Title
Seasonal and Interannual Patterns and Controls of Hydrological Fluxes in an Amazon Floodplain Lake With a Surface-Subsurface Process Model

Journal
Water Resources Research, 55(4)

ISSN
0043-1397

Authors
Ji, Xinye
Lesack, Lance FW
Melack, John M
et al.

Publication Date
2019-04-01

DOI
10.1029/2018wr023897

Peer reviewed

Seasonal and Interannual Patterns and Controls of Hydrological Fluxes in an Amazon Floodplain Lake With a Surface-Subsurface Process Model

Xinye Ji[1], Lance F. W. Lesack[2], John M. Melack[3], Shilong Wang[1], William J. Riley[4], and Chaopeng Shen[1]

[1] Civil and Environmental Engineering, Pennsylvania State University, University Park, PA, USA, [2] Geography and Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, [3] Earth Research Institute, University of California, Santa Barbara, CA, USA, [4] Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Correspondence to: C. Shen, cshen@engr.psu.edu

Abstract

Floodplain lakes represent important aquatic ecosystems, and field-based estimates of their water budgets are difficult to obtain, especially over multiple years. We examine the hydrological fluxes for an Amazon floodplain lake connected to the Solimões River using a process-based hydrologic model. Water exchanges between the river and lake agree well with field estimates, including the timing of different hydrological phases. However, beyond available field data, modeling results show that the seven simulated years all differed from each other. These interannual differences were caused by the interplay between phases when water levels were rising with river-water flowing into the lake (RWRI), versus rising with lake-water flowing out to the river (RWLO). This exchange determines the river-water content in the lake (C_L). Maximum C_L occurred before river levels peaked because local catchment contributions can be sufficient to push lake-water out to the river, even as river levels rise. Numerical experiments show that the seasonal distribution of local rainfall, local catchment size, and interannual variability in both climate and river stage can contribute to differing dynamics of C_L in a floodplain lake. Their impacts vary among phases: river-rise dominates the RWRI, whereas local hydrological processes dominate the RWLO and receding-water phases. Intermediate-to-long-term rainfall accumulation controls C_L during the RWLO phase, whereas annual precipitation accumulation is important for C_L during low water. Our model generalizes beyond limited available field studies and offers potential to better understand floodplain lakes in other areas and how regional versus local changes in climate may affect their hydrological dynamics.

1 Introduction

The hydrologic behavior of lakes on large river floodplains is particularly complex and is not well understood, despite the ecological importance (Thorp et al., 2006) and widespread occurrence of floodplain lakes from the tropics (Melack & Coe, 2012) to the Arctic (Lesack & Marsh, 2010). Water balances in floodplain lakes are driven by interactions, temporal variations,

and timing of variations among precipitation and runoff in the upstream main stem-river basin, local precipitation, local runoff, and effects of antecedent conditions in prior years. In floodplains along rivers in semiarid and arctic regions, rising river levels drive a strong pulse of river-water into the floodplain. This pulse can serve as a primary source of water renewal for perennial water bodies and intermittent ones. By contrast, in floodplain lakes of the humid tropics, such as the Amazon basin, the climate has superimposed a strong effect of local catchment hydrology in combination with the basin-driven pulse of river-water into the lakes. Such a combination may generate more complex interannual variability than would otherwise be present. Mertes (1997) introduced the concept of the perirheic zone to explicitly recognize that the relative roles of regional versus local water differ among floodplains. Connectivity dynamics of aquatic communities has also been of considerable interest in riverine ecosystems (Amoros & Bornette, 2002; Stanford & Ward, 1993; Thorp et al., 2006) but has mostly been investigated in temperate systems. The roles of regional versus local water, river-to-lake connectivity, and their temporal variabilities in floodplain systems all contribute to aquatic habitat diversity and facilitate important biogeochemical functions of these systems. However, the complex temporal dynamics and interactions among these factors have not been well documented because fieldwork is very labor-intensive and realistic models have been difficult to implement.

Floodplains in the Amazon basin cover up to about 20% of the lowland area and have large seasonal fluctuations in stage and inundated area (Hess et al., 2015; Paiva et al., 2013). Floodplain lakes are an important component of the inundated areas, and as water levels vary, the proportion of open water relative to other aquatic habitats changes. They exchange water (Alsdorf et al., 2010; Rudorff et al., 2014a, 2014b), nutrients and other solutes (Melack et al., 2009; Melack & Forsberg, 2001), and sediments (Dunne et al., 1998; Rudorff et al., 2018) with the rivers. These lakes are a source of methane to the troposphere (Melack et al., 2004), release a large amount of carbon dioxide derived from respiration of organic matter from uplands and floodplains (Forsberg et al., 2017; Melack, 2016), and support an important inland fishery (Bayley & Petrere, 1989).

Though thousands of floodplain lakes are spread across the central Amazon basin (Sippel et al., 1991), hydrological studies of these lakes are rare, and only three lakes have had full suites of inputs and losses of water measured or modeled. These include Lake Janauacá (Bonnet et al., 2017), and a large, shallow complex of lakes (Curuai floodplain) in the eastern Amazon (Bonnet et al., 2008; Rudorff et al., 2014b). The most complete measurements of the hydrological balance of an Amazon floodplain lake have been made in Lake Calado (Lesack, 1993; Lesack & Melack, 1991, 1995). The lake volume and proportion of water in Lake Calado contributed by its water inputs vary over the hydrological cycle in a complex way. In the year with available observations, river water enters the lake at the start of rising water in the

Solimões River, by the middle of the rising river period, lake water flows from the lake into the Solimões River, while river levels continued to rise (Lesack & Melack, 1995; hereafter referred to as LM95). However, it is uncertain whether such patterns generalize to other years.

Lake Calado is representative of ria-type lakes (inundated former stream valleys; Gourou, 1949) on the Amazon floodplain, where such lakes represent ~1,000 of the 8,000 lakes present in this system and about half of its lake surface area (Sippel et al., 1991). Besides local climate and interannual variability, an important factor affecting the comparability of Lake Calado with other lakes on the Amazon floodplain is the ratio of its local drainage basin area to lake area (BA:LA), which is an important control on the magnitude of local water supplied to the lake relative to main-stem river water. This factor was qualitatively corroborated by Forsberg et al. (1988) based on alkalinity levels (high alkalinity equated to high river-water content in the lake-water) from the occasional sampling of 51 lakes. By the end of low water, lakes with BA:LA < 20 appeared to have a mixture of river and local water, whereas lakes with BA:LA > 20 contained primarily local water. However, subsequent water balance investigations in Lakes Calado and Curuai showed that Forsberg et al.'s inferred BA:LA threshold was not fully consistent with the observations, and in particular, that a simple threshold ratio did not account for the dynamic evolution of lake water composition (Bonnet et al., 2008, LM95). We expect that this type of hydrological behavior occurs widely both within the Amazon floodplain, and in high-rainfall river systems elsewhere, but they have not been well documented. Due to the complex spatiotemporal variability, we could not observe lakes with different BA:LA at different representative times and could not assess how this ratio affect lake water balances across the Amazon, together with other physical factors.

The field-based investigations described above are limited because of the labor-intensive nature of the work and short periods of observations, which has constrained our understanding of this complex system. Particular hydrological processes have been challenging to model for Amazon floodplain lakes in a way that would facilitate an understanding of system dynamics beyond a single year or so of empirical data. These processes include (i) dynamic water exchange between the river and lake, (ii) the backflow inundation of the lake and conditions when this occurs, (iii) the switching of lake inflow and outflow from channelized connection to over-bank flow, (iv) seepage inflow during rising-water and outflow during falling-water from the changing edge of the lake, and (v) variation in the extent of decoupling between regionally driven river levels and local rainfall driving upland hydrology into the lake.

In this paper, we adapted a process-based hydrological model with surface and subsurface components configured to elucidate the dynamical patterns and physical controls of river water content in an Amazon floodplain lake, Lake Calado. With this model, we first evaluated (a) its behavior for periods

where field data are available. Then we employed this model to (b) disentangle the impacts of different hydrological processes including rainfall distributions and main-stem river stage on the river-water content within the lake (i.e., the fraction of water mass derived from the Solimões River), a key hydrological variable in river floodplain lakes that determines their biogeochemical character. This analysis was conducted for a range of hydrological conditions outside the period with direct measurements to obtain a broader window of temporal variability. Lastly, we used the model to (c) experiment with differing local catchment sizes, along with differing combinations of climates and stage patterns to infer how lakes in differing floodplain settings than Calado may differ in their river-water content.

2 Data and Methods

2.1 Study Site

The lowland Amazon basin is warm and humid. Regional differences in amounts and timing of rainfall occur (Marengo, 2004). Seasonality of rainfall is related to conditions in the tropical Atlantic and movement of the intertropical convergence zone leading to seasonal precipitation maxima in the northern Amazon from March to May and in the southern Amazon from January to April. Further information about the climate system in the Amazon basin is summarized in Marengo et al. (2009), Melack and Coe (2012), and Nobre et al. (2009).

Upstream from its confluence with the Negro River in the central Brazilian Amazon, the Amazon River is called the Solimões River (simply referred to as *the river* in this paper, as compared to *upland streams* for streams originating in a local catchment). Lake Calado lies on a Solimões River floodplain 100 km above the Solimões-Negro confluence, near the city of Manacapuru (Figure 1). A 100- to 200-m-wide natural levee separates the lake and the Solimões. The river and lake are connected year-round by a channel on the southeastern edge of the lake. The area included in our analysis, the upland catchment, the lake, and the levee, is 73 km^2. We divided the lake into three zones: outlet, west, and upstream (Figure 2a), for which we calculated hydrological terms. More detailed descriptions of Lake Calado and the upland catchment can be found in Lesack (1993) and LM95.

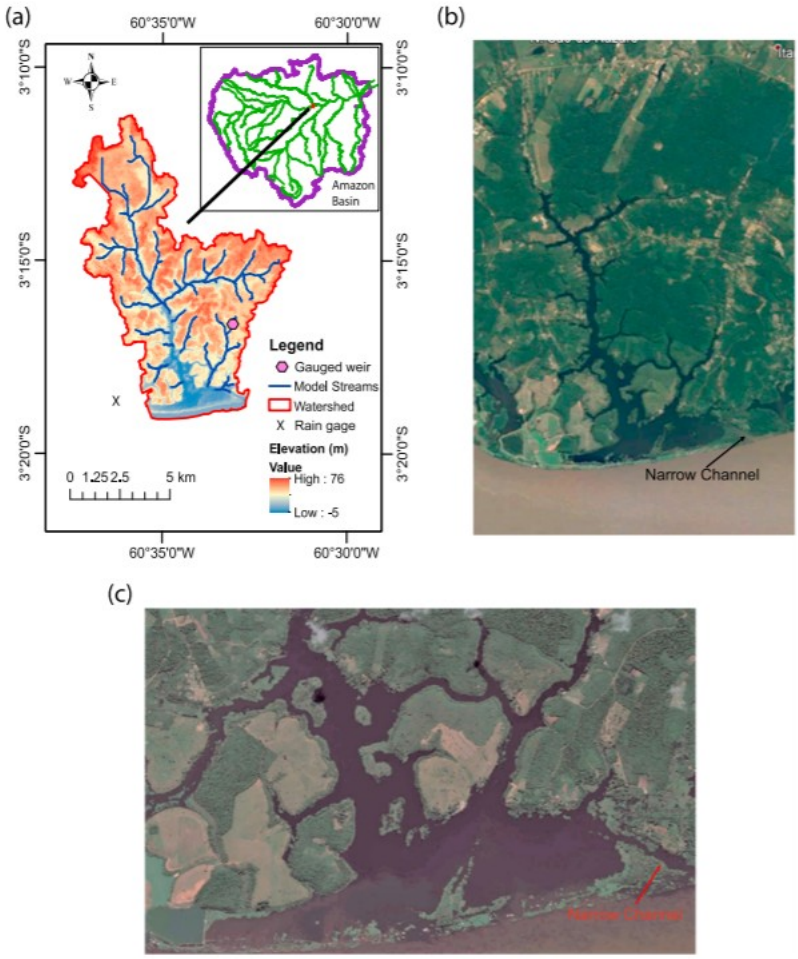

Figure 1. (a) Map of Lake Calado and upland basin. The streamflow gauge is annotated as *gauged weir* in the figure; (b) Google Earth satellite image for the basin taken on 8 June 2016, when there was only one connection between the lake and the Solimões River at the eastern end of the lake. (c) Google Earth satellite image taken on 14 June 2015, when the levee was inundated by high water. Image brightness has been increased.

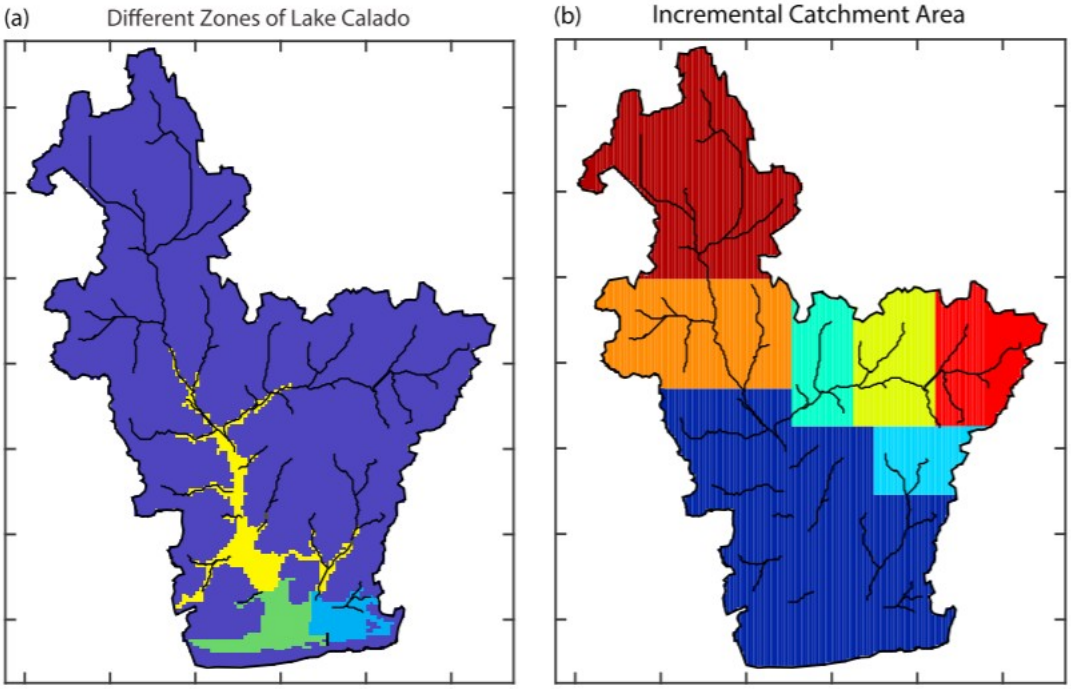

Figure 2. (a) Different zones in Lake Calado: Outlet zone (light blue), West zone (green) and Upstream zone (yellow). (b) Catchment areas incrementally removed in the order of brown, red, orange, yellow, green, and light blue. The corresponding basin area to lake area ratios in that order are 8.3, 7.4, 5.7,4.6, 4, and 3.4.

LM95 measured runoff using a weir-based gauge for a subbasin and extrapolated the runoff rates to the entire Calado basin. They estimated evapotranspiration using the Penman-Monteith equation with locally measured weather data. The lake volume was estimated using the measured water stage and the depth-sounding-based bathymetry data. The exchange between the Solimões River and the lake was estimated as the residual of the water balance equation. Therefore, this estimate includes uncertainty, especially related to the spatial extrapolation of seepage and runoff and evaporation.

2.2 Model Description

The Process-based Adaptive Watershed Simulator coupled with the Community Land Model (PAWS+CLM) represents hydrologic, energy, elemental cycles, and vegetation dynamics based on physically based equations (Shen et al., 2013; Shen & Phanikumar, 2010). PAWS simulates above and below ground water flows, along with infiltration, while CLM simulates the land surface energy-water balance that partitions the amount of water available from rainfall to drive the PAWS hydrology. PAWS models two-dimensional overland flow using dynamic or diffusive wave equations, a one-dimensional river flow cascade governed by the diffusive wave equation, and 3-D variably saturated subsurface flow and their exchanges (Ji et al., 2015; Ji & Shen, 2018; Shen et al., 2016). The subsurface flow in the model is governed by Richards equation, but PAWS+CLM simplified this equation by

assuming lateral flow does not occur in the unsaturated zone. The model solves a series of 1-D Richards equation coupled to quasi-3-D saturated flow in a physically consistent manner. This simplification allows significant savings in computational demand. PAWS+CLM has been applied to midwestern U.S. basins with thick unconsolidated soil or regoliths (Niu et al., 2017; Pau et al., 2016; Riley & Shen, 2014; Shen et al., 2014, 2016), as occur in the Amazon basin. Niu et al. (2017) configured the model for an Amazonian basin, approximately 80 km from the present one, to study the control of evapotranspiration by annual precipitation, and streamflow generation mechanisms. The model had good performance when compared to gauge-measured streamflow and satellite-based estimates of evapotranspiration and terrestrial water storage. They reported a Nash-Sutcliffe model efficient coefficient ranging from 0.4 to 0.80 for daily streamflow compared to observed streamflow data. Their results highlighted the dominance of groundwater contributions in streamflow generation.

2.2.1 Model Adaptations

The gridded model simulates the following processes associated with floodplain lakes: Lake inundation is driven by rising river levels, either as (a) flow from the river into the lake due to free surface gradient or from (b) local runoff contributions backing up against rising lake levels that (c) may (or may not) cause lake levels to rise faster than the river and drive water out of the lake prior to river levels switching to the receding phase. Seepage outflow (d) from the lake to groundwater is driven by rising lake levels, then (e) from groundwater to the lake, with infiltrated rainfall recharging groundwater, as lake-levels fall.

As a summary, the overland flow in our model (with an 80-m horizontal grid size) is governed by the dynamic wave equation (or St. Venant equation) with time-dependent boundary conditions in the south to represent the stage of the Solimões River. The St. Venant equation resolves water inertia, pressure gradients, and friction and thus automatically addresses processes (a), (b), and (c) above. Water contributions to the lake in the model come from direct overland flow, inputs from the local stream network, direct precipitation, and groundwater exchange. In reality, groundwater-lake exchange (d and e described above) could occur laterally as seepage through the banks. In the model, overland flow exchanges with groundwater as described by Darcy's law-based leakance concept (Gunduz & Aral, 2005). This exchange was conceptualized to occur through the bottom, but numerically, it is simply driven by the head difference between the regional groundwater system and the overland flow free surface. Our formulation should capture this seepage as long as groundwater flow and lake levels are modeled adequately.

Previously, PAWS solved a two-dimensional diffusive wave equation that treated lake inundation as equivalent to overland flow with an explicit Runge Kutta finite volume scheme. The diffusive wave equation allows flow

direction to be determined by the flow stage, but it does not resolve the effects of inertia. While this assumption has been adequate for describing upland water convergence, the solver encounters numerical instability in lakes. In addition, the inertia term becomes more important under low slope, slow-moving conditions. Another constraint is that the lake flow is integrated with the hydrologic model so that the time step needs to be large enough to enable multiyear hydrologic simulations. Hence, we adopted a semiimplicit, semiLagrangian (SISL) scheme (Casulli, 1990; Martin & Gorelick, 2005a, 2005b) to solve the full dynamic wave equation (also called the St. Venant equation or the shallow water equation):

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -g\frac{\partial \eta}{\partial x} + \varepsilon\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) - g\frac{\sqrt{u^2 + v^2}}{C_z^2}u$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -g\frac{\partial \eta}{\partial y} + \varepsilon\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) - g\frac{\sqrt{u^2 + v^2}}{C_z^2}v$$

$$\frac{\partial \eta}{\partial t} + \frac{\partial (hu)}{\partial x} + \frac{\partial (hv)}{\partial y} = 0.$$

where h is the water depth, η is the free surface elevation ($\eta = E + h$ where E is elevation), u and v are the x- and y-directional depth-averaged velocities respectively, g is the gravitational constant, ε is the horizontal eddy viscosity coefficient, and C_z is the Chezy roughness coefficient. C_z is related to Manning's coefficient, n, via $C_z = \frac{h^{1/6}}{n}$. The Manning's coefficient n for the overland flow was determined from a landuse-based Manning's coefficient from the TR-55 manual (Mays, 2010; Natural Resources Conservation Service, 1986). The composite n is an areal average of the n from different land uses. For the purpose of this study, we used a viscosity of 1×10^{-3} m^2/s and have run sensitivity tests for the impact of this value. The choice of the shallow water equation is justified as the horizontal characteristic length scale (~4 to 6 km north-south and ~1 to 3 km east-west during low or high water, respectively), is much larger than the vertical scale (2 to 11 m at low or high water). This choice reduces computational demand and facilitates long-term integration and is consistent with nearshore modeling in Lake Michigan (Liu et al., 2006). We ignored the Coriolis and wind terms. The inertia terms often cause numerical instabilities for this hyperbolic equation, and upwinding schemes are preferred. The SISL scheme solves the inertia terms using a semi-Lagrangian approach, which traces back and retrieves velocities from where gridpoints were in the last time step. It solves the pressure gradient term in an implicit fashion. Both semi-Lagrangian and semi-implicit treatments allow larger time steps (~5 min for 80-m grid spacing for the surface flow modules) in the hydrologic model. While adopting some of the implementations of Martin and Gorelick (2005a), we changed the tracing scheme from their departure point tracking scheme (Martin & Gorelick, 2005b) to a more traditional 4-stage Runge Kutta scheme (Zheng & Bennett, 2002) as the latter was found to be more stable.

The SISL scheme described above was used to model both the local catchment overland flow and the floodplain lake. In the local catchment, upland streams are represented as a cascaded network of one-dimensional objects, which were delineated from the digital elevation model (DEM). Overland flow first converges to the valleys and then contributes to the streams via land gridcells intersecting the streams. However, when the stage in the streams is higher than land elevation, it will flood the land gridcells. Baseflow is modeled as a head-dependent first-order exchange between the groundwater level and the stream stage (Gunduz & Aral, 2005; Shen et al., 2016; Shen & Phanikumar, 2010) and is calculated after each hourly time step. To describe such two-way interactions, our preprocessing utility records the mapping relations between stream grid cells and land grid cells (Shen et al., 2014).

At the mouths of the streams in the lake, the free surface elevation in the interacting overland flow grid cell (the lake level) is passed to the river flow module as a downstream boundary condition. Therefore, if the lake level rises, water can backfill into streams and propagate upstream. Meanwhile, evaporation from the lake is modeled as occurring at the potential rate estimated by the Penman-Monteith equation.

2.2.2 Modeling River-Water Content in the Lake

River-water content in the lake is an important quantity as a measure of how much water entered the lake and transported materials contained in that water. In LM95, the lake was treated as a fully mixed system where both river-water and local-water from rainfall-runoff became lake-water upon entering the lake. Lake-surface evaporation represented a loss of the mixed lake-water but would not alter the relative contribution or mixture of water types (and their materials content) from the various sources within the lake. In our present model, we represent the river-water content in the lake as a spatially distributed quantity. Prior work has shown Lake Calado to be laterally mixed to varying degrees, with some gradients (Lesack, 1988). Thus, the system lies somewhere between fully mixed, and our model allows for the modeling of lateral lake gradients.

To represent the amount of river water that enters the lake, we model the vertically integrated concentration of an inert tracer, C. The concentration C is assumed to be 1 at the river boundary. Therefore, with this assumption, C at any location in the lake could be interpreted as the fraction of lake water that comes from the Solimões River, which we define herein as the *river-water content*. C is governed by the following equation:

$$\frac{\partial(h\mathrm{C})}{\partial t} = -\frac{\partial(h\mathrm{C}u)}{\partial x} - \frac{\partial(h\mathrm{C}v)}{\partial y} + k\mathrm{C}$$

where C is the river-water content and k represents a decay rate ([1/T]). For a conservative tracer, $k = 0$. We also calculated the average river-water

content over the whole lake: $C_L = \int_A hCdA / \int_A hdA$, where *A* is the area of the lake.

2.3 Data Sources and Model Configuration

The 90-m-resolution National Aeronautics and Space Administration (NASA) Shuttle Radar Topography Mission (SRTM) was used as the DEM to delineate the upland watershed and generate average cell elevation, slope, and stream routing properties. Since Carabajal and Harding (2006) reported SRTM elevations are located, on average, at 40% of canopy height below the canopy top in vegetated areas, we adjusted our DEM using the global 1-km-resolution canopy height data (Simard et al., 2011). The Calado basin has mature and second growth Amazonian forest. The secondary forest was delineated manually, and the height of the forest was estimated by calculating height differences across the boundary of mature and secondary forests. The adjustment lowered the elevation by ~3.8 m, on average, and in some locations by as much as 12.8 m. We digitized the bathymetry of the lake based on depth soundings. The southeast channel of Lake Calado is not represented in SRTM data; we manually altered the elevation of grid cells to create the channel. PAWS represents the upland streams as a network of 1-D objects. Upland stream networks were derived by watershed delineation and flow accumulation software using the DEM. The streams that overlapped with Lake Calado were removed from the model. A 80-m × 80-m horizontal grid was used for spatial discretization. Twenty vertical layers were applied for the vadose zone, and the vertical discretization adaptively depends on local soil thickness. The land surface and groundwater processes ran on 1-hr time step, while the overland flow and stream routing used a maximum time step of 5 min. Based on the Courant number, the flow modules can adaptively choose to use smaller time steps.

Climate forcing data, except rainfall, were extracted from the default Climate Research Unit – National Centers for Environmental Prediction data set (CRUNCEP; Viovy, 2018) in 6-hourly and a 0.5° × 0.5° spatial resolution from 1901 to 2010, which includes temperature, short and long wave solar radiation, humidity, air pressure, and wind speed. The CRUNCEP is a combination of the Climatic Research Unit's monthly climatology from 1901 to 2009, and the NCEP reanalysis from 1948 to 2013. However, it is known that CRUNCEP's precipitation tends to underestimate large precipitation events (Niu et al., 2017). Hence, we used precipitation data from the nearby town of Manacapuru (Station code: 00360001; ~8 km from L. Calado) collected by Agência Nacional de Águas (ANA, Brazilian National Water Agency). ANA's measurements indicated that annual precipitation from 1980 to 1984 was ~1997 mm, close to the CRUNCEP value of ~1,926 mm, but the standard deviation of monthly rainfall in every year was higher in the ANA data than CRUNCEP's, suggesting that the CRUNCEP is not able to capture large-magnitude storms. Previously, in a nearby basin, using precipitation from the Tropical Rainfall Measurement Mission (TRMM) resulted in better

and different streamflow simulations than CRUNCEP (Niu et al., 2017). The local ANA data should be more realistic than TRMM.

Land use and land cover data were obtained from default CLM input data that incorporate the historical record and remote sensing land surface data (Oleson et al., 2013). The coarse default CLM Plant Functional Types were mapped to the simulation grid. Apart from the water surface of the lake, the primary land cover type is the tropical broadleaf deciduous tree, with some tropical grassland. The 0.083°-resolution CLM soil texture and organic matter input data (Bonan et al., 2002) were processed by CalcPTF code (Guber et al., 2006, 2009) to calculate the initial van-Genuchten soil property parameters. Due to data constraints, this pedotransfer function utilizes the sand, silt, and clay percentages to infer water retention and hydraulic conductivity parameters.

Both soil and groundwater parameters were adjusted via *global multipliers*; that is, the parameters in the domain are adjusted by a uniform factor. These global multipliers were obtained previously by calibrating simulated streamflow to observations in a basin nearby, as reported in Niu et al. (2017). The same parameter set that was calibrated and validated to produce a good match with streamflow and evapotranspiration observations in the nearby basin is used in this paper.

The Solimões River stage data came from ANA's record (station code: 14100000). The gauge station was in Manacapuru, approximately 8 km upstream from Lake Calado; river stage measured on the Solimões River in Manacapuru is highly correlated with the stage measured at Lake Calado ($R^2 > 0.999$; from LM95). The free surface elevation gradient in the Solimões River is very small, around 1×10^{-5} in this part of the domain. We set river stage data for the grid cells that represented (or belonged to) the Solimões River as a boundary condition. The model records the inflow flux from the cells representing the river boundary condition to the lake. The lake has been divided into several zones for the purpose of reporting spatially distributed river-water content (Figure 2a).

2.4 Experiments

To disentangle the separate impacts of seasonal rainfall distribution and river stage on C_L, in particular in 1984 and 1987, we conducted several different hypothetical simulations, where we (a) applied the 1984 climate forcing to other years, to demonstrate relative influence of river stage or the climate forcing, (b) scaled January–May rainfall in different years to have the same total amount of rainfall as in 1984, (c) increased rainfall only in April and May to 300 mm/month, and (d) migrated the river stage time series of 1981 to other years; (e) migrated the slow seasonal river rise in the fourth quarter of 1986 (Q4, Oct-Dec) to Q4 of 1981; and (f) migrated the climate forcing from 1986 to other years. These simulations examine various hypotheses regarding C_L behaviors to be described later. For example, comparing (b) to (c), we can examine whether the system has a longer term memory for the

impacts of rainfall. The differences between the default case and (e) or (f) would reveal the separate impacts of river stage and weather forcing on C_L.

To study how C_L is influenced by the upland catchment area, we conducted perturbation analyses where we incrementally removed parts of the basin (Figure 2b). These parts were found by identifying control points in the drainage network and removing their corresponding contributing areas. Perturbed numerical experiments simulate the same lake morphology with these different upland basin sizes (BSs). Besides BS, we are also interested in how spatial variability in climate patterns and interannual variability in river stage and weather affect C_L. Understanding the relative importance of these factors can shed light on how well our findings may generalize into other ria lakes. Thus, we employed analysis of variance (ANOVA) to attribute the variability in C_L to four factors: climate types (CTs), BS, the year of river stage data (SY), and the year of weather data (WY). For CT, we ran simulations with atmospheric forcing extracted from the CRUNCEP data set (Viovy, 2018) for Manaus, Tefe, Obidos, and Leticia. Each site is linked to a CT code, which serves as a categorical variable in ANOVA. These locations represent spatial heterogeneity in CTs that are found along the Amazon River. Tefe and Leticia are in the western lowland Amazon basin (upstream). Manaus is centrally located near Lake Calado, and Obidos is to the east (more downstream). SY and WY, coded as 1981 through 1988, were used as categorical variables to represent interannual variability in river stage and local climate, respectively. For each CT and BS, we ran one experiment with default forcing and another one with the year of weather forcing randomly permuted. For example, the forcing in 1987 could be migrated to 1981, while that of 1981 could be migrated to 1984. Without the permutation, WY and SY both naturally went through the years in 1981–1988 in each simulation year. With the permutation, the WY codes are reordered in exactly the same way as how the corresponding weather forcings are permuted. These permutation experiments allowed the impacts of river stage and interannual variability to be separated. Thus, given seven BSs and four CTs, we ran a total of $7 \times 4 \times 2 = 56$ experiments. To be able to run this number of simulations, we employed 200-m resolution simulations, but we note that the C_L curves, although showing some differences in the values of peaks and troughs, do not fundamentally change their patterns from the 80-m resolution runs. The dependent variable for the ANOVA is the C_L in January and June each year, when C_L is relatively steady. BS was used as a continuous variable. The ANOVA partitions the variance in the dependent variable into within-group and between-group variances. A four-way ANOVA test (assessed at the 1% confidence level) with interactions was employed.

While the WY and SY codes represent interannual variabilities, they do not represent an interpretable, quantitative value or a physical force that dominates the interannual changes. To provide more insights, we conduct ANOVA analyses with WY and SY replaced by continuous variables. To replace SY, we tested the average rate of river rise in the 30 days prior to

the reading of C_L ($\partial S/\partial t$). We also tested the peak and trough stages of the river. In place of WY, we tested the rainfall accumulated in m months prior to the reading of C_L, denoted by P_m. We varied m during the ANOVA and looked for the m value that explains the most variance. The idea is to identify the length of the system memory, and whether interannual variability in climate forcings can be mostly explained by differences in rainfall accumulation alone. For example, if P_1 explains more variance than P_{10}, it suggests that the system has a short-term memory, as long-term rainfall accumulation is not as relevant.

3 Result and Discussion

We first examine how well the model describes the hydrologic exchanges between Lake Calado and the Solimões River (section 3.1). Then we discuss the different phases of lake-river exchange and the evolution of river-water content (section 3.2). In section 3.3, we study the potential hydrologic control, including BS, rainfall seasonality, river stage, CT, and interannual variability via numerical experiments.

3.1 Comparing Model Estimated Fluxes With Field Data

The lake level time series matched well with the observed data (Figure 3a), and the maximum measured depth agreed with that simulated. This comparison merely shows that the enforced Solimões River boundary condition propagated into the domain and controlled the stage of the lake. The simulated river inflow to the lake from November 1983 to mid-March 1984 was 31.2×10^6 m^3, roughly 20% lower than LM95's estimate (27×10^6 m^3). LM95 estimated runoff by extrapolating from stream discharge data recorded at a weir gauging a 23.4-ha subcatchment. Spatial heterogeneity in topography and land cover could contribute to systematic deviations.

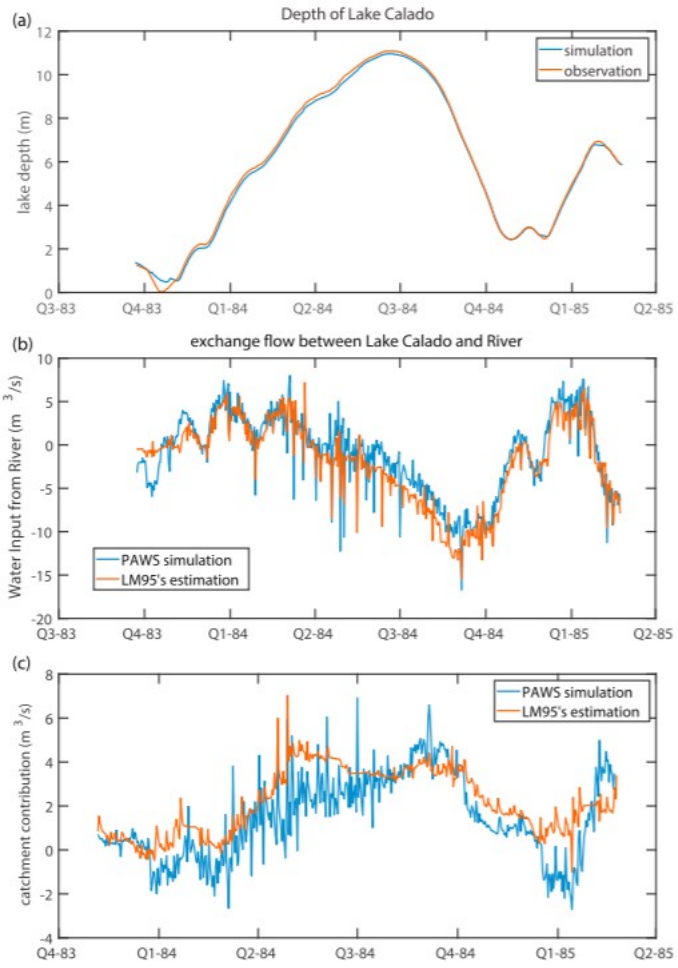

Figure 3. (a) Comparison between Process-based Adaptive Watershed Simulator (PAWS)-simulated lake depth and observational data. (b) Comparison between Lesack and Melack's (1995) estimation and PAWS simulated results for exchange flow between Lake Calado and Amazon River (NASH = 0.795). The tick on *x* axis represented the first day of the corresponding quarter in a year (Q1 starts with January). (c) Comparison of PAWS simulated catchment contributions to the lake and runoff + seepage estimated by Lesack and Melack (1995).

The exchange flow between Lake Calado and the Solimões River from October 1983 to mid-January 1985 is illustrated in Figure 3b. No calibration was performed to obtain this result. The simulation generally matched the LM95 estimates, producing a high Nash-Sutcliffe model efficiency coefficient (NASH = 0.795). The two estimates of the seasonal trend agreed well. Simulated river water input was ~2 m^3/s higher (less water going from the lake to the river) than LM95 estimated from Q2 to Q4 in 1984. The discrepancy resulted from a smaller simulated runoff compared to the measured (Figure 3c). Another smaller source of discrepancy was an intermittent connection to upstream Lake Miriti, which was not included in the model.

There were daily fluctuations in the river-lake exchange that were superimposed on the seasonal trend (Figure 3b). Most of the short-term spikes were negative, as they were introduced by local storms that caused an abrupt increase in lake level on a daily basis, which caused outflows (or

reduced inflows). Some spikes were captured well by the model; some were not, perhaps resulting from rainfall errors. Some of the daily spikes were shared by PAWS+CLM and LM95; some were not. Especially, during Q1 and early Q2 in 1984, PAWS+CLM tended to predict larger outflows than LM95. In particular, these low flow periods occurred when the exchange flux was sensitive to local storms.

In PAWS+CLM simulations the lake volume was not imposed as a boundary condition but determined by the river stage and catchment inputs. In contrast, in LM95 lake volume was observed and utilized in calculating the river-lake exchanges. The simulated river water inputs were extracted directly from exchanges between cells representing the river and the lake, whereas in LM95 they were estimated as residuals. The runoff contributions in LM95 were extrapolated from a gauge in a small catchment, while it was simulated in PAWS+CLM from the whole basin. In summary, because PAWS+CLM and LM95 rely on different mechanisms and information sources, their approximate agreement is encouraging.

3.2 Seasonal Trends in River Water Inflow and Lake Water Sources

The river-lake exchange is influenced by the combination of change in river stage, seasonality in local rainfall, and hydrologic modulation of local inputs. Beyond the period with field data, our 7 years of simulated water exchange for Lake Calado (Figure 4) illustrate that each year is different (Figure 4). River stage is omitted for clarity in Figure 4, and lake volume serves as a surrogate for river stage. A key aspect of the interannual differences is the interplay between phases when water levels are rising with river-water flowing into the lake (RWRI; shaded blue in Figure 4) versus the phase with rising water-levels and lake-water flowing out to the river (RWLO; shaded red in Figure 4), and how this interplay affects the river-water content in the lake (C_L). RWLO is a counterintuitive behavior relative to river-floodplain water exchanges, and it occurred in varying magnitudes and frequencies in 6 of the 7 years. This phase had substantial effects on the C_L in the lake, driving a range in annual peak values from 71% (1981) to 23% (1987). C_L during RWLO phases of a given year was also driven down, ranging from a decline from 64% to 48% in 1983, to a large decline from 52% to 6% in 1985. Switches between RWRI versus RWLO during a given water year generated multiple episodes of flow reversal between the river and lake, as illustrated in 1983, 1986, and 1987. Moreover, in 1987, an important antecedent effect that may have caused lower-than-average levels of C_L throughout the rising water period may have been that the lake began the water year about half full of water from the prior year.

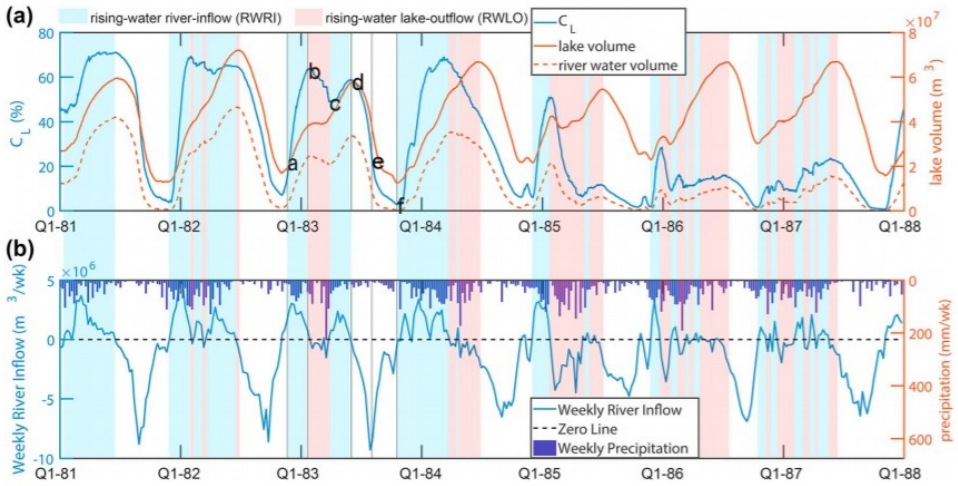

Figure 4. (a) Daily time series of fraction of lake water from the river, called the river-water content (blue solid line), lake volume (orange solid line, serving as an indicator for the lake stage), and volume of river water (orange dash line); the annotation of a–f represent 6 times in the 1982–1983 inundation cycle. (b) Time series of weekly river inflow to the lake (blue line) and precipitation in the watershed and Lake Calado (blue bar). The blue shading down through panels (a) and (b) represents the hydrological phase of rising-water with river inflow to the lake; the red shading represents rising-water with lake-outflow to the river.

We can thus distinguish four behavioral phases in the hydrological dynamics of the lake, but they do not necessarily occur in a fixed sequence within a given year or even necessarily recur every year in the case of RWLO. In the following, we describe the dynamics of each phase, using 1983 as an example with reversing flow during rising-water.

3.2.1 Low Water Phase

The seasonal low water phase starts in Q4 (*e* to *f*, 1983) and is characterized by a low and stable river stage. The river water inflow was around zero or slightly negative (Figure 4b; negative meaning lake water is flowing to the river). The connection between the lake and the river occurred through the narrow channel. Both river-water content and river water volume were at seasonal low values. At points a or f, the location-specific *C* value was highest near the channel (the outlet zone), around 30%, but it was near its lowest value (Figure 5). The low water level can vary significantly among years (note the effect on 1987 in Figure 4).

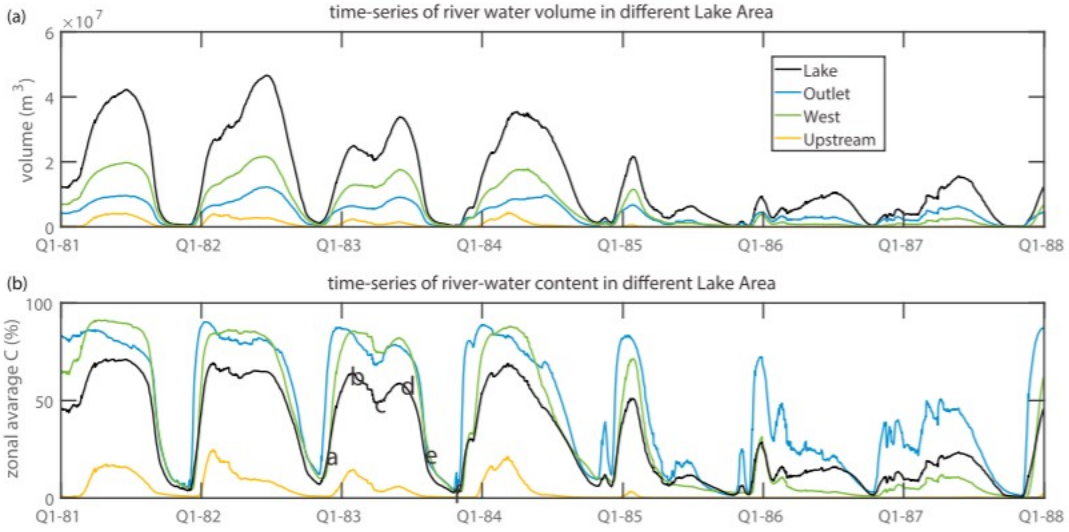

Figure 5. Time series of the river water volume (upper panel) and river-water content (lower panel) for different zones in the lake.

3.2.2 Rising-Water River-Inflow (RWRI)Phase

The low phase ends when the rainy season arrives in the western Amazon basin, around December each year. The floodwave descends downstream, commonly raising river stage ~0.1 m/day, though in some years the stage rises more slowly. For example, from October 1986 through early-January 1987, the rate of stage rise was only ~0.03 m/day. Thus, C_L within the lake rose slowly in this year. During early rising-water, the local catchment generally begins to receive increased rain. Simulated soil moisture storage is low at the beginning of the rainy season and can absorb a significant amount of rainfall via infiltration and evapotranspiration. At this time, the local basin did not produce enough runoff to raise the lake level at the same rate as the river rises.

Around midway through the rising-water phase (point b, 1983), the river-water content reached its seasonal maximum. The river water inflow rate peaked earlier and decreased to nearly 0 at point b. At the beginning of Q1 1983, the peak inflow rate corresponded to increased rainfall in the local basin. Integrated over the lake, around 64% of the lake water had come from the river at point b (Figure 4a). At the same time, *C* in the West zone had risen to the same level as the Outlet zone (Figure 5b). *C* was the highest and was nearly 1 at the center of the lake. Point *b* was also the time with the highest *C* in the Upstream zone. *C* rose to around 9% near the outlets of the upland streams. While the levee had narrowed because some of its areas had been submerged, there is a solid boundary between the lake and the river, and inflow occurred only through the channel connection to the river.

3.2.3 Rising-Water Lake-Outflow (RWLO)Phase

From points b to c, as the local rainfall strengthened, local water input was able to raise lake levels at a higher rate than the rise of the river. As a result, while the river stage and the total lake volume continued to rise at a rate from <0.01 to ~0.07 m/day, river-inflow ceased as lake-water flowed out to the river until point c. Consequently, C_L declined from ~64 to 50% as a result of the increase in the total lake volume. C in all zones of the lake declined during this period (Figure 5b). During 1983, there is a secondary peak in C_L near point d, caused by a switch from lake-outflow back to river-inflow (RWRI) that then persisted until the high-water point d. Throughout the rising-water period (RWRI + RWLO phases), the river is rising and acts as a *moving dam* relative to the lake, but concurrently, local water is adding water to the lake. When rates of local water input are higher than river rise rate, river inflow ceases, and C_L is either diluted by further addition of local water (i.e., no flow in either direction) or purged from the lake as water flows back out to the river. The balance between the two water sources is dynamic and fluctuating, as shown over the 7-year simulation period (Figure 4).

At point d of 1983, most of the levee had been submerged (as shown in the satellite image in Figure 1c). This submergence was captured by the model, which predicted a full contact between the river and the lake. C in the West zone remained high, around 80% (Figure 5b), and was only around 3% in the Upstream zone. Though the river stage determines the volume of the lake, the composition of lake water is determined by the combination of upland rainfall and runoff and river inflow.

3.2.4 Receding Phase and the Complete Cycle

Generally, beginning in June to July (*d* to *e*, 1983), the river stage falls, and the lake drains to the river as C_L continues to decline throughout the lake during this phase, driven by ongoing input of local water. The Outlet zone is more variable than other areas (Figure 5), and C in this zone declines faster than other zones. The West zone lags until point b, C declines in the Outlet and West zones during the RWLO and receding phases.

To examine the impacts of potential errors with inputs and spin-up, we ran tests with large biases, which showed that even with >1 m bias, the potential impacts remained limited (Figure S1 in the supporting information). This result is related to the large seasonal variability in the Solimões River. We also ran sensitivity tests where we looped the model through an additional cycle of the climate forcings and have not found significant impacts due to this spin up (Figure S2). Similarly, our sensitivity test showed that the eddy viscosity coefficient in the shallow water equation appears to have a negligible impact on C_L (Figure S3).

3.3 Hydrologic Controls of River Water Volume and Fraction

3.3.1 Basin Size Relative to the Lake Area

The time-averaged BA:LA ratio for Lake Calado and its local watershed is 10.8. Using the tested model, we performed simulation experiments to

examine how the river-water content changed if BA:LA were smaller (section 2.4). The volume of the lake is controlled by the river stage at the boundary. The river-water volume is, in turn, determined by how much water in the lake is contributed by upland sources. A smaller BA:LA ratio led to higher C_L (Figure 6a) and larger river water volume (Figure 6b). However, both curves were nonlinear and differ among the phases. At the end of the receding phase (point e) and the end of the second RWRI phase (*d*), reducing the BA:LA ratio caused the strongest increase in C_L among the phases. This behavior is because a lowered supply of upland runoff would weaken the dilution and flushing of river-water from the lake during the receding phase and would allow river-water to more strongly enter the lake during the RWRI phase.

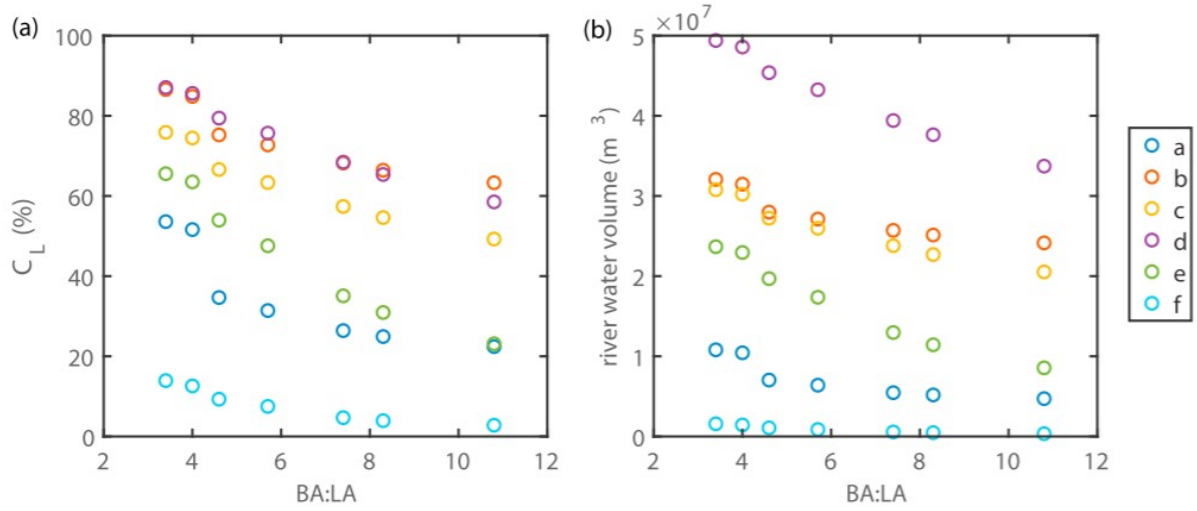

Figure 6. The impact of time-averaged basin to lake area (BA:LA) ratio on lake-integrated river-water content (left panel) and river water volume (right panel) at six points (a–f) during the 1983 water year (refer to Figure 4). Sensitivity of C_L to BA:LA (~slope of the coded points in the left panel) decline in the following order: point e, middle of the receding water > b, end of receding water > d, end of second RWRI ≫ d, end of RWLO > b, start of RWLO > *a*, start of initial RWRI ≫ *f*, end of low-water. Sensitivity of river water volume to BA:LA (~slope of the coded points in the right panel) as BA:LA is reduced decline in the following order: point e, middle of the receding water ≈ d, end of second RWRI > b, end of receding water ≫ c, end of RWLO > b, start of RWLO ≫ a, start of initial RWRI > f, end of low water.

At the start (point b) and end (c) of the RWLO phase during 1983, reducing BA:LA increased C_L by a lesser rate compared to *e* and *d*. More generally, however, as the supply of local water to the lake diminishes, the frequency and duration of RWLO phases should decline. When BA:LA fell below ~5 in the 1983 example, the peak annual C_L value (Figure 6a) started to occur at the end of the second RWRI phase (marked by point b) rather than at the start of the RWLO phase. When BA:LA was sufficiently small, RWLO phases might not occur at all, though the supply of local water may remain sufficient to still represent a significant portion (i.e., C_L < 100%) of the composite lake-water. In addition to the water supply from local runoff, RWLO is also affected by the rate of river rise and its peak level. Thus, the effect of reduced BA:LA would be stronger in water years such as 1984 and 1985,

when either per unit area runoff rates were higher because of higher rainfall (1984) or the rate of river rise progressed more slowly (1985).

At the end of low-water (point f) and at the start of initial RWRI (point a), the effect of reducing BA:LA on river-water volume in the lake was relatively weak. At these low-water times upland runoff flowed out of the lake and the lake is already dominated by local water. In these cases, reducing the supply of local water reduced the rate of flushing within the lake, but it appeared to have only a minor effect on C_L during low-water times. At the start of initial RWRI (*a*), the BA:LA effect was somewhat stronger on C_L than on the river-water volume because C_L was also affected by river water already in the lake, which differs over the 7 years shown in Figure 4.

3.3.2 Effects of Seasonal Rainfall Distribution and River Stage

As described earlier, we conducted numerical experiments to explore how seasonal distributions of rainfall and the Solimões River stage separately influence the sources of water. In 1984, the river-water content dropped rapidly from its peak to around 30%, when the lake volume was at the maximum (Figure 4a). There was no secondary peak of C_L in that year. Several hypotheses may explain the missing secondary peak: (1) The secondary peak was caused by the Solimões River stage and not rainfall. (2) The first-five-month rainfall in 1984 was large (1,553 mm, compared to 948, 1,391, and 1,049 from 1981 to 1983, respectively), generating sufficient runoff to reduce the input of river water. (3) Regardless of the previous month's precipitation, if there were large rainfall in April and May, then there would not be a secondary peak.

Our simulations (section 2.4) tested these hypotheses. When we repeated the rainfall of 1984 for all years, no year maintained a high level at the point with the highest lake volume. Therefore, hypothesis (1) is not correct. When we increased the rainfall from January to May in all years to be of the same total rainfall as in 1984, again C_L decreased. However, when we only boosted the rainfall in April and May, the same did not occur: C_L remained high and was different from the pattern in 1984. Therefore, hypothesis (3) is incorrect, while hypothesis (2) appears to hold.

In contrast, 1987 was the only year where C_L rose only slowly through the initial RWRI phase, but it was also a year with high river stage. The trough of the river stage in Q4 1986 was higher than that of the rest of the simulated years, with the lake starting the 1987 water-cycle nearly half full of water. Therefore, the low C_L peak in 1987 needs explanation. Either the rate of river rise was too slow or the river stage at the beginning of the water cycle was high and confined local water in the lake or the precipitation in the previous months accumulates enough runoff to match the rise of river stage. Toward the end of 1986, the rainy season came earlier for this water year. The rainfall from October to December in 1986 was 272, 343, and 161 mm respectively, significantly higher than other years. When we migrated the rate of river stage rise from 1987 to 1981, C_L in 1981 rose to a level similar

to 1987, suggesting that the river stage was a significant reason for the low C_L peak in 1987. However, in the experiment when we repeated precipitation of 1984 (less rainfall in Q4), the C_L peaks in 1986 and 1987 were notably higher than those in the default simulation (Figure 7). Therefore, the peak of C_L in these 2 years appeared to be controlled by the combination of rainfall and rate of river stage rise: the larger-than-normal rainfall in the months prior effectively caused an early RWLO phase, while the river rise in 1987 was too slow to dominate over local contributions.

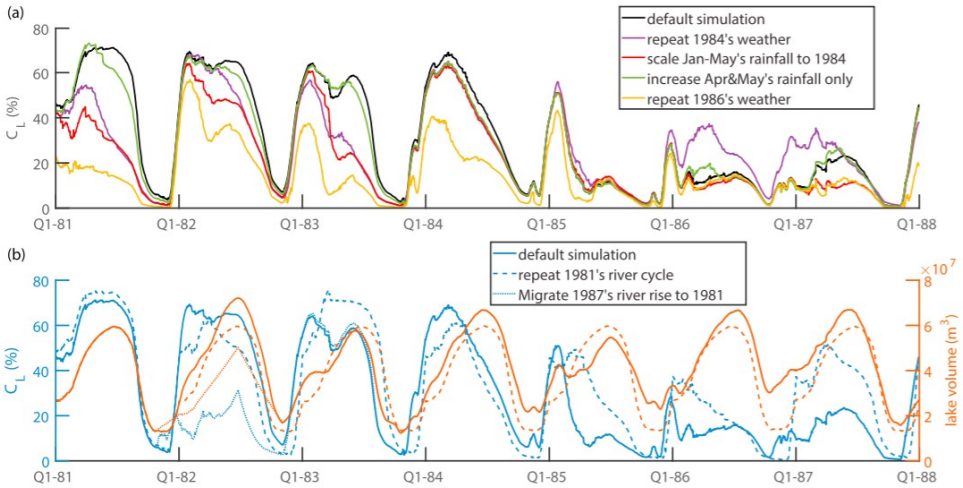

Figure 7. (a and b) Experimental examination of the influence of seasonal rainfall distribution, inter-annual precipitation variability, and Solimões River stage. In (a) the black line represents the default simulation, the purple line denotes the simulation when the rainfall in 1984 is repeated in all years, the red line is the simulation that linearly scales the storm events so that precipitation from January and May of each year sums up to be the amount in the same months in 1984, the green line is the simulation that scales the precipitation in April and May to 300 mm/month, and the yellow line repeated the forcings in 1986 to all other years. In (b), the solid line denotes the default simulation, the dashed line represents the simulation that repeats 1981 river stages, and the dotted line is the simulation that migrates the slow river rise in Q4-1986 to Q4-1981 (then assumes a linear increase until the stage reaches the river peak in 1981).

To examine how the river stage influences the river-water content, we migrated the river stage data from 1981 to all years. The peak river stage of 1981 was lower than in 1982 and 1984 but was similar to 1983. The changes in river stage have modified the river-water content, which would be higher when the river stage is higher. The river stage in 1981 had a characteristic rapid rise near the beginning of Q1, and then the rate of river rise slowed. This pattern led to a short-lived rapid rise phase in C_L and a delay of the C_L peak, which did not arrive until the beginning of Q2. However, the resulting C_L patterns were notably different between different years, indicating both rainfall distributions and river stages play important roles.

3.3.3 Relative Importance of Climate Type, Interannual Variability, and Basin Size

Earlier we showed that rainfall distributions, BS, and interannual variability in both climate and river stage could all contribute to different dynamics in river-water content. Our ANOVA analysis of the simulations was designed to offer a quantitative analysis of their relative importance for the lake-integrated river-water content. The ANOVA analysis indicates that CT, BS, and year can explain a majority of the variance for January, June, and

October monthly average C_L (Figure 8a). January is near the end of the initial RWRI phase that most commonly coincides with the peak annual C_L, June is close to the peak of the river water level at the end of the RWLO phase, and October is at the end of an annual water cycle, or just before the river stage starts to rise. The small percentage of variance associated with error and interaction terms suggest the chosen factors are effective controls of the C_L variability, and the nonlinear effects appear to be weak. For January C_L, the influence of CTs and river stage's variability (SY) dominated over BS and climatic interannual variability. This behavior was not surprising, as the rate at which the river stage rose was shown to have strong control in the initial RWRI phase. Meanwhile, C_L shows substantially different behaviors under different CTs (Figure 9). Obidos, in the eastern Amazon, has a drier climate than other sites. The total rainfall is 1,823 mm/year in Obidos compared to the 2,461 mm/year in Leticia. Thus, this site had the highest C_L. Leticia had the highest rainfall and the lowest C_L. The second important factor is BS. BS appeared to be more influential for June than January C_L. During the high-water period of June, the lake may be in the RWLO phase, and C_L is a balance between upland contribution and river stage. Therefore, runoff processes (and thus BA) are more important for this phase (section 3.3.1). For January, interannual variability in both climate forcing and river stage are of similar importance to the BS. For June, however, the interannual variability has a smaller role. October is similar to June except for the effect of SY is larger, because in some years, for example, 1983 and 1986, the river stage began to increase in October.

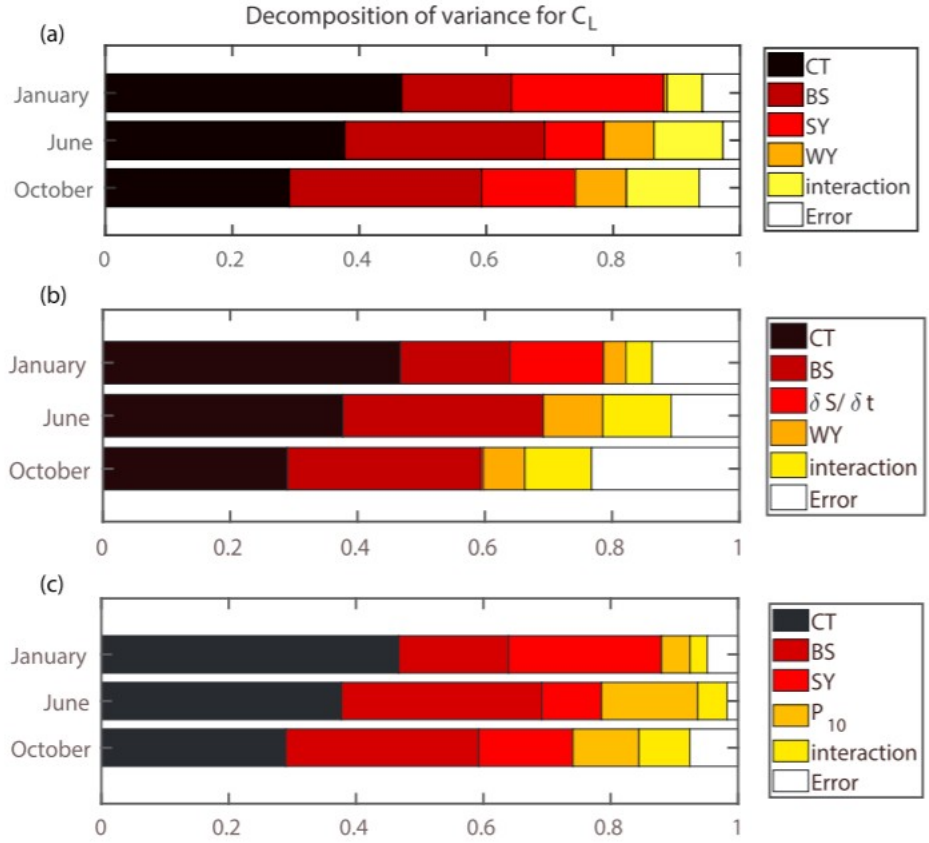

Figure 8. (a) Decomposition of variances with an analysis of variance, which compares the effects of interannual variability of river stage (SY) and weather (WY), versus basin size (BS), versus climate types (CT). Different climate types considered include Climate Research Unit – National Centers for Environmental Prediction (CRUNCEP) weather data from Manaus, Leticia, Tefe, and Obidos. In (b) we replaced SY by $\delta S/\delta t$, the average rate of river rise in the 30 prior days, and in (c) we replaced WY by P_{10}, cumulative rainfall in 10 prior months.

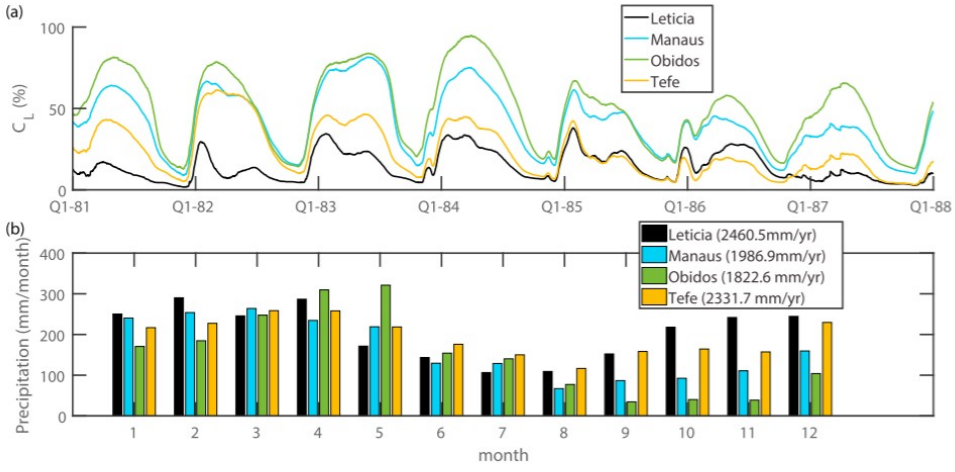

Figure 9. (a) C_L time series under different climate types (Leticia, Manaus, Obidos, and Tefe, all from CRUNCEP data set). (b) Monthly rainfall distribution of these four climate types (average from 1981 to 1988).

We have also attempted an ANOVA analysis where SY was replaced with the average rate of river rise in the 30 prior days ($\partial S/\partial t$). According to our previous observations, it seemed the faster river stage rises, more river water can flow into the lake, forming a higher peak in January. However, with this replacement, the total explained variance (100%-error) was reduced (Figure 8b). $\partial S/\partial t$ explained around 60% of the variance that has previously been assigned to SY in January. This observation suggests that the influence of SY is a complex and nonlinear, but the rate of river water rise is indeed an important aspect of the interannual river stage variability, and the RWRI phase is indeed dependent on how fast the river rises. However, it cannot fully represent the interannual variability in the river stage. Other characteristics of the river stage dynamics, for example, lowest and highest river stage, may also play roles. For June and October, however, $\partial S/\partial t$ has no explanatory power as these phases were controlled by climate forcings and upland runoff processes.

We tested replacing the categorical variable WY with P_1 through P_{16} as a continuous variable. We found when $m > 6$, P_m became a powerful explanatory variable (Figure 10a), even supporting more total explainable variance than WY (Figure 10b), which suggests that precipitation is a more important control in this tropical climate than other climatic variables. P_{10} explained the most variance in ANOVA, but P_9 to P_{10} were quite similar. This result indicates a long system memory, which is related to the groundwater-dominated streamflow generation in this part of the Amazon (Niu et al., 2017) and a thick soil acting as a large buffer for precipitation. However, P_{16} explained much less variance than P_{12}, which means there is little carryover memory from longer than a year, perhaps due to the annual flushing of the lake. Since this ANOVA analysis was carried out with different BA:LA ratios and the result showed only a small interaction, it means a long system memory could be found even with much smaller local basins. Niu et al. (2017) showed that in a nearby 9,000-km^2 basin with similar characteristics, groundwater contributions fluctuate between 60% and 85% of basin outflows. As a result, P_1–P_3 explained much less variance than P_{10}. Consistent with earlier discussion, in January, P_m had much less explanatory power because at this time the lake water content was dominated by inflows from the rising river. However, interestingly, June and October also show different characteristics. At the end of the RWLO phase, June was influenced by monthly scale to annual scale precipitation, whereas only annual-scale precipitation mattered for October (end of an annual water cycle).

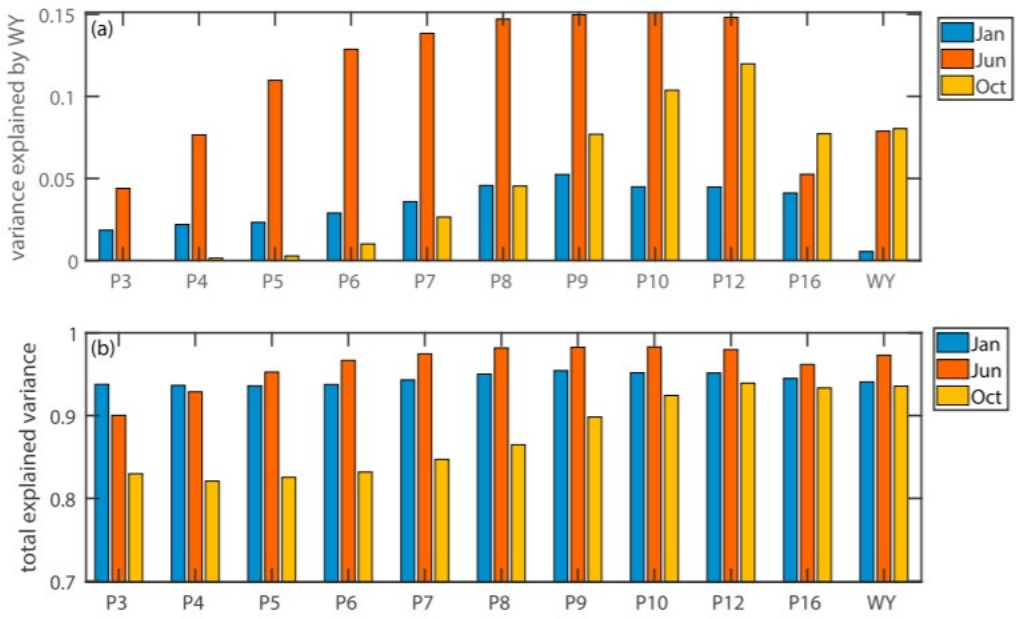

Figure 10. (a) The fraction of variance explained by m months of accumulated rainfall (as indicated by P_m on the x axis) from analysis of variance (ANOVA). WY means the weather year code was used as a categorical variable. For example, P_{10} means that the 10-month accumulated rainfall (prior to the reading of C_L) was used as a continuous variable in the ANOVA analysis. The peak of explained variance occurred around 8–10 months, indicating the system has a long memory. (b) The total explainable variance by ANOVA with linear terms (CT, BS, SY, and P_m) and first-order interaction terms. Note that the y-axis lower limit of (b) is not 0.

3.4 Further Discussion

Our model experiments provide guidance for modeling Amazon floodplain lakes at large scales with reduced model complexity. First, the system is highly transient and depends on CT and river stage each year. Judging from Figures 8 and 9, spatial and interannual variabilities have strong impacts on the water mixture and could not be predicted by simple metrics like 5-month precipitation. The nonlinear and intertwined influences of climate and river stage make it difficult to devise simplistic reduced-order models that rely on simple predictors like BS and annual precipitation. Thus, a dynamical modeling system considering real-world forcing and river stages is preferred over methods assuming fixed seasonal patterns. Second, for the RWLO phase, it is important to consider the runoff generation process in the upland catchments realistically. Third, given that the system has a long memory (rainfall of the preceding months affects C_L), a long-term continuous hydrologic model is required to simulate the effects of hydrologic buffering of the soil and groundwater system.

Overall results show that the mixture of waters in Lake Calado differed in each year that was modeled. Our results also illustrate that the hydrological connectivity between the river channel and floodplain lakes, even in a setting where a physical connection appears to be sustained year round, can be complicated. Thus far, hydrological connectivity in large river floodplains has mostly been approached by identification of disconnected periods from

analysis of high- and low-water levels extracted from river-stage records (e.g., Lesack & Marsh, 2010; Thorp et al., 2006). Our simulations of lengthy and annually recurring hydrological phases where water flows from the floodplain to the river during periods of rising river levels (RWLO) are counterintuitive and, thus far, a poorly documented type of hydrological connectivity. One could argue that during such phases, the river is physically connected to the lake, but the lake is not functionally connected to the river if primary interest is in water, sediments, solutes, or plankton in the lake. Most motile organisms such as fish would be an exception because they can move through the connection despite the direction of flow. RWLO is a biogeochemically important type of hydrological connectivity, and its presence and recurrence rates should be further explored in other river ecosystems within humid climates.

4 Limitations and Future Work

The presented model did not include the effects of longitudinal shear associated with riverine inflows, whose impact is limited to high water times when the river stage rose above the natural levees. We have assumed vertical homogeneity in modeling flow and water fractions in the lake. During high water, we expect periods of stratification of the lake (MacIntyre & Melack, 1988), which should have minor impacts on water flow, but may lead to differences in the distribution of solutes. Hence, we could temporarily underestimate river water fraction in one vertical layer of the lake while overestimating it in another. However, as we are mostly concerned with lake-integrated C_L, we do not expect our results and conclusions to be affected by this simplification. We recognize that net water exchange between the river and lake versus hydrodynamically mixed fractions of water are not equivalent. Our C_L values agree well with LM95 (i.e., net exchange), but actual river water fractions could be different depending on extent of mixing. Despite these limitations, the good agreement between observed and simulated river-water exchange suggests that the modeling is capturing the salient feature of the exchange process.

The ANOVA tests were done using CRUNCEP data, though these rainfall data are known to miss large storms. This choice was made because no station data are available across Amazon basin, and TRMM data were not available for our study period. However, we also note that due to the buffering effects of the hydrologic system, the ANOVA tests are not as sensitive to large storms as total rainfall over a few months. Also, the regional differences between sites are considerable.

Collection of field data to develop hydrological balances and validate models has evolved significantly since the measurements made at Lake Calado in the 1980s. Sensors linked to digital data loggers are available to record relevant meteorological and stage data. Acoustic Doppler current profiles allow integrated measurements of discharge. Remote sensing data and approaches can also be utilized as inputs to hydrological models of Amazon

floodplain lakes, e.g., (Alsdorf et al., 2000; Hamilton et al., 2002; Hess et al., 2003; Policelli et al., 2018; Rudorff et al., 2014b). Advances in modeling and measuring hydrodynamic processes permit mixing to be examined mechanistically (Augusto-Silva et al., 2018; MacIntyre et al., 2014; Tedford et al., 2014; Vidal et al., 2013). In combination, these advances will improve understanding of the complex processes in floodplain systems.

5 Conclusions

A physically based surface-subsurface process model was modified to model integrated catchment hydrology and river inundation in an Amazon floodplain lake. The model allowed multiyear simulations of the seasonal variations in the hydrology of the system and represented well four hydrological phases: low water, rising-water with river-inflow to the lake (RWRI), rising-water with lake-outflow to the river (RWLO), and receding phases. Water exchanges between the river and lake agreed well with field estimates, including the timings of different phases, and extended the 1 year of field estimates; However, the multiyear simulation revealed surprisingly strong interannual variability and some phases could be shortened or weakened in years with certain river stage and climate setups. The seven simulated years all differed from each other. These strong interannual differences are driven by the interplay between local climate and river stage fluctuations and their interannual variabilities and determine the river-water content in the lake (C_L). C_L during 5 of the 7 years peaked well before the peak river stage, and the interannual C_L peaks varied from 73 to 23% of the water in the lake. These results highlight a highly dynamical and evolving system conditioned by multiple forcings (river stage and climate).

The cumulative rainfall in the previous 6–12 months influences C_L. These antecedent effects differ among seasons as seasonal to annual rainfall accumulation, but not longer than a year, controls C_L at peak river level (end of the RWLO phase), whereas only the annual precipitation accumulation affects C_L during October (low water). The rate of seasonal river rise at the beginning of the annual flood cycle also had a strong impact during the RWRI phase. The faster the river rises in a year, the more river water flows into the lake. A particularly slow rise will result in an earlier than usual switch from the RWRI to RWLO phase or multiple switches between these two phases and a lack of rapid filling of the lake. During the RWLO phase, C_L is diluted and purged from the lake, as local water input raises lake-water levels at a higher rate than the rise of the river. The importance and recurrence of this phase is a function of the area of the local basin. Analysis of variance found that CTs along the Amazon River is a significant control of river water content, followed by BS and interannual variability in rainfall and river stage. However, factors reflecting the influence of local hydrology, for example, BS and CT, are more important for the RWLO phase than the RWRI phase. Moreover, the system has a subannual memory as approximately 10 months of accumulated rainfall could explain the most variance associated with interannual variability in local climate.

Acknowledgments

This research was supported by Office of Biological and Environmental Research of the U.S. Department of Energy under contract DE-SC0010620. Shen was partially supported by U.S. Department of Energy under contract DE-SC0016605. Additional support for J. M. M. during the manuscript preparation was provided by NASA grant NNX17AK49G. The preprocessed data used in this work, including observations of lake stage, climate forcing, river-lake exchanges, and runoff estimated in LM95, can be accessed in the supporting information of this paper. The river stage and daily precipitation were obtained from ANA, Brazilian National Water Agency. The CRUNCEP forcing data were obtained from Viovy (2018). The lake stage, river-lake exchange, and runoff estimates were obtained from Lesack's PhD thesis.

References

Alsdorf, D. E., Han, S., Bates, P., & Melack, J. (2010). Seasonal water storage on the Amazon floodplain measured from satellites. *Remote Sensing of Environment*, 114(11), 2448– 2456. https://doi.org/10.1016/j.rse.2010.05.020

Alsdorf, D. E., Melack, J. M., Dunne, T., Mertes, L. A. K., Hess, L. L., & Smith, L. C. (2000). Interferometric radar measurements of water level changes on the Amazon flood plain. *Nature*, 404(6774), 174– 177. https://doi.org/10.1038/35004560

Amoros, C., & Bornette, G. (2002). Connectivity and biocomplexity in waterbodies of riverine floodplains. *Freshwater Biology*, 47(4), 761– 776. https://doi.org/10.1046/j.1365-2427.2002.00905.x

Augusto-Silva, P. B., MacIntyre, S., Rudorff, C. M., Cortes, A., & Melack, J. M. (2018). Stratification and mixing in large shallow lakes along the lower Amazon River floodplain. *Journal of Great Lakes Research*. https://doi.org/10.1016/j.jglr.2018.11.001

Bayley, P. B., & Petrere, M. (1989). Amazon fisheries: Assessment methods, current status and management options. In D. P. Dodge (Ed.), *Proceedings of the International Large River Symposium* (pp. 385– 398). Ottawa: Canadian Special Publication of Fisheries and Aquatic Sciences.

Bonan, G. B., Levis, S., Kergoat, L., & Oleson, K. W. (2002). Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles*, 16(2), 1021. https://doi.org/10.1029/2000GB001360

Bonnet, M. P., Barroux, G., Martinez, J. M., Seyler, F., Moreira-Turcq, P., Cochonneau, G., Melack, J. M., Boaventura, G., Maurice-Bourgoin, L., León, J. G., Roux, E., Calmant, S., Kosuth, P., Guyot, J. L., & Seyler, P. (2008). Floodplain hydrology in an Amazon floodplain lake (Lago Grande de Curuaí). *Journal of Hydrology*, 349(1–2), 18– 30. https://doi.org/10.1016/j.jhydrol.2007.10.055

Bonnet, M. P., Pinel, S., Garnier, J., Bois, J., Resende Boaventura, G., Seyler, P., & Motta Marques, D. (2017). Amazonian floodplain water balance based on modelling and analyses of hydrologic and electrical conductivity data. *Hydrological Processes*, 31(9), 1702– 1718. https://doi.org/10.1002/hyp.11138

Carabajal, C. C., & Harding, D. J. (2006). SRTM C-band and ICEsat laser altimetry elevation comparisons as a function of tree cover and relief. *Photogrammetric Engineering & Remote Sensing*, 72(3), 287– 298. https://doi.org/10.14358/PERS.72.3.287

Casulli, V. (1990). Semi-implicit finite difference methods for the two-dimensional shallow water equations. *Journal of Computational Physics*, 5674, 56– 74.

Dunne, T., Mertes, L. A. K., Meade, R. H., Richey, J. E., & Forsberg, B. R. (1998). Exchanges of sediment between the flood plain and channel of the Amazon River in Brazil. *Geological Society of America Bulletin*, 110(4), 0450–0467. https://doi.org/10.1130/0016-7606(1998)110<0450:EOSBTF>2.3.CO;2

Forsberg, B. R., Devol, A. H., Richey, J. E., Martinelli, L. A., & dos Santos, H. (1988). Factors controlling nutrient concentrations in Amazon floodplain lakes. *Limnology and Oceanography*, 33(1), 41– 56. https://doi.org/10.4319/lo.1988.33.1.0041

Forsberg, B. R., Melack, J. M., Richey, J. E., & Pimentel, T. P. (2017). Regional and seasonal variability in planktonic photosynthesis and planktonic community respiration in Amazon floodplain lakes. *Hydrobiologia*, 800(1), 187– 206. https://doi.org/10.1007/s10750-017-3222-3

Gourou, P. (1949). Observações geográficas na Amazônia. *Revista Brasileira de Geografia*, 11(3), 355– 408.

Guber, A. K., Pachepsky, Y. A., van Genuchten, M. T., Rawls, W. J., Simunek, J., Jacques, D., Nicholson, T. J., & Cady, R. E. (2006). Field-scale water flow simulations using ensembles of pedotransfer functions for soil water retention. *Vadose Zone Journal*, 5(1), 234. https://doi.org/10.2136/vzj2005.0111

Guber, A. K., Pachepsky, Y. A., van Genuchten, M. T., Simunek, J., Jacques, D., Nemes, A., Nicholson, T. J., & Cady, R. E. (2009). Multimodel simulation of water flow in a field soil using pedotransfer functions. *Vadose Zone Journal*, 8(1), 1. https://doi.org/10.2136/vzj2007.0144

Gunduz, O., & Aral, M. M. (2005). River networks and groundwater flow: A simultaneous solution of a coupled system. *Journal of Hydrology*, 301(1–4), 216– 234. https://doi.org/10.1016/j.jhydrol.2004.06.034

Hamilton, S. K., Sippel, S. J., & Melack, J. M. (2002). Comparison of inundation patterns among major South American floodplains. *Journal of Geophysical Research*, 107(D20), 8038. https://doi.org/10.1029/2000JD000306

Hess, L. L., Melack, J. M., Affonso, A. G., Barbosa, C., Gastil-buhl, M., & Novo, E. M. L. M. (2015). Wetlands of the lowland Amazon basin: Extent, vegetative cover, and dual-season inundated area as mapped with JERS-1 Synthetic Aperture Radar. *Wetlands*, 35(4), 745– 756. https://doi.org/10.1007/s13157-015-0666-y

Hess, L. L., Melack, J. M., Novo, E. M. L., Barbosa, C. C., & Gastil, M. (2003). Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. *Remote Sensing of Environment*, 87(4), 404– 428. https://doi.org/10.1016/j.rse.2003.04.001

Ji, X., & Shen, C. (2018). The introspective may achieve more: Enhancing existing geoscientific models with native-language structural reflection. *Computers and Geosciences*, 110, 32– 40. https://doi.org/10.1016/j.cageo.2017.09.014

Ji, X., Shen, C., & Riley, W. J. (2015). Temporal evolution of soil moisture statistical fractal and controls by soil texture and regional groundwater flow. *Advances in Water Resources*, 86, 155– 169. https://doi.org/10.1016/j.advwatres.2015.09.027

Lesack, L. F. W. (1988). Mass balance of nutrients, major solutes, and water in an Amazon floodplain lake and biogeochemical implications for the Amazon Basin. University of California, Santa Barbara.

Lesack, L. F. W. (1993). Export of nutrients and major ionic solutes from a rain forest catchment. *Water Resources Research*, 29(3), 743– 758. https://doi.org/10.1029/92WR02372

Lesack, L. F. W., & Marsh, P. (2010). River-to-lake connectivities, water renewal, and aquatic habitat diversity in the Mackenzie River Delta. *Water Resources Research*, 46, W12504. https://doi.org/10.1029/2010WR009607

Lesack, L. F. W., & Melack, J. M. (1991). The deposition, composition, and potential sources of major ionic solutes in rain of the Central Amazon Basin. *Water Resources Research*, 27(11), 2953– 2977. https://doi.org/10.1029/91WR01946

Lesack, L. F. W., & Melack, J. M. (1995). Flooding hydrology and mixture dynamics of lake water derived from multiple sources in an Amazon floodplain lake. *Water Resources Research*, 31(2), 329– 345. https://doi.org/10.1029/94WR02271

Liu, L., Phanikumar, M. S., Molloy, S. L., Whitman, R. L., Shively, D. A., Nevers, M. B., Schwab, D. J., & Rose, J. B. (2006). Modeling the transport and inactivation of E. coli and Enterococci in the near-shore region of Lake Michigan. *Environmental Science & Technology*. https://doi.org/10.1021/ES060438K

MacIntyre, S., & Melack, J. M. (1988). Frequency and depth of vertical mixing in an Amazon floodplain lake (L. Calado, Brazil). *Verhandlugen Internationale*

Vereinigen Limnologia, 23, 80– 85. https://doi.org/10.1080/03680770.1987.11897906

MacIntyre, S., Romero, J. R., Silsbe, G. M., & Emery, B. M. (2014). Stratification and horizontal exchanges in Lake Victoria, East Africa. *Limnology and Oceanography*, 59, 1805– 1838.

Marengo, J., Nobre, C. A., Betts, R. A., Cox, P. M., Sampaio, G., & Salazar, L. (2009). Global warming and climate change in Amazonia: Climate-vegetation feedback and impacts on water resources. *Amazonia and Global Change*, 186, 273– 292. https://doi.org/10.1029/2008GM000743

Marengo, J. A. (2004). Interdecadal variability and trends of rainfall across the Amazon basin. *Theoretical and Applied Climatology*, 78(1–3), 79– 96. https://doi.org/10.1007/s00704-004-0045-8

Martin, N., & Gorelick, S. M. (2005a). MOD_FreeSurf2D: A MATLAB surface fluid flow model for rivers and streams. *Computers & Geosciences*, 31(7), 929– 946. https://doi.org/10.1016/j.cageo.2005.03.004

Martin, N., & Gorelick, S. M. (2005b). Semi-analytical method for departure point determination. *International Journal for Numerical Methods in Fluids*, 47, 121– 137.

Mays, L. W. (2010). *Water Resources Engineering* (2nd ed.). Tempe, AZ: Wiley.

Melack, J. M. (2016). Aquatic ecosystems. In L. Nagy, B. R. Forsberg, & P. Artaxo (Eds.), *Interactions between biosphere, atmosphere and human land use in the Amazon Basin* (pp. 119– 148). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49902-3_7

Melack, J. M., & Coe, M. T. (2012). Climate change and the floodplain lakes of the Amazon basin. In *Climatic change and global warming of inland waters* (pp. 295– 310). West Sussex, UK: Wiley-Blackwell. https://doi.org/10.1002/9781118470596.ch17

Melack, J. M., & Forsberg, B. R. (2001). Biogeochemistry of Amazon floodplain lakes and associated wetlands. In *The biogeochemistry of the Amazon Basin* (pp. 235– 274). New York, NY: Oxford University Press.

Melack, J. M., Hess, L. L., Gastil, M., Forsberg, B. R., Hamilton, S. K., Lima, I. B. T., & Novo, E. M. L. M. (2004). Regionalization of methane emissions in the Amazon Basin with microwave remote sensing. *Global Change Biology*, 10(5), 530– 544. https://doi.org/10.1111/j.1365-2486.2004.00763.x

Melack, J. M., Novo, E. M. L. M., Forsberg, B. R., Piedade, M. T. F., & Maurice, L. (2009). Floodplain ecosystem processes. In J. Gash, M. Keller, & P. Silva-Dias (Eds.), *Amazonia and global change*, Geophysical Monograph Series 186 (pp. 525– 541). Washington, DC: American Geophysical Union. https://doi.org/10.1029/2008GM000721

Mertes, L. A. K. (1997). Documentation and significance of the perirheic zone on inundated floodplains. *Water Resources Research*, 33(7), 1749– 1762. https://doi.org/10.1029/97WR00658

Niu, J., Shen, C., Chambers, J., Melack, J. M., & Riley, W. J. (2017). Interannual variation in hydrologic budgets in an Amazonian watershed with a coupled subsurface - land surface process model. *Journal of Hydrometeorology*, 18(9), 2597– 2617. https://doi.org/10.1175/JHM-D-17-0108.1

Nobre, C. A., Obregón, G. O., Marengo, J. A., Fu, R., & Poveda, G. (2009). Characteristics of Amazonian climate: Main features. In M. Keller, M. Bustamante, J. Gash, & P. Silva Dias (Eds.), *Amazonia and global change, Geophysical Monograph Series* (Vol. 186, pp. 149– 162). Washington, DC: American Geophysical Union.

Natural Resources Conservation Service (1986). Urban hydrology for small watersheds -- TR-55.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, J. -F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D. M., Sacks, W., Sun, Y., Tang, J., & Yang, Z. -L. (2013). NCAR/TN-503+STR NCAR Technical Note, (July).

Paiva, R. C. D., Buarque, D. C., Collischonn, W., Bonnet, M.-P., Frappart, F., Calmant, S., & Bulhões Mendes, C. A. (2013). Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research*, 49, 1226– 1243. https://doi.org/10.1002/wrcr.20067

Pau, G. S. H., Shen, C., Riley, W. J., & Liu, Y. (2016). Accurate and efficient prediction of fine-resolution hydrologic and carbon dynamic simulations from coarse-resolution models. *Water Resources Research*, 52, 791– 812. https://doi.org/10.1002/2015WR017782

Policelli, F., Hubbard, A., Jung, H., Zaitchik, B., & Ichoku, C. (2018). Lake Chad total surface water area as derived from land surface temperature and radar remote sensing data. *Remote Sensing*, 10(2), 252. https://doi.org/10.3390/rs10020252

Riley, W. J., & Shen, C. (2014). Characterizing coarse-resolution watershed soil moisture heterogeneity using fine-scale simulations. *Hydrology and Earth System Sciences*, 18(7), 2463– 2483. https://doi.org/10.5194/hess-18-2463-2014

Rudorff, C. M., Dunne, T., & Melack, J. M. (2018). Recent increase of river–floodplain suspended sediment exchange in a reach of the lower Amazon River. *Earth Surface Processes and Landforms*, 43(1), 322– 332. https://doi.org/10.1002/esp.4247

Rudorff, C. M., Melack, J. M., & Bates, P. D. (2014a). Flooding dynamics on the lower Amazon floodplain: 1. Hydraulic controls on water elevation,

inundation extent, and river-floodplain discharge. *Water Resources Research*, 50, 619– 634. https://doi.org/10.1002/2013WR014091

Rudorff, C. M., Melack, J. M., & Bates, P. D. (2014b). Flooding dynamics on the lower Amazon floodplain: 2. Seasonal and interannual hydrological variability. *Water Resources Research*, 50, 635– 649. https://doi.org/10.1002/2013WR014714

Shen, C., Niu, J., & Fang, K. (2014). Quantifying the effects of data integration algorithms on the outcomes of a subsurface–land surface processes model. *Environmental Modelling & Software*, 59, 146– 161. https://doi.org/10.1016/j.envsoft.2014.05.006

Shen, C., Niu, J., & Phanikumar, M. S. (2013). Evaluating controls on coupled hydrologic and vegetation dynamics in a humid continental climate watershed using a subsurface - land surface processes model. *Water Resources Research*, 49, 2552– 2572. https://doi.org/10.1002/wrcr.20189

Shen, C., & Phanikumar, M. S. (2010). A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling. *Advances in Water Resources*, 33(12), 1524– 1541. https://doi.org/10.1016/j.advwatres.2010.09.002

Shen, C., Riley, W. J., Smithgall, K. M., Melack, J. M., & Fang, K. (2016). The fan of influence of streams and channel feedbacks to simulated land surface water and carbon dynamics. *Water Resources Research*, 52, 880– 902. https://doi.org/10.1002/2015WR018086

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research*, 116, G04021. https://doi.org/10.1029/2011JG001708

Sippel, S. J., Hamilton, S. K., & Melack, J. M. (1991). Inundation area and morphometry of lakes on the Amazon River floodplain, Brazil. *Archiv für Hydrobiologie*, 123, 385– 400.

Stanford, J. A., & Ward, J. V. (1993). An ecosystem perspective of alluvial rivers: Connectivity and the hyporheic corridor. *Journal of the North American Benthological Society*, 12(1), 48– 60. https://doi.org/10.2307/1467685

Tedford, E. W., MacIntyre, S., Miller, S. D., & Czikowsky, M. J. (2014). Similarity scaling of turbulence in a small temperate lake during fall cooling. *Journal of Geophysical Research: Oceans*, 119, 4689– 4713. https://doi.org/10.1002/2014JC010135

Thorp, J. H., Thoms, M. C., & Delong, M. D. (2006). The riverine ecosystem synthesis: Biocomplexity in river networks across space and time. *River Research and Applications*, 22(2), 123– 147. https://doi.org/10.1002/rra.901

Vidal, J., MacIntyre, S., McPhee-Shaw, E. E., Shaw, W. J., & Monismith, S. G. (2013). Temporal-spatial variability of the internal wave field in a lake with complex morphometry. *Limnology and Oceanography*, 58, 1557– 1580.

Viovy, N. (2018). CRUNCEP Version 7—Atmospheric forcing data for the Community Land Model. Retrieved from http://rda.ucar.edu/datasets/ds314.3/

Zheng, C., & Bennett, G. D. (2002). *Applied contaminant transport modeling* (2nd ed.). New York, NY, USA: John Wiley.