Climate model biases in seasonality of continental water storage revealed by satellite gravimetry

Sean Claude Swenson[1] and P. C. D. Milly[2]

Received 30 September 2005; revised 1 November 2005; accepted 17 January 2006; published 1 March 2006.

[1] Satellite gravimetric observations of monthly changes in continental water storage are compared with outputs from five climate models. All models qualitatively reproduce the global pattern of annual storage amplitude, and the seasonal cycle of global average storage is reproduced well, consistent with earlier studies. However, global average agreements mask systematic model biases in low latitudes. Seasonal extrema of low-latitude, hemispheric storage generally occur too early in the models, and model-specific errors in amplitude of the low-latitude annual variations are substantial. These errors are potentially explicable in terms of neglected or suboptimally parameterized water stores in the land models and precipitation biases in the climate models.

Citation: Swenson, S. C., and P. C. D. Milly (2006), Climate model biases in seasonality of continental water storage revealed by satellite gravimetry, *Water Resour. Res.*, *42*, W03201, doi:10.1029/2005WR004628.

1. Introduction

[2] Water storage is a major term in river basin water balances at subannual timescales, and interseasonal storage plays a critical role in the seasonal and annual partitioning of precipitation into runoff and evapotranspiration [*Milly*, 1994]. That partitioning, in turn, controls the moistening and sensible heating of the atmosphere, thereby exerting a major influence on climate [*Milly and Dunne*, 1994]. For this reason, it is desirable that numerical models of the climate system be evaluated with respect to their ability to reproduce the seasonal cycle of continental water storage.

[3] The intense spatial variability of terrestrial water storage prohibits comprehensive global-scale analysis based directly on measurements [*Robock et al.*, 1998]. Indirect estimates of storage from large-scale (10^5–10^6 km^2 basin area) differencing of river discharge and atmospheric water-vapor flux convergence provide useful information where atmospheric sounding networks are sufficiently dense [*Seneviratne et al.*, 2004]. However, such analyses must be seriously questioned when, as is common, long-term convergence fails to agree with surface discharge, and the disagreement is too large to explain by transient storage or subsurface discharge. Apparently, budget methods and measurements do not account adequately for all important spatial and temporal scales of flux variability.

[4] Changes in continental water storage are balanced by changes in the other global water reservoirs. If the problem of model evaluation is limited to global average storage, indirect estimates can be obtained by estimation of mass changes in the global ocean and atmosphere, typically with neglect of the changes in mass of the Greenland and Antarctic Ice Sheets. Such analyses have shown generally favorable agreement between models and observations [*Chen et al.*, 1998], but provide no information on the spatial structure of temporal variations of storage.

[5] The land models in global climate models all track water storage in snowpack and the subsurface. Many models at least implicitly track storage in the river network and some include storage in wetlands and/or lakes. Subsurface water is represented on a domain that typically extends 1–10 m downward from the surface. If infiltration of precipitation consistently exceeds evapotranspiration at some time of year, excess water will accumulate; during a dry season, storage will be depleted. If the annual mean balance is such that net accumulation tends to occur, then during the wet season water will leave the soil domain under the force of gravity. Such drainage may be parameterized as a downhill slope flow to the stream network or as a vertical flux that recharges a deeper groundwater reservoir, which in turn drains to the stream network. In either case, seasonal changes in drainage fluxes will lead to seasonal storage changes whose magnitude is controlled by the lateral resistance to flow. This representation is inadequate in arid regions, where the water table may be tens or hundreds of meters below the surface. In this study, however, the signals examined are dominated by humid regions, and this inadequacy is not expected to be a problem.

[6] The Gravity Recovery and Climate Experiment (GRACE) satellite mission was launched jointly by the National Aeronautics and Space Administration and the Deutsches Zentrum für Luft- und Raumfahrt in March 2002 [*Tapley et al.*, 2004]. The month-to-month gravity variations obtained from GRACE provide information about changes in the distribution of mass within the Earth and at its surface. In general, the largest time variable gravity signals are the result of changes in the distribution of water and snow stored on land [*Wahr et al.*, 1998]. The seasonal effects of human water management activities and seasonal biomass changes are both negligible compared to the effect of water storage. GRACE can thus provide global observations of changes in total water storage (vertically integrated water content), averaged over scales of a few

[1]Department of Physics and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.
[2]U.S. Geological Survey, Princeton, New Jersey, USA.

0043-1397/06/2005WR004628$09.00

Table 1. Models Evaluated in This Study and Water Stores Used[a]

Institution	Model	Water Store				
				Subsurface		
		Canopy	Snow	Root Zone	Groundwater	River
GFDL (USA)	GFDL-CM2.0	0	X	X	X	0
GFDL (USA)	GFDL-CM2.1	0	X	X	X	0
CCSR-NIES-FRCGC (Japan)	MIROC3.2(hires)	X	X	X[b]	X[b]	X
CCSR-NIES-FRCGC (Japan)	MIROC3.2(medres)	X	X	X[b]	X[b]	X
MRI (Japan)	MRI-CGCM2.3.2	NO	X	X[b]	X[b]	X

[a]"X" indicates presence of term in model and inclusion in our analysis; "0" indicates absence from model; and "NO" indicates presence in model but absence from our analysis. GFDL, Geophysical Fluid Dynamics Laboratory; CCSR-NIES-FRCGC, Center for Climate System Research (University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change; MRI, Meteorological Research Institute.

[b]In these models, root zone and groundwater stores are represented in a unified manner.

hundred km and greater [*Wahr et al.*, 2004; *Swenson et al.*, 2003].

[7] *Ramillien et al.* [2005], *Ellett et al.* [2005], and *Chen et al.* [2005] have made preliminary comparisons of GRACE water storage estimates with estimates from stand-alone land model simulations. The objective of this paper is to provide a quantitative evaluation of continental water storage changes computed in five climate models by comparison with storage changes estimated from GRACE observations. The comparison accounts for internal variability of the climate system and GRACE observational errors, so that the statistical significance of our results is apparent.

2. Data and Methods

[8] Twenty-one monthly GRACE gravity fields spanning from August 2002 to July 2004 were analyzed. The temporal coverage of this data set is not continuous; solutions are not available for December 2002, January 2003, or June 2003. Each gravity field is composed of a set of spherical harmonic coefficients complete to degree and order 70. Degree-one and degree-two zonal (C_20) coefficients are estimated independently via the Global Positioning System (GPS) and satellite laser ranging (SLR), respectively [*Cheng and Tapley*, 2003; *Chen et al.*, 2004; *Chambers et al.*, 2004; X. Wu et al., Seasonal and interannual global surface mass variations from geodetic data combinations, submitted to *Journal of Geophysical Research*, 2005]. After removal of the temporal mean and conversion of the gravity field anomalies to an equivalent water thickness, each monthly field was smoothed using a Gaussian filter with a half-width corresponding to 750 km [*Wahr et al.*, 1998]. Spatial averaging (smoothing) is required to reduce uncertainty in the water storage estimates due to errors in the short-wavelength gravity field coefficients [*Swenson and Wahr*, 2002]. After smoothing in the spectral domain, the data were transferred to a 1° by 1° global grid.

[9] To characterize the geographic distribution of temporal changes in storage, we used the phase and amplitude of the best fit sine wave having a period of one year, which we refer to as the annual cycle. To characterize the temporal changes of large-area averages of storage, we used the average monthly march of storage anomalies, obtained by averaging across the available years of data.

[10] Accordingly, from the 21-month GRACE time series at each grid point, the annual phase and amplitude were determined, and an average seasonal cycle was computed by averaging values for those months that had measurements in 2 years. From the resulting 12-month time series, spatial averages were computed for the global land area and for the following zonal land regions: Northern Hemisphere (NH) high latitudes (defined as 45°–90°N), NH middle latitudes (defined as 30°–45°N), NH low latitudes (defined as 0°–30°N), and Southern Hemisphere (SH) low latitudes (defined as 0°–30°S).

[11] For each regional water storage time series, standard errors were computed by the method of *Wahr et al.* [2004]. In brief, error fields for each month were estimated by scaling the formal error estimates provided by the GRACE Project to match the root-mean-square variability about the annual cycle in the GRACE gravity fields. From these error estimates, uncertainties for each spatial average time series were computed [*Swenson and Wahr*, 2002], representing one standard deviation from the expected monthly water storage.

[12] In preparation for Assessment Report 4 of the Intergovernmental Panel on Climate Change, 23 climate models recently were run with estimates of historical external forcing (i.e., atmospheric concentrations of radiatively active gases and aerosols) corresponding to the late 19th and full 20th centuries. When the experiments were executed, some of the modeled terrestrial water storage variables were not saved from some models. For those models that saved all of their storage variables, not all of those variables were reported publicly or otherwise easily accessible. We identified only five models for which we could obtain all significant continental water storage components that were represented in the models (Table 1). For the Meteorological Research Institute (MRI) model [*Yukimoto and Noda*, 2002], we did not obtain canopy storage, but the magnitude of this neglected term is very small. The Geophysical Fluid Dynamics Laboratory (GFDL) models [*Delworth et al.*, 2006] include neither canopy nor river stores and represent subsurface storage as separate root zone and groundwater stores. In contrast, the Model for Interdisciplinary Research on Climate (MIROC) [*K-1 Model Developers*, 2004] and MRI models have soil domains that are sufficiently deep to represent groundwater storage in the humid regions that dominate the signals in our analysis. The MRI model also tracks, and we included in our analysis, dynamic stores of water in the Caspian and Aral Seas and in Lakes Balkhash, Chad and

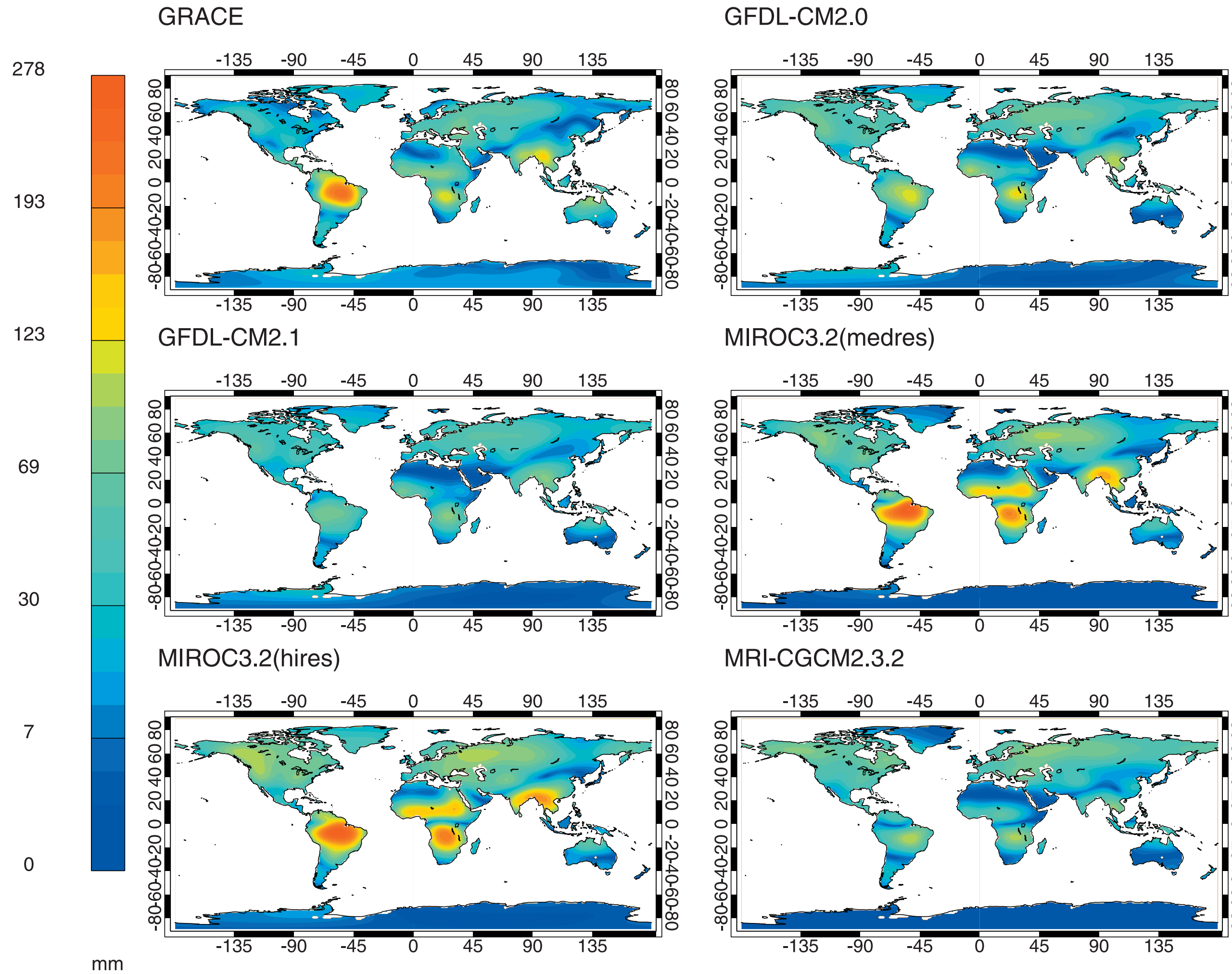

Figure 1. Global map of amplitude (mm) of annual cycle of land water storage from GRACE and from five climate models.

Eyre, but these contribute negligibly to the overall signal. We analyzed only one replicate of the experiment from each model, because temporal sampling errors were found to be small.

[13] The model data were processed in the same manner as the GRACE data. Each model monthly field was converted from its native grid to a set of spherical harmonic coefficients, smoothed in the spectral domain with a 750-km Gaussian filter, and placed on a 1° output grid. The annual amplitude and phase were computed for each grid point. Each (overlapping) 3-year period in the 1900–1999 model time series was mapped onto the time period 2002–2004 and sampled according to the 21-month schedule of GRACE data. From each of the 98 resulting samples and for each area of interest, the spatially averaged seasonal cycle was computed in exactly the same manner as for the GRACE data. Results were then processed to obtain, for each month of the year, the mean and standard deviation of storage from these 98 samples of the seasonal cycle.

[14] Differences between GRACE observations and model outputs differ for the following four reasons: (1) the models do not exactly reproduce the behavior of the real climate system; (2) the GRACE observations include measurement error; (3) the GRACE observations are a short-time sample of a climate system having substantial internal (that is, natural, unforced) variability; and (4) the temporal sampling windows for GRACE and the model outputs differ, while climate is changing. To evaluate the importance of reason 4, we repeated much of our analysis with model outputs only from 1980 to 1999 (instead of 1900–1999) and found results that were virtually identical. When we attempt herein to evaluate the importance of reason 1 by comparing GRACE observations to model outputs, we represent the effects of reasons 2 and 3 by use of the estimated GRACE errors and the model-based estimate of variability across the 98 3-year samples, both mentioned earlier. When we plot errors associated with reason 3, we attach the error bars to the model outputs, because they vary from one model to the next, even though they are actually characterizing temporal sampling error in GRACE observations.

3. Results

[15] The global distribution of the amplitude of the annual cycle of land water storage inferred from GRACE has regional maxima in both the low and high latitudes (Figure 1). In the low latitudes, maxima are present in northern South America, in two bands straddling the equa-

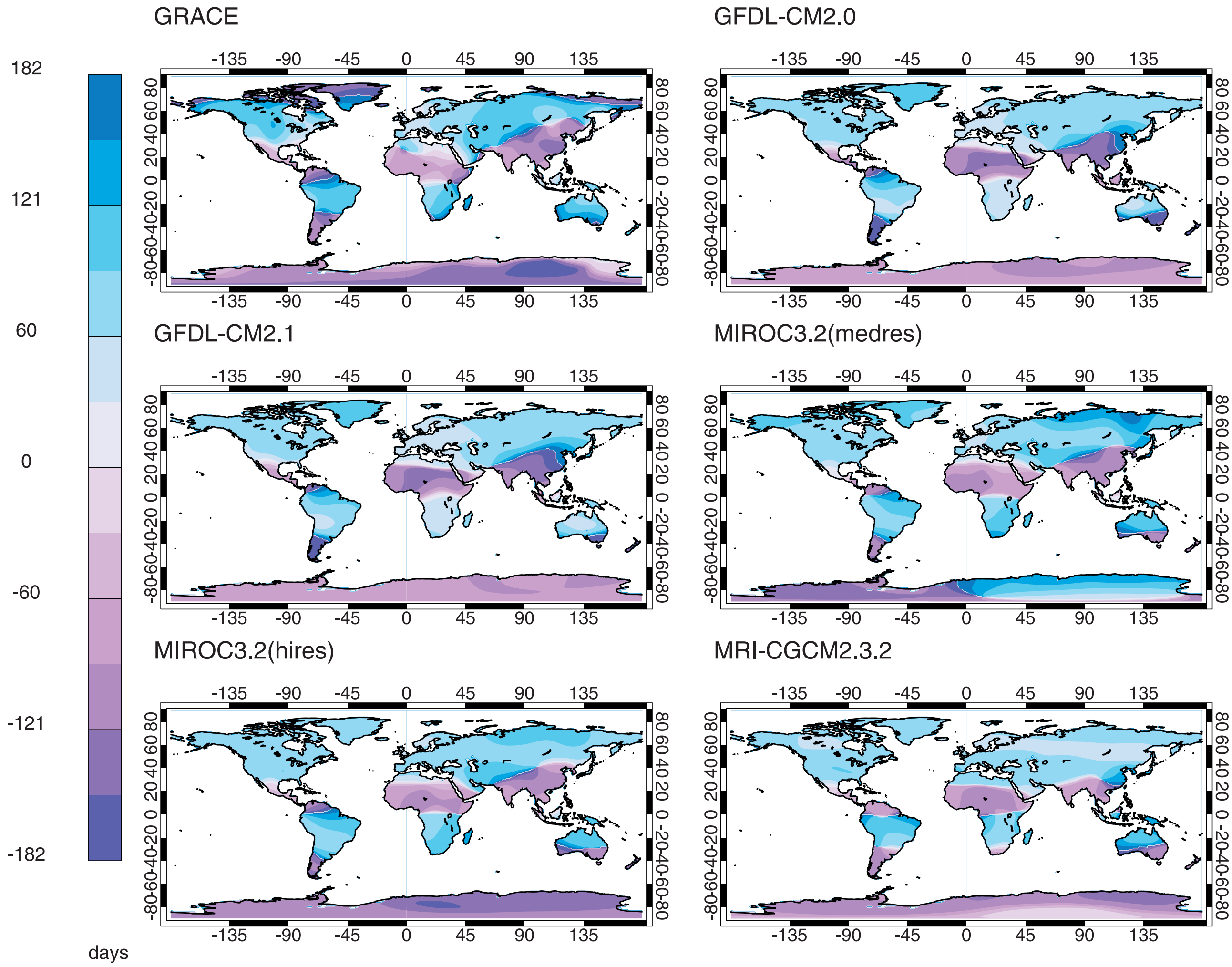

Figure 2. Global map of phase (time of maximum as day of year) of annual cycle of land water storage from GRACE and from five climate models.

tor in Africa, and in southern Asia and northern Australia. Each of these regions experiences a strong seasonal cycle of precipitation, and the phase of the annual storage signal (Figure 2) is consistent with the expectation that storage maxima should generally follow that of the intertropical convergence zone. In the high latitudes, the maximum amplitudes of the annual water storage signal inferred from GRACE are found in northwestern Eurasia and in northwestern North America. These correspond generally to areas of relatively high precipitation. The phasing of the annual storage signal is determined predominantly by the strong seasonal variation in the surface energy balance, which drives snowmelt and evapotranspiration. In arid regions, the range of storage inferred from GRACE is small.

[16] All five models qualitatively agree with GRACE on the spatial and temporal patterns of terrestrial water storage (Figures 1 and 2). However, the two GFDL models and the MRI model fail to generate the large storage amplitude inferred by GRACE for northern South America. In contrast, the two MIROC models succeed in generating a large amplitude there, but generate excessive amplitudes in the other tropical centers of storage. Thus all five models fail to reproduce the gradient in storage amplitude from South America to Africa and Southeast Asia.

[17] The annual cycle of global land water storage computed by each of the five models compares favorably with the GRACE data (Figure 3). The fitted maximum (19 mm) in the GRACE data occurs on day 99 of the year. In the two MIROC models, the maximum occurs within two days of this time, but it occurs about 15 days earlier in both GFDL models and 25 days earlier in the MRI model. The observed value of the maximum is matched by GFDL CM2.1 but is higher by 5–12 mm in the other four models.

[18] Examination of low-latitude results for separate hemispheres reveals large systematic biases in the models relative to GRACE (Figure 3). In all five models, the maximum storage occurs earlier, by 13–54 days, than in the GRACE data, with the GFDL models showing the earliest maxima. Compared to GRACE, the GFDL models and the MRI model have maxima that are too small (by 9–48%) and the MIROC models have maxima that are too large by 45–94%.

[19] As a summary error index, we use the root-mean-square value of the difference between monthly model and GRACE storage values, normalized by the root-mean-square value of the GRACE monthly storage signal (Figure 4). By this measure, the high-latitude storage signal is best reproduced by the two GFDL models and the

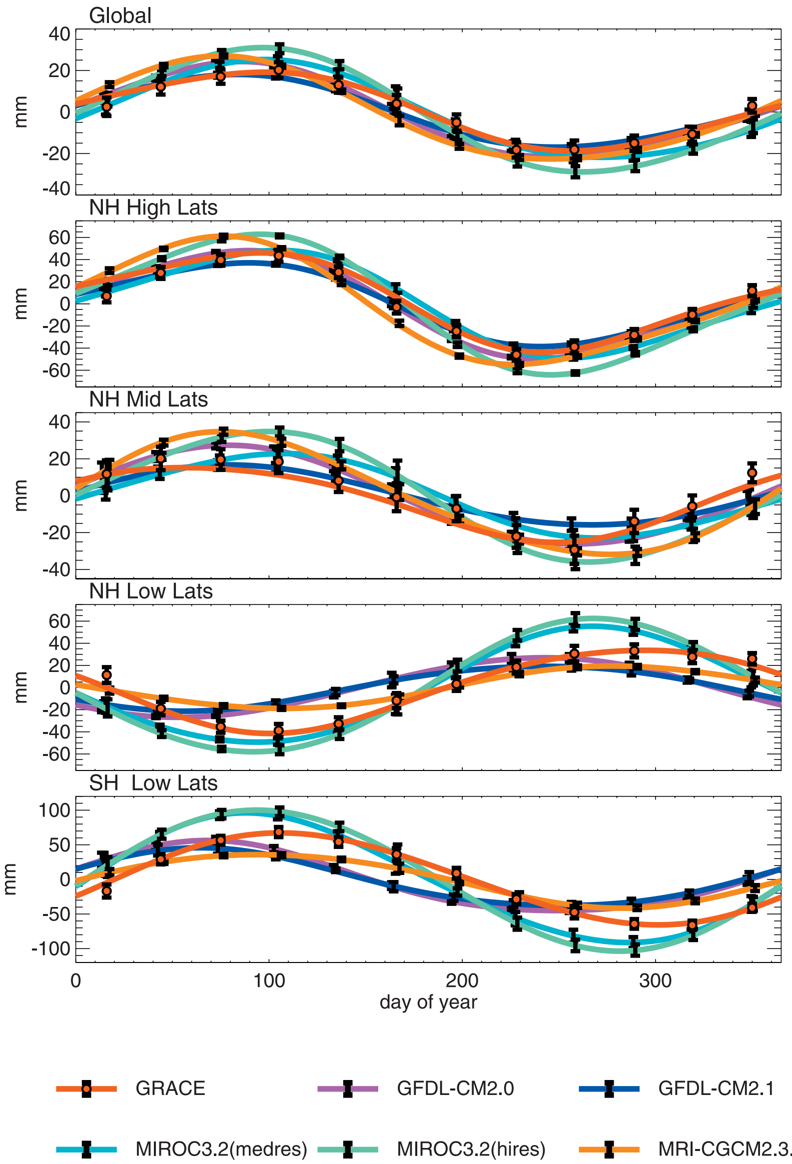

Figure 3. Monthly time series of land water storage anomaly (mm) estimated from GRACE (red) and by climate models for global land and various latitude zones. Continuous curves are based on fit of annual and semiannual cycles. For GRACE, each symbol represents one or two analyzed months of data, and error bars are estimated observational error (±1 standard error). For each model, the error bars (horizontally offset slightly for visibility) represent ±1 standard deviation of the model-derived estimate of the temporal sampling error in the GRACE measurements. NH/SH denotes Northern/Southern Hemisphere.

MIROC medium-resolution model. In the low latitudes, the MRI model generates the smallest errors. No model is "best overall."

4. Discussion

[20] From a land-modeling perspective, one can decompose errors in modeled terrestrial water storage into those that are caused by errors in the model climate and those that are caused by errors in the land model formulation, including its parameter values. In reality, of course, the errors of the second type may feed back to cause errors of the first type. A comprehensive model-by-model analysis of either type of error is far beyond the scope of this report. Instead, the objective of this discussion is to suggest a small number of potentially fruitful areas for further exploration.

[21] Differences in storage errors among models may be partially associated with differences in precipitation timing and magnitude. An examination of the average monthly march of zonal precipitation in the tropics revealed very similar phasing in all the models, with extrema very near the time of the solstices. However, in the MRI and the GFDL CM2.1 models, the seasonal range of precipitation rate was smaller by about 30% than in the other models. This relative deficiency in amplitude of precipitation seasonality would tend to make the storage range smaller for these two models than for the other models.

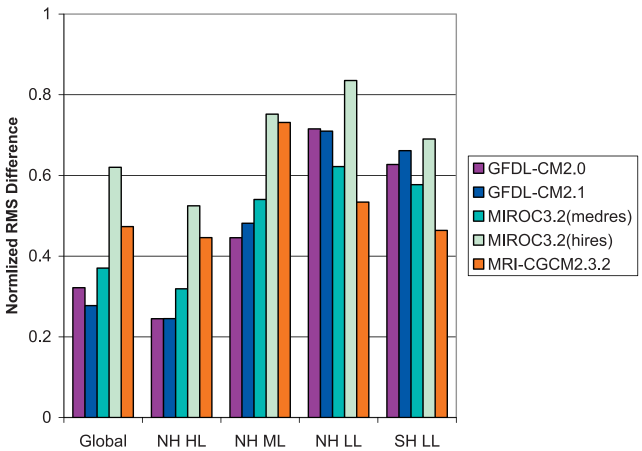

Figure 4. Normalized root-mean-square difference between GRACE and model estimates of zonal mean monthly water storage for each model. Differences of storage are normalized by the root-mean-square value of the GRACE anomaly. NH/SH denotes Northern/Southern Hemisphere; HL/ML/LL denotes high/medium/low latitudes.

[22] The absence of river storage in the GFDL models would tend to advance the phase and reduce the annual amplitude of storage in those models. The very rapid routing of river water in the MRI model [*Yukimoto et al.*, 2001] would have the same effect.

[23] The MIROC models are the only models in this study with significant river storage. These models used an effective river velocity parameter of 0.3 m s^{-1} [*K-1 Model Developers*, 2004], which is close to the one that *Miller et al.* [1994] estimated by global optimization of river discharge simulations of another model. That other model employed a shallow (i.e., root zone) bucket model for the soil, with no additional groundwater reservoir. Accordingly, groundwater storage processes were presumably aliased into the river velocity value, as noted by *Miller et al.* [1994]. When the same river parameterization is used in conjunction with a more advanced and deeper domain soil model, it is likely that some redundancy of groundwater storage occurs, and storage may be overestimated. We speculate that this may be a factor in the excessive storage range of the MIROC models relative to the GRACE measurements. However, this does not explain why the MIROC models would produce maximum storage earlier than the GRACE measurements.

[24] None of the models considered here has explicit representation of storage associated with seasonal inundation of river valleys. This deficiency would tend to make the storage amplitude too small and the phase too early in models. *Chapelon et al.* [2002] have shown, on the basis of comparisons between modeled and observed streamflow, that inundation may be an important unaddressed process in land models. In a typical year, the stage of the Amazon main stem varies seasonally over a range (Δh) of 7–10 m, causing inundation at maximum flood stage (in June and July) of an area (A) of 47,000 km^2 [*Hamilton et al.*, 2002]. Assuming a linear relation between flooded area and stage, we infer that maximum flood volume on the main stem is (AΔh)/2, or about 2×10^{11} m^2. Spread over the Amazon drainage basin area of approximately 5×10^6 km^2, this is equivalent to a storage range of 40 mm water depth; part of this is channel storage, and part is overbank storage. Inundation along tributaries would increase this value.

[25] We also note that the neglect of seasonal inundation in climate models might help to explain their inability to reproduce the contrast in storage amplitude between South America and other tropical regions. Another possible explanation is the tendency for climate models to fail to generate the observed contrast in precipitation and runoff [*Milly et al.*, 2005] between South America and other tropical regions. Of course, these two potential explanations may be linked through land-atmosphere feedbacks.

5. Summary and Conclusions

[26] We have used water storage changes inferred by satellite gravimetry (GRACE) to evaluate the quality of water storage simulations by five climate models. The seasonal cycle of global land water storage was reproduced well by all models, and this finding is consistent with earlier analyses (of other models) wherein water storage was deduced from global mean sea level variations. However, the use of satellite gravimetry allowed us to assess model performance at regional scale. In so doing, we showed that the observed high-latitude signal is reasonably reproduced by the models, but that low-latitude storage simulations contain large errors. Because of the shifted phasing of Southern and Northern Hemisphere processes, these errors are undetectable in the global mean.

[27] In general, we find that the annual maximum flow occurs too early in the models in low latitudes and that substantial model-dependent bias is present in the amplitude of the annual cycle. We speculatively identify several factors that could explain these errors. These include errors in climate model precipitation amounts and neglect and/or suboptimal parameterization of river water storage and seasonal inundation.

[28] **Acknowledgments.** We gratefully acknowledge the advice and technical assistance of Krista A. Dunne (U.S. Geological Survey), Seita Emori (National Institute for Environmental Studies, Japan), and Masahiro Hosaka and Seiji Yukimoto (Meteorological Research Institute, Japan) and the thoughtful reviews of Jianli Chen (University of Texas), Kevin Ellett (U.S. Geological Survey), David Lawrence (National Center for Atmospheric Research), and John Wahr (University of Colorado). We also acknowledge the institutions listed in Table 1 for providing model output; the Program for Climate Model Diagnosis and Intercomparison for collecting and archiving the model output; the World Climate Research Programme's Joint Scientific Committee/Climate Variability and Predictability Project Working Group on Coupled Modeling and its Coupled Model Intercomparison Project and Climate Simulation Panel for organizing analysis of model output; and the Intergovernmental Panel on Climate Change (IPCC) Working Group 1 Technical Support Unit. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy. This investigation was carried out within the IPCC Global Hydroclimates Analysis Project. The work was supported by NASA (NNG04GF02G) and NSF (EAR-0309678).

References

Chambers, D. P., J. Wahr, and R. S. Nerem (2004), Preliminary observations of global ocean mass observations with GRACE, *Geophys. Res. Lett.*, *31*, L13310, doi:10.1029/2004GL020461.

Chapelon, N., H. Douville, P. Kosuth, and T. Oki (2002), Off-line simulation of the Amazon water balance: A sensitivity study with implications for GSWP, *Clim. Dyn.*, *19*, 141–154.

Chen, J. L., C. R. Wilson, D. P. Chambers, R. S. Nerem, and B. D. Tapley (1998), Seasonal global water mass budget and mean sea level variations, *Geophys. Res. Lett.*, *25*, 3555–3558.

Chen, J. L., C. R. Wilson, B. D. Tapley, and J. Ries (2004), Low degree gravitational changes from GRACE: Validation and interpretation, *Geophys. Res. Lett.*, *31*, L22607, doi:10.1029/2004GL021670.

Chen, J. L., M. Rodell, C. R. Wilson, and J. S. Famiglietti (2005), Low degree spherical harmonic influences on Gravity Recovery and Climate Experiment (GRACE) water storage estimates, *Geophys. Res. Lett.*, *32*, L14405, doi:10.1029/2005GL022964.

Cheng, M. K., and B. D. Tapley (2003), Temporal variations in the Earth's gravity field from SLR and CHAMP GPS data, in *Gravity and Geoid 2002*, edited by I. N. Tziavos, pp. 424–431, Ziti, Thessaloniki, Greece.

Delworth, T. D., et al. (2006), GFDL's CM2 global coupled climate models—Part 1: Formulation and simulation characteristics, *J. Clim.*, in press.

Ellett, K. M., J. P. Walker, M. Rodell, J. Chen, and A. W. Western (2005), GRACE gravity fields as a new measure for assessing large-scale hydrological models, in *MODSIM 2005 International Congress on Modelling and Simulation*, edited by A. Zerger and R. M. Argent, pp. 2911–2917, Modell. and Simul. Soc. of Aust. and N. Z., Canberra.

Hamilton, S. K., S. J. Sippel, and J. M. Melack (2002), Comparison of inundation patterns among major South American floodplains, *J. Geophys. Res.*, *107*(D20), 8038, doi:10.1029/2000JD000306.

K-1 Model Developers (2004), K-1 coupled GCM (MIROC) description, edited by H. Hasumi and S. Emori, *K-1 Tech. Rep. 1*, Cent. for Clim. Syst. Res., Univ. of Tokyo, Tokyo. (Available at http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/tech-repo.pdf)

Miller, J. R., G. L. Russell, and G. Caliri (1994), Continental-scale river flow in climate models, *J. Clim.*, *7*, 914–928.

Milly, P. C. D. (1994), Climate, soil water storage, and the average annual water balance, *Water Resour. Res.*, *30*, 2143–2156.

Milly, P. C. D., and K. A. Dunne (1994), Sensitivity of the global water cycle to the water-holding capacity of land, *J. Clim.*, *7*, 506–526.

Milly, P. C. D., K. A. Dunne, and A. V. Vecchia (2005), Global pattern of trends in streamflow and water availability in a changing climate, *Nature*, *438*, 347–350.

Ramillien, G., F. Frappart, A. Cazenave, and A. Günter (2005), Time variations of land water storage from an inversion of 2 years of GRACE geoids, *Earth Planet. Sci. Lett.*, *235*, 283–301.

Robock, A. C., C. A. Schlosser, K. Y. Vinnikov, N. A. Speranskaya, and J. K. Entin (1998), Evaluation of AMIP soil moisture simulations, *Global Planet. Change*, *19*, 181–208.

Seneviratne, S. I., P. Viterbo, D. Lüthi, and C. Schär (2004), Inferring changes in terrestrial water storage using ERA-40 reanalysis data: The Mississippi River basin, *J. Clim.*, *17*, 2039–2057.

Swenson, S., and J. Wahr (2002), Methods for inferring regional surface-mass anomalies from GRACE measurements of time-variable gravity, *J. Geophys. Res.*, *107*(B9), 2193, doi:10.1029/2001JB000576.

Swenson, S., J. Wahr, and P. C. D. Milly (2003), Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *39*(8), 1223, doi:10.1029/2002WR001808.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004), The Gravity Recovery and Climate Experiment: Mission overview and early results, *Geophys. Res. Lett.*, *31*, L09607, doi:10.1029/2004GL019920.

Wahr, J., M. Molenaar, and F. Bryan (1998), Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res.*, *103*(B12), 30,205–30,229.

Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna (2004), Time-variable gravity from GRACE: First results, *Geophys. Res. Lett.*, *31*, L11501, doi:10.1029/2004GL019779.

Yukimoto, S., and A. Noda (2002), Improvements of the Meteorological Research Institute Global Ocean-Atmosphere Coupled GCM (MRI-CGCM2) and its climate sensitivity, *CGER's Supercomput. Activ. Rep. 10*, pp. 37–44, Natl. Inst. for Environ. Stud., Tsukuba, Japan.

Yukimoto, S., A. Noda, A. Kitoh, M. Sugi, Y. Kitamura, M. Hosaka, K. Shibata, S. Maeda, and T. Uchiyama (2001), The new Meteorological Research Institute coupled GCM (MRI-CGCM2): Model climate and variability, *Pap. Meteorol. Geophys.*, *51*, 47–88.

P. C. D. Milly, U.S. Geological Survey, P.O. Box 308, Route 1, Princeton, NJ 08542, USA. (cmilly@usgs.gov)

S. C. Swenson, Department of Physics, University of Colorado, Boulder, CO 80309, USA.