

Supplement of

State-of-the-art hydrological datasets exhibit low water balance consistency globally

Huang et al.

Correspondence to: Junguo Liu (junguo.liu@gmail.com)

Text S1. Performance calculations at daily resolution

Before the performance calculations, we selected the precipitation (P), evapotranspiration (ET), runoff (R), and soil moisture (SM) datasets that originally provide daily or eight-daily data. There remain

- 12 P datasets (i.e., CPC, EM-EARTH, CHIRPS v2.0, CMAP, CMORPH v1, GPCP(D) v1.3, GPM IMERG v07, PERSIANN-CDR, MSWEP v2.8, 20CR v3, MERRA-2, and ERA5),
- 10 ET datasets (i.e., X-BASE, MODIS, PML-v2, GLASS, GLEAM v4.1, MERRA-2, GLDAS-2.0, GLDAS-2.1, GLDAS-2.2, and ERA5-land),
- 6 R datasets (i.e., SMAP L4 v7, MERRA-2, GLDAS-2.0, GLDAS-2.1, GLDAS-2.2, and ERA5-land), and
- 9 SM datasets (i.e., SoMo.ml, SMAP L4 v7, CCI v08.1, GLEAM v4.1, MERRA-2, GLDAS-2.0, GLDAS-2.1, GLDAS-2.2, and ERA5-land)

for further calculations. We applied the same exclusion rules detailed in Methods, and obtained 3,647 independent combinations. Consistent with the calculations based on monthly P , ET , R , and changes in SM (ΔSM), the model of equation (1) was built in each grid cell of each combination, while the ΔSM values at each daily timestep are now the differences between the next day and this day. By calculating the adjusted R^2 score for each daily-based model, we follow the same way as monthly analysis to obtain the dataset performance.

Text S2. Performance calculations at yearly resolution

Unlike Text S1, the performance calculations on a yearly basis are almost the same as monthly since all the datasets considered in the main text were also considered here. The only differences were using yearly amounts of P , ET , and R , as well as yearly differences of SM between the last day and the first day of each year in equation (1).

Table S1. Information on precipitation (P) datasets used in the study. Colors of the abbreviation are tied to Fig. 1, namely orange for gauge-based, green for satellite-based, and blue for reanalysis datasets.

Product	Abbreviation	Temporal resolution	Characteristics
NOAA Climate Prediction Center Unified Gauge-based Analysis of Global Daily Precipitation	CPC	Jan. 2000 to Dec. 2022	CPC utilizes the gauge-observed P reports transmitted through the World Meteorological Organization's (WMO) Global Telecommunication System (GTS) (Xie et al., 2010). The GTS gauge network is relatively dense over the United States, Western Europe, and the east coasts of Australia and China, while it is very sparse over several regions including like equatorial Africa (Fig. S7a).
Climatic Research Unit gridded Time Series v4.06	CRU TS v4.06	Jan. 2000 to Dec. 2022	CRU TS v4.06 has several principal monthly sources for P : information through the GTS, Monthly Climatic Data for the World (MCDW) summaries, and adhoc collections of stations (Harris et al., 2020).
Terrestrial Precipitation: 1900-2017 Gridded Monthly Time Series v5.01 from University of Delaware	UDel v5.01	Jan. 2000 to Dec. 2017	Station data for UDel v5.01 was compiled from several updated sources including a recent version of the Global Historical Climatology Network dataset (GHCN2), a version of the Daily Global Historical Climatology Network (GHCN-Daily), an Atmospheric Environment Service/Environment Canada archive, data from the Hydrometeorological Institute in Russia, GC-Net data, Greenland station records from the Automatic Weather Station Project, the National Center for Atmospheric Research (NCAR) daily India data, Sharon Nicholson's archive of African P data, Webber and Webber and Willmott's South American monthly P station records; and daily records from the Global Surface Summary of Day (GSOD) (https://psl.noaa.gov/data/gridded/data.UDel_AirT_Precip.html).
the Ensemble	EM-EARTH	Jan. 2000 to	The two major inputs of EM-Earth are a

Meteorological Dataset for Planet Earth		Dec. 2019	station-based Serially Complete Earth (SC-Earth) and ERA5, where the SC-Earth used station data from GHCN-Daily and GSOD (Tang et al., 2022).
Global Precipitation Climatology Centre Full Data Monthly Product Version 2022	GPCC v2022	Jan. 2000 to Dec. 2020	GPCC v2022 is based on GTS reports and other available global and regional collections like GHCN, CRU, Nicholson's African data, etc. (Schneider et al., 2022).
Precipitation Reconstruction over Land	PREC/L	Jan. 2000 to Dec. 2022	PREC/L is based on gauge observations from GHCN2 and the Climate Anomaly Monitoring System (CAMS) dataset (Chen et al., 2002).
Climate Hazards Group InfraRed Precipitation with Station data	CHIRPS v2.0	Jan. 2000 to Dec. 2022	CHIRPS v2.0 uses the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis version 7 (TMPA 3B42 v7) to calibrate its Cold Cloud Duration (CCD)-based rainfall estimates, which is then combined with GTS gauge data to produce the product (Funk et al., 2015).
CPC Merged Analysis of Precipitation	CMAP	Jan. 2000 to Dec. 2022	The standard version of CMAP was used in this study, which was produced by merging gauge observations from GPCC and estimates inferred from satellite observations, including infrared (IR)-based GOES Precipitation Index, outgoing longwave radiation-based Precipitation Index, Microwave Sounding Unit-based Spencer, and microwave scattering- and emission-based products from the Special Sensor Microwave/Imager (SSM/I) (Xie and Arkin, 1997).
Bias-Corrected CPC Morphing technique Climate Data Record	CMORPH v1	Jan. 2000 to Dec. 2022	The raw CMORPH is constructed through integrating P information from multiple passive microwave (PMW) sensors aboard low-Earth-orbiting satellites. Then, the biases in the raw CMORPH P estimates were reduced by calibrating against the GPCC monthly gauge analysis (Xie et al., 2017).
Global Precipitation Climatology	GPCP(M) v2.3	Jan. 2000 to Dec. 2022	GPCP(M) v2.3 is a merger of various satellite-based estimates and the P gauge analyses over land from the

Project Monthly
Analysis v2.3

Global
Precipitation
Climatology
Project Climate
Data Record
(CDR) -Daily
Analysis, Version
1.3

GPCP(D) v1.3

Jan. 2000 to
Dec. 2022

GPCC. Over global land, the satellite-based estimates are a combination of PMW estimates and estimates from IR/microwave sounders, including Special Sensor Microwave Imager/Sounder (SSMIS; different from SSM/I) and the Atmospheric Infrared Sounder (AIRS)/Advanced Microwave Sounding Unit flying on NASA's Aqua satellite (Adler et al., 2018).

The GPCP daily product blends data from polar-orbit PMW satellites (SSM/I, SSMIS), polar orbit IR sounders (TIROS Operational Vertical Sounder (TOVS), AIRS), geostationary IR satellites (such as GOES) and the GPCP monthly analysis (Huffman et al., 2001).

The Integrated
Multi-satelliE
Retrievals for the
Global
Precipitation
Measurement
mission

GPM IMERG
v07

Jun. 2000 to
Dec. 2022

The GPM IMERG intended to intercalibrate, merge, and interpolate microwave P estimates from TRMM (used in CHIRPS v2.0) and GPM satellites, together with microwave-calibrated IR estimates, and potentially other P estimators at fine time and space scales over the entire globe. Compared to TRMM, the GPM can increase orbital inclination to afford more coverage and enhance the sensitivity to light and solid P . Its bias correction technique was adopted from CMORPH, and gauge analysis from CPC and GPCC v2022 was used (Huffman et al., 2023).

Precipitation
Estimation from
Remotely Sensed
Information using
Artificial Neural
Networks-CDR

PERSIANN-
CDR

Jan. 2000 to
Dec. 2022

PERSIANN-CDR developed a different algorithm from the aforementioned products. That is, they trained artificial neural networks with NCEP stage IV hourly P data to obtain nonlinear regression parameters, and then ran the model with Gridded Satellite (GridSat)-B1 IR satellite data, eliminating the need for PMW observations. The final product was adjusted by using bias correction methods similar to CMORPH and GPM IMERG v07, as

Multi-Source Weighted-Ensemble Precipitation, version 2	MSWEP v2.8	Jan. 2000 to Dec. 2022	well as the gauge information from GPCC (Ashouri et al., 2015). MSWEP v2.8 was produced by merging gauge observations from various sources (GSOD, GHCN-D, and several national databases), two satellite-based <i>P</i> products (GPM IMERG and GridSat), and ERA5 (shown in the following) (see its technical documentation: https://www.dropbox.com/s/5r4nnicfe3ft12d/MSWEP_V2_doc.pdf?dl=1).
NOAA-CIRES-DOE Twentieth Century Reanalysis project	20CR v3	Jan. 2000 to Dec. 2015	The 20CR v3 used modern weather forecast models to generate the atmospheric background fields given prescribed sea surface temperature and sea ice concentration fields, and then used an 80-member ensemble Kalman filter assimilates historical observations (on surface pressure values) to update the background fields, yielding analysis fields (Slivinski et al., 2019).
Japanese 55-year Reanalysis	JRA-55	Jan. 2000 to Dec. 2022	The JRA-55 applied a four-dimensional variational data assimilation (4D-Var), which attempts to obtain the best estimate of the ocean state over a finite time interval (assimilation window) by using all available observations and a numerical model to dynamically interpolate information in space and time. The observational data adopted for JRA-55 were primarily those used in ERA-40 in addition to information archived by Japan Meteorological Agency (https://jra.kishou.go.jp/JRA-55/leaflet/JRA-55_leaflet_display.pdf).
The second Modern-Era Retrospective analysis for Research and Applications	MERRA-2	Jan. 2000 to Dec. 2022	MERRA-2 also provided observation-corrected model-generated <i>P</i> by applying the Goddard Earth Observing System Model, Version 5 (GEOS-5) data assimilation system, where GPCP v2.,1 CPC, and CMAP were used (Reichle and Liu, 2014). As a result, the land surface in MERRA-2 was forced primarily by the corrected estimates at low to mid latitudes, by the MERRA-2

the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis	ERA5	Jan. 2000 to Dec. 2022	model-generated P at high latitudes, and by a weighted mixture of corrected and model-generated estimates for the regions in between to prevent spatial discontinuities in climatological means (Gelaro et al., 2017). ERA5 also used 4D-Var, similarly to JRA-55, to assimilate satellite, P radar, and gauge measurements with the modeled P from a numerical weather prediction (NWP) model (Lavers et al., 2022; Hersbach et al., 2020).
NCEP/NCAR Reanalysis Project	NCEP-NCAR R1	Jan. 2000 to Dec. 2022	The central module in NCEP-NCAR R1 is the spectral statistical interpolation, a three-dimensional variational (3D-Var) data assimilation which was earlier developed than 4D-Var. The assembly of observational databases include global rawinsonde data, surface marine data, aircraft data, surface land synoptic data, satellite sounder data, SSM/I surface wind speeds, and satellite cloud drift winds (Kistler et al., 2001).
NCEP-DOE Reanalysis 2 project	NCEP-DOE R2	Jan. 2000 to Dec. 2022	NCEP-DOE R2 should not be considered as a next-generation reanalysis of NCEP-NCAR R1, but should be regarded as an updated and human error-fixed version. They used similar raw observational data, but R2 has new components like simple rainfall assimilation (with CMAP pentad P) over land surfaces for improved soil wetness, smoothed orography (to prevent Gibbs phenomena-like P, especially over the Amazon basin, and sensible and latent heat fluxes over ocean near steep orography), and treatment of snow (Kanamitsu et al., 2002).

Table S2. Information of evapotranspiration (*ET*) datasets used in the study. Colors of the abbreviation are tied to Fig. 1, namely orange for gauge-based, green for satellite-based, and blue for reanalysis datasets.

Product	Abbreviation	Temporal resolution	Characteristics
the initial “basic” set of products from FLUXCOM-X framework	X-BASE	Jan. 2001 to Dec. 2021	X-BASE applied gradient boosted regression trees from the XGBoost library, and treated the <i>ET</i> fluxes directly observed by eddy covariance measurement systems (FLUXNET) as training targets. For a global up-scaling, its predictors were remotely sensed vegetation indices and land surface temperatures from the Moderate Resolution Imaging Spectroradiometer along with meteorological variables (Nelson et al., 2024).
The Terra Moderate Resolution Imaging Spectroradiometer MOD16A2GF Version 6.1 Evapotranspiration product	MODIS	Jan. 2000 to Dec. 2022	The MOD16 algorithm for the product we used is based on the logic of the Penman-Monteith equation which uses daily meteorological reanalysis data and 8-day remotely sensed vegetation property dynamics that remotely observed by MODIS as inputs (Running et al., 2019).
Priestley-Taylor Jet Propulsion Laboratory	PT-JPL	Jan. 2002 to Apr. 2017	PT-JPL was using the Priestley–Taylor model with the inputs of net radiation, normalized difference vegetation index (NDVI), soil adjusted vegetation index (SAVI), maximum air temperature, and water vapor pressure. It was globally driven by datasets from the International Satellite Land-Surface Climatology Project, Initiative II (ISLSCP-II) and the Advanced Very High Resolution Spectroradiometer (AVHRR) (Fisher et al., 2008).
Penman-Monteith-Leuning Evapotranspiration	PMLv2	Mar. 2000 to Dec. 2022	PML developed a Penman-Monteith-Leuning model, and takes MODIS data (leaf area index, albedo, and emissivity) together with GLDAS (whose <i>ET</i> products are shown in the following) meteorological forcing data as model inputs (Zhang et al., 2019).
Global Land	GLASS	Mar. 2000 to	GLASS <i>ET</i> product algorithm is based on

Surface Satellite product		Dec. 2018	the multi-model ensemble method, that is the Bayesian model averaging method which merges five process-based <i>ET</i> algorithms. They are MOD16, revised remote-sensing-based Penman-Monteith ET algorithm (RRS-PM), PT-JPL, modified satellite-based Priestley-Taylor ET algorithm (MS-PT) and semi-empirical Penman ET algorithm of the University of Maryland (UMD-SEMI) (http://www.glass.umd.edu/introduction.html).
Global Land Evaporation Amsterdam Model v4.1a	GLEAM v4.1	Jan. 2000 to Dec. 2022	GLEAM estimates <i>ET</i> based on Penman's equation, instead of the Priestley and Taylor's equation used in its v3 products. Accordingly, it requires satellite observations of surface net radiation, near-surface air temperature, wind speed, vegetation height, and vapour pressure deficit to calculate potential <i>ET</i> . Converting potential <i>ET</i> into <i>ET</i> , evaporative stress based on hybrid learning with eddy-covariance and sapflow data was used (https://www.gleam.eu/).
The second Modern-Era Retrospective analysis for Research and Applications	MERRA-2	Jan. 2000 to Dec. 2022	MERRA-2 uses observation-corrected <i>P</i> data (Table S1) as forcing for the land surface parameterization. To maintain realistic balances between variations in total mass and total water content in previous reanalyzes, it conserves atmospheric dry mass and guarantees that the net source of water from <i>P</i> and surface <i>E</i> equals the change in total atmospheric water. (Gelaro et al., 2017).
Component 2.0 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.0	Jan. 2000 to Dec. 2014	GLDAS-2.0 was forced entirely with the Princeton meteorological forcing input data, and it used Daily Catchment model with land cover data based on AVHRR (Rui et al., 2022).
Component 2.1 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.1	Jan. 2000 to Dec. 2022	GLDAS-2.1 is forced with a combination of model (NOAA/Global Data Assimilation System (GDAS) atmospheric analysis fields) and observation data (GPCP(D) v1.3). The simulation from March 1, 2001 onwards

			also includes the Air Force Weather Agency's AGRicultural METeorological modeling system (AGRMET) radiation fields as inputs. The simulations are using the Noah Model 3.6 in Land Information System (LIS) Version 7 (Rui et al., 2022).
Component 2.2 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.2	Feb. 2003 to Dec. 2022	GLDAS-2.2 uses Daily Catchment model as GLDAS-2.0, but is forced with the meteorological analysis fields from ECMWF Integrated Forecasting System (IFS). In addition, The total terrestrial water anomaly observation from the Gravity Recovery and Climate Experiment (GRACE) was assimilated (Rui et al., 2022).
the land component of the fifth generation of European ReAnalysis	ERA5-land	Jan. 2000 to Dec. 2022	Differences between ERA5 (mentioned in Table S1 for <i>P</i>) and ERA5-land are not so obvious. They both share quite similar parameterizations of land processes; the main improvement of ERA5-land is due to the non-linear dynamical downscaling with corrected thermodynamic input. However, through the use of common reanalysis forcing, the higher realism of the more complex model in ERA5-Land is evident. That is Carbon Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land (CHTESSEL) land surface model, and ERA5 uses IFS (+ CHTESSEL) as the land surface model (Muñoz-Sabater et al., 2021).

Table S3. Information of runoff (R) datasets used in the study. Colors of the abbreviation are tied to Fig. 1, namely orange for gauge-based, green for satellite-based, and blue for reanalysis datasets.

Product	Abbreviation	Temporal resolution	Characteristics
Global Runoff Reconstruction	GRUN	Jan. 2000 to Dec. 2014	GRUN employed the random forest algorithm to train <i>in situ</i> streamflow observations, with precipitation and temperature data from the Global Soil Wetness Project Phase 3 (GSWP3) dataset, a dynamically downscaled and bias-corrected version of 20CR (mentioned in Table S1). The target observations were derived from the Global Streamflow Indices and Metadata Archive (GSIM), and observations in the Global Runoff Data Centre (GRDC) were used for validation (Ghiggi et al., 2019).
NASA Soil Moisture Active Passive mission Level-4 Surface and Root Zone Soil Moisture Geophysical Data, Version 7	SMAP L4 v7	Apr. 2015 to Dec. 2022	SMAP L4 v7 is derived from the assimilation of Earth's L-band (1.4 GHz) passive microwave brightness temperature observations from SMAP mission into a land surface model driven with observed P forcing. L-band retrievals is highly sensitive to the water content in the top few centimeters of the soil and in the vegetation (Reichle et al., 2019). Therefore, the NASA Catchment land surface model is used, and P data is from CPC and GPCP. Due to encapsulating the L-band constraints into the land surface model, the SMAP L4 v7 reduced its uncertainties of surface soil moisture estimates and runoff estimates (Reichle et al., 2019).
The second Modern-Era Retrospective analysis for Research and Applications	MERRA-2	Jan. 2000 to Dec. 2022	In addition to the MERRA-2 information mentioned in Table S2, MERRA-2 provides an explicit representation of snow densification, meltwater R , percolation, refreezing, and a prognostic surface albedo, but does not provide runoff over land ice which lacks ablation areas (where the annual surface mass balance is negative) along the periphery

of the ice sheet (Gelaro et al., 2017).

Component 2.0 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.0	Jan. 2000 to Dec. 2014	See Table S2.
Component 2.1 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.1	Jan. 2000 to Dec. 2022	See Table S2.
Component 2.2 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.2	Feb. 2003 to Dec. 2022	See Table S2.
the land component of the fifth generation of European ReAnalysis	ERA5-land	Jan. 2000 to Dec. 2022	CHTESSEL, the land surface model used in ERA5-land, did not directly produce river discharge at the river basin scale. Alternatively, gridded surface and subsurface runoff from CHTESSEL is coupled to the LISFLOOD hydrological and channel routing model, where the GRDC observations were considered (Muñoz-Sabater et al., 2021).

Table S4. Information of soil moisture (*SM*) datasets used in the study. Colors of the abbreviation are tied to Fig. 1, namely orange for gauge-based, green for satellite-based, and blue for reanalysis datasets.

Product	Abbreviation	Temporal resolution	Characteristics
A long-term dataset of soil moisture derived through machine learning trained with <i>in situ</i> measurements	SoMo.ml	Jan. 2000 to Dec. 2019	SoMo.ml generated with a Long Short-Term Memory neural network ingesting <i>in situ</i> <i>SM</i> measurements and ERA5 meteorological forcing (whose <i>P</i> data is shown in Table S1). The dataset did not consider the entire soil layers due to the depth limited availability of <i>in situ</i> data (O and Orth, 2021).
NASA Soil Moisture Active Passive mission Level-4 Surface and Root Zone Soil Moisture Geophysical Data, Version 7	SMAP L4 v7	Apr. 2015 to Dec. 2022	See Table S3.
European Space Agency Climate Change Initiative Soil Moisture v08.1 product	ESA CCI v08.1	Jan. 2000 to Dec. 2022	ESA CCI v08.1 is fusing a group of scatterometer and radiometer soil moisture sensors' estimates to provide remote-sensed surface <i>SM</i> estimates globally. The scatterometers and radars of instruments measure the radar backscattering coefficient, whose products are active because they use their own source of electromagnetic energy for the measurement, while the radiometers measuring the brightness temperature is passive instruments because they measure energy that is reflected or emitted from the earth surface (Gruber et al., 2019).
Global Land Evaporation Amsterdam Model v4.1a	GLEAM v4.1	Jan. 2000 to Dec. 2022	GLEAM v4.1 <i>SM</i> estimates were calculated using a multi-layer running water balance that describes the infiltration of observed precipitation through the vertical soil profile, including a linear reservoir model to account for plant access to groundwater. To correct for random forcing errors, microwave observations of surface soil moisture are

The second Modern-Era Retrospective analysis for Research and Applications	MERRA-2	Jan. 2000 to Dec. 2022	assimilated into the soil profile (https://www.gleam.eu/). No more information other than Tables S1–S3.
Component 2.0 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.0	Jan. 2000 to Dec. 2014	See Table S2.
Component 2.1 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.1	Jan. 2000 to Dec. 2022	See Table S2.
Component 2.2 in Version 2 of the NASA Global Land Data Assimilation System	GLDAS-2.2	Feb. 2003 to Dec. 2022	See Table S2.
the land component of the fifth generation of European ReAnalysis	ERA5-land	Jan. 2000 to Dec. 2022	CHTESSEL improved the formulation of the soil hydrologic conductivity and diffusivity, compared to the land component of ERA-Interim (Muñoz-Sabater et al., 2021).

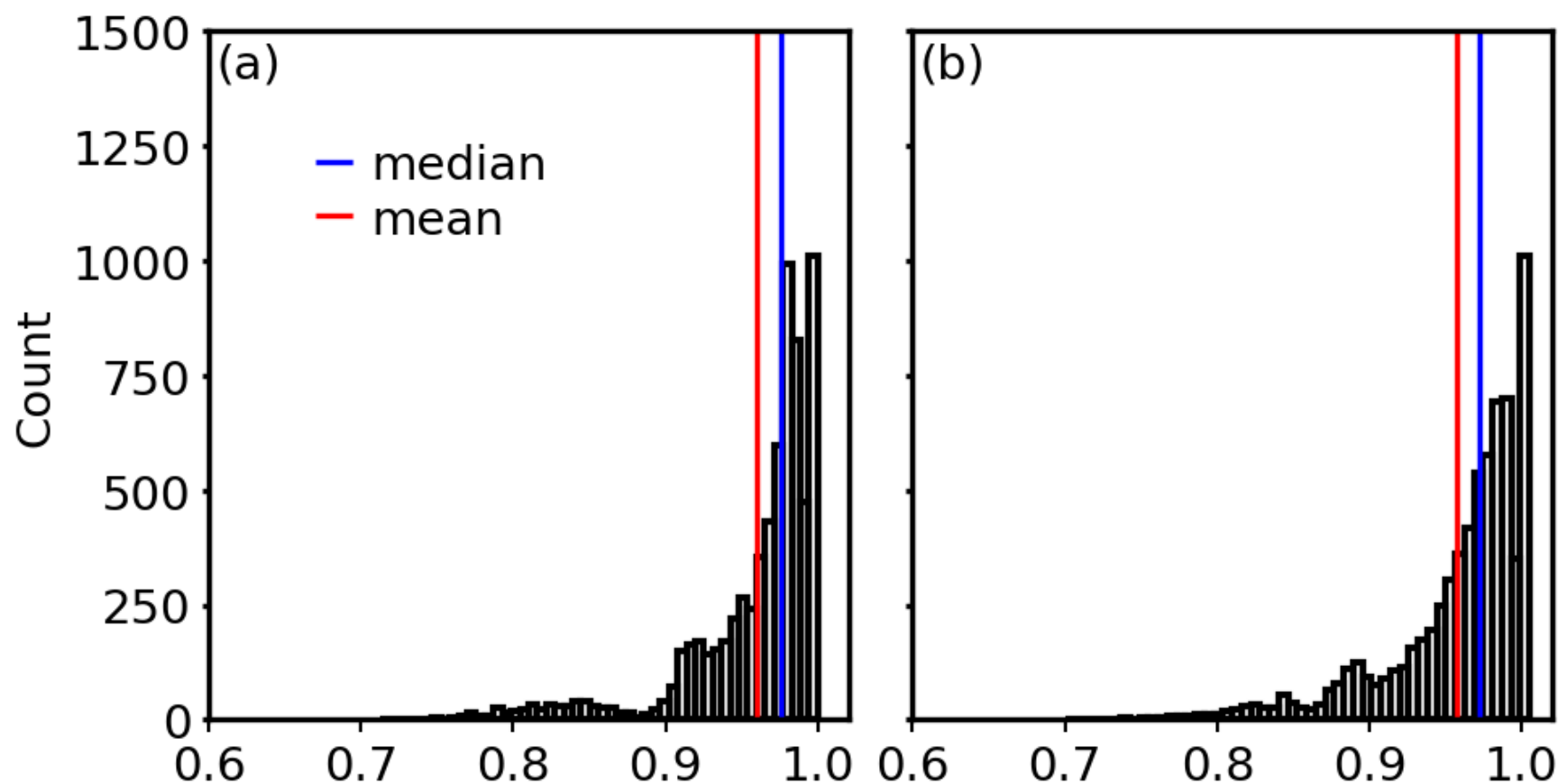


Fig. S1. Histograms of (a) R^2 scores and (b) slopes that were used for the comparisons between using varying M and using fixed M in equation (1)'s calculation (Methods). There are 7,001 independent combinations available for both varying M and fixed M . For each of them, we calculated two performance maps of using varying M and fixed M in equation (1), and obtained a linearly regressed R^2 score and a slope to evaluate if varying M would lead to substantial differences compared to fixed M .

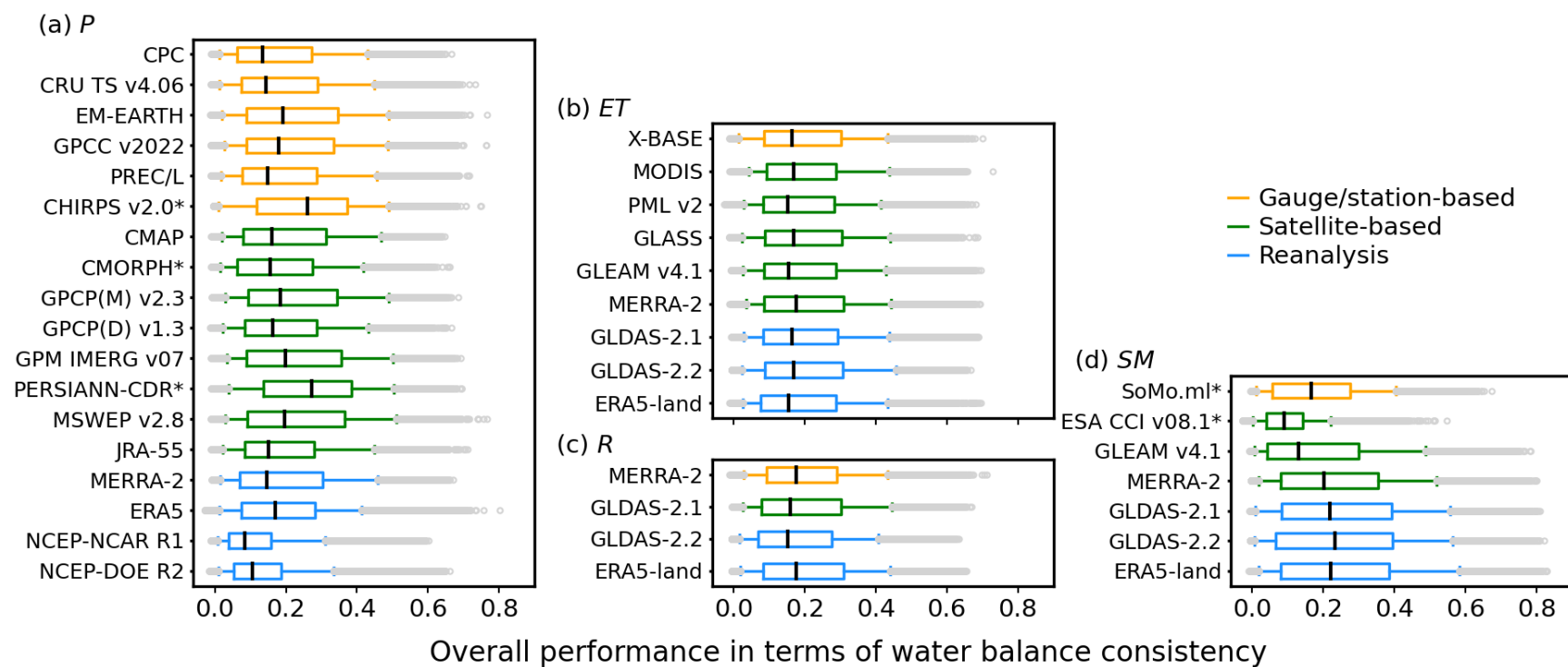


Fig. S2. Similar to Fig. 1 but using the combinations for temporal changes analysis, instead of all the 8,294 independent combinations because they have similarly large numbers of monthly data.

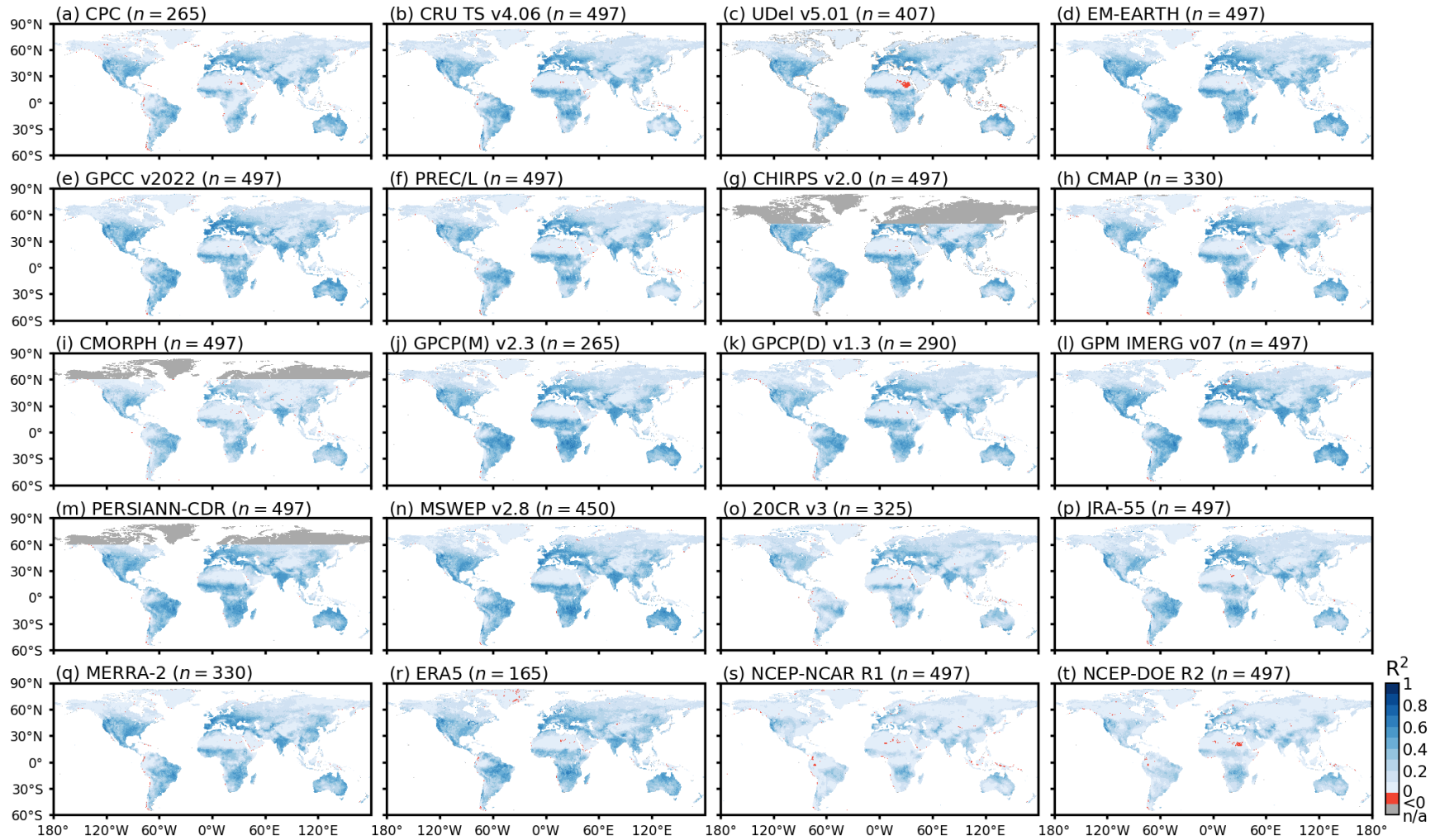


Fig. S3. Overall performance of P datasets in terms of water balance consistency. The overall performance of each dataset was calculated by averaging the adjusted R^2 scores of all the independent combinations that include this dataset. n for each dataset denotes the number of independent combinations that include this dataset.

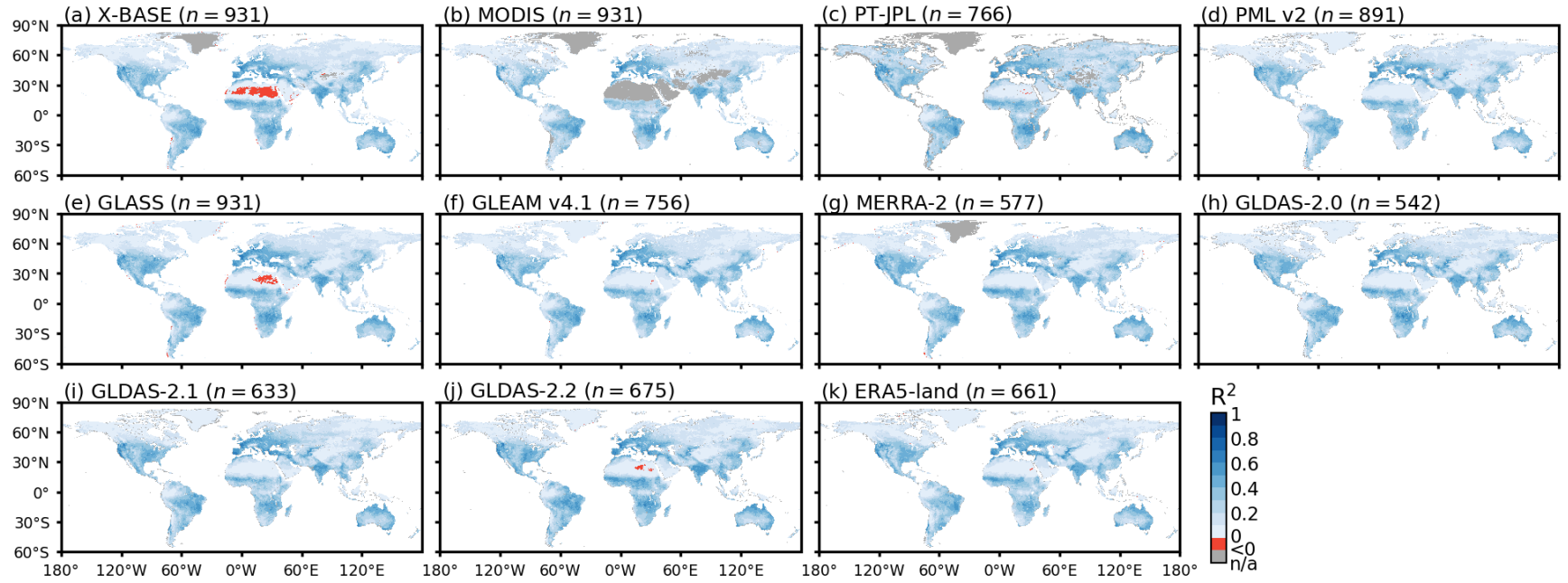


Fig. S4. Overall performance of *ET* datasets in terms of water balance consistency. The overall performance of each dataset was calculated by averaging the adjusted R^2 scores of all the independent combinations that include this dataset. n for each dataset denotes the number of independent combinations that include this dataset.

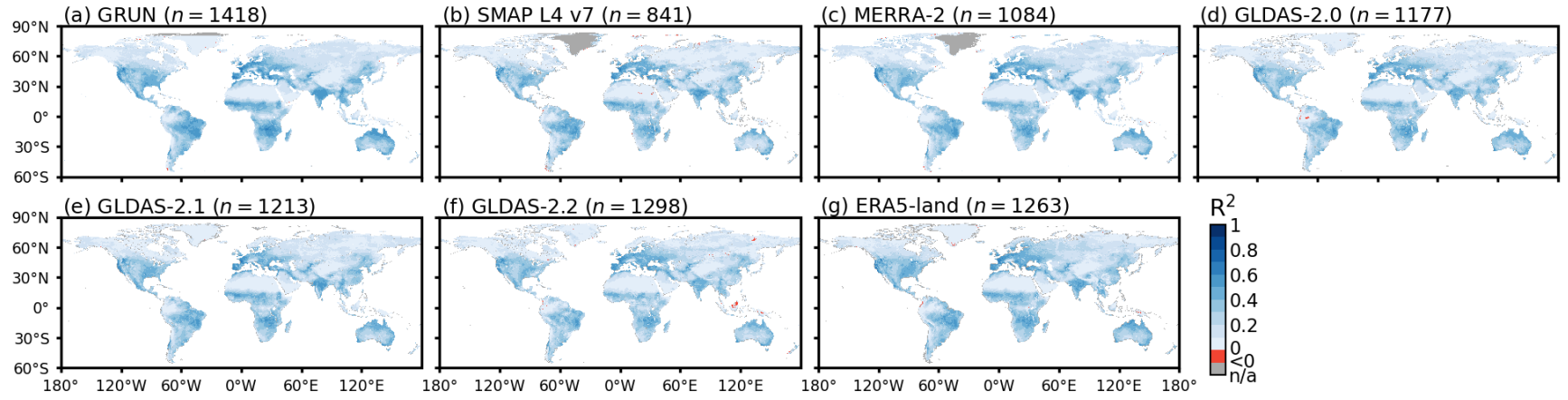


Fig. S5. Overall performance of R datasets in terms of water balance consistency. The overall performance of each dataset was calculated by averaging the adjusted R^2 scores of all the independent combinations that include this dataset. n for each dataset denotes the number of independent combinations that include this dataset.

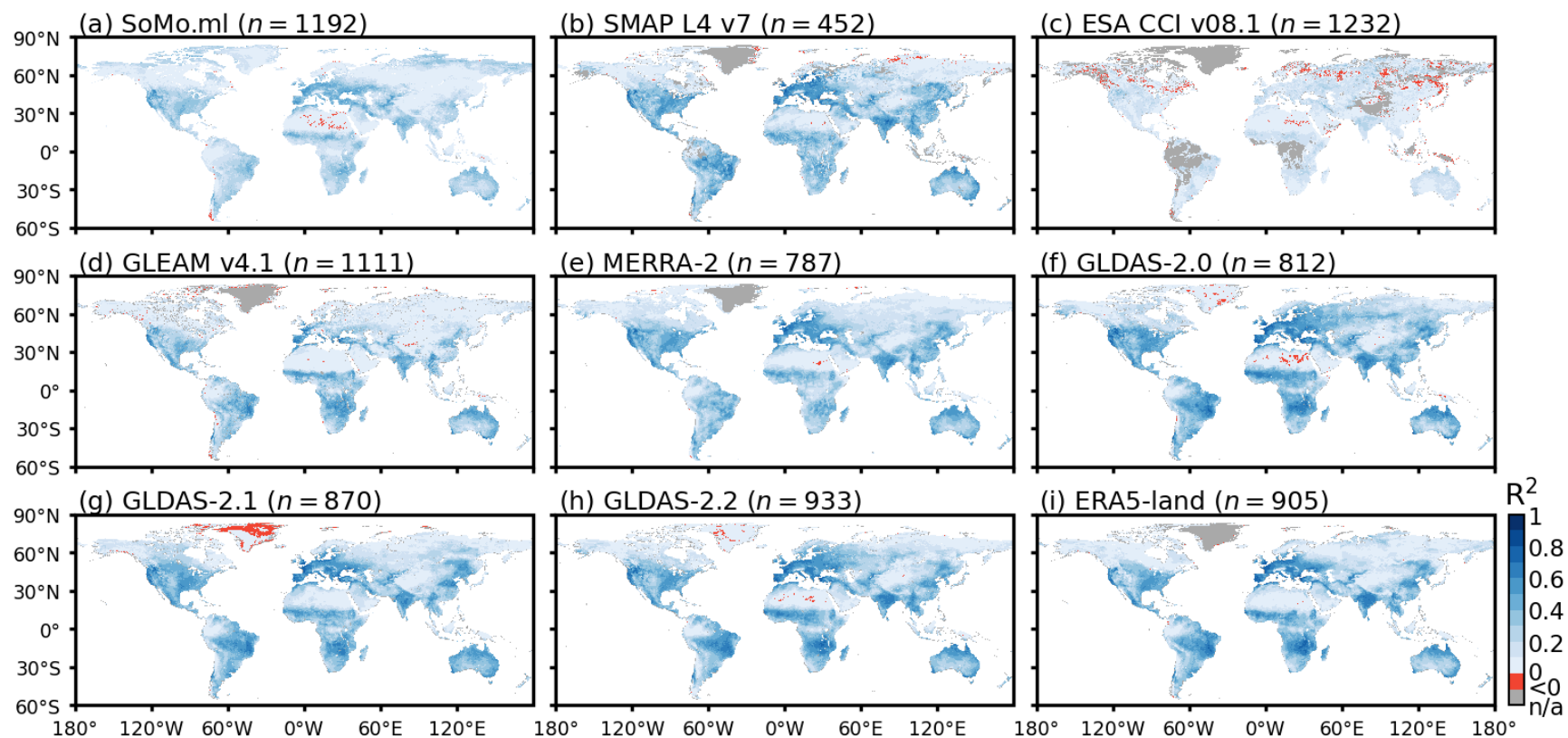


Fig. S6. Overall performance of *SM* datasets in terms of water balance consistency. The overall performance of each dataset was calculated by averaging the adjusted R^2 scores of all the independent combinations that include this dataset. n for each dataset denotes the number of independent combinations that include this dataset.

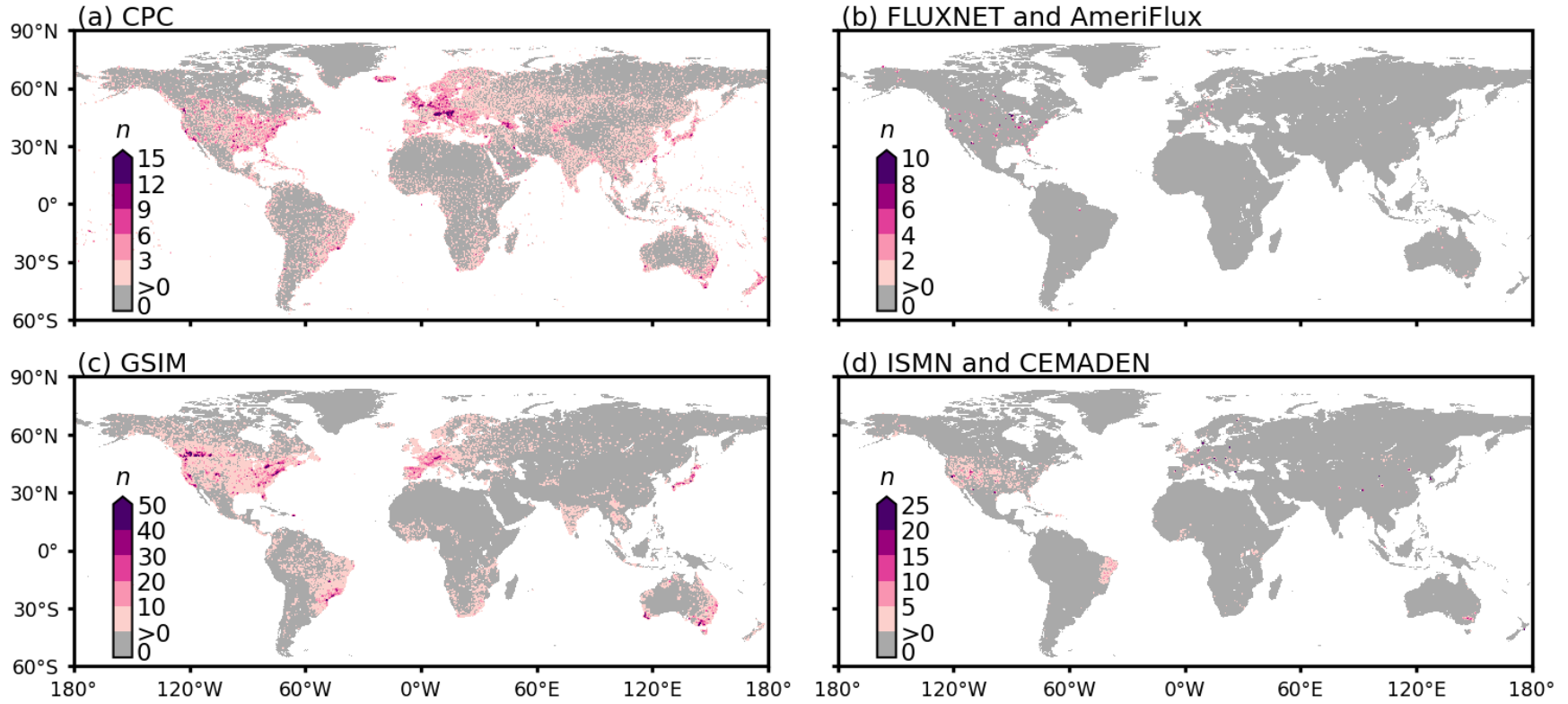


Fig. S7. Maps showing stations/sites density for different observation networks of the water variables, including (a) CPC global stations for P , eddy covariance sites in FLUXNET2015 (Pastorello et al., 2020) and AmeriFlux for ET , streamflow stations in the Global Streamflow Indices and Metadata Archive (GSIM) (Do et al., 2018) for R , and sites of *in-situ* measurement in the International Soil Moisture Network (ISMN) (Dorigo et al., 2011) and the National Center for Monitoring and Early Warning of Natural Disasters of Brazil (CEMADEN) (Zeri et al., 2020) for SM . The density in each grid cell refers to the sum of stations/sites located in each grid cell and its eight neighboring grid cells (Ruiz-Vásquez et al., 2022).



Fig. S8. Correlation matrices of the seven predictors for the random forest model applied on each *P* dataset. The calculations of each predictor map are detailed in Methods.

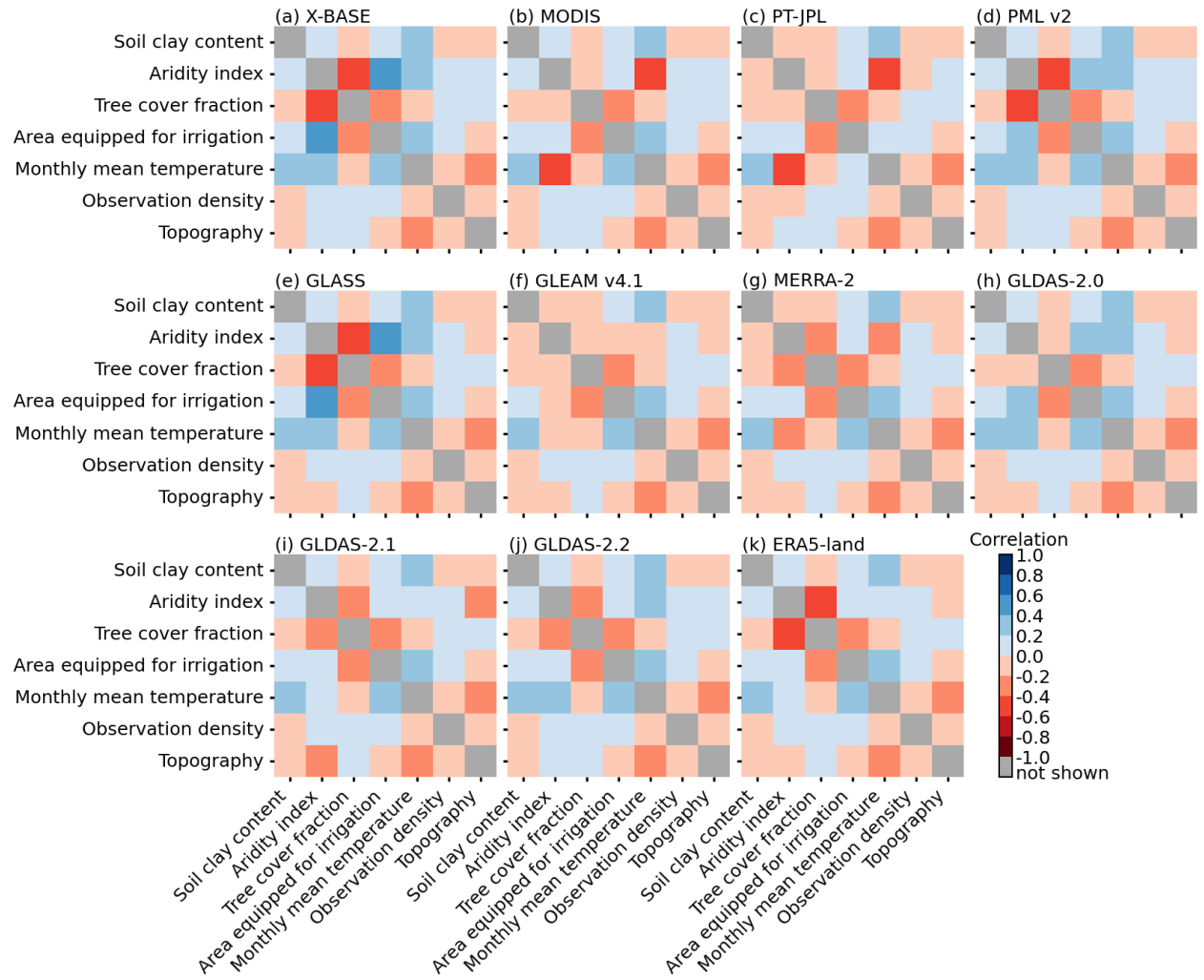


Fig. S9. Correlation matrices of the seven predictors for the random forest model applied on each *ET* dataset. The calculations of each predictor map are detailed in Methods.

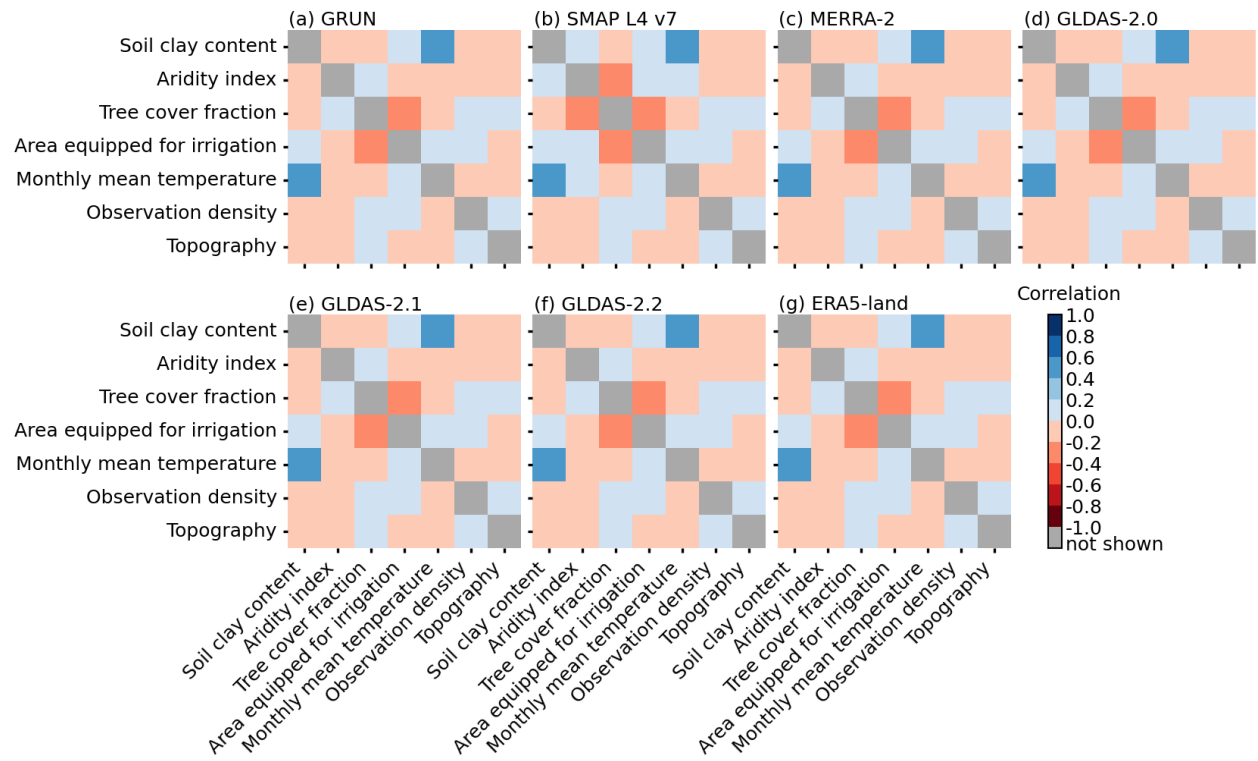


Fig. S10. Correlation matrices of the seven predictors for the random forest model applied on each *R* dataset. The calculations of each predictor map are detailed in Methods.

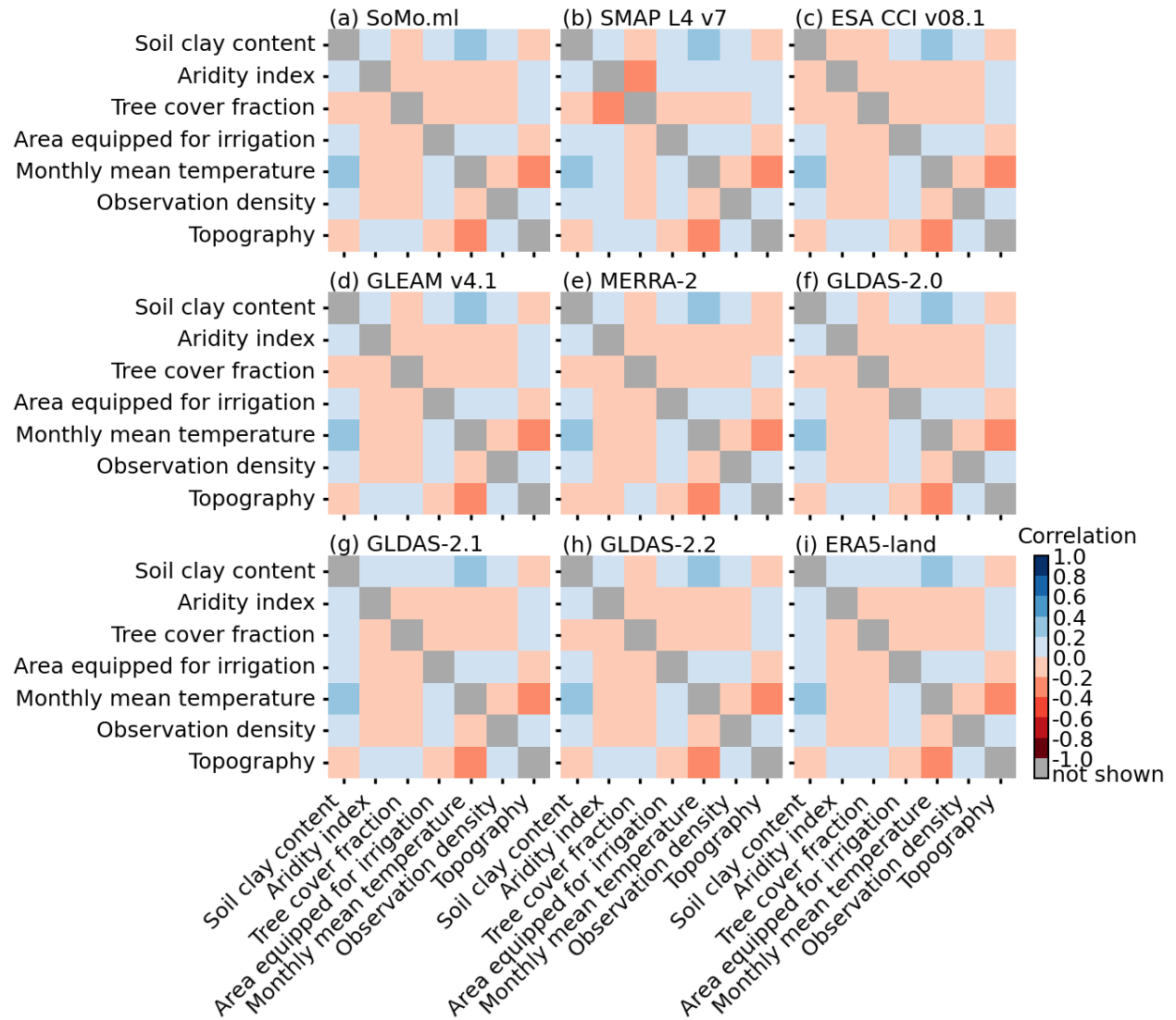


Fig. S11. Correlation matrices of the seven predictors for the random forest model applied on each *SM* dataset. The calculations of each predictor map are detailed in Methods.

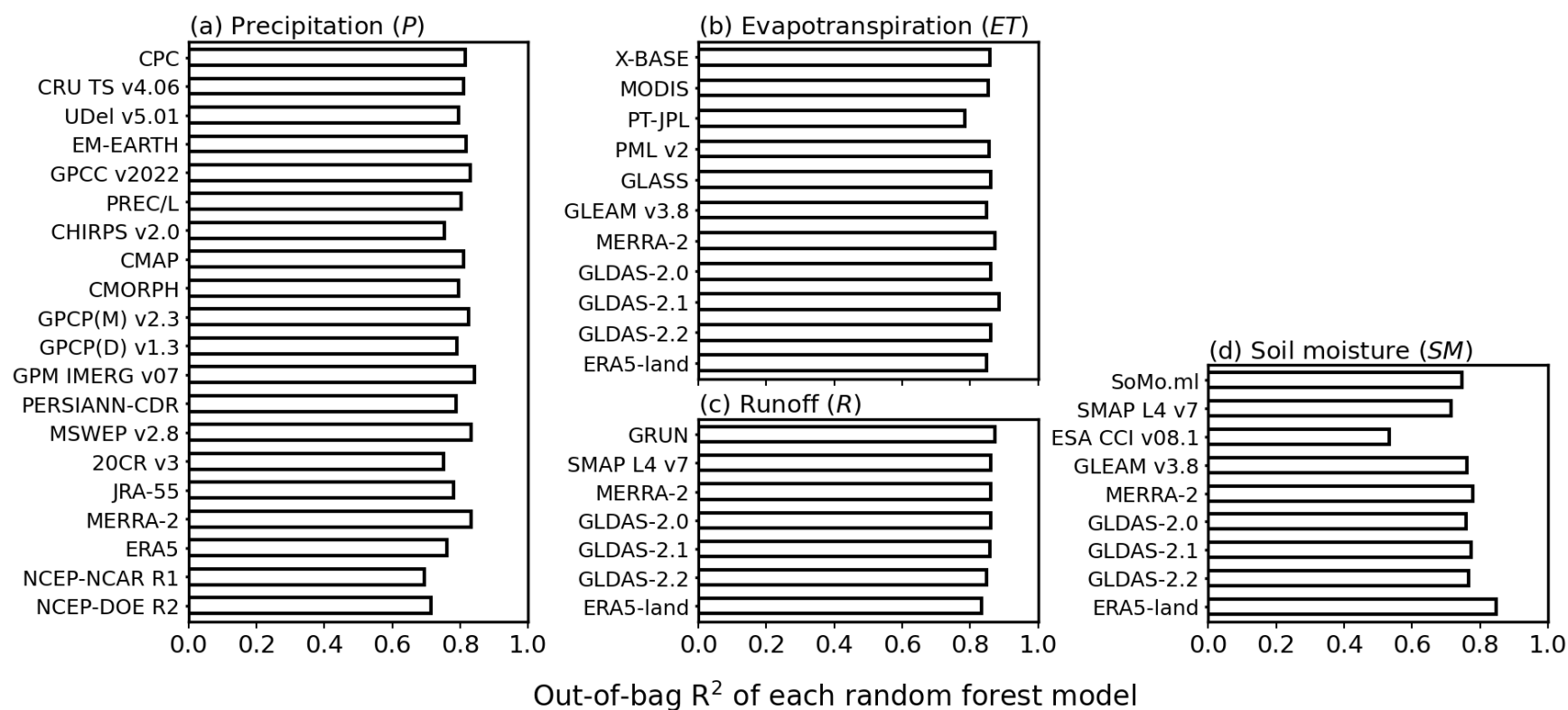


Fig. S12. Out-of-bag R^2 of the random forest models, which were trained to evaluate the relative importance of different factors to the water balance consistency of each dataset.

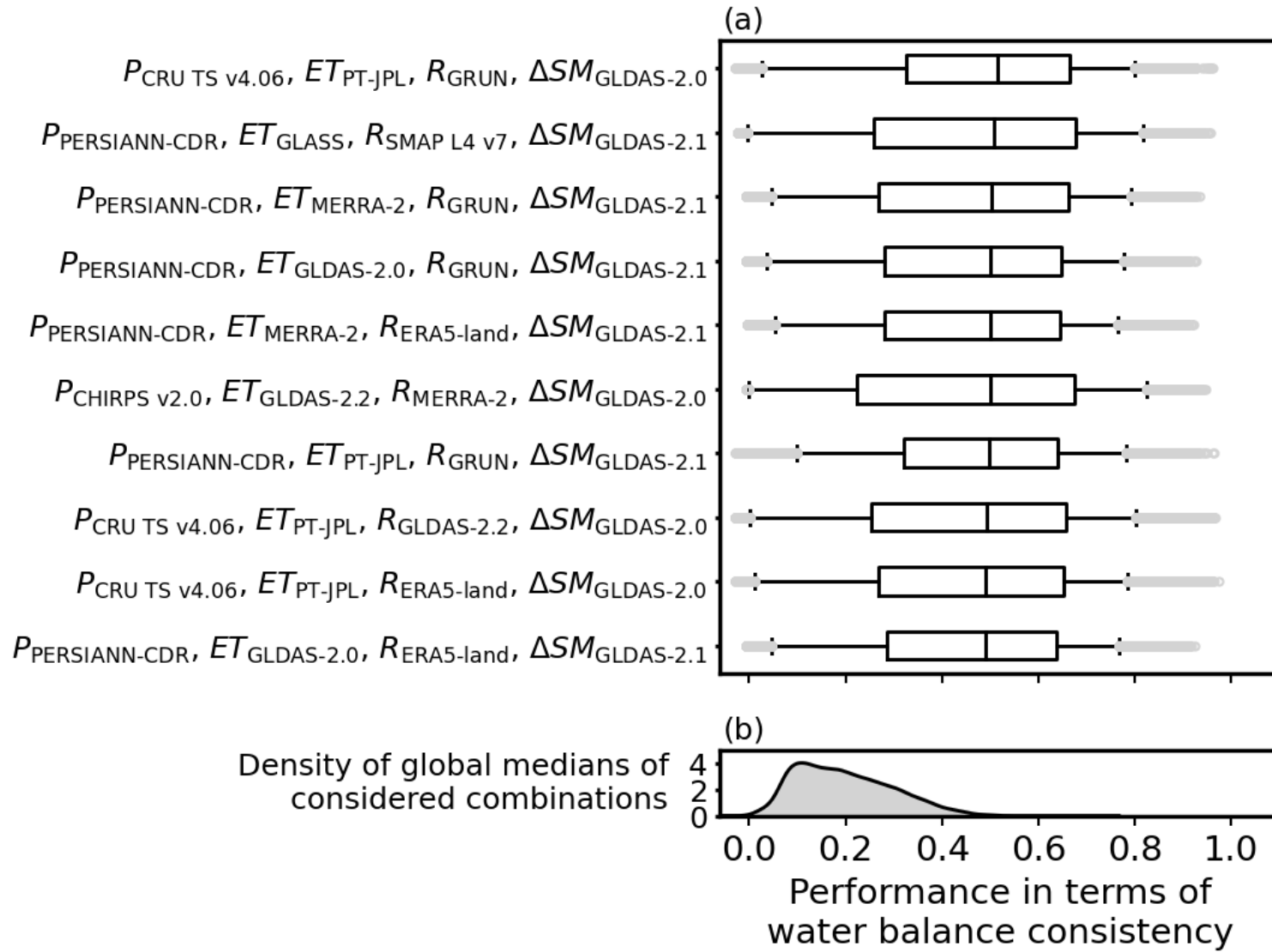


Fig. S13. The combinations with top ten performance in terms of global medians, and kernel density estimation of global medians of all the 8,294 combinations. Each box encompasses the 25th to 75th percentiles, the vertical line in each box marks the median value, and the whiskers are the 5th to 95th percentiles.

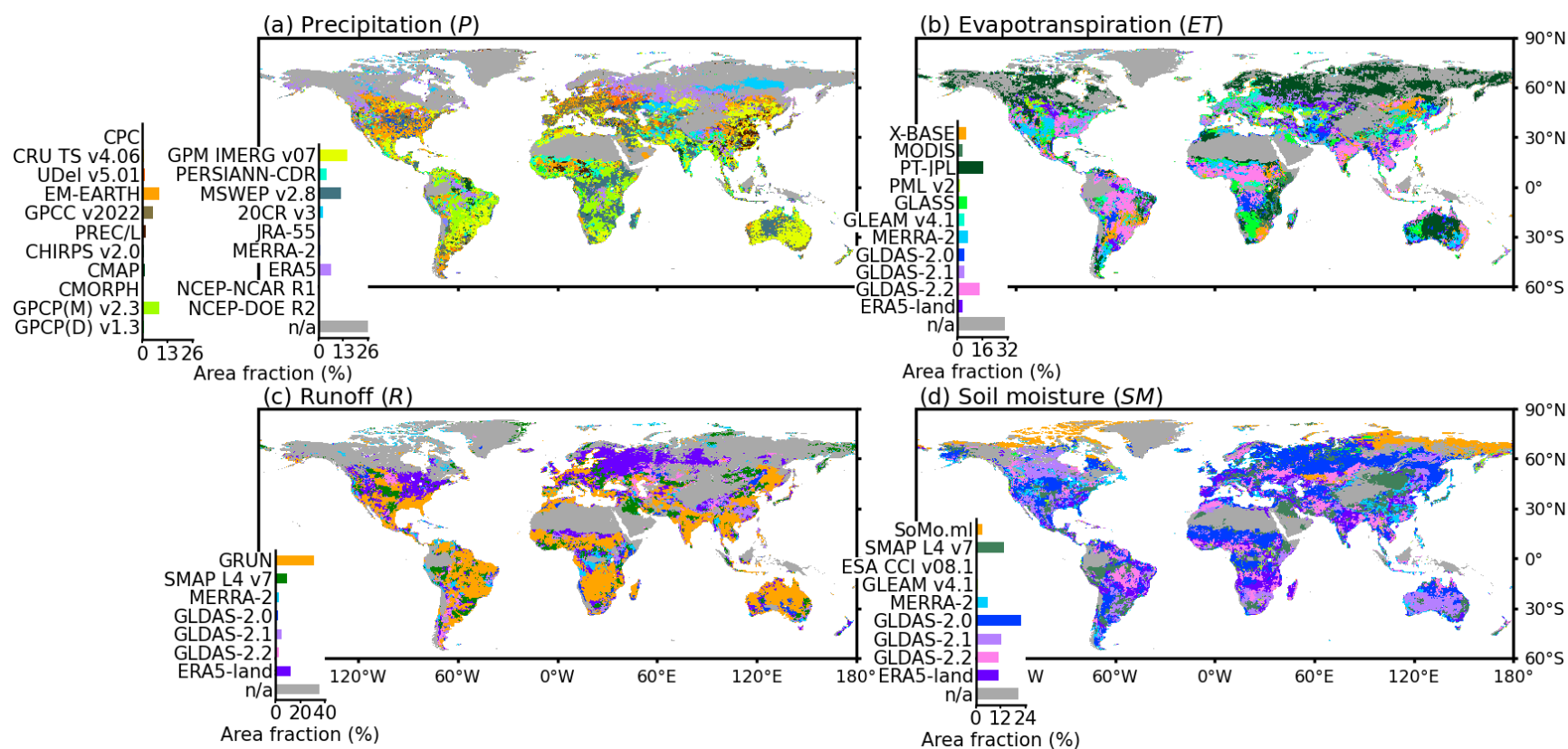


Fig. S14. Maps showing the dataset with the best performance inferred from our water-balance approach. ‘*the best*’ indicates that this dataset had the highest water balance consistency. Grid cells where the performance scores of all datasets are lower than 0.2 are also shown in gray.

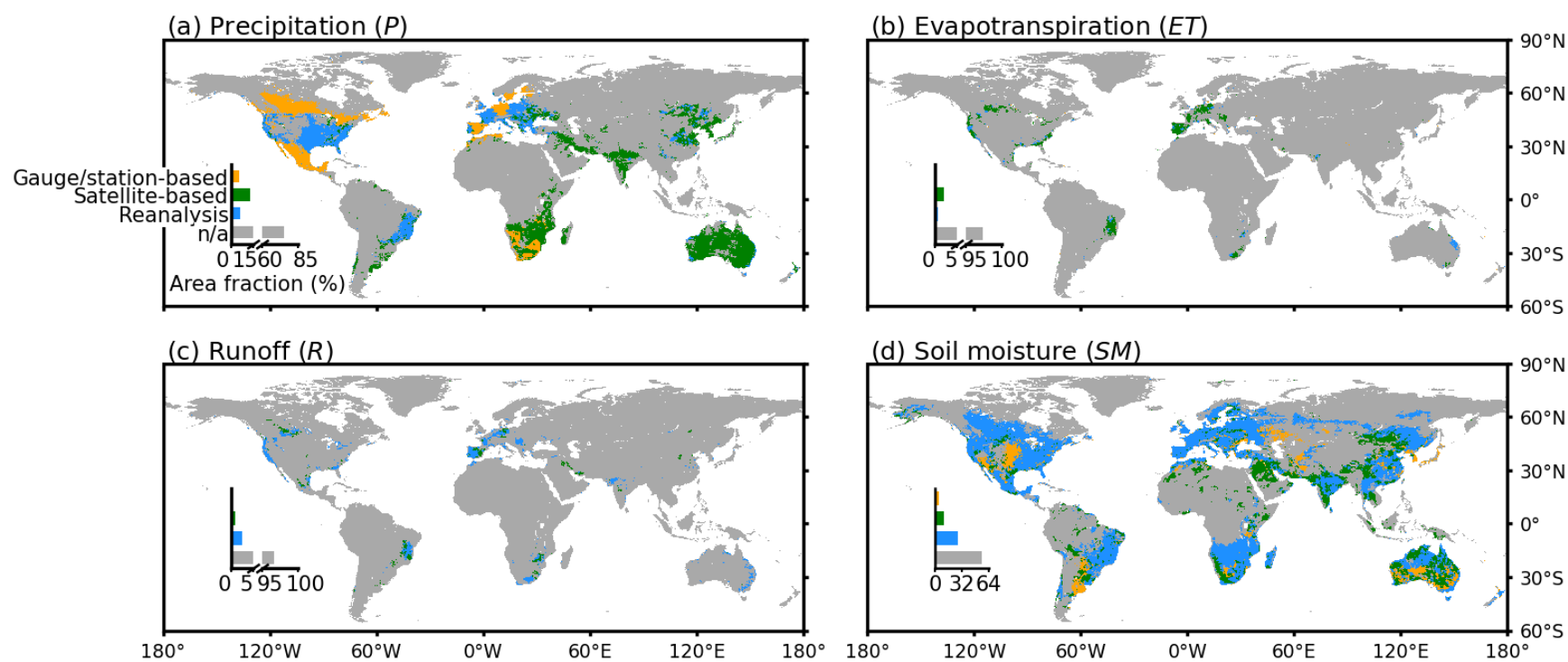


Fig. S15. Similar to Fig. 2, but calculating water balance consistency based on daily data (Text S1).

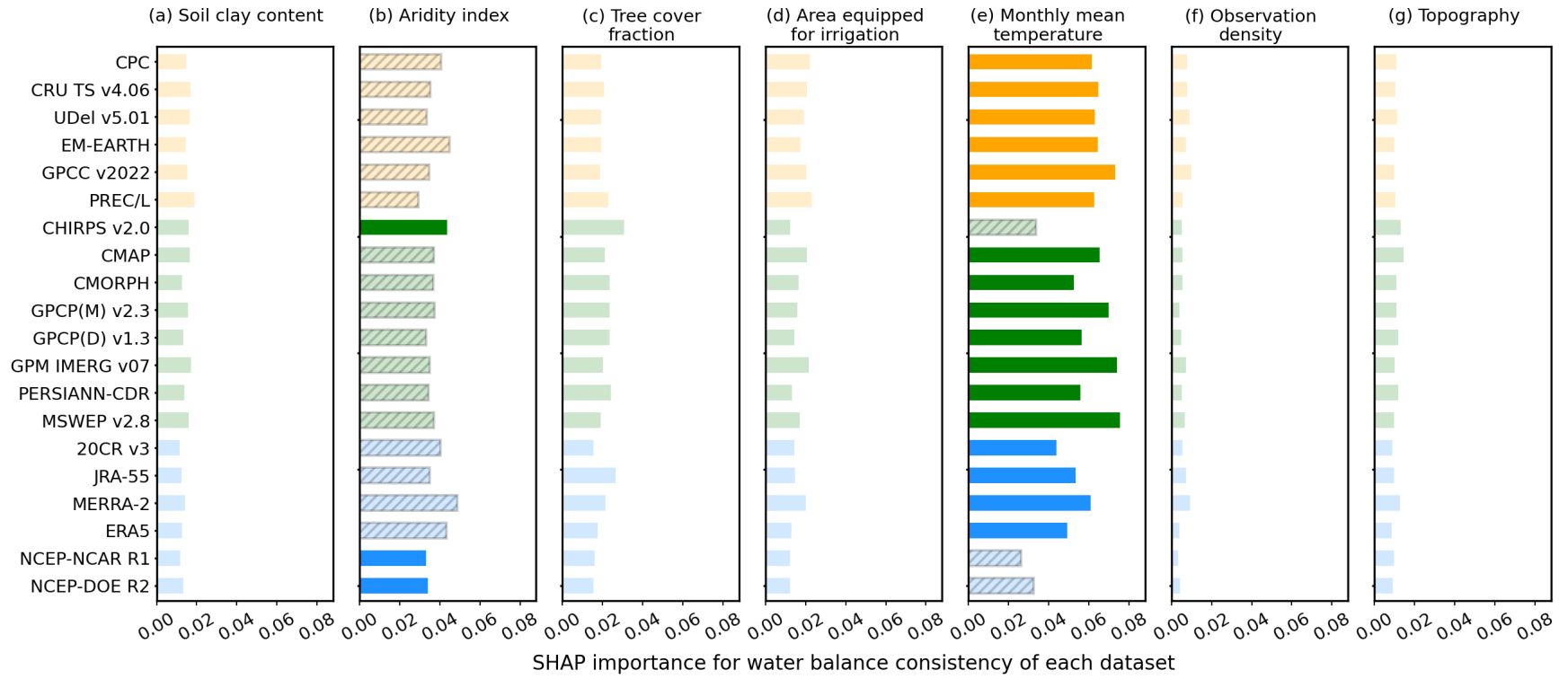


Fig. S16. Importance of (a) soil clay content, (b) aridity index, (c) tree cover fraction, (d) area equipped for irrigation, (e) monthly mean temperature, (f) observation density, and (g) topography to water balance consistency of each P dataset. The importance is quantified by global averaged absolute SHAP values (Methods). Bars with darkened color and hatch, respectively, indicate the firstly and the secondly important factor for the water balance consistency of each P dataset.

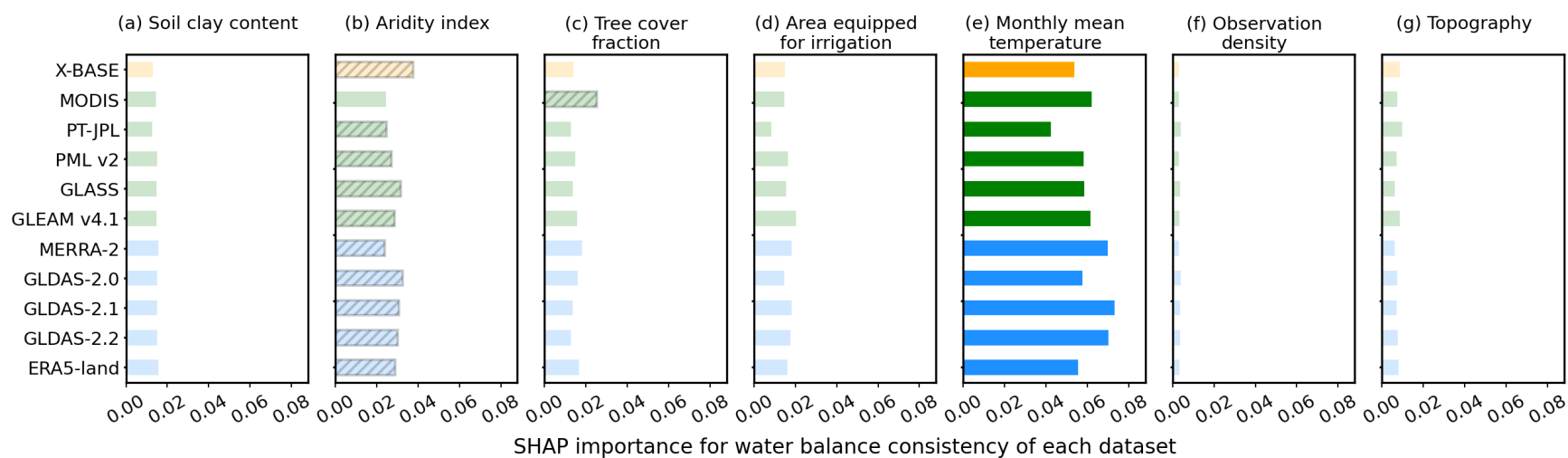


Fig. S17. Importance of (a) soil clay content, (b) aridity index, (c) tree cover fraction, (d) area equipped for irrigation, (e) monthly mean temperature, (f) observation density, and (g) topography to water balance consistency of each *ET* dataset. The importance is quantified by global averaged absolute SHAP values (Methods). Bars with dark color and hatch, respectively, indicate the firstly and the secondly important factor for the water balance consistency of each *ET* dataset.

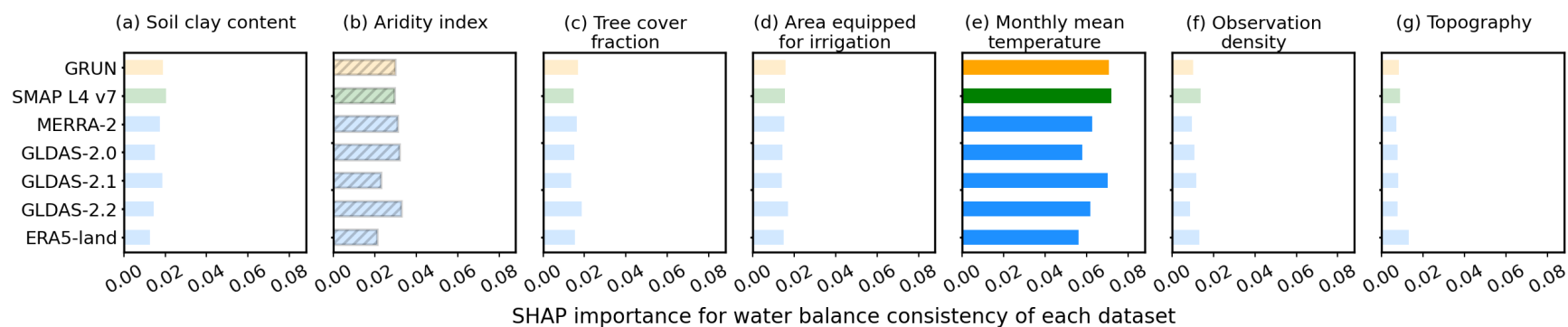


Fig. S18. Importance of (a) soil clay content, (b) aridity index, (c) tree cover fraction, (d) area equipped for irrigation, (e) monthly mean temperature, (f) observation density, and (g) topography to water balance consistency of each *R* dataset. The importance is quantified by global averaged absolute SHAP values (Methods). Bars with darken color and hatch, respectively, indicate the firstly and the secondly important factor for the water balance consistency of each *R* dataset.

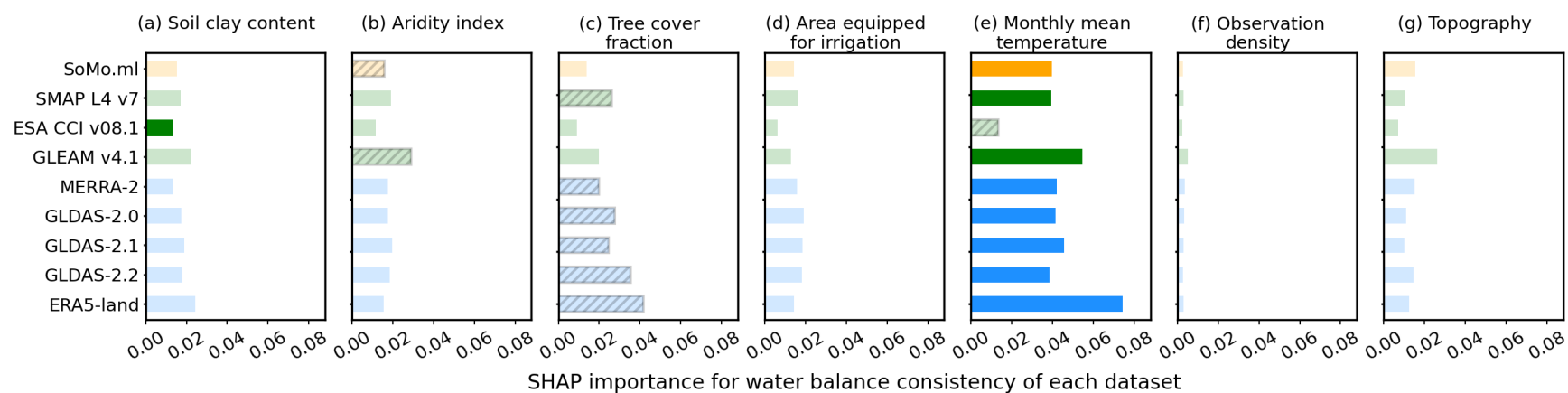


Fig. S19. Importance of (a) soil clay content, (b) aridity index, (c) tree cover fraction, (d) area equipped for irrigation, (e) monthly mean temperature, (f) observation density, and (g) topography to water balance consistency of each *SM* dataset. The importance is quantified by global averaged absolute SHAP values (Methods). Bars with darkened color and hatch, respectively, indicate the firstly and the secondly important factor for the water balance consistency of each ΔSM dataset.

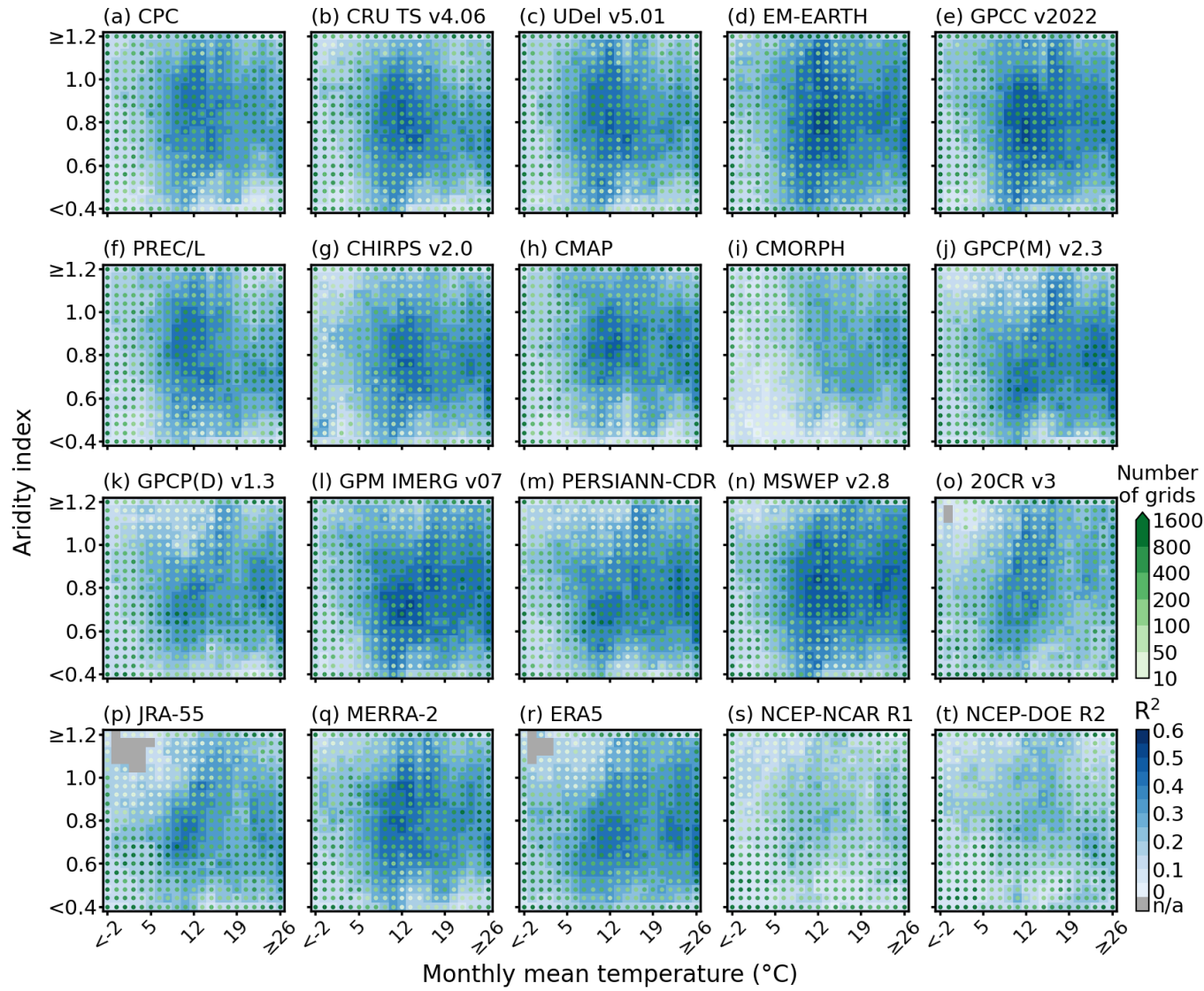


Fig. S20. Influence of aridity index and monthly mean temperature on water balance consistency of each P dataset. The consistency through water balance is quantified by R^2 scores (Fig. 1), and heatmaps show means across grid cells in different temperature and aridity regimes. The numbers of grid cells belonging to a certain aridity group are shown when there are at least ten.

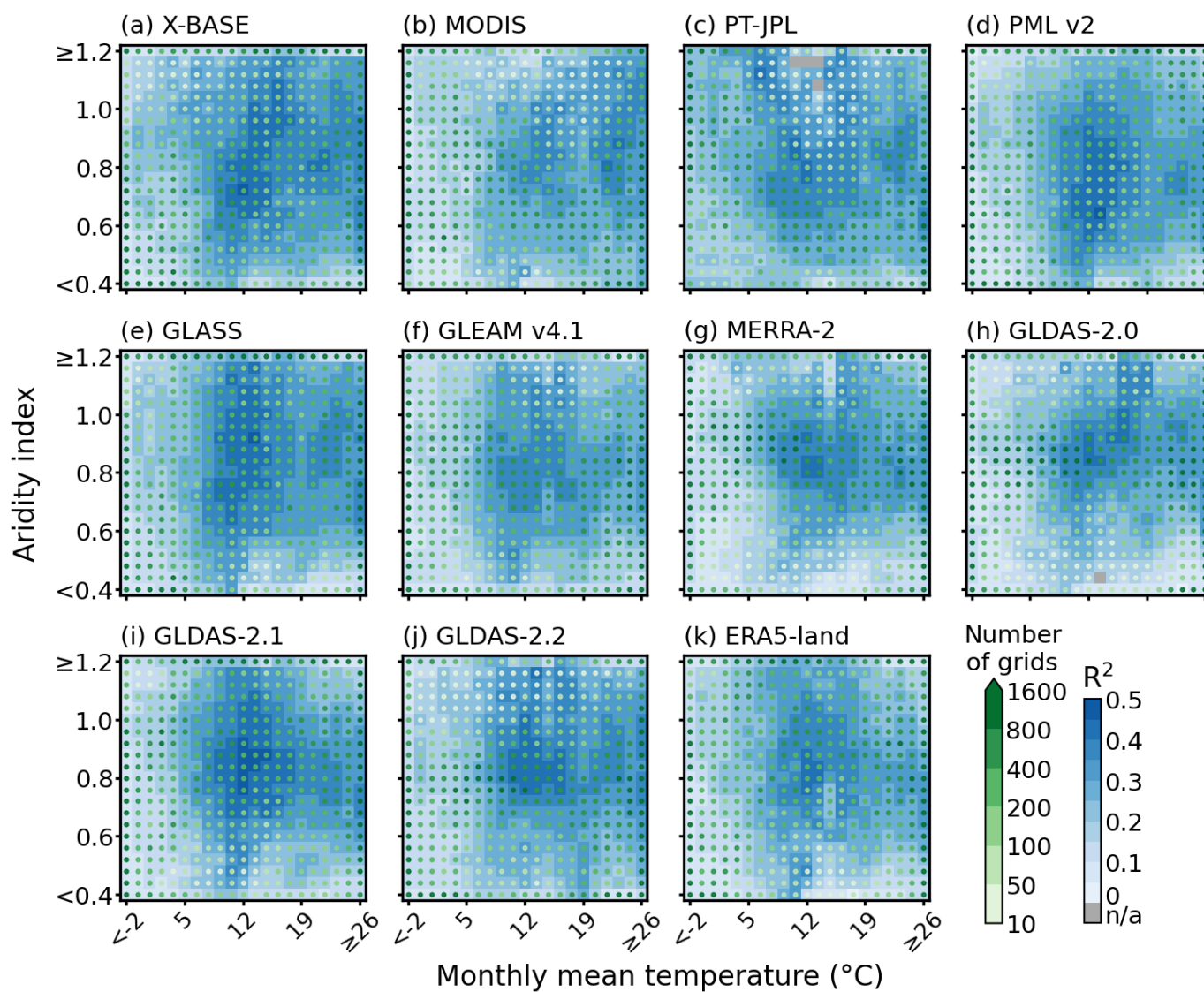


Fig. S21. Influence of aridity index and monthly mean temperature on water balance consistency of each *ET* dataset. The consistency through water balance is quantified by R^2 scores (Fig. 1), and heatmaps show means across grid cells in different temperature and aridity regimes. The numbers of grid cells belonging to a certain aridity group are shown when there are at least ten.

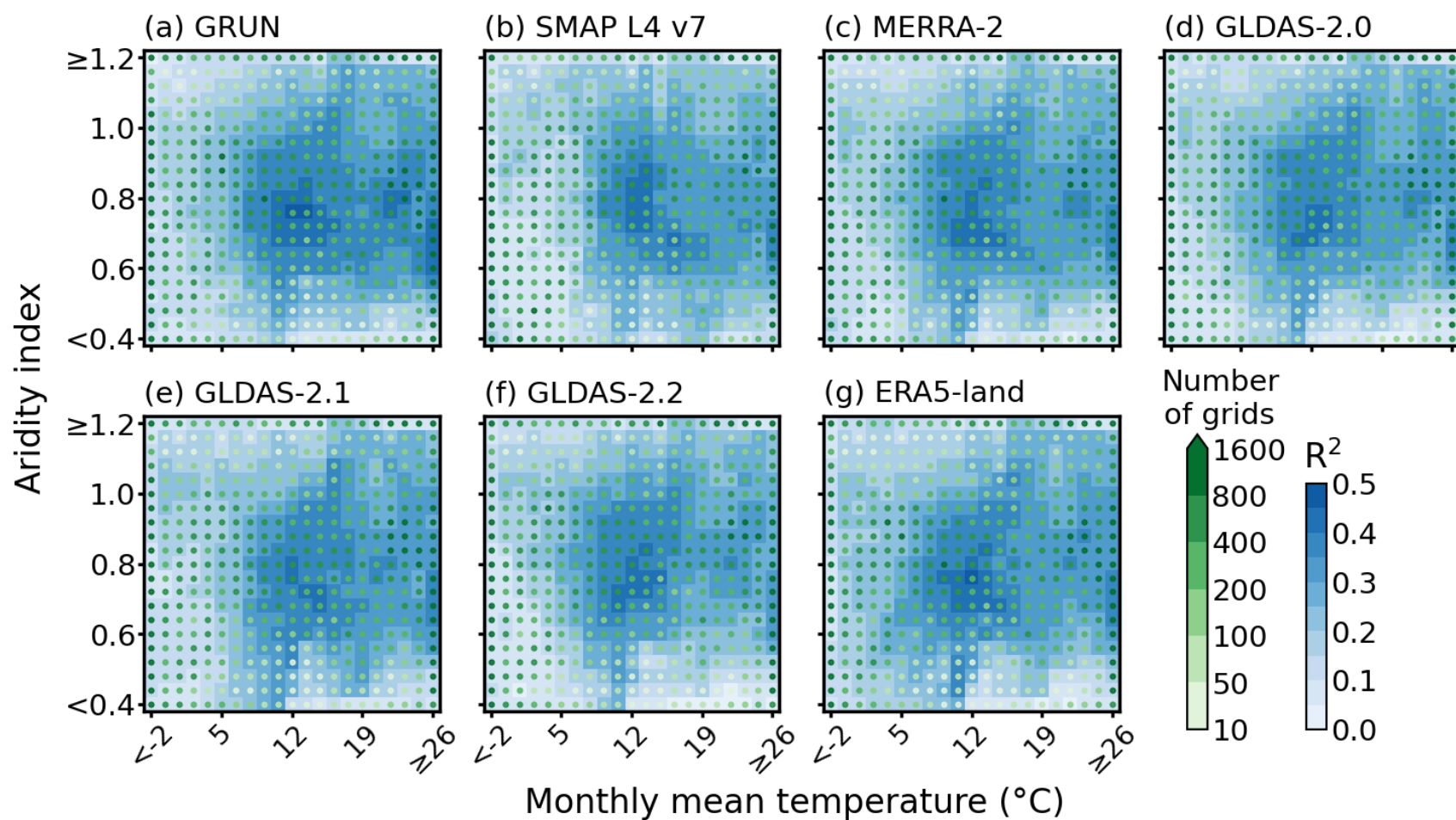


Fig. S22. Influence of aridity index and monthly mean temperature on water balance consistency of each R dataset. The consistency through water balance is quantified by R^2 scores (Fig. 1), and heatmaps show means across grid cells in different temperature and aridity regimes. The numbers of grid cells belonging to a certain aridity group are shown when there are at least ten.

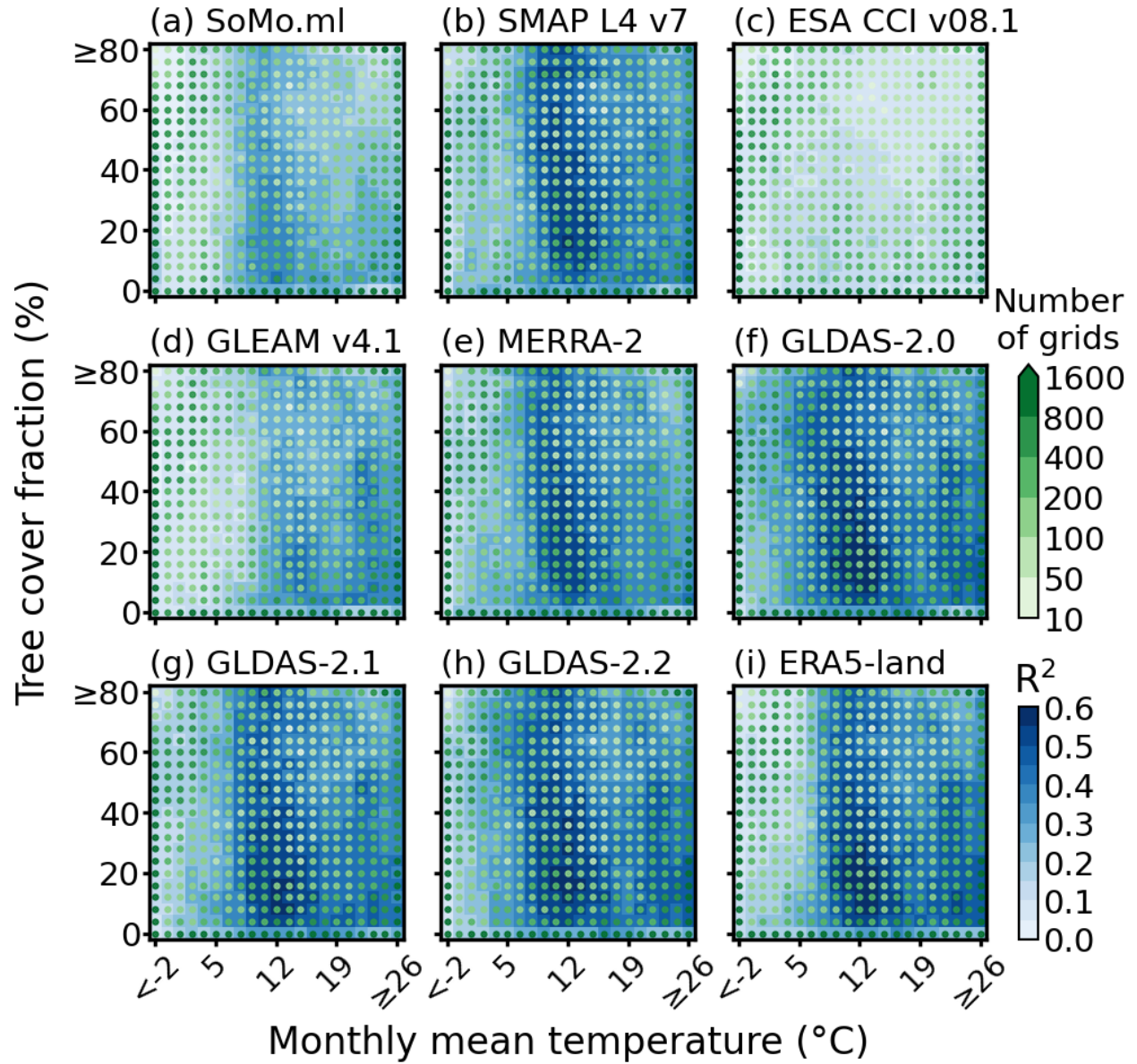


Fig. S23. Influence of aridity index and monthly mean temperature on water balance consistency of each *SM* dataset. The consistency through water balance is quantified by R^2 scores (Fig. 1), and heatmaps show means across grid cells in different temperature and aridity regimes. The numbers of grid cells belonging to a certain aridity group are shown when there are at least ten.

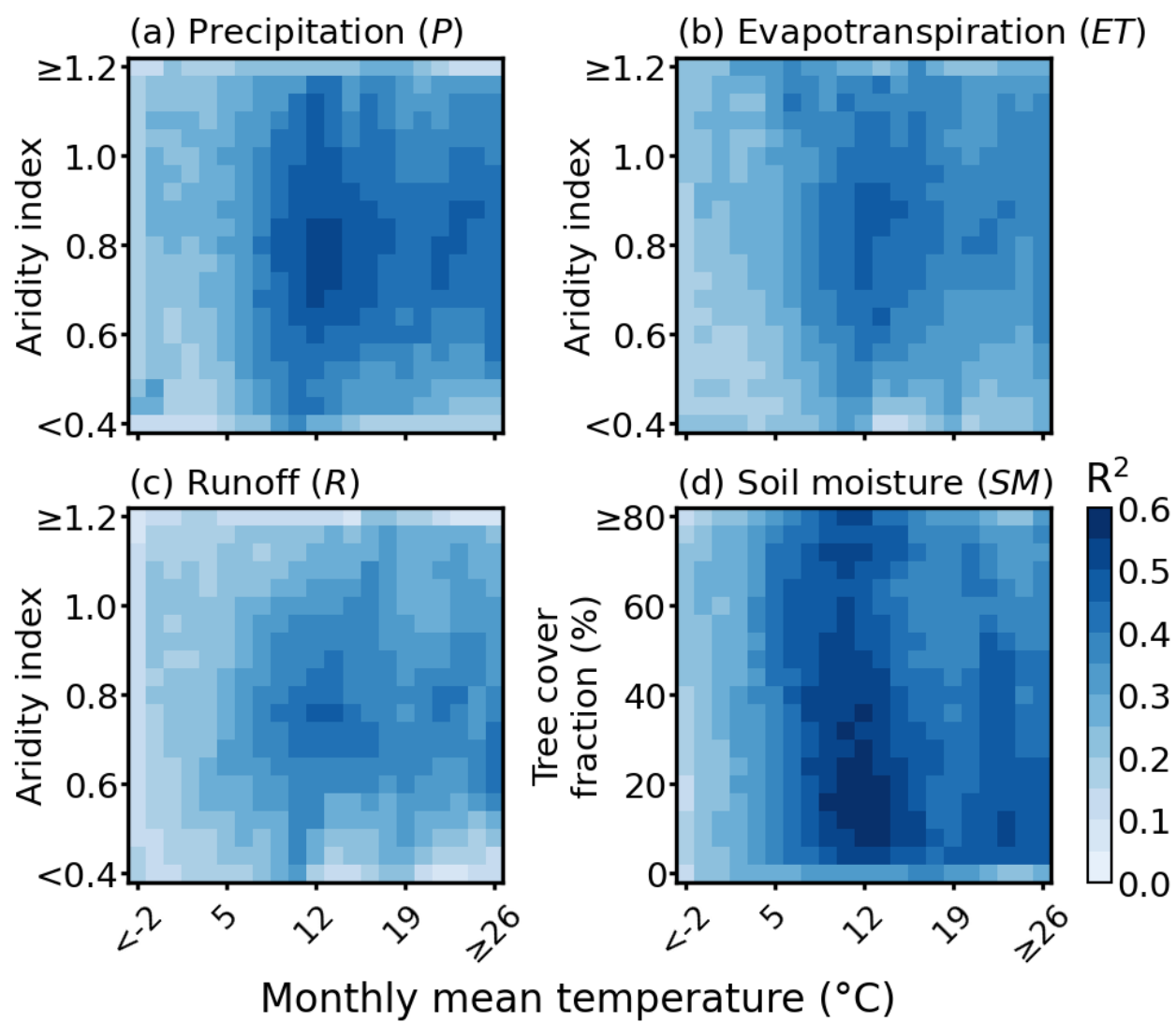


Fig. S24. Similar to Fig. 3 but showing maximums in each pixel of the heatmap.

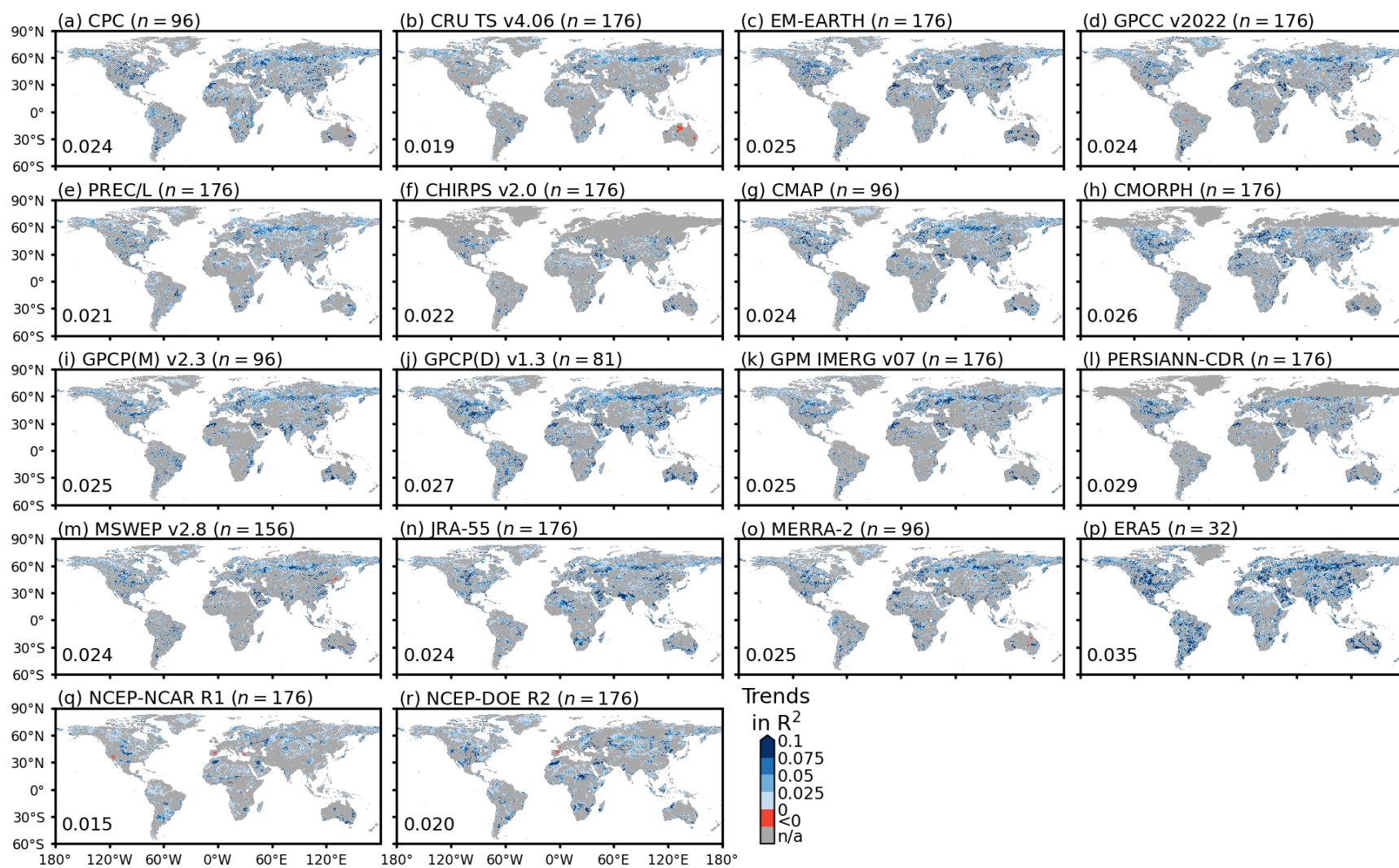


Fig. S25. Temporal changes in water balance consistency of P datasets from 2000–2010 to 2011–2022. The consistency through water balance is quantified by R^2 scores in the two periods, and differences to indicate temporal changes are obtained (Methods). By accounting for bootstrap-based uncertainties, only the significant temporal changes are shown in blue/red (Methods). Global medians are shown in the left-lower corner.

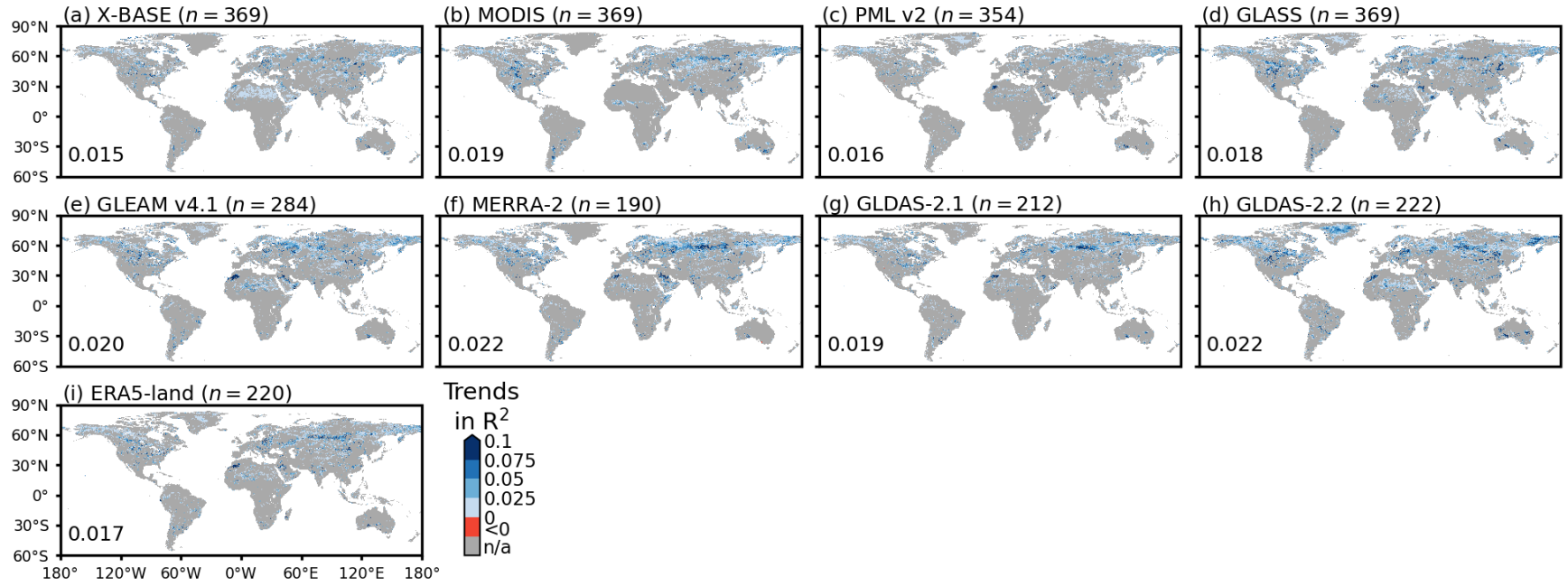


Fig. S26. Temporal changes in water balance consistency of *ET* datasets from 2000–2010 to 2011–2022. The consistency through water balance is quantified by R^2 scores in the two periods, and differences to indicate temporal changes are obtained (Methods). By accounting for bootstrap-based uncertainties, only the significant temporal changes are shown in blue/red. Global medians are shown in the left-lower corner.

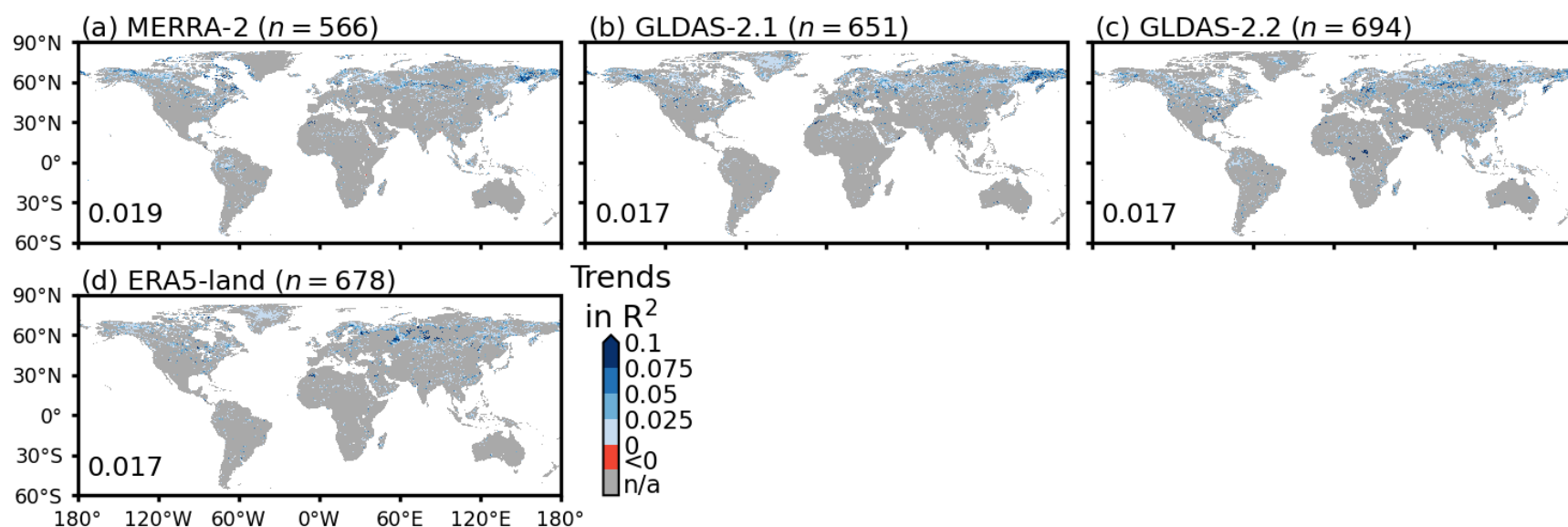


Fig. S27. Temporal changes in water balance consistency of R datasets from 2000–2010 to 2011–2022. The consistency through water balance is quantified by R^2 scores in the two periods, and differences to indicate temporal changes are obtained (Methods). By accounting for bootstrap-based uncertainties, only the significant temporal changes are shown in blue/red. Global medians are shown in the left-lower corner.

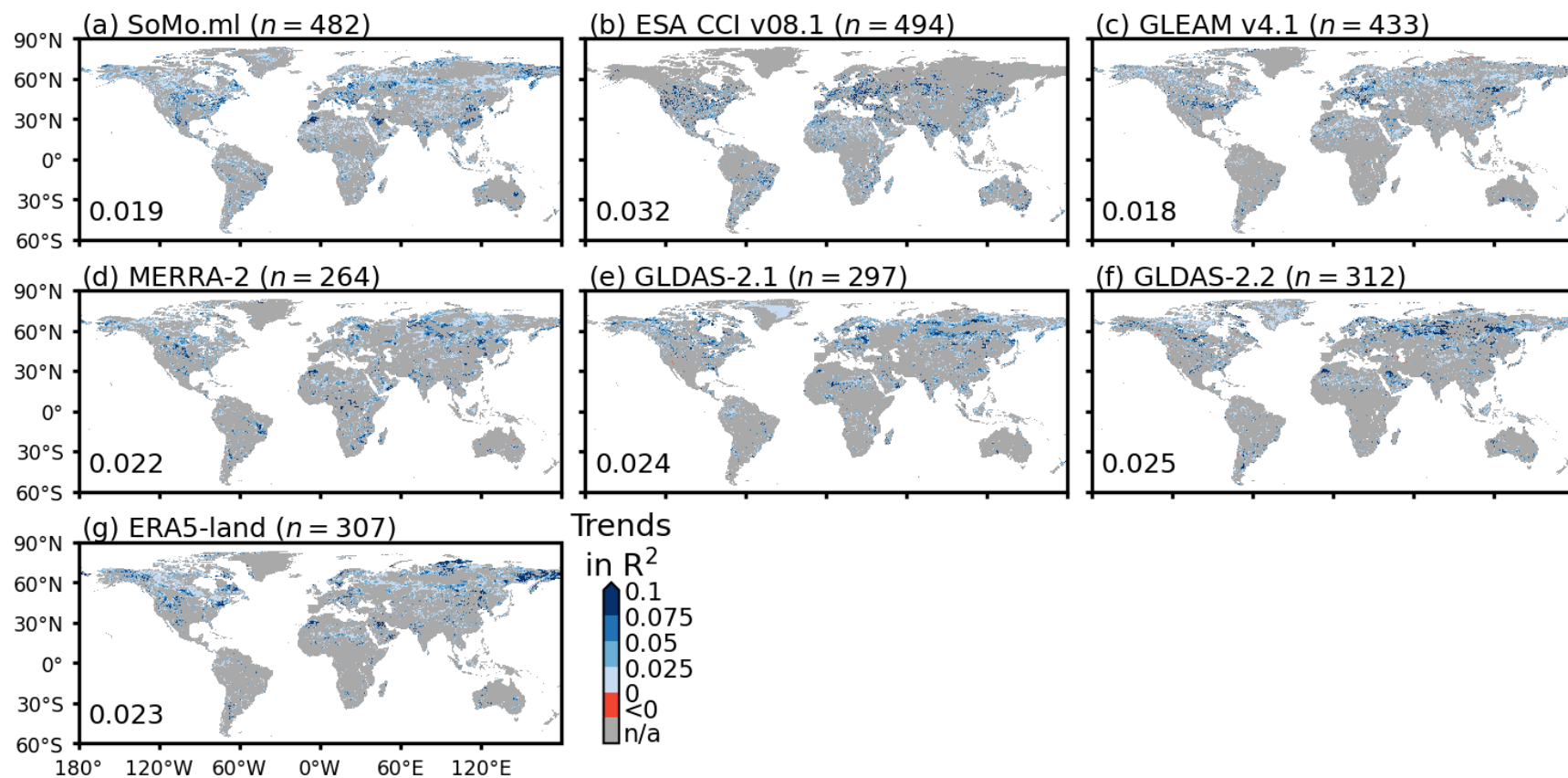


Fig. S28. Temporal changes in water balance consistency of *SM* datasets from 2000–2010 to 2011–2022. The consistency through water balance is quantified by R^2 scores in the two periods, and differences to indicate temporal changes are obtained (Methods). By accounting for bootstrap-based uncertainties, only the significant temporal changes are shown in blue/red. Global medians are shown in the left-lower corner.

References

- Adler, R. F., Sapiiano, M., Huffman, G. J., Wang, J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., Xie, P., Ferraro, R., and Shin, D. B.: The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation, *Atmosphere (Basel)*, 9, 10.3390/atmos9040138, 2018.
- Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies, *Bulletin of the American Meteorological Society*, 96, 69–+, 10.1175/Bams-D-13-00068.1, 2015.
- Chen, M. Y., Xie, P. P., Janowiak, J. E., and Arkin, P. A.: Global land precipitation: A 50-yr monthly analysis based on gauge observations, *Journal of Hydrometeorology*, 3, 249–266, Doi 10.1175/1525-7541(2002)003<0249:Glpaym>2.0.Co;2, 2002.
- Do, H. X., Gudmundsson, L., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM) - Part 1: The production of a daily streamflow archive and metadata, *Earth System Science Data*, 10, 765–785, 10.5194/essd-10-765-2018, 2018.
- Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, *Hydrology and Earth System Sciences*, 15, 1675–1698, 10.5194/hess-15-1675-2011, 2011.
- Fisher, J. B., Tu, K. P., and 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 Sensing of Environment*, 112, 901–919, 10.1016/j.rse.2007.06.025, 2008.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes, *Scientific Data*, 2, 10.1038/sdata.2015.66, 2015.
- Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), *Journal of Climate*, Volume 30, 5419–5454, 10.1175/JCLI-D-16-0758.1, 2017.
- Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: an observation-based global gridded runoff dataset from 1902 to 2014, *Earth System Science Data*, 11, 1655–1674, 10.5194/essd-11-1655-2019, 2019.
- Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology, *Earth System Science Data*, 11, 717–739, 10.5194/essd-11-717-2019, 2019.
- Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, *Scientific Data*, 7, 109, 10.1038/s41597-020-0453-3, 2020.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger,

- L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, *Q J Roy Meteor Soc*, 146, 1999-2049, 10.1002/qj.3803, 2020.
- Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: Global precipitation at one-degree daily resolution from multisatellite observations, *Journal of Hydrometeorology*, 2, 36-50, Doi 10.1175/1525-7541(2001)002<0036:Gpaodd>2.0.Co;2, 2001.
- Huffman, G. J., Bolvin, D. T., Joyce, R., Nelkin, E., Tan, J., Braithwaite, D., Hsu, K., Kelley, O. A., Nguyen, P., Sorooshian, S., Watters, D. C., West, B. J., and Xie, P.: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) Version 07, 2023.
- Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP-DOE AMIP-II reanalysis (R-2), *Bulletin of the American Meteorological Society*, 83, 1631-1643, 10.1175/Bams-83-11-1631(2002)083<1631:Nar>2.3.Co;2, 2002.
- Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, *Bulletin of the American Meteorological Society*, 82, 247-267, 10.1175/1520-0477(2001)082<0247:Tnnymr>2.3.Co;2, 2001.
- Lavers, D. A., Simmons, A., Vamborg, F., and Rodwell, M. J.: An evaluation of ERA5 precipitation for climate monitoring, *Q J Roy Meteor Soc*, 148, 3152-3165, 10.1002/qj.4351, 2022.
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, *Earth System Science Data*, 13, 4349-4383, 10.5194/essd-13-4349-2021, 2021.
- Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M., Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R. L., Zhang, W., Hamdi, Z. M., Reichstein, M., Aranda-Barranco, S., Ardö, J., Op de Beeck, M., Billesbach, D., Bowling, D., Bracho, R., Brümmer, C., Camps-Valls, G., Chen, S., Cleverly, J. R., Desai, A., Dong, G., El-Madany, T. S., Euskirchen, E. S., Feigenwinter, I., Galvagno, M., Gerosa, G. A., Gielen, B., Goded, I., Goslee, S., Gough, C. M., Heinesch, B., Ichii, K., Jackowicz-Korczynski, M. A., Klosterhalfen, A., Knox, S., Kobayashi, H., Kohonen, K.-M., Korkiakoski, M., Mammarella, I., Gharun, M., Marzuoli, R., Matamala, R., Metzger, S., Montagnani, L., Nicolini, G., O'Halloran, T., Ourcival, J.-M., Peichl, M., Pendall, E., Ruiz Reverter, B., Roland, M., Sabbatini, S., Sachs, T., Schmidt, M., Schwalm, C. R., Shekhar, A., Silberstein, R., Silveira, M. L., Spano, D., Tagesson, T., Tramontana, G., Trotta, C., Turco, F., Vesala, T., Vincke, C., Vitale, D., Vivoni, E. R., Wang, Y., Woodgate, W., Yezpe, E. A., Zhang, J., Zona, D., and Jung, M.: X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, *FLUXCOM-X, Biogeosciences*, 21, 5079-5115, 10.5194/bg-21-5079-2024, 2024.
- O, S. and Orth, R.: Global soil moisture data derived through machine learning trained with in-situ measurements, *Scientific Data*, 8, 170, 10.1038/s41597-021-00964-1, 2021.
- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., Poindexter, C.,

Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardo, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. V., Bonal, D., Bonnefond, J. M., Bowling, D. R., Bracho, R., Brodeur, J., Brummer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K. J., Cinti, B., Grandcourt, A., Ligne, A., De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., Tommasi, P. D., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrene, E., Dunn, A., Dusek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Goded, I., Goekede, M., Goldstein, A. H., Gough, C. M., Goulden, M. L., Graf, A., Griebel, A., Gruening, C., Grunwald, T., Hammerle, A., Han, S., Han, X., Hansen, B. U., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hortnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janous, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W. L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J. M., Lion, M., Liska, A. J., Lohila, A., Lopez-Ballesteros, A., Lopez-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Luers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H. A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Matthes, J. H., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A. M. S., Merbold, L., Meyer, W., Meyers, T., Miller, S. D., Minerbi, S., Moderow, U., Monson, R. K., Montagnani, L., Moore, C. E., Moors, E., Moreaux, V., Moureaux, C., Munger, J. W., Nakai, T., Neirynck, J., Nesic, Z., Nicolini, G., Noormets, A., Northwood, M., Nosoetto, M., Nouvellon, Y., Novick, K., Oechel, W., Olesen, J. E., Ourcival, J. M., Papuga, S. A., Parmentier, F. J., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R. P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S. M., Rambal, S., Rannik, U., Raz-Yaseef, N., Rebmann, C., Reed, D., Dios, V. R., Restrepo-Coupe, N., Reverter, B. R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S. R., Sanchez-Canete, E. P., Sanchez-Mejia, Z. M., Schmid, H. P., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R. L., Sedlak, P., Serrano-Ortiz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Sigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R. M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J. P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J. P., Walter-Shea, E., Wang, H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., and Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, *Sci Data*, 7, 225, 10.1038/s41597-020-0534-3, 2020.

Reichle, R. H. and Liu, Q.: Observation-Corrected Precipitation Estimates in GEOS-5, 2014.

Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J. M., Kimball, J. S., Ardizzone, J. V., Bosch, D., Colliander, A., Cosh, M., Kolassa, J., Mahanama, S. P., Prueger,

- J., Starks, P., and Walker, J. P.: Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product, *J Adv Model Earth Sy*, 11, 3106-3130, 10.1029/2019ms001729, 2019.
- Rui, H., Beaudoin, H., and Loeser, C.: README Document for NASA GLDAS Version 2 Data Products, 2022.
- Ruiz-Vásquez, M., O, S., Brenning, A., Koster, R. D., Balsamo, G., Weber, U., Arduini, G., Bastos, A., Reichstein, M., and Orth, R.: Exploring the relationship between temperature forecast errors and Earth system variables, *Earth System Dynamics*, 13, 1451-1471, 10.5194/esd-13-1451-2022, 2022.
- Running, S. W., Mu, Q., Zhao, M., and Moreno, A.: User's Guide MODIS Global Terrestrial Evapotranspiration (ET) Product, 2019.
- Schneider, U., Hänsel, S., Finger, P., Rustemeier, E., and Ziese, M.: GPCC Full Data Monthly Product Version 2022 at 0.25°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data, 10.5676/DWD_GPCC/FD_M_V2022_025, 2022.
- Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X. G., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Bronnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Dominguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Jourdain, S., Kaplan, A., Kubota, H., Le Blancq, F., Lee, T. C., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C. J., Moore, G. W. K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V. C., Smith, C. A., Tinz, B., Trewin, B., Valente, M. A., Wang, X. L., Wilkinson, C., Wood, K., and Wyszynski, P.: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system, *Q J Roy Meteor Soc*, 145, 2876-2908, 10.1002/qj.3598, 2019.
- Tang, G. Q., Clark, M. P., and Papalexiou, S. M.: EM-Earth The Ensemble Meteorological Dataset for Planet Earth, *Bulletin of the American Meteorological Society*, 103, E996-E1018, 10.1175/Bams-D-21-0106.1, 2022.
- Xie, P. P. and Arkin, P. A.: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bulletin of the American Meteorological Society*, 78, 2539-2558, 10.1175/1520-0477(1997)078<2539:Gpayma>2.0.Co;2, 1997.
- Xie, P. P., Chen, M., and Shi, W.: CPC unified gauge-based analysis of global daily precipitation, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc.,
- Xie, P. P., Joyce, R., Wu, S. R., Yoo, S. H., Yarosh, Y., Sun, F. Y., and Lin, R.: Reprocessed, Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998, *Journal of Hydrometeorology*, 18, 1617-1641, 10.1175/Jhm-D-16-0168.1, 2017.
- Zeri, M., Costa, J., Urbano, D., Cuartas, L., Ivo, A., Marengo, J., and Alvala, R.: A soil moisture dataset over the Brazilian semiarid region, *Mendeley Data Version*, 2, 2020.
- Zhang, Y. Q., Kong, D. D., Gan, R., Chiew, F. H. S., McVicar, T. R., Zhang, Q., and Yang, Y. T.: Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002-2017, *Remote Sensing of Environment*, 222, 165-182, 10.1016/j.rse.2018.12.031, 2019.