Improving Satellite Remote Sensing Estimates of the Global Terrestrial Hydrologic Cycle via Neural Network Modeling

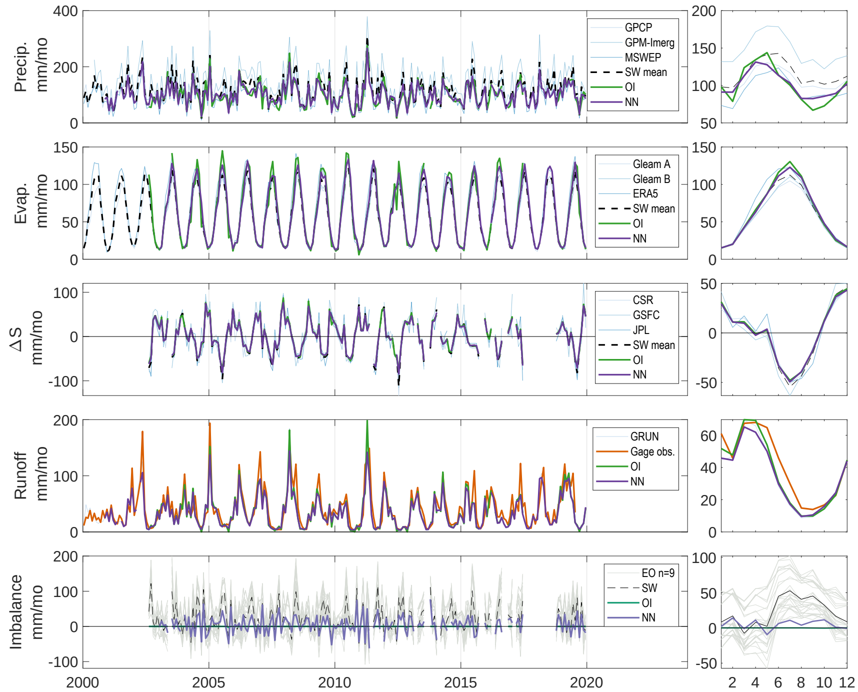
Matthew G Heberger[1], Victor Pellet[2], and Filipe Aires[3]

[1]Paris Observatory
[2]The University of Tokyo
[3]LERMA

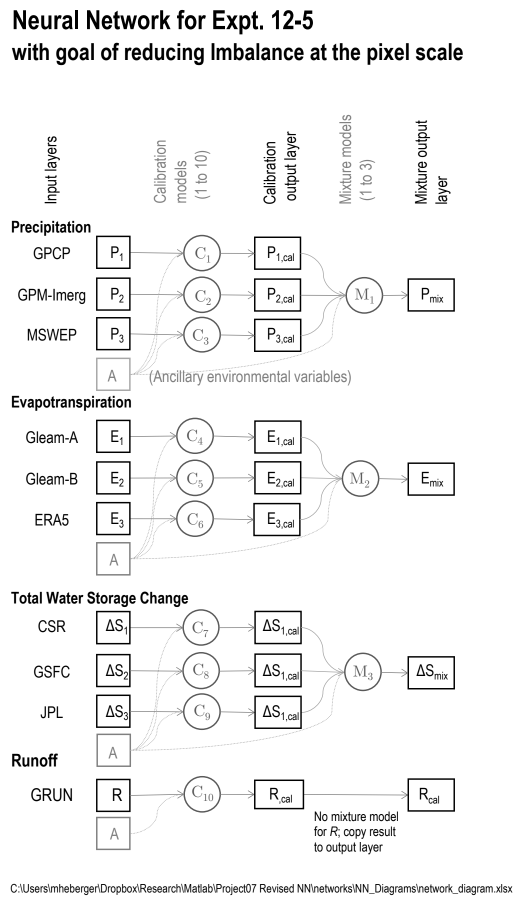
July 27, 2023

Abstract

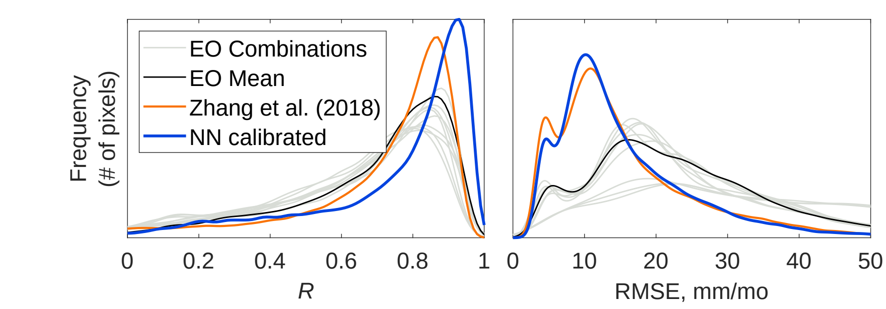
Satellite remote sensing is commonly used to observe the hydrologic cycle at spatial scales ranging from river basins to the globe. Yet it remains difficult to obtain a balanced water budget using remote sensing data, which highlights the errors and uncertainties in earth observation (EO) data. Various methods have been proposed to correct EO datasets to make them more coherent, so that they result in a more balanced water budget. This study aimed to improve estimates of water budget components (precipitation, evapotranspiration, runoff, and total water storage change) at the global scale using the methods of optimal interpolation (OI) and neural network (NN) modeling. We trained a set of NNs on a set of 1,358 river basins and validated them on an independent set of 340 basins and in-situ observations of evapotranspiration and river discharge. We extended the models to make pixel-scale predictions in 0.5° grid cells for near-global coverage. Calibrated datasets result in lower water budget residuals in validation basins: the mean and standard deviation of the imbalance is 11 $\pm$ 44 mm/mo when calculated with uncorrected EO data and 0.03 $\pm$ 24 mm/mo after calibration by the NN models. This study suggests to data producers where corrections should be made to the EO datasets, and demonstrates the benefits of physically-driven NN models for studying the hydrologic cycle at the global scale.



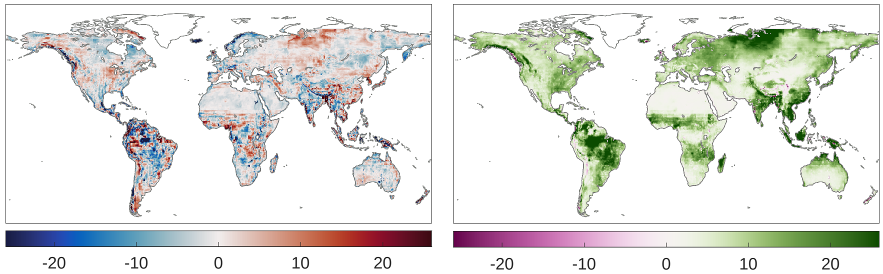
2023-06-29 15:06:45 C:\Users\mheberger\Dropbox\Research\Matlab\Project12 S with Runoff\expt4\m30_ts_plots_for_paper.m



Neural Network for Expt. 12-5
with goal of reducing Imbalance at the pixel scale
Input layers
Calibration models (1 to 10)
Calibration output layer
Mixture models (1 to 3)
Mixture output layer
Precipitation
GPCP
GPM-Imerg
MSWEP
(Ancillary environmental variables)
Evapotranspiration
Gleam-A
Gleam-B
ERA5
Total Water Storage Change
CSR
GSFC
JPL
Runoff
GRUN
No mixture model for R; copy result to output layer
C:\Users\mheberger\Dropbox\Research\Matlab\Project07 Revised NN\networks\NN_Diagrams\network_diagram.xlsx

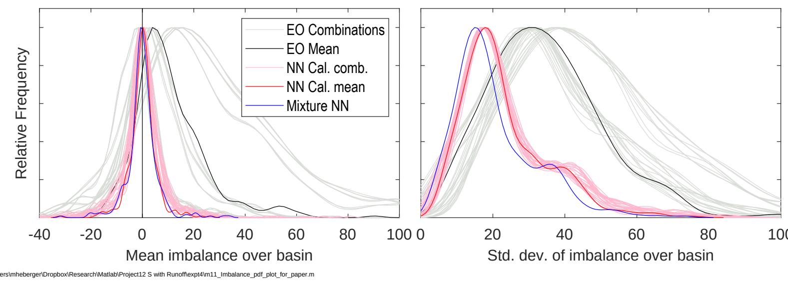


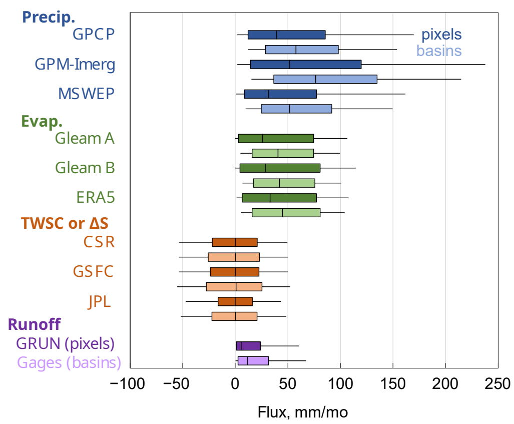
EO Combinations
EO Mean
Zhang et al. (2018)
NN calibrated
Frequency (# of pixels)
0
0.2
0.4
0.6
0.8
1
R
0
10
20
30
40
50
RMSE, mm/mo
2023-06-29 15:44:41 C:\Users\mheberger\Dropbox\Research\Matlab\Project12 S with Runoff\expt9\m12_comparison_my_vs_zhang_smoothed.m



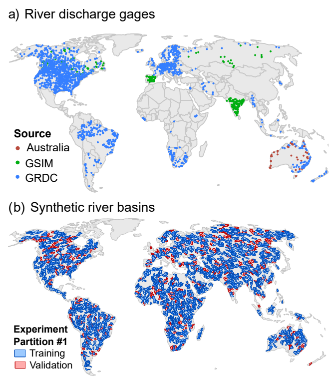
(a) Mean imbalance in pixels (mm/mo)

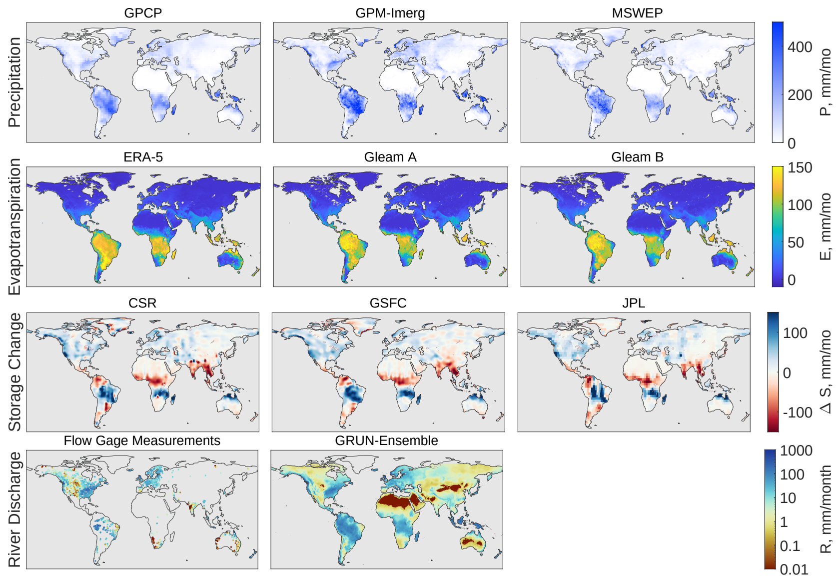
(b) Mean imbalance improvement in pixels (mm/mo)

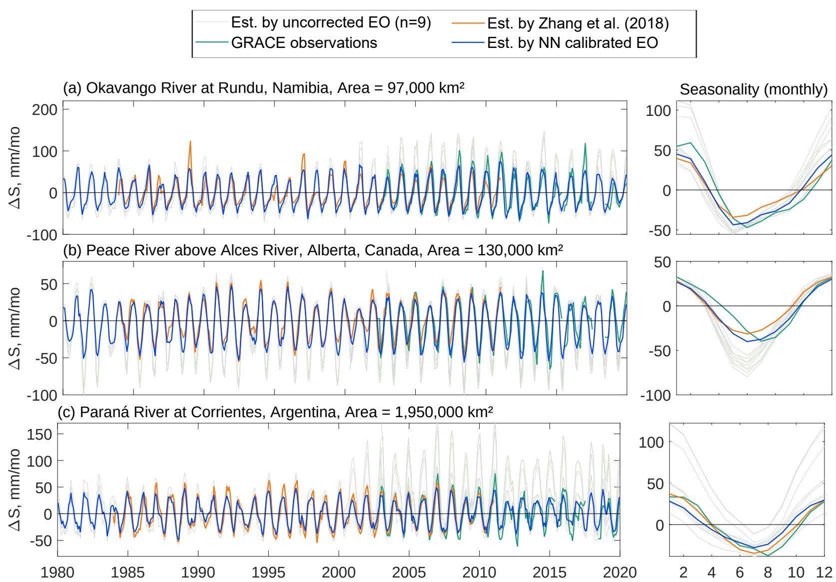




Precip.
GPCP
GPM-Imerg
MSWEP
Evap.
Gleam A
Gleam B
ERA5
TWSC or ΔS
CSR
GSFC
JPL
Runoff
GRUN (pixels)
Gages (basins)
pixels
basins
−100
−50
0
50
100
150
200
250
Flux, mm/mo


a) River discharge gages
Source
Australia
GSIM
GRDC
(b) Synthetic river basins
Experiment
Partition #1
Training
Validation


GPCP
GPM-Imerg
MSWEP
Precipitation
P, mm/mo
400
200
0
ERA-5
Gleam A
Gleam B
Evapotranspiration
E, mm/mo
150
100
50
0
CSR
GSFC
JPL
Storage Change
Δ S, mm/mo
100
0
-100
Flow Gage Measurements
GRUN-Ensemble
River Discharge
R, mm/month
1000
100
10
1
0.1
0.01


Est. by uncorrected EO (n=9)
Est. by Zhang et al. (2018)
GRACE observations
Est. by NN calibrated EO
(a) Okavango River at Rundu, Namibia, Area = 97,000 km²
Seasonality (monthly)
(b) Peace River above Alces River, Alberta, Canada, Area = 130,000 km²
(c) Paraná River at Corrientes, Argentina, Area = 1,950,000 km²
ΔS, mm/mo
1980
1985
1990
1995
2000
2005
2010
2015
2020
2
4
6
8
10
12

Improving Satellite Remote Sensing Estimates of the Global Terrestrial Hydrologic Cycle via Neural Network Modeling

Matthew Heberger[1,2,3], Filipe Aires[2,3], Victor Pellet[2,3]

[1

]Sorbonne University, Paris, France

[2

]LERMA, Paris Observatory

[3

]ESTELLUS

Key Points:

- Uncorrected remote sensing datasets cannot be combined for a balanced water budget.
- Optimal interpolation and neural network modeling can be used together to readjust datasets and reduce the water budget imbalance.
- Results can be used to show where remote sensing datasets are biased, to fill in missing data, and for hindcasting.

Corresponding author: M. Heberger, matthew.heberger@obspm.fr

Abstract

Satellite remote sensing is commonly used to observe the hydrologic cycle at spatial scales ranging from river basins to the globe. Yet it remains difficult to obtain a balanced water budget using remote sensing data, which highlights the errors and uncertainties in earth observation (EO) data. Various methods have been proposed to correct EO datasets to make them more coherent, so that they result in a more balanced water budget. This study aimed to improve estimates of water budget components (precipitation, evapotranspiration, runoff, and total water storage change) at the global scale using the methods of optimal interpolation (OI) and neural network (NN) modeling. We trained a set of NNs on a set of 1,358 river basins and validated them on an independent set of 340 basins and in-situ observations of evapotranspiration and river discharge. We extended the models to make pixel-scale predictions in 0.5° grid cells for near-global coverage. Calibrated datasets result in lower water budget residuals in validation basins: the mean and standard deviation of the imbalance is 11 ± 44 mm/mo when calculated with uncorrected EO data and 0.03 ± 24 mm/mo after calibration by the NN models. This study suggests to data producers where corrections should be made to the EO datasets, and demonstrates the benefits of physically-driven NN models for studying the hydrologic cycle at the global scale.

Plain Language Summary

Today, satellite remote sensing can measure all the major flows in the water cycle. This includes precipitation, evaporation, and changes in the amount of water stored underground and in lakes and reservoirs. These flows are related to one another via the water cycle: flows into and out of any region should be balanced. Yet, we cannot calculate a balanced water budget with satellite data. This shows that further improvements to these data are possible. We used a method called optimal interpolation to make corrections to satellite datasets, so that they better balance the water budget. However, this method only works at the scale of river basins, where river discharge data are available. To extend our findings to other locations, we created a statistical model based on machine learning. This model can make predictions at the pixel scale over most of the earth's land surface. Often, our model can improve satellite observations of the water cycle. Further, it provides useful information about when and where corrections are most needed. Our study shows that machine learning methods can help improve data from satellites related to the global water cycle.

1 Introduction

The water cycle, or hydrologic cycle (HC), is an important field of study for earth scientists — changes to the HC have broad societal implications, with effects on drought, flooding, agriculture, and water supply. And while enormous progress has been made in monitoring the water cycle via remote sensing, capturing a complete picture from space remains a difficult goal. The usefulness of earth observation (EO) datasets has not been fully achieved because of "incoherence" among various data products – studies have demonstrated that the water budget cannot be closed using remote sensing data without significant errors (Hegerl et al., 2015; Rodell et al., 2015; McCabe et al., 2017). This leads to the conclusion that satellite datasets still suffer from systematic bias or random errors.

In this study, we seek to simultaneously optimize multiple satellite-estimated datasets of hydrologic fluxes, reconciling them to create a balanced water budget. A simplified water budget for any land area (e.g., river basin, grid cell) includes the four main fluxes or *HC components*: precipitation, P, evapotranspiration E, total water storage change (TWSC in the text and ΔS in equations), and runoff, R. By conservation of mass, the water budget can be stated:

$$P - E - \Delta S - R = 0 \quad (1)$$

Conventional optimization methods have focused on calibrating individual components of the HC one at a time (for example fitting P or E to ground-based observations). This approach is fundamental, but as one author wrote, it "ignores the interdependencies and relationships inherent in observed responses" (McCabe et al., 2017). In other words, we may be able to correct P by exploiting valuable information in the variables E, ΔS, and R, as they are related via the HC (Equation 1).

Recent research has shown that simultaneously optimizing multiple HC components can result in a more balanced water budget. Many of these approaches focus on "assimilation" of EO into hydrological models (see e.g., Yilmaz et al., 2011; Zhang et al., 2016; Wong et al., 2021). Several recent studies focused on closing the water budget with a more data-driven approach. One class of studies estimates a single component as a function of the other three. Some authors have made the simplifying assumption that, over sufficiently long time periods, $\Delta S = 0$ (i.e. no trend in storage), allowing one to estimate total runoff (including subsurface flow) as $R = P - E$ (Liu et al., 2020). Rodell et al. (2011) estimated evapotranspiration over seven large river basins via the relation $E = P - R - \Delta S$, using the output of several land surface and atmospheric models for the right side of this equation. The authors concluded that the uncertainty in ΔS measured by the GRACE satellites is too high to produce useful monthly estimates of E, but that predicted seasonal patterns are fairly reliable. In another example, Lehmann et al. (2022) estimated $\Delta S = P - E - R$, and compared predictions to GRACE observations. They performed this analysis over 189 large river basins covering 90% of the continental land area. Rather than seeking to optimize the datasets, the authors looked for the *best combination* of inputs, and reduced the imbalance through "cancellation of errors in poor estimates of water budget components."

Aires (2014) introduced an integration method called optimal interpolation (OI). This closed-form analytical solution imposes a HC budget closure constraint. It forces the imbalance to zero, and modifies each of the HC components by an amount inversely proportional to its uncertainty. Aires showed that this constraint improves the estimation of the HC components in some places and times. In a related paper Munier et al. (2014) applied OI over the 3 million km^2 Mississippi River basin, revising satellite estimates for P, E, R, and ΔS. Later, Munier and Aires (2018) applied OI over 11 large river basins. OI has also been shown to work well in optimizing satellite observations of the hydrologic cycle over river basins in the Mediterranean (Pellet et al., 2018), South Asia (Pellet, Aires, Papa, et al., 2019), and the Amazon (Pellet et al., 2021).

A major limitation of OI is that it can only be used at the basin scale, where observations of river discharge are available. Yet, the vast majority of the world's rivers and streams are ungaged, and gage data are particularly sparse in less-developed countries. Furthermore, to make truly global predictions, a model must be able to make predictions at the pixel scale. Munier and Aires (2018) extrapolated the balanced solution from OI to the global scale with the help of auxiliary environmental information. However, environmental datasets were used in a rather simple way; the authors did not use them as explanatory variables, but rather to divide basins into classes based on climate regime.

We hypothesized that OI solutions could be extrapolated to new locations more accurately with a more complex model and with more inputs to describe the environment. We attempt to do this here by using environmental data as input variables to a flexible neural network (NN) model. Our approach involves two main steps. First, we use OI over a pre-defined set of river basins. The solution is an optimized set of HC components which satisfy the closure constraint. Next, we train a NN model to *calibrate* EO datasets, with a goal of making them closer to the optimized version calculated by OI.

We further hypothesized that the necessary adjustments are complex and non-linear, and vary by time and location. This makes the problem well suited to approaches based on machine learning.

Our method uses *supervised learning*, where a target is provided to *train* the NN. The purpose of training is to find the set of model parameters to the function that best relates the input(s) to the target(s). Here, we use the output of the OI algorithm as the target. The NN output is a set of calibrated monthly estimates for P, E, R, and ΔS in each basin. Again, these are not calibrated in the conventional sense of fitting to in situ observations, but by combining information from multiple remote sensing datasets while seeking to satisfy the HC closure constraint (Equation 1). We used environmental data (elevation, slope, vegetation, etc.) as inputs to the model, hypothesizing that these additional inputs will help the NN to find an optimal solution under varying conditions.

This study has two main objectives. First, to optimize hydrologic EO datasets and to calculate a balanced water budget at the river basin scale from these data. Second, to train a NN model based on these results to make improved estimates at the pixel scale. A third, stretch goal for the study was to test the model's ability to estimate missing data via inference. For example, we can estimate GRACE-like TWSC by rearranging Equation 1 to give $\Delta S = P - E - R$. This would allow us to fill in missing data or to estimate water storage from before GRACE was launched in 2002, or similarly, to estimate runoff in ungaged basins.

2 Datasets

We created a database of earth observations (EO) based on satellite remote sensing, with datasets that quantify each of the four major fluxes in Equation 1. The EO datasets are summarized in Table 1. All fluxes are expressed in area-normalized units of depth per time in millimeters per month (mm/mo). Our database covers the 20-year time period from January 2000 to December 2019, a total of 240 months. This time period was chosen to overlap with the availability of observations of TWSC from the GRACE satellites, launched in 2002.

As can be seen in Table 1, the EO datasets vary in terms of their spatial and temporal resolution, posing a challenge to their integration. We put all EO data into the same 0.5° equirectangular grid, based on latitude and longitude, rescaling and projecting as necessary. In theory, the analysis could be performed at any time scale (daily, weekly, etc.) and at any spatial resolution. We chose 0.5° resolution to be compatible with the runoff dataset GRUN. We computed monthly averages for all variables where needed. We chose a monthly time scale to be compatible with GRACE TWSC data. Finally, we evaluated the quality and completeness of each dataset and discarded anomalous observations.

2.1 Total Water Storage Change

Information on total water storage (TWS), comes from the GRACE (Gravity Recovery and Climate Experiment) satellites. The first pair of satellites were in operation from 2002-2017, and a follow-on mission began in 2018. GRACE makes detailed measurements in changes to the Earth's gravity field over time. Most short-term changes are due to the movement of water on land and underground (Tapley et al., 2004). GRACE data have been used in groundbreaking studies to analyze the terrestrial water budget, drought, climate change, and water management (see e.g., Famiglietti et al., 2011; Richey et al., 2015; McCabe et al., 2017; Rodell et al., 2018).

GRACE provides the monthly TWS *anomaly*, expressed as a liquid water equivalent thickness, in units of cm or mm. GRACE does not estimate the total volume or

Table 1. Datasets compiled for the four major fluxes in the hydrologic cycle

Dataset	Begin	End	Temporal res.	Spatial res.	Citation
Water Storage Anomaly					
GRACE-CSR	2002	present	month	1.0°	Save (2020)
GRACE-JPL	2002	present	month	1.0°	Landerer and Cooley (2021)
GRACE-GSFC	2002	present	month	1.0°	Loomis et al. (2019)
Precipitation					
GPCP v2.3	1979	present	day	2.5°	Adler et al. (2018)
GPM IMERG	2000	present	day	0.10°	Huffman et al. (2020)
MSWEP	1979	present	day	0.10°	Beck et al. (2019)
Evapotranspiration					
GLEAM v3.5a	1980	present	day	0.25°	Martens et al. (2017)
GLEAM v3.5b	2003	present	day	0.25°	idem
ERA5	1950	present	3 hour	0.25°	Hersbach et al. (2018)
Observed E for validation					
E: FluxNet 2015	2002	2010	hour	point	Pastorello et al. (2020)
Observed River Discharge - in situ					
GRDC	varies	varies	day	gage	BfG (2020)
Australia	1970	2020	day	gage	Australia BOM (2020)
GSIM	varies	2016	month	gage	Gudmundsson et al. (2018)
Runoff - synthetic					
G-RUN	1902	2019	month	0.5°	Ghiggi et al. (2021)

mass of water in a region, but rather its change with respect to a historical baseline. Nevertheless, the observations encompass water in all its forms and "represent the full magnitude of land hydrology and land ice" (Landerer, 2021). We obtained three different GRACE products (see Table 1). Each is based on the mass concentration solution developed by the Jet Propulsion Laboratory (JPL), known as *mascon*. This technique employs a gravity field basis function to separate the contributions in the signal from unequal distribution of the earth's mass from other factors such as water storage variations.

We calculated the month-over-month *rate of change* in water storage to provide the flux in mm/mo. This converts the TWS anomaly to a flux, TWSC or ΔS, and creating the link between GRACE data and the other variables in the HC. There are several methods for calculating the rate of change, but most researchers in this field use simple finite difference methods (see e.g., Landerer & Swenson, 2012; Biancamaria et al., 2019). We used the backwards finite difference method.

$$\frac{\Delta S}{\Delta t} = \frac{S_t - S_{t-1}}{t - (t-1)} \tag{2}$$

The results we obtained from more complex methods such as fitting a cubic spline or using an "equivalent smoothing filter" (see e.g., Landerer et al., 2010) were comparable to those obtained with the simpler methods but often resulted in more missing observations, therefore we used the simple method in Equation 2.

Modeled TWSC - As a source of validation data, we also collected predictions from a recent modeling study that reconstructed GRACE-like TWSC via other water cycle components. Zhang et al. (2018) used a land surface model and data assimilation techniques, first estimating the errors in each water budget component by comparison to in situ observations, then using a constrained Kalman filter to merge the datasets based on their error information, with a goal of minimizing the imbalance. This study produced global gridded datasets at 0.5° resolution, with monthly P, E, R, and ΔS for 1984–2010.

2.2 Precipitation

We obtained data from three sources (Table 1).

GPCP - The Global Precipitation Climatology Project (GPCP) has the longest time record, beginning in 1979 (Adler et al., 2018). It also has the coarsest spatial resolution, at 2.5°. This dataset, produced by an international consortium of researchers, is based on multiple satellite observations that are merged to estimate precipitation at the global scale. We used version 2.3 of this dataset, which was updated in 2018.

GPM-Imerg - GPM-Imerg is the multi-satellite precipitation product from NASA. IMERG combines data from multiple low-earth orbit satellites and geosynchronous orbiting infrared satellites, using morphing techniques and a Kalman filter, to provide accurate satellite-based precipitation estimates, supplemented by precipitation gauge analyses.

MSWEP - The Multi-Source Weighted-Ensemble Precipitation (MSWEP) is not a pure remote sensing product, but an “optimal merging” of gage observations, satellite observations, and reanalysis model output (Beck et al., 2019). MSWEP has been shown to be more accurate over mountainous regions, where many products consistently underestimate P.

2.3 Evapotranspiration

Evapotranspiration, E, is the upward flux of water from the land to the atmosphere, combining free-surface *evaporation* with *transpiration*, the flux of water from plant leaves to the atmosphere. It is an important driver of the global climate, responsible for the exchange of water and energy from the land and sea surface to the atmosphere. It has been estimated that as a global average, E is about 60-75% of precipitation (Shiklomanov, 2009). E cannot be measured directly via remote sensing. Rather, scientists measure land surface temperature or near-surface air temperature and use empirical relationships to estimate E.

Gleam - The Global Land Evaporation Amsterdam Model (GLEAM) is a set of algorithms that estimates the various components that contribute to total E: transpiration, bare-soil evaporation, interception loss, open-water evaporation, and sublimation (Martens et al., 2017; Miralles et al., 2011; Hersbach et al., 2018). The authors used an empirical relationship, the Priestley-Taylor equation, to calculate potential E based on satellite observations of surface net radiation and near-surface air temperature. GLEAM version 3.5a used reanalysis rather than satellite observation, and covers 1980 to present. The updated version 3.5b relies more on remote sensing data, and has a more limited temporal coverage of 2003 to present.

ERA5 - We also included a dataset that is not a purely remote-sensing based product, but based on the assimilation model ERA5, from the European Centre for Medium-Range Weather Forecasts (ECMWF). The model combines historical estimates (from both remote sensing and in situ observations) using an advanced modeling and assimilation system. ERA5 produces many variables describing the atmosphere, land, and ocean, at a resolution of up to a 30 km grid (Guillory, 2022). ERA5 estimates of E have been used in many recent hydroclimatic studies (see e.g., Tarek et al., 2020; Singer et al., 2021; Lu et al., 2021).

Observed evapotranspiration at flux towers - In order to validate our results, we obtained in situ measurements of E from flux towers, where latent heat flux and other measurements are obtained via eddy covariance methods. We selected data for 117 towers from the FluxNet2015 dataset, which compiles data from 212 global towers (Pastorello et al., 2020). The majority of selected towers are in Europe (51 towers) or North America (46), with fewer in Africa (2), Asia (6), Australia (9), and South America (3).

Care must be taken in comparing E observed at flux towers to gridded hydro-climatic data. The value in a single grid cell (or pixel) represents an average for an area over which conditions can vary widely. At the scale of our model grid, a single 0.5° pixel has an area of about 3,000 km^2 near the equator. Land cover, vegetation, and topography over a grid cell may vary drastically from those at the flux tower site. This limits the meaningfulness of comparisons between gridded model estimates flux tower observations.

2.4 Runoff

River flow, or *discharge*, is an in situ measurement, measured at *gages*, typically operated by public agencies. The gage location is considered the outlet of a river basin, and the discharge is the sum of basin runoff. The terms *runoff* and *river discharge* are frequently the source of confusion, and care must be taken to distinguish these related but distinct quantities. While definitions vary among sources, here we follow the definitions used by Ghiggi et al. (2019). Runoff is defined as all the water draining from a small land area, and cannot be observed directly. River discharge, by contrast, is measured at a single point on a river. One may estimate river discharge from the runoff in the upstream area by spatially averaging the gridded runoff data. As the travel time of water in the river system is neglected, this method can only be assumed to be correct over long time scales. We assume that on a monthly scale, the effect of water routing is negligible for small- to mid-size basins.

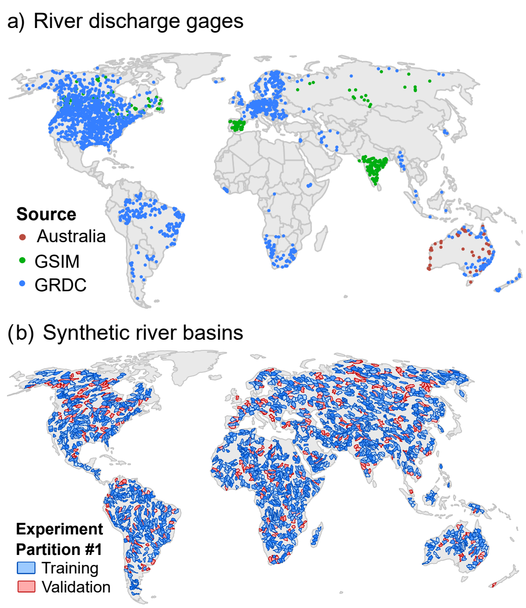


Figure 1. Map of this study's river basins: (a) 2,056 river flow gaging stations, corresponding to the basin outlets (basin boundaries not shown); (b) 1,698 synthetic river basins created for training and validating the neural network model.

We sought to develop a large database of global gaged basins that would represent a range of geographic locations, environments, and basin sizes to better sample the global water cycle on Earth. We selected gages for our analysis based on data quality, geographic and temporal coverage, and location. We considered gages with an upstream area ≥2,500 km^2.

We obtained river dischage data from 3 sources. First, we selected 1,737 gages from the Global Runoff Data Center (GRDC) and supplemented it with information from two other sources to fill in blank spaces on the map (notably Asia and Australia). The GRDC database contains historical mean daily and monthly discharge data from 159 countries (WMO, 1989; BfG, 2020). The GRDC database contained 10,361 stations when we acquired data, however the majority of these did not fit our criteria for spatial and temporal coverage. Second, we obtained data for 272 gages from the Global Streamflow Indices and Metadata (GSIM) archive (Do et al., 2018; Gudmundsson et al., 2018). Finally, we obtained runoff data for 47 gages in Australia from their Bureau of Meteorology (BOM)'s Hydrologic Reference Stations (Australia BOM, 2020).

We calculated monthly average runoff for months with at least 25 days of data. Volumetric flow rates in m^3/s were converted to area-normalized fluxes in mm/mo by dividing by the land surface area in km^2 and multiplying by an appropriate conversion factor. The spatial coverage of our final 2,056 river gages (and their basins) is uneven across the globe (see Figure 1). North America is over-represented with 1,111 gages (more than half the total), as is Europe with 393 gages, while we have only 70 gages in Africa, 178 in Asia, and 195 in South America.

Runoff observations are limited, as they are only available at gaged locations. As an alternative, indirect estimates of runoff are available from several sources. For our experiments in closing the HC, we used estimated runoff from GRUN Ensemble (Ghiggi et al., 2021). The authors created a global gridded dataset of runoff with a random forest model using P and near-surface temperature as predictor variables. For the 2021 GRUN Ensemble project, the authors used input data from 21 different sources, "including a set of atmospheric reanalysis, post-processed reanalysis and interpolated-stations data."

In order to check the quality of the GRUN dataset, we performed an independent evaluation against our 2,056 gages and found that GRUN is a relatively good fit to observed discharge. We first estimated the monthly discharge at the basin outlet by calculating the spatially averaged mean of gridded GRUN runoff. Then we calculated fit statistics comparing the observed and modeled flow time series. We found that a median correlation $R = 0.84$ and median root mean square error, RMSE = 11.8 mm/mo, and 75% of gages had RMSE $<$ 19 mm/mo. We also calculated a common fit indicator for modeled discharge, the Kling-Gupta Efficiency (KGE). Median KGE is 0.53, and 81% of gages have KGE > -0.41, the point at which a model's predictions are better than the mean of observations (Knoben et al., 2019).

2.5 Environmental Indices

We also collected observations of ancillary environmental data as inputs to our NN model. Our hypothesis is that errors in EO data (that the NN will attempt to correct) are the consequence of certain environmental conditions. For example, precipitation estimates are often biased in mountainous regions, or in relation to snow cover. The environmental data are listed with their source/citation below. In all cases, we rescaled and reprojected datasets as necessary and calculated spatial means for river basins as described above. The 12 environmental indices are:

1. Aridity index (Trabucco & Zomer, 2019)
2. Mean elevation (Amatulli et al., 2018)
3. Median slope (Amatulli et al., 2018)

4. Basin centroid latitude (calculated)
5. Enhanced vegetation index (Didan, 2015)
6. Vegetation growth/senescence (calculated)
7. Irrigated area (Siebert et al., 2015)
8. Fire: burned area (Giglio et al., 2020)
9. Snow cover (Hall & Riggs, 2021)
10. Solar radiation (Hogan, 2015)
11. Land surface temperature (Wan et al., 2021)

2.6 Preliminary Analysis of EO Datasets

Figure 2 shows a snapshot of one month (January 2005) of the EO datasets used as input in our analysis. The different datasets share many similarities in terms of the overall patterns, but there are many differences. For example, GPCP precipitation appears smoother, while the other two datasets, which have a higher spatial resolution, show finer-grained patterns of rainfall. This is particularly evident over the Amazon and southern Africa. Similarly, one can see differences in the spatial patterns and range of magnitude of E and ΔS. River discharge, measured at gages, has a sparser coverage, and the distribution of flows is highly skewed, with measured discharges covering several orders of magnitude from 0 to nearly 1,000 mm/mo.

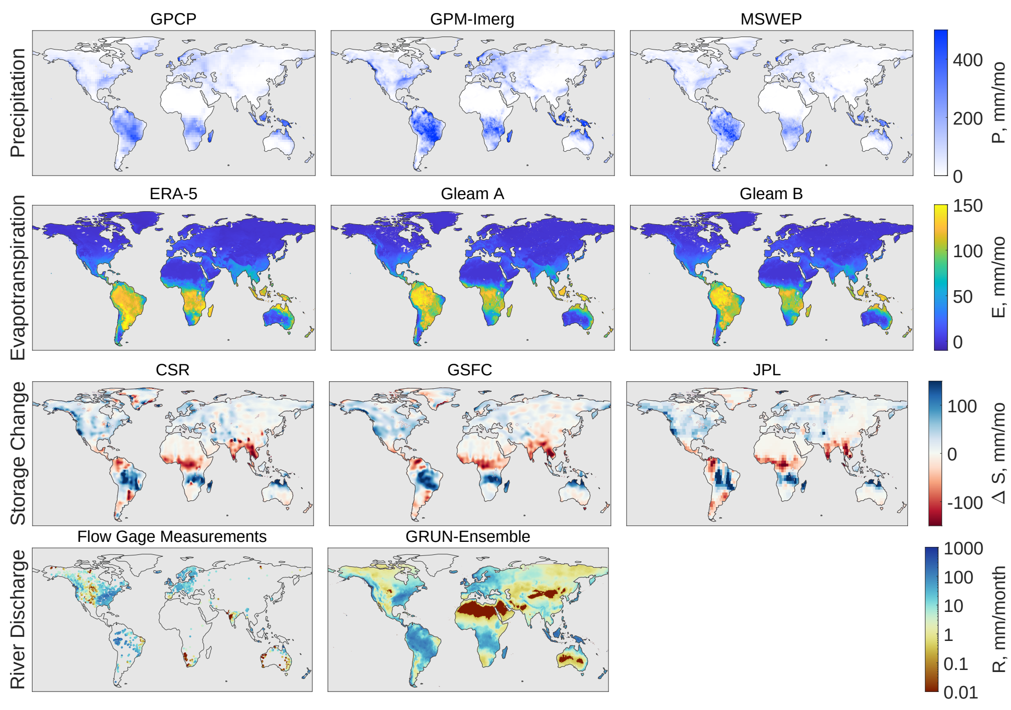


Figure 2. Snapshot of EO data for a single month, January 2005.

Figure 3 shows the distribution of values in the EO datasets used as input in our model using standard boxplot conventions (boxes = interquartile range, whiskers = 10%ile and the 90%-ile). The top boxplot in each set of observations is for all pixels over land, while the lower box shows the distribution across our 2,056 gaged basins. For most variables, the distribution of fluxes is greater over the pixels compared to the basins, with higher highs and lower lows. This is particularly the case for P, but is also seen with E.

When we calculate the mean flux over a basin, it tends to smooth out the extremes and compress the distribution of observed fluxes. There are also differences in the distributions within each category of fluxes. For example, GPM-Imerg contains higher observations of P, with a higher 75- and 90-percentile than the other two datasets. The monthly water storage change, ΔS, is centered at about zero for each dataset. This is expected, as the storage in pixels and basins tends to fluctuate seasonally, and any long-term trend is small compared to the annual variations in storage. Runoff has the smallest magnitude of any of the hydrologic fluxes, with a low of 0 mm/mo (no observed flow) to a 90%-ile of 68 mm/mo, lower than the 90%-ile of P or E. We conclude that the lack of consensus among datasets is further evidence of the need for them to be reconciled.

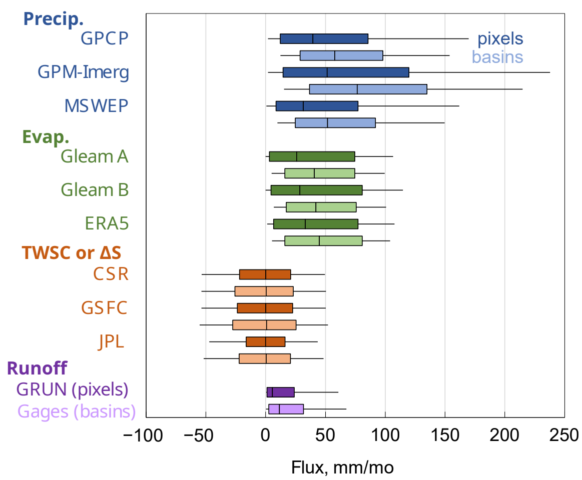


Figure 3. Boxplots showing the distribution of values in the EO datasets at the pixel scale over continents (except Antarctica and Greenland) and averaged over gaged river basins used in this study.

3 Methodology

3.1 Training database at basin scale over the world

We analyzed the water balance at the scale of two geographic units: river basins and pixels. For the pixel-scale analysis, we used a 0.5° equirectangular grid. We excluded Antarctica, Greenland, and the Arctic north of 77°. Near the equator, a single pixel is about 56 km on a side and has an area of about 3,100 km^2. The disadvantage to studying the water balance at the pixel scale is the lack of observations of horizontal inflow or outflow. In this regard, there are advantages to working at the scale of river basins, or watersheds. A watershed is defined as the area on the Earth's surface where water drains to a common outlet and is determined by the topography of the land surface.

We obtained basin geodata in shapefile format from the GRDC, which covered many of our gaged basins (Lehner et al., 2008). However, some of these basin boundaries appeared to be inaccurate. Therefore, we created a new set of boundaries for every watershed using the best-available global hydrographic data. We used a hybrid method that uses both vector- and raster-based data (Heberger, 2022). Our method uses the vector dataset MERIT-Basins (Lin et al., 2019, 2021), where rivers are encoded as polylines and catchment boundaries as polygons.

Our resulting 2,056 basins vary in size from 2,500 km^2 to 4.7 million km^2 for the Amazon basin. The distribution of basin sizes is highly skewed, meaning that we have many small- and medium-sized basins, and fewer large basins. Of these, 119 watersheds have an area of less than 3,000 km^2. Our set of gaged basins covers 47 million km^2, about 35% of our study domain, the land surface below 77° North and excluding Greenland and Antarctica.

For each basin, we calculated the average for each of EO variables in Tables 1 and the environmental variables described above. To calculate the spatial weighted mean, we converted each basin polygon to a grid "mask," where each pixel is a floating-point number representing the fraction of the pixel's area that is inside the basin (from 0 to 1). Because the surface area of pixels varies by latitude, we used the pixel's area in our calculation of the weighted mean.

When working with a gridded runoff dataset, we are free from the constraint of using gaged basins, and we may define river basins of any size and at any location for better sampling of many environmental conditions. To take full advantage of this, we created a set of 1,698 *synthetic* river basins, shown in Figure 1(b). These represent real, physical basins, but their outlets do not correspond to a gage. We created the synthetic basins by using a gridded dataset of flow direction created by the developers of the GloFaS-LISFLOOD model (Harrigan et al., 2020) and the Python library *pysheds* (Bartos et al., 2023). The basins range in size from 20,000 to 50,000 km^2; the relatively small size allows fits the hypothesis that we may neglect water travel time at the monthly time scale. The color coding in Figure 1 shows the experimental partition we created for the training and validation of the NN model, with 80% of basins for training (in blue), and 20% of basins for validation (in red). The result is a set of 1,698 basins shown in Figure 1(b).

3.2 Optimal Interpolation

We follow previous studies (Aires, 2014; Pellet, Aires, Munier, et al., 2019) in using OI to integrate EO datasets and balance the water budget at the river basin scale. We refer to the water balance residual as the *imbalance*, calculated by:

$$I = P - E - \Delta S - R. \tag{3}$$

The OI approach is based on forcing I to equal zero, or minimizing I, while distributing the errors among the inputs in inverse proportion to each variable's uncertainty. These methods, well described by Rodgers (2000), are referred to as "inverse methods" and are widely used in remote sensing. The goal of OI is to combine these multiple estimates to obtain the best consensus of the HC state. For a detailed explanation of the mathematics behind OI, see Aires (2014) and Pellet, Aires, Munier, et al. (2019).

We begin by defining an initial best guess for each of the four HC components. This is done by calculating a weighted average of the inputs for each component. Simple weighting (SW), as described by Aires (2014), is an application of the method of inverse variance weighting to the problem of calculating the "best estimate" that combines multiple satellite observed fluxes. It is a form of weighted averaging where the weight on the each observation is the inverse of the variance or uncertainty of that observation. The uncertainty must be estimated a priori for each observation or dataset. After calculating the best first guess of the water budget based on the SW mean, we apply a post-filter to enforce the water balance. The post filter is a linear transformation based on the uncertainty in each component. Aires (2014) derived a solution for determining the linear combination of variables that satisfies the water budget constraint, weighting the contribution such that variables with lower error variance receive greater weight.

The OI method is simple and effective. Further, it has the advantage of not relying on any model. When it is applied strictly (e.g., without an optional relaxation factor described by Pellet, Aires, Munier, et al. (2019)), it will always result in a balanced water budget. However, this strict requirement can also produce unrealistic results. The OI method does not guard against returning negative values, which is obviously unrealistic for precipitation or runoff. Or it may produce values outside of the range that has been observed in a region.

For this study, we altered how OI is applied compared to previous applications, by recalculating the post-filter matrix in every river basin and at every time step. The OI algorithm requires an a priori estimate of the error covariance matrix for our input variables, the hydrologic fluxes estimated by remote sensing. In practice, this information is rarely available, and therefore uncertainties are estimated by expert judgment or by computational experiments. Previous applications of OI assumed constant values for uncertainties, regardless of the season or the location. Such an assumption is defensible when analyzing a single river basin (the Mississippi, in Munier et al., 2014), a single region, (Southeast Asia, in Pellet, Aires, Papa, et al., 2019), or the analysis is restricted to very large basins (Munier & Aires, 2018). However, we aimed for global coverage, and our river basins cover a wide range of climates and hydrologic conditions, from highly arid to tropical rainforest. We estimated the uncertainty for each estimated flux as the minimum of 6 mm/mo or 20% of the absolute value of the flux.

3.3 Neural Network Model

The OI method works well at the basin scale but require all HC components to be present. To improve the accuracy of applying OI findings to new locations, we aim to use a more complex model that includes additional inputs to describe the environment. We attempt to achieve this by utilizing environmental data as input variables in a flexible neural network (NN) model.

We chose a particular type of NN, a multi-layered perceptron (Rumelhart et al., 1987). The neurons are organized in successive layers, each neuron first performs a weighted average of their inputs using synaptic weights. A non-linear sigmoid function g such as a *tanh* or *tansig* function is then applied on the weighted average. The final output of a neuron i is then given by: $y_i = g\left(\sum_{j=1}^{N} w_{ji} x_j\right)$, where $(x_j; j = 1, \cdots, N)$ are the N inputs of neuron i, and w_{ji} is the synaptic weight between neuron j and i (Bishop, 1996).

More generally, a NN is a flexible model that can simulate complex nonlinear relationships. Given the correct model form and proper training, it can fit any arbitrary function. Often, classical NN architecture is fully connected, meaning that every neuron has a connection with all the neurons of the previous layer. This is not the case here, where we are operating multiple independent NNs for calibration and mixture. We experimented with a number of NN architectures. While the one shown in Figure 4 is among the simpler models that we tried, it performed the best. On the left are the model inputs, the uncorrected EO datasets, and on the right are the targets, the solution from OI that results in a balanced water budget. We chose a modular architecture with separate calibration and mixture steps that allows us to investigate the outputs of individual layers as we may gain useful information from each:

- First, a set of NNs serves to *calibrate* the individual inputs, or to transform them such that they more closely match the OI solution that satisfies the water balance constraint. For example, the output of the first calibration sub-model in Figure 4, $P_{1,cal}$, is a function of P_1 and the ancillary variables. In this way, each EO product can be optimized independently to each other. This allows running the NN in various configurations with different numbers of input variables (e.g., when one input variable is missing).

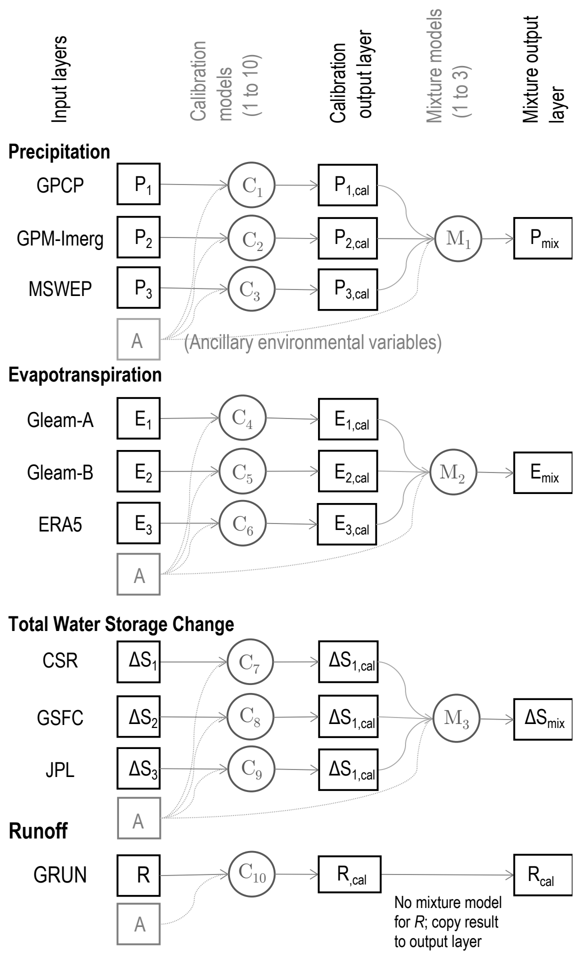


Figure 4. Neural network model architecture for calibration then mixture of EO datasets.

- Next, the *mixture* NNs combine information output by the calibration layer to estimate P, E, ΔS, and R. The NN seeks the best compromise among the calibrated EO datasets to fit the target, the OI solution.

A database with paired input and target data is required to train and test the NN model, as well as to select the best model architecture and find the best set of model parameters. For the set of NNs shown in Figure 4, each of the 10 calibration networks has 13 inputs (1 EO variable and 12 ancillary environmental variables), 10 neurons in the hidden layer, and 1 neuron in the output layer. The outputs of the calibration layer are calibrated EO datasets, which are useful in their own right, as they should better balance the water budget. Further, they are inputs to the mixture model layers. These layers also have 10 neurons in the hidden layer and 1 neuron in the output layer. For example, the inputs to the precipitation mixture model are calibrated P from each of the three calibration models plus the ancillary variables. Again the target is the OI solution for P calculated previously. In the following section, we evaluate the results of the 10 calibration NNs (1 calibration per EO dataset), and the output of 4 mixture NNs (1 mixture per HC component).

The number of neurons in the hidden layers and the number of hidden layers controls the complexity of the model. We experimented with a range of network sizes and configurations, and found that the fit does not improve with more neurons. Estimation of the optimal parameters of the NN was performed during the training stage using the back-propagation Levenberg-Marquardt algorithm (Rumelhart et al., 1987). We trained the model on a set of 1,358 basins and validated the model over a set of 340 independent basins (for an 80/20 split between training and validation). We corrected any physically implausible negative values for P or R by setting them zero. Finally, outputs for ΔS and R were smoothed with a 3-month moving mean filter to remove high-frequency noise from the predictions. We also performed the equivalent smoothing on validation datasets in order to ensure a fair comparison.

4 Results

Here, we evaluate the results of our optimization procedure for EO data using OI and NN modeling. The best model will be one that reconciles the inputs and results in a lower water budget residual, I. It should yield results that are plausible while changing the inputs as little as necessary.

Figure 5 is an example showing the inputs and outputs of our method over one river basin. The data is for the White River at Petersburg, Indiana, United States, with a drainage area of 29,000 km^2. While no river basin is typical, this location does a good job demonstrating the output from our calculations as it has a long record of river discharge. The corrections made in this basin are relatively modest; over this region of the eastern United States, remote sensing datasets tend to be more reliable and well-calibrated due to the density and availability of in situ calibration data.

The time series plots in Figure 5 show the inputs (EO datasets, in gray), the outputs of OI (green) and the outputs of the mixture NN (purple). There is significant disagreement among the 3 P datasets as their seasonality differs. E for this location is more consistent. The three GRACE datasets for TWSC or ΔS are highly correlated with one another, as expected since they are derived from the same satellite data. The bottom plot shows the HC residual or *imbalance*, I. The gray lines show each of the 27 possible combinations of the datasets ($3P \times 3E \times 3\Delta S \times 1R$). The imbalances based on uncorrected EO data are significant: the seasonal I can reach ± 50 mm/mo depending on the combination of datasets. The objective of our integration technique is to reduce this imbalance as much as possible. I based on the OI solution (in blue) is equal to zero by definition.

The NN optimization of P, E and R results in a significant improvement in I. One of the key features of our model is that it should make minimal modifications to the inputs while moving closer to a solution that balances the water budget. In particular, we note that discharge R is changed less by the NN optimization than it is by OI. This means that the NN optimization acts mostly on P and E towards a better coherency with R and ΔS.

4.1 Evaluation of Water Budget Closure

Figure 6 shows the distribution of the HC imbalance in the 340 validation basins. The empirical PDFs are kernel density plots showing the mean (left) and the standard deviation (right) of I in each basin. The gray lines show the imbalance calculated from the original uncorrected EO datasets (27 cases). The OI solution is not shown, as $I =$ 0. We again calculated I using each of the 27 combinations of datasets output by the *calibration* NNs (shown in pink), and the I resulting from the mean for each component (in red). Finally, the blue line shows the result of the final NN mixture model. Each step in the optimization process reduces both the bias and the variance of I. The mean and standard deviation of the I with uncorrected EO data is 11 ± 44 mm/mo. Simply averaging multiple datasets significantly improves the water balance. A great deal more improvement comes from the NN calibration models ($I = 0.12 \pm 27$ mm/mo). The NN

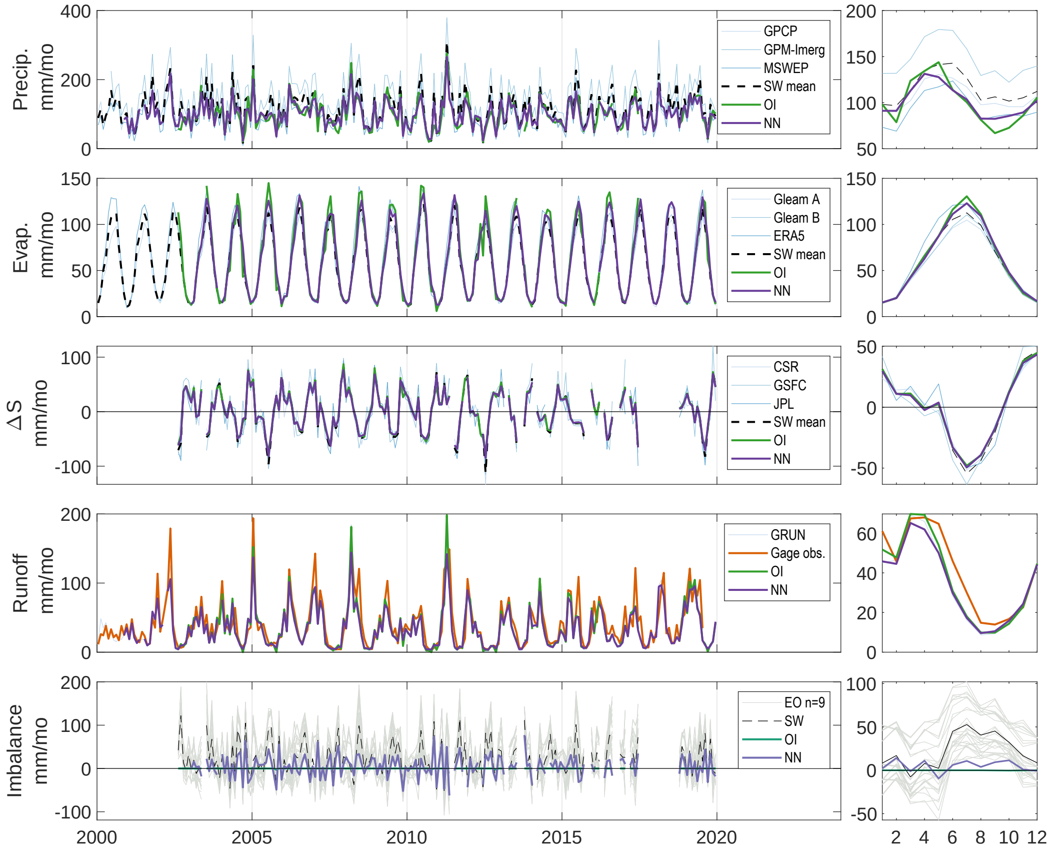


Figure 5. Time series plots of EO data over the White River basin in Indiana, US (GRDC gage 4123202) on left. Datasets are: observed (light colors), the simple-weighted mean of observations (SW, dashed black), OI solution (green), and estimated by the NN model (purple). At right is the corresponding seasonality (monthly averages).

mixture model has a slight positive impact ($I = -0.03 \pm 24$ mm/mo). It appears therefore that most of the improvement comes from the initial calibration layer with an additional but minor improvement from the mixture layer.

We next applied the trained NN model at the pixel scale, making monthly predictions of P, E, ΔS and R in 0.5° grid cells over land. We then calculated imbalance, I, in every pixel. Figure 7(a) shows I_{MIX}, the long-term average imbalance based on the output of the NN mixture model. We visualize how much the NN has improved the imbalance at the pixel scale in Figure 7(b), where we have calculated an "improvement factor," comparing I_{MIX} to I_{SW}, the imbalance based on the SW mean of EO datasets. The improvement factor is a convergence metric that measures how much closer I is to zero after optimization, and is calculated as $|I_{SW}| - |I_{MIX}|$. A positive value indicates that the imbalance is closer to zero (our desired result), while a negative value means that the imbalance is further from zero (negative result). The NN model results in a lower water budget residual in nearly all locations, with particularly large improvements over parts of the Amazon and southeast Asia. The imbalance is made worse in a few locations, notably near the extreme western coasts of Canada, Chile, England, and Norway. These more difficult locations can be related to coastal contamination on the EO data, elevation, and ice presence. Furthermore, our model may not adequately capture the dynamics in high mountain regions; such environments are not well sampled in our dataset as we set a minimum threshold for the basin area.

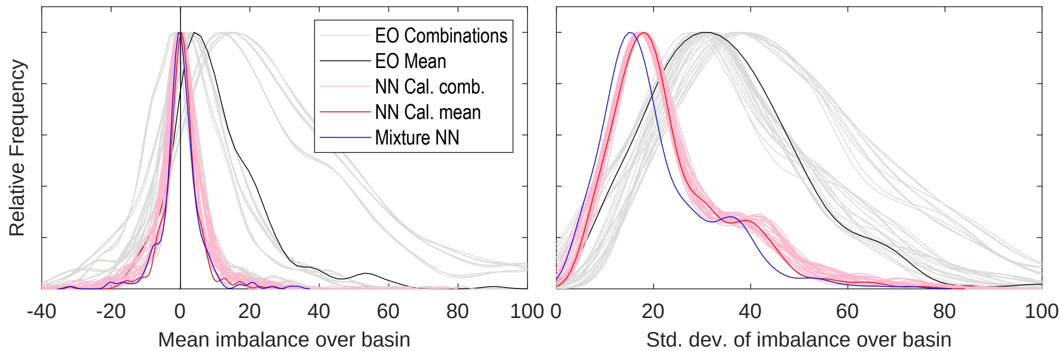


Figure 6. Empirical probability distribution plots of the HC imbalance, showing the mean and standard deviation of the imbalance over the 340 validation basins.

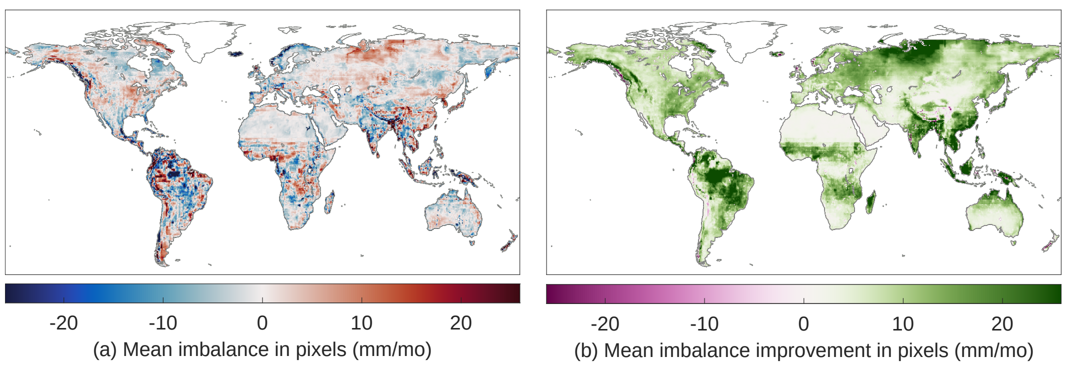


Figure 7. Map of the average HC imbalance in 0.5° pixels over the years 2000 - 2019: (a) the imbalance calculated by fluxes calibrated by the NN mixture model, and (b) the average improvement from EO observations.

4.2 Evaluation of the calibration EO data

As an additional assessment of our optimization, we compared the output of our NN model to observations where available, seeking to answer the following questions: Are we improving the fit to observations, or moving further away from them? Are we able to improve EO data more in certain locations or under certain conditions?

For this analysis, we first compared EO estimates of E to observed E at 117 global flux towers. Then we compared the outputs of the calibration and mixture NN models to these same observations. We repeated the same procedure for R, comparing NN predictions to discharge measured at gages. We calculated fit statistics comparing the observed and predicted time series at each flux tower or gage. Table 2 reports the median of the fit statistic. For example, we calculated 117 values for the correlation coefficient, R. For Gleam-A, the first row in the table, these values ranged from -0.11 to 0.98, with a median of 0.91. The models denoted with *cal.* have undergone calibration using the NN model. Entries in bold text highlight the best value of each indicator within its class.

For E, the NN models generally improved the fit to observations collected at flux towers. The improvements are not very big, and may not be important considering the caveats related to comparing point estimates to grid cell values. Nevertheless, it is a positive sign that our model does not degrade the signal, and in fact may be improving it.

Table 2. Validation of the NN model predictions for E and R, showing the impact of the NN calibration and mixture model on the goodness of fit to observations.

Dataset	Corr. R	RMSE, mm/mo
Evapotranspiration, at 117 flux towers		
GLEAM-A	0.91	21.4
Gleam-A cal.*	0.92	19.0
Gleam-B	**0.93**	20.1
Gleam-B cal.	0.92	**18.5**
ERA5	0.91	19.9
ERA5 cal.	0.91	19.4
Mixture NN	0.92	19.4
Runoff, at 1,781 gages		
GRUN	**0.90**	**9.26**
GRUN cal.	0.89	9.34

* cal. = calibrated by NN model

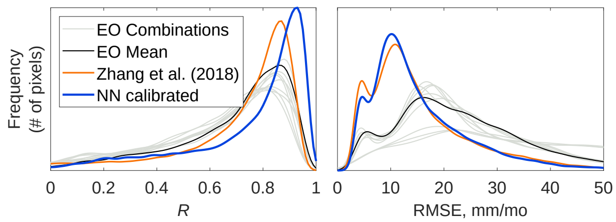
The situation with discharge is largely reversed, and it appears that NN calibration is degrading the signal somewhat, albeit only slightly. Here, we calculated fit statistics against a set of gages with a strong runoff signal (we excluded gages in arid regions where runoff is often at or near zero, leaving 1,781 gages). The changes made to runoff data, and fits to observations are not evenly distributed. Based on the change in RMSE, there is an improved fit to observations in 47% of basins, and a slight degradation in the fit in 53% of basins. Maps of the changes in each fit indicator (not shown in this paper) reveal that the most improvement occurs in arid regions, while the worst degradation occurs at gages north of 70° latitude, in the Arctic regions of North America, Europe, and Asia.

5 Reconstruction of Total Water Storage Change

An advantage to the NN architecture described in Figure 4 is that it is modular. Each step (calibration, mixture) results in an improvement to EO datasets, in terms of producing a balanced water budget, as seen in Figure 6. This is very valuable when faced with missing data: A missing HC component can be estimated by inference from the other three. Indeed, several studies have exploited this relationship (see e.g., Rodell et al., 2011; Munier et al., 2014; Liu et al., 2020; Pellet, Aires, Papa, et al., 2019; Lehmann et al., 2022).

We used this approach for indirect estimation of ΔS. This allows us to estimate GRACE-like TWSC for the time period before 2002 when the satellites were launched, or to fill in missing data. If we assume a balanced water budget, rearranging Equation 1 gives $\Delta S = P - E - R$. Estimating missing components using indirect observations should be improved when using the optimized water components of the previous section. This is therefore an indirect evaluation of the water budget obtained by our integration framework.

Overall, we obtained a significantly improved fit to GRACE observations with ΔS obtained from the three other NN-calibrated fluxes, compared to similar estimation with uncorrected EO data. At the pixel scale, our new NN-inferred ΔS compare favorably to those predicted by Zhang et al. (2018). Figure 8(b) shows the empirical probability distribution function (PDF) for two fit indicators over land pixels. While reconstructing TWSC was not the main goal of this study, this experiment shows the improved agree-



Indirect estimator of ΔS	Median Corr., R	Median RMSE, mm/mo
EO combinations (n=9)	0.71 ± 0.17	24 ± 12
EO Mean	0.75	22
Zhang et al. (2018)	0.79	13
NN calibrated (this study)	0.84	13

Figure 8. Empirical PDF of the correlation (left) and RMSE (right) between GRACE observations and indirect estimates for ΔS over 57,286 land pixels.

ment of the water components, which should be beneficial for future applications. The fact that our NN model performs well under most conditions is encouraging.

Figure 9 shows a reconstruction of GRACE-like monthly TWSC over 3 river basins of varying size. Here, it is estimated indirectly from the other three components of the water cycle, $\Delta S_{est.} = P - E - R$. The gray lines show ΔS estimated by uncorrected EO datasets. After 2000, there are 9 different combinations shown ($3P\times3E\times3\times1R$). Before 1980, there are fewer combinations, as some datasets have limited temporal coverage (see Table 1). The green line shows TWSC from GRACE, where available (average of the three solutions in Table 1). The orange line is our reconstruction of ΔS. Finally, the dashed purple line is ΔS from the study by Zhang et al. (2018). Over the selected basins, the reconstructed time series of TWSC do a good job recreating the seasonal patterns observed by GRACE over river basins of a range of sizes. Further, both reconstructed time series of TWSC are a significantly better fit to observations compared to estimates based on uncorrected EO data. As shown in 8, the reconstruction based on this study's NN is a slightly better fit to observations compared to the results from Zhang et al. (2018). This study's indirect estimates of TWSC are able to cover a longer time period; the modular nature of the calibration NN model allows us to use whichever dataset(s) are available in a given time period for estimation. In general, estimates are more robust when more datasets are available. As fewer datasets are available from 1980 to 2000, this is an additional source of uncertainty for hindcast estimates of TWSC.

There are also other limitations to the reconstructed datasets of TWSC. It can be shown that even a very small bias makes it impossible to calculate the trend in TWS with any degree of accuracy. We are computing TWSC from climate data only, while it has been shown that human activities like groundwater pumping and the filling and draining of reservoirs have a major impact on TWS (Rodell et al., 2018).

6 Conclusions

We explored novel methods of analyzing and combining earth observation datasets describing major hydrologic fluxes, with the goal of reducing the overall error in estimating the water budget. We applied a closed-form analytical solution, optimal interpolation, which forces the water budget residual to zero. This approach has several advan-

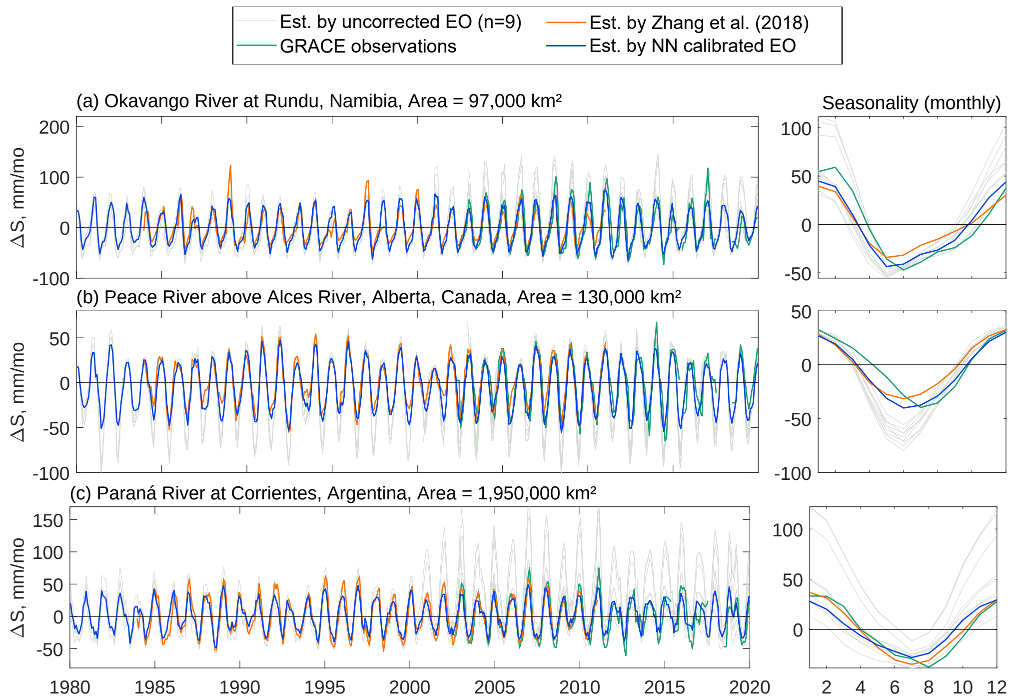


Figure 9. GRACE-like TWSC reconstructed by indirect estimation over three river basins

tages – it is simple to implement and has a basis in theory and existing practice, as it seeks to allocate errors in observations in inverse proportion to their uncertainty. Nevertheless, this approach has limitations that prevent us from applying it globally. Most importantly, OI requires observations of river discharge (only available on a few gaged river basins) and change in water storage (only available via the GRACE satellites in operation since 2002).

Previous research in this area has demonstrated the utility of the OI approach. In this paper, we expand upon previous work in two important ways. First, we applied the method at a larger scale, optimizing observed fluxes in over 1,654 river basins on every continent except Greenland and Antarctica. Second, we demonstrated the ability of a neural network model to reproduce the results of OI with reasonable accuracy over relatively large river basins ($> 2,500$ km^2). The model fit varies by location; it tends to be better over humid regions, and less accurate over the Arctic or over parts of Asia and South America. The NN model can be used over river basins nearly anywhere on the globe, globally and at the pixel scale. We showed that calibrating EO data with our NN at the pixel scale results in improved coherency among datasets and a lower HC residual over most continental land surfaces.

Our set of NN is modular, with separate models for calibration of individual datasets, and for mixture of different datasets of the same water component. This allows us to make estimations in the absence of one or more of the four main fluxes in the hydrologic cycle. We validated our NN model by comparing the output against in situ observations and found that the calibration generally improves the fit to E measured at flux towers, and does not seriously degrade the fit to observed river discharge. We tested the ability of the NN model to estimate missing HC components by inference. Estimates based on NN calibrated fluxes are a major improvement over uncorrected EO data. Nevertheless, estimating TWS indirectly via the three other HC components is not accurate enough for trend detection or for hindcasting TWS anamolies in the decades before the launch of the GRACE satellites.

The NN framework introduced by Aires (2014) and expanded upon in this paper opens new doors for the integration of satellite data to study the HC. The NN model we developed is original in the field of water budget closure studies, and has some special features that allow us to integrate satellite observations. Our model is nested, featuring independent calibration and mixture models to stay closer to the physical treatment that we intend to produce. Our approach optimizes EO datasets and closes the HC without the use of a simulation model. Rather, our data-driven approach can be set up to rely only on data from satellite returns. This makes it valuable for the calibration and validation of climate models and hydrologic models, among other applications.

Future research in this area could experiment with using different NN architectures. The fit of the NN may also be improved by providing more input data. Our hypothesis is that providing the model with more information about the hydrologic conditions allows it to customize parameters for different climate zones, plant communities, and hydrologic conditions. Our results confirm that ancillary environmental data improves the fit of the model, although the improvement is modest. Further research may find a combination of environmental data and model configuration that helps the model differentiate zones with a different hydrologic response, such as deserts or tropical rainforests.

Open Research

Availability Statement - The input datasets used in this analysis can be freely obtained via the sources listed in Section 2, Datasets. The compiled data and scripts (in Matlab format) needed to perform the analysis described in this paper can be downloaded from: `https://doi.org/10.5281/zenodo.8101659`

Acknowledgments

This research was supported by ESTELLUS and by a contract with the European Space Agency (contract #4000136793/21/I-DT-lr). We would like to thank Espen Volden for accepting/managing this project at ESA.

CRediT author statement

M. Heberger: Software, Formal analysis, Writing - Original Draft, Writing - Review & Editing. **F. Aires**: Conceptualization, Supervision, Project administration, Funding acquisition, Writing - Review & Editing. **V. Pellet**: Conceptualization, Writing - Review & Editing

References

Adler, R. F., Sapiano, M. R. P., Huffman, G. J., Wang, J.-J., Gu, G., Bolvin, D., ... Shin, D.-B. (2018, April). The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. *Atmosphere*, *9*(4), 138. doi: 10.3390/atmos9040138

Aires, F. (2014, April). Combining datasets of satellite-retrieved products. Part I: Methodology and water budget closure. *Journal of Hydrometeorology*, *15*(4), 1677–1691. Retrieved 2020-03-19, from `https://journals.ametsoc.org/doi/10.1175/JHM-D-13-0148.1` doi: 10.1175/JHM-D-13-0148.1

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., & Jetz, W. (2018, March). A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data*, *5*(1), 180040. Retrieved 2021-11-16, from https://www.nature.com/articles/sdata201840 doi: 10.1038/sdata.2018.40

Australia BOM. (2020). *Hydrologic Reference Stations update 2020.* Retrieved 2022-12-22, from http://www.bom.gov.au/water/hrs/update_2020.shtml

Bartos, M., Smith, T. J., Itati01, Debbout, R., Kraft, P., & Huard, D. (2023, May). *pysheds: 0.3.5.* Zenodo. Retrieved 2023-06-13, from https://zenodo.org/record/3822494 doi: 10.5281/ZENODO.3822494

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I., ... Adler, R. F. (2019). MSWEP V2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. doi: 10.1175/BAMS-D-17-0138.1

BfG. (2020). *BfG - The GRDC.* Retrieved 2022-12-22, from https://www.bafg.de/GRDC/EN/01_GRDC/grdc_node.html

Biancamaria, S., Mballo, M., Le Moigne, P., Sánchez Pérez, J. M., Espitalier-Noël, G., Grusson, Y., ... Sauvage, S. (2019, August). Total water storage variability from GRACE mission and hydrological models for a 50,000 km2 temperate watershed: the Garonne River basin (France). *Journal of Hydrology: Regional Studies*, *24*, 100609. Retrieved 2021-11-09, from https://www.sciencedirect.com/science/article/pii/S2214581818303574 doi: 10.1016/j.ejrh.2019.100609

Bishop, C. M. (1996). *Neural Networks for Pattern Recognition.* Oxford University Press.

Didan, K. (2015). *MOD13C2 MODIS/Terra Vegetation Indices Monthly L3 Global 0.05Deg CMG V006.* NASA EOSDIS Land Processes DAAC. Retrieved 2023-02-27, from https://lpdaac.usgs.gov/products/mod13c2v006/ doi: 10.5067/MODIS/MOD13C2.006

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018, April). The Global Streamflow Indices and Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata. *Earth System Science Data*, *10*(2), 765–785. Retrieved 2021-06-29, from https://essd.copernicus.org/articles/10/765/2018/ doi: 10.5194/essd-10-765-2018

Famiglietti, J. S., Bijoor, N., Reager, J. T., & Lo, M. (2011). *Using GRACE and isotopes for Hydrology: Combining techniques for improved observation.* Monaco.

Ghiggi, G., Humphrey, V., Seneviratne, S. I., & Gudmundsson, L. (2019, November). GRUN: an observation-based global gridded runoff dataset from 1902 to 2014. *Earth System Science Data*, *11*(4), 1655–1674. Retrieved 2021-06-29, from https://essd.copernicus.org/articles/11/1655/2019/ doi: 10.5194/essd-11-1655-2019

Ghiggi, G., Humphrey, V., Seneviratne, S. I., & Gudmundsson, L. (2021). G-RUN ENSEMBLE: A Multi-Forcing Observation-Based Global Runoff Reanalysis. *Water Resources Research*, *57*(5). Retrieved 2021-06-18, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020WR028787 doi: 10.1029/2020WR028787

Giglio, L., Schroeder, W., & Hall, J. V. (2020). *MODIS Collection 6 Active Fire Product User's Guide Revision C.*

Gudmundsson, L., Do, H. X., Leonard, M., & Westra, S. (2018, April). The Global Streamflow Indices and Metadata Archive (GSIM) – Part 2: Quality control, time-series indices and homogeneity assessment. *Earth System Science Data*, *10*(2), 787–804. Retrieved 2021-06-29, from https://essd.copernicus.org/articles/10/787/2018/ doi: 10.5194/essd-10-787-2018

Guillory, A. (2022, May). *ERA5* [Text]. Retrieved 2023-03-10, from https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5

Hall, D., & Riggs, G. (2021). *MODIS/Terra Snow Cover Monthly L3 Global 0.05Deg*

CMG, version 6.1. NASA National Snow and Ice Data Center Distributed Active Archive Center. Retrieved from https://doi.org/10.5067/MODIS/MOD10CM.061 doi: 10.5067/MODIS/MOD10CM.061

Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., ... Pappenberger, F. (2020, September). GloFAS-ERA5 operational global river discharge reanalysis 1979–present. *Earth System Science Data*, *12*(3), 2043–2060. Retrieved 2023-06-05, from https://essd.copernicus.org/articles/12/2043/2020/ doi: 10.5194/essd-12-2043-2020

Heberger, M. (2022, November). *mheberger/delineator: 1.0.* Zenodo. Retrieved 2022-11-11, from https://zenodo.org/record/7314287 doi: 10.5281/ZENODO.7314287

Hegerl, G. C., Black, E., Allan, R. P., Ingram, W. J., Polson, D., Trenberth, K. E., ... Zhang, X. (2015, July). Challenges in Quantifying Changes in the Global Water Cycle. *Bulletin of the American Meteorological Society*, *96*(7), 1097–1115. Retrieved 2023-03-01, from https://journals.ametsoc.org/view/journals/bams/96/7/bams-d-13-00212.1.xml doi: 10.1175/BAMS-D-13-00212.1

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., ... Thépaut, J-N.\ (2018). *Copernicus Climate Data Store |.* Retrieved 2021-04-21, from https://cds.climate.copernicus.eu/cdsapp doi: 10.24381/cds.adbb2d47

Hogan, R. (2015). *Radiation Quantities in the ECMWF model and MARS.* ECMWF. Retrieved from https://www.ecmwf.int/sites/default/files/elibrary/2015/18490-radiation-quantities-ecmwf-model-and-mars.pdf

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., ... Xie, P. (2020). *Algorithm Theoretical Basis Document (ATBD) Version 06: NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG)* (Tech. Rep.). Greenbelt, MD: National Aeronautics and Space Administration. Retrieved from https://gpm.nasa.gov/sites/default/files/2020-05/IMERG_ATBD_V06.3.pdf

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019, October). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. Retrieved 2020-09-28, from https://hess.copernicus.org/articles/23/4323/2019/ doi: 10.5194/hess-23-4323-2019

Landerer, F. W. (2021). *CSR TELLUS GRACE Level-3 Monthly Land Water-Equivalent-Thickness Surface Mass Anomaly Release 6.0 version 04 in netCDF/ASCII/GeoTIFF Formats.* NASA Physical Oceanography DAAC. Retrieved 2022-12-22, from https://podaac.jpl.nasa.gov/dataset/TELLUS_GRAC_L3_CSR_RL06_LND_v04 doi: 10.5067/TELND-3AC64

Landerer, F. W., & Cooley, S. S. (2021). *Gravity Recovery and Climate Experiment Follow-on (GRACE-FO) Level-3 Data Product User Handbook.* Jet Propulsion Laboratory, California Institute of Technology.

Landerer, F. W., Dickey, J. O., & Güntner, A. (2010). Terrestrial water budget of the Eurasian pan-Arctic from GRACE satellite measurements during 2003–2009. *Journal of Geophysical Research: Atmospheres*, *115*(D23). Retrieved 2021-11-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2010JD014584 doi: 10.1029/2010JD014584

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, *48*(4). Retrieved 2021-04-06, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR011453 doi: 10.1029/2011wr011453

Lehmann, F., Vishwakarma, B. D., & Bamber, J. (2022). How well are we able to close the water budget at the global scale? *Hydrology and Earth System Sciences*, *26*(1). Retrieved 2022-03-02, from https://hess.copernicus.org/

articles/26/35/2022/ doi: doi.org/10.5194/hess-26-35-2022

Lehner, B., Verdin, K., & Jarvis, A. (2008). New Global Hydrography Derived from Spaceborne Elevation Data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94. Retrieved 2022-11-21, from http://onlinelibrary.wiley.com/doi/abs/10.1029/2008EO100001 doi: 10.1029/2008EO100001

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., ... Wood, E. F. (2019, August). Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. *Water Resources Research*, *55*(8), 6499–6516. Retrieved 2020-02-22, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019WR025287 doi: 10.1029/2019WR025287

Lin, P., Pan, M., Wood, E. F., Yamazaki, D., & Allen, G. H. (2021, January). A new vector-based global river network dataset accounting for variable drainage density. *Scientific Data*, *8*(1), 28. Retrieved 2021-07-26, from https://www.nature.com/articles/s41597-021-00819-9 doi: 10.1038/s41597-021-00819-9

Liu, Y., Wagener, T., Beck, H. E., & Hartmann, A. (2020, September). What is the hydrologically effective area of a catchment? *Environmental Research Letters*, *15*(10), 104024. Retrieved from https://iopscience.iop.org/article/10.1088/1748-9326/aba7e5 doi: 10.1088/1748-9326/aba7e5

Loomis, B. D., Luthcke, S. B., & Sabaka, T. J. (2019, September). Regularization and error characterization of GRACE mascons. *Journal of Geodesy*, *93*(9), 1381–1398. Retrieved 2021-11-29, from http://link.springer.com/10.1007/s00190-019-01252-y doi: 10.1007/s00190-019-01252-y

Lu, J., Wang, G., Chen, T., Li, S., Hagan, D. F. T., Kattel, G., ... Su, B. (2021, December). A harmonized global land evaporation dataset from model-based products covering 1980–2017. *Earth System Science Data*, *13*(12), 5879–5898. Retrieved 2023-02-24, from https://essd.copernicus.org/articles/13/5879/2021/ doi: 10.5194/essd-13-5879-2021

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A., Fernández-Prieto, D., ... Verhoest, N. E. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. Retrieved from https://gmd.copernicus.org/articles/10/1903/2017/gmd-10-1903-2017.pdf doi: 10.5194/gmd-10-1903-2017

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., ... Wood, E. F. (2017). The future of Earth observation in hydrology. *Hydrology and Earth System Sciences*, *21*(7), 3879–3914. Retrieved 2021-03-22, from https://hess.copernicus.org/articles/21/3879/2017/ doi: 10.5194/hess-21-3879-2017

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. Retrieved from https://hess.copernicus.org/articles/15/453/2011/ doi: 10.5194/hess-15-453-2011

Munier, S., & Aires, F. (2018, February). A new global method of satellite dataset merging and quality characterization constrained by the terrestrial water budget. *Remote Sensing of Environment*, *205*, 119–130. Retrieved 2020-03-19, from http://www.sciencedirect.com/science/article/pii/S0034425717305321 doi: 10.1016/j.rse.2017.11.008

Munier, S., Aires, F., Schlaffer, S., Prigent, C., Papa, F., Maisongrande, P., & Pan, M. (2014). Combining datasets of satellite-retrieved products for basin-scale water balance study: 2. Evaluation on the Mississippi Basin and closure correction model. *Journal of Geophysical Research: Atmospheres*, *119*(21), 12,100–12,116. Retrieved 2020-03-19, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD021953 doi:

10.1002/2014JD021953

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., … Zhang, L. (2020, July). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 225. Retrieved 2023-05-04, from https://www.nature.com/articles/s41597-020-0534-3 doi: 10.1038/s41597-020-0534-3

Pellet, V., Aires, F., Mariotti, A., & Fernández-Prieto, D. (2018, November). Analyzing the Mediterranean Water Cycle Via Satellite Data Integration. *Pure and Applied Geophysics*, *175*(11), 3909–3937. Retrieved 2020-03-19, from https://doi.org/10.1007/s00024-018-1912-z doi: 10.1007/s00024-018-1912-z

Pellet, V., Aires, F., Munier, S., Fernández Prieto, D., Jordá, G., Dorigo, W. A., … Brocca, L. (2019, January). Integrating multiple satellite observations into a coherent dataset to monitor the full water cycle–application to the Mediterranean region. *Hydrology and Earth System Sciences*, *23*(1), 465–491. Retrieved 2020-03-19, from https://www.hydrol-earth-syst-sci.net/23/465/2019/ doi: 10.5194/hess-23-465-2019

Pellet, V., Aires, F., Papa, F., Munier, S., & Decharme, B. (2019, September). Long-term Total Water Storage Change from a Satellite Water Cycle Reconstruction over large south Asian basins. *Hydrology and Earth System Sciences Discussions*, 1–30. Retrieved 2020-03-19, from https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-262/ doi: 10.5194/hess-2019-262

Pellet, V., Aires, F., Yamazaki, D., & Papa, F. (2021). Coherent Satellite Monitoring of the Water Cycle Over the Amazon. Part 1: Methodology and Initial Evaluation. *Water Resources Research*, *57*(5), 21. doi: 10.1029/2020WR028647

Richey, A. S., Thomas, B. F., Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., … Rodell, M. (2015). Quantifying renewable groundwater stress with GRACE. *Water Resources Research*, *51*(7), 5217–5238. doi: 10.1002/2015WR017349

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., … Chambers, D. (2015). The observed state of the water cycle in the early twenty-first century. *Journal of Climate*, *28*(21), 8289–8318. Retrieved from https://journals.ametsoc.org/view/journals/clim/28/21/jcli-d-14-00555.1.xml doi: 10.1175/JCLI-D-14-00555.1

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M.-H. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(7707), 651–659. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6077847/ doi: 10.1038/s41586-018-0123-1

Rodell, M., McWilliams, E. B., Famiglietti, J. S., Beaudoing, H. K., & Nigro, J. (2011). Estimating evapotranspiration using an observation based terrestrial water budget. *Hydrological Processes*, *25*(26), 4082–4092. doi: 10.1002/hyp.8369

Rodgers, C. D. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practice* (Vol. 2). World Scientific. Retrieved 2022-03-15, from https://www.worldscientific.com/worldscibooks/10.1142/3171 doi: 10.1142/3171

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1987). Learning Internal Representations by Error Propagation. In D. E. Rumelhart & J. L. McLelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations* (pp. 318 – 362). Cambridge, Massachusetts: MIT Press. Retrieved from https://ieeexplore.ieee.org/servlet/opac?bknumber=6276825

Save, H. (2020). *CSR GRACE and GRACE-FO RL06 Mascon Solutions v02.* Retrieved from https://www2.csr.utexas.edu/grace/RL06_mascons.html doi: 10.15781/cgq9-nh24

Shiklomanov, I. A. (2009). World Water Balance. In *Hydrological Cycle, Volume 2* (p. 6). UNESCO - Encyclopedia Life Support Systems (UNESCO-EOLSS). Retrieved from https://books.google.fr/books?id=KAiCCwAAQBAJ

Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., & Scanlon, B. R. (2015, March). A global data set of the extent of irrigated land from 1900 to 2005. *Hydrology and Earth System Sciences*, *19*(3), 1521–1545. Retrieved 2022-12-05, from https://hess.copernicus.org/articles/19/1521/2015/ doi: 10.5194/hess-19-1521-2015

Singer, M. B., Asfaw, D. T., Rosolem, R., Cuthbert, M. O., Miralles, D. G., MacLeod, D., ... Michaelides, K. (2021, August). Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981-present. *Scientific Data*, *8*(1), 224. Retrieved 2021-12-08, from https://www.nature.com/articles/s41597-021-01003-9 doi: 10.1038/s41597-021-01003-9

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., & Watkins, M. M. (2004). GRACE measurements of mass variability in the Earth system. *Science*, *305*(5683), 503–505. doi: 10.1126/science.1099192

Tarek, M., Brissette, F., & Arsenault, R. (2020, September). Large-Scale Analysis of Global Gridded Precipitation and Temperature Datasets for Climate Change Impact Studies. *Journal of Hydrometeorology*, *21*, 1–54. doi: 10.1175/JHM-D-20-0100.1

Trabucco, A., & Zomer, R. (2019, January). *Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2.* CGIAR. Retrieved 2023-02-27, from https://figshare.com/articles/dataset/Global_Aridity_Index_and_Potential_Evapotranspiration_ET0_Climate_Database_v2/7504448/3 doi: 10.6084/m9.figshare.7504448.v3

Wan, Z., Hook, S., & Hulley, G. (2021). *MODIS/Terra Land Surface Temperature/Emissivity Monthly L3 Global 0.05Deg CMG V061.* NASA EOSDIS Land Processes DAAC. Retrieved 2023-05-22, from https://lpdaac.usgs.gov/products/mod11c3v061/ doi: 10.5067/MODIS/MOD11C3.061

WMO. (1989). *The Global Water Runoff Data Project: Workshop on the Global Runoff Data Set and Grid Estimation* (Tech. Rep.). Koblenz, Germany: World Meteorological Organization.

Wong, J. S., Zhang, X., Gharari, S., Shrestha, R. R., Wheater, H. S., & Famiglietti, J. S. (2021). Assessing Water Balance Closure Using Multiple Data Assimilation–and Remote Sensing–Based Datasets for Canada. *Journal of Hydrometeorology*, *22*(6), 1569–1589. Retrieved from https://journals.ametsoc.org/view/journals/hydr/22/6/JHM-D-20-0131.1.xml doi: 10.1175/JHM-D-20-0131.1

Yilmaz, M. T., DelSole, T., & Houser, P. R. (2011). Improving land data assimilation performance with a water budget constraint. *Journal of Hydrometeorology*, *12*(5), 1040–1055. doi: 10.1175/2011JHM1346.1

Zhang, Y., Pan, M., Sheffield, J., Siemann, A. L., Fisher, C. K., Liang, M., ... Wood, E. F. (2018, January). A Climate Data Record (CDR) for the global terrestrial water budget: 1984–2010. *Hydrology and Earth System Sciences*, *22*(1), 241–263. Retrieved 2020-09-29, from https://hess.copernicus.org/articles/22/241/2018/ doi: 10.5194/hess-22-241-2018

Zhang, Y., Pan, M., & Wood, E. F. (2016). On creating global gridded terrestrial water budget estimates from satellite remote sensing. In (pp. 59–78). Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-32449-4_4

Figure 1.

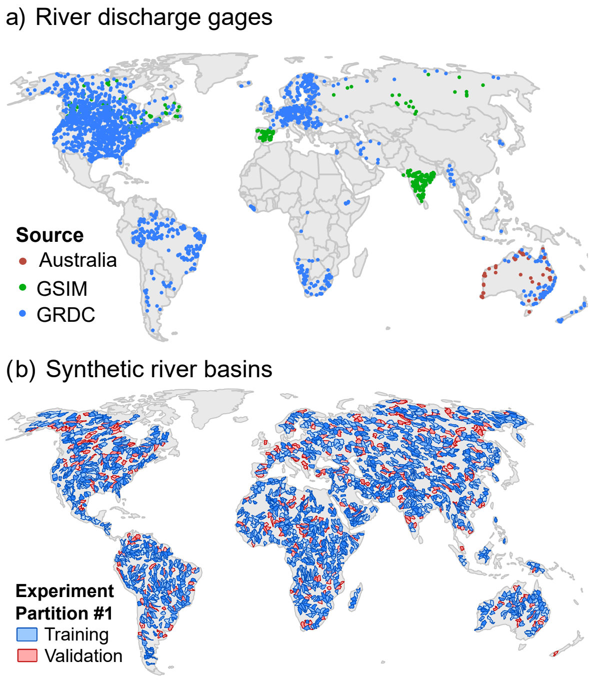
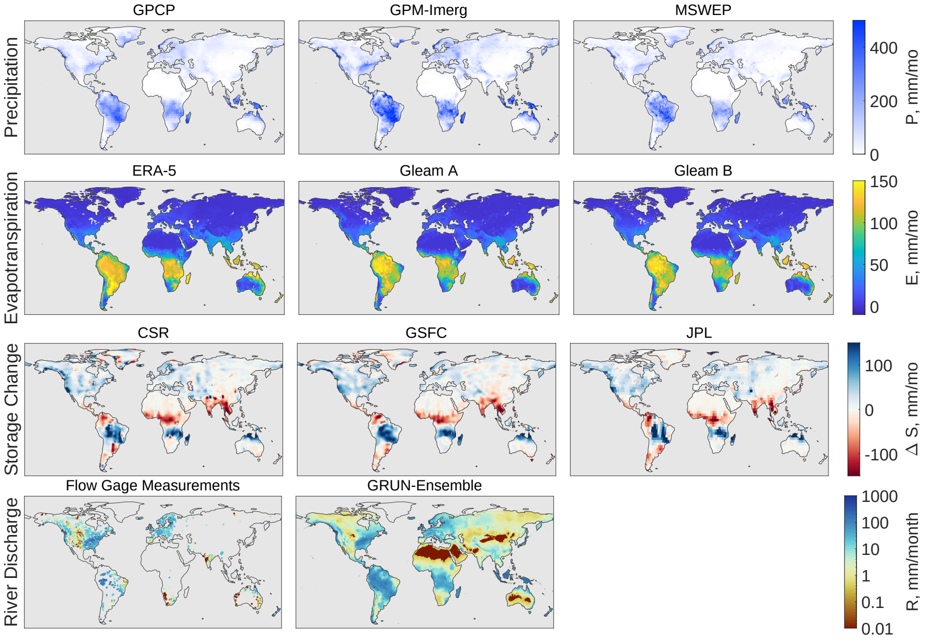
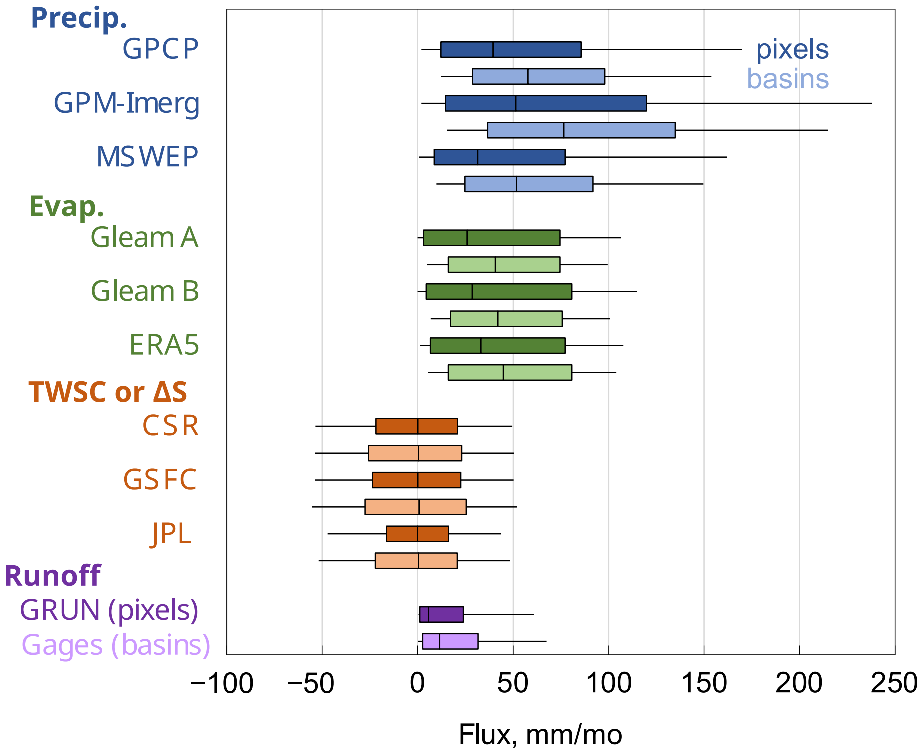

a) River discharge gages
Source
Australia
GSIM
GRDC
(b) Synthetic river basins
Experiment
Partition #1
Training
Validation

Figure 2.



GPCP
GPM-Imerg
MSWEP
Precipitation
400
200
0
P, mm/mo
ERA-5
Gleam A
Gleam B
Evapotranspiration
150
100
50
0
E, mm/mo
CSR
GSFC
JPL
Storage Change
100
0
-100
Δ S, mm/mo
Flow Gage Measurements
GRUN-Ensemble
River Discharge
1000
100
10
1
0.1
0.01
R, mm/month

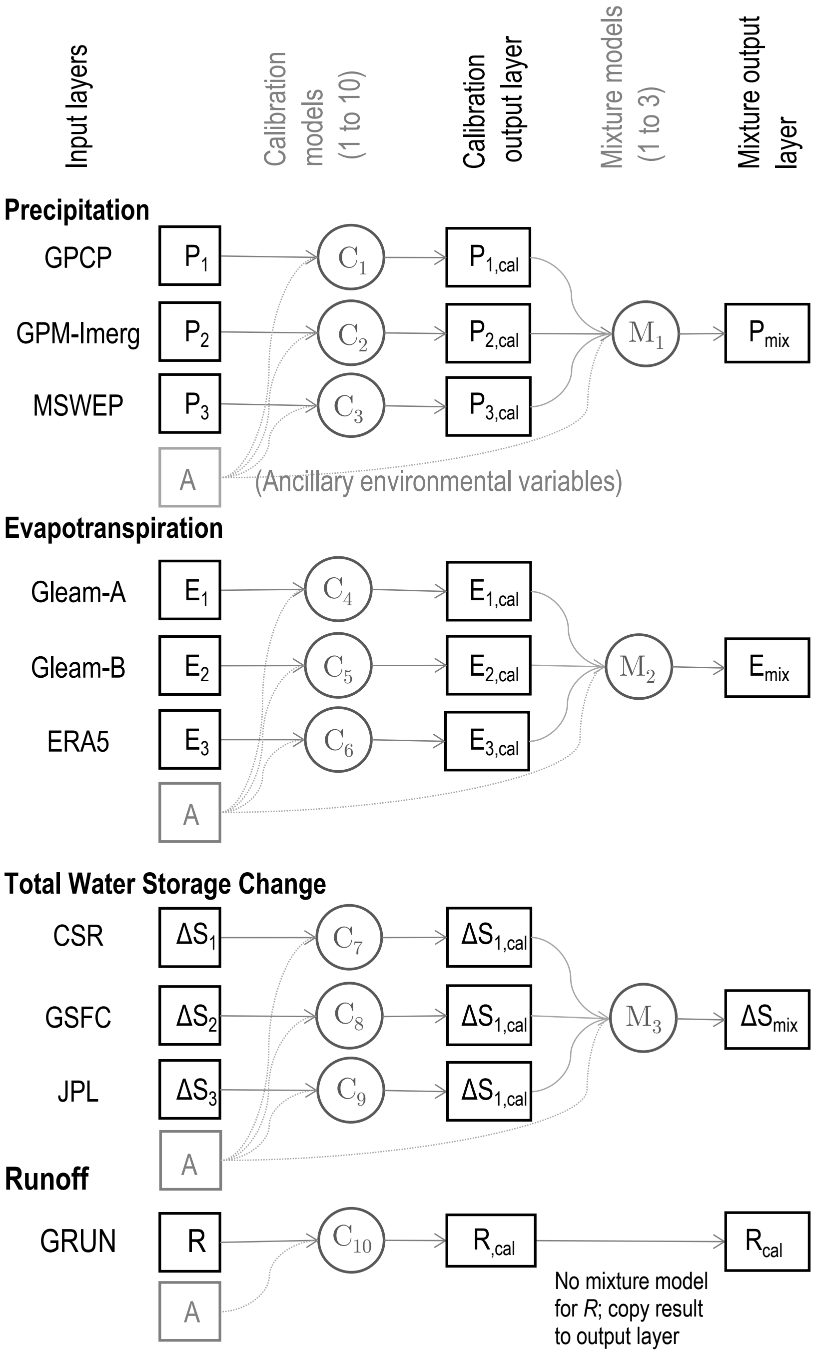
Figure 3.



Precip.
GPCP
GPM-Imerg
MSWEP
Evap.
Gleam A
Gleam B
ERA5
TWSC or ΔS
CSR
GSFC
JPL
Runoff
GRUN (pixels)
Gages (basins)
pixels
basins
−100
−50
0
50
100
150
200
250
Flux, mm/mo

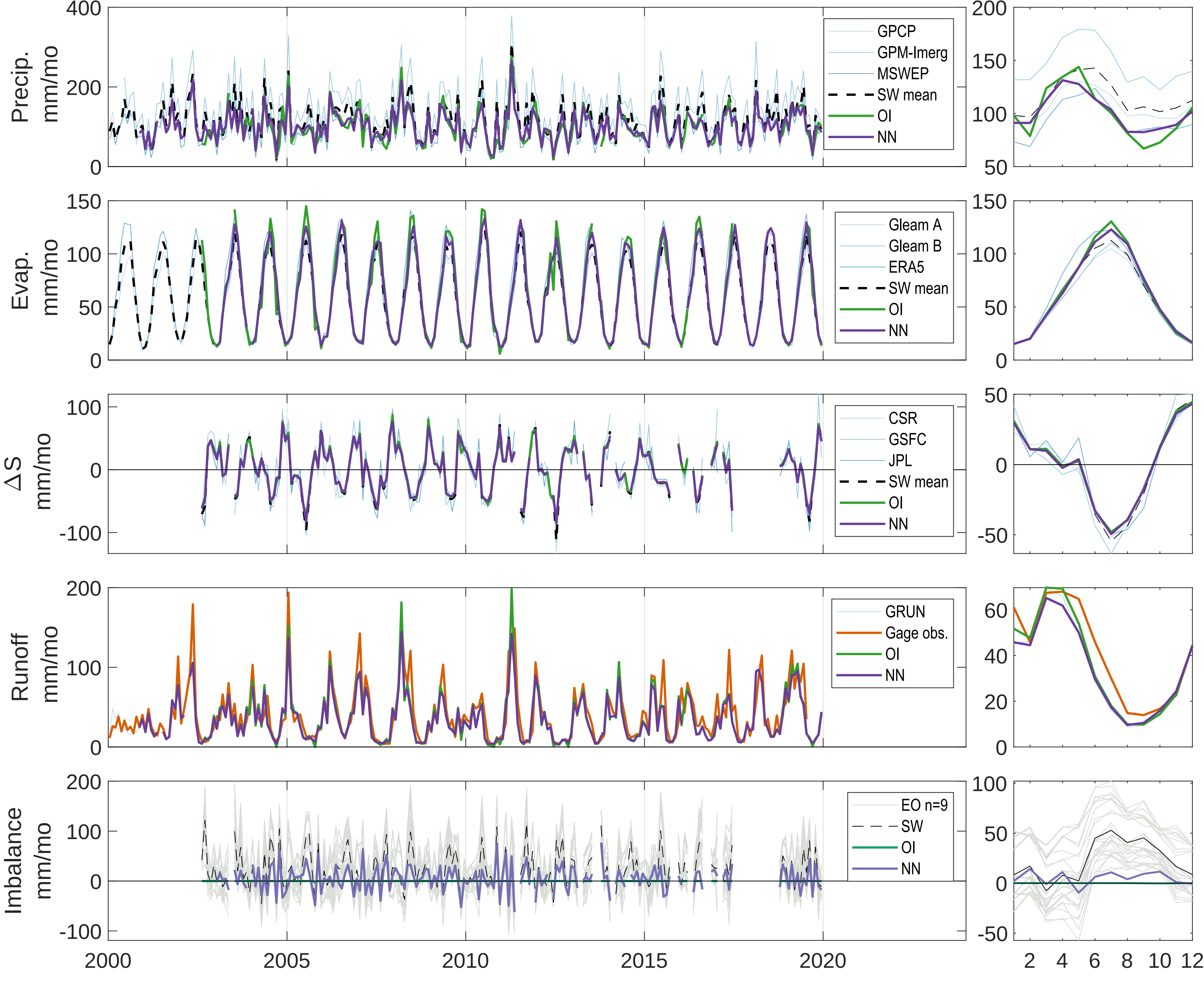
Figure 4.

Neural Network for Expt. 12-5
with goal of reducing Imbalance at the pixel scale



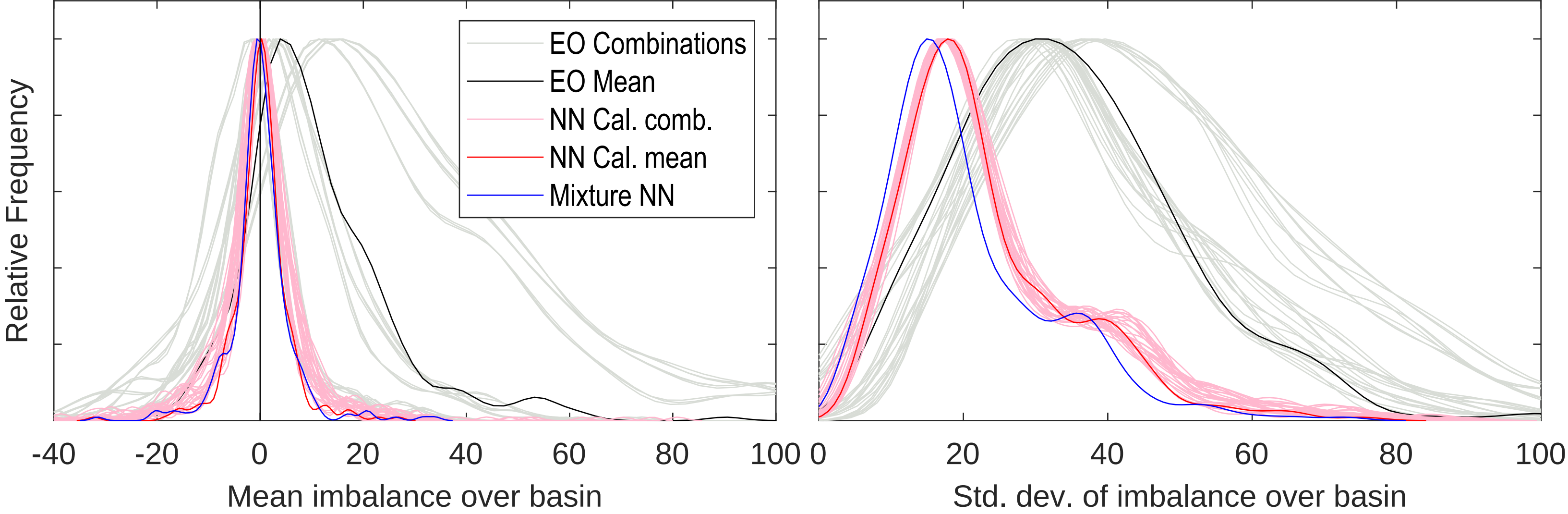
C:\Users\mheberger\Dropbox\Research\Matlab\Project07 Revised NN\networks\NN_Diagrams\network_diagram.xlsx

Figure 5.



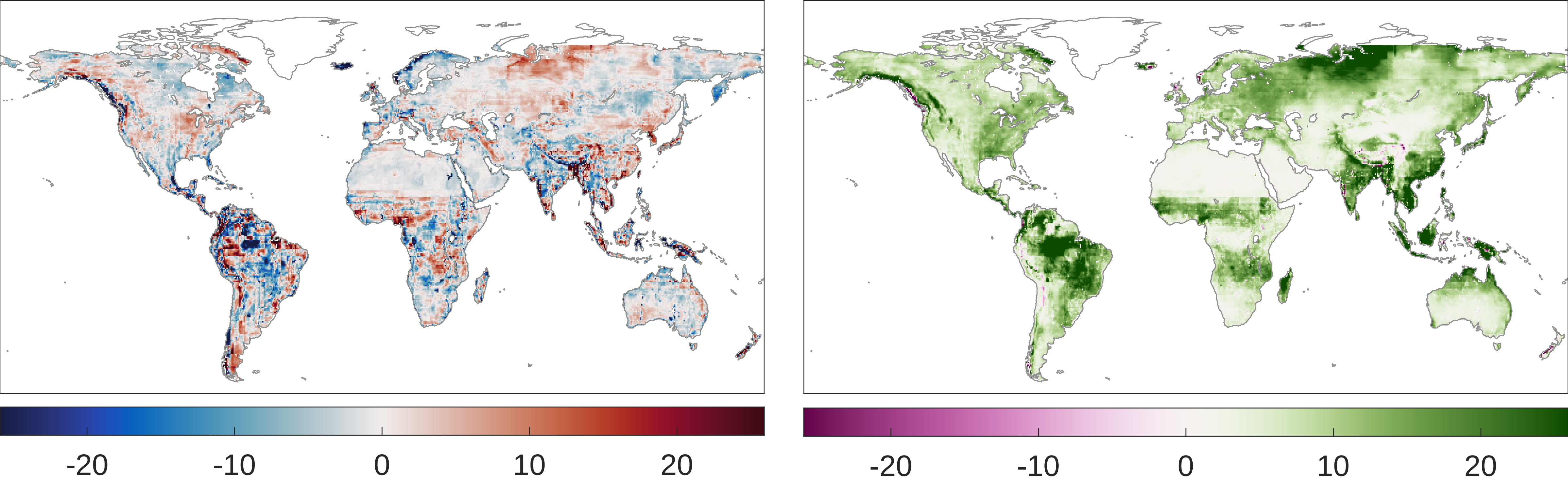
2023-06-29 15:06:45 C:\Users\mheberger\Dropbox\Research\Matlab\Project12 S with Runoff\expt4\m30_ts_plots_for_paper.m

Figure 6.



2023-06-30 13:48:45 C:\Users\mheberger\Dropbox\Research\Matlab\Project12 S with Runoff\expt4\m11_Imbalance_pdf_plot_for_paper.m

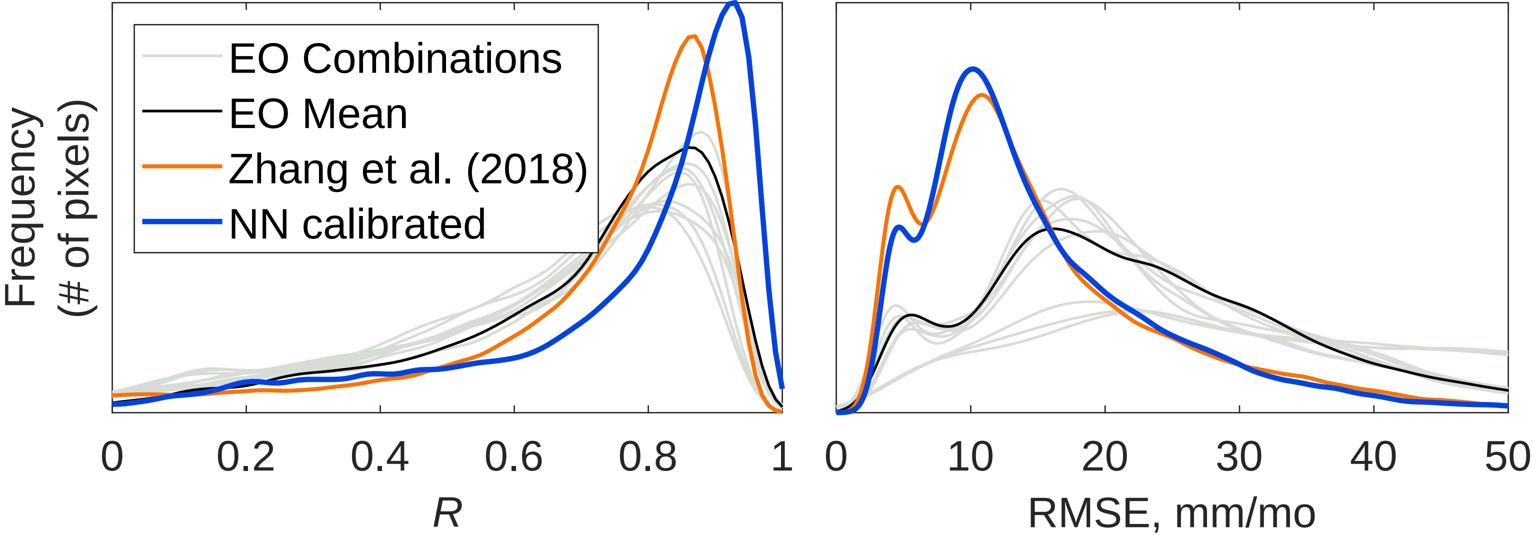
Figure 7.



(a) Mean imbalance in pixels (mm/mo)

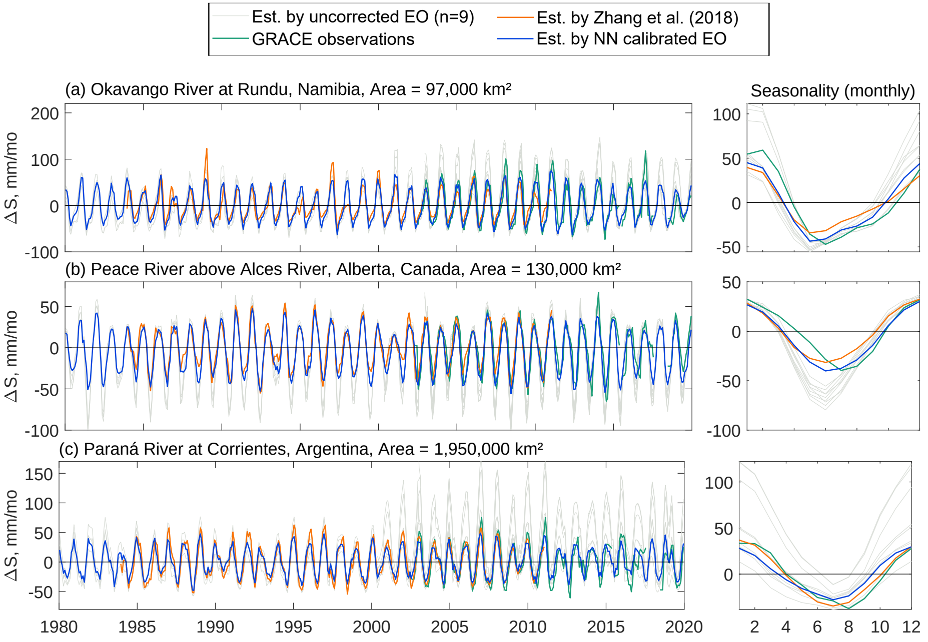
(b) Mean imbalance improvement in pixels (mm/mo)

Figure 8.



2023-06-29 15:44:41 C:\Users\mheberger\Dropbox\Research\Matlab\Project12 S with Runoff\expt9\m12_comparison_my_vs_zhang_smoothed.m

Figure 9.



Est. by uncorrected EO (n=9)
GRACE observations
Est. by Zhang et al. (2018)
Est. by NN calibrated EO
(a) Okavango River at Rundu, Namibia, Area = 97,000 km²
Seasonality (monthly)
(b) Peace River above Alces River, Alberta, Canada, Area = 130,000 km²
(c) Paraná River at Corrientes, Argentina, Area = 1,950,000 km²
ΔS, mm/mo
200
100
0
-100
50
-50
150
1980
1985
1990
1995
2000
2005
2010
2015
2020
2
4
6
8
10
12