INVESTIGATION OF MODIS SNOW MAPPING IN A BOREAL FOREST WATERSHED

Keith N. Abbott, Technician
Amir Ali Khan, Manager
Department of Environment & Conservation
Water Resources Management Division
4th Floor, Confederation Building, West Block
PO Box 8700, St, John's, NL, A1B 4J6
keithnabbott@gov.nl.ca

ABSTRACT

The Water Resources Management Division (WMRD), of the Newfoundland and Labrador Department of Environment and Conservation, is responsible for forecasting flood events for one of its largest provincial river systems, the Humber River. Snow cover maps derived from MODIS satellite imagery have proven to be beneficial for measuring percent snow cover within the Humber River Watershed and have improved forecasting river flow rates during springtime snow melt. This paper compares a valuable and freely-distributed snow cover map, provided by the National Snow and Ice Data Center (i.e. MOD10A1), with a WRMD in-house variation of this product. Results based solely on visual comparisons of these snow products with corresponding MODIS composite images show that snow is often misclassified as cloud in the MOD10A1 product. The more labor intensive method of creating the WRMD product excludes the use of a cloud mask and was found to produce more reliable snow cover maps.

INTRODUCTION

During the 2008 snow season (November-June), the Newfoundland and Labrador Department of Environment and Conservation, Water Resources Management Division (WRMD) utilized snow cover maps, derived from MODIS satellite imagery, for forecasting flood risk within the Humber River Watershed. These maps were used in addition to a dynamic regression statistical model designed specifically to generate flow forecasts for the Humber River (Picco 1997).

Intensive flood forecasting for the Humber River is performed in springtime when melting snow causes river flow rates to reach flooding levels. The snow maps were used for measuring percent snow cover in the watershed, which was a useful indicator of the amount of water stored on the land in the form of snow.

The National Snow and Ice Data Center (NSIDC) provide a valuable service of distributing snow cover maps to the public, free-of-charge and on a global scale. The goal of this study was to compare the MOD10A1 product, distributed by the NSIDC, with a WRMD in-house variation of this product.

SNOW MAPPING ALGORITHMS

Table 1 provides a list of data types used as input to create the WRMD and MOD10A1 products. The MODIS Level 1B data type (MOD02HKM) is common to both products. It contains the reflectance data used to identify a pixel as snow covered. Three additional data types are unique to the MOD10A1 product: (i) MOD021KM, (ii) MOD03, and (iii) MOD35_L2. Riggs *et al.* (2006) provide details on the utility of these additional data types. Their general use is to exclude unwanted areas from the snow classification procedure (e.g. cloud cover, salt pans, and sandy beaches), as well as to help select pixels of higher quality (e.g. pixels that are close to nadir and acquired near the time of local solar noon). The additional data types were not integrated into the WRMD algorithm, since much of this work is done manually during the image selection process.

Table 1. Input data used for the NSIDC and WRMD snow products (adapted from Riggs *et al.* 2006).

Data type	Long name	Data used*	Snow products
MOD02HKM	MODIS Level 1B calibrated & geolocated radiances	Reflectance for MODIS bands: Band 1 (0.645 µm) - Red** Band 2 (0.865 µm) - NIR Band 4 (0.555 µm) - Green Band 6 (1.640 µm) - SWIR	NSIDC & WRMD
MOD021KM	MODIS Level 1B calibrated & geolocated radiances	Band 31 (11.28 µm) - TIR Band 32 (12.27 µm) - TIR	NSIDC
MOD03	MODIS geolocation	Land/water mask Solar zenith angles Sensor zenith angles Latitude Longitude	NSIDC
MOD35_L2	MODIS cloud mask	Cloud mask flag Unobstructed field of view flag Day/night flag	NSIDC

* NIR = Near InfraRed; SWIR = ShortWave InfraRed; TIR = Thermal InfraRed.
** The red reflectance band was not used as input for the WRMD product.

MOD10A1 Algorithm

The MOD10A1 product (Collection 5) was downloaded from the NSIDC website (http://nsidc.org). It contains four image datasets including the *snow cover daily tile* used here. This tile is comprised of an 8-bit image that categorizes spatial information into 11 classes, including cloud, lake ice, and snow. Each MOD10A1 tile is gridded in the sinusoidal projection, covers a ground area of approximately 1200km x 1200km, and is stored in Hierarchical Data Format (Riggs *et al.* 2006).

The process of creating the MOD10A1 product first involves culling through the MODIS data to select only pixels that sample land or inland water features, are in daylight, are unobstructed by clouds, and have a temperature less than 283K (Riggs *et al.* 2006). It is important to note that Top-of-the-Atmosphere (TOA) reflectance is used in the MOD10A1 algorithm. It can be computed from the MOD02HKM data type by first converting the scaled digital number of that data set to reflectance values using equation [1], with the scale and offset values extracted from the MOD02HKM metadata files:

$$\text{reflectance value} = \text{scale value} \times (\text{MOD02HKM digital number value - offset value}) \qquad [1]$$

The reflectance value is then converted to TOA reflectance using equation [2], with the solar zenith angle extracted from the MOD03 data type:

$$\text{TOA reflectance} = \frac{\text{reflectance value}}{\cos\text{ (solar zenith angle)}} \qquad [2]$$

The Normalized Difference Vegetation Index (NDVI) and Normalized Difference Snow Index (NDSI) are calculated using the TOA reflectance values as follows:

$$\text{NDVI} = \frac{\text{NIR (band 2) - Red (band 1)}}{\text{NIR (band 2) + Red (band 1)}} \qquad [3]$$

$$\text{NDSI} = \frac{\text{Green (band 4) - SWIR (band 6)}}{\text{Green (band 4) + SWIR (band 6)}} \qquad [4]$$

Three sets of conditions are then applied to determine if a pixel is mapped as snow cover. The first set will map a pixel as snow cover if conditions [5], [6] and [7] are all satisfied:

$$\text{NDSI} \geq 0.40 \qquad [5]$$

$$\text{NIR (band 2)} > 0.11 \qquad [6]$$

$$\text{Green (band 4)} > 0.10 \qquad [7]$$

The 0.40 threshold used in condition [5] was determined based on Landsat-TM data acquired over Sierra Nevada, California, where pixels with approximately 50% or more snow cover were found to have NDSI values greater or equal to 0.40 (Hall *et al.* 2001). The NIR threshold of 0.11 is used to prevent water pixels from being classified as snow and the green band threshold of 0.10 is used to prevent pixels with very low visible reflectance (e.g. black spruce stands) from being classified as snow (Klein *et al.* 1998).

To improve snow cover mapping in forested regions, the MOD10A1 product uses the relationship between NDSI and NDVI. The method involves selecting pixels that fall inside a specific polygon region of the NDSI and NDVI graph and labeling these pixels as snow (Klein *et al.* 1998). Poon and Valeo (2006) have defined this polygon region in the form of conditions; i.e., a forest pixel will be label as snow if conditions [6], [7], [8] and [9] are satisfied, or if conditions [6], [7], [10] and [11] are satisfied:

$$0.1 \leq \text{NDVI} < 0.25 \qquad [8]$$

$$\text{NDSI} \geq \frac{\text{NDVI} - 0.2883}{-0.4828} \qquad [9]$$

$$\text{NDVI} \geq 0.25 \qquad [10]$$

$$\text{NDSI} \geq 0.0652\, e^{(1.8069 \times \text{NDVI})} \qquad [11]$$

Hall and Riggs (2007) reported that the overall absolute accuracy of the MOD10A1 product is 93%, but varies by land-cover type and snow condition. They list the most frequent errors are due to snow/cloud discrimination problems and detection of very thin snow (< 1 cm thick).

WRMD Algorithm

The WRMD algorithm is loosely based on the MOD10A1 algorithm. Data used as input for the WRMD algorithm was extracted from the MOD02HKM data type. The digital numbers were converted to reflectance values using equation [1] and then corrected for atmospheric effects using a variation of the Dark-Object Subtraction technique proposed by Chavez (1988). The WRMD algorithm then classifies a pixel as snow if conditions [12] and [13] are satisfied:

$$\text{NDSI} \geq 0 \qquad [12]$$

$$\text{NIR (band 2)} > 0.03\text{-to-}0.09 \qquad [13]$$

A range of NIR values (0.03-to-0.09) are given in condition [13]. The actual value used is determined based on the histogram display of the NIR band; the NIR value having the lowest frequency in this range is selected as the threshold value. This is an interactive approach to exclude open water from being classified as snow, since illumination and reflectance levels vary from date-to-date.

The green band threshold presented in condition [7] and applied by the MOD10A1 algorithm was not used in the WRMD algorithm, primarily because the 0.10 threshold value would not have worked with the atmospherically corrected data and selecting an appropriate value would require fieldwork, which was not done at this time. Also a comparison of reflectance between the two major forest types in our study area found that black spruce (*Picea mariana* (Mill.) B.S.P.) dominant stands (i.e. comprise ≥ 75% of the basal area with > 50% crown closure) had a significantly higher reflectance at the $\alpha = 0.05$ level (n = 960, P-value < 0.001) in the green (mean = 0.01095, ±SD = 0.00207) and SWIR (mean = 0.0461, ±SD = 0.0138) bands than did balsam fir (*Abies balsamea* (L.) Mill.) dominant stands (i.e. comprise ≥ 75% of the basal area with > 50% crown closure) in the green (mean = 0.00953, ±SD = 0.00205) and SWIR (mean = 0.0409, ±SD = 0.0123) bands. These results were based on two-sample t-tests using

MOD02HKM reflectance data (corrected for atmospheric effects) that were acquired October 25, 2008, prior to snowfall accumulation, but after deciduous vegetation have dropped their leaves. The higher reflectance observed for the black spruce stands was attributed to surrounding vegetation, since black spruce stands are often found in low-lying areas and interspersed with wetlands. Based on these findings, there appeared to be little need to apply condition [7].

STUDY AREA

The Humber River is the second largest river system on the Island of Newfoundland (49.254° latitude, -57.335° longitude). Its watershed covers an area approximately 7800 km^2 with the headwaters located in the highlands of the Long Range Mountains (maximum elevation 836m) and its outlet located at sea level near the town of Corner Brook. The watershed is within the Boreal Shield Ecozone and is comprised mainly of balsam fir and black spruce forests (62.6%), water bodies (12.5%), wetlands (9.7%), barrens (7.2%) and other vegetation types (6.9%); landcover data derived from the Earth Observation for Sustainable Development of Forests (EOSD) project (Wulder *et al.* 2003).

METHODS

The WRMD and MOD10A1 snow cover products were visually compared to their corresponding true-color composite image to determine product accuracy. Data for the comparison was collected over the Humber River Valley for nine separate dates, starting November 19, 2007 and ending June 6, 2008. The data acquired on these dates were assumed to have minimal cloud cover, based on a visual inspection of the MODIS true-color composite image and the SWIR image. Indeed, it was difficult to obtain 100% cloud-free conditions and some clouds were observed for six of the nine dates, but were assumed to have a negligible effect on the results due to their lack of abundance and thickness.

The snow cover products were developed using the WRMD and MOD10A1 algorithms described above. The true color composite images were created from MODIS MOD02HKM data, using bands 1 (red), 3 (blue) and 4 (green) that were corrected for atmospheric effects. The MOD02HKM data were downloaded from NASA's online Web service (http://ladsweb.nascom.nasa.gov/data/search.html). The geographic extent of the MOD10A1 map does not overlap the MODIS composite image in the figures that follow, thus, the northern portion of the MOD10A1 map is missing snow cover data. This information exists on a separate file that was not included here.

Data used in this report were processed using the PCI Geomatica (v.10.1.3), MRTSwath (v.2.1) and HDF Viewer (v.2.4) software packages.

RESULTS & DISCUSSION

Table 2 lists the percent snow and ice coverage within the Humber River Watershed that was mapped using WRMD and MOD10A1 algorithms for nine dates during the 2007-2008 snow season. The difference in snow and ice cover between the WRMD and MOD10A1 products is also listed, with larger differences displayed in bold type.

The percent snow and ice cover results for both the WRMD and MOD10A1 products were similar for four of the nine dates. Five out of the nine dates show a larger difference between products (i.e. 7%-to-32%). With the exception of the November 20, 2007 image, the differences correspond well with the amount of cloud cover computed for the MOD10A1 product.

Figure 1 displays the snow map products for February 8, 2008, with the corresponding MODIS composite image, as well as a MODIS composite image acquired 52 days later on March 31, 2008. A visual comparison of the snow map products with the corresponding MODIS composite image confirms that the MOD10A1 product classifies much of WRMD snow cover as cloud. The snow and ice cover difference between the WRMD and MOD10A1 maps on February 8, 2008 is 20.64%, which is approximately the same value that the MOD10A1 product classified as cloud cover (20.92%). Figures 1c and 1d shows the outline of the MOD10A1 cloud mask overlaid on top of the MODIS composite image for February 8, 2008 and for March 31, 2008, respectively. The cloud mask matched well the pattern of white areas that appear in both images. The March 31, 2008 image was estimated by the MOD10A1 product to have approximately 2% cloud cover. In this case, it appears the MOD10A1 product misclassified patterns of snow as cloud on February 8, 2008; since it is unlikely similar patterns of cloud could remain or be replicated after 52 days.

Table 2. Percent snow and cloud cover associated with each image.

Date (Y-M-D)	WRMD snow & ice cover (%)	NSIDC snow & ice cover (%)	Snow cover difference: WRMD - NSIDC (%)	NSIDC cloud cover (%)
November 19, 2007	12.81	24.15	**-11.34**	0.45
February 8, 2008	95.72	75.08	**20.64**	20.92
March 1, 2008	97.12	80.40	**16.72**	16.68
March 31, 2008	95.78	93.87	1.91	2.04
April 7, 2008	99.52	95.68	3.85	0.06
April 27, 2008	87.29	79.86	**7.42**	9.24
May 4, 2008	55.08	23.26	**31.82**	22.01
May 14, 2008	26.29	25.88	0.41	2.22
June 6, 2008	3.80	0.40	3.40	1.05

A similar phenomenon occurred on three other dates (i.e. March 1, 2008, April 27, 2008, and May 4, 2008). The most notable of these was the image acquired on May 4, 2008, where the difference in snow and ice cover between the WRMD and MOD10A1 products was 32%. Figure 2 displays the snow and ice cover maps for this date, as well as its composite image, and an image acquired 10 days later on May 14, 2008. Cloud cover for the May 14, 2008 image was estimated by the MOD10A1 algorithm to be less than 1%. Despite the large amounts of snow that had melted between these dates (i.e. 29% reduction according to the WRMD algorithm), it is clear from the odd pattern of cloud cover as estimated by the MOD10A1 algorithm, as well as the similar pattern of white areas in the two composite images, the MOD10A1 product had misclassified snow as cloud. The snow cover in these areas is presumed to be thinner, which highlights some confusion the cloud detection algorithm has with discriminating thin snow from cloud.

Discrimination problems between snow and cloud have been documented in the literature (e.g. Hall and Riggs, 2007). Because the WRMD algorithm requires a manual approach of selecting cloud free data for analysis, the troubles with snow and cloud confusion can be avoided to some extent. Indeed, cloud cover discrimination based on visual analysis can be difficult at times when snow cover is near 100% and also when examining higher elevation areas, where there is little or no vegetation or topographic patterns to separate the appearance of bright snow from cloud. The need to manually select cloud free data adds to image processing time, and for that reason, such an approach to cloud detection would be limited to smaller scale areas, unlike the MOD10A1 product that generates snow cover maps on a global scale.

It is at times difficult to acquire cloud free data in order to produce accurate snow cover maps. This is often the case with the Humber River Watershed, where a month could pass without receiving cloud free data of the Watershed. To help overcome this problem the NSIDC distributes an 8-day global snow map, which is created from any cloud-free pixel sampled 8-days prior to map creation. The product is free to download from its website (http://nsidc.org/daac/). It has an advantage over the WRMD product due to its ability to display a complete snow cover map of the study area by piecing together multi-date partial-cloudy images. Using this idea, there are plans to incorporate multi-date MODIS imagery into the WRMD algorithm.

The MODIS products collected on November 19, 2007 tell a different story as to why there was an 11% difference in the amount of snow and ice estimated by the MOD10A1 and WRMD products. Unlike the other image dates, snow cover estimates were much higher for the MOD10A1 product (24%) compared to the WRMD product (13%). Figure 3 shows the MODIS composite data acquired on November 19, 2007 and the corresponding snow map products. It appears the WRMD omitted thinner snow layers along the edge of the existing snow cover mask, as well as to the east of Grand Lake at the higher elevations (Figure 3*c*). The MOD10A1 product did a better job identifying these areas as snow cover; however, it appears the MOD10A1 product may have misclassified dense forest regions as snow cover. An example of this would be the classification of snow along the eastern edge of Grand Lake, which in the composite image (Figure 3*a*) appears to have a dark green reflectance and no snow. The pattern of snow in this area matches the pattern of dense forest displayed in Figure 3*e*, which is a product of the EOSD project (Wulder *et al.* 2003). The two forest classes shown in this figure include: (i) dense coniferous forests, with crown closure greater than 60% crown closure and coniferous trees consisting of 75% or more of total basal area; and (ii) mixedwood dense forest, with crown closure greater than 60% and neither coniferous or deciduous trees account for more than 75% or more of total basal area.

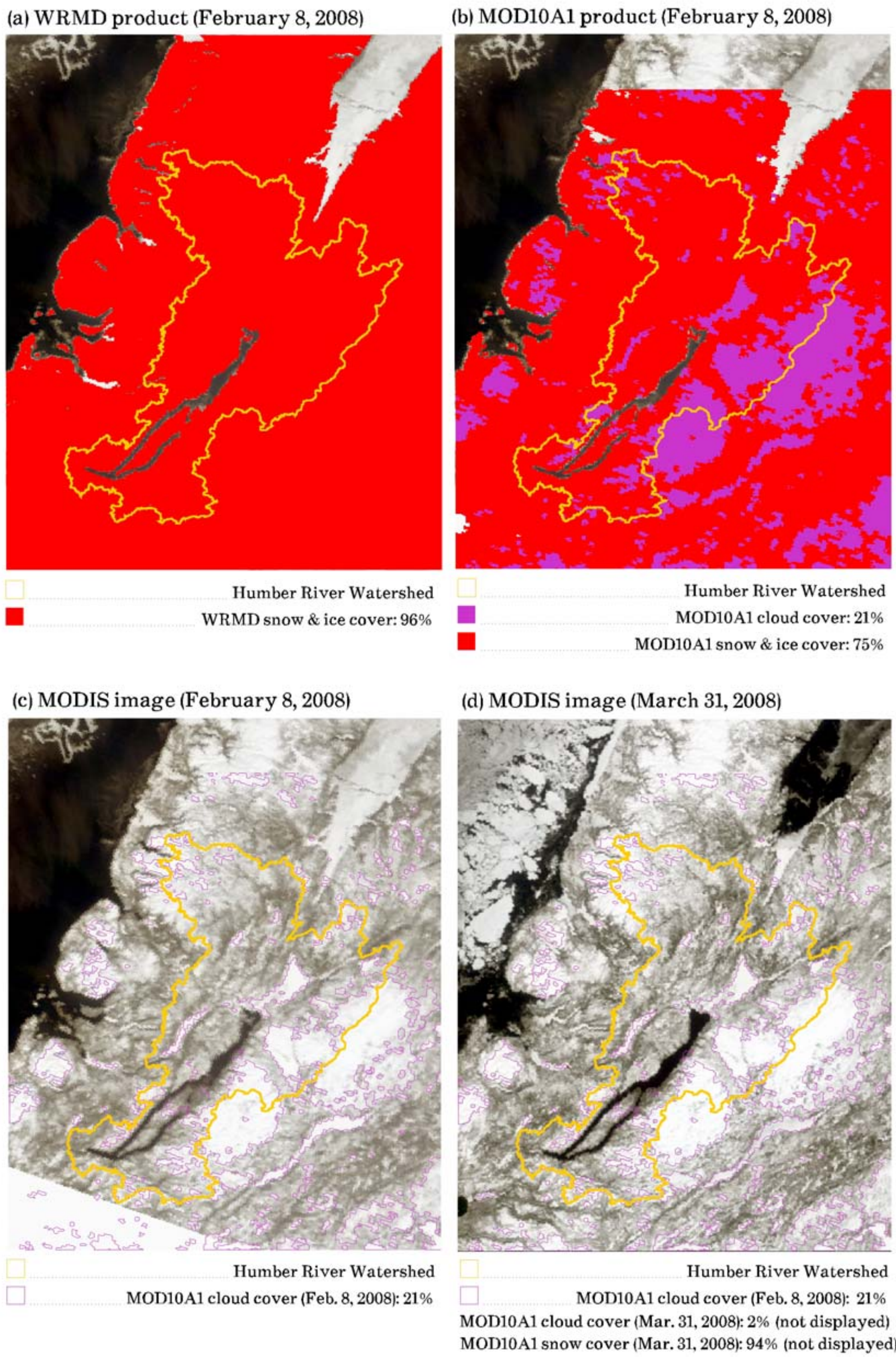

Figure 1. Snow cover products created from the (a) WRMD algorithm and (b) MOD10A1 algorithm using data acquired on February 8, 2008, with their (c) corresponding true-color composite image and (d) a MODIS image acquired March 31, 2008, when cloud cover was negligible (~2%).

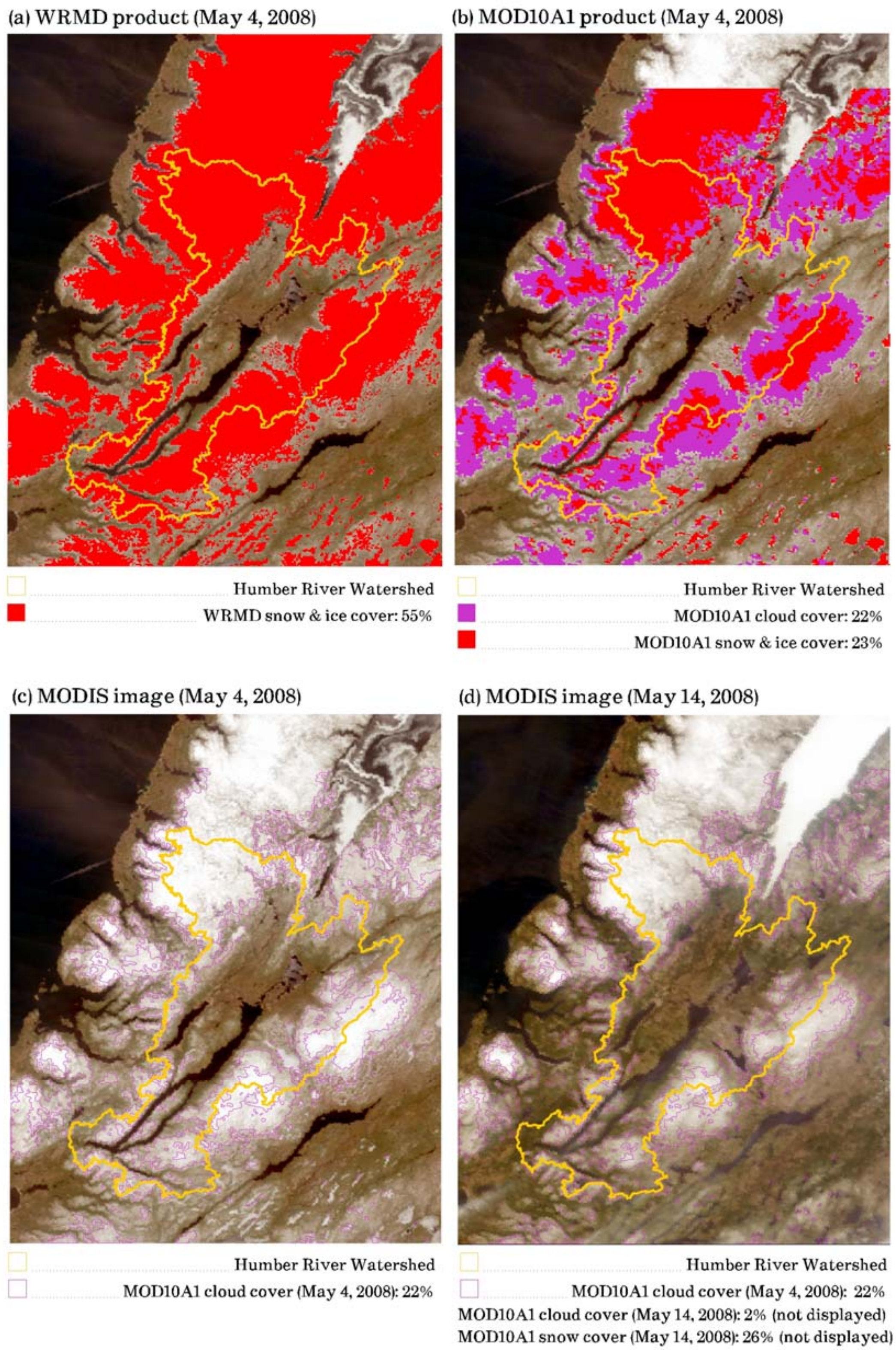

Figure 2. Snow cover products created from the (a) WRMD algorithm and (b) MOD10A1 algorithm using data acquired on May 4, 2008, with their (c) corresponding true-color composite image and (d) a MODIS image acquired on May 14, 2008, when cloud cover was negligible (~2%).

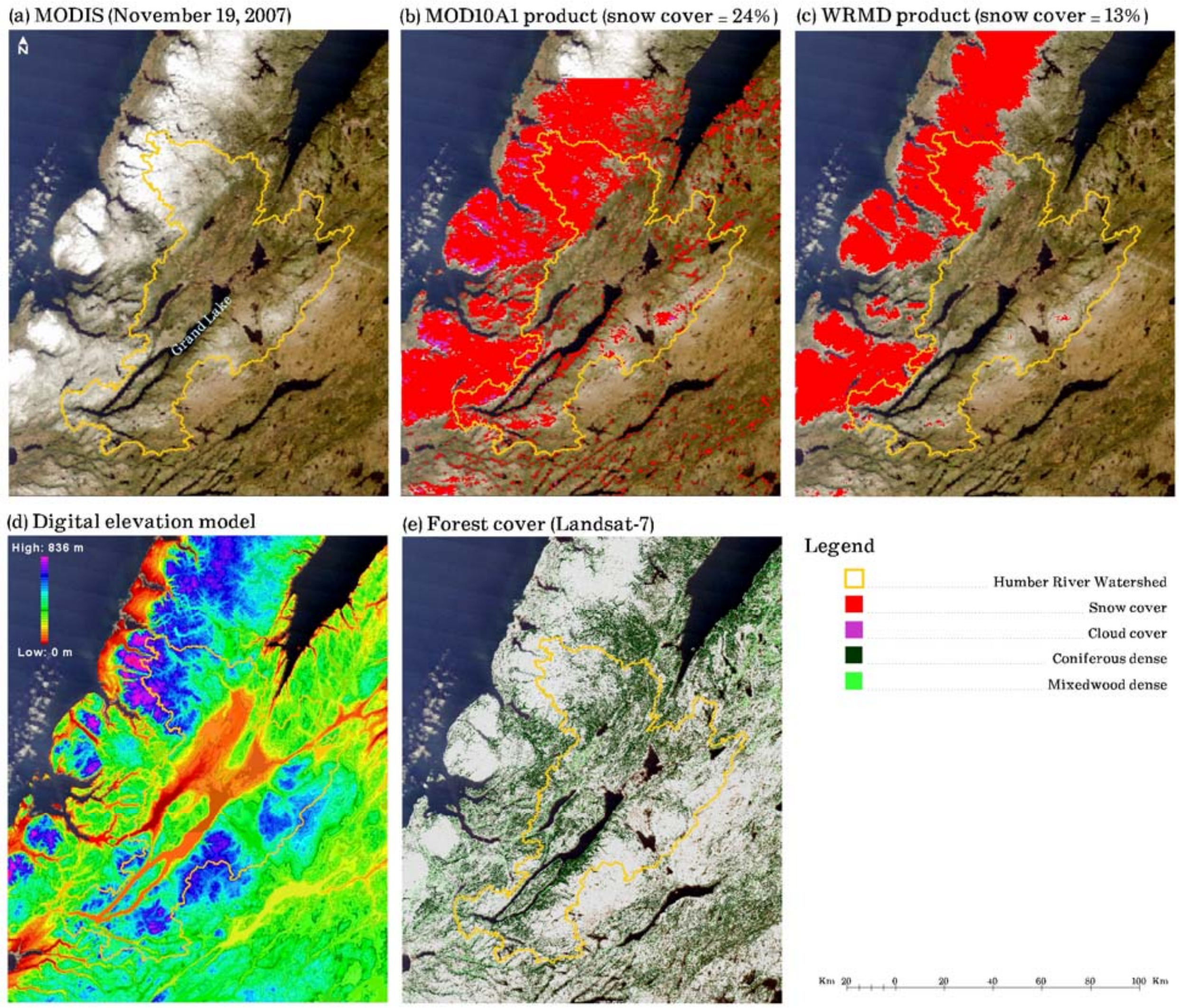

Figure 3. Spatial data of the Humber River Watershed includes (a) the MODIS true-color composite image acquired November 19, 2007, the corresponding (b) WRMD product and (c) MOD10A1 product, as well as (d) a digital elevation model and (c) a dense forest mask created from EOSD project data (Wulder *et al.* 2003).

Higher elevation areas (Figure 3*d*) adjacent to these dense forest patches are often snow free. Examples of this can be seen along the southeastern shore of Grand Lake, as well as in the southwest and southeast regions of the image where snow and forest patterns are narrow in appearance. According to the principal of lapse rate, it should be expected that snowfall will more likely be presence at higher elevations, due to a decrease in air temperature with altitude. It is uncertainty as to whether this is an actual error since no ground truthing was conducted.

CONCLUSIONS & RECOMENDATIONS

The goal of this study was to compare the WRMD and MOD10A1 snow cover products to a corresponding composite image to determine product accuracy. Errors seemed apparent for the MOD10A1 product primarily due to snow being misclassified as cloud, and to a much smaller extent, forests were misclassified as snow. Errors were also evident in the WRMD product where areas of thinner snow cover or areas where snow is mixed with vegetation cover were omitted from its snow cover mask. Work will be required to adjust the snow algorithm to map these areas as snow cover.

The WRMD product showed the most promise for mapping snow cover in the Humber River Valley. Indeed, this manual method of creating a snow map is more labor intensive than the MOD10A1 product, and is not feasible on a global scale. However, for smaller scale mapping the WRMD method of generating snow cover maps was more reliable than the MOD10A1 product, which is essential for flood forecasting.

Another point worth mentioning is related to the temporal availability of the snow cover products. Data used for the WRMD product is typically available one day after the MODIS collection date and one day before the release of the MOD10A1 product. At times of high flood risk, it is crucial that up-to-date information is available for improved decision making. Thus the WRMD product has a slight temporal advantage over the MOD10A1 product.

There needs to be further work done to ground truth the WRMD product. Based on concerns dealt with in developing the MOD10A1 product, it would be beneficial to conduct a field survey to determine whether black spruce stands without snow cover (dark reflectance pixels) are confused as snow using the WRMD algorithm. A field survey would also be beneficial in determining how well the WRMD product maps snow cover under forest canopy. Incorporating new rules into the WRMD algorithm for identifying snow under forest canopy would be worthwhile based on the omission errors observed on November 19, 2007. The idea of using a NDSI/NDVI relationship for this purpose, as applied by the MOD10A1 algorithm, would be advisable, since reflectance ratios, like the NDSI and NDVI, have less of a temporal dependence than absolute reflectance thresholds, as used in conditions [6] and [7].

REFERENCES

Chavez, P.S.J., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data, *Remote Sensing of Environment*, *24*: 459-479.

Hall, D.K., and G.A. Riggs, 2007. Accuracy assessment of the MODIS snow products, *Hydrological Processes*, *21*: 1534-1547.

Hall, D.K., G.A. Riggs, and V.V. Salomonson, 2001. Algorithm theoretical basis document (ATBD) for the MODIS snow and sea ice-mapping algorithms. (p. 45), Greenbelt, MD: NASA/Goddard Space Flight Center (Online: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod10.pdf).

Klein, A.G., D.K. Hall, and G.A. Riggs, 1998. Improving snow cover mapping in forests through the use of a canopy reflectance model, *Hydrological Processes, 12*: 1723-1744.

Picco, R.C., 1997. A comparative study of flow forecasting in the Humber River Basin using a deterministic hydrological model and a dynamic regression statistical model, *Faculty of Engineering and Applied Science*, (p. 102), St. John's: Memorial University of Newfoundland.

Poon, S.K.M., and C. Valeo, 2006. Investigation of the MODIS snow mapping algorithm during snowmelt in the northern boreal forest of Canada, *Canadian Journal of Remote Sensing, 32*: 254-267.

Riggs, G.A., D.K. Hall, and V.V. Salomonson, 2006. MODIS snow products user guide to Collection 5, (p. 80). Greenbelt, MD: NASA (Online: http://modis-snow-ice.gsfc.nasa.gov/userguides.html).

Wulder, M.A., J.A. Dechka, M.A. Gillis, J.E. Luther, R.J. Hall, A. Beaudoin, and S.E. Franklin, 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program, *The Forestry Chronicle*, *79*: 1075-1083.