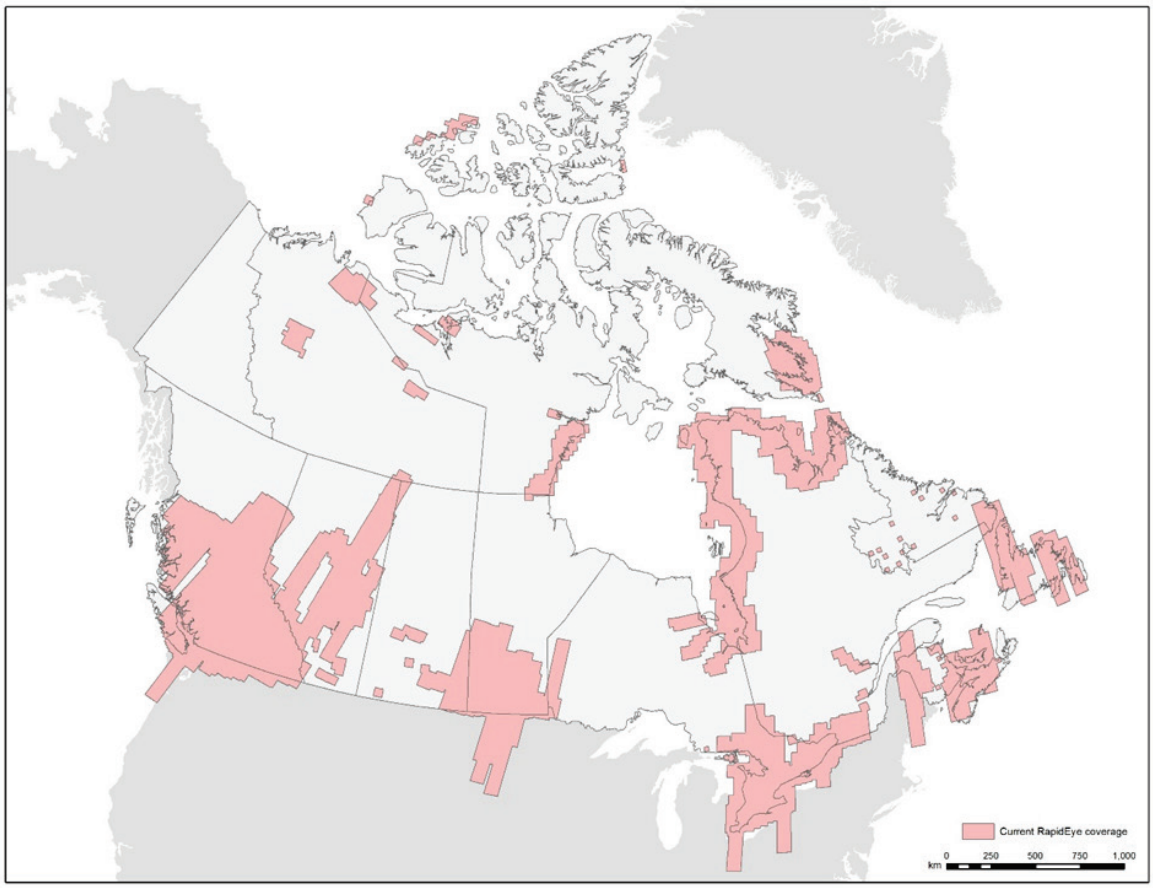
1. INTRODUCTION

Mapped surface water extents represent basic geospatial information required by many stakeholders in Canada for public safety, to inform land use, and to monitor ecosystem integrity among other needs. Surface water maps must be updated frequently due to water's seasonal dynamics as well as extents that may vary inter-annually because of climate change. Optical remote sensing data has been shown to be well suited for surface water mapping (Frazier and Page, 2000). Depending on the application, the required frequency and detail of surface water information may vary. For example, historical surface water extents with high revisit frequency can be used to help delineate floodplains, and can only be reliably obtained at moderate spatial resolution (10-30 m) due to data availability in image archives going back more than four decades. More recently, a number of fine resolution sensors (< 5 m) have been launched that can provide better detail to map watercourses and small water bodies. Regardless of application, automated surface water extraction from EO data is desirable due to large data volumes contained in the spatial domain of fine resolution imagery, and in the temporal domain when using archived image data going back several decades.

The Emergency Geomatics Services (EGS) at the Canada Centre for Mapping and Earth Observation, Natural Resources Canada maps surface water during flood events to provide situational awareness to the public and emergency responders. Automated surface water extraction has been developed by EGS for a number of moderate resolution optical and radar sensors using historical inundation frequency and machine learning (Olthof, 2017; Olthof and Tolszczuk-Leclerc, 2018). The EGS method uses historical inundation frequency maps to sample scene-specific signatures representing permanent land where historical inundation frequency is 0 %, and permanent water where inundation frequency is 100 %. Once sampled, scene-specific signatures representing permanent land and water are input into machine learning to classify the scene and map surface water extents. Because the method relies on signatures sampled from the same scene used for classification, it is sensor-independent and has been shown to perform well on a number of optical and radar satellite sensor data including Landsat, Sentinel 1 and 2, RadarSat-2 and TerraSAR-X. The medium resolution data mentioned, which include SWIR optical bands and radar wavelengths that are most reliable for surface water extraction (Frazier and Page, 2000; Brisco et al., 2009) have spatial resolutions between 12.5 and 30 m. This resolution is generally sufficient for the purpose for which the automated surface water extraction method was developed; to provide emergency situational awareness during a flood. However, surface water base mapping is also needed in Canada to update national hydrographic base maps at a finer resolution than the 12.5-30 m provided by these sensors.

RapidEye is a constellation of five satellites that acquires data in five spectral bands at a 5 m spatial resolution and 16-bit pixel depth. In addition to visible red (630-685 nm), green (520-590 nm) and blue (440-510 nm) as well as a red edge (690-730 nm) band, it also includes a NIR (760-850 nm) band that is particularly useful for surface water mapping. In these longer wavelengths, a high contrast exists between land and water while they are also relatively transparent to atmospheric effects and minimally affected by suspended sediment (Olthof et al., 2015). An added benefit of using this sensor for surface water mapping in Canada is the availability of purchased data currently covering more than 30 % of Canada's provinces from 2008-present to the Government of Canada through the Earth Observation Data Management

System (EODMS) (Figure 1). The five-metre resolution of RapidEye is also a good compromise between resolution, coverage and data volume since many water bodies missed at moderate resolution will be captured, while data volume is still manageable using current hardware and processing over imaged regions of Canada.



	Area (km^2)	RapidEye Coverage (km^2)	Missing Area (km^2)	Percent Uncovered (%)
Provinces	5,944,047	1,883,334	4,060,713	68.32
Canada	9,843,345	2,054,164	7,789,181	79.13

Figure 1. Rapideye coverage of Canada contained in the Earth Observation Data Management System (EODMS) catalogue as of March 2018.

Because RapidEye is an optical sensor and a fully automated method is desirable due to the large data volume needed to map Canada, an automated cloud and cloud shadow detection routine was required to mask no data areas prior to water extraction. A cloud detection algorithm called fmask has been developed for Landsat that uses all available channels including thermal (Zhu et

al., 2012). Since RapidEye does not include a thermal channel, individual bands as well as the Haze Optimized Transform (HOT) algorithm (Zhang et al., 2002) that uses blue and red bands were evaluated for separability between cloud, cloud shadow and land prior to development of a new cloud screening approach for RapidEye. The HOT algorithm relies on a high correlation between blue and red reflectance under clear-sky conditions represented by a clear-sky line in a scatter plot between the two bands. Atmospheric contamination will migrate pixels away from the clear-sky line in the direction of higher blue reflectance. Thus, residuals between predicted blue reflectance based on a regression between red and blue under clear sky, and measured blue reflectance can be used to produce a HOT image depicting different levels of haze from transparent to opaque cloud.

2. METHODS

2.1. Study Areas

Surface water maps were generated from archived RapidEye imagery over three separate study regions in Canada where other water studies have been conducted, including parts of the St-John NB River basin, Red River MB and Richelieu QC River basins. The three regions represent different geographical settings, with St-John being primarily forested with some mixed land use including agriculture and urban, the Red River traversing the prairies being predominately surrounded by agriculture and the Richelieu being the most diverse including the large urban centre of Montreal. Tests were conducted on several images acquired over all three regions to ensure the developed methods are robust to location and imaging conditions. All three regions experience frequent springtime flooding and are monitored almost annually by the EGS using Canada's RadarSat-2 satellite. Because of frequent water mapping in these areas, an extensive RapidEye image archive covering both regions has also been purchased and maintained through EODMS.

2.2. Imagery

The RapidEye dataset over the Richelieu includes four sensor level scenes, each containing all five spectral bands. Individual dates include years 2012 and 2013 and anniversary dates from September 3 to September 29, with multiple scenes covering certain areas. All images are more than 24000 pixels (120 km) by 34000 lines (170 km) including surrounding no-data areas, with the largest scene having data covering an area of 13979 km^2.

The St-John NB RapidEye dataset includes six sensor level scenes with five spectral bands and twelve dates from years 2012 and 2013 and anniversary dates from July 23 to September 28. Images range in size between 21000-27000 pixels (105-135 km) by 28000-56000 lines (140-280 km) with the largest covering an area of 22397 km^2 with data.

The Red River dataset contains twelve sensor level scenes with five spectral bands from years 2013, 2016 and 2017 and anniversary dates from July 13 to September 3. Images are between 17000-23000 pixels (85-115 km) by 13000-31000 lines (65-155 km) with the largest covering an area of 11321 km^2 with data.

2.3. Processing chain overview

PCI Geomatica was used to orthorectify sensor scenes using the rational polynomial functions provided with each scene, Canadian Digital Elevation Data (CDED) DEMs and nearest neighbor resampling to preserve original radiometry.

The cloud masking and classification procedures are otherwise implemented in R statistics. Image metadata from each scene is first read to determine if cloud is present, in which case cloud and cloud shadow screening is performed prior to surface water extraction. All spectral bands are used in the full procedure, with bands and their uses listed in Table 1.

Table 1. RapidEye bands and their application for cloud / cloud shadow screening and water extraction

Band	Use
440 - 510 nm (Blue)	Haze Optimized Transform (HOT) / Water Classification
520 - 590 nm (Green)	Water Classification
630 - 685 nm (Red)	Haze Optimized Transform (HOT) / Water Classification
690 - 730 nm (Red Edge)	Final Cloud and Cloud Shadow Detection / Water Classification
760 - 850 nm (Near IR)	Water Classification / Water Region Growing

2.4. Cloud and cloud shadow

A single RapidEye image was used to determine the spectral band that best separates cloud and cloud shadow from land, acquired on September 28, 2013 over the Richelieu region with clouds present. Masks representing cloud, cloud shadow, land and water were manually digitized using this image as a backdrop. Data were extracted beneath each class mask and spectral band as well as a Haze Optimized Transform (HOT) image computed from the blue and red bands. The two-sample Kolmogorov–Smirnov (KS) distance was computed between cloud and land, and separately between cloud shadow and land to determine the channel with the highest separability between classes in order to mask cloud and shadow. The K-S test is a nonparametric test that can be used to compare two samples and represents the distance between distributions ranging from zero to one with one corresponding to the maximum distance as a measure of separability. Distance measures have been grouped into four classes in other studies (Mohammadimanesh et al., 2019) representing poor separability (K-S < 0.5), some separability ($0.5 \leq$ K-S ≤ 0.7), good separability ($0.7 <$ K-S ≤ 0.85) and very good separability (K-S > 0.85).

The HOT image was generated by regressing the blue band against the red band using Theil-Sen robust regression that is insensitive to outliers. The regression represents the relation between the two bands in clear-sky areas as long as clouds that are outliers to the regression represent less than 29 % of the data (Kendall and Stuart, 1967). HOT coefficients are represented by regression slopes and offsets calculated from each scene. Slope and offsets are applied to predict blue reflectance from the red band under clear-sky conditions, and HOT images represent the residuals (or difference) between the blue image and predicted blue. Higher HOT levels correspond to larger residuals representing more atmospheric contamination and at the highest HOT levels, opaque clouds (Figure 2).

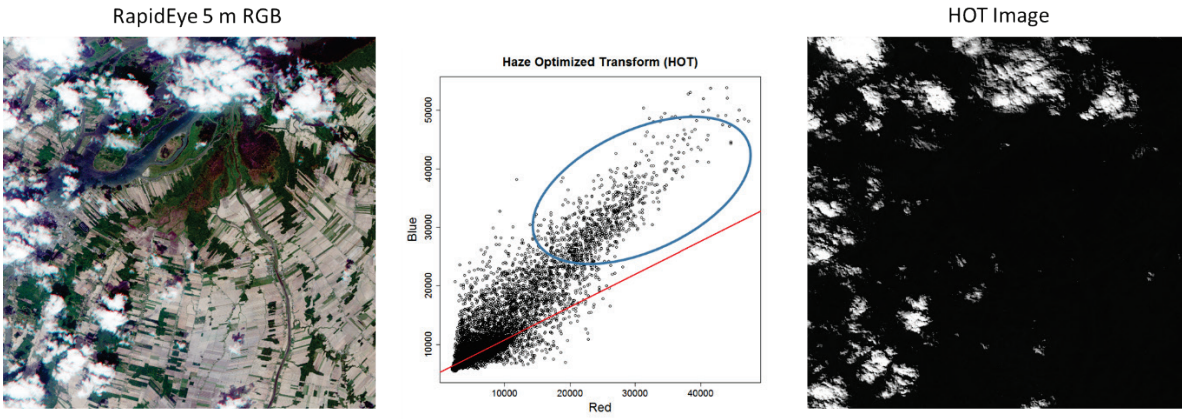


Figure 2. RapidEye image (left) displayed as R, G, B with the robust regression fit between red and blue bands as the red line in the scatterplot (middle) and haze and cloud contaminated pixels circled in blue. The corresponding HOT image representing regression residuals is on the right.

The channel with the highest separability between cloud, shadow and background land determined previously was thresholded to separate cloud from clear-sky. Upon inspection of image frequency histograms, it was noted that the brightness level at the maximum curvature of the histogram beyond peak frequency adequately separates cloud from clear-sky (Figure 3) with some omission along the edge of clouds. After applying this bright threshold value, a sieve filter is applied to the cloud sieved to remove small bright objects that have similar reflectance as clouds, including certain urban infrastructure built of materials such as concrete. A second sieving procedure can be applied optionally to remove bright agricultural fields and other manmade objects that are falsely detected as cloud. This procedure samples the border around the perimeter of each image object in turn and calculates the local slope and offset of the border at the sample location. Objects are removed from the cloud image if their border has more than 50 identical slope and offset values indicating at least one straight edge suggesting the object is manmade. Remaining large cloud objects are region grown to ensure entire clouds are masked including edges that were omitted in the initial thresholding. Region growing is performed iteratively by including adjacent values that are greater than 85 percent of the initial threshold, to a maximum of 1000 iterations or until no adjacent cloud pixels meet the brightness criterion. The threshold value for region growing was determined through trial and error and represents a

conservative value that grows clouds marginally while not including other bright features in the image in order to better match the shape and size of cloud shadows for subsequent matching.

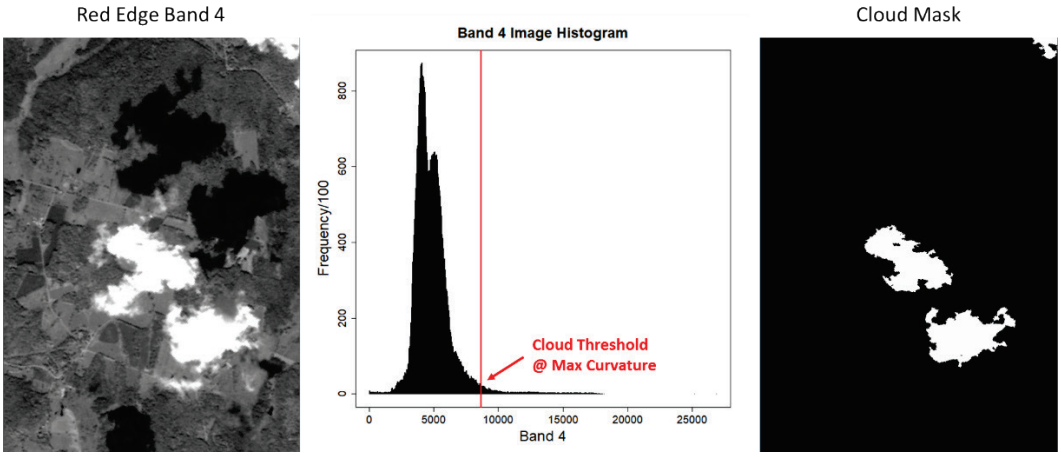


Figure 3. HOT image (left), frequency histogram of HOT levels with threshold used to separate cloud from clear-sky determined as the point of maximum curvature beyond the highest frequency HOT level (middle), and corresponding cloud mask after application of the threshold to the HOT image.

Cloud shadow is mapped next by determining an offset in x and y that best matches cloud with cloud shadow, and applying these offsets to shift the cloud mask to the location of corresponding shadow. The method assumes a fairly constant cloud height across the scene, and this assumption generally holds over the area covered by a scene as long as there are no weather fronts moving across the orbit image at the time of acquisition. Shadow cast from clouds located outside the image frame will be missed using this approach; however overlap among scenes provides the opportunity to identify and remove cloud shadow falsely identified as water. The band that has the darkest shadows and best contrast between shaded and illuminated land cover based on KS separability is selected to map cloud shadow.

To determine the offset between cloud and corresponding shadow, both the cloud mask and corresponding shadow spectral band are aggregated by a factor of ten to reduce computation time. After permanent water is masked using Pekel's occurrence layer, a crude shadow mask is obtained by applying a dark threshold calculated from the cloud shadow spectral band quantile representing the cloud fraction obtained from the cloud mask. The resampled cloud mask is then systematically displaced across the resampled shadow mask to determine a best match between cloud and shadow. A simple distance measure is used obtain a best match calculated as the number of cloud and shadow pixels that intersect at each offset iteration. The maximum number of intersecting pixels between the two images is used to obtain offset values in x and y representing a best match. Offset value are applied to the full resolution cloud mask to generate the final cloud shadow mask. The complete no data mask is obtained by combining

corresponding cloud and cloud shadow masks for each scene. Final masks are subsequently buffered using a 60 pixels radius circular kerned to ensure that edge pixels of clouds and shadow not captured by initial masks are included as no data (Figure 4).

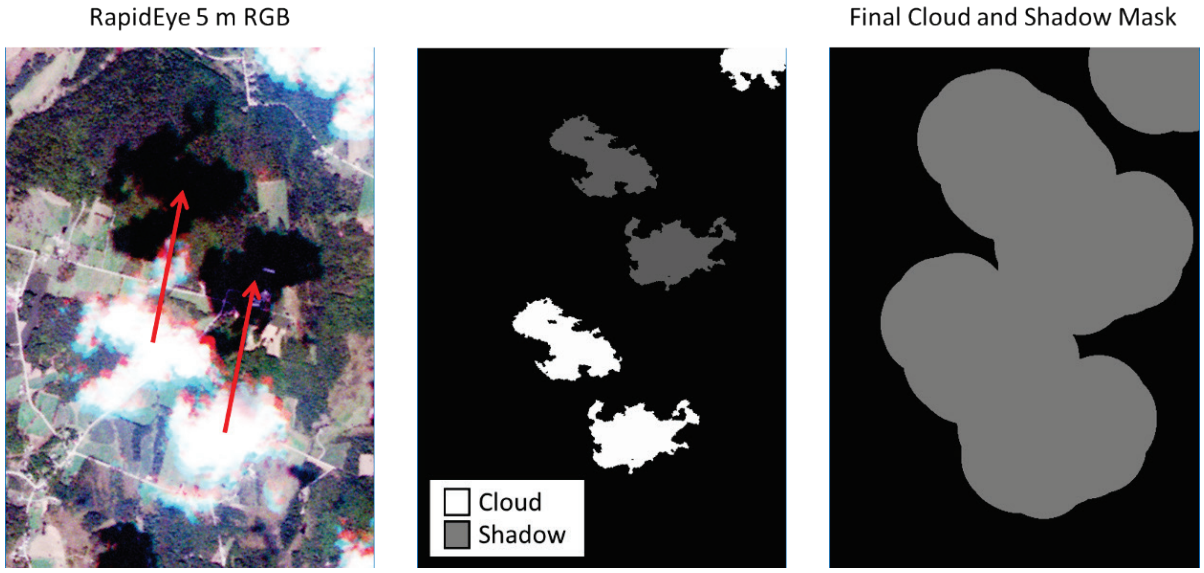


Figure 4. Cloud shadow mask obtained by determining the offset value (left) that best matches the cloud mask and corresponding cloud shadow in the red edge (band 4) image. The final cloud and shadow mask is obtained by combining cloud and shadow (middle) and then buffering (right).

2.5. Surface water extraction

Once cloud and cloud shadow masks are generated and set to no data in the each band, surface water is extracted using machine learning classification. Pekel's occurrence layer (Pekel et al., 2016) is used as a mask to sample RapidEye multispectral signatures representing permanent land and water. This layer was generated from historical Landsat imagery to map the percentage each location on earth was observed to be inundated, from permanent land with 0 % inundation frequency to permanent water with 100 % inundation.

Pekel's occurrence is first resampled to the extents and resolution of the RapidEye image, and then thresolded so that occurrence above 80 % represents permanent water, and occurrence at 0 % permanent land. The 80 % threshold is selected for permanent water because no occurrences of 100 % exist in the data due to some omission error in the historical water classifications, while 80 % was deemed acceptable based on visual inspection and other base water layers such as Canada's National Hydrographic Network (NHN). RapidEye mutipspectral data representing permanent land and water is extracted using stratified sampling with 25000 samples per stratum, and input as training to the C50 classification algorithm using the default parameters of 10 trials, a confidence factor of 0.25 and no winnowing. For water, the stratified sample is further reduced by removing the ten percent brightest samples in band 5 to avoid confusion between shallow or

turbid water that is relatively bright, and dark land features such as shadow and dark, fallow fields. Machine learning has been shown to be extremely robust to training error (Olthof and Tolszczuk-Leclerc, 2018), and therefore minor discrepancies between RapidEye surface water and Pekel's occurrence due to resolution differences or changes in water extents between Landsat and RapidEye dates have a negligible effect on classification performance.

Once the initial classification is generated, a refinement process is implemented to improve the results. Region growing is performed using band 5 to merge pixels adjacent to water obtained previously if their digital number is less than a threshold value obtained by overlaying the initial C50 classification on band 5 and calculating three times the mean brightness of water. This value was obtained by examining multiple scenes with shallow, turbid or haze obscured water that is brighter than the majority of water in the scene. Region growing is completed after a maximum of one thousand iterations or until no adjacent pixels meet the threshold criterion (Figure 5).

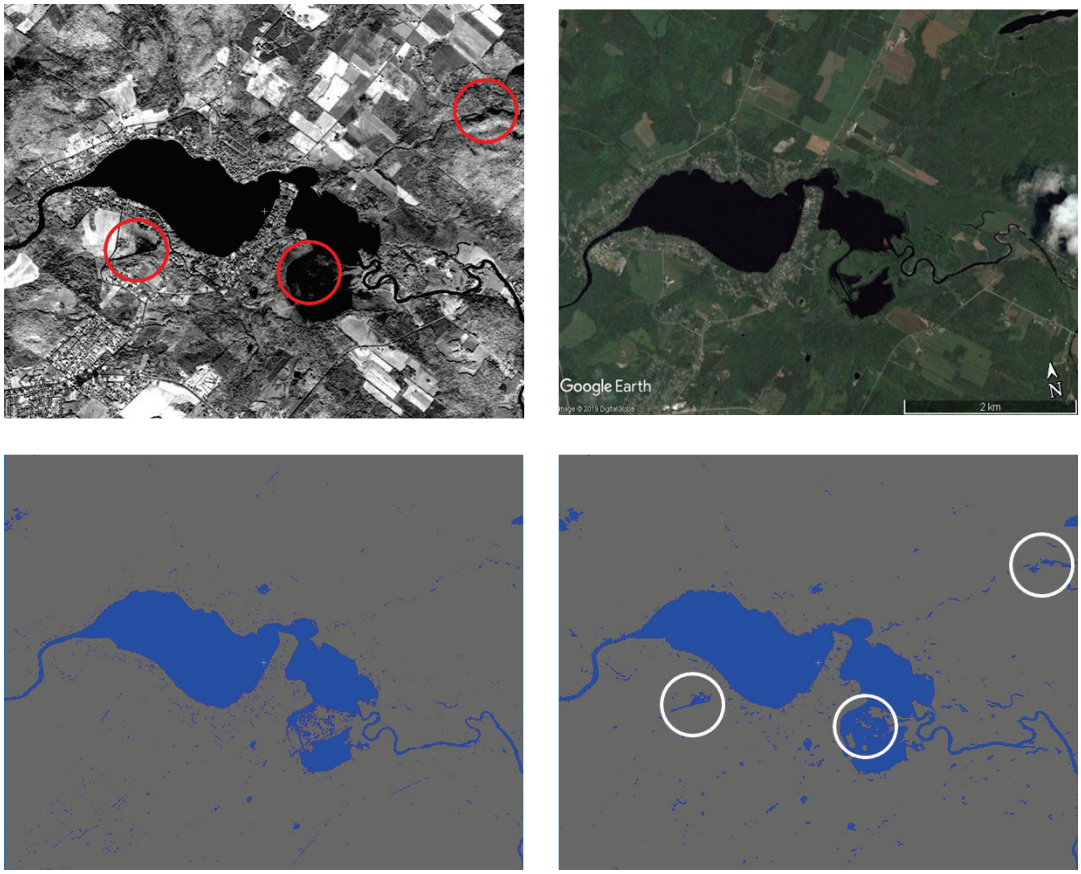


Figure 5. NIR (band 5) image (top left), Google Earth image (top right), initial water classification (bottom left) and final classification after region growing with major infilled areas circled in white (bottom right).

Multiple scenes acquired on different dates may cover the same area and surface water extents are captured for each date. A final product is generated by stacking images acquired on different dates over common areas and calculating the percentage inundation for each pixel in the stack. The final products therefore represents inundation frequency calculated as the percentage each pixel location is covered by water in the input images (Figure 6).

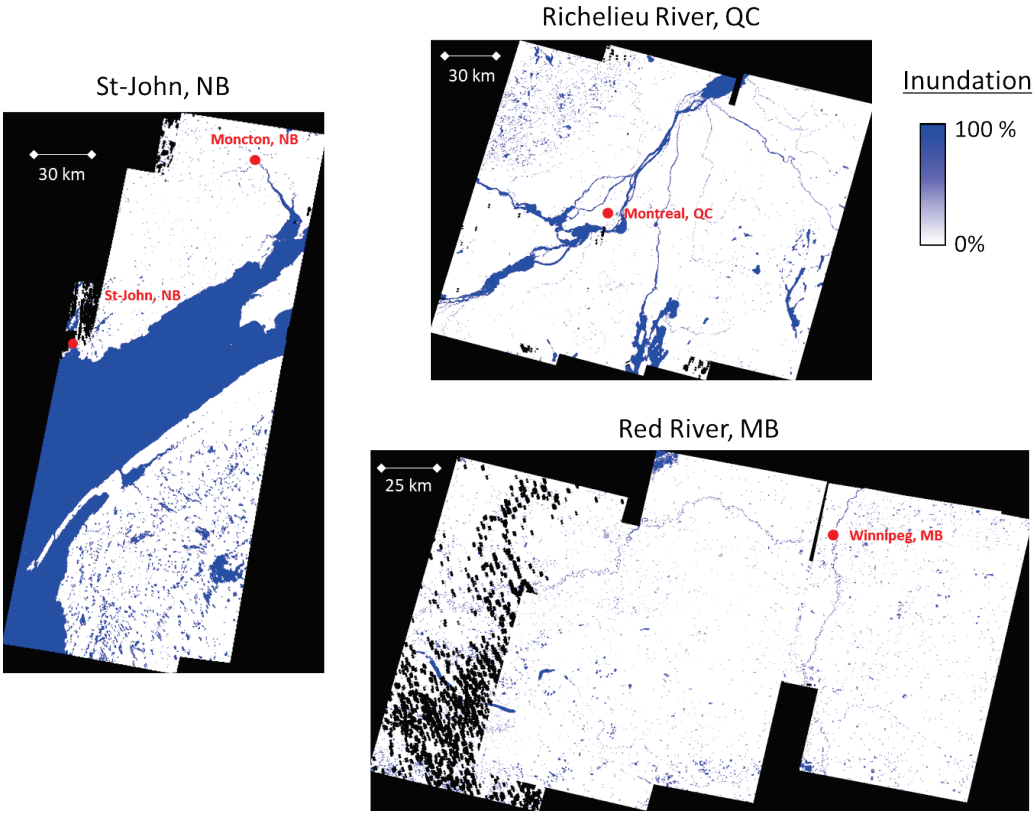


Figure 6. Years 2012-2013 inundation frequency of St-John, NB and Richelieu, QC Rivers, and years 2013-2017 inundation frequency of the Red, MB River from multi-temporal RapidEye imagery.

3. RESULTS AND DISCUSSION

K-S test statistics between land and cloud, and land and shadow are presented for channels representing each band plus the HOT image in Table 2, with a statistic closer to one representing higher separability. In the case of land versus cloud, all channels exhibited very good separability (K-S > 0.85), with the HOT image being marginally lower than individual bands. The band with the highest separability between land and cloud was the red edge band 4. In all bands, confusion between cloud and land occurred in urban and other similarly bright barren areas such as bare fields, quarries and logging roads.

Table 2. Two-sample Kolmogorov–Smirnov (KS) separability distances between land, cloud and shadow for five RapidEye spectral bands and Haze Optimized Transform (HOT) image

	band1	band2	band3	band4	band5	HOT
land - cloud	0.994	0.994	0.997	0.998	0.995	0.988
land - shadow	0.658	0.867	0.857	0.957	0.932	0.388

For cloud shadow processing, the initial shadow mask is generated by thresholding at the brightness distribution percentile represented by the percent cloud cover in the scene. More overlap between cloud shadow and land will produce a poorer initial cloud mask, thereby reducing the ability to produce accurate offsets that match clouds with their corresponding cloud shadows. Of all the channels, only band 1 and the HOT image exhibit less than very good separability (K-S < 0.85) between land and shadow, with band 4 being the channel with the highest discriminatory power. Therefore, band 4 was used to produce both the cloud mask and offsets to shift the cloud mask to the location of corresponding cloud shadows.

No formal accuracy assessment has been performed on any of the products to date. Even at moderate spatial resolution, optical imagery is routinely able to achieve greater than 90 % water classification accuracy. However, overall accuracy of a binary classification such as land vs water may not reflect the overall product quality. Challenges in classifying water normally occur near the margins between water and shoreline whereas under random sampling, these areas are rarely sampled. Such areas include shallow or tubid water, periodically inundated or wetted land and shaded beaches and shorelines from nearby vegetation; all conditions that cause spectral confusion with the opposite class.

In landscapes with large, regularly shaped water bodies, a majority of randomly sampled reference points will fall well within water bodies or on land in areas that are most often classified correctly due to high spectral separability between land and water. As water bodies become narrower and more irregularly shaped, the amount of shoreline increases at a greater rate than the surface water area, and therefore more randomly sampled points will tend to occur closer to shorelines in areas that are more difficult to classify correctly. Therefore, classification accuracy not only reflects product quality, but also the configuration of water bodies including size and shape. An alternative sampling scheme such as purposive sampling near shoreline margins may produce a more meaningful assessment of product quality; however, the final accuracy assessed in this manner still depends on the configuration of water bodies, making comparisons with products from other areas difficult.

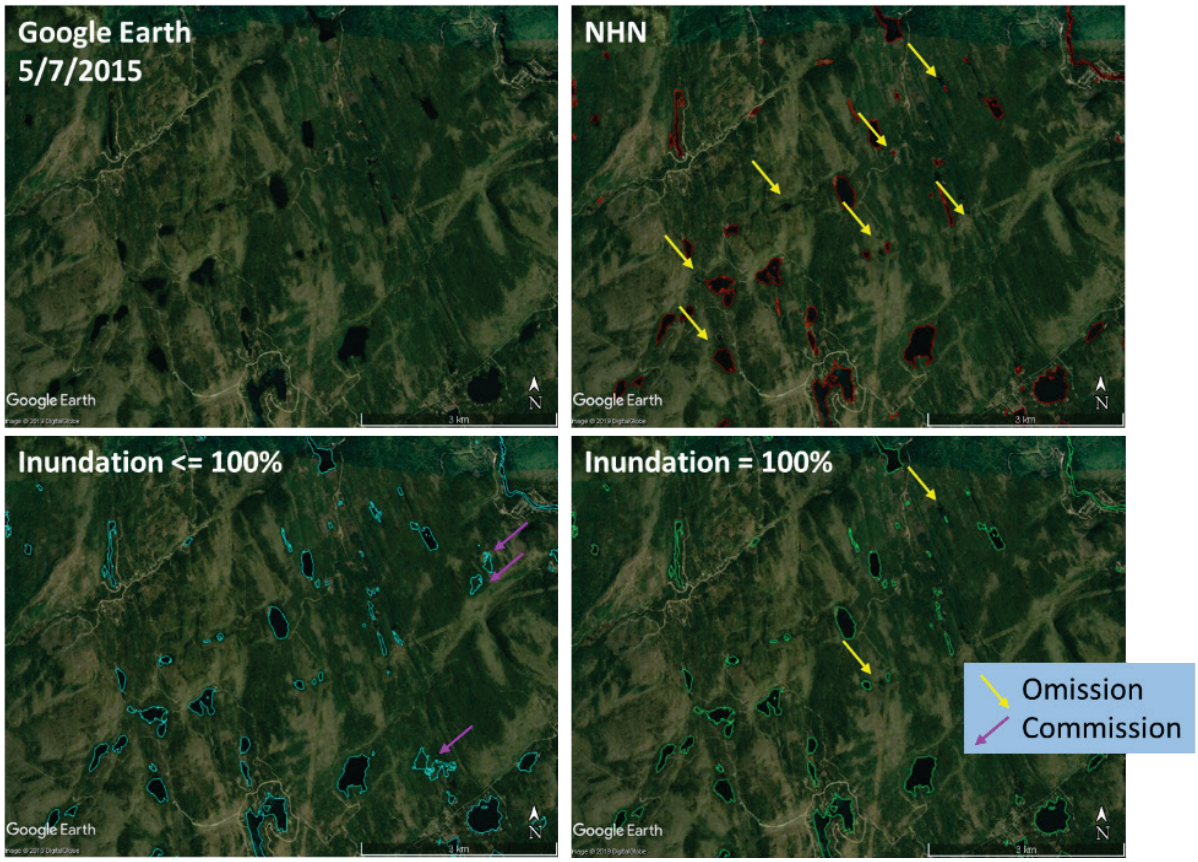


Figure 7. Example of qualitative assessment of a subset of the RapidEye September 28 and 29, 2013 inundation frequency product compared to the National Hydrographic Network generated between 1998 and 2003, and 2015 Google Earth imagery. Omission refers to lakes observed in Google Earth that were not mapped, while commission refers to mapped lakes not present in Google Earth.

To assess the quality of the Richelieu product in order to identify sources of error, water detected in the map was visually compared to features detected in Google Earth (Figure 7). Also, because one of the objectives of this work is to assist in NHN updating, a visual comparison with existing NHN data was performed to determine where NHN performs poorly or is out-of-date. Comparisons were made between Google Earth imagery acquired July 5, 2015, water extracted from two RapidEye scenes acquired September 28 and 29, 2013 and NHN generated between 1998 and 2003. Water bodies that were present in both overlapping scenes (100 % inundation) were considered, and water bodies present in either scene (<= 100% inundation) were also evaluated separately. While the evaluation was conducted over a large area, conclusions can be drawn looking at a relatively small area shown in Figure 7 that is located northwest of Lake St-Pierre. In this region, at least seven small lakes were seen to be missing from NHN as confirmed by Google Earth. All seven of these lakes were mapped at least once by both RapidEye scenes, while two were absent in one scene or the other (inundation < 100%). Although all lakes present in this area were mapped in at least one of two RapidEye scenes, false lake detections also

existed when combining lakes from both scenes in part due to residual cloud shadow not removed by the cloud / shadow masking procedure. It should be noted that the time difference between all sources of lake information may lead to errors in the assessment due to changes that may have taken place between dates. However, due to the location of the assessment area in the Canadian Shield and the relatively short time difference particularly between Google Earth and RapidEye, entire lake appearance or disappearance with no evidence of a transition should be rare to non-existent.

In agricultural regions, some fallow fields that appeared dark due to high moisture and / or organic matter in Google Earth were falsely classified as water in the map (Figure 8). This confusion appeared in a small percentage of agricultural areas as well as in at least one area that appeared to be a sod farm. While these areas do not represent permanent water bodies, this does not imply that they were dry when imaged, as standing water may persist after a rain event on fine textured soils before percolation or evaporation. The vast majority of these fields were not permanent based on a limited temporal stack of images; in other words their inundation frequency was less than 100 %.

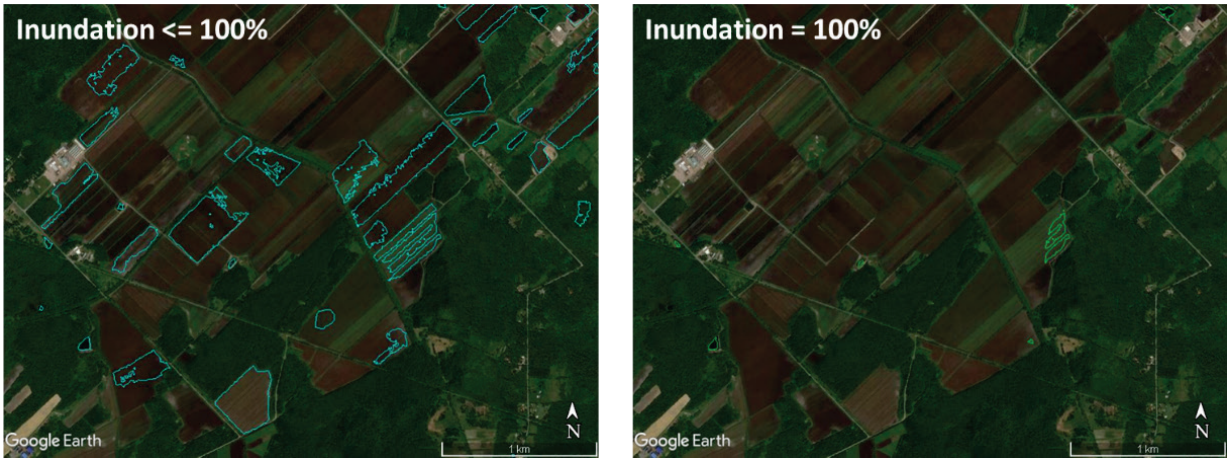


Figure 8. Example of potential commission error in the map due to dark soil with high moisture and / or organic matter. Inundation frequency suggests that if water was present in either RapidEye scene covering this region, it was not permanent.

A second significant source of error was in dense urban areas particularly in downtown Montreal. This confusion is understandable given the extent and depth of shadow cast by tall buildings. Both sources of confusion due to dark fields and shadow can be minimized or eliminated with high-resolution ancillary agriculture and urban masks. Correctly classified features include quarries, small ponds, water reservoirs and wetlands. Some of these features are on the order of 1-2 pixels across and are difficult to detect in natural colour imagery because of aquatic vegetation but were verified in Google Earth upon careful inspection. Many wetland types are represented by areas with ephemeral water; however not all ephemeral water is wetland, for example standing water in agricultural fields. In combination with other optical or

radar data, inundation frequency may help to interpret and classify wetlands. While both areas include a limited time-series of images in any one location up to a maximum of six over Richelieu, water dynamic can nonetheless be informative in some areas. For example, multiple observations in some locations along the Richelieu River portray different water levels and therefore different water extents due to overland flooding.

The approach developed here was designed to be as robust as possible to different surface and atmospheric conditions such as haze. The methodology initially classifies deep, dark water and avoids classifying shallow or turbid water, or water obstructed by transparent haze in order to minimize commission error with other dark targets such as those described above. Once seeded by classification, dark water is extended using region growing into adjacent shallow, turbid and haze-obstructed water. By growing into adjacent areas only, commission error due to shadow in urban areas as well as other dark targets is minimized as long as these targets are not immediately adjacent to water. Tuning certain classification parameters is necessary to ensure that at least one water pixel exists in each water body as the seed for region growing to their full extents and avoid omitting that water body altogether. Conversely, if the initial classification prior to region growing captures too much water, then invariably some commission error will exist.

The methodology is not intended to fully replace visual interpretation of remote sensing imagery for base surface water mapping. The spatial resolution is too coarse to detect the smallest water bodies, but should be sufficient at the 1:50 k scale that most NHN tiles are currently published at. Inspection of NHN data in in the Richelieu region suggests much of it is out of date, particularly in mapping small water bodies that have high errors of omission and commission in some locations. Products generated from RapidEye appear to have small omission errors, but some commission due to errors stated above. Products generated from RapidEye can be used as a start to manually update NHN, whereby an analyst can review and remove falsely detected water bodies manually in urban and agriculture regions.

Sensors with spectral bands similar to RapidEye and higher spatial resolution provide opportunities to generate more detailed surface water extents in the future by adapting current methods to these sensors. For example, data from the WorldView constellation already purchased by the Government of Canada is available at < 2 m spatial resolution and includes bands with similar spectral ranges to RapidEye's five bands, in addition to two extra bands in the yellow and NIR ranges for WordView 1 and 2. WorldView 3 include the same bands as WorldView 1 and 2 at a higher spatial resolution, plus eight SWIR bands at a slightly coarser resolution that should enhance surface water extraction greatly (Table 3).

Table 3. WorldView sensors with spectral bands similar to RapidEye imaged at a higher than RapidEye's 5 m spatial resolution

RapidEye			WorldView-2			WorldView-3		
Band	Wavelength (nm)	Resolution	Band	Wavelength (nm)	Resolution	Band	Wavelength (nm)	Resolution
			Coastal	400-450	1.84 m	Coastal	400-450	1.24 m
Blue	440-510	5 m	Blue	450-510	1.84 m	Blue	450-510	1.24 m
Green	520-590	5 m	Green	510-580	1.84 m	Green	510-580	1.24 m
			Yellow	585-625	1.84 m	Yellow	585-625	1.24 m
Red	630-685	5 m	Red	630-690	1.84 m	Red	630-690	1.24 m
Red Edge	690-730	5 m	Red Edge	705-745	1.84 m	Red Edge	705-745	1.24 m
NIR	760-850	5 m	NIR1	770-895	1.84 m	NIR1	770-895	1.24 m
			NIR2	860-1040	1.84 m	NIR2	860-1040	1.24 m
						SWIR-1	1195 - 1225	3.7 m
						SWIR-2	1550 - 1590	3.7 m
						SWIR-3	1640 - 1680	3.7 m
						SWIR-4	1710 - 1750	3.7 m
						SWIR-5	2145 - 2185	3.7 m
						SWIR-6	2185 - 2225	3.7 m
						SWIR-7	2235 - 2285	3.7 m
						SWIR-8	2295 - 2365	3.7 m

4. CONCLUSIONS

A fully automated surface water extraction methodology has been developed for RapidEye imagery to potentially assist ongoing NHN updating in Canada. The method has shown itself to be reasonable efficient and can be deployed on high-end workstations to exploit the RapidEye image archive already purchased by the Government of Canada covering ~30 % of provinces. Qualitative assessment of a sample product generated over the Richelieu River showed sources of map error including confusion with fallow fields with dark organic soils and high moisture content as well as shadow in urban areas. Both sources of confusion can be minimized with the application of ancillary data such as high resolution agriculture and urban masks. While RapidEye's 5 m spatial resolution currently provides a good compromise between higher spatial detail available in sub-metre resolution sensors and medium resolution data volume, similar spectral bands available in higher resolution data (e.g. < 2 m from WorldView) suggests more detailed information may become available through the adaptation of current methods.

5. REFERENCES

Brisco, B.; Short, N.; van der Sanden, J.; Landry, R.; Raymond, D. A semi-automated tool for surface water mapping with Radarsat-1. Can. J. Remote Sens. **2009**, 35, 336–344.

Dabboor, M.; Montpetit, B.; Howell, S. Assessment of the High Resolution SAR Mode of the RADARSAT Constellation Mission for First Year Ice and Multiyear Ice Characterization. Remote Sensing **2018**, 10, 594.

Espeseth, M.M.; Brekke, C.; Johansson, A.M. Assessment of RISAT-1 and Radarsat-2 for sea ice observations from a hybrid-polarity perspective. Remote Sensing **2017**, 9, 1088.

Frazier, P. S.; Page, K. J. Water body detection and delineation with Landsat TM data. Photogrammetric Engineering and Remote Sensing, **2000**, 66, 1461–1467.

Ghosh, A.; Manwani, N.; Sastry, P.S. On the Robustness of Decision Tree Learning under Label Noise. arXiv, **2016**, arXiv:1605.06296.

Geldsetzer, T.; Arkett, M.; Zagon, T.; Charbonneau, F.; Yackel, J.J.; Scharien, R.K. All-season compact-polarimetry C-band SAR observations of sea ice. Canadian Journal of Remote Sensing. **2015**, 41, 485–504.

Kendall, M. G.; Stuart, A. S. Advanced Theory of Statistics, vol. 2. London7 Charles Griffin and Company, **1967**.

Massey Jr, F.J. The Kolmogorov-Smirnov test for goodness of fit. Journal of the American statistical Association **1951**, 46, 68–78.

Mohammadimanesh, F.; Salehi, B.; Mahdianpari, M.; Brisco, B.; Gill, E. Full and simulated polarimetry SAR responses to Canadian wetlands : separability analysis and classification. Remote Sensing. **2019**, 11, 516. doi:10.3390/rs11050516.

Olthof, I.; Fraser, R.H.; Schmitt,C. Landsat-based mapping of thermokarst lake dynamics on the Tuktoyaktuk Coastal Plain, Northwest Territories, Canada since 1985. Remote Sensing of Environment. **2015**, 168, 194-204.

Olthof, I. Mapping seasonal inundation frequency (1985-2016) along the St-John River, New Brunswick, Canada using the Landsat archive. Remote Sensing. **2017**, 9, 143. doi:10.3390/rs9020143.

Olthof, I.; Tolszczuk-Leclerc, S. Comparing Landsat and RADARSAT for current and historical dynamic flood mapping. Remote Sensing. **2018**, 10, 780. doi:10.3390/rs10050780.

Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A. S. High-resolution mapping of global surface water and its long-term changes. Nature**. 2016**, 540, 418–422. doi:10.1038/nature20584.

Zhang Y.; Guindon B.; Cihlar J. An image transform to change characterize and compensate for spatial variability in thin cloud contamination of Landsat images. Remote Sensing of Environment. **2002**, 82, 173-187. doi:10.1016/S0034-4257(02)00034-2.

Zhu, Z.; Woodcock, C. E. Object-Based Cloud and Cloud Shadow Detection in Landsat Imagery. Remote Sensing of Environment. **2012**, 118, 83-94, doi:10.1016/j.rse.2011.10.028 (paper for Fmask version 1.6).