

Quantifying water volumes of ungauged lakes using optical satellite imagery and high-resolution LiDAR DEMs

N. Bulovic^a , **D.R. Sari**^b  and **S. Sabaghy**^c 

^a Sustainable Minerals Institute, The University of Queensland, QLD 4072, ^b School of Earth and Environmental Sciences, The University of Queensland, QLD 4072, ^c Department of Civil Engineering, Monash University, Clayton Campus, VIC 3800
Email: n.bulovic@uq.edu.au

Abstract: Surface water bodies such as lakes, rivers and wetlands are critical water resources to both human and ecological systems, but are under increasing pressure from competing users. Monitoring of the location, quantity and movement of water is crucial in effectively managing these resources and implementing a sustainable water management strategy for the future. However, continuous information on the quantity and distribution of water across the landscape is limited in some regions because of the high cost of traditional in-situ monitoring. As such, global remote sensing datasets are being used more frequently to complement these sparse networks. This paper aims to develop new methods to estimate the quantity of water (volume) in open water storages such as lakes, using remote sensing data.

Lake Menindee, part of the greater Menindee Lakes complex in the Murray-Darling Basin, was selected as the case study of this research because of its geographic location and data availability. Water management in the Murray-Darling Basin has been under increasing scrutiny partly due to exposure of water theft by irrigators. As such, there is a pressing need for large scale monitoring of water resources in the region using novel data and methods.

This paper developed three methods to estimate water volumes in a lake, all of which only used a high-resolution (5m) LiDAR DEM in conjunction with optical imagery. As an initial preprocessing step, the water observations from space (WofS) algorithm (Mueller et al. 2016) was applied to Landsat optical imagery to detect areas of surface water in the lake which was used as an input to all the methods. The first method derived a relationship between lake inundated surface area and volume using the DEM. Subsequently, this relationship was used to convert WofS-derived surface areas to volumes. The second method evaluated the quality of match between the WofS spatial inundation pattern and DEM-modelled inundation patterns at 0.1m water level increments, from which an optimal match and the respective DEM-derived volume was picked. Quality of match was quantified with three metrics commonly used in weather forecasting. In the third method, the elevation of the WofS water body edge was derived from the DEM, and a volume was estimated by “filling” the lake DEM to this height.

Water volumes by all three methods were estimated using 19 years of high-quality Landsat data equivalent to 209 scenes, and daily gauged measurements were used for validation. A combination of scatterplots and statistical metrics were used for evaluation.

Initial findings show that all methods have reasonable skill in estimating water volumes with high Pearson correlation coefficients, and estimates from methods 2 and 3 have relative biases of less than 10 percent. No single method performed consistently better across all ranges of volumes, with method 3 having poorest performance for low volumes while method 1 substantially overestimated high volumes. Additionally, estimation errors were volume-dependent, with medium-range of volumes having highest accuracy estimates while prediction skill consistently worsened at higher volumes across all methods. Future research should further investigate drivers of the volume-dependent errors, expand the evaluations to multiple case studies, including the large on-farm water storages across the Murray-Darling Basin, and test other remote sensing data sources such as radar altimetry. These results clearly demonstrate the potential of remote sensing based methods for lake volume estimation.

Keywords: *Water volume, ungauged lakes, Water observations from space (WofS), LiDAR DEM*

1. INTRODUCTION

Water is an essential resource for every aspect of human life, used for drinking, agriculture, transportation, electricity production, and other industrial activities. Additionally, water is critical for preserving biodiversity and ecosystems. Surface water bodies such as lakes, rivers, and wetlands are critical water resources (Huang et al. 2018) that need to be properly managed to balance human use and ecosystem protection (Vörösmarty et al. 2010). Therefore, observation and monitoring of these water bodies is crucial for ensuring effective water management strategies. However, this is hampered in some regions by a lack of data, as traditional in-situ observation and monitoring networks are not cost-effective in capturing the temporal dynamics of numerous water bodies over large regions.

Remote sensing provides effective and efficient opportunities to continuously observe and monitor global water dynamics at various spatial and temporal scales (Huang et al. 2018) and to supplement in-situ networks. Previous studies have focused on the use of optical sensors such as Landsat for detecting and mapping surface water boundaries, with extensive research investigating the suitability, accuracy and limitations of several water detection indices and algorithms. Fisher et al. (2016) compared several water indices for automatic mapping of surface water across Qld and NSW using Landsat imagery. This study found that no single index performed best across all water types, and misclassification was most common in murky / green water or water bodies with large mixed pixel count. Similarly, Mueller et al. (2016) developed a water detection algorithm called Water Observations from Space (WOfS) to map surface water extents across Australia using the entire Landsat record, providing important multi-decadal information on water dynamics across the continent.

Even though water extent mapping provides valuable information on the distribution of water in the landscape, estimates of water quantity are necessary for effective management. Water extent information has been coupled with satellite altimetry data to estimate lake volume dynamics in the U.S.A., Europe and Africa (Baup et al. 2014; Duan and Bastiaanssen 2013; Tong et al. 2016). This approach couples coincident water level (altimetry) and surface area (optical imagery) data, from which a level-area-volume relationship is derived and applied to historical imagery. Despite this method performing well in most case study regions, it is limited to areas with altimetry data and has typically been applied to large reservoirs. More research is necessary to develop and assess volume estimation methods that are not dependent on altimetry data and are applicable to smaller reservoirs.

Therefore, this paper aims to develop water volume estimation methods using optical imagery in conjunction with other commonly available remote sensing datasets, such as DEMs. Three potential methods are evaluated utilising the WOfS algorithm for water detection and a LiDAR-derived high-resolution DEM. This case study is on Menindee Lakes in the Murray-Darling Basin, New South Wales, Australia, are used as a case study, where water management has been under increasing scrutiny due to water theft by irrigators and mass fish kills.

2. THE MENINDEE LAKES COMPLEX

The Menindee Lakes complex was chosen as the study location for our research due to its geographic position in the Murray-Darling Basin and data availability in an otherwise relatively data sparse region. The complex is located alongside the Darling River towards the western end of the basin and consists of four main lakes, namely lakes Wetherell, Pamamaroo, Menindee (the largest lake) and Cawndilla (Figure 1). The lakes are situated in a semi-arid environment (244 mm/yr mean rainfall) and are shallow with a large surface area. The NSW government manages the Lakes system under the Murray-Darling Basin Agreement, to provide water to ecosystems, local community needs and downstream water users. This study focuses on Lake Menindee in particular, which has a maximum storage volume of 900 GL and a long historical record of daily water volume measurements making it ideal for method validation (Figure 1b).

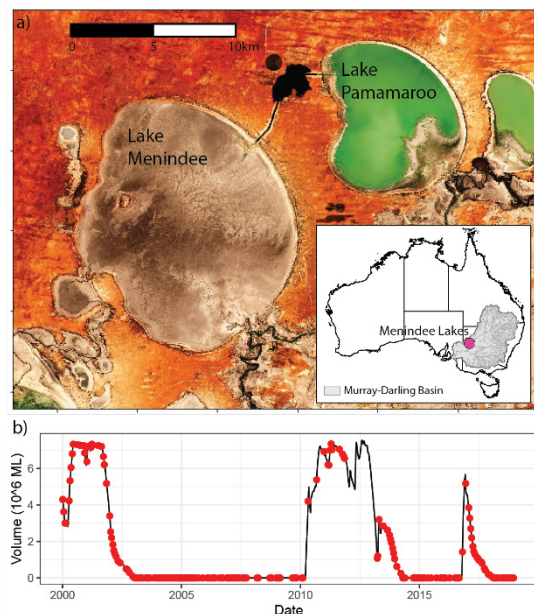


Figure 1. a) Location of the Menindee Lakes complex showing Lakes Menindee and Pamamaroo; b) Daily time series of Lake Menindee water volumes since 2000. Red dots indicate timing of Landsat images used in methods assessment.

3. METHODS

The research approach employed in this paper is made up of three main components which are discussed in further detail in the following sections: 1) data and data preprocessing, 2) water volume estimation, and 3) results validation. A flow chart depiction of the entire process is shown in Figure 2 with the three main subsections outlined in orange, green and blue respectively.

3.1. Data and data preprocessing

Water detection

For the purposes of quantifying reservoir water volumes at Lake Menindee, we used the WOfS algorithm (Mueller *et al.* 2016) as the primary means of surface water detection. WOfS is a nationally calibrated algorithm developed to classify decades of satellite imagery data at the continental scale into wet or dry classes. WOfS was applied to 19 years of Landsat data with a 25m pixel resolution (1 January 2000 to 31 December 2018) covering the Lake Menindee area. Only images with a minimum of 95% high quality pixels were considered for the volume calculations. High quality pixels are those without cloud cover, shadow and band saturation issues. Instead of using the ready-made WOfS product, we preferred this approach as it reduced spurious “not-wet” classifications due to pixel quality issues rather than pixels being genuinely dry. A total of 209 scenes of suitable quality were identified (Figure 1b), where at least some water was present in the lake in 198 scenes. Volumes were assumed to be zero if no wet pixels were identified.

LiDAR-derived DEM

A 5m cell-size LiDAR-derived DEM of Lake Menindee, produced by Geoscience Australia, was obtained from ELVIS. The lake was empty at the time of the LiDAR survey, making it ideal for quantifying storage volumes. To quantify the expected inundated surface area, storage volume and the spatial inundation pattern (pixel classification into wet or dry) at different water depths, the DEM was “filled” in 10 cm increments from empty (53.5 m ASL) to 1.5 m above the maximum recorded water height (63 m ASL), resulting in 96 water depths. “Filling” was carried out starting at the lake intake point with the “r.lake” tool in QGIS. DEM outputs were then used in Methods 1 and 2 (Section 3.2) for water volume estimations (Figure 2).

Lake Menindee gauged data

Historical daily measurements of water depth and volume at Lake Menindee were obtained from WaterNSW (Site no.: 425022). Gauge data was firstly used to check that the DEM-derived water storage volumes were reasonable. Next, the historical readings shown in Figure 1b were used to validate water volume estimates obtained from the different methods. As DEM-derived volumes had a small but consistent systematic bias when compared to the gauge data, historical readings were adjusted accordingly to remove this bias effect in the validation process. Lake rating curves were not freely accessible to determine the cause of this discrepancy, but we suspect that they were constructed with older and lower resolution data than the LiDAR DEM.

3.2. Water volume estimation

Method 1: Surface area – volume relationship

DEM-derived surface area and volume estimates (across different water depths) were used to develop an empirical surface area – volume relationship (SAV), and used as the first method of estimating water volumes in storage. This approach is fundamentally very similar to other altimetry-based methods such as that of Duan and Bastiaansen (2013) except that the SAV relationship is derived with a DEM as opposed to coincident

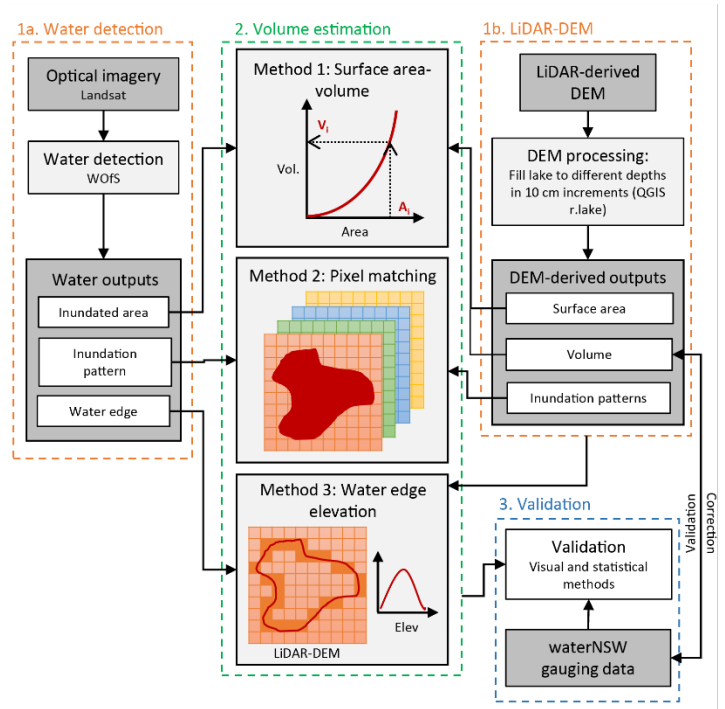


Figure 2. Flow chart of the research process with sections outlined in different colours: 1. data preprocessing (orange), 2. volume estimation (green) and 3. validation (blue).

altimetry readings and optical imagery. For each WOfS scene, lake surface area was calculated by quantifying the area of wet pixels within the lake extent and converted into a respective volume using the SAV relationship.

Method 2: Pixel matching

The second water volume estimation method focuses on the spatial pattern of lake inundation (i.e. which pixels are being classified as wet or dry) as opposed to just the total inundated area (irrespective of whether the same pixels are wet). This “matching” of pixels may be important in relatively flat regions where the “wetted” area could be inflated at low water levels due to shallow disconnected pools forming after heavy localized rain or receding waters. Additionally, in comparison to area-based methods, this approach should be far less susceptible to biases induced from poor quality and thus misclassified pixels (e.g. clouds) or missing data.

As this is a new approach, three metrics commonly used in weather forecasting were trialed and applied to quantify the similarity in inundation patterns (i.e. wet / dry pixel classifications) between a WOfS scene and a DEM-derived inundation pattern: 1) accuracy index (ACC), 2) critical success index (CSI), and 3) Heidke skill score (HSS). For the purposes of calculating scores, the WOfS scene was assumed to be the ‘observation’ while the DEM pattern the ‘simulation’. ACC measures the proportion of pixels within the lake that were correctly ‘simulated’ (as wet or dry), the CSI measures the proportion of correctly ‘simulated’ wet pixels to the total number of ‘observed’ or ‘simulated’ wet pixels, and the HSS quantifies the skill of the pixel matches accounting for matches due to random chance. Each of the metrics have a perfect score of one and are defined as follows:

$$ACC = \frac{h + n}{h + f + m + n} \tag{1}$$

$$CSI = \frac{h}{h + f + m} \tag{2}$$

$$HSS = \frac{2 \cdot (h \cdot n - f \cdot m)}{(h + f) \cdot (f + n) + (h + m) \cdot (m + n)} \tag{3}$$

where ‘h’ represents number of hits (pixels classified as wet by both WOfS scene and DEM), ‘m’ is number of misses (pixels classified as wet by DEM but as dry by WOfS scene), ‘f’ is number of false alarms (pixels classified as wet by WOfS scene but as dry by DEM) and ‘n’ is number of correct negatives (pixels classified as dry by both WOfS scene and DEM).

These metrics were applied as follows to estimate water volume of a WOfS scene: 1) Similarities in inundation patterns between the WOfS scene and each of the 96 DEM-derived patterns (corresponding to each 10 cm water depth increment) were quantified per metric. 2) Water volume was estimated by assigning the volume of the depth increment with the highest (optimal) score. 3) Steps 1 and 2 were repeated for each individual metric resulting in three estimates per scene, that are referred to from here on as method 2 (ACC, CSI or HSS).

Method 3: Water edge elevation

Lastly, elevation of optical imagery-derived water boundary was used to derive storage water volumes using a similar approach to that of Tseng *et al.* (2016). The water body boundary (edge) of Lake Menindee was derived in each WOfS scene as the outside of the wet pixels. DEM elevations of overlapping pixels were extracted and boundary elevation statistics (median and 25th and 75th percentiles) quantified. The QGIS “r.lake” tool was used to “fill” the lake to the three respective water heights, from which three volumes for each WOfS scene were estimated, that are referred to from here on as method 3 (median, 25th or 75th percentiles).

3.3. Validation

Corrected daily WaterNSW gauge readings were used to validate water volume estimates over the 19 year analysis period. Coincident gauge readings were extracted for each WOfS scene using the Landsat image timestamp. The first part of the validation procedure assessed whether the estimation methods were able to determine if the lake was empty or not (referred to as “dry lake detection”) using a contingency table. In addition, volume estimates of scenes falsely identified as not dry were investigated to determine the magnitude of errors. The second part of the validation procedure evaluated and compared estimates from the three approaches using scatter plots and statistical metrics including relative bias (RB), normalised root-mean-square-error (nRMSE), mean absolute bias (MAE), Pearson correlation coefficient (r), and fitted slope of a linear regression. Only estimates of WOfS scenes correctly identified as “not dry” were considered. The influence of volume and hydrograph position on estimates was also explored by categorising the data into three (low ≤ 100,000 ML < medium < 600,000ML ≤ high) and two classes (rising-falling limb) respectively.

4. RESULTS

4.1. Dry lake detection

This section assessed how accurately the estimation methods could determine whether the lake was empty (dry) or not using a contingency table approach shown in Table 1. The lake was considered ‘dry’ if the volume was 0ML and ‘wet’ if it was greater than 0ML. The columns show the classification based on WaterNSW data and the rows are based on the estimation methods, where all methods with same results were lumped together for brevity. All volume estimation methods show good skill at detecting if there is any water in the lake, with only one scene being wrongly classified as ‘dry’ while actually ‘wet’ according to WaterNSW (Table 1). On the other hand, all methods have a tendency to classify lakes as ‘wet’ when they are ‘dry’ according to WaterNSW.

Table 1. Dry lake detection contingency table. All methods with same results are lumped together for brevity.

		WaterNSW data		
		Dry	Wet	Total
Methods 1, 2 and 3 (75th percentile)	Dry	10	1	11
	Wet	73	125	198
Method 3 (25th percentile)	Dry	47	1	48
	Wet	36	125	161
Method 3 (median)	Dry	15	1	16
	Wet	68	125	193
Total		83	126	

Table 2 shows the volume estimate statistics of misclassified scenes for each method, indicating the number of scenes that estimated a volume below / above a certain threshold and also the maximum estimated volume. When interpreting these results,

Table 2. Volume estimate statistics of misclassified scenes (lake was classified as ‘dry’ by WaterNSW and ‘wet’ by the methods) for each method.

		1	2: ACC	2: CSI	2: HSS	3: 25th perc.	3: Median	3: 75th perc.
Number of data points:	All	73	73	73	73	36	68	73
	Vol≤10ML	63	73	69	69	28	41	56
	Vol≥100ML	0	0	1	1	1	2	10
Max volume (ML)		73	8	5,858	5,858	25,819	59,139	129,711

lower volumes indicate smaller errors while larger volumes indicate larger errors as the lake is ‘dry’ according to the WaterNSW data. The results indicate that method 2 (ACC), followed by method 1, have the smallest volume estimates, i.e. the smallest errors. Overall, nearly all volume estimates are very small (less than 100 ML), other than for method 3 (75th percentile) which has the poorest performance (10 scenes have volume >100ML and a maximum value of 129,711ML). A visual inspection of the WofS scenes with larger errors shows that these are as a result of small patches of water distributed at higher points on the lake floor.

4.2. Water volume estimation

Next, accuracy of the volume estimates from the three estimation methods was evaluated. For this purpose, only scenes with water estimates greater than 0ML were considered (i.e. scenes classified as ‘wet’ in the previous section). Figure 3 shows a scatterplot of observed versus predicted volumes for all three methods and metrics. Overall, all methods exhibit relatively good skill in estimating volumes across the low and middle classes. However, there is a general divergence from the observations at higher values across all methods. This could be indicative of poorer prediction skill at higher volumes, but it might also be attributed to errors in the lake rating curve conversions from water depth to volume at higher volumes. This could not be checked as information on the rating curve and its derivation is not accessible.

Figure 4 shows the suite of statistical metrics that were computed for all three methods for the entire dataset and the individual volumetric and hydrograph classes. For the entire dataset, these metrics confirm that all methods have overall satisfactory estimation skill, with reasonably low relative biases (+10 to -20%), slopes (0.75 to 1.25), nRMSEs (0.15 to 0.3), and very high Pearson correlation coefficients (~ 0.99). Most of the methods tend to underestimate volumes, other than for method 1 which overestimates volumes by 10%.

Overall, method 3 (75th percentile) has the highest accuracy, with an RB of -3% and the best scores across four of the five measures (RB, nRMSE, MAE, slope). All three method 2 metrics are the next most accurate, with nearly identical nRMSE, r, and slope scores when compared to method 3 (75th percentile), however CSI minutely outperforms ACC, followed by HSS as the least accurate metric. Conversely, method 3 has a definite hierarchy with performance substantially improving from the 25th percentile to the median, and from the median to the 75th percentile.

Next, method performance was evaluated and compared across the three volume classes (low, medium, and high). A negative RB score and slope less than 1 indicates that there is a general tendency to underestimate

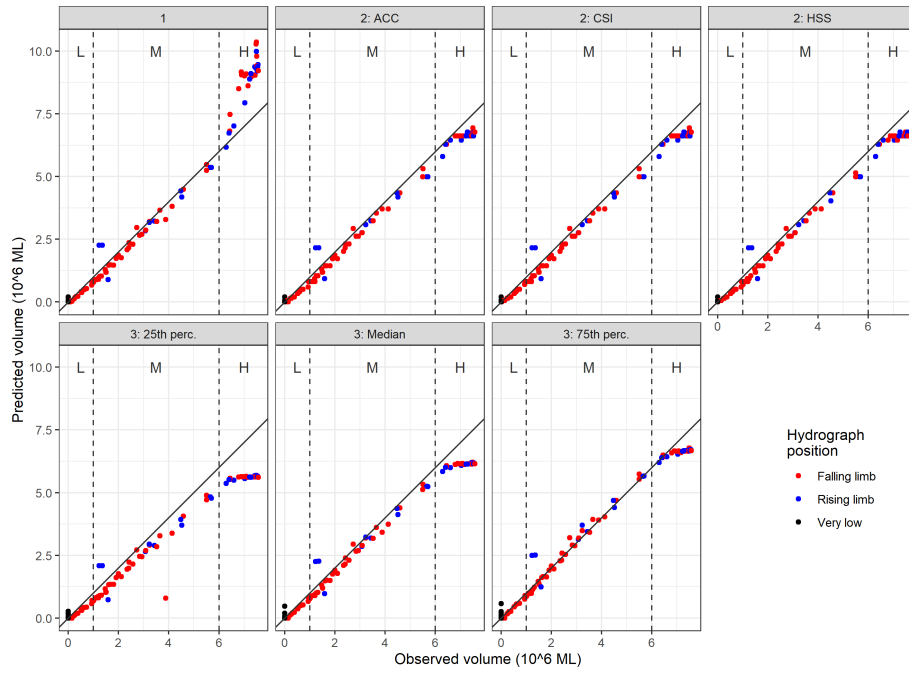


Figure 3. Scatterplot of observed versus predicted volume for all three methods and metrics. Boundaries of volumetric classes are indicated with vertical dashed lines, the solid line shows the 1:1 relationship, while points are coloured based on their hydrograph position.

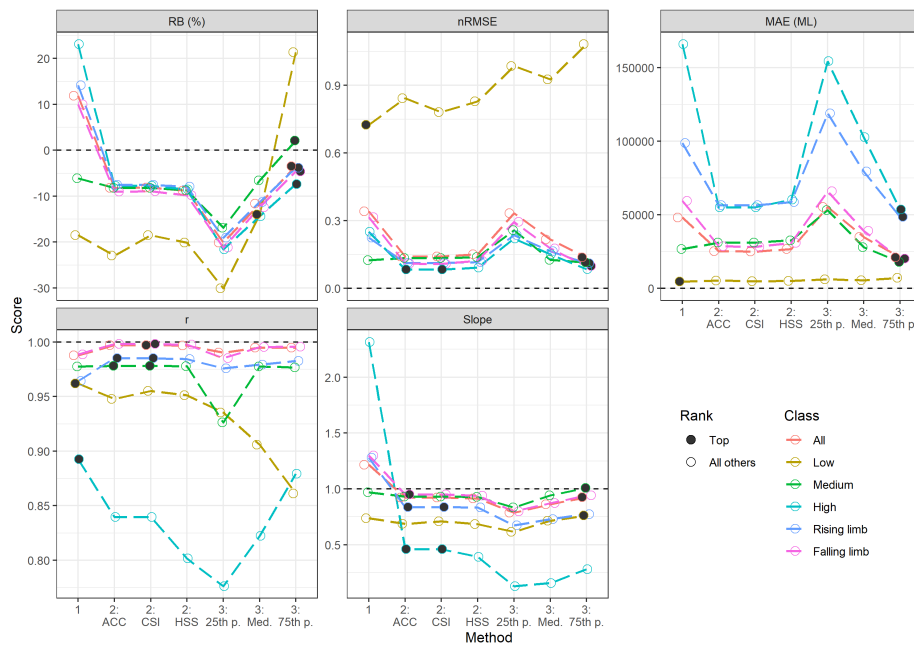


Figure 4. Plots of statistical metrics for all three methods and metrics for the entire dataset (all) and also split into volumetric (low, medium and high) and hydrograph (rising and falling limb) classes. Black dots highlight best performing method for each class.

volumes across all classes and methods, with the exception of “high” volumes being overestimated by method 1, and method 3 (75th percentile) overestimating “low” volumes while having nearly unbiased estimates of “medium” volumes (RB ~0%). Estimates of “medium” volumes generally have the highest accuracy and consistency across all the different methods, other than for method 3 (25th percentile). Interestingly, even though method 1 had the second lowest prediction accuracy when using the entire dataset, this result was not reflected across individual volume classes. Method 1 “low” volume estimates had the highest scores across 3 of the 5 measures (nRMSE, MAE, and r). On the other hand, the opposite trend was observed for method 3

(75th percentile) which had relatively poor accuracy at “low” volumes (lowest nRMSE, MAE, and r). Lastly, differences in accuracy were not as evident between hydrograph classes (rising or falling limb).

5. DISCUSSION AND CONCLUSIONS

In this study, we investigated the potential of using a combination of optical imagery (Landsat) and DEMs for estimating the quantity of water being stored in lakes / reservoirs over time without any in-situ monitoring. We evaluated the accuracy of three different estimation methods. To the authors’ best knowledge, this is the first time the method 2 has been applied to estimate lake water volumes. Lake Menindee, a steep-sided lake with a large surface area, was used as the case study due to data availability and its locally typical morphology.

A summary of the main findings to date are: 1) All methods have reasonable water volume estimation skill with high correlation scores and the higher accuracy methods have relative biases of less than 10%. 2) Most methods have a tendency to slightly underestimate volumes. 3) Accuracy consistently degrades at high volumes, this could be linked to potential errors in the lake rating curve at high volumes but this data could not be verified. 4) Not all methods perform equally well across all volume ranges, with methods 1 and 2 exhibiting better skill at low volumes (for both dry lake detection and rate estimation), while method 3 does best in the medium volume range.

The findings of this paper have potentially significant implications not only for scientists but also for water practitioners such as managers and compliance officers, who are expected to equitably manage a resource that is under increasing stress and is a point of conflict for competing users. To improve the usefulness of our approach, future research should investigate the source of errors of the proposed methods (especially for higher volumes), and test them on multiple case studies of ideally differing characteristics (volume, surface area, etc.). Furthermore, the current methods use Landsat data which has two main limitations: 1) relatively infrequent readings compounded by cloud obstruction (our study averaged a good image every 33 days), and 2) coarser resolution (25m) which may be inadequate in capturing water dynamics of smaller features. Newer products derived from Sentinel 1 and 2 may prove useful in addressing these issues. Future work should aim to merge multiple data sources, and even use more frequent ancillary datasets associated with water presence such as land surface temperature. Addressing these limitations is important before this approach can be confidently applied at larger scales.

ACKNOWLEDGEMENTS

Thanks to Albert van Dijk, OzEWEX and the 2018-19 ACAWSI sponsors for making the program and this project possible. This project was also partially funded by an AusAID grant. The authors acknowledge support from the Geoscience Australia DEA team and Nathan Clough (MDBA).

REFERENCES

- Baup, F, F. Frappart, and J. Maubant (2014). Combining high-resolution satellite images and altimetry to estimate the volume of small lakes. *Hydrology and Earth System Sciences*, 18(5), 2007-2020.
- Duan, Z., and W.G.M. Bastiaanssen (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.
- Fisher, A., N. Flood, and T. Danaher (2016). Comparing Landsat water index methods for automated water classification in eastern Australia. *Remote Sensing of Environment*, 175, 167-182.
- Huang, C, Y. Chen, S. Zhang, and J. Wu (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: A review. *Reviews of Geophysics*, 56(2), 333-360.
- Mueller, N., A. Lewis, D. Roberts, S. Ring, R. Melrose, J. Sixsmith, L. Lymburner, A. McIntyre, P. Tan, and A. Ip (2016). Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341-352.
- Vorosmarty, C.J., P.B. McIntyre, M.O. Gessner, D. Dudgeon, A. Prusevich, P. Green, S. Glidden, S.E. Bunn, C.A. Sullivan, C.R. Liermann, and P.M. Davies (2010). Global threats to human water security and river biodiversity. *Nature*, 467(7315), 555-561.
- Tong, X., H. Pan, H. Xie, X. Xu, F. Li, L. Chen, X. Luo, S. Liu, P. Chen, P and Y. Jin (2016). Estimating water volume variations in Lake Victoria over the past 22 years using multi-mission altimetry and remotely sensed images. *Remote Sensing of Environment*, 187, 400-413.
- Tseng, K., C.K. Shum, J. Kim, X. Wang, K. Zhu, and X. Cheng (2016). Integrating Landsat imageries and digital elevation models to infer water level change in Hoover Dam. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9, 1696-1709.