Department of Applied Mathematics

Universitat d'Alacant
Universidad de Alicante

4th September 2020

Dear editor Zhenghui Xie,

Please find below a point-by-point response to the reviews, a list of all relevant changes made in the manuscript, and a marked-up manuscript version.

Yours sincerely,

David García-García, Ph. D.

Tel. +34 965 90 97 14 / Fax +34 965 90 97 07
Campus de Sant Vicent del Raspeig
Ap.99. E-03080 Alicante, Spain
e-mail: d.garcia@ua.es

Answer to Reviewer 1

We thank the reviewer for thorough reading and thoughtful comments and suggestions. A detailed discussion of the changes that we made in response to the reviewer's comments is given below. In what follows, we state the reviewer's comment in boldface, and describe our response in plain text. Text in the manuscript is represented in italics. The text that has been modified/included in the new version has been highlighted in red.

"Overall this is an excellent manuscript, presenting a new result about the Earth's ocean-land-atmosphere mass exchange, using a unique combination of satellite and reanalysis datasets, and a clear easy-to-follow methodology."

We appreciate the positive overall comment about the manuscript.

"The only major concern/question I have is this: the interbasin ocean transport N is a small residual of differencing large numbers. I see that each set of numbers is followed by a 95% confidence range, and I read without quite understanding that the confidence interval is computed by a bootstrap method on the data itself. I don't believe the re-analysis data have their own error estimates; I believe the GRACE data do but those did not seem to be used in the confidence interval estimation. I wonder whether estimating uncertainties in the transports by propagating uncertainties in the inputs would give intervals consistent with those of the bootstrap method. Upper bounds on the uncertainties in the inputs can be estimated, for example, by comparing UT-CSR mascons to JPL or GSFC mascons, by comparing ECMWF reanalysis to NCEP or another model's reanalyses, etc. I say this because the lack of correlation between the inter- annual transports and ANY index of ocean-atmosphere interaction (ENSO, SOI, etc) is suspicious."

The following changes have been included to address the issues raised by the referee:

1. **Bootstrap:** We have included an intuitive description of the bootstrap method for time series and a reference to a paper on bootstrap method for time series. Besides, we have provided extended details about how confidence intervals have been evaluated:

   The reported 95% confidence intervals and the correlation coefficients are evaluated using the stationary bootstrap scheme of Politis and Romano (1994) (with optimal block length selected according to Patton et al., 2009), and the percentile method. The intuition underlying the bootstrap is simple. Suppose that the observed time series $x_1, \ldots, x_n$ is a realization of the random vector $(X_1, \ldots, X_n)$ with joint distribution P_n and which is assumed to be part of a stationary stochastic process. Given $\mathbf{X}_n$, we first build and estimate $\hat{P}_n$ of P_n. Then B random vectors $(X_1^*, \ldots, X_n^*)$ are generated from $\hat{P}_n$. If $\hat{P}_n$ is a good approximation of P_n, then the relation between $(X_1^*, \ldots, X_n^*)$ and $\hat{P}_n$ should well reproduce the relation between $(X_1, \ldots, X_n)$ and P_n (for an introduction of bootstrap methods for time series see Kreiss and Lahiri (2012) and the references therein). Here, the number of bootstrap replications was set to B=2000. In general, half length of the confidence interval can be very well approximated by twice the standard deviation of the sample mean estimated from the bootstrap replications. Prior to applying the bootstrap to a time series, least-squares estimated linear/quadratic trend and sinusoid with the most relevant frequencies are removed from it to meet the stationarity conditions of the method. In particular, each series

Answer to Reviewer 1

We thank the reviewer for thorough reading and thoughtful comments and suggestions. A detailed discussion of the changes that we made in response to the reviewer's comments is given below. In what follows, we state the reviewer's comment in boldface, and describe our response in plain text. Text in the manuscript is represented in italics. The text that has been modified/included in the new version has been highlighted in red.

"Overall this is an excellent manuscript, presenting a new result about the Earth's ocean-land-atmosphere mass exchange, using a unique combination of satellite and reanalysis datasets, and a clear easy-to-follow methodology."

We appreciate the positive overall comment about the manuscript.

"The only major concern/question I have is this: the interbasin ocean transport N is a small residual of differencing large numbers. I see that each set of numbers is followed by a 95% confidence range, and I read without quite understanding that the confidence interval is computed by a bootstrap method on the data itself. I don't believe the re-analysis data have their own error estimates; I believe the GRACE data do but those did not seem to be used in the confidence interval estimation. I wonder whether estimating uncertainties in the transports by propagating uncertainties in the inputs would give intervals consistent with those of the bootstrap method. Upper bounds on the uncertainties in the inputs can be estimated, for example, by comparing UT-CSR mascons to JPL or GSFC mascons, by comparing ECMWF reanalysis to NCEP or another model's reanalyses, etc. I say this because the lack of correlation between the inter- annual transports and ANY index of ocean-atmosphere interaction (ENSO, SOI, etc) is suspicious."

The following changes have been included to address the issues raised by the referee:

1. **Bootstrap:** We have included an intuitive description of the bootstrap method for time series and a reference to a paper on bootstrap method for time series. Besides, we have provided extended details about how confidence intervals have been evaluated:

   The reported 95% confidence intervals and the correlation coefficients are evaluated using the stationary bootstrap scheme of Politis and Romano (1994) (with optimal block length selected according to Patton et al., 2009), and the percentile method. The intuition underlying the bootstrap is simple. Suppose that the observed time series $x_1, \ldots, x_n$ is a realization of the random vector $(X_1, \ldots, X_n)$ with joint distribution P_n and which is assumed to be part of a stationary stochastic process. Given $\mathbf{X}_n$, we first build and estimate $\hat{P}_n$ of P_n. Then B random vectors $(X_1^*, \ldots, X_n^*)$ are generated from $\hat{P}_n$. If $\hat{P}_n$ is a good approximation of P_n, then the relation between $(X_1^*, \ldots, X_n^*)$ and $\hat{P}_n$ should well reproduce the relation between $(X_1, \ldots, X_n)$ and P_n (for an introduction of bootstrap methods for time series see Kreiss and Lahiri (2012) and the references therein). Here, the number of bootstrap replications was set to B=2000. In general, half length of the confidence interval can be very well approximated by twice the standard deviation of the sample mean estimated from the bootstrap replications. Prior to applying the bootstrap to a time series, least-squares estimated linear/quadratic trend and sinusoid with the most relevant frequencies are removed from it to meet the stationarity conditions of the method. In particular, each series *has been decomposed into trend, seasonal and residual components. The bootstrap is applied to the residual component producing bootstrap samples of the residuals. For the evaluation of confidence intervals for the different components of WT, the trend and seasonal terms are added back (to the bootstrap sample of the residuals) producing bootstrapped time series of the component of interest. These samples are then used for further analysis. As an illustration, for the WT N component we proceed as follows: (i) a model with linear, annual, and semiannual signals is fitted to the data. The fitted linear trend and annual and semiannual signals are subtracted from the original time series; (ii) the stationary bootstrap is then applied to the residuals producing 2000 bootstrap samples of the residuals; (iii) The estimated trend and seasonal components are added back to each bootstrap sample of the residuals obtaining an ensemble of 2000 bootstrapped time series for the N component; (iv) these 2000 bootstrapped time series are used to obtain 95% confidence intervals for the mean fluxes (average of N over the 14 year period of study) and for the amplitude and phase of the annual component using the percentile method. For the mean fluxes, the average of N for each of the 2000 bootstrapped time series was first evaluated and then the 0.025 and 0.975 percentiles of these 2000 averages were reported as 95% confidence interval. For the study of the climatology, a linear trend model with annual and semiannual components was fitted to the 2000 bootstrapped time series producing corresponding estimates of the annual amplitude and phase. The 0.025 and 0.975 percentiles of these estimates were reported as 95% confidence intervals. In order to study the robustness of the results with respect to the model choice, the analysis is rerun using 11 alternative models obtained considering different forms for the trend component (quadratic or constant) and including higher frequencies in the harmonic regression (up to 5). The results are robust. The relative difference with respect to the reported values is smaller than 1.2% for point estimates and smaller than 3.3% for the extremes of the 95% confidence intervals.*

2. **Confidence intervals of the correlation coefficients**. More details are provided:

   Note that for the study of correlation the bootstrap was applied to the bivariate time series of the residuals of the two variables of interest producing an ensemble of 2000 bivariate time series of residuals. For each bivariate time series of residuals the correlation between the two components of the series was first evaluated. The average and the 0.025 and 0.975 percentiles of these 2000 estimates were reported as point estimate and confidence limits for the correlation between the two variables of interest (correlation between residual components is used to avoid spurious correlation).

3. **Bootstrap Vs Error propagation:** The confidence intervals estimated from bootstrap have been compared to those estimated from error propagation of the mascon. As CSR mascon solution does not provide such error estimates, we have used the JPL mascon solution for the comparison. An explanation of why bootstrap confidence intervals contains, as expected, the error propagation confidence interval has been also provided. In the description of the bootstrap method we have included the following text:

   As an independent check of the bootstrap, confidence intervals for the mean value of N have been also evaluated by propagating the error estimate in GRACE data (using the JPL GRACE mascon solution for which error estimates are available). The resulting intervals were *consistent with those of the bootstrap method. In particular (see Section 4 for details), we show that in all cases the bootstrap intervals contain the intervals obtained from error propagation. In this respect, the CI_{95} from bootstrap analysis can be considered a conservative estimate. This should be expected, since the residual component underlying the bootstrap approach includes measurement errors and other type of errors (related, for example, with the estimate of the trend and seasonal terms). As a result, the uncertainties in the transports estimated by the bootstrap should be larger than the corresponding uncertainties estimated by error propagation.*

We have included a new section 4, entitled “Comparison with other datasets”, which includes the comparison between error propagation and bootstrap confidence intervals for the *N* component estimated from JPL data:

CSR GRACE mascon solution is replaced by the JPL GRACE mascon solution provided by the Jet Propulsion Laboratory/NASA (Watkins et al., 2015; Wiese et al., 2019). Similarly to CSR data, JPL are corrected for GIA effects, C_{20} Stoke coefficients are replaced by a solution from SLR, and data are reduced to 1º regular grids from 0.5º regular grids. Besides, we have applied the degree-0 Stoke coefficients correction. However, CSR and JPL mascon solutions are not directly comparable. The main reason is that an estimate of degree-1 coefficients has been added to JPL mascon solutions, and the GAD product has not been added back. The corrections applied by JPL are not supplied separately and we cannot do/undo any of the corrections to process JPL data as we did with CSR data. In particular, the GAD product is not available for JPL. In any case, the JPL solution is useful here since it provides an error estimate of the mascon solution that can be propagated to obtain confidence intervals of N, which are independent from those estimated with the bootstrap analysis. Table 2 shows the CI_{95} of the mean values of the N component for different ocean basin estimated from error propagation and bootstrap analysis. It is observed that in all cases the CI_{95} from error propagation are included in those from bootstrap analysis, meaning that the latter are a conservative estimate of the error. JPL propagated error can be expected to be similar to that propagated from CSR error estimates (which are not available), and then we can assume that the reported CI_{95} for N calculated from CSR data are a conservative estimate. Besides, comparing Tables 1 and 2, it is observed that the mean values of N are quite similar and that the CI_{95} largely overlap. Regarding to the time variability, the values of the N component from CSR and JPL mascon solutions show Pearson correlation coefficients greater than 0.85 (p-value $< 10^{-3}$), except for the Atlantic (0.70). Thus, despite the different processing of CSR and JPL data, the reported analysis for the N component is robust with respect to the choice of GRACE datasets.

Table 2. Mean net WT from JPL mascon for different ocean basins according to Equation 2 . CI_{95} are estimated as propagation of mascon errors provided by JPL, and from bootstrap analysis. Units are Gt/month.

		Mean (CI_{95} from error propagation)	Mean (CI_{95} from bootstrap)
Outflows	Pacific	1182 (1143, 1220)	1182 (1062, 1306)
	Arctic	735 (713, 757)	735 (711,761)
	Pacific + Arctic	1917 (1872, 1961)	1917 (1806, 2036)
Inflows	AIA	1183 (1092, 1274)	1183 (1077, 1282)
	Atlantic	919 (866, 972)	919 (845, 985)
	Indian	999 (980, 1018)	999 (928, 1067)
	Atlantic + Indian	1918 (1862, 1974)	1918 (1838, 2003)

4. **Other P and E datasets:** According to ERA5 documentation, there exists error estimates. Unfortunately, they are not available for the general public as us. In any case, we have included new computations with several *P* and *E* datasets. It is included in the second point of the new section "Comparison with other datasets":

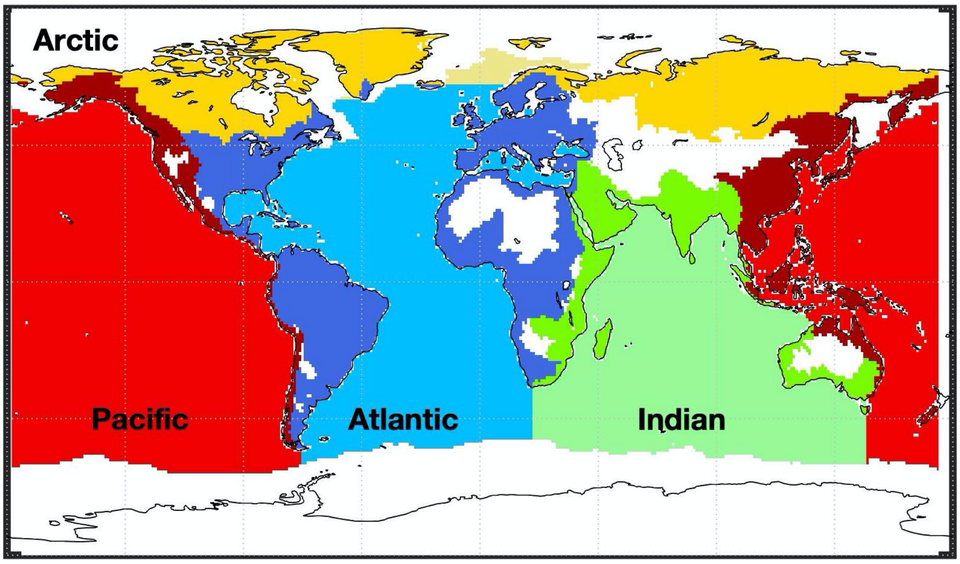
   ERA5 P and E data are replaced by several datasets for comparison purposes. The objective is not to be exhaustive in the selection, but rather to show that the reported features of the N component are quite robust with respect to the choice of the P and E datasets. The data sets considered are:

   (i) Continental P from GPCC (Schneider et al., 2011), GPCP (Adler et al., 2018), CMAP (Xie and Arkin, 1997), UDel (Willmott and Matsuura, 2001), and GLDAS/Noah (Rodell et al., 2004; Beaudoing and Rodell, 2016).
   (ii) Ocean P from GPCP and CMAP.
   (iii) Continental E from GLEAM (Miralles et al., 2011; Martens et al., 2017) and GLDAS/Noah.
   (iv) Ocean E from OAFlux (Yu et al., 2008) and HOAPS/CM SAF (Schulz et al., 2009).

   The Pacific outflow is estimated with the 162 possible combinations of P and E, including ERA5. The time period is 2003-2016, except for HOAPS/CM SAF and GPCP, which span from 2003 to 12/2014 and 10/2015, respectively. The degree-0 corrections in GRACE data is made for each combination. Note that only ERA5 includes P and E for both continents and oceans. All grids have been homogenized to 1° regular grids. The main concern here is the heterogeneity of the spatial coverage among datasets. To make the results comparable among datasets, the computations are restricted to the common grid points, which do not cover the entire Earth (Figure 8a). However, in spite of the fact that due to the partial coverage the principle of water mass conservation is not accomplished, the Pacific outflow obtained in the common grid points from ERA5 (black line in Figure 8b) is quite in agreement with the same signal obtained with global coverage (red line in Figure 3 which is also reported as red line in Figure 8b). The Pearson correlation coefficient between the two signals is 0.994 (p-values < 10^{-3}) with an average difference around 50 Gt/month. In general, the Pacific outflows estimated from all the P and E dataset combinations show qualitatively the same signal than the one reported in Figure 3. For each of the 162 estimates of the Pacific outflows corresponding to the possible P and E dataset combinations, we evaluated the average outflow (over the period of study), which is 968 Gt/month (STD: 489), and the correlation with the Pacific outflows in Figure 3, which is 0.82 (STD: 0.06; p-values < 10^{-3}).

   These experiments show that the reported net WT are physically consistent among datasets, at least qualitatively."

a) Common spatial coverage to all P and E datasets



b) (-1) × $N_{Pacific}$

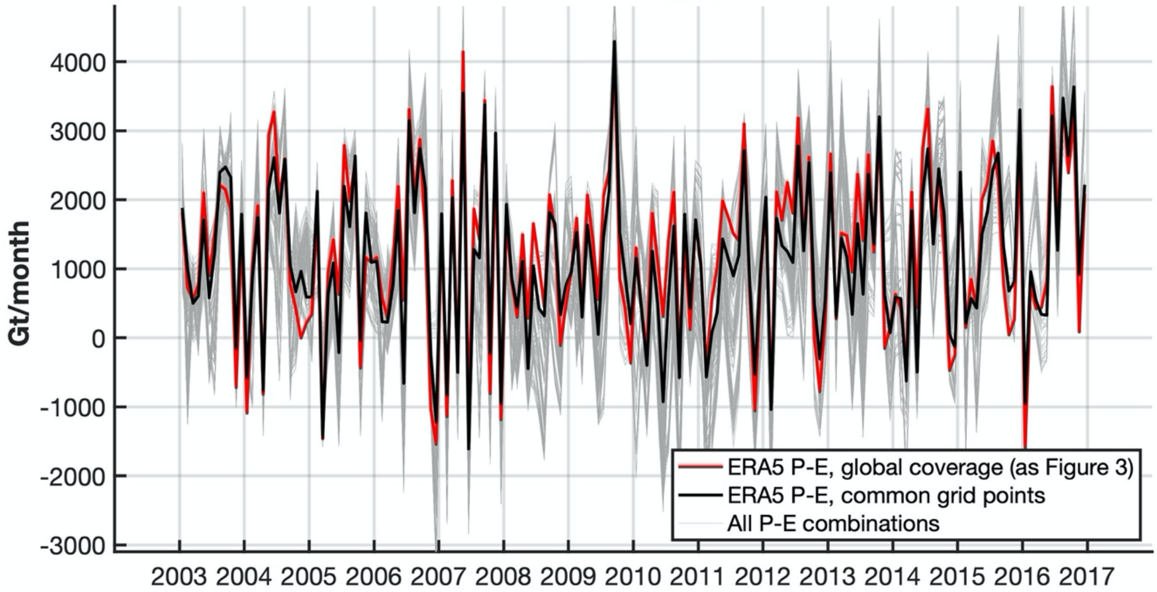


Figure 8. Monthly time series of (the opposite of) the Pacific outflow estimated from 162 combinations of *P* and *E* datasets. a) Spatial coverage common to all datasets. b) Pacific outflows: Gray thin curves are the 162 Pacific outflows estimated in the common grid points to all datasets (no global coverage); black and red curves are based on ERA5 *P* and *E* and are obtained using either only the grid points common to all datasets (black curve) or global coverage (red curve). Note that the red curve is the same as in Figure 3.

5. **Lack of correlation.** We have included a discussion on the lack of correlation between the inter-annual transports and the indices of ocean-atmosphere interaction. In particular we propose the two following explanations:

   "To explore this lack of correlation, we have estimated the correlation coefficient between each climatic index and each WT component (Figure 7b).

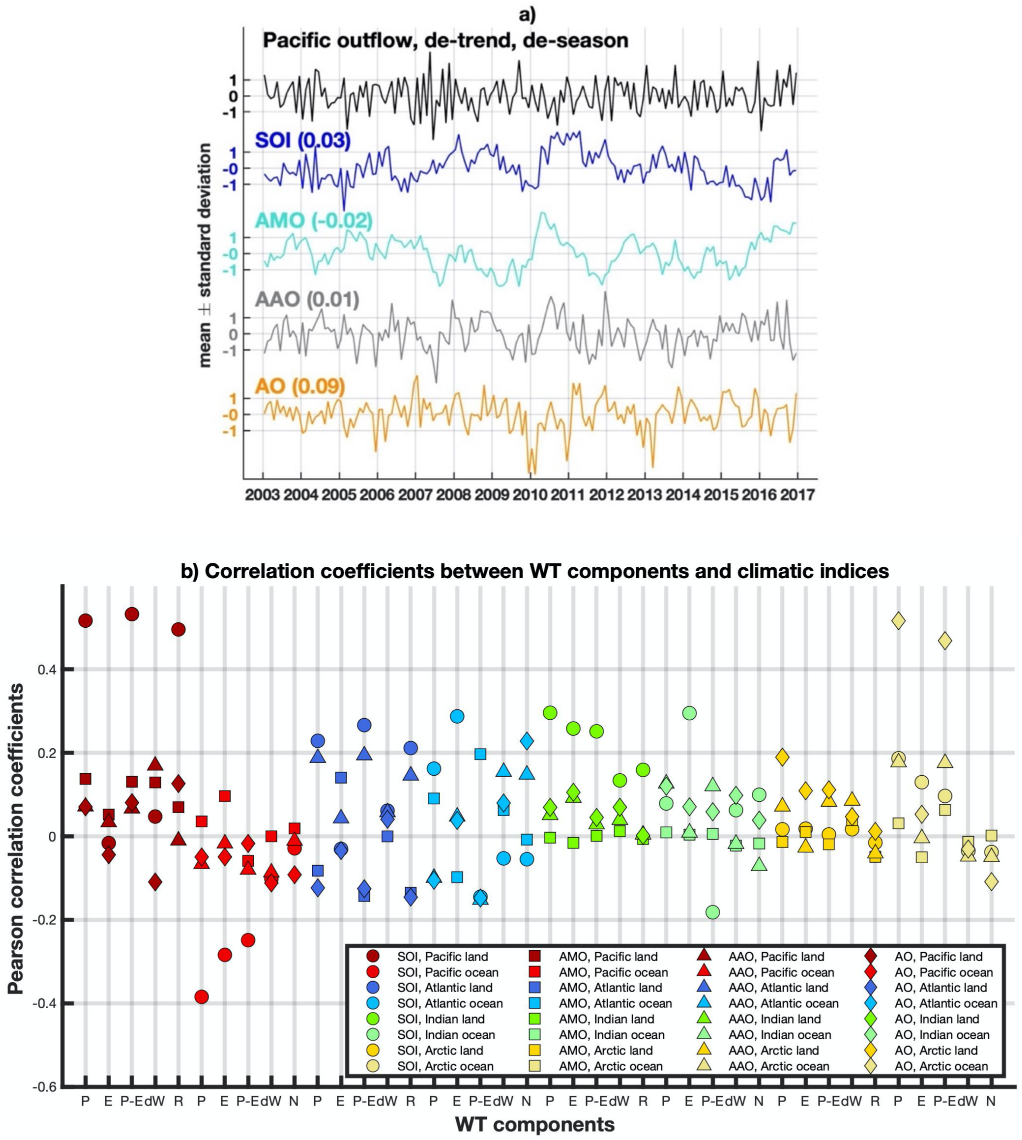


Figure 7. Pacific outflow and climatic indices for ENSO, AMO, AO, and AAO. a) Time series of Pacific outflow is de-trend and de-season. All time series are normalized to have unit variance. Values in the parenthesis are the correlation coefficient between the corresponding climatic index and the Pacific outflow. b) Correlation coefficients between de-trend and de-season WT components of different regions and the climatic indices.

All of them are lower than 0.3 except for 6 cases in 2 regions. In the Arctic, P and P−E in the drainage basins of the Arctic show a correlation of ~0.5 with the AO. This correlation is natural since that is the area of influence of the AO. The other region is the Pacific, where, as expected, the SOI shows a correlation around 0.5 with P, P−E, and R in the drainage basins, and around −0.4 with P in the ocean. However, this individual correlation does not extend to the Pacific outflow. In order to understand why this is the case, it is convenient to express the N component of the water transport as a function of (P-E) and dW. According to Equations 1 and 2 we have:

$$N = -(P-E)_{ocean} - R + dW_{ocean} = \underbrace{-(P-E)_{ocean}}_{X_1} \underbrace{-(P-E)_{land}}_{X_2} + \underbrace{dW_{land}}_{X_3} + \underbrace{dW_{ocean}}_{X_4} \qquad (3)$$

It can be shown that the correlation between N and a given index can be express as follows

$$corr(N, Index) = \sum_{i=1}^{4} corr(X_i, Index) \cdot \frac{std(X_i)}{std(N)}, \qquad (4)$$

where corr denotes the correlation coefficient, and std stands for standard deviation. As shown in equation (4), the correlation between N and a given index is a linear combination of the correlation between each component and the index. The coefficients of the linear combination std(X_i)/std(N) are proportional to the standard deviation of each component. The components of equation (4) for the Pacific outflow and the SOI index are shown in Table 3. Despite the fact that some of the individual component exhibits significant correlation with SOI (in particular P−E in land and ocean) when combined with the corresponding coefficients their effects are canceled out yielding to a negligible correlation between water transport and SOI (below 0.03 in magnitude).

Another possible reason for the lack of correlation resides in the definition of the studied regions, for which the presence of subregions with positive and negative influence of an index results in an overall negligible/attenuated influence of the index in the overall region. For example, a positive phase of the AMO is related to an increase of P in western Europe (Sutton and Hodson, 2005), and the Sahel (Folland et al., 1986; Knight et al., 2006; Zhang and Delworth, 2006; Ting et al., 2009), but to a decrease of P in the U.S. (Enfield et al., 2001; Sutton and Hodson, 2005), and northeast Brazil (Knight et al., 2006; Zhang and Delworth, 2006). All these regions are included in the Atlantic drainage basin, and then the influence of a positive phase of the AMO is attenuated."

Table 3. Correlation coefficients between SOI and de-trend and de-season WT components involved to estimate the Pacific outflow according to Equations 3 and 4.

	$std(X_i)$ (Stand. Deviation)	$corr(X_i, SOI)$ (Correlation between X_i with SOI)	$\frac{std(X_i)}{std(N)}$ (Coefficients)	$corr(X_i, SOI) \cdot \frac{std(X_i)}{std(N)}$ (Correlation ·Coefficient)
$X_1 = -(P-E)_{ocean}$	605	0.25	0.57	0.14
$X_2 = -(P-E)_{land}$	212	-0.53	0.20	-0.11
$X_3 = dW_{land}$	96	0.048	0.09	0.004
$X_4 = dW_{ocean}$	711	-0.10	0.67	-0.07
Corr(N,SOI)				-0.03

Note that table 3 provides also some insights about the causes of the interannual variability of Pacific Ocean outflow. The largest standard deviation of $P-E$ and dW in the ocean suggests that these two components might drive the interannual variability of the Pacific Ocean outflow. This is confirmed by a correlation analysis. The correlation between N and the $(P-E)_{Ocean}$ is -0.70. The correlation between N and the dW_{ocean} is 0.84. The correlation of N with the corresponding land components is below 0.18. In all cases, prior to the evaluation of the correlation the corresponding time series have been de-trend and de-season.

Now addressing some details:

(1) Figure 1: I would have liked to see a row with P-E- R next to the row for dW in Figure 1.

Figure 1 probably means Figure 2. Including *P-E-R*, in our opinion, is not very useful since, by definition of *R*, *P-E-R* will perfectly match *dW*. The comparison would be interesting with an independent dataset of R.

(2) Figures 1 and 3: I am sure the authors know better smoothers than the running mean (Hanning, Kaiser, etc). I recommend they use one.

We have replaced the running mean by a low pass filter defined by a Hann function of 24 months (the resulting smoothed curve is quite in agreement with the one previously obtained by running mean smoothing)

(3) Line 27: Clark reference missing. Recheck all your references, I did not do an exhaustive check.

Thank you. We have checked all the references.

(4) Line 93: tectonic signals in the gravity field do not 'masquerader as mascons'. Mascons are a simple mathematical representation of the gravity field with a physical interpretation. Tectonics "would be incorrectly interpreted as water mass flux"

Thank you. It is better expressed in this way. We have re-written the sentence: *"Any other non-surficial effect such as long-term tectonics would be incorrectly interpreted as water mass fluxes…"*

(5) Lines 124 et seq: see my concern above. A physical interpretation of this mathematical approach to confidence intervals would be useful.

We have extended the description of the bootstrap - see point 3 (Bootstrap Vs Error propagation) in page 1 of this response.

(6) Line 164: and loses 'to the atmosphere' 879 Gt/month. . .

The sentence has been re-written:
"On average, the Atlantic Ocean receives 926 Gt/month (CI_{95}=[876, 980]; or 0.36 Sv) of salty water, and loses to the atmosphere 879 Gt/month (CI_{95}=[828, 930]) via P−E+R."

(7) Line 188: I think 'The Atlantic/Arctic inflow 'mirrors this behaviour' is a better phrase in English.

Thank you. We have re-written the sentence: *"The Atlantic/Arctic inflow mirrors this behaviour."*

(8) Somewhere: W. T. Liu et al (GRL 2006, on South American water balance) did a similar estimation of water flux between an ocean basin and the land, without using any numerical model data.

Thank you. We agree that it is a pertinent reference. We have included it in the last paragraph of the introduction, which now is:

"In this work we propose a new methodology devised to estimate the net WT through the boundaries of a given oceanic region. A defining feature of the proposed approach is the use of the time-variable gravity data from the GRACE (Gravity Recovery and Climate Experiment) satellite mission to estimate the change of water content. We apply the methodology, in conjunction with conventional meteorological data of general hydrologic budget schemes, to estimate the time evolution over the period 2003-2016 of the net WT and exchanges among the four major ocean basins – namely Pacific, Atlantic, Indian, and Arctic. We analyse and report our results of the seasonal climatology as well as the interannual variability of WT. Such information, not available previously, is of valuable importance. For example, in closed regions, net WT through the boundaries on the surface must be counteracted by moisture fluxes through the same boundaries in the atmosphere to match GRACE measurements. Such approach has been successfully applied to study the hydrological cycle of South America (Liu et al., 2006). At ocean basin scale, knowledge about net WT not only would help elucidate the role of the oceans within the water cycle, but it will also impose restrictions on moisture advection in the atmosphere that would help to improve atmospheric models. On the other hand, ocean models usually deal with inflows and outflows of a given ocean region (Warren, 1983; Rahmstorf, 1996; Emile-Geay et al., 2003; de Vries and Weber, 2005; Dijkstra, 2007). Net WT estimates for such ocean region would be useful to impose constraints to the relationship between its inflows and outflows, which would improve the reliability of the models. Better models will improve our knowledge of the Earth's WT dynamics and its evolution in the future, which is critical in the present scenario of climate change."

(9) There are a few more minor language errors (lines 255, 267 and possibly others). Please go over the manuscript and clean up.

Done. Thank you.

Answer to Reviewer 2

We thank the reviewer for thorough reading and thoughtful comments and suggestions. A detailed discussion of the changes that we made in response to the reviewer's comments is given below. In what follows, we state the reviewer's comment in boldface, and describe our response in plain text. Text in the manuscript is represented in italics. The text that has been modified/included in the new version has been highlighted in red.

Specific comment 1:

I feel a strong motivation for estimating the lateral water transports from oceanic boundaries is lacking. It is not clear why we require measurements of lateral transports, given that the overall transport among different oceanic basin is known to a reasonable degree of certainty, such as by the studies noted by the authors in the introduction and discussion sections. The authors should write a concise and clear paragraph of why it is important to estimate the water fluxes through boundaries with this novel approach.

The last paragraph of the introduction has been re-written:

> *"In this work we propose a new methodology devised to estimate the net WT through the boundaries of a given oceanic region. A defining feature of the proposed approach is the use of the time-variable gravity data from the GRACE (Gravity Recovery and Climate Experiment) satellite mission to estimate the change of water content. We apply the methodology, in conjunction with conventional meteorological data of general hydrologic budget schemes, to estimate the time evolution over the period 2003-2016 of the net WT and exchanges among the four major ocean basins – namely Pacific, Atlantic, Indian, and Arctic. We analyse and report our results of the seasonal climatology as well as the interannual variability of WT. Such information, not available previously, is of valuable importance. For example, in closed regions, net WT through the boundaries on the surface must be counteracted by moisture fluxes through the same boundaries in the atmosphere to match GRACE measurements. Such approach has been successfully applied to study the hydrological cycle of South America (Liu et al., 2006). At ocean basin scale, knowledge about net WT not only would help elucidate the role of the oceans within the water cycle, but it will also impose restrictions on moisture advection in the atmosphere that would help to improve atmospheric models. On the other hand, ocean models usually deal with inflows and outflows of a given ocean region (Warren, 1983; Rahmstorf, 1996; Emile-Geay et al., 2003; de Vries and Weber, 2005; Dijkstra, 2007). Net WT estimates for such ocean region would be useful to impose constraints to the relationship between its inflows and outflows, which would improve the reliability of the models. Better models will improve our knowledge of the Earth's WT dynamics and its evolution in the future, which is critical in the present scenario of climate change."*

Specific comment 2:

In addition, the authors note in their introduction that their method improves upon the previous estimates. The literature on previous estimates and how (and how much) the new method improves upon them is not discussed in detail. In addition, a critical comparison of previous estimates of water transports and the estimates provided in this study is lacking.

As far as we are applying a new methodology, there are not many studies to compare with. The only two studies, up to our knowledge, doing something similar are discussed in the third paragraph of the section "Discussion and Conclusions":

"The results presented here are consistent with the well-known salinity asymmetry between the Pacific and Atlantic Oceans (Reid, 1953; Warren, 1983; Broecker et al., 1985; Zaucker et al., 1994; Rahmstorf, 1996; Emile-Geay et al., 2003; Lagerloef et al., 2008; Czaja, 2009; Reul, 2014). However, they are in contrast to previous GRACE-based studies where a simple seesaw WT between the Pacific and the Atlantic/Indian oceans was reported (Chambers and Willis, 2009; Wouters et al., 2014). In those studies, the P−E+R term in Equation 2 in both Pacific and Atlantic/Indian Oceans was approximated by that from the global ocean mean. However, the mean freshwater flux in the Pacific (1261 Gt/month) quite mis-matches that in the Atlantic/Indian Oceans (−1837 Gt/month), meaning that the approximation was quite poor and hence the N term was not properly estimated in these studies (see Appendix for further discussion)."

As stated in the text, in the Appendix we explain in detail why the proposed methodology overcomes some important limitations of previous approaches which will always show a seesaw of water transport, even if it does not exist.

Specific comment 3:

The authors observe that loss through E-P is much more in AIA as compared to the Pacific, even though the surface area is same. However, the reasons for such disparity is not discussed. Similarly, the potential reasons for other important results are not discussed. I hope to see some discussion on the results from this study.

The *P*, *E*, *P-E*, and *R* components are auxiliar in this study. However, we understand the reviewer's concern and we have added some more references for comparison purposes in the last paragraph of Section 3.1:

"Corresponding analyses have been performed for the Atlantic, Indian, and Arctic Oceans separately. The time evolution of the WT components in Eqs. 1 and 2 are shown in Figure 4, and a diagram of the water-mass fluxes is shown in Figure 5. On average, the Atlantic Ocean receives 926 Gt/month (CI_{95}=[876, 980]; or 0.36 Sv) of salty water, and loses to the atmosphere 879 Gt/month (CI_{95}=[828, 930]) via P−E+R. The latter is equivalent to a freshwater deficit of 0.34 Sv, which increases the near-surface salt concentration and enables water to sink in North Atlantic producing deep water. These values are close to the 0.13-0.32 Sv estimated from ocean models, as needed to keep

salinity balance in the Atlantic Ocean (Zaucker et al., 1994). Similarly, the Indian Ocean loses 957 Gt/month (CI_{95}=[894, 1022]) of freshwater that is restored by 991 Gt/month (CI_{95}=[907, 1073]) of salty water. The freshwater lost via P−E+R by the Atlantic and Indian Oceans goes to the Pacific (1261 Gt/month, CI_{95}=[1171, 1347]) and Arctic (730 Gt/month, CI_{95}=[712, 747]) Oceans, which return 1194 (CI_{95}=[1096, 1291]) and 723 (CI_{95}=[708, 739]) Gt/month of salty water through the ocean, respectively. Then, the Pacific presents a surplus of freshwater that reduces near-surface salt concentration, which prevents the formation of deep water. Together, the Pacific and Arctic Oceans supply 1917 Gt/month (CI_{95}=[1812, 2021]) of water to the Atlantic and Indian Oceans, where it is reincorporated into the water cycle via net E−P. As in previous studies (see Craig et al., 2017 for a synthesis), the freshwater lost in the Indian Ocean is similar to that in the Atlantic Ocean. In those studies, P−E+R is close to zero in the Pacific Ocean, producing a difference of 0.4 Sv between Atlantic and Pacific Oceans. In this study, P−E+R is 1261 Gt/month in the Pacific Ocean and the difference with the Atlantic increases to ~0.8 Sv. Some of these differences would be expected as far as the ocean basins are not defined in exactly the same way. On the other hand, the global R is 3781 Gt/month (or 3781 × 12 = 45368 km^3/year), close to the 41867 km^3/year reported by the Global Runoff Data Centre (GRDC, 2014). At basin scale, R is 16834 km^3/year in the Pacific, greater than the 11826 km^3/year reported by GRDC. In the Atlantic, Indian, and Arctic, R is 18228, 4479, and 5827 km^3/year, respectively, which is closer to the GRDC values: 20772, 5238, and 4080 km^3/year. Finally, according to the diagram in Figure 5, the water content in the atmosphere decreases 178 Gt/month (and it is gained by Earth's surface), but this amount is not realistic as discussed in Section 2 since it should increase a few Gt/month (Nilsson and Elgered, 2008). This value differs from the 188 Gt/month mentioned in Section 2 because the endorheic regions are not accounted here."

More importantly, we have extended our analysis to other datasets. The objective is to show that our main results concerning the *N component*, are not an artifact of CSR GRACE and ERA5 datasets. As a result, there is new section entitled "Comparison with other datasets":

"Equations 1 and 2 are applied to estimate the Pacific outflow using different datasets:

(1) CSR GRACE mascon solution is replaced by the JPL GRACE mascon solution provided by the Jet Propulsion Laboratory/NASA (Watkins et al., 2015; Wiese et al., 2019). Similarly to CSR data, JPL are corrected for GIA effects, C_{20} Stoke coefficients are replaced by a solution from SLR, and data are reduced to 1º regular grids from 0.5º regular grids. Besides, we have applied the degree-0 Stoke coefficients correction. However, CSR and JPL mascon solutions are not directly comparable. The main reason is that an estimate of degree-1 coefficients has been added to JPL mascon solutions, and the GAD product has not been added back. The corrections applied by JPL are not supplied separately and we cannot do/undo any of the corrections to process JPL data as we did with CSR data. In particular, the GAD product is not available for JPL. In any case, the JPL solution is useful here since it provides an error estimate of the mascon solution that can be propagated to obtain confidence intervals of N, which are independent from those estimated with the bootstrap analysis. Table 2 shows the CI_{95} of the mean values of the N component for different ocean basin estimated from error propagation and bootstrap analysis. It is observed that in all cases the CI_{95} from error

propagation are included in those from bootstrap analysis, meaning that the latter are a conservative estimate of the error. JPL propagated error can be expected to be similar to that propagated from CSR error estimates (which are not available), and then we can assume that the reported CI$_{95}$ for N calculated from CSR data are a conservative estimate. Besides, comparing Tables 1 and 2, it is observed that the mean values of N are quite similar and that the CI$_{95}$ largely overlap. Regarding to the time variability, the values of the N component from CSR and JPL mascon solutions show Pearson correlation coefficients greater than 0.85 (p-value $< 10^{-3}$), except for the Atlantic (0.70). Thus, despite the different processing of CSR and JPL data, the reported analysis for the N component is robust with respect to the choice of GRACE datasets.

Table 2. Mean net WT from JPL mascon for different ocean basins according to Equation 2 . CI$_{95}$ are estimated as propagation of mascon errors provided by JPL, and from bootstrap analysis. Units are Gt/month.

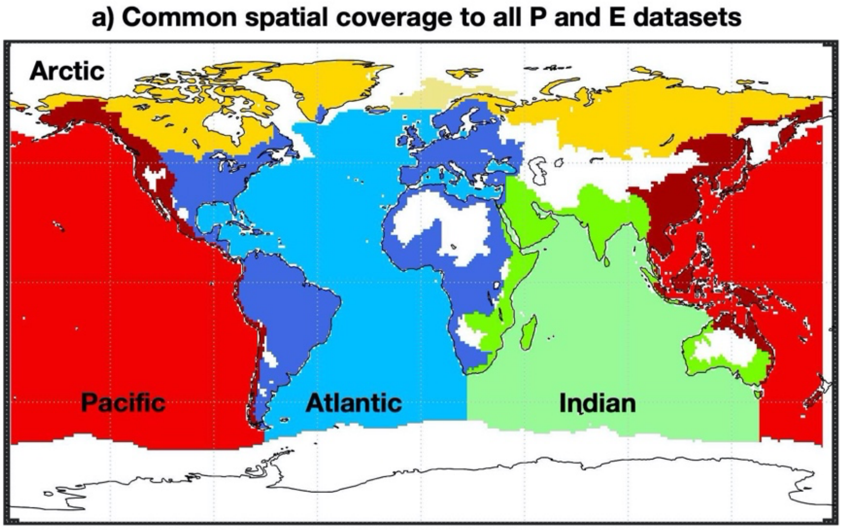
		Mean (CI$_{95}$ from error propagation)	Mean (CI$_{95}$ from bootstrap)
Outflows	Pacific	1182 (1143, 1220)	1182 (1062, 1306)
	Arctic	735 (713, 757)	735 (711,761)
	Pacific + Arctic	1917 (1872, 1961)	1917 (1806, 2036)
Inflows	AIA	1183 (1092, 1274)	1183 (1077, 1282)
	Atlantic	919 (866, 972)	919 (845, 985)
	Indian	999 (980, 1018)	999 (928, 1067)
	Atlantic + Indian	1918 (1862, 1974)	1918 (1838, 2003)

(2) ERA5 P and E data are replaced by several datasets for comparison purposes. The objective is not to be exhaustive in the selection, but rather to show that the reported features of the N component are quite robust with respect to the choice of the P and E datasets. The data sets considered are:

- *(i) Continental P from GPCC (Schneider et al., 2011), GPCP (Adler et al., 2018), CMAP (Xie and Arkin, 1997), UDel (Willmott and Matsuura, 2001), and GLDAS/Noah (Rodell et al., 2004; Beaudoing and Rodell, 2016).*
- *(ii) Ocean P from GPCP and CMAP.*
- *(iii) Continental E from GLEAM (Miralles et al., 2011; Martens et al., 2017) and GLDAS/Noah.*
- *(iv) Ocean E from OAFlux (Yu et al., 2008) and HOAPS/CM SAF (Schulz et al., 2009).*

The Pacific outflow is estimated with the 162 possible combinations of P and E, including ERA5. The time period is 2003-2016, except for HOAPS/CM SAF and GPCP, which span from 2003 to 12/2014 and 10/2015, respectively. The degree-0 corrections in GRACE data is made for each combination. Note that only ERA5 includes P and E for both continents and oceans. All grids have been homogenized to 1° regular grids. The main concern here is the heterogeneity of the spatial coverage among datasets. To make the results comparable among datasets, the computations are restricted to the common grid points, which do not cover the entire Earth (Figure 8a). However, in spite of the fact that due to the partial coverage the principle of water mass conservation is not accomplished, the Pacific outflow obtained in the common grid points from ERA5 (black line in Figure 8b) is quite in agreement with the same signal obtained with global coverage (red line in Figure 3 which is also reported as red line in Figure 8b). The Pearson correlation coefficient between the two signals is 0.994 *(p-values < 10^{-3}) with an average difference around 50 Gt/month. In general, the Pacific outflows estimated from all the P and E dataset combinations show qualitatively the same signal than the one reported in Figure 3. For each of the 162 estimates of the Pacific outflows corresponding to the possible P and E dataset combinations, we evaluated the average outflow (over the period of study), which is 968 Gt/month (STD: 489), and the correlation with the Pacific outflows in Figure 3, which is 0.82 (STD: 0.06; p-values < 10^{-3}).*

These experiments show that the reported net WT are physically consistent among datasets, at least qualitatively."



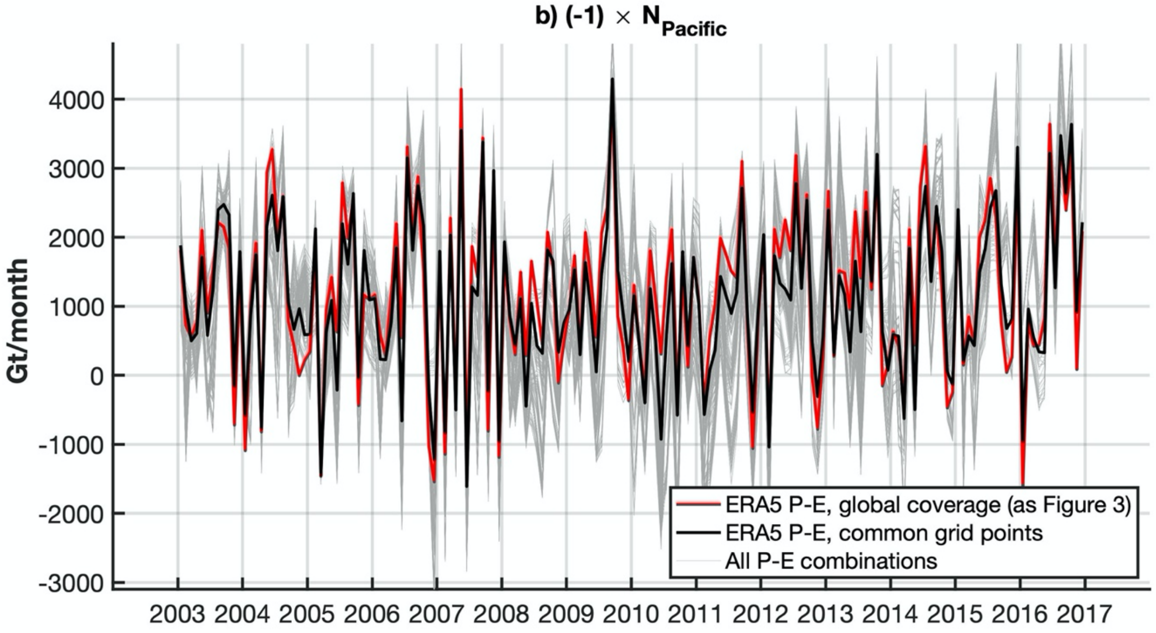


Figure 8. Monthly time series of (the opposite of) the Pacific outflow estimated from 162 combinations of *P* and *E* datasets. a) Spatial coverage common to all datasets. b) Pacific outflows: Gray thin curves are the 162 Pacific outflows estimated in the common grid points to all datasets (no global coverage); black and red curves are based on ERA5 *P* and *E* and are obtained using either only the grid points common to all datasets (black curve) or global coverage (red curve). Note that the red curve is the same as in Figure 3.

Specific comment 4:

Results section (L160-170). As of now, when I read the number and where the losses and gains take place, it is difficult for me to visualize the transfers among different basins. I strongly suggest the authors to present this information in terms of a multi-panel graph/map, where each map shows specific water transfer-related variable (such as N, R, etc.) with thick arrows giving the direction of transport, their color showing the magnitude of transport or we can just add text (number) inside the arrows to show magnitude.

Following the suggestion of the reviewer, we have included a new Figure with a diagram of the mean WT components to ease the reading:

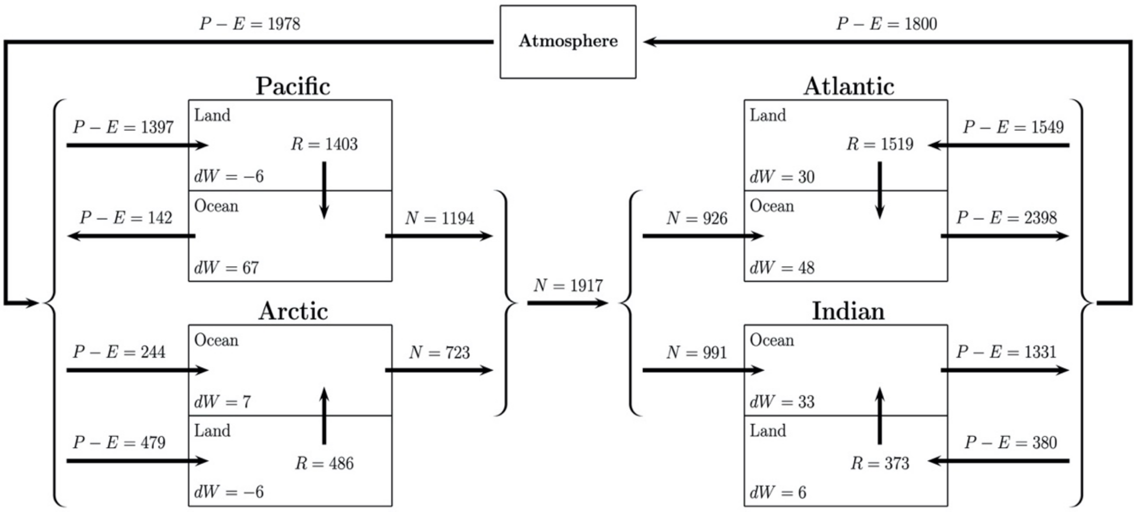


Figure 5. Diagram of the mean values of the WT of the studied regions. Units are Gt/month.

Specific comment 5:

L218: This is related to my comment#3. If none of the major indices shows strong correlation with the Pacific outflow, we do not have a confidence in what causes the interannual variability of Pacific Ocean outflow. Perhaps, more detailed insights from P and E series might help and/or some literature review on this might guide the authors in understanding the likely causes of the interannual variability of the Pacific Ocean outflow. Likewise, it would be useful to perform the same analysis on other basins for understanding their interannual variability of outflows.

Following the suggestion of the reviewer we have looked in more details at the *P* and *E* series. This has provided some insight in both the lack of the correlation of the Pacific outflow with the most important climatic indices and the interannual variability of the Pacific Ocean outflow. In particular:

We have extended the analysis about the lack of correlation and we give two possible explanations:

“ To explore this lack of correlation, we have estimated the correlation coefficient between each climatic index and each WT component (Figure 7b).

All of them are lower than 0.3 except for 6 cases in 2 regions. In the Arctic, P and P−E in the drainage basins of the Arctic show a correlation of ~0.5 with the AO. This correlation is natural since that is the area of influence of the AO. The other region is the Pacific, where, as expected, the SOI shows a correlation around 0.5 with P, P−E, and R in the drainage basins, and around -0.4 with P in the ocean. However, this individual correlation does not extend to the Pacific outflow. In order to understand why this is the case, it is convenient to express the N component of the water transport as a function of (P-E) and dW. According to Equations 1 and 2 we have:

$$N = -(P - E)_{ocean} - R + dW_{ocean} = \underbrace{-(P-E)_{ocean}}_{X_1} \underbrace{-(P-E)_{land}}_{X_2} + \underbrace{dW_{land}}_{X_3} + \underbrace{dW_{ocean}}_{X_4}. \qquad (3)$$

It can be shown that the correlation between N and a given index can be express as follows

$$corr(N, Index) = \sum_{i=1}^{4} corr(X_i, Index) \cdot \frac{std(X_i)}{std(N)}, \qquad (4)$$

where corr denotes the correlation coefficient, and std stands for standard deviation. As shown in equation (4), the correlation between N and a given index is a linear combination of the correlation between each component and the index. The coefficients of the linear combination std(Xi)/std(N) are proportional to the standard deviation of each component. The components of equation (4) for the Pacific outflow and the SOI index are shown in Table 3. Despite the fact that some of the individual component exhibits significant correlation with SOI (in particular P-E in land and ocean) when combined with the corresponding coefficients their effects are cancelled out yielding to a negligible correlation between water transport and SOI (below 0.03 in magnitude).

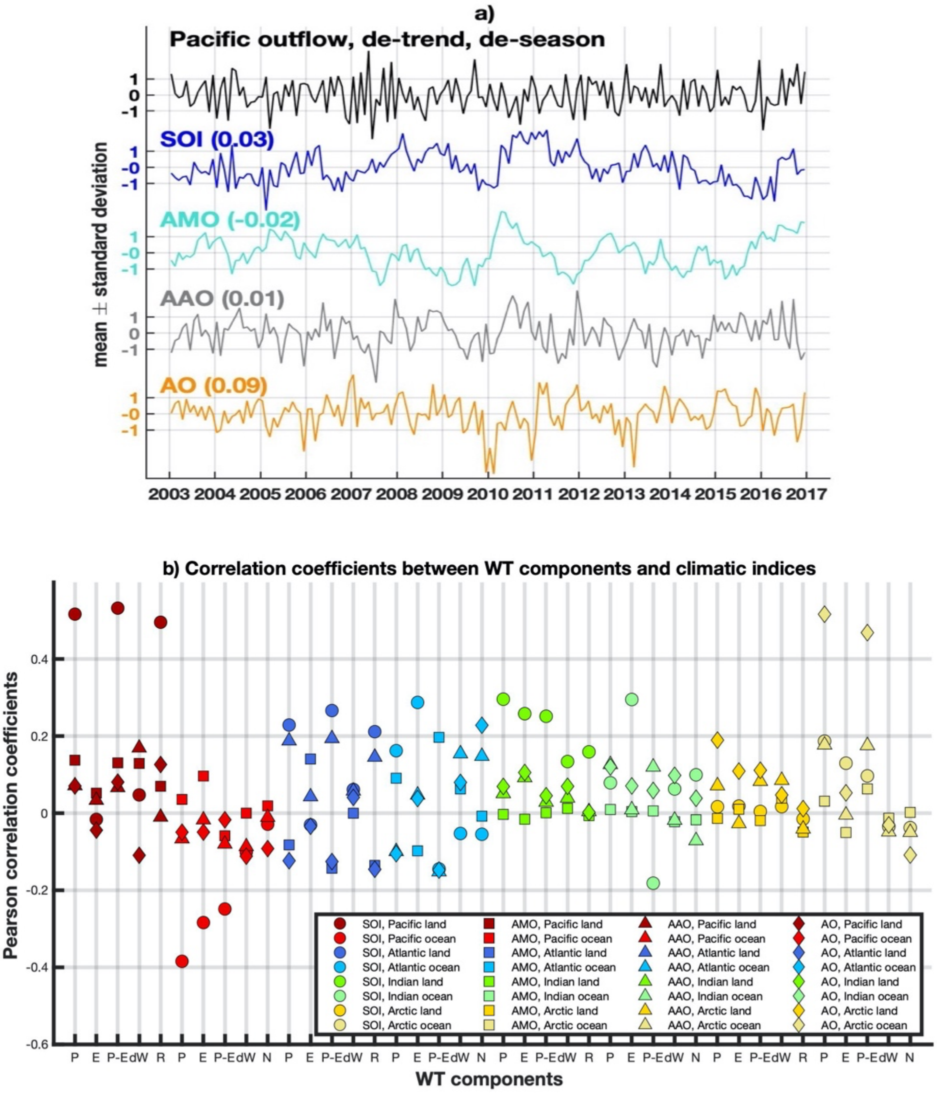


Figure 7. Pacific outflow and climatic indices for ENSO, AMO, AO, and AAO. a) Time series of Pacific outflow is de-trend and de-season. All time series are normalized to have unit variance. Values in the parenthesis are the correlation coefficient between the corresponding climatic index and the Pacific outflow. b) Correlation coefficients between de-trend and de-season WT components of different regions and the climatic indices.

Another possible reason for the lack of correlation resides in the definition of the studied regions, for which the presence of subregions with positive and negative influence of an index results in an overall negligible/attenuated influence of the index in the overall region. For example, a positive phase of the AMO is related to an increase of P in western Europe (Sutton and Hodson, 2005), and the Sahel (Folland et al., 1986; Knight et al., 2006; Zhang and Delworth, 2006; Ting et al., 2009), but to a decrease of P in the U.S. (Enfield et al., 2001; Sutton and Hodson, 2005), and northeast Brazil (Knight et al., 2006; Zhang and Delworth, 2006). All these regions are included in the Atlantic drainage basin, and then the influence of a positive phase of the AMO is attenuated."

Table 3. Correlation coefficients between SOI and de-trend and de-season WT components involved to estimate the Pacific outflow according to Equations 3 and 4.

	$std(X_i)$ (Stand. Deviation)	$corr(X_i, SOI)$ (Correlation between X_i with SOI)	$\frac{std(X_i)}{std(N)}$ (Coefficients)	$corr(X_i, SOI) \cdot \frac{std(X_i)}{std(N)}$ (Correlation ·Coefficient)
$X_1 = -(P-E)_{ocean}$	605	0.25	0.57	0.14
$X_2 = -(P-E)_{land}$	212	-0.53	0.20	-0.11
$X_3 = dW_{land}$	96	0.048	0.09	0.004
$X_4 = dW_{ocean}$	711	-0.10	0.67	-0.07
			Corr(N,SOI)	-0.03

Note that table 3 provides also some insights about the causes of the interannual variability of Pacific Ocean outflow. The largest standard deviation of $P-E$ and dW in the ocean suggests that these two components might drive the interannual variability of the Pacific Ocean outflow. This is confirmed by a correlation analysis. The correlation between N and the $(P-E)_{ocean}$ is -0.70. The correlation between N and the dW_{ocean} is 0.84. The correlation of N with the corresponding land components is below 0.18. In all cases, prior to the evaluation of the correlation the corresponding time series have been de-trend and de-season.

Technical corrections

L44: Correct 'de' here

Done, thank you.

Many details are missing: Such as how were E and R computed? Are these also taken from ERA5?

We have clarified that P and E data are both from ERA5 at the beginning of the description of ERA5 dataset:

> *"The P and E data we use are adopted from the ERA5 reanalysis…"*

We have also clarified that R is estimated as a residual in Equation 1 (in the description of Equatin 1):

> *"The R component will be estimated as a residual in Equation 1."*

Figure 1: What is the source of Figure 1? Or how was this figure obtained?

The source of Figure 1 is the runoff pathways scheme of Oki and Sud (1998) as stated in lines 75-76 of the original manuscript:

"The land is divided into their associated drainages according to the global continental runoff pathways scheme of Oki and Sud (1998)"

To avoid any confusion, we added in the caption of Figure 1 the reference to runoff pathways scheme of Oki and Sud (1998). The new caption of Figure 1 is:

Figure 1. Pacific, Atlantic, Indian, and Arctic Ocean basins and their associated continental drainage basins according to the global continental runoff pathways scheme of Oki and Sud (1998). Within each basin, darker colour represents the continental basin, lighter colour the ocean basin. White regions represent endorheic basins.

Figure 1: Please add color labels to the figure. I also think the figure would look better is you add the basin names in the figure itself.; perhaps major continental basins as well. In addition, major known pathways of water transport can be added to the figure.

The names of the basins have been added to the figure, and in figure caption has been modified.

A diagram of net water transport has been included as a new figure 5 (see answer to specific comment 4). The goal of figure 1 is just to show how the regions have been defined. For this reason, we believe that additional partitions (as major continental basins) might be distracting. Although we agree that showing major known pathways of water transport could enrich the figure we decided not to include it to avoid confusion with the reported net water transport, which is the target of this study.

The results of the changes described above is the new Figure 1

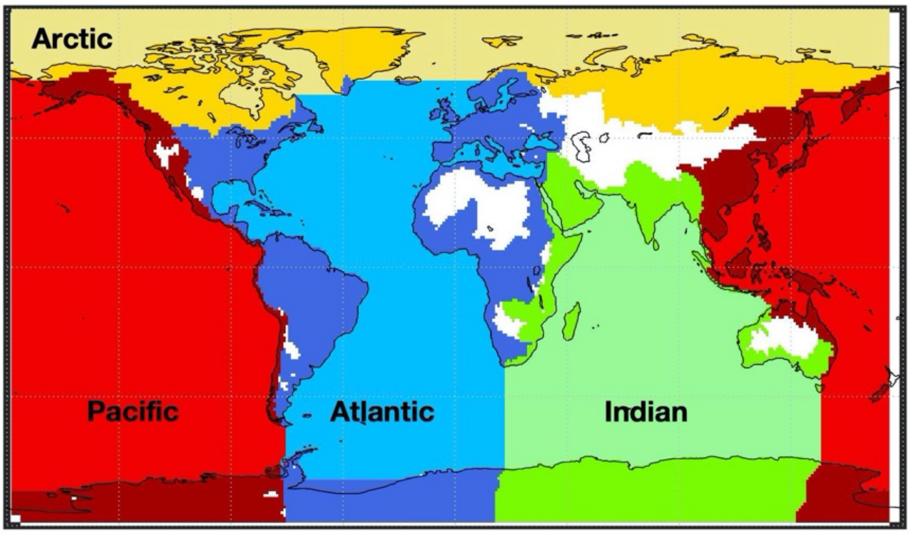


Figure 1. Pacific, Atlantic, Indian, and Arctic Ocean basins and their associated continental drainage basins according to the global continental runoff pathways scheme of Oki and Sud (1998). Within each basin, darker colour represents the continental basin, lighter colour the ocean basin. White regions represent endorheic basins

L93: Elaborate more.

The sentence has been re-written:

> *"Any other non-surficial effect such as long-term tectonics would be incorrectly interpreted as water mass fluxes (Chao, 2016) but they may only have importance in the determination of secular trends; so are the non-climatic sources such as the rare, local earthquake events."*

L94-L95: Please provide more details as to why this is required.

The sentence has been re-written:

> *"As the C_{20} Stokes coefficient is not well determined from GRACE mission, it is replaced with a more accurate solution from Satellite Laser Ranging (SLR) (Cheng and Ries, 2017)."*

L98: Have you defined "GAD" earlier? Please provide details. Also, provide detailed justification of effects ignored here.

The GAD acronym has not been defined. It stands for:
G: Geopotential coefficients;
A: Averaged of any background model over a time period;
D: Bottom pressure over oceans, zero over land.
However, in the GRACE community it is not usually defined (we have had to check its meaning in the GRACE user's manual). On the other hand, it is very common and it is usually referred just as "GAD product". In any case, we give a description of what represents the GAD product for readers no familiarized with GRACE jargon.

L97-100: This sentence is not very clear, rephrase for better clarity.
L99: What is MPIOM?

This part has been re-written:

> "On the other hand, *the atmospheric, and some oceanic, effects on gravity change had beforehand been removed from the processing of the GRACE data by CSR, for de-aliasing purposes, according to the operational numerical weather prediction (NWP) model from ECMWF and to an unconstrained simulation with the global ocean general circulation model MPI-OM* -Max-Planck-Institute Global Ocean/Sea-Ice Model- *(Dobslaw et al. 2017). To recover the "true" ocean mass variability, we restore the removed signal on the oceans adding back the GAD product, which is set to zero on the continents.*"

L105 to 115: I am not able to quite follow these sentences, perhaps because of my lack of expertise in GRACE. I'd suggest writing them in more details for readers who are not well- versed with GRACE estimates. In fact, I feel the entire paragraph can be re-written for more clearity.

We have re-written these sentences:

> *"GRACE's degree-0 Stokes coefficients □C_{00} is set identically to zero on the recognition that Earth's total mass (including the atmosphere) is constant. Then, any increase (decrease) of the water-mass budget of the atmosphere will be counteracted by a decrease (increase) of the same amount of water-mass in the surface. However, after the atmospheric and dynamic oceanic mass changes are corrected in GRACE data, the GRACE □C_{00} are still set to zero even though they should match the opposite of the removed signals. To restore the lost degree-0 signal, the GAD product (which is set to zero on the continents) must be added back to GRACE with averaged ocean signal set to zero, and then, the □C_{00} from an atmospheric model must be subtracted from GRACE data to force the Earth's total mass to be constant. Doing so, the GRACE data will account for the global exchange of water-mass between the Earth surface and atmosphere."*

L126: Bootstrap replications of what variables? Are these timeseries data? L128-130: Please write these in more details. "subtracted" what from what? Please clarify.

We have re-written this part:

> *"The reported 95% confidence intervals and the correlation coefficients are evaluated using the stationary bootstrap scheme of Politis and Romano (1994) (with optimal block length selected according to Patton et al., 2009), and the percentile method. The intuition underlying the bootstrap is simple. Suppose that the observed time series $x_1, \ldots, x_n$ is a realization of the random vector $(X_1, \ldots, X_n)$ with joint distribution P_n and which is assumed to be part of a stationary stochastic process. Given $\mathbf{X}_n$, we first build and estimate $\hat{P}_n$ of P_n. Then B random vectors $(X_1^*, \ldots, X_n^*)$ are generated from $\hat{P}_n$. If $\hat{P}_n$ is a good approximation of P_n, then the relation between $(X_1^*, \ldots, X_n^*)$ and $\hat{P}_n$ should well reproduce the relation between $(X_1, \ldots, X_n)$ and P_n (for an introduction of bootstrap methods for time series see Kreiss and Lahiri (2012) and the references therein). Here, the number of bootstrap replications was set to B=2000. In general, half length of the confidence interval can be very well approximated by twice the standard deviation of the sample mean estimated from the bootstrap replications. Prior to applying the bootstrap to a time series, least-squares estimated linear/quadratic trend and sinusoid with the most relevant frequencies are removed from it to meet the stationarity conditions of the method. In particular, each series has been decomposed into trend, seasonal and residual components. The bootstrap is applied to the residual component producing bootstrap samples of the residuals. For the evaluation of confidence intervals for the different components of WT, the trend and seasonal terms are added back (to the bootstrap sample of the residuals) producing bootstrapped time series of the component of interest. These samples are then used for further analysis. As an illustration, for the WT N component we proceed as follows: (i) a model with linear, annual, and semiannual signals is fitted to the data. The fitted linear trend and annual and semiannual signals are subtracted from the original time series; (ii) the stationary bootstrap is then applied to the residuals producing 2000 bootstrap samples of the residuals; (iii) The estimated trend and seasonal components are added back to each bootstrap sample of the residuals*

obtaining an ensemble of 2000 bootstrapped time series for the N component; (iv) these 2000 bootstrapped time series are used to obtain 95% confidence intervals for the mean fluxes (average of N over the 14 year period of study) and for the amplitude and phase of the annual component using the percentile method. For the mean fluxes, the average of N for each of the 2000 bootstrapped time series was first evaluated and then the 0.025 and 0.975 percentiles of these 2000 averages were reported as 95% confidence interval. For the study of the climatology, a linear trend model with annual and semiannual components was fitted to the 2000 bootstrapped time series producing corresponding estimates of the annual amplitude and phase. The 0.025 and 0.975 percentiles of these estimates were reported as 95% confidence intervals. In order to study the robustness of the results with respect to the model choice, the analysis is rerun using 11 alternative models obtained considering different forms for the trend component (quadratic or constant) and including higher frequencies in the harmonic regression (up to 5). The results are robust. The relative difference with respect to the reported values is smaller than 1.2% for point estimates and smaller than 3.3% for the extremes of the 95% confidence intervals.

As an independent check of the bootstrap, confidence intervals for the mean value of N have been also evaluated by propagating the error estimate in GRACE data (using the JPL GRACE mascon solution for which error estimates are available). The resulting intervals were consistent with those of the bootstrap method. In particular (see Section 4 for details), we show that in all cases the bootstrap intervals contain the intervals obtained from error propagation. In this respect, the CI_{95} *from bootstrap analysis can be considered a conservative estimate. This should be expected, since the residual component underlying the bootstrap approach includes measurement errors and other type of errors (related, for example, with the estimate of the trend and seasonal terms). As a result, the uncertainties in the transports estimated by the bootstrap should be larger than the corresponding uncertainties estimated by error propagation.*

Note that for the study of correlation the bootstrap was applied to the bivariate time series of the residuals of the two variables of interest producing an ensemble of 2000 bivariate time series of residuals. For each bivariate time series of residuals the correlation between the two components of the series was first evaluated. The average and the 0.025 and 0.975 percentiles of these 2000 estimates were reported as point estimate and confidence limits for the correlation between the two variables of interest (correlation between residual components is used to avoid spurious correlation)."

L135: All these can be written concisely in a table.

The text in line 135 summarizes the robustness of the estimation of the main feature of the *N* component of WT, with respect to the model choice (trend + seasonality):

L135: *In order to study the robustness of the results with respect to the model choice, the analysis is rerun using 11 alternative models obtained considering different forms for the* *trend component (quadratic or constant) and including higher frequencies in the harmonic regression (up to 5). The results are robust. The relative difference with respect to the reported values is smaller than 1.2% for point estimates and smaller than 3.3% for the extremes of the 95% confidence intervals.*

Although we see the point of including a table we decided to maintain line L135 in its original form since, in our opinion, it provides a better summary of the robustness than a table. The table, in fact, would be quite large (with height 11 times the height of the actual table 1) and should include details of each model (which might be different for the different basins). For each of the quantities of interest in table 1 the reader should compare the 11 solutions provided by the different models (each solution comprises a point estimate and a confidence interval). All these comparisons, in our opinion, are more effectively summarized in the two lines:

"The relative difference with respect to the reported values is smaller than 1.2% for point estimates and smaller than 3.3% for the extremes of the 95% confidence intervals."

Figure 2: It might be useful to show N in right panels. Perhaps, you can only keep P- E, rather than showing both P and E.

In the new version of Figure 2 N has been included in right panels.

Figure 2: Should be made clearer. Shouldn't "R" be same in both left and right columns. Perhaps, you can remove the space between columns and extended the x-axis of the plots.

Yes, *R* is the same in both columns. We have removed the R time series in the right column of Figures 2 and 4. With respect to the "design" of the figures, we have tried many options and selected the one that seems to us most clear (see the new Figure 2 below).

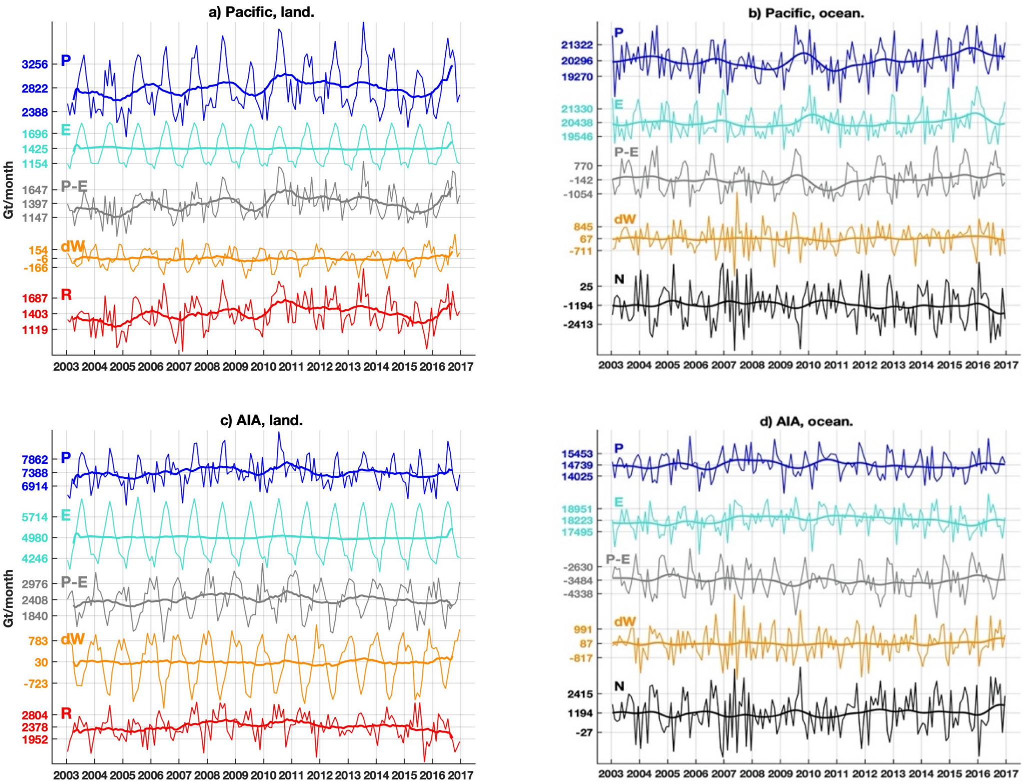


Figure 2. WT of Equations (1) and (2) in the Pacific (first row), Atlantic/Indian/Arctic (AIA) oceans collectively (second row), and their drainage basins. First column: associated land drainage basins; second column: ocean basins. Labels in the vertical axis correspond to the mean ± standard deviation of the associated curve. Thick lines are the low pass filtered signal by a Hann function of 24 months. All curves in the same panel are plotted on the same scale. *P*, *E*, and *P–E* are from ERA5 dataset; *dW* is estimated from GRACE; *R* and *N* are estimated as a residual in equations 1 and 2, respectively.

L145: The P-E estimate of 142 Gt/month should be corrected for final presentation, since it was earlier noted that ERA5 may have some internal inconsistency when comparing previously estimated values from the literature.

The global average of $P-E$ is 188 Gt/month and it should be [−4.3, 0.9]. The 142 Gt/month is for the Pacific Ocean and it may be corrected according to its area. However, we have not applied this correction for two reasons: (1) It is irrelevant for the estimation of *N* after the degree-0 correction of GRACE data is applied; (2) We do not know how it would affect the data locally and how errors would spread on computations.

L145: How does the results on fluxes in this manuscript compare with the values in the literature.

See our answer to specific comments 2 and 3.

Figure 3: I'm not sure if Figure 3 is required. I think it can be removed without loss of any critical details.

We are very interested in showing the agreement between the Pacific outflow and the AIA inflow, which is not trivial at all. That is the reason to keep them in a separate figure.

In our opinion, the comparison between the Pacific outflow and the AIA inflow is critical as a warranty of consistency in the processing of the data.

L145 to 155: I would suggest the values to be put in a graphical format, which might look more appealing that writing the numbers.

It can be seen graphically now in the new diagram inserted as Figure 5 (see the answer to the specific comment 4).

L164: What is basis of deciding the salinity of water in this study?

We provide the results in the form of Gt/month. Freshwater from ERA5 is 1000 kg/m^3, and GRACE data are directly kg/m^2. When converting Gt/month into Sv, it does not matter if we choose a density 1020 or 1035 kg/m^3 since we only show two significant digits (in Sv).

L181: I though the data is monthly; if so, how can we say day of month also?

Yes, data are monthly. The day of the maximum annual signal is obtained from the annual component fitted to the data, which is an analytical function.

L198: Again, it seems important that authors add N time series also in Figures 2 and 4.

N component was already shown in Figure 4 and has been added in Figure 2.

Figure 6: In addition to showing the timeseries, you can show the correlation values of each index with Pacific outflow.

The correlation coefficients are shown in the new version of the Figure (see the answer to the specific comment 5).

L225: The importance of these changes has not been discussed before in the manuscript.

A previous comment has been added in the description of GRACE data:

> *"Any other non-surficial effect such as long-term tectonics would be incorrectly interpreted as water mass fluxes (Chao, 2016) but they may only have importance in the determination of secular trends; so are the non-climatic sources such as the rare, local earthquake events."*

Appendix:
Have we referend to Appendix anywhere in the main text?

It was mis-referenced at the end of the third paragraph of the section “Discussion and Conclusions” as Supplementary Material. It has been corrected.

> *“However, the mean freshwater flux in the Pacific (1261 Gt/month) quite mis-matches that in the Atlantic/Indian Oceans (−1837 Gt/month), meaning that the approximation was quite poor and hence the N term was not properly estimated in these studies (see Appendix for further discussion).”*

Answer to Reviewer 3

We thank the reviewer for thorough reading and thoughtful comments and suggestions. A detailed discussion of the changes that we made in response to the reviewer's comments is given below. In what follows, we state the reviewer's comment in boldface, and describe our response in plain text. Text in the manuscript is represented in italics. The text that has been modified/included in the new version has been highlighted in red.

Review for esd-2020-54 "Water transport among the world ocean basins within the water".
In general, this study is interesting to me, with the global ocean basins to study water mass transport based on GRACE and ERA5. The conclusions are generally supported by the data analyses in this study, but more validations/evaluations are necessary to improve the reliability of the results. At least, a careful intercomparsion between this study and previous studies/literatures can be discussed to enhance our understanding.

We have included several improvements in the manuscript:

1. **Comparison with similar studies:** As far as we are applying a new methodology, there are not many studies to compare with. The only two studies, up to our knowledge, doing something similar are discussed in the third paragraph of the section "Discussion and Conclusions":

   "The results presented here are consistent with the well-known salinity asymmetry between the Pacific and Atlantic Oceans (Reid, 1953; Warren, 1983; Broecker et al., 1985; Zaucker et al., 1994; Rahmstorf, 1996; Emile-Geay et al., 2003; Lagerloef et al., 2008; Czaja, 2009; Reul, 2014). However, they are in contrast to previous GRACE-based studies where a simple seesaw WT between the Pacific and the Atlantic/Indian oceans was reported [Chambers and Willis, 2009; Wouters et al., 2014]. In those studies, the P−E+R term in Equation 2 in both Pacific and Atlantic/Indian Oceans was approximated by that from the global ocean mean. However, the mean freshwater flux in the Pacific (1261 Gt/month) quite mis-matches that in the Atlantic/Indian Oceans (−1837 Gt/month), meaning that the approximation was quite poor and hence the N term was not properly estimated in these studies (see Appendix for further discussion)."

   As stated in the text, in the Appendix we explain in details why the proposed methodology overcomes some important limitations of previous approaches which will always show a seesaw of water transport, even if it does not exist.

2. **Other datasets:**

   The *P*, *E*, *P*−*E*, and *R* components are auxiliar in this study. However, we have added some more references for comparison purposes in the last paragraph of Section 3.1:

“Corresponding analyses have been performed for the Atlantic, Indian, and Arctic Oceans separately. The time evolution of the WT components in Eqs. 1 and 2 are shown in Figure 4, and a diagram of the water-mass fluxes is shown in Figure 5. On average, the Atlantic Ocean receives 926 Gt/month (CI_{95}=[876, 980]; or 0.36 Sv) of salty water, and loses to the atmosphere 879 Gt/month (CI_{95}=[828, 930]) via P−E+R. The latter is equivalent to a freshwater deficit of 0.34 Sv, which increases the near-surface salt concentration and enables water to sink in North Atlantic producing deep water. These values are close to the 0.13-0.32 Sv estimated from ocean models, as needed to keep salinity balance in the Atlantic Ocean (Zaucker et al., 1994). Similarly, the Indian Ocean loses 957 Gt/month (CI_{95}=[894, 1022]) of freshwater that is restored by 991 Gt/month (CI_{95}=[907, 1073]) of salty water. The freshwater lost via P−E+R by the Atlantic and Indian Oceans goes to the Pacific (1261 Gt/month, CI_{95}=[1171, 1347]) and Arctic (730 Gt/month, CI_{95}=[712, 747]) Oceans, which return 1194 (CI_{95}=[1096, 1291]) and 723 (CI_{95}=[708, 739]) Gt/month of salty water through the ocean, respectively. Then, the Pacific presents a surplus of freshwater that reduces near-surface salt concentration, which prevents the formation of deep water. Together, the Pacific and Arctic Oceans supply 1917 Gt/month (CI_{95}=[1812, 2021]) of water to the Atlantic and Indian Oceans, where it is reincorporated into the water cycle via net E−P. As in previous studies (see Craig et al., 2017 for a synthesis), the freshwater lost in the Indian Ocean is similar to that in the Atlantic Ocean. In those studies, P-E+R is close to zero in the Pacific Ocean, producing a difference of 0.4 Sv between Atlantic and Pacific Oceans. In this study, P-E+R is 1261 Gt/month in the Pacific Ocean and the difference with the Atlantic increases to ~0.8 Sv. Some of these differences would be expected as far as the ocean basins are not defined in exactly the same way. On the other hand, the global R is 3781 Gt/month (or 3781 × 12 = 45368 km^3/year), close to the 41867 km^3/year reported by the Global Runoff Data Centre (GRDC, 2014). At basin scale, R is 16834 km^3/year in the Pacific, greater than the 11826 km^3/year reported by GRDC. In the Atlantic, Indian, and Arctic, R is 18228, 4479, and 5827 km^3/year, respectively, which is closer to the GRDC values: 20772, 5238, and 4080 km^3/year. Finally, according to the diagram in Figure 5, the water content in the atmosphere decreases 178 Gt/month (and it is gained by Earth’s surface), but this amount is not realistic as discussed in Section 2 since it should increase a few Gt/month [Nilsson and Elgered, 2008]. This value differs from the 188 Gt/month mentioned in Section 2 because the endorheic regions are not accounted here.”

More importantly, we have extended our analysis to other datasets. The objective is to show that our main results concerning the *N component*, are not an artifact of CSR GRACE and ERA5 datasets. As a result, there is new section entitled “Comparison with other datasets”:

“Equations 1 and 2 are applied to estimate the Pacific outflow using different datasets:

(1) CSR GRACE mascon solution is replaced by the JPL GRACE mascon solution provided by the Jet Propulsion Laboratory/NASA (Watkins et al., 2015; Wiese et al., 2019). Similarly to CSR data, JPL are corrected for GIA effects, C_{20} Stoke coefficients

are replaced by a solution from SLR, and data are reduced to 1º regular grids from 0.5º regular grids. Besides, we have applied the degree-0 Stoke coefficients correction. However, CSR and JPL mascon solutions are not directly comparable. The main reason is that an estimate of degree-1 coefficients has been added to JPL mascon solutions, and the GAD product has not been added back. The corrections applied by JPL are not supplied separately and we cannot do/undo any of the corrections to process JPL data as we did with CSR data. In particular, the GAD product is not available for JPL. In any case, the JPL solution is useful here since it provides an error estimate of the mascon solution that can be propagated to obtain confidence intervals of N, which are independent from those estimated with the bootstrap analysis. Table 2 shows the CI_{95} *of the mean values of the N component for different ocean basin estimated from error propagation and bootstrap analysis. It is observed that in all cases the* CI_{95} *from error propagation are included in those from bootstrap analysis, meaning that the latter are a conservative estimate of the error. JPL propagated error can be expected to be similar to that propagated from CSR error estimates (which are not available), and then we can assume that the reported* CI_{95} *for N calculated from CSR data are a conservative estimate. Besides, comparing Tables 1 and 2, it is observed that the mean values of N are quite similar and that the* CI_{95} *largely overlap. Regarding to the time variability, the values of the N component from CSR and JPL mascon solutions show Pearson correlation coefficients greater than 0.85 (p-value <* 10^{-3}*), except for the Atlantic (0.70). Thus, despite the different processing of CSR and JPL data, the reported analysis for the N component is robust with respect to the choice of GRACE datasets.*

Table 2. Mean net WT from JPL mascon for different ocean basins according to Equation 2 . CI_{95} are estimated as propagation of mascon errors provided by JPL, and from bootstrap analysis. Units are Gt/month.

		Mean (CI_{95} from error propagation)	Mean (CI_{95} from bootstrap)
Outflows	Pacific	1182 (1143, 1220)	1182 (1062, 1306)
	Arctic	735 (713, 757)	735 (711,761)
	Pacific + Arctic	1917 (1872, 1961)	1917 (1806, 2036)
Inflows	AIA	1183 (1092, 1274)	1183 (1077, 1282)
	Atlantic	919 (866, 972)	919 (845, 985)
	Indian	999 (980, 1018)	999 (928, 1067)
	Atlantic + Indian	1918 (1862, 1974)	1918 (1838, 2003)

(2) ERA5 P and E data are replaced by several datasets for comparison purposes. The objective is not to be exhaustive in the selection, but rather to show that the reported features of the N component are quite robust with respect to the choice of the P and E datasets. The data sets considered are:

- *(i) Continental P from GPCC (Schneider et al., 2011), GPCP (Adler et al., 2018), CMAP (Xie and Arkin, 1997), UDel (Willmott and Matsuura, 2001), and GLDAS/Noah (Rodell et al., 2004; Beaudoing and Rodell, 2016).*
- *(ii) Ocean P from GPCP and CMAP.*
- *(iii) Continental E from GLEAM (Miralles et al., 2011; Martens et al., 2017) and GLDAS/Noah.*

(iv) Ocean E from OAFlux (Yu et al., 2008) and HOAPS/CM SAF (Schulz et al., 2009).

The Pacific outflow is estimated with the 162 possible combinations of P and E, including ERA5. The time period is 2003-2016, except for HOAPS/CM SAF and GPCP, which span from 2003 to 12/2014 and 10/2015, respectively. The degree-0 corrections in GRACE data is made for each combination. Note that only ERA5 includes P and E for both continents and oceans. All grids have been homogenized to 1 °regular grids. The main concern here is the heterogeneity of the spatial coverage among datasets. To make the results comparable among datasets, the computations are restricted to the common grid points, which do not cover the entire Earth (Figure 8a).

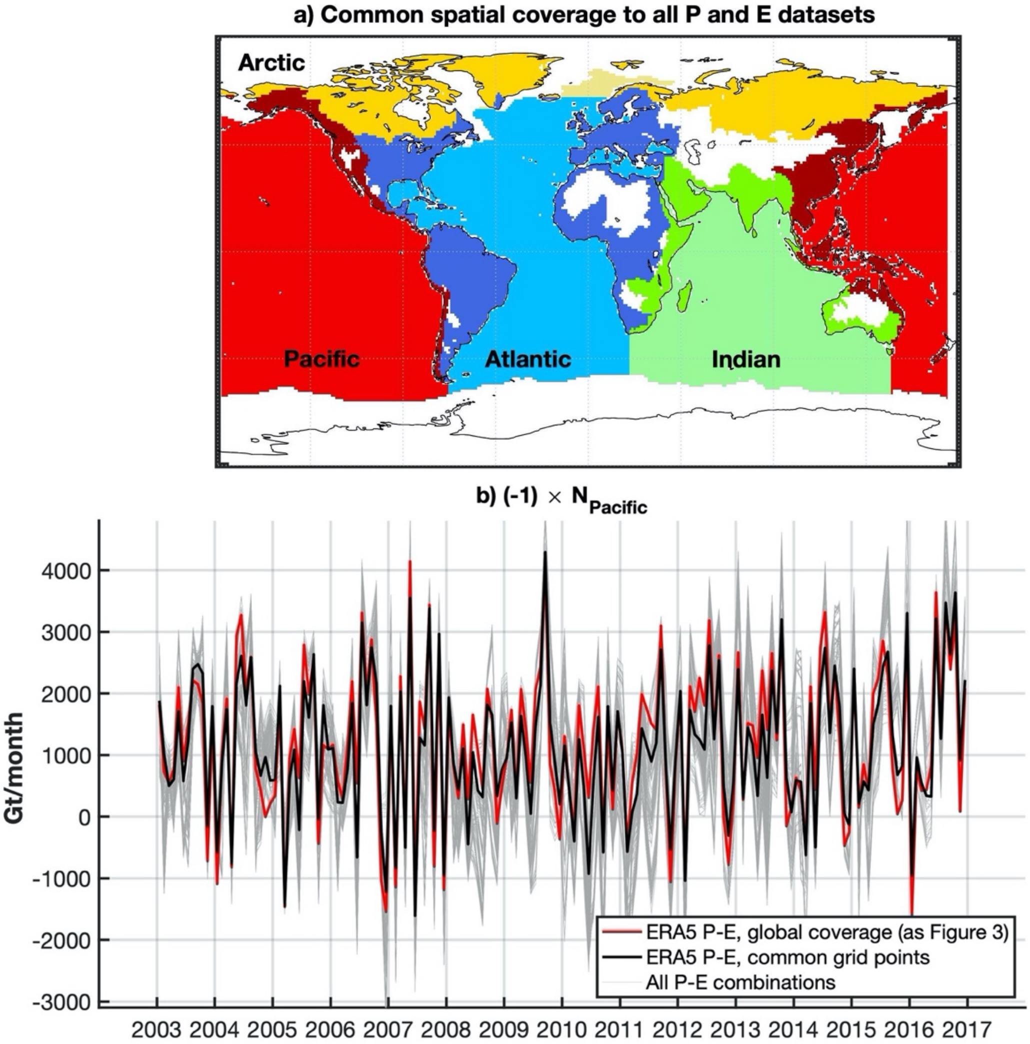


Figure 8. Monthly time series of (the opposite of) the Pacific outflow estimated from 162 combinations of *P* and *E* datasets. a) Spatial coverage common to all datasets. b) Pacific outflows: Gray thin curves are the 162 Pacific outflows estimated in the common grid points to all datasets (no global coverage); black and red curves are based on ERA5 *P* and *E* and are obtained using either only the grid points common to all datasets (black curve) or global coverage (red curve). Note that the red curve is the same as in Figure 3.

However, in spite of the fact that due to the partial coverage the principle of water mass conservation is not accomplished, the Pacific outflow obtained in the common grid points from ERA5 (black line in Figure 8b) is quite in agreement with the same signal obtained with global coverage (red line in Figure 3 which is also reported as red line in Figure 8b). The Pearson correlation coefficient between the two signals is 0.994 (p-values < 10^{-3}) with an average difference around 50 Gt/month. In general, the Pacific outflows estimated from all the P and E dataset combinations show qualitatively the same signal than the one reported in Figure 3. For each of the 162 estimates of the Pacific outflows corresponding to the possible P and E dataset combinations, we evaluated the average outflow (over the period of study), which is 968 Gt/month (STD: 489), and the correlation with the Pacific outflows in Figure 3, which is 0.82 (STD: 0.06; p-values < 10^{-3}).

These experiments show that the reported net WT are physically consistent among datasets, at least qualitatively."

3. **Bootstrap Vs Error propagation:** The confidence intervals estimated from bootstrap have been compared to those estimated from error propagation of the GRACE mascon. As CSR mascon solution does not provide such error estimates, we have used the JPL mascon solution for the comparison. An explanation of why bootstrap confidence intervals contains, as expected, the error propagation confidence interval has been also provided. In the description of the bootstrap method we have included the following text:

As an independent check of the bootstrap, confidence intervals for the mean value of N have been also evaluated by propagating the error estimate in GRACE data (using the JPL GRACE mascon solution for which error estimates are available). The resulting intervals were consistent with those of the bootstrap method. In particular (see Section 4 for details), we show that in all cases the bootstrap intervals contain the intervals obtained from error propagation. In this respect, the CI_{95} from bootstrap analysis can be considered a conservative estimate. This should be expected, since the residual component underlying the bootstrap approach includes measurement errors and other type of errors (related, for example, with the estimate of the trend and seasonal terms). As a result, the uncertainties in the transports estimated by the bootstrap should be larger than the corresponding uncertainties estimated by error propagation.

The details are shown in the new section "Comparison with other datasets", that can be found above.

4. About the lack of correlation. We have included a discussion on the lack of correlation between the inter-annual transports and the indices of ocean-atmosphere interaction. In particular we propose the two following explanations:

> *"To explore this lack of correlation, we have estimated the correlation coefficient between each climatic index and each WT component (Figure 7b).*

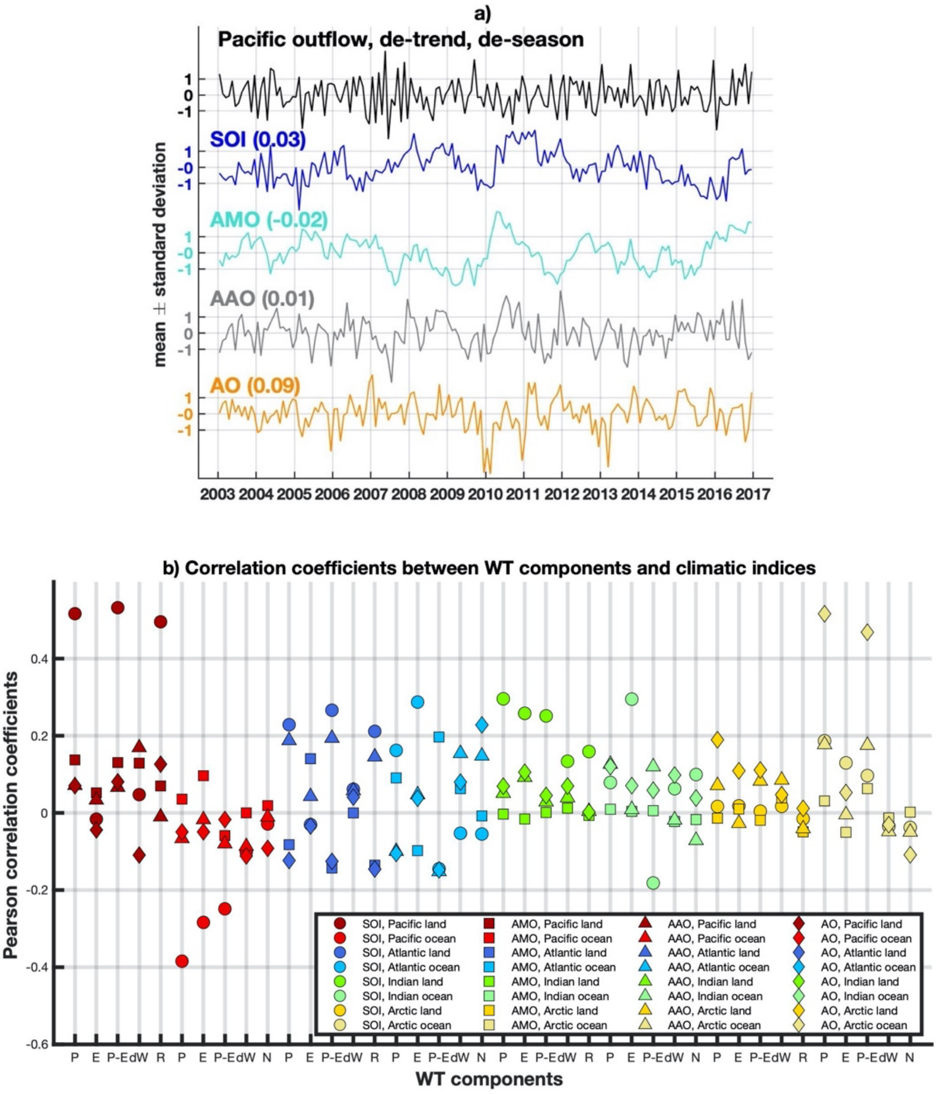


Figure 7. Pacific outflow and climatic indices for ENSO, AMO, AO, and AAO. a) Time series of Pacific outflow is de-trend and de-season. All time series are normalized to have unit variance. Values in the parenthesis are the correlation coefficient between the corresponding climatic index and the Pacific outflow. b) Correlation coefficients between de-trend and de-season WT components of different regions and the climatic indices.

> *All of them are lower than 0.3 except for 6 cases in 2 regions. In the Arctic, P and P−E in the drainage basins of the Arctic show a correlation of ~0.5 with the AO. This correlation is natural since that is the area of influence of the AO. The other region is the Pacific, where, as expected, the SOI shows a correlation around 0.5 with P, P−E, and R in the drainage basins, and around -0.4 with P in the ocean. However, this individual correlation does not extend to the Pacific outflow. In order to understand why this is the case, it is convenient to express the N component of the water transport as a function of (P-E) and dW. According to Equations 1 and 2 we have:*

$$N = -(P-E)_{ocean} - R + dW_{ocean} = \underbrace{-(P-E)_{ocean}}_{X_1} \underbrace{-(P-E)_{land}}_{X_2} + \underbrace{dW_{land}}_{X_3} + \underbrace{dW_{ocean}}_{X_4}. \qquad (3)$$

It can be shown that the correlation between N and a given index can be express as follows

$$corr(N, Index) = \sum_{i=1}^{4} corr(X_i, Index) \cdot \frac{std(X_i)}{std(N)}, \qquad (4)$$

where corr denotes the correlation coefficient, and std stands for standard deviation. As shown in equation (4), the correlation between N and a given index is a linear combination of the correlation between each component and the index. The coefficients of the linear combination std(X_i)/std(N) are proportional to the standard deviation of each component. The components of equation (4) for the Pacific outflow and the SOI index are shown in Table 3. Despite the fact that some of the individual component exhibits significant correlation with SOI (in particular P-E in land and ocean) when combined with the corresponding coefficients their effects are canceled out yielding to a negligible correlation between water transport and SOI (below 0.03 in magnitude).

Another possible reason for the lack of correlation resides in the definition of the studied regions, for which the presence of subregions with positive and negative influence of an index results in an overall negligible/attenuated influence of the index in the overall region. For example, a positive phase of the AMO is related to an increase of P in western Europe (Sutton and Hodson, 2005), and the Sahel (Folland et al., 1986; Knight et al., 2006; Zhang and Delworth, 2006; Ting et al., 2009), but to a decrease of P in the U.S. (Enfield et al., 2001; Sutton and Hodson, 2005), and northeast Brazil (Knight et al., 2006; Zhang and Delworth, 2006). All these regions are included in the Atlantic drainage basin, and then the influence of a positive phase of the AMO is attenuated."

Table 3. Correlation coefficients between SOI and de-trend and de-season WT components involved to estimate the Pacific outflow according to Equations 3 and 4.

	$std(X_i)$ (Stand. Deviation)	$corr(X_i, SOI)$ (Correlation between X_i with SOI)	$\frac{std(X_i)}{std(N)}$ (Coefficients)	$corr(X_i, SOI) \cdot \frac{std(X_i)}{std(N)}$ (Correlation ·Coefficient)
$X_1 = -(P-E)_{ocean}$	605	0.25	0.57	0.14
$X_2 = -(P-E)_{land}$	212	-0.53	0.20	-0.11
$X_3 = dW_{land}$	96	0.048	0.09	0.004
$X_4 = dW_{ocean}$	711	-0.10	0.67	-0.07
Corr(*N*,SOI)				-0.03

Note that table 3 provides also some insights about the causes of the interannual variability of Pacific Ocean outflow. The largest standard deviation of *P-E* and *dW* in the ocean suggests that these two components might drive the interannual variability of the Pacific Ocean outflow. This is confirmed by a correlation analysis. The correlation between *N* and the $(P-E)_{ocean}$ is -0.70. The correlation between *N* and the dW_{ocean} is 0.84. The correlation of *N* with the corresponding land components is below 0.18. In all cases, prior to the evaluation of the correlation the corresponding time series have been de-trend and de-season.

5. **Improvement in the visualization of main results**. We have included a new Figure with a diagram of the mean WT components to ease the reading:

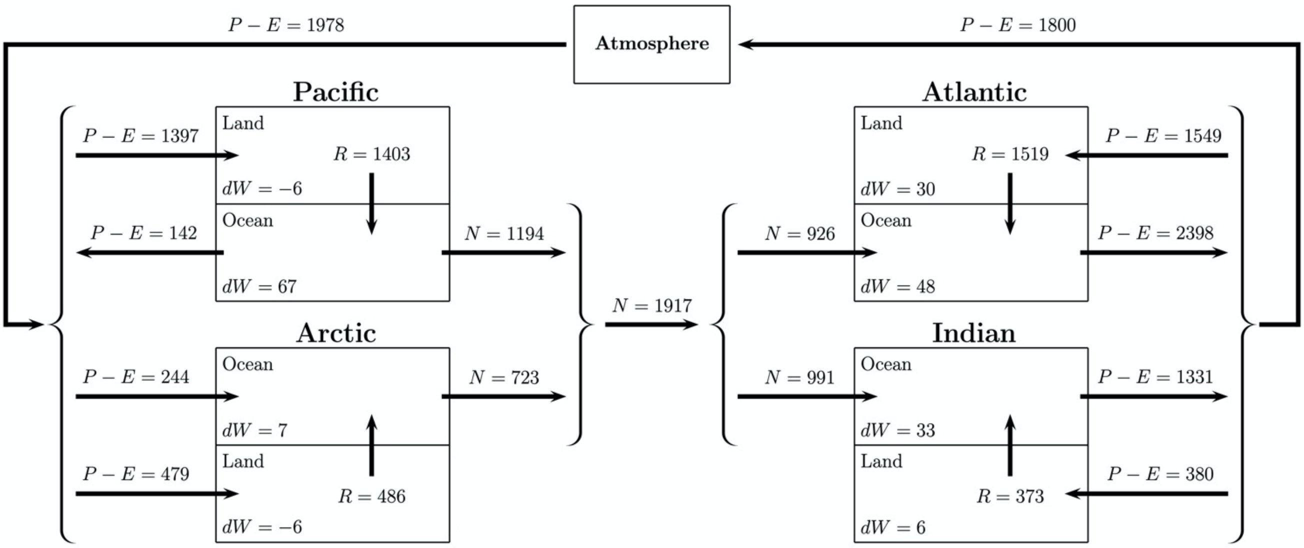


Figure 5. Diagram of the mean values of the WT of the studied regions. Units are Gt/month.

Other comments:
L44, change 'de' to 'the'.

Typo has been corrected.

In section 2: Methodology and Data: please use subtitles to re-organize the section, and improve the readability.

Section 2 is now divided in 4 subsections:
2.1 Methodology
2.2 Precipitation and Evaporation data
2.3 Time-variable GRACE data
2.4 Confidence intervals

Fig 1: suggest to add legends to indicate the locations of different ocean basins.

Figure 1 has been modified, now it includes the names of the basins in the figure itself.

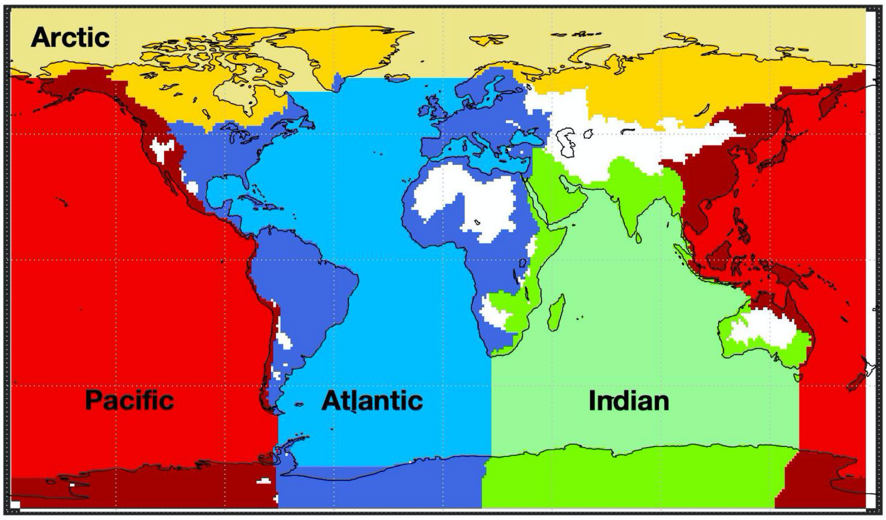


Figure 1. Pacific, Atlantic, Indian, and Arctic Ocean basins and their associated continental drainage basins according to the global continental runoff pathways scheme of Oki and Sud (1998). Within each basin, darker colour represents the continental basin, lighter colour the ocean basin. White regions represent endorheic basins.

Fig 2: are these values obtained from GRACE? or a combination of ERA5 and GRACE? please briefly clarify this in the figure caption.

The next sentence has been added to the caption:
"P, E, and P–E are from ERA5 dataset; dW is estimated from GRACE; R and N are estimated as a residual in equations 1 and 2, respectively."

Fig 3: black curve is the AIA inflow: ’if’ should be ’is’.

Typo has been corrected.

Fig 4: N component should be briefly explained in the caption.

The last sentence of the caption is now:
“Black lines are the WT N component, which are estimated as residuals in Equation 2.”

Fig 6: please indicate the correlation coefficient R between each climate index and the Pacific outflow, at each time series.

It has been included. See the new Figure 7b in the point 4 (About the lack of correlation) of this response.

List of relevant changes

The revised version of the manuscript includes the following relevant changes:

- Lines 47-56: Extension of the motivation to estimate the net water transport in the ocean.
- Section 2.4: More details about the bootstrap methodology.
- Lines 216-226: New comparison with previous studies.
- Lines 273-306: New discussion about the lack of correlation between the Pacific outflow and some climatic indices.
- Section 4: This section is new. It compares the obtained results with those obtained from other datasets.
- Figures 1, 2, 3, and 4: Minor changes.
- Figures 5, 7b, and 8: They are new.
- Tables 2 and 3: They are new.

Water transport among the world ocean basins within the water cycle

David García-García[1], Isabel Vigo[1], Mario Trottini[2]

[1]Applied Mathematics Department, University of Alicante, San Vicente del Raspeig, 03690, Spain
[2]Mathematics Department, University of Alicante, San Vicente del Raspeig, 03690, Spain

Correspondence to: David García-García (d.garcia@ua.es)

Abstract. Global water cycle involves water-mass transport on land, atmosphere, ocean, and among them. Quantification of such transport, and especially its time evolution, is essential to identify footprints of the climate change and helps to constrain and improve climatic models. In the ocean, net water-mass transport among the ocean basins is a key, but poorly estimated parameter presently. We propose a new methodology that incorporates the time-variable gravity observations from the GRACE satellite (2003-2016) to estimate the change of water content, and that overcomes some fundamental limitations of existing approaches. We show that the Pacific and Arctic Oceans receive an average of 1916 (95% confidence interval [1812, 2021]) Gt/month (~0.72 ± 0.02 Sv) of excess freshwater from the atmosphere and the continents that gets discharged into the Atlantic and Indian Oceans, where net evaporation minus precipitation returns the water to complete the cycle. This is in contrast to previous GRACE-based studies, where the notion of a seesaw mass exchange between the Pacific and Atlantic/Indian Oceans has been reported. Seasonal climatology as well as the interannual variability of water-mass transport are also reported.

1 Introduction

The water-mass transport (henceforth WT for brevity) in the oceans is a deciding factor of the world climate system. Quantification of such transport, and especially its time evolution, is essential to better understand the climate change. Atlantic Ocean presents notably a deficit of freshwater flux, in contrast to the Pacific Ocean. This produces a salinity asymmetry that explains why deep waters are formed in the North Atlantic and not in the North Pacific (Warren, 1983; Broecker et al., 1985; Rahmstorf, 1996; Emile-Geay et al., 2003; Czaja, 2009). Upper layers of North Atlantic flow northward, while deep waters flow southward, forming the Atlantic Meridional Overturning Circulation (AMOC), which distributes heat within the Earth system and influences temperature and precipitation patterns worldwide (Vellinga and Wood, 2002). While small changes in hydrological cycle may have caused changes in AMOC during the last glaciation that led to abrupt climate changes (Clark et al., 2002), different models project a weakening of the AMOC in the 21st century that would lead to profound climatic and ecological changes in vast regions (Collins et al., 2019). The Antarctic Circumpolar Current (ACC) receives deep water injected by AMOC with excess salinity, which in turn gets transported into the Indian and Pacific Oceans (Warren, 1981). The Indian Ocean returns saltier water, but Pacific and Arctic Oceans return less-salty waters, producing a salinity imbalance in the Atlantic. To restore the balance, freshwater must be transported outside the Atlantic at the rate of 0.13-0.32 Sv through the atmosphere (Zaucker et al., 1994). This WT produces an excess of freshwater in other ocean regions, as in the Pacific and Arctic Oceans, that must discharge out through the ocean.

Meanwhile, conventional observations on the lateral WT of world ocean climatology have been sparse. In fact, measuring such WT in an open ocean region proves difficult as it amounts only to a few tenths Sv, several orders of magnitude smaller than the total ocean water inflow/outflows in such regions. For example, the Pacific is believed to receive regularly an inflow of 157 ± 10 Sv to south of Australia (Ganachaud and Wunsch, 2000), against three outflows: 0.7-1.1 Sv through the Bering Strait (Woodgate et al., 2012), 16 ± 5 Sv through the Indonesian Strait (Ganachaud and Wunsch, 2000), and 140-175 Sv through the Drake Passage (Ganachaud and Wunsch, 2000; Donohue et al., 2016; Colin de Verdière and Ollitrault, 2016; Vigo et al., 2018).

In this work we propose a new methodology devised to estimate the net WT through the boundaries of a given oceanic region. A defining feature of the proposed approach is the use of the time-variable gravity data from the GRACE (Gravity Recovery and Climate Experiment) satellite mission to estimate the change of water content. We apply the methodology, in conjunction with conventional meteorological data of general hydrologic budget schemes, to estimate the time evolution over the period 2003-2016 of the net WT and exchanges among the four major ocean basins – namely Pacific, Atlantic, Indian, and Arctic. We analyse and report our results of the seasonal climatology as well as the interannual variability of WT. Such information, not available previously, is of valuable importance. For example, in closed regions, net WT through the boundaries on the surface must be counteracted by moisture fluxes through the same boundaries in the atmosphere to match GRACE measurements. Such approach has been successfully applied to study the hydrological cycle of South America (Liu et al., 2006). At ocean basin scale, knowledge about net WT not only would help elucidate the role of the oceans within the water cycle, but it will also impose restrictions on moisture advection in the atmosphere that would help to improve atmospheric models. On the other hand, ocean models usually deal with inflows and outflows of a given ocean region (Warren, 1983; Rahmstorf, 1996; Emile-Geay et al., 2003; de Vries and Weber, 2005; Dijkstra, 2007). Net WT estimates for such ocean region would be useful to impose constraints to the relationship between its inflows and outflows, which would improve the reliability of the models. Better models will improve our knowledge of the Earth's WT dynamics and its evolution in the future, which is critical in the present scenario of climate change.

2. Methodology and Data

2.1 Methodology

The general hydrologic budget equation states that, at any continental location and any moment in time, the change of water content dW equals the precipitation P minus evapotranspiration E (as vertical transport) minus the net runoff R (as horizontal transport):

$$dW = P - E - R \qquad \text{for land.} \tag{1}$$

Under the conservation of water mass, the global net $P-E$ over ocean is negative [e.g., Hartmann, 1994]. That amount of water gets transported to land through atmosphere and returns to the ocean as R completing the water cycle. The general R for a river basin connected to the ocean consists of river runoff, land ice melting, and submarine groundwater discharge to ocean. The R component will be estimated as a residual in Equation 1.

For an ocean region, R represents the inflow from adjacent land regions plus an extra additive term, call it N, accounting for water exchange between neighbouring ocean regions through boundaries, as (positive) inflow or (negative) outflow:

$$dW = P - E + R + N \qquad \text{for ocean.} \tag{2}$$

The ocean water flux N is the target quantity that we shall solve for as a residual in Equation 2, which up till now has been infeasible to estimate directly [Rodell et al., 2015]. Note that N represents the integrated WT over the total-column depth of ocean, including deep-water flows. This is a strength of the GRACE observation for the oceans, compared to in-situ or other remote-sensing measurements typically targeting only the surface layer.

Our targeted four ocean basins are largely separated geographically with designated continental boundaries and restricted water throughways. The land is divided into their associated drainages according to the global continental runoff pathways scheme of Oki and Sud (1998). There are no direct water exchanges in the form of R among land drainages (see Figure 1). The WT component R is estimated through Equation (1) over each continental region, then input to Equation (2) to estimate N in the associated ocean basins.

2.2 Precipitation and Evaporation data

The P and E data we use are adopted from the ERA5 reanalysis [Hersbach et al., 2018], which assimilates observations into general-circulation modelling provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). They are given at 0.25° latitude/longitude regular grids and monthly (and hourly) intervals for global coverage of both continents and oceans. In order to match the spatial resolution of the above-mentioned continental runoff pathways data, we homogenise the grid to 1°x1° by averaging the corresponding 0.25° grid points.

2.3 Time-variable GRACE data

The critical knowledge needed in Equations (1) and (2), now obtainable from GRACE monthly data, is *dW* (Tapley et al., 2004, 2019), the month-to-month difference of the stored water. Note that the GRACE mass variability pertains to WT directly, as opposed to, for example, altimetric sea level measurements that also contain non-WT, steric effects. We use the GRACE "mascon" (mass concentration) solutions that have already been converted into surficial mass from the original time-variable gravity observations (in our case the GRACE RL06 mascon dataset provided by the Center of Space Research (CSR) of University of Texas; see Save et al., 2016, Save, 2019). The non-surficial gravity change due to the glacial isostatic adjustment (GIA) has been removed to the extent of the ICE6G-D model (Peltier et al., 2018). Any other non-surficial effect such as long-term tectonics would be incorrectly interpreted as water mass fluxes (Chao, 2016) but they may only have importance in the determination of secular trends; so are the non-climatic sources such as the rare, local earthquake events. As the C_{20} Stokes coefficient is not well determined from GRACE mission, it is replaced with a more accurate solution from Satellite Laser Ranging (SLR) (Cheng and Ries, 2017). GRACE is not sensitive to the geocenter variations, and its degree-1 Stokes coefficients are set to zero. We had tried adding to GRACE data an estimate of geocenter variations due to modelled water-mass variability (Swenson et al., 2008), and our reported results would change less than 1%. On the other hand, the atmospheric, and some oceanic, effects on gravity change had beforehand been removed from the processing of the GRACE data by CSR, for de-aliasing purposes, according to the operational numerical weather prediction (NWP) model from ECMWF and to an unconstrained simulation with the global ocean general circulation model MPI-OM -Max-Planck-Institute Global Ocean/Sea-Ice Model- (Dobslaw et al. 2017). To recover the "true" ocean mass variability, we restore the removed signal on the oceans adding back the GAD product, which is set to zero on the continents. Data are provided on a 0.25º regular grid; we reduce it to 1º regular grids, still finer than the spatial resolution of GRACE (~300 km), to match the spatial resolution of the continental drainage basin data as above.

GRACE's degree-0 Stokes coefficients ΔC_{00} is set identically to zero on the recognition that Earth's total mass (including the atmosphere) is constant. Then, any increase (decrease) of the water-mass budget of the atmosphere will be counteracted by a decrease (increase) of the same amount of water-mass in the surface. However, after the atmospheric and dynamic oceanic mass changes are corrected in GRACE data, the GRACE ΔC_{00} are still set to zero even though they should match the opposite of the removed signals. To restore the lost degree-0 signal, the GAD product (which is set to zero on the continents) must be added back to GRACE with averaged ocean signal set to zero, and then, the ΔC_{00} from an atmospheric model must be subtracted from GRACE data to force the Earth's total mass to be constant. Doing so, the GRACE data will account for the global exchange of water-mass between the Earth surface and atmosphere. Such correction has recently proved to improve the agreement between the GRACE global ocean mass change and non-steric sea level variation estimates from altimetry and ARGO data (Chen et al., 2019). Looking for consistency between the GRACE and ERA5 datasets, we use ΔC_{00} from $P–E$ to restore degree-0 signal in *dW*. This ΔC_{00} accounts for uniform mass variations in the global surface equivalent to a global averaged signal for $P-E$, at 188 Gt/month (95% confidence interval CI_{95}=[136, 243], see below). As global $-(P-E)$

represents the variability of global total-column water (TCW), it should match the time derivative of the global TCW. However, the average rate of change of the global TWC in ERA5 is 1.5 Gt/month (CI$_{95}$=[−9.2, 12.7]), although in the range of previously reported values of [−0.9, 4.3] Gt/month [Nilsson and Elgered, 2008] departs far from the global −($P-E$) value. This reveals some internal inconsistency within the ERA5 dataset. However, while artificially increasing the dW estimate, the excessive value of $P-E$ does not affect the WT components R and N estimated from Equations (1) and (2), since the degree-0 signal vanishes due to the residual estimate between dW and $P-E$. In fact, adding ΔC_{00} from $P-E$ to GRACE is numerically equivalent to setting $P-E$ ΔC_{00} to zero as far as Equations (1) and (2) are concerned.

2.4 Confidence intervals

The reported 95% confidence intervals and the correlation coefficients are evaluated using the stationary bootstrap scheme of Politis and Romano (1994) (with optimal block length selected according to Patton et al., 2009), and the percentile method. The intuition underlying the bootstrap is simple. Suppose that the observed time series $x_1, \dots, x_n$ is a realization of the random vector $(X_1, \dots, X_n)$ with joint distribution P_n and which is assumed to be part of a stationary stochastic process. Given $\boldsymbol{X}_n$, we first build and estimate $\hat{P}_n$ of P_n. Then B random vectors $(X_1^*, \dots, X_n^*)$ are generated from $\hat{P}_n$. If $\hat{P}_n$ is a good approximation of P_n, then the relation between $(X_1^*, \dots, X_n^*)$ and $\hat{P}_n$ should well reproduce the relation between $(X_1, \dots, X_n)$ and P_n (for an introduction of bootstrap methods for time series see Kreiss and Lahiri (2012) and the references therein). Here, the number of bootstrap replications was set to B=2000. In general, half length of the confidence interval can be very well approximated by twice the standard deviation of the sample mean estimated from the bootstrap replications. Prior to applying the bootstrap to a time series, least-squares estimated linear/quadratic trend and sinusoid with the most relevant frequencies are removed from it to meet the stationarity conditions of the method. In particular, each series has been decomposed into trend, seasonal and residual components. The bootstrap is applied to the residual component producing bootstrap samples of the residuals. For the evaluation of confidence intervals for the different components of WT, the trend and seasonal terms are added back (to the bootstrap sample of the residuals) producing bootstrapped time series of the component of interest. These samples are then used for further analysis. As an illustration, for the WT N component we proceed as follows: (i) a model with linear, annual, and semiannual signals is fitted to the data. The fitted linear trend and annual and semiannual signals are subtracted from the original time series; (ii) the stationary bootstrap is then applied to the residuals producing 2000 bootstrap samples of the residuals; (iii) The estimated trend and seasonal components are added back to each bootstrap sample of the residuals obtaining an ensemble of 2000 bootstrapped time series for the N component; (iv) these 2000 bootstrapped time series are used to obtain 95% confidence intervals for the mean fluxes (average of N over the 14 year period of study) and for the amplitude and phase of the annual component using the percentile method. For the mean fluxes, the average of N for each of the 2000 bootstrapped time series was first evaluated and then the 0.025 and 0.975 percentiles of these 2000 averages were reported as 95% confidence interval. For the study of the climatology, a linear trend model with annual and semiannual components was fitted to the 2000 bootstrapped time series producing corresponding estimates of the annual amplitude and phase. The 0.025 and 0.975 percentiles of these estimates were reported as 95% confidence intervals. In order to study the robustness of the results with respect to the model choice, the analysis is rerun using 11 alternative models obtained considering different forms for the trend component (quadratic or constant) and including higher frequencies in the harmonic regression (up to 5). The results are robust. The relative difference with respect to the reported values is smaller than 1.2% for point estimates and smaller than 3.3% for the extremes of the 95% confidence intervals.

As an independent check of the bootstrap, confidence intervals for the mean value of N have been also evaluated by propagating the error estimate in GRACE data (using the JPL GRACE mascon solution for which error estimates are available). The resulting intervals were consistent with those of the bootstrap method. In particular (see Section 4 for details), we show that in all cases the bootstrap intervals contain the intervals obtained from error propagation. In this respect, the CI_{95} from bootstrap analysis can be considered a conservative estimate. This should be expected, since the residual component underlying the bootstrap approach includes measurement errors and other type of errors (related, for example, with the estimate of the trend and seasonal terms). As a result, the uncertainties in the transports estimated by the bootstrap should be larger than the corresponding uncertainties estimated by error propagation.

Note that for the study of correlation the bootstrap was applied to the bivariate time series of the residuals of the two variables of interest producing an ensemble of 2000 bivariate time series of residuals. For each bivariate time series of residuals the correlation between the two components of the series was first evaluated. The average and the 0.025 and 0.975 percentiles of these 2000 estimates were reported as point estimate and confidence limits for the correlation between the two variables of interest (correlation between residual components is used to avoid spurious correlation).

3. Results

The various WT components of the Pacific and its associated land drainage regions are shown in Figure 2 in units of Gt/month (1 Sv ≈ 2600 Gt/month; 1 Gt = 10^{12} kg, the weight of 1 km^3 of freshwater). The same analysis is applied to the rest of the ocean basins, i.e. the AIA oceans individually and collectively, with its associated land drainages (see Figure 1).

3.1 Mean fluxes

Averaged over the studied 14 years, the Pacific Ocean loses water through the atmospheric $P-E$ at the average rate of 142 Gt/month (CI_{95}=[48, 243]), which is greatly over-compensated by inflow R from land of 1403 Gt/month (CI_{95}=[1370, 1436]). From this surplus, a minor (if any) amount of 67 Gt/month (CI_{95}=[25, 108]) stays (and accumulates) in the Pacific, while 1194 Gt/month (CI_{95}=[1096,1291]) is transported horizontally to the "non-Pacific" Atlantic/Indian/Artic (AIA) oceans, which will be called the "Pacific outflow" hereafter.

In the AIA Oceans, the situation is found to be markedly distinct, given the fact that the AIA oceans together have surface area comparable to the Pacific (177x10^6 m^2). The AIA oceans collectively lose 3484 Gt/month (CI$_{95}$=[3406, 3560]) through the atmospheric $P-E$, that is ~25 times more than does the Pacific. This water deficit is only ~68% compensated by land R inflow of 2378 Gt/month (CI$_{95}$=[2337, 2419]). With the nominal minor amount of water accumulation at 87 Gt/month (CI$_{95}$=[44, 130]), the AIA oceans thus presents an average inflow of 1194 Gt/month (CI$_{95}$=[1102, 1284]) from the Pacific, which will be called the "AIA inflow".

As expected from the overall conservation of water mass inherent in our methodology, the estimated Pacific outflow and AIA inflow match (Figure 3). It is worth mentioning that a difference of 188 Gt/month would exist between the two mean flux values if the degree-0 correction were not applied.

Corresponding analyses have been performed for the Atlantic, Indian, and Arctic Oceans separately. The time evolution of the WT components in Eqs. 1 and 2 are shown in Figure 4, and a diagram of the water-mass fluxes is shown in Figure 5. On average, the Atlantic Ocean receives 926 Gt/month (CI$_{95}$=[876, 980]; or 0.36 Sv) of salty water, and loses to the atmosphere 879 Gt/month (CI$_{95}$=[828, 930]) via $P-E+R$. The latter is equivalent to a freshwater deficit of 0.34 Sv, which increases the near-surface salt concentration and enables water to sink in North Atlantic producing deep water. These values are close to the 0.13-0.32 Sv estimated from ocean models, as needed to keep salinity balance in the Atlantic Ocean (Zaucker et al., 1994). Similarly, the Indian Ocean loses 957 Gt/month (CI$_{95}$=[894, 1022]) of freshwater that is restored by 991 Gt/month (CI$_{95}$=[907, 1073]) of salty water. The freshwater lost via $P-E+R$ by the Atlantic and Indian Oceans goes to the Pacific (1261 Gt/month, CI$_{95}$=[1171, 1347]) and Arctic (730 Gt/month, CI$_{95}$=[712, 747]) Oceans, which return 1194 (CI$_{95}$=[1096, 1291]) and 723 (CI$_{95}$=[708, 739]) Gt/month of salty water through the ocean, respectively. Then, the Pacific presents a surplus of freshwater that reduces near-surface salt concentration, which prevents the formation of deep water. Together, the Pacific and Arctic Oceans supply 1917 Gt/month (CI$_{95}$=[1812, 2021]) of water to the Atlantic and Indian Oceans, where it is reincorporated into the water cycle via net $E-P$. As in previous studies (see Craig et al., 2017 for a synthesis), the freshwater lost in the Indian Ocean is similar to that in the Atlantic Ocean. In these studies, $P-E+R$ is close to zero in the Pacific Ocean, producing a difference of 0.4 Sv between Atlantic and Pacific Oceans. In our study, $P-E+R$ is 1261 Gt/month in the Pacific Ocean and the difference with the Atlantic increases to ~0.8 Sv. Some of these differences would be expected as far as the ocean basins are not defined in exactly the same way. On the other hand, the global R is 3781 Gt/month (or 3781 × 12 = 45368 km^3/year), close to the 41867 km^3/year reported by the Global Runoff Data Centre (GRDC, 2014). At basin scale, R is 16834 km^3/year in the Pacific, greater than the 11826 km^3/year reported by GRDC. In the Atlantic, Indian, and Arctic, R is 18228, 4479, and 5827 km^3/year, respectively, which is closer to the GRDC values: 20772, 5238, and 4080 km^3/year. Finally, according to the diagram in Figure 5, the water content in the atmosphere decreases 178 Gt/month (and it is gained by Earth's surface), but this amount is not realistic as discussed in Section 2 since it should increase a few Gt/month (Nilsson and Elgered, 2008). This value differs from the 188 Gt/month mentioned in Section 2 because the endorheic regions are not accounted here.

3.2 Annual climatology

The WT climatology of the N component is estimated in two ways: (1) averaging the 14 N values for each months of the year (Figure 6a); and (2) fitting a linear trend plus annual and semiannual components model as described in Section 2. Annual amplitudes and phases are plotted in Figure 6b and reported, with corresponding 95% quantile-based confidence intervals, in Table 1.

The Pacific and Arctic Oceans show an overall outflow throughout the year, unlike the Atlantic and Indian Oceans, which show an inflow for every month. The Pacific outflow shows a prominent seasonal undulation peaked around August 3 and a peak-to-peak WT variation of ~2000 Gt/month from boreal summer to November, when a near-zero minimum occurs. The Arctic Ocean show half of the Pacific variability and a less pronounced seasonal undulation. A minimum outflow of ~320 Gt/month is reached in March and April, and a maximum ~1300 Gt/month in July. Together, the Pacific and Arctic Oceans send ~3000 Gt/month of seawater to the Atlantic and Indian Oceans during boreal summer, and a minimum amount five times lower, around 600 Gt/month, in November. The annual maximum is reached on August 8th. The Atlantic/Arctic inflow mirrors this behaviour. Separately, the Atlantic and Indian inflows show a similar peak-to-peak variation of ~2000 Gt/month, reaching the maxima in August and May, respectively. The Indian maximum seems to be related to a local maximum of the Pacific outflow. The annual maxima of net WT of the four basins are reached between August 3rd and September 9th , although the annual signals of the Pacific and Indian Oceans almost triple those from Arctic and Atlantic Oceans (Table 1 and Figure 6b).

3.3 Interannual variability

Interannually, the Pacific outflow shows remarkable variability, mainly produced by P on the continents, which is inherited by R, and $P-E$ in the oceans (Figure 2). For example, the Pacific outflow shows a maximum around 1372 Gt/month in 2009 that matches with a $P-E$ maximum in the Pacific, $P-E$ minimum in the AIA oceans, and P minima in the continental basins draining to both Pacific and AIA oceans. The opposite behaviour, that is a minimum around 939 Gt/month is observed in 2010. The difference, 433 Gt/month, is comparable to the discharge of Amazon (Lorenz et al., 2014). In the tropical Pacific, the El Niño/Southern Oscillation (ENSO) is the strong recurring climate pattern involving changes in the temperature of seawater and air pressure in the tropical Pacific Ocean. The ENSO had a mild El Niño phase in 2009 followed by a strong La Niña phase in 2010, that may be related to the interannual variability of the Pacific outflow. To elucidate this, we conduct a correlation study of the interannual Pacific outflow with respect to the major climate oscillations in the Earth's atmosphere-ocean: ENSO, Atlantic Multi-decadal Oscillation (AMO), Antarctic Oscillation (AAO), and Arctic Oscillation (AO). The climatic oscillation is represented by monthly time series of its indices, which are non-dimensional functions of time derived from relevant meteorological observations; their values indicate the polarity and strength of the oscillation at a given epoch. The ENSO oscillations are measured here with the Southern Oscillation Index (SOI), which represents the sea level pressure differences between Tahiti and Darwin, Australia. The AMO is a coherent mode of natural variability based upon the average anomalies of sea surface temperatures, with AMO Index to reflect the non-secular multi-decadal sea surface temperature pattern variability in the North Atlantic basin. The AAO describes the intensity of westerly wind belt surrounding the Antarctic, quantified by the AAO Index, which is the leading principal component of the 700 hPa atmospheric geopotential height anomalies poleward of 20°S. The AO is to be interpreted as the surface signature of modulations in the strength of the polar vortex aloft the Arctic, while the AO Index is constructed by projecting the 1000 hPa height anomalies poleward of 20°N. Figure 7a show all indices with amplitudes normalized to one standard deviation, as well as the de-trend, de-season, standard deviation normalized Pacific outflow. The correlation analysis between the Pacific outflow and the SOI shows no overall correlation (Pearson coefficient of 0.03) during the whole period, meaning that the influence of ENSO on the Pacific outflow may be restricted to the strong phases of ENSO as in 2009 and 2010. A similar lack of correlation (lower than 0.1) is observed with respect to the AMO, AAO, and AO.

To explore this lack of correlation, we have estimated the correlation coefficient between each climatic index and each WT component (Figure 7b). All of them are lower than 0.3 except for 6 cases in 2 regions. In the Arctic, P and $P-E$ in the drainage basins of the Arctic show a correlation of ~0.5 with the AO. This correlation is natural since that is the area of influence of the AO. The other region is the Pacific, where, as expected, the SOI shows a correlation around 0.5 with P, $P-E$, and R in the drainage basins, and around −0.4 with P in the ocean. However, this individual correlation does not extend to the Pacific outflow. In order to understand why this is the case, it is convenient to express the N component of the water transport as a function of $P-E$ and dW. According to Equations 1 and 2 we have:

$$N = -(P-E)_{ocean} - R + dW_{ocean} = \underbrace{-(P-E)_{ocean}}_{X_1} \underbrace{-(P-E)_{land}}_{X_2} + \underbrace{dW_{land}}_{X_3} + \underbrace{dW_{ocean}}_{X_4}. \qquad (3)$$

It can be shown that the correlation between N and a given index can be express as follows

$$corr(N, SOI) = \sum_{i=1}^{4} corr(X_i, SOI) \cdot \frac{std(X_i)}{std(N)}, \qquad (4)$$

where corr denotes the correlation coefficient, and std stands for standard deviation. As shown in equation (4), the correlation between N and a given index is a linear combination of the correlation between each component and the index. The coefficients of the linear combination, $std(X_i)/std(N)$, are proportional to the standard deviation of each component. The components of equation (4) for the Pacific outflow and the SOI index are shown in Table 3. Despite the fact that some of the individual component exhibits significant correlation with SOI (in particular $P-E$ in land and ocean) when combined with the corresponding coefficients their effects canceled out yielding a negligible correlation between water transport and SOI (below 0.03 in magnitude). Note that table 3 provides also some insights about the causes of the interannual variability of Pacific Ocean outflow. The largest standard deviation of $P-E$ and dW in the ocean suggests that these two components might drive the interannual variability of the Pacific Ocean outflow. This is confirmed by a correlation analysis. The correlation between N and $(P-E)_{ocean}$ is -0.70. The correlation between N and dW_{ocean} is 0.84. The correlation of N with the corresponding land components is below 0.18. In all cases, prior to the evaluation of the correlation the corresponding time series have been detrend and de-season.

Another possible reason for the lack of correlation resides in the definition of the studied regions, for which the presence of subregions with positive and negative influence of an index results in an overall negligible/attenuated influence of the index in the overall region. For example, a positive phase of the AMO is related to an increase of P in western Europe (Sutton and Hodson, 2005), and the Sahel (Folland et al., 1986; Knight et al., 2006; Zhang and Delworth, 2006; Ting et al., 2009), but to a decrease of P in the U.S. (Enfield et al., 2001; Sutton and Hodson, 2005), and northeast Brazil (Knight et al., 2006; Zhang and Delworth, 2006). All these regions are included in the Atlantic drainage basin, and then the influence of a positive phase of the AMO is attenuated.

4. Comparison with other datasets

In this section, we perform a comparisons using alternative datasets. In particular:

(1) CSR GRACE mascon solution is replaced by the JPL GRACE mascon solution provided by the Jet Propulsion Laboratory/NASA (Watkins et al., 2015; Wiese et al., 2019). Similarly to CSR data, JPL are corrected for GIA effects, C_{20} Stoke coefficients are replaced by a solution from SLR, and data are reduced to 1º regular grids from 0.5º regular grids. Besides, we have applied the degree-0 Stoke coefficients correction. However, CSR and JPL mascon solutions are not directly comparable. The main reason is that an estimate of degree-1 coefficients has been added to JPL mascon solutions, and the GAD product has not been added back. The corrections applied by JPL are not supplied separately and we cannot do/undo any of the corrections to process JPL data as we did with CSR data. In particular, the GAD product is not available for JPL. In any case, the JPL solution is useful here since it provides an error estimate of the mascon solution that can be propagated to obtain confidence intervals of N, which are independent from those estimated with the bootstrap analysis. Table 2 shows the CI_{95} of the mean values of the N component for different ocean basin estimated from error propagation and bootstrap analysis. It is observed that in all cases the CI_{95} from error propagation are included in those from bootstrap analysis, meaning that the latter are a conservative estimate of the error. JPL propagated error can be expected to be similar to that propagated from CSR error estimates (which are not available), and then we can assume that the reported CI_{95} for N calculated from CSR data are a conservative estimate. Besides, comparing Tables 1 and 2, it is observed that the mean values of N are quite similar and that the CI_{95} largely overlap. Regarding to the time variability,

the values of the *N* component from CSR and JPL mascon solutions show Pearson correlation coefficients greater than 0.85 (p-value $< 10^{-3}$), except for the Atlantic (0.70). Thus, despite the different processing of CSR and JPL data, the reported analysis for the *N* component is robust with respect to the choice of GRACE datasets.

(2) ERA5 *P* and *E* data are replaced by several datasets for comparison purposes. The objective is not to be exhaustive in the selection, but rather to show that the reported features of the *N* component are quite robust with respect to the choice of the *P* and *E* datasets. The data sets considered are:

(i) Continental *P* from GPCC (Schneider et al., 2011), GPCP (Adler et al., 2018), CMAP (Xie and Arkin, 1997), UDel (Willmott and Matsuura, 2001), and GLDAS/Noah (Rodell et al., 2004; Beaudoing and Rodell, 2016).
(ii) Ocean *P* from GPCP and CMAP.
(iii) Continental *E* from GLEAM (Miralles et al., 2011; Martens et al., 2017) and GLDAS/Noah.
(iv) Ocean *E* from OAFlux (Yu et al., 2008) and HOAPS/CM SAF (Schulz et al., 2009).

The Pacific outflow is estimated with the 162 possible combinations of *P* and *E*, including ERA5. The time period is 2003-2016, except for HOAPS/CM SAF and GPCP, which span from 2003 to 12/2014 and 10/2015, respectively. The degree-0 corrections in GRACE data is made for each combination. Note that only ERA5 includes *P* and *E* for both continents and oceans. All grids have been homogenized to 1° regular grids. The main concern here is the heterogeneity of the spatial coverage among datasets. To make the results comparable among datasets, the computations are restricted to the common grid points, which do not cover the entire Earth (Figure 8a). However, in spite of the fact that due to the partial coverage the principle of water mass conservation is not accomplished, the Pacific outflow obtained in the common grid points from ERA5 (black line in Figure 8b) is quite in agreement with the same signal obtained with global coverage (red line in Figure 3 which is also reported as red line in Figure 8b). The Pearson correlation coefficient between the two signals is 0.994 (p-values $< 10^{-3}$) with an average difference around 50 Gt/month. In general, the Pacific outflows estimated from all the *P* and *E* dataset combinations show qualitatively the same signal than the one reported in Figure 3. For each of the 162 estimates of the Pacific outflows corresponding to the possible *P* and *E* dataset combinations, we evaluated the average outflow (over the period of study), which is 968 Gt/month (STD: 489), and the correlation with the Pacific outflows in Figure 3, which is 0.82 (STD: 0.06; p-values $< 10^{-3}$).

These experiments show that the reported net WT are physically consistent among datasets, at least qualitatively.

5. Discussion and Conclusions

In this work we present a new methodology that combines GRACE data with the general hydrologic budget equation to estimate the horizontal water-mass convergence/divergence for any oceanic region. We have assumed that the gravity changes are produced by mass changes on the Earth surface, such as in the oceans, so that the mascon solution is physically meaningful (Chao, 2016). Any mis-modelling of the ocean basin "container" volume change due to GIA and other non-surficial changes would masquerade as WT variations. However, they are not critical as far as our non-secular analysis is concerned.

We use the proposed methodology to estimate the net WT and exchanges among the Pacific, Atlantic, Indian, and Arctic Oceans, for the period of 2003 – 2016. Our main finding is that the Pacific and Arctic Oceans, while replenished with precipitation and land runoff, are nearly continuously losing water to the Atlantic and Indian Oceans. In particular, the WT climatology is such that the Pacific Ocean loses water at a rate from near zero to up to the peak of 2000 Gt/month during the boreal summer, which coincides with the maximum of the global atmosphere water content. On top of the climatology, the interannual Pacific water loss varies significantly between ~950 to ~1450 Gt/month annual means during the studied period, but seemingly uncorrelated with ENSO.

The results presented here are consistent with the well-known salinity asymmetry between the Pacific and Atlantic Oceans (Reid, 1953; Warren, 1983; Broecker et al., 1985; Zaucker et al., 1994; Rahmstorf, 1996; Emile-Geay et al., 2003; Lagerloef et al., 2008; Czaja, 2009; Reul, 2014). However, they are in contrast to previous GRACE-based studies where a simple seesaw WT between the Pacific and the Atlantic/Indian oceans was reported (Chambers and Willis, 2009; Wouters et al., 2014). In those studies, the $P-E+R$ term in Equation 2 in both Pacific and Atlantic/Indian Oceans was approximated by that from the global ocean mean. However, the mean freshwater flux in the Pacific (1261 Gt/month) quite mis-matches that in the Atlantic/Indian Oceans (−1837 Gt/month), meaning that the approximation was quite poor and hence the N term was not properly estimated in these studies (see Appendix for further discussion).

Differences in freshwater fluxes between the Pacific and Atlantic Oceans produce salinity contrasts, and in turn contrasts on deep water formation. Nevertheless, there are other factors influencing these contrasts such as the narrower extent of the Atlantic (de Boer et al., 2008; Jones and Cessi, 2017), the meridional span of the African and American continents (Nilsson et al., 2013; Cessi and Jones, 2017), and the salty WT from the Indian Ocean to the Atlantic (Gordon, 1986; Marsh et al., 2007). AMOC is also influenced by WT through Bering Strait (Reason and Power, 1994; Goosse et al., 1997; Wadley and Bigg, 2002), and by surface processes of temperature, precipitation and evaporation at low-latitudes of Pacific and Indian Oceans (Newsom and Thompson, 2018). The relative importance among the multiple drivers influencing the AMOC is an open problem (Ferreira, 2018). The net WT estimated here provides information for differences between oceanic inflows and outflows, which can be useful to elucidate on this problem.

Net WT in the open oceans can alternatively be estimated using global ocean models, which simulate observational data based on physical principles. However, these models are not necessarily sensitive to the WT specifically given the data types, and the geography and topography resolutions involved in the models. Knowledge about three-dimensional global ocean circulation could also elucidate on the net WT. However, the small ratio between the net and the total WT hinders the estimation of the former from the latter.

We have applied our WT estimation scheme to the four major ocean basins. The methodology can of course be applied to any extensive ocean region of interest as long as it is much larger than the GRACE resolution. The findings reported here will be useful for a better understanding of the global climate system in terms of its climatology and spatio-temporal variations.

Appendix: Apparent net mass exchange between Pacific and Atlantic/Indian oceans

We shall show here that the net water mass exchange between the Pacific and Atlantic/Indian Oceans reported by Chambers and Willis (2009) was a mathematical artefact. Their Equation (2) approximated the freshwater flux, i.e. $P-E+R$, of the Pacific (Pcf) and Atlantic/Indian (AI) oceans by the global ocean (GO) mean. However, from Figures 2 and 4 we get very different $(P-E+R)_{Pcf}=1261$ and $(P-E+R)_{AI}=-1837$ Gt/month, meaning that the approximation in Chambers and Willis (2009) and hence their resultant estimates of WT are rather poor. In addition, under their approximation an apparent net mass exchange will always arise, since

$$
\begin{aligned}
(P-E+R)_{GO} &= \sum_{x\in GO} \frac{(P-E+R)_x \cdot Area(x)}{Area(GO)} = \\
&= \sum_{x\in Pcf} \frac{(P-E+R)_x \cdot Area(x)}{Area(GO)} + \sum_{x\in AI} \frac{(P-E+R)_x \cdot Area(x)}{Area(GO)} = \\
&= \sum_{x\in Pcf} \frac{(P-E+R)_x \cdot Area(x)}{Area(Pcf)} \cdot \frac{Area(Pcf)}{Area(GO)} + \sum_{x\in Atl/Indian} \frac{(P-E+R)_x \cdot Area(x)}{Area(AI)} \cdot \frac{Area(AI)}{Area(GO)} \approx \\
&\approx (P-E+R)_{Pcf} \cdot \frac{1}{2} + (P-E+R)_{AI} \cdot \frac{1}{2},
\end{aligned}
$$

where x are disjoint grid cells in the ocean basins, the areas of the Pcf, AI, and GO are around 177, 160, and 351 x 10^6 km^2, and the ratios 177/351 and 160/351 have been approximated by 1/2. Then, multiplying by 2 and rearranging the equation we get,

$$(P-E+R)_{Pcf} - (P-E+R)_{GO} \approx -[(P-E+R)_{AI} - (P-E+R)_{GO}].$$

Thus, wherever the signal is in the Pacific and Atlantic/Indian oceans, the anomalies with respect to the global ocean mean will always mirror each other, showing an apparent net mass exchange between them, even if such exchange does not exist.

Data availability

All datasets used in this study are publicly available.

Author contributions

DGG designed the study, processed datasets, and wrote the first draft. MT and IV evaluated the bootstrap confidence intervals, and contributed to other aspects of the data analysis including the robustness with respect to the use of alternative datasets. IV also provided funding for the research. All the authors discussed and interpreted the results, contributed to the writing and revision of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We would like to thank Ben Chao for sowing the seeds of the idea of this work, during a research stay of D. García-García in Goddard Space Flight Center/NASA in 2004, and for all the subsequent fruitful discussions and advice. We thank the organizations that provide the data used in this work, which are publicly available: ERA5 data provided by the Copernicus Climate Change Service Climate Data Store (CDS), https://cds.climate.copernicus.eu/cdsapp#!/home; GPCC, CMAP, and UDel precipitation data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, https://www.esrl.noaa.gov/psd/; GPCP Precipitation data by the Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center, https://precip.gsfc.nasa.gov/; OAFlux evaporation data provided by the WHOI OAFlux project (http://oaflux.whoi.edu) funded by the NOAA Climate Observations and Monitoring (COM) program; HOAPS Evaporation data provided by CM SAF/EUMETSAT, https://www.cmsaf.eu; GLDAS/Noah data by GSFC/NASA, https://disc.gsfc.nasa.gov/; GLEAM evaporation data available at https://www.gleam.eu/; GRACE time-variable gravity data provided by CSR, University of Texas, http://www2.csr.utexas.edu/grace, and by JPL/NASA, https://podaac.jpl.nasa.gov/dataset/TELLUS_GRAC-GRFO_MASCON_CRI_GRID_RL06_V2; Southern Oscillation Index (SOI) from Bureau of Meteorology de Australia; Atlantic Multidecadal Oscillation (AMO) from the NOAA Earth System Research Laboratory (ESRL); Arctic Oscillation (AO) and Antarctic Oscillation (AAO) from NOAA Climate Prediction Center (CPC). This research is funded by Spanish Ministry of Science, Innovation and Universities grant number RTI2018-093874-B-100.

References

Adler, R. F., Sapiano, M., Huffman, G. J., Wang, J.-J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., Xie, P., Ferraro, R., and Shin, D.-B.: The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. Atmosphere, 9(4), 138, doi:10.3390/atmos9040138, 2018.

Beaudoing, H., and Rodell, M.: NASA/GSFC/HSL (2016), GLDAS Noah Land Surface Model L4 monthly 1.0 x 1.0 degree V2.1, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 28/05/2019, doi:10.5067/LWTYSMP3VM5Z, 2016.

Broecker, W. S., Peteet, D. M., and Rind, D.: Does the ocean-atmosphere system have more than one stable mode of operation? Nature 315, 21–26, https://doi.org/10.1038/315021a0, 1985.

Cessi, P., and Jones, C. S.: Warm-route versus cold-route interbasin exchange in the meridional overturning circulation. J. Phys. Oceanogr. 47, 1981–97, https://doi.org/10.1175/JPO-D-16-0249.1, 2017.

Chambers, D. P., and Willis, J. K.: Low-frequency exchange of mass between ocean basins. J. Geophys. Res. 114, C11008, https://doi.org/10.1029/2009JC005518, 2009.

Chao, B. F.: Caveats on the equivalent water thickness and surface mascon solutions derived from the GRACE satellite-observed time-variable gravity. Journal of Geodesy, 90, 807-813, https://doi.org/10.1007/s00190-016-0912-y, 2016.

Chen, J., Tapley, B., Seo, K.-W., Wilson, C., and Ries, J.: Improved Quantification of Global Mean Ocean Mass Change Using GRACE Satellite Gravimetry Measurements. Geophys. Res. Lett. 46, 13984-13991, https://doi.org/10.1029/2019GL085519, 2019.

Cheng, M. K., and Ries, J. C.: The unexpected signal in GRACE estimates of C20. J. Geodesy 91, Issue 8, 897-91, https://doi.org/10.1007/s00190-016-0995-5, 2017.

Clark, P. U., Pisias, N. G., Stocker, T. F., and Weaver, A. J.: The role of the thermohaline circulation in abrupt climate change. , 2002. Nature, 415, 863-869, https://doi.org/10.1038/415863a, 2002.

Colin de Verdière, A., and Ollitrault, M.: A direct determination of the World Ocean barotropic circulation. J. Phys. Oceanogr. 46, 255-273, https://doi.org/10.1175/JPO-D-15-0046.1, 2016.

Collins, M., Sutherland, M., Bouwer, L., Cheong, S.-M., Frölicher, T., Jacot Des Combes, H., Koll Roxy, M., Losada, I., McInnes, K., Ratter, B., Rivera-Arriaga, E., Susanto, R. D., Swingedouw, D., and Tibig, L.: Extremes, Abrupt Changes and Managing Risk. In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. In press., 2019.

Craig, P. M., Ferreira, D., and Methven, J.: The contrast between Atlantic and Pacific surface water fluxes. Tellus A: Dynamic Meteorology and Oceanography 69:1, 1330454, https://doi.org/10.1080/16000870.2017.1330454, 2017.

Czaja, A.: Atmospheric control on the thermohaline circulation. J. Phys. Oceanogr. 39, 234–247, https://doi.org/10.1175/2008JPO3897.1, 2009.

de Boer, A. M., Toggweiler, J. R., and Sigman, D. M.: Atlantic dominance of the meridional overturning circulation. J. Phys. Oceanogr. 38, 435–450, https://doi.org/10.1175/2007JPO3731.1, 2008.

de Vries, P., and Weber, S. L.: The Atlantic freshwater budget as a diagnostic for the existence of a stable shut down of the meridional overturning circulation. Geophys. Res. Lett. 32, L09606, https://doi.org/10.1029/2004GL021450, 2005.

Dijkstra, H. A.: Characterization of the multiple equilibria regime in a global ocean model. Tellus 59A, 695–705, https://doi.org/10.1111/j.1600-0870.2007.00267.x, 2007.

Dobslaw, H., Bergmann-Wolf, I., Dill, R., Poropat, L., and Flechtner, F.: Product Description Document for AOD1B Release 06. GRACE 327-750 document, Rev. 6.1, October 19, 2017.

Donohue, K. A., Tracey, K. L., Watts, D. R., Chidichimo, M. P., and Chereskin, T. K.: Mean Antarctic Circumpolar Current transport measured in Drake Passage. Geophys. Res. Lett. 43, 11760-11767, https://doi.org/10.1002/2016GL070319, 2016.

Emile-Geay, J., Cane, M. A., Naik, N., Seager, R., Clement, A. C., and van Geen, A.: Warren revisited: Atmospheric freshwater fluxes and ‘‘Why is no deep water formed in the North Pacific’’. J. Geophys. Res. 108 (C6), 3178, https://doi.org/10.1029/2001JC001058, 2003.

Enfield, D. B., Mestas-Nuñez, A. M., and Trimble, P.: The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S., Geophys. Res. Lett., 28, 2077–2080, doi:10.1029/2000GL012745, 2001.

Ferreira, D., Cessi, P., Coxall, H. K., de Boer, A., Dijkstra, H. A., Drijfhout, S. S., Eldevik, T., Harnik, N., McManus, J. F., Marshall, D. P., Nilsson, J., Roquet, F., Schneider, T., and Wills, R. C.: Atlantic- Pacific asymmetry in deep water formation. Annual Reviews of Earth and Planetary Sciences 46, 327–352, https://doi.org/10.1146/annurev-earth-082517-010045, 2018.

Folland, C. K., Palmer, T. N., and Parker, D. E.: Sahel rainfall and worldwide sea temperatures, 1901–85, Nature, 320, 602–607, doi:10.1038/320602a0, 1986.

Ganachaud, A., and Wunsch, C.: Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. Nature 408, 453-457, https://doi.org/10.1038/35044048, 2000.

GRDC (2014): Global Freshwater Fluxes into the World Oceans / Online provided by Global Runoff Data Centre. 2014 ed. Koblenz: Federal Institute of Hydrology (BfG), 2014.

Goosse, H., Campin, J. M., Fichefet, T., and Deleersnijder, E.: Sensitivity of a global ice-ocean model to the Bering Strait throughflow. Clim. Dyn. 13, 349 – 358, https://doi.org/10.1007/s003820050170, 1997.

Gordon, A. L.: Interocean exchange of thermocline water, J. Geophys. Res. 91, 5037–5046, https://doi.org/10.1029/JC091iC04p05037, 1986.

Hartmann, D. L. Global Physical Climatology. International Geophysics v. 56 (Academic Press, 1994).

Hersbach, H., de Rosnay, P., Bell, B., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Alonso-Balmaseda, M., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J.-R., de Boisséson, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P., Dee, D., Dragani, R., Diamantakis, M., Flemming, J., Forbes, R., Geer, A. J., Haiden, T., Hólm, E., Haimberger, L., Hogan, R., Horányi, A., Janiskova, M., Laloyaux, P., Lopez, P., Munoz-Sabater, J., Peubey, C., Radu, R., Richardson, D., Thépaut, J.-N., Vitart, F., Yang, X., Zsótér, E., and Zuo, H..: Operational global reanalysis: progress, future directions and synergies with NWP. ERA Report Series 27, doi: 10.21957/tkic6g3wm, 2018.

Jones, C. S., and Cessi, P. Size Matters: Another Reason Why the Atlantic Is Saltier than the Pacific. J. Phys. Oceanogr. 47, 2843-2859, https://doi.org/10.1175/JPO-D-17-0075.1, 2017.

Knight, J. R., Folland, C. K., and Scaife, A. A.: Climate impacts of the Atlantic Multidecadal Oscillation, Geophys. Res. Lett., 33, L17706, doi:10.1029/2006GL026242, 2006.

Kreiss, J.-P., and Lahiri, S. N.: Bootstrap Methods for Time Series, Handbook of Statistics 30. Editor(s): Tata Subba Rao, Suhasini Subba Rao, C.R. Rao. Elsevier, Volume 30, pages3-26, ISBN 9780444538581, https://doi.org/10.1016/B978-0-444-53858-1.00001-6, 2016.

Lagerloef, G., Colomb, F. R., Le Vine, D., Wentz, F., Yueh, S., Ruf, C., Lilly, J., Gunn, J., Chao, Y., deCharon, A., Feldman, G., and Swift, C.: The Aquarius/SAC-D Mission: Designed to meet the salinity remote-sensing challenge. Oceanography 21 (1), 68-81, https://doi.org/10.5670/oceanog.2008.68, 2008.

Liu, W. T., Xie, X., Tang, W., and Zlotnicki, V.: Spacebased observations of oceanic influence on the annual variation of South American water balance. Geophys. Res. Lett., 33, L08710, doi:10.1029/2006GL025683, 2006.

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N., and Riegger, J.: Large-Scale Runoff from Landmasses: A Global Assessment of the Closure of the Hydrological and Atmospheric Water Balances. J. Hydrometeor. 15, 2111-2139, https://doi.org/10.1175/JHM-D-13-0157.1, 2014.

Marsh, R., Hazeleger, W., Yool, A., and Rohling, E. J.: Stability of the thermohaline circulation under millennial CO2 forcing and two alternative controls on Atlantic salinity. Geophys. Res. Lett. 34, L03605, https://doi.org/10.1029/2006GL027815, 2007.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903–1925, doi: 10.5194/gmd-10-1903-2017, 2017.

Miralles, D. G., Holmes, T. R. H., de Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrology and Earth System Sciences, 15, 453–469, doi: 10.5194/hess-15-453-2011, 2011.

Newsom, E. R., and Thompson, A. F.: Reassessing the role of the Indo-Pacific in the ocean’s global overturning circulation. Geophys. Res. Lett. 45, 12422–12431, https://doi.org/10.1029/2018GL080350, 2018.

Nilsson, J., Langen, P. L., Ferreira, D., and Marshall, J.: Ocean basin geometry and the salinification of the Atlantic Ocean. J. Clim. 26, 6163–6184, https://doi.org/10.1175/JCLI-D-12-00358.1, 2013

Nilsson, T., and Elgered, G.: Long-term trends in the atmospheric water vapor content estimated from ground-based GPS data. J. Geophys. Res. 113, D19101, https://doi.org/10.1029/2008JD010110, 2008.

Oki, T., and Sud, Y. C.: Design of Total Runoff Integrating Pathways (TRIP) - A Global River Channel Network. Earth Interactions 2, Paper No. 1, https://doi.org/10.1175/1087-3562(1998)002<0001:DOTRIP>2.3.CO;2, 1998.

Patton, A., Politis, D. N, and White, H.: Correction to "Automatic Block-Length Selection for the Dependent Bootstrap", Econometric Reviews 28, 372-375, https://doi.org/10.1080/07474930802459016, 2009.

Peltier, W. R., Argus, D. F., and Drummond, R.: Comment on "An assessment of the ICE-6G_C (VM5a) glacial isostatic adjustment model" by Purcell et al. J. Geophys. Res. Solid Earth 123, 2019– 2028, https://doi.org/10.1002/2016JB013844, 2018.

Politis, D. N., and Romano, J. P.: The stationary bootstrap. J. Amer. Statist. Assoc. 89, 1303-1313. https://doi.org/10.1080/01621459.1994.10476870, 1994.

Rahmstorf, S. On the freshwater forcing and transport of the Atlantic thermohaline circulation. Clim. Dyn. 12, 799–811, https://doi.org/10.1007/s003820050144, 1996.

Reason, C. J. C., and Power, S. B.: The influence of the Bering Strait on the circulation in a coarse resolution global ocean model. Clim. Dyn. 9, 363–369, https://doi.org/10.1007/BF00223448, 1994.

Reid, J. L.: On the temperature, salinity, and density differences between the Atlantic and Pacific oceans in the upper kilometre. Deep Sea Research 7, Issue 4, 265-275, https://doi.org/10.1016/0146-6313(61)90044-2, 1953.

Reul, N., Fournier, S., Boutin, J., Hernandez, O., Maes, C., Chapron, B., Alory, G., Quilfen, Y., Tenerelli, J., Morisset, S., Kerr, Y., Mecklenburg, S., and Delwart, S.: Sea Surface Salinity Observations from Space with the SMOS Satellite: A New Means to Monitor the Marine Branch of the Water Cycle. Surv. Geophys. 35, 681-722, https://doi.org/10.1007/s10712-013-9244-0, 2014.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bull. Amer. Meteor. Soc., 85, 381-394, doi:10.1175/BAMS-85-3-381, 2004.

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., Adler, R., Bosilovich, M. G., Clayson, C. A., Chambers, D., Clark, E., Fetzer, E. J., Gao, X., Gu, G., Hilburn, K., Huffman, G. J., Lettenmaier, D. P., Liu, W. T., Robertson, F. R., Schlosser, C. A., Sheffield, J., and Wood, E. F.: The Observed State of the Water Cycle in the Early Twenty-First Century, J. Climate 28, 8289-8318, https://doi.org/10.1175/JCLI-D-14-00555.1, 2015.

Ropelewski, C. F., and Halpert, M. S.: Global and regional precipitation patterns associated with El Niño-Southern Oscillation. Mon. Weather Rev. 115, 1606-1625, https://doi.org/10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2, 1987.

Save, H., Bettadpur, S., and Tapley, B. D.: High resolution CSR GRACE RL05 mascons. J. Geophys. Res. Solid Earth 121, https://doi.org/10.1002/2016JB013007, 2016.

Save, H.: "CSR GRACE RL06 Mascon Solutions", https://doi.org/10.18738/T8/UN91VR, Texas Data Repository Dataverse, V1, https://doi.org/10.18738/T8/UN91VR, 2019.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Ziese, M.: GPCC Full Data Reanalysis Version 6.0 at 1.0°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_M_V7_100, 2011.

Schulz, J., Albert, P., Behr, H.-D., Caprion, D., Deneke, H., Dewitte, S., Dürr, B., Fuchs, P., Gratzki, A., Hechler, P., Hollmann, R., Johnston, S., Karlsson, K.-G., Manninen, T., Müller, R., Reuter, M., Riihelä, A., Roebeling, R., Selbach, N., Tetzlaff, A., Thomas, W., Werscheck, M., Wolters, E., and Zelenka, A.: Operational climate monitoring from space: the EUMETSAT Satellite Application Facility on Climate Monitoring (CM-SAF), Atmos. Chem. Phys., 9, 1687-1709, https://doi.org/10.5194/acp-9-1687-2009, 2009.

Sutton, R. T., and Hodson, D. L. R.: Atlantic Ocean forcing of North American and European summer climate, Science, 309, 115–118, doi: 10.1126/science.1109496, 2005.

Swenson, S., Chambers, D., and Wahr, J.: Estimating geocenter variations from a combination of GRACE and ocean model output, J. Geophys. Res., 113, B08410, https://doi.org/10.1029/2007JB005338, 2008.

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

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J. S., Landerer, F. W., Chambers, D. P., Reager, J. T., Gardner, A. S., Save, H., Ivins, E. R., Swenson, S. C., Boening, C., Dahle, C., Wiese, D. N., Dobslaw, H., Tamisiea, M. E., and Velicogna, I.: Contributions of GRACE to understanding climate change. Nat. Clim. Change 9, 358-369, https://doi.org/10.1038/s41558-019-0456-2, 2019.

Ting, M., Kushnir, Y., Seagar, R., and Li, C.: Forced and internal twentieth-century SST trends in the North Atlantic, J. Clim, 22, 1469–1481, doi:10.1175/2008JCLI2561.1, 2009.

Vellinga, M., and Wood, R. A.: Global climate impacts of a collapse of the Atlantic thermohaline circulation. Clim. Change 54, 251–267, https://doi.org/10.1023/A:1016168827653, 2002.

Vigo, M. I., García-Garcia, D., Sempere, M. D., and Chao, B. F.: 3D Geostrophy and Volume Transport in the Southern Ocean. Remote Sens. 10, 715, https://doi.org/10.3390/rs10050715, 2018.

Wadley, M. R., and Bigg, G. R.: Impact of flow through the Canadian Archipelago and Bering Strait on the North Atlantic and Arctic circulation: An ocean modelling study. Q. J. R. Meteorol. Soc. 128, 2187 – 2203, https://doi.org/10.1256/qj.00.35, 2002.

Warren, B. A.: Deep circulation of the World Ocean. In: Warren, B.A., Wunsch, C. (eds) Evolution of physical oceanography. MIT Press, Cambridge, Ma., USA, https://doi.org/10.4319/lo.1982.27.2.0397a, 1981.

Warren, B. A.: Why is no deep water formed in the North Pacific? J. Mar. Res. 41, 327-347, https://doi.org/10.1357/002224083788520207, 1983.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE. J. Geophys. Res. Solid Earth 120, 2648–2671, https://doi.org/10.1002/2014JB011547, 2015.

Willis, J. K., Chambers, D. P., and Nerem, R. S.: Assessing the globally averaged sea level budget on seasonal to interannual time-scales. J. Geophys. Res. 113, C06015, https://doi.org/10.1029/2007JC004517, 2008.

Wiese, D. N., Yuan, D.-N., Boening, C., Landerer, F. W., and Watkins, M. M.: JPL GRACE and GRACE-FO Mascon Ocean, Ice, and Hydrology Equivalent Water Height Coastal Resolution Improvement (CRI) Filtered Release 06 Version 02. Ver. 02. PO.DAAC, CA, USA. Dataset accessed [2020-08-20] at https://doi.org/10.5067/TEMSC-3JC62, 2019.

Willmott, C. J., and Matsuura, K.: Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999), http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html, 2001.

Woodgate, R. A., Weingartner, T. J., and Lindsay, R.: Observed increases in Bering Strait oceanic fluxes from the Pacific to the Arctic from 2001 to 2011 and their impacts on the Arctic Ocean water column. Geophys. Res. Lett. 39, L24603, https://doi.org/10.1029/2012GL054092, 2012.

Wouters, B., Bonin, J. A., Chambers, D. P., Riva, R. E. M., Sasgen, I., Wahr, J.: GRACE, time-varying gravity, Earth system dynamics and climate change. Rep. Prog. Phys. 77, 116801, doi:10.1088/0034-4885/77/11/116801, 2014.

Xie, P., and Arkin, P.A.: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bull. Amer. Meteor. Soc., 78, 2539 – 2558, 1997.

Yu, L., Jin, X., and Weller, R. A.: Multidecade Global Flux Datasets from the Objectively Analyzed Air-sea Fluxes (OAFlux) Project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables. Woods Hole Oceanographic Institution, OAFlux Project Technical Report. OA-2008-01, 64pp. Woods Hole. Massachusetts, 2008.

Zhang, R., and Delworth, T. L.: Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes, Geophys. Res. Lett., 33, L17712, doi:10.1029/2006GL026267, 2006.

Zaucker, F., Stocker, T. F., and Broecker, W. S. Atmospheric freshwater fluxes and their effect on the global thermohaline circulation. J. Geophys. Res. 99, 12443–12457, https://doi.org/10.1029/94JC00526, 1994.

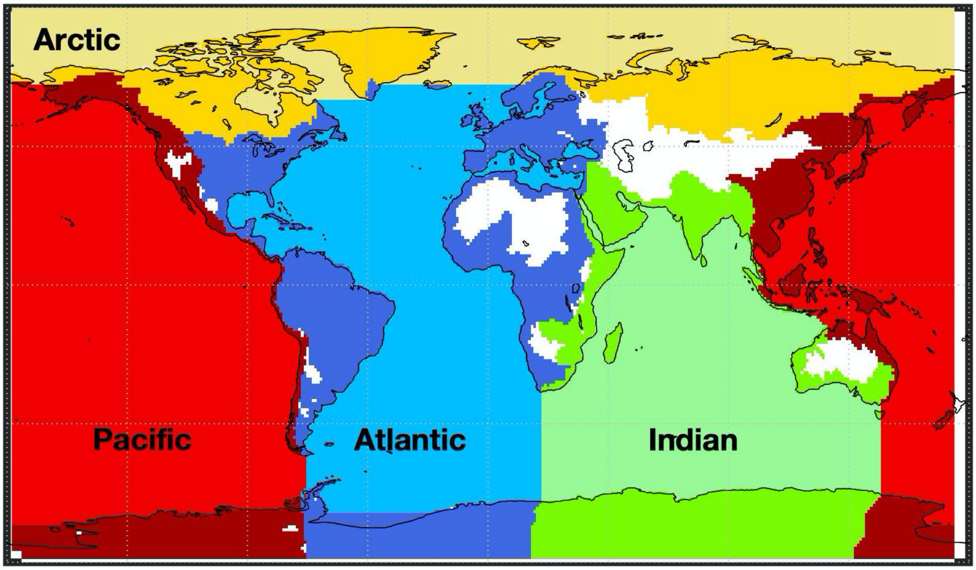
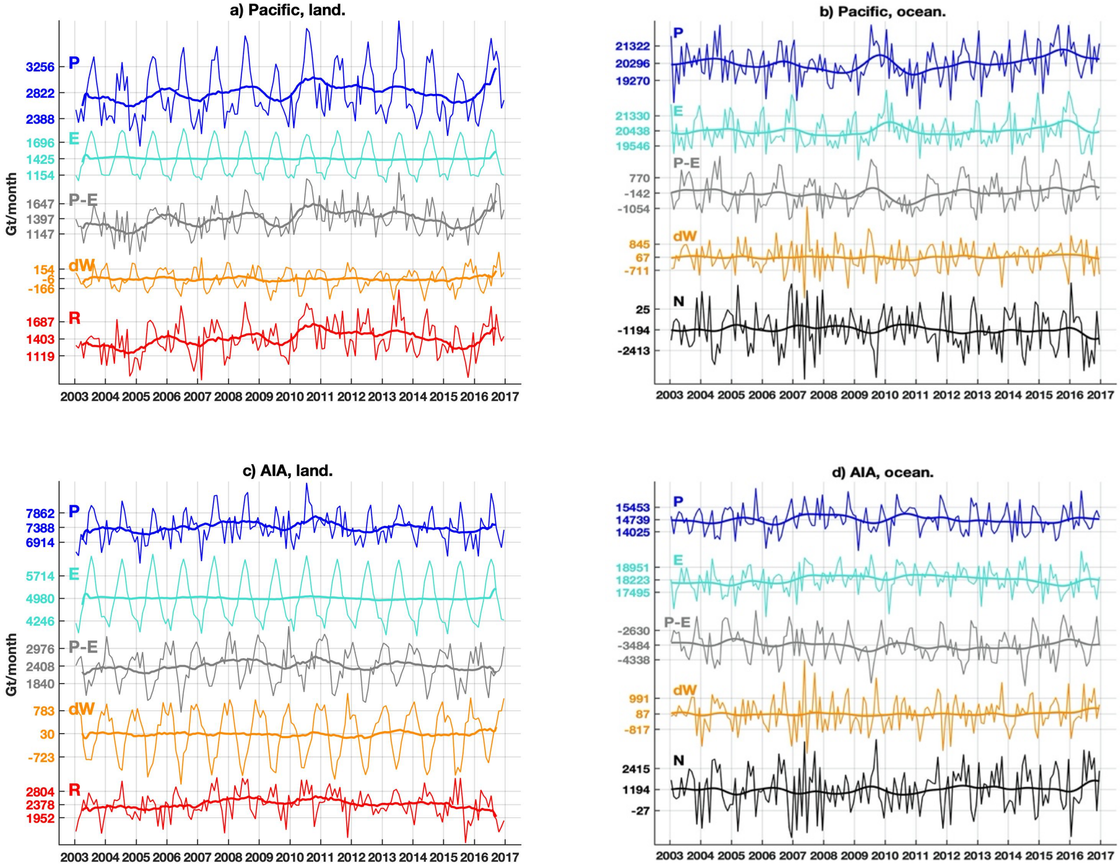


Figure 1. Pacific, Atlantic, Indian, and Arctic Ocean basins and their associated continental drainage basins according to the global continental runoff pathways scheme of Oki and Sud (1998). Within each basin, darker colour represents the continental basin, lighter colour the ocean basin. White regions represent endorheic basins.


a) Pacific, land.

b) Pacific, ocean.

c) AIA, land.

d) AIA, ocean.

Figure 2. WT of Equations (1) and (2) in the Pacific (first row), Atlantic/Indian/Arctic (AIA) oceans collectively (second row), and their drainage basins. First column: associated land drainage basins; second column: ocean basins. Labels in the vertical axis correspond to the mean ± standard deviation of the associated curve. Thick lines are the low pass filtered signal by a Hann function of 24 months. All curves in the same panel are plotted on the same scale. *P*, *E*, and *P–E* are from ERA5 dataset; *dW* is estimated from GRACE; *R* and *N* are estimated as a residual in equations 1 and 2, respectively.

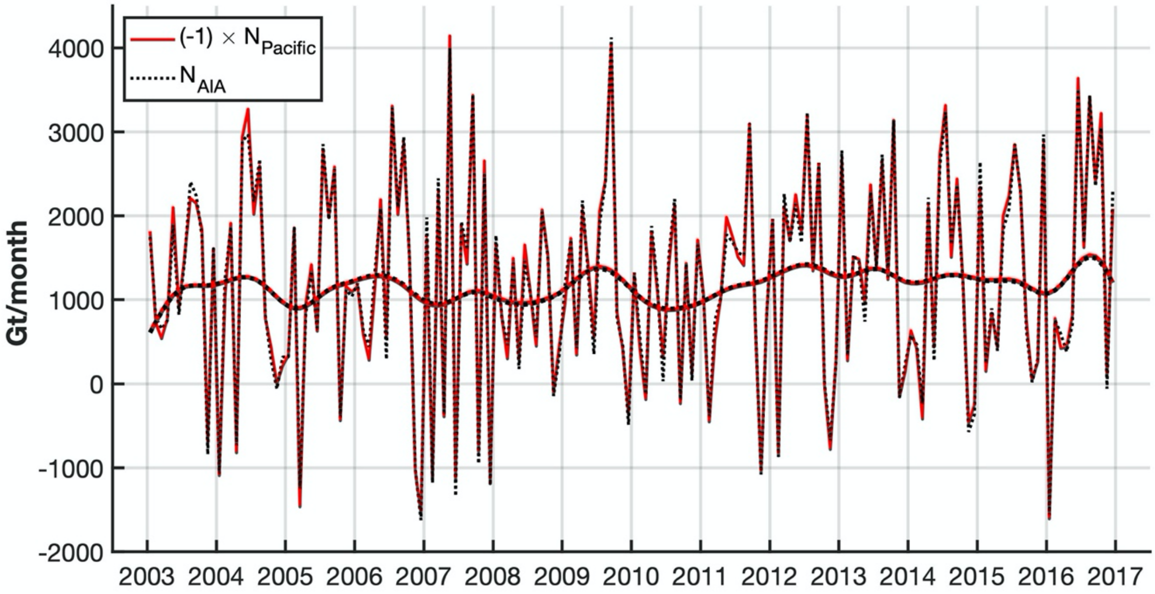
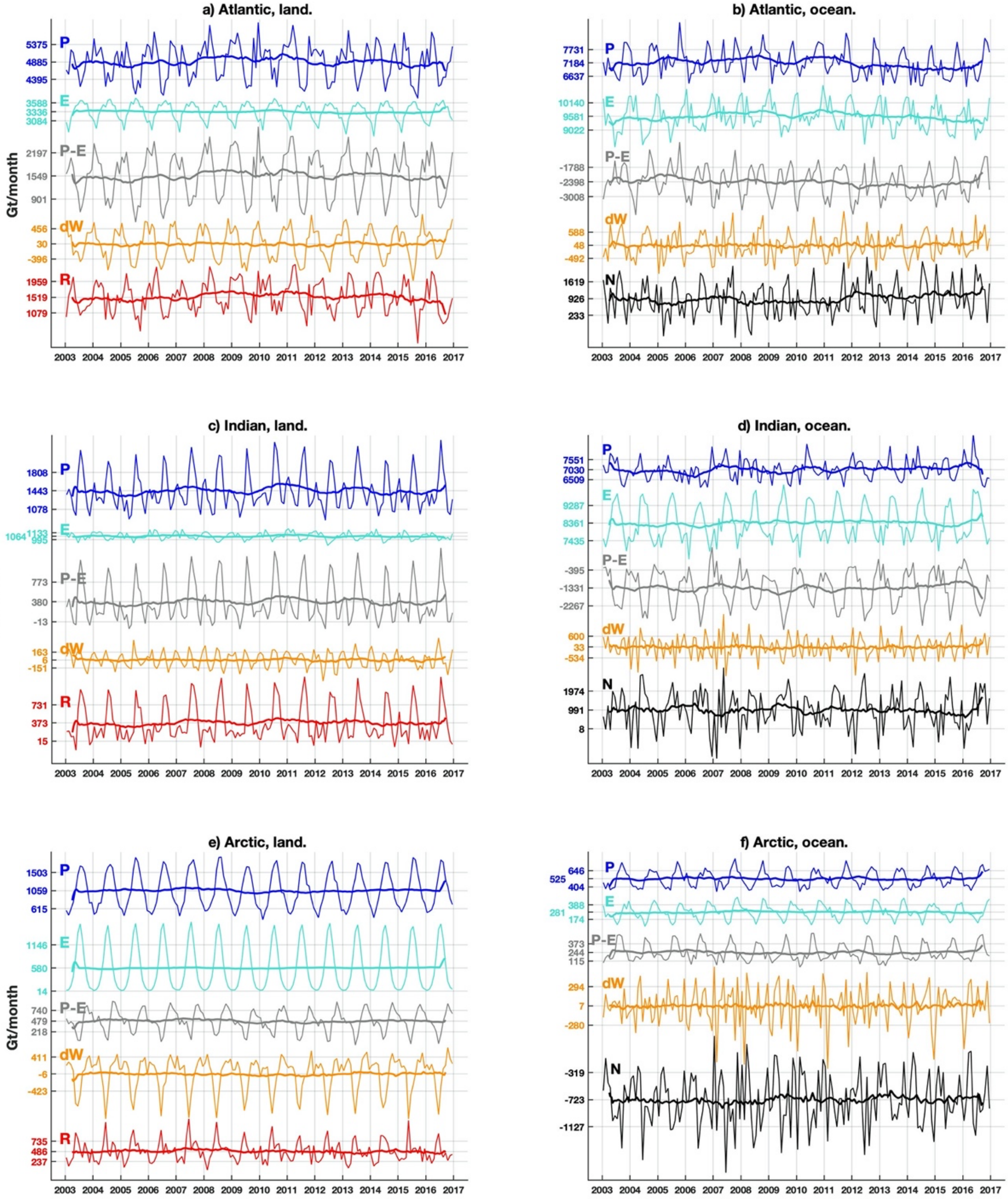


Figure 3. Monthly time series of WT flux from the Pacific to the AIA Oceans. Red curve is (the opposite of) the Pacific outflow, and black curve is the AIA inflow. Thick lines are the low pass filtered signal by a Hann function of 24 months.



735 **Figure 4. As Figure 2 but for Atlantic, Indian, and Arctic Oceans.**

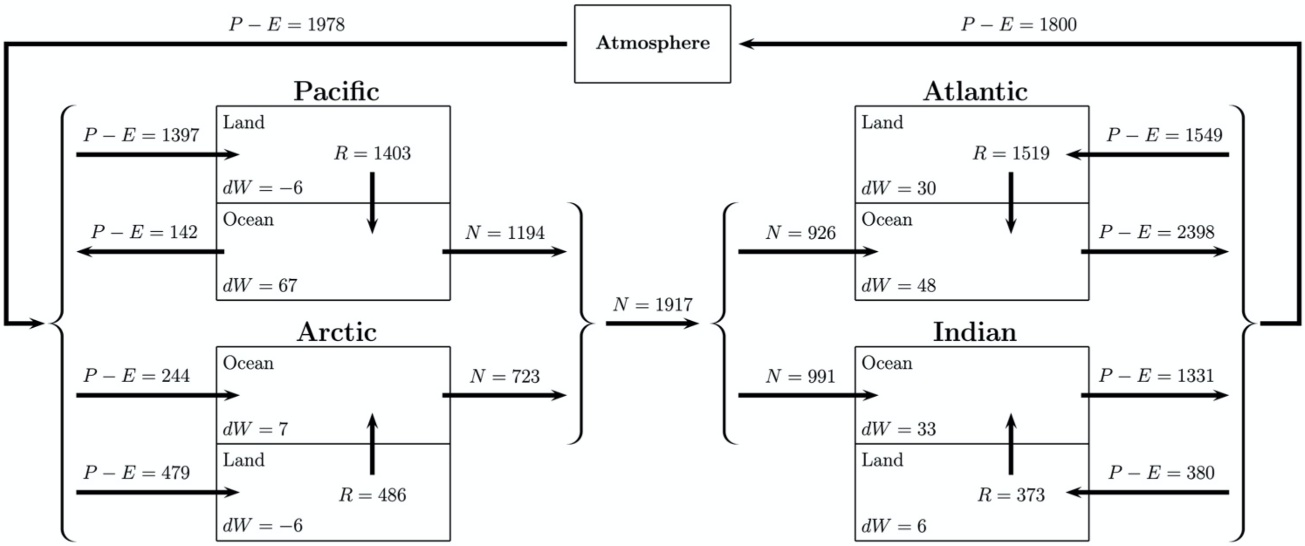


Figure 5. Diagram of the mean values of the WT of the studied regions. Units are Gt/month.

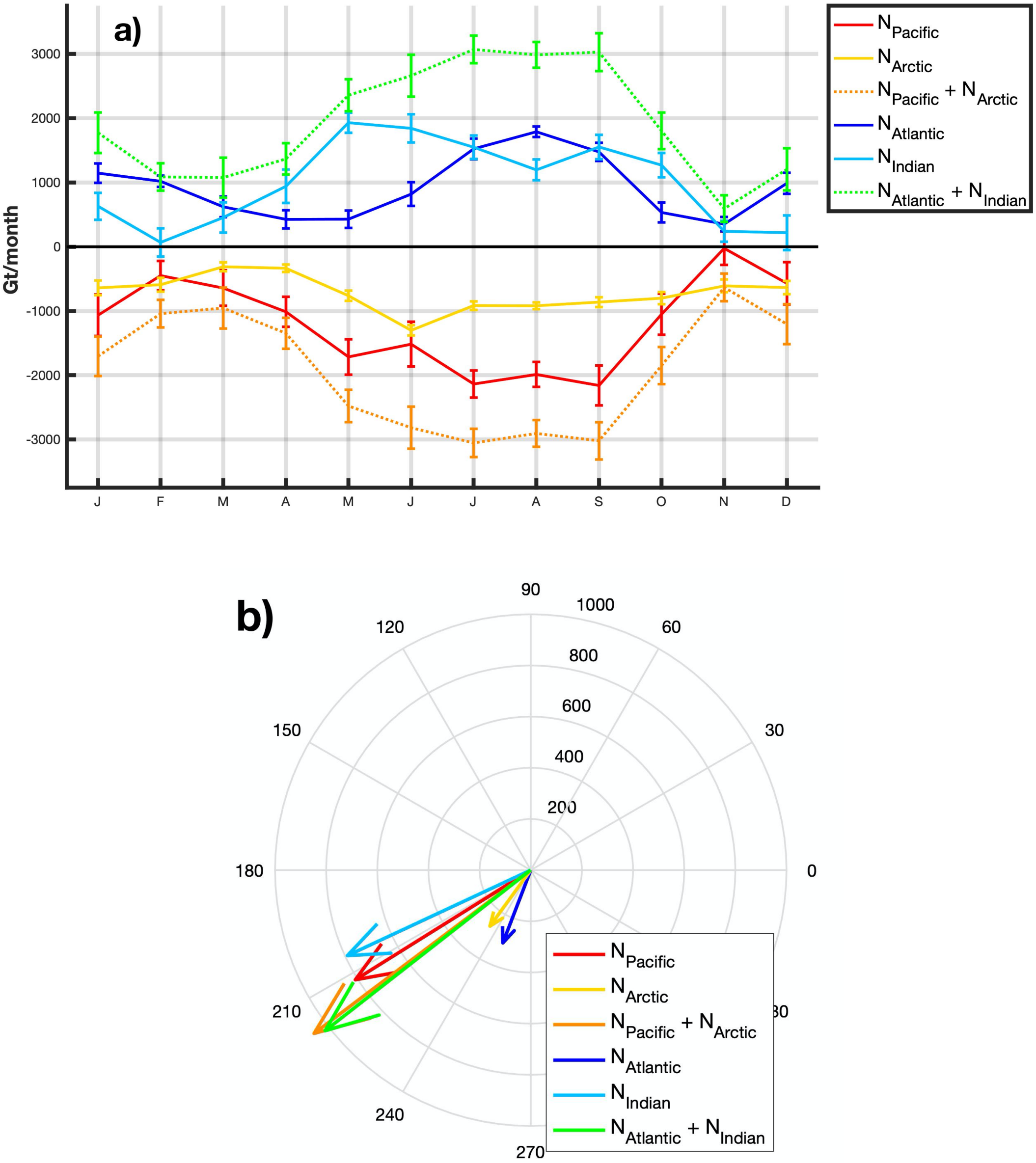


Figure 6. (a) Annual climatology time series (error bar is one standard deviation), and (b) phasor diagram (amplitude in unit of Gt/month, phase angle according to Equation 3) of the inflow/outflow WT of the ocean basins.

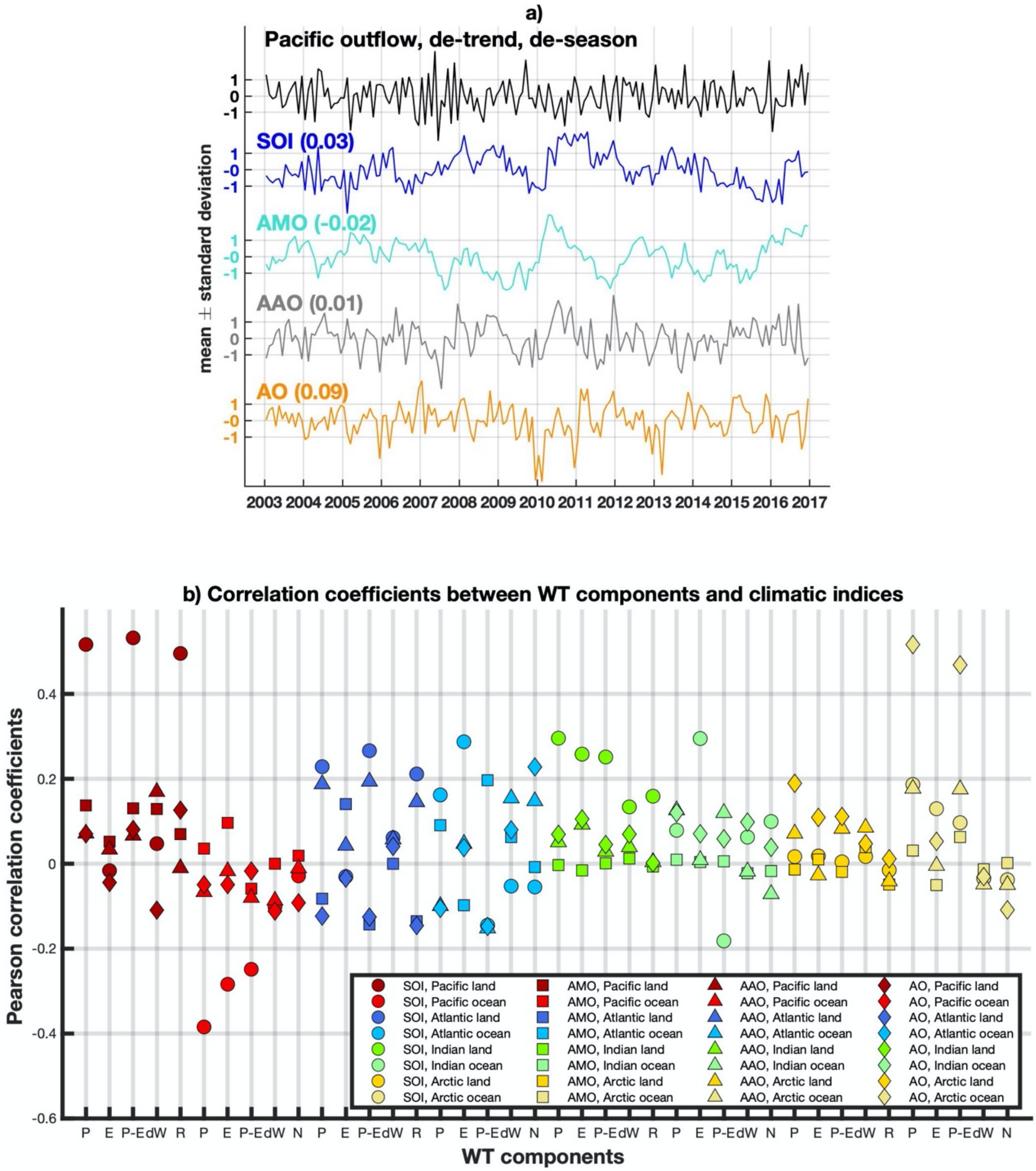
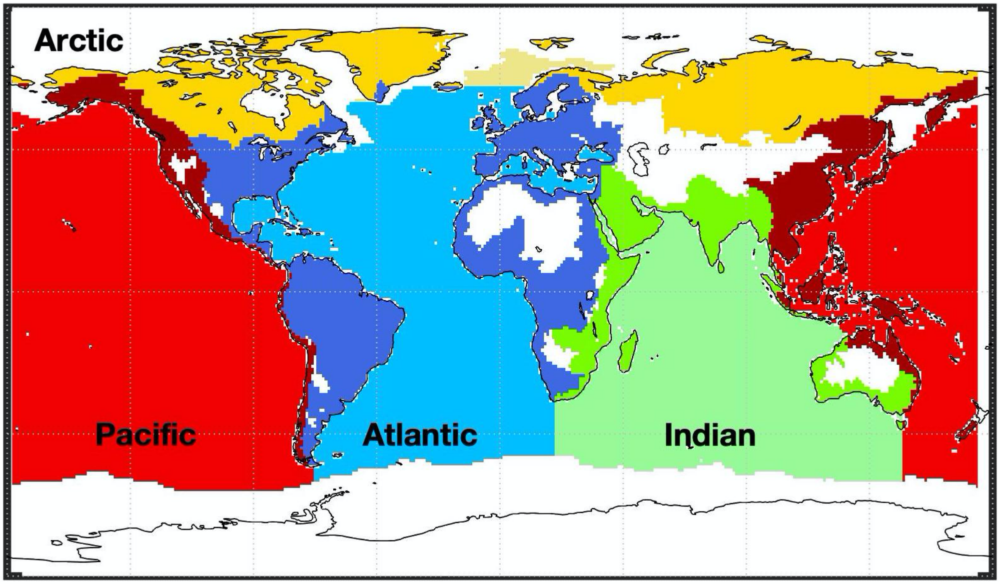
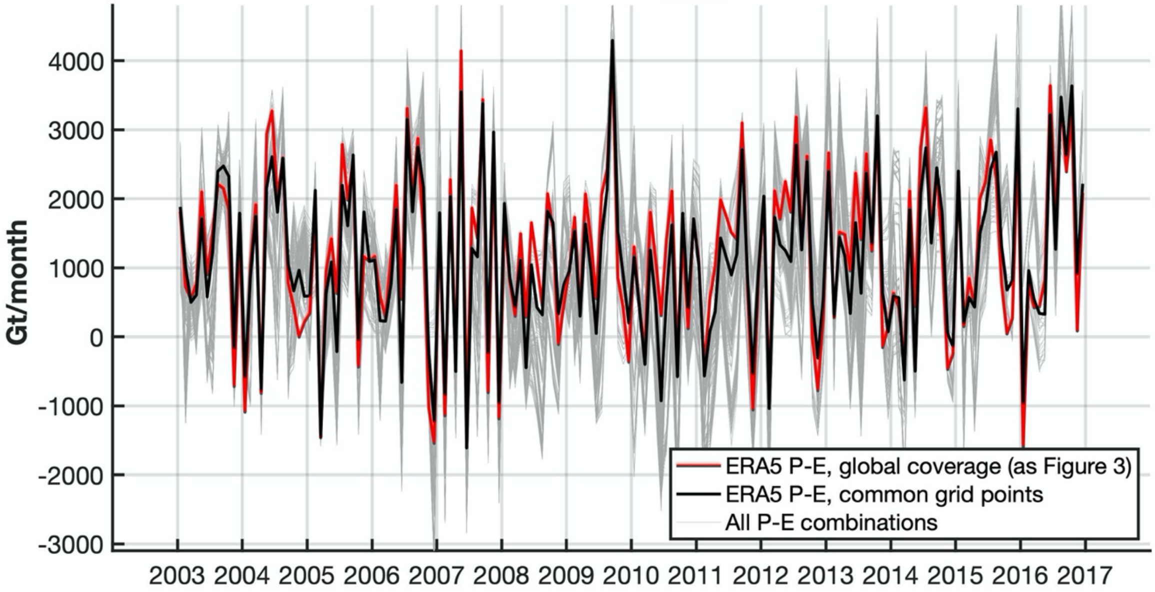


Figure 7. Pacific outflow and climatic indices for ENSO, AMO, AO, and AAO. a) Time series of Pacific outflow is de-trend and de-season. All time series are normalized to have unit variance. Values in the parenthesis are the correlation coefficient between the corresponding climatic index and the Pacific outflow. b) Correlation coefficients between de-trend and de-season WT components of different regions and the climatic indices.



a) Common spatial coverage to all P and E datasets



b) (-1) × $N_{Pacific}$

Figure 8. Monthly time series of (the opposite of) the Pacific outflow estimated from 162 combinations of *P* and *E* datasets. a) Spatial coverage common to all datasets. b) Pacific outflows: Gray thin curves are the 162 Pacific outflows estimated in the common grid points to all datasets (no global coverage); black and red curves are based on ERA5 *P* and *E* and are obtained using either only the grid points common to all datasets (black curve) or global coverage (red curve). Note that the red curve is the same as in Figure 3.

Table 1. Mean and annual signals of the *N* component as estimated from CSR mascon solution for different ocean basins according to Equation 2.

		Mean (CI_{95})	Annual signal (CI95)		
		(Gt/month)	Amplitude (Gt/month)	Phase (degree)	Peak date
Outflows	Pacific	1194 (1087, 1308)	809 (637, 975)	212 (200, 224)	August 3rd
	Arctic	723 (709, 738)	271 (242, 302)	234 (228, 240)	August 25th
	Pacific + Arctic	1917 (1826, 2010)	1061 (904, 1216)	217 (209, 225)	August 8th
Inflows	AIA	1194 (1086, 1304)	767 (610, 926)	212 (199, 224)	August 3rd
	Atlantic	926 (863, 991)	305 (219, 384)	249 (234, 266)	September 9th
	Indian	991 (911, 1067)	791 (664, 918)	205 (196, 214)	July 27th
	Atlantic + Indian	1917 (1821, 2015)	1020 (876, 1172)	218 (209, 226)	August 9th

Table 2. Mean of the *N* component as estimated from JPL mascon solution for different ocean basins according to Equation 2 . CI$_{95}$ are estimated as propagation of mascon errors provided by JPL, and from bootstrap analysis. Units are Gt/month.

		Mean (CI$_{95}$ from error propagation)	Mean (CI$_{95}$ from bootstrap)
Outflows	Pacific	1182 (1143, 1220)	1182 (1062, 1306)
	Arctic	735 (713, 757)	735 (711,761)
	Pacific + Arctic	1917 (1872, 1961)	1917 (1806, 2036)
Inflows	AIA	1183 (1092, 1274)	1183 (1077, 1282)
	Atlantic	919 (866, 972)	919 (845, 985)
	Indian	999 (980, 1018)	999 (928, 1067)
	Atlantic + Indian	1918 (1862, 1974)	1918 (1838, 2003)

Table 3. Correlation coefficients between SOI and de-trend and de-season WT components involved to estimate the Pacific outflow according to Equations 3 and 4.

	$std(X_i)$ (Stand. Deviation)	$corr(X_i, SOI)$ (Correlation between Xi with SOI)	$\frac{std(X_i)}{std(N)}$ (Coefficients)	$corr(X_i, SOI) \cdot \frac{std(X_i)}{std(N)}$ (Correlation ·Coefficient)
X1= –(P–E)ocean	605	0.25	0.57	0.14
X2= –(P–E)land	212	-0.53	0.20	-0.11
X3= dWland	96	0.048	0.09	0.004
X4= dWocean	711	-0.10	0.67	-0.07
Corr(N,SOI)				-0.03