

Earth's Future

RESEARCH ARTICLE

10.1029/2022EF003211

Special Section:

CMIP6: Trends, Interactions, Evaluation, and Impacts

Key Points:

- Despite resolution, CMIP6 models struggle to reproduce amplified synoptic patterns over the US, with a robust bias toward zonal patterns
- Most CMIP6 models produce too few cyclones over the northeastern US, while model errors are more varied when considering cyclone intensity
- Increased resolution increases the number and intensity of simulated extratropical cyclones but does not improve model representation of synoptic patterns

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. J. Gore,
mjj6459@psu.edu

Citation:

Gore, M. J., Zarzycki, C. M., & Gervais, M. M. (2023). Connecting large-scale meteorological patterns to extratropical cyclones in CMIP6 climate models using self-organizing maps. *Earth's Future*, 11, e2022EF003211. <https://doi.org/10.1029/2022EF003211>

Received 16 SEP 2022

Accepted 17 JUL 2023

Connecting Large-Scale Meteorological Patterns to Extratropical Cyclones in CMIP6 Climate Models Using Self-Organizing Maps

Michelle J. Gore¹ , Colin M. Zarzycki¹ , and Melissa M. Gervais^{1,2} 

¹Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, PA, USA, ²The Institute for Computational and Data Sciences, The Pennsylvania State University, University Park, PA, USA

Abstract Extratropical cyclones (ETCs) are responsible for the majority of cool-season extreme events in the northeastern United States (NEUS), often leading to high-impact weather conditions that can have wide-ranging socioeconomic impacts. Evaluating the ability of climate models to adequately simulate ETC dynamics is essential for improving model performance and increasing confidence in future projections used by stakeholders and policymakers. ETCs are traditionally studied using techniques such as case studies and synoptic typing, however, these approaches can be time-consuming, require subjective analysis, and do not necessarily identify the coincident large-scale meteorological patterns (LSMPs). Here, we apply self-organizing maps (SOMs) as an automated machine-learning approach to characterize the LSMPs and associated frequency and intensity of discrete ETC events over NEUS. The dominant patterns of geopotential height variability are identified through SOM analysis of five reanalysis products during the last four decades. ETC events are tracked using TempestExtremes and are integrated with SOMs to classify the accumulated cyclone activity (ACA) associated with each pattern. We then evaluate the skill of CMIP6 historical experiments in simulating the LSMPs and ETC events identified in the SOM. Our results identify a robust bias toward more zonal patterns, with models struggling to reproduce the more amplified patterns typically associated with the highest cyclone activity. While model resolution has some impact on simulation credibility, model configuration appears to be more important in LSMP representation. The vast majority of CMIP6 models produce too few ETCs, although model errors are distributed around historical reanalyses when ACA is normalized by storm frequency.

Plain Language Summary Winter storms can have devastating socioeconomic impacts across the United States. Climate models are used to understand and predict these winter storms and how they might change in the future, therefore helping to inform climate policy and emergency services. It is essential that we evaluate these climate models so that their performance and accuracy can be optimized. One way of evaluating climate models is to assess their ability in reproducing the large-scale atmospheric conditions which occur during extreme events. Here we apply a cyclone tracking algorithm and self-organizing maps, a machine-learning approach, to automate the process of identifying key patterns associated with historical winter storm activity. We then use this approach to assess if CMIP6 climate models are able to reproduce the same patterns including the winter storm frequency and intensity. Our results show that, regardless of resolution, most CMIP6 models struggle to simulate the more extreme patterns and tend to favor weaker patterns that are not as conducive to cyclone formation. As a result, most models simulate too few winter storms, however, the error in storm intensity is more varied, with some models producing stronger storms and others producing weaker storms on average.

1. Introduction

Extratropical cyclones (ETCs) are a key feature of mid-latitude circulation and are vital to the distribution of heat and moisture outside of the tropics. They are particularly important to the regional climate of the northeastern United States (NEUS), with over 80% of winter precipitation in the region attributed to ETCs (Hawcroft et al., 2012). Furthermore, ETCs are responsible for the majority of cool-season extreme events in NEUS (Agel et al., 2015; Pfahl & Wernli, 2012), often leading to high-impact weather conditions such as strong winds, storm surge, flooding, heavy precipitation, and snowstorms (Zielinski, 2002). These can cause significant

© 2023 The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

damage with wide-ranging socioeconomic impacts, making it necessary to accurately predict ETC events and their variability.

While extreme events, such as those related to ETCs, are primarily driven by local and mesoscale processes, they are typically organized, enhanced, and/or triggered by distinct large-scale meteorological patterns (LSMPs, Barlow et al., 2019). LSMPs are recurrent patterns defined relative to the occurrence of a specific event (e.g., extreme precipitation, heat waves, ETCs), with a focus on the synoptic scale dynamics that produce favorable conditions for these events. Large-scale synoptic conditions are better represented in climate models and are therefore more predictable and useful in forecasting extremes (Barlow et al., 2019). It is worth noting that, while many studies have identified teleconnection patterns associated with ETCs (e.g., Seierstad et al., 2007), LSMPs have been defined as distinct from climate modes of variability, in terms of frequency and duration (Grotjahn et al., 2016). However, that is not to say that there is no relationship or modulating effect between climate modes and LSMPs (Barlow et al., 2019).

Grotjahn et al. (2016) describe various approaches to identifying LSMPs associated with extreme events, namely case studies, compositing, or pattern-based analysis of circulation or thermodynamic fields during these events. Many studies have identified the LSMPs associated with ETCs using the first two approaches: case studies (e.g., Mass & Dotson, 2010) and compositing based on manual synoptic typing (e.g., Davis et al., 1993; Kocin & Uccellini, 2004; Yoshida & Asuma, 2004). However, these approaches can be time-consuming, especially when dealing with large data sets, and require subjective analysis (Barlow et al., 2019). Pattern-based analysis methods have become increasingly popular for identifying LSMPs associated with extremes as they offer a more efficient and objective approach to manual synoptic typing. A review of some of these methods is contained in Grotjahn et al. (2016), including regression, empirical orthogonal function or principal component analysis, cluster analysis, and self-organizing maps (SOMs), discussing the advantages and limitations of each. Many studies have implemented these methods to identify the LSMPs associated with extreme events (e.g., temperature and precipitation) and weather phenomena. For example, Agel et al. (2018) applied k-means clustering (KMC) and SOMs to tropopause height to identify the LSMPs associated with extreme precipitation over NEUS. They found that extreme precipitation days were typically accompanied by a collocated deep upper-level trough and pronounced ridge downstream. Similarly, Sukhdeo et al. (2022) applied linear orthogonal decomposition to identify the large-scale atmospheric drivers of precipitation over the NEUS. While they did tie ETC activity into their analysis, the focus is primarily on the dynamics behind monthly precipitation variability.

While other clustering techniques have been extensively used to classify circulation patterns (e.g., Catalano & Broccoli, 2018), SOMs have been of growing interest in climatology applications. SOMs are a type of artificial neural network which allow for the nonlinear unsupervised classification of multivariate data into a finite set of representative patterns (Brereton, 2012; Hewitson & Crane, 2002; Kohonen, 2013; Reusch et al., 2005; Shieh & Liao, 2012). They have been effectively used in various climate studies identifying LSMPs associated with extreme events (such as temperature and winds (Cassano et al., 2006), and extreme precipitation (Agel et al., 2018; Cavazos, 2000)). SOMs have also been successfully applied as a tool for climate model evaluation (e.g., Bu et al., 2022; Herbst & Casper, 2008; Sheridan & Lee, 2010; Taylor et al., 2022), climate change studies (e.g., Gervais et al., 2016; Gervais et al., 2020), and decadal climate variability and predictability (e.g., Gu & Gervais, 2021; Gu & Gervais, 2022). The key advantage of SOMs over other clustering techniques is that it is able to capture nonlinear features as it does not assume stationarity or orthogonality, producing more realistic results which are easily interpretable (Gervais et al., 2020; Hewitson & Crane, 2002; Van Hulle, 2012). Furthermore, SOMs are able to represent the data as a continuum of patterns, rather than discrete realizations (Hewitson & Crane, 2002; Reusch et al., 2005). While previous studies have effectively used SOMs to study the synoptic patterns over the US (e.g., Agel et al., 2018; Loikith et al., 2017; Swales et al., 2016), they have typically focused on extreme temperature and precipitation events. Here, we aim to integrate discrete cyclone tracking with the SOM method to identify the synoptic evolution of ETCs over the NEUS.

Climate models have become a useful tool for analyzing present and future synoptic conditions associated with ETCs, providing crucial information to stakeholders and policymakers (Terando et al., 2020). Therefore, evaluating the ability of climate models to adequately simulate ETC dynamics is essential for improving model performance and increasing confidence in future projections. Previous studies have used data from the Coupled Model Intercomparison Project (CMIP) to investigate the mean synoptic conditions associated with ETCs. Several studies found that the climate models in both phase 5 (CMIP5) and Phase 6 (CMIP6) frequently under-represent

blocking patterns in the northern hemisphere, resulting in an underestimation in ETC frequency and intensity with a storm track that is too zonal, particularly over the North Atlantic (Priestley et al., 2020; Zappa et al., 2013). Zappa et al. (2013) found that these biases are only weakly associated with errors in the jet stream and are more likely a direct result of subgrid parameterizations within the models. In addition, Priestley et al. (2020) focused on the ability of CMIP models to reproduce high-intensity, bomb cyclones, finding that most models struggled to capture the rapid intensification rates of these events, resulting in a notable decrease in their frequency. Generally, there has been an improvement in the representation of ETCs in CMIP6 compared to CMIP5, which has been mostly attributed to increased horizontal resolution (Harvey et al., 2020; Priestley et al., 2020). It is suggested that higher resolution (i.e., finer grid spacing) could be improving the representation of mean-flow interactions with orography and mesoscale moist processes within ETCs, leading to more realistic ETC cyclogenesis, intensification, and trajectories (Berckmans et al., 2013; Priestley et al., 2020). However, this is not always the case, as shown by J. Lee et al. (2022), where some high-resolution models still show notable ETC biases around regions of high elevation. Therefore, it is possible that other non-resolution factors, such as orographic drag parameterization (Pithan et al., 2016), play an important role in ETC representation. Despite these advances, there are still significant biases in the CMIP6 climate models associated with ETC dynamics which requires further investigation. While many CMIP6 evaluation studies only consider biases associated with the mean climatology, there is merit in applying a method like SOMs to understand the distribution of errors across different patterns of circulation (e.g., Bu et al., 2022; Taylor et al., 2022).

In this paper, SOMs are applied as an objective approach to characterize the LSMPs over the NEUS and the associated frequency and intensity of discrete ETC storm events using gridded climate data. Thereafter, the ability of CMIP6 climate models to simulate these LSMPs and associated NEUS ETCs is also investigated.

2. Data and Methods

2.1. Data

Synoptic conditions associated with ETCs are commonly characterized by variations in geopotential height and surface pressure. Therefore this study used 500-hPa geopotential height (Z500) over the boreal winter season (December, January, and February; DJF) from 1980 to 2019 to define the LSMPs over NEUS. In addition, 6-hourly sea level pressure (SLP) is used to track ETCs. Reanalysis products are frequently utilized when evaluating climate models due to their extensive spatial and temporal coverage (Fasullo, 2020; Priestley et al., 2020; Zappa et al., 2013). Due to subtle differences between different global reanalyses, five products are considered in this analysis: the European Center for Medium-range Weather Forecasting (ECMWF) fifth generation reanalysis (ERA5, Hersbach et al., 2020), the Japanese Meteorological Agency's (JMA) 55-year reanalysis (JRA-55, Kobayashi et al., 2015), the National Aeronautics and Space Administration's (NASA) Modern Era Retrospective Reanalysis for Research and Application version 2 (MERRA2, Koster et al., 2015), the National Oceanic and Atmospheric Administration's (NOAA) Climate Forecasting System Reanalysis (CFSR, Saha et al., 2010a, 2014), and the NOAA 20th Century Reanalysis version 3 (20CRv3; only available for 1980–2015, Slivinski et al., 2019b). Combining multiple reanalysis products ensures that limitations of individual products have a minimal impact on the LSMP analysis, thus producing a more robust result. When evaluating ETC dynamics, only one reanalysis product is used as the reference data set to simplify the analysis. ERA5 will be used as the reference data set for comparison due to its higher horizontal and vertical resolutions.

Phase 6 of the Coupled Model Intercomparison Project (CMIP6) provides a comprehensive multi-model framework for assessing the LSMPs associated with ETCs over NEUS. In this study, we consider historical simulations from 26 global climate model contributions to the core CMIP6 experiments (Eyring et al., 2016), including both standard (100 km) and low (250 km) resolution, and 8 simulations from the High Resolution Model Intercomparison Project (HighResMIP, Haarsma et al., 2016) CMIP6 experiments. Institution and reference information for each reanalysis product and climate model is provided in Table 1. These climate models are fully coupled to a dynamical ocean with interactive sea ice, and atmospheric forcings are based on observations. Table 2 summarizes the horizontal and vertical resolutions for the atmospheric and ocean components of each climate model. To compare all of the models equally and for consistency, only the first ensemble member (i.e., r1i1p1f1) from the historical experiments (“historical” for CMIP6 and “hist-1950” for HighResMIP) were used for each model with the exception of CESM2 which only had 6-hourly data available in the second ensemble member (r2i1p1f1) of the naturalized historical experiment (“hist-nat”). To be comparable to the reanalyses, analysis of the 6-hourly data for SLP and Z500 are restricted to 1980–2014.

Table 1

Reanalysis Products and CMIP6 Climate Models Used in This Study

Name	Institution	Reference
Reanalysis Products:		
20CRv3	NOAA; National Oceanic and Atmospheric Administration, United States	Slivinski et al. (2019b)
CFSR	NOAA; National Oceanic and Atmospheric Administration, United States	Saha et al. (2014)
ERA5	ECMWF; European Centre for Medium-Range Weather Forecasts, United Kingdom	Hersbach et al. (2020)
JRA-55	JMA; Japanese Meteorological Agency, Japan	Kobayashi et al. (2015)
MERRA2	NASA; National Aeronautics and Space Administration, United States	Koster et al. (2015)
CMIP6 Climate Models:		
ACCESS-CM2	CSIRO-ARCCSS; Commonwealth Scientific and Industrial Research Organisation, Australian Research Council Centre of Excellence for Climate System Science, Australia	Dix et al. (2019)
ACCESS-ESM1.5	CSIRO; Commonwealth Scientific and Industrial Research Organisation, Australia	Ziehn et al. (2019)
AWI-ESM-1-1-LR	AWI; Alfred Wegener Institute, Germany	Danek et al. (2020)
BCC-CSM2-HR	BCC; Beijing Climate Center, China	Jie et al. (2020)
BCC-CSM2-MR	BCC; Beijing Climate Center, China	Wu et al. (2018)
CESM2	NCAR; National Center for Atmospheric Research, Climate and Global Dynamics Laboratory, United States	Danabasoglu (2019)
CMCC-CM2-HR4	CMCC; Centro Euro-Mediterraneo per i Cambiamenti, Italy	Scoccimarro et al. (2020)
CMCC-CM2-SR5	CMCC; Centro Euro-Mediterraneo per i Cambiamenti, Italy	Lovato and Peano (2020)
CMCC-CM2-VHR4	CMCC; Centro Euro-Mediterraneo per i Cambiamenti, Italy	Scoccimarro et al. (2018)
E3SM-1-0-LR	DoE; Department of Energy, United States	Bader et al. (2019)
E3SM-1-0-HR	DoE; Department of Energy, United States	Caldwell et al. (2019)
EC-Earth3	EC-Earth-Consortium	EC-Earth (2019a)
EC-Earth3-Veg	EC-Earth-Consortium	EC-Earth (2019b)
EC-Earth3-Veg-LR	EC-Earth-Consortium	EC-Earth (2020b)
EC-Earth3-AerChem	EC-Earth-Consortium	EC-Earth (2020a)
ECMWF-IFS-HR	ECMWF; European Centre for Medium-Range Weather Forecasts, United Kingdom	Roberts et al. (2017)
ECMWF-IFS-MR	ECMWF; European Centre for Medium-Range Weather Forecasts, United Kingdom	Roberts et al. (2018b)
ECMWF-IFS-LR	ECMWF; European Centre for Medium-Range Weather Forecasts, United Kingdom	Roberts et al. (2018a)
GFDL-CM4	NOAA-GFDL; National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, United States	H. Guo et al. (2018)
GFDL-CM4C192	NOAA-GFDL; National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, United States	Zhao et al. (2018)
GISS-E2-1-G	NASA-GISS; Goddard Institute for Space Studies, United States	NASA/GISS (2018)
IPSL-CM6A-LR	IPSL; Institut Pierre Simon Laplace, France	Boucher et al. (2018)
KIOST-ESM	KIOST; Korea Institute of Ocean Science and Technology, Republic of Korea	Kim et al. (2019)
MIROC6	MIROC; Model for Interdisciplinary Research on Climate Consortium (JAMSTEC, AORI, NIES, R-CCS), Japan	Tatebe and Watanabe (2018)
MPI-ESM-1-2-HAM	HAMMOZ-Consortium	Neubauer et al. (2019)
MPI-ESM1-2-HR	MPI-M, DWD, DKRZ; Max Planck Institute for Meteorology, Deutscher Wetterdienst, Deutsches Klimarechenzentrum, Germany	Jungclaus et al. (2019)
MPI-ESM1-2-LR	MPI-M, AWI; Max Planck Institute for Meteorology, Alfred Wegener Institute, Germany	Wieners et al. (2019)
MPI-ESM1-2-XR	MPI-M; Max Planck Institute for Meteorology, Germany	von Storch et al. (2018)
MRI-ESM2-0	MRI; Meteorological Research Institute, Japan	Yukimoto et al. (2019)
NESM3	NUIST; Nanjing University of Information Science and Technology, China	Cao and Wang (2019)
NorESM2-LM	NCC; Norwegian Climate Centre, Norway	Seland et al. (2019)

Table 1
Continued

Name	Institution	Reference
NorESM2-MM	NCC; Norwegian Climate Centre, Norway	Bentsen et al. (2019)
SAM0-UNICON	SNU; Seoul National University, Republic of Korea	Park and Shin (2019)
TaiESM1	AS-RCEC; Academia Sinica Research Center for Environmental Changes, Taiwan	W.-L. Lee and Liang (2020)

2.2. ETC Tracking and Metrics

ETC events were tracked using TempestExtremes (Ullrich & Zarzycki, 2017; Ullrich et al., 2021; Zarzycki & Ullrich, 2017), an automated Lagrangian feature tracker commonly applied for extracting individual weather events from gridded climate data. Candidate ETCs are identified by finding local minima in the 6-hourly SLP field that are surrounded by a closed isobar within a 6° radius and at least 2 hPa greater than the minimum SLP. Trajectories are then stitched together by detecting consecutive candidate cyclones that lie within a 9° radius of each other and persist for at least 24 hr (Zarzycki, 2018). ETCs were tracked for all reanalysis products and CMIP6 climate models on their native grid resolutions. While some studies (e.g., Hoskins & Hodges, 2002; Zappa et al., 2013) use 850-hPa vorticity to track ETCs, Colle et al. (2013) found that tracking by SLP performed similarly to vorticity-tracking methods. A comprehensive comparison of different ETC tracking methodologies can be found in Neu et al. (2013).

From these trajectories, two metrics were calculated. “ETC hits” is the number of 6-hourly timesteps identified by the tracker as an ETC, used as an indicator of cyclone frequency. Further, “accumulated cyclone activity” (ACA, Y. Guo et al., 2017) combines both ETC frequency and mean ETC intensity. ACA is calculated as the sum of the pressure anomaly associated with all cyclones centers located within a 2° radius of each grid point and can be written as:

$$\text{ACA} = \sum (P_{\text{mean}} - P_{\text{etc}}) \quad (1)$$

where P_{mean} is the seasonal (DJF) mean pressure of the whole eastern US and P_{etc} is the minimum pressure of the ETC at that time step. P_{mean} is calculated separately for each climate model to account for differences in mean climatology, and varies by approximately 5-hPa across all products (shown in supporting information). The eastern US is defined as 24°N – 60°N and 60°W – 95°W as indicated in Figure 3. ACA is calculated at 6-hourly time steps, and therefore a single ETC can be counted multiple times and slow-moving ETCs can strongly influence the total ACA (Yau & Chang, 2020). Additionally, we calculated the percent contribution of each LSMP to total ACA, which is the sum of ACA in a given LSMP divided by the total winter ACA for the eastern US. A latitude weighting is applied to this calculation.

In this study, we tracked ETCs across the entire northern hemisphere, however, we are only interested in the stronger ETCs occurring in the NEUS. Therefore, after tracking, we filtered the storm tracks by only considering ETC days which occur over the eastern US (as defined above) and ETCs with a minimum pressure less than the 75th percentile of cyclone intensities. Therefore, in some cases, only part of the cyclone trajectories are used. Neu et al. (2013) found that discrepancies between tracking algorithms persisted with the detection of weaker storms (with more agreement when tracking stronger storms), therefore by filtering our ETCs we are slightly reducing the sensitivity to tracking algorithms. Additionally, ACA is weighted by storm intensity and will, therefore, further reduce any sensitivity to differences in frequency between methods. It is possible for an ETC to contribute negatively toward ACA (if a weak cyclone happens to have $P_{\text{etc}} > P_{\text{mean}}$) which is further motivation to filter the ETCs by an intensity threshold. The 75th percentile of cyclone intensity is calculated for each data set and is shown in supporting information ($\sim 1,005$ -hPa for ERA5).

2.3. Self-Organizing Maps

The application of SOMs is a versatile method that identifies a set of archetypal patterns within a data set. The SOM algorithm results in patterns being spatially ordered, with similar patterns closer together in the map, and allows for subtle variations to be distinguished (Gervais et al., 2016; Kohonen, 2013; Reusch et al., 2005). It effectively captures non-linear features as it does not assume stationarity or orthogonality, producing more realistic results compared to other clustering techniques (Gervais et al., 2020; Hewitson & Crane, 2002; Van Hulle, 2012). A detailed description of the SOM method can be found in Kohonen (2001).

Table 2
Resolution of Reanalysis Products and CMIP6 Climate Models Used in This Study

Name	Atmospheric resolution		Ocean resolution	
	Nominal resolution (km)	Vertical levels	Nominal resolution (km)	Vertical levels
Reanalysis Products:				
1 ERA5	30	137	—	—
2 JRA-55	55	60	—	—
3 MERRA2	50	72	—	—
4 CFSR	38	64	—	—
5 20CRv3	75	64	—	—
HighResMIP Models:				
6 BCC-CSM2-HR	50	56	50	40
7 CMCC-CM2-VHR4	25	26	25	50
8 E3SM-1-0-HR	25	72	8–16	80
9 ECMWF-IFS-HR	25	91	25	75
10 ECMWF-IFS-MR	50	91	25	75
11 ECMWF-IFS-LR	50	91	100	75
12 GFDL-CM4C192	25	33	25	75
13 MPI-ESM1-2-XR	50	95	50	40
CMIP Standard Resolution Models:				
14 BCC-CSM2-MR	100	46	50	40
15 CESM2	100	32	100	60
16 CMCC-CM2-HR4	100	30	25	50
17 CMCC-CM2-SR5	100	26	100	50
18 E3SM-1-0-LR	100	72	50	60
19 EC-Earth3	100	91	100	75
20 EC-Earth3-Veg	100	91	100	75
21 EC-Earth3-AerChem	100	91	100	75
22 GFDL-CM4	100	33	25	75
23 MPI-ESM1-2-HR	100	95	50	40
24 MRI-ESM2-0	100	80	100	61
25 NorESM2-MM	100	32	100	70
26 SAM0-UNICON	100	30	100	60
27 TaiESM1	100	30	100	60
CMIP Low Resolution Models:				
28 ACCESS-CM2	250	85	100	50
29 ACCESS-ESM1.5	250	85	100	50
30 AWI-ESM-1-1-LR	250	47	50	46
31 EC-Earth3-Veg-LR	250	62	100	75
32 GISS-E2-1-G	250	40	100	40
33 IPSL-CM6A-LR	250	79	100	75
34 KIOST-ESM	250	32	100	52
35 MIROC6	250	81	100	63
36 MPI-ESM-1-2-HAM	250	95	250	40
37 MPI-ESM1-2-LR	250	47	250	40

Table 2
Continued

Name	Atmospheric resolution		Ocean resolution	
	Nominal resolution (km)	Vertical levels	Nominal resolution (km)	Vertical levels
38 NESM3	250	47	100	46
39 NorESM2-LM	250	32	100	70

Note. hist-1950 experiment r1i1p1f1 used for all HighResMIP models and historical experiment r1i1p1f1 used for all CMIP models, except CESM2 where the hist-nat experiment r2i1p1f1 was used due to availability.

A SOM is made up of a predetermined number of patterns (nodes) in a two-dimensional grid, which are adjusted and modified as successive random samples of data are processed by the SOM algorithm, eventually resulting in a stabilized map that is representative of the data (Brereton, 2012; Hewitson & Crane, 2002; Reusch et al., 2005). The SOM nodes are randomly initialized with data vectors and then the training algorithm consists of a competitive stage and a cooperative stage (Van Hulle, 2012). During the competitive stage, the similarity between a random sample vector and each node is measured by Euclidean distance (Brereton, 2012; Hewitson & Crane, 2002; Kohonen, 2013; Reusch et al., 2005). The node which is closest to that sample is the “winning” node, also known as the best match unit (BMU). During the cooperative stage, the BMU and surrounding nodes are then adjusted according to a defined learning rate, neighborhood function, and radius of influence to reduce the variation between the SOM nodes and the input vector (Brereton, 2012; Hewitson & Crane, 2002; Van Hulle, 2012). This process can be represented as follows:

$$m_i(t+1) = m_i(t) + \alpha(t)h_{ci}(t)[x(t) - m_i(t)] \quad (2)$$

where $m_i(t)$ is the vector associated with the winning node i at training time step t , $\alpha(t)$ is the learning rate, $h_{ci}(t)$ is the neighborhood function, and $x(t)$ is the input data vector. The learning rate determines the degree of adjustment (Hewitson & Crane, 2002), which is defined as the inverse function of the training time. The neighborhood function determines the influence that the input data vector has on the BMU and surrounding nodes at each time step (Brereton, 2012; Van Hulle, 2012). In this study, the Epanechnikov neighborhood function is used, as it has been shown to perform better than other neighborhood functions. Liu et al. (2006) compared four different neighborhood functions and found that the Epanechnikov neighborhood function produced the smallest SOM errors (indicating that it is well organized) and had less smoothing than other functions, resulting in a more accurate and detailed map. The Epanechnikov neighborhood function can be represented as follows:

$$h_{ci}(t) = \max\left(0, 1 - \frac{d_{ci}^2}{\sigma(t)^2}\right) \quad (3)$$

where d_{ci} is the distance between the BMU c and each node i , and $\sigma(t)$ is the radius of influence. In this case, only nodes within the radius of influence are modified with the majority of the modification focused on the BMU with smaller modifications as you move further away. This process results in the nodes being spatially ordered, with similar nodes closer together and different nodes further apart (Gervais et al., 2016; Kohonen, 2013; Reusch et al., 2005).

The nodes need to be sufficiently modified to ensure that they are well organized and stable. Therefore, the SOM algorithm is repeatedly iterated to ensure that each sample is chosen numerous times (Brereton, 2012; Kohonen, 2013). The training length multiple determines the number of times each sample is selected during the training process. With increasing iterations, the radius of influence and the learning rate decrease, until the SOM nodes are sufficiently modified and stabilize (Brereton, 2012; Hewitson & Crane, 2002). Once the training is complete, the data is compared to the SOM nodes once more (without any adjustment of the nodes) to determine the final BMUs for each data point. This enables composite patterns of other variables to be created for each node (Gervais et al., 2020).

In this study, a SOM was used to identify 20 dominant patterns of the NEUS winter climatology and associated ETCs. The SOM was trained on Z500 over the winter season (DJF) from 1980 to 2019, which will henceforth be referred to as the “master SOM.” Since we are interested in the large-scale synoptics, the SOM was trained over the entire continental US (20°N – 60°N and 60°W – 130°W). All five reanalysis products (ERA5, JRA-55, MERRA2, CFSR, and 20CRv3) were included to maximize the training data set (196 years total). Due to computational restrictions, the reanalysis products were regressed to a 1° spatial resolution and the temporal resolution was reduced from 6-hourly data to daily data by selecting the 12Z analysis for each calendar day. In addition, the data was normalized and multiplied by the cosine

Table 3
Summary of the Self-Organizing Map Parameters Used in This Study

Grid size	4 × 5
Initialization	Random
Training Periods	Two
Training Length Multiple	Ten
Radius of Influence	5 (T1); 2 (T2)
Learning Rate	0.1 (T1); 0.01 (T2)
Neighborhood Function	Epanechnikov

Note. T1 and T2 represent values chosen for first and second training periods respectively.

of the latitude to ensure equal area weighting within the SOM algorithm. Various SOM sizes were tested (shown in Supporting Information S1) and the 4×5 grid was selected as the optimal size which maximized pattern variance, whilst maintaining significance between patterns. Although the SOM algorithm consists of several user-defined parameters, if it is well-trained, these parameters (with the exception of size) have a negligible influence on the final result (Gervais et al., 2020; Gibson et al., 2017; Grotjahn et al., 2016). The effectiveness of the SOM training can be determined by the Sammon map, quantization error (QE), and topographic error (TE). A Sammon map is an approximation of the Euclidean distance between the SOM nodes, indicating their similarity. QE is the average error for a SOM node relative to the input data, while TE is the percentage of data whose second BMU is not a neighboring node. A well-trained SOM will have a “flat” Sammon map and low error (in this case QE = 40.2 and TE = 0.09). These metrics (shown in Supporting Information S1) were used to optimize the parameters and the SOM size chosen for this study, specified in Table 3.

Once the master SOM was created, the BMUs were used to identify the ETC days and ACA associated with each node. A schematic of this step is shown in Figure 1a which shows how ETC tracking and SOMs were combined to create the master SOM patterns and identify the LSMPs. The percent of total days and ETC hits that fall within each node was calculated to indicate the frequency of each pattern and its associated ETCs. Additionally, the contribution of each node to the overall ACA was determined by comparing the ACA in each pattern to the total winter ACA. The significance of node frequency, ETC frequency, and ACA contribution was tested using a permutation test (Wilks, 2011) to ensure that they are significantly different from random activity and can be classified as LSMPs. For testing node frequency significance, a null distribution was established by assigning each day a new BMU (random number between 1 and 20) and calculating the resampled node frequency, repeating this 1,000 times. The actual node frequency was compared to this null distribution and any value outside the 5th (2.5th) and 95th (97.5th) percentiles was considered significant at the 10% (5%) level. Similarly, to test the significance of ETC frequency and ACA contribution of each SOM pattern, ETC hits were randomly distributed between all winter days. A null distribution was established by calculating the ETC frequency and ACA contribution for each pattern from this resampled data and repeating 1,000 times.

2.4. CMIP6 Application

In order to evaluate the ability of climate models to reproduce these SOM patterns, the CMIP6 Z500 data for each model was organized according to its similarity with the patterns in the master SOM. For each data vector (i.e., time step) in the model, the similarity with the master SOM nodes was measured by Euclidean distance and a BMU was determined for each day in the model data. This enabled composite patterns of Z500 anomalies, node frequency, ETC frequency, and ACA to be calculated for each CMIP6 model, creating a pseudo-SOM based on the master SOM patterns. Please note that the CMIP6 data is simply composited based on the master SOM patterns, and does not influence the actual master SOM training or patterns, similar to the analysis done in Bu et al. (2022) and Taylor et al. (2022). Figure 1b summarizes this step, showing the classification of climate model data into the patterns identified by the master SOM in Figure 1a. Model errors were defined by differences between the CMIP6 patterns and the reanalysis SOM patterns. Four different types of error were analyzed, including node frequency error, spatial error, ETC frequency error, and ETC intensity error. Node frequency errors are calculated as the difference between the percentage of occurrences of a given node in the model composite relative to the percentage of occurrences of that node in the master SOM, giving an indication of discrepancies in the pattern frequency. Similarly, spatial errors are determined by root mean square error (RMSE) which is calculated by comparing the spatial map of a model composite to the corresponding spatial map in the master SOM. The mean climatology is removed and latitude weighting is applied to the RMSE calculation for more accurate interpretation. ETC frequency and intensity errors are calculated in a similar manner. ETC frequency error is the difference between the model composite and the master SOM for the average number of ETC hits per day within a given node. Lastly, ETC intensity error is defined as the difference between the model composite and the master SOM for average ACA per ETC hit. The ETC statistics are normalized relative to ERA5 in order to remove the mean bias. A detailed explanation of these calculations and the normalization method can be found in supporting information.

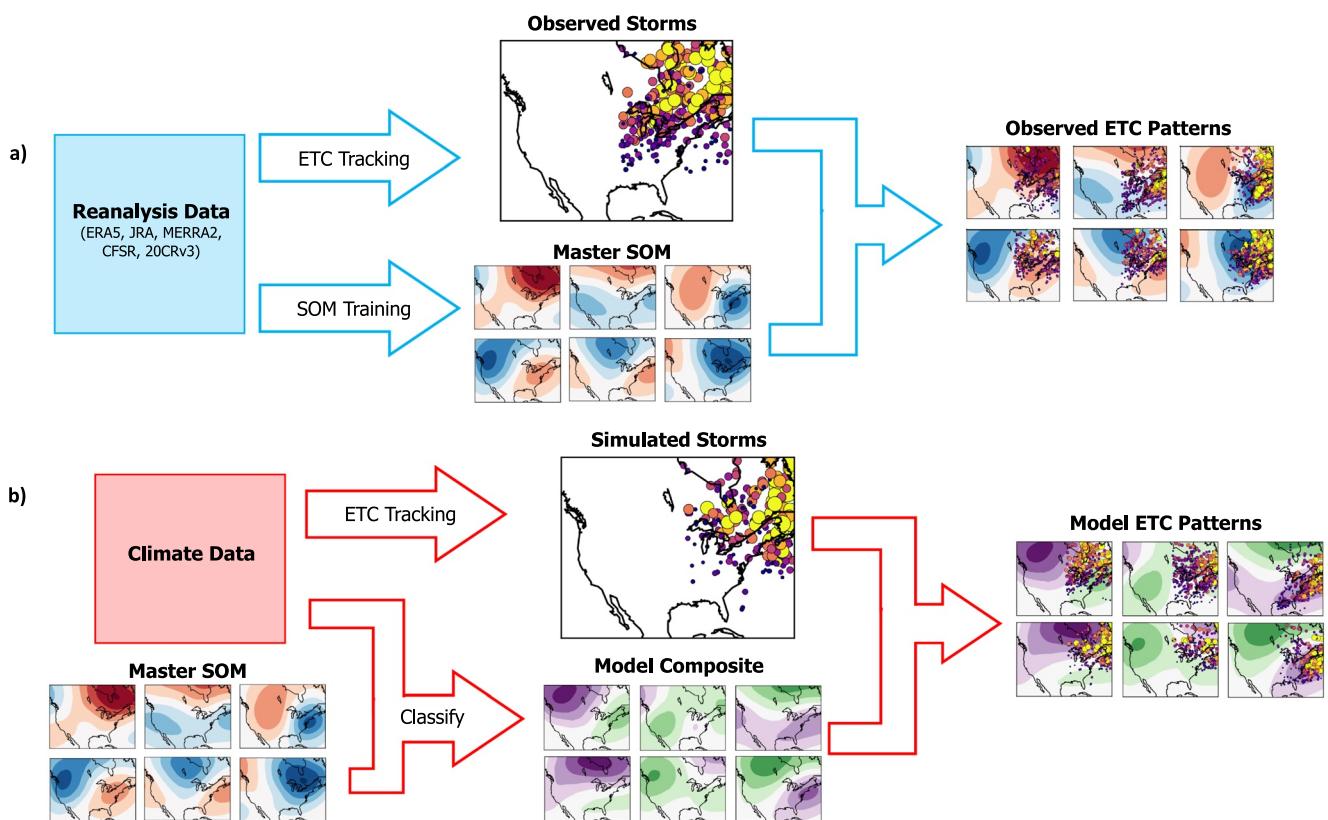


Figure 1. Schematic showing (a) the application of self-organizing maps (SOMs) and extratropical cyclone tracking to produce the master SOM patterns and (b) the classification and organization of climate data into patterns based on the master SOM.

Significance testing was conducted using a permutation test, which has been effectively used in other SOM analyses (e.g., Gervais et al., 2020; Gu & Gervais, 2021). For each CMIP6 model, the BMU data from the reanalysis SOM and model were combined and randomly separated into two new sets of data with the same number of data points as the original (Wilks, 2011). The RMSE and differences in node frequency, ETC frequency, and ACA between the two sets of resampled data were then computed for each SOM node. The number of days per node and the number of ETC hits per node was kept consistent when resampling for ETC frequency and ACA, respectively. This was repeated 1,000 times to establish a null distribution for each error type. If the actual difference in node frequency, ETC frequency, or ACA between the model and the master SOM lies outside the 5th (2.5th) and 95th (97.5th) percentiles of the null distribution, then the difference is regarded as significant at the 10% (5%) level. Similarly, the RMSE is regarded as significant at the 10% (5%) level if it lies outside the 90th (95th) percentile.

To calculate the CMIP6 and HighResMIP mean errors we concatenate all of the model BMUs for each group, respectively. We can then determine the frequency of each node pattern in the concatenated total CMIP6/HighResMIP data set (i.e., the mean CMIP6/HighResMIP node frequency). Similar to the individual models, we determine statistical significance through a permutation test, however this time we are permuting the concatenated model BMU data set with the reanalysis SOM BMUs to create a null distribution. Similarly, the same approach is used to test significance of RMSE, ETC frequency error, and ACA error for CMIP6 and HighResMIP means.

3. Results and Discussion

3.1. Patterns of Geopotential Height Variability

Figure 2 shows the historical large-scale synoptic conditions and corresponding ETC events associated with each pattern in the master SOM. The underlying color shading shows anomalies (relative to the DJF climatological mean) in the Z500 field, where positive (red) and negative (blue) anomalies indicate mid-atmospheric ridges and troughs, respectively. The mean climatology for each pattern is available in supporting information. The

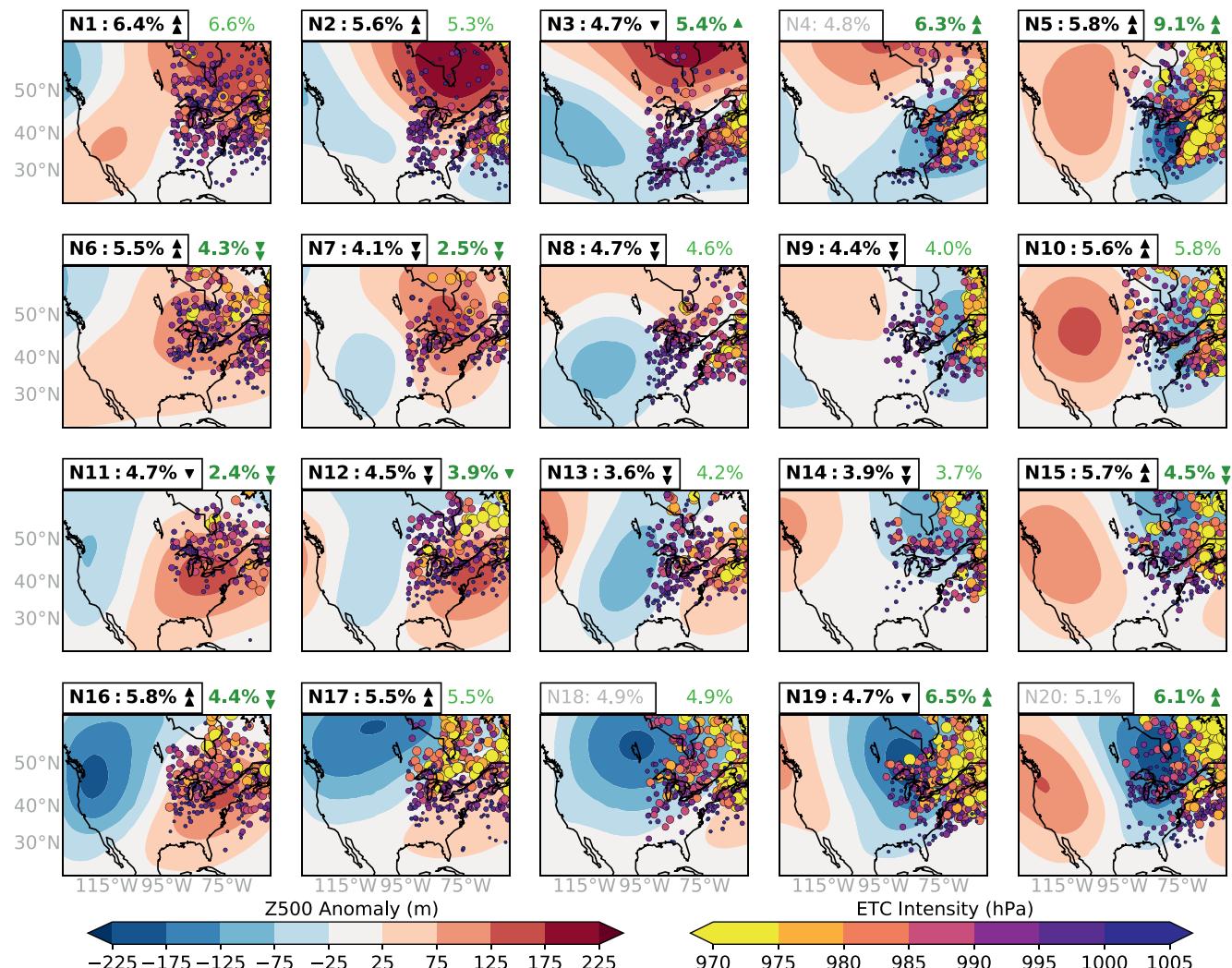


Figure 2. Master self-organizing map composite of DJF 500-hPa geopotential height anomalies (m) over NEUS. The location of extratropical cyclone (ETC) events associated with each pattern is indicated with circles, which are color-coded and vary in size according to intensity. The node frequency (black) and ETC frequency (green) are indicated above each node. Bold values indicate significance with up (down) arrows indicating significantly high (low) values at the 10% (single arrow) and 5% (double arrow) significance levels.

frequency of each SOM node in the winter climatology is indicated above each panel (black text), with bold values indicating LSMPs with significantly high or low occurrence relative to a random distribution. The circles indicate the location of ETCs which coincide with each SOM node. The size and color vary according to the minimum pressure intensity of the cyclone, with larger, brighter dots indicating more intense ETC events as measured by SLP. The frequency of ETC hits that occur in each SOM node is indicated above each panel (green text), with bold values indicating significantly high or low storm occurrence relative to the SOM pattern frequency.

The SOM effectively breaks down the winter climatology into distinct patterns of Z500 variability. Overall, the SOM represents different phases of a propagating Rossby wave, emphasizing the advantage of using SOMs to represent the data as a continuum. The outer edges of the SOM show more amplified anomalies indicating what can be interpreted as more “extreme” patterns, whilst the interior nodes show weaker anomalies, indicating more zonal, progressive, transitional patterns. Nodes 1, 2, 6, 11, and 16 show anomalous ridging over the eastern US (henceforth referred to as the *ridging patterns*), which is furthest north and more pronounced in node 2. Similarly, nodes 5, 10, 15, 19, and 20 show a deepened trough over the eastern US (henceforth referred to as the *trough patterns*), which is furthest north and more pronounced in node 20. Node 1 is the most frequent pattern (6.4%) and node 13 is the least frequent (3.6%). Nodes 7, 8, 9, 12, 13, and 14 will henceforth be referred to as the *zonal patterns*. Generally, all of the nodes have a similar frequency, indicating no single pattern dominates the winter

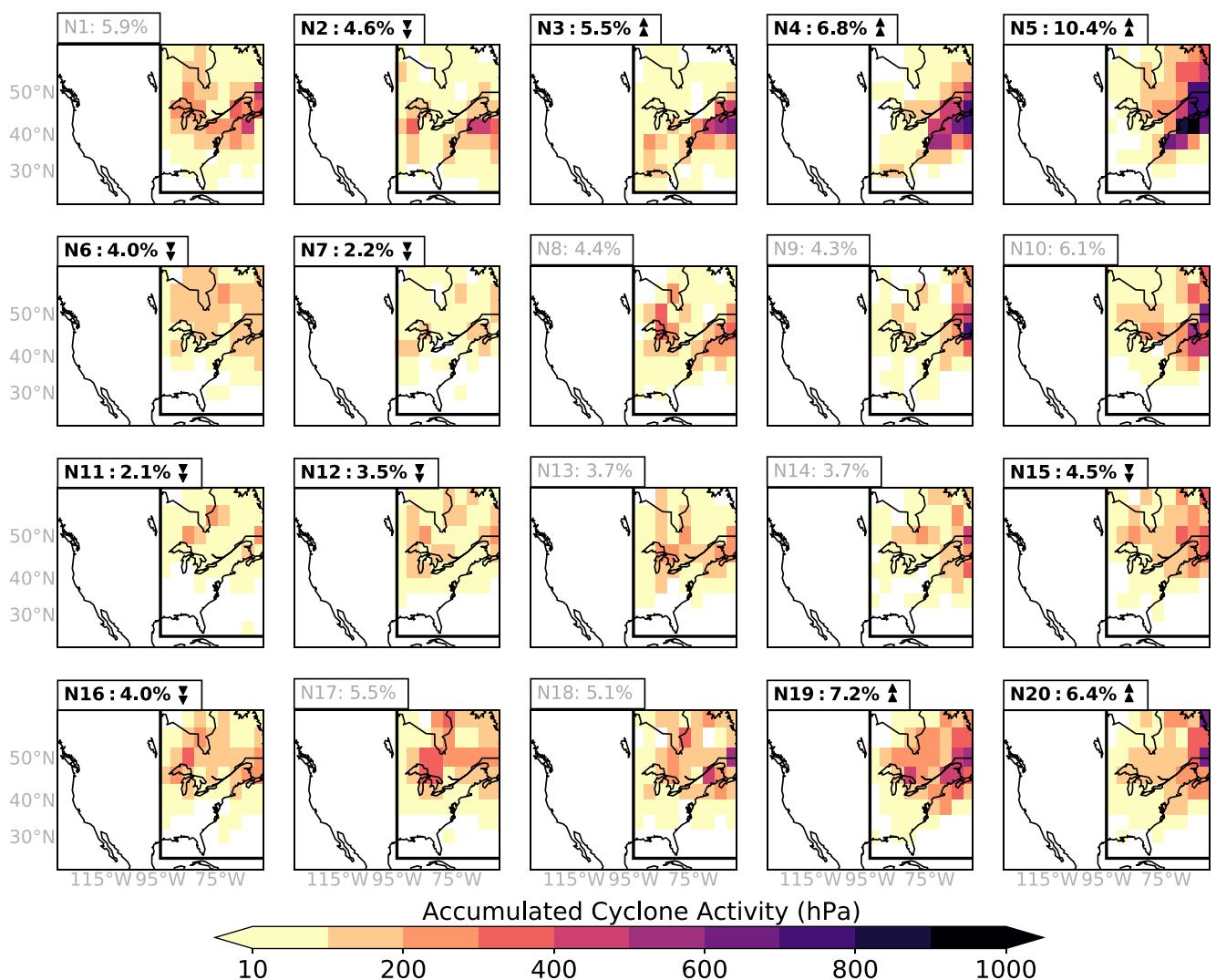


Figure 3. Accumulated cyclone activity (ACA, hPa) for each self-organizing map node calculated at a 4° grid spacing. Percent contribution to overall winter ACA indicated above each map. The black box represents the bounding of the eastern US. Bold values indicate significant ACA contribution relative to node frequency, with up (down) arrows indicating significantly high (low) values at the 10% (single arrow) and 5% (double arrow) significance levels.

climatology of the NEUS, although all of the patterns have frequencies statistically distinct from purely random activity, with the exception of N4, N18 and N20.

Geopotential height is dynamically related to SLP, therefore it is expected that ETCs will be strongly associated with troughs in the Z500 fields. This is particularly evident in nodes 4, 5, 19, and 20 (henceforth referred to as the *stormy patterns*), where there are significantly high ETC frequencies and noticeably more intense ETCs (<975-hPa) over the eastern US. In most cases, ETCs are situated slightly eastward of the Z500 minimum, attributed to the typically observed westward tilt associated with cold-core cyclones in the Northern Hemisphere (Dacre et al., 2012; Hoskins, 1990). ETC frequencies are higher in the nodes with amplified meridional flow (nodes 5, 19, and 20), which are typically associated with increased temperature advection leading to high baroclinic instability (Catto, 2016), producing favorable conditions for ETC formation. Interestingly, node 15 has significantly low ETC frequency despite being a trough pattern. It is possible that ETCs are falling outside of the domain due to the more northerly position of the trough, as surface low pressure centers tend to form downstream of upper level troughs in a baroclinic atmosphere (Hoskins, 1990). In contrast, there are fewer ETCs in the interior, *zonal patterns* (nodes 7, 9, 12, and 14, although only significant in nodes 7 and 12) which are characteristic of more transitional patterns, as well as in the *ridging patterns* (nodes 6, 11, and 16). Nevertheless, it is important to note that at least some ETCs do occur in all patterns, although they tend to be fairly weak in unfavorable environments.

3.2. Accumulated Cyclone Activity

Figure 3 shows the ACA associated with each SOM node. ACA is the sum of pressure anomalies associated with all ETCs (Equation 1) which occur during that pattern for each 4° grid box, therefore taking into consideration both ETC frequency and intensity. The fractional contribution of each SOM pattern to total winter ACA is indicated above each panel. In agreement with Figure 2, the *trough patterns*, associated with an east coast trough and high baroclinic instability, also experience the highest cyclone activity over the NEUS. In particular, the *stormy patterns* (nodes 4, 5, 19, and 20) account for approximately 30.8% of winter ACA. In contrast, the *ridging patterns* (nodes 2, 6, 11, and 16) and *zonal patterns* (nodes 7, 12, 13, and 14) are associated with the least ACA. Here we see the advantage of using ACA over a discrete count of ETC days. For example, Figure 2 indicates that node 1 has higher ETC frequency (6.6%) than node 20 (6.1%), however, when incorporating ETC intensity in Figure 3, node 20 accounts for more winter ACA (6.4%) than node 1 (5.9%). This differentiation can be critical as information about more intense cyclones is generally of greater value to stakeholders. ACA also highlights some spatial differences in the regional distribution of ETCs between the patterns. For example, some nodes have ETCs with zonal-like trajectories (e.g., nodes 1 and 2), whilst others have ETCs with more meridional-like trajectories (e.g., nodes 4 and 5). This implies that the SOM could be discriminating patterns conducive to the evolution of different “types” of ETCs, such as those described in Zielinski (2002) and Colle et al. (2015), although further analysis is needed to confirm this.

3.3. CMIP6 Analysis

The master SOM can be applied as a tool to evaluate the fidelity of historical simulations from CMIP6 climate models. It is important to note that the aim of this work is not to completely disentangle the LSMPs from ETCs, but rather we are interested in the synoptic evolution and representation of these events in the climate models. We analyze four different types of error. First, we consider the bias in node frequency. Is the model able to reproduce the pattern with the same frequency as in reanalysis? Second, we consider the RMSE. Is the model able to reproduce the correct spatial structure of a given pattern? Third, we consider the bias in ETC frequency. Given a particular pattern, is the model able to reproduce the correct number of ETC hits? Last, we consider the bias in ACA. Given an ETC in a particular pattern, is the model able to reproduce the correct ETC intensity? Figure 4 provides an example of how we present the frequency error for a given model. This is repeated for all CMIP6 climate models and consolidated into a heat map shown in Figure 5.

Figure 5 indicates the CMIP6 climate model errors in simulating the frequency associated with each SOM node (columns). Each row represents the difference in frequency between the climate data (reanalysis or climate model) and the master SOM frequency shown in the first row. Blue (red) shading represents models which have negative (positive) frequency biases indicating that model under (over) represents that pattern in the data. For example, node 1 has a frequency of 6.43% in the master SOM and a frequency of 6.16% in BCC-CSM2-HR, therefore, that pattern occurs less frequently in the model with a difference of -0.27% (as shown in Figure 5). If a similar bias persists across most climate models (i.e., down the column), this indicates a systematic error across the CMIP6 ensemble. The second last column shows the mean pattern error (MPE) which is the arithmetic mean of the absolute values across nodes 1 to 20. This provides a general overview of the model's performance across the nodes, with high values indicating that the model struggles to reproduce the SOM patterns, either producing too many or too few occurrences relative to the master SOM. The last column shows the absolute total error (ATE) for each model, which is zero in Figure 5 because all models add up to 100% frequency. Data products are clustered as reanalyses followed by model groups in ascending order of grid spacing that is, HighResMIP ($\sim 25\text{--}50\text{ km}$), CMIP6 standard resolution ($\sim 100\text{ km}$), and CMIP6 low resolution ($\sim 250\text{ km}$) from top to bottom.

The master SOM is trained using the reanalysis products, therefore it is expected that the reanalysis shows relatively low error. Further, reanalyses broadly assimilate the same data and therefore should be constrained synoptically in similar fashions. Nevertheless, 20CRv3 does show some marginally larger errors in comparison to the other reanalysis products. These differences are likely a result of the 20CRv3 data only assimilating surface observations (Slivinski et al., 2019b). In general, the CMIP6 climate models tend to favor the more *zonal patterns* with common positive frequency errors seen in nodes 6, 7, 9, 10, 13, and 17. This agrees with previous studies (e.g., Priestley et al., 2020; Zappa et al., 2013) which found that CMIP6 models typically under-represent blocking patterns with a tendency toward zonal storm tracks. In contrast, there are persistent negative frequency errors associated with the more amplified *trough* (nodes 5, and 20) and *ridging patterns* (nodes 1, 11, 12, and 16). These

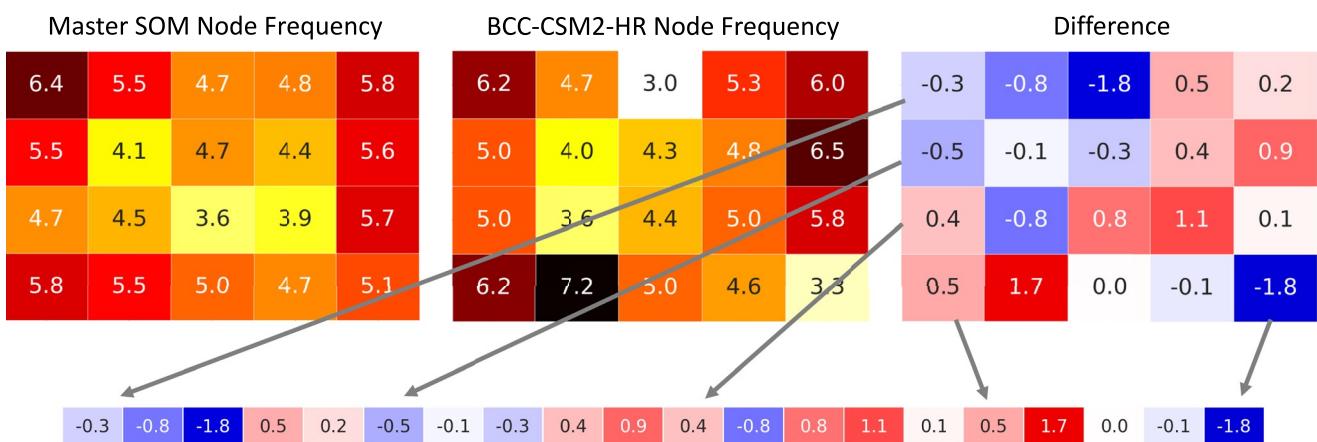


Figure 4. Example showing how the node frequency error is calculated for BCC-CSM2-HR and deconstructed into a linear form. Note that these are the same values as in Figure 5, except rounded to the nearest tenth of a percent.

frequency errors are seen across most of the CMIP6 models, indicating a systematic error where models consistently underrepresent (overrepresent) the amplified (zonal) patterns. These errors persist in the core standard and low-resolution CMIP6 climate models as well as the HighResMIP models, indicating that model resolution isn't the key driver of a model's ability to reproduce large-scale synoptic patterns. The mean CMIP6 error is shown as a heat map in Figure 6a to visualize the systematic model errors with respect to their relationship with surrounding nodes. Here we see the errors which persisted across the models and did not average out, showing negative frequency biases associated with the most extreme, outer nodes and positive frequency biases associated with the interior nodes. This reaffirms the shift in node frequency from the *trough/ridging patterns* to the *zonal patterns*.

Similarly, Figure 7 shows the spatially averaged RMSE between each of the CMIP6 climate models and the master SOM for each SOM node. The dark (light) shading specifies high (low) spatial error, indicating structural deficiencies (precision) within the patterns of Z500 anomalies. Here, the last column shows the normalized total error which is the total error of all days after the mean climatology has been removed. In contrast, ATE is the total error of all days before the climatology was removed. While a climate model may be able to capture the node frequency, the actual pattern may be spatially incorrect (and vice versa). For instance, node 12, which showed significant frequency error in Figure 5, has relatively lower spatial error (although still significant). This leads to the interpretation that the models represent the pattern adequately from a structural standpoint, but simulate their occurrence too frequently. As expected, there is higher spatial error associated with the more amplified *trough patterns* and *ridging patterns*, especially nodes 3, 4, 5, 11, 16, and 17, which also experienced high frequency errors. This indicates that the climate models particularly struggle to reproduce these specific patterns in terms of both frequency and spatial structure. These errors are systematic, persisting across most models, resulting in a bias in the CMIP6 mean, as seen in Figure 6b. As above, there is no noticeable difference between the performance of the low-resolution (250 km), standard-resolution (100 km) CMIP6 climate models, and the HighResMIP models (25–50 km) in the simulation of the large-scale patterns. Although high-resolution models have been shown to improve flow interacting with orography (e.g., Jung et al., 2012) and that large-scale mountain ranges play a role in shaping the mid-latitude jet stream (Babaei et al., 2021), these results indicate that factors other than atmospheric resolution have a significant influence on the representation of LSMPs in coupled climate models.

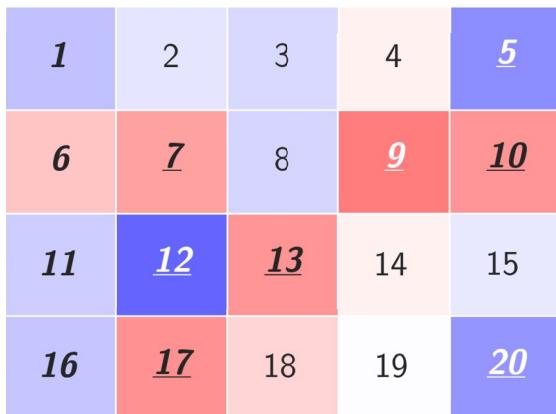
Shifting to an ETC (i.e., storm-scale) perspective, we now consider the error in ETC frequency and intensity. Figure 8 shows the bias in the average number of ETC hits on a given day in a node. This has been normalized relative to the number of days in each node to remove any bias from the node frequency error discussed above (i.e., assuming a model simulates the large-scale climatology correctly, what errors are due to deficiencies in representing ETC climatology?). For example, node 5 in ERA5 has high ETC frequency with an average of 2.96 ETC hits per day, whilst node 11 has less ETC activity with an average of 1.03 ETC hits per day. Previous studies (e.g., Priestley et al., 2020; Zappa et al., 2013) have identified mean biases in the frequency and intensity of simulated ETCs in CMIP5 and CMIP6 climate models, confirmed here by the ATE, therefore we normalize ETC hits relative to ERA5 statistics to remove the mean model bias and to analyze discrepancies within the patterns which are not a remnant

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	MPE	ATE
MASTER	6.43	5.55	4.73	4.78	5.80	5.53	4.06	4.67	4.42	5.63	4.69	4.46	3.62	3.89	5.72	5.75	5.52	4.95	4.72	5.10	5.00	100.00
ERA5	-0.20	0.07	-0.02	-0.05	0.11	-0.16	0.15	-0.02	-0.12	0.14	-0.14	0.05	0.07	0.10	-0.07	0.04	0.18	-0.07	-0.09	0.02	0.09	0.00
JRA-55	-0.03	0.10	-0.13	-0.07	0.11	-0.07	-0.13	-0.10	0.21	-0.09	0.02	0.11	-0.02	0.07	0.10	0.04	-0.07	-0.04	0.02	-0.03	0.08	0.00
MERRA2	-0.00	0.04	-0.02	-0.10	-0.01	-0.24	0.04	-0.07	0.10	0.14	-0.06	0.11	-0.05	0.07	0.02	0.04	0.13	-0.13	-0.09	0.08	0.08	0.00
CFSR	-0.22	0.04	0.07	-0.07	0.05	0.01	-0.07	0.12	0.04	-0.09	-0.01	0.08	-0.05	-0.04	0.21	0.04	-0.07	-0.02	0.02	-0.06	0.07	0.00
20CRv3	0.50	-0.29	0.11	0.33	-0.29	0.50	0.03	0.07	-0.26	-0.12	0.21	-0.40	0.04	-0.23	-0.27	-0.18	-0.20	0.29	0.17	-0.02	0.22	0.00
BCC-CSM2-HR	-0.27	-0.82	-1.77	0.49	0.20	-0.48	-0.06	-0.32	0.38	0.91	0.36	-0.84	0.83	1.06	0.06	0.50	1.71	0.01	-0.12	-1.83	0.65	0.00
CMCC-CM2-VHR4	0.12	0.40	1.35	1.29	-1.4	0.58	-0.58	-1.47	0.62	0.58	-0.57	-0.47	0.97	-0.12	0.30	-0.94	-0.43	1.32	-0.42	-1.14	0.75	0.00
E3SM-1-0-HR	-0.40	-0.28	0.07	-0.66	-0.56	0.72	-1.17	-0.73	-0.03	2.22	0.93	-0.97	0.41	-0.30	-0.22	0.22	0.57	0.74	0.36	-0.91	0.62	0.00
ECMWF-IFS-HR	-0.19	-0.04	0.24	0.16	-1.46	1.05	1.61	-0.43	0.11	-0.02	-1.39	-0.31	0.09	1.02	0.84	-0.97	-0.05	0.31	-0.51	-0.03	0.54	0.00
ECMWF-IFS-MR	-0.95	-0.42	-0.07	0.35	-1.14	1.02	1.35	0.55	1.19	0.04	-1.2	-0.95	0.56	-0.47	0.49	-0.43	0.05	-0.58	0.38	0.25	0.62	0.00
ECMWF-IFS-LR	0.12	-0.58	-0.77	-0.10	-0.41	0.49	1.58	-0.17	0.14	0.58	-1.05	-1.01	0.50	-0.47	0.43	0.20	0.40	0.24	0.60	-0.7	0.53	0.00
GFDL-CM4C192	0.24	-0.89	-0.60	-0.56	-0.56	-0.13	0.67	-0.16	0.57	0.66	-0.12	-0.27	0.48	-0.37	-0.64	0.57	0.54	0.64	0.45	-0.53	0.48	0.00
MPI-ESM1-2-XR	-0.07	-0.46	0.84	0.79	-1.78	0.93	-0.32	-0.68	1.69	-0.34	-1.01	-0.88	1.26	-0.57	1.34	-0.94	0.24	0.97	0.03	-1.05	0.81	0.00
HRMIP Mean	-0.17	-0.39	-0.09	0.22	-0.89	0.52	0.38	-0.43	0.58	0.58	-0.51	-0.71	0.64	-0.03	0.32	-0.22	0.38	0.46	0.10	-0.74	0.42	0.00
BCC-CSM2-MR	-0.46	1.43	-1.55	-0.02	0.27	-0.99	0.19	-0.96	1.74	0.41	-1.1	-1.83	0.92	0.01	-0.03	0.03	2.1	0.51	0.39	-1.07	0.8	0.00
CESM2	0.37	0.00	-0.38	-1.13	-1.13	-0.32	0.99	-0.16	0.79	0.95	0.04	-1.03	0.83	0.21	-0.29	0.47	-0.38	0.07	0.42	-0.31	0.51	0.00
CMCC-CM2-HR4	-0.43	0.96	-0.63	0.77	-0.65	-0.39	0.13	0.06	0.82	0.82	-0.12	-1.6	0.6	0.55	-0.76	-0.07	1.17	-0.25	-0.34	-0.66	0.59	0.00
CMCC-CM2-SR5	0.17	-0.16	-0.50	0.68	-0.37	1.01	0.04	-0.29	0.38	0.37	-0.72	-1.41	0.95	0.81	-0.54	0.28	0.57	-0.57	-0.15	-0.56	0.53	0.00
E3SM-1-0-LR	-0.56	0.45	-0.50	0.52	-0.33	0.69	0.13	-1.11	0.09	0.91	0.20	-0.68	0.32	0.30	-0.76	-0.04	-0.03	0.42	0.71	-0.72	0.47	0.00
EC-Earth3	-1.08	0.18	-1.4	0.12	-1.05	0.83	1.61	0.46	0.62	0.07	0.31	-0.60	0.34	0.35	-0.02	-1.03	0.33	0.09	0.28	-0.41	0.56	0.00
EC-Earth3-Veg	-0.26	0.72	-0.61	1.1	-1.55	0.74	1.01	-1.06	0.17	0.14	-0.7	-0.38	0.72	0.64	0.46	-0.69	0.21	0.85	-0.10	-1.4	0.67	0.00
EC-Earth3-AerChem	-0.67	-0.39	-0.77	0.63	-0.98	0.71	0.75	-0.17	2.23	0.39	-0.73	-1.17	1.41	1.05	-0.46	-0.53	0.43	-0.45	-0.38	-0.89	0.76	0.00
GFDL-CM4	-1.44	-0.85	0.32	-0.02	0.36	-0.39	1.31	-1.46	0.09	0.09	0.87	-0.52	0.98	0.01	-0.35	-0.04	0.25	-0.28	0.29	0.77	0.54	0.00
MPI-ESM1-2-HR	-1.81	-0.36	0.53	0.25	-0.73	0.83	0.06	0.55	1.69	0.33	-0.45	-0.60	-0.39	0.10	0.46	-0.88	0.93	0.53	-0.16	-0.89	0.63	0.00
MRI-ESM2-0	-0.76	-0.08	2.4	-0.60	-1.62	1.78	0.28	-1.16	1.6	0.74	-0.73	-0.22	-0.23	0.48	0.20	-1.73	0.08	0.53	0.09	-1.05	0.82	0.00
NorESM2-MM	-0.81	-0.19	0.42	-0.02	-0.81	-0.13	0.64	-0.16	0.22	1.74	-0.31	-0.75	0.16	-0.43	0.13	0.03	0.54	0.74	0.07	-1.07	0.47	0.00
SAM0-UNICON	0.78	0.86	0.13	-1.16	0.65	-0.55	0.83	-0.57	0.95	1.33	-0.78	-1.44	0.13	-0.59	-0.86	-0.8	1.71	-0.12	0.74	-1.23	0.81	0.00
TaiESM1	0.71	-1.11	-0.63	0.14	-0.18	-0.07	1.02	-1.18	1.08	0.69	-0.56	-0.30	0.98	-0.05	-0.64	0.19	0.83	-0.57	-0.24	-0.12	0.56	0.00
ACCESS-CM2	0.69	-0.74	1.7	1.01	-1.65	0.39	0.44	0.08	0.49	-0.15	-0.10	-0.06	-0.23	0.32	-0.27	-1.38	-0.21	0.53	0.38	-1.24	0.6	0.00
ACCESS-ESM1-5	-0.35	-0.71	-0.83	-0.26	0.19	0.8	2.21	0.46	0.21	-0.34	-0.64	-1.04	0.59	-0.53	0.11	-0.78	-0.05	0.97	-0.07	0.06	0.56	0.00
AWI-ESM-1-1-LR	-1.21	-0.58	0.12	0.03	-0.35	-0.62	-0.01	-0.71	-0.68	1.66	-0.19	-0.76	1.73	-0.09	0.08	0.9	1.09	-0.14	0.34	-0.60	0.59	0.00
EC-Earth3-Veg-LR	0.22	0.05	-1.02	0.47	-1.02	0.68	0.47	-0.97	1.47	0.36	-0.45	-0.60	0.12	-0.12	-0.37	-0.59	0.59	1.26	0.15	-0.7	0.58	0.00
GISS-E2-1-G	0.14	-0.92	1.31	1.12	-3.22	0.09	-0.09	1.23	-0.29	-0.64	0.39	-0.78	0.57	-0.27	1.52	0.09	0.41	-0.60	-0.05	-0.02	0.69	0.00
IPLS-CM6A-LR	-1.62	-0.27	-0.55	-0.03	0.44	0.74	0.94	-0.87	0.52	0.42	-0.41	-1.49	0.40	0.10	-0.84	0.55	1.98	0.05	-0.04	-0.03	0.61	0.00
KIOST-ESM	-0.75	-0.06	0.48	-0.53	-1.54	1.61	1.62	-0.96	1.24	1.45	-0.18	-1.6	0.86	-0.02	0.89	-2.07	-0.06	0.51	-0.88	-0.02	0.87	0.00
MIROC6	0.16	-0.11	0.46	-0.13	-1.46	-0.02	0.53	1.28	0.43	0.04	-0.60	0.07	0.37	-0.76	0.36	-0.37	0.21	0.24	-0.19	-0.51	0.41	0.00
MPI-ESM-1-2-HAM	0.35	-0.49	-1.05	-0.07	0.22	-0.56	-1.4	-0.02	0.59	0.01	0.41	-0.73	1.79	0.83	-0.30	0.20	1.19	-0.04	-0.10	-0.83	0.56	0.00
MPI-ESM-1-2-LR	-0.54	0.15	-1.15	-0.64	0.09	0.49	-0.13	-0.17	1.31	1.62	-0.35	-1.01	0.66	-0.69	-1.16	-0.08	1.6	0.88	0.15	-1.02	0.69	0.00
NESM3	-0.38	-1.12	-0.96	0.12	0.31	0.74	-1.02	0.17	1.0	1.69	-0.51	-1.11	1.19	-0.31	-0.27	0.52	1.47	0.62	-1.05	-1.11	0.78	0.00
NorESM2-LM	-0.08	-0.92	-1.11	-0.34	-1.03	0.50	1.59	1.36	0.89	1.3	0.04	-1.6	0.51	0.11	0.63	-0.99	-0.00	0.51	-0.59	-0.78	0.74	0.00
CMIP6 Mean	-0.37	-0.16	-0.22	0.08	-0.66	0.33	0.54	-0.24	0.76	0.63	-0.28	-0.89	0.63	0.08	-0.12	-0.34	0.65	0.24	-0.01	-0.63	0.39	0.00

Figure 5. Frequency error (%) of the reanalysis, HighResMIP climate models, and CMIP6 core climate models for each self-organizing map (SOM) node. The first row acts as a reference showing the node frequency of the master SOM, while the subsequent rows show the deviation from this reference. The HighResMIP and CMIP6 mean errors are also shown for each node. The second last column shows the mean pattern error (MPE) for each model, which is the arithmetic mean of the absolute values across nodes 1 to 20. The last column shows the absolute total error (ATE) for each model. Shading indicates significant values, where blue indicates that the model has a negative frequency bias and red indicates that the model has a positive frequency bias. Bold, italicized values indicate significance at the 10% level (5th and 95th percentiles) and underlined values indicate significance at the 5% level (2.5th and 97.5th percentiles). Bold vertical lines indicate the start of a new row in the SOM structure. The CMIP6 core models have been ordered from standard resolution (100 km) to lower resolution (250 km).

of the mean bias. This is explained in more detail in supporting information. For example, BCC-CSM2-HR shows a positive ATE toward the right of the table indicating that it generates (on a per day basis) more ETC events than ERA5. When

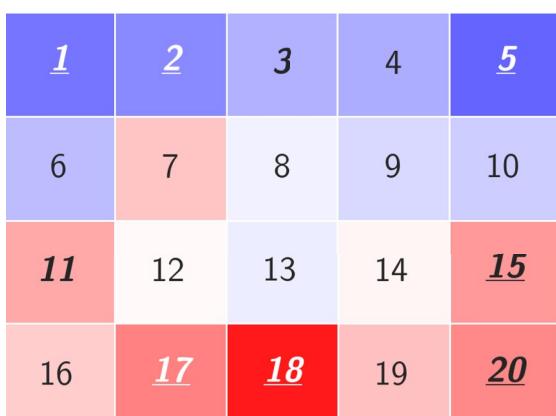
a) Node frequency error



b) RMSE



c) ETC frequency error



d) ACA error

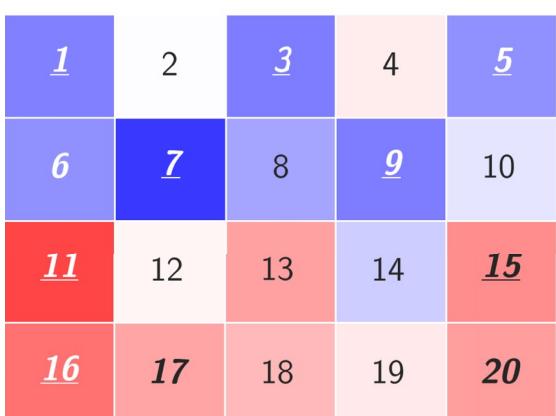


Figure 6. Heat map showing the mean CMIP6 error for (a) node frequency, (b) root mean square error, (c) extratropical cyclone frequency, and (d) accumulated cyclone activity. Note that this is equivalent to the last row in Figures 5, and Figures 7–9, however, here the self-organizing map structure has been maintained to show the relationship between nodes. Bold, italicized values indicate significance at the 10% level (5th and 95th percentiles) and underlined values indicate significance at the 5% level (2.5th and 97.5th percentiles).

total ETCs across all the CMIP6 climate models, which is most significant in nodes 11, 15, 17, 18, and 20, as well as a significant negative bias in nodes 1, 2, 3, and 5. There is disagreement in ETC frequency error between the CMIP6 climate models in nodes 8, 12, 13, and 14, resulting in relatively low error in the CMIP6 mean (Figure 6c). In contrast to node frequency error and RMSE, Figure 8 shows that the other reanalyses have some notable ETC frequency errors when compared to ERA5 (although not significant). However, this is consistent with previous studies which found discrepancies in ETC dynamics between reanalysis products due to variations in both horizontal and vertical resolutions, as well as the representation of orography which is important for cyclogenesis, and how much of the data are assimilated into the reanalysis system (Hodges et al., 2011; Rohrer et al., 2018; Tilinina et al., 2013). This highlights that the CMIP6 errors could vary if another reanalysis was used as the reference.

Lastly, Figure 9 shows the average ACA per ETC hit (i.e., 6-hourly snapshot) in a given node. This removes any bias associated with the node frequency error and ETC frequency error, providing insight into the relative error in ETC intensity (i.e., assuming a well-simulated pattern and ETC frequency, what errors are apportioned to storm strength?). For example, node 5 in ERA5 has particularly strong ETCs with an average ACA of 25.25 hPa per ETC hit, whilst node 11 has weaker ETCs with an average ACA of 19.07 hPa per ETC hit. Again, there is a significant ATE in most models, therefore model ACA has been normalized relative to ERA5 to remove the average ACA bias (see supporting information for more information). For example, the ETCs in nodes 5 and 11 in BCC-CSM2-HR are stronger by 0.21 and 1.94 hPa, respectively. That is, when the ACA of BCC-CSM2-HR is normalized to ERA5 to remove the mean model bias, it relatively overproduces the strength of ETCs in these

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	MPE	ATE	NTE
MASTER	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ERA5	1.31	1.12	1.64	1.59	1.65	1.8	1.33	1.24	1.75	1.06	1.58	1.54	1.95	1.52	1.41	1.87	1.1	2.05	1.84	1.39	1.54	2.43	0.0
JRA-55	1.63	1.28	2.19	1.42	1.24	1.76	1.72	1.8	1.67	1.5	1.13	1.71	1.59	1.99	1.85	1.35	1.36	1.72	1.53	0.95	1.57	2.87	0.0
MERRA2	1.11	1.53	2.0	2.65	1.86	1.64	1.78	1.73	1.78	1.17	1.89	1.64	1.98	2.2	1.29	1.85	1.78	2.33	1.9	1.18	1.76	1.66	0.0
CFSR	1.22	1.49	1.75	1.49	2.11	1.49	1.85	1.45	1.75	1.48	1.73	1.31	1.74	1.22	2.0	1.37	1.11	1.98	1.45	1.36	1.57	2.03	0.0
20CRv3	3.62	4.15	6.05	4.74	6.32	4.66	4.77	3.88	3.12	4.02	4.85	4.14	5.34	4.75	6.72	5.94	3.83	6.16	5.91	3.91	4.84	4.36	0.0
BCC-CSM2-HR	14.76	13.26	13.5	17.15	17.49	12.75	8.96	11.66	15.52	9.53	11.37	10.98	9.14	12.91	9.18	19.84	17.14	11.91	7.63	15.0	12.98	26.1	0.0
CMCC-CM2-VHR4	7.9	9.6	14.69	16.27	21.49	12.53	13.35	13.43	11.96	18.84	17.06	11.57	8.61	11.43	14.64	17.71	13.07	11.71	12.07	13.19	13.56	26.48	0.0
E3SM-1-0-HR	11.5	14.95	22.98	9.17	12.46	12.83	19.19	16.44	15.92	11.16	11.17	14.09	14.01	11.17	12.68	14.57	15.19	9.63	12.05	11.69	13.64	19.36	0.0
ECMWF-IFS-HR	10.1	8.95	12.61	8.27	11.23	5.94	9.71	12.09	9.72	10.08	14.66	9.44	10.59	6.96	7.77	13.22	11.75	5.55	10.03	7.11	9.79	18.98	0.0
ECMWF-IFS-MR	9.15	11.65	12.96	13.04	8.57	6.14	12.44	5.58	6.79	12.41	11.71	7.52	12.24	7.01	7.01	10.86	9.92	5.79	5.96	7.8	9.23	17.81	0.0
ECMWF-IFS-LR	11.76	10.68	10.74	14.27	8.41	6.47	15.58	11.93	6.91	14.08	11.95	11.3	10.11	9.64	11.17	12.85	12.47	11.28	13.61	5.99	11.06	31.4	0.0
GFDL-CM4C192	12.1	11.42	15.48	15.35	15.08	12.42	11.61	8.76	7.95	11.44	19.43	13.49	9.0	12.54	10.45	17.09	9.3	8.57	6.8	10.62	11.94	12.83	0.0
MPI-ESM1-2-XR	11.27	15.74	14.73	14.69	14.88	16.93	19.7	8.29	6.5	18.12	16.55	13.54	15.09	13.11	12.73	18.46	18.88	15.12	12.37	11.08	14.39	31.85	0.0
HRMIP Mean	7.79	7.36	6.94	6.89	9.5	8.11	11.25	6.69	7.13	9.7	10.37	7.68	8.55	7.68	8.9	12.03	10.7	7.2	6.84	4.94	8.31	14.08	0.0
BCC-CSM2-MR	15.53	18.0	18.99	16.21	20.64	15.68	17.39	16.16	20.41	11.31	16.11	13.93	17.78	17.5	10.82	30.4	26.59	20.9	14.36	16.62	17.76	12.1	0.0
CESM2	11.31	7.31	13.15	14.65	12.81	6.18	11.08	5.0	8.89	9.19	10.93	10.84	15.89	5.07	8.79	11.97	5.4	8.79	4.66	11.27	9.66	17.56	0.0
CMCC-CM2-HR4	15.19	11.96	12.96	18.73	16.62	17.01	14.34	11.86	22.15	9.39	18.4	10.81	14.58	16.1	12.45	27.41	19.86	15.32	13.7	11.78	15.53	41.87	0.0
CMCC-CM2-SR5	13.23	12.03	10.01	11.98	18.15	6.48	10.78	10.92	10.65	12.13	9.93	10.48	9.72	12.2	10.28	14.36	13.89	11.33	8.85	12.29	11.48	27.6	0.0
E3SM-1-0-LR	7.65	7.32	22.89	12.6	10.87	9.86	10.83	10.02	9.97	11.75	13.39	13.94	9.58	9.78	7.27	18.9	13.28	7.64	7.7	13.94	11.46	18.3	0.0
EC-Earth3	11.2	6.45	12.14	10.2	6.97	6.73	10.79	11.01	5.84	9.4	13.79	11.38	13.28	12.83	14.18	7.39	14.12	8.04	9.73	10.34	22.2	0.0	
EC-Earth3-Veg	9.93	8.03	7.3	15.66	11.6	6.14	9.13	7.53	8.3	10.17	15.16	6.18	10.06	8.38	12.72	12.14	8.62	11.38	10.9	9.67	9.95	23.33	0.0
EC-Earth3-AerChem	11.43	8.39	14.41	9.74	12.82	10.32	7.35	8.93	8.46	12.67	8.94	8.64	14.23	7.78	12.63	12.84	9.73	11.53	8.1	7.54	10.32	26.08	0.0
GFDL-CM4	12.37	10.45	11.16	17.82	8.41	11.79	11.54	7.99	12.88	8.71	12.12	7.16	7.35	17.1	12.7	17.72	16.27	11.52	9.21	7.19	11.57	34.19	0.0
MPI-ESM1-2-HR	15.38	18.34	17.28	14.54	12.39	13.89	19.72	9.23	17.3	11.99	24.87	11.74	9.45	10.73	15.1	24.73	21.68	15.53	11.89	11.42	15.36	21.6	0.0
MRI-ESM2-0	12.99	9.75	21.63	12.57	13.58	9.31	9.31	9.66	9.64	9.63	11.28	8.78	9.39	5.62	11.45	18.8	15.63	10.98	12.05	10.61	11.63	29.21	0.0
NorESM2-MM	12.71	8.92	12.42	11.76	10.44	10.69	11.22	14.91	7.99	9.76	9.54	8.88	14.12	9.31	10.21	8.41	7.62	9.08	11.74	9.07	10.44	15.36	0.0
SAM0-UNICON	9.41	9.99	17.95	11.5	19.11	10.46	7.64	9.39	11.28	15.6	17.99	9.05	14.98	9.41	9.03	13.59	19.03	9.42	17.27	13.08	12.76	24.13	0.0
TaiESM1	9.93	10.53	16.1	11.85	14.73	8.51	11.47	12.09	10.98	7.54	16.83	10.56	13.02	6.26	9.77	11.1	13.14	8.62	9.0	11.35	11.17	21.59	0.0
ACCESS-CM2	10.24	9.67	13.32	14.61	13.65	12.88	14.9	7.28	10.65	11.36	10.24	7.97	12.46	8.44	9.88	14.69	7.38	9.91	9.66	9.88	10.95	34.64	0.0
ACCESS-ESM1-5	12.65	9.37	16.16	12.72	7.95	6.46	9.55	8.55	11.88	11.66	12.53	11.5	12.71	5.57	10.78	18.63	13.54	8.61	9.02	10.26	11.0	33.91	0.0
AWI-ESM-1-1-LR	14.22	13.33	15.29	17.93	27.15	17.81	12.61	17.29	15.22	14.56	21.8	14.56	12.02	13.27	13.07	20.1	17.61	15.35	7.39	12.52	15.65	31.09	0.0
EC-Earth3-Veg-LR	12.07	13.96	10.86	10.09	9.38	11.91	10.37	12.05	8.22	10.89	11.87	10.51	19.54	4.36	10.61	13.83	9.01	8.03	11.77	11.3	11.03	33.14	0.0
GISS-E2-1-G	12.83	11.86	21.22	19.27	18.4	14.95	9.4	16.16	10.98	20.74	21.81	7.58	12.13	8.72	16.11	12.64	10.97	19.87	12.7	13.76	14.61	60.24	0.0
IPSL-CM6A-LR	11.17	19.67	13.83	22.58	21.06	13.6	16.28	17.07	15.55	14.26	16.91	11.64	10.71	11.74	12.28	21.46	19.27	13.04	14.16	11.34	15.38	54.39	0.0
KIOST-ESM	12.11	18.78	28.83	19.37	10.7	9.68	9.11	14.1	10.28	10.87	10.31	13.1	9.22	9.98	10.01	19.37	15.3	13.56	11.28</b				

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	MPE	ATE	NTE
ERA5	2.01	1.81	2.18	2.54	2.96	1.53	1.12	1.9	1.76	1.94	1.03	1.66	2.17	1.77	1.53	1.46	1.85	1.91	2.69	2.28	1.91	1.91	1.91
JRA-55	-0.05	-0.22	0.13	-0.0	0.05	-0.15	0.18	-0.08	0.05	0.09	-0.0	-0.02	0.0	-0.11	0.1	0.12	0.14	-0.11	-0.07	-0.04	0.09	-0.03	0.0
MERRA2	0.02	-0.01	-0.01	-0.07	0.0	-0.11	-0.06	0.05	0.09	0.01	0.12	-0.02	0.09	0.07	0.04	0.01	0.02	-0.01	-0.02	-0.15	0.05	0.03	0.0
CFSR	-0.1	-0.2	0.09	0.06	-0.1	-0.07	0.03	-0.13	0.04	-0.02	0.09	0.06	0.28	-0.02	0.15	0.06	-0.04	0.07	-0.07	-0.02	0.08	-0.14	0.0
20CRv3	0.07	0.01	-0.09	0.04	-0.18	-0.15	0.24	0.02	-0.13	0.05	-0.07	-0.04	0.09	-0.14	-0.02	0.08	-0.05	-0.21	0.1	0.38	0.11	-0.21	0.0
BCC-CSM2-HR	-0.43	-0.3	-0.12	-0.03	-0.24	0.08	0.14	0.27	-0.1	0.15	0.37	-0.18	-0.17	0.12	0.28	0.02	0.1	0.23	-0.25	0.25	0.19	0.07	0.0
CMCC-CM2-VHR4	0.25	-0.05	0.18	-0.41	-0.56	-0.4	0.53	0.63	-0.14	-0.37	0.2	0.1	0.05	-0.54	-0.04	0.02	0.14	0.83	-0.01	-0.36	0.29	-0.07	0.0
E3SMHR	-0.27	-0.5	-0.41	-0.12	-0.43	-0.01	0.57	-0.35	0.21	-0.01	0.32	0.07	-0.06	0.2	0.4	-0.23	-0.07	0.58	0.17	0.44	0.27	0.05	0.0
ECMWF-IFS-HR	-0.37	0.18	-0.14	-0.21	-0.51	-0.19	0.2	0.23	0.26	-0.05	0.02	0.21	-0.18	0.14	0.19	0.17	-0.15	0.4	-0.01	0.27	0.2	0.08	0.0
ECMWF-IFS-MR	-0.15	-0.11	-0.06	-0.53	-0.69	-0.18	0.23	0.08	-0.16	-0.04	0.45	-0.12	0.26	0.03	0.34	0.11	0.01	0.33	0.22	0.28	0.22	-0.03	0.0
ECMWF-IFS-LR	-0.1	0.07	-0.4	-0.06	-0.47	-0.09	0.26	0.06	0.18	-0.27	0.34	-0.22	0.0	-0.09	0.06	0.06	0.14	0.52	0.13	0.03	0.18	-0.01	0.0
GFDL-CM4C192	-0.13	0.05	-0.71	-0.14	-0.45	-0.09	0.23	0.37	-0.14	-0.07	0.31	0.29	-0.35	0.01	0.1	0.04	0.16	0.37	-0.19	0.56	0.24	0.15	0.0
MPI-ESM1-2-XR	-0.23	-0.27	-0.52	-0.29	-0.42	-0.25	-0.11	-0.05	-0.01	-0.25	0.12	-0.39	0.45	0.09	0.15	0.23	0.28	0.89	0.25	0.33	0.28	-0.03	0.0
HRMIP Mean	-0.18	-0.11	-0.27	-0.24	-0.46	-0.15	0.24	0.15	0.0	-0.1	0.28	-0.01	0.0	0.01	0.18	0.05	0.08	0.53	0.04	0.24	0.17	0.03	0.0
BCC-CSM2-MR	0.89	-0.73	-0.52	-0.19	-0.61	-0.23	0.21	0.17	-0.25	-0.16	0.08	0.16	0.24	0.03	0.29	0.01	0.37	0.79	0.68	0.86	0.37	-0.45	0.0
CESM2	-0.17	0.28	-0.36	-0.37	-0.61	-0.2	0.53	0.33	-0.21	-0.36	0.34	0.58	0.44	-0.22	0.25	0.25	0.26	0.16	-0.34	-0.0	0.31	0.07	0.0
CMCC-CM2-HR4	-0.29	-0.58	-0.54	-0.07	-0.14	0.02	0.04	0.06	-0.07	0.27	0.16	-0.16	-0.29	0.19	0.38	0.12	0.23	0.92	-0.25	0.0	0.24	-0.4	0.0
CMCC-CM2-SR5	-0.35	-0.14	0.07	0.21	-0.31	-0.05	0.25	0.45	-0.32	-0.39	0.16	-0.3	-0.22	-0.22	0.2	0.14	0.11	0.6	0.32	0.01	0.24	-0.38	0.0
E3SM-1-0	-0.48	-0.1	-0.71	-0.23	-0.92	-0.29	0.16	0.03	0.23	0.06	0.4	0.18	-0.28	0.33	0.41	0.34	0.45	0.33	-0.02	0.28	0.31	0.01	0.0
EC-Earth3	-0.53	-0.31	0.08	-0.05	0.04	-0.08	0.1	0.22	0.0	-0.42	0.28	-0.08	0.12	0.06	0.17	0.23	0.06	0.55	0.09	0.22	0.18	-0.08	0.0
EC-Earth3-Veg	-0.3	-0.22	0.23	-0.22	-0.15	-0.13	0.05	-0.37	0.03	-0.14	0.28	-0.07	-0.14	-0.18	0.25	0.27	0.19	0.49	0.18	0.45	0.22	-0.01	0.0
EC-Earth3-AerChem	-0.13	-0.24	0.28	-0.23	-0.39	-0.28	0.03	-0.12	-0.14	-0.11	0.24	-0.02	-0.14	0.34	-0.07	-0.23	0.66	0.34	0.31	0.43	0.24	-0.01	0.0
GFDL-CM4	-0.58	-0.5	-0.43	-0.38	-0.35	-0.14	0.19	-0.71	0.16	-0.08	0.2	0.18	-0.21	-0.03	0.42	0.02	0.28	0.73	0.19	0.75	0.33	0.05	0.0
MPI-ESM1-2-HR	-0.37	-0.44	-0.38	-0.4	-0.6	-0.32	0.45	-0.56	-0.35	0.01	0.17	-0.08	0.15	0.26	0.53	0.34	0.22	0.92	0.47	0.46	0.37	-0.1	0.0
MRI-ESM2-0	-0.22	-0.38	-0.21	-0.72	-0.24	-0.26	0.36	-0.01	-0.19	0.03	0.09	-0.01	-0.04	-0.06	0.27	0.3	0.37	0.81	0.19	0.43	0.26	-0.04	0.0
NorESM2-MM	-0.16	-0.3	-0.03	-0.3	-0.08	-0.03	-0.01	0.09	-0.09	-0.29	0.09	0.12	0.04	0.01	0.36	-0.0	0.33	0.26	-0.09	0.47	0.16	-0.07	0.0
SAM0-UNICON	-0.22	-0.39	0.05	-0.11	-0.21	-0.08	0.23	0.14	0.04	-0.25	0.08	-0.1	-0.04	-0.15	0.39	-0.22	0.17	0.55	0.09	0.0	0.18	-0.39	0.0
TaiESM1	-0.43	-0.26	0.27	-0.44	-0.27	0.42	-0.32	0.26	-0.16	0.5	-0.0	-0.42	-0.11	0.0	0.36	0.2	0.7	-0.35	0.02	0.29	-0.42	0.0	
ACCESS-CM2	-0.3	-0.13	-0.06	-0.35	-0.11	-0.05	0.21	0.09	-0.01	-0.19	0.12	-0.21	0.14	-0.19	0.0	-0.03	0.3	0.47	0.33	0.29	0.18	-0.31	0.0
ACCESS-ESM1-5	-0.62	-0.38	-0.73	-0.17	-0.54	-0.21	0.53	0.05	0.02	-0.49	0.23	-0.0	0.12	-0.05	0.24	-0.14	0.54	0.79	0.43	0.42	0.33	-0.46	0.0
AWI-ESM-1-1-LR	-0.73	-0.74	-0.87	-0.42	-0.79	-0.15	-0.02	-0.35	-0.21	-0.15	0.31	0.27	-0.13	-0.07	0.49	0.36	0.69	0.91	0.76	0.53	0.45	-0.29	0.0
EC-Earth3-Veg-LR	-0.38	-0.46	-0.19	-0.32	-0.38	0.29	0.11	-0.21	0.13	-0.18	0.3	-0.16	-0.08	0.13	0.47	0.08	0.25	0.75	0.08	0.54	0.27	-0.19	0.0
GISS-E2-1-G	-0.79	-0.36	0.5	0.07	-0.22	-0.34	-0.24	0.49	-0.26	-0.64	0.14	-0.35	0.14	0.42	-0.14	0.02	0.54	1.11	0.2	0.32	0.37	-0.66	0.0
IPSL-CM6A-LR	-0.42	-0.61	-0.35	-0.35	-0.86	-0.13	0.16	-0.01	0.11	0.18	0.26	-0.02	-0.0	-0.06	0.43	0.22	0.33	0.73	0.07	0.35	0.28	-0.01	0.0
KIOST-ESM	-0.17	-0.06	0.26	0.15	-0.76	-0.28	0.09	0.2	-0.1	-0.23	0.18	0.06	-0.33	0.17	0.34	0.05	0.61	0.69	-0.21	-0.01	0.25	-0.46	0.0
MIROC6	-0.19	-0.05	-0.11	-0.36	-0.81	-0.54	0.09	-0.09	-0.23	0.13	0.18	-0.04	0.05	0.05	0.53	0.0	0.2	0.54	0.37	0.57	0.26	-0.1	0.0
MPI-ESM-1-2-HAM	-0.49	-0.27	-0.85	-0.65	-0.26	-0.56	0.06	0.21	-0.37	0.01	0.31	0.17	-0.44	-0.33	0.23	-0.04	0.69	1.27	0.42	0.68	0.42	-0.54	0.0
MPI-ESM1-2-LR	-0.38	-0.44	-0.88	-0.36	-0.25	-0.11	-0.01	-0.14	-0.24	-0.12	0.45	0.11	-0.12	0.23	0.06	0.39	0.62	0.33	0.55	0.11	0.3	-0.31	0.0
NESM3	-0.52	-0.91	-0.89	-0.25	-0.81	-0.08	-0.1	-0.05	-0.41	-0.33	0.21	0.05	0.28	-0.03	0.68	0.32	0.74	1.42	-0.3	0.53	0.45	-0.32	0.0
NorESM2-LM	-0.49	-0.5	-0.54	-0.17	-0.51	-0.04	0.09	-0.24	-0.42	-0.18	0.65	-0.24	0.05	0.2	0.46	0.48	0.67	0.91	0.2	0.07	0.35	-0.49	0.0
CMIP6 Mean	-0.4	-0.34	-0.23	-0.24	-0.45	-0.19	0.17	-0.04	-0.11	-0.15	0.25	0.02	-0.06	0.03	0.29	0.15	0.37	0.68	0.18	0.34	0.23	-0.24	0.0

Figure 8. Same as Figure 5, except showing the error in extratropical cyclone (ETC) frequency as the average ETC hits per day that is, discrete count of 6-hourly ETC snapshots per node divided by the number of days in that node. Mean pattern error (MPE), absolute total error (ATE), and normalized total error (NTE) are shown in the last three columns. Bold, italicized values indicate significance at the 10% level (5th and 95th percentiles) and underlined values indicate significance at the 5% level (2.5th and 97.5th percentiles).

Although there is no direct comparison for the SOM patterns, previous CMIP (Phase 5 and Phase 6) studies investigating mean synoptic conditions can provide some insight into the source of these ETC errors. In agreement with these results, many studies have found that ETC frequency is generally underestimated by the climate models in both CMIP5 and CMIP6, with a tendency for them to be weaker cyclones with a storm track that is too zonal (Priestley et al., 2020; Zappa et al., 2013). This work has provided further insight into identifying the specific patterns which models struggle to reproduce in terms of ETC dynamics. Priestley et al. (2020) also considered high intensity, bomb cyclones in their analysis, which is most analogous to the ETCs seen in the

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	MPE	ATE	NTE
ERA5	19.96	19.23	22.75	24.21	25.25	20.58	20.31	21.36	23.96	23.2	19.07	19.82	19.98	22.55	22.24	20.13	22.04	23.25	24.46	23.35	21.89	22.22	22.22
JRA-55	-0.48	0.11	0.0	0.24	0.08	0.28	-0.68	-1.05	-0.19	0.11	0.64	0.08	-0.13	0.39	0.04	0.18	-0.15	0.56	-0.3	0.02	0.28	0.49	0.0
MERRA2	0.15	-0.18	-0.27	0.33	0.13	0.11	-0.72	-0.57	0.49	0.55	0.17	-0.07	-0.65	0.59	0.27	-0.01	-0.05	-0.83	-0.19	0.47	0.34	0.61	0.0
CFSR	-0.24	-0.09	0.07	-0.28	-0.06	0.64	-0.01	-0.68	0.06	0.18	0.66	-0.23	0.2	-0.01	0.34	0.49	-0.45	-0.3	-0.16	0.34	0.27	0.28	0.0
20CRv3	-0.69	0.5	0.52	1.13	0.33	0.44	0.63	-0.56	-1.47	0.43	-0.52	-0.67	-0.22	-1.47	0.02	-0.0	-0.46	-1.2	1.13	0.37	0.64	0.76	0.0
BCC-CSM2-HR	-1.16	-0.14	-0.77	2.01	0.21	-0.99	-2.95	-1.25	-1.34	-0.26	1.94	-0.81	-0.93	-0.47	0.28	2.29	-0.14	0.18	1.17	-0.15	0.97	0.71	0.0
CMCC-CM2-VHR4	-1.4	0.44	-1.83	-0.82	-1.92	-0.52	-2.31	0.25	-1.36	-1.91	0.92	0.98	2.09	1.22	1.51	2.4	1.21	2.23	-0.22	3.22	1.44	1.02	0.0
E3SMHR	-1.16	0.04	-0.79	-0.57	-2.51	-0.5	0.26	-1.2	-1.88	0.46	3.13	1.56	1.85	-1.59	0.78	0.72	1.04	-0.08	0.12	0.65	1.04	1.73	0.0
ECMWF-IFS-HR	-1.51	0.78	-1.43	1.04	-0.03	0.58	-1.58	-1.78	0.51	0.62	-0.71	0.31	2.03	1.37	-0.13	1.68	-0.28	-0.72	-1.29	1.74	1.01	2.05	0.0
ECMWF-IFS-MR	-1.9	-0.99	-1.46	1.45	0.38	-0.6	-1.52	-0.32	-1.53	-1.28	2.18	-0.41	1.31	-0.92	0.66	0.95	1.77	0.71	1.08	0.75	1.11	2.39	0.0
ECMWF-IFS-LR	-0.97	0.74	0.03	0.99	-0.72	-1.13	-2.28	-0.2	-1.33	-0.51	2.49	1.04	1.06	0.7	0.4	1.09	1.15	0.48	-0.5	0.14	0.9	2.51	0.0
GFDL-CM4C192	-1.84	0.86	-1.27	0.77	-2.42	-0.01	-0.32	-1.41	2.76	0.01	1.57	0.09	-0.81	-0.68	1.0	1.42	0.33	0.16	1.26	1.2	1.01	0.71	0.0
MPI-ESM1-2-XR	-1.21	-1.11	-0.82	0.9	-0.33	-0.24	-1.36	-1.24	-0.41	-0.29	0.74	0.48	0.97	-0.71	0.41	0.12	0.4	0.35	0.03	1.24	0.67	3.31	0.0
HRMIP Mean	-1.41	0.14	-1.1	0.74	-1.01	-0.39	-1.53	-0.94	-0.49	-0.38	1.64	0.32	1.06	-0.17	0.64	1.3	0.62	0.48	0.27	1.11	0.79	1.79	0.0
BCC-CSM2-MR	-1.65	-1.06	-3.95	-1.74	-1.5	-0.78	-1.94	-0.76	-3.26	-0.83	2.52	2.46	-0.02	0.97	3.13	2.0	0.24	0.42	-0.93	3.02	1.66	1.7	0.0
CESM2	1.11	0.57	-2.05	0.29	-2.64	-0.26	-0.69	-1.33	-2.5	0.86	2.18	1.11	2.31	-0.51	0.79	1.64	0.48	0.85	1.39	1.0	1.23	0.71	0.0
CMCC-CM2-HR4	-1.21	-0.24	-3.09	-0.48	-2.62	0.13	-1.61	-0.17	-3.01	-1.43	2.84	-0.4	1.86	-1.01	0.8	1.63	3.31	2.16	0.49	0.13	1.43	-0.74	0.0
CMCC-CM2-SR5	-1.03	0.32	1.47	0.47	-0.25	-1.61	-1.01	-1.19	-2.11	-0.24	-0.28	-2.62	0.45	-0.51	1.06	0.39	0.19	1.37	-0.13	2.22	0.95	-0.85	0.0
E3SM-1-0	-0.44	0.86	-1.22	-0.51	-0.58	-0.81	-0.52	-0.93	-1.38	-0.2	1.82	0.93	0.86	-0.56	1.43	2.31	1.18	-0.11	-0.25	-1.16	0.9	0.04	0.0
EC-Earth3	-0.59	-0.25	-2.33	0.56	-0.21	-1.04	-1.29	-1.06	0.92	0.11	1.59	0.15	0.71	-0.86	0.73	2.28	-0.17	-0.22	0.49	0.33	0.79	1.21	0.0
EC-Earth3-Veg	-2.55	-1.16	-0.55	0.75	-0.18	-2.07	-1.74	-1.8	-0.56	-0.03	1.71	0.13	0.47	-0.93	1.94	1.08	1.08	1.69	-0.47	1.68	1.13	1.32	0.0
EC-Earth3-AerChem	-1.73	2.2	-2.64	-0.68	0.83	-1.76	-2.34	-1.0	-1.13	-0.75	2.19	0.65	0.88	-0.11	2.76	-0.04	0.48	-0.29	1.41	0.73	1.23	1.52	0.0
GFDL-CM4	0.38	0.6	-1.52	0.06	-0.45	-0.87	-2.0	-1.95	0.37	-0.41	0.53	-0.59	-0.72	-0.47	1.24	1.56	0.5	-0.13	-1.28	0.88	0.83	0.26	0.0
MPI-ESM1-2-HR	-1.24	-0.73	0.27	1.12	-0.84	-0.62	-1.32	-1.01	-0.24	0.78	0.8	-0.01	0.36	-0.24	0.65	0.45	-0.02	0.88	-1.33	-0.26	0.66	2.89	0.0
MRI-ESM2-0	-1.24	-0.12	-0.7	-1.24	0.15	-1.23	-2.37	-0.74	-0.17	-0.56	0.54	-0.13	0.76	0.51	1.04	0.09	0.86	1.55	0.26	0.75	0.75	1.96	0.0
NorESM2-MM	-0.94	-0.19	-1.12	0.65	-0.2	0.92	-1.85	-1.77	0.21	-1.27	0.81	-0.3	1.14	0.31	0.12	0.63	0.5	-0.31	0.75	1.56	0.78	0.11	0.0
SAM0-UNICON	0.38	0.59	-2.28	-0.7	-2.29	0.28	-1.91	-1.32	-1.6	-0.79	1.09	-0.99	0.75	1.14	0.57	1.08	1.75	1.14	0.99	1.43	1.15	-0.85	0.0
TaiESM1	-0.93	-0.61	0.4	1.11	-1.08	0.35	-1.19	-0.1	-1.55	-0.73	1.28	0.02	0.02	-0.38	1.55	0.35	-0.39	0.89	1.75	-0.1	0.74	-1.23	0.0
ACCESS-CM2	-0.24	1.11	-0.73	0.42	0.79	-1.33	-0.76	-0.27	-0.43	0.96	-1.53	0.13	0.9	-1.27	0.98	-1.01	-0.68	0.14	0.74	0.08	0.73	1.68	0.0
ACCESS-ESM1-5	-1.63	-0.8	-1.6	0.34	-1.12	-1.64	-1.51	0.42	-2.2	0.25	0.57	0.31	1.88	0.61	1.18	1.5	1.47	0.23	0.04	0.22	0.98	-0.4	0.0
AWI-ESM-1-1-LR	-2.07	-0.77	-2.71	-1.84	-1.49	0.66	-1.4	-0.86	-1.09	-0.03	2.03	0.42	0.86	0.76	1.98	1.06	1.19	0.53	-0.8	-0.06	1.13	1.73	0.0
EC-Earth3-Veg-LR	-1.8	-1.19	-1.27	0.92	-1.79	-0.41	-1.69	-1.17	-0.22	0.12	0.6	0.25	0.62	-1.29	-0.87	0.47	1.34	1.48	0.54	1.01	0.95	1.06	0.0
GISS-E2-1-G	-1.33	-0.06	-1.32	1.75	0.13	-0.95	-1.83	-0.67	-0.22	-1.86	3.02	-0.36	0.12	-0.21	0.64	-0.49	-0.32	-0.06	1.41	0.79	0.88	1.86	0.0
IPSL-CM6A-LR	-1.87	-2.09	-2.03	-1.42	-1.2	-0.8	-1.3	-1.13	-2.59	0.21	2.76	-0.34	0.1	-0.08	0.18	0.53	2.0	1.5	0.89	1.72	1.24	1.43	0.0
KIOST-ESM	-1.25	0.56	0.97	1.74	0.07	-0.48	-0.78	0.94	-0.44	1.03	1.81	-0.35	1.67	-1.87	-1.94	1.65	0.82	0.91	-1.7	-1.81	1.14	-3.36	0.0
MIROC6	0.28	2.09	1.67	0.9	0.51	-1.75	-0.39	0.96	0.43	0.45	1.17	-1.94	0.37	0.43	0.03	0.76	-0.29	-1.93	-0.77	-0.75	0.89	0.06	0.0
MPI-ESM-1-2-HAM	-1.41	0.36	-0.85	0.57	-1.02	-2.71	-2.4	-0.66	0.95	-0.04	0.84	-0.78	-1.14	-0.98	1.67	0.27	1.13	0.91	-0.18	0.59	0.97	0.23	0.0
MPI-ESM1-2-LR	-0.59	1.12	0.58	-0.23	-1.02	-1.8	-1.5	-0.84	-1.17	-0.77	1.5	1.54	1.05	-0.22	0.46	1.78	-0.64	-0.57	0.28	2.28	1.0	2.03	0.0
NESM3	-1.05	-0.6	-0.11	0.07	-1.66	-0.58	-1.74	0.31	-2.03	-0.49	2.15	0.61	0.78	-2.24	0.9	1.9	0.66	-0.11	-1.11	1.92	1.05	3.35	0.0
NorESM2-LM	-1.04	-0.64	-0.83	2.02	-1.0	-1.25	-1.62	0.18	-0.54	0.08	3.08	-0.87	0.28	-1.3	0.35	1.29	-0.14	-0.78	1.37	1.36	1.0	-0.44	0.0
CMIP6 Mean	-0.99	-0.0	-1.02	0.15	-0.86	-0.86	-1.56	-0.69	-1.01	-0.21	1.44	0.07	0.73	-0.39	0.89	1.09	0.72	0.54	0.18	0.76	0.71	0.72	0.0

Figure 9. Same as Figure 5, except showing the error in extratropical cyclone (ETC) intensity as the average accumulated cyclone activity (ACA) (hPa) per ETC hit that is, total ACA per node divided by the number of 6-hourly ETC hits in that node. Mean pattern error (MPE), absolute total error (ATE), and normalized total error (NTE) are shown in the last three columns. Bold, italicized values indicate significance at the 10% level (5th and 95th percentiles) and underlined values indicate significance at the 5% level (2.5th and 97.5th percentiles).

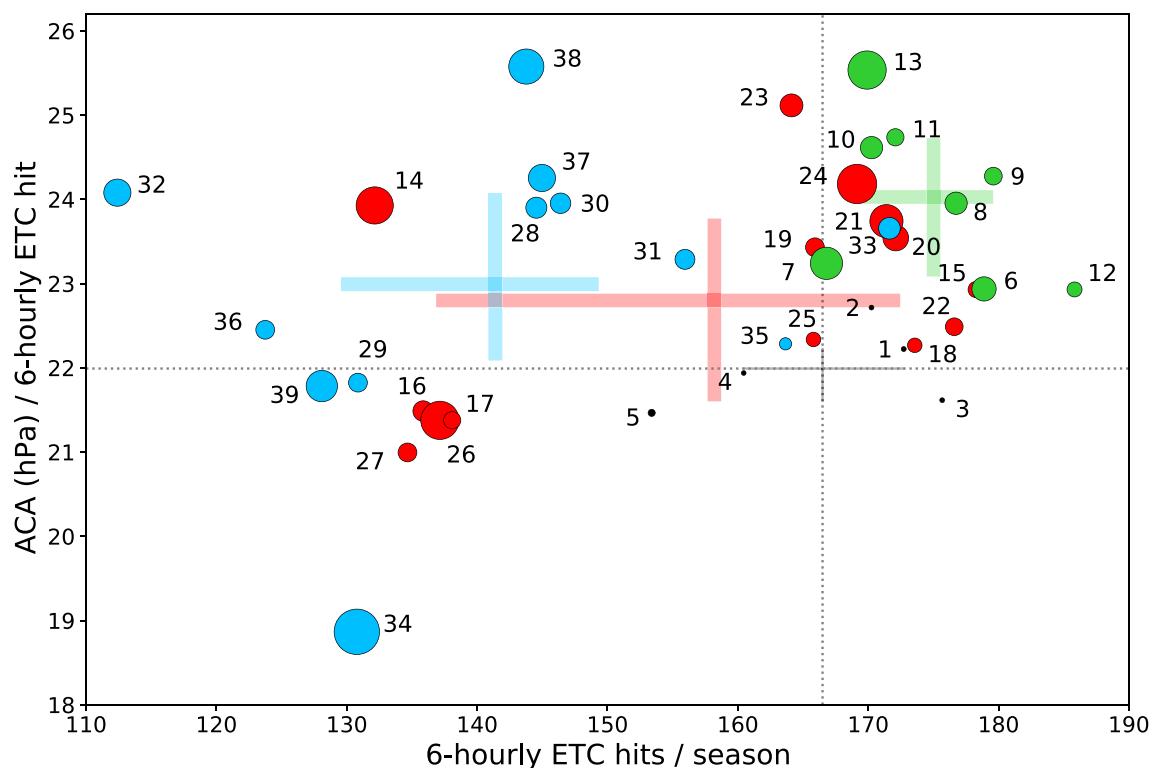


Figure 10. Averaged extratropical cyclone (ETC) hits per winter season (x-axis) and averaged accumulated cyclone activity per ETC hit (y-axis) associated with reanalyses (black) and CMIP6 climate models (blue = 250 km resolution; red = 100 km resolution; green = HighResMIP). The size of the dots varies according to the mean pattern error in node frequency as measured by the self-organizing map (second last column in Figure 5). The crossbars indicate the interquartile range for each model category.

Project (AMIP)—which apply specified sea surface temperatures, thereby eliminating errors associated with ocean state biases—concluding that the errors are likely a result of biases in the atmospheric dynamics.

A compact synthesis of these errors is shown in Figure 10, which plots the average error in ETC frequency and intensity over NEUS associated with each of the reanalysis products, HighResMIP climate models, and CMIP6 climate models. The discrete count of 6-hourly ETC hits per cold season is shown on the x-axis and is proportional to ATE in Figure 8, whilst the average ACA per ETC hit is shown on the y-axis and is proportional to the ATE in Figure 9. Here we are looking at biases in the mean ETC statistics and have not normalized relative to ERA5. The black dots represent the ETC frequency and intensity of each of the five reanalysis products, with the black dotted lines showing the average ETC frequency and intensity from the reanalyses. The CMIP6 climate models are represented by colored dots, namely the 250 km resolution CMIP models (blue), 100 km resolution CMIP models (red), and the HighResMIP models (green). In addition to ETC dynamics, the size of the dots varies according to the model's MPE in the SOM node frequency (Figure 5), indicating its ability to simulate the occurrence of large-scale patterns. Therefore, models with larger dots have difficulty reproducing the frequency of LSMPs identified in the SOM. The crosses indicate the interquartile range of ETC frequency and intensity for the reanalysis products and each model group, whilst the thickness indicates the mean node frequency error for each group. This provides a benchmark for comparison; the closer a model is to the intersection of the black dotted lines, and the smaller the associated dot, the better its overall performance relative to reanalyses.

We generally see that as resolution increases from 250 km to 100 km to 25 km, there is an increase in the number of ETCs simulated (x-axis), with the HighResMIP models being more skillful in this regard. While ACA per ETC hit (y-axis) is not overly different between 250 and 100 km resolution CMIP6 models, we do see a high intensity bias (per storm) develop in the ETCs simulated by the HighResMIP models. We interpret this to show that while there is improved simulation of ETC frequency in the HighResMIP models, there are some significant positive errors in terms of ETC intensity (ACA). Lastly, there is little difference in the node frequency error (size of dots) between the low resolution (250 km), standard resolution (100 km), and high resolution (25–50 km) CMIP6

climate models signifying that resolution has minimal influence on the simulation of the large-scale patterns themselves as measured by the SOM, rather factors such as model formulation and forcing exert a larger influence. This suggests that differences in ETC climatology in models with finer grid spacing relative to their coarser counterparts arise from the increased resolution of the atmosphere near the storm scale.

4. Conclusion

This paper used SOMs to characterize the LSMPs of winter Z500 anomalies and the associated discrete ETC events over NEUS using reanalysis data sets. Differing from previous studies (e.g., Agel et al., 2018) which filter the data by extreme events before applying the SOM analysis, this work applies a top-down approach, first identifying the large-scale synoptic conditions and then tying in ETC events to identify LSMPs associated with extreme events. The results show that no single pattern dominates the winter climatology and categorized patterns are representative of Rossby wave propagation. Whilst ETCs occur in all patterns, the patterns which are characterized by amplified meridional flow tend to be associated with more frequent ETCs, particularly those associated with a mid-level trough, in agreement with established synoptic theory. This result is accentuated when analyzing ACA, which takes into consideration both ETC frequency and intensity. Here, the trough patterns stand out as the more stormy nodes. This agrees with what has been typically referenced in the literature (e.g., Dacre et al., 2012) through case studies and manual synoptic typing, which highlights the ability and suitability of the SOM analysis for identifying LSMPs. In addition, there is generally strong agreement between the different reanalysis products in representing the LSMPs (although less agreement in representing cyclone frequency), providing further confidence and robustness to the technique described here.

The SOM patterns were then applied to CMIP6 and HighResMIP climate models to evaluate their ability in reproducing the synoptic patterns associated with ETCs. The CMIP6 models tend to favor the more zonal patterns, with increased node frequency in the interior nodes and decreased node frequency associated with the exterior nodes. There is also higher spatial error associated with the exterior nodes, indicating that the models particularly struggle to reproduce the meridionally amplified (i.e., extreme) patterns. Overall, the CMIP6 models underestimate ETC frequency with mixed performance reproducing ETC intensity. Moreover, they particularly struggle to reproduce ETC frequency in the patterns associated with the highest cyclone activity (nodes 5 and 20). These errors tend to be systematic as they are seen across all of the CMIP6 climate models, with only a few exceptions. However, it is worth noting that these ETC errors could vary if another reanalysis product was used as the reference. Generally, there is improved performance in the HighResMIP climate models with respect to cyclone frequency, although the impact of resolution on LSMPs is far more marginal, if not non-existent. While there may be an improvement in overall ETC climatology, ETC intensity is typically exaggerated, which may indicate overparameterization of moist processes important for storm intensification (Seiler et al., 2018; Willison et al., 2013). Providing a breakdown of errors across different LSMPs helps to pinpoint and inform model developers of the exact environments with which climate models struggle to reproduce. It can also reveal potential errors which would have otherwise been lost when considering mean-state biases. Climate model evaluations are fundamental to improving model performance and increasing confidence in future climate projections, which are crucial for policymakers and the future of many economic sectors, especially when considering potentially devastating events such as ETCs.

It is worth noting that understanding the physical mechanisms behind these biases at the individual model level is beyond the scope of this work. Further, it is not our intention to rank models or define particular metrics as more desirable than others when evaluating model credibility. We do aim to provide a process-level evaluation of models that provides insight into the dynamical behavior that goes beyond traditional analyses that assess mean climatology. Further investigation is needed to identify the cause of these errors, as it has been shown that errors are not consistent across the different LSMPs. Analyzing differences in the thermodynamics and moisture transport of each pattern could provide useful insight, as well as exploring discrepancies in ETC cyclogenesis and trajectories. The role of the ocean model in producing these biases is also not clear. Comparing these results with historical experiments that follow AMIP protocols by applying prescribed sea surface temperature fields could provide further insight into isolating the role of the ocean and atmospheric components in producing these CMIP6 biases. Nonetheless, it is easy to envision model developers using this approach to reduce model biases by targeting model components most closely tied to said errors.

It is important to note some caveats of this work. First is the subjective nature of selecting the SOM parameters, however, these parameters generally have minimal influence on the final SOM product when sufficiently trained (Gervais et al., 2020; Gibson et al., 2017; Grotjahn et al., 2016). The SOM size could exert influence on the

results as there is an interplay between node frequency error and spatial pattern error and the optimal size varies based on the problem at hand. Second, only one cyclone tracking algorithm was used in this study. Whilst Neu et al. (2013) generally found consistency between tracking methods for strong ETCs, testing other algorithms or implementing an ensemble approach could be beneficial. Third, while several reanalyses products were used in this study, they are all representing just one realization of the climate. Therefore, it is possible that there are model biases that are due to internal variability and not necessarily model errors. Similarly, we only used one ensemble member for each climate model which could be under-/overestimating the significance and limits our confidence in the model results. Future work should consider using large ensemble climate model simulations if and when available, especially if this analysis is to be used for improving model formulations. Nevertheless, this work has shown that SOMs are a useful tool for linking discrete ETC events to synoptic conditions and demonstrates the potential for this technique to be applied to other traceable phenomena (e.g., tropical cyclones, atmospheric rivers, mesoscale convective systems, etc.) around the world.

Data Availability Statement

ERA5 reanalysis is available through the Copernicus Climate Data Store (CDS), accessible at Hersbach et al. (2023). JRA-55, CFSR, and 20CRv3 reanalyses are available through the Research Data Archive at the NCAR Computational and Information Systems Laboratory (CISL). JRA-55 is accessible at JMA (2013), CFSR 1980–2010 is accessible at Saha et al. (2010b), and CFSR 2011–2019 is accessible at Saha et al. (2011). 20CRv3 reanalysis is supported by the U.S. Department of Energy, Office of Science Biological and Environmental Research, by the NOAA Climate Program Office, and by the NOAA Earth System Research Laboratory Physical Sciences Laboratory, accessible at Slivinski et al. (2019a). MERRA2 reanalysis is available through the Modeling and Assimilation Data and Information Services Center (MDISC), managed by the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC), accessible at GMAO (2015). CMIP6 data (CMIP and HighResMIP) is publicly available through the Department of Energy-supported Earth System Grid Federation (<https://esgf-node.llnl.gov/projects/cmip6/>). The SOM software package is also publicly available from Vesanto et al. (2000). In this manuscript, ETCs are tracked using TempestExtremes which is available from Ullrich (2022). All data and scripts used for the analysis in this manuscript are available for download on Zenodo (Gore et al., 2023).

Acknowledgments

Gore and Zarzycki were supported by the Regional and Global Modeling and Analysis (RGMA) program area of the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research as part of the multi-program, collaborative Integrated Coastal Modeling project. Additional funding for Zarzycki to develop ETC datasets over the United States came from NOAA MAPP award NA19OAR4310288 “Future changes in the frequency of winter snowstorms and their impact on snowfall and snow water equivalent.” The authors acknowledge the World Climate Research Program, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups (listed in Table 1) 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 authors thank two anonymous reviewers for their feedback that improved this manuscript.

References

- Agel, L., Barlow, M., Feldstein, S., & Gutowski, W. (2018). Identification of large-scale meteorological patterns associated with extreme precipitation in the US northeast. *Climate Dynamics*, 50(5–6), 1819–1839. <https://doi.org/10.1007/s00382-017-3724-8>
- Agel, L., Barlow, M., Qian, J., Colby, F., Douglas, E., & Eichler, T. (2015). Climatology of daily precipitation and extreme precipitation events in the northeast United States. *Journal of Hydrometeorology*, 16(6), 2537–2557. <https://doi.org/10.1175/jhm-d-14-0147.1>
- Babaei, M., Alizadeh, O., & Irandejad, P. (2021). Impacts of orography on large-scale atmospheric circulation: Application of a regional climate model. *Climate Dynamics*, 57(7–8), 1–20. <https://doi.org/10.1007/s00382-021-05790-0>
- Bader, D. C., Leung, R., Taylor, M., & McCoy, R. B. (2019). *E3SM-Project E3SM1.0 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4497>
- Barlow, M., Gutowski, W., Gyakum, J., Katz, R., Lim, Y., Schumacher, R., et al. (2019). North American extreme precipitation events and related large-scale meteorological patterns: A review of statistical methods, dynamics, modeling, and trends. *Climate Dynamics*, 53(11), 6835–6875. <https://doi.org/10.1007/s00382-019-04958-z>
- Bentsen, M., Olivie, D. J. L., Seland, y., Tonizzo, T., Gjermundsen, A., & Graff, L. S. (2019). *NCC NorESM2-MM model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.8040>
- Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L., & Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: Influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, 14(1), 34–40. <https://doi.org/10.1002/asl2.412>
- Boucher, O., Deniel, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., et al. (2018). *IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.5195>
- Brereton, R. (2012). Self organising maps for visualising and modelling. *Chemistry Central Journal*, 6(S1), S1. <https://doi.org/10.1186/1752-153x-6-s2-s1>
- Bu, L., Zuo, Z., & An, N. (2022). Evaluating boreal summer circulation patterns of cmip6 climate models over the Asian region. *Climate Dynamics*, 58(1–2), 1–15. <https://doi.org/10.1007/s00382-021-05914-6>
- Caldwell, P. M., Mametjanov, A., Tang, Q., Van Roekel, L. P., Golaz, J.-C., Lin, W., et al. (2019). The DOE E3SM coupled model version 1: Description and results at high resolution. *Journal of Advances in Modeling Earth Systems*, 11(12), 4095–4146. <https://doi.org/10.1029/2019ms001870>
- Cao, J., & Wang, B. (2019). *NUIST NESMv3 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.8769>
- Cassano, J., Uotila, P., & Lynch, A. (2006). Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part 1: Arctic. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 26(8), 1027–1049. <https://doi.org/10.1002/joc.1306>
- Catalano, A. J., & Broccoli, A. J. (2018). Synoptic characteristics of surge-producing extratropical cyclones along the northeast coast of the United States. *Journal of Applied Meteorology and Climatology*, 57(1), 171–184. <https://doi.org/10.1175/jamc-d-17-0123.1>

- Catto, J. (2016). Extratropical cyclone classification and its use in climate studies. *Reviews of Geophysics*, 54(2), 486–520. <https://doi.org/10.1002/2016rg000519>
- Cavazos, T. (2000). Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans. *Journal of Climate*, 13(10), 1718–1732. [https://doi.org/10.1175/1520-0442\(2000\)013<1718:usomti>2.0.co;2](https://doi.org/10.1175/1520-0442(2000)013<1718:usomti>2.0.co;2)
- Colle, B. A., Booth, J. F., & Chang, E. K. (2015). A review of historical and future changes of extratropical cyclones and associated impacts along the US east coast. *Current Climate Change Reports*, 1(3), 125–143. <https://doi.org/10.1007/s40641-015-0013-7>
- Colle, B. A., Zhang, Z., Lombardo, K. A., Chang, E., Liu, P., & Zhang, M. (2013). Historical evaluation and future prediction of eastern North American and western Atlantic extratropical cyclones in the CMIP5 models during the cool season. *Journal of Climate*, 26(18), 6882–6903. <https://doi.org/10.1175/jcli-d-12-00498.1>
- Dacre, H., Hawcroft, M., Stringer, M., & Hodges, K. (2012). An extratropical cyclone atlas: A tool for illustrating cyclone structure and evolution characteristics. *Bulletin of the American Meteorological Society*, 93(10), 1497–1502. <https://doi.org/10.1175/bams-d-11-00164.1>
- Danabasoglu, G. (2019). NCAR CESM2 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.7627>
- Danek, C., Shi, X., Stepanek, C., Yang, H., Barbi, D., Hegewald, J., & Lohmann, G. (2020). AWI AWI-ESM1.ILR model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.9328>
- Davis, R. E., Demme, G., & Dolan, R. (1993). Synoptic climatology of Atlantic coast north-easters. *International Journal of Climatology*, 13(2), 171–189. <https://doi.org/10.1002/joc.3370130204>
- Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., Harman, I., & Law, R. (2019). CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4271>
- EC-Earth (2019a). EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4700>
- EC-Earth (2019b). EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4706>
- EC-Earth (2020a). EC-Earth-Consortium EC-Earth3-AerChem model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4701>
- EC-Earth (2020b). EC-Earth-Consortium EC-Earth3-Veg-LR model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4707>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model inter-comparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Fasullo, J. T. (2020). Evaluating simulated climate patterns from the CMIP archives using satellite and reanalysis datasets using the Climate Model Assessment Tool (CMATv1). *Geoscientific Model Development*, 13(8), 3627–3642. <https://doi.org/10.5194/gmd-13-3627-2020>
- Gervais, M., Atallah, E., Gyakum, J., & Tremblay, L. (2016). Arctic air masses in a warming world. *Journal of Climate*, 29(7), 2359–2373. <https://doi.org/10.1175/jcli-d-15-0499.1>
- Gervais, M., Shaman, J., & Kushnir, Y. (2020). Impact of the North Atlantic warming hole on sensible weather. *Journal of Climate*, 33(10), 4255–4271. <https://doi.org/10.1175/jcli-d-19-0636.1>
- Gibson, P., Perkins-Kirkpatrick, S., Uotila, P., Pepler, A., & Alexander, L. (2017). On the use of self-organizing maps for studying climate extremes. *Journal of Geophysical Research: Atmospheres*, 122(7), 3891–3903. <https://doi.org/10.1002/2016jd026256>
- GMAO (2015). MERRA-2 inst3_3d_asm_Np: 3d, 3-hourly, instantaneous, pressure-level, assimilation, assimilated meteorological fields v5.12.4 [Dataset]. Global Modeling and Assimilation Office, Goddard Earth Sciences Data and Information Services Center (GES DISC). <https://doi.org/10.5067/QBZ6MG944HW0>
- Gore, M. J., Zarzycki, C. M., & Gervais, M. M. (2023). Data from “Connecting large-scale meteorological patterns to extratropical cyclones in CMIP6 climate models using self-organizing maps” [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.8179303>
- Grotjahn, R., Black, R., Leung, R., Wehner, M., Barlow, M., Bosilovich, M., et al. (2016). North American extreme temperature events and related large scale meteorological patterns: A review of statistical methods, dynamics, modeling and trends. *Climate Dynamics*, 46(1151–1184), 2638–2646. <https://doi.org/10.1007/s00382-015-2638-6>
- Gu, Q., & Gervais, M. (2021). Exploring North Atlantic and North Pacific decadal climate prediction using self-organizing maps. *Journal of Climate*, 34(1), 123–141. <https://doi.org/10.1175/jcli-d-20-0017.1>
- Gu, Q., & Gervais, M. (2022). Diagnosing two-way coupling in decadal north Atlantic SST variability using time-evolving self-organizing maps. *Geophysical Research Letters*, 49(8), e2021GL096560. <https://doi.org/10.1029/2021gl096560>
- Guo, H., John, J. G., Blanton, C., McHugh, C., Nikonorov, S., Radhakrishnan, A., et al. (2018). NOAA-GFDL GFDL-CM4 model output. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.1402>
- Guo, Y., Shinoda, T., Lin, J., & Chang, E. K. (2017). Variations of northern hemisphere storm track and extratropical cyclone activity associated with the Madden-Julian Oscillation. *Journal of Climate*, 30(13), 4799–4818. <https://doi.org/10.1175/jcli-d-16-0513.1>
- Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al. (2016). High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6. *Geoscientific Model Development*, 9(11), 4185–4208. <https://doi.org/10.5194/gmd-9-4185-2016>
- Harvey, B., Cook, P., Shaffrey, L., & Schiemann, R. (2020). The response of the northern hemisphere storm tracks and jet streams to climate change in the CMIP3, CMIP5, and CMIP6 climate models. *Journal of Geophysical Research: Atmospheres*, 125(23), e2020JD032701. <https://doi.org/10.1029/2020jd032701>
- Hawcroft, M., Shaffrey, L., Hodges, K., & Dacre, H. (2012). How much Northern Hemisphere precipitation is associated with extratropical cyclones? *Geophysical Research Letters*, 39(24). <https://doi.org/10.1029/2012gl053866>
- Herbst, M., & Casper, M. (2008). Towards model evaluation and identification using self-organizing maps. *Hydrology and Earth System Sciences*, 12(2), 657–667. <https://doi.org/10.5194/hess-12-657-2008>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., & Muñoz-Sabater, J. (2023). ERA5 hourly data on pressure levels from 1940 to present [Dataset]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.bd0915c6>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hewitson, B., & Crane, R. (2002). Self-organizing maps: Applications to synoptic climatology. *Climate Research*, 22(1), 13–26. <https://doi.org/10.3354/cr022013>
- Hodges, K. I., Lee, R. W., & Bengtsson, L. (2011). A comparison of extratropical cyclones in recent reanalyses ERA-Interim, NASA MERRA, NCEP CFSR, and JRA-25. *Journal of Climate*, 24(18), 4888–4906. <https://doi.org/10.1175/2011jcli4097.1>
- Hoskins, B. J. (1990). Theory of extratropical cyclones. *Extratropical Cyclones: The Erik Palmén Memorial*, 63–80.

- Hoskins, B. J., & Hodges, K. I. (2002). New perspectives on the northern hemisphere winter storm tracks. *Journal of the Atmospheric Sciences*, 59(6), 1041–1061. [https://doi.org/10.1175/1520-0469\(2002\)059<1041:nptnh>2.0.co;2](https://doi.org/10.1175/1520-0469(2002)059<1041:nptnh>2.0.co;2)
- Jie, W., Zhang, J., Wu, T., Shi, X., Zhang, F., Li, J., & Wei, M. (2020). *BCC BCC-CSM2HR model output prepared for CMIP6 HighResMIP hist-1950*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.2921>
- JMA (2013). *JRA-55: Japanese 55-year reanalysis, daily 3-hourly and 6-hourly data [Dataset]*. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/10.5065/D6HH6H41>
- Jung, T., Miller, M., Palmer, T., Towers, P., Wedi, N., Achuthavarier, D., et al. (2012). High-resolution global climate simulations with the ECMWF model in Project Athena: Experimental design, model climate, and seasonal forecast skill. *Journal of Climate*, 25(9), 3155–3172. <https://doi.org/10.1175/jcli-d-11-00265.1>
- Jungclaus, J., Bittrner, M., Wiemers, K.-H., Wachsmann, F., Schupfner, M., Legutke, S., & Roeckner, E. (2019). *MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.6594>
- Kim, Y., Noh, Y., Kim, D., Lee, M.-I., Lee, H. J., Kim, S. Y., & Kim, D. (2019). *KIOST KIOST-ESM model output prepared for CMIP6 CMIP*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.1922>
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, 93(1), 5–48. <https://doi.org/10.2151/jmsj.2015-001>
- Kocin, P. J., & Uccellini, L. W. (2004). *Northeast snowstorms* (Vol. 644). Springer.
- Kohonen, T. (2001). An overview of SOM literature. In *Self-organizing maps* (pp. 347–371). Springer.
- Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, 37, 52–65. <https://doi.org/10.1016/j.neunet.2012.09.018>
- Koster, R. D., Bosilovich, M. G., Akella, S., Lawrence, C., Cullather, R., Draper, C., et al. (2015). *MERRA-2: Initial evaluation of the climate*. (Tech. Rep. No. 43). NASA.
- Lee, J., Hwang, J., Son, S.-W., & Gyakum, J. R. (2022). Future changes of East Asian extratropical cyclones in the CMIP5 models. *Journal of Climate*, 35(21), 3311–3321. <https://doi.org/10.1175/jcli-d-21-0945.1>
- Lee, W.-L., & Liang, H.-C. (2020). *AS-RCEC TaiESM1.0 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.9755>
- Liu, Y., Weisberg, R., & Mooers, C. (2006). Performance evaluation of the self-organizing map for feature extraction. *Journal of Geophysical Research*, 111(C5), C05018. <https://doi.org/10.1029/2005jc003117>
- Loikith, P. C., Lintner, B. R., & Sweeney, A. (2017). Characterizing large-scale meteorological patterns and associated temperature and precipitation extremes over the northwestern United States using self-organizing maps. *Journal of Climate*, 30(8), 2829–2847. <https://doi.org/10.1175/jcli-d-16-0670.1>
- Lovato, T., & Peano, D. (2020). *CMCC CMCC-CM2-SR5 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.3825>
- Mass, C., & Dotson, B. (2010). Major extratropical cyclones of the northwest United States: Historical review, climatology, and synoptic environment. *Monthly Weather Review*, 138(7), 2499–2527. <https://doi.org/10.1175/2010mwr3213.1>
- NASA/GISS. (2018). *NASA-GISS GISS-E2.1G model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.7127>
- Neu, U., Akperov, M. G., Bellenbaum, N., Benestad, R., Blender, R., Caballero, R., et al. (2013). IMILAST: A community effort to inter-compare extratropical cyclone detection and tracking algorithms. *Bulletin of the American Meteorological Society*, 94(4), 529–547. <https://doi.org/10.1175/bams-d-11-00154.1>
- Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stoll, J., Folini, D. S., Tegen, I., & Lohmann, U. (2019). *HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.5016>
- Park, S., & Shin, J. (2019). *SNU SAM0-UNICON model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.7789>
- Pfahl, S., & Wernli, H. (2012). Quantifying the relevance of cyclones for precipitation extremes. *Journal of Climate*, 25(19), 6770–6780. <https://doi.org/10.1175/jcli-d-11-00705.1>
- Pithan, F., Shepherd, T. G., Zappa, G., & Sandu, I. (2016). Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag. *Geophysical Research Letters*, 43(13), 7231–7240. <https://doi.org/10.1002/2016gl069551>
- Priestley, M. D., Ackerley, D., Catto, J. L., Hodges, K. I., McDonald, R. E., & Lee, R. W. (2020). An overview of the extratropical storm tracks in CMIP6 historical simulations. *Journal of Climate*, 33(15), 6315–6343. <https://doi.org/10.1175/jcli-d-19-0928.1>
- Reusch, D., Alley, R., & Hewitson, B. (2005). Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. *Polar Geography*, 29(3), 188–212. <https://doi.org/10.1080/789610199>
- Roberts, C. D., Senan, R., Molteni, F., Boussetta, S., & Keeley, S. (2017). *ECMWF ECMWF-IFS-HR model output prepared for CMIP6 HighResMIP hist-1950*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4981>
- Roberts, C. D., Senan, R., Molteni, F., Boussetta, S., & Keeley, S. (2018a). *ECMWF ECMWF-IFS-LR model output prepared for CMIP6 HighResMIP hist-1950*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4982>
- Roberts, C. D., Senan, R., Molteni, F., Boussetta, S., & Keeley, S. (2018b). *ECMWF ECMWF-IFS-MR model output prepared for CMIP6 HighResMIP hist-1950*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4983>
- Rohrer, M., Brönnimann, S., Martius, O., Raible, C. C., Wild, M., & Compo, G. P. (2018). Representation of extratropical cyclones, blocking anticyclones, and Alpine circulation types in multiple reanalyses and model simulations. *Journal of Climate*, 31(8), 3009–3031. <https://doi.org/10.1175/jcli-d-17-0350.1>
- Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., & Nadiga, S. (2011). NCEP Climate Forecast System version 2 (CFSv2) 6-hourly products (updated daily) [Dataset]. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/10.5065/D61C1TXF>
- Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., et al. (2010a). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8), 1015–1057. <https://doi.org/10.1175/2010BAMS3001.1>
- Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., et al. (2010b). NCEP climate forecast system reanalysis (CFSR) 6-hourly products, January 1979 to December 2010 [Dataset]. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/10.5065/D69K487J>
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208. <https://doi.org/10.1175/JCLI-D-12-00823.1>
- Scoccimarro, E., Bellucci, A., & Peano, D. (2018). *CMCC CMCC-CM2-VHR4 model output prepared for CMIP6 HighResMIP hist-1950*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.3818>

- Scoccimarro, E., Bellucci, A., & Peano, D. (2020). *CMCC CMCC-CM2-HR4 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.3823>
- Seierstad, I., Stephenson, D., & Kvamstø, N. (2007). How useful are teleconnection patterns for explaining variability in extratropical storminess? *Tellus A: Dynamic Meteorology and Oceanography*, 59(2), 170–181. <https://doi.org/10.1111/j.1600-0870.2007.00226.x>
- Seiler, C., Zwiers, F., Hodges, K. I., & Scinocca, J. (2018). How does dynamical downscaling affect model biases and future projections of explosive extratropical cyclones along North America's Atlantic coast? *Climate Dynamics*, 50(1), 677–692. <https://doi.org/10.1007/s00382-017-3634-9>
- Seland, y., Bentzen, M., Olivè, D. J. L., Tonizzi, T., Gjermundsen, A., & Graff, L. S. (2019). *NCC NorESM2-LM model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.8036>
- Sheridan, S. C., & Lee, C. C. (2010). Synoptic climatology and the general circulation model. *Progress in Physical Geography*, 34(1), 101–109. <https://doi.org/10.1177/030913309357012>
- Shieh, S., & Liao, I. (2012). A new approach for data clustering and visualization using self-organizing maps. *Expert Systems with Applications*, 39(15), 11924–11933. <https://doi.org/10.1016/j.eswa.2012.02.181>
- Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., et al. (2019b). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 145(724), 2876–2908. <https://doi.org/10.1002/qj.3598>
- Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., & Wyszyński, P. (2019a). NOAA-CIRES-DOE twentieth Century reanalysis version 3 [Dataset]. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/10.5065/H93G-WS83>
- Sukhdeo, R., Ullrich, P. A., & Grotjahn, R. (2022). Assessing the large-scale drivers of precipitation in the northeastern United States via linear orthogonal decomposition. *Climate Dynamics*, 59(11–12), 3657–3681. <https://doi.org/10.1007/s00382-022-06289-y>
- Swales, D., Alexander, M., & Hughes, M. (2016). Examining moisture pathways and extreme precipitation in the US Intermountain West using self-organizing maps. *Geophysical Research Letters*, 43(4), 1727–1735. <https://doi.org/10.1002/2015gl067478>
- Tatebe, H., & Watanabe, M. (2018). *MIROC MIROC6 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.5603>
- Taylor, G. P., Loikith, P. C., Aragon, C. M., Lee, H., & Waliser, D. E. (2022). CMIP6 model fidelity at simulating large-scale atmospheric circulation patterns and associated temperature and precipitation over the Pacific Northwest. *Climate Dynamics*, 60(7–8), 1–20. <https://doi.org/10.1007/s00382-022-06410-1>
- Terando, A., Reidmiller, D., Hostetler, S. W., Littell, J. S., Beard, T. D., Jr., Weiskopf, S. R., et al. (2020). *Using information from global climate models to inform policymaking: The role of the US Geological Survey*. US Department of the Interior, US Geological Survey.
- Tilinina, N., Gulev, S. K., Rudeva, I., & Koltermann, P. (2013). Comparing cyclone life cycle characteristics and their interannual variability in different reanalyses. *Journal of Climate*, 26(17), 6419–6438. <https://doi.org/10.1175/jcli-d-12-00777.1>
- Ullrich, P. A. (2022). TempestExtremes [Software]. <https://github.com/ClimateGlobalChange/tempesteextremes>
- Ullrich, P. A., & Zarzycki, C. M. (2017). Tempestextremes: A framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geoscientific Model Development*, 10(3), 1069–1090. <https://doi.org/10.5194/gmd-10-1069-2017>
- Ulrich, P. A., Zarzycki, C. M., McCleary, E. E., Pinheiro, M. C., Stansfield, A. M., & Reed, K. A. (2021). TempestExtremes v2. 1: A community framework for feature detection, tracking and analysis in large datasets. *Geoscientific Model Development Discussions*, 1–37(8), 5023–5048. <https://doi.org/10.5194/gmd-14-5023-2021>
- Van Hulle, M. (2012). Self-organizing maps. In *Handbook of natural computing* (pp. 585–622). Springer.
- Vesanto, J., Hinberg, J., Alhoniemi, E., Parhankangas, J., Team, S., & Oy, L. (2000). SOM toolbox for Matlab 5 [Software]. <http://www.cis.hut.fi/projects/somtoolbox/>
- von Storch, J.-S., Putrasahan, D., Lohmann, K., Gutjahr, O., Jungclaus, J., Bittner, M., & Roeckner, E. (2018). *MPI-M MPI-ESM1.2-XR model output prepared for CMIP6 HighResMIP hist-1950*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.10307>
- Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., & Roeckner, E. (2019). *MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.6595>
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.
- Willison, J., Robinson, W. A., & Lackmann, G. M. (2013). The importance of resolving mesoscale latent heating in the North Atlantic storm track. *Journal of the Atmospheric Sciences*, 70(7), 2234–2250. <https://doi.org/10.1175/jas-d-12-0226.1>
- Wu, T., Chu, M., Dong, M., Fang, Y., Jie, W., Li, J., & Zhang, Y. (2018). *BCC BCC-CSM2MR model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.2948>
- Yau, A. M.-W., & Chang, E. K.-M. (2020). Finding storm track activity metrics that are highly correlated with weather impacts. Part I: Frameworks for evaluation and accumulated track activity. *Journal of Climate*, 33(23), 10169–10186. <https://doi.org/10.1175/jcli-d-20-0393.1>
- Yoshida, A., & Asuma, Y. (2004). Structures and environment of explosively developing extratropical cyclones in the northwestern pacific region. *Monthly Weather Review*, 132(5), 1121–1142. [https://doi.org/10.1175/1520-0493\(2004\)132<1121:saeoed>2.0.co;2](https://doi.org/10.1175/1520-0493(2004)132<1121:saeoed>2.0.co;2)
- Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., & Urakawa, S. (2019). *MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.6842>
- Zappa, G., Shaffrey, L. C., & Hodges, K. I. (2013). The ability of CMIP5 models to simulate North Atlantic extratropical cyclones. *Journal of Climate*, 26(15), 5379–5396. <https://doi.org/10.1175/jcli-d-12-00501.1>
- Zarzycki, C. M. (2018). Projecting changes in societally impactful northeastern US snowstorms. *Geophysical Research Letters*, 45(21), 12–067. <https://doi.org/10.1029/2018gl079820>
- Zarzycki, C. M., & Ullrich, P. A. (2017). Assessing sensitivities in algorithmic detection of tropical cyclones in climate data. *Geophysical Research Letters*, 44(2), 1141–1149. <https://doi.org/10.1002/2016gl071606>
- Zhao, M., Blanton, C., John, J. G., Radhakrishnan, A., Zadeh, N. T., & McHugh, C. (2018). *NOAA-GFDL GFDL-CM4C192 model output prepared for CMIP6 HighResMIP hist-1950*. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.8567>
- Ziehn, T., Chamberlain, M., Lenton, A., Law, R., Bodman, R., Dix, M., & Druken, K. (2019). CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.4272>
- Zielinski, G. (2002). A classification scheme for winter storms in the eastern and central United States with an emphasis on nor'easters. *Bulletin of the American Meteorological Society*, 83(1), 37–52. [https://doi.org/10.1175/1520-0477\(2002\)083<0037:acsfs>2.3.co;2](https://doi.org/10.1175/1520-0477(2002)083<0037:acsfs>2.3.co;2)