Response to Reviewer #2

General comments: This well-organized manuscript presents a valuable data set and an interesting method to monitor surface water fraction/extent variation in large spatial scale and high temporal resolution. The combination of a medium resolution MODIS surface reflectance product and a well-tuned Cubist regression model provided a sub-pixel estimation of water fraction with satisfactory performance compared with GSW and previous MODIS product on surface water. Apparently, the authors have made their efforts to further improve the quality of the data set by using ancillary data such as DEM and Land Cover to eliminate possible contaminated pixels. Besides the cross validation with GSW, the authors also compared the generated data set with altimetry data in certain lakes. Given the importance of high frequency monitoring of surface water in water management, it is positive that this work will benefit the scientific community as well as public decision makers.

In revising the manuscript, following issues were encountered and I suggest authors to provide further explanation:

We thank the reviewer for her/his positive feedbacks and valuable comments and suggestions to improve the quality of the manuscript. Please find below our responses and how we have incorporated these comments in the manuscript.

(1) **Comment:** In Page 11, Table 2, the author listed the input predictors for Cubist regression model. However, it seems to me that TCWI and TCBI might be redundant variables because they are linear combination of MODIS individual bands and the regression model is also linear. Have you ever tried to remove these two predictors to simplify the model input? If so, how did it work?

Response: The predictor variables are used in two ways in the Cubist regression model. Firstly, they set rule conditions to split the samples into smaller subsets. Secondly, they are used to build linear regression models related to the rule conditions. The reviewer's concern only relates to the latter purpose, and in this regard, MODIS individual bands can indeed replace TCWI and TCBI to build linear regression models. However, TCWI and TCBI, especially their temporal characteristics, are important variables to set rule conditions. This was demonstrated in our previous paper (Li et al. 2018), in which we measured the relative importance of all variables for estimating surface water fraction in two small regions based on its usage in the rule conditions and in the linear regression models. The results (please see the figure below) showed that the temporal characteristics of TCBI and TCWI such as the annual max/min/mean are frequently used for setting rule conditions following TWI (Topographic Wetness Index) and NIR. In addition, our previous paper also showed that models using all variables (inclusive of temporal variables) achieved higher prediction accuracies as compared to simpler models (exclusive of temporal variables). Therefore, we conclude that TCBI and TCWI are not redundant but instead are important variables in the Cubist regression model. We realized that the text did not specify this, but have now revised in section 4.1.3 to clarify this.

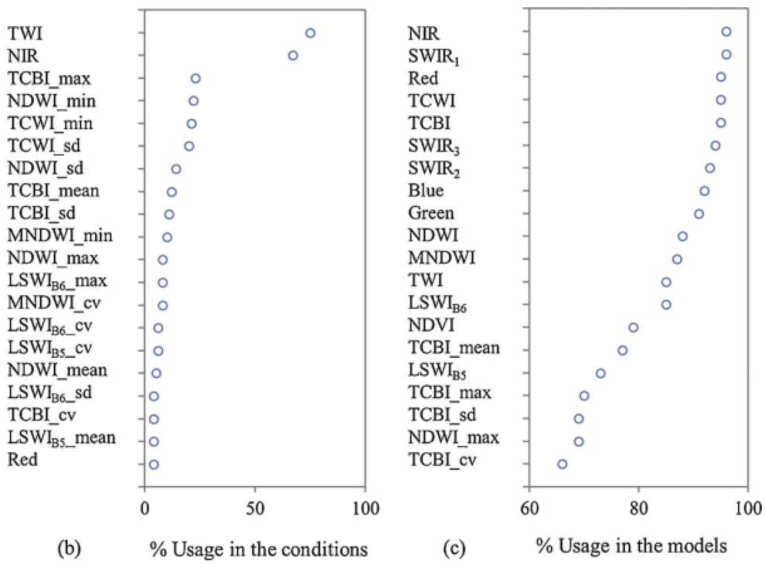

Figure 1. Twenty predictor variables with the highest relative importance for estimation of surface water fraction. Importance is measured as variable usage (%) in the rule conditions (b) and in the linear models (c) with the Cubist model (Li et al. 2018)

Reference: Li, L., Vrieling, A., Skidmore, A., Wang, T., & Turak, E. (2018). Monitoring the dynamics of surface water fraction from MODIS time series in a Mediterranean environment. *International Journal of Applied Earth Observation and Geoinformation*, 66, 135-145.

Changes in manuscript:

P11L1: added "also" and ". These temporal summaries were demonstrated to be an important input"

P11L2: added "in our previous study (Li et al. 2018)."

2) **Comment:** In Page 17, Table 5 listed the comparison results between water extent determined from GSW and MODIS at different thresholds. Though the author mentioned that the generated MODIS surface water product tends to overestimate the water extend due to mixed pixel effects at low water fraction thresholds, it's still confusing that the difference between the two products (MODIS and GSW) doesn't decrease with the increasing threshold higher than 40% or 50%. In fact, the generated MODIS product tends to underestimate the water extent compared with GSW when using larger thresholds.

Response: We thank the reviewer for raising this point. In fact, machine learning approaches such as Cubist regression model and random forest often overestimate small values and underestimate large values when estimating fractional cover of land surface components (e.g. Huang et al., 2014; Li et al. 2018; Wang et al., 2017). As a consequence, our MODIS surface water product tends to overestimate water extent at low water fraction and underestimate water extent at high water fraction. In the previous version of the manuscript, we only discussed the overestimation but not the underestimation. We now have added the reasons why large water fraction covers were underestimated by our MODIS product.

Changes in manuscript:

P16 L26, replaced "The large discrepancy between MODIS surface water fraction and GSW when including low surface water fraction (i.e. threshold =20% and threshold =10%) is probably due to the corresponding mixed pixel effects as described above and also stated by Klein et al. (2017)." with "Nonetheless, our MODIS product detects less surface water compared to GSW for larger thresholds (≥50%), whereas it detects much more surface water than GSW for small thresholds (≤20%). This confirms an earlier finding that machine learning approaches such as Cubist and random forest often underestimate large values and overestimate small values when estimating fractional cover of land surface (e.g. Huang et al. 2014; Li et al. 2018; Wang et al. 2017). In addition to effects of mixed pixels (Klein et al. 2017), the most obvious reason is because regression techniques used in such approaches fit linear equations to relationships that may not be linear over the entire range of values."

3) **Comment:** In Page 20, figure 7 (c) and Page 21, figure 8 (c), the altimetry water levels and MODIS generated water areas were compared. However, it could be more convincing to compare two time series with same physical meaning. The altimetry water levels can be transformed into water areas using hypsometric curves (some can be found in existing data sets such as Hydroweb), vice versa. By doing this you can calculate some metrics to better describe the agreement of two data sources.

Response: We agree with the reviewer that it would be preferable to compare water area with water area, and that hypsometric curves are a way to transform our in-situ water level data in water area. Unfortunately, such curves are to the best of our knowledge not presently available for the lakes presented here. Although arguably these could be constructed using remote sensing data, we prefer to do a direct comparison of our water area estimates with a directly-measured in-situ quantity, i.e. water level, despite that we acknowledge their different physical meanings. Our illustration is mainly intended to demonstrate that the temporal behavior of our estimates with water level correspond, which is to be expected. Rather than a normal regression analysis, a Spearman rank correlation is a better way to assess their relationship. We have now calculated the Spearman rank correlation between water level and water area derived from our MODIS product and JRC's GSW, and added the results in Figure 7 (d) and 8 (d).

Changes in manuscript: P21-22: updated Figure 7-8.

P14L15: added "We calculated the Spearman rank correlation (ρ) between water level and water area derived from MODIS SWF and JRC's GSW data to assess the correspondence between these datasets."

P22L7: added "ρ represents the Spearman rank correlation between water level and water area".

4) **Comment:** The potential of the data set in water management could be better illustrated. High temporal resolution surface water areas can benefit some studies that use water area information as input for hydrological modeling in ungauged basin, e.g., Huang et al. (2018) used river widths generated from Landsat and Sentinel-2 to calibrate the parameters of a distributed hydrology model in Upper Brahmaputra River.

Reference: Huang, Q., Long, D., Du, M., Zeng, C., Qiao, G., Li, X., Hou, A., and Hong, Y.: Discharge estimation in high-mountain regions with improved methods using multisource remote sensing: A case study of the Upper Brahmaputra River, Remote Sensing of Environment, 219, 115-134, 2018

Response: We have incorporated this suggestion and added more content in the Discussion section to better illustrate the potential and application of this dataset. We have also added the suggested paper.

Changes in manuscript: P26L14: added "For example, it could be used as a monitoring tool for analyzing hydrologic extremes such as floods and droughts, detecting abnormal changes of wetland hydrology, capturing short-duration events, identifying newly-formed and disappearing water bodies, and estimating global water loss."

P26L18: added "For example, the water area can help to estimate a series of hydrological parameters such as water discharge (Huang et al. 2018) and water volume (Busker et al. 2019; Cael et al. 2017; Duan and Bastiaanssen 2013; Tong et al. 2016). This would be particularly useful for areas where in situ measurements are sparse or inaccessible."

P26L21: replaced "It may provide new insights" with "Closely monitoring hydrological variability is important"

P26L23: added "It may also provide new insights for understanding how surface water dynamics further influence climate. For example, lake expansion and creation of new dams can alter local and regional precipitation patterns (Ekhtiari et al. 2017; Hossain et al. 2009; Mohamed Degu et al. 2011)."