Global Evapotranspiration Datasets Assessment Using Water Balance in South America

Anderson Ruhoff 1,* , Bruno Comini de Andrade 1, Leonardo Laipelt 2, Ayan Santos Fleischmann 2, Vinicius Alencar Siqueira 1, Adriana Aparecida Moreira 1, Rafael Barbedo 1*, Gabriele Leão Cyganski 1, Gabriel Matte Rios Fernandez 1, João Paulo Lyra Filho Bréda 1, Rodrigo Cauduro Dias de Paiva 1, Adalberto Meller 3, Alexandre de Amorim Teixeira 1, Alexandre Abdalla Araújo 3, Marcus André Fuckner 3 and Trent Biggs 4

1 Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre 91501-970, RS, Brazil; bruno.comini@ufrgs.br (B.C.d.A.); leonardo.laipelt@ufrgs.br (L.L.); vinicius.siqueira@ufrgs.br (V.A.S.); adriana.moreira@ufrgs.br (A.A.M.); rafael.barbedo@ufrgs.br (R.B.); gabriele.leao@ufrgs.br (G.L.C.); matte.rios@ufrgs.br (G.M.R.F.); joao.breda@ufrgs.br (J.P.L.F.B.); rodrigo.paiva@ufrgs.br (R.C.D.d.P.)
2 Mamirauá Institute for Sustainable Development, Tefe 69553-225, AM, Brazil; ayan.fleischmann@mamiraua.org.br
3 Agência Nacional de Águas e Saneamento Básico (ANA), Brasília 70610-200, DF, Brazil; adalberto.meller@ana.gov.br (A.M.); alexandre.amorim@ana.gov.br (A.d.A.T.); alexandre.araujo@ana.gov.br (A.A.A.); marcus.fuckner@ana.gov.br (M.A.F.)
4 Department of Geography, San Diego State University, San Diego, CA 92182, USA; tbiggs@sdsu.edu
* Correspondence: anderson.ruohoff@ufrgs.br

Abstract: Evapotranspiration (ET) connects the land to the atmosphere, linking water, energy, and carbon cycles. ET is an essential climate variable with a fundamental importance, and accurate assessments of the spatiotemporal trends and variability in ET are needed from regional to continental scales. This study compared eight global actual ET datasets (ETgl) and the average actual ET ensemble (ETens) based on remote sensing, climate reanalysis, land-surface, and biophysical models to ET computed from basin-scale water balance (ETwb) in South America on monthly time scale. The 50 small-to-large basins covered major rivers and different biomes and climate types. We also examined the magnitude, seasonality, and interannual variability of ET, comparing ETgl and ETens with ETwb. Global ET datasets were evaluated between 2003 and 2014 from the following datasets: Breathing Earth System Simulator (BESS), ECMWF Reanalysis 5 (ERA5), Global Land Data Assimilation System (GLDAS), Global Land Evaporation Amsterdam Model (GLEAM), MOD16, Penman–Monteith–Leuning (PML), Operational Simplified Surface Energy Balance (SSEBop) and Terra Climate. By using ETwb as a basis for comparison, correlation coefficients ranged from 0.45 (SSEBop) to 0.60 (ETens), and RMSE ranged from 35.6 (ETens) to 40.5 mm month−1 (MOD16). Overall, ETgl estimates ranged from 0 to 150 mm month−1 in most basins in South America, while ETwb estimates showed maximum rates up to 250 mm month−1. ETgl varied by hydroclimatic regions: (i) basins located in humid climates with low seasonality in precipitation, including the Amazon, Uruguay, and South Atlantic basins, yielded weak correlation coefficients between monthly ETgl and ETwb; and (ii) tropical and semi-arid basins (areas where precipitation demonstrates a strong seasonality, as in the São Francisco, Northeast Atlantic, Paraná/Paraguay, and Tocantins basins) yielded moderate-to-strong correlation coefficients. An assessment of the interannual variability demonstrated a disagreement between ETgl and ETwb in the humid tropics (in the Amazon), with ETgl showing a wide range of interannual variability. However, in tropical, subtropical, and semi-arid climates, including the Tocantins, São Francisco, Paraná, Paraguay, Uruguay, and Atlantic basins (Northeast, East, and South), we found a stronger agreement between ETgl and ETwb for interannual variability. Assessing ET datasets enables the understanding of land–atmosphere exchanges in South America, to improvement of ET estimation and monitoring for water management.
1. Introduction

South America is a major source of water to the atmosphere, being responsible for 21.7% of global actual evapotranspiration (ET) from the continents [1]. It is also the home of important ecosystems with high productivity, such as the Amazon, with tight feedbacks among ET, land surface, and the atmosphere [2], with ET playing a direct role on vegetation development [3]. Recycled precipitation (P) through ET regulates the continent’s climate and water cycles [4]. Furthermore, South American countries are major world food providers, and cropland expansion—in association with land cover changes—has significantly altered the surface energy partition, and thus ET fluxes [5,6]. As an essential climate variable [7,8], monitoring ET is relevant for both supporting new water-infrastructure projects and for regulating water uses to avoid water conflicts, as well as to achieve the United Nations (UN) Sustainable Development Goals [9]. There is a need to estimate water fluxes in the context of national- to continental-scale water management, e.g., to fulfill the needs of Environmental and Economical Water Accounting frameworks [10,11]. Understanding fluxes at the land–atmosphere interface contributes to better estimation of land-boundary conditions in earth-surface models [12,13], consequently improving the accuracy of short- to long-term climate modeling [14]. Accurate estimates of ET may be also useful for the calibration of hydrological models [15,16], which ultimately lead to better prediction of anomalously wet and dry periods and increased confidence regarding climate-change projections.

Actual ET, the combined process of evaporation of water from the surface and transpiration from vegetation, is a major component of the land-energy and hydrological cycles. Currently, local ET measurements can be obtained through a wide range of techniques, including—but not limited to—Bowen ratio systems, weighing lysimeters, and eddy covariance flux towers [17]. Some initiatives, such as the FLUXNET network [18,19], for example, have developed standards for data collection and created a global eddy covariance dataset from regional networks. Yet, like in most parts of the world, South America faces a very sparse coverage of field ET observations. In addition, recent years have seen a boom of global datasets that allow mapping ET from regional to global scales, made possible due to an increasing computational capacity, combined with the launching of several satellite missions [20]. Information obtained through remote-sensing datasets, from visible-to-thermal infrared bands, can be used to retrieve land-surface temperature, vegetation phenology, and other surface properties [21], which in turn are used as inputs to algorithms specifically developed for estimating ET. Typical algorithms for this purpose are the surface energy balance (SEB) approaches [22–24], physical equations based on vegetation phenology and potential ET equations (e.g., Penman–Monteith and Priestley–Taylor) [25,26], statistical or machine-learning models [27,28], as well as water-balance methods [29–31]. In parallel, land-surface models (LSMs) simulate water and energy fluxes through physical and conceptual representations of the terrestrial hydrological system [32], with increasing focus on the depiction of biophysical and biogeochemical processes related to plant and carbon dynamics [33]. LSMs are able to provide global estimates of ET when forced with meteorological data produced in combination with atmospheric models, reanalysis, and data-assimilation techniques [34,35].

Despite the significant advances in remote sensing and LSM, the estimation of ET is challenging and subject to major uncertainties. For instance, thermal remote-sensing data required for the application of SEB models are only reliable for cloud-free conditions, and such methods are very sensitive to the estimation of surface and near-surface temperature [36]. ET models, in general, are dependent on the accuracy of surface net radiation, which usually explains up to 80% of the ET variability [37], while methods that
use physical equations are also subject to errors in additional meteorological variables [36], as well as to estimation of surface and canopy resistance as in the case of the Penman-Monteith approach [38]. Similarly, substantial uncertainties in LSMs are attributed to the existence of a large number of parameters, errors in forcing data, and deficiencies in model structure [32,33].

The increasing availability of global datasets, with different degrees of accuracy, urges effective methods to understand their uncertainties in both spatial and temporal dimensions. While global-scale assessments of ET datasets have been frequently reported over the last decade [39–42], in South America it has been only recently conducted, but with a focus on specific biomes, such as the Amazon basin [43–45]. In particular, Sörensson and Ruscica [46] evaluated the similarity among nine global ET datasets based on remote sensing, LSM, and reanalysis over continental South America, and provided a comprehensive overview of the level of uncertainty in the water and energy limitations of ET. However, detailed validation of global ET datasets relative to reference data in multiple basins of South America has not been explored so far, and knowledge of their regional performance still remains a gap in the scientific literature.

When direct measurements of ET are not available, large-scale ET datasets derived from remote sensing (RS) and LSMs can be compared to ET computed as the water-budget residual for a given basin [47–51]. In this approach, the water-balance equation is solved by using observations of discharge ($Q$), $P$, and terrestrial water-storage variations ($dS/dt$). The latter is currently estimated by the Gravity Recovery and Climate Experiment (GRACE) mission with median accuracy of 20–30 mm of equivalent water thickness for basins smaller than 100,000 km$^2$ [52]. As errors in data from multiple sources may introduce significant biases in the water-budget closure, the uncertainty of individual water-balance components should be evaluated at the basin scale in such a way that the associated ET uncertainties can be properly quantified [30,53,54].

The objective of this paper is to evaluate multiple global actual ET datasets ($ET_{gl}$), freely available and easily accessible, over South America using ET computed from water budget ($ET_{wb}$) as a proxy of observations. The following questions are addressed: (i) How consistent are ET estimates from global datasets based on remote sensing, reanalysis, land-surface, and biophysical models, when compared to $ET_{wb}$? (ii) Do they coherently represent the ET magnitude and seasonality? (iii) Do they agree in terms of seasonal and interannual variability? The assessment, following [30,48,53,55], was performed on monthly time scales for 50 medium-to-large basins from 2003 to 2014, covering eight major South American rivers (Figure 1), including the Amazon, Tocantins, São Francisco, Parana, Uruguay, and Atlantic basins (Northeast, East, and South).

2. Materials and Methods
2.1. Study Area

South America is characterized by a wide range of climate regimes and biomes, with three major climate zones (tropical, arid, and temperate climates), from the humid tropics with equatorial climates to temperate and cold arid zones in the southern part of the continent [56], encompassing several major biomes, including the Amazon, Andes, Atlantic Forest, Caatinga, Cerrado, Chaco, Pampa, Pantanal and Patagonia [57,58]. The largest biome in South America is the Amazon, which has undergone in recent years high rates of deforestation and significant changes in surface–atmosphere feedback [59–61]. The Amazon biome is characterized by a tropical humid climate and low seasonal temperature variability [62]. The Cerrado biome is characterized by savanna vegetation with a tropical seasonal (wet-to-dry) climate [63]. The Caatinga biome, located in northeast Brazil, is described as a seasonally dry tropical forest with arid and semiarid hot climates [63]. In addition, the Chaco biome, located mostly in Argentina but also in Paraguay and Bolivia, is also described as a dry tropical forest (the largest in South America), with a semiarid climate [64]. While most of this biome is still native vegetation, around 25% of its area has been cleared for agriculture and livestock production [65,66].
Figure 1. (a) River gauges and associated drainage areas (river basins) used in this study to assess water balance and remote-sensing-based evapotranspiration estimates in South America. The red dots represent the gauge stations used to assess water balance (the bigger dots refer to the gauges used to illustrate the water-balance assessment in the following figures). (b) South American biomes from the Terrestrial Ecoregions of the World [58], with modifications based on Turchetto-Zolet et al. [57]. (c) Koppen’s climate zones in South America [56,63], denoted as follows: tropical rainforest (Af), tropical monsoon (Am), tropical wet savannah (Aw), arid hot (BWh), arid cold (BWk), semiarid hot (BSh), semiarid cold (BSk), temperate dry hot summer (Csa), temperate dry and warm summer (Csb), temperate dry and hot summer (Cwa), temperate dry winter and warm summer (Cwb), temperate dry winter and cold summer (Cwc), temperate without dry season and warm summer (Cfb), temperate without dry season and cold summer (Cfc), temperate with dry and warm summer (Dsb), temperate without dry season and cold summer (Dfc), polar tundra (ET), and polar frost (EF).

Some of South American biomes, such as the Cerrado, the Pantanal (one of the largest wetland complexes on Earth), and the Atlantic Forest (which has both tropical and subtropical humid climates) are facing environmental and economic threats, including deforestation, fire, cropland expansion, and drought occurrences [67–69]. The Pampa biome, located in Southern Brazil, Uruguay, and Argentina, has a seasonal (cold and hot) subtropical climate and temperate grassland vegetation [70], playing an important role in cropland production. Finally, Patagonia reaches the extreme south of South America, and is mainly covered by temperate grassland and steppe vegetation interspersed with shrubs characterized by temperate and arid cold climates [56].
2.2. Global Evapotranspiration Datasets

A total of eight global datasets were used in this study (Table 1), which were divided into two approaches: (i) remote-sensing-based, and (ii) comprehensive-model-based, including land-surface models, biophysical processes, and simplified water-balance approaches. We also computed the ensemble mean \( \bar{ET}_{gl} \) as the arithmetic average of all available \( ET_{gl} \) datasets. Given data availability, we used the 2003–2014 period to assess \( ET_{gl} \) estimates in South America.

2.2.1. Remote-Sensing-Based Datasets

Four \( ET_{gl} \) datasets relied primarily on remote sensing, which can be divided into vegetation (\( VI \))-based and surface temperature (\( T_s \))-based methods. \( VI \)-based methods include the Global Land Evaporation Amsterdam Model (GLEAM) \([26,71]\), MOD16 \([25,72]\) and Penman-Monteith-Leuning (PML) \([73]\), while \( T_s \)-based ones are represented here by the operational Simplified Surface Energy Balance (SSEBOP) algorithm \([24,74]\).

<table>
<thead>
<tr>
<th>Model</th>
<th>Methods</th>
<th>Spatial Resolution</th>
<th>Temporal Resolution</th>
<th>Remote Sensing Source</th>
<th>Remote Sensing Drivers</th>
<th>Meteorological Drivers</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BESS</td>
<td>Biophysical process-based model</td>
<td>5 km</td>
<td>8 days</td>
<td>MODIS</td>
<td>Atmospheric data (aerosol, water vapor, cloud, atmospheric profile)</td>
<td>Global meteorology (MODIS, ERA Interim and NCEP/NCAR)</td>
<td>Jiang and Ryu \cite{75}; Ryu et al. \cite{18}</td>
</tr>
<tr>
<td>ERA5 (v. Land)</td>
<td>Land-surface model (CHTESSEL)</td>
<td>0.1 degree</td>
<td>Hourly</td>
<td>MODIS, SPOT-Vegetation</td>
<td>Vegetation phenology (leaf area index, vegetation index)</td>
<td>Global meteorology (ERA5)</td>
<td>Hersbach et al. \cite{35}; Nogueira et al. \cite{76}</td>
</tr>
<tr>
<td>GLDAS (v. 2.1)</td>
<td>Land-surface model (NOAH)</td>
<td>0.25 degree</td>
<td>3-hourly</td>
<td>MODIS GPCP</td>
<td>Surface properties Precipitation</td>
<td>Global meteorology (AGRMET) and global data assimilation (GDAS)</td>
<td>Rodel et al. \cite{34}</td>
</tr>
<tr>
<td>Model</td>
<td>Methods</td>
<td>Spatial Resolution</td>
<td>Temporal Resolution</td>
<td>Remote Sensing Source</td>
<td>Remote Sensing Drivers</td>
<td>Meteorological Drivers</td>
<td>References</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------</td>
<td>--------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>---------------------------------------------</td>
<td>------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>GLEAM (v. 3.3b)</td>
<td>Remote sensing (Priestley–Taylor equation)</td>
<td>0.25 degree</td>
<td>Daily</td>
<td>AIRS, CERES, MODIS, ESA-CCI, SMOS</td>
<td>Atmospheric data (radiation, precipitation, air temperature, lightning frequency) Surface properties (snow water equivalent, multisource soil moisture, vegetation cover fraction, vegetation optical depth)</td>
<td>–</td>
<td>Martens et al. [71]; Miralles et al. [26]</td>
</tr>
<tr>
<td>MOD16 (v. 6)</td>
<td>Remote sensing (Penman–Monteith equation)</td>
<td>500 m</td>
<td>8 days</td>
<td>MODIS</td>
<td>Vegetation phenology (leaf area index) Surface properties (land cover, albedo, emissivity)</td>
<td>Global meteorology (MERRA-2)</td>
<td>Mu et al. [25,77]</td>
</tr>
<tr>
<td>PML (v. 2)</td>
<td>Remote sensing (Penman–Monteith equation)</td>
<td>500 m</td>
<td>8 days</td>
<td>MODIS</td>
<td>Vegetation phenology (leaf area index, fraction of photosynthetically active radiation) Surface properties (land cover, albedo)</td>
<td>Global meteorology (GLDAS)</td>
<td>Zhang et al. [73]</td>
</tr>
<tr>
<td>SSEBOP (v. 4)</td>
<td>Remote Sensing (Simplified surface energy balance)</td>
<td>1 km</td>
<td>Monthly</td>
<td>MODIS</td>
<td>Thermal data (land surface temperature) Multispectral data (surface reflectance)</td>
<td>Global meteorology (GLDAS)</td>
<td>Senay et al. [24,78]</td>
</tr>
<tr>
<td>Terra Climate</td>
<td>One-dimensional water balance (modified Thornthwaite–Mather equation)</td>
<td>2.5 arcmin</td>
<td>Monthly</td>
<td>–</td>
<td>–</td>
<td>Global meteorology (WorldClim and JRA55)</td>
<td>Abatzoglou et al. [29]</td>
</tr>
</tbody>
</table>

GLEAM is based on the Priestley–Taylor equation and was developed to estimate surface evaporative fluxes at a spatial resolution of 0.25° and daily temporal resolution [26,71]. GLEAM estimates canopy transpiration, bare-soil evaporation, canopy interception and open-water evaporation, combining multiple sources of data, including snow-water equiv-
alent, soil moisture, and vegetation optical depth and vegetation cover fraction. The algo-
rithm also provides a root-zone soil-moisture profile, infiltration, and evaporative stress
conditions, used as constraints on soil evaporation and canopy transpiration. GLEAM
version 3.3.b adopted here uses only remote-sensing observations (optical and microwave
datasets, and dynamic land-cover classification), including data from CERES (Clouds
and Earth’s Radiant Energy System), AIRS (Atmospheric Infra-Red Sounder), SMOS (Soil
Moisture and Ocean Salinity) and ESA-CCI (European Space Agency—Climate Change
Initiative) [26,71].

MOD16 [25,72] is based on the Penman–Monteith equation and uses vegetation indices
to estimate canopy conductance. The total $ET_{T}$ includes soil and canopy evaporation, and
canopy transpiration at a spatial resolution of 500 m and temporal resolution of 8 days. To
estimate $ET_{T}$, the algorithm includes remote-sensing and global meteorological reanalysis
data. Remote-sensing inputs that represent land-surface processes are derived from the
Moderate Resolution Imaging Spectroradiometer (MODIS) datasets, including land-cover
classification, leaf area index (LAI), fraction of photosynthetically active radiation (fPAR),
and albedo. As meteorological inputs, the algorithm uses reanalysis from MERRA-2
(Modern-Era Retrospective analysis for Research and Applications), with spatial resolution
of 0.5° × 0.65°. To combine meteorological and remote-sensing inputs, meteorological data
are interpolated from coarse spatial resolution to match THE MODIS spatial resolution and
to remove abrupt changes from the meteorological reanalysis data [79].

PML [73] is a physically based model that estimates $ET_{T}$ using the Penman–Monteith
equation. Input data are from MODIS, including LAI, albedo, and emissivity, while mete-
orological inputs are from the Global Land Data Assimilation System (GLDAS) rescaled
to 500 m using a bilinear interpolation. PML computes total $ET_{T}$ by accounting for the
evaporation from soil and canopy interception and vegetation transpiration. Spatial and
temporal resolution are 500 m and 8 days, respectively.

SSEBOP [24,78] is based on a simplified SEB approach with a simple parameteriza-
tion for operational applications. The algorithm combines $ET_{T}$ fractions derived from $T_{s}$
and surface reflectance from MODIS and meteorological reanalysis from GLDAS. The
one-source algorithm estimates total $ET_{T}$ at a spatial resolution of 1 km and monthly
temporal resolution.

2.2.2. Comprehensive-Model-Based Approaches

Four $ET_{T,i}$ datasets were based on land-surface models, including ERA5 [35,80] and
GLDAS [34]; biophysical-process-based models, including the Breathing Earth System
Simulator (BESS) [18,75]; and simplified water-balance approaches from Terra Climate [29].
ERA5 is a global atmospheric-reanalysis dataset [35,80] generated by the European Centre
for Medium-Range Weather Forecasts (ECMWF). ERA5-Land uses the ECMWF land-surface
model Carbon Hydrology-Tiled Scheme for Surface Exchanges over Land (CHTESSEL) to
compute hydrological fluxes [76]. Vegetation characteristics are derived from global land-
cover datasets, while vegetation density and phenology are obtained from lookup tables as
a function of the respective dominant type of vegetation [76]. The dataset is available at
hourly, daily, and monthly temporal resolutions, with a spatial resolution of 0.1°.

GLDAS reanalysis dataset [34] is based on in situ and remote-sensing datasets, in-
cluding MODIS and Global Precipitation Climatology Project (GPCP), forced with the
Global Data Assimilation System (GDAS) and Agricultural Meteorological Modeling Sys-
tem (AGRMET). GLDAS includes four different land-surface models, and in this study we
used NOAH, which uses a Penman–Monteith approach to compute potential $ET_{T}$ ($PET$)
with a parameterization based on surface-atmosphere interactions, vegetation conditions,
and water availability to estimate actual $ET_{T}$. GLDAS is available at 3-hourly, daily, and
monthly temporal resolutions, with a spatial resolution of 0.25°.

The Breathing Earth System Simulator (BESS) is a process-based model designed to es-
imate $ET_{T}$ and gross primary productivity at global scale with fine spatial resolution [18,75].
It couples a one-dimensional atmospheric radiative transfer module, a two-leaf canopy radiative transfer module, and an integrated carbon assimilation and energy balance module, fully based on MODIS land-surface \((T_s, \text{ land cover, LAI, albedo})\) and atmospheric (aerosol, water vapor, cloud, atmospheric profile) datasets. It uses global meteorology from ERA-Interim to gap-fill missing atmospheric data, and global reanalysis from the National Center for Environmental Prediction–National Center for Atmospheric Research (NCEP-NCAR) to compute aerodynamic resistances. The spatial and temporal resolution of the adopted BESS dataset is 5 km and 8 days, respectively.

Terra Climate [29] is a monthly climate dataset that spans from 1958 to present, based on the high-spatial-resolution WorldClim climate data, in addition to Climate Research Unit (CRU) and Japanese 55-year Reanalysis (JRA55) datasets. \(ET\) is computed from a one-dimensional simplified water balance based on a modified Thornthwaite–Mather approach, that uses \(P, PET\) (calculated using Penman–Monteith equation), soil moisture, and snowpack water storage. The model data are available at a monthly temporal resolution and 2.5 arcminutes spatial resolution.

2.3. Water-Balance Estimation of Evapotranspiration

We computed \(ET_{wb}\) from 2003 to 2014, as the residual of the water-balance equation (Equation (1)).

\[
ET_{wb} = P - Q - \frac{dS}{dt}
\]  

(1)

where \(P\) was taken from Multi-Source Weighted-Ensemble Precipitation (MSWEP) version 2.2 [81,82], \(dS/dt\) from Gravity Recovery and Climate Experiment (GRACE) [83–85] and \(Q\) measurements from the Brazilian Water and Sanitation Agency (ANA) and the Argentinian Hydrological Database System (SNIH).

MSWEP dataset was designed specifically for hydrological modeling and assessment, with a spatial resolution of 0.1° and temporal resolution from 3 h to 30 days. MSWEP merges multiple high-quality \(P\) datasets, including measurements, satellite observations, and reanalysis data [81,82].

Monthly \(Q\) was integrated from daily streamflow measurements from ANA (available at [https://www.snirh.gov.br], last accessed on 18 August 2021) and SNIH (available at [https://snih.hidricosargentina.gob.ar/], last accessed on 18 August 2021). We did not gap-fill the data, so we selected basins with complete data, and months with missing data were excluded from analysis. All stations used to compute the water balance are presented in Supplementary Information Table S1.

Changes in storage \((dS/dt)\) were computed using changes in terrestrial water storage \((TWS)\) from GRACE, which detects changes in the Earth’s gravitational field by measuring the distance between two orbiting satellites [83–85] (Equation (2)). The available GRACE TWS data are anomalies relative to the 2004–2009 time-mean baseline, with a spatial resolution of 1° in both latitude and longitude, and monthly temporal resolution (approximately). Considering the irregular time intervals, we computed these changes as the difference between two GRACE observations \((t + 1\) and \(t)\), representing the average change in \(TWS\) [86].

\[
\frac{dS}{dt} \approx \frac{\Delta TWS}{\Delta t} \approx \frac{TWS_{t+1} - TWS_t}{\Delta t}
\]  

(2)

We used a simple average of the three different GRACE spherical harmonic solutions (from GFZ (GeoForschungsZentrum), CSR (Center of Space Research from the University of Texas), and NASA JPL (Jet Propulsion Laboratory)) to reduce the noise in the gravity fields [87].

2.4. Seasonal and Interannual Assessments of Evapotranspiration

To remove the seasonal cycle of the \(ET\) estimations and compute \(ET\) anomalies, we standardized the time series by subtracting the monthly long-term average from each monthly \(ET\) estimate, dividing by the standard deviation [88]. To assess the agreement in
terms of average, phase, and amplitude between the ET time series derived from multiple global datasets, we computed the similarity index ($\Omega$) \cite{46,89} for monthly ET and monthly ET anomaly (Equation (3)).

$$\Omega = \frac{m\sigma_b^2 - \sigma^2}{(m - 1)\sigma^2}$$

(3)

where $m$ is the number of ensemble members ($m = 8$), $\sigma_b^2$ is the ensemble variance of the time series, and $\sigma^2$ is the total variance of all ensemble members concatenated. A value of $\Omega$ equal to 1 indicates identical series, whereas values around 0 indicate poor-to-no agreement. Although values below zero are mathematically possible, they are very unlikely to happen, especially with an increasing number of members ($m$) \cite{89}.

2.5. Assessment of the Accuracy of Global Evapotranspiration

To compare $ET_{gl}$ and $ET_{wb}$, we resampled all global ET datasets to monthly time scales and then we aggregated the basin-wide ET using a simple average, since this scheme has the best performance in preserving the mass balance \cite{90}. To assess the accuracy of the $ET_{gl}$ estimates, we used nonprobabilistic metrics \cite{88} of Pearson’s coefficient of correlation ($r$), bias, and root-mean-squared error (RMSE) between $ET_{wb}$ and $ET_{gl}$.

3. Results and Discussion

3.1. Evapotranspiration Comparison between Water Balance and Global Datasets

To assess how consistent $ET_{gl}$ estimates are on a monthly time scale, a comparison between the water-balance approach and global datasets is presented in Figure 2. Results showed moderate correlation coefficients, with values ranging from 0.45 (SSEBOP) to 0.60 ($ET_{ens}$), while RMSE ranged between 35.6 ($ET_{ens}$) and 40.5 (MOD16) mm·month$^{-1}$. $ET_{gl}$ rates ranged between 0 and 150 mm·month$^{-1}$ in most basins, while $ET_{wb}$ estimates presented maximum rates up to 250 mm·month$^{-1}$. Overall, $ET_{ens}$ performed slightly better than the other global datasets, yielding correlations around 0.60 and RMSE of 35.6 mm·month$^{-1}$. From an assessment based on major basins (Supplementary Information Table S2), two main results can be highlighted: (i) basins located in humid climates (including the Amazon, Uruguay, and South Atlantic basins) exhibited weaker correlation, with average values ranging between 0.37 and 0.48 (between 0.17 and 0.56 in the Amazon, and between 0.37 and 0.51 in the Uruguay and South Atlantic basins), with average RMSE around 39.6 mm·month$^{-1}$ and bias between 11.6 and 23.4 mm·month$^{-1}$; and (ii) tropical and semiarid basins located in areas where $P$ has a strong seasonality (including the Sào Francisco, Northeast Atlantic, Paraná/Paraguay, and Tocantins basins), exhibited stronger correlation, with average values ranging between 0.59 and 0.81, average RMSE between 32.5 and 43.1 mm·month$^{-1}$, and smaller negative bias ranging between 1 and $-14$ mm·month$^{-1}$. These differences may be related to different ET driving factors. While in the Amazon and other basins located in Southern Brazil, $ET$ is mainly driven by available energy (with global radiation showing a strong seasonality in the Uruguay and South Atlantic basins) \cite{91–93}, in the northeast and central areas of Brazil, water availability is the main driver \cite{91,94,95}. Therefore, our results suggest that global ET datasets have closer agreement with $ET_{wb}$ in basins with stronger $P$ seasonality, where $ET$ is mainly driven by water availability, while in basins with humid climate, global ET products yield a low range of seasonal variation due to abundant water availability.
A statistical assessment based on all 50 basins is summarized in Figure 3. Overall, our results indicate a slight overestimation by most $ET_{gl}$ models compared to $ET_{wb}$, with median bias ranging between $-2.4$ and $11.2$ mm-month$^{-1}$ and an interquartile range between $-20$ (for MOD16) and $30$ (for ERA5) mm-month$^{-1}$. In addition, median RMSE smoothly ranged between $34.8$ and $39.7$ mm-month$^{-1}$, with some models showing a higher interquartile range (e.g., ERA5, GLDAS, GLEAM), with errors up to $60$ mm-month$^{-1}$. However, higher correlation values were found for ERA5, GLEAM, MOD16, and $ET_{ens}$ (median higher than $0.7$), while the other $ET_{gl}$ models showed median correlations between $0.51$ and $0.65$). Overall, our results suggest small discrepancies between global datasets, although some models performed slightly better than others, similarly to other reported results [41]. In the long term, all $ET$ datasets presented moderate-to-low bias, with higher bias in humid climates, despite some of the datasets having high monthly errors, similarly to results found in other studies [41,51].
Figure 3. Statistical assessment of monthly evapotranspiration derived from global datasets, relative to the water-balance approach. The analysis is performed for the period 2003–2014 considering 50 medium-to-large basins located in South America. The black dots represent the statistics for all assessed basins, the blue boxes represent the inter-quartile range (IQR) and the median, the whiskers indicate variability outside the upper and lower quartiles (1.5 IQR), whilst the black diamonds represent the outliers.

3.2. Seasonal and Interannual Assessment of Evapotranspiration

The long-term ET seasonal average for the 2003–2014 period is presented in Figure 4 for major South American basins, where ET_{wb} standard deviations are highlighted as shaded areas. Most ET_{gl} presented a similar seasonal pattern in the assessed basins, with ET estimates generally falling within the ET_{wb} deviation range. In basins where P presents a stronger seasonality (i.e., Tocantins, Northeast Atlantic, São Francisco, and Paraná), our results indicate a closer agreement between ET_{wb} and ET_{gl}, yielding a similar seasonal pattern with smaller ET values during the dry season and its smaller deviation. In these basins we found a predominant underestimation (during the wet season and wet-to-dry seasons transition), with average bias between $-30$ (at Tocantins, for MOD16) and
19 mm·month$^{-1}$ (at Northeast Atlantic, for SSEBOP). In addition, SSEBOP estimated higher \( ET \) rates during the dry season in semiarid climates, especially in the São Francisco and Northeast Atlantic basins, while all other models agreed and yielded lower rates during the dry season. Some overestimation was also found in subtropical humid basins such as the Uruguay and South Atlantic basins, mostly during summer months, with average bias between 7 and 36 mm·month$^{-1}$ and up 29 mm·month$^{-1}$, respectively.

Figure 4. Long-term monthly evapotranspiration averages (2003–2014) based on estimates from water balance and global datasets. The shaded areas highlight the water-balance evapotranspiration standard deviation.

In the humid equatorial climate zone, especially in the Amazon Basin, where water availability is not a limiting factor, since \( P \) is relatively well-distributed throughout the year, and available energy is the main driver for \( ET \) [92], the seasonal pattern shown by the \( ET_{wb} \) was not followed by most of the \( ET_{gl} \), with models showing different timing and magnitude when compared to \( ET_{wb} \), as previously discussed [92]. Overall, our results show that \( ET_{gl} \) is overestimated in the humid tropics (mainly in central-west Amazon, such as Itapeua and Obidos, with bias between 2.5 and 51.8 mm·month$^{-1}$) when compared to \( ET_{wb} \), even though an underestimation is seen in eastern parts of the basin (at Itaituba) mostly during the wet season (average bias between $-8$ and $46$ mm·month$^{-1}$). This could be due
to the largest \( P \) contribution on the terrestrial water balance; consequently, \( P \) variations will strongly affect the \( ET_{wb} \) patterns [47]. As in the mainly humid Amazon region, \( ET \) is not only controlled by \( P \) and water availability, but also by radiation and vegetation phenology [92]; consequently, the water-balance approach will be less accurate in humid tropics [51].

Based on an assessment of the seasonal and interannual variability (Figure 5), a strong disagreement between \( ET_{wb} \) and \( ET_{gl} \) estimates was found in the Amazon Basin, leading to a wide range of variability. On the other hand, in tropical, subtropical, and semiarid basins (including the Tocantins, São Francisco, Paraná, Paraguay, Uruguay and Atlantic basins), a stronger agreement was found. Despite the wide range of variability, most models agreed in positive anomalies during the drought of 2010 (Amazon at Itapeua) [96,97] and during the anomalously wet year of 2009 (at Itaituba) [98–100], while \( ET_{wb} \) showed negative anomalies in the same period.

Figure 5. Standardized evapotranspiration anomalies with six-month moving average, derived from water balance and global datasets. Periods with intense floods and droughts reported are highlighted with vertical blue and orange shades, respectively.
In tropical and semiarid climates (Tocantins, São Francisco, Northeast Atlantic, Paraná and Paraguay basins), we found moderate-to-strong agreements between global $ET_{gl}$ and $ET_{wb}$ estimates for most of the $ET$ models and the ensemble mean. Most of the $ET_{gl}$ datasets accurately represented wet and dry anomalous events in these climates. For the long-lasting floods of 2009–2010 that affected the Amazon region, including the Tocantins (at Tucuruí) [98] and the Paraná (at Porto São José) [101] basins, all models consistently estimated a positive $ET$ anomaly over the flooding period, in agreement with water-balance estimates. Another hydrological extreme event well-captured by global models was the long drought that affected the Northeast region in Brazil over the 2012–2016 period [100], in the Northeast Atlantic and São Francisco basins (at Luzilandia and Traipu), with negative strong anomalies being observed since 2012. Furthermore, in subtropical humid climates, drought events that occurred in Southern Brazil (Uruguay and South Atlantic basins) over 2005, 2009, and 2012 [96], were only partially represented by $ET_{gl}$ datasets. Overall, most of the global $ET$ datasets yielded a moderate agreement with $ET_{wb}$; however, in some situations, MOD16, GLDAS, and SSEBOP showed a higher amplitude of anomalies than other models in most of the basins. SSEBOP showed higher variations than other models in the Northeast Atlantic basin (at Luzilandia) for both positive and negative anomalies. MOD16 showed negative anomalies before 2007/2008 and positive anomalies for 2009–2012 in Paraguay (at Puerto Pilcomayo), and Paraná (at Timbues) basins, in contrast to other global models. This MOD16 behavior resulted in a positive $ET$ trend. Similar results were previously reported in the Pantanal biome, with changes in the interannual variability from 2001–2006 to 2007–2012 [102]. MOD16 errors may arise from the algorithm structure [103], meteorological forcing data and possible biases [25,104], and land-cover parameterization [105], and further research is needed to identify this source of uncertainties.

From a climatic viewpoint of water-limited versus energy-limited $ET$ estimates in South America, an assessment based on the agreement among the global datasets, represented by the similarity index $\Omega$ (Figure 6a), indicates (i) higher similarity ($\Omega$ higher than 0.7) from the northeast to the southeast parts of the continent, including the Cerrado, Caatinga, Pantanal, Atlantic Forest, and Pampa biomes, where $ET$ is driven by water and radiation combined [46]; (ii) lower similarity ($\Omega$ between 0.2 and 0.4) in the extreme south of the continent, in the Patagonia biome, where $ET$ is mainly driven by water availability [46]; and (iii) very low similarity ($\Omega$ lower than 0.2) in the Amazon biome and Andes regions, where $ET$ is mainly driven by radiation [46], but also by a combination with radiation and vegetation phenology [92]. The similarity of seasonal variability ($ET$ monthly anomalies, Figure 6b) is relatively higher in northeast Brazil (Caatinga biome) and in central parts of Brazil (Cerrado biome), while moderate similarity was found in southern Brazil (Pampa biome, including areas in Central Argentina, and Uruguay), and lower similarity was found in the Amazon.

In some regions, as in southeastern, central, and northeastern Brazil, the global $ET$ datasets consistently agree on monthly $ET$, but to a lower degree on $ET$ anomalies. Some opposite behavior was found for the Amazon and Patagonia, with higher agreement on monthly anomalies. Considering different dataset sources (remote-sensing-driven models, including GLEAM, MOD16, PML, and SSEBOP; and reanalysis, land-surface, and process-based models, including BESS, ERA5, GLDAS, and Terra Climate), a slightly higher similarity was found for the second group of assessed models for both monthly $ET$ and $ET$ anomalies in the Amazon (Supplementary Information Figure S1). Our findings are similar with another study [46], even though different $ET$ datasets were used. As they indicate, when the similarity is higher, less uncertainty is found, and thus the selection of a particular dataset is less important, while when the similarity of $ET$ anomalies is higher, seasonal variability is better represented by the models, which can be important for studies where this is of particular interest.
3.3. Water-Balance Closure

We assessed the water-balance closure (imbalance) using monthly averages from 2003 to 2014 as \( P - ET - Q - dS/dt \), considering \( ET \) estimates from the global datasets. Figure 7 summarizes the spatial variability of the long-term average (2003–2014) imbalance for major South American hydrological regions. There is a larger imbalance in the Amazon basin, and most global models showed a negative imbalance in the western Amazon (Itapeua and Obidos) and a positive imbalance in the eastern Amazon (Itaituba). The water balance using BESS, PML, SSEBOP, and Terra Climate models led to the lowest imbalance range in the Amazon. The other tropical basins exhibited a smaller positive imbalance. For the Tocantins basin, most global models showed an imbalance average of around 13 mm·month\(^{-1}\), ranging between 6 and 17.8 mm·month\(^{-1}\), except MOD16, whose median value was 30.2 mm·month\(^{-1}\). In the Paraná Basin, the median imbalance ranged between \(-18\) mm·month\(^{-1}\) and 11.5 mm·month\(^{-1}\), with BESS, \( ET_{ens} \), PML, SSEBOP, and Terra Climate yielding imbalances lower than 5 mm·month\(^{-1}\). In the São Francisco basin, \( ET_g \) showed an average imbalance from 3 to 11 mm·month\(^{-1}\), except MOD16, which showed a larger imbalance (median of 21 mm·month\(^{-1}\)). In the Northeast Atlantic basin, SSEBOP showed an average imbalance around \(-19\) mm·month\(^{-1}\), while for the other models the average imbalance ranged from \(-8\) to 9 mm·month\(^{-1}\). Basins located in Southern Brazil (South Atlantic and Uruguay) showed lower imbalance values for BESS, PML, and SSEBOP models (average imbalance from \(-7\) to \(-17\) mm·month\(^{-1}\)), while with the other models a higher imbalance was verified (higher than \(-20\) mm·month\(^{-1}\)).
Remote Sens. 2022, 14, x FOR PEER REVIEW 17 of 23

$\Delta ET$ showed an average imbalance from 3 to 11 mm·month$^{-1}$, except MOD16, which showed a larger imbalance (median of 21 mm·month$^{-1}$). In the Northeast Atlantic basin, SSEBOP showed an average imbalance around −19 mm·month$^{-1}$, while for the other models the average imbalance ranged from −8 to 9 mm·month$^{-1}$. Basins located in Southern Brazil (South Atlantic and Uruguay) showed lower imbalance values for BESS, PML, and SSEBOP models (average imbalance from −7 to −17 mm·month$^{-1}$), while with the other models a higher imbalance was verified (higher than −20 mm·month$^{-1}$).

Figure 7. Spatial variability of the long-term annual average imbalance (2003–2014) for major river basins in South America, considering the difference between precipitation and the sum of terrestrial water storage changes, discharge, and evapotranspiration from global ET datasets.

From the spatial assessment of the long-term water-balance closure (Figure 7, three main results can be highlighted: (i) in the Amazon basin, our results showed higher negative imbalances, ranging from −60 up to 20 mm·month$^{-1}$ using most of the $ET_{gl}$ datasets; (ii) in the southern basins (Uruguay and South Atlantic), higher negative imbalances were also found, mainly for ERA5, GLDAS, GLEAM, and MOD16; and (iii) most of the $ET_{gl}$ datasets yielded lower imbalance, ranging between −10 to 20 mm·month$^{-1}$ in basins.
located in tropical and semi-arid climates (Northeast Atlantic and São Francisco basins), with MOD16 showing higher positive imbalances. Overall, despite some differences among all global models, the long-term water-balance closure demonstrated similar patterns over South American basins, with higher imbalance in the humid tropics and lower imbalance in tropical and semi-arid climates. Overall, our results agree with other research findings, indicating that ET datasets demonstrate hydrological consistency, especially in water-limited conditions, while in humid climates, biases are relatively higher [51]. Furthermore, our results also indicate similarities between different approaches to estimate ET, with no obvious divergences, since most of the models yielded similar statistics. Some models performed relatively better than others, which was concluded in previous studies [41].

4. Conclusions

The main goal of this research paper was to assess publicly available and easily accessible global ET datasets across multiple climate conditions in South America. We evaluated eight global ET datasets based on remote-sensing, reanalysis, land-surface, and biophysical-process-based modeling, as well as the ensemble mean, and compared them with ET estimated from water balance in 50 medium-to-large basins to answer the following questions:

i. How consistent are ET estimates from global datasets based on remote-sensing, reanalysis, land-surface, and biophysical models when compared to ET\(_{wb}\)? Our results indicate similar errors when ET\(_{gl}\) are compared with the water-balance approach (ET\(_{wb}\)), with moderate correlations and a slight tendency of overestimation in most of South American climates. However, basins located in tropical (and subtropical) humid climates presented lower accuracy and weaker correlations when compared to basins located in tropical seasonal (wet-to-dry) and semi-arid climates, suggesting that global ET datasets have closer agreement with ET\(_{wb}\) in basins with stronger P seasonality, when ET is mainly driven by water availability.

ii. Do they coherently represent its magnitude and seasonality? Most ET\(_{gl}\) presented a similar seasonal pattern in the assessed basins, with ET estimates generally falling within the ET\(_{wb}\) deviation range, with basins located in tropical seasonal (wet-to-dry) climates achieving a closer agreement. Nevertheless, in humid climates (especially in the tropics) we found higher differences, with ET\(_{gl}\) yielding different timing and magnitude when compared to ET\(_{wb}\), with an overall tendency of overestimation.

iii. How do they agree in terms of seasonal and interannual variability? Our results indicate a strong disagreement in the seasonal and interannual variability between ET\(_{wb}\) and ET\(_{gl}\) in the humid tropics. However, in tropical seasonal (wet-to-dry) and subtropical climates, we found higher agreement between ET\(_{gl}\) for detection of wet and dry anomalous events. These findings also agree with the similarity assessment, with lower similarity in the humid tropics (Amazon biome) and higher similarity in the Cerrado, Caatinga, Pantanal, Atlantic Forest and Pampa biomes. Overall, similarities were higher for monthly ET than monthly anomalies in these biomes.

The spatial and temporal assessment of the water-balance closure indicated a moderate hydrological consistency in the long-term annual average for most of the ET\(_{gl}\), including ET\(_{ens}\), however, higher imbalances were found in the humid tropics and subtropics. Improvements are needed to achieve lower uncertainties and higher accuracy of ET estimates mostly in the humid tropics, especially for seasonal and interannual variability. The use of ET\(_{ens}\) can reduce uncertainties related to single models and their forcing data [41], improving the quality of ET estimates at regional and continental scales for water management purposes [1]. ET estimates based on multiple models can also be enhanced based on spatial refinements of the water-balance approach and the use of eddy covariance networks to better understand ET drivers and thus improve models’ calibration and validation [55].

Achieving a consistent large-scale water-balance closure still remains a challenge in the humid tropics, and optimization strategies could be used to improve ET estimates [106], and its seasonal and interannual variability [107]. Assessing the sources of uncertainties in
the water-balance approach from all hydrological variables is a critical issue that deserves further investigation, especially to achieve a higher accuracy on each water cycle component and to consistently close the water balance [55]. However, results from this study provided important information about the $ET$ dynamics at basin scale and a basis for the future development of optimized $ET$ based on the combination of multiple datasets at continental scale in South America. Therefore, future research is needed to improve $ET$ estimates for water resources management and hydrological applications.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/rs14112526/s1: Figure S1: Similarity index for all global evapotranspiration datasets, remote-sensing-based models, including GLEAM, MOD16, PML, and SSEBOP, and land-surface, process-based models, and simplified water balance, including BESS, ERA5, GLDAS, and Terra Climate. Maps are provided for monthly absolute and monthly anomaly evapotranspiration estimates. Table S1: Summary of the discharge stations used to assess the water-balance evapotranspiration in major South American river basins. Table S2: Summary of the statistical assessment by evapotranspiration model and by major basins in South America.


Funding: The authors would like to gratefully acknowledge the financial support provided by the Brazilian Agency for the Improvement of Higher Education (CAPES) and the Brazilian National Water and Sanitation Agency (ANA) in the context of the projects “Estimating land surface evapotranspiration using remote sensing models for water management in Brazil”, and “Technological Cooperation for Hydrological Assessments in Brazil—Environmental and Economical Water Accounting”, numbers 88881.178687/2018-01 and TED-05/2019-ANA, respectively. A.A.M., A.S.F. and V.A.S were supported by the National Council for Scientific and Technological Development (CNPq). L.L and J.P.L.F.B were supported by Higher Education Improvement Coordination (CAPES). T.B. was funded in part by National Science Foundation award number BCS-1825046.

Data Availability Statement: The data that support the findings of this study, including discharge stations, river basins and measured discharge, are freely available at https://metadados.snirh.gov.br/ or upon request to the corresponding author.

Acknowledgments: The authors would like to thank the Federal University of Rio Grande do Sul (UFRGS) and the Brazilian Water and Sanitation Agency (ANA) for institutional support.

Conflicts of Interest: The authors declare no conflict of interest.

References

Remote Sens. 2022, 14, 2526


9. Sachs, J.D. From millennium development goals to sustainable development goals. Lancet 2012, 379, 2206–2211. [CrossRef]


41. Liu, W.; Wang, L.; Zhou, J.; Li, Y.; Sun, F.; Fu, G.; Li, X.; Sang, Y.-F. A worldwide evaluation of basin-scale evapotranspiration estimates against the water balance method. J. Hydrol. 2016, 538, 82–95. [CrossRef]