


Article

Big Data and Multiple Methods for Mapping Small Reservoirs: Comparing Accuracies for Applications in Agricultural Landscapes

Sarah K. Jones ^{1,2,*}, Alexander K. Fremier ³ , Fabrice A. DeClerck ², David Smedley ¹,
Aline Ortega Pieck ³ and Mark Mulligan ¹

¹ Department of Geography, King's College London, Strand, London WC2R 2LS, UK; david.smedley@kcl.ac.uk (D.S.); mark.mulligan@kcl.ac.uk (M.M.)

² Bioversity International, Parc Scientifique Agropolis II, 34397 Montpellier, France; f.declerck@cgiar.org

³ School of the Environment, Washington State University, Pullman, WA 99164, USA; alex.fremier@wsu.edu (A.K.F.); alineortega@gmail.com (A.O.P.)

* Correspondence: s.jones@cgiar.org; Tel.: +33-648-813-912

Received: 29 August 2017; Accepted: 7 December 2017; Published: 13 December 2017

Abstract: Whether or not reservoirs contain water throughout the dry season is critical to avoiding late season crop failure in seasonally-arid agricultural landscapes. Locations, volumes, and temporal dynamics, particularly of small (<1 Mm³) reservoirs are poorly documented globally, thus making it difficult to identify geographic and intra-annual gaps in reservoir water availability. Yet, small reservoirs are the most vulnerable to drying out and often service the poorest of farmers. Using the transboundary Volta River Basin (~413,000 sq km) in West Africa as a case study, we present a novel method to map reservoirs and quantify the uncertainty of Landsat derived reservoir area estimates, which can be readily applied anywhere in the globe. We applied our method to compare the accuracy of reservoir areas that are derived from the Global Surface Water Monthly Water History (GSW) dataset to those that are derived when surface water is classified on Landsat 8 OLI imagery using the Normalised Difference Water Index (NDWI), Modified NDWI with band 6 (MNDWI1), and Modified NDWI with band 7 (MNDWI2). We quantified how the areal accuracies of reservoir size estimates vary with the water classification method, reservoir properties, and environmental context, and assessed the options and limitations of using uncertain reservoir area estimates to monitor reservoir dynamics in an agricultural context. Results show that reservoir area estimates that are derived from the GSW data are 19% less accurate for our study site than MNDWI1 derived estimates, for a sample of 272 reservoir extents of 0.09 to 72 ha. The accuracy of Landsat-derived estimates improves with reservoir size and perimeter-area ratio, while accuracy may decline as surface vegetation increases. We show that GSW derived reservoir area estimates can provide an upper limit for current reservoir capacity and seasonal dynamics of larger reservoirs. Data gaps and uncertainties make GSW derived reservoir extents unsuitable for monitoring reservoirs that are smaller than 5.1 ha (holding ~49,759 m³), which constitute 674 (56%) reservoirs in the Volta basin, or monitoring seasonal fluctuations of most small reservoirs, limiting its utility for agricultural planning. This study is one of the first to test the utility and limitations of the newly available GSW dataset and provides guidance on the conditions under which this, and other Landsat-based surface water maps, can be reliably used to monitor reservoir resources.

Keywords: small reservoirs; surface water dynamics; sustainable development; Google Earth Engine; West Africa

1. Introduction

Freshwater scarcity is a major constraint to food production in agricultural regions of the world with variable intra- and inter-annual rainfall patterns and poor water storage infrastructures [1]. In sub-Saharan Africa, seasonal rainfall fluctuations and shortages cause up to 53% crop failure in smallholder farming systems [2]. Monitoring water resource availability in such areas is critical to limit food shortages and the subsequent sometimes far-reaching social, political, and economic implications [3].

Small and large reservoirs are a common development investment to avert or reduce water shortages and boost production in seasonally dry, agriculture-dependent regions [4]. Reservoirs capture and store runoff to provide farmers with a source of freshwater during dry spells and annual dry seasons, increasing the viable extent, productivity, and resilience of cropping, fishery, and livestock production systems [5]. “Small” reservoirs are engineered surface water bodies with a capacity of less than one million m³ (Mm³) [6]. Small reservoirs are less expensive to construct than larger ones, and are therefore perceived as low cost, high return development devices [4]. Decades of investments to construct small reservoirs remain largely undocumented at the river basin level and above [7–10], making their impact on human well-being or environmental outcomes impossible to accurately assess. Information on reservoir locations, volumes, and seasonality (water presence-absence) can guide policies to allocate and manage reservoir water sustainably and avert agricultural water and associated food shortages. Access to this information would therefore support donors, governments, and NGOs in efforts to better understand the value of reservoirs and target reservoir investments and maintenance to achieve local to global sustainable development objectives.

Information on small reservoirs is challenging to compile and to keep updated at the national, regional, or global level [11]. Ground-based assessments of small reservoir locations, capacities, and seasonal volumes are time-consuming to conduct because of the spatial dispersion of these reservoirs and decentralized decision making regarding reservoir investments and maintenance [12]. Many previous studies have characterized inland surface water resources using free satellite imagery, such as from Landsat [13–17] and the Moderate Resolution Imaging Spectroradiometer (MODIS) [18–22]. Where the spatial and temporal resolution are sufficiently high, remotely sensed imagery provides a practical approach to small water body mapping and monitoring [23,24]. The highest resolution, freely available imagery that is collected on an intra-annual timestep over long time scales comes from the Landsat satellite series [25]. Instruments on Landsat satellites provide near-complete global coverage of multispectral imagery at 30 m resolution and 16-day time steps, from 1982 to present. Specifically, Landsat 4 Thematic Mapper provides imagery from 1982 to 1993; Landsat 5 Thematic Mapper from 1984 to 2012; Landsat 7 Enhanced Thematic Mapper from 1999 to present; and, Landsat 8 Operational Land Imager (OLI) from February 2013 to present, providing 35 years of almost continuous data [26]. Accurately mapping waterbodies from Landsat imagery is a non-trivial task, since water can be misclassified as urban areas, shadows, and other objects with similar spectral signatures, or these non-water objects can be falsely classified as water [27]. Sediment or vegetation in water—particularly common in West Africa [16]—alters the spectral signature of water, while cloud and dust particles in the atmosphere obscure or distort information about where land is water-covered [28].

Researchers have successfully mapped water body locations and extents from Landsat imagery for several decades through the use of spectral indices [23,24,29]. Spectral indices are created by calculating the difference, ratio, or normalised difference of two multispectral image bands and identifying the threshold that enhances the reflectance of wavelengths for objects of interest, such as water [28]. The most effective index or threshold identified to map water is rarely the same between studies [29]. More recently, Pekel J.-F. et al. [16] used multi-spectral analysis to classify water and non-water over all of the images in the Landsat database from 1984 to 2015, resulting in pixel-level measures of water occurrence at monthly time steps over the entire globe. Their Global Surface Water Monthly Water History (GSW) dataset represents an attempt to establish a globally applicable method for water detection from Landsat imagery. Uncertainties in GSW and other surface water maps are

generally reported in terms of pixel-level classification accuracy [30] rather than the impact on practical applications. However, the level of uncertainty that is acceptable will depend on the end-use, and knowledge of this uncertainty may help to ensure its effective use in policy [31]. For example, water managers that unknowingly use incorrect information on where a small reservoir exists or when reservoirs run dry may make decisions on water allocation that have serious consequences for the agricultural sector and farmers who rely on reservoirs for their livelihoods. Intra-annual surface water maps derived from Landsat satellite imagery, and particularly the globally available GSW dataset, may be useful for monitoring spatial and seasonal dynamics of small reservoirs in agricultural landscapes opening up opportunities for improved global reservoir data, but their accuracy for this end-use is currently untested.

The primary objective of this paper was to compare a range of methods for rapid and low-cost monitoring of reservoirs, and to establish levels of accuracy in reservoir characterization using these methods, as reservoir properties and environmental conditions vary. For this purpose, we developed a new method for mapping reservoir extents rapidly across large spatio-temporal extents using free, globally available datasets and tools. We compare the effect of accuracy, which is the level of uncertainty in reservoir surface area and equivalent volume estimates, on information about temporal and spatial reservoir water availability to highlight where improved accuracies may be important for agricultural applications. Using the Volta Basin in West Africa as a case study, we focus on three specific questions: (i) What is the accuracy of reservoir areal extents digitized manually from high resolution Google Earth imagery compared to those derived computationally from GSW or from commonly used spectral water indices applied to Landsat 8 OLI imagery? (ii) How does the accuracy of Landsat-based reservoir area estimates vary with environmental factors? (iii) What information on reservoir size and seasonality can be reliably determined from the GSW and what cannot? Our study represents the first attempt to test the limits of the GSW dataset for monitoring reservoirs of varied size and across a range of environmental conditions in a West African context. Establishing reliable end-uses for the GSW is particularly important given that this dataset is global, publically available, and easy to use. The global applicability of our approach makes it useful to a wide range of stakeholders interested in surface water resources or low-cost environmental monitoring.

We structure the remainder of the paper as follows. Section 2.1 presents the study site. Sections 2.2 and 2.3 describe the input remote sensing datasets and processing techniques that were used to identify reservoirs and prepare surface water maps. Section 2.4 explains how reservoir extents were extracted, and Section 2.5 describes the validation data used to determine the accuracy of these reservoir extents. Sections 2.6 and 2.7 describes the comparison of accuracies and analysis of covariance with environmental factors, while Section 2.8 explains how we used the GSW-derived reservoir extent data to illustrate potential policy-relevant applications. Sections 3.1 and 3.2 reports the accuracy of reservoir area estimates across water classification methods and how accuracies vary with environmental conditions. Section 3.3 presents reservoir volume and seasonality derived from GSW where sufficient data were available over the 1200 Volta basin reservoirs. Section 4 critically analyses the results and the proposed approach for reservoir monitoring. In Section 5, we conclude by summarising the utility of our approach for agricultural applications and in a broader context.

2. Materials and Methods

2.1. Study Site

The 413,000 km² Volta basin, which drains parts of Benin, Burkina Faso, Cote d'Ivoire, Ghana, Mali, and Togo, has a mean annual precipitation level relatively high when compared to other major global basins at 953 mm/year [32]. This rainfall is unevenly distributed from north to south and seasonally skewed across the basin, with parts of the south receiving over 1700 mm/year compared to under 400 mm/year in the northern extremes (Figure 1) and rain typically falling between May and

September in the south and June and September in the north, based on WorldClim data [33]. Between 11% (Ghana) and 64% (Burkina Faso), of the population are estimated to live in severe poverty [34].

The Volta basin is a pertinent region to use as a case study given the absence of consistent information on reservoir locations and functionality between and within the basin's six countries [35], and the livelihood dependencies on the water stored in the basin's small reservoirs which open up opportunities for dry season food production and supplemental irrigation to avert losses during dry spells. [36]. A lack of accessible data make it impossible to robustly monitor the basin's water resources, a problem that is emblematic of Africa's water resources [3]. Identifying transferable, practical methods for water resource monitoring in the Volta basin and Africa more generally is important to enable better targeting of interventions to manage water resources under the continent's rapid population growth [37] and shift towards more resource-intensive diets [38].

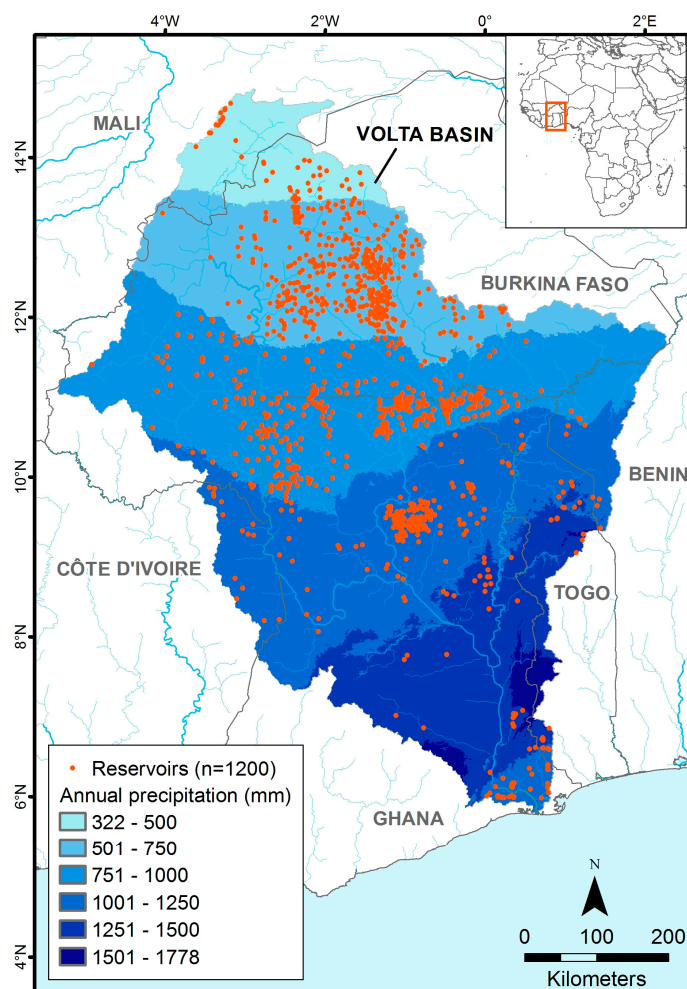


Figure 1. Annual precipitation over the Volta basin study site based on 1980–2010 WorldClim data, with reservoirs identified in this study from Google Earth imagery.

2.2. Reservoir Locations

We made use of the most recent imagery hosted in Google Earth in September 2015 and a 100 km × 100 km grid to systematically, manually identify, and map all of the visible engineered reservoirs across the Volta basin, generating a georeferenced point dataset of existing reservoir locations (Figure 1). For each of the 1200 identified reservoirs, we placed a point inside the reservoir boundary near the dam wall and where reservoir water was consistently present in months that the reservoir contained water, as shown on imagery within Google Earth's historical imagery collection.

2.3. Landsat-Derived Surface Water Maps

We used Google Earth Engine (GEE) to source, generate, and analyse surface water maps that were derived from Landsat imagery. GEE is an online coding environment enabling relatively rapid, server-based analysis of large spatial datasets [39].

GEE provides access to the EC JRC/Google Monthly Water History V1.0 dataset [16], which contains maps of Global Surface Water created from decadal analysis of Landsat 4, 5, and 7 imagery and contains 30 m × 30 m pixel-level measures of water presence-absence on a monthly time step from March 1984 to October 2015. We worked in GEE to create additional monthly surface water maps from three spectral indices applied to Landsat 8 OLI imagery, for use in subsequent analyses. We sourced 1291 Landsat 8 OLI images that were acquired over the Volta basin (Landsat paths 192 to 197; rows 050 to 056) between 1 May 2013 and 31 October 2015, corresponding to the earliest complete month of data available in GEE and the temporal limit of GSW data. We used imagery that was pre-processed by USGS to Surface Reflectance and for each image, computed the Normalised Difference Water Index (NDWI) [40], Modified NDWI [27] using band 6 (referred to here as MNDWI1), and using band 7 (MNDWI2), indices that are commonly applied in peer-reviewed literature for surface water mapping [14,41–43]. The relevant bands in Landsat 8 OLI imagery used to compute the three water indices are:

$$NDWI = \frac{Band\ 3 - Band\ 5}{Band\ 3 + Band\ 5} \quad (1)$$

$$MNDWI1 = \frac{Band\ 3 - Band\ 6}{Band\ 3 + Band\ 6} \quad (2)$$

$$MNDWI2 = \frac{Band\ 3 - Band\ 7}{Band\ 3 + Band\ 7} \quad (3)$$

The OLI on Landsat 8 collects data in slightly different bandwidths to that collected by sensors on earlier Landsat satellites [44]. This can lead to substantially different reflectance values [45], and therefore spectral index outputs [46–48] between these two sensor groups. Limiting this study to Landsat 8 OLI imagery, rather than including images from several sensors in the Landsat satellite series, allowed for a simpler analysis that ensured consistency in spectral index values over water and non-water features.

We classified pixels on Landsat 8 OLI spectral indices as water, non-water or non-valid (masked), consistent with GSW. Non-valid pixels correspond to those classified as cloud in the Landsat ‘CFmask’ layer [49]. To separate water from non-water pixels, we followed [29] who recommend testing several indices and thresholds to identify the index and class boundaries that are most effective for the images and the area of interest. We computed surface water maps using a “0” threshold, and 0.1 increments either side of this up to +/−0.5 (i.e., −0.5, −0.4, −0.3, . . . , 0.5), for each index.

2.4. Landsat-Derived Reservoir Area and Volume Estimates

To derive reservoir area estimates from surface water maps, we used the `connectedPixelCount` function in GEE [50] to extract a count of connected pixels classified as water by the GSW, NDWI, MNDWI1, and MNDWI2 over each of the 1200 identified reservoirs. The `connectedPixelCount` algorithm identifies adjacent pixels of the same value that share an edge, termed “4-way” connected, or adjacent pixels which share an edge or a corner, termed “8-way” connected. We used 4-way rather than 8-way connections to reduce the possibility of stretches of river being included in the `connectedPixelCount`. The trade-off in this approach is under-estimates in area for reservoirs that have an irregular edge.

Reservoir water volumes can be estimated from reservoir extents by determining the empirical relationship between these two variables for a given reservoir [51]. In this study, volume equivalents were computed using the empirical method for relating reservoir surface area to volume derived by [23], as per Equation (4). Liebe and colleagues [23] carried out bathymetric surveys at 41 small

reservoirs in Upper-East Ghana, part of the Volta basin, and found that the following expression could explain 97.5% of observed variance between Landsat-derived surface areas of between 1 and 35 ha and measured reservoir volumes:

$$Volume = 0.00857 \times Area^{1.4367} \quad (4)$$

The Upper East region of Ghana where Liebe's study focused has a mean of 1.1% slopes, similar to the basin-wide mean of 0.9% slopes (USGS 15s elevation data).

2.5. Validation Data

To validate the Landsat-derived reservoir area estimates we used a dataset of reservoir extents that were digitised manually from Google Earth imagery. Google Earth provides high resolution (<15 m) images from multiple sources with most images in the collection are sourced from Digital Globe's satellites (~2 m resolution). We derived the validation dataset by, first, randomly selecting 250 reservoirs from the 1200 identified in this study. Second, we digitised 347 reservoir extents corresponding to every date that imagery were available in the Google Earth historical imagery collection over the 250 reservoirs within the period May 2013 to October 2015. Google Earth images covering the entire reservoir extent were available for more than one month during this period at some reservoirs, and were not available for any month at other reservoirs. Of the 347 digitised reservoir surface extents, 75 were excluded from subsequent analyses because: (i) they exceeded the neighbourhood search area in GEE's `connectedPixelCount` function, which is limited to 1024 pixels (92 ha); or (ii) they were smaller than 0.09 ha, equivalent to one Landsat pixel; or (iii) there were masked pixels in the underlying Landsat imagery, and therefore no Landsat-derived extent estimates against which to compare the validation data. Of the 250 randomly selected reservoirs, no suitable imagery were available over 48 reservoirs, while all of the digitized surface extents associated with a further 31 reservoirs were part of the 75 extents excluded from the validation dataset for one of the aforementioned reasons. The final validation dataset therefore contained 272 reservoir extents, from imagery acquired in different months across three hydrological seasons over 171 reservoirs. These validation extents ranged from 0.09 ha to 72.4 ha with a mean of 7.1 ha and median of 2.3 ha (lower QR: 0.8, upper QR: 7.1 ha), dispersed spatially and seasonally across the basin (Appendix A Figure A1). As expected, the validation data were skewed towards dry months (October through March for most of the basin) when cloud-free images are more likely to be available in the Google Earth historical imagery collection. Reservoirs are smaller in dry months as water levels recede, which partly explain the high proportion of small reservoirs in the validation dataset. Because drying patterns can also depend on reservoir depth and catchment size, we calculated the catchment area for each reservoir and confirmed that the distribution of our validation dataset was representative of the distribution of basin-wide reservoir catchment sizes using the WaterWorld [52] zones of interest tool.

2.6. Accuracy Assessment

To assess the accuracy of reservoir area estimates derived from each surface water map (GSW, NDWI, MNDWI1, and MNDWI2), we compared the Landsat-derived reservoir areal extents and equivalent volumes to those in the validation dataset for corresponding image dates. We used the difference to compute the mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) of Landsat-derived reservoir area and volume estimates. We identified the optimal threshold for classifying reservoir water using NDWI, MNDWI1, or MNDWI2 as that which provides reservoir area estimates with the lowest mean area percentage error.

2.7. Analysis of Environmental Covariates

Accurate mapping of reservoir extents from Landsat imagery can be hindered by the methodological approach as well as environmental factors. Research on environmental sources of error

in water classifications from Landsat data shows that green-brown water, arising for example because of suspended sediment or high chlorophyll content, reduces classification accuracy [53]. Our validation data were based on images from multiple sources available in Google Earth (e.g., WorldView, IKONOS, GeoEye, SPOT), making it challenging to visibly determine water colours of reservoirs in our validation dataset in a consistent manner. However, the Normalised Difference Vegetation Index (NDVI) [54] is sensitive to fractional green vegetation cover [55] and may be a suitable indicator of surface or sub-surface vegetation. The geometry of the reservoir has also been identified as an important factor, with errors in water classification being higher on small or narrow reservoirs, or those with long perimeters [53], since this increases the number of mixed water and non-water pixels, which are more susceptible to misclassification. We hypothesize that seasonal rainfall patterns may also introduce errors in reservoir extent analysis on a monthly time-step, since rainfall events can significantly alter reservoir extents overnight, increasing discrepancies between area estimates derived from Landsat and validation data, while cloud-cover during wetter periods of the year can obscure water pixels and lead to underestimates in reservoir area.

We ran a random forest regression tree analysis in R for each of the four methods that were used to generate reservoir area estimates from Landsat data, corresponding to GSW and the optimal threshold identified (see Section 3.1) from applying NDWI, MNDWI1, and MNDWI2. We used percentage errors in reservoir area estimates as the dependent variable, and indicators of reservoir surface vegetation, reservoir geometry, and rainfall patterns as independent variables. In particular, as an indicator of reservoir surface vegetation, we used the mean NDVI from all pixels intersecting the reservoir extent, calculated in GEE from Landsat 8 OLI Surface Reflectance images acquired in the month each reservoir area estimate is made. We used reservoir area in hectares and reservoir perimeter-area ratio as recorded in our validation dataset, as indicators of reservoir size and shape. We used month and latitude as indicators of rainfall patterns. We selected the Random Forest approach [56] as a statistical method for checking the relationship between a dependent and multiple independent variables, which accepts categorical and continuous data as well as correlated variables (such as reservoir size and perimeter-area ratio) as inputs. Random forest regression works by splitting the dataset into 63% test (bagged) data and 37% validation (out of bag) data, and then constructing multiple trees from random samples of the bagged data, such that the tree nodes represent decreasingly good predictors of the response variable [57]. Residual errors for each estimate are computed by comparing the predicted response against actual response for out of bag data. The trees are combined into a single tree whose nodes are ordered according to the importance of each variable in predicting the response. The “Importance” is a measure of prediction error that is divided by the standard error.

Results from the Random Forest were used to compare which variables were associated with errors in areal estimates, irrespective of the water classification method. In addition we assessed variation in MAPE for validation data stratified by 0.1 percentiles (ratio data) or classes (categorical data) for each factor.

2.8. Data Applications in Agricultural Landscapes

We used our GEE approach to extract area estimates for the 1200 identified Volta basin reservoirs over the period 1984 to 2015 from GSW data, to test what information on reservoir volumes and seasonality could be determined given the limitations of data coverage and estimation errors.

We identified non-valid estimates, corresponding to months where pixels were masked or where no information was available in the underlying GSW data, and used these to assess the inter-annual, intra-annual and spatial availability of monthly reservoir area estimates. This information is used to evaluate whether there is sufficient data to obtain information on annual and intra-annual fluctuations in reservoir volumes and thus seasonality, which is useful for agricultural planning.

Next, we estimated the current maximum capacity for 1117 of the 1200 Volta basin reservoirs by identifying the largest extent recorded at these reservoirs between 1984 and 2015. We used one of the GSW derivative layers, the GSW Maximum Extent (GSW-MX), to get the maximum extent for the

remaining 83 reservoirs whose maxima derived from GSW were \geq (92 ha—RMSE of area estimates), and therefore cannot be measured using our GEE approach, which is restricted to extents of less than 1024 pixels (~92 ha). The GSW-MX dataset highlights all of the pixels that have contained water at any time between 1984 and 2015. It was not available through GEE at the time of this analysis and so we extracted estimates from the GSW-MX in a desktop GIS. We assumed that reservoirs where no water was ever recorded are smaller than the minimum areal unit that can be reliably mapped using GSW data. Based on their maximums, we classified reservoirs as “Small” ($<1 \text{ Mm}^3$), “Large” ($\geq 1 \text{ Mm}^3$ and $<100 \text{ Mm}^3$), “Very large” ($\geq 100 \text{ Mm}^3$) or “Unknown” (no water identified in any month).

Finally, we analysed monthly fluctuations in area estimates for reservoirs where GSW-derived reservoir extents were available throughout the dry season in 2014–2015, the most recent hydrological year of data available, to distinguish the reservoirs that are perennial, ephemeral with 6–11 months water, and ephemeral with <6 months water within estimate uncertainties. June, July, and August are rainfall months across the entire basin, and likely due to cloud cover, reservoir extent data were often missing over these months. Reservoirs with no valid observations for June, July and August were assumed to contain water during these months. Reservoirs with data missing for any month outside this period were excluded from the analysis, as were reservoirs with no water ever recorded during the 32 years covered by the dataset. The monthly rate of water loss can be obtained from GSW-derived reservoir area estimates by subtracting the annual minima from the maxima, and dividing this by the number of months between the two extremes. Where the monthly water loss was less than the uncertainty contained in the reservoir area estimates, as estimated by the RMSE, we assumed the estimates cannot be reliably used to monitor the loss in reservoir water through the year and classified its seasonality as “uncertain”.

GEE codes that were used in our analyses are available here: <https://earthengine.googleusercontent.com/SurfaceWaterMapping>. R codes used to calculate estimation errors and the run random forest analyses are available on request.

3. Results

3.1. Accuracy of Reservoir Area Estimates

Comparing reservoir area estimates from GSW, NDWI, MDNWI1, and MNDWI2 against those in our validation dataset, we find that estimates varied sustainably in accuracy across methods. Careful selection of the threshold for water classification is critical for minimizing percentage errors in reservoir area estimates derived from NDWI, MNDWI1, and MNDWI2. The lowest MAPE was achieved using a threshold of -0.2 on both the NDWI and MNDWI1, and a 0 threshold on MNDWI2 (see Appendix C Table A1 for comparative accuracies of NDWI, MNDWI1, and MNDWI2 across all thresholds). We subsequently refer to results from NDWI, MNDWI1 and MNDWI2 corresponding to these optimal thresholds.

Comparison of absolute errors in reservoir area estimates indicates MNDWI1 produced slightly better estimates than the other three approaches. Estimates using MNDWI1 had an RMSE of 3.0 ha equivalent to $22,581 \text{ m}^3$, and an MAE of 1.7 ha ($10,085 \text{ m}^3$), while estimates from the GSW had a nearly two-fold higher RMSE of 5.1 ha ($49,376 \text{ m}^3$) and an MAE of 2.8 ha ($20,786 \text{ m}^3$). Small absolute errors can mask large percentage errors in area estimates. Moreover, a high percentage area error can translate to an even higher percentage volume error, since reservoir area and volume are related through a power relationship. Comparing percentage errors indicates that MNDWI1 still out-performed other approaches, producing estimates with a mean absolute percentage error at 51%, which equates to a mean volume percentage error of 58%. The GSW method had a higher MAPE than all three water indices tested, producing estimates with a mean area percentage error of 70%, which equate to a 75% mean volume percentage error. In other words, reservoir area estimates that were derived from the GSW dataset using our method were 19% less accurate than those that can be derived by

applying MNDWI1 to Landsat 8 OLI imagery, resulting in 17% less accurate reservoir volume estimates (see Table 1).

Table 1. Accuracy of 272 reservoir areal extents and volume equivalents derived from Global Surface Water Monthly Water History (GSW), Normalised Difference Water Index (NDWI), Modified NDWI with band 6 (MNDWI1), and Modified NDWI with band 7 (MNDWI2).

Method	Threshold	Mean Error (ha)	SD (ha)	RMSE (ha)	RMSE (m ³)	MAE (ha)	MAE (m ³)	MAPE (% Area)	MAPE (% Volume)
GSW	–	–2.7	4.3	5.1	49,376	2.8	20,786	70.3	75.1
NDWI	–0.2	–2.1	3.5	4.0	35,605	2.4	16,325	64.4	69.6
MNDWI1	–0.2	–1.3	2.6	3.0	22,581	1.7	10,085	51.2	58.3
MNDWI2	0	–1.3	2.8	3.0	24,041	1.8	10,705	52.7	60.1

The additional overall inaccuracy of the GSW method when compared to MNDWI1 was associated with a higher number of omission errors, i.e., 100% under-estimates (Table 2). In total 140 (51%) reservoir areas were omitted using the GSW-derived estimates, as compared to 82 using those from MNDWI1. These 100% GSW-derived underestimates occurred over very small reservoirs: 75% were smaller than 2.9 ha (~21,788 m³), and the median reservoir area was 1.0 ha. In contrast, 100% underestimates in MNDWI1-derived reservoir areas occurred on reservoirs with a median area of 0.7 ha and 75% of which were smaller than 1.5 ha (~8396 m³).

Table 2. Number and type of errors in reservoir area estimates (n = 272) derived from GSW, NDWI, MNDWI1, and MNDWI2.

Method	Omissions (100% Under-Estimate)	Under-Estimates (by <100%)	Over-Estimates
GSW	140	121	11
NDWI	124	123	25
MNDWI1	82	153	37
MNDWI2	86	148	38

3.2. Environmental Covariates

We analysed how the accuracy of reservoir area estimates from GSW, MNDWI1, MNDWI2, and NDWI vary with environmental conditions. Factors included in the random forest regression analysis explained 63% of variance in reservoir area percentage estimation errors for GSW when compared to 49% for MNDWI1, 41% for MNDWI2 and 72% for NDWI. Mean NDVI over the reservoir was identified as the most important variable in all of the cases, followed by reservoir extent and perimeter-area ratio (i.e., reservoir geometry). Latitude was the next most important factor, while month of year was of least importance to improving accuracy under all of the methods (Figure 2).

Reviewing the MAPE for reservoir area estimates from GSW, NDWI, MNDWI1, and MNDWI2 as environmental factors vary gives an indication of when area estimates are most accurate and when estimates from GSW are comparable in accuracy to other methods. We stratified reservoir area estimates into 0.1 percentile classes for each factor that was included in the random forest analysis and compared variance in mean absolute percentage errors across each class (see Figure 3).

Results show that area estimates from MNDWI1 and MNDWI2 were more accurate than those from GSW, while accuracies of GSW and NDWI are similar, under all of the conditions. For all of the methods, percentage errors increased with NDVI and with reservoir perimeter-area ratio, and reduced with increasing reservoir size. Percentage errors were lower at mid-latitudes across all of the methods, and relatively stable through the year except in July and September where there was high variability across methods. Errors in GSW estimates were lower than average when mean NDVI \leq 0.09; reservoir area is $>$ 3.64 ha, reservoir perimeter-area ratio is \leq 0.32.

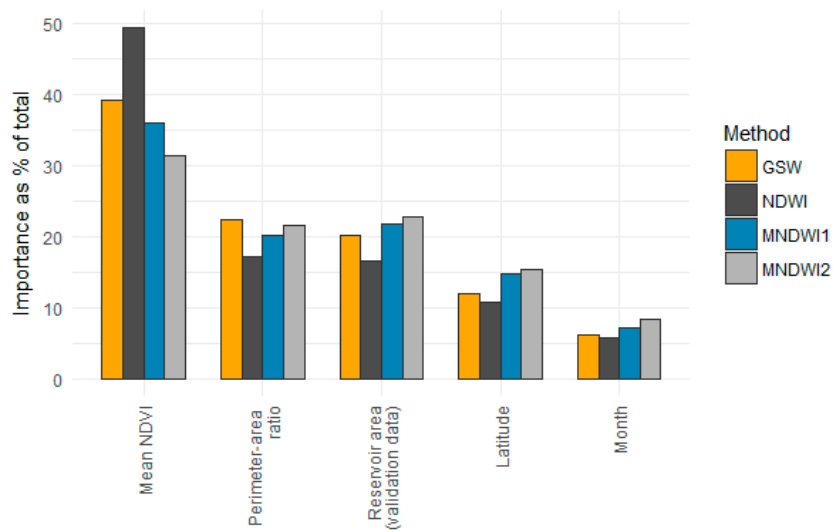


Figure 2. Importance (as percentage of total across all variables) of five environmental variables in producing accurate reservoir area estimates, as calculated by a random forest regression analysis of percentage errors in reservoir area estimates from GSW, MNDWI1, MNDWI2, and NDWI on mean NDVI, perimeter-area ratio, reservoir area, latitude and month.

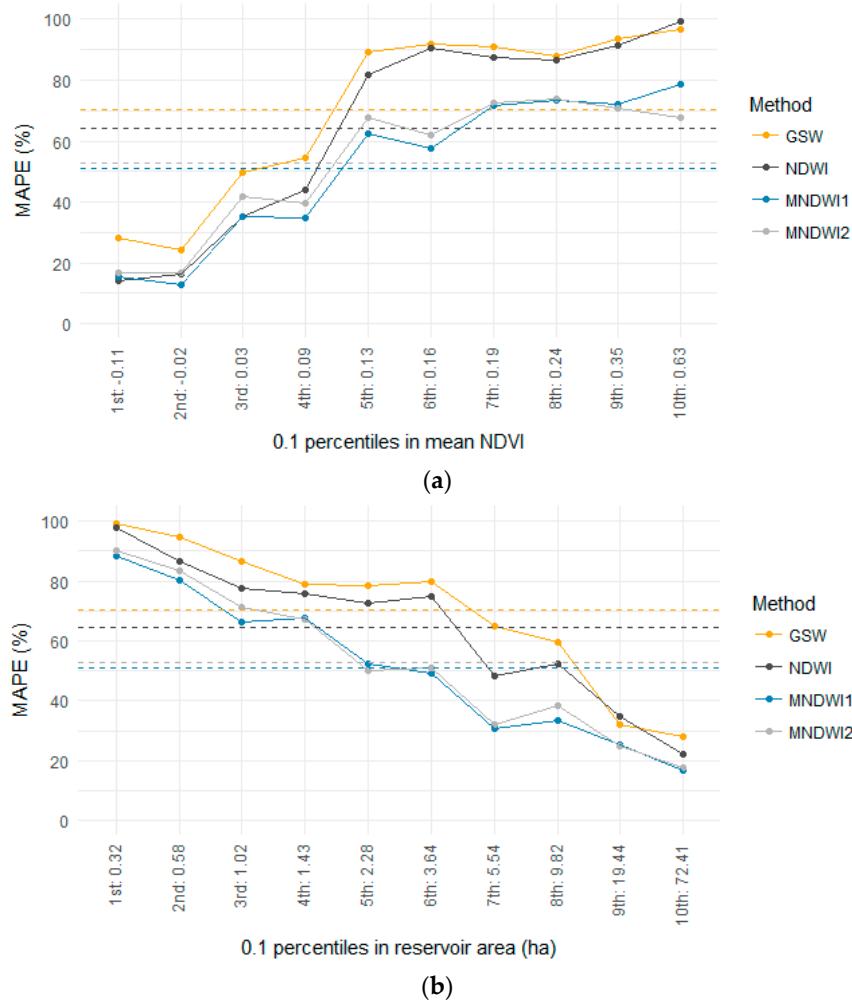


Figure 3. Cont.

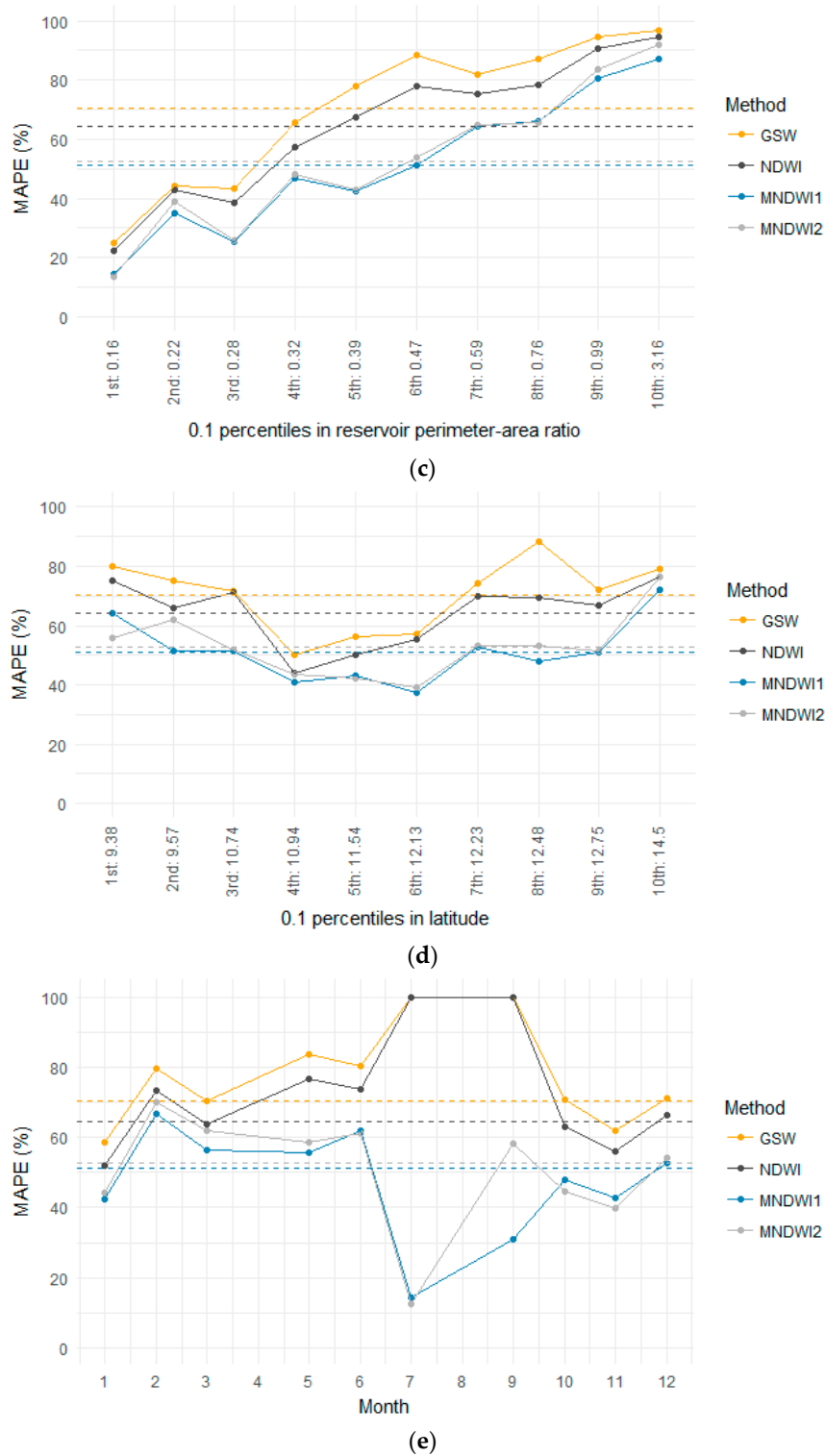


Figure 3. MAPE in reservoir area estimates derived from the four water classification methods (GSW, NDWI, MNDWI1, MNDWI2) when validation reservoirs are stratified by (a) mean NDVI, (b) reservoir area, (c) perimeter-area ratio, (d) latitude, and (e) month. Dashed lines indicate the overall MAPE corresponding to each method.

3.3. Applications of Reservoir Extent Data in Agricultural Landscapes

In total, 380 months of data are embedded in the March 1984 to October 2015 GSW dataset. However, there were substantial gaps in coverage over specific reservoirs and time periods in our derivative reservoir extent layer. Individual reservoirs had between 47% and 87% missing data across the 380 months; or, 74–98% pre-2000 and 18–77% from 2000 onwards. Gaps were concentrated in the period 1984–1999, indicative of gaps in the underlying USGS Landsat image archive [58] specifically over West and north Africa [25]. Data availability increased with latitude, likely being a result of decreasing cloud cover. Data gaps of over 50% were present in most years during peak rainfall months, from June to September, when cloud-cover may prevent accurate water classification. The bias towards reservoir area estimates in dry season months means annual mean and maximum reservoir areas will be underestimated, especially for reservoirs that dry up quickly after rains cease (e.g., very small reservoirs). Individual reservoirs had one or more valid area estimate per year (i.e., one month or more of data) for between 19 and 28 of the 32 years, meaning there were whole years of no data at every reservoir. While at most reservoirs the years of missing data are pre-2000, data coverage was very low in some subsequent years, including 2003 and 2012. Even in years with valid observations if these are from the dry season then they can be 0 ha, indicating an absence of water. Further details on data coverage are provided in Appendix B Figure A2.

Reservoir capacities as indicated by the maximum extents ever recorded show that at 339 reservoirs, the reservoir areas were permanently recorded as 0 ha, and were therefore assumed to be reservoirs smaller than the minimum mapping unit for GSW. Including these 339 reservoirs, 674 (56%) of the basin's reservoirs were smaller than 5.1 ha, the RMSE for GSW-derived estimates, while an estimated 1055 (88%) of the basin's reservoirs are smaller than 41.2 ha (1 Mm³) (see Table 3). At their maximums, small reservoirs cover 6618 ha of the Volta basin, collectively holding 1476 Mm³ of water. This equates to about 0.1% of the total water stored in reservoirs in the Volta basin, or 16% if the 13 very large reservoirs (>100 Mm³) are excluded. The mean area that is covered by these small reservoirs at their maximum extents is 9.2 ha (SD = 10.0 ha), with each estimated to contain 116,780 m³ water, equivalent to 47 Olympic swimming pools.

Table 3. Reservoir number and size based on their maximum GSW-derived extents, providing an indication of reservoir capacity and an upper limit for current reservoir size. Volumes are calculated using Equation (4).

Type	Size	Number	Mean Area (ha)	Mean Volume (m ³)	Total Area (ha)	Total Volume (m ³)
Small	Unknown (likely very small)	339	-	-	-	-
	<41.2 ha (1 Mm ³)	716	9.2 (SD = 10.0)	116,780 (SD = 130,197)	6618	1475.8 M
Large	41.2 (1 Mm ³)– 1016.3 ha (100 Mm ³)	132	157.0 (SD = 163.1)	6.8 M (7.2 M)	20,728	7610.0 M
	>1016.3 ha (100 Mm ³)	13	63,511.7 (SD = 197,167.2)	38,022.5 M (SD = 193,584.2 M)	825,652	1,515,106.4 M
All	-	1200	-	-	852,998	1,524,192.0 M

Over 92% of the identified reservoirs are located in Burkina Faso or Ghana. Reservoir density is higher (>50 per sq km) in the southern tip and central corridor of the basin, while reservoirs are, on average, substantially larger towards the north of the basin, in Burkina Faso, and in pockets of southern and western Ghana. Small reservoirs dominate the central plains (see Figure 4). While the volumes that are contained in small reservoirs are relatively small, their spatial dispersion means they make water accessible to people in many otherwise unserved parts of the basin, which includes some of the basin's poorest households and driest landscapes.

The presence of a reservoir does not assure that it contains water throughout the dry season, information that is important for water scarcity mapping and agricultural planning. The large

intra-annual data gaps and uncertainties in GSW derived reservoir extent data make it challenging to monitor the rates of change in reservoir area and equivalent volumes, for example to identify the month a reservoir runs dry. However, this is possible at reservoirs where data exist throughout the dry season in a single hydrological year. We find that, excluding June, July, and August, which correspond to rainfall months throughout the basin, and excluding reservoirs where no water was ever recorded between 1984 and 2015, GSW derived reservoir areas are available for every month for 350 reservoirs in the 2014–2015 hydrological year. Area estimates from September 2014 to May 2015 show that 256 of these reservoirs had a monthly area loss equal to or smaller than the 5.1 ha RMSE for GSW-derived estimates, and therefore their seasonality cannot be reliably determined. Of these 256 reservoirs, 255 were classified as Small and one as Large based on their maximum volumes (see Table 3). For the remaining 94 reservoirs, of which 30 were Small and the others Large or Very Large, and 25 were perennial while 69 were ephemeral running dry for up to seven months of the year (see Figure 5).

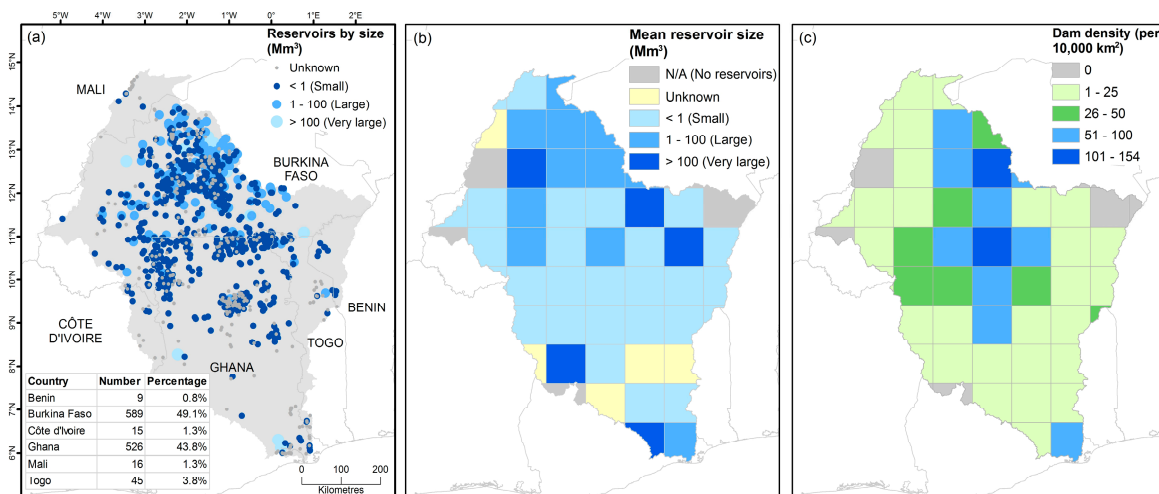


Figure 4. (a) Reservoir locations and size, (b) mean reservoir size, and (c) reservoir density across the six Volta basin countries.

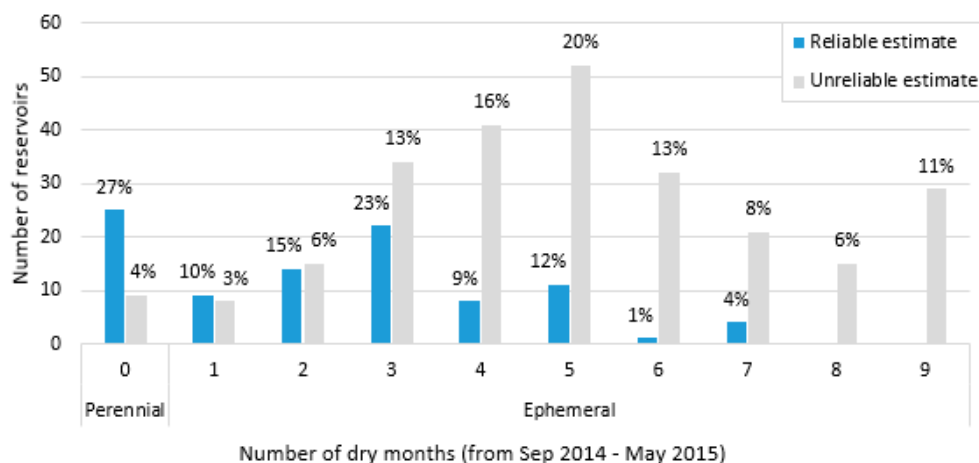


Figure 5. Seasonality of a subset of reservoirs (n = 350), for which valid GSW derived reservoir area estimates were available from September 2014 to May 2015. The figure shows the number of months a reservoir was recorded as dry distinguishing reservoirs where the mean monthly change in area was larger than the RMSE in GSW derived area estimates (“Reliable estimate”, n = 94) from reservoirs where the monthly area change was equal to or smaller than the RMSE (“Unreliable estimate”, n = 256).

4. Discussion

4.1. Accuracy of Reservoir Extents Derived from Landsat-Based Surface Water Maps

Landsat-derived reservoir area estimates contain uncertainty with even MNDWI1, the best performing of the water classification methods that were tested, producing estimates with a MAPE of 51%, for reservoirs between 0.09 ha and 72 ha in extent. Reservoir area estimates from MNDWI1, which uses Landsat 8 band 6, are 1.6% more accurate than those from MNDWI2, which uses Landsat 8 OLI band 7, and 13% more accurate than those derived from the NDWI. In contrast, GSW produces area estimates with a MAPE of 71%, or 19% less accurate than those derived from MNDWI1.

These results indicate that exploiting the difference between visible and short-wave infrared light, rather than visible and infrared, is preferable for surface water mapping in the Volta basin. This may be because the difference in the amount of infrared light reflected from pure water and from vegetation is small relative to the difference in short-wave infrared light reflected [27], and crops and other waterline vegetation are prevalent around dryland reservoirs. The results further indicate that bands with shorter wavelengths slightly out-perform those with longer as short-wave infrared inputs to the MNDWI for mapping surface water extents from Landsat satellite imagery, as consistent with [14,29]. The higher uncertainty in GSW-derived area estimates indicates that the water classification algorithm used in creating the GSW datasets is less effective at distinguishing water from non-water pixels than NDWI or MNDWI using Landsat 8 OLI imagery, for our West African study site. Further research is required to check whether this result holds for other sensors and dryland regions.

For NDWI, MNDWI1, and MNDWI2, we find that careful selection of the threshold for water classification is critical to reduce errors in reservoir area estimates. Across all three spectral indices, the accuracy of reservoir area estimates varies significantly with thresholds between -0.5 and 0.5 . The commonly used zero-threshold for water classification is sub-optimal for two of the three water indices that were tested, NDWI and MNDWI1, for which use of a slightly negative (-0.2) threshold produces superior reservoir area estimates. This may be because while pure water normally reflects little or no infrared or short-wave infrared light, the inverse is true for both vegetation and soil [59]. Therefore, water with suspended sediment or high chlorophyll concentrations, likely to be common in dryland reservoirs in agricultural landscapes that are exposed to both accelerated erosion and nutrient runoff [60], may reflect more infrared and short-wave infrared light than clear water. This increases the possibility of a slightly negative difference between visible and infrared or short-wave infrared light.

Reservoir area estimates from MNDWI1 and MNDWI2 are more accurate than those from GSW under all of the conditions tested in this study, while estimates from NDWI slightly outperform those from GSW, except in a few cases, such as for larger reservoirs (see Figure 3). Percentage errors in reservoir area estimates increase with NDVI and with reservoir perimeter-area ratio, and reduce with an increasing reservoir size, for all four water classification methods. Since NDVI responds to green vegetation cover [55], it is likely that surface or sub-surface vegetation is responsible for an increase in errors with mean NDVI. We expect the poor performance of all the methods on very small reservoirs and those with a high perimeter to area ratio is due to the 30 m resolution of Landsat data, limitations in pixel-based water-classification algorithms, and our automated approach that extracts connected water features above a pre-defined point. Reservoir area estimates have lower percentage errors at mid-latitudes of the Volta basin across all of the methods, which might be a result of higher annual cloud cover in the south and higher airborne sand and dust levels in the north hindering water classification. Percentage errors in reservoir area estimates from GSW and NDWI show a sudden increase in wetter months, while those from MNDWI1 and MNDWI2 show a sharp decline. This result is partly a reflection of the superior performance of MNDWI1 and MNDWI2 over very small reservoirs, since all of the reservoirs from July to September are <7.6 ha with a median of 3.5 ha, however it may also be symptomatic of shortcomings in GSW and NDWI water classification approaches over the turbid waters that are associated with heavy rainfall.

4.2. Conditions Under Which GSW Can Provide Reliable Information on Reservoir Size and Seasonality

Results of the random forest and subsequent analysis of percentage error variance with environmental conditions show that controlling for surface water vegetation (as indicated by NDVI) and reservoir size and shape would improve the reliability of area estimates from GSW, since these are the main factors underpinning 63% of the variance in percentage errors. Users can apply thresholds for NDVI and reservoir size and shape to identify when reservoir area estimates have an unacceptable error depending on the end-use, or when errors are lower than average and therefore reliability increases. For example, in our study, percentage errors in GSW estimates were lower than average when mean NDVI ≤ 0.09 , reservoir area >3.64 ha, and reservoir perimeter-area ratio ≤ 0.32 .

Correctly identifying when a reservoir contains water during the year is necessary for seasonality analyses and for determining reservoir locations if these are unknown. Our analysis shows that GSW produces reservoir area estimates with a RMSE of 5.1 ha, meaning that if a reservoir extent changes by ≤ 5.1 ha, this change may not be detected in GSW-based analyses of reservoir size. For reservoirs ≤ 5.1 ha in extent, the equivalent loss in volume is $\leq 49,759$ m³, while for larger reservoirs, any loss in area would occur in the shallowest regions and therefore represent a smaller equivalent loss in volume. In contrast, reservoir areas estimated from MNDWI1 in this study had an RMSE of 3.0 ha, equivalent to 24,041 m³. Further, if a reservoir extent is equal to or smaller than the RMSE for any given month, the reservoir may be incorrectly identified as dry introducing additional uncertainty to seasonality analyses (see Figure 5). Indeed, we find that the presence of water in very small reservoirs, specifically those under 2.9 ha for GSW estimates and under 1.5 ha for MNDWI1 estimates, is often entirely missed using Landsat-based water classification approaches. The omission of many small waterbodies in national or global inventories will lead to inaccurate surface water accounting and may hinder government or NGO ability to target dam construction and maintenance effectively or allocate water resources in a socially and environmentally sustainable manner. Moreover small surface waterbodies can have surprisingly large-scale cumulative effects on hydrological and ecosystem processes [10], including altering downstream water supplies, trapping sediment, and impacting on global greenhouse gas emissions [61,62].

Knowing the reservoir size constraints that are associated with a selected method is important for agricultural and water resource management planning, particularly in landscapes with small reservoirs. For example, our results indicate there are 674 reservoirs in the Volta basin with a GSW-estimated maximum capacity of 49,759 m³ (5.1 ha), for which GSW-derived size estimates are not reliable and complete omissions of reservoir water presence are likely. Further, of 350 reservoirs where continuous monthly area estimates were derived from GSW through the 2014–2015 dry season, reliable estimates of intra-annual dry periods could be obtained for only 30 of the 285 reservoirs that were classified as small since all of the other small reservoirs had a monthly area loss of less than 5.1 ha (see Figure 5). However, even when the quantity of water cannot be reliably measured, the presence or absence of water may still be correctly identified since the RMSE will be inflated by errors at larger reservoirs. As a minimum, end-users should indicate the expected uncertainty in area or volumes derived from GSW, for example, based on the RMSE or a similar measure. This can be used to indicate when the size of a reservoir or change in reservoir size drops below the size of the estimation error and thus estimates are unreliable.

4.3. Value of a Mixed-Methods Approach

We developed a method for remotely mapping small reservoirs, monitoring reservoir extents through time, and quantifying uncertainty in these extent estimates, that uses freely available data and tools. The added value of this method over previous attempts to monitor surface water dynamics is that it can be easily repeated anywhere to improve global reservoir data and establish rates of uncertainty in Landsat-derived reservoir extent estimates for other contexts and regions of interest. Our results indicate studies that rely solely on Landsat or coarser resolution data to map surface water dynamics will omit smaller reservoirs. Integrating data from freely available high

resolution imagery or ground-based monitoring systems is a practical, low-cost approach to ensuring reservoirs of all sizes are captured in Landsat-based water resource assessments. We show that reservoirs can be mapped manually through freely available Google Earth imagery—which could be completed using crowd-sourcing techniques (for example, see <http://geodata.policysupport.org/geowiki-databases>)—allowing for reservoir capture over large spatial extents.

As well as reducing gaps in reservoir inventories, mapping reservoirs prior to analysis of surface water maps enables application of simple waterbody extraction algorithms, such as the connectedPixelCount in GEE, significantly reducing the time required to extract data over large spatio-temporal extents. Using a server-based approach also avoids the significant computer storage and processing requirements that create challenges to analyses of multiple Landsat images [16,24,39]. GEE is particularly useful for the analysis of many Landsat images, since each Landsat image is ~1 GB in size with a footprint of approximately 185 km² and collected on a 16-day time step, so temporal analyses over large areas require hundreds of gigabytes of imagery. Access to Landsat satellite imagery, GSW datasets, and image processing algorithms through server-based tools such as GEE, once the software learning curve is surpassed, significantly reduces the time and cost that is involved in deriving and comparing the accuracy of waterbody extents across time and space.

Data gaps in the reservoir area estimates from the GSW Monthly Histories dataset over the Volta basin limit the types of information that can be gathered from this resource. Most of these gaps are caused by missing imagery in the underlying Landsat archive [25], and therefore would not be avoided by applying an alternative water classification method on Landsat satellite imagery. Other gaps represent areas that are classified as non-valid by [16], e.g., cloud or shadow. Time series interpolation at distinct reservoirs, for example using Amelia [63], can produce estimates of missing monthly reservoir areas, but levels of uncertainty are likely to be high over many reservoirs and years due to the large data gaps. The most promising solution to filling time-series data gaps—in the absence of access to additional Landsat and other historical satellite data [64]—is probably by integrating water classifications from Landsat data with those produced from lower resolution optical imagery, such as from MODIS [20], while seasonal gaps may be reduced by integration with water classified using synthetic aperture radar data, such as from Sentinel-1 [65], since these are not sensitive to periods of cloud cover.

4.4. Policy Applications in Agricultural Landscapes

In the Volta basin, applying our approach shows that the availability of small reservoir water is currently highly uneven across space and through the seasons. For example, dam density is higher in central Burkina Faso and northern Ghana, while some of the northern, western, and eastern regions of Burkina Faso, despite being among the driest in the Volta basin, are served by far fewer reservoirs. However, the extent of the gaps in GSW-derived data, which cover entire years for many reservoirs, limit its use for monitoring long-term trends in reservoir water availability. For some policy applications, this may not present a constraint, since long-term trends will be of less interest than analyses that show current conditions and short-term fluctuations, which can be derived from data post-2012 when underlying coverage is much higher in the GSW dataset for our West African site, and imagery from Landsat 8 OLI are available. Even without data gaps, agricultural policy applications in landscapes with small reservoirs, such as the Volta basin, where our analysis indicates 88% of reservoirs are <1 Mm³, remains a challenge since the level of uncertainty in GSW and other Landsat derived reservoir areas hinders reliable seasonality analyses and can result in omissions of whole reservoirs.

4.5. Limitations of Our Approach

While our approach to obtaining time series reservoir extents can be rapidly applied over large spatio-temporal extents through GEE, it relies on the prior identification of dam locations. For this study, we mapped dams from the most recent imagery available in Google Earth in 2015. Dams that were subsequently constructed, or previously constructed and are now out of use, are not captured in

our analysis. Automating the identification of reservoir locations would be a faster approach, however we have demonstrated that water classification maps from Landsat imagery are unable to reliably detect reservoirs <2 ha, which constitute over a third of reservoirs in the Volta basin. A useful future step would be to explore the potential for using higher resolution imagery, particularly from the Sentinel-2 satellite (10 m resolution), to automate the identification of small reservoirs.

Gaps in the 32-year reservoir extent data derived from GSW-MH arise from masked pixels (e.g., cloud or haze-covered) or a lack of imagery in the underlying Landsat collection [16]. Additional gaps may arise from our automated area-extraction method where the reservoir point is outside the reservoir water extent. Missing data in monthly reservoir area estimates distorts individual and population reservoir storage means, maximums, and minimums, and therefore limits the utility of the dataset for monitoring changes in reservoir surface areas through time.

In addition to missing values in the GSW-MH dataset, a limitation of our method is that the `connectedPixelCount` algorithm in GEE used to extract reservoir extents can handle up to 1024 connected pixels, meaning that reservoirs that cover a larger area will be underestimated. For Landsat data at 30 m resolution, this area limit is equivalent to 92.16 ha. This is substantially larger than the 41.2 ha upper size limit for small reservoirs, but creates a problem for the analysis of larger waterbodies. Reservoir estimates for these reservoirs can either be obtained by reducing the resolution (and therefore the number of pixels covered by a reservoir) on underlying Landsat imagery over larger reservoirs to enable use of the `connectedPixelCount` algorithm, or performing the analysis in a desktop GIS.

5. Conclusions

Reservoirs can provide a lifeline to the basin's rural poor during dry spells and seasons, generating stable sources of food and income through livestock, fish, and crop production. But households that depend on small reservoirs for their livelihoods do so at substantial risk. In years with low rainfall or heavy withdrawals, there may not be enough water to complete the dry season cropping season, to water livestock, or to sustain fish populations. Crop failure or fish decline means the loss of food and income at the household and community level. Lack of livestock water forces seasonal migration often to areas with their own land and water resource management challenges. Monitoring the distribution and seasonality of reservoir water on a regular basis through remote methods in the Volta basin and elsewhere is a cost-effective way for governments and non-governmental organisations to identify high risk zones. This knowledge may facilitate the implementation of safeguards to minimise water shortages and food losses, and can be used to inform decisions on future dam investments.

We developed a semi-automated method for mapping reservoirs and their extents through time, and assessing uncertainty in Landsat-derived reservoir size estimates, which can be readily applied anywhere in the globe using freely available data and tools. We used our method to compare estimates of reservoir area that were derived from four approaches to classifying surface water from Landsat data. These include the approach that was used to create the Global Surface Water Monthly Water History (GSW) datasets [16], and surface water maps created in this study from Landsat 8 OLI imagery by computing the Normalised Difference Water Index (NDWI) and two variations of the Modified NDWI, which employ band 6 (MNDWI1) and band 7 (MNDWI2) as the short-wave infrared inputs, testing effects of classifying water across 11 thresholds. We find the mean absolute percentage error is 71% for reservoir area estimates derived from GSW data, tested over 272 reservoirs between 0.09 ha and 72 ha in extent. The accuracy of these reservoir areal estimates can be improved by up to 19% by classifying water pixels on Landsat 8 OLI imagery using MNDWI1 with a carefully selected threshold, and improved to a lesser degree using NDWI or MNDWI2. Estimates that are derived from MNDWI1 consistently out-perform estimates from GSW as reservoir geometry, vegetation characteristics, and measurement season vary. Our results imply that the expert system classifier used to identify water pixels in creation of the GSW is sub-optimal to using any of the three water indices tested here, for images from the Landsat 8 OLI and over our West African study site. Further research

is required to check whether this result holds in other contexts, and for imagery that is collected by the Thematic Mappers and Enhanced Thematic Mapper onboard Landsat satellites 4, 5, and 7.

Our study provides new information on the reliability of reservoir size and seasonality from GSW and other Landsat-based surface water maps, which is important given that these data are freely and globally available and use of information on reservoir resources for policy making without the knowledge of inherent uncertainties may have serious consequences. Whether or not the 19% increase in accuracy in reservoir area estimates obtained from using surface water maps generated by applying MNDWI1 to Landsat 8 OLI images rather than using the pre-prepared GSW data, depends on the end-use. For agricultural planning and seasonal water resource management, we recommend that the use of GSW estimates be restricted to reservoirs with a maximum volume and monthly water loss of $>49,759 \text{ m}^3$ (5.1 ha), to avoid masked or false detection of water shortages. Further, Landsat-based approaches are often unable to detect any water in very small reservoirs, namely those <1.5 ha using MNDWI1 and <2.9 ha using GSW. New opportunities for the remote monitoring of small reservoirs are opening up with the availability of Sentinel-1 radar and high resolution (up to 10 m) Sentinel-2 optical data, which together should help to reduce spatial and seasonal data gaps, and improve the accuracy of derived reservoir area estimates. In the meantime, adopting integrated approaches to mapping small reservoirs remotely, such as manual digitising reservoir locations from high resolution imagery combined with automated reservoir extent extraction from Landsat imagery, is essential to avoid entirely omitting numerous water bodies.

Acknowledgments: This research was supported by a two-year CGIAR Water, Land and Ecosystems project called Targeting agricultural innovation and ecosystem services in the northern Volta basin (TAI). We would like to thank all donors who supported this research through their contributions to the CGIAR Fund. Sarah Jones is grateful to Bioversity International and King's College London for providing her with additional funding that helped take this research to completion. David Smedley would like to thank the NERC DTP for supporting his work through award NE/L002485/1. We gratefully acknowledge Simone Turner and Lambert Ngenzi for their input to creating the validation dataset, Elise Devine for helping snap the dam point dataset, and Bradley Luff for technical support. Bioversity International covered the costs of open access publishing.

Author Contributions: A.F., M.M. and S.J. conceived and designed the experiments. D.S., M.M. and S.J. performed the experiments. S.J. analyzed the data. All authors helped write the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A

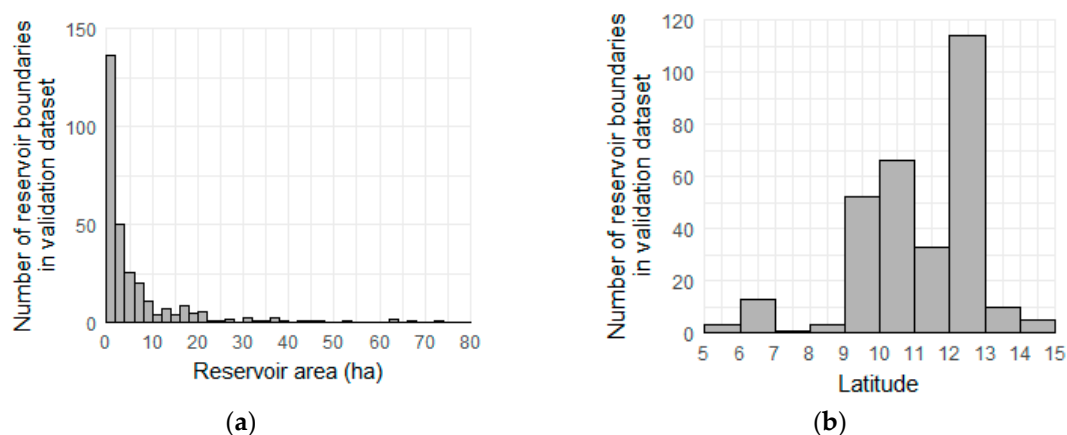


Figure A1. Cont.

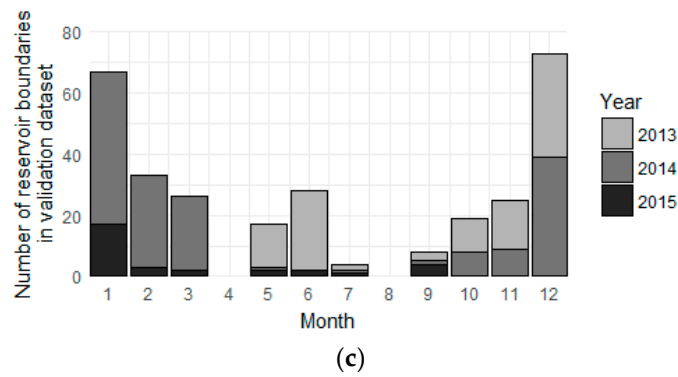


Figure A1. Google Earth derived reservoir areas data (n = 272), used as validation to assess accuracy of GSW and Landsat 8 OLI reservoir area estimates, distributed by (a) reservoir area, (b) reservoir latitude, (c) imagery month and year used to digitize reservoir boundary.

Appendix B

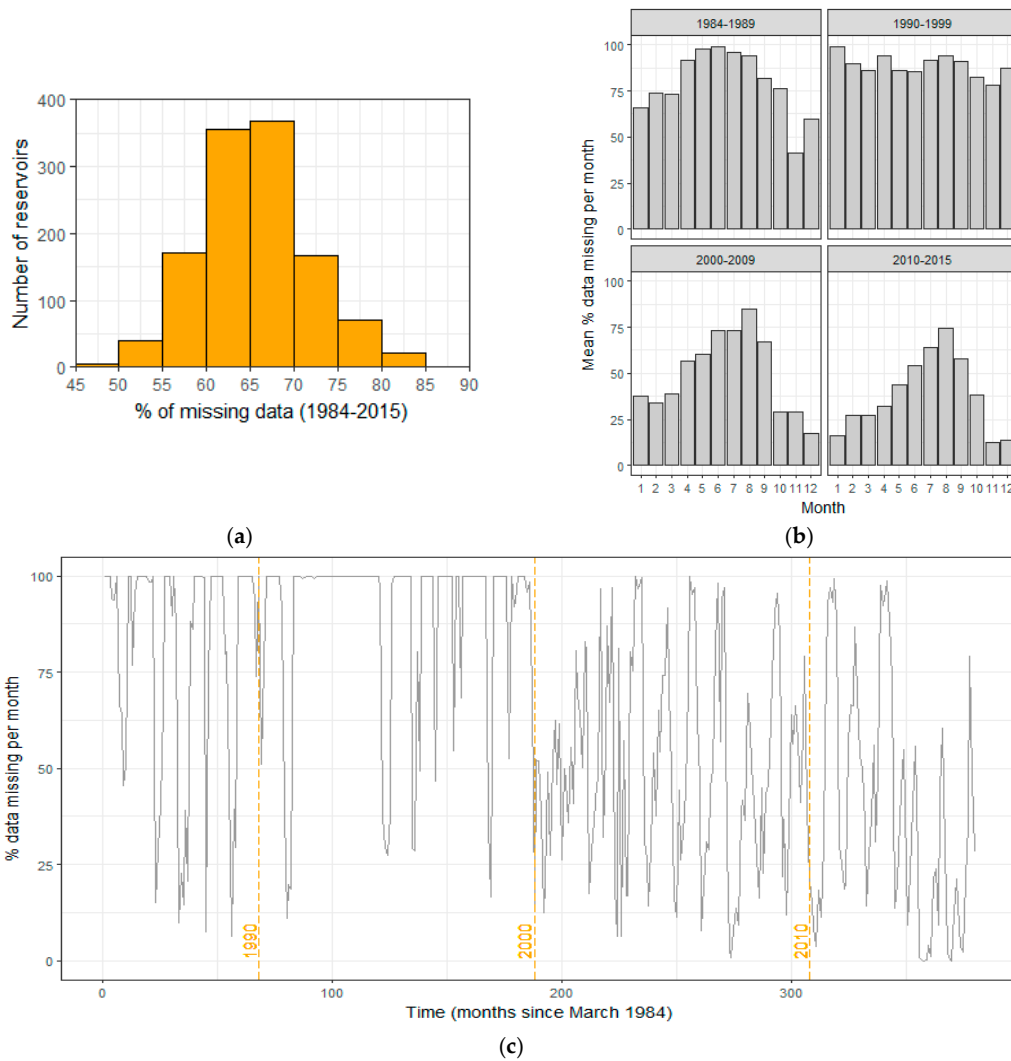


Figure A2. Data gaps in GSW-MH derived monthly reservoir area estimates for 1200 Volta basin reservoirs by (a) percentage of missing data for each reservoir, (b) mean percentage of missing data for each reservoir per month for each decade, and (c) percentage of reservoirs missing data for individual months.

Appendix C

Table A1. Accuracy of reservoir area estimates (n = 272) derived from Landsat 8 OLI imagery across the 11 thresholds tested for NDWI, MNDWI1 and MNDWI2. Highlighted rows represent optimal thresholds, i.e., those with the lowest mean absolute percentage error for reservoir area estimates.

Method	Threshold	Mean Error (ha)	SD (ha)	RMSE (ha)	RMSE (m ³)	MAE (ha)	MAE (m ³)	MAPE (% Area)	MAPE (% Volume)
NDWI	−0.2	−2.10	3.46	4.04	35,604.71	2.35	16,325.01	64.36	69.63
NDWI	−0.3	0.18	5.20	5.19	50,983.33	2.33	16,097.32	80.94	170.03
NDWI	−0.1	−3.25	4.72	5.72	58,641.61	3.30	26,624.84	75.92	79.38
NDWI	0	−4.19	5.83	7.17	81,063.24	4.20	37,658.41	85.93	88.36
NDWI	0.1	−5.35	8.41	9.95	129,900.56	5.35	53,278.05	92.77	94.28
NDWI	0.2	−6.28	10.14	11.91	168,116.85	6.28	67,039.37	97.21	97.84
NDWI	0.3	−6.87	11.43	13.32	197,431.62	6.87	76,284.01	99.24	99.42
NDWI	0.4	−7.05	11.70	13.64	204,164.64	7.05	79,058.93	99.87	99.92
NDWI	0.5	−7.09	11.80	13.75	206,626.22	7.09	79,844.29	100.00	100.00
NDWI	−0.4	28.00	39.07	48.01	1,245,499.69	28.98	603,229.78	4025.68	49,456.09
NDWI	−0.5	64.76	37.03	74.56	2,344,505.15	65.06	1,927,217.86	7244.77	84,773.09
MNDWI1	−0.3	−0.28	2.62	2.63	19,241.17	1.43	7989.78	56.81	123.11
MNDWI1	−0.2	−1.32	2.64	2.95	22,581.41	1.68	10,085.23	51.17	58.30
MNDWI1	−0.1	−2.29	3.58	4.25	38,206.12	2.45	17,316.77	63.29	68.82
MNDWI1	0	−3.15	4.71	5.66	57,726.14	3.21	25,599.05	73.34	77.38
MNDWI1	0.1	−4.16	6.88	8.02	95,314.52	4.19	37,427.90	81.80	84.90
MNDWI1	0.2	−5.16	9.33	10.64	143,043.85	5.18	50,788.42	88.38	90.61
MNDWI1	0.3	−5.75	9.90	11.43	158,542.46	5.76	59,134.43	92.56	94.04
MNDWI1	0.4	−6.47	11.21	12.93	189,063.76	6.47	69,929.69	96.03	96.91
MNDWI1	0.5	−6.71	11.50	13.29	196,860.03	6.71	73,649.05	97.77	98.23
MNDWI1	−0.4	3.82	13.17	13.69	205,231.30	4.24	38,102.16	497.09	5788.66
MNDWI1	−0.5	35.12	39.54	52.83	1,429,033.43	35.18	796,790.10	4424.54	55,829.47
MNDWI2	0	−1.33	2.78	3.08	24,040.58	1.75	10,704.68	52.74	60.11
MNDWI2	0.1	−2.30	3.56	4.23	38,001.46	2.45	17,381.79	64.08	69.66
MNDWI2	−0.1	0.10	4.39	4.39	40,035.31	1.77	10,841.12	74.68	200.54
MNDWI2	0.2	−3.21	4.97	5.91	61,382.62	3.29	26,462.21	74.09	78.32
MNDWI2	0.3	−4.73	8.99	10.15	133,549.18	4.77	45,100.89	83.62	86.37
MNDWI2	0.4	−5.40	9.67	11.06	151,062.42	5.41	54,122.27	89.85	91.70
MNDWI2	0.5	−6.11	10.28	11.94	168,735.67	6.11	64,469.66	94.60	95.68
MNDWI2	−0.2	3.74	14.56	15.01	234,362.69	4.40	40,225.86	560.81	8282.32
MNDWI2	−0.3	14.44	27.85	31.32	674,344.73	14.60	225,195.18	1568.41	18,858.55
MNDWI2	−0.4	55.46	38.07	67.23	2,020,361.07	55.50	1,534,017.70	6123.54	71,805.25
MNDWI2	−0.5	84.73	12.88	85.70	2,863,350.91	84.73	2,817,038.71	10,128.18	118,618.67

References

1. Rockström, J.; Falkenmark, M. Agriculture: Increase water harvesting in Africa. *Nature* **2015**, *519*, 283–285. [[CrossRef](#)] [[PubMed](#)]
2. Hyman, G.; Fujisaka, S.; Jones, P.; Wood, S.; de Vicente, M.C.; Dixon, J. Strategic approaches to targeting technology generation: Assessing the coincidence of poverty and drought-prone crop production. *Agric. Syst.* **2008**, *98*, 50–61. [[CrossRef](#)]
3. Vörösmarty, C.J.; Douglas, E.M.; Green, P.A.; Revenga, C. Geospatial indicators of emerging water stress: An application to Africa. *Ambio* **2005**, *34*, 230–236. [[CrossRef](#)] [[PubMed](#)]
4. Venot, J.P.; Krishnan, J. Discursive framing: Debates over small reservoirs in the Rural South. *Water Altern.* **2011**, *4*, 316–324.
5. Douxchamps, S.; Ayantunde, A.; Barron, J. Taking stock of forty years of agricultural water management interventions in smallholder systems of Burkina Faso. *Water Resour. Rural Dev.* **2015**, *3*, 1–13. [[CrossRef](#)]
6. ICLD. Dictionary: Large Dam. Available online: <http://www.icold-cigb.net/GB/Dictionary/dictionary.asp> (accessed on 23 September 2016).
7. Sawunyama, T.; Mhizha, A. Estimation of small reservoir storage capacities in Limpopo River Basin using geographical information systems (GIS) and remotely sensed surface areas: Case of Mzingwane catchment. *Phys. Chem. Earth Parts A/B/C* **2006**, *31*, 935–943. [[CrossRef](#)]
8. Wisser, D.; Frohking, S.; Douglas, E.M.; Fekete, B.M.; Schumann, A.H.; Vörösmarty, C.J. The significance of local water resources captured in small reservoirs for crop production—A global-scale analysis. *J. Hydrol.* **2010**, *384*, 264–275. [[CrossRef](#)]
9. Downing, J.A.; Prairie, Y.T.; Cole, J.J.; Duarte, C.M.; Tranvik, L.J.; Striegl, R.G.; McDowell, W.H.; Kortelainen, P.; Caraco, N.F.; Melack, J.M.; et al. The global abundance and size distribution of lakes, ponds, and impoundments. *Limnol. Oceanogr.* **2006**, *51*, 2388–2397. [[CrossRef](#)]
10. Downing, J.A. Emerging global role of small lakes and ponds: Little things mean a lot. *Limnetica* **2010**, *29*, 9–24.
11. Lemoalle, J.; Condappa, D.D. *Atlas de l'eau du Bassin de la Volta = Water Atlas of the Volta Basin*; Challenge Program on Water and Food and Institut de Recherche pour le Développement: Colombo, Sri Lanka; Marseille, France, 2009; pp. 1–96.
12. Birner, R.; McCarthy, N.; Robertson, R.; Waale, D.; Schiffer, E. Increasing access to irrigation: Lessons learned from investing in small reservoirs in Ghana. In Proceedings of the Agricultural Services, Decentralization, and Local Governance Workshop, Accra, Ghana, 3 June 2010.
13. Sawaya, K.E.; Olmanson, L.G.; Heinert, N.J.; Brezonik, P.L.; Bauer, M.E. Extending satellite remote sensing to local scales: Land and water resource monitoring using high-resolution imagery. *Remote Sens. Environ.* **2003**, *88*, 144–156. [[CrossRef](#)]
14. Du, Z.; Li, W.; Zhou, D.; Tian, L.; Ling, F.; Wang, H.; Gui, Y.; Sun, B. Analysis of Landsat-8 OLI imagery for land surface water mapping. *Remote Sens. Lett.* **2014**, *5*, 672–681. [[CrossRef](#)]
15. Mueller, N.; Lewis, A.; Roberts, D.; Ring, S.; Melrose, R.; Sixsmith, J.; Lymburner, L.; McIntyre, A.; Tan, P.; Curnow, S.; et al. Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sens. Environ.* **2016**, *174*, 341–352. [[CrossRef](#)]
16. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [[CrossRef](#)] [[PubMed](#)]
17. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [[CrossRef](#)]
18. Pekel, J.-F.; Vancutsem, C.; Bastin, L.; Clerici, M.; Vanbogaert, E.; Bartholomé, E.; Defourny, P. A near real-time water surface detection method based on HSV transformation of MODIS multi-spectral time series data. *Remote Sens. Environ.* **2014**, *140*, 704–716. [[CrossRef](#)]
19. Ogilvie, A.; Belaud, G.; Delenne, C.; Bailly, J.-S.; Bader, J.-C.; Oleksiak, A.; Ferry, L.; Martin, D. Decadal monitoring of the Niger Inner Delta flood dynamics using MODIS optical data. *J. Hydrol.* **2015**, *523*, 368–383. [[CrossRef](#)]
20. Khandelwal, A.; Karpatne, A.; Marlier, M.E.; Kim, J.; Lettenmaier, D.P.; Kumar, V. An approach for global monitoring of surface water extent variations in reservoirs using MODIS data. *GISci. Remote Sens.* **2016**, 1–24. [[CrossRef](#)]

21. D'Andrimont, R.; Defourny, P. Monitoring African water bodies from twice-daily MODIS observation. *GISci. Remote Sens.* **2017**, *1*–24. [[CrossRef](#)]
22. Klein, I.; Gessner, U.; Dietz, A.J.; Kuenzer, C. Global WaterPack—A 250 m resolution dataset revealing the daily dynamics of global inland water bodies. *Remote Sens. Environ.* **2017**, *198*, 345–362. [[CrossRef](#)]
23. Liebe, J.; van de Giesen, N.; Andreini, M. Estimation of small reservoir storage capacities in a semi-arid environment. *Phys. Chem. Earth* **2005**, *30*, 448–454. [[CrossRef](#)]
24. Ogilvie, A.; Belaud, G.; Massuel, S.; Mulligan, M.; Le Goulven, P.; Calvez, R. Assessing Floods and Droughts in Ungauged Small Reservoirs with Long-Term Landsat Imagery. *Geosciences* **2016**, *6*, 42. [[CrossRef](#)]
25. Wulder, M.A.; White, J.C.; Woodcock, C.E.; Belward, A.S.; Cohen, W.B.; Fosnight, E.A.; Shaw, J.; Masek, J.G.; Roy, D.P. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* **2016**, *185*, 271–283. [[CrossRef](#)]
26. USGS. Landsat Surface Reflectance Higher-Level Data Products. Available online: <https://landsat.usgs.gov/landsat-surface-reflectance-high-level-data-products> (accessed on 31 May 2017).
27. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [[CrossRef](#)]
28. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*, 2nd ed.; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2007; ISBN 0-13-188950-8.
29. Ji, L.; Zhang, L.; Wylie, B. Analysis of Dynamic Thresholds for the Normalized Difference Water Index. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1307–1317. [[CrossRef](#)]
30. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [[CrossRef](#)]
31. Bradshaw, G.A.; Borchers, J.G. Uncertainty as Information: Narrowing the Science-policy Gap. *Conserv. Ecol.* **2000**, *4*, 7. [[CrossRef](#)]
32. Mulligan, M.; Fisher, M.; Sharma, B.; Xu, Z.X.; Ringler, C.; Mahé, G.; Jarvis, A.; Ramírez, J.; Clanet, J.-C.; Ogilvie, A.; et al. The nature and impact of climate change in the Challenge Program on Water and Food (CPWF) basins. *Water Int.* **2011**, *36*, 96–124. [[CrossRef](#)]
33. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [[CrossRef](#)]
34. United Nations Development Programme (UNDP). *2015 Human Development Report*; UNDP: New York, NY, USA, 2015.
35. Venot, J.-P.; de Fraiture, C.; Acheampong, E.N. *Revisiting Dominant Notions: A Review of Costs, Performance and Institutions of Small Reservoirs in Sub-Saharan Africa*; International Water Management Institute (IWMI): Colombo, Sri Lanka, 2012.
36. Douchamps, S.; Ayantunde, A.; Panyan, E.K.; Ouattara, K.; Kaboré, A.; Karbo, N.; Sawadogo, B. Agricultural water management and livelihoods in the crop-livestock systems of the Volta Basin. *Water Resour. Rural Dev.* **2015**, *6*, 92–104. [[CrossRef](#)]
37. Gerland, P.; Raftery, A.E.; Ševčíková, H.; Li, N.; Gu, D.; Spoorenberg, T.; Alkema, L.; Fosdick, B.K.; Chunn, J.; Lalic, N.; et al. World population stabilization unlikely this century. *Science* **2014**, *346*, 234–237. [[CrossRef](#)] [[PubMed](#)]
38. Godfray, H.C.J.; Garnett, T. Food security and sustainable intensification. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2014**, *369*, 20120273. [[CrossRef](#)] [[PubMed](#)]
39. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**. [[CrossRef](#)]
40. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
41. Singh, K.V.; Setia, R.; Sahoo, S.; Prasad, A.; Pateriya, B. Evaluation of NDWI and MNDWI for assessment of waterlogging by integrating digital elevation model and groundwater level. *Geocarto Int.* **2014**, *30*, 650–661. [[CrossRef](#)]
42. Rokni, K.; Ahmad, A.; Selamat, A.; Hazini, S. Water feature extraction and change detection using multitemporal landsat imagery. *Remote Sens.* **2014**, *6*, 4173–4189. [[CrossRef](#)]
43. Li, W.; Du, Z.; Ling, F.; Zhou, D.; Wang, H.; Gui, Y.; Sun, B.; Zhang, X. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sens.* **2013**, *5*, 5530–5549. [[CrossRef](#)]

44. USGS. *Landsat 8 (L8) Data Users Handbook*; Version 2; United States Geological Survey (USGS): Sioux Falls, SD, USA, 2016.
45. Flood, N. Continuity of reflectance data between landsat-7 ETM+ and landsat-8 OLI, for both top-of-atmosphere and surface reflectance: A study in the Australian landscape. *Remote Sens.* **2014**, *6*, 7952–7970. [[CrossRef](#)]
46. Roy, D.P.; Kovalsky, V.; Zhang, H.K.; Vermote, E.F.; Yan, L.; Kumar, S.S.; Egorov, A. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* **2015**. [[CrossRef](#)]
47. Holden, C.E.; Woodcock, C.E. An analysis of Landsat 7 and Landsat 8 underflight data and the implications for time series investigations. *Remote Sens. Environ.* **2016**, *185*, 16–36. [[CrossRef](#)]
48. Li, P.; Jiang, L.; Feng, Z. Cross-comparison of vegetation indices derived from landsat-7 enhanced thematic mapper plus (ETM+) and landsat-8 operational land imager (OLI) sensors. *Remote Sens.* **2014**, *6*, 310–329. [[CrossRef](#)]
49. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [[CrossRef](#)]
50. Google Earth Engine. Google Earth Engine API Guides: Object-Based Methods. Available online: https://developers.google.com/earth-engine/image_objects (accessed on 2 November 2017).
51. Li, W.; Qin, Y.; Sun, Y.; Huang, H.; Ling, F.; Tian, L.; Ding, Y. Estimating the relationship between dam water level and surface water area for the Danjiangkou Reservoir using Landsat remote sensing images. *Remote Sens. Lett.* **2016**, *7*, 121–130. [[CrossRef](#)]
52. Mulligan, M. WaterWorld: A self-parameterising, physically based model for application in data-poor but problem-rich environments globally. *Hydrol. Res.* **2013**, *44*, 748. [[CrossRef](#)]
53. Fisher, A.; Flood, N.; Danaher, T. Comparing Landsat water index methods for automated water classification in eastern Australia. *Remote Sens. Environ.* **2016**, *175*, 167–182. [[CrossRef](#)]
54. Rouse, J.W.; Hass, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. *Third ERTS Symp.* **1973**, *1*, 309–317.
55. Carlson, T.N.; Ripley, D.A. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.* **1997**, *62*, 241–252. [[CrossRef](#)]
56. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
57. Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random Forests for Classification in Ecology. *Ecology* **2007**, *88*, 2783–2792. [[CrossRef](#)] [[PubMed](#)]
58. Goward, S.N.; Masek, J.G.; Williams, D.L.; Irons, J.R.; Thompson, R.J. The Landsat 7 mission: Terrestrial research and applications for the 21st century. *Remote Sens. Environ.* **2001**, *78*, 3–12. [[CrossRef](#)]
59. Baldrige, A.M.; Hook, S.J.; Grove, C.I.; Rivera, G. The ASTER spectral library version 2.0. *Remote Sens. Environ.* **2009**, *113*, 711–715. [[CrossRef](#)]
60. Lal, R. Soil degradation by erosion. *L. Degrad. Dev.* **2001**, *12*, 519–539. [[CrossRef](#)]
61. Messenger, M.L.; Lehner, B.; Grill, G.; Nedeva, I.; Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **2016**, *7*, 13603. [[CrossRef](#)] [[PubMed](#)]
62. Holgerson, M.A.; Raymond, P.A. Large contribution to inland water CO₂ and CH₄ emissions from very small ponds. *Nat. Geosci.* **2016**, *9*, 222–226. [[CrossRef](#)]
63. Honaker, J.; King, G. What to Do about Missing Values in Time-Series Cross-Section Data. *Am. J. Political Sci.* **2010**, *54*, 561–581. [[CrossRef](#)]
64. Wulder, M.A.; Coops, N.C. Satellites: Make Earth observations open access. *Nature* **2014**, *513*, 30–31. [[CrossRef](#)] [[PubMed](#)]
65. Pham-Duc, B.; Prigent, C.; Aires, F. Surface Water Monitoring within Cambodia and the Vietnamese Mekong Delta over a Year, with Sentinel-1 SAR Observations. *Water* **2017**, *9*, 366. [[CrossRef](#)]

