S1 AIC and BIC

In Fig. 1 in the paper, we showed the derivative of the AIC and BIC relative to the number of mixtures versus the number of mixtures. Here, we show the AIC and BIC versus the number of mixtures (Fig. S1). As discussed in the main text, we chose the number of mixtures based on where the gradient of the criteria relative to the number of mixtures flattens. For robustness, we also tested the method with a range of different mixture number choices and found the results were insensitive to this choice (see Table 2).
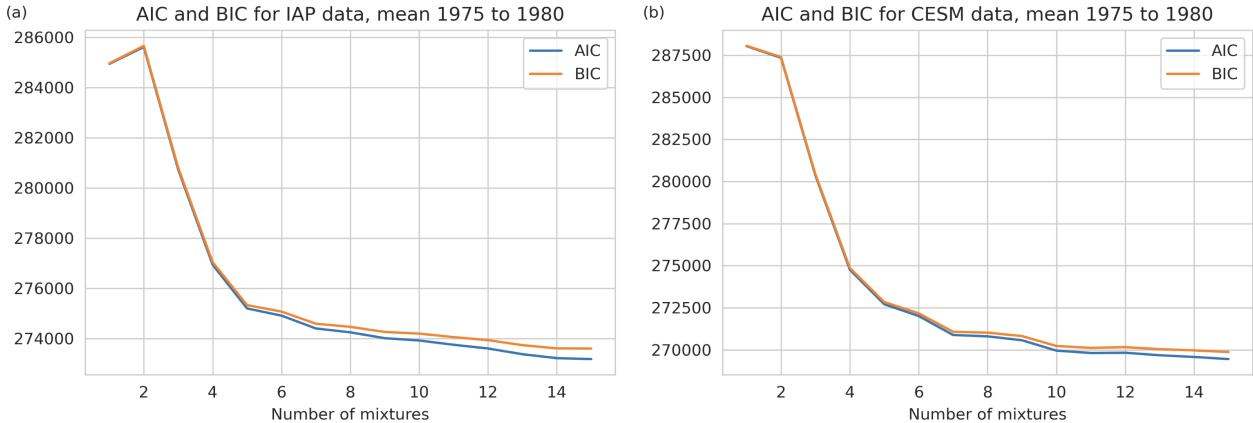
Figure S1: The AIC and BIC versus number of mixtures. Panel a uses the IAP surface salinity distribution south of 65N over the period 1975 to 1980, while panel b uses the CESM ensemble mean surface salinity distribution over the same region and time period.

S2 Equivalents to Figures 4 and 5 with IAP regions

In the main text, we evaluated the hypothesis that (heat flux induced) circulation change imprints regionally on the surface salinity pattern using regions found by a GMM fit to CESM ensemble mean data (Fig. 4). We also used these regions to evaluate the linearity assumption that the separate imposition of freshwater flux, heat flux and wind stress forcings are approximately equivalent to all forcings imposed at once (Fig. 5). Here, we show analogous plots but for regions found by the GMM fit to the IAP data.

In Fig. S2, we show the change in salinity and temperature due to individual forcings in regions found by a GMM fit to the IAP data. Similar to the plot with regions based on the GMM fit to CESM data (Fig. 4), we find that heat fluxes have the strongest imprint in the saltiest regions (mixture 6) while freshwater fluxes makes fresher (saltier) regions fresher (saltier).

In Fig. S3, we assess the linearity assumption that the separate imposition of forcings ('faf-stress+faf-water+faf-heat') is equivalent to all forcings applied at once ('faf-all') using the regions found by a GMM fit to the IAP data. As in text (Fig. 5), we find that there is good agreement between the sum of individual forcings and forcings applied together except for in the sixth mixture for the HadOM3 model, where there is a non-linear response.

S3 Change in temperature in each region due to different forcings

In the main text, we showed the change in salinity in each region between the end of forced experiments and the end of control run for the three ocean-only FAFMIP models (Fig. 4). Here, in Fig. S4, we show the same results for the surface temperature. We find that different forcings affect the pattern of surface temperature distinctly. However, unlike salinity, the scale of change due to temperature is an order of magnitude larger than other forcings, whereas for salinity the scale of change due to heat flux and freshwater flux are the same in the saltiest regions.

S4 Additional testing on individual CESM members

In the main text, we showed the comparison of the true freshwater fluxes, as determined by the target metric defined in Eq. (10), and the linear response theory results for ensemble members for which the salinity trends met the significance criteria. Here, in Fig. S5, we show the application to all ensemble members regardless of
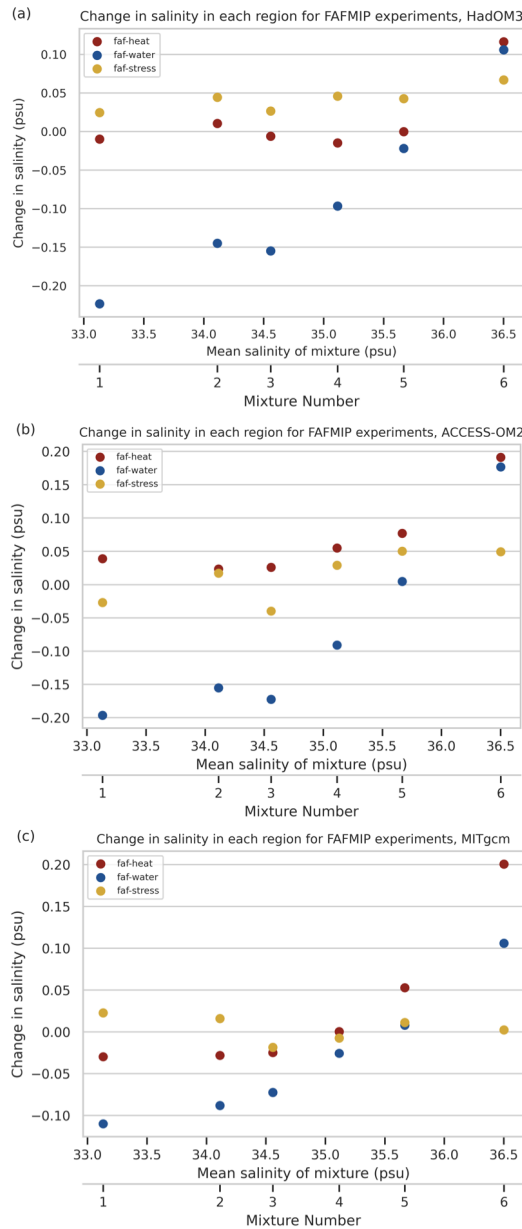
Figure S2: The change in surface salinity in each region from the GMM applied to the IAP data (Fig. 3d) for each individual forcing experiment. The heat flux experiment (faf-heat) is shown in red, freshwater flux (faf-water) in blue, and wind stress perturbation (faf-stress) in yellow. The response is defined as the difference between the last decade of a forced run and the last decade of the control run. Panel a-c show the results for ocean models HadOM3, ACCESS-OM2, and MITgcm respectively.

the significance of their salinity trends. We see that many members, but not all, capture the true response. As shown in text, if we restrict to members with sufficiently large signal to noise ratios, the correct response can be recovered.

We also test this method on the CESM ensemble data over the period 2011 to 2055. The forcing over this period is stronger, and thus more ensemble members meet the significance criteria determined in the main text. In Fig. S6, we show a comparable plot to Fig. 9 in the main text but over the period 2011 to 2055. Here, we apply the salinity trend significance criteria and exclude 3 members. We find that the true response is captured with the exception of members 4, 11, and 34 – thus, it is captured in 90.3% of members.
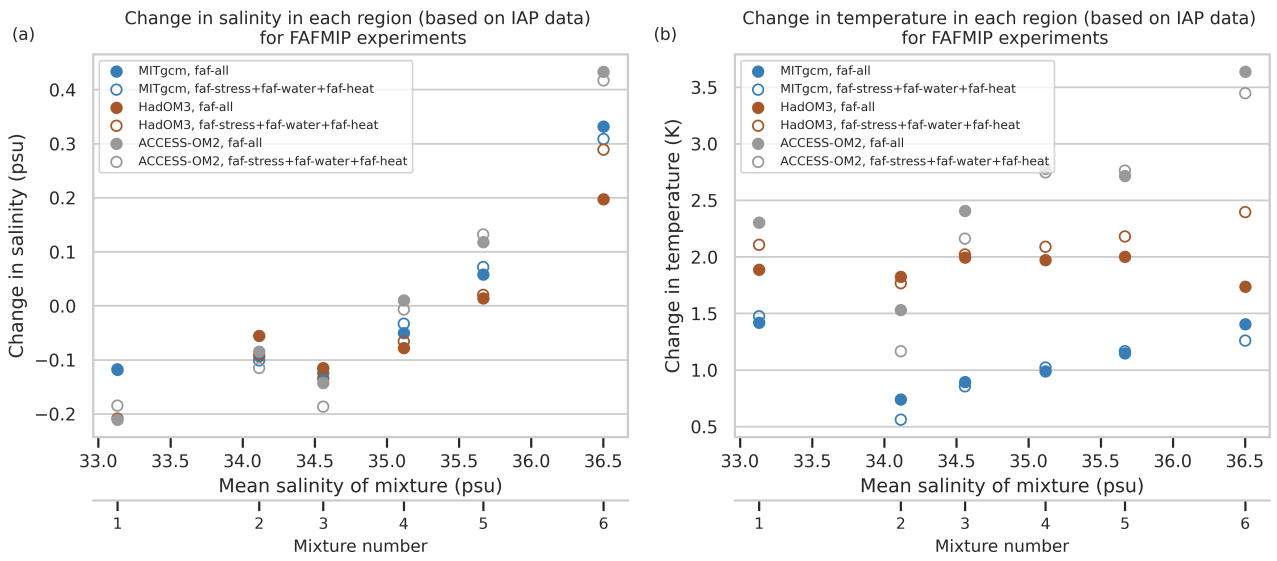
Figure S3: Comparing the surface salinity response (a) and surface temperature response (b) in each region between two cases: the sum of the response when forcings are applied individually (faf-stress+faf-water+faf-heat) and the response to all forcings applied at once (faf-all). The response is defined as the difference between the last decade of a forced run and the last decade of the control run. The response is largely linear with the exception of the sixth mixture for the HadOM3 model. Here the regions are found by a GMM fit to the IAP data.
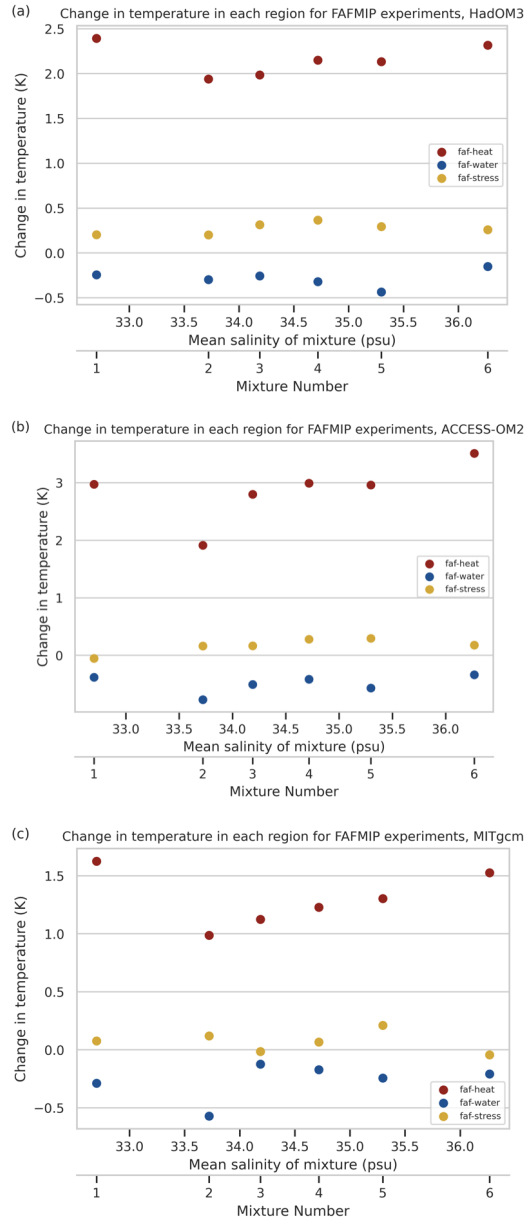
Figure S4: The change in surface temperature in each region from the GMM applied to the CESM data for each individual forcing experiment. The heat flux experiment (faf-heat) is shown in red, freshwater flux (faf-water) in blue, and wind stress perturbation (faf-stress) in yellow. The response is defined as the difference between the last decade of a forced run and the last decade of the control run. Panel a-c show the results for ocean models HadOM3, ACCESS-OM2, and MITgcm respectively.
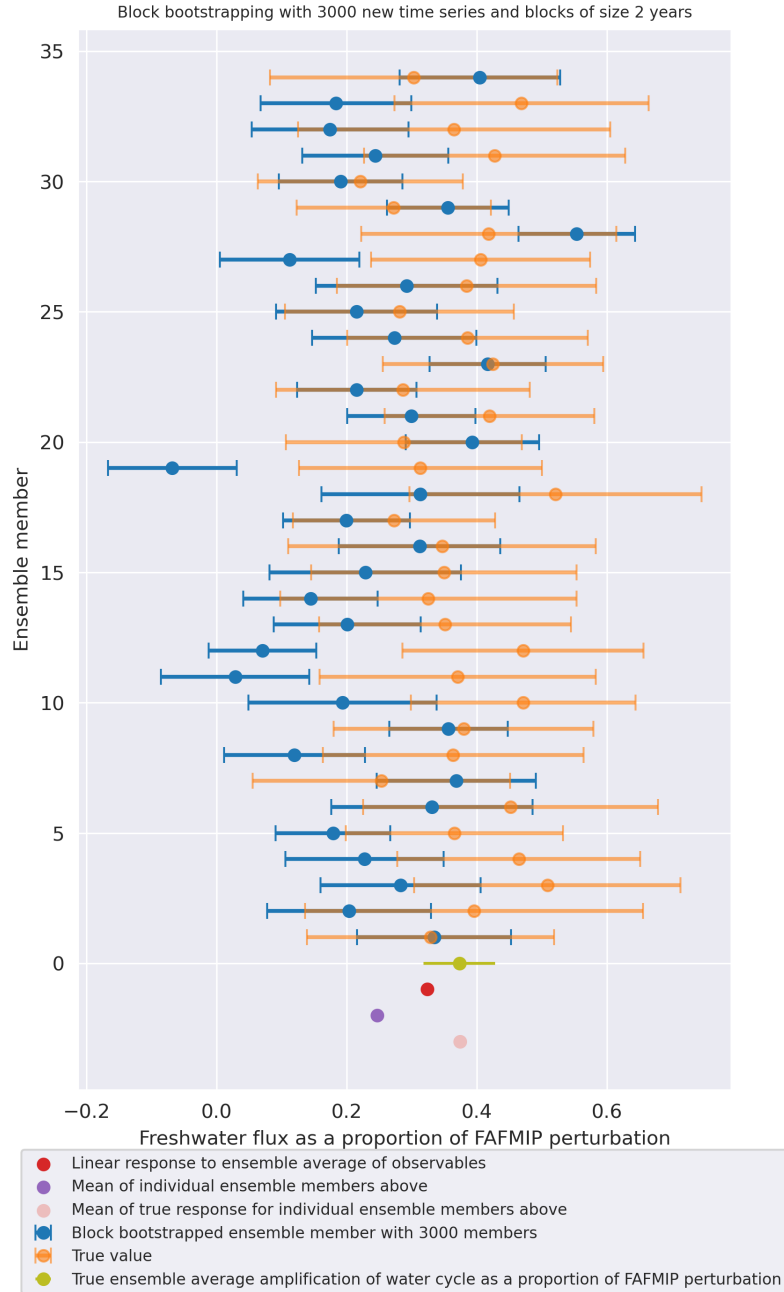
Figure S5: Freshwater flux responses over the period 1975 to 2019 for individual CESM ensemble members following our methodology described in text (blue) compared to truth from E-P model fields (orange). This plot can be compared to Fig. 9 in the main text which plots only members which met the significance criteria for salinity trends.
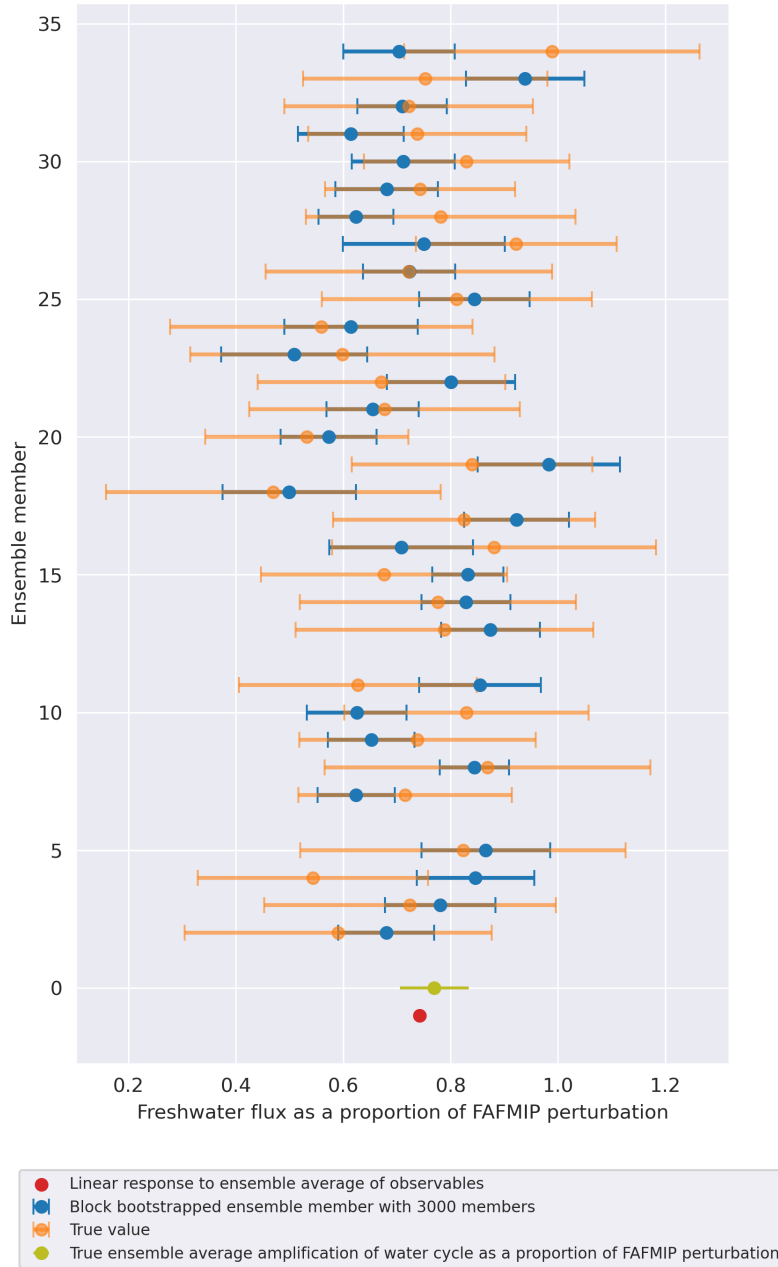
Figure S6: Freshwater flux responses over the period 2011 to 2055 (RCP8.5 pathway) for individual CESM ensemble members following our methodology described in text (blue) compared to truth from E-P model fields (orange). Here, we only plot members which met salinity significance criteria, and thus dropped 3 members.