Capstone

Examining Thresholding and Factors Impacting Snow Cover Detection Using Nighttime Images Captured by the Visible Infrared Imaging Radiometer Suite Day/Night Band

Author:

Renato Stopic – Amsterdam University College

renato.stopic@student.auc.nl

1 June 2022

Supervisor:

dr. Eduardo Simao Da Graca Dias – Vrije Unversiteit Amsterdam

e.simaodagracadias@vu.nl

Reader:

dr. Eric Koomen

e.koomen@vu.nl

Tutor:

dr. Ydwine Zanstra

Word Count: 7079

Sciences

Abstract

Using nighttime remote sensing data from the Visible Infrared Imaging Radiometer Suite Day/Night band (VIIRS DNB) enables snow cover extent detection in areas with limited daytime hours due to their latitude or in regions covered by clouds during the day. The main question in this field remains, determining at what threshold a pixel is considered covered by snow and what factors cause false estimates. This study tested seven automated thresholding algorithms and compared the results to the MODIS MOD10A1 snow cover extent product. Besides that, both the false positives and false negatives were examined by splitting the results into discrete categories and visual analysis. The results found that Li thresholding gives the most consistently good results, with an overall accuracy between 65% and 81%, while Mean thresholding performed best in mountainous regions but struggled in other areas. Most false negatives are caused by forests, especially closed and evergreen forests. The analysis of NDVI data matches these findings, with the NDVI of false negatives being higher than true positives. False positives appear to be primarily located in or around built-up areas. This study shows that using nighttime VIIRS DNB data can supplement snow cover extent data, especially in regions with a low number of forests. This approach could be helpful in high latitude regions where there are few daytime hours in winter or in areas covered by clouds during the day and aid in creating better spatial and temporal coverage of snow cover extent data.

Keywords: Remote Sensing, Snow Extent, Nighttime, Moonlight, VIIRS, Thresholding

Introduction

Knowing which areas of the planet are covered by snow or ice is of major importance to many fields of earth science, including climatology, meteorology, hydrology, and others. The importance of snow comes from the fact that it has an impact on both local environments and on global climate. The local impacts are especially notable in hydrology because molten snow is frequently the main source of fresh water for many parts of the world (Barnett et al., 2005; Musselman et al., 2017). On the global level, snow has a significant impact on the earth's energy balance and temperature due to its high albedo (Groisman et al., 1994; Levis et al., 2007). Therefore, knowing what regions are and are not covered by snow will only become more important in the coming years as climate change will heavily impact snow and ice cover across the world, especially in high latitude regions and mountainous areas (Anisimov et al., 2007).

One way to determine which areas are covered by snow is by using remote sensing. Since the beginning of remote sensing in the early 1970s, snow was one of the first things that were investigated because of its high albedo. This characteristic meant that snow was relatively easy to detect, which was especially important for early satellites due to their limited sensor quality (Foster & Hall, 1991). Snow detection continued throughout the years with new satellite launches, but significant advancement in the field came with the launch of satellites equipped with the Moderate Resolution Imaging Spectroradiometer (MODIS). The introduction of MODIS enabled the creation of daily snow cover products on a global scale. Due to its high temporal resolution and global coverage, the MODIS snow cover product is widely used in many scientific fields, including climatology and hydrology (Da Ronco et al., 2020). The MODIS snow cover extent product continues to be actively released, and the most recent version of the product is version 6.1 of MOD10A1 (Hall & Riggs, 2021).

While the MODIS snow cover product creates daily snow extent maps, that does not necessarily mean that there is data available for all regions every day. There are factors that can potentially prevent the satellites from acquiring the necessary data to determine if there is or is not snow. The most notable factors are limited light and cloud coverage. The limitation of low light is especially impactful in high latitude regions that receive few hours of daylight in winter months due to their position on the globe

(Miller et al., 2013). Even when there is sufficient light, it is often possible that the view of the ground, and whether it is covered by snow or not, is blocked by the presence of clouds. Unlike available light, the impacts of clouds are noticeable in both high latitude regions and other potentially snow-covered areas in mid-latitude regions.

One way in which both of these potential obstacles can be avoided is by using satellite images captured during the night. The illumination from moonlight during certain lunar phases can be enough to detect things with a high albedo, which would also include snow (Huang et al., 2022; Lee et al., 2006). The ability to use illumination from moonlight means that the data collection would not be limited by the low number of daytime hours in winter in high latitude regions.

The use of nighttime data would also increase the possibility of data being captured in cloud-free conditions anywhere on the planet. Currently, all regions which were covered by clouds during the day would have no data available. However, if the clouds clear during the night, it is possible that the necessary data still gets captured to determine if there is or is not snow. Besides that, another benefit of remote sensing at night would come from the fact that there are assumed to be fewer clouds at night due to the reduced convection (Foster, 1983a; Wiesnet et al., 1981).

The idea of snow detection with nighttime images has been around since the early days of nighttime satellites. The first attempts were made using the Defense Meteorological Satellite Program's Operational Line-Scan System (DMSP-OLS) in the 1980s and 1990s, but their usefulness was limited due to their coarse resolution. While the resulting data might not have been extremely useful, these studies still served as a proof of concept, and they showed that this approach to snow detection was possible (Foster, 1983b; Foster & Hall, 1991). It was clear that in order to improve and further the field, a new satellite equipped with a newer and higher resolution sensor would be necessary. The need for a new satellite was further exemplified when the launch and specifications of the Visible Infrared Imaging Radiometer Suite (VIIRS) were announced, which sparked discussions about what this improved sensor could enable, which of course, included snow detection (Lee et al., 2006; Miller et al., 2005). Shortly after the launch of the Suomi National Polar Partnership (S-NPP) satellite equipped with VIIRS, the first study detecting snow cover was conducted which showed that this approach would be feasible with

the sensor (Miller et al., 2013). Further research into this field has come in recent years, with the two most notable studies being Huang et al. (2022) and Liu et al. (2021).

In the same way that there are limitations and challenges with snow detection during the day, there are also challenges with snow detection during the night. One of these challenges is determining the threshold at which any given pixel of the VIIRS image is assumed to be covered by snow. While some studies approached this challenge with visual identification (Miller et al., 2013), it is also possible to calculate the threshold using different automated thresholding algorithms. So far, only one thresholding algorithm has been applied, that being the Minimum Error Thresholding done by Huang et al. (2022). While the number of used thresholding methods for this specific purpose is limited, this is not the case in other areas of remote sensing. Even for snow detection during the day, there have been studies that examined multiple different thresholding algorithms, including the study done by Yin et al. (2013), which included nine different algorithms.

Another challenge to nighttime snow detection is that other factors besides snow cover can impact the image and, therefore, the results. Huang et al. (2022) looked at what the adequate lunar angle would be to determine snow cover extent and if there was a difference between different lunar phases. However, the lunar angle is not the only thing that can impact the data; ground features can also have a significant impact on the captured image. Land cover can play a crucial role in VIIRS Day/Night Band (DNB) data since different land covers will have different albedos and, therefore, different pixel values when captured by the sensor. The impact of land cover on VIIRS data has already been documented in a study that showed that seasonal land cover change impacted the overall brightness of VIIRS DNB images (Levin, 2017).

While there are already existing ways to determine snow cover extent, there are some limitations, like lack of light or cloud blockage, which could be addressed by using data captured at night. Because nighttime snow extent estimation is still in its early phases, some questions still need to be answered before researchers can more widely adopt it. The two most notable questions are; what thresholding methods are the best and what factors impact the data besides snow coverage.

Therefore, this study examined seven different automated thresholding algorithms in order to determine which of them performs the best for snow detection using nighttime imagery captured by the S-NPP VIIRS DNB. After identifying the best automated threshold algorithm, the false estimates were analyzed to determine what factors could possibly be the cause of the misclassification. Improvements in this field could lead to a better understanding of what areas and under which conditions could be most suitable for snow cover detection at night. Furthermore, snow cover detection at night could further enable better spatial and temporal coverage of snow cover extent data, which could play a significant role in furthering the understanding of the impact snow coverage has on weather, climate, climate change, and hydrology.

Case Study Area

For the purposes of this study, four different case study areas were selected. The selection was made to incorporate different latitudes, elevations, topologies, and land covers. The four case study areas are spread throughout North America, including both the United States of America and Canada. The case study areas are located in Colorado, USA (a), Ontario, Canada (b), Alaska, USA (c), and Saskatchewan, Canada (d), and their locations are pictured in *Figure 1*.

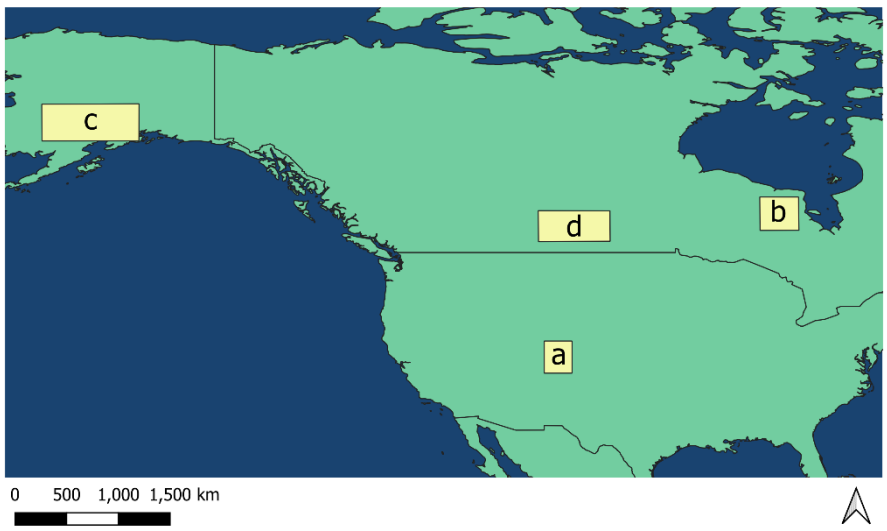

Figure. 1. Locations of case study areas. Colorado, USA (a), Ontario, Canada (b), Alaska, USA (c), and in Saskatchewan, Canada (d).

The case study areas in Colorado (a) and Alaska (c) are primarily mountainous, with the area in Colorado encompassing a part of the Rocky Mountains, while on the other hand, the area in Alaska encompasses part of the Alaska Range. The major difference between the two areas lies in their latitude. The Colorado case study area is in a mid-latitude region, while the Alaska case study area is in a high latitude region. Even though the Colorado case study area is not limited by the number of daytime hours due to it being located in a mid-latitude region, it could still experience cloud coverage during the day, and therefore it is included in this study as snow cover extent data could be added if nights are clear.

The other two case study areas are located in Canada and are mostly on flat terrain. The major difference between the two areas lies in their land cover, with the area in Saskatchewan (d) being primarily covered in cropland; in contrast, the area in Ontario (b) is covered mainly by different types of natural land cover including several types of forests, shrubs, and herbaceous vegetation.

Besides the geographic location of the case study areas, another important factor was selecting the exact date and time at which the data was captured. Three major factors determined the appropriate date for the data capture. The first factor was ensuring that the data was captured during a full moon. While Huang et al. (2022) suggest that other lunar phases are also suitable for nighttime snow detection, only full moon phases were selected for this study in order to reduce the potential impact different lunar phases could have on the study. Besides the data being captured during a date that had a full moon, it was also essential to guarantee a sufficient amount of day and night hours. Since this study uses both MODIS data captured during the day and VIIRS DNB data captured at night, the selected dates ended up being in March and October, during both of which there usually is snow cover in the selected case study areas, and both the day and night are sufficiently long for there to be enough time to capture the data. Finally, the last factor was ensuring that there was no or little cloud coverage during the data capture to ensure that the captured data could be used for this study. The exact dates of data collection for the four case study areas can be seen in *Table 1*.

Data Collection

For this study, multiple different data sources were needed for the four case study areas. The data includes both VIIRS DNB data for the nighttime snow extent estimates and the MODIS snow cover product to compare the resulting snow cover extent map gotten from the VIIRS data. Besides those two, data was needed to determine what factors could explain the wrong estimates. In this case, that data includes land cover data and necessary red and near-infrared band data to calculate Normalized Difference Vegetation Index (NDVI) values.

The VIIRS DNB data was captured by the Suomi National Polar Partnership (S-NPP) satellite. The captured data was accessed through the Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center (LAADS DAAC), operated by the National Aeronautics and Space Administration (NASA). The data was part of the NPP_VDNES_L1 - VIIRS/NPP Day/Night Band 6-Min L1 Swath SDR 750m product, and for this study, the raw radiance values from the DNB were used (LAADS DAAC, 2018). The exact date and time (UTC) of the data used in this study can be found in *Table 1.*

Location	**Date**	**Time (UTC)**
Colorado	March 2 2018	08:30
Ontario	March 2 2018	08:24
Alaska	October 13 2019	14:06
Saskatchewan	March 21 2019	10:06

Table 1. Date and time of data collection by case study area

Besides the VIIRS data, MODIS snow cover extent data was also used in this study, specifically the latest version 6.1 of the MOD10A1 snow cover extent product (Hall & Riggs, 2021). This data source was selected because it is commonly used in many hydrology and climate studies and has relatively high accuracy (Da Ronco et al., 2020). The snow cover extent data was selected for the appropriate date and case study area as indicated in *Table 1*, but since the product is released daily, no exact time of capture is given.

Besides snow data, other data sources were also used in this study. One of which is land cover data. Since this study covers different countries, it was important to select land cover data that has global coverage. Other than what countries are covered, the level of detail of the included data was also an important factor, as a more detailed breakdown of forest and herbaceous vegetation types could lead to a clearer understanding of what kind of land cover impacts snow detection. For this reason, the Copernicus Global Land Service data was used for both 2018 and 2019 (Buchhorn et al., 2020a, 2020b).

Finally, in order to calculate the Normalized Difference Vegetation Index (NDVI), near-infrared and red band data was needed for the case study areas. For the purposes of this study, another MODIS product was used to guarantee consistency and data availability. The product in question is version 6.1 of MOD09GA, which includes the two necessary bands for NDVI calculation (Vermote & Wolfe, 2021).

Data Preparation

In order to maximize the potential of accurate threshold calculations, certain parts of the VIIRS DNB data were masked as they are either not considered in this study, like bodies of water, or because they emit light which could skew the data, like built-up areas (Huang et al., 2022). To determine which parts of the VIIRS data had to be masked, the Copernicus land cover data was used. Areas that overlapped with pixels marked as Urban/Built-Up, Permanent Water Bodies, and Open Sea were those that were masked in the VIIRS DNB data. In this step, any clouds present would also have been masked, but since the case study areas were explicitly chosen to avoid cloud coverage, this step was not necessary for this study.

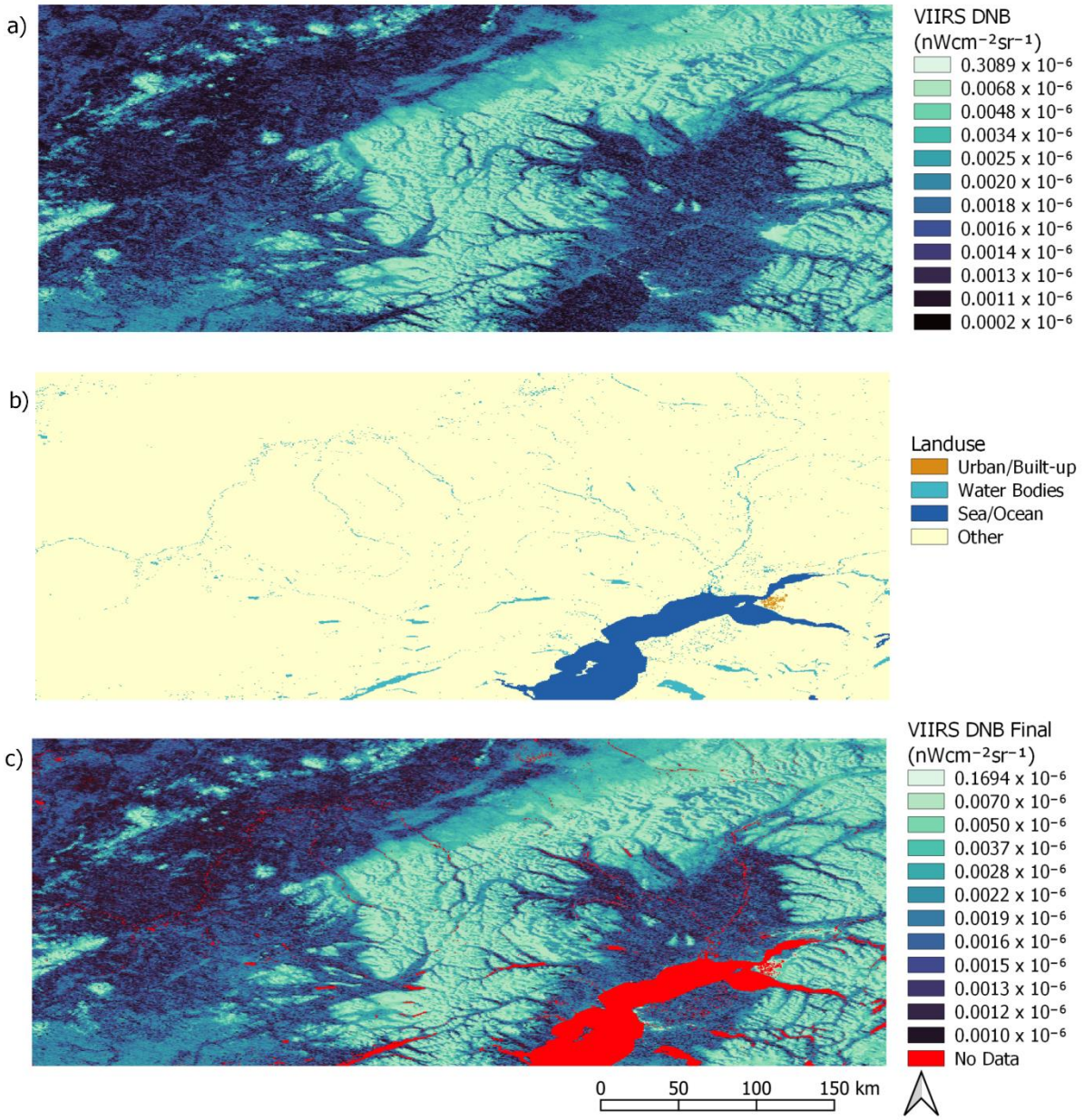

Figure 2. Example of VIIRS data preparation in the Alaska case study area. (a) Raw VIIRS Data (b) Copernicus Land Cover data with separated Open Sea, Water Bodies, and Built-Up/Urban area (c) Final preprocessed VIIRS data with Open Sea, Water Bodies, and Built-Up areas masked

The MODIS snow cover extent data was also preprocessed in order to create a binary snow cover map. All pixels were marked as snow-covered, snow-free, or as no data in case of cloud coverage or other data gaps. This was done following the steps outlined in previous studies in order to ensure consistency between them (Huang et al., 2018, 2022). While the VIIRS data was completely cloud-free,

some cloud coverage was present in the MODIS data. Therefore, all cloud-covered pixels and any other missing data were masked and were not considered in the calculations.

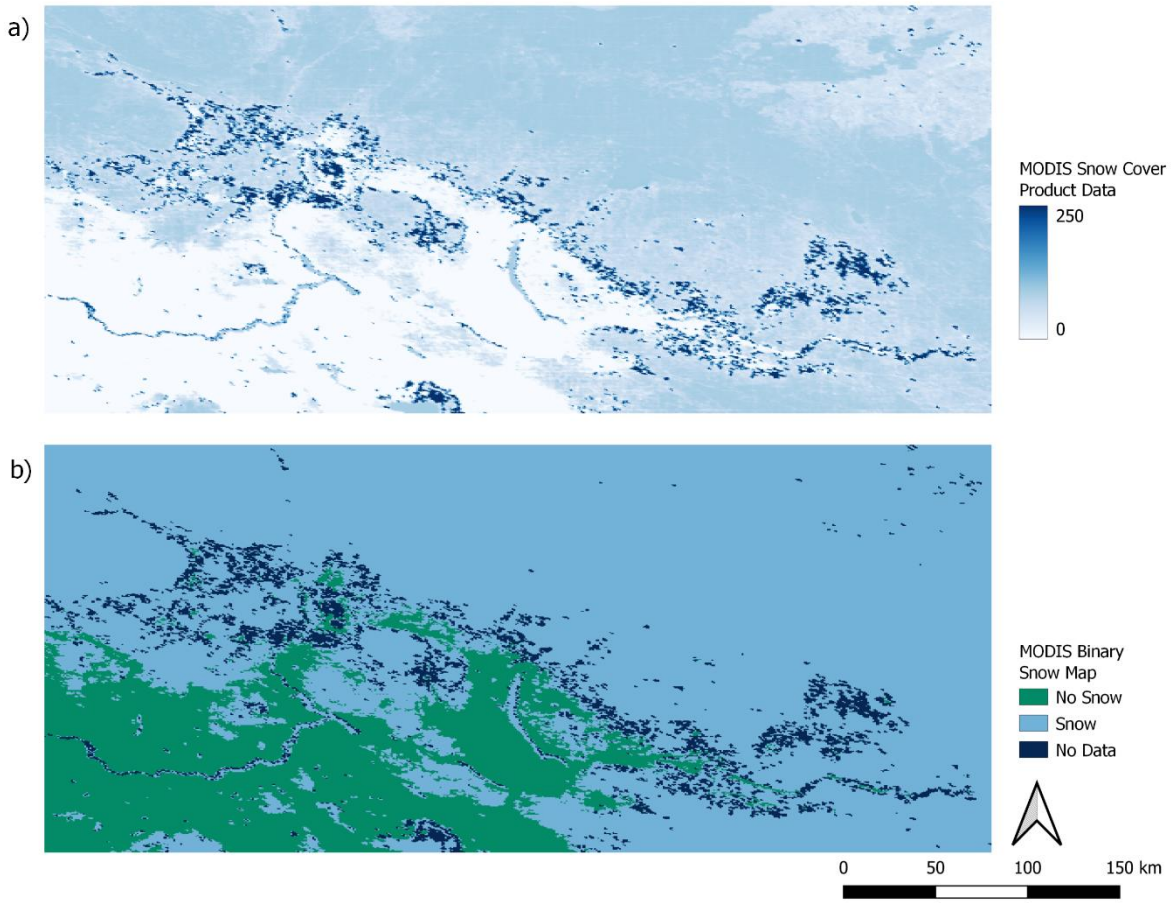

Figure 3. Saskatchewan Case Study Area: (a) Raw MODIS snow cover product data. (b) Final binary snow cover product data

Finally, the NDVI values per case study area were calculated using the MOD09GA data source. For this, the standard NDVI formula was used where MOD09GA Band 1 is the red band while MOD09GA Band 2 is the near-infrared (NIR) band. The formula used (*Equation 1*) resulted in an NDVI map for the case study areas with values between -1 and 1, with values closer to 1 indicating a higher vegetation presence.

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

Equation 1. NDVI

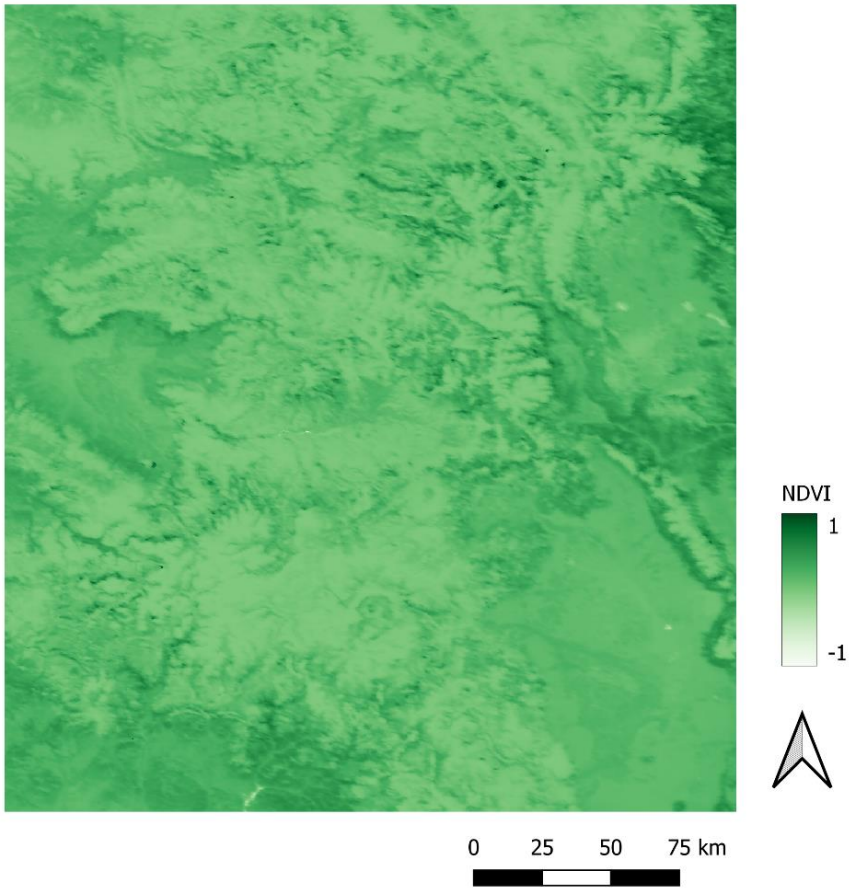

Figure 4. NDVI Values for the Colorado Case Study Area

Methodology

The first step in this study was calculating the threshold values in order to create binary snow cover maps from the preprocessed VIIRS data for all four case study areas. The thresholding in this study was implemented using the *Python* programming language and the *scikit-image* image processing plugin (Walt et al., 2014). Seven different thresholding algorithms were selected and applied to the preprocessed VIIRS data. During the calculations, values that were previously masked during the preprocessing phase were not taken into consideration. The seven thresholding algorithms that were used are Otsu (Otsu, 1979), Li (Li & Tam, 1998), Yen (Yen et al., 1995), Triangle (Zack et al., 1977), Minimum (Glasbey, 1993), Mean (Glasbey, 1993), and Isodata (Ridler & Calvard, 1978). While the thresholding algorithms were initially created for different purposes, like general computer vision, text identification, or microbiology, they have since been used in many fields, including remote sensing (Sekertekin, 2021).

The exact way the threshold is calculated varies slightly from algorithm to algorithm, but they are all based on the same basic principle. From the image, a histogram is created using all the values that were not masked during the preprocessing step. The resulting histogram is the basis of the thresholding calculation. Even though the different algorithms consider different aspects of the histogram and have different approaches to calculating the threshold, they still have the same output of one singular threshold value. All the values smaller or equal to the threshold are marked as not being snow covered, while all the values larger than the threshold are marked as snow covered.

Following this step, the binary snow cover maps that were made by thresholding were compared to the MODIS binary snow cover data for the same case study area. This was done by comparing the results on a pixel-by-pixel basis. All pixels were put into one of the four possible categories, except for areas where there was no available data in either of the two binary snow cover maps. The four potential values are true positive, true negative, false positive, and false negative. The different values indicate the relationship between the two binary snow cover maps. A visual representation of the four results and their relationship to the two binary snow cover models can be seen in *Table 2.*

	VIIRS Snow	VIIRS No Snow
MODIS Snow	True Positive	False Negative
MODIS No Snow	False Positive	True Negative

Table 2. Confusion Matrix

There are two distinct ways in which these results can be displayed, both of which serve a different purpose. One way is by calculating the percentage value for each category in the whole case study area, creating a confusion matrix. The confusion matrix is especially useful in order to determine the overall accuracy of the thresholding results in comparison to the MODIS snow cover product. Specifically, the higher the sum of the true positive and the true negative percentages, the more accurate the thresholding snow cover extent is. Besides the confusion matrix, the per pixel results can also be shown by their position on the map. This is especially important to determine which factors cause both

false negatives and false positives. By mapping their position, it is possible to perform a visual identification and determine the performance per discrete category on the basis of land cover and NDVI.

Splitting the results into discrete categories enables the possibility of determining if any particular characteristic impacts the overall accuracy. For this study, this entails the creation of confusion matrices per land cover category. This technique is based upon a study by Yang et al. (2015), where a similar approach of splitting the results per land cover type was used. If land cover has no or a limited impact on the results, the accuracy should be relatively consistent between different land covers. On the other hand, if land cover does impact the results, it is highly likely that the performance will be measurably worse with certain land covers.

Furthermore, because version five of the MODIS snow cover product treated areas with high NDVI values separately from regions with low NDVI values (Hall et al., 2006), this study also investigated that factor. In version five of the MODIS snow cover product, areas with high NDVI values were given a lower threshold in order to be considered snow-covered compared to areas with low NDVI values (Hall et al., 2006). If the impact of NDVI on VIIRS matches the impact it had on MOIDS, it is expected that the NDVI values of pixels marked as false negatives would be higher than those marked as true positives, indicating that the potential reason for the misclassification would be the NDVI. This factor was examined by looking at the distribution and mean values of NDVI of true positives and false negatives per case study area.

Finally, some of the false positive and false negative values might not be explainable using the methods outlined above; therefore, a visual analysis is also performed. While not giving nearly as conclusive results as the more analytical approaches, this step still enables the creation of possible hypotheses to explain some of the results and potentially create the basis for further research.

Results

The resulting confusion matrix of the seven different thresholding algorithms for the four case study areas can be seen in *Table 3*. The results are expressed as percentages, with the number representing the total fraction of the classification relative to the total number of all classified pixels. The Overall Accuracy indicates the sum of both the true positive and true negative values, while the Overall Error indicates the sum of both false positives and false negatives.

Colorado Confusion Matrix (in %)	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Otsu	28.42	31.32	0.14	40.13	59.73	40.27
Li	36.86	31.19	0.26	31.68	68.05	31.95
Yen	0.03	31.44	0.01	68.51	31.47	68.53
Triangle	7.98	31.42	0.04	60.57	39.40	60.60
Minimum	0.01	31.45	0.00	68.54	31.45	68.55
Mean	38.36	31.15	0.30	30.19	69.51	30.49
Isodata	29.73	31.30	0.16	38.82	61.02	39.98

Ontario Confusion Matrix (in %)	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Otsu	62.34	0.00	0.00	37.66	62.34	37.66
Li	69.17	0.00	0.00	30.83	69.17	30.83
Yen	77.91	0.00	0.00	22.09	77.91	22.09
Triangle	0.07	0.00	0.00	99.93	0.07	99.93
Minimum	0.00	0.00	0.00	100.00	0.00	100.00
Mean	54.75	0.00	0.00	45.25	54.75	45.25
Isodata	64.97	0.00	0.00	35.03	64.94	35.03

Alaska Confusion Matrix (in %)	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Otsu	19.17	40.25	0.22	40.36	59.42	40.58
Li	25.72	40.16	0.31	33.81	65.88	34.12
Yen	0.01	40.42	0.05	59.53	40.42	59.58
Triangle	0.20	40.39	0.08	59.33	40.59	59.41
Minimum	0.00	40.47	0.00	59.53	40.47	59.53
Mean	30.50	40.03	0.44	29.03	70.53	29.47
Isodata	23.38	40.19	0.27	36.15	63.57	36.43

Saskatchewan Confusion Matrix (in %)	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Otsu	55.10	22.26	0.17	22.47	77.36	22.64
Li	59.37	22.21	0.22	18.20	81.58	18.42
Yen	0.02	22.38	0.04	77.55	22.41	77.59
Triangle	0.15	22.34	0.08	77.42	22.50	77.50
Minimum	0.00	22.42	0.00	77.57	22.42	77.58
Mean	52.68	22.27	0.16	24.89	74.95	25.05
Isodata	55.10	22.26	0.17	22.47	77.36	22.64

Table 3. Confusion Matrix and Overall Accuracy and Error for the seven thresholding algorithms and the four case study areas.

Per case study area, the thresholding method with the best overall accuracy was selected in order to map the resulting confusion matrix. For the Colorado and Alaska case study areas, the Mean thresholding was used; for the Ontario Case Study Area, Yen thresholding was used; and finally, the Li thresholding results were used for the Saskatchewan case study area. The resulting four maps for the case study area show the geographic location of the True Positive, True Negative, False Positive, and False Negative values.

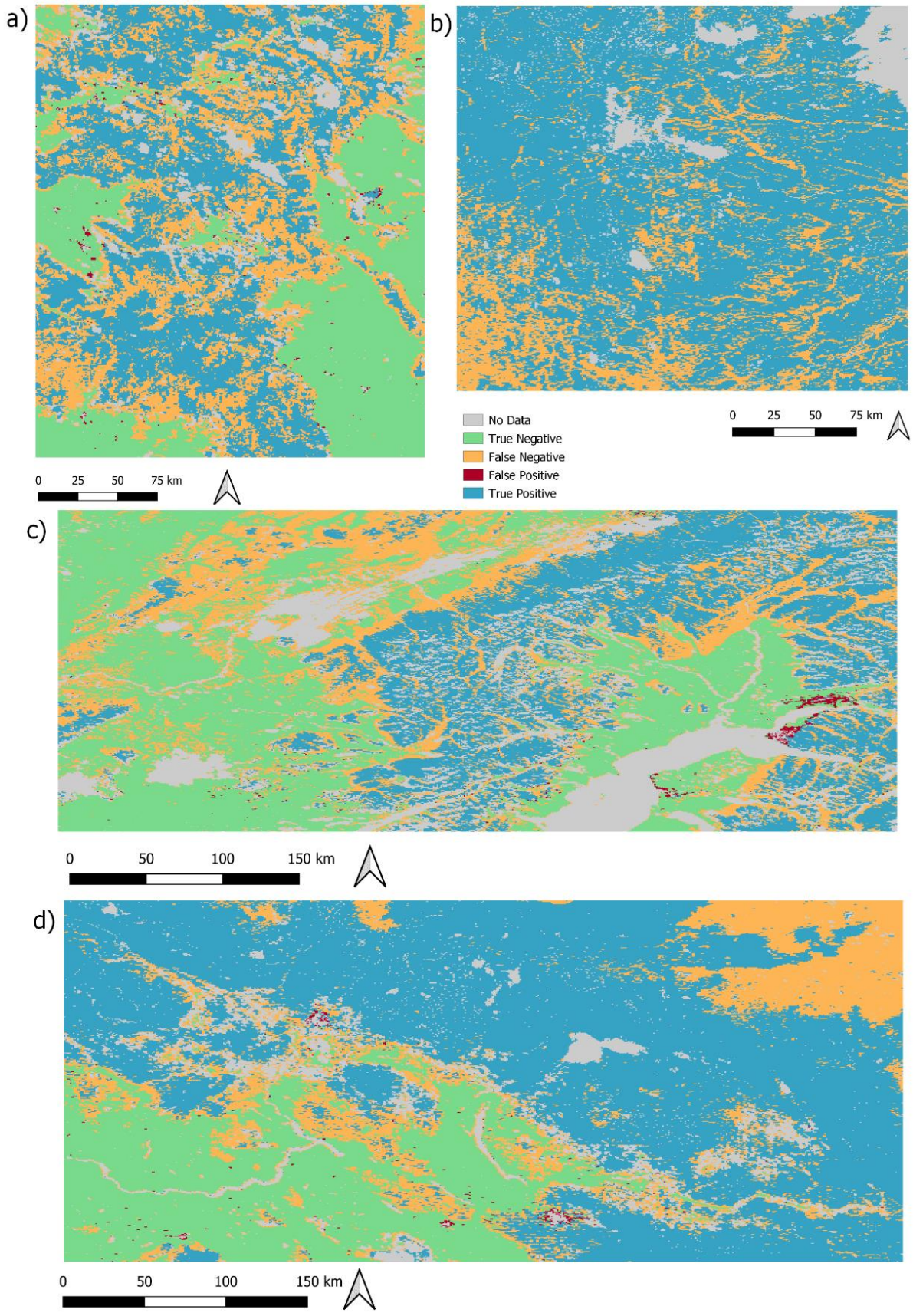

Figure 5. Maps of confusion matrices. (a) Colorado – Mean thresholding. (b) Ontario – Yen thresholding. (c) Alaska – Mean thresholding. (d) Saskatchewan – Li thresholding

The mapping of the confusion matrix values onto their geographic location also enabled the creation of discrete results per land cover type. For this, the maps from *Figure 5* were compared per land cover type of the Copernicus data for the appropriate year (Buchhorn et al., 2020a, 2020b). The resulting confusion matrices have the same six categories now split per land cover type. As the four case study areas cover different geographic regions, there are differences between the land cover types present in them. Besides that, some land cover types appeared relatively infrequently in any given case study area, and therefore any land cover with under 500 appearances was removed from the results.

Colorado	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Shrubs	25.45	58.43	0.73	15.39	83.88	16.12
Herbaceous vegetation	46.45	41.22	0.46	11.88	87.67	12.33
Cultivated and managed vegetation/agriculture (cropland)	7.72	86.68	0.64	4.95	94.40	5.60
Bare/sparse vegetation	73.37	24.09	0.13	2.37	97.50	2.50
Closed forest, evergreen needle leaf	22.81	15.33	0.08	61.77	38.15	61.85
Closed forest, deciduous broad leaf	70.13	0.11	0.03	29.74	70.24	29.76
Closed forest, unknown	43.42	20.74	0.14	35.69	64.16	35.84
Open forest, evergreen needle leaf	34.72	21.94	0.18	43.15	56.66	43.34
Open forest, deciduous broad leaf	72.95	0.07	0.00	26.99	73.01	26.99
Open forest, unknown	43.14	29.40	0.24	27.22	75.86	24.14

Ontario	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Shrubs	85.81	0.00	0.00	14.19	85.81	14.19
Herbaceous vegetation	94.17	0.00	0.00	5.83	94.17	5.83
Herbaceous wetland	97.09	0.00	0.00	2.91	97.09	2.91
Closed forest, evergreen needle leaf	36.94	0.00	0.00	63.06	36.94	63.06
Open forest, evergreen needle leaf	63.92	0.00	0.00	36.08	63.92	36.08
Open forest, unknown	74.73	0.00	0.00	25.27	74.73	25.27

Alaska	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Shrubs	20.65	56.49	0.25	22.61	77.14	22.86
Herbaceous vegetation	48.50	33.71	0.31	17.48	82.21	17.79
Bare/sparse vegetation	84.51	1.14	0.28	14.06	85.66	14.34
Snow and Ice	78.48	0.18	0.03	21.30	78.67	21.33
Herbaceous wetland	3.37	82.83	2.28	11.52	86.20	13.80
Closed forest, evergreen needle leaf	6.13	38.31	0.35	55.22	44.44	55.56
Closed forest, mixed	0.95	69.95	2.47	26.63	70.90	29.10
Closed forest, unknown	6.66	56.01	0.74	36.59	62.67	37.33
Open forest, evergreen needle leaf	11.60	31.55	0.91	55.93	43.16	56.84
Open forest, unknown	12.24	58.13	0.45	29.18	70.37	29.63

Saskatchewan	True Positive	True Negative	False Positive	False Negative	Overall Accuracy	Overall Error
Herbaceous vegetation	48.16	37.93	0.59	13.33	86.08	13.92
Cultivated and managed vegetation/agriculture(cropland)	67.79	20.76	0.13	11.31	88.55	11.45
Herbaceous wetland	62.41	20.66	0.18	16.75	83.07	16.93
Closed forest, evergreen needle leaf	6.43	0.08	0.00	93.50	6.50	93.50
Closed forest, deciduous broad leaf	29.14	0.22	0.00	70.63	29.37	70.63
Closed forest, mixed	3.46	0.07	0.00	96.48	3.52	96.48
Closed forest, unknown	35.92	5.77	0.00	58.31	41.69	58.31
Open forest, evergreen needle leaf	57.12	0.92	0.00	41.96	58.04	41.96
Open forest, deciduous broad leaf	58.94	0.26	0.00	40.80	59.20	40.80
Open forest, unknown	60.25	7.63	0.08	32.03	67.88	32.12

Table 4. Confusion matrix per land cover type for the four case study areas

Besides land use, NDVI values were also investigated in this study. As previous snow extent models that were using daytime data had a lower threshold for areas with high NDVI, this study also investigated this factor. If areas that were marked as false negative have a higher NDVI than true positives, this may indicate that the thresholding value should be lower for the areas with higher NDVI. In order to determine if the NDVI values are higher, two boxplots per case study area were created, one for true positives and one for false negatives. The boxplots indicate mean values, the first and third quantile of the data, and maximum values making analysis relatively straightforward.

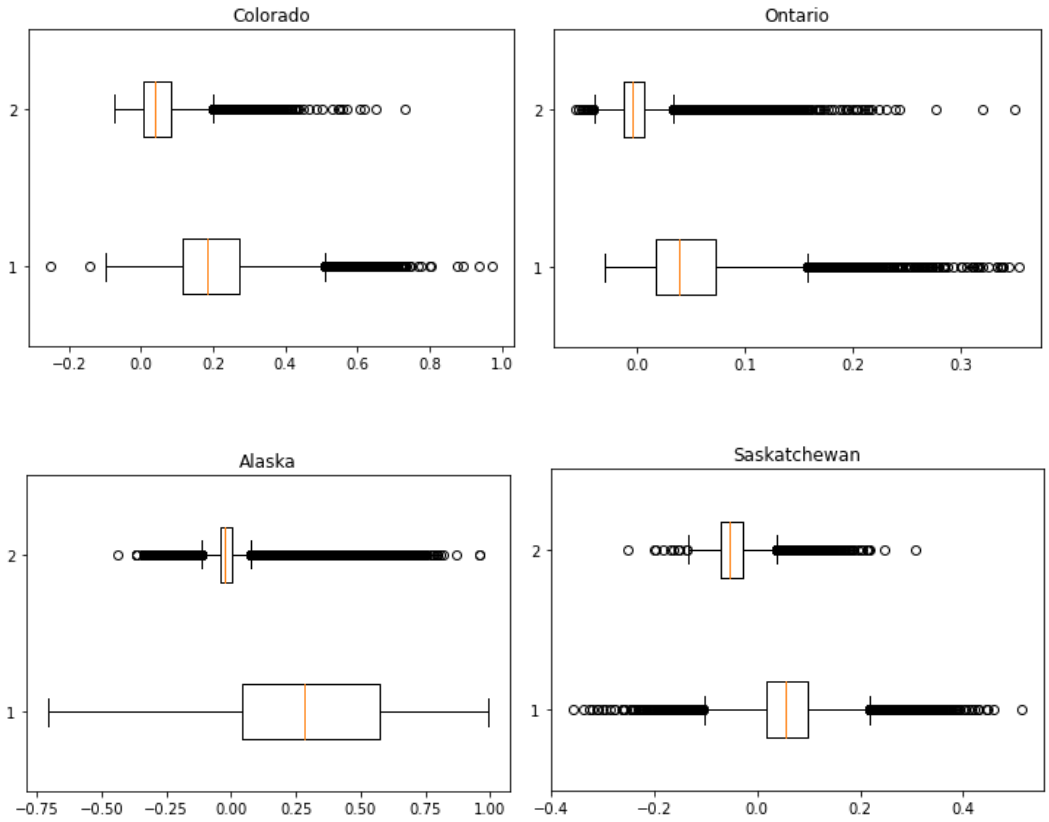

Figure 6. Boxplots of NDVI values per case study area. NDVI values for true positives (2) and false negatives (1)

Finally, because false positives make up less than 1% of all results, it is relatively difficult to make any statistical explanation for them due to their small number. Therefore, to explain them, a visual identification was attempted. For this, especially notable was the spatial distribution of the false positives, which in most cases appear to be in the proximity of urban/built-up areas. Examples of this phenomenon can be seen in *Figure 7*, where the built/up areas from the Copernicus data are presented as well as false positives. The Ontario case study area had no false positives, and therefore it is not included in this part.

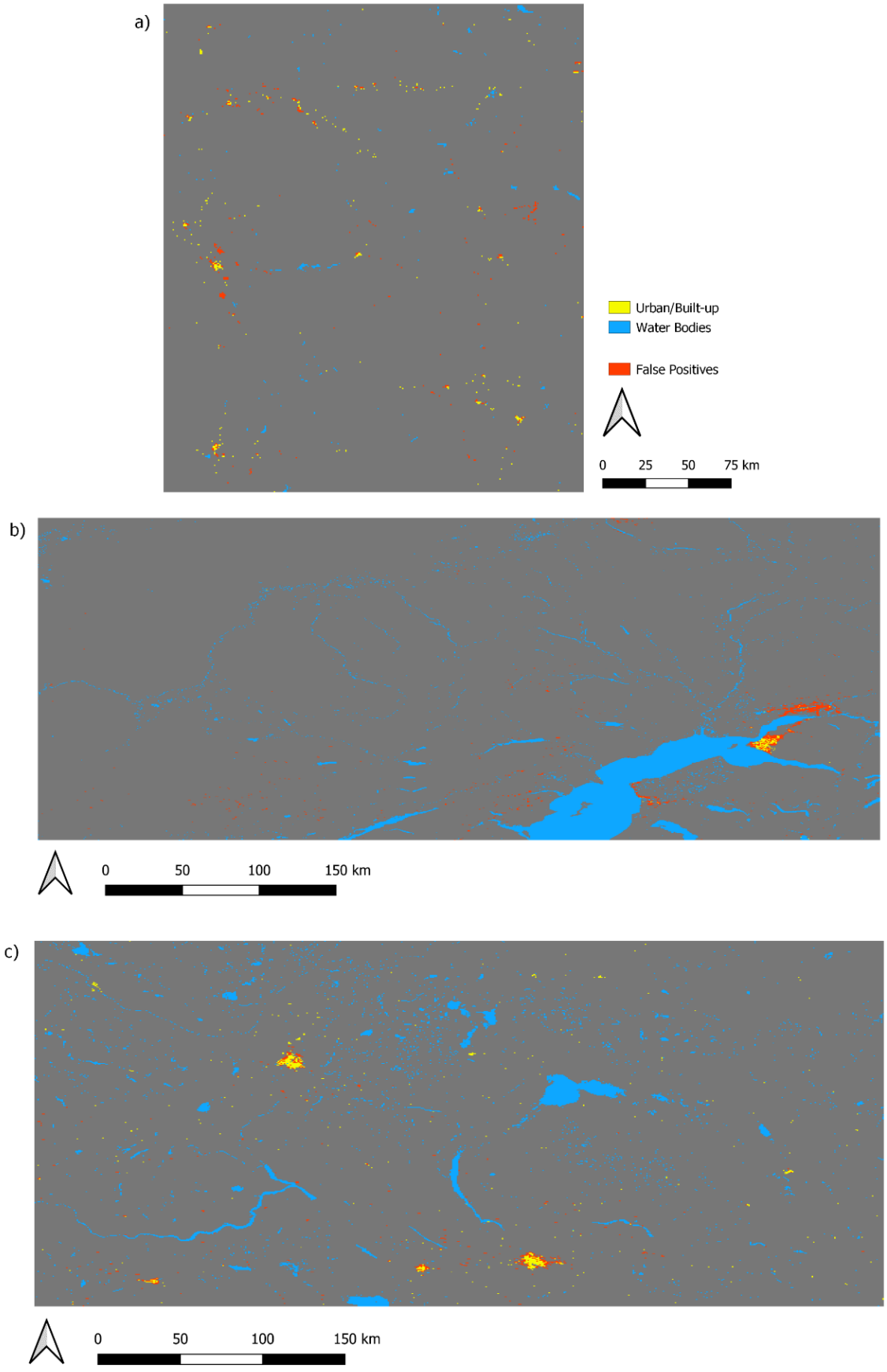

Figure 7. False Positive Values and Urban / Built-up Areas in the Colorado (a), Alaska (b), and Saskatchewan (c) case study areas

Discussion

The seven thresholding algorithms investigated in this study had notably different performances. The overall accuracy varied for each algorithm and between the four case study areas, but some overall conclusions can still be drawn from them. The overall accuracy of the algorithms had a quite wide spread of values, with the lowest overall accuracy being 0%, while the highest overall accuracy was 81.58%.

The two thresholding algorithms that performed the worst were Triangle thresholding and Minimum thresholding, both of which had their highest overall score at only around 40%. Similarly, Yen thresholding also had results of less than or approximately 40%, except in the Ontario case study area, where it performed the best with a 77.91% overall accuracy. Otsu and Isodata thresholding had mediocre performances in all four case study areas, with none of them being the best performing one in any case study area. Their overall accuracies ranged between 59% and 77%. Another algorithm, Mean thresholding, had a similar range of overall accuracy, with its lowest result being 54.75% and the highest being 74.95%. However, what is notable is that Mean thresholding was the best performing algorithm in both the Colorado and Alaska case study areas. However, it underperformed in the Ontario case study area, where it achieved only a 54.75% overall accuracy. These results indicate that Mean thresholding is the best performing algorithm in mountainous regions, as both the Colorado and Alaska case study areas are predominantly mountainous.

While Mean thresholding performed the best in mountainous areas, the algorithm with the best range of overall accuracy and the highest individual accuracy in the Saskatchewan case study area was Li thresholding. This algorithm always had an overall accuracy above 65% and was either the highest or second highest result in every case study area. Therefore, even though Mean thresholding did have the best result in 2 out of the 4 case study areas, it does not necessarily mean it is the best choice in all circumstances. The fact that Mean thresholding measurably underperformed in the Ontario case study area means that in order to ensure the best overall and consistently high overall accuracy, the Li thresholding algorithm is the best choice of all seven algorithms tested in this study. If the thresholding

is applied exclusively to mountainous areas, using Mean thresholding could yield better results. However, if various terrains are present, Li thresholding will give good overall and consistent results.

Determining the best thresholding algorithm is not the only conclusion that can be drawn from the results. What is especially noteworthy is the difference between false positives and false negatives. The percentage of false negatives is significantly larger than the false positives for all thresholding algorithms and in all case study areas. In fact, false positives in every case study area and thresholding algorithm make up less than 0.5%. This indicates that the vast majority of the overall error comes from the underestimation of the snow extent by the VIIRS DNB snow extent model when compared to the MODIS snow cover product.

The causes of this underestimation and the locations where it occurs can be potentially explained by the land cover results. While the four case study areas had different land cover types present, some overall trends are still noticeable. The two land covers which had the lowest overall accuracy in all case study areas are open and closed evergreen needle leaf forests. Besides those two, all other types of forests, when present in the case study area, had lower overall accuracy than non-forest land covers like shrubs or herbaceous vegetation. When it comes to the different forest types, as already mentioned, evergreen needle leaf has the lowest overall accuracy, but when comparing open and closed forests, closed forests generally have lower overall accuracy. In all cases, the low overall accuracy can be entirely explained by false negatives, with false positives being rare or nonexistent for most forests. The remaining non-forested areas performed well with high true positive and true negative values, resulting in high overall accuracy. The false negatives were mainly in the 10% - 25% range, and false positives were present but usually made up less than 1%.

The effect that forests have on the accuracy of snow cover extent is not unexpected. Forests have been a factor that limited the accuracy of snow cover extent estimates in many different products (Dietz et al., 2012). The presence of tree canopies would effectively block the view of the ground and, therefore, any snow that is present, causing the underestimations (Dietz et al., 2012). Furthermore, while it is possible that the canopy itself would capture some of the snow and therefore increase its albedo, studies suggest that even when that does occur, the overall albedo is still lower (Webster & Jonas, 2018).

The differences between closed and open forests would also be explained mainly by the canopies. In the Copernicus land cover data, closed forests are all those that have more than 70% of the ground covered by canopies, while open forests are those that have 15%-70% canopy cover (Buchhorn et al., 2020c). Since in closed forests, the canopies would cover more of the view of the ground and any snow present there, they would have more false negative estimates than open forests, where more of the ground is visible.

The impact of canopies would also explain the lower overall accuracy of evergreen forests compared to other forest types. In most cases, the evergreen forests performed roughly 20-30% worse than other types of forests. Since the images were captured during March and October, during which evergreen trees would still retain their full canopy while deciduous trees would have lost most of their canopy. The lack of canopies in deciduous forests would mean that the ground, and therefore any potential snow on the ground, would be more visible and therefore easier to detect.

Another way in which this effect could possibly be measured is by investigating the NDVI. As tree canopies, and especially evergreen trees, would have a high NDVI value at the time the data was captured, it would be expected that the NDVI values of the false negatives would skew higher than the true positives. As can be seen from the box plots in *Figure 6*, the mean, quantiles, and maximum values of the NDVI are higher for false negatives than for true positives. This would indicate that higher NDVI would be correlated with false negative identification. This factor, combined with the previous discussion of the impact of forest canopies, would suggest that high density and especially evergreen forests have the most negative impact on identification, with their low albedo causing the areas to be classified as not covered in snow.

Even though the majority of false negatives did occur in forested areas, other land covers also had false negatives. The distribution and cause of these can not be explained by tree canopies; therefore, something else has to be their cause. The two different false negatives can be best be seen in the Saskatchewan case study area, pictured in *Figure 8*. While the false negative values in the north and northeast can be explained by the forest present there, the false negative in the center and south-west are present in a mix between cropland and herbaceous vegetation.

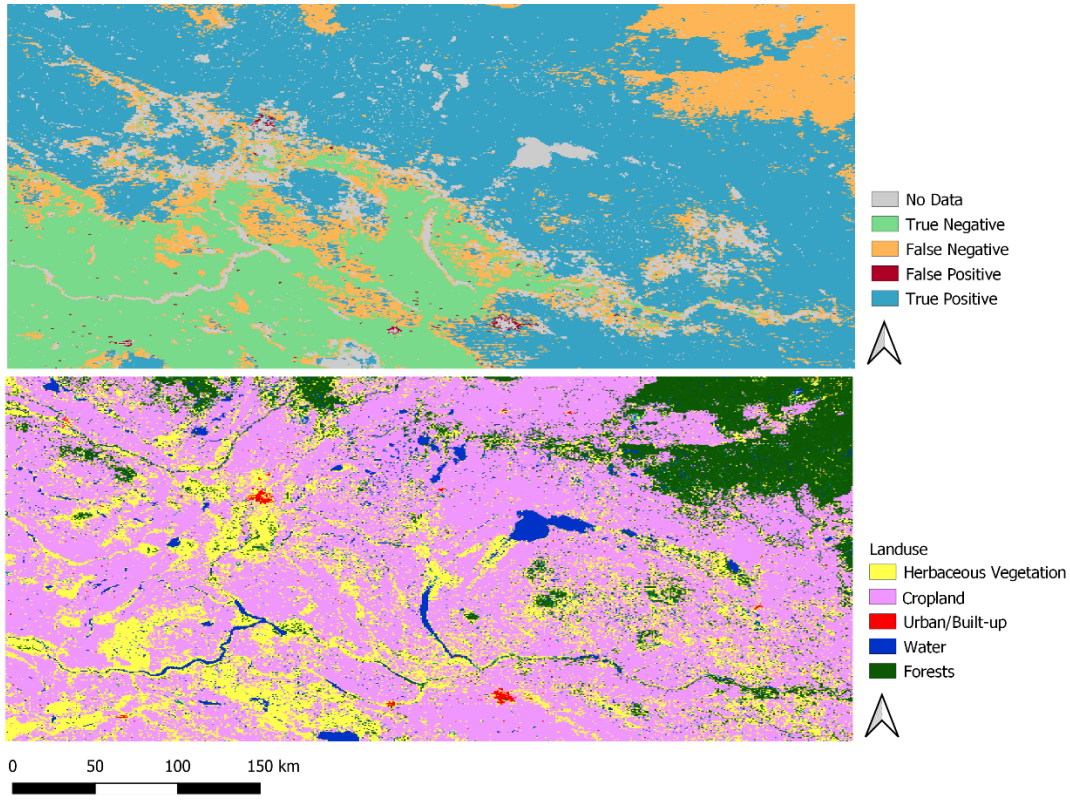

Figure 8. Map of confusion matrix for Saskatchewan case study area (top) and Land use for Saskatchewan case study area (bottom)

The cause of the false negatives could be explained by their location. Because these areas mark the transition between snow and no snow, they may have relatively shallow snow depth. Studies have shown that even during the day, the accuracy of remote sensing of snow decreases with lower snow depth (Yang et al., 2015). Therefore, nighttime snow detection accuracy may also be susceptible to underestimations with low snow depth, explaining the underestimates in those regions. The snow depth could have also slightly changed in the period between when the MODIS and VIIRS DNB data was captured, which could explain some of the discrepancies between them. While in the vast majority of non-forested areas, the VIIRS DNB data perform well, the accuracy of the boundaries will only become apparent with an in situ study, which could analyze the impact of snow depth on the accuracy.

Even though they make up the majority, false negatives were not the only type of wrong estimates. While making up a notably smaller amount of the overall error, false positives are still present and worth investigating. Due to their low numbers, it is relatively difficult to conclude anything statistically significant about them; therefore, a visual inspection of the false negatives was performed.

The spatial distribution, which can be seen in *Figure 7*, clearly shows that the distribution of the false negatives is closely related to urban/built-up areas. The connection between these two can especially be seen near some of the larger built-up areas present in the case study areas like Anchorage in Alaska or Saskatoon and Regina in Saskatchewan. While the specific built-up areas were masked out during the data preparation step, especially so that their presence does not impact the calculations, it is clear that the surrounding areas are also much brighter due to human activity.

This phenomenon is not entirely unexpected, as studies have suggested that suburban and rural areas do emit more artificial light toward the horizon, meaning that the surrounding area would also be brighter because of that (Tong et al., 2020). Besides those false positives obviously near built-up areas, there are some which are not close to any pixels marked as such. Many of these false positives do appear to still be due to human activity, but due to the nature of the land cover data, they were not marked as built-up areas. This is primarily the case for smaller communities, like the town of Beechy in Saskatchewan, which due to its small size and population, got marked as cropland, which completely surrounds it. While being relatively small, it is clear that there is still enough artificial light emitted there to skew the results into being false positives. Other areas in the three case study areas with false positive results follow the same pattern. While this approach does seem to explain the false positives, it still remains only a visual identification of the pattern, and in order to determine if it is the main reason behind it, a more rigorous and analytical study would be necessary. One way to potentially avoid these false positives is by creating a mask layer not based on land use but instead on light. Since the NASA Black Marble product suite releases monthly and annual data, it could be used to create a mask layer as all the brightest areas would be due to human activity (Román et al., 2018).

Naturally, this study has certain limitations which could have some impact on the results. The main limitation is that the VIIRS snow cover extent is compared to another snow cover model and not to in situ measurements. While the selected MODIS snow cover product is not perfect, and like any other model, it has its own known limitations (Riggs et al., 2019), it is still a widely used model, and agreement between the models would still indicate that it is acceptable for many studies in hydrology, climate science, or other areas where MODIS data is commonly used (Da Ronco et al., 2020).

There are other factors that can be examined in subsequent studies that could further the potential of VIIRS DNB data for snow cover extent calculations. The main potential lies in investigating other automated thresholding algorithms. While this study examined seven different algorithms, there are multiple different ones that were not included in this study that could perform even better.

Other avenues for further research also include trying to increase the accuracy in evergreen or all forested areas. Since other daytime snow cover extent products have dealt with similar issues before (Dietz et al., 2012), improving the results in those areas could be possible. This improvement could be achieved by lowering the threshold for areas with high NDVI or by calculating a separate threshold for forested areas.

Finally, since future Joint Polar Satellite System (JPSS) missions are already planned, and will be equipped with VIIRS, the data availability in the future will only increase (Levin et al., 2020). Hopefully, this will reduce the impact cloud coverage has on VIIRS DNB data as multiple satellites will be capturing the data, enabling even better coverage in the future.

Conclusion

Identifying snow cover extent using VIIRS DNB data is possible and achievable with relatively high overall accuracy. The false estimates are usually underestimations, with overestimations present but usually limited to areas near human activity. The false negatives are usually present in forested areas, especially in closed and evergreen forests. Some false negatives also occur on the border between areas covered and not covered by snow. While this could impact the usefulness of the VIIRS DNB data for some purposes, this is an area where the results could be improved with in situ research.

Of the seven tested thresholding algorithms, Li thresholding achieved the most consistently high overall accuracy in all case study areas, while the Mean thresholding algorithm performed the best in mountainous regions, but it did not have high overall accuracy in other areas. While some limitations are present, the high overall accuracy and the distribution of accurate identification should still prove useful to fill in some spatial or temporal data gaps for snow cover extent maps, especially in regions with few forests.

Bibliography

Anisimov, O. A., Vaughan, D. G., Callaghan, T., Furgal, C., Marchant, H., Prowse, T. D., Vilhjálmsson, H., Walsh, J. E., Christensen, T. R., Forbes, D. L., Nelson, F. E., Nuttall, M., Reist, J. D., Rose, G. A., Vandenberghe, J., Barry, R., Jefferies, R., & Stone, J. (2007). *Polar regions (Arctic and Antarctic)* (Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change).

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309. https://doi.org/10.1038/nature04141

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Herold, M., & Fritz, S. (2020a). *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2018: Globe* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3518038

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Herold, M., & Fritz, S. (2020b). *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3939050

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Li, L., & Tarko, A. (2020). *Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual* (Dataset v3.0, doc issue 3.3). Zenodo. https://doi.org/10.5281/ZENODO.3938963

Da Ronco, P., Avanzi, F., De Michele, C., Notarnicola, C., & Schaefli, B. (2020). Comparing MODIS snow products Collection 5 with Collection 6 over Italian Central Apennines. *International Journal of Remote Sensing*, *41*(11), 4174–4205. https://doi.org/10.1080/01431161.2020.1714778

Dietz, A. J., Kuenzer, C., Gessner, U., & Dech, S. (2012). Remote sensing of snow – a review of available methods. *International Journal of Remote Sensing*, *33*(13), 4094–4134. https://doi.org/10.1080/01431161.2011.640964

Foster, J. L. (1983a). *Observations Of The Earth Using Nighttime Visible Imagery* (G. W. Aitken, Ed.; pp. 187–193). https://doi.org/10.1117/12.935885

Foster, J. L. (1983b). Night-time observations of snow using visible imagery. *International Journal of Remote Sensing*, *4*(4), 785–791. https://doi.org/10.1080/01431168308948597

Foster, J. L., & Hall, D. K. (1991). Observations of snow and ice features during the polar winter using moonlight as a source of illumination. *Remote Sensing of Environment*, *37*(2), 77–88. https://doi.org/10.1016/0034-4257(91)90020-7

Glasbey, C. A. (1993). An Analysis of Histogram-Based Thresholding Algorithms. *CVGIP: Graphical Models and Image Processing*, *55*(6), 532–537. https://doi.org/10.1006/cgip.1993.1040

Groisman, P. Ya., Karl, T. R., & Knight, R. W. (1994). Observed Impact of Snow Cover on the Heat Balance and the Rise of Continental Spring Temperatures. *Science*, *263*(5144), 198–200. https://doi.org/10.1126/science.263.5144.198

Hall, D. K., & Riggs, G. A. (2021). *MODIS/Aqua Snow Cover Daily L3 Global 500m Grid, Version 61* [Data set]. NASA National Snow and Ice Data Center DAAC. https://doi.org/10.5067/MODIS/MOD10A1.061

Hall, D. K., Riggs, G. A., & Salomonson, V. V. (2006). *MODIS/Terra Snow Cover 5-Min L2 Swath 500m, Version 5* [Data set]. NASA National Snow and Ice Data Center DAAC. https://doi.org/10.5067/ACYTYZB9BEOS

Huang, Y., Liu, H., Yu, B., Wu, J., Kang, E. L., Xu, M., Wang, S., Klein, A., & Chen, Y. (2018). Improving MODIS snow products with a HMRF-based spatio-temporal modeling technique in the Upper Rio Grande Basin. *Remote Sensing of Environment*, *204*, 568–582. https://doi.org/10.1016/j.rse.2017.10.001

Huang, Y., Song, Z., Yang, H., Yu, B., Liu, H., Che, T., Chen, J., Wu, J., Shu, S., Peng, X., Zheng, Z., & Xu, J. (2022). Snow cover detection in mid-latitude mountainous and polar regions using nighttime light data. *Remote Sensing of Environment*, *268*, 112766. https://doi.org/10.1016/j.rse.2021.112766

LAADS DAAC. (2018, 2019). *VIIRS/NPP Day/Night Band 6-Min L1 Swath SDR 750m—LAADS DAAC*. https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/NPP_VDNES_L1#overview

Lee, T. E., Miller, S. D., Turk, F. J., Schueler, C., Julian, R., Deyo, S., Dills, P., & Wang, S. (2006). The NPOESS VIIRS Day/Night Visible Sensor. *Bulletin of the American Meteorological Society*, *87*(2), 191–200. https://doi.org/10.1175/BAMS-87-2-191

Levin, N. (2017). The impact of seasonal changes on observed nighttime brightness from 2014 to 2015 monthly VIIRS DNB composites. *Remote Sensing of Environment*, *193*, 150–164. https://doi.org/10.1016/j.rse.2017.03.003

Levin, N., Kyba, C. C. M., Zhang, Q., Sánchez de Miguel, A., Román, M. O., Li, X., Portnov, B. A., Molthan, A. L., Jechow, A., Miller, S. D., Wang, Z., Shrestha, R. M., & Elvidge, C. D. (2020). Remote sensing of night lights: A review and an outlook for the future. *Remote Sensing of Environment*, *237*, 111443. https://doi.org/10.1016/j.rse.2019.111443

Levis, S., Bonan, G. B., & Lawrence, P. J. (2007). Present-day springtime high-latitude surface albedo as a predictor of simulated climate sensitivity. *Geophysical Research Letters*, *34*(17). https://doi.org/10.1029/2007GL030775

Li, C. H., & Tam, P. K. S. (1998). An iterative algorithm for minimum cross entropy thresholding. *Pattern Recognition Letters*, *19*(8), 771–776. https://doi.org/10.1016/S0167-8655(98)00057-9

Liu, D., Zhang, Q., Wang, J., Wang, Y., Shen, Y., & Shuai, Y. (2021). The Potential of Moonlight Remote Sensing: A Systematic Assessment with Multi-Source Nightlight Remote Sensing Data. *Remote Sensing*, *13*(22), 4639. https://doi.org/10.3390/rs13224639

Miller, S. D., Lee, T. F., Turk, F. J., Kuciauskas, A. P., & Hawkins, J. D. (2005). Shedding new light on nocturnal monitoring of the environment with the VIIRS day/night band. *Atmospheric and Environmental Remote Sensing Data Processing and Utilization: Numerical Atmospheric Prediction and Environmental Monitoring*, *5890*, 201–209. https://doi.org/10.1117/12.619534

Miller, S. D., Straka, W., Mills, S. P., Elvidge, C. D., Lee, T. F., Solbrig, J., Walther, A., Heidinger, A. K., & Weiss, S. C. (2013). Illuminating the Capabilities of the Suomi National Polar-

Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band. *Remote Sensing*, *5*(12), 6717–6766. https://doi.org/10.3390/rs5126717

Musselman, K. N., Clark, M. P., Liu, C., Ikeda, K., & Rasmussen, R. (2017). Slower snowmelt in a warmer world. *Nature Climate Change*, *7*(3), 214–219. https://doi.org/10.1038/nclimate3225

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Ridler, T. W., & Calvard, S. (1978). Picture Thresholding Using an Iterative Selection Method. *IEEE Transactions on Systems, Man, and Cybernetics*, *8*(8), 630–632. https://doi.org/10.1109/TSMC.1978.4310039

Riggs, G. A., Hall, D. K., & Román, M. O. (2019). *MODIS Snow Products Collection 6.1 User Guide*. 66.

Román, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., Schultz, L., Bell, J., Stokes, E. C., Pandey, B., Seto, K. C., Hall, D., Oda, T., Wolfe, R. E., Lin, G., Golpayegani, N., Devadiga, S., Davidson, C., Sarkar, S., … Masuoka, E. J. (2018). NASA's Black Marble nighttime lights product suite. *Remote Sensing of Environment*, *210*, 113–143. https://doi.org/10.1016/j.rse.2018.03.017

Sekertekin, A. (2021). A Survey on Global Thresholding Methods for Mapping Open Water Body Using Sentinel-2 Satellite Imagery and Normalized Difference Water Index. *Archives of Computational Methods in Engineering*, *28*(3), 1335–1347. https://doi.org/10.1007/s11831-020-09416-2

Tong, K. P., Kyba, C. C. M., Heygster, G., Kuechly, H. U., Notholt, J., & Kolláth, Z. (2020). Angular distribution of upwelling artificial light in Europe as observed by Suomi–NPP satellite. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *249*, 107009. https://doi.org/10.1016/j.jqsrt.2020.107009

Vermote, E., & Wolfe, R. (2021). *MODIS/Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V061* [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD09GA.061

Walt, S. van der, Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: Image processing in Python. *PeerJ*, *2*, e453. https://doi.org/10.7717/peerj.453

Webster, C., & Jonas, T. (2018). Influence of canopy shading and snow coverage on effective albedo in a snow-dominated evergreen needleleaf forest. *Remote Sensing of Environment*, *214*, 48–58. https://doi.org/10.1016/j.rse.2018.05.023

Wiesnet, D. R., Mcginnis, J., Matson, M., & Pritchard, J. A. (1981). *Evaluation of hcmm satellite data for estuarine tidal circulation patterns and thermal inertia soil moisture measurements. Interim; final report* (E-82-10064). National Oceanic and Atmospheric Administration, Washington, DC (USA). https://www.osti.gov/biblio/6558581

Yang, J., Jiang, L., Ménard, C. B., Luojus, K., Lemmetyinen, J., & Pulliainen, J. (2015). Evaluation of snow products over the Tibetan Plateau. *Hydrological Processes*, *29*(15), 3247–3260. https://doi.org/10.1002/hyp.10427

Yen, J.-C., Chang, F.-J., & Chang, S. (1995). A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*, *4*(3), 370–378. https://doi.org/10.1109/83.366472

Yin, D., Cao, X., Chen, X., Shao, Y., & Chen, J. (2013). Comparison of automatic thresholding methods for snow-cover mapping using Landsat TM imagery. *International Journal of Remote Sensing*, *34*(19), 6529–6538. https://doi.org/10.1080/01431161.2013.803631

Zack, G. W., Rogers, W. E., & Latt, S. A. (1977). Automatic measurement of sister chromatid exchange frequency. *Journal of Histochemistry & Cytochemistry*, *25*(7), 741–753. https://doi.org/10.1177/25.7.70454