https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

Improving the freshwater fluxes in global oceanic reanalyses

Thomas Vaujour[1], Louis Kern[1], Andrea Storto[2], Julia Pfeffer[1], Chloé Thenoz[1], Chunxue Yang[2], and Stéphanie Guinehut[3]

[1]Magellium, Earth Observation Unit, 1 rue Ariane 31520 Ramonville-Saint-Agne FRANCE
[2]CNR-ISMAR, Rome ITALY
[3]Mercator Ocean international, Toulouse FRANCE

Correspondence: Thomas Vaujour (thomas.vaujour@magellium.fr)

Abstract. Freshwater fluxes have a notable impact on the ocean by adding water mass and changing local salinity. However, most ocean modeling systems currently rely on climatological estimates of the river discharge, which rely on approximations leading to unrealistic estimates. This study presents F3O (Freshwater Fluxes in Ocean Forecasts), a hybrid river discharge dataset developed using a recurrent deep learning neural network that corrects the river discharge from the Copernicus Emergency Service Global Flood Awareness System (GloFAS), based on in situ and satellite measurements. The resulting F3O dataset provides a hybrid river discharge estimate with high accuracy, high temporal and spatial resolution, and global coverage. The F3O dataset improves the representation of discharge biases, seasonal cycles, and interannual variability, providing enhanced freshwater forcing for ocean models. Implemented in ocean model systems that represent reanalyses products, F3O significantly improves sea surface salinity skill scores, primarily through the correction of the mean discharge rather than its interannual variability. Reducing the freshwater-induced salinity bias also leads to a deepening of the mixed layer and thermocline, and the resulting weakening of stratification contributes to stronger Atlantic subtropical gyres and Antarctic Circumpolar Current.

Short summary: Freshwater fluxes impact the ocean by adding water and changing salinity. To improve the freshwater forcing of ocean models, this study uses artificial intelligence techniques to correct the errors of a hydrological model. The resulting F3O dataset provides a hybrid river discharge estimate with high accuracy, high temporal resolution, and global coverage. It has been implemented in ocean model systems and the results show a significant improvement in the model performance.

Keywords. LSTM, River discharge, Oceanic reanalyses, Freshwater fluxes

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

1 Introduction

The ocean plays a dominant role the global water cycle (Schmitt (1995)), exchanging water with the atmosphere, the continents and the cryosphere. Freshwater enters the ocean through precipitation, river discharge, submarine groundwater discharge and ice melt, while leaving through evaporation and, indirectly, through atmospheric water vapor transport and sea ice formation (Sohail and Zika (2025)). The balance between freshwater inputs and outputs determines ocean freshwater content and influences Sea Surface Salinity (SSS), which provides an important constraint on the oceanic water cycle. By modifying salinity and, consequently, seawater density, freshwater fluxes affect upper-ocean stratification and the density gradients that drive ocean circulation. Changes in stratification also regulate vertical mixing and heat exchange between the ocean and atmosphere, thereby influencing climate variability (Thouvenin-Masson et al. (2024)). Through these interconnected processes, freshwater fluxes contribute to the regulation of ocean circulation, global water transport, and climate variability (Trenberth et al. (2007), Schanze et al. (2010), Durack (2015)).

Despite this crucial importance, current ocean reanalyses often rely on river discharge products that exhibit large spatial and temporal biases compared to independent observations (Sohail and Zika (2025)). Many state-of-the-art Ocean General Circulation Models (OGCMs) still utilize climatological river discharge products—such as those by Dai et al. (2009)—which do not contain the interannual variability necessary for realistic SSS and density simulations (Thouvenin-Masson et al. (2024)). Several recent studies have demonstrated the benefits of replacing static climatological runoff with time-varying river discharge in ocean model simulations. For instance, Feng et al. (2021) showed, within the ECCOv4 framework, that daily-varying discharge leads to markedly improved sea surface salinity near major river plumes and more realistic subsurface stratification. At the regional scale, Thouvenin-Masson et al. (2024) and Tajouri et al. (2024) reinforced these findings, highlighting improved representations of SSS variability and sea level when time-varying freshwater inputs are used. These results provide a background motivation for developing improved river discharge datasets and assessing their impact within ocean reanalysis systems.

Existing river discharge datasets have known limitations. For example, in situ data are often sparse, delayed, or unavailable in some regions. They provide accurate discharge measurements but cannot be used in operational frameworks. While systems like the Global Flood Awareness System (GloFAS) provide global coverage and temporal availability, they can suffer from linear trends inherited from atmospheric reanalysis forcing (e.g., ERA5) and may lack the accuracy needed for local-scale coastal dynamics (Senent-Aparicio et al. (2021)). These atmospheric reanalyses can also be used in the resolution of the water-mass balance equation (Syed et al. (2005), Chen et al. (2020)) to estimate the river discharge at basin scale. However, even though this method can yield accurate estimates for some major river basins, such as the Amazon, the above-mentioned atmospheric biases can be propagated and produce unreliable estimates for other rivers, such as the Congo basin (Kern et al. (2026)). Correcting the biases and errors in GloFAS would combine the strengths of these different datasets.

In recent years, Deep Learning (DL) approaches have increasingly been applied to river discharge estimation, leveraging ancillary atmospheric and hydrological datasets to produce high-resolution discharge estimates. Most applications, however, focus on improving discharge estimates inland for flood monitoring rather than constraining freshwater fluxes at river mouths,

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



where they directly enter the ocean. Traditional physics-based hydrological models are computationally expensive and require substantial calibration, particulaly at global scales. In contrast, DL models, can efficiently learn the relationship between atmospheric forcing and river discharge without requiring explicit parameterization of the underlying hydrological processes(Fathi et al. (2025)). Among these approaches, Long Short Term Memory (LSTM) have proven effective for river discharge estimation because their recurrent architecture can represent temporal dependencies and nonlinear relationships in time series (Leščešen et al. (2025), Yang et al. (2025), Fathi et al. (2025), Ruparell et al. (2025)). LSTMs can therefore capture the delayed and nonlinear response of runoff to atmospheric forcing, particularly precipitation. For example, Leščešen et al. (2025) demonstrate that LSTMs achieve significantly higher Nash-Sutcliffe Efficiency (NSE) values (often > 0.90) compared to traditional models because they better capture non-linear, long-term temporal dependencies between precipitation and runoff without requiring explicit physical parameterization. At the global scale, Yang et al. (2025) developed the GRADES-hydroDL framework, combining a Grid LSTM with the RAPID model to estimate daily discharge for 2.94 million rivers. The framework achieves improved performance compared with previous global discharge products, with relatively consistent skill across regions. The median KGEs for their test experiments however stay around 0.7, which is similar to GloFAS for the world major rivers. Although such performance is promising for hydrological applications, further improvements in discharge accuracy, particularly at river mouths, are needed for applications requiring realistic freshwater fluxes as inputs to global ocean reanalyses.

This study aims to improve the river discharge datasets used as freshwater forcing in ocean models. To achieve this goal, LSTM networks are used to correct GloFAS discharge estimates at the river outlet and generate a global river discharge dataset at a daily resolution for ocean applications. A key feature of our approach is its focus on the discharge at the outlet and on assessing its impact on oceanic simulations. The F3O (Freshwater Fluxes in Ocean Forecasts) dataset corrects GloFAS river discharge estimates based on in situ measurements and atmospheric data. It provides river discharge estimates for 41 rivers worldwide. With continuous daily coverage, the F3O dataset extends the temporal coverage of in situ observations and provides a consistent daily resolution. It also enables the characterization of freshwater fluxes variability beyond the seasonal cycle, accounting for intra-seasonal changes, interannual changes and decadal variations. Its impact on ocean reanalyses has been estimated using free ocean model simulations configured to represent the CIGAR reanalysis system (Storto and Yang, 2024), comparing the standard Dai et al. (2009) monthly runoff climatology with the new hybrid LSTM-based discharge dataset. The experiments use a free-running ocean configuration enabling a direct assessment of the response of the ocean to freshwater forcing, without the effect of the observation-driven corrections brought by data assimilation.

2 Data

2.1 Rivers under study

The objective is to assess the impact of freshwater fluxes on the global ocean and to develop a global daily river discharge dataset designed to improve freshwater forcing in ocean reanalyses relative to standard climatological estimates.

A total of 41 major rivers were selected based on several criteria. The first criterion was direct discharge into the ocean, excluding rivers flowing into enclosed seas as well as tributaries of large rivers. In addition, since we wanted to look at the

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

global scale, we selected only the most important rivers in terms of discharge. The mean river discharge had to be above 2000 m^3/s. The only exception to this rule is the Orange river that was selected because its basin is widely spread over the south of Africa, making it a major African river.

Then, there was a need for availability of relatively recent in situ measurements close to the river outlet. To do so, for each selected river, we looked at the available in situ data at the most downstream in situ stations. There was also a criterion of temporal overlap with other geophysical datasets (described hereafter). In practice, if no in situ data was available from 1989 to present (temporal coverage of GloFAS), the river was discarded.

In the end, 41 major rivers worldwide flowing directly into the ocean have been selected. Their basin are displayed on the map in Figure 1.

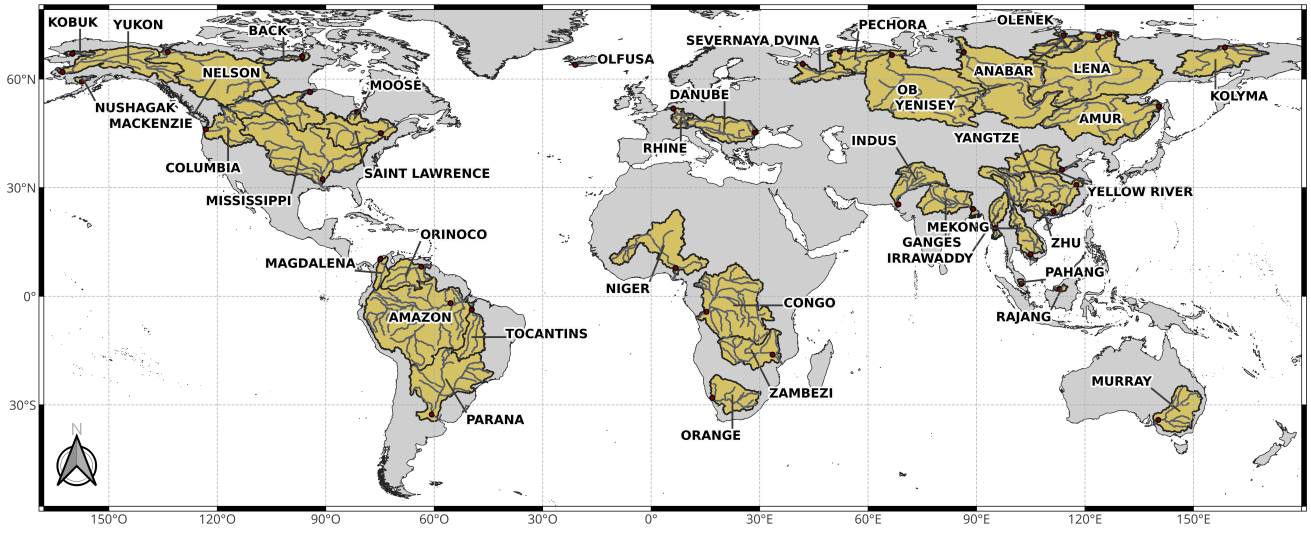

Figure 1. The 41 selected river basins. The red points represent the most downstream in situ station for each river.

2.2 In situ measurements

In situ measurement were taken from three different sources described in Kern et al. (2026): the Global Runoff Data Center (GRDC (2020)), the HYBAM project (HYdro-géochimie du Bassin AMazonien, noa) and Dai and Trenberth dataset (Dai (2017)).

The temporal coverage of in situ measurements can significantly vary from one river to another. The median temporal coverage from 1989 is 17 years, with substantial variability among rivers, ranging from 1 year for the Zhu River to 34 years for the Saint Lawrence River. Figure A1 in Appendix A shows the high variability of the temporal coverage in in situ measurements availability for the 41 selected rivers.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.


For each river, one in situ gauge station has been selected, as close as possible to the river outlet. As explained in Section 2.1, the stations' selection also relied on data availability and temporal coverage.

2.3 GloFAS

The Global Flood Awareness System (GloFAS), jointly developed by the European Commission and the European Centre for Medium-Range Weather Forecasts (ECMWF), is a global hydrological forecast and monitoring system. It is a combination of weather forecasts and a hydrological model. GloFAS produces daily flood forecasts and historical daily discharge data, as well as other hydrological variables such as the snow depth water equivalent and the soil wetness index. The system has been fully operational as part of the Copernicus Emergency Management Service since April 2018. GloFAS data were accessed from the Copernicus Emergency Management Service (CEMS) Early Warning Data Store (EWDS) (Harrigan et al. (2020), last accessed on 01-04-2026).

In this study, the GloFAS reanalysis version 4.0 (Grimaldi et al. (2022)) has been used to estimate the river discharge at a daily resolution at the location of the in situ gauge stations as well as close to the rivers' outlet. GloFAS v4.0 uses the LISFLOOD hydrological model (PA et al. (2013)). As forcing, the LISFLOOD model uses historical hydro-meteorological records from ERA5 (Hersbach et al. (2020)). GloFAS v4.0 uses the ECMWF-C3S ERA5 reanalysis as a proxy meteorological observation and a range of hydrological observational records, managed jointly by the JRC and the Computational centre (COMP). The above-mentioned physical variables are available daily on a 0.05° spatial grid from 1979 to 2025.

We used the last available version of GloFAS at the beginning of the study, v4.0, taking the mean discharge in the last 24h, with a temporal coverage going from January 1989 to December 2023.

In order to derive the 1D river discharge timeseries, we selected the points on the GloFAS grid that were the closest to the location of the gauge station and the location of the outlet.

2.4 ERA5

ERA5 (Hersbach et al. (2020)) is a global atmospheric reanalysis providing estimations of geophysical variables from 1940 to present.

In this study, we used the ERA5 post-processed daily statistics on single levels dataset from Hersbach et al. (2023). It also has a daily temporal resolution and covers January 1989 to December 2023. We selected the following variables: 2m dewpoint temperature, 2m temperature, 10m u component of wind, 10m v component of wind, evaporation, snowfall, surface net solar radiation, surface net thermal radiation, surface pressure and total precipitation. These geophysical variables have been selected based on several similar studies from the literature (Nearing et al. (2023), Tao et al. (2024), Snieder and Khan (2025), Börgel et al. (2025), Hunt et al. (2022)). These studies most of the time use this list of variables because they are the main drivers of the river discharge evolution.

We applied some preprocessing to the data. First, the variables were either summed or averaged, depending on their intensive or extensive physical nature, over the basin area. For example, the precipitations were summed but the air temperature was

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

averaged. Then, a low-pass filtering with a cutoff frequency of 30 days was applied to the time series in order to alleviate the high-frequency noise, having a detrimental impact on the neural network performance.

3 Methods

3.1 Water Mass Balance approach

The Water Mass Balance (WMB) approach consists in estimating the river discharge by solving the water mass balance equation: $\frac{dS}{dt} = P - ET - R$, where R is the basin river discharge, $\frac{dS}{dt}$ are the changes in land water storage with time, P the precipitations and ET the evapotranspiration. The land water storage changes are estimated from observations provided by the GRACE and GRACE-FO satellite missions (Tapley et al. (2004) and Landerer et al. (2020)), while precipitation and evapotranspiration are estimated with the ERA5 atmospheric reanalysis integrated over the river basin.

A detailed description and evaluation of the WMB approach is provided inKern et al. (2026). The WMB approach provides reliable river discharge estimates for a limited number of basins, notably the Amazon and Tocantins, but its performance is not sufficiently consistent across basins to justify its systematic use in the present study. It is, however, retained for the Amazon River because of its exceptional contribution to global freshwater discharge to the ocean. As the largest river in terms of freshwater discharge, even relatively small errors in the Amazon discharge can have a disproportionate impact on the simulated global ocean state. The Amazon WMB estimate is therefore included as an additional constraint and is described in Section 3.3.

3.2 LSTM

The LSTM model corrects GloFAS discharge estimates with respect to in situ measurements used as ground truth (details in Section 3.3). Both the GloFAS river discharge and atmospheric forcing are used as inputs

The final objective is to correct GloFAS discharge at the outlet. However, we do not have in situ measurements at the outlet, we only have in situ measurements at gauge stations. This means we can only generate ground truth data at the stations. The strategy therefore consists in first training the neural network at the station and then adapting its inputs to the outlet to make it predict the outlet discharge. This method is illustrated in Figure 2 with the Amazon basin. As we see in panel **a**, the catchment area of the most downstream in situ station (Obidos in the case of the Amazon river) is significantly (about 20 percent) smaller than the total catchment area of the river. As a result, the model's input variables will be different for these two areas. Examples are shown in panel **b** for GloFAS discharge and in panel **c** for ERA5 evaporation. Since our ground truth data is only available at Obidos, we have to train the LSTM model at Obidos. The network will learn how to correct GloFAS with ERA5 atmospheric variables. Then, in order for the network (trained at the last in situ station) to predict the river discharge at the outlet, we use as inputs the GloFAS discharge and the ERA5 variables corresponding to the outlet.

The LSTM model was implemented with the Pytorch library. It is a model built as a combination of a single LSTM layer followed by a fully connected layer. The network takes time sequences of 10 days as inputs and returns the river discharge

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

estimation for the day following the 10-day sequence. If there are data gaps longer than 10 days, the corresponding sequences are not used for training. For each river, the time series have been split: the first two thirds were dedicated to the training and the last third to the test.

Some sensitivity tests were conducted to assess the impact of the model hyperparameters and the choice of ERA5 inputs. The results showed that the hyperparameters have a negligible effect on LSTM performance whereas the choice of ERA5 inputs significantly affects the network's ability to correct GloFAS. The most useful atmospheric variables were the evaporation, the precipitation and the solar radiations whereas the snowfall and the air temperature were found less relevant. One LSTM network was trained for each river. While the same hyperparameters were used for all rivers, the training period and the selection of ERA5 inputs were tailored to each river. As a result, each LSTM model is river-specific and is intended to be applied only to the river for which its was trained.

As we train one model per river, the models are light, with relatively few training data, so the complete process, i.e. data processing, training and inference, lasts less than 300 s in CPU time, which is relatively quick compared to many Deep Learning applications requiring several days of training.

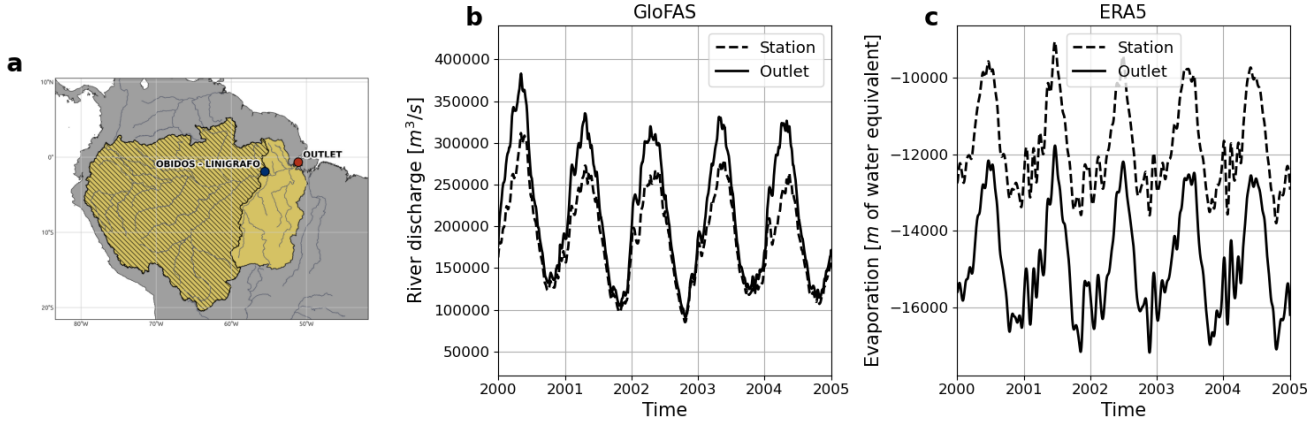

Figure 2. The LSTM strategy to generate discharge data at the outlet. The examples are given for the Amazon river for illustration purposes. Panel **a** describes the difference between the gauge station (Obidos) drainage area and the total basin area, panel **b** the difference between the GloFAS discharge at the station and at the outlet and panel **c** the difference in ERA5 evaporation data (summed over the basin area).

3.3 Ground truth: adjusted monthly means and WMB

Our objective is to generate daily river discharge estimates to use as a ground truth in our DL model. Our reference data are the monthly in situ measurements at the most downstream in situ stations. For the ground truth data, we want time series as close as possible to in situ measurements, but with a daily temporal resolution. In order to get daily data that fit to the monthly means of in situ, we adjust the monthly means of GloFAS on in situ ones while keeping the high-frequency variations. This new method is called the Adjusted Monthly Means (AMM) method. In summary, the AMM method combines the monthly

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



accuracy of in situ discharge measurements with the high-frequency variability captured by GloFAS, producing daily discharge estimates that remain consistent with observed monthly means.

The AMM method is illustrated in Figure 3. The first step (panel **a**) is to compute the monthly mean values of the GloFAS time series and to oversample at a daily resolution the in situ data. The second step (panel **b**) is to compute the ratio between the monthly in situ time series and the monthly mean values from GloFAS. This ratio is discontinous so we smooth it with a 31-day moving average (panel **b**). Then, this smoothed ratio is multiplied with the original daily GloFAS time series to get the corrected AMM time series (panel **c**), that we use as the ground truth to train the LSTM network.

For some Arctic rivers, this method produces high discharge peaks due to occasional very low GloFAS discharge. It is a purely numerical artifact caused by a denominator close to zero. Indeed, a low GloFAS discharge implies a high ratio that, when smoothed, is multiplied with high discharge values (corresponding to the beginning of the melting season), leading to unrealistically high discharge AMM values. To avoid this phenomenon, the ratio is set to 1 before being smoothed if the GloFAS discharge is below 10 m^3/s. For the other rivers, this patch is transparent and does not modify the AMM discharge time series.

The main restriction of the AMM correction is its temporal coverage. As the ratio requires in situ measurements to be computed, the correction can only be applied over periods for which in situ data are available, which is very limited for some rivers.

The Amazon River requires specific processing. The river discharge estimated with GloFAS at Obidos, which is the most downstream station, is subject to substantial biases and errors, with a different source and behavior than for the rest of the basin. As a result, the LSTM discharge at the outlet exhibit substantial errors in the representation of the annual cycle (see Figure 6, explained later). We therefore developped a specific processsing for the Amazon river to obtain a reliable estimate of the discharge at the outlet. Kern et al. (2026) validated the WMB approach for the Amazon basin through comparison with independent in situ measurements at Obidos. The WMB approach can therefore be used to estimate accurately the Amazon river discharge at the outlet, where no in situ station is available. This WMB time series were smoothed with a low-pass filter with a cutoff period of 5 months in order to remove high-frequency noise, before being used in the AMM method to create a ground truth discharge at the outlet. The LSTM network was then trained to correct GloFAS at the outlet using the low-pass-filtered WMB river discharge processed with AMM, thereby providing a reliable daily estimate of the total Amazon river discharge. Illustrations are provided and described in Section 4, Figure 6.

3.4 Ocean model simulations

3.4.1 Numerical model configuration

The ocean model configuration adopted in this study is the same as described in Kern et al. (2026), and briefly recalled here for clarity. Simulations are performed with the NEMO ocean model (version 4.0.7, Madec et al., 2017), coupled with the SI3 sea-ice model, configured on the ORCA1 tripolar grid at approximately 1° horizontal resolution with 75 vertical levels. Surface boundary conditions are derived from ERA5 atmospheric reanalysis at hourly frequency using CORE bulk formulas (Large

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

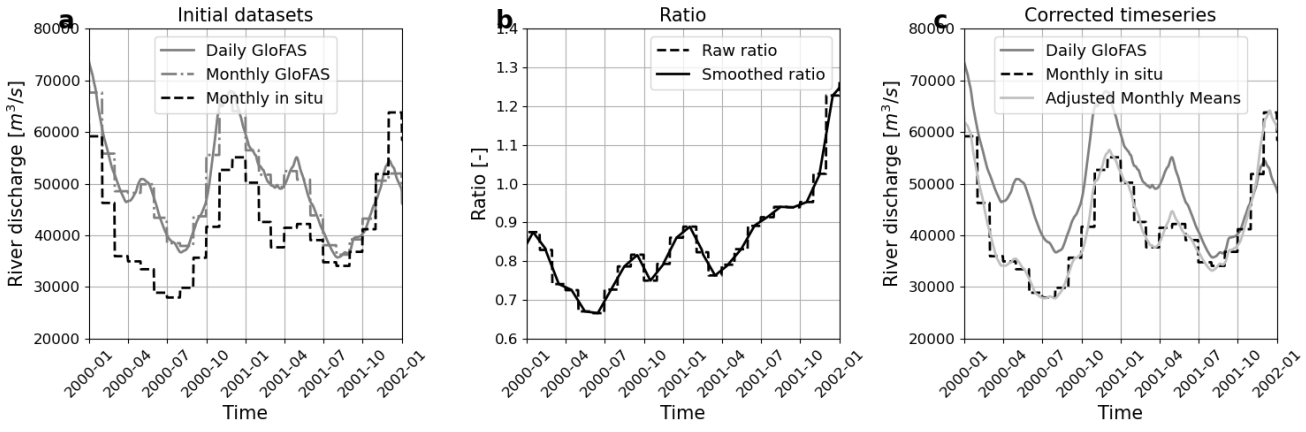

Figure 3. The Adjusted Monthly Means method. This method consists in computing the ratio between monthly in situ measurements and monthly GloFAS discharge values (panel **a**), then smoothing this ratio (panel **b**) before multiplying it with the GloFAS daily data to get the corrected time series (panel **c**). The examples are given for the Congo river.

and Yeager, 2009) as implemented by the AEROBULK package (Brodeau et al., 2017). The setup is identical to that of the CIGAR reanalysis (Storto and Yang, 2024), including a 3-band RGB shortwave radiation scheme and TKE vertical mixing parameterization. The main difference with respect to Kern et al. (2026) is the simulation period, which here spans the period from January 1991 to December 2023, enabling the assessment of the full temporal extent of the hybrid discharge dataset. No data assimilation is applied, so that changes in the simulated ocean state can be unambiguously attributed to differences in the freshwater forcing.

3.4.2 Experiments

Six experiments are performed to isolate the impact of different river discharge configurations, with all other aspects of the model setup held identical. The reference experiment (REF) uses the standard NEMO monthly climatology from Dai et al. (2009) applied globally to all rivers. A second experiment (REF41) retains the same climatological runoff, but restricts it to the 41 rivers included in the hybrid dataset, allowing the effect of river selection to be separated from the effect of the discharge dataset itself. A third experiment (GLOFAS) replaces the climatological discharge for the 41 rivers with the raw (uncorrected) GloFAS daily time series, providing a benchmark for the LSTM-based correction. The main experiment of this study (HYBRID) uses the forcing for the 41 rivers from the new hybrid LSTM-corrected daily discharge dataset, named the F3O dataset. To disentangle the role of temporal variability from mean bias correction, a fifth experiment (HYBRID-CLIM) uses the hybrid dataset averaged into a monthly climatology, so that only the corrected mean seasonal cycle (i.e., not the interannual variability) is introduced. Finally, a sixth experiment (JRA55) uses the daily varying runoff from the JRA55-do dataset (Tsujino et al., 2018) as an alternative time-varying reference, enabling comparison with an existing external product.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

3.4.3 Remapping

Before being ingested into NEMO, the river discharge data must be remapped from the nominal river outlet locations onto the model grid. This is achieved through a dedicated tool (see Code availability section) that follows a four-step procedure. First, a coastal mask is generated by flagging all ocean grid cells adjacent to land, restricting discharge injection to physically meaningful land-ocean boundary points. Second, each river mouth is matched to the nearest coastal gridpoint using the Haversine great-circle distance. Third, since at the ORCA1 resolution the river mouth and its associated freshwater plume are subgrid-scale features, injecting the full discharge into a single grid cell would cause unrealistic local salinity drops and potential numerical instabilities. Discharge is therefore spatially distributed using a Gaussian decay function centred on the river mouth gridpoint, with a decorrelation length scale equal to 150 km in a circle of radius equal to 6 gridpoints, providing a physically motivated parameterization of the lateral spreading of riverine freshwater in the coastal ocean (Horner-Devine et al., 2015). Finally, discharge values (m^3/s) are converted to the NEMO flux units ($kg/m^2/s$) through conservative remapping that preserves the total freshwater mass flux.

4 Quantification of freshwater fluxes

4.1 Comparison with in situ, GloFAS and GloFAS empirically corrected for individual rivers

The quality assessment of our AI-based hybrid dataset is performed through two different methods. First, we evaluate the performance of the neural network at the gauge stations by comparing the hybrid discharge dataset with the in situ measurements. Second, we update the freshwater forcing from rivers in an oceanic reanalysis and compare the model's outputs with salinity data (see Section 5).

For the comparison with in situ measurements, we use three different metrics commonly used in hydrology to estimate the quality of the hybrid discharge dataset: the modified KGE (modified Kling-Gupta Efficiency, named KGE' hereafter, Kling et al. (2012)), the NSE (Nash Sutcliffe Efficiency, Nash and Sutcliffe (1970)) and the NSE_1 (modified version of the NSE, Legates and McCabe Jr. (1999)). The KGE' is a comparison metric ranging between $-\inf$ and 1, 1 being the best possible score, that looks at correlation, bias and variability at the same time. The NSE and the NSE_1 both focus on the variance of the signal, the NSE_1 being more adapted to dataset with outliers (Legates and McCabe Jr. (1999)). Combining the KGE' with the NSE_1 can be useful to study discharge time series from Arctic rivers, with high discharge peaks following low discharge periods. The metrics for the 41 rivers are listed in the Table in Appendix B.

Table 1 presents a summary of the performance metrics. The LSTM outperforms GloFAS for most rivers (29 rivers when looking at the KGE' only, 27 when looking at NSE and 32 when looking at NSE_1), with a median KGE' of 0.76 compared to 0.65 for GloFAS. There are however some exceptions: the Moose, Zhu, Kolyma and Anabar rivers. In the case of the Moose, Kolyma and Anabar rivers, the ERA5 inputs exhibit high frequency noise, directly affecting the LSTM river discharge output. In the case of the Zhu river, very few in situ data are available for training, limiting the performance of the LSTM. The median NSE (resp. NSE_1) of the LSTM is 0.56 (resp. 0.44) compared to 0.37 (resp. 0.34) for GloFAS. In addition, metrics values across

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



rivers show less dispersion around the mean for the LSTM than for GloFAS. Figure 4 presents a map of the KGE' difference between GloFAS and the LSTM, indicating that the LSTM performs better than or similarly to GloFAS for all rivers, except for the Moose, Zhu, Kolyma and Anabar rivers.

Overall, the LSTM provides higher discharge estimation performance than GloFAS across the selected rivers. Among the 41 selected rivers, KGE' improves significantly for 26 rivers, remains similar for 11 and decreases for 4 (Zhu, Kolyma, Moose, Anabar), see the Table in Appendix B for more details. On average, the KGE' increases by approximately 26 % with respect to GloFAS, the NSE by 28 % and the NSE_1 by 14 %. The mean bias is reduced by 58 %, while the difference between the LSTM and in situ estimates of the annual cycle amplitude is reduced by 51 % .

The river discharge time series from GloFAS, the LSTM, and in situ observations were compared to assess the LSTM ability to reproduce the overall temporal dynamics and variability of river discharge across multiple time scales, ranging from the temporal mean and long-term trends to interannual and annual variability, as well as higher-frequency fluctuations (Figure 5). Overall, the LSTM estimates are consistent with the in situ observations at time scales larger than the month, while reproducing the high-frequency variability present in GloFAS. Figure 5 illustrates several aspects of the LSTM performance. The model is able to correct the discharge bias (panel **c**), the phase of the annual cycle (panels **a**, **e** and **b** for the low flow conditions), and the amplitude of the annual cycle (panels **a** and **d**). More generally, the comparison indicates that the LSTM can adjust the large-scale temporal dynamics of GloFAS while retaining shorter-term variability. Panel **f** is an example of a river that the LSTM struggles to correct. Here, the combination of low discharge values and noisy input data leads to erroneous discharge estimates.

As explained in Section 3.3, we applied a specific training strategy to the Amazon river basin. Figure 6 shows the comparison between the LSTM river discharge trained at the Obidos station with in situ data and trained at the river outlet with the WMB approach. In panel **a**, we can see that the classic LSTM (trained at the station) generates an outlet discharge that is close to the in situ discharge at the station. In particular, the maximum annual discharge is approximately the same whereas it should be much higher, taking into account the downstream tributaries. In addition, the amplitude of the LSTM outlet signal is small compared to the reliable WMB outlet discharge. It is closer to the amplitude of the in situ station discharge whereas it should be closer to the amplitude of the WMB outlet discharge. One can also notice that the interannual variations are barely visible in the LSTM time series, and we know it is not correct when looking at the in situ data. All these factors led us to consider this classic LSTM outlet discharge as unreliable. After setting the specific strategy described in Section 3.3, we get the results presented in panel **b**. We can see that the new LSTM, trained at the outlet with WMB data (called LSTM-WMB), provides a discharge time series that is much closer to the WMB one, with more realistic amplitude and interannual variations. We can also remark that the GloFAS discharge has way bigger annual maxima compared to the classic LSTM, which is comforting regarding the tributaries downstream of Obidos. To summarise, training the LSTM at the outlet based on river discharge estimated with the WMB approach increases performance compared to the standard procedure (involving a training based on in situ measurements). The poor performance from the standard procedure is likely due to (1) the absence of in situ station near the outlet, (2) GloFAS inability to correctly reproduce the change in discharge dynamics between the outlet and the station for the Amazon river.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

Table 1. Summary of the performance metrics comparing GloFAS and the LSTM to the in situ measurements. The mean, median and standard deviation of each metric, computed for the 41 rivers, are displayed.

Metrics	KGE'		NSE		NSE_1	
Dataset	GloFAS	LSTM	GloFAS	LSTM	GloFAS	LSTM
Mean	0.44	0.66	-1.43	0.40	-0.10	0.39
Median	0.65	0.76	0.37	0.56	0.34	0.44
Std	0.53	0.31	4.59	0.71	1.09	0.31

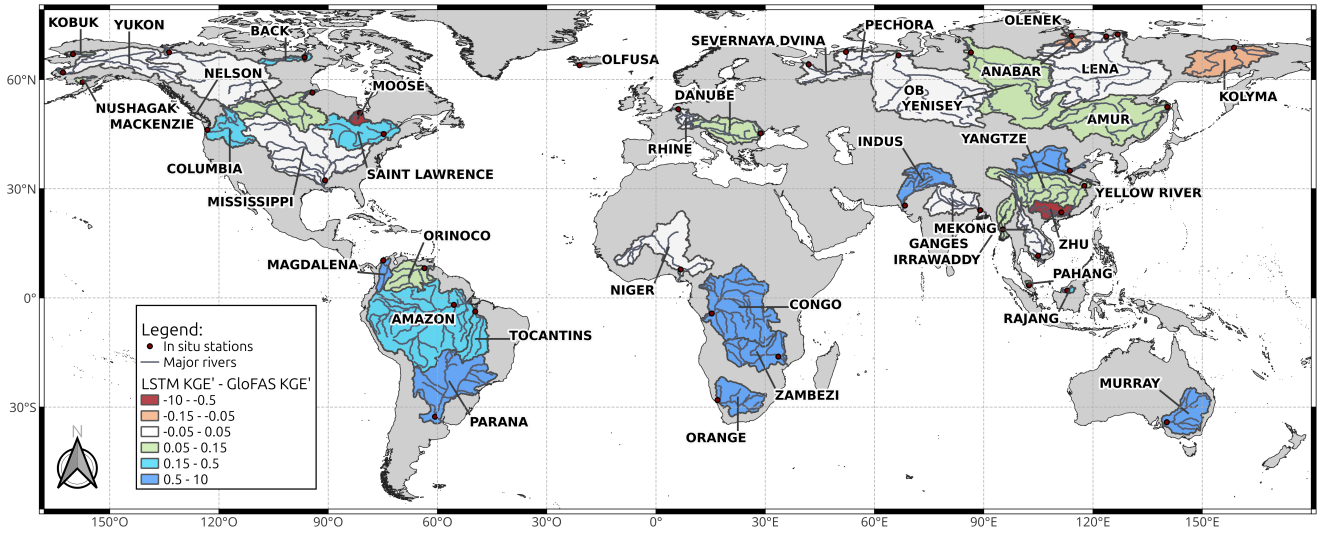

Figure 4. Difference between LSTM KGE' and GloFAS KGE'. Blue and green basins represent an improvement compared to GloFAS.

4.2 Difference between LSTM discharge at the station and at the outlet

The differences between the LSTM discharge at the outlet and at the station, that represent the actual discharge difference between the station and the outlet are less than 10 % for most of the major rivers, except for the Amazon. This is due to the fact that we chose the in situ stations as close to the outlet as possible.

However, for each river, we made sure that the outlet discharge was on average above the station discharge. We also looked at the difference in variability between the LSTM discharge at the station and at the outlet to ensure it was relatively close to the difference of GloFAS discharge variability between the station and the outlet.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

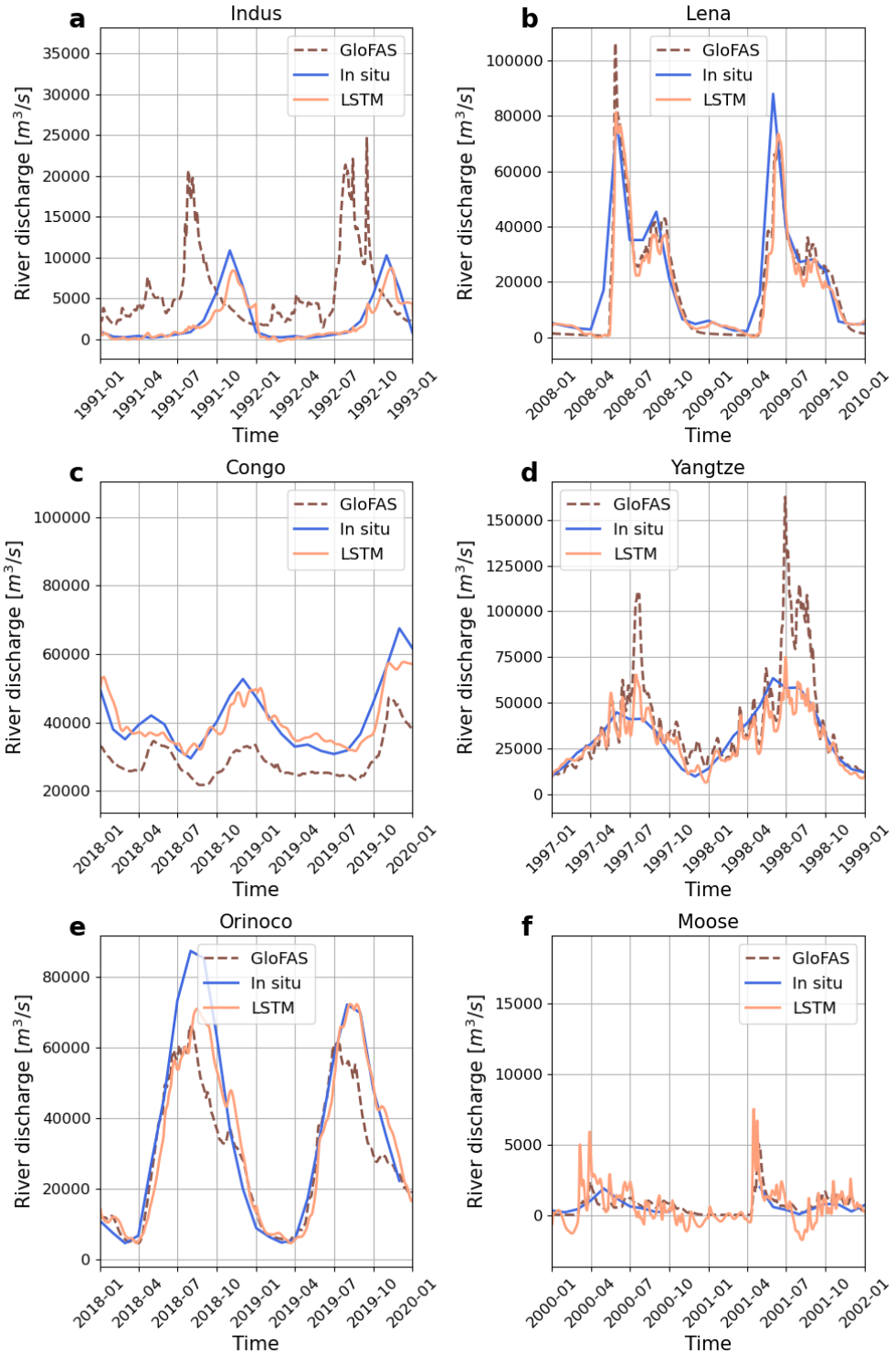

Figure 5. Comparison between GloFAS and LSTM discharge for several rivers.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

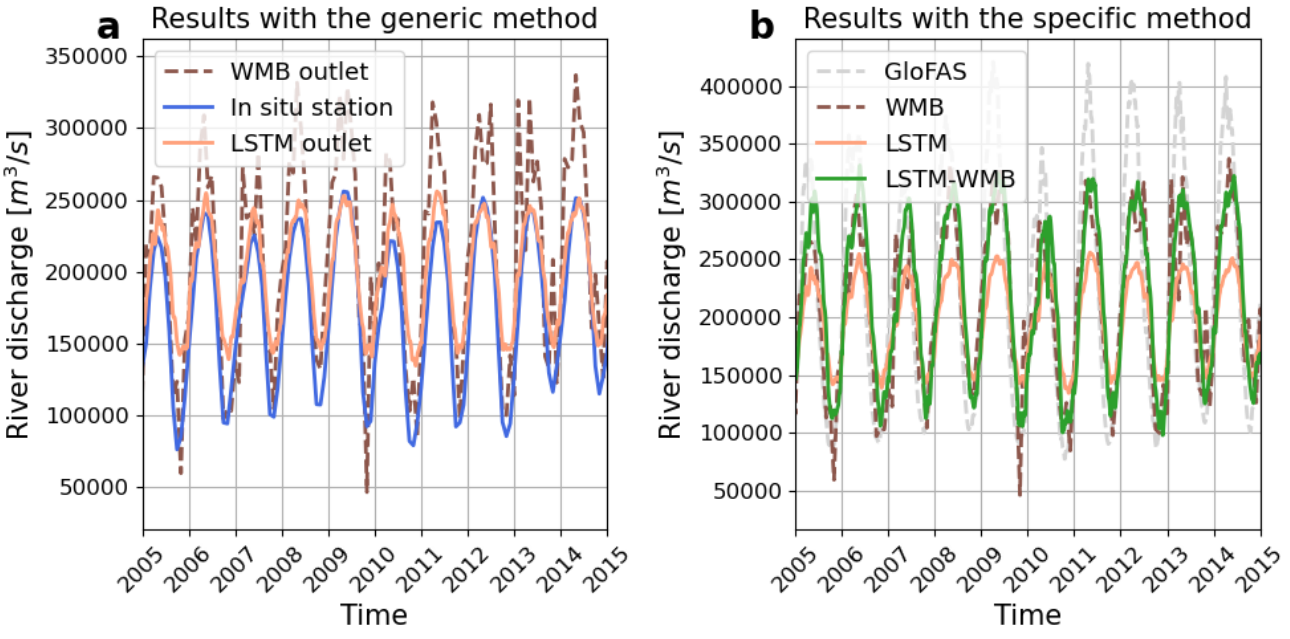

Figure 6. The specific strategy for the Amazon discharge at the outlet. Panel **a** presents a comparison between the in situ discharge at the Obidos station and the classic LSTM (trained at the station) and WMB discharge estimates at the outlet. Panel **b** adds to the comparison the GloFAS discharge at the outlet as well as the adapted LSTM (trained at the outlet on the filtered WMB data) discharge at the outlet.

5 Impact of freshwater fluxes on the ocean

Figure 7 (left panel) presents the time series of the total global river discharge integrated across all rivers included in each experiment. Three groups of datasets emerge clearly. REF and JRA55 deliver comparable total runoff of about 1.3 Sv, consistent with a global-ocean freshwater input that includes the full set of rivers worldwide; REF appears as a flat annual cycle by construction (climatology), while JRA55 exhibits interannual variability. REF41, which retains the climatology but applies it only to the 41 rivers in our study, remains at approximately 0.8 Sv, illustrating the cumulative contribution of these major rivers to global discharge. Most notably, HYBRID and GLOFAS produce a substantially smaller total discharge of about 0.5–0.6 Sv for these same 41 rivers, which is roughly 25–35% less than the Dai et al. (2009) climatology applied to the identical set of catchments. This indicates that the standard climatology systematically overestimates the discharge of several major rivers, an interpretation supported by the in situ data comparison in the following paragraphs, where the hybrid dataset was shown to reduce mean biases in many basins. The hybrid and GloFAS time series also exhibit a pronounced seasonal cycle together with a weak negative trend over the period.

This diagnosis directly explains the SSS RMSE time series shown in the right panel of Figure 7, computed over coastal grid points within 500 km from the coastline. The HYBRID experiment consistently achieves the lowest RMSE throughout 1991–2023, confirming that the reduction of the excess freshwater input translates into a more realistic surface salinity field.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



HYBRID-CLIM tracks HYBRID very closely, indicating that the improvement is primarily driven by corrections to the mean discharge rather than by the introduction of interannual variability. GLOFAS and REF41 lead to intermediate skill, both outperforming the global climatological reference REF, while JRA55 performs comparably to or slightly worse than REF. A common feature of all experiments is a gently increasing RMSE trend over the period, which may reflect a progressive amplification of the global freshwater cycle and associated salinity contrasts that no experiment fully captures. The fact that the offsets between experiments remain approximately constant in time further reinforces the conclusion that the dominant effect of the hybrid dataset is on the mean salinity state rather than on its temporal variability.

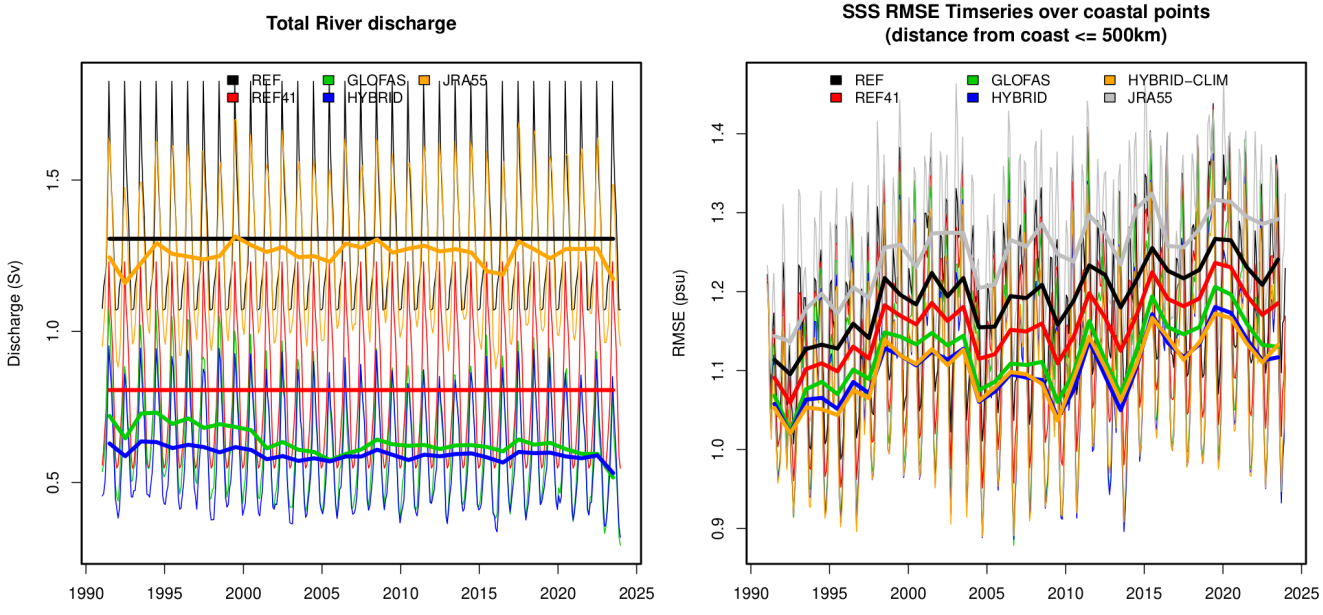

Figure 7. Left panel: time series of the total global river discharge integrated across all rivers included in each experiment. Right panel: SSS RMSE time series computed over coastal grid points within 500 km from the coastline, for the period 1991–2023, against EN4 objective analyses.

Figure 8 presents the SSS RMSE and its reduction relative to REF as a function of distance from the coast. All experiments except JRA55 reduce RMSE with respect to REF within the first 500 km, where the impact of river forcing is most direct. HYBRID and HYBRID-CLIM deliver the largest improvement, remaining positive out to approximately 2000–3000 km from the coast, consistent with the large-scale propagation of freshwater anomalies by ocean circulation. The near-identity of HYBRID and HYBRID-CLIM curves across all distances confirms, again, that the benefit is predominantly associated with correcting the mean discharge rather than its variability. The decomposition of RMSE reduction into bias and random error components shows that systematic bias accounts for roughly 40–60% of the total RMSE decrease, highlighting that the climatological runoff introduced a significant mean error in the modelled salinity that the hybrid dataset effectively corrects.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

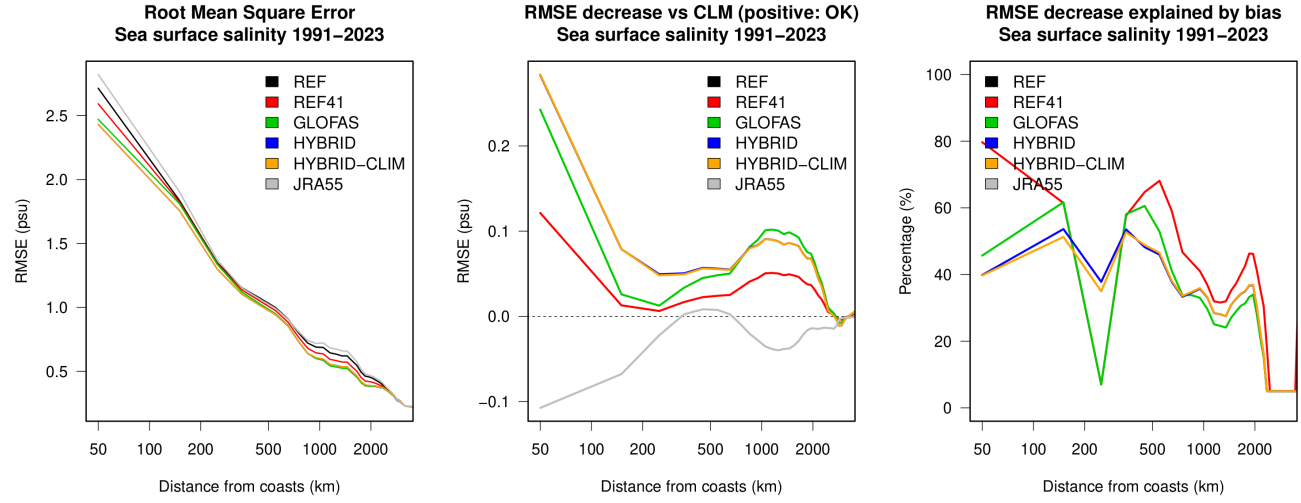

Figure 8. SSS RMSE (left), RMSE decrease relative to REF (centre, positive means improvement), and percentage of RMSE decrease explained by bias (right), as a function of distance from the coast, for the period 1991–2023. Verification is against the EN4 objective analyses.

The spatial maps of SSS RMSE decrease with respect to REF (Figure 9) reveal where these improvements occur. REF41 already shows widespread moderate gains simply from the selection of 41 major rivers, indicating that the spatial targeting of the most influential discharge points matters even before any correction is applied. GLOFAS improves on REF across broad regions but introduces some degradation in specific areas, reflecting its known biases. HYBRID produces the most extensive and consistent improvement, with the largest gains concentrated in the North Atlantic, encompassing the Amazon plume and Gulf of Mexico outflow region, and in the Maritime Continent and South China Sea, areas strongly influenced by large tropical rivers. HYBRID-CLIM shows a very similar pattern to HYBRID, again confirming the dominance of mean bias correction. The direct comparison of HYBRID against GLOFAS (bottom-left panel) highlights locally significant additional improvements, notably near the Amazon mouth and across South and East Asian river systems, demonstrating the added value of the LSTM-based correction beyond the raw hydrological model output. The comparison of HYBRID against HYBRID-CLIM (bottom-right panel) shows only very marginal differences, consistent with the interpretation that interannual variability in discharge plays a secondary role in the global SSS error budget at this model resolution.

To assess the impact of the enhanced freshwater forcing on the upper ocean structure, we computed a set of stratification diagnostics from the monthly NEMO output, following common practice in the physical oceanography literature. The Mixed Layer Depth (MLD) is defined using a potential density threshold criterion, namely the depth at which the density exceeds its value at 10 m by 0.03 kg m^{-3} (de Boyer Montégut et al., 2004). The squared Brunt-Väisälä frequency $N^2 = -(g/\rho_0)(\partial\rho/\partial z)$ quantifies the local strength of vertical density stratification (Vallis, 2017); it is computed here through the TEOS-10 framework (IOC et al., 2010), and reported both as a vertical average over the 0–200 m layer (mean stratification) and as the maximum N^2

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

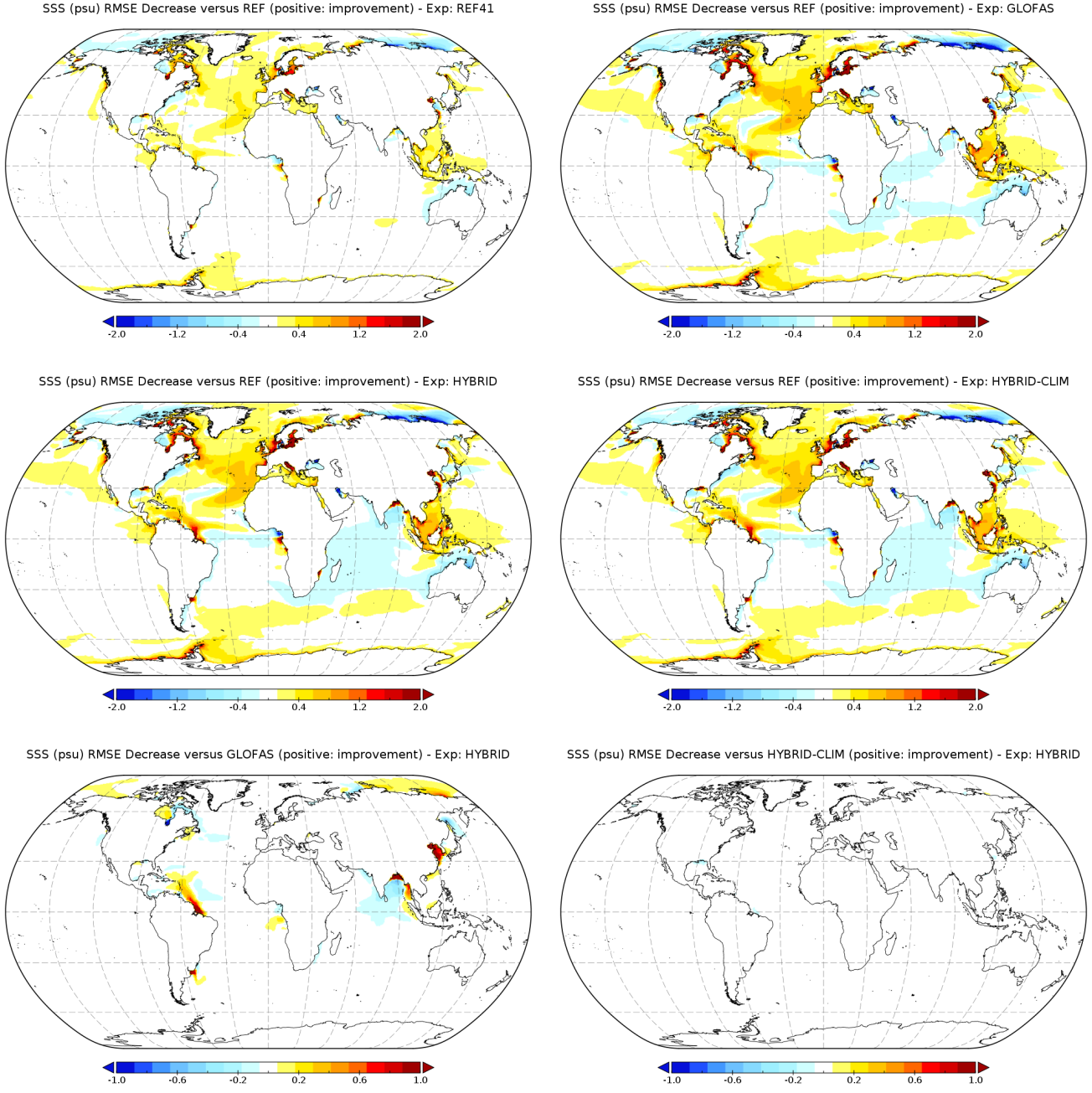

Figure 9. Maps of SSS RMSE decrease (psu) relative to REF over the period 1991–2023, against EN4 objective analyses. Positive values indicate improvement. Top row: REF41 and GLOFAS versus REF. Middle row: HYBRID and HYBRID-CLIM versus REF. Bottom row: HYBRID versus GLOFAS (left) and HYBRID versus HYBRID-CLIM (right).

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



in the same layer (peak stratification, corresponding to the thermocline location). The thermocline depth is defined as the depth of maximum N^2 within 0–200 m, providing a measure of the vertical position of the pycnocline. Finally, the Potential Energy Anomaly (PEA) of Simpson et al. (1978), $\varphi = -(g/H)\int_0^H (\bar{\rho} - \rho(z)) \cdot z\, dz$, represents the amount of mechanical energy per unit volume required to fully mix a water column; larger positive values indicate stronger stratification.

Figure 10 shows the differences between the HYBRID and REF experiments for these diagnostics, averaged over 1991–2023. Consistent with the SSS validation discussed previously, the upper ocean in HYBRID is saltier and therefore denser at the surface, leading to a generalised weakening of vertical density gradients. The MLD difference fields confirm this picture: in March, the mixed layer deepens (positive differences of several tens of metres) over large portions of the northern hemisphere, with the strongest signals in the subpolar North Atlantic, the Nordic Seas, the Mediterranean, and the northwestern Pacific, which are areas where REF was identified as too fresh in the SSS analysis. In September, the deepening signal shifts to the Southern Hemisphere, particularly across the Southern Ocean and the southern subtropical Atlantic, again following the seasonality of deep convection. Both the mean and peak N^2 fields exhibit predominantly negative differences across the global ocean, indicating that the corrected, less freshwater-loaded surface layer is associated with reduced stratification both on average over the upper 200 m and around the thermocline. The PEA difference reinforces this interpretation, with widespread negative values (i.e., less energy required to homogenise the water column) particularly evident in the tropical Atlantic, the Indian Ocean, and the Maritime Continent, where the SSS improvement was also most pronounced.

The response of the thermocline depth is spatially more heterogeneous. A general deepening (positive differences of several metres) emerges over much of the tropics, the Indo-Pacific warm pool, the Maritime Continent, and parts of the Southern Ocean, reflecting the downward shift of the maximum density gradient in response to the saltier and denser surface layer. Notably, in regions such as the Maritime Continent and the South China Sea, the thermocline deepens even where the MLD signal is weak, suggesting that the freshwater correction influences the vertical structure of the upper ocean beyond the directly mixed surface layer. This result is consistent with the broad open-ocean RMSE improvements observed up to several thousand kilometres offshore. Localised opposite signals (shallower thermocline) are found in a few subpolar North Atlantic regions, likely associated with the strong MLD deepening there, which tends to erode the seasonal pycnocline. The stratification diagnostics provide a coherent dynamical interpretation of the salinity changes induced by the hybrid runoff: a more realistic, less freshwater-biased coastal input propagates into a measurable adjustment of the upper-ocean density structure, with concrete implications for mixed-layer dynamics, air-sea exchange, and the vertical exchange of heat and tracers.

Finally, we examine how the freshwater correction propagates into the large-scale ocean dynamics. Figure 11 shows the time-mean barotropic streamfunction in the REF experiment (left panel) and its difference between HYBRID and REF (right panel). While the absolute pattern in REF reproduces the canonical global gyre structure, the Antarctic Circumpolar Current (ACC) being the dominant feature, together with the western boundary currents and the subtropical gyres, the difference field reveals a small but spatially coherent enhancement of the barotropic circulation, of order of 0.5–3 Sv, organised in two main signals: a basin-wide intensification across the North and South Atlantic subtropical gyres and a clear strengthening along the path of the ACC, particularly downstream of Drake Passage and south of Africa.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

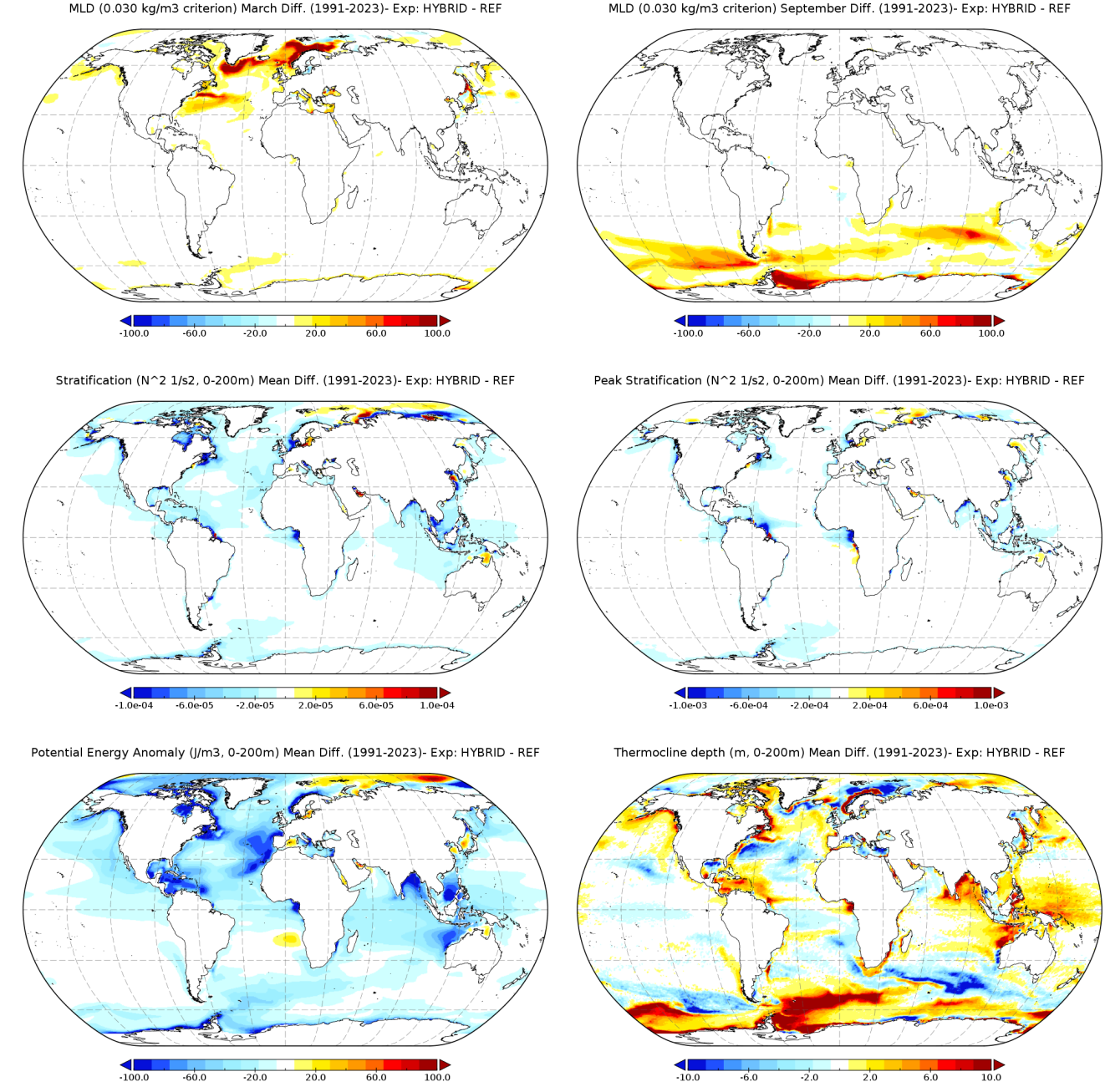

Figure 10. Differences between the HYBRID and REF experiments, averaged over 1991–2023, for: (top) MLD in March and September; (middle) mean and peak N^2 in the 0–200 m layer; (bottom) Potential Energy Anomaly in the 0–200 m layer and thermocline depth.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

This response is dynamically consistent with the salinity and stratification changes documented in the previous paragraphs. The reduction of excess freshwater input in HYBRID produces a saltier and denser surface ocean, especially in the North Atlantic, the tropical Atlantic, and the southern subtropics, which in turn lowers the halosteric component of sea surface height (SSH) downstream of major river mouths. The resulting modification of horizontal SSH gradients alters the surface geostrophic balance (see, e.g., Aroucha et al., 2025). In the Atlantic Ocean, the combined effect of reduced freshwater inputs from multiple major outlets (Amazon, Orinoco, Mississippi, Congo, Niger, and North European rivers) reinforces the meridional SSH gradient, with a corresponding amplification of the subtropical gyres' strength. The enhancement of the ACC may reflect a complementary mechanism: even small density changes in the Southern Ocean can modify the meridional density gradient that supports the ACC's geostrophic eastward flow (Hogg, 2010).

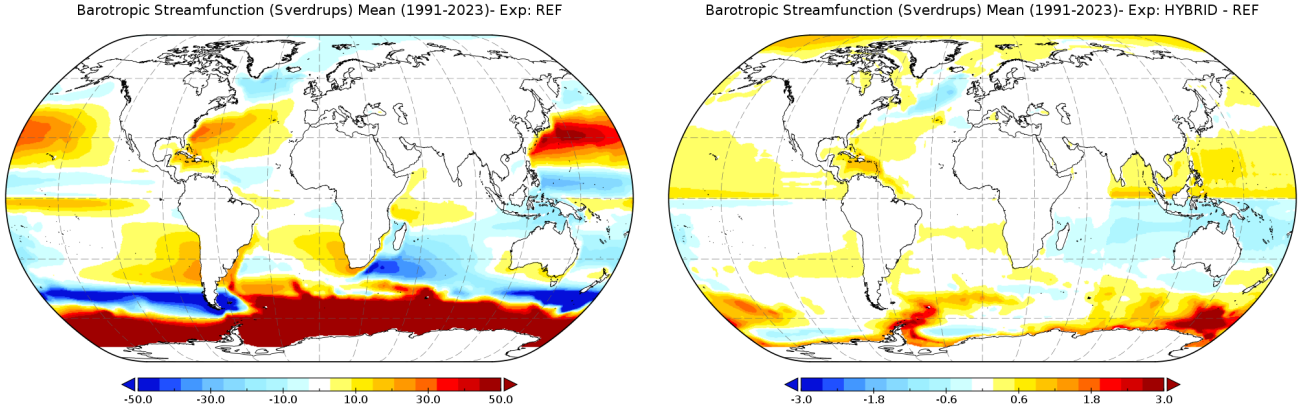

Figure 11. Time-mean barotropic streamfunction (Sv) over 1991–2023: REF experiment (left) and HYBRID minus REF difference (right).

The Pacific response is weaker and more spatially scattered. We attribute this to two complementary factors. First, the Pacific basin receives a smaller cumulative correction from the 41-river ensemble, while the Atlantic concentrates the world's largest discharges; the SSS validation already showed that salinity improvements in the open Pacific were modest. Second, the dominant Pacific salinity signals in our experiments are confined to the Maritime Continent and the South China Sea, not projecting onto the basin-scale gyre circulation.

Consistently, the meridional transport diagnostics (Figure 12) show that the Atlantic exhibits a small but systematic increase in the meridional heat and salt transports, particularly in the tropics and subtropics, accompanied by a modest reinforcement of the volume transport, which is consistent with the gyre-scale intensification visible in the barotropic streamfunction. This signal extends to the overturning component as well: the Atlantic Meridional Overturning Circulation (AMOC) in HYBRID is strengthened compared to REF (not shown), with an enhancement of the maximum streamfunction of about 0.8 to 1 Sv depending on the latitude (peaking near 1 Sv around the Equator and reaching approximately 0.8 Sv at 26°N, with the maximum increase located at depths of about 200 m near the Equator and 400 m around 26°N, consistent with a denser, saltier upper Atlantic favouring slightly enhanced downward branches of the overturning). The global meridional transports are nearly

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

indistinguishable between REF and HYBRID, indicating that the Atlantic signal is partially compensated at the global scale, due to the negligible impact of the hybrid runoff dataset in the Pacific Ocean.

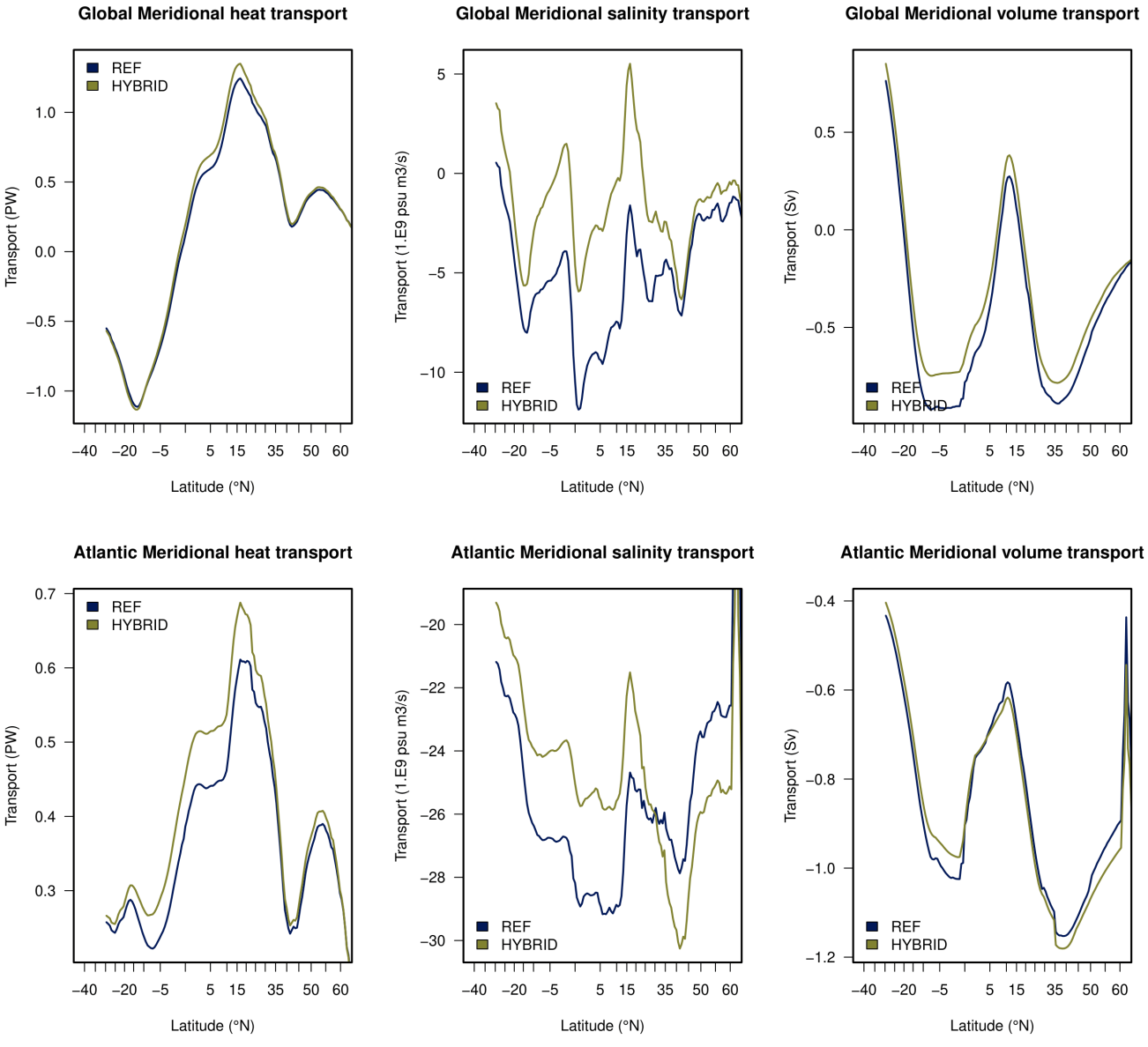

Figure 12. Meridional heat (left), salinity (centre), and volume (right) transports across the Global (top) and the Atlantic (bottom) Oceans for the REF and HYBRID experiments, averaged over 1991–2023.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



6 Conclusion and perspectives

This study led to the production of the F3O dataset, a new daily discharge dataset going from 1989 to 2023, for 41 rivers worldwide, at the last gauge station and at the outlet. This new dataset can be seen as a temporal extension of in situ datasets as well as an improvement in temporal resolution as most in situ measurements are monthly. This new dataset has better performance than GloFAS for several indicators such as the KGE' or the phase and amplitude of the annual cycle. It provides accurate, daily, continuous, and globally consistent river discharge estimates, thereby enabling a more comprehensive characterization of river discharge dynamics at the global scale. It adresses all the challenges associated with individual datasets: accuracy for GloFAS, coverage and resolution for in situ. The AI-based method developed in this work is applicable as long as there are GloFAS discharge estimates and in-situ measurements.

The hybrid LSTM-corrected discharge dataset was finally evaluated in a set of NEMO ocean simulations covering 1991–2023 and mimicking the CIGAR reanalysis configuration, in which data assimilation was disabled to isolate the response to the freshwater forcing, across six experiments using the standard climatology, the raw GloFAS output, the new hybrid product, or the JRA55-do alternative. The HYBRID experiment systematically outperformed all other configurations against EN4 SSS gridded observations, with the largest improvements in the North Atlantic and the Maritime Continent, and reductions in coastal RMSE extending well beyond the first 500 km offshore. The close similarity between HYBRID and HYBRID-CLIM further indicated that the bulk of the improvement stems from correcting the mean discharge rather than its interannual variability, consistently with the finding of Kern et al. (2026). However, with higher spatial resolution than those used in the present study (e.g., fully eddy-resolving), the impact of inter-annual variability might emerge more clearly. Following the reduction of an excess fresh bias, the upper ocean in HYBRID is denser at the surface, leading to a generalised weakening of vertical stratification, a deepening of the mixed layer and thermocline, and a coherent reduction in the potential energy anomaly in a large part of the global ocean, with potential implications for mixed-layer dynamics, air-sea exchange, and the vertical exchange of heat and tracers. Finally, these density changes project onto the large-scale circulation as a small but systematic intensification of the Atlantic subtropical gyres and the Antarctic Circumpolar Current, accompanied by modestly enhanced meridional heat and salt transports in the Atlantic, which confirms the value of accurate river runoff forcing for ocean reanalysis applications.

Code availability. The river discharge remapping tool onto the NEMO grid is freely available at the following link: https://baltig.cnr.it/nemo_ismar-rm/nemo_runoff_remapping

Data availability. The F3O outlet river discharge dataset is freely available at the following link: https://doi.org/10.5281/zenodo.23062293 (Vaujour et al. (2026)).

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

Appendix A: Data temporal coverage

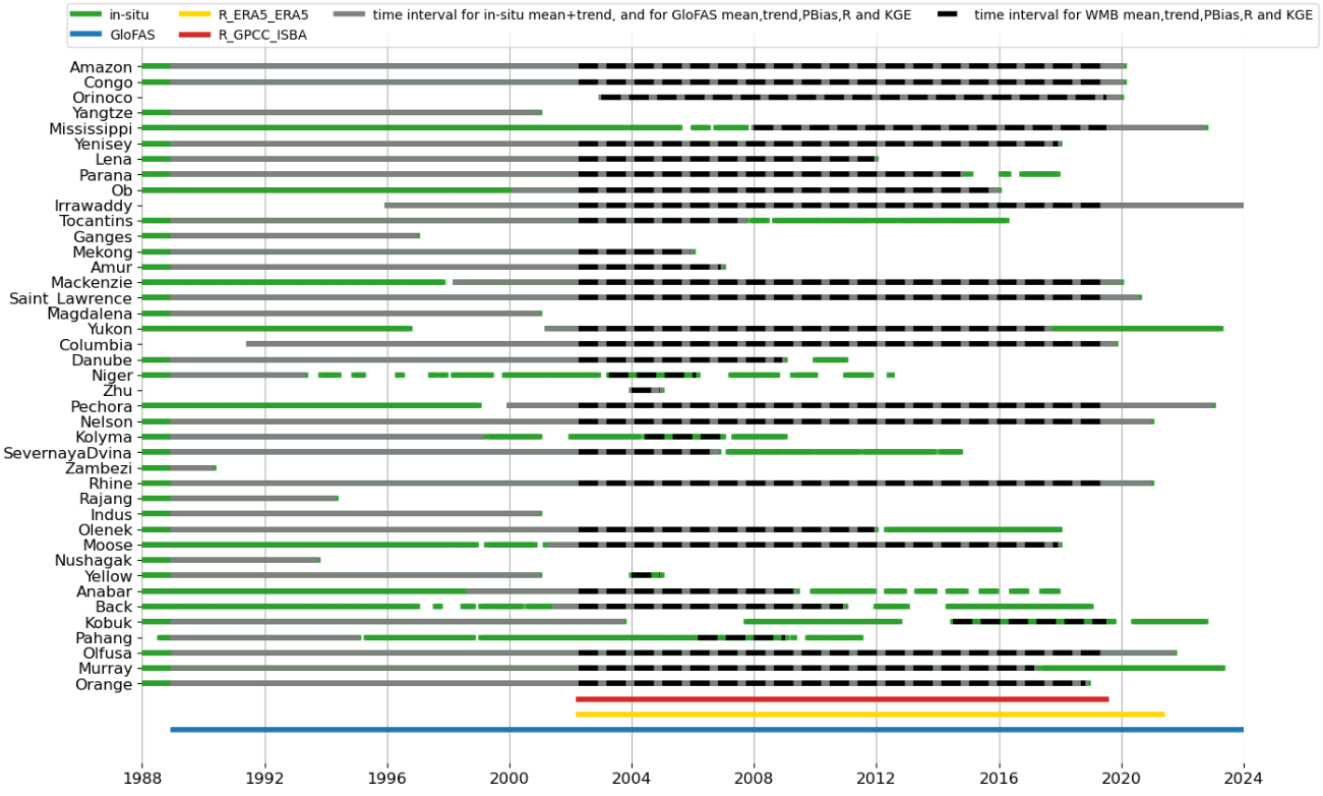

Figure A1. The data availability for GloFAS the WMB datasets and the in situ measurements. The labels *R ERA5 ERA5* and *R GPCC ISBA* respectively represent the WMB estimates computed with ERA5 atmospheric variables and GPCC and ISBA atmospheric variables.

Appendix B: Performance metrics of GloFAS and the LSTM for all rivers

Author contributions. Thomas Vaujour, Louis Kern and Andrea Storto contributed to the technical work. Julia Pfeffer, Chloé Thénoz and Chunxue Yang had respectively a supervisor, technical expert and thematic expert role. Stéphanie Guinehut funded the study.

Competing interests. The authors declare that they have no conflict of interest.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

Table B1. Comparison of KGE', NSE and NSE_1 scores for the GloFAS and LSTM time series

Metrics	KGE'		NSE		NSE_1	
River	GloFAS	LSTM	GloFAS	LSTM	GloFAS	LSTM
Amazon	0.70	0.90	0.19	0.87	0.16	0.69
Congo	0.08	0.78	-2.56	0.55	-0.86	0.34
Orinoco	0.84	0.92	0.88	0.91	0.71	0.72
Yangtze	0.27	0.36	-0.95	0.12	-0.10	0.30
Mississippi	0.76	0.74	0.48	0.44	0.34	0.30
Yenisey	0.73	0.79	0.65	0.62	0.43	0.49
Lena	0.87	0.83	0.79	0.74	0.62	0.63
Parana	-0.09	0.69	-14.70	0.40	-2.98	0.20
Ob	0.89	0.87	0.82	0.80	0.63	0.66
Irrawaddy	0.82	0.90	0.76	0.81	0.62	0.66
Tocantins	0.70	0.86	0.40	0.73	0.31	0.56
Ganges	0.87	0.88	0.76	0.82	0.68	0.71
Mekong	0.87	0.89	0.82	0.82	0.67	0.69
Amur	0.69	0.83	0.44	0.65	0.42	0.58
Mackenzie	0.84	0.84	0.73	0.69	0.60	0.55
Saint Lawrence	0.64	0.82	0.35	0.69	0.23	0.46
Magdalena	-0.45	0.68	-21.15	0.56	-4.40	0.38
Yukon	0.85	0.86	0.74	0.72	0.64	0.65
Columbia	0.42	0.73	-0.22	0.40	-0.08	0.29
Danube	0.73	0.79	0.42	0.59	0.24	0.35
Niger	0.82	0.86	0.67	0.79	0.51	0.59
Zhu	0.65	0.08	0.45	0.39	0.37	0.14
Pechora	0.74	0.73	0.52	0.45	0.43	0.44
Nelson	0.67	0.78	0.22	0.54	0.10	0.32
Kolyma	0.86	0.80	0.75	0.66	0.66	0.60
Severnayadvina	0.75	0.72	0.57	0.60	0.40	0.49
Zambezi	-0.09	0.55	-4.04	0.34	-1.05	0.37
Rhine	0.55	0.53	-0.00	0.16	0.07	0.17
Rajang	0.35	0.55	-1.63	0.09	-0.40	0.12
Indus	-0.93	0.76	-4.28	0.72	-1.05	0.60
Olenek	0.59	0.57	0.37	0.37	0.54	0.46
Moose	0.54	-0.69	0.00	-0.82	0.09	-0.50
Nushagak	0.74	0.84	0.59	0.72	0.44	0.52
Yellow	-0.79	0.66	-12.97	0.43	-2.52	0.35
Anabar	0.53	0.41	0.16	0.04	0.43	0.36
Back	0.61	0.77	0.38	0.49	0.48	0.44
Kobuk	0.51	0.64	0.26	0.18	0.42	0.31
Pahang	0.41	0.48	-0.95	-0.15	-0.37	0.04
Olfusa	-1.09	-0.20	-5.56	-3.52	-1.87	-0.96
Murray	-0.52	0.72	-2.81	0.65	-0.63	0.51
Orange	-0.07	0.48	-0.90	0.25	-0.12	0.32

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Acknowledgements. This work has been carried out as part of the Copernicus Marine Service WAMBOR and F3O projects. Copernicus Marine Service is implemented by Mercator Ocean in the framework of a delegation agreement with the European Union.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



References

Catalogue – Hybam, https://hybam.obs-mip.fr/catalogue/?currentSelection=a6e2e07a-de59-30f5-9a97-e6e7ef0cdc98.

Aroucha, L. C., Lübbecke, J. F., Brandt, P., Schwarzkopf, F. U., and Biastoch, A.: River discharge impacts coastal southeastern tropical Atlantic sea surface temperature and circulation: a model-based analysis, Ocean Science, 21, 661–678, https://doi.org/10.5194/os-21-661-2025, 2025.

Brodeau, L., Barnier, B., Gulev, S. K., and Woods, C.: Climatologically Significant Effects of Some Approximations in the Bulk Parameterizations of Turbulent Air–Sea Fluxes, Journal of Physical Oceanography, 47, 5–28, https://doi.org/10.1175/JPO-D-16-0169.1, 2017.

Börgel, F., Karsten, S., Rummel, K., and Gräwe, U.: From weather data to river runoff: using spatiotemporal convolutional networks for discharge forecasting, Geoscientific Model Development, 18, 2005–2019, https://doi.org/10.5194/gmd-18-2005-2025, 2025.

Chen, J., Tapley, B., Rodell, M., Seo, K., Wilson, C., Scanlon, B. R., and Pokhrel, Y.: Basin-Scale River Runoff Estimation From GRACE Gravity Satellites, Climate Models, and In Situ Observations: A Case Study in the Amazon Basin, Water Resources Research, 56, https://doi.org/10.1029/2020wr028032, 2020.

Dai, A.: Dai and Trenberth Global River Flow and Continental Discharge Dataset, https://doi.org/10.5065/D6V69H1T, artwork Size: 25.581 Mbytes Pages: 25.581 Mbytes, 2017.

Dai, A., Qian, T., Trenberth, K. E., and Milliman, J. D.: Changes in Continental Freshwater Discharge from 1948 to 2004, Journal of Climate, 22, 2773–2792, https://doi.org/10.1175/2008JCLI2592.1, 2009.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, Journal of Geophysical Research: Oceans, 109, https://doi.org/https://doi.org/10.1029/2004JC002378, 2004.

Durack, P. J.: Ocean Salinity and the Global Water Cycle, Oceanography, issue_volume, https://doi.org/10.5670/oceanog.2015.03, 2015.

Fathi, M. M., Liu, Z., Fernandes, A. M., Hren, M. T., Terry, Jr., D. O., Nataraj, C., and Smith, V.: An Integrated Deep Learning Framework Enables Rapid Spatiotemporal Morphodynamic Predictions Toward Long-Term Simulations, EGUsphere, 2025, 1–22, https://doi.org/10.5194/egusphere-2025-3368, 2025.

Feng, Y., Menemenlis, D., Xue, H., Zhang, H., Carroll, D., Du, Y., and Wu, H.: Improved representation of river runoff in Estimating the Circulation and Climate of the Ocean Version 4 (ECCOv4) simulations: implementation, evaluation, and impacts to coastal plume regions, Geoscientific Model Development, 14, 1801–1819, https://doi.org/10.5194/gmd-14-1801-2021, 2021.

GRDC: GRDC Major River Basins. Global Runoff Data Centre. 2nd, rev. ed. Koblenz: Federal Institute of Hydrology (BfG), 2020.

Grimaldi, S., Salamon, P., Disperati, J., Zsoter, E., Russo, C.,Ramos, A., Carton De Wiart, C., Barnard, C., Hansford, E.,, and Gomes, G., Prudhomme, C.: River discharge and related historical data from the Global Flood Awareness System. v4.0. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2022.

Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H., and Pappenberger, F.: GloFAS-ERA5 operational global river discharge reanalysis 1979–present, Earth System Science Data, 12, 2043–2060, https://doi.org/10.5194/essd-12-2043-2020, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Vil-

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



laume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 monthly averaged data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/https://doi.org/10.24381/cds.f17050d7, 2023.

Hogg, A. M.: An Antarctic Circumpolar Current driven by surface buoyancy forcing, Geophysical Research Letters, 37, https://doi.org/https://doi.org/10.1029/2010GL044777, 2010.

Horner-Devine, A. R., Hetland, R. D., and MacDonald, D. G.: Mixing and Transport in Coastal River Plumes, Annual Review Fluid Mechanics, 47, https://doi.org/https://doi.org/10.1146/annurev-fluid-010313-141408, 2015.

Hunt, K. M. R., Matthews, G. R., Pappenberger, F., and Prudhomme, C.: Using a long short-term memory (LSTM) neural network to boost river streamflow forecasts over the western United States, Hydrology and Earth System Sciences, 26, 5449–5472, https://doi.org/10.5194/hess-26-5449-2022, 2022.

IOC, SCOR, and IAPSO: The international thermodynamic equation of seawater – 2010: Calculation and use of thermodynamic properties. Intergovernmental Oceanographic Commission, Manuals and Guides No. 56, unesco edn., 2010.

Kern, L., Vaujour, T., Pfeffer, J., Storto, A., Szczypta, C., Garric, G., Sirere, C., Larnicol, G., Yang, C., Bourdalle-Badie, R., and Guinehut, S.: Assessing the Impact of Freshwater Fluxes from Major Rivers on the Atlantic Ocean, EGUsphere, 2026, 1–25, https://doi.org/10.5194/egusphere-2025-6351, 2026.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424-425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Landerer, F. W., Flechtner, F. M., Save, H., Webb, F. H., Bandikova, T., Bertiger, W. I., Bettadpur, S. V., Byun, S. H., Dahle, C., Dobslaw, H., Fahnestock, E., Harvey, N., Kang, Z., Kruizinga, G. L. H., Loomis, B. D., McCullough, C., Murböck, M., Nagel, P., Paik, M., Pie, N., Poole, S., Strekalov, D., Tamisiea, M. E., Wang, F., Watkins, M. M., Wen, H.-Y., Wiese, D. N., and Yuan, D.-N.: Extending the Global Mass Change Data Record: GRACE Follow-On Instrument and Science Data Performance, Geophysical Research Letters, 47, e2020GL088 306, https://doi.org/https://doi.org/10.1029/2020GL088306, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL088306, 2020.

Large, W. G. and Yeager, S. G.: The global climatology of an interannually varying air–sea flux data set, Climate Dynamics, 33, 341–364, https://doi.org/10.1007/s00382-008-0441-3, 2009.

Legates, D. R. and McCabe Jr., G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resources Research, 35, 233–241, https://doi.org/https://doi.org/10.1029/1998WR900018, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/1998WR900018, 1999.

Leščešen, I., Tanhapour, M., Pekárová, P., Miklánek, P., and Bajtek, Z.: Long Short-Term Memory (LSTM) Networks for Accurate River Flow Forecasting: A Case Study on the Morava River Basin (Serbia), Water, 17, https://doi.org/10.3390/w17060907, 2025.

Madec, G., Bourdallé-Badie, R., Pierre-Antoine Bouttier, Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., Ethé, C., Flavoni, S., Graham, T., Harle, J., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Paul, J., Rousset, C., Storkey, D., Storto, A., and Vancoppenolle, M.: NEMO ocean engine, https://doi.org/10.5281/ZENODO.1472492, version Number: v3.6, 2017.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Nearing, G., Cohen, D., Dube, V., Gauch, M., Gilon, O., Harrigan, S., Hassidim, A., Klotz, D., Kratzert, F., Metzger, A., Nevo, S., Pappenberger, F., Prudhomme, C., Shalev, G., Shenzis, S., Tekalign, T., Weitzner, D., and Matias, Y.: AI Increases Global Access to Reliable Flood Forecasts, https://doi.org/10.48550/arXiv.2307.16104, arXiv:2307.16104, 2023.

PA, B., J, V. D. K., and A, D. R.: LISFLOOD - Distributed Water Balance and Flood Simulation Model - Revised User Manual 2013, https://doi.org/10.2788/24982 (print), 10.2788/24719 (online), iSBN: 978-92-79-33191-6 (print), 978-92-79-33190-9 (pdf), 2013.

Ruparell, K., Marks, R., Wood, A., Hunt, K., Cloke, H., Prudhomme, C., Pappenberger, F., and Chantry, M.: Hydra-LSTM: A Semi-Shared Machine Learning Architecture for Prediction across Watersheds, Artificial Intelligence for the Earth Systems, 4, https://doi.org/10.1175/AIES-D-24-0103.1, 2025.

Schanze, J., Schmitt, R., and Yu, L.: The global oceanic freshwater cycle: A State-of-the-art quantification, Journal of Marine Research, 68, https://doi.org/10.1357/002224010794657164, 2010.

Schmitt, R. W.: The ocean component of the global water cycle, Reviews of Geophysics, 33, 1395–1409, https://doi.org/https://doi.org/10.1029/95RG00184, 1995.

Senent-Aparicio, J., Blanco-Gómez, P., López-Ballesteros, A., Jimeno-Sáez, P., and Pérez-Sánchez, J.: Evaluating the Potential of GloFAS-ERA5 River Discharge Reanalysis Data for Calibrating the SWAT Model in the Grande San Miguel River Basin (El Salvador), Remote Sensing, 13, 3299, https://doi.org/10.3390/rs13163299, 2021.

Simpson, J. H., Allen, C. M., and Morris, N. C. G.: Fronts on the continental shelf, Journal of Geophysical Research: Oceans, 83, 4607–4614, https://doi.org/https://doi.org/10.1029/JC083iC09p04607, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/JC083iC09p04607, 1978.

Snieder, E. and Khan, U. T.: A diversity-centric strategy for the selection of spatio-temporal training data for LSTM-based streamflow forecasting, Hydrology and Earth System Sciences, 29, 785–798, https://doi.org/10.5194/hess-29-785-2025, 2025.

Sohail, T. and Zika, J. D.: Global air-sea heat and freshwater fluxes constrained by ocean observations, Earth System Science Data Discussions, 2025, 1–24, https://doi.org/10.5194/essd-2024-545, 2025.

Storto, A. and Yang, C.: Acceleration of the ocean warming from 1961 to 2022 unveiled by large-ensemble reanalyses, Nature Communications, 15, 545, https://doi.org/10.1038/s41467-024-44749-7, 2024.

Syed, T. H., Famiglietti, J. S., Chen, J., Rodell, M., Seneviratne, S. I., Viterbo, P., and Wilson, C. R.: Total basin discharge for the Amazon and Mississippi River basins from GRACE and a land-atmosphere water balance, Geophysical Research Letters, 32, https://doi.org/10.1029/2005GL024851, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2005GL024851, 2005.

Tajouri, S., Llovel, W., Sévellec, F., Molines, J.-M., Mathiot, P., Penduff, T., and Leroux, S.: Simulated Impact of Time-Varying River Runoff and Greenland Freshwater Discharge on Sea Level Variability in the Beaufort Gyre Over 2005–2018, Journal of Geophysical Research: Oceans, 129, e2024JC021 237, https://doi.org/10.1029/2024JC021237, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JC021237, 2024.

Tao, H., Abba, S. I., Al-Areeq, A. M., Tangang, F., Samantaray, S., Sahoo, A., Siqueira, H. V., Maroufpoor, S., Demir, V., Dhanraj Bokde, N., Goliatt, L., Jamei, M., Ahmadianfar, I., Bhagat, S. K., Halder, B., Guo, T., Helman, D. S., Ali, M., Sattar, S., Al-Khafaji, Z., Shahid, S., and Yaseen, Z. M.: Hybridized artificial intelligence models with nature-inspired algorithms for river flow modeling: A comprehensive review, assessment, and possible future research directions, Engineering Applications of Artificial Intelligence, 129, 107 559, https://doi.org/10.1016/j.engappai.2023.107559, 2024.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M. M.: GRACE Measurements of Mass Variability in the Earth System, Science, 305, 503–505, https://doi.org/10.1126/science.1099192, 2004.

https://doi.org/10.5194/essd-2026-853
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Thouvenin-Masson, C., Boutin, J., Échevin, V., Lazar, A., and Vergely, J.-L.: Influence of river runoff and precipitation on the seasonal and interannual variability of sea surface salinity in the eastern North Tropical Atlantic, Ocean Science, 20, 1547–1566, https://doi.org/10.5194/os-20-1547-2024, 2024.

Trenberth, Jones, P., Ambenje, Bojariu, R., Easterling, D., Tank, K., Parker, D., Rahimzadeh, F., Renwick, J., Rusticucci, M., Soden, B., Zhai, P., and Folland, C.: Observations: Surface and Atmospheric Climate Change, pp. 235–336, 2007.

Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., Danabasoglu, G., Suzuki, T., Bamber, J. L., Bentsen, M., Böning, C. W., Bozec, A., Chassignet, E. P., Curchitser, E., Boeira Dias, F., Durack, P. J., Griffies, S. M., Harada, Y., Ilicak, M., Josey, S. A., Kobayashi, C., Kobayashi, S., Komuro, Y., Large, W. G., Le Sommer, J., Marsland, S. J., Masina, S., Scheinert, M., Tomita, H., Valdivieso, M., and Yamazaki, D.: JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do), Ocean Modelling, 130, 79–139, https://doi.org/10.1016/j.ocemod.2018.07.002, 2018.

Vallis, G. K.: Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation, Cambridge University Press, 2 edn., 2017.

Vaujour, T., Kern, L., Storto, A., Pfeffer, J., Thenoz, C., Yang, C., and Guinehut, S.: F3O outlet river discharge dataset, https://doi.org/https://doi.org/10.5281/zenodo.23062294, 2026.

Yang, Y., Feng, D., Beck, H. E., Hu, W., Abbas, A., Sengupta, A., Delle Monache, L., Hartman, R., Lin, P., Shen, C., and Pan, M.: Global Daily Discharge Estimation Based on Grid Long Short-Term Memory (LSTM) Model and River Routing, Water Resources Research, 61, e2024WR039 764, https://doi.org/https://doi.org/10.1029/2024WR039764, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024WR039764, 2025.