



OPEN

DATA DESCRIPTOR

GLEAM4: global land evaporation and soil moisture dataset at 0.1 resolution from 1980 to near present

Diego G. Miralles¹✉, Olivier Bonte¹, Akash Koppa^{1,2}, Oscar M. Baez-Villanueva¹, Emma Tronquo¹, Feng Zhong¹, Hylke E. Beck³, Petra Hulsman¹, Wouter Dorigo⁴, Niko E. C. Verhoest¹ & Shekoofeh Haghdoust¹

Terrestrial evaporation plays a crucial role in modulating climate and water resources. Here, we present a continuous, daily dataset covering 1980–2023 with a 0.1° spatial resolution, produced using the fourth generation of the Global Land Evaporation Amsterdam Model (GLEAM). GLEAM4 embraces developments in hybrid modelling, learning evaporative stress from eddy-covariance and sapflow data. It features improved representation of key factors such as interception, atmospheric water demand, soil moisture, and plant access to groundwater. Estimates are inter-compared with existing global evaporation products and validated against *in situ* measurements, including data from 473 eddy-covariance sites, showing a median correlation of 0.73, root-mean-square error of 0.95 mm d⁻¹, and Kling–Gupta efficiency of 0.49. Global land evaporation is estimated at $68.5 \times 10^3 \text{ km}^3 \text{ yr}^{-1}$, with 62% attributed to transpiration. Beyond actual evaporation and its components (transpiration, interception loss, soil evaporation, etc.), the dataset also provides soil moisture, potential evaporation, sensible heat flux, and evaporative stress, facilitating a wide range of hydrological, climatic, and ecological studies.

Background & Summary

Terrestrial evaporation (E) or ‘*evapotranspiration*’¹ plays a crucial role in the climate system as a nexus between the water, carbon, and energy cycles, reacting to changes in anthropogenic emissions and propagating their influence throughout the global hydrological cycle. It regulates long-term precipitation and temperature projections through its influence on the water vapour, lapse rate and cloud feedbacks, and it influences the occurrence of extreme events, such as droughts, floods and heatwaves^{2,3}. For water management, E is a net loss of available resources that must be monitored, and for agriculture, crop transpiration determines irrigation needs⁴. Despite this importance, E is highly uncertain at regional and global scales, especially regarding long-term trends and responses to short-term climate anomalies^{5–8}. This uncertainty arises because E is (i) rarely measured in the field, (ii) challenging to model accurately (as it involves both plant physiological responses and complex turbulent atmospheric processes), and (iii) invisible to satellite sensors (despite its imprint on surface water and energy balance)⁹. The critical yet uncertain nature of E has spurred innovative attempts to combine *in situ*, satellite, and reanalysis data to estimate global E . Since over a decade ago, myriad approaches to derive global E datasets have been proposed^{10–12}, often based on the application of prognostic models originally designed for regional scales^{13–15}. Machine learning models trained on eddy-covariance measurements have also been used to represent E globally, leveraging both satellite and *in situ* data^{16,17}. However, pure machine learning-based approaches do not explicitly obey physical limits, and their black-box nature complicates interpretability and process understanding¹⁸. An emerging research direction is to combine both physics-based and machine learning models in

¹Hydro-Climate Extremes Lab (H-CEL), Ghent University, Ghent, Belgium. ²Laboratory of Catchment Hydrology and Geomorphology, École Polytechnique Fédérale de Lausanne (EPFL), Sion, Switzerland. ³Physical Sciences and Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. ⁴Department of Geodesy and Geoinformation, TU Wien, Vienna, Austria. ✉e-mail: diego.miralles@ugent.be

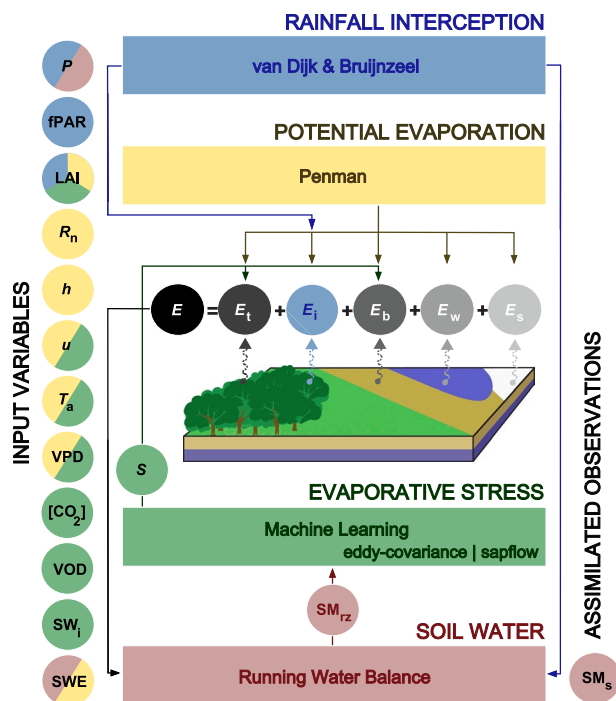


Fig. 1 Schematic of GLEAM4. Variable names and data sources are listed in Table 1, and their colour denotes the module in which they are employed.

Variable	Source	Type	Resolution
Net radiation (R_n) and shortwave incoming (SW_i)	MSWX v1.00 ³⁷	Downscaled reanalysis	0.1°
	CERES ^{51,68}	Satellite	0.5°
Air temperature (T_a)	MSWX v1.00 ³⁷	Downscaled reanalysis	0.1°
	AIRS ^{38,69}	Satellite	1°
Precipitation (P)	MSWEP v2.8 ⁵⁰	Observational merger	0.1°
	IMERG Final V07 ^{52,70}	Satellite	0.1°
Wind speed (u)	ERA5 ^{71,72}	Reanalysis	0.25°
Vapour pressure deficit (VPD)	MSWX v1.00 ³⁷	Downscaled reanalysis	0.1°
	AIRS ^{38,69}	Satellite	1°
Carbon dioxide concentration (CO_2)	CAMS ^{73,74}	Reanalysis	0.75°
Snow water equivalent (SWE)	GlobSnow ⁷⁵ /NSIDC ⁷⁶	Satellite	25 km
Surface soil moisture (SM_s)	ESA CCI ^{48,77}	Satellite	0.25°
Vegetation optical depth (VOD)	VODCA v2 ^{78,79}	Satellite	0.25°
Fraction absorbed photos ynthetic radiation (fPAR)	MOD15A3H ⁸⁰	Satellite	500 m
Leaf area index (LAI)	MOD15A3H ⁸⁰	Satellite	500 m
Vegetation height (h)	GEDI/Landsat ³⁶	Satellite	30 m
Land cover fractions	MEaSURES ⁸¹ /MOD44B ⁸²	Downscaled reanalysis	0.05°/ 250 m
Soil properties	HiHydroSoil ⁸³	Observation driven	250 m

Table 1. Sources of data used in GLEAM4. When two datasets are available for the same variable, the top one refers to data archive ‘a’ and the bottom one to data archive ‘b’.

a synergistic manner, to yield what is frequently referred to as ‘hybrid models’ and which has already met some success in E modelling^{19–22}.

The Global Land Evaporation Amsterdam Model (GLEAM¹⁰), one of the first prognostic approaches developed to estimate E globally using satellite data, remains freely available and widely used, with over 10,000 independent users over the past decade²³. GLEAM data are updated to near present at least once a year, and include not just E , but also its different component fluxes (or sources): transpiration (evaporation of water within the leaves), interception loss (from wet surfaces), bare soil evaporation (within soil pores), evaporation from inland water bodies, and evaporation from snow-covered surfaces (typically, yet inaccurately, referred to as ‘sublimation’¹). The dataset also includes other related variables, such as (surface and root-zone) soil moisture, potential evaporation, and evaporative stress. GLEAM data have been used for a wide range of purposes, including

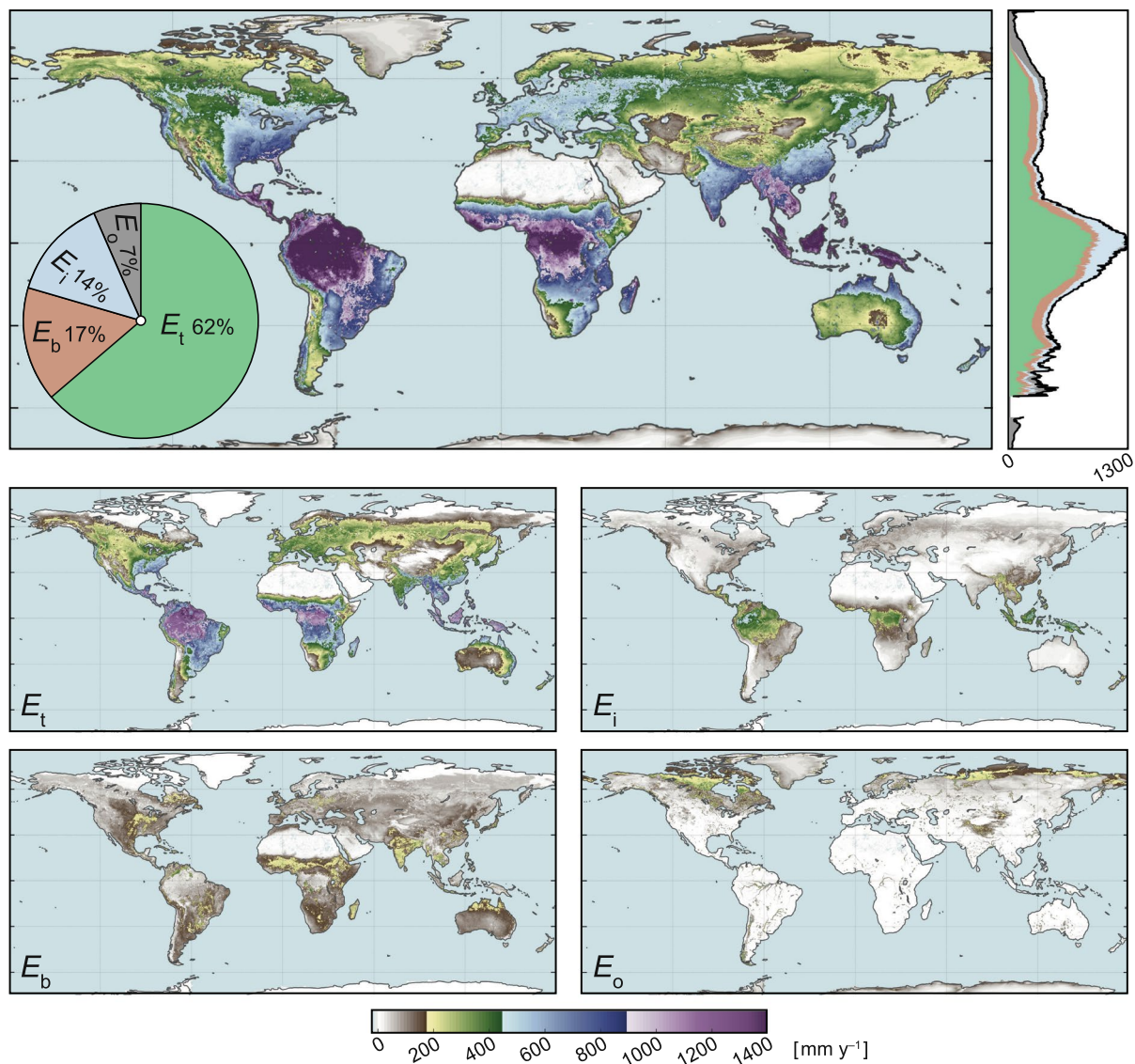


Fig. 2 Mean evaporation and its components. Long-term (1980–2023) mean terrestrial evaporation (E) (top). The global decomposition of its component fluxes is shown in the pie diagram and the latitudinal profile on the right. Long-term means for each of the components (bottom): transpiration (E_t), interception loss (E_i), soil evaporation (E_b), and other sources (E_o), including snow sublimation (E_s), open water evaporation (E_w), and condensation (E_c). All fluxes are in mm yr^{-1} .

the quantification of water resources, driving basin-scale hydrological models, studying global climate trends, and benchmarking climate models²³. Over the years, extensive evaluations of GLEAM against *in situ* observations^{24,25} and alternative gridded datasets^{26,27} have evidenced the consistent performance of the dataset in a wide range of applications²³. Recent improvements in GLEAM have concentrated on the use of machine learning to represent evaporative stress²¹, the depiction of groundwater access by vegetation²⁸, and the characterisation of interception loss²⁹. Here, we unify these efforts and present the fourth generation of the GLEAM algorithm and datasets, GLEAM4, which also features an improved spatial resolution (from 0.25 to 0.1°) and extended record length (1980–2023). In the following sections, the approach is explained with specific emphasis on the novel aspects compared to its predecessor, driving data are introduced, and the resulting global datasets are analysed in terms of spatiotemporal consistency and performance.

Methods

The rationale behind GLEAM is to focus exclusively on processes that directly impact E while maintaining a parsimonious approach. The aim is to extract the most relevant information about E from existing Earth observations, incorporating new processes only if they are both crucial and can be effectively constrained by observations. Following this rationale, GLEAM calculates E (and its components) through four sequential steps (or ‘modules’) targeting the computation of (i) interception, (ii) potential evaporation, (iii) soil water content, and (iv) evaporative stress¹⁰. Figure 1 provides a schematic of GLEAM4 together with the input and output variables

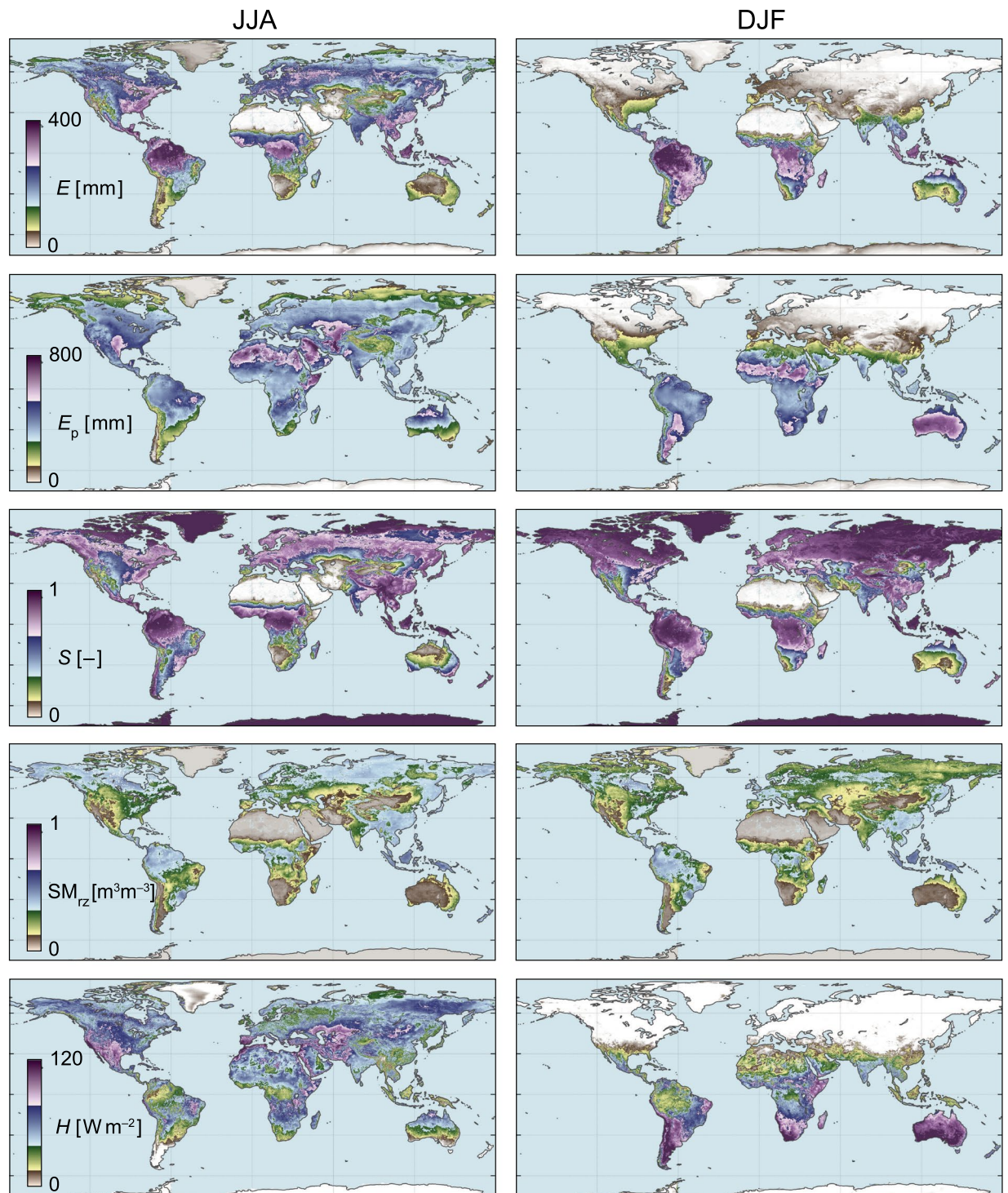


Fig. 3 Seasonal patterns. Long-term (1980–2023) means for June, July, and August (JJA, left) and December, January, and February (DJF, right) for terrestrial evaporation (E , mm), potential evaporation (E_p , mm), evaporative stress ($S = E/E_p$, –), root-zone soil moisture (SM_{rz} , $m^3 m^{-3}$), and surface sensible heat flux (H , $W m^{-2}$).

of each module. First, rainfall interception loss over vegetated surfaces (E_i) and potential evaporation (E_p) are computed. E_p is then converted into actual evaporation (E) using a multiplicative evaporative stress factor (S) that is based on root-zone soil moisture (SM_{rz}), among other variables²¹. This is done independently for the fraction of bare soil, and the fractions of tall and short vegetation within each pixel, yielding estimates of (actual) bare soil evaporation (E_b) and transpiration (E_t), respectively. To consider understory bare soil evaporation, GLEAM4 computes the transmission of incoming radiation through the canopy using the Beer-Lambert law

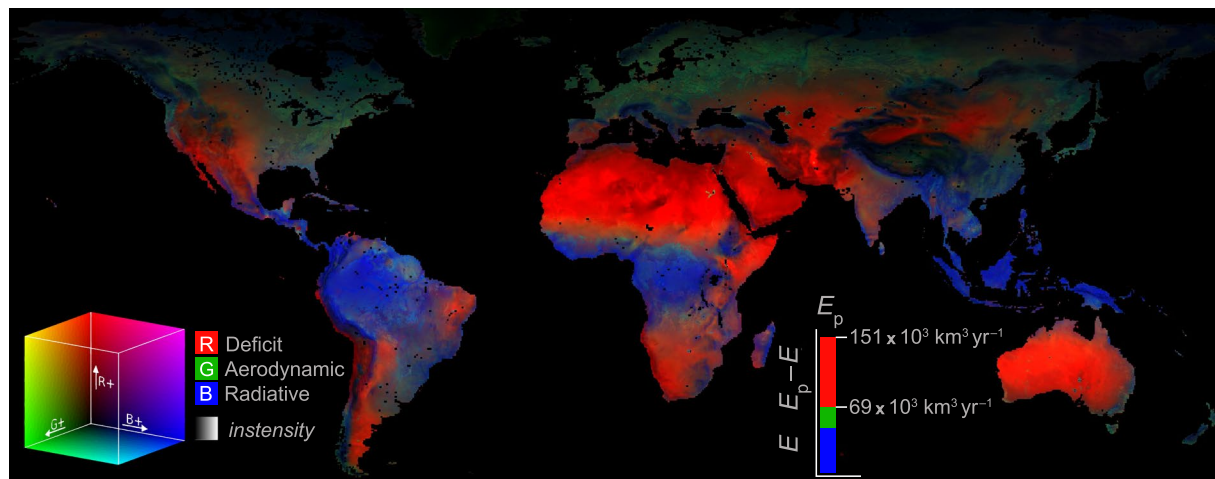


Fig. 4 Evaporation control factors. Relative importance of radiative energy (left-hand term in Eq. 1), aerodynamics (right-hand term in Eq. 1), and evaporative deficit ($E_p - E$), based on long-term (1980–2023) mean fluxes. The bar diagram indicates the global (area-weighted) mean of each term.

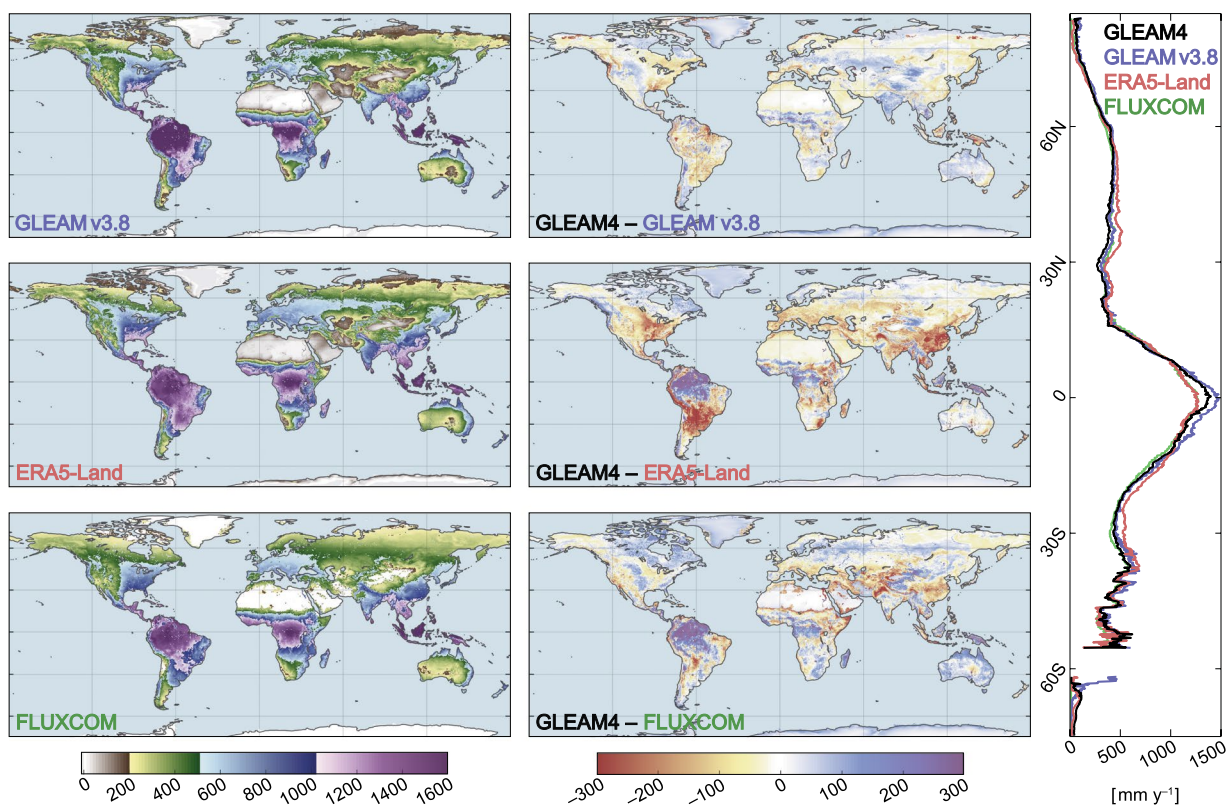


Fig. 5 Global inter-comparison of E estimates from different datasets. The left column shows long-term means for GLEAM v3.8³¹, ERA5-Land⁵⁹, and FLUXCOM⁵⁶ for their common period (1980–2020). The difference between them and GLEAM4 is shown on the right, along with the mean latitudinal profile for all four datasets.

based on leaf area index (LAI)³⁰. Finally, E is considered to equate E_p in regions covered with snow and ice (E_s), and open water (E_w), using specific parametrisations for these surfaces¹⁰.

Since the publication of GLEAM version 3 (v3)³¹, several research efforts have concentrated on improving multiple aspects of the modelling framework. These improvements concern each of the four modules in GLEAM, as illustrated in Fig. 1, and have been documented in individual publications over the past few years^{21,28,29}. Moreover, (i) datasets were extended to near-present, and (ii) spatial resolution was increased from 0.25° to 0.1° thanks to the updates in forcing data. The following provides a general summary of the methodology

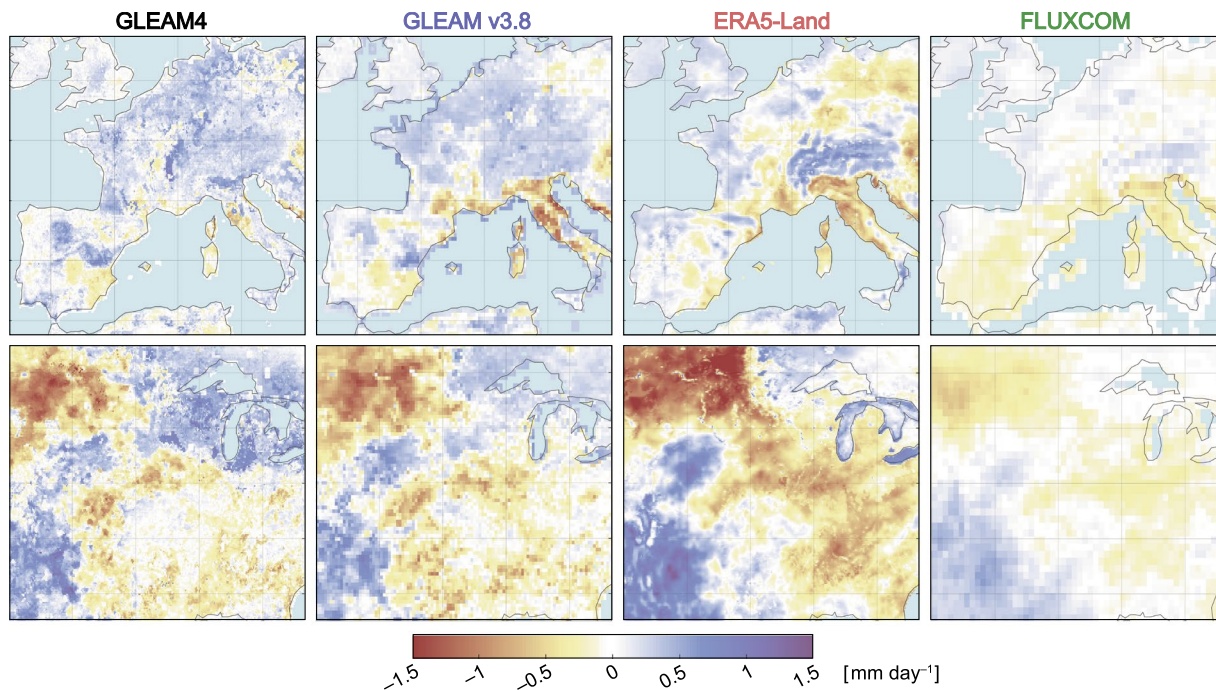


Fig. 6 Regional inter-comparison during summer droughts. Top figures refer to the E anomalies during the 2003 summer (JJA) European drought. Bottom figures show E anomalies during the 1988 summer (JJA) North American drought. From the left to right: GLEAM4, GLEAM v3.8³¹, ERA5-Land⁵⁹, and FLUXCOM⁵⁶ at their original resolution (i.e., 0.1°, 0.25°, 0.1°, 0.5°, respectively).

behind GLEAM4, concentrating on the improvements upon GLEAM v3 and how these affect the estimates of E , its component fluxes, and other related hydroclimatic variables.

Rainfall interception. E_i is computed on rainy days, contributing to E while influencing effective precipitation and soil water content. It remains one of the most uncertain fluxes in the global water cycle, mainly due to the limited availability of *in situ* campaign data for parameterising (or training) universal models³². In GLEAM4, the previous approach based on Gash's analytical model³³ is replaced by the approach in ref. ²⁹, which performed a synthesis of interception data from past field experiments conducted worldwide, including campaigns in 166 forest sites and 17 agricultural plots. Based on this meta-analysis, a global van Dijk–Bruijnzeel interception model³⁴ was constrained using satellite-observed vegetation dynamics — i.e., fraction of absorbed photosynthetically active radiation ($fPAR$) and LAI — potential evaporation, and precipitation data²⁹. This E_i formulation accounts for sub-grid heterogeneity, computing the flux for tall and short vegetation fractions separately, thus improving upon previous GLEAM versions that only included interception for tall vegetation. This improved performance has been demonstrated in validation experiments against field data²⁹ as well as product inter-comparisons³⁵.

Potential evaporation. Unlike previous GLEAM versions, which were based on Priestley and Taylor's equation¹⁴, GLEAM4 uses Penman's equation¹³ to explicitly reflect the influence of wind speed (u , $m\ s^{-1}$), vegetation height (h , m), and vapour pressure deficit (VPD, Pa) on E_p . The original motivation for using Priestley and Taylor was its minimum input requirements (i.e., net radiation and air temperature), making it well-suited for satellite data applications¹⁰. However, recent advances in satellite remote sensing and climate reanalysis have yielded observational datasets of h ³⁶, u ³⁷, and VPD³⁸, making Penman's approach increasingly suited for purpose. As such, E_p ($mm\ s^{-1}$) in GLEAM4 is calculated for each land cover fraction within each pixel as:

$$E_p = \frac{\Delta \times (R_n - G) + \rho_a \times c_p \times g_a \times VPD}{\lambda \times (\Delta + \gamma)}, \quad (1)$$

where Δ is the slope of saturation vapour pressure curve ($Pa\ ^\circ C^{-1}$), R_n is the surface net radiation ($W\ m^{-2}$), G the ground heat flux ($W\ m^{-2}$), ρ_a is the air density ($kg\ m^{-3}$), c_p is the specific heat at constant pressure ($J\ kg^{-1}\ ^\circ C^{-1}$), g_a is the bulk aerodynamic conductance ($m\ s^{-1}$), λ is the latent heat of vaporisation of water ($J\ kg^{-1}$), and γ is the psychrometric constant ($Pa\ ^\circ C^{-1}$). Δ , ρ_a , λ and γ are computed as a function of air temperature¹⁰, R_n is partitioned per land cover fraction³¹, and the G/R_n ratio is considered an inverse function of LAI³⁹. g_a is approximated using Thom's equation⁴⁰, assuming a neutral atmosphere and accounting for the excess resistance of the transfer of vapour compared to momentum⁴¹:

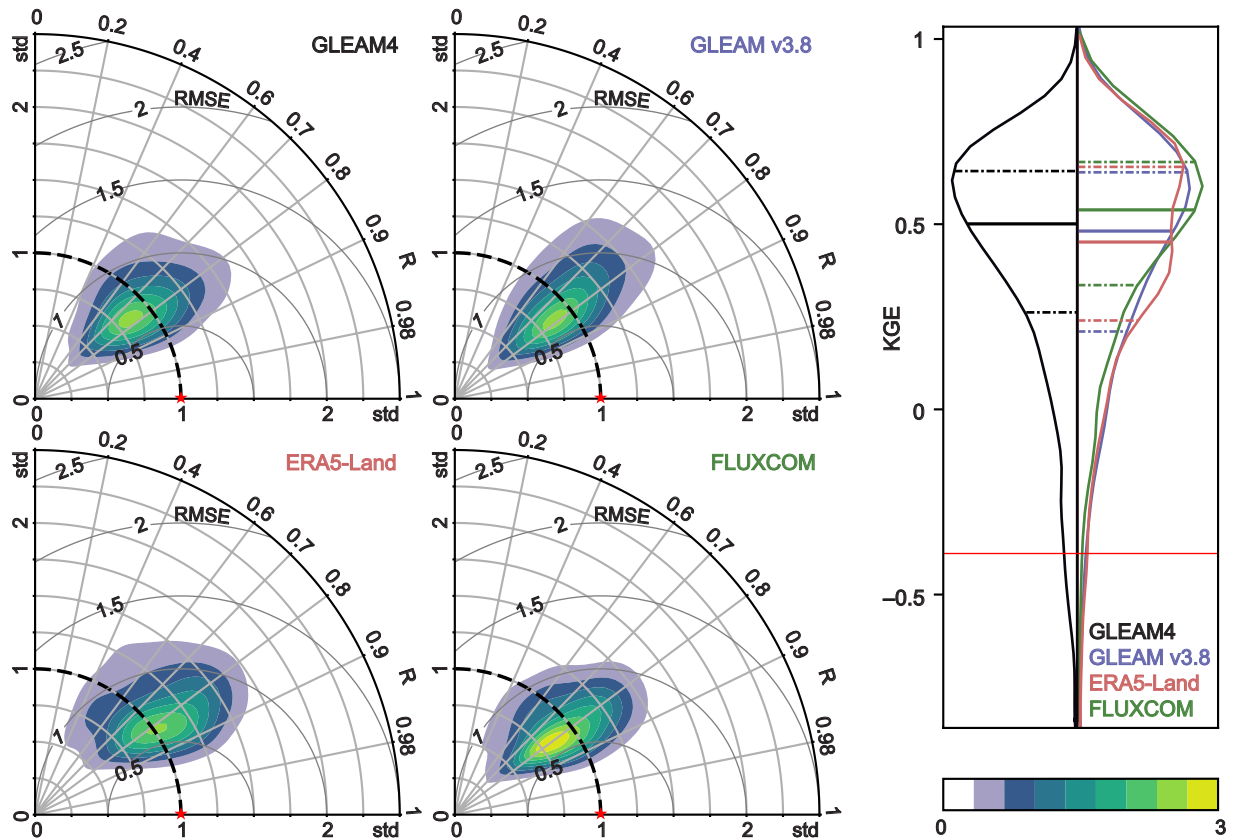


Fig. 7 Validation using *in situ* eddy-covariance measurements. Density Taylor diagrams for the validation of GLEAM4, GLEAM v3.8³¹, ERA5-Land⁵⁹, and FLUXCOM⁵⁶ against *in situ* data (left). Normalized standard deviation (std), Pearson's correlation (R) and centred pattern root mean square error (RMSE) are shown. The ideal performance is marked with a red star. Gaussian Kernel density functions are used to estimate the probability density of data points in the polar space (defined by R and std), thus contour levels indicate unitless thresholds of density values. Violin plots of Kling–Gupta Efficiency for the four datasets (right). Median and interquartile ranges are marked with continuous and discontinuous lines, respectively. Validation based on data from 473 sites.

$$g_a = k^2 \times u \times \ln\left(\frac{z-d}{z_{0m}}\right)^{-1} \times \ln\left(\frac{z-d}{z_{0v}}\right)^{-1}, \quad (2)$$

where k is von Kármán's constant (0.41), z is the height of the u observations (m), d is the zero-plane displacement height (m), and z_{0m} is the roughness length for momentum (m); d and z_{0m} are assumed to be $2/3$ and $1/10$ of h respectively for the non-vegetated fractions⁴², with an additional dependence on LAI for the vegetated fractions⁴¹. The roughness length for vapour (z_{0v}), which is assumed to be the same for heat (z_{0h}), is calculated from z_{0m} via the so-called kB^{-1} approach, where $kB^{-1} = \ln(z_{0m}/z_{0h})$ is a vegetation specific value taking values of 5 and 8 for short and tall vegetation, respectively⁴³.

Soil water. Soil water content across the root depth is required for later computation of evaporative stress (see below). GLEAM uses a multi-layer running water balance driven by precipitation data (and E), which considers a constant root depth per land cover fraction³¹. Microwave soil moisture (SM_s) and/or backscatter observations are assimilated^{44,45} in the top soil layer. Plants are assumed to be able to extract water from where it is more easily accessible within the soil profile; thus the wettest soil layer is selected for computing evaporative stress. In nature, groundwater can also be an important source for E , especially during dry conditions and in ecosystems where vegetation has deep roots⁴⁶. Previous versions of GLEAM did not explicitly account for plant access to groundwater. The approach adopted by GLEAM4 uses a linear reservoir model to represent groundwater, and introduces a partitioning of transpiration to estimate groundwater-sourced E ²⁸. Validations against field observations of E , soil moisture, discharge and groundwater levels demonstrated a realistic representation of E under water-limited conditions²⁸, enabling the future assimilation of satellite gravimetry data into the model⁴⁷. GLEAM4 assimilates surface soil moisture data from the European Space Agency (ESA) Climate Change Initiative (CCI)⁴⁸ (see Table 1). The data assimilation is based on a Newtonian Nudging scheme where soil moisture is first decomposed into anomalies, and then uncertainties in the latter are computed based on triple collocation³¹.

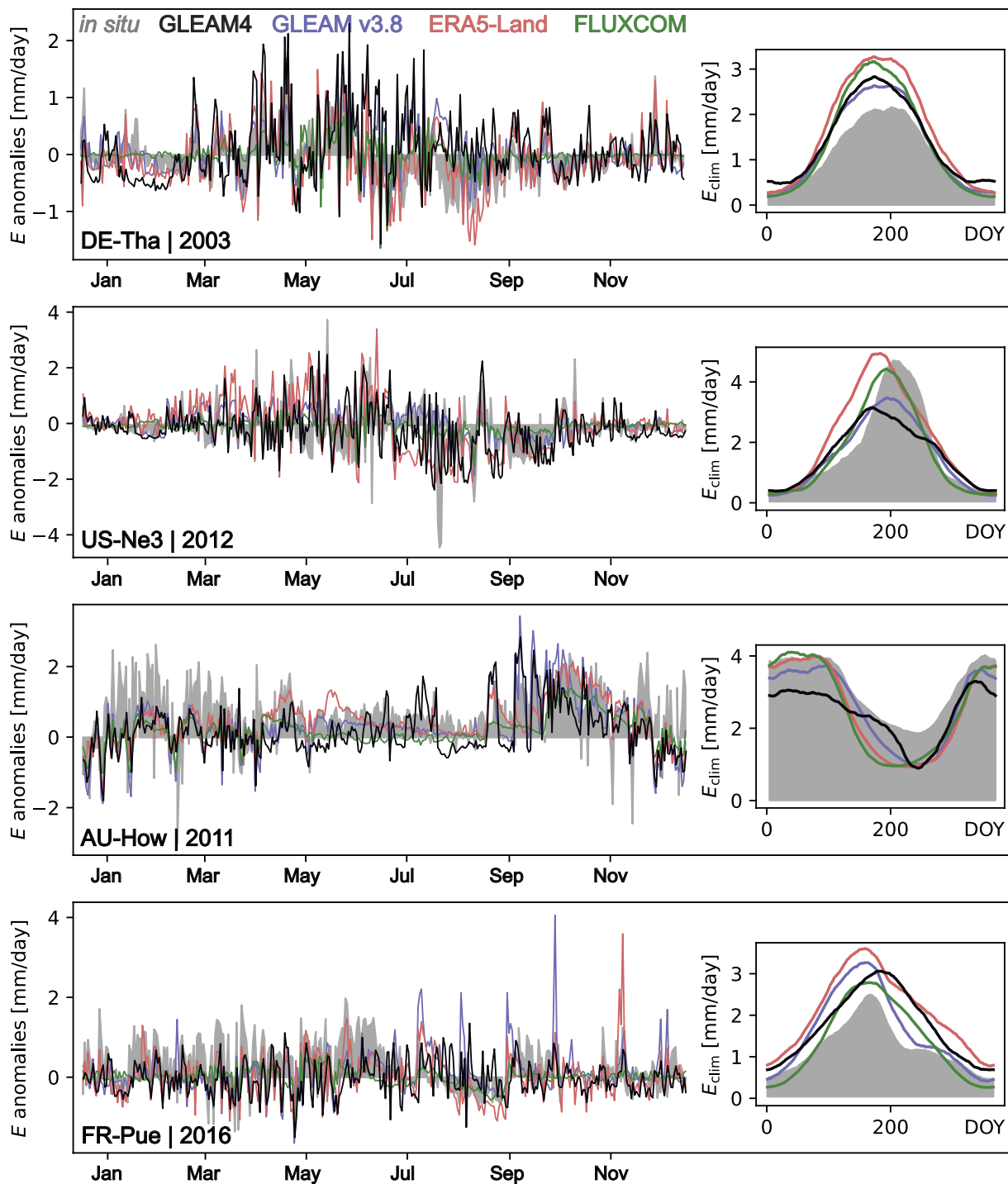


Fig. 8 Example time series. Time series for GLEAM4, GLEAM v3.8³¹, ERA5-Land⁵⁹, and FLUXCOM⁵⁶ against *in situ* data from four sample sites during a selection of extreme years. Left plots indicate anomalies in E while the right plots show the seasonal climatology for each day of the year (DOY).

Evaporative stress. As mentioned above, to constrain E below E_p , GLEAM uses a multiplicative stress factor (S) that ranges from 0 (maximum stress) to 1 (no stress). S is expected to capture all factors that restrain the supply of water to the atmosphere below the atmospheric demand (i.e., E_p). In GLEAM4, the original semi-empirical computation of S based on soil moisture (see above) and vegetation optical depth (VOD)^{10,31}, is replaced by the deep neural network approach presented in ref. ²¹. The latter acknowledges that the ratio between actual and potential transpiration can be controlled by numerous environmental variables that interact non-linearly, including not just soil moisture and VOD, but also VPD, incoming solar radiation (SW_i), air temperature (T_a), CO_2 concentration, u , and LAI — see Fig. 1. Global eddy-covariance and sapflow data are used to learn universal transpiration stress functions, separately for tall and short vegetation. The neural network formulations of S

are embedded within GLEAM4, enabling bidirectional coupling with the process-based model and influencing both E and the soil moisture used to compute S in the next time step. Comparisons against *in situ* data and satellite-based proxies demonstrated that the hybrid (AI–process-based) approach has an enhanced ability to estimate S and E for most ecosystems compared to previous GLEAM versions²¹.

Input data. Table 1 lists all the input variables and datasets used in the generation of GLEAM4 datasets. Input data have been resampled to a common 0.1° resolution by means of bilinear interpolation, when needed.

Data Records

Data archive. The GLEAM4 dataset currently amounts to approximately 1.1TB. Data are freely available in a public SFTP server under a CC BY licence, and can be accessed through <https://www.gleam.eu/#downloads>. For a detailed description of the dataset, we refer readers to the technical notes at <https://doi.org/10.5281/zenodo.14056079>. The E dataset from the exact sub-version used in this manuscript (v4.2a) can also be found in this repository. Dataset specifications are also found on the GLEAM website and through its Digital Object Identifier (DOI)⁴⁹. The period covered by GLEAM4 is currently 1980–2023 at daily temporal resolution and 0.1° spatial resolution. GLEAM datasets are updated annually (in March–April) and extended until the end of the previous year, as input data become available.

Like in previous versions of GLEAM³¹, two distinct data archives are available that differ on their temporal coverage and reliability on observational data. The archive ‘a’ relies mostly on the reanalysis from Multi-Source Weather (MSWX³⁷) and precipitation from Multi-Source Weighted-Ensemble Precipitation (MSWEP⁵⁰) as forcing data. It covers the entire period 1980–2023, and it is intended for climatological studies requiring longer record lengths. The archive ‘b’ has a more observational nature and a lower reliance on reanalysis due to its use of radiation fluxes from the Clouds and the Earth’s Radiant Energy System (CERES⁵¹), precipitation from Integrated Multi-satellite Retrievals for the Global Precipitation Mission (IMERG⁵²), and temperature and VPD from the Atmospheric Infrared Sounder (AIRS³⁸). Its record length is however shorter, currently spanning the period 2003–2023. Table 1 lists the specific datasets used to generate both data archives.

The following 12 variables are available for each of the two data archives (‘a’ and ‘b’):

- Actual evaporation (E , mm day⁻¹)
- Transpiration (E_t , mm day⁻¹)
- Bare soil evaporation (E_b , mm day⁻¹)
- Interception loss (E_i , mm day⁻¹)
- Open-water evaporation (E_w , mm day⁻¹)
- Condensation (E_c , mm day⁻¹)
- Evaporation over snow and ice (E_s , mm day⁻¹)
- Potential evaporation (E_p , mm day⁻¹)
- Evaporative stress (S , unitless)
- Surface soil moisture (SM_s , m³ m⁻³)
- Root-zone soil moisture (SM_{rz} , m³ m⁻³)
- Surface sensible heat flux (H , W m⁻²)

Data structure. Data are organized into netCDF files, with one file per variable per year. Each daily file contains a 3D array with dimensions $n \times 1800 \times 3600$, where n is the number of days in the respective year, 1800 is the number of grid cells in the latitudinal dimension, and 3600 is the number of grid cells in the longitudinal dimension. The first cell in each file corresponds to January 1st of that year, centred at latitude 89.95° and longitude -179.95°. In addition to daily data, netCDF files containing monthly (dimensions 12 \times 1800 \times 3600) and annual (dimensions 1 \times 1800 \times 3600) means are also available.

These datasets are stored on the public server in the following directory structure: <ARCHIVE>/<TEMPORAL_RESOLUTION>/, where <ARCHIVE> refers to either GLEAM4.2a (v4.2a) or GLEAM4.2b (v4.2b), <TEMPORAL_RESOLUTION> indicates the temporal aggregation level (‘daily’, ‘monthly’, or ‘yearly’), and ‘v4.2’ indicates the subversion of the GLEAM4 dataset. Daily datasets are organized by year, while monthly and yearly datasets are organized by variable.

Daily files follow this naming convention: <VARIABLE>_<YEAR>_GLEAM_<ARCHIVE>.nc, where <VARIABLE> corresponds to the variable names listed in the previous section: ‘E’, ‘Et’, ‘Eb’, ‘Ei’, ‘Ew’, ‘Ec’, ‘Es’, ‘Ep’, ‘S’, ‘SMrz’, ‘SMs’, ‘H’, and <YEAR> is the four-digit year. For example, a file containing daily evaporation data for 2023 in the ‘a’ dataset would be named:

v4.2a/daily/2023/E_2023_GLEAM_v4.2a.nc.

Monthly and yearly files follow this naming convention: <VARIABLE>_<YEAR>_GLEAM_<ARCHIVE>_<TEMPORAL_RESOLUTION>.nc, where <TEMPORAL_RESOLUTION> is abbreviated as ‘MO’ for monthly or ‘YR’ for yearly. For monthly and yearly files, flux units are mm month⁻¹ and mm year⁻¹, respectively, instead of mm day⁻¹. For example, a file containing monthly root-zone soil moisture data for 2010 from the ‘b’ archive would be named:

v4.2b/monthly/SMrz/SMrz_2010_GLEAM_v4.2b_MO.nc.

Technical Validation

Global patterns. Figure 2 explores mean global E patterns (1980–2023), along with the absolute and relative contributions from different component fluxes. Unless otherwise noted, GLEAM4 corresponds to the v4.2a data, the latest subversion at the time of writing this manuscript. As expected, E_t dominates the flux globally, especially in densely vegetated humid tropics due to year-round soil water availability and high incoming radiation. The

global proportion of E originating from E_t is 62%, falling within the envelope of current global estimates⁵³, but below the 74% from GLEAM v3³¹, which was on the high end of the spectrum of globally available products. The reduction of E_t from GLEAM v3 to GLEAM4, results from the consideration of short vegetation interception loss and understorey bare soil evaporation in GLEAM4, among other methodological improvements (see Methods). E_t constitutes 14% of the global flux, being larger in forested regions, as expected, while E_p amounts to 17% of global E and is larger in sparsely vegetated regions. Negative estimates in GLEAM4 indicate condensation and are illustrated together with other minor components (i.e., snow and open-water evaporation) in Fig. 2. The mean E in GLEAM4 is $68.5 \times 10^3 \text{ km}^3 \text{ yr}^{-1}$, which agrees with state-of-the-art water cycle appraisals based on extensive literature meta-analysis⁸ ($69.2 \pm 7 \times 10^3 \text{ km}^3 \text{ yr}^{-1}$) and with other global observational datasets (see below).

The seasonal dynamics of some of the key variables from GLEAM4 are portrayed in Fig. 3. The sub-panels showcase global multi-year (1980–2023) mean E , E_p , S , SM_{rz} , and H during boreal summer (June, July, August), and boreal winter (December, January, February). The seasonal pattern of E_p aligns primarily with the cycle of net radiation, while E is additionally determined by the seasonality of SM_{rz} and thus precipitation. Subtropical regions with sufficient JJA precipitation (e.g., India, Northern Australia, parts of Southern Africa, or the east coast of the United States) exhibit the most significant variations in E , with summer E often being an order of magnitude larger than winter levels. In more arid regions, such as central Australia or the Arabian Peninsula, where rainfall occurrences are rare, the seasonal volumes of E remain persistently low throughout the year and unaffected by the E_p cycle. In these areas, the dissipation of available energy primarily occurs through H due to limited SM_{rz} and E . Likewise, E is persistently low in permanent snow regions, despite higher values during the high-radiation season. Overall, these seasonal patterns agree with relevant literature^{26,54–57}.

Understanding the dependency of E on different driving factors can provide crucial insights into the seasonal behaviour of E in specific regions, and potentially into the main controls on long-term E trends⁵⁸. Figure 4 provides an overview of these driving factors, leveraging from the partitioning of potential evaporation into an aerodynamic and a radiative term in Penman's combination equation (see Eq. 1), and taking advantage of the separate calculation of evaporative stress in GLEAM4 (see Methods). Red tones in semiarid regions indicate the dominance of evaporative deficit ($E - E_p$); in these regions, precipitation supply is insufficient to satisfy the high atmospheric demand for water. On the other hand, in temperate and boreal forests, precipitation supply is sufficient to meet the atmospheric demand; in particular, the aerodynamic component of E_p (which depends on wind, turbulence, ecosystem height and VPD) shows a higher relevance, as shown by the green tones. In the tropics, E_p is primarily satisfied by the radiative component of E (blue tones), which is high due to the high incoming radiation and low albedo of rainforests. Globally, the atmosphere demand for water, or E_p , adds up to $151 \times 10^3 \text{ km}^3 \text{ yr}^{-1}$, of which $\sim 55\%$ is unsatisfied (evaporative deficit), and the remaining is satisfied by E through radiative ($\sim 31\%$) and aerodynamic ($\sim 14\%$) processes.

Validation and inter-comparison. Figure 5 shows a comparison of GLEAM4 E against frequently used global E datasets, including its immediate predecessor GLEAM v3.8a³¹, the ERA5-Land reanalysis⁵⁹, and FLUXCOM (RS-METEO)⁵⁶. Long-term means for all datasets portray similar geographical patterns, with GLEAM4 showing greater agreement with ERA5-Land and FLUXCOM in the tropics compared to GLEAM v3.8a, largely due to a decrease in transpiration estimates over rainforests (as seen in the comparison between Fig. 2 and the results in ref. 31). Regional differences indicate relatively high values of GLEAM4 compared to other products in temperate and boreal forests in the Northern Hemisphere, where the consideration of the aerodynamic term of Penman's Eq. (1) in the new version results in higher E_p and subsequently higher E than in GLEAM v3.8a. Relatively low estimates by GLEAM4 are concentrated in semiarid ecosystems — such as western United States, southern Africa or the Mediterranean region — especially when compared to FLUXCOM and ERA5-Land. This reflects the fact that the evaporative stress in GLEAM4 under water-limited conditions is greater than for GLEAM v3.8a, regardless of the generally larger atmospheric demand for water in the former (as seen in the comparison between E_p in Fig. 3 and the results in ref. 55).

The global mean estimates of E from the four datasets are comparable, ranging from highest to lowest: 72.8, 71.8, 68.5 and $67.7 \times 10^3 \text{ km}^3 \text{ yr}^{-1}$, for ERA5-Land, GLEAM v3.8a, GLEAM4 and FLUXCOM, respectively. As indicated above, these estimates fall within the range of a recent meta-analysis⁸ that reported $69.2 \pm 7 \times 10^3 \text{ km}^3 \text{ yr}^{-1}$.

At regional scales, the patterns of GLEAM4 appear realistic and highlight the value of transitioning to higher spatial resolutions to better capture the influence of complex topography and land use changes. Figure 6 compares the estimates from GLEAM4 to those from the other three datasets during two of the most significant summer droughts in the historical record: the 1988 North American drought⁶⁰ and the 2003 European drought⁶¹, both of which were compounded by severe heatwaves. During drought and heatwave events, E tends to exhibit positive anomalies in the early stages, as long as soil moisture remains sufficiently available, due to the high atmospheric demand for water (E_p). However, as these events progress and soil moisture depletion leads to increased evaporative stress (S), anomalies typically become negative, triggering feedback mechanisms that can further intensify the events². Whether anomalies are overall positive or negative when integrated across the event largely depends on initial soil moisture conditions and the duration and severity of the event⁷. The four datasets evaluated in Fig. 6 appear to capture this complex interplay between water supply and demand, showing good agreement in the regional distribution of positive and negative anomalies. The coarser resolution of FLUXCOM is evident, and so are its reported difficulties in capturing the magnitude of anomalies. However, the recently released next generation of FLUXCOM datasets employs higher resolution and may offer improved capabilities for capturing anomalies during such events¹⁷.

To evaluate the skill of GLEAM4 in capturing temporal dynamics at the ecosystem scale, E estimates are validated against *in situ* eddy-covariance data from a wide range of global networks, including FLUXNET La Thuile, FLUXNET2015, FLUXNET-CH4, AmeriFlux, ICOS, and EFDC²¹. For duplicate stations across sources, the longest record was retained. Subsequently, sites with fewer than 250 days were excluded. This yielded a

final sample of 473 sites and 2511 years of data. Figure 7 compares the overall performance of the four datasets GLEAM4 and the other three datasets (i.e., GLEAM v3.8a, ERA5-Land, and FLUXCOM) in simulating E . Results are illustrated through Taylor density diagrams and violin plots of Kling-Gupta Efficiency (KGE), both displaying the distribution of the validation metrics that are calculated per site. The normalised standard deviation (std) in the Taylor diagrams indicates a mild tendency of all datasets (except for ERA5-Land) to underestimate the variability of E times series, which could relate to their large pixel coverage compared to the more reduced tower footprint. Median root-mean-square error (RMSE) and Pearson's correlations (R) are similar for all datasets, but slightly better for FLUXCOM (0.89 mm d^{-1} , 0.77) than for GLEAM4 (0.95 mm d^{-1} , 0.73) and ERA5-Land (1.0 mm d^{-1} , 0.77). Meanwhile, KGE values — integrating correlation, variability and bias — show higher median values (and thus better performance) for GLEAM4 than for ERA5-Land (0.49 vs. 0.45), and a slight improvement upon GLEAM v3.8a (0.48).

Figure 8 zooms into example time series at four specific sites. The sites correspond to the Tharandt spruce forest in Eastern Germany (DE-Tha), a rainfed maize-soybean rotation site in Nebraska (US-Ne3), the Australian open woodland savanna site in Howard Springs (AU-How), and the French evergreen Mediterranean forest in Puechabon (FR-Pue). These sites were selected based on their long records (>10 years) which allow the reliable computation of seasonal climatologies. Years were selected based on extreme conditions: droughts for FR-Pue⁶² and US-Ne3⁶³, heatwaves for DE-Tha⁶⁴, and pluvial events for AU-How⁶⁵. The left time series indicate anomalies in E , computed by subtracting the seasonal climatology of the corresponding dataset, using a multiyear mean for each calendar day and a 31-day moving average to compute that climatology⁶⁶. Time series of anomalies show significant fluctuations across seasons and climatic events, and often differences among the different datasets, but also remarkable similarities among them and when compared to the *in situ* data. This pattern is observed at all four stations. While simulating correctly the seasonal cycles may in principle seem like a trivial task compared to simulating temporal anomalies accurately, the smaller panels on the right indicate that the different products still struggle to simulate the seasonality of E . This could relate to differences in land cover from the tower footprint to the coarser-resolution pixels, and it is also influenced by biases in the input data used by each of the models. Nonetheless, all datasets capture the timing of the seasonal cycle, but overestimate its amplitude in DE-Tha and FR-Pue, while they show a larger divergence in general and a tendency to underestimate in both US-Ne3 and AU-How. Overall, GLEAM4 aligns well with observed seasonal cycles, and despite the heterogeneous performance across sites, it reproduces E anomalies across diverse climates and ecosystems successfully, even during extreme events like the ones depicted in Fig. 8.

Code availability

MATLAB and Python code for synthesizing the results and generating all figures of the article can be obtained from the public repository at <https://doi.org/10.5281/zenodo.14056593>⁶⁷. Python code used to generate the GLEAM4 products can be requested via info@gleam.eu for reproducibility purposes. However, the code has been tailored to the computing and storage system of the Flemish Supercomputer Center (VSC) and is therefore not suitable for use in other environments.

Received: 19 November 2024; Accepted: 11 February 2025;

Published online: 10 March 2025

References

- Miralles, D. G., Brutsaert, W., Dolman, A. J. & Gash, J. H. On the use of the term “Evapotranspiration”. *Water Resour. Res.* **56**, e2020WR028055, <https://doi.org/10.1029/2020wr028055> (2020).
- Miralles, D. G., Gentile, P., Seneviratne, S. I. & Teuling, A. J. Land-atmospheric feedbacks during droughts and heatwaves: State of the science and current challenges. *Annals New York Acad. Sci.* **1436**, 19–35, <https://doi.org/10.1111/nyas.13912> (2019).
- Teuling, A. J. *et al.* Evapotranspiration amplifies European summer drought. *Geophys. Res. Lett.* **40**, 2071–2075, <https://doi.org/10.1002/grl.50495> (2013).
- Brocca, L. *et al.* A Digital Twin of the terrestrial water cycle: a glimpse into the future through high-resolution Earth observations. *Front. Sci.* **1**, 1190191, <https://doi.org/10.3389/fsci.2023.1190191> (2024).
- Miralles, D. G. *et al.* El Niño–La Niña cycle and recent trends in continental evaporation. *Nat. Clim. Chang.* **4**, 122–126, <https://doi.org/10.1038/nclimate2068> (2014).
- Vicente-Serrano, S. M. *et al.* The uncertain role of rising atmospheric CO₂ on global plant transpiration. *Earth-Science Rev.* **230**, 104055, <https://doi.org/10.1016/j.earscirev.2022.104055> (2022).
- Zhao, M., A. G., Liu, Y. & Konings, A. G. Evapotranspiration frequently increases during droughts. *Nat. Clim. Chang.* **1–7**, <https://doi.org/10.1038/s41558-022-01505-3> (2022).
- Dorigo, W. *et al.* Closing the water cycle from observations across scales: Where do we stand? *Bull. Am. Meteorol. Soc.* **1–95**, <https://doi.org/10.1175/bams-d-19-0316.1> (2021).
- Ezenne, G. I., Eyibio, N. U., Tanner, J. L., Asoiro, F. U. & Obalum, S. E. An overview of uncertainties in evapotranspiration estimation techniques. *J. Agrometeorology* **25**, <https://doi.org/10.54386/jam.v25i1.2014> (2023).
- Miralles, D. G. *et al.* Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci.* **15**, 453–469, <https://doi.org/10.5194/hess-15-453-2011> (2011).
- Fisher, J. B., Tu, K. P. & Baldocchi, D. D. Global estimates of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. *Remote Sens. Environ.* **112**, 901–919 (2008).
- Mu, Q., Zhao, M. & Running, S. W. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* **115**, 1781–1800, <https://doi.org/10.1016/j.rse.2011.02.019> (2011).
- Penman, H. L. Natural evaporation from open water, bare soil and grass. *Proc. Royal Society of London Series A- Mathematical and Physical Sciences* **193**, 120–145 (1948).
- Priestley, C. H. B. & Taylor, R. J. On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* **100**, 81–92 (1972).
- Monteith, J. L. Evaporation and environment. *Symp. Soc. for Exp. Biol.* **19**, 4 (1965).
- Jung, M., Reichstein, M. & Bondeau, A. Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences* **6**, 2001–2013 (2009).

17. Nelson, J. A. *et al.* X-BASE: The first terrestrial carbon and water flux products from an extended data-driven scaling framework. *FLUXCOM-X. EGU Sphere* **2024**, 1–51, <https://doi.org/10.5194/egusphere-2024-165> (2024).
18. Reichstein, M. *et al.* Deep learning and process understanding for data-driven earth system science. *Nature* **566**, 195–204, <https://doi.org/10.1038/s41586-019-0912-1> (2019).
19. Zhao, W. L. *et al.* Physics-constrained machine learning of evapotranspiration. *Geophys. Res. Lett.* **46**, 14496–14507, <https://doi.org/10.1029/2019gl085291> (2019).
20. Kraft, B., Jung, M., Körner, M., Koira, S. & Reichstein, M. Towards hybrid modeling of the global hydrological cycle. *Hydrol. Earth Syst. Sci. Discuss.* **2021**, 1–40, <https://doi.org/10.5194/hess-2021-211> (2021).
21. Koppa, A., Rains, D., Hulsman, P., Poyatos, R. & Miralles, D. G. A deep learning-based hybrid model of global terrestrial evaporation. *Nat. Commun.* **13**, 1912, <https://doi.org/10.1038/s41467-022-29543-7> (2022).
22. ElGhawi, R. *et al.* Hybrid modeling of evapotranspiration: Inferring stomatal and aerodynamic resistances using combined physics-based and machine learning. *Environ. Res. Lett.* **18**, 034039, <https://doi.org/10.1088/1748-9326/acbbe0> (2023).
23. Jahromi, M. N. *et al.* Ten years of GLEAM: A review of scientific advances and applications. In Bozorg-Haddad, O. & Zolghadr-Asli, B. (eds.) *Computational Intelligence for Water and Environmental Sciences*, 525–540, https://doi.org/10.1007/978-981-19-2519-1_25 (Springer Nature Singapore, Singapore, 2022).
24. Michel, D. *et al.* The WACMOS-ET project – Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms. *Hydrol. Earth Syst. Sci.* **20**, 803–822, <https://doi.org/10.5194/hess-20-803-2016> (2016).
25. Talsma, C. J. *et al.* Partitioning of evapotranspiration in remote sensing-based models. *Agric. For. Meteorol.* **260**, 131–143, <https://doi.org/10.1016/j.agrformet.2018.05.010> (2018).
26. Miralles, D. G. *et al.* The WACMOS-ET project – part 2: Evaluation of global terrestrial evaporation data sets. *Hydrol. Earth Syst. Sci.* **20**, 823–842, <https://doi.org/10.5194/hess-20-823-2016> (2016).
27. McCabe, M. F. *et al.* The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally gridded forcing data. *Geosci. Model. Dev.* **9**, 283–305, <https://doi.org/10.5194/gmd-9-283-2016> (2016).
28. Hulsman, P., Keune, J., Koppa, A., Schellekens, J. & Miralles, D. G. Incorporating plant access to groundwater in existing global, satellite-based evaporation estimates. *Water Resour. Res.* **59**, <https://doi.org/10.1029/2022WR033731> (2023).
29. Zhong, F. *et al.* Revisiting large-scale interception patterns constrained by a synthesis of global experimental data. *Hydrol. Earth Syst. Sci.* **26**, 5647–5667, <https://doi.org/10.5194/hess-26-5647-2022> (2022).
30. van Dijk, A. I. J. M. & Bruijnzeel, L. A. Modelling rainfall interception by vegetation of variable density using an adapted analytical model. part 1. model description. *J. Hydrol.* **247**, 230–238, [https://doi.org/10.1016/S0022-1694\(01\)00392-4](https://doi.org/10.1016/S0022-1694(01)00392-4) (2001).
31. Martens, B. *et al.* GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geosci. Model. Dev.* **10**, 1903–1925, <https://doi.org/10.5194/gmd-10-1903-2017> (2017).
32. Miralles, D. G., Gash, J. H., Holmes, T. R. H., de Jeu, R. A. M. & Dolman, A. J. Global canopy interception from satellite observations. *J. Geophys. Res. Atmospheres* **115**, D16122, <https://doi.org/10.1029/2009jd013530> (2010).
33. Gash, J. H. C. & Morton, A. J. An application of the Rutter model to the estimation of the interception loss from Thetford Forest. *J. Hydrol.* **38**, 49–58, [https://doi.org/10.1016/0022-1694\(78\)90131-2](https://doi.org/10.1016/0022-1694(78)90131-2) (1978).
34. van Dijk, A. I. J. M. & Bruijnzeel, L. A. Modelling rainfall interception by vegetation of variable density using an adapted analytical model. part 2. model validation for a tropical upland mixed cropping system. *J. Hydrol.* **247**, 239–262, [https://doi.org/10.1016/S0022-1694\(01\)00393-6](https://doi.org/10.1016/S0022-1694(01)00393-6) (2001).
35. Zhong, F. *et al.* Multi-Decadal Dynamics of Global Rainfall Interception and Their Drivers. *Geophys. Res. Lett.* **51**, <https://doi.org/10.1029/2024gl109295> (2024).
36. Potapov, P. *et al.* Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote. Sens. Environ.* **253**, 112165, <https://doi.org/10.1016/j.rse.2020.112165> (2021).
37. Beck, H. E. *et al.* MSWX: Global 3-hourly 0.1° bias-corrected meteorological data including near-real-time updates and forecast ensembles. *Bull. Am. Meteorol. Soc.* **103**, E710–E732, <https://doi.org/10.1175/bams-d-21-0145.1> (2022).
38. Aumann, H. H. *et al.* AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Transactions on Geosci. Remote. Sens.* **41**, 253–264, <https://doi.org/10.1109/TGRS.2002.808356> (2003).
39. Allen, R. G., Tasumi, M. & Trezza, R. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—model. *J. Irrigation Drainage Eng.* **133**, 380–394, [10.1061/\(asce\)0733-9437\(2007\)133:4\(380\)](https://doi.org/10.1061/(asce)0733-9437(2007)133:4(380)) (2007).
40. Thom, A. S. Momentum, mass and heat exchange of plant communities. *Q. J. Royal Meteorol. Soc.* **18** (1975).
41. Shuttleworth, W. J. *Whole-Canopy Interactions*, chap. **22**, 316–333 <https://doi.org/10.1002/9781119951933.ch22> (John Wiley & Sons, Ltd, 2012).
42. Garrat, J. R. *Surface roughness and local advection*, chap. **4**, 85–114 (Cambridge University Press, 1994).
43. Riggden, A. J. & Salvucci, G. D. Evapotranspiration based on equilibrated relative humidity (ETRHEQ): Evaluation over the continental U.S. *Water Resour. Res.* **51**, 2951–2973, <https://doi.org/10.1002/2014WR016072> (2015).
44. Rains, D. *et al.* Sentinel-1 backscatter assimilation using support vector regression or the water cloud model at european soil moisture sites. *IEEE Geosci. Remote. Sens. Lett.* **PP**, 1–5, <https://doi.org/10.1109/lgrs.2021.3073484> (2021).
45. Lievens, H. *et al.* Assimilation of global radar backscatter and radiometer brightness temperature observations to improve soil moisture and land evaporation estimates. *Remote. Sens. Environ.* **189**, 194–210, <https://doi.org/10.1016/j.rse.2016.11.022> (2017).
46. Maxwell, R. M. & Condon, L. E. Connections between groundwater flow and transpiration partitioning. *Science* **353**, 377–380, <https://doi.org/10.1126/science.aaf7891> (2016).
47. Haghdoost, S., Koppa, A., Lievens, H. & Miralles, D. G. Improving global evaporation estimation using GRACE and GRACE-FO satellite data assimilation. In *EGU General Assembly 2024, Vienna, Austria*, EGU24-17588, <https://doi.org/10.5194/egusphere-egu24-17588> (2024).
48. Dorigo, W. *et al.* ESA CCI Soil Moisture for improved earth system understanding: State-of-the art and future directions. *Remote. Sens. Environ.* **203**, 185–215, <https://doi.org/10.1016/j.rse.2017.07.001> (2017).
49. Miralles, D. G. *et al.* GLEAM4 <https://doi.org/10.5281/zenodo.14056079> (2024).
50. Beck, H. E. *et al.* MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull. Am. Meteorol. Soc.* **100**, 473–500, <https://doi.org/10.1175/bams-d-17-0138.1> (2018).
51. Wielicki, B. A. *et al.* Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment. *Bull. Am. Meteorol. Soc.* **77**, 853–868, [10.1175/1520-0477\(1996\)077<0853:CATERE>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0853:CATERE>2.0.CO;2) (1996).
52. Hou, A. Y. *et al.* The Global Precipitation Measurement Mission. *Bull. Am. Meteorol. Soc.* **95**, 701–722, <https://doi.org/10.1175/bams-d-13-00164.1> (2014).
53. Wei, Z. *et al.* Revisiting the contribution of transpiration to global terrestrial evapotranspiration. *Geophys. Res. Lett.* **44**, 2792–2801, <https://doi.org/10.1002/2016gl072235> (2017).
54. Miralles, D. G., de Jeu, R. A. M., Gash, J. H., Holmes, T. R. H. & Dolman, A. J. Magnitude and variability of land evaporation and its components at the global scale. *Hydrol. Earth Syst. Sci.* **15**, 967–981, <https://doi.org/10.5194/hess-15-967-2011> (2011).
55. Singer, M. B. *et al.* Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981–present. *Sci. Data* **8**, 224, <https://doi.org/10.1038/s41597-021-01003-9> (2021).
56. Jung, M. *et al.* The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Sci. Data* **6**, 74, <https://doi.org/10.1038/s41597-019-0076-8> (2019).

57. Zhang, K. A global dataset of terrestrial evapotranspiration and soil moisture dynamics from 1982 to 2020. *Scientific Data* **11**, 445 (2024).
58. Teuling, A. J. *et al.* A regional perspective on trends in continental evaporation. *Geophys. Res. Lett.* **36**, <https://doi.org/10.1029/2008gl036584> (2009).
59. Muñoz-Sabater, J. *et al.* ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **13**, 4349–4383, <https://doi.org/10.5194/essd-13-4349-2021> (2021).
60. Trenberth, K. E., Branstator, G. W. & Arkin, P. A. Origins of the 1988 North American Drought. *Science* **242**, 1640–1645, <https://doi.org/10.1126/science.242.4886.1640> (1988).
61. García-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J. & Fischer, E. M. A Review of the European Summer Heat Wave of 2003. *Critical Rev. Environ. Sci. Technol.* **40**, 267–306, <https://doi.org/10.1080/10643380802238137> (2010).
62. García-Herrera, R. *et al.* The European 2016/17 drought. *J. Clim.* **32**, 3169–3187, <https://doi.org/10.1175/JCLI-D-18-0331.1> (2019).
63. Kerr, E. Brutal drought depresses agriculture, thwarting US and Texas economies. *Southwest Econ.* 10–13 (2012).
64. Ciais, P. *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* **437**, 529–533, <https://doi.org/10.1038/nature03972> (2005).
65. Australian Government Bureau of Meteorology. Annual climate summary 2010. Tech. Rep., Australian Government Bureau of Meteorology (2011).
66. Miralles, D. G., Crow, W. T. & Cosh, M. H. Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations. *J. Hydrometeorol.* **11**, 1423–1429, <https://doi.org/10.1175/2010JHM1285.1> (2010).
67. Miralles, D. G. *et al.* h-cel/GLEAM4: First submission. *Zenodo [Code]* <https://doi.org/10.5281/zenodo.14056593> (2024).
68. NASA/LARC/SD/ASDC. CERES and GEO-Enhanced TOA, Within-Atmosphere and Surface Fluxes, Clouds and Aerosols 1-Hourly Terra-Aqua Edition4A, https://doi.org/10.5067/TERRA+AQUA/CERES/SYN1DEG-1HOUR_L3.004A (2017).
69. Goddard Earth Sciences Data and Information Services Center. Aqua/AIRS L3 Daily Standard Physical Retrieval (AIRS-only) 1 degree × 1 degree V7.0, <https://doi.org/10.5067/UO3Q64CTTS1U> (2019).
70. Huffman, G., Stocker, E., Bolvin, D., Nelkin, E. & Tan, J. GPM IMERG Final Precipitation L3 1 day 0.1 degree × 0.1 degree V07, <https://doi.org/10.5067/GPM/IMERGDF/DAY/07> (2023).
71. Hersbach, H. *et al.* The ERA5 global reanalysis. *Q. J. Royal Meteorol. Soc.* **146**, 1999–2049, <https://doi.org/10.1002/qj.3803> (2020).
72. Hersbach, H. *et al.* ERA5 hourly data on single levels from 1940 to present, <https://doi.org/10.24381/cds.adbb2d47> (2023).
73. Inness, A. *et al.* The CAMS reanalysis of atmospheric composition. *Atmospheric Chem. Phys.* **19**, 3515–3556, <https://doi.org/10.5194/acp-19-3515-2019> (2019).
74. Copernicus Atmosphere Monitoring Service Atmosphere Data Store. CAMS global greenhouse gas reanalysis (EGG4), <https://doi.org/10.24380/8fck-9w87> (2021).
75. Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J. & Moisander, M. *GlobSnow v3.0 snow water equivalent (SWE)* <https://doi.org/10.1594/PANGAEA.911944> (2020).
76. Armstrong, R., Brodzik, M. J., Knowles, K. & Savoie, M. *Global monthly EASE-Grid snow water equivalent climatology, version 1* <https://doi.org/10.5067/KJVERY3MIBPS> (2005).
77. Dorigo, W. *et al.* Soil moisture gridded data from 1978 to present, v201706.0.0., <https://doi.org/10.24381/cds.d7782f18> (2017).
78. Zotta, R.-M. *et al.* VODCA v2: Multi-sensor, multi-frequency vegetation optical depth data for long-term canopy dynamics and biomass monitoring. *Earth Syst. Sci. Data* **16**, 4573–4617, <https://doi.org/10.5194/essd-16-4573-2024> (2024).
79. Moesinger, L. *et al.* The Global Long-term Microwave Vegetation Optical Depth Climate Archive VODCA, <https://doi.org/10.5281/zenodo.2575599> (2019).
80. Myneni, R., Knyazikhin, Y. & Park, T. Mcd15a3h modis/terra+aqua leaf area index/fpar 4-day l4 global 500m sin grid v006 [data set]. *nasa eosdis land processes distributed active archive center*. accessed 2024-11-19. <https://doi.org/10.5067/MODIS/MCD15A3H.006>.
81. Hansen, M. & Song, X. Vegetation continuous fields (VCF) yearly global 0.05 deg. 2018, distributed by NASA EOSDIS Land Processes Distributed Active Archive Center, <https://doi.org/10.5067/MEaSURES/VCF/VCF5KYR.001> (2017).
82. DiMiceli, C. *et al.* MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006, distributed by NASA EOSDIS Land Processes Distributed Active Archive Center <https://doi.org/10.5067/MODIS/MOD44B.006> (2015).
83. Simons, G., Koster, R. & Droogers, P. Hihydrosoil v2.0-high resolution soil maps of global hydraulic properties. *Futur. Work. Available from* <https://www.futurewater.eu/projects/hihydrosoil> (2020).

Acknowledgements

GLEAM4 has been developed in recent years thanks to recurrent funding from the Belgian Science Policy Office (BELSPO) STEREO program (SR/02/402, SR/02/377, SR/00/373) and the European Space Agency (ESA) (4000129870/20/I-NB, 4000136272/21/I-EF, 000136272/21/I-EF). DGM acknowledges the support of the European Research Council (ERC) via the HEAT Consolidator grant (101088405). We are grateful to Martin Jung and Ulrich Weber for providing FLUXCOM data for this study, and to all the ground data providers. The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation, Flanders (FWO), and the Flemish Government.

Author contributions

D.G.M. has led the development of GLEAM as P.I. since 2009. F.Z., D.G.M., P.H. and A.K. coordinated the development of the new versions of the interception, potential evaporation, soil water, and evaporative stress modules, respectively. D.G.M., A.K., O.B., O.B.-V., E.T., P.H., F.Z. and S.H. contributed to the development of the code and decisions regarding the algorithm and dataset structure. O.B. streamlined the model code. O.B.-V., A.K. and F.Z. collected and pre-processed the forcing data. H.E.B. developed and provided the meteorological forcing. W.D. developed and provided the VOD and surface soil moisture input data. A.K. and O.B.-V. executed the model. E.T. led the validation against *in situ* data. D.G.M. analysed the output, prepared the figures, and wrote the first draft. All authors contributed to the discussion and interpretation of the results, as well as the editing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.G.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025