Hydrology and Earth System Sciences

Open Access

Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis

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

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[2]LERMA, Observatoire de Paris, Paris, France
[3]LSCE, UMR CEA-CNRS, Gif-sur-Yvette, France
[4]George Mason University, Fairfax, Virginia, USA
[5]VU University Amsterdam, the Netherlands
[6]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
[7]Max Planck Institute for Biogeochemistry, Jena, Germany
[8]Wageningen University, Wageningen, the Netherlands
[9]School of Geographical Sciences, University of Bristol, UK
[10]Water Desalination and Reuse Center, King Abdullah University of Science and Technology, Kingdom of Saudi Arabia
[11]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA
[12]College of Global Change and Earth System Science, Beijing Normal University, Beijing, China
[13]CSIRO Land and Water, Canberra, Australia

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

Abstract. Land evapotranspiration (ET) estimates are available from several global data sets. Here, monthly global land ET synthesis products, merged from these individual data sets 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 data sets while those over the longer period are based on a total of 14 data sets. In the individual data sets, ET is derived from satellite and/or in situ observations (diagnostic data sets) or calculated via land-surface models (LSMs) driven with observations-based forcing or output from atmospheric reanalyses. Statistics for four merged synthesis products are provided, one including all data sets and three including only data sets 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 (0.13 mm yr^{-2} in our merged product) followed by a significant decrease in this trend (-0.18 mm yr^{-2}), 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 data sets) is 493 mm yr^{-1} (1.35 mm d^{-1}) 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 263 mm yr^{-1} (34 406 km^3 yr^{-1}) for a total land area of 130 922 000 km^2. Precipitation, being an important driving factor and input to most simulated ET data sets, presents uncertainties between single data sets 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 data sets based on in situ observations or satellite retrievals of different indirect variables have been derived. In Mueller et al. (2011b), an evaluation of their characteristics and agreement within the LandFlux-EVAL (see 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 data sets over a shorter (3 yr) period but also assessed the radiative and sensible fluxes. These studies considered dedicated data sets that derive ET from combinations of observations or observations-based estimates together with targeted algorithms (referred to as diagnostic data sets), 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). 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 data sets are currently available and have been analyzed in these studies, a global benchmark for ET is missing. Such a benchmark data set 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 ET. Besides hydrological applications, changes in ET are also relevant for temperature variability and hot extremes (e.g., Seneviratne et al., 2006, 2010; Hirschi et al., 2011; Mueller and Seneviratne, 2012). 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 data sets have been developed in the present study with the provision of different estimates of uncertainty.

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 data sets 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. As the variability of water vapour is negligible at the global scale, the rates of precipitation and evaporation are similar. Within a changing climate, changes in the hydrological cycle are expected, but very difficult 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). Increases of land evaporation over these periods are therefore expected. While some publications relate this behavior to a possible intensification of the hydrological cycle, this term is not well defined. Wild et al. (2008) explained the acceleration of the hydrological cycle since the mid-1980s with an increase in solar heating (global brightening) and in thermal heating (enhanced greenhouse gas forcing). While evaporation from ocean surfaces is likely 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 0.7 mm per yr^2, using 1120 globally distributed stations (Wang et al., 2010a). Jung et al. (2010) performed a trend analysis based on a global data set 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 data set), but also including eight other data sets. 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., 2012). 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 data sets used to derive such ET trends are large and may entail spurious features 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).

Besides precipitation and temperature, radiation is an important forcing for ET. Studies on trends in radiative forcing do not agree well. Wild and Liepert (2010) for example noted a brightening since the 1990s, while Romanou et al. (2007) found a slight dimming. An evaluation of radiation data sets

Table 1. Number and type of data sets included in the 8 different merged synthesis products.

Merged synthesis products based on	Number of data sets 1989–1995 (denoted short)	Number of data sets 1989–2005 (denoted long)
All data set categories	40	14
Diagnostic	7	5
LSMs	29	5
Reanalyses	4	4

can be found in e.g. Troy and Wood (2009) or Jimenez et al. (2011).

The benchmark synthesis products presented in this study are used to assess the interannual variations of ET on the global scale and encompasses the largest number of ET products to date. Besides the evaluation of the temporal variability of the benchmark products and the underlying single data sets, 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 Data sets and method

2.1 Merged benchmark synthesis products of evapotranspiration

We present here new multi-year merged synthesis products based on the analyses of existing land ET data sets that were available to us at the time of writing and fully cover one or both of the two studied time periods. A first product spans the time period 1989–1995 and includes 40 data sets, while the second is available for the longer time period 1989–2005 and includes 14 data sets (Table 1, all data set categories). Consistent with a previous analysis (Mueller et al., 2011b), the type of data sets included can be classified as diagnostic data sets, LSMs and reanalyses (see Sect. 1). Besides the two merged synthesis products based on all types of data sets, merged synthesis products from each of the individual data set 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 data sets. 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).

2.2 Overview of included data sets

An overview of the diagnostic data sets, LSMs and reanalyses considered for the preparation of the merged synthesis products is provided in Table 2. All data sets are available for the time period 1989–1995, while a subset is available over the period 1989–2005 (cross in 5th column in Table 2) and 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 data sets, such as the ET schemes, the number of soil layers used in the case of LSMs, the precipitation forcing data sets and other forcing variables used for the derivation of the respective data sets or, in the case of reanalyses, the land-surface schemes.

We considered here several additional data sets compared to the earlier analysis of Mueller et al. (2011b). The additional data sets included in this study are the diagnostic data set 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 Supplement Fig. A1 and Rui, 2011), which arose because the source of forcing data changed several times over the GLDAS-I time period (Matt Rodell, personal communication). However, we included GLDAS-II simulations (see Rui, 2011) from one of these models (NOAH Version 3.3) which was produced recently with a consistent forcing data set (Princeton forcing, see Sheffield et al., 2006).

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 data set (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

www.hydrol-earth-syst-sci.net/17/3707/2013/

Hydrol. Earth Syst. Sci., 17, 3707–3720, 2013

Table 2. Overview of ET data sets, including their ET scheme or land-surface schemes (LSS), along with the number of soil layers, precipitation forcing data set 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. Note that PT-JPL was referred to as UCB in Mueller et al. (2011b).

	Group	Name	ET scheme/ LSS for reanalyses	89-05	No. soil layers	Precipitation forcing data sets	Atmosph. forcing variables	Reference
Diagnostic		PT-JPL	Priestley–Taylor		0	Not required	T, Q, R, red/NIR reflectances	Fisher et al. (2008)
		MAUNI	Empirical		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)	P, T, W, Q, R, SP	Dirmeyer et al. (2006)
		GS-NOAH	Penman–Monteith		4		P, T, W, Q, R, SP	
		GS-NSIPP	Aerodynamic		3		P, T, W, Q, R, SP	
		GS-VISA	Aerodynamic		10		P, T, W, Q, R, SP	
		GS-ISBA	Aerodynamic		3		P, T, W, Q, R, SP	
		GS-BUCKET	Aerodynamic		1		P, T, W, Q, R, SP	
		GS-CLMTOP	Aerodynamic		10		P, T, W, Q, R, SP	
		GS-HY-SSIB	Aerodynamic		3		P, T, W, Q, R, SP	
		GS-LAD	Aerodynamic		1		P, T, W, Q, R, SP	
		GS-MOSAIC	Penman–Monteith		3		P, T, W, Q, R, SP	
		GS-MOSES2	Penman–Monteith		4		P, T, W, Q, R, SP	
		GS-SIBUC	Aerodynamic		3		P, T, W, Q, R, SP	
		GS-SWAP	Penman–Monteith		2		P, T, W, Q, R, SP	
	WaterMIP	WM-GWAVA*	Penman–Monteith		multi	WATCH Weedon et al. (2011)	P, T, W, Q, R, SP	Haddeland et al. (2011)
		WM-H08	Aerodynamic		1		P, T, W, Q, R, SP	
		WM-HTESSEL	Penman–Monteith		4		P, T, W, Q, R, SP	
		WM-JULES	Penman–Monteith		4		P, T, W, Q, R, SP	
		WM-LPJmL*	Priestley–Taylor		2		P,T,R	
		WM-MacPDM*	Penman–Monteith		multi		P,T,W,Q,R	
		WM-MATSIRO	Aerodynamic		5		P, T, W, Q, R, SP	
		WM-MPI*	Thornthwaite		1		P,T	
		WM-Orchidee	Aerodynamic		11		P, T, W, Q, R, SP	
		WM-VIC	Penman–Monteith		2		P, T, W, Q, R, SP	
		WM-WaterGAP*	Priestley–Taylor		multi		P,T,R	
	ORCH	EI-ORCH	Aerodynamic	x	2	ERA-Interim	P, T, W, Q, R, SP	Krinner et al. (2005)
		CRU-ORCH	Aerodynamic	x	11	CRU, NCEP	P, T, W, Q, R, SP	
	VIC	VIC	Penman–Monteith	x	2	obs. and NCEP	P, T, W, Q, R, SP	Sheffield and Wood (2007)
	NOAH-PF	GL-NOAHPF	Penman–Monteith	x	4	Sheffield et al. (2006)	P, T, W, Q, R, SP	
	MERRA-LAND	M-LAND	Penman–Monteith	x	na	Replay of MERRA-reanalysis	P, T, W, Q, R, SP	Reichle et al. (2011)
Reanalyses		ERAINT (ERA-Interim)	TESSEL	x	na			Dee and Uppala (2008)
		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)

information on all other data sets, the reader is referred to Mueller et al. (2011b) and Jimenez et al. (2011).

2.3 Processing of ET data sets and merged synthesis products

In order to prepare the merged synthesis products, we first interpolated all data sets 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 data sets was done, and if one grid point was covered by less than 70 % of the data sets, it was excluded from the final synthesis product (for the number of data sets originally available, see Supplement Fig. A2). Antarctica is missing in the synthesis product. Some of the data sets 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 and for monthly values. For each grid point of the merged synthesis products, we calculated long-term monthly maxima of net radiation (from the Surface Radiation Budget (SRB) version 3) based on all

available years (1984–2007). Monthly ET values exceeding the long-term monthly maximum net radiation of that month by more than 25 % were excluded, unless ET was smaller than $0.35\,\mathrm{mm\,d^{-1}}$ ($128\,\mathrm{mm\,yr^{-1}}$), since for such small values, ground heat flux cannot be neglected (Bennett et al., 2008). Note, however, that if interception plays an important role, such as during winter time, ET can be several times larger than radiation due to additional energy input through advection. 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 runon could provide additional water input for ET. In order to exclude single data set values that were very different from those of the other data sets, 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 data sets at each grid point 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 data sets 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 data sets, LSMs and reanalyses) and from all categories (see Table 1) coincide to a large extent in their global land mean ET (Fig. 1), with highest values in the merged product based on reanalyses only ($563\,\mathrm{mm\,yr^{-1}}$) and lowest in that based on LSMs only ($423\,\mathrm{mm\,yr^{-1}}$). The interquartile ranges (IQRs, 75th-percentile minus 25th-percentile) are largest in the merged products based on diagnostic data sets 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 data set categories is shown in Fig. 2 (median of yearly values). The long merged product shows slightly higher values. The largest difference in the list of data sets included 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 an indication that including a large number of dependent data sets (i.e. model simulations driven with the same forcing data, such as GSWP and WaterMIP runs) does not have a strong influence.
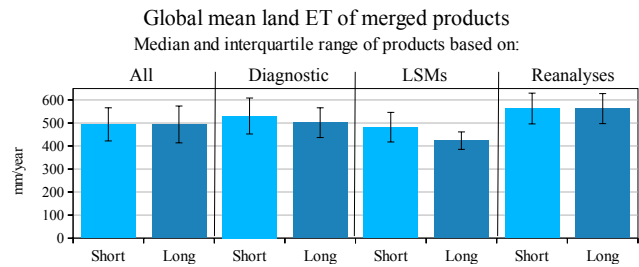

Fig. 1. Global mean ET of merged synthesis products based on all data sets, only the diagnostic, only LSMs and only reanalyses. The medians and interquartile ranges of the multi-year values for the short (1989–1995) and long (1989–2005) merged products are shown.

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 interannual variation as that found in the MPIBGC data set 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.

The reason for the large IQRs and standard deviations are the large differences in the absolute ET values of the single data sets. The IQRs and standard deviations based on the yearly anomalies of the underlying data sets (i.e. setting the mean of all data sets 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).

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

3.2 Single data sets

The temporal variations of the 14 single data sets contributing to the long merged synthesis products are shown in Fig. 3 (top right and bottom panels). In these analyses of single data sets, we excluded unrealistically high ET values, setting a threshold of $4560\,\mathrm{mm\,yr^{-1}}$ ($12.5\,\mathrm{mm\,d^{-1}}$). The LSMs (bottom left) and reanalyses (bottom right) are more consistent amongst one another in their yearly variations than the diagnostic data sets (top right). The ET time series of all LSMs and reanalyses peak between 1997 and 1999. Some of the diagnostic data sets 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 data sets display a significant negative trend (italic font) for
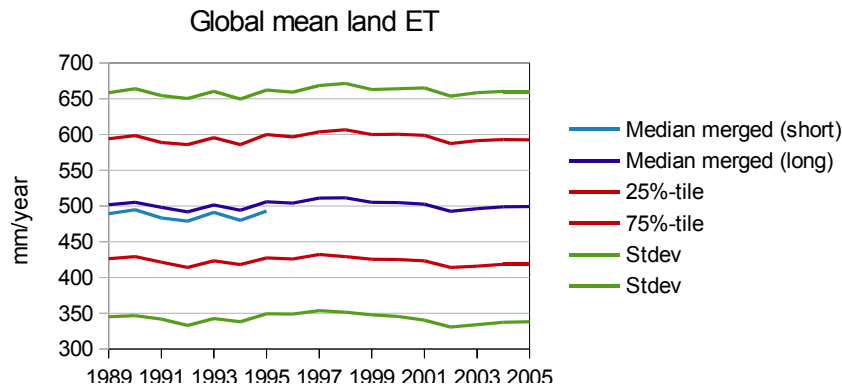
Fig. 2. Variation of global mean ET of merged synthesis products (based on all data sets). The median for both short (1989–1995) and *long* (1989–2005) products as well as the interquartile range and the standard deviation (median ± 1 standard deviation) of the long product's annual values are shown.
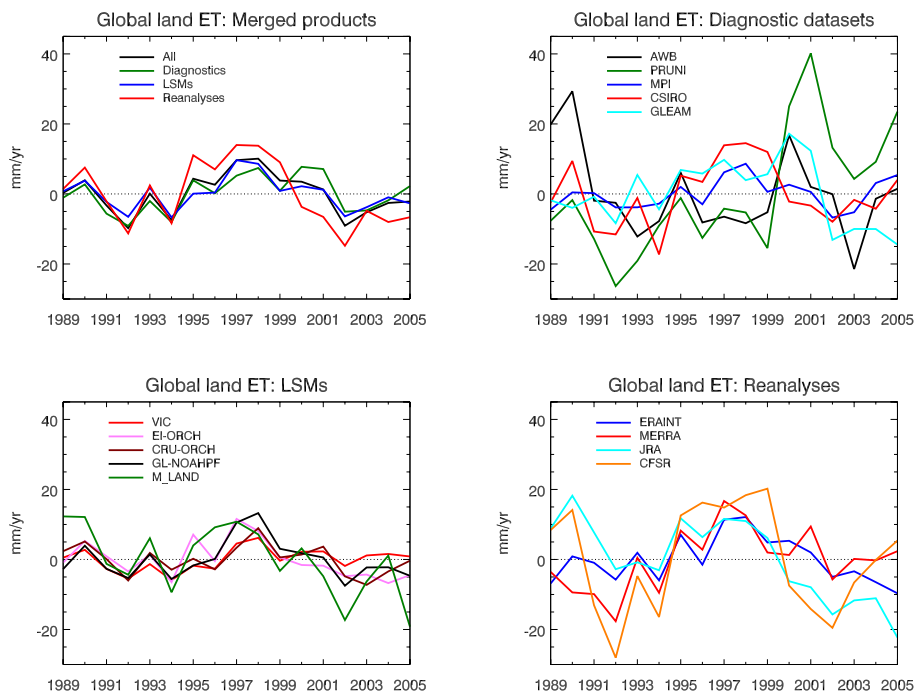

Fig. 3. Anomaly time series (1989–2005) of the four merged synthesis benchmark products (top left) and the individual diagnostic data sets (top right), LSMs (bottom left) and reanalyses (bottom right) that contribute to the long merged synthesis product.

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 data set categories) as well as precipitation data (average of CRU, GPCC, GPCP and CPC, for ref-

erences and information on these data sets, 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.

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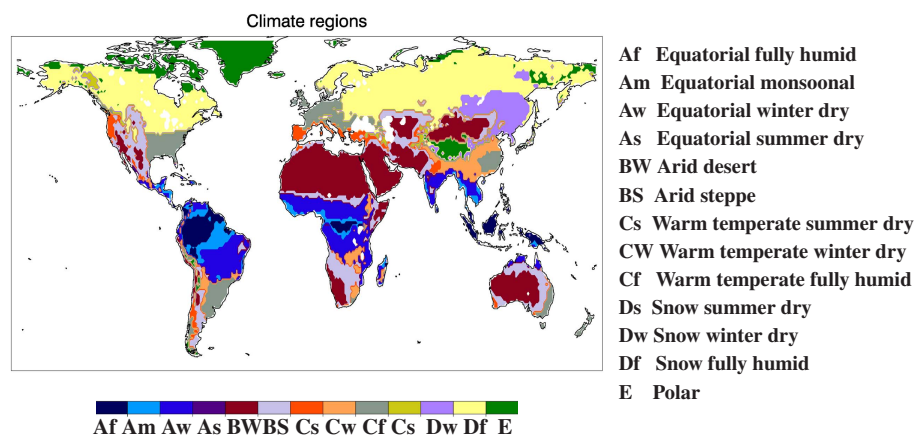
Fig. 4. Climate regions (Koeppen–Geiger classification).

Table 3. Slope of trends for the two time periods 1989–1997 and 1998–2005 of the merged (all) product and the single data sets. The slopes are estimated with the Theil–Sen estimator, which is robust against outliers. Significant values (non parametric Mann–Kendal two-sided test at 90 % level) are printed in italic font.

Data set	Trend 1989–1997 $\mathrm{mm\,yr^{-2}}$	Trend 1998–2005 $\mathrm{mm\,yr^{-2}}$
Merged (all)	0.13	*−0.18*
AWB	−0.24	0.12
PRUNI	0.08	0.55
MPIBGC	0.04	0.01
CSIRO	0.20	−0.18
GLEAM	*0.19*	−0.37
VIC	−0.01	−0.03
EI-ORCH	0.09	*−0.16*
CRU-ORCH	−0.02	−0.16
GL-NOAHPF	0.08	*−0.16*
M-LAND	−0.02	−0.29
ERAINT	0.19	*−0.37*
MERRA	0.38	−0.06
JRA	−0.01	*−0.51*
CFSR	0.27	−0.10

areas of all climate regions (third column) represents nearly 90 % of the global land area.

The global mean ET from the merged synthesis product amounts to 493 mm yr^{-1} for both the 1989–2005 and 1989–1995 products. This value is generally well within the range, and sometimes at the lower boundary, of other studies. For example Trenberth et al. (2009) reported a range of 504 to 664 mm yr^{-1}, Haddeland et al. (2011) 416 to 588, Wang and Dickinson (2012) 438 to 548 and Dirmeyer et al. (2006) a mean of 497 mm yr^{-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 val-

ues. Note, however, that values from different studies cannot be compared directly due to differences in land areas. The mean value of precipitation (average of CRU, GPCC, GPCP and CPC) amounts to 756 mm yr^{-1}. The difference between global precipitation and land ET corresponds to the water that leaves the continents as runoff and amounts to 263 mm yr^{-1} (34 406 km^3 yr^{-1}). 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 km^3 yr^{-2} (−0.14 mm 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 hemispheric parts of the BS and BW regions can also be found in the precipitation data sets. Furthermore, the table shows that global ET has decreased much more strongly than global precipitation over the period 1998–2005.

3.4 Precipitation forcing

The 1989–1995 global mean land ET of each data set 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 data set as listed in Table 2. If precipitation was not available (for some

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	Abbre-viation	Area [10^3 km^2]	Mean ET synthesis long [mm yr^{-1}]	Mean ET synthesis short [mm yr^{-1}]	Mean precip CRU,GPCC, GPCP,CRU [mm yr^{-1}]	ET trend 1989–2005 [mm yr^{-2}]	ET trend 1989–2005 [km^3 yr^{-2}]	Mean preci-pitation trend 1989–2005 [mm yr^{-2}]	Mean preci-pitation trend 1989–2005 [km^3 yr^{-2}]
Equatorial fully humid	Af	5914	1218	1179	2473	*−0.4*	*−2.6*	0.0	−0.1
Equatorial monsoonal	Am	4822	1155	1101	2027	*−0.3*	*−1.6*	−0.5	−2.4
Equatorial winter dry	Aw	16 687	918	907	1266	*−0.5*	*−9.0*	−0.3	−4.6
Equatorial summer dry	As	745	749	764	1012	0.0	0.0	1.3	1.0
Arid desert	BW north	19 247	98	99	104	0.1	1.9	0.3	5.9
Arid desert	BWsouth	4766	219	222	253	−0.9	−4.1	−1.2	−5.8
Aridsteppe	BS north	9993	332	360	376	0.3	3.4	0.2	2.2
Arid steppe	BS south	6455	454	456	527	−1.0	−6.5	−1.1	−6.9
Warm temp. summer dry	Cs	3901	446	453	591	0.1	0.4	0.3	1.3
Warm temp. winter dry	Cw	5802	741	720	1077	−0.1	−0.7	−0.5	−2.8
Warm temp. fully humid	Cf	11 533	682	740	1137	−0.1	−0.8	−0.7	−8.0
Snow summer dry	Ds	1060	318	305	419	0.2	0.2	0.7	0.7
Snow winter dry	Dw	4777	382	393	487	0.1	0.5	0.8	3.7
Snow fullyhumid	Df	26 207	353	348	544	0.0	0.3	0.1	3.7
Polar	E	9012	168	153	383	0.0	−0.4	*0.5*	*4.9*
Sum/average of above	Global land	130'922	493	493	756	−0.14	−18.9	−0.05	−7.2

diagnostic data sets), the average of four currently available observational data sets (CRU, GPCP, GPCC and CPC) was taken. Global mean values of these four precipitation data sets range from 730 to 803 mm yr^{-1}. The data set median of the merged synthesis ET product is indicated with a solid line, and the IQRs with dash-dotted lines. The single data sets are indicated with different symbols (groups) and colors (ET schemes).

We first compare simulations from the GSWP and the WaterMIP projects, which are each based on common forcing data sets (filled circles and stars/rhombi, respectively). The spread within the GSWP and WaterMIP simulations is similar, both globally and in most climate regions (see Supplement 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 clearly distinct 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.

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 Supplement 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 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 data sets exhibits an ET value of 550 mm yr^{-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 data sets, we only included those pixels of the merged product that were also available in all other data sets. Table 4, on the other hand, includes all land pixels (see Supplement Movie for data coverage).

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 data sets reveal an ET exceedance over precipitation (see bisecting line through origin in Supplement Fig. A3). The reasons could be that (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 data sets. This is not only the case for the global mean values, but also for single climate regions (Supplement Fig. A3). Large uncertainties in precipitation data sets have also recently been highlighted in Lorenz and Kunstmann (2012). The reason for smaller uncertainties in ET than in precipitation could
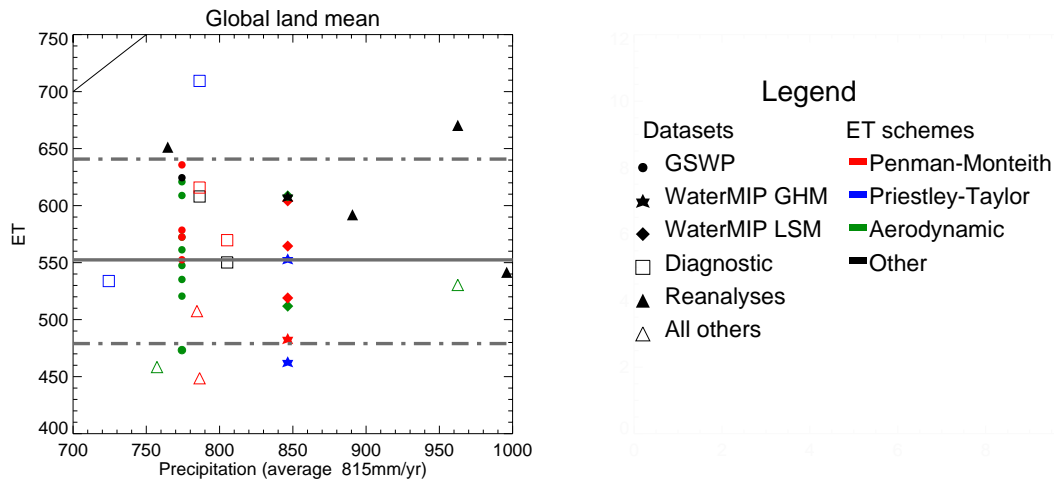
Fig. 5. Scatter plot of ET (in mm yr^{-1}) from each data set that is included in the short merged product (1989–1995) versus precipitation from the corresponding forcing data set. If no precipitation data is used for the derivation of ET, the average of CRU, GPCP, CPC and GPCC has been used instead (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 data sets.

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 data sets is nearly as large as the one in ET.

4 Conclusions

The intensity of the hydrological cycle determines the water availability and influences the climate system in various ways. Despite these important implications of possible changes in the hydrological cycle and with that, in ET, a global benchmark ET data set has long been missing. In the framework of the LandFlux-EVAL initiative (www.iac.ethz.ch/url/LandFlux-EVAL), several ET data sets based on observations (diagnostic data sets, 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 data sets. Global merged benchmark synthesis products of ET are derived and trends are analyzed in single LandFlux-EVAL data sets as well as in the merged ET products.

The benchmark synthesis products provide monthly, yearly and multi-year ensemble statistics for the time periods 1989–1995 (short) and 1989–2005 (long), respectively. For the creation of the short benchmark products, 7 diagnostic data sets, 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 data sets, we created short and long merged synthesis products based on all data sets 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. The synthesis products include different statistics of the multi-data set ensemble, such as median, mean, 25th-percentile, 75th-percentile, interquartile range, standard deviation and minimum and maximum values.

Evapotranspiration from the merged benchmark synthesis products shows realistic interannual variations that correspond to those found in a previous study based on a smaller number of ET data sets (Jung et al., 2010). The negative trend in global land ET[1] between 1998–2005 amounts to $-18.9 \, \text{km}^3 \, \text{yr}^{-2}$ ($-0.14 \, \text{mm yr}^{-2}$). 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 globe. Dividing these arid desert and steppe regions into 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 data sets on atmospheric input data sets, the influence of spurious trends in these data sets cannot be excluded.

Large uncertainties in absolute values of ET are found, which can partly be related to uncertainties in precipitation.

---

[1]After a space and time matching of all data sets, data coverage of roughly 90 % of the land surface was attained. Note that different land masks are employed in different studies and make a comparison difficult.
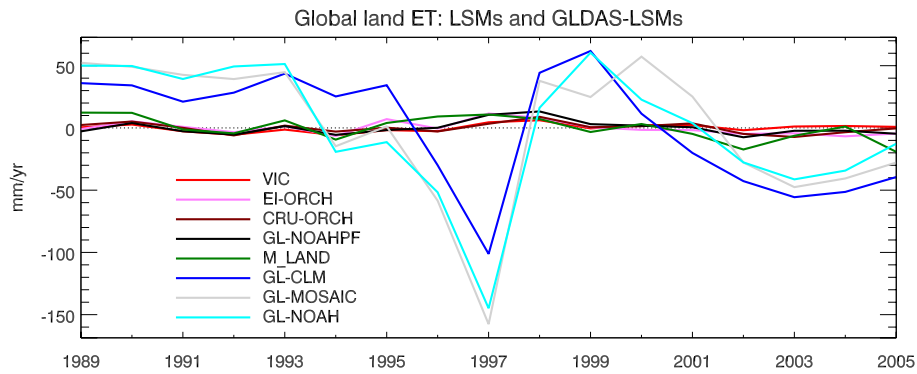
Fig. A1. Time series 1989–2005 of LSMs. In addition to the LSMs that contribute to the long-merged synthesis product, GLDAS-I simulations from the models CLM, MOSAIC and NOAH are shown.
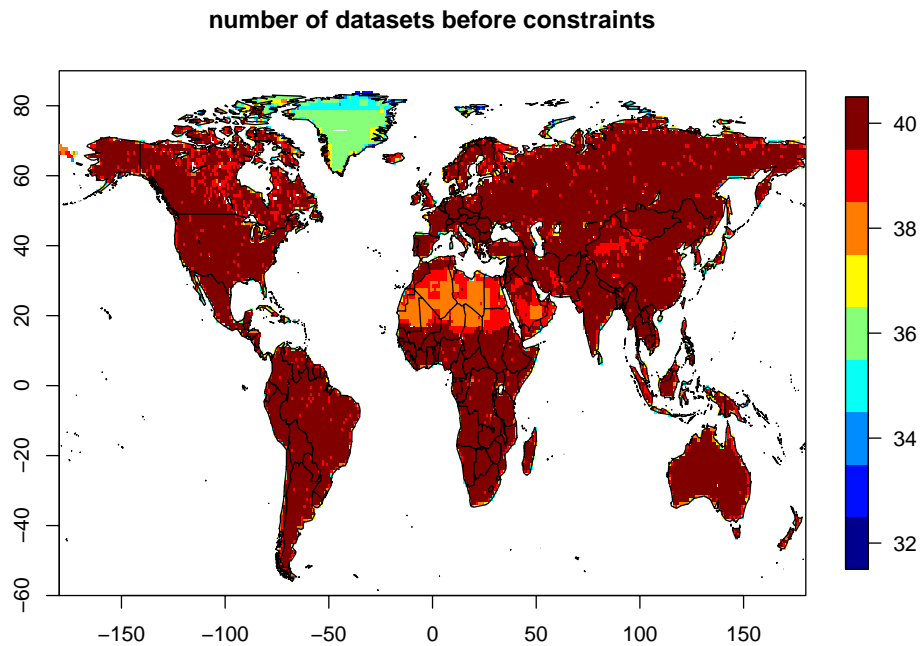

Fig. A2. Number of data sets available before applying the constraints at each pixel. For the number of data sets included in the final merged synthesis product (after constraints) at each month, see movie in Supplement.

Precipitation is 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 data sets and LSMs. Indeed, the spread in ET data sets is smaller than the spread in the corresponding precipitation data sets in our global analyses as well as in most climatic regions, which indicates that ET, as expected, 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 493 mm yr^{-1} (1.35 mm d^{-1}),

while precipitation to 756 mm yr^{-1} (2.07 mm d^{-1}) (average of four observations-based data sets). The difference of 263 mm yr^{-1} (34 406 km^3 yr^{-1}, runoff) is in agreement with estimates from previous studies (an overview can be found in Syed et al., 2009). In dry regions, ET exceeds precipitation in several data sets. The merged synthesis product's (median) ET is always lower than average precipitation.

Another important factor for the estimation of ET is the land cover type, especially in vegetated areas where transpiration is high. Even though the differences in land cover in the individual data sets may account for a large part of the uncertainties in the final merged synthesis products, including this information lies beyond the scope of our study.
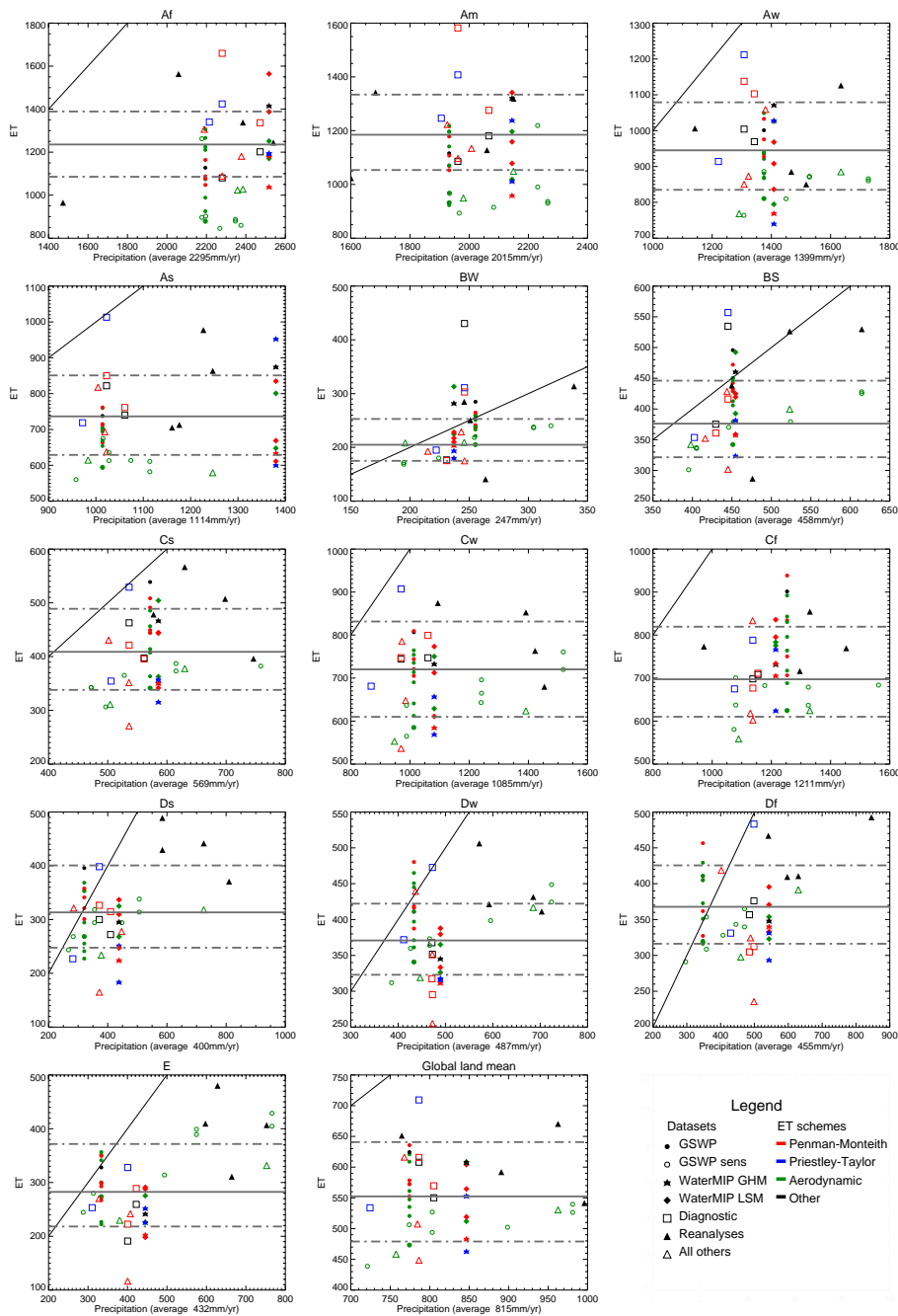
Fig. A3. Scatter plot for all different climate regions of ET (in mm yr^{-1}) from each data set that is included in the short merged products (1989–1995) as well as the GSWP sensitivity runs versus precipitation from the corresponding forcing data set. If no precipitation data is used for the derivation 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.

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 data sets are not totally independent from each other (e.g. same forcing 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 data sets

The observation-based precipitation data sets 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 data sets 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 data sets employed for the diagnostic ET data sets 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 data set 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 data set 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 GPCC monitoring product for the period 1986 to present is based on quality-controlled data from 7000 stations, which are interpolated into monthly area averages. 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).

Supplementary material related to this article is available online at http://www.hydrol-earth-syst-sci.net/ 17/3707/2013/hess-17-3707-2013-supplement.zip.

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.

Bengtsson, L., Hagemann, S., and Hodges, K. I.: Can climate trends be calculated from reanalysis data?, J. Geophys. Res.-Atmos., 109, D11111, doi:10.1029/2004JD004536, 2004.

Bennett, W. B., Wang, J., and Bras, R. L.: Estimation of global ground heat flux, J. Hydrometeorol., 9, 744–759, doi:10.1175/2008JHM940.1, 2008.

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.

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, 26 September, 2008.

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, D04110, doi:10.1029/2007JD009132, 2008.

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.

Dee, D. and Uppala, S.: Variational bias correction in ERA-Interim, ECMWF Technical Memoranda 575, European Centre for Medium-Range Weather Forecasts, available at: http://dss.ucar.edu/datasets/ds627.0/docs/IFS_documentation/ifs.2b.variational_bias_correction.pdf, 2008.

Dirmeyer, P. A., Gao, X. A., Zhao, M., Guo, Z. C., Oki, T. K., and Hanasaki, N.: GSWP-2 – Multimodel anlysis and implications for our perception of the land surface, B. Am. Meteorol. Soc., 87, 1381–1397, 2006.

Douville, H., Ribes, A., Decharme, B., Alkama, R., and Sheffield, J.: Anthropogenic influence on multidecadal changes in reconstructed global evapotranspiration, Nat. Clim. Change, 3, 59–62, doi:10.1038/nclimate1632, 2013.

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.

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.

Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B., and Stepanek, P.: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, Nat. Geosci., 4, 17–21, doi:10.1038/NGEO1032, 2011.

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.

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.

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.

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., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspira-

tion trend due to limited moisture supply, Nature, 467, 951–954, doi:10.1038/nature09396, 2010.

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.

Lorenz, C. and Kunstmann, H.: The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis, J. Hydrometeor., 13, 1397–1420, doi:10.1175/JHM-D-11-088.1, 2012.

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.

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.

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.

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.

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.

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, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2, 2000.

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

Onogi, K., Tslttsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kaalhori, 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.

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.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, P., Lohmann, D., and Toll, D.: The global land data assimilation system, B. Am. Meteorol. Soc., 85, 381–394, 2004.

Romanou, A., Liepert, B., Schmidt, G. A., Rossow, W. B., Ruedy, R. A., and Zhang, Y.: 20th century changes in surface solar

irradiance in simulations and observations, Geophys. Res. Lett., 34, L05713, doi:10.1029/2006GL028356, 2007.

Rui, H.: README Document for Global Land Data Assimilation System Version 2 (GLDAS-2) Products, GES DISC, 2011.

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.

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.

Seneviratne, S. I.: Climate science: Historical drought trends revisited, Nature, 491, 338–339, doi:10.1038/491338a, 2012.

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.

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C.: Land-atmosphere coupling and climate change in Europe, Nature, 443, 205–209, doi:10.1038/nature05095, 2006.

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.

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, doi:10.1029/2006JD008288, 2007.

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.

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.

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.

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.

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.

Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth'S Global Energy Budget, B. Am. Meteorol. Soc., 90, 311–323, 2009.

Troy, T. J. and Wood, E.: Comparison and evaluation of gridded radiation products across northern Eurasia, Environ. Res. Lett., 4, 045008, doi:10.1088/1748-9326/4/4/045008, 2009.

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.

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.

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.

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.

Weedon, G. P.: WATCH Technical Report Number 37, Creation of the WATCH 20th Century Ensemble Product, Tech. rep., UK Met Office, 2011.

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.

Wild, M. and Liepert, B.: The Earth radiation balance as driver of the global hydrological cycle, Environ. Res. Lett., 5, 025003, doi:10.1088/1748-9326/5/2/025203 , 2010.

Wild, M., Grieser, J., and Schaer, C.: Combined surface solar brightening and increasing greenhouse effect support recent intensification of the global land-based hydrological cycle, Geophys. Res. Lett., 35, L17706, doi:10.1029/2008GL034842, 2008.

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, Int. J. Dig. Earth, 5, 299–318, doi:10.1080/17538947.2011.598953, 2012.

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.