

**Evapotranspiration  
benchmark products**

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



# Benchmark products for land evapotranspiration: LandFlux-EVAL multi-dataset synthesis

**B. Mueller<sup>1</sup>, M. Hirschi<sup>1</sup>, C. Jimenez<sup>2</sup>, P. Ciais<sup>3</sup>, P. A. Dirmeyer<sup>4</sup>, A. J. Dolman<sup>5</sup>,  
J. B. Fisher<sup>6</sup>, M. Jung<sup>7</sup>, F. Ludwig<sup>8</sup>, F. Maignan<sup>3</sup>, D. Miralles<sup>9</sup>, M. F. McCabe<sup>10</sup>,  
M. Reichstein<sup>7</sup>, J. Sheffield<sup>11</sup>, K. C. Wang<sup>12</sup>, E. F. Wood<sup>11</sup>, Y. Zhang<sup>13</sup>, and  
S. I. Seneviratne<sup>1</sup>**

<sup>1</sup>Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

<sup>2</sup>LERMA, Observatoire de Paris, Paris, France

<sup>3</sup>LSCE, UMR CEA-CNRS, Gif-sur-Yvette, France

<sup>4</sup>George Mason University, Fairfax, Virginia, USA

<sup>5</sup>VU University Amsterdam, The Netherlands

<sup>6</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA

<sup>7</sup>Max Planck Institute for Biogeochemistry, Jena, Germany

<sup>8</sup>Wageningen University, Wageningen, The Netherlands

<sup>9</sup>School of Geographical Sciences, University of Bristol, UK

<sup>10</sup>Water Desalination and Reuse Center, King Abdullah University of Science and Technology,  
Saudi Arabia

<sup>11</sup>Department of Civil and Environmental Engineering, Princeton University, Princeton, USA

<sup>12</sup>College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

<sup>13</sup>CSIRO Land and Water, Canberra, Australia

Received: 19 December 2012 – Accepted: 6 January 2013 – Published: 17 January 2013

Correspondence to: B. Mueller (brigitte.mueller@env.ethz.ch),  
S. I. Seneviratne (sonia.seneviratne@env.ethz.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

**HESSD**

10, 769–805, 2013

---

## Evapotranspiration benchmark products

B. Mueller et al.

---

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Abstract

Land evapotranspiration (ET) estimates are available from several global datasets. Here, monthly global land ET synthesis products, merged from these individual datasets over the time periods 1989–1995 (7 yr) and 1989–2005 (17 yr), are presented. The merged synthesis products over the shorter period are based on a total of 40 distinct datasets while those over the longer period are based on a total of 14 datasets. In the individual datasets, ET is derived from satellite and/or in-situ observations (diagnostic datasets) or calculated via land-surface models (LSMs) driven with observations-based forcing and atmospheric reanalyses. Statistics for four merged synthesis products are provided, one including all datasets and three including only datasets from one category each (diagnostic, LSMs, and reanalyses). The multi-annual variations of ET in the merged synthesis products display realistic responses. They are also consistent with previous findings of a global increase in ET between 1989 and 1997 (1.15 mm yr<sup>-2</sup> in our merged product) followed by a decrease in this trend (–1.40 mm yr<sup>-2</sup>), although these trends are relatively small compared to the uncertainty of absolute ET values. The global mean ET from the merged synthesis products (based on all datasets) is 1.35 mm per day for both the 1989–1995 and 1989–2005 products, which is relatively low compared to previously published estimates. We estimate global runoff (precipitation minus ET) to 34 406 km<sup>3</sup> per year for a total land area of 130 922 km<sup>2</sup>. Precipitation, being an important driving factor and input to most simulated ET datasets, presents uncertainties between single datasets as large as those in the ET estimates. In order to reduce uncertainties in current ET products, improving the accuracy of the input variables, especially precipitation, as well as the parameterizations of ET are crucial.

## 1 Introduction

In recent years, several global multi-year evapotranspiration datasets based on in-situ observations or satellite retrievals of different indirect variables have been derived. In

HESSD

10, 769–805, 2013

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Mueller et al. (2011b), an evaluation of their performance within the LandFlux-EVAL (see [www.iac.ethz.ch/url/LandFlux-EVAL](http://www.iac.ethz.ch/url/LandFlux-EVAL)) initiative over the time period 1989–1995 was presented, while the study of Jimenez et al. (2011) assessed a subset of these datasets over a shorter period but also assessing the radiative and sensible fluxes.

5 These studies considered dedicated datasets that derive ET from combinations of observations or observations-based estimates together with targeted algorithms (referred to as *diagnostic datasets*), ET from land surface model (*LSM*) *simulations driven with observations-based forcing* as well as ET from atmospheric *reanalyses* (i.e. computed with LSMs within a global model assimilating mostly atmospheric observations).  
10 The general main geographical structures related to the principal climatic regimes are present in all products, but relatively large differences in the absolute values among some of the products were observed (Mueller et al., 2011b; Jimenez et al., 2011).

Even though a large number of ET datasets is currently available and has been analyzed in these studies, a global benchmark for ET is missing. Such a benchmark dataset would be useful for several purposes. Land-surface modellers and hydrologists often use ET to validate their model output, because it is one of the main components in the land water and energy budgets as well as a key driver for droughts (e.g. Sheffield et al., 2012; Seneviratne, 2012). Furthermore, agricultural and water-management communities estimate the water needed for irrigation with information on  
15 ET. Apart from mean ET values, corresponding uncertainty estimates are necessary for all kinds of applications. For these reasons, benchmark synthesis products of ET derived from existing datasets have been developed in the present study with the provision of different estimates of uncertainty.

25 The previous studies by Mueller et al. (2011b) and Jimenez et al. (2011) focused on spatial patterns of multi-year means and seasonal variations, respectively. However, the behavior of the LandFlux-EVAL datasets with respect to ET trends or multi-annual variations has not yet been investigated. Knowledge of the temporal changes of ET is important since it is a major component of the global water cycle. Within a changing climate, changes in the hydrological cycle are also expected, but very difficult

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



to determine. Observations indicate that precipitation over land increased by about 2.4 mm per decade from 1900 to 1988 (Dai et al., 1997, excluding North Africa in their analysis). Extending the analysis to the entire 20th century indicates a similar large trend (reduced by about 25 %, New et al., 2001). While some publications relate this behavior to a possible *intensification* of the hydrological cycle, this term is not well defined. Indeed, while evaporation from ocean surfaces is expected to increase with increasing temperature (as warmer air can hold more water vapour), it is unclear whether ET from land surfaces could similarly increase due to possible limitations imposed by soil moisture content and vegetation physiology. Due to a lack of relevant observations, respective trends of land ET could not be assessed until recently. The studies by Wang et al. (2010b) and Jung et al. (2010) are the first to investigate this issue over a relatively short time span from 1982 to 2002 and 1982 to 2008, respectively. Wang et al. (2010b) found an increase in global land ET of 15 mm per year, using 1120 globally distributed stations (Wang et al., 2010a). Jung et al. (2010) performed a trend analysis based on a global dataset empirically derived from in-situ measurements of ET from the FLUXNET project and satellite remote sensing and surface meteorological data (Jung et al., 2009, hereafter referred to as *MPIBGC* dataset), but also including eight other datasets. A tendency of increasing ET was found for the years 1982 to 1997, which indicates a possible intensification of the hydrological cycle. However, this trend was found to vanish after 1998. The decline in global land ET trend after 1998 was attributed to a decrease in moisture availability in Southern Hemisphere supply- (i. e. water-) limited evaporative regimes, which might indicate that a limit to the temperature-driven acceleration of the hydrological cycle was reached during the 1998–2008 time period. Nonetheless, the article also mentioned that whether this tendency was related to a long-term trend or only decadal variability could not be assessed given the short time period considered (see also Douville et al., 2013). Another study based on satellite retrievals also found that the increasing trend in global land ET disappeared after 2000 (Yao et al., 2012). However, it is important to note that uncertainties in forcing datasets used to derive such ET trends are large and may entail spurious features

---

**Evapotranspiration  
benchmark products**B. Mueller et al.

---

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

linked to the use of reanalyses products assimilating non-homogeneous satellite products or variations in the density of stations considered in gridded precipitation products (e.g. Bengtsson et al., 2004; Seneviratne et al., 2004; Lorenz and Kunstmann, 2012; Sheffield et al., 2012).

5 Decreasing ET in soil moisture-limited regions would lead to a further increase in air temperature, since more available energy would be partitioned into sensible heat. Thus, understanding changes in the hydrological cycle is not only important to reduce the uncertainty in climate change projections, but also to assess the impacts of these changes on water availability, as well as for the occurrence of droughts, floods and hot  
10 extremes (see, e.g. Sheffield et al., 2012; Mueller and Seneviratne, 2012; IPCC, 2011; Seneviratne et al., 2010).

The benchmark synthesis products presented in this study are used to assess the inter-annual variations of ET on the global scale and encompasses the largest number of ET products to date. Besides the evaluation of temporal variability of the benchmark  
15 products and the single datasets contributing to them, the present study also compares these to precipitation, which is one of the most important drivers of ET, especially in soil moisture-limited regions (see, e.g. Teuling et al., 2009; Seneviratne et al., 2010).

## 2 Datasets and method

### 2.1 Merged benchmark synthesis products of evapotranspiration

20 We present here new multi-year merged synthesis products based on the analyses of existing land ET datasets. A first product spans the time period 1989–1995 and includes 40 products, while the second is available for the longer time period 1989–2005 and includes 14 products. Consistent with a previous analysis (Mueller et al., 2011b), the type of datasets included can be classified as diagnostic datasets, LSMs and reanalyses (see Sect. 1). Besides the two merged synthesis products based on  
25 all types of datasets, merged synthesis products from each of the individual dataset

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



types are also produced (see Table 1). The output statistics for each of the merged synthesis products are: Mean, median, 25th-percentile, 75th-percentile, interquartile range, standard deviation and minimum and maximum values of the ensemble of underlying datasets. All products are available in monthly and yearly temporal resolution, and as multi-year statistics. All merged synthesis products are made available through the internet ([www.iac.ethz.ch/url/LandFlux-EVAL](http://www.iac.ethz.ch/url/LandFlux-EVAL)).

## 2.2 Overview of included datasets

An overview of the diagnostic datasets, LSMs and reanalyses considered for the preparation of the merged synthesis products is provided in Table 2. The subset of datasets available over the period 1989–2005 (cross in 5th column in Table 2) forms the basis of the merged synthesis products over this longer time period (see also Sect. 2.1). The table also lists information on the single datasets, such as the ET schemes, the number of soil layers used in the case of LSMs, the precipitation forcing datasets and other forcing variables used for the derivation of the respective datasets or, in the case of reanalyses, the land-surface schemes.

We considered here several additional datasets compared to the earlier analysis of Mueller et al. (2011b). These additional datasets are the diagnostic dataset GLEAM (Global Land-surface Evaporation: The Amsterdam Methodology, Miralles et al., 2011a), as well as LSM estimates from the Water Model Intercomparison Project WaterMIP (Haddeland et al., 2011). Simulations from the Global Land Data Assimilation System I (GLDAS-I, Rodell et al., 2004) were included in Mueller et al. (2011b) but excluded in the present study because of spurious trends (see Fig. A1 and Rui, 2011), which arised because the source of forcing data changed several times over the GLDAS-I time period (Matt Rodell, personal communication, 2012). However, we included GLDAS-II simulations (see Rui, 2011) from one of these models (NOAH version 3.3) which has been produced recently with a consistent forcing dataset (Princeton forcing, see Sheffield et al., 2006).

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



---

**Evapotranspiration  
benchmark products**


---

B. Mueller et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

In GLEAM, the calculation of ET is based on the Priestley-Taylor equation and the Gash analytical model of forest rainfall interception (Miralles et al., 2011b). The model discriminates the different evapotranspiration components, i. e. interception, bare soil evaporation, transpiration and sublimation, and ET is coupled to soil moisture (Miralles et al., 2011a). Note that not all diagnostic estimates separately calculate these components or account for all of them, which leads to large differences especially in the Amazon region. The forcing data for GLEAM were all obtained from remote sensing products and synthesis of rain gauges (CPC, see Appendix A).

All WaterMIP simulations are driven with the same forcing dataset (WATCH forcing, see Weedon et al., 2011), but the employed forcing variables and time steps differ. For a list of these variables as well as references for each model, see Haddeland et al. (2011). The differences between the WaterMIP models are large. Some models, for example, solve both the water and the energy balances at the land surface and are classified as (classical) *LSMs*, while others solve the water balance only and are classified as *global hydrological models*, *GHMs* (following the classification proposed by Haddeland et al., 2011, note that for simplicity, we refer to both as *LSMs* in most of the present article). Further, the WaterMIP models vary substantially in their complexity in the representation of ET (e.g. including or excluding interception and transpiration), runoff, groundwater, snow or frozen soil (for more details, see Haddeland et al., 2011). For more information on all other datasets, the reader is referred to Mueller et al. (2011b) and Jimenez et al. (2011).

### 2.3 Processing of ET datasets and merged synthesis products

In order to prepare the merged synthesis products, we first interpolated all datasets on a common global grid of 1 degree longitude and latitude and aggregated daily values to monthly values where necessary. A spatial matching of the datasets was done, and if one gridpoint was covered by less than 70 % of the datasets, it was excluded from the final synthesis product (for the number of datasets originally available, see Fig. A2). Some of the datasets exhibit unrealistically large values (especially in the northern

latitudes due to the viewing angle of satellites). For the merged synthesis products, we applied a physical constraint to exclude such values. An upper limit to the latent heat flux is given by the energy balance, i. e. ET should not exceed net surface radiation at a scale as large as our grid cells. For each gridpoint of the merged synthesis products, we calculated a seasonal cycle of net radiation (from the Surface Radiation Budget (SRB) version 3) based on the monthly maxima of all available years (1984–2007). Monthly ET values exceeding the seasonal maxima cycle's net radiation of that month by more than 25 % were excluded, unless ET was smaller than  $0.3 \text{ mm day}^{-1}$ , since for such small values, ground heat flux cannot be neglected. Note, however, that if interception plays an important role, such as during winter time, ET can be larger than radiation. A further possible constraint might be applied from the assumption that ET should not exceed precipitation over a longer time period. However, we did not apply such a constraint because soil moisture depletion might play a role in some regions, and based on a small scale analysis (such as single pixels), atmospheric water fluxes or runoff could provide additional water input for ET. In order to exclude single dataset values that were very different from those of the other datasets, we performed a statistical outlier detection after the application of the physical constraint, similar to that described in Weedon (2011), but applied on monthly values. A movie in the Supplement shows the number of datasets at each gridpoint and time step after all these steps. Finally, the mean, median, 25th-percentile, 75th-percentile, interquartile range, standard deviation and minimum and maximum statistics of the ensemble of underlying datasets are derived and provided as monthly, yearly and multi-year statistics.

### 3 Results

#### 3.1 Merged synthesis products

The different merged synthesis products created from single categories only (*diagnostic datasets*, *LSMs* and *reanalyses*) and from all categories (see Table 1) coincide to

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



5 a large extent in their global land mean ET (Fig. 1), with highest values in the merged product based on reanalyses only (563 mm per year) and lowest in that based on LSMs only (423 mm per year). The interquartile ranges (IQRs, 75th-percentile minus 25th-percentile) are largest in the merged products based on diagnostic datasets and reanalyses. The variation of global mean ET for the 1989–2005 (*long*) as well as for the 1989–1995 (*short*) merged product created from all dataset categories (median) is shown in Fig. 2. The long merged product shows slightly higher values. The largest difference in the list of datasets in the short and long merged synthesis products is the inclusion of 28 LSMs (*short*) versus only 5 LSMs (*long*). WaterMIP and GSWP simulations are not available for the *long* version, and are therefore, due to their consistently low ET values (see Mueller et al., 2011b), the main reason for lower ET in the *short* product. The small difference in the temporal variations between the *short* and the *long* merged synthesis products is a strong indication that including a large number of dependent datasets (i.e. model simulations driven with the same forcing data, such as GSWP and WaterMIP runs) does not have a strong influence.

10 Global mean ET shows a slight increase between 1989 and 1997 followed by a decrease until 2005 (Fig. 2). The merged synthesis product (*long*) shows a nearly identical inter-annual variation as that found in the MPIBGC dataset in Jung et al. (2010). However, if we consider this variation in relation to the IQRs or the standard deviations, both shown in Fig. 2, the absolute ET trend change is very small and the interannual variations nearly vanish.

15 The reason for the large IQRs and standard deviations are the large differences in the absolute ET values of the single datasets. The IQRs and standard deviations based on the yearly anomalies of the underlying datasets (i. e. setting the mean of all datasets to zero before calculating the statistics), which is the quantity shown in Jung et al. (2010), are much smaller (can partly be seen also in Fig. 3). Note also that we consider more estimates than in the previous analyses from Jung et al. (2010).

20 The ET anomalies from all long merged synthesis products are shown in Fig. 3 (top left). The comparison reveals a very similar temporal evolution of ET in all four merged

---

**Evapotranspiration  
benchmark products**B. Mueller et al.

---

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

synthesis products. Therefore, in the remainder of this study, only the merged products based on all dataset categories (long and short) will be analysed.

### 3.2 Single datasets

The temporal variations of the 14 single datasets contributing to the long merged synthesis products are shown in Fig. 3. In these analyses of single datasets, we excluded unrealistically high ET values, setting a threshold of  $12.5 \text{ mm day}^{-1}$ . The LSMs (bottom left) and reanalyses (bottom right) are more consistent amongst one another in their yearly variations than the diagnostic datasets (top right). The ET timeseries of all LSMs and reanalyses peak between 1997 and 1999. Some of the diagnostic datasets peak in other years, such as 2001 in the case of PRUNI and 2000 in GLEAM and AWB. The trends for the two time periods 1989–1997 and 1998–2005 are listed in Table 3. The merged product as well as 5 single datasets display a significant negative trend (italic font) for 1998–2005, indicating a decrease in global ET during that period. The positive trend found in Jung et al. (2010) for the previous period is only significantly positive in GLEAM. The reason for this could be that we calculate the trends over a shorter time period compared to Jung et al. (2010), who calculated them over 1982–1997 and 1998–2008.

### 3.3 Analyses of climate regions

We analyze here the two merged products (i.e. short and long, based on all dataset categories) as well as precipitation data (average of CRU, GPCC, GPCP and CPC, for references and information on these datasets, see Appendix A and Biemans et al., 2009) in climate regions using the classification of Koeppen-Geiger (data available from <http://koeppen-geiger.vu-wien.ac.at>). In order to facilitate the interpretation of the results, subregions have been merged to larger regions. The regions considered are displayed in Fig. 4.

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Mean ET and precipitation are listed for the various climate regions in Table 4. Also included are the ET and precipitation trends (Theil-Sen estimator) from 1998–2005, i. e. for the period for which a decline in ET trend was found in Jung et al. (2010, see also previous section). The sum of the areas of all climate regions (third column) represents over 90 % of the global land area.

The global mean ET from the merged synthesis amounts to  $1.35 \text{ mm day}^{-1}$  for both the 1989–2005 and 1989–1995 products. This value is well within the range, and sometimes at the lower boundary, of other studies. For example Trenberth et al. (2009) reported a range of  $1.38$  to  $1.82 \text{ mm day}^{-1}$ , Haddeland et al. (2011)  $1.14$  to  $1.61$ , Wang and Dickinson (2012)  $1.2$  to  $1.5$  and Dirmeyer et al. (2006) a mean of  $1.36 \text{ mm day}^{-1}$  for different time periods. The larger values from Trenberth et al. (2009) can be explained by their reliance on reanalysis products, which were found here to display a tendency for high ET values. The mean value of precipitation (average of CRU, GPCC, GPCP and CPC) amounts to  $2.07 \text{ mm day}^{-1}$ . The difference between global precipitation and land ET corresponds to the water that leaves the continents as runoff and amounts to  $34\,406 \text{ km}^3$  per year. This value is in good agreement with values from other studies summarized in Syed et al. (2009).

The largest contribution to the global ET trend over 1998–2005, which amounts to  $18.9 \text{ km}^3 \text{ yr}^{-2}$ , stems from the equatorial winter dry (Aw), arid desert (BW) and arid steppe (BS) climate regions, even though the latter two are characterized by very low per area values of ET. The study of Jung et al. (2010) showed that the decline in trend change is mainly due to Southern Hemisphere dry regions. We therefore treated the northern and Southern Hemisphere of these regions (BW and BS) separately. Indeed, we find that even though they belong to the same climate regions, the Southern Hemisphere parts of the arid steppe (BS) and arid desert (BW) regions exhibit a large negative trend, while the Northern Hemisphere parts show very small (and positive) trends.

The signs of the trends in precipitation agree with the signs of the ET trends, except for the polar climate region (E). The opposite trends in the northern and southern

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



hemisphere parts of the BS and BW regions can also be found in the precipitation datasets. Furthermore, the table shows that global ET has decreased much stronger than global precipitation over the period 1998–2005.

### 3.4 Precipitation forcing

5 The 1989–1995 global mean land ET of each dataset contributing to the synthesis product (short) is plotted against precipitation in Fig. 5. The precipitation value was taken from the forcing data of the respective ET dataset as listed in Table 2. If precipitation was not available (for some diagnostic datasets), the average of four currently available observational datasets (CRU, GPCP, GPCC and CPC) was taken. Global mean values of these four precipitation datasets range from 2 to 2.2 mm day<sup>-1</sup>. The dataset-median of the merged synthesis ET product is indicated with a solid line, and the IQRs with dash-dotted lines. The single datasets are indicated with different symbols (groups) and colors (ET schemes).

10 We first compare simulations from the GSWP and the WaterMIP projects, which are each based on common forcing datasets (filled circles and stars/rhombi, respectively). The spread within the GSWP and WaterMIP simulations is similar, both globally and in most climate regions (see Fig. A3). However, the spread in the WaterMIP ensemble is smaller in some dry regions (Cs, Dw and Df), and larger in wetter regions (all equatorial regions). Looking at the WaterMIP *GHMs* and *LSMs* separately, we find that the *GHMs* (stars) are not separated from the *LSMs* (rhombi), which supports the findings from Haddeland et al. (2011), that this classification does not fully account for differences among the WaterMIP models.

15 In order to compare the influence of uncertainties in precipitation forcing to model structure, sensitivity simulations using the same model (here, the COLA model) and different precipitation forcing have been performed in the framework of GSWP (Schlosser and Gao, 2010) and are included in the Fig. A3. Evapotranspiration from simulations with differing precipitation (GSWP sens, noted with empty circles) shows a smaller range than from GSWP simulations from different models using the same forcing (filled

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



circles), which has also been shown in Schlosser and Gao (2010). However, note that global mean ET from these sensitivity simulations is relatively low, indicating dry conditions in the COLA model, even if a forcing with high precipitation was employed. This possibly points to a dry bias of the model independently of the applied precipitation forcing, which could be the reason for the separation of this GSWP model in the cluster analyses reported in Mueller et al. (2011b).

The merged synthesis product based on all datasets exhibits a value of  $1.5 \text{ mm day}^{-1}$ . Note that the global mean values in the analyses for Fig. 5 are higher than the ones given in Table 4. The reason is that for the analyses of single datasets, we only included those pixels of the merged product that were also available in all other datasets. Table 4, on the other hand, includes all land pixels.

The largest exceedance of precipitation over ET, on average, is found in the wettest climate regimes (Af, Am, Aw, Cw, Cf and Df), as expected. In several dry regions, especially the arid desert (BW) and arid steppe (BS) regions, some datasets reveal an ET exceedance over precipitation (see bisecting line through origin in Fig. A3). The reasons could be (1) ET is too high, (2) precipitation is too low, (3) both ET and precipitation are correct, but the net depletion of soil water storage is larger than the volume of runoff generated over the period 1989–1995.

A comparison of the range between the lowest and highest values in precipitation and ET shows that the uncertainties in precipitation are larger than those of the ET datasets. This is not only the case for the global mean values, but also for single climate regions (Fig. A3). Large uncertainties in precipitation datasets have also recently been highlighted in Lorenz and Kunstmann (2012). The reason for smaller uncertainties in ET than in precipitation could be that ET estimates are constrained not only by the water, but also by the energy balance. This indicates that the uncertainty range in ET estimates will be difficult to reduce as long as the uncertainties in precipitation and radiation are not reduced. Jimenez et al. (2011), e.g. showed that the spread in net radiation datasets is nearly as large as the one in ET.

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## 4 Conclusions

The intensity of the hydrological cycle determines the water availability and influences the climate system in various ways. Despite the important impacts of the hydrological cycle and one of its key variable, ET, a global benchmark ET dataset has long been missing. In the framework of the LandFlux-EVAL initiative ([www.iac.ethz.ch/url/LandFlux-EVAL](http://www.iac.ethz.ch/url/LandFlux-EVAL)), several ET datasets based on observations (diagnostic datasets, LSMs and reanalyses) have been evaluated in previous studies (Mueller et al., 2011b; Jimenez et al., 2011), focusing on multi-annual means and seasonal cycles. The present study further investigates ET datasets. Global merged benchmarking ET products are derived and trends are analyzed in single LandFlux-EVAL datasets as well as the merged ET products.

The benchmark synthesis products provide monthly estimates for the time periods 1989–1995 (*short*) and 1989–2005 (*long*), respectively. For the creation of the *short* benchmark products, 7 diagnostic datasets, 29 LSMs and 4 reanalyses are considered, for the *long* products 5 diagnostic, 5 LSMs and 4 reanalyses. In order to address several demands on benchmark datasets, we created *short* and *long* merged synthesis products based on all datasets as well as based on each category. Monthly radiation is used as a physical constraint on maximum ET, and a statistical outlier detection is applied on the monthly ET estimates.

Evapotranspiration from the merged synthesis benchmark products shows realistic interannual variations that correspond to those found in a previous study based on a smaller number of ET datasets (Jung et al., 2010). The negative trend in *global land ET*<sup>1</sup> between 1998–2005 amounts to  $18.9 \text{ km}^3 \text{ yr}^{-2}$ . Most of this trend is attributed to the equatorial winter dry, arid desert and arid steppe regions. The latter two regions are determined by low per area ET and precipitation, but cover very large areas of the northern and Southern Hemisphere. Dividing these arid desert and steppe regions into

<sup>1</sup>After a space and time matching of all datasets, data coverage of roughly 90 % of the land surface was attained.

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



northern and Southern Hemisphere fractions, we find that the negative trend change arises from the southern part only, which is consistent with the results of Jung et al. (2010). However it is important to note that the signal is very small compared to the overall global land ET as well as compared to the uncertainty of absolute ET values (interquartile range or standard deviations of the merged synthesis products). In addition, it is still unclear whether this signal corresponds to a long-term trend or decadal variability. Finally, because of the reliance of all ET datasets on atmospheric input datasets, the influence of spurious trends in these datasets cannot be excluded.

Large uncertainties in absolute values of ET are found, which can partly be related to uncertainties in precipitation. Precipitation is both one of the main drivers for ET in water-limited evaporation regimes and overall in forests where interception can be large. As a consequence, it belongs to one of the main forcing variables for ET used in most diagnostic datasets and LSMs. Indeed, the spread in ET datasets is smaller than the spread in the corresponding precipitation datasets in our global analyses as well as in most climatic regions, which indicates that ET is not only constrained by precipitation, but also by other variables such as radiation. In general, the absolute values of precipitation are higher than ET, as expected, globally and in wet climate regions. Global mean ET in the merged synthesis product amounts to 1.35 mm per day, while precipitation to 2.07 mm per day (average of four observations-based datasets). The difference of  $34\,406\text{ km}^3\text{ yr}^{-1}$  (runoff) is in agreement with previous studies (an overview can be found in Syed et al., 2009). In dry regions, ET exceeds precipitation in several datasets. The merged synthesis product's (median) ET is always lower than average precipitation.

In summary, we have presented here the first benchmark synthesis products for monthly, global land ET estimates. A reproduction of a negative trend in global ET during 1998–2005 with these benchmark synthesis products supports previous findings of a declining global ET trend over that period. However, caution is necessary when analyzing trends, because the considered time period is very short for trend analyses, the analyzed ET datasets are not totally independent from each other (e.g. same forcing

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



data, similar methodologies), and agreement between them is not necessarily an indicator of their validity. Furthermore, spurious trends can be introduced through changes in the observing systems for the forcing variables (e.g. precipitation, radiation) of ET. In order to gain more confidence in ET estimates, not only are improvements in model parameterizations necessary, but so is a reduction of uncertainties in precipitation and radiation data in order to better constrain ET.

## Appendix A

### Precipitation datasets

A The observation-based precipitation datasets are from the Climate Research Unit (CRU) at the University of East Anglia, the Global Precipitation Climatology Centre (GPCC), the Global Precipitation Climatology Project (GPCP) and the unified gauge-based analysis of global daily precipitation from the climate prediction center (CPC) from the National Oceanic and Atmospheric Administration (NOAA Chen et al., 2008). These datasets are chosen for this investigation because (a) they are mainly based on observations, (b) they cover the period 1989–2005, and (c) they are forcing datasets employed for the diagnostic ET datasets used in this study.

The CRU precipitation data are based on rain gauge data, whose number varies over time between around 5000 and nearly 15 000 stations. The CRU TS3.1 dataset covers the period 1901–2009. It has not been corrected for gauge biases, which vary with gauge type and can result in inhomogeneities in the records (New et al., 2000).

The NOAA CPC unified precipitation dataset is created from quality-controlled daily precipitation gauge data, taking advantage of the optimal interpolation objective analysis technique (Chen et al., 2008). The retrospective version, covering 1979 to 2005, includes more than 30 000 gauge station data.

The GPCP monitoring product for the period 1986 to present is based on quality-controlled data from 7000 stations, which are interpolated into monthly area averages.

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



This product delivers the in-situ component for the satellite (microwave and infrared)-gauge combination GPCP (Huffman et al., 1995; Adler et al., 2003). The GPCP product includes gauge-bias corrections, but due to the limited length of satellite records, inhomogeneities arise (Adler et al., 2003).

5 **Supplementary material related to this article is available online at:**  
**<http://www.hydrol-earth-syst-sci-discuss.net/10/769/2013/hessd-10-769-2013-supplement.zip>**

10 *Acknowledgements.* The LandFlux-EVAL project was coordinated under the Global Energy and Water Exchanges (GEWEX) LandFlux initiative. We also acknowledge the support of the Integrated Land Ecosystem-Atmosphere Processes Study, iLEAPS, in the development of the synthesis products. CRU data were obtained from the University of East Anglia Climate Research Unit (CRU), British Atmospheric Data Centre, 2008, available from <http://badc.nerc.ac.uk/data/cru>. The GPCP combined precipitation data were developed and computed by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres as a contribution to the GEWEX Global Precipitation Climatology Project. GPCC precipitation data are available from the GPCC Homepage: <http://gpcc.dwd.de>. CPC merged analysis of precipitation data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>. We acknowledge the Global Modeling and Assimilation Office and the GES DISC for the dissemination of MERRA and MERRA-LAND, and the ECMWF for the dissemination of ERA-Interim data. The CFSR data are from the Research Data Archive which is maintained by the Computational and Information Systems Laboratory at the National Center for Atmospheric Research (NCAR). NCAR is sponsored by the National Science Foundation. The original data are available from the RDA (<http://dss.ucar.edu>) in dataset number ds093.0. We would further like to acknowledge the Japanese 25-yr ReAnalysis and JMA Climate Data Assimilation System (JCDAS) for the dissemination of JRA-25 data. JBF contributed to this work at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. In addition, several authors were partially funded by EU-PF7 projects. We would like to thank Zhichang Guo from COLA for providing the GSWP sensitivity simulations.

---

**Evapotranspiration benchmark products**

B. Mueller et al.

---

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## References

- Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P. P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E.: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present), *J. Hydrometeorol.*, 4, 1147–1167, 2003. 786
- 5 Bengtsson, L., Hagemann, S., and Hodges, K. I.: Can climate trends be calculated from re-analysis data?, *J. Geophys. Res.-Atmos.*, 109, D11111, doi:10.1029/2004JD004536, 2004. 774
- Biemans, H., Hutjes, R. W. A., Kabat, P., Strengers, B. J., Gerten, D., and Rost, S.: Effects of precipitation uncertainty on discharge calculations for main river basins, *J. Hydrometeorol.*, 10, 1011–1025, doi:10.1175/2008JHM1067.1, 2009. 779
- 10 Bosilovich, M. G.: NASA's modern era retrospective-analysis for research and applications: Integrating Earth observations, available at: [www.earthzine.org/2008/09/26/nasas-modern-era-retrospective-analysis](http://www.earthzine.org/2008/09/26/nasas-modern-era-retrospective-analysis) (last access: 20 December 2012), 26 September 2008. 795
- 15 Chen, M. Y., Shi, W., Xie, P. P., Silva, V. B. S., Kousky, V. E., Higgins, R. W., and Janowiak, J. E.: Assessing objective techniques for gauge-based analyses of global daily precipitation, *J. Geophys. Res.-Atmos.*, 113, D04 110. doi:10.1029/2007JD009132, 2008. 785
- Dai, A., Fung, I. Y., and DelGenio, A. D.: Surface observed global land precipitation variations during 1900–88, *J. Climate*, 10, 2943–2962, doi:10.1175/1520-0442(1997)010<2943:SOGLPV>2.0.CO;2, 1997. 773
- 20 Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteorol. Soc.*, 137, 553–597, 2011. 795
- 25 Dirmeyer, P. A., Gao, X. A., Zhao, M., Guo, Z. C., Oki, T. K., and Hanasaki, N.: GSWP-2 – multimodel analysis and implications for our perception of the land surface, *B. Am. Meteorol. Soc.*, 87, 1381–1397, 2006. 780, 794
- 30

# HESSD

10, 769–805, 2013

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



---

**Evapotranspiration  
benchmark products**


---

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Douville, H., Ribes, A., Decharme, B., Alkama, R., and Sheffield, J.: Anthropogenic influence on multidecadal changes in reconstructed global evapotranspiration, *Nature Climate Change*, 3, 59–62, doi:10.1038/nclimate1632, 2013. 773
- 5 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 Sens. Environ.*, 112, 901–919, doi:10.1016/j.rse.2007.06.025, 2008. 794
- Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voss, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., and Yeh, P.: Multimodel Estimate of the global terrestrial water balance: setup and first results, *J. Hydrometeorol.*, 12, 869–884, doi:10.1175/2011JHM1324.1, 2011. 775, 776, 780, 781, 794
- 10 Huffman, G. J., Adler, R. F., Rudolf, B., Schneider, U., and Keehn, P. R.: Global precipitation estimates based on a technique for combining satellite-based estimates, rain-gauge analysis, and NWP model precipitation information, *J. Climate*, 8, 1284–1295, doi:10.1175/1520-0442(1995)008<1284:GPEBOA>2.0.CO;2, 1995. 786
- IPCC: Summary for policymakers, in: Intergovernmental Panel on Climate Change Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, Cambridge University Press, edited by: Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M., Cambridge, UK and New York, NY, USA, 2011. 774
- 20 Jimenez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., Rossow, W. B., Balsamo, G., Betts, A. K., Dirmeyer, P. A., Fisher, J. B., Jung, M., Kanamitsu, M., Reichle, R. H., Reichstein, M., Rodell, M., Sheffield, J., Tu, K., and Wang, K.: Global intercomparison of 12 land surface heat flux estimates, *J. Geophys. Res.-Atmos.*, 116, D02102, doi:10.1029/2010JD014545, 2011. 772, 776, 782, 783
- 25 Jung, M., Reichstein, M., and 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, doi:10.5194/bg-6-2001-2009, 2009. 773, 794
- 30 Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J. Q., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q. Z., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Rouspard, O., Running, S., Tomelleri, E., Viovy, N.,

---

**Evapotranspiration  
benchmark products**


---

B. Mueller et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, *Nature*, 467, 951–954, doi:10.1038/nature09396, 2010. 773, 778, 779, 780, 783, 784

5 Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cy.*, 19, GB1015, doi:10.1029/2003GB002199, 2005. 795

10 Lorenz, C. and Kunstmann, H.: The hydrological cycle in three state-of-the-art reanalyses: inter-comparison and performance analysis, *J. Hydrometeor.*, 13, 1397–1420, doi:10.1175/JHM-D-11-088.1, 2012. 774, 782

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J.: Magnitude and variability of land evaporation and its components at the global scale, *Hydrol. Earth Syst. Sci.*, 15, 967–981, doi:10.5194/hess-15-967-2011, 2011a. 775, 776, 794

15 Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, *Hydrol. Earth Syst. Sci.*, 15, 453–469, doi:10.5194/hess-15-453-2011, 2011b. 776, 794

Mueller, B. and Seneviratne, S. I.: Hot days induced by precipitation deficits at the global scale, *P. Natl. Acad. Sci. USA*, 109, 12398–12403, doi:10.1073/pnas.1204330109, 2012. 774

20 Mueller, B., Hirschi, M., and Seneviratne, S. I.: New diagnostic estimates of variations in terrestrial water storage based on ERA-Interim data, *Hydrol. Process.*, 25, 996–1008, doi:10.1002/hyp.7652, 2011a. 794

Mueller, B., Seneviratne, S. I., Jimenez, C., Corti, T., Hirschi, M., Balsamo, G., Ciais, P., Dirmeyer, P., Fisher, J. B., Guo, Z., Jung, M., Maignan, F., McCabe, M. F., Reichle, R., Reichstein, M., Rodell, M., Sheffield, J., Teuling, A. J., Wang, K., Wood, E. F., and Zhang, Y.: Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophys. Res. Lett.*, 38, L06402, doi:10.1029/2010GL046230, 2011b. 772, 774, 775, 776, 778, 782, 783

25 New, M., Hulme, M., and Jones, P.: Representing twentieth-century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate, *J. Climate*, 13, 2217–2238, 2000. 785

30 New, M., Todd, M., Hulme, M., and Jones, P.: Precipitation measurements and trends in the twentieth century, *Int. J. Climatol.*, 21, 1899–1922, 2001. 773

---

**Evapotranspiration  
benchmark products**


---

B. Mueller et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Onogi, K., Tsltsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kaalhoru, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Taira, R.: The JRA-25 reanalysis, *J. Meteorol. Soc. Jpn.*, 85, 369–432, 2007. 795
- 5 Reichle, R., Koster, R., De Lannoy, G., Forman, B., Liu, Q., Mahanama, S., and Toure, A.: Assessment and enhancement of MERRA land surface hydrology estimates, *J. Climate*, 24, 6322–6338, doi:10.1175/JCLI-D-10-05033.1, 2011. 795
- Rodell, M., Famiglietti, J. S., Chen, J., Seneviratne, S. I., Viterbo, P., Holl, S., and Wilson, C. R.: Basin scale estimates of evapotranspiration using GRACE and other observations, *Geophys. Res. Lett.*, 31, L20504, doi:10.1029/2004GL020873, 2004. 775
- 10 Rui, H.: README Document for Global Land Data Assimilation System Version 2 (GLDAS-2) Products, GES DISC, 2011. 775
- Saha, S., Moorthi, S., Pan, H. L., Wu, X. R., Wang, J. D., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H. X., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y. T., Chuang, H. Y., Juang, H. M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H. L., Yang, R. Q., Lord, S., Van den Dool, H., Kumar, A., Wang, W. Q., Long, C., Chelliah, M., Xue, Y., Huang, B. Y., Schemm, J. K., Ebisuzaki, W., Lin, R., Xie, P. P., Chen, M. Y., Zhou, S. T., Higgins, W., Zou, C. Z., Liu, Q. H., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, *B. Am. Meteorol. Soc.*, 91, 1015–1057, doi:10.1175/2010BAMS3001.1, 2010. 795
- 15 Schlosser, C. A. and Gao, X. A.: Assessing Evapotranspiration Estimates from the Second Global Soil Wetness Project (GSWP-2) Simulations, *J. Hydrometeorol.*, 11, 880–897, doi:10.1175/2010JHM1203.1, 2010. 781, 782
- 25 Seneviratne, S. I.: Climate science: historical drought trends revisited, *Nature*, 491, 338–339, doi:10.1038/491338a, 2012. 772
- Seneviratne, S. I., Viterbo, P., Lüthi, D., and Schär, C.: Inferring changes in terrestrial water storage using ERA-40 reanalysis data: the Mississippi River basin, *J. Climate*, 17, 2039–2057, 2004. 774
- 30 Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture-climate interactions in a changing climate: a review, *Earth-Sci. Rev.*, 99, 125–161, doi:10.1016/j.earscirev.2010.02.004, 2010. 774

---

**Evapotranspiration  
benchmark products**


---

B. Mueller et al.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


- Sheffield, J. and Wood, E. F.: Characteristics of global and regional drought, 1950–2000: analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle, *J. Geophys. Res.-Atmos.*, 112, D17115-1-21, doi:10.1029/2006JD008288, 2007. 795
- 5 Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *J. Climate*, 19, 3088–3111, doi:10.1175/JCLI3790.1, 2006. 775, 794, 795
- Sheffield, J., Wood, E. F., and Munoz-Arriola, F.: Long-term regional estimates of evapotranspiration for Mexico based on downscaled ISCCP data, *J. Hydrometeorol.*, 11, 253–275, doi:10.1175/2009JHM1176.1, 2010. 794
- 10 Sheffield, J., Wood, E. F., and Roderick, M. L.: Little change in global drought over the past 60 years, *Nature*, 491, 435–438, doi:10.1038/nature11575, 2012. 772, 774
- Syed, T. H., Famiglietti, J. S., and Chambers, D. P.: GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales, *J. Hydrometeorol.*, 10, 22–40, doi:10.1175/2008JHM993.1, 2009. 780, 784
- 15 Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., Buchmann, N., Ammann, C., Montagnani, L., Richardson, A. D., Wohlfahrt, G., and Seneviratne, S. I.: A regional perspective on trends in continental evaporation, *Geophys. Res. Lett.*, 36, L02404, doi:10.1029/2008GL036584, 2009. 774
- Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's global energy budget, *B. Am. Meteorol. Soc.*, 90, 311–323, 2009. 780
- 20 Wang, K. C. and Dickinson, R. E.: A review of global terrestrial evapotranspiration: observation, modeling, climatology, and climatic variability, *Rev. Geophys.*, 50, RG2005, doi:10.1029/2011RG000373, 2012. 780
- Wang, K. C. and Liang, S. L.: An improved method for estimating global evapotranspiration based on satellite determination of surface net radiation, vegetation index, temperature, and soil moisture, *J. Hydrometeorol.*, 9, 712–727, doi:10.1175/2007JHM911.1, 2008. 794
- 25 Wang, K. C., Dickinson, R. E., Wild, M., and Liang, S. L.: Evidence for decadal variation in global terrestrial evapotranspiration between 1982 and 2002: 1. Model development, *J. Geophys. Res.-Atmos.*, 115, D20112, doi:10.1029/2009JD013671, 2010a. 773
- 30 Wang, K. C., Dickinson, R. E., Wild, M., and Liang, S. L.: Evidence for decadal variation in global terrestrial evapotranspiration between 1982 and 2002: 2. Results, *J. Geophys. Res.-Atmos.*, 115, D20113, doi:10.1029/2010JD013847, 2010b. 773

- Weedon, G. P.: WATCH Technical Report Number 37, Creation of the WATCH 20th Century Ensemble Product, Tech. rep., UK Met Office, Exeter, UK, 2011. 777
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Osterle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *J. Hydrometeorol.*, 12, 823–848, doi:10.1175/2011JHM1369.1, 2011. 776, 794
- 5 Yao, Y. J., Liang, S. L., Qin, Q. M., Wang, K. C., Liu, S. M., and Zhao, S. H.: Satellite detection of increases in global land surface evapotranspiration during 1984–2007, *International Journal of Digital Earth*, 5, 299–318, doi:10.1080/17538947.2011.598953, 2012. 773
- 10 Zhang, Y. Q., Leuning, R., Hutley, L. B., Beringer, J., McHugh, I., and Walker, J. P.: Using long-term water balances to parameterize surface conductances and calculate evaporation at 0.05 degrees spatial resolution, *Water Resour. Res.*, 46, W05512, doi:10.1029/2009WR008716, 2010. 794

# HESSD

10, 769–805, 2013

## Evapotranspiration benchmark products

B. Mueller et al.

<a href="#">Title Page</a>	
<a href="#">Abstract</a>	<a href="#">Introduction</a>
<a href="#">Conclusions</a>	<a href="#">References</a>
<a href="#">Tables</a>	<a href="#">Figures</a>
<a href="#">I◀</a>	<a href="#">▶I</a>
<a href="#">◀</a>	<a href="#">▶</a>
<a href="#">Back</a>	<a href="#">Close</a>
<a href="#">Full Screen / Esc</a>	
<a href="#">Printer-friendly Version</a>	
<a href="#">Interactive Discussion</a>	





Evapotranspiration benchmark products

B. Mueller et al.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

**Table 2.** Overview of ET datasets, including their ET scheme or land-surface schemes (LSS), along with the number of soil layers, precipitation forcing dataset and atmospheric forcing variables. Model names with a star are classified as global hydrological models (GHMs, see text). Forcing variables are *P*: precipitation; *T*: air temperature; *W*: wind speed; *Q*: specific humidity; *R*: radiation; SP: surface pressure. “na” denotes either not applicable or information currently not available. Note that GS-VISA and GS-CLMTOP cannot strictly be classified as aerodynamic approaches, since they include a carbon cycle and photosynthetic control on transpiration. Models with an x are included in the 1989–2005 merged synthesis product.

Group	Name	ET scheme/ LSS for reanalyses	89-05	No. soil layers	Precipitation forcing datasets	Atmosph. forcing variables	Reference
Diagnostic	PT-JPL*	Priestley-Taylor	0	na	Not required	<i>T, Q, R, red/NIR</i> reflectances	Fisher et al. (2008)
	MAUNI	Empirical	na	na	Not required	na	Wang and Liang (2008)
	PRUNI	Penman-Monteith	x	na	Sheffield et al. (2006)	na	Sheffield et al. (2010)
	MPIBGC	Empirical	x	na	GPCC	na	Jung et al. (2009)
	CSIRO	Penman-Monteith	x	na	GPCC	na	Zhang et al. (2010)
	GLEAM v1A	Priestley-Taylor	x	3	CPC unified precip	na	Miralles et al. (2011b,a)
	AWB	None	x	na	GPCP	na	Mueller et al. (2011a)
LSMs and GHMs	GSWP	GS-COLA	Aerodynamic	6	NCEP, GPCC, GPCP (CRU for spin-up)	<i>P, T, W, Q, R, SP</i>	Dirmeyer et al. (2006)
		GS-NOAH	Penman-Monteith	4		<i>P, T, W, Q, R, SP</i>	
		GS-NSIPP	Aerodynamic	3		<i>P, T, W, Q, R, SP</i>	
		GS-VISA	Aerodynamic	10		<i>P, T, W, Q, R, SP</i>	
		GS-ISBA	Aerodynamic	3		<i>P, T, W, Q, R, SP</i>	
		GS-BUCKET	Aerodynamic	1		<i>P, T, W, Q, R, SP</i>	
		GS-CLMTOP	Aerodynamic	10		<i>P, T, W, Q, R, SP</i>	
		GS-HY-SSIB	Aerodynamic	3		<i>P, T, W, Q, R, SP</i>	
		GS-LAD	Aerodynamic	1		<i>P, T, W, Q, R, SP</i>	
		GS-MOSAIC	Penman-Monteith	3		<i>P, T, W, Q, R, SP</i>	
		GS-MOSES2	Penman-Monteith	4		<i>P, T, W, Q, R, SP</i>	
		GS-SIBUC	Aerodynamic	3		<i>P, T, W, Q, R, SP</i>	
		GS-SWAP	Penman-Monteith	2		<i>P, T, W, Q, R, SP</i>	
		WaterMIP	WM-GWAVA*	Penman-Monteith		multi	
	WM-H08		Aerodynamic	1	<i>P, T, W, Q, R, SP</i>		
	WM-HTESSEL		Penman-Monteith	4	<i>P, T, W, Q, R, SP</i>		
	WM-JULES		Penman-Monteith	4	<i>P, T, W, Q, R, SP</i>		
WM-LPJmL*	Priestley-Taylor		2	<i>P, T, R</i>			
WM-MacPDM*	Penman-Monteith	multi	<i>P, T, W, Q, R</i>				
WM-MATSIRO	Aerodynamic	5	<i>P, T, W, Q, R, SP</i>				
WM-MPI*	Thornthwaite	1	<i>P, T</i>				
WM-Orchidee	Aerodynamic	11	<i>P, T, W, Q, R, SP</i>				
WM-VIC	Penman-Monteith	2	<i>P, T, W, Q, R, SP</i>				
WM-WaterGAP*	Priestley-Taylor	multi	<i>P, T, R</i>				

\* Referred to as UCB in Mueller et al. (2011a).

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Evapotranspiration  
benchmark products

B. Mueller et al.

Table 2. Continued.

Group	Name	ET scheme/ LSS for reanalyses	89-05	No. soil layers	Precipitation forcing datasets	Atmosph. forcing variables	Reference	
ORCH	EI-ORCH	Aerodynamic	x	2	ERA-Interim CRU, NCEP	<i>P, T, W, Q, R, SP</i>	Krinner et al. (2005)	
		Aerodynamic	x	11				
	VIC	VIC	Penman-Monteith	x	2	obs. and NCEP	<i>P, T, W, Q, R, SP</i>	Sheffield and Wood (2007)
	NOAH-PF	GL-NOAHPF	Penman-Monteith	x	4	Sheffield et al. (2006)	<i>P, T, W, Q, R, SP</i>	
	MERRA-LAND	M-LAND	Penman-Monteith	x	na	Replay of MERRA-reanalysis	<i>P, T, W, Q, R, SP</i>	Reichle et al. (2011)
Reanalyses	ERAINT (ERA-Interim)	TESSEL	x	na			Dee et al. (2011)	
	CFSR (CFSR-NCEP)	NOAH	x	na			Saha et al. (2010)	
	JRA (JRA-25)	SIB	x	na			Onogi et al. (2007)	
	MERRA	GEOS-5 Catchment LSM	x	na			Bosilovich (2008)	

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I ◀

▶ I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## Evapotranspiration benchmark products

B. Mueller et al.

**Table 4.** Mean ET of merged synthesis products 1989–2005 (*long*), 1989–1995 (*short*), mean precipitation 1989–2005 (average of CRU, GPCC, GPCP and CPC) and ET and precipitation trends 1998–2005 in climate regions. Slope of trends (Theil-Sen) and significance (italic font, Mann-Kendal) estimated as for Table 3.

Climate region	Abbr- viation	Area [10 <sup>3</sup> km <sup>2</sup> ]	Mean ET synthesis long [mm d <sup>-1</sup> ]	Mean ET synthesis short [mm d <sup>-1</sup> ]	Mean precip CRU,GPCC, GPCP,CPC [mm d <sup>-1</sup> ]	ET trend 1998–2005 [km <sup>3</sup> yr <sup>-2</sup> ]	Mean preci- pitation trend 1998–2005 [km <sup>3</sup> yr <sup>-2</sup> ]
Equatorial fully humid	Af	5914	3.34	3.23	6.78	<i>-2.6</i>	<i>-0.1</i>
Equatorial monsoonal	Am	4822	3.16	3.02	5.55	<i>-1.6</i>	<i>-2.4</i>
Equatorial winter dry	Aw	16 687	2.52	2.49	3.47	<i>-9.0</i>	<i>-4.6</i>
Equatorial summer dry	As	745	2.05	2.09	2.77	0.0	1.0
Arid desert	BW north	19 247	0.27	0.27	0.29	1.9	5.9
Arid desert	BW south	4766	0.60	0.61	0.69	<i>-4.1</i>	<i>-5.8</i>
Arid steppe	BS north	9993	0.91	0.99	1.03	3.4	2.2
Arid steppe	BS south	6455	1.24	1.25	1.44	<i>-6.5</i>	<i>-6.9</i>
Warm temp. summer dry	Cs	3901	1.22	1.24	1.62	0.4	1.3
Warm temp. winter dry	Cw	5802	2.03	1.97	2.95	<i>-0.7</i>	<i>-2.8</i>
Warm temp. fully humid	Cf	11 533	1.87	2.03	3.11	<i>-0.8</i>	<i>-8.0</i>
Snow summer dry	Ds	1060	0.87	0.84	1.15	0.2	0.7
Snow winter dry	Dw	4777	1.05	1.08	1.33	0.5	3.7
Snow fully humid	Df	26 207	0.97	0.95	1.49	0.3	3.7
Polar	E	9012	0.46	0.42	1.05	<i>-0.4</i>	<i>4.9</i>
Global land		130 922	1.35	1.35	2.07	<i>-18.9</i>	<i>-7.2</i>

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

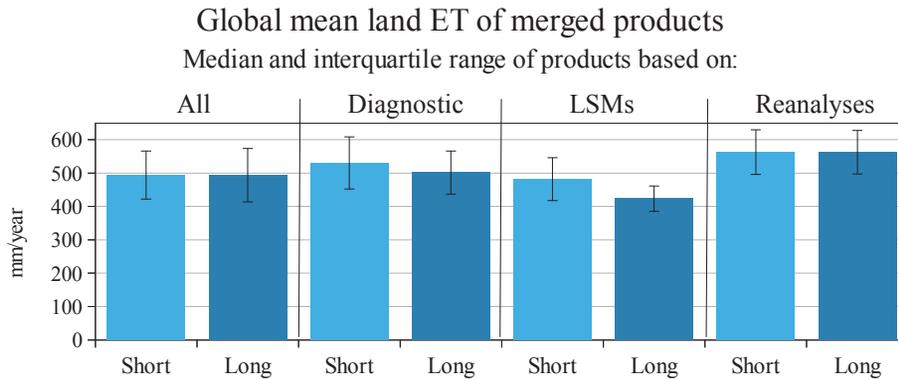
Printer-friendly Version

Interactive Discussion



Evapotranspiration benchmark products

B. Mueller et al.



**Fig. 1.** Global mean ET of merged synthesis products based on all datasets, only the diagnostic, only LSMs and only reanalyses. The medians and interquartile ranges for the short (1989–1995) and long (1989–2005) merged products are shown.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

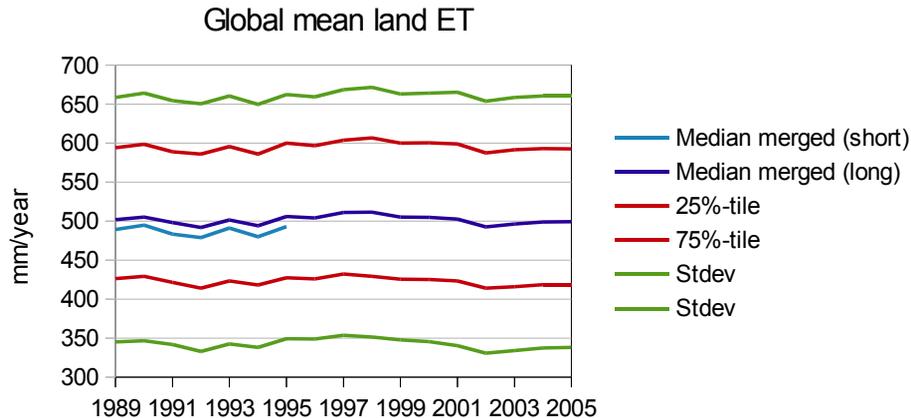
Printer-friendly Version

Interactive Discussion



## Evapotranspiration benchmark products

B. Mueller et al.



**Fig. 2.** Variation of global mean ET of merged synthesis products (based on all datasets). The median for both short (1989–1995) and long (1989–2005) products as well as the interquartile range and the standard deviation (median  $\pm 1$  standard deviation) of the long product are shown.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

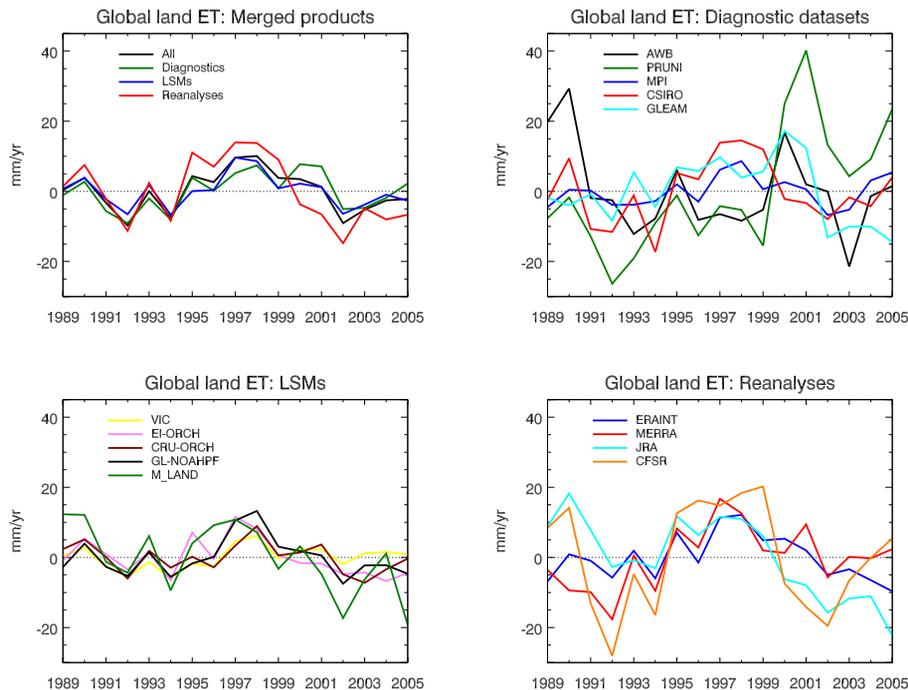
Printer-friendly Version

Interactive Discussion



## Evapotranspiration benchmark products

B. Mueller et al.



**Fig. 3.** Anomaly timeseries (1989–2005) of the four merged synthesis benchmark products (top left) and the individual diagnostic datasets (top right), LSMs (bottom left) and reanalyses (bottom right) that contribute to the long merged synthesis product.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Evapotranspiration benchmark products

B. Mueller et al.

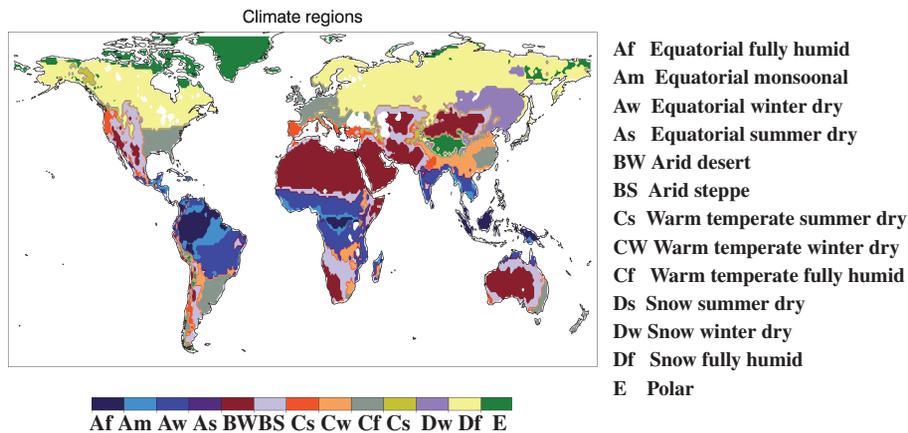


Fig. 4. Climate regions (Koeppen–Geiger classification).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

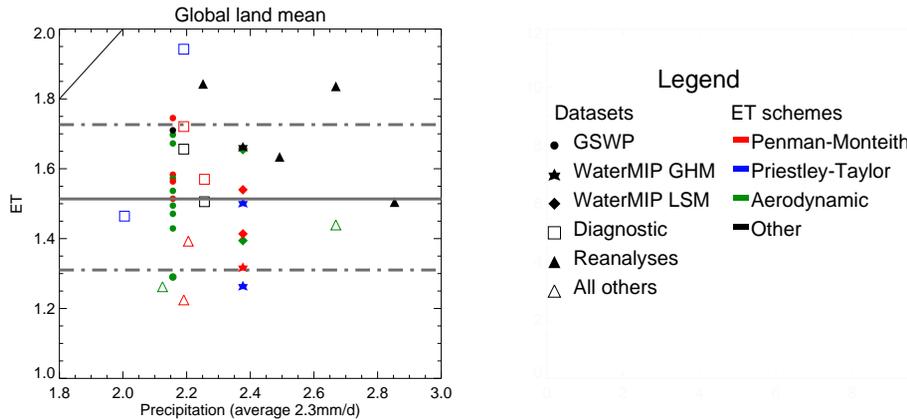
Printer-friendly Version

Interactive Discussion



## Evapotranspiration benchmark products

B. Mueller et al.



**Fig. 5.** Scatter plot of ET (in  $\text{mm day}^{-1}$ ) from each dataset that is included in the short merged product (1989–1995) versus precipitation from the corresponding forcing dataset. If no precipitation data is used for the derivation of ET, the average of CRU, GPCP, CPC and GPCC has been used instead (see Table 2). The merged synthesis product's median is indicated with a full line, the IQR with dash-dotted lines. The precipitation value indicated is the average of all datasets.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

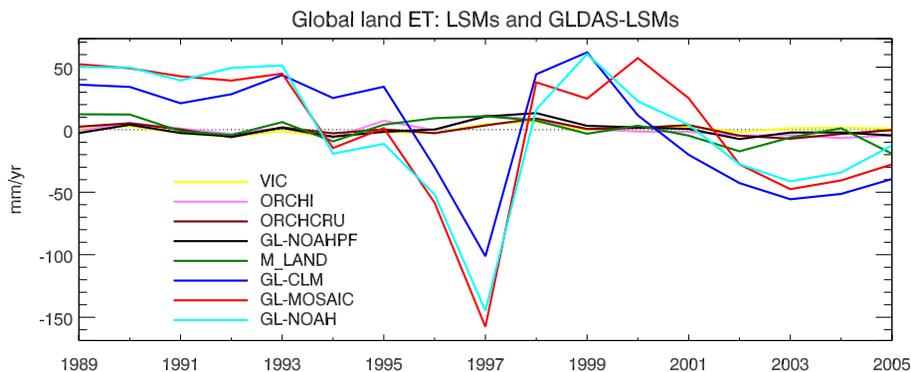
Printer-friendly Version

Interactive Discussion



## Evapotranspiration benchmark products

B. Mueller et al.



**Fig. A1.** Timeseries 1989–2005 of LSMs. In addition to the LSMs that contribute to the long merged synthesis product, GLDAS-I simulations from the models CLM, MOSAIC and NOAH are shown.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

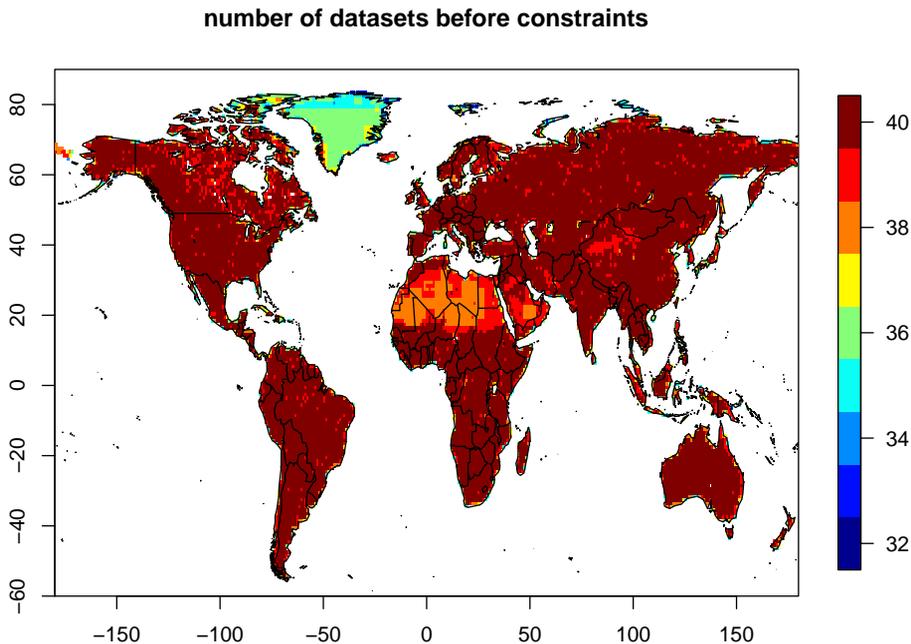
Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Fig. A2.** Number of datasets available before applying the constraints at each pixel. For the number of datasets included in the final merged synthesis product (after constraints) at each month, see movie in Supplement online information.

## Evapotranspiration benchmark products

B. Mueller et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

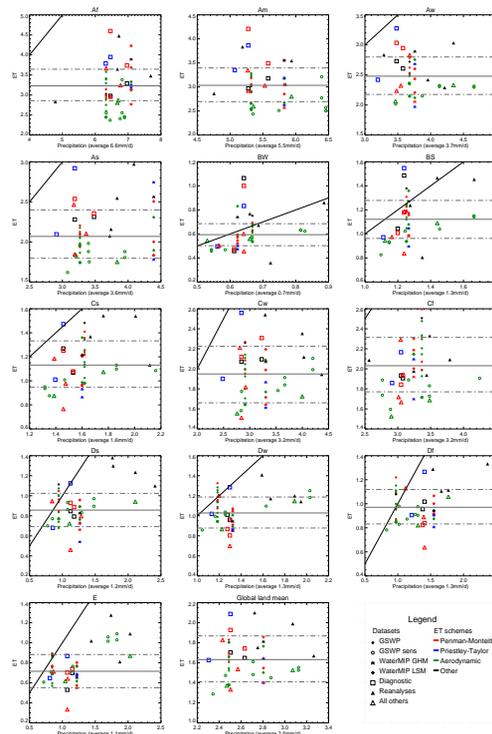
Printer-friendly Version

Interactive Discussion



## Evapotranspiration benchmark products

B. Mueller et al.



**Fig. A3.** Scatter plot for all different climate regions of ET (in  $\text{mm day}^{-1}$ ) from each dataset that is included in the short merged products (1989–1995) as well as the GSWP sensitivity runs versus precipitation from the corresponding forcing dataset. If no precipitation data is used for the derivation of ET, the average of CRU, GPCP, CPC and GPCC has been used instead (see Table 2). The merged synthesis product’s median is indicated with a full line, the IQR with dash-dotted lines. For abbreviations of climate regions, see Table 4.