Supporting Information for "Forced Changes in the Arctic Freshwater Budget Emerge in the Early 21st Century"

Alexandra Jahn[1], Rory Laiho[1]

[1]Department of Atmospheric and Oceanic Sciences and Institute of Arctic and Alpine Research, University of Colorado Boulder, 450 UCB, Boulder, 80309-0450, USA.

Contents of this file

Section S1. Climatological Arctic FW Budget in the CESM1.1

The Arctic FW budget is calculated from the CESM1.1 model simulations, relative to a reference salinity of 34.8. The surface fluxes (net precipitation and runoff) and the FW storage terms are calculated over the shaded region shown in Fig. 1a, which is delineated by Bering Strait, Barrow Strait, Nares Strait, Fram Strait, and the BSO (shown in Fig. 1a). The liquid FW fluxes through those gateways are given as net FW fluxes over the full depth of the gateways, and for solid FW combine the FW contained in sea ice and in snow on sea ice. The liquid FW storage shown throughout the paper is calculated down to the 34.8 isohaline, following previous conventions (Serreze et al., 2006; Haine et al., 2015). Davis Strait is shown for reference as an additional strait that is often used in Arctic freshwater studies (e.g., Haine et al., 2015; Wang et al., 2016b, 2016a; Shu et al., 2018), but is not part of the Arctic domain over which the FW budget is calculated.

Compared to the observational Arctic FW fluxes for the late 20th century, we find that the largest biases in the CESM1.1 FW budget compared to observations are found in the liquid FW exports from the Arctic Ocean. In particular the Fram Strait liquid FW export is much smaller than observed, while the net BSO FW flux is too large. The total solid FW exports on the other hand are slightly too large compared to the observations, except in Barrow Strait, where they are too small. However, the net simulated FW export from the Arctic (7066 km^3/yr) is within the observational uncertainty of the observed net FW export from the Arctic ($8324 \pm 1263 km^3/yr$), so the biases in the fluxes represent a combination of a bias in FW export routes (e.g., more FW export through the BSO, at the expense of the Fram Strait) and a bias between solid and liquid FW export (i.e., more solid FW export than observed, at the expense of the liquid FW export), rather than an overall too small FW exchange between the Arctic and North Atlantic. The bias in the liquid versus solid FW fluxes goes along with a larger than observed solid FW storage in the CESM1.1 (see Table S1), indicating that in the late 20th century more FW is stored in the solid versus liquid component in the CESM1.1 compared to observations. Note that while this means there is more solid FW stored in the CESM1.1 over 1980–2000 than observed, the simulated decrease in the solid FW storage over the first decade of the 21st century is not too large and agrees well with estimates based on PIOMAS: Haine et al. (2015) found a decrease of 6,900 km^3 in the solid FW storage based on PIMOAS between 1980–2000 and 2011, compared to 6,387 km^3 in the ensemble mean from the CESM1.1 if calculated over the same period and domain as used in Haine et al. (2015) (using Davis Strait rather than Nares and Barrow Straits as boundary west of Greenland; for the smaller Arctic domain used here the simulated decrease over this period is slightly less, at 5,869 km^3). Hence, the liquid FW storage increase in the CESM1.1 over the early 21st century that leads to the simulated emergence is not unduly driven by a concurrent too large decline in the solid FW storage over the early 21st century period. Eventually, however, the bias in the solid FW storage over the historical period will lead to a too large contribution from sea ice melt, compared to the real world.

While there are clear biases in the CESM1.1, it is important to note that limited observations make it challenging to even know what some of the details of the Arctic FW budget should be (as also discussed by Haine et al., 2015; Lique et al., 2016). In particular, the liquid FW export west of Greenland has in the past been assumed to be strongly dominated by Barrow Strait/Lancaster Sound, based on the available data at the time (Jahn et al., 2012). However, more recent data from Nares Strait has raised the expected contribution from Nares Strait liquid and solid FW export, due to the inclusion of the surface layer, as well as revealed large, previously unknown interannual variability (Münchow, 2016). This new data suggests that the two main channels west of Greenland may in fact be exporting approximately equal amounts of FW from the Arctic (Table S1). Similarly, there is a wide range of estimated solid FW storage (Haine et al., 2015), due to uncertain Arctic wide sea ice thickness data, in particular prior to the 2000s.

Section S2. IVT Sensitivity to Different Threshold Choices

We here chose an IVT of ±3.5 standard deviations, as for normally distributed processes the range between the upper and lower IVT captures 99.95% of values due to unforced internal variability. For most of the FW budget terms, this means that all values in the 1800 year long control simulation fall within this ±3.5 standard deviation range. However, for a few terms (Fram Strait liquid FW flux, runoff, liquid FW storage, and solid BSO), the IVT threshold is crossed a few times during the 1800 years of the control simulation (Fig. S1). Such isolated occurrences outside the ±3.5 standard deviation over 1800 years are consistent with the fact that individual very rare (<0.05% probability) events can potentially lead to departures outside the IVT range, even for an IVT range of ±3.5 standard deviations. For the BSO solid term, it is also a reflection that this flux is not normally distributed (as it is close to but does not cross the zero line), so different probabilities apply; however, this term is small and it is only included for completeness as part of the Arctic FW budget. All results presented also generally hold if we do not assume normally distributed processes but instead use the maximum and minimum values of each FW budget terms during the 1800 year long control simulation plus an

0094-8276/20/$5.00

extra margin of 10% of the flux to exclude any unsampled rare natural variability-driven events (Fig. S3).

As we are using at least 11 ensemble members for the 20th and 21st centuries, it would be extremely unlikely to see rare events with a probability of <0.05% occurring for all ensemble members over the 181 years of the 20th and 21st centuries simulation. This means that the detection of spurious complete shifts is highly unlikely. Spurious emergence is not statistically possible, as emergence requires sustained changes outside the pre-industrial IVT range. Hence, this methodology and IVT choice is able to detect truly forced changes in the Arctic FW budget terms. Smaller/larger thresholds than 3.5 standard deviations and a non-gaussian approach lead to qualitatively similar results, but some changes in the specific shifts and emergence years due to the change of the probability of events outside the chosen range (see Fig. S3).

Note that our methodology to determine time of emergence differs from several other "time of emergence" methods (e.g., Hawkins & Sutton, 2012; Mora et al., 2013; Lehner et al., 2017). In particular, we look at annual mean values outside the IVT range rather than considering when the ensemble mean first exceeds the background variability by a certain factor (a typical signal/noise ratio definition of emergence). This approach is most similar to the determination of shift and emergence of Arctic open water days in Barnhart, Miller, Overeem, and Kay (2015), who demonstrated that there can be substantial differences between the emergence time of a variable's ensemble mean versus its unsmoothed trajectory. As we want to be able to assess when we can expect to observe fluxes and storage that are fully outside the background state, we prefer this time of emergence methodology of using the unsmoothed variables, as that is what we will be able to observe in the real world.

Section S3: Effect of a shorter base period and of sampling a non-steady state system

To provide insights into how shift and emergence detection would look different for observations of the Arctic FW budget, we have repeated our emergence analysis for 20-year periods from the control as well as from the historical simulation (Fig. S4). This allows us to assess how the results presented here are affected by using a shorter base period as well as a base period that covers a period where forced changes are starting to affect some of the budget terms. We find that the shorter base period by itself does not affect the main results on emergence, but does change the start and end years by a few years (see Fig. S4b and c versus Fig. S4a). Shift periods on the other hand are more strongly affected by a shorter base period, with some changes of several decades in either direction. This behavior is expected, as emergence detects a sustained, forced change while shift is triggered by an individual event, so a small change in the IVT will affect shift more strongly than emergence. Sampling a non-steady state system for 20 years, however, has a big effect on detecting emergence. Emergence patterns similar to the ones based on the full length of the control simulation are found primarily for a 20 year period from the historical simulation that ends before 2000 (see Fig. S4a, d, g). Once the base period extends past 2000, emergence is reached later, in particular for the terms that show early emergence (Fig. S4e, f, h, i). Nonetheless, the general order of emergence of FW budget terms remains the same even for base periods that extend to 2009. For base periods that extend past 2009, however, even the order of emergence changes, as the base period from 2000-2019 now samples the already very different FW storage terms, leading to a much later emergence of these terms compared to their already very different base state. Hence, it is the changing nature of the FW budget terms during the base period, rather than the much shorter base period itself, that complicates the diagnosis of emergence and shift from observations that, at best, extend to the mid or late 1990s.

References

Barnhart, K. R., Miller, C. R., Overeem, I., & Kay, J. E. (2015). Mapping the future expansion of Arctic open water. *Nature Clim. Change*, *6*, 280–285. doi: 10.1038/nclimate2848

Haine, T. W., Curry, B., Gerdes, R., Hansen, E., Karcher, M., Lee, C., ... Woodgate, R. (2015). Arctic freshwater export: Status, mechanisms, and prospects. *Global and Planetary Change*, *125*, 13 - 35. doi: 10.1016/j.gloplacha.2014.11.013

Hawkins, E., & Sutton, R. (2012). Time of emergence of climate signals. *Geophys. Res. Lett.*, *39*(1). doi: 10.1029/2011GL050087

Jahn, A., Aksenov, Y., de Cuevas, B., de Steur, L., Häkkinen, S., Hansen, E., ... Zhang, J. (2012). Arctic Ocean freshwater - How robust are model simulations? *J. Geophys. Res.*, *117*. doi: 10.1029/2012JC007907

Kwok, R., Maslowski, W., & Laxon, S. W. (2005). On large outflows of Arctic sea ice into Barents Sea. *Geophys. Res. Lett.*, *32*. doi: 10.1029/2005GL024485

Lehner, F., Deser, C., & Terray, L. (2017). Toward a new estimate of time of emergence of anthropogenic warming: Insights from dynamical adjustment and a large initial-condition model ensemble. *J. Climate*, *30*(19), 7739-7756. doi: 10.1175/JCLI-D-16-0792.1

Lique, C., Holland, M. M., Dibike, Y. B., Lawrence, D. M., & Screen, J. A. (2016). Modeling the Arctic freshwater system and its integration in the global system: Lessons learned and future challenges. *J. Geophys. Res. Biogeosciences*, *121*(3), 540-566. doi: 10.1002/2015JG003120

Mora, C., Frazier, A. G., Longman, R. J., Dacks, R. S., Walton, M. M., Tong, E. J., ... Giambelluca, T. W. (2013). The projected timing of climate departure from recent variability. *Nature*, *502*, 183-187. doi: 10.1038/nature12540

Münchow, A. (2016). Volume and freshwater flux observations from Nares Strait to the West of Greenland at daily time scales from 2003 to 2009. *J. Phys. Oceanogr.*, *46*(1), 141-157. doi: 10.1175/JPO-D-15-0093.1

Prinsenberg, S., & Hamilton, J. (2005). Monitoring the volume, freshwater and heat fluxes passing through Lancaster Sound in the Canadian Arctic Archipelago. *Atmos.-Ocean*, *43*(1), 1-22.

Serreze, M. C., Barrett, A. P., Slater, A. G., Woodgate, R. A., Aagaard, K., Lammers, R. B., … Lee, C. M. (2006). The large-scale freshwater cycle of the Arctic. *J. Geophys. Res.*, *111*. doi: 10.1029/2005JC003424

Shu, Q., Qiao, F., Song, Z., Zhao, J., & Li, X. (2018). Projected freshening of the Arctic Ocean in the 21st century. *J. Geophys. Res: Oceans*, *123*(12), 9232-9244. doi: 10.1029/2018JC014036

Wang, Q., Ilicak, M., Gerdes, R., Drange, H., Aksenov, Y., Bailey, D. A., … Yeager, S. G. (2016a). An assessment of the Arctic Ocean in a suite of interannual CORE-II simulations. Part II: Liquid freshwater. *Ocean Modelling*, *99*, 86 - 109. doi: 10.1016/j.ocemod.2015.12.009

Wang, Q., Ilicak, M., Gerdes, R., Drange, H., Aksenov, Y., Bailey, D. A., … Yeager, S. G. (2016b). An assessment of the Arctic Ocean in a suite of interannual CORE-II simulations. Part I: Sea ice and solid freshwater. *Ocean Modelling*, *99*, 110-132. doi: 10.1016/j.ocemod.2015.12.008

Woodgate, R. A., & Aagaard, K. (2005). Revising the Bering Strait freshwater flux into the Arctic Ocean. *Geophys. Res. Lett.*, *32*. doi: 10.1029/2004GL021747

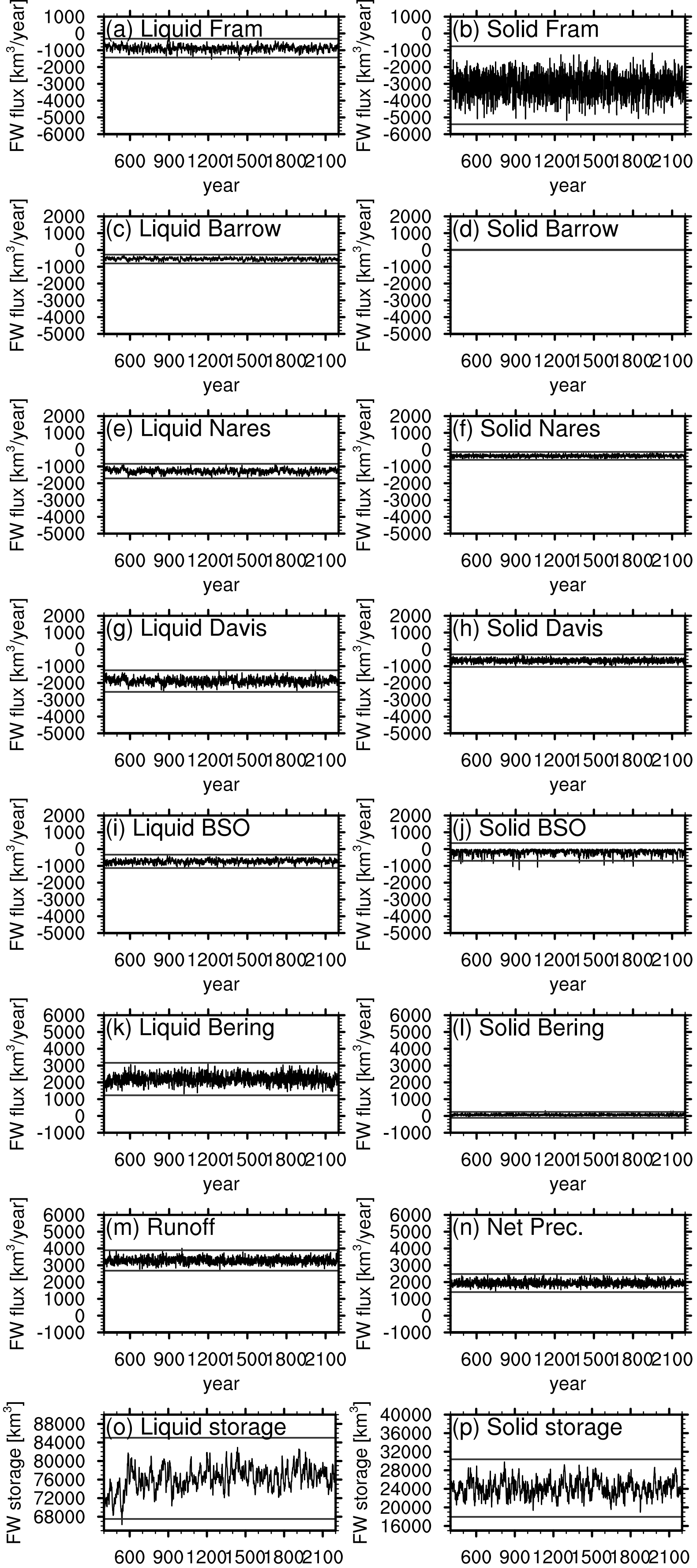

Figure S1: **Variability in the control simulation**. The ±3.5 standard deviation threshold for each variable (which is used to determine shift and emergence in the 20^{th} and 21^{st} centuries simulations) is shown as solid dark grey lines. Flux terms (a-n) and storage terms (o, p) are labeled in the panels. Note that all flux panels (a-n) and all storage panels (o-p) each have the same y-axis range, but that the axis are offset from each other.

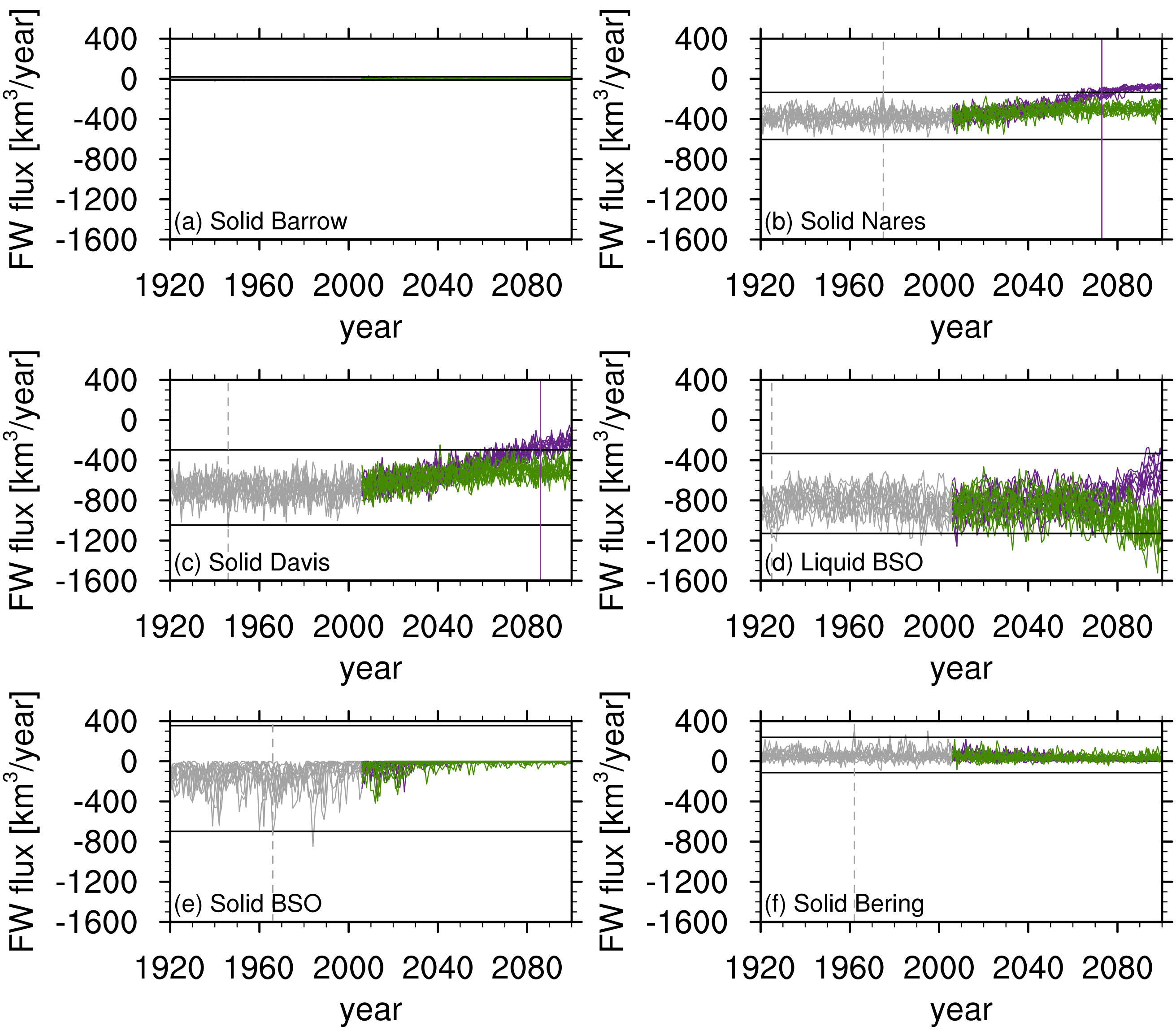

Figure S2: **Small FW budget terms over time.** As in Fig. 2, but for the small (less than 300 km^3/yr in the observed net fluxes) FW fluxes not shown in Fig. 2. Note that the y-axis is the same for all panels, but is different from Fig. 2 to allow a more meaningful depiction of these small fluxes.

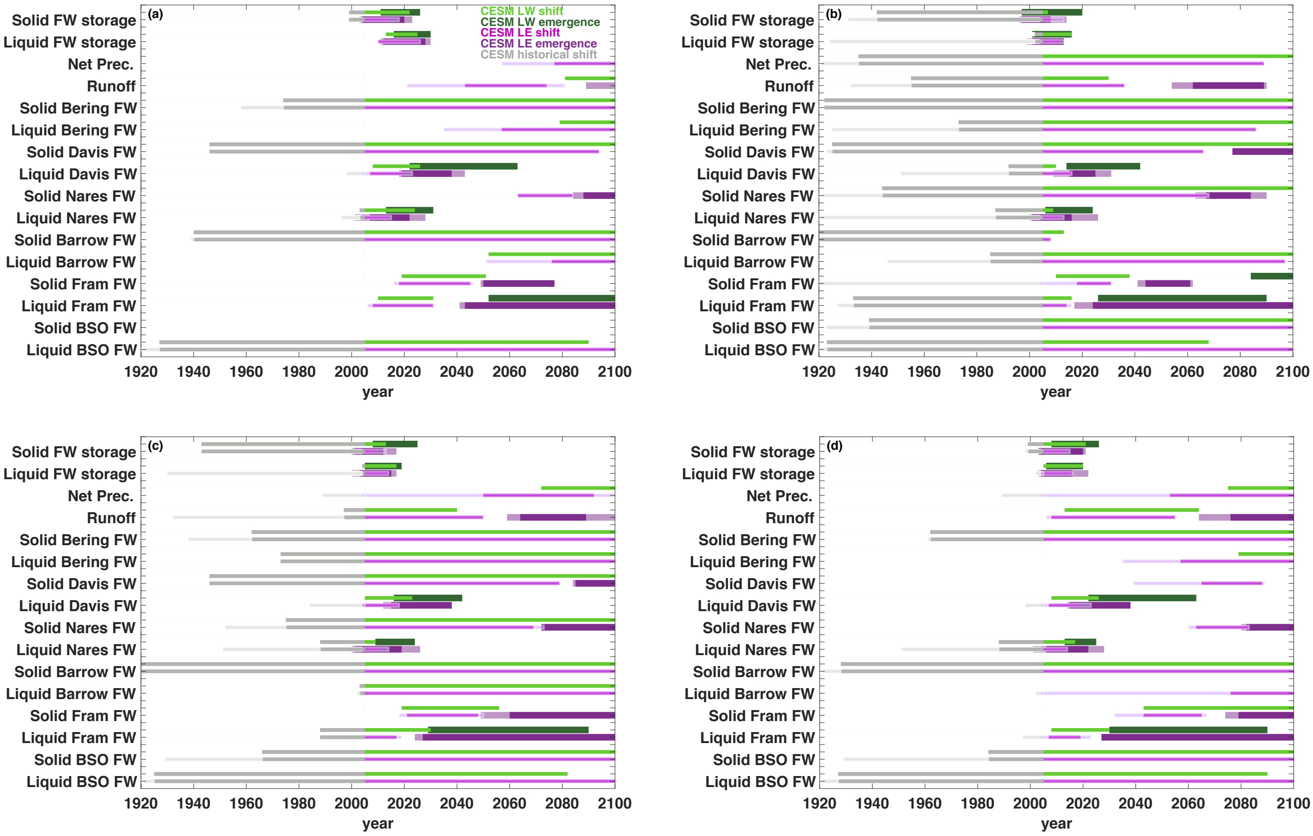

Figure S3: **Sensitivity of results to different IVT choices.** As Fig. 3, but for an IVT defined as (a) the maximum/minmum values in the control ±10% of the mean, (b) ±3 standard deviations, (c) ±3.5 standard deviations, and (d) ±4 standard deviations. This figure also includes the FW fluxes with a net observed flux of less than 300 km^3/yr, which were not shown in Fig. 3. These different IVT choices (a, b, d) show qualitatively similar results as for 3.5 standard deviations (c), with the largest changes primarily in the start dates of the shift periods, due to the smaller/larger IVT range. None of the main conclusions are affected by the choice of the IVT, as they mainly focus on the emergence of the forced signal.

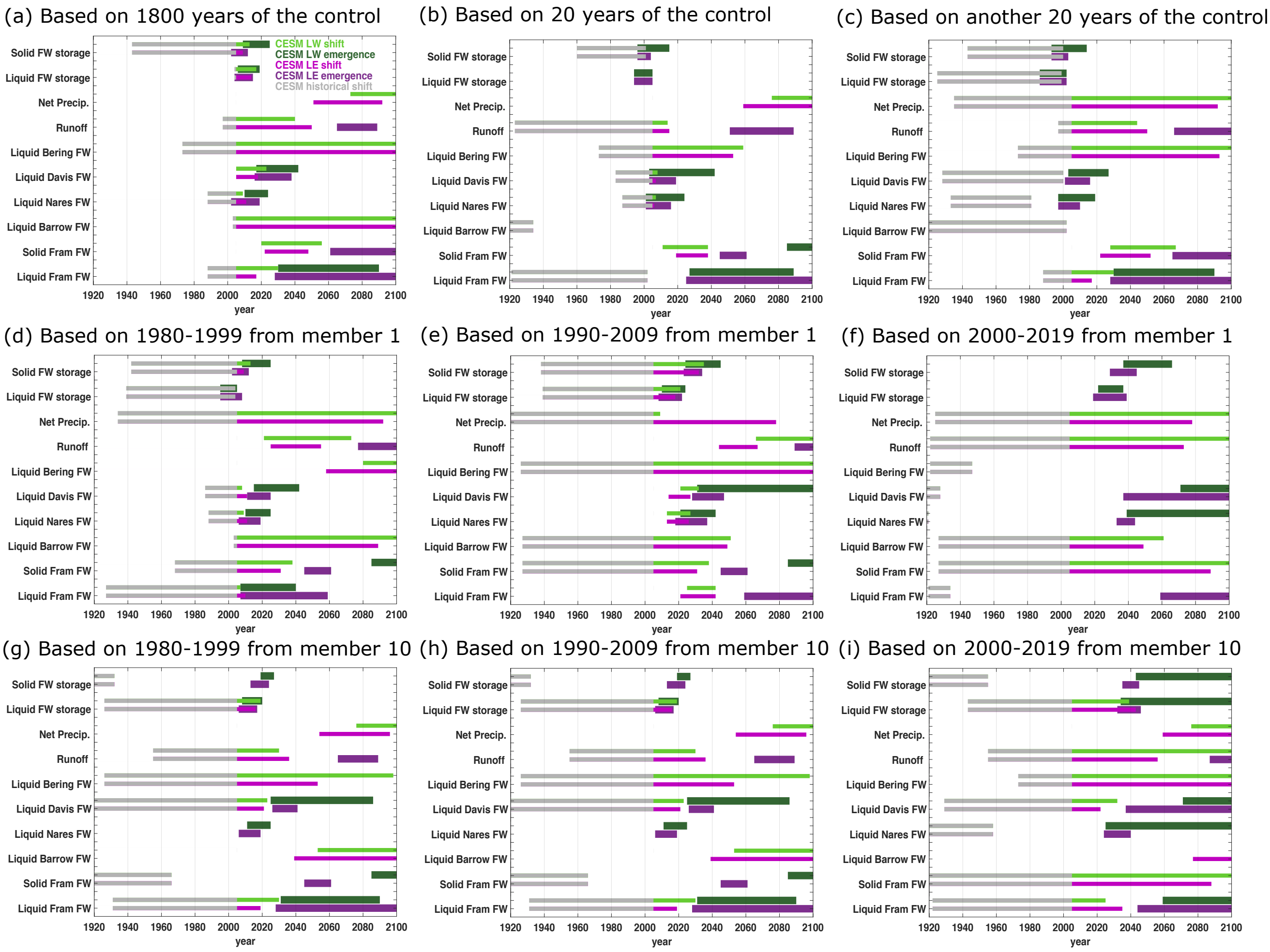

Figure S4: **Influence of different base periods on shift and emergence.** Shift and emergence, as shown in Fig. 3, but using different base periods in the different panels to determine the IVT, to assess the impact of a shorter base period, as would be available from observations. (a) Using the full 1800 years of the control simulation, same as Fig. 3, (b) using only 20 years of the control (here years 1000-1019), (c) using a different set of 20 years from the control (here years 400-419), (d/g) using years 1980-1999 from ensemble member 1/10 of the CESM LE, (e/h) using years 1990–2009 from ensemble member 1/10, (f/i) using years 2000-2019 from ensemble member 1/10. This shows that sampling of a system more and more affected by climate change if years after 2000 are included in the base period affects the results more than using a shorter base period, in particular for emergence (with shift sensitive to both). Results are similar for 30 year instead of 20 year periods. Members 1 and 10 are shown in panels d–f and g–i, respectively, to illustrate the effects of sampling different 20-yr periods under the same external forcing but with different internal variability. Other ensemble members show slightly different patterns, but changes are qualitatively similar to the difference between the two members shown here.

Table S1: **Climatological Arctic Ocean freshwater (FW) budget (1980–2000)**. Observational values are partially taken from the compilation by Serreze et al. (2006) (indicated by * in the table). Terms that are not from Serreze et al. (2006) are: Bering Strait solid FW fluxes (Woodgate & Aagaard, 2005), BSO solid FW fluxes (Kwok et al., 2005), Nares Strait liquid and solid FW fluxes (Münchow, 2016), Barrow Strait solid and liquid fluxes (Prinsenberg & Hamilton, 2005), and Davis Strait solid and liquid FW fluxes (Haine et al., 2015). The solid FW storage in the Arctic is shown as range, based on the values given in Serreze et al. (2006) and Haine et al. (2015). These two estimates differ in the assumed mean ice thickness (thinner ice assumed in Serreze et al. (2006) than Haine et al. (2015)) as well as in their Arctic domain, with the Arctic domain in Serreze et al. (2006) smaller than our domain (entirely excluding the CAA) and the domain in Haine et al. (2015) larger than our domain (including Baffin Bay down to Davis Strait). In the CESM1.1, the impact of these domain differences compared to the Arctic domain used here is an additional solid FW storage of 1,868 km3 for the domain of Haine et al. (2015) and 2988 km^3 less solid FW storage for the domain of Serreze et al. (2006), which does not change the fact that the CESM1.1 has too much solid FW storage. However, note that the solid FW flux and storage includes FW from the snow on sea ice as well as from the ice itself while the observational estimates typically only include the FW in the sea ice, which leads to a difference of about 10%. Further note that the Nares and Barrow Strait values are from the early 2000s, rather than the late 20^{th} century, as no earlier data exists. If available, error estimates for the observations are included. Model values show the ensemble mean values, and the ± indicates the standard deviation of the 40-member CESM LE in the 21-yr averages. All FW fluxes are quoted in km^3/year, and the FW storage is quoted in km^3. All values are annual mean net fluxes, for oceanic fluxes over the full depth of each channel, combining negative and positive fluxes through a strait, where applicable. Positive values indicate FW sources and negative values indicate FW sinks for the Arctic Ocean. Note that Davis Strait is included here for reference only, with the surface fluxes and storage calculated over the Arctic Ocean domain delineated by Nares Strait and Barrow Strait west of Greenland (see Fig. 1a).

FW fluxes	**Observations**	**CESM LE**
River runoff	3200 ±110*	3358±55
Net precipitation	2000 ±200*	1958±32
Bering Strait liquid FW	2400 ±300*	2159±66
Bering Strait solid FW	140 ±40	56±14
Barrow Strait liquid FW	−1510	−567±25
Barrow Strait solid FW	−76	2±1
Nares Strait liquid FW	−1356±236	−1439±69
Nares Strait solid FW	−252±63	−395±15
Davis Strait liquid FW	−3200±320	−2044±69
David Strait solid FW	−160	−701±24
Fram Strait liquid FW	−2700 ±530*	−948±68
Fram Strait solid FW	−2300 ±340*	−2776±174
BSO liquid FW	−90 ±94*	−852±50
BSO solid FW	−40	−91±41
Liquid FW storage	74,000 ±7400*	77,485±1562
Solid FW storage	10,000* – 17,800	21,931±1011