Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations

B. Mueller[1], S.I. Seneviratne[1], C. Jimenez[2], T. Corti[1,3], M. Hirschi[1,4], G. Balsamo[5], P. Ciais[6], P. Dirmeyer[7], J.B. Fisher[8], Z. Guo[7], M. Jung[9], F. Maignan[6], M.F. McCabe[10], R. Reichle[11], M. Reichstein[9], M. Rodell[11], J. Sheffield[12], A. J. Teuling[1,13], K. Wang[14], E.F. Wood[12], and Y. Zhang[15]

T. Corti, M. Hirschi, B. Mueller, S.I. Seneviratne, A.J. Teuling, Institute for Atmospheric and Climate Science, ETH Zurich, 8092 Zurich, Switzerland. (brigitte.mueller@env.ethz.ch, sonia.seneviratne@env.ethz.ch)

G. Balsamo, ECMWF, Shinfield Park, Reading, UK.

P. Ciais, F. Maignan, LSCE, UMR CEA-CNRS, CE, L'Orme des Merisiers, Gif-sur-Yvette, France.

P. Dirmeyer, Center for Ocean-Land-Atmosphere Studies, Calverton, MD, USA.

J.B. Fisher, Water and Carbon Cycles Group, NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.

C. Jimenez, LERMA, Observatoire de Paris, Paris, France.

M. Jung, M. Reichstein, Max Planck Institute for Biogeochemistry, Jena, Germany.

M.F. McCabe, School of Civil and Environmental Engineering, University of NSW, Sydney, Australia.

R. Reichle, M. Rodell, NASA-GSFC, Greenbelt, MD, USA.

J. Sheffield, E. F. Wood, Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.

K. Wang, Department of Geography, University of Maryland, College Park, MD, Maryland, USA.

Y. Zhang, CSIRO Land and Water, Canberra, Australia.

[1]Institute for Atmospheric and Climate

Quantification of global land evapotranspiration (ET) has long been associated with large uncertainties due to the lack of reference observations. Several recently developed products now provide the capacity to estimate ET at global scales. These products, partly based on observational data, include satellite-based products, land surface model (LSM) simulations, atmospheric reanalysis output, estimates based on empirical upscaling of eddy-covariance flux measurements, and atmospheric water balance datasets.

The LandFlux-EVAL project aims to evaluate and compare these newly developed datasets. Additionally, an evaluation of IPCC AR4 global climate model (GCM) simulations is presented, providing an assessment of their capacity to reproduce flux behavior relative to the observations-based products. Though differently constrained with observations, the analyzed reference datasets display similar large-scale ET patterns. ET from the IPCC AR4 simulations was significantly smaller than that from the other products for India (up to 1 mm/d) and parts of eastern South America, and larger in the western USA, Australia and China. The inter-product variance is lower across the IPCC AR4 simulations than across the reference datasets in several regions, which indicates that uncertainties may be underestimated in the IPCC AR4 models due to shared biases of these simulations.

Science, ETH Zurich, Zurich, Switzerland.

1. Introduction

Land evapotranspiration (ET) is a common component in the water and energy cycles, and provides a link between the surface and the atmosphere. Accurate global-scale estimates of ET are critical for better understanding climate and hydrological interactions. Local scale ET observations are available from the FLUXNET project [*Baldocchi et al.*, 2001]. However, dense global coverage by such point measurements is not feasible and the representativeness of point-scale in-situ measurements for larger areas is a subject of active research.

To address this limitation, several alternative global multi-year ET datasets have been derived in recent years. These datasets include satellite-based estimates, land surface models driven with observations-based forcing, reanalysis data products, estimates based on empirical upscaling of point observations, and atmospheric water balance estimates. The LandFlux-EVAL project (see `www.iac.ethz.ch/url/research/LandFlux-EVAL`) aims at evaluating and comparing these currently available ET datasets. The effort forms a key component of the Global Energy and Water Cycle Experiment (GEWEX) LandFlux initiative, a GEWEX Radiation Panel program that seeks to develop a consistent and high-quality global ET dataset for climate studies. Knowledge of the uncertainties in available ET products is a prerequisite for their use in many applications, in particular for the evaluation of climate-change projections [e.g. *Boe and Terray*, 2008; *Seneviratne et al.*, 2010]. We provide here an analysis of 30 observations-based multi-year global ET datasets

---

[2]LERMA, Observatoire de Paris, Paris,

for the 1989–1995 time period, focusing on inter-product spread in various river basins. In addition, we analyze ET in 11 coupled atmosphere-ocean-land GCMs from the IPCC Fourth Assessment Report (AR4). A complementary analysis for a three-year period (1993–1995) by *Jimenez et al.* [2010] focuses on sensible and latent heat fluxes in a subset of twelve satellite-based, LSM and reanalysis datasets.

2. Data and Methods

The analyzed datasets are subdivided into four categories (Table 1). In the 'diagnostic datasets' category, we include datasets that specifically derive ET from combinations of observations or observations-based estimates, together with relatively simple or empirically-derived formulations. The remaining categories provide ET estimates as a byproduct. The second category includes LSM products driven with observations-based surface meteorological data, while the third includes several atmospheric reanalyses. These first three categories are referred to collectively as 'reference datasets' in the context of assessing the IPCC AR4 estimates. IPCC AR4 simulations from 11 GCMs form the fourth category. An overview of the datasets can be found in Table 1. For a detailed description, the reader is referred to the auxiliary material.

The subdivision of the datasets in the first three categories is somewhat arbitrary, since they are all based to some degree on observations and modeling assumptions. Thus, it cannot be inferred a priori that one category of datasets may be closer to actual ET. In addition, several datasets are not independent, since they use common calibration or forcing datasets, and/or common model assumptions (ET parametrization).

France.

The analyses are performed for the common period 1989–1995. The calculation of the interquartile ranges (IQR) and standard deviations presented below are based on the categories (see Table 1), giving each dataset equal weight. Only land pixels that are common to all datasets (excluding Greenland and the Sahara, where ET values are generally low) are considered for the analyses.

3. Results and Discussion

3.1. Annual means and global patterns

Figure 1 (a) displays the mean annual land ET values of each analyzed dataset, as well as the means and the standard deviations within each category. The values are around 1.59 +/-0.19 mm/d (46 +/-5 $\mathrm{W\,m^{-2}}$), a value close to the reanalyses estimates given in *Trenberth et al.* [2009] for two different time periods. The standard deviation of the IPCC AR4 simulations (0.16 mm/d or 4.6 $\mathrm{W\,m^{-2}}$) is lower than those of the reference datasets (standard deviations ranging from 0.17 to 0.19 mm/d or 4.9 to 5.6 $\mathrm{W\,m^{-2}}$). The standard deviation of the GSWP LSMs is still smaller (0.12 mm/d or 3.6 $\mathrm{W\,m^{-2}}$) than that of the IPCC AR4 simulations.

Global patterns of ET for 1989–1995 are displayed in Figure 1 (b–p). The mean values of the four categories (first column) reveal high congruence (for example high ET in the tropics, and lower ET in higher latitudes), and nearly no regions with significant differences (5% Level, Wilcoxon Rank-Sum Test) are found in respective comparisons with the mean of all reference datasets (third column), except for the IPCC AR4 category. In this

---

[3]now at Center for Climate Systems

category, the ET values compare well with the reference datasets in many regions (1k–o), but ET values are significantly lower in the IPCC AR4 simulations in India and South America, and significantly higher in semi-arid regions such as western Australia, western China and the western USA. Overall, the IPCC AR4 simulations appear to underestimate ET gradients within continents (e.g. in North and South America, in Asia north and south of the Himalaya, and in Australia), which could be related to the generally coarse resolution of the models.

The relative IQR (IQR divided by the median, second column) of the LSMs is lower than those of the other categories in Australia and in tropical regions, probably because many of the LSMs share a common forcing (GSWP, GLDAS), but higher in e.g. most of Europe. The IQR of the diagnostic datasets is, compared to the other reference datasets, high in e.g. Australia, and southern and central Africa, but much smaller in Europe.

The IPCC AR4 simulations display higher inter-model deviations than the reference datasets in semi-arid regions such as Australia, India, South Africa, and parts of the Tibetan plateau (1l,o). Accordingly, the IQR of the models in these regions (1p) is much higher. On the other hand, some regions show markedly less inter-model spread across the IPCC AR4 simulations than would be expected based on the uncertainties inferred from the reference datasets (e.g. tropical Africa, East Asia, central Europe, eastern USA). Thus, climate models may share common biases in these regions, either related to biases in forcing (precipitation, clouds, radiation) or in the representation of land hydrology.

Modeling, ETH Zurich, Zurich, Switzerland.

3.2. Basin-scale analysis

Multi-year ET values of all analyzed datasets are displayed in Figure 2 as the deviation from the reference datasets' mean for selected basins (Mississippi, Amazon, central European basins, Volga, Nile, Changjiang, Murray-Darling). The catchment definitions from *Hirschi et al.* [2006] are used for the computation (see Figure 2 bottom). Plots for individual seasons (May to June (MAM), June to August (JJA), September to November (SON), and December to February (DJF)) are provided in the auxiliary material. Datasets are sorted into the four categories (separate bars). Additionally, ET estimated from the difference between precipitation (P) derived from the Global Precipitation Climatology Project (GPCP) and runoff (R) from local measurements is shown for multi-year means (ET=P−R is not generally valid for shorter time scales) in the Mississippi, central European, Volga, Changjiang and Murray-Darling basins. The P−R values can be seen as a long-term constraint on ET (indicated with red lines where available), although multi-year anomalies of terrestrial water storage cannot be excluded in some regions. Overall, the P−R values are found to be close to the reference datasets in the Mississippi, central European and Murray-Darling basins.

The absolute intra-category spreads are largest in the Amazon basin, where the highest ET rates occur. The second largest spreads are found in the Murray-Darling basin during SON and DJF, most pronounced in the IPCC AR4 simulations (see auxiliary material). Comparing the four dataset categories, the intra-category spreads are similar. However,

---

[4]MeteoSwiss, Zurich, Switzerland.

the values can differ largely between basins. In the Changjiang basin for example, the reanalyses and IPCC AR4 simulations display notably higher ET rates than the other dataset categories (up to 0.75 mm/d on average during MAM; see auxiliary material). The intra-category spreads of the IPCC AR4 simulations are much larger than the other categories in the semi-arid Nile and Murray-Darling basins. ET is water (precipitation) limited in these regions, and since the calculation of ET in the IPCC AR4 simulations is based on modeled precipitation (as compared to observed precipitation in the case of reference datasets), the high variability of ET may be partly explained by the large uncertainties in modeled precipitation.

Despite overall similarities of the ET values within these analyzed dataset categories, individual datasets stand out in some regions and seasons. For example, during MAM and in the annual mean, the NCEP reanalysis exhibits above average ET values in the Mississippi, central European, Volga and the Amazon basins. The GFDL IPCC simulation stands out in the Amazon basin during SON (auxiliary material). Note that outliers among the reference datasets are not necessarily erroneous. Indeed, congruence across ET datasets may be induced by the use of common data forcing or model algorithms, rather than the correct representation of ET, as several of the considered products are not independent (see next section).

3.3. Cluster analysis

---

[5]ECMWF, Reading, UK.

In order to study the inter-relationship between the individual datasets, a hierarchical cluster analysis of the multi-year mean ET values is performed (Figure 3). The cluster analysis sorts the datasets into groups in a way that the degree of association between two datasets belonging to the same group is maximal. The criterion used for our analysis is the Euclidean distance between datasets on each land grid cell. Datasets in the same branch of the cluster tree share similar global patterns. The strongest dataset cluster is built by the GSWP simulations (with GS-COLA being the only GSWP model outside the cluster). Most of the IPCC models also form a common branch in the cluster tree. However, the diagnostic datasets and reanalyses are separated into two different main branches of the cluster tree. This indicates that these datasets, although based on observations, exhibit distinct spatial patterns. All the reanalysis datasets are constrained by different exogenous data and some of them are on different main branches of the tree. Note also that simulations using the same model but a different forcing (Mosaic, driven with both GSWP and GLDAS forcing) are separated into two main branches. These findings suggest that forcing can be critical for the resulting ET patterns.

4. Conclusions

This study provides an overview and evaluation of 41 global land ET datasets for the 1989–1995 time period. Comparing IPCC AR4 GCM simulations with datasets which include some observational information (reference datasets), similarities can be found regarding their global patterns and level of uncertainty (interquartile ranges) in most

---

[6]LSCE, UMR CEA-CNRS,

regions. In their global average, the IPCC AR4 simulations show a smaller spread than the categories and groups that are partly based on observations, except for LSMs from the GSWP, which are driven with common forcing data. In addition, climate models display narrower inter-model range than the reference datasets in some regions, which may suggest shared biases. However the uncertainty of the observational datasets prevents evaluation of the magnitude of this bias.

To reduce uncertainty in ET estimates, besides improving ET models, further collection of 'ground truth' observations to validate and force the models continues to be essential, especially in data-poor regions. More refined analyses may allow a reduction of the uncertainty range in observations-based ET products, by identifying whether given outliers can be excluded based on physical considerations [e.g. *McCabe et al.*, 2008]. Such analyses should nonetheless also consider the lack of independence among certain products, which may lead to an underestimation of ET uncertainty. This is well illustrated by the analysis of the GSWP simulations, which e.g. form a strong cluster in the cluster analysis performed for global ET values of all datasets. Further analyses of the datasets collected as part of the LandFlux-EVAL project will allow addressing some of these questions.

Acknowledgments. The GPCP precipitation data were provided by the NASA GSFC's Laboratory for Atmospheres. NCEP reanalysis data were retrieved from `www.esrl.noaa.gov/psd`. The JRA-25 data are from the Japan Meteorological Agency and the Central Research Institute of Electric Power Industry. We acknowledge the Global Modeling and Assimilation Office and the GES DISC for the dissemination of MERRA.

Gif-sur-Yvette, France.

We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling for making available the WCRP CMIP3 dataset. Runoff data were obtained from the GRDC in Koblenz (Germany), the U.S. Geological Survey, the South Australian Surface Water Archive and the Changjiang Water Resource Commission in Wuhan, China.

References

Baldocchi, D., et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Am. Meteorol. Soc.*, *82*(11), 2415–2434.

Boe, J., and L. Terray (2008), Uncertainties in summer evapotranspiration changes over Europe and implications for regional climate change, *Geophys. Res. Lett.*, *35*(5), L05,702, doi:10.1029/2007GL032417.

Bosilovich, M. G. (2008), NASA's modern era retrospective-analysis for research and applications: Integrating Earth observations., *www.earthzine.org/2008/09/26/nasas-modern-era-retrospective-analysis*, *September 26th*.

Dee, D., and S. Uppala (2008), Variational bias correction in ERA-Interim, *ECMWF Technical Memoranda 575*, European Centre for Medium-Range Weather Forecasts.

Dirmeyer, P. A., X. A. Gao, M. Zhao, Z. C. Guo, T. K. Oki, and N. Hanasaki (2006), GSWP-2 - Multimodel anlysis and implications for our perception of the land surface,

---

[7]Center for Ocean-Land-Atmosphere

Bull. Am. Meteorol. Soc., *87*(10), 1381–1397.

Fisher, J. B., K. P. Tu, and D. D. Baldocchi (2008), Global estimates of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, *Rem. Sens. Env.*, *112*(3), 901–919, doi:10.1016/j.rse.2007.06.025.

Hirschi, M., S. I. Seneviratne, and C. Schär (2006), Seasonal variations in terrestrial water storage for major midlatitude river basins, *J. Hydrometeorol.*, *7*(1), 39–60.

Jimenez, C., et al. (2010), Global inter-comparison of selected 1993-1995 monthly averaged land surface heat fluxes, *J. Geophys. Res.*, *in press.*

Jung, M., et al. (2010), Recent decline in the global land evapotranspiration trend due to limited moisture supply, *Nature*, *467*(7318), 951–954, doi:10.1038/nature09396.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*(3), 437–471.

Krinner, G., N. Viovy, N. de Noblet-Ducoudre, J. Ogee, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. C. Prentice (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Gob. Biogeo. Cyc.*, *19*(1), GB1015, doi:1029/2003GB002199.

McCabe, M. F., E. F. Wood, R. Wojcik, M. Pan, J. Sheffield, H. Gao, and H. Su (2008), Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, *Rem. Sens. Env.*, *112*(2), 430–444., doi:10.1016/j.rse.2007.03.027.

Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor (2007), The WCRP CMIP3 multimodel dataset - A new era

Studies, Calverton, USA.

in climate change research, *Bull. Am. Meteorol. Soc.*, *88*(9), 1383–1394.

Mueller, B., M. Hirschi, and S. I. Seneviratne (2010), New diagnostic estimates of variations in terrestrial water storage based on ERA-Interim data, *Hydrol. Process.*, *in press*, doi:10.1002/hyp.7652.

Onogi, K., et al. (2007), The JRA-25 reanalysis, *J. Meteorol. Soc. Jpn.*, *85*(3), 369–432.

Reichle, R., and coauthors (2010), Assessment and enhancement of MERRA land surface hydrology estimates, *submitted to J. Climate*.

Rodell, M., et al. (2004), The global land data assimilation system, *Bull. Am. Meteorol. Soc.*, *85*(3), 381–394, doi:10.1175/BAMS-85-3-381.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth-Sci. Rev.*, *99*(3-4), 125–161, doi:10.1016/j.earscirev.2010.02.004.

Sheffield, J., and E. F. Wood (2007), 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*(17), D17,115–1–21, doi:10.1029/2006JD008288.

Sheffield, J., E. F. Wood, and F. Munoz-Arriola (2010), Long-term regional estimates of evapotranspiration for Mexico based on downscaled ISCCP data, *J. Hydrometeorol.*, *11(2)*, 253–275, doi:10.1175/2009JHM1176.1.

---

[8]NASA Jet Propulsion Laboratory,

Trenberth, K. E., J. T. Fasullo, and J. Kiehl (2009), Earth's global energy budget, *Bull. Am. Meteorol. Soc.*, *90*(3), 311–323.

California Institute of Technology, Pasadena, USA.

[9]Max Planck Institute for Biogeochemistry, Jena, Germany.

[10]School of Civil and Environmental Engineering, University of NSW, Sydney, Australia.

[11]NASA-GSFC, Greenbelt, MD, USA.

[12]Department of Civil and Environmental Engineering, Princeton University, Princeton, USA.

[13]Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, NL.

[14]Department of Geography, University of Maryland, College Park, USA.

[15]CSIRO Land and Water, Canberra, Australia.

Wang, K. C., and S. L. Liang (2008), An improved method for estimating global evapotranspiration based on satellite determination of surface net radiation, vegetation index, temperature, and soil moisture, *J. Hydrometeorol.*, *9*(4), 712–727, doi: 10.1175/2007JHM911.1.

Zhang, Y. Q., R. Leuning, L. B. Hutley, J. Beringer, I. McHugh, and J. P. Walker (2010), Using long-term water balances to parameterize surface conductances and calculate evaporation at 0.05 degrees spatial resolution, *Water Resour. Res.*, *46*, W05,512, doi: 10.1029/2009WR008716.

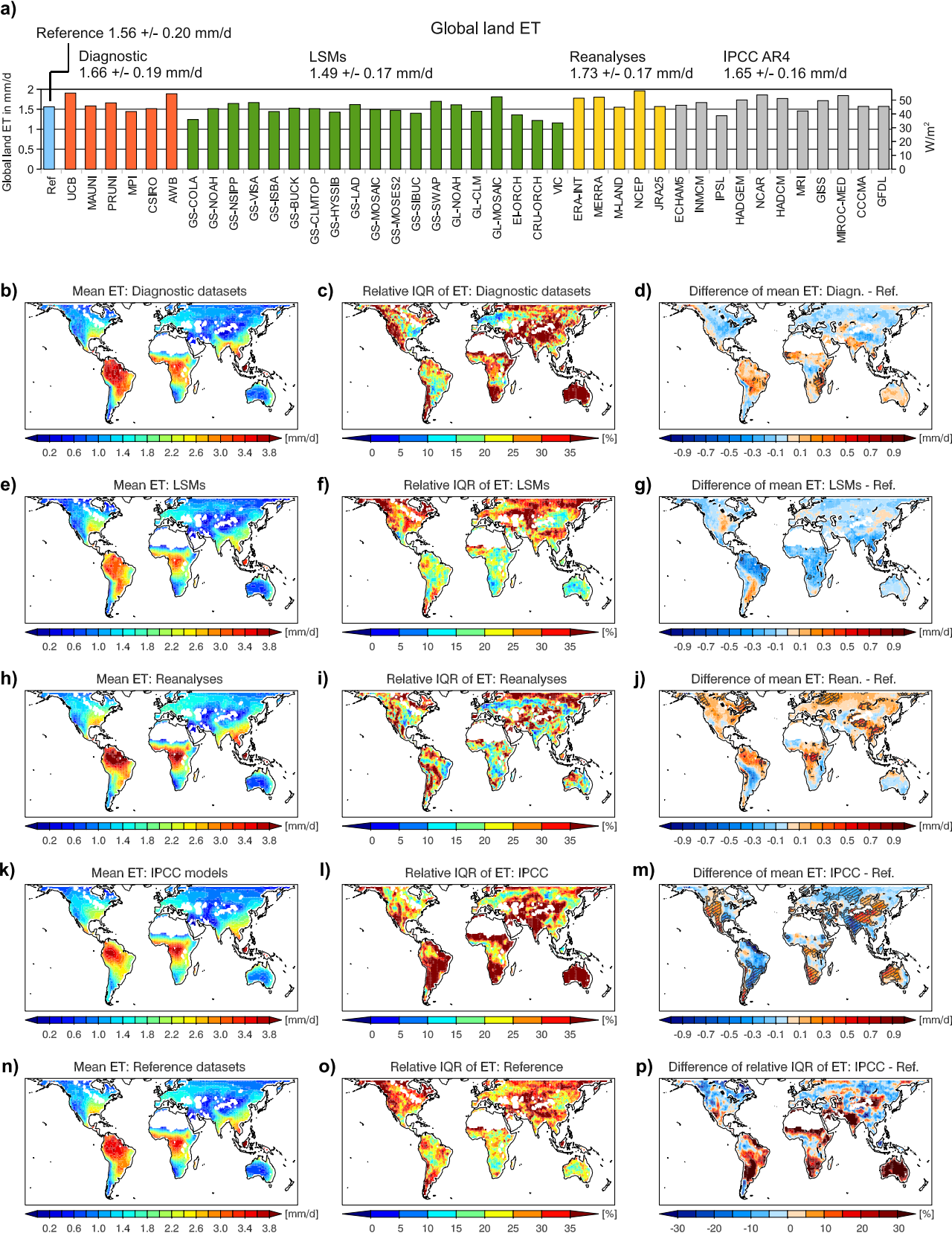
Table 1. Overview of employed ET datasets. Ct: Category.

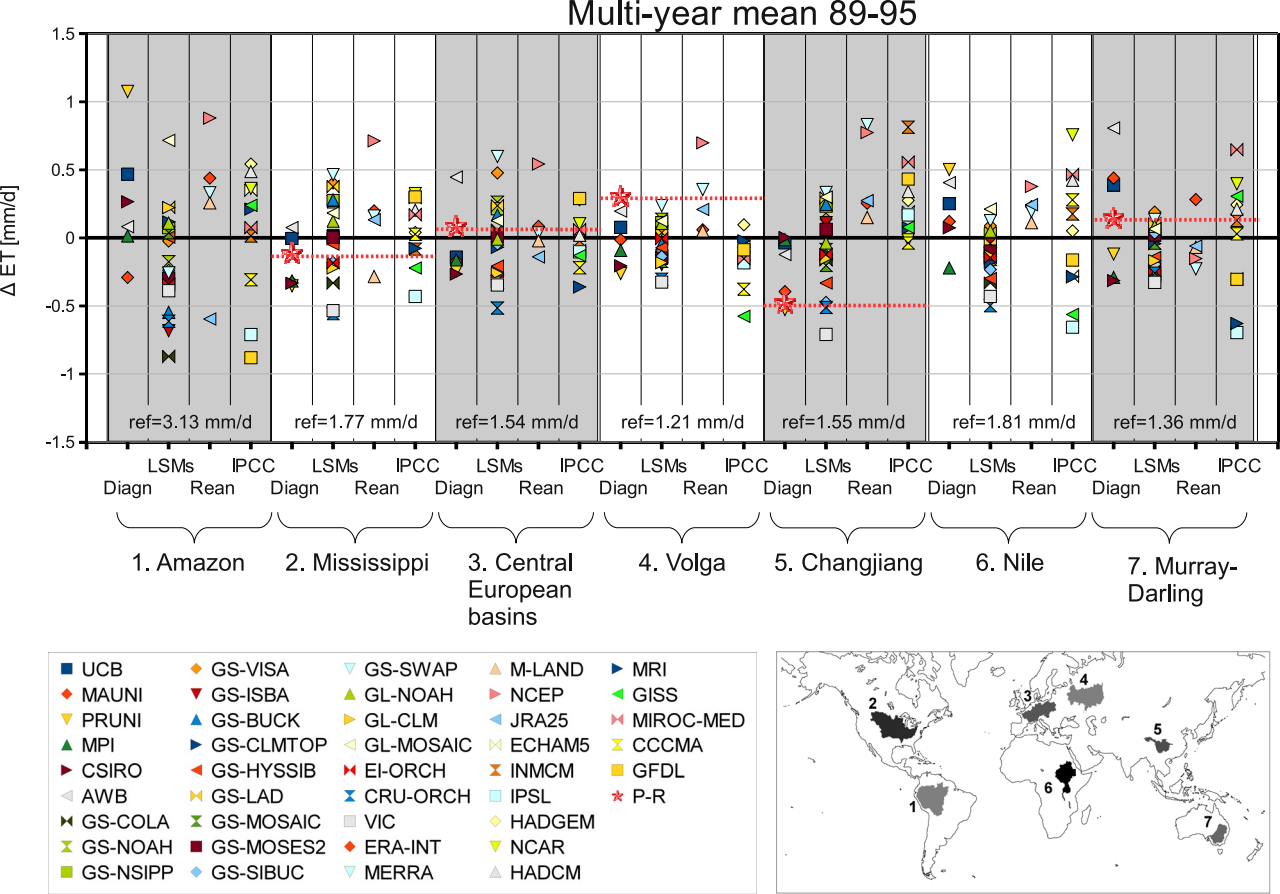
	Ct	Group	Dataset	Reference	Information
Observations-based datasets	Diagnostic datasets		UCB	Fisher et al. (2008)	Priestley-Taylor, ISLSCP-II (SRB, CRU, AVHRR)
			MAUNI	Wang and Liang (2008)	Empirical, calibrated with Ameriflux, ISLSCP-II (SRB, CRU, AVHRR)
			PRUNI	Sheffield et al. (2010)	Penman-Monteith ET, ISCCP, AVHRR
			MPI	Jung et al. (2010)	Empirical upscaling of FLUXNET, CRU, GPCC, AVHRR
			CSIRO	Zhang et al. (2010)	Penman-Monteith-Leuning ET
			AWB	Mueller et al. (2010)	Atmospheric water balance (GPCP, ERA-Interim)
	LSMs	GSWP	GS-COLA, GS-NOAH, GS-NSIPP, GS-VISA, GS-ISBA, GS-BUCK, GS-CLMTOP, GS-HYSSIB, GS-LAD, GS-MOSAIC, GS-MOSES2, GS-SIBUC, GS-SWAP	Dirmeyer et al. (2006)	13 GSWP LSM simulations, forced with ISLSCP-II and/or reanalysis data
		GLDAS	GL-NOAH, GL-CLM, GL-MOSAIC	Rodell et al. (2004)	
		ORCH	EI-ORCH	Krinner et al. (2005)	ORCHIDEE LSM with ERA-Interim forcing
			CRU-ORCH		ORCHIDEE LSM with CRU-NCEP forcing
		VIC	VIC	Sheffield and Wood (2007)	LSM
	Reanalyses		ERA-INT	Dee and Uppala (2008)	ERA-Interim Reanalysis
			MERRA	Bosilovich (2008)	Reanalysis
			M-LAND	Reichle et al. (2010)	MERRA-Land Reanalysis
			NCEP	Kalnay et al. (1996)	Reanalysis
			JRA25	Onogi et al. (2007)	Reanalysis
			Reference datasets		Statistics (mean, IQR, standard deviation) of ensemble of single observations-based datasets (30 in total)
GCMs	IPCC AR4		ECHAM5, INMSM, IPSL, HADGEM, NCAR, HADCM, MRI, GISS, MIROC-MED, CCCMA, GFDL	Meehl et al. (2007)	AR4 simulations (20c3m) from 11 global climate models

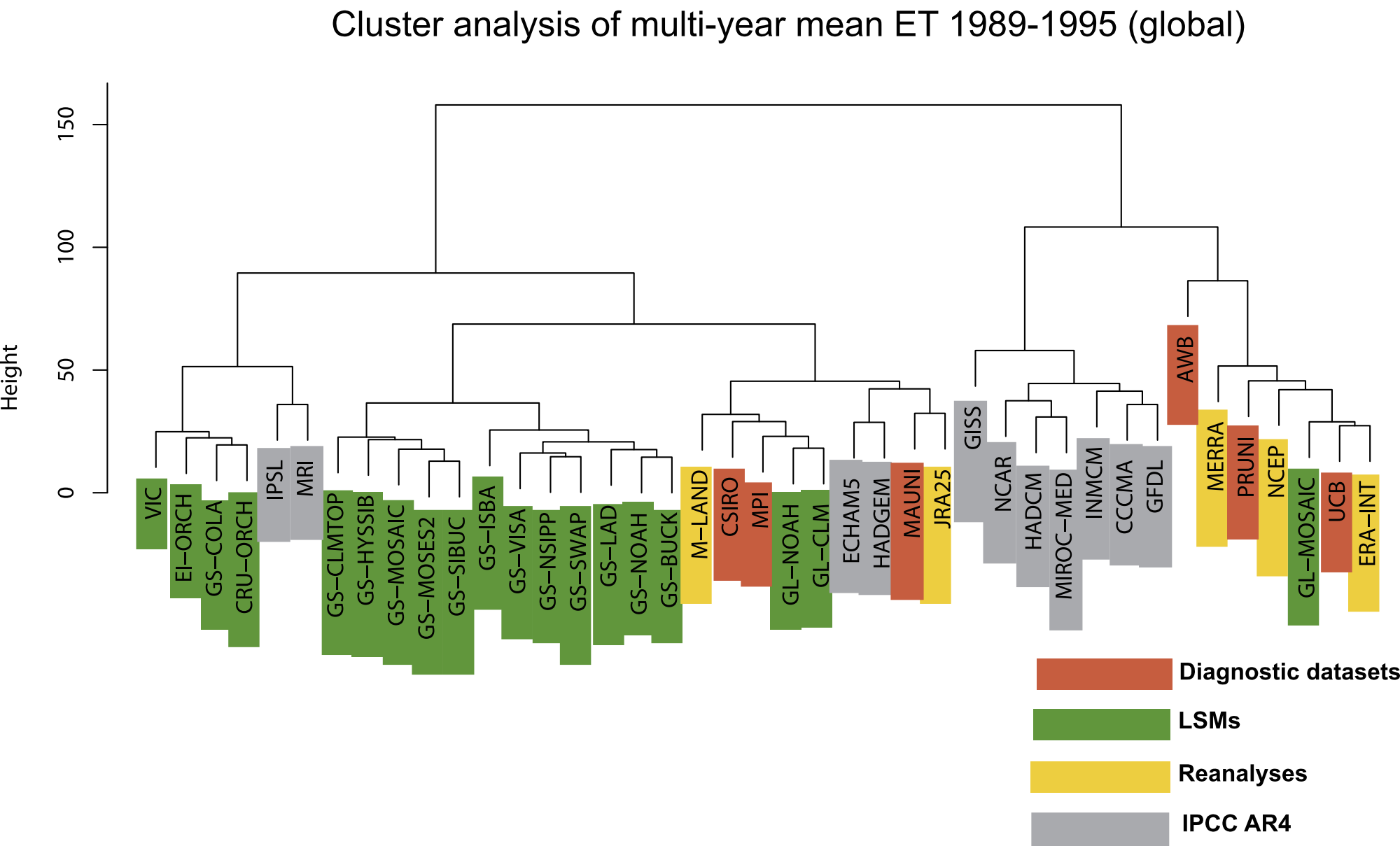
Figure 1. Mean global land ET values for each dataset (a) with mean and standard deviation for each category (numbers). Mean, relative interquartile range (IQR) and difference of mean to mean of reference datasets (Ref.) of the diagnostic datasets (b-d), LSMs (e-g), reanalyses (h-j), and IPCC AR4 simulations (k-m). Mean (n) and relative IQR (o) of the reference datasets and difference of relative IQRs IPCC AR4 to reference datasets (p). Hatched areas in d,g,j,m show a nominal 5%-significance level as heuristic descriptive indicator (Wilcoxon Rank-Sum test).

Figure 2. Deviations of each dataset from the reference datasets' mean (displayed on the bottom as 'ref') over 1989–1995 (multi-year means) for the selected river basins (1) Amazon, (2) Mississippi, (3) Central European basins, (4) Volga, (5) Changjiang, (6) Nile, (7) Murray-Darling. The datasets are grouped into diagnostic datasets (Diagn), LSMs (LSMs), reanalyses (Rean) and IPCC AR4 simulations (IPCC). P-R values are marked with red stars and dashed lines. Location of river basins (right bottom).

Figure 3. Hierarchical cluster analysis of global ET values, averaged over 1989–1995, using Euclidean distance matrix. Diagnostic datasets (red), LSMs (green), reanalyses (yellow) and IPCC AR4 simulations (grey).


a)
Global land ET
Reference 1.56 +/- 0.20 mm/d
Diagnostic
1.66 +/- 0.19 mm/d
LSMs
1.49 +/- 0.17 mm/d
Reanalyses
1.73 +/- 0.17 mm/d
IPCC AR4
1.65 +/- 0.16 mm/d
Global land ET in mm/d
W/m²
Ref
UCB
MAUNI
PRUNI
MPI
CSIRO
AWB
GS-COLA
GS-NOAH
GS-NSIPP
GS-VISA
GS-ISBA
GS-BUCK
GS-CLMTOP
GS-HYSSIB
GS-LAD
GS-MOSAIC
GS-MOSES2
GS-SIBUC
GS-SWAP
GL-NOAH
GL-CLM
GL-MOSAIC
EI-ORCH
CRU-ORCH
VIC
ERA-INT
MERRA
M-LAND
NCEP
JRA25
ECHAM5
INMCM
IPSL
HADGEM
NCAR
HADCM
MRI
GISS
MIROC-MED
CCCMA
GFDL
b) Mean ET: Diagnostic datasets
c) Relative IQR of ET: Diagnostic datasets
d) Difference of mean ET: Diagn. - Ref.
e) Mean ET: LSMs
f) Relative IQR of ET: LSMs
g) Difference of mean ET: LSMs - Ref.
h) Mean ET: Reanalyses
i) Relative IQR of ET: Reanalyses
j) Difference of mean ET: Rean. - Ref.
k) Mean ET: IPCC models
l) Relative IQR of ET: IPCC
m) Difference of mean ET: IPCC - Ref.
n) Mean ET: Reference datasets
o) Relative IQR of ET: Reference
p) Difference of relative IQR of ET: IPCC - Ref.
[mm/d]
[%]


Multi-year mean 89-95
Δ ET [mm/d]
1.5
1
0.5
0
-0.5
-1
-1.5
ref=3.13 mm/d
ref=1.77 mm/d
ref=1.54 mm/d
ref=1.21 mm/d
ref=1.55 mm/d
ref=1.81 mm/d
ref=1.36 mm/d
LSMs
IPCC
Diagn
Rean
1. Amazon
2. Mississippi
3. Central European basins
4. Volga
5. Changjiang
6. Nile
7. Murray-Darling
UCB
MAUNI
PRUNI
MPI
CSIRO
AWB
GS-COLA
GS-NOAH
GS-NSIPP
GS-VISA
GS-ISBA
GS-BUCK
GS-CLMTOP
GS-HYSSIB
GS-LAD
GS-MOSAIC
GS-MOSES2
GS-SIBUC
GS-SWAP
GL-NOAH
GL-CLM
GL-MOSAIC
EI-ORCH
CRU-ORCH
VIC
ERA-INT
MERRA
M-LAND
NCEP
JRA25
ECHAM5
INMCM
IPSL
HADGEM
NCAR
HADCM
MRI
GISS
MIROC-MED
CCCMA
GFDL
P-R
1
2
3
4
5
6
7


Cluster analysis of multi-year mean ET 1989-1995 (global)
Height
0
50
100
150
VIC
EI–ORCH
GS–COLA
CRU–ORCH
IPSL
MRI
GS–CLMTOP
GS–HYSSIB
GS–MOSAIC
GS–MOSES2
GS–SIBUC
GS–ISBA
GS–VISA
GS–NSIPP
GS–SWAP
GS–LAD
GS–NOAH
GS–BUCK
M–LAND
CSIRO
MPI
GL–NOAH
GL–CLM
ECHAM5
HADGEM
MAUNI
JRA25
GISS
NCAR
HADCM
MIROC–MED
INMCM
CCCMA
GFDL
AWB
MERRA
PRUNI
NCEP
GL–MOSAIC
UCB
ERA–INT
Diagnostic datasets
LSMs
Reanalyses
IPCC AR4