Journal of Hydrology 603 (2021) 126781

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

A comprehensive calibration and validation of SWAT-T using local datasets, evapotranspiration and streamflow in a tropical montane cloud forest area with permeable substrate in central Veracruz, Mexico

Sergio Miguel López-Ramírez[a,*], Alex Mayer[b], Leonardo Sáenz[a], Lyssette Elena Muñoz-Villers[c], Friso Holwerda[c], Nathaniel Looker[d], Christoph Schürz[e], Z. Carter Berry[f], Robert Manson[g], Heidi Asbjornsen[h], Randall Kolka[i], Daniel Geissert[g], Carlos Lezama[g]

[a] Michigan Technological University, Civil and Environmental Engineering, 870 Dow Environmental Sciences, 1400 Townsend Drive, Houghton, MI 49931, USA
[b] University of Texas at El Paso, Civil Engineering, 500 West University Avenue, El Paso, TX 79968, USA
[c] Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Ciudad de México, México, Coyoacán 04510, Mexico
[d] University of Minnesota, Department of Soil, Water and Climate, St. Paul, MN 55108, USA
[e] Institute of Water Management, Hydrology and Hydraulic Engineering, University of Natural Resources and Life Science Vienna, Gregor-Mendel-Straße 33, 1180 Wien, Austria
[f] Chapman University, Schmid College of Science and Technology, Orange, CA 92866-1005, USA
[g] Instituto de Ecología, A.C., Carretera antigua a Coatepec 351, Xalapa 91073, Veracruz, Mexico
[h] University of New Hampshire, Department of Natural Resources and the Environment, Durham, NH 03824, USA
[i] USDA Forest Service, Northern Research Station, Grand Rapids, MN 55744, USA

ARTICLE INFO

This manuscript was handled by Marco Borga, Editor-in-Chief, with the assistance of Andres Iroume, Associate Editor

Keywords:
Streamflow
Catchment hydrological modeling
Potential evapotranspiration
Rainfall interception
Tropical montane forests

ABSTRACT

Tropical montane cloud forests (TMCF) are threatened ecosystems despite their capacity to maintain high dry-season baseflow. A number of conservation policies, including payments for hydrological services, have been implemented to protect these forests. However, since most of the modeling tools used to assess the impacts of these policies were developed for temperate zones, more work is needed to understand and improve the applicability of popular models in tropical contexts. This study uses local evapotranspiration and streamflow datasets to calibrate and validate an improved version of the Soil and Water Assessment Tool model for the Tropics (SWAT-T). Vegetation growth and canopy water storage capacity were calibrated using field data. Three methods provided by SWAT-T to calculate potential evapotranspiration (PET) were compared: Penman-Monteith (SWAT-T-PM), Hargreaves (SWAT-T-HA), and Priestly-Taylor (SWAT-T-PT). Sensitivity analysis and calibration of daily streamflow were conducted at the catchment scale (34 km^2). Furthermore, the calibrated models were validated at three sites with evapotranspiration data, and at four distinct micro-catchments (0.137–0.446 km^2) with gauged streamflow data. Overall, SWAT-T satisfactorily simulated streamflow during the calibration period producing acceptable goodness of fit indices. However, the model incorrectly predicted the dominance of lateral flow instead of the deep groundwater flow observed from isotope-based studies. SWAT-T-HA performed better than SWAT-T-PM and SWAT-T-PT, but all models underestimated the influence of rainfall interception losses since evaporation is limited by daily PET in forests. Finally, SWAT-T largely over- and underestimated mean annual daily low flow in pastures and forests, respectively. Taken together, these results indicate that improvements in the parametrization of rainfall interception and deep subsurface flow dynamics in SWAT-T are required to improve applicability of this modeling tool in tropical montane areas underlain by permeable substrates.

* Corresponding author.
E-mail addresses: slopezra@mtu.edu (S.M. López-Ramírez), amayer2@utep.edu (A. Mayer), christoph.schuerz@boku.ac.at (C. Schürz), robert.manson@inecol.mx (R. Manson), Heidi.Asbjornsen@unh.edu (H. Asbjornsen), randall.k.kolka@usda.gov (R. Kolka).

https://doi.org/10.1016/j.jhydrol.2021.126781
Received 8 March 2021; Received in revised form 25 July 2021; Accepted 30 July 2021
Available online 17 August 2021

1. Introduction

Integrating scientific knowledge into hydrological models is key for scaling up and improving catchment policy design (Naeem et al., 2015; Wright et al., 2018). Hydrological models have been used to assess the efficiency of management schemes for conserving hydrological services (Quintero et al., 2009; Bremer et al., 2020). Nonetheless, in most cases, models are calibrated using only statistical criteria, comparing model output to stream gauge data at catchment outlets, largely due to the lack of adequate data for other hydrological processes (Arnold et al., 2015; Hrachowitz et al., 2014). This issue remains a major challenge in hydrology, especially in tropical montane areas characterized by high loss of forested area, limited data to constrain ecohydrological processes, less suitable modeling tools, and high spatial variability in climate, vegetation, and soils (Hansen et al., 2013; Hamel et al., 2017; Wright et al., 2018). Better understanding and modeling of hydrological responses with a focus on the incorporation of local datasets for multiple sites and processes in tropical montane settings is key to the effective implementation, monitoring, and success of conservation policies (López-Ramírez et al., 2020; Guswa et al., 2014).

The Soil and Water Assessment Tool (SWAT) is a process-based, continuous daily hydrological model that allows the simultaneous simulation of hydrology and plant growth (Arnold et al., 1998; Heidari et al., 2019). In SWAT, a basin is partitioned into sub-basins, which in turn are subdivided into hydrological response units (HRUs) that represent a unique combination of land use, soil type, and slope class (Neitsch et al., 2011). SWAT includes five storage types to calculate the water balance: snow, the canopy, the soil profile, and the shallow and deep aquifers (Neitsch et al., 2011). SWAT uses empirical Potential Evapotranspiration (PET) methods to estimate actual ET, including the temperature-based Hargreaves (HA) (Hargreaves and Samani, 1985), radiation-based Priestley-Taylor (PT) (Priestley and Taylor, 1972), and the combined energy-mass transfer Penman-Monteith (PM) (Monteith, 1965) methods. A comprehensive outline of ET calculations in SWAT is presented by Abiodun et al. (2018). The literature suggests that different methods perform better under certain climatic conditions, especially in data-scarce regions (i.e., Alemayehu et al., 2015; Samadi, 2017). Moreover, while the PET is a commonly used index in estimating ET, forest PET values at the landscape level are often indirectly estimated using models that were developed for short crops and under specific climatic conditions (Lu et al., 2005), and more work is needed to better understand their applicability in tropical montane regions (Archibald and Walter, 2013).

Applications across a wide range of topographic, soils, and land use conditions have demonstrated the broad utility of SWAT for simulating streamflow responses (Francesconi et al., 2016; Shrestha et al., 2018; Van Liew et al., 2007; Quintero et al., 2009; Francesconi et al., 2016; Tuppad et al., 2010; Plesca et al., 2012), including in Tropical montane cloud forests (TMCF) in Mexico, underlain by impermeable soils with high dominance of surface runoff (Salas-Martínez et al., 2014; Sánchez-Galindo et al., 2017). However, most model evaluations have relied only on statistics comparing simulated and measured flow at the outlet (Arnold et al., 2015). A recent introduction of signature metrics for different parts of the flow duration curve has shown that it is possible to achieve a balanced hydrograph with the SWAT model (Haas et al., 2015; Pfannerstill et al., 2014). Nonetheless, calibrating SWAT simultaneously for very high and very low flows is still challenging (Shrestha et al., 2018; Pfannerstill et al., 2014). Assessing and improving the applicability of the SWAT model across the spectrum of catchment characteristics (e.g., substrate permeability, streamflow generation mechanisms) present in tropical montane areas requires more thorough calibration procedures and multi-site validation with ancillary data and local understanding of the main water budget components.

Recently, Strauch and Volk (2013) and Alemayehu et al. (2017) suggested improvements to SWAT to better represent vegetation growth in tropical areas (SWAT-T) and hence improve water balance simulations. SWAT assumes that vegetation enters a dormant period at the end of each growing season (i.e., during the autumn and winter), at which point the Leaf Area Index (LAI) is set to a minimum value close to zero (Neitsch et al., 2011). LAI is assumed to decline linearly to its specified minimum when leaf senescence exceeds leaf growth (Neitsch et al., 2011). These assumptions do not realistically represent the seasonal dynamics (phenology) of foliage in the tropics in general or in the study area in particular (Williams-Linera, 1999; Alemayehu et al., 2017), since many tropical plant species exhibit drought-controlled dormancy or continuous growth throughout the year (Strauch and Volk, 2013). In SWAT-T the plant growth subroutine was modified to better represent the phenological patterns encountered at low latitudes (20° N to 20° S; Alemayehu et al., 2017). SWAT-T incorporates the logistic decline curve to model leaf senescence as suggested by Strauch and Volk (2013). Additionally, in this model, the beginning of the new growth cycle for trees and perennials is triggered by a Soil Moisture Index, which occurs within a period defined by the user as the months at the end of the dry season (SOS1) and the beginning of the rainy season (SOS2). Results from Alemayehu et al. (2017) in the Mara Basin (Kenya/Tanzania) showed that SWAT-T achieved higher correlations between observed and simulated LAI than the standard SWAT and improved the simulation of evapotranspiration and streamflow of the studied period by around 12%. As SWAT-T has not been evaluated in other contexts, we sought to apply this model in areas characterized by TMCF.

In this study, we calibrated and validated the SWAT-T model to simulate discharge and evapotranspiration in areas influenced by TMCF located in Central Veracruz, Mexico. We hypothesize that by contrasting calibrated SWAT-T models (using local vegetation and ecohydrological parameters, such as LAI and canopy storage capacity) against hydrologic observations (e.g. streamflow and evapotranspiration), we can identify model strengths and areas of improvement for analyzing the water budget response in these environments. Specifically, (a) we validate calibrated SWAT-T models with local evapotranspiration data in three sites, with a focus on the performance of the PET methods available in the model, (b) we assess the accuracy of the SWAT-T model to simulate streamflow at the outlet of the Gavilanes catchment (34 km^2), and (c) we validate the calibrated SWAT-T models with streamflow observed in four nested micro-catchments (0.137–0.446 km^2) within the study area. Two particularly novel aspects of our work are that (1) it validates the SWAT-T model with evapotranspiration and streamflow data at HRU level in areas with contrasting land covers and that (2) it identifies areas for improving the applicability of SWAT-T for science and policy applications in tropical montane areas underlain by permeable soils.

2. Material and methods

2.1. Study site

The research was carried out in the Gavilanes catchment (34 km^2) and four micro-catchments (0.137–0.446 km^2) located within the catchments of the Pixquiac and Gavilanes rivers (106 and 34 km^2, respectively), which are tributaries of the upper Antigua River basin (1565 km^2). The Gavilanes catchment is the main source of water for the city of Coatepec, Veracruz, Mexico (García et al. 2004), while the Pixquiac catchment provides 38% of the water supply for the Veracruz state capital of Xalapa (Paré and Gerez, 2012); numerous water extractions upstream of the gauging station in the Pixquiac river prohibited the inclusion of this catchment in our analysis. The Gavilanes catchment was originally dominated by TMCF and is located in central Veracruz, Mexico. Elevations in the catchment range from 1226 m a.s.l. to 2962 m a.s.l. (Fig. 1). Its general climate is temperate humid (Garcia, 2004) with about 80% of the annual rainfall occurring during the wet season (May-October). Mean annual rainfall ranges from 1385 mm to 3185 mm and 1900 mm as elevation increases from 1210 m a.s.l. to 2100 m a.s.l. and 3000 m a.s.l., respectively. Mean daily temperatures decreases from 19° to 5 °C from 1200 to 3000 m.a.s.l. (Muñoz-Villers et al., 2016). Annual

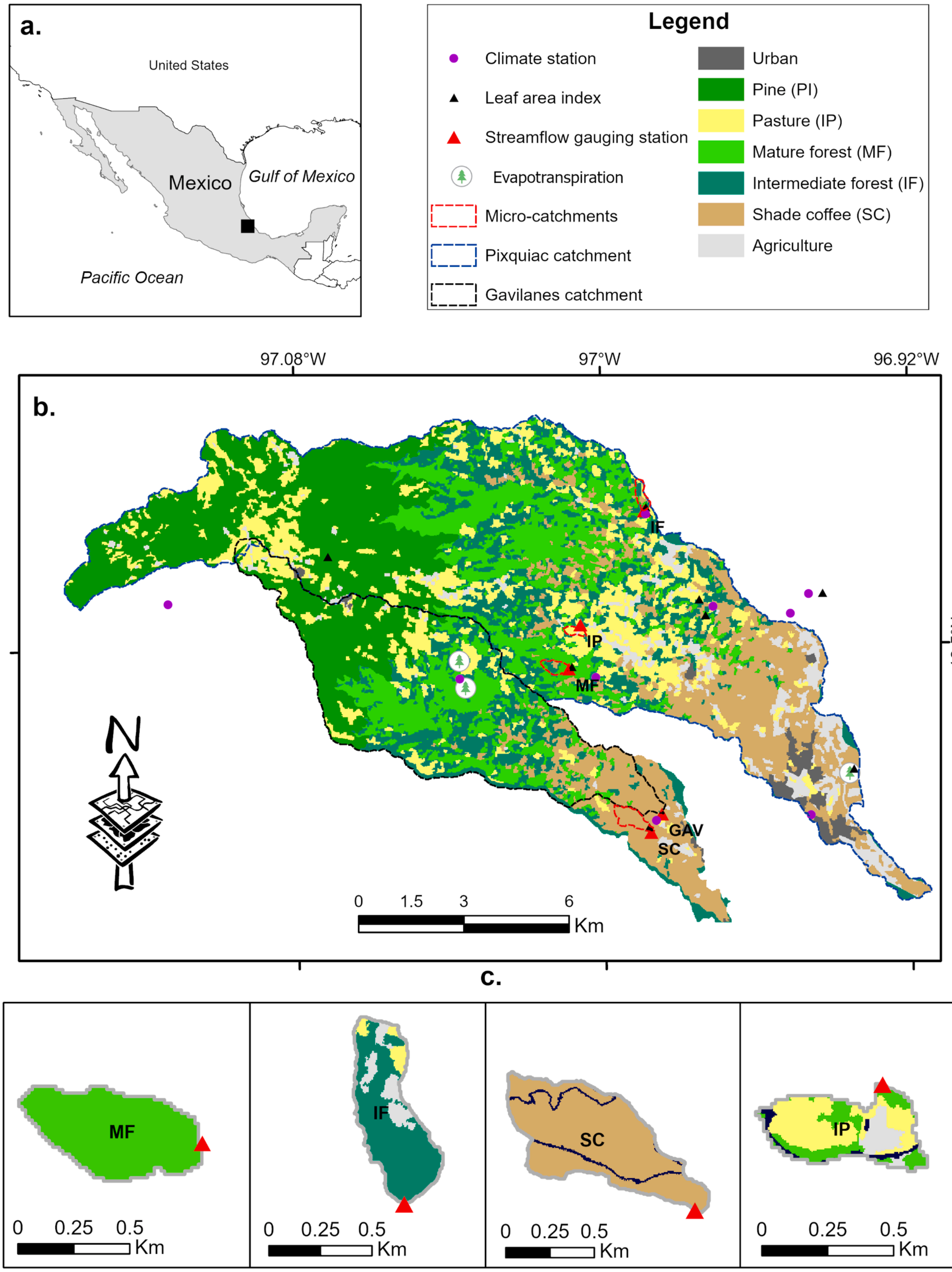

Fig. 1. General location (a), the study area, including the Gavilanes and Pixquiac catchments (b), and four micro-catchments with different land use: mature forest (MF), intermediate forest (IF), shade coffee (SC), and pasture (IP) (c). The map includes sites where the leaf area index and evapotranspiration were observed. Land use in (b) may look different than (c) because in the micro-catchments land use was refined with Rapid Eye imagery.

values of cloud water interception account for<2% of the total rainfall in the region (Holwerda et al., 2010).

We selected micro-catchments based on the dominance of four land use/land cover (LULC) types within each micro-catchment, (see the LULC map in Fig. 1b) including mature forest (MF), intermediate forest (IF), shade coffee (SC), and pasture (IP). Together, these four LULC types comprise around 70% of LULC in this study area (Von Thaden et al., 2019) with much of the remaining area covered by pine forest (PI). A detailed description of the characteristics of the micro-catchments is provided in Table 1, and more details can be found in López-Ramírez et al. (2020).

The MF micro-catchment is dominated by old-growth TMCF (>50 years old) with low disturbance. IF is mainly covered by 40-year-old TMCF (>77% of the area) and scattered pastureland and annual crops

Table 1
Description and daily hydrologic indices observed at four micro-catchments. Table includes land use and slope class corresponding to the HRUs used for model validation at the micro-catchment scale.

	MF	IF	SC	IP
Area (km^2)[a]	0.242	0.224	0.446	0.137
Mean elevation (m a.s.l.)[a]	1756	1604	1284	1655
Mean slope (%)[a]	36%	25%	21%	21%
Percent forest or coffee cover[a]	100%	77%	94%	29%
Percent of pasture and crops cover[a]	0%	23%	0%	63%
Percent of urban and roads[a]	0%	0%	6%	8%
Leaf area index (LAI) (m^2 m^{-2})[a]	5.93 ± 0.9	5.96 ± 1.22	4.3 ± 1.71	–
Soil saturated hydraulic conductivity (Kfs) at 20-cm depth (mm hr^{-1})[a]	125 ± 2.1	127 ± 2.1	48 ± 3.4	26 ± 2
N (day) [a]	469	642	360	467
Average daily rainfall (mm day^{-1})[a]	6.47	4.64	6.67	6
Average daily runoff (mm day^{-1})[a]	4.35	2.67	3.05	4.14
Ratio between runoff and rainfall, Q/P (–)[a]	0.67	0.58	0.46	0.69
Mean annual high flow, Q_5 (±SD) (mm day^{-1})[a]	10.7 ± 0.25	4.7 ± 0.05	6.5 ± 0.12	12.5 ± 0.35
Mean annual low flow, Q_{95} (±SD) (mm day^{-1})[a]	1.29 ± 0	1.18 ± 0	0.54 ± 0	0.37 ± 0
HRU Slope (%)[b]	35–99	0–35	0–35	0–35

Note: Where available, the standard deviation (SD) is provided.
Abbreviations :a.s.l., above sea level IF, intermediate forest IP, pasture MF, mature forest SC, shade coffee.
[a] López-Ramírez et al. (2020).
[b] Estimated from SWAT-T setup (See section 2.2 in text for explanation).

(mainly maize) located in the upper parts of the catchment. The mean tree height in these two micro-catchments is similar and ranges between 20 and 25 m (Vizcaíno-Bravo et al., 2020). They also have a similar mean leaf area index (LAI), see Table 1. The SC micro-catchment has been completely covered by shade coffee agroforestry for>80 years (Marín-Castro et al., 2016), with 94% of the land area dominated by this land cover. This production system retains some tree cover to provide shade to the coffee, but it exhibits a lower LAI (4.3 m^2 m^{-2}). The IP micro-catchment was largely cleared more than 40 years ago (López-Ramírez et al., 2020). Since then, pastures have been heavily grazed by sheep, goats, and cows (63% of the land area); patches of young secondary forest and built infrastructure (roads and dwellings) occupy the remainder (29 and 8%, respectively).

The soils in these areas are mainly classified as Umbric Andosols derived from volcanic ash, with clay and silty clay as dominant textures (Campos, 2010; Paré and Gerez, 2012). Mineral soil horizons are characterized by low bulk densities (<0.7 g cm^{-3}) across land cover types due to the abundance of non-crystalline materials and organic matter; thick organic horizons (5–15 cm) often overlie the mineral soil in TMCFs, but typically not in pastures and coffee plantations. Soil profiles are generally deep (A + B horizons > 1 m and C + Cr horizons > 10 m on ridges and backslopes) and moderately well developed (Karlsen, 2010), favoring good water storage. The soils in the region are generally underlain by andesitic saprolite, with high permeabilities ranging from 0.05 to 0.08 mm hr^{-1} (Karlsen, 2010; Muñoz-Villers and McDonnell, 2012). Although we did not measure bedrock hydraulic properties here, we observed the presence of saturated saprolite on various road cuts in our study sites. Field-saturated soil hydraulic conductivities are generally higher in TMCF areas in comparison to pasture and shade coffee (López-Ramírez et al., 2020).

2.2. Model set up and data preparation

The SWAT-T model was set up for land use conditions representing the year 2013 (Von Thaden et al., 2019). The land cover map contained seven classifications: pine, mature forest, intermediate forest, crop, pasture, coffee, and human settlements. These land cover classes were input into the model setup using the categories from the SWAT database that are presented in Table B1 in Appendix B. The complete set of characteristics of each category can be reviewed in the SWAT plant database (Neitsch et al. 2011). We used the soil classes from the INEGI database (INEGI, 2007) and field data to populate the soil database in SWAT-T (see Table 2, and Appendix A for more details).

The model was set up to capture the elevation, slope, land use, and soil classes corresponding to the predominant conditions in the Gavilanes catchment and in the four monitored micro-catchments (Table 1). ArcSWAT version 2012 (Winchell et al., 2013) was used to set up the SWAT-T model. A digital elevation model (DEM) with a 15-m resolution (INEGI, 2012) was used to delineate the catchments. The study area was divided into 17 sub-catchments. Two slope classes were defined, using the mean slope as the dividing threshold, with slope ranges of 0–35 % and >35 %, respectively. The sub-catchments were further subdivided into hydrological response units (HRUs). Small HRUs were removed from the model setup and their areas were apportioned to the remaining HRUs (Her et al., 2015); specifically, for land cover, HRUs with shares of the total area below thresholds of 5% for soils, land cover, and soils combinations with shares of the area below thresholds of 20% and for slope, land cover, soils, and slope combinations with shares of the area below thresholds of 20%. These thresholds were selected for retaining the dominant land covers, maintaining individual HRUs for the four monitored micro-catchments, and keeping the computational costs manageable. In total, the model setup resulted in 140 HRUs which represented around 80% of the total area.

We set up three SWAT-T models to evaluate three empirical potential evapotranspiration (PET) methods available in the model to estimate the actual ET. SWAT-T-HA used the Hargreaves method (Hargreaves and Samani, 1985), SWAT-T-PT the Priestley-Taylor formulation (Priestley and Taylor, 1972), and SWAT-T-PM the Penman-Monteith method (Monteith, 1965).

The list of hydroclimatological and spatial data used to set up the SWAT-T models is presented in Table 2. Maximum and minimum daily temperature and daily precipitation data were compiled from ten weather stations: four from the Mexican National Weather Service (SMN, 2020), three from López-Ramírez et al. (2020), and three

Table 2
Input data for the SWAT-T models setup, the data sources, and data processing steps.

Input dataset	Source	Data preparation
Topography	INEGI (2012)	Digital Elevation Model for Mexico in 15 m resolution.
Land use	Von Thaden et al. (2019)	Supervised classification of Landsat imagery from the dry season (>500 ground-based land cover reference data)
Soil data	INEGI (2007); Daniel Geissert (unpublished data, 2010); Nathaniel Looker (unpublished data, 2018)	Soil classes from (INEGI, 2007) and pooling soil data collected in>100 sites.
Climate	SMN (2020); López-Ramírez et al. (2020); UNAM (unpublished data, 2020)	Daily min and max temperature and precipitation. Monthly average of solar radiation, wind speed and relative humidity. (Jan 2014- Dec 2017)
Discharge Gavilanes	Lyssette E. Muñoz-Villers (Unpublished data, 2019)	10 min discharge, resampled (mean) to daily (May 2015-Apr 2017)
Discharge micro catchments	López-Ramírez et al. (2020)	3 min discharge, resampled (mean) to daily (Jul 2016- Aug 2019)
Leaf area index	LAI ground measurements with LI-COR Biosciences (2019); MODIS time series (Myneni et al., 2015)	LAI measurements using LI-COR Canopy Analyzer. MOD15A2 (500 m/8-day from Jun 2014 to Aug 2019)

operated by the National Autonomous University of Mexico (UNAM). Monthly average relative humidity, solar radiation, and wind speed were available only for the UNAM and López-Ramírez (2020) stations, which were assigned to the rest of the weather stations, considering the similarity in terms of elevation and the shortest distance. This monthly data was used to inform the SWAT weather generator.

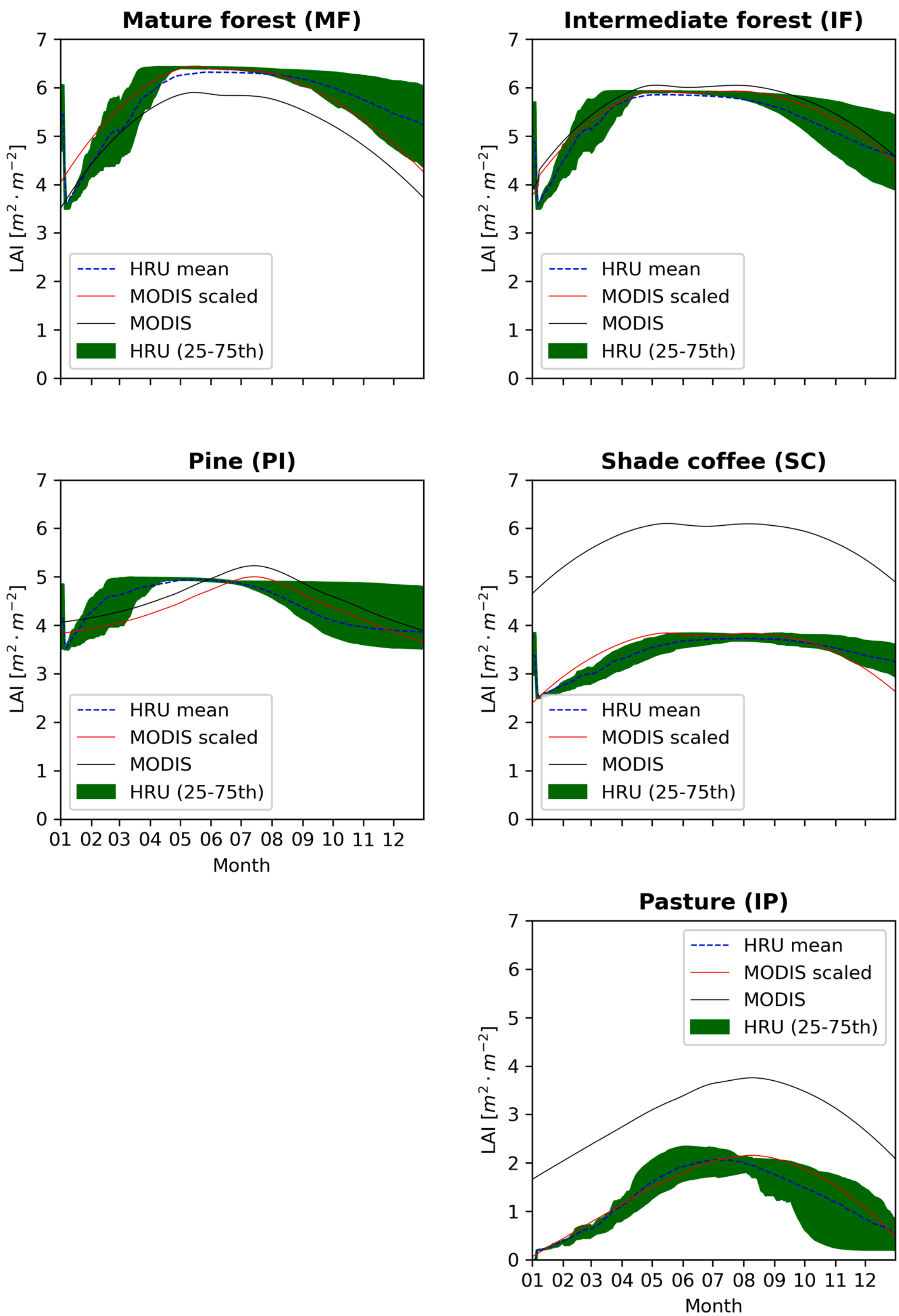

Fig. 2. Annual LAI as simulated by SWAT-T plant growth using calibrated SWAT-T parameters and the rescaled MODIS annual growth model (red line). The black line represents the original smoothed MODIS annual growth model. The green band denotes the area between the 25 and 75th quantile of the corresponding HRUs. (see explanations in the text). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.3. Plant growth and canopy interception calibration

2.3.1. Leaf area index

We combined ground-based leaf area index (LAI) measurements with MODIS satellite products. Maximum LAI was measured with the LAI-2200C Plant Canopy Analyzer (LI-COR, 2019) over a 500-m transect in eight plots with dominant land covers at different elevations: two MF, two IF, two SC, one IP, and one PI (Fig. 1b). These measurements were taken during the rainy season (July 8 to August 20, 2019) in the early morning hours (6:45 a.m. to 8:00 a.m.), before sunset, or under cloudy conditions to avoid scattering effects. The MODIS satellite provide estimates of LAI on a 500-m grid every eight days at a world scale (Myneni et al., 2015). We extracted and analyzed 231 MODIS LAI composites from June 2014 to August 2019 to estimate the average daily LAI values for 8 pixels with homogeneous land covers representative of our LAI sampling sites. Clouds were masked and mean LAI values were predicted for each day of the year by fitting local polynomial regressions (Ripley, 1998) in the *stats* Package (R Core Team, 2019) to the eight-day MODIS LAI estimates pooled across years (2014–2019).

Finally, we rescaled the smoothed daily LAI using the observed LAI for each of our sites. In pine forests, we used the default SWAT max LAI value (5.0) since we did not have an accurate needle-to-shoot area ratio to adjust the observed LAI. In general, we observed that the MODIS algorithm (Myneni et al., 2015) produced acceptable results in the broadleaf forest (MF and IF), while it significantly overestimated LAI in the SC and in the IP (Fig. 2). The smoothed and rescaled MODIS LAI models were used to calibrate the SWAT-T vegetation growth module for simulating LAI (Fig. 2).

Table B1 (Appendix B) presents the SWAT-T model parameters that were adjusted during the manual calibration process. Initially, the minimum LAI (ALAI_MIN) and maximum potential LAI (BLAI) for each land cover class were based on the ground-based MODIS LAI model, which corresponds well with independent measurements in the study area (González-Martínez and Holwerda, 2018). The shape coefficients for the LAI curve (FRGRW1, FRGRW2, LAIMX1, LAIMX2) and the remaining parameters were adjusted by a trial-and-error process, based on values reported in the literature (i.e., Strauch and Volk, 2013; Alemayehu et al., 2017), such that the SWAT-T-simulated LAI mimics the rescaled MODIS LAI.

2.3.2. Plant growth and canopy interception calibration

In the study region, the TMCF consists of a mix of deciduous, semi-deciduous (e.g., leaf-exchanging) and evergreen species of tropical and temperate biogeographical origins (Williams-Linera, 1999; Williams-Linera et al., 2013). Most of the semi-deciduous species are of temperate origin, and the timing of the leaf exchange is not controlled by water availability. Instead, these species shed their leaves during the early and mid-dry season and leaf out during the mid- to late dry season (Borchert et al., 2005). We manually defined the end of the dry season and the beginning of the rainy season (SOS1 = January, and SOS2 = February, respectively), where a new vegetation growth cycle takes place. Two parameters are used to avoid unrealistic results such as a false starts of the growing cycle during the dry season due to short-spell rainfall (Alemayehu et al., 2017).

The SWAT-T parameters that control the shape, the magnitude, and the temporal dynamics of LAI were manually adjusted to reproduce the values of the rescaled MODIS LAI for each land cover class (Table B1 in Appendix B). This procedure was repeated for the parameterization of the maximum canopy storage capacity (CANMX). CANMX was determined for mature forest (3.3 mm), intermediate forest (1.8 mm), pine (1.2 mm), and shade coffee (1.6 mm). These values were obtained from measurements of throughfall and stemflow in these land covers at different elevations in the study catchments (González-Martínez and Holwerda, 2018; Holwerda et al., 2010; Holwerda et al., 2013; Muñoz-Villers et al., 2012, 2015).

2.4. Sensitivity analysis and calibration for streamflow at the Gavilanes catchment

2.4.1. Streamflow at the Gavilanes catchment

We used a two-year daily streamflow series from the Gavilanes catchment streamflow gauge (2 May 2015–30 April 2017). Water levels were measured every 10 min using water level sensors paired with barometric pressure recorders. Recorded levels were converted to streamflow (L s^{-1}) using experimental stage-discharge relationships based on rating curves derived from salt dilution measurements of discharge (Guzmán-Huerta, 2019). The streamflow data were resampled (mean) to daily timesteps (Fig. 3).

2.4.2. Selection of sensitive parameters

In a parameter screening, we ran a global sensitivity analysis (GSA) for the simulation of discharge at the Gavilanes catchment outlet for the SWAT-T models using each PET method to identify the most influential parameters. We explored 23 model parameters that are frequently calibrated in SWAT to simulate discharge (see Arnold et al., 2012; Mehdi et al., 2018; Schürz et al., 2019; Meins, 2013, for a description of the relevant model parameters controlling the water balance). The ranges and types of parameter changes represent typical procedures often found in the SWAT literature. An overview of the model parameters that were identified as influential and that were further used in the model is provided in Table B2 (Appendix B). We employed the *Fourier Amplitude Sensitivity Test* (FAST) algorithm (Cukier et al., 1973) included in the R *fast* package (Reusser, 2015) and integrated as part of the *SWATplusR* workflow (Schürz, 2019) to screen and rank the model parameters. A parameter was selected as important if it was within the most sensitive parameters, employing the Kling–Gupta Efficiency criterion (KGE) for daily streamflow, including its three components (Gupta et al., 2009). The GSA performed for the model parameters of the three SWAT-T models yielded similar results for the first 13 independent parameters. Therefore, we employed 13 parameters during the calibration of the three models (Table B2 in Appendix B).

2.4.3. Identification of non-unique parameter sets (calibration)

To facilitate the analysis of the SWAT-T model performance and separate the effects of particular PET methods on model output, we selected a series of simulations with non-unique parameter sets (Schürz et al., 2019). We simulated daily discharge for each SWAT-T model, using the Latin Hypercube Sampling method with the R package *lhs* (Carnell, 2019), to select 3000 random samples for the parameter combinations of the 13 most influential parameters that we had previously identified. We evaluated simulations using signature metrics and a statistical performance metric (KGE). The ratio of the root mean square error to the standard deviation (RSR, Moriasi et al., 2007) was calculated for five segments of the flow duration curve (FDC): very high (0–5 % days of exceedance), high (5–20%), medium (20–70%), low (70–95%), and very low flows (95–100%), as suggested by Pfannerstill et al. (2014). RSR varies from an optimal value of 0, which indicates a perfect model simulation, to a large positive value. This evaluation method has been shown to lead to an improved selection of good calibration runs since it captures different parts of the hydrograph (Pfannerstill et al., 2014). We simultaneously identified the best 1500 runs out of 3000 for the five-signature metrics (Minimize RSR), followed by the selection and comparison of the best ten simulations with maximum KGE for each SWAT-T model (SWAT-T-HA, SWAT-T-PT, and SWAT-T-PM).

2.4.4. Temporally distributed parameter sensitivity analysis

The temporally distributed parameter sensitivity analysis (TDSA) is used to identify deficiencies in the model structure as a function of time of year (Pfannerstill et al., 2015; Haas et al., 2015; Guse et al., 2014; Guse et al., 2016; Guse et al., 2019). A TDSA was applied to the three SWAT-T models to obtain daily temporal parameter sensitivities, which

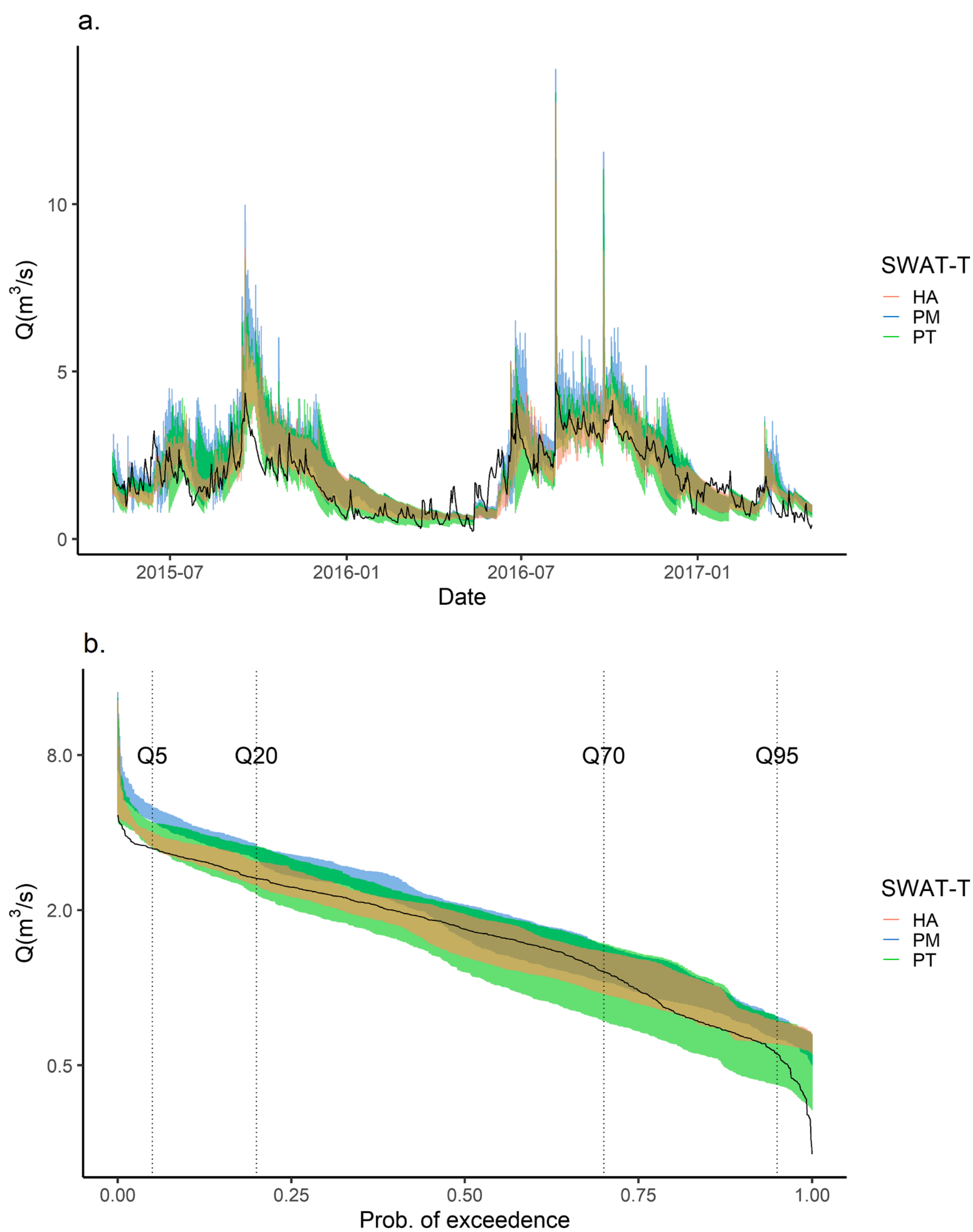

Fig. 3. Calibrated SWAT-T simulations of discharge at the Gavilanes outlet using the best 10 sets of non-unique parameters for each model (SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT) (a) and their corresponding flow duration curves (b), black line represents observed discharge. Figures include an alpha factor = 0.5, adding transparency to the color bands with darker tones indicating areas of overlap. Dotted lines denote the beginning and the end of the FDC segments.

were later summarized at monthly intervals (Guse et al., 2016) to derive information about the dominant hydrological processes during the wet and dry seasons.

We simulated daily discharge for each model in the Gavilanes catchment, again using the Latin Hypercube Sampling to select 1500 random samples for the combinations of eight parameters (Table B3 in Appendix B). From the thirteen parameters used for calibration (Table B2 in Appendix B), seven parameters were selected to assess parameter sensitivity for each hydrologic process: surface runoff, lateral flow, groundwater flow, evaporation, and soil water storage. CANMX was added in the TDSA for rainfall interception. We did not consider the routing parameters (CH_K2, CH_N2, ALPHA_BNK) as they are related to channels. For lateral flow, groundwater, and evaporation we used only parameters previously reported in other SWAT studies (e.g. Guse et al., 2014; Guse et al., 2019), leaving HRU_SLP, DEP_IMP, and EPCO out of the TDSA analysis. The parameters that control surface runoff, soil water storage, lateral flow, and rainfall interception were modified spatially, while the parameters controlling groundwater were altered globally due to the lack of spatially distributed empirical data. For a detailed description of the model parameters and the associated processes, please refer to Neitsch et al. (2011). To study the temporal dynamics of the parameterization of daily discharge simulations for the three PET methods we employed the R package *temPAWN* (Schürz, 2020), which calculates the daily PAWN (Pianosi and Wagener, 2015; the name of the method was derived from the authors names) sensitivity index according to Pianosi and Wagener (2018) for each time step of a simulation. PAWN is moment-independent and is a suitable method for asymmetrically distributed outputs of models (Zadeh et al., 2017). PAWN (Pianosi and Wagener, 2015) has shown to be effective for parameter ranking, including various applications with the SWAT model (Zadeh et al., 2017; Pianosi and Wagener, 2018; Schürz et al., 2019). A dummy variable is artificially introduced that cannot affect the output but provides an indication of the extent of approximation errors. If the sensitivity index of a parameter is smaller than the dummy variable, then we can conclude that this factor is not influential (Pianosi and Wagener, 2018).

2.5. Validation of streamflow in four micro-catchments

2.5.1. Streamflow at four micro-catchments

Within each micro-catchment (Fig. 1c), streamflow and rainfall were measured from 2015 to 2019. For streamflow, V-notch weirs were installed at each micro-catchment outlet, and the water level was logged every 1.5 min using a Solinst water level sensor (model 3001) paired with a barometric pressure recorder (model 3001). We calculated the streamflow (L s^{-1}) using field-derived rating curves generated via volumetric and salt dilution measurements of discharge (Moore, 2005). Please see López-Ramírez et al. (2020) for more details.

2.5.2. Validation of streamflow in four micro-catchments

We evaluated the performance of the three calibrated SWAT-T models (SWAT-T-HA, SWAT-T-PT, and SWAT-T-PM) for streamflow simulation at four sites. All calibrated parameters (identified non-unique parameter sets) were subsequently applied individually (a parameter set at a time) to the SWAT-T models (Mehdi et al., 2018), and the model simulations were evaluated in four HRUs representing the dominant land cover in the corresponding micro-catchments (Table 1): MF (100%), IF (77%), SC (94 %), and IP (63%). The location and weather data of these HRUs are the same as in the corresponding micro-catchments and the HRUs represent the dominant soil, slope, and elevation conditions observed in the micro-catchments. First, we visually compared the simulated and observed FDCs. Second, we estimated the percentage bias (PBIAS) criteria for the annual ratio between average daily runoff and average daily rainfall (Q/P), for the mean annual daily high flow(Q_5), and for the mean annual daily low flow (Q_{95}).

2.6. Validation of evapotranspiration at three sites

2.6.1. Evapotranspiration at three sites

Water and energy fluxes were measured at three sites over different time periods to calculate evapotranspiration (ET). ET was calculated as the sum of rainfall interception (I) and dry-canopy evaporation or transpiration (E_t). In a mature and an intermediate montane cloud forest located at an altitude of about 2100 m (Fig. 1b) E_t values were derived from continuous sapflow measurements from January-November 2007 (37 trees for mature forest and 13 trees for intermediate forest), and I values were quantified using throughfall and stemflow measurements in plots between September 2006 and June 2007 in the intermediate forest and between April 2007 and April 2008 in the mature forest; see Holwerda et al. (2010) and Muñoz-Villers et al. (2012) for more details. In a shaded coffee plantation located at an altitude of 1210 m (Fig. 1b), E_t and I were calculated using the eddy covariance method and measurements of throughfall and stemflow, respectively (September 2006-August 2008); see Holwerda et al. (2013) for more details. In these studies, the E_t results were also expressed as a fraction of the FAO Penman-Monteith reference evapotranspiration (ET_0); this metric facilitates comparison of independent site (vegetation) transpiration rates relative to the evaporative demand of the atmosphere (Allen et al., 1998). The interception results were expressed as a fraction of the mean annual rainfall (P). For the mature forest E_t was 787 mm (92 % of ET_0) and I was 613 mm (18% of P), in the intermediate forest E_t was 788 mm (92% of ET_0) and I was 315 mm (9% of P), and in the shade coffee E_t was 992 mm (89% of ET_0) while I was 110 mm (8% of P).

2.6.2. Validation of evapotranspiration at three sites

We compared the evapotranspiration (ET) predicted by the three SWAT-T models with ET estimated from meteorological data and locally derived indices of interception divided by precipitation (I/P) and transpiration divided by reference evapotranspiration (E_t/ ET_0). First, we calculated the mean annual precipitation (mP) and the mean reference evapotranspiration (mET_0) following Allen et al. (1998) and using the available meteorological data for the study period (4 years) at three sites. Next, the interception (I) was estimated as the product of I/P and mP, and the E_t was obtained by multiplying mET_0 by E_t/ET_0. Finally, ET was estimated by adding I and E_t (Table 5). These ET values were compared with ET predicted in the corresponding HRUs by the calibrated SWAT-T models, using the best set of previously identified parameters (Table B4, in Appendix B. Model parameters). Additionally, we estimated the simulated rainfall interception (Sim I SWAT-T) as the difference between ET predicted by the calibrated SWAT-T models and ET predicted by the calibrated SWAT-T models when the canopy storage capacity (CANMX) is set to zero (Table 5).

3. Results

3.1. Leaf area index calibration

We evaluated the degree of agreement between daily MODIS LAI and the LAI simulated by the calibrated SWAT-T using visual comparisons and statistical measures. From the qualitative inspection (see Fig. 2), it is apparent that the annual growth cycle of each land cover class from the calibrated SWAT-T model corresponds well with the MODIS LAI model rescaled with field data. In quantitative terms, the calibrated models exhibited generally strong correlations—0.67 (MF), 0.89 (IF), 0.92 (IP), 0.47 (PI), and 0.64 (SC) — during the calibration period. The greatest mismatch in LAI was found for PI, most likely because the pine trees have bundles of needles rather than broad leaves. The 25-75th percentiles in Fig. 2 show the effect of temperature, primarily related to the elevation gradient, on the computation of LAI. SC exhibited narrower bands than the other land covers because it is mainly distributed between 1100 and 1500 m a.s.l. It can be observed that the LAI bands become wider at the end of the annual cycle because the growth model uses the accumulated number of heat units.

3.2. Streamflow calibration at the Gavilanes catchment

Visual comparisons of daily simulated streamflow hydrographs with observed streamflow over the calibration period for the three models (SWAT-T-HA, SWAT-T-PT, and SWAT-T-PM) suggest a good fit (Fig. 3a). The quantitative results presented in Table 3 indicate that SWAT-T-HA had a higher KGE and NSE and a lower PBIAS than the other two models. The SWAT-T-PM model had the poorest ability of the calibrated models to reproduce observed streamflow (lower NSE and KGE and larger PBIAS). These statistical results were consistent with the results from the signature metric analysis, where SWAT-T-HA consistently had the lowest RSR for the five parts of the FDC, and SWAT-T-PM the largest. These results are also evident in Fig. 3b, where we can observe narrower uncertainty bands for SWAT-T-HA. At very low flows, however, the three models produced similar results, exhibiting high RSR values.

Table 3 indicates that the three SWAT-T models overestimated the observed discharge (mean PBIAS for SWAT-T-PM = 24%, mean PBIAS for SWAT-T-PT = 12%, mean PBIAS for SWAT-T-HA = 5%). The water balance presented in Table C1 (Appendix C) confirms these findings, and shows that SWAT-T-PM predicted the lowest PET (807 mm) and ET (665 mm) and the highest total streamflow (1407 mm), followed by SWAT-T-PT (PET = 977 mm, ET = 926 mm, total streamflow = 1203 mm) and SWAT-T-HA (PET = 1179 mm, ET = 1098 mm, total streamflow = 1085 mm). Overall, the water balance (Table C1) indicates a low influence of surface runoff (4 – 16 % of total streamflow in all models) and a dominance of lateral flow (40 to 72% of total streamflow in all models) in the streamflow response.

Table 3
Final selection of best model calibration runs after applying 5 FDC performance metric-based selection. The mean value for all indices of fit is presented for each model (SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT) for comparison of performance.

Sim	PET	Calibration run	KGE	NSE	PBIAS	R^2	RSR				
							Q_5	Q_{5-20}	Q_{20-70}	Q_{70-95}	Q_{95}
1	PM	831	0.51	0.07	22.70	0.61	5.36	3.39	1.03	0.88	3.29
2	PM	2293	0.48	0.13	19.60	0.67	5.67	3.87	0.85	0.53	2.75
3	PM	1146	0.42	−0.08	27.10	0.64	5.97	4.26	1.22	0.57	2.24
4	PM	525	0.49	0.06	31.00	0.70	6.27	3.78	1.31	1.91	2.96
5	PM	1480	0.44	0.06	21.00	0.67	6.44	4.70	0.61	0.92	3.33
6	PM	875	0.38	−0.05	28.30	0.71	7.13	4.93	1.06	1.08	2.74
7	PM	2216	0.41	0.03	21.90	0.69	7.17	4.26	0.85	0.78	3.10
8	PM	2871	0.44	0.00	28.70	0.68	7.31	3.57	1.26	1.48	1.93
9	PM	2822	0.42	0.12	15.60	0.69	8.31	3.35	0.53	0.66	2.86
10	PM	515	0.45	−0.07	28.50	0.63	8.52	2.86	1.11	1.86	3.27
	Mean		0.44	0.03	24.44	0.67	6.81	3.90	0.98	1.07	2.85
1	PT	2953	0.63	0.20	12.20	0.48	1.24	2.14	0.72	0.64	2.87
2	PT	2311	0.58	0.34	18.70	0.71	3.51	3.80	0.79	0.62	2.52
3	PT	2873	0.59	0.21	27.20	0.66	3.52	3.50	1.12	2.07	3.02
4	PT	764	0.64	0.47	15.50	0.75	3.96	3.00	0.56	0.76	2.86
5	PT	1813	0.71	0.48	−16.20	0.70	4.23	0.99	1.15	1.44	0.90
6	PT	1938	0.68	0.45	9.80	0.66	4.36	1.59	0.40	0.78	3.33
7	PT	716	0.63	0.47	11.90	0.73	5.34	2.38	0.40	0.83	3.30
8	PT	456	0.58	0.23	23.20	0.66	5.74	2.63	0.94	1.77	3.33
9	PT	1544	0.66	0.42	10.40	0.67	6.60	1.30	0.30	0.97	3.32
10	PT	600	0.59	0.36	9.90	0.68	6.86	1.77	0.37	0.60	2.73
	Mean		0.63	0.36	12.26	0.67	4.54	2.31	0.67	1.05	2.82
1	HA	2059	0.74	0.47	14.90	0.62	1.36	1.42	0.62	1.69	2.55
2	HA	2312	0.80	0.61	8.20	0.70	1.51	1.31	0.31	1.02	2.57
3	HA	1880	0.83	0.65	−1.10	0.71	2.57	0.50	0.49	0.69	3.28
4	HA	770	0.71	0.51	8.10	0.69	2.79	2.05	0.44	0.42	2.52
5	HA	1513	0.84	0.69	3.80	0.74	2.84	0.58	0.22	0.88	3.33
6	HA	2787	0.72	0.44	6.00	0.61	2.91	1.44	0.40	0.73	3.16
7	HA	546	0.71	0.47	1.30	0.64	3.79	1.00	0.47	0.44	2.11
8	HA	541	0.71	0.52	7.90	0.69	5.50	1.18	0.27	0.92	3.43
9	HA	2638	0.70	0.51	3.20	0.69	5.54	0.84	0.39	0.48	2.63
10	HA	686	0.72	0.51	−5.90	0.67	5.75	0.45	0.69	0.47	2.03
	Mean		0.75	0.54	4.64	0.67	3.46	1.08	0.43	0.77	2.76

3.3. Temporally distributed parameter sensitivity analysis

The monthly averages of sensitivity results (Fig. 4) for the six model processes were very similar for the three models (SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT). A marked dominance of lateral flow was detected throughout the year, with an important contribution of groundwater. The highest significance of groundwater parameters occurred in October and November, during the transition between the wet and dry seasons (hydrograph recession). A low influence of surface runoff was observed, except in the middle of the rainy season (July-September) when this runoff component plays a slightly more significant role. Evapotranspiration, interception, and soil water storage parameters had a low influence on streamflow (Fig. 4); this result was expected because the water deficit for evapotranspiration is low throughout the year.

3.4. Validation of streamflow in the four micro-catchments

The selection of a specific PET method influences the calculations of ET and in turn affects the water balance. Qualitatively, the three models (SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT) showed good performance in simulating different phases of the FDC for MF and SC. Additionally, all three models accurately simulated the high, mid and low flows in the IF (Fig. 5). However, SWAT-T clearly failed to capture the very low flows observed in IF. An unexpected result was that the model largely overestimated the low and very low flow dynamics in the IP, probably due to the global calibration of groundwater parameters. The statistical results (Table 4) confirm these findings, where the three SWAT-T models exhibited extremely high mean PBIAS (>400) for mean annual daily low flow (Q_{95}) in the IP micro-catchment. Additionally, the three models underestimated Q_{95} in the MF and IF HRUs. For mean annual daily high flow (Q_5) the SWAT-T-PM and SWAT-T-PT models exhibited slightly better performance than SWAT-T-HA. For average water balance conditions (Q/P), the SWAT-T-HA model produced a general underestimation, while SWAT-T-PT showed mixed results, and the SWAT-T-PM method produced a general overestimation.

3.5. Validation of evapotranspiration at three sites

Comparisons of ET as predicted by the best calibrated models (SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT) with ET estimated based on I/P and E_t/ET_0 derived from local measurements (see section 2.6 for explanation) revealed that ET simulated by SWAT-T-HA more closely matched ET based on local data than the other two models (Table 5). However, our results also show that SWAT-T-HA overestimates ET in SC by 247 mm (22 %) at lower elevations (1210 m a.s.l.), while underestimating ET in MF by 310 mm (22 %) and in IF by 58 mm (5%) at about 2100 (m a.s.l.). SWAT-T-PM and SWAT-T-PT caused a higher underestimation of ET for the same forest sites. The simulated rainfall interception capacity varied from 36 mm yr^{-1} (Sim I SWAT-T-PT) to 109 mm yr^{-1} (Sim I SWAT-T-PM), representing<4% of the water balance for all sites and models (Table 5).

4. Discussion

4.1. Validation of calibrated SWAT-T models with local evapotranspiration data at three sites, with focus on the performance of the PET methods available in the model

The adequate selection of a PET method and the accurate simulation of leaf area index and rainfall interception influence the computation of

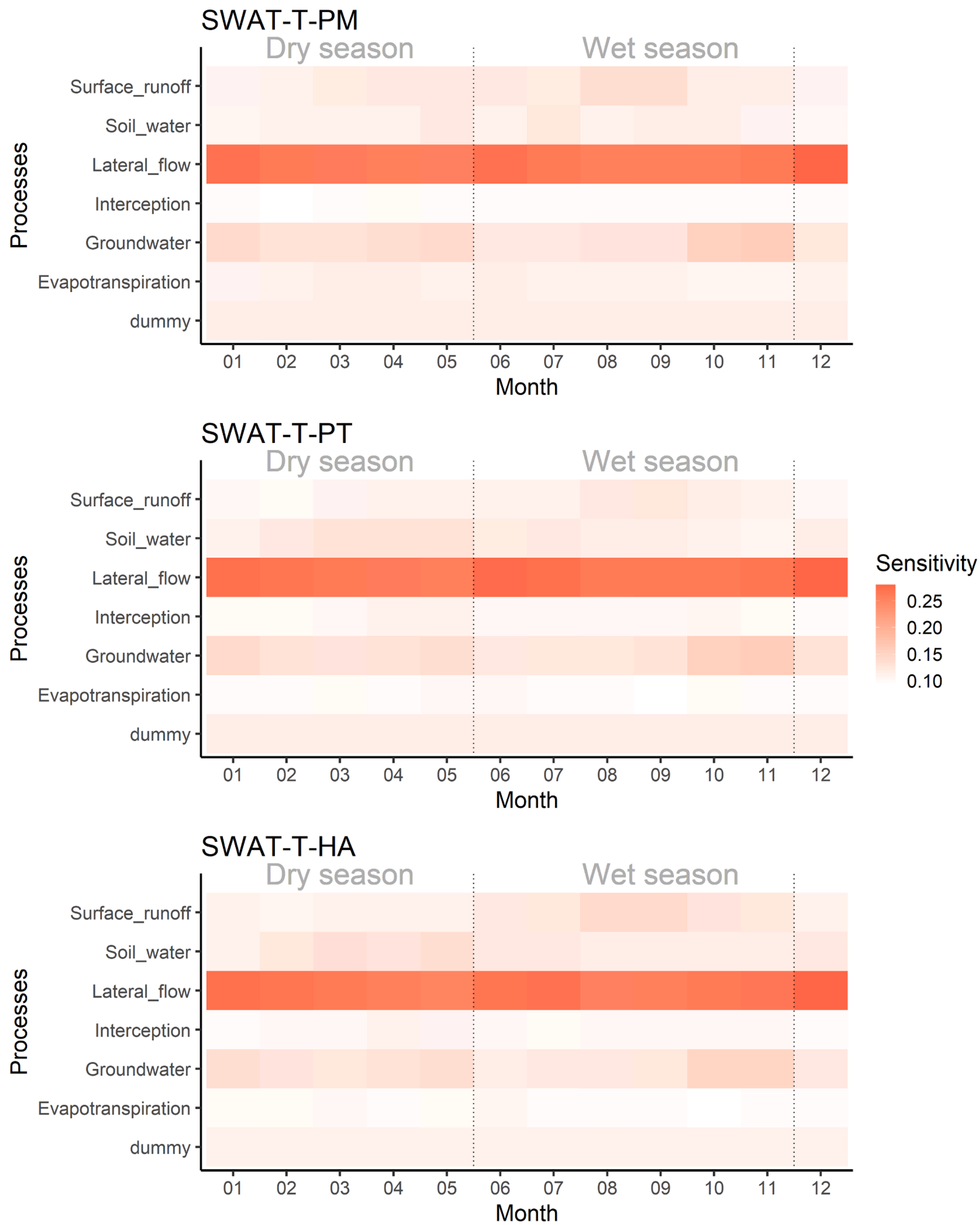

Fig. 4. Mean monthly averaged parameter sensitivities for six model processes and for the three SWAT-T models evaluated at the Gavilanes catchment. Processes sensitivity is presented relative to a dummy variable, processes with a darker tone had a correspondingly more significant effect on streamflow. Dotted lines denote the beginning and the end of the wet and dry seasons.

evapotranspiration in the SWAT-T model and affect the water balance. SWAT-T (Alemayehu et al., 2017) was capable of accurately simulating LAI in the study area (Fig. 2). Overall, the rescaled MODIS LAI model effectively represented the annual growth dynamics when compared with independent LAI observations. In mature forest, the average LAI measured for the dry and wet seasons was 4.1 and 5.9, respectively (González-Martínez and Holwerda, 2018). The LAI model also provided good results when compared with shade coffee plantations where the minimum LAI was 2.7 in December 2016 and the maximum LAI was 3.9 in July 2017 (Holwerda and Meesters, 2019). Furthermore, the SWAT-T-simulated LAI corresponded well with the MODIS LAI annual growth model (an average R^2 of 0.71 for all land covers).

The formulation of SWAT-T did not allow realistic simulation of the effect of rainfall interception in forested areas. Rainfall interception in the study region has been reported in the literature to be about 18 and 9% of P in mature and intermediate forests, respectively (Holwerda et al., 2010; Muñoz-Villers et al., 2012) and around 8% of P in a shade coffee plantation (Holwerda et al., 2013). These values strongly suggest that in these environments, interception is an important component of the water balance for mature forests and that the conversion of mature forests to other land covers such as shade coffee is associated with a decrease in rainfall interception as high as 10% of P. All of the SWAT-T models in our study underestimated rainfall interception ($I < 4\%$ of P, for all land covers, see Table 5). Furthermore, the temporally distributed sensitivity analysis (Fig. 4) indicated that interception has a low effect on streamflow for all the SWAT-T models. These results can be explained

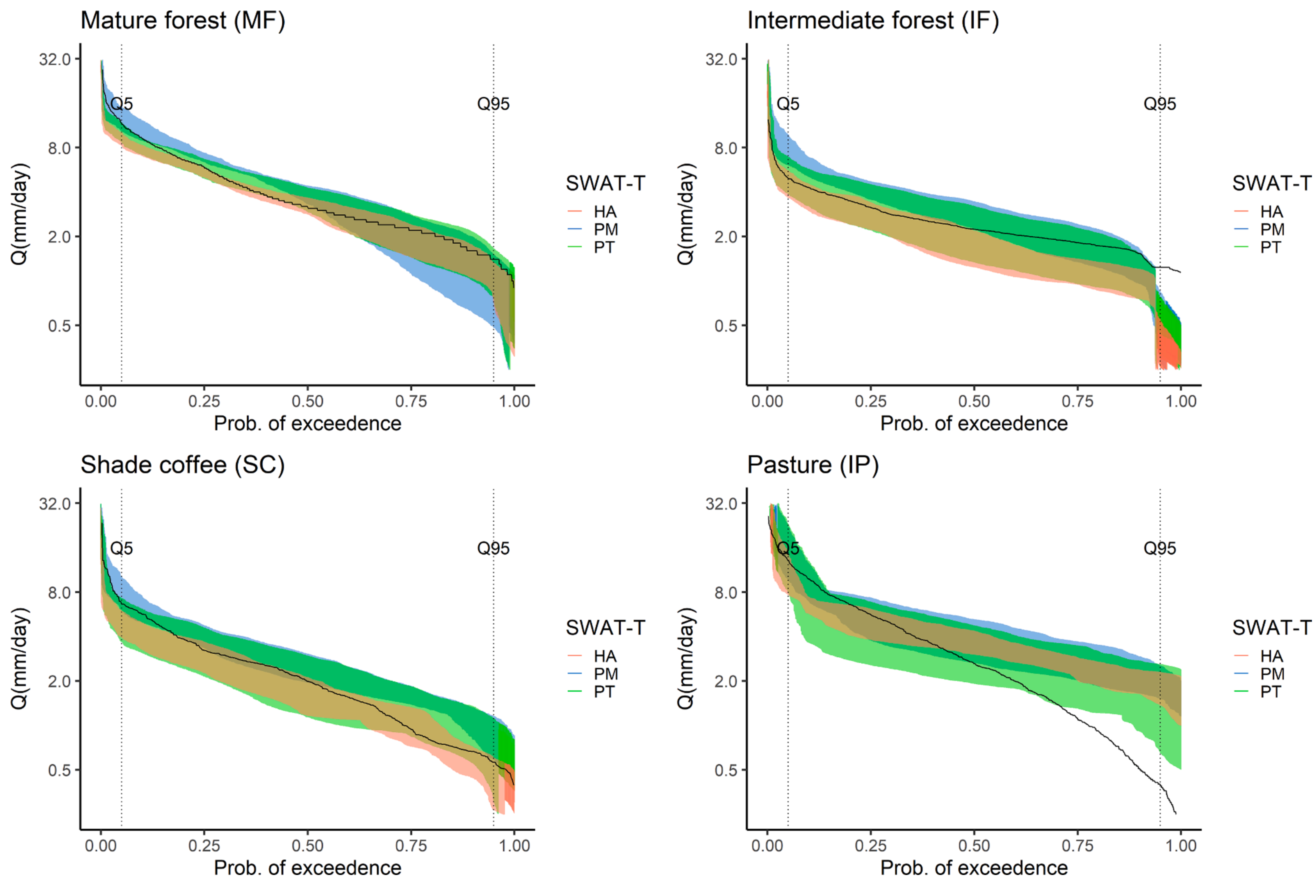

Fig. 5. Validation of streamflow simulations with calibrated models (SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT) at four micro-catchments with contrasting land covers, black line represents observed discharge. Figure includes an alpha factor = 0.5, this adds transparency to the color bands and darker tones of the colors represent overlapping areas. Dotted lines denote the mean annual daily high (Q_5) and low flows(Q_{95}).

Table 4
PBIAS for different streamflow metrics in four monitored micro-catchments. Positive values indicate overestimation, values close to 0 are considered excellent and up to 25% are considered satisfactory (Moriasi et al., 2007).

	Metric	MF	IF	SC	IP
SWAT-T-PM	PBIAS Q/P	4.30	5.19	24.08	2.83
	PBIAS Q_5	11.53	47.59	9.91	19.45
	PBIAS Q_{95}	−17.01	−55.12*	53.12	476.00*
SWAT-T-PT	PBIAS Q/P	−7.82	−14.97	0.98	−9.62
	PBIAS Q_5	−6.08	19.18	−12.03	−2.88
	PBIAS Q_{95}	−11.42	−62.59*	24.54	420.80*
SWAT-T-HA	PBIAS Q/P	−20.15	−39.12	−23.99	−18.64
	PBIAS Q_5	−16.35	−5.52	−26.90	−2.29
	PBIAS Q_{95}	−23.43	−72.17*	−20.66	410.20*

* Systematic errors.

by the fact that SWAT-T limits daily evaporation from the canopy by daily PET (Fig. 1 of Abiodun et al., 2018; Neitsch et al., 2011) and thus underestimates the actual evaporation from the wet forest canopy (see below for further discussion). This point requires further review, and we suggest that the SWAT-T model should allow the canopy storage capacity to empty daily, as assumed by simpler, widely accepted interception models such as the Liu model (Liu, 2001).

More attention should be given to the appropriate selection of PET methods for SWAT model applications. Although we found low influence of evapotranspiration (ET), interception (I), and soil water storage parameters in the daily sensitivity analysis (Fig. 4), we observed that the PET method impacted model performance (Table 3) and significantly affected the water balance (Table C1). ET accounts for 30–40% of rainfall in the headwater catchments where most recharge occurs (Muñoz-Villers et al., 2012; Muñoz-Villers et al., 2015; Muñoz-Villers et al., 2016). These ET results agree with values of ET (19–50%) reported in other tropical montane mesoscale catchments (i.e., Crespo et al., 2012; Beck et al., 2013). We attribute the low influence of ET, I, and soil water storage parameters to the availability of water throughout the year and the limited capacity of the SWAT-T model to simulate interception in forested areas. Nevertheless, our findings support that the choice of the PET estimation method influences the SWAT-T model performance to simulate discharge, as suggested by Alemayehu et al. (2015).

The models parameterized with radiation-based methods, SWAT-T-PT and SWAT-T-PM, significantly underestimated ET in the upper parts of the cloud forest zone (Table 5) and consequently limited the models' capacity to accommodate evaporation from the forests. In the lower areas fog occurrence is practically absent (only 2% of the time, Holwerda 2010). Nonetheless, fog is more frequent (around 32%, Alvarado-Barrientos et al., 2014) in the middle and upper sections of the cloud forest belt, and it is probable that a higher presence of fog is associated with higher epiphyte densities that may contribute to higher canopy storage capacities in these areas (González-Martínez and Holwerda, 2018). These high interception capacities may in turn enhance wet canopy evaporation in forested areas driven by negative (downward) sensible heat flux rather than sustained by radiant energy alone (Stewart, 1977; Mizutani et al., 1997). The aerodynamic term in the PM equation accounts for sensible heat advection. However, in SWAT-T the PM method will underestimate wet-canopy evaporation, because it uses the surface resistance of alfalfa (100 s m^{-1}) (Neitsch et al., 2011), while the actual surface resistance is zero under such conditions. Moreover, this PET method calculates aerodynamic resistance considering a crop

Table 5
Comparison of simulated ET (best calibration run for each model: SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT) with ET derived from local measurements.

	Mature forest	Intermediate forest	Shade coffee
Elevation (m a.s.l.)	2060	2140	1210
mP (mm yr^{-1})	2841	2841	1898
mET_0 (mm yr^{-1})	919	919	1090
I/P	0.18	0.09	0.08
E_t/ET_0	0.92	0.92	0.89
I (mm yr^{-1})	508	255	151
E_t (mm yr^{-1})	845	847	968
ET (mm yr^{-1})	1354	1102	1119
Sim ET SWAT-T-PM[c] (mm yr^{-1})	726	708	845
Sim ET SWAT-T-PT[d] (mm yr^{-1})	934	933	1138
Sim ET SWAT-T-HA[e] (mm yr^{-1})	1044	1044	1366
Sim I SWAT-T-PM[f] (mm yr^{-1})	109	72	30
Sim I SWAT-T-PT[f] (mm yr^{-1})	36	37	63
Sim I SWAT-T-HA[f] (mm yr^{-1})	57	48	57

mP = mean annual precipitation over the studied period (n = 4 years); mET_0 = mean annual FAO Penman-Monteith reference evapotranspiration over the studied period (n = 4 years); I/P and Et/ET_0 derived from local measurements see section 2.3 for explanation; I = mP × I/P; $E_t = mET_0 \times E_t/ET_0$; ET = I + E_t.
[a] Holwerda et al. (2010) and Muñoz-Villers et al. (2012).
[b] Holwerda et al. (2013).
[c] Calibration run 831, see Table B4 in Appendix B for details.
[d] Calibration run 2953, see Table B4 in Appendix B for details.
[e] Calibration run 2059, see Table B4 in Appendix B for details.
[f] Calculated as the difference between ET simulated by the calibrated SWAT-T models and ET simulated by the calibrated SWAT-T models with a CANMX set to 0, no canopy storage capacity.

height of 40 cm (Neitsch et al., 2011). This yields a value for aerodynamic resistance about 5 times larger than that of forest, contributing to the underestimation of PET. In the PT method, the aerodynamic term is represented by the PT coefficient (1.28, Neitsch et al., 2011). However, Shuttleworth and Calder (1979) suggested that this is not sufficient to account for the high evaporation rates from wet forest canopies (i.e., energy input from sensible heat advection can be several times greater than that from radiation).

The model parameterized with the temperature-based method SWAT-T-HA generally produced higher ET estimates (Table 5). Higher values of PET compensate for the low capacity of the model to represent the evaporation of intercepted rainfall in forests but also lead to overestimation of transpiration rates in other land covers. It is likely that the PT method also overestimates forest transpiration. When using the PT and HA methods, SWAT-T does not explicitly model the effect of the LAI on transpiration for LAI > 3.0 (Neitsch et al., 2011). Therefore, the ET values simulated by SWAT-T for MF and IF are very similar when using the HA and PT methods (Table 5). Both methods underestimate the effect of vegetation cover on hydrological flows. In conclusion, the SWAT-T-HA model performed better, but not for the right reasons, and the use of this approach overestimated ET in areas with lower interception capacity such as shade coffee, especially in the lower parts of the catchment where the presence of fog is less important.

4.2. Assessment of the accuracy of the SWAT-T model to simulate streamflow at the outlet of the Gavilanes catchment (34 km^2)

The SWAT-T model was able to satisfactorily simulate streamflow during the calibration period as shown in Fig. 3 and confirmed by acceptable goodness of fit indices (Table 3). The incorporation of a multi-metric framework that considers five segments of the FDC (Pfannerstill et al., 2014) improved the identification of the best parameter sets. However, calibrating the models was challenging for very high and low flow conditions. In our study, we were able to improve the model calibration by relaxing the conditions for the simultaneous selection of the best 50% quantile for the five segments of the FDC. While other researchers have used more restrictive conditions for selecting simulations, including the best 20% quantile by Pfannerstill et al. (2014), and the best 25% quantile by Guse et al. (2020), we observed that our more relaxed approach was effective in enhancing overall model performance and parameter identification.

The temporally distributed sensitivity analysis suggested a low influence of surface runoff, an intermediate influence of groundwater, and the dominance of lateral flow in estimating streamflow response (Fig. 4). This finding agrees with the water balance (Table C1), which indicates that surface runoff accounts for about 4 to 16% of the total streamflow, while lateral subsurface flow controls around 40 to 72% of the streamflow response. This finding agrees with a recent study in a neighboring catchment of similar size (60.6 km^2), where López-Hernández (2019) examined 159 storm events and found that only 3–7% of total streamflow were contributed by surface runoff. Nevertheless, Muñoz-Villers et al. (2016) reported long mean baseflow transit times (between 1.2 and 2.7 years) for the Gavilanes catchment and 11 nested subcatchments (0.1 to 34 km^2), suggesting that deep subsurface flow paths, rather than shallow lateral flow, is the dominant flow path for runoff generation (Muñoz-Villers and McDonnell, 2012, 2013). These results agree with the high permeability observed at the soil–bedrock interface (5–30 mm h^{-1}).

A structural deficit in the SWAT-T model can explain the contradiction in between the SWAT-T prediction of lateral flow as the most important process, when evidence suggests that deep subsurface flow paths are dominant. In SWAT-T, groundwater is divided into a shallow and a deep aquifer. The shallow aquifer represents an unconfined aquifer that may discharge into the channel. The deep aquifer is described as a confined aquifer that does not contribute to the streamflow (Neitsch et al., 2011). The SWAT-T model's reliance on a single active groundwater storage compartment is insufficient for describing both the fast recession phase and the low flow with the same parameter set (Luo et al., 2012; Guse et al., 2014; Pfannerstill et al., 2014). In this regard, the parameter GWDELAY presents a tradeoff between both discharge phases (Table B3 and Table B4 in Appendix B). If GWDELAY represents a fast recession with a short number of days, the model produces a large amount of water available for contribution to streamflow. On the other end, a longer GWDELAY is required to represent the slow water release during the long low flow periods (Pfannerstill et al., 2014). Streamflow measurements in our study catchment indicated a long recession phase, so the calibration resulted in a large value for GWDELAY (around 400 days, see Table B4 Appendix B). Consequently, the shallow aquifer releases the water more slowly, and the system behaves more like a semipermeable layer, providing conditions that inaccurately imply the dominance of lateral flow (Hu and Li, 2018). In summary, local evidence suggested that discharge in our catchment is determined by the volume of water in storage and should be characterized by nonlinear rate coefficients (Kirchner, 2009), while SWAT-T assumes that the variation in groundwater flow is linearly related to the rate of change in water table height (Neitsch et al., 2011).

4.3. Validation of the calibrated SWAT-T models with streamflow observed in four nested micro-catchments (0.137–0.446 km^2) within the study area

Examining the performance of the SWAT-T model simulations of streamflow in micro-catchments with contrasting land cover is useful, since these results can help to better understand the strengths and limitations of the model at HRU scale. Some authors have noted that SWAT shows a tendency to underperform in either very low or very high discharge events, particularly in forested areas (Qiu et al., 2012). The three models (SWAT-T-HA, SWAT-T-PM, and SWAT-T-PT) were capable of satisfactorily estimating annual runoff coefficients and mean annual daily high flows (Q_5) in all land covers, although the best results were obtained for MF and SC (Fig. 5, Table 4). Major discrepancies were observed for the mean annual daily low flows, especially in the IF and IP micro-catchments (Table 4).

Previous studies in this region suggest that managed land covers such as pasture can increase the variability of the streamflow response, but indicators such as the mean annual runoff coefficient (Q/P) and the mean annual daily high flow (Q_5) tend to be largely determined by elevation and slope (López-Ramírez et al., 2020; Muñoz-Villers et al., 2013). This phenomenon was captured well by SWAT-T models, where higher runoff coefficients were observed at higher elevations (Ramírez et al., 2017; Sáenz et al., 2014). For slope, SWAT-T predicted higher peak flows in steeper slopes, which is consistent with the higher mean annual daily high flows (Q_5) observed in steeper slopes by López-Ramírez et al. (2020) and Muñoz-Villers et al. (2013). Steeper slopes have also been associated with higher quickflow and lower time to peak responses (Mu et al., 2015; Nainar et al., 2018). Overall, the calibrated SWAT-T models produced acceptable Q/P and mean annual daily high flows (Q_5) predictions (PBIAS < 25) for most of the land covers. We attribute this result to the dense network of climate stations available (Fig. 1) and the use of slope in the definition of HRUs.

In terms of dry-season baseflow, local studies indicate that both mature and intermediate forests exhibit higher mean annual daily low flows(Q_{95}) than pastureland and shade coffee (López-Ramírez et al., 2020; Muñoz-Villers et al., 2013), which is consistent with the "sponge-effect hypothesis" for forests (Bruijnzeel, 2004; Muñoz-Villers and McDonnell, 2013; Ogden et al., 2013; Muñoz-Villers et al., 2015). Our SWAT-T results, however, point in the opposite direction, with the model substantially overestimating mean annual daily low flow in the IP micro-catchment for all models (PBIAS > 400, see Table 4), and underestimating mean annual daily low flow for all the models in the MF and IF micro-catchments (Table 4). This finding has serious implications when using hydrological models to analyze the effects of land use changes (e.g., deforestation, reforestation, and natural regeneration) on water budgets. We attribute this result to the dominance of global calibration of parameters controlling lateral and groundwater flow (Fig. 4). Spatially distributed empirical data is rarely available for these processes.

4.3.1. *Potential limitations of this approach.*

In our global sensitivity analysis, we did not consider the impact of model setup (number of sub-basins and HRUs). However, our study addresses the uncertainty associated with model parameterizations, allowing the visualization of the effects of non-unique parameter sets. Many authors recognize that the impact of this factor is generally more significant than model setup by several orders of magnitude (Schürz et al., 2019; Jha et al., 2004). The selection of HRUs considers a land cover threshold of 5% given the significant computation time required to run this analysis. This threshold was used to reduce the computational costs without significantly compromising the simulation accuracy. However, we recognize that this practice reduces the effects of small land covers at the outlet (Schürz et al., 2019). Uncertainty in data such as the rainfall interception and evaporation used in this study may be very high due to the complex terrain. Finally, we were not able to completely isolate the response of unique HRUs because the four studied micro-catchments were not entirely covered by a unique land use and slope class (Table 1). These heterogeneities may have influenced our results.

5. Conclusions

SWAT-T was capable of accurately mimicking the annual vegetation growth, simulating streamflow at the main outlet, and predicting a low influence of surface runoff in zones of tropical montane cloud forest in central Veracruz, Mexico. However, the model incorrectly predicted the dominance of lateral flow versus deep groundwater flow as observed by isotope-based studies. Our results thus support the hypothesis that the single active aquifer formulation of SWAT is unable to adequately reproduce the complex response in groundwater dominated catchments.

The model parameterized with the HA method for PET computation (SWAT-T-HA) produced the best goodness-of fit-results for streamflow and evapotranspiration. All the SWAT-T model variants resulted in an underestimation of forest evaporation under the wet conditions of the cloud forest zone of central Veracruz. Temperature-based PET methods (SWAT-T-HA) perform better in these areas, but for the wrong reasons since the use of these approaches overestimates PET and ET rates in areas with lower interception capacity (i.e. non-forested). The incorporation of simpler, and more widely-accepted, interception models such as Liu model (Liu 2001) may improve these results for the right reasons.

SWAT-T satisfactorily reproduced the Q/P and mean annual daily high flows in contrasting land covers (MF, IF, SC, IP). However, the model overestimated the mean annual daily low flow in the IP micro-catchment and underestimated it in the MF and IF micro-catchments. This result suggests that the use of the SWAT-T models in cloud forests underlain by permeable substrates, where ground water flow controls the hydrological response, requires improvement in calibration procedures to incorporate local knowledge regarding the land use effects on the baseflow sustenance. Overall, structural improvements informed by field data are required to better understand and model the hydrology in SWAT-T.

Our findings showed that the incorporation of local datasets from multiple locations and processes during the modeling calibration and validation is key to advance hydrological modeling tools.

6. Funding sources

- This research was founded by a US National Science Foundation Award Number: 1313804. This National Science Foundation grant has been awarded to the University of New Hampshire with a subcontract agreement with Michigan Technological University.
- Funding was also provided by Michigan Technological University through the Doctoral Finishing Fellowship Award.
- The research was partially supported by SEP-CONACYT- Mexico (Number 106788) grant to Daniel Geissert.

CRediT authorship contribution statement

Sergio Miguel López-Ramírez: . **Alex Mayer:** Conceptualization, Supervision, Methodology, Writing – original draft, Funding acquisition. **Leonardo Sáenz:** Conceptualization, Supervision, Writing – original draft. **Lyssette Elena Muñoz-Villers:** Conceptualization, Writing – original draft. **Friso Holwerda:** Validation, Formal analysis, Writing – original draft. **Nathaniel Looker:** Formal analysis, Investigation, Writing - review & editing. **Christoph Schürz:** Software, Resources, Writing - review & editing. **Z. Carter Berry:** Investigation, Writing - review & editing. **Robert Manson:** Project administration, Writing - review & editing. **Heidi Asbjornsen:** Funding acquisition, Writing - review & editing. **Randall Kolka:** Conceptualization, Writing - review & editing. **Daniel Geissert:** Investigation, Writing - review & editing. **Carlos Lezama:** . **:** Investigation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the landowners of the micro-catchments for permitting us to work on their land. We would also like to thank León Rodrigo Gómez Aguilar and Víctor Vásquez for their great help in the field. Finally, we would like to acknowledge the valuable comments of Tadesse Alemayehu during the calibration of the SWAT-T model.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2021.126781.

References

Abiodun, O.O., Guan, H., Post, V.E.A., Batelaan, O., 2018. Comparison of MODIS and SWAT evapotranspiration over a complex terrain at different spatial scales. Hydrol. Earth Syst. Sci. 22, 2775–2794. https://doi.org/10.5194/hess-22-2775-2018.

Alemayehu, T., Griensven, A.V., Bauwens, W., 2015. Evaluating CFSR and WATCH Data as Input to SWAT for the Estimation of the Potential Evapotranspiration in a Data-Scarce Eastern-African Catchment. J. Hydrol. Eng. 21 (3), 05015028. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001305.

Alemayehu, T., Griensven, A., Woldegiorgis, B.T., Bauwens, W., 2017. An improved SWAT vegetation growth module and its evaluation for four tropical ecosystems. Hydrol. Earth Syst. Sci. 21, 4449–4467. https://doi.org/10.5194/hess-21-4449-2017.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56. Last consulted on November 4, 2020 at: http://www.fao.org/docrep/X0490E/X0490E00.htm.

Alvarado-Barrientos, M.S., Holwerda, F., Asbjornsen, H., Dawson, T.E., Bruijnzeel, L.A., 2014. Suppression of transpiration due to cloud immersion in a seasonally dry Mexican weeping pine plantation. Agric. For. Meteorol. 186, 12–25. https://doi.org/10.1016/j.agrformet.2013.11.002.

Archibald, J.A., Walter, M.T., 2013. Do Energy-Based PET Models Require More Input Data than Temperature-Based Models? -An Evaluation at Four Humid FluxNet Sites. J. Am. Water Resour. Assoc. 50 (2), 497–508. https://doi.org/10.1111/jawr.12137.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., Griensven, A.V., Van Liew, M.W., Kannan, N., Jha, M.K., 2012. Swat: Model Use, Calibration, and Validation. Trans. ASABE 55, 1491–1508.

Arnold, J.G., Youssef, M.A., Yen, H., White, M.J., Sheshukov, A.Y., Sadeghi, A.M., Moriasi, D.N., Steiner, J.L., Amatya, D.M., Skaggs, R.W., Haney, E.B., Jeong, J., Arabi, M., Gowda, P.H., 2015. Hydrological processes and model representation: impact of soft data on calibration. Am. Soc. Agric. Biol. Eng. 58 (6), 1637–1660. https://doi.org/10.13031/trans.58.10726.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modelling and assessment part I: model development. J. Am. Water Resour. Assoc. 34, 73–89. https://doi.org/10.1111/j.1752-1688.1998.tb05961.x.

Beck, H.E., Bruijnzeel, L.A., van Dijk, A.I.J.M., McVicar, T.R., Scatena, F.N., Schellekens, J., 2013. The impact of forest regeneration on streamflow in 12 mesoscale humid tropical catchments. Hydrol. Earth Syst. Sci. 17, 2613–2635. https://doi.org/10.5194/hess-17-2613-2013.

Borchert, R., Robertson, K., Schwartz, M.D., Williams-Linera, G., 2005. Phenology of temperate trees in tropical climates. Int. J. Biometeorol. 50, 57–65. https://doi.org/10.1007/s00484-005-0261-7.

Bremer, L.L., Hamel, P., Ponette González, A.G., Pompeu, P.V., Saad, S.I., Brauman, K.A., 2020. Who are we measuring and modeling for? Supporting multilevel decision-making in watershed management. e2019WR026011 Water Resour. Res. 56. https://doi.org/10.1029/2019WR026011.

Bruijnzeel, L.A., 2004. Hydrological functions of tropical forests: Not seeing the soil for the trees? Agric. Ecosyst. Environ. 104, 185–228. https://doi.org/10.1016/j.agee.2004.01.015.

Campos, C.A., 2010. Response of soil inorganic nitrogen to land use and topographic position in the Cofre de Perote Volcano (Mexico). Environ. Manage. 46, 213. https://doi.org/10.1007/s00267-010-9517-z.

Carnell, R., 2019. Lhs: Latin hypercube samples. [online] Available from: https://CRAN.R-project.org/package=lhs.

Crespo, P., Bücker, A., Feyen, J., Vaché, K.B., Frede, H., Breuer, L., 2012. Preliminary evaluation of the runoff processes in a remote montane cloud forest basin using Mixing Model Analysis and Mean Transit Time. Hydrol. Process. 26, 3896–3910. https://doi.org/10.1002/hyp.8382.

Cukier, R.I., Fortuin, C.M., Shuler, K.E., 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. J. Chem. Phys. 59, 3873. https://doi.org/10.1063/1.1680571.

Francesconi, W., Srinivasan, R., Pérez-Miñana, E., Willcock, S.P., Quintero, M., 2016. Using the Soil and Water Assessment Tool (SWAT) to model ecosystem services: A systematic review. J. Hydrol. 535, 625–636. https://doi.org/10.1016/j.jhydrol.2016.01.034.

García, C.I., Martínez, A.A., Ramírez, A., Cruz, A.N., Rivas A.J., Domínguez, L., 2004. La relación agua-bosque: delimitación de zonas prioritarias para pago de servicios ambientales hidrológicos en la cuenca del río Gavilanes, Coatepec, Veracruz. Last consulted on November 4, 2020 at: http://www2.inecc.gob.mx/publicaciones2/libros/528/relacion.pdf.

Garcia, E., 2004. Modificaciones al sistema de clasificación climática de Köppen, 5th ed. Instituto de Geografía UNAM, Mexico D. F, p. 90.

González-Martínez, T.M., Holwerda, F., 2018. Rainfall and fog interception at the lower and upper altitudinal limits of cloud forest in Veracruz, Mexico. Hydrol. Processes 32 (25), 3717–3728. https://doi.org/10.1002/hyp.13299.

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

Guse, B., Reusser, D.E., Fohrer, N., 2014. How to improve the representation of hydrological processes in SWAT for a lowland catchment – Temporal analysis of parameter sensitivity and model performance. Hydrol. Process. 28, 2651–2670. https://doi.org/10.1002/hyp.9777.

Guse, B., Pfannerstill, M., Strauch, M., Reusser, D.E., Volk, M., Gupta, H.V., Fohrer, N., 2016. On characterizing the temporal dominance patterns of model parameters and processes. Hydrol. Process. 30, 2255–2270. https://doi.org/10.1002/hyp.10764.

Guse, B., Pfannerstill, M., Kiesel, J., Strauch, M., Volk, M., Fohrer, N., 2019. Analysing spatio-temporal process and parameter dynamics in models to characterise contrasting catchments. J. Hydrol. 579, 863–874. https://doi.org/10.1016/j.jhydrol.2018.12.050.

Guse, B., Kiesel, J., Pfannerstill, M., Fohrer, N., 2020. Assessing parameter identifiability for multiple performance criteria to constrain model parameters. Hydrol. Sci. J. 65 (7), 1158–1172. https://doi.org/10.1080/02626667.2020.1734204.

Guswa, A.J., Brauman, K.A., Brown, C., Hamel, P., Keeler, B.L., Sayre, S.S., 2014. Ecosystem services: Challenges and opportunities for hydrologic modeling to support decision making. Water Resour. Res. 50, 4535–4544. https://doi.org/10.1002/2014WR015497.

Guzmán-Huerta, A., 2019. Respuesta hidrológica de una cuenca de bosque tropical húmedo de montaña y su asociación con eventos meteorológicos a diferentes escalas temporales (MSc Thesis). Universidad Nacional Autónoma de México, México City, México, p. 27.

Haas, M.B., Guse, B., Pfannerstill, M., Fohrer, N., 2015. Detection of dominant nitrate processes in ecohydrological modeling with temporal parameter sensitivity analysis. Ecol. Model. 314, 62–72. https://doi.org/10.1016/j.ecolmodel.2015.07.009.

Hamel, P., Riveros-Iregui, D., Ballari, D., Browning, T., Célleri, R., Chandler, D., Chun, K., Destouni, G., Jacobs, S., Jasechko, S., Johnson, M., Krishnaswamy, J., Poca, M., Vieira, P.P., Rocha, H., 2017. Watershed services in the humid tropics: Opportunities from recent advances in ecohydrology. Ecohydrology 11 (3), e1921. https://doi.org/10.1002/eco.1921.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science 342, 850–853. https://www.jstor.org/stable/42620097.

Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture 1(2), 96-99. (doi: 10.13031/2013.26773) @1985 https://elibrary.asabe.org/abstract.asp?aid=26773.

Heidari, A., Mayer, A., Watkins, D., 2019. Hydrologic impacts and trade-offs associated with forest-based bioenergy development practices in a snow-dominated watershed, Wisconsin, USA. J. Hydrol. 574, 421–429. https://doi.org/10.1016/j.jhydrol.2019.04.067.

Her, Y., Frankenberger, J., Chaubey, I., Srinivasan, R., 2015. Threshold Effects in HRU Definition of the Soil and Water Assessment Tool. Transactions of the ASABE 58(2), 367-378. (doi:10.13031/trans.58.10805) @2015 https://elibrary.asabe.org/abstract.asp?aid=45434&t=2&redir=&redirType.

Holwerda, F., Bruijnzeel, L.A., Muñoz-Villers, L.E., Equihua, M., Asbjornsen, H., 2010. Rainfall and cloud water interception in mature and secondary lower montane cloud forests of Central Veracruz, Mexico. J. Hydrol. 384, 84–96. https://doi.org/10.1016/j.jhydrol.2010.01.012.

Holwerda, F., Bruijnzeel, L.A., Barradas, V.L., Cervantes, J., 2013. The water and energy exchange of a shaded coffee plantation in the lower montane cloud forest zone of central Veracruz, Mexico. Agric. For. Meteorol. 173, 1–13. https://doi.org/10.1016/j.agrformet.2012.12.015.

Holwerda, F., Meesters, A.G.C.A., 2019. Soil evaporation in a shaded coffee plantation derived from eddy covariance measurements. J. Geophys. Res. Biogeosci. 124 https://doi.org/10.1029/2018JG004911.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H.H., Gascuel-Odoux, C., 2014. Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. Water Resour. Res. 50, 7445–7469. https://doi.org/10.1002/2014WR015484.

Hu, G.R., Li, X.Y., 2018. Subsurface Flow. In: Li, X., Vereecken, H. (Eds.), Observation and Measurement. Ecohydrology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47871-4_9-1.

Jha, M., Gassman, P.W., Secchi, S., Gu, R., Arnold, J., 2004. Effect of Watershed Subdivision on SWAT Flow, Sediment, and Nutrient Predictions. J. Am. Water Resour. Assoc. 40, 811–825. https://doi.org/10.1111/j.1752-1688.2004.tb04460.x.

Karlsen, R., 2010. Stormflow processes in a mature tropical montane cloud forest catchment, Coatepec, Veracruz, Mexico. MSc. Thesis. VU Univ., Amsterdam, Netherlands, 110 pp.

Kirchner, J.W., 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. Water Resour. Res. 45, 1–35. https://doi.org/10.1029/2008WR006912.

LI-COR, 2019. LAI-2200C Plant Canopy Analyzer Instruction Manual. Copyright©2013–2019, LI-COR, Inc., Publication Number: 984–14112. Last consulted on November 4, 2020 at: https://licor.app.boxenterprise.net/s/fqjn5mlu8c1a7zir5qel.

Liu, S.G., 2001. Evaluation of the Liu model for predicting rainfall interception in forests world-wide. Hydrol. Process. 15 (12), 2341–2360. https://doi.org/10.1002/hyp.264.

López-Hernández, J., 2019. Comportamiento hidrológico a varias escalas temporales de una cuenca periurbana, centro de Veracruz, México (BSc Thesis). Universidad Nacional Autónoma de México, México City, México, p. 45.

López-Ramírez, S.M., Sáenz, L., Mayer, A., Muñoz-Villers, L.E., Asbjornsen, H., Berry, Z. C., Looker, N., Manson, R., Gómez-Aguilar, L.R., 2020. Land use change effects on catchment streamflow response in a humid tropical montane cloud forest region, central Veracruz, Mexico. Hydrol. Process. 1–16 https://doi.org/10.1002/hyp.13800.

Lu, J., Ge, S., Steven, M., Devendra, M.A., 2005. A Comparison of Six Potential Evapotranspiration Methods for Regional Use in the Southeastern United States. J. Am. Water Resour. Assoc. 41 (3), 621–633. https://doi.org/10.1111/j.1752-1688.2005.tb03759.x.

Luo, Y., Arnold, J., Allen, P., Chen, X., 2012. Baseflow simulation using SWAT model in an inland river basin in Tianshan Mountains, Northwest China. Hydrol. Earth Syst. Sci. 16, 1259–1267. https://doi.org/10.5194/hess-16-1259-2012.

Marín-Castro, B.E., Geissert, D., Negrete-Yankelevich, S., Gómez-Tagle, C.A., 2016. Spatial distribution of hydraulic conductivity in soils of secondary tropical montane cloud forests and shade coffee agroecosystems. Elsevier Geoderma 283, 57–67. https://doi.org/10.1016/j.geoderma.2016.08.002.

Mehdi, B., Schulz, K., Ludwig, R., 2018. Evaluating the Importance of Non-Unique Behavioural Parameter Sets on Surface Water Quality Variables under Climate Change Conditions in a Mesoscale Agricultural Watershed. Water Resour. Manage. 32, 619–639. https://doi.org/10.1007/s11269-017-1830-3.

Meins, F.M., 2013. Evaluation of spatial scale alternatives for hydrological modelling of the Lake Naivasha basin, Kenya, Master thesis, University of Twente. Last consulted on November 5, 2020 at: https://www.utwente.nl/en/et/wem/education/msc-thesis/2013/meins.pdf.

Mizutani, K., Yamanoi, K., Ikeda, T., Watanabe, T., 1997. Applicability of the eddy correlation method to measure sensible heat transfer to forest under rainfall conditions. Agric. For. Meteorol. 86, 193-203. PIZ SO168-1923(97)00012-9.

Monteith, J.L., 1965. Evaporation and Environment.19th Symposia of the Society for Experimental Biology, 19. University Press, Cambridge, pp. 205–234. Last consulted on November 5, 2020 at: https://repository.rothamsted.ac.uk/item/8v5v7/evaporation-and-environment.

Moore, D., 2005. Slug injection using salt in solution. Streamline watershed Management Bulletin, vol. 8, No. 2, printed in Canada, ISSN 1705-5989.

Moriasi, D., Arnold, J., Van Liew, M., Binger, R., Harmel, R., Veith, T., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. Am. Soc. Agric. Biol. Eng. 50, 885–900. https://doi.org/10.13031/2013.23153.

Mu, W., Yu, F., Li, C., Xie, Y., Tian, J., Liu, J., Zhao, N., 2015. Effects of rainfall intensity and slope gradient on runoff and soil moisture content on different growing stages of spring maize. Water 7, 2990–3008. https://doi.org/10.3390/w7062990.

Muñoz-Villers, L.E., McDonnell, J.J., 2012. Runoff generation in a steep, tropical montane cloud forest catchment on permeable volcanic substrate. Water Resour. Res. 48, W09528. https://doi.org/10.1029/2011WR011316.

Muñoz-Villers, L.E., McDonnell, J.J., 2013. Land use change effects on runoff generation in a humid tropical montane cloud forest region. Hydrol. Earth Syst. Sci. 17, 3543–3560. https://doi.org/10.5194/hess-17-3543-2013.

Muñoz-Villers, L.E., Holwerda, F., Gómez-Cárdenas, M., Equihua, M., Asbjornsen, H., Bruijnzeel, L.A., Marín-Castro, B.E., Tobón, C., 2012. Water balances of old-growth and regenerating montane cloud forests in central Veracruz, Mexico. J. Hydrol. 462–463, 53–66. https://doi.org/10.1016/j.jhydrol.2011.01.062.

Muñoz-Villers, L.E., Holwerda, F., Alvarado-Barrientos, M.S., Geissert, D., Marín-Castro, B., Gómez-Tagle, A., McDonnell, J.J., Asbjornsen, H., Dawson, T., Bruijnzeel, L. A., 2015. Hydrological effects of cloud forest conversion in central Veracruz. Mexico. Bosque 36 (3), 395–407. https://doi.org/10.4067/S0717-92002015000300007.

Muñoz-Villers, L.E., Geissert, D.R., Holwerda, F., McDonnell, J.J., 2016. Factors influencing stream baseflow transit times in tropical montane watersheds. Hydrol. Earth Syst. Sci. 20, 1621–1635. https://doi.org/10.5194/hess-20-1621-2016.

Myneni, R., Knyazikhin, Y., Park, T., 2015. MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD15A2H.006 (accessed July 30, 2019).

Naeem, S., Ingram, J.C., Varga, A., Agardy, T., Barten, P., Bennett, G., Bloomgarden, E., Bremer, L.L., Burkill, P., Cattau, M., Ching, C., Colby, M., Cook, D.C., Costanza, R., DeClerck, F., Freund, C., Gartner, T., Goldman-Benner, R., Gunderson, J., Jarrett, D., Kinzig, A.P., Kiss, A., Koontz, A., Kumar, P., Lasky, J.R., Masozera, M., Meyers, D., Milano, F., Naughton-Treves, L., Nichols, E., Olander, L., Olmsted, P., Perge, E., Perrings, C., Polasky, S., Potent, J., Prager, C., Quétier, F., Redford, K., Saterson, K., Thoumi, G., Vargas, M.T., Vickerman, S., Weisser, W., Wilkie, D., Wunder, S., 2015. Get the science right when paying for nature's services. Science 347 (6227), 1206–1207. https://doi.org/10.1126/science.aaa1403.

Nainar, A., Tanaka, N., Bidin, K., Annammala, K.V., Ewers, R.M., Reynolds, G., Walsh, R. P.D., 2018. Hydrological dynamics of tropical streams on a gradient of land-use disturbance and recovery: A multi-catchment experiment. J. Hydrol. 566, 581–594. https://doi.org/10.1016/j.jhydrol.2018.09.022.

National Institute of Statistics, Geography and Informatics (INEGI, by its name in Spanish). 2007. Conjunto de datos vectoriales edafológicos. Escala 1:250 000 Serie II. Last consulted on November 3, 2020 at: https://www.inegi.org.mx/temas/edafologia/.

National Institute of Statistics, Geography and Informatics (INEGI, by its name in Spanish). 2012. Continuo de Elevaciones Mexicano 3.0, 15m resolution. Last consulted on November 1, 2020 at: https://www.inegi.org.mx/app/geo2/elevacionesmex/.

National Weather Service (SMN, by its name in Spanish). 2020. Información estadística climatológica. Last consulted on November 2, 2020 at: https://smn.conagua.gob.mx/es/climatologia/informacion-climatologica/informacion-estadistica-climatologica.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009, Texas Water Resources Institute. Last consulted on November 3, 2020 at: https://swat.tamu.edu/media/99192/swat2009-theory.pdf.

Ogden, F.L., Crouch, T.D., Stallard, R.F., Hall, J.S., 2013. Effect of land cover and use on dry season river runoff, runoff efficiency, and peak storm runoff in the seasonal tropics of Central Panama. Water Resour. Res. 49, 8443–8462. https://doi.org/10.1002/2013WR013956.

Paré, L., Gerez, P., 2012. Al filo del agua: Cogestión de la subcuenca del río Pixquiac, Veracruz, 1st ed. Delegación Tlalpan, México, D.F., INE-Semarnat, ISBN 978-607-7908-89-0.

Pfannerstill, M., Guse, B., Fohrer, N., 2014. Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. J. Hydrol. 510, 447–458. https://doi.org/10.1016/j.jhydrol.2013.12.044.

Pfannerstill, M., Guse, B., Reusser, D., Fohrer, N., 2015. Process verification of a hydrological model using a temporal parameter sensitivity analysis. Hydrol. Earth Syst. Sci. 19, 4365–4376. https://doi.org/10.5194/hess-19-4365-2015.

Pianosi, F., Wagener, T., 2015. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environ. Modell. Software 67, 1–11. https://doi.org/10.1016/j.envsoft.2015.01.004.

Pianosi, F., Wagener, T., 2018. Distribution-based sensitivity analysis from a generic input-output sample. Environ. Modell. Software 108, 197–207. https://doi.org/10.1016/j.envsoft.2018.07.019.

Plesca, I., Timbe, E., Exbrayat, J.F., Windhorst, D., Kraft, P., Crespo, P., Vaché, K., Frede, H.G., Breuer, L., 2012. Model intercomparison to explore catchment functioning: results from a remote montane tropical rainforest. Ecol. Model. 239, 3–13. https://doi.org/10.1016/j.ecolmodel.2011.05.005.

Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100, 81–92. https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2.

Qiu, L., Zhen, F., Yin, R., 2012. SWAT-based runoff and sediment simulation in a small watershed, the loessial hilly-gullied region of China: capabilities and challenges. Int. J. Sedim. Res. 27 (2), 226–234. https://doi.org/10.1016/S1001-6279(12)60030-4.

Quintero, M., Wunder, S., Estrada, R.D., 2009. For services rendered? Modeling hydrology and livelihoods in Andean payments for environmental services schemes. For. Ecol. Manage. 258, 1871–1880. https://doi.org/10.1016/j.foreco.2009.04.032.

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/.

Ramírez, B.H., Teuling, A.J., Ganzeveld, L., Hegger, Z., Leemans, R., 2017. Tropical montane cloud forests: Hydrometeorological variability in three neighboring catchments with different forest cover. J. Hydrol. 552, 151–167. https://doi.org/10.1016/j.jhydrol.2017.06.023.

Reusser, D., 2015. Fast: Implementation of the fourier amplitude sensitivity test (fast). Available from: https://CRAN.R-project.org/package=fast.

Ripley B.D., 1998. Based on the cloess package of Cleveland, Grosse and Shyu. Available from: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/loess.html.

Sáenz, L., Mulligan, M., Arjona, F., Gutierrez, T., 2014. The role of cloud forest restoration on energy security. Ecosyst. Serv. 9, 180–190. https://doi.org/10.1016/j.ecoser.2014.06.012.

Salas-Martínez, R., Ibáñez-Castillo, L.A., Arteaga-Ramírez, R., Martínez-Menes, M.R., Fernández-Reynoso, D.S., 2014. Hydrological modelling of mixteco river watershed in the state of Oaxaca, México. Agrociencia 48, 1–15. http://www.colpos.mx/agrocien/Bimestral/2014/ene-feb/art-1.pdf.

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

Sánchez-Galindo, M., Fernández-Reynoso, D.S., Martínez-Menes, M., Rubio-Granados, E., Ríos-Berber, J.D., 2017. Hydrological model of the Sordo River watershed, Oaxaca, México, using SWAT. Water Technol. Sci. (in Spanish) 8 (5), 141–156. https://doi.org/10.24850/j-tyca-2017-05-10.

Schürz, C., Hollosi, B., Matulla, C., Pressl, A., Ertl, T., Schulz, K., Mehdi, B., 2019. A comprehensive sensitivity and uncertainty analysis for discharge and nitrate-nitrogen loads involving multiple discrete model inputs under future changing conditions. Hydrol. Earth Syst. Sci. 23, 1211–1244. https://doi.org/10.5194/hess-23-1211-2019.

Schürz C., 2019. SWATplusR: Running SWAT2012 and SWAT+ Projects in R. doi: 10.5281/zenodo.3373859, R package version 0.2.7. Available from: https://github.com/chrisschuerz/SWATplusR.

Schürz C., 2020. temPAWN: Analyze SWATplusR simulations with temporal PAWN sensitivity analysis, doi: 10.5281/zenodo.4103887, R package version 0.1.0,). Available from: https://github.com/chrisschuerz/temPAWN.

Shrestha, S., Shrestha, M., Shrestha, P.K., 2018. Evaluation of the SWAT model performance for simulating river discharge in the Himalayan and tropical basins of Asia. Hydrol. Res. 49 (3), 846–860. https://doi.org/10.2166/nh.2017.189.

Shuttleworth, W.J., Calder, I.R., 1979. Has the Priestley-Taylor equation any relevance to forest evaporation? J. Appl. Meteorol. 18, 639–646. https://doi.org/10.1175/1520-0450(1979)018%3C0639:HTPTEA%3E2.0.CO;2.

Stewart, J.B., 1977. Evaporation from the wet canopy of a pine forest. Water Resour. Res. 13 (6), 915–921. https://doi.org/10.1029/WR013i006p00915.

Strauch, M., Volk, M., 2013. SWAT plant growth modification for improved modeling of perennial vegetation in the tropics. Ecol. Model. 269, 98–112. https://doi.org/10.1016/j.ecolmodel.2013.08.013.

Tuppad, P., Kannan, N., Srinivasan, R., Rossi, C.G., Arnold, J.G., 2010. Simulation of agricultural management alternatives for watershed protection. Water Resour. Manage. 24 (12), 3115–3144. https://doi.org/10.1007/s11269-010-9598-8.

Van Liew, M.W., Veith, T.L., Bosch, D.D., Arnold, J.G., 2007. Suitability of SWAT for the Conservation Effects Assessment Project: Comparison on USDA Agricultural Research Service Watersheds. J. Hydrol. Eng. 12 (2) https://doi.org/10.1061/(ASCE)1084-0699(2007)12:2(173).

Vizcaíno-Bravo, Q., Williams-Linera, G., Asbjornsen, H., 2020. Biodiversity and carbon storage are correlated along a land use intensity gradient in a tropical montane forest watershed, Mexico. Basic Appl. Ecol. 44, 24–34. https://doi.org/10.1016/j.baae.2019.12.004.

Von Thaden, J., Manson, R.H., Congalton, R.G., López-Barrera, F., Salcone, J., 2019. A regional evaluation of the effectiveness of Mexico's payments for hydrological services. Reg. Environ. Change 19 (6), 1751–1764. https://doi.org/10.1007/s10113-019-01518-3.

Williams-Linera, G., 1999. Leaf Dynamics in a Tropical Cloud Forest: Phenology, Herbivory, and Life Span. Retrieved from Selbyana 20 (1), 98–105. https://journals.flvc.org/selbyana/article/view/120452.

Williams-Linera, G., Toledo-Garibaldi, M., Gallardo-Hernández, C., 2013. How heterogeneous are the cloud forest communities in the mountains of central Veracruz, Mexico? Plant Ecol. 213, 685–701. https://doi.org/10.1007/s11258-013-0199-5.

Winchell, M., Srinivasan, R., Di Luzio, M., Arnold, J., 2013. ArcSWAT interface for SWAT2012: user's guide. Blackland Research and Extension Center, Texas Agrilife Research. Grassland, Soil and Water Research Laboratory, USDA Agricultural Research Service, Texas, USA. Available from: http://swat.tamu.edu/software/arcswat/.

Wright, C., Kagawa-Viviani, A., Gerlein-Safdi, C., Mosquera, G.M., Poca, M., Tseng, H., Chun, K.P., 2018. Advancing ecohydrology in the changing tropics: Perspectives from early career scientists. Ecohydrology 11 (3). https://doi.org/10.1002/eco.1918.

Zadeh, F.K., Nossent, J., Sarrazin, F., Pianosi, F., Griensven, A., Wagener, T., Bauwens, W., 2017. Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model. Environ. Modell. Software 91, 210–222. https://doi.org/10.1016/j.envsoft.2017.02.001.