

Tracking Multi-Decadal Lake Dynamics using Optical Imagery, Digital Elevation Models, and Bathymetric Datasets

© 2020

David Weekley

M.A., University of Nebraska-Lincoln, 2009

B.S., University of Nebraska-Lincoln, 2007

Submitted to the graduate degree program in Geography and Atmospheric Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Xingong Li

Jude H. Kastens

Stephen L. Egbert

Ting Lei

Leigh Stearns

Date Defended: 29 April 2020

The dissertation committee for David Weekley certifies that this is the approved version of the following dissertation:

**Tracking Multi-Decadal Lake Dynamics using Optical Imagery,
Digital Elevation Models, and Bathymetric Datasets**

Chair: Xingong Li

Date Approved: 29 April 2020

Abstract

The goal of this research is to review the current state of long-term, multi-decadal lake dynamic monitoring and develop novel techniques for scalable analysis at local, regional, and global levels. This dissertation is comprised of three chapters formatted as journal manuscripts with each chapter progressively addressing some key limitation in current lake dynamic monitoring methodologies.

Chapter 1 tracks lake dynamics (surface elevation, surface area, volume, and volume change) for a single water body, Lake McConaughy, which is the largest lake and reservoir in the state of Nebraska, using the cloud-based geospatial analysis platform Google Earth Engine. Lake dynamics were estimated using bathymetric survey data, the Shuttle Radar Topography Mission 30-meter digital elevation model, and Landsat 5 image composites for 100 time periods between 1984 and 2009. Water surface elevation was estimated and assessed for 5,994 different combinations of water indices, segmentation thresholds, water boundaries, and statistics and produced elevations as accurate as 0.768 m $CI_{95\%}$ [0.657, 0.885] root-mean-square-error. The method also detected seasonal and long-term trends which would have major implications for regional agriculture, recreation, and water quality. Chapter 1 was published as an article in the peer-reviewed journal *Water Resources Research* in October 2019.

Chapter 2 expands and improves upon the techniques explored in Chapter 1 in multiple ways. First, the techniques were improved to remove image contamination sources such as snow, ice, cloud cover, shadow, and sensor error for individual images using the Pixel Quality Assurance (QA) band available as a part of the Landsat 4, 5, 7, and 8 Top-of-Atmosphere Tier-1 Collection-1 archives. Using the Pixel QA band information, image contamination was removed from each image between August 1982 and December 2017 and water surface elevation was estimated with the remaining visible water boundary extents overlaying merged National

Elevation Dataset digital elevation model and bathymetric survey data resampled to 30-meters which resulted in enhanced temporal resolution compared to the techniques used in Chapter 1. Second, the analysis was expanded from a single water body to fifty-two lakes/reservoirs to provide a better understanding of how the techniques generalize to imagery and water bodies encompassing a wide range of ecotypes, geologies, climates, and management strategies. A variety of common water indices, such as the Modified Normalized Difference Water Index, naïve and dynamic water indices, water boundary types, and filtering strategies were tested and individual lake accuracies are as low as 0.191m RMSE CI95% [0.129, 0.243], and 45 of the 52 lakes produced sub-meter root-mean-squared-error accuracies. Furthermore, accuracy of surface elevation estimates is highly correlated with the mean slope of surrounding terrain with low-slope shorelines having greater accuracy than high-slope shorelines such as those in canyon-filled reservoirs or in mountainous regions. Overall, the improved techniques extend our ability to track long-term lake dynamics to lakes with bathymetric datasets while lacking in-situ hydrological stations, provide a framework for scale-able analysis in Google Earth Engine, and balance a need between high-accuracy estimates and maximum temporal resolution.

Bathymetric survey data, such as that used in Chapters 1 and 2 is, unfortunately, not available for most water bodies at regional and global scales. Chapter 3 introduces a method of tracking long-term lake dynamics without bathymetry data and only using available digital elevation models such as Shuttle Radar Topography Mission, the National Elevation Dataset, and Advanced Land Observing Satellite. In digital elevation models, the water surface is often, but not always, hydroflattened producing a flat surface approximating the surface of the water at the time of the data capture which precludes using water boundaries like those in Chapter 1 and Chapter 2 to estimate water level when it is lower than the hydroflattened surface in the digital elevation model. However, using hypsometric relationships developed from the digital elevation

models, subsurface water dynamics can still be estimated by extrapolating the low water levels using regression, albeit with increased uncertainty compared to levels above the hydroflattened surface. Using multiple digital elevation models, the lowest hydroflattened surface can be identified for each water body which reduces uncertainty for low water levels by reducing the extrapolation distance to those values while simultaneously increasing the number of above hydroflattened surface estimates. In addition to low-level uncertainty, hypsometric techniques are highly impacted by image contamination such as cloud, cloud shadow, snow, ice, and sensor error which reduces the observable water surface area resulting in erroneous surface elevation, volume, and volume change estimates. To help alleviate this issue, a technique of using proportional hypsometry was developed to remove contamination effects. Together, using the lowest hydroflattened surface and proportional hypsometry, this research produced 12,680 additional water surface elevation estimates for 46 lakes in comparison to traditional hypsometric techniques, reduced the number of sub-hydroflattened water surface estimates by 549 or more compared to individually using any of the three digital elevation models assessed, and lays the groundwork for regional and global scale surface water dynamic research without bathymetric survey data.

Acknowledgments

This dissertation would not have been possible without the contributions and support of my family, friends, colleagues, committee members, and mentors. My committee members, Dr. Xingong Li, Dr. Jude Kastens, Dr. Ting Lei, Dr. Stephen Egbert, and Dr. Leigh Stearns, thank you for the research suggestions, recommendations, and guidance over the years. This research would not be where it is today without your contributions.

Special thanks to Stephen Egbert for the nudge to go back to school and pursue a PhD, this has been a fun and exciting journey. Special thanks to Jude Kastens as well, who on more than a few occasions impromptu hosted me in his office to hash out thoughts and point me towards hidden datasets. Most of all, thanks to my advisor, Xingong Li, who suggested research, tirelessly read my rough drafts, pushed me to present my research at conferences, and always encouraged progress. Also thank you to Jim Coll, Regina Biswah Thomas, Kenneth Ekpeter, and Chen Liang for enduring my papers and presentations in our weekly research group meetings.

Additionally, I would like to thank the staff at the Kansas Biological Survey; Jude Kastens, Dana Peterson, and my team members Adnan Zahoor, Dustin Fross, Elizabeth Wesley, and Mike Houts. Each of you has been a pleasure to work with and supportive in my pursuit of a PhD. Special thank you to Mike Houts for allowing a flexible work schedule enabling me to run over to campus for classes and meetings as needed.

Finally, thank you to my friends and family. It is your support that has made this all possible. Thank you to my parents, who fed my curiosity and taught me perseverance and so many valuable life lessons. Most of all though, thank you to my wife Hannah, who has remained gracious, patient, encouraging, and loving since day one. I would not be here without you.

An edited version of Chapter 2, *Tracking Multidecadal Lake Water Dynamics with Landsat and Topography/Bathymetry*, was published by AGU. Copyright (2019) American Geophysical Union.

Weekley, D., & Li, X. (2019). Tracking Multidecadal Lake Water Dynamics with Landsat Imagery and Topography / Bathymetry. *Water Resources Research*, 1–18.
<https://doi.org/10.1029/2019WR025500>

Table of Contents

1	Introduction.....	1
2	Tracking Multi-Decadal Lake Water Dynamics with Landsat Imagery and Topography/Bathymetry	4
2.1	Abstract.....	4
2.2	Introduction.....	4
2.3	Study Area and Data.....	9
2.4	Methods	11
2.4.1	Constructing a Time Series from Landsat 5 Images.....	12
2.4.2	Lake Water Identification	14
2.4.3	Lake Water Surface Elevation Estimation.....	15
2.4.4	Validation	16
2.5	Results.....	17
2.6	Discussions	21
2.6.1	Composite Images and Cloud Cover	21
2.6.2	Segmentation Thresholds, Boundaries, and Statistics	21
2.6.3	Water Index Performance	23
2.6.4	Identification of Water Body Interest	24
2.6.5	Water Volume Change.....	24
2.6.6	Surface Area to Elevation Relationships	25
2.6.7	Time Series Implications	25
2.7	Conclusions.....	27
2.8	Figures and Tables.....	29
3	Maximizing Multi-Decadal Water Surface Elevation Estimates with Landsat Imagery and Elevation/Bathymetry Datasets	40
3.1	Abstract.....	40
3.2	Introduction.....	41
3.3	Study Area and Data.....	44
3.4	Methods	46
3.4.1	Digital Elevation Model and Bathymetry Merger	46
3.4.2	Water Surface Elevation Estimation.....	47
3.4.3	Water Detection and Shoreline Extraction	48
3.4.4	Lake Surface Area and Volume.....	50
3.4.5	Contamination/Impacted Area.....	50

3.4.6	Lake Morphological Metrics	51
3.4.7	Model Selection and Validation.....	51
3.5	Results.....	52
3.6	Discussion.....	55
3.6.1	Terrain Slope.....	55
3.6.2	Image Contamination.....	56
3.6.3	Water Index Performance and Filtering.....	57
3.6.4	QA Band Limitations.....	58
3.6.5	Segmentation Threshold	58
3.6.6	Time Series	59
3.7	Conclusions.....	60
3.8	Figures and Tables.....	62
4	Tracking Lake Surface Elevations with Proportional Hypsometric Relationships, Landsat Imagery, and Multiple DEMs	75
4.1	Abstract.....	75
4.2	Introduction.....	76
4.3	Data and Study Area.....	80
4.4	Methods	81
4.4.1	Lake Seeding.....	82
4.4.2	Water and Shoreline Detection	82
4.4.3	Image Contamination.....	83
4.4.4	Hypsometric Relationships	84
4.4.5	Model Selection and Accuracy	85
4.5	Results.....	85
4.6	Discussion.....	87
4.6.1	Digital Elevation Model Differences	87
4.6.2	Image Contamination.....	88
4.6.3	Time Series and Individual Lake Accuracy.....	88
4.7	Conclusions.....	89
4.8	Figures and Tables.....	92
5	Conclusions.....	104
6	References.....	108

List of Figures

Figure 2-1: Lake McConaughy and its location in the state of Nebraska along with orbital paths of altimetry missions commonly used for inland water surface elevation monitoring within the vicinity of the reservoir. Of the available sensors, only RA-2 (orange color) onboard Envisat routinely crosses the main body of the reservoir but contains significant noise. The background image is a natural-color USGS Landsat 5 TM Collection 1 Tier 1 Raw Scene from September 2 nd , 2001.	29
Figure 2-2: SRTM DEM (left) and bathymetry merged with SRTM DEM (right).	30
Figure 2-3: An example of binary water index image. The top image is before the vector intersect approach has been applied and includes all water bodies within the study area.	30
Figure 2-4: Three potential types of lake water boundaries are possible with a binary water image: interior, exterior, and combination. Water boundaries were identified using a one pixel radius square kernel to erode and/or dilate the water areas using morphological operations. Considering the water body as a whole, each of the water boundary types has a different focus and therefore uses different elevation values from the underlying DEM when estimating water surface elevation. Interior boundaries use submerged water pixels along the shoreline, exterior boundaries use land pixels along the shoreline, and combination boundaries use both land and water pixels. For a given image, interior boundaries will produce the lowest water surface elevation estimate and exterior boundaries will produce the highest estimate. Combination boundary water surface elevation estimates will fall between those of interior and exterior boundaries.	31
Figure 2-5: RMSE (Root-Mean-Square-Error) curves by water index types. (a) Normalized Difference Water Index (NDWI); (b) Modified Normalized Difference Water Index (MNDWI); (c) Automatic Water Extraction Index – no shadow (AWEInsh); (d) Automatic Water Extraction Index – shadow (AWEIsh); (e) NDWI + MNDWI (B1 & B4); (f) NDWI + MNDWI (B2 & B5).	33
Figure 2-6: Lake water surface elevation, surface area, and volume over time. (a) Surface elevation; (b) Surface area; (c) Water volume.	35
Figure 2-7: (a) Reservoir maximum, minimum, median summer surface elevations and lake depth along a centralized line bisecting the reservoir longitudinally. It shows the vast change in water quantity between the maximum and minimum reservoir levels. Note that the y-axis is not in the same scale as the x-axis; False-color infrared composites of peak reservoir level in June, 1986 (b) and minimum reservoir level in September, 2006 (c). The western end of the reservoir retreated nearly 14.5 kilometers during that time span.	35
Figure 2-8: Surface cover frequency of Lake McConaughy during the entire study period from 1985-2009. Blue regions indicate areas where water was always or almost always detected. Red regions are areas where water was detected as little as one time.	36
Figure 3-1: Spatial distribution of lakes used in the study.	62
Figure 3-2: General processing procedure for estimating water surface elevation for any given lake with merged DEM/bathymetry data. This procedure was implemented in GEE and applied to each image within the Landsat image stack.	63
Figure 3-3: Comparison of boundaries with and without the dilated contamination mask for an image with approximately 50% localized contamination. Without the mask, the contaminated areas are identified as non-water which generates a boundary crossing the center of the lake in several locations (A). However, using the Landsat QA band, not only can contaminated areas be identified, but they can be used to remove erroneous shorelines as well (B).	63

Figure 3-4: Three different types of shoreline water boundaries were used to estimate lake water surface elevation (Weekley and Li, 2019).....	64
Figure 3-5: RMSE (A) and MAE (B) accuracy for each lake plotted from lowest to highest accuracy using the best overall model. For comparison, the best individual lake models are also shown.....	65
Figure 3-6: A) General model absolute error (estimated surface elevation – ground truth elevation) vs shoreline length. B) Lake-specific model absolute error vs shoreline length. C) General model estimate counts per 7.5 km bin. D) Lake-specific model estimate counts per 7.5 km bin. The charts for absolute error were each capped at 20 meters absolute error and all charts were capped at 350 km shoreline. ~1% of measurements exceeded these thresholds.....	66
Figure 3-7: Monthly RMSE accuracies and image counts for the best performing overall model and the best lake-specific models with shoreline filtering and mean-mode difference filtering.....	67
Figure 3-8: Selected water surface elevation time series. Most lakes showed good agreement between the estimated water surface elevations and the ground-truth. Note, the y-scale varies for each chart due to differences in the estimated water levels.	69
Figure 3-9: A) False-color infrared image of Big Hill Reservoir, Kansas, USA using Landsat 5 TOA Tier-1 Collection-1 image from July 16, 1993. The lake appears to be free of image contamination at first glance, yet no water was detected for this image. B) The same image with the QA Band coverage displaying in grey. C) The impacted areas (dark areas) and clear areas (light areas). All pixels with non-clear attribution, including no data regions, are considered impacted for the purposes of this study.	69
Figure 4-1: Forty-six lakes and reservoirs utilized within the study.....	92
Figure 4-2: The general processing procedure used in this study to dynamically estimate water surface elevation using hypsometric relationships.....	93
Figure 4-3: Clear and contaminated Landsat 5 image examples for Clinton Lake, Kansas, USA from (a) September 3, 1985 and (b) April 27, 2008. Each example shows the lake as it appears in (1) false-color infrared, (2) false-color infrared with transparency overlaying a hill shade with detected water in blue and image contamination in yellow, and (3) the resulting hypsometric relationship.....	94
Figure 4-4: Accuracy comparison as the amount of image contamination increases for 1) all images above the hydroflattened water level, 2) the LHFS DEM model with all images (3 rd -degree fit for above hydroflattened surface images and 2 nd -degree fit for sub-hydroflattened surface images), and 3) the LHFS DEM model with all above-hydroflattened surface images (3 rd -degree fit) and images with <=1% local image contamination (2 nd -degree fit).	94
Figure 4-5: Time-series charts for selected lakes and reservoirs. Blue points represent the estimated water surface elevations while the gray line represents in-situ water surface elevation measurements.....	97
Figure 4-6: Common water levels and their accompanying surface areas for selected water bodies.....	98
Figure 4-7: Comparison of the 30 meters above sea level inundation extent on the (a) NED, (b) SRTM, and (c) ALOS digital elevation models for B.A. Steinhagen Lake in Texas, U.S.A and 1,278 meters above sea level on (d) NED, (e) SRTM, and (f) ALOS digital elevation models for Anderson Ranch Reservoir in Idaho, USA.	98
Figure 4-8: Large-scale interconnectivity of water bodies in the area of Big Sandy Lake, Minnesota, USA. False-color infrared Landsat 5 image from August 16, 1999 overlaid with a NED 30m hillshade to highlight the elevation differences. The blue area is the detected water for Big Sandy Lake and the green area is the connected area for the DEM at the August 16, 1999 in-situ elevation level. For this lake, the interconnected water body issue was present for all tested DEMs as well as the full-resolution NED.....	99

List of Tables

Table 2-1: Band designations and additional information for the Landsat 5 Thematic Mapper.	37
Table 2-2: Water index formulas using Landsat 5 Thematic Mapper bandwidths.	37
Table 2-3: Statistical formulas used to assess lake dynamic estimate accuracy.	38
Table 2-4: All parameters and values tested.	38
Table 2-5: The number of combinations by water index types which exceed selected water surface elevation RMSEs. NDWI + MNDWI (B1 & B4), for example, had 245 combinations (24.5%) with RMSEs better (lower) than 1.0 meters and 10 combinations (1.0%) better than 0.8 meters.	38
Table 2-6: Number of analyses with RMSEs less than 1.0 meter (water surface elevation). Table is broken down by water index types and boundary / statistics types. Of the tested parameter combinations, interior and combination boundaries outperformed exterior boundary in terms of water surface elevation accuracy.	39
Table 2-7: Accuracy statistics for water surface elevation, surface area, volume, and volume change using Landsat 5 image composites and merged bathymetry/SRTM data. Also, error statistics for water volume change estimated the pyramidal frustum method are also included for comparison.	39
Table 3-1: Sensor and mission details for each Landsat mission included in this study including available image date ranges, spatial resolution, temporal resolution, and radiometric resolution.	70
Table 3-2: Bathymetric survey information and vertical adjustment used (where applicable).	71
Table 3-3: Overall, 594 different image processing parameter combinations were tested including six water indices, 11 segmentation thresholds (nine pre-set, two dynamic), three boundary types, and three statistic types.	72
Table 3-4: Accuracy metrics for the best performing general and lake-specific models using various combinations of minimum shoreline lengths and maximum mean-mode difference thresholds to remove erroneous data points. Mean % impacted corresponds to localized image contamination (image contamination within 500 meters of median lake extent), filtered valid images is the number of images remaining after the shoreline length and mean-mode difference filters have been applied, and total estimates are the number of images which produced an elevation estimate.	73
Table 3-5: Pearson's product-moment correlation r^2 for RMSE, MAE, and potential error drivers.	74
Table 3-6: Accuracy metrics for each Landsat mission. Accuracy statistics for total overall, SLC-On, and SLC-Off are provided for Landsat 7 ETM+.	74
Table 4-1: Sensor and mission details for each Landsat mission used in this study including available image date ranges, spatial resolution, temporal resolution, and radiometric resolution.	100
Table 4-2: Number of Landsat images included in the analysis. <1% contamination represents images available to conventional hypsometric techniques. <100% contamination are the number images with at least one clear image pixel within 500 meters of the JRC GSW observed water extents for each lake (the <100% Contaminated column also includes images with <1% contamination).	101
Table 4-3: Comparison of hypsometric relationships generated from ALOS, NED, SRTM, and the LHFS DEM for all clear images (<1% contamination).	101
Table 4-4: Accuracy assessment for all 32,899 images with observable shorelines and in-situ surface elevation measurements.	102
Table 4-5: Accuracy comparison for sub-hydroflattened surface images with <1% local image contamination.	102

Table 4-6: Accuracy comparison for all images above hydroflattened water surface levels. An image is above hydroflattened if the estimated water surface area is larger than the surface area of the DEM's hydroflattened water surface..... 103

Table 4-7: Accuracy comparison for all DEM models using all above hydroflattened surface images fit with a 3rd-degree polynomial and sub-hydroflattened surface images with <1% local image contamination fit with a 2nd-degree polynomial. 103

1 Introduction

To the limits of our current knowledge, liquid water is universally recognized as one of the key ingredients for life. In astrobiology, scientists theorize life could exist in the liquid water beneath the icy crust of moons like Europa and Enceladus or in liquid water stored beneath the surface of Mars (McKay, 2014). Such a discovery, life from another world, would fundamentally change our perception and place in the universe, yet, despite its known importance to the only type of life we know and the planet that harbors it, Earth, our knowledge of our planet's surface water is limited. Approximately 97.2% of Earth's surface water is contained within the oceans with just the remaining 2.8% being considered freshwater. Of that 2.8%, 2.1% is stored as glacial ice, 0.6% is stored as groundwater and soil moisture, and the remaining 0.1% is split amongst the atmosphere, lakes, and rivers (California, 2017; USGS, 2018a). In fact, just 0.01% of the global water supply is available as liquid freshwater and yet that water directly supports 100,000 species including humans and human related activities (Dudgeon et al., 2005). We understand that freshwater is critically important, in fact, freshwater resources impact geology, climatology, ecology, and numerous other fields yet we know little of how these resources vary in time and space (J F Crétaux et al., 2016).

Historically, our primary source of knowledge concerning surface water dynamics has been in-situ hydrological stations installed on-site for lakes, reservoirs, and streams. Unfortunately, the spatial distribution of these stations is severely limited which restricts their applicability when it comes to large-scale studies. Even in the United States, which has a dense water monitoring network, hydrological stations are only available for a small subset of the total number of water bodies. The station availability issue multiplies when working internationally as entire regions may have just a handful of stations and bureaucratic policies may restrict access to the data (Alsdorf, Rodríguez, & Lettenmaier, 2007). This data limitation has severe implications when it comes to understanding surface water dynamics over large scales. Remote sensing has emerged as a viable alternative means of monitoring surface water dynamics with techniques being developed using optical sensors like Landsat and MODIS, synthetic aperture

radars, and satellite altimetry. More information and in-depth review of these topics will be covered in Chapters 2, 3, and 4.

The research presented in this dissertation advances our understanding of surface water dynamics – how water surface elevation, surface area, volume, and volume change vary spatiotemporally – through the development of new techniques for monitoring these resources using widely available optical image archives, digital elevation models, and bathymetric datasets. These techniques were developed primarily within Google Earth Engine (GEE) (Gorelick et al., 2016) to be readily scalable from local, to regional, to global scales but can also be implemented in a variety of GIS programs.

This dissertation is organized into three journal article style chapters with a common purpose; estimating long-term lake/reservoir dynamics using available datasets. Chapter 2, *Tracking Multi-Decadal Lake Water Dynamics with Landsat Imagery and Topography/Bathymetry* was published in Water Resources Research in 2019 (Weekley & Li, 2019). This chapter develops a technique for estimating water dynamics using Landsat 5 image composites, water boundary extents, and merged Shuttle Radar Topography Mission (SRTM)/bathymetry data for Lake McConaughy in Nebraska, USA from 1985 to 2009. Chapter 3, *Maximizing Multi-Decadal Water Surface Elevation Estimates with Landsat Imagery and Elevation/Bathymetry Datasets*, builds upon and improves upon the technique used in Chapter 2 by leveraging the Pixel Quality Assurance band packaged with the Landsat 4, 5, 7, and 8 Tier-1 Collection-1 Top-of-Atmosphere image archives to identify and remove the effects of image contamination such as snow, ice, cloud, cloud shadow, and sensor error. Removing the effects of image contamination enables individual images to be used, rather than composites such as those used in Chapter 2, which greatly increases the temporal resolution of the analysis. Furthermore, the analysis in Chapter 3 is performed on 52 lakes with varying geologies, climates, ecotypes, and management practices to better understand how the method generalizes to unseen images and water bodies. Chapter 4, *Proportional Hypsometric Relationships for Maximizing Lake Dynamic Tracking Temporal Resolution and Reducing Low-Level Uncertainty with Landsat Imagery and Digital Elevation Models*, takes a different approach to

the long-term water dynamic monitoring issue by using hypsometric relationships derived from digital elevation models only, no bathymetry, to estimate water surface elevation from water surface area estimates. While bathymetry may be available at regional levels, such as some states in the USA which have large bathymetric survey efforts and data catalogs (deNoyelles & Kastens, 2016), bathymetry is relatively sparse at the global scale. This chapter explores methods of using available digital elevation models to estimate water surface elevation, and therefore water surface dynamics, using multiple digital elevation models to minimize sub-hydroflattened surface estimates while using a novel proportional hypsometry technique generated using the Pixel Quality Assurance band to remove image contamination.

Overall, these three chapters represent a unified theme of method development and refinement for tracking long-term lake dynamics using optical imagery and available elevation data products, are scalable for analysis at local, regional, and global scales, and provide insight into trends and patterns at seasonal, annual, and decadal timeframes.

2 Tracking Multi-Decadal Lake Water Dynamics with Landsat Imagery and Topography/Bathymetry

2.1 Abstract

Water resource management is of critical importance due to its close relationship with nearly every industry, field, and lifeform on this planet. The success of future water management will rely upon having detailed data of current and historic water dynamics. This research leverages Google Earth Engine and uses Landsat 5 imagery in conjunction with bathymetry and Shuttle Radar Topography Mission digital elevation model data to analyze long-term lake dynamics (water surface elevation, surface area, volume, volume change, and frequency) for Lake McConaughy in Nebraska, USA. Water surface elevation was estimated by extracting elevation values from underlying bathymetry and digital elevations models using 5,994 different combinations of water indices, water boundaries, and statistics for 100 time periods spanning 1985-2009. Surface elevation calculations were as accurate as 0.768 meters RMSE $CI_{95\%}$ [0.657, 0.885]. Water volume change calculations found a maximum change of 1.568 km³ and a minimum total volume of only 23.97% of the maximum reservoir volume. Seasonal and long-term trends were identified which have major affects regarding regional agriculture, local recreation, and lake water quality. This research fills an existing gap in optical remote sensing-based monitoring of lakes and reservoirs, is more robust and outperforms other commonly used monitoring techniques, increases the number of water bodies available for long-term studies, introduces a scalable framework deployable within Google Earth Engine, and will enable data collection of both gauged and un-gauged water bodies which will substantially increase our knowledge and understanding of these critical ecosystems.

2.2 Introduction

Water is one of the most abundant resources upon Earth and is also one of the most critical to life. While water is massively abundant when considering the Earth's surface as a whole, its distribution is uneven across both time and space (J F Crétau et al., 2016) leading to extensive impacts and

implications for biology, ecology, economy, and human welfare. Ocean waters aside, global surface freshwater dynamics are especially vital considering freshwater constitutes just 0.01% of the global water supply and yet it supports at least 100,000 different species including humans (Dudgeon et al., 2005). Yet, despite its universal importance, our understanding of continental surface water dynamics is limited.

At the most basic level, water dynamics describe where, when, and how much water is present on the landscape. For lakes and reservoirs, water dynamics are observed through water surface elevation, surface area, volume, and volume change measurements. Historically, most water dynamic information has been derived from in-situ gauge networks which measure water surface elevation (height or stage) that can be combined with bathymetric survey data (pre- or post-impoundment for reservoirs) to create storage curves relating elevation to surface area and/or volume. Unfortunately, most water bodies lack in-situ monitoring and the data for many others are unavailable due to legal or institutional restrictions (Alsdorf et al., 2007). This problem exists even in developed nations with dense gauge networks. For example, in the State of Kansas, 60% of the population relies upon 80 reservoirs for their primary or back-up drinking water supply (Rahmani et al., 2018), most of which are unmonitored or do not have publically available datasets.

To alleviate the issues with in-situ monitoring, several remote sensing techniques have been developed to estimate water dynamics. As mentioned earlier, water dynamics are typically monitored through changes in water surface elevation, surface area, volume, and volume change. Water surface area is often the simplest metric to estimate as it can often be directly measured using optical sensors, such as the multi-spectral sensors onboard MODIS (Moderate Resolution Imaging Spectroradiometer) and Landsat, or using SAR (synthetic aperture radar) systems. Water surface elevation can also be directly measured, if not available via in-situ gauges, through the use of satellite altimeters such as the dual- or single-frequency altimeters onboard NASA/CNES's Topex/Poseidon satellite ("TOPEX/Poseidon Fact Sheet," n.d.) or GLAS (Geoscience Laser Altimetry System) onboard NASA's IceSat (Zwally, n.d.). Unlike water surface elevation and surface area, volume and volume change cannot be directly measured.

Total volume estimates require knowledge of underlying bathymetry while volume change, at the bare minimum, requires both water surface elevation and surface area measurements on two separate dates to estimate volume change via pyramidal or conical frustum equations (Alsdorf et al., 2007; J F Crétau et al., 2016; Gao, 2015).

The aforementioned techniques have been used in numerous studies to analyze water dynamic changes for several water bodies around the globe including, but not limited to, the following examples. Gao et al. (2012) used MODIS (Moderate Resolution Imaging Spectroradiometer) and satellite altimetry to study storage variations for 34 global reservoirs via elevation/surface area/volume relationships. Crétau et al. (2015) used Landsat imagery, satellite altimetry, and the pyramidal frustum formula to investigate regional water dynamics in the Sydarya River region of Central Asia. Moradi et al. (2014) used DEMs and sub-pixel reprocessed MODIS imagery to estimate total water volume and mean lake level for the Caspian Sea. Zhang et al. (2016) used TanDEM-X DEMs to explore bathymetry and storage of unmonitored reservoirs in Brazil. Tseng et al. (2016) used Landsat imagery and DEMs to track water level changes in Lake Mead by estimating subsurface bathymetry from surrounding topography. Cai et al. (2016) used MODIS and area-based water storage estimation models to analyze 15 years of spatiotemporal water storage dynamics for large lakes and reservoirs in the Yangtze River Basin. Additional works in this area include, but are not limited to, studies by Yuan et al. (2017), Wang et al. (2013), Duan and Bastiaanssen (2013), Liang et al. (2017), El-Shazli and Hoermann (2016), Avisse et al. (2017), and Jiang et al. (2017).

While the methods used in each of the examples above have added substantially to the existing body of knowledge and undoubtedly will continue to do so, each also possesses a distinct set of limitations. Gao et al. (2015) found seasonal effects in MODIS surface area estimates due to subgrid spatial heterogeneity related to mixed pixels along the water shoreline because of the coarse spatial resolution (250 m to 500 m) of MODIS. Moradi et al. used sub-pixel reprocessing to minimize mixed pixels resulting from the coarse resolution of MODIS imagery. Unfortunately, sub-pixel reprocessing

requires increased computation time and accuracy of water placement within the overall pixel is algorithm dependent (Moradi et al., 2014; Xiong et al., 2018). Both the works of Avisse et al. (2017) and Zhang et al. (2016) are reliant upon water levels being low enough at the time of observation to capture a complete or near-complete bathymetric profile. Unfortunately, most lakes and reservoirs pre-date modern remote sensing and/or lack the natural variability that would make this technique widely applicable. Tseng et al. (2016), on the other hand, attempt to alleviate the issue by extending the slope of surrounding terrain to estimate bathymetry. While the method performed reasonably well for Lake Mead, DEM resolution dependencies and increasing uncertainty at low-water levels must be considered before applying the technique elsewhere. Techniques such as those employed by Cai et al. (2016) use empirical models which estimate storage capacities based on regional measurements. Other studies, such as those by Messenger et al. (2016) and Heathcote et al. (2015), which also estimated storage and bathymetric features using regional data, found that the techniques worked well at the regional level but cautioned that the results of individual lakes could contain significant error. Satellite altimetry based studies, on the other hand, which estimate water surface elevation by measuring the return time of signals reflected off the surface are limited due to sensor constraints (Solander, Reager, & Famiglietti, 2016). Most altimetry missions were designed for ocean or cryosphere monitoring and lack the resolution needed to discern smaller bodies of water without substantial land contamination. Furthermore, the application of altimetry to inland water monitoring is also limited due to orbital characteristics such as large ground-track spacings which prevent observation of many water bodies (Solander et al., 2016). While altimeter-based calculations are capable of sub-decimeter accuracy, those types of results are generally limited to large lakes with favorable shape and conditions (Asadzadeh Jarihani, Callow, Johansen, & Gouweleeuw, 2013a; J F Crétaux et al., 2016). Even studies which use elevation/area/volume relationships derived from bathymetric surveys can have issues in cases where surface area is used to estimate water surface elevation or volume. In this case, surface area is represented by a single value and any errors or deviations in that number, such as those caused by local erosional and depositional processes captured in the imagery but not in the bathymetry, directly effect the resulting elevation or volume calculation.

In 2021, the SWOT (Surface Waters and Ocean Topography) mission is set to launch and will feature increased spatial and temporal resolution over its SAR and altimeter predecessors while also offering approximately 90% coverage of the Earth's surface (Biancamaria, Lettenmaier, & Pavelsky, 2016). The SWOT mission is expected to offer water surface elevations within 10 – 25 cm of actual height (depending upon water body size) and surface area measurements within 15% of the actual area over the course of its 3-year mission life (Solander et al., 2016). While the SWOT mission will provide the most detailed and accurate assessment of the world's surface waters to date, its short mission life will, unfortunately, prevent it from providing climate-scale observations. This shortcoming creates a need for better strategies in assessing currently available datasets to analyze beyond the 3-year SWOT window as well as to truly assess the long-term dynamics of these critical water systems.

This chapter builds upon previous work by leveraging the Landsat 5 image archive in combination with bathymetric and digital elevation model data within Google Earth Engine (GEE), a cloud-based geospatial processing platform (Gorelick et al., 2016), to assess summer water surface elevation, surface area, and water volume from 1985 to 2009 for Lake McConaughy in Nebraska. Unlike many previous studies which assessed water indices and segmentation thresholds based upon their ability to correctly identify water from non-water, we use the water/land interface to estimate water surface elevation. While mixed water/land pixels are of concern in those studies, they are critical to identifying the water-land edge/boundary in our approach. Several common water indices (NDWI, MNDWI, AWEI) as well as some index combinations (NDWI + MNDWI) were analyzed and evaluated for their ability to predict water surface elevation from composite Landsat imagery using elevation statistics from lake water boundaries across a range of segmentation thresholds. The most accurate combination of variables and their thresholds, determined in comparison to daily gauged water surface elevation measurements, were then used to calculate surface elevation, surface area, volume, volume change, and lake cover frequency for four summer time periods (June to September) each year from 1985 to 2009. Additionally, linear regression was used to assess whether any long-term trends were present for the lake. Finally, this

chapter compares the results of the proposed technique to those using established techniques. To do so, water surface elevation was estimated using surface area via elevation/surface area relationships developed from the underlying bathymetry and compared to the proposed technique. All in all, this chapter comprehensively tests a wide range of water dynamic monitoring parameters, provides recommendations for application to other study areas, and compares the results to established techniques.

2.3 Study Area and Data

Lake McConaughy, the largest lake and reservoir in the state of Nebraska, is highly variable in both lake extent and volume over time (CNPPID, n.d.; Joeckel & Diffendal, 2004). The lake, located in western Nebraska (Figure 2-1), was formed following the completion of Kingsley Dam in 1941 as a means of storing irrigation water for what would eventually become the Central Nebraska Public Power and Irrigation District (CNPPID). Kingsley Dam is currently the second largest hydraulic-fill dam in the world and, in addition to providing irrigation water to the Tri-County region, contributes electricity via a hydroplant completed in 1984 (“Lake History,” 2016).

Beyond the hydroelectric and irrigation benefits provided by the reservoir, Lake McConaughy, or ‘Big Mac’ as it is otherwise known, is also a highly popular recreation destination. Depending upon the water level, the reservoir features white-sand beaches and numerous swimming, water sports, boating, fishing, hunting, and camping opportunities. As mentioned previously, the water-level of the reservoir can vary significantly from season to season or even month to month. This is due to variable inflows from the North Platte River as well as seasonal water requirements, chiefly for irrigation during the summer growing season. Despite being the largest water body in the state of Nebraska, Lake McConaughy is a poor candidate for monitoring via satellite altimetry as it has received limited crossovers from altimeter instruments (Figure 2-1). Of the available sensors, only RA-2 (Radar Altimeter 2) onboard Envisat in its 2010 to 2012 (end of mission life) orbit routinely crosses the main body of the reservoir (“Pass locator: Aviso+,” n.d.). However, altimeters are non-imaging, profiling instruments

which collect information from all the objects within their footprint simultaneously (Sulistioadi et al., 2015). As a result, all altimeter measurements for Lake McConaughey, including Envisat's RA-2 which has a 20 km IFOV (instantaneous field of view), would likely contain significant non-water contamination ("RA-2 - Earth Online - ESA," n.d.).

The primary source of data for this project was Google Earth Engine (GEE) which maintains the entire USGS Landsat 5 TM Collection 1 Tier 1 Raw Scenes image archive (1984-2013) with a moderate spatial resolution of 30 meter as well as the 30-m Shuttle Radar Topography Mission (SRTM) DEM which provides integerized elevation data for 80% of Earth's landmass (Farr et al., 2007). The study area, while smaller than an individual Landsat scene, is split between two scene paths (WRS 31/31 and 31/32) and requires imagery from both to cover the entire reservoir. Daily lake elevation gauge measurements were retrieved from the Central Nebraska Public Power and Irrigation District in tabular form (H. Rahmann, personal communication, January 9, 2017) and Lake McConaughey bathymetric contours, which were collected as part of a study performed in 2003, were downloaded from the United States Geological Survey (USGS) (Kress, Sebree, Littin, Drain, & Kling, 2005).

Relatively low water levels at the time of the bathymetry survey necessitated extending the topographic model of Lake McConaughey via merging bathymetry with a supplemental DEM. Our tests indicated slightly better agreement along the land/water interface between the bathymetry and SRTM DEM than the more detailed National Elevation Dataset (NED). To merge the bathymetry with the SRTM DEM, Lake McConaughey bathymetric contours were downloaded, processed, and resampled into a 30 meter raster to match the spatial resolution of the SRTM DEM. The overlap area in the SRTM DEM was replaced by the USGS bathymetry data (Figure 2-2).

Ideally, data for both bathymetry and the surrounding terrain elevation should be collected simultaneously to ensure continuous coverage from under-water to above water. In reality, the time difference between the SRTM mission (February 2000) and the bathymetric survey (Spring 2003) resulted in lake elevation dropping from a February 2000 average of 992.73 meters to a Spring 2003

average of 983.52 meters. Due to the higher water level in the SRTM DEM than the bathymetric survey, some stretches of land that were exposed in 2003 but submerged in 2000 were left unsampled and required filling via inverse distance weighting (IDW).

2.4 Methods

Images intersecting the study area (Figure 2-1) were first retrieved from the USGS Landsat 5 TM Collection 1 Tier 1 Raw Scenes image archive within Google Earth Engine. Images were then filtered based upon their cloud score (USGS, 2018b), sorted into time periods, and processed from digital numbers (DN) into top-of-atmosphere (TOA) composites with per-pixel cloud filtering for each time period. Once the composite images were generated, lake water was identified based on water index images which were then segmented to create binary images with values of 1 representing water and values of 0 representing land areas. Since the focus of this research is upon one singular body of water, disconnected water bodies were eliminated from the analysis.

Following the removal of disconnected water bodies, a one pixel radius square kernel was applied over the binary image to extract lake water boundary, which is a ring of pixels around the edge of the lake water body representing the interface between land and water. Water surface elevation was then calculated with statistics (mean, median, or mode) from the bathymetry/elevation values of the pixels within the water boundary.

Once the water surface elevation was determined, surface area, volume, and volume change were then calculated. Surface area calculation was completed with the binary water image by first counting the number of water pixels and then multiplying the count by the area of a pixel. Similarly, volume calculation was done by masking out the non-water areas from the merged DEM leaving just the pixels covered by water. The elevation of each pixel was then subtracted from the previously calculated water surface elevation to determine the water depth at each pixel. The depth at each pixel was then multiplied by its surface area to determine the water volume of each pixel before being summed to determine total

water volume. Finally, water volume change was calculated by subtracting the total water volume of each time period from the successive time period.

While the general processing workflow is outlined above, several key details deserve further explanation including: constructing a time series from Landsat 5 images, lake water identification, and water surface elevation estimation.

2.4.1 Constructing a Time Series from Landsat 5 Images

This research spanned twenty-five years from 1985 – 2009 and utilized most of the available images in the Landsat 5 archive. Due to the climate of the study area, the analysis was limited to warm weather months between May 1st and October 31st in order to avoid complications from ice and snow which would hinder water detection efforts in many early spring, late fall, and winter images. A 3-month temporal window was also used to create four analysis time periods each year: June (May 1st – July 31st), July (June 1st – August 31st), August (July 1st – September 30th), and September (August 1st – October 31st). While using a time window means that some images might be used in multiple time periods in a year, it was expected that the compositing process would have enough images spread across the entire three-month window to capture a median value reflective of the central month for that time period (more on the GEE compositing process in subsequent paragraphs). The results of the analysis confirmed this expectation as the values calculated for each time period are quite distinct.

Beyond that concern, the benefit of the temporal window was two fold. First, a larger date range was necessary to ensure that enough images were available for analysis within each time period once images were filtered for cloud cover (discussed in the following section). This is especially important not only due to Landsat's 16 day revisit period but also because the study area crosses more than one Landsat scene. In an early test using single months (June, July, August, and September) as time periods, half of the reservoir was completely devoid of imagery in several months once scenes were filtered by their cloud scores. Secondly, by overlapping the temporal windows of the time periods, a greater number of

analysis periods are available each year which allows for a more complete picture of water dynamics within the reservoir.

Clouds represent a potentially large source of noise in most remote sensing applications using Landsat images. To address this concern, two cloud removal steps were implemented during the image composition process. First, images with an image cloud score greater than 50% were automatically excluded from the analysis. The image cloud score is a value assigned to each Landsat image that indicates the percentage of a scene which is covered by cloud. While this metric is useful in identifying cloudy scenes, it does not assess the location or distribution of cloud cover within an image meaning that in some images the object of interest could still be observed and should be included in the analysis even though other areas of the image are obscured by cloud. For this research, tests indicated slightly better results when scenes with image cloud scores greater than or equal to 50% were excluded from the analysis as images with larger values were much more likely to obscure large portions of the reservoir.

As mentioned, the image cloud score alone is a poor measure of the cloud distribution within a Landsat scene. While removing images with cloud scores greater than or equal 50% eliminated a large amount of cloudy images, the remaining images could still contain extensive amounts of cloud cover capable of negatively impacting the analysis. To further address this, a per-pixel cloud score generated in GEE during the image composite process was utilized to assess the relative likelihood of a pixel representing cloud cover using a combination of brightness, temperature, and the Normalized Difference Snow Index (NDSI) (Google Earth Engine, n.d.; Gorelick et al., 2016). While this method is not a robust cloud detection algorithm, it does serve as a simple method for assessing the likelihood that an individual pixel is cloudy. In this research, pixels with relative cloud scores greater than 10 were eliminated leaving the remaining pixels available for image composition. Per-band percentile values at each pixel were then computed from all the remaining pixels to form the final composite image. For this research, all composite images for the time periods were assigned to the the 50th percentile value to represent the median central tendency for each time period.

Ideally, per-pixel cloud scores would eliminate the need to filter individual scenes by their image cloud scores. However, early tests which used per-pixel cloud score exclusively to address cloud cover were less accurate overall than using a combination of the two methods where the cloudiest scenes were completely eliminated from consideration (more in Discussion section).

2.4.2 Lake Water Identification

In comparison to water identification techniques requiring extensive field observations, training, and validation (such as supervised image classification), segmentation based techniques, such as the water indices used in this analysis, are very computationally efficient which is especially useful for analyzing multiple images spanning long time periods. Several water indices were analyzed in this research including Normalized Difference Water Index (NDWI) (McFeeters, 1996), Modified Normalized Difference Water Index (MNDWI) (Xu, 2006), Automated Water Extraction Index (AWEIsh and AWEInsh) (Feyisa, Meilby, Fensholt, & Proud, 2014) and two combined indices formed from NDWI + MNDWI (B1 & B4) and NDWI + MNDWI (B2 & B5). These water indices rely upon the spectral properties of water, most notably its strong absorption in near-infrared and shortwave-infrared wavelengths. Table 2-1 provides the band designations for Landsat 5 Thematic Mapper while Table 2-2 provides the formulas for each of the water indices.

In spite of computational efficiency, selecting an optimum segmentation threshold for a given analysis can be a difficult process. Atmospheric conditions, water/land composition such as water depth, turbidity and emergent/submerged vegetation even temporal changes such as seasons or even time of day can all influence the optimum segmentation threshold for any single image. This problem compounds itself in cases where multiple images in a long time period are used such as in this study. In addition to this, the different water indices used in this analysis also vary in their requirements for selecting an optimum segmentation threshold. For example, AWEI attempts to produce a stable segmentation threshold (at or near zero when additional atmospheric corrections are undertaken) (Feyisa et al., 2014) while NDWI and MNDWI do not.

While a number of factors affect optimum segmentation threshold selection, we assume that segmentation thresholds could be identified which minimize measurement error over the entire analysis period rather than attempting to minimize error for a singular time period. In other words, our goal was to identify a threshold or thresholds to accurately estimate water dynamics for the entire image collection in the study period using minimal computational effort rather than find the optimum segmentation threshold for each individual image composite. In order to accomplish this, while searching for the optimum thresholds, the analysis was completed with an increment of .01 ranging from -0.35 to +0.75 for each analysis parameter setting.

Once the water index images were segmented using the selected thresholds, disconnected water bodies were eliminated from binary images using a vector intersect approach which was performed by seeding a small polygon (a point or line could also be used) within the main body of the reservoir in an area known to contain water throughout the entire study period. Water areas in the binary water images were vectorized and intersected with the seeded-polygon to select the water body of interest which was then re-rasterized to create a new binary water image (Figure 2-3).

2.4.3 Lake Water Surface Elevation Estimation

We used kernel-based morphological operations to delineate the lake water edge with the goal of using the water/land boundary to determine water surface elevation. Three potential types of boundaries were identified; interior boundary, exterior boundary, and combination boundary (Figure 2-4). Three boundary types were analyzed because approximating the actual land/water interface on a raster grid is difficult. On a binary water image, when ignoring the effect of mixed pixels, interior boundaries represent the first ring of water pixels while exterior boundaries represent the first ring of land pixels. The combination boundary, on the other hand, uses both land and water pixels to better approximate the land/water interface. With that said, a variety of factors such as local slope, mixed pixels, and water detection accuracy could all impact the ideal water boundary type making a full testing of the available types necessary. Interior boundaries were created by eroding the water pixels on the binary land/water

image by one pixel while exterior boundaries were created by dilating the water pixels by one pixel. Combination boundaries were extracted using both methods. While the exact steps varied to some degree depending upon the boundary types used, the boundaries were used to retrieve the elevation from the merged bathymetry/DEM terrain model to capture water surface elevation within one pixel of the water's edge.

The standard statistics of mean, median, and mode were used to calculate a single water surface elevation from the elevations retrieved from the land/water boundary. Shoreline topography can be very diverse. In some cases the interface between land and water is a very gentle slope, such as in many beach areas, while other areas can have significant changes over short distances. In areas with low-slopes, the depth of water is very shallow making the water surface elevation and the ground elevation essentially the same. Conversely, elevation can differ substantially in areas of steep slope. As such, by using all of the values along the shoreline, a single representative value for the water surface elevation can be estimated.

2.4.4 Validation

Analysis results were assessed by root-mean-squared-error (RMSE) (Table 2-3) to determine the best performing parameter combination's ability to estimate surface water elevation for the entire analysis time period. For consistency, and to compare a similar number of data points, the true surface elevations for each time period were calculated using only the dates common to both the daily gauged surface elevations and the Landsat 5 image archive (there are 90+ gauged measurements for each time period versus 12 or fewer Landsat images). The median daily elevation of those common dates was then used as the true surface elevation for each particular time period. This approach eliminated the influence from events not captured in the Landsat archive which could have unfairly weighted the gauged surface elevations due to their greater observation density (which increases their chances of capturing events, such as floods, which may or may not be visible within the Landsat image archive due to the 16-day revisit period). RMSE was then calculated using the water surface elevation estimated from image analysis and the true water surface elevation. Additionally, linear regression was used to assess whether

any long-term trends were present for Lake McConaughy over the analysis period. Additional error statistics such as mean absolute error (MAE), mean bias error (MBE), and mean absolute percentage error (MAPE) were calculated for the best performing parameter combination (Table 2-3). Bootstrap resampling (1,000 repetitions) was used to calculate each statistic and generate 95% confidence intervals.

Water volume change was evaluated in a similar manner by comparing it to ground-truth using error statistics. Additionally, water volume change was also validated against the water volume change calculated using the following pyramidal frustum equation which is a common technique used in altimeter based studies (J F Crétau et al., 2016):

$$\Delta V = ((H1 - H0) \times (A1 + A0 + \sqrt{(A1 \times A0)})) / 3 \quad (1)$$

where ΔV is the change in volume between two dates, $H1$ and $H0$ are the water surface elevations for the two dates, and $A1$ and $A0$ are the corresponding surface areas for those same dates.

Finally, water surface elevation accuracy was compared to elevation estimates derived from surface area to elevation relationships to establish the performance of this method relative to that of an established alternative. First-order (linear), 2nd-order polynomial, and 3rd-order polynomial equations (Duan & Bastiaanssen, 2013) were developed from the USGS provided elevation/surface area/volume table. Water surface areas for each composite image was then used to estimate elevation using the surface area to elevation relationship.

2.5 Results

All in all, we tested over 5,994 different parameter combinations (water index, segmentation threshold, boundary type, statistical type) to determine the best parameter combination for predicting the water surface elevation, surface area, and volume of Lake McConaughy (Table 2-4). Each set of parameters was mapped over 100 composite images that were generated from 597 unique Landsat scenes (after cloud score filtering) from a total of 651 scenes available in the study area during the study period.

Images were counted a single time even though many were used more than once due to the overlapping time windows.

RMSE was calculated for each parameter combination to allow for accuracy assessment and comparison. Overall, the best performing parameter combination for predicting water surface elevation was NDWI + MNDWI (B1 & B4) with a segmentation threshold of 0.06 using a combination boundary and the mean statistic which produced a RMSE of 0.768 meters $CI_{95\%}$ [.657, .885]. Table 2-5 shows the number of combinations by water index with RMSEs better (lower) than some selected values. This table indicates that of the 5,994 total combinations tested, 803 of them produced RMSEs less than 1.0 meter (about 13.4%). While this is a low percentage overall, a closer look indicates that the majority of these low RMSEs were concentrated in the NDWI + MNDWI (B1 & B4) and NDWI + MNDWI (B2 & B5) indices which together account for 519 (64.6%) of the 803 total parameter combinations which have a RMSE below 1 meter. NDWI + MNDWI (B1 & B4) proved to be the most accurate index overall with 10 parameter combinations having RMSEs better than 0.80 meters.

In addition to the table, the RMSEs for all parameter combinations were compiled into a series of graphs to further highlight and illustrate the impact each parameter has on overall water surface elevation accuracy and allow the comparison between the different water indices, segmentation thresholds, and boundary types (Figure 2-5). The graphs span the entire range of segmentation thresholds tested (-0.35 to +0.75) but were capped at a RMSE of 2.0 meters in order to focus upon the most accurate combinations. The graphs reveal some key patterns and relationships that will be useful in future research. One such pattern is the relationship between boundary types and their lowest RMSE values (or the curve in general). Interior boundaries have their lowest RMSEs at lower segmentation thresholds than exterior boundaries while combination boundaries fall in the middle. For example, the bottom of the mean and median NDWI curves are centered around a segmentation threshold of -0.05 while the exterior boundary is centered around 0.10 with the combination boundary falling in between around 0.0. The graphs also highlight a maximum threshold limit of +0.19 for AWEIsh (Figure 2-5d) regardless of the boundary or

statistic type used as well as the relatively poor performance of AWEInsh in water surface elevation estimation.

Another key observation stems from the type of statistics used in the analysis. Each graph reveals roughly the same pattern where mean has the narrowest RMSE curve, followed by median, and then mode with the widest RMSE curve. Table 2-6, which shows the number of segmentation thresholds broken down by water index and boundary/statistic type with RMSEs less than 1.0 meters, further confirms this observation. Mode, while generally being more stable across a range of thresholds, often performs at a lower accuracy than mean or median. To further illustrate, there are 3 median combinations, 14 mean combinations, and zero mode combinations with a RMSE better than 0.80 meters. Additionally, Table 2-6 indicates better accuracy with interior (372 sub-meter combinations) and combination (336 sub-meter combinations) boundary types compared to exterior boundaries (155 sub-meter combinations) across the range of tested parameter combinations.

Once the best performing parameter combination (NDWI + MNDWI (B1 & B4), combination boundary, mean statistic, 0.06 segmentation threshold) for predicting water surface elevation was determined additional water estimates for the reservoir were also calculated using those parameters. Figure 2-6 plots the water surface elevation, surface area, and volume of the reservoir for each month during the study period. Water surface elevation was plotted against the gauged-measurements provided by the CNPPID while surface area and volume were plotted against surface area and volumes provided or interpolated, where necessary, in a Lake McConaughy look-up table from the United States Geological Survey developed from the same bathymetry data used in this study (USGS) (Kress et al., 2005). Each figure was also fit with a trend-line which indicates a moderate negative trend in each of the measurements over the course of the study period. It also be noted that each year has its own localized trend with the peak elevation, surface area, and volume occurring early summer and then decreasing steadily until the fall.

In spite of the differences, Figure 2-6 indicates a good relationship between the predicted values and the actual gauge-based values (median values calculated from the daily gauge values and the look-up table). Table 2-7 provides accuracy assessment statistics for the each of the three estimates.

The estimated water surface elevation, surface area, and volume during the study period peaked in June, 1986 at a water surface elevation of 994.54 meters. However, the actual peak water surface elevation, surface area, and volume occurred in June, 1997 at a water surface elevation of 995.02 meters. This error occurs around the SRTM/bathymetry interface and is likely a result of undersampling in those areas as well as estimation bias (see discussion). At the other end of the spectrum, September, 2006 had the lowest lake levels for both the estimated (975.91 m) and actual (975.21 m) measurements. Figure 2-7 provides a longitudinal view of the reservoir to further illustrate the water disparity between the maximum and minimum surface elevations. From Lake McConaughy's peak surface elevation calculated in June, 1986 to its minimum in September, 2006 the western portion of the reservoir retreated nearly 14.5 kilometers (visible in Figure 2-7b).

In terms of surface cover frequency, 44.04% (~52.25 km²) of the max reservoir extents was covered by water during all 100 time periods (blue color in Figure 2-8). Conversely, 0.76 km² (0.64%) was covered by detected water just a single time.

Finally, water volume change was calculated. Table 2-7 shows error statistics for our method, as well as for water volume changes calculated using the pyramidal frustum method (Eq. 1) using our estimated surface area and surface elevation. Overall, we calculated a total volume change between the maximum and minimum lake levels of 1.568 km³ compared to an actual change of 1.659 km³. At its minimum surface level, the reservoir contained just 23.97% of its maximum volume.

For one final comparison, water surface elevation for each time period using surface area to elevation relationships built from the underlying bathymetry. Using this method, a maximum accuracy of

0.824m RMSE $CI_{95\%}$ [0.706, 0.934] was achieved using linear regression (1st-order), a 0.10 segmentation threshold, and the NDWI + MNDWI (B1 & B4) water index combination.

2.6 Discussions

2.6.1 Composite Images and Cloud Cover

During the data exploration stage of this project, we noticed reduced accuracy when using all of the available imagery during the study period. We originally thought per-pixel cloud scores would eliminate the need to filter individual images by their overall image cloud scores, but eliminating cloudier images actually improved the results. This is due, in part, to the relationship between the composite images and the ground-truth data. The ground-truth for each time period is the median gauged surface elevation for each image date. When cloudier images are included in the analysis, the lake is more likely to be obscured resulting in fewer pixels from that date being included in the analysis which weights the analysis in favor of less cloudy images. The ground-truth on the other hand is not subject to cloud-cover. In other words, cloudier images result in more data points being used to generate the ground-truth value than are being used in the estimation process.

2.6.2 Segmentation Thresholds, Boundaries, and Statistics

As mentioned previously, the RMSEs graphs (Figure 2-5) reveal several interesting relationships between peak accuracy, segmentation threshold, and boundary types. There appears to be a slight shift in the segmentation thresholds which produce accurate results among the three different boundary types with interior boundaries typically peaking at a slightly lower threshold than combination or exterior boundaries. Figure 2-4 illustrates the cause of this difference. Assuming the same index, segmentation threshold, and statistic are applied, an exterior boundary will be displaced one pixel outward from an interior boundary. In this case, a displacement of one pixel outward for the exterior boundary will likely extract pixels of a higher elevation than would be extracted for the interior boundary. For an exterior boundary and an interior boundary to accurately calculate water surface height, an exterior boundary

requires a smaller water surface area to extract the same set of pixels as the interior boundary. Therefore, the peak performance of an exterior boundary would be shifted to slightly higher segmentation threshold than an interior boundary. The segmentation threshold with the best accuracy for combination boundaries typically falls between the best segmentation thresholds for interior and exterior boundaries. As mentioned, one implication of this relationship is that when both boundary types are at the same surface elevation, the interior boundary will therefore provide a larger surface area and a larger lake volume than the exterior boundary, while again, the combination boundary type results would fall in the middle.

Figure 2-5 also reveals an interesting relationship between segmentation threshold and statistic types. Of the three central tendency statistics, mean is the most susceptible to outlier influence followed by median. Due to varying topography around the reservoir, elevation values captured by the water boundary are sure to include values significantly above or below the actual water surface. Mode, on the other hand, assesses central tendency by determining the most frequently occurring value within the boundary. In this study, while not necessarily accurate overall, mode statistics often produce a significantly wider RMSE curve than mean or median. Mode is able to better leverage the common pixels within the water boundary at each segmentation threshold increment (.01) to a much greater degree than mean or median resulting in a greater range of acceptable segmentation thresholds that can be used to estimate water surface elevation. However, one caveat of this comes into play when attempting to calculate surface area and volume. For example, using mode, the estimated surface elevation may be very similar using a segmentation threshold of 0.05 or 0.15. However, the 0.05 segmentation threshold still means a greater number of pixels were identified as water when compared to the 0.15 segmentation threshold which would result in a larger surface area and volume in spite of similar surface elevation. Another potential caveat of using mode statistics could arise when using higher resolution bathymetry/DEMs. With sufficient elevation measurement precision fewer values would repeat potentially leading to less stable mode estimates.

Additionally, it should be noted that this method may be ill-suited for canyon-filled lakes or reservoirs. If the topography near the shoreline is sufficiently steep then the water surface elevation can diverge significantly from the land elevation above or below water if the DEM/bathymetry does not possess a high enough spatial resolution since this technique does not directly measure the water level itself (such as is done with an altimeter), but rather it assesses the elevation of the shoreline (above or below the water).

2.6.3 Water Index Performance

Table 2-5, Table 2-6, and Figure 2-5 each highlight a disparity in water index performance across the range of tested parameter combinations. As mentioned, the combined indices, NDWI + MNDWI (B1 & B4) and NDWI + MNDWI(B2 & B5), had better accuracy across a wide range of segmentation thresholds in comparison to MNDWI, NDWI, AWEInsh, and AWEIsh. The improved results of the combined NDWI and MNDWI indices may be due to increased separation between the water and non-water classes within the image as noted by Acharya et al. (2017) and Lu et al. (2011) in their study of combined water indices. Conversely, the poor performance of AWEInsh and AWEIsh is likely due to a few factors. In terms of this research, the primary cause could be a lack of the additional preprocessing and atmospheric corrections steps undertaken by Feyisha et al. (2014) to improve image quality and produce a stable segmentation threshold at or near 0. In our study, AWEIsh and AWEInsh had the narrowest RMSE curves (Figure 2-5) and produced the fewest number of sub-meter surface elevation accuracies of any of the tested water indices. Peak performance for AWEIsh occurred around -0.10 segmentation threshold while AWEInsh generally improved in accuracy until an abrupt decrease in accuracy at +0.19 segmentation threshold. In any case, the peak accuracy performance thresholds for the two AWEI algorithms are significantly different than the stable 0 threshold envisioned by Feyisha et al. (2014) Further research is needed to determine whether or not the additional preprocessing procedures undertaken by Feyisha et al. would improve the results from the indices.

2.6.4 Identification of Water Body Interest

A few potential methods to identify the main water body of Lake McConaughy were examined. In addition to the vector intersect method ultimately used in this study, the cumulative cost (or cost distance) method, was also tested and found to be a functional, capable method. The cumulative cost method within GEE does require some prior knowledge about the max dimensions of the area being investigated as it requires a maximum distance parameter to perform the calculation. If the max distance parameter is set too small, the analysis would stop short of capturing the entire water body. Also, if the water body were of sufficient size, using a large max distance could result in memory errors (Gorelick et al., 2016).

The vector intersect method used in this project is a multi-step process where all of the water bodies within the study area are vectorized. Like the cumulative cost method, the vectorization process can be memory intensive so one potential issue with this method could stem from particularly large water bodies with sufficiently complex shorelines.

2.6.5 Water Volume Change

The water volume change calculated using our method agrees well with using the pyramidal frustum equation. Despite identical error statistics, our method estimates a maximum volume change of 1.568 km³ between the maximum and minimum water levels during the study period compared to a pyramidal frustum volume change estimate of 1.541 km³ and an actual volume change of 1.659 km³. The decreased accuracy of the pyramidal frustum method may be due to the oversimplification of the lake bottom morphology assumed by the method (Hollister & Milstead, 2010). Also, with an RMSE of 0.074 km³ CI_{95%} [0.064, 0.084], our method outperformed the the 0.6 km³ RMSE Crétaux et al. (2015) obtained with a combination of Landsat, MODIS, and altimetry , for a reservoir with overlapping water volume ranges with Lake McConaughy. While some of the error differences between our method and that of Crétaux et al. (2015) may be attributable to differences in the lakes themselves, the Crétaux et al. method is also subject to error due to the surface area measurements used to construct the surface area/elevation

relationship (coarse MODIS spatial resolution), temporal misalignment of altimeter measurements and surface area measurements from Landsat imagery (passovers occurred on different days), and/or the varying accuracy of the multiple altimeters used in the study. In our study, volume estimation has a MAPE of 4.856% which is greater than the MAPE of either surface area (3.095%) or surface elevation (0.061%) indicating error propagation throughout the calculations. Increased accuracy for either surface elevation or surface area should improve the accuracies on water volume and water volume change.

2.6.6 Surface Area to Elevation Relationships

A common method for estimating water surface elevation using optical or SAR imagery is through the use of surface area to elevation relationships developed using the underlying bathymetry. In this scenario, using regression equations developed from the bathymetry, elevation can be estimated using surface area. The results of this analysis indicate that our method using water boundaries outperforms surface area to elevation relationships (0.768 meters $CI_{95\%}$ [0.657, 0.885] for our method versus 0.824m RMSE $CI_{95\%}$ [0.706, 0.934] for surface area to elevation relationships). We expect that this is the result of using many values (thousands of shoreline elevation values) compared to a single surface area value for estimation. Over long periods of time, erosional and depositional processes may change the underlying bathymetry and cause areas of the lake to expand (increased surface area) or contract (decreased surface area). This change is then captured in the remote sensing imagery, but is not accounted for in the bathymetry unless additional surveys are conducted. In this scenario, areas of change have a lower impact on the elevation estimate using shoreline boundaries as they are buoyed by the remaining lake shore compared to surface area to elevation relationships where the lost surface area will immediately result in lower elevation and volume estimates.

2.6.7 Time Series Implications

Over the course of the study period, two key patterns are visible in Figure 2-6. First, each year there is a drawdown of the reservoir levels over the course of the summer. The annual summer drawdown occurs because Lake McConaughy was built for irrigation and the CNPPID is required to release water to

irrigators upon request regardless of other recreational or environmental needs (Commission, n.d.). In terms of the summer drawdown, Figure 2-6 reveals some biases in the optimal segmentation threshold results in comparison to the measured ground-truth data. In the first low period (1989-1994), surface elevation, surface area, and volume are overestimated early in the year before, generally, becoming more accurate in later months. This bias pattern then again repeats itself at the beginning of the second, larger prolonged low period (2000-2009) before the bias pattern reverses once the reservoir reservoir falls to lower levels (2003-2008). These biases, as well as some of the error at the highest water levels, likely occur because a single segmentation threshold is being used for the entire analysis period. At high levels, the mixed pixels along the shoreline consist of shoreline and emergent vegetation which interfere with water classification. At low lake levels, the spectral properties of the shoreline change in response to increasing beach area and changing water quality conditions. In essence, the selected model performs better at some water levels than others due to changing spectral properties related to the surrounding non-water landcover types. Future research, which will include dynamic thresholding and other strategies, should address this issue. As covered in the results, several other parameter combinations also performed well with nine combinations exceeding 0.80m RMSE and a total of 803 combinations better than 1.0m RMSE. The total number of combinations with sub-meter performance illustrates some flexibility in the model selection. However, it is important to acknowledge that each parameter combination has specific water levels in which the model is most accurate and that other water levels will exhibit some bias by either overestimating or underestimating water surface elevation. The strong performance of multiple models and segmentation thresholds further indicates the gains that could be achieved through dynamic thresholding techniques.

Figure 2-6 also shows a second, long-term drawdown of Lake McConaughy as a result of extended drought periods which result in decreased inflows and increased irrigation requirements (Joeckel & Diffendal, 2004). Similar to the annual drawdown, but on a larger scale, the long-term drawdown has a direct impact upon recreational activities and ecological habitat. The number and location of available boat docks, as well as the lake's carrying capacity of sportfish, change in response to lake water levels.

At the lowest levels, fish mortality is a serious concern as poorer water conditions and potential toxic blue-green algae blooms place increased physical stress on fish and other organisms (Commission, n.d.).

2.7 Conclusions

The methods developed in this chapter have shown great promise for studying long-term lake water dynamics. While the ability to accurately estimate water surface elevation relies upon water indices and selecting an appropriate segmentation threshold, the use of appropriate boundary types and statistical measures can increase the range of acceptable values considerably. Regardless of water indice or boundary type, estimates were least sensitive to changes in segmentation threshold using the mode statistic, followed by median, and finally mean; however at optimal segmentation thresholds mean and median provide significantly better accuracy. NDWI + MNDWI (B1 & B4) had the highest accuracy of the tested indices with 10 combinations having a RMSE better than 0.80 meter. Overall, 803 of the tested parameter combinations produced RMSEs within 1.0 meter of in-situ gauge measurements with the lowest RMSE being 0.768 meters $CI_{95\%}$ [0.657, 0.885] produced from NDWI + MNDWI (B1 & B4) with a combination boundary, mean statistic, and 0.06 segmentation threshold. In general, if applying this method to ungauged water bodies we would recommend either of the NDWI + MNDWI water indices along with either an interior or combination boundary and the mode statistic. However, if ground-truth is available, we would recommend training the model on the available data to further optimize the selected model. Although altimeters are capable of calculating surface elevation at the sub-decimeter level, this level of accuracy is generally reserved for significantly larger water bodies with favorable shape, area, and topography. While reliant upon existing bathymetry, the proposed methods in this chapter would expand lake dynamic studies to additional lakes for which bathymetry exists but altimeter observations are unavailable. In the state of Kansas, for example, bathymetric surveys have been conducted for approximately 80 lakes and reservoirs, yet relatively few of these lakes are suitable to altimeter based studies (“Kansas Lakes and Reservoirs,” 2016). The method could also be used for independent verification of new sensors and missions studying inland water dynamics. Furthermore, this method

modestly outperformed surface area to elevation relationships with our method producing accuracies as high as 0.768 meters $CI_{95\%}$ [0.657, 0.885] compared to a high accuracy of 0.824m RMSE $CI_{95\%}$ [0.706, 0.934] for surface area to elevation relationships.

Additionally, our study successfully revealed the seasonal patterns of Lake McConaughy (highest level in the Spring and lowest in the late Fall in response to irrigation water diversions during the Summer), as well as recognized the moderate long-term trend present over the 25 year time period (R -squared = 0.547 for lake surface elevation) which has major implications for not only agriculture in the surrounding area but also recreational activities, fish habitat, and water quality within the lake itself. All in all, the results of lake surface elevation, surface area, and volume were in excellent agreement with ground-truth values.

While the methods showed promise in tracking long-term lake dynamics, it is expected that higher resolution imagery, DEMs, and bathymetry would increase the accuracy of our method as well as improve the ability to calculate volume change between time periods. One potential source of error in this particular analysis likely occurs along the interface between the SRTM DEM and bathymetry data due to the bathymetry being collected at a lower lake surface elevation than was present during the SRTM mission. Additionally, improved cloud filtering techniques would also boost the final analysis and dynamic segmentation thresholding techniques could improve accuracy across various surface elevation zones. Finally, this study also assumes a static reservoir bottom for volume calculations. In reality, this reservoir, like many around the world, has been in-filling with sediment over the entirety of its life.

Future research will focus upon improving DEM/bathymetry merging techniques, more robust cloud filtering techniques including FMask (available in the Landsat QA bands), methods of estimating sediment in-filling, as well as utilizing higher-resolution and/or temporal resolution imagery products to increase the number of available observations and estimation accuracy. Overall, the long-term monitoring of water dynamics has the potential to improve water resource management as well as

increase our understanding of temporal changes in water quantity distribution and its impacts upon water-dependent phenomena.

2.8 Figures and Tables

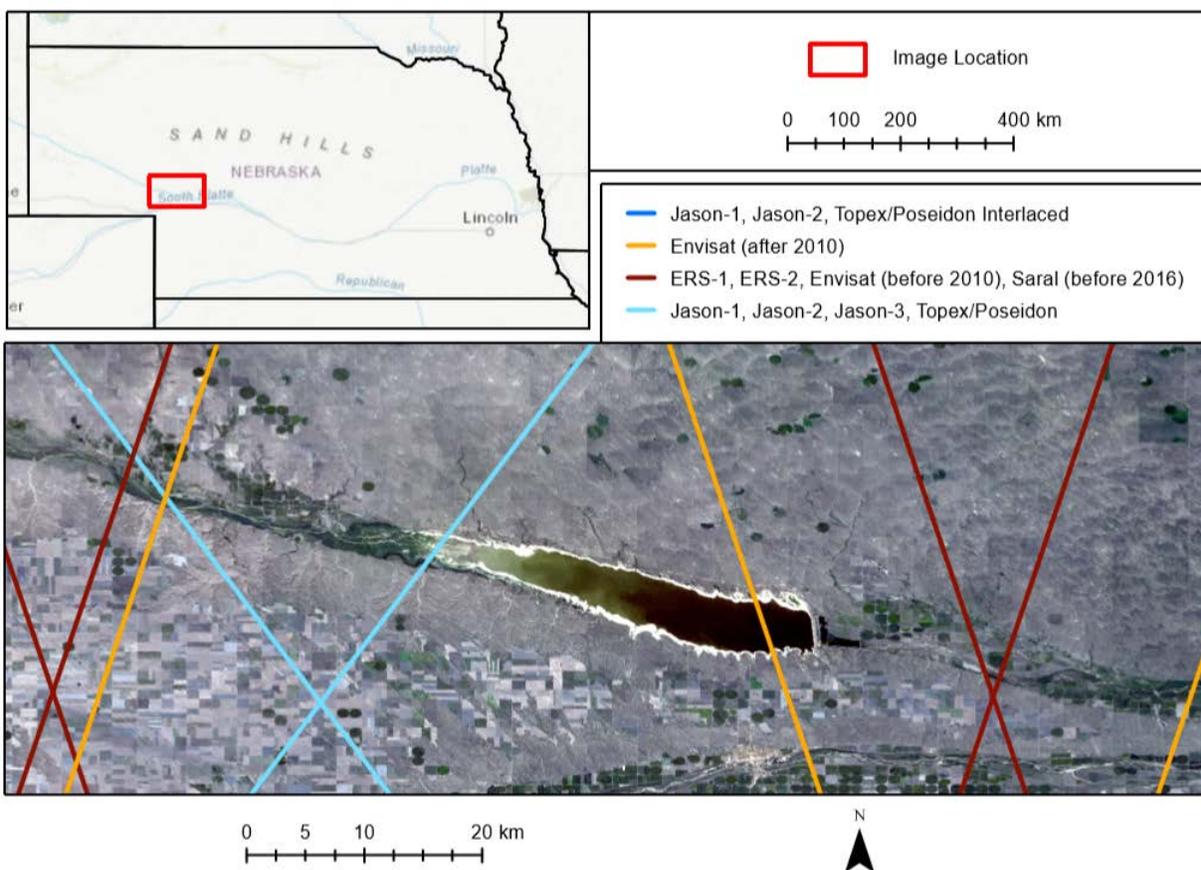


Figure 2-1: Lake McConaughy and its location in the state of Nebraska along with orbital paths of altimetry missions commonly used for inland water surface elevation monitoring within the vicinity of the reservoir. Of the available sensors, only RA-2 (orange color) onboard Envisat routinely crosses the main body of the reservoir but contains significant noise. The background image is a natural-color USGS Landsat 5 TM Collection 1 Tier 1 Raw Scene from September 2nd, 2001.

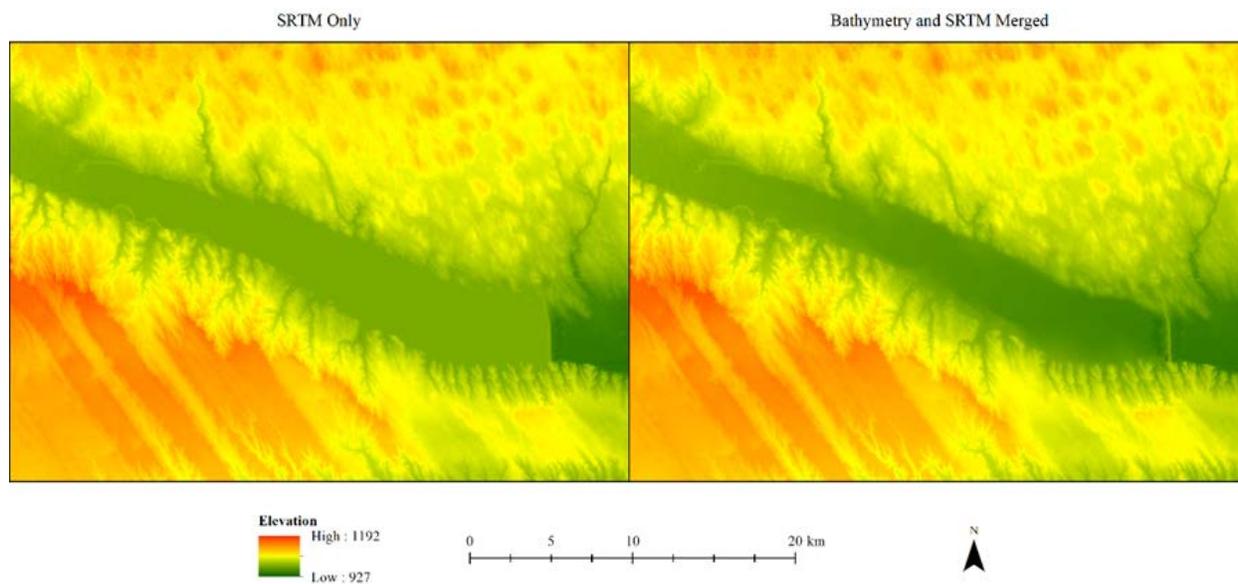


Figure 2-2: SRTM DEM (left) and bathymetry merged with SRTM DEM (right).

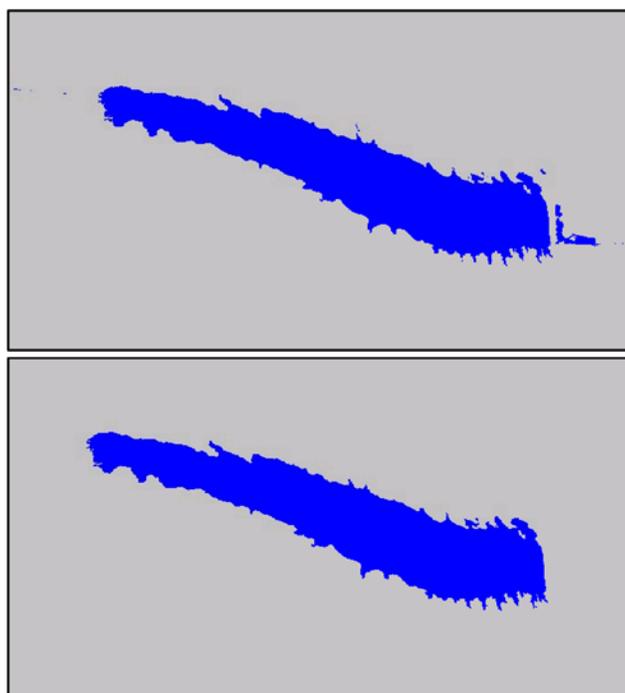
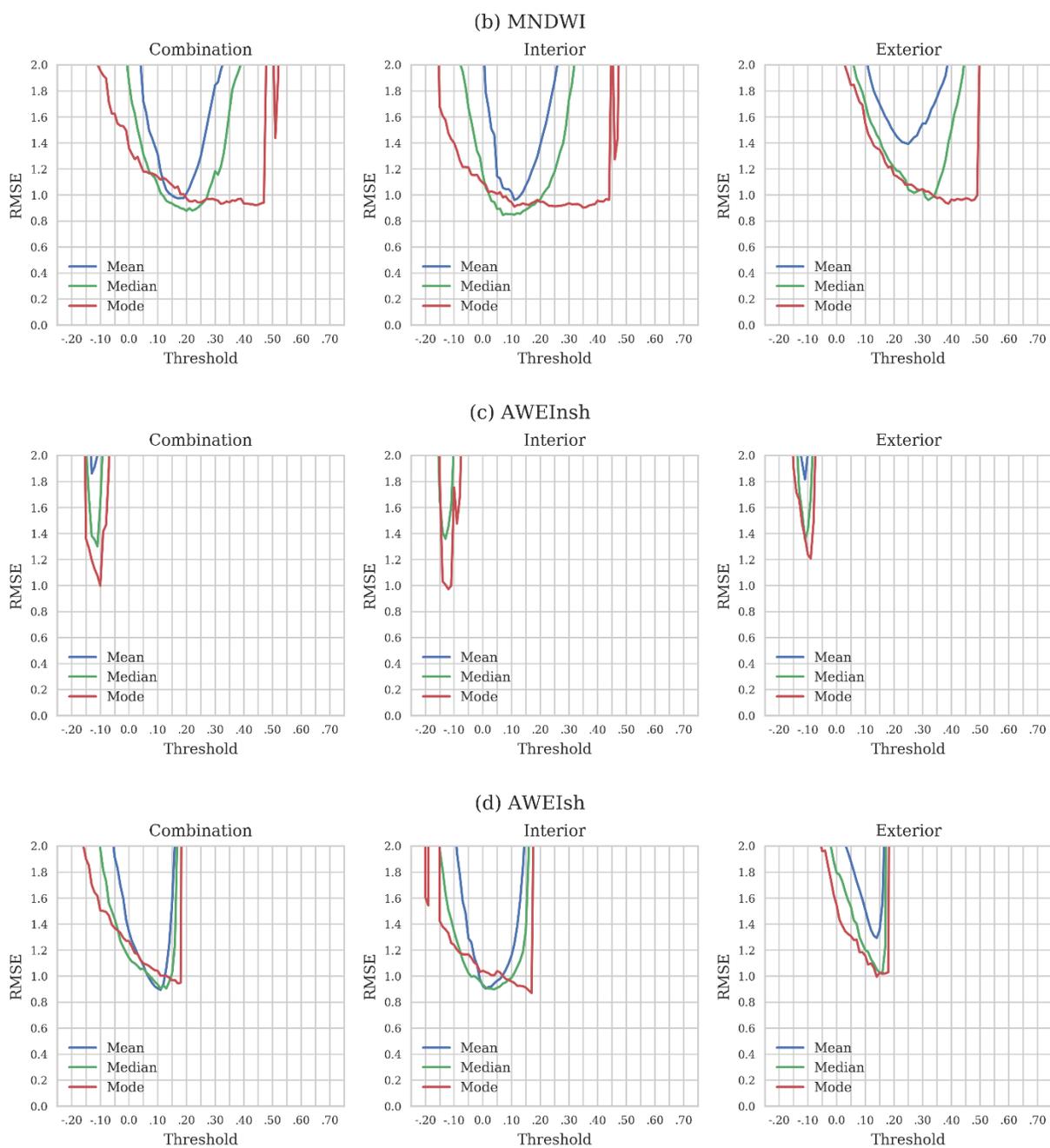


Figure 2-3: An example of binary water index image. The top image is before the vector intersect approach has been applied and includes all water bodies within the study area.



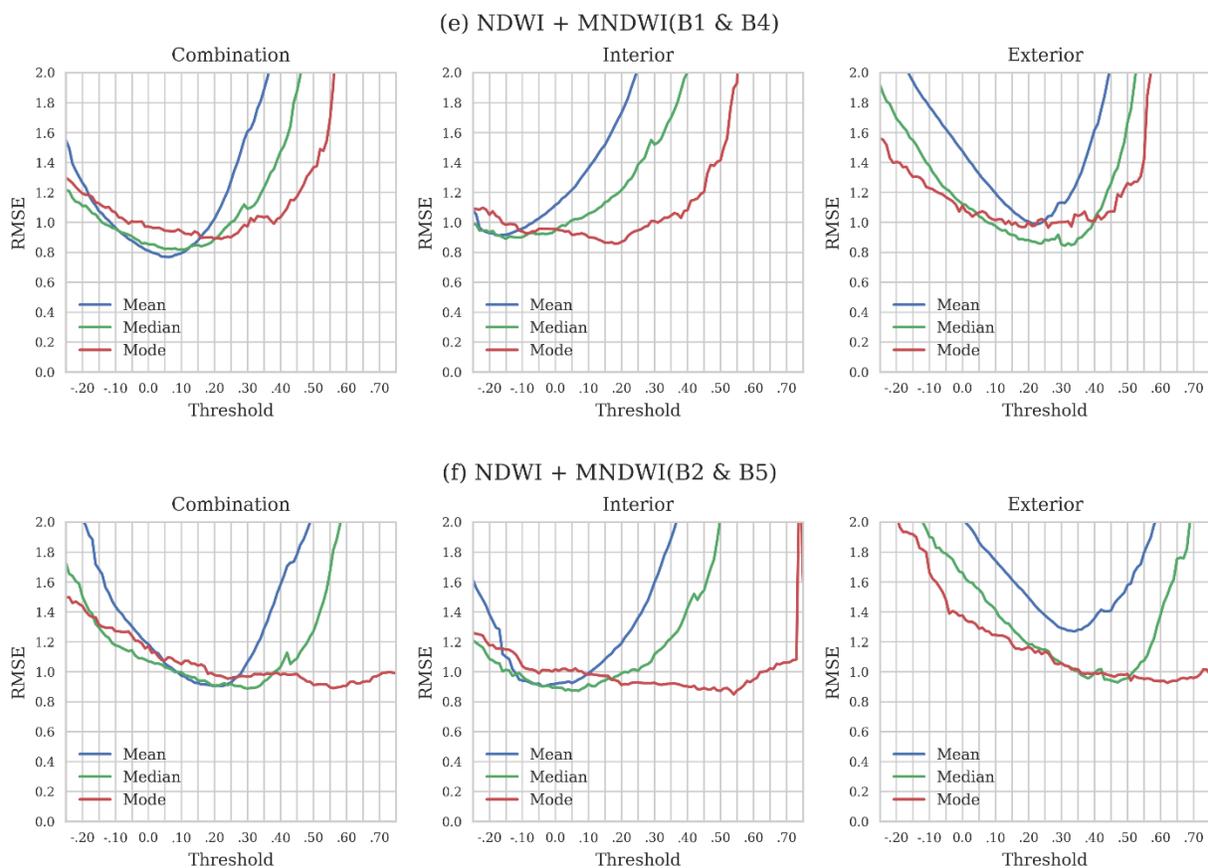
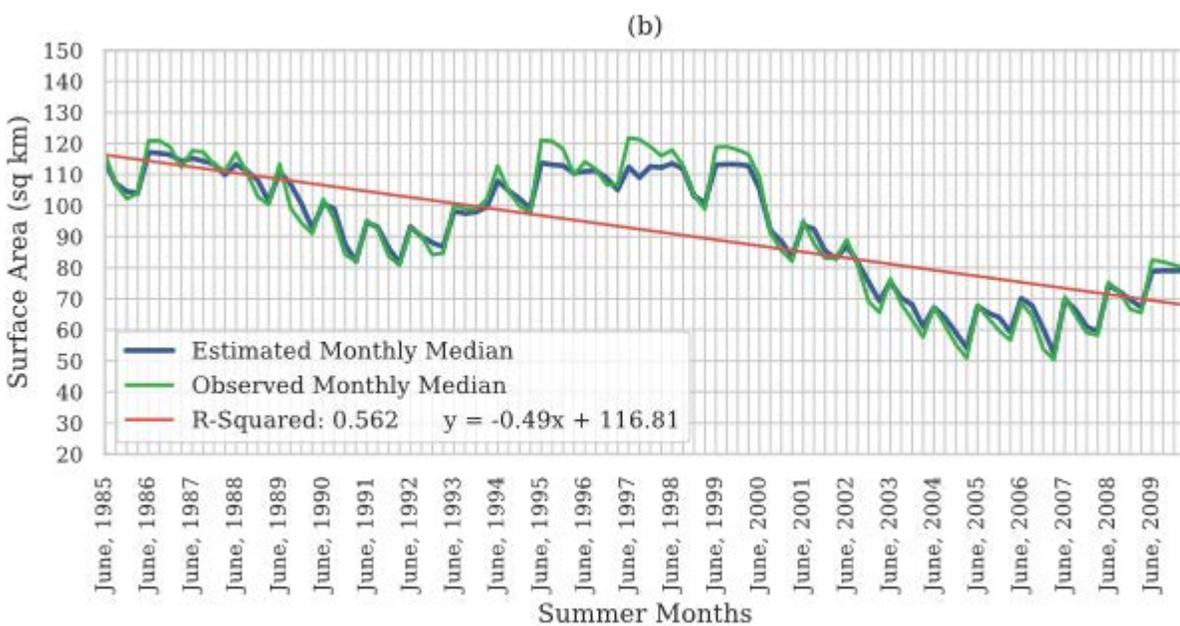
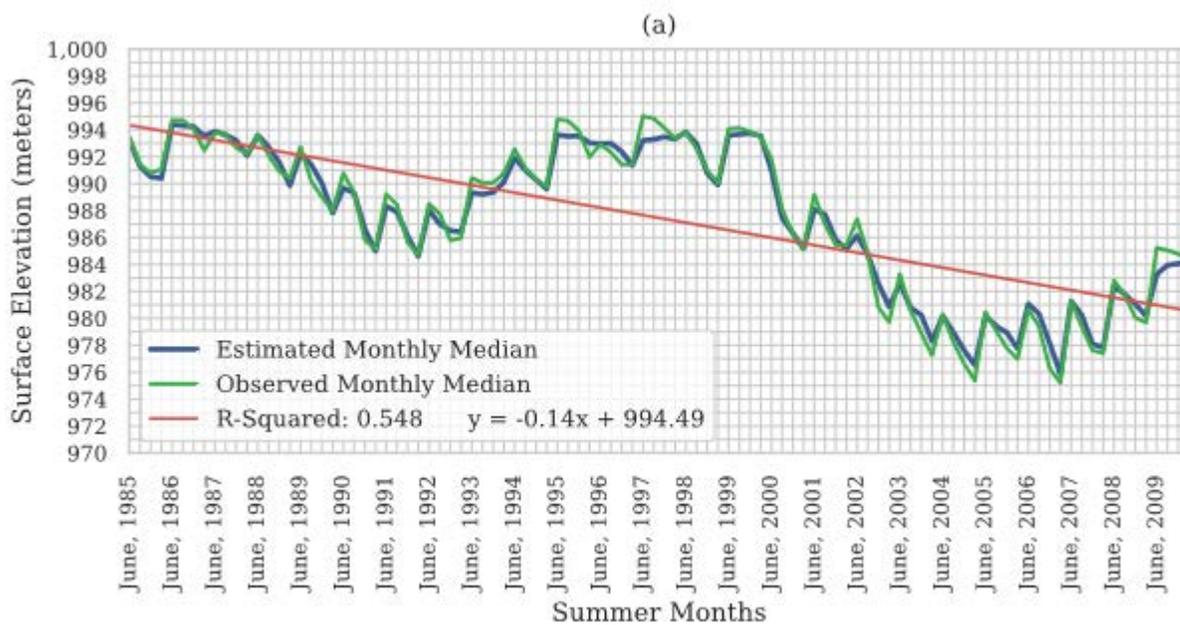


Figure 2-5: RMSE (Root-Mean-Square-Error) curves by water index types. (a) Normalized Difference Water Index (NDWI); (b) Modified Normalized Difference Water Index (MNDWI); (c) Automatic Water Extraction Index – no shadow (AWEInsh); (d) Automatic Water Extraction Index – shadow (AWEIsh); (e) NDWI + MNDWI (B1 & B4); (f) NDWI + MNDWI (B2 & B5).



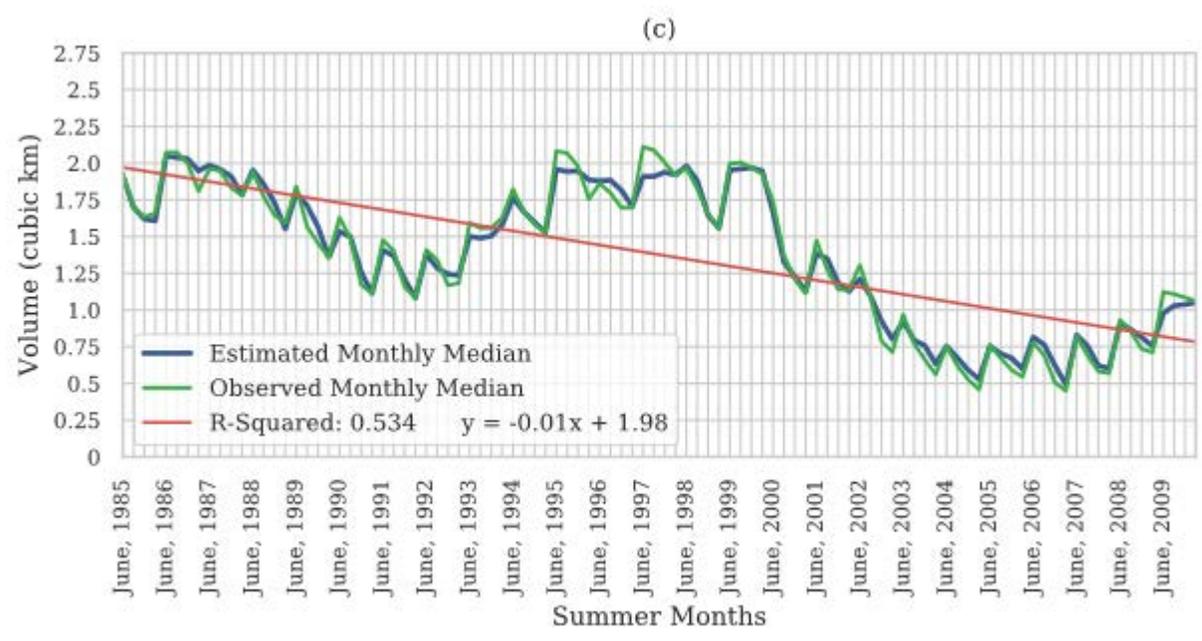


Figure 2-6: Lake water surface elevation, surface area, and volume over time. (a) Surface elevation; (b) Surface area; (c) Water volume.

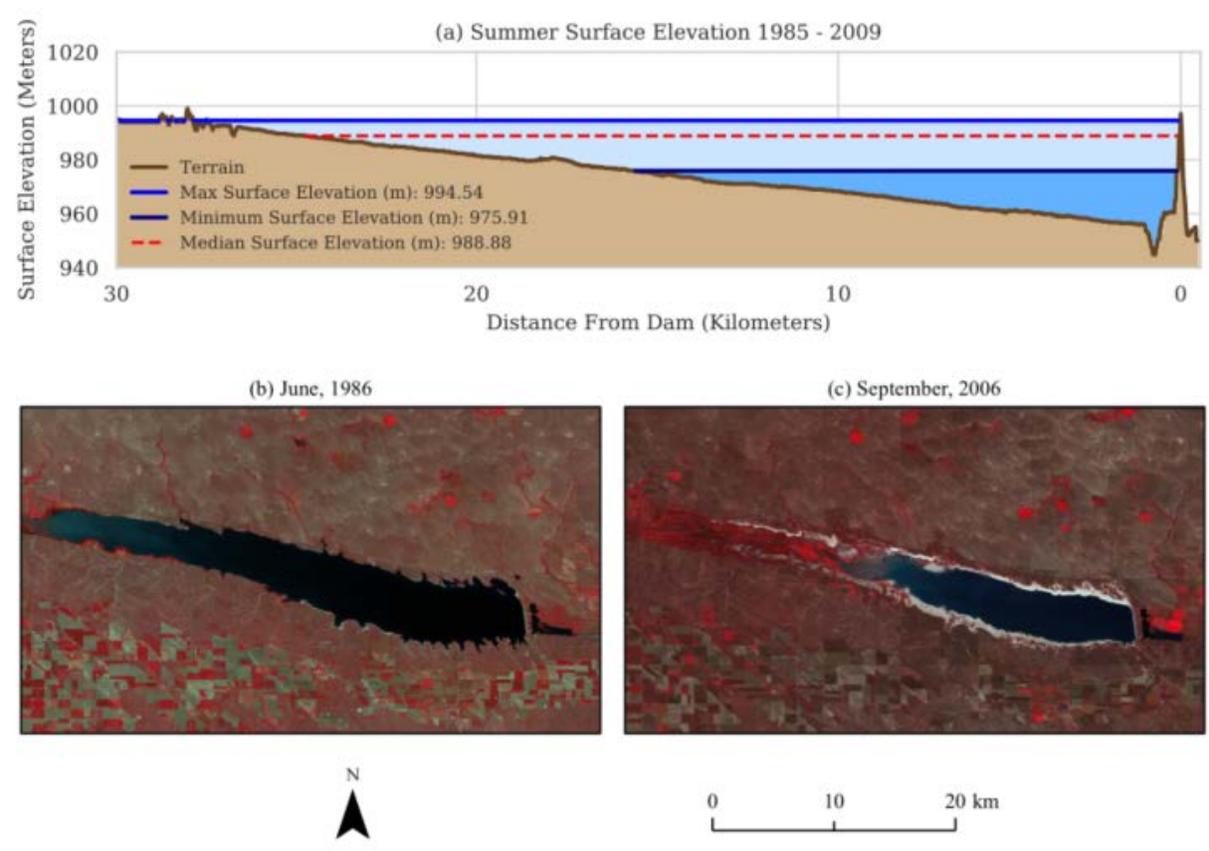


Figure 2-7: (a) Reservoir maximum, minimum, median summer surface elevations and lake depth along a centralized line bisecting the reservoir longitudinally. It shows the vast change in water quantity between the maximum and minimum reservoir

levels. Note that the y-axis is not in the same scale as the x-axis; False-color infrared composites of peak reservoir level in June, 1986 (b) and minimum reservoir level in September, 2006 (c). The western end of the reservoir retreated nearly 14.5 kilometers during that time span.

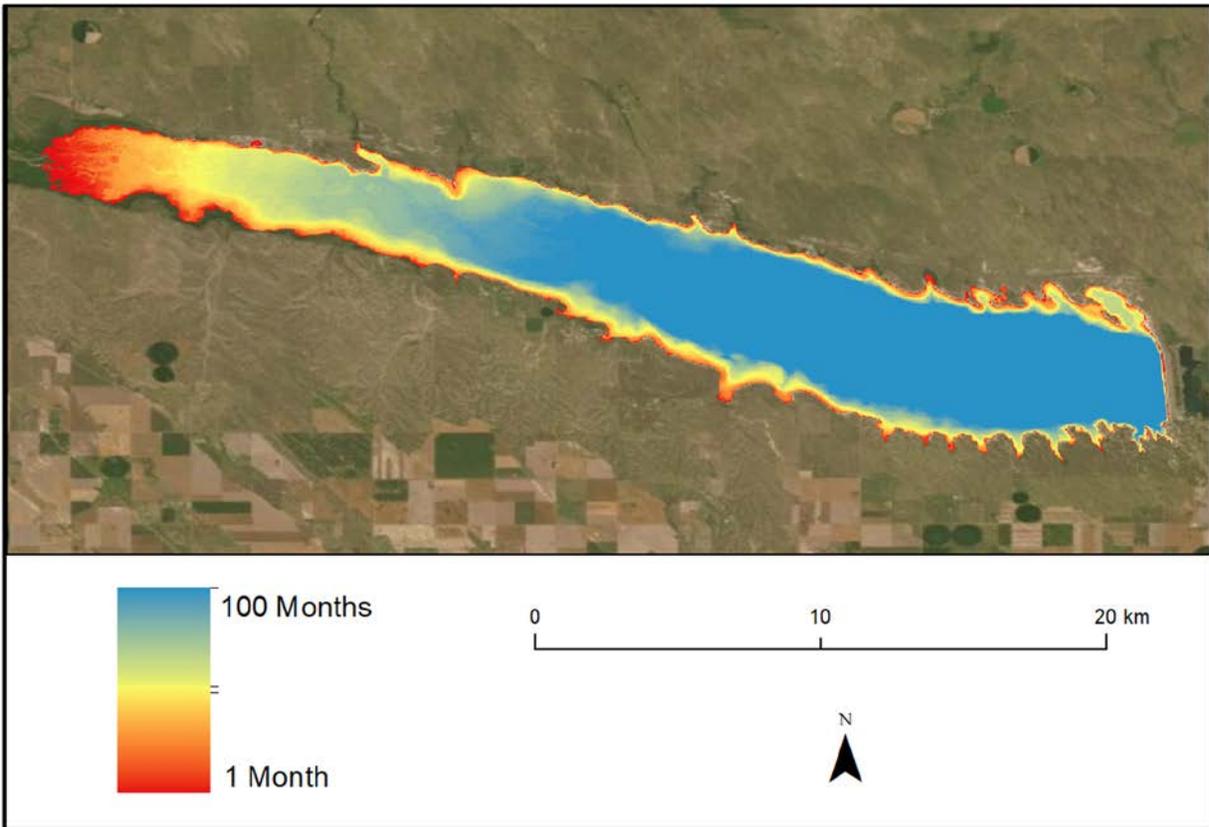


Figure 2-8: Surface cover frequency of Lake McConaughy during the entire study period from 1985-2009. Blue regions indicate areas where water was always or almost always detected. Red regions are areas where water was detected as little as one time.

Table 2-1: Band designations and additional information for the Landsat 5 Thematic Mapper.

Mission	Landsat 5	
Sensor	Thematic Mapper	
GEE Date Range	03/16/1984 - 05/05/2012	
Band 1	Blue (30m)	0.45 - 0.52 μm
Band 2	Green (30m)	0.52 - 0.60 μm
Band 3	Red (30m)	0.63 - 0.69 μm
Band 4	Near Infrared (30m)	0.76 - 0.90 μm
Band 5	Shortwave Infrared 1 (30m)	1.55 - 1.75 μm
Band 6	Thermal Infrared (60m)	10.40 - 12.50 μm
Band 7	Shortwave Infrared 2 (30m)	2.08 - 2.35 μm
Radiometric Resolution	8 bit (256 levels)	
Revisit Period	16 days	

Table 2-2: Water index formulas using Landsat 5 Thematic Mapper bandwidths.

Water Indices		
Water Index Name	Abbreviation	Formula
Normalized Difference Water Index	NDWI	$(B2 - B4) / (B2 + B4)$
Modified Normalized Difference Water Index	MNDWI	$(B2 - B5) / (B2 + B5)$
Automated Water Extraction Index - No Shadow	AWEInsh	$4 \times (B2 - B5) - (0.25 \times B4 + 2.75 \times B7)$
Automated Water Extraction Index - Shadow	AWEIsh	$B1 + 2.5 \times B2 - 1.5 \times (B4 + B5) - 0.25 \times B7$

Water Index Combinations	
Combination	Formula
NDWI + MNDWI (B1 & B4)	$((B2 - B4) / (B2 + B4)) + ((B1 - B4) / (B1 + B4))$
NDWI + MNDWI (B2 & B5)	$((B2 - B4) / (B2 + B4)) + ((B2 - B5) / (B2 + B5))$

Table 2-3: Statistical formulas used to assess lake dynamic estimate accuracy.

Equation	Formula
Root-mean-squared-error (RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$
Mean absolute error (MAE)	$MAE = \frac{1}{n} \sum_{j=1}^n y_j - \hat{y}_j $
Mean bias error (MBE)	$MBE = \frac{\sum_{j=1}^N (y_j - \hat{y}_j)}{N}$
Mean absolute percentage error (MAPE)	$MAPE = \frac{100\%}{n} \sum_{j=1}^n \left \frac{y_j}{\hat{y}_j} \right $

Table 2-4: All parameters and values tested.

Name	Number of Values	Values
Water Indices	6	NDWI, MNDWI, AWEInsh, AWEIsh, NDWI+MNDWI (B1 & B4), NDWI+MNDWI (B2 & B5)
Segmentation Thresholds	111	-0.35 to 0.75 with an increment of 0.01
Boundary Types	3	Interior, Exterior, Combination
Statistical Types	3	Mode, Median, Mean

Table 2-5: The number of combinations by water index types which exceed selected water surface elevation RMSEs. NDWI + MNDWI (B1 & B4), for example, had 245 combinations (24.5%) with RMSEs better (lower) than 1.0 meters and 10 combinations (1.0%) better than 0.8 meters.

	Less Than 1.25		Less Than 1.0		Less Than 0.9		Less Than 0.8		Total Tested Combinations
	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage	
NDWI	212	21.2%	102	10.2%	29	2.9%	0	0.0%	999
MNDWI	228	22.8%	128	12.8%	17	1.7%	0	0.0%	999
AWEInsh	10	1.0%	3	0.3%	0	0.0%	0	0.0%	999
AWEIsh	136	13.6%	51	5.1%	4	0.4%	0	0.0%	999
NDWI + MNDWI (B1 & B4)	481	48.1%	245	24.5%	86	8.6%	10	1.0%	999
NDWI + MNDWI (B2 & B5)	488	48.8%	274	27.4%	35	3.5%	0	0.0%	999
Totals	1,555	25.9%	803	13.4%	171	2.9%	10	0.2%	5,994

Table 2-6: Number of analyses with RMSEs less than 1.0 meter (water surface elevation). Table is broken down by water index types and boundary / statistics types. Of the tested parameter combinations, interior and combination boundaries outperformed exterior boundary in terms of water surface elevation accuracy.

	Combination			Interior			Exterior		
	Mode	Median	Mean	Mode	Median	Mean	Mode	Median	Mean
NDWI	19	17	12	23	14	7	11	12	0
MNDWI	28	16	5	38	18	3	17	4	0
AWEInsh	0	0	0	1	0	0	0	0	0
AWEIsh	8	8	6	14	15	9	2	0	0
NDWI + MNDWI (B1 & B4)	39	39	32	48	31	18	20	32	5
NDWI + MNDWI (B2 & B5)	55	31	21	70	39	24	37	15	0
Subtotals	149	111	76	194	117	61	87	63	5
Total by Boundary	336			372			155		

Table 2-7: Accuracy statistics for water surface elevation, surface area, volume, and volume change using Landsat 5 image composites and merged bathymetry/SRTM data. Also, error statistics for water volume change estimated the pyramidal frustum method are also included for comparison.

Metric	Root Mean Square Error	Mean Absolute Error	Mean Bias Error	Mean Absolute Percentage Error
Surface Elevation (m)	0.768 CI _{95%} [0.657, 0.885]	0.601 CI _{95%} [0.513, 0.698]	-0.071 CI _{95%} [-0.223, 0.073]	0.06% CI _{95%} [0.052, 0.071]
Surface Area (km²)	3.527 CI _{95%} [3.000, 4.085]	2.744 CI _{95%} [2.333, 3.198]	-0.009 CI _{95%} [-0.709, 0.710]	3.10% CI _{95%} [2.584, 3.667]
Volume (km³)	0.069 CI _{95%} [0.060, 0.077]	0.055 CI _{95%} [0.047, 0.063]	-0.012 CI _{95%} [-0.026, 0.002]	4.86% CI _{95%} [3.922, 5.876]
Volume Change (km³)	0.074 CI _{95%} [0.064, 0.084]	0.058 CI _{95%} [0.050, 0.067]	0.00 CI _{95%} [-0.014, 0.015]	- -
Volume Change (km³) Pyramidal Frustum	0.074 CI _{95%} [0.065, 0.084]	0.059 CI _{95%} [0.050, 0.068]	0.00 CI _{95%} [-0.014, 0.015]	- -

3 Maximizing Multi-Decadal Water Surface Elevation Estimates with Landsat Imagery and Elevation/Bathymetry Datasets

3.1 Abstract

Identifying patterns and trends in long-term lake dynamics is essential to establish effective water management procedures and boost our understanding of inland water's role in the global water cycle. This research leverages Google Earth Engine to estimate multi-decadal water surface elevations for 52 lakes and reservoirs with varying physical properties (shape, size, origin, ecoregion, underlying geology, and management practices). Water surface elevation was estimated using the entire Landsat 4, 5, 7, and 8 Landsat Top-of-Atmosphere Tier-1 Collection-1 archive from August 1982 through December 2017 via shoreline boundary statistics extracted from the National Elevation Dataset merged with lake bathymetry. Using the Pixel Quality Assessment (QA) band, image contamination (cloud, shadow, snow, ice, sensor error, etc.) was identified and removed to provide data returns for images with varying levels of image contamination. To improve accuracy, data filtering techniques were identified which retained over 70% of images with detectable water boundaries producing 26 lakes with sub-meter root-mean-square-error accuracy and 40 lakes with sub-meter mean-absolute-error-accuracy using a generalize overall parameter model. Additionally, lake-specific locally optimized models were also determined with 45 of the 52 lakes producing sub-meter root-mean-square-error accuracies and 49 with sub-meter mean-absolute-errors with individual lake accuracy as low as 0.191m RMSE CI95% [0.129, 0.243]. In general, individual lake accuracy is highly correlated with mean slope of surrounding terrain with low-slope shorelines having greater accuracy than high-slope shorelines. Seasonal patterns in estimate accuracy were also identified. This research extends our ability to track lake dynamics over long time periods to lakes lacking traditional in-situ monitoring, enables rapid assessment of lake dynamics across large areas, and balances a need for both high-accuracy measurements and maximum temporal resolution.

3.2 Introduction

Over the last few decades, inland surface waters have been increasingly recognized as critically important to global climate, biodiversity, and security (Vörösmarty et al., 2010). Historically, our primary source of water surface elevations has been in-situ hydrological gauge station networks installed at individual lake and reservoir sites. Unfortunately, due to costs and logistics, in-situ monitoring stations are only available for a small subset of lakes globally and suffer greatly from uneven spatial and temporal distribution.

Remote sensing has long been used to supplement in-situ networks by providing measurements of water surface area via single-band thresholds, water indices, tasseled cap wetness, supervised and unsupervised classifications, and other strategies (Crist, 1985; Feyisa et al., 2014; McFeeters, 1996; Rokni, Ahmad, Selamat, & Hazini, 2014; Zhou et al., 2017). Using these techniques, numerous studies have used optical sensors, like Landsat Thematic Mapper and MODIS (Moderate Resolution Imaging Spectroradiometer), to assess long-term water surface area changes including Tulbure and Broich (2013), Tulbure et al. (2016), and Kang and Hong (2016). In 2016, Pekel et al. published the Global Water Surface Explorer, developed using Google Earth Engine (Gorelick et al., 2016), which analyzed multi-decadal Landsat imagery and provided the largest, most complete view of global surface waters, water surface areas, occurrence, and coverage trends to date.

While insightful, water surface area is a two-dimensional measurement. To better our understanding of inland water dynamics, three-dimensional measurements (volume and volume change) are needed. Achieving three-dimensional surface water measurements is dependent upon having accurate water surface elevations which can then be combined with either water surface area measurements or hypsometric curves (volume/area/elevation relationships) to estimate water volume and/or volume change. Using remote sensing, several methods of measuring water surface elevation have been developed. The primary method utilizes satellite altimeters (non-imaging, profiling instruments) which

measure the response time of radar/lidar signals reflected off the water's surface to estimate water surface elevation. While capable of high accuracy (3 – 10 cm) in ideal conditions (Asadzadeh Jarihani, Callow, Johansen, & Gouweleeuw, 2013b; J F Crétaux et al., 2016), the current fleet of altimeters were primarily designed for ocean and cryosphere monitoring and typically have wide ground-track spacings and poor spatial resolution which limit their ability to monitor small to medium sized water bodies (Solander et al., 2016). In fact, of the 52 lakes used in this study, only 28 receive direct flyovers from satellite altimeters (excluding altimeters with long-return periods allowing denser orbital paths like CryoSat) with most being less than 1 mile in length and potentially containing significant land contamination (“Pass locator: Aviso+,” n.d.). Significant works blending optical imagery surface area measurements and satellite altimetry elevations include, but are not limited to, the works of Alsdorf et al. (2007), Crétaux and Birkett (2006), Duan and Bastiaanssen (2013), and Crétaux et al. (2016). Recently, Busker et al. (2019) combined the satellite altimetry dataset DAHITI (Database for Hydrological Time Series over Inland Waters) (Schwatke, Dettmering, Bosch, & Seitz, 2015) with Global Surface Water Explorer (Pekel, Cottam, Gorelick, & Belward, 2016) monthly water surface area measurements to assess water volume change for 137 global reservoirs.

In addition to satellite altimetry, water surface elevation, total water volume, and volume change have also been commonly measured using optical remote sensing imagery based on area/elevation relationships derived from topographic and/or bathymetric datasets (Avisse et al., 2017; Zhang et al., 2016). Like in-situ hydrological stations, bathymetric data is only available for a small subset of lakes and reservoirs globally. Some studies have attempted to estimate water surface elevation and/or volume by simulating bathymetry using data from surrounding and/or regional terrain (Tseng et al., 2016; van Bemmelen, Mann, de Ridder, Rutten, & van de Giesen, 2016), but these methods have several uncertainties limiting scalability such as site limitations in estimating slope from surrounding terrain or availability of similar watersheds for virtual dam placement. However, in recent years, several governments have initiated large bathymetric survey campaigns to better understand and address growing

reservoir sedimentation concerns which threaten regional water supplies (Kress et al., 2005; Mcalister, Fox, Wilcox, & Srinivasan, 2013; Rahmani et al., 2018). While not available on a global scale, these bathymetric surveys can be used to improve our understanding of climate and seasonal scale water dynamics at the local and regional level.

The Landsat mission, including the satellites and multi-spectral instruments from Landsat 4, 5, 7, and 8, represents the longest, continuous remote sensing monitoring dataset and is ideally suited for this analysis (Wulder et al., 2016). Despite possessing a modest 16-day revisit period, overlap along scene edges will provide additional measurements for several water bodies, and the 30-m spatial resolution will enable observation of smaller water bodies with greater accuracy than possible with other sensors featuring better temporal resolution but decreased spatial resolution, like MODIS (Keys & Scott, 2018; Moradi et al., 2014). Furthermore, while many remote sensing studies have avoided using contaminated imagery (cloud, cloud shadow, snow, ice, sensor failure, etc.) careful analysis of image contamination can retain and maximize useful data (Zhao & Gao, 2018).

Most previous research on estimating water surface elevation uses the area/elevation relationships. This approach, however, depends on clear or composite imagery for water surface area estimate, which severely limits its temporal resolution. Expanding upon the earlier work of Weekley and Li (2019) and leveraging Google Earth Engine (Gorelick et al., 2016), we have developed a method of estimating water surface elevation by detecting water shoreline boundaries which are then used to extract elevation values from an underlying digital elevation model merged with bathymetric survey data. While previous studies have avoided using contaminated imagery (cloud, cloud shadow, snow, ice, sensor failure, etc.), our method uses pixel quality to identify and remove image contamination and other artifacts from the detected water shoreline boundary to enable water surface elevation estimation from images with varying levels of contamination. Using the method, we analyzed the complete Landsat 4, 5, 7, and 8 Top-of-Atmosphere Tier-1 Collection-1 archives for 52 lakes and reservoirs throughout the Western and Midwestern United States from August 1982 through December 2017. Water surface

elevation estimate accuracy was assessed for lakes and reservoirs with wide ranging sizes, shapes, climates, environments, and geologies. Given bathymetric data availability, the method used in this study is rapidly deployable and scalable within Google Earth Engine, can be used on lakes with and without existing in-situ elevation data, offers high-accuracy elevation estimates, and provides valuable insights into multi-decadal lake dynamics with high temporal resolution.

3.3 Study Area and Data

Fifty-two lakes and reservoirs spread throughout the Western and Midwestern United States were selected to ensure a wide range of shapes, sizes, geologic origins, management practices, and climate zones (Figure 3-1).

Four primary types of data were used in this analysis; Landsat Collection 1 Tier 1 TOA imagery, National Elevation Dataset digital elevation models, lake/reservoir bathymetric contours, and in-situ gauge stations. The Global Surface Water Explorer (Pekel et al., 2016) was used as a supplementary dataset to aid in lake seed siting and assessing localized image contamination (the amount of image contamination within 500 meters of the median lake extent).

Table 3-1 summarizes the sensor bands, resolutions, and other characteristics for each Landsat mission included in this study. The complete Collection 1 Tier 1 TOA (Top-of-Atmosphere) archive from August 1982 through December 2017 for Landsat 4, 5, 7, and 8 was included in the study. Tier 1 Landsat scenes have the highest available data quality with Level-1 Precision and Terrain (L1TP) correction, well-characterized radiometry, and inter-calibration across the various Landsat sensors making the collections well-suited for time series analysis (Wulder et al., 2016). In addition to each sensor band, Level-1 Landsat products also include a Quality Assessment (QA) (USGS, 2018b, 2019) band which provides per-pixel bit values useful for identifying surface, atmospheric, and sensor conditions which could impact image analysis such as cloud cover, snow/ice, shadow, dropped pixels, and radiometric saturation (USGS EROS, 2017). Surface reflectance is also available for Landsat; however, the USGS

reports that algorithms currently used to produce surface reflectance have not been optimized to work over surface water (“Landsat Surface Reflectance,” n.d.).

As shown in Table 3-1, Landsat 8 differs from the earlier missions in several respects. First, Landsat 8 has several additional bands as well as different band designations. For Landsat 8, Band 1 is an ultra-blue band to detect coastal aerosols while Band 2 is a blue band like Band 1 on the earlier missions. Additionally, the radiometric resolution of Landsat 8 is 16-bit providing greater quantization and bit depth to the images (USGS, 2019). All in all, despite the band and radiometric resolution differences the same general processing procedure was used for each sensor.

Two types of elevation data were used in this research. First, the National Elevation Dataset (NED), now a part of the 3D Elevation Program (3DEP), is a 1/3 arc-second (approximately 10 meters north/south) seamless DEM with full coverage of the lower 48 states and was used to represent surrounding topography elevations for each water body (Archuleta et al., 2017). The second elevation data source was lake bathymetric contours. Bathymetric surveys were obtained from each of the following agencies or organizations; The United States Bureau of Reclamation (USBOR), the Texas Water Development Board (TWDB), the Kansas Biological Survey (KBS) in conjunction with the Kansas Water Office (KWO), and the Minnesota Department of Natural Resources (MNDR) (“Kansas Lakes and Reservoirs,” 2016; “Lakes Data for Minnesota,” n.d.; “Technical Service Center | Reservoir Surveys,” n.d.; “Texas Water Development Board,” n.d.).

Unlike the bathymetric surveys from the USBOR, KBS, and TWDB, the survey data provided by MNDR lacked the reference elevation of the water’s surface at the time of acquisition and instead only reported contour depth. Fortunately, the edges of the hydro-flattened surfaces in NED were often close enough to the outer contour line to approximate a surface elevation which allowed the lakes to be included despite lacking this critical piece of information.

In-situ gauge station data were obtained from multiple sources including the United States Geological Survey (USGS), the MNDR, and the California Department of Water Resources (CDWR) (“California Data Exchange Center,” n.d.; “Lakes Data for Minnesota,” n.d.; “USGS Lakes and

Reservoirs,” n.d.). Like the bathymetric surveys, gauge data were also reported in a variety of vertical datums. Both the in-situ gauge data and bathymetry data were adjusted to NAVD88 as needed (“VERTCON,” n.d.). Table 3-2 provides additional information regarding each bathymetric survey and vertical datum adjustments.

3.4 Methods

The objective of this research is to obtain water surface elevation estimates from as many Landsat images as possible while maintaining high accuracy. To do this, a multi-step process was developed to eliminate image contamination from the water shorelines allowing underlying elevation values to be extracted from merged bathymetry/topography and compiled into a single water surface elevation estimate.

3.4.1 Digital Elevation Model and Bathymetry Merger

The NED, like most DEMs and for most water bodies, is limited to a hydro-flattened surface approximately representing the water surface level at the time of the DEM data acquisition. Additionally, many lakes and reservoirs have fluctuating water levels making the hydro-flattened surface insufficient for tracking long-term lake dynamics as the DEM cannot be directly used for water surface elevation estimates when water levels fall below the hydro-flattened elevation level (Avisse et al., 2017; Zhang et al., 2016). To deal with this issue and establish a full elevation profile for each lake, the NED and bathymetric surveys were merged into a single DEM for each lake. Despite differences in water level between the DEM hydro-flattened water surface and the bathymetry survey surface contour for each lake, which could result in gaps or overlaps in the two datasets, the basic merging process remained the same. First, hydro-flattened lake surfaces were masked and removed from the NED. The unmasked NED raster was then converted into point features (each pixel is a point) and then merged with the bathymetry contours using the Topo to Raster tool in ArcGIS. Adjustments to this procedure were made on a lake by lake basis to ensure a clean merger. For consistency, in areas of raster point and bathymetric contour

overlap, priority was given to the bathymetric contours as this yielded a cleaner land/water interface in the resulting DEM. It should also be noted that in many cases the vertical datum of the NED and the bathymetric survey data are different. While the NED uses the North American Vertical Datum of 1988 (NAVD88), many of the bathymetric surveys used either NGVD29 (National Geodetic Vertical Datum of 1929) or a local, site-specific vertical datum. Surveys using NGVD29 were adjusted to match NAVD88 using VERTCON (“VERTCON,” n.d.). Other lakes required custom adjustments based upon reported datum differences within the bathymetric survey documents.

3.4.2 Water Surface Elevation Estimation

To estimate water surface elevation, the methodological procedure in Figure 3-2 was applied to every image in the Landsat stack intersecting the target lake. In short, an image is first selected from the stack, a water index is applied, the water index image is segmented into water/non-water regions, water areas are vectorized and then filtered using a multi-point or linear seed (representing the permanent water body) to eliminate disconnected water bodies from the targeted lake body, and the targeted water body is then converted back into a raster. Once the targeted water body has been isolated, the shoreline boundary is identified using kernel morphological operations (dilation and/or erosion) which is then used to mask the merged DEM/bathymetry data leaving a “ring” of elevation values corresponding to the shoreline (Weekley & Li, 2019).

Unfortunately, image contamination, such as clouds, ice/snow, pixel drop-out, shadow, Scan Line Corrector failure, and other contamination/noise sources interfere not only with water detection but also shoreline detection. To alleviate this issue, an additional round of kernel dilation is needed prior to the DEM/bathymetry mask. In this case, the contaminated regions are identified, dilated, and used as a mask to remove erroneous shorelines adjacent to contaminated areas (Figure 3).

Once contaminated sections of the shoreline have been removed, statistical measures of central tendency (mean, median, and mode) were calculated using the remaining pixels to estimate water surface elevation. Although not covered in detail within this chapter, from elevation it is a relatively simple

process to estimate surface area and volume for each lake using an elevation/surface area/volume relationship formula derived from the merged DEM and bathymetry. The full process effectively estimates water surface elevation, surface area, and volume despite image contamination which would otherwise hinder their measurements.

3.4.3 Water Detection and Shoreline Extraction

As discussed in the introduction, water detection has a long history in remote sensing. Some methods, such as supervised image classification, require extensive field observations to train and test the classification while others, such as those involving decision trees, require rules which can be difficult to develop and may not be universally applicable to all data types (Huang, Chen, Zhang, & Wu, 2018). In comparison to these techniques, water indices, which are calculated from two or more bands based on the spectral characteristics of water and non-water targets, are computationally efficient which makes them ideal for time-series analysis if a suitable segmentation threshold can be identified. This research analyzes several common water indices including Normalized Difference Water Index (NDWI) (McFeeters, 1996), Modified Normalized Difference Water Index (MNDWI) (Xu, 2006), Automated Water Extraction Index (AWEI_{sh} and AWEI_{insh}) (Feyisa et al., 2014) and two combined indices formed from NDWI + MNDWI (Blue Band & B4) and NDWI + MNDWI (B2 & B5).

The accuracy of water indices is highly dependent upon selecting an optimum segmentation threshold which can be a difficult process. Multiple factors impact water identification including atmospheric conditions, water color (sediment load, dissolved organics, depth to bottom, etc.), time of day, and land cover composition. These impacts are amplified in time-series analysis where conditions may change over time due to cyclical or seasonal effects. Two strategies were utilized in this research. First, a series of pre-set naïve segmentation thresholds ranging between -0.15 to +0.25 in increments of 0.05 were used. This strategy, naïve segmentation thresholding, treats long-term water surface elevation estimation like linear regression, where the goal is to minimize the error for the entire analysis rather than for a single point in time. In other words, the best naïve threshold is the one that minimizes the error over

the length of study period and as such may perform well under certain lake conditions while performing sub-optimally as conditions change. The second strategy utilizes Otsu's method (Otsu, 1979) to dynamically threshold each image. Otsu's method is an adaptive image processing technique used to split an image or dataset into two classes based on the observed distribution of pixel values. For this research, global Otsu thresholding was implemented for each lake by restricting the operation to within 500 meters of observed water extents in the Global Water Surface Explorer dataset (same polygons as used to estimate local percent impacted) (Pekel et al., 2016). Otsu's method assumes the two classes (water and non-water) have roughly the same number of pixels. Uneven distribution in pixel values, due to limited water surface area in comparison to non-water, or significant image contamination can affect the threshold calculation. To alleviate some of these issues, local Otsu thresholding was also implemented. Local Otsu thresholding is a multi-step procedure requiring an initial water detection step using a naïve threshold value to identify an approximate shoreline location which is then used to isolate and buffer the shoreline to restrict the Otsu calculations to a specific region in order to provide a more precise segmentation threshold. Furthermore, image contamination sources such as clouds, shadows, snow, ice, etc. were masked from the water index image prior to application of both the global and local Otsu's method so that the dynamic threshold value would be restricted to clean land and water pixels. Without the image contamination masking step, Otsu's method may inappropriately separate other image features such as clouds and cloud shadow instead of land and water. Table 3-3 summarizes the water indices and thresholds used in the analysis.

With the segmented binary water/non-water image, the target water body was then isolated within the image to ensure all subsequent processing steps are applied only to the target lake. In GEE, three isolation techniques are available including region group, cumulative cost, and vector filtering. Since all three methods produce the same result, vector filtering was used in this research as it was easier to implement and scale within the analysis. To do this, all water bodies within the water/non-water binary image are vectorized, then using a seed (a point, line, or polygon placed within the water body manually), the target water body is selected, and all other water bodies are discarded. The target water body is then

converted back to a raster. Once isolated, the shoreline is identified using kernel morphological operations (dilation and erosion). Figure 3-4 illustrates the three boundary types used to estimate lake water surface elevation.

As mentioned, one of the primary objectives of this chapter is to maximize the temporal resolution of the water surface elevation estimates. Many images contain some form of image contamination or noise which has a direct effect on the accuracy and identification of water pixels and the water boundary. Unmitigated, this contamination produces erroneous water boundaries which may cross non-boundary water or land pixels incorporating false elevations into the calculations. To alleviate this issue, contaminated areas identified using the Pixel QA Band (USGS, 2018b, 2019) were dilated using kernel morphological operations and then used as an additional mask to remove unwanted areas from the water boundary (Figure 3-3).

3.4.4 Lake Surface Area and Volume

While not covered extensively in this chapter due to the primary focus upon estimating water surface elevations and their accuracies across a wide range of environments and conditions, water surface area and volume estimates can also be estimated for contaminated images via elevation/surface area/volume relationships constructed from the merged DEM/bathymetry models. It should be noted, however, that this method could result in measurement errors for lakes that have significantly changed over time as the relationship would reflect the lake state at the time of the topographic and/or bathymetric survey. One lake in this study that this phenomenon could affect is Tuttle Creek Reservoir in Kansas, USA which has lost significant portions of its northern reach due to sediment infilling (Denoyelles & Kastens, 2016; Rahmani et al., 2018).

3.4.5 Contamination/Impacted Area

The localized contamination for each image was determined using the maximum water extent within the Global Water Surface Water Explorer (Pekel et al., 2016) for each lake plus 500 meters. All non-clear pixels (cloud, cloud shadow, ice, snow, etc.) as flagged in each image's Pixel QA band were

considered contamination (USGS, 2018b, 2019). Additionally, gaps resulting from the Scan-line Corrector failure onboard Landsat 7 after May 31st, 2003 as well as uncovered portions of the lake where the Landsat scene only covers a portion of the water body were also considered contamination/impacted area for the purposes of this study. The percent impact for each image is the total contamination/impacted area divided by the maximum lake extent plus 500 meters. For most water bodies in this study, the maximum water extent plus 500 meters could be used as-is, but a few water bodies sourced from larger rivers, which would be detectable in the Landsat imagery, required some manual cleanup to limit the inclusion of upstream, non-lake areas in the analysis.

3.4.6 Lake Morphological Metrics

Three additional lake morphological metrics were also calculated for analysis purposes: mean slope, median surface area, and median lake shape (surface area to perimeter ratio roughly approximating the shape as serpentine or spherical). Like percent impacted, these three metrics are also calculated using the JRC Global Surface Water Explorer Mapping Layers (Pekel et al., 2016). Median surface area and water extent were both determined using the Global Survey Water Explorer's occurrence layer with values exceeding 50% occurrence. Median lake shape was approximated using the ratio of median surface area to median lake shoreline perimeter, and, finally, mean slope was calculated from the NED using a 250 m buffered annulus around the median water occurrence extent in the Global Surface Water Explorer Mapping Layers.

3.4.7 Model Selection and Validation

Model selection was determined by comparing estimated lake surface elevations to in-situ lake elevation measurements using root-mean-squared-error (RMSE) and mean-absolute-error (MAE) calculated using bootstrap resampling (1,000 iterations). For each lake (lake-specific locally optimized model) and all lakes (general models), the best model was determined using the lowest combined MAE and RMSE (ordinally ranked and summed). Testing accuracy was evaluated using the out-of-bag samples in each bootstrap iteration. Additionally, the total number of results was also considered in the

model selection process. Only models generating lake surface elevation estimates for more than 25% of the intersecting images were considered (additional details in discussion).

3.5 Results

The highest overall accuracy for all 52 lakes together is 5.746m RMSE CI_{95%}[5.208, 6.265] and 1.810m MAE CI_{95%}[1.739, 1.886] using NDWI, mode statistic, outside boundary, and localized Otsu dynamic thresholding (Table 3-4). Using the overall best model, four lakes achieved sub-1m RMSE accuracy while 45 lakes total have RMSEs less than the overall 5.746m RMSE. For MAE, 27 lakes achieved sub-1m MAE accuracy and 41 have MAEs lower than the overall 1.81m MAE. The imbalance in the number of lakes with accuracy metrics lower than the overall accuracy indicates that most of the accuracy error comes from just a small number of lakes while the considerable difference between RMSE and MAE indicates large error variance. The accuracy of each lake using the best overall general model is shown in Figure 3-5a.

In addition to the best overall general model, given ground-truth data exist for a lake, lake-specific locally optimized models can also be determined which deliver equal to or greater accuracy than the best overall general model. Figure 3-5 shows that the best overall general model was also the best lake-specific locally optimized model for just one lake (Gibson Reservoir) and all other lake-specific models produced higher accuracies. Overall, the lake-specific locally optimized accuracy was 4.327m RMSE CI_{95%}[3.932, 4.742] and 1.172m MAE CI_{95%}[1.126, 1.220] which is 1.419m RMSE better than the best overall general model. Using this method, 25 lakes have sub-meter RMSE accuracies and 40 lakes have sub-meter MAE accuracies. While the accuracy gains in this approach are promising, the low overall RMSE accuracy (4.362m) relative to the number of lakes with sub-meter accuracies, as evidenced in Figure 3-5, indicates that much of the error is being driven by just a few lakes.

To investigate potential reasons for increased estimate error within the small subset of lakes, as well as general error drivers overall, five potential error drivers were analyzed: mean slope within 250m

of the median shoreline, median surface area, median shape, mean percent impacted, and the total number of images used in the accuracy assessment process (the number of estimates with corresponding ground-truth measurements) (Table 3-5). Mean slope had the highest correlation with accuracy in all four cases (RMSE vs MAE and Best General Model vs Lake-Specific models) with r^2 values ranging from 0.793 CI95% [0.680, 0.899] to 0.892 CI95% [0.799, 0.948] indicating a very strong link between surface elevation estimate accuracy and the slope of the surrounding terrain. The other four potential error drivers showed weak correlation to overall accuracy.

Despite the weak correlation between image contamination (percent impacted) and median shoreline length to overall model accuracy at a lake by lake level, these two metrics could still significantly affect water surface elevation estimates on an image by image basis. Image contamination restricts shoreline visibility effectively shortening the useable shoreline and reducing the number of elevation values available for water surface elevation estimates. Figure 3-6A and B, which plot observed shoreline length against estimate error, shows increased error and greater uncertainty for short shorelines compared to longer shorelines. For shorelines less than 7.5 km in length, the best general overall model accuracy approaches 10m RMSE (9.53 m RMSE) while the lake-specific models exceed 7.5m RMSE. While the bin widths used in this analysis were somewhat arbitrarily assigned and could be adjusted smaller or larger, Figure 3-6 indicates a substantial portion of the error in both the general overall model and the lake-specific models comes from images with significant local contamination which results in shorter observable shorelines.

In addition to the link between shoreline length and estimate error, Figure 3-6 also reveals a second phenomenon impacting water surface elevation accuracy. In the figure, numerous high error points appear sporadically across a range of shoreline lengths. Investigation of the affected images reveals these errors are often caused by unidentified image contamination within the Pixel QA band causing deviations in the detected shoreline. By comparing the mean estimated water surface elevation to that of the mode estimated water surface elevation, many measurements affected by the missing

contamination can be identified and removed. This issue can also occur in areas with significant topographic relief where elevation changes rapidly.

Using the details revealed in Figure 3-6, the analysis was repeated using several preset thresholds to eliminate erroneous measurements using shoreline length and/or mean-mode difference (Table 3-4). The results in Table 3-4 indicate that Mean-Mode difference thresholds and minimum shoreline lengths can substantially improve analysis accuracy while still retaining a high proportion of the available imagery and water surface elevation estimates. Of the combinations and thresholds tested, the lake-specific model with shorelines greater than or equal to 7.5km and Mean-Mode differences less than 5.0 meters produces overall accuracies of 0.68m RMSE CI95% [0.616, 0.662] and 0.385m MAE CI95% [0.387, 0.391]. Additionally, this model produced sub-meter RMSE and MAE accuracies for nearly all tested lakes while retaining approximately 75% of all the estimates. Unfortunately, due to its small size, at this level of filtering, all but six estimates for Flatiron Reservoir in Wyoming, USA are completely filtered. However, even just using Mean-Mode difference filtering without minimum shoreline filtering, Flatiron Reservoir still retains sub-meter accuracy with 0.984m RMSE CI95% [0.875, 1.098] and 0.747m MAE CI95% [0.664, 0.839]. Other small water bodies may be similarly affected with only high-water levels remaining in the results at this level of filtering.

Two final tests for model accuracy were performed. First, water surface elevation estimate accuracy was tested for each Landsat satellite using both the best general overall model and lake-specific locally optimized models with minimum shoreline length (>7.5 km) and Mean-Mode difference filtering (<5.0 m) (Table 3-6). As expected, Landsat 7, which suffered Scan-Line Corrector failure (SLC-off) in May 2003 (Avisse et al., 2017), had the lowest overall accuracy. However, Landsat 7 imagery prior to SLC-off performed quite well and had the second highest accuracy of all the Landsat missions. Unexpectedly, though, Landsat 8 had the lowest accuracy of all the non-SLC-off imagery. It should be noted that (excluding the SLC-off imagery) Landsat 8 imagery also had the highest level of local image contamination which could be contributing to the decreased accuracy. Finally, monthly accuracy was

assessed using both the best overall general model and the lake-specific locally optimized model with filtering. In both models, the winter months had the greatest uncertainty and highest error of all months with general accuracy improving in the early spring months before a slight dip in accuracy occurred in late spring. Late summer to mid-fall, which have the greatest number of useable images, produced the highest accuracy estimates (Figure 3-7).

Figure 3-8 plots the results of the lake-specific locally optimized model with shoreline and Mean-Mode filtering for selected lakes. In general, the plots indicate good agreement between the estimated water surface elevations and the ground-truth measurements with long-term trends, cyclical patterns, seasonal patterns, droughts, and flooding events coinciding with available imagery are detectable. In fact, even the lakes with lower accuracy, such as Anderson Ranch Reservoir in Idaho match the ground-truth data very well in time series. The charts also reveal several locations, such as Lake Minnewaska in Minnesota, where several water surface elevation estimates are likely incorrect, possibly as result of unmitigated snow/ice contamination. Improved data filtering, or even manual removal of these data points could significantly improve the overall estimate accuracy.

All told, due to the incomplete nature of the ground-truth records for many of these water bodies the methods used in this study generated 8,000 – 11,000 water surface elevation data points lacking ground-truth counterparts (Table 3-4). Overall, between 31,000+ to 42,000+ water surface elevations were estimated using this technique with only 11,663 – 11,798 (depending on model and filtering) from images with less than 1% local image contamination (contamination within 500 m of the median lake extent).

3.6 Discussion

3.6.1 Terrain Slope

High slope drives error due to the limited spatial resolution of the Landsat imagery. In the case of high-slope areas, the 30-meter Landsat pixels are insufficient to resolve the water boundary with

enough detail to accurately place it on the ground. In canyon-filled or mountainous reservoirs the elevation can change by ten meters or more in little to no distance which is a significant change within a single Landsat pixel. When working with Landsat, aside from increased vertical accuracy, higher spatial resolution DEM products would likely yield little improvement over coarse resolution DEM since the placement of the boundary is based upon the 30-meter Landsat pixel. The issue is further compounded due to the mixed boundary pixels. The classification as water or non-water for mixed pixels can make a large difference in the elevation value used in the estimation process. Conversely, the effect of mixed pixels for low slope water bodies have a lesser effect since the rate of elevation change is much less. Additionally, many canyon-filled and mountain-based reservoirs, including those used in this study, are riverine/serpentine in shape which results in a higher number of mixed pixels versus the total surface area of the reservoir compared to more elliptical shaped lakes and reservoirs. Together, the steep slopes and relative long shorelines result in greater variance within the shoreline elevation values leading to greater uncertainty and decreased accuracy compared to gentler sloped water bodies.

3.6.2 Image Contamination

Despite the steps taken to eliminate image contamination it still affects many images causing large errors. In many cases, however, problematic images can be identified and removed from the analysis by filtering the results using the local percent impacted or by setting a minimum shoreline length threshold. The applicability of either method (local percent impacted or shoreline length) can be lake dependent due to some subtle differences between the two methods. As covered earlier, local percent impacted considers image contamination within 500 meters of the maximum lake extent within the Global Surface Water Explorer water layers (Pekel et al., 2016). Depending on the lake shape and size though, the lake contamination may have a limited effect on the water boundary where the water surface elevation estimate is being generated. Therefore, it is possible for an image to have a relatively high local percent impacted value and yet the contamination to have a limited effect on the results such as cloud cover over the middle of a lake. Water boundary length, on the other hand, directly influences the

number of elevation values used in the elevation estimate process. For most water bodies, shorter shorelines indicate greater local contamination directly affecting the water boundary, which results in fewer elevation values being used in the estimate which increases its susceptibility to outlier influence. All told, shorter shorelines experience greater uncertainty and removing these boundaries improves estimate accuracy considerably (Table 3-4, Figure 3-6). However, as we all know, shoreline length is also a physical property of the lake itself directly tied to its shape and size. Smaller water bodies may require different minimum shoreline lengths than medium to larger water bodies to prevent being completely filtered from the results as mentioned earlier for Flatiron Reservoir (Wyoming, USA).

3.6.3 Water Index Performance and Filtering

While filtering the results using Mean-Mode differencing and/or minimum shoreline lengths can significantly improve the results it is important to not over-filter the data. As mentioned earlier, the results in this chapter were restricted to models which retained a minimum of 25% of the total number of images with both estimated water surface elevations and ground-truth data. Without the 25% minimum, several lakes did produce higher accuracy water surface elevations but did so because the models were over-filtered and only retained a handful of images with highly accurate results. The 25% threshold in this study was arbitrarily determined and any future application of this technique should take this into consideration. Fortunately, the general overall models we recommend are less susceptible to this effect. Some models, almost exclusively those using AWEInsh, were particularly susceptible to this issue. In general, we would not recommend use of AWEInsh without careful calibration beyond what was utilized in this study as its performance was significantly worse than almost all other tested model combinations. Of the 594 model combinations tested in the general overall model analysis (no-filtering), the highest AWEInsh combination ranked 450th and occupied 95 of the 100 lowest ranks. All the other water indices placed at least one model combination in the top 17. Furthermore, in the lake-specific locally optimized models, all the water indices except for AWEInsh were represented by multiple lakes.

3.6.4 QA Band Limitations

The QA bands processed and added to the Landsat imagery are an invaluable data source for rapid assessment of image contamination within a given image. In this study, the QA bands were used to identify multiple forms of image contamination and noise sources which would negatively affect water surface elevation estimates. While use of the QA bands to identify and eliminate image contamination greatly increased the number of valid water surface elevations, there are cases, both spurious and systematic, which impacted estimate accuracy. As mentioned previously, the water boundary technique used in this study enables water surface elevation estimation from even partial imagery. However, in several cases, even in otherwise clear imagery along boundaries, water boundaries were not detected due to a lack of Pixel QA data along the scene boundaries (Figure 3-9). Additionally, multiple cases of missed clouds, cloud shadow, snow, ice, and other noise sources were identified which affected the water detection and boundary identification processes. Incorrect water and water boundaries result in false elevation values being included in the estimate process resulting in greater error.

3.6.5 Segmentation Threshold

Table 3-4 indicates that local Otsu thresholding offers the best overall performance across all four of the tested general overall models. This is expected as Otsu's method is a dynamic segmentation threshold calculated using a balance of land and water pixels along the shoreline. Local Otsu thresholding was also the most commonly chosen segmentation threshold in all four versions of the lake-specific locally optimized models. Interestingly though, most individual lake models preferred naïve segmentation thresholds. This could be for a few reasons such as peculiarities with the merged DEM/bathymetry where, despite accurate thresholding and water detection, deviations in the elevation values themselves would dictate the use of naïve thresholds which better estimate water surface elevation for that lake by over- or under-estimating the water area. The bathymetry data used in this study come from a variety of sources with varying resolution and quality. Additionally, it was previously noted that the Minnesota lake bathymetry lacked a water surface elevation reference value for the bathymetric

contours. If the estimated water surface elevations applied to those contours during the DEM/bathymetry merger process were incorrect, then that would affect selection of optimal segmentation thresholds which are determined by their ability to minimize water surface elevation estimate errors. Secondly, because Otsu's method is dynamic, issues can arise where unmitigated image contamination, especially ice and snow in and around the water body, affect the threshold calculation process by pushing the class separation in one direction or the other. However, in general, local Otsu's thresholding was most adaptable and accurate water surface elevation strategy.

3.6.6 Time Series

Most time series in Figure 3-8 show good agreement with ground-truth. Some errors are obvious, improved filtering methods or manual removal would significantly improve the accuracy, especially for Pinewood Reservoir (Colorado, USA) where most of the error comes from a single data point. Some water bodies have limited results, so while the available points match up, significant portions of the time series are missing which results in an incomplete lake dynamic record. For example, Lake Waha (Oregon, USA) misses all the minimum lake elevations visible in the ground-truth record. In some cases, the missing data points may be due to the elevation result filtering being too stringent for that water body but could also be due to local conditions which prevent detection of the water boundary during those time periods. Relatedly, the winter months have lower accuracy and greater uncertainty compared to summer and late fall in addition to having fewer valid images and estimates (Figure 3-7). The lack of estimates and greater uncertainty are reflected in many of the time series. For example, Beulah and several others consistently show dense elevation points in the summer drawdown months leading into fall and then relatively few elevation points during the winter and early spring when the lake/reservoir is refilling. One other interesting observation is visible in the time-series charts. Flood event peaks are seldom fully captured in the analysis, this is in part due to timing, and if the satellite is not overhead at the flood peak then it simply can't capture those data as is the case with Kanopolis Reservoir (Kansas, USA) on July 25th, 1993. In that case, Landsat 5 passed overhead on July 20th (pre-flood peak) and on July 27th (post-

flood peak). Furthermore, it's possible that water boundaries during flood events could be subject to increased emergent vegetation such as submerged tree stands which alter the already mixed pixels along the shoreline.

3.7 Conclusions

The data integration, image processing, and data analysis techniques used in this chapter have shown great promise for long-term lake dynamic studies. Use of the Pixel QA band to locate, assess, and remove image noise and contamination sources such as cloud, cloud shadow, ice, snow, etc. not only improves the accuracy of the analysis but also improves the temporal resolution as well. Overall, the analysis added, at minimum, 8,000 additional elevation estimates not currently available from any other data source including in-situ hydrological stations and the method dramatically improved upon the techniques, the accuracy, and the relevancy of the earlier work by Weekley and Li (2019).

While application of this technique is dependent upon having lake bathymetry, either as a separate data source or through pre-impoundment DEM data, it has proven to be an accurate means of estimating water surface elevation across a wide range of environments and lake types. Model accuracy is strongly linked to the mean slope near the lake shoreline but is also affected by high levels of local image contamination limiting the number of DEM/bathymetry elevation values available for estimating water surface elevation. Using shoreline length and/or Mean-Mode differencing can significantly improve model accuracy for both general overall models and lake-specific locally optimized models. If in-situ elevation data are available, lake-specific models offer the highest accuracy with 45 of the tested lakes producing sub-meter RMSE elevation accuracies and 49 producing sub-meter MAE accuracies. Local Otsu dynamic thresholding was the best performing segmentation threshold for each of the general overall models tested (no-filtering and various combinations of minimum shoreline length and maximum Mean-Mode difference) and for most of the best lake-specific locally optimized models.

The results of this research can significantly improve our knowledge and understanding of inland water dynamics at a variety of scales, from local to regional, and offers a scalable mechanism for rapid water dynamic assessment within the Google Earth Engine cloud-computing environment. Additionally, SWOT (Surface Water and Ocean Topography Mission) (Solander et al., 2016), which launches in 2021, will provide lake elevation data at a global scale. While SWOT itself is limited to a 3-year mission life, the elevation data it provides can be used to train lake-specific locally optimized models using the method in this chapter to capture high-accuracy, multi-decadal water dynamics for many lakes and reservoirs.

Future research will investigate means of estimating lake dynamics from surrounding topography (no bathymetry), improved image contamination removal methodologies, improved erroneous result filtering, integration of higher resolution satellite imagery, such as Sentinel-2 (10m), and integration with future SWOT elevation data. All told, the method used in this chapter has significantly improved our understanding and monitoring ability of lakes and reservoirs of varying types, sizes, and environments.

3.8 Figures and Tables

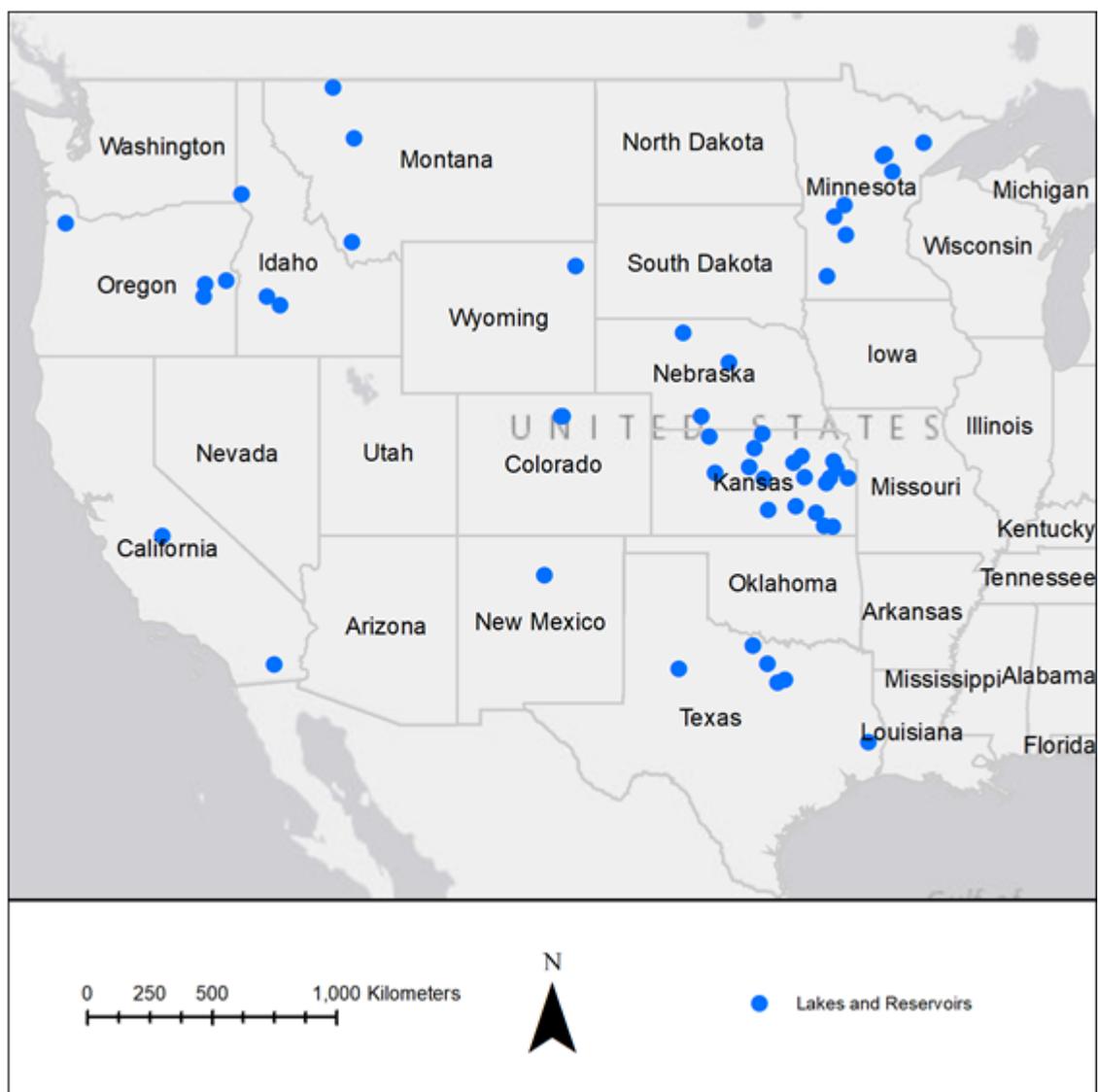


Figure 3-1: Spatial distribution of lakes used in the study.

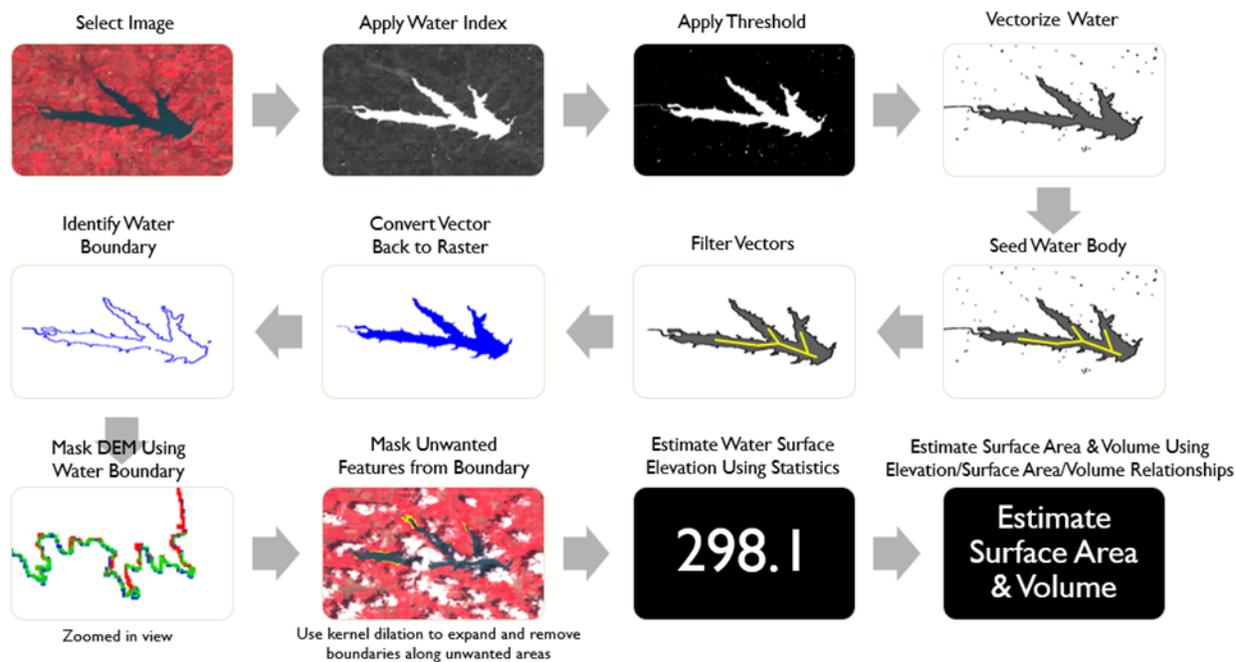


Figure 3-2: General processing procedure for estimating water surface elevation for any given lake with merged DEM/bathymetry data. This procedure was implemented in GEE and applied to each image within the Landsat image stack.

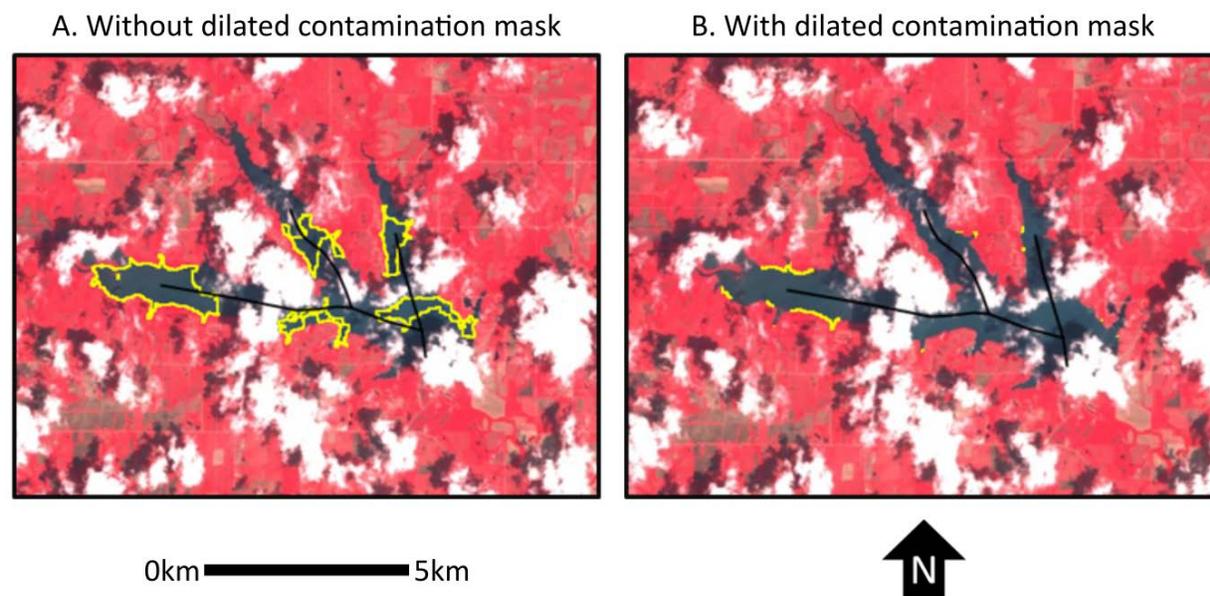


Figure 3-3: Comparison of boundaries with and without the dilated contamination mask for an image with approximately 50% localized contamination. Without the mask, the contaminated areas are identified as non-water which generates a boundary crossing the center of the lake in several locations (A). However, using the Landsat QA band, not only can contaminated areas be identified, but they can be used to remove erroneous shorelines as well (B).

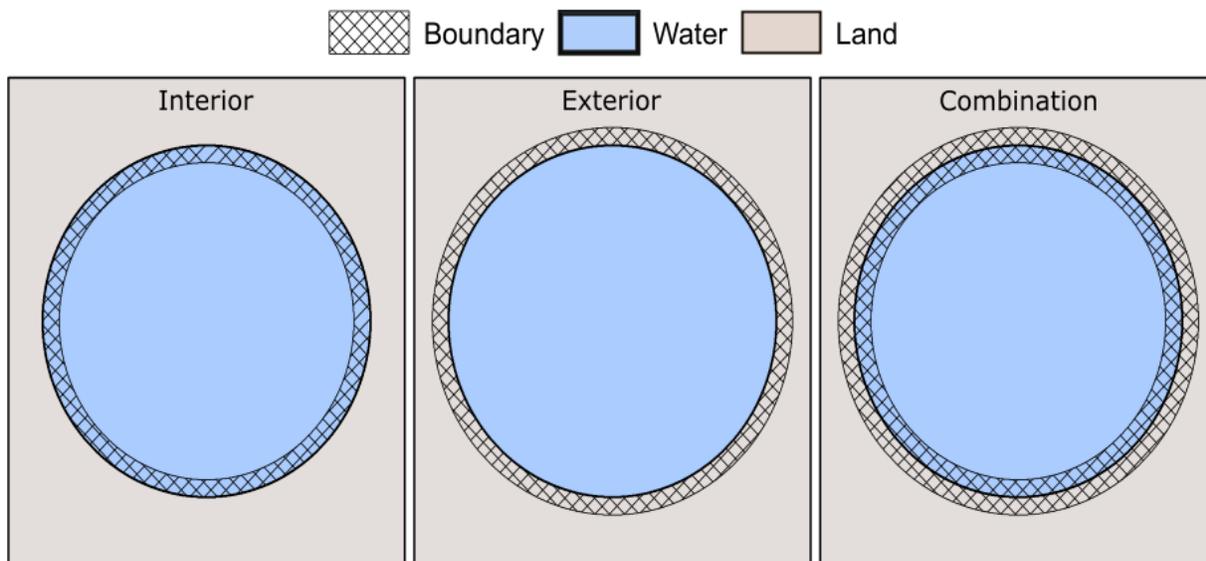
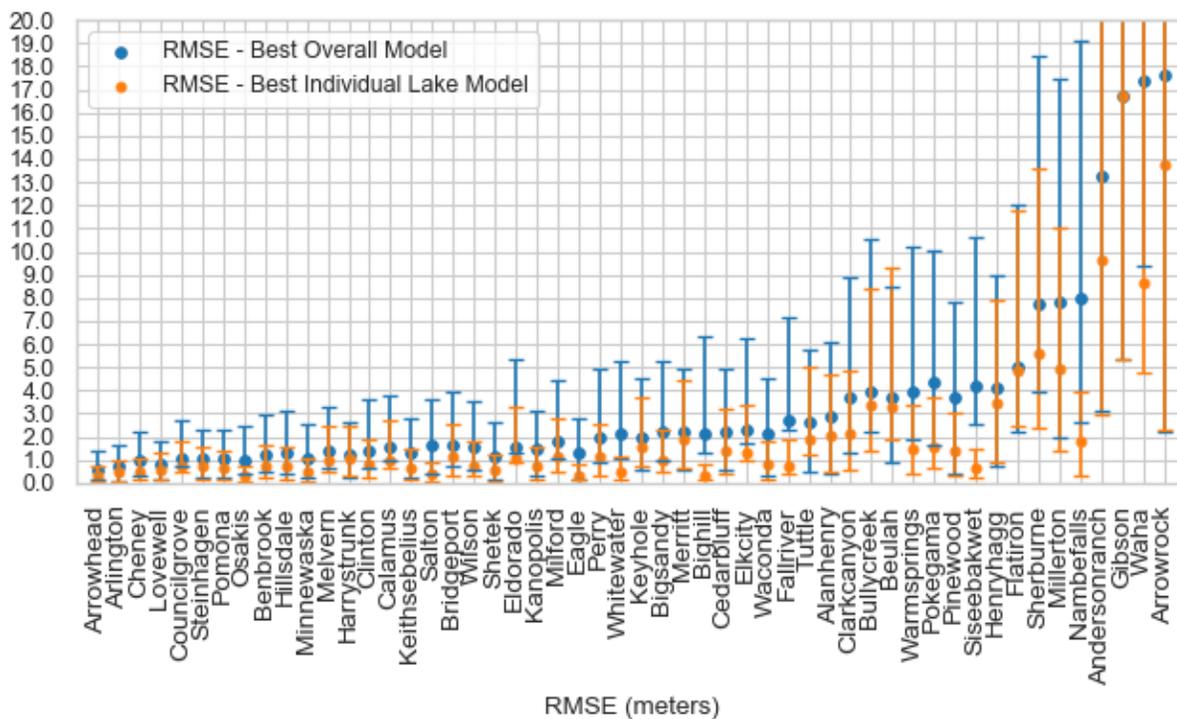


Figure 3-4: Three different types of shoreline water boundaries were used to estimate lake water surface elevation (Weekley & Li, 2019).



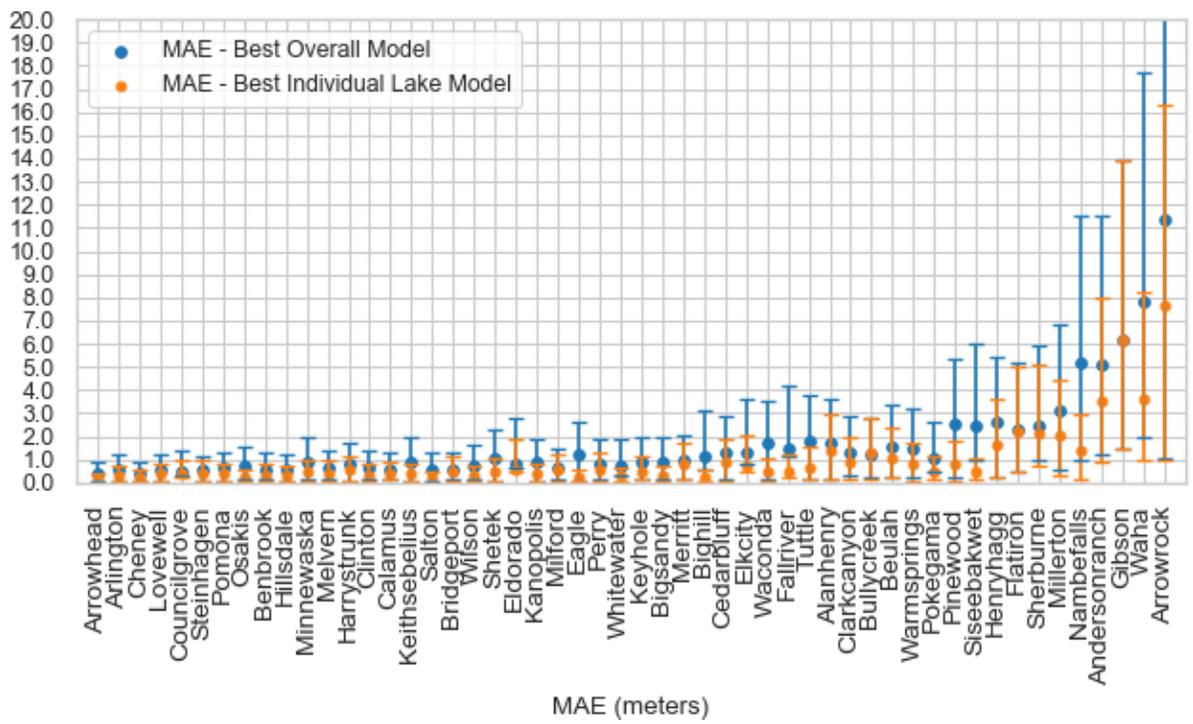


Figure 3-5: RMSE (A) and MAE (B) accuracy for each lake plotted from lowest to highest accuracy using the best overall model. For comparison, the best individual lake models are also shown.

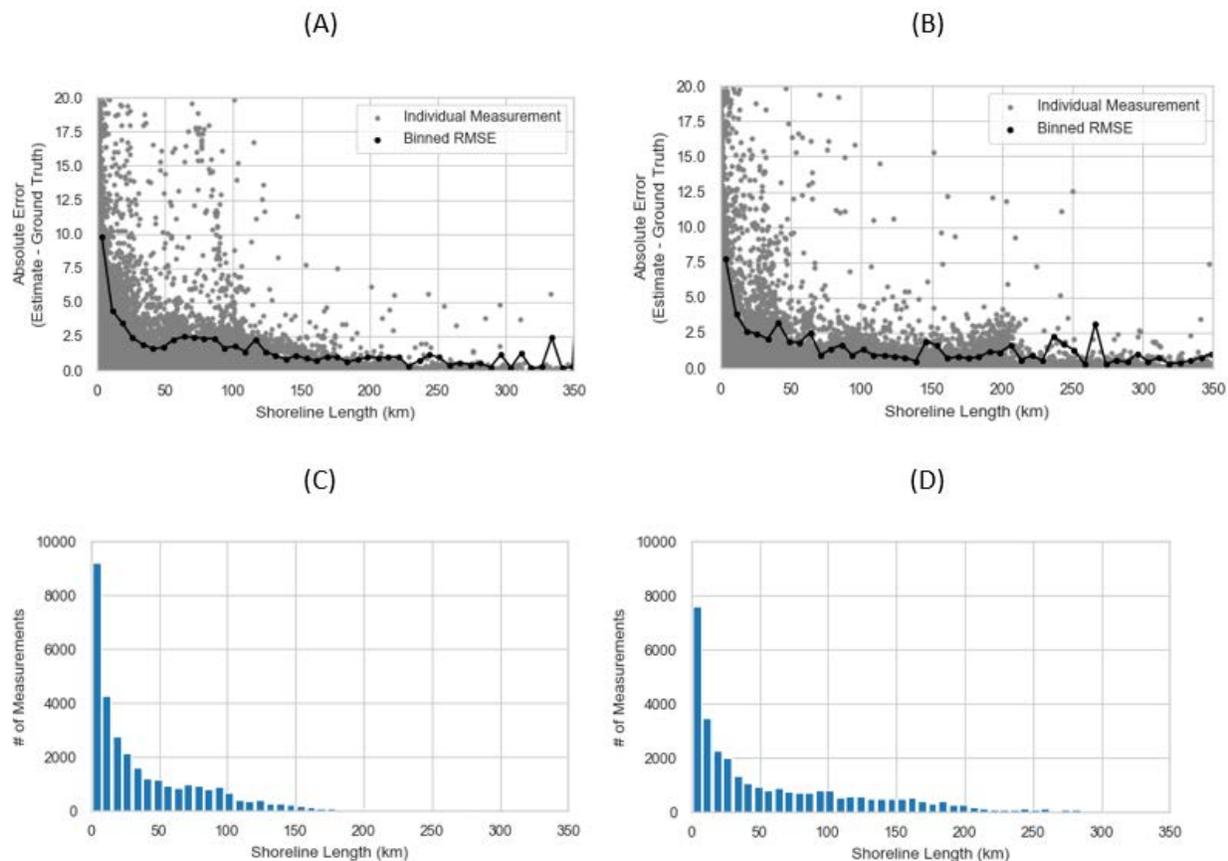


Figure 3-6: A) General model absolute error (estimated surface elevation – ground truth elevation) vs shoreline length. B) Lake-specific model absolute error vs shoreline length. C) General model estimate counts per 7.5 km bin. D) Lake-specific model estimate counts per 7.5 km bin. The charts for absolute error were each capped at 20 meters absolute error and all charts were capped at 350 km shoreline. ~1% of measurements exceeded these thresholds.

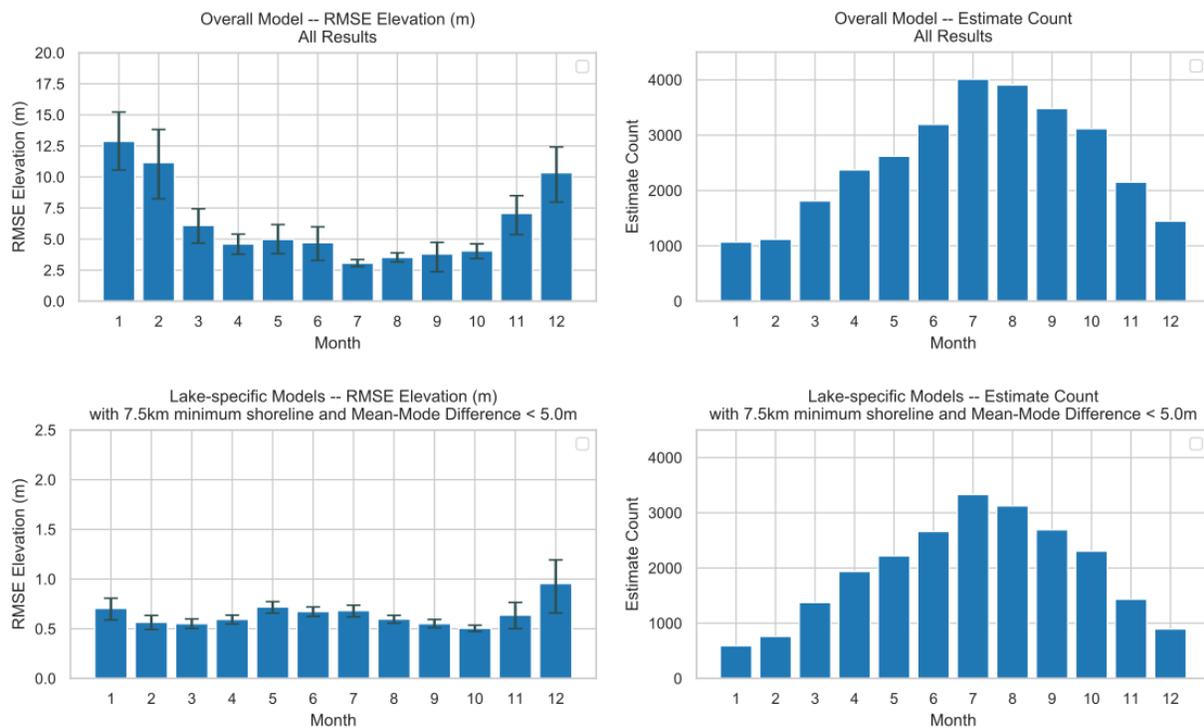
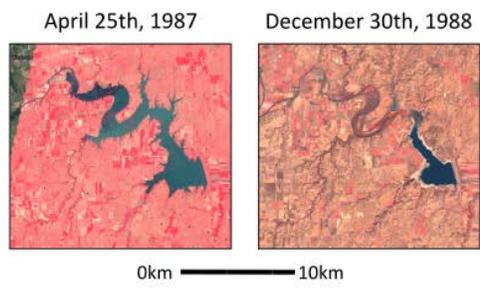
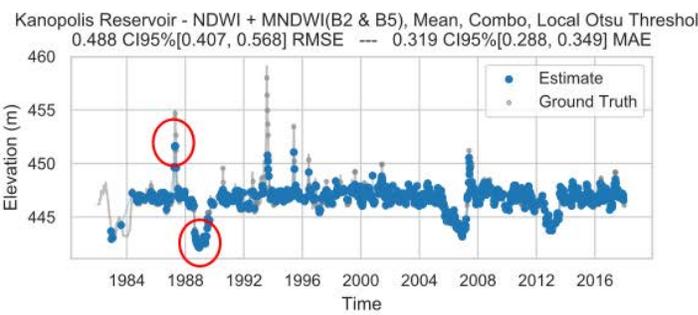
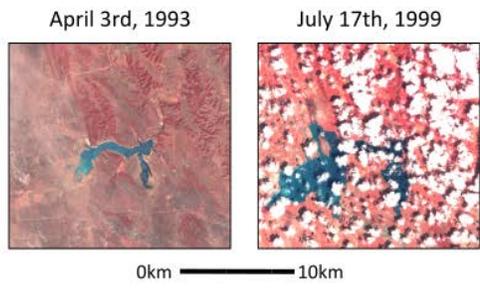
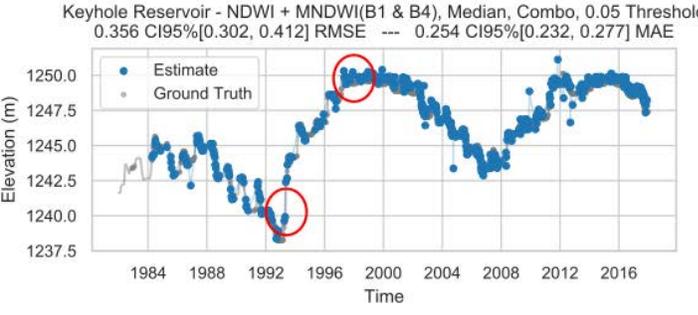
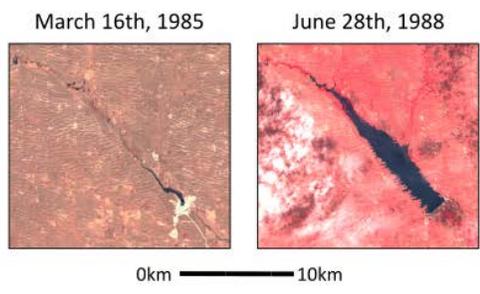
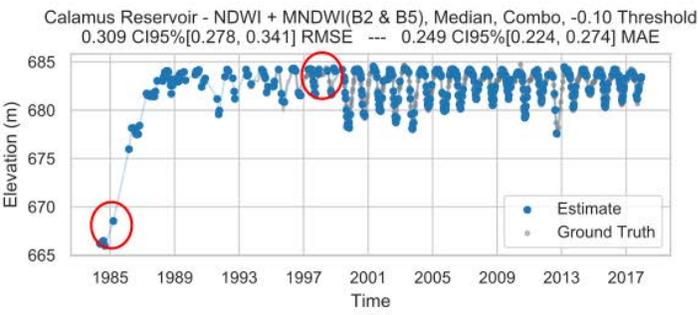
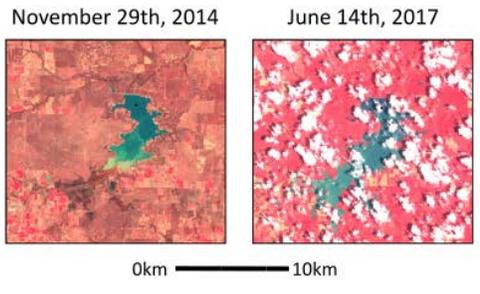
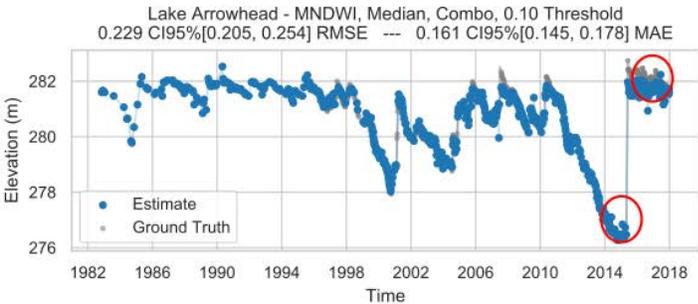
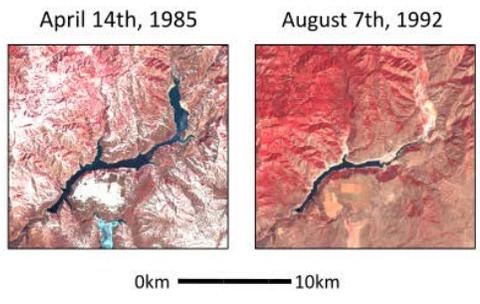
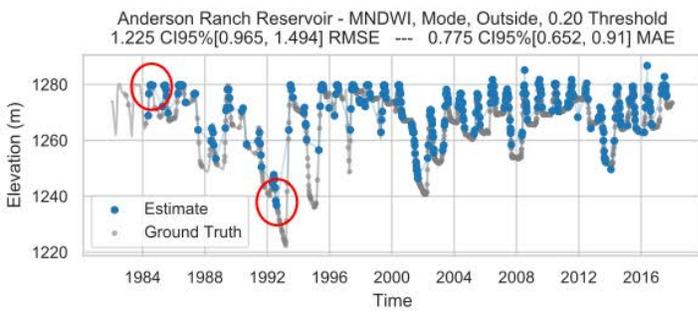


Figure 3-7: Monthly RMSE accuracies and image counts for the best performing overall model and the best lake-specific models with shoreline filtering and mean-mode difference filtering.



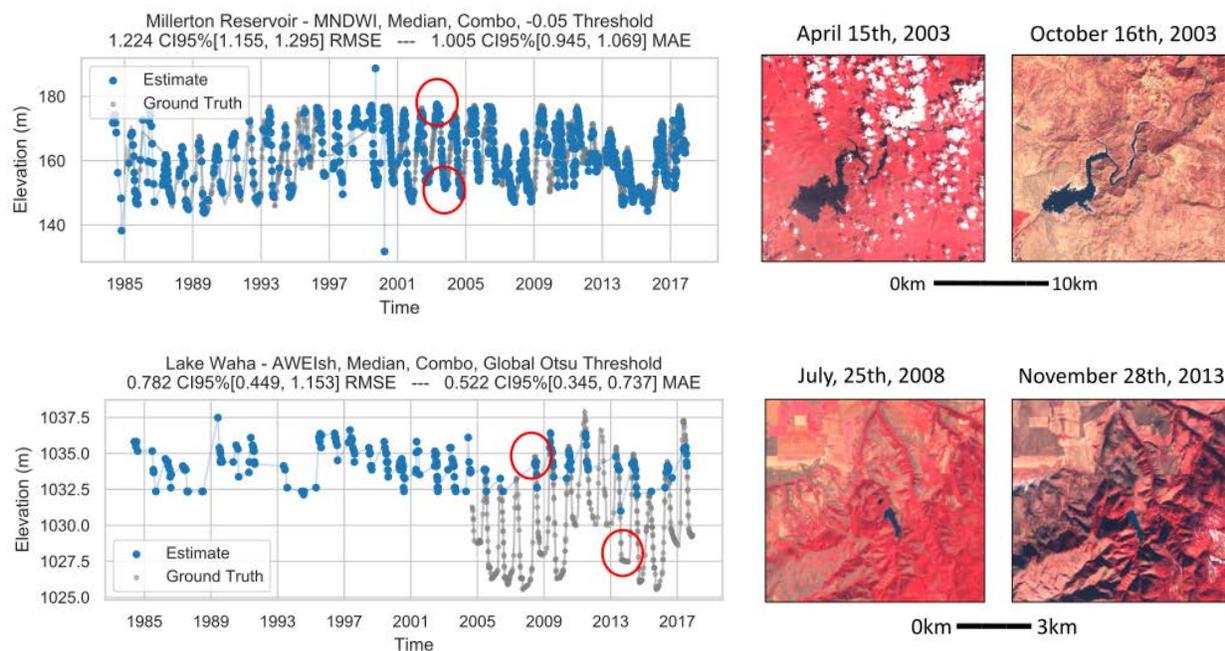


Figure 3-8: Selected water surface elevation time series. Most lakes showed good agreement between the estimated water surface elevations and the ground-truth. Note, the y-scale varies for each chart due to differences in the estimated water levels.

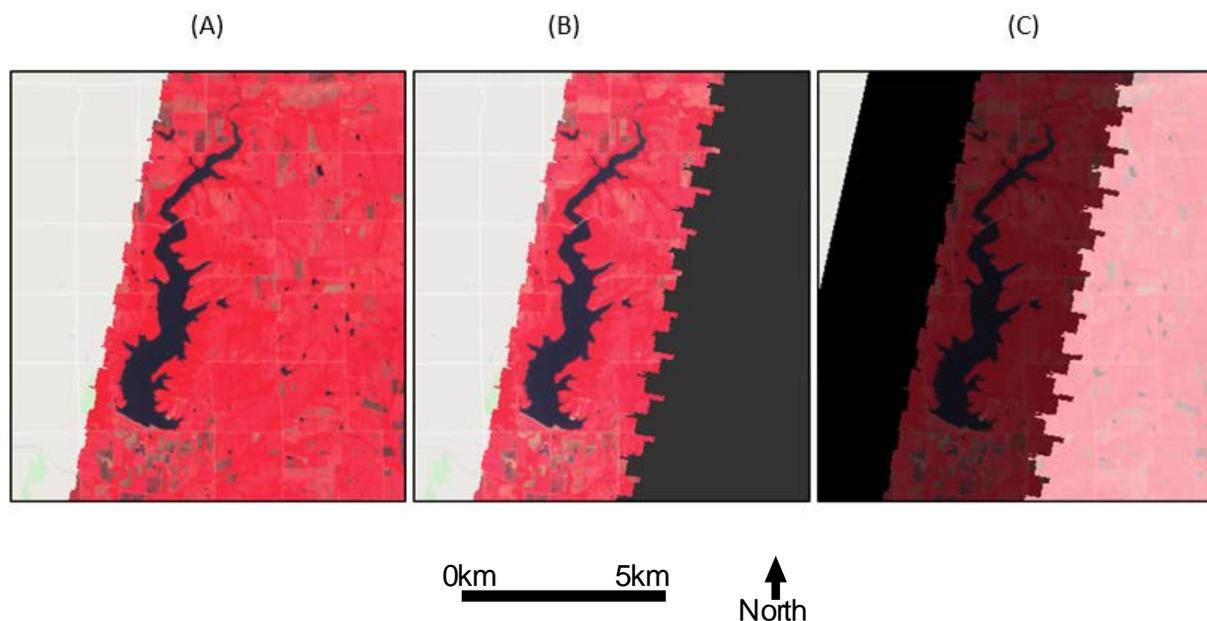


Figure 3-9: A) False-color infrared image of Big Hill Reservoir, Kansas, USA using Landsat 5 TOA Tier-1 Collection-1 image from July 16, 1993. The lake appears to be free of image contamination at first glance, yet no water was detected for this image. B) The same image with the QA Band coverage displaying in grey. C) The impacted areas (dark areas), clear areas (light areas), and unmasked underlying footprint (black). All pixels with non-clear attribution, including no data regions, are considered impacted for the purposes of this study.

Table 3-1: Sensor and mission details for each Landsat mission included in this study including available image date ranges, spatial resolution, temporal resolution, and radiometric resolution.

Mission	Landsat 4	Landsat 5	Landsat 7	Landsat 8
Sensor	Thematic Mapper	Thematic Mapper	Enhanced Thematic Mapper Plus	Operational Land Imager
GEE Date Range	08/22/1982 - 11/18/1993	03/16/1984 - 05/05/2012	05/28/1999 - Present*	03/18/2013 - Present
Band 1	Blue (30m) 0.45 - 0.52 μm	Blue (30m) 0.45 - 0.52 μm	Blue (30m) 0.45 - 0.52 μm	Coastal (30m) 0.43 - 0.45 μm
Band 2	Green (30m) 0.52 - 0.60 μm	Green (30m) 0.52 - 0.60 μm	Green (30m) 0.52 - 0.60 μm	Blue (30m) 0.45 - 0.51 μm
Band 3	Red (30m) 0.63 - 0.69 μm	Red (30m) 0.63 - 0.69 μm	Red (30m) 0.63 - 0.69 μm	Green (30m) 0.53 - 0.59 μm
Band 4	NIR ¹ (30m) 0.76 - 0.90 μm	NIR ¹ (30m) 0.76 - 0.90 μm	NIR ¹ (30m) 0.77 - 0.90 μm	Red (30m) 0.64 - 0.67 μm
Band 5	SWIR ² 1 (30m) 1.55 - 1.75 μm	SWIR ² 1 (30m) 1.55 - 1.75 μm	SWIR ² 1 (30m) 1.55 - 1.75 μm	NIR ¹ (30m) 0.85 - 0.88 μm
Band 6	TIR ³ (60m) 10.40 - 12.50 μm	TIR ³ (60m) 10.40 - 12.50 μm	TIR ³ (60m) 10.40 - 12.50 μm	SWIR ² 1 (30m) 1.57 - 1.65 μm
Band 7	SWIR ² 2 (30m) 2.08 - 2.35 μm	SWIR ² 2 (30m) 2.08 - 2.35 μm	SWIR ² 2 (30m) 2.08 - 2.35 μm	SWIR ² 2 (30m) 2.11 - 2.29 μm
Band 8	-	-	Panchromatic (15m) 0.52 - 0.90 μm	Panchromatic (15m) 0.52 - 0.90 μm
Band 9	-	-	-	Cirrus (15m) 1.36 - 1.38 μm
Band 10	-	-	-	TIR ³ 1 (100m) 10.60 - 11.19 μm
Band 11	-	-	-	TIR ³ 2 (100m) 11.50 - 12.51 μm
QA Band	Collection 1 QA Bitmask			
Radiometric Resolution	8 bit (256 levels)	8 bit (256 levels)	8 bit (256 levels)	16 bit (65,536 levels)**
Revisit Period	16 days	16 days	16 days	16 days

¹Near Infrared, ²Shortwave Infrared, ³Thermal Infrared
 *Operation despite Scan Line Corrector (SLC) failure May 31, 2003
 **When processed into Level-1 data products.

Table 3-2: Bathymetric survey information and vertical adjustment used (where applicable).

Lake or Reservoir	Survey Project Datum	Vertical Adjustment to NAVD88 (meters)	State	Impoundment Year	Survey Year
Alan Henry Reservoir	NGVD29	0.28968	Texas	1993	2017
Anderson Ranch Reservoir	NGVD29	1.01592	Idaho	1950	1998
Lake Arlington	NGVD29	0.00624	Texas	1957	2007
Lake Arrowhead	NGVD29	0.05976	Texas	1966	2013
Arrowrock Reservoir	NGVD29	1.03452	Idaho	1915	1997
Benbrook Lake	NGVD29	-0.01027	Texas	1952	1998
Beulah Reservoir	Local	3.38328	Oregon	1935	2000
Big Hill Lake	NGVD29	0.1396	Kansas	1981	2010
Big Sandy Lake	NGVD29	0.19003	Minnesota	Natural Lake	1998
Lake Bridgeport	NGVD29	0.09306	Texas	1932	2010
Bully Creek Reservoir	Local	0.734568	Oregon	1963	2000
Calamus Reservoir	NGVD29	0.25615	Nebraska	1985	2012
Cedar Bluff Reservoir	NGVD29	0.2751	Kansas	1951	2000
Cheney Reservoir	NGVD29	0.14569	Kansas	1965	2010
Clark Canyon Reservoir	NGVD29	1.21555	Montana	1964	2000
Clinton Lake	NGVD29	0.0841	Kansas	1975	2009
Council Grove Lake	NGVD29	0.14369	Kansas	1964	2008
Eagle Lake	NGVD29	0.17539	Minnesota	Natural Lake	N/A
El Dorado Lake	NGVD29	0.13601	Kansas	1981	2010
Elk City Lake	NGVD29	0.14136	Kansas	1966	2010
Fall River Lake	NGVD29	0.14318	Kansas	1949	2010
Flatiron Reservoir	Local	2.298192	Colorado	1953	2012
Gibson Reservoir	Local	2.286	Montana	1929	2009
Harry Strunk Lake	NGVD29	0.30389	Nebraska	1949	2006
Henry Hagg Lake	NGVD29	1.0582	Oregon	1975	2001
Hillsdale Lake	NGVD29	0.12622	Kansas	1982	2009
Kanopolis Lake	NGVD29	0.13184	Kansas	1948	2007
Keith Sebelius Lake	NGVD29	0.26026	Kansas	1964	2000
Keyhole Reservoir	Local	0.64008	Wyoming	1952	2003
Lovewell Reservoir	NGVD29	0.16896	Kansas	1957	2011
Melvern Lake	NGVD29	0.15444	Kansas	1970	2009
Merritt Reservoir	NGVD29	0.3162	Nebraska	1964	2003
Milford Lake	NGVD29	0.12916	Kansas	1967	2009
Millerton Lake	NGVD29	0.75957	California	1944	2004
Lake Minnewaska	NGVD29	0.19972	Minnesota	Natural Lake	1963
Nambe Falls Reservoir	NGVD29	1.05604	New Mexico	1976	2013
Lake Osakis	NGVD29	0.19487	Minnesota	Natural Lake	1964

Perry Lake	NGVD29	0.08266	Kansas	1966	2009
Pinewood Reservoir	NGVD29	1.09816	Colorado	1952	2012
Lake Pokegama	NGVD29	0.24636	Minnesota	Natural Lake	1940
Pomona Lake	NGVD29	0.13858	Kansas	1965	2009
Salton Sea	NGVD29	0.59741	California	1905	1995
Lake Sherburne	Local	4.191	Montana	1919	2002
Shetek	NGVD29	0.18971	Minnesota	Natural Lake	1959
Siseebakwet	NGVD29	0.25532	Minnesota	Natural Lake	1951
B. A. Steinhagen Lake	NGVD29	0.00128	Texas	1951	2011
Tuttle Creek Lake	NGVD29	0.12448	Kansas	1962	2009
Waconda Lake	Local	-0.19502	Kansas	1969	2001
Lake Waha	NGVD29	1.10815	Idaho	Natural Lake	2006
Warm Springs Reservoir	NGVD29	1.12218	Oregon	1919	2000
Whitewater Lake	NGVD29	0.20048	Minnesota	Natural Lake	N/A
Wilson Lake	NGVD29	0.17543	Kansas	1964	2008

Table 3-3: Overall, 594 different image processing parameter combinations were tested including six water indices, 11 segmentation thresholds (nine pre-set, two dynamic), three boundary types, and three statistic types.

Name	Number of Values	Values
Water Indices	6	NDWI, MNDWI, AWEInsh, AWEIsh, NDWI+MNDWI (Blue & NIR), NDWI+MNDWI (Green & SWIR)
Pre-set Thresholds	9	-0.15 to 0.25 with an increment of 0.05
Otsu Dynamic Thresholding	2	Global Otsu's Method, Local Otsu's Method
Boundary Types	3	Interior, Exterior, Combined
Statistical Types	3	Mode, Median, Mean
Total Tested Combinations	594	

Table 3-4: Accuracy metrics for the best performing general and lake-specific models using various combinations of minimum shoreline lengths and maximum mean-mode difference thresholds to remove erroneous data points. Mean % impacted corresponds to localized image contamination (image contamination within 500 meters of median lake extent), filtered valid images is the number of images remaining after the shoreline length and mean-mode difference filters have been applied, and total estimates are the number of images which produced an elevation estimate.

Model	Minimum Shoreline Length (km)	Maximum Mean-Mode Difference (m)	RMSE (m)	MAE (m)	Mean % Impacted	Filtered Valid Images	Total Valid Images	Filtered Estimates	Total Estimates	Lakes with RMSE <1.0m	Lakes with MAE <1.0m
General Overall #1	0	0	5.746	1.81	20.13%	31,438	31,438	42,249	42,249	4	27
			C195%[5.208, 6.265]	C195%[1.739, 1.886]							
General Overall #2	7.5	0	2.776	0.954	12.21%	22,423	31,857	30,127	42,905	27	32
			C195%[2.209, 3.413]	C195%[0.909, 1.000]							
General Overall #3	0	5	2.328	0.883	18.65%	28,568	31,924	38,972	42,991	15	34
			C195%[1.773, 2.889]	C195%[0.801, 0.866]							
General Overall #4	7.5	5	1.366	0.669	14.46%	24,406	31,924	33,062	42,991	26	40
			C195%[1.181, 1.606]	C195%[0.649, 0.687]							
Lake-Specific #1	0	0	4.327	1.172	19.17%	31,657	31,665	42,669	42,669	25	40
			C195%[3.932, 4.742]	C195%[1.126, 1.220]							
Lake-Specific #2	7.5	0	1.896	0.604	13.00%	23,676	31,807	31,584	42,851	42	44
			C195%[1.564, 2.338]	C195%[0.584, 0.627]							
Lake-Specific #3	0	5	1.418	0.585	17.87%	28,461	31,803	39,036	42,870	39	46
			C195%[1.223, 1.650]	C195%[0.570, 0.601]							
Lake-Specific #4	7.5	5	0.638	0.385	13.04%	23,350	31,457	31,497	42,138	45	49
			C195%[0.616, 0.662]	C195%[0.387, 0.391]							

Table 3-5: Pearson's product-moment correlation r^2 for RMSE, MAE, and potential error drivers.

	Best Overall Model (General Model)		Best Lake Specific Models (Locally Optimized)	
	RMSE	MAE	RMSE	MAE
Mean Slope	0.892 CI95%[0.799, 0.948]	0.824 CI95%[0.714, 0.912]	0.801 CI95%[0.659, 0.908]	0.793 CI95%[0.680, 0.899]
Mean % Impacted	0.165 CI95%[0.012, 0.393]	0.128 CI95%[0.000, 0.372]	0.17 CI95%[0.026, 0.390]	0.162 CI95%[0.009, 0.377]
Median Lake Area	0.059 CI95%[0.007, 0.203]	0.063 CI95%[0.010, 0.225]	0.048 CI95%[0.009, 0.159]	0.052 CI95%[0.009, 0.180]
Shoreline Perimeter Length	0.074 CI95%[0.006, 0.182]	0.087 CI95%[0.006, 0.225]	0.048 CI95%[0.000, 0.139]	0.048 CI95%[0.001, 0.145]
Surface Area to Perimeter Ratio	0.174 CI95%[0.070, 0.302]	0.178 CI95%[0.067, 0.315]	0.145 CI95%[0.058, 0.27]	0.171 CI95%[0.079, 0.296]
# of Images with Estimates and Ground-Truth	0.024 CI95%[0.000, 0.108]	0.018 CI95%[0.00, 0.073]	0.05 CI95%[0.001, 0.167]	0.041 CI95%[0.000, 0.130]

Table 3-6: Accuracy metrics for each Landsat mission. Accuracy statistics for total overall, SLC-On, and SLC-Off are provided for Landsat 7 ETM+.

Best Overall Model - No filtering				
Mission	RMSE	MAE	Mean % Impacted	Valid Image Count
LT04	4.544m CI95%[2.699, 6.484]	2.152m CI95%[1.463, 2.959]	14.75%	112
LT05	5.415m CI95%[4.74, 6.142]	1.588m CI95%[1.502, 1.675]	13.98%	14432
LE07 (All)	6.148m CI95%[5.736, 6.6]	2.208m CI95%[2.112, 2.313]	27.64%	12125
LE07 (SLC-On)	5.159m CI95%[4.194, 6.121]	1.519m CI95%[1.35, 1.71]	14.61%	2451
LE07 (SLC-Off)	6.37m CI95%[5.864, 6.875]	2.381m CI95%[2.265, 2.501]	30.94%	9674
LC08	6.318m CI95%[4.62, 8.197]	1.737m CI95%[1.545, 1.96]	16.66%	3630

Best Lake-Specific Models with Filtering (>7.5km shorelines and Mean-Mode Difference < 5.0m)				
Mission	RMSE	MAE	Mean % Impacted	Valid Image Count
LT04	0.485m CI95%[0.375; 0.602]	0.346m CI95%[0.279; 0.414]	8.74%	95
LT05	0.594m CI95%[0.574; 0.618]	0.368m CI95%[0.36; 0.377]	10.38%	12347
LE07 (All)	0.695m CI95%[0.652; 0.747]	0.41m CI95%[0.397; 0.422]	19.31%	8008
LE07 (SLC-On)	0.55m CI95%[0.519; 0.583]	0.355m CI95%[0.336; 0.372]	10.91%	2175
LE07 (SLC-Off)	0.742m CI95%[0.685; 0.808]	0.43m CI95%[0.415; 0.445]	22.44%	5833
LC08	0.655m CI95%[0.586; 0.737]	0.387m CI95%[0.369; 0.407]	12.04%	2887

4 Tracking Lake Surface Elevations with Proportional Hypsometric Relationships, Landsat Imagery, and Multiple DEMs

4.1 Abstract

Multi-decadal inland surface water dynamics are of increasing interest due to their widespread influence and importance to climate, agriculture, ecology, industry, and society yet, despite their importance, several key challenges impede long-term monitoring of inland surface waters globally. First, the most accurate lake dynamic tracking methods utilize bathymetric survey data to construct hypsometric relationships, also known as storage curves, to estimate water quantities. This research investigates two novel methods, one to address sub-hydroflattened surface estimate uncertainty issue, and a second to address limited temporal resolution issues, using 46 lakes and reservoirs spread across the western United States. First, low water level estimate uncertainty was reduced using multiple digital elevation models (ALOS, SRTM, and NED) and the hypsometric relationship for each lake was derived from the digital elevation model with the lowest hydroflattened water surface. This technique reduced the number of images with sub-hydroflattened water surfaces by at least 549 over the best individual DEM leading to increased accuracy by reducing the extrapolation distance for any remaining sub-hydroflattened water surface levels while simultaneously increasing the number of estimates above the hydroflattened surface. Second, this chapter introduces proportional hypsometry which dynamically generates surface area/elevation relationships for every image using clear pixel observation areas only (contamination is removed from both the image and underlying DEM). Proportional hypsometry was found to be ill-suited for sub-hydroflattened water surface levels but produced comparable accuracy to clear images for above hydroflattened water levels. Overall, using the lowest hydroflattened surface

along with proportional hypsometry improved the temporal resolution by enabling the analysis of nearly 10,000 additional images while maintaining an accuracy level similar to images with <1% contamination (2.35 m RMSE and 1.61 m MAE vs 2.17 m RMSE and 1.53 m MAE). Overall, this research decreases hypsometric analysis uncertainty while increasing temporal resolution and is scalable within cloud computing platforms such as Google Earth Engine enabling wide-scale regional and global water dynamic analysis.

4.2 Introduction

Multi-decadal inland surface water dynamics are of increasing interest due to their universal importance and widespread impact on climate, industry, agriculture, ecology, and society (Prigent et al., 2012). At the most basic level, water surface dynamics represent the spatiotemporal distribution of water on the landscape and describe where, when, and how much water is present using four key metrics: surface elevation, surface area, volume, and volume change (Weekley & Li, 2019). However, despite its universal importance, inland surface waters are poorly understood, and several key challenges impede long-term monitoring of inland surface waters globally.

The first issue with large-scale, long-term lake dynamic monitoring is obtaining each of the necessary measurements. As mentioned, the four common lake dynamic measurements are surface elevation, surface area, volume, and volume change. Of the four metrics, surface elevation and surface area are the easiest to obtain and are the only two that are directly and independently measurable using current instruments and techniques. Volume and volume change, on the other hand, are more difficult to obtain. Volume change requires, at minimum, surface elevation and surface area measurements from two separate dates to estimate water volume change using conical or pyramidal frustum equations (Abileah & Vignudelli, 2011; J F Crétaux et al., 2016; Gao, 2015) while total water volume requires knowledge of underlying topography (bathymetry) which is only available for a small subset of global lakes and reservoirs (Alsdorf et al., 2007; Peng, Guo, Liu, & Liu, 2006).

Despite the difficulties, several methods have been developed to measure lake water dynamics at various scales with most of the focus, at least historically, being on surface elevation and surface area measurements. Surface elevation is commonly measured using gauges installed on-site. As a result of their manual installation requirements, in-situ hydrological stations are only available for a small subset of global water bodies (Alsdorf et al., 2007). Surface elevation is also dependent upon several factors, such as weather conditions which can cause variations in surface elevation across a single water body meaning that individual hydrological stations, which are point measurements, may not be representative of the water body as a whole (Alsdorf et al., 2007). Due to the uneven distribution and limited availability of in-situ monitoring stations, remote sensing has commonly been used to supplement or outright replace in-situ measurements. The most common method of obtaining water surface elevation via remote sensing is through satellite altimeters such as the RA (radar altimeter) sensor onboard European Space Agency's ERS-1 (European Remote Sensing) satellite ("ERS-1: Aviso+," n.d.) or the Poseidon-2 altimeter onboard the joint NASA (National Aeronautics and Space Administration) and CNES (French Space Agency) Jason-1 satellite. Altimeters, which can be either radar- or laser-based, are non-imaging, nadir-looking, profiling instruments which estimate height (distance from sensor) by precisely measuring the return time of signals reflected off the surface. Satellite altimetry, which has been used in numerous studies including works by Medina et al. (2008), Duan and Bastiaanssen (2013), Zhu et al. (2014), Crétaux et al. (2015), Schwatke et al. (2015), Dettmering et al. (2016), Göttl et al. (2016), and Okeowo et al. (2017), has repeatedly demonstrated high-accuracy water surface elevation estimates while certain conditions are met. Satellite altimetry works best for large water bodies with long pass-overs with minimal contamination from surrounding terrain and under such conditions is capable of sub-decimeter accuracy (3 – 10 cm) (Solander et al., 2016). Unfortunately, the current/historic altimeter fleet was designed for oceanographic and/or cryosphere mapping and therefore has large ground-track spacings and poor spatial resolution which limit its applicability for most water bodies. In fact, it's estimated that only 60% of water bodies larger than 100 km² can be monitored for any appreciable period using current and historic satellite altimeters (Solander et al., 2016).

While direct water surface elevation measurements may not be available for most water bodies due to the lack of in-situ monitoring or altimetry-based measurements, water surface area, assuming accurate detection methods are employed, is relatively easy to obtain and can be directly measured using optical and synthetic aperture radar imagery. Water surface area measurements are also scale-able to regional and global levels (Pekel et al., 2016; Tulbure & Broich, 2013; Tulbure et al., 2016). While useful in a wide range of applications, water surface area remains a two-dimensional measurement whereas, at minimum, water volume change estimates are needed to improve our understanding of surface water dynamics and its effects upon climate and other fields. To extend the usefulness of water surface area measurements and fill in the gap of missing water surface elevations due to the limitations of in-situ and altimeter-based monitoring, several studies have used elevation data products such as digital elevation models and/or bathymetry to estimate water dynamics using hypsometric relationships relating surface area to volume and/or surface elevation (Avisse et al., 2017; Pan, Liao, Li, & Guo, 2013; Zhang et al., 2016) or elevations extracted from water surface boundaries (Tseng et al., 2016; Weekley & Li, 2019). Unfortunately, each of these methods has limitations in terms of expanding and scaling their applications. The water boundary approach requires full or estimated bathymetric profiles to ensure water surface elevation estimates can be made if water levels are lower than they were at the time the digital elevation model was collected. Like altimetry, bathymetric surveys are only available for a small subset of lakes globally which limits its application for global analysis (Alsdorf et al., 2007) while efforts to simulate bathymetric profiles from surrounding terrain have reported resolution scale dependencies and large uncertainties from lake to lake (Heathcote et al., 2015; Messenger et al., 2016; Tseng et al., 2016). Hypsometric relationships, similarly, can be extrapolated to lower water levels from higher water levels by extending the slope of the storage curve using linear/polynomial regression and/or power relationships (Pan et al., 2013); however, concerns about accuracy uncertainty at increasingly lower water levels, relationship fit, and application to varying water bodies remain. Furthermore, most hypsometry implementations rely upon having full, clear observations to accurately measure water surface area and any image contamination (clouds, cloud shadow, snow, ice, sensor error) reduces the estimated water

surface area resulting in erroneous elevation, volume, and volume change estimates. This issue is substantial as many areas of the planet have significant cloud cover and/or other image contamination for long periods each year and restricting water surface dynamic estimates to clear images only results in poor, uneven temporal resolution (Huang et al., 2018) and methods of filling data gaps and cloud contamination using data-filling or temporal windows are best suited only for larger water-bodies (Ogilvie et al., 2018).

This study presents a novel approach to improve digital elevation model (DEM) hypsometry-based surface water dynamic monitoring. First, this research uses multiple digital elevation models (ALOS – Advanced Land Observing Satellite, SRTM – Shuttle Radar Topography Mission, and NED – National Elevation Dataset) (Archuleta et al., 2017; Farr et al., 2007; Tadono et al., 2014) to select the lowest hydroflattened water surface for each lake which reduces the total number of sub-hydroflattened water surface estimates and decreases the uncertainty of the remaining sub-hydroflattened water surface estimates by reducing the extrapolation distance to those water levels. Second, this research introduces proportional imagery/DEM hypsometric relationships which extend traditional hypsometry to contaminated imagery by removing image contamination from both the imagery and the underlying digital elevation model and constructs a new hypsometric surface area/elevation relationship for each satellite image. Forty-six lakes throughout the western and midwestern United States were used to test the method which uses the entire Landsat 4, 5, 7, and 8 Tier-1 Collection-1 Top-of-Atmosphere image archives. This research addresses the following gaps in hypsometric surface water dynamic monitoring: 1) analyzes and compares regression methods of extending hypsometric curves from above hydroflattened surface to sub-surface levels; 2) investigates methods of reducing sub-hydroflattened water surface elevation estimate uncertainty; 3) analyzes hypsometric relationship estimate techniques for a variety of water bodies with varying geologies, shape, sizes, water qualities, climate regimes, management practices, and ecotypes; 4) identifies and mitigates a variety of image contamination sources to substantially improve the temporal resolution of hypsometry-based water surface dynamic monitoring;

and, finally, 5) this chapter pioneers a scalable solution using hypsometric relationships within Google Earth Engine (Gorelick et al., 2016) enabling global water dynamic monitoring using the entire Landsat 4, 5, 7, and 8 archives.

4.3 Data and Study Area

This research was applied to forty-six lakes and reservoirs throughout the western and midwestern United States (Figure 4-1). These forty-six lakes encompass a wide range of climate regimes, ecotypes, management practices, geologies and are diverse in their size, shape, and water qualities.

As mentioned, this research utilized three different digital elevation models: ALOS DEM (Advanced Land Observing Satellite) (Tadono et al., 2014), SRTM (Shuttle Radar Topography Mission) (Farr et al., 2007), and NED (National Elevation Dataset) (Archuleta et al., 2017). ALOS and SRTM are global datasets featuring 30-meter spatial resolution while NED is a 1/3 arc-second seamless DEM covering the lower 48 states (United States). Additionally, elevation values in ALOS and NED are decimal-point data, while SRTM elevations are stored as integers. For the purposes of this research, NED data was resampled to 30 meters to match the spatial resolution of ALOS and SRTM. NED is vertically referenced to NAVD88 (North American Vertical Datum of 1988) while SRTM and ALOS are referenced to the EGM96 geoid. Due to the close approximation of NAVD88 to the geoid used in EGM96, no vertical datum transformation was performed and each was used as-is (“SRTM NED Vertical Differencing,” n.d.), however, the vertical accuracy of each DEM varies spatially with lower accuracy typical in areas with greater vertical relief such as mountainous areas (Alganci, Besol, & Sertel, 2018; “What is the vertical accuracy of the 3D Elevation Program (3DEP) DEMs?,” n.d.). Other global DEMs, like TanDEM-X (Rizzoli et al., 2017; Zhang et al., 2016), were not included in this study as they are not currently available in GEE (Gorelick et al., 2016). Future analyses will incorporate more DEM options.

In-situ gauge data, which was used for model selection and accuracy assessment, was obtained from multiple sources including the United States Geological Survey (USGS), the Minnesota Department

of Natural Resources (MNRD), and the California Department of Water Resources (CDWR) (“California Data Exchange Center,” n.d.; “Lakes Data for Minnesota,” n.d.; “USGS Lakes and Reservoirs,” n.d.). Depending upon the lake, in-situ gauge data are provided in a variety of vertical datums. Where necessary, these data were adjusted to NAVD88 using VERTCON (“VERTCON,” n.d.).

The entire Landsat 4, 5, 7, and 8 Collection 1 Tier 1 Top-of-Atmosphere (TOA) image archives (Table 4-1) from August 1982 through September 2019 were used in this research which provides multi-decadal, medium resolution (30 meter), and moderate temporal resolution analysis for most water bodies around the globe. Other sensors, such as Moderate Resolution Imaging Spectrometer (MODIS) (Gao, 2015) or RapidEye (Heine, Stüve, Kleinschmit, & Itzerott, 2015), may offer improved temporal and/or spatial resolution but lack the long-term data record required for climate scale analysis. Tier 1 Landsat scenes have the highest available data quality with Level-1 Precision and Terrain (L1TP) correction, well-characterized radiometry, and inter-calibration across the various Landsat sensors making the collections well-suited for time-series analysis (Wulder et al., 2016). In addition to the sensors bands, Level-1 Landsat products also include a Quality Assessment (QA) (USGS, 2018b, 2019) band which provides per-pixel bit values corresponding to surface, atmospheric, and sensor conditions like cloud cover, snow/ice, shadow, dropped pixels, and radiometric saturation (USGS EROS, 2017). Surface reflectance is also available for Landsat, however, the USGS reports that algorithms currently used to produce surface reflectance have not been optimized to work over surface water (“Landsat Surface Reflectance,” n.d.).

4.4 Methods

The general processing procedure is illustrated in Figure 4-2 and consists of three distinct sections. The first section (orange colored) involves selecting the target water body and generating a seed from the JRC Global Surface Water Explorer occurrence layer (Pekel et al., 2016). The second section (green colored) involves pre-processing the Landsat image using a pre-set, naïve segmentation threshold (0.15)

to determine if the image is suitable for further processing. If water is detected within the image and a water boundary can be determined, then the process can continue to the next section. The final section (blue colored) involves dynamically estimating an image specific segmentation threshold using Otsu's method (Otsu, 1979), then repeating the necessary steps to isolate the targeted water body, remove image contamination, estimate the observable water surface area, generate an image specific surface area/elevation relationship, and finally estimate water surface elevation. Further explanation for several key steps is provided below.

4.4.1 Lake Seeding

In order to isolate the targeted water body for subsequent analyses, two types of seeds were utilized. First, the targeted water body was manually seeded using a line placed within the water body (in the future, this step will be automated). Using the manual seed, a larger, more complete seed was generated using the JRC Global Surface Water Explorer occurrence layer (Pekel et al., 2016). In contaminated images, the larger seed is more likely to intersect disconnected water areas than the smaller linear seed. For each lake, areas with $\geq 85\%$ occurrence were selected and then isolated using the manual seed to keep each lake separate, and then converted to a vector for future use. Ideally, areas with 100% occurrence would be used, however, for some lakes the 100% occurrence areas were patchy, so a lower threshold was chosen.

4.4.2 Water and Shoreline Detection

For this research, water detection was completed using the Modified Normalized Water Difference Index (MNDWI) (Xu, 2006) expressed as the following:

$$MNDWI = \frac{Green - MIR}{Green + MIR} \quad (1)$$

Water indices such as MNDWI, in comparison to other methods such as supervised or unsupervised image classification techniques, are computationally efficient algorithms making them suitable for large-

scale analysis if an optimal water segmentation threshold can be determined. Multiple water detection strategies exist within the literature, including NDWI (Normalized Water Difference Index) (McFeeters, 1996), AWEI (Automated Water Extraction Index) (Feyisa et al., 2014), and Tasseled Cap Wetness (Baig, Zhang, Shuai, & Tong, 2014; Crist & Cicone, 1984). Rather than comparing the effectiveness of different water detection algorithms, this research is primarily focused upon integrating multiple digital elevation models which can be done using a single water detection method. MNDWI was chosen as it is a popular water index with a long publication history (Duan & Bastiaanssen, 2013; Lu, Ouyang, Wu, Wei, & Tesemma, 2013; Ma et al., 2019; Rokni et al., 2014; Tulbure et al., 2016). Unfortunately, when utilizing water index-based water detection methods such as MNDWI, no one ideal segmentation threshold exists as a result of changing land, water, and atmospheric conditions. To alleviate this constraint, a two-pronged approach was developed for this research to account for these changing conditions and improve the water thresholding process. First, the image pre-processing section used a naïve segmentation threshold of 0.15 to preliminarily determine if water, and more importantly a water boundary, was observable within the image. If a water boundary was detectable, Otsu's method (Otsu, 1979), which is an adaptive image processing technique used to split an image or dataset into two classes based upon the observed distribution of pixel values, was then used to determine an image specific local segmentation threshold. Otsu's method works best when each class (water and non-water in this case) has roughly the same number of pixels. In this implementation, the classes (water and non-water) were rebalanced for each image by buffering the naively detected water boundary by 150 meters, masking and removing image contamination from the water index image (also aggressively expanded by 150 meters to ensure boundaries along contaminated edges were completely removed), and then applying Otsu's method to the remaining water index pixels.

4.4.3 Image Contamination

Common image contamination sources include cloud, cloud shadow, ice, snow, and terrain shadow but also include sensor issues such as oversaturation, pixel drop out, and issues such as the scan-

line corrector error which affected Landsat 7 images after May 31st, 2003. Several image contamination steps were utilized in this research to quantify and alleviate the effects. As mentioned earlier, the Landsat Tier-1 Collection-1 TOA datasets include Quality Assessment bands providing per-pixel bit values useful for identifying clear and non-clear pixels within a given image (USGS, 2018b, 2019). To supplement this information and improve analysis in areas affected by terrain shadow, date/time-specific terrain shadows were generated using the applicable digital elevation model. In short, all non-clear image pixels, terrain shadows, and uncovered areas within 500 meters of the JRC GSW (Pekel et al., 2016) observed water extent were treated as image contamination.

4.4.4 Hypsometric Relationships

One of the primary goals of this research was to improve elevation estimate accuracy by selecting the DEM with the lowest water surface elevation. This was done using manually created linear seeds to estimate the water surface elevation in that region. The JRC GSW-derived seeds were not used here as their extent may extend beyond the hydroflattened surface for some lakes. For the purposes of this chapter, the lowest water surface DEM model will be referred to as the best combined DEM model. While water surfaces in SRTM and NED are generally hydroflattened, these surfaces often have some variation in the elevation. Furthermore, visual inspection of the target lakes revealed much rougher water surfaces in the ALOS dataset than in the NED and SRTM datasets. To deal with the variations in the surface elevations, the mean elevation along the linear seeds was calculated for each lake and each DEM and were then rounded up to the nearest integer value (this value could be adjusted on a lake by lake basis but was sufficient for the lakes used in this study). Using the estimated water surface level, water surface area was then estimated at 1-meter intervals. To extend the method to contaminated images, the hypsometric relationships were re-created for every image with an observable water shoreline by masking the image contamination from the applicable DEM. Using these relationships, water surface elevation can be estimated using the estimated water surface area in each remaining image. Figure 4-3 provides examples of clear and contaminated image hypsometric relationships.

4.4.5 Model Selection and Accuracy

Model selection and accuracy were determined by comparing estimated lake surface elevations to in-situ lake elevation measurements using root-mean-squared-error (RMSE) and mean-absolute-error (MAE) calculated using a 60/40 train-test split repeated with 1,000 iterations. The best model was determined using the lowest combined MAE and RMSE (ordinally ranked and summed) from the training datasets. Testing accuracy is reported using both MAE and RMSE from the test samples.

4.5 Results

Table 4-2 summarizes the Landsat images used in the analysis. All told, 71,598 Landsat images were found to overlap the 46 targeted water bodies partially or totally. Of those images, 41,459 (57.9%) had observable water shorelines and 32,899 (45.9%) of those with observable shorelines had corresponding in-situ elevation measurements suitable for model selection and accuracy analysis. Additionally, 244 of the 14,170 images with <1% image contamination had no observable shorelines due to spurious pixels along the water shoreline being flagged in the Pixel QA band resulting in the complete removal of the shoreline.

As stated, one of the primary goals of this study was to improve water surface elevation estimate accuracy and reduce low water level estimate uncertainty by constructing hypsometric relationships from the DEM (ALOS, NED, or SRTM) with the lowest water level for each lake. Table 4-3 compares the elevation estimate accuracy for each DEM and the lowest hydroflattened surface DEM (LHFS DEM) for each lake fit with 1st, 2nd, and 3rd-order polynomials for all images with less than 1% image contamination. This analysis represents traditional hypsometry where a single relationship model is used, and the analysis is constrained to contamination-free imagery. Using those criteria, the best combined DEM model fit with a 2nd-degree polynomial had the highest combined RMSE and MAE accuracy.

Unfortunately, restricting hypsometric relationships to clear images only (<1% local contamination) severely limits the temporal resolution of the analysis and in many cases would eliminate

all winter images from potential analysis in areas with frequent snow, ice, and terrain shadow conditions (Table 4-2). To improve upon this, proportional hypsometric relationships which allow analysis of contaminated imagery were also analyzed using all images with observable water boundaries and in-situ surface elevation measurements (Table 4-4). Using these criteria, the LHFS DEM model once again had the highest overall accuracy, however, the performance gain was limited to 1st-order equations due to significant uncertainty at low water levels for images at sub-hydroflattened water levels. This effect is visible in the Figure 4-3 examples where the 2nd and 3rd-order relationships experience significant curving at sub-hydroflattened surface water levels (the curve direction could be downwards or upwards depending upon the amount and distribution of image contamination within a given image).

To further improve the overall accuracy of the analysis, elevation estimates for sub-hydroflattened water levels (an image is considered sub-hydroflattened if the estimated surface area is less than the surface area of the lowest DEM level) (Table 4-5) and above hydroflattened surfaces (Table 4-6) were analyzed separately. Using the results from Table 4-5 and Table 4-6, a single model using all images above the hydroflattened water surface and fitting them with a 3rd-degree polynomial to maximize the fit but restricting sub-hydroflattened surface estimates to just those with <1% image contamination and fitting them with a 2nd-degree polynomial was proposed and analyzed (Table 4-7). While this model produces fewer estimates than the models used in Table 4-4, it provides significantly higher accuracy than using all the images and maintains most of the temporal resolution improvements gained from using proportional hypsometry. The accuracy improvements of this model and the proportional hypsometric technique are illustrated in Figure 4-4. Figure 4-4 shows a rapid decline in elevation estimate accuracy when including all images above and below the hydroflattened water surface (line 2-green) but also shows consistent accuracy when using proportional hypsometry to include all above hydroflattened surface images and limiting sub-hydroflattened surface estimates to images with $\leq 1\%$ local image contamination.

While the results of the proposed methods improve the overall accuracy for all 46 lakes, at the individual lake level the results are much more mixed. Overall, just nine lakes produced sub-meter RMSE accuracies and 21 lakes produced sub-meter MAE accuracies using the LHFS DEM method. Conversely, 13 lakes had RMSEs and 11 had MAE accuracies greater than 2.5 meters. Figure 4-5 provides a time-series look at several lakes ranging from high accuracy to low accuracy using the Table 4-7 model with the LHFS DEM and proportional hypsometry where all images above the hydroflattened water surface fit with a 3rd-degree polynomial and sub-hydroflattened surface estimates with <1% image contamination fit with a 2nd-degree polynomial.

4.6 Discussion

4.6.1 Digital Elevation Model Differences

One of the key uncertainty sources which affected the accuracy of this analysis involved noted differences between the three digital elevation models beyond just the height of the water level (hydroflattened surface) on the DEM. Figure 4-6 provides several examples highlighting the differences between the measured surface areas at common surface elevations. Several lakes display significant difference in the measured surface areas such as Lake B.A. Steinhagen and Anderson Ranch Reservoir. In these cases, water surface elevations estimated using the NED are tens of meters less accurate than their SRTM or ALOS counterparts based upon comparison with in-situ measurements. Figure 4-7, which compares hillshades and inundated surface areas for B.A. Steinhagen Lake at 30 meters along with hillshades for Anderson Ranch Reservoir at 1,278 meters, shows the large difference in surface area associated with each DEM. For B.A. Steinhagen, compared to ALOS or SRTM, the NED DEM is significantly smoother due to its 1/3 arc-second native resolution, however, the elevations are also significantly lower across large areas leading to substantially larger surface areas at each elevation while for Anderson Ranch Reservoir, the southwestern end of the reservoir near the dam is non-hydroflattened in the NED resulting in massive surface area differences between the DEMs. As mentioned earlier, the vertical accuracy of each of these data sets varies spatially. While determining which DEM is most

accurate in each location is difficult without external reference data, in general, these differences illustrate the strength of using multiple DEMs to select the best DEM for each water body. In fact, some of the bias visible in time series of some lakes may be due to vertical accuracy deviations in that specific area and could be corrected given supplementary elevation data such as elevation measurements from the upcoming SWOT (Surface Water and Ocean Topography) mission set to launch in 2021 (Biancamaria et al., 2016; Solander et al., 2016).

4.6.2 Image Contamination

The quality assessment bands available for each image within the Landsat archive make rapid assessment of local conditions easy and effective, however, limitations do exist. In many cases, unidentified and misclassified image contamination affects the estimated water surface elevation by changing the amount of water detected within the target area. Any erroneous or missed water affects not only the estimated water surface area, which alone impacts the water surface elevation estimate, but is also unlikely to be accounted for in the proportional hypsometric relationship further leading to increased error and uncertainty.

4.6.3 Time Series and Individual Lake Accuracy

The time series for each lake explored in Figure 4-5 highlighted some of the difficulties of using digital elevation models to estimate water surface elevation. As noted, some lakes, especially isolated lakes with relatively consistent slopes were more accurate than lakes in high slope areas (canyons or mountains) and, as mentioned, the DEMs themselves exhibit spatially variable vertical accuracies dependent upon the type of terrain which may be responsible for much of the error seen in some areas. Other factors, such as DEM resolution can also influence accuracy as insufficient resolution can allow connections across barriers, such as dams and levees, that do not actually exist. Additionally, areas with many inter-connected lakes, like those in Minnesota, USA, often suffered from low accuracy as well. In these cases, the inter-connectivity may cause issues. At some points in time the channel may be large enough to be detectable in the Landsat imagery and that additional surface area will be included in the

analysis. If, in those cases, the water bodies are also connected in the DEM then the estimate will be accurate. However, if the channel is not captured in the DEM, or the other water body is at a different elevation, or the interconnection is not detectable in the Landsat imagery then the relationship can quickly become skewed and the estimate accuracy will diminish. This is evidenced by the Figure 4-5 time-series for Big Sandy Lake in Minnesota, USA. The elevation estimates for Big Sandy Lake are significantly lower than the in-situ elevation measurements due to wide-scale interconnectivity within the DEMs between water bodies in that region as illustrated in Figure 4-8. In the case of Big Sandy Lake, the elevations are routinely underestimated because the DEMs connect Big Sandy Lake to Aitken Lake (and others) while the water channel itself does not fully connect within the Landsat imagery.

4.7 Conclusions

Overall, despite low general accuracy in comparison to some previous papers which used DEMs to estimate water surface elevation using hypsometric relationships, the strategies developed in this chapter were successful on multiple fronts. The first and second goals of this research were: 1) to analyze and compare methods of extending hypsometric curves from the above hydroflattened surface to sub-surface levels and 2) investigate methods of reducing sub-hydroflattened water surface elevation estimation uncertainty. These goals were accomplished by analyzing the accuracy of 1st, 2nd, and 3rd-order polynomials and by using multiple DEMs to lower the hydroflattened water surface using the Lowest Hydroflattened Surface (LHFS) DEM at each location. For clear images (<1% image contamination), using the lowest hydroflattened surface DEM and 2nd-order polynomials had the highest accuracy for at 2.52 m RMSE CI95%[2.42, 2.61] and 1.74 m MAE CI95%[1.67, 1.8] and reduced the number of remaining subsurface estimates between 549 and 3,597 images versus using a single DEM. The LHFS DEM model also increased accuracy by 1.2 meters over the best single DEM model. Additional accuracy gains and decreases in low-level uncertainty should be expected as more DEMs are included in the analysis.

The third goal of this research was to assess the technique's performance across water bodies of various shape, sizes, water qualities, climate regimes, management practices, and ecotypes. This was accomplished using 46 lakes and reservoirs across the western and midwestern United States. While limited information was provided regarding individual lake accuracies, it was found that accuracy did vary significantly from location to location with deviations being driven by DEM quality issues and interconnected water bodies in some regions. As such, future research will investigate methods of correcting for DEM related errors using in-situ or satellite sensor data, like SWOT, to apply a local terrain bias correction. The fourth goal, 4) identify and mitigate a variety of image contamination sources to substantially improve the temporal resolution of hypsometry-based water surface dynamic monitoring was also successful. Typical hypsometric relationships are restricted to clear imagery, but proportional hypsometry utilizes the quality assessment band to remove image contamination from the underlying DEM for each Landsat image and then generates a unique hypsometric relationship for each image with an observable water shoreline (the shoreline is necessary to ensure the surface area and elevation intersect). Using proportional hypsometry, the temporal resolution of the analysis improved substantially. Unfortunately, proportional hypsometry suffers from increased uncertainty at sub-hydroflattened water surface elevations (images with surface levels remaining below the best combined DEM) and should be limited to use when water levels are above the hydroflattened surface to enable higher degree (3rd-degree) polynomial line fitting. Despite this limitation, proportional hypsometry of above hydroflattened water surface images (combined with images with <1% local image contamination for sub-hydroflattened water levels) enabled analysis of 23,341 out of 32,899 images with observable water boundaries and corresponding in-situ measurements. Furthermore, proportional hypsometry maintained comparable accuracy to clear imagery (<1% local image contamination) while providing 12,680 additional elevation estimates (23,341 vs 10,661) for the 46 lakes used in this study.

At the individual lake level, accuracy error was primarily driven by limitations in the DEMs. Spatially, the vertical accuracy of each DEM varies, especially in areas with substantial vertical relief.

Furthermore, lake inter-connectivity in some regions caused significant deviations as well. Areas of inter-connectivity would likely be better suited to techniques using water boundary extents, such as Weekley and Li (2019), as the elevation values would be pulled directly from the DEM. In theory, estimate bias could be reduced through the inclusion of supplementary elevation data such as a SWOT.

Finally, this research was developed using Google Earth Engine using readily available datasets. While future research will work to improve the overall accuracy of the analysis and remove detected biases, the method as-is is readily deployable and scalable for lakes across the globe. For future hypsometric based studies, we would recommend using multiple digital elevation models to reduce the number of sub-hydroflattened surface water levels. For the remaining sub-hydroflattened surface images we would recommend limiting the analysis to clear images (<1% local image contamination) and using a 2nd-degree polynomial to fit the relationship. Furthermore, we would recommend using proportional hypsometry for all above-hydroflattened surface water levels and fitting the relationship using a 3rd-degree polynomial. Future research will include analysis of additional DEMs such as the ASTER global DEM (version 3) and incorporation of remotely-sensed surface area/elevation pairs (such as altimetry in combination with optical imagery) to improve the analysis as well investigate means of incorporating elevation data to correct the biases resulting from DEM vertical accuracy error in mountainous and other high error regions.

4.8 Figures and Tables

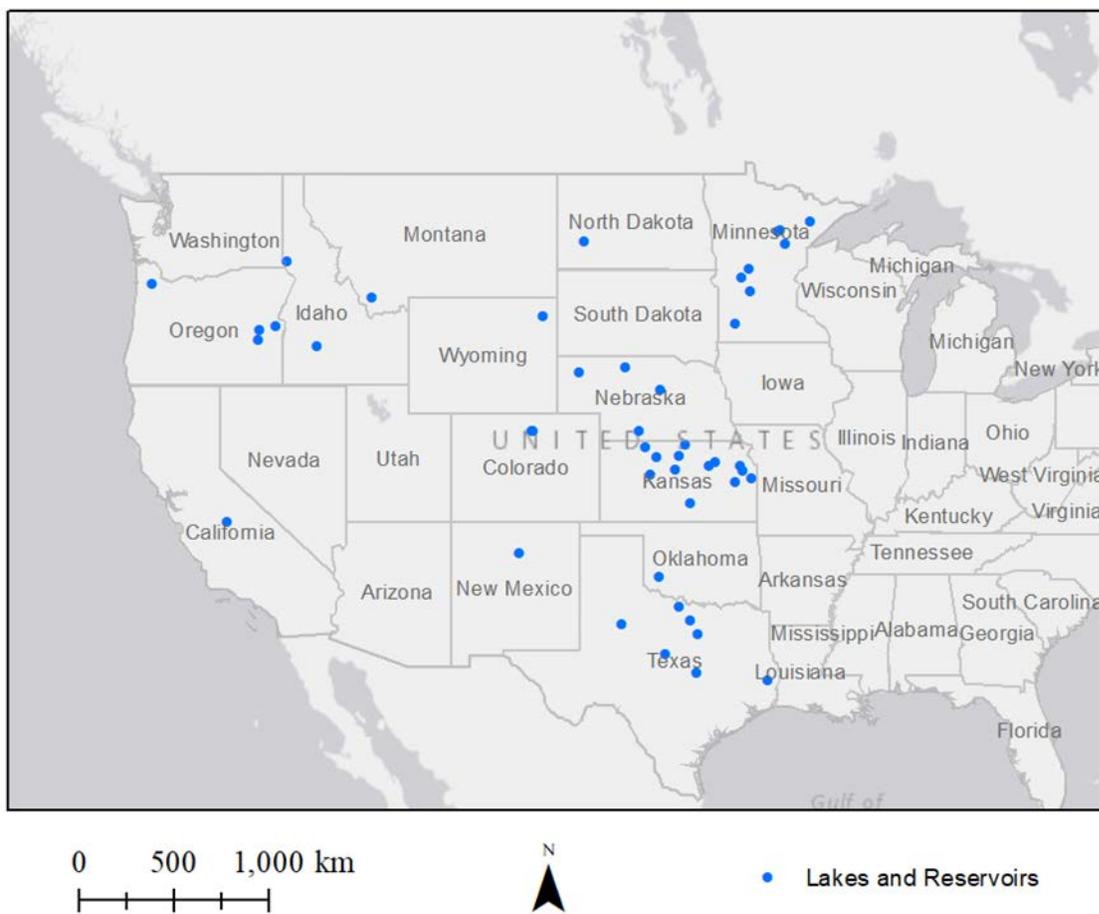


Figure 4-1: Forty-six lakes and reservoirs utilized within the study.

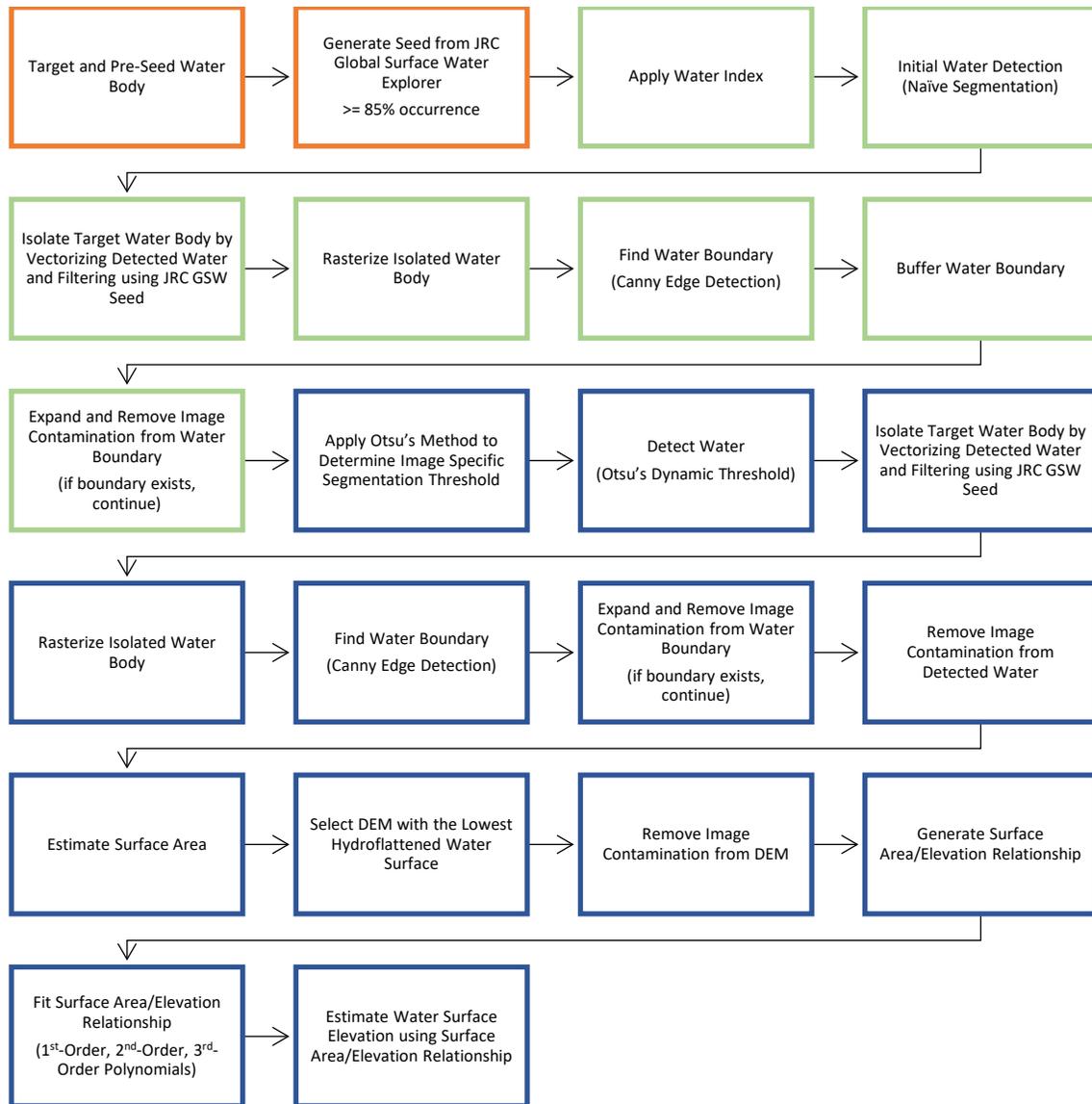


Figure 4-2: The general processing procedure used in this study to dynamically estimate water surface elevation using hypsometric relationships.

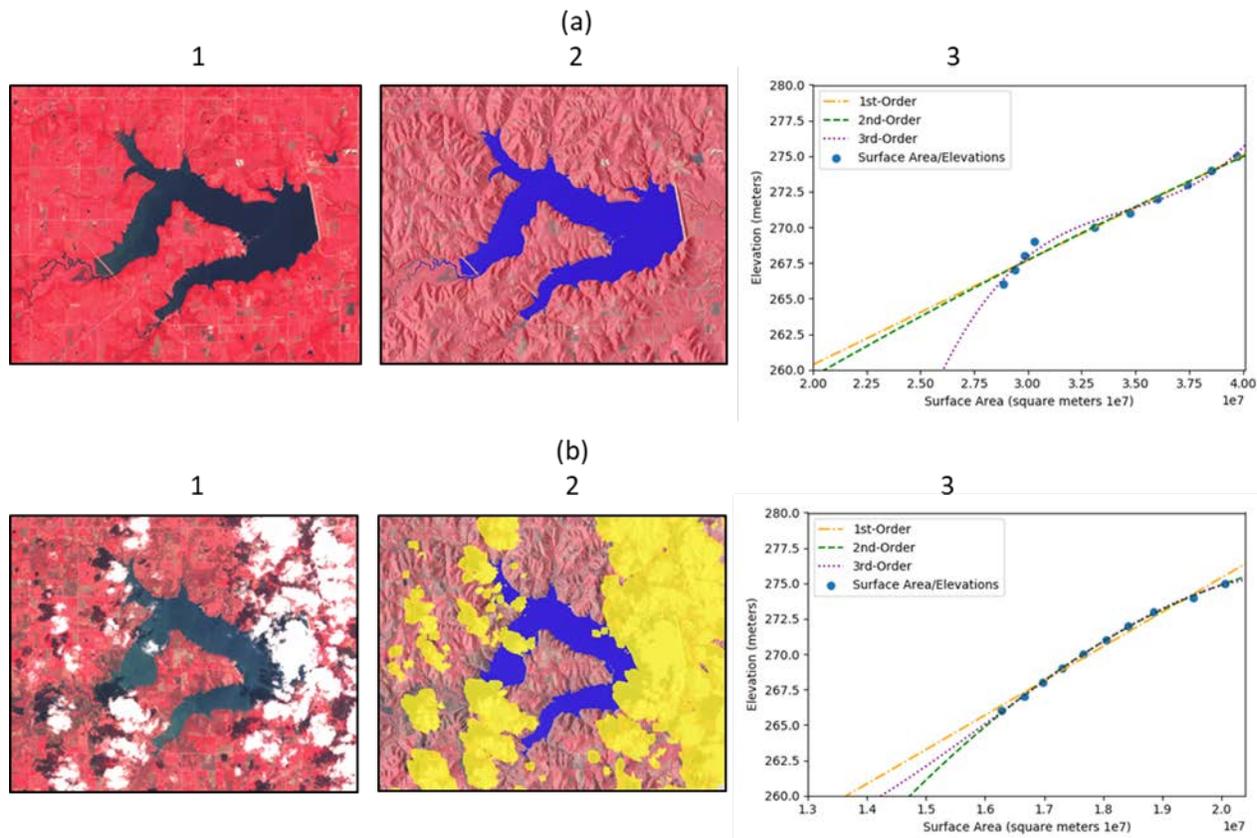


Figure 4-3: Clear and contaminated Landsat 5 image examples for Clinton Lake, Kansas, USA from (a) September 3, 1985 and (b) April 27, 2008. Each example shows the lake as it appears in (1) false-color infrared, (2) false-color infrared with transparency overlaying a hill shade with detected water in blue and image contamination in yellow, and (3) the resulting hypsometric relationship.

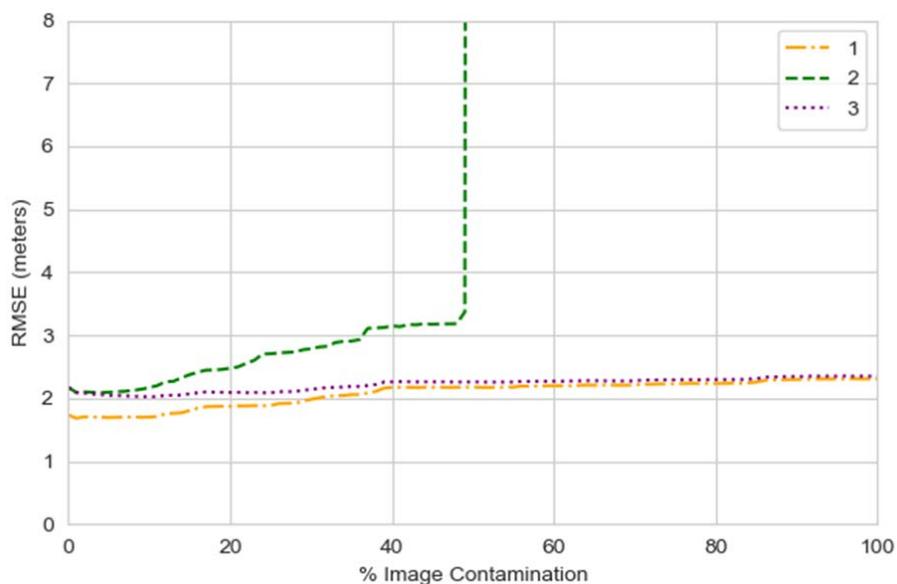
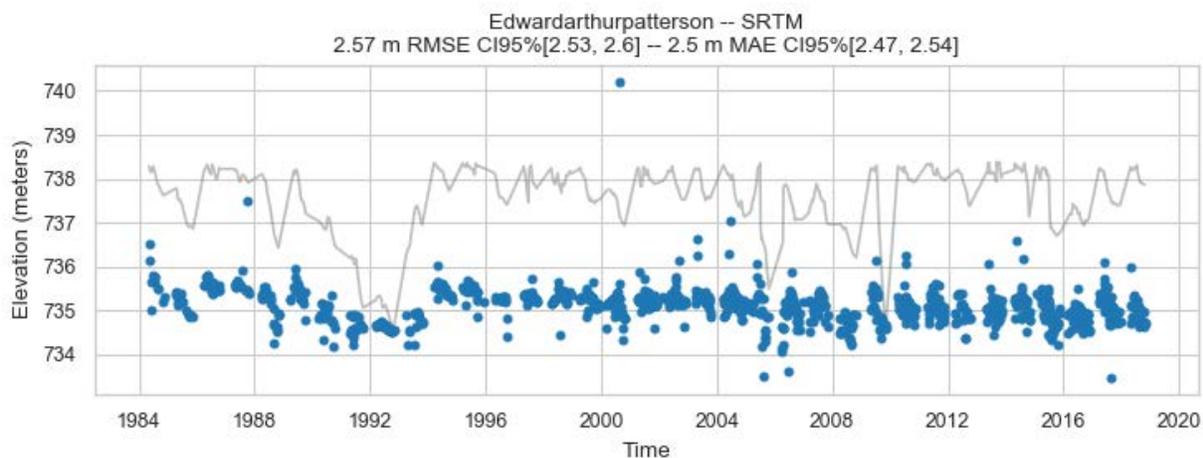
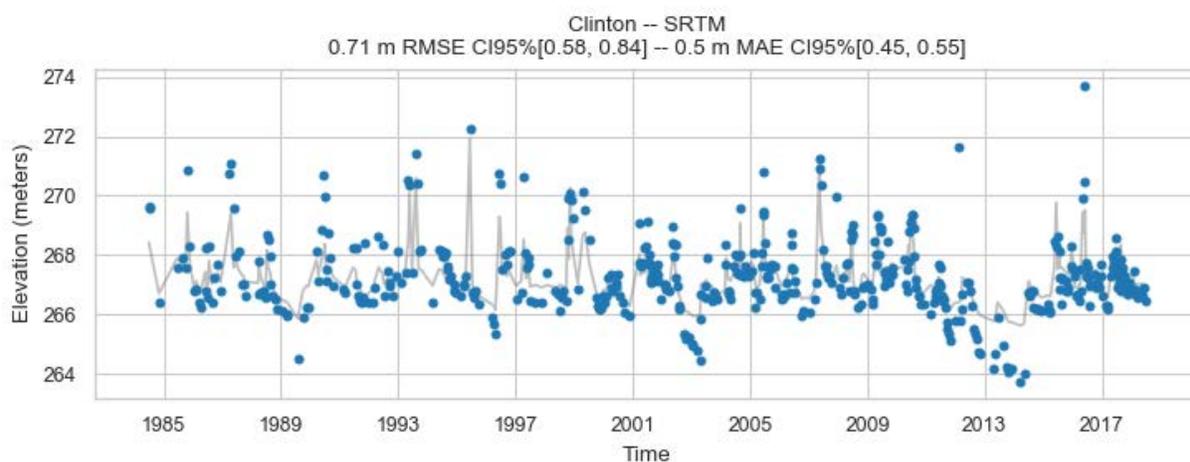
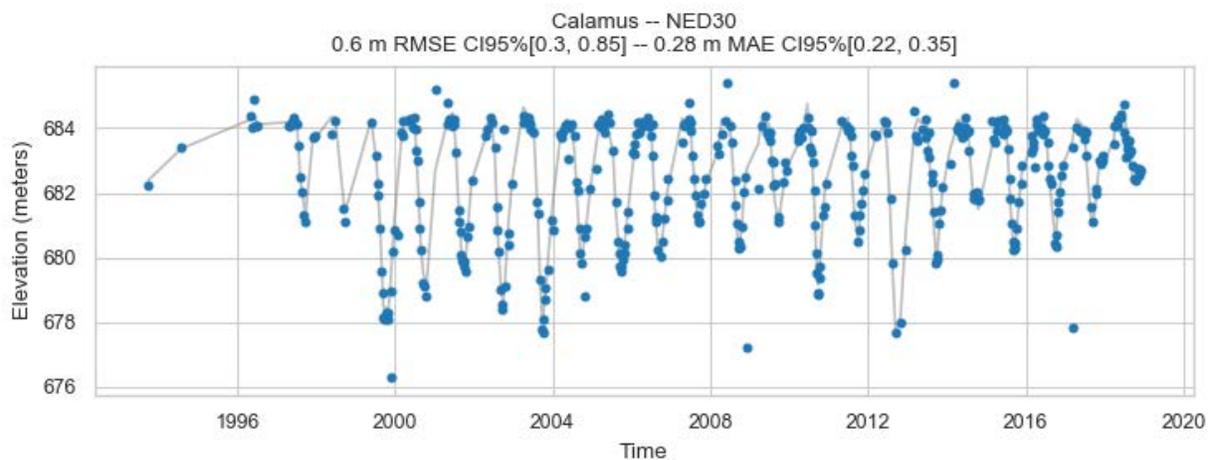
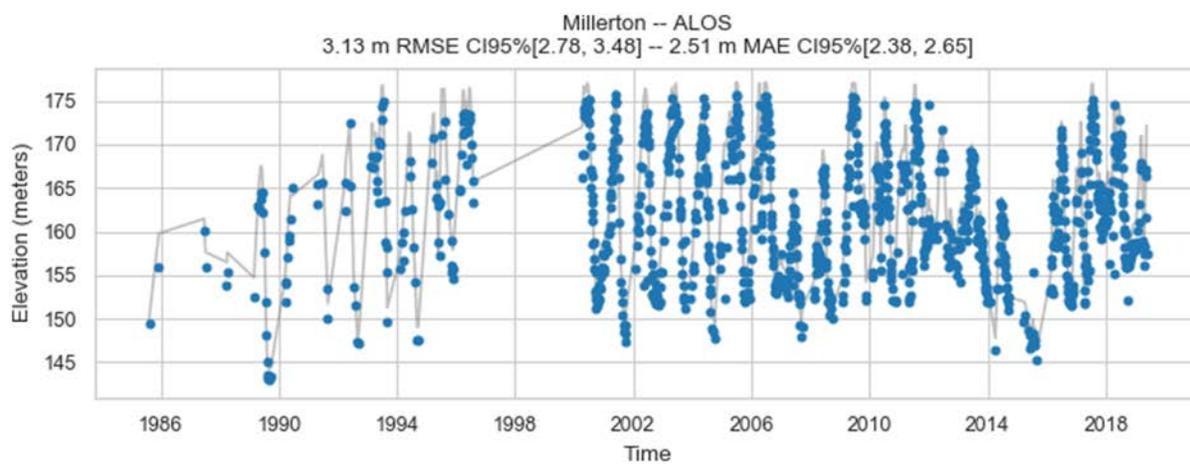
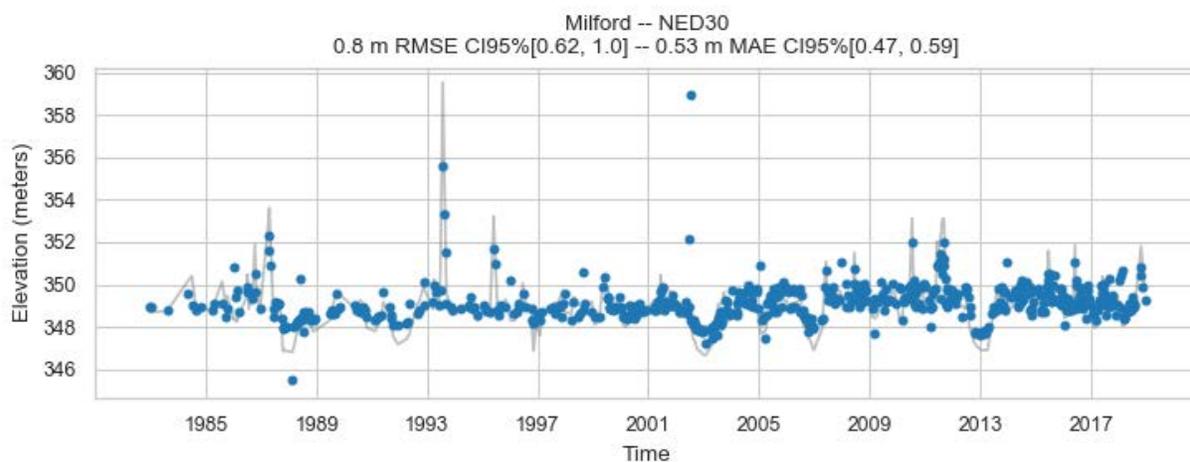
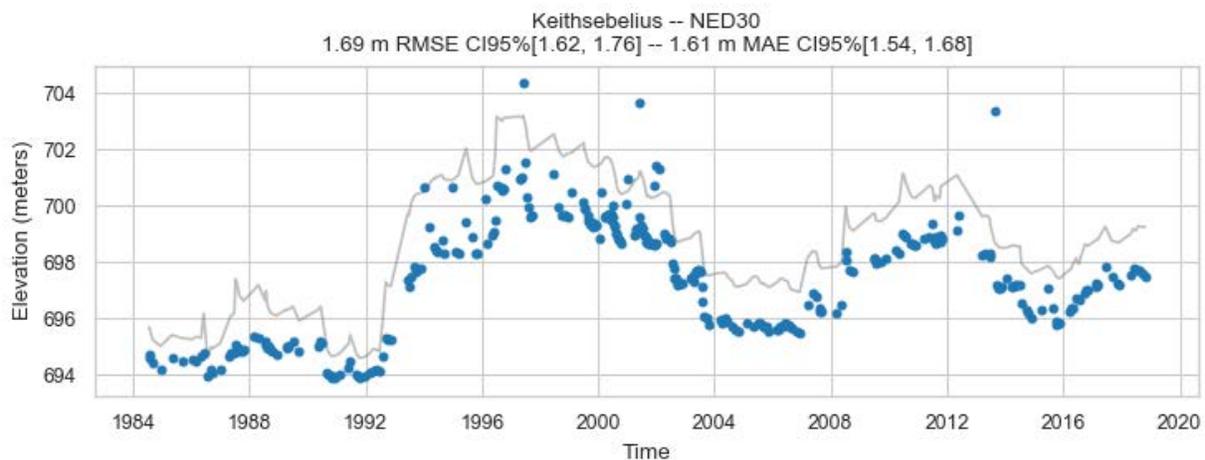


Figure 4-4: Accuracy comparison as the amount of image contamination increases for 1) all images above the hydroflattened water level, 2) the LHFS DEM model with all images (3rd-degree fit for above hydroflattened surface images and 2nd-degree fit

for sub-hydroflattened surface images), and 3) the LHFS DEM model with all above-hydroflattened surface images (3rd-degree fit) and images with $\leq 1\%$ local image contamination (2nd-degree fit).





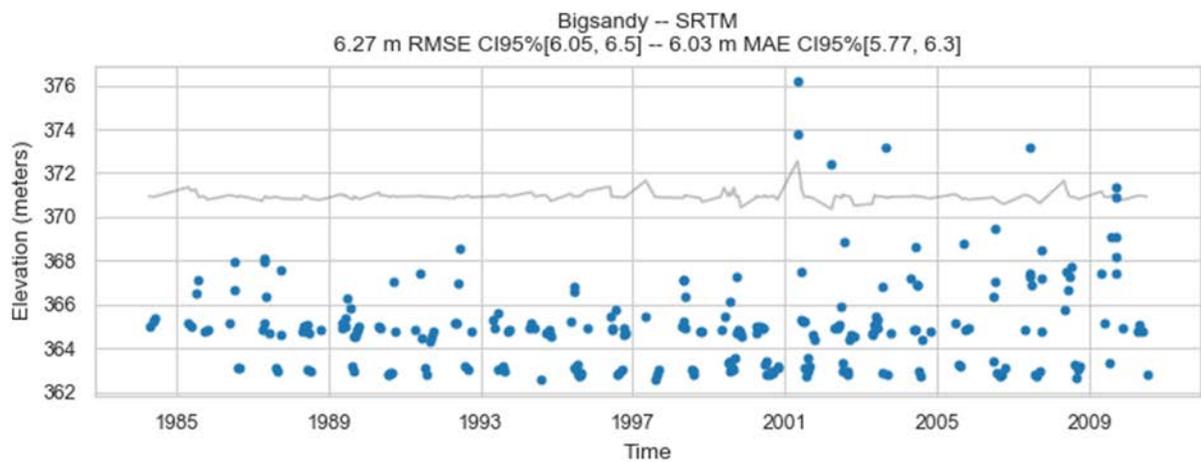
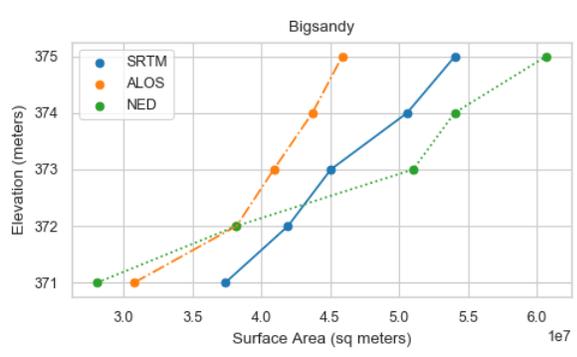
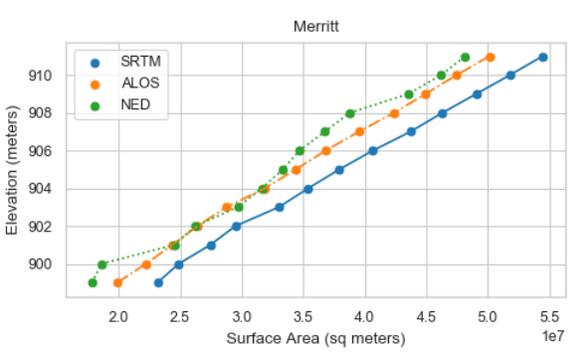
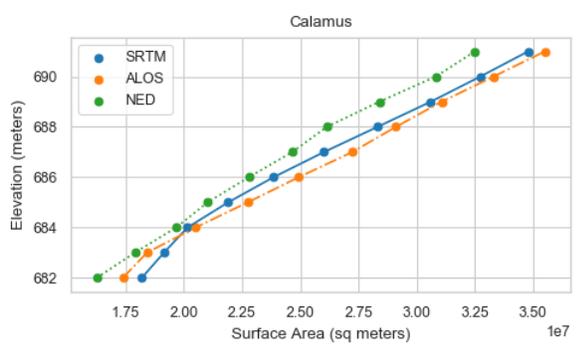
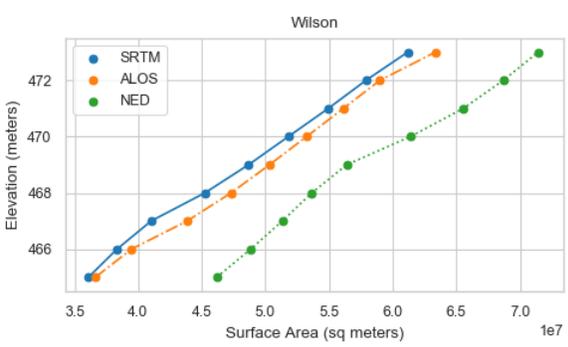


Figure 4-5: Time-series charts for selected lakes and reservoirs. Blue points represent the estimated water surface elevations while the gray line represents in-situ water surface elevation measurements.



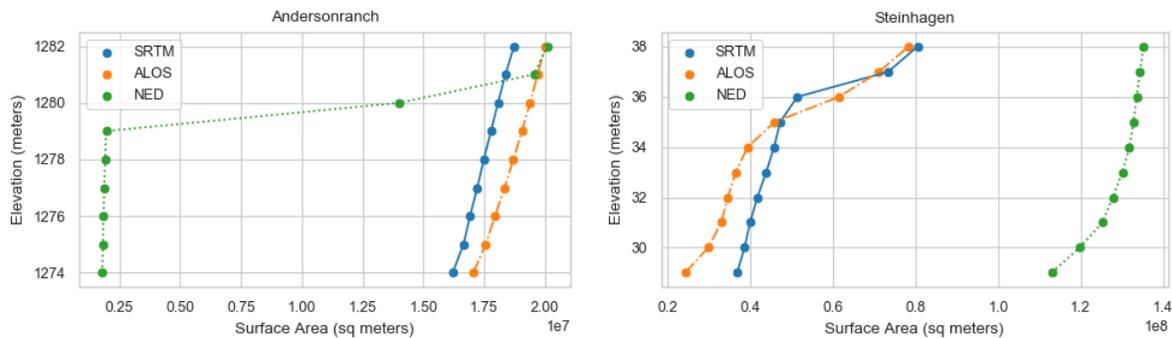


Figure 4-6: Common water levels and their accompanying surface areas for selected water bodies.

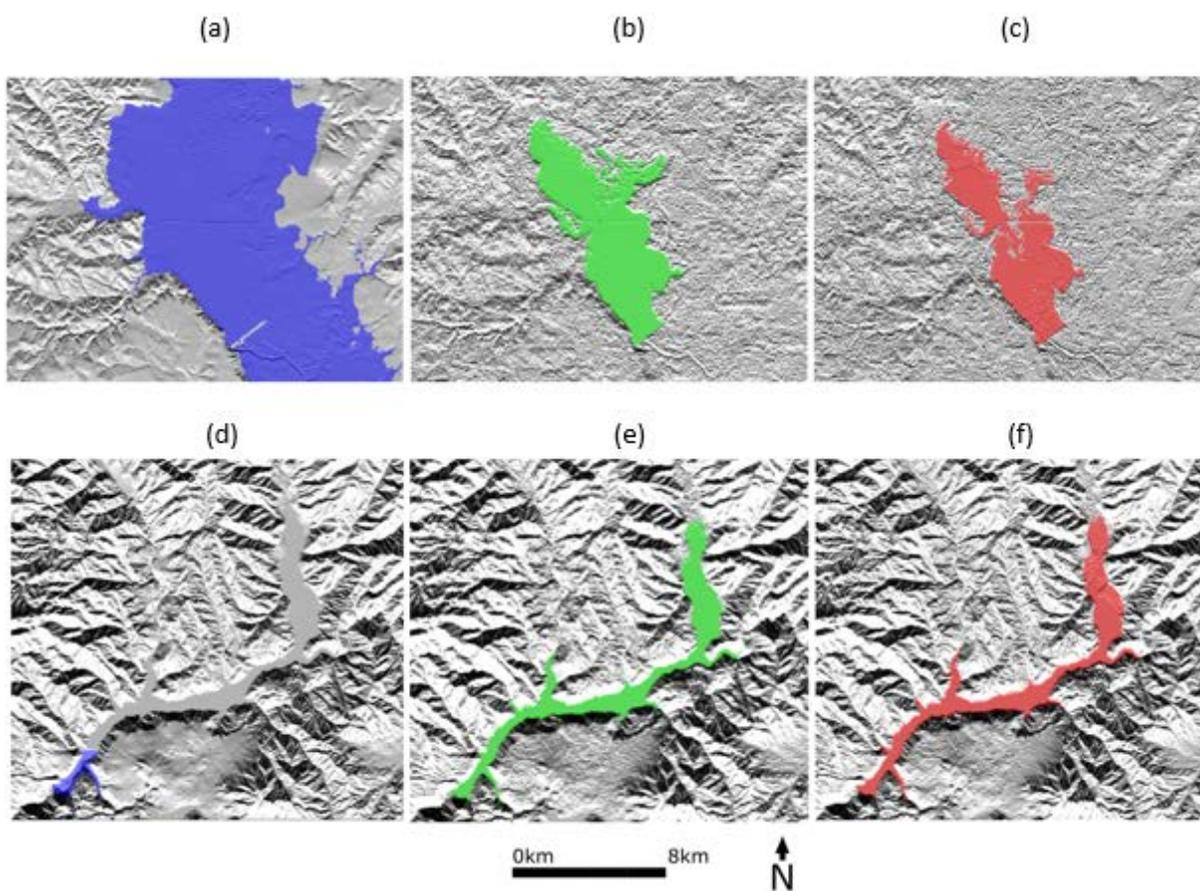


Figure 4-7: Comparison of the 30 meters above sea level inundation extent on the (a) NED, (b) SRTM, and (c) ALOS digital elevation models for B.A. Steinhagen Lake in Texas, U.S.A and 1,278 meters above sea level on (d) NED, (e) SRTM, and (f) ALOS digital elevation models for Anderson Ranch Reservoir in Idaho, USA.

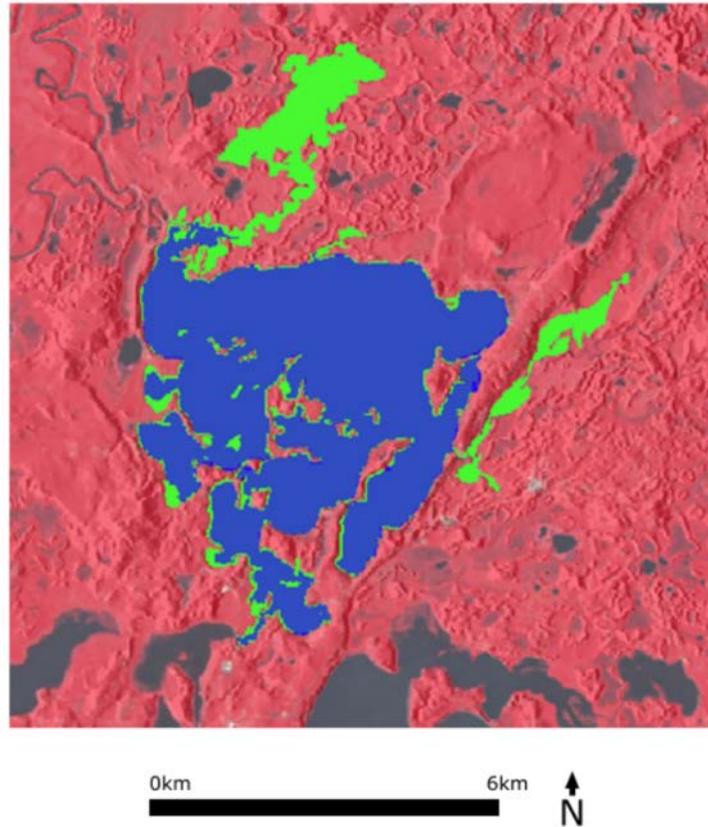


Figure 4-8: Large-scale interconnectivity of water bodies in the area of Big Sandy Lake, Minnesota, USA. False-color infrared Landsat 5 image from August 16, 1999 overlaid with a NED 30m hillshade to highlight the elevation differences. The blue area is the detected water for Big Sandy Lake and the green area is the connected area for the DEM at the August 16, 1999 in-situ elevation level. For this lake, the interconnected water body issue was present for all tested DEMs as well as the full-resolution NED.

Table 4-1: Sensor and mission details for each Landsat mission used in this study including available image date ranges, spatial resolution, temporal resolution, and radiometric resolution.

Mission	Landsat 4	Landsat 5	Landsat 7	Landsat 8
Sensor	Thematic Mapper	Thematic Mapper	Enhanced Thematic Mapper Plus	Operational Land Imager
GEE Date Range	08/22/1982 - 11/18/1993	03/16/1984 - 05/05/2012	05/28/1999 - Present*	03/18/2013 - Present
Band 1	Blue (30m) 0.45 - 0.52 μ m	Blue (30m) 0.45 - 0.52 μ m	Blue (30m) 0.45 - 0.52 μ m	Coastal (30m) 0.43 - 0.45 μ m
Band 2	Green (30m) 0.52 - 0.60 μ m	Green (30m) 0.52 - 0.60 μ m	Green (30m) 0.52 - 0.60 μ m	Blue (30m) 0.45 - 0.51 μ m
Band 3	Red (30m) 0.63 - 0.69 μ m	Red (30m) 0.63 - 0.69 μ m	Red (30m) 0.63 - 0.69 μ m	Green (30m) 0.53 - 0.59 μ m
Band 4	NIR ¹ (30m) 0.76 - 0.90 μ m	NIR ¹ (30m) 0.76 - 0.90 μ m	NIR ¹ (30m) 0.77 - 0.90 μ m	Red (30m) 0.64 - 0.67 μ m
Band 5	SWIR ² 1 (30m) 1.55 - 1.75 μ m	SWIR ² 1 (30m) 1.55 - 1.75 μ m	SWIR ² 1 (30m) 1.55 - 1.75 μ m	NIR ¹ (30m) 0.85 - 0.88 μ m
Band 6	TIR ³ (60m) 10.40 - 12.50 μ m	TIR ³ (60m) 10.40 - 12.50 μ m	TIR ³ (60m) 10.40 - 12.50 μ m	SWIR ² 1 (30m) 1.57 - 1.65 μ m
Band 7	SWIR ² 2 (30m) 2.08 - 2.35 μ m	SWIR ² 2 (30m) 2.08 - 2.35 μ m	SWIR ² 2 (30m) 2.08 - 2.35 μ m	SWIR ² 2 (30m) 2.11 - 2.29 μ m
Band 8	-	-	Panchromatic (15m) 0.52 - 0.90 μ m	Panchromatic (15m) 0.52 - 0.90 μ m
Band 9	-	-	-	Cirrus (15m) 1.36 - 1.38 μ m
Band 10	-	-	-	TIR ³ 1 (100m) 10.60 - 11.19 μ m
Band 11	-	-	-	TIR ³ 2 (100m) 11.50 - 12.51 μ m
QA Band	Collection 1 QA Bitmask 8 bit (256 levels)	Collection 1 QA Bitmask 8 bit (256 levels)	Collection 1 QA Bitmask 8 bit (256 levels)	Collection 1 QA Bitmask 16 bit (65,536 levels)**
Radiometric Resolution				
Revisit Period	16 days	16 days	16 days	16 days

¹Near Infrared, ²Shortwave Infrared, ³Thermal Infrared

*Operation despite Scan Line Corrector (SLC) failure May 31, 2003

**When processed into Level-1 data products.

Table 4-2: Number of Landsat images included in the analysis. <1% contamination represents images available to conventional hypsometric techniques. <100% contamination are the number images with at least one clear image pixel within 500 meters of the JRC GSW observed water extents for each lake (the <100% Contaminated column also includes images with <1% contamination).

Image Counts	< 1% Contamination	<100% Contaminated	Not Used (100% Contaminated)	Total Image Count
All Overlapping Images	14,170 (19.8%)	59,920 (83.7%)	11,678 (16.3%)	71,598 (100%)
With an Observable Shoreline Boundary	13,926 (19.5%)	41,459 (57.9%)	30,139 (42.1%)	
With an Observable Shoreline Boundary and In-Situ Elevation	10,661 (14.9%)	32,899 (45.9%)	38,699 (54.1%)	

Table 4-3: Comparison of hypsometric relationships generated from ALOS, NED, SRTM, and the LHFS DEM for all clear images (<1% contamination).

Rank	DEM	Fit Type	RMSE Train	MAE Train	RMSE Test	MAE Test
1	LHFS	2nd	2.17 m CI95%[2.14, 2.21]	1.53 m CI95%[1.51, 1.56]	2.17 m CI95%[2.12, 2.23]	1.53 m CI95%[1.5, 1.57]
2	LHFS	1st	2.29 m CI95%[2.26, 2.32]	1.74 m CI95%[1.71, 1.76]	2.29 m CI95%[2.24, 2.33]	1.74 m CI95%[1.7, 1.77]
3	ALOS	2nd	3.3 m CI95%[3.26, 3.33]	2.63 m CI95%[2.6, 2.66]	3.29 m CI95%[3.24, 3.35]	2.63 m CI95%[2.58, 2.68]
4	LHFS	3rd	3.82 m CI95%[3.59, 4.04]	1.92 m CI95%[1.86, 1.97]	3.82 m CI95%[3.48, 4.15]	1.91 m CI95%[1.84, 1.99]
6	SRTM	1st	3.76 m CI95%[3.67, 3.85]	2.71 m CI95%[2.67, 2.75]	3.76 m CI95%[3.63, 3.89]	2.71 m CI95%[2.65, 2.77]
6	ALOS	1st	3.39 m CI95%[3.35, 3.42]	2.77 m CI95%[2.74, 2.8]	3.39 m CI95%[3.34, 3.44]	2.77 m CI95%[2.73, 2.82]
6	ALOS	3rd	3.95 m CI95%[3.76, 4.13]	2.6 m CI95%[2.55, 2.64]	3.95 m CI95%[3.67, 4.23]	2.59 m CI95%[2.52, 2.66]
8	SRTM	2nd	4.27 m CI95%[4.12, 4.41]	2.78 m CI95%[2.73, 2.83]	4.27 m CI95%[4.04, 4.48]	2.78 m CI95%[2.71, 2.86]
9	NED30	1st	8.88 m CI95%[8.43, 9.31]	3.54 m CI95%[3.41, 3.66]	8.86 m CI95%[8.18, 9.5]	3.53 m CI95%[3.34, 3.72]
10	SRTM	3rd	14.27 m CI95%[12.27, 15.93]	4.71 m CI95%[4.49, 4.91]	14.12 m CI95%[11.29, 16.8]	4.69 m CI95%[4.39, 5.02]
11	NED30	2nd	27.29 m CI95%[24.61, 29.6]	5.66 m CI95%[5.23, 6.05]	27.24 m CI95%[23.46, 30.92]	5.65 m CI95%[5.06, 6.29]
12	NED30	3rd	188.99 m CI95%[169.22, 206.79]	27.99 m CI95%[25.12, 30.87]	188.4 m CI95%[159.0, 215.55]	27.93 m CI95%[23.62, 32.24]

Table 4-4: Accuracy assessment for all 32,899 images with observable shorelines and in-situ surface elevation measurements.

Rank	DEM	Fit Type	RMSE Train	MAE Train	RMSE Test	MAE Test
1	LHFS	1st	6.91 m CI95%[4.35, 8.49]	2.16 m CI95%[2.11, 2.22]	6.75 m CI95%[3.87, 9.73]	2.16 m CI95%[2.08, 2.25]
2	ALOS	1st	5.39 m CI95%[4.52, 6.02]	3.05 m CI95%[3.01, 3.09]	5.34 m CI95%[4.32, 6.5]	3.05 m CI95%[2.99, 3.11]
3	SRTM	1st	35.55 m CI95%[7.36, 49.1]	3.5 m CI95%[3.16, 3.77]	33.9 m CI95%[6.25, 59.79]	3.51 m CI95%[3.11, 4.02]
4	ALOS	2nd	113.96 m CI95%[32.22, 157.33]	4.3 m CI95%[3.27, 5.14]	101.92 m CI95%[12.41, 189.01]	4.25 m CI95%[3.0, 5.8]
5	NED30	1st	44.19 m CI95%[13.84, 58.46]	5.04 m CI95%[4.64, 5.35]	39.65 m CI95%[13.09, 70.78]	5.02 m CI95%[4.56, 5.62]
6	LHFS	2nd	275.15 m CI95%[25.29, 376.45]	5.1 m CI95%[2.41, 6.98]	245.8 m CI95%[7.48, 460.08]	5.03 m CI95%[2.21, 9.06]
7	ALOS	3rd	7991.95 m CI95%[968.07, 11461.68]	85.06 m CI95%[14.61, 134.4]	6849.26 m CI95%[406.36, 13993.02]	81.44 m CI95%[7.43, 187.11]
8	LHFS	3rd	22974.55 m CI95%[338.78, 34431.52]	200.43 m CI95%[8.38, 330.07]	19766.45 m CI95%[132.19, 42167.99]	199.87 m CI95%[5.4, 487.94]
9	SRTM	2nd	433636.35 m CI95%[81.22, 711846.09]	3103.4 m CI95%[6.93, 5091.96]	341048.81 m CI95%[60.99, 871796.55]	2988.91 m CI95%[6.3, 7633.27]
10	SRTM	3rd	511043.61 m CI95%[32768.06, 729089.4]	4864.3 m CI95%[292.98, 8003.9]	444612.89 m CI95%[2568.3, 892015.81]	4785.93 m CI95%[76.88, 11642.39]
11	NED30	3rd	1590315476.64 m CI95%[38149.21, 2654868924.93]	11343470.52 m CI95%[530.34, 18937166.06]	1303881828.05 m CI95%[18670.82, 3251413059.19]	11389901.51 m CI95%[227.24, 28403013.67]
12	NED30	2nd	3040102322.72 m CI95%[319.4, 5075295596.25]	21682424.22 m CI95%[13.9, 36197695.24]	2492498117.29 m CI95%[106.22, 6215705124.79]	21771257.04 m CI95%[11.67, 54292391.12]

Table 4-5: Accuracy comparison for sub-hydroflattened surface images with <1% local image contamination.

Rank	DEM	Fit Type	Image Count	RMSE Train	MAE Train	RMSE Test	MAE Test
1	LHFS	2 nd	4690	2.52 m CI95%[2.45, 2.58]	1.74 m CI95%[1.7, 1.78]	2.52 m CI95%[2.42, 2.61]	1.74 m CI95%[1.67, 1.8]
2	LHFS	1 st	4690	2.64 m CI95%[2.58, 2.69]	1.97 m CI95%[1.93, 2.0]	2.64 m CI95%[2.56, 2.72]	1.96 m CI95%[1.91, 2.02]
3	ALOS	2 nd	5239	3.74 m CI95%[3.68, 3.79]	3.06 m CI95%[3.02, 3.11]	3.74 m CI95%[3.66, 3.82]	3.06 m CI95%[3.0, 3.13]
4	LHFS	3 rd	4690	5.44 m CI95%[5.07, 5.77]	2.68 m CI95%[2.56, 2.78]	5.45 m CI95%[4.92, 5.97]	2.68 m CI95%[2.52, 2.85]
5	SRTM	1 st	5701	4.42 m CI95%[4.29, 4.56]	3.13 m CI95%[3.06, 3.19]	4.41 m CI95%[4.2, 4.6]	3.12 m CI95%[3.03, 3.22]
6	ALOS	1 st	5239	3.8 m CI95%[3.75, 3.85]	3.15 m CI95%[3.1, 3.19]	3.8 m CI95%[3.73, 3.88]	3.15 m CI95%[3.08, 3.22]
7	ALOS	3 rd	5239	4.99 m CI95%[4.71, 5.25]	3.1 m CI95%[3.01, 3.18]	4.97 m CI95%[4.55, 5.37]	3.09 m CI95%[2.96, 3.21]
8	SRTM	2 nd	5701	5.32 m CI95%[5.12, 5.53]	3.46 m CI95%[3.38, 3.55]	5.31 m CI95%[4.99, 5.61]	3.46 m CI95%[3.33, 3.59]
9	NED	1 st	8287	9.9 m CI95%[9.36, 10.37]	4.09 m CI95%[3.94, 4.24]	9.91 m CI95%[9.18, 10.68]	4.1 m CI95%[3.88, 4.33]
10	SRTM	3 rd	5701	19.3 m CI95%[16.38, 21.54]	7.11 m CI95%[6.72, 7.47]	19.27 m CI95%[15.51, 23.11]	7.12 m CI95%[6.59, 7.71]
11	NED	2 nd	8287	30.88 m CI95%[27.8, 33.44]	6.85 m CI95%[6.31, 7.38]	30.86 m CI95%[26.72, 35.1]	6.88 m CI95%[6.08, 7.69]
12	NED	3 rd	8287	213.65 m CI95%[190.93, 232.8]	35.37 m CI95%[31.56, 38.99]	214.7 m CI95%[183.63, 245.61]	35.6 m CI95%[30.16, 41.32]

Table 4-6: Accuracy comparison for all images above hydroflattened water surface levels. An image is above hydroflattened if the estimated water surface area is larger than the surface area of the DEM's hydroflattened water surface.

Rank	DEM	Fit Type	Image Count	RMSE Train	MAE Train	RMSE Test	MAE Test
1	LHFS	3rd	18651	2.31 CI95%[2.24, 2.38]	1.58 CI95%[1.57, 1.6]	2.31 CI95%[2.21, 2.41]	1.58 CI95%[1.55, 1.61]
2	LHFS	1st	18651	2.32 CI95%[2.29, 2.35]	1.74 CI95%[1.73, 1.76]	2.32 CI95%[2.27, 2.37]	1.74 CI95%[1.72, 1.77]
3	LHFS	2nd	18651	2.39 CI95%[2.25, 2.5]	1.62 CI95%[1.6, 1.64]	2.4 CI95%[2.23, 2.6]	1.62 CI95%[1.59, 1.65]
4	SRTM	2nd	14381	2.82 CI95%[2.67, 2.94]	2.04 CI95%[2.02, 2.07]	2.82 CI95%[2.64, 3.05]	2.04 CI95%[2.0, 2.08]
5	SRTM	1st	14381	2.85 CI95%[2.81, 2.88]	2.19 CI95%[2.17, 2.22]	2.84 CI95%[2.79, 2.9]	2.19 CI95%[2.15, 2.23]
6	SRTM	3rd	14381	4.15 CI95%[3.09, 4.83]	2.1 CI95%[2.05, 2.14]	4.04 CI95%[2.87, 5.38]	2.09 CI95%[2.03, 2.17]
7	ALOS	3rd	17297	2.87 CI95%[2.82, 2.92]	2.23 CI95%[2.21, 2.25]	2.87 CI95%[2.8, 2.95]	2.23 CI95%[2.2, 2.27]
8	ALOS	2nd	17297	3.01 CI95%[2.98, 3.05]	2.33 CI95%[2.31, 2.35]	3.01 CI95%[2.96, 3.07]	2.33 CI95%[2.3, 2.37]
9	NED30	2nd	7318	6.07 CI95%[4.53, 7.04]	2.18 CI95%[2.08, 2.28]	5.97 CI95%[4.33, 7.89]	2.18 CI95%[2.04, 2.34]
10	NED30	1st	7318	4.49 CI95%[4.28, 4.69]	2.22 CI95%[2.15, 2.29]	4.48 CI95%[4.16, 4.79]	2.22 CI95%[2.12, 2.33]
11	ALOS	1st	17297	3.14 CI95%[3.1, 3.17]	2.5 CI95%[2.48, 2.52]	3.14 CI95%[3.09, 3.18]	2.5 CI95%[2.47, 2.53]
12	NED30	3rd	7318	23.01 CI95%[7.04, 30.89]	2.82 CI95%[2.38, 3.16]	21.17 CI95%[6.57, 37.41]	2.81 CI95%[2.3, 3.48]

Table 4-7: Accuracy comparison for all DEM models using all above hydroflattened surface images fit with a 3rd-degree polynomial and sub-hydroflattened surface images with <1% local image contamination fit with a 2nd-degree polynomial.

Rank	DEM	Image Count	RMSE Train	MAE Train	RMSE Test	MAE Test
1	LHFS	23341	2.35 m CI95%[2.3, 2.41]	1.62 m CI95%[1.6, 1.63]	2.35 m CI95%[2.27, 2.43]	1.61 m CI95%[1.59, 1.64]
2	ALOS	22536	3.1 m CI95%[3.05, 3.14]	2.42 m CI95%[2.4, 2.45]	3.1 m CI95%[3.03, 3.16]	2.42 m CI95%[2.39, 2.46]
3	SRTM	20082	4.51 m CI95%[3.85, 4.98]	2.48 m CI95%[2.44, 2.52]	4.48 m CI95%[3.72, 5.37]	2.48 m CI95%[2.42, 2.55]
4	NED30	15605	27.92 m CI95%[22.19, 31.94]	4.97 m CI95%[4.64, 5.32]	27.58 m CI95%[20.76, 34.97]	4.95 m CI95%[4.43, 5.46]

5 Conclusions

This dissertation consists of three components which share a common, overarching theme related to the exploration and development of scalable long-term surface water dynamic monitoring using available optical imagery, digital elevation models, and bathymetric datasets.

Chapter 2, *Tracking Multi-Decadal Lake Water Dynamics with Landsat Imagery and Topography/Bathymetry* (Weekley & Li, 2019) provides a methodology for tracking surface water dynamics using Lake McConaughy in Nebraska, USA as an example to demonstrate the techniques. The major findings and contributions can be summarized as follows:

1. This research introduces a scalable framework deployable within Google Earth Engine (Gorelick et al., 2016) and enables rapid assessment of gauged and un-gauged water bodies increasing our knowledge and understanding of critical water systems.
2. Surface water boundaries are a viable method of estimating surface water elevation from an underlying merged digital elevation model and bathymetric survey dataset, especially on small to medium-sized water bodies not currently observable using other common surface water monitoring sensors such as satellite altimeters.
3. Surface water elevation estimate accuracy varies depending upon the water indice, segmentation threshold, and water boundary type, and statistic used. In this study, the highest accuracy parameter combination produced water surface elevation estimates as accurate as 0.768 meters RMSE CI95% [0.657, 0.885] using NDWI + MNDWI (B1 & B4) with a combination boundary, mean statistic, and 0.06 segmentation threshold. In total, 5,994 different parameter model combinations were evaluated.
4. The method identified seasonal and long-term patterns in Lake McConaughy's surface elevation, surface area, and volume. These trends have significant impact on regional agriculture, recreational activities, fish habitat, and overall water quality.

Chapter 3, *Maximizing Multi-Decadal Water Surface Elevation Estimates with Landsat Imagery and Elevation/Bathymetry Datasets* builds and expands upon the work in Chapter 2. In Chapter 2, the analysis was limited by a single study area as well as the use of image composites which limited the temporal resolution. Chapter 3 improves upon the techniques developed in Chapter 2 with the following major contributions and findings:

1. The analysis uses the entire Landsat 4, 5, 7, and 8 Top-of-Atmosphere Tier-1 Collection-1 image archives between August 1982 and December 2017 providing 35 years of surface water dynamic data.
2. Temporal resolution was maximized using the Pixel Quality Assurance band to eliminate image contamination such as cloud, cloud shadow, ice, snow, and sensor error. Nearly 43,000 water surface elevations were produced including at least 8,000+ additional elevation estimates that are not currently available from any other data source including in-situ hydrological records.
3. Fifty-two lakes and reservoirs across the Western and Midwestern United States were analyzed to determine how the analysis could be applied to water bodies with varying shapes, sizes, geologies, climate regimes, and management practices.
4. The analysis integrates bathymetric survey data of various resolutions and ages with the National Elevation Dataset. In general, model accuracy appears to be strongly linked to mean slope near the lake shoreline due to the spatial resolution of the digital elevation models, bathymetry survey data, and most importantly the Landsat 30-meter image pixel.
5. Elevation accuracy uncertainty increased for images with fewer observable shoreline pixels (increased image contamination). Filtering the results based upon shoreline length and/or Mean-Mode differencing could significantly improve model accuracy.
6. The accuracy of water surface elevation estimates varies from water body to water body. While a general overall model was identified, this model is optimized for the 52 lakes used explicitly in this analysis. Given the availability of some ground-truth data, such as in-situ ground station data, or the elevations which will be available from the upcoming SWOT (Surface Water and

Ocean Topography) mission (Biancamaria et al., 2016), locally optimized models can be developed for each individual water body. Using this technique, in addition to filtering of shoreline lengths and Mean-Mode difference, 45 lakes produced sub-meter RMSE elevation accuracies and 49 produced sub-meter MAE accuracies.

7. 594 parameter combinations were tested. Local Otsu dynamic thresholding was the best performing segmentation threshold strategy for each of the general models as well as most of the lake-specific locally optimized models.

Chapter 4, *Tracking Lake Surface Elevations with Proportional Hypsometric Relationships, Landsat Imagery, and Multiple DEMs* is a further step in extending lake dynamic research to regional and global levels. While the analysis here itself focuses upon 46 water bodies, it introduces several novel techniques for providing large-scale lake dynamic analysis. The following are key findings and major contributions for Chapter 4:

1. Hypsometric relationships can be used to estimate sub-hydroflattened water surface elevations albeit at a lower temporal resolution, slightly lower accuracy, and increased uncertainty versus above hydroflattened water levels. However, these estimates enable lake dynamic tracking, including surface area and water volume change, for water bodies lacking bathymetric survey data.
2. Multiple Digital Elevation Models can be used to reduce sub-hydroflattened water surface elevation estimate uncertainty by using the Lowest Hydroflattened Surface model (the DEM with the lowest hydroflattened surface and therefore a shorter extrapolation distance to low water levels and fewer sub-hydroflattened water surface levels overall).
3. Proportional hypsometry, which enables analysis of contaminated imagery by removing it from both the image and underlying DEM, can be used to estimate water surface elevation for above-hydroflattened water surface levels with just a small decrease in estimate accuracy in comparison to traditional hypsometric techniques. Proportional hypsometry produced 12,680

additional water surface elevation estimates in comparison to traditional hypsometry techniques which require clear image conditions.

4. While the overall accuracy is significantly less than that produced using bathymetric data in Chapters 2 and 3, the techniques developed in this chapter provide critical groundwork for feasible path to multi-decadal lake dynamic research at the global scale.

Overall, the research conducted for this dissertation produced multiple lake dynamic monitoring strategies. For lakes and water bodies with bathymetry, the methods developed in Chapters 2 and 3 provide high to moderate accuracy. For lakes without bathymetry, the techniques developed in Chapter 4 have shown promise for large-scale analysis. While the techniques currently lack the accuracy available in Chapters 2 and 3, future research could reduce this uncertainty and improve estimate accuracy.

Future research opportunities for long-term lake dynamic monitoring may include the following topics and applications:

1. Development of web application interfaces for estimating long-term lake dynamics from user selected water bodies and user provided data.
2. Expansion of analysis to full regional and global scales.
3. Collect, compare, and evaluate additional available digital elevation models, especially in relation to the LHFS model used in Chapter 4.
4. Assessment of higher resolution image datasets such as Sentinel as higher resolution imagery may increase estimate accuracy.
5. Investigate use of SWOT data as a bias adjustment for LHFS and proportional hypsometry techniques.

6 References

- Abileah, R., & Vignudelli, S. (2011). Bathymetry from fusion of multi-temporal Landsat and radar altimetry. *2011 6th International Workshop on the Analysis of Multi-Temporal Remote Sensing Images, Multi-Temp 2011 - Proceedings*, 189–192. <https://doi.org/10.1109/Multi-Temp.2011.6005080>
- Acharya, T. D., Subedi, A., Yang, I. T., & Lee, D. H. (2017). Combining Water Indices for Water and Background Threshold in Landsat Image. *Proceedings*, 2(3), 143. <https://doi.org/10.3390/ecsa-4-04902>
- Alganci, U., Besol, B., & Sertel, E. (2018). Accuracy assessment of different digital surface models. *ISPRS International Journal of Geo-Information*, 7(3), 1–16. <https://doi.org/10.3390/ijgi7030114>
- Alsdorf, D. E., Rodríguez, E., & Lettenmaier, D. P. (2007). Measuring surface water from space. *Reviews of Geophysics*, 45(2). <https://doi.org/10.1029/2006rg000197>
- Archuleta, C.-A. M., Constance, E. W., Arundel, S. T., Lowe, A. J., Mantey, K. S., & Phillips, L. A. (2017). The National Map Seamless Digital Elevation Model Specifications: U.S. Geological Survey Techniques and Methods, book 11, chap. B9, 39. Retrieved from <https://pubs.usgs.gov/tm/11b9/tm11B9.pdf>
- Asadzadeh Jarihani, A., Callow, J. N., Johansen, K., & Gouweleeuw, B. (2013a). Evaluation of multiple satellite altimetry data for studying inland water bodies and river floods. *Journal of Hydrology*, 505, 78–90. <https://doi.org/10.1016/j.jhydrol.2013.09.010>
- Asadzadeh Jarihani, A., Callow, J. N., Johansen, K., & Gouweleeuw, B. (2013b). Evaluation of multiple satellite altimetry data for studying inland water bodies and river floods. *Journal of Hydrology*, 505, 78–90. <https://doi.org/10.1016/j.jhydrol.2013.09.010>
- Avisse, N., Tilmant, A., François Müller, M., & Zhang, H. (2017). Monitoring small reservoirs' storage with satellite remote sensing in inaccessible areas. *Hydrol. Earth Syst. Sci*, 21, 6445–6459. <https://doi.org/10.5194/hess-21-6445-2017>
- Baig, M. H. A., Zhang, L., Shuai, T., & Tong, Q. (2014). Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sensing Letters*, 5(5), 423–431. <https://doi.org/10.1080/2150704X.2014.915434>
- Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016, March 27). The SWOT Mission and Its Capabilities for Land Hydrology. *Surveys in Geophysics*. Springer Netherlands. <https://doi.org/10.1007/s10712-015-9346-y>
- Busker, T., De Roo, A., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., ... Cottam, A. (2019). A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry. *Hydrology and Earth System Sciences*, 23(2), 669–690. <https://doi.org/10.5194/hess-23-669-2019>
- Cai, X., Feng, L., Hou, X., & Chen, X. (2016). Remote Sensing of the Water Storage Dynamics of Large Lakes and Reservoirs in the Yangtze River Basin from 2000 to 2014. *Scientific Reports*, 6(June), 1–9. <https://doi.org/10.1038/srep36405>
- California, B. of R. (2017). Water Facts - Worldwide Water | Bureau of Reclamation. Retrieved March 7, 2020, from <https://www.usbr.gov/mp/arwec/water-facts-ww-water-sup.html>
- California Data Exchange Center. (n.d.). Retrieved October 15, 2019, from

- <http://cdec.water.ca.gov/dynamicapp/wsSensorData>
- CNPPID. (n.d.). Nebraska Lake Elevations. Retrieved August 6, 2019, from <https://www.cnppid.com/news-info/reservoirriver-data/>
- Commission, N. G. and P. (n.d.). *Lake McConaughy Lake Ogallala Master Plan*. Retrieved from https://outdoornebraska.gov/wp-content/uploads/2017/03/Lake-McConaughy-Master-Plan_final.pdf
- Crétaux, J.-F., Biancamaria, S., Arsen, A., Bergé-Nguyen, M., & Becker, M. (2015). Global surveys of reservoirs and lakes from satellites and regional application to the Syrdarya river basin. *Environmental Research Letters*, *10*(1), 15002. <https://doi.org/10.1088/1748-9326/10/1/015002>
- Crétaux, J F, Abarca-del-Río, R., Bergé-Nguyen, M., Arsen, A., Drolon, V., Clos, G., & Maisongrande, P. (2016). Lake Volume Monitoring from Space. *Surveys in Geophysics*, *37*(2), 269–305. <https://doi.org/10.1007/s10712-016-9362-6>
- Crétaux, Jean François, & Birkett, C. (2006). Lake studies from satellite radar altimetry. *Comptes Rendus - Geoscience*, *338*(14–15), 1098–1112. <https://doi.org/10.1016/j.crte.2006.08.002>
- Crist, E. P. (1985). A TM Tasseled Cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, *17*(3), 301–306. [https://doi.org/10.1016/0034-4257\(85\)90102-6](https://doi.org/10.1016/0034-4257(85)90102-6)
- Crist, E. P., & Cicone, R. C. (1984). A Physically-Based Transformation of Thematic Mapper Data—The TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-22*(3), 256–263. <https://doi.org/10.1109/TGRS.1984.350619>
- deNoyelles, F., & Kastens, J. H. (2016). Reservoir Sedimentation Challenges Kansas. *Transactions of the Kansas Academy of Science*, *119*(1), 69–81. <https://doi.org/10.1660/062.119.0110>
- Denoyelles, F., & Kastens, J. H. (2016). Reservoir Sedimentation Challenges Kansas. *Source: Transactions of the Kansas Academy of Science*, *119*(1), 69–81. <https://doi.org/10.1660/062.119.0110>
- Dettmering, D., Schwatke, C., Boergens, E., & Seitz, F. (2016). Potential of ENVISAT radar altimetry for water level monitoring in the Pantanal wetland. *Remote Sensing*, *8*(7). <https://doi.org/10.3390/rs8070596>
- Duan, Z., & Bastiaanssen, W. G. M. (2013). Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sensing of Environment*, *134*, 403–416. <https://doi.org/10.1016/j.rse.2013.03.010>
- Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., ... Sullivan, C. A. (2005). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, *81*(02). <https://doi.org/10.1017/s1464793105006950>
- El-Shazli, A., & Hoermann, G. (2016). Development of storage capacity and morphology of the Aswan High Dam Reservoir. *Hydrological Sciences Journal*, *61*(14), 2639–2648. <https://doi.org/10.1080/02626667.2016.1151979>
- ERS-1: Aviso+. (n.d.). Retrieved October 23, 2019, from <https://www.aviso.altimetry.fr/en/missions/past-missions/ers-1.html>
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., ... Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, *45*(2). <https://doi.org/10.1029/2005rg000183>
- Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new

- technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, 140(October 2016), 23–35. <https://doi.org/10.1016/j.rse.2013.08.029>
- Gao, H. (2015). Satellite remote sensing of large lakes and reservoirs: from elevation and area to storage. *Wiley Interdisciplinary Reviews: Water*, 2(2), 147–157. <https://doi.org/10.1002/wat2.1065>
- Gao, H., Birkett, C., & Lettenmaier, D. P. (2012). Global monitoring of large reservoir storage from satellite remote sensing. *Water Resources Research*, 48(9). <https://doi.org/10.1029/2012WR012063>
- Google Earth Engine. (n.d.). Simple Cloud Score. Retrieved August 6, 2019, from <https://code.earthengine.google.com/5523bdba1566eb5809d8429a1dccfa34>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2016). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Göttl, F., Dettmering, D., Müller, F. L., & Schwatke, C. (2016). Lake level estimation based on CryoSat-2 SAR altimetry and multi-looked waveform classification. *Remote Sensing*, 8(11). <https://doi.org/10.3390/rs8110885>
- Heathcote, A. J., del Giorgio, P. A., Prairie, Y. T., & Brickman, D. (2015). Predicting bathymetric features of lakes from the topography of their surrounding landscape. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(5), 643–650. <https://doi.org/10.1139/cjfas-2014-0392>
- Heine, I., Stüve, P., Kleinschmit, B., & Itzerott, S. (2015). Reconstruction of lake level changes of groundwater-fed lakes in Northeastern Germany using rapideye time series. *Water (Switzerland)*, 7(8), 4175–4199. <https://doi.org/10.3390/w7084175>
- Hollister, J., & Milstead, W. B. (2010). Using GIS to estimate lake volume from limited data. *Lake and Reservoir Management*, 26(3), 194–199. <https://doi.org/10.1080/07438141.2010.504321>
- Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A Review. *Reviews of Geophysics*, 56, 333–360. <https://doi.org/10.1029/2018RG000598>
- Jiang, Nielsen, L., Andersen, K., Bauer-Gottwein, O. B., & Peter. (2017). CryoSat-2 radar altimetry for monitoring freshwater resources of China. *Remote Sensing of Environment*, 200, 125–139. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0034425717303681>
- Joeckel, R. M., & Diffendal, R. F. (2004). Geomorphic and environmental change around a large, aging reservoir: Lake C. W. McConaughy, Western Nebraska, USA. *Environmental and Engineering Geoscience*, 10(1), 69–90. <https://doi.org/10.2113/10.1.69>
- Kang, S., & Hong, S. Y. (2016). Assessing Seasonal and Inter-Annual Variations of Lake Surface Areas in Mongolia during 2000-2011 Using Minimum Composite MODIS NDVI. *PLOS ONE*, 11(3), e0151395. <https://doi.org/10.1371/journal.pone.0151395>
- Kansas Lakes and Reservoirs. (2016). Retrieved March 30, 2019, from <https://kars.ku.edu/maps/kansaslakes/>
- Keys, T. A., & Scott, D. T. (2018). Monitoring volumetric fluctuations in tropical lakes and reservoirs using satellite remote sensing. *Lake and Reservoir Management*, 34(2), 154–166. <https://doi.org/10.1080/10402381.2017.1402226>
- Kress, W. H., Seabee, S. K., Littin, G. R., Drain, M. A., & Kling, M. E. (2005). Comparison of Preconstruction and 2003 Comparison of Preconstruction and 2003 Bathymetric and Topographic

Surveys Bathymetric and Topographic Surveys of Lake McConaughy, Nebraska. *U.S. Geological Survey Scientific Investigations Report 2005-5040*.

Lake History. (2016). Retrieved from <https://ilovelakemac.com/lake-info/lake-history/>

Lakes Data for Minnesota. (n.d.). Retrieved October 15, 2019, from https://www.mngeo.state.mn.us/chouse/water_lakes.html

Landsat Surface Reflectance. (n.d.). Retrieved October 15, 2019, from https://www.usgs.gov/land-resources/nli/landsat/landsat-surface-reflectance?qt-science_support_page_related_con=0#qt-science_support_page_related_con

Liang, K., & Yan, G. (2017). Application of landsat imagery to investigate lake area variations and relict Gull Habitat in Hongjian Lake, Ordos Plateau, China. *Remote Sensing*, 9(10). <https://doi.org/10.3390/rs9101019>

Lu, S., Ouyang, N., Wu, B., Wei, Y., & Tesemma, Z. (2013). Lake water volume calculation with time series remote-sensing images. *International Journal of Remote Sensing*, 34(22), 7962–7973. <https://doi.org/10.1080/01431161.2013.827814>

Lu, S., Wu, B., Yan, N., & Wang, H. (2011). Water body mapping method with HJ-1A/B satellite imagery. *International Journal of Applied Earth Observation and Geoinformation*, 13(3), 428–434. <https://doi.org/10.1016/j.jag.2010.09.006>

Ma, Y., Xu, N., Sun, J., Wang, X. H., Yang, F., & Li, S. (2019). Estimating water levels and volumes of lakes dated back to the 1980s using Landsat imagery and photon-counting lidar datasets. *Remote Sensing of Environment*, 232(November 2018), 111287. <https://doi.org/10.1016/j.rse.2019.111287>

Mcalister, J. R., Fox, W. E., Wilcox, B., & Srinivasan, R. (2013). Reservoir volumetric and sedimentation survey data: A necessary tool for evaluating historic sediment flux and appropriate mitigation response. *Lakes and Reservoirs: Research and Management*, 18(3), 275–283. <https://doi.org/10.1111/lre.12036>

McFeeters, S. K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432.

McKay, C. P. (2014). Requirements and limits for life in the context of exoplanets. *Proceedings of the National Academy of Sciences of the United States of America*, 111(35), 12628–12633. <https://doi.org/10.1073/pnas.1304212111>

Medina, C. E., Gomez-Enri, J., Alonso, J. J., & Villares, P. (2008). Water level fluctuations derived from ENVISAT Radar Altimeter (RA-2) and in-situ measurements in a subtropical waterbody: Lake Izabal (Guatemala). *Remote Sensing of Environment*, 112(9), 3604–3617. <https://doi.org/10.1016/J.RSE.2008.05.001>

Messenger, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7. <https://doi.org/10.1038/ncomms13603>

Moradi, A., Metivier, L., de Viron, O., Calmant, S., & Mering, C. (2014). Evaluation of MODIS data for improved monitoring of the Caspian Sea. *International Journal of Remote Sensing*, 35(16), 6060–6075. <https://doi.org/10.1080/01431161.2014.943324>

Ogilvie, A., Belaud, G., Massuel, S., Mulligan, M., Le Goulven, P., & Calvez, R. (2018). Surface water monitoring in small water bodies: Potential and limits of multi-sensor Landsat time series. *Hydrology and Earth System Sciences*, 22(8), 4349–4380. <https://doi.org/10.5194/hess-22-4349->

2018

- Okeowo, M. A., Lee, H., Hossain, F., & Getirana, A. (2017). Automated Generation of Lakes and Reservoirs Water Elevation Changes From Satellite Radar Altimetry. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(8), 3465–3481. <https://doi.org/10.1109/jstars.2017.2684081>
- 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>
- Pan, F., Liao, J., Li, X., & Guo, H. (2013). Application of the inundation area-lake level rating curves constructed from the SRTM DEM to retrieving lake levels from satellite measured inundation areas. *Computers and Geosciences*, *52*, 168–176. <https://doi.org/10.1016/j.cageo.2012.09.024>
- Pass locator: Aviso+. (n.d.). Retrieved February 27, 2019, from <https://www.aviso.altimetry.fr/en/data/tools/pass-locator.html>
- Pekel, J. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. <https://doi.org/10.1038/nature20584>
- Peng, D., Guo, S., Liu, P., & Liu, T. (2006). Reservoir storage curve estimation based on remote sensing data. *Journal of Hydrologic Engineering*, *11*(2), 165–172. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2006\)11:2\(165\)](https://doi.org/10.1061/(ASCE)1084-0699(2006)11:2(165))
- Prigent, C., Papa, F., Aires, F., Jimenez, C., Rossow, W. B., & Matthews, E. (2012). Changes in land surface water dynamics since the 1990s and relation to population pressure. *Geophysical Research Letters*, *39*(8), 2–6. <https://doi.org/10.1029/2012GL051276>
- Rahmani, V., Kastens, J. H., de Noyelles, F., Jakubauskas, M. E., Martinko, E. A., Huggins, D. H., ... Blackwood, A. J. (2018). Examining storage capacity loss and sedimentation rate of large reservoirs in the Central U.S. great plains. *Water (Switzerland)*, *10*(2), 1–17. <https://doi.org/10.3390/w10020190>
- Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Borla Tridon, D., Bräutigam, B., ... Moreira, A. (2017). Generation and performance assessment of the global TanDEM-X digital elevation model. *ISPRS Journal of Photogrammetry and Remote Sensing*, *132*, 119–139. <https://doi.org/10.1016/j.isprsjprs.2017.08.008>
- Rokni, K., Ahmad, A., Selamat, A., & Hazini, S. (2014). Water feature extraction and change detection using multitemporal landsat imagery. *Remote Sensing*, *6*(5). <https://doi.org/10.3390/rs6054173>
- Schwatke, C., Dettmering, D., Bosch, W., & Seitz, F. (2015). DAHITI-an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrol. Earth Syst. Sci*, *19*, 4345–4364. <https://doi.org/10.5194/hess-19-4345-2015>
- Solander, K. C., Reager, J. T., & Famiglietti, J. S. (2016). How well will the Surface Water and Ocean Topography (SWOT) mission observe global reservoirs? *Water Resources Research*, *52*(3), 2123–2140. <https://doi.org/10.1002/2015wr017952>
- SRTM NED Vertical Differencing. (n.d.). Retrieved December 12, 2019, from https://www.usgs.gov/land-resources/eros/topchange/science/srtm-ned-vertical-differencing?qt-science_center_objects=0#qt-science_center_objects
- Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., & Iwamoto, H. (2014). Precise Global DEM Generation by ALOS PRISM. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial*

- Information Sciences*, II-4(May), 71–76. <https://doi.org/10.5194/isprsannals-ii-4-71-2014>
- Technical Service Center | Reservoir Surveys. (n.d.). Retrieved October 15, 2019, from <https://www.usbr.gov/tsc/techreferences/reservoir.html>
- Texas Water Development Board. (n.d.). Retrieved October 15, 2019, from <http://www.twdb.texas.gov/surfacewater/surveys/completed/index.asp>
- TOPEX/Poseidon Fact Sheet. (n.d.). Retrieved August 5, 2019, from <https://sealevel.jpl.nasa.gov/missions/topex/topexfactsheet/>
- Tseng, K.-H., Shum, C. K., Kim, J.-W., Wang, X., Zhu, K., & Cheng, X. (2016). Integrating Landsat Imageries and Digital Elevation Models to Infer Water Level Change in Hoover Dam. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(4), 1696–1709. <https://doi.org/10.1109/jstars.2015.2500599>
- Tulbure, M. G., & Broich, M. (2013). Spatiotemporal dynamic of surface water bodies using Landsat time-series data from 1999 to 2011. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79, 44–52. <https://doi.org/10.1016/j.isprsjprs.2013.01.010>
- Tulbure, M. G., Broich, M., Stehman, S. V., & Kommareddy, A. (2016). Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. *Remote Sensing of Environment*, 178, 142–157. <https://doi.org/10.1016/j.rse.2016.02.034>
- USGS. (2018a). How much of the Earth ' s water is stored in glaciers ? Retrieved March 7, 2020, from https://www.usgs.gov/faqs/how-much-earths-water-stored-glaciers?qt-news_science_products=0#qt-news_science_products
- USGS. (2018b). *Landsat 4-7 Surface Reflectance (Ledaps) Product Guide*. *Usgs*. [https://doi.org/10.1016/0042-207X\(74\)93024-3](https://doi.org/10.1016/0042-207X(74)93024-3)
- USGS. (2019). Landsat 8 Surface Reflectance Code (LASRC) Product Guide. (No. LSDS-1368 Version 2.0). Retrieved from <https://www.usgs.gov/media/files/landsat-8-surface-reflectance-code-lasrc-product-guide>, (December), 40.
- USGS EROS. (2017). *Landsat Collection 1 Level 1 Product Definition*. Retrieved from https://prd-wret.s3-us-west-2.amazonaws.com/assets/palladium/production/atoms/files/LSDS-1656_Landsat_Collection1_L1_Product_Definition-v2.0.pdf
- USGS Lakes and Reservoirs. (n.d.). Retrieved October 15, 2019, from <https://waterdata.usgs.gov/wa/nwis/current/?type=lakes>
- van Bemmelen, C. W. T., Mann, M., de Ridder, M. P., Rutten, M. M., & van de Giesen, N. C. (2016). Determining water reservoir characteristics with global elevation data. *Geophysical Research Letters*, 43(21), 11,278-11,286. <https://doi.org/10.1002/2016GL069816>
- VERTCON. (n.d.). Retrieved October 15, 2019, from <https://www.ngs.noaa.gov/TOOLS/Vertcon/vertcon.html>
- Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., ... Davies, P. M. (2010). Global threats to human water security and river biodiversity. <https://doi.org/10.1038/nature09440>
- Wang, X., Chen, Y., Song, L., Chen, X., Xie, H., & Liu, L. (2013). Analysis of lengths, water areas and volumes of the Three Gorges Reservoir at different water levels using Landsat images and SRTM

- DEM data. *Quaternary International*, 304, 115–125. <https://doi.org/10.1016/j.quaint.2013.03.041>
- Weekley, D., & Li, X. (2019). Tracking Multidecadal Lake Water Dynamics with Landsat Imagery and Topography / Bathymetry. *Water Resources Research*, 1–18. <https://doi.org/10.1029/2019WR025500>
- What is the vertical accuracy of the 3D Elevation Program (3DEP) DEMs? (n.d.). Retrieved December 12, 2019, from https://www.usgs.gov/faqs/what-vertical-accuracy-3d-elevation-program-3dep-dems?qt-news_science_products=0#qt-news_science_products
- Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., ... Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. <https://doi.org/10.1016/j.rse.2015.11.032>
- Xiong, L., Deng, R., Li, J., Liu, X., Qin, Y., Liang, Y., & Liu, Y. (2018). Subpixel surface water extraction (SSWE) using Landsat 8 OLI data. *Water (Switzerland)*, 10(5), 1–15. <https://doi.org/10.3390/w10050653>
- Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. <https://doi.org/10.1080/01431160600589179>
- Yuan, Lee, T., Jung, H., Aierken, H. C., Beighley, A., Alsdorf, E., ... Donghwan. (2017). Absolute water storages in the Congo River floodplains from integration of InSAR and satellite radar altimetry. *Remote Sensing of Environment*, 201, 57–72. <https://doi.org/10.1016/j.rse.2017.09.003>
- Zhang, S., Foerster, S., Medeiros, P., de Araújo, J. C., Motagh, M., & Waske, B. (2016). Bathymetric survey of water reservoirs in north-eastern Brazil based on TanDEM-X satellite data. *Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2016.07.024>
- Zhao, G., & Gao, H. (2018). Automatic Correction of Contaminated Images for Assessment of Reservoir Surface Area Dynamics. *Geophysical Research Letters*, 45(12), 6092–6099. <https://doi.org/10.1029/2018GL078343>
- Zhou, Y., Dong, J., Xiao, X., Xiao, T., Yang, Z., Zhao, G., ... Qin, Y. (2017). Open surface water mapping algorithms: A comparison of water-related spectral indices and sensors. *Water (Switzerland)*, 9(4). <https://doi.org/10.3390/w9040256>
- Zhu, W., Jia, S., & Lv, A. (2014). Monitoring the Fluctuation of Lake Qinghai Using Multi-Source Remote Sensing Data. *Remote Sensing*, 6(12), 10457–10482. <https://doi.org/10.3390/rs61110457>
- Zwally, D. H. J. (n.d.). NASA: ICESat. Retrieved August 5, 2019, from <https://icesat.gsfc.nasa.gov/icesat/glas.php>