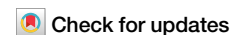**Article**

Exploring multiyear-to-decadal North Atlantic sea level predictability and prediction using machine learning

Check for updates

Qinxue Gu [1] ✉, Liping Zhang [2,3], Liwei Jia[2], Thomas L. Delworth [2], Xiaosong Yang [2], Fanrong Zeng[2], William F. Cooke [2] & Shouwei Li[1]

Coastal communities face substantial risks from long-term sea level rise and decadal sea level variations, with the North Atlantic and U.S. East Coast being particularly vulnerable under changing climates. Employing a self-organizing map-based framework, we assess the North Atlantic sea level variability and predictability using 5000-year sea level anomalies (SLA) from two preindustrial control model simulations. Preferred transitions among patterns of variability are identified, revealing long-term predictability on decadal timescales related to shifts in Atlantic meridional overturning circulation phases. Combining this framework with model-analog techniques, we demonstrate prediction skill of large-scale SLA patterns and low-frequency coastal SLA variations comparable to that from initialized hindcasts. Moreover, additional short-term predictability is identified after the exclusion of low-frequency signals, which arises from slow gyre circulation adjustment triggered by the North Atlantic Oscillation-like stochastic variability. This study highlights the potential of machine learning to assess sources of predictability and to enable long-term climate prediction.

Sea level rise (SLR) poses significant risks to coastal communities due to its impact on flooding, erosion, and seawater intrusion in estuaries and aquifers[1–7]. Therefore, accurate predictions of coastal sea levels are crucial because they enable communities to prepare more effectively for extreme sea level events and the associated risk of coastal flooding. Reliable sea level predictions hinge on a thorough understanding of sea level variability and its predictability.

Coastal sea level varies on a broad range of timescales from hourly to centennial, in addition to the global SLR associated with thermal expansion of seawater and melting of glaciers and ice sheets[8,9]. Sea level changes on hourly to daily timescales can primarily be attributed to storm surges, astronomical tides, and waves[10,11]. On seasonal-to-multidecadal timescales, variations in sea level are often linked to climate variability (e.g., El Niño-Southern Oscillation [ENSO][12–14]) and large-scale ocean dynamics (e.g., Atlantic meridional overturning circulation [AMOC][15,16]). Multiyear-to-decadal sea level variations establish a background state that modulates the impacts of high-frequency storm surges and tides[17,18].

The U.S. East Coast has been shown as a hotspot of accelerated SLR under present and future climates[15,16,19–23]. Specifically, the sterodynamic sea level in the North Atlantic is projected to rise rapidly in a warmer climate, especially in the northwestern Atlantic including the U.S. Coast north of Cape Hatteras[19,20]. This SLR is primarily attributed to steric effect and mass redistribution, both closely related to a weakened AMOC[19,20]. The resulting sharp gradient of sea level change across the Gulf Stream and North Atlantic current is balanced by a deceleration of these currents. This Gulf Stream and AMOC-related mechanism is a prevailing argument to explain the acceleration of SLR along the U.S. Northeast Coast over the past several decades, together with other processes such as alongshore wind forcing and the inverse barometer effect[24–30].

In addition to the SLR along the U.S. Northeast Coast, recent studies have identified an accelerated SLR since 2010 along the U.S. Southeast Coast region, south of Cape Hatteras, and the Gulf of Mexico[16,21–23,31–33]. This acceleration is not seen to the north of Cape Hatteras, indicating that mechanisms differing from these discussed above should be responsible. In addition to the anthropogenic external forcing, multiple processes are demonstrated as potential mechanisms driving this rapid SLR, yet without a consensus being reached. These proposed mechanisms include variations in the Gulf Stream strength and position[33], warming of the Florida Current[31,33], large-scale heat divergence linked to AMOC and low-frequency North Atlantic Oscillation (NAO)[23,32], a lagged response to the 2009–2010 AMOC slowdown[16], cumulative effect of NAO and ENSO[21], and wind-forced Rossby waves in the tropical North Atlantic with additional contributions from river discharge and coastal winds[22].

[1]The Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA. [2]NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA. [3]University Corporation for Atmospheric Research, Boulder, CO, USA. ✉e-mail: qg4783@princeton.edu

Due to the limited spatial coverage of tide gauges and the short duration of satellite observations, capturing a full range of observed sea level variability in the North Atlantic on decadal timescales is challenging. This limitation also complicates our understanding of the underlying mechanisms and potential decadal predictability of sea level changes. To better understand the multiyear-to-decadal sea level variability and predictability in the North Atlantic, we leverage an unsupervised machine learning classification method named the self-organizing map (SOM)[34] and apply it to long climate model simulations. The SOM is an effective tool for classifying high-dimensional datasets into clusters[34] and identifying physically relevant patterns of variability[35]. It has been widely applied across a range of studies in climate science, including characterizing atmospheric circulation[36–40], probing teleconnection patterns[35,41,42], investigating climate variability[43–47], and conducting climate predictions[45].

In this study, we classify 5000-year annual mean sea level anomalies (SLA) in the North Atlantic from two preindustrial control (piControl) simulations from the **S**eamless System for **P**rediction and **EA**rth System **R**esearch (SPEAR)[48] coupled global climate model using the SOM. Applying the SOM method to two long piControl simulations enables an investigation of a full range of internal sea level variability in the North Atlantic, revealing patterns that resemble the observational dynamic sea level patterns from recent decades. Additionally, we assess the multiyear-to-decadal predictability of the sea level by tracing the temporal evolution of the classified patterns across the SOM. Decadal predictions of large-scale North Atlantic sea level patterns and low-frequency coastal sea level variations are also conducted using a model-analog method[49–52] based on the SOM framework, leading to comparable skill to that from initialized decadal hindcasts.

Results

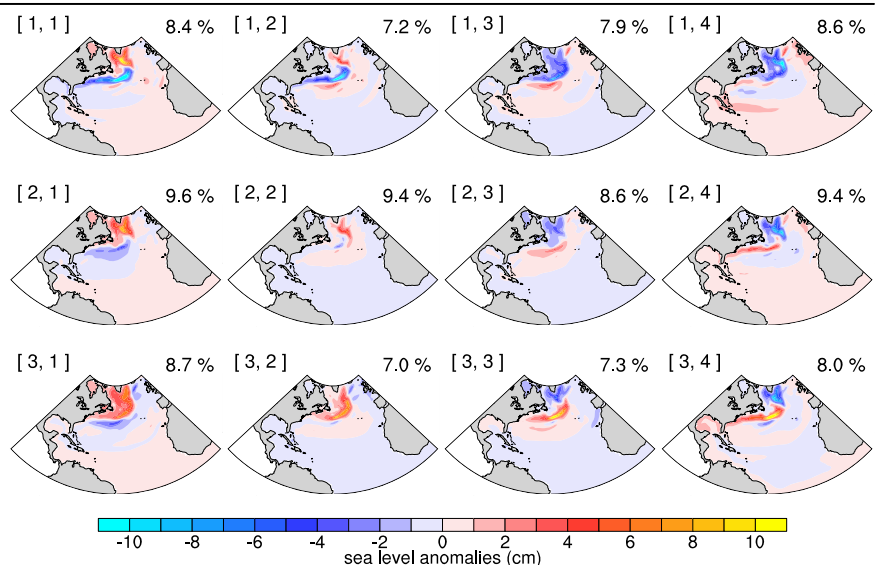### Multiyear-to-decadal North Atlantic sea level variability in SPEAR piControl simulations

To examine the sea level variability and predictability in the North Atlantic Ocean, we first utilize the SOM to classify the 5000-year annual mean North Atlantic SLA from the piControl simulations from two versions of the SPEAR model. This classification is conducted by repeatedly training the SOM to learn the SLA patterns from the simulations (see Methods for details). These two versions, SPEAR_MED and SPEAR_LO, have the same oceanic resolution but different atmospheric resolution: SPEAR_LO with a 1° horizontal resolution and SPEAR_MED with a 0.5° horizontal resolution. Using these simulations enriches the data for training the SOM and also allows us to assess the influence of atmospheric horizontal resolution on SLA variability and predictability. Classification using the SOM method results in a predefined number of generalized patterns (a 3 × 4 array in this
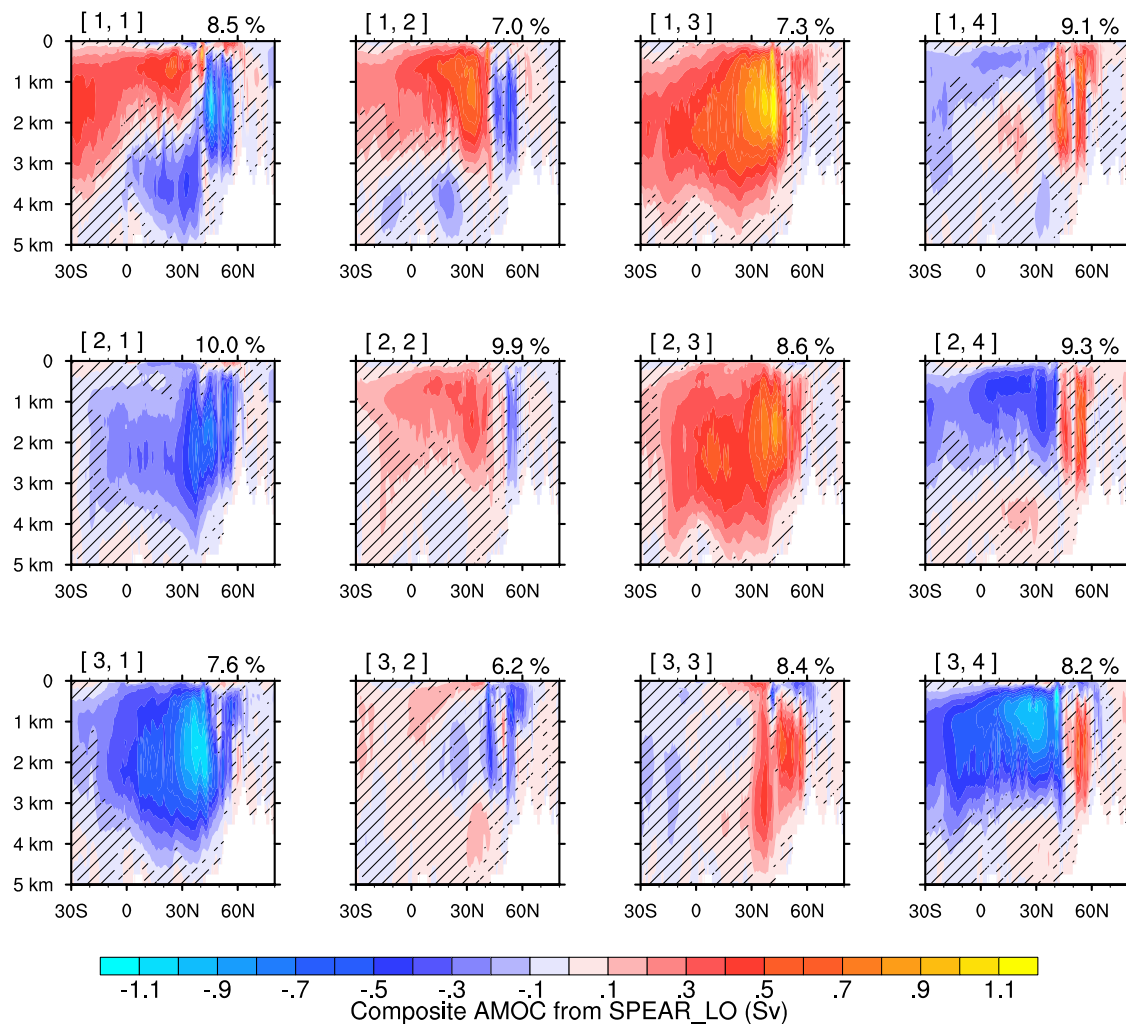
instance), referred to as SOM nodes (Fig. 1). Each node represents a dominant pattern of sea level variability in the North Atlantic, with the frequency of occurrence of each pattern denoted in the top right corner of the corresponding subplot. These dominant patterns of variability are robust to different SOM sizes (see Methods for details). As expected from a well-trained SOM[34], similar patterns appear in nearby nodes, while different patterns are found in nodes that are farther apart. The composite SLA patterns associated with each SOM pattern from SPEAR_LO (Supplementary Fig. 1) and SPEAR_MED (Supplementary Fig. 2) piControl simulations are similar to the corresponding SOM pattern, indicating that 1) the dominant patterns from both simulations are similar, despite the different resolutions of their atmospheric components; and 2) the SOM training process is robust, effectively capturing the patterns of SLA variability in both simulations.

Among the identified SOM patterns, several delineate pronounced SLAs along the U.S. East Coast. Notably, node [3,1] characterizes robust positive anomalies in the western subpolar North Atlantic and the Labrador Sea, extending to the U.S. Northeast Coast, north of Cape Hatteras. South of these pronounced positive anomalies, a zone of moderate negative anomalies is seen around 35°N near Cape Hatteras with limited meridional expansion. This pattern resembles the changes in dynamic sea level (deviation from the global mean sea level) projected for future climates due to a weakened AMOC under global warming[19,20]. Additionally, it aligns with the most predictable component of North Atlantic SLA identified using average predictability time by Zhang et al.[15]. Conversely, node [1,3] characterizes a pattern largely opposite to [3,1], with lower-than-average sea levels along the U.S. Northeast Coast. The composite Atlantic meridional overturning streamfunction patterns associated with the SLA in nodes [3,1] and [1,3] reveal a mature negative and positive AMOC phase, respectively (Fig. 2 for the SPEAR_LO piControl simulation and Supplementary Fig. 3 for the SPEAR_MED piControl simulation; see Methods for details). These AMOC states are consistent with the underlying physical mechanisms driving their respective SLA patterns, encompassing mass redistribution, density variations, and circulation dynamics[19,20,24,25].

In addition to the nodes exhibiting pronounced SLAs along the U.S. Northeast Coast, nodes [1,1] and [3,4] display distinct loadings along the U.S. Southeast Coast. Node [1,1] is characterized by a dipole-like pattern, with positive anomalies in the western subpolar North Atlantic and Labrador Sea, situated north of 50°N, and negative anomalies near the Gulf Stream path. These negative anomalies extend to the U.S. Southeast Coast. Conversely, node [3,4] presents an opposite pattern, with positive anomalies along the U.S. Southeast Coast. These two patterns bear resemblance to the second most predictable component of North Atlantic SLA, as revealed by

Fig. 1 | Generalized 3 × 4 SOM for the 5000-year annual mean SLA (cm; shading) in the North Atlantic from the SPEAR_LO and SPEAR_MED piControl simulations. The top left corner of each subplot denotes a two-index notation with the first (second) index representing the row (column) of the SOM node. The top right corner of each subplot displays the overall frequency of occurrence of the corresponding node in both simulations.

Fig. 2 | Composites of annual mean Atlantic meridional overturning stream-function patterns (Sv; shading) from the SPEAR_LO piControl simulation for all SOM nodes (with the SOM shown in Fig. 1). The absence of black hatching indicates that the shaded value is significant at the 95% confidence level using a two-tailed *t* test. The top left corner of each subplot denotes a two-index notation with the first (second) index representing the row (column) of the node. The top right corner of each subplot displays the frequency of occurrence of the corresponding node in the SPEAR_LO piControl simulation, which slightly differs from the overall frequency shown in Fig. 1.
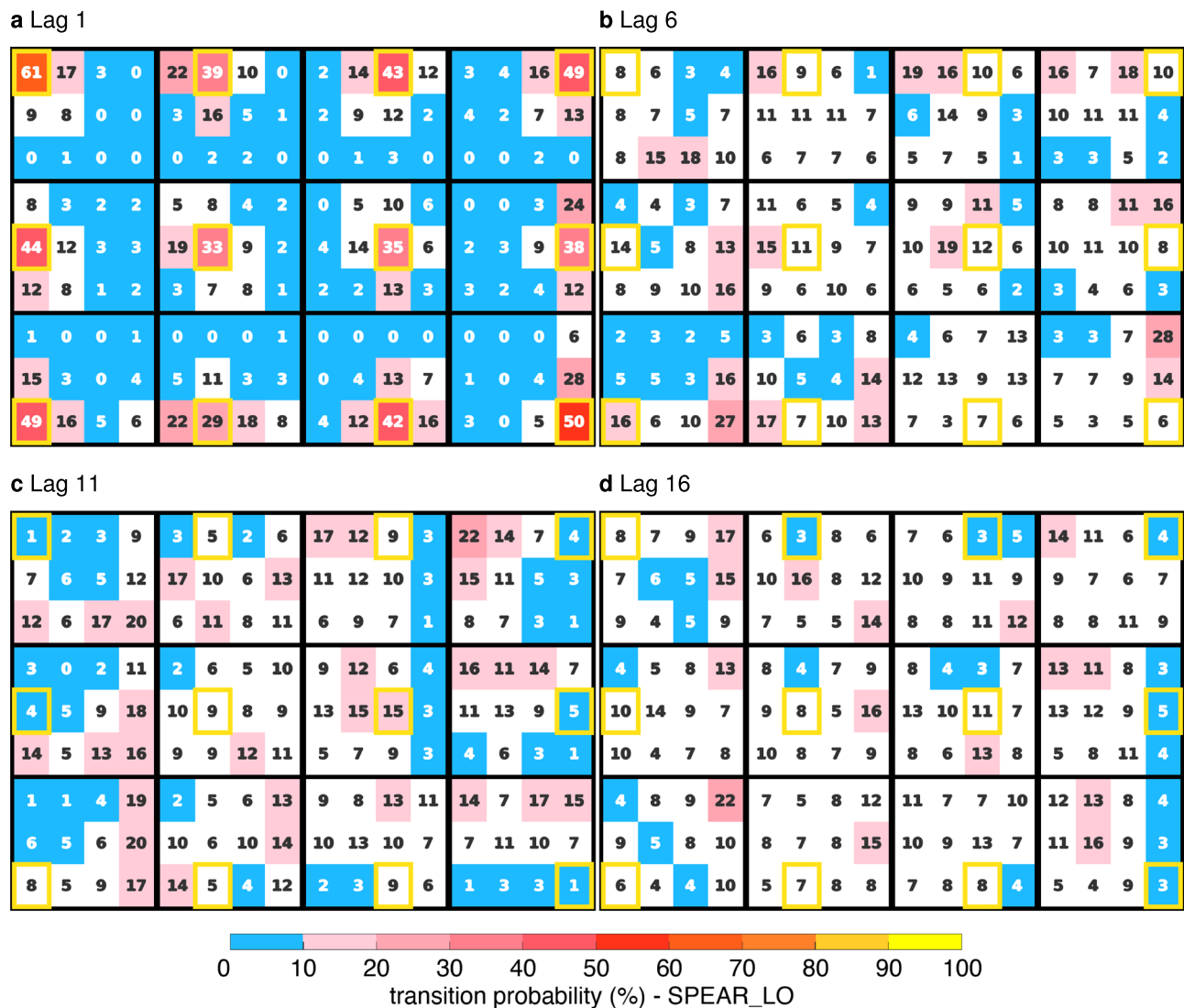
Zhang et al.[15]. The underlying mechanisms, as detailed in their work, involve heat convergence/divergence due to AMOC transition between its mature positive and negative phases[15,53,54]. Consistently, the composites of AMOC patterns associated with these two nodes show transition states with different signs of anomalies in the north and south of 40°N (Fig. 2 and Supplementary Fig. 3).

The SOM is designed to span the data space, thereby depicting different phases of North Atlantic SLA variability. Hence, the remaining nodes can be interpreted as transition states between the aforementioned four nodes or as different "flavors" of these nodes, each displaying a similar frequency of occurrence to these four nodes but with relatively minor loadings along the coastal regions of the United States. For example, node [2,3] represents an SLA pattern similar to that of [1,3], albeit with weaker anomalies (Fig. 1), consistent with the comparison between their AMOC states (Fig. 2). Nodes [3,3] and [2,4] exhibit a dipole pattern similar to that of node [3,4], but with one pole exhibiting less pronounced anomalies. A similar relationship exists between [1,2], [2,1], and [1,1]. While it is common for most nodes to have a counterpart with a similar pattern but an opposite phase across the SOM (e.g., [1,3] versus [3,1]), there are exceptions. For instance, the extended band of modest positive SLA located south of Florida in node [1,4] does not have a corresponding counterpart in other nodes, even when using a larger SOM that displays

more phases of SLA variability (Supplementary Fig. 4). This discrepancy may be attributed to certain nonlinear processes.

Having identified the variability of SLA in the North Atlantic, we proceed to examine their multiyear-to-decadal predictability by employing a SOM-based lagged transition probability framework[45]. Specifically, we calculate the probability of each node transitioning to any of the 12 nodes, including itself, after 1–30 years (hereafter denoted as lags 1–30 years; see Methods for details). This approach enables us to trace the temporal evolution of SLAs across the SOM nodes.

The resulting transition probability tables are similar between the SPEAR_LO (Fig. 3 for lags 1, 6, 11, 16 years and Supplementary Movie 1 for lags 1–30 years) and SPEAR_MED piControl simulations (Supplementary Movie 2). Within the first 3 years, most nodes show strong persistence, as evidenced by the highest transition probability remaining within the starting node itself (e.g., Fig. 3a). In addition to persistence, most nodes along the edge of the SOM display distinct anticlockwise tendencies within the first 16 years (Fig. 3). That is, nodes tend to transition anticlockwise to adjacent nodes, rather than clockwise. This tendency is more clearly depicted in Fig. 4 for node [3,1] over lags 1–24, and in Supplementary Movies 1, 2 for all nodes over lags 1–30, where the reddish shadings gradually spread out from the starting nodes (denoted by yellow rectangles) and transition in an anticlockwise direction across the SOM. For some nodes (e.g., node [3,1]), this

Fig. 3 | Transition probability tables used to track the temporal evolution of SLA from the SPEAR_LO piControl simulation across the SOM. The small shaded cells display the conditional probability of their occurrence (%) at **a** lag 1, **b** lag 6, **c** lag 11, and **d** lag 16 years, given the occurrence of the larger cell at lag 0. The positions of these cells correspond to the positions of the SOM 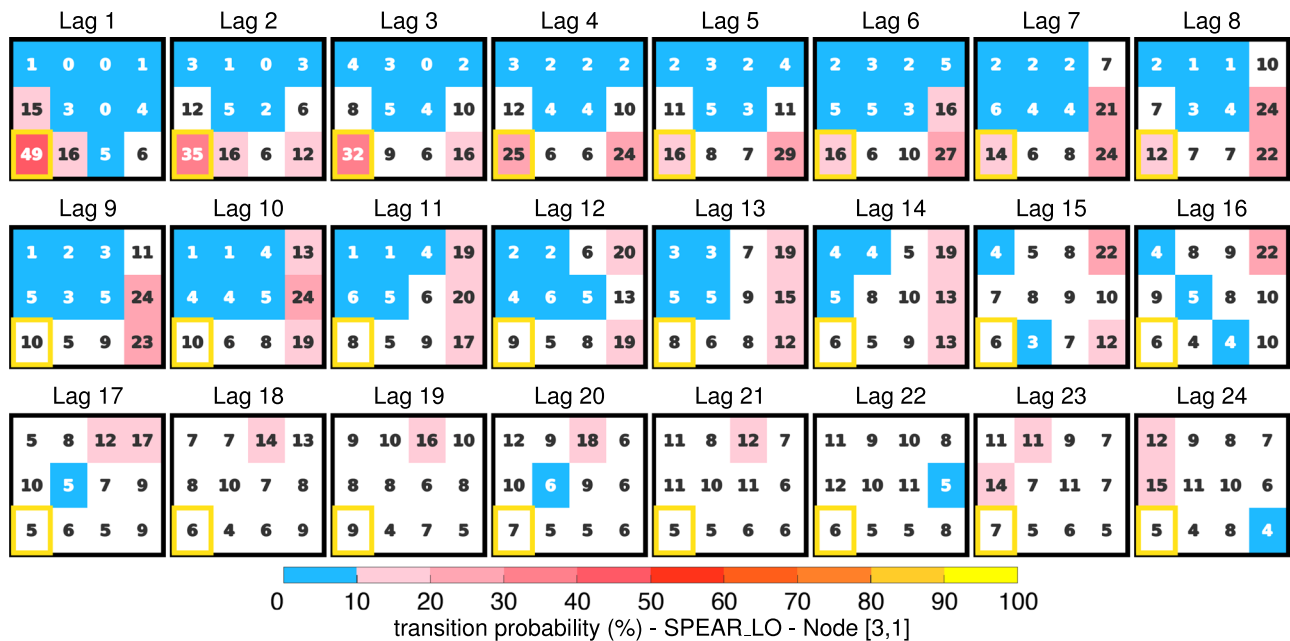nodes in Fig. 1, with the yellow rectangle denoting the starting node (the state at lag 0). Cells shaded with warm (cold) colors represent significantly larger (smaller) values than the unconditional probabilities according to a Monte Carlo significance test, while white cells represent values not significantly different from the unconditional probabilities (see Methods for details).

anticlockwise tendency persists throughout the entire 30-year lag period. The persistence and directional transition of SLA patterns are attributable to the low-frequency buoyancy-driven AMOC variability, as evidenced by their corresponding relationship (Figs. 1 and 2).
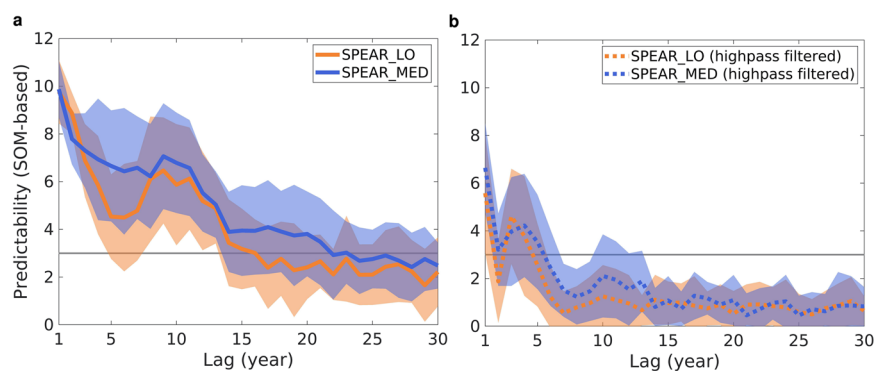
The tendency for SLA patterns to remain in or transition to certain nodes, as signified by higher-than-normal probabilities (reddish shadings in Figs. 3 and 4), or to avoid certain nodes, as marked by lower-than-normal probabilities (bluish shadings in Figs. 3 and 4), serves as an indicator of elevated predictability relative to climatological frequencies of occurrence. Therefore, the aforementioned behavior—the tendency of persistence and preferential transitioning to adjacent nodes along a specific direction (anticlockwise here)—denotes high predictability of North Atlantic SLA on decadal timescales for the first 15–20 years. Over time, the probability of all nodes becomes more evenly distributed (Supplementary Movies 1 and 2), converging towards their climatological frequencies of occurrence, which indicates a loss of predictability. Following this logic, we can therefore quantify the predictability of SLA over time based on the number of transition probabilities that are significantly larger or smaller than normal within the transition probability tables starting from each node over lags[45]

(see Methods for details). As shown in Fig. 5a, the large-scale North Atlantic SLA evolutions are considered predictable for about 16 and 21 years, on average, in SPEAR_LO and SPEAR_MED piControl simulations, respectively. We speculate that the long theoretical predictability identified here arises from the persistence of low-frequency AMOC states and their directional shifts, along with the close association of the SLA patterns with various AMOC phases, as discussed above. Spectral analysis reveals that the AMOC indices in both models have peaks at a period of 25–40 years, with the amplitude in SPEAR_MED exceeding that in SPEAR_LO (Supplementary Fig. 5). This difference may explain the higher SLA predictability in SPEAR_MED compared to SPEAR_LO.

Based on previous discussions, low-frequency signals dominated by buoyancy-driven AMOC variability are identified as critical to North Atlantic sea level variability and predictability in the piControl simulations. To further substantiate their role from an additional perspective, we generate a SOM using 15-year low-pass filtered SLA patterns from the piControl simulations to characterize the low-frequency sea level variability. The resulting patterns (Supplementary Fig. 6) closely resemble the unfiltered SOM (Fig. 1), with slight

Fig. 4 | Transition probability tables used to track the temporal evolution of SLA starting from node [3,1]. Same as Fig. 3 but for node [3,1] across lags 1–24.

Fig. 5 | Predictability measured in the SOM-based framework. a The number of significant transition probabilities averaged in all nodes (large cells) of the transition probability tables from lags 1–30 years for SPEAR_LO (orange line) and SPEAR_MED (blue line) piControl simulations, with their corresponding transition probability tables shown in Supplementary Movies 1 and 2, respectively. The corresponding shadings are calculated as the average number of significant transition probabilities of all nodes ± one standard deviation across all nodes. The mean and standard deviations are weighted with the frequency of occurrence of each node.
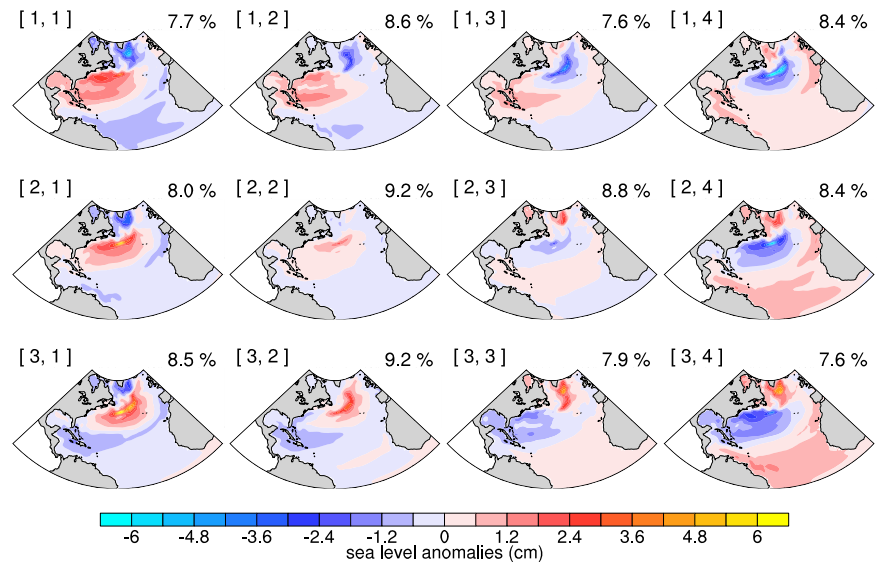
A greater number of significant probabilities indicates higher predictability, with 12 (the total number of SOM nodes) being perfectly predictable. If the number of significant transition probabilities reaches or exceeds the threshold marked by the gray line at 3, the associated SLA state is viewed as theoretically predictable in this study (see Methods for details). **b** Same as **a** but for the transition probability tables associated with the SOM generated from 15-year high-pass filtered North Atlantic SLA patterns (Supplementary Movies 3 and 4).

differences in magnitude and detail. This similarity can be attributed to the greater total sea level variability explained by low-frequency variability compared to high-frequency variability. Specifically, as shown in Supplementary Fig. 7a, the total sea level variability is much higher in the subpolar and subtropical North Atlantic compared to the tropics, with the maximum variability occurring along the Gulf Stream and North Atlantic Current. The variability explained by the 15-year high-pass filtered SLA is primarily concentrated in the tropical region (Supplementary Fig. 7b), where the total variability is minimal. In contrast, the low-pass filtered patterns account for a much larger share of the sea level variability in mid-to-high latitude regions (Supplementary Fig. 7c), where the total sea level variability across all timescales is also greater. This additional evidence further underscores the importance of low-frequency signals, dominated by the buoyancy-driven AMOC, in influencing North Atlantic internal sea level variability.

North Atlantic sea level variability and predictability on shorter timescales

The low-frequency buoyancy-driven AMOC variability has been identified as the dominant source of internal decadal predictability for the North Atlantic sea level in fully coupled control simulations, as discussed above and in previous studies[15,23]. However, the limited length of observational records poses challenges in reliably isolating the low-frequency AMOC signals to inform decadal SLA predictability. Therefore, our investigation extends to exploring SLA variability and predictability on shorter timescales. This is achieved by applying a 15-year high-pass Butterworth filter to the annual mean SLA, filtering out longer-timescale signals dominated by the low-frequency buoyancy-driven AMOC, followed by a SOM analysis that produces a 3 × 4 SOM (Fig. 6; see Methods for details on the choice of the SOM size). Notably, the 15-year high-pass filter does not eliminate wind-driven AMOC signals on shorter timescales, which typically range from seasonal to interannual[55,56].

Fig. 6 | Generalized 3 × 4 SOM for the 15-year high-pass filtered annual mean SLA (cm; shading) in the North Atlantic from the SPEAR_LO and SPEAR_MED piControl simulations. The top left corner of each subplot denotes a two-index notation with the first (second) index representing the row (column) of the SOM node. The top right corner of each subplot displays the overall frequency of occurrence of the corresponding node in both simulations.

In general, the SOM patterns derived from the high-pass filtered SLA are characterized by anomalies of relatively uniform magnitude that display tripole-like features (Fig. 6). This finding contrasts with the SOM patterns associated with unfiltered SLA variability (Fig. 1), which feature strong bands of anomalies with limited meridional extent, tied to the low-frequency buoyancy-driven AMOC in various phases as discussed above. In Fig. 6, node [1,1] represents a tripole pattern, with negative SLA in the subpolar and tropical North Atlantic and positive SLA in the subtropical region, extending towards the U.S. Southeast Coast. This pattern resembles the sea surface temperature (SST) tripole typically associated with positive NAO-related atmospheric forcing and the consequent wind-driven ocean circulation[46,57]. Furthermore, it aligns with the SLA pattern induced by NAO-like wind stress as evidenced by Zhang et al.[23]. Conversely, node [3,4] presents an opposite pattern to that of [1,1], leading to lower-than-normal sea levels in the U.S. Southeast Coast region. The rest of the patterns identified in this SOM exhibit lower loadings in coastal areas.

Next, we explore the temporal evolution of the high-pass filtered SLA across this SOM and infer the predictability of all nodes based on the transition probability (Fig. 7 for lags 1, 3, 5, 7 years from the SPEAR_LO piControl simulation, and Supplementary Movies 3 and 4 for lags 1–30 years from the SPEAR_LO and SPEAR_MED simulations, respectively). Consistent with findings from the transition probability tables of the SOM with unfiltered SLA, high-pass filtered patterns in both simulations exhibit similar temporal evolution. At lag 1 year (1 year after the initial state), the SLA pattern tends to stay in the starting node or transition to adjacent nodes (Fig. 7a). The close alignment between the SLA patterns (Fig. 6) and their corresponding SST composites (Supplementary Figs. 8 and 9) indicates that the SLA variability is dominated by its thermosteric component. Therefore, the 1-year persistence of SLA may be attributed to the persistence of SST, possibly due to the ocean memory[58,59], the transient-eddy feedback mechanism[46,47,60–64], or both.

Unlike the prolonged persistence and directional transition as shown in the transition probability tables for the SOM with unfiltered SLA (Fig. 3), the probabilities for most high-pass filtered SLA patterns become relatively evenly distributed at lag 2 year (Supplementary Movies 3 and 4). Consequently, tracing the evolution of SLA patterns across the SOM becomes challenging, signaling a decrease in predictability. However, predictability reemerges at lags 3–4 years for most nodes, owing to an increased probability of transitions to nodes that are distant from the starting node (Fig. 7b). This reemergence might be attributed to longer timescale processes associated with slow gyre circulation adjustment[23]. From lag 6 year onwards, the return to evenly distributed transition probabilities that mirror the climatological frequencies of o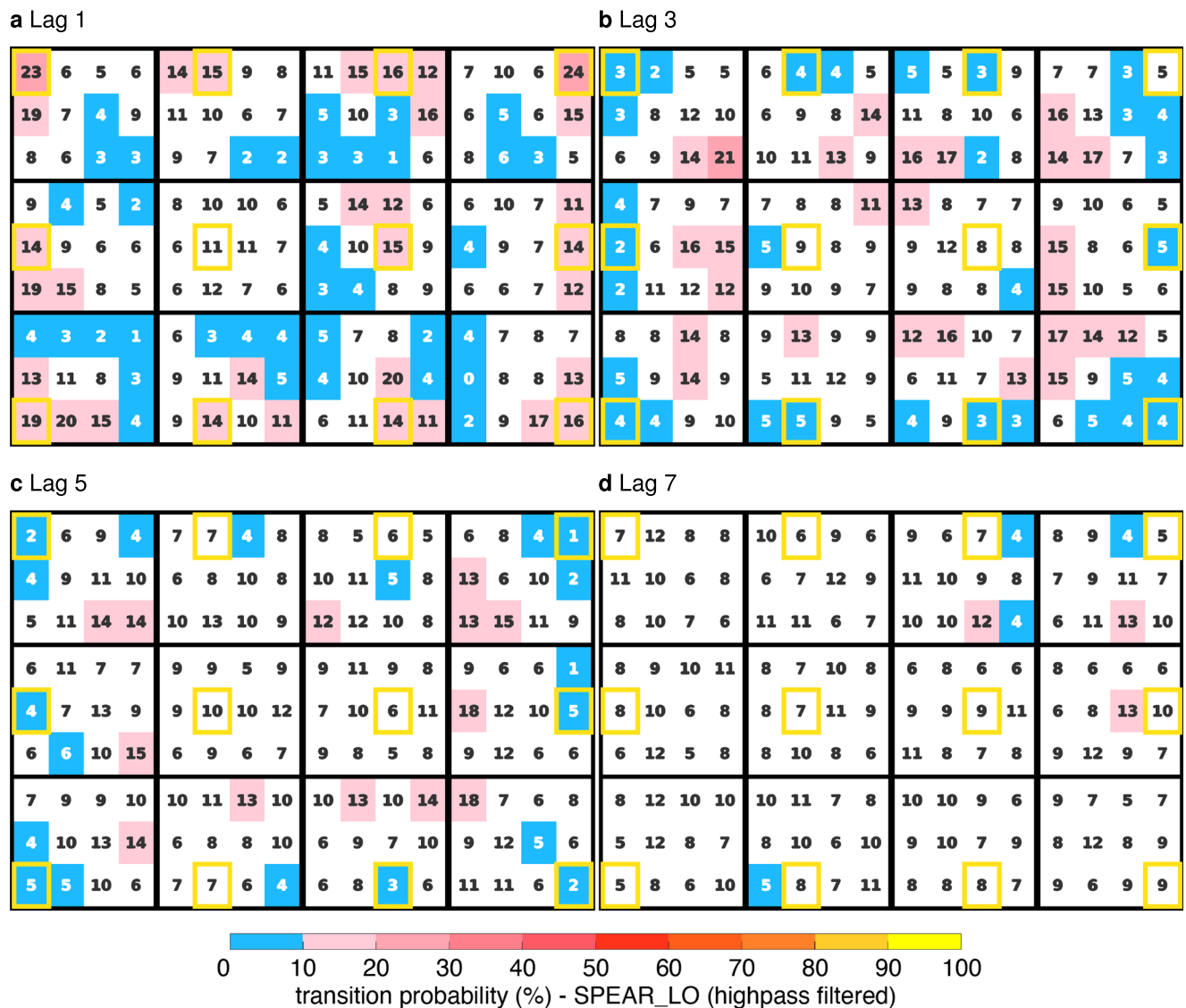ccurrence once again signifies a loss of predictability. The evolution of the SOM-based predictability for the high-pass filtered SLA patterns over 30 years is depicted in Fig. 5b. The overall predictability lasts for 4–5 years for both simulations, much shorter than that of the unfiltered SLA patterns, due to the exclusion of longer timescale signals.

Observational multiyear-to-decadal SLA variability

To examine whether the patterns characterized by the SOM from high-pass filtered piControl simulations using SPEAR (Fig. 6) also exist in the observed record, we next classify the dynamic sea level (deviation from the global mean sea level) anomalies in satellite observations over the 1993–2021 period against the existing SOM nodes by identifying their most similar SOM node (see Methods for details). There are two nodes, [1,4] and [2,2], that do not have satellite observations assigned, potentially due to the limited duration of the satellite record (Fig. 8). Composites of the observational SLA patterns associated with the SOM nodes show similar large-scale features to their corresponding SLA patterns from piControl simulations, albeit with larger magnitude and more small-scale structures (Fig. 8 versus Fig. 6). Particularly, the enlarged anomalies along the Gulf Stream and North Atlantic Current resemble those seen in the SOM from the unfiltered SLA (Fig. 1). Such discrepancies were anticipated due to several factors: 1) the absence of high-pass filtering on satellite observations due to its short record, 2) the climate model's limitation in resolving subgrid-scale processes, and 3) the presence of other processes in the observations, such as the glacial isostatic adjustment, which might influence regional sea levels but are not accounted for in the model simulations. Due to the short duration of the satellite observations, the longer timescale buoyancy-driven AMOC signals may not be well-captured. Consequently, the composites of observational SLA on the SOM associated with unfiltered piControl SLA patterns (with the SOM displayed in Fig. 1) show much less similarity to the SLA patterns from unfiltered piControl simulations, since the latter are closely linked to low-frequency AMOC signals.
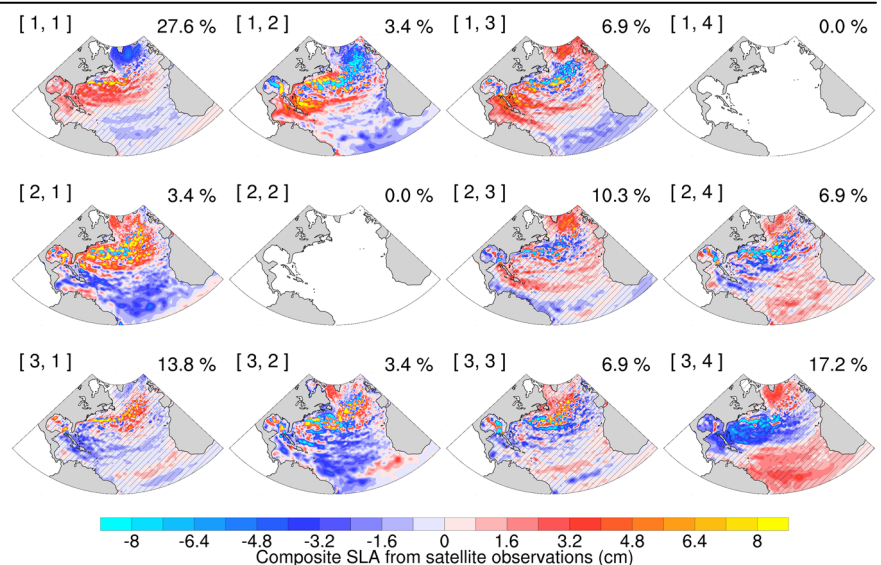
It is noteworthy that the frequency of observational SLA classified into node [1,1] is exceptionally high, reaching 27.6% (8 out of 29 years). A closer examination indicates that the satellite observation is persistently classified as node [1,1] from 2015 to 2021. This recent increase in the occurrence of node [1,1] suggests an accelerated increase in the coastal sea level in the U.S. Southeast, relative to other regions, which aligns with findings documented in recent studies[16,21–23,32].

The persistence of node [1,1], shown in the observations, contrasts sharply with its representation in high-pass filtered SLA from the piControl simulations. In these simulations, the probability of remaining in node [1,1] five years after starting from this node itself is only 2% and 1% for the
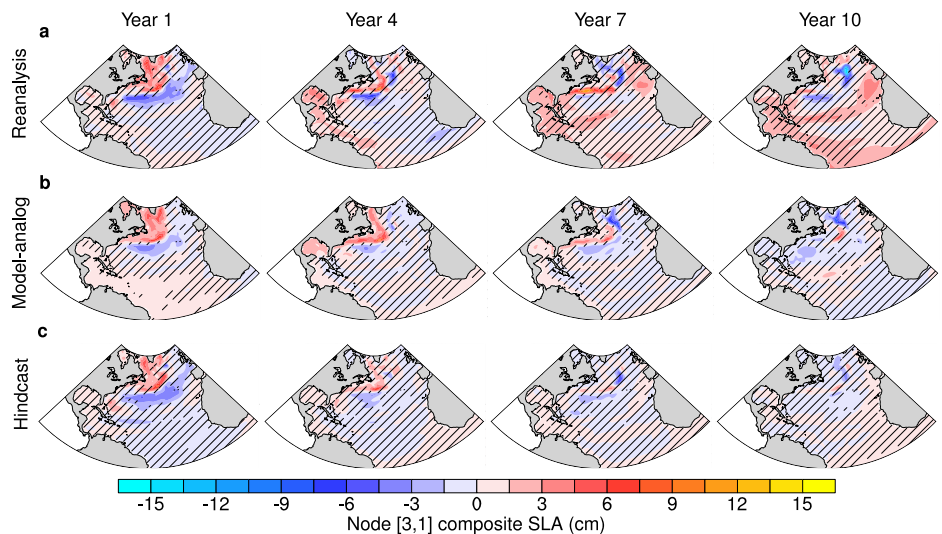
Fig. 7 | Transition probability tables used to track the temporal evolution of high-pass filtered SLA. Same as Fig. 3 but for high-pass filtered SLA patterns from the SPEAR_LO piControl simulation shown in Fig. 6 at **a** lag 1, **b** lag 3, **c** lag 5, and **d** lag 7 years.
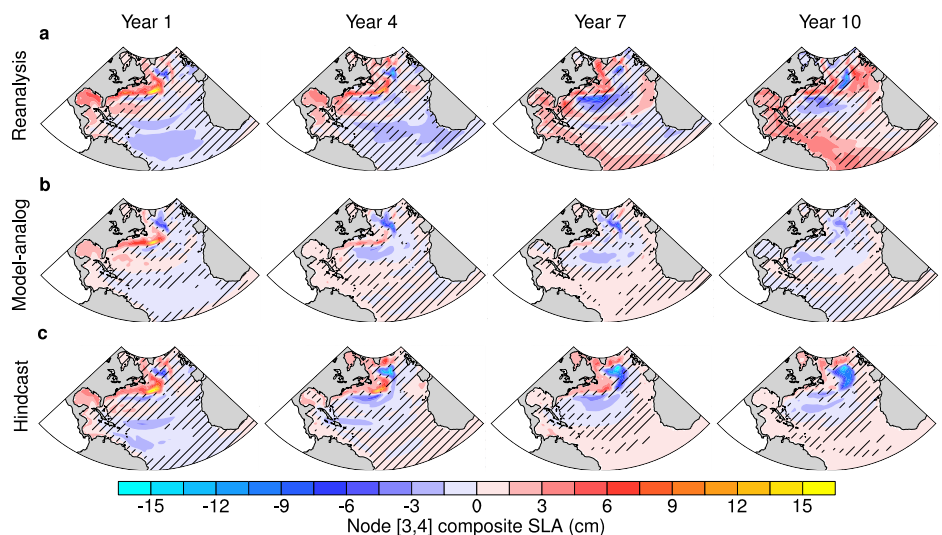
Fig. 8 | Composites of annual mean dynamic sea level anomalies from satellite observations (cm; shading) for all SOM nodes generated from high-pass filtered SLA patterns in the piControl simulations (with the SOM shown in Fig. 6). The top right corner of each subplot displays the frequency of occurrence of the corresponding node in satellite observations. The absence of black hatching in nodes indicates that the shaded value is significant at the 95% confidence level using a two-tailed *t* test, except for nodes [1,2], [2,1], and [3,2] where only one year composited, precluding a significance test.

Fig. 9 | SOM-based predictions of SLA patterns from initialized decadal hindcasts and model-analogs for node [3,1]. a Composites of SLA patterns from the five reanalysis years classified in node [3,1] (Year 1) and their following evolutions on Years 4, 7, and 10. **b** Composites of model-analog SLA patterns including ten most similar SLA patterns from piControl simulations for each reanalysis year assigned to node [3,1] (Year 1), and their subsequent evolutions on Years 4, 7, and 10. **c** Composites of SLA patterns from decadal hindcasts, initialized on 1 January of the years that correspond to those classified in node [3,1] from the reanalysis (Year 1), and the subsequent hindcasts on Years 4, 7, and 10. The absence of black hatching indicates that the shaded value is significant at the 95% confidence level using a two-tailed *t* test.
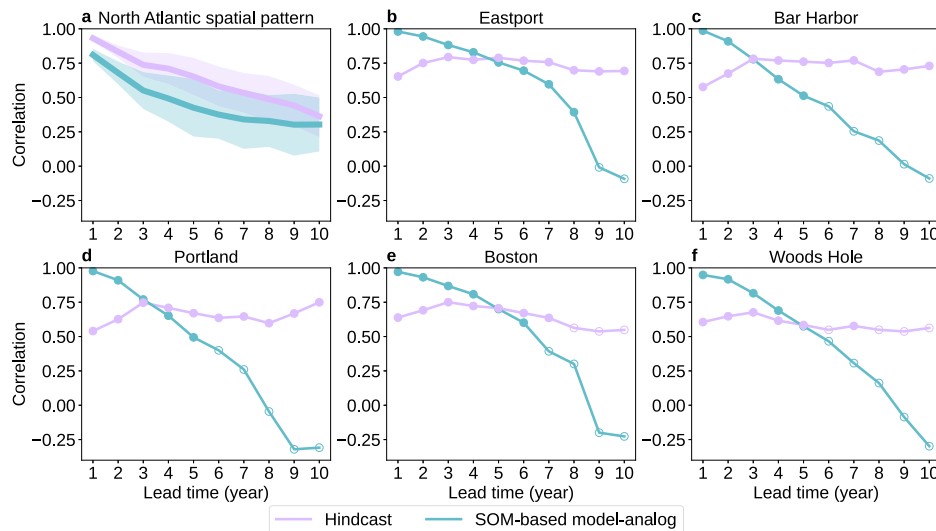

Fig. 10 | SOM-based predictions of SLA patterns from initialized decadal hindcasts and model-analogs for node [3,4]. a, **b**, and **c** are the same as those in Fig. 9 but for node [3,4]. There are eight reanalysis years classified in [3,4], which are included in the composite shown in **a**.

SPEAR_LO and SPEAR_MED piControl simulations, respectively, as demonstrated in Supplementary Movies 3 and 4. This substantial difference between the observations and the high-pass filtered piControl simulations could be attributed to the impact of external forcing, which may not be fully removed by subtracting the global mean sea level from the observations, or to longer timescale processes retained in the satellite observations due to the lack of high-pass filtering. External forcing may lead to the persistence of node [1,1], either through its direct influence, reflected as a trend, or indirectly by modulating internal variability, even after the trend is removed[65,66]. To explore both possibilities, we employ high-pass filtered SLA patterns from 1993 to 2021 in the SPEAR large ensemble[48], conducting one analysis with the ensemble mean retained to assess the direct impact of external forcing, and another with the ensemble mean removed from each grid point in each year to evaluate its modulation effect on internal variability. Specifically, we composite these high-pass filtered SLA patterns onto the SOM generated from high-pass filtered piControl simulations to evaluate how external forcing influences patterns of high-frequency variability and the persistence of node [1,1]. Surprisingly, the resulting composite SLA patterns from the large ensemble, whether the ensemble mean is retained or removed, closely resemble those from the piControl simulations (Supplementary Figs. 10 and 11). Moreover, based on the transition probabilities

associated with these composite patterns from the SPEAR large ensemble, the persistence of node [1,1] after five years is 4% for the patterns with ensemble mean (Supplementary Fig. 12) and 0% for the patterns without ensemble mean (Supplementary Fig. 13), similar to the piControl simulations and much lower than the observations. These results suggest a limited role of external forcing on short timescale sea level variability. However, given model biases, such as different sensitivities to external forcing compared to observations, we cannot completely eliminate the potential impact of external forcing on the observed persistence of node [1,1]. Nonetheless, the notable persistence of node [1,1] seen only in the observations, coupled with the presence of features resembling AMOC SLA fingerprints, suggests that the observed variability cannot be solely attributed to short-term atmospheric variability, its impact on wind-driven ocean circulation, or external forcing. Instead, it is also shaped by mechanisms operating on longer timescales, likely involving the low-frequency buoyancy-driven AMOC[16,23,32,67–69]. Consequently, as some features of the observational dynamic sea level variability resemble high-pass filtered variability while others align with unfiltered variability in the piControl simulations, understanding their predictability requires consideration of processes operating across both longer and shorter timescales (a combination of Fig. 5a, b).

Fig. 11 | Skill assessment for the SLA predictions over a lead time of 1–10 years.
a Pattern correlation (ACC) between the SLA patterns from the reanalysis and from
predictions through initialized decadal hindcasts (purple) and the SOM-based
model-analog method (cyan). The solid thick lines denote the mean ACC over all
nodes, with shadings showing the mean ACC ± one standard deviation of ACC
across all nodes. **b** Correlation between the low-pass filtered, detrended tide gauge
record at Eastport and the predicted time series through initialized decadal hind-
casts (purple) and the SOM-based model-analog method (cyan) on the
corresponding closest model grid point. Solid circles denote significantly positive
correlations, while open circles indicate correlations that are either not significant or
significantly negative at the 90% confidence level, with the statistical significance
determined using a *t* test based on an effective sample size[99–101] (see Methods for
details). **c–f** Same as **b**, but for the tide gauge stations and the corresponding closest
model grid points in Bar Harbor, Portland, Boston, and Woods Hole, respectively.
Locations of the tide gauge stations and the closest model grid points are shown in
Supplementary Fig. 15.

Prediction of SLA using the SOM-based model-analog method

After assessing the variability and predictability of the North Atlantic SLA
using SOM-based methods, we delve into the potential of perfect-model
predictability, as identified in the piControl simulations, to foster skillful
decadal predictions. To achieve this, we apply a model-analog method[49–52]
on a per-SOM node basis associated with the SOM from the unfiltered SLA
patterns in the piControl simulations (Fig. 1). We refer to this as a SOM-
based model-analog prediction. Specifically, we use composites of a certain
number of SLA patterns from the piControl simulations, associated with the
SOM, and their subsequent evolutions as the model-analog predictions of
the target benchmark SLA patterns. Here we focus on the SOM derived from
the unfiltered piControl SLA patterns, as these unfiltered patterns, with both
low- and high-frequency signals retained, capture much longer theoretical
predictability compared to the high-pass filtered patterns (Fig. 5).

The benchmark used to evaluate predicted North Atlantic SLA pat-
terns is from an observationally constrained reanalysis dataset by SPEAR,
with the linear trend removed to focus on internal variability. Here, we opt
for reanalysis over satellite observations because 1) the reanalysis dataset
offers a substantially longer historical record, and 2) it provides a reasonable
representation of AMOC evolution, facilitated by quasi-observational
boundary conditions[15,70] (see Methods for details). To create a SOM-based
benchmark, we map the SLA patterns from the reanalysis dataset to their
most similar SOM node from the piControl simulations (Fig. 1), based on a
minimum Euclidean distance principle (see Methods for details). Then we
construct the composite of the SLAs from the reanalysis for each SOM,
including those with time lags (Fig. 9a for node [3,1] and Fig. 10a for [3,4]).

The SOM-based model-analog prediction for the reanalysis composite
in each SOM node involves selecting the ten most similar piControl
simulation analogs from the corresponding SOM node for each reanalysis
year assigned to this node. These selected analogs and their subsequent
evolutions are then composited to form the SOM-based model-analog
prediction for each node (Fig. 9b for node [3,1] and Fig. 10b for [3,4]). It is
notable that the model-analog prediction skill exhibits low sensitivity both
to the number of analogs used (ten in this instance) and to whether the
analogs are confined to the specific SOM node associated with the reana-
lysis. Increasing the number of analogs, however, can decrease the

prediction skill during the early lead times. As a comparison, we also assess
the skill of a widely used method, the initialized decadal hindcasts produced
by SPEAR[70] (with the linear trend removed for each lead time; see Methods
for details). Specifically, the 10-year decadal hindcasts, initialized from the
same years as the reanalysis associated with each SOM node, are composited
onto the corresponding node (Fig. 9c for node [3,1] and Fig. 10c for [3,4]).
The hindcast model is initialized on 1 January of each year, with outputs for
the remainder of that year considered as predictions. Therefore, we denote
the year of initialization as a lead time of 1 year for the hindcasts. Following
this naming convention, a lead time of 1 year in predictions from the model-
analog method signifies the composites of piControl SLA patterns onto the
SOM without lags.

Here we present the predictions of nodes [3,1] (Fig. 9) and [3,4]
(Fig. 10) as examples, as they initialize from mature and transition AMOC
phases, respectively. Node [3,1] starts from a mature negative AMOC phase,
leading to higher-than-normal SLA in the U.S. Northeast Coast (Fig. 9a).
This feature is well-captured in both model-analogs and initialized hindcasts
(Fig. 9b, c). After several years, the increase in SLA to the north gradually
diminishes, and a narrow band of positive SLA extending to the U.S.
Southeast Coast emerges as shown in the reanalysis (Fig. 9a). Surprisingly,
the model-analogs identify this change more distinctly compared to the
hindcast, which exhibits much weaker anomalies than that in the reanalysis
(Year 7, Fig. 9). In contrast to node [3,1], the prediction skill of the model-
analog method for node [3,4] falls short of the hindcast (Fig. 10), with the
signal gradually diminished from Year 7. This comparison of model-analog
prediction skills between nodes [3,1] and [3,4] highlights the importance of
the initial AMOC state to the prediction skill using this method. That is, a
prediction starting from a mature AMOC state has higher skill than one
from a transition state, closely aligning with a previous study[15].

To evaluate the predictions associated with all SOM nodes, we further
quantify the pattern correlations between the reanalysis and predictions
from the model-analogs and hindcasts using the Anomaly Correlation
Coefficient (ACC; see Methods for details). In general, the hindcast out-
performs the model-analog prediction for most nodes in the first several
years (Fig. 11a), which may be attributed to the effective initialization in the
hindcast[70]. Then the spread of the prediction skill of hindcast starts to

overlap with that of the model-analog prediction. Although the overall skill is lower in the model-analog method, it can exhibit notable skill for SLA patterns starting from certain states (e.g., node [3,1]). Another advantage of the model-analog method is its ability to provide reasonably good prediction skill over longer lead times for certain initial states. Specifically, nodes [1,2], [1,3], [2,1], [2,2], [3,1] exhibit an ACC greater than 0.5 after a lead time of 10 years (Supplementary Fig. 14). Therefore, this method can provide a cost-effective preliminary assessment of the internal variability in future SLA states before hindcasts, which conventionally extend up to 10 years, reach those points.

Evaluating predictions against the reanalysis enables an examination of the prediction skill across the entire North Atlantic basin. However, the SPEAR reanalysis cannot be considered an independent dataset for model-analogs and hindcasts. Therefore, we extend our evaluation of prediction skill to the U.S. coastal regions, utilizing annual mean tide gauge records. At each station, the tide gauge record undergoes linear detrending to focus on internal variability and is subjected to an 11-year running mean low-pass filter to remove short timescale variations, such as the inverse barometer effect, thereby focusing on decadal timescale variability. In a manner similar to the reanalysis, the record from each tide gauge station for each year is matched to a SOM node based on its similarity to the nearest model grid point. For the model-analog method, rather than selecting the ten most similar analogs, we employ all SLA from the piControl simulations that are categorized within the SOM node corresponding to the tide gauge data. While the differences are subtle, using a broader set of analogs enhances the prediction skill. Notably, restricting analogs to the SOM node linked with the tide gauge data substantially boosts the prediction skill, as opposed to merely comparing the tide gauge records with individual model grid points. This approach incorporates the information associated with large-scale patterns, which are often linked to low-frequency variability.

Figure 11b–d shows the correlation between a number of tide gauge records along the U.S. Northeast Coast and predictions from the hindcasts and SOM-based model-analogs at the closest model grid point to the corresponding tide gauge station (see the locations in Supplementary Fig. 15). Here we focus on the Northeast Coast, where low-frequency variability from the AMOC and the Gulf Stream enhances the predictability of low-pass filtered tide gauge records[15,19,20,24,25]. In general, the SOM-based model-analogs demonstrate a significantly positive correlation with tide gauge records during the initial 5–8 lead years, after which the correlation decreases (Fig. 11b–d). Unlike Fig. 11a, which measures the pattern correlation across the entire North Atlantic basin, the model-analogs perform better than the hindcasts in the first several lead years at tide gauge stations along the U.S. Northeast Coast. However, the skill of the hindcasts remains consistently high, maintaining a positive correlation exceeding 0.5 throughout. We speculate that this high skill is due to an accurate representation of low-frequency variability in the tide gauge records across all lead times. In contrast, the model-analogs exhibit a decline in skill as lead time increases. This decline is anticipated to some degree due to the predominant timescale of low-frequency AMOC variability in the piControl simulations, which is 25–40 years, being shorter than the timescale in the observations and hindcasts.

Discussion

In summary, we utilize a SOM-based framework to identify dominant patterns of SLA variability in the North Atlantic and assess the associated predictability on multiyear-to-decadal timescales within the SPEAR_LO and SPEAR_MED piControl simulations. These dominant patterns of internal variability are closely linked to different phases of AMOC, corroborating findings from previous studies focusing on specific AMOC phases[15,19,25]. This low-frequency buoyancy-driven AMOC variability associated with the SLA leads to the persistence of SLA patterns and their directional transition across the SOM, resulting in prolonged decadal predictability that lasts for 15–20 years. Moreover, SLA variability excluding longer timescale signals dominated by buoyancy-driven AMOC is characterized using the SOM after applying a 15-year high-pass filter to the annual mean SLA patterns. Additional predictability on shorter timescales is then identified, which we speculate is associated with ocean memory[58,59], the transient-eddy feedback[46,47,60–64], or both, within the first year. Moreover, gyre circulation adjustments (e.g., through oceanic Rossby wave propagation[23,71,72]) triggered by the initial stochastic forcing (e.g., NAO) can extend the predictability to 4–5 years.

Composites of dynamic sea level from satellite observations, when mapped onto the SOM, reveal large-scale patterns similar to the SOM patterns associated with the 15-year high-pass filtered SLAs from the piControl simulations. This is due to the short observational record that inadequately captures the longer timescale buoyancy-driven AMOC signals. Nevertheless, the signal of buoyancy-driven AMOC emerges in these observational composites (e.g., enlarged SLAs along the Gulf Stream and North Atlantic Current), resembling those in the SOM patterns from unfiltered SLAs. This suggests that when considering the predictability of dynamic sea level in observations, it is crucial to integrate predictability sources on both longer and shorter timescales, affirming the results from Zhang et al. (2024)[23].

Given the prolonged predictability identified in the piControl simulations, we further conduct decadal predictions of the internal SLA variability using a SOM-based model-analog method. Compared to a benchmark derived from an observationally constrained reanalysis dataset, the overall predictions from model-analogs are less accurate in predicting large-scale North Atlantic SLA patterns than conventional initialized hindcasts. However, predictions that start from certain SLA states (e.g., those associated with a mature AMOC state) can achieve reasonably accurate prediction skills over an extended period. Importantly, this model-analog method is cost-effective, allowing for long-lead predictions when necessary. Therefore, it can serve as a useful tool for early estimation of large-scale SLA states, bridging the gap until the initialized decadal predictions from dynamical models, which typically cover up to 10 years, become available. This ability is useful for informing policy and decision makers. However, determining whether the AMOC is in a state that can provide high predictability to SLA patterns may be challenging due to the limited duration of direct AMOC observations and the projected future changes in both the AMOC[73–77] and North Atlantic internal variability[66]. Long-term proxy data[78–80] may help identify periods when the AMOC is in such a state; however, caution is needed in light of the evolving impacts of climate change.

Our focus here is on the large-scale North Atlantic SLA patterns, which may have implications for coastal regions. On decadal timescales, it is typically the large-scale patterns that offer some level of predictability. This predictability is reflected in the ability of our SOM-based model-analogs to predict low-frequency U.S. Northeast coastal SLA up to 5–8 years in advance, where the analogs are selected based on large-scale patterns. This SOM-based coastal sea level prediction serves as a testbed to demonstrate the linkage between large-scale SLA patterns and low-frequency coastal SLA. It offers improved prediction skill in the first several lead years in a cost-effective way compared to initialized decadal hindcasts. Future research could enhance prediction skill over longer timescales by training a SOM on model simulations that exhibit a predominant timescale of AMOC variability similar to that in observations. Furthermore, coastal sea level is subject to many other processes that are not represented in the model, such as glacial isostatic adjustment, vertical land motion, land ice melting, and gravitational, rotational, and deformational effects. Future research could also investigate whether long-term predictability exists within these processes, potentially contributing to even more informed decadal predictions.

In this study, we examine two piControl simulations, each spanning 2500 years but with different atmospheric resolutions. The SLA variability captured by these simulations is comparable. However, SPEAR_MED exhibits slightly higher predictability, and we speculate that this is due to its larger low-frequency AMOC variability compared to SPEAR_LO. Future work could explore the impact of horizontal ocean grid resolution on SLA variability and predictability by leveraging long simulations with higher-resolution ocean components.

Methods

Model simulations

In this study, we use simulations from the National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Laboratory (GFDL) SPEAR model[48], which include two long piControl simulations from SPEAR_LO and SPEAR_MED, with the atmospheric composition fixed at levels representative of the year 1850. SPEAR_LO is fully coupled and consists of ocean and ice components from Modular Ocean Model version 6 (MOM6)[81] and atmosphere and land components from AM4.0-LM4.0[82,83], with nominal 1° horizontal resolution in all components. SPEAR_MED is the same as SPEAR_LO but with 0.5° horizontal resolution in its atmosphere and land components. A total of 2500 years from both piControl simulations (model years 201–2700) are used to characterize the multiyear-to-decadal variability of SLA patterns in the North Atlantic, after omitting the first 200 years to account for model spin-up. All data from both simulations are linearly detrended to remove the model drift.

In addition to the piControl simulations, the SPEAR large ensemble[48], which contains 30 members, is used to assess the role of external forcing on internal variability. Simulations are performed with SPEAR_MED, initialized from the piControl simulation with a 20-year interval between each member. These simulations are conducted under the historical forcing from 1921 to 2014, and the Shared Socioeconomic Pathway 5–8.5 (SSP5–8.5) scenario[84,85] from 2015 to 2100.

Furthermore, the SPEAR reanalysis and initialized decadal hindcasts based on SPEAR_LO[70] are used to evaluate the prediction skill of the SOM-based model-analogs (described more thoroughly in the main text). In the SPEAR reanalysis, the 6-hourly atmospheric temperature and winds were restored towards the 55-year Japanese Reanalysis (JRA55)[86], and the SSTs were restored towards the NOAA Extended Reconstructed Sea Surface Temperature version 5 (ERSSTv5)[87]. These atmospheric and SST constraints ensure that the ocean component of the reanalysis adheres to quasi-observational boundary conditions, leading to reasonable AMOC multidecadal variability. Here we use the 30 ensemble members contained in this reanalysis over the 1961–2020 period. Before they are compared with the predicted SLA patterns from the model-analog method, linear detrending is applied to each ensemble member, after which the ensemble mean is computed and used as the benchmark.

The decadal hindcasts used contain 20 members, each initialized from the SPEAR reanalysis on 1 January every year from 1961 to 2020, and integrated for 10 years with realistic time-evolving anthropogenic and natural forcing conditions. To remove the systematic model drift, the lead time-dependent climatology is subtracted from the hindcasts. Linear detrending is then applied to each lead time of each ensemble member. Similar to the reanalysis, here we use the ensemble mean across the 20 members.

Observations

We utilize the gridded monthly mean global SLA estimates based on satellite altimetry measurements over the 1993–2021 period, released by the Copernicus Marine Environment Monitoring Service and the Copernicus Climate Change Service[88]. This dataset has a horizontal resolution of 0.25° × 0.25°. Before comparing it with the SLA patterns from the piControl simulations, the annual mean SLA is calculated from the monthly mean measurements, the global mean sea level is subtracted from each grid point to derive the dynamic sea level, and the 1993–2021 climatology is subtracted to obtain the anomalies.

Additionally, we utilize coastal sea level records from tide gauge stations, processed and distributed by the Permanent Service for Mean Sea Level (PSMSL)[89], to evaluate sea level predictions along the U.S. Northeast Coast. We select data from 1956 to 2023 for all the stations used, remove the linear trend, and apply an 11-year running mean to focus on the low-frequency internal variability. After excluding the first and last five years affected by the low-pass filter, the record used is refined to 1961–2018.

Self-organizing maps

In the SOM method, a predetermined number of generalized patterns, known as SOM nodes, are trained to approximate the distribution of the input data. In a well-constructed SOM, these SOM nodes are spatially organized, conserving topological order. That is, adjacent SOM nodes represent similar patterns, while distant ones exhibit different features. Unlike other classification methods that maximize the difference between clusters (e.g., k-means clustering[90]), SOM views the input data distribution as a continuum. Consequently, SOM is trained to span the input data space. This characteristic renders SOM especially effective for examining fields that vary continuously, such as sea level pressure[35] and SST[45]. Additionally, SOM can capture nonlinear behaviors and is not constrained by orthogonality or stationarity. This endows SOM an advantage over principal component analysis, which is designed to maximize the variance and might merge physically distinct patterns of variability[91,92]. In contrast, SOM can identify physically relevant patterns[35], thereby offering insights into the mechanisms and processes underlying these patterns.

To create a SOM, the SOM size (i.e., the number and arrangement of nodes) and other relevant parameters are first specified. All SOM nodes have the same size as an input data vector. These nodes are initialized with random values, and then trained by repeatedly being compared to input data vectors and modified. We use N to denote the number of times each input data vector is compared to the nodes per training process. Therefore, for a training process, the total training steps L is the number of input data vectors times N. At each training step n, spanning from 1 to L, an input data vector $\mathbf{x}(n)$ is compared against each SOM node $\mathbf{m}_i(n)$, where i is the SOM node index and $\mathbf{m}_i$ denotes the SOM node vector associated with node i. The Euclidean distance between $\mathbf{x}(n)$ and $\mathbf{m}_i(n)$ is computed for each SOM node, and the node with the minimum Euclidean distance is selected as the best matching unit (BMU) c for this training step. The SOM nodes are then updated towards the current input data vector in this training step as follows:

$$\mathbf{m}_i(n+1) = \mathbf{m}_i(n) + \alpha(n) \cdot h_{ci}(n) \cdot (\mathbf{x}(n) - \mathbf{m}_i(n)),$$

where $\mathbf{m}_i(n+1)$ denotes the data vector associated with SOM node i at training step $n+1$, updated from its previous state $\mathbf{m}_i(n)$ at step n, $\alpha(n)$ is the learning rate parameter at step n, and $h_{ci}(n)$ denotes the neighborhood function at step n[93]. $\alpha(n)$ denotes the extent to which SOM nodes are updated at each training step. Here the inverse parameter is utilized, defined as follows:

$$\alpha(n) = \alpha_0 / \left(1 + 100\frac{n}{L}\right),$$

where α_0 is the initial α value that is specified before training. $\alpha(n)$ is designed to facilitate larger updates at the beginning of each training process and more subtle tuning later. The neighborhood function $h_{ci}(n)$ dictates the shape of the influence of input data on SOM nodes. The Epanechikov function is applied here as it has been shown to outperform other neighborhood functions[94]:

$$h_{ci} = \max\left(0, 1 - \frac{d_{ci}^2}{\sigma(n)^2}\right).$$

Here d_{ci} denotes the distance between the BMU c and each SOM node i, $\sigma(n)$ represents the radius of influence at training step n. According to this equation, h_{ci} is maximized when d_{ci} is 0, indicating that the impact of each input vector on the nodes is maximized at its respective BMU in the neighborhood function. The impact of an input data vector on SOM nodes diminishes as d_{ci} increases, with nodes beyond $\sigma(n)$ remaining unaffected. In this study, we start with a larger σ that impacts a wide array of nodes and decreases its value with training time. This neighborhood function establishes the topological order of the SOM, where adjacent nodes are similar and distant nodes are different, thereby leading to the key difference between the SOM method and the k-means clustering. Further details on parameters

and the quality of SOM are available in Kohonen (2021)[93] and Gu et al. (2023)[47]. The SOM Program Package (SOM_PAK) is available online[95] (http://www.cis.hut.fi/research/som-research/).

In this study, we use the SOM method to investigate the multiyear-to-decadal variability and predictability of SLA in the North Atlantic (0-65°N, 100°W-0), with and without the buoyancy-driven AMOC-related longer timescale signal. Specifically, we train three 3 × 4 SOMs, one for the annual mean SLA patterns in the 5000-year piControl simulations from the SPEAR_MED and SPEAR_LO, a second for the same simulations but with a 15-year high-pass Butterworth filter applied to the annual mean SLA field before training, and a third with a 15-year low-pass Butterworth filter applied. We note that the first and last 15 years are removed from both simulations for the second and third SOMs to mitigate the edge effects of the filtering, resulting in a total of 4940 years. For all three SOMs, we conduct two sets of trainings, with each data vector presented to the nodes 50 times per training (N equals 50). The radii of influence σ_t are set to 5 and 2 and α_0 are set to 0.1 and 0.01 for the first and second trainings, respectively. Well-trained SOMs lead to similar patterns even with different tuning parameters. We also test different SOM sizes for both unfiltered and high-pass filtered SLA patterns, as shown in Supplementary Figs. 4 and 16–18. A larger SOM size leads to more phases of SLA patterns that can be viewed as transition states of the existing patterns in the 3 × 4 SOMs, while a smaller SOM size shows fewer phases and can merge physically distinct patterns of variability. Overall, SOMs of various sizes display similar dominant patterns of variability. The 3 × 4 SOMs are used in this study as they encompass all the physically relevant patterns needed for our purpose, as well as an appropriate number of transition states.

After the training process, each input data vector from both simulations is assigned to its final BMU, that is, the SOM node most closely resembling the input data vector on the basis of minimum Euclidean distance. The frequency of each SOM node within each simulation or across both simulations can be calculated as the percentage of input data vectors whose final BMU is the corresponding SOM node in the corresponding simulation(s). With the final BMU identified, we then create composites of other variables (e.g., SST, meridional overturning streamfunction) associated with SOM nodes for each piControl simulation. This enables us to investigate the processes associated with each SLA classification.

In addition to the variables from the same simulations used to train the SOM, we composite the SLA patterns from satellite observations to assess whether they exhibit similar patterns of variability. To accomplish this, we follow these steps: (1) Remove the global mean sea level from each grid point to derive the dynamic sea level, and then subtract the 1993–2021 climatology to obtain the SLA, making it more comparable to the SLA patterns from piControl simulations. (2) Interpolate the observational SLA onto the model grid. (3) Calculate the Euclidean distance between each observational SLA pattern and all SOM nodes, selecting the node with the minimum distance as its BMU. (4) Composite the original (not interpolated) observational SLA associated with each SOM node, based on their BMUs. Using a similar approach as in steps 3–4, the SLA patterns from the reanalysis dataset are composited onto the SOM as the benchmark for the SOM-based decadal predictions.

To evaluate the role of external forcing on short timescale SLA variability, we also composite the SLA patterns from the 30-member SPEAR large ensemble[48] onto the SOM generated from high-pass filtered SLA patterns in the SPEAR piControl simulations. Specifically, two sets of composites are conducted: one with the ensemble mean retained and another with the ensemble mean removed from each grid point in each year. To focus on short timescales, we apply a 15-year high-pass filter and then use the SLA patterns over the 1993–2021 period for the composite analysis to align with the period used for the satellite observations.

Transition probability table
To assess the temporal evolution of SLA patterns across the SOM, we create a set of transition probability tables for SOMs following Gu and Gervais[45].

Each input data vector (an annual mean SLA pattern) associated with a SOM has its own BMU. The BMU series for all input data in each simulation can be represented as a time series of discrete variables with values ranging from 1 to 12 for a 3 × 4 SOM map (one value for each node). These 12 states form a mutually exclusive and collectively exhaustive partition of the SLA patterns.

The temporal evolution of the SLA patterns can then be represented using transition probability, defined as:

$$p_{ij}(n) = \Pr\{\mathbf{x}_{t+n} = j | \mathbf{x}_t = i\},$$

where $\mathbf{x}_t$ is the SLA pattern at a specific year t, $\mathbf{x}_{t+n}$ is the SLA pattern at year $t + n$ with n being the time lag, and p_{ij} is the conditional probability of the occurrence of node j at lag n year given the occurrence of node i at lag 0. The transition probabilities are calculated over lags for each simulation within each SOM (Supplementary Movies 1–4, Figs. 3, 4, and 7, and Supplementary Figs. 12 and 13). In each table, the 3 × 4 large cells represent the initial node i and the 3 × 4 small cells within each large cell represent the node j. The number in each small cell is p_{ij}.

A Monte Carlo method is applied to evaluate the significance of the transition probabilities. The null hypothesis is that the transition probability equals the unconditional probability, which is the frequency of the occurrence of each SOM node without any given information. A null distribution is created for each SOM by repeatedly calculating the unconditional probability from a number of randomly selected years 100,000 times. The numbers of years selected are 208 and 206 for each piControl simulation associated with the SOM generated from unfiltered and high-pass filtered SLA patterns, respectively. For the SPEAR large ensemble composites associated with the SOM generated from high-pass filtered piControl simulations, the selected number of years is 73. These values represent the average number of input data from each simulation classified within each SOM node. Transition probabilities greater (smaller) than the 97.5th (2.5th) percentile of the null distribution are considered to be significantly higher (lower) than normal, denoted by warm (cold) shadings in the transition probability tables (Supplementary Movies 1–4, Figs. 3, 4, and 7, and Supplementary Figs. 12 and 13).

The global significance theory[96,97] is then utilized to determine the overall significance level of an initial node (global significance) based on the number of significant transition probabilities when starting from this node (local significance)[45]. Specifically, the global null hypothesis, which asserts that all local null hypotheses are true, is rejected if there are m or more locally significant probabilities:

$$\Pr(M \geq m) \leq \alpha_{global},$$

where M represents the number of local tests that reject their null hypotheses, m denotes the threshold number of rejections needed among the local tests to reject the global null hypothesis, and α_{global} is the global significance level.

Assuming that the local tests are mutually independent, the probability of $M \geq m$ can be calculated using the binomial distribution,

$$\Pr(M \geq m) = \sum_{i=m}^{K} \frac{K!}{i!(K-i)!} (\alpha_{local})^i (1 - \alpha_{local})^{K-i},$$

where K is 12, and α_{local} is equal to 0.05. For a global significance level of 0.05, m is 3. That is, if the number of significant probabilities starting from one initial node is greater than or equal to 3 at a given time lag, the transition probabilities associated with the whole initial node are viewed as globally significant at this lag. The global significance of a node at a certain time lag suggests that this initial node has a significant number of nodes to which it does or does not tend to transition to at this lag. Therefore, if a given node at a given lag is globally significant, we view the pattern starting from this node as theoretically predictable at that lag.

Pattern correlation

We verify the prediction of North Atlantic SLA patterns from the model-analogs and initialized hindcasts using the ACC, which is widely used to quantify the spatial similarity between patterns[98]. The ACC between two patterns **P** and **Q** is calculated as follows:

$$ACC \equiv \frac{\sum_{g=1}^{G} w_g P_g Q_g}{\sqrt{\sum_{g=1}^{G} w_g P_g{}^2 \sum_{g=1}^{G} w_g Q_g{}^2}}, (-1 \le ACC \le 1),$$

where P_g and Q_g denote the anomalies for each grid cell g with respect to their climatology in **P** and **Q**, respectively; w_g is the weight, calculated as the horizontal area for each grid cell g; G is the total number of grid cells within the domain.

Effective sample size

To account for the autocorrelation in the data when computing the correlations between the tide gauge records and predictions, we use a t test based on an effective sample size to determine the statistical significance of correlations. Specifically, we estimate the effective sample size S^* using the following equation[99,100]:

$$S^* = \frac{S}{1 + 2\sum_{\tau=1}^{max\_lag}\left(1 - \frac{\tau}{S}\right)\rho_{xx}(\tau)\rho_{yy}(\tau)},$$

where S is the original sample size, $\rho_{xx}(\tau)$ and $\rho_{yy}(\tau)$ are the autocorrelation coefficients of the two time series x and y at lag τ. Here we consider the autocorrelation up to a maximum lag of 5.

Data availability

The global estimates of SLA based on satellite altimetry measurements are available through Copernicus Marine Environment Monitoring Service and the Copernicus Climate Change Service, at https://doi.org/10.24381/cds.4c328c78. The tide gauge sea level records from the PSMSL are available at https://psmsl.org/.

Code availability

The source code of ocean component MOM6 of SPEAR_LO model is available at https://github.com/NOAA-GFDL/MOM6.

References

1. Nicholls, R. J. et al. Coastal systems and low-lying areas. Climate change 2007: impacts, adaptation and vulnerability. 315–356 (2007).
2. FitzGerald, D. M., Fenster, M. S., Argow, B. A. & Buynevich, I. V. Coastal impacts due to sea-level rise. *Annu. Rev. Earth Planet. Sci.* **36**, 601–647 (2008).
3. Nicholls, R. J. & Cazenave, A. Sea-level rise and its impact on coastal zones. *Science* **328**, 1517–1520 (2010).
4. Kirwan, M. L. & Megonigal, J. P. Tidal wetland stability in the face of human impacts and sea-level rise. *Nature* **504**, 53–60 (2013).
5. Woodruff, J. D., Irish, J. L. & Camargo, S. J. Coastal flooding by tropical cyclones and sea-level rise. *Nature* **504**, 44–52 (2013).
6. Cazenave, A. & Cozannet, G. Le. Sea level rise and its coastal impacts. *Earth's Future* **2**, 15–34 (2014).
7. Moftakhari, H. R. et al. Increased nuisance flooding along the coasts of the United States due to sea level rise: past and future. *Geophys. Res. Lett.* **42**, 9846–9852 (2015).
8. Meehl, G. A. et al. How much more global warming and sea level rise? *Science* **307**, 1769–1772 (2005).
9. Tebaldi, C. et al. Extreme sea levels at different global warming levels. *Nat. Clim. Change* **11**, 746–751 (2021).
10. Pugh, D. T. Tides, surges and mean sea-level (reprinted with corrections). pp, 486 (1996).
11. Rahmstorf, S. Rising hazard of storm-surge flooding. *Proc. Natl Acad. Sci. USA* **114**, 11806–11808 (2017).
12. Landerer, F. W., Jungclaus, J. H. & Marotzke, J. El Niño-Southern Oscillation signals in sea level, surface mass redistribution, and degree-two geoid coefficients. *J. Geophys. Res. Oceans* **113**, 1–16 (2008).
13. Zhang, X. & Church, J. A. Sea level trends, interannual and decadal variability in the Pacific Ocean. *Geophys. Res. Lett.* **39**, 1–8 (2012).
14. Holbrook, N. J. et al. ENSO-driven ocean extremes and their ecosystem impacts. *Geophys. Monogr. Ser.* **253**, 409–428 (2020).
15. Zhang, L., Delworth, T. L., Yang, X. & Zeng, F. Skillful multiyear to decadal predictions of sea level in the North Atlantic Ocean and US east coast. *Commun. Earth Environ.* **4**, (2023).
16. Yin, J. Rapid decadal acceleration of sea level rise along the US east and gulf coasts during 2010–22 and its impact on hurricane-induced storm surge. *J. Clim.* **36**, 4511–4529 (2023).
17. Kemp, A. C. & Horton, B. P. Contribution of relative sea-level rise to historical hurricane flooding in New York City. *J. Quat. Sci.* **28**, 537–541 (2013).
18. Wdowinski, S., Bray, R., Kirtman, B. P. & Wu, Z. Increasing flooding hazard in coastal communities due to rising sea level: case study of Miami Beach, Florida. *Ocean Coast. Manag.* **126**, 1–8 (2016).
19. Yin, J., Schlesinger, M. E. & Stouffer, R. J. Model projections of rapid sea-level rise on the northeast coast of the United States. *Nat. Geosci.* **2**, 262–266 (2009).
20. Yin, J., Griffies, S. M. & Stouffer, R. J. Spatial variability of sea level rise in twenty-first century projections. *J. Clim.* **23**, 4585–4607 (2010).
21. Valle-Levinson, A., Dutton, A. & Martin, J. B. Spatial and temporal variability of sea level rise hot spots over the eastern United States. *Geophys. Res. Lett.* **44**, 7876–7882 (2017).
22. Dangendorf, S. et al. Acceleration of US. Southeast and Gulf coast sea-level rise amplified by internal climate variability. *Nat. Commun.* **14**, 1935 (2023).
23. Zhang, L. et al. Causes and multiyear predictability of the rapid acceleration of US. Southeast Sea level rise after 2010. *npj Clim. Atmos. Sci.* **7**, 113 (2024).
24. Ezer, T., Atkinson, L. P., Corlett, W. B. & Blanco, J. L. Gulf Stream's induced sea level rise and variability along the US mid-Atlantic coast. *J. Geophys. Res. Oceans* **118**, 685–697 (2013).
25. Ezer, T. Detecting changes in the transport of the Gulf Stream and the Atlantic overturning circulation from coastal sea level data: The extreme decline in 2009-2010 and estimated variations for 1935-2012. *Glob. Planet. Change* **129**, 23–36 (2015).
26. Piecuch, C. G. & Ponte, R. M. Inverted barometer contributions to recent sea level changes along the northeast coast of North America. *Geophys. Res. Lett.* **42**, 5918–5925 (2015).
27. Goddard, P. B., Yin, J., Griffies, S. M. & Zhang, S. An extreme event of sea-level rise along the Northeast coast of North America in 2009-2010. *Nat. Commun.* **6**, 1–9 (2015).
28. Piecuch, C. G., Dangendorf, S., Ponte, R. M. & Marcos, M. Annual sea level changes on the North American northeast coast: Influence of local winds and barotropic motions. *J. Clim.* **29**, 4801–4816 (2016).
29. Zhu, Y., Han, W. & Alexander, M. A. Nonstationary roles of regional forcings in driving low-frequency sea level variability along the US. East coast since the 1950s. *Geophys. Res. Lett.* **50**, (2023).
30. Zhu, Y., Han, W., Alexander, M. A. & Shin, S. I. Interannual sea level variability along the U.S. East coast during the satellite altimetry era: local versus remote forcing. *J. Clim.* **37**, 21–39 (2024).
31. Domingues, R., Goni, G., Baringer, M. & Volkov, D. What caused the accelerated sea level changes along the U.S. east coast during 2010–2015? *Geophys. Res. Lett.* **45**, 13,367–13,376 (2018).

32. Volkov, D. L., Lee, S. K., Domingues, R., Zhang, H. & Goes, M. Interannual sea level variability along the Southeastern seaboard of the United States in relation to the gyre-scale heat divergence in the North Atlantic. *Geophys. Res. Lett.* **46**, 7481–7490 (2019).

33. Ezer, T. Regional differences in sea level rise between the Mid-Atlantic Bight and the South Atlantic Bight: is the gulf stream to blame? *Earth's Future* **7**, 771–783 (2019).

34. Kohonen, T. Self-organized formation of topologically correct feature maps. *Biol. Cybernetics* **43**, 59–69 (1982).

35. Johnson, N. C., Feldstein, S. B. & Tremblay, B. The continuum of northern hemisphere teleconnection patterns and a description of the NAO shift with the use of self-organizing maps. *J. Clim.* **21**, 6354–6371 (2008).

36. Hewitson, B. C. & Crane, R. G. Self-orginizing maps, application to synoptic climatology. *Clim. Res.* **22**, 13–26 (2002).

37. Sheridan, S. C. & Lee, C. C. The self-organizing map in synoptic climatological research. *Prog. Phys. Geogr.* **35**, 109–119 (2011).

38. Gervais, M., Atallah, E., Gyakum, J. R. & Bruno Tremblay, L. Arctic air masses in a warming world. *J. Clim.* **29**, 2359–2373 (2016).

39. Gervais, M., Shaman, J. & Kushnir, Y. Impact of the North Atlantic warming hole on sensible weather. *J. Clim.* **33**, 4255–4271 (2020).

40. Horvath, S., Stroeve, J., Rajagopalan, B. & Jahn, A. Arctic sea ice melt onset favored by an atmospheric pressure pattern reminiscent of the North American-Eurasian Arctic pattern. *Clim. Dyn.* **57**, 1771–1787 (2021).

41. Reusch, D. B., Alley, R. B. & Hewitson, B. C. North Atlantic climate variability from a self-organizing map perspective. *J. Geophys. Res. Atmos.* **112**, (2007).

42. Li, X., Li, C., Ling, J. & Tan, Y. The relationship between contiguous El Niño and La Niña revealed by self-organizing maps. *J. Clim.* **28**, 8118–8134 (2015).

43. Leloup, J. A., Lachkar, Z., Boulanger, J. P. & Thiria, S. Detecting decadal changes in ENSO using neural networks. *Clim. Dyn.* **28**, 147–162 (2007).

44. Morioka, Y., Tozuka, T. & Yamagata, T. Climate variability in the southern Indian Ocean as revealed by self-organizing maps. *Clim. Dyn.* **35**, 1075–1088 (2010).

45. Gu, Q. & Gervais, M. Exploring North Atlantic and north pacific decadal climate prediction using self-organizing maps. *J. Clim.* **34**, 123–141 (2021).

46. Gu, Q. & Gervais, M. Diagnosing two-way coupling in decadal North Atlantic SST variability using time-evolving self-organizing maps. *Geophys. Res. Lett.* **49**, 1–13 (2022).

47. Gu, Q. et al. Investigating atmospheric responses to and mechanisms governing North Atlantic sea surface temperatures over 10-year periods. *J. Clim.* **36**, 8601–8618 (2023).

48. Delworth, T. L. et al. SPEAR: the next generation GFDL modeling system for seasonal to multidecadal prediction and projection. *J. Adv. Model. Earth Syst.* **12**, 1–36 (2020).

49. Ding, H., Newman, M., Alexander, M. A. & Wittenberg, A. T. Skillful climate forecasts of the tropical indo-Pacific ocean using model-analogs. *J. Clim.* **31**, 5437–5459 (2018).

50. Ding, H., Newman, M., Alexander, M. A. & Wittenberg, A. T. Diagnosing secular variations in retrospective ENSO seasonal forecast skill using CMIP5 model-analogs. *Geophys. Res. Lett.* **46**, 1721–1730 (2019).

51. Lou, J., Newman, M. & Hoell, A. Multi-decadal variation of ENSO forecast skill since the late 1800s. *npj Clim. Atmos. Sci.* **6**, 89 (2023).

52. Zhang, L. et al. Skillful decadal prediction skill over the Southern Ocean based on GFDL SPEAR Model-Analogs. *Environ. Res. Commun.* **5**, (2023).

53. Zhang, R. Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophys. Res. Lett.* **35**, 1–6 (2008).

54. Zhang, R. Latitudinal dependence of Atlantic meridional overturning circulation (AMOC) variations. *Geophys. Res. Lett.* **37**, 1–6 (2010).

55. Srokosz, M. A. & Bryden, H. L. Observing the atlantic meridional overturning circulation yields a decade of inevitable surprises. *Science* **348**, 1255575 (2015).

56. Larson, S. M., Buckley, M. W. & Clement, A. C. Extracting the buoyancy-driven atlantic meridional overturning circulation. *J. Clim.* **33**, 4697–4714 (2020).

57. Visbeck, M. et al. The ocean's response to North Atlantic oscillation variability. *Geophys. Monogr. Ser.* **134**, 113–145 (2003).

58. Alexander, M. A., Timlin, M. S. & Scott, J. D. Winter-to-winter recurrence of sea surface temperature, salinity and mixed layer depth anomalies. *Prog. Oceanogr.* **49**, 41–61 (2001).

59. Deser, C., Alexander, M. A. & Timlin, M. S. Understanding the persistence of sea surface temperature anomalies in midlatitudes. *J. Clim.* **16**, 57–72 (2003).

60. Peings, Y. & Magnusdottir, G. Forcing of the wintertime atmospheric circulation by the multidecadal fluctuations of the North Atlantic ocean. *Environ. Res. Lett.* **9**, (2014).

61. Peings, Y. & Magnusdottir, G. Wintertime atmospheric response to Atlantic multidecadal variability: effect of stratospheric representation and ocean–atmosphere coupling. *Clim. Dyn.* **47**, 1029–1047 (2016).

62. Barnes, E. A. & Hartmann, D. L. Dynamical feedbacks and the persistence of the NAO. *J. Atmos. Sci.* **67**, 851–865 (2010).

63. Gervais, M., Shaman, J. & Kushnir, Y. Impacts of the North Atlantic warming hole in future climate projections: mean atmospheric circulation and the North Atlantic jet. *J. Clim.* **32**, 2673–2689 (2019).

64. Athanasiadis, P. J. et al. Decadal predictability of North Atlantic blocking and the NAO. *npj Clim. Atmos. Sci.* **3**, 20 (2020).

65. O'Brien, J. P. & Deser, C. Quantifying and understanding forced changes to unforced modes of atmospheric circulation variability over the North Pacific in a coupled model large ensemble. *J. Clim.* **36**, 19–37 (2023).

66. Gu, Q. et al. Wide range of possible trajectories of North Atlantic climate in a warming world. *Nat. Commun.* **15**, 4221 (2024).

67. Årthun, M., Wills, R. C. J., Johnson, H. L., Chafik, L. & Langehaug, H. R. Mechanisms of decadal North Atlantic climate variability and implications for the recent cold anomaly. *J. Clim.* **34**, 3421–3439 (2021).

68. Nnamchi, H. C. et al. Pan-Atlantic decadal climate oscillation linked to ocean circulation. *Commun. Earth Environ.* **4**, 1–9 (2023).

69. Bingham, R. J. & Hughes, C. W. Signature of the Atlantic meridional overturning circulation in sea level along the east coast of North America. *Geophys. Res. Lett.* **36**, 1–5 (2009).

70. Yang, X. et al. On the development of GFDL's decadal prediction system: initialization approaches and retrospective forecast assessment. *J. Adv. Model. Earth Syst.* **13**, 1–30 (2021).

71. Zhang, H. & Wu, L. Predicting North Atlantic sea surface temperature variability on the basis of the first-mode baroclinic Rossby wave model. *J. Geophys. Res. Oceans* **115**, 1–15 (2010).

72. Famooss Paolini, L. et al. Non-stationarity in the NAO–Gulf Stream SST Front Interaction. *J. Clim.* **37**, 1629–1650 (2024).

73. Meehl, G. et al. Global climate projections. In: *Climate Change* vol. 3495 (2007).

74. Drijfhout, S., van Oldenborgh, G. J. & Cimatoribus, A. Is a decline of AMOC causing the warming hole above the North Atlantic in observed and modeled warming patterns? *J. Clim.* **25**, 8373–8379 (2012).

75. Woollings, T., Gregory, J. M., Pinto, J. G., Reyers, M. & Brayshaw, D. J. Response of the North Atlantic storm track to climate change shaped by ocean-atmosphere coupling. *Nat. Geosci.* **5**, 313–317 (2012).

76. Marshall, J. et al. The ocean's role in the transient response of climate to abrupt greenhouse gas forcing. *Clim. Dyn.* **44**, 2287–2299 (2015).

77. Gervais, M., Shaman, J. & Kushnir, Y. Mechanisms governing the development of the North Atlantic warming hole in the CESM-LE future climate simulations. *J. Clim.* **31**, 5927–5946 (2018).

78. Caesar, L., McCarthy, G. D., Thornalley, D. J. R., Cahill, N. & Rahmstorf, S. Current atlantic meridional overturning circulation weakest in last millennium. *Nat. Geosci.* **14**, 118–120 (2021).

79. Kilbourne, K. H. et al. Atlantic circulation change still uncertain. *Nat. Geosci.* **15**, 165–167 (2022).

80. Caesar, L., McCarthy, G. D., Thornalley, D. J. R., Cahill, N. & Rahmstorf, S. Reply to: Atlantic circulation change still uncertain. *Nat. Geosci.* **15**, 168–170 (2022).

81. Adcroft, A. et al. The GFDL global ocean and sea ice model OM4.0: model description and simulation features. *J. Adv. Model. Earth Syst.* **11**, 3167–3211 (2019).

82. Zhao, M. et al. The GFDL global atmosphere and land model AM4.0/LM4.0: 1. simulation characteristics with prescribed SSTs. *J. Adv. Model. Earth Syst.* **10**, 691–734 (2018).

83. Zhao, M. et al. The GFDL global atmosphere and land model AM4.0/LM4.0: 2. model description, sensitivity studies, and tuning strategies. *J. Adv. Model. Earth Syst.* **10**, 735–769 (2018).

84. Riahi, K. et al. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Glob. Environ. Change* **42**, 153–168 (2017).

85. Kriegler, E. et al. Fossil-fueled development (SSP5): an energy and resource intensive scenario for the 21st century. *Glob. Environ. Change* **42**, 297–315 (2017).

86. Kobayashi, S. et al. The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteorol. Soc. Jpn.* **93**, 5–48 (2015).

87. Huang, B. et al. Extended reconstructed Sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J. Clim.* **30**, 8179–8205 (2017).

88. Taburet, G. et al. DUACS DT2018: 25 years of reprocessed sea level altimetry products. *Ocean Sci.* **15**, 1207–1224 (2019).

89. Holgate, S. J. et al. New data systems and products at the permanent service for mean sea level. *J. Coast. Res.* **29**, 493–504 (2013).

90. MacQueen, J. Some methods for classification and analysis of multivariate observations. In: *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability,* vol. 1, pp 281–296 (1967).

91. Reusch, D. B., Alley, R. B. & Hewitson, B. C. Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. *Polar Geogr.* **29**, 188–212 (2005).

92. Wills, R. C., Battisti, D. S., Hartmann, D. L. & Schneider, T. Extracting modes of variability and change from climate model ensembles. *In:* Proceedings of the 7th International Workshop on Climate Informatics: CI, 25–28 (2017).

93. Kohonen, T. *Self-organizing maps.* (Springer-Verlag, 2001).

94. Liu, Y., Weisberg, R. H. & Mooers, C. N. K. Performance evaluation of the self-organizing map for feature extraction. *J. Geophys. Res. Oceans* **111**, (2006).

95. Kohonen, T. K., Hynninen, J., Kangas, J. & Laaksonen, J. SOM PAK: the self-organizing map program package. *Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science* (1996).

96. Livezey, R. E. & Chen, W. Y. Statistical field significance and its determination by Monte Carlo techniques. *Mon Weather Rev.* **111**, 46–59 (1983).

97. Wilks, D. S. On 'field significance' and the false discovery rate. *J. Appl. Meteorol. Climatol.* **45**, 1181–1189 (2006).

98. Levine, R. A. & Wilks, D. S. Statistical methods in the atmospheric sciences. *J. Am. Stat. Assoc*. 95 (Academic Press, 2000).

99. Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M. & Bladé, I. The effective number of spatial degrees of freedom of a time-varying field. *J. Clim.* **12**, 1990–2009 (1999).

100. Yang, X. et al. A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *J. Clim.* **26**, 650–661 (2013).

101. Trenberth, K. E. Some effects of finite sample size and persistence on meteorological statistics. Part I: autocorrelations. *Mon. Weather Rev.* **112**, 2359–2368 (1984).

Author contributions
L.Z. and Q.G. conceived the initial idea. Q.G. designed the study, performed the analysis, and wrote the manuscript with input from L.Z. and L.J. T.L.D., X.Y., F.Z., and L.Z. led the development of the SPEAR decadal prediction system. T.L.D. and W.F.C. conducted the piControl and large ensemble simulations. All authors, including Q.G., L.Z., L.J., T.L.D., X.Y., F.Z., W.F.C., and S.L., contributed to the improvement of the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-024-00802-2.

Correspondence and requests for materials should be addressed to Qinxue Gu.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.