Published version available at: https://rdcu.be/5rec
DOI: 10.1038/s41586-018-0424-4

Title: Sensitivity of atmospheric CO_2 growth rate to observed changes in terrestrial water storage

Authors: Vincent Humphrey[1]*, Jakob Zscheischler[1], Philippe Ciais[2], Lukas Gudmundsson[1], Stephen Sitch[3], Sonia I. Seneviratne[1]*

Affiliations:

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Universitaetstrasse 16, 8092 Zurich, Switzerland.

[2]Laboratoire des Sciences du Climat et de l'Environnement CEA CNRS UVSQ, 91191, Gif-sur-Yvette, France.

[3]College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QF, United Kingdom.

*Correspondence to: vincent.humphrey@env.ethz.ch, sonia.seneviratne@ethz.ch.

Land ecosystems absorb on average 30% of the anthropogenic CO_2 emissions, thereby tempering the growth of the CO_2 concentration in the atmosphere[1]. Year-to-year variations in the atmospheric CO_2 growth rate are mostly due to fluctuating carbon uptake by land ecosystems[1]. While the sensitivity of these fluctuations to changes in tropical temperature has been well documented[2-6], identifying the role of global water availability has proven more elusive. To date, only time-lagged precipitation anomalies and drought indices have served as proxies for water availability[3-5] due to a lack of direct observations. Here, we use recent observations of terrestrial water storage changes derived from satellite gravimetry[7] to investigate land water effects on carbon cycle variability at global to regional scales. We show that the CO_2 growth rate is strongly sensitive to observed changes in terrestrial water storage, drier years being associated with faster atmospheric CO_2 growth. We demonstrate that this global relationship is independent from known temperature effects and is underestimated in current carbon cycle models. Our results indicate that inter-annual fluctuations in terrestrial water storage strongly impact the land carbon sink and highlight the important role of interactions between the water and carbon cycles.

Acquiring accurate estimates of the land carbon sink is a key requirement for monitoring global CO_2 emissions on a year-to-year basis[8] and for reducing significant uncertainties in projections of future carbon cycle-climate feedbacks[9,10]. One critical aspect is to understand the sensitivity of the CO_2 growth rate (CGR) to natural climate variability. At the global scale, it was found that the inter-annual variability (IAV) of the CGR is coupled with the El Niño Southern

Oscillation (ENSO) and more specifically with variations in mean tropical temperature[3,4,6,11]. In addition, the role of water availability has been widely documented at the regional scale. Major droughts have been shown to cause drastic regional reductions in the land carbon sink[12,13] and photosynthesis is limited by water scarcity over most of the globe[14]. Previous attempts to quantify the response of CGR IAV to water scarcity have used proxies to represent the amount of water available to ecosystems, such as yearly means of precipitation anomalies[6], time-lagged and low-pass filtered monthly precipitation[3,5] or standardized drought indices[4]. Although convenient, these proxies are limited since they only consider water inputs and either omit or model water losses due to evapotranspiration and runoff. From a process perspective, plants and micro-organisms respond however to the amount of water stored on land rather than to precipitation fluxes (Extended Data Fig. 1). Here, we overcome these limitations by using direct satellite observations of terrestrial water storage (TWS) anomalies to investigate links between the carbon and water cycles.

From 2002 to 2017, the twin satellites of the Gravity Recovery and Climate Experiment (GRACE) have measured monthly anomalies of the Earth's gravity field[7] that can be used to retrieve net changes in TWS including groundwater, soil moisture, surface waters, snow and water stored in the biosphere (also see Methods). We isolate the monthly TWS IAV from GRACE by subtracting the mean seasonal cycle, and remove the long-term trend using linear regression. Measurements of atmospheric CGR IAV from the National Oceanic and Atmospheric Administration (NOAA) are compared with the satellite-based TWS IAV over the overlapping period, revealing a significant negative correlation at both monthly ($r = -0.65$, $n=158$) and yearly ($r = -0.85$, $n=15$) scales (Fig. 1a,b). The sign of this relationship indicates that drier years, characterized by a negative anomaly in TWS, are associated with higher rates of atmospheric CO_2 growth and therefore a weakening of the land carbon sink (Fig. 1b). Composite TWS maps associated with high (Fig. 1c) and low (Fig. 1d) monthly CGR primarily reflect winter-spring water storage anomalies in South America and tropical regions in general. Given the relatively short observational record provided by the GRACE satellites, we investigate the robustness of this coupling and the associated spatial patterns using alternative estimates of TWS, which offer longer temporal coverage (Methods, Supplementary Information Tables 1 & 3). Although comparable, these estimates are based on model simulations that are considered less reliable than the actual GRACE observations. First, we use a statistical model of climate-driven water storage variability that is trained with GRACE observations (GRACE-REC)[15] (Fig. 1a). We exclude the years following the eruption of Mt. Pinatubo (1991-1993)[16-18] (Methods), and find a significant negative coupling between

GRACE-REC TWS and CGR at both monthly (r = -0.59, n=408) and yearly (r = -0.61, n=34) scales over the period 1980-2016 (Fig. 2a-b). Using TWS simulated by process-based land surface models, we tend to find lower correlations as we move from global hydrological models, which usually have the most complete or well calibrated representation of water reservoirs (WaterGAP[19]), to land surface models (GLDAS2-Noah[20]) or Dynamic Global Vegetation Models (DGVMs), which often only consider root-zone soil moisture (TRENDY ensemble, version 3[9]). Nevertheless, these model estimates confirm the existence of a coupling between water storage and observed CGR. Unlike precipitation anomalies, water storage changes integrate the history of variations in both water supply and water demand over time. Therefore, looking at precipitation alone (with an optimal 4 months lag[5]) underestimates the strength of the coupling between water storage and carbon fluxes, in particular at the monthly scale (Fig. 2a,c). The strength of the link between CGR and water storage is comparable to that of the link between ENSO and CGR with a lag of about 4 months (ENSO leading CGR[5]). ENSO is a key mode of variability in global atmospheric circulation and is associated with large-scale fluctuations in precipitation patterns, which ultimately translate into water storage anomalies[6,21] (Extended Data Fig. 3).

As documented in previous studies[3,5], the correlation between CGR and temperature is more pronounced in the tropical domain and at yearly time scale (Fig. 2c-d). Individual effects of temperature and water storage on CGR may be difficult to disentangle because these two drivers co-vary. Warmer years generally coincide with drier years (Fig. 3a), raising the question of whether the TWS signal might implicitly contain some response to temperature. However, our results show that GRACE TWS can be almost entirely reconstructed from precipitation anomalies alone (Extended Data Fig. 4) with only little impact from temperature variability. Partial correlations indicate that the global CGR-TWS relationship remains significant after controlling for the effect of either global or tropical temperature (partial correlations of -0.72) (Fig. 3b, blue bars). This means that most of the information on CGR variations that is contained in TWS cannot be found in temperature. On the opposite, controlling for the effect of TWS strongly decreases partial correlations between CGR and temperature (Fig. 3b, orange bars). Using univariate linear regression (Methods), we find a global yearly sensitivity of -1.33 (95% confidence interval spanning from -1.85 to -1.07) Gt of carbon per year for each additional Tt of water stored on land (Fig. 3c). This corresponds to a ratio of roughly 1.3 g C yr^{-1} kg^{-1} H_2O. When including both TWS and temperature in a bivariate regression, the sensitivity to TWS is reduced to -0.93 (-1.50 to -0.48) Gt C yr^{-1} Tt^{-1} H_2O (28% decrease). For temperature, the univariate sensitivity is 3.89 (2.44 to 5.16) Gt C yr^{-1} $°C^{-1}$ and is largely reduced in the bivariate

case to 1.99 (0.66 to 3.59) Gt C yr^{-1} $°C^{-1}$ (49% decrease), which is much lower than previous estimates[2,4] (Fig. 3d).

Our findings provide strong observational evidence that the CGR is coupled to changes in both temperature and water storage at the global scale. The role of water storage is also stronger than what can be diagnosed from precipitation[5] (Extended Data Fig. 5) or precipitation conditional on ENSO phase[6]. However, these findings differ from recent results of Jung and colleagues[22], who suggested that the global mean net ecosystem exchange (NEE) simulated by statistical models (FluxCom[23]) and physical carbon cycle models (DGVMs[9]) responds to temperature rather than to water storage. In order to investigate this discrepancy, we reproduce the approach of Jung and colleagues (Methods) and find that, while our observations indicate that CGR is highly correlated to global water storage changes (Fig. 4a, circle), modelled NEE fails to reproduce this pattern and is instead mostly correlated to temperature (Fig. 4a, squares). Here, we suggest that this occurs because models underestimate the magnitude of water-driven NEE variations at the global scale (Fig. 4b,c). We find that the water-driven NEE of a given model is (except for one model) directly correlated to its simulated global mean water storage (Fig. 4d, Supplementary Information Fig. 1). This internal model relationship indicates that a link exists between global mean water storage and its resulting global effect on NEE, which directly supports our observation-based results. We note that this global relationship also holds for the temperature-driven response (Fig. 4d). Therefore, the correlations reported in Fig. 4a for total model NEE are directly controlled by the relative importance of the temperature-driven and water-driven NEE components (Fig. 4b,c). Our observations (Fig. 4a, circle) thus suggest that simulated global NEE may appear dominated by temperature effects because the amplitude of water-driven NEE is underestimated. This might indicate that the modelled NEE response is not sensitive enough to soil moisture or point towards the role of non-modelled processes that are strongly regulated by other types of water storage changes (e.g. the access of deep roots to groundwater[24] or the response of inland waters and wetland ecosystems[25]). In addition, inaccuracies in the precipitation forcing as well as missing water reservoirs in model hydrology might also affect water-driven NEE signals. Compared to GRACE observations, models display a widespread tendency towards underestimating the importance of low-frequency (inter-annual) water storage anomalies and are dominated by short-term fluctuations (Methods, Extended Data Fig. 6-8). This is likely explained by the limited (or absent) representation of deep soil layers, groundwater, wetlands and surface waters which respond more slowly to climate forcing and have a much longer residence time than root-zone soil moisture. Interestingly, we find that the fraction of IAV (Methods) in modelled water storage imposes a

strong upper limit on how much IAV can ultimately be found in modelled water-driven NEE (Extended Data Fig. 9). As a result, the amplitude of water-driven NEE at the inter-annual time scale (Fig. 4b) is limited by a lack of long-term memory in the underlying water storage signal. By partitioning the water storage signal among six land cover classes (Supplementary Information Fig. 2), we find that GRACE observations and models agree that semi-arid regions dominate the global mean water storage signal (Extended Data Fig. 10), even though models do not correlate very well with the actual signal observed by GRACE (Supplementary Information Fig. 3). These findings support recent results suggesting that semi-arid (and thus water-limited) ecosystems are responsible for most of the CGR IAV[26,27]. However, while we find that GRACE water storage in semi-arid regions is well correlated with CGR, our analysis also suggests a possible role of tropical forests (Fig.1c-d, Supplementary Information Fig. 4).

In summary, we have provided for the first time observational evidence that the inter-annual variability of the CO_2 growth rate is tightly coupled to terrestrial water storage changes. The sensitivities derived here represent the aggregated response of processes that operate at smaller spatial scales[22]. For this reason, they are not directly transferable to the ecosystem scale but may still provide a valuable metric for evaluating and constraining Earth system models[2,6,10]. Our results suggest that current models might underestimate the response of ecosystems to global changes in water availability. Models typically only respond to shallow soil moisture and are therefore less sensitive to inter-annual variability in water storage. They might also miss the response to changes in non-modelled water reservoirs such as wetlands or surface waters. The presented findings offer new perspectives on the use of satellite observations of water storage for global carbon cycle research. Projections of inter-annual as well as long-term water storage changes from hydrological models still display large uncertainties[28,29], and will need to be better assessed in order to reduce uncertainties in projections of future land carbon uptake. As an additional complexity, estimates of future terrestrial water storage are themselves very dependent on how transpiration will be regulated by vegetation in a world of rising CO_2 concentrations[30]. Such evidence of the interplay between the water and carbon cycles also highlights the need for stronger interactions between the hydrological and biogeochemical research communities.

References:

1 Le Quéré, C. et al. Global Carbon Budget 2017. Earth Syst Sci Data 10, 405-448, doi:10.5194/essd-10-405-2018 (2018).

2 Cox, P. M. *et al.* Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature* **494**, 341-344, doi:10.1038/nature11882 (2013).

3 Wang, W. *et al.* Variations in atmospheric CO2 growth rates coupled with tropical temperature. *Proceedings of the National Academy of Sciences* **110**, 13061-13066, doi:10.1073/pnas.1219683110 (2013).
4 Wang, X. *et al.* A two-fold increase of carbon cycle sensitivity to tropical temperature variations. *Nature* **506**, 212-215, doi:10.1038/nature12915 (2014).
5 Wang, J., Zeng, N. & Wang, M. R. Interannual variability of the atmospheric CO2 growth rate: roles of precipitation and temperature. *Biogeosciences* **13**, 2339-2352, doi:10.5194/bg-13-2339-2016 (2016).
6 Fang, Y. *et al.* Global land carbon sink response to temperature and precipitation varies with ENSO phase. *Environ Res Lett* **12**, 064007, doi:10.1088/1748-9326/aa6e8e (2017).
7 Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F. & Watkins, M. M. GRACE measurements of mass variability in the Earth System. *Science* **305**, 503-505, doi:10.1126/science.1099192 (2004).
8 Peters, G. P. *et al.* Towards real-time verification of CO2 emissions. *Nature Climate Change* **7**, 848-850, doi:10.1038/s41558-017-0013-9 (2017).
9 Sitch, S. *et al.* Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* **12**, 653-679, doi:10.5194/bg-12-653-2015 (2015).
10 Friedlingstein, P. *et al.* Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. *J Climate* **27**, 511-526, doi:10.1175/JCLI-D-12-00579.1 (2014).
11 Keeling, C. D., Whorf, T. P., Wahlen, M. & Vanderplicht, J. Interannual Extremes in the Rate of Rise of Atmospheric Carbon-Dioxide since 1980. *Nature* **375**, 666-670, doi:10.1038/375666a0 (1995).
12 Ciais, P. *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* **437**, 529-533, doi:10.1038/nature03972 (2005).
13 Phillips, O. L. *et al.* Drought sensitivity of the Amazon rainforest. *Science* **323**, 1344-1347, doi:10.1126/science.1164033 (2009).
14 Beer, C. *et al.* Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. *Science* **329**, 834-838, doi:10.1126/science.1184984 (2010).
15 Humphrey, V., Gudmundsson, L. & Seneviratne, S. I. A global reconstruction of climate-driven subdecadal water storage variability. *Geophys Res Lett* **44**, 2300-2309, doi:10.1002/2017GL072564 (2017).
16 Lucht, W. *et al.* Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science* **296**, 1687-1689, doi:10.1126/science.1071828 (2002).
17 Trenberth, K. E. & Dai, A. Effects of Mount Pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering. *Geophys Res Lett* **34**, doi:10.1029/2007GL030524 (2007).
18 Mercado, L. M. *et al.* Impact of changes in diffuse radiation on the global land carbon sink. *Nature* **458**, 1014-1017, doi:10.1038/nature07949 (2009).
19 Döll, P., Müller Schmied, H., Schuh, C., Portmann, F. T. & Eicker, A. Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resour Res* **50**, 5698-5720, doi:10.1002/2014WR015595 (2014).
20 Rodell, M. *et al.* The global land data assimilation system. *B Am Meteorol Soc* **85**, 381-394, doi:10.1175/BAMS-85-3-381 (2004).
21 Ni, S. *et al.* Global Terrestrial Water Storage Changes and Connections to ENSO Events. *Surv Geophys* **39**, 1-22, doi:10.1007/s10712-017-9421-7 (2017).
22 Jung, M. *et al.* Compensatory water effects link yearly global land CO2 sink changes to temperature. *Nature* **541**, 516-520, doi:10.1038/nature20780 (2017).
23 Tramontana, G. *et al.* Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences* **13**, 4291-4313, doi:10.5194/bg-13-4291-2016 (2016).
24 Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B. & Otero-Casal, C. Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences* **114**, 10572-10577, doi:10.1073/pnas.1712381114 (2017).
25 Battin, T. J. *et al.* The boundless carbon cycle. *Nat Geosci* **2**, 598-600, doi:10.1038/ngeo618 (2009).
26 Poulter, B. *et al.* Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* **509**, 600-603, doi:10.1038/nature13376 (2014).
27 Ahlstrom, A. *et al.* The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink. *Science* **348**, 895-899, doi:10.1126/science.aaa1668 (2015).
28 Orlowsky, B. & Seneviratne, S. I. Elusive drought: uncertainty in observed trends and short- and long-term CMIP5 projections. *Hydrol Earth Syst Sc* **17**, 1765-1781, doi:10.5194/hess-17-1765-2013 (2013).
29 Scanlon, B. R. *et al.* Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences* **115**, E1080-E1089, doi:10.1073/pnas.1704665115 (2018).

30 Swann, A. L. S., Hoffman, F. M., Koven, C. D. & Randerson, J. T. Plant responses to increasing CO2 reduce estimates of climate impacts on drought severity. *Proceedings of the National Academy of Sciences* **113**, 10019-10024, doi:10.1073/pnas.1604581113 (2016).

Supplementary Information:

Supplementary Information Tables 1-3

Supplementary Information Figures 1-10

Acknowledgments: All datasets supporting the results of this paper are openly accessible from the references listed in Supplementary Information Table 1. This research was funded by the European Research Council DROUGHT-HEAT project (contract 617518). P. C. was supported by the European Research Council Synergy grant ERC-2013-SyG-610028 IMBALANCE-P. We thank Martin Jung and Ulrich Weber for providing the water availability index used in FluxCom, Richard Wartenburger for technical support. We gratefully thank the following data providers and model developers for their continuous efforts and for sharing their data: NASA Jet Propulsion Laboratory, NOAA Earth System Research Laboratory, the Global Carbon Project, WaterGAP Global Hydrology Model (WGHM), the Global Land Data Assimilation System (GLDAS), Multi-Source Weighted-Ensemble Precipitation (MSWEP), Global Precipitation Climatology Project (GPCP), University of East Anglia Climatic Research Unit (CRU), Berkeley Earth, and all contributors as well as data providers to the FluxCom initiative and the TRENDY experiment version 3, which included the models CABLE, CLM, ISAM, JSBACH, JULES, LPJ, LJP-GUESS, LPX-Bern, ORCHIDEE, VEGAS and VISIT.

Author contributions: V.H., S.I.S., J.Z., and P.C. designed the study. V.H. conducted the data analysis with support from L.G., J.Z., S.S. and S.I.S., and wrote the manuscript. The interpretation, final text and figures resulted from the contributions of all co-authors.

Author information: Reprints and permissions information is available at www.nature.com/reprints. The authors declare that they have no competing interest. Correspondence and requests for materials should be addressed to V.H. (vincent.humphrey@env.ethz.ch) and S.I.S (sonia.seneviratne@ethz.ch).

Figure 1. Inter-annual variability in CGR and TWS. (a) Monthly de-seasonalised and de-trended CGR, TWS from satellite observations (GRACE) and TWS from a statistical model (GRACE-REC[15]). The vertical axis is inverted for CGR so that positive (downwards) CGR anomalies indicate a weaker land carbon sink. A 6-month moving average was applied to GRACE data for readability. **(b)** Yearly CGR versus GRACE TWS anomalies. **(c-d)** Composite TWS anomalies associated with the 5% highest (c) and 5% lowest (d) monthly CGR (n=8, see Source Data). Inset bar-plots indicate the season of the corresponding months. Composites based on GRACE-REC show similar patterns (Extended Data Fig. 2).

Figure 2. Correlations between CGR and meteorological drivers over different spatial domains at monthly and yearly scale. The years 1991-1993 affected by the eruption of Mt Pinatubo are excluded (see Methods). Observations (circles) are distinguished from model-

based estimates (squares). A black cross indicates a non-significant correlation (alpha = 5%). Horizontal lines correspond to the 95% confidence interval of the correlation coefficient (see Methods). The different products and their temporal coverage as well as the number of data points used to generate these results are listed in Supplementary Information Tables 1 & 3.

Figure 3. Confounding effects of water storage and temperature on correlations with CGR. (a) Yearly co-variation between global mean GRACE TWS and global mean temperature over the period 2002-2016. **(b)** Partial correlations between GRACE TWS and CGR (in blue) remain high and significant after controlling for the effect of global or tropical temperature. Significance indicated with asterisks (alpha = 5%). **(c, d)** Probability distributions of the yearly sensitivities of CGR to TWS and to temperature estimated with a Monte Carlo approach (Methods).

Figure 4. Observed and modelled relations between global water storage, temperature and carbon fluxes. (a) Correlations of the land carbon sink with observed global mean temperature and global mean water storage (based on simulated soil moisture when correlating with model NEE and based on GRACE when correlating with observed CGR). Gray shading indicates the 95% confidence intervals (from Fig. 2b) for the observed relationships (circle). Solid and hollow squares indicate the relationships obtained with DGVMs and FluxCom models respectively. **(b)** Global mean water-driven (NEE^{Water}) and **(c)** temperature-driven (NEE^{Temp}) NEE signals in Gt C yr^{-1} (Methods). **(d)** The global mean NEE^{Water} of a given model is correlated to its simulated soil moisture signal (blue bars), while NEE^{Temp} is correlated to global mean temperature (orange bars), indicating an internal consistency between the global means of these two climatic drivers and their associated NEE response.

Methods:

GRACE water mass changes

From 2002 to 2017, the twin satellites of the Gravity Recovery and Climate Experiment (GRACE) have measured monthly anomalies of the Earth's gravity field[7,31] that can be used to retrieve relative changes in water storage, both on land and in the ocean, at a spatial resolution of about 300 km[32,33]. Over land, these observations reflect net changes in terrestrial water storage (TWS), including groundwater, soil moisture, surface waters, snow, land ice, and water stored in the biosphere. These observations of water mass redistribution are consistent with other observed geophysical constraints such as changes in sea level[34] and polar motion[35] and correlate with satellite observations of surface soil moisture[36] as well as with changes in precipitation and temperature[37]. Here, we use the Jet Propulsion Laboratory GRACE mascon solution[38,39] and exclude the contribution of Greenland and Antarctica in

order to obtain a global mean TWS signal for all available months between April 2002 and December 2016. We isolate the inter-annual variability (IAV) by subtracting the mean seasonal cycle and remove the linear trend using simple linear regression. Because we focus on global and regional averages over very large spatial domains, using different GRACE solutions (Supplementary Information Table 2) has very little impact (Supplementary Information Fig. 5. A comprehensive comparison can be found for example in Scanlon et al.[40]. We would recommend checking different solutions in the case of local case-studies. The GRACE Follow-On satellites, which were launched in May 2018, will replace the GRACE satellites and are expected to extend the gravity record by another 5-10 years.

Derivation of the CO_2 growth rate (CGR)

We use monthly time series of atmospheric CO_2 concentration from the Greenhouse Gas Marine Boundary Layer Reference (MBL) of the National Oceanic and Atmospheric Administration (NOAA/ESRL)[41,42]. This dataset compiles measurements of weekly air samples from the Cooperative Global Air Sampling Network since 1980. Similarly to Wang, et al. [3], we derive monthly CGR as the first-order difference of CO_2 concentrations between two successive months. We then remove the mean seasonal cycle and apply a 12-month moving sum to convert monthly values into annual CGR. For completeness, we also repeat the analysis at yearly scale using estimates of the Residual Land Sink (RLS) from the Global Carbon Project[1], and show that this does not affect the conclusions of the paper (Supplementary Information Fig. 6 & 7).

GRACE-REC (statistical reconstruction of GRACE)

The statistical approach used to generate the reconstruction of past TWS anomalies is explained in detail in Humphrey and colleagues[15]. In summary, a statistical model forced with daily precipitation and temperature anomalies is trained with GRACE observations and used to reconstruct past changes in water storage. In Humphrey and colleagues[15], the precipitation forcing is based on the ERA-Interim reanalysis, which has some limitations in representing tropical precipitation compared to other datasets. In this study, we reconstruct past TWS anomalies with the same approach but using a recently published merged daily precipitation product[43]. Using this new precipitation dataset leads to a small improvement in model performance, but there may still be limitations in the accuracy of the precipitation data, in particular over tropical regions. This updated TWS reconstruction is publicly accessible as part of this publication (Supplementary Information Table 1).

Global and regional land averages

The contribution of Greenland and Antarctica is removed for all analyses. Global and regional averages are weighted according to the land area of grid cells. The tropical domain definition used in this paper ranges from 24°S to 24°N, as in Wang and colleagues[3]. Information on the datasets[44-49] used to generate land averages can be obtained from Table 1 of the Supplementary Information. Land cover classes are based on MODIS MCD12C1 (Supplementary Information Fig. 2).

Monte Carlo estimate of correlation significance and uncertainty intervals

We estimate the 95% confidence interval of correlation coefficients (such as confidence intervals reported in Fig. 2), null hypothesis distributions for two-tailed significance testing, as well as distributions for univariate and bivariate linear sensitivities in Fig. 3c-d using moving-block bootstrapping[50]. The selection of the block length is a compromise between accounting for the effect of autocorrelation in the time series and keeping a sufficiently large block sample size so that random resamples stay independent. Based on different block length

selectors[50], we defined a block length of 12 months for monthly analyses (block length = 12). For yearly analyses, the block length was defined as 1 year (block length = 1, which is equivalent to a simple bootstrap approach), because the autocorrelation of time series was not significant at the yearly scale. The same procedure was applied to all considered datasets with 10'000 random samples.

Time intervals considered for various datasets and exclusion of years 1991-1993

The correlations reported in Figure 2 and Figure 4 are computed over heterogeneous time intervals in order to make use of as much data as is currently available (Supplementary Information Table 3). Pairs of time series are de-trended over their common time interval. In order to assess these correlations over a time period as homogeneous as possible, we repeat the analysis for these figures over the period 2002-2013 only and find that our conclusions remain unchanged (Supplementary Information Figure 8 & 9). The eruption of Mt. Pinatubo strongly affected radiation budgets, which perturbed both CGR[16,18] and the water cycle[17], explaining the de-coupling between CGR and water storage changes. We also reproduce Figure 2 without discarding the years following the eruption of Mt. Pinatubo (1991-1993) and find that our main conclusions remain unchanged (Supplementary Information Figure 10), although the correlation between CGR and water-related variables decreases (as can be expected from Figure 1).

Approach of Jung et al to separate temperature-driven and water-driven NEE signals

In a recent paper, Jung and colleagues[22] performed a global analysis of the drivers of NEE IAV using DGVMs and statistical models trained on flux tower measurements. They estimated the sensitivity of NEE IAV to climate drivers by fitting local multivariate regressions to the model outputs. With this approach, the simulated soil moisture and the observed temperature forcings are used to fit a linear statistical model of the monthly carbon flux response calculated by the more complex DGVMs and upscaling models. We replicated the analysis performed in Jung and colleagues[22] on DGVMs using the same set of seven DGVMs (from TRENDY v3/S2)[9] as well as upscaling models (FluxCom[23]).

Fraction of inter-annual variability (IAV)

The fraction of IAV quantifies the importance of low frequency variability in the overall variance of a given signal. It is computed as:

$$F_{IAV} = \frac{Var(X_{yearly})}{Var(X_{monthly})}$$

where, Var() denotes the variance estimator, $X_{monthly}$ is the de-seasonalised and de-trended monthly time series and X_{yearly} is the yearly time series (computed from $X_{monthly}$). This indicator is also illustrated in Extended Data Fig. 6. While being much simpler in practice, this approach shares the same idea as analyzing the relative importance of low frequencies in a signal's power spectrum. When adding GRACE measurement errors to the GRACE-REC estimates of IAV fraction in Extended Data Fig. 8, we use the measurement errors provided with the original JPL Mascons, but without applying a conservative scale factor of 2 to the diagonal elements of the formal covariance matrix (see Wiese and colleagues[38]).

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

32 Wahr, J., Molenaar, M. & Bryan, F. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *J Geophys Res-Sol Ea* **103**, 30205-30229, doi:10.1029/98JB02844 (1998).

33 Wouters, B. *et al.* GRACE, time-varying gravity, Earth system dynamics and climate change. *Rep Prog Phys* **77**, doi:10.1088/0034-4885/77/11/116801 (2014).
34 Cazenave, A. *et al.* The rate of sea-level rise. *Nature Climate Change* **4**, 358-361, doi:10.1038/nclimate2159 (2014).
35 Adhikari, S. & Ivins, E. R. Climate-driven polar motion: 2003-2015. *Sci Adv* **2**, doi:10.1126/sciadv.1501693 (2016).
36 Abelen, S., Seitz, F., Abarca-del-Rio, R. & Güntner, A. Droughts and floods in the La Plata basin in soil moisture data and GRACE. *Remote Sensing* **7**, 7324-7349, doi:10.3390/rs70607324 (2015).
37 Humphrey, V., Gudmundsson, L. & Seneviratne, S. I. Assessing Global Water Storage Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes. *Surv Geophys* **37**, 357-395, doi:10.1007/s10712-016-9367-1 (2016).
38 Wiese, D. N., Landerer, F. W. & Watkins, M. M. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resour Res* **52**, 7490-7502, doi:10.1002/2016WR019344 (2016).
39 Watkins, M. M., Wiese, D. N., Yuan, D. N., Boening, C. & Landerer, F. W. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research-Solid Earth* **120**, 2648-2671, doi:10.1002/2014JB011547 (2015).
40 Scanlon, B. R. *et al.* Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resour Res* **52**, 9412-9429, doi:10.1002/2016wr019494 (2016).
41 Masarie, K. A. & Tans, P. P. Extension and Integration of Atmospheric Carbon-Dioxide Data into a Globally Consistent Measurement Record. *Journal of Geophysical Research-Atmospheres* **100**, 11593-11610, doi:10.1029/95JD00859 (1995).
42 Dlugokencky, E. & Tans, P. *Trends in atmospheric carbon dioxide, National Oceanic & Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL)*, <http://www.esrl.noaa.gov/gmd/ccgg/trends/> (2014).
43 Beck, H. E. *et al.* MSWEP: 3-hourly 0.25 degrees global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrol Earth Syst Sc* **21**, 589-615, doi:10.5194/hess-21-589-2017 (2017).
44 Adler, R. F. *et al.* The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *J Hydrometeorol* **4**, 1147-1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2 (2003).
45 Huffman, G. J., Adler, R. F., Bolvin, D. T. & Gu, G. J. Improving the global precipitation record: GPCP Version 2.1. *Geophys Res Lett* **36**, doi:10.1029/2009GL040000 (2009).
46 Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int J Climatol* **34**, 623-642, doi:10.1002/joc.3711 (2014).
47 Rohde, R. *et al.* A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011. *Geoinformatics & Geostatistics: An Overview* **01**, doi:10.4172/2327-4581.1000101 (2013).
48 Wolter, K. & Timlin, M. S. Measuring the strength of ENSO events: How does 1997/98 rank? *Weather* **53**, 315-324, doi:10.1002/j.1477-8696.1998.tb06408.x (1998).
49 Friedl, M. A. *et al.* MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens Environ* **114**, 168-182, doi:10.1016/j.rse.2009.08.016 (2010).
50 Mudelsee, M. *Climate Time Series Analysis : Classical Statistical and Bootstrap Methods*. 2nd edn, (Springer, 2014).

Data availability statement

All datasets generated or analysed during this study are available from the links listed in Table 1 of the Supplementary Information. The source data for Fig. 1a–b, Fig. 2, Fig. 3 and Fig. 4 are additionally provided as spreadsheets with the online version of the paper.

Extended Data Figure 1. Ecosystems respond to water storage. Water storage is more relevant than precipitation when investigating the impacts of changes in water availability on ecosystems.

Extended Data Figure 2. Reproduction of main Figure 1c-d with GRACE-REC. Composite TWS anomalies associated with the 5% highest (c) and 5% lowest (d) monthly

CGR (n=20 months in each case) based on GRACE-REC (i.e. covering the 1980-2016 time period). Inset bar-plots indicate the season of the selected months.

Extended Data Figure 3. ENSO, precipitation and terrestrial water storage. Because it integrates precipitation anomalies, water storage is slightly phase shifted with respect to ENSO and precipitation time series. Here, El Niño (La Niña) conditions correspond to the periods where the Multivariate ENSO Index (MEI) exceeds 0.5 (-0.5). Strongest ENSO events (MEI >1 or <-1) are shown in darker color.

Extended Data Figure 4. Dominant contribution of precipitation to terrestrial water storage anomalies. (a) Global means of GRACE-REC and GRACE-REC driven only with precipitation anomalies. The statistical reconstruction of GRACE (GRACE-REC) is calibrated with both precipitation and temperature information[15]. We use this model to predict the precipitation-driven component of the TWS signal (by setting temperature variability to zero). Most of the global TWS signal can be reconstructed based on precipitation anomalies only. **(b)** Performance of the GRACE-REC model at the grid scale. **(c)** Contribution of precipitation to the locally reconstructed TWS. A comparison between GRACE-REC, global hydrological models and GRACE can also be found in Humphrey and colleagues[15].

Extended Data Figure 5. Reproduction of main Figure 3 with mean precipitation. Same as Figure 3, but using yearly precipitation from GPCP (with a 4-month lag) instead of water storage from GRACE.

Extended Data Figure 6. Illustration of soil moisture signals with different fractions of inter-annual variability (IAV). The fraction of IAV quantifies the importance of low frequency variability in the overall variance of a given signal. Here, it is defined as the ratio between the variance of the yearly (de-trended) time series **(b)** and the variance of the monthly anomalies **(a)** (see Methods). The fraction of IAV tends to increase when deeper soil layers are included. This is because deeper layers have a longer residence time (or memory) and thus respond more slowly to changes in the meteorological forcing. Illustrative data based on GLDAS2-Noah, extracted for Spain (4.25°W, 40.25°N).

Extended Data Figure 7. Fraction of IAV in water storage changes. (a) Average fraction of IAV in water storage changes simulated by DGVMs and FluxCom (which typically only include root-zone soil moisture). **(b)** Fraction of IAV in water storage changes observed by GRACE (which include all water reservoirs). In order to ensure comparability between models and GRACE, model outputs were first averaged to the spatial resolution of GRACE. Note that unlike modelled soil moisture, GRACE observations suffer from measurement errors that tend to increase the high-frequency (month-to-month) variability. Therefore, the fraction of IAV retrieved from GRACE would be even higher if there was no measurement error in GRACE.

Extended Data Figure 8. Distribution of the fraction of IAV by land cover classes. This compares the values shown in the maps of Extended Data Figure 7 for different land cover classes. The fraction of IAV found in GRACE TWS (dark blue) is higher compared to models (green). Because GRACE observations are contaminated by high-frequency measurement errors, the fraction of IAV found in GRACE is shifted towards lower values. Here, the fraction of IAV derived from GRACE-REC (light blue) may provide a more robust estimate of the actual fraction of IAV in TWS. Adding GRACE measurement errors (as provided with

GRACE NASA-JPL data) to the GRACE-REC data reproduces very well the overall shift (dashed light blue) towards lower values that occurs with original GRACE data.

Extended Data Figure 9. Relationship between the fraction of IAV in water storage and the fraction of IAV in NEE_{Water}. Left panels **(a,c)** show the mean fraction of IAV obtained at all grid cells (for TRENDY and FluxCom), with point cloud density indicated by the color shading. The fraction of IAV in NEE_{Water} is directly limited by the fraction of IAV present in the underlying water storage signal. **(b,d)** Same as (a,c) stratified by land cover class. In land cover classes that are typically moisture-limited (e.g. semi-arid), the fraction of IAV in NEE_{Water} is potentially strongly limited by the fraction of IAV in water storage. **(e)** This relationship is also found for the global mean signals of the individual models.

Extended Data Figure 10. Contribution of six different land cover types to the global water storage signal. (a) GRACE TWS anomalies by land cover type, smoothed with a 6-month moving average and offset for readability. **(b)** Regional contributions to the global water storage signal. High values indicate that a region bears a high contribution to the overall global mean water storage signal. This metric is based on the definition proposed in Ahlström and colleagues[27] for analyzing regional contributions to global net biome production (NBP). The value reported for the models is the mean across all models.

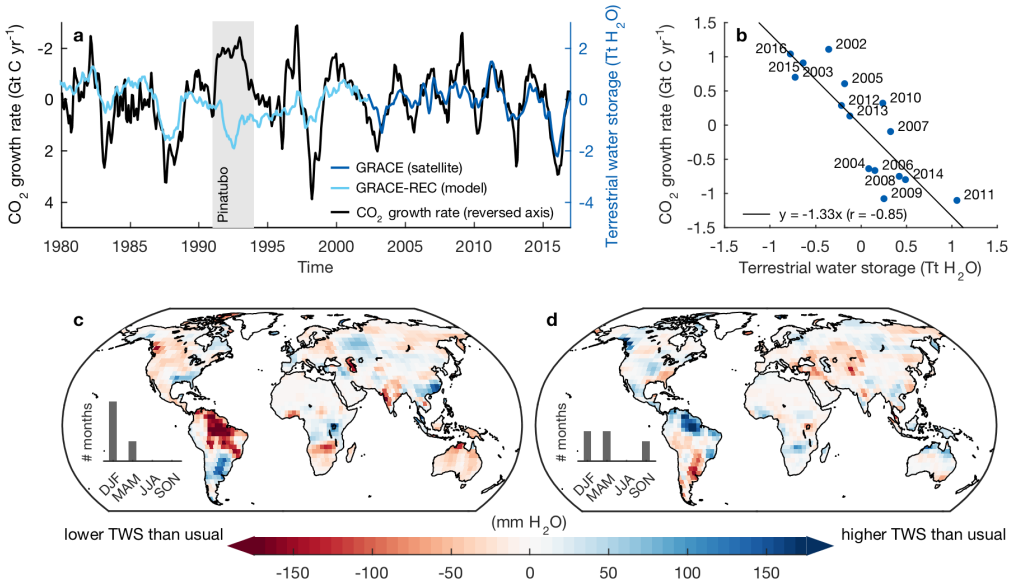


Figure 1. Inter-annual variability in CGR and TWS. (a) Monthly de-seasonalised and de-trended CGR, TWS from satellite observations (GRACE) and TWS from a statistical model (GRACE-REC[15]). The vertical axis is inverted for CGR so that positive (downwards) CGR anomalies indicate a weaker land carbon sink. A 6-month moving average was applied to GRACE data for readability. **(b)** Yearly CGR versus GRACE TWS anomalies. **(c-d)** Composite TWS anomalies associated with the 5% highest (c) and 5% lowest (d) monthly CGR (n=8, see Source Data). Inset bar-plots indicate the season of the corresponding months. Composites based on GRACE-REC show similar patterns (Extended Data Fig. 2).

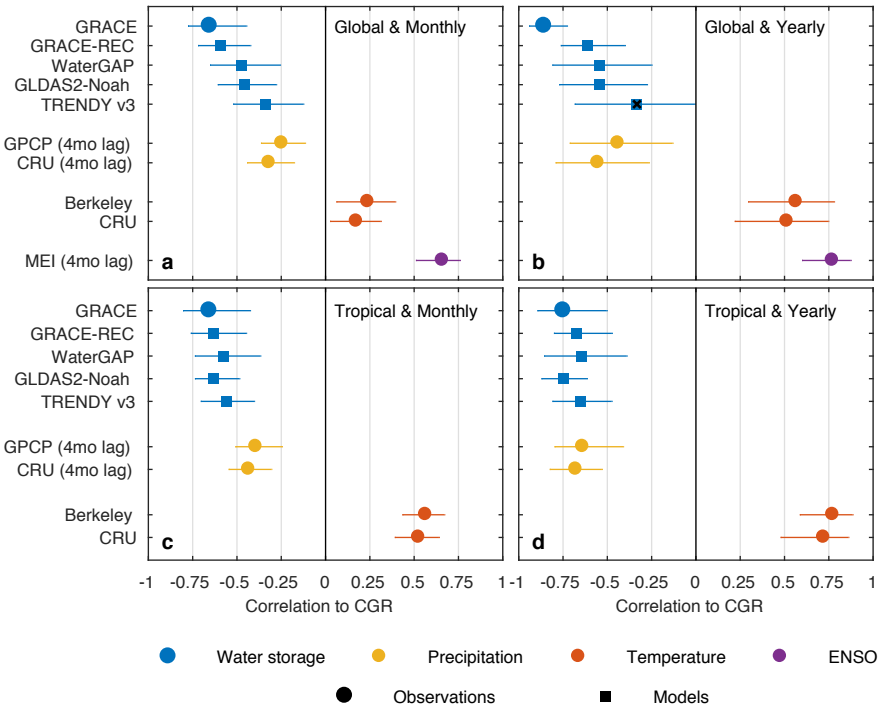


Figure 2. Correlations between CGR and meteorological drivers over different spatial domains at monthly and yearly scale. The years 1991-1993 affected by the eruption of Mt Pinatubo are excluded (Methods). Observations (circles) are distinguished from model-based estimates (squares). A black cross indicates a non-significant correlation (alpha = 5%). Horizontal lines correspond to the 95% confidence interval of the correlation coefficient (Methods). The different products as well as the number of data points used to generate these results are listed in Supplementary Information Tables 1 & 3.

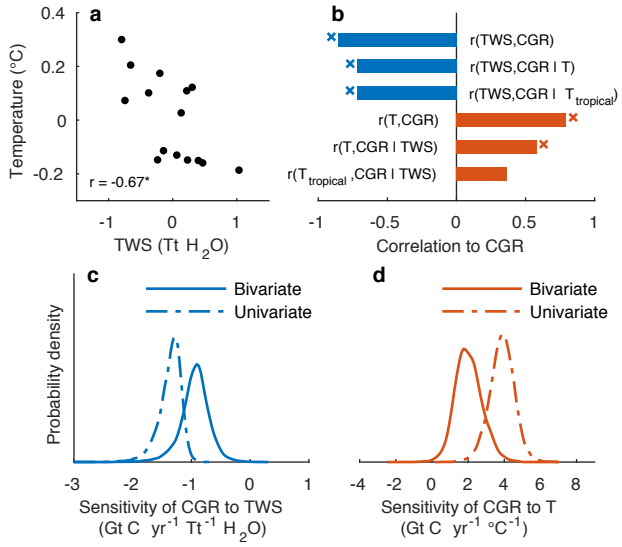


Figure 3. Confounding effects of water storage and temperature on correlations with CGR. (a) Yearly co-variation between global mean GRACE TWS and global mean temperature over the period 2002-2016. **(b)** Partial correlations between GRACE TWS and CGR (in blue) remain high and significant after controlling for the effect of global or tropical temperature. Significance indicated with asterisks (alpha = 5%). **(c, d)** Probability distributions of the yearly sensitivities of CGR to TWS and to temperature (Methods).

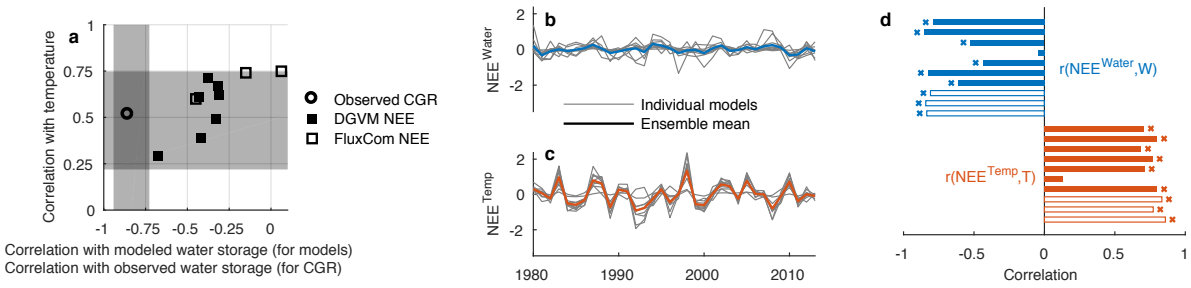
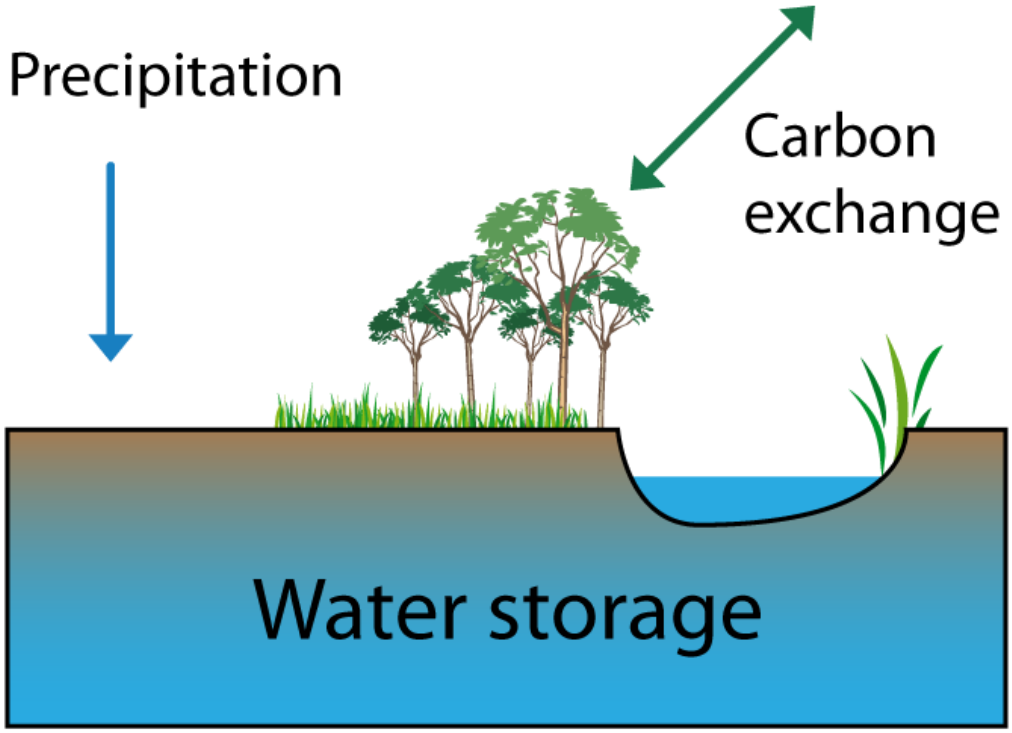
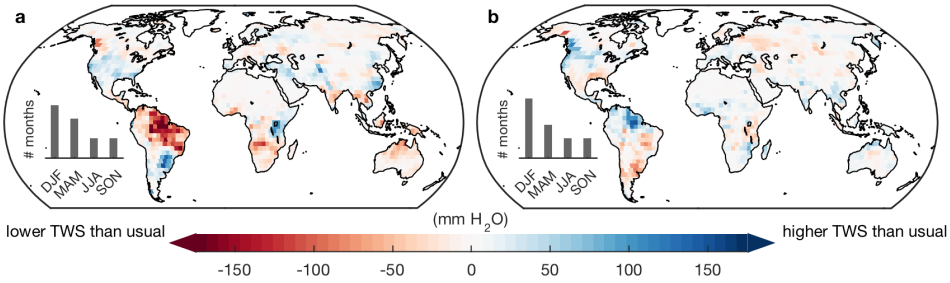


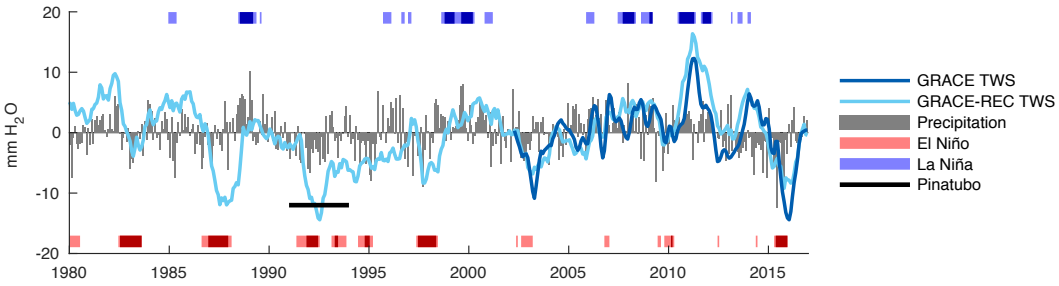
Figure 4. Observed and modelled relations between global water storage, temperature and carbon fluxes. (a) Correlations of the land carbon sink with global mean temperature and global mean water storage (based on simulated soil moisture when correlating with model NEE and based on GRACE when correlating with observed CGR). Gray shading indicates the 95% confidence intervals (from Fig. 2b) for the observed relationships (circle). Solid and hollow squares indicate the relationships obtained with DGVMs and FluxCom models respectively. **(b)** Global mean water-driven (NEE^{Water}) and **(c)** temperature-driven (NEE^{Temp}) NEE signals in Gt C yr^{-1} (Methods). **(d)** The global mean NEE^{Water} of a given model is correlated to its simulated soil moisture signal (blue bars), while NEE^{Temp} is correlated to global mean temperature (orange bars), indicating an internal consistency between the global means of these two climatic drivers and their associated NEE response.



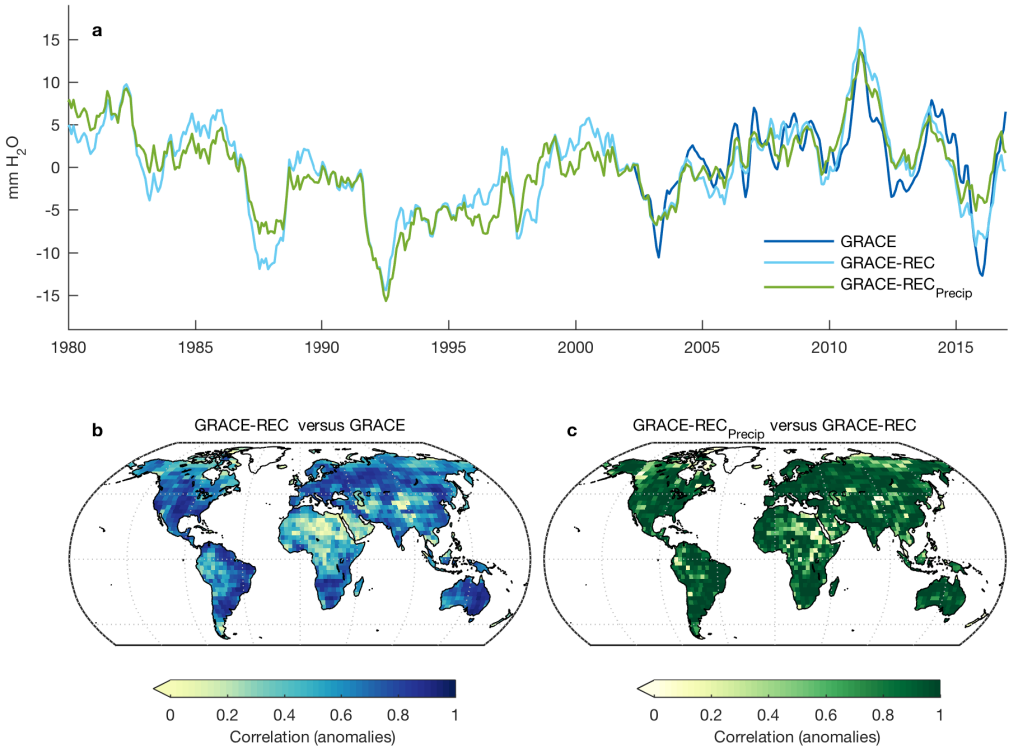
Extended Data Figure 1. Ecosystems respond to water storage. Water storage is more relevant than precipitation when investigating the impacts of changes in water availability on ecosystems.



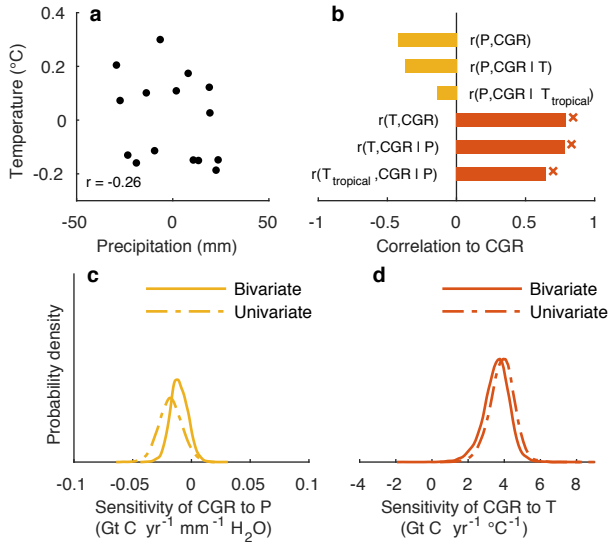
Extended Data Figure 2. Reproduction of main Figure 1c-d with GRACE-REC. Composite TWS anomalies associated with the 5% highest (c) and 5% lowest (d) monthly CGR (n=20 months in each case) based on GRACE-REC (i.e. covering the 1980-2016 time period). Inset bar-plots indicate the season of the selected months.



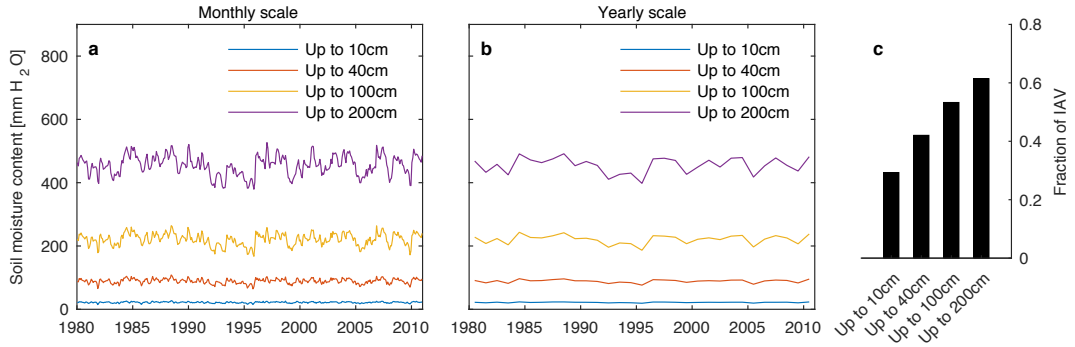
Extended Data Figure 3. ENSO, precipitation and terrestrial water storage. Because it integrates precipitation anomalies, water storage is slightly phase shifted with respect to ENSO and precipitation time series. Here, El Niño (La Niña) conditions correspond to the periods where the Multivariate ENSO Index (MEI) exceeds 0.5 (-0.5). Strongest ENSO events (MEI >1 or <-1) are shown in darker color.



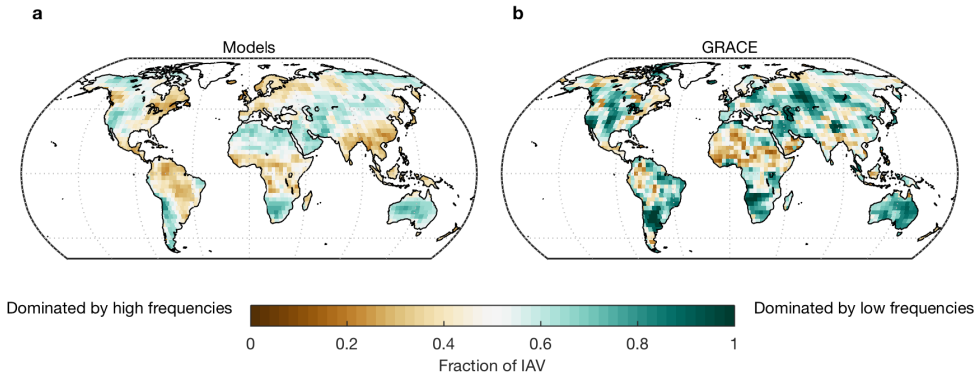
Extended Data Figure 4. Dominant contribution of precipitation to terrestrial water storage anomalies. (a) Global means of GRACE-REC and GRACE-REC driven only with precipitation anomalies. The statistical reconstruction of GRACE (GRACE-REC) is calibrated with both precipitation and temperature information[15]. We use this model to predict the precipitation-driven component of the TWS signal (by setting temperature variability to zero). Most of the global TWS signal can be reconstructed based on precipitation anomalies only. **(b)** Performance of the GRACE-REC model at the grid scale. **(c)** Contribution of precipitation to the locally reconstructed TWS. A comparison between GRACE-REC, global hydrological models and GRACE can also be found in Humphrey and colleagues[15].



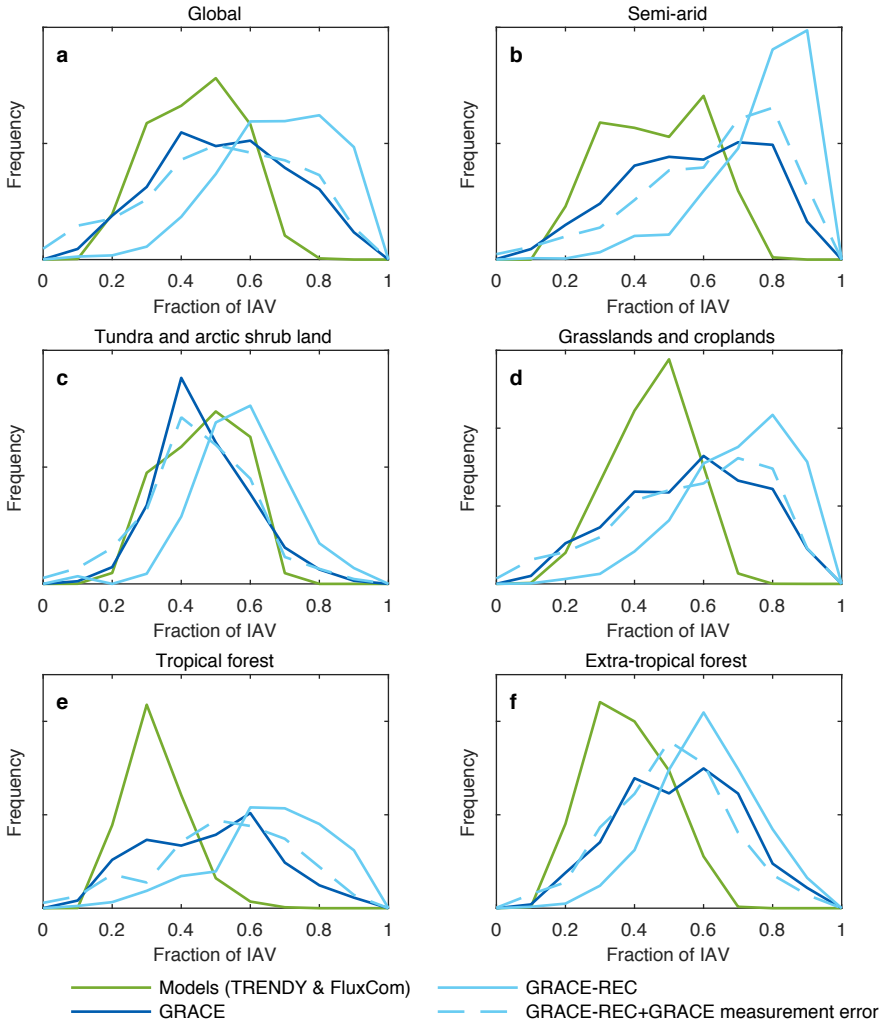
Extended Data Figure 5. Reproduction of main Figure 3 with mean precipitation. Same as Figure 3, but using yearly precipitation from GPCP (with a 4-month lag) instead of water storage from GRACE.



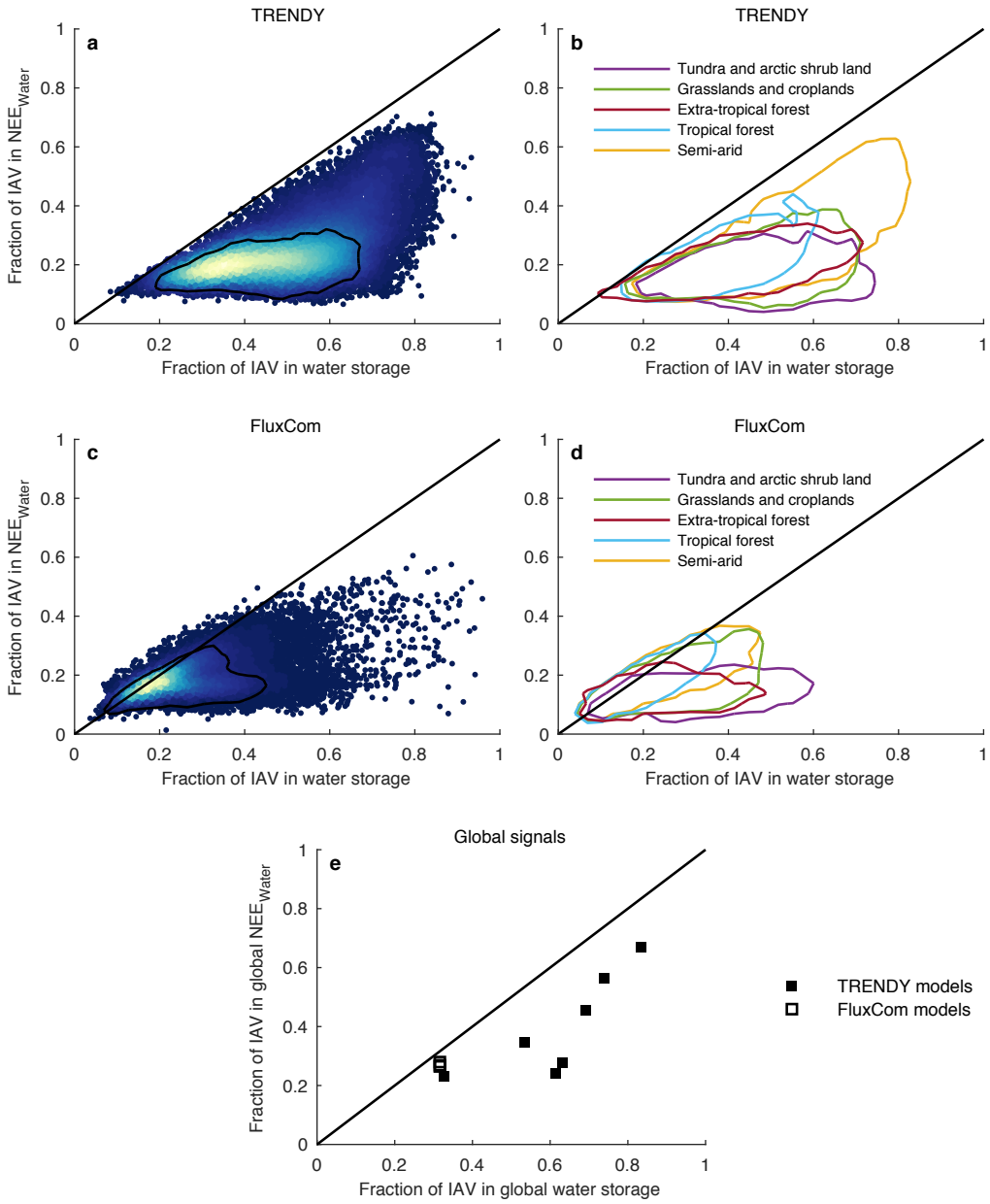
Extended Data Figure 6. Illustration of soil moisture signals with different fractions of inter-annual variability (IAV). The fraction of IAV quantifies the importance of low frequency variability in the overall variance of a given signal. Here, it is defined as the ratio between the variance of the yearly (de-trended) time series **(b)** and the variance of the monthly anomalies **(a)** (see Methods). The fraction of IAV tends to increase when deeper soil layers are included. This is because deeper layers have a longer residence time (or memory) and thus respond more slowly to changes in the meteorological forcing. Illustrative data based on GLDAS2-Noah, extracted for Spain (4.25°W, 40.25°N).



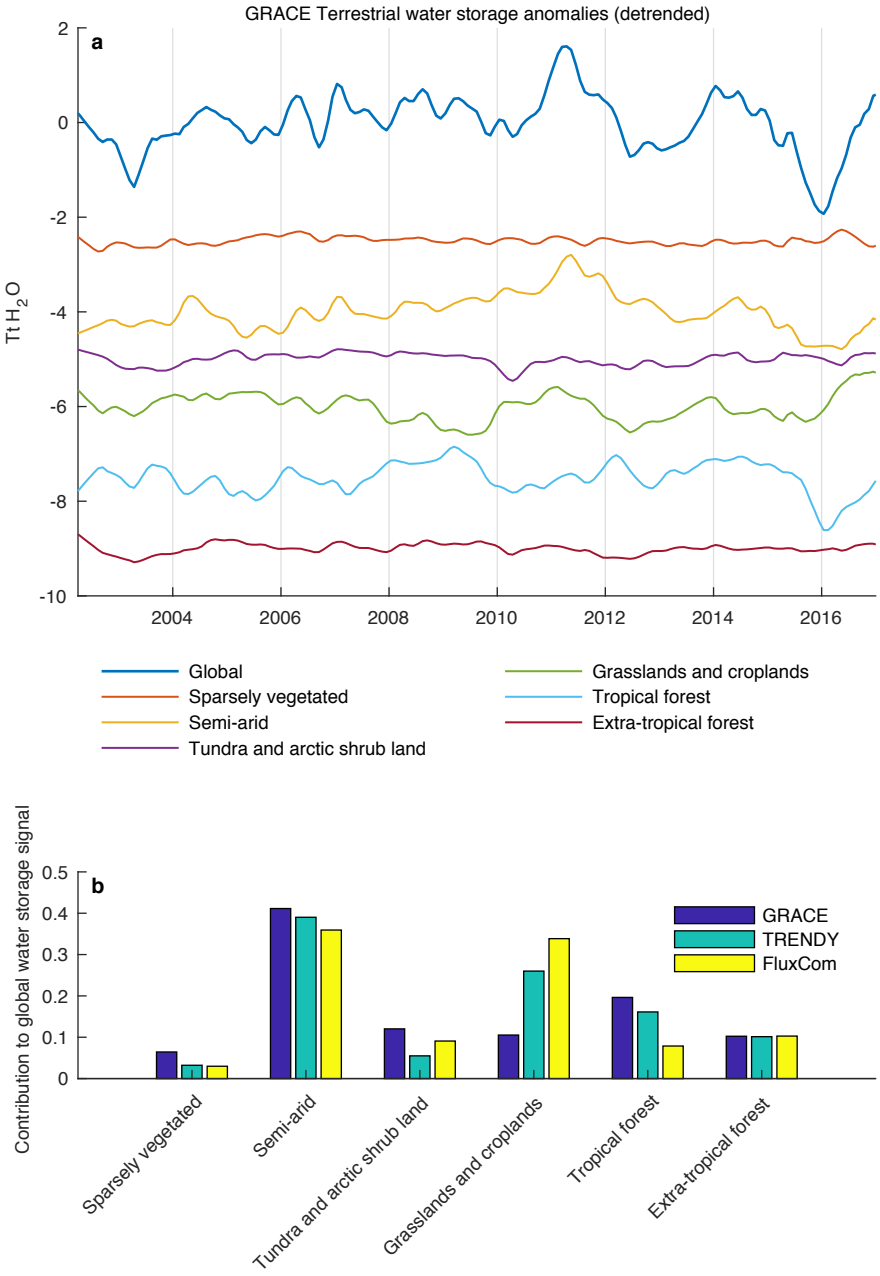
Extended Data Figure 7. Fraction of IAV in water storage changes. (a) Average fraction of IAV in water storage changes simulated by DGVMs and FluxCom (which typically only include root-zone soil moisture). **(b)** Fraction of IAV in water storage changes observed by GRACE (which include all water reservoirs). In order to ensure comparability between models and GRACE, model outputs were first averaged to the spatial resolution of GRACE. Note that unlike modelled soil moisture, GRACE observations suffer from measurement errors that tend to increase the high-frequency (month-to-month) variability. Therefore, the fraction of IAV retrieved from GRACE would be even higher if there was no measurement error in GRACE.



Extended Data Figure 8. Distribution of the fraction of IAV by land cover classes. This compares the values shown in the maps of Extended Data Figure 7 for different land cover classes. The fraction of IAV found in GRACE TWS (dark blue) is higher compared to models (green). Because GRACE observations are contaminated by high-frequency measurement errors, the fraction of IAV found in GRACE is shifted towards lower values. Here, the fraction of IAV derived from GRACE-REC (light blue) may provide a more robust estimate of the actual fraction of IAV in TWS. Adding GRACE measurement errors (as provided with GRACE NASA-JPL data) to the GRACE-REC data reproduces very well the overall shift (dashed light blue) towards lower values that occurs with original GRACE data.



Extended Data Figure 9. Relationship between the fraction of IAV in water storage and the fraction of IAV in NEE_{Water}. Left panels **(a,c)** show the mean fraction of IAV obtained at all grid cells (for TRENDY and FluxCom), with point cloud density indicated by the color shading. The fraction of IAV in NEE_{Water} is directly limited by the fraction of IAV present in the underlying water storage signal. **(b,d)** Same as (a,c) stratified by land cover class. In land cover classes that are typically moisture-limited (e.g. semi-arid), the fraction of IAV in NEE_{Water} is potentially strongly limited by the fraction of IAV in water storage. **(e)** This relationship is also found for the global mean signals of the individual models.



Extended Data Figure 10. Contribution of six different land cover types to the global water storage signal. (a) GRACE TWS anomalies by land cover type, smoothed with a 6-month moving average and offset for readability. **(b)** Regional contributions to the global water storage signal. High values indicate that a region bears a high contribution to the overall global mean water storage signal. This metric is based on the definition proposed in Ahlström and colleagues[27] for analyzing regional contributions to global net biome production (NBP). The value reported for the models is the mean across all models.