

# Systematic land climate and evapotranspiration biases in CMIP5 simulations

B. Mueller<sup>1,2</sup> and S. I. Seneviratne<sup>1</sup>

Received 25 September 2013; revised 27 November 2013; accepted 3 December 2013; published 8 January 2014.

[1] Land climate is important for human population since it affects inhabited areas. Here we evaluate the realism of simulated evapotranspiration (ET), precipitation, and temperature in the CMIP5 multimodel ensemble on continental areas. For ET, a newly compiled synthesis data set prepared within the Global Energy and Water Cycle Experiment-sponsored LandFlux-EVAL project is used. The results reveal systematic ET biases in the Coupled Model Intercomparison Project Phase 5 (CMIP5) simulations, with an overestimation in most regions, especially in Europe, Africa, China, Australia, Western North America, and part of the Amazon region. The global average overestimation amounts to 0.17 mm/d. This bias is more pronounced than in the previous CMIP3 ensemble (overestimation of 0.09 mm/d). Consistent with the ET overestimation, precipitation is also overestimated relative to existing reference data sets. We suggest that the identified biases in ET can explain respective systematic biases in temperature in many of the considered regions. The biases additionally display a seasonal dependence and are generally of opposite sign (ET underestimation and temperature overestimation) in boreal summer (June–August). **Citation:** Mueller, B., and S. I. Seneviratne (2014), Systematic land climate and evapotranspiration biases in CMIP5 simulations, *Geophys. Res. Lett.*, *41*, 128–134, doi:10.1002/2013GL058055.

## 1. Introduction

[2] The multimodel climate experiments conducted as part of the different phases of the Coupled Model Intercomparison Project (CMIP) are essential contributions for the evaluation of past and future climate change in recent and on-going reports of the Intergovernmental Panel on Climate Change (IPCC) [e.g., Solomon *et al.*, 2007]. It is thus of critical importance to evaluate the reliability of these simulations.

[3] While there exist good-quality observations for several climate variables, large uncertainties remain in the

evaluation of land-surface fluxes [e.g., Seneviratne *et al.*, 2010; Mueller *et al.*, 2011; Jimenez *et al.*, 2011; Sheffield *et al.*, 2012; Seneviratne, 2012]. In particular, observations of evapotranspiration (ET) have traditionally been scarce, and global-scale data sets can only be derived indirectly. Nonetheless, there have been substantial advances in this research area in recent years, and new data sets are now available, each combining observational information and numerical algorithms in alternative ways [Mueller *et al.*, 2011; Jimenez *et al.*, 2011, see [www.iac.ethz.ch/url/research/LandFlux-EVAL](http://www.iac.ethz.ch/url/research/LandFlux-EVAL)]. As part of the Global Energy and Water Exchanges Project (GEWEX)-sponsored LandFlux activity, the LandFlux-EVAL project recently compiled several synthesis data sets based on existing global ET data sets [Mueller *et al.*, 2013], providing for the first time the “best-guess” estimate of this quantity.

[4] In the present article, we use one of these recently compiled data sets to evaluate the realism of land ET fields in simulations of the fifth phase of the CMIP project (CMIP5), which serve as basis for the fifth assessment report of the IPCC. These analyses are complemented with comparisons to the earlier CMIP3 multimodel ensemble (which served as basis for the fourth assessment report of the IPCC). Several recent studies have shown that soil moisture and ET have a strong influence on temperature in many regions of the world [Seneviratne *et al.*, 2006; Koster *et al.*, 2006; Mueller and Seneviratne, 2012], since latent heat flux associated with ET uses a substantial fraction of the available net radiation on land. Whenever ET is reduced, the corresponding energy is instead used by sensible heat flux, leading to higher air temperatures. Due to this strong relation between evaporation and temperature, we validate also temperature fields over land and assess where their biases can be related to biases in ET. We also include an evaluation of precipitation in the CMIP simulations in order to investigate the causes of the ET biases.

## 2. Data and Methods

### 2.1. Observational Estimates

[5] As reference ET data, we use here the recently compiled LandFlux-EVAL synthesis product based on all data set categories and computed for the years 1989–1995 [see Mueller *et al.*, 2013]. A total of 40 observation-based data sets, land-surface model output, and atmospheric reanalyses are included in this synthesis product. The statistics used are the median and interquartile range from the long-term mean or monthly values. Note that the median of most included data sets agreed well with observational estimates of precipitation minus runoff for large river basins [Mueller *et al.*, 2011].

Additional supporting information may be found in the online version of this article.

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

<sup>2</sup>Climate Research Division, Environment Canada, Toronto, Ontario, Canada.

Corresponding author: B. Mueller, Climate Research Division, Environment Canada, Toronto, ON M3H 5T4, Canada. ([brigitte.mueller@ec.gc.ca](mailto:brigitte.mueller@ec.gc.ca))

©2013. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made. 0094-8276/14/10.1002/2013GL058055

**Table 1.** List of Included CMIP Models, Their Original Resolution, and the Respective Root-Mean-Square Error (From Monthly Values) of ET, Precipitation, and Temperature (Excluding Greenland and Antarctica for ET) During 1989–1995

	Model Name	Model Resolution	RMSE ET (mm/d)	Precipitation (mm/d)	Temperature (K)	Biases ET (mm/d)
CMIP5	BCC_CSM	128 × 64	1.11	2.37	8.79	0.00
	CanESM2	128 × 64	0.77	1.70	8.09	−0.03
	CNRM_CM5	256 × 128	2.37	2.34	10.29	0.15
	CSIRO_Mk3	192 × 96	1.70	2.11	5.40	0.10
	GFDL_CM3	144 × 90	3.42	2.47	8.29	0.26
	GISS E2H	144 × 90	3.70	3.87	23.10	0.28
	HadCM3	96 × 73	2.84	2.67	7.88	0.20
	HadGEM2_CC	192 × 145	3.06	2.90	6.88	0.28
	Inmcm4	180 × 120	3.86	3.13	10.43	0.28
	IPSL_CM5A_LR	96 × 96	0.95	1.39	6.84	0.07
	MIROC_ESM	128 × 64	4.07	3.93	15.50	0.30
	MPI_ESM_LR	192 × 96	2.24	1.90	9.05	0.19
	MRI_CGCM3	320 × 160	1.73	2.90	9.18	0.07
	NorESM1	144 × 96	4.11	4.04	5.83	0.25
	Mean of CMIP5	360 × 180	2.42	2.18	5.26	0.17
CMIP3	cccma_cgcm3	96 × 48	1.13	1.61	7.67	0.12
	mpi_echam5	192 × 96	1.12	1.75	8.83	0.09
	gfdl_cm2_0	144 × 90	1.33	1.58	13.29	−0.12
	giss_model_e_r	72 × 46	1.85	1.69	6.58	0.24
	ukmo_hadcm3	96 × 73	2.56	2.05	15.32	0.20
	ukmo_hadgem1	192 × 145	2.59	2.34	9.78	0.20
	inmcm3_0	72 × 45	1.67	1.88	11.75	−0.02
	ipsl_cm4	96 × 72	1.27	1.97	5.44	0.03
	miroc3_2_medres	128 × 64	2.99	2.67	5.73	0.23
	mri_cgcm2_3_2a	128 × 64	0.63	1.46	10.14	−0.03
	ncar_pcm1	128 × 64	3.36	3.54	14.37	0.04
	Mean of CMIP3	360 × 180	1.51	1.30	6.64	0.09

[6] We further use combined precipitation data from the Global Precipitation Climatology Project (GPCP) [Adler *et al.*, 2003] and precipitation from the Climate Research Unit (CRU) [Mitchell and Jones, 2005] Version 3.1 as a reference for the analysis of the CMIP simulations. For temperature, we use the Willmott and Matsuura data set [Willmott and Robeson, 1995, University of Delaware] based on Global Historical Climatology Network and Legates and Willmott data [Legates and Willmott, 1990], as well as data from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis ERA-Interim [Dee *et al.*, 2011].

## 2.2. CMIP3 and CMIP5 Simulations

[7] Simulations of 14 models contributing to the CMIP5 [Taylor *et al.*, 2012] and 11 models contributing to the CMIP3 multimodel experiments have been used (see Table 1 for details). We evaluate the variable “surface latent heat flux” (i. e., ET, including both evaporation and sublimation), near-surface air temperature, and precipitation from the historical CMIP5 and the twentieth century CMIP3 simulations, respectively. All data have been linearly interpolated to a  $1 \times 1^\circ$  resolution, corresponding to the resolution of the LandFlux-EVAL synthesis ET product.

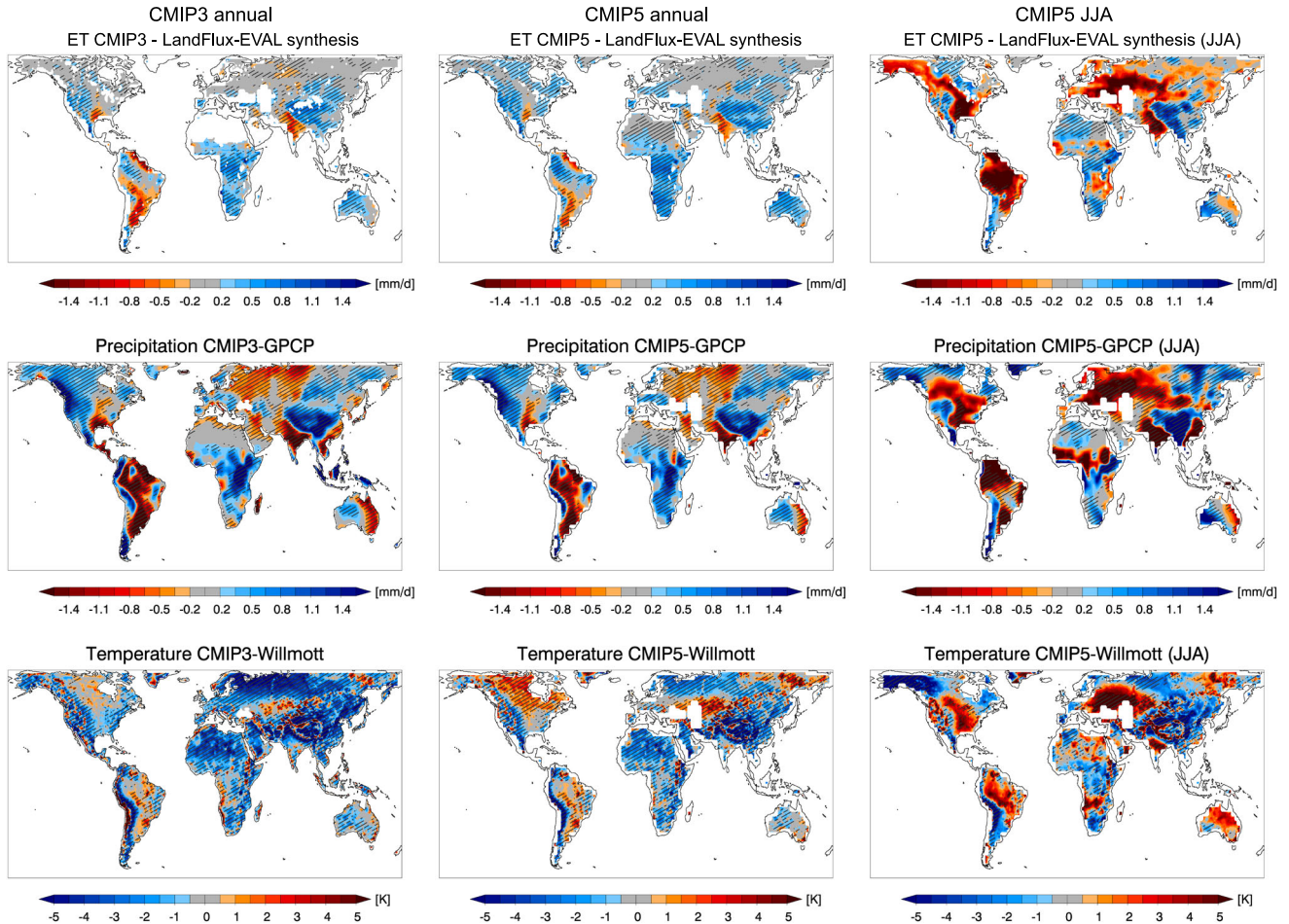
## 3. Results

### 3.1. Biases in ET and Precipitation

[8] Figure 1 (top row) displays the difference in ET between the CMIP simulations and the LandFlux-EVAL synthesis product for annual means (Figure 1, left column: CMIP3 and Figure 1, middle column: CMIP5) and June to August means (JJA; Figure 1, right column: CMIP5, see

Figure S1 in the supporting information for relative differences). Regions where at least 66% of the models (i. e., 10 out of 14 for CMIP5 and 8 out of 11 for CMIP3) agree on the sign of the difference are overlaid with stripes. In the annual average, the CMIP simulations display an overall systematic overestimation of ET in most regions, in particular in Europe, Africa, China, Australia, Western North America, and part of the Amazon region. This bias is slightly more pronounced in CMIP5 than in the previous CMIP3 ensemble (Figure 1), (see also Mueller *et al.* [2011]). India is one of the few regions where the CMIP simulations are too dry. The CMIP5 simulations display consistent (and thus shared) biases in more regions than the CMIP3 simulations.

[9] A more in-depth evaluation of the CMIP5 and CMIP3 ET fields (annual means) for 25 regions based on the definition of the IPCC Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX) [see Seneviratne *et al.*, 2012] is provided in Figure 2. This regional analysis compares the respective spreads of the CMIP3 and CMIP5 ensemble (box plots) and the median (full lines) and interquartile range (dotted lines) of the LandFlux-EVAL synthesis product (red lines for CMIP5 underestimation, blue lines for overestimation). The overall systematic overestimation of ET is confirmed for most considered regions. CMIP simulations (medians) are consistent with the LandFlux-EVAL estimates (interquartile ranges) in 16 (CMIP3) and 12 (CMIP5) out of the 25 regions. The intermodel spread is on average similar for the CMIP3 and the CMIP5 ensemble, indicating no systematic improvement in the newer multimodel ensemble in terms of ET. This fact is confirmed by the average root-mean-square errors for the respective ensembles (see Table 1).



**Figure 1.** (left column) Difference in median of CMIP3 annual means, (middle column) CMIP5 annual means, and (right column) CMIP5 June to August means and reference data sets. The reference product for evapotranspiration (ET; top row) is the LandFlux-EVAL merged synthesis product, for precipitation (middle row) the GPCP data set, and for temperature (bottom row) the Willmott and Matsuura data set. Stripes indicate where at least 66% (i. e., 8 out of 11 for CMIP3 and 10 out of 14 and for CMIP5) of the models agree on the sign of the differences. Note the reversed color schemes for ET and precipitation versus temperature.

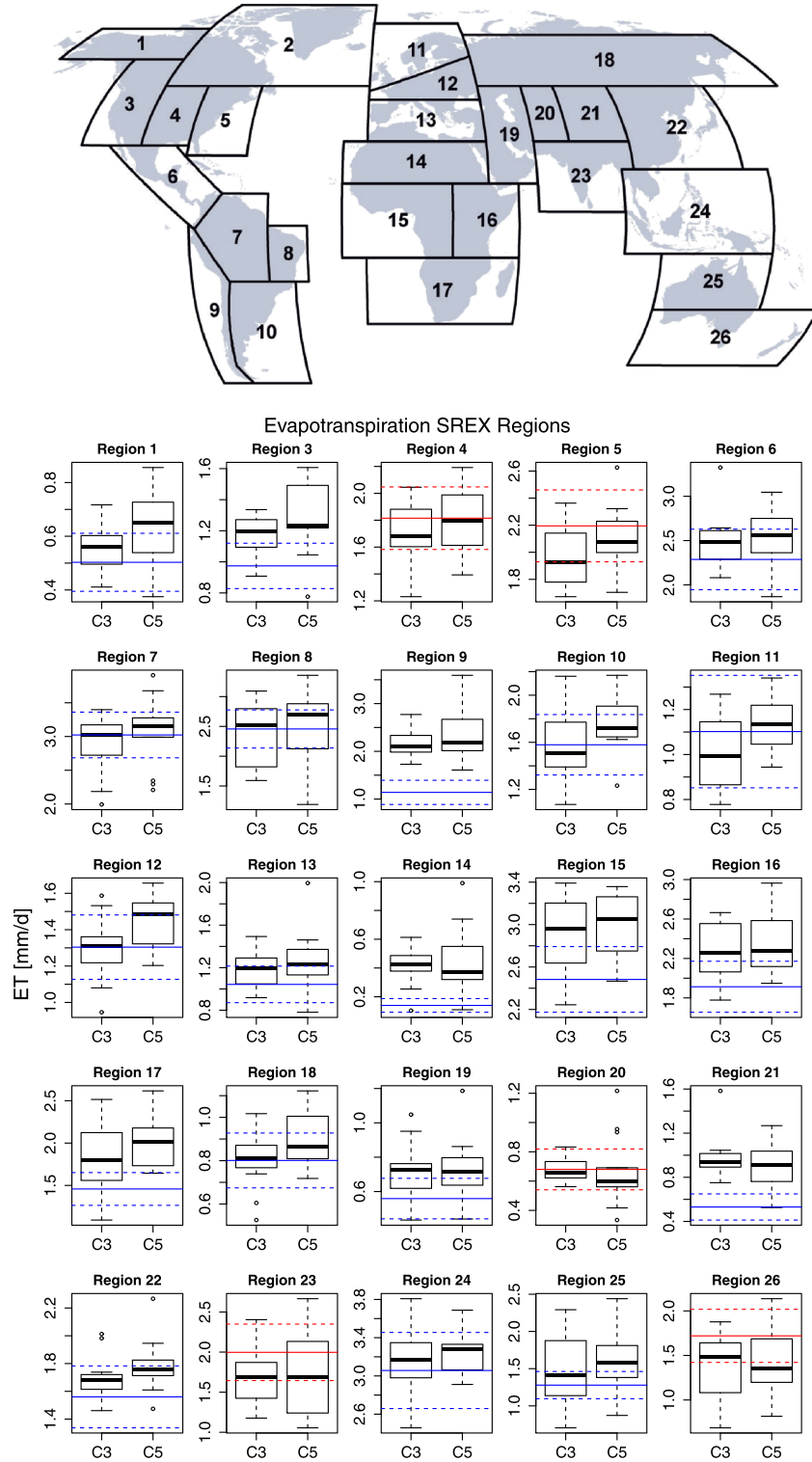
[10] To assess the possible causes and implications of this systematic bias, we also compare the CMIP5 and CMIP3 simulations with observational data sets of precipitation and temperature in Figure 1 (middle and bottom rows). This analysis reveals a systematic overestimation of precipitation in the CMIP simulations in the annual mean values (Figure 1, left and middle columns, see also Figures S2 and S3) in most of the regions displaying a systematic overestimation of ET, thus highlighting precipitation biases as likely driver for the ET biases. Indeed, most of these regions are characterized by a soil moisture-limited ET regime (Figure S4), and thus additional precipitation is generally used by ET in these regions [e.g., *Teuling et al.*, 2009; *Seneviratne et al.*, 2010]. Nonetheless, a few tropical regions display an underestimation of precipitation despite an overestimation of ET, which is realistic in regions where ET is not limited by water availability, but rather by the incoming net radiation [*Nemani et al.*, 2003]. It should also be noted that a few regions display systematic dry biases rather than systematic wet biases (i. e., underestimation of ET and precipitation), such as India, the Mississippi River Basin, and Argentina.

[11] Hence, though there are some exceptions in a few regions, the identified overestimation of ET over a large

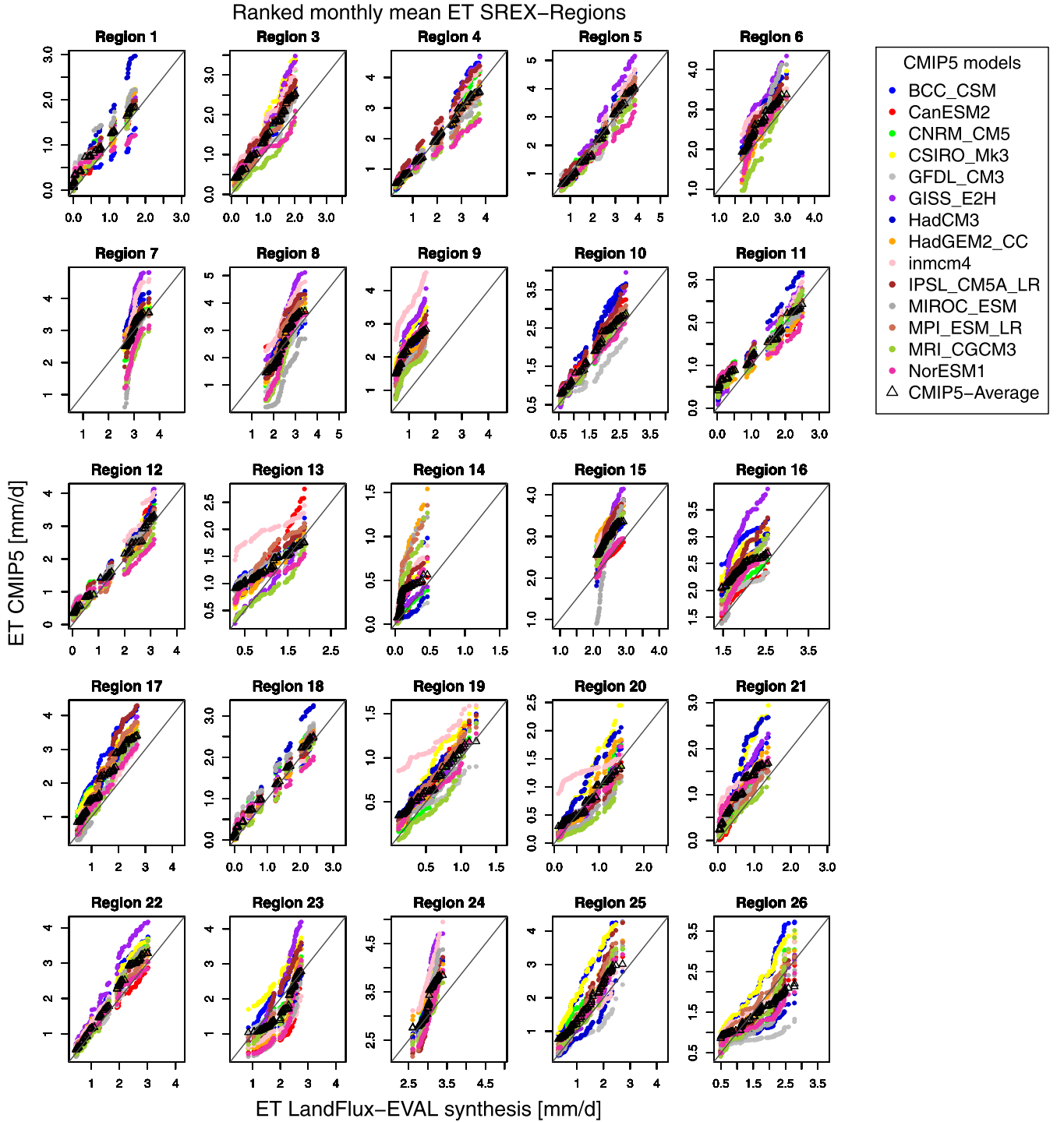
fraction of land areas is found to be driven by a corresponding overestimation of precipitation in these regions. In a global average, the overestimation of ET amounts to 0.17 mm/d (CMIP5) and 0.09 mm/d (CMIP3). The biases for precipitation are similar (0.16 mm/d for CMIP5 and 0.05 mm/d for CMIP3) when GPCP is employed as a reference data set. However, precipitation in GPCP is relatively high compared to other reference data sets. For instance, the precipitation bias of the CMIP5 and CMIP3 ensemble relative to the Climate Research Unit (CRU) precipitation data set is 0.27 mm/d and 0.22 mm/d. Overall, higher overestimation of precipitation than ET appears consistent with the assumption that the former drives the latter (although the presence of feedback cannot be excluded).

### 3.2. Possible Relation to Temperature Biases

[12] The evaluation of temperature in the CMIP5 simulations (Figure 1, bottom row) suggests a tendency for a systematic cold bias in the multimodel ensemble (0.4 and 1.3 K colder than the reference on a global average for CMIP5 and CMIP3, respectively). The temperature biases are even more pronounced when compared to a different reference data set (ERA-Interim, not shown). In regions with



**Figure 2.** (Top) IPCC SREX regions: Domains considered in regional analyses based on the definition used in the IPCC Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Seneviratne *et al.*, 2012]. (Bottom) Box plots for the IPCC SREX regions (except region 2) displaying evapotranspiration (ET) in each region in the CMIP5 (“C5”) and CMIP3 (“C3”) model simulations (statistics for multiyear averages (1989–1995) for the ensemble of analyzed models, see Table 1). The respective ET estimates of the LandFlux-EVAL merged synthesis data set’s median (full line) and interquartile ranges (dashed lines) are indicated in blue (red) in cases of overestimation (underestimation) of the CMIP5 models.



**Figure 3.** Ranked monthly mean ET from CMIP5 versus LandFlux-EVAL synthesis reference. Different colors denote different CMIP5 models.

water-limited regimes (see Figure S4), the underestimation of temperature is likely related to the identified systematic ET overestimation, linked with enhanced evaporative cooling in the model simulations [e.g., Seneviratne *et al.*, 2010]. For similar reasons, it is likely that the temperature overestimation in South America (Figure 1 bottom left and middle) is related to the ET underestimation there (Figure 1 top left and middle, i.e., underestimation of evaporative cooling). One of the few regions where the temperature and ET biases are not of opposite sign is India, where previous

studies disagreed on the existence of strong soil moisture-temperature coupling [Koster *et al.*, 2006; Mueller and Seneviratne, 2012].

[13] Analyses of ET and temperature biases for Northern Hemispheric summer (JJA) are provided in Figure 1 (right column). Interestingly, these analyses reveal biases of opposite sign compared to those on annual time scale in some regions (ET underestimation and temperature overestimation in most regions during JJA) and more pronounced biases in others. Results for the boreal winter are provided



in Figure S5. They are generally more consistent with the annual biases but also display a tendency for low ET and high temperatures in Southern Hemisphere midlatitudes.

[14] As has been shown in recent studies [Boberg and Christensen, 2012; Christensen and Boberg, 2012], CMIP5 simulations display a tendency to overestimate warm temperatures, which is consistent with our seasonal results despite the cold and wet biases identified on annual time scale. Quantile-quantile plots of ET and temperature from CMIP5 versus reference data sets (Figures 3 and S6) show an overestimation of ET in the wettest months as well as an underestimation (overestimation) of temperature in cold (warm) months. A tendency to underestimate ET in dry months is present in a few regions only, which may be due to the bounded lower limit of ET (zero flux). We note that most CMIP5 models display these patterns, but that they are more pronounced in some of the models. These analyses confirm the presence of systematic biases in both the ET and temperature distributions in most regions.

#### 4. Summary and Conclusions

[15] This article provides an evaluation of ET in the recent CMIP5 and CMIP3 multimodel ensembles. These first analyses reveal systematic land hydrological and climate biases in the CMIP simulations, characterized by an overestimation of ET in the yearly average in many land regions. These biases can be generally related to a corresponding overestimation of precipitation in most of the identified regions, which is driving the ET bias. Concomitantly, the excess evaporative cooling leads to a cold bias in most of the identified regions on the annual time scale. These biases are generally found in regions with soil moisture-limited ET regimes, with a few exceptions, and tend to change sign in the warm season in midlatitude regions.

[16] The possibly underlying causes for the identified biases as well as consequences for CMIP5 projections would need to be investigated in follow-up studies. Direct interaction between precipitation (clouds) and temperature may also be possible. Furthermore, the direction of causality cannot always be fully established, and it is possible that in some regions the diagnosed biases are unrelated. For instance, in energy-limited regions, one would expect the opposite relationship, i. e., temperature biases would cause ET biases (of same sign rather than opposite sign). However, these results highlight the presence of remaining systematic biases in state-of-art climate models, which are similar in magnitude for the CMIP5 and the CMIP3 ensembles (and even larger for the former). Previous studies pointed out the fact that extensive model developments were undertaken between the two series of experiments [e.g., Knutti and Sedlacek, 2013]. Our results suggest that these new model developments and ensembles did not necessarily lead to a better performance for key variables such as ET, precipitation, and temperature over land.

[17] **Acknowledgments.** This work was funded by the European Commission's 7th Framework Programme, under grant agreement 282672, EMBRACE project. We also thank the WCRP GEWEX project for support of the LandFlux initiative. The LandFlux-EVAL project (<http://www.iac.ethz.ch/url/research/LandFlux-EVAL>) was additionally supported by the Integrated Land Ecosystem-Atmosphere Process Study (ILEAPS). We acknowledge the WCRP's Working Group on Coupled Modelling and the climate modeling groups for producing and making available their model output. We also thank Jan Sedlacek, Urs Beyerle, and Thierry Corti (IAC

ETH and C2SM) for support with downloading, storage, and regridding of CMIP5 data. GPCP combined precipitation data were developed by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres as a contribution to the GEWEX Global Precipitation Climatology Project. We acknowledge the ECMWF for the dissemination of ERA-Interim data. CRU data were obtained from the University of East Anglia Climate Research Unit, British Atmospheric Data Centre, 2008, available from <http://badc.nerc.ac.uk/data/cru>. Willmott air temperature (University of Delaware) data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>.

[18] The Editor thanks Ray Anderson and an anonymous reviewer for their assistance in evaluating this paper.

#### References

- Adler, R. F., et al. (2003), The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present), *J. Hydrometeorol.*, 4(6), 1147–1167.
- Boberg, F., and J. H. Christensen (2012), Overestimation of Mediterranean summer temperature projections due to model deficiencies, *Nat. Clim. Change*, 2(6), 433–436, doi:10.1038/nclimate1454.
- Christensen, J. H., and F. Boberg (2012), Temperature dependent climate projection deficiencies in CMIP5 models, *Geophys. Res. Lett.*, 39, L24705, doi:10.1029/2012GL053650.
- Dee, D. P., et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, 137(656), 553–597, doi:10.1002/qj.828.
- Jimenez, C., et al. (2011), Global intercomparison of 12 land surface heat flux estimates, *J. Geophys. Res.*, 116, D02102, doi:10.1029/2010JD014545.
- Knutti, R., and J. Sedlacek (2013), Robustness and uncertainties in the new CMIP5 climate model projections, *Nat. Clim. Change*, 3, 369–373, doi:10.1038/nclimate1716.
- Koster, R. D., et al. (2006), GLACE: The global land-atmosphere coupling experiment. Part I: Overview, *J. Hydrometeorol.*, 7(4), 590–610, doi:10.1175/JHM510.1.
- Legates, D. R., and C. J. Willmott (1990), Mean seasonal and spatial variability in global surface air temperature, *Theor. Appl. Climatol.*, 41(1–2), 11–21, doi:10.1007/BF00866198.
- Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, 25(6), 693–712, doi:10.1002/joc.1181.
- Mueller, B., and S. I. Seneviratne (2012), Hot days induced by precipitation deficits at the global scale, *Proc. Natl. Acad. Sci. U.S.A.*, 109(31), 12,398–12,403, doi:10.1073/pnas.1204330109.
- Mueller, B., et al. (2011), Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophys. Res. Lett.*, 38, L06402, doi:10.1029/2010GL046230.
- Mueller, B., et al. (2013), Benchmark products for land evapotranspiration: LandFlux-EVAL multi-dataset synthesis, *Hydrol. Earth Syst. Sci. Discuss.*, 10, 769–805, doi:10.5194/hessd-10-769-2013.
- Nemani, R. R., C. D. Keeling, H. Hashimoto, W. M. Jolly, S. C. Piper, C. J. Tucker, R. B. Myneni, and S. W. Running (2003), Climate-driven increases in global terrestrial net primary production from 1982 to 1999, *Science*, 300(5625), 1560–1563, doi:10.1126/science.1082750.
- Seneviratne, S. I. (2012), Climate science: Historical drought trends revisited, *Nature*, 491(7424), 338–339.
- Seneviratne, S. I., et al. (2006), Soil moisture memory in AGCM simulations: Analysis of global land-atmosphere coupling experiment (GLACE) data, *J. Hydrometeorol.*, 7(5), 1090–1112.
- Seneviratne, S. I., et al. (2012), Changes in climate extremes and their impacts on the natural physical environment, in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, (IPCC SREX Report)*, edited by C. B. Field et al., pp. 109–230, Cambridge Univ. Press, Cambridge, U. K., and New York.
- 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., E. F. Wood, and M. L. Roderick (2012), Little change in global drought over the past 60 years, *Nature*, 491(7424), 435–438.
- Solomon, S., et al. (Eds.) (2007), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, 996 pp., Cambridge Univ. Press, Cambridge, U. K., and New York.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.

- Teuling, A. J., et al. (2009), A regional perspective on trends in continental evaporation, *Geophys. Res. Lett.*, *36*, L02404, doi:10.1029/2008GL036584.
- Willmott, C. J., and S. M. Robeson (1995), Climatologically aided interpolation (CAI) of terrestrial air temperature, *Int. J. Climatol.*, *15*(2), 221–229, doi:10.1002/joc.3370150207.