

Supplementary Materials: Emerging Changes in Terrestrial Water Storage Variability as a Target for Future Satellite Gravity Missions

Laura Jensen ^{1,*} , Annette Eicker ¹ , Henryk Dobslaw ² , and Roland Pail ³ 

1. Regions influenced by groundwater, surface water, or glaciers

In some regions TWS from GRACE observations and mTWS from CMIP6 models is not directly comparable since the CMIP6 ESMs do not represent groundwater, surface water storage, and glacier mass changes explicitly. These discrepancies could in principle be reduced by either introducing groundwater, surface water and glacier representations into the ESMs or by removing these components from the integral GRACE signal. The former would be a rewarding effort for future model development as these components have significant impact on land-atmosphere interactions influencing both moisture and energy fluxes and thereby altering climate. The later depends on additional observations or on model results, which introduce uncertainties into the residual GRACE estimates. To identify regions with particularly large differences between TWS and mTWS we compare the full GRACE TWS signal to a TWS signal with the effect of surface water storage and groundwater storage changes removed. In addition, we consider glaciated regions to be affected by discrepancies in TWS and mTWS.

To estimate the effect of surface water storage we make use of an observational data set [1] containing mass change time series of 283 large lakes and reservoirs derived from combining surface water levels (from satellite altimetry) with surface water extent (from remote sensing). By forward modeling these surface water changes to spherical harmonics and applying a spatial DDK3 filtering [2] the data were made comparable to the spatial resolution of GRACE and subsequently mapped to the 2° grid used in this study.

For groundwater long in-situ records are sparse and models exhibit large uncertainties, making a sensible global separation from GRACE-derived TWS unfeasible. However, for natural groundwater variability we consider the discrepancies between models and observations to be minor. Even though the ESMs do not represent groundwater-soil water interactions, they implicitly contain large fractions of the groundwater within their deeper soil layers because the water balance in the models is largely closed and the mass transport to the ocean and atmosphere is limited [3]. Therefore, here we focus on anthropogenic groundwater abstractions only, which are definitely not included in the climate models. We access data from the hydrological model WaterGAP 2.2d [4] where net abstraction groundwater is defined as groundwater withdrawals minus return flow from irrigation. Since irrigation is partly taken from surface waters, net abstraction from groundwater lead to mass increase. We convert the monthly (2003/01 – 2016/12) global grids from fluxes (in mm/s) to monthly accumulated water storage change (EWH in mm) and remap it to 2° spatial resolution. We then apply a GRACE-like spatial filtering (DDK3 filter) [2].

Afterward, we compute the RMS over 2003/01 – 2016/12 of the annual plus interannual signal (the main signal components discussed in the paper) and compare (1) the full GRACE time series and (2) the GRACE time series minus the surface water storage and net abstraction time series. The relative difference of these two RMS values is displayed in Figure S1. The differences are mostly positive, which means that the signal gets smaller after removing estimates of surface water storage and (anthropogenic) groundwater change, which is expected. For surface water storage (Figure S1a) the RMS reduction is generally larger than for the net abstraction (Figure S1b). Smaller values for the net abstraction are due to the fact that they mainly occur as a linear mass trends (which is not investigated here) whereas seasonal and year-to-year variations are minor. The relative RMS difference for the combined surface water and net abstraction effect is displayed in Figure S1c. Regions where

the difference exceeds 10% of the total (annual and interannual) signal are shown in red in Figure S1d. These regions make up 8% of the land surface (Greenland and Antarctica excluded).

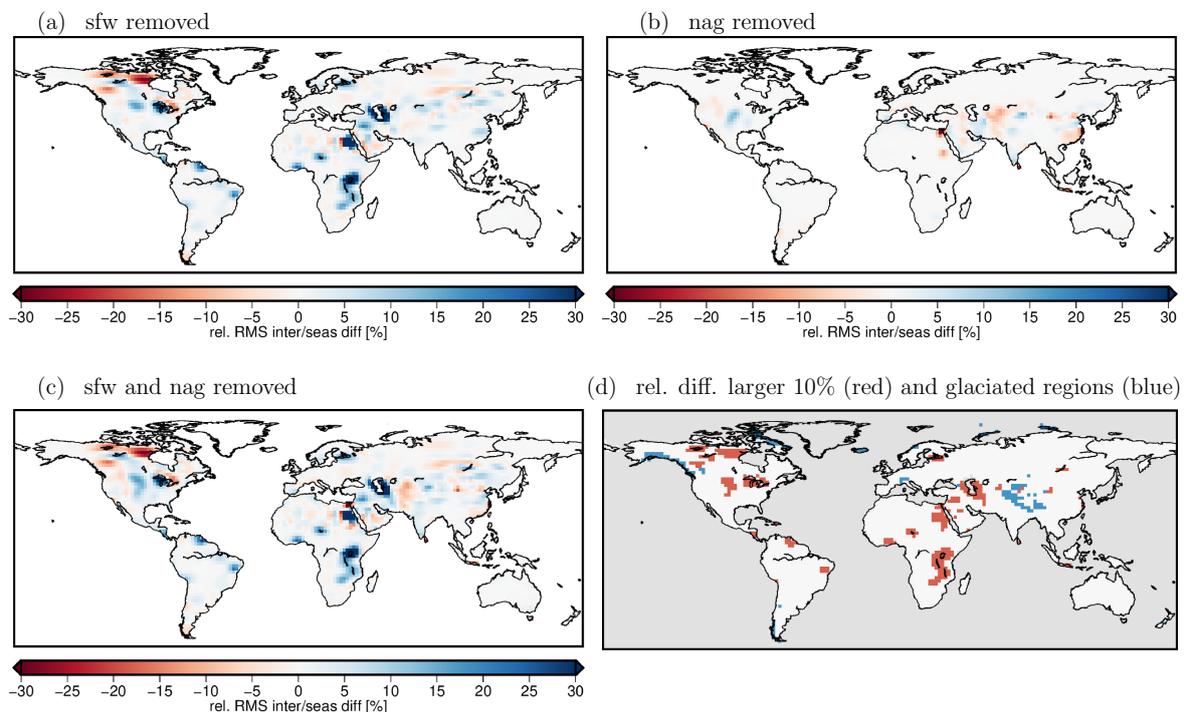


Figure S1. (a) Relative difference between the interannual plus annual RMS of GRACE and GRACE reduced by surface water storage changes for the time span 2003 – 2016. (b) same as (a) for net groundwater abstraction. (c) same as (a) for combined surface water and net abstraction changes. (d) regions where relative difference from (c) exceeds 10% (red) and regions partly covered by glaciers (blue).

The reduction of the GRACE time series by the two data sets described above provides a good estimate which regions are most affected by TWS and mTWS discrepancies. However, the currently available data set for surface water changes is still affected by large uncertainties and cannot capture all surface water bodies due to a spatial undersampling of, e.g., rivers and smaller lakes. This can be expected to improve once the SWOT data [5] become available. Also groundwater abstractions from hydrological modeling have uncertainties that are difficult to quantify due to sparse observations. Accurate quantification of both components still requires more research and with respect to the impractical quantification of errors we abstained from actually removing the surface water and abstraction data from the GRACE observations in the actual comparison with CMIP6 models in the main article. We rather indicate in Figure S1d regions where the results have to be interpreted with care as the discrepancies between observed and modeled TWS are probably larger here than in the remaining land area.

In addition to groundwater and surface water, also glacier mass changes are not contained in CMIP6 models, but observed by GRACE. In glaciated regions, moisture dynamics are not purely driven by soil moisture and snow variability but also influenced by ice mass changes, constituting a discrepancy between TWS and mTWS. To identify these regions, we access the GLIMS (Global Land Ice Measurements from Space) Glacier Database [6], currently containing the outlines of about 546300 glaciers from around the world. To quantify the degree to which a 2° grid cell (the resolution of the maps in this study) is covered by glaciers, we raster the GLIMS glacier polygons to a rather fine resolution of 0.025° and afterward count the number of 0.025° glacier grid cells within a 2° grid cell. In Figure S1d all 2° grid cells where more than 100 small 0.025° grid cells are glaciated are marked in blue. These regions make up 3% of the land surface.

2. Identification of independent CMIP6 models

At the time of writing, the CMIP6 data base contains mrso and snw data from 25 models. However, some of these models share the same sub-models (land, ocean, or atmosphere) or are extensions and sub-versions from each other. Thus, not all model results can be considered to be independent from each other. If all 25 models would contribute to the multi-model median, it would be biased towards particular models because their data would be included multiple times. Therefore, we exclude dependent models. Two models are considered not to be independent from each other when their long-term (2000 – 2100) linear trend patterns are very similar. This criterion was also applied in Jensen *et al.* [7] and has been shown to effectively identify models with common components. The trend pattern of each model is obtained from the ensemble mean of all runs belonging to the respective model via least squares adjustment. The similarity between the trend maps is measured by the pattern correlation, which is calculated as Pearson's product-moment correlation coefficient of the vectorized trend maps (excluding ocean grid cells). The pattern correlation of the trend maps between all 25 model ensemble means is displayed in Figure S2. A correlation of 70% or more was defined as the threshold for a model to be excluded. The criteria for the selection of a specific model from the highly correlated models were its degree of specialization (most general), spatial resolution (closest to 2°), or the number of ensemble members (most members).

The models excluded by these criteria are marked in light gray in Figure S2. Detailed information and references for the 17 models remaining (bold black) can be accessed, e.g., via <https://esgf-data.dkrz.de/projects/cmip6-dkrz/> (last visit: 9/24/2020).

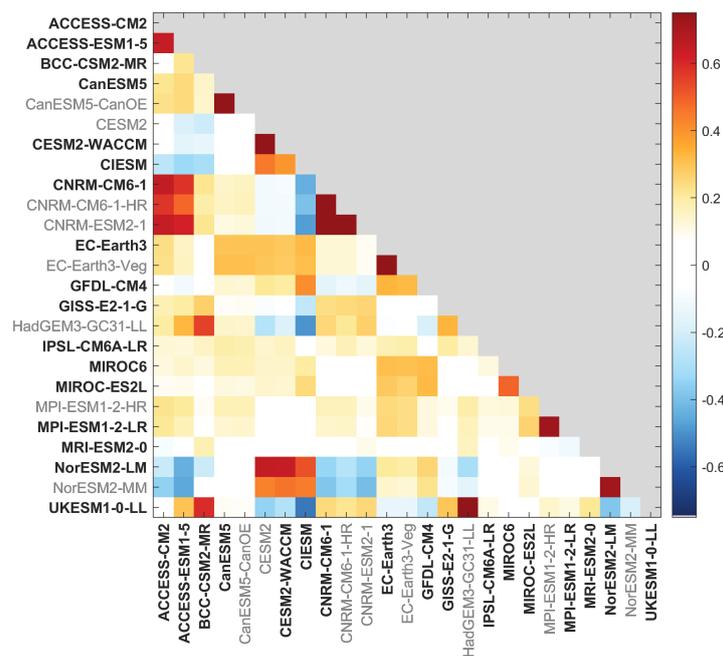


Figure S2. Correlations of long-term (2000 – 2100) linear trend patterns from 25 CMIP6 model ensemble means. Models excluded from the analysis are marked in light gray, models used in the study are highlighted in bold black.

3. Example for signal decomposition of GRACE TWS time series

An example TWS time series from GRACE and its decomposition into long-term, seasonal and sub-seasonal components is shown in Figure S3. Compared to the modeled time series (Figure 1 of main text) the sub-seasonal component is larger in GRACE. This is probably partly due to observational noise, but also due to real signals captured by GRACE on a month-to-month time scale. It has been shown that GRACE can observe TWS signals even on time scales down to a few days [8], and it is not clear to which extent these variations are reproduced by the ESMs. Therefore, we concentrate in our study on seasonal and interannual variations. Please note that the outlier in the GRACE data (January 2015) is due to a repeat-orbit constellation leading to a degraded monthly solution. However, this does only marginally influence the estimation of the seasonal and interannual parameters.

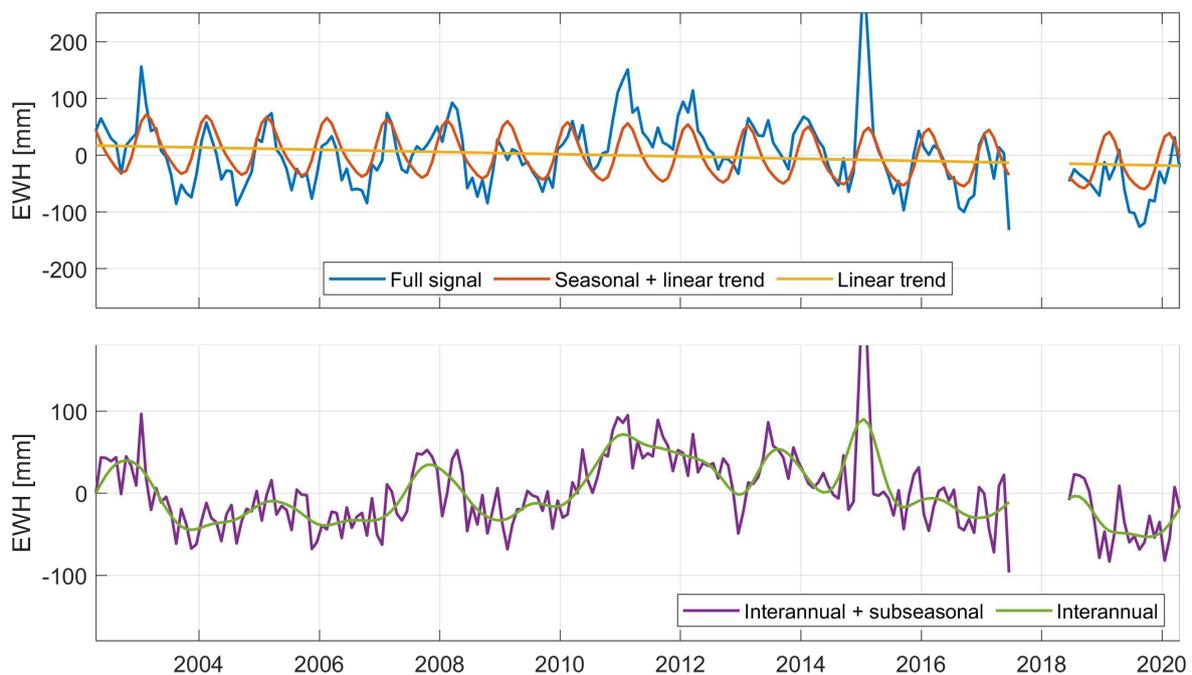


Figure S3. Example for the decomposition of a GRACE TWS time series into linear trend, seasonal, subseasonal and interannual signal. The location is 13°E and 52.5°N (Potsdam, Germany).

4. Statistics for different climate zones

Table S1. Comparison of annual amplitude from GRACE and CMIP6 MMed (2002/04 – 2020/04) for different climate zones. Column 1: percentage of land area lying in the respective climate zone; column 2: percentage of area with a SNR < 1; column 3: percentage of land area lying in the respective climate zone after excluding areas with SNR < 1; columns 4 and 5: area percentage of overestimation (ratio > 1) and underestimation (ratio ≤ 1) of the CMIP6 amplitude w.r.t. the GRACE amplitude (after excluding regions with SNR < 1); columns 6 and 7: median of the ratios in the over- and underestimation areas.

Amplitude	% area	% SNR < 1	% area w/o SNR < 1	% over	% under	median over	median under
global	100.0	24.0	100.0	57.5	42.5	1.38	0.79
equatorial	21.7	10.0	25.8	50.6	49.4	1.46	0.76
arid	36.0	49.5	24.0	55.2	44.8	1.51	0.70
temperate	16.1	10.1	19.0	57.0	43.0	1.41	0.74
polar	26.2	9.9	31.2	65.4	34.6	1.28	0.87

Table S2. As table S1, but for the phase of the annual cycle. Columns 1 and 2: area percentage of positive time shift (models earlier) and negative time shift (models later) of the CMIP6 phase w.r.t. the GRACE phase (after excluding regions with SNR < 1); columns 3 and 4: median of the phase shift in the positive and negative areas.

Phase	% earlier	% later	median earlier [months]	median later [months]
global	71.9	28.1	0.50	-0.32
equatorial	62.1	37.9	0.43	-0.36
arid	72.5	27.5	0.59	-0.48
temperate	68.2	31.8	0.66	-0.27
polar	81.8	18.3	0.45	-0.20

Table S3. As table S1, but for the RMS of the interannual signal. Column 1: percentage of area with a SNR < 1; column 2: percentage of land area lying in the respective climate zone after excluding areas with SNR < 1; columns 3 and 4: area percentage of overestimation (ratio > 1) and underestimation (ratio ≤ 1) of the CMIP6 interannual RMS w.r.t. the GRACE interannual RMS (after excluding regions with SNR < 1); columns 5 and 6: median of the ratios in the over- and underestimation areas.

Inter. RMS	% SNR < 1	% area w/o SNR < 1	% over	% under	median over	median under
global	14.7	100.0	40.2	59.9	1.30	0.69
equatorial	12.4	22.6	46.5	53.5	1.39	0.76
arid	29.6	29.6	32.2	67.8	1.38	0.56
temperate	3.6	18.1	42.7	57.3	1.30	0.67
polar	3.1	29.7	41.7	58.3	1.21	0.77

Table S4. Analysis of CMIP6 MMed annual amplitude changes (2000/01 – 2100/12) for different climate zones. Upper part, column 1: the percentage of land area where > 75% of the models agree on the sign of the trend (high consensus); columns 2 and 3: area percentage of positive and negative trends; columns 4 and 5: median trend calculated separately over the positive and negative areas. Lower part, column 1: percentage of land area lying in the respective climate zone when restricting to regions of high model consensus; columns 2 to 5: as upper part but for high consensus regions only.

Amp. Change	% highcons	% pos	% neg	median pos [mm/yr]	median neg [mm/yr]
global	44.5	56.0	44.0	0.12	-0.11
equatorial	34.0	51.3	48.8	0.13	-0.18
arid	41.9	47.3	52.7	0.06	-0.04
temperate	48.2	61.5	38.5	0.13	-0.15
polar	54.3	68.5	31.5	0.18	-0.21
high consensus	% area				
global	100.0	66.3	33.7	0.21	-0.26
equatorial	16.6	58.6	41.4	0.26	-0.37
arid	33.9	59.9	40.1	0.11	-0.15
temperate	17.4	70.3	29.7	0.20	-0.31
polar	32.0	74.8	25.2	0.27	-0.42

Table S5. As table S4, but for changes of the phase of the annual cycle. Upper part, column 1: the percentage of land area where > 75% of the models agree on the sign of the phase shift (high consensus); columns 2 and 3: area percentage of positive (later) and negative (earlier) phase shifts; columns 4 and 5: median phase shift calculated separately over the positive and negative areas. Lower part, column 1: percentage of land area lying in the respective climate zone when restricting to regions of high model consensus; columns 2 to 5: as upper part but for high consensus regions only.

Phase shift	% highcons	% pos/later	% neg/earlier	median pos [months]	median neg [months]
global	36.7	54.8	45.2	0.39	-0.35
equatorial	43.8	74.8	25.3	0.49	-0.38
arid	32.3	55.3	44.7	0.46	-0.38
temperate	39.4	49.3	50.7	0.35	-0.31
polar	35.1	40.9	59.1	0.25	-0.33
high consensus	% area				
global	100.0	60.7	39.3	0.70	-0.78
equatorial	25.9	85.0	15.0	0.72	-0.77
arid	31.7	66.6	33.4	0.94	-0.95
temperate	17.2	55.6	44.4	0.56	-1.15
polar	25.1	31.8	68.2	0.45	-0.65

Table S6. As table S4, but for changes of the RMS of the interannual signal. Upper part, column 1: the percentage of land area where > 75% of the models agree on the sign of the change (high consensus); columns 2 and 3: area percentage of positive and negative changes; columns 4 and 5: median interannual RMS change calculated separately over the positive and negative areas. Lower part, column 1: percentage of land area lying in the respective climate zone when restricting to regions of high model consensus; columns 2 to 5: as upper part but for high consensus regions only.

Inter. RMS change	% highcons	% pos	% neg	median pos [mm]	median neg [mm]
global	22.6	54.1	45.9	7.02	-5.75
equatorial	14.6	50.0	50.0	8.14	-6.42
arid	31.3	53.9	46.2	5.40	-3.92
temperate	24.8	56.3	43.7	8.02	-7.12
polar	16.1	56.6	43.4	7.98	-6.50
high consensus	% area				
global	0.0	77.8	22.2	13.49	-23.79
equatorial	14.2	75.6	24.4	16.28	-25.94
arid	49.7	82.3	17.7	9.48	-15.31
temperate	17.6	76.3	23.7	19.37	-26.95
polar	18.6	68.6	31.4	18.50	-31.69

Table S7. As table S4, but for the linear trend and for both, CMIP6 and CMIP5. Upper part, column 1: the percentage of land area where > 75% of the models agree on the sign of the change (high consensus); columns 2 and 3: area percentage of positive and negative trends; columns 4 and 5: median trend calculated separately over the positive and negative areas. Lower part, column 1: percentage of land area lying in the respective climate zone when restricting to regions of high model consensus; columns 2 to 5: as upper part but for high consensus regions only.

Trend CMIP6	% highcons	% pos	% neg	median pos [mm/yr]	median neg [mm/yr]
global	47.2	42.9	57.1	0.42	-0.42
equatorial	57.3	49.9	50.1	0.60	-1.73
arid	47.3	54.0	46.0	0.31	-0.30
temperate	45.9	35.6	64.4	0.42	-0.44
polar	39.3	26.5	73.5	0.37	-0.40
high consensus	% area				
global	100.0	30.1	70.0	1.01	-0.89
equatorial	26.4	34.1	65.9	1.13	-2.76
arid	36.2	44.0	56.1	0.90	-0.58
temperate	15.6	15.5	84.5	1.10	-1.00
polar	21.9	12.6	87.4	1.01	-0.71
Trend CMIP5	% highcons	% pos	% neg	median pos [mm/yr]	median neg [mm/yr]
global	34.6	40.8	59.2	0.18	-0.36
equatorial	39.7	46.5	53.5	0.27	-0.54
arid	33.1	53.3	46.7	0.12	-0.27
temperate	43.9	27.3	72.7	0.28	-0.35
polar	26.8	27.4	72.7	0.30	-0.44
high consensus	% area				
global	100.0	24.0	76.1	0.62	-0.91
equatorial	24.4	33.0	67.0	0.62	-1.19
arid	34.6	31.6	68.4	0.51	-0.73
temperate	20.5	14.8	85.2	0.68	-0.80
polar	20.6	9.4	90.6	0.67	-1.02

5. Averaging phases

Due to the fact that phases of the annual cycle are not normally distributed, the calculation of statistical measures like mean, median, or standard deviation is not straight forward. To overcome this, we apply an iterative algorithm that is illustrated in Figure S4. The upper panel shows the histogram of an artificial data set of 13 phase values (in month of the year, i.e., 0, . . . , 1 = January, 1, . . . , 2 = February, and so on) for which the mean, median, and standard deviation shall be calculated. The arithmetic mean of the phase values is shown as a vertical red line, and corresponding numbers for median, mean and standard deviation are given on the right of the panel. The value 6.8462 (i.e., July) is obviously not the correct mean phase, which should be between December and January. Now we iteratively shift all phase values smaller or equal than k months (with $k = 1, \dots, 6$) by 12 months and calculate median, mean, and standard deviation again (panels 2 – 4 in Figure S4). We define the mean and median where the standard deviation of the (shifted) sample gets minimal as the best estimates for the statistical measures (panel 4). If the computed mean or median from the shifted sample happens to be larger than 12, we subtract 12, in order to keep the range of values from January to December. In this example, the mean is 0.3077 (i.e., January, vertical yellow line in panel 4), and the median is 12 (i.e. December), with a standard deviation of 1.6525 months.

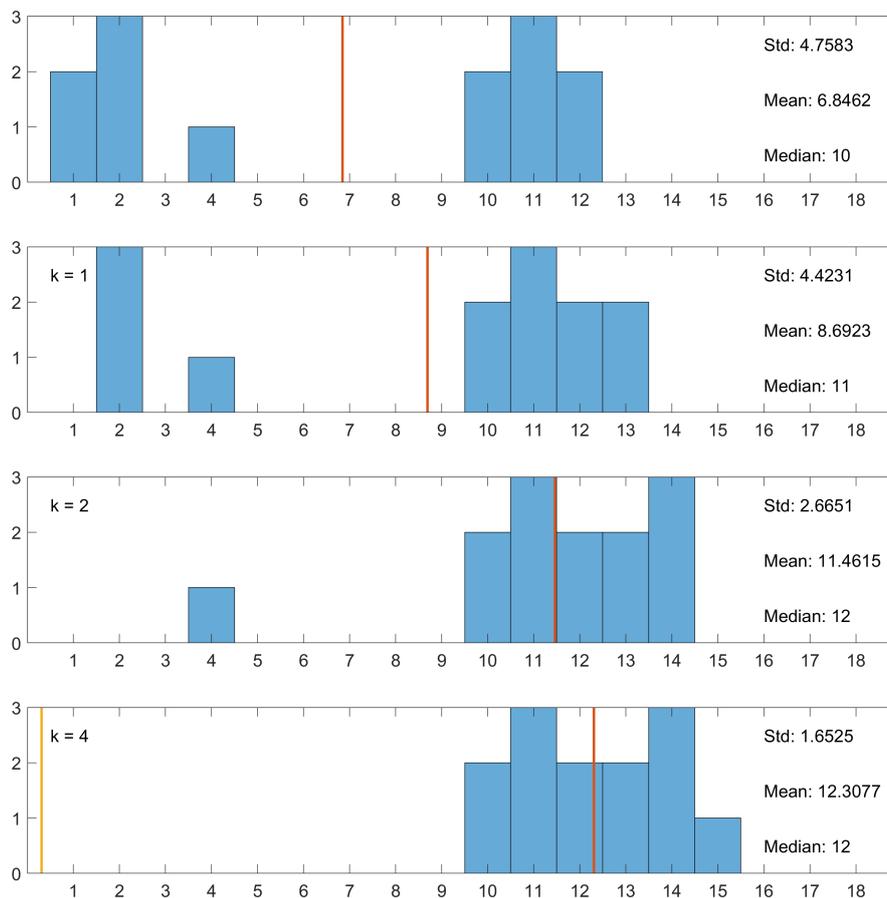


Figure S4. Illustration for the computation of the mean and median phase.

6. Interference of mrso and snw phases

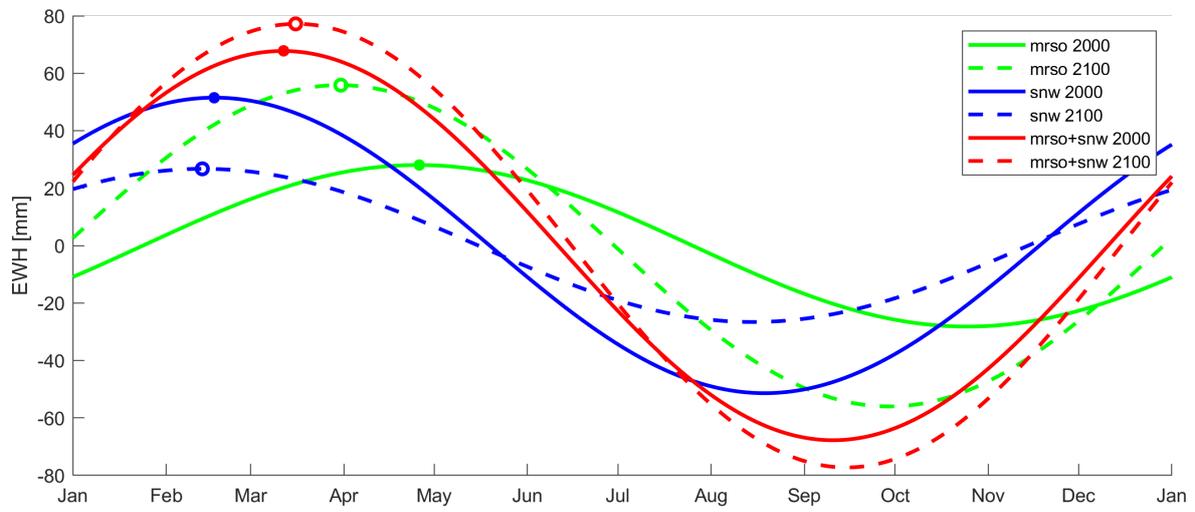


Figure S5. Example for a time series where the time shift from 2000 to 2100 is negative for mrso and snw, but positive for mTWS (i.e. the sum of mrso and snw). The location is 75.0°E and 57.0°N (Russia).

7. Ranking of ensemble members

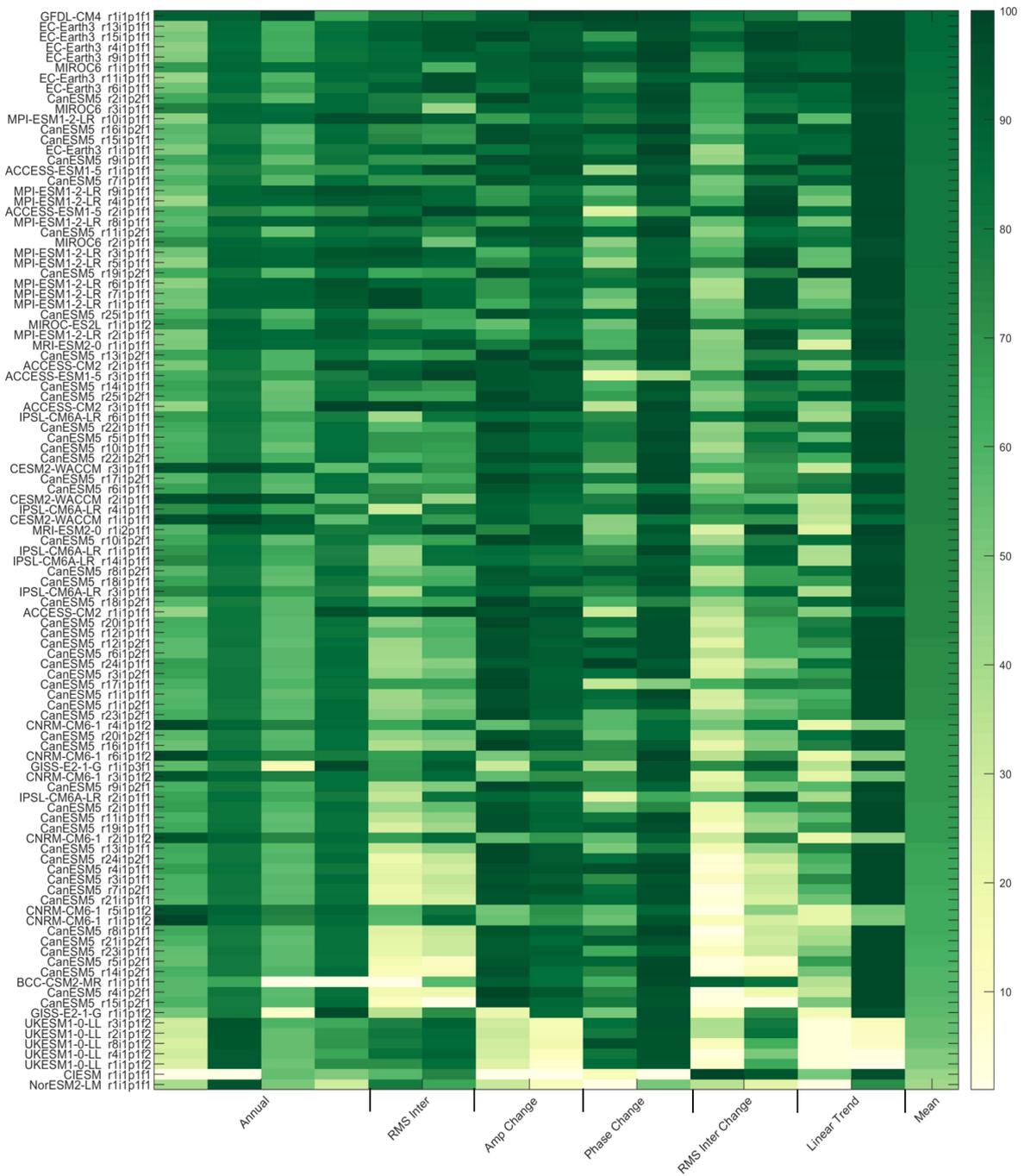


Figure S6. Ranking of the ensemble members according to the classes assigned with pattern correlation (odd columns) and RMSD of ECDF (even columns) of annual cycle and interannual RMS with GRACE (columns 1–6) and amplitude change, phase change, interannual RMS change and linear trend with the MMed (columns 7–14).

8. Detectability of annual cycle changes for a specific CMIP6 model run

In Figure S7 we compare the amplitude and phase change patterns of the r1i1p1f1 run of the GFDL-CM4 model (selected as a representative model run in Section 3.4 of main text) to (1) the current GRACE accuracy and (2) a five times smaller accuracy of a possible NGGM (cf. Section 3.3 of main text). The area percentages of detectable changes are 40% (amplitude) and 27% (phase) for the GRACE accuracy and 77% (amplitude) and 68% (phase) for the NGGM accuracy.

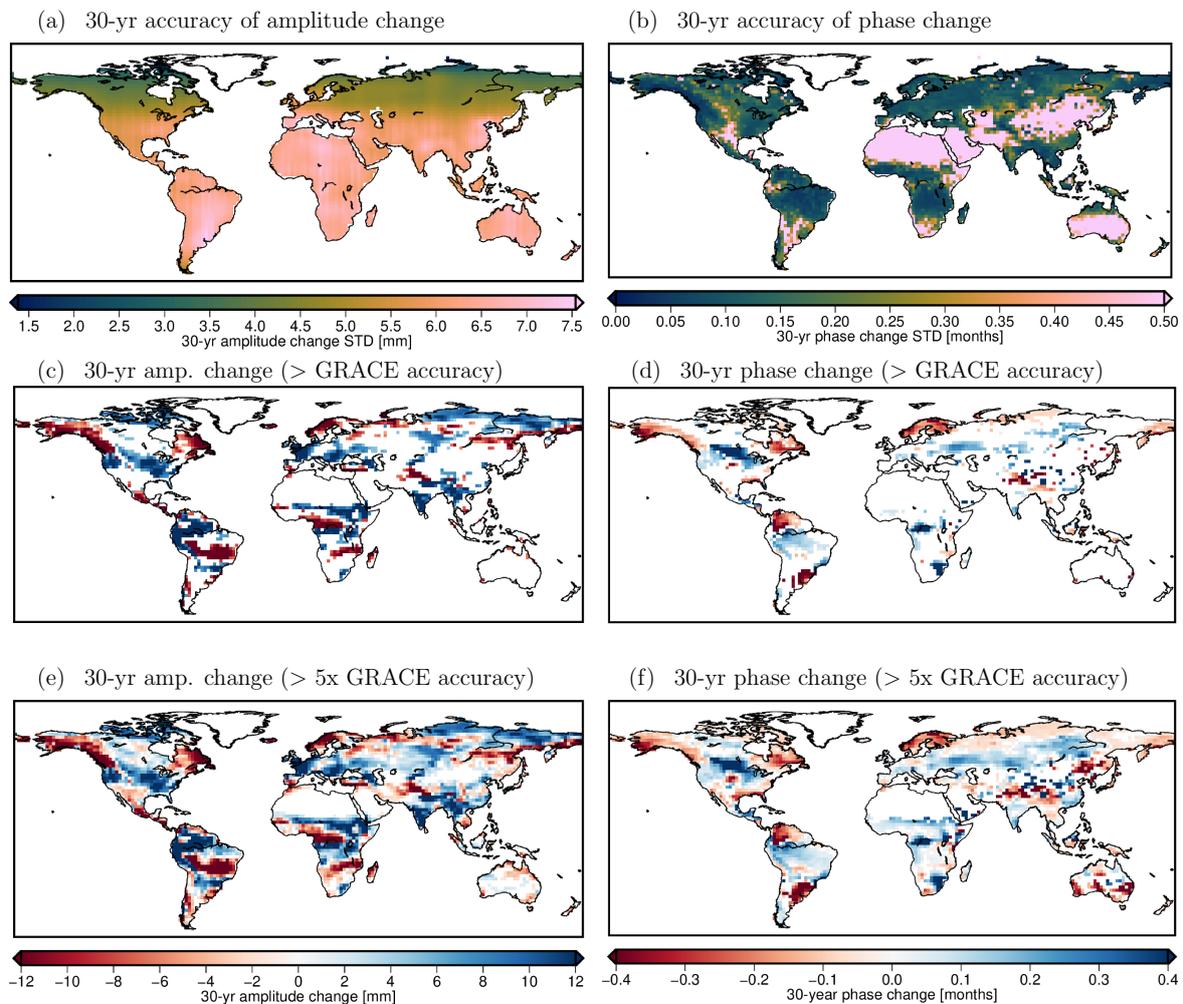


Figure S7. (a) standard deviation of GRACE TWS annual amplitude change over 30 years. (b) standard deviation of GRACE TWS phase change of annual cycle over 30 years. (c) mTWS annual amplitude change of the GFDL-CM4 r1i1p1f1 over 30 years that exceeds the GRACE accuracy (given in (a)). (d) same as (c) but for phase change. (e,f) same as (c,d) but assuming the standard deviation of GRACE (given in (a) and (b)) being five times smaller.

References

1. Deggim, S.; Eicker, A.; Schawohl, L.; Gerdener, H.; Schulze, K.; Engels, O.; Kusche, J.; Saraswati, A.T.; van Dam, T.; Ellenbeck, L.; Dettmering, D.; Schwatke, C.; Mayr, S.; Klein, I.; Longuevergne, L. RECOG RL01: Correcting GRACE total water storage estimates for global lakes/reservoirs and earthquakes. *Earth System Science Data Discussions* **2020**, pp. 1–30. Publisher: Copernicus GmbH, doi:https://doi.org/10.5194/essd-2020-256.
2. Kusche, J. Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. *Journal of Geodesy* **2007**, *81*, 733–749. doi:10.1007/s00190-007-0143-3.

3. Liepert, B.G.; Lo, F. CMIP5 update of 'Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models'. *Environmental Research Letters* **2013**, *8*, 029401. doi:10.1088/1748-9326/8/2/029401.
4. Müller Schmied, H.; Cáceres, D.; Eisner, S.; Flörke, M.; Herbert, C.; Niemann, C.; Peiris, T.A.; Popat, E.; Portmann, F.T.; Reinecke, R.; Schumacher, M.; Shadkam, S.; Telteu, C.E.; Trautmann, T.; Döll, P. The global water resources and use model WaterGAP v2.2d: Model description and evaluation. *Geoscientific Model Development Discussions* **2020**, pp. 1–69. Publisher: Copernicus GmbH, doi:https://doi.org/10.5194/gmd-2020-225.
5. Alsdorf, D.; Rodriguez, E.; Morrow, R.; Mognard, N.; Lambin, J.; Vaze, P.; Lafon, T. THE SURFACE WATER AND OCEAN TOPOGRAPHY (SWOT) MISSION **2010**.
6. GLIMS.; NSIDC. Global Land Ice Measurements from Space glacier database. Compiled and made available by the international GLIMS community and the National Snow and Ice Data Center, Boulder CO, U.S.A. **2005, updated 2020**. Publisher: NSIDC, doi:10.7265/N5V98602.
7. Jensen, L.; Eicker, A.; Dobslaw, H.; Stacke, T.; Humphrey, V. Long-Term Wetting and Drying Trends in Land Water Storage Derived From GRACE and CMIP5 Models. *Journal of Geophysical Research: Atmospheres* **2019**, *124*, 9808–9823. doi:10.1029/2018JD029989.
8. Eicker, A.; Jensen, L.; Wöhnke, V.; Dobslaw, H.; Kvas, A.; Mayer-Gürr, T.; Dill, R. Daily GRACE satellite data evaluate short-term hydro-meteorological fluxes from global atmospheric reanalyses. *Scientific Reports* **2020**, *10*, 4504. Number: 1 Publisher: Nature Publishing Group, doi:10.1038/s41598-020-61166-0.