Catalogue of Strong Nonlinear Surprises in ocean, sea-ice, and atmospheric variables in CMIP6

Joran R. Angevaare[1] and Sybren S. Drijfhout[1,2,3]

[1]Royal Netherlands Meteorological Institute (KNMI), de Bilt, The Netherlands
[2]Institute for Marine and Atmospheric Research, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands
[3]Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, SO14 3ZH, UK

Correspondence: Joran R. Angevaare (joran.angevaare@knmi.nl)

Abstract. The Coupled Model Intercomparison Project Phase 6 (CMIP6) archive was analysed for the occurrence of Strong Nonlinear Surprises (SNS) in future climate-change projections. To this end, we built an automated detection algorithm to identify SNS in a reproducible manner. Two different types of SNS were addressed: abrupt changes measured over decadal timescales and state transitions developing over multiple decades, too large to be explained by the forcing without invoking strong internal feedbacks in the climate system. We also discuss when one SNS case features a cascade from one variable to another (where sea-ice, deep convection, and temperature and salinity changes interact) and when two SNS interact to form a cascade, which occurs for the subpolar mixed-layer collapse triggering an Atlantic Meridional Overturning Circulation (AMOC) shutdown. Data of 54 models were analysed for five shared socio-economic pathways for ocean, sea ice, and atmospheric variables that are involved in the coupling to ocean and sea ice. The algorithm isolates regions of at least $10^6\,\mathrm{km}^2$ and utilizes stringent criteria to select SNS. In total 73 SNS were found, divided in 11 categories of which 4 apply to abrupt change and 7 to state transitions. Of the identified SNS $45\,\%$ relate to sea-ice cover, $19\,\%$ to ocean currents, $29\,\%$ to mixed-layer depth, and $7\,\%$ to atmospheric systems like the Intertropical Convergence Zone. For each category, probability density functions for time-windows of maximal change indicate SNS occurring earlier and at lower global temperature rise than assessed in previous reviews, in particular the ones associated with winter Arctic Sea ice disappearance, northern North Atlantic winter mixed-layer collapse and subsequent transition of the AMOC to a weak state in which the cell associated with North Atlantic Deep Water involved has vanished.

1 Introduction

Based on a literature review and an expert elicitation workshop the concept of Climate Tipping Points was introduced and formalized by Lenton et al. (2008). They recognized that elements of the Earth System can qualitatively change their mode of operation when passing a critical threshold, the latter being popularized as a tipping point. Recent updates of this overview were given in Armstrong McKay et al. (2022), relating the possible tipping to the 1.5 °C global warming threshold, and in Loriani et al. (2023) based on a new assessment by a team of experts. After 2008 a large body of literature appeared discussing specific climate tipping points and subsequent socio-economic tipping points, associating the climate tipping point concept with anthropogenic climate change, its impacts and risks, and policy (Kopp et al., 2025).

In the latest Report (AR6) of the Intergovernmental Panel for Climate Change (IPCC), the focus was slightly altered from tipping points to potential abrupt and irreversible change of Earth System elements (Lee et al., 2021), although the related definition of a tipping point was also used, recognizing the difficulty in proving the existence of a tipping point other than in often simplified climate or Earth System element models. One step earlier than what now can be seen as the "tipping point controversy" was taken in Drijfhout et al. (2015), where a catalogue of abrupt changes occurring in simulations of the Climate Model Intercomparison Project Phase 5 (CMIP5) (Taylor et al., 2012) was presented. The naming "abrupt shifts" was chosen deliberately, to avoid discussion whether a tipping point was involved, and if so, where the tipping point exactly was located.

Since Drijfhout et al. (2015) a new CMIP6 has been carried out and a new IPCC AR6 Report has been written, and it seems pertinent to also develop a new catalogue based on CMIP6 simulations. These simulations are driven by new emissions and land use scenarios based on new pathways of societal development (shared socio-economic pathways or SSPs) (O'Neill et al., 2016). CMIP6 simulations thus differ from previous CMIP projections, being produced with updated climate models as well as being driven with updated emissions and land-use scenarios. In addition, because new updates and overviews of tipping points have been published since CMIP5 (Armstrong McKay et al., 2022; Loriani et al., 2023) it seems pertinent to benchmark these overviews with how CMIP6 models simulate the associated events. Also, a recent paper by Terpstra et al. (2024) followed a slightly different approach and applied Canny edge detection (Canny, 1986) to CMIP6 data from model simulations forced with the rapid 1ptCO2 (1 % annual CO_2 increase) scenario, similar to the work of Bathiany et al. (2020) who applied similar methods to CMIP5 data. Apart from comparing our results with those of Terpstra et al. (2024), it is of interest to investigate whether CMIP6 models support recent publications on approaching tipping points for e.g. the Atlantic Meridional Overturning Circulation (AMOC) (Boers, 2021; Ditlevsen and Ditlevsen, 2023).

Here, we follow the same philosophy as in Drijfhout et al. (2015) and to avoid confusion about applying the term "abrupt change" to changes that involve a qualitative change in mode of operation but evolve over many decades, we will coin here the name Strong Nonlinear Surprises (SNS), as from CMIP runs we cannot deduce whether these SNS involve the passing of a critical threshold, or tipping point, or whether the SNS is reversible and if so, on which timescale. New with respect to Drijfhout et al. (2015) is that, with improved computer power and developments in the realm of Machine Learning, we have been able to adjust the method of finding and classifying SNS used in Drijfhout et al. (2015), to become fully automatic and reproducible. It should be noted that the classification criteria themselves always will contain an element of subjectivity that is open for discussion or adjustment, as there exists no fundamental theory or set of equations that uniquely defines those criteria. Instead, they are motivated to capture well-known tipping points as defined in the literature (e.g., Armstrong McKay et al. (2022); Loriani et al. (2023)), to detect model behaviour that is truly surprising and features large changes in climate variables and climate elements. With the word truly we mean that state changes in transient climate runs with changing forcing are by themselves trivial results. However, when such state change involves large additional effects from internal feedbacks and can no longer be explained by the change in forcing alone, we call this state change a state transition and a surprising effect, hence SNS. To clarify as such, the change in a certain subsystem must be much larger than witnessed elsewhere in the climate system and/or feature an almost total disappearance of that climate subsystem that is an essential ingredient of our present-day climate. In addition, using our automated and reproducible method other users can simply adjust these criteria to their own preference.

Also new is that for each SNS we indicate the moment (time window) of maximum change, depending on the timescale of smoothing applied to the relevant time series and timescale over which the change is calculated. Using this method, we construct for each SNS and each model a probability density function (PDF) that quantifies the conditional likelihood of the global warming level at which the maximum change occurs. This PDF is derived from applying multiple plausible smoothing and change windows to the same time series, to reflect the methodical uncertainty. We then average these conditional PDFs across all models in each category to obtain a category-level distribution. Importantly, these PDFs do not describe the probability that an SNS occurs in a model, but rather the distribution of the warming level at which the maximum change is detected, conditional on the presence of an SNS. With this approach we aim to prevent our results becoming subject to a politicized scientific debate and to avoid pitfalls of oversimplifying the concept of thresholds which, after being exceeded, would turn climate change into a new category of harmfulness. However, we do believe that SNS imply elements of climate change that generally contain larger, potentially more harmful changes than anticipated from current ongoing changes in climate. Cataloguing these changes as they occur in climate simulations, together with their projected timing of maximum change, should be of interest to the climate science community, as well as for a wider audience including policymakers. However, we emphasize, and discuss per case when relevant, that climate models contain severe biases and errors, and that both the type of SNS as well as its timing is subject to large uncertainties and often model-dependent. But if they do occur in model simulations, the implication often is that they might occur in nature as well and should be reckoned with.

This paper is organized as follows: In section 2 (methods) we define the categories of SNS and how they are found. Details of the classification criteria have been moved to the Appendix, also containing tables and supporting figures. In section 3 the SNS are discussed and illustrated per category and model simulation in which they appear. Section 4 contains a discussion focusing on robustness of our results, followed by conclusions (section 5).

2 Methods

In our method we focus on decadal-timescale changes of yearly mean values and discern two types of SNS: abrupt changes and transitions to a new equilibrium. The precise criteria may depend on the variable involved and the length of the time series investigated. Additionally, we have chosen to do this in an automated and reproducible manner to improve replicability. To achieve this, we follow a phased approach as in Drijfhout et al. (2015), but the three phases identified in that paper are now combined in one workflow without relying on visual inspections as in Drijfhout et al. (2015). The phases are schematically depicted in Figure 1.

We applied this method to data of 54 models using the SSP119, SSP126, SSP245, SSP370, and SSP585 scenarios and all available ensemble-members. We investigated 12 CMIP6 variables in this study (see Table A1): hfds (Downward Heat Flux at the Sea Water Surface), mlotst (Ocean Mixed-Layer Thickness Defined by Sigma-T), msftbarot (Ocean Barotropic Mass Streamfunction), msftyz (Ocean Y Overturning Overturning Mass Streamfunction), msftmz (Ocean Meridional Overturning Mass Streamfunction), siconc (Sea-Ice Area Percentage on ocean grid), siconca (only if siconc is not available, Sea-Ice Area Percentage on atmospheric grid), sos (Sea Surface Salinity), tas (Near-Surface Air Temperature), tos (Sea Surface Temperature), wfo (Water Flux into Sea Water) and zos (Sea Surface Height Above Geoid). We focused on these ocean, sea-ice, and atmospheric variables involved in the coupling to ocean and sea ice as they were the largest contributor to the SNS in Drijfhout et al. (2015), catalogues for SNS in other variables will follow in a separate paper. The variables msftyz, msftmz are used to identify SNS in the AMOC. These variables have a different format from other (surface) variables and are therefore handled differently (as discussed in (Drijfhout et al., 2025)).

2.1 Pre-processing

As shown in Figure 1, we take the following steps to identify SNS. First, monthly CMIP6 data are averaged to obtain yearly data[1]. Thereafter, the data are re-gridded to a Gaussian N90 grid (grid with 180 latitudes and 360 longitudes) using `cdo` (Schulzweida, 2023). We also combine the historical run with the scenario runs to obtain a dataset that spans the years 1850 to 2100 or later. Combining these time-intervals allows for more effective use of several statistical properties (like the dip-test (Hartigan and Hartigan, 1985)) that will be used to select SNS (see subsection 2.3 and subsection A2. The historical run is matched to the scenario runs via the metadata available within CMIP6-standardized metadata, and the match is validated to assure that there is no statistically unlikely transition at exactly the year 2014 (the transition from historical to scenario run) to assure no artificial SNS are identified due to a mismatch between historical and scenario. Only data where there is a match between historical and scenario runs is taken into account in this work, of which an overview is listed in Table A2. We split the ocean and land fractions of the data and evaluate both separately as several features used below or in subsection A2 may greatly differ above ocean and land. In total, 13266 different datasets (unique combinations of model, SSP, variable, ensemble-member, land/ocean) are taken into account in this work. In this catalogue, this is only relevant for tas as all other variables are only ocean variables.

[1]Except for mlotst, where we took the maximal value from monthly data as the yearly maximum is more relevant for deep convection than the yearly average.

2.2 Region isolation

In the second step, we isolate regions using four different approaches. The aim of each approach is to isolate regions that have a higher likelihood of showing an SNS, which will be evaluated in the next phase. Two competing interests were taking into account in the design of this step:

- We want to select regions that are not too large as that might smooth out potential signals of SNS.
- We want regions to be large enough (at least $10^6\,\mathrm{km}^2$) to focus on large-scale SNS, rather than local effects.

We achieve both, by selecting regions that are slightly larger than $10^6\,\mathrm{km}^2$ by using a range of thresholds as explained below (also see subsection A1). The region-isolation only serves to reduce the computational complexity and is therefore designed to rather select too many than too few regions. For the second step, we use the scenario run only (years 2015 to 2100 or later), and always use the 10-year running mean of time-series.

1. The first region-finding approach selects regions that have a large difference between their start and end values. In this approach we select regions that surpass a percentile threshold T. To avoid selecting too large regions, we first try a very high threshold of $T \geq 99.99\,\%$. We select all resulting regions that surpass this threshold and span at least $10^6\,\mathrm{km}^2$. We employ the unsupervised clustering algorithm HDBSCAN (Campello et al., 2013; McInnes and Healy, 2017) to cluster grid cells that surpass the threshold, forming continuous regions. We then select regions (that were not selected in the previous iteration) that surpass $T \geq 99.74\,\%$ (and are $\geq 10^6\,\mathrm{km}^2$). We keep repeating this process for successively lower values of T until (after a total of 61 iterations) we select regions with the lowest threshold $T \geq 85\,\%$. Regions ($\geq 10^6\,\mathrm{km}^2$) where $T < 85\,\%$ are not selected. In other words, we repeat the process for a total of 61 values of T, which are linearly spaced between 99.99 % and 85 %. This approach is also exemplified in subsection A1.

2. The second approach computes the percentiles of the standard-deviation of the (linearly) detrended time series of 10-year running mean (σ_{det}) and the 10-year maximum jump of the 10-year running mean (MJ_{10}). Both percentiles must lie above the percentile threshold T, which is linearly spaced between 99.99 % and 85 % in 61 steps.

3. The third approach is very similar to the second, and multiplies the percentile scores of σ_{det} and MJ_{10} and then selects regions where this product is above threshold T_p, which is linearly spaced between 99.9 % and 85 %. In comparison to the second approach, it selects regions where both the σ_{det} and MJ_{10} have high percentiles simultaneously. The fact that the isolated regions of the second and third approaches often similar does not hinder the goal of this step (isolating regions of interest to be evaluated for in the next phase), and only result in an increase in the number of regions to be evaluated which is solely a small increase in computational effort.

4. Finally, the last approach compares the values of σ_{det} and MJ_{10} with the values in the preindustrial-control, henceforth pi-control dataset. The ratio of the scenario values divided by the respective pi-control values has to surpass a threshold of T_{PiC}, linearly spaced between $8\times$ and $2.5\times$ in 61 steps.

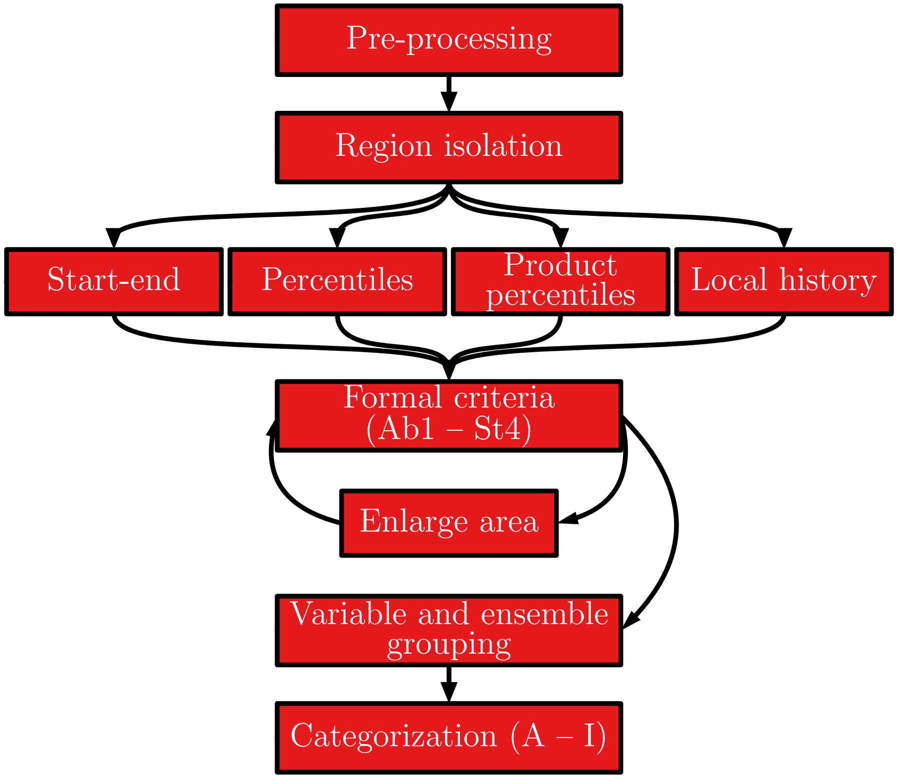

Figure 1. The workflow that leads to SNS identification. At the first phase, monthly CMIP6 data is pre-processed by calculating annual means and regridding any of the original coordinate systems to a gaussian N90 grid. At the next step, promising regions are isolated based on four algorithms (see text), which are subsequently evaluated using formal criteria to identify the regions with an SNS. Regions where an SNS occurs are enlarged by merging with adjacent or overlapping regions or groups of $5^\circ \times 5^\circ$ square regions. Finally, we first group the SNS by combining ensemble-members and variables of one model that occur in overlapping regions and then manually categorize these SNS into 11 categories.

The region isolation phase isolates $\mathcal{O}(10-100)$ regions of at least $10^6\,\mathrm{km}^2$ are often adjacent and may overlap when found by different methods.

2.3 Formal SNS criteria and area enlargement

In the third phase, we evaluate which regions from the previous phase fulfil formal criteria, further explained below. We take all the regions from the previous phase, without distinction between the different approaches. We iteratively evaluate all regions if they fulfil any of the formal criteria. If this is the case for a given region, any adjacent or overlapping regions are merged with it, as long as the combined region also fulfils the formal criteria. Additionally, to flexibly increase the size of an identified region, we also attempt merging adjacent $5^\circ \times 5^\circ$ regions (again, as long as the combined region also fulfils the formal criteria).

These strict requirements formalize the role of expert judgment of earlier work (Drijfhout et al., 2015). We used six different formal sets of criteria to detect various SNS events:

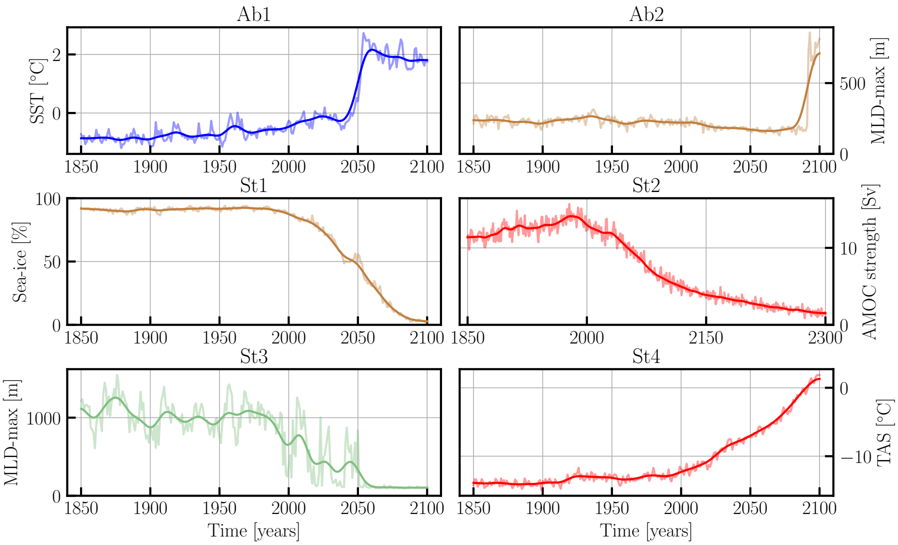

Figure 2. Examples of the six formal classification criteria (Ab1-St4), see text for further details. Panels Ab1, Ab2, St1, St2, St3 and St4 refer to SNS 61, 60, 6, 50, 66, and 18 respectively in Table A3. The solid lines indicate the smoothed yearly time series (which are illustrated as shaded lines). The data is smoothed using a Lowess filter with a window of 40 years.

Ab1: Abrupt change.

Ab2: Abrupt change at the end of a dataset.

St1: Almost complete disappearance of sea-ice in an area of at least $5 \times 10^6\,\mathrm{km}^2$.

St2: Transition to an extremely weak state in the Atlantic meridional overturning circulation (AMOC)[2].

St3: Transition from a deep to an extremely shallow ocean mixed-layer, or vice versa.

St4: Transition to a completely new state in other ocean/atmosphere variables.

An example of a time series associated with each criterion is shown in Figure 2, and the definition of each criterion is discussed in the appendix (subsection A2). The six different criteria-sets can be motivated by the moment the SNS occurs relative to the length of the timeseries and/or the nature of the climate system showing a state (phase) transition, the timescale of the inherent feedbacks, and the noise level of the system. We define two criteria related to abrupt changes; Ab1 and Ab2, where Ab2 is specifically developed for abrupt shifts at the end of a timeseries which implies that these SNS can never fulfil the demand of a bimodal distribution, otherwise applied. While having to relax this demand we somewhat sharpened other ones compared to Ab1 in the Ab2-set to compensate. We defined four different criteria-sets associated with state transitions. St4 is our standard set, applying to most ocean variables for which ocean dynamics and feedbacks dominate. St1 is developed for sea-ice and more specifically the disappearance of year-round (or winter) sea-ice in the Arctic. As this is a circumpolar phenomenon,

[2] As discussed above, this criterion follows the methodology described in Drijfhout et al. (2025) to isolate SNS rather than the workflow as in Figure 1.

we made the minimum area demand more stringent and demanded a 95% decline of sea-ice in this enlarged area to exclude trivial sea-ice decline in a warmer world which occurs everywhere where sea-ice developed. For the AMOC we used a slightly different set (St2) as we deal here with a phenomenon that is integrated over depth and its zonal dimension, making it to stand out from latitude-longitude patterns. For Mixed-layer depth we also adapted the criteria-set of St3 as due to the much higher noise (variability) in its timeseries on the decadal timescales, the SNS we find with the purpose build St3 would not have been isolated with St4 due to this higher variability. This adaptation was motivated by the large impact of atmospheric noise on mixed-layer depth and relevant for phenomena dominated by mixed-layer dynamics. This criteria-set will also be used for Ocean Biogeochemistry variables in a follow-up catalogue. Criteria-sets Ab1, Ab2 and St2 each detect several types of SNS (described below), while St1, St2, and St3 are each tailored to one specific type of SNS in this catalogue.

The criteria-sets Ab1,2 and St1-4 have been formulated in such a general way that they cover all types of SNS that apply to mean states and are dominated by slow (order of ten years or more) processes relevant for the coupled ocean-sea-ice-atmosphere system. Other types of SNS events not considered here are changes in variability, extremes, and fast atmospheric dynamics, likely excluding monsoon-related SNS events, although it should be noted that CMIP-style climate models are probably unable to simulate such SNS events or even reliably simulate recent precipitation trends in monsoon systems (e.g. (Xin et al., 2020)). In addition, tipping of monsoon-systems had not been found in CMIP models so far and it was recently debated whether global warming would be able to induce tipping of the Indian (Armstrong McKay et al., 2022). In the same paper it was hypothesized, however, that the West African Monsoon might show tipping associated with AMOC shutdown. SNS events in hydrology, vegetation, carbon stores and ocean biogeochemistry will be discussed in a follow-up paper.

The equations related to each of the six sets of criteria are available in the appendix (subsection A2). We evaluate the formal criteria for each variable and each ensemble-member available in our dataset. In Table A2 we list all the investigated datasets. After combining ensemble-members, variables, and scenarios that overlap, we end up with 73 SNS, as will be discussed in section 3. We group the 73 SNS into 11 categories (as in Figure 3) by combining cases of different models that are in the same region and occur because of the same underlying physical processes. Another type of SNS is a possible cascade of tipping points. The possible occurrence of such interactions between tipping elements was recently discussed in Wunderling et al. (2024). If occurring within the timeline of CMIP6 simulations such cascades will implicitly be detected as our criteria-sets would detect both the primal tipping element and successive other tipping elements without explicitly making the link between the two. In our analysis we aim to identify cascades in SNS when they occur and discuss and table them as such later in this work.

3 Results

In theory abrupt changes and state (phase) transitions could coincide, making the distinction in the formal criteria (Ab1-St4, subsection 2.3) look arbitrary. In practice, however, all abrupt changes involve a fast, nonlinear change in the state of the pertinent climate subsystem, that is too small to stand out as a state (phase) transition mainly caused by system feedbacks instead of a response to the forcing, while the state transitions that we do classify as such always evolve too gradual to classify

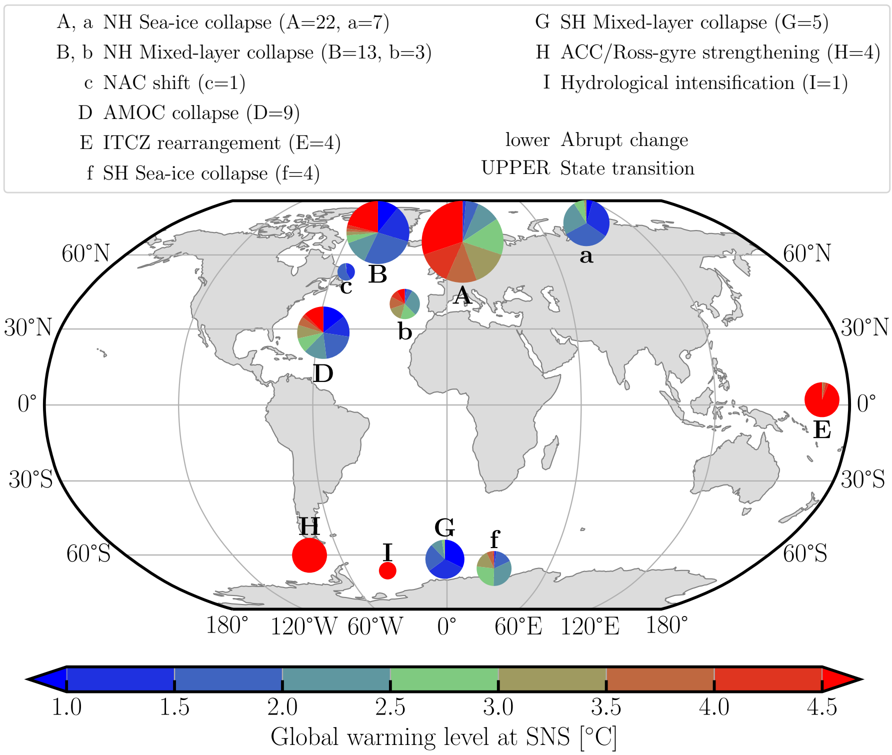

Figure 3. Overview of strong nonlinear surprise categories. The physical processes relate to Northern Hemisphere (NH) sea-ice collapse, Northern Hemisphere mixed-layer collapse, North Atlantic Current (NAC) shift, Atlantic Meridional Overturning Circulation (AMOC) collapse, Intertropical Convergence Zone (ITCZ) rearrangement, Southern Hemisphere (SH) sea-ice collapse, Southern Hemisphere mixed-layer collapse, Antarctic Circumpolar Current (ACC)/Ross Gyre strengthening, and Hydrological intensification. Capital letters denote large state transitions, while lowercase letters indicate abrupt changes. If a physical process causes both state transitions and abrupt changes in the same region, the same letter is used (e.g. A and a). The pie-charts are scaled in size to the number of models in which the pertinent SNS are found (corresponding to Table A3). The associated colours show the average global warming level with respect to 1850-1880 using the method described in subsection 3.10. The locations of the pie-charts are indicative of the median location of the SNS within each category, which might be slightly shifted within the standard-error of latitude and longitude for illustration purposes.

as abrupt transitions. While this distinction often is a result of systems response to the different forcing applied in the various SSP scenario's and could be explained by strong local feedbacks in case of abrupt shifts and moderate forcing that do not upgrade to strong basin-scale or global feedbacks involved with a state transition in case of strong forcing, we will discuss them together but keep the distinction between the two. Figure 3 illustrates the locations of the 11 categories (A-I). Some categories have related abrupt change and state transition SNS, for which we use the same uppercase letter for their state transitions and lowercase letters for abrupt changes to highlight their relation. The position of each of the pie-charts in Figure 3 is indicative of the mean global position of the SNS in that category. Each pie-chart indicates the global warming level at the time of the maximum change, we use the methodology described in subsection 3.10 to build CDFs (cumulative distribution functions) of the SNS-categories according at global warming levels (resultant in Figure 18). The pie-charts in Figure 3 corresponds to binning the probability of Figure 18 between $1 - 4\,°C$ global temperature rise, with overflow bins for SNS at global warming levels below 1 °C and above 4 °C. Figure 3 emphasizes the relative number of models that features a certain category of SNS, as the size of each pie-chart is scaled to the number of SNS.

In the remaining sections, we discuss each of the 11 categories alphabetically (corresponding from North to South). For each category, we provide at least one exemplary case to demonstrate the related physical processes. However, categories may contain a multitude of models, SSPs, ensemble members, and variables. Demonstrating all possible permutations is beyond the scope of a single paper, we therefore provide the following additional resources and informational datasets:

1. The example of each category (subsection 3.1 - subsection 3.9).
2. Table A3 list all cases, with categories, models, SSPs, variables and regions.
3. In Figure A5, we show the first ensemble member for each SNS in each category, with each associated variable displayed independently.
4. Much information may still be outside the scope of this finite list of information sources, we therefore make the data relating to all cases available under (Angevaare, 2025b). Here one can find a figure with isolated region and spatially averaged time series for each SNS, per model, per SSP-scenario, per variable, per ensemble member. Additionally, we also include the data for reproducing these figures and demonstrate how to extract several types of information from those data.

3.1 Abrupt year-round sea-ice loss in the Arctic (a) and transition to large, year-round ice-free areas in the Arctic (A)

3.1.1 Abrupt (Cat. a, 7 models)

As in Drijfhout et al. (2015), we find a few cases of abrupt sea-ice decrease in the Arctic. These cases are no longer restricted to the SSP585 scenario, as in CMIP5 for the RCP8.5 scenario, but also occur for weaker forced scenarios. Still, the forcing mechanism is in each case the global warming and polar amplification of this signal, mainly caused by the sea-ice albedo feedback (Davy and Griewank, 2023) In these cases the trajectories in the SSP126 and in the SSP245 scenario are qualitatively similar to the transition to a large, year-round ice-free area in the Arctic in the more strongly forced SSP370 and SSP585 scenarios. An example of this behaviour can be found in the CanESM5. Before 2050 the change proceeds abruptly but after 2050 it stabilizes unlike the evolution in the more strongly forced scenarios, which become part of Category A (transition to large, year-round ice-free areas in the Arctic) also see the difference between Figure 4 and Figure 5).

A qualitatively similar story holds for other models but depending on the model's sea-ice sensitivity to global warming, abrupt changes may also occur for the SSP370 scenario. These cases are: CanESM5-1 (SSP119 and SSP245); EC_EARTH3-Veg-LR (SSP126); MIROC6 (SSP126, SSP245 and SSP370); and GISS-E2-1-H (SSP370). In one scenario the abrupt sea-ice loss induces equally abrupt changes in ocean and atmospheric variables, namely TAS (near-surface air temperature) in CanESM5-1 (SSP119).

The relevant feedback mechanisms were already discussed in Drijfhout et al. (2015). This type of abrupt change was compared to the feedback-driven abrupt shifts in sea-ice edge found in an energy balance model (Rose and Marshall, 2009). This

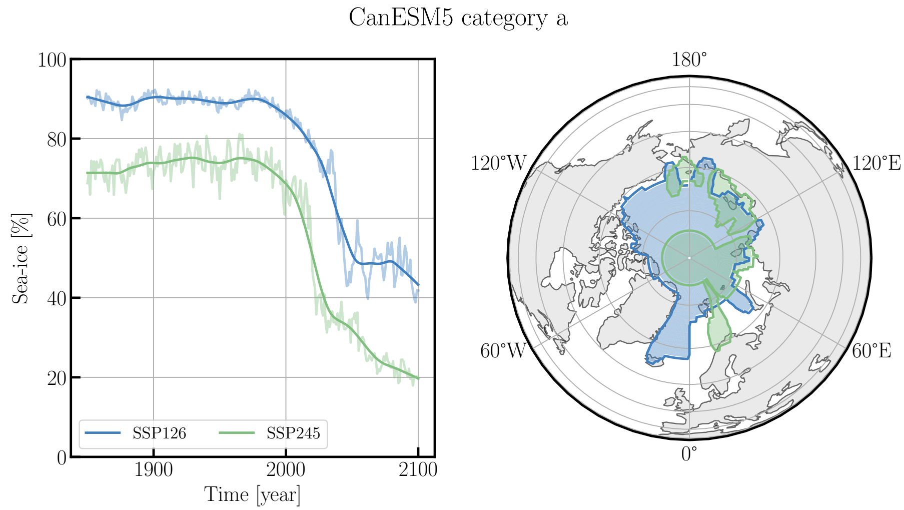

Figure 4. Category "a" for the CanESM5 model (SSP126: r25i1p2f1, SSP245: r1i1p1f1). The left panel shows the yearly averaged time series (shaded) and the smoothed time series (solid), where we use the same smoothing as in Figure 2. The right panel shows the corresponding region. This figure is one example of the category "a" abrupt changes, please see Table A3 for all models, as well as (Angevaare, 2025b) for a complete overview with time series and locations for all SNS from all categories.

category's similarity to the more state transition to ice-free conditions in more strongly forced scenarios (subsubsection 3.1.2) points to anthropogenic warming as the driver, but internal variability may affect the timing of the abrupt event. Hankel and Tziperman (2023) demonstrated the existence of a winter (and thus year-round) sea-ice tipping point in a state-of-the-art climate model. In a previous study (Hankel and Tziperman, 2021) they identified a positive feedback cycle associated with this tipping point in which, at the start of winter, more open ocean water allows for moistening and warming of the lower atmosphere, which in turn increases downward longwave radiation, which reduces ocean freezing. The diminished winter sea-ice growth results in lower surface albedo, leading to increased shortwave absorption during springtime. In addition, increased ocean heat uptake causes even warmer ocean conditions that lead to further sea-ice reduction. Abrupt loss of Arctic sea-ice could also be driven by other positive feedback mechanisms (Abbot et al., 2009; Kay et al., 2012; Feldl et al., 2020)).

Large changes in sea ice that themselves do not classify as an SNS can trigger an SNS in ocean or atmospheric variables. One such an example is the very large surface air temperature increase over the Pacific sector of the Arctic in UKESM-0-LL in its SSP126 scenario, which appears driven by sea-ice loss over that area allowing for atmospheric warming due to enhanced ocean heat loss (not shown). The sea-ice area involved is too small to classify as a category A transition and the change in sea-ice is not abrupt enough in sea-ice cover either, yet still causes an abrupt change in air temperature. This is an example of a cascade, not between tipping points themselves but between different variables associated with the same tipping event. In this case disappearance of sea-ice drives large and abrupt changes in the ocean-atmosphere interface. In this case sea-ice loss cascades into abrupt surface to temperature increases.

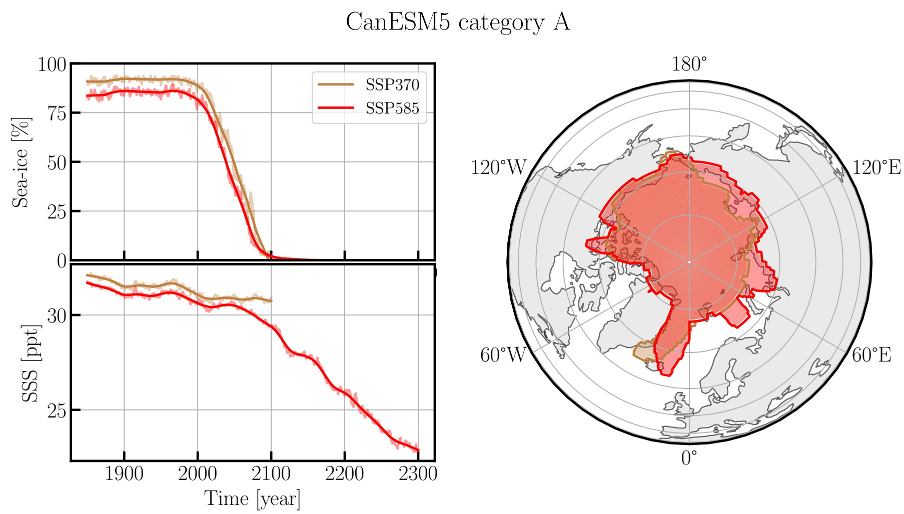

Figure 5. Category "A" for the CanESM5 model (r1i1p1f1), where the SSP370 model run was evaluated until the year 2100 and SSP585 was extended to 2300. The sea surface salinity (SSS) is expressed in parts per thousand (ppt). The region (right) corresponds to the region wherein the sea-ice SNS is found, and, for consistency, the left panels correspond to these regions. For SSP585, an SNS was also found in a (smaller) sub-region for SSS. See caption of Figure 4 for details.

Another such example is the abrupt increase in mixed-layer depth after the disappearance of sea-ice leading to enhanced heat loss from the ocean. This occurs in the Arctic where the yearly averaged mixed-layer depth from 50 to 350 m (CNRM-ESM2-1 in SSP245), north of 80 ° N and ranging from 10-120 ° E, roughly north of the Barents and Laptev Seas.

3.1.2 State Transitions (Cat. A, 7 models)

In Drijfhout et al. (2015) five models were found in which the winter sea-ice cover (siconc) largely collapsed. In CMIP6 many more of such cases appear, namely in 21 models and 29 model runs (ACCESS-CM2, ACESS-ESM1, CanESM5, CanESM5-1, CanESM5-CanOE, CESM2-WACCM, CMCC-CM2, CMCC-ESM2, E3SM-1-0, E3SM-1-1, E3SM-1-1-ECA, EC-Earth3-Veg, FIO-ESM-2-0, GISS-E2-1-H, HadGem3-GC31-LL, HadGem3-GC31-MM, IPSL-CM6A-LR, MRI-ESM2-0, NorESM2-MM, UKESM1-0-LL, UKESM1-1-LL), although it should be noted that we focus on annual mean siconc and allow for arbitrary areas of more than five million square kilometres, rather than restricting the analysis to the entire Arctic north of 75 ° N as in Drijfhout et al. (2015). An illustration of this type of SNS in CMIP6 is given by Figure 5. In that paper the Arctic winter siconc collapse only occurred in RCP8.5 extended runs. A model-dependent temperature threshold varying between 4.5 °C and 8.2 °C was needed for the onset of the winter siconc collapse. Such thresholds are much earlier reached in CMIP6 models ($2.5-4.1$ °C 68% confidence interval, see Table 1 and subsection 3.10). Also see Angevaare et al. (2025) for further discussion.

Likewise as for some of the abrupt cases large-scale sea-ice loss cascades into state transitions into variables affected by enhanced ocean-atmosphere exchange. In these cases, the disappearance of sea-ice allows for a much stronger ocean-atmosphere coupling with enhanced heat and freshwater exchange that was otherwise shielded by the sea-ice lid and that now ultimately leads to a new ocean and/or atmospheric climate in the Arctic. Most of these cases are state transitions sea surface salinity and atmospheric surface temperature, and sometimes circulation changes exemplified by state transition in sea surface height. They appear in SSP585 scenarios in ACCESS-CM2 (sea surface salinity; SSS), ACCESS-ESM1-5 (near surface air temperature; TAS), CanESM5 (SSS) see also Figure 5, CESM2-WACCM (SSS and Sea Surface Height; SSH), FIO-ESM-2-0 (TAS), GISS-E2-1-H (SSS), MIROC6 (TAS) MRI-ESM2-0 (SSS), UKESM1-0-LL (TAS), and UKESM1-1-LL (TAS).

3.2 Abrupt Northern Hemisphere winter mixed-layer weakening (b), and State transitions to extremely shallow Northern mixed-layers (B)

3.2.1 Abrupt (Cat. b, 3 models)

In Drijfhout et al. (2015) several cases of abrupt change in Northern Hemisphere deep convection were identified. These were all cases where deep convection and mixed-layer depths collapsed. In CMIP6 we do not find any such case. This is in apparent contradiction with Swingedouw et al. (2021) who identified three CMIP6 models in which abrupt cooling events and abrupt transitions to very low mixed-layer depths in the North Atlantic subpolar gyre (SPG) occurred. Here, we do not classify those events as abrupt changes, as such changes are ubiquitous in the time series, also occurring in the historical and pi-control runs preceding the abrupt events that in Swingedouw et al. (2021) were identified (see their Figures 1 and 6). This is partly due to the more stringent classification criteria we apply, among others by also comparing with the variability in the historical and pi-control runs, and due to the very large internal variability in mixed-layer depth, making it difficult for mixed-layer changes to classify as abrupt. Also note that Swingedouw et al. (2021) classified their abrupt events by demanding the associated temperature change to exceed a certain threshold. This discrepancy should not be interpreted that either we or Swingedouw et al. (2021) are right or wrong, but simply that what one finds as an abrupt shift depends on the criteria applied, and unfortunately, the criteria for an abrupt shift or SNS will always remain somewhat subjective, as for instance for what is coined extreme weather.

While we do not find abrupt changes in mixed-layer depth, we do find in three models abrupt changes in the North Atlantic SPG in ocean variables which concur with a significant, but non-abrupt mixed-layer depth decrease. These occur in scenarios SSP245 and SSP370 for cases where the decrease in mixed-layer depth in the SSP585 scenario over the same region is large enough to classify as a State Transition (category B). For this reason we include them in this section. In GISS-E2-1-G the abrupt change occurs for sea surface salinity (SSS) in SSP245 and SSP370, and for sea-surface temperature (SST) in SSP370 only; Figure 6. Since both SSS and SST decrease, the driver for the mixed-layer decrease must come from SSS decreasing the surface density as freshening, only partly counteracted by the cooling that increases the surface density. This is again an example of a cascade of SNS variables in the same tipping point case, but now the large decrease in SSS is triggering a mixed-layer collapse, see also Drijfhout et al. (2025). Similar mixed-layer decrease-induced abrupt changes in ocean variables occur for GISS-E2-1-H and for CESM2-WACCM (SSS in SSP585).

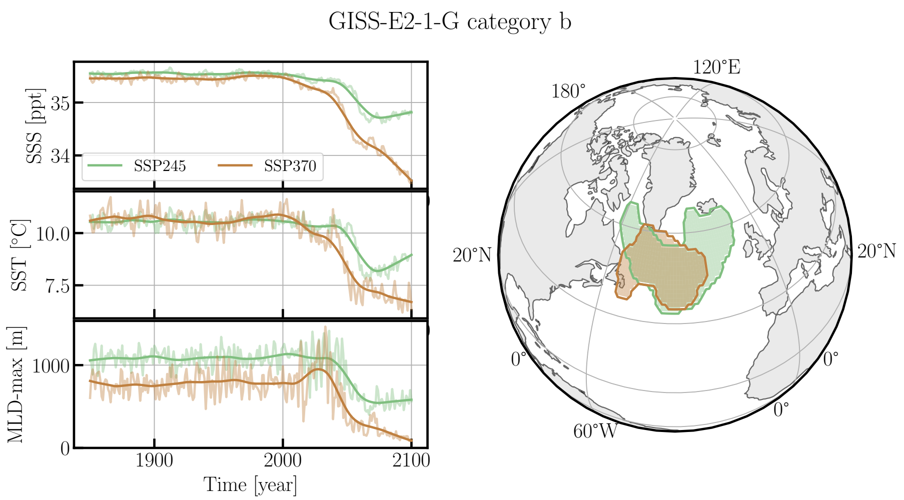

Figure 6. Category "b" for the GISS-E2-1-G model (SSP245: r2i1p5f1, SSP370: r4i1p5f1). The region SSP245 belongs to the SNS detected in SSS, while the SNS for the SSP370 region is detected in SST. See caption of Figure 4 for details.

3.2.2 State Transitions (Cat. B, 13 models)

While we do not find abrupt changes in mixed-layer depth, various models do show significant transitions from deep convection in the SPG to a state of no deep convection. This becomes more evident when data after year 2100 is included in the analysis. In our analysis, 10 models are identified featuring a collapse in mixed-layer depth in the Atlantic SPG and/or Nordic Seas. These are ACCESS-CM2, ACCESS-ESM1-5, CESM2, CESM2-WACCM, GISS-E2-1-G, GISS-E2-2-G, HadGEM3-GC31-MM, MRI-ESM2-0, NorESM2-MM, and UKESM-0-LL.

For all of those models the collapse occurs in the SSP585 scenario, but for CESM2 and MRI-ESM2-0, it also occurs in the SSP126, SSP245, and SSP370 scenarios; in CESM2-WACCM also in all scenarios but the SSP370 scenario was not run there. Curiously, in the GISS-E2-2-G model, the collapse occurs only in the SSP126 and SSP245 scenarios. These dramatic shifts in mixed-layer depth are, due to our requirement that they occur over an area of at least $10^6\,\mathrm{km}^2$, more often occurring over larger areas than just the Labrador Sea; often over the whole SPG and parts of the Greenland, Iceland and/or Norwegian Sea as well. In the UKESM-0-LL the collapse occurs in the Nordic Seas and develops between years 2020 and 2060; in the MRI-ESM2-0 and GISS-E2-1-G the collapse also includes the mixed-layer depth in Labrador and Irminger Seas. In the other models the collapse occurs in the Labrador and Irminger Seas only (e.g. Figure 7). In all models the North Atlantic mixed-layer collapse occurs during the early to mid 21$^{\mathrm{st}}$ century and often already starts around 2020. In Drijfhout et al. (2025), the collapse is analysed for the much smaller deep convection area and for the maximum mixed-layer (March) only, showing a far more abrupt signal with steepest decline around 2050. There, it is argued that the North Atlantic mixed-layer collapse precedes the AMOC collapse, and that this appears to be the main cause for the AMOC winding down to extremely weak states. As a result, this is a clear example of a tipping point cascade.

NorESM2-MM category B

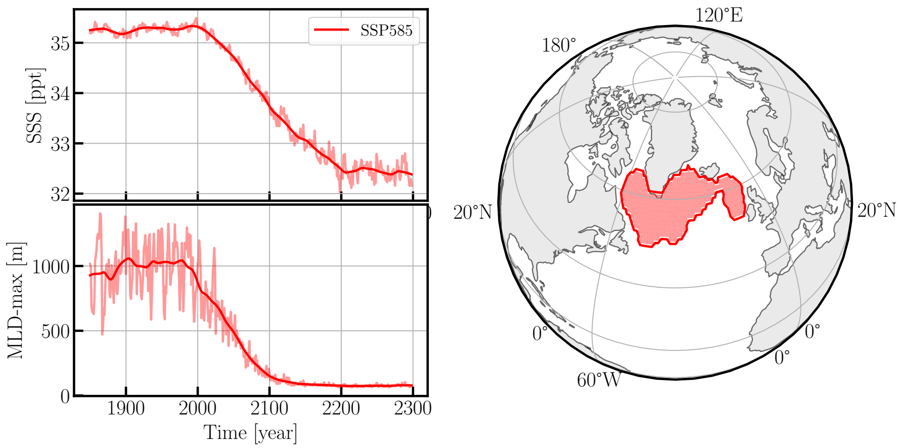

Figure 7. Category "B" for the NorESM2-MM model (r1i1p1f1). The region (left) belongs to the region of the SNS detected in MLD-max. In this model run, the SSS also showed an SNS, that extended further in the Baffin Bay and less far eastwards. For consistency, the left panels correspond to the region in the right panel. See caption of Figure 4 for details.

It should be noted that these collapses can occur in all SSP scenarios with an onset that sometimes already starts at a 1 °C warming level in the beginning of the 21st century, suggesting a strongly increased sensitivity in CMIP6 of mixed-layer depth and deep convection to global warming levels compared to CMIP5. Given the spread in warming level for the SPG- and Nordic Seas-convection collapse to occur, it is unlikely that global warming level is the right metric to measure its possible onset. As was already discussed in Drijfhout et al. (2015); Sgubin et al. (2017); Swingedouw et al. (2021), the onset of the collapse primarily occurs when the density stratification crosses a certain threshold, and this is more often caused by upper-layer freshening than by upper-layer warming, see also Drijfhout et al. (2025).

It should be noted however, that this freshening is brought about by a declining AMOC transporting less salty water northward, since increased meltwater release from a decreasing Greenland Ice Sheet was not included in CMIP6 simulations. The trigger for AMOC decrease and freshening must come from the forcing and in these models is supplied by warming over the SPG reducing ocean heat loss and even reversing the sign of the surface heat flux in the SPG. The impact of decreasing surface heat loss to buoyancy gain, is however always smaller than the impact of freshening in terms of buoyancy at the surface, see Drijfhout et al. (2025) and van Westen et al. (2024) for more discussion on this topic.

In all models the mixed-layer collapse is associated with substantial freshening, but in some models this change in SSS even qualifies as a large state transition to much fresher surface waters. In ACCESS-CM2, CanESM5, NorESM-MM (Figure 7) and UKESM1-0-LL this occurs in the Baffin Bay and Labrador Sea with sometimes an extension into the Hudson Bay and along the Labrador Current. Note that (Figure 7) nicely illustrates that although the mixed-layer collapse can ultimately be caused by salinity decrease it needs another trigger (i.e. changes in surface heat flux) as the start of the salinity decline lags the onset of the mixed-layer collapse. In CanESM5, the mixed-layer collapse itself does not qualify as an SNS, but the large-scale freshening in this region does. In GISS-E2-1-H and MRI-ESM2-0 the SSS transition occurs over the whole SPG and large parts of the

Arctic. The freshening of the Arctic, however, must be associated with the melt and year-round disappearance of nearly all sea-ice there. In CESM2, the large-scale freshening occurs preferably in the eastern SPG.

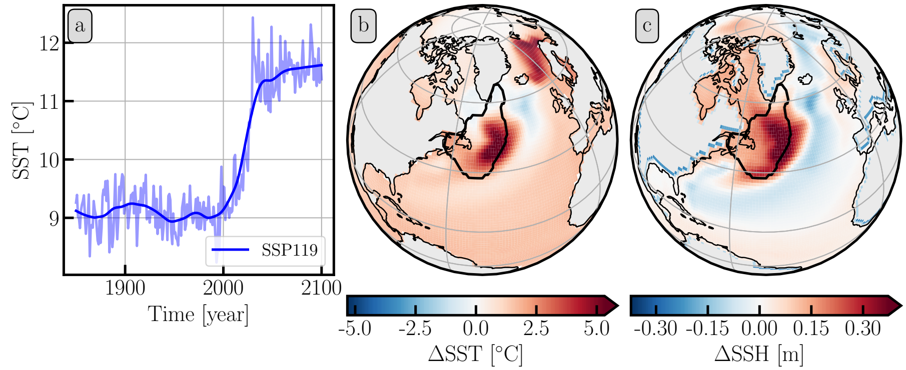

Figure 8. The time series of SST for category "c" (panel a), together with the change in SST (panel b) and SSH (panel c) are shown for GISS-E2-1-H (r1i1p3f1). The change for SST and SSH is calculated between the averages for the 1850 – 1880 and 2070 – 2100 periods. The black outlined region in panels b and c illustrates the location of the SNS region. Since the global average SSH over time in GISS-E2-1-H does not equal zero (that would be the dynamical sea-level rise), we manually subtract the global average SSH from each year's data. The yearly data in panel a are illustrated by the shaded lines while the solid lines depict the smoothed time series, where the smoothing is the same as in Figure 2.

3.3 Abrupt shift in the North Atlantic Current (Cat. c, 1 model)

Many low-resolution coupled models feature a strong cold bias east of Newfoundland due to an overly zonal North Atlantic Current (NAC) (Moreno-Chamarro et al., 2021). Associated with this bias in the current pathway a cold anomaly appears in sea-surface and surface atmospheric temperatures, which is often called the northwest corner bias. This bias is also present in the GISS-E2-1-H model, but surprisingly, NAC and surface temperatures feature an abrupt shift to a more northward pathway for the NAC through the central/eastern SPG, and subsequently warmer temperatures east of Newfoundland under the SSP119 scenario. The temperature change in the ocean is almost 3 °C and in the surface atmosphere over 2 °C, the upward heat flux doubles and sea surface height increases by almost half a meter. The change in SST is most abrupt, occurring within a decade around 2030 (Figure 8). Why this change in NAC occurs most sharply in the SSP119 scenario in this model remains unclear, but it is remarkable that the bias can abruptly change in response to a modest increase in radiative forcing, much more than when that forcing becomes stronger in a higher-emission scenario. That the NAC can be steered by applying advanced forms of local flux correction was demonstrated by Drews et al. (2015), who argued that this also must be applied to the subsurface to adjust the flow field by modifying density and pressure gradients over a large vertical column. We hypothesize that in this model, perhaps by chance, the temperature and salinity adjustment in response to an SSP119 forcing conspire to change the pathway of the NAC. The GISS-E2-1-H model suggests the NAC can feature abrupt pathway changes that resemble a tipping point, but whether such a tipping point exists in nature remains unclear. Further investigation of this abrupt change, however, could be valuable in guiding bias reduction efforts in models with a northwest corner problem.

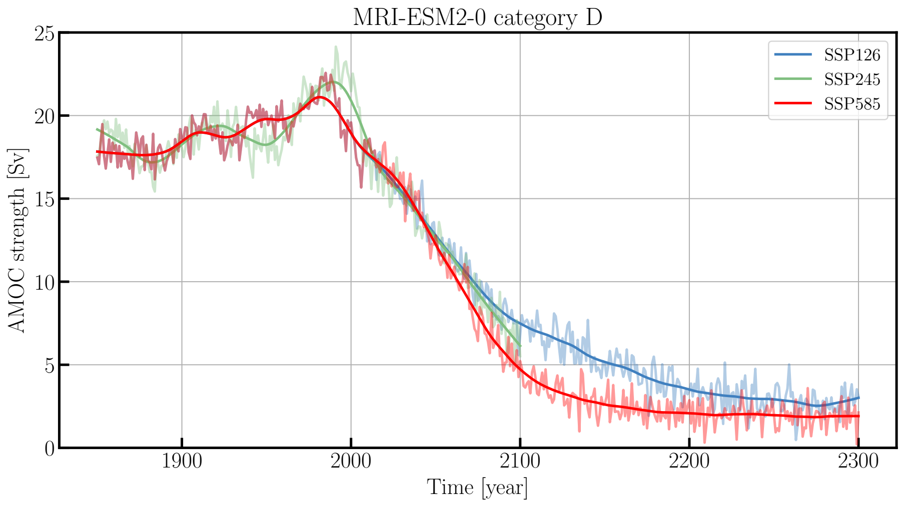

Figure 9. Category “D” for the MRI-ESM2-0 model (SSP126: r1i1p1f1, SSP245: r4i1p1f1, SSP585: r1i1p1f1). The yearly data is illustrated by the shaded lines while the solid lines depict the smoothed time series, where the smoothing is the same as in Figure 2.

3.4 Transition to extremely weak AMOC states (Cat. D, 9 models)

Given the widespread convection collapse in large parts of the northern North Atlantic in 10 models, it is unsurprising that an AMOC collapse also appears in seven of these and in two additional models. In the three runs in which the mixed-layer collapsed, but the AMOC did not, for one model (CESM2) the overturning streamfunction was not available, and for the other two models (GISS-E2-2-G and HadGEM3-GC31-MM) the respective runs stopped in year 2100, which is long enough to allow detection of a mixed-layer depth collapse, but too short to detect an AMOC collapse. Note however that in GISS-E2-2-G the AMOC declined in the 21st century from ~ 25 to ~ 12 Sv and in HadGEM3-GC31-MM from ~ 18 to ~ 7 Sv, implying that, were these runs extended to 2300, they probably would have shown a collapsed AMOC, almost certainly for HadGEM3-GC31-MM. Because an AMOC collapse tends to develop more slowly than the SPG convection collapse, it most is often identified in runs that were extended to year 2300. These include the previously discussed models showing an accompanying mixed-layer depth collapse being the models ACCESS-CM2, ACCESS-ESM1-5, CESM2-WACCM, GISS-E2-1-G, MRI-ESM2-0, NorESM2-MM, and UKESM-0-LL. In all seven models the AMOC collapse occurs in extended SSP585 runs, but in CESM2-WACCM and MRI-ESM2-0 the collapse also occurs in the extended SSP126 runs. In addition, we identify an AMOC collapse in the CanESM5 and MIROC-ES2L in their SSP585 runs. For these two models the mixed-layer depth collapse was not observed. For MIROC-ES2L, the data was not available and in CanESM5 a mixed-layer collapse can be seen but did not qualify as a category B SNS because all mixed-layers in this model show a large decline. In MIROC-ES2L, the AMOC collapse proceeds rapidly enough to be detected in a run that ends in year 2100. The onset of AMOC decline occurs in these models between years 1980 and 2020, corresponding to global warming levels between 1.0 and 1.5 °C, which is then completed between years 2150 and 2250 with an accelerated weakening until 2100 - 2150 and a slower approach to the new

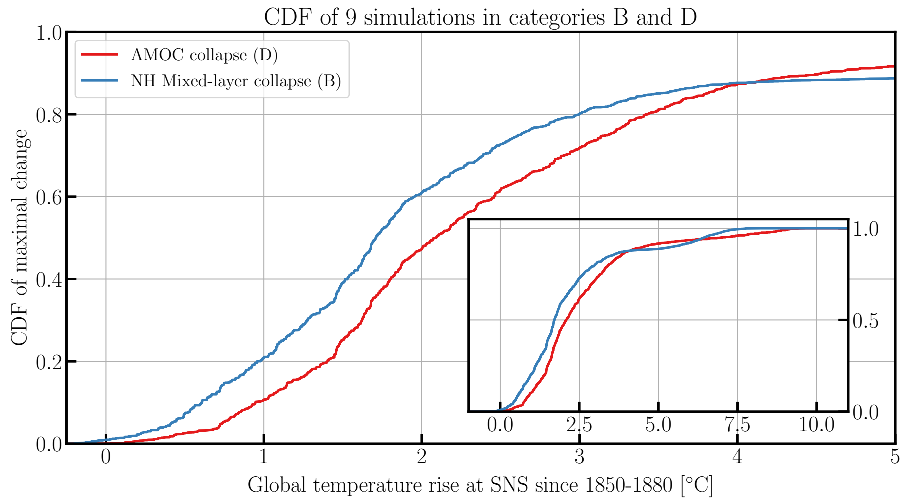

Figure 10. The change in category B SNS precedes the change in category D SNS. We show this using the cumulative distribution function (CDF) of the change being maximal before a global temperature rise for nine model runs that are in category B and in category D. We select the nine simulations (ACCESS-CM2, ACCESS-ESM1-5, CESM2-WACCM, GISS-E2-1-G, MRI-ESM2-0 (SSP126 and SSP585), NorESM2-MM (SSP126 and SSP585), and UKESM1-0-LL) that have the same ensemble-member in both categories B and D. The main panel shows that the collapse of the mixed-layer depth occurs 0.3–0.4 °C before the collapse of the AMOC. The inset shows the CDF over the full temperature range. The approach to obtain this figure is similar to that of Figure 17 (see text for details subsection A4).

(quasi)-equilibrium. While a wind- and buoyancy-driven shallow overturning cell of ~ 3 Sv always remains in these runs, the 5 Sverdrup (Sv) boundary is crossed between 2100 and 2140 in most models. Only in the CESM-WACCM model in the SSP126 scenario the AMOC stabilizes around 5 Sv.

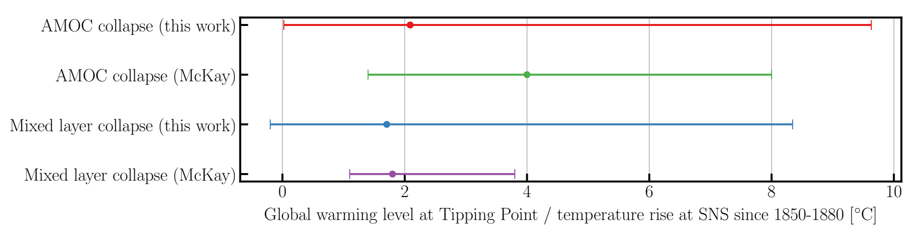

Figure 11. Global warming levels thresholds for SNS/tipping points comparison with respect to (Armstrong McKay et al., 2022), where we compare the the 0, 0.5, and 1.0 probabilities of Figure 10, to the Min, Estimate, and Max global warming level thresholds described in (Armstrong McKay et al., 2022). We compare our result AMOC collapse (D) to the Atlantic M.O. Circulation collapse, and our result NH Mixed-layer collapse to the Labrador-Irminger Seas / SPG Convection collapse.

In the MRI-ESM2 model the AMOC declines to ~ 3 Sv in both SSP126 SSP245 and SSP585 scenarios, only a bit slower in the SSP126 and SSP245 scenarios. In this model the AMOC decline is similar in all scenarios until year 2050 (Figure 9); in CESM2-WACCM even until year 2080. The AMOC collapse under an SSP126 scenario is a novel finding. It underscores the increased AMOC-sensitivity to warming in the multimodel mean Ensemble of CMIP6 compared to CMIP5 (see also Jesse et al. (2024) and (Mecking and Drijfhout, 2023)). It also suggests that in the MRI-ESM2 model the AMOC collapse can possibly be identified with a real tipping point as it largely becomes independent of the forcing and may already be committed to collapse decades earlier in the 21st century.

The much earlier onset of SPG convection and AMOC collapse than previously estimated (Fox-Kemper et al., 2021; Armstrong McKay et al., 2022; Loriani et al., 2023) is illustrated by Figure 11, urging for adjusting the burning embers in these reviews. This figure, together with Figure 10 also supports that we see here a tipping point cascade between a mixed-layer collapse triggering the AMOC collapse in the models.

3.5 Transition to new ITCZ state (Cat. E, 4 models)

Large changes associated with a reorganization of the ITCZ are found in large-scale SSS transitions to much fresher surface waters in the western tropical Pacific in ACCESS-CM2, ACCESS-ESM1-5, CESM2-WACCM and IPSL-CM6A-LR. An example of this transition is given in Figure 12. These changes occur exclusively in the SSP585 scenario and are associated with model bias. Many low-resolution coupled models feature a double ITCZ, especially over the Pacific Ocean (Moreno-Chamarro et al., 2021). This is also the case in the IPSL-CM6A-LR model and while the bias does not disappear in the SSP585 scenario, the signal over the equator, especially in the west Pacific, decreases abruptly. This is associated with a strong increase in rainfall over the region and reduced evaporation, caused by the double ITCZ becoming less pronounced. This change in ITCZ pattern causes a drop in salinity in the Pacific across the equator and most notably north of it, which occurs most abruptly along the western Pacific equatorial zone in one member[3] of IPSL-CM6A-LR (Figure 12). It is doubtful whether this change can be associated with a tipping point in the real world, as it originates from a strongly biased climate system, namely the double ITCZ, which is a common model bias, but not a feature of the real world (Tian and Dong, 2020). On the other hand, it is related to strong (positive) feedbacks between clouds, atmospheric convection and atmospheric/oceanic circulation that also are present in nature (Hwang and Frierson, 2013; Zhang et al., 2019). In principle, a large shift in the ITCZ could trigger (cascade) a Monsoon-related SNS (Geen et al., 2020). While we see a large increase in precipitation over the Himalayas, the area involved was too small to classify as an SNS in this model.

3.6 Abrupt year-round sea-ice loss in the Southern Ocean, associated with abrupt change in ocean and atmospheric variables (Cat. f, 4 models)

We find four cases of abrupt sea-ice loss in the Southern Ocean. These cases are, like in the Arctic, probably triggered by warming of the overlying atmosphere, although the polar amplification / warming signal in the Southern Hemisphere is weaker. The abruptness of the event, however, needs strong nonlinear feedbacks from within the ocean and the most likely candidate is the onset of (deep) convection after the disappearance of sea ice allows for large changes in surface heat flux. One such case was also found in Drijfhout et al. (2015) where the forcing mechanism resulting from the interaction of sea ice and deep convection was found similar to described in Lenderink and Haarsma (1996) The cases here can also be explained by the

[3]The algorithm also detected the change in the r6i1p1f1 member, which runs only up to 2100. For this ensemble-member, this region fulfilled criterion Ab2, which would strictly speaking classify it as an abrupt shift. However, as the evolution is so similar to that observed in r1i1p1f1, we decided against discussing r6i1p1f1 separately as an abrupt change. Note that r1i1p1 is the only member that was extended to year 2300.

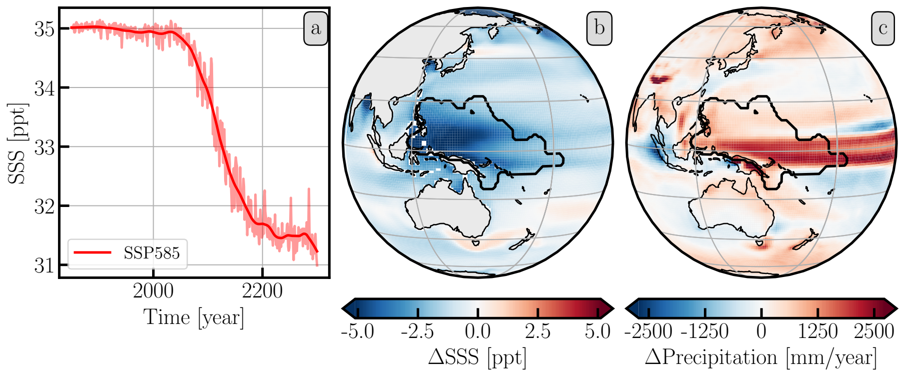

Figure 12. The time series (panel a) of category "E" in IPSL-CM6A-LR (r1i1p1f1) is shown together with the change in SSS (panel b) and precipitation (panel c). The change for SSS and precipitation is calculated between the averages for the 1850 – 1880 and 2270 – 2300 periods. See the caption of Figure 8 for further details.

same mechanism. This interaction of feedbacks can only be triggered if the stratification under the sea ice is weak enough and therefore depends on subtle differences in the ocean mean state at the end of the historical period. Abrupt sea-ice loss in the eastern Ross Sea and Amundsen Sea occurs in the SSP585 scenario in HadGem3-GC31-MM. Similar abrupt sea-ice loss, but more in the central Ross Sea, is seen in GISS-E2-1-H, accompanied with abrupt increases in upward heat flux and sea surface temperature for both the SSP245 and SSP370 scenarios, consistent with the onset of deep convection mechanism that we propose. In the same model also an abrupt sea-ice loss occurs east of the Weddell Sea (Lazarev Sea) in most scenarios (SSP126, SSP245, SSP370, SSP585) which shows an almost complete collapse of perennial sea-ice there. Subsequent abrupt changes in the Lazarev Sea also occur for ocean and atmospheric temperature and surface heat flux (Figure 13).

Depending on the variable, the abrupt change is seen in different scenarios; always in SSP126 but the change in SST is also abrupt in SSP245 and SSP370, and even in SSP119 where sea-ice itself disappears less abruptly. These changes all occur in the second half of the 21$^{\mathrm{st}}$ century. A similar scenario for the Lazarev Sea occurs in CESM-WACCM, but the change in that model is only abrupt in sea-ice in the SSP585 scenario. This example illustrates that a fast, but non-abrupt change in sea-ice, can, if coupled to fast changes in convective mixing, trigger abrupt changes in ocean or atmospheric temperature.

A similar scenario is simulated in the BCC-CSM2 model, but here we only see an abrupt change in mixed-layer depth without accompanying abrupt changes in other variables. The yearly averaged mixed-layer depth in the Ross Sea, in the middle of the Ross Gyre, suddenly increases from less than 100 to 1000 m between 2090 and 2095 in the SSP370 scenario. This abrupt change is associated with the disappearance of sea-ice, which itself does not occur abruptly. It should be noted that the feedback between sea-ice disappearance and the onset of deep convection associated with global warming has been demonstrated to occur for the Arctic (Våge et al., 2018), but not (yet) for the Southern Ocean. It should be noted, however, that this very same feedback was demonstrated to be instrumental in the multi-decadal variability in Southern Ocean sea-ice cover as sea-ice increase/decrease coincided with less or more deep mixed-layers Morioka et al. (2024). Instead of the increased warming resulting from SSP-scenarios these changes were triggered by the Southern Annular Mode, which, by changing

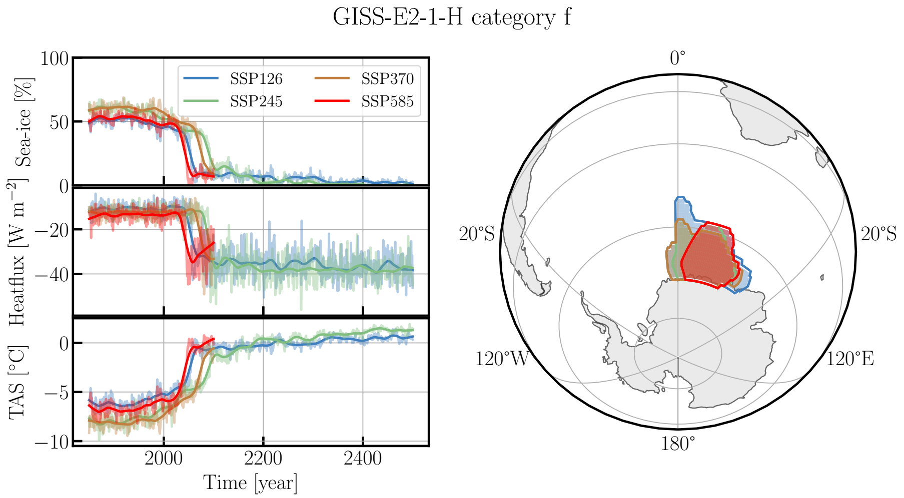

Figure 13. Category "f" for the GISS-E2-1-H model (SSP126: r4i1p3f1, SSP245: r2i1p1f2, SSP370: r3i1p1f2, SSP585: r4i1p3f1). For consistency, the averages (left panels) are calculated over the region (right panel) wherein the sea-ice SNS are detected. See caption of Figure 4 for details.

upwelling of subsurface saline water, modulates the ocean stratification and can abruptly enhance or onset deep convection and subsequent sea-ice cover changes.

3.7 Transition to extremely shallow Southern Hemisphere winter mixed-layers (Cat. G, 5 models)

Mixed-layer collapses also occur in the Southern Hemisphere, but less often than in the Northern Hemisphere. We identified five such cases for different models (ACCESS-CM2, ACCESS-ESM1-5, CNRM-CM6-1, NorESM2-LM, NorESM2-MM). CNRM-CM6-1 shows a collapse in an atypical area: the Davis Sea in the Indian Ocean sector of the Southern Ocean. This site is not particularly known for deep convection and this SNS is therefore unrealistic and a result of model bias, although the type of SNS is in itself realistic, when occurring at other sites. The other models show a mixed-layer collapse in, or just east of the Weddell Sea. In ACCESS-ESM1-5, this occurs in the Lazarev Sea east of the Weddell Sea. In the other three models the Weddell Sea is the site where the mixed-layer collapse occurs (Figure 7). As with the North Atlantic cases, these collapses all occur during the 21st century. Since open ocean convection in the Southern Ocean is a widespread feature of lower-resolution ocean and climate models, while in the observations, this type of deep convection is rare, and most of the Antarctic Bottom Water and its even denser precursors is formed on the Antarctic shelves, we do not assign much significance to these cases (Heuzé, 2020). The processes identified here, however, could be representative of the Weddell Polynya, where open ocean deep convection sometimes occurs, if it is opened (De Lavergne et al., 2014; Rheinlænder et al., 2021; Mchedlishvili et al., 2022). Anthropogenic-driven shifts in the Amundsen Low might play a role in causing abrupt changes on the Antarctic shelves, basal melt of the West Antarctic Ice Sheet and possibly the Antarctic Slope Current (Thompson et al., 2018), open ocean sea-ice and deep mixing (Holland et al., 2019; Drijfhout et al., 2024).

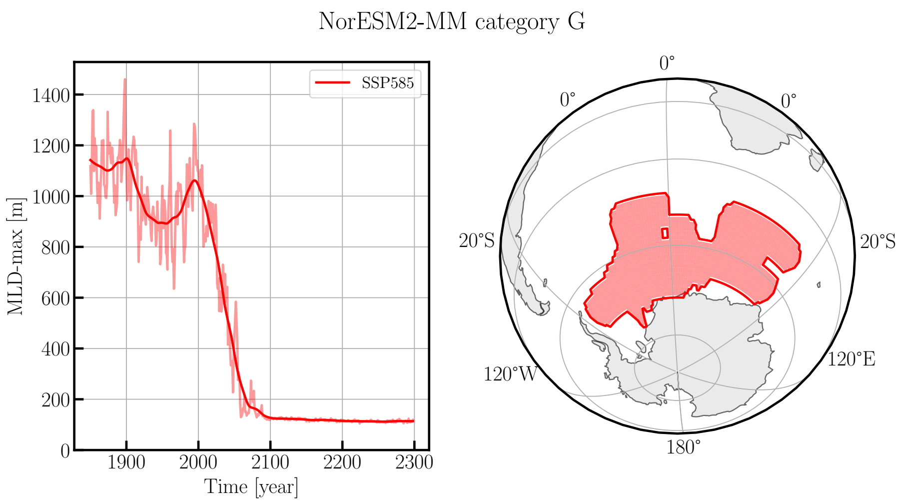

Figure 14. Category "G" for the NorESM2-MM model (r1i1p1f1). See caption of Figure 4 for details.

3.8 Transition to much increased Antarctic Circumpolar/Ross Gyre (Cat. H, 4 models)

Three models show a very strong drop in sea surface height (SSH) south of the Antarctic Circumpolar Current (ACC) associated with an increase in ACC mass transport of approximately 50 % in the SSP585 scenario (CESM2-WACCM, MRI-ESM2-0, UKESM1-0-LL). This is most likely related to the ACC response to the intensification of the Southern Hemisphere westerlies (Bracegirdle et al., 2013). This response, however, is probably spurious, in the sense that higher-resolution models show a much weaker ACC-sensitivity to winds, caused by what has been named "eddy saturation" (Munday et al., 2013), consisting of an intensification of the eddy field in response to increasing winds, rather than an increase in mean ACC transport. As a result, both SROCC (Meredith et al., 2019) and AR6 (Fox-Kemper et al., 2021) have assessed that the ACC is likely to remain only weakly sensitive to increasing winds. Moreover, such strong ACC response to winds may be enhanced by model bias explaining a large part of the spread in projections of CMIP5 models (Bracegirdle et al., 2013), which most likely also holds for CMIP6 models.

One model (IPSL-CM6A-LR) shows a very strong drop in SSH just east of the Ross Gyre in the SSP585 scenario, associated with an 80 % increase in its regional depression. This appears to be a sign of an eastward expansion of the Ross gyre, which also occurs in the UKESM1-0-LL (but less extreme than in IPSL-CM6A-LR), and was analysed by Gómez-Valdivia et al. (2023). They argued that sea-ice loss in the Amundsen-Bellingshausen sector enhances the wind stress curl on ocean water driving the eastward expansion of the Ross Gyre. This scenario was also described in a higher-resolution regional model of the region (Holland et al., 2019). Furthermore, the same process was already found to occur during the historical period (using an ensemble of high-resolution regional models forced with reanalysis data (Naughten et al., 2022)). As a result, this particular state transition is more credible than the projected increase in the ACC in some models.

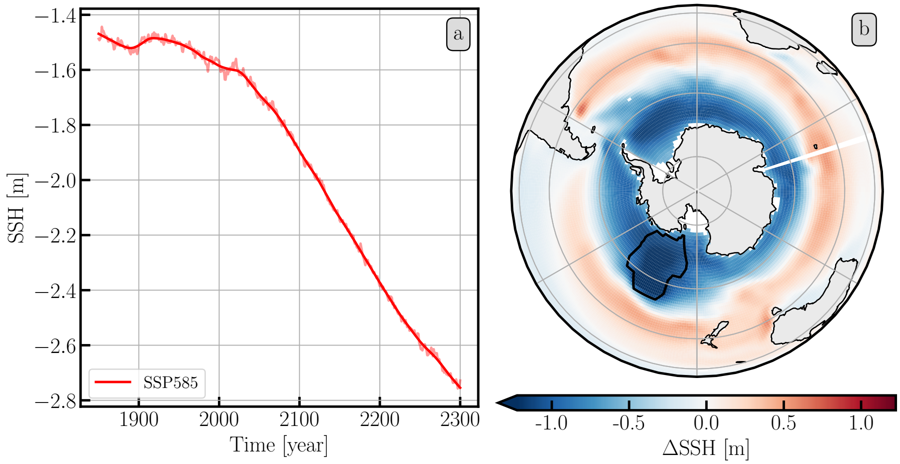

Figure 15. Time series (panel a) of category "H" in IPSL-CM6A-LR (r1i1p1f1) together with the change in SSH (panel b). The change of SSH is calculated between the averages for the 1850 – 1880 and 2270 – 2300 periods. See the caption of Figure 8 for further details.

IPSL-CM6A-LR also features a strong drop in SSH in other areas south of the ACC (Figure 15), but here the circumpolar drop in SSH is slightly too small to classify the entire region as a state transition. The other models (CESM2-WACCM, MRI-ESM2-0, UKESM1-0-LL) that do identify the entire circumpolar region south of the ACC as a state transition in SSH show a less pronounced drop in SSH in Ross Gyre than in IPSL-CM6A-LR. It is unclear whether the drop in SSH in the eastward extension of the Ross Gyre would classify as a state transition in a more realistic eddy-resolving model. In such a model, the ACC typically exhibits eddy saturation, and the circumpolar drop in SSH south of the ACC is either absent or remains modest (Stewart et al., 2023). Nevertheless, the response of the Ross Gyre appears robust and is corroborated in forced high-resolution ocean-sea-ice models.

3.9 State transition to much more humid climate over and in the Southern Ocean associated with strong intensification of the hydrological cycle (Cat. I, 1 model)

Another state transition is the large increase in surface freshwater flux in the Southern Ocean. This occurs in the CanESM5 model in the SSP585 scenario for the Amundsen-Bellingshausen sector (Figure 16) and the wider Weddell gyre.

Although sea-ice cover dramatically decreases over these two areas, the increase in net water flux from atmosphere to ocean is not related to this (Figure 16). It is shown that the increase in river run-off from the Antarctic continent (read snow melt and iceberg calving) plus precipitation (rain and snow) over the area (which also includes rain and snow over sea-ice) almost completely explains the increase in net water flux and when the increased evaporation (which acts as a negative contribution to this water flux) is included, the balance is closed. The reason for this large increase in water flux is thus the very strong increase in the hydrological cycle and transport of atmospheric water (vapor) from further north to Southern Hemisphere subpolar areas, in particular the eastern Pacific and western Atlantic sectors. It should be noted that CanESM5 features one of

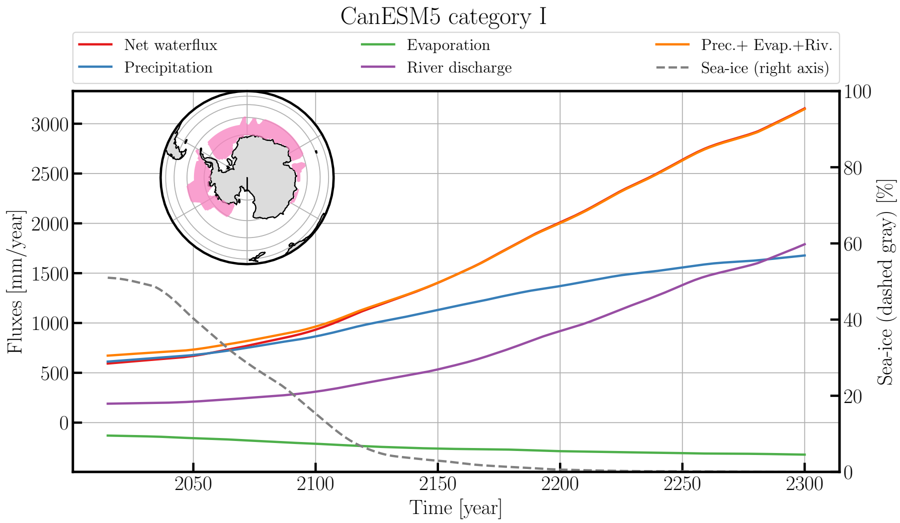

Figure 16. Case "I" for the CanESM5 model (r1i1p1f1) as found in the net water flux, split by precipitation, evaporation and river discharge (purple). By summing (Prec. + Evap. + Riv.) the contributions of precipitation, evaporation, and river discharge we are able to close the water budget in this region. The right axis additionally shows the sea-ice cover in this region. The inset shows the region of the category I case in the CanESM5 model, where the region of the SNS is illustrated as the pink shaded region.

the highest Equilibrium Climate Sensitivities of CMIP6 (5.64 (Zelinka et al., 2020)), and as a result the intensification of the hydrological cycle is particularly strong in this model.

3.10 Global temperature rise at SNS

In the previous subsections, we discussed each of the 11 categories of SNS one by one. To see how often SNS occur in the available data, we first calculated the number of simulations that reach a given global warming level (since 1850-1880). To emphasize the contributions of the different SSPs, we considered each SSP for each model separately. That way, we have roughly 200 model-SSP combinations. Next, we needed to obtain the period in which each shift occurs. This is not always easily identifiable, for example panel iii in Figure 2 shows a sharp decline around year 2035, but also in year 2055, one might therefore argue that the SNS occurs at either of these years (and the temperatures associated with these years) or at a time in between these years. To overcome this ambigity, we used a flexible window w of 20-50 years to find the years (and the temperatures associated with them) at which the difference in w years in the w-year running mean is maximal, also see subsection A4 for more information. The window w should be long enough to prevent measuring short term fluctuations as those apparent in panel v in Figure 2, where before 1900 there are large, decadal time-scale changes (while the real change of interest evolves slowly in 2000-2050). The same window, on the other hand, should not be so long that it will span the entire scenario run duration since then all global warming temperatures will be associated with the SNS (such as Figure 2 panel i, which, using such a long window, always occurs in 2050). The flexible window w of 20-50 years strikes a balance between these two opposing interests and gives a PDF for the timing (and temperature) of maximum change.

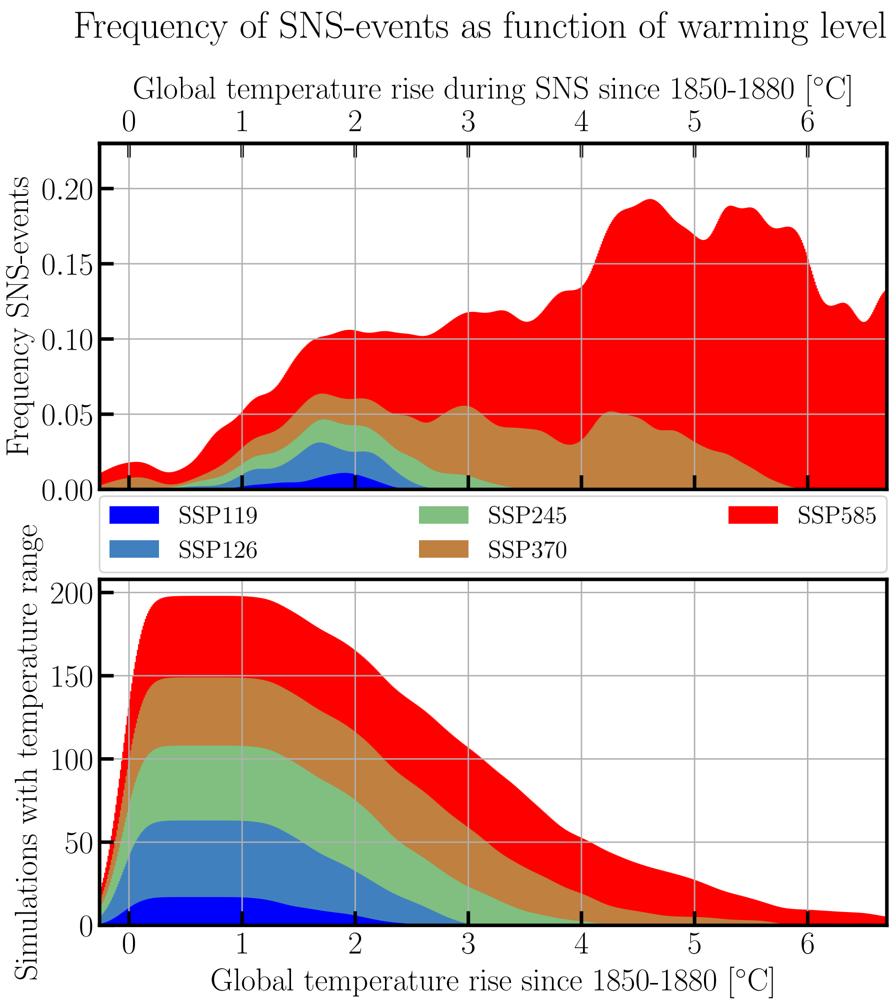

Figure 17. Frequency of SNS (top) as a fraction of the number of available simulations that reach such a temperature range (bottom). The frequency of a shift is calculated by dividing the summed fractional probability of an SNS to occur in a temperature interval, divided by the number of available simulations that cover that interval (see text for details). The data is smoothed using a Gaussian convolution with a standard deviation of $0.1\,^{\circ}$C. The temperature scale (x-axis) is limited to $-0.3-6.7\,^{\circ}$C, where at least five simulations are available (bottom panel).

We then calculate the frequency at which temperature the SNS occur by summing the PDFs of all SNS and dividing that by the number of models that reach that global warming level. This leads to Figure 17 where the bottom shows the number of available simulations for each temperature rise, while the top panel shows the frequency of SNS at each warming level. We normalize this distribution so that each change is once accounted for (even if it is observed in more than one variable). Additionally, models may have more than one ensemble-member, if this is the case, each of their temperature distributions is divided by the total number of ensemble-members. The top panel therefore shows (as a function of global temperature rise

ΔT):

$$\text{Freq.}(\Delta T) = \frac{\sum\limits_{\text{SNS}} \sum\limits_{i}^{n_{\text{variables}}} \frac{1}{n_{\text{variables}}} \sum\limits_{j}^{n_{\text{ensemble}}} \frac{1}{n_{\text{ensemble}}} \cdot \text{PDF}_{i,j}(\Delta T)}{\sum\limits_{\text{model, SSP}} \sum\limits_{j}^{n_{\text{ensemble}}} \frac{1}{n_{\text{ensemble}}} \text{simulation at } \Delta T}. \tag{1}$$

The frequency of SNS events (Figure 17) increases steeply at global warming levels below 2,°C and remains relatively high between 2–4,°C. While the frequency of SNS rises further between 4–6,°C, this is based on a rapidly decreasing number of models (dropping from ∼54 to ∼10 in this interval), which correspondingly decreases the robustness of the inferred frequency. We should emphasize that Figure 17 illustrates the warming level at which the changes in the SNS are maximal, meaning that the initialization of the SNS always precedes this time. If a tipping point is associated to the SNS, it is passed *before* the temperature noted here, when the transition to the new state is occurring.

We limit Figure 17 to $-0.3 - 6.7\,°\mathrm{C}$ global warming levels, where at least five simulations are available. The small, unexpected bump near $0°\mathrm{C}$ warming is an artifact (despite the method explained in subsection A4) of the large maximum window size (50-years) that includes up to 25 years before (or after) the year of maximum change, combined with the very few simulations (bottom panel) that pass this temperature range.

We should note that the extended model runs (beyond the year 2100) are disproportionately represented in the SNS. The number of extended models (each normalized by the number of ensemble-members available to that model) compared to the total number of available models clearly illustrates this overrepresentation. The number of normalized extended models / number of available models is: 0.3/17 (2.0 %) for SSP119, 4.5/46 (10 %) for SSP126, 1.6/45 (3.6 %) for SSP245, 0.0/41 (0.0 %) for SSP370, 3.0/49 (6.2 %) for SSP585. The relative frequency of extended runs in those models that show an SNS is much higher: 0.0/2 (0.0 %) for SSP119, 5.0/10 (50 %) for SSP126, 1.8/10 (18 %) for SSP245, 1.0/16 (6.3 %) for SSP370, 7.0/26 (27 %) for SSP585. The larger availability of extended runs compared to Drijfhout et al. (2015) is one additional reason why the frequency of SNS appears higher in our analysis.

In Figure 18 we illustrate the CDF of the SNS occurring as a function of a global temperature rise. We see that the state transitions (uppercase SNS categories) generally occur at later timescales compared to the abrupt shifts. There are several reasons for this difference between the categories. As seen in e.g. subsection 3.1, the change in sea-ice is sometimes observed as abrupt in the low-emission scenarios, while it is seen as a state transition in the high emission scenario where the disappearance of sea-ice is complete (see e.g. #6 and #27 in Table A3). Additionally, many state transitions are observed in the extended datasets. From Figure 17, we can infer the read of the temperature at which, for each category, the change in the SNS is maximal, which we include in Table 1.

4 Discussion

In this work, we have searched for SNS in CMIP6 data, focusing on ocean, sea-ice, and atmospheric variables. We identified 73 SNS that we grouped into 11 categories. Compared to Drijfhout et al. (2015), where 37 SNS were identified, this represents

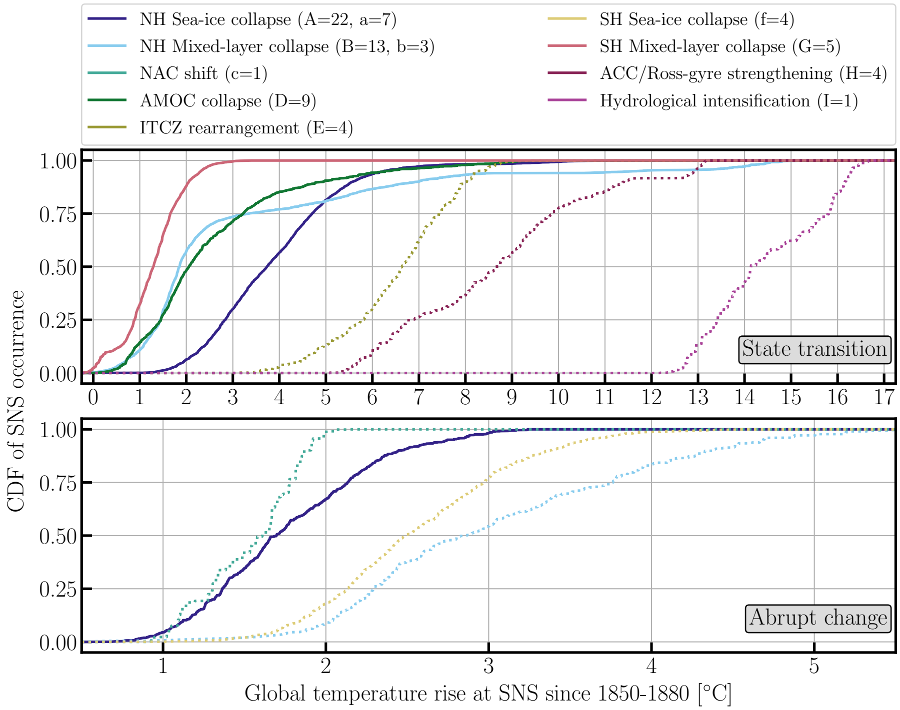

Figure 18. Conditional CDF for SNS occurring per category as function of global warming level. The CDF quantifies the conditional probability that the change of the SNS is maximal (i.e. has a maximal derivative) at that global warming level, as explained in subsection 3.10 and subsection A4. The top panel gives the CDF for the state transitions (uppercase) categories while the bottom panel gives the CDF of the abrupt change SNS (lowercase) categories. The state transitions are often observed at (much) higher global warming levels, which therefore are displayed in a different global temperature rise range (x-axis) than the abrupt changes. The legend lists the number of models for each category. The categories that are observed for < 5 models, have a higher uncertainty associated to their CDF (therefore displayed by dotted lines), compared to the categories with ≥ 5 models (solid lines).

a large increase. It is unclear however, to what extent this increase is related to differences between CMIP5 and CMIP6 or due to our improved detection method.

The method used here made substantial changes to the approach discussed in Drijfhout et al. (2015). It allows for the isolation of arbitrarily shaped regions, while the region-selection in Drijfhout et al. (2015) was performed manually in rough $5° \times 5°$ rectangular longitude-latitude boxes and strict selection was already implied by the expert making the region-selection. In this work we now also rely more on formal, stringent criteria-sets and apply those to the arbitrarily shaped regions. Compared to the method used in Drijfhout et al. (2015) our method has become fully reproducible, as it no longer relies on expert judgment of time series.

In three cases we find it unlikely that the change in method can explain the increased numbers of SNS found here. Those are category A/a (Northern Hemisphere sea-ice collapse); category D (AMOC collapse), where we now find cases for lower emission scenarios as evident Drijfhout et al. (2015) and category B/b (Northern Hemisphere mixed-layer collapse).

Category	Description	Median [°C]	Range (68% C.I.) [°C]
A	NH Sea-ice collapse	4	2.5 – 5
a	NH Sea-ice collapse	1.7	1.3 – 2.3
B	NH Mixed-layer collapse	1.9	1.2 – 6
b	NH Mixed-layer collapse	2.9	2.2 – 4
c	NAC shift	1.6	1.1 – 1.8
D	AMOC collapse	2.1	1.1 – 4
E	ITCZ rearrangement	7	5 – 8
f	SH Sea-ice collapse	2.5	2.0 – 3
G	SH Mixed-layer collapse	1.3	0.7 – 1.9
H	ACC/Ross-gyre strengthening	9	6 – 11
I	Hydrological intensification	14	13 – 16

Table 1. Global temperature rise at which each SNS categories occurs. The global temperature identified for each SNS indicates the the time at which the derivative is highest, like in Figure 18). The median corresponds to 0.5 in Figure 18, while the 68% confidence interval (C.I.) are the values of 0.16 and 0.84.

We found abrupt changes and state transitions in 11 categories. For Arctic sea ice there is a clear shift of the onset of a year-round ice-free Arctic to lower temperature increases, implying this could occur already at the end of the 21st century. This shift is associated with a reduced bias in CMIP6 in sea-ice thickness in the historical period associated with a too cold Arctic, see also Angevaare et al. (2025). A collapse of convective deep mixing also occurs faster and earlier in CMIP6 than in CMIP5 and may already develop between now and 2050, see also Drijfhout et al. (2025). One reason for this likely is a reduced stratification in the North Atlantic SPG. It is non-trivial to check the role of model bias for each individual case, but in general model bias is stronger in models not showing a collapse than in model showing a collapse, see Sgubin et al. (2017); Swingedouw et al. (2021). Associated with the mixed-layer collapse also the AMOC features earlier shutdowns, a collapse of the North Atlantic Deep Water cell leaving a residual largely wind-driven overturning that does not extend northward of 30°N, see Drijfhout et al. (2025). In some cases, the shutdown even manifests itself in an SSP126 scenario that is in agreement with the Paris Agreements. Model spread in the AMOC response to global warming remains difficult to relate to model bias. One may argue that many models are too sensitive to aerosol forcing (Robson et al., 2022). The role of anthropogenic aerosol forcing in the 1850–1985 strengthening of the AMOC in CMIP6 historical simulations. Implying an AMOC decrease too strongly forced by aerosol removal since the nineties in the historical simulations and especially in the SSP126 scenario afterwards. On the other hand, it was recently shown that using observational constraints on model biases the AMOC decline in year 2100 became $\sim 70\%$ larger than using the unconstrained multi-model ensemble, implying that the better, less-biased models showed larger declines Portmann (2025). Also, CMIP models do not include the freshwater flux from Greenland Ice Sheet melting and thus underestimate the freshwater forcing that enables an AMOC shutdown via increasing stratification in the North Atlantic SPG and inhibiting convective mixing there.

We also showed that a collapse of deep mixing in the SPG precedes the AMOC shutdown and is a clear example of cascading tipping points. The interaction between deep convection, regional sea-ice cover, air/sea fluxes and changes in ocean and atmospheric surface temperature sea surface salinity also shows elements of a cascade, but we would argue these are not separate tipping points but a cascade of tipping variables within a common tipping point/element.

A few cases we found were clearly associated with model biases but still contained valuable information for more general tipping point behaviour of the pertinent climate subsystem and possible routes how to improve model bias. These examples were an abrupt shift in the North Atlantic Current in a model where its pathway was biased towards a more realistic pathway, a shift from a double ITZC (model bias) to a more single ITCZ (realistic), large freshwater flux increases in parts of the Southern Ocean neighbouring the Antarctic continent in a model with an extremely high climate sensitivity and in an extended SSP585 simulation, abrupt changes in deep convection at places in the Southern Ocean where open ocean convection is not observed, and a strong response of the ACC to future increasing Southern Hemisphere winds, which is deemed unrealistic as CMIP models do not include the break of mesoscale eddies on such increase due to their coarse resolution. This underscores one of the main findings in a previous CMIP5 catalogue of SNS (Drijfhout et al., 2015), namely that many SNS appear stochastic and only occur in some models and/or even in some ensemble members of one model, implying that small changes in background state and or atmospheric variability (noise) may determine whether an SNS occurs or not. A nice example of the stochastic bifurcation between a 10-member ensemble simulation featuring two members with an AMOC collapse and eight members in which the AMOC recovered under an SSP245 scenario illustrates this point (Romanou et al., 2023). This also implies that an SNS found in one or a few models may not have a counterpart in the present or future climate of our planet, but that we cannot rule this out unless the model at hand is severely biased and the SNS depends on this biased climate in a crucial way. This is probably true for the SNS associated with a shift in the North Atlantic Current and with a shift from a double to a single ITCZ in the tropical Pacific.

Some SNS in the Southern Ocean, although occurring in one or only a few model simulations, could be realistic and cannot be simply discarded as results from model bias. These are sudden disappearances of sea ice in regions where deep convection is triggered, a collapse of deep convection and mixed-layer depth in other Southern Ocean regions, and an eastward expansion and strengthening of the Ross gyre. The first two types of SNS are associated with a well-known convection-sea ice interaction demonstrated to occur in the Arctic (in observations and in models), and playing a role in low-frequency changes in sea ice and mixed-layer depth in the Southern Ocean, demonstrating the realism of the mechanism in the Southern Ocean as well. The eastward expansion of the Ross gyre is demonstrated in various eddy-resolving ocean-only models and is not affected by the inability to resolve eddies.

It is impossible to compare our results on a one-by-one basis with those obtained in Terpstra et al. (2024) who performed a somewhat similar but more restricted analyse as we have done here, also (partly) focusing on other variables than we do, but more importantly, analysing results from another scenario, namely the 1ptCO2 scenario, which starts in the preindustrial and increases every year with 1% its CO_2 concentration until the preindustrial value is doubled. What is immediately apparent, is that our and their approach both lead to the conclusion that in CMIP6 thresholds for abrupt shifts/SNS are earlier passed and at lower temperatures for all variables that we both analysed. In that respect our results confirm each other. In this work, we have focused on ocean, sea-ice, and atmospheric variables that are involved in the coupling to ocean and sea-ice changes. A logical next step would be the assessment of vegetation, carbon-cycle and related CMIP6 variables, as also Drijfhout et al. (2015)

found a substantial number of SNS in this regime, as well as assessing SNS in fast atmospheric systems including extreme weather.

5 Conclusions

We have searched for SNS in CMIP6 data, with a focus on ocean, sea-ice, and atmospheric variables. We aimed to identify both abrupt changes over a decadal timescale and more gradual state transitions too large to be explained by increasing forcing only. We isolated 73 SNS that we categorized in 11 categories, which we each discussed. A large number $45\,\%$ of SNS is attributed to sea-ice loss (categories A, a and f) relates to sea-ice cover, $29\,\%$ to mixed-layer depth (categories B, b and G), $19\,\%$ to ocean currents (categories c, D and H), and $7\,\%$ to atmospheric systems like the Intertropical Convergence Zone (categories E and I). The number of SNS steeply increases between global warming levels below $2\,^{\circ}\mathrm{C}$, stays at that level between $2 – 4\,^{\circ}\mathrm{C}$ before rising further until around $6\,^{\circ}\mathrm{C}$, few models reach such high global warming levels. In general, the CMIP6 data contain various SNS that occur at lower temperature increase and earlier in time than estimated in previous assessments of SNS (abrupt changes) and tipping points, pointing to a increased sensitivity of the newest generation of climate models to anthropogenic forcing and an increased risk to strong nonlinear surprises, especially those associated with Arctic sea ice, the AMOC and convective mixing in the North Atlantic SPG. This finding signifies a strong warning to limit further future global warming as much as possible.

Appendix A: Appendix

A1 Region isolation example

In this section, we provide an example of the first region-isolation approach discussed in subsection 2.2. The approach aims to select regions that have a high probability to fulfil any of the formal SNS criteria of subsection 2.3. In this example we perform the region isolation for SSS, GISS-E2-1-G, SSP245 (ensemble r1i1p5f1), corresponding to case number #44 in Table A3. This example, as well as the other three region isolation approaches, is also available online (Angevaare, 2025a), specifically https://github.com/JoranAngevaare/optim_esm_tools/blob/master/notebooks/advanced_methods.ipynb.

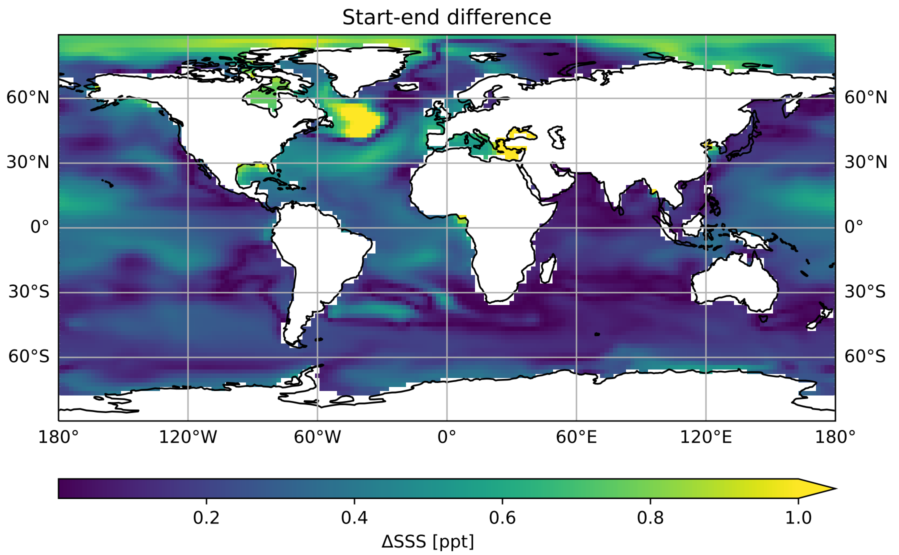

Figure A1. Difference between the start (year 1850) and end (year 2100) of the dataset (GISS-E2-1-G, SSP245, ensemble r1i1p5f1) for SSS, measured in parts per thousand (ppt). Since we use the 10-year running mean, we use years 1855 and 2095 as the first and last year, respectively. In the region finding approach, we select regions that have a high start-end difference compared to the rest of the dataset, such as for example the polar-latitudes, the subpolar-gyre. The result of the region-selection is illustrated in Figure A2.

The difference between the start and end of the time series (years 1855 and 2095 respectively, because of the 10-year running mean) is illustrated in Figure A1. We now describe step by step how the start-end difference region isolation approach (subsection 2.2) selects regions from this 2D-map. The method first evaluates if any region passes a very high threshold of the 99.99%-percentile. For didactic purposes, let's assume this threshold corresponds to 1.0 parts per thousand (ppt). We produce a Boolean map if the start-end difference exceeds this threshold. Using an unsupervised clustering algorithm (Campello et al., 2013), we build continuous regions of the Boolean map. This yields one or more clusters on the map with regions higher than the 1.0 ppt criterion. Now, we only keep those clusters that also exceed the area-requirement of $10^6\,\mathrm{km}^2$. It happens rarely that the first threshold (99.99%) yields any regions that are sufficiently large (since only 0.01% of the grid-cells should exceed it). So let us assume we did not find any cluster at the first iteration. If we lower the threshold, and run the algorithm again (for

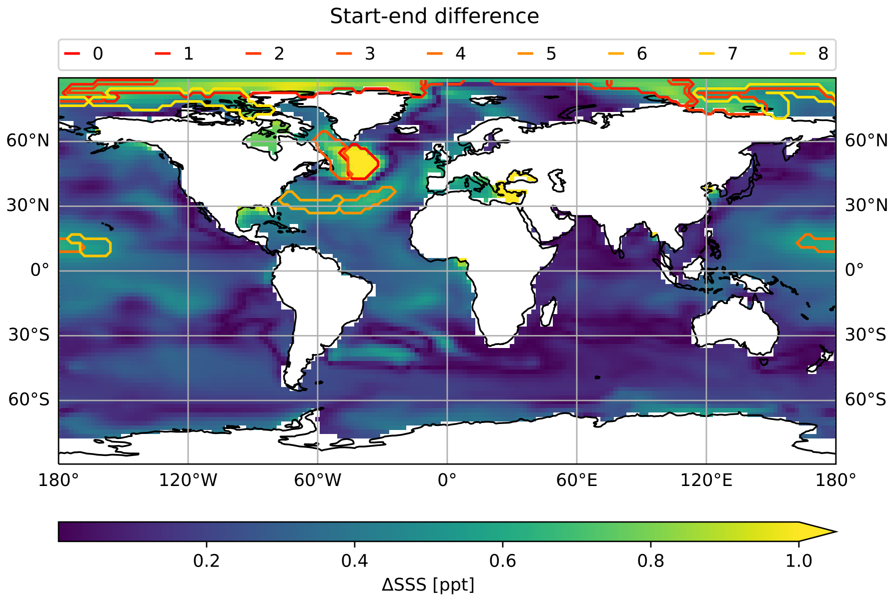

Figure A2. Difference between the start and end of the dataset (GISS-E2-1-G, SSP245, ensemble r1i1p5f1) for SSS, like Figure A2, with the nine identified regions (labelled from 0-8) overlaid. The order of the regions is determined by the order in which the regions are selected. This means that region 0, the sub-polar gyre, was first identified as it was the first region to surpass (high) a percentile threshold and be of $\geq 10^6\,\mathrm{km}^2$.

example, with a threshold of 0.9 ppt difference), the clusters are generally larger than in previous step. It may now happen that a clustered region is sufficiently large. We save this region, but also check again if a lower threshold (for example, 0.8 ppt) yields any new clusters. We emphasize that it should be new, since we exclude the first selected cluster from the Boolean mask (since we already identified that region). We keep repeating these steps, so at the lowest percentile (85%), we have most likely already identified some regions, usually in the order of 1-10 regions. These 1-10 regions are then the input for the next phase (subsection 2.3). To illustrate the contrast to the explained method, we could have instead only applied the 85%-percentile threshold. We would have gotten for example, two regions that $\sim 5 - 10 \times 10^6\,\mathrm{km}^2$. These two regions might then wash out any SNS, which is not in line with the goal of this step. This is why we iterate the thresholds to get smaller regions ($\geq 10^6\,\mathrm{km}^2$) such that we obtain sufficiently fine-grained regions to identify SNS.

The iterative approach described above, applied to Figure A1, will first select the large start-end difference near the subpolar gyre, followed by regions near the arctic. When the actual thresholds of the method are used, this method isolates nine distinct regions to be passed on to the next step, where we check each region for the occurrence of SNS. The nine distinct regions are illustrated in Figure A2.

The other three approaches (see subsection 2.2) to identify regions work similarly and are all available online https://github.com/JoranAngevaare/optim_esm_tools/blob/master/notebooks/advanced_methods.ipynb (Angevaare, 2025a). Furthermore, this material also illustrates how the regions from multiple approaches are joined in the next step of the algorithm where the formal criteria are applied and regions are iteratively expanded (subsection 2.3).

A2 Formal criteria

Below we give the six criteria and their respective thresholds. The criteria were designed to isolate regions with spatially averaged time series that are extreme for the available CMIP6 data. Criteria Ab1 and Ab2 are specifically designed to select regimes that feature a single abrupt change. These two criteria are general purpose and are applied to all CMIP6 variables in Table A1. Criteria St1-St4 are instead formalized to select large-scale state transitions, which occur on longer timescales. Criteria St1-St3 are for specific variables, while criterion St4 is general purpose.

The criteria use slightly different metrics to achieve these aims. Below, we discuss the criteria one by one. Most metrics are related, e.g. the abrupt criteria (Ab1 and Ab2) use a maximum jump in 10 years in the time series in places where the state transition (criterion St4) uses the difference between the start and end of the time series. The selection of each threshold in the criteria below was an iterative process and based on careful selection by the authors. They are therefore subjective, and while criteria Ab1, Ab2 and St4 are designed to be general purpose, CMIP6 variables that were not investigated in this work may require better suited thresholds or optimized criteria.

A2.1 Criterion Ab1

Criterion Ab1 is designed to select abrupt changes. To this end we calculate the spatially averaged time series (below, we will simply call this the time series) of an isolated variable (Figure 1). As we are interested in large changes in the time series, we calculate the maximum jump (MJ) in a 10-year window, after smoothing the time series with a 10-year running mean. We want the maximum jump to be large with respect to the pi-control dataset and the average variability in the scenario dataset. To this end, we calculate the standard deviation of the 10-year running mean of the detrended time series of the pi-control dataset within the region, which we call σ_{PiC}. We require that the jump seen in the scenario is at least $4.5 \times \sigma_{\text{PiC}}$ as in the first line of Equation A1.

Additionally, we calculate the average standard deviation of all cells, based on their individual 10-year running mean of the detrended values. This gives us a measure of the average global variability in the pi-control. However, since the tropical latitudes have different average standard deviations, we only take the average over grid cells within the $23.5°\text{S} : 23.5°\text{N}$ band if $50\,\%$ of the grid cells of the selected region are in that latitude range. In case the selected region has less than $50\,\%$ in the $23.5°\text{S} : 23.5°\text{N}$-band, we instead take the average over grid cells that are outside this band. We call this average standard-deviation of the pi-control dataset $\bar{\sigma}_{\text{PiC trop}}$. We require that MJ is extreme, compared to the average variability in that zone (either tropics or outside of the tropics) by imposing $MJ/\bar{\sigma}_{\text{PiC trop}} > 10$ (second line of Equation A1).

Similar to $\bar{\sigma}_{\text{PiC trop}}$, we also require that the MJ is extreme with respect to the average standard-deviation of the 10-year running mean of the detrended time series in a large-scale zone (tropics or outside tropics). Completely analogous to $\bar{\sigma}_{\text{PiC trop}}$, we call this average $\bar{\sigma}_{\text{trop}}$ require that it is $5\times$ smaller than MJ.

Furthermore, we require that the (un-smoothed, yearly) time series is not unimodal, but bi- or multimodal by applying the dip test for unimodality (Hartigan and Hartigan, 1985) and requiring that its value is below $1\,\%$.

Finally, we select time series that show one clearly identifiable jump. We compare the largest jump (MJ) to the second largest jump of the same sign (that is at least 10 years removed from MJ). We call the ratio between these two jumps $J2$. As some time series have both increasing and decreasing jumps, we also explicitly calculate the largest jump that has the opposite sign from the largest jump anywhere in the time series. The ratio between these jumps we similarly call $J2_{\text{min}}$.

Together, the thresholds in Equation A1 select time series with single, abrupt (in the timescale of 10-year) jumps in the time series that are unique for this region, as well as large when compared to the standard deviation of either averages of the pi-control or scenario datasets. The abruptness should cause a detectable bi/multi-modality to pass the Diptest. The full criterion reads:

$$\begin{aligned}
MJ/\sigma_{\text{PiC}} &> 4.5 \wedge \\
MJ/\bar{\sigma}_{\text{PiC trop}} &> 10 \wedge \\
MJ/\bar{\sigma}_{\text{trop}} &> 5 \wedge \\
p_{\text{dip}} &< 1\,\% \wedge \\
\text{min}(J2, J2_{\text{min}}) &> 3.0 \wedge \\
\text{max}(J2, J2_{\text{min}}) &> 4.5\,.
\end{aligned} \tag{A1}$$

A2.2 Criterion Ab2

Criterion Ab2, similar to i, aims to detect abrupt changes. Criterion Ab1 often fails to identify abrupt changes when they occur at the end (or, theoretically, start) of the time series. When this occurs, a second state is not sufficiently sampled to pass the Diptest. By dropping the Diptest and increasing the requirement on the $J2$ and $J2_{\text{min}}$ parameters, we are able to select this type of time series that fail criterion Ab1. This leads to:

$$\begin{aligned}
MJ/\sigma_{\text{PiC}} &> 4.5 \wedge \\
\text{min}(J2, J2_{\text{min}}) &> 4.5 \wedge \\
MJ/\bar{\sigma}_{\text{trop}} &> 5 \wedge \\
MJ/\bar{\sigma}_{\text{PiC trop}} &> 10\,.
\end{aligned} \tag{A2}$$

A2.3 Criterion St1

Criterion St1 aims to select (almost) complete disappearance of sea-ice in a large-scale area. For an arbitrary region, this could easily be achieved, as a peripheral region of sea-ice may melt in a warming climate (e.g. starting from $50\,\%$ sea-ice cover en ending with $0\,\%$). Therefore, we require the regions to be at least $5 \times 10^6\,\mathrm{km}^2$. Additionally, we require that the disappearance is at least $95\,\%$ of the maximal value in an area that is at least as large. We compute the maximal average in a connected area (if it is of the same size) with the maximal sea-ice values. The ratio between the start and end time of the region of interest and this maximum value is what we call $ES_{S.I.}$, which should exceed $95\,\%$:

$$\begin{aligned} &\text{variable} \in \{\text{siconc, siconca}\} \wedge \\ &ES_{S.I.} > 95\,\%\,. \end{aligned} \tag{A3}$$

A2.4 Criterion St2

Criterion St2 is similar to criterion St1 and focuses on a large percentage drop in the time series. This criterion applies only to msftyz, msftmz, and it uses the workflow as described in (Drijfhout et al., 2025) to analyze this data format. We calculate the decline between the maximum in the 10-year running mean ($\max(v_{\text{rm 10}}(t))$) and the last value in the 10-year running mean ($v_{\text{rm 10}}(t = t_{\text{end}})$), which we call $ME = \max(v_{\text{rm 10}}(t)) - v_{\text{rm 10}}(t = t_{\text{end}})$. We require that there is a loss of at least $70\,\%$ of the maximal value, as in the equation below:

$$\begin{aligned} \text{variable} \in \{\text{msftyz, msftmz}\} \wedge \\ ME/\max\left(v_{\text{rm 10}}(t)\right) > 70\,\%\,. \end{aligned} \tag{A4}$$

A2.5 Criterion St3

Criterion St3 is meant to isolate regions of dramatic and persistent loss of mixed-layer depth. We utilize a longer running mean of 50 years since the mixed-layer depth often shows large inter-annual variability. We require, analogous to criterion St2, that the change of the 50-year running mean of the time series since the maximum is $90\,\%$ of that maximum. Additionally, we use $\sigma_{\text{PiC rm 50}}$ to filter out regions that also have large variability in the pi-control dataset (similar to criteria i and ii). Finally, we also use $ME_{\text{trop rm 50}}$, which is the difference between the maximum and end of the spatially averaged time series in the tropics/outside the tropics. With this last metric, we also filter out regions that are showing large declines when considering only that region but are not standing out significantly when considering a large zonal average. Combining these metrics leads to:

$$\begin{aligned} \text{variable} = \text{mlotst-max} \wedge \\ ME_{\text{rm 50}}/ME_{\text{trop rm 50}} > 15 \wedge \\ ME_{\text{rm 50}}/\sigma_{\text{PiC rm 50}} > 15 \wedge \\ ME_{\text{rm 50}}/\max\left(v_{\text{rm 50}}(t)\right) > 90\,\%\,. \end{aligned} \tag{A5}$$

A2.6 Criterion St4

Criterion St4 is applied to variables not yet covered by criteria iii-iv. As a general-purpose search for state transitions, it uses the difference between start and end of the time series (SE), which has to be extreme compared to certain metrics. We compare SE to $\sigma_{\text{PiC trop}}$ and σ_{PiC} which have been discussed above. Similar to $ME_{\text{rm 50}}/ME_{\text{trop rm 50}}$ (criterion St3), we also here formulate a requirement on SE/SE_{trop} (although in contrast to mlotst-max, this value may be negative, so an absolute value is required). Finally, we filter out time series with a large decadal variability. We use the standard-deviation of the difference in the 50 and 10 year running mean ($\text{std}\,(v_{\text{rm 50}}(t) - v_{\text{rm 10}}(t))$) and require that SE is large with respect to this number. Combined, this leads to the following requirement:

$$\begin{aligned}
\text{variable} \notin \{\text{siconc}, \text{siconca}, \text{msftyz, msftmz, mlotst-max}\} \wedge \\
SE/\sigma_{\text{PiC}} > 50 \wedge \\
SE/\text{std}\,(v_{\text{rm 50}}(t) - v_{\text{rm 10}}(t)) > 40 \wedge \\
SE/\bar{\sigma}_{\text{PiC trop}} > 60 \wedge \\
\text{abs}\,(SE/SE_{\text{trop}}) > 4\,.
\end{aligned} \tag{A6}$$

A3 Variables and available data

We used CMIP6 variables in this work which we list, together with their short notations, in Table A1.

In Table A2 we list all the available data used for making this catalogue. The data are limited to those available at the time of writing, sets where a matching pi-control, historical and scenario run is complete and available.

A4 Temperature at maximal change of an SNS

In section 4 and Figure 10 we discuss the PDF/CDF (Cumulative Density Function) of shifts occurring at a certain global temperature rise. Figure A3 illustrates the method for obtaining the PDF/CDF for an example time series (SNS #6 of Table A3). To do this, we analyze the spatially averaged time series of an SNS, as shown in panel a. First, we determine the time at which the change occurs. We find the largest change in a w-year window in the time series that we smooth by a w-year running mean. We assign an equal probability to all years within that w-year window. We repeat this for w from 20 to 50 year, as SNS typically evolve on such timescales. This results in the PDF and CDF in panel b of Figure A3.

To account for this, we take all temperatures in a window (w) into account around the year where the w-year difference in the time series is maximal (a w-year running mean). Each temperature within that time region is treated with equal probability. We repeat this for w from 20 to 50 year to build a temperature distribution at which the change is maximal. Panel b shows the PDF and CDF of the timing at which the change is maximal. By weighing the temperature change since 1850-1880 (panel c) with the PDF from panel b, we are able to build a PDF of the shift occurring at a certain temperature rise (panel d).

CMIP6 variable	Short notation	CMIP6 description
hfds	Heatflux	Downward Heat Flux at the Sea Water Surface
mlotst	MLD	Ocean Mixed Layer Thickness Defined by Sigma-T
mlotst_max	MLD-max	Yearly maximum of mlotst
msftbarot	Barotropic Streamfunction	Ocean Barotropic Mass Streamfunction
msftmz	AMOC strength	Ocean Meridional Overturning Mass Streamfunction
msftyz	AMOC strength	Ocean Y Overturning Mass Streamfunction
siconc	Sea-ice, sic	Sea-Ice Area Percentage
siconca	Sea-ice, sic	Sea-Ice Area Percentage (Atmospheric Grid)
sos	SSS	Sea Surface Salinity
tas	TAS	Near-Surface Air Temperature
tos	SST	Sea Surface Temperature
wfo	Net waterflux	Water Flux into Sea Water
zos	SSH	Sea Surface Height Above Geoid

Table A1. The CMIP6 variables along with their sort notation (used throughout this paper) as well as the CMIP6 description. The variables msftyz and msftmz have a different data-format than the other CMIP6 variables, which are surface variables. The handling of msftyz and msftmz is therefore different as discussed in section 2 and only criterion St2 is applied to this data.

Model	Available data [Variable: SSP119, SSP245, SSP370, SSP585]
ACCESS-CM2	hfds: 0, 5, 5, 5, 5, mlotst_max: 0, 10, 10, 9, 10, msftbarot: 0, 4, 5, 5, 5, msftmz: 0, 10, 10, 9, 10, siconc: 0, 5, 5, 5, 10, siconca: 0, 10, 10, 10, 10, sos: 0, 4, 4, 4, 4, tas: 0, 5, 5, 5, 10, tos: 0, 5, 5, 5, 10, wfo: 0, 10, 10, 9, 10, zos: 0, 10, 10, 9, 10
ACCESS-ESM1-5	hfds: 0, 37, 40, 40, 39, mlotst_max: 0, 40, 40, 40, 38, msftbarot: 0, 38, 40, 40, 37, msftmz: 0, 39, 22, 40, 40, siconc: 0, 40, 40, 40, 40, siconca: 0, 40, 40, 40, 40, sos: 0, 39, 40, 40, 40, tas: 0, 40, 40, 40, 40, tos: 0, 40, 40, 40, 40, wfo: 0, 40, 40, 40, 39
BCC-CSM2-MR	mlotst_max: 0, 1, 1, 1, 1, msftbarot: 0, 1, 0, 0, 1, siconc: 0, 1, 1, 1, 1, sos: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 0, 1, zos: 0, 1, 1, 1, 1
CAMS-CSM1-0	hfds: 2, 2, 2, 2, 2, mlotst_max: 2, 2, 2, 2, 2, siconc: 2, 2, 2, 2, 2, siconca: 2, 2, 2, 2, 2, tas: 2, 2, 2, 2, 2, tos: 2, 2, 2, 0, 2, zos: 2, 2, 2, 2, 2
CAS-ESM2-0	siconc: 0, 2, 2, 2, 2, sos: 0, 1, 2, 2, 2, tas: 0, 2, 2, 2, 2, tos: 0, 1, 2, 0, 2, zos: 0, 1, 2, 2, 2
CESM2	hfds: 0, 3, 3, 3, 3, mlotst_max: 0, 3, 3, 3, 3, msftbarot: 0, 3, 3, 3, 3, siconc: 0, 3, 3, 3, 3, siconca: 0, 3, 3, 3, 3, sos: 0, 3, 3, 3, 3, tas: 0, 3, 3, 3, 3, tos: 0, 3, 3, 3, 3, zos: 0, 3, 3, 3, 3
CESM2-WACCM	hfds: 0, 1, 5, 1, 5, mlotst_max: 0, 1, 5, 1, 5, msftbarot: 0, 1, 5, 1, 5, msftmz: 0, 1, 5, 1, 5, siconc: 0, 1, 5, 1, 5, siconca: 0, 1, 3, 1, 5, sos: 0, 1, 5, 1, 5, tas: 0, 1, 5, 1, 5, tos: 0, 1, 5, 1, 5, zos: 0, 1, 5, 1, 5
CMCC-CM2-SR5	hfds: 0, 1, 1, 1, 1, mlotst_max: 0, 1, 1, 1, 1, msftbarot: 0, 1, 1, 1, 1, msftyz: 0, 1, 1, 1, 1, siconc: 0, 1, 1, 1, 1, sos: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 1, 1, wfo: 0, 1, 1, 1, 1, zos: 0, 1, 1, 1, 1
CMCC-ESM2	hfds: 0, 1, 1, 1, 1, mlotst_max: 0, 1, 1, 1, 1, msftbarot: 0, 1, 1, 1, 1, msftyz: 0, 1, 1, 1, 1, siconc: 0, 1, 1, 1, 1, sos: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 1, 1, wfo: 0, 1, 1, 1, 1, zos: 0, 1, 1, 1, 1
CNRM-CM6-1	hfds: 0, 6, 6, 6, 6, mlotst_max: 0, 6, 6, 6, 6, msftyz: 0, 6, 6, 6, 6, siconc: 0, 6, 6, 6, 6, sos: 0, 6, 6, 6, 6, tas: 0, 6, 6, 6, 6, tos: 0, 6, 6, 6, 6, wfo: 0, 6, 6, 6, 6, zos: 0, 6, 6, 6, 6
CNRM-CM6-1-HR	hfds: 0, 1, 1, 0, 0, siconc: 0, 1, 1, 1, 1, sos: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 1, 1, wfo: 0, 1, 1, 1, 1, zos: 0, 1, 1, 1, 1
CNRM-ESM2-1	hfds: 5, 5, 10, 5, 5, mlotst_max: 5, 5, 10, 5, 5, msftyz: 5, 5, 10, 5, 5, siconc: 1, 1, 1, 1, 1, siconca: 0, 0, 1, 0, 0, sos: 5, 5, 10, 5, 5, tas: 5, 5, 10, 5, 5, tos: 5, 5, 10, 5, 5, wfo: 5, 5, 10, 5, 5, zos: 5, 5, 10, 5, 5
CanESM5	hfds: 46, 50, 50, 50, 50, mlotst_max: 5, 50, 50, 50, 50, msftbarot: 5, 48, 48, 48, 47, msftmz: 5, 50, 50, 50, 49, siconc: 5, 50, 50, 50, 50, sos: 5, 50, 50, 50, 50, tas: 50, 50, 50, 50, 50, tos: 4, 50, 50, 50, 50, wfo: 5, 50, 50, 50, 50, zos: 5, 50, 50, 50, 49
CanESM5-1	hfds: 0, 0, 0, 0, 16, mlotst_max: 0, 0, 0, 0, 16, msftbarot: 0, 0, 0, 0, 17, msftmz: 0, 0, 0, 0, 18, siconc: 1, 1, 1, 1, 20, siconca: 2, 2, 20, 20, 18, sos: 0, 0, 0, 0, 18, tas: 2, 1, 2, 9, 20, tos: 1, 1, 1, 0, 18, wfo: 0, 0, 0, 0, 18, zos: 1, 1, 1, 1, 20
CanESM5-CanOE	hfds: 0, 3, 3, 3, 3, mlotst_max: 0, 3, 3, 3, 3, msftbarot: 0, 3, 3, 3, 3, msftmz: 0, 3, 3, 3, 3, siconc: 0, 3, 3, 3, 3, sos: 0, 3, 3, 3, 3, tas: 0, 3, 3, 3, 3, tos: 0, 3, 3, 0, 3, wfo: 0, 3, 3, 3, 3, zos: 0, 3, 3, 3, 3
E3SM-1-0	siconc: 0, 0, 0, 0, 4, tas: 0, 0, 0, 0, 5, tos: 0, 0, 0, 0, 5, zos: 0, 0, 0, 0, 5
E3SM-1-1	siconc: 0, 0, 0, 0, 1, sos: 0, 0, 0, 0, 1, tas: 0, 0, 0, 0, 1, tos: 0, 0, 0, 0, 1, wfo: 0, 0, 0, 0, 1, zos: 0, 0, 0, 0, 1
E3SM-1-1-ECA	siconc: 0, 0, 0, 0, 1, tas: 0, 0, 0, 0, 1
E3SM-2-0	zos: 0, 0, 0, 1, 0
EC-Earth3	hfds: 2, 8, 23, 8, 7, mlotst_max: 2, 8, 23, 8, 7, msftbarot: 2, 8, 23, 8, 7, msftyz: 0, 0, 15, 0, 0, siconc: 2, 6, 22, 8, 7, sos: 2, 8, 23, 8, 7, tas: 2, 8, 23, 8, 8, tos: 2, 8, 23, 8, 6, wfo: 2, 8, 23, 8, 7, zos: 2, 8, 23, 8, 7
EC-Earth3-AerChem	hfds: 0, 0, 0, 2, 0, mlotst_max: 0, 0, 0, 2, 0, msftbarot: 0, 0, 0, 2, 0, msftyz: 0, 0, 0, 2, 0, siconc: 0, 0, 0, 2, 0, siconca: 0, 0, 0, 2, 0, sos: 0, 0, 0, 2, 0, tas: 0, 0, 0, 2, 0, tos: 0, 0, 0, 2, 0, wfo: 0, 0, 0, 2, 0, zos: 0, 0, 0, 2, 0
EC-Earth3-CC	hfds: 0, 0, 9, 0, 1, mlotst_max: 0, 0, 9, 0, 1, msftbarot: 0, 0, 9, 0, 1, msftyz: 0, 0, 8, 0, 0, siconc: 0, 0, 9, 0, 1, siconca: 0, 0, 9, 0, 1, sos: 0, 0, 9, 0, 1, tas: 0, 0, 9, 0, 1, tos: 0, 0, 9, 0, 1, wfo: 0, 0, 9, 0, 1, zos: 0, 0, 9, 0, 1
EC-Earth3-Veg	hfds: 2, 7, 8, 4, 7, mlotst_max: 2, 6, 7, 4, 6, msftbarot: 2, 7, 8, 4, 7, msftyz: 0, 0, 1, 0, 0, siconc: 3, 7, 7, 5, 6, sos: 2, 7, 7, 4, 7, tas: 3, 7, 7, 6, 7, tos: 3, 7, 8, 5, 7, wfo: 2, 7, 8, 4, 7, zos: 3, 7, 7, 5, 7
EC-Earth3-Veg-LR	hfds: 3, 3, 3, 3, 3, mlotst_max: 3, 3, 3, 3, 3, msftbarot: 3, 3, 3, 3, 3, siconc: 3, 3, 3, 3, 3, siconca: 3, 3, 3, 3, 3, sos: 3, 3, 3, 3, 3, tas: 2, 2, 2, 2, 2, tos: 3, 3, 3, 3, 3, wfo: 3, 3, 3, 3, 3, zos: 3, 3, 3, 3, 3
FGOALS-f3-L	hfds: 0, 3, 3, 3, 3, mlotst_max: 0, 3, 2, 3, 2, msftbarot: 0, 3, 3, 3, 2, sos: 0, 3, 2, 3, 3, tas: 0, 1, 1, 1, 1, tos: 0, 3, 3, 0, 3, wfo: 0, 3, 3, 3, 3
FGOALS-g3	hfds: 1, 4, 4, 5, 4, mlotst_max: 1, 4, 4, 5, 4, msftbarot: 1, 4, 4, 5, 4, msftmz: 1, 4, 2, 5, 4, siconc: 1, 4, 4, 5, 4, tos: 1, 4, 4, 0, 4, wfo: 1, 4, 4, 5, 4, zos: 1, 4, 4, 5, 4
FIO-ESM-2-0	hfds: 0, 3, 3, 0, 3, siconc: 0, 1, 1, 0, 1, tas: 0, 3, 3, 0, 3, tos: 0, 1, 1, 0, 1, wfo: 0, 3, 3, 0, 2, zos: 0, 1, 1, 0, 1
GFDL-CM4	hfds: 0, 0, 1, 0, 1, mlotst_max: 0, 0, 0, 0, 1, siconc: 0, 0, 1, 0, 1, sos: 0, 0, 1, 0, 1, tas: 0, 0, 1, 0, 1, tos: 0, 0, 1, 0, 1, wfo: 0, 0, 1, 0, 1, zos: 0, 0, 0, 0, 1
GFDL-ESM4	hfds: 1, 0, 1, 1, 1, mlotst_max: 1, 1, 1, 1, 1, msftyz: 1, 1, 1, 1, 1, siconc: 1, 1, 3, 1, 1, sos: 1, 1, 1, 1, 1, tas: 1, 1, 3, 1, 1, tos: 1, 1, 3, 1, 1, wfo: 1, 1, 3, 1, 1, zos: 1, 1, 1, 1, 1
GISS-E2-1-G	hfds: 6, 10, 20, 19, 9, mlotst_max: 6, 10, 20, 19, 9, msftbarot: 6, 10, 20, 19, 9, msftmz: 6, 10, 20, 19, 9, siconca: 6, 10, 11, 19, 9, sos: 6, 10, 20, 19, 9, tas: 6, 10, 20, 19, 9, tos: 6, 10, 20, 19, 9, wfo: 6, 10, 20, 19, 9
GISS-E2-1-G-CC	zos: 0, 1, 1, 0, 0
GISS-E2-1-H	hfds: 1, 9, 9, 5, 9, siconc: 1, 8, 9, 5, 9, siconca: 1, 9, 9, 5, 9, sos: 1, 8, 8, 5, 8, tas: 2, 8, 10, 6, 10, tos: 2, 10, 10, 6, 10
GISS-E2-2-G	hfds: 0, 5, 5, 5, 5, mlotst_max: 0, 5, 5, 5, 5, msftbarot: 0, 5, 5, 5, 5, msftmz: 0, 5, 5, 5, 5, siconca: 0, 5, 5, 5, 5, sos: 0, 5, 5, 5, 5, tas: 0, 5, 5, 5, 5, tos: 0, 5, 5, 5, 5, wfo: 0, 5, 5, 5, 5
HadGEM3-GC31-LL	hfds: 0, 1, 5, 0, 3, mlotst_max: 0, 1, 5, 0, 4, msftbarot: 0, 1, 5, 0, 3, msftyz: 0, 1, 5, 0, 4, siconc: 0, 1, 5, 0, 4, sos: 0, 1, 5, 0, 3, tas: 0, 1, 5, 0, 4, tos: 0, 1, 5, 0, 4, wfo: 0, 1, 5, 0, 4, zos: 0, 1, 5, 0, 4
HadGEM3-GC31-MM	mlotst_max: 0, 1, 0, 0, 4, msftyz: 0, 1, 0, 0, 4, siconc: 0, 1, 0, 0, 4, sos: 0, 1, 0, 0, 1, tas: 0, 1, 0, 0, 4, tos: 0, 1, 0, 0, 4, wfo: 0, 1, 0, 0, 4, zos: 0, 1, 0, 0, 4
IITM-ESM	tas: 0, 1, 1, 0, 1
INM-CM4-8	msftmz: 0, 1, 1, 1, 1, siconc: 0, 1, 1, 1, 1, sos: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 1, 1, zos: 0, 1, 1, 1, 1
INM-CM5-0	msftmz: 0, 1, 1, 1, 1, siconc: 0, 1, 1, 1, 1, sos: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 1, 1, zos: 0, 1, 1, 1, 1
IPSL-CM5A2-INCA	siconc: 0, 1, 0, 1, 0, sos: 0, 1, 0, 1, 0, tas: 0, 1, 0, 1, 0, tos: 0, 1, 0, 1, 0, wfo: 0, 1, 0, 1, 0
IPSL-CM6A-LR	hfds: 6, 6, 11, 11, 6, mlotst_max: 6, 6, 11, 11, 6, msftbarot: 6, 6, 11, 11, 6, siconc: 6, 6, 11, 11, 6, sos: 6, 6, 11, 11, 6, tas: 6, 6, 11, 11, 6, tos: 6, 6, 11, 11, 6, wfo: 6, 6, 11, 11, 6, zos: 6, 6, 11, 11, 6
KACE-1-0-G	tas: 0, 3, 3, 3, 3
KIOST-ESM	hfds: 0, 1, 1, 0, 1, mlotst_max: 0, 0, 1, 0, 0, siconc: 0, 1, 1, 0, 1, sos: 0, 1, 1, 0, 1, tas: 0, 1, 1, 0, 1, tos: 0, 1, 1, 0, 1, zos: 0, 1, 1, 0, 1
MCM-UA-1-0	tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 1, 1
MIROC-ES2L	msftmz: 10, 8, 30, 10, 10, siconc: 10, 10, 30, 10, 10, sos: 10, 10, 30, 10, 10, tas: 10, 10, 30, 10, 10, tos: 10, 10, 30, 10, 10, wfo: 10, 10, 30, 10, 10, zos: 10, 10, 30, 10, 10
MIROC6	hfds: 0, 1, 1, 1, 1, mlotst_max: 0, 1, 1, 1, 1, msftbarot: 0, 1, 1, 1, 1, msftmz: 0, 3, 3, 3, 3, siconc: 50, 50, 50, 50, 50, sos: 50, 50, 50, 50, 50, tas: 50, 24, 50, 50, 50, tos: 50, 50, 50, 50, 50, wfo: 0, 41, 41, 41, 41, zos: 50, 50, 3, 3, 50
MPI-ESM1-2-HR	hfds: 0, 1, 1, 1, 1, mlotst_max: 0, 1, 1, 1, 1, msftbarot: 0, 1, 1, 1, 1, msftmz: 0, 1, 1, 1, 1, siconc: 0, 1, 1, 1, 1, siconca: 0, 1, 1, 1, 1, sos: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, tos: 0, 1, 1, 1, 1, wfo: 0, 1, 1, 1, 1, zos: 0, 1, 1, 1, 1
MPI-ESM1-2-LR	hfds: 29, 29, 29, 29, 29, mlotst_max: 30, 27, 30, 29, 30, msftbarot: 30, 28, 30, 29, 30, msftmz: 30, 28, 30, 29, 30, siconc: 30, 30, 30, 30, 30, siconca: 30, 30, 30, 30, 30, sos: 30, 27, 30, 29, 30, tas: 30, 30, 30, 30, 30, tos: 30, 30, 30, 30, 30, wfo: 30, 27, 30, 29, 30, zos: 30, 30, 30, 29, 30
MRI-ESM2-0	hfds: 5, 5, 5, 5, 6, mlotst_max: 1, 1, 1, 1, 2, msftbarot: 1, 1, 1, 1, 2, msftmz: 4, 4, 4, 4, 5, msftyz: 1, 1, 1, 1, 2, siconc: 5, 5, 5, 5, 5, siconca: 5, 5, 5, 5, 5, sos: 5, 5, 5, 5, 6, tas: 5, 5, 5, 5, 6, tos: 5, 5, 5, 5, 6, wfo: 5, 5, 5, 5, 6, zos: 1, 1, 1, 1, 2
NESM3	hfds: 0, 1, 1, 0, 1, mlotst_max: 0, 1, 1, 0, 1, siconc: 0, 1, 1, 0, 1, sos: 0, 1, 1, 0, 1, tas: 0, 2, 2, 0, 2, tos: 0, 1, 1, 0, 1, wfo: 0, 1, 1, 0, 1, zos: 0, 1, 1, 0, 1
NorESM2-LM	hfds: 0, 1, 3, 1, 1, mlotst_max: 0, 1, 3, 1, 1, msftbarot: 0, 1, 3, 1, 1, msftmz: 0, 1, 3, 1, 1, siconc: 0, 1, 3, 1, 1, sos: 0, 0, 2, 0, 0, tas: 0, 1, 3, 1, 1, tos: 0, 1, 3, 1, 1
NorESM2-MM	hfds: 0, 1, 2, 1, 1, mlotst_max: 0, 1, 2, 1, 1, msftbarot: 0, 0, 2, 1, 0, siconc: 0, 1, 2, 1, 1, sos: 0, 1, 2, 1, 1, tas: 0, 1, 2, 1, 1, tos: 0, 1, 2, 1, 1, wfo: 0, 1, 2, 1, 1
TaiESM1	hfds: 0, 1, 1, 1, 1, siconc: 0, 1, 1, 1, 1, tas: 0, 1, 1, 1, 1, zos: 0, 1, 1, 1, 1
UKESM1-0-LL	hfds: 5, 16, 5, 10, 5, mlotst_max: 5, 15, 13, 16, 5, msftbarot: 5, 15, 5, 10, 5, msftyz: 0, 16, 5, 16, 5, siconc: 5, 16, 12, 10, 5, siconca: 5, 16, 6, 16, 5, sos: 4, 14, 5, 10, 5, tas: 5, 16, 14, 14, 5, tos: 5, 16, 5, 10, 5, wfo: 5, 15, 5, 16, 5, zos: 5, 15, 5, 16, 5
UKESM1-1-LL	hfds: 0, 1, 0, 1, 0, mlotst_max: 0, 1, 0, 1, 0, msftbarot: 0, 1, 0, 1, 0, msftyz: 0, 1, 0, 1, 0, siconc: 0, 1, 0, 1, 0, siconca: 0, 1, 0, 1, 0, sos: 0, 0, 0, 1, 0, tas: 0, 1, 0, 1, 0, tos: 0, 1, 0, 1, 0, zos: 0, 1, 0, 1, 0

Table A2. Number of ensemble-members per model, per variable, per SSP. For each variable we list the five SSPs separated by commas such that e.g. "hfds: 0, 0, 1, 2, 1" indicates 0×SSP119, 0×SSP126, 1×SSP245, 2×SSP370, 1×SSP585 for hdfs.

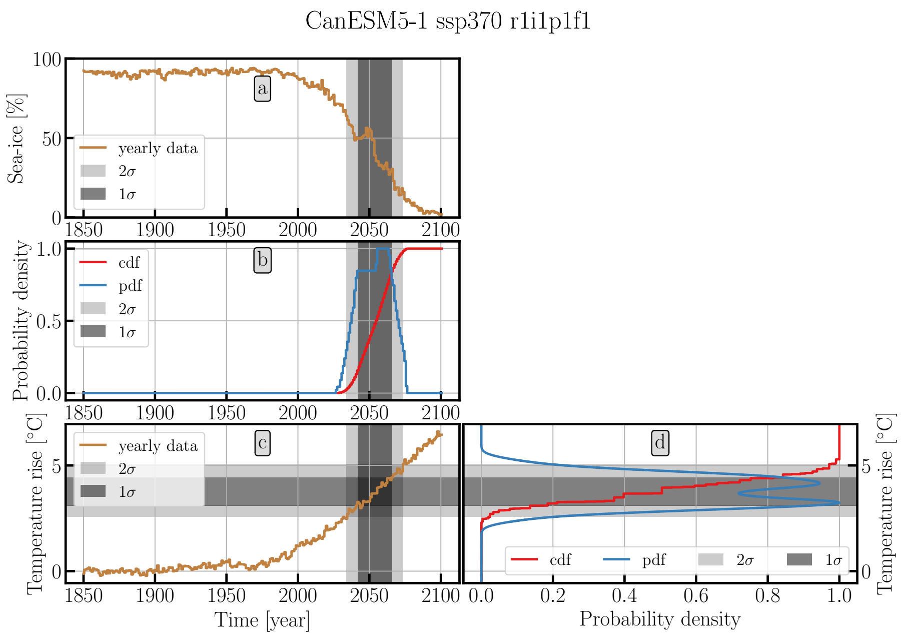

Figure A3. Method of obtaining the global temperature rise at shift for an SNS. We analyze the yearly data of panel a (see text for details) to determine when the change is maximal. The resulting CDF and PDF are shown in panel b. Based on the CDF in panel b, we draw the vertical 1σ and 2σ probability intervals. In panel c, we add the global temperature rise relative to the 1850-1880 average. By weighing the temperature rise values with the PDF of panel B we can calculate panel d, where we show the CDF and the PDF of the shift to occur at a certain temperature rise. The PDF in panel d is smoothed by a Gaussian with a standard deviation of 0.2 °C for illustration purposes. The horizontal 1σ and 2σ probability intervals are based on the CDF of panel d.

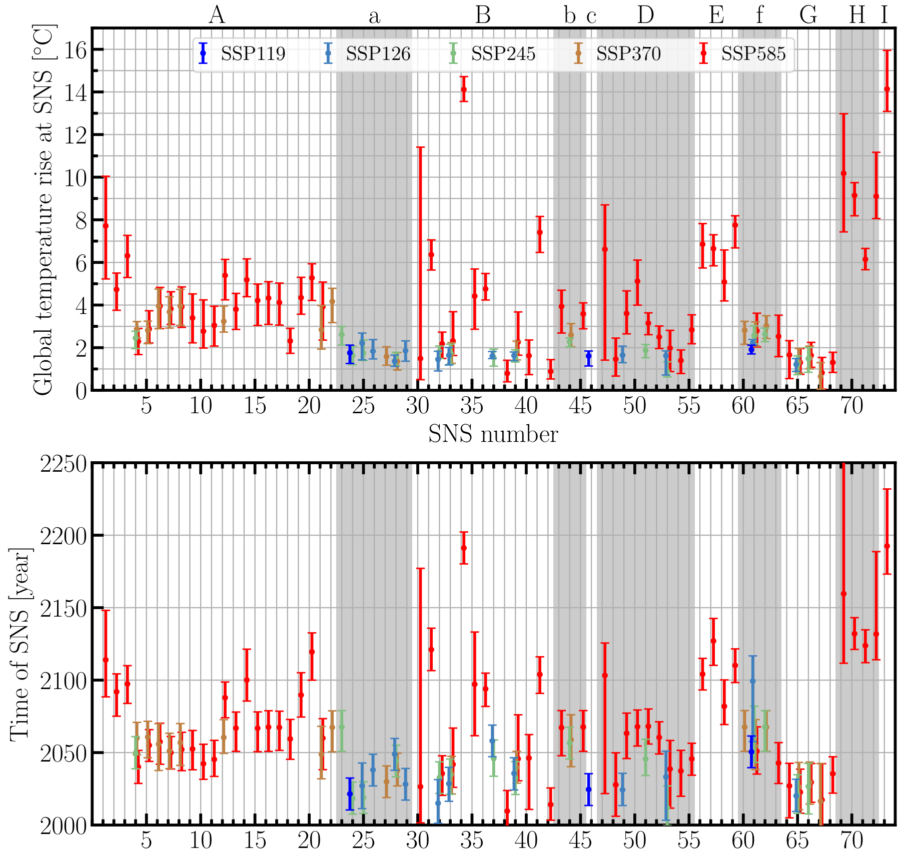

Figure A4. Median global temperature rise at SNS for all 73 SNS. The error bars indicate the 1σ confidence interval. Per case, the contributions of the SSPs are split (and slightly shifted horizontally for illustration purposes). The letters (top) denote the category to which the SNS belong and the categories are also illustrated using alternating background shades.

A5 SNS

In Table A3 we include each of the 73 SNS. As apparent from the table, one SNS is the occurrence of an event of one of the 11 SNS categories in a model. One SNS can be manifested in multiple variables or SSPs, as long as their associated regions are spatially overlapping. Some categories span a large spatial region, for this reason one model can result in more than one SNS per category, for example #8 and #9 in Table A3.

The global temperature rise at the time of the SNS within a given category may vary significantly per model (as we have seen in Figure 18). We can instead also calculate the CDF per SNS, as shown in Figure A4. Each of the 73 cases is plotted separately, and split per SSP. As expected from Figure A3, the 1σ-confidence interval in Figure A4 can be skewed, depending on the evolution of the time series and global temperature rise within the model.

Figure A5 illustrates the first ensemble-member per case, where we show each variable per SNS category in a single panel.

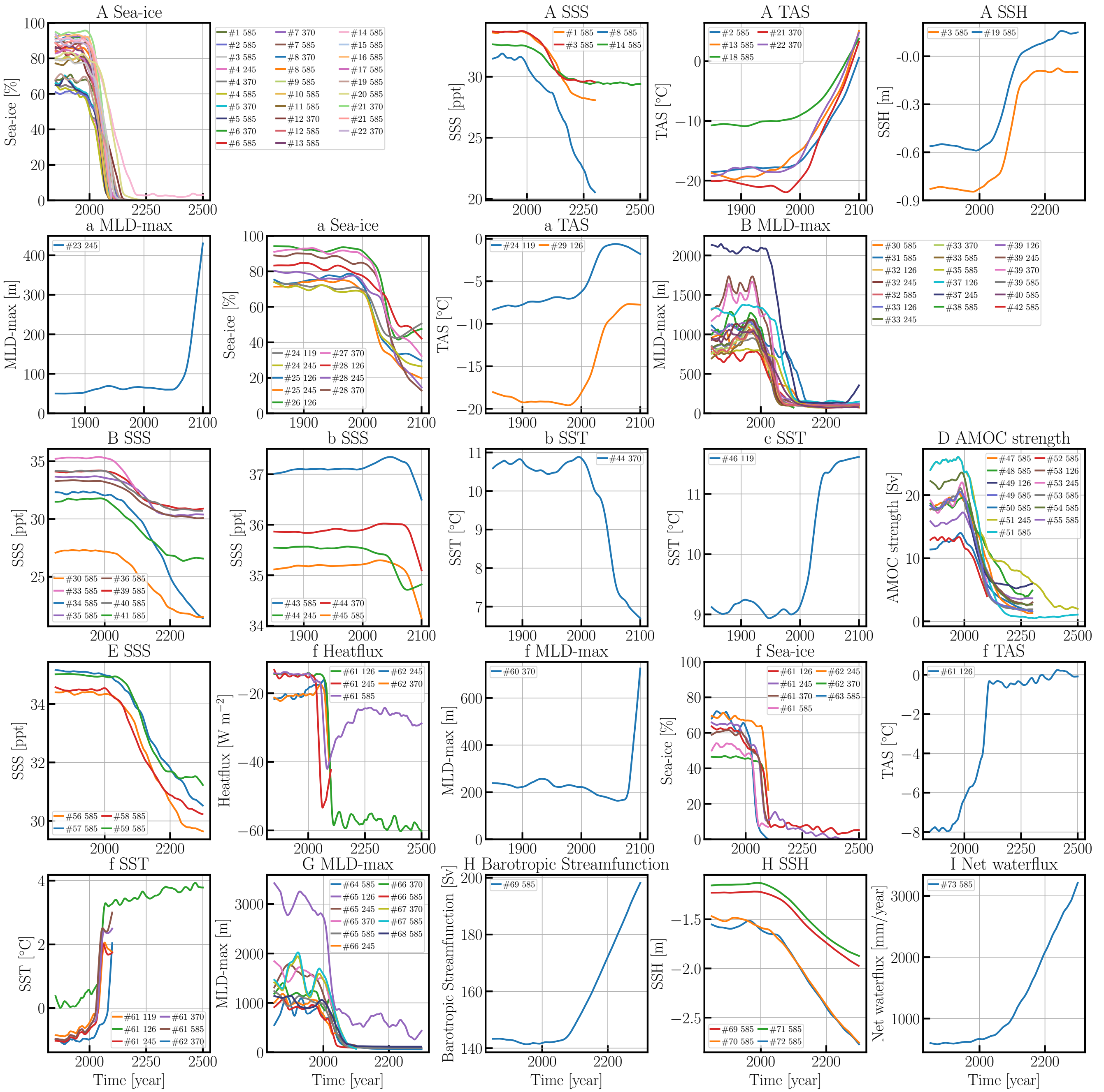

Figure A5. Time series of the first ensemble-member of each SNS case from Table A3. The time series are split per SNS category, and per variable within the SNS category, represented by the title of each panel. The legend labels refer to the SNS number in Table A3 and the SSP (119, 126, 245, 370 or 585). For illustration purposes, the time series are smoothed using a Lowess filter with a window of 40 year. Also see (Angevaare, 2025b) for each case separately.

#	Cat.	Model	SSP	Crit	St2ars	Region
1	A	ACCESS-CM2	585	St1, St4	siconc, sos	Arctic Ocean
2	A	ACCESS-ESM1-5	585	St1, St4	siconc, tas	Arctic Ocean
3	A	CESM2-WACCM	585	St1, St4	siconc, sos, zos	Arctic Ocean
4	A	CMCC-CM2-SR5	245, 370, 585	St1	siconc	Arctic Ocean
5	A	CMCC-ESM2	370, 585	St1	siconc	Arctic Ocean
6	A	CanESM5-1	370, 585	St1	siconc(a)	Arctic Ocean
7	A	CanESM5-CanOE	370, 585	St1	siconc	Arctic Ocean
8	A	CanESM5	370, 585	St1, St4	siconc, sos	Arctic Ocean
9	A	E3SM-1-0	585	St1	siconc	Arctic Ocean
10	A	E3SM-1-1-ECA	585	St1	siconc	Arctic Ocean
11	A	E3SM-1-1	585	St1	siconc	Arctic Ocean
12	A	EC-Earth3-Veg	370, 585	St1	siconc	Arctic Ocean
13	A	FIO-ESM-2-0	585	St1, St4	siconc, tas	Arctic Ocean
14	A	GISS-E2-1-H	585	St1, St4	siconc, sos	Arctic Ocean
15	A	HadGEM3-GC31-LL	585	St1	siconc	Arctic Ocean
16	A	HadGEM3-GC31-MM	585	St1	siconc	Arctic Ocean
17	A	Ab1PSL-CM6A-LR	585	St1	siconc	Arctic Ocean
18	A	MIROC6	585	St4	tas	Arctic Ocean
19	A	MRI-ESM2-0	585	St1, St4	siconc, zos	Arctic Ocean
20	A	NorESM2-MM	585	St1	siconc	Arctic Ocean
21	A	UKESM1-0-LL	370, 585	St1, St4	siconc(a), tas	Arctic Ocean
22	A	UKESM1-1-LL	370	St1, St4	siconc, tas	Arctic Ocean
23	a	CNRM-ESM2-1	245	Ab2	mlotst_max	Arctic Ocean
24	a	CanESM5-1	119, 245	Ab1	siconca, tas	Arctic Ocean
25	a	CanESM5	126, 245	Ab1	siconc	Arctic Ocean
26	a	EC-Earth3-Veg-LR	126	Ab1	siconc	Arctic Ocean
27	a	GISS-E2-1-H	370	Ab1	siconc	Arctic Ocean
28	a	MIROC6	126, 245, 370	Ab1	siconc	Arctic Ocean
29	a	UKESM1-0-LL	126	Ab1	tas	Arctic Ocean
30	B	ACCESS-CM2	585	St2, St4	mlotst_max, sos	North Atlantic Ocean
31	B	ACCESS-ESM1-5	585	St2	mlotst_max	North Atlantic Ocean
32	B	CESM2-WACCM	126, 245, 585	St2	mlotst_max	North Atlantic Ocean
33	B	CESM2	126, 245, 370, 585	St2, St4	mlotst_max, sos	North Atlantic Ocean
34	B	CanESM5	585	St4	sos	Baffin Bay
35	B	GISS-E2-1-G	585	St2, St4	mlotst_max, sos	Arctic Ocean
36	B	GISS-E2-1-H	585	St4	sos	Arctic Ocean
37	B	GISS-E2-2-G	126, 245	St2	mlotst_max	North Atlantic Ocean
38	B	HadGEM3-GC31-MM	585	St2	mlotst_max	North Atlantic Ocean
39	B	MRI-ESM2-0	126, 245, 370, 585	St2, St4	mlotst_max, sos	Arctic Ocean
40	B	NorESM2-MM	585	St2, St4	mlotst_max, sos	North Atlantic Ocean
41	B	UKESM1-0-LL	585	St4	sos	Baffin Bay
42	B	UKESM1-0-LL	585	St2	mlotst_max	Norwegian Sea
43	b	CESM2-WACCM	585	Ab2	sos	North Atlantic Ocean
44	b	GISS-E2-1-G	245, 370	Ab1, Ab2	sos, tos	North Atlantic Ocean
45	b	GISS-E2-1-H	585	Ab2	sos	North Atlantic Ocean
46	c	GISS-E2-1-H	119	Ab1	tos	North Atlantic Ocean
47	D	ACCESS-CM2	585	St2	msftmz	North Atlantic Ocean
48	D	ACCESS-ESM1-5	585	St2	msftmz	North Atlantic Ocean
49	D	CESM2-WACCM	126, 585	St2	msftmz	North Atlantic Ocean
50	D	CanESM5	585	St2	msftmz	North Atlantic Ocean
51	D	GISS-E2-1-G	245, 585	St2	msftmz	North Atlantic Ocean
52	D	MIROC-ES2L	585	St2	msftmz	North Atlantic Ocean
53	D	MRI-ESM2-0	126, 245, 585	St2	msftmz	North Atlantic Ocean
54	D	NorESM2-MM	585	St2	msftmz	North Atlantic Ocean
55	D	UKESM1-0-LL	585	St2	msftyz	North Atlantic Ocean
56	E	ACCESS-CM2	585	St4	sos	South Pacific Ocean
57	E	ACCESS-ESM1-5	585	St4	sos	South Pacific Ocean
58	E	CESM2-WACCM	585	St4	sos	South Pacific Ocean
59	E	Ab1PSL-CM6A-LR	585	St4	sos	North Pacific Ocean
60	f	BCC-CSM2-MR	370	Ab2	mlotst_max	Southern Ocean
61	f	GISS-E2-1-H	119, 126, 245, 370, 585	Ab1, Ab2	hfds, siconc, tas, tos	Southern Ocean
62	f	GISS-E2-1-H	245, 370	Ab2	hfds, siconc, tos	Southern Ocean
63	f	HadGEM3-GC31-MM	585	Ab1	siconc	Southern Ocean
64	G	ACCESS-CM2	585	St2	mlotst_max	Southern Ocean
65	G	ACCESS-ESM1-5	126, 245, 370, 585	St2	mlotst_max	Southern Ocean
66	G	CNRM-CM6-1	245, 370, 585	St2	mlotst_max	Ab1ndian Ocean
67	G	NorESM2-LM	370, 585	St2	mlotst_max	Southern Ocean
68	G	NorESM2-MM	585	St2	mlotst_max	Southern Ocean
69	H	CESM2-WACCM	585	St4	msftbarot, zos	Southern Ocean
70	H	Ab1PSL-CM6A-LR	585	St4	zos	Southern Ocean
71	H	MRI-ESM2-0	585	St4	zos	Southern Ocean
72	H	UKESM1-0-LL	585	St4	zos	Southern Ocean
73	Ab1	CanESM5	585	St4	wfo	Southern Ocean

Table A3. List of all 73 SNS and attributes of each SNS. The associated SNS category (cat.) and SSP(s) are listed. The criteria (crit.) and variables (Vars) that are found at least once for each case and associated region are also listed. See text for more information. Also see (Angevaare, 2025b) for each case separately.

Code availability. The framework (Angevaare, 2025a) for this work is available at https://github.com/JoranAngevaare/optim_esm_tools. The methods for identifying the criterion iv (subsubsection A2.4, subsection 3.4) are available at https://github.com/jmecki/AMOCcollapse

Data availability. The CMIP6 data used in this work is freely available online at https://esgf-node.llnl.gov/search/cmip6/. The identified SNS cases from this work are available at (Angevaare, 2025b).

Author contributions. JRA and SSD designed the study, JRA implemented the study, JRA produced the figures, JRA and SSD wrote and edited the paper.

Competing interests. The authors declare no competing interests.

Acknowledgements. We acknowledge the World Climate Research Programs Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modelling groups (listed in Table A2 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. The CMIP6 data was accessed and analysed using the Netherlands super-computer-cluster SURFsara.

SSD and JRA received funding from OptimESM and SSD from TipESM. “TipESM “Exploring Tipping Pints and Their Impacts Using Earth System Models” is funded by the European Union. Grant Agreement number: 101137673.” OptimESM “Optimal High Resolution Earth System Models for Exploring Future Climate Change” is funded by the European Union. Grant Agreement number:101081193. SSD received funding from ISOTIPIC. “ISOTIPIC” Interacting ice Sheet and Ocean Tipping - Indicators, Processes, Impacts and Challenges” is funded by the National Environmental Research Council, UK, under Grant Number: NE/Y503320/1.

References

Abbot, D. S., Huber, M., Bousquet, G., and Walker, C. C.: High-CO2 cloud radiative forcing feedback over both land and ocean in a global climate model, Geophysical Research Letters, 36, 2009.

Angevaare, J. R.: JoranAngevaare/optim_esm_tools: v4.0.0, https://doi.org/10.5281/zenodo.15275185, 2025a.

Angevaare, J. R.: JoranAngevaare/Catalogue-of-Strong-Nonlinear- Surprises-in-ocean-sea-ice-and-atmospheric- variables-in-CMIP6, https://doi.org/10.5281/zenodo.17867115, 2025b.

Angevaare, J. R., Drijfhout, S., , and Notz, D.: Newest climate models project much faster disappearance of year-round sea ice in the Arctic, manuscript in preperation, 2025.

Armstrong McKay, D. I., Staal, A., Abrams, J. F., Winkelmann, R., Sakschewski, B., Loriani, S., Fetzer, I., Cornell, S. E., Rockström, J., and Lenton, T. M.: Exceeding 1.5 C global warming could trigger multiple climate tipping points, Science, 377, eabn7950, 2022.

Bathiany, S., Hidding, J., and Scheffer, M.: Edge detection reveals abrupt and extreme climate events, Journal of Climate, 33, 6399–6421, 2020.

Boers, N.: Observation-based early-warning signals for a collapse of the Atlantic Meridional Overturning Circulation, Nature Climate Change, 11, 680–688, 2021.

Bracegirdle, T. J., Shuckburgh, E., Sallee, J.-B., Wang, Z., Meijers, A. J., Bruneau, N., Phillips, T., and Wilcox, L. J.: Assessment of surface winds over the Atlantic, Indian, and Pacific Ocean sectors of the Southern Ocean in CMIP5 models: Historical bias, forcing response, and state dependence, Journal of Geophysical Research: Atmospheres, 118, 547–562, 2013.

Campello, R. J. G. B., Moulavi, D., and Sander, J.: Density-Based Clustering Based on Hierarchical Density Estimates, in: Advances in Knowledge Discovery and Data Mining, edited by Pei, J., Tseng, V. S., Cao, L., Motoda, H., and Xu, G., pp. 160–172, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-642-37456-2, 2013.

Canny, J.: A computational approach to edge detection, IEEE Transactions on pattern analysis and machine intelligence, pp. 679–698, 1986.

Davy, R. and Griewank, P.: Arctic amplification has already peaked, Environmental Research Letters, 18, 084 003, https://doi.org/10.1088/1748-9326/ace273, 2023.

De Lavergne, C., Palter, J. B., Galbraith, E. D., Bernardello, R., and Marinov, I.: Cessation of deep convection in the open Southern Ocean under anthropogenic climate change, Nature Climate Change, 4, 278–282, 2014.

Ditlevsen, P. and Ditlevsen, S.: Warning of a forthcoming collapse of the Atlantic meridional overturning circulation, Nature Communications, 14, 1–12, 2023.

Drews, A., Greatbatch, R. J., Ding, H., Latif, M., and Park, W.: The use of a flow field correction technique for alleviating the North Atlantic cold bias with application to the Kiel Climate Model, Ocean Dynamics, 65, 1079–1093, 2015.

Drijfhout, S., Bathiany, S., Beaulieu, C., Brovkin, V., Claussen, M., Huntingford, C., Scheffer, M., Sgubin, G., and Swingedouw, D.: Catalogue of abrupt shifts in Intergovernmental Panel on Climate Change climate models, Proceedings of the National Academy of Sciences, 112, E5777–E5786, 2015.

Drijfhout, S., Angevaare, J. R., Mecking, J., van Westen, R. M., and Rahmstorf, S.: Shutdown of northern Atlantic overturning after 2100 following deep mixing collapse in CMIP6 projections, Environmental Research Letters, 20, 094 062, https://doi.org/10.1088/1748-9326/adfa3b, 2025.

Drijfhout, S. S., Bull, C. Y., Hewitt, H., Holland, P. R., Jenkins, A., Mathiot, P., and Garabato, A. N.: An Amundsen Sea source of decadal temperature changes on the Antarctic continental shelf, Ocean Dynamics, 74, 37–52, 2024.

Feldl, N., Po-Chedley, S., Singh, H. K., Hay, S., and Kushner, P. J.: Sea ice and atmospheric circulation shape the high-latitude lapse rate feedback, NPJ climate and atmospheric science, 3, 41, 2020.

Fox-Kemper, B., Hewitt, H. T., Xiao, C., Adalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E., Krinner, G., et al.: Ocean, cryosphere, and sea level change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 1211–1361, Cambridge University Press, 2021.

Geen, R., Bordoni, S., Battisti, D. S., and Hui, K.: Monsoons, ITCZs, and the Concept of the Global Monsoon, Reviews of Geophysics, 58, e2020RG000 700, https://doi.org/https://doi.org/10.1029/2020RG000700, e2020RG000700 10.1029/2020RG000700, 2020.

Gómez-Valdivia, F., Holland, P. R., Siahaan, A., Dutrieux, P., and Young, E.: Projected West Antarctic ocean warming caused by an expansion of the Ross Gyre, Geophysical Research Letters, 50, e2023GL102 978, 2023.

Hankel, C. and Tziperman, E.: The role of atmospheric feedbacks in abrupt winter Arctic sea ice loss in future warming scenarios, Journal of Climate, 34, 4435–4447, 2021.

Hankel, C. and Tziperman, E.: Assessing the Robustness of Arctic Sea Ice Bi-Stability in the Presence of Atmospheric Feedbacks, Journal of Geophysical Research: Atmospheres, 128, e2023JD039 337, 2023.

Hartigan, J. A. and Hartigan, P. M.: The dip test of unimodality, The annals of Statistics, pp. 70–84, 1985.

Heuzé, C.: Antarctic bottom water and North Atlantic deep water in CMIP6 models, Ocean Science Discussions, 2020, 1–38, 2020.

Holland, P. R., Bracegirdle, T. J., Dutrieux, P., Jenkins, A., and Steig, E. J.: West Antarctic ice loss influenced by internal climate variability and anthropogenic forcing, Nature Geoscience, 12, 718–724, 2019.

Hwang, Y.-T. and Frierson, D. M.: Link between the double-Intertropical Convergence Zone problem and cloud biases over the Southern Ocean, Proceedings of the National Academy of Sciences, 110, 4935–4940, 2013.

Jesse, F., Le Bars, D., and Drijfhout, S.: Processes explaining increased ocean dynamic sea level in the North Sea in CMIP6, Environmental Research Letters, 19, 044 060, 2024.

Kay, J. E., Holland, M. M., Bitz, C. M., Blanchard-Wrigglesworth, E., Gettelman, A., Conley, A., and Bailey, D.: The influence of local feedbacks and northward heat transport on the equilibrium Arctic climate response to increased greenhouse gas forcing, Journal of Climate, 25, 5433–5450, 2012.

Kopp, R. E., Gilmore, E. A., Shwom, R. L., Adams, H., Adler, C., Oppenheimer, M., Patwardhan, A., Russill, C., Schmidt, D. N., and York, R.: ‘Tipping points’ confuse and can distract from urgent climate action, Nature Climate Change, 15, 29–36, 2025.

Lee, J.-Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J., Engelbrecht, F., Fischer, E., Fyfe, J., Jones, C., Maycock, A., Mutemi, J., Niaye, O., Panickal, S., Zhou, T., and Christensen, H.: Future global climate: scenario-based projections and near-term information, pp. 553–672, IPCC Assessment Reports, Cambridge University Press, 2021.

Lenderink, G. and Haarsma, R. J.: Modeling Convective Transitions in the Presence of Sea lce, Journal of Physical Oceanography, 26, 1448 – 1467, https://doi.org/10.1175/1520-0485(1996)026<1448:MCTITP>2.0.CO;2, 1996.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth’s climate system, Proceedings of the national Academy of Sciences, 105, 1786–1793, 2008.

Loriani, S., Aksenov, Y., Armstrong McKay, D., Bala, G., Born, A., Chiessi, C. M., Dijkstra, H., Donges, J. F., Drijfhout, S., England, M. H., et al.: Tipping points in ocean and atmosphere circulations, EGUsphere, 2023, 1–62, 2023.

Mchedlishvili, A., Spreen, G., Melsheimer, C., and Huntemann, M.: Weddell Sea polynya analysis using SMOS–SMAP apparent sea ice thickness retrieval, The Cryosphere, 16, 471–487, 2022.

McInnes, L. and Healy, J.: Accelerated hierarchical density based clustering, in: 2017 IEEE international conference on data mining workshops (ICDMW), pp. 33–42, IEEE, 2017.

Mecking, J. V. and Drijfhout, S. S.: The decrease in ocean heat transport in response to global warming, Nature Climate Change, 13, 1229–1236, 2023.

Meredith, M. P., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A. A., Hollowed, A. B., Kofinas, G., Mackintosh, A. N., Muelbert, M. M. C., Melbourne-Thomas, J., et al.: Polar regions, in: The ocean and cryosphere in a changing climate: Summary for policymakers, pp. 3–1, Intergovernmental Panel on Climate Change, 2019.

Moreno-Chamarro, E., Caron, L.-P., Loosveldt Tomas, S., Gutjahr, O., Moine, M.-P., Putrasahan, D., Roberts, C. D., Roberts, M. J., Senan, R., Terray, L., et al.: Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models, Geoscientific Model Development Discussions, 2021, 1–33, 2021.

Morioka, Y., Manabe, S., Zhang, L., Delworth, T. L., Cooke, W., Nonaka, M., and Behera, S. K.: Antarctic sea ice multidecadal variability triggered by Southern Annular Mode and deep convection, Communications Earth & Environment, 5, 633, 2024.

Munday, D. R., Johnson, H. L., and Marshall, D. P.: Eddy saturation of equilibrated circumpolar currents, Journal of Physical Oceanography, 43, 507–532, 2013.

Naughten, K. A., Holland, P. R., Dutrieux, P., Kimura, S., Bett, D. T., and Jenkins, A.: Simulated twentieth-century ocean warming in the Amundsen Sea, West Antarctica, Geophysical Research Letters, 49, e2021GL094 566, 2022.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., et al.: The scenario model intercomparison project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, 2016.

Portmann, V.: Observational constraints applied to the AMOC projections indicates a slowdown of half before 2100, Science Advances, under review, 2025.

Rheinlænder, J. W., Smedsrud, L. H., and Nisanciouglu, K. H.: Internal ocean dynamics control the long-term evolution of Weddell Sea Polynya activity, Frontiers in Climate, 3, 718 016, 2021.

Robson, J., Menary, M. B., Sutton, R. T., Mecking, J., Gregory, J. M., Jones, C., Sinha, B., Stevens, D. P., and Wilcox, L. J.: The Role of Anthropogenic Aerosol Forcing in the 1850–1985 Strengthening of the AMOC in CMIP6 Historical Simulations, Journal of Climate, 35, 6843 – 6863, https://doi.org/10.1175/JCLI-D-22-0124.1, 2022.

Romanou, A., Rind, D., Jonas, J., Miller, R., Kelley, M., Russell, G., Orbe, C., Nazarenko, L., Latto, R., and Schmidt, G. A.: Stochastic Bifurcation of the North Atlantic Circulation under a Midrange Future Climate Scenario with the NASA-GISS ModelE, Journal of Climate, 36, 6141 – 6161, https://doi.org/10.1175/JCLI-D-22-0536.1, 2023.

Rose, B. E. and Marshall, J.: Ocean heat transport, sea ice, and multiple climate states: Insights from energy balance models, Journal of the atmospheric sciences, 66, 2828–2843, 2009.

Schulzweida, U.: CDO User Guide, https://doi.org/10.5281/zenodo.10020800, 2023.

Sgubin, G., Swingedouw, D., Drijfhout, S., Mary, Y., and Bennabi, A.: Abrupt cooling over the North Atlantic in modern climate models, Nature Communications, 8, 14 375, 2017.

Stewart, A. L., Neumann, N. K., and Solodoch, A.: "Eddy" Saturation of the Antarctic Circumpolar Current by Standing Waves, Journal of Physical Oceanography, 53, 1161 – 1181, https://doi.org/10.1175/JPO-D-22-0154.1, 2023.

Swingedouw, D., Bily, A., Esquerdo, C., Borchert, L. F., Sgubin, G., Mignot, J., and Menary, M.: On the risk of abrupt changes in the North Atlantic subpolar gyre in CMIP6 models, Annals of the New York Academy of Sciences, 1504, 187–201, 2021.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American meteorological Society, 93, 485–498, 2012.

Terpstra, S., Falkena, S. K., Bastiaansen, R., Bathiany, S., Dijkstra, H. A., and von der Heydt, A. S.: Assessment of Abrupt Shifts in CMIP6 Models using Edge Detection, arXiv preprint arXiv:2410.19498, 2024.

Thompson, A. F., Stewart, A. L., Spence, P., and Heywood, K. J.: The Antarctic Slope Current in a changing climate, Reviews of Geophysics, 56, 741–770, 2018.

Tian, B. and Dong, X.: The Double-ITCZ Bias in CMIP3, CMIP5, and CMIP6 Models Based on Annual Mean Precipitation, Geophysical Research Letters, 47, e2020GL087 232, https://doi.org/https://doi.org/10.1029/2020GL087232, e2020GL087232 2020GL087232, 2020.

Våge, K., Papritz, L., Håvik, L., Spall, M., and Moore, G.: Ocean convection linked to the recent ice edge retreat along east Greenland. Nat. Commun., 9, 1287, 2018.

van Westen, R. M., Vanderborght, E. Y., Kliphuis, M., and Dijkstra, H. A.: Substantial Risk of 21st Century AMOC Tipping even under Moderate Climate Change, arXiv preprint arXiv:2407.19909, 2024.

Wunderling, N., von der Heydt, A. S., Aksenov, Y., Barker, S., Bastiaansen, R., Brovkin, V., Brunetti, M., Couplet, V., Kleinen, T., Lear, C. H., Lohmann, J., Roman-Cuesta, R. M., Sinet, S., Swingedouw, D., Winkelmann, R., Anand, P., Barichivich, J., Bathiany, S., Baudena, M., Bruun, J. T., Chiessi, C. M., Coxall, H. K., Docquier, D., Donges, J. F., Falkena, S. K. J., Klose, A. K., Obura, D., Rocha, J., Rynders, S., Steinert, N. J., and Willeit, M.: Climate tipping point interactions and cascades: a review, Earth System Dynamics, 15, 41–74, https://doi.org/10.5194/esd-15-41-2024, 2024.

Xin, X., Wu, T., Zhang, J., Yao, J., and Fang, Y.: Comparison of CMIP6 and CMIP5 simulations of precipitation in China and the East Asian summer monsoon, International Journal of Climatology, 40, 6423–6440, https://doi.org/https://doi.org/10.1002/joc.6590, 2020.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of higher climate sensitivity in CMIP6 models, Geophysical Research Letters, 47, e2019GL085 782, 2020.

Zhang, G. J., Song, X., and Wang, Y.: The double ITCZ syndrome in GCMs: A coupled feedback problem among convection, clouds, atmospheric and ocean circulations, Atmospheric Research, 229, 255–268, 2019.