CAETÊ hydrology analysis report

Comparison between two hydrological modules

Gabriel Banstarck Marandola

2022-08-20

Introduction

During my undergraduate studies, I worked with trait-based vegetation modeling (Bodegom *et al.* (2014)) focused on answering ecosystem-level questions about the Amazon Rainforest future. For this purpose, the research group I work with develops a trait-based dynamic vegetation model called CAETÊ, an acronymn for "CArbon and Ecosystem Trait-based Evaluation model", which is a software built in Python and Fortran that simulates the vegetation in the Pan-Amazon region based on five climatic inputs and several empirical and deterministic equations that can estimate everything that we need to create insights on the vegetation future for different climatic scenarios (Darela Filho (2021); Prado (2017); Rius (2017)).

CAETÊ was initially built upon code from an older vegetation model known as CPTEC-PVM (Oyama and Nobre (2004)). This information is important for the analysis conducted here, as what I proposed to do for this project was a comparison between the CPTEC-PVM hydrological module previously used on CAETÊ - which will be referred as CAETÊ-PVM - and the new module that I developed during my undergraduate research project - CAETÊ-IC. It is also important to notice that the water pools are modeled based on the concept of the bucket model described in Hartmann (2015). Here, I will list some of the main differences between both hydrology modules:

Model	Soil Layers	Water Capacity	Runoff Calculation
CAETÊ-PVM	One (100 cm)	Always 500 mm	Water that exceeds 500 mm capacity
CAETÊ-IC	Two (0-30 cm; 30 - 100 cm)	Derived from soil texture	Different rates for each layer

Table 1: the main differences between the two model versions. The water capacity in CAETE-IC model is estimated based on the equations defined by Saxton and Rawls (2006).

Raw Data Preparation

CAETÊ returns data in two ways: as python dictionaries for each variable saved as `.pkl` files and as `.nc4` files for each variable and for each year of simulation. The analysis conducted here used results from a "historical run", which inputs measured values of precipitation, surface temperature, atmospheric carbon dioxide concentration, relative humidity and photosynthetically active radiation for the time simulated, which was from 01-01-1979 to to 12-31-2016.

The first step for the analysis was to separate only soil water and water runoff data from all the outputs variables provided, as they are the outputs mostly directly affected by the hydrology module. Then, I used the package Climate Data Operators (CDO) to merge data for each year into a single `.nc4` file, which then was used to obtain the mean values of the two variables for each grid cell simulated in a given time (spatial mean) and for each time among all grid cells (temporal mean). These mean values were the inputs that I used in R to generate my figures.

Processed Data Analysis

Using the processed raw data, I generated three different comparisons for each variable (soil water and runoff), resulting in four graphs that show spatial mean data ("spatial_mean") and four maps created with temporal mean files ("time_mean"). The graphs created are (*i*) scatter plots comparisons between CAETÊ-IC and CAETÊ-PVM values for the same variable and (*ii*) line charts containing variable rolling mean values along time for each model. The maps created display the mean variable value for each grid cell for the whole simulation.

Results

Soil Water

The soil water comparison in general revealed the overestimation of CAETE-PVM values in relation to CAETE-IC. I would say this is an overestimation from PVM because of previous analysis that I have conducted with satellite data (Rodell *et al.* (2004)), which were more close to CAETE-IC values. In the scatter plot (Figure 1), all points fell above the y = x line, indicating the overestimation of values. The line plot (Figure 2) gives us the same information, but in relation to time: the seasonality is the same, but both the total values and the variation between seasons are much bigger on CAETE-PVM data. The map (Figure 3) provides the analysis a new dimension: the spatial differences. It is evident that, for the majority of the regions, the mean values from CPTEC-PVM are higher, except for the Central Amazon region and for other minor patches of grids closer to the border of the simulated region. It also gives us insight on the spatial variability difference between two models, which is most certainly caused by soil texture differences incorporated in CAETE-IC, providing a more patchy environment for the simulation.

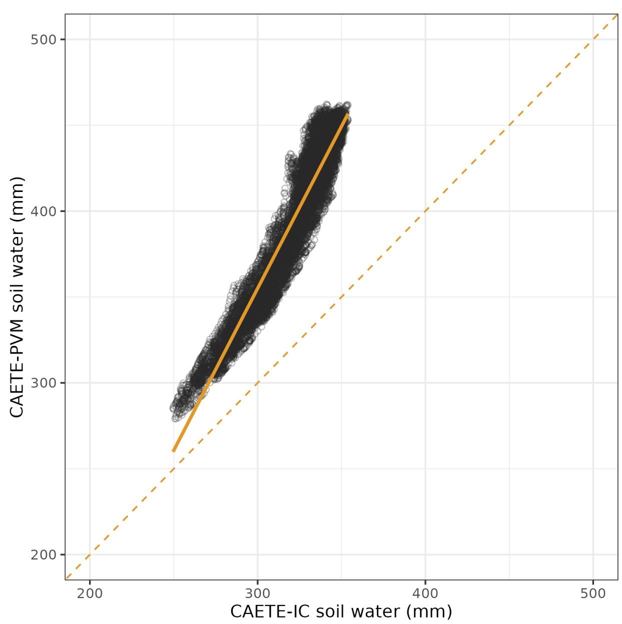

Figure 1: scatter plot with soil water values in milimeters for the spatial means of CAETÊ-IC on the x axis and CAETÊ-PVM on the y axis. The orange solid line is a linear regression between the data and the dashed line is the identity function y = x.

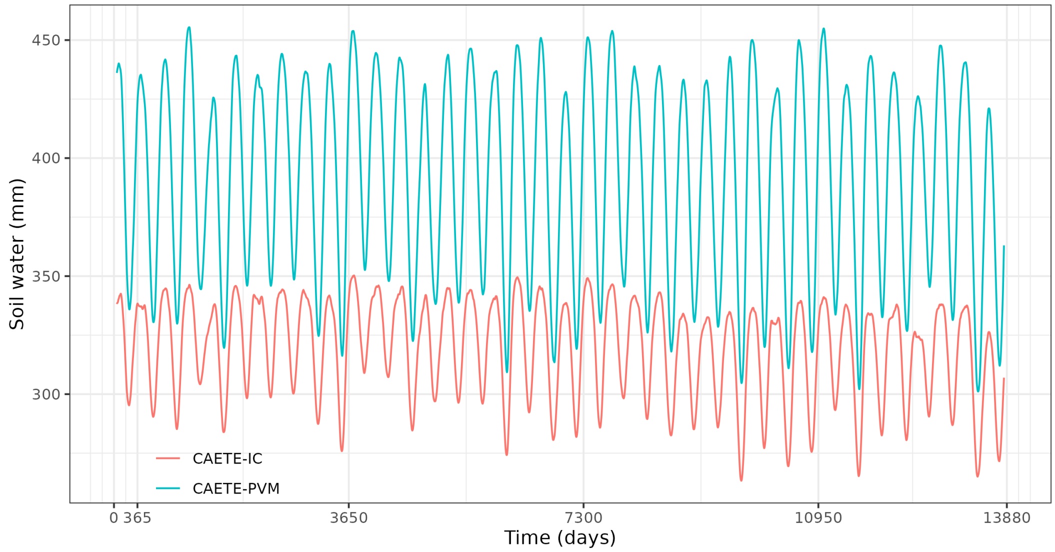

Figure 2: rolling mean of 90 days for soil water mean daily values of each module.

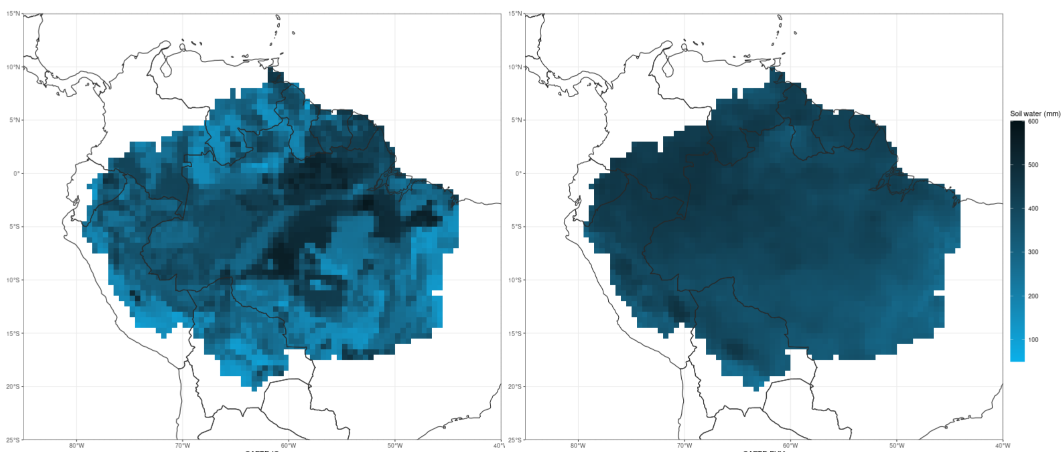

Figure 3: maps of the mean water runoff values for each grid cell during the whole simulation period of the two model versions compared.

Water Runoff

For this variable, the comparison is a little more complex: the scatter plot shown in Figure 4 shows that, for intermediate values, both modules seem to behave similarly, but CAETÊ-IC presents higher values of runoff that do not appear on CAETÊ-PVM. For this reason, I chose a logarhitmic scale to the comparison, so the variation along the whole scale could be better visualized. The line plot in Figure 5 shows that seasonality is conserved between both modules, although CAETÊ-IC have shown higher values of runoff, which makes sense due to the variable soil maximum water capacity implemented, variation that resulted in an overall reduced water capacity for the majority of grid cells. For lower capacities, it is expected to observe higher runoff values, as the amount of precipitation is the same. The map in Figure 6 shows that spatial patterns are kept similar for runoff values, but with a general slight increase in values. Some new patches are present in the newer module that were not seen in CAETÊ-PVM, although not as much as the difference observed for soil water in Figure 3.

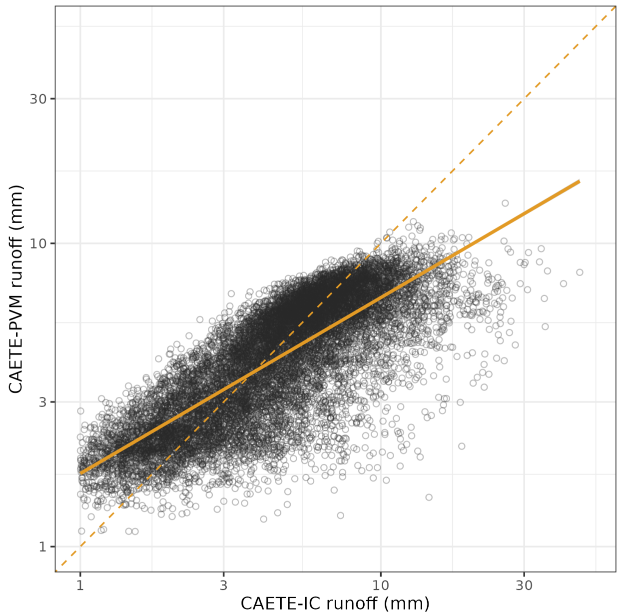

Figure 4: scatter plot with water runoff values in milimeters for the spatial means of CAETÊ-IC on the x axis and CAETÊ-PVM on the y axis. The orange solid line is a linear regression between the data and the dashed line is the identity function y = x.

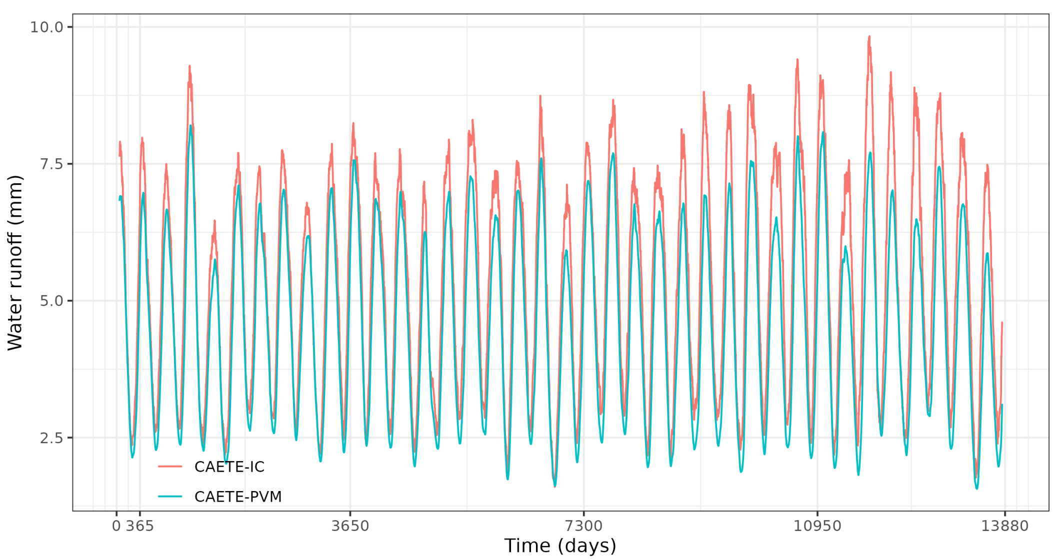

Figure 5: rolling mean of 90 days for water runoff mean daily values of each model version.

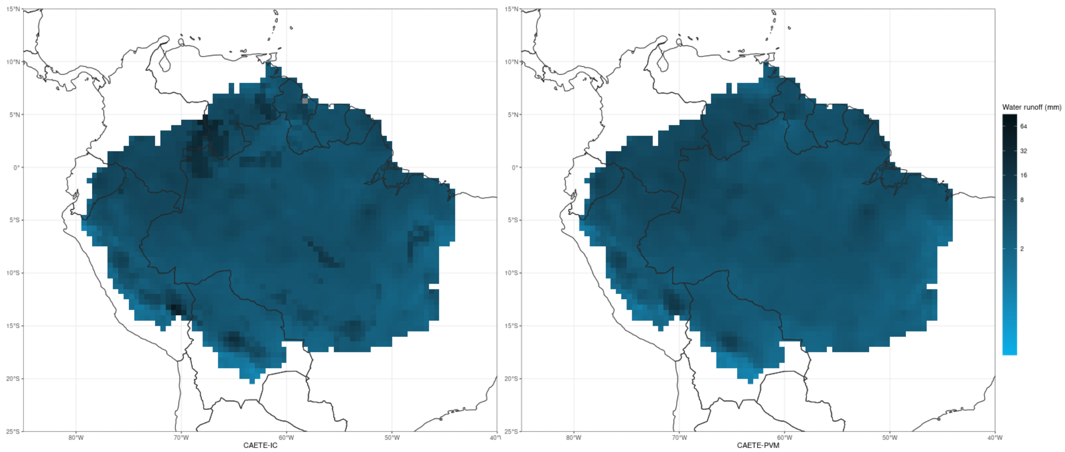

Figure 6: maps of the mean water runoff values for each grid cell during the whole simulation period of the two model versions compared.

Discussion

Based on the results of the comparison, it is evident that the two modules behave differently in terms of soil water and water runoff dynamics. Without a proper benchmarking method and sensibility test data, it is not possible to precisely determine if the complexity added to CAETÊ-IC in comparison to the previous module is really beneficial for the purpose of CAETÊ and if it translates to real observations, which are also a problem to obtain for such a large area of study. However, it is possible to compare direct aspects between the two versions from the analysis here conducted. Firstly, it is evident that CAETÊ-IC generates grid cell patches with higher variability in values, fact that can be related to the different soil water maximum capacity for each space. It is also clear that CAETÊ-IC outputs higher spikes of water runoff that do not occur in CAETÊ-PVM, as it is observable in Figure 4 and 5. The seasonality between both models is conserved, which is a positive and expected result as both versions only differ in one module, that hardly would break the entire seasonality of the model dynamics. At last, a significant and promising result of this analysis is that CAETÊ-IC presented lower mean soil water values and less abrupt spikes between dry season and wet season levels compared to CAETÊ-PVM. This is a promising result as there is a high change that the previous module was overestimating the amount of soil water due to its simplicity and to previous study of the GLDAS dataset (Rodell *et al.* (2004)). The reduction in seasonal changes is also more coherent with GLDAS values and with an equatorial rainforest climate.

Conclusion

The comparison presented here incites that the developments realized to create a better hydrology model for CAETÊ were successful in reaching this goal. Of course, as said previously, a more robust analysis involving variables of the model indirectly related to the hydrology are necessary to achieve a proper benchmarking study, but it was already possible to obtain initial insights on the differences between the two versions with the data studied here.

References

Bodegom, P. M. van, J. C. Douma, and L. M. Verheijen. 2014. A fully traits-based approach to modeling global vegetation distribution. Proceedings of the National Academy of Sciences. 111: 13733–13738.

Darela Filho, J. P. 2021. Modeling the impacts of increased concentration of CO2 and nutrients over the functional composition of the amazon forest. PhD thesis, Universidade Estadual Paulista "Jũlio de Mesquita Filho", Rio Claro. pp.

Hartmann, D. 2015. Global Physical Climatology, 2nd ed. Elsevier Science.

Oyama, M. D., and C. A. Nobre. 2004. A simple potential vegetation model for coupling with the simple bioesphere model (SIB). Revista Brasileira de Meteorologia. 19: 203–216.

Prado, H. A. do. 2017. Produtividade primária, uso da água e diversidade funcional a partir de um novo modelo vegetacional: Primeira aplicação sob CO2 elevado. Master's thesis, Universidade Estadual Paulista "Jũlio de Mesquita Filho", Rio Claro. pp.

Rius, B. 2017. Modelagem da alocação, tempo de residência do carbono e diversidade funcional em florestas tropicais: Uma abordagem "trait-based" e os efeitos do CO2 elevado. Master's thesis, Universidade Estadual Paulista "Jũlio de Mesquita Filho", Rio Claro. pp.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll. 2004. The global land data assimilation system. Bulletin of the American Meteorological Society. 85: 381–394.

Saxton, K. E., and W. J. Rawls. 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Science Society of America Journal. 70: 1569–1578.