Hydro temporal variability - technical documentation

ABMI Geospatial Centre

December, 2017

ABMI

Contents

1 Introduction

Surface water can vary on time scales from hourly to seasonally to decadally. Monitoring changes in surface water resources is important in Alberta as it appears the Prairie Provinces may be heading towards a future water crisis (Schindler and Donahue, 2006). Some monitoring of surface water fluctuation is done by the National Hydrological Service (Environment and Climate Change Canada, 2015) but the program cannot possibly cover all waterbodies in Alberta. The main source of surface water information in Alberta comes from the Government of Alberta Base Features Hydrography Polygons (hereafter "hydropolys") (Government of Alberta, 2004). This dataset provides great spatial coverage of Alberta's waterbodies but are limited to a certain point in time. Consistent, repeated satellite data collection may be the best way to achieve large spatial coverage and frequent (weekly) monitoring for surface water. The best data for this is Synthetic Apeture Radar (SAR) data as it is capable of detecting standing water (Malenovsky *et al.,* 2012; Brisco, 2015) at all times and all weather conditions (day, night, cloud, sun, rain). Sentinel-1 data (Copernicus [2014, 2015, 2016, 2017]) is a great source for this as it has 10m resolution, minimum 6-day revisit time (as of October 2016), and is freely available. The goal of this project is to estimate the weekly, seasonal, and yearly surface water fluctuation in Alberta, Canada using multi-temporal Sentinel-1 SAR data. A peer-reviewed paper describing this methodology is currently in press. in the Canadian Journal of Remote Sensing (DeLancey *et al.,* 2018) and the final product is on the ABMI website here: http://abmi.ca/home/data-analytics/da-top/da-product-overview/GIS-Human-Footprint-Land-Cover-Data/Hydro-Temporal-Variability.html

2 Methods

2.1 Data and SAR processing

The HTV dataset was calculated with Sentinel-1 C-band (S1) SAR data (Copernicus Sentinel data [2014, 2015, 2016, 2017]). All S1 images were gathered and processed in Google Earth Engine (GEE) (Google Earth Engine Team, 2015). GEE stores S1 ground range detected scenes which have been pre-processed with the Sentinel-1 Toolbox (Sentinel Application Platform – Sentinel-1 Toolbox). These pre-processing steps include thermal noise removal, radiometric calibration, and terrain correction (Google Earth Engine Team, 2015). S1 images were further processed in the GEE environment by performing an incidence angle correction (Gauthier *et al.,* 1998) and smoothing with a 3x3 Sigma Lee filter (Lee *et al.,* 2009) (credit to Guido Lemoine for GEE code).

S1 images intersecting with Alberta during ice free months were gathered for the time period April, 1st, 2014 – August, 5th, 2017 (see Table 1 for what defines ice free months). Winter months were not included as most lakes in Alberta are frozen from October/November to March/April. Additionally, only images with a 10m resolution were used, which resulted in the exclusion of HH or HV polarizations as these images are only available in 40m resolution. The VV polarization mode was used for the analysis as it had far more revisits over Alberta when compared to the other polarization modes (VV-VH, HH, HH-HV), however Brisco, 2015 and Bolanos *et al*., 2016 state that other polarizations are more suitable for water detection but the VV polarization is still very useful (Kasichke and Bourgeau-Chavez, 1997). This resulted in a temporal pixel stack of anywhere from 1 to 100 across Alberta.

The processing of the HTV dataset was split into two regions (Figure 1). The first region is the grasslands region of southeast Alberta which is predominately covered by native grasslands or agriculture. The second region is the boreal (boreal forest) region of northern and western Alberta which is dominated by forests. These regions were delineated using the Alberta natural regions dataset (Natural Regions Committee, 2006). The grassland and parkland regions were merged into the grassland region and the boreal, Canadian Shield, foothills, and Rocky Mountain regions were merged into the boreal region. This was done because the differentiation between water and land is distinctly different for forested areas versus low biomass grassland areas due to the lower backscatter of grasslands (Quegan *et al.*, 2000) (see Figures 2, 3, and 4).

Four ancillary datasets were used in the generation of the HTV layer or analysis of the results. Daily wind speed data from the NCEP Climate Forecast System Version 2 (Saha *et al.*, 2014) was used to remove windy days. The Shuttle Radar Topography Mission (SRTM) 30m DEM (USGS, 2006) was used to derive slope for a slope mask while the Alberta Biodiversity Monitoring Institute (ABMI) Human Footprint Inventory 2014 (HFI2014) (ABMI, 2017) was used for an agriculture and major roads mask. The Grasslands Vegetation Inventory (Alberta Environment and Parks, 2011), a polygon based inventory describing land cover in the grasslands regions of Alberta, was used to mask out low biomass grasslands. Lastly, the Government of Alberta hydropolys (Government of Alberta, 2004) were used as a training and analysis dataset.

2.2 Sentinel-1 thresholds

To generate the HTV layer two important decisions needed to be made:

1. Senintel-1 VV intensity threshold where a pixel is considered water
2. Wind speed threshold above which data should be removed from the HTV algorithm

Three Sentinel-1 images from low wind speed days in 2017 (Figure 1) were analyzed to assess which Sentinel-1 backscatter threshold resulted in the lowest error rate for classifying water from land. Training water/land data was derived from the hydropoly layer. All permanent lakes were considered as water and land was considered any area without a hydropoly feature.

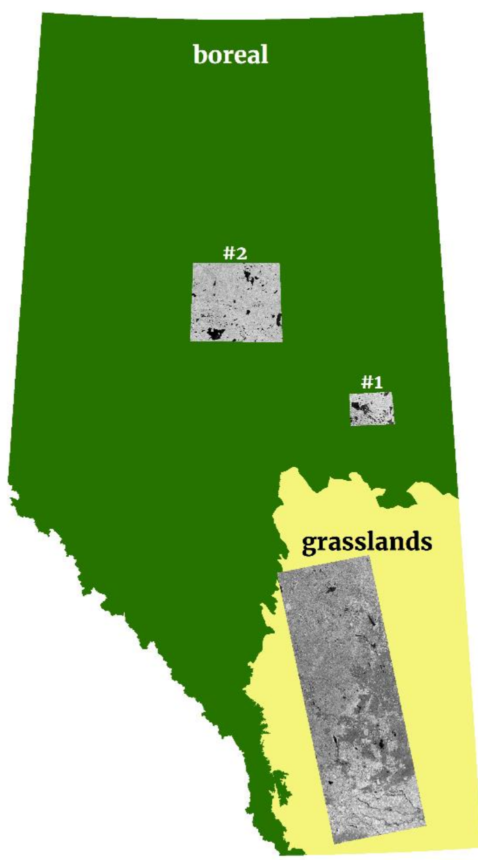

Figure 1: The spatial delineation of the boreal and grasslands processing regions with the three Sentinel-1 images used for thresholding decisions overlaid. The

boreal and grasslands regions are delineated using the Natural Regions of Alberta.

Figure 2 shows the error rate of different Sentinel-1 backscatter thresholds for differentiating water and land in the grasslands region. The lowest error rate (0.101) is seen at a threshold of -17.5 decibels (dB) with most of the error coming from false land errors. Figure 3 shows the error rate for a boreal image (boreal 1) with mixed forest and agriculture. The lowest error rate (0.049) is seen at a threshold of -15.1 dB with most of the error coming from false water pixels. Figure 4 shows the error rate for a boreal image (boreal 2) with continuous forest and minimal human footprint. A threshold of -13.7 dB generates the lowest error rate (0.048).

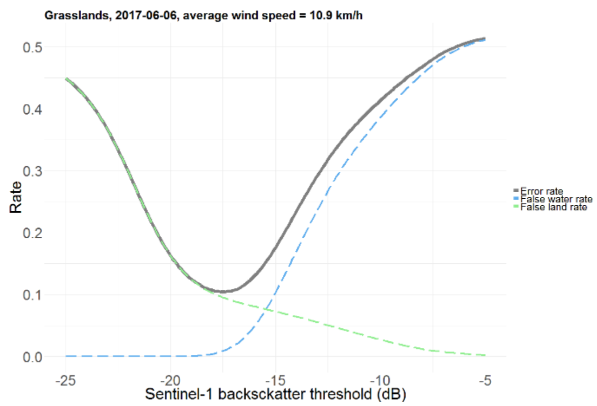

Figure 2: Errors rate of different Sentinel-1 backscatter thresholds for the grasslands test image. Image date from 2017-06-06 with an average wind speed of 10.9 km/h (from the Brooks Environment Canada weather station). The grey line represents the total error rate (false water + false land), the green line represents the false land rate, and the blue line represents the false water rate.

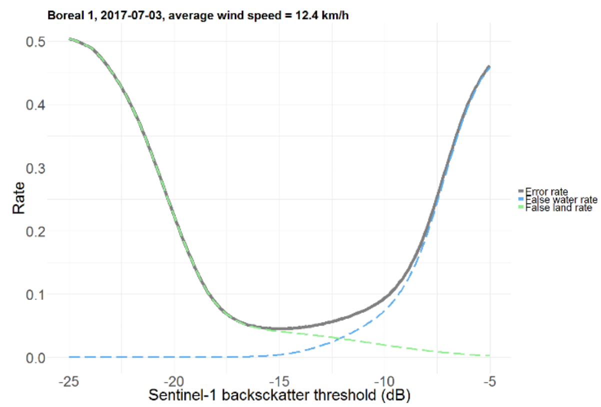

Figure 3: Errors rate of different Sentinel-1 backscatter thresholds for the boreal 1 test image. Image date from 2017-07-03 with an average wind speed of 12.4 km/h (from the Lac Le Biche Environment Canada weather station). The grey line represents the total error rate (false water + false land), the green line represents the false land rate, and the blue line represents the false water rate.

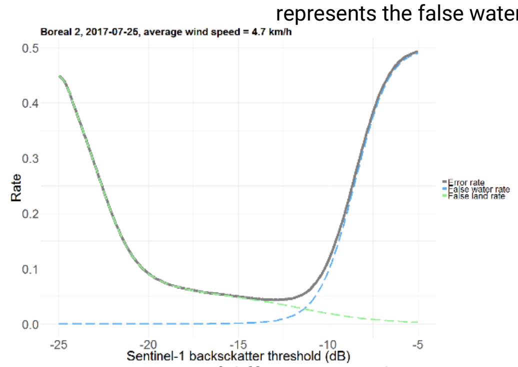

Figure 4: Errors rate of different Sentinel-1 backscatter thresholds for the boreal 2 test image. Image date from 2017-07-25 with an average wind speed of 4.7 km/h (from the Red Earth Creek Environment Canada weather station). The grey line represents the total error rate (false water + false land), the green line represents the false land rate, and the blue line represents the false water rate.

The data in Figure 2 show that an acceptable threshold for the grasslands region is -17.5 db. For this region it is expected that there will be an approximate error rate of 10% when classifying water and land. Figures 3 and 4 give different optimal thresholds for the boreal region (-15.1 and -13.7 dBs respectively). Since false land error is preferred over false water error for this algorithm, the lower -15.1 dB threshold was chosen for the boreal region. Even with this lower threshold the boreal 2 image still has an error rate of only 5%. Therefore in the boreal region we can expect an error rate of about 5% for classifying water.

Figure 5 shows the trend of increasing Sentinel-1 backscatter values with increased wind speed. A wind speed greater than 9 km/h was chosen as the threshold where data would be removed from the HTV algorithm. This was chosen since 90% of the data points in this threshold were below the (-17.5) dB threshold. We acknowledge that this modeled wind data is too coarse to pick up on wind gusts or increase in speed over large lakes and thus we see a poor fit between the relation in wind speed and Sentinel-1 backscatter. However, the main goal of this threshold is to use "potentially" calm days but it is known that wind at the time of image acquisition may not have actually been calm.

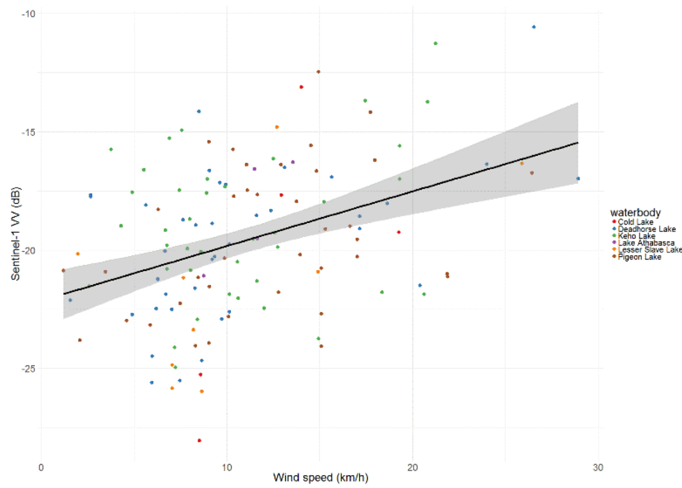

Figure 5: Wind speed (km/h) versus Sentinel-1 VV backscatter (dB) for six waterbodies across Alberta. With a fitted trend line (R2 = 0.15).

2.3 HTV algorithm

To calculate HTV, each S1 image was first turned into a binary "water" (1) and "non-water" (0) image. Any pixel below -17.5, and -15.1 dB for the grasslands and boreal zone respectively was considered water. To account for lake waves causing higher backscatter values, any pixel where maximum wind speed, for the day of acquisition, above 9 km/h was removed (see Figure 5). To account for low backscatter values on the lee side of mountain slopes, all pixels with a slope value greater than 15 were removed from the analysis using SRTM DEM (USGS, 2006). This was done based on visual interpretation of low backscatter values on parallel ridges corresponding to slopes over 15 degrees. All pixels overlapping with cultivation or major roads (the ABMI Human Footprint Inventory for 2014 conditions, ABMI, 2016) were assigned a value of zero. Finally, a grasslands mask was applied by masking out all grassland polygon habitat types from the Government of Alberta's Grassland Vegetation Inventory (Alberta Environment and Parks, 2011). After all the thresholds and decisions, the binary water images were summed to get the number of times each pixel was classified as water. This was then divided by the number of pixels in the total pixel stack (after masking), multiplied by 100, and turned into integer format to get the percent of time a pixel was identified as water.

This method of assigning percent pixel values can be seen in Figure 6. The four panels of Figure 6 show the raw Sentinel-1 VV backscatter values. Generally vegetation is seen as bright and water is seen as dark. In Figure 6 panels a) April, 14th and c) May 24th the two lakes on the right appear to be filled with water (dark) while in panel b) April, 30th and d) October, 21st the lake appears to have been drained (same values as the surrounding vegetation). In the HTV layer, most of the pixels in the lake would be assigned a value of 50 as water was seen twice in a pixel stack of four images.

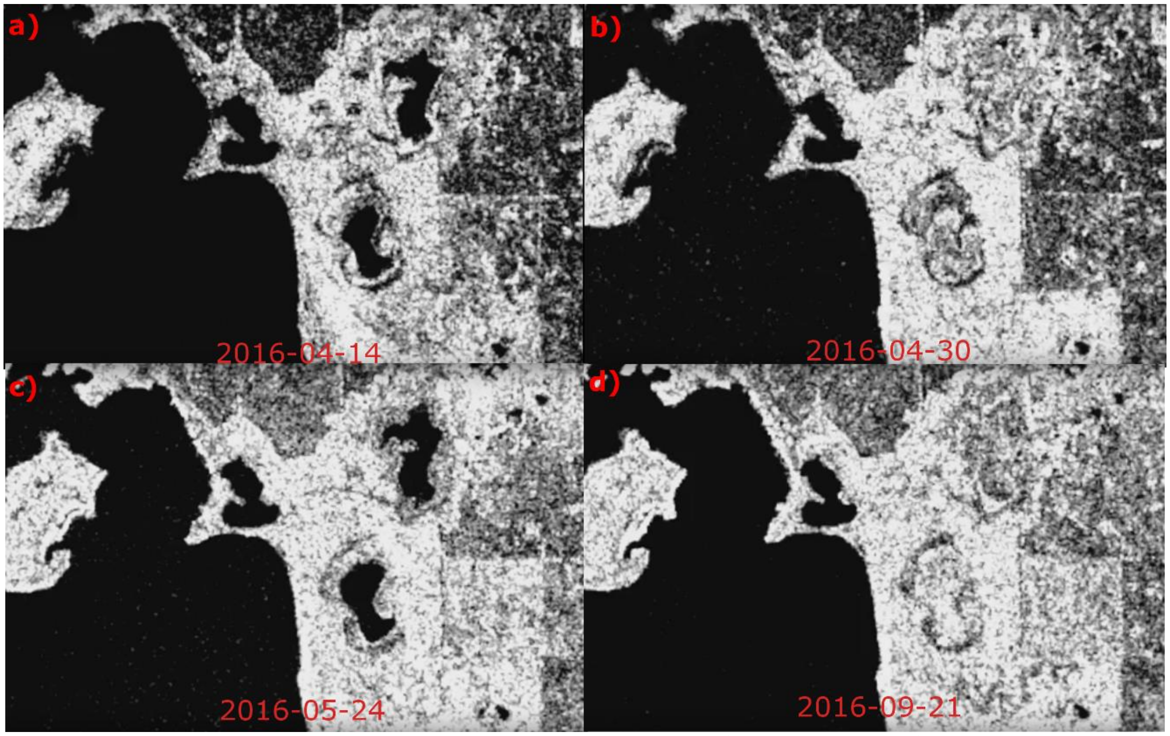

Figure 6: Four raw Setninel-1 images showing the variation of the two lakes (Lynn and Boss Lake) on the right side of the image throughout the ice free 2016 period. These images summed together would give a HTV value of 50 for most of Lynn and Boss Lake. Each day shown in these images had a wind speed below 10 km/h.

2.4 Comparison to Alberta Base Features Hydrography Polygons dataset

A comparison was done between the hydropoly layer and the HTV dataset. To see how the HTV data represented different waterbody types, permanent lakes, recurring lakes, and rivers were extracted from the hydropoly dataset. HTV values were extracted for each waterbody type and the distribution of values was plotted for each water body type.

3 Results

In total, the HTV product created in the GEE environment used 125 billion pixels in the calculation (temporal pixel stack x number of 10m pixels in Alberta). Figure 7 shows the pixel count used for calculations in the HTV dataset. This result shows a high of 65 for in the southwest and 0 for a strip along Lake Athabasca.

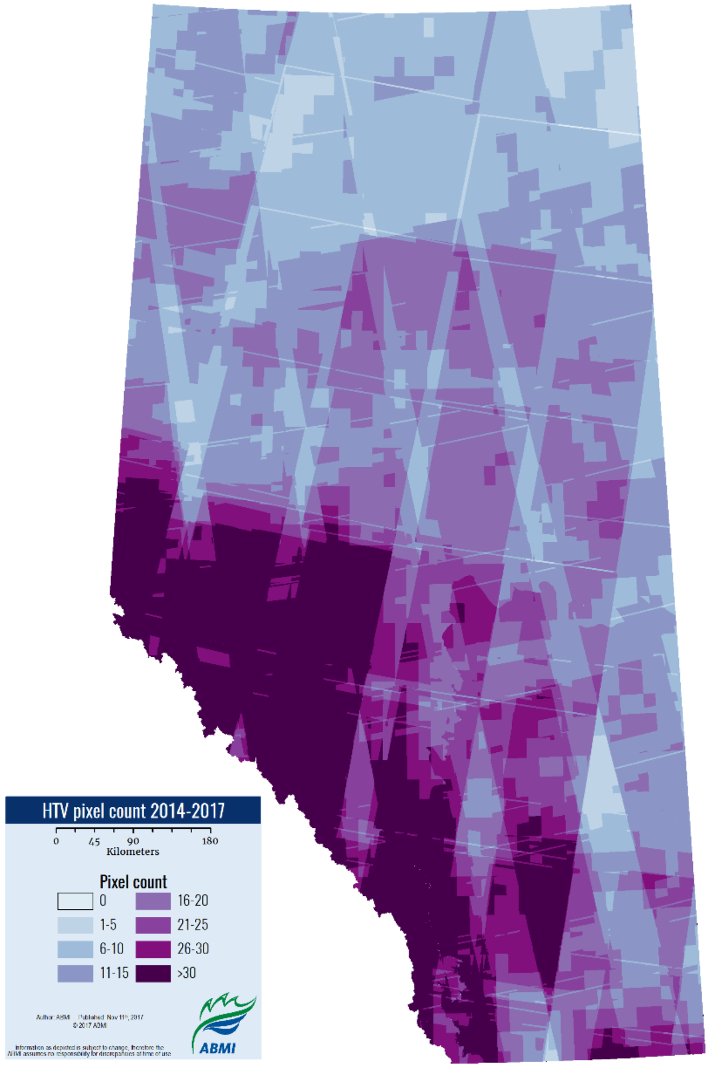

Figure 7: Pixel count of 10m pixels used in the HTV calculation (wind speeds over 9 km/h removed).

The HTV layer is shown in Figure 8. This represents the percentage of time each pixel was identified as water. This layer maps the boundaries of waterbodies and gives an idea of how permanent they are and how much they fluctuate yearly or seasonally. Generally, land can be visualized with values of 0-10, recurring waterbodies with values of 11-65, and permanent lakes with values of 66-100. With this visualization it can be determined if each lake is permanent or recurring and it can also be used for delineating the permanent boundaries of lakes and the dynamic zones of lakes.

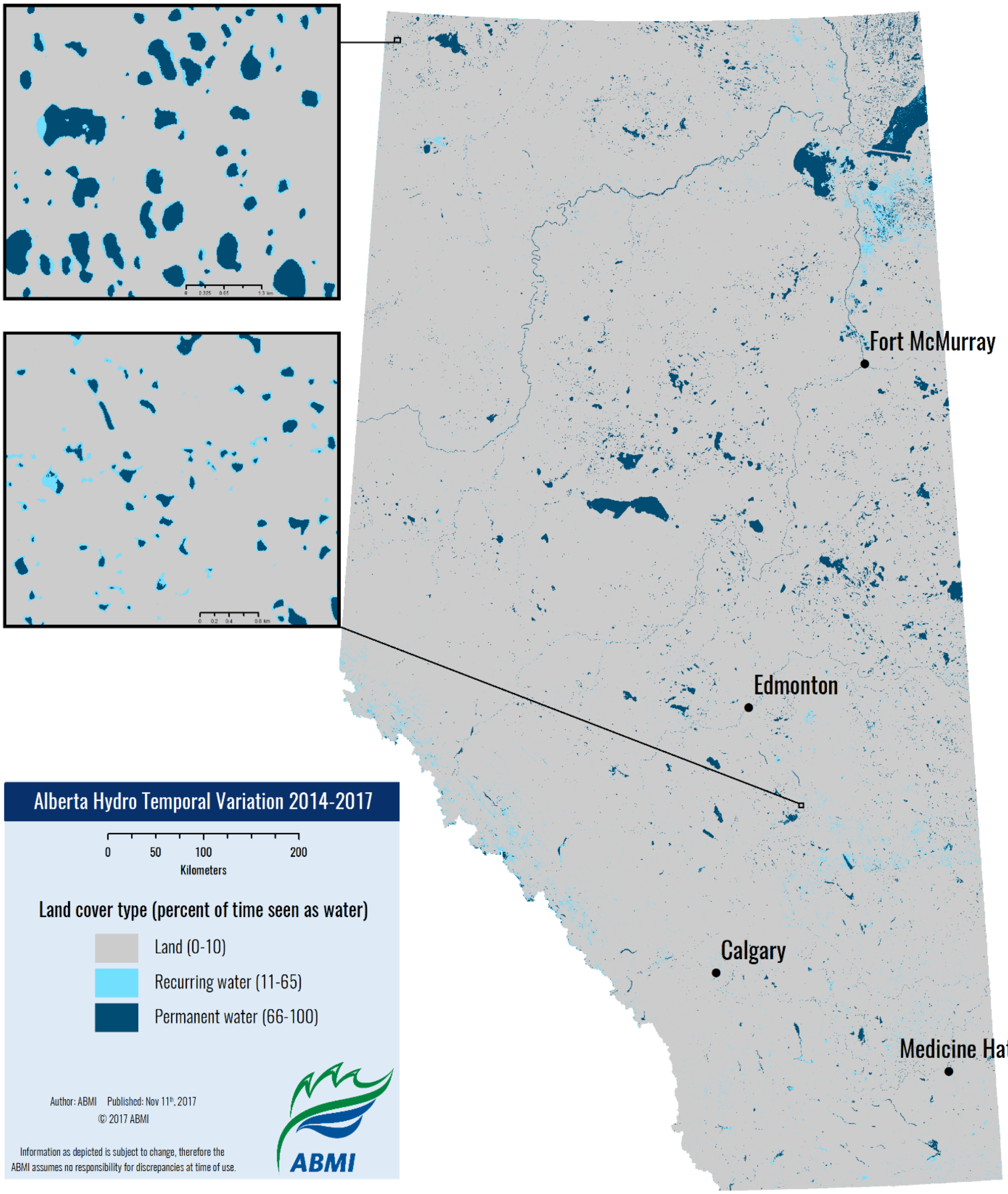

Figure 8: HTV for all of Alberta for the 2014 to 2017 time period. HTV values represented in three classes: Land (HTV values 0-10) – where water is never or rarely seen, Recurring water (11-65) - lakes which are seasonal or the area of lake level fluctuation around a permanent water body, and Permanent water (66- 100) - areas with consistent water. Inset areas provide more detail to show the dynamic regions around permanent lakes.

In a comparison to the hydropoly layer, the HTV layer was shown to have a higher value for permanent lakes with most permanent lakes having a HTV value ranging from 62-87 (Figure 9). Recurring lakes had a peak distribution of HTV values at about 20 and very few with values over 60 (Figure 9). The majority of rivers had a low HTV value of 20-30 but there were some rivers with HTV values of 95-100 which was very rare in permanent and recurring lakes. The problematic areas for this dataset are the low biomass grassland areas of southeast Alberta, the Rocky Mountains (southwest Alberta), and the sand dunes around Lake Athabasca (northeast Alberta). These areas have visible false water error but slope and grassland masking have helped minimize this error.

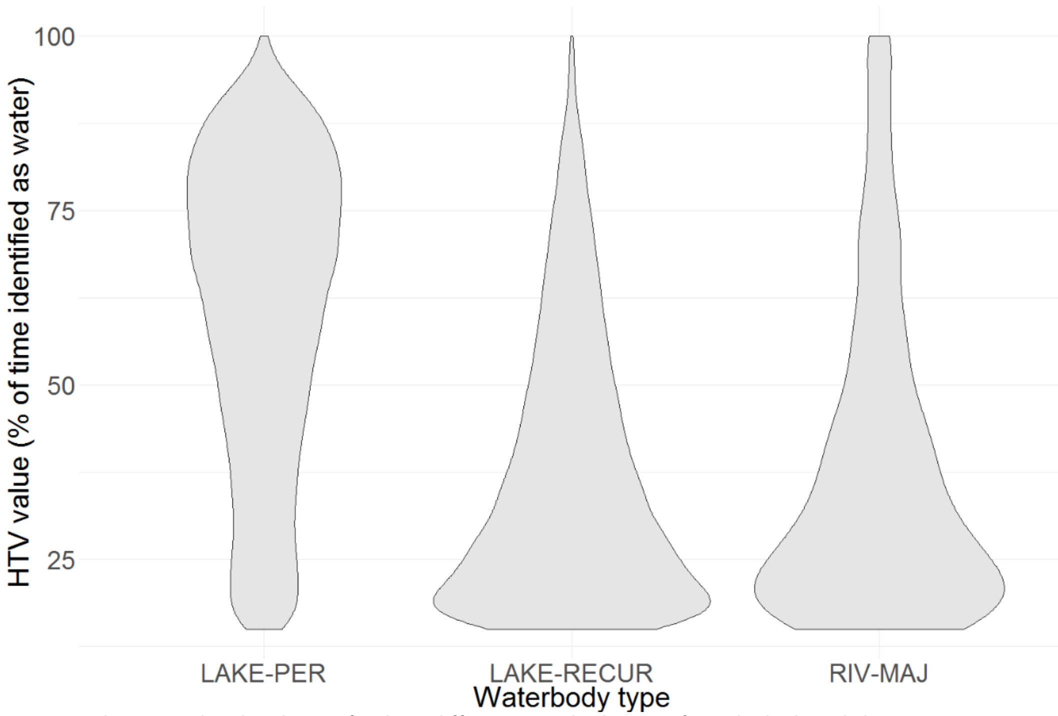

Figure 9: The HTV value distribution for three different waterbody types from the hydropoly layer. LAKE-PER = permanent lakes, LAKE-RECUR = recurring lakes, and RIV-MAJ = rivers. Lakes less than 1ha were removed due to inaccurate mapping of the boundaries in many cases and lakes/rivers with HTV values below 15 were removed as these are likely lakes which no longer exist or recurring lakes which did not have water in them during the imaging time period. Note Lake Athabasca was eliminated from these statistics due to a pixel count of 0 in some areas of the lake.

4 References

Alberta Biodiversity Monitoring Institute Geospatial Centre. 2017. "Human Footprint Inventory 2014 – Public Version, Version 2." Edmonton, Alberta.

Alberta Environment and Parks, Government of Alberta. 2004. "Base Water Polygon." Edmonton, Alberta, Canada.

Alberta Environment and Parks, Government of Alberta. 2011. "Grassland Vegetation Inventory (GVI) Specifications." Edmonton, Alberta, Canada.

Bolanos, S., Stiff, D., Brisco, B, and Pietroniro, A. 2016. "Operational Surface Water Detection and Monitoring Using Radarsat 2." *Remote sensing*, Vol. 8(No.4): pp. 285.

Brisco, B. 2015. "Mapping and Monitoring Surface Water and Wetlands with Synthetic Aperture Radar." In: *Remote Sensing of Wetlands: Applications and Advances*. Bocca Raton, FL, USA. Taylor and Francis Group.

Copernicus Sentinel-1 data [2014, 2015, 2016, 2017], European Space Agency.

DeLancey, E.R., Kariyeva, J., Cranston, J., and Brisco, B. 2018. "Monitoring hydro temporal variability in Alberta, Canada with multi-temporal Sentinel-1 SAR data." *Canadian Journal of Remote Sensing,* Vol. 44(No.1): pp.1-10.

Environment and Climate Change Canada. 2015. "The National Hydrological Service. Ottawa, Ontario, Canada.

Gauthier, Y., Bernier, and Fortin, J-P. 1998. "Aspect and incidence angle sensitivity in ERS-1 SAR data." *International Journal of Remote Sensing*, Vol. 19(No.10): pp. 2001-2006.

Google Earth Engine Team. 2015. "Google Earth Engine: A planetary-scale geospatial analysis platform." https://eathengine.google.com.

Lee, J-S., Wen, J-H., Ainsworth, T.L., Chen, K-S., and Chen, A.J. 2008. "Improved Sigma Filter for Speckle Filtering of SAR Imagery." *IEEE Transactions on Geosciences and Remote Sensing*, Vol. 47(No.1): pp.202-213.

Malenovsky, Z., Rott, H., Chihar, J., Schaepman, M.E., Garcia-Santos, G., Fernandes, R., and Berger, M. 2012. "Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land." *Remote Sensing of Environment*, Vol. 120: pp. 91-101.

Natural Regions Committee. 2006. "Natural Regions and Subregions of Alberta." Compiled by D.J. Downing and W.W. Pettapiece. Government of Alberta. Pub. No. T/852.

Quegan, S., Le Toan, T., Yu, J.J., Ribbes, F., and Floury, N. 2000. "Multitemporal ERS SAR analysis applied to forest mapping." *IEEE Transactions of Geoscience and Remote Sensing,* Vol 38(No. 2): pp. 741-753.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nagdiga, S., Tripp, P., Behringer, D., Chuang, H-Y., Iredell, M., Ek, M., Yang, R., Mendez, M.P., Dool, H.v.d., Zhang, Q., Wang, W., Chen, M., and Becker, E. 2014. "The NCEP Climate Forecast System Version 2." *Journal of Climate*, Vol. 27: pp. 2185-2208.

Santoro, M., and Wugmüller, U. 2014. "Multi-temporal Synthetic Aperture Radar Metrics Applied to Map Open Water Bodies." *IEEE Journal of Selected Topics in Applied Earth observations and Remote Sensing,* Vol. 7(No. 8): pp. 3225-3238.

USGS. 2006. "Shuttle Radar Topography Mission." Global Land Cover Facility, University of Maryland.