remote sensing

MDPI

Article

Vertical Accuracy of Global DEMs in the Pantanal Floodplain and Implications for Flood Inundation Modelling

Larisse Rodrigues de Jesus [1], Rejane Ennes Cicerelli [1,2], Guilherme Gomes Pessoa [2], Leonardo Zandonadi Moura [3], Nadja Gomes Machado [4], Edson Eyji Sano [1,5] and Tati de Almeida [1,*]

[1] Graduate Program in Applied Geosciences and Geodynamics, Institute of Geosciences, University of Brasília, Brasília 70910-900, Brazil; larisseerodrigues@gmail.com (L.R.d.J.); rejaneig@unb.br (R.E.C.); edson.sano@embrapa.br (E.E.S.)
[2] Institute of Geosciences, University of Brasília, Brasília 70910-900, Brazil; guilherme.pessoa@unb.br
[3] Graduate Program in Environmental Technology and Water Resources, University of Brasília, Brasília 70910-900, Brazil; lzmoura@unb.br
[4] Federal Institute of Mato Grosso, Cuiabá 78043-400, Brazil; nadja.machado@ifmt.edu.br
[5] Embrapa Cerrados, Planaltina 73310-970, Brazil
* Correspondence: tati_almeida@unb.br

Highlights

What are the main findings?

- GEDTM, FABDEM, and ANADEM showed the highest vertical accuracy among the evaluated DEMs.
- DEMs with high vertical accuracy can still produce large variations in simulated inundation extents, especially under low-water conditions.

What are the implications of the main findings?

- DEM selection for flood inundation modelling should account for vertical accuracy, terrain morphology, and land cover characteristics, especially in low-gradient floodplains such as the Pantanal.
- HEC-RAS-simulated inundation extents and Sentinel-2-derived inundation extents should not be considered directly equivalent.

Abstract

Floodplains are highly sensitive to small variations in terrain elevation, making accurate digital elevation models (DEMs) essential for hydrological modelling and flood mapping. However, the performance of global DEMs in tropical floodplains remains poorly evaluated. This study assessed the vertical accuracy of 11 DEMs in the Brazilian Pantanal floodplain and the influence of DEM selection on simulated inundation extents. DEM-derived elevations were compared with ICESat-2 ATL08 observations and multi-source ground control points using mean error (ME), mean absolute error (MAE), and root mean square error (RMSE). The three DEMs with the lowest elevation errors (ANADEM, FABDEM, and GEDTM), together with a Sentinel-1-derived DEM, were subsequently used in HEC-RAS simulations under three contrasting hydrological scenarios. Simulated inundation extents were compared with independent Sentinel-2-derived surface water masks. Spatial agreement between simulated inundation extents and Sentinel-2-derived water masks was relatively low. The DEM with the highest vertical accuracy did not produce the highest spatial agreement with satellite-derived inundation extents. These findings demonstrate that conventional vertical accuracy metrics alone are insufficient to determine DEM suitability for flood inundation modelling in large tropical wetlands.

Check for updates

Academic Editors: Dehua Zhu, Yunqing Xuan, Dingzhi Peng and Gaoxu Wang

Received: 28 August 2026
Revised: 9 November 2026
Accepted: 12 November 2026
Published: 28 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: digital elevation model; floodplain topography; HEC-RAS; hydrological modelling; ICESat-2 altimetry; Pantanal wetlands; vertical accuracy

1. Introduction

Floodplains and wetlands are among the most dynamic ecosystems on Earth, where hydrological processes, geomorphological characteristics, and ecological interactions influence the spatial and temporal distribution of surface water. These environments provide important ecosystem services, including water storage, biodiversity conservation, and climate regulation. Because of their ecological importance, several wetlands are protected under international agreements, such as the Ramsar Convention on Wetlands [1]. The Pantanal, one of the world's largest tropical wetland systems, is characterized by seasonal flood pulses that regulate hydrological connectivity and influence the distribution of aquatic and terrestrial habitats [2]. This highly dynamic ecosystem supports exceptional biodiversity and plays a key role in regional hydrological regulation [3]. However, wetlands worldwide, including the Pantanal, have experienced increasing environmental degradation in recent decades due to land use and land cover (LULC) changes, hydrological alterations, and climate variability [2]. In this context, remote sensing has become an important tool for monitoring wetland ecosystems, enabling large-scale observations of surface water dynamics and flooding processes that are difficult to capture through conventional field measurements [2–6].

The Pantanal is one of the world's largest and most ecologically significant tropical wetland systems, covering approximately 150,000 km^2 across Brazil, Bolivia, and Paraguay. The region is characterized by seasonal flood pulses that regulate hydrological connectivity and influence the distribution of aquatic and terrestrial habitats. These dynamics create a complex mosaic of rivers, shallow channels, floodplain lakes, and seasonally inundated areas. The alternation between wet and dry periods shapes habitat distribution and regulates ecological processes throughout the floodplain [3,7]. Given the low topographic gradients across much of the Pantanal, even small variations in terrain elevation can significantly influence the spatial extent and seasonal dynamics of inundation [3,8,9].

Understanding flood dynamics in such environments requires accurate topographic data. Digital elevation models (DEMs) represent terrain elevation as gridded datasets and are widely used in geomorphological, hydrological, and environmental modelling applications. They provide key topographic information for hydrological simulations, terrain analysis, and floodplain inundation mapping [10,11]. However, the reliability of these applications depends on the vertical accuracy of the data [12–14].

Over the past few decades, advances in remote sensing technologies have enabled the development of numerous global DEMs derived from radar interferometry, optical stereoscopy, and multi-source data integration, including SRTM, NASADEM, ASTER GDEM, AW3D30, TanDEM-X, Copernicus DEM, FABDEM, GEDTM, and ANADEM. These datasets support terrain analysis in regions where detailed topographic surveys are unavailable. However, the vertical accuracy of global DEMs varies with terrain characteristics, LULC, and sensor properties, highlighting the importance of evaluating their accuracy before applying them to hydrological modelling and flood inundation mapping [12,14].

Remotely sensed elevation measurements often include contributions from vegetation canopies, potentially leading to systematic overestimation of ground elevations in densely vegetated wetlands [14,15]. In low-gradient environments such as floodplains, even small DEM elevation errors can distort flow pathways, alter hydrological connectivity, and affect the delineation of inundation boundaries [5,12,13]. Consequently, assessing DEM vertical

accuracy is essential for determining its suitability for hydrological modelling and flood inundation mapping.

Several studies have demonstrated that globally available DEMs can support large-scale flood simulations and inundation mapping, particularly in regions where detailed topographic surveys are unavailable [12]. However, their performance depends on terrain characteristics, spatial resolution, vertical accuracy, and the representation of hydrologically relevant topographic features [12,13,15]. Therefore, DEM selection should consider not only elevation errors but also the ability of each product to represent floodplain morphology and support reliable simulations of water depth and flood propagation [13].

Previous studies have shown that DEM-based hydrodynamic models can reproduce large-scale floodplain dynamics when appropriate modelling approaches are applied. For example, hydrodynamic simulations in the Pantanal successfully reproduced seasonal flood pulses and spatial inundation patterns over 140,000 km^2 using DEM-derived terrain information [11]. These findings demonstrated the potential of hydrodynamic modelling for simulating regional-scale flood inundation, despite the uncertainties related to DEM-derived elevations in low-gradient environments.

In this context, the present study provides a comprehensive assessment of DEM performance in a sector of the Pantanal floodplain by integrating three complementary approaches that have rarely been combined in previous studies of tropical wetlands. First, we assess the vertical accuracy of 11 DEM products, including recently developed terrain-corrected datasets (GEDTM, ANADEM, and FABDEM) that have not been systematically evaluated in the Pantanal, using two independent reference datasets: ICESat-2 ATL08 satellite altimetry observations and multi-source ground control points. Second, we investigate variations in DEM vertical accuracy across different LULC classes, especially considering the influence of vegetation on remotely sensed elevation estimates. Third, we extend the assessment beyond conventional point-based accuracy metrics by incorporating selected DEMs into the Hydrologic Engineering Center's River Analysis System (HEC-RAS) hydrodynamic simulations and comparing the resulting inundation extents with independent Sentinel-2-derived surface water masks under contrasting hydrological conditions. This integrated approach allows us to investigate whether conventional vertical accuracy metrics adequately reflect DEM suitability for flood inundation modelling. By examining the relationship between DEM vertical accuracy and flood inundation simulation results in a low-gradient wetland, this study provides practical insights into DEM selection for hydrodynamic modelling in large tropical floodplains, where subtle topographic variations can strongly influence hydrological connectivity and inundation dynamics.

2. Materials and Methods

2.1. Study Area

The study area is located along the Paraguay River floodplain within the Pantanal biome, encompassing portions of the municipalities of Corumbá and Ladário in Mato Grosso do Sul State, and Cáceres and Poconé in Mato Grosso State, Brazil, as well as adjacent areas of Bolivia. It covers approximately 5060 km^2 (Figure 1).

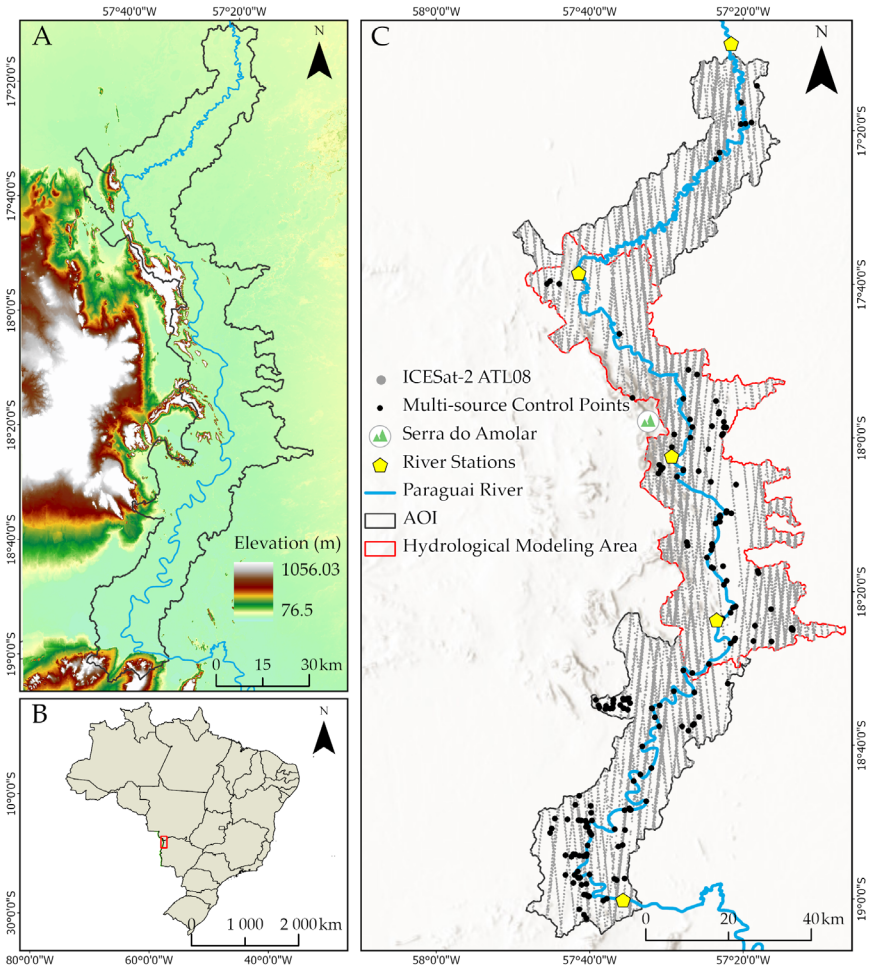

Figure 1. Location and characteristics of the study area in the Upper Paraguay River Basin (Pantanal, Brazil). (**A**) Elevation map; (**B**) location of the study area in Brazil (red rectangle); and (**C**) detailed view of the study area along the Paraguay River. The black outline delineates the study area in the Paraguay River floodplain. Black points indicate ICESat-2 observations used for digital elevation model (DEM) vertical accuracy assessment, while the Paraguay River channel is shown in cyan. Data sources: National Water and Sanitation Agency (ANA) and Brazilian Institute of Geography and Statistics (IBGE).

The Pantanal is one of the world's largest wetland systems and is characterized by complex hydrological dynamics driven by seasonal flood pulses [2,5,7,8]. The Paraguay River serves as the main hydrological axis of the basin, strongly influencing the spatial distribution of inundation across the floodplain [10]. Seasonal variations in river discharge result in alternating periods of flooding and drying, which regulate ecological processes and habitat distribution throughout the region [16,17]. Gentle topographic gradients and subtle microtopographic variations strongly influence hydrological connectivity among river channels, floodplain lakes, and vegetated wetlands. Consequently, accurate terrain representation is essential for reliable hydrological analysis and flood inundation modelling in Pantanal [2,11].

LULC in the study area is dominated by natural floodplain vegetation characteristic of the Pantanal biome. Grasslands and seasonally flooded savannas are the predominant LULC types, accounting for 40–50% of the study area, followed by riparian forests and other woody vegetation formations occurring on slightly elevated terrain (20–25%). Wetlands dominated by aquatic macrophytes cover 10–15% of the area, while permanent water bodies, including the Paraguay River and connected floodplain lakes, account for 10–12%. Anthropogenic land use is relatively limited, consisting of areas used for extensive cattle ranching (5–10%),

whereas urban areas occupy less than 1% [18]. These LULC patterns reflect the strong influence of seasonal flooding on vegetation distribution, hydrological connectivity, and habitat organization across the Pantanal floodplain [2,18].

2.2. Digital Elevation Models

To assess terrain representation across the study area, ten DEM products were evaluated. These datasets were derived from different acquisition technologies and processing approaches, including radar interferometry, optical stereoscopy, and multi-source data integration. The evaluated products comprised SRTM GL1, ALOS PALSAR, FABDEM, Copernicus DEM, NASADEM, ASTER GDEM, AW3D30, TanDEM-X, GEDTM, and ANADEM (Table 1) [19,20]. Their nominal horizontal spatial resolutions range from 10 to 30 m. These DEMs are widely used in geomorphological and hydrological applications.

Table 1. Main characteristics of the elevation datasets used in this study. Rows 1–10 correspond to publicly available global digital elevation model (DEM) products evaluated for vertical accuracy; row 11 presents the Sentinel-1-derived DEM generated in this study; and row 12 describes the ICESat-2 ATL08 dataset, used exclusively as reference data for vertical accuracy assessment.

Product	Pixel Size (m)	Vertical Accuracy (RMSE) [1]	Horizontal/Vertical Datum	Source	Reference
ALOS PALSAR	12.5	~5 m	WGS84/WGS84 ellipsoid	L-band InSAR	[21]
AW3D30	30	3.63 m	WGS84/EGM96	Stereo pan imagery	[22]
ANADEM	30	6.99 m	SIRGAS 2000/EGM2008	Multi-source optical/SAR fusion	[23]
ASTER GDEM	30	10.3 m	WGS84/EGM96	Stereo NIR imagery	[21]
Copernicus DEM	30	<1.68 m	WGS84/EGM2008	X-band single-pass InSAR	[24]
FABDEM	30	~3–7 m	WGS84/EGM2008	Copernicus DEM	[25]
GEDTM	30	~10 m	WGS84/EGM2008	Multisource data fusion	[26]
NASADEM	30	4.76 m	WGS84/EGM96	Reprocessed SRTM C-band InSAR	[27]
SRTM GL1	30	8.52 m	WGS84/EGM96	C-band InSAR	[28]
TanDEM-X	30	<1.4 m	WGS84/EGM2008	X-band single-pass InSAR	[29]
Sentinel-1	10	9.5 m	WGS84/WGS84	C-band SAR interferometry (InSAR)	[30]
ICESat-2 ATL08	100	0.73 m	WGS84/WGS84 ellipsoid	Photon-counting laser altimetry (ATLAS)	[31]

[1] RMSE values were obtained from product documentation or previously published assessments, as indicated in the Reference column. Because the reference datasets, geographic areas, and validation methods vary among sources, these values are provided for descriptive purposes only.

The DEM products evaluated in this study were classified according to their data acquisition technologies and processing approaches to facilitate comparative analysis. Three main groups were established: Group I—optical stereoscopy-derived DEMs, including ASTER GDEM and AW3D30; Group II—synthetic aperture radar (SAR)-derived DEMs, comprising SRTM, ALOS PALSAR, TanDEM-X, Copernicus DEM, and the Sentinel-1 inter-

ferometric DEM generated in this study; and Group III—multi-source or post-processed DEMs, comprising NASADEM, FABDEM, ANADEM, and GEDTM. This classification provides a framework for analyzing potential differences in DEM performance associated with sensor characteristics and processing techniques.

A DEM was generated from Sentinel-1 SAR data to assess the potential of interferometric SAR (InSAR) processing for representing floodplain topography. Two Sentinel-1A Single Look Complex (SLC) images acquired on 16 September and 10 October 2024 (dry season) were selected to form an interferometric pair [32–34]. Image selection considered both the temporal and perpendicular baselines, which influence interferometric coherence and the reliability of elevation estimates. The selected pair had a 24-day temporal baseline and a perpendicular baseline of approximately 162 m. Interferometric processing was performed using the Sentinel Application Platform (SNAP), version 13.0.0 [35]. The processing workflow included image co-registration, interferogram generation, coherence estimation, terrain observation by progressive scans (TOPS) debursting, Goldstein phase filtering, phase unwrapping, and conversion of the unwrapped interferometric phase into elevation estimates. These procedures followed the standard TOPS InSAR processing workflow implemented in SNAP. Terrain correction and geocoding were performed using the Copernicus DEM (30 m) as a reference elevation dataset to correct geometric distortions and generate the georeferenced DEM product.

2.3. Reference Elevation Data

The vertical accuracy of the DEMs was assessed using two independent reference datasets. The first consisted of 210 ground control points (GCPs) obtained from official Brazilian geodetic and surveying databases, including benchmarks from the Brazilian Geodetic System (SGB), georeferenced cadastral points from the Land Management System (SIGEF; https://sigef.incra.gov.br/sobre/acesso/; accessed on 10 September 2025), and elevation measurements from surveys conducted by the National Department of Transport Infrastructure (DNIT) and the Brazilian Institute of Geography and Statistics (IBGE; http://www.bdg.ibge.gov.br/appbdg/, accessed on 10 September 2025). These reference points provided elevation measurements for evaluating DEM vertical accuracy across the study area.

The second reference dataset initially comprised 15,515 elevation observations obtained from the ICESat-2 ATL08 product, which provides photon-counting laser altimetry estimates of terrain elevation and vegetation structure. To minimize noise and reduce potential vegetation-related biases, only observations satisfying the predefined elevation uncertainty and vegetation influence criteria were retained. Approximately 37.4% of the initial observations were excluded during preprocessing based on the adopted uncertainty threshold, resulting in 9713 elevation measurements retained for subsequent DEM vertical accuracy assessments. As reported by Weifeng et al. [21], terrain slope can influence the vertical accuracy of ICESat-2 elevation measurements, increasing elevation uncertainty in steep terrains. In the present study, a large proportion of the excluded observations were located in mountainous areas characterized by steep slopes, suggesting that terrain-related effects may have contributed to the higher uncertainty values observed in these locations.

The evaluated DEMs differed in nominal spatial resolution, ranging from 10 m for the Sentinel-1-derived DEM and 12.5 m for ALOS PALSAR to 30 m for the remaining products. These differences were considered when interpreting the vertical accuracy assessments and hydrodynamic simulation results. Because vertical accuracy was assessed using independent point-based reference datasets, the resulting statistics reflect elevation errors at the sampled locations rather than differences in spatial resolution alone.

2.4. Accuracy Assessment

The vertical accuracy of the DEMs was assessed by comparing DEM elevations extracted using ArcMap version 10.8 with independent reference elevations. All elevation values were converted to a common ellipsoidal height reference based on the WGS84 ellipsoid. For each ICESat-2 reference point, the corresponding DEM elevation and geoid undulation were extracted and compiled in Microsoft Excel for Microsoft 365. For DEMs originally referenced to orthometric heights, geoid undulations were obtained from the EGM2008 (egm08_25.gtx) or EGM96 (egm96_15.gtx) geoid models, according to the vertical datum of each dataset (Table 1). Orthometric heights were converted to ellipsoidal heights using the equation h = H + N, where h is the ellipsoidal height, H is the orthometric height, and N is the geoid undulation. Datasets originally georeferenced to ellipsoidal heights, including ICESat-2 ATL08, were retained in their original vertical reference system. The normality of the DEM elevation residuals was assessed using the Shapiro–Wilk test at a significance level of 0.05. Systematic vertical bias was evaluated using the one-sample Student's *t*-test and the nonparametric Wilcoxon signed-rank test.

The vertical accuracy of the DEM products was also evaluated according to the Brazilian Cartographic Accuracy Standard for Digital Cartographic Products (PEC-PCD), which establishes criteria for assessing vertical positional accuracy [36]. This standard defines four accuracy classes (A–D) based on acceptable elevation errors and root mean square error (RMSE) thresholds for specific mapping scales. The PEC-PCD framework and its application to DEM accuracy assessment were presented by Souza et al. [37]. The assessment involves comparison of DEM elevations with independent reference measurements to quantify vertical errors and determine compliance with the prescribed accuracy thresholds. To minimize the influence of reference data uncertainty on the assessment results, the reference measurements should preferably be at least three times more accurate than the DEM being evaluated [38,39].

The classification of DEMs into PEC-PCD vertical accuracy classes was based on two criteria derived from elevation residuals. The first criterion was the PEC threshold, which specifies the maximum acceptable absolute elevation error that at least 90% of the evaluated points must satisfy. The second criterion was the standard error (EP) threshold, which was compared with the RMSE of the elevation residuals. To qualify for a given accuracy class, a DEM had to meet both the PEC and EP thresholds established by the standard. Within the PEC-PCD framework, vertical accuracy thresholds are defined according to contour intervals associated with specific cartographic mapping scales. Contour intervals of 5, 10, 20, 50, and 100 m correspond approximately to mapping scales of 1:10,000, 1:25,000, 1:50,000, 1:100,000, and 1:250,000, respectively. Each DEM was assigned to the applicable PEC-PCD accuracy class by comparing its absolute elevation residuals and RMSE values with the corresponding PEC and EP thresholds.

Vertical accuracy was assessed using the mean error (ME), mean absolute error (MAE), RMSE, and standard deviation (SD) of elevation residuals. RMSE was adopted as the primary indicator of vertical accuracy because it quantifies the overall magnitude of elevation errors while assigning higher weight to larger discrepancies [37]. MAE provides a complementary measure of error magnitude by representing the average absolute difference between DEM-derived and reference elevations, whereas ME indicates systematic vertical bias. Elevation residuals (ΔZ) were first calculated as the differences between DEM-derived elevations and their corresponding reference values (Equation (1)). Subsequently, SD, ME, MAE, and RMSE were calculated using the expressions presented in Equations (2)–(5).

$$\Delta Z = Z_{DEM} - Z_{ATL08} \quad (1)$$

$$SD = \sqrt{\frac{\sum_{i=1}^{n}((Z_{MDE} - Z_{ATL08}) - ME)^2}{n-1}} \quad (2)$$

$$ME = \frac{1}{n}\sum_{i=1}^{n}(Z_{DEM} - Z_{ATL08}) \quad (3)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Z_{DEM} - Z_{ATL08}| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Z_{DEM} - Z_{ATL08})^2} \quad (5)$$

where ΔZ is the elevation difference between the DEM and the control point elevation; Z_{DEM} is the elevation value derived from the DEM; Z ATL08 is the elevation value of the control point; and n is the number of observations.

DEM vertical accuracy was also assessed across different LULC classes using data from the MapBiomas Project. MapBiomas Collection 10 provides annual LULC maps for Brazil from 1985 to 2024 at a spatial resolution of 30 m [18]. The 2024 LULC map was used to stratify the vertical accuracy assessment by LULC class. DEM-derived elevations were also compared with ICESat-2 canopy-top elevations, calculated as the sum of terrain elevation and canopy height (h_canopy).

Visual assessment was conducted to complement statistical validation and evaluate terrain characteristics not fully captured by conventional accuracy metrics, including surface smoothness, noise patterns, the representation of channels and paleochannels, elevation artifacts, and differences associated with sensor characteristics and processing methods. These features are particularly relevant in low-relief floodplains, where subtle topographic variations can influence drainage networks, surface water flow pathways, and hydrological connectivity.

2.5. Flood Inundation Simulation

To investigate the hydrological implications of DEM selection, flood inundation simulations were conducted using HEC-RAS version 7.0. This hydrodynamic modelling simulates water flow and inundation extent based on terrain elevation, hydraulic parameters, and boundary conditions [11,40]. The three DEM products with the lowest RMSE values (ANADEM, FABDEM, and GEDTM) were selected as the primary terrain inputs for the simulations. The Sentinel-1-derived DEM was also included as an experimental terrain input to assess its ability to reproduce inundation patterns consistent with independently derived Sentinel-2 surface water maps. The inclusion of the Sentinel-1-derived DEM was motivated not only by its vertical accuracy but also by the opportunity to evaluate its suitability for representing floodplain topography and simulating inundation patterns under low-water conditions. Before hydrodynamic modeling, all DEMs were converted to a common vertical reference using the Brazilian hgeoHNOR quasigeoid model.

Monthly mean water level and discharge records from five gauging stations operated by the National Water and Sanitation Agency (ANA) were analyzed to evaluate data consistency and interstation relationships through correlation analysis. This preliminary assessment was used to determine the suitability of the available hydrological records for HEC-RAS simulations. The modelling domain was restricted to the river reach between the Bela Vista and São Francisco gauging stations. This delimitation allowed the establishment of upstream and downstream boundary conditions using observations from these two stations.

Based on the 1960–2024 hydrological records, two contrasting hydrological scenarios were selected for flood inundation simulations. The high-water scenario corresponded to April 1983, one of the most severe flood events recorded during the study period, whereas the low-water scenario corresponded to October 2024, when exceptionally low monthly

mean water levels were observed along the Paraguay River. Because a Sentinel-2 satellite image was unavailable for the April 1983 flood event, an additional high-water scenario corresponding to May 2023 was included.

For the April 1983 scenario, the observed monthly mean discharges at the upstream and downstream gauging stations were approximately 1760 and 2961 $m^3 s^{-1}$, respectively, with corresponding water levels of 94.9 and 95.4 m. For the May 2023 scenario, the monthly mean discharges were 1598 and 1905 $m^3 s^{-1}$, with corresponding water levels of 93.4 and 90.5 m at the upstream and downstream stations, respectively. The October 2024 scenario was characterized by significantly lower monthly mean discharges of 443 and 441 $m^3 s^{-1}$, with corresponding water levels of 90.3 and 86.2 m at the upstream and downstream stations, respectively.

The hydraulic geometry was represented by 54 cross-sections distributed along a 168.60 km reach of the Paraguay River. The cross-sections were oriented approximately perpendicular to the main flow direction, with longitudinal spacing ranging from 1 to 5 km and widths varying from approximately 5 to 30 km to encompass the areas potentially contributing to flood conveyance and floodplain storage. Because the evaluated DEMs do not adequately represent submerged riverbed topography, the channel geometry was refined using bathymetric data provided by the DNIT. These data were incorporated into the cross-sectional profiles to represent the submerged channel bed, whereas the surrounding floodplain elevations were extracted from each evaluated DEM. The same bathymetric dataset and channel geometry were used in all DEM-based simulations to minimize differences related to channel representation and isolate, as far as possible, the influence of floodplain topography on simulated inundation patterns.

This procedure ensured that observed river stages were expressed in the same vertical reference system as the DEMs. Hydraulic resistance was represented using Manning's roughness coefficient ($n = 0.035$), which was held constant for all simulations. For each hydrological scenario, channel geometry, hydraulic roughness, hydrological boundary conditions, and numerical model configuration remained unchanged across all DEM-based simulations. Floodplain terrain elevation was the primary variable distinguishing the model runs, allowing differences in simulated inundation extent and water depth to be attributed mainly to variations in the topography represented by each DEM.

Model outputs were used to generate inundation extent and water-depth maps for each hydrological scenario. The total inundated area was calculated for each DEM-based simulation to evaluate the sensitivity of inundation extent to differences in terrain representation under contrasting high- and low-water conditions.

2.6. Flood Extent Validation Using Sentinel-2 Water Masks

To evaluate the spatial agreement between inundation extents simulated by the HEC-RAS hydrodynamic model and satellite-derived surface water, independent reference maps were generated from Sentinel-2 images processed in Google Earth Engine (GEE). Sentinel-2 Level 2A surface reflectance data covering the study area (tiles 21KVA, 21KVB, 21KVU, and 21KVV) were accessed through GEE. Sentinel-2 images were selected because of their relatively high spatial resolution (10–20 m) and spectral characteristics suitable for detecting surface water and distinguishing it from surrounding LULC types. Two hydrological conditions were considered for satellite-based validation: the June 2023 flood event and the October 2024 low-water season. The April 1983 flood scenario was excluded from this assessment because Sentinel-2 observations were unavailable for that period.

For each hydrological scenario, a Sentinel-2 image acquired on a single date was selected to ensure temporal consistency and complete coverage of the study area across all four tiles. For the 2024 dry-season scenario, images acquired between 1 September and 10

October 2024 with scene-level cloud cover below 10% were initially screened. The screening identified 15 scenes acquired on 2, 7, 17, and 22 September 2024. No October acquisition met the selection criteria, as scenes satisfying the cloud-cover threshold covered only 3.1% of the study area. Therefore, a suitable image was available only for September 2024. Among the available dates, 7 September 2024 was selected because the image showed the lowest apparent smoke contamination. For the 2023 flood scenario, no Sentinel-2 acquisition in May provided complete coverage of the study area under the same cloud-cover criterion. Therefore, 5 June 2023 was selected as the earliest subsequent acquisition date providing complete coverage of the study area. Cloud and cirrus pixels were masked using the Sen2Cor version 2.12 SCL (classes 8, 9, and 10). Pixels classified as cloud shadows (class 3) were retained to minimize the potential exclusion of dark water surfaces misclassified as shadows.

The hydrological boundary conditions used in the HEC-RAS simulations were defined using monthly mean water levels recorded at the ANA gauging stations for May 2023 and October 2024, consistent with the temporal resolution of the available hydrological data. Consequently, the Sentinel-2-derived surface water masks and HEC-RAS-simulated inundation extents did not correspond to the same dates or hydrological conditions. This temporal mismatch constitutes a limitation of the satellite-based validation, as short-term fluctuations in water levels and their effects on inundation extent were not explicitly accounted for.

Surface water masks were generated using a rule-based classification approach that combined spectral indices, scene classification information, and spectral reflectance thresholds. The Normalized Difference Water Index (NDWI) and Modified Normalized Difference Water Index (MNDWI) were used to identify open-water surfaces. However, the effectiveness of these indices may be reduced in optically complex aquatic environments, particularly in dark or turbid waters characterized by high concentrations of dissolved organic matter and humic substances, as commonly observed in the Pantanal during flooding events. To address this limitation, a composite surface water mask was generated using the logical union (Boolean OR) of four complementary criteria: MNDWI > 0; NDWI > 0; pixels classified as water by the Sen2Cor SCL (SCL $= 6$); and pixels satisfying both B8 (near-infrared) < 1000 and B11 (shortwave infrared) < 400, using surface reflectance values scaled to the range of 0–10,000. This combined approach was designed to improve the detection of surface water across different aquatic environments, including river channels, permanent lakes, and seasonally inundated areas. The same classification criteria were applied to both acquisition dates to ensure methodological consistency.

To evaluate the spatial agreement between simulated inundation extents and satellite-derived surface water, the inundation maps generated by HEC-RAS using each DEM were compared with the corresponding Sentinel-2-derived surface water masks. Spatial agreement was quantified using the Intersection over Union (IoU) metric, calculated as the ratio of the intersection area to the union area of the simulated inundation extents and Sentinel-2-derived surface water masks. IoU values range from 0 (no spatial overlap) to 1 (perfect spatial agreement) and were expressed as percentages.

Because HEC-RAS was implemented as a one-dimensional hydraulic model, the lateral extent of simulated inundation was constrained by the cross-section endpoints, which did not extend into the surrounding higher-elevation terrain. In contrast, Sentinel-2-derived surface water masks detected water bodies regardless of their hydraulic connectivity to the main river channel. To account for this difference in spatial coverage, a second reference mask was generated by clipping the Sentinel-2-derived surface water mask to a polygon encompassing all cross-sections (hereafter referred to as the cross-section envelope). Only the Sentinel-2-derived reference mask was clipped, as all simulated inundation extents were already contained within this envelope. The same envelope was used for comparisons in-

volving all four DEM-based simulations, allowing a consistent spatial domain for evaluating spatial agreement. Both the original and clipped Sentinel-2-derived reference masks were retained. The original mask was used to assess spatial agreement across the entire study area, whereas the clipped mask was used to evaluate agreement within the spatial domain covered by the hydraulic model. Spatial discrepancies were further quantified in terms of false positives (FP), corresponding to areas simulated as inundated but classified as non-water in the reference mask, and false negatives (FN), corresponding to areas classified as water in the reference mask but not simulated as inundated.

3. Results

3.1. DEM Accuracy Assessment Using ICESat-2 Data

ME indicates the direction and magnitude of systematic elevation bias in relation to ICESat-2 terrain elevations. In this study, the smallest absolute biases were observed for the multi-source DEMs (Group III) and the Sentinel-1-derived interferometric DEM. The Sentinel-1-derived DEM showed the lowest absolute bias (0.07 m), followed by ANADEM (0.31 m), indicating minimal systematic elevation differences in relation to the ICESat-2 reference data. In contrast, ASTER GDEM (3.90 m) and SRTM (3.44 m) presented larger positive biases, suggesting that their elevation estimates may be affected by vegetation canopy effects and may partially reflect vegetation surface elevations rather than consistently representing the underlying terrain. GEDTM was the only DEM showing a negative bias (−0.46 m), indicating a slight tendency to underestimate elevations compared with the ICESat-2 ATL08 reference data.

The distributions of elevation residuals differed among the evaluated DEMs, as illustrated by the histograms in Figure 2. The Shapiro–Wilk test rejected the null hypothesis of normality for all DEMs ($p < 0.05$), indicating statistically significant deviations from normality (Table 2). Because all DEMs were evaluated using the same 9713 ICESat-2 reference points, their elevation residuals were paired across DEM products. The Wilcoxon test was applied separately to each DEM to assess whether its elevation residuals were centered around zero. Statistically significant deviations from zero were detected for all DEMs ($p < 0.05$). However, given the large sample size, even small deviations may reach statistical significance without necessarily indicating significant systematic elevation biases. Therefore, these statistical test results should not be interpreted as indicators of the magnitude of systematic vertical errors. Instead, the magnitude and direction of vertical bias were characterized using ME, while the dispersion of elevation residuals was quantified using SD.

The SD of elevation residuals quantifies error dispersion around the ME, providing an indication of the variability of vertical errors after accounting for systematic bias. The multi-source DEMs presented the lowest residual variability, with FABDEM (1.30 m), GEDTM (1.31 m), and ANADEM (1.39 m) presenting the smallest SD values. In contrast, ASTER GDEM showed the highest SD (5.51 m), indicating higher dispersion of elevation errors. This higher variability may be partly attributed to limitations inherent in optical stereoscopic elevation retrieval, including sensitivity to cloud contamination, terrain shadows, and reduced image matching reliability under unfavorable surface and imaging conditions.

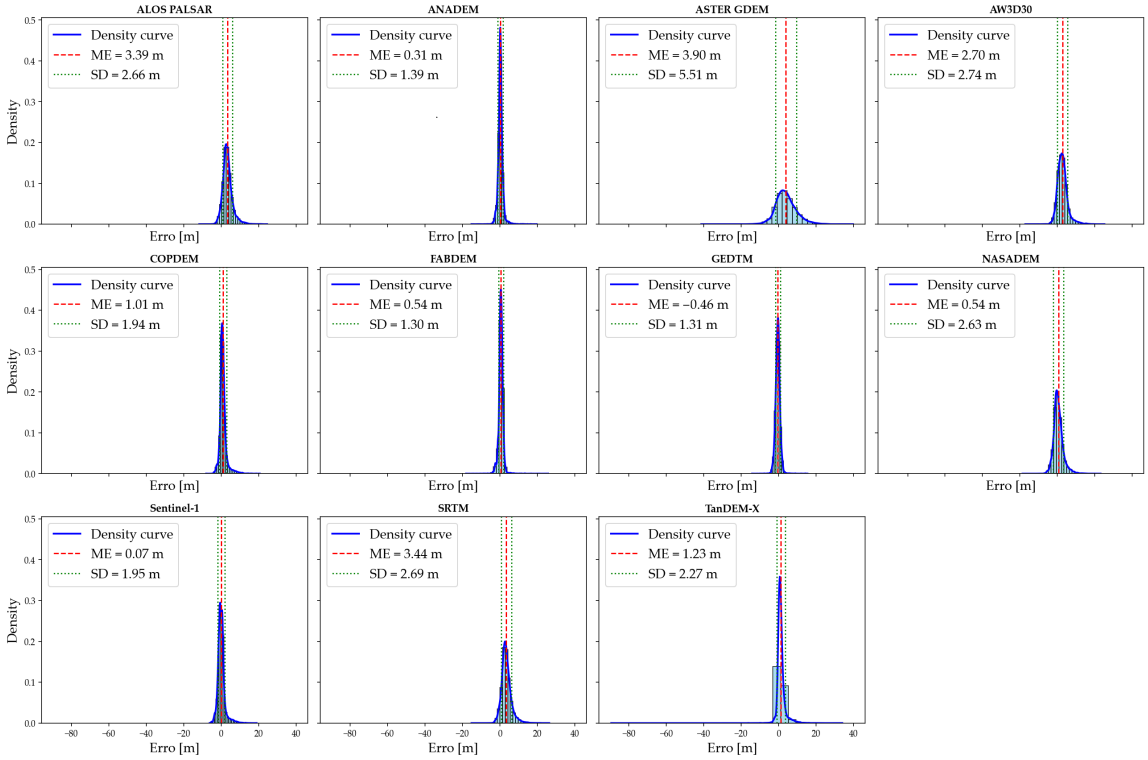

Figure 2. Distributions of elevation residuals for the evaluated digital elevation models (DEMs) compared with ICESat-2 ATL08 reference terrain elevations. Blue bars represent the frequency distribution of elevation residuals across predefined intervals (bins). Blue curves represent kernel density estimates of elevation residuals (DEM-ICESat-2). Red dashed lines indicate mean error (ME), whereas green dashed lines represent the standard deviation (SD) of the residuals. Differences in distribution shape, central tendency, and dispersion illustrate variations in vertical bias and error variability among the DEMs.

Table 2. Results of the Shapiro–Wilk normality and Wilcoxon signed-rank tests applied to elevation residuals (DEM-ICESat-2) for the evaluated digital elevation models (DEMs).

DEM	N	Shapiro–Wilk W	*p*-Value	Wilcoxon W	*p*-Value
ALOS PALSAR	9713	0.933	4.450×10^{-63}	2,052,130	$<1.0 \times 10^{-3}$
ANADEM	9713	0.830	5.144×10^{-83}	41,795,255	1.33×10^{-237}
ASTER GDEM	9713	0.977	1.408×10^{-43}	17,164,972	$<1.0 \times 10^{-3}$
AW3D30	9713	0.930	5.708×10^{-64}	6,279,835	$<1.0 \times 10^{-3}$
Copernicus DEM	9713	0.792	1.310×10^{-87}	19,865,729	$<1.0 \times 10^{-3}$
FABDEM	9713	0.817	1.335×10^{-84}	25,578,583	$<1.0 \times 10^{-3}$
GEDTM	9713	0.903	9.096×10^{-71}	33,071,693	$<1.0 \times 10^{-3}$
NASADEM	9713	0.924	8.402×10^{-66}	49,646,318	3.73×10^{-79}
Sentinel-1	9713	0.858	6.660×10^{-79}	55,152,544	3.13×10^{-19}
SRTM	9713	0.923	6.765×10^{-66}	195,412	$<1.0 \times 10^{-3}$
TanDEM-X	9713	0.693	5.482×10^{-97}	13,429,998	$<1.0 \times 10^{-3}$

Based on RMSE, the DEMs were ranked in descending order of vertical accuracy as follows: GEDTM, FABDEM, ANADEM, Sentinel-1, Copernicus DEM, TanDEM-X, NASADEM, AW3D30, ALOS PALSAR, SRTM, and ASTER GDEM. GEDTM and FABDEM showed the lowest RMSE values (1.39 m and 1.41 m, respectively), indicating the highest overall vertical accuracy among the evaluated products. Among the SAR-derived DEMs, Sentinel-1 (1.95 m) and Copernicus DEM (2.18 m) presented considerably lower RMSE values than ALOS PALSAR (4.31 m) and SRTM (4.37 m). RMSE is a widely used measure of overall

vertical error magnitude because it incorporates both systematic bias and residual variability. However, its sensitivity to large residuals and outliers should be considered when interpreting differences in DEM performance.

The comparison between RMSE and MAE provides complementary insights into the magnitude and distribution of elevation residuals (Figure 3). A smaller difference between RMSE and MAE suggests a lower influence of large residuals on the overall error magnitude, whereas a larger difference indicates greater sensitivity to high-magnitude errors. GEDTM (0.76 m), ANADEM (0.87 m), and FABDEM (0.89 m) presented the smallest RMSE–MAE differences, suggesting that large residuals had a comparatively limited influence on their RMSE values. However, the RMSE–MAE difference alone does not fully characterize DEM vertical accuracy or the distribution of elevation errors and should therefore be interpreted in conjunction with the individual error metrics.

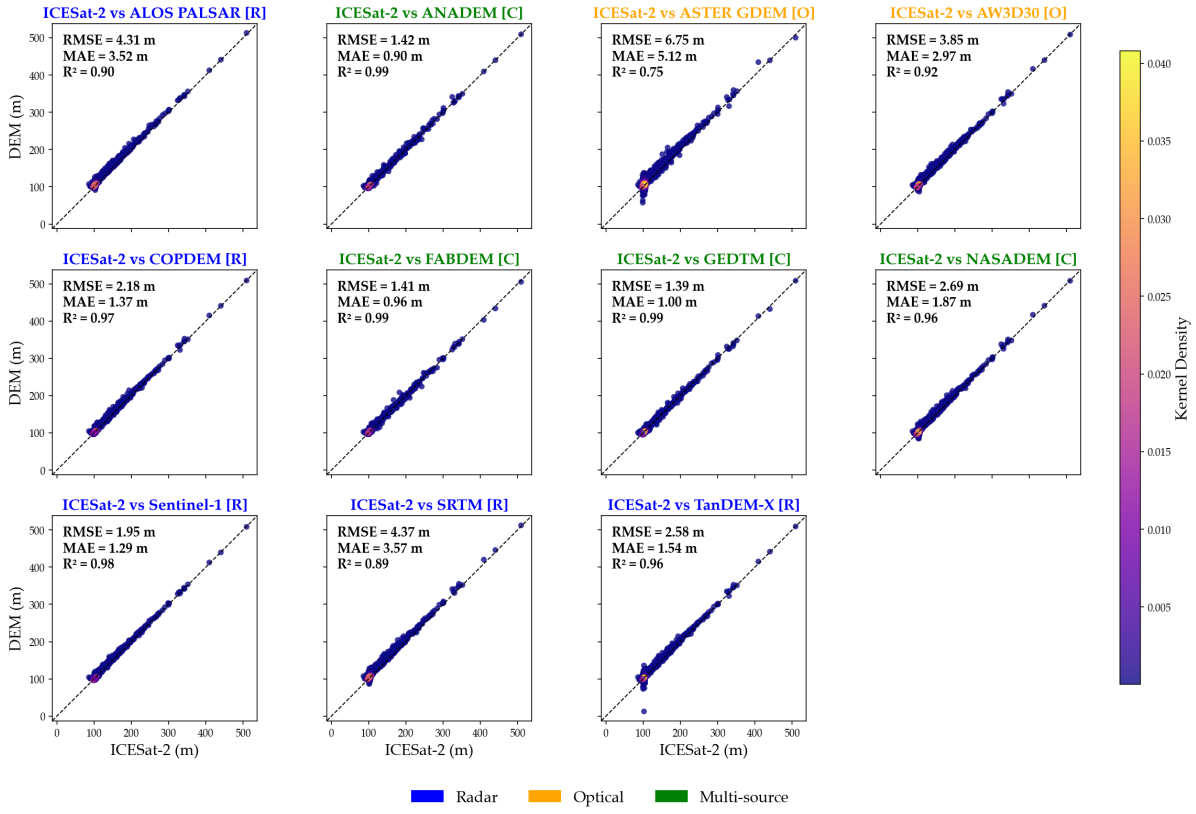

Figure 3. Scatter plots comparing elevations derived from the evaluated digital elevation models (DEMs) with ICESat-2 ATL08 reference terrain elevations. Each panel represents a DEM product, and the dashed line indicates perfect agreement (1:1 line). Point colors represent kernel density estimates, highlighting areas with higher concentrations of observations. Statistical metrics reported in each panel include root mean square error (RMSE), mean absolute error (MAE), and R^2. DEMs are grouped according to their data sources and processing approaches: radar-derived (blue), optical stereoscopy-derived (orange), and multi-source (green).

ASTER GDEM and TanDEM-X presented larger RMSE-MAE differences (1.63 m and 1.04 m, respectively), suggesting a higher contribution of high-magnitude elevation residuals to their overall errors. In low-relief environments such as floodplains, localized elevation errors may create artificial topographic barriers or depressions, potentially disrupting hydrological connectivity and affecting the delineation of inundated areas [41,42]. FABDEM and GEDTM showed both low RMSE values and small RMSE–MAE differences, indicating high overall vertical accuracy and a limited contribution of large residuals to their RMSE values. These characteristics are desirable for hydrological applications in low-relief envi-

ronments, where even small elevation errors can significantly affect simulated inundation patterns [13].

The R^2 was used to evaluate the strength of the linear relationship between DEM-derived elevations and ICESat-2 reference elevations, rather than as a direct measure of vertical accuracy. The multi-source DEMs (ANADEM, FABDEM, and GEDTM) showed the strongest linear relationships with the reference elevations ($R^2 = 0.99$), indicating that they closely reproduced the elevation variability observed by ICESat-2. In contrast, ASTER GDEM showed the lowest R^2 value (0.75), reflecting a weaker linear relationship with the reference elevations. However, a high R^2 does not necessarily imply high vertical accuracy, as systematic elevation biases and other discrepancies may persist even when the linear relationship is strong.

3.2. Validation Using Multi-Source Ground Control Points

Validation based on ground-based reference points from multiple sources indicated larger elevation discrepancies (Figure 4) than those observed using ICESat-2 altimetry. The ground-based reference dataset comprised geodetic benchmarks and cadastral control points obtained from the SIGEF, DNIT, and IBGE databases. Although these datasets provide independent reference measurements for evaluating DEM vertical accuracy, they are also subject to uncertainties. Differences in surveying methods, acquisition dates, vertical reference systems, measurement accuracy, and spatial distribution may have contributed to the higher variability in elevation residuals observed in the ground-based validation.

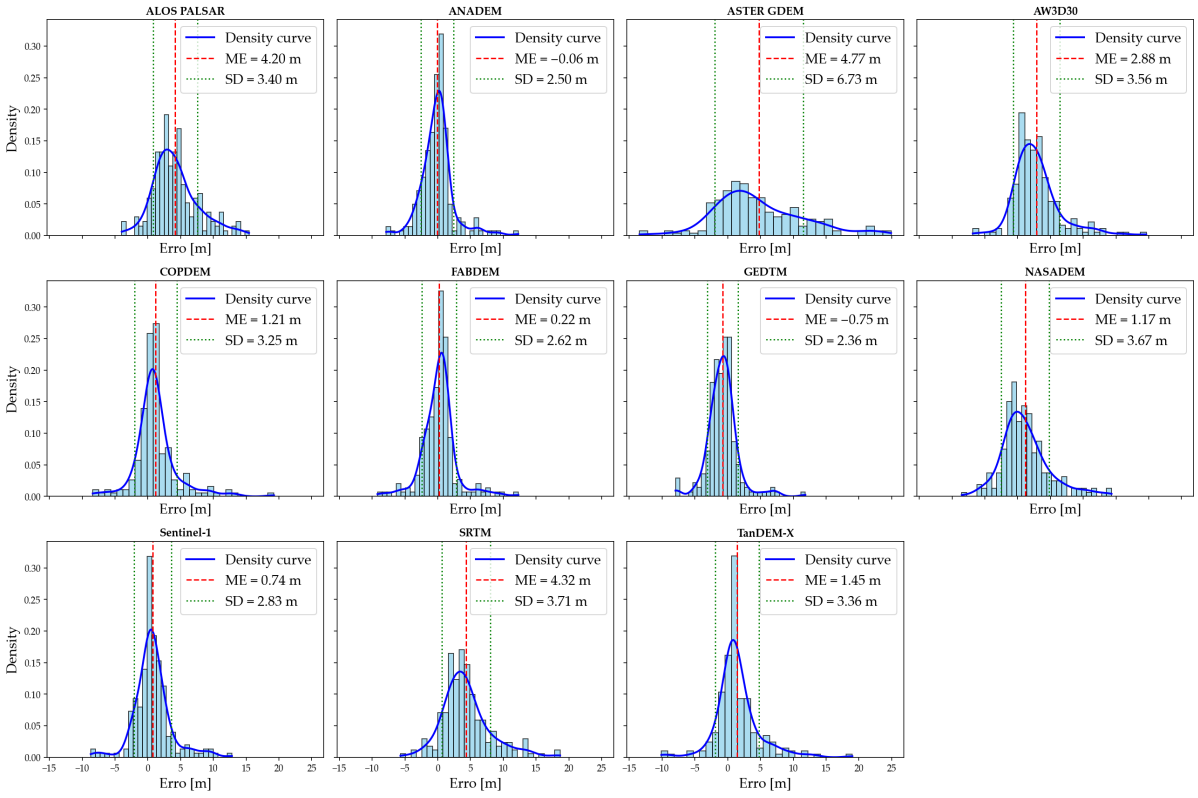

Figure 4. Distribution of elevation residuals (DEM-derived elevation minus reference elevation) for the evaluated digital elevation models (DEMs), based on the multi-source ground control dataset. Each panel presents a histogram of elevation residuals with an overlaid kernel density estimate curve, illustrating the frequency distribution and variability of elevation errors for each DEM. Blue bars represent the frequency distribution of elevation residuals across predefined intervals (bins). Blue curves represent kernel density estimates of elevation residuals. Red dashed lines indicate mean error (ME), whereas green dashed lines represent the standard deviation (SD) of the residuals.

GEDTM and ANADEM showed the lowest RMSE values (approximately 2.50 m), indicating the highest vertical accuracy when compared with ground-based reference points. In contrast, ASTER GDEM showed the largest elevation errors (RMSE = 8.25 m) and a high positive mean error (ME = 4.77 m), indicating a systematic tendency to overestimate terrain elevations at the sampled locations. Despite differences in error magnitudes between the two reference datasets, the broadly consistent results from ICESat-2 altimetry and ground-based reference points reinforce the accuracy ranking of the evaluated DEMs. GEDTM, the most accurate DEM in the ICESat-2-based assessment, presented an RMSE of 1.39 m, compared with 2.48 m when evaluated using ground-based reference points. Despite differences in error magnitude between the two validation datasets, GEDTM, ANADEM, and FABDEM consistently ranked among the three most accurate DEMs.

The consistency between the results obtained from the two independent validation datasets provides evidence of the comparatively high vertical accuracy of these DEMs in the study area. Their superior performance may be partly attributed to improvements in DEM processing, including vegetation-related elevation corrections and terrain refinement procedures. GEDTM and ANADEM emerged as especially suitable for representing terrain elevations in the study area. In contrast, ASTER GDEM consistently showed the lowest vertical accuracy, with RMSE values of 6.75 m and 8.25 m when compared with the ICESat-2 and ground-based reference datasets, respectively. Table 3 summarizes the RMSE values obtained from the two independent reference datasets. Overall, the results indicate broadly consistent patterns of vertical accuracy across both validation approaches, with more recently developed or refined DEM products showing the lowest elevation errors.

Table 3. Comparison of the vertical accuracy of the evaluated digital elevation models (DEMs) using ICESat-2 ATL08 observations and multi-source ground control points.

DEMs	ICESat-2 ATL08 RMSE (m)	Ground Control Points RMSE (m)
GEDTM	1.39	2.48
FABDEM	1.41	2.62
ANADEM	1.42	2.50
Sentinel-1	1.95	2.93
Copernicus DEM	2.18	3.47
TanDEM-X	2.58	3.66
NASADEM	2.69	3.85
AW3D30	3.85	4.58
ALOS PALSAR	4.31	4.40
SRTM	4.37	5.70
ASTER GDEM	6.75	8.25

3.3. Influence of Land Cover Classes

Due to the limited spatial density of ground-based reference points available from the DNIT, IBGE, and SIGEF databases, the assessment of DEM vertical accuracy across different LULC classes was conducted exclusively using ICESat-2 reference elevations. Eight LULC classes were considered: forest formation, savanna formation, grassland formation, pasture, wetland area, urban area, non-vegetated area, and water bodies (rivers and lakes).

DEM vertical errors varied considerably across LULC classes (Figure 5). RMSE values ranged from 1.09 to 10.28 m across all evaluated DEMs and LULC classes, with lower errors over water bodies and wetlands and higher errors in forest formations. However, the spatial distribution of ICESat-2 reference points was highly uneven among LULC classes. Wetlands (37.7%), grasslands (26.8%), and water bodies (26.7%) accounted for 91.2% of the reference points, followed by forest formations (6.6%) and savanna formations (1.4%).

In contrast, urban areas (0.4%), pastures (0.3%), and non-vegetated areas (<0.1%) were underrepresented, accounting for 0.7% of the sample. Given their limited sample sizes, accuracy estimates for these three classes were interpreted with caution and excluded from the comparative ranking of DEM performance across LULC classes.

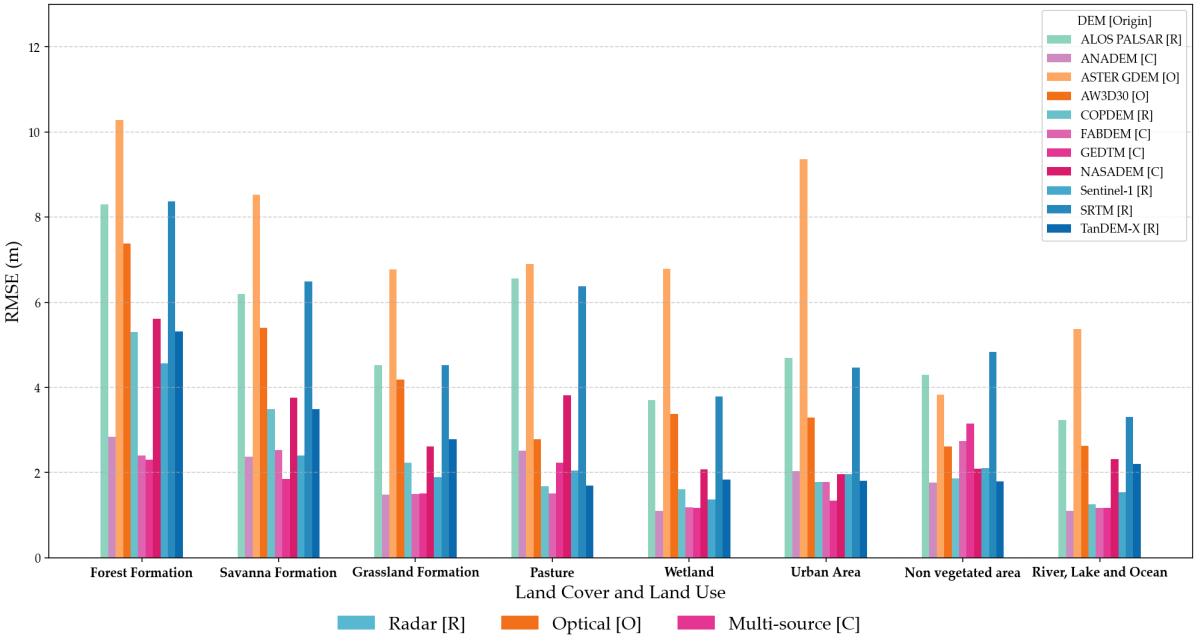

Figure 5. Root mean square error (RMSE) of the evaluated digital elevation models (DEMs) across eight land use and land cover (LULC) classes, based on ICESat-2 ATL08 reference elevations. Validation was performed using only ground-classified photons to minimize vegetation-related effects.

Among LULC classes with large reference samples, GEDTM, FABDEM, and ANADEM showed the lowest RMSE values. In forest formations, GEDTM (2.31 m), FABDEM (2.40 m), and ANADEM (2.83 m) were the most accurate DEMs, whereas GEDTM (1.85 m) and ANADEM (2.37 m) presented the lowest RMSE values in savanna formations. In grasslands, GEDTM, FABDEM, and ANADEM presented similar RMSE values, ranging from 1.48 to 1.52 m. In wetlands and water bodies, ANADEM achieved the lowest RMSE values (1.10 m and 1.09 m, respectively), closely followed by GEDTM and FABDEM (1.17–1.18 m). In contrast, ASTER GDEM presented the highest RMSE values in seven of the eight LULC classes, with an RMSE value of 10.28 m in forest formations. SRTM and ALOS PALSAR also showed comparatively high elevation errors across several LULC classes.

ME statistics suggest that systematic elevation biases contributed considerably to differences in DEM performance across vegetation classes. In forest formations, where the mean canopy height was 8.4 m, ASTER GDEM, SRTM, and ALOS PALSAR showed high positive elevation biases of 8.15, 7.59, and 7.59 m, respectively, equivalent to 90–97% of the mean canopy height. AW3D30 also presented a considerable positive bias of 6.43 m (77% of mean canopy height), whereas Sentinel-1, Copernicus DEM, TanDEM-X, and NASADEM showed smaller positive biases ranging from 3.20 to 4.45 m (38–53%). In contrast, FABDEM, ANADEM, and GEDTM showed lower biases of 1.03, 1.13, and −0.50 m, respectively. A similar pattern was observed in savanna formations, where the mean canopy height was 6.9 m. When DEM elevations were compared with ICESat-2 canopy-top elevations rather than terrain elevations, RMSE values for ASTER GDEM, SRTM, and ALOS PALSAR in forest formations decreased by 2.2–2.7 m, while their mean elevation differences, when compared with the canopy-top elevations, remained within ±0.8 m. Conversely, RMSE values for GEDTM, FABDEM, and ANADEM increased by 7.0–8.9 m.

These patterns suggest that ASTER GDEM, SRTM, ALOS PALSAR, and AW3D30 tend to represent elevations closer to the vegetation canopy surface, whereas GEDTM, FABDEM, and ANADEM closely approximate terrain elevations, consistent with the vegetation-correction procedures applied to these products. However, these results indicate associations rather than establishing a causal relationship between vegetation cover and elevation errors, for four main reasons. First, high positive elevation biases were also observed in sparsely vegetated LULC classes. In wetlands, where the mean canopy height was 1.0 m, ASTER GDEM, ALOS PALSAR, and SRTM presented positive biases ranging from 3.1 to 4.1 m. Similarly, positive biases of 2.4–3.0 m were observed over water bodies, where the mean canopy height was only 0.4 m. Because these biases exceeded the corresponding mean canopy heights, vegetation effects alone cannot fully explain the observed elevation errors. Other factors, including differences in water levels between DEM acquisition dates and ICESat-2 observations, may also have contributed to these discrepancies.

Second, canopy-top and terrain reference elevations were derived from the same ICESat-2 observations, implying that uncertainties in terrain detection may affect both reference datasets. In addition, canopy-height estimates may be less reliable in areas dominated by low-height vegetation. Third, the estimated mean canopy heights of 5.7 m in pastures and 10.2 m in non-vegetated areas were inconsistent with the expected characteristics of these LULC classes, suggesting potential spatial or temporal mismatches between the 30 m LULC map and ICESat-2 observations. Fourth, RMSE values related to canopy-top elevations remained high (6.1–7.6 m) for DEMs showing canopy-surface-like characteristics in forest formations, suggesting that local variations in canopy height were not adequately represented. Therefore, comparisons with canopy-top elevations provide additional evidence of potential vegetation-related elevation biases but do not allow their magnitude to be reliably quantified or their underlying causes to be conclusively identified. These findings highlight the importance of considering LULC characteristics when selecting DEMs for hydrological modelling and floodplain applications. Therefore, the class-specific accuracy estimates should be interpreted as indicators of DEM performance under different LULC conditions rather than as direct measures of vegetation-induced elevation errors.

3.4. PEC-PCD Classification

The results indicate that recently developed or post-processed DEMs presented the highest cartographic accuracy classifications (Table 4). ANADEM and FABDEM presented the best performance, satisfying the requirements for Class A at a 10 m contour interval and Class B at a 5 m contour interval. These findings suggest that both DEMs are suitable for cartographic applications requiring the corresponding levels of vertical accuracy.

Established DEM products, such as SRTM and ALOS PALSAR, presented lower cartographic accuracy classifications, meeting only Class D requirements at a 10 m contour interval or failing to satisfy the classification criteria at a 5 m interval. These results suggest that, despite their widespread use in regional-scale analyses, these DEMs may require additional elevation corrections to meet more rigorous vertical accuracy standards, particularly in low-relief floodplain environments, where even small elevation errors can significantly affect hydrological modeling results.

As the contour interval increased to 20, 50, and 100 m, most DEMs satisfied the requirements for Class A, reflecting the progressively less demanding vertical accuracy thresholds associated with larger contour intervals. Smaller contour intervals (5 and 10 m) indicated more pronounced differences in vertical accuracy among the evaluated DEMs, highlighting the limitations of several global elevation products for detailed terrain mapping.

Overall, these findings highlight the importance of considering both DEM vertical accuracy and the cartographic scale of the intended application. For applications requiring high vertical accuracy, such as hydrodynamic modeling and floodplain connectivity assessment, ANADEM and FABDEM emerged as promising terrain elevation datasets for the study area. However, their suitability for hydrodynamic applications cannot be determined solely from vertical accuracy metrics and cartographic classifications. Additional evaluation is necessary to assess how accurately these DEMs represent floodplain topography and how their terrain characteristics influence simulated inundation patterns.

Table 4. Classification of the evaluated digital elevation model (DEMs) according to the Brazilian Cartographic Accuracy Standard for Digital Cartographic Products (PEC-PCD) at contour intervals of 5, 10, 20, 50, and 100 m. Classes A–D represent progressively lower levels of vertical accuracy, with Class A indicating the highest accuracy.

DEM	5 m	10 m	20 m	50 m	100 m	Legend:
ANADEM	Class B	Class A	Class A	Class A	Class A	
FABDEM	Class B	Class A	Class A	Class A	Class A	
GEDTM	Class B	Class A	Class A	Class A	Class A	
Copernicus DEM	Class D	Class B	Class A	Class A	Class A	Class A
TanDEM-X	Not classified	Class B	Class A	Class A	Class A	Class B
Sentinel-1	Class C	Class B	Class A	Class A	Class A	Class C
NASADEM	Not classified	Class B	Class A	Class A	Class A	Class D
AW3D30	Not classified	Class C	Class B	Class A	Class A	Not classified
SRTM	Not classified	Class D	Class B	Class A	Class A	
ALOS PALSAR	Not classified	Class D	Class B	Class A	Class A	
ASTER GDEM	Not classified	Not classified	Class C	Class A	Class A	

3.5. Visual Comparison of DEMs

Two representative areas (C and D) along the Paraguay River floodplain were selected to illustrate differences in terrain representation, channel delineation, and floodplain microtopography among the evaluated DEMs (Figure 6). Visual inspection indicated noticeable differences in the representation of topographic features. Copernicus DEM provided the most visually coherent terrain representation of floodplain morphology, with clearly delineated main and secondary channels, depressions, and natural levees. FABDEM and the Sentinel-1-derived DEM also provided satisfactory representations of these features, although their elevation surfaces appeared slightly smoother, with less pronounced topographic detail along channel margins.

GEDTM presented notable limitations in its visual representation of floodplain topography. Despite its high overall vertical accuracy metrics, the DEM appeared to smooth subtle topographic features, resulting in less clearly defined river channels and reduced representation of local terrain variability. These findings highlight that point-based vertical accuracy metrics, such as RMSE, do not necessarily reflect a DEM's ability to preserve the spatial configuration and geomorphological characteristics of floodplain terrain.

ASTER GDEM showed a poor visual representation of floodplain topography, with limited ability to resolve subtle terrain features, whereas AW3D30 showed intermediate performance. Overall, Copernicus DEM, FABDEM, and the Sentinel-1-derived DEM provided more detailed and spatially coherent representations of floodplain morphology, while ASTER GDEM and GEDTM showed notable limitations in preserving local topographic features. These differences are especially relevant in low-relief floodplain environments, where even subtle elevation variations can significantly influence hydrological connectivity and inundation dynamics.

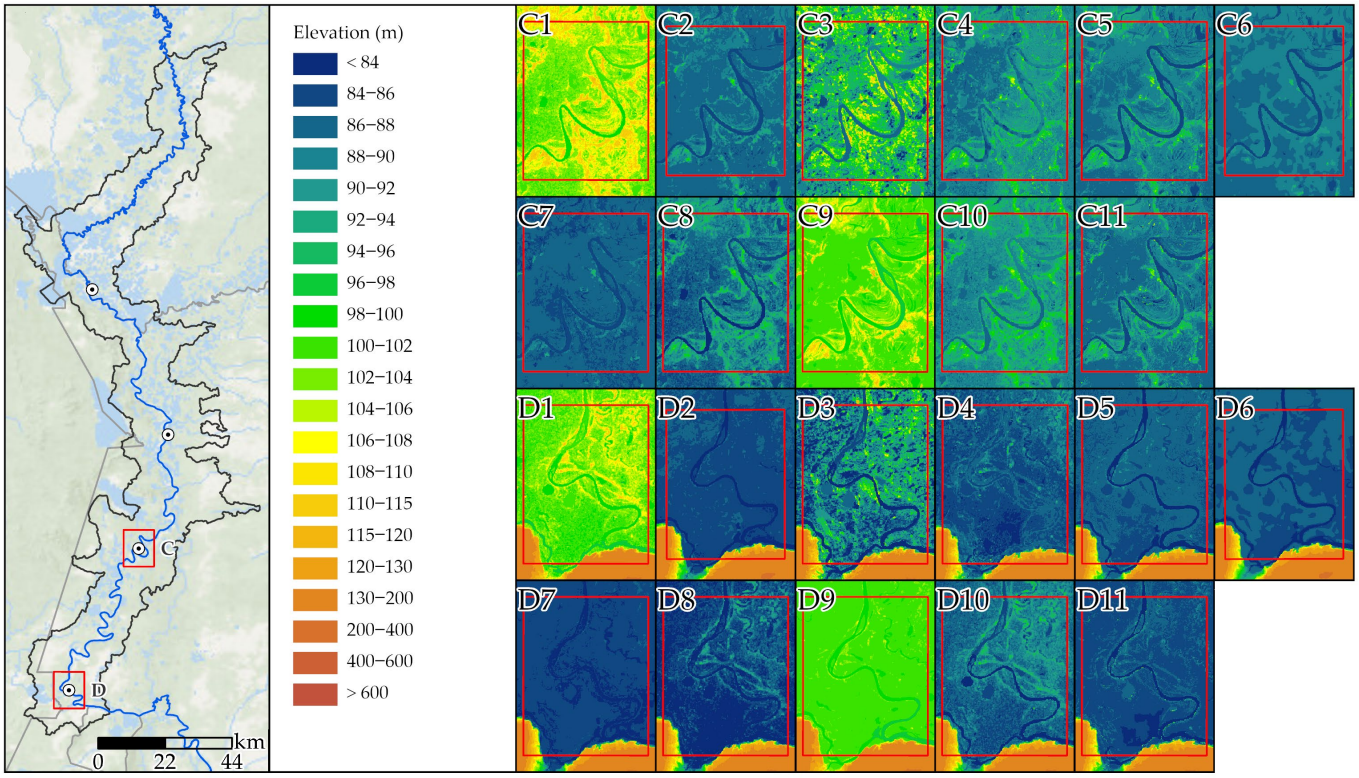

Figure 6. Visual comparison of the 11 evaluated digital elevation models (DEMs) in two representative areas (C and D) along the Paraguay River floodplain. The circles indicate the locations of staff gauges at fluviometric monitoring stations. The location map (**left**) identifies the selected areas, while the enlarged panels (**right**) illustrate differences in terrain elevation patterns, river channel delineation, and floodplain microtopography. Numbered panels correspond to the following DEMs: 1 = ALOS PALSAR; 2 = ANADEM; 3 = ASTER GDEM; 4 = AW3D30; 5 = Copernicus DEM; 6 = FABDEM; 7 = GEDTM; 8 = NASADEM; 9 = Sentinel-1; 10 = SRTM; and 11 = TanDEM-X.

3.6. Flood Inundation Simulation and DEM Influence

The influence of DEM selection on simulated inundation extent varied markedly across the hydrological scenarios (Figure 7). Under the 1983 flood conditions, the four DEMs produced similar inundation extents, ranging from 1637.84 km^2 (Sentinel-1-derived DEM) to 1664.83 km^2 (GEDTM), corresponding to a maximum difference of less than 2%. Because a satellite image was unavailable for this historical event, the simulated inundation extents could not be validated based on an independently observed flood boundary. To enable a quantitative comparison with satellite-derived inundation extents, the same modeling framework was applied to the May 2023 flood event, for which a Sentinel-2 image was available. Under these conditions, simulated inundation extents ranged from 1000.92 km^2 (FABDEM) to 1324.34 km^2 (GEDTM), representing a difference of approximately 32% in comparison with FABDEM. ANADEM (1084.81 km^2) and the Sentinel-1-derived DEM (1288.53 km^2) produced intermediate estimates (Table 5). Under the October 2024 low-water conditions, DEM selection resulted in the largest differences in simulated inundation extent, ranging from 90.55 km^2 (FABDEM) to 143.14 km^2 (Sentinel-1-derived DEM), corresponding to a difference of 58% when compared with FABDEM (Figure S1). In contrast, ANADEM (91.06 km^2) and GEDTM (98.07 km^2) produced inundation extents comparable to that obtained with FABDEM. Overall, these results indicate that the sensitivity of simulated inundation extent to DEM selection increased significantly as inundation extent decreased.

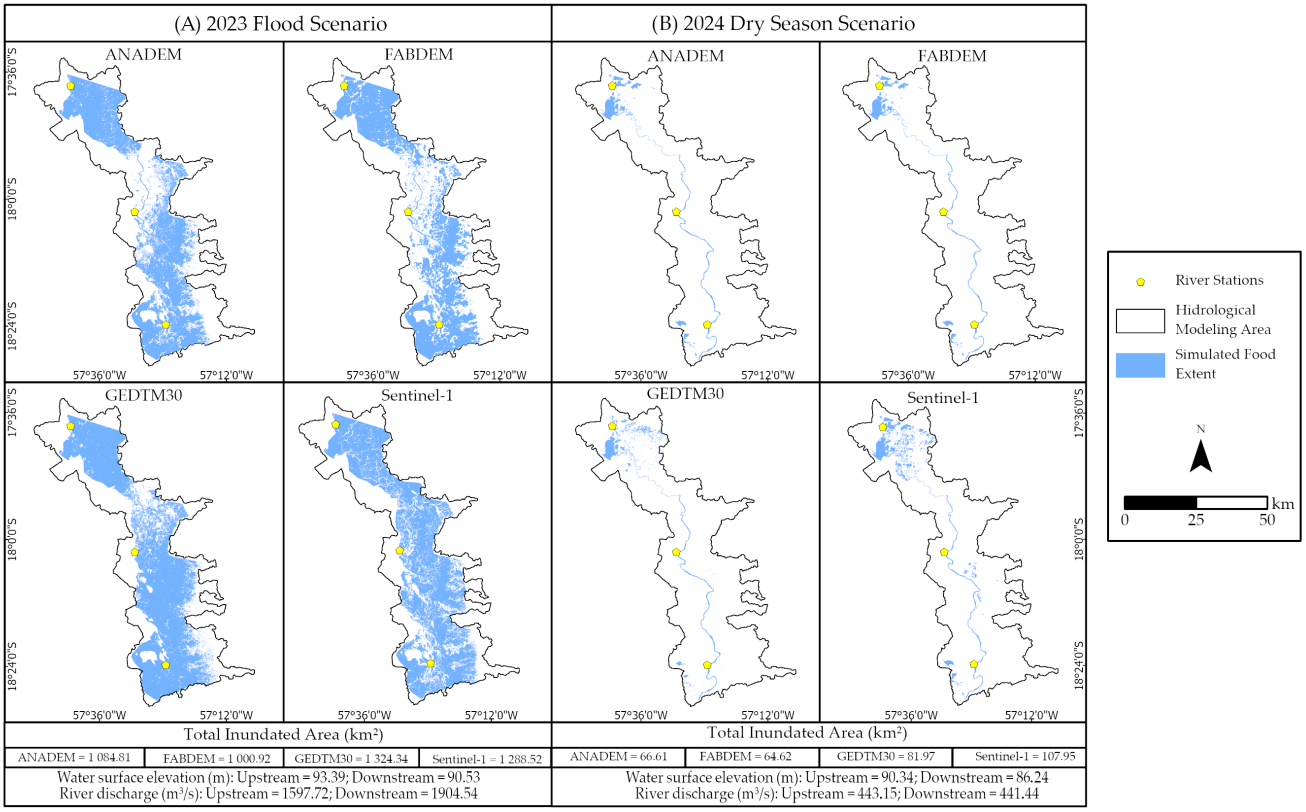

Figure 7. Influence of digital elevation model (DEM) selection on simulated inundation extents along the Paraguay River floodplain under contrasting hydrological conditions. (**A**) Simulated inundation extents for the 2023 flood scenario using ANADEM, FABDEM, GEDTM, and the Sentinel-1-derived DEM as terrain inputs to the HEC-RAS hydrodynamic model. (**B**) Corresponding inundation extents for the October 2024 low-water scenario using the same DEMs.

For the 2023 flood event, spatial agreement between simulated and satellite-derived inundation extents was moderate but significantly higher than under low-water conditions, with IoU values ranging from 23.71% to 29.71% (Table 6). The Sentinel-1-derived DEM presented the highest spatial agreement (IoU = 29.71%), followed by FABDEM (26.25%), ANADEM (25.03%), and GEDTM (23.71%). The satellite-derived reference inundation extent was approximately 610 km^2, whereas all four simulations produced larger inundated areas, indicating a systematic tendency to overestimate inundation extent. The spatial overlap between simulated and reference inundation ranged from 334.88 km^2 (FABDEM) to 434.77 km^2 (Sentinel-1-derived DEM). The Sentinel-1-derived DEM captured 71% of the reference inundation extent, compared with 55–61% for the other DEMs. Differences in IoU reflected variations in both correctly simulated inundated areas and FP predictions. FABDEM achieved higher IoU values than ANADEM and GEDTM despite producing the smallest simulated inundation extent, suggesting that its lower degree of overestimation contributed to better spatial agreement with the satellite-derived reference.

To account for the lateral extent limitations of the one-dimensional HEC-RAS model, the Sentinel-2-derived inundation mask was also clipped to the area bounded by the modeled cross-sections, reducing the reference inundation extent from approximately 610 to 500 km^2. Using this spatially constrained reference, IoU values increased by 1.8–2.5 percentage points across all DEMs, ranging from 25.5% (GEDTM) to 32.1% (Sentinel-1-derived DEM), without altering their rankings (Table 6). The proportion of the reference inundation extent correctly identified by the simulations increased to 87% for the Sentinel-1-derived DEM and 67–74% for the other DEMs. Nevertheless, simulated inundation extents remained 2.0–2.6 times

larger than the spatially constrained reference extent, indicating that overestimation of inundated areas remained a major contributor to the observed spatial disagreement.

Table 5. Simulated inundation extents and their spatial overlap with Sentinel-2-derived reference inundation masks for the 2023 flood and 2024 low-water scenarios, together with simulated inundation extents for the 1983 historical flood scenario. Simulated extent represents the total inundated area predicted by each digital elevation model (DEM). Intersection represents the area classified as inundated in both the simulation and the Sentinel-2 reference mask. Union full represents the combined inundated area using the complete Sentinel-2 reference mask, whereas Union clipped represents the corresponding area using the reference mask restricted to the polygon enclosing the Hydrologic Engineering Center's River Analysis System (HEC-RAS) cross-sections. Intersection values are identical for both reference masks because all simulated inundated areas fall within the clipping polygon. For the 1983 flood scenario, only simulated inundation extents are reported because no corresponding satellite-derived reference data were available (–).

Season	DEM	Simulated Extent (km^2)	Intersection (km^2)	Union Full (km^2)	Union Clipped (km^2)
Dry 2024	ANADEM	91.06	72.60	431.90	333.96
Dry 2024	FABDEM	90.55	72.89	431.10	333.16
Dry 2024	GEDTM	98.07	61.04	450.47	352.52
Dry 2024	Sentinel-1	143.14	101.82	454.76	356.82
Flood 2023	ANADEM	1084.81	339.27	1355.39	1245.48
Flood 2023	FABDEM	1000.92	334.88	1275.90	1165.99
Flood 2023	GEDTM	1324.34	370.66	1563.54	1453.63
Flood 2023	Sentinel-1	1288.53	434.77	1463.61	1353.70
Flood 1983	ANADEM	1659.54	-	-	-
Flood 1983	FABDEM	1663.33	-	-	-
Flood 1983	GEDTM	1664.83	-	-	-
Flood 1983	Sentinel-1	1637.84	-	-	-

For the 2024 dry-season scenario, spatial agreement between simulated and satellite-derived inundation extents was considerably lower than that observed for the 2023 flood scenario, with IoU values ranging from 13.55% to 22.39%. The Sentinel-1-derived DEM again achieved the highest spatial agreement (IoU = 22.39%), followed by FABDEM (16.91%) and ANADEM (16.81%), which presented comparable performance. GEDTM showed the lowest agreement (IoU = 13.55%). In contrast to the 2023 flood scenario, all four DEMs underestimated the satellite-derived reference water extent (413 km^2), producing simulated inundation extents ranging from 90.55 to 143.14 km^2. The spatial overlap between simulated and reference inundation extents ranged from 61.04 to 101.82 km^2, corresponding to 15–25% of the reference water extent.

Clipping the satellite-derived reference inundation mask to the area bounded by the modeled cross-sections reduced its extent from 413 to 316 km^2. Consequently, IoU values increased by 3.8–6.1 percentage points, ranging from 17.32% to 28.53%. The Sentinel-1-derived DEM again showed the highest spatial agreement (IoU = 28.53%), followed by FABDEM (21.88%), ANADEM (21.74%), and GEDTM (17.32%). Nevertheless, the simulations correctly identified only 19–32% of the spatially constrained reference inundation extent, while their simulated inundation extents (91–143 km^2) remained smaller than the reference extent. These findings indicate that the limited lateral extent of the modeled cross-sections alone cannot fully explain the low spatial agreement observed under dry-season conditions.

Table 6. Spatial agreement between Hydrologic Engineering Center's River Analysis System (HEC-RAS)-simulated inundation extents and Sentinel-2-derived reference inundation masks for the 2023 flood and 2024 low-water scenarios, using four digital elevation models (DEMs). Agreement was evaluated using the complete (Full) and spatially constrained (Clipped) reference masks. Intersection over Union (IoU) represents the ratio of the spatial intersection to the union of simulated and reference inundation extents, expressed as a percentage. False positives (FP) represent areas classified as inundated in the simulations but not in the reference masks, whereas false negatives (FN) represent areas classified as inundated in the reference masks but not in the simulations. FP values are identical for both reference masks because all simulated inundated areas fall within the clipping polygon, whereas FN values vary according to the spatial extent of the reference mask.

Reference	Season	DEM	IoU (%)	False Positive (km^2)	False Negative (km^2)
Full	Dry 2024	ANADEM	16.81	18.46	340.85
Full	Dry 2024	FABDEM	16.91	17.66	340.56
Full	Dry 2024	GEDTM	13.55	37.03	352.40
Full	Dry 2024	Sentinel-1	22.39	41.32	311.63
Full	Flood 2023	ANADEM	25.03	745.54	270.58
Full	Flood 2023	FABDEM	26.25	666.05	274.98
Full	Flood 2023	GEDTM	23.71	953.69	239.19
Full	Flood 2023	Sentinel-1	29.71	853.75	175.08
Clipped	Dry 2024	ANADEM	21.74	18.46	242.90
Clipped	Dry 2024	FABDEM	21.88	17.66	242.61
Clipped	Dry 2024	GEDTM	17.32	37.03	254.45
Clipped	Dry 2024	Sentinel-1	28.53	41.32	213.68
Clipped	Flood 2023	ANADEM	27.24	745.54	160.67
Clipped	Flood 2023	FABDEM	28.72	666.05	165.07
Clipped	Flood 2023	GEDTM	25.50	953.69	129.28
Clipped	Flood 2023	Sentinel-1	32.12	853.75	65.17

Decomposing spatial disagreement into FP and FN showed contrasting error patterns between the flood and dry-season scenarios. During the 2023 flood, FP areas predominated, ranging from 666.05 km^2 (FABDEM) to 953.69 km^2 (GEDTM), compared with FN areas of 65.17–165.07 km^2 when evaluated against the clipped reference inundation extent. Consequently, FP areas were approximately 4–13 times larger than FN areas. The Sentinel-1-derived DEM showed the lowest FN area (65.17 km^2), but not the lowest FP area (853.75 km^2), indicating that its higher IoU was mainly associated with larger spatial overlap with the reference inundation extent rather than reduced overprediction. In contrast, GEDTM presented the largest FP area and the lowest IoU. During the 2024 dry season, the opposite pattern was found, with relatively small FP areas (17.66–41.32 km^2) and significantly larger FN areas (213.68–254.45 km^2), 5–14 times higher than the corresponding FP areas. Because all simulated inundation extents were contained within the polygon enclosing the modeled cross-sections, clipping the reference mask affected only FN areas, reducing them by approximately 110 km^2 in 2023 and 98 km^2 in 2024 for each DEM, while leaving intersection and FP areas unchanged. Nevertheless, FN areas still accounted for 68–81% of the clipped reference inundation extent under dry-season conditions, highlighting the substantial underrepresentation of satellite-observed surface water in the simulations (Figure S2).

The poor spatial agreement observed during the dry season may not be attributable to DEM characteristics alone but may also reflect limitations of the steady-flow modeling approach under low-water conditions. During low flows, hydraulic connectivity between the main river channel and adjacent floodplain areas may be reduced or interrupted, allowing water to remain stored in isolated lagoons and topographic depressions independently of the prevailing river discharge. This residual water may originate from previous flooding

events and gradually decrease through infiltration and evaporation. Because the steady-flow model estimates water-surface elevations based on river discharge and specified boundary conditions, it cannot explicitly simulate the temporal dynamics of water storage and gradual depletion in hydraulically disconnected floodplain depressions. Consequently, the model may underestimate the extent of residual surface water during the dry season, contributing to the observed discrepancies between simulated and satellite-derived inundation extents. The simulated HEC-RAS cross-sectional profiles provide insights into this potential limitation (Figure 8).

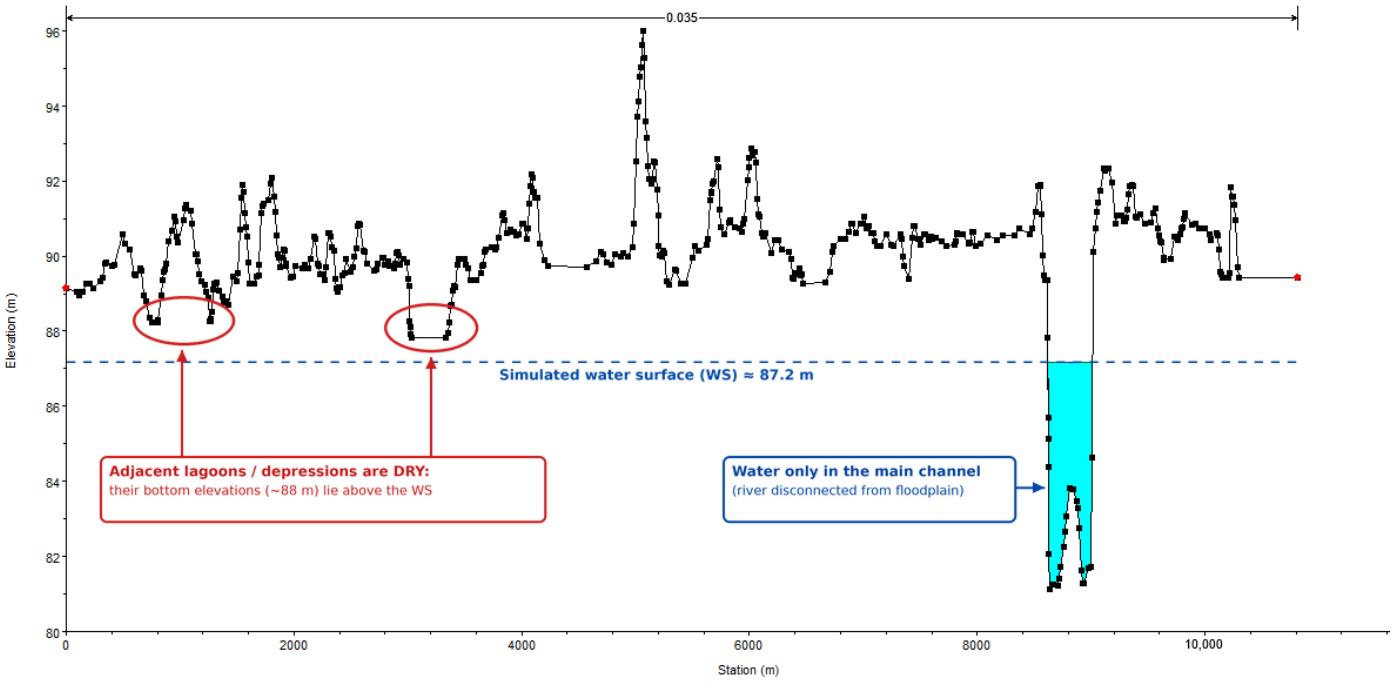

Figure 8. Representative Hydrologic Engineering Center's River Analysis System (HEC-RAS) cross-sectional profile illustrating simulated water-surface elevation under the 2024 dry-season steady-flow scenario. The simulated water surface (87.2 m; blue dashed line) remains confined to the main river channel, while adjacent floodplain lagoons and topographic depressions are not inundated because their bottom elevations (approximately 88–90 m) exceed the simulated water level.

Under the 2024 dry-season conditions, the simulations indicated that inundation was largely confined to the main river channel, while adjacent lakes and topographic depressions remained dry because the simulated water-surface elevation was below their bottom elevations (88–90 m). However, the extent of residual surface water in floodplain depressions during the dry season may depend more strongly on water levels reached during preceding flood events and subsequent water retention than on the prevailing river stage. Reproducing these conditions would require an unsteady-flow simulation initiated during a preceding flood or high-water period and continued through the recession into the dry season. Such an approach would allow the simulation of floodplain depressions filling during high-water conditions, becoming hydraulically disconnected from the main river channel as water levels decline, and progressively losing stored water through infiltration and evaporation, provided these processes are adequately represented in the model. In contrast, steady-flow simulations may reproduce observed inundation extents more accurately under high-water conditions, when elevated river stages promote hydraulic connectivity between the main channel and adjacent floodplain depressions, as observed during the 2023 flood (Figure S3). This interpretation is consistent with the higher spatial agreement obtained for that event,

although other sources of uncertainty may also have contributed to the differences between hydrological scenarios.

Finally, IoU quantifies the spatial agreement between simulated and satellite-derived inundation extents and should not be interpreted as a direct measure of DEM vertical accuracy. IoU values are also influenced by the hydrodynamic model formulation, boundary conditions, and uncertainties in the satellite-derived reference data. In addition, temporal mismatches between satellite image acquisition dates and the hydrological conditions represented in the simulations introduce additional uncertainty. Because each simulation represents a specific hydrological condition rather than the temporal evolution of inundation, IoU values are more appropriate for assessing the performance of DEMs under a consistent modeling framework than for determining the absolute accuracy of simulated inundation extents. Therefore, these comparisons provide a basis for evaluating DEM-dependent differences in simulated inundation patterns within each hydrological scenario but should not be considered definitive validation of their absolute spatial accuracy.

4. Discussion

The validation results indicate that DEM vertical accuracy in the Pantanal floodplain varies significantly depending on the elevation data source and LULC characteristics. Among the evaluated products, GEDTM, FABDEM, and ANADEM consistently presented the lowest elevation errors, suggesting that recently developed multi-source and terrain-corrected DEMs can provide more accurate estimates of terrain elevations in low-relief floodplain environments than several established global elevation products. These findings are consistent with recent advances in global terrain modeling, in which multi-source data integration, vegetation-related elevation corrections, and terrain refinement techniques have been applied to improve the accuracy of terrain elevation estimates [12,26].

ICESat-2 photon-counting laser altimetry provided a valuable reference for DEM validation owing to its high vertical precision and extensive spatial coverage. Nevertheless, ICESat-2-derived terrain elevations require careful interpretation, as terrain slope, LULC characteristics, vegetation structure, and uncertainties in photon classification can affect the accuracy of ground-elevation estimates. Rigorous data filtering and quality control procedures are essential to identify reliable ground-elevation observations, reduce measurement uncertainties, and ensure consistency between the reference elevations and the terrain surfaces represented by the evaluated DEMs [14].

Validation using the multi-source ground control dataset yielded higher RMSE values than validation using ICESat-2 terrain elevations. However, these differences should not necessarily be interpreted as evidence of lower reliability of ground-control observations. The ground control dataset comprises measurements collected by different institutions using various surveying techniques, acquisition periods, and spatial sampling strategies, whereas ICESat-2 provides a more internally consistent reference derived from a single satellite altimetry system. In addition, although ICESat-2 and most of the evaluated DEMs rely on remotely sensed observations, they employ different sensors and measurement principles. Consequently, discrepancies between the two validation approaches may reflect differences in acquisition methods, spatial sampling distributions, reference data uncertainties, and the terrain characteristics represented by each dataset. Nevertheless, the consistent ranking of GEDTM, FABDEM, and ANADEM among the most accurate DEMs across both validation approaches reinforces confidence in their comparatively high vertical accuracy.

The hydrodynamic simulations showed that the influence of DEM selection on simulated inundation extent varied significantly across hydrological conditions. Under the historical 1983 flood scenario, the four DEMs produced similar inundation extents, with differences of less than 2%. In contrast, the largest differences were observed during the 2024

dry-season scenario, when simulated inundation extents ranged from 90.55 km^2 (FABDEM) to 143.14 km^2 (Sentinel-1-derived DEM), corresponding to a difference of 58% in relation to that from FABDEM. These findings suggest that simulated inundation extent shows higher sensitivity to DEM selection under low-water conditions, when subtle variations in floodplain topography may exert a stronger influence on hydraulic connectivity and the spatial distribution of surface water. This interpretation is consistent with previous large-scale hydrodynamic studies highlighting the importance of low topographic gradients and complex floodplain storage processes in regulating floodwater propagation and redistribution across the Pantanal [11].

Under high-water conditions, elevated river stages promote hydraulic connectivity among channels, topographic depressions, and floodplain storage areas, potentially reducing the influence of local terrain elevation differences on simulated inundation extent. During the dry season, however, hydraulic connectivity may become increasingly dependent on subtle elevation thresholds separating adjacent floodplain compartments. Consequently, even small differences in DEM-derived elevations can affect whether channels and depressions remain hydraulically connected or become isolated from the main river system. This interpretation is consistent with previous studies highlighting the sensitivity of flood inundation simulations to terrain representation, particularly in low-relief environments where subtle topographic variations can significantly influence drainage pathways and floodplain connectivity [13,43].

GEDTM showed the highest overall vertical accuracy, whereas the Sentinel-1-derived DEM produced the largest simulated inundation extent during the 2024 dry season and achieved the highest spatial agreement with the Sentinel-2-derived surface water extent. In contrast, GEDTM showed the lowest spatial agreement. This apparent discrepancy highlights that point-based vertical accuracy metrics and inundation-based performance assessments capture different aspects of DEM quality. While RMSE quantifies the magnitude of vertical errors at sampled locations, simulated inundation patterns also depend on the spatial configuration and continuity of terrain elevations, which influence drainage pathways, topographic depressions, channel–floodplain connectivity, and hydraulic thresholds. These results are consistent with recent studies suggesting that conventional vertical accuracy metrics alone may be insufficient to assess DEM suitability for hydrological connectivity analysis and flood modeling in low-relief environments [13].

The low IoU values obtained using the full reference masks (23.71–29.71% for the 2023 flood and 13.55–22.39% for the 2024 dry season) should be interpreted in the context of differences between the processes represented by hydrodynamic simulations and satellite-derived surface water observations. Sentinel-2 images provide a snapshot of detectable surface water at the time of acquisition, potentially including water stored in isolated depressions, residual lakes, and temporarily disconnected floodplain compartments. In contrast, steady-flow HEC-RAS simulations represent inundation under specified discharge and boundary conditions, accounting for terrain geometry and hydraulic resistance within the modeled domain but not explicitly simulating the temporal dynamics of water storage and depletion in hydraulically disconnected floodplain areas.

Consequently, discrepancies between simulated and satellite-derived inundation extents cannot be attributed solely to DEM inaccuracies or hydrodynamic modelling errors. Instead, satellite-derived surface water observations and hydraulic simulations provide complementary but not necessarily equivalent representations of floodplain inundation dynamics [43]. Part of the spatial disagreement may therefore reflect the coexistence of hydraulically connected inundation and isolated surface water storage during flood recession and dry-season conditions in the Pantanal. Alternating phases of hydraulic connectivity, water storage, and isolation are characteristic of the flood-pulse dynamics of large river–

floodplain systems [44]. In the Pantanal, surface water may persist in floodplain lakes and topographic depressions even after hydraulic connections with the main river network have weakened or been interrupted. Consequently, satellite images may detect residual surface water that is not reproduced by steady-flow hydraulic simulations, which do not explicitly represent the temporal dynamics of water storage and gradual depletion in hydraulically disconnected floodplain areas.

The results obtained with the Sentinel-1-derived DEM warrant particular attention in this context. Its larger simulated inundation extent during the 2024 dry-season suggests that its terrain representation preserves low-elevation pathways or connections between floodplain depressions that are less clearly represented or are interrupted in the other DEMs. However, the higher IoU presented by the Sentinel-1-derived DEM should not be interpreted as evidence of superior vertical accuracy or overall DEM quality compared with GEDTM. Rather, it suggests that certain characteristics of its elevation surface may be more favorable for reproducing the observed spatial distribution of surface water under low-water conditions within the specific modeling framework adopted in this study. Further investigations involving elevation profiles, topographic depressions, and channel–floodplain connectivity are needed to identify the terrain characteristics that may explain these differences.

Overall, GEDTM, FABDEM, and ANADEM showed the highest vertical accuracy among the evaluated DEMs, although their relative rankings varied depending on the reference dataset. Validation based on ICESat-2-derived terrain elevations ranked GEDTM, FABDEM, and ANADEM in ascending order of RMSE, whereas validation using a multi-source ground control dataset ranked GEDTM, ANADEM, and FABDEM. According to the PEC-PCD classification, all three products achieved Class A at the 10 m contour interval and Class B at the 5 m interval. However, visual inspection indicated limitations in GEDTM's representation of local topographic features, including apparent surface irregularities and inadequate preservation of floodplain geomorphological features, despite its superior point-based vertical accuracy.

Regarding hydrodynamic performance, none of the evaluated DEMs consistently reached high spatial agreement with satellite-derived reference inundation extents, precluding the identification of a single optimal product. The Sentinel-1-derived DEM reached the highest IoU values during both the 2023 flood (29.7%) and the 2024 dry season (22.4%), although these values still indicate substantial spatial disagreement. This DEM was included in the hydrodynamic simulations partly because it was generated from relatively recent Sentinel-1 observations acquired in 2024, providing a terrain representation based on observations temporally closer to the Sentinel-2 reference image than those underlying some of the other DEMs. However, the publication years of ANADEM (2024), FABDEM (2022), and GEDTM (2025) do not necessarily reflect the acquisition dates of their underlying elevation data. The comparatively higher IoU values obtained with the Sentinel-1-derived DEM may partly reflect differences in the temporal characteristics of the elevation datasets, although this possibility remains speculative and requires further investigation.

Taken together, these findings highlight the importance of extending DEM evaluation for floodplain hydrodynamic modeling beyond conventional point-based vertical accuracy metrics. In low-relief environments such as the Pantanal, a comprehensive assessment should integrate quantitative validation of vertical accuracy, evaluation of terrain morphology and spatial continuity, and analysis of simulated inundation patterns under contrasting hydrological conditions [12,19]. The larger differences among DEM-based inundation extents during the 2024 dry season suggest that low-water conditions may provide an especially sensitive test of the influence of terrain representation on simulated floodplain connectivity. Under these conditions, subtle differences in terrain and topographic thresh-

olds may exert a stronger influence on the spatial distribution of inundation than during high-water periods.

Limitations of the Study

The findings of this study should be interpreted in light of several limitations associated with the datasets and methods employed. The evaluated DEMs, ICESat-2 ATL08 reference elevations, and MapBiomas LULC data were not acquired contemporaneously. The underlying elevation observations used to generate the evaluated DEMs span different periods, including SRTM acquisitions from 2000, whereas the publication dates of FABDEM (2022), ANADEM (2024), and GEDTM (2025) do not necessarily reflect the acquisition dates of their source elevation data. Similarly, ICESat-2 observations and LULC information represent different periods within the study interval. In a floodplain characterized by seasonal inundation and dynamic surface conditions, these temporal differences may contribute to elevation discrepancies, although their magnitude and potential influence on DEM rankings remain uncertain. The reference datasets are also subject to measurement uncertainties. The reported ICESat-2 vertical uncertainty of approximately 0.73 m is not negligible compared with the RMSE values of the best-performing DEMs (1.39–1.42 m). In addition, the multi-source ground-control dataset combines surveys conducted by different institutions using various measurement techniques, acquisition periods, and original vertical reference systems. This heterogeneity may contribute to the higher RMSE values obtained from ground control validation, although the relative contributions of these factors cannot be determined from the present analysis. Finally, the Sentinel-2-derived surface water masks were generated from individual satellite acquisitions using rule-based spectral thresholds. Because no independent classification accuracy assessment was conducted, uncertainties associated with surface water detection could not be quantified separately from discrepancies arising from the hydrodynamic simulations when interpreting the IoU values.

The hydrodynamic modeling approach imposes additional limitations on the interpretation and generalizability of the findings. Although the steady-flow configuration provides a simplified representation of inundation under high-water conditions, it cannot explicitly simulate water storage and gradual depletion in floodplain depressions following hydraulic disconnection from the main river channel. As discussed in Section 3.6, this limitation may have contributed significantly to the low spatial agreement observed during the dry season, although its relative importance compared with DEM-related uncertainties could not be quantified. In addition, the one-dimensional model limits the lateral extent of simulated inundation to the domain represented by the modeled cross-sections, while the cross-section envelope used to constrain the reference inundation masks provides only an approximation of this domain. Regarding generalizability, the analysis was conducted within a 5060 km^2 sector of the Paraguay River floodplain characterized by low topographic gradients and specific vegetation and hydrological conditions. Although similar DEM performance patterns may occur in other low-relief sectors of the Pantanal, the rankings obtained in this study should not be extrapolated to other regions, especially those characterized by steeper terrain or contrasting hydrological regimes, without independent validation. Likewise, the higher sensitivity of simulated inundation extent to DEM selection under low-water conditions was identified from a single dry-season scenario and requires confirmation across additional hydrological years and flow conditions.

5. Conclusions

This study provides an integrated assessment of DEM performance in the Pantanal floodplain by combining vertical accuracy validation using two independent reference

datasets, land cover-based analysis, and hydrodynamic simulations under contrasting hydrological conditions. The findings demonstrate that DEM suitability for applications in low-relief wetlands depends not only on overall vertical accuracy but also on the representation of floodplain terrain morphology and the specific hydrological conditions and processes under investigation.

GEDTM, FABDEM, and ANADEM consistently showed the lowest vertical errors across both independent reference datasets, indicating their superior point-based vertical accuracy within the study area. The consistency of these results between the ICESat-2 photon-counting laser altimetry and multi-source ground-control assessments reinforces confidence in the performance of these DEMs, despite differences in the acquisition methods and spatial distributions of the reference observations. DEM accuracy also varied across LULC classes, with generally higher vertical errors in densely vegetated areas and lower errors in wetlands and other open floodplain environments.

The hydrodynamic simulations indicated that the sensitivity of simulated inundation extent to DEM selection varied significantly across hydrological conditions. Differences among DEM-based inundation estimates remained below 2% under the 1983 high-water scenario but increased to 32% during the 2023 flood and 58% under the 2024 dry-season conditions. These findings suggest that simulated inundation extent shows higher sensitivity to DEM selection under low-water conditions, when subtle elevation differences and topographic thresholds may exert a stronger influence on hydraulic connectivity among channels, topographic depressions, and floodplain storage areas.

Importantly, the DEM with the highest vertical accuracy did not necessarily achieve the highest spatial agreement with satellite-derived inundation extents. Although GEDTM presented the lowest overall vertical errors, the Sentinel-1-derived DEM produced the largest simulated inundation extent during the 2024 dry season and reached the highest spatial agreement with the Sentinel-2-derived surface water extent, whereas GEDTM presented the lowest agreement. These results highlight the complementary nature of the three assessment approaches adopted in this study, each capturing distinct aspects of DEM performance. GEDTM showed the highest point-based vertical accuracy, Copernicus DEM and FABDEM provided the most coherent visual representations of floodplain morphology, and the Sentinel-1-derived DEM achieved the highest IoU values in the hydrodynamic comparisons. No single DEM consistently outperformed the others across all three assessment criteria, indicating that DEM selection should be guided by the specific requirements of the intended application. In low-relief floodplains, the spatial configuration and continuity of terrain elevations, especially the representation of potential hydraulic pathways and topographic barriers, may be as important as overall vertical accuracy for accurately simulating inundation patterns.

Overall, these findings highlight the importance of extending DEM evaluation for hydrodynamic applications in large tropical wetlands beyond conventional vertical accuracy metrics. A comprehensive assessment should integrate quantitative validation of vertical accuracy, analysis of LULC-related variations in DEM performance, assessment of terrain morphology and spatial continuity, and evaluation of simulated inundation patterns under contrasting hydrological conditions. Such an integrated approach provides a more robust basis for application-specific DEM selection in the Pantanal and other low-relief floodplain environments. Future research should investigate the relationships among terrain representation, hydraulic connectivity, and inundation dynamics by incorporating higher-resolution topographic and bathymetric data, temporally matched satellite-derived surface-water observations, and two-dimensional unsteady-flow hydrodynamic simulations.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18234116/s1, Figure S1: HEC-RAS inundation extents simulated for the 2024 dry season; Figure S2: Comparison between the HEC-RAS inundation extents simulated for the 2024 dry season and the Sentinel-2 surface-water mask; Figure S3: HEC-RAS inundation extents simulated for the 2023 flood season.

Author Contributions: Conceptualization, L.R.d.J., T.d.A. and R.E.C.; methodology, L.R.d.J., R.E.C. and T.d.A.; software, L.R.d.J.; validation, L.R.d.J., R.E.C. and G.G.P.; formal analysis, L.R.d.J., E.E.S., L.Z.M. and N.G.M.; investigation, L.R.d.J.; resources, T.d.A.; data curation, L.R.d.J.; writing—original draft preparation, L.R.d.J.; writing—review and editing, R.E.C., G.G.P., T.d.A., E.E.S., L.Z.M. and N.G.M.; visualization, L.R.d.J.; supervision, T.d.A. and R.E.C.; project administration, T.d.A. and R.E.C.; funding acquisition, T.d.A. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially financed by the Coordination for the Improvement of Higher Education Personnel (CAPES), Finance Code 001, and by the National Council for Scientific and Technological Development (CNPq), Brazil (Process No. 406339/2023-6). CNPq also provided research productivity fellowships for R.E.C. (Process No. 307055/2025-6), T.A. (Process No. 302843/2026-4) and E.E.S. (Process No. 307850/2023-4).

Data Availability Statement: The ICESat-2 ATL03 and ATL08 data are publicly available from NSIDC DAAC through NASA Earthdata Search. The multi-source validation data are available from DNIT, IBGE and INCRA. The Sentinel-2 imagery products are available at https://browser.dataspace.copernicus.eu/ (accessed on 10 June 2025) and processed through SNAP. Sentinel-2 images were accessed and processed through Google Earth Engine.

Acknowledgments: We thank Luana Oliveira Sales and Eduarda Araujo for their valuable support with the project's methodology.

Conflicts of Interest: Author Edson Sano was employed by Embrapa Cerrados. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANA	National Water and Sanitation Agency
DEM	Digital Elevation Model
DNIT	National Department of Transport Infrastructure
EP	Erro-Padrão (Standard Error)
FN	False Negative
FP	False Positive
GCP	Ground Control Point
GEE	Google Earth Engine
HEC-RAS	Hydrologic Engineering Center's River Analysis System
IBGE	Brazilian Institute of Geography and Statistics
ICESat-2	Ice, Cloud, and Land Elevation Satellite-2
INCRA	National Institute for Colonization and Agrarian Reform
IoU	Intersection over Union
LULC	Land Use and Land Cover
MAE	Mean Absolute Error
ME	Mean Error
NDWI	Normalized Difference Water Index
NIR	Near-Infrared
MNDWI	Modified Normalized Difference Water Index
PEC-PCD	Padrão de Exatidão Cartográfica para Produtos Cartográficos Digitais
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar
SD	Standard Deviation
SGB	Sistema Geodésico Brasileiro

SIGEF	Sistema de Gestão Fundiária (Land Management System)
SLC	Single Look Complex
SNAP	Sentinel Application Platform
TOPS	Terrain Observation by Progressive Scans

References

1. Bridgewater, P.; Kim, R.E. The Ramsar Convention on wetlands at 50. *Nat. Ecol. Evol.* **2021**, *5*, 268–270. [CrossRef] [PubMed]
2. da Costa Nunes, N.; Machado, N.G.; Barros-Rosa, L.; dos Santos, L.O.F.; Fernandes, F.A.C.; Biudes, M.S. Land use change and water loss in the Upper Paraguay River Basin: Trends, future scenarios, and implications for the Pantanal. *J. Environ. Manag.* **2025**, *394*, 127666. [CrossRef] [PubMed]
3. Ramos, M.D.; Merino, E.R.; Montes, C.R.; Gastmans, D.; Melfi, A.J. Flood frequency characterization of the Nhecolândia region (Pantanal) based on high spatial and temporal resolution PlanetScope satellite images. *Prog. Phys. Geogr. Earth Environ.* **2025**, *49*, 732–752. [CrossRef]
4. Guo, M.; Li, J.; Sheng, C.; Xu, J.; Wu, L. A Review of wetland remote sensing. *Sensors* **2017**, *17*, 777. [CrossRef] [PubMed]
5. Bergier, I.; Assine, M.L.; McGlue, M.M.; Alho, C.J.; Silva, A.; Guerreiro, R.L.; Carvalho, J.C. Amazon rainforest modulation of water security in the Pantanal wetland. *Sci. Total Environ.* **2018**, *619–620*, 1116–1125. [CrossRef] [PubMed]
6. Junk, W.J.; Piedade, M.T.F.; Schöngart, J.; Wittmann, F. A classification of major natural habitats of Amazonian white-water river floodplains (várzeas). *Wetl. Ecol. Manag.* **2012**, *20*, 461–475. [CrossRef]
7. Pompeu, J. Cross-boundary drivers of water cover reduction in the Pantanal wetland and implications for its conservation. *Wetlands* **2025**, *45*, 32. [CrossRef]
8. Boin, M.N.; Martins, P.C.S.; da Silva, C.A.; Salgado, A.A.R. Pantanal: The Brazilian wetlands. In *The Physical Geography of Brazil*; Springer: Cham, Switzerland, 2019; pp. 75–91. [CrossRef]
9. Assine, M.L.; Merino, E.R.; Pupim, F.N.; Warren, L.V.; Guerreiro, R.L.; McGlue, M.M. Geology and geomorphology of the Pantanal Basin. In *Dynamics of the Pantanal Wetland in South America*; Springer: Cham, Switzerland, 2015; pp. 23–50. [CrossRef]
10. Milien, E.J.; Nunes, G.M.; Pierre, G.; Hamilton, S.K.; Da Cunha, C.N. Hydrological dynamics of the Pantanal, a large tropical floodplain in Brazil, revealed by analysis of Sentinel-2 satellite imagery. *Water* **2023**, *15*, 2180. [CrossRef]
11. da Paz, A.R.; Collischonn, W.; Tucci, C.E.M.; Padovani, C.R. Large-scale modelling of channel flow and floodplain inundation dynamics and its application to the Pantanal (Brazil). *Hydrol. Process.* **2011**, *25*, 1498–1516. [CrossRef]
12. Lin, Y.; Fan, C.; Liu, K.; Liu, L.; Deng, X.; Xu, P.; Zhan, P.; Fu, X.; Song, C. Characterizing the capability of public DEMs for mapping global floodplain bathymetry. *J. Hydrol.* **2025**, *658*, 133205. [CrossRef]
13. Fragozo Arevalo, J.M.; Escobar Villanueva, J.R.; Pérez-Montiel, J.I. Beyond vertical accuracy: Benchmarking global DEMs for hydrologic connectivity and flood sensitivity in flat coastal plains. *Hydrology* **2026**, *13*, 74. [CrossRef]
14. Shahabi, E.; Jafari, M.; Taheri Dehkordi, A. A new framework for vertical accuracy assessment and spatial rrror mapping of global digital elevation models using ICESat-2. *J. Geovis. Spat. Anal.* **2026**, *10*, 2. [CrossRef]
15. Huang, J.; Yang, X. Evaluation and improvement of the vertical accuracy of the global open DEM under forest environment. *Geocarto Int.* **2025**, *40*, 2453024. [CrossRef]
16. de Morais, R.G.; da Cruz, I.F.; da Silveira, S.W.G.; Zeilhofer, P. Mapping flood agents in the Northern Pantanal Wetland using multiple spatio-temporal GIS and remote sensing techniques. *Wetlands* **2026**, *46*, 70. [CrossRef]
17. de Magalhães Neto, N.; Evangelista, H. Human activity behind the unprecedented 2020 wildfire in Brazilian wetlands (Pantanal). *Front. Environ. Sci.* **2022**, *10*, 888578. [CrossRef]
18. MapBiomas Project. MapBiomas Collection 9: Brazilian Land Use and Land Cover Map Series (1985–2023). 2024. Available online: https://mapbiomas.org/ (accessed on 10 June 2025).
19. Toledo Salgado, S.R.; da Silva Carvalho, E.M.; Viseu, M.T.; de Oliveira, O.F. Comparative analysis of global DEMs for dam-break flood modeling and inundation mapping. *Water Resour. Manag.* **2026**, *40*, 98. [CrossRef]
20. Helmy, M.; Mandanici, E.; Vittuari, L.; Bitelli, G. Global digital terrain models: Evaluation of vertical accuracy against different morphological contexts. *Prog. Phys. Geogr. Earth Environ.* **2025**, *50*, 640–654. [CrossRef]
21. Weifeng, X.; Jun, L.; Dailiang, P.; Jinge, J.; Hongxuan, X.; Hongyue, Y.; Jun, Y. Multi-source DEM accuracy evaluation based on ICESat-2 in Qinghai-Tibet Plateau, China. *Int. J. Digit. Earth* **2024**, *17*, 2297843. [CrossRef]
22. Talchabhadel, R.; Nakagawa, H.; Kawaike, K.; Yamanoi, K.; Thapa, B.R. Assessment of vertical accuracy of open source 30 m resolution space-borne digital elevation models. *Geomat. Nat. Hazards Risk* **2021**, *12*, 939–960. [CrossRef]
23. Laipelt, L.; Comini de Andrade, B.; Collischonn, W.; de Amorim Teixeira, A.; de Paiva, R.C.D.; Ruhoff, A. ANADEM: A digital terrain model for South America. *Remote Sens.* **2024**, *16*, 2321. [CrossRef]
24. Airbus Defence and Space GmbH. *Copernicus DEM—Global and European Digital Elevation Model (COP-DEM)*; GLO-30, Version 2022_1; European Space Agency: Paris, France, 2022. [CrossRef]

25. Hawker, L.; Uhe, P.; Paulo, L.; Sosa, J.; Savage, J.; Sampson, C.; Neal, J. A 30 m global map of elevation with forests and buildings removed. *Environ. Res. Lett.* **2022**, *17*, 024016. [CrossRef]
26. Ho, Y.-F.; Grohmann, C.H.; Lindsay, J.; Reuter, H.I.; Parente, L.; Witjes, M.; Hengl, T. GEDTM30: Global ensemble digital terrain model at 30 m and derived multiscale terrain variables. *PeerJ* **2025**, *13*, e19673. [CrossRef] [PubMed]
27. NASA JPL. *NASADEM Merged DEM Global 1 arc Second V001*; NASA EOSDIS Land Processes Distributed Active Archive Center: Sioux Falls, SD, USA.
28. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. [CrossRef]
29. Rizzoli, P.; Martone, M.; Gonzalez, C.; Wecklich, C.; Tridon, D.B.; Bräutigam, B.; Bachmann, M.; Schulze, D.; Fritz, T.; Huber, M.; et al. Generation and performance assessment of the global TanDEM-X digital elevation model. *ISPRS J. Photogramm. Remote Sens.* **2017**, *132*, 119–139. [CrossRef]
30. Mohammadi, A.; Karimzadeh, S.; Jalal, S.J.; Kamran, K.V.; Shahabi, H.; Homayouni, S.; Al-Ansari, N. A Multi-sensor comparative analysis on the suitability of generated DEM from Sentinel-1 SAR interferometry using statistical and hydrological models. *Sensors* **2020**, *20*, 7214. [CrossRef] [PubMed]
31. Neuenschwander, A.; Guenther, E.; White, J.C.; Duncanson, L.; Montesano, P. Validation of ICESat-2 terrain and canopy heights in boreal forests. *Remote Sens. Environ.* **2020**, *251*, 112110. [CrossRef]
32. Braun, A. One-day repeat pass interferometry highlights the role of temporal baseline on digital elevation models retrieved from Sentinel-1. *Geosci. Instrum. Methods Data Syst.* **2026**, *15*, 39–52. [CrossRef]
33. Yan, L.; Li, H.; Hao, X.; Wang, J.; Yin, Z.; Liu, J. Generation and assessment of digital elevation models by combining Sentinel-1A and Sentinel-1B data in mountain glacier area. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 642–654. [CrossRef]
34. Salameh, E.; Frappart, F.; Turki, I.; Laignel, B. Intertidal topography mapping using the waterline method from Sentinel-1 & -2 images: The examples of Arcachon and Veys Bays in France. *ISPRS J. Photogramm. Remote Sens.* **2020**, *163*, 98–120. [CrossRef]
35. European Space Agency (ESA). SNAP—Sentinel Application Platform. Available online: https://step.esa.int/main/toolboxes/snap/ (accessed on 16 February 2025).
36. Brasil. *Norma da Especificação Técnica Para Controle de Qualidade de Dados Geoespaciais (ET-CQDG)*; Ministério da Defesa: Brasília, Brazil, 2016.
37. de Souza, M.; Siqueira, H.L.; Araujo, M.S.; Oliveira, L.; Gonçalves, W.N.; Ramos, A.P.M.; Junior, J.M. Accuracy of high resolution digital cartographic products with elevation control points. *Anu. Inst. Geociênc.* **2025**, *48*, e59100. [CrossRef]
38. de França, L.L.S.; da Penha, A.d.L.T.; de Carvalho, J.A.B. Comparison between absolute and relative positional accuracy assessment—A case study applied to digital elevation models. *Bol. Ciênc. Geodésicas* **2019**, *25*, e2019003. [CrossRef]
39. Bettiol, G.M.; Ferreira, M.E.; Motta, L.P.; Cremon, É.H.; Sano, E.E. Conformity of the NASADEM_HGT and ALOS AW3D30 DEM with the altitude from the Brazilian geodetic reference stations: A case study from Brazilian Cerrado. *Sensors* **2021**, *21*, 2935. [CrossRef] [PubMed]
40. Engelhardt, R.C.B.; da Silva, A.F.; da Nobrega, S.L. Simulação de inundação utilizando software HEC-RAS e modelo de elevação MERIT, análise de situação em Boa Vista, Roraima, Brasil. *Contrib. Cienc. Soc.* **2024**, *17*, e6101. [CrossRef]
41. Bates, P.D.; Horritt, M.S.; Fewtrell, T.J. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *J. Hydrol.* **2010**, *387*, 33–45. [CrossRef]
42. Hawker, L.; Neal, J.; Bates, P. Accuracy assessment of the TanDEM-X 90 digital elevation model for selected floodplain sites. *Remote Sens. Environ.* **2019**, *232*, 111319. [CrossRef]
43. Schumann, G.; Bates, P.D.; Horritt, M.S.; Matgen, P.; Pappenberger, F. Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. *Rev. Geophys.* **2009**, *47*, RG4001. [CrossRef]
44. Junk, W.; Bayley, P.B.; Sparks, R.E. The flood pulse concept in river-floodplain systems. In *Proceedings of the International Large River Symposium (LARS)*; Dodge, D.P., Ed.; Department of Fisheries and Oceans: Ottawa, ON, Canada, 1989; pp. 110–127.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.