Response to Community Comment #1 (CC1)

Manuscript:

Long-term hydro-sediment dynamics of the Ucayali River (Amazon Basin) revealed through combined observations, remote sensing, and SWAT-Amazon modelling

Hydrology and Earth System Sciences – egusphere-2025-4101

William Santini et al.

30 March 2026

We thank Dr. Dhruv Sehgal for a careful and constructive review of our manuscript. His comments were particularly helpful in sharpening the statement of research objectives, clarifying the granulometric representation adopted in the framework, and strengthening the calibration and validation strategy. We have revised the manuscript accordingly and address each comment in detail below.

The referee comments are reproduced in italic, followed by our responses.

Sincerely,
W. Santini et al.

General Comments

Thank you for the opportunity to review this manuscript. The paper presents an integrative framework to quantify fine and sand-sized sediment fluxes in a large floodplain river system by combining discharge modelling, floodplain-aware routing, and satellite-derived suspended sediment estimates. The topic is timely and potentially high impact. I have listed my section-based comments below.

Abstract

1- *While I understand the large goal of the paper is to highlight the role of floodplain in sediment transport, and the paper is attempting to build an integrated approach using discharge estimates to calculate fine sediment load and land surface model to simulate sand load. I cannot understand from the latter half of the Abstract what critical research questions this paper attempts to solve, except the role of floodplain in sediment transport.*

2- *The abstract misses in providing uncertainty or targeted application of this study, as the last three lines reads very generic. Until the reader reads the full paper, he/she may not understand the latter part of Abstract.*

We thank the reviewer for these helpful comments. We agree that the key research questions were not sufficiently explicit in the original abstract. The abstract has therefore been revised to clearly state the main objectives of the study, namely the development of an integrated framework combining long-term observations, remote sensing, and hydrological–hydraulic modelling to quantify multi-decadal hydro-sedimentary budgets and analyze the role of floodplains in sediment dynamics in large Amazonian rivers, and to better specify the associated uncertainties and intended applications of the approach.

The revised abstract reads as follows:

Since the early 1970s, the Amazon basin has experienced growing local and global changes, potentially approaching a climatic tipping point in the coming decades. Understanding how these changes affect water and sediment fluxes is essential for assessing the evolution of biogeochemical cycles in the basin. However, due to cost constraints and limited access, conventional hydrological networks struggle to provide the spatial resolution and temporal extent required to accurately quantify water and sediment budgets.

To address this limitation, we develop an integrated framework combining long-term observations, remote sensing, and hydrological–hydraulic modelling (SWAT-Amazon) to quantify multi-decadal hydro-sedimentary budgets and investigate how floodplain inundation controls sediment dynamics in large Amazonian rivers. Focusing on the Ucayali River, a major foreland tributary of the Amazon, this study provides the first long-term hydro-sedimentary balances for the Upper Amazon at sub-basin scale, distinguishing fine sediment fluxes from sand loads.

Results indicate that the Andean Ucayali River exports 455 10^6 t yr^{-1} (± 10%) of suspended sediment (40% sand). Floodplains trap 36% of Andean sediment inputs, with sand accounting for 65% of total deposition, mostly through tectonic subsidence, resulting in a downstream export

of 290 10^6 t yr^{-1} (± 19%) to the Amazon River (26% sand). Floodplain recycling represents an important secondary sediment source (22% of the Ucayali load), while floodplain water storage peaks at 19.1 km^3 (± 20%) in March (38% of discharge). A previously undocumented floodplain-controlled sand sedimentation process occurs during the flooding period, capturing up to 14% [10%, 20%] of the sand flux at peak discharge and partially decoupling sediment transport from water discharge. Although sediment flux estimates carry GLUE-based parameter and retrieval uncertainties of 10–20%, no significant long-term trends were detected in flood duration, discharge, or sediment fluxes over the study period.

These results highlight the central role of floodplain dynamics in regulating water and sediment transport in large Amazonian rivers and demonstrate the value of combining long-term observations, remote sensing, and modelling to quantify hydro-sedimentary budgets in poorly monitored tropical floodplain systems. They also emphasize the need for regionally constrained approaches, as large-scale assessments may overlook key processes and introduce substantial uncertainties.

Introduction

3- ***Section 1.4:*** *What are the key research questions that the authors are trying to address in this research? Role of floodplain in sediment transport? The paper should state 2–3 explicit objectives/hypotheses early (end of Abstract + end of Intro). Kindly add a 1–2 sentence novelty statement distinguishing this from prior Amazon basin sediment studies*

We thank the reviewer for this suggestion. The end of the introduction (Section 1.4) has been revised to explicitly state the main objectives of the study and to clarify its novelty compared with previous Amazon sediment studies:

Revised text:

However, no study has yet combined remote sensing with modelling to investigate sediment dynamics in detail in the Amazon. Building on long-term CZO HyBAm conventional observations, this study proposes an integrated framework that combines field calibration campaigns, satellite remote sensing, and hydraulic–hydrological modelling. It combines satellite altimetry and satellite-derived fine sediment estimates with a hydrological–hydraulic modelling scheme (SWAT-Amazon) to simulate water and sand fluxes. In this model, water and sediments are routed hydraulically, and floodplains are represented using a simplified reservoir approach.

The objectives are to (i) quantify long-term hydro-sedimentary budgets at sub-basin scale in a large Amazonian river, (ii) investigate the impact of floodplain inundation on water and sediment fluxes, thereby identifying key processes controlling sediment transport, storage, and recycling that remain poorly constrained by conventional river monitoring approaches, and (iii) distinguish the respective contributions of fine sediment fluxes, associated with organic matter and pollutant transfer, and sand loads related to river dynamics.

Unlike previous basin-scale sediment studies in the Amazon, which largely relied on sparse gauging networks or large-scale modelling approaches, this framework provides process-based hydro-sedimentary budgets by combining multi-source observations. Given the key role of sediment dynamics in biogeochemical cycles, it also contributes to improving the understanding of the Amazon's role in global material fluxes and of the potential impacts of environmental changes on its hydrology and sediment transport.

Integrative strategy

4- Figure 1 is scattered. How a), b) & c), and d) are linked with each other is not easy to comprehend with the figure. Kindly sync them

The figure caption has been clarified to better explain the logical connection between panels.

Modified figure caption:

Figure 1: General schematic overview of the proposed methodology. Panels (a–c) illustrate the integrative approach. (a) Types of stations. (b) Typical bimodal particle size distribution (PSD) in the large Amazonian rivers, identifying two main size groups: 1- fine sediments that can be monitored by satellite but not modelled; 2- fine sands in graded suspension, invisible to satellites but whose transport capacity can be modelled. (c) Integrated approach combining remote sensing, modelling, and calibration-validation campaigns. (d) SWAT-Amazon, a tailored version of the SWAT model for simulating water and sand fluxes. This modelling framework consists of a Fortran-based executable (SWAT-Amazon.exe) and an R notebook (Run-SWAT-Amazon.Rmd) used for model runs, simulation analysis, interactive visualization, sensitivity analysis and calibration with the SWATrunR package (Schürz et al., 2019).

5- I doubt the assumption of assuming mean diameter for fine sediment in the range of 10-20 micrometer. This because Firstly, are there any records of PSD profile of the basin. Secondly, 10- 20 micrometer sits at the lower range of very fine silt. Thus, underrepresenting coarse silt or large composition of fine sediments in the range of 20- 100 Micrometer. The sediment transport of both the group large differs, with former influenced by flocculation, organic matter, and high suspension. And latter having conventional particle dynamics. To simplify, why don't the authors use at least 3 particle size to better represent the PSD, clay, silt and sand.

We thank the reviewer for this insightful comment. The representation adopted in this study is supported by both field observations and methodological considerations. Indeed, particle-size distribution (PSD) measurements available for suspended sediments in the Ucayali River and other white-water rivers of the Amazon Basin (Fig. CC1.1a) consistently show a predominantly bimodal structure, with a dominant fine mode composed mainly of silts and clay aggregates and a sand mode centred in the very fine sand range. This typical bimodal structure has been documented in several previous studies (e.g. Vanoni, 1979, 1980; Bouchez et al., 2011; Armijos et al., 2016; Santini et al., 2019; Martinelli, 2022). Thus, the mean diameter ranges reported in the manuscript (10–20 µm for fine sediments and 80–120 µm for sands) should be understood as

order-of-magnitude estimates of the mean diameters of these dominant grain-size classes, rather than strict granulometric bounds of the same groups.

The separation between these two dominant size groups is defined by the sieving at 63 µm of the suspended sediment samples (see the new section in the Supplementary Material (Section S2), where the HyBAm monitoring protocol is now detailed). This threshold provides a consistent and reproducible basis for sediment flux measurements across the basin. Therefore, intermediate particle sizes of the PSD continuum, in particular the coarser silts (~20–63 µm), are not neglected, but are represented by one of these two mean diameters.

For the fine sediment group, several previous studies in the Amazon Basin have demonstrated that their concentration and flux can be robustly quantified from satellite reflectance (e.g. Martinez et al., 2015; Espinoza-Villar et al., 2012; Park and Latrubesse, 2014). Despite processes such as flocculation, aggregation with organic matter, or variations in particle properties (shape, density, mineral composition, or Si/Al ratios), the bulk of fine sediments generally behaves at the reach scale as a wash-load fraction. In the remaining subgroup of the coarse silts size (~20–63 µm), as described by the reviewer, **t**wo distinct particle behaviours may occur. First, aggregates, which have a lower density than non-cohesive particles, tend to behave as passive scalars in the flow. A smaller fraction of non-cohesive silt particles with higher density may exhibit graded suspension behaviour; however, their contribution to the uncertainty in fine sediment fluxes is likely negligible compared to the bulk of the fine fraction. Moreover, given the range of shear velocities in large Amazonian rivers (roughly 0.01–0.4 m s^{-1}), the Rouse number of non-cohesive silts close to the 63 µm threshold is expected to remain low (typically ~0.01–0.2 depending on particle density and flow conditions), consistent with a wash-load to fully suspended regime. Therefore, even if a very small clay fraction can be detected in PSD measurements, and even if this subgroup dominates satellite reflectance according to Mie scattering theory, the fine sediment group can be robustly represented as a single wash-load fraction at the scale considered in this study.

From the 63 µm threshold, non-cohesive coarse silts are considered equivalent to very fine sands transported in a graded suspension regime, and represented by a single effective mean diameter in the transport capacity equations. Even if some coarse aggregates may temporarily resist turbulent shear, their contribution to the suspended sand mass remains negligible. For non-cohesive particles, given the quadratic dependence of settling velocity on particle size, the graded suspension regime is rapidly reached, and the uncertainty induced by the small volume of particles behaving closer to wash load than to graded suspension is expected to be negligible.

Overall, field observations indicate a predominantly bimodal PSD structure, suggesting that a separation into three size groups is not warranted in this context. Such a division would increase

model complexity and parameter uncertainty without clear improvement in budget robustness. The two-fraction representation is therefore considered sufficient, with only second-order impacts on simulated fluxes.

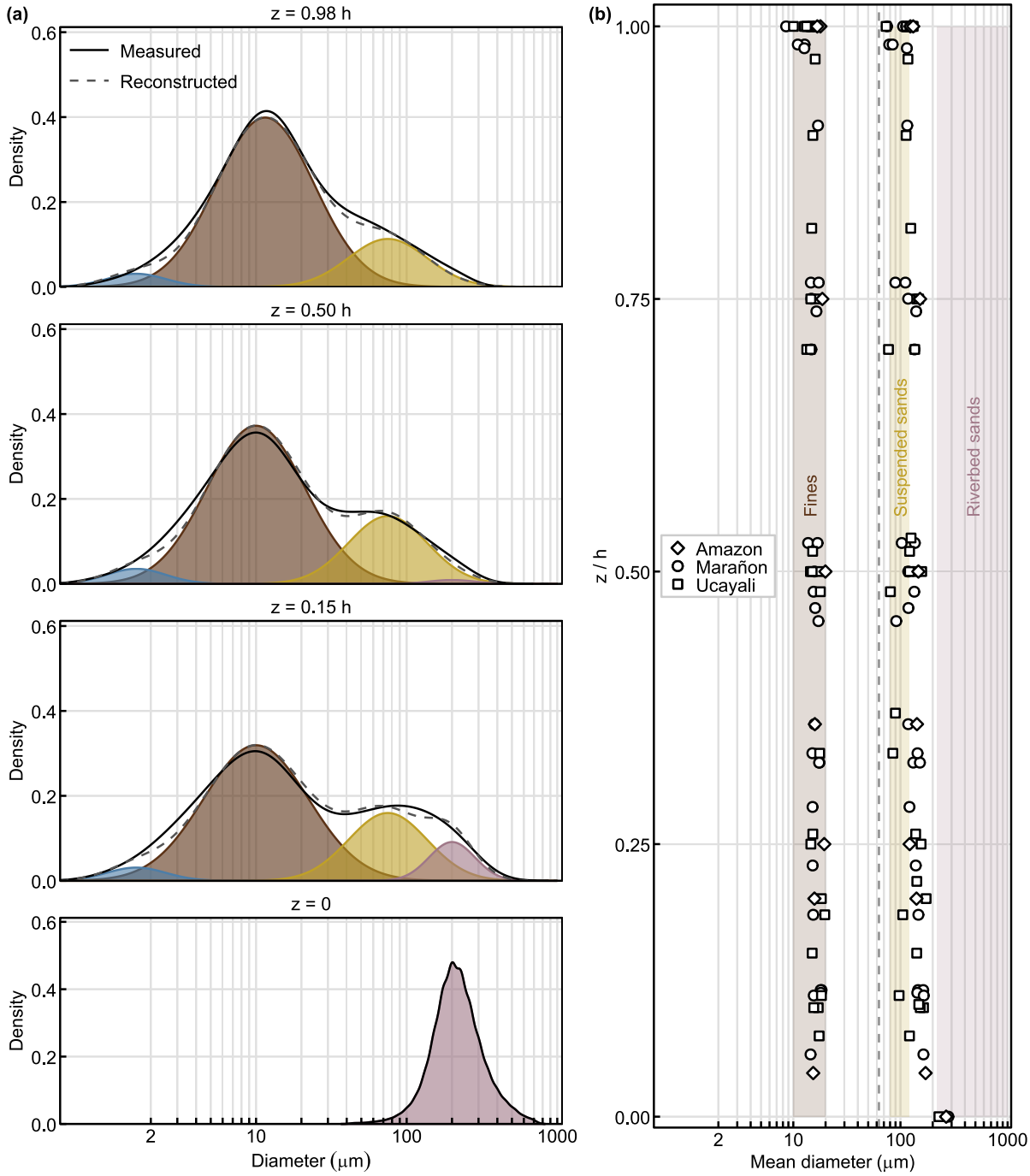

Figure CC1.1: Typical PSD of large white-water rivers in the Amazon Basin (adapted from Santini et al., 2019). (a) Particle-size distributions (PSD) measured in suspension at different relative depths and in the riverbed (Requena station, Ucayali River, 16 March 2015). The grey line shows the measured distribution, while the dashed line represents a reconstructed PSD obtained from a mixture of log-normal components corresponding to clay, silt, and sand fractions. **(b)** Vertical profiles of representative grain diameters for the fine and sand fractions. Symbols correspond to measurements at different stations. The dashed vertical line indicates the conventional 63 µm boundary between silt and sand. This figure illustrates the predominance of two main modes in the PSD, supporting the two-fraction representation adopted in this study.

We have added the following paragraph in Section 2 to clarify this point:

These ranges of mean diameter should be understood as representative values of the dominant PSD modes, rather than strict granulometric bounds. Intermediate particle sizes (e.g. coarse silts in the 20–63 µm range) are not neglected but are included within the fine sediment fraction defined by the 63 µm threshold used in the HyBAm monitoring protocol. Depending on their physical properties, coarse silts may either behave as wash load (e.g. aggregates) or exhibit transport dynamics closer to those of very fine sands when non-cohesive. This supports a first-order bimodal representation of suspended sediment transport at the basin scale, consistent with PSD measurements and vertical grain-size profiles presented in the Supplementary Material (Fig. S1).

6- I feel the current setup may underrepresent sand load in the model. Could you please clarify and provide support?

We respectfully do not find evidence supporting this concern and provide several lines of justification below:

First, in the present framework, suspended sand fluxes are directly constrained by long-term field observations, including cross-sectional sediment sampling, and are therefore not solely determined by the model structure. Model simulations were calibrated and evaluated against these observations and successfully reproduce both the magnitude and temporal variability of measured sand fluxes at the main gauging stations, which indicates that no systematic underestimation of sand load is observed.

Second, from a sediment transport perspective, the use of a representative grain diameter is consistent with the transport formulation adopted in the model. The Camenen and Larson equations are designed to use a representative diameter for the sand fraction and are known to provide robust and reliable estimates under a wide range of hydraulic conditions. Observations of particle-size distributions in the Amazon Basin indicate that suspended sediments exhibit a predominantly bimodal structure, with a clear separation between fine sediments and a single dominant suspended sand mode, which supports the use of a representative sand diameter in the model.

Third, the calibrated representative diameter ($d_s \approx 80$ µm), close to the silt–sand transition, likely integrates part of the coarsest non-cohesive silts, whose hydraulic behaviour approaches that of very fine sands. This suggests that the effective transported fraction represented in the model may be slightly broader than the strict sand class, further reducing the risk of underestimating sand fluxes.

Finally, a secondary coarser mode may exist close to the riverbed under mixed suspension–bedload conditions. However, this fraction is largely associated with bed material and contributes only intermittently to the suspended load, likely through short hydraulic bursts linked to bedform dynamics. In addition, flow velocities decrease in the near-bed region, and this coarser fraction remains concentrated within the near-bed layer (0–0.2h), limiting its contribution to depth-integrated fluxes. Therefore, its contribution to the depth-integrated suspended sand flux is expected to remain minor at the temporal and spatial scales considered in this study.

Taken together, these elements indicate that the two-fraction representation does not lead to an underestimation of sand load but rather provides a physically consistent and observation-constrained approximation of suspended sediment transport at the basin scale.

Conventional data

7- ***Section 3.1:*** *Kindly introduce the term "super stations" in this section as you talk about long data stations, the virtual stations and low data station in the next station.*

The term "super stations" is already introduced and defined in Section 2 (Integrative strategy), where the three types of stations used in the framework (low-data stations, virtual stations, and super stations) are described. No modification to the manuscript was considered necessary.

8- ***Figure 2:***

- The blue color interferes with the color used in the map. Kindly use some other color to mark the stations. Perhaps red. It is not needed to write the years in the map.

*- "The text box details observation periods for water level (**h**) discharge (**Q**), and suspended sediment concentration (**C**)." It is not easy to comprehend.*

- River mapping is not clear

The figure has been revised to improve readability: station markers now include a bolder border to increase visibility, wetland colours have been lightened to improve contrast with other map elements, and river network line width has been increased from 1 to 3 points. The observation periods were retained in the figure as they provide important contextual information on the temporal coverage of the datasets; the abbreviated notation (h, Q, C) is defined in the figure caption, as space constraints preclude spelling out full variable names directly on the map.

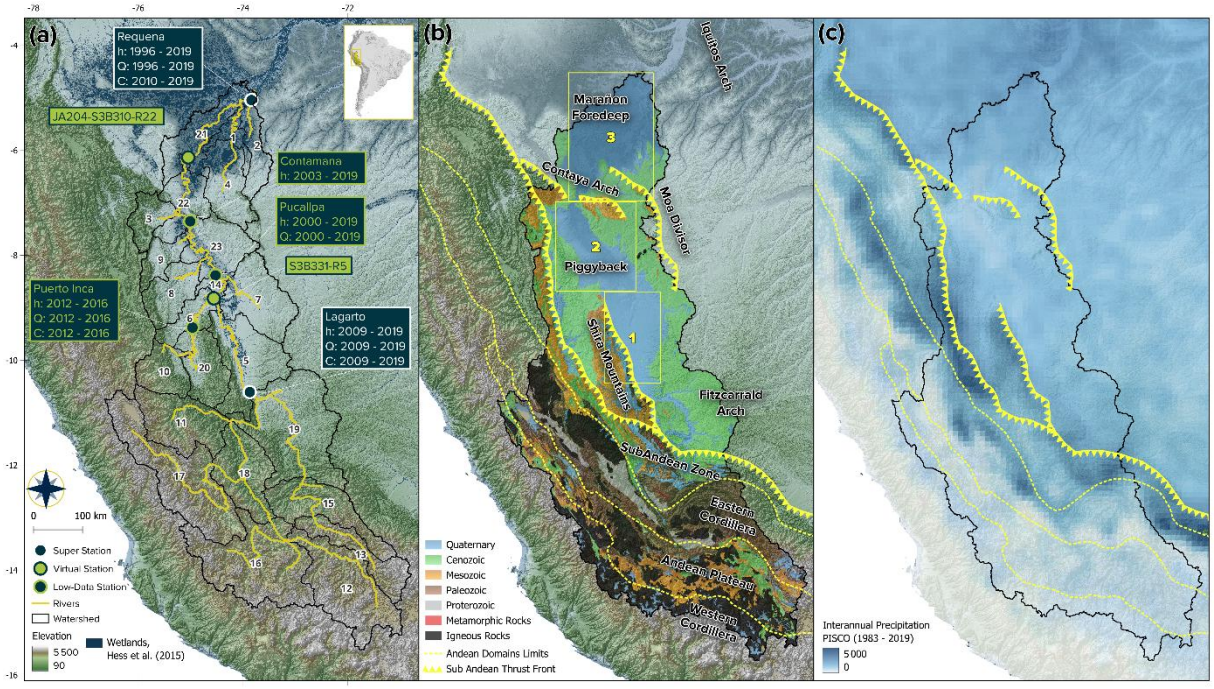

Figure CC1-2: Revised Figure 2.

9- ***Section 3.3.2:*** *Could you please clarify, how total, fine, and sand concentration was determined. Were the samples dry or wet sieved. Was gravimetric filtration used or else?*

10- ***Section 3.3.3:*** *How the vertical gradient in sediment concentration was considered in the field measurements?*

11- *Does the site also have lateral change in the suspended sediment concentration? Was the mean concentration representative vertically and/or laterally? How would author try to convince this?*

This point is clarified in the revised manuscript, and additional details on the measurement protocols are now provided in the **Supplementary Material S6**. A reference to this new supplementary section has been added in **Section 3.1**.

Suspended sediment concentrations were determined using standard methods. After field collection, water samples were transported to the HyBAm laboratory in Lima, where the sand fraction was separated from the fine fraction using a 63 µm sieve following ASTM D3977 procedures. Both fractions were then filtered through 0.45 µm cellulose acetate filters, dried at 50 °C for 24 h, and weighed to determine sediment concentrations.

Vertical concentration gradients were measured during dedicated field campaigns conducted at the HyBAm gauging stations. Suspended sediment samples were collected using a point-sampling method along three verticals across the river section. The sampling protocol evolved during the monitoring period. Early campaigns typically included three samples per vertical (near-bed, mid-depth, and near-surface). Later campaigns increased the vertical resolution, with sampling levels typically located at 0.98 h, 0.75 h, 0.5 h and 0.25 h above the riverbed (occasionally including additional levels). As a result, between 9 and 15 samples per cross-section were generally used to estimate cross-sectionally averaged concentrations of total suspended sediments, fine sediments and sands.

Lateral variability was addressed by sampling along three verticals distributed across the channel, allowing cross-sectional mean concentrations to be derived from multiple vertical profiles. These measurements were used to establish empirical relationships between surface index concentrations and cross-sectionally averaged concentrations.

The dataset used in this study relies on more than two decades of systematic monitoring within the HyBAm observatory and 82 dedicated campaigns conducted between 2003 and 2019, representing an unprecedented hydro-sedimentary observational effort for large Amazonian rivers.

Model development

12- ***Section 4.1:*** *Please explain how reach-wise inputs are derived and how this can be reproduced in other basins. Additionally, given floodplain activation is event-scale, please discuss how a daily routing framework affects peak timing/magnitude and floodplain exchange, and whether it biases sediment flux estimation.*

Reach-wise inputs follow the standard SWAT routing framework, where each reach receives routed inflows from upstream reaches through the predefined reach network structure. In the sand routing module, this incoming routed sediment flux represents the sediment supply entering the reach. The effective sand flux exported from the reach is subsequently recalculated according to the local transport capacity. The resulting sand dynamics are governed by the mass balance formulated at the reach scale (Eq. 25), which accounts for riverbed erosion when transport capacity exceeds the incoming supply, as well as exchanges with the floodplain through overbank deposition and floodplain-derived inputs. Consequently, the exported sand flux is not imposed directly by the incoming routed flux but results from the balance between sediment supply, transport capacity, and floodplain–channel exchanges as described in Section 4.2.4.

We add to the manuscript the following sentence (section 4.2.5):

The routed sediment inflow provided by the SWAT-Amazon reach network represents the sediment supply entering each reach and is incorporated into the reach-scale sand mass balance described in Eq. (25).

Regarding the routing time step, the reviewer raises an important point. Although SWAT-Amazon simulations are reported at a daily time step, the routing equations for water and sediments are internally solved at a sub-daily time step dynamically determined according to the Courant–Friedrichs–Lewy (CFL) stability condition. This ensures that the computational time step remains smaller than the time required for the fastest wave allowed by the governing equations to travel across a reach, thereby maintaining numerical stability.

Consequently, the routing module iterates internally over intraday time steps to satisfy the CFL criterion, and the daily outputs correspond to the model state at the end of each simulation day. This adaptive routing ensures that flood-wave propagation and attenuation remain correctly resolved despite the daily reporting of model outputs. In large low-gradient Amazonian rivers, where flood-wave propagation occurs over time scales of several days to weeks, the daily reporting time step therefore does not significantly affect peak timing, floodplain exchange, or sediment flux estimates.

We added the following clarifications in the manuscript:

Section 4.1.3:

Although SWAT simulations are reported at a daily time step, river routing is internally solved at a sub-daily time step dynamically determined by the Courant–Friedrichs–Lewy (CFL) stability condition. The routing module automatically adjusts the internal time step to satisfy the CFL criterion and ensure numerical stability (Bates et al., 2010). Daily outputs therefore correspond to the model state at the end of each simulation day.

Section 4.2.5:

As for water routing, sand routing is internally solved at a sub-daily time step to satisfy the CFL stability condition.

13- ***Section 4.2****: The approach relies on an effective bed-material sand diameter; please justify its spatial representativeness (Amazon has strong upstream–downstream contrasts) and provide a sensitivity test to plausible diameter ranges (or consider a simple multi-class sand representation).*

In the Upper Amazon, suspended sand transported by the mainstem largely originates from Andean sources (e.g. Baby et al., 2009). Downstream of the Sub-Andean zone, strong hydraulic sorting and long transport distances tend to produce relatively homogeneous sand grain-size characteristics along the lowland reaches. Available particle size measurements from the HyBAm observatory indicate a consistent sand mode along the Ucayali mainstem, with mean diameters of approximately 260 µm at Lagarto, 243 µm at Pucallpa, and 228 µm at Requena.

This relatively limited variability in the Ucayali is consistent with previous observations in Amazonian rivers, where riverbed and suspended sand fractions tend to remain relatively uniform along long lowland reaches due to hydraulic sorting and mixing processes (e.g. Guyot et al., 1999). More pronounced grain-size contrasts may occur further downstream in the central and lower Amazon due to inputs from tributaries draining distinct geological units, such as the Brazilian and Guiana shields (e.g. the Madeira or Negro rivers; Nordin et al., 1980), or locally in systems influenced by major geomorphological structures such as the Pastaza megafans in the Marañón basin. However, such contrasts remain limited within the Upper Amazon basin considered here.

We agree that the sensitivity to the representative sand diameter should be assessed. In the present study, this aspect is already addressed through the global Sobol sensitivity analysis presented in Section 6.1.2. The parameter d_b is identified as one of the most influential parameters controlling sand routing. To ensure a robust exploration of model sensitivity, a

deliberately wide parameter range (180–300 µm) was explored, substantially exceeding the variability indicated by the available field measurements. Interestingly, the calibrated values converge toward the narrower range suggested by observations along the Ucayali mainstem (Lagarto, Pucallpa, and Requena), which provides additional support for the physical consistency of the adopted parameterization.

Future developments of the framework could incorporate a multi-class sand representation if required. However, given the relatively narrow range of sand diameters observed in the Ucayali system and the scale of the present modelling framework, the use of a representative diameter provides a parsimonious and physically consistent representation of sand transport.

Finally, the transport formulation used in the model (Camenen and Larson, 2005, 2008) relies on a representative grain diameter for the sand fraction, which supports the use of a single bed-material sand diameter in the present framework. The formulation could also be applied to several sand classes if required, but given the limited variability observed in the Ucayali system, such an extension was not considered necessary in the present study.

Results and discussion

14- I recommend strengthening the separation of calibration and validation (including timing metrics), making 'not shown' key comparisons available in Supplement, and adding a clearer treatment of uncertainty.

This is an important point that deserves clarification. We acknowledge that the wording in the original manuscript was ambiguous regarding the distinction between calibration and the reported performance periods, which may have given the impression that calibration and validation fully overlap. We also acknowledge that the available monitoring dataset limits the definition of a fully independent validation period for sand routing in the conventional sense. Nevertheless, model robustness is supported by: (i) consistent performance metrics over periods outside the calibration window (Table CC1.1 and CC1.2), with reduced performance at Lagarto attributable to increased forcing and observational uncertainties rather than model degradation; and (ii) the comparison with direct gauging measurements (Fig. CC1.3), which constitute a quasi-independent validation as these observations are not explicitly used in calibration. Residual uncertainty is further addressed through the GLUE analysis presented in Section 6. **We address a complete response below:**

For water routing (Section 5.1), the 2000–2016 period does not correspond to a calibration period, but rather to a common observation window used to compute performance metrics across stations, depending on data availability. Model calibration for water discharge (Table CC1-1) was conducted over the 2010–2015 period (except for Puerto Inca, for which the period was 2012–2014), selected based on data quality and availability. Importantly, calibration does not rely on the optimization of a single time series, but on multiple hydraulic diagnostics, including water levels (Fig. 4a), velocities (Fig. 4b), and stage–discharge and stage–velocity relationships (Fig. 5). Moreover, independent hold-out periods were used to further evaluate model performance (Table CC1-1), showing consistent performance metrics across all periods.

Table CC1-1. Discharge model performance over calibration and independent validation periods. Calibration was performed over 2010–2015 (except for Puerto Inca, for which the period was 2012–2014), while independent validation periods correspond to time windows outside the calibration interval, selected based on data availability at each station. Metrics are computed using observed daily discharge records.

Station	Period type	Period	N (days)	NSE	KGE	PBIAS (%)
Lagarto	Independent validation	01/2009–12/2009	287	0.9	0.94	-1.3
Lagarto	Calibration	01/2010–12/2015	2191	0.89	0.94	-1.7
Lagarto	Independent validation	01/2016–12/2019	1162	0.84	0.90	-1.7
Puerto Inca	Calibration	09/2012–08/2014	730	0.73	0.86	-1.1
Puerto Inca	Independent validation	09/2015–08/2016	731	0.69	0.83	-7.1*
Pucallpa	Independent validation	01/2000–09/2009	3653	0.94	0.94	-4.8
Pucallpa	Calibration	01/2010–12/2015	2161	0.92	0.89	1.1
Pucallpa	Independent validation	01/2016–09/2019	1389	0.91	0.94	-1.2
Requena	Independent	01/2000–09/2009	3653	0.92	0.93	-3.8
Requena	Calibration	01/2010–12/2015	2191	0.9	0.95	0.1
Requena	Independent validation	01/2016–12/2019	1461	0.85*	0.83*	-9.5*

** Lower performance during these periods is associated with reduced quality of observed discharge data and/or precipitation inputs.*

For sand routing (Section 5.3), the 09/2009–08/2015 period shown in Fig. 6 corresponds to the interval with the most reliable and dense observations, as explained in the manuscript, although data quality remains heterogeneous depending on the period and sampling protocols. This period was therefore selected for model calibration. However, outside this interval, uncertainties in sand flux observations (derived from surface sediment concentration monitoring and sediment gauging; see Supplementary Material S1) increase, which precludes the definition of a robust and fully independent validation period. In particular, at Lagarto, the NSE for the validation period decreases to 0.45 (Table CC1-2). When excluding flood events, the NSE remains high (0.80), indicating robust model performance under low- and intermediate-flow conditions. The lower performance is therefore mainly associated with rapid Andean flood events and related uncertainties in rainfall forcing.

Table CC1-2. Sand routing model performance over calibration and independent validation periods. Calibration was performed over 2009–2015, while independent validation periods correspond to time windows outside the calibration interval, selected based on data availability at each station. Metrics are computed using observed sand flux at surface concentration sampling time step.

Station	Period type	Period	N (days)	NSE	KGE	PBIAS (%)
Lagarto	Calibration	09/2009–08/2015	323	0.80	0.87	7.9
Lagarto	Independent validation	09/2015–08/2018	181	0.45	0.60	27.7
Requena	Calibration	09/2009–08/2015	404	0.86	0.92	-2.3
Requena	Independent validation	09/2015–08/2018	77	0.80	0.70	-24.0*

Overall, reduced performance after the calibration period reflects both increased observational and forcing uncertainties rather than a degradation of model performance.

Therefore, to further evaluate model performance, we performed an additional comparison between daily simulations (discharge and suspended sand fluxes) and direct gauging measurements over the 09/2009–08/2016 period (Fig. CC1-3). This period starts with the implementation of improved sediment sampling protocols during field measurements, while

simulations after 2016 were not considered due to increased uncertainties in rainfall inputs at Requena (Fig. 3d). This comparison bypasses rating curves and provides a more direct evaluation of model performance. Although not fully independent, it offers a complementary and partially independent validation, as direct gauging measurements are not explicitly used in the calibration, even though they contribute to the construction of rating curves over the full observation period. The agreement between simulations and gauging measurements supports model robustness, with bias-corrected R^2 (bR^2) values of 0.90 for discharge and 0.88 for sand fluxes (Fig. CC1-3).

The corresponding results have been added in the Supplementary Material (Fig. S9) and are briefly discussed in the revised manuscript. We have also revised the text to explicitly distinguish between calibration, performance assessment, and the different sources of observational constraints.

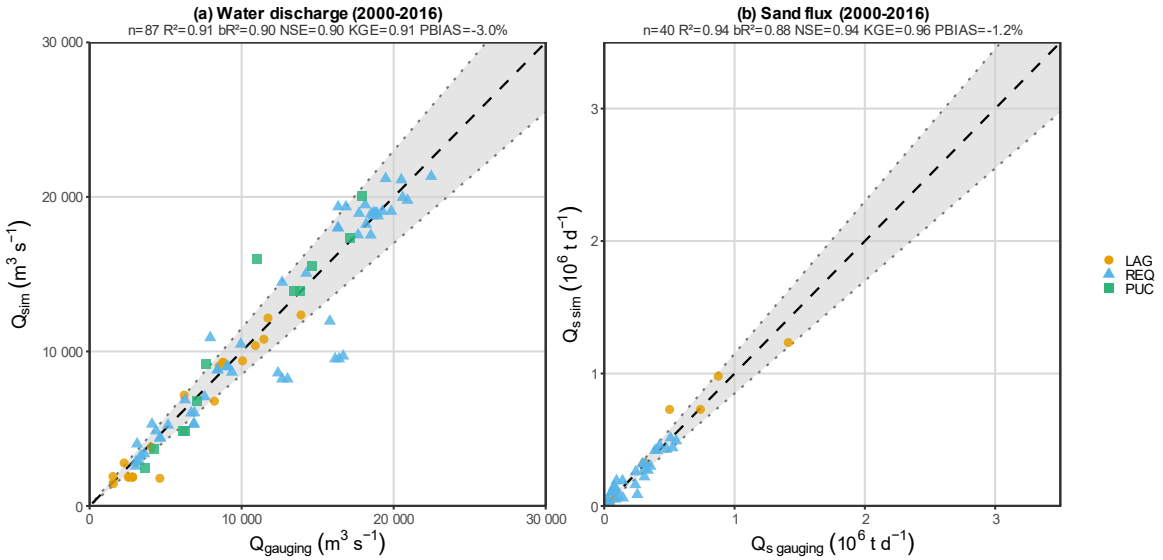

Figure CC1-3: Comparison of simulated and measured river discharge (a) and suspended sediment flux (b) at gauging stations in the Ucayali basin. Simulated values (Q_{sim}, $Q_{s\ sim}$) correspond to daily outputs from the SWAT-Amazon model extracted at the dates of field measurements. Observed values ($Q_{gauging}$, $Q_{s\ gauging}$) are direct field measurements, independent of rating curve derivation. The dashed line indicates the 1:1 relationship, and the shaded band represents the ±15% envelope. The Puerto Inca station was excluded, as its sub-daily flood pulse dynamics are not adequately captured at the model's daily time step. Axis ranges reflect the observed variability of Q and Q_s across all stations. Statistical indicators are computed over all stations pooled: n = number of paired observations; R^2 = coefficient of determination; bR^2 = bias-corrected coefficient of determination; NSE = Nash–Sutcliffe efficiency; KGE = Kling–Gupta efficiency (Gupta et al., 2009); PBIAS = percent bias.

The manuscript was revised as follows:

Section 5.1

For water discharge, model calibration was performed over the 2010–2015 period based on multiple hydraulic diagnostics (water levels, velocities, and stage–discharge relationships). Model performance was further evaluated using independent hold-out periods and direct comparisons with gauging measurements that bypass rating curves, as detailed in the Supplementary Material (Sx). These complementary evaluations show consistent performance across periods and support the temporal robustness of the model. At a daily time step, SWAT-Amazon simulations at the watershed outlet show excellent performance (Fig. 3d): NSE (Nash-Sutcliffe Efficiency) = 0.92, KGE (Kling–Gupta efficiency) = 0.95, PBIAS (Percent Bias) = -1.8%, LogNSE (NSE on the logarithms of the series) = 0.92 over the 2000–2016 observation period used here for inter-station evaluation (see Moriasi et al. (2007) for details on these metrics).

Section 5.3

Calibration focused on the 09/2009–08/2015 period, when sediment monitoring protocols were enhanced, including higher sampling frequency at Requena between 11/2012 and 06/2013, where one sample was collected each two days plus three sampling repetitions each ten days. Beyond, sampling was conducted at five-day intervals during the wet period between July 2013 and September 2015. Additionally, the concentration gaugings were performed in all sites with a higher number of samples collected throughout the cross-section, particularly in the first half of the water column, to ensure more accurate sand concentration calculations. Outside this interval, uncertainties in sand flux observations increase, which complicates the definition of a robust and fully independent validation period (see Supplementary Material S9 for a summary of model performance across periods). The lower performance outside the calibration period primarily reflects uncertainties in rainfall forcing and observations, particularly during rapid Andean flood events, rather than a degradation of model performance. To further evaluate model performance, direct comparisons between simulations and gauging measurements were performed (Supplementary Material, Fig. S9). These complementary evaluations support the temporal robustness of the model despite observational limitations and data heterogeneity.

Supporting analyses and details are now provided in the Supplementary Material (S9).

15- The Sobol analysis is performed on 1 sub-basin. Would the authors clarify how this is representative and ideally include at least one additional contrasting sub-basin.

The Sobol sensitivity analysis was conducted on the most downstream sub-basin of the model domain, which is representative of the lowland floodplain system analyzed in this study. This reach is characteristic of large Amazonian rivers, with a very wide floodplain, a predominantly unimodal annual flood pulse, and frequent overbank flooding except during extreme drought years. Because the same routing equations and similar hydraulic conditions govern water and sediment transport along the lowland mainstem, comparable parameter sensitivities would be expected in other lowland sub-basins. Extending the analysis to all sub-basins would therefore not modify the physical interpretation of parameter sensitivities, given the hydraulic homogeneity of the lowland reaches considered here.

16- In sand routing sensitivity, I would like to clarify whether the PSD is stationary or seasonally varying.

In the present modelling framework, the particle size distribution (PSD) of the sand PSD is assumed to remain stationary over the seasonal cycle.

The manuscript was clarified accordingly in section 6.1.2:

The sensitivity analysis, using the parameter set ($\beta_s, C_{bk}, d_b, d_s, \eta, k_{ch}$), shows that d_b is the most critical calibration parameter (Fig. 9d, f), as previously highlighted by Fagundes et al. (2023). In the present framework, d_b and d_s are assumed to remain constant over the seasonal cycle.

17- They state discharge calibration is insensitive to and under some conditions. kindly specify it is sensitive to which regimes (low flow vs floodplain-active) and show quantitatively.

The comment likely refers to Section 6.2. We agree that the original wording was not sufficiently explicit and may have suggested a regime-dependent sensitivity. This was not our intention.

In the hydraulic routing method (Section 2.4), the influence of channel width B and Manning roughness n on simulated discharge Q is strongly limited by compensating hydraulic effects. As explained in the manuscript, reducing n increases flow velocity u and discharge Q (Eq. 5), which reduces the water volume stored in the reach and therefore water depth h (Eq. 1). The resulting decrease in cross-sectional area leads to a proportional reduction of u and Q. This compensation mechanism limits the influence of B and n on routed discharge, while these parameters remain important for reproducing water depth and flow velocity.

The manuscript has been revised (section 6.2) to clarify this point and avoid any ambiguity:

In the hydraulic routing method (Section 2.4), compensating hydraulic effects limit the influence of B *and* n *on routed discharge. Reducing* n *increases flow velocity* u *and discharge* Q *(Eq. 5), which decreases the water volume stored in the reach and therefore* h *(Eq. 1). The resulting decrease in cross-sectional area leads to a proportional reduction in* u *and* Q*.*

18- The authors argue that sand resuspension is driven by floodplain recession and banks/bars erosion and that bed erosion is negligible because allowing bed erosion does not fit peaks as well. Isn't it a strong process statement to make, just based on calibration.

We agree that such an interpretation should not rely solely on calibration results. The manuscript has therefore been revised to moderate the wording and clarify that the calibration outcomes only provide indirect support for the dominant role of C_{bk} over K_{bed}.

In the modelling framework, K_{bed} controls the susceptibility of the riverbed to erosion when the simulated sand transport capacity exceeds the available sand load in the reach. As clarified in Section 4.2.4, this parameter was calibrated together with the other sediment routing parameters and therefore allows the potential contribution of bed erosion to be tested within the model structure. The calibration experiments indicate that when bed erosion is activated ($K_{bed} > 0$), the simulated sand transport capacity can rapidly entrain bed material once the transport threshold is exceeded, producing sharp and abrupt peaks in simulated sand flux. These peaks occur earlier and appear more intense than the resuspension signal observed in the measurements. In contrast, adjusting the parameter C_{bk}, which represents sediment inputs associated with bank and bar erosion during floodplain recession, produces a smoother and more progressive increase in sand concentration that better reproduces the secondary peaks observed during the falling limb of the hydrograph.

In addition, the calibrated values of K_{bed} remain very small ($K_{bed} \ll 1$), suggesting that the contribution of riverbed erosion to the simulated suspended sand flux in the main stem is limited under the conditions represented in the model. These results therefore suggest a dominant contribution of floodplain-related recycling processes to the observed resuspension signal, although a minor contribution from riverbed erosion cannot be fully excluded.

Revised text (Section 6.3):

Lastly, the distinction between E_{bk} *and* E_{bed} *is partly supported by the calibration experiments. When bed erosion is activated (*$k_{bed} > 0$*), the model tends to generate rapid and abrupt peaks in simulated sand flux once the transport capacity exceeds the available sand load, whereas adjusting*

C_{bk} *produces a smoother and more progressive increase in sand concentration that better reproduces the secondary peaks observed during the recession phase. The calibrated values of* K_{bed} *remain very small (*$K_{bed} \ll 1$*), suggesting that riverbed erosion contributes only weakly to the simulated suspended sand flux in the main stem, although a minor contribution cannot be fully excluded.*

Additional references

Armijos, E., Crave, A., Espinoza, R., Fraizy, P., Dos Santos, A. L. M. R., Sampaio, F., De Oliveira, E., Santini, W., Martinez, J. M., Autin, P., Pantoja, N., and Oliveira, M.: Measuring and modeling vertical gradients in suspended sediments in the Solimões/Amazon River, Hydrol. Process., 31, 654–667, https://doi.org/10.1002/hyp.11059, 2016.

Bouchez, J., Métivier, F., Lupker, M., Maurice-Bourgoin, L., Perez, M., Gaillardet, J., and France-Lanord, C.: Prediction of depth-integrated fluxes of suspended sediment in the Amazon River: particle aggregation as a complicating factor, Hydrol. Process., 25, 778–794, https://doi.org/10.1002/hyp.7868, 2011.

Guyot, J. L., Jouanneau, J. M., and Wasson, J. G.: Characterisation of river bed and suspended sediments in the Rio Madeira drainage basin (Bolivian Amazonia), Journal of South American Earth Sciences, 12, 401–410, https://doi.org/10.1016/S0895-9811(99)00020-8, 1999.

Nordin, C. F., Meade, R. H., Curtis, W. F., Bosio, N. J., and Landim, P. M. B.: Size distribution of Amazon River bed sediment, Nature, 286, 52–53, https://doi.org/10.1038/286052a0, 1980.

Park, E. and Latrubesse, E. M.: Modeling suspended sediment distribution patterns of the Amazon River using MODIS data, Remote Sensing of Environment, 147, 232–242, https://doi.org/10.1016/j.rse.2014.03.013, 2014.

Vanoni, V. A.: Sediment studies in the Brazilian Amazon River basin, World Meteorological Organization, Geneva, Switzerland, 1979.

Vanoni, V. A.: Sediment studies in the Brazilian Amazon River basin, Report KH-P-168, W.M. Keck Laboratory of Hydraulic and Water Resources, California Institute of Technology, Pasadena, California, USA, 1980.