Moose Lake

Algae Bloom Detection using LANDSAT data 1984-2017

Eric Dilligeard, Ph.D.

10/18/2018

Acknowledgement: I would like to thank Dr. Dörte Köster for her insight in statistical analysis and for her review of the report.

TABLE OF CONTENTS

INTRODUCTION

Algal blooms are a major water quality concern for residents and visitors of Moose Lake in north-eastern Alberta. The Moose Lake Watershed Society, a group of concerned citizens working towards the vision of a healthy and functioning watershed, wishes to better understand patterns in algal blooms at the lake. The objective of this project was to look for trends over time in algal blooms in Moose Lake using available satellite data.

METHODOLOGY

DATASETS

We used 30 m-spatial resolution Landsat5 TM, Landsat7 ETM+ and Landsat8 OLI data (table 1) to identify algal blooms in Moose Lake. This approach was recently successfully applied to Lake Erie (Ho *et al.*, 2017) and to Pigeon Lake (Eric Dilligeard and Dörte Köster, 2007). We gathered all available pseudo color Landsat images (figure 1) and associated Quality Assurance (QA) products of Moose Lake, acquired between June and October (915 images from 1984 to 2017, courtesy of the U.S. Geological Survey). We performed a preliminary data selection using the QA products to remove images with more than 90% cloud cover over the lake as these images can't be used for our analysis. We then downloaded the Top Of Atmosphere Reflectance (TOAR) and QA products (courtesy of the U.S. Geological Survey) for the 499 Landsat images with less than 90% cloud cover. We later used the more reliable QA products associated with the TOAR to further filter images on cloud cover.

BLOOM DETECTION

The TOAR were used to generate the Near Infrared with Simple Atmospheric Correction (NIR with SAC) products that had been used for tracking Lake Erie blooms (Ho *et al.*, 2017). This product is the result of a per-pixel basis algorithm based on the Near Infrared band (NIR) signal and corrected for the atmosphere effect with a simple approach using the Shortwave Infrared Band (SWIR1):

$$NIR\ with\ SAC = NIR - 1.03 * SWIR1 \quad (1)$$

Where NIR is the near infrared band and SWIR1 is the Shortwave Infrared channel 1

Ho *et al.*, 2017 optimized the coefficient of 1.03 used for the atmospheric correction for the local conditions of their study, which may not be optimal for Moose Lake. Optimizing the coefficient for Moose Lake was outside the scope of this study; we therefore used the same value of 1.03.

Ho *et al.*, 2017 used a threshold of > 0.0235 on the NIR with SAC algorithm to identify blooms on Landsat data. They optimized this threshold using the MERIS Cyanobacterial Index (CI) product (Stumpf *et al.*, 2012) acquired from Lake Erie as well. They maximized the number of LANDSAT pixels classified as blooms to match the corresponding MERIS CI bloom classification. Again, optimizing this threshold for the Moose Lake conditions was outside the scope of this study; we therefore used the same threshold of 0.0235.

We consulted three additional pieces of information to increase confidence in bloom detections. First, we used the Quality Assurance (QA) index provided by USGS to classify each pixel as cloudy or not. Second, we looked at each individual pseudo-color image to see where clouds and their shadows were located. Third, we implemented the Hue algorithm defined by Ho *et al.*, 2017. We used this algorithm to confirm blooms previously identified by NIR with SAC method based on the dominance of the green color.

$$Hue = \begin{cases} \frac{g-b}{r+g-2*b} & if\ b = min(r,g,b) \\ \frac{b-r}{g+b-2*r}+1 & if\ r = min(r,g,b) \\ \frac{r-g}{b+r-2*g}+2 & if\ g = min(r,g,b) \end{cases} \quad (2)$$

Where b, g and r are blue, green and red Landsat bands, respectively.

For the Hue algorithm, we used the threshold of <1.6 as implemented by Ho *et al.*, to confirm a bloom.

We selected the Moose Lake areas with a depth larger than 3.2 m for this study to reduce the number of misclassifications caused by aquatic plants or by the bottom of the lake that were sometimes visible in the shallow water and misdiagnosed as algal blooms. For this purpose we used depth contours provided by AEP. The resulting area of study is presented in figure 2. The study area represents about 50 % of the lake surface.

Overall the NIR with SAC and Hue algorithm provided a clear identification of blooms. We looked at pixels that were mistakenly classified as ice in the QA product. We concluded that they were pixels with a very intense signal caused by clouds. Therefore ice pixels were classified as cloud for the time period from June to October.

This visual inspection of the images was also useful to rule out some false positive bloom identifications caused by the presence of shadows. This manual inspection allowed us to be less strict on the automatic filtering on cloud cover percentage and to keep more images for the analysis. Therefore we fixed the maximum cloud cover to 75%. It resulted in a total of 444 usable images out of 915, with 17 out of a total of 170 months (10%) without any usable image (table2).

TEMPORAL TRENDS IN BLOOM DETECTION

The chance of detecting a bloom was uneven among years and months (table 2) because of the highly variable number of usable data. This highly variable number was due to the variability of the cloud cover and due to the increased number of operational satellites after 1999. Therefore, a temporal analysis of the total number of bloom events among years would be biased, with a larger chance of detecting blooms in recent years due to a larger number of available imagery.

We used a statistical approach based on trend tests to deal with this potential bias. We calculated the percentage of images with identified blooms out of the total number of usable imagery (cloud cover < 75%) per month and year. We grouped the results by month and performed 3 statistical trend tests on each group. We then grouped the data per year and performed the tests again. We can thus explore seasonal and yearly trends.

The trend tests were the following: the Mann Kendall Trend test, the Chi squared test for proportions and the logit regression. The Mann Kendall trend test has been used previously to detect trends in Alberta Lake water quality (Casey 2011). Since the dataset consisted of percentages, however, analytical methods tailored to proportion data, which have a binomial distribution, are more appropriate. We therefore conducted the Chi squared test of proportions and a logit regression. The results of the 3 tests were inspected for evidence of a trend through the p value, where a strong likelihood at the 95% confidence level that a trend is present was indicated by $p < 0.05$ and good likelihood at the 90% confidence level with $p < 0.1$. The Mann Kendall Trend test and the logit regression will also determine the direction of any trend, i.e., if the trend is an increase or a decrease. We considered that a trend is significant if at least 2 of the 3 tests resulted in a significant trend.

RESULTS

TEMPORAL TRENDS IN BLOOM DETECTION

We identified algal blooms on 156 out of the 444 images with a cloud cover over the lake less than 75% (figure 3) using the NIR with SAC and the Hue algorithms. This indicates that algal blooms were a common occurrence in Moose Lake during the study period. 40% of the algal blooms were detected in August, 27% in July, 24% in September, 8% in June and 1% in September. 1985 is the only year with no bloom detected at all. The majority (66%) of the detected blooms covered less than 10% of the study area (figure 4). However the method couldn't be applied to shallow areas where blooms may also have occurred or been accumulated by the wind.

To identify trends over time, we looked at the results of three statistical methods applied to the same dataset. The trend analyses indicated that there was no trend over time in the frequency of algal blooms in any of the individual months from June to October (Figure 5, Table 3). The trend analyses on the annual datasets also did not detect any trend over time in algal bloom frequency from 1984 to 2017 (Figure 6, Table 4). These results are in concordance with trend analyses recently completed by the Alberta Lake Management Society, who also did not identify any significant trends over time in chlorophyll-a, a water quality indicator of algae occurrence (ALMS 2018).

CONCLUSION

With the analysis of the Landsat datasets, we demonstrated that there have been algal blooms in Moose Lake every year from 1984 to 2017. Many of these blooms only covered a portion of the lake. Trend analyses on annual and monthly datasets indicated that the number of detected blooms varied from year to year without any significant trend over time.

REFERENCE

Alberta Lake Management Society (ALMS) 2018. Prepared by Laura E. Redmond.

Casey, R. 2011. Water quality conditions and long-term trends in Alberta lakes. Alberta Environment and Water,. Edmonton, AB. 419pp

Dilligeard, E and D Köster, 2017. Using Landsat Data for Macro-scale Algae Bloom Detection in Pigeon Lake: 1984-2016. Presentation given at Alberta Lake Management Society Workshop, Lac La Biche. September 2017.

Ho, JC, RP Stumpf, TB Bridgeman, AM Michalak, 2017. Using Landsat to extend the historical record of lacustrine phytoplankton blooms: A Lake Erie case study. Remote Sensing of Environment 191, 273-285

Stumpf RP, Wynne TT, Baker DB, Fahnenstiel GL, 2012. Interannual Variability of Cyanobacterial Blooms in Lake Erie. PLoS ONE 7(8): e42444.

U.S. Geological Survey 2018. http://glovis.usgs.gov/next/

TABLES

Satellite	Launch Date	Termination Date
Landsat 5	March 1984	June 2013
Landsat 7	April 1999	Still Active
Landsat 8	Feb 2013	Still Active

Table 1: Landsat satellites operational times.

Year	Jun	Jul	Aug	Sept	Oct	Total	Year	Jun	Jul	Aug	Sept	Oct	Total
1984	3	1	3	1	1	9	2002	4	3	2	1	5	15
1985	2	2	2	2	0	8	2003	3	5	2	2	1	13
1986	2	1	1	2	0	6	2004	3	4	3	0	5	15
1987	1	2	1	2	0	6	2005	2	3	4	3	6	18
1988	3	2	3	1	2	11	2006	4	3	5	1	4	17
1989	5	3	2	1	2	13	2007	3	5	2	4	5	19
1990	2	2	2	2	0	8	2008	3	2	1	5	2	13
1991	1	3	3	0	1	8	2009	4	3	5	4	3	19
1992	3	2	1	0	0	6	2010	2	3	6	4	5	20
1993	2	1	2	1	0	6	2011	4	6	7	5	4	26
1994	1	1	1	0	0	3	2012	2	1	3	3	1	10
1995	3	0	0	3	2	8	2013	0	5	4	6	5	20
1996	1	1	0	0	1	3	2014	2	5	4	6	3	20
1997	1	2	1	1	1	6	2015	2	6	5	5	2	20
1998	2	1	4	1	2	10	2016	5	6	4	2	0	17
1999	2	4	5	6	4	21	2017	2	5	5	5	1	18
2000	3	2	3	1	3	12							
2001	6	3	5	4	2	20	Total	88	98	101	84	73	444

Table 2: Summary of number of usable Landsat data per year and per month (cloud cover less than 75%). The usable dataset represents less than 50 % of the total number of available images.

Month	Logit regression		Chi Squared Test		Mann-Kendall trend test		
	P value	Slope	P value	χ-squared	P value	Tau	Score
June	0.964	0.001	0.724	0.124	0.939	0.013	5
July	0.566	-0.025	***0.030***	4.681	0.458	-0.100	-47
August	0.852	0.008	0.222	1.490	0.621	-0.069	-30
September	0.763	0.014	0.509	0.436	0.795	0.040	14
October	0.491	0.011	0.355	0.856	0.505	0.122	11

Table 3: Summary of the results of 3 statistical methods for trend analysis based on the annual percentage of images with a bloom per month and for the annual datasets. P values below 0.05 are highlighted in bold and italic font.

Logit regression		Chi Squared Test		Mann-Kendall trend test		
P value	Slope	P value	χ-square	P value	Tau	Score
0.830	-0.004263	0.1618	1.9569	0.914866	-0.00659	-64

Table 4: Summary of the results of 3 statistical methods for trend analysis based on the annual percentage of images with a bloom per year.

FIGURES

Figure 1: Example of pseudo color images where blooms are potentially present.

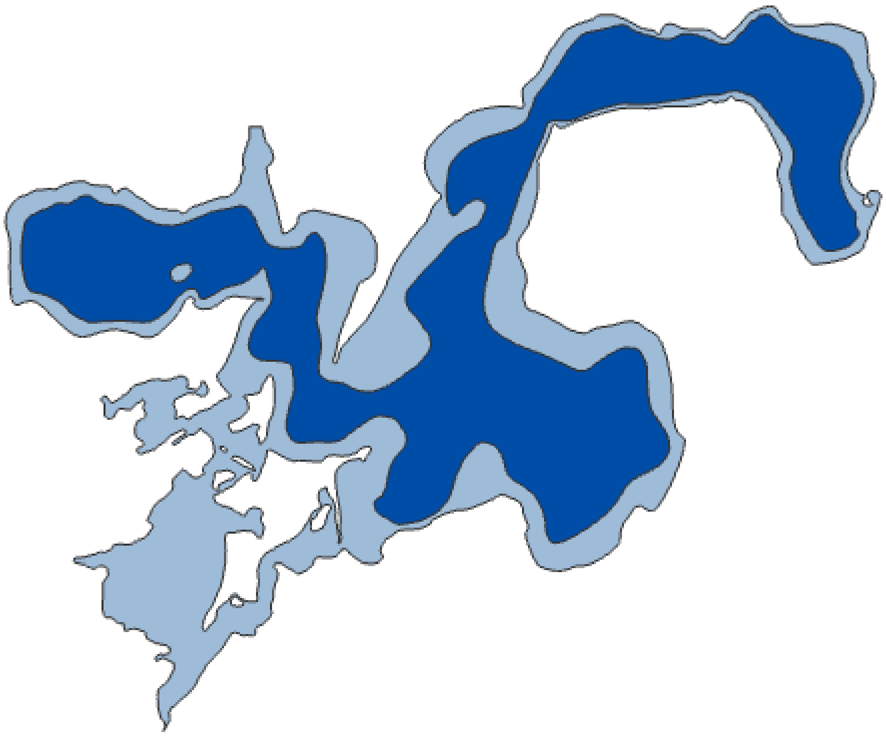

Figure 2: Moose Lake (light blue) and Study area (dark blue). The study area represents areas with a depth larger than 3.2 m.

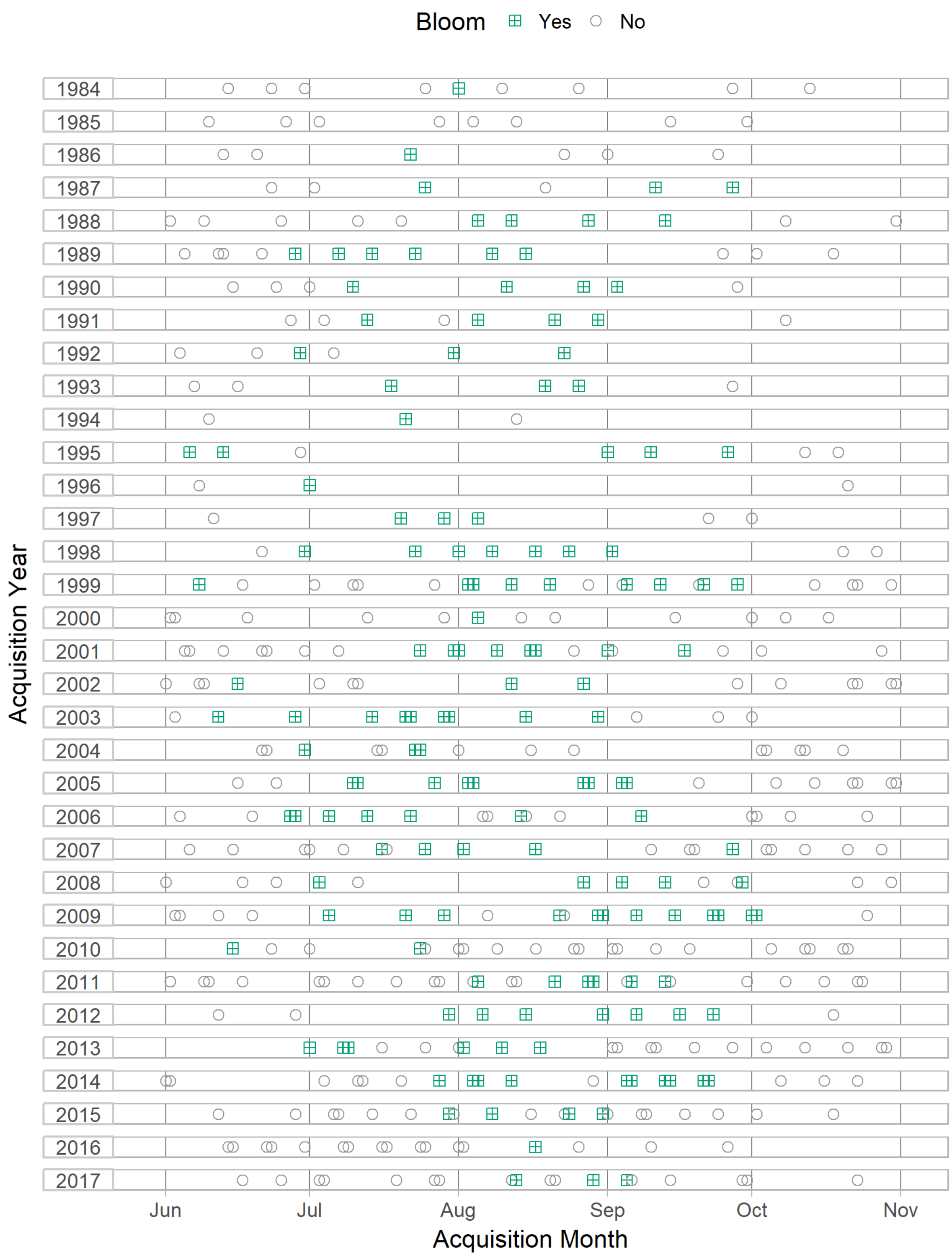

Figure 3: Summary of bloom identification on 34 years of Landsat images based on the analysis of the NIR with SAC and Hue products calculated for images with a cloud cover less than 75%.

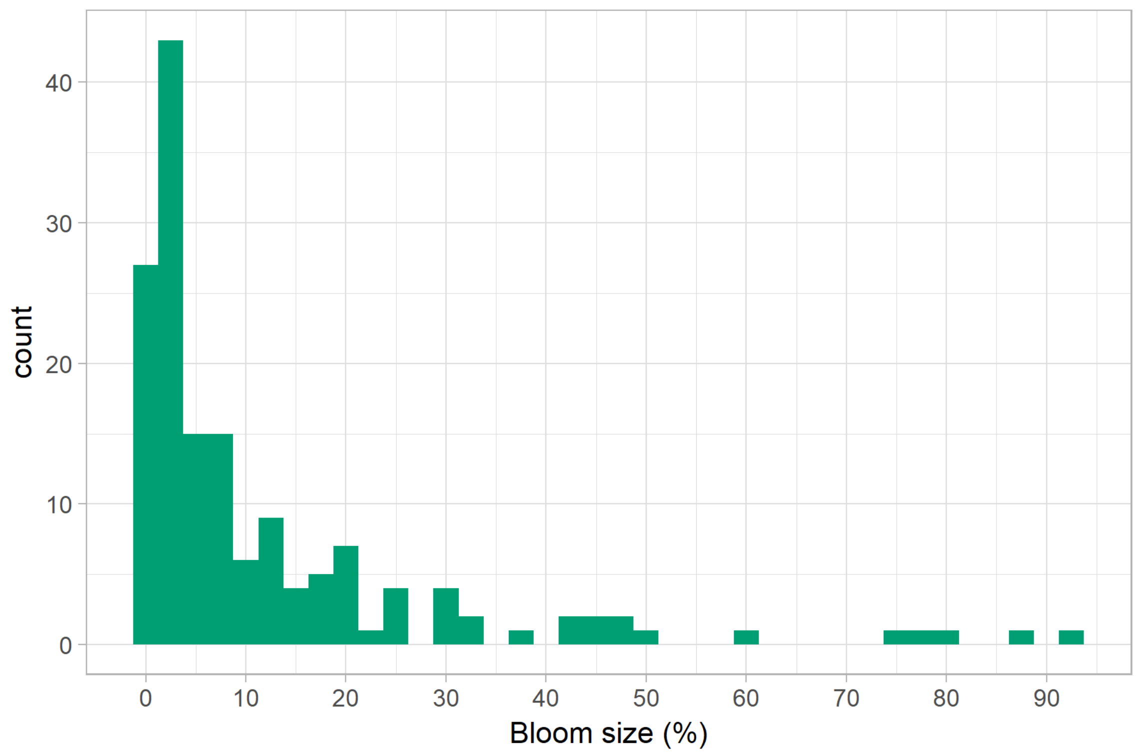

Fig 4: Histogram on bloom size as percentage of the lake study area.

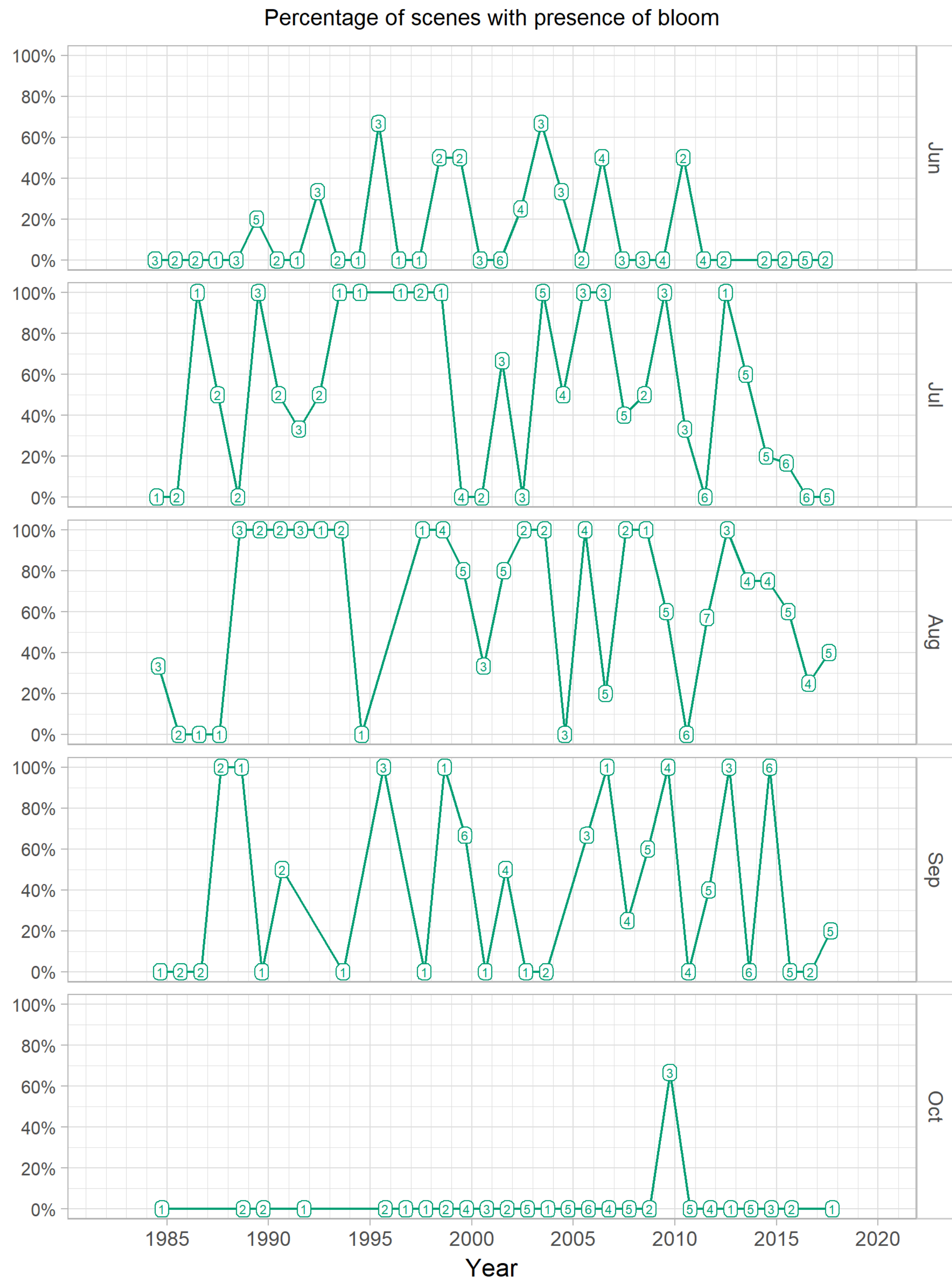

Figure 5: Monthly percentage of images with presence of bloom per year. The green label represents the total of usable image for a month and a year.

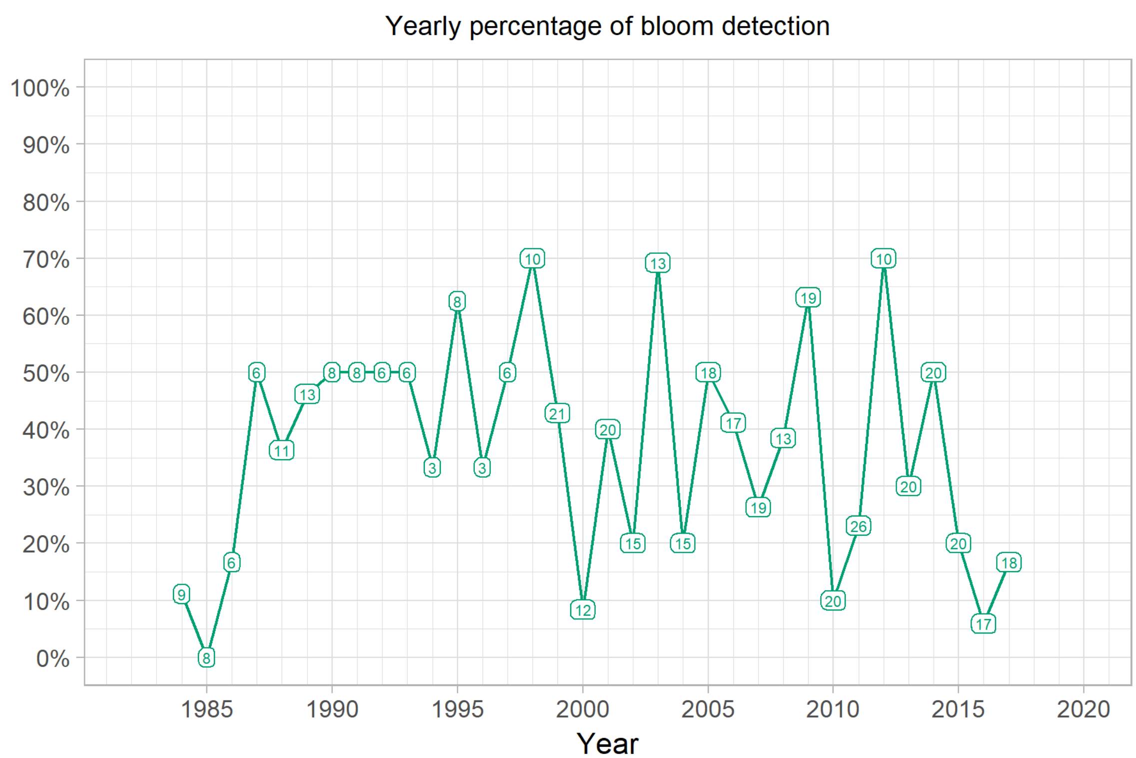

Figure 6: Yearly percentage of images with presence of bloom per year. The green label represents the total number of usable images for a month and a year.

APPENDIX 1: PSEUDO COLOR IMAGES

APPENDIX 2: RESULTS OF SAC AND HUE ALGORITHMS: FIGURES

APPENDIX 3: RESULTS OF SAC AND HUE ALGORITHMS: TABLE