Reply to comments of Anonymous Referee #1

We thank anonymous referee #1 for reviewing our manuscript. We are especially grateful for the many insightful comments and valuable suggestions, as these comments will lead us to improve the paper. We have to the best of our abilities responded to them and address the referee's comments in the following point by point response. Note the following conventions: RC = referee comments, AC = authors comments (replies) printed in italic.

Major comments

RC 1: This paper calibrates the SWAT model using 2 available ET global products, a simple remote sensing ET equation (MOD16) and a more complex water balance model forced by remote sensing data (GLEAM). MOD16 does not explicitly account for transient water stress (as, say, derived from TIR data); how does this impact the results?

AC1: We agree with the referee that MOD 16 does not explicitly account for transient water stress because it is not directly derived from Thermal Infrared Remote sensing (TIR) data. Some of the reasons for not fully using TIR data at the global scale for the MOD16 product are: a) a changing relationship between TIR based land surface temperature (LST) and NDVI when moving from mid to high latitudes; b) LST as derived from TIR is not equal to the aerodynamic surface temperature (which is driving the sensible heat flux), potentially leading to non-accurate ET estimations under various conditions (Mu et al., 2007, 2013). MOD16 applies the Penman-Monteith (PM) equation to calculate ET on a global scale by using variables and parameters needed from VIS/NIR remote sensing (land cover, LAI, albedo, FPAR) and from daily meteorological reanalysis data (radiation, T_{air}, pressure, rel. humidity; NASA's global modeling and assimilation office, GMAO). In principle, the surface resistance (r_s) parameter in the PM equation accounts for any direct effect on ET due to limitations in available water. The MOD16 scheme however, does not include any soil water content data directly. The way r_s is derived in the MOD16 scheme only considers an indirect effect via a non-linear dependency of r_s with the water vapor pressure deficit (VPD) in the atmosphere. VPD under daytime conditions often represents a proxy for soil moisture conditions and therefore r_s. The impact of the not-explicit consideration of transient water stress in the MOD16 product on our SWAT-model calibration is difficult estimate.

Transient water stress is not a main challenge in the present study area, which is located in the humid region of south western Nigeria with a mean Aridity Index of 0.75 from the period 1989 to 2012 (A.I. > 0.65 value, which is considered to be a humid region; UNEP, 1997).

We also reviewed the literature of the MOD16 ET with measured (EC) flux data at sites climatically similar to our catchment and found an agreement between our catchment and the result obtained between MOD16 ET and the measurements for a study area in the tropical region, Brazil (PDG site in the Rio Grande Basin) conducted by Ruhoff et al. (2013) also located in a tropical region. While not being comprehensive, the comparisons were an indication that MOD16 behave similarly having a positive PBIAS (MOD 16 overestimating AET in both sites). Also, Trambauer et al. (2014) compared different evaporation products for Africa, in their paper they stated that MOD16 evaporation do not show a good agreement with other products in most part of Africa, while the rest (GLEAM, ERAL, PCR-GLOBWB hydrological model simulated AET) are more consistence.

We will fully indicate this comparison in the manuscript and will discuss this point in more detail in the revised version of the manuscript, and also provide the same level of detail for the description of ET in the GLEAM product.

RC2: It is unclear to me whether the SWAT model used here uses the plant growth model.

AC2: The Soil and Water Assessment Tool (SWAT) is an eco-hydrological model that uses at its core the plant growth model EPIC (Williams et al., 1989) that is able to simulate the growth (including nutrient and water uptake) of many types of crops and trees as land cover. The plant growth component of SWAT is a simplified version of the EPIC plant growth model. We will include this sentence in the paper.

RC2b: How is the vegetation taken into account?

AC2b: SWAT is a physically based model that requires a land use map as one input data source that represents the spatial distribution of vegetation in the watershed. SWAT categorizes plants into seven different types: warm season annual legume, cold season annual legume, perennial legume, warm season annual, cold season annual, perennial and trees.

Plant growth is modeled by simulating leaf area development, light interception and conversion of intercepted light into biomass assuming a plant species-specific radiation-use efficiency. Hence, In SWAT, phenological plant development is based on daily accumulated heat units.

The plant growth model is used to assess removal of water and nutrients from the root zone, transpiration, and biomass/yield production. In SWAT the plant growth can be inhibited by a minimum or maximum temperature, available water, nitrogen and phosphorus stress. The potential biomass is based on a method developed by Monteith in 1977 (a radiation model, that uses solar radiation as its decisive factor of crop production, while temperature and water are two other important factors); and a harvest index is used to calculate the final yield. We will include this sentence in the revised manuscript for more clarity.

RC3: Two additional important performance metrics are needed as a reference for the six calibrations:

1. A reference run with default (uncalibrated) parameters – this is needed absolutely!
2. A focus on stressed /unstressed periods as defined by GLEAM ET product, with metric specific for each periods; this would allowed help analyze whether model improvement comes from better ETP formulation or a better simulation of stress.

AC3: 1. the referee is right to point out that a reference run with uncalibrated parameters is needed. We agree with the referee and will add the reference scenario results in the revised manuscript.

AC3: 2. it is true that stressed /unstressed periods should be considered. We included both dry and the wet periods of growth in our calibration and validation. In this paper, we followed the split-sample test as presented by Klemes (1986), which is a model calibration and validation approach that consists of equally splitting the available data, when the record is sufficiently long to represent different climatic conditions. Also, Gan et al. (1997) stated that data are most frequently split by time periods, carefully ensuring that the climate data used for both calibration and validation are not substantially different, i.e., wet, moderate, and dry years occur in both periods. In this study, this we achieved by comparing the absolute AET values of the calibration and validation years, with the aim of having a minimal difference by including wet, moderate, and dry years in both periods. We further checked the linear

regression of the calibration and the validation years. With these two indices, we successfully included wet, moderate and dry years in each calibration and the validation periods. Since the now newly added reference run will include both dry and wet periods, we believe having another separate performance metric accounting for stressed/unstressed periods will be of no need.

RC4: The description of the calibrated parameters (which, I assumes, follow the SWAT terminology) is lacking: there is only a Table; equations showing where those parameters appear should be provided in, say, an annex, to improve the paper standalone readability.

AC4: We believe that the list of the calibrated parameters and where they can be found, as well as the description of the selected parameters and their calibrated optimal values, are of importance and these are provided in Table 2. Including the equations showing where these parameters appear will be too ambitious, because several different equations are formulated for different hydrological conditions for most of the parameters, and this information will be difficult to present in a tabular form. We agree to direct the reader to SWAT documentation by stating that:

"For further reading on the model equations in which the calibrated parameters in SWAT appear, the reader is referred to the in SWAT theoretical documentation version 2009 by Nietsch et al. (2011) (http://swatmodel.tamu.edu)". We also agree to add equations related to the evapotranspiration estimations in the manuscript

Minor comments:

RC5: Figure 2: why use a half-half split sample for MOD16 but only a 1/11 split sample for GLEAM?

AC5: The referee is right to point out this issue and we agreed that this needs clarification in the manuscript. MOD16 is a global dataset spanning the 13-year period 2000-2012 and the splitting of calibration period (2000-2006) and validation period (2007-2012) followed the split-sample test as presented by Klemes (1986) and Gan et al. (1997). While the GLEAM_v3.0a is a global dataset spanning the 35-year period 1980-2014. For this study, we used GLEAM_v3.0a dataset spanning 24-year period 1984-2012 because the SWAT simulation output was from 1989-2012. The splitting of calibration period (1989-2000) and validation period (2001-2012) for GLEAM_v3.0a also followed the split-sample test as presented by Klemes (1986) and Gan et al. (1997). The splitting by time periods was carefully done by ensuring that both MOD16 and GLEAM AET available years dataset used for calibration and validation are not substantially different, i.e., wet, moderate, and dry years occur in both periods.

RC6: Equation 5: the square root should extend to the third quadratic term.

AC6: We agree with the referee and we will make the suggested changes.

RC7: Page 10 line 18: use the term "ratio"

AC7: Thank you! We will make the change on page 10 line 18.

RC8: Page 13 line 22: predicted>predict

AC8: Thank you! We will make the change on page 13 line 22

RC9: Page 15 line 11: Runoff > Ruhoff?

AC9: Thank you! We will make the change on page 15 line 11

RC10: Page 15 line 33: Therefore, the Hargreaves…….periods": I don't understand this sentence

AC9: We agree with the referee that the sentence needs clarification. We will make the changes on page 15 from lines 28-33 to read in the following way:

"The Hargreaves PET equation uses ground measured minimum and maximum temperature to estimate the corresponding AET, whereas GLEAM uses an algorithm to convert PET into AET using a multiplicative evaporation stress factor (S). The derivation of S is based on microwave observation of the vegetation optical depth that is used as a proxy for the vegetation water content and used in simulations of root zone soil moisture. Therefore, the Hargreaves equation in SWAT directly uses measured data and radiation estimates of the area, whereas GLEAM algorithms separately estimate different component of terrestrial evaporation, which consist of evaporation, soil-water balance, stress and rainfall interception (Martens et al., 2016).

To this end, the good GS1 model performance is attributed to the use of Hargreaves equation (that uses the measured temperature data) and the complex water balance model algorithm of GLEAM that takes into account soil-water balance, bare-soil evaporation and open water evaporation, stress and rainfall interception. All of these components have temporal and spatial variations and assist in simulating the dynamic hydrological process, especially the AET, which is the variable of most concern in this study."

Reference

Gan, T. Y., Dlamini, E. M., and Biftu, G. F.: Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. J. Hydrol. 192(1): 81-103, 1997.

Klemes, V.: Operational testing of hydrological simulation models. Hydrological Sciences Journal, v. 31, n. 1, p. 13-24, doi: 10.1080/02626668609491024, 1986.

Courault, D., Seguin,B., and Olioso, A. : Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches, Irrig. Drain. Syst., 19, 223–239 , 2005.

Martens, B., Miralles, D., Lievens, H., Fernández-Prieto, D., and Verhoest, N.: Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture, Int. J. Appl. Earth Obs., 48, 146–162, 2016.

Monteith, J.L.: Climate and the efficiency of crop production in Britain. Philosophical Transactions of the Royal Society of London (Series B-Biological Sciences), 281, 277–294, 1977.

Moran, M. S.: Thermal infrared measurement as an indicator of plant ecosystem health, in Thermal Remote Sensing in Land Surface Processes. Taylor and Francis, Philadelphia, Pa, pp. 257–282, 2003.

Neitsch, S., Arnold, J., Kiniry, J., and Williams, J.: Soil & Water Assessment Tool Theoretical Documentation Version 2009, Texas Water Resour. Inst., 1–647, doi:10.1016/j.scitotenv.2015.11.063, 2011.

UNEP: United Nation Environmental Programme, world atlas of desertification 2ED.UNEP, London, 1997

Ruhoff, A. L., Paz, A. R., Aragao, L. E. O. C., Mu, Q., Malhi, Y., Collischonn, W., Rocha, H. R. and Running, S. W.: Assessment of the MODIS global evapotranspiration algorithm using eddy covariance measurements and hydrological modelling in the Rio Grande basin, Hydrol. Sci. J., 58(8), 1658–1676,

doi:10.1080/02626667.2013.837578, 2013.

Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., Van Beek, L. P. H. and Uhlenbrook, S.: Comparison of different evaporation estimates over the African continent, Hydrol. Earth Syst. Sci., 18(1), 193–212, doi:10.5194/hess-18-193-2014, 2014.

Williams J. R., Jones, C. A., Kiniry, J. R., and Spanel, D. A.: The EPIC Crop Growth Model, Transactions of the American Society of Agricultural Engineers, Vol. 32, No. 2, pp. 497-511, 1989.