https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Assessing historic water extents in rapidly changing lakes: a hybrid remote sensing classification approach

Connor Mullen [1] and Marc F. Muller [1]

[1]Department of Civil and Environmental Engineering and Earth Sciences at University of Notre Dame

Correspondence: Marc F. Muller (mmuller1@nd.edu)

Abstract. The empirical attribution of rapid hydrologic change presents a unique data availability challenge in terms of establishing baseline prior conditions. On the one hand, one cannot go back in time to collect the necessary in situ data if it were not serendipitously collected when the change was taking place. On the other hand, modern satellite monitoring missions are often too recent to capture changes that are ancient enough to provide sufficient observations for adequate statistical inference. In that context, the four decades of continuous global high resolution monitoring enabled by the Landsat missions are an unrivaled source of information to study hydrologic change globally. However, extracting the relevant time series information in a systematic way across Landsat missions remains a monumental challenge. Cloud masking and inconsistent image quality often complicate the automatized interpretation of optical imagery.

Focusing on the monitoring of lake water extents, we address this challenge by coupling supervised and unsupervised image classification techniques. Unsupervised classification is first used to detect water on unmasked (cloudless and high quality) pixels. Classification results are then compiled across images to estimate the inundation frequency of each pixel, hinging on the assumption that different pixels will be masked at different times. Inundation frequency is then leveraged to infer the inundation status of masked pixels on individual images through supervised classification. Applied to a representative set of global and rapidly changing lakes, the approach successfully captured water extent fluctuations obtained from in situ gauges (when applicable), or from other Landsat missions during overlapping time periods.

1 Introduction

Large lakes often sustain fragile ecosystems while providing critical ecological services to local communities (i.e. fisheries, irrigation and water supply (Kummu et al., 2010). Despite their importance, many lakes are undergoing rapid change, both in their volume and seasonal fluctuation. Once imposing bodies of water have declined to a small fraction of their historical volume in many parts of the world, with the Aral Sea standing out as an iconic example (Micklin, 2007). In addition to dessication, the seasonality of lakes in humid climates are also rapidly changing. Seasonal fluctuations in the Tonle Sap, which are vital to local fisheries, are rapidly changing as a consequenceof shifting flow regimes in the Mekong River Basin (Kummu

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

and Sarkkula, 2008). These changes often emerge as a result of complex interplay of natural (e.g., changing temperatures and precipitations) and anthropogenic (damming and irrigation) factors (Haddeland et al., 2014). Proper attribution of their drivers is critical to inform policy, particularly in complex transboundary basins where the lake often plays a key strategic role (Wine et al., 2019). By providing a consistent space-time representation of the earth system, satellite remote sensing is a promising data source for causal attribution of historic change (Sheffield et al., 2018; Müller and Levy, 2019). Satellite imagery with global coverage provides the unique ability to study change *ex post* in situations where in situ observations are nonexistent, unavailable or disputed.

A substantial body of recent research has focused on monitoring surface water extents using satellite data, with applications ranging from small reservoirs (Avisse et al., 2017; Müller et al., 2016) to large continental water bodies (Mercier et al., 2002), paddy rice fields (Islam et al., 2010), tidal floods (Yan et al., 2010) and global-scale assessments (Busker et al., 2019). Active remote sensing at microwave frequencies has long been used to monitor the elevation (radar altimetry) and extent (synthetic aperture radar - SAR) of surface water bodies (Van Den Hoek et al., 2019). These approaches are unimpeded by clouds and do not rely on reflected sunlight. Space-borne radar altimeters (e.g., Topex/Posidon, GFO, Jason 1 and 2) can detect changes in surface water elevation on the order of decimeters, with return times of 10-35 days (Avisse et al., 2017). Although the limitations normally associated with radar altimetry (narrow swath, coarse cross-track spacing and large along-track path length (see Yale et al., 1998)) have limited implications for large lakes, radar altimetry cannot directly be used to map water coverage or estimate storage volumes, which are often the considered policy-relevant variables. These characteristics can only be determined when combined with digital elevation models (DEMs) (with vertical accuracies on the order of meters (Yan et al., 2015)) or water extent observations from other remote sensing sources (Busker et al., 2018). As an alternative to radar altimeters, SAR have long been used to monitor water extents (as opposed to elevation) at a large scale (Bioresita et al., 2018). These approaches leverage the fact that areas of open, smooth water bodies exhibit lower back-scatter coefficients in the X, L or C band. A number of recently launched SAR missions (e.g. COSMO-SkyMed, TerraSAR-X and Sentinel-1) allow for detection of water at resolutions and return times that are appropriate to capture rapidly unfolding local changes (Pérez Valentín and Müller, 2020). For instance, Sentinel 1 was launched in 2014 and has a 6-day revisit time and a spatial resolution of 20 m. However, while promising in their ability to monitor ongoing changes, these recent sensors are unable to capture events that happened before their launch. Studying changes that occurred sufficiently far in the past is often necessary for enough (post-event) observations to be available to attribute the change with adequate statistical power.

In this context, optical satellite imagery remains an important source of information for the study of historic changes. Passive mulispectral sensors from the Landsat missions have collected high resolution (30 m) images at the global scale since the early 1990s at weekly to bi-weekly return times. A number of spectral indices have been proposed to detect water on multispectral imagery, including the Modified Normalized Difference Water Index (MNDWI) (Xu, 2006) used in this study. These indices leverage the high contrast between land and water at specific frequencies of the electromagnetic spectrum, and a range of techniques have been developed to systematically classify pixels (as "wet" or "dry") based on their spectral index. These methods range from manual thresholding using a raw MNDWI image (Xu, 2006) to complex and dynamically calibrated machine learning algorithms including maximum-likelihood classification, decision trees , artificial neural networks,

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

fuzzy clustering method and k-means clustering (Lu and Weng, 2007). These approaches are limited in their ability to identify water in situation when surface reflectance is impeded by poor image quality (e.g., Landsat 7 Scan Line Corrector failure on images taken after May 2003) or clouds. Clouds in particular obstruct the electromagnetic signal at the frequencies necessary to identify inundated surfaces, and cloud shadows have a similar spectral signature to that of open water (Zhu and Woodcock, 2012). These limitations are particularly restrictive for large lakes in highly seasonal climates which are likely to be at least partially masked by clouds at any given time. For instance, our analysis suggests that the Tonle Sap Lake in Cambodia did not have a single cloud-free Landsat 7 image during 14 of the last 20 rainy seasons (here cloud-free is interpreted as less then 5% cloud coverage and rainy seasons extend between May 1 and October 31).

A variety of statistical interpolation approaches have been developed to infer the status of masked categorical pixels (see Shen et al., 2015, for an extensive review). For lake area monitoring, these approaches often build on the intuition that a masked pixel within the lake's footprint will likely be inundated if its elevation is lower than that of an *unmasked* pixel that is inundated. Avisse et al. (2017) used a DEM to infer the inundation status of masked pixels through non-linear regressions that accommodate the typically large vertical uncertainty of DEMs. The approach was predominantly applied to man-made reservoirs that were once empty, which allowed DEMs to be used to estimate bathymetry. More recently, Van Den Hoek et al. (2019) measured lake surface elevation via radar altimetry and utilized a regression framework to infer area-volume relationships and classify masked pixels.

Here we combine supervised and unsupervised classification approaches to infer the inundation status of each masked (i.e. cloudy) pixels based their inundation history. The proposed approach relies on the crucial assumption (further discussed in Sect. 4) that clouds are randomly distributed in time and space and that no pixel is permanently covered by clouds. Unlike previous efforts, the approach does not rely on secondary sources of information, which may be unavailable at the appropriate frequency or study period (radar altimetry as in Van Den Hoek et al. (2019)) or prone to high levels of uncertainty (DEM as in Avisse et al. (2017)). Instead, the proposed method relies exclusively on Landsat imagery, making it scalable and accessible (i.e., through Google Earth Engine (Gorelick et al., 2017)), and a promising tool to study long-term change through the analysis of historic imagery from older Landsat missions.

2 Methods

2.1 Water detection algorithm

The proposed algorithm consists of four main steps, depicted on Fig. 1.

- **Pre-processing**. A rudimentary cloud-scoring algorithm available on Google Earth Engine (SimpleCloudScore) is used to detect and mask clouds based on "Top of Atmosphere" Landsat reflectance images (see Supplementary Information (SI) S3). Pixels indicated as faulty (e.g., due to the Landsat 7 Scan Line Corrector failure) are also masked out. Individual images can then optionally be aggregated at the desired observation frequency (here monthly) through Normalized Difference Vegetation Index (NDVI) based maximum value compositing (Chen et al., 2003).

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

- **Stage 1: Unsupervised classification.** The inundation status ("wet" or "dry") of the *unmasked* pixels of each monthly image is estimated through unsupervised classification.
- **Stage 2: Supervised classification.** Classified images are then assembled to produce an Inundation Frequency (IF) image indicating the historic probability of a pixel being classified as "inundated". The IF raster is then used to infer the status of the *masked* pixels in each image through supervised classification, using classified (*unmasked*) pixels as training.
- **Post-processing.** A time series of lake area is generated by counting, on each monthly classified image, the number of inundated pixels within a predetermined polygon encompassing the maximum historical extent of the lake. Outliers are removed from the time series using the automatic detection approach suggested in Chen and Liu (1993).

Implementation details of the two classification steps (Stages 1 and 2) are discussed in the following paragraphs. The algorithm is implemented on Google Earth Engine with operational JavaScript source code openly accessible at the url: "https://code.earthengine.google.com/bed0dacc981bb883533f4958e80b50be".

2.1.1 Unsupervised classification

Cloud-free pixels of each monthly image were classified as wet or dry based on their modified normalized difference water index (MNDWI) value (Xu, 2006):

$$MNDWI = \frac{G - MIR}{G + MIR} \tag{1}$$

where G and MIR are the green and mid-infrared range of the electromagnetic spectrum, respectively (see Table 1 for corresponding bands in the considered imagery). The MNDWI enhances water/land contrasts by leveraging the ability of open water (compared to dry land) to preferentially absorb and reflect in the MIR and G regions of the electromagnetic spectrum, respectively. A clustering algorithm is applied to each image to identify the MNDWI threshold that partitions its pixels into two sets, so as to minimizes the MNDWI variance within each set. Because it can dynamically separate dry and wet pixels in cloud-free images, unsupervised classification stands as a promising (and somewhat less arbitrary) alternative to the manual determination of classification thresholds implemented in past studies (e.g., Müller et al. (2016) among others). However, by minimizing within-cluster variance, k-means tends to favor clusters of comparable sizes (Jain, 2010), which is problematic for cloudy images with preferential cloud covers on either land or water. As an extreme example, if all unmasked pixels are covered by water, a two-cluster k-means classification will not be able to distinguish water from land. We address this issue by computing the median value from the set of MNDWI thresholds obtained from the classification of individual images. This single median MNDWI threshold is then used to classify all unmasked pixels from all monthly images. Assuming the unsupervised classification can distinguish water from dry land on *most* images, the median threshold will allow for the identification of all unmasked pixels from the above extreme example as "wet". Using a single threshold comes at the cost of

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

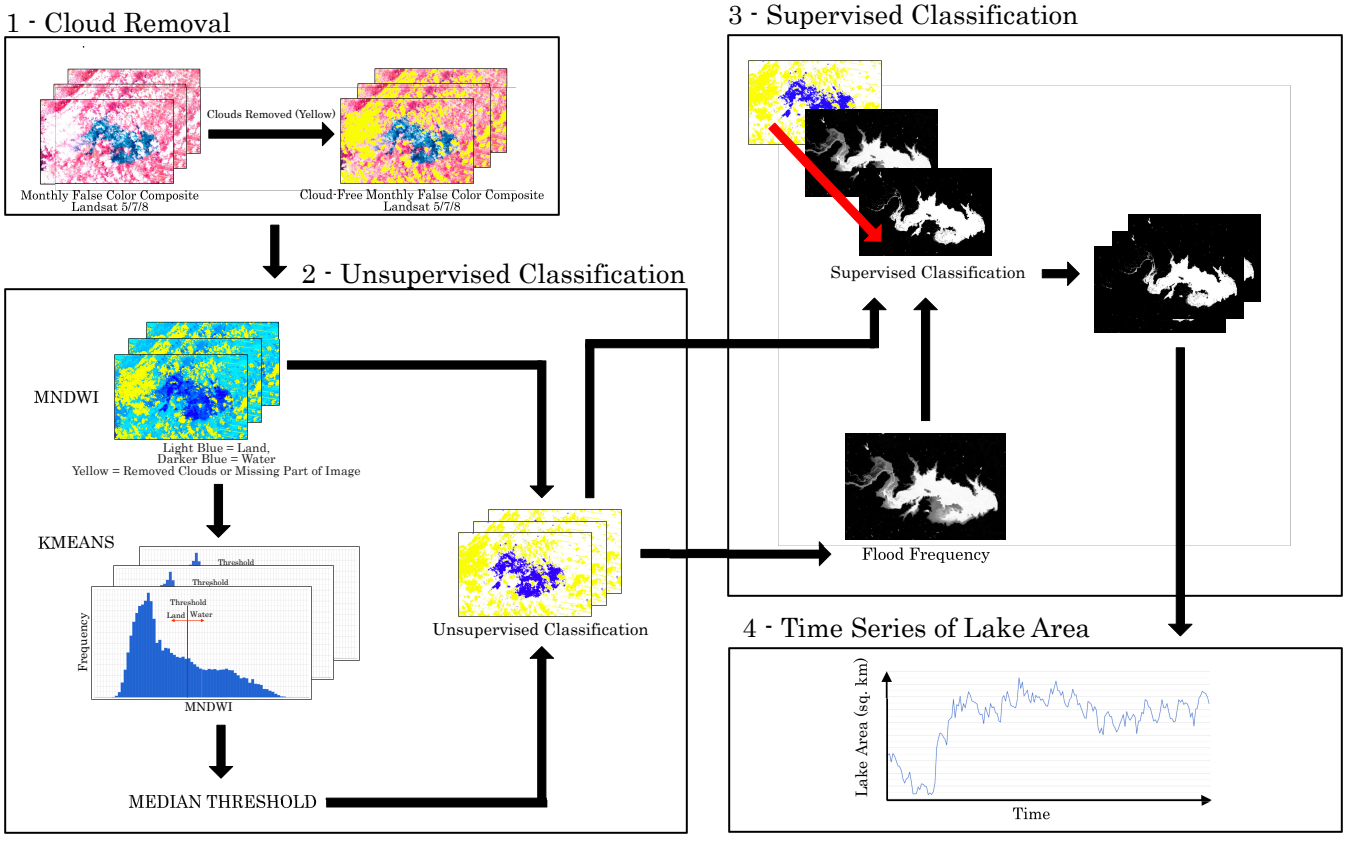

Figure 1. Lake surface monitoring algorithm in four steps: (i) Pre-processing, (ii) Stage 1: Unsupervised classification, (iii) Stage 2: Supervised classification and (iv) Post-processing: Aggregation and noise removal.

induced classification errors on individual images, e.g., due to changing land cover or atmospheric conditions. This important trade-off is further investigated in Sect. 4.

2.1.2 Supervised classification

A supervised classification approach was used to infer the wet/dry status of masked (cloudy) pixels based on their inundation history. To do so, space-time information on historic water extents are compiled into a single IF image representing the historic probability of each pixel i being identified as "wet" by the threshold-based unsupervised classification:

$$IF_i = \frac{N_i^{(wet)}}{N} = 1 - \frac{N_i^{dry} + N_i^{masked}}{N} \quad (2)$$

where $N_i^{(wet)}$, $N_i^{(dry)}$ and $N_i^{(masked)}$ are (respectively) the number of times pixel i is classified as "wet", "dry" or masked (i.e. cloudy) over the N considered monthly images. The IF image is then used to classify masked images by relying on two

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

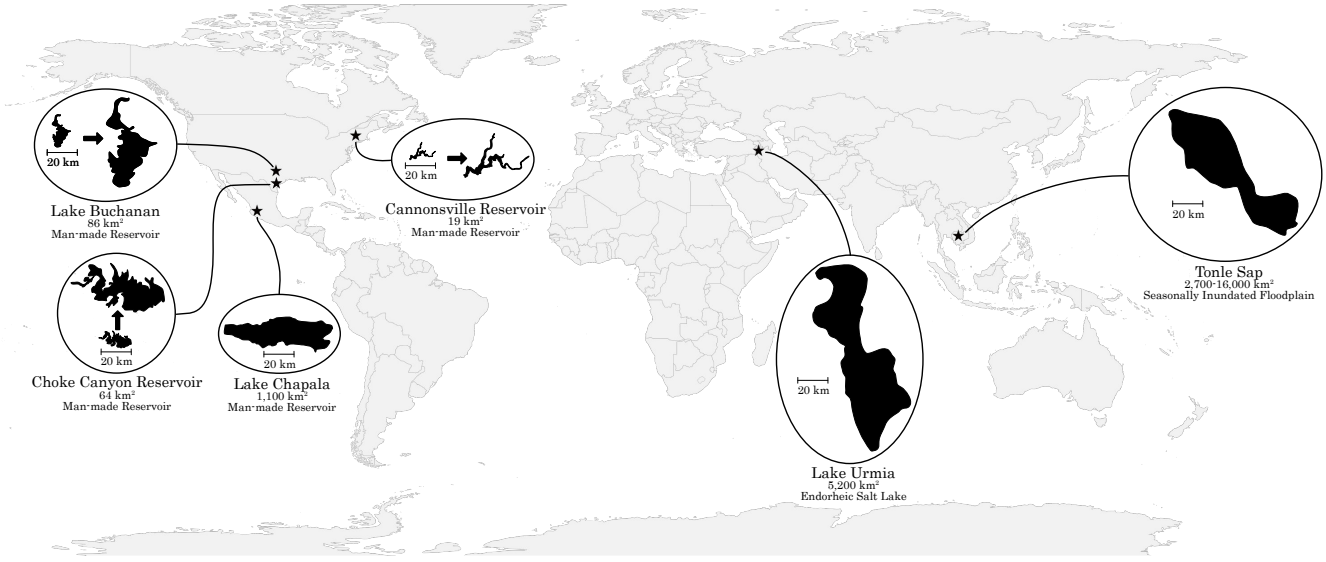

Figure 2. Location and characteristics of the considered water bodies.

critical assumptions (further discussed in 4.2): (i) cloud coverage is not affected by the inundation status of a pixel; and (ii) a more often inundated pixel is "lower" topographically, and therefore more likely to be inundated. More specifically, if cloud-free pixels associated with a certain IF value are inundated in a given image, it is very likely that pixels associated with an equal or higher IF (i.e. pixels of equal or lower elevation) are also inundated. We instrumentalize this principle in a supervised classification process (random-forest algorithm Pelletier et al. (2016)), that estimates the statistical association between the (known) classification status of unmasked pixels and their inundation frequency given by the IF image. The algorithm then uses the estimated relationship to infer the status of the masked pixels of that image based on their own inundation frequency. Classification noise that emerges from the uncertainty of the estimated statistical relationship are then removed through morphological filtering (Schowengerdt, 2006). The supervised classification algorithm is run independently on each individual image using a different set of unmasked classified pixels as training, but using the same IF image as the independent variable. This implies that the relative inundation frequency of each pixel (i.e. the bathymetry of the lake) is assumed to not change over the duration of the study period, which is a corollary of assumption (ii) discussed above.

2.2 Data sources and validation

The approach was validated on Lake Buchanan (TX, USA), Choke Canyon Reservoir (TX, USA) and Cannonsville Reservoir (NY, USA), for which long (17 to 47 years) time series of daily lake levels were obtained from the United States Geological Survey and the Texas Water Development Board. The two water bodies in Texas are emblematic of rapidly changing lake conditions that emerge in intensively managed lakes in semi-arid areas. Cannonsville Reservoir represents the complex topography and strongly seasonal climate and land-cover (including snow and ice) that emerges in mountainous regions and affects

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Satellite	Spectral Bands	Resolution	Return Time	Coverage
Landsat 5	B2-Green (0.52-0.60 μm) B5-MIR1 (1.55-1.75 μm)	30 m	16 days	1984 - 2013
Landsat 7	B2-Green (0.52-0.60 μm) B5-MIR1 (1.55-1.75 μm)	30 m	16 days	1999 - Present
Landsat 8	B2-Green (0.53-0.59 μm) B6-MIR1 (1.57-1.65 μm)	30 m	16 days	2013 - Present

Table 1. Properties of Data Sources (Survey and , 2015)

the performance of our approach (see Sect. 4). For each lake, monthly water extents were determined based on daily water levels using the provided elevation-area-capacity tables (see S2 for detailed procedures) and corrected for additive bias. Indeed, visual inspection of satellite classification output during cloudless days when water and land can be clearly distinguished (Fig. 3) reveal a near-perfect detection of water extents, which area does not match the area obtained from in situ water level observation. The source of this bias is unclear (and beyond the scope of this paper) and may be related to temporal changes in the elevation-area-capacity curves, for instance due to sedimentation processes (Raje and Mujumdar, 2010). Bias correction was implemented as follows:

$$A_i^* = A_i - \overline{A_i} + \overline{A_i^{(RS)}} \tag{3}$$

where A_i indicates a (biased) in situ area estimate, $\overline{A_i}$ its average value across observations, and $\overline{A_i^{(RS)}}$ indicates the average lake area from the remote sensing observation.

Additive biases notwithstanding, the above validation demonstrates the ability of the approach to capture changes in lake water extent with no in situ data requirements. We illustrate this capability by applying the approach to reconstruct the historic water extents of three rapidly changing lakes that are of major regional significance: Lakes Tonle Sap (Cambodia), Urmia (Iran) and Chapala (Mexico). We did not have access to in situ data from these lakes for direct validation. Instead, we compared estimates from the Landsat 7 to estimates from Landsats 5 and 8 during respective overlapping periods. This provides a reasonable estimate of the error in data-scarce regions as error sources across Landsat missions are close to independent (different sensors on different space platforms taking images at different times, see S1).

3 Results

3.1 Validation against in situ observation

Validation results are presented in Fig. 3, where lake area outputs from Landsats 5 (red) and 7 (blue) are compared to bias-corrected in situ observations. Outliers removed by the automatic detection procedure (Chen and Liu, 1993) are displayed

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

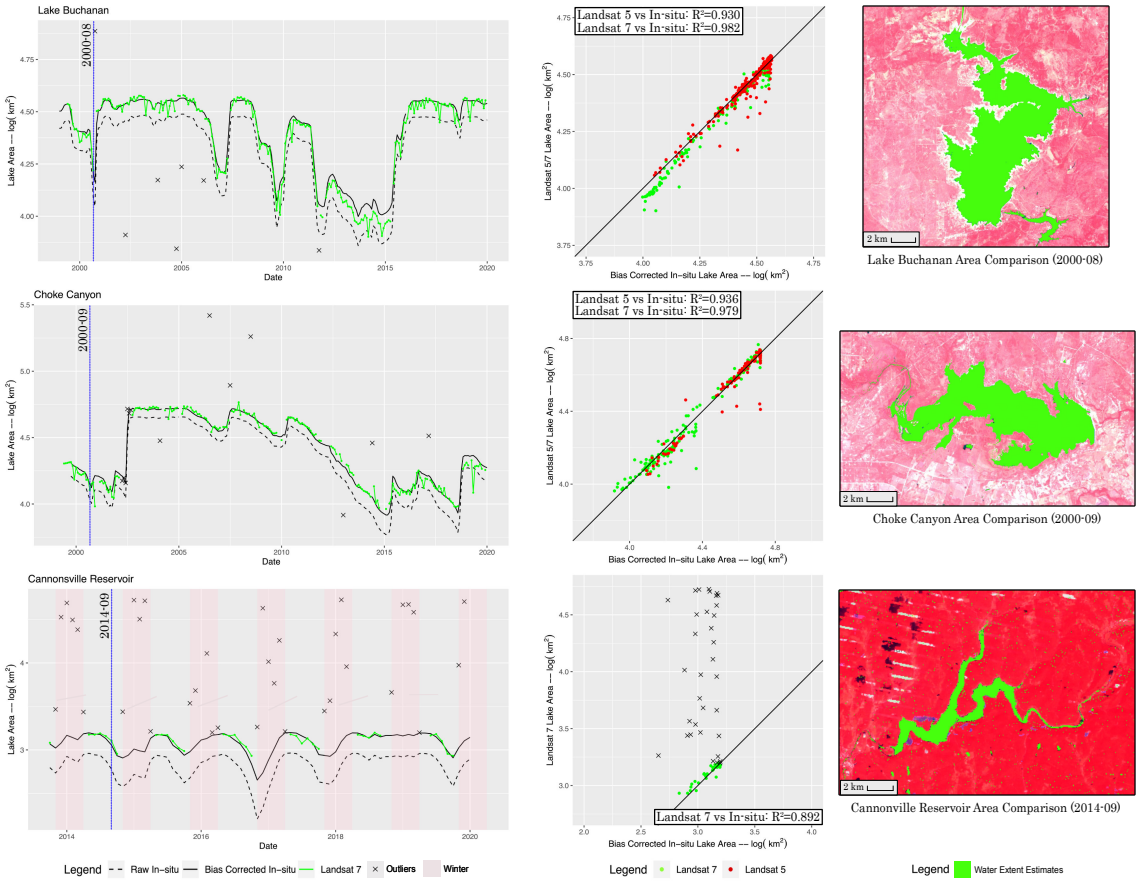

Figure 3. Validation of the Approach. *Left*: Time series representation water extent from in situ observation (green) and Landsat 7 (blue). Automatically removed outliers (crosses) and raw (non-bias corrected) in situ time series (dashed) are also displayed for indicative purpose. Winter (ice-covered) periods are shaded out for Cannonville to illustrate the limitation of the approach. Corresponding observations are also marked as outliers. *Middle* Scatter-plot of Landsat 5 (red) and Landsat 7 (green) water extent estimates against bias-corrected in situ observations. Winter observations are represented as empty symbols for Cannonville. *Left* False color composite of a Landsat 7 image of each lake on a cloud-free day. For all three lakes, the area estimated by the approach (green) matches the open water visible on the image. Yet their areas are substantially larger than that given by in situ observation (see difference between the blue and dashed time series at the vertical line drawn on the *left* panel). This is indicative of biases in the in situ observations.

as crosses for indicative purposes on the time series. Note that the outlier detection algorithm is fully automatic and does not require any user input. Results show a strong correlation between in situ and Landsat-determined lake extents for Lake Buchanan ($R^2 = 0.982$ on normal scale), Choke Canyon ($R^2 = 0.979$ on normal scale) and Cannonsville Reservoir ($R^2 = 0.892$ on normal scale). In all cases, the remote sensing algorithm was able to capture the strong temporal change in water extent of these intensively managed reservoirs. Performance was good for the Cannonsville Reservoir after winter is removed, which illustrates the main limitations of the approach under ice-covered conditions (keeping winter observations decreases the

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

175 coefficient of determination to $R^2 = 0.01$ on normal scale) and small lake sizes (see Sect. 4). These two limitations are further discussed Sect. 4.1.

3.2 Application to data-scarce, rapidly changing lake

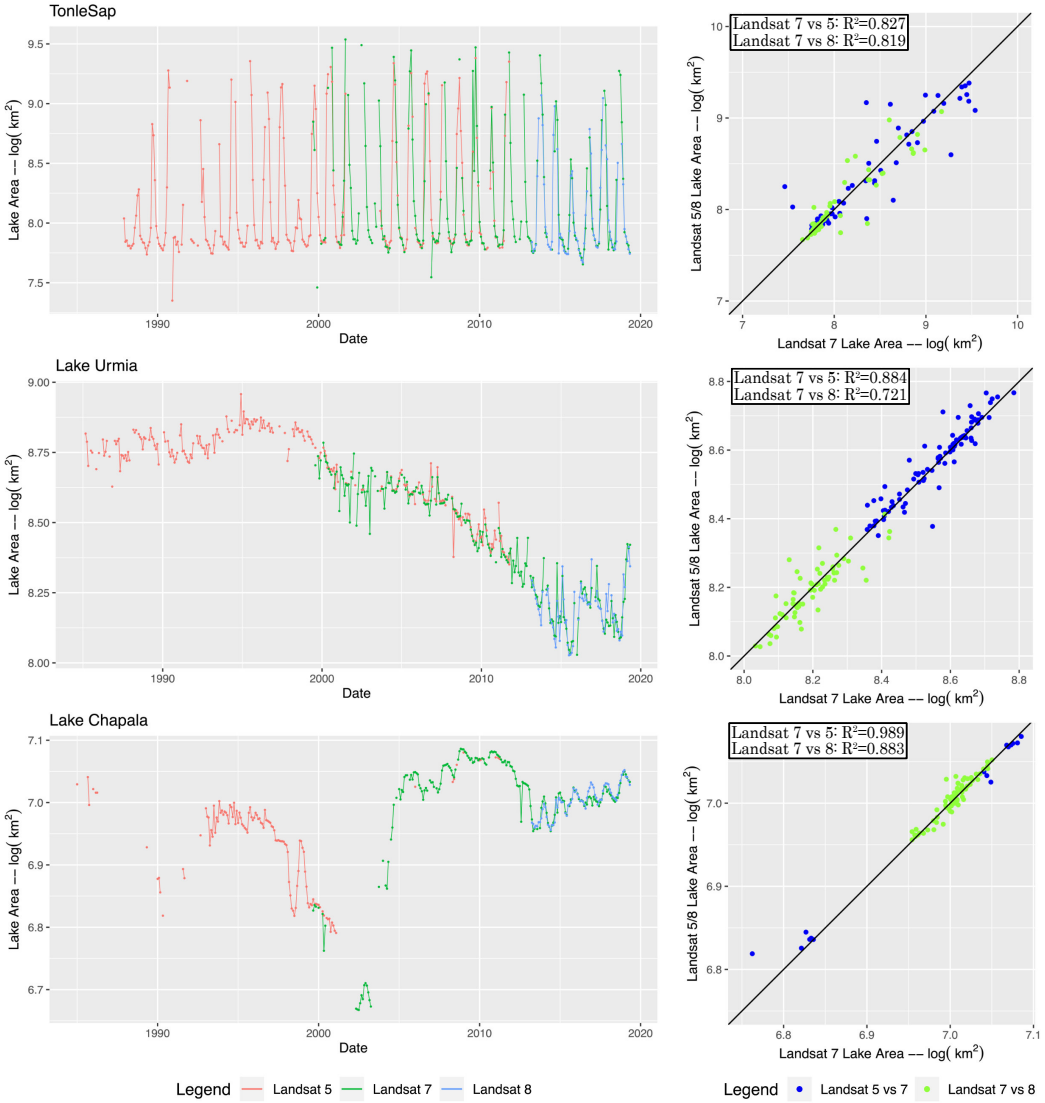

Figure 4. Implementation of the approach on lakes with documented changes. *Left*: Time series of monthly lake extent estimates from Landsat 5 (red), Landsat 7 (green) and Landsat 8 (blue) for Lakes Tonle Sap (*Top*), Urmia (*Right*) and Chapala (*Bottom*). *Right Column*: Scatterplots representing Landsat 8 (green) and Landsat 5 (blue) water extent estimates against Landsat 7 during overlapping periods.

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Application to Lakes Tonle Sap, Urmia and Chapala, where no in situ observations were obtained, shows a high level of agreement across Landsat missions during overlapping periods (Fig. 4) with normal scale R^2 values ranging between 0.83 and 0.99 (or between 0.72 and 0.88) when comparing Landsat 7 and 5 (or 7 and 8), respectively. The approach provides consistent and reliable time series of historic water extent that are consistent with documented historical patterns. The analysis suggests that recent fluctuations in the amplitude of the seasonal inundation cycles of the Tonle Sap, which are critical to maintain its function as a regional biodiversity and food security hotspot. This is consistent with recent modeling simulations that predict decreased seasonal variations owing to flow regime alterations in the Mekong tributary region (Yu et al., 2019; Kummu and Sarkkula, 2008). The dramatic desiccation of Lake Urmia, once among the world's largest freshwater lakes, is also clearly visible in our analysis. Lake extent has declined steadily since the late 1990's to reach a low point in August 2014, which is consistent with existing estimates (AghaKouchak et al., 2015). Similarly, large water fluctuations in Lake Chapala, a strategic and historically overexploited reservoir in Central Mexico (Wester, 2008; Godinez-Madrigal et al., 2019) in the 1990's and early 2000's can be seen in our analysis, along with the effects of the dramatic (albeit controversial (Godinez-Madrigal et al., 2019)) remediation policies that were implemented thereafter to restore lake levels (Wester, 2008).

4 Discussion

A key feature of the water detection algorithm is its reliance on an inundation frequency (IF) image to infer the inundation status of clouded pixels. The IF image is constructed based on the historic probability of each pixel being covered by water under cloud-free conditions, as determined by the unsupervised classification procedure in the first stage of the algorithm. This presumes the existence of two random binary processes – cloud coverage and inundation status – that are (i) stationary, (ii) independent from each other, and (iii) can be reliably identified on Landsat imagery. These three important assumptions potentially constrain the applicability of the method and merit further discussion.

4.1 Assumption 1: Stationarity

A threat to the stationarity requirement may for instance emerge if erosion or sedimentation processes substantially alter the bathymetry of the lake shores. We used a split sample approach to assess this possibility. Two IF images were constructed using the first (1998-2009) and second (2010-2020) half of the available Landsat 7 dataset. The inundation frequencies given by the first and second image were then collected for a random sample of 5000 pixel which IF value was above zero and below one for both images. The sample pixels were then ranked according to their IF value for each images. The stationary assumption implies that the ranks of the pixels do not vary between observation periods: if bathymetry did not change, a pixel that is more often inundated than another given pixel during the 1998-2009 period should still be more often inundated during 2010-2020. Results on Fig. 5 (top) suggest that the effect of bathymetric change on the classification outcome is negligible. To interpret this, note that classification outcomes are only affected by bathymetric changes that concern those pixels that lie within the range of variability of water extent. This excludes pixels that are permanently covered (IF=1), where bathymetry may be most affected by sedimentation processes.

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

4.2 Assumption 2: Independence

A threat to the independence requirement may emerge if the inundation status of a pixel determines its cloud coverage. For instance, fog can be produced by the micro-climatic conditions associated with open water (Koračin et al., 2014). We test this hypothesis by comparing the inundation frequency of pixels during cloudless days (the output of the first stage of the algorithm), with their unconditional inundation frequency (during cloudy and cloudless days). The latter was determined by computing the probability of inundation predicted by the second stage of the algorithm (supervised classification), which includes cloudy days. As before, we sampled 4000 pixels with IF values between (and excluding) 0 and 1 for both images (cloudless and unconditional). We then ranked the pixels according to their IF value for each image. The independence assumption implies that the pixels rank is not affected by its cloud coverage status: a pixel with a higher inundation frequency than another for a subset of observations that had cloudless conditions should also have a higher inundation frequency if the full sample of observations (cloudless and cloudy) is considered. Results, shown in Fig. 5 (bottom), suggests that the ranking of inundation frequency does not depend on cloud coverage. Again, to interpret this, consider that preferential cloud coverage will only affect classification output if it concerns pixels near shores (i.e. where $0 < IF < 1$), whereas lake-induced fog tends to occur over permanently inundated pixels (Koračin et al., 2014).

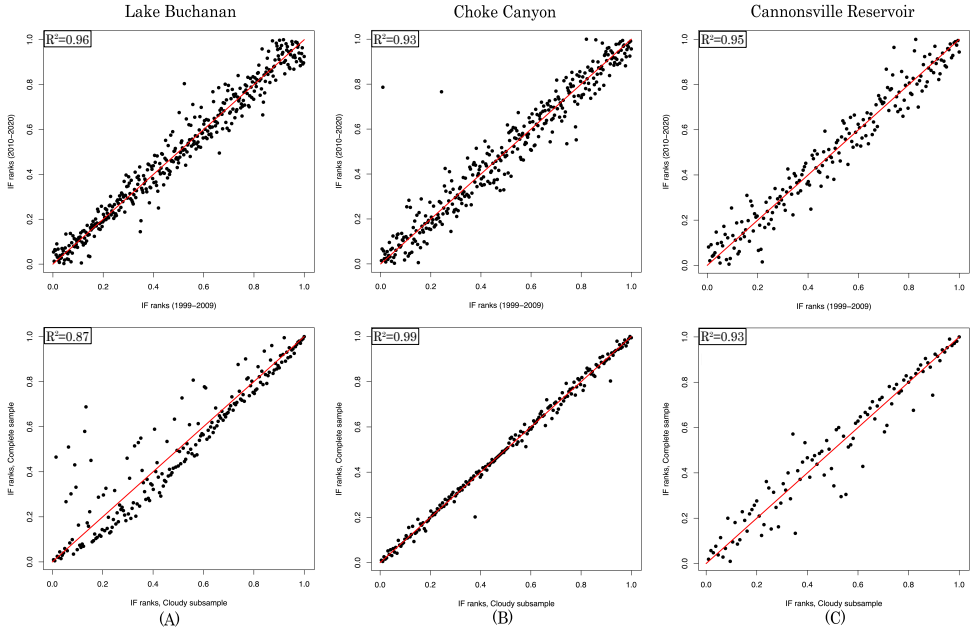

Figure 5. Stationary and independence assumptions. *Top*: Inundation Frequency ranks per pixel estimated under cloudless condition (unsupervised classification, y-axis) plotted against corresponding ranks estimated using the full sample of observations (combined supervised-unsupervised classification, x-axis) for Lake Buchannan (*left*), Choke Canyon Reservoir (*middle*) and Cannonville Reservoir (*right*). *Bottom*: Inundation frequencies ranks per pixel estimated using the first (x-axis) and second (y-axis) half of the Landsat 7 observation period (1999-2019).

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

4.3 Assumption 3: Detection

The procedure used to detect open water under cloudless conditions is widely used and its performance well established (see, e.g., MNDWI (Xu, 2006)). Nonetheless, two important limitations of MNDWI-based water detection may affect the performance of our approach. The first issue concerns the difficulty in distinguishing open water from cloud shadows which have a similar MNDWI reflectance (see Zhang et al., 2015). Cloud shadows are visible on Landsat imagery when cloud cover is fragmented and, if shadows dry land in the vicinity of the lake, may cause an overestimation of the surface area of the lake. This error will be amplified by the supervised classification step of our approach (Stage 2), which will mistakenly associate low IF pixels to inundated conditions. This issue is particularly salient for smaller lakes with surface area on the order of the typical size of cloud fragments. To the extreme, a small lake completely covered by a cloud shadow would lead to a very large overestimation of its surface area. Note that a lake completely covered by clouds would be removed from the analysis in the pre-processing step. While problematic, such large individual deviations in water extent estimations are typically captured by the automatic outlier detection implemented in the post-processing stage.

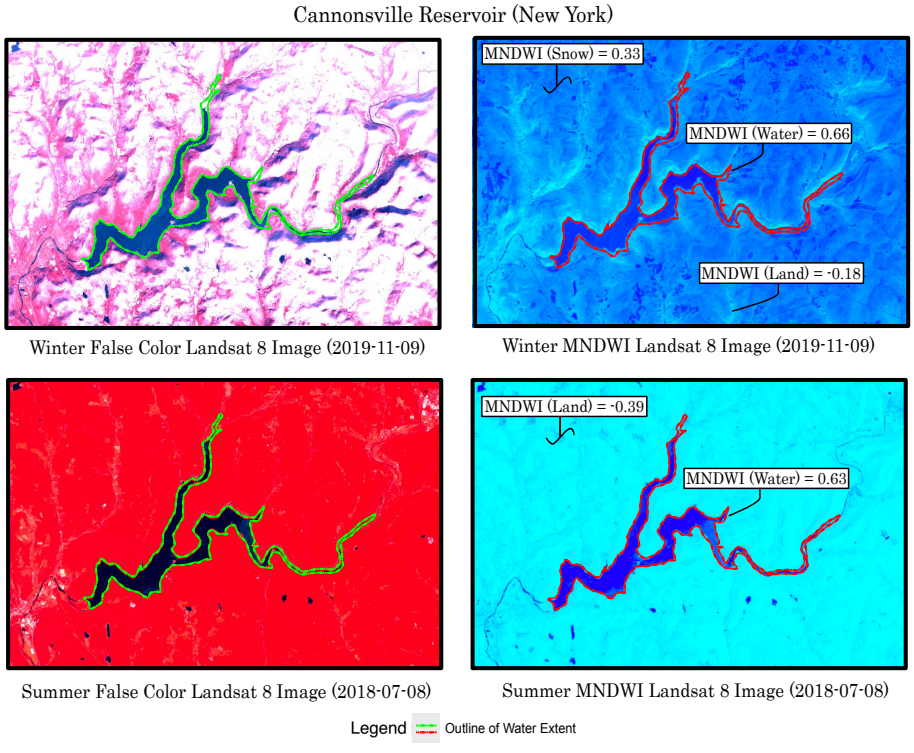

Figure 6. Illustrative example of the effect of snow and ice cover on water detection for the Cannonville reservoir. The contrast between land and water on the MNDWI image (right) is much larger during the summer (bottom) than during the winter, when snow cover (in white on the false color image, top left) has a similar MNDWI value to water.

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

The second issue arises from the fact that snow and ice have a MNDWI reflectance which is much larger than land and closer to that of open water. Under these conditions, a predetermined MNDWI threshold (here the median of k-means classification on individual images) would mistakenly classify snow (or ice) -covered dry land as open water. Similar to cloud shadows, the ensuing overestimation of water cover in the first stage will then be amplified by the supervised classification in the second stage. This phenomenon is visible for Cannonville Reservoir (Fig. 6), where water cover is substantially overestimated during winter months. Because these errors persist throughout the winter season, they are typically not captured by the automatic outlier detection and must be removed manually based on the image date.

5 Conclusion

The level and seasonality of lake water is rapidly changing globally and long-term, high frequency time series observations of lake water extents are essential to ascertain and attribute these changes. This data is rarely collected in situ and challenging to obtain from satellite imagery due to cloud coverage, missing images and the limited temporal coverage of recent remote sensing missions. The proposed method addresses these challenges using a hybrid unsupervised-supervised classification approach that is scalable in that it does not require any local ancillary data. It also seamlessly integrates images from the three most recent Landsat missions (Landsats 5, 7 and 8), which provide consistent high frequency (bi-weekly to monthly) inundation estimates with high spatial resolution (30 m) between the late 1980's and today.

The method was successfully validated against in situ observations for three lakes within the US with rapidly changing water extents. Application to ungauged lakes was demonstrated on a sample of three regionally significant lakes: Tonle Sap (Cambodia), Lake Urmia (Iran), and Lake Chapala (Mexico). For all three lakes, water extent estimates were consistent with data from Landsat missions during overlapping periods, and the observed time series reproduced documented historical patterns. Key assumptions of the approach – stationarity, independance and detection – are discussed, pointing to potential limitations of the approach. In particular, the approach hinges on the fact that relevant information used by the algorithm is most salient for partially inundated pixels near the shores of the lake, whereas potential threats to the underlying assumptions are most salient away from the shores (e.g. sedimentation, fog-on-water). The application of the method is therefore limited for small lakes where shore-to-surface ratio is large.

Despite its limitations, the algorithm stands as a valuable tool to provide reliable observation on the historical extent of rapidly changing lakes, at a temporal frequency and over an observation period that allows for modern statistical approaches (e.g., Müller and Levy, 2019; Penny et al., 2020; Pérez Valentín and Müller, 2020) to be leveraged to understand the underlying causes.

Data availability. Lake level datasets for validation are publicly made available by the United States Geological Survey (https://waterdata.usgs.gov/nwis) and the Texas Water Development Board (https://www.waterdatafortexas.org/reservoirs/statewide)

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Code and data availability. The remote sensing datasets and classification algorithms are permanently available at https://code.earthengine.google.com/bed0dacc981bb883533f4958e80b50be

Author contributions. C.M. and M.F.M designed the research, C.M. conducted the analysis, C.M. and M.F.M wrote the paper

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. We acknowledge financial support from the US National Science Foundation (NSF) under grant ICER 1824951.

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

References

AghaKouchak, A., Norouzi, H., Madani, K., Mirchi, A., Azarderakhsh, M., Nazemi, A., Nasrollahi, N., Farahmand, A., Mehran, A., and Hasanzadeh, E.: Aral Sea syndrome desiccates Lake Urmia: Call for action, Journal of Great Lakes Research, 41, 307–311, https://doi.org/10.1016/J.JGLR.2014.12.007, https://www.sciencedirect.com/science/article/pii/S0380133014002688, 2015.

Avisse, N., Tilmant, A., François Müller, M., and Zhang, H.: Monitoring small reservoirs' storage with satellite remote sensing in inaccessible areas, in: Hydrol. Earth Syst. Sci, vol. 21, pp. 6445–6459, https://doi.org/10.5194/hess-21-6445-2017, https://doi.org/10.5194/hess-21-6445-2017, 2017.

Bioresita, F., Puissant, A., Stumpf, A., and Malet, J.-P.: A method for automatic and rapid mapping of water surfaces from sentinel-1 imagery, Remote Sensing, 10, 217, 2018.

Busker, T., Roo, A., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., Pekel, J.-F., and Cottam, A.: A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry, Hydrology and Earth System Sciences Discussions, pp. 1–32, https://doi.org/10.5194/hess-2018-21, 2018.

Busker, T., Roo, A. d., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., Pekel, J.-F., and Cottam, A.: A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry, Hydrology and Earth System Sciences, 23, 669–690, 2019.

Chen, C. and Liu, L.-M.: Joint estimation of model parameters and outlier effects in time series, Journal of the American Statistical Association, 88, 284–297, 1993.

Chen, P.-Y., Srinivasan, R., Fedosejevs, G., and Kiniry, J.: Evaluating different NDVI composite techniques using NOAA-14 AVHRR data, International Journal of Remote Sensing, 24, 3403–3412, 2003.

Godinez-Madrigal, J., Van Cauwenbergh, N., and van der Zaag, P.: Production of competing water knowledge in the face of water crises: Revisiting the IWRM success story of the Lerma-Chapala Basin, Mexico, Geoforum, 103, 3–15, 2019.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote sensing of Environment, 202, 18–27, 2017.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., et al.: Global water resources affected by human interventions and climate change, Proceedings of the National Academy of Sciences, 111, 3251–3256, 2014.

Islam, A., Bala, S., and Haque, M.: Flood inundation map of Bangladesh using MODIS time-series images, Journal of Flood Risk Management, 3, 210–222, 2010.

Jain, A. K.: Data clustering: 50 years beyond K-means, Pattern recognition letters, 31, 651–666, 2010.

Koračin, D., Dorman, C. E., Lewis, J. M., Hudson, J. G., Wilcox, E. M., and Torregrosa, A.: Marine fog: A review, Atmospheric Research, 143, 142–175, 2014.

Kummu, M. and Sarkkula, J.: Impact of the Mekong River flow alteration on the Tonle Sap flood pulse, AMBIO: A Journal of the Human Environment, 37, 185–192, 2008.

Kummu, M., Ward, P. J., de Moel, H., and Varis, O.: Is physical water scarcity a new phenomenon? Global assessment of water shortage over the last two millennia, Environmental Research Letters, 5, 034 006, 2010.

Lu, D. and Weng, Q.: A survey of image classification methods and techniques for improving classification performance, International Journal of Remote Sensing, 28, 823–870, https://doi.org/10.1080/01431160600746456, https://www.tandfonline.com/doi/full/10.1080/01431160600746456, 2007.

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Mercier, F., Cazenave, A., and Maheu, C.: Interannual lake level fluctuations (1993–1999) in Africa from Topex/Poseidon: connections with ocean–atmosphere interactions over the Indian Ocean, Global and Planetary Change, 32, 141–163, 2002.

Micklin, P.: The Aral sea disaster, Annu. Rev. Earth Planet. Sci., 35, 47–72, 2007.

Müller, M. F. and Levy, M. C.: Complementary vantage points: Integrating hydrology and economics for sociohydrologic knowledge generation, Water Resources Research, 55, 2549–2571, 2019.

Müller, M. F., Yoon, J., Gorelick, S. M., Avisse, N., and Tilmant, A.: Impact of the Syrian refugee crisis on land use and transboundary freshwater resources, Proceedings of the national academy of sciences, 113, 14 932–14 937, 2016.

Pelletier, C., Valero, S., Inglada, J., Champion, N., and Dedieu, G.: Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas, Remote Sensing of Environment, 187, 156–168, 2016.

Penny, G., Mondal, M. S., Biswas, S., Bolster, D., Tank, J. L., and Müller, M. F.: Using natural experiments and counterfactuals for causal assessment: River salinity and the Ganges water agreement, 2020.

Pérez Valentín, J. M. and Müller, M. F.: Impact of Hurricane Maria on beach erosion in Puerto Rico: remote sensing and causal inference, Geophysical Research Letters, p. e2020GL087306, 2020.

Raje, D. and Mujumdar, P.: Reservoir performance under uncertainty in hydrologic impacts of climate change, Advances in Water Resources, 33, 312–326, 2010.

Schowengerdt, R. A.: Remote sensing: models and methods for image processing, Elsevier, 2006.

Sheffield, J., Wood, E., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., and Verbist, K.: Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions, Water Resources Research, 54, 9724–9758, 2018.

Shen, H., Li, X., Cheng, Q., Zeng, C., Yang, G., Li, H., and Zhang, L.: Missing information reconstruction of remote sensing data: A technical review, IEEE Geoscience and Remote Sensing Magazine, 3, 61–85, 2015.

Survey, U. S. G. and : Landsat—Earth observation satellites, Tech. rep., Reston, VA, https://doi.org/10.3133/fs20153081, http://pubs.er.usgs.gov/publication/fs20153081, 2015.

Van Den Hoek, J., Getirana, A., Jung, H. C., Okeowo, M. A., and Lee, H.: Monitoring Reservoir Drought Dynamics with Landsat and Radar/Lidar Altimetry Time Series in Persistently Cloudy Eastern Brazil, Remote Sensing, 11, 827, 2019.

Wester, P.: Shedding the waters: institutional change and water control in the Lerma-Chapala Basin, Mexico, 2008.

Wine, M. L., Rimmer, A., and Laronne, J. B.: Agriculture, diversions, and drought shrinking Galilee Sea, Science of The Total Environment, 651, 70–83, https://doi.org/10.1016/J.SCITOTENV.2018.09.058, https://www.sciencedirect.com/science/article/pii/S0048969718334946, 2019.

Xu, H.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, International journal of remote sensing, 27, 3025–3033, 2006.

Yale, M. M., Sandwell, D. T., and Herring, A. T.: What are the limitations of satellite altimetry?, The leading edge, 17, 73–76, 1998.

Yan, K., Di Baldassarre, G., Solomatine, D. P., and Schumann, G. J.-P.: A review of low-cost space-borne data for flood modelling: topography, flood extent and water level, Hydrological processes, 29, 3368–3387, 2015.

Yan, Y.-E., Ouyang, Z.-T., Guo, H.-Q., Jin, S.-S., and Zhao, B.: Detecting the spatiotemporal changes of tidal flood in the estuarine wetland by using MODIS time series data, Journal of Hydrology, 384, 156–163, 2010.

Yu, W., Kim, Y., Lee, D., and Lee, G.: Hydrological assessment of basin development scenarios: Impacts on the Tonle Sap Lake in Cambodia, Quaternary International, 503, 115–127, 2019.

https://doi.org/10.5194/hess-2020-198
Preprint. Discussion started: 30 June 2020
© Author(s) 2020. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

345 Zhang, Q., Li, B., Thau, D., and Moore, R.: Building a better urban picture: Combining day and night remote sensing imagery, Remote Sensing, 7, 11 887–11 913, 2015.

Zhu, Z. and Woodcock, C. E.: Object-based cloud and cloud shadow detection in Landsat imagery, Remote sensing of environment, 118, 83–94, 2012.