**Reply to C. Cudennec - Referee #2**

We wish to thank C. Cudennec for the valuable and relevant comments provided. Please find below a point by point reply to the issues raised and the steps taken to modify the manuscript. References not in the additional manuscript have been specified here. We look forward to finalising the revised manuscript based on these comments and additions.

**Ogilvie et al. develop and present an important methodology which values available remote sensing informations and opens low cost applications in the future. It allows assessing and mapping filling of reservoirs spread across a semiarid area, with strong implications for monitoring resources availability across the territory where water is crucial in the Water-Food-Energy-Development nexus; as well as agregation of hydrological impacts on the functioning of the whole basin. The methodology is developed and tested thanks to accurate field campaigns, over a pilot basin in central semiarid Tunisia which is emerging as a strong reference since 15 years. The methodology is based on the use of 7 well known spectral indices; and their assessment thanks to the available Landsat images is well justified and discussed, against field difficulties such as shallow waters, vegetation development, frequency of images regarding rapid dynamic of floods.**

**Litterature review about these 7 indices, and their previous applications and limits could be expanded in section 2.4 as it is too implicit so far. It is refered to a "widely used" status whereas it is mentioned at some places that these are more or less relevant (rural gently sloping watershed, design to detect mixed water reflectance, MNDWI outperformed AWEI when extracting narrow water bodies...).**

This section has been modified to provide additional detail of the applications and specificities of the spectral water indices. The following text has been inserted in the manuscript:

“Six widely used water detection indices (Table 1) were compared to assess their suitability to monitor floods on small reservoirs in terms of detection accuracy and threshold stability. These include NDWI (Normalised Difference Water Index) developed by McFeeters (1996) exploiting the difference in reflectance of water in the green and NIR bands. Xu (2006) proposed a modified NDWI (MNDWI) where the NIR band was substituted by the SWIR band to improve distinction of built up features over water. In parallel, Lacaux et al. (2007) developed the NDPI (Normalised Difference Pond Index) which also exploits the low reflectance of water in SWIR and its contrast with the green band. NDPI is effectively the opposite of MNDWI and was therefore not investigated separately. Gao (1996) developed a NDWI but using NIR and SWIR bands. Xu (2006) later defined this as NDMI (Normalised Difference Moisture Index). NDTI (Normalised Difference Turbidity Index) was also developed using the red and green bands, and exploits the principle that turbid water reacts like bare soils, i.e. low reflectance in green and high in red. The NDVI (Normalised Difference Vegetation Index) (Rouse et al. 1973) is one of the most well known band ratio index which exploits the contrast between the peak reflectance in the infrared band and the low reflectance in the red to monitor vegetation. It has also been used to detect water bodies (Ma et al. 2007; Mohamed et al. 2004) as the index becomes positive in presence of vegetation and negative for water bodies. Finally, AWEI (Automated Water Extraction Index) (Feyisa et al. 2014) was empirically developed to discriminate water over several large lakes using wavelengths within 5 Landsat bands. Two variants exist, AWEInsh and AWEIsh, the latter being optimised to remove (urban and topographic) shadow pixels. Both can be used in succession in the presence of highly reflective areas (snow, roofs, etc.) but in this rural, semi-arid gently sloping watershed, the AWEInsh was used. For SWIR, band 5 for TM/ETM+ and band 7 for OLI were used in accordance with other studies (Ji et al. 2009; Ouma and Tateishi 2006). For NVDI and NDTI, values below the threshold were classified as water, and conversely for the four other indices."

**Also a rapid positioning of this approach regarding spatial altimetry and geodesy applied to lakes would be welcome to better assess the challenges of small reservoirs in data scarce regions (See for instance Cretaux et al., 2011, 2015, 2016).**

This relevant comment has led us to add the following text in the introduction of the manuscript.

"Satellite altimetry originally developed to monitor the ocean's surface has increasingly been exploited and adapted since the 1990s to monitor height variations of continental surface waters (Cretaux et al., 2016). Used across rivers and large lakes, recent works have sought to transpose these to smaller water bodies (from 50 ha) but highlight several limitations relating essentially to the poor density of altimetry tracks, the long along-track path lengths, and low revisit times (Baup et al., 2016, Avisse et al. 2017). Altimeters aboard Topex, Envisat, Jason-1 and Jason-2 have temporal resolutions ranging from 10 to 35 days and high vertical accuracy however their narrow swaths and long along track path lengths (1 km) restrict their application essentially to large lakes (> 100 km2, Avisse et al., 2017). Other altimeters have improved spatial resolution and cover more of the globe but at the expense of low revisit times (368 days for Cryostat-2), removing any monitoring possibilities. The Dahiti altimetry database (Schwatke et al., 2015) employed by Busker et al., 2018 combines the tracks of numerous altimeters to optimise the sites covered, temporal resolution and the length of observations. These provide data for 168 sites in Africa, however none in Tunisia and lakes must have a minimum 300 m diameter (circa 7 ha). The Sentinel-3a and 3b constellation provide major improvements in their along track resolution (300 m) making them potentially suitable for lakes around 4 ha but inter-tracks of 52 km mean many lakes are not covered by their trajectories (Cretaux et al., 2016). Furthermore, radar altimetry provides an estimate of the altitude of the water surface, based on the two-way travel time of the radar pulse and the known altitude of the satellite, but assessing absolute volumes requires site specific data such as stage height or topographic data (Baup et al., 2016). Avisse et al., 2017 showed that DEM data taken before lakes were built or when it was empty could be used but on larger lakes (59 ha to 379 ha) with 30 m data."

Crétaux, J. F., Abarca-del-Río, R., Berge-Nguyen, M., Arsen, A., Drolon, V., Clos, G., & Maisongrande, P. (2016). Lake volume monitoring from space. Surveys in Geophysics, 37(2), 269-305.

Schwatke, C., Dettmering, D., Bosch, W., & Seitz, F. (2015). DAHITI–an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. Hydrology and Earth System Sciences, 19(10), 4345-4364.

Busker, T., de 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, Hydrol. Earth Syst. Sci. Discuss., in review, 2018.

**My major concern is the need to better explicit the hydrometric approach: remote sensing allows to assess water surface. Deducing water storage / volume / availability / resources needs the use of a volume-stage-surface / rating / bathymetric curve. This should be made more explicit in principle in the Introduction and then the Methods section. Further, the bathymetric curves of small reservoirs**

**in this region are not stationnary, as erosion-silting is important, yet very heterogeneous across regions and so reservoirs. The Authors say page 19 that the curve of the biggest reservoir (Mora) is obsolescent because not updated over the past 20 years. This points the need to precisely address the issue of the exact availability and accuracy of bathymetric curves for every considered reservoirs (beyond the short statement on P5, L15 referring to old Albergel and Rejeb, 1997 reference); as well as consequences in terms of uncertainties in the overall method.**

Both reviewers rightly highlighted that further details on the bathymetric curves, their evolution over time from silting and the associated uncertainties are required here. To assess the long-term performance of Landsat imagery to quantify flooded surface areas and volumes, stage data converted using site specific hypsometric (stage-surface-volume) relationships were used. These were acquired over 1990-2007 through previous research projects and additional levelling of Hoshas as part of this research in 2014. A figure has been added to the manuscript to explicit the number of relationships for each lake. To overcome the absence of regular surveys on some lakes (e.g. Morra), silting was modelled based on research on silting in 15 lakes in and around the Merguellil catchment (Albergel et al., 2003). These showed that the decline in capacity over time from silting could be modelled through linear regression. Analysis of these 70 rating curves highlighted the progressive shift in the parameters of the rating curve power relation ($V = B * S^{beta}$). Beta is shown to increase gradually over time, reflecting the decreasingly concave nature of the lakes floor. The evolution over time of the site-specific power relations was therefore calculated based on a gradual annual increase of the beta parameter and an associated decrease in maximum capacity ($V_{max}$). Initial Vmax were here known based on the inventories and used to calculate the initial $S_{max}$. By supposing that $S_{max}$ at the spillway does not evolve over time, which is acceptable based on the rating curves in our possession, the resulting B is then calculated over time. In practice, silting is heterogeneous and occurs through sudden, discrete events not a linear, incremental process but local studies confirmed the difficulties in modelling sediment transport in these small catchments (Hentati et al., 2010).

These additional details on the hypsometric relationships applied will be added to the manuscript. Furthermore, accuracy assessments of power relations updated over time to account for silting against the available updated hypsometric rating curves (as per Ogilvie et al. 2016b) as well as additional GPS contours acquired on Morra and Guettar in 2014 will be integrated. The potential to use regular Landsat derived surface area estimates at multiple water levels to create and correct the site-specific hypsometric relationships will also be discussed.

Hentati, A., Kawamura, A., Amaguchi, H., & Iseri, Y. (2010). Evaluation of sedimentation vulnerability at small hillside reservoirs in the semi-arid region of Tunisia using the Self-Organizing Map. Geomorphology, 122(1-2), 56-64.

**Minor issues:**

**-P1, L22: Reservoirs do not reduce soil loss -but sediment transfer once in the network.**

This sentence has been modified as : "These have been built to reduce sediment transfer and silting of downstream dams, as well as harvest scarce and unreliable water resources for local users (Habi and Morsli, 2011 ; Wisser et al., 2010)"

**-P2, L20: Okavango and Mekong Deltas.**

The text has been modified accordingly.

**-P3, L33: Localise instead of localised.**

The text has been modified accordingly.

**-Section 2.5: What are the exact dates of the images -and what are the characteristics of the rainfalls over that particular period?**

The precise dates of the images (29.03.2013, 24.05.2013, 09.06.2013) have been included in table 2. The following sentence has also been added: "The gradual decline in water surface area observed on all lakes across these 3 images are also coherent with the rainfall characteristics. Rainfall recorded over this period (March-June 2013) ranged between 27 mm and 44 mm across the lakes concentrated on 1 event on the 24.04.2013."

**-P9, L17: Provide references about Gouazine basin and reservoir (Nasri et al., for instance).**

Two references have been added here. "Gouazine lake (Nasri et al., 2004, Al Ali et al., 2008) which possessed both the longest time series (over 15 years) and the most accurate data and rating curves (updated 6 times) was used to optimise the thresholds."

**-P13, L6: Duplication of "in".**

The text has been modified accordingly.

**-P16, L2: The Merguellil catchment.**

The text has been modified accordingly.

**-P33, L4: author Calvez duplicated.**

The text has been modified accordingly.