Response to the Editor for hess-2018-170

Dear Editor, thank you very much for the opportunity to resubmit a revised copy of our paper on ‘Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data sparse catchment in southwestern Nigeria’. The constructive criticism, comments and suggestions offered by the reviewers have been immensely helpful. We greatly appreciate their insightful comments on revising the paper. The manuscript has been revised to address the reviewer’s concerns and a point-by-point reply to the reviewer’s comments has been made. The changes arising from the comments have clearly improved our manuscript, which you find uploaded alongside this document. All the modifications are highlighted in yellow in the paper as you requested. We look forward to hearing from you in due time regarding our submission and to respond to any further questions and comments you may have.

Response to the Reviewers Comments for hess-2018-170

We thank the anonymous referees #1 and #2 for reviewing our manuscript. We are especially grateful for the many insightful and constructive comments and their valuable suggestions; these changes have clearly improved the quality of the manuscript. We have to the best of our abilities responded to them and address the referees' comments in the following point by point response. Note the following conventions: RC = referee comments, AC = authors comments (replies) printed in italic. All the modifications are highlighted in yellow in the revised manuscript.

Reply to comments of Anonymous Referee #1

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 MOD16 AET 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 AET estimations under various conditions (Mu et al., 2007, 2013). MOD16 applies the Penman-Monteith (PM) equation to calculate AET 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 AET due to limitations in available water. The MOD16 AET scheme however, does not include any soil water content data directly. The way r_s is derived in the MOD16 AET 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 AET 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 have added this paragraph to indicate minimal impact of non-explicit consideration of transient water stress of MOD16 AET on our results in the revised manuscript ***(Page 18, line 24 – 30))***

We also reviewed the literature of the MOD16 AET 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 AET and the measurements for a study area in the tropical region, covered by natural savannah vegetation site in the Rio Grande Basin, Brazil conducted by Ruhoff et al. (2013) also located in a tropical region. While not being comprehensive, the comparisons were an indication that MOD16 AET 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, ECMWF reanalysis ERAL-Land and PCR-GLOBWB hydrological model simulated AET) are more consistence.*

We have fully indicated this comparison and have discussed this point in more detail in the revised version of the manuscript ***(page 18, line 31 – 34 and page 19, line 1- 11)*** *and have also provided the same level of detail for the description of GLEAM AET product* ***(page 5, line 1-7 and page 19, line 22 - 24***

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 have included this sentence in the revised manuscript ***(page 7, line 6 - 13).***

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 have included this sentence in the revised manuscript for more clarity ***(page 7, line 6- 13).***

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. We agree with the referee and we have added the reference scenario results in the revised manuscript ***(page 15, line 8 – 12 and page 33 (Figure 3))***

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 correlation of the calibration and the validation years which shows high value. 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 has included both dry and wet periods, we believe having another separate performance metric accounting for stressed/unstressed periods will not be necessary.*

RC4: The description of the calibrated parameters (which, I assume, 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 important and these are provided in Table 2. Including detailed equations showing where these parameters appear will be too ambitious, because 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. To this effect, we have added in the revised manuscript some equations (equations taken from SWAT theoretical documentation) showing where the 11 most sensitive parameters used in the calibration appear in SWAT ***(Appendix C in page 45 - 47)***

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 for clarification. 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. we have rightly clarified this point in the revised manuscript ***(page 12, line 6 -9, page 12, line 28-33 and page 14, line 2-3)***

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

AC6: We agreed with the referee and we have made the suggested changes ***in page 43: KGE***

RC7: Page 10 line 18: use the term “ratio”

AC7: Thank you! We have made the changes ***in page 43, line 8-9.***

RC8: Page 13 line 22: predicted>predict

AC8: Thank you! We have made the changes in ***page 16, line 11***

RC9: Page 15 line 11: Runoff > Ruhoff?

AC9: Ruhoff et al. (2013) cited in the manuscript is correct. They author a Journal we cited in our manuscript.

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

AC10: We agree with the referee that the sentence needs clarification. We have made the clarification and the changes ***in page 19, line 16-21*** *to read in the following way:*

"The better SWAT model performance in GS1 is attributed to the selection of the Hargreaves equation, which is based on available observed precipitation and maximum and minimum temperature to obtained AET, while the Penman-Monteith and the Priestly-Taylor equations are driven by simulated variables (wind speed, relative humidity and solar radiation) in this study. Also the complex water balance model algorithm of GLEAM takes into account soil-water balance, bare-soil evaporation and open water evaporation, evaporative stress factor and rainfall interception, all of which assist in simulating the dynamic hydrological components, especially the AET"

Reply to comments of Anonymous Referee #2

Major comments

RC (a): Literature Review: it lacks significant contributions in the context of large-scale hydrological model simulation in data scarce area and it mainly focused on previous studies based on SWAT. It could be good to mention and discuss other approaches even if performed in different study areas but with the same problems (data scarce areas) (Kim et al., 2008; Kim and Kaluarachchi, 2009; Gebremicael et al., 2013; Tekleab et al., 2011; Abera et al., 2016 which applied a different hydrological modeling approach (Formetta et al., 2014))

AC (a): Agreed. We have added other related references and discussed other approaches in the context of large-scale hydrological model simulation in data scarce area as suggested ***in page 3, line 31 – 34 and page 4, line 1 - 13.***

RC (b): I feel that the authors should acknowledge explicitly that the analysis presented needs to be tested against observed data and that the satellite data are them self-based on modelling assumptions, which may or may not be plausible in some areas. Of course, they provide a huge help and the way in which they are used in the paper nicely show it, but probably assuming them as "measured" can be misleading. At least can be specified once in the text that "measured AET" doesn't mean proper eddy-covariance data

AC (b): We thank the reviewer for this suggestion. In the revised manuscript, we have strengthened the fact that we only make use of two satellite derived AET products. We have also emphasized that we do not have any e.g. EC-based local measurements within our catchment. However, the satellite products have been tested elsewhere and we have briefly summarised study results that are relevant to our study (similar climate conditions) in a revised version of the manuscript ***in page 5, line 18 – 20 and page 13, line 20 – 22.***

RC (c): In the paper is claimed the importance of the Curve Number parameter but nothing is said about soil moisture evolution and runoff. I wonder why the authors do not use runoff-measured data as independent validation. This will show the effects of the different ET calibration on the runoff dynamic. The two processes are strongly related, and the sensitivity of the CN parameter confirms this. This will be an important added value to the paper. Again, the authors claim: "The average long-term annual of the water balance at the outlet of the study area shows a satisfactory percentage error of closure". Is this referred to modelled data or modelled and measured? The use of measured streamflow data would help to better understand this part as well.

RC (c): It is true that we emphases importance on the curve number parameter because it is found to be sensitive for all the six calibrations. As suggested by the referee, we have mentioned and discussed other parameters relating to soil moisture and runoff considered during the calibration ***in page 17, line 20-32 and in page 18, line 1-16.***

We agree that AET calibration and runoff dynamics are strongly related. In the manuscript, we do not consider runoff-measured data as independent validation because it is not available for the study area and that is main reason we considered AET derived from satellite products as an alternative option for SWAT hydrologic model calibration. We believe using a freely available AET products (GLEAM & MOD16), that have been heavily tested in the past in calibration and validation studies undertaken by a number of scientists is one solution in setting up a hydrological model that will be used as a decision support tool in such a data scarce region. These points have been added and emphased in the revised manuscript ***in page 13, line 13 to 18****. The average long-term annual water balance at the outlet of the study area is referred to the SWAT modelled data with only precipitation and temperature. as measured input data. We assessed the water balance component of the model inorder to ascertain, examine and verify that, SWAT numerical technique and computer code truly represents the conceptual model and that there are no inherent numerical problems with obtaining a solution. We have clarified this in the manuscript that the long-term annual water balance is referred to the SWAT modelled data. The added sentence can be found* ***in page 14, line 22 to 24 and page 30.***

RC (d): Because one of the main points in discussion/conclusion is the fact that: "Hargreaves equation had a superior model performance of the Penman Monteith and the Priestly-Taylor" the authors should add their equations in the text. This would help to visualize the variables in input for each method, the variables that have been chosen for calibration and the variables that have been excluded.

AC (d): We have included the Hargreaves, Penman-Monteith and Priestley-Taylor PET equations in the paper for visualizing of each variables input into each ***method in page 8, Eq.2, Eq.3 and Eq.4***

Specific comments:

RC1: Page 1 line 20: remove space in the number: River Basin (20 292 km2)

AC1: We have left this unchanged. By checking the HESS manuscript preparation guidelines for authors, it is mentioned in the last sentence (h.) under heading "physical dimensions and units", that:

Numerals should also be typeset using upright fonts. The symbol for the decimal marker is the dot. To facilitate reading, numbers may be divided in groups of three using a thin space (e.g. 12 345.6), starting with the ten-thousand digit. Neither dots nor commas are permitted as group separators.

RC2: Page 1 line 21: "The novelty of the study is the use of freely available satellite derived AET data for calibration/validation of each of the SWAT delineated subbasins, thereby obtaining a better performing model at the local scale as well as at the whole watershed level": sounds like this is the first time the gleam dataset have been used to validate/calibrate swat, which is a strong sentence. May be in the study area?

*AC2: We agree with the referee that it is a strong sentence and we acknowledge that the novelty of the study needs further clarification. To this effect we have changed the statement in the revised manuscript **in page 2, line 2- 4** to read as follows:*

"The novelty of the study is the use of these freely available satellite derived AET datasets to calibrate and validate three different SWAT simulated AET for each of the delineated subbasins, to improve the hydrological model performance at both the local and watershed scales for a data-scarce catchment."

*We have also added the paragraph below in the reversed manuscript **in page 5, line 26 to 30** because we believe this is a new contribution both to research community and in the study area:*

"Although the three different PET equations and the corresponding AET simulations from SWAT have been tested for their performance before (Wang et al. (2006); Franco and Bonumá (2017); Samadi (2017); Ha et al. (2018)), the study of calibrating each of the three SWAT simaulated AET variables with two different available remotely sensed derived AET products for each delineated subbasin within SWAT to determine the highest performing model for a particular region has not been undertaken"

RC3: Page 1 line 24: "Three different structures of the SWAT model were used in which each model structure was a set-up of SWAT with a different potential evapotranspiration (PET) equation": I would say that three different PET equations are tested: the model setup (in term of all the single components is the same except the pet).

*AC3: We agree with the referee. As suggested, we have modified the sentence in **page 1, line 20-23, page 9, line 29 - 30** and **page 17, line 17** and throughout the manuscript.*

RC4: Page 2: mechanistic, what the authors mean? Please explain.

*AC4: We agree with the referee that the terminology needs further clarification in the manuscript. we have simplified the terminology in **page 2, line 22 - 23** to read as follows:*

"Numerous physically based distributed (PBD), continuous models that aim to describe which driving processes are present in a system and are able to make detailed predictions in both time and space"

RC5: Page 3 line 25: results showed a good Nash-Sutcliffe efficiency (NSE) and Coefficient of determination (R2) value for mothly average: quantify what good means for the authors and the values obtained.

*AC5: We agree with the referee that the statement needs further quantification and clarification. To this effect we have changed the statement in the revised manuscript **in page 3, line 25 - 28**, to read as follows:*

"The model results showed a high Nash-Sutcliffe efficiency (NSE) of 0.72 and Coefficient of determination (R^2) of 0.76 during the calibration periods. For the validation periods, a high model performance result showing R^2 of 0.71 and NSE of 0.78 for monthly average streamflow were also obtained"

RC6: Page 5: The mean annual rainfall for the watershed is 1224 mm year-1 and the mean annual temperature is about 27o C. Mean annual potential evapotranspiration (PET) estimated by Hargreaves method (Hargreaves and Samani, 1985) is 1720 mm year-1 and the mean AET is about 692 mm year-1. Are this value based on measured or modeled data? Please specify it.

AC6: Many thanks for highlighting this points for clarification. We have changed the statement in the revised manuscript ***in page 6, line 8-12*** *to read as follows:*

The mean annual rainfall (1984-2012) obtained from measured data of Ogun watershed is 1224 mm yr^{-1} *and the mean annual temperature(1984-2012) obtained from measured data is about* 27^{o} *C. Mean annual potential evapotranspiration (PET) estimated by Hargreaves method (Hargreaves and Samani, 1985) using measured minimum and maximum temperature is 1720 mm* yr^{-1} *and the mean AET obtained from SWAT output(1989-2012) for this study area is 692 mm* yr^{-1}*.*

RC7: Page 5 typos: is 1224 mm year-1

AC7: Thank you. We have made the correction ***in page 6, line 9*** *and throughout the manuscript.*

RC8: Page 5: of 4103 ha, and please convert in km^2 because all the other areas are in km

AC8: Thank you. We have made the changes as suggested in the revised manuscript in ***page 6, line 21****.*

RC9: 30 m spatial resolution digital elevation model (DEM), 17 soil classes, 17 landuse classes, 3 slope categories, meteorological data and landuse with its management (Table 1). Please specify if those data are available, from which web-site, and the accessed date

AC9: Thank you. We have added the additional information to ***Table 1 in page 28*** *and in* ***the references in page 23 line 13 - 14, page 23 line 33 - 34, page 24 line 1 - 2 and page 26 line 27 - 28*** *in the revised manuscript*

RC10: Page 7 line 10: “The topHRU program allows the identification of a pareto-optimal threshold which minimizes the spatial error to 0.01 ha for a given number of HRUs and thereby minimizes the trade-off between SWAT computation time and number of HRUs. In this case, topHRU determined the optimum number of HRUs to be 1397 for the Ogun River basin. Thresholds of 0 ha for landuse, 150 ha for soil and 250 ha for slope were used in the SWAT set-up”. What are the physical consequences of the thresholds? What happens if you use larger or lower values? How you define them?

AC10: As explained in the manuscript SWAT uses HRUs, whereby a watershed is subdivided into homogenous hydrologic response units having unique soil, slope, and land use properties as the basic unit of all SWAT model calculations. For SWAT, threshold specification of landcover, soil and slope is allowed and the physical consequences of the thresholds is to improve the computational efficiency of simulations while keeping key landscape features and information of the watershed in the hydrologic modelling. In the paper, we selected thresholds of 150 ha for soil and 250 ha for the slope. This means that HRUs should be created for all the area occupied by the landuse classes. For soil it means that any homogenous soil class occupying less than 150 ha should not be considered when determining HRUs. For slope, it means that any homogenous slope class that occupies less than 250 ha should not be considered. This allows us to define how detail the watershed will be represented by selecting the desired threshold values.

If we select larger values, then we eliminate key landscape features and their processes out of the system which may lead to considerable loss of information about the watershed landscape, resulting in model output that are less representative of the watershed as a whole and if we select lower values then we retain as many landscape features (spatial data) in the model thereby increasing the computational time of SWAT.

Therefore, we defined this threshold by selecting the desired threshold values using topHRU tool and its concept as explained in the paper while minimizing the spatial error to 0.01 ha for a given number *of HRUs. Not selecting a threshold for landuse was based on our desire to retain all of the landuse classes for future landuse change research needs.*

*We have explained the process of threshold selection more clearly **in page 9, line 22-26 in** the revised manuscript.*

RC11: Pae 7: “delineated into 53 subbasins, with the main outlet in Abeokuta”. Can you please give some summary statistics about them: min max average area, elevation, etc. Daily precipitation 5 data (1984-2012) and minimum and maximum temperature data (1984-2012) at four weather stations (Fig. 1) were used as observed input data. Are you only using 4 stations for the whole basin (20292 km2)? Why not considering satellite products for a variable (precipitation), which sometimes could be even more important than etp? The authors should include this in the discussion. The missing values of daily precipitation and minimum and maximum temperatures, along with solar radiation, wind speed and relative humidity were simulated by the ArcSWAT CSFR_World weather generator: it is clear that the ArcSWAT CSFR_World is used for gap filling of precipitation and temperature. The authors should specify: 1) how did you use the dataset for solar radiation, wind speed and relative humidity? 2) At which time resolutions are you specifying that input? 3) For which hydrological processes did you use these “simulated” forcing variables and how this affects your results?

AC11: Many thanks for raising this comment for clarification.

*In SWAT, the Ogun River Basin was delineated into 53 subbasin for this study. The summary statistics of the 53 delineated subbasins is as follows; the minimum and maximum elevation are 23 m and 624 m respectively, while the mean elevation is 289.1 m. The minimum and maximum subbasin area are 72.4 km^2 and 853.1 km^2 respectively, while the mean is 382.8 km^2. We have added this information **in page 9 line 2, 3, 6, and 7** in the manuscript.*

*Daily precipitation data (1984-2012) and minimum and maximum temperature data (1984-2012) obtained from the Nigerian Meteorological Agency for four weather stations (Fig. 1) were used as observed input data. Since the weather stations are more or less evenly distributed in or around the watershed, and the weather data obtained from stations located in the same proximity show the same rise and fall dynamics we were satisfied with the data. No orographic effect correction is needed for correcting the precipitation values. The reasons for using only 4 weather stations in this study have been added **in page 9, line 8 - 11** in the revised manuscript.*

SWAT requires daily values of solar radiation, relative humidity and wind speed in addition to the daily precipitation, minimum and maximum temperature as weather input in SWAT. One out of many options in SWAT to generate this input variables, is to use “WGEN_CFSR_World (ArcSWAT CSFR_World weather generator)” which is an MS Access file containing long-term monthly weather statistics covering the entire globe. In this study, the long-term monthly weather statistics developed using The National Centres for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) global dataset were used to simulate daily solar radiation, relative humidity and wind speed using the WGEN CFSR World. It is also used for filling gaps in measured climate data.

The simulated variables were used as input variables into Penman-Monteith and Priestly-Taylor equations for obtaining the different PET estimates from SWAT.

*The simulated variables allow options for different evaporation estimates which actually affect the results of the model performance during the calibration and validation period as shown in the manuscript. We have added this points **in page 9 line 13 to 18** in the revised manuscript*

RC12: Page7 line22-27: it sounds slightly repetitive: please consider to write the full sentence only of one model structure and to generalize for the other 2.

*AC12: We have made the suggested changes in **page 9, line 29 – 31** and we further explain the meaning of the **acronyms used in figure 2 in page 9, line 30 -32 and page 10, line 1 – 13** (The figure 2 has also been updated to effect the changes **in page 32)** in the revised manuscript to read as follows:*

The SWAT model was set-up once for the entire Ogun River Basin and then run three times, where each model run is composed of a different PET equation available in SWAT (HG, P-M or P-T). Figure 2 shows the framework in which the three SWAT model runs (SWAT_HG, SWAT_P-T, and SWAT_P-M) were used to evaluate the model performance by:

(i) comparing the three uncalibrated SWAT simulations of AET with the two global AET products (GLEAM and MOD16), thus allowing for six reference runs of SWAT (RGS1 through RMS6). SWAT_HG represents the SWAT run using the Hargreaves PET equation to simulate uncalibrated AET, these results were compare with the AET from GLEAM_V3.0a (RGS1) and MOD16 (RMS4). SWAT_P-T represents the SWAT run using the Priestley-Taylor PET equation to simulate uncalibrated AET and the results were compared with the AET from GLEAM_V3.0a (RGS2) and MOD16 (RMS5). SWAT_P-M represents the SWAT run using the Penman-Monteith PET equation to simulate uncalibrated AET and the results were compared with GLEAM_v3.0a (RGS3) and MOD16 (RMS6) and,

(ii) (ii) comparing the calibrations/validations implemented with two global AET products (GLEAM and MOD16), thus allowing for six calibration results of SWAT (GS1 through MS6). SWAT_HG represents the SWAT run using the Hargreaves PET equation to simulate AET and that was calibrated and validated with the AET from GLEAM_v3.0a (GS1) and MOD16 (MS4). SWAT_P-T represents the SWAT run using the Priestley-Taylor PET equation to simulate AET and that was calibrated and validated with the AET from GLEAM_v3.0a (GS2) and MOD16 (MS5). SWAT_P- M represents the SWAT run using the Penman-Monteith PET equation to simulate AET and that was calibrated and validated with the AET from GLEAM_v3.0a (GS3) and MOD16 (MS6).

This procedure enabled the SWAT model run with the highest performing simulated AET to be chosen for further.

RC13: Page 8 line 10: please explain what are the main difference s between the two dataset GLEAM_v3.0a and GLEAM_v3.0b and justify why you selected one of the two.

AC13: Many thanks for raising this comment for more clarification.

The two datasets differ in their forcing variables and their temporal coverage (Martens et al., 2016).

GLEAM_v3.0a is a global dataset that is based on reanalysis net radiation and air temperature, satellite-based vegetation optical length, and a combination of gauge-based, reanalysis satellite-based precipitation. It is a dataset spanning the 35-year period 1980-2014. For this study, we preferred and selected GLEAM_v3.0a dataset spanning 24-year period 1984-2012 because of its long-term availability that allows reasonably selection and splitting of calibration and validation periods that are not substantially different in climatic condition i.e., wet, moderate, and dry years occur in both periods and which covers our SWAT simulation output period (1989-2012).

GLEAM_v3.0b is a global dataset driven by satellite data only and spanning 13-year period 2003-2015. We considered this dataset for the verification of SWAT simulated AET because there are no ground truth AET data in the study area and, because of its different forcing variable, which categories it as an independent dataset not considered in the calibration and validation period.

*We have clarified the differences between GLEAM_v3.0a and GLEAM_v3.****0b in page 5, line 3 - 7*** *and justified why we selected one for model calibration/validation* ***in page 12, line 28 - 33*** *and the other for model verification* ***in page 14, line 19 - 21*** *in the revised manuscript.*

RC14: Page 8 line 25: "was implemented in SWAT-CUP. SWAT-CUP (Abbaspour, 2015)" move the citation when you firstly introduce SWAT-CUP.

AC14: Thank you. We have made the change in ***page 12, line 3*** *as suggested.*

RC15: Pages 8 line 28 to page 9 line 6: Please specify the parameter set that you started the sensitivity analysis with, at least the processes to which they are related. Moreover, specify the list of the parameters that resulted sensitive and how you define a parameter as "sensitive".

AC15: We started the global sensitivity analysis for each of the six-calibration run with the same 50 parameters as shown in the table below. The table below only represent the result of the GS1 parameter sensitivity analysis but the same 50 parameters were used for GS2 through MS6 sensitivity analysis.

In this study as described in the manuscript, the parameter sensitivity is determined by numerous rounds of Latin Hypercube sampling and we defined a parameter as being sensitive when (considering the absolute values) large values of t-stat and smaller values of p-value were obtained, then, the more sensitive the parameter.

The processes to which the parameters are related hydrological are runoff, evaporation, interception, transpiration. In short, each of the hydrological process are represented in the 50 parameters we started the initial global sensitivity with, as shown in table 1 below.

Table 1: The 50 parameters consider in the initial global sensitivity analysis and their relative significance. The table below only show the global sensitivity analysis result of the GS1

Parameter Name	t-Stat	P-Value
45:V__BMX_TREES{..}.plant.dat	0.68	0.62
23:R__SOL_CBN(..).sol	-0.99	0.50
47:V__TMPMX(..).wgn	1.84	0.32
48:V__PCPMM(..).wgn	-2.27	0.26
35:V__LAI_INIT.mgt	2.28	0.26
31:V__SURLAG.bsn	-3.03	0.20
28:R__ALPHA_BF_D.gw	3.21	0.19
19:A__GWQMN.gw	-3.32	0.19
24:R__SOL_ALB(..).sol	-3.82	0.16
37:V__FLOWFR.mgt	-3.86	0.16
40:R__SOL_ZMX.sol	-4.06	0.15
50:R__SOL_Z(..).sol	4.08	0.15
39:V__TLAPS.sub	4.27	0.15
15:V__ESCO.hru	-4.35	0.14
27:V__RCHRG_DP.gw	4.55	0.14
42:V__ALPHA_BF_D.gw	-4.70	0.13
25:R__USLE_K(..).sol	-5.73	0.11
26:V__SOLARAV(..).wgn	-5.96	0.11

38:V__TDRAIN.mgt	6.02	0.10
34:V__BIO_MIN.mgt	6.03	0.10
46:V__TMPMN(..).wgn	-6.18	0.10
3:V__REVAPMN.gw	-6.19	0.10
43:V__RADINC(..).sub	-6.20	0.10
36:V__BIO_INIT.mgt	6.25	0.10
13:V__EVRCH.bsn	-6.91	0.09
20:R__HRU_SLP.hru	6.94	0.09
21:V__GW_DELAY.gw	6.96	0.09
14:V__GW_REVAP.gw	-7.01	0.09
44:V__HUMINC(..).sub	-7.22	0.09
29:V__SHALLST.gw	-7.34	0.09
16:V__CH_N2.rte	7.59	0.08
22:V__ALPHA_BNK.rte	7.66	0.08
17:V__CH_K2.rte	-7.66	0.08
41:V__CH_N1.sub	7.70	0.08
49:V__DEEPST.gw	7.98	0.08
12:V__EVLAI.bsn	8.03	0.08
11:V__OV_N.hru	8.12	0.08
18:V__SFTMP.bsn	8.49	0.07
33:V__GWHT.gw	8.74	0.07
10:V__FFCB.bsn	-8.78	0.07
7:R__SOL_BD(..).sol	-8.91	0.07
6:V__EVRSV.res	-8.95	0.07
32:V__GSI{..}.plant.dat	-9.78	0.06
2:V__EPCO.hru	10.14	0.06
4:R__SOL_K(..).sol	10.51	0.06
9:V__ALPHA_BF.gw	-10.52	0.06
1:V__ESCO.hru	-10.71	0.06
30:V__CANMX.hru	11.52	0.06
5:R__SOL_AWC(..).sol	12.37	0.05
8:R__CN2.mgt	15.79	0.04

After choosing the 11 most sensitive parameters based on the t-stat and p-values. We tried many combinations of the most sensitive analysis e.g. we started with P < 0.09 (37 parameters) the result was not as good at that of the 11 parameters combination. So, we decided to take the 11 most sensitive parameters and we run another global sensitivity analysis to further identify the relative significance of each parameter before calibration. Using 11 parameters gave the most reasonable results. This methodology was extended to all the remaining calibration runs.

We have further clarified the number of parameters we started with, the process to which they are related and how we define the parameter sensitive in the revised manuscript on page ***in page12, line 11 - 19.***

Table 2: The 11 most sensitive parameters consider in the final calibration and validation of all the six calibration runs. The table below only show the sensitivity analysis of the GS1 calibration results.

Parameter Name	t-Stat	P-Value
9:R__SOL_AWC(..).sol	0.01	0.99
2:V__EPCO.hru	-0.21	0.83
7:V__FFCB.bsn	-0.27	0.79
4:V__GSI{..}.plant.dat	0.40	0.69
6:V__EVRSV.res	-0.54	0.59
10:R__SOL_K(..).sol	-0.89	0.37
5:V__ALPHA_BF.gw	1.43	0.15
11:R__SOL_BD(..).sol	-2.01	0.04
3:V__CANMX.hru	2.19	0.03
1:V__ESCO.hru	-2.86	0.00
8:R__CN2.mgt	-23.93	0.00

RC16: Page 9 line 16: A metric among the six can be considered an objective function if it is optimized in the calibration procedure; it can be considered as goodness of fit metric if it is used to quantify how well or bad the model reproduces the measured data. Are those goodness of fit metrics? Which one of these six metrics has been optimized in the calibration procedure? Have you used all of them also as objective function? This is not fully clear.

AC16: Many thanks for raising this comment for more clarification. The Nash-Sutcliffe is the selected objective function that was optimized during the calibration process. This statement has been added for clarity ***in page 14, line 2*** *in the revised manuscript.*

RC17 Page 9 line 20 –Page 10 line 20: Consider to: i) just spell in the text the statistics used, their ranges and their optimal values and ii) move in appendix the explanation of each statistics because they are well known.

AC17: Thank you. We have made the changes as suggested in ***page 13, line 28 – 34 and page 14, line 1****. The equations and their description has been moved to* ***Appendix A in page 43.***

RC18: Page 10 line 23-26: please specify how the uncertainty is quantified: what are the parameters that are changed/sampled the LHS, what are their ranges?

AC18: Thank you for highlighting this point for further clarification. We have emphasized on how the uncertainty was quantified in ***page 14, line 7 – 13.***

As described in the manuscript the uncertainty was quantified by: 1.) assessing the percentage of GLEAM_3.0a AET data bracketed by the model output 95% predictive uncertainty band, the index used for the quantification is P-factor, and 2.) assessing the ratio of the average width of the 95ppu and the standard deviation of the GLEAM_3.0a, the index used for the quantification is R-factor.

The parameters that are changed are listed in Table 2 above by implementing numerous rounds of Latin Hypercube sampling. The optimum values are presented in the revised manuscript ***in Table 2 (page 29)****. In this study, the ranges in which we started the calibration with are presented in* ***Appendix B in the revised manuscript in page 44.***

Table 3a: Eleven parameters and their minimum and maximum range used in this study for the 1st iteration (1000 simulations) for all the six calibration runs.

Parameter name	*Minimum range*	*Maximum range*
v__ESCO.hru	0.00	1.00
v__EPCO.hru	0.00	1.00
v__CANMX.hru	0.00	100.00
v__GSI{2,4,5}.plant.dat	0.00	5.00
v__ALPHA_BF.gw	0.00	1.00
v__EVRSV.res_________17,50	0.00	1.00
v__FFCB.bsn	0.00	1.00
r__CN2.mgt	-0.25	0.85
r__SOL_AWC().sol	0.23	0.95
r__SOL_K().sol	-0.06	0.95
r__SOL_BD().sol	-0.41	0.95

RC19: Page 11: please show a figure of the river basin with the subbasin polygons and the pixel of MODIS and GLEAM. This will help the reader to understand how many pixels of MODIS and GLEAM cover your basins. Each sub-basins has its own model AET. How did you choose the MODIS or GLEAM pixel to compare with and compute the NSE, R^2, etc.

AC19: Thanks for the suggestion, we prepared a figure showing the river basin with the subbasin polygon and the pixel of MOD16 and GLEAM AET. We have inserted the figures as suggested in ***Appendix D in page 48.***

To compare MODIS pixel value (Fig.1) to SWAT simulated AET values from each subbasin for computing the NSE, R^2, PBIAS, KGE, an area-weighted averaging scheme was performed in ArcGIS to create aggregated monthly time-series of AET data of each subbas

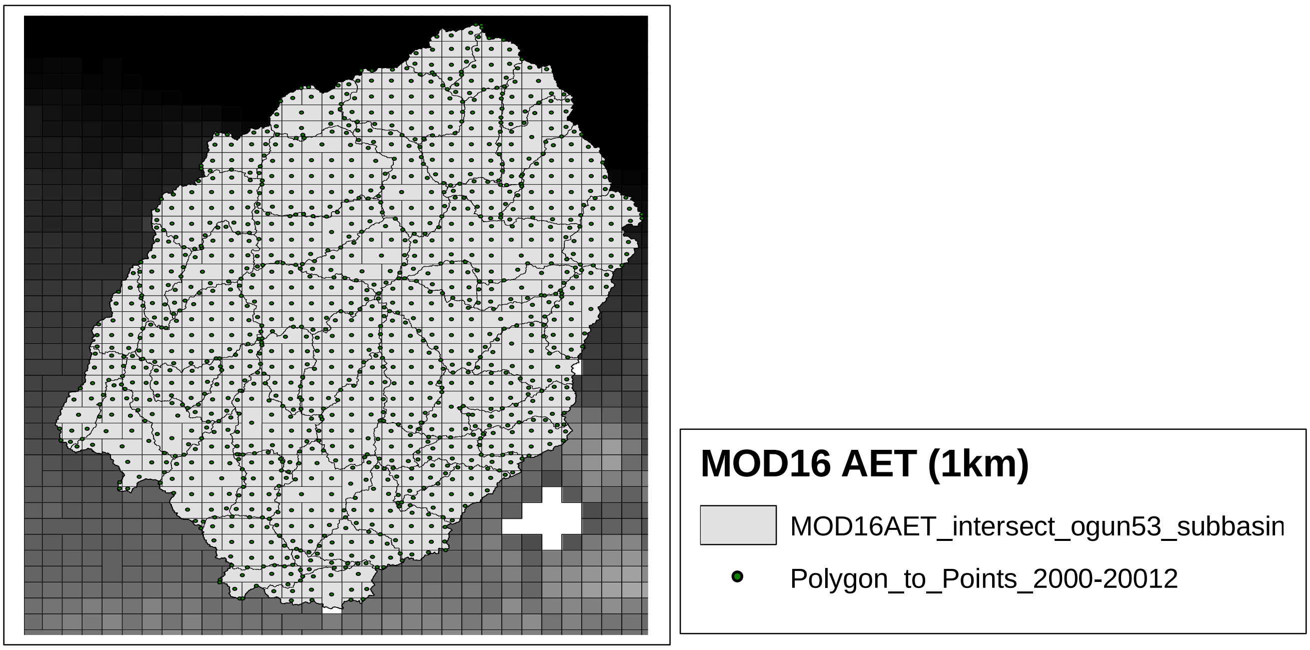

Figure 1: Ogun River Basin with its 53-subbasin polygons intersected by the pixel of MOD16 AET.

The GLEAM_3.0a and GLEAM_3.0b AET is provided in netcdf format, with one file per year and variable. The datasets are available on a 0.25^{0} latitude- longitude regular grid and at daily temporal resolution.

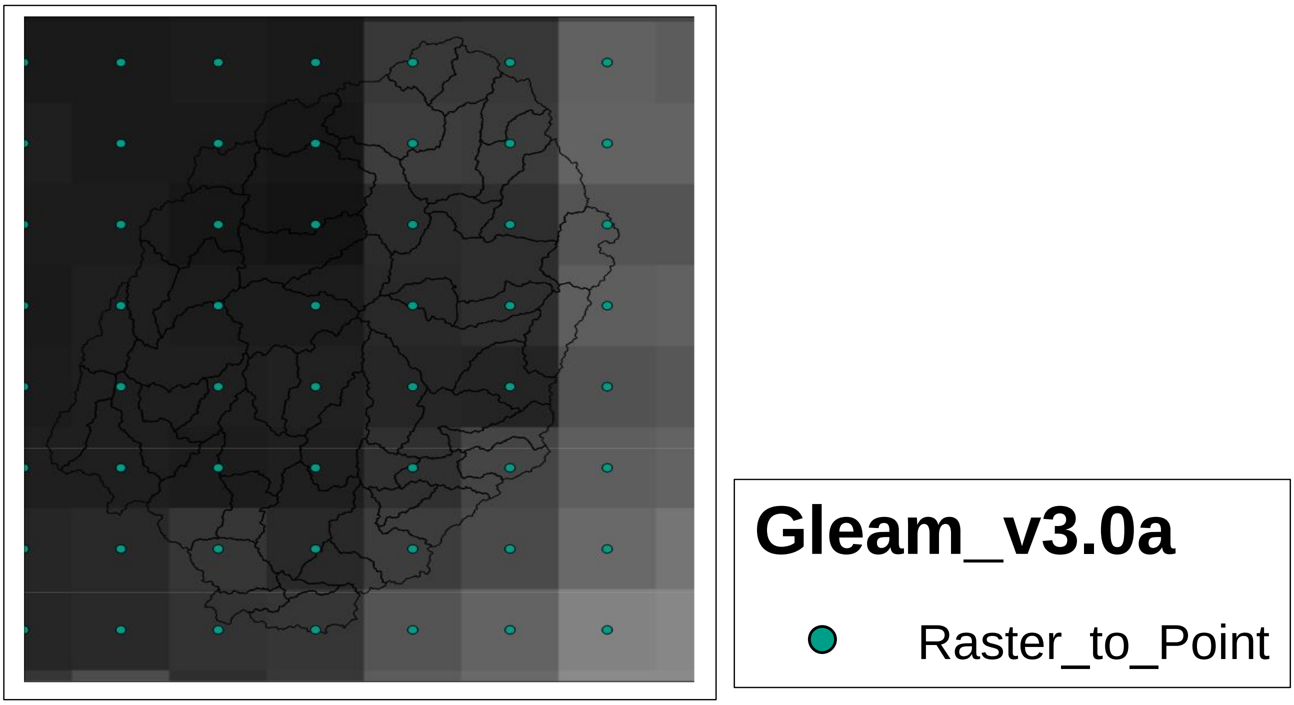

Figure 2: Ogun River Basin with its 53-subbasin polygons intersected by the pixel of GLEAM AET

We used "make NetCDF raster layer" tool in ArcGIS to convert the NetCDF file into a raster layer to view how many pixels of GLEAM cover our subbasins (Fig 2). We realized some points from which the data will be extracted from the pixel are not located in some subbasins (Fig.2). Therefore, we decided to create a point at the center of each subbasin in ArcGIS (Fig. 3). The coordinates of each point at the center of the subbasins were obtained.

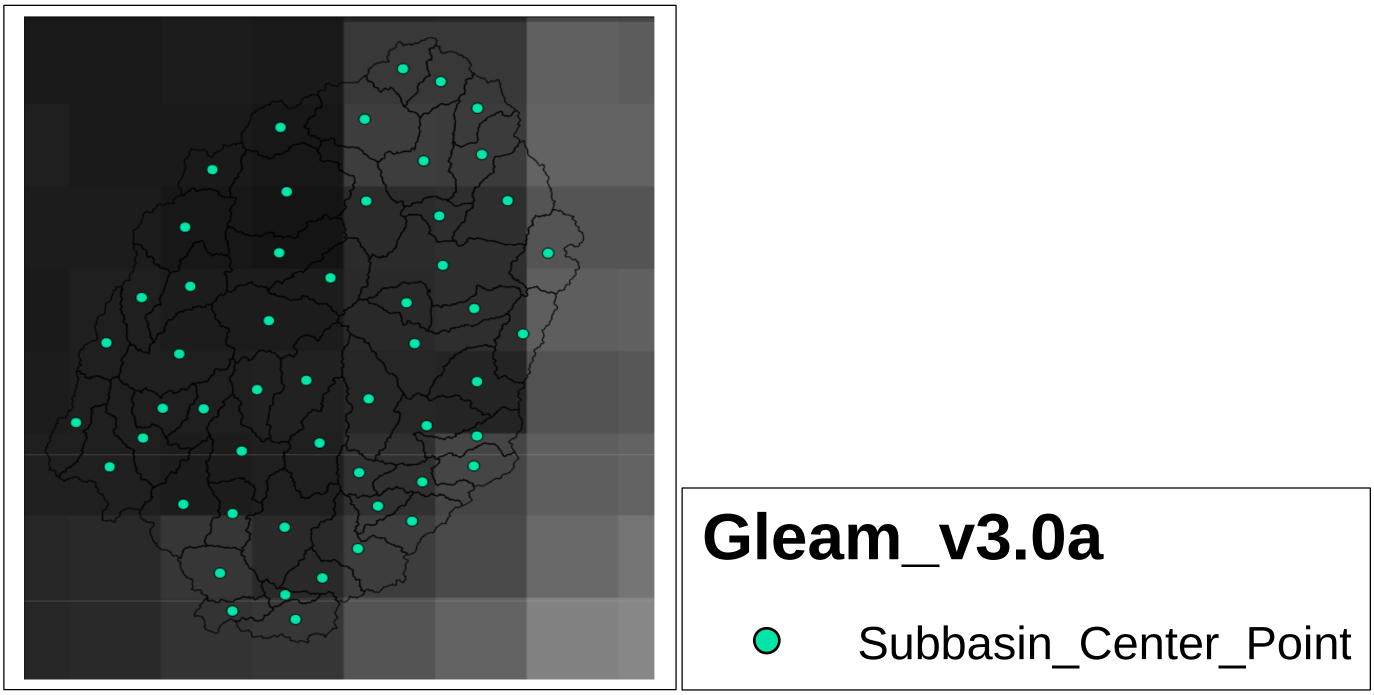

Figure 3: Ogun River Basin with its 53-subbasin polygons intersected by the pixel of GLEAM AET with a point at the center of each subbasin.

To enable us to compare GLEAM pixel value to SWAT simulated AET values for each subbasin for computing the NSE, R^2, PBIAS, KGE, we went back to the NetCDF file of the daily time series (1989-2012) and extracted the GLEAM AET value for each subbasin using "Make NetCDF table view. The values extracted were aggregated to monthly values. The same procedure was carried out for GLEAM_3.0b version. All the clarification has been added in the revised manuscript ***in page 12, line 22 - 28 and page 13, line 3 – 6.***

RC20: Page 11: figure 3, 4, 5, and 6 please consider to add a 4th class for the range of KGE and NSE (<0). This indicates where just the mean of the observed data will be more performing than the model itself

AC20: Many thanks for the suggestion that needs further clarification.

We divided the results into 5 classes. Wish we think the performance rating classes represented with figure 3,4,5,6 for each subbasins are well represented and sufficient

-6.0 – 0 = red

0.10- - 0.50 = purple

0.51 – 0.60 = yellow

0.61-0.70 = orange

0.71 -1.0 = green

The KGE and NSE figure that do not contain a class -6.0 – 0 (<0) do not contain values <0 (fig.3, 4 in the manuscript) and figures that contains -6.0 – 0 (<0) has a model performance that is less than 0 (fig. 5 and fig.6 in the manuscript)

We do not think is necessary to include a class in the legend without having its value in the subbasin polygon.

RC21: Page 10 line 15: equation is not correct please revise it.

AC21: Many thanks for the point. We have made the changes in the revised manuscript in page 43: KGE.

RC22: Page 11: figure 3, 4, 5, and 6 please add the North symbol and the scale bar in the maps.

AC22: Many thanks for the suggestion. We have included North symbol and the scale bar in the revised manuscript ***in page 34 – 37.***

RC23: Page 11 line 23: "The results of global sensitivity analysis revealed that the SCS runoff": please specify from where the reader can see this.

AC23: Thank you. We have included Table 2 at the end of the statement in **page 17, line 17** in the manuscript

RC24: Page 12: all the page can be summarized by just one figure reporting on the x-axis the model configurations and on the y axis the percentage of sub-basin for a given class of the goodness of fit index (NSE, R^2, etc).

AC24: Many thanks for the suggestion. The results presented in page 12 are already summarized in Figures 3, 4, 5,6,7,8 and 9 for clarity. We believe having figures that summarize well the calibration/ validation results of the Ogun River Basin 53 subbasins and that of the whole catchment is necessary for detailed information. We believe having just one figure to report all the results of the calibration/validation procedure might not be sufficient in this case.

To this effect, we have reduced the length of the result presentation described in page **15 and 16** because the figures already shown them.

RC25: Page 13 the paper goes from section 3 to subsection 3.3: 3.1 and 3.2 are missing.

AC25: Thank you for highlighting this point. We have included the subsection 3.1 and 3.2 in the revised manuscript in **page 16, line 9 and page 16, line 22**

RC26: Page 14: increases in February from 55mm to 76mm as the space between 55 and mm

AC26: Thank you for highlighting this point. We have made the corrections in the revised manuscript **in page 16, line 27**.

RC27: Page 15 line 5: "Using the guidelines in Moriasi et al. (2007, 2015) and Kouchi et al. (2017) for" probably these guidelines were drawn for runoff? Is it correct to use it for others hydrological processes? Is this been done in the past? If yes, please add a citation otherwise just clarify this aspect.

AC27: *Many thanks for raising this point for further clarification. The general hydrologic model performance ratings for recommended statistics (NSE, PBIAS, R^2) performed at a monthly time and recommended by Moriasi et al. (2007, 2015) and Gupta et al are mostly relevant for runoff, sediment and nutrients because when these articles were published, the use of remotely sensed evapotranspiration datasets for hydrologic model calibration/validation has not gained much ground.*

In this paper, we also conducted reviewed literatures on model evaluation methods and ratings for model calibration using satellite or non-satellite derived evapotranspiration. Ha et al. (2018) presented a study of calibration of spatially distributed hydrological processes and model parameters in SWAT using remote sensing data and an autocalibration procedure. NSE, R^2, and KGE criteria were used to assess the model performances.

Djman (2016) in their study of evaluation, calibration and validation of six reference ET_0 equation for Senegal River Delta using Penman-Monteith derived ET_0 obtained at saint louis station (1960-2012), uses R^2 and other statistical measures to perform the evaluation. Lopez et al. (2017), calibrated a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. They evaluated the model performance using NSE, PBIAS, KGE and R. Samadi et al. (2016) presented a study on assessing the sensitivity of SWAT physical parameters to potential evapotranspiration estimation *methods over a coastal plain watershed in the southeastern United States, NSE, and KGE statistical measure were used for the model performance assessment.*

All the reviewed literatures set their performance ratings for recommended statistics (NSE, PBIAS, R, KGE, R^2) based on Moriasi et al. (2007, 2015) and kouchi et al. (2017) guidelines.

In this study, we follow Lopez et al. (2017) and others as reviewed above to base our performance rating criteria for judging the model performance by using NSE, R^2, PBIAS and KGE. These points have added in the revised manuscript ***in page 13, line 23 - 31.***

RC28: Page 15 line 20: "From our results, we agree that the AET from MOD16 tends to overestimate AET". Overestimate against what? This is a strong statement mainly because there is not direct comparison against measured AET data

AC28: Thank you for highlighting this point. We agree this is a strong statement that needs to be revised. Actually, we meant that AET from MOD 16 values are higher than that of GLEAM AET and SWAT simulated AET and that this finding agrees with other studies carried out in tropical regions. Since we are using the satellite based MOD16 to calibrate the SWAT AET simulation, the statement needs correction and we have made the corrections in ***page 18, line 24 – 34 and page 19, line 1 – 11.***

Reference

Arnold, J. G., Srinivasan, R., Muttiah, R. S. and Williams, J. R.: Large area hydrologic modeling and assesment Part I: Model development, JAWRA J. Am. Water Resour. Assoc., 34(1), 73–89, doi:10.1111/j.1752-1688.1998.tb05961.x, 1998.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L. J., Donigian, A. S. J. and Johanson, R. C.: Hydrological Simulation Program--Fortran: User's manual for version 11, , 755, doi:EPA/68/C-01/037, 1997.

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.

Djaman,K., Tabari, H., Baide,A.B., Diop,L., Futakuchi,K., and Irmak,S.: Analysis, calibration, and validation of evapotranspiration model to predict grass reference evapotranspiration in Senegal River Delta. Journal of Hydrology: Regional Studies 8, 82-94,2016.

Ewen, J., Parkin, G. and O'Conell, P. E.: SHETRAN : Distributed River Basin Flow Modeling System, J. Hydrol. Eng., 5(JULY), 250–258, doi:doi:10.1061/(ASCE)1084-0699(2000)5:3(250), 2000.

Franco, A. L. and Bonumá, N.B.: Multi-variable SWAT model calibration with remotely sensed evapotranspiration and observed flow. *RBRH*, *22*, e35, 2017.

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.

Ha,L.T, Bastiaanssen, W.G.M, van Griensven, A, van Dijk, A.I.J.M, Senay, G.B.: Calibration of Spatially Distributed Hydrological Processes and Model Parameters in SWAT Using Remote Sensing Data and an Auto-Calibration Procedure: A Case Study in a Vietnamese River Basin. *Water*. 10(2):212, doi: 10.3390/w10020212, 2018.

Hargreaves, G. H. and Samani, Z. A.: Reference Crop Evapotranspiration from Temperature, Appl. Eng. Agric., 1(2), 96–99, doi:10.13031/2013.26773, 1985.

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

Kouchi, D. H., Esmaili, K., Faridhosseini, A., Sanaeinejad, S. H., Khalili, D. and Abbaspour, K. C.: Sensitivity of calibrated parameters and water resource estimates on different objective functions and optimization algorithms, Water (Switzerland), 9(6), 1–16, doi:10.3390/w9060384, 2017.

Lopez Lopez, P., Sutanudjaja, E., Schellekens, J., Sterk, G. and Bierkens, M.: Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products, Hydrol. Earth Syst. Sci. Discuss., (January), 1–39, doi:10.5194/hess-2017-16, 2017.

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.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Binger, R. L., Harmel, R. D. and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Trans. ASABE, 50(3), 885–900, doi:10.13031/2013.23153, 2007.

Moriasi, D. N., Gitau, M. W., Pai, N. and Daggupati, P.: Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria, Trans. ASABE, 58(6), 1763–1785, doi:10.13031/trans.58.10715, 2015.

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.

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

Samadi,S.Z.: Assessing the sensitivity of SWAT physical parameters to potential evapotranspiration estimation methods over a coastal plain watershed in the southeastern United States. IWA, 48 (2) 395-415, doi: 10.2166/nh.2016.034, 2017.

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.

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

Wang, X., Melesse, A. M. and Yang, W.: Influences of potential evapotranspiration estimation methods on SWAT's hydrologic simulation in a northwestern Minnesota watershed, Trans. ASABE, 49(6), 1755–1771, doi:10.13031/2013.22297, 2006.

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.

Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data-sparse catchment in southwestern Nigeria

Abolanle E. Odusanya[1], Bano Mehdi[1, 2], Christoph Schürz[1], Adebayo O. Oke[3], Olufiropo S. Awokola[4], Julius A. Awomeso[5], Joseph O. Adejuwon[5] and Karsten Schulz[1]

[1]Institute of Water Management, Hydrology and Hydraulic Engineering, University of Natural Resources and Life Sciences, Vienna (BOKU), 1190 Vienna, Austria
[2]Department of Crop Sciences, University of Natural Resources and Life Sciences, Vienna (BOKU), 3430 Tulln, Austria
[3]Institute of Agricultural Research and Training, Land and Water Resources Management Programme, Obafemi Awolowo University P.M.B 5029, Moor Plantation, Ibadan, Nigeria
[4]Department of Civil Engineering, College of Engineering, University of Agriculture, P.M.B. 2240, Abeokuta, Nigeria
[5]Department of Water Resources Management and Agrometeorology, College of Environmental Resources Management, University of Agriculture, P.M.B.2240, Abeokuta, Nigeria

Correspondence to: Bano Mehdi (bano.mehdi@boku.ac.at)

Abstract. The main objective of this study was to calibrate and validate the eco-hydrological model Soil and Water Assessment Tool (SWAT) with satellite-based actual evapotranspiration (AET) data from the Global Land Evaporation Amsterdam Model (GLEAM_v3.0a) and from the Moderate Resolution Imaging Spectroradiometer Global Evaporation (MOD16) for the Ogun River Basin (20 292 km^2) located in southwestern Nigeria. Three potential evapotranspiration (PET) equations (Hargreaves, Priestley-Taylor and Penman-Monteith) were used for the SWAT simulation of AET. As reference simulations, the three AET variables simulated before SWAT model calibration, were compared with the GLEAM_v3.0a and MOD16 products. The Sequential Uncertainty Fitting technique (SUFI-2) was used for the SWAT model sensitivity analysis, calibration, validation, and uncertainty analysis. The GLEAM_v3.0a and MOD16 products were subsequently used to calibrate the three SWAT simulated AET variables, thereby obtaining six calibrations/validations at a monthly time scale. The model performance for the three SWAT model runs was evaluated for each of the 53 subbasins against the GLEAM_v3.0a and MOD16 products, which enabled the best model run with the highest performing satellite-based AET product to be chosen. A verification of the simulated AET variable was carried out by: (i) comparing the simulated AET of the calibrated model to GLEAM_v3.0b AET, this is a product that has a different forcing data to version of GLEAM used for the calibration, and (ii) assessing the long-term average annual and average monthly water balances at the outlet of the watershed. Overall, the SWAT model composed of Hargreaves PET equation and calibrated using the GLEAM_v3.0a data performed well for the simulation of AET and provided a good level of confidence for using the SWAT model as a decision support tool. The 95% uncertainty of the SWAT simulated variable bracketed most of the satellite based AET data in each subbasin. The SWAT model also proved efficient in capturing the seasonal variability of the water balance components at the outlet of the watershed. This study demonstrated the

potential to use remotely sensed evapotranspiration data for hydrological model calibration and validation in a sparsely gauged large river basin with reasonable accuracy. The novelty of the study is the use of these freely available satellite derived AET datasets to calibrate and validate three different SWAT simulated AET for each of the delineated subbasins, to improve the hydrological model performance at both the local and watershed scales for a data-scarce catchment.

1. Introduction

Hydrological modelling in data sparse catchments has always been a challenging task due to lack of ground observations, and insufficient or poor quality data. Data scarcity is the main limitation in tropical regions for setting up hydrological models for watershed simulations, which could be used as significant decision support tools for sustainable water resources management. Water resources globally are becoming increasingly vulnerable as a result of escalating water demand arising from population growth, expanding industrialisation, increased food production and pollution due to various anthropogenic activities, climate and land use change impacts (Carroll et al., 2013; McDonald et al., 2014; Goonetilleke et al., 2016). The situation is more evident and critical in many developing countries where no water resources monitoring plans or water management strategies are in place for the future. Like many developing countries, Nigeria cannot satisfy its domestic water needs as only 47% of the total population have access to water from improved sources (Ishaku et al., 2012).

The Ogun River is the main source of public water supply for the people living in the States of Lagos and Ogun in southwestern Nigeria. The prevalent situation of insufficient hydrological data associated with lack of up to date streamflow data (Sobowale and Oyedepo, 2013) and the poor level of data quality in this watershed can be attributed to a gradual decline in hydrological stations number and their management. Water management planners are facing considerable uncertainties in terms of future availability and quality of the water resource. Therefore, a clear understanding of the on-going challenges and innovative management approaches are needed. One of the many ways to tackle this task is by using hydrological models as tools coupled with the use of increasingly available global and regional datasets to run the models.

Numerous physically based distributed (PBD), continuous models that aim to describe which driving processes are present in a system and are able to make detailed predictions in both time and space are available to simulate water quantity and quality variables and these include, among others: the Soil and Water Assessment Tool (SWAT; Arnold et al., 1998), which is able to represent detailed agricultural management practices and simulate water quality and quality variables; or the Hydrologic Simulation Program Fortran (HSPF, Bicknell et al., 1997) that is used in predicting hydrology with in-stream nutrient transport processes; or SHETRAN (Ewen et al., 2000), which has capabilities for modelling subsurface flow and transport. These PBD models attempt to explain hydrological phenomena through their underlying physical mechanisms, and explicitly represent (through mathematical equations) the biological, chemical and physical processes of a basin.

Schuol et al. (2008) have successfully applied the hydrological model SWAT to quantify the freshwater availability for the whole of Africa at a detailed subbasin level and on a monthly time scale. Using the SUFI-2 (Sequential Uncertainty Fitting Algorithm) program with three different objective functions, the model was calibrated and validated at 207 discharge stations.

They reported the models' inability to simulate runoff adequately in some areas in the East and South Africa, but also reported that the model results were quite satisfactory for such a large-scale application although containing large prediction uncertainties in some areas. Many of the limitations reported within this continental modelling study in Africa were data related. Abaho et al. (2009) applied an uncalibrated SWAT model to evaluate the impacts of climate change on river flows and groundwater recharge in Sezibwa catchment, Uganda. They observed a 40% increase in groundwater recharge for the period of 2070-2100 and a 47% increase in average river flow. However, there are high levels of uncertainty associated with the model predictions since the model was not calibrated due to insufficient data.

In West Africa, the SWAT model has been widely applied to different river basins with satisfactory results. For example, Schuol and Abbaspour (2006) applied SWAT to model a 4×10^6 km^2 area; mainly the basins of the Niger, Volta and Senegal, addressing calibration and uncertainty issues. Measured river discharges at 64 stations to which many of these stations available data doesn't cover the whole simulation period were used for annual and monthly calibration using SUFI-2 algorithm. Although the results obtained are preliminary with basis for discussion of further improvement, Schuol and Abbaspour (2006) reported that the annual and monthly simulations with the calibrated SWAT model for West Africa showed promising results for the freshwater quantification despite the modelling shortcomings of lack of dams management operation long-term dataset. They also pointed out the importance of evaluating the conceptual model uncertainty as well as the parameter uncertainty. Laurent and Ruelland (2010) successfully calibrated SWAT for the Bani catchment (1×10^6 km^2) in Mali, a major tributary of the upper Niger River. The calibration and validation results were satisfactory at the catchment outlet and also in various gauging stations located in tributaries. They showed the model performance by reporting discharge and biomass calibration results but did not assess the model prediction uncertainty.

In northwestern Nigeria, Xie et al. (2010) evaluated the SWAT model performance in a large watershed (30 300 km^2). Due to the short data period available, all the data obtained were used for calibration. In their study, the model parameters were first optimized with a genetic algorithm, and the uncertainty in the calibration was further analysed using the generalized likelihood uncertainty estimation (GLUE) method; the study presented a reasonably good calibrated model performance without validation. Adeogun et al. (2014) successfully calibrated and validated the SWAT model for the prediction of streamflow at the upstream watershed of Jebba reservoir (area: 12 992 km^2) located in north central Nigeria. The model results showed a high Nash-Sutcliffe efficiency (NSE) of 0.72 and Coefficient of determination (R^2) of 0.76 during the calibration period. In the validation period, a high model performance result showing R^2 of 0.71 and NSE of 0.78 for monthly average streamflow were also obtained, but the model prediction uncertainty was not quantified.

The findings from these past studies call for continued improvement in the hydrological model performances in Africa, especially in data-sparse regions. One solution is to use freely available global datasets to improve the model performance.

In the context of large scale hydrological model simulation in data scarce areas, Lopez Lopez et al. (2017) investigated alternative ways to calibrate the large scale hydrological model PCRaster GLOBAL Water Balance (PCR-GLOBWB) using satellite-based evapotranspiration (GLEAM) and surface soil moisture (ESA CCI) for the data poor catchments Oum er Rbia in Morocco with the aim to improve discharge estimates. In their study, different calibration scenarios are inter-compared, and the results show that GLEAM evapotranspiration and ESA CCI soil moisture used for model calibration resulted in reasonable discharge estimates (NSE ranges from -0.22 to 0.68 and -0.31 to 0.66, respectively). Although better model performance was achieved when the model was calibrated with in-situ streamflow observations resulting in NSE values from -0.15 to 0.75. Their results showed the possibility of using globally available Earth observations datasets in large scale hydrological models to estimate discharge at a river basin scale. Abera et al. (2017) developed a methodology that can improve the state of the art by using available, but sparse, hydrometeorological data and satellite products to obtain the estimates of all the components of the hydrological cycle (precipitation, evapotranspiration, discharge, and storage) in the Upper Blue Nile Basin. To obtain a water-budget closure, Abera et al. (2017) used the JGrass-NewAge hydrological model calibrated with observed discharge (1994-1999) using the particle swarm (PS) optimization method. The simulation of each hydrological component by JGrass-NewAge model was verified using available in-situ and remote sensing data. GLEAM (Miralles et al., 2011a) and MOD 16 AET were used as independent data sets to assess the JGrass-NewAge estimated AET. Overall, the AET simulations showed that the correlation and PBIAS obtained between JGrass-NewAge and GLEAM AET had a better agreement (very low bias and acceptable correlation) compared to JGrass-NewAge and MOD 16.

Recently, Ha et al. (2018) used remotely sensed precipitation, actual evapotranspiration (AET) and leaf area index (LAI) from open access data sources to calibrate the SWAT model for the Day Basin, a tributary of the Red River in Vietnam. The calibration was performed in SWAT-CUP using the Sequential Uncertainty Fitting algorithm (SUFI-2). In this study simulated monthly AET correlations with remote sensing estimates showed an R^2 of 0.71. Pomeon et al. (2018) set up a hydrological modelling framework for sparsely gauged catchments in West Africa using SWAT model whilst largely relied on remote sensing and reanalysis inputs. In their study, validation of the model was conducted to further investigate its performance, where simulated actual evapotranspiration, soil moisture, and total water storage were evaluated using remote sensing data. The validation result reveals good agreement between predictions and the remotely sensed data (R^2 calibration: 0.52 and 0.51; R^2 validation: 0.63 and 0.61)

Remote sensing technologies offer large scale spatially distributed observations and have opened up new opportunities for calibrating hydrologic models. This advancement enables several global evapotranspiration products to be used. Extensive reviews of earth observation based methods for deriving AET have been carried out by several research groups (Anderson et al., 2012; Bateni et al., 2013; Li et al., 2013; Savoca et al., 2013; Senay et al., 2013; Nouri et al., 2015; Wang-Erlandsson et al., 2016).

Two global-scale AET products derived from satellite observation have become available and these two AET products were used in this study. The Global Land Evaporation Amsterdam Model (GLEAM, http://www.gleam.eu) and Moderate Resolution Imaging Spectroradiometer Global Evaporation (MOD16). GLEAM is an evapotranspiration product developed by the VU University of Amsterdam (Miralles et al., 2011a, 2011b) and contains a set of algorithms that separately estimate the different components of terrestrial evaporation (i.e. transpiration, interception loss, bare soil evaporation, snow sublimation and open water evaporation), as well as variables such as the evaporative stress factor, potential evaporation, root-zone soil moisture and surface soil moisture by using satellite-based climatic and environmental observations (Miralles et al., 2011a;

Martens et al., 2017). Recently, the GLEAM_v3.0 AET has been validated against measurements from 64 eddy-covariance towers and 2338 soil moisture sensors across a broad range of ecosystems with varying levels of success (Martens et al., 2017). In this study, GLEAM_v3.0a and v3.0b were used. These two datasets differ only in their forcing variables and spatial-temporal coverage. GLEAM_v3.0a is a dataset spanning the 35-year period 1980-2014 and is based on reanalysis net radiation and air temperature, a combination of gauged-based, reanalysis and satellite-based precipitation and satellite-based vegetation optical depth. GLEAM_v3.0b is a dataset spanning the 13-year period 2003-2015 and is derived by satellite data only (Miralles et al., 2011a; Martens et al., 2017).

The MOD16 global evapotranspiration data is based on a $1\ \mathrm{km}^2$ grid of land surface AET that was developed with an energy balance model using satellite data as input (Mu et al., 2011). The MOD16 product estimates evapotranspiration using Moderate Resolution Imaging Spectroradiometer, landcover, albedo, LAI, an Enhanced Vegetation Index (EVI), and a daily meteorological reanalysis data set from NASA's Global Modelling and Assimilation office (GMAO). The non-satellite input data are NASA's MERRA GMAO (GEOS-5) daily meteorological reanalysis data. MOD 16 has been validated using measurement from eddy covariance station in different tropical sites (Ruhoff et al., 2013; Ramoelo et al., 2014). Ruhoff et al. (2013) validated MOD16 AET using ground-based measurements of energy fluxes obtained from eddy covariance sites installed in tropical sites in the Rio Grande basin Brazil. Likewise, Ramoelo et al. (2014) validated MOD16 using data from two eddy covariance flux towers installed in a savannah and woodland ecosystem within the Kruger National Park, South Africa.

In this study, we make use of two satellite derived AET products even though there are no eddy covariance based AET measurements within our catchment. However, the satellite-based products we use (GLEAM_v3.0 and MOD16 AET) have been tested in similar climate conditions to ascertain their performances as briefly summarised above.

The objective of our study was to obtain a high performing eco-hydrological model for the Ogun River Basin in southwestern Nigeria that can be used as a decision-support tool. To this effect, the specific objectives were to calibrate/validate the Soil and Water Assessment Tool (SWAT) model with remotely sensed actual evapotranspiration products; namely the Global Land Evaporation Amsterdam Model (GLEAM_v3.0a) and the Moderate Resolution Imaging Spectroradiometer Global Evaporation (MOD16), and also to verify the results of the model AET simulations and the water balance components. Although the three different PET equations and the corresponding AET simulations from SWAT have been tested for their performance before (Wang et al. (2006); Franco and Bonumá (2017); Samadi (2017); Ha et al. (2018)), the study of calibrating each of the three SWAT simaulated AET variables with two different available remotely sensed derived AET products for each delineated subbasin within SWAT to determine the highest performing model for a particular region has not been undertaken.

Hence, the contribution of this study include: (i) the calibration/validation of SWAT simulated AET from the three SWAT model run using the satellite derived AET data sets and (ii) the use of satellite-based actual evapotranspiration data for calibration/validation of the SWAT hydrological model in each of the SWAT delineated subbasins.

2. Materials and methods

2.1 Description of the study site

The study area is a sub-watershed (20 292 km^2) of the Ogun River Basin (23 700 km^2) located in southwestern Nigeria (Fig. 1), bordered geographically by latitudes 7^o 7' N and 8^o 59' N and longitudes 2^o 4' E and 4^o 9' E. About 2 % of the catchment area is located in the Benin Republic. The study area encompasses the Sepeteri, Iseyin, Olokemeji, Oyan and Abeokuta catchments and cut across the Oyo and Ogun state administrative boundaries. The Ogun River, which literarily means the River of Medicine, springs from Igaran Hills in Oyo state, near Saki, at an elevation of about 624 m above the mean sea level. The elevation ranges from 624 m to 23 m. The mean annual rainfall (1984-2012) obtained from measured data of Ogun watershed is 1224 mm yr^{-1} and the mean annual temperature (1984-2012) obtained from measured data is about 27^o C. Mean annual potential evapotranspiration (PET) estimated by Hargreaves method (Hargreaves and Samani, 1985) using measured minimum and maximum temperature is 1720 mm yr^{-1} and the mean AET obtained from SWAT output (1989-2012) for this study area is 692 mm yr^{-1}. Two seasons are distinguishable in the watershed, a dry season from November to March and a wet season between April and October. The watershed area is characterized by strong climatic variation and an irregular rainfall (Eruola et al. 2012). The geology of the study area can be described as a rock sequence that starts with Precambrian Basement; which consists of quartzites and biotite schist, hornblende-biotite, granite and gneisses (Bhattacharya and Bolaji, 2010). The major soils of the basin are sandy clayey loam, sandy loam, clayey loam and silt loam. The landuse in the watershed is primarily forest (75 %), cropland (24 %), and urban (1 %).

The basin, in which two large dams (Oyan and Ikere Gorge dams) are located, is of great importance for the economic advancement both at the federal and state level. The dams are the main principal provider of water to Lagos and Ogun States Water Corporation for municipal drinking water production. The Oyan reservoir is located at the confluence of Oyan and Ofiki rivers at an elevation of 43.3 m above mean sea level and was built in 1984, it has a surface area of 40 km^2, and a catchment area of 9 × 10^3 km^2, with a dead storage capacity of 16 × 10^6 m^3, a gross storage capacity of 270 × 10^6 m^3, an embankment crest length of 1044 m, a height of 30.4 m, four spillway gates (each 15 m wide and 7 m high) and three outlet valves (each 1.8 m diameter). The Ikere Gorge is an uncontrolled dam, which started operation in 1991. The dam crosses Ogun River in Iseyin local government area of Oyo state. Ikere Gorge has a capacity of 690 × 10^6 m^3. The reservoir is adjacent to the Old Oyo National Park, providing recreational facilities for tourists, and the river flows through the park (Oyegoke and Sojobi, 2012). Twenty-five local government areas fall within the study area. In densely populated areas, the Ogun River is used for bathing, washing and drinking.

Fig.1

2.2 SWAT model description

The Soil and Water Assessment Tool (Arnold et al., 1998) is an open source eco-hydrological model developed for the USDA Agricultural Research Services. SWAT is a semi-distributed, process based, continuous model that uses weather, soil, topography and landuse for hydrologic modelling of a basin and runs at a daily time step. It was developed to predict the impact of agricultural land management practices on discharge, sediments, nutrients, bacteria, pesticides and biomass in large complex watersheds with varying soils, land use and management conditions over long periods of time. The SWAT model uses at its core the plant growth model EPIC (Williams et al., 1989) to simulate the growth (including nutrient and water uptake) of many types of crops and trees as land cover. 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 the 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.

For modelling purpose in SWAT, the watershed is divided into subbasins which are then further subdivided into hydrologic response units (HRUs) that consist of homogeneous landuse, soil types and slope (Arnold et al., 1998). The soil water balance (WB) is conducted for each HRU and the equation comprises six variables and is estimated in SWAT using the following Eq. (1):

$$SW_t = SW_0 + \sum_{i=1}^{t}\left(R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}\right) \tag{1}$$

Where SW_t is the final soil water content (mm of water), SW_o is the initial soil water content on day i (mm of water), t is the time (days), R_{day} is the amount of precipitation on day i (mm of water), Q_{surf} the amount of surface runoff on day i (mm of water), E_a the amount of evapotranspiration on day i (mm of water), W_{seep} amount of water entering the vadose zone from the soil profile on day i (mm of water), and Q_{gw} is the amount of return flow on day i (mm of water)

2.2.1 Evaporation estimation in SWAT

Evapotranspiration is a key process of the water balance and one of the more difficult components to determine. Although different empirical methods for the estimation of PET are widely adopted, AET is difficult to quantify and it usually requires the reduction of PET through a factor that describes the level of stress experienced by plants. This relationship has been described in detail by several researches (e.g. Morton, 1986; Hobbins et al., 1999; Wang et al., 2006). Numerous methods have been developed to estimate PET (Lu et al., 2005) and SWAT offers three PET estimation options from which the user can choose depending on e.g. the data availability: the Penman-Monteith method (P-M), the Priestley-Taylor method (P-T), or the Hargreaves method (HG). Any one of these three PET equations can be chosen to run in SWAT, but they vary in the amount of input data required. The Hargreaves method (Hargreaves et al., 1985) is temperature-based and requires only average daily air temperature as input Eq. (2):

$$\lambda E_0 = 0.0023 \times H_0 \times (T_{max} - T_{min})^{0.5} \times (T_{mean} + 17.8) \qquad (2)$$

Where **λ** is the latent heat of vaporization (MJ kg^{-1}), E_0 is the potential evapotranspiration (mm d^{-1}), H_0 is the extraterrestrial radiation (MJ m^{-2} d^{-1}) , T_{max} is the maximum air temperature for a given day (°C), T_{min} is the maximum air temperature for a given day (°C), T_{mean} is the mean air temperature for a given day (°C).

The Penman-Monteith method (Monteith, 1965; Allen, 1986; Allen et al., 1989) requires air temperature, solar radiation, relative humidity and wind speed as input Eq. (3):

$$\lambda E = \frac{\Delta \times (H_{net} - G) + \rho_{air} \times C_P \times (e_z^o - e_z)/r_a}{\Delta + \gamma \times (1 + r_c/r_a)} \qquad (3)$$

Where λE is the latent heat flux density (MJ m^{-2} d^{-1}), E is the depth rate evaporation (mm d^{-1}), Δ is the slope of the saturation vapor pressure-temperature curve, de/dT (kPa °C^{-1}), H_{net} is the net radiation (MJ m^{-2} d^{-1}), G is the heat flux density to the ground (MJ m^{-2} d^{-1}), ρ_{air} is the air density (kg m^{-3}); C_p is the specific heat at constant pressure (MJ kg^{-1} °C^{-1}), e_z^o is the saturation vapor pressure of air at height z (kPa), e_z is the water vapor pressure of air at height z (kPa), γ is the psychrometric constant (kPa °C^{-1}), r_c is the plant canopy resistance (s m^{-1}), and r_a is the aerodynamic resistance (s m^{-1}).

The Priestley-Taylor equation (Priestley and Taylor, 1972) is a radiation-based method and it provides PET estimates for low advective conditions. The P-T method requires solar radiation, air temperature and relative humidity as input (Eq. 4):

$$\lambda E_o = \alpha_{pet} \times \frac{\Delta}{\Delta + \gamma} \times (H_{net} - G) \qquad (4)$$

Where α_{pet} is a coefficient, Δ is the slope of the saturation vapour pressure-temperature curve, de/dT (kPa °C^{-1}), γ is the psychometric constant (kPa °C^{-1}), H_{net} is the net radiation (MJ m^{-2} d^{-1}), and G is the heat flux density to the ground (MJ m^{-2} d^{-1}).

Once PET is determined, AET is estimated in SWAT, whereby first, SWAT evaporates any rainfall intercepted by the plant canopy. Second, it calculates the maximum amount of transpiration and sublimation/soil evaporation. Finally, the actual amount of sublimation and evaporation from the soil surface is calculated. If snow is presented in the HRU, sublimation can occur. When there is no snow (such as this case study), only evaporation from the soil surface is calculated. A complete description of the SWAT model and the model equations can be found in Neitsch et al. (2002, 2005) and in Arnold et al. (1998).

2.3 Model set-up

The ArcView GIS interface for SWAT2012 (Winchell et al., 2013) was used to configure and parameterize the SWAT model. SWAT model inputs included a 30 m spatial resolution digital elevation model (DEM) with minimum, maximum and mean value of 23 m, 624 m and 289.1 m respectively (Fig 1), 17 soil classes, 17 landuse classes, 3 slope categories, meteorological data and landuse with its management (Table 1).

For the SWAT model set-up, the watershed was delineated into 53 subbasins, with the main outlet in Abeokuta. The minimum and maximum subbasin areas are 72.4 km^2 and 853.1 km^2 respectively, while the mean is 382.8 km^2. Daily precipitation data (1984-2012) and minimum and maximum temperature data (1984-2012) obtained from the Nigerian Meteorological Agency for four weather stations (Fig. 1) were used as observed input data. Since the weather stations are more or less evenly distributed in or around the watershed, and the weather data obtained from stations located in the same proximity show the same rise and fall dynamics we were satisfied with the data. No orographic effect correction is needed for correcting the precipitation values. The missing values of daily measured precipitation and minimum and maximum temperatures, were simulated by the WGEN _CFSR _World. The WGEN_CFSR_World weather database is an input into SWAT (ArcSWAT CSFR_World weather generator), containing long-term monthly weather statistics covering the entire globe and developed using the National Centres for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) global dataset. The ArcSWAT CSFR_World weather generator was used to simulate daily solar radiation, wind speed, relative humidity and wind speed. The simulated variables were used as input variables into the Penman-Monteith and Priestley-Taylor equations for obtaining the different PET estimates from SWAT.

The topHRU program (Strauch et al., 2017) was used to determine the optimum number of HRUs to use in the watershed. The topHRU program allows the identification of a pareto-optimal threshold which minimizes the spatial error to 0.01 ha for a given number of HRUs and thereby minimizes the trade-off between SWAT computation time and number of HRUs. In this case, topHRU determined the optimum number of HRUs to be 1397 for the Ogun River Basin. Thresholds of 150 ha for soil and 250 ha for slope were used in the SWAT set-up. The physical consequences of the thresholds is to improve the computational efficiency of simulations while keeping key landscape features and information of the watershed in the hydrologic modelling. Not selecting a threshold for landuse was based on our desire to retain all of the landuse classes for future landuse change research needs. The surface runoff in SWAT was estimated using the modified Soil Conservation Society Curve Number method. The SWATfarmR program (Schürz et al., 2017) was used to write the management files in SWAT. All SWAT simulations included a warm-up period of 5 years for the simulation period from 1984 to 2012.

The SWAT model was set-up once for the entire Ogun River Basin and then run three times, where each model run is composed of a different PET equation available in SWAT (HG, P-M or P-T). Figure 2 shows the framework in which the three SWAT model runs (SWAT_HG, SWAT_P-T, and SWAT_P-M) were used to evaluate the model performance by: (i) comparing the three uncalibrated SWAT simulations of AET with the two global AET products (GLEAM and MOD16), thus allowing for six reference runs of SWAT (RGS1 through RMS6). SWAT_HG represents the SWAT run using the Hargreaves PET equation to simulate uncalibrated AET, these results were compare with the AET from GLEAM_V3.0a (RGS1) and MOD16 (RMS4). SWAT_P-T represents the SWAT run using the Priestley-Taylor PET equation to simulate uncalibrated AET and the results were compared with the AET from GLEAM_V3.0a (RGS2) and MOD16 (RMS5). SWAT_P-M represents the SWAT run using the Penman-Monteith PET equation to simulate uncalibrated AET and the results were compared with GLEAM_v3.0a (RGS3) and MOD16 (RMS6) and, (ii) comparing the calibrations/validations implemented with two global AET products (GLEAM and MOD16), thus allowing for six calibration results of SWAT (GS1 through MS6). SWAT_HG represents the SWAT run using the Hargreaves PET equation to simulate AET and that was calibrated and validated with the AET from GLEAM_v3.0a (GS1) and MOD16 (MS4). SWAT_P-T represents the SWAT run using the Priestley-Taylor PET equation to simulate AET and that was calibrated and validated with the AET from GLEAM_v3.0a (GS2) and MOD16 (MS5). SWAT_P-M represents the SWAT run using the Penman-Monteith PET equation to simulate AET and that was calibrated and validated with the AET from GLEAM_v3.0a (GS3) and MOD16 (MS6). This procedure enabled the SWAT model run with the highest performing simulated AET to be chosen for further use.

Table 1

Fig. 2

2.4 Satellite evaporation dataset

Due to unavailability of discharge measurements in the watershed, two satellites based AET products (GLEAM_v3.0a and MOD16) were used for the SWAT calibration and validation. The criteria for choosing GLEAM and MOD16 products are based on their temporal and spatial resolution and the fact that they are freely available and because these two AET data sets have been validated in several countries in Africa.

2.4.1 GLEAM

The Global Land Evaporation Amsterdam Model (GLEAM) developed in 2011, has been continuously revised and updated. The Priestley and Taylor equation (1972) used in GLEAM calculates the potential evaporation (mm d^{-1}) based on remotely sensed observation of surface net radiation and near-surface air temperature (Eq. 4). Since GLEAM separately derives the different components of terrestrial evaporation (Eq. 5), the estimates of potential evaporation for the land fractions of bare soil, open water, tall canopy and short canopy derived are converted into actual evaporation using a multiplicative evaporative stress factor (Eq. 6) obtained from observations of microwave Vegetation Optical Depth (VOD) used as a proxy for vegetation water content and simulations of root-zone soil moisture. Interception loss is estimated separately based on the Gash analytical model of rainfall interception driven by observations of precipitation and both vegetation and rainfall characteristics.

Two of the three version of the datasets produced in 2016 using GLEAM_v3.0 were downloaded for this study (GLEAM_v3.0a and GLEAM_v3.0b). In this study, GLEAM_v3.0a was used for SWAT calibration and validation while GLEAM_v3.0b was used for the verification of the SWAT simulated AET.

$$E = Et + Eb + Ew + Ei + Es \tag{5}$$

$$S = \frac{(E - Ei)}{Ep} \tag{6}$$

Where E is the actual evaporation (mm d^{-1}), Et is the transpiration (mm d^{-1}), Eb is bare-soil evaporation (mm d^{-1}), Ew is the open-water evaporation (mm d^{-1}), Ei is the interception loss (mm d^{-1}), Es is the snow sublimation (mm d^{-1}), S is the evaporative stress factor (-) and Ep is potential evaporation (mm d^{-1}).

The datasets are provided on a 0.25^0 by0.25^0 regular grid. For more information on GLEAM and the different forcing variables used to produce GLEAM_v3.0a and GLEAM_v3.0b dataset the reader is referred to Miralles et al. (2011b) and Martens et al. (2017).

2.4.2 MOD16

The MOD16 retrieval algorithm (Mu et al., 2007, 2011) is based on the Penman–Monteith framework (Monteith, 1965) with modifications to account for parameters not readily available from space (Cleugh et al., 2007). Terrestrial evapotranspiration includes evaporation from wet and moist soil, evaporation from rain water intercepted by the canopy before it reaches the ground, the sublimation of water vapor from ice and snow and the transpiration through stomata on plant leaves and stems (Mu et al., 2011). Mu et al. (2007) derived actual evaporation from potential evaporation data by using multipliers to halt soil evaporation and plant transpiration through transpiration flow that was limited by water stress and low temperatures and a complementary relationship which defines land-atmospheric interactions from relative humidity and vapour pressure deficit (Mu et al., 2007). Mu et al. (2011) apply the Penman-Monteith (P-M) equation (Eq. 3) to calculate PET 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 from NASA's global modeling and assimilation office (radiation, T_{air}, pressure, relative humidity;). In principle, the surface resistance (r_s) parameter in the P-M equation accounts for any direct effect on evapotranspiration due to limitations in available water. The way r_s is derived in the MOD16 evapotranspiration 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. MOD16 AET is described in detail by Mu et al. (2007, 2011).

2.5 SWAT calibration, validation and uncertainty analysis

A multi-objective calibration and validation of SWAT simulated AET using satellite derived AET from GLEAM_v3.0a and MOD16 was implemented in SWAT-CUP (Abbaspour, 2015). SWAT-CUP is a package used to carry out sensitivity analysis, calibration and validation of the SWAT model. SUFI-2 (Abbaspour et al., 2004) is one of the programs available in SWAT-CUP that is a multi-site, semi-automated, inverse modelling procedure used for calibrating parameters. SUFI-2 is based on a stochastic procedure for drawing independent parameter sets using Latin Hypercube sampling (LHS). In this paper, we followed the split-sample test as presented by Klemes (1986) and Gan et al. (1997), using 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 i.e. wet, moderate, and dry years in both periods. An initial pre-selection of parameters based on literature research (Bicknell et al., 1997; Wang et al., 2006; Rafiei Emam et al., 2016; Ha et al., 2017; Lopez Lopez et al., 2017) was undertaken to choose the most sensitive parameters to AET, and making sure that each of the hydrological processes (runoff, evaporation, interception, transpiration and percolation) are represented in the 50 parameters of the global sensitivity analysis. The initial parameter ranges were based on Neitsch et al.(2002, 2005, 2011). The global sensitivity analysis based on multiple regression method (Abbaspour, 2015) was carried out in which parameter sensitivities are determined by numerous rounds of LHS to obtain the most sensitive parameters by examining the resulting p-value and the t-stat value. The p-value determines the significance of the sensitivity (a value close to zero has more significance) and the t-stat provides a measure of parameter sensitivity (a larger absolute value is more sensitive). Based on the sensitivity analysis, 11 of the most sensitive parameters were selected and altered during calibration process using SUFI-2. The ranking and the calibrated values of the 11 parameters for each of the six calibration procedures are listed in Table 2. The equations written in SWAT theoretical documentation (Neitsch et al. (2011)) showing were the selected 11 sensitive parameters appear are presented in Appendix C.

In this study, the first three calibrations/validations GS1, GS2 and GS3 use the AET from GLEAM_v3.0a for SWAT calibration (1989-2000) and validation (2001-2012). To compare SWAT simulated AET value of each subbasin to the GLEAM_v3.0a and GLEAM_v3.0b AET pixel value and compute their NSE, R^2, PBIAS, KGE for each subbasin, a NetCDF raster layer was created in ArcGIS to view how many pixels of GLEAM covered each of Ogun River subbasin polygon (Fig. D2). Thereafter, a raster to point tool in ArcGIS was used to convert raster cell that intersect each subbasin to point. Since GLEAM AET product is in NetCDF format (daily resolution), the pixel value for each subbasin was extracted using the converted raster to points geographical co-ordinate and by using the ‘Make NetCDF table view’ tool in ArcGIS. The value extracted were aggregated to monthly values. We preferred and selected GLEAM_v3.0a AET for the calibration/validation because of its long-term availability that allows reasonably selection and splitting of calibration and validation periods, which are not substantially different in climatic condition i.e., wet, moderate, and dry years in both periods and which covers our SWAT simulation output period (1989-2012). GLEAM_v3.0a dataset spanning 24-year period 1989-2012 was used 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 AET followed the split-sample test as presented by Klemes (1986) and Gan et al. (1997).

The last three calibration/validation MS4, MS5 and MS6 use the MOD16 AET for SWAT calibration (2000-2006) and validation (2007-2012). Considering MOD16 AET available time-series period, the splitting of calibration and validation period also followed the split sample test as presented by Klemes (1986). Since, MOD16 AET is a raster in geotiff format, to compare the AET pixel value to SWAT simulated AET values for each subbasin for computing the NSE, R^2, PBIAS, KGE, an area-weighted averaging scheme was performed in ArcGIS to create aggregated monthly time-series of AET data for each subbasin.

The three model runs were calibrated (GS1 through MS6) by adjusting the 11 most sensitive parameters found in SUFI-2. In the calibration of SWAT with the AET from GLEAM a sample size of 1000 was chosen for the first iteration and a sample size of 500 for the second iteration, resulting in 1500 simulations. In the calibration of SWAT with AET from MOD16 a sample size of 1000 was chosen for two iterations of LHS, resulting in 2000 simulations. The validation process involved running the model using parameter values that were determined during the calibration process and comparing the SWAT AET simulations to satellite based AET data.

In this study, we do not consider runoff-measured data for an independent validation because it is not available for the study basin and this is the main reason we considered AET derived from satellite products as an alternative option for the SWAT model calibration and validation. We believe using available AET products (GLEAM & MOD16), that have been tested in the past in various calibration and validation studies undertaken by a number of scientists (Roy et al., 2017, Herman et al., 2017, Lopez Lopez et al., 2017, Ha et al., 2018, Pomeon et al., 2018) is one solution in setting up a hydrological model that will be used as a decision support tool in such a data scarce region.

The three SWAT model run calibrated and validated using GLEAM and MOD16 AET (GS1 through MS6) were evaluated with four objective functions, in which their mathematical formulations are presented in Appendix A. It should be noted again that the AET in the GLEAM and MOD16 products does not stem from measured data obtained from eddy covariance instruments, but instead are based on global earth observation products (satellite).

Presently, the general hydrologic model performance ratings for recommended statistics (NSE, PBIAS, R^2) performed at a monthly time and mentioned by Moriasi et al. (2007, 2015) are mostly relevant for runoff, sediment and nutrients. For this study, the literature was searched on model evaluation methods for hydrologic model calibration using satellite or non-satellite derived evapotranspiration. The reviewed literature (Djman, 2016; Samadi et al., 2017; Lopez Lopez et al., 2017; Ha et al., 2018) showed that these studies also set their performance ratings for recommended statistics (NSE, PBIAS and, R^2) based on Moriasi et al. (2007, 2015) guidelines. In this study, we followed Lopez Lopez et al. (2017) and others to base our performance rating criteria for judging the SWAT model performance (GS1 to MS6) by using Nash-Sutcliffe efficiency (NSE, (Nash and Sutcliffe, 1970)), Kling-Gupta efficiency (KGE,(Gupta et al., 2009)), the percent bias (PBIAS) and the coefficient of determination (R^2).

NSE ranges from $-\infty$ to 1, where NSE > 0.5 indicates a good agreement (Moriasi et al., 2007, 2015) between simulated and satellite based evapotranspiration and NSE of 1 being the optimal value. R^2 ranges from 0 to 1 with higher values indicating less error variance and 1 being the optimal value. KGE ranges from $-\infty$ to 1, where KGE of 1 is the optimal value. PBIAS ranges from $-\infty$ to ∞, where low magnitude values indicate better simulation. The optimum value of PBIAS is 0. In this paper, NSE is the selected objective function that was optimised during the calibration process.

The recommended statistics for a monthly time step based on Kouchi et al.(2017) and Moriasi et al.(2007, 2015), states that NSE>0.50, R^2>0.60, KGE≥0.50 and PBIAS ≤ ±25% are the required satisfactory threshold. SUFI-2 was also used for the uncertainty analysis of the AET modelling process. In this step, the procedure depicts the 95% prediction uncertainty (95PPU) of the model compared with satellite based AET. The 95PPU was estimated at the 2.5% and 97.5% levels of the cumulative distribution of the AET simulated output variable derived through LHS. The uncertainties were quantified by two indices referred to as P-factor and R-factor (Abbaspour et al., 2004). The P-factor represents the percentage of observed data plus its error bracketed by the 95% predictive uncertainty (95PPU) band and varies from 0 to 1. Where 1 indicates a 100% bracketing of the observed data within model simulations. While the R- factor is the ratio of the average width of the 95PPU and the standard deviation of the observed variable, this value ranges between 0 and infinity. These two indices were also used to judge the strength of the calibration and validation in which the ideal situation would be to account for 100% of the satellite AET data in the 95PPU while at the same time have an R-factor close to 0.

2.6 SWAT Model Verification

In some modelling studies (EPA, 2013; Faramarzi et al., 2017), the term model verification is used to refer to the examination of the numerical technique and computer code to ascertain that it truly represents the conceptual model and that there are no inherent numerical problems with obtaining a solution. In this study, to further examine the accuracy of the calibrated SWAT model, a verification of the simulated variables was carried out by: (i) a graphical comparison of calibrated SWAT simulated AET to GLEAM_v3.0b AET time-series (2003-2012). We considered GLEAM_v3.0b dataset for the verification of SWAT simulated AET because there are no ground truth AET data in the study area and, because of its different forcing variable, which categorizes it as an independent dataset not considered in the calibration and validation,; and (ii) assessment of the long-term average annual and average monthly water balances at the outlet of the watershed. The long-term average monthly and annual water balance assessment was based on SWAT simulated output with only precipitation and temperature as measured input data. The SWAT water balance equations used for the assessment are:

$$WYLD = SURQ + LAT\_Q + GW\_Q - Q\_TLOSS \tag{7}$$

$$PRECIP = WYLD + AET + \Delta SW + PERC - GW\_Q \tag{8}$$

Where *PRECIP* is the observed precipitation; *AET* is the actual evapotranspiration; *WYLD* is the net amount of water that leaves the subbasin and contributes to stream flow in the reach; *SURQ* is the surface runoff contribution to stream flow; *GW_Q* is the groundwater contribution to stream flow; ΔSW is the change in soil water content; *PERC* is the water percolating past the root zone; *LAT_Q* is the lateral flow contribution to stream flow and *Q_TLOSS* is the transmission loss. The soil water content for both monthly and annual output is the average soil water content for the time period. Hence, the initial soil water content is the average for the time period of 25 years.

3. Results

The results of global sensitivity analysis revealed that the SCS runoff curve number (CN2.mgt) is the most sensitive parameter to SWAT simulations of AET for all the six calibrations (Table 2). The sensitivity ranking of the remaining 10 parameters varies significantly.

Table 2

Figure 3 show the performance of the uncalibrated SWAT (RGS1, RGS2, RGS3, RMS4, RMS5, and RSM6), which represent the reference runs. Results indicate that the uncalibrated SWAT model underestimated AET (positive PBIAS) and has a high percentage deviation from the GLEAM and MOD16 AET. The NSE, KGE, PBIAS and R^2 all depict a low model performance. Interestingly, the reference runs with MOD16 tend to have higher R^2 compared to reference runs with GLEAM. The results from all the six-reference runs justify the need to further improve the SWAT model performance.

Fig. 3

In this paper, only the spatial representations of the SWAT_HG with the highest (GS1) and the lowest (MS6) model performance are included for the purpose of showing the two extreme results obtained (Fig. 4, Fig. 5, Fig. 6 and Fig. 7). The calibration/validation results for GS1 show a model performance of NSE>0.50, KGE>0.50, R^2 >0.6 in more than half of the 53 subbasins and a PBIAS < ±15% in all of the 53 subbasins (Fig.4 and Fig 5). The calibration/validation results for MS6 (Fig.6 and Fig.7) show the lowest model performance.

Fig.4

Fig 5

Fig.6

Fig.7

The remaining calibration/validation results of GS2, GS3, MS4, MS5 (Fig. S1, Fig S2, Fig S3, Fig, S4, Fig S5, Fig S6, Fig S7 and Fig S8) show a lower model performance to GS1.

Figure 8 and 9 summarize the model performance results of the SWAT model runs when calibrated/validated with GLEAM_v3.0a (GS1, GS2 and GS3) and MOD16 AET (MS4, MS5 and MS6). Overall results indicate that the calibration/validation of GS1 (Fig.8 and Fig.9) exhibits a model performance superior to the remaining two model runs for AET simulation (through GS1 to MS6) judging by the four objective functions except for the validation period, where a lower NSE (average value of 0.45) was obtained (GS1, Fig. 9). NSE>0.50 was achieved in 32 out of 53 subbasins during the model validation (GS1), meaning that, more than half of the 53 subbasin have a satisfactory model performance, therefore the average NSE value of 0.45 obtained can be considered acceptable.

Fig.8

Fig.9

3.1 Uncertainty analysis of SWAT model

The SWAT model performance results of the SWAT-HG when calibrated/validated with the AET from GLEAM_v3.0a (GS1) proved to be the most efficient of the three model run (through GS1 to MS6), therefore, it was used to further predict the uncertainty associated with the AET simulations for each of the 53 subbasins to map error sources. In the calibration period, the values of the P-factor obtained were between 0.50 and 0.90 and the values of the R-factor were between 1.40 and 2.4. In the validation period, the values of P-factor were between 0.6 and 0.88, and that of the R-factor were between 1.43 and 2.5. The P-factor values revealed that more than half of the earth observation derived AET plus its error are bracketed by the 95% predictive uncertainty. The predictive uncertainties were adequate in the 53 subbasins and had a satisfactory performance for monthly AET simulations using the Hargreaves equation, though the R-factor was quite large in all the 53 subbasins, indicating large model uncertainties. Extracts of the monthly calibration/validation results showing the 95% prediction uncertainty intervals along with the satellite based AET (GLEAM_v3.0a) are presented in Fig.10.

Fig.10

3.2 Model verification result

It was found that the AET from GLEAM_v3.0b was bracketed within the 95 percent uncertainty prediction (Fig. 10). The long-term average monthly water balance assessment performed at the outlet of the watershed shows a seasonal fluctuation which agrees with previous water balance studies conducted at the outlet of the of the study area located in Abeokuta (Ufoegbune et al., 2011; Eruola et al., 2012; Ufoegbune et al., 2012; Sobowale and Oyedepo, 2013), namely: (i) the study area is characterized by bimodal rainfall pattern, (ii) the AET increases in February from 55 mm to 76 mm as the wet season approaches and decreases in October from 72 mm to 54 mm in November as the dry season approaches (Ufoegbune et al., 2011), (iii) rainfall commences in March (66 mm) and is plentiful in June (165 mm) and September (167 mm), (iv) in August there is a decrease in precipitation to 96 mm and a decrease in AET to 94 mm, the dry spell often experienced in August is termed “August break” (Ufoegbune et al., 2011), (iii) The dry period extends from November to March, the months of low rainfall, AET, and soil moisture values (Ufoegbune et al., 2011), (v) with moderate rain in March soil water increases from 83 mm to 200 mm in July (Ufoegbune et al., 2011), (vi) as dry season commences, the soil water gradually declines.

The differences in the long term mean monthly water balance values obtained in the past studies conducted within the catchment are due to the variation in duration of years considered. Also, Eruola et al. (2012) revealed the two peak rainfalls in July and September agree with the current study, while Ufoegbune et al. (2011) showed the two peak rainfall to be in the month of June and September. All these previous studies and the current study water balance results are in the same range. Figure 12 shows the seasonal fluctuation of the SWAT estimated water balance components at the outlet of watershed, located in Abeokuta town. Our results show, the average long-term annual water balance estimated by SWAT to be within a reasonable percentage error of closure of ±15% (Table 3).

Table 3

Fig. 10

Fig. 11

Fig. 12

4. Discussion

The global sensitivity analysis revealed that for the three SWAT model set-up calibrations (GS1 to MS6), the same 11 SWAT hydrologic parameters governing AET were sensitive (Table 2). When different PET equations were tested in SWAT, different simulated AET values were obtained and the overall sensitivity ranking of the parameters varied significantly. Since parameters represent processes, the significant variation in the sensitivity ranking of the parameters implies that the impact of the selected PET methods in simulating the AET in the study area is relatively high. The SCS runoff curve number for moisture condition II (CN2.mgt) which is one of the parameters that controls the overland processes and a function of the soil's permeability, landuse and antecedent soil water condition was used to determine the surface runoff generation in the basin, is found to be the most sensitive parameter of the three model runs (through GS1 to MS6), indicating that it is also the dominant parameter controlling the AET processes in SWAT for the Ogun River basin (Table 2). The soil evaporation compensation coefficient (ESCO.hru) depends on soil characteristics, controls soil evaporation and it allows modification of depth distribution of water used to meet the soil evaporative demand. As the value of ESCO is reduced, the model is able to extract more of the evaporative demand from the soil lower layer. ESCO is the second most sensitive parameter for GS1 with a very low value of 0.02, but for other results (GS2 to MS6), the sensitivity ranking varies. The maximum canopy storage (CANMAX.hru) was used to account for the amount of water that can be trapped in the canopy when the canopy is fully developed, and this parameter significantly affects infiltration, surface runoff, and evapotranspiration processes. It was more sensitive when the model was calibrated with GLEAM than when the model was calibrated with MOD16 AET. The soil bulk density (SOL_BD. sol) defines the relative amounts of pore space and its overall sensitivity ranking is between 2nd to 4th position (GS1 through MS6). It was more sensitive when SWAT was calibrated with MOD16 than when SWAT was calibrated with GLEAM AET. The baseflow alpha factor (ALPHA_BF.gw) which is the direct index of groundwater flow response to changes in recharge influences the baseflow simulation and its sensitivity varies (GS1 to MS6). The saturated hydraulic conductivity (SOL_K.sol) determines the shallow sub-surface flow and groundwater recharge and it affect the surface runoff response. In order to correctly account for volume of water lost to evaporation from the two reservoirs (Oyan dam and Ikere George dam), the evaporation coefficient (EVRSV.res) was calibrated. Interestingly, we observed that GS1, GS3 and MS6 has higher EVRSV.res values that agrees with the expected values in such a humid tropical region compared to EVRSV.res values obtained from GS2, MS4 and MS5. It is observed that GS1 tends to have the highest (4.7) maximum stomatal conductance (GSI.plant.dat) that denotes the maximum conductance of a leaf when the canopy resistance term is modified to reflect the impact of high vapour pressure deficit, during the calibration. The MS4 has the highest value (0.99) of initial soil water storage expressed as a fraction of field capacity water content (FFCB). The MS4 also obtained the highest value (0.95) of plant uptake compensation factor (EPCO.hru), meaning that the calibration procedure makes the model allows for more of the water uptake demand to be met by lower layers in the soil, and as EPCO approaches 0 (GS3 with the value 0.07) the model allows less variation from the depth distribution. The soil available storage capacity (SOL_AWC.sol) values for GS1 to MS6 were allowed to vary by a factor of 0.8 to 0.96, meaning there is an increase in SOL_AWC for the soil in the model set-up for Ogun River Basin.

Assessing the model performance with the objective function and their satisfactory threshold values used in this study, the calibration/validation with the AET from GLEAM_v3.0a showed an overall satisfactory SWAT model performance when the Hargreaves PET equation was used in SWAT to simulate AET (GS1), compared to the other model calibration results (GS2 to MS6). The calibration/validation with the AET from MOD16 yielded a lower SWAT model performance regardless which of the three PET equations was tested in SWAT.

Using the guidelines in Moriasi et al. (2007, 2015) and Kouchi et al. (2017) for evaluating the SWAT model performance at a monthly time-step, the PBIAS values showed a satisfactory model performance (PBIAS≤±25) in the six calibrations/validations of the three SWAT model runs (Fig. 8 and Fig.9). The positive PBIAS obtained in the calibration/validation of the three SWAT model run using the AET from MOD16 (MS4, MS5 and MS6) indicated a tendency for the SWAT model to underestimate monthly AET at the Ogun River Basin site. An intuitive reason for this may be due to transient water stress occurring in the basin, however, transient water stress is not the main challenge in the study area, which is located in the humid region of southwestern Nigeria with a mean Aridity Index of 0.75 (for the period 1989-2012). The careful consideration of equal wet and dry years in the calibration and validation years has accounted for similar climatic conditions in both periods.

The positive PBIAS result obtained using MOD16 for calibrating agrees with previous studies conducted at a site in tropical region. Ruhoff et al. (2013) validated MOD16 AET using ground-based measurements of energy fluxes obtained from eddy covariance sites installed in tropical sites in the Rio Grande basin, Brazil from a hydrological model (MGB-IPH) at both local and regional scales and found that at the natural savannah vegetation site, the annual AET estimate derived by the MOD16 algorithm was 19% higher than the measured amount. Ruhoff et al. (2013) found that misclassification of land use and land cover was identified as the largest contributor to the error from the MOD16 algorithm. Ramoelo et al. (2014) validated MOD16 using data from two eddy covariance flux towers installed in a savannah and woodland ecosystem within the Kruger National Park, South Africa and found that one flux tower results showed inconsistent comparisons with MOD16 AET and the other site achieved a poor comparison with MOD16 ET. In their study, they found that, the inconsistent comparison of MOD16 and flux tower-based AET can be attributed to the parameterization of the Penman-Monteith model, flux tower measurement errors, and flux tower footprint vs MODIS pixel. Also, Trambauer et al. (2014) compared different evaporation products in Africa and found that MOD16 evaporation does not show a good agreement with other products in most part of Africa, while other evaporation datasets (GLEAM, ECMWF reanalysis ERA-LAND and PCR-GLOBWB hydrological model simulated AET) are more consistent. From our results, we found that when the SWAT model was calibrated with MOD16 AET, the SWAT simulations tend to underestimate AET.

The satisfactory SWAT model GS1 performance was achieved for all objective functions, except for the average NSE value of 0.45 in the validation period, however NSE values >0.50 were obtained in 60% of the subbasins. The KGE result revealed the SWAT-HG model validation (GS1) to be satisfactory (Fig 9). Also, the low PBIAS result of -0.02% and 0.45% (GS1, Fig 8 and Fig 9) corresponded to a performance rating “very good” indicating predictive capability of accurate model simulation. The better SWAT model performance in GS1 is attributed to the selection of the Hargreaves equation, which is based on available observed precipitation and maximum and minimum temperature to obtained AET, while the Penman-Monteith and the Priestly-Taylor equations are driven by simulated variables (wind speed, relative humidity and solar radiation) in this study. Also the complex water balance model algorithm of GLEAM takes into account soil-water balance, bare-soil evaporation and open water evaporation, evaporative stress factor and rainfall interception, all of which assist in simulating the dynamic hydrological components, especially the AET.

The differences in GLEAM and MOD16 products are due to their input and forcing data (Trambauer et al. 2014). Our results agree with another study in which AET from GLEAM performed satisfactorily for the calibration of a large-scale hydrological model set up in Morocco (Lopez Lopez et al., 2017). The 95% predictive uncertainty of the highest SWAT model performance (GS1) was quantified, and the 95% predictive uncertainty bracketed most of the satellite based AET, although the R-factor was quite large in all of the subbasins signifying a large model uncertainty which can be ascribed to the uncertainty in satellite derived AET, the forcing climate data, the conceptual model and the model parameters. The 95PPU are the combined outcome of the uncertainties, these uncertainty sources are not separately evaluated in SUFI-2 but attributed as a total model uncertainty to the parameters which are visualized through the simulated model output ranges. A first verification of the SWAT model run with the best model performance (GS1) was carried out using GLEAM_v3.0b as an independent dataset and found the results to be bracketed within the 95PPU of GS1 (Fig. 11).

The second verification of the SWAT model structure with the best model performance (GS1) was carried out by assessing the output of SWAT water balance components (Table 3 and Fig. 12). The results obtained from the long-term mean monthly water balance agrees with the previous water balance studies conducted within the study area. The differences in the water balance components values of the past and the current study are due to variation in the length of years considered. The average long-term annual of the water balance at the outlet of the study area shows a satisfactory percentage error of closure of ±15% (Table 3) .

5. Conclusion

This study examined an alternative method to calibrate the SWAT eco-hydrological model using available satellite-based AET products for the data-sparse Ogun River Basin in southwestern Nigeria. The calibration approach open-up a new direction for solving the calibration problem of hydrological models in sparsely-gauged and ungauged basins.

Due to the different retrieval algorithms of both the satellite-based AET and SWAT simulated AET, two global evaporation products ((GLEAM and MOD16) were used to calibrate the three SWAT simulated AET on a monthly time scale. The use of Hargreaves, Priestley-Taylor and Penman-Monteith equations in SWAT cause the different AET values obtained. Six different calibrations were implemented with the global AET products with the aim to obtain a high performing model for Ogun-River Basin. Results showed that, an alternative approach to calibrate the SWAT model in tropical ungauged and sparsely-gauged basins can be the use of global AET products. Specifically, when SWAT model was used with the Hargreaves PET equation and was calibrated using the GLEAM_v3.0a AET product, the highest model performance was obtained with an acceptable predictive uncertainty.

Statistical analysis of the model performance shows that global AET datasets used for the calibration were significantly different from each other, which was expected because of their different retrieval algorithms.

Our findings suggest that the SWAT model run using the Hargreaves equation can be used as a potential decision support tool for further studies and predictions on basin hydrology in the Ogun River Basin.

There is still a need for further research on: (i) improving the model calibration performance in those subbasins where the performances are unsatisfactory and (ii) validation of other simulated variable (e.g. stream flow) of the calibrated SWAT model using observed datasets when these are available.

The results from this research contribute to a better understanding of the ease and suitability of using freely available satellite based AET datasets for model calibration in a tropical ungauged basins where the main limitation of setting-up hydrological model for watershed simulations is lack of measured streamflow data, and in a sparsely gauged catchments where the traditional calibration on a limited number of measured streamflow datasets exist. Furthermore, a new contribution of this study is the better understanding of calibration of the three different estimated AET in SWAT to derive the model with the best goodness of fit and a satisfactory predictive capability.

We recommend testing the three PET equations in SWAT to simulated AET whenever SWAT calibration is carried out with any satellite-based AET products. The work presented here is a first step in the hydrological modelling study that will set the basis for future modelling applications within the study basin.

Author Contributions: Abolanle E. Odusanya, Bano Mehdi and Karsten Schulz designed the methodological framework and advised and contributed to the entire strategic and conceptual framework of the study. Abolanle E. Odusanya performed the simulations, analyzed the results and prepared the manuscript under supervision of Bano Mehdi and Karsten Schulz. Christoph Schürz prepared the landuse and soil maps, wrote the SWATfarmR script, and modified topHRU code for this study. Adebayo O. Oke, carried out the field work for point source water pollution data collection used in the SWAT configuration. Olufiropo. S. Awokola, Julius A. Awomeso and Joseph O. Adejuwon carried out the field work for the necessary data input for the two reservoirs used in the SWAT configuration.

6. Acknowledgements

The authors wish to thank those persons, institutions, authorities and agencies, especially Olusegun Orekoya (NIMET) that assisted in accessing research data and information.

Conflicts of Interest: The authors declare no conflict of interest

Reference

Abaho, P., Amanda, B., Kigobe, M., Kizza, M. and Rugumayo, A.: Climate Change and its Impacts on River Flows and Recharge in the Sezibwa Catchment, Uganda, Second Int. Conf. Adv. Eng. Technol., 572–578, 2009.

Abbaspour, K. C.: SWAT-CUP: SWAT Calibration and Uncertainty Programs- A User Manual,Department of Systems Analysis,Intergrated Assessment and Modelling (SIAM),EAWAG. Swiss Federal Institute of Aqualtic Science and Technology,Duebendorf,Switzerland., User Man., 100p, doi:10.1007/s00402-009-1032-4, 2015.

Abbaspour, K. C., van Genuchten, M. T., Schulin, R. and Schläppi, E.: A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters, Water Resour. Res., 33(8), 1879, doi:10.1029/97WR01230, 1997.

Abbaspour, K. C., Johnson, C. A. and van Genuchten, M. T.: Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure, Vadose Zo. J., 3(4), 1340–1352, doi:10.2136/vzj2004.1340, 2004.

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H. and Kløve, B.: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model, J. Hydrol., 524, 733–752, doi:10.1016/j.jhydrol.2015.03.027, 2015.

Abera, W., Formetta, G., Brocca, L., and Rigon, R.: Modeling the water budget of the Upper Blue Nile basin using the JGrass-NewAge model system and satellite data, Hydrol. Earth Syst. Sci., 21, 3145-3165, doi:/10.5194/hess-21-3145-2017, 2017.

Adeogun, A. G., Sule, B. F., Salami, A. W. and Okeola, O. G.: Gis- Based Hydrological Modelling Using SWAT : Case Study of Upstream Watershed of Jebba Reservoir in Nigeria, Niger. J. Technol., 33(3), 351–358, doi:10.4314/njt.v33i3.13, 2014.

Allen, R. G.: A Penman for all seasons, J. Irrig. Drain. Eng., 112(4), 348–368, doi:http://dx.doi.org/10.1061/(ASCE)0733-9437(1986)112:4(348), 1986.

Allen, R.G., Jensen, M.E., Wright, J.L., and Burman, R.D., 1989. Operational estimates of reference evapotranspiration. Agron. J., 8 1: 650-662, 1989.

Anderson, M. C., Allen, R. G., Morse, A. and Kustas, W. P.: Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources, Remote Sens. Environ., 122, 50–65, doi:10.1016/j.rse.2011.08.025, 2012.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. and Williams, J. R.: Large area hydrologic modeling and assesment Part I: Model development, JAWRA J. Am. Water Resour. Assoc., 34(1), 73–89, doi:10.1111/j.1752-1688.1998.tb05961.x, 1998.

Bateni, S. M., Entekhabi, D. and Castelli, F.: Mapping evaporation and estimation of surface control of evaporation using remotely sensed land surface temperature from a constellation of satellites, Water Resour. Res., 49(2), 950–968, doi:10.1002/wrcr.20071, 2013.

Bhattacharya, A. K. and Bolaji, G. A.: Fluid flow interactions in Ogun River, Nigeria, Int. J. Res. Rev. Appl. Sci., 2(2), 173–180, 2010.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L. J., Donigian, A. S. J. and Johanson, R. C.: Hydrological Simulation Program--Fortran: User’s manual for version 11, 755, doi:EPA/68/C-01/037, 1997.

Franco, A. L., and Bonumá, N.B.: Multi-variable SWAT model calibration with remotely sensed evapotranspiration and observed flow. *RBRH*, *22*, e35, doi:10.1590/2318-0331.011716090, 2017.

Carroll, S., Liu, A., Dawes, L., Hargreaves, M. and Goonetilleke, A.: Role of Land Use and Seasonal Factors in Water Quality Degradations, Water Resour. Manag., 27(9), 3433–3440, doi:10.1007/s11269-013-0356-6, 2013.

Djaman,K., Tabari, H., Baide,A.B., Diop,L., Futakuchi,K., and Irmak,S.: Analysis, calibration, and validation of evapotranspiration model to predict grass reference evapotranspiration in Senegal River Delta. J. Hydrol.Reg. Stud., 8, 82-94, doi: 10.1016/j.ejrh.2016.06.003, 2016.

Cleugh, H. A., Leuning, R., Mu, Q. and Running, S. W.: Regional evaporation estimates from flux tower and MODIS satellite data, Remote Sens. Environ., 106(3), 285–304, doi:10.1016/j.rse.2006.07.007, 2007.

EPA: Watershed Modeling, EPA's Watershed Acad. Web, 1–35, 2013.

Eruola, A.O., Ufeogbune, G.C., Eruola, A.A., Idowu, O.A., Oluwasanya, G.O., Ede, V. A.: Effect of Climate Change on Water Balance of Lower Ogun River Basin, , 360–367, 2012.

ESA CCI LC (2014): European Space Agency Climate Change Initiative Land Cover Maps Project. <https://www.esa-landcover-cci.org/?q=node/158> (accessed 20.09.2016)

Ewen, J., Parkin, G. and O'Conell, P. E.: SHETRAN : Distributed River Basin Flow Modeling System, J. Hydrol. Eng., 5(JULY), 250–258, doi:doi:10.1061/(ASCE)1084-0699(2000)5:3(250), 2000.

Faramarzi, M., Abbaspour, K. C., Adamowicz, W. L. V., Lu, W., Fennell, J., Zehnder, A. J. B. and Goss, G. G.: Uncertainty based assessment of dynamic freshwater scarcity in semi-arid watersheds of Alberta, Canada, J. Hydrol. Reg. Stud., 9, 48–68, doi:10.1016/j.ejrh.2016.11.003, 2017.

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, doi: 10.1016/S0022-1694(96)03114-9, 1997.

Goonetilleke, A., Liu, A. and Gardner, T.: Urban Stormwater Reuse : an Agenda for Sustainable, , 1–4, 2016.

Gupta, Hosshin Vijai; Soroosh, S. and Patrice Ogou, Y.: Status of Automatic Calibration For Hydrologic Models: Comparison with Multilevel Expert Calibration, , doi:10.1061/(ASCE)1084-0699(1999)4, 1999.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377(1–2), 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Ha, L.T., Bastiaanssen, W.G.M., van Griensven, A., van Dijk, A.I.J.M., and Senay, G.B.: Calibration of Spatially Distributed Hydrological Processes and Model Parameters in SWAT Using Remote Sensing Data and an Auto-Calibration Procedure: A Case Study in a Vietnamese River Basin. *Water*. 10(2):212, doi: 10.3390/w10020212, 2018.

Hargreaves, G. H. and Samani, Z. A.: Reference Crop Evapotranspiration from Temperature, Appl. Eng. Agric., 1(2), 96–99, doi:10.13031/2013.26773, 1985.

Hengl, T., Heuvelink, G.B.M., Kempen, B., Leenaars, J.G.B., Walsh, M.G. and Shepherd, K.D., et al. (2015).: Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. PLoS ONE 10(6):

e0125814,doi:10.1371/journal.pone.0125814.<http://www.isric.org/projects/soil-property-maps-africa-250-m-resolution> (accessed 05.10.16)

Herman, M.R., Nejadhashemi, Abouali, A.P., Hernandez-Suarez, J.S., Daneshvar, F., Zhang, F., Anderson, M.C., Sadeghi, A.M., Hain, C.R., and Sharif, A.: Evaluating the role of evapotranspiration remote sensing data in improving hydrological modeling predictability. J.Hydrol. 556: 39–49, doi: 10.1016/j.jhydrol.2017.11.009, 2017.

Hobbins, M. T., Ramírez, J. a and Brown, T. C.: The complementary relationship in regional evapotranspiration: the CRAE model and the Advection-Aridity approach, Hydrol. Days, 37(5), 1–16, doi:10.1029/2000WR900359, 1999.

Ishaku, H. T., Majid, M. R. and Johar, F.: Rainwater Harvesting: An Alternative to Safe Water Supply in Nigerian Rural Communities, Water Resour. Manag., 26(2), 295–305, doi:10.1007/s11269-011-9918-7, 2012.

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

Kouchi, D. H., Esmaili, K., Faridhosseini, A., Sanaeinejad, S. H., Khalili, D. and Abbaspour, K. C.: Sensitivity of calibrated parameters and water resource estimates on different objective functions and optimization algorithms, Water (Switzerland), 9(6), 1–16, doi:10.3390/w9060384, 2017.

Laurent, Ruelland, F. L.: Modelisation ` a base physique de la variabilit ´ e hydroclimatique ` a l ' ´ echelle d ' un grand bassin versant tropical To cite this version : Modélisation à base physique de la variabilité hydroclimatique à l ' échelle d ' un grand bassin versant trop, 2010.

Li, Z. L., Tang, B. H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F. and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sens. Environ., 131, 14–37, doi:10.1016/j.rse.2012.12.008, 2013.

López López, P., Sutanudjaja, E. H., Schellekens, J., Sterk, G., and Bierkens, M. F. P.: Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products, Hydrol. Earth Syst. Sci., 21, 3125-3144, doi.org/10.5194/hess-21-3125-2017, 2017.

Lu, Jianbiao, Ge Sun, Steven G, McNulty, and D. M. A.: A Comparison of Six Potential Evapotranspiration Methods For Regional USe in the Southeastern United States, 2005.

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A. and Verhoest, N. E. C.: GLEAM v3: Satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10(5), 1903–1925, doi:10.5194/gmd-10-1903-2017, 2017.

McDonald, R. I., Weber, K., Padowski, J., Fl??rke, M., Schneider, C., Green, P. A., Gleeson, T., Eckman, S., Lehner, B., Balk, D., Boucher, T., Grill, G. and Montgomery, M.: Water on an urban planet: Urbanization and the reach of urban water infrastructure, Glob. Environ. Chang., 27(1), 96–105, doi:10.1016/j.gloenvcha.2014.04.022, 2014.

Miralles, D. Gonzalez, van den Berg, M.J., Gash, J.H., Parinussa, R.M., de Jeu, R.A.M. , Beck, H.E., Holmes, T.R.H., Jimenez, C., Verhoest, N.E.C. , Dorigo, W.A., Teuling, A.J., Dolman, A. J.: El Nino-La Nina cycle and recent trends in continental evaporation, 2014.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A. and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci., 15(2), 453–469, doi:10.5194/hess-15-453-2011, 2011a.

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H. and Dolman, A. J.: Magnitude and variability of land evaporation and its components at the global scale, Hydrol. Earth Syst. Sci., 15(3), 967–981, doi:10.5194/hess-15-967-2011, 2011b.

Monteith, J. L.: Evaporation and environment, Symp. Soc. Exp. Biol., 19, 205–234, doi:10.1613/jair.301, 1965.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Binger, R. L., Harmel, R. D. and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Trans. ASABE, 50(3), 885–900, doi:10.13031/2013.23153, 2007.

Moriasi, D. N., Gitau, M. W., Pai, N. and Daggupati, P.: Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria, Trans. ASABE, 58(6), 1763–1785, doi:10.13031/trans.58.10715, 2015.

Morton, F. I.: Pratical Estimates of Lake Evaporation, 371–387, 1986.

Mu, Q., Heinsch, F. A., Zhao, M. and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sens. Environ., 106(3), 285–304, doi:10.1016/j.rse.2006.07.007, 2007.

Mu, Q., Zhao, M. and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sens. Environ., 115(8), 1781–1800, doi:10.1016/j.rse.2011.02.019, 2011.

Nash, I. E. and Sutcliffe, I. V: River flow forecasting through conceptual models, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.

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.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Williams, J. R. . R. and King, K. W.: Soil and water assessment tool theoretical documentation, Texas Water Resour. Inst., 494 [online] Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-0011239709&partnerID=tZOtx3y1, 2002.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R. and Williams., J. R.: Soil and Water Assessment Tool User’s Manual Version 2005., Diffus. Pollut. Conf. Dublin, 494, 2005.

Nouri, H., Beecham, S., Anderson, S., Hassanli, A. M. and Kazemi, F.: Remote sensing techniques for predicting evapotranspiration from mixed vegetated surfaces, Urban Water J., 12(5), 380–393, doi:10.1080/1573062X.2014.900092, 2015.

Oyegoke, S. . and Sojobi, A. .: Developing Appropriate Techniques to Alleviate the Ogun River Network Annual Flooding Problems, Int. J. Sci. Eng. Res., 3(2), 1–7, 2012.

Poméon, T., Diekkrüger, B., Springer, A., Kusche, J., and Eicker, A.: Multi-Objective Validation of SWAT for Sparsely-Gauged West African River Basins—A Remote Sensing Approach. *Water*. 10(4):451, doi: 10.3390/w10020212, 2018.

Priestley, C. H. B. and Taylor, R. J.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Mon. Weather Rev., 100(2), 81–92, doi:10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, 1972.

Rafiei Emam, A., Kappas, M., Hoang Khanh Nguyen, L. and Renchin, T.: Hydrological Modeling in an Ungauged Basin of Central Vietnam Using SWAT Model, Hydrol. Earth Syst. Sci. Discuss., (February 2016), 1–33, doi:10.5194/hess-2016-44, 2016.

Ramoelo, A., Majozi, N., Mathieu, R., Jovanovic, N., Nickless, A. and Dzikiti, S.: Validation of global evapotranspiration product (MOD16) using flux tower data in the African savanna, South Africa, Remote Sens., 6(8), 7406–7423, doi:10.3390/rs6087406, 2014.

Roy, T., Gupta, H. V., Serrat-Capdevila, A., and Valdes, J. B.: Using satellite-based evapotranspiration estimates to improve the structure of a simple conceptual rainfall–runoff model, Hydrol. Earth Syst. Sci., 21, 879-896, doi: 10.5194/hess-21-879-2017, 2017.

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.

Savoca, M. E., Senay, G. B., Maupin, M. A., Kenny, J. F. and Perry, C. A.: Actual Evapotranspiration Modeling Using the Operational Simplified Surface Energy Balance (SSEBop) Approach, U.S Geol. Surv. Sci. Investig. Rep. 2013–5126, 16 p., 2013.

Samadi, S.Z.: Assessing the sensitivity of SWAT physical parameters to potential evapotranspiration estimation methods over a coastal plain watershed in the southeastern United States. IWA, 48 (2) 395-415, doi: 10.2166/nh.2016.034, 2017.

Schuol, J. and Abbaspour, K. C.: Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa, Adv. Geosci, 9, 137–143, doi:10.5194/adgeo-9-137-2006, 2006.

Schuol, J., Abbaspour, K. C., Srinivasan, R. and Yang, H.: Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model, J. Hydrol., 352(1–2), 30–49, doi:10.1016/j.jhydrol.2007.12.025, 2008.

Schürz, C., Strauch, M., Mehdi, B. and Schulz, K.: SWATfarmR : A simple rule-based scheduling of SWAT management operations,in: Proceedings of the 2017 Int. SWAT Conf. Warsaw Univ. Life Sci. Poland, 28-30 June 2017, 97-98, 2017.

Senay, G. B., Bohms, S., Singh, R. K., Gowda, P. H., Velpuri, N. M., Alemu, H. and Verdin, J. P.: Operational Evapotranspiration Mapping Using Remote Sensing and Weather Datasets: A New Parameterization for the SSEB Approach, J. Am. Water Resour. Assoc., 49(3), 577–591, doi:10.1111/jawr.12057, 2013.

SRTM (2015), Shuttle Radar Topography Mission Digital Elevation Model Courtesy of the US Geological Survey. <https://lta.cr.usgs.gov/SRTM1Arc> (accessed 27.09.16)

Sobowale, A. and Oyedepo, J. A.: Status of flood vulnerability area in an ungauged basin, South-west Nigeria, Int. J. Agric. Biol. Eng., 6(2), doi:10.3965/j.ijabe.20130602.00?, 2013.

Strauch, M., Schürz, C. and Schweppe, R.: topHRU - threshold optimization for HRUs in SWAT, 2017.

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.

Ufoegbune, G.C., Bello, N.J., Dada, O.F., Eruola, A.O.Makinde, A.A & Amori, A. .: Estimating Water Availability for Agriculture in Abeokuta, South Western Nigeria, , 12(9), 2012.

Ufoegbune, G. C., Yusuf, H. O., Eruola, A. O. and Awomeso, J. A.: Estimation of Water Balance of Oyan Lake in the North West Region of Abeokuta , Nigeria, Br. J. Environ. Clim. Chang., 1(1), 13–27, doi:10.5281/ZENODO.8060, 2011.

Wang-Erlandsson, L., Bastiaanssen, W. G. M., Gao, H., Jägermeyr, J., Senay, G. B., Van Dijk, A. I. J. M., Guerschman, J. P., Keys, P. W., Gordon, L. J. and Savenije, H. H. G.: Global root zone storage capacity from satellite-based evaporation, Hydrol. Earth Syst. Sci., 20(4), 1459–1481, doi:10.5194/hess-20-1459-2016, 2016.

Wang, X., Melesse, A. M. and Yang, W.: Influences of potential evapotranspiration estimation methods on SWAT's hydrologic simulation in a northwestern Minnesota watershed, Trans. ASABE, 49(6), 1755–1771, doi:10.13031/2013.22297, 2006.

Winchell, M., Srinivasan, R., Di Luzio, M. and Arnold, J.: SWAT Help, Texas Agrilife Res. United States Dep. Agric. Agric. Reseach Serv., 2013.

Xie, H., Nkonya, E. and Wielgosz, B.: Evaluation of the swat model in hydrologic modeling of a large watershed in Nigeria, in Proceedings of the 3rd IASTED African Conference on Water Resource Management, AfricaWRM 2010, pp. 71–76. [online] Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84858637919&partnerID=tZOtx3y1, 2010.

Table 1. Description and sources of input data used to configure SWAT model for the Ogun River Basin

Data type	Description/Resolution	Sources
Topography	30m Resolution, Digital Elevation Model 1 arc-second global coverage	Shuttle Radar Topography Mission (SRTM, 2015) https://lta.cr.usgs.gov/SRTM1Arc
Soil	250m Resolution, soil property maps of Africa	Soil property maps of Africa (Hengl et al., 2015) http://www.isric.org/projects/soil-property-maps-africa-250-m-resolution
Landuse	300m Resolution, landuse classification Year 2010	European Space Agency global land cover map (ESA CCI LC, 2014) https://www.esa-landcover-cci.org/?q=node/158
Weather	Daily precipitation, max. and min. temperature (1984-2012)	Nigerian Metrological Agency
Reservoir outflow	Reservoir daily discharge (Oyan:2007-2012)	Ogun-Oshun River Basin Authority Nigeria
Reservoir Water level	Daily water level (Oyan:1984-2012)	Ogun-Oshun River Basin Authority Nigeria
Management practices	Major crop management practices	Ogun state Agricultural Development Authority, Nigeria Oyo state Agricultural Development Authority, Nigeria

Table 2. Sensitivity rank and calibrated parameters with their optimal value of the three SWAT model run through the six calibrations

		Rank (optimal value)					
SWAT Parameter	**Description**	**GS1**	**GS2**	**GS3**	**MS4**	**MS5**	**MS6**
r__CN2.mgt	SCS runoff curve number for soil moisture condition II	1 (-0.01)	1 (-0.13)	1 (0.08)	1 (-0.48)	1 (-0.48)	1 (-0.47)
v__ESCO.hru	Soil evaporation compensation coefficient	2 (0.02)	4 (0.20)	3 (0.20)	4 (0.23)	8 (0.33)	5 (0.50)
v__CANMX.hru	Maximum canopy storage	3 (6.96)	2 (0.61)	2 (3.86)	5 (82.11)	9 (33.9)	4 (15.6)
r__SOL_BD,sol	Moist bulk density	4 (-0.19)	3 (0.11)	4 (-0.20)	3 (-0.82)	3(-0.005)	2 (-0.07)
v__ALPHA_BF.gw	Baseflow alpha factor	5 (0.66)	5 (0.62)	7 (0.13)	6 (0.42)	6 (0.9)	8 (0.14)
r__SOL_K.sol	Saturated hydraulic conductivity	6 (0.23)	10 (-0.26)	8 (0.24)	10 (0.49)	10 (-0.19)	10 (0.26)
v__EVRSV.res	Lake evaporation coefficient	7 (0.59)	7 (0.55)	10 (0.62)	8 (0.22)	7 (0.23)	7 (0.74)
v__GSI.plant.dat	Maximum stomatal conductance	8 (4.7)	11 (1.66)	11 (3.4)	7 (2.34)	5 (1.9)	6 (0.34)
v__FFCB.bsn	Initial soil water storage expressed as a fraction of field capacity water content	9 (0.59)	6 (0.82)	5 (0.15)	11 (0.99)	11 (0.4)	11 (0.83)
v__EPCO.hru	Plant uptake compensation factor	10 (0.47)	9 (0.61)	9 (0.07)	9 (0.95)	4 (0.88)	9 (0.47)
r__SOL_AWC.sol	Soil available water storage capacity	11 (0.8)	8 (0.92)	6 (0.77)	2 (0.96)	2 (0.89)	3 (0.93)

“v_” means a replacement (initial or existing parameter value is to be replaced by a given value);

“r_” means a relative change (initial or existing parameter value is multiplied by 1+ given value within the range)

Table 3. Average annual water balance at the outlet of the watershed in Abeokuta Town based on SWAT simulated output

Year	PRECP (mm)	AET (mm)	SW (mm)	PERC (mm)	SURQ (mm)	GW_Q (mm)	WYLD (mm)	LAT Q (mm)	ΔSW (mm)	*Estimated PRECP	Balance Year	PBIAS (%)
1989	1357	941	57	188	294	147	456	5	5	1442	-85	-6
1990	1094	882	82	69	145	52	207	4	-25	1081	13	1
1991	1161	881	54	117	228	84	321	4	27	1263	-101	-9
1992	1066	806	57	113	177	86	274	4	-3	1104	-38	-4
1993	1185	862	63	55	305	38	351	4	-5	1225	-41	-3
1994	870	768	47	34	96	17	118	3	16	918	-48	-6
1995	1166	858	55	116	225	83	317	4	-8	1200	-34	-3
1996	1457	885	45	201	460	148	621	5	10	1569	-112	-8
1997	1341	851	110	151	342	122	478	5	-65	1292	50	4
1998	1107	767	81	124	290	93	394	4	29	1222	-114	-10
1999	1515	900	100	223	458	183	656	5	-19	1577	-62	-4
2000	1198	814	55	175	306	143	463	4	45	1355	-157	-13
2001	841	738	35	27	108	12	128	3	20	900	-60	-7
2002	1241	758	64	146	375	108	492	4	-29	1260	-19	-2
2003	1456	845	56	216	488	177	681	5	8	1572	-117	-8
2004	1156	922	44	90	186	69	265	4	12	1220	-64	-6
2005	915	792	41	27	114	14	134	3	3	942	-27	-3
2006	1153	804	46	128	263	94	365	4	-5	1198	-45	-4
2007	1600	910	50	229	552	175	742	6	-4	1702	-103	-6
2008	1395	832	55	221	416	174	605	5	-4	1480	-85	-6
2009	1338	872	65	185	334	151	500	5	-10	1397	-59	-4
2010	1609	928	91	232	519	189	722	6	-26	1667	-58	-4
2011	1264	815	64	172	367	134	515	5	27	1395	-130	-10
2012	1409	839	60	265	386	205	609	6	4	1512	-103	-7

PRECIP: precipitation; AET: actual evapotranspiration; SW: soil water; PERC: percolation; SURQ: surface runoff; GW_Q: groundwater recharge; WYLD: water yield; LAT_Q: lateral flow; SW: change in soil moisture; * Estimated PRECP is WYLD+AET+ΔS+PERC-GW_Q, expressed in mm.

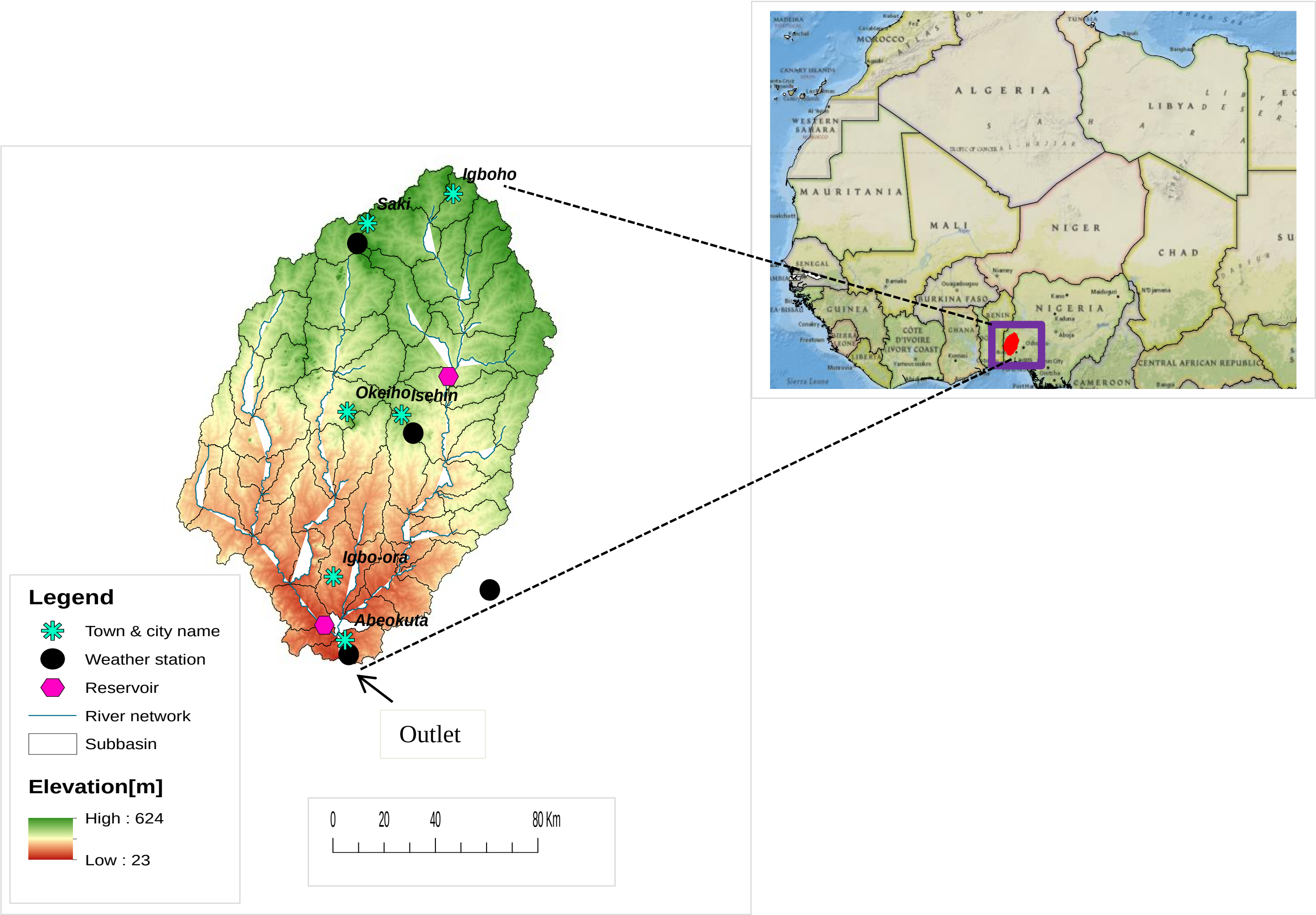

Figure 1. The Ogun River Basin located in Nigeria showing the SWAT-delineated subbasins, weather stations and river network

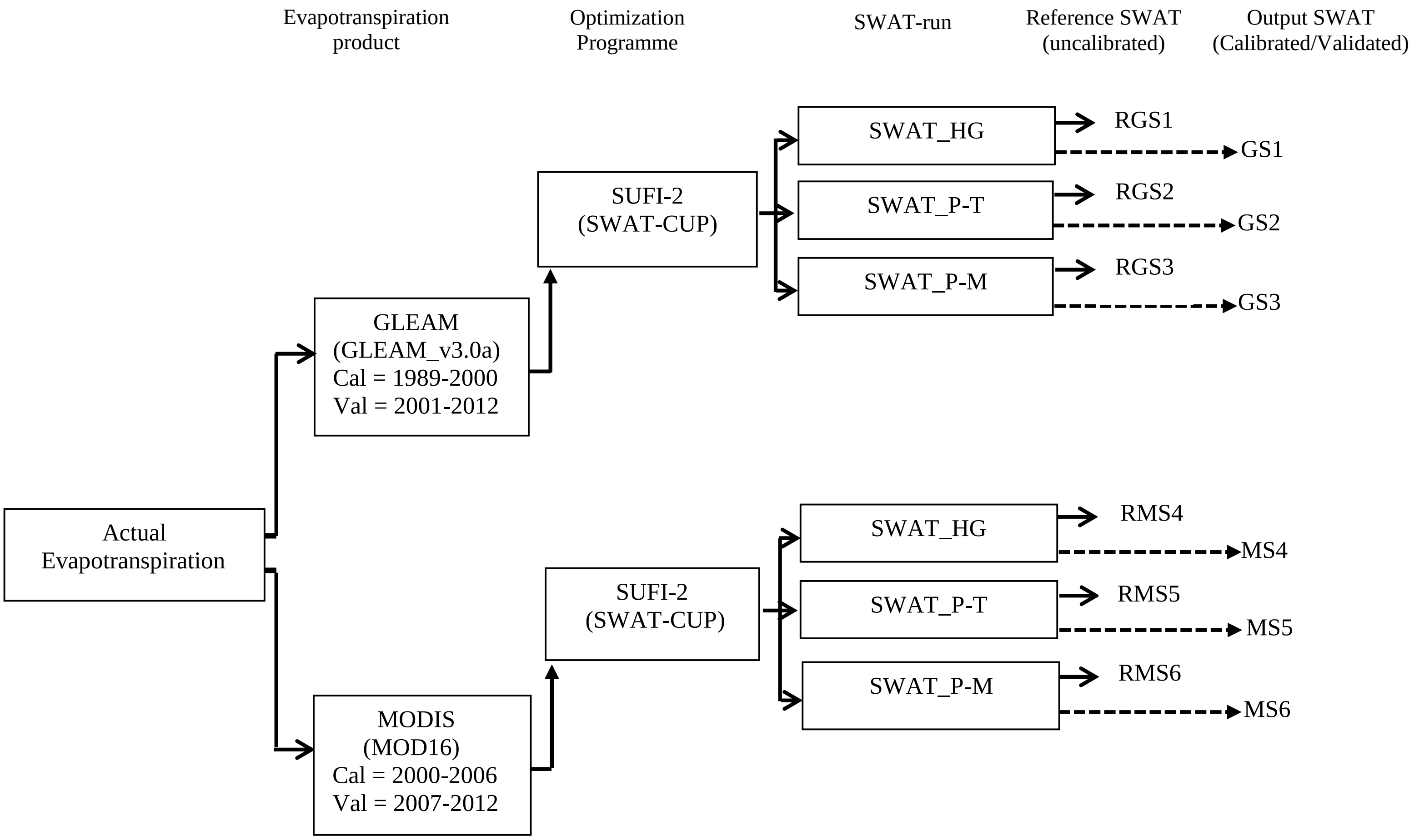

Figure 2. Schematic diagram showing the set-up of the SWAT model, the two global AET products, the resulting six SWAT reference runs, and calibration and validation procedures for the Ogun River Basin.

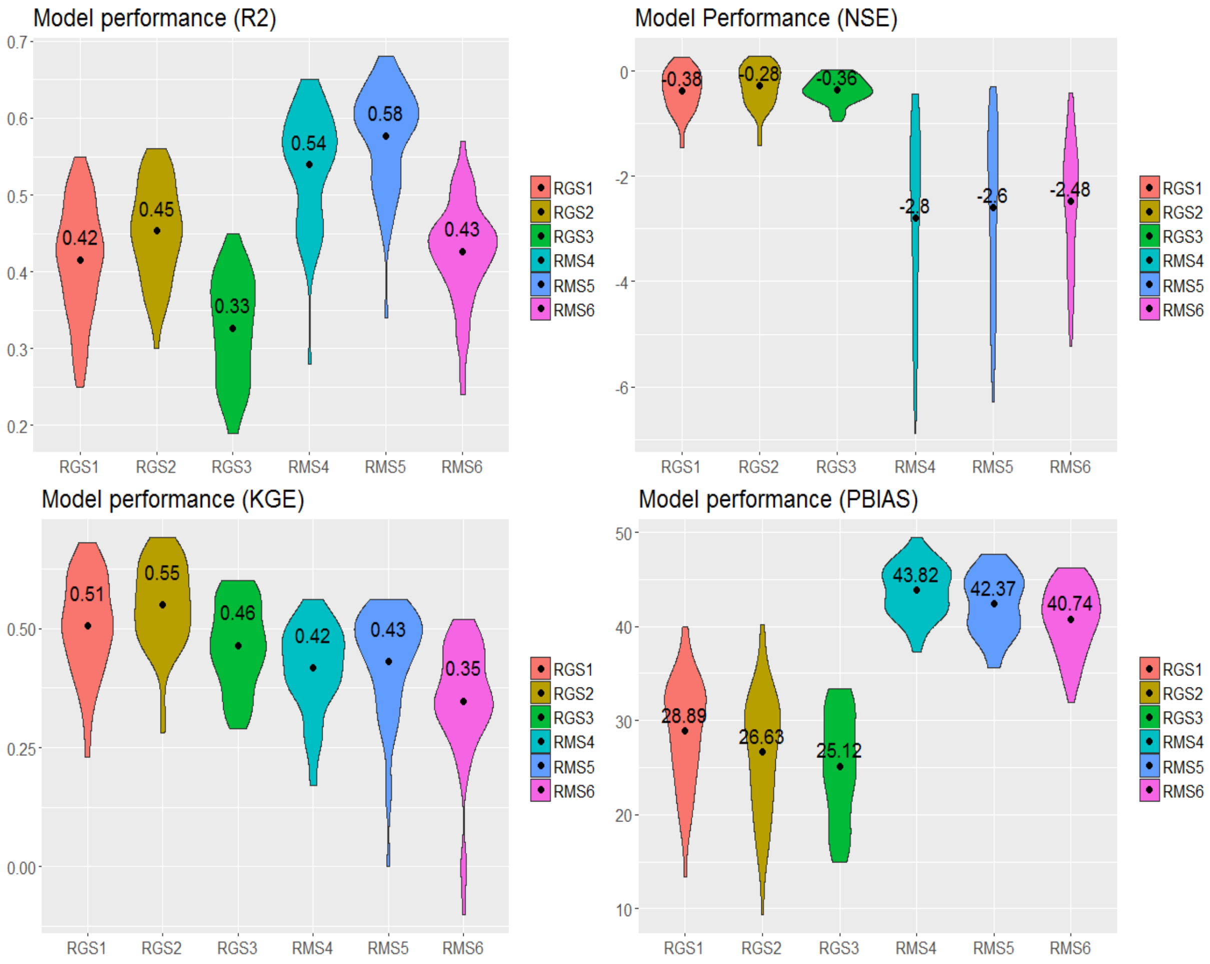

Figure 3. The plots of the performance results of the uncalibrated SWAT in simulating actual evapotranspiration. The values and the black dot symbol ("•") depicts the average value of, R^2, NSE, KGE and PBIAS obtained for each of the reference runs.

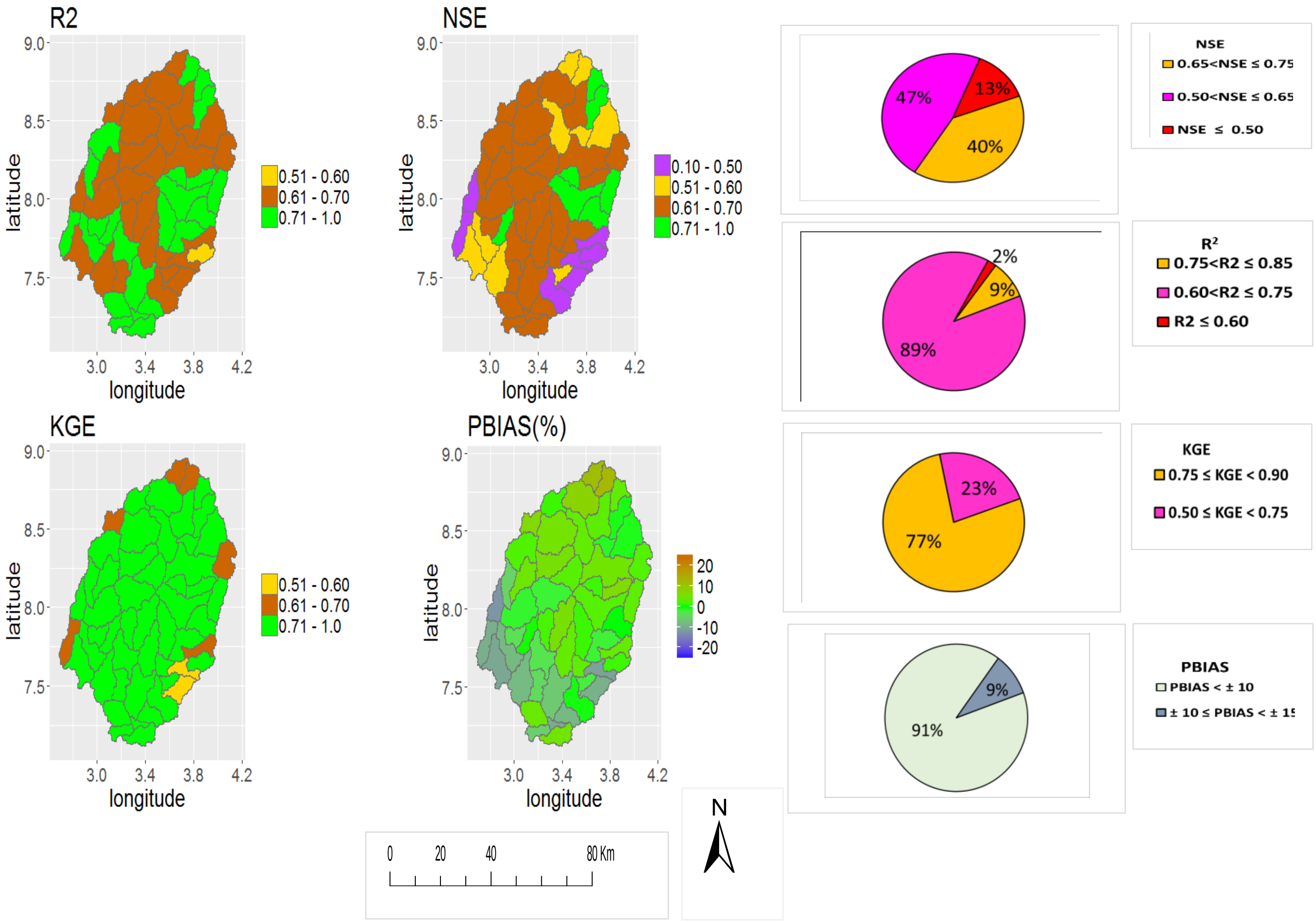

Figure 4. Performance metrics (NSE, KGE, R^2, and PBIAS) of SWAT (SWAT_HG) when calibrated with GLEAM_v3.0a (GS1).

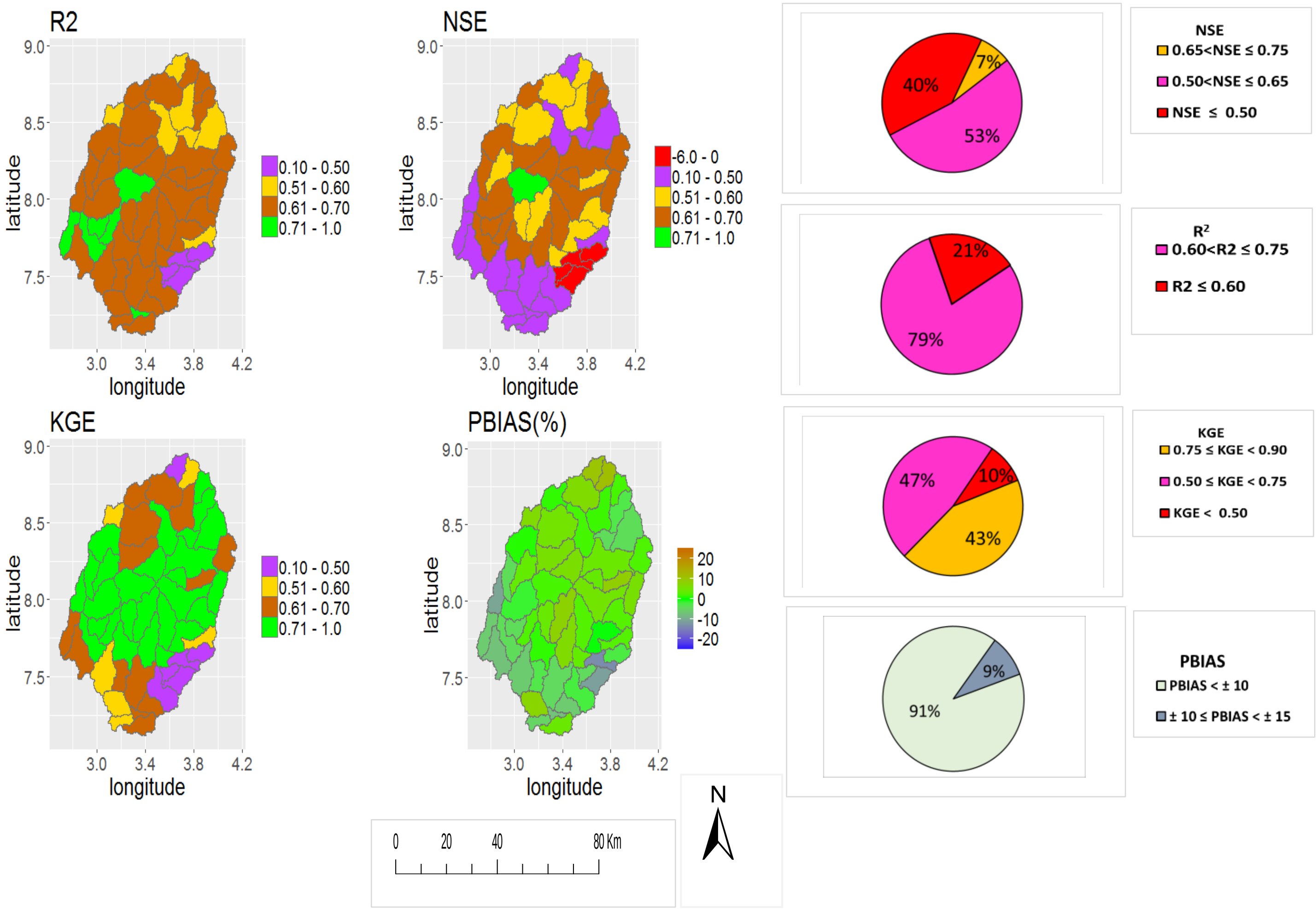

Figure 5. Performance metrics (NSE, KGE, R^2, and PBIAS) of SWAT (SWAT_HG) when validated with GLEAM_v.3.0a (GS1)

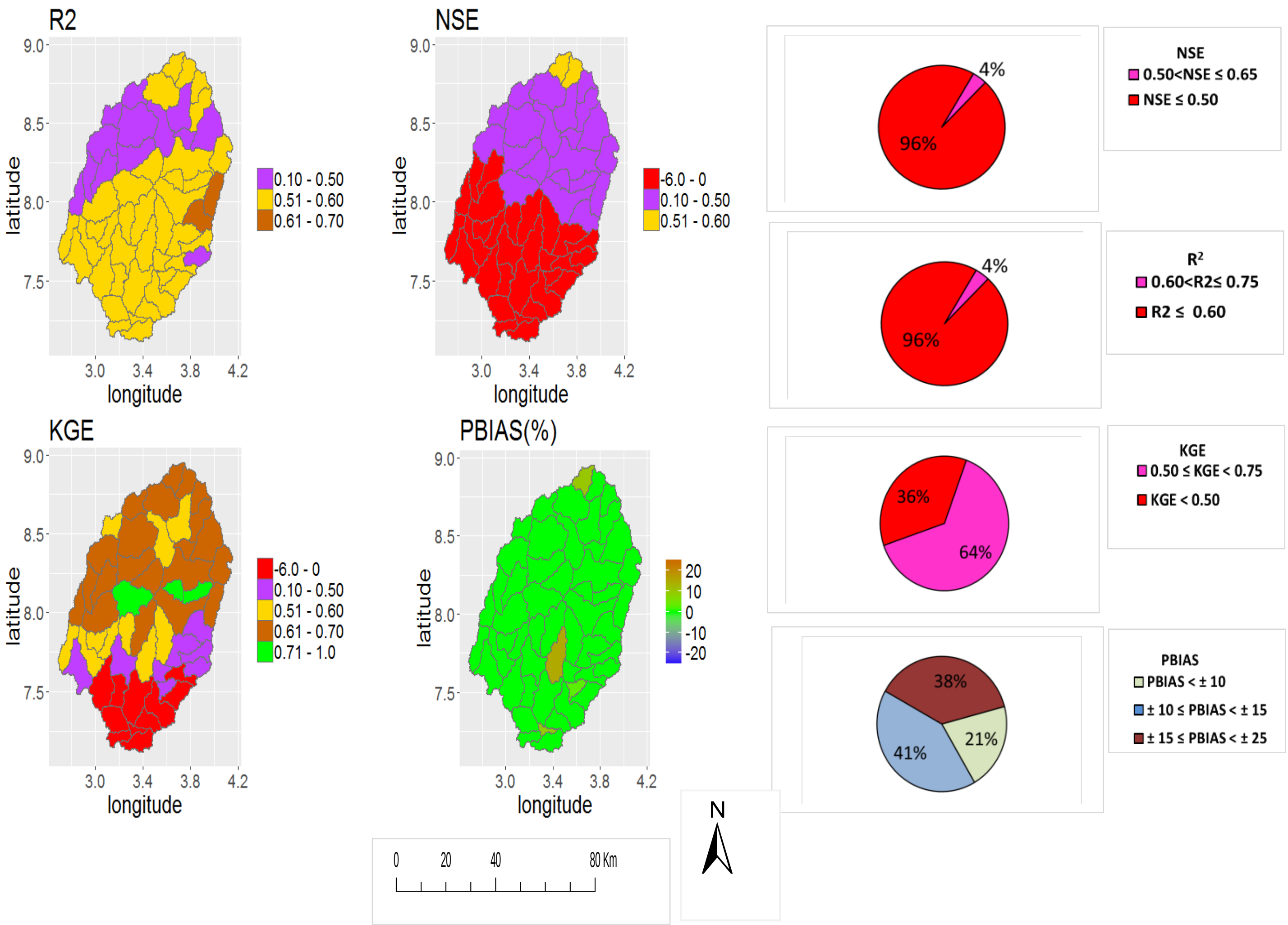

Figure 6. Performance metrics (NSE, KGE, R^2, and PBIAS) of SWAT (SWAT_P-M) when calibrated with MOD16 (MS6)

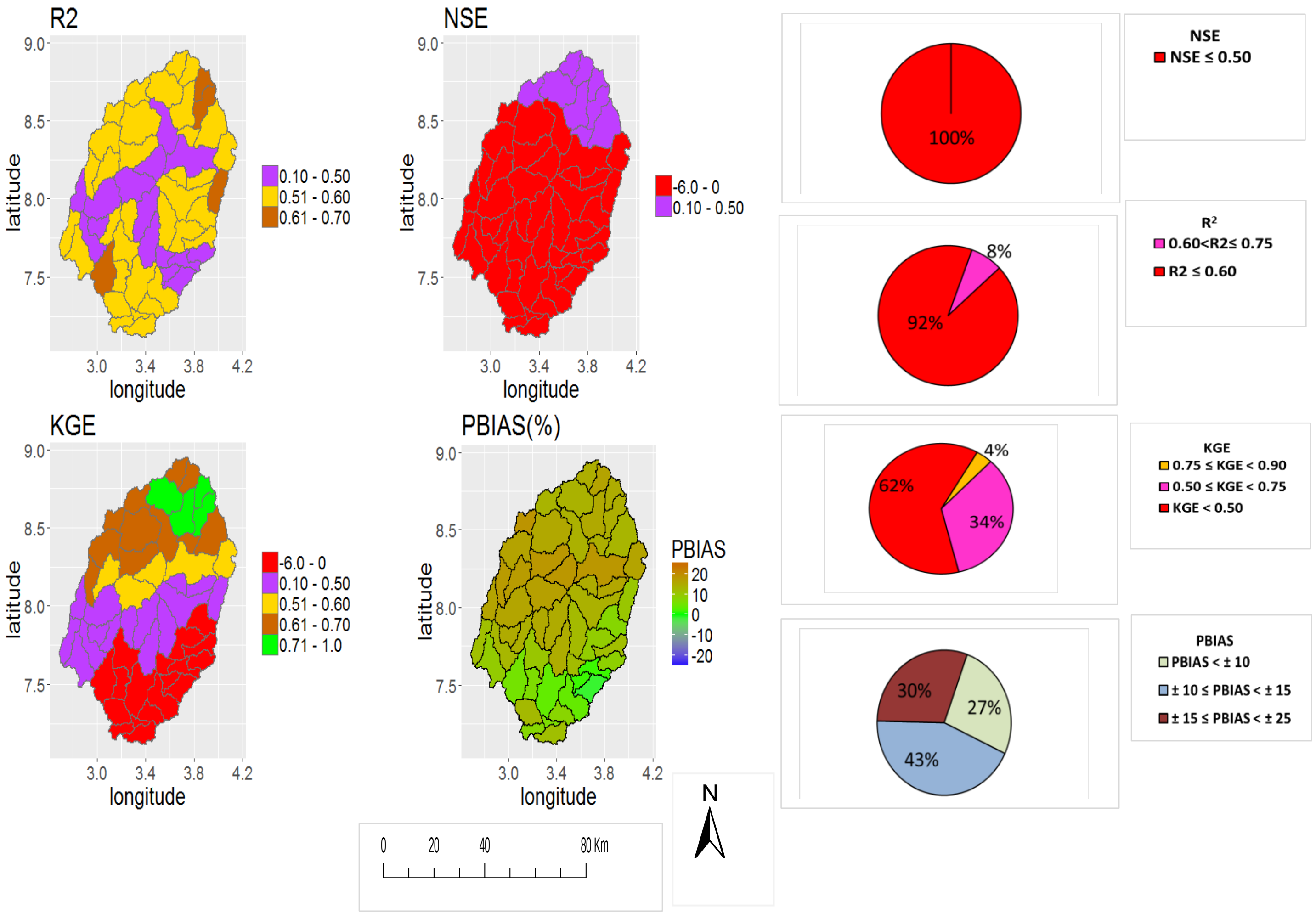

Figure 7. Performance metrics (NSE, KGE, R^2, and PBIAS) result of SWAT (SWAT_P-M) when validated MOD16 (MS6)

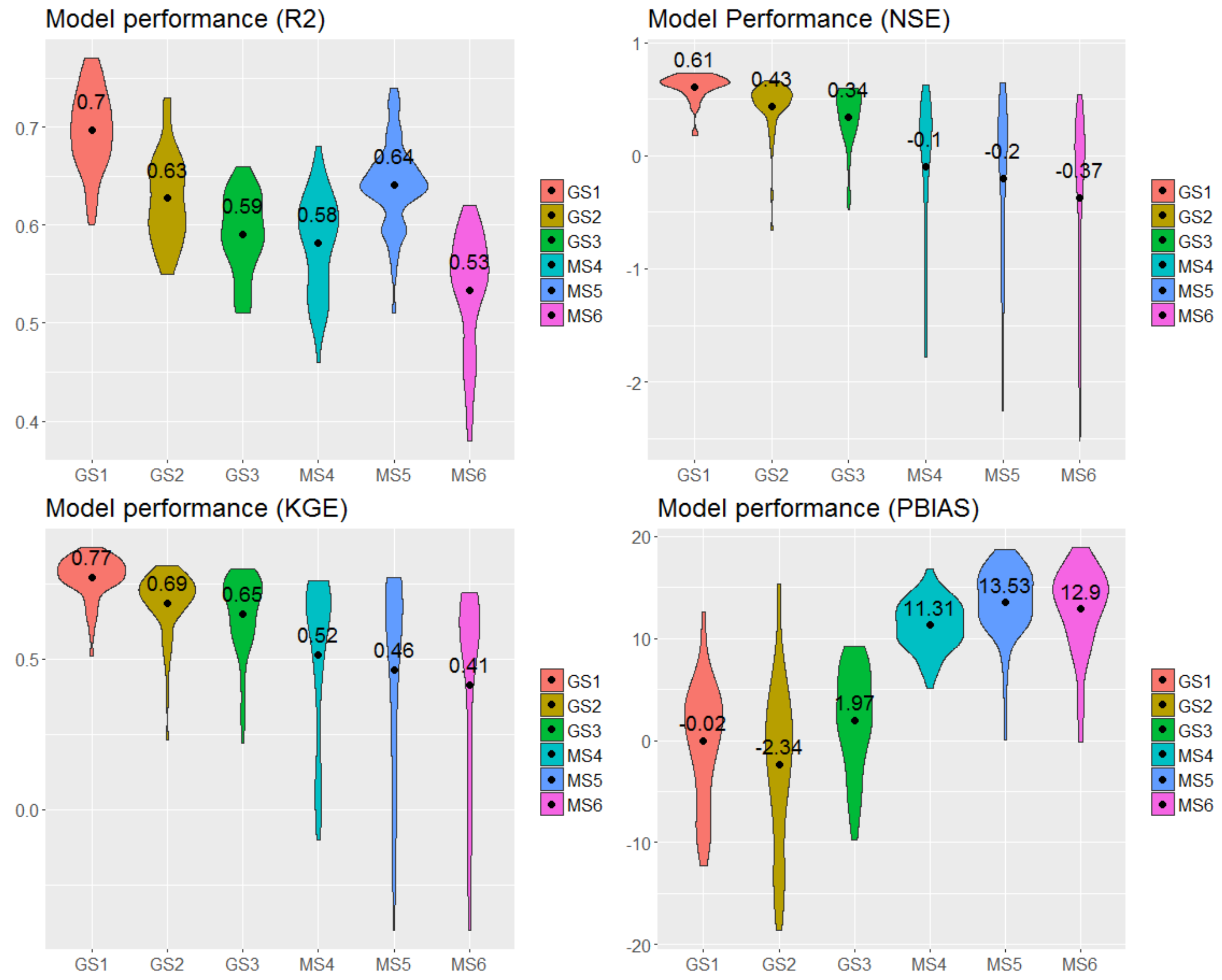

Figure 8. The plots of the performance result of SWAT in simulating actual evapotranspiration. The values and the black dot symbol ("•") depicts the average value of, R^2, NSE, KGE and PBIAS obtained for each calibration.

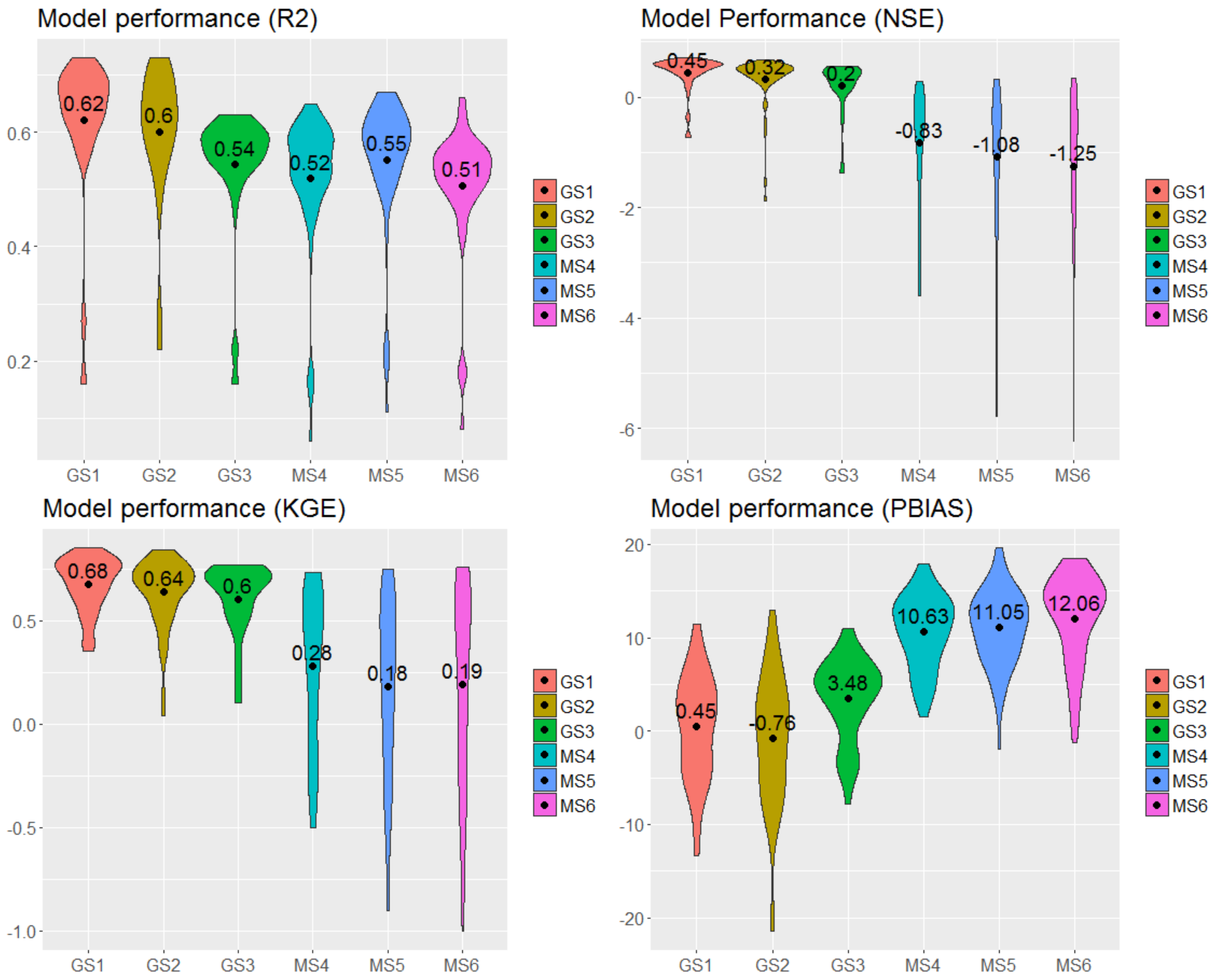

Figure 9. The plots of the performance result of SWAT in simulating actual evapotranspiration. The values and the black dot symbol (“•”) depicts the average value of, R^2, NSE, KGE and PBIAS obtained for each validation.

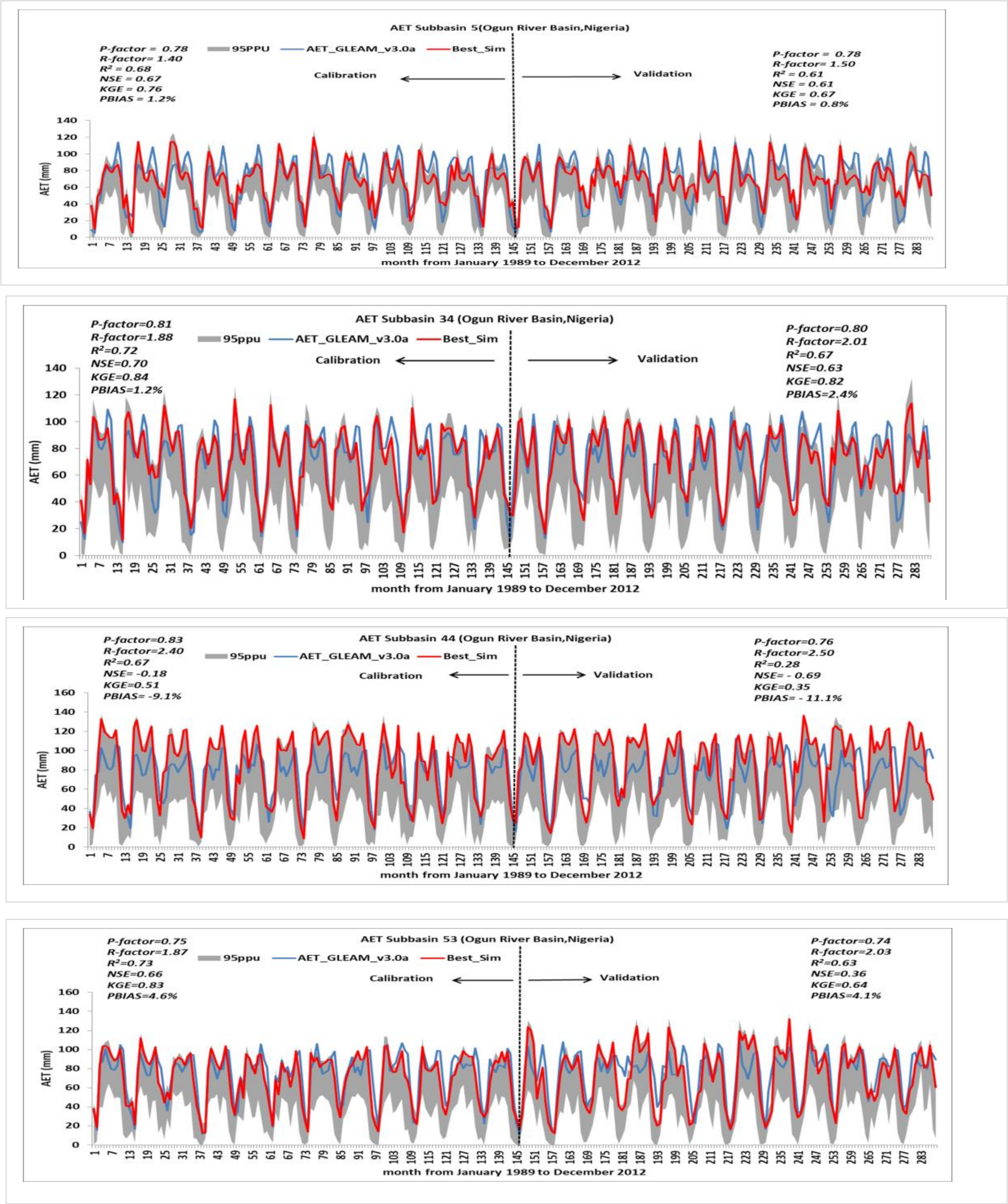

Figure 10. Extracts of the monthly calibration and validation results (GSI) showing the 95% prediction uncertainty interval along with the best SWAT simulated actual evapotranspiration and the satellite based actual evapotranspiration (GLEAM-v3.0a).

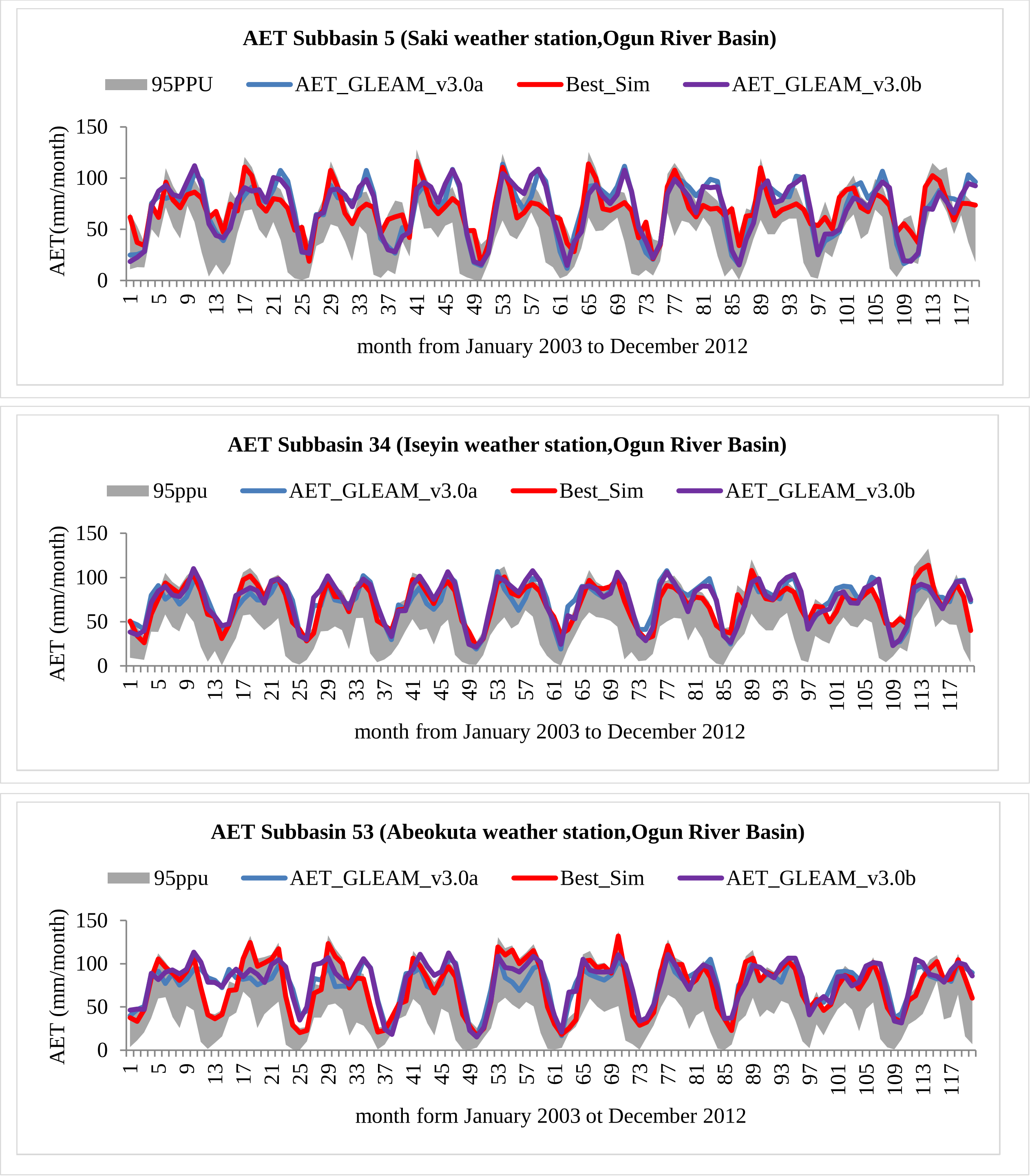

Figure 11. SWAT model verification results showing the satellite based AET GLEAM_v3.0a used for the model calibration/validation, the best SWAT simulated actual evapotranspiration (GS1), and an independent GLEAM_v3.0b time series bracketed by 95% predictive uncertainty.

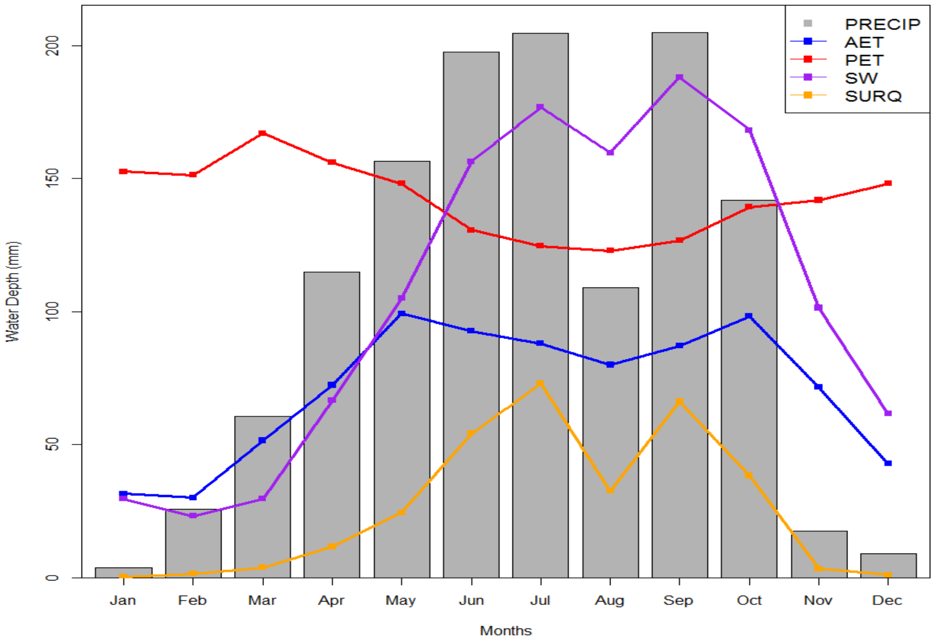

Figure 12. Seasonal fluctuation of water balance components at the outlet of the watershed located in Abeokuta Town

Appendix A: Performance metrics and their equations

Table A1. Table showing the performance metrics and their equations used to evaluate the model performance in this study.

Criterion	Mathematical equation	Description
R^2	$R^2 = \left(\frac{\sum_{i=1}^{n}(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2 \sum_{i=1}^{n}(P_i - \bar{P})^2}}\right)^2$	The percent of variance explained by the model. It is a statistical measure of how close the data are to the fitted regression line.
NSE	$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}$	Quantifies the relative magnitudes of the residual variance (noise) compared to the observed data variance
KGE	$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2}$	The goodness of fit measure provides an analysis of the relative importance of different components (correlation, bias and variability) in hydrologic simulation.
PBIAS	$PBIAS = 100\, x \frac{\sum_{i=1}^{n}(O_i - P_i)}{\sum_{i=1}^{n} O_i}$	The deviation of data being evaluated expressed in percentage. It measures the average tendency of the simulated data to be larger or smaller than the observation. Negative values indicate model overestimating (overprediction) and positive values indicate model underestimating (underprediction).

O_i are satellite based AET values; P_i are simulated AET values; $\bar{O}$ are mean satellite based AET values; $\bar{P}$ are mean simulated AET values; r is the Pearson product correlation coefficient between satellite-based AET and the simulated AET, α is the standard deviation of the simulated AET over the standard deviation of the satellite-based AET, and β is the ratio of the mean simulated AET to the satellite based AET.

Appendix B: SWAT model calibrated parameters

Table B1. Eleven parameters and their minimum and maximum range used in this study for the 1st iteration (1000 simulations) of all the six calibrations.

Parameter name	Minimum range	Maximum range
v__ESCO.hru	0.00	1.00
v__EPCO.hru	0.00	1.00
v__CANMX.hru	0.00	100.00
v__GSI{2,4,5}.plant.dat	0.00	5.00
v__ALPHA_BF.gw	0.00	1.00
v__EVRSV.res________17,50	0.00	1.00
v__FFCB.bsn	0.00	1.00
r__CN2.mgt	-0.25	0.85
r__SOL_AWC().sol	0.23	0.95
r__SOL_K().sol	-0.06	0.95
r__SOL_BD().sol	-0.41	0.95

Appendix C: Equations (taken from SWAT theoretical documentation (Neitsch et al., 2011)) showing the selected 11 most sensitive SWAT parameters used in this study

Maximum stomata conductance

The canopy resistance term is modified to reflect the impact of high vapor pressure deficit on leaf conductance when calculating actual evapotranspiration (Stockle et al,1992). The adjusted leaf conductance in which parameter GSI appears is calculated in SWAT using equation C1 and C2:

$$g_e = g_{e,mx} \times \left[1 - \Delta g_{e,dcl}(vpd - vpd_{thr})\right] \quad if\ vpd > vpd_{thr} \qquad \text{(C1)}$$

$$g_e = g_{e,mx} \qquad if\ vpd \leq vpd_{thr} \qquad \text{(C2)}$$

Where g_e is the conductance of a single leaf (m s^{-1}); $g_{e,mx}$ is the parameter GS1 which is the maximum stomatal conductance of a single leaf ((m s^{-1}); $\Delta g_{e,dcl}$ is the rate of decline in leaf conductance per unit increase in vapor pressure deficit (m s^{-1} kPa^{-1}), vpd is the vapor pressure deficit (kPa), and vpd$_{thr}$ is the threshold vapor pressure deficit above which a plant will exhibit reduced leaf conductance (kPa). The rate of decline in leaf conductance per unit increase in vapor pressure is calculated by solving equation C1.

The SCS curve number for soil moisture condition II

Three antecedent moisture conditions are defined by SCS curve number: I -dry (wilting point), II -average moisture, and III -wet (field capacity). The SCS curve numbers II is calculated from either SCS moisture condition I or from SCS moisture III in equation C3 and C4:

$$CN_1 = CN_2 - \frac{20\times(100-CN_2)}{(100-CN_2+exp[2.533-0.0636\times(100-CN_2)])} \qquad \text{(C3)}$$

$$CN_3 = CN_2 \times exp[0.00673 \times (100 - CN_2)] \qquad \text{(C4)}$$

Where CN1 is the moisture condition I curve number; CN2 is the moisture condition II curve number, and CN3 is the is the moisture condition III curve number.

Maximum canopy storage

The maximum amount of water that can be held in canopy storage varies from day to day as a function of the leaf area index in SWAT model and is estimated with equation C5 in which CANMX parameter appears:

$$can_{day} = can_{mx} \times \frac{LAI}{LAI_{mx}} \qquad \text{(C5)}$$

Where can$_{day}$ is the maximum amount of water than can be trapped in the canopy on a given day (mm H_20), can$_{mx}$ is the CANMAX parameter and is the maximum amount of water than can be trapped in the canopy when the canopy is fully developed (mm H_20), LAI is the leaf area index for a given day, and LAI$_{mx}$ is the maximum leaf area index for the plant.

Bulk density

Bulk density is calculated using equation C6:

$$\rho_b = \frac{M_S}{V_T} \qquad \text{(C6)}$$

Where ρ_b is the bulk density (Mg m^{-3}), Ms is the mass of solids (Mg) and VT is the total volume (m3). The total volume is calculated as:

$$V_T = V_A + V_W + V_S \tag{C7}$$

Where V_A is the volume of air (m^3), V_W is the volume of water (m^3), and V_S is the volume of solids (m^3).

Soil available water storage capacity

Soil available water storage capacity is calculated by subtracting the fraction of water present at permanent wilting point from that present at field capacity.

$$AWC = FC - WP \tag{C8}$$

Where AWC is the plant available water content, FC is the water content at field capacity, and WP is the water content at permanent wilting point.

Saturated hydraulic conductivity

The equation in which the parameter saturated hydraulic conductivity (SOL_K) appears is given in C9:

$$TT_{perc} = \frac{SAT_{ly} - FC_{ly}}{K_{sat}} \tag{C9}$$

Where TT_{perc} is the travel time for percolation (hrs), SAT_{ly} is the amount of water in the soil layer when completely saturated (mm H_20), FC_{ly} is the water content of the soil layer at field capacity (mm H_20), and K_{sat} is the saturated hydraulic conductivity for the layer (mm h^{-1}).

Baseflow alpha factor

The baseflow recession constant (Baseflow alpha factor) is α_{gw} . The α_{gw} is calculated using equation C10:

$$\alpha_{gw} = \frac{1}{N} \times ln\left[\frac{Q_{gw,N}}{Q_{gw,0}}\right] \tag{C10}$$

Where α_{gw} is the ALPHA_BF parameter, N is the time lapsed since the start of the recession (days), $Q_{gw,N}$ is the groundwater flow on day N (mm H_20), Qgw,0 is the groundwater flow at the of the start of the recession (mm H_20).

Lake evaporation coefficient

The equation in which the Reservoir evaporation coefficient (EVRSV.res) appears is shown in C11:

$$V_{evap} = 10 \times \eta \times E_0 \times SA \tag{C11}$$

Where V_{evap} is the volume of water removed from the water body by evaporation during the day (m^3 H_20), η is an evaporation coefficient with a default value of 0.6 (EVRSV), E_0 is the potential evapotranspiration for a given day (mm H_20), and SA is the surface a rea of the water body (ha).

Plant uptake compensation factor

The equation in which plant uptake compensation factor (EPCO) appears (C13) is used to calculate the adjusted potential water uptake when the upper layers in the soil profile do not contain enough water to meet the potential water uptake (C12):

$$W_{up,ly} = W_{up,zl} - E_{up,zu} \tag{C12}$$

Where $W_{up,ly}$ is the potential water uptake for layer ly (mm H_20), $W_{up,zl}$ is the potential water uptake for the profile to the lower boundary of the soil layer (mm H_20), $E_{up,zu}$ is the potential water uptake for the profile to the upper boundary of the soil layer (mm H_20).

$$W'_{up,ly} = W_{up,ly} + W_{demand} \times epco \tag{C13}$$

Where $W'_{up,ly}$ is the adjusted potential water uptake for layer ly (mm H_20), W_{demand} is the water uptake demand not met by overlying soil layers (mm H_20), and epco is the plant uptake compensation factor.

Soil evaporation compensation coefficient

The modified equation of the amount of evaporative demand for a soil layer which is determined by taking the difference between the evaporative demands calculated at the upper and lower boundaries of the soil layer incorporate a coefficient called ESCO for depth distribution modification. The modified equation is:

$$E_{soil,ly} = E_{soil,zl} - E_{soil,zu} \times esco \qquad \text{(C14)}$$

Where $E_{soil,ly}$ is the evaporative demand or layer ly (mm H20), $E_{soil,zl}$ is the evaporative demand at the lower boundary of the soil layer (mm H20), $E_{soil,zu}$ is the evaporative demand at the upper boundary of the soil (mm H20) and esco is the soil evaporative compensation coefficient.

Initial soil water storage expressed a fraction of field capacity water content

The estimation of field capacity water content is the equation in which the initial soil water storage expressed as a fraction of field capacity water content (FFCB) appears. The equation is C15:

$$FC_{ly} = WP_{ly} + AWC_{ly} \qquad \text{(C15)}$$

Where FC_{ly} is the water content at field capacity expressed as a fraction of the total soil volume (FFCB), WP_{ly} is the water content at wilting point expressed as a fraction of the total soil volume, and AWC_{ly} is the available water capacity of the soil layer expressed as a fraction of the total soil volume.

Appendix D: Ogun River Basin with its 53 subbasins intersected by the satellite-based AET pixels

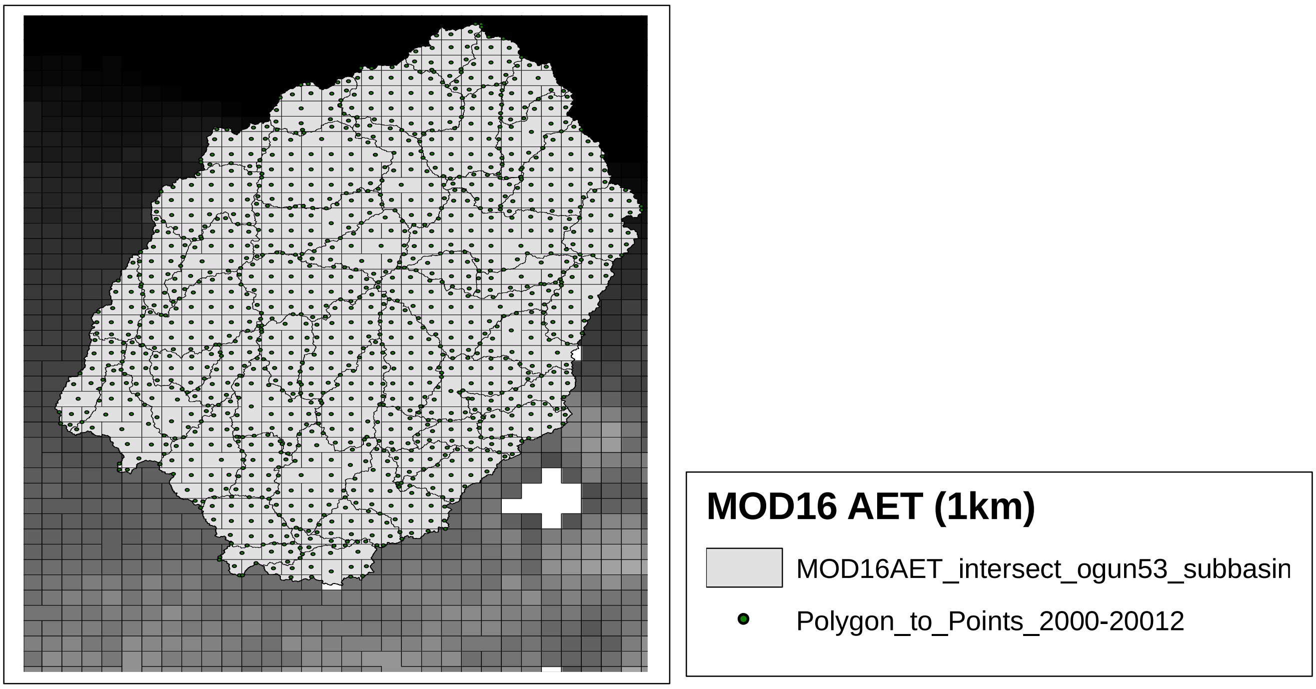

Figure D1. Ogun River Basin with its 53 subbasin polygons intersected by the pixel of MOD16 AET and showing the number of MOD16 AET pixels covering each subbasin polygon

Figure D2. Ogun River Basin with its 53 subbasin polygons intersected by the pixel of GLEAM AET and showing the number of GLEAM AET pixels covering each subbasin polygon.