

1 **Exploring a data-driven approach to identify regions of**
2 **change associated with future climate scenarios**

3 **Zachary M. Labe¹, Thomas L. Delworth², Nathaniel C. Johnson², and William**
4 **F. Cooke²**

5 ¹Atmospheric and Oceanic Sciences Program, Princeton University, NJ, USA

6 ²NOAA/OAR/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

7 **Key Points:**

- 8 • A neural network applied to large ensembles can link annual mean maps of cli-
9 mate variables to a range of radiative forcing scenarios
- 10 • Information extracted from regional change patterns is used to distinguish between
11 climate scenarios, even those with similar global warming
- 12 • Radiative forcing scenario classifications for the later 21st century are sensitive
13 to a difference in the timing of mitigation by ten years

Abstract

A key consideration for evaluating climate projections is uncertainty in radiative forcing scenarios. Although it is straightforward to monitor greenhouse gas concentrations and compare those observations with specified climate scenarios, it remains less obvious on how to connect regional climate patterns with these scenarios in real time. Here we introduce a machine learning approach for linking patterns of climate change with radiative forcing scenarios and use an attribution method to understand how these linkages are made. We train a neural network using output from the SPEAR Large Ensemble to classify whether temperature or precipitation maps are most likely to originate from one of several potential radiative forcing scenarios. The neural network learns to identify “fingerprint” patterns that associate signals of climate change with the scenarios. We illustrate this using output from additional mitigation experiments and highlight regions that are critical for associating the new simulations with likely radiative forcing scenarios.

Plain Language Summary

There are several sources of uncertainties when considering future projections of climate change. This includes uncertainty related to natural climate variations, uncertainties related to biases and climate sensitivity among different models, and finally the uncertainty related to the trajectory of greenhouse gas emissions. We focus on this third source of uncertainty, which is typically considered by running a climate model with a range of scenarios that include varying amounts of greenhouse gases. Although comparing real-world greenhouse gas levels with each climate scenario is a relatively simple task, it is harder to compare which climate scenario is most closely aligned with year-to-year patterns of weather and climate anomalies. In this study, we introduce a machine learning approach that learns to associate yearly maps of global temperature and precipitation with individual climate scenarios. We also compare how these future predictions of climate scenarios may change over time depending on the introduction of climate mitigation efforts and show regions that are particularly sensitive to this change. Our results indicate that starting aggressive mitigation efforts a decade earlier can lead to the lowest greenhouse gas emission scenario being predicted by the machine learning model at the end of the century using this climate model.

1 Introduction

The evolution of future greenhouse gas pathways, such as those developed using integrated assessment models, remains one of the dominant drivers of uncertainty in climate change projections (Hawkins & Sutton, 2009; Lehner et al., 2020; S. Zhang et al., 2023). In the near term, it is even more difficult to identify which climate change scenario is most closely aligned with real-world observations due to the similarities in greenhouse gas concentrations (Meinshausen et al., 2020; Pedersen et al., 2021; Huard et al., 2022) and the outsized influence of internal climate variability (Maher et al., 2020). Although it is possible to track changes in global emissions through the carbon and methane budgets (e.g., Saunio et al., 2020; Sognaes et al., 2021; Friedlingstein et al., 2022, 2023; Liu et al., 2023) and further quantify the time-mean, long-term warming signal using historical records (e.g., Stott et al., 2013; Dong et al., 2020; Hausfather et al., 2020) or with observational constraint-like approaches (e.g., Brunner et al., 2020; Liang et al., 2020; Tokarska et al., 2020; Ribes et al., 2021), it is less clear on how to monitor whether interannual patterns of weather and climate are consistent with particular climate change scenarios. This is made uniquely difficult due to the modulating effect of internal climate variability on the forced response (Deser et al., 2012; Medhaug et al., 2017; Wills et al., 2020; Sippel et al., 2021; Jain et al., 2023; Lehner & Deser, 2023), which can even delay detection of climate mitigation efforts as well (Tebaldi & Friedlingstein, 2013; Marotzke, 2019; Samset et al., 2020). At the same time, recent data-driven results have shown that fingerprints of forced change are now detectable in any single day of observational data (Sippel et al., 2020), but this framing does not necessarily address the question of which climate change pathway is more realistic or probable from year-to-year. Our research letter begins to investigate this question by building off developments in applications of machine learning for climate science (Huntingford et al., 2019; Irrgang et al., 2021; Sonnewald et al., 2021; Rolnick et al., 2022) that are then applied to a collection of large ensemble simulations from a high-resolution, fully-coupled climate model.

Here, we design an artificial neural network (ANN) to learn to associate yearly maps of simulated surface temperature or precipitation with several possible climate scenarios that consist of either natural forcing, historical forcing, or one of three possible future anthropogenic climate change trajectories. Then we input data from two overshoot scenarios that feature aggressive climate mitigation efforts beginning in either 2031 or 2040. The purposes of evaluating these additional simulations are to: 1) use this neu-

78 ral network detection framework to examine hypothetical futures that could be analo-
79 gous to inputting data from the real world, and 2) identify whether there are differences
80 in the temporal evolution of climate scenario classifications, given a 10-year difference
81 in the onset of climate mitigation. This is especially relevant given the growing inter-
82 est in alternative pathways for achieving climate mitigation strategies (IPCC, 2022), such
83 as through the development of carbon dioxide removal for net negative emissions (Davis
84 et al., 2018; Fuss et al., 2018; Minx et al., 2018; de Kleijne et al., 2022). In all cases, we
85 apply attribution methods from explainable artificial intelligence (XAI) to attempt to
86 understand which climate features the neural network is using to make its scenario clas-
87 sifications. Ultimately, we show that an ANN can skillfully detect which climate scenario
88 is associated with simulated fields of global temperature or precipitation by learning in-
89 formation from regional climate anomalies, largely over the subpolar North Atlantic and
90 portions of land areas across the tropics.

91 **2 Data and Methods**

92 To begin this data-driven approach, we employ a collection of large ensemble ex-
93 periments from a single modeling system - the Seamless System for Prediction and EArth
94 System Research (SPEAR; Delworth et al., 2020) by the Geophysical Fluid Dynamics
95 Laboratory (GFDL). We include these SPEAR simulations as inputs to the neural net-
96 works, which are used for the purpose of distinguishing between individual climate sce-
97 narios (Figure S1). This includes several future projections from the Shared Socioeco-
98 nomic Pathways (SSPs; O’Neill et al., 2014, 2016). Since ANNs can learn nonlinear in-
99 formation across a given geographic domain (Irrgang et al., 2021; de Burgh-Day & Leeuwen-
100 burg, 2023), recent work has discovered that they can be powerful tools for comparing
101 across different GCMs and climate change scenarios (e.g., Labe & Barnes, 2022; Labe,
102 Barnes, & Hurrell, 2023; Bône et al., 2023; Brunner & Sippel, 2023) and for use in ex-
103 tracting patterns of forced change from the background noise of internal variability (e.g.,
104 Rader et al., 2022; Po-Chedley et al., 2022; Gordon et al., 2023). This can be especially
105 advantageous when compared to traditional methods that require local gridpoint and
106 time-mean statistics (Barnes et al., 2020). Although our current detection framework
107 is therefore limited to a single GCM, this subsequently eliminates any uncertainties re-
108 lated to model structural biases, which Labe and Barnes (2022) showed can influence the
109 results because the machine learning model can instead begin to discern mean state bi-

ases for its classifications. SPEAR also provides a large number of individual ensemble members for training each different climate scenario, while most other GCM large ensembles only provide enough data for a single SSP projection, at least given what is publicly available (NCAR, 2020; Deser et al., 2020). Lastly, SPEAR has a relatively high horizontal resolution, which a recent study found can improve machine learning prediction skill since the model can learn to recognize relevant smaller scale features, like near topography (Labe et al., 2024).

2.1 GFDL SPEAR Large Ensemble Experiments

We use the medium resolution configuration of the fully-coupled (atmosphere-ocean-sea ice-land) SPEAR model (also referred to as SPEAR_MED). This version has 33 vertical levels in the atmosphere with a model top at 1 hPa and uses a land-atmosphere grid spacing of 0.5° and a coarser ocean-sea ice grid spacing of approximately 1° (telescoping to 0.33° near the equator). SPEAR features the same model components as GFDL CM4 (Held et al., 2019), which includes AM4, LM4, MOM6, and SIS2 (Zhao et al., 2018a, 2018b; Adcroft et al., 2019). However, SPEAR has been tuned for the study of seasonal to multidecadal predictability and projection, and more details on this can be found in Delworth et al. (2020).

SPEAR offers 30 ensemble members for each climate scenario evaluated here, which are listed in Table S1 and shown in Figure 1. To sample different phases of internal climate variability, each ensemble member of SPEAR is branched using initial conditions from an 1850 control run at 20 year intervals, but using the same land initial conditions starting in 1921. Every ensemble member is then prescribed with historical radiative forcing from the years 1921 to 2014, which includes aerosols, greenhouse gases, land use/land change, and solar irradiance (Meinshausen et al., 2017; Hurtt et al., 2020). Note that to balance the number of years in each climate scenario class (see Text S1), we only analyze the years of 1929 to 2014 from the SPEAR historical large ensemble. Thereafter, SPEAR is prescribed with radiative forcing following either future projections from the SSP5-8.5 scenario (extreme, outlier greenhouse gas emissions), SSP2-4.5 (moderate emission scenario), or SSP1-1.9 (lowest emission scenario with net zero by 2050) (Kriegler et al., 2017; Ritchie & Dowlatabadi, 2017; Riahi et al., 2017; Burgess et al., 2020; Peters & Hausfather, 2020; Hausfather & Peters, 2020; Tebaldi et al., 2021; Pielke et al., 2022). Again, 30 ensemble members are available for each of the three SSP scenarios over

142 the years of 2015 to 2100, which are the basis for training and testing the ANN. Two at-
143 mospheric variables from SPEAR are considered for this work: 2 m height air temper-
144 ature (“temperature”) and total precipitation rate (“precipitation”).

145 Along with the future climate change projections, we examine a natural forcing-
146 only scenario over the period of 2015 to 2100. For this counterfactual climate experiment,
147 all external forcings including anthropogenic aerosols, land use/land change, and green-
148 house gases are maintained at 1921 levels. Solar irradiance is then prescribed toward a
149 hypothetical estimate based on the solar cycle taken from observations. Volcanic aerosols
150 after 2024 are set to the long-term mean over the 1850 to 2014 period (Delworth et al.,
151 2022). Thus, without external anthropogenic forcing, there are generally no pronounced
152 long-term trends in this climate scenario (Figures 1 and S3a,e).

153 We also analyze two rapid climate mitigation scenarios that are used for out-of-
154 sample inferences after the ANN training process is complete. The first follows SSP5-
155 3.4OS, which is an overshoot scenario (OS) that closely emulates SSP5-8.5 until the year
156 2040 and thereafter includes a rapid reduction in greenhouse gas levels (Figure S2) due
157 to bioenergy crops and other carbon capture and storage-like technology (Melnikova et
158 al., 2022). This leads to large net negative emissions by 2100 (Meinshausen et al., 2020).
159 We also conducted an additional idealized mitigation scenario, which again follows SSP5-
160 3.4OS, but this time is scaled to start in 2031 following a similar rate of decay in the lev-
161 els of carbon dioxide and methane (Figure S2a-b). All other forcings are kept to SSP5-
162 3.4OS (e.g., ozone, aerosols, and nitrous oxide (Figure S2c)). This scenario, which we
163 denote as SSP5-3.4OS_10ye (i.e., 10ye for 10 years earlier), is meant to imitate an ear-
164 lier start to rapid climate mitigation, and thus comparing the SSP5-3.4OS and SSP5-
165 3.4OS_10ye climate scenarios can provide a hypothetical comparison for revealing how
166 the climate system could respond to different timings of aggressive future climate mit-
167 igation.

168 Figure 1 compares the responses of global mean annual temperature and precip-
169 itation for each of the climate scenarios used in this work. In contrast to the higher emis-
170 sions simulated under SSP5-8.5 and SSP2-4.5, there is a maximum in global surface tem-
171 perature by the 2030s under SSP1-1.9 radiative forcing that is followed by a slow cool-
172 ing through the end of the 21st century (Figures 1 and S3). The overshoot mitigation
173 scenarios, which are similar to SSP5-8.5 until either 2031 or 2040, show ensemble mean

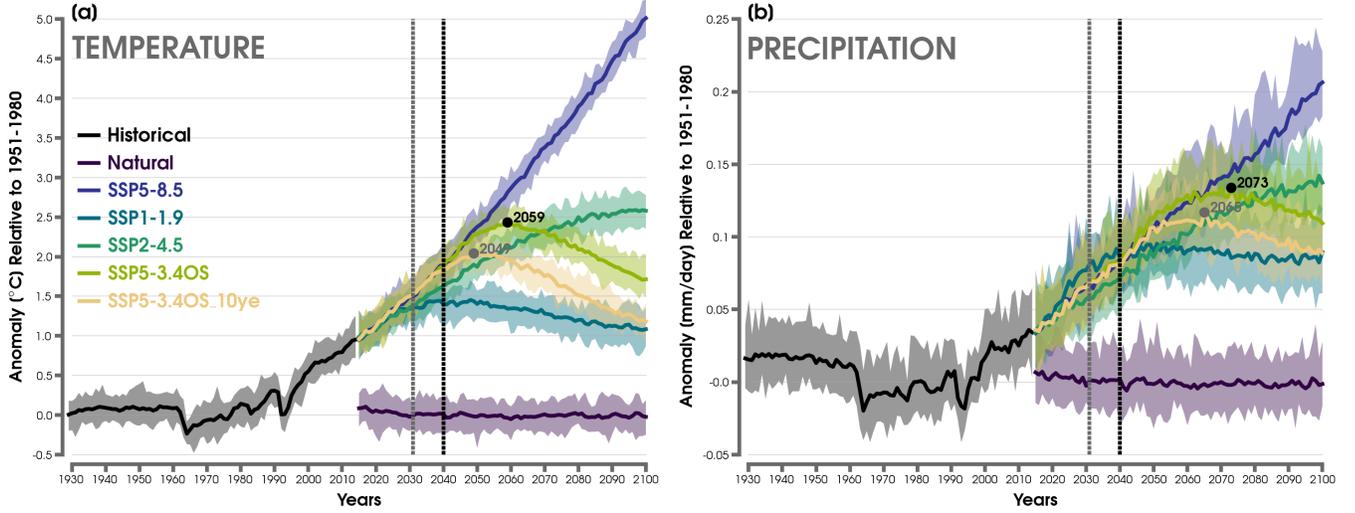


Figure 1. (a) Time series of annually-averaged global mean temperature anomalies for the ensemble mean of the SPEAR historical scenario from 1929 to 2014 (black line), a natural-only forcing scenario experiment with SPEAR from 2015 to 2100 (purple line), a future scenario experiment with SPEAR following SSP1-1.9 from 2015 to 2100 (light blue line), a future scenario experiment with SPEAR following SSP2-4.5 from 2015 to 2100 (dark green line), a future scenario experiment with SPEAR following SSP5-8.5 from 2015 to 2100 (dark blue line), a future mitigation scenario experiment with SPEAR following SSP5-3.4OS from 2015 to 2100 (light green line), and a future mitigation scenario experiment with SPEAR following SSP5-3.4OS but starting mitigation 10 years earlier (SSP5-3.4OS_10ye; tan line). The spread across the 30 ensemble members is indicated by the lighter shading for each climate scenario experiment. All anomalies are computed from their respective 1921-1950 climatological time means (historical or natural forcing). The black and gray markers note the highest ensemble mean temperature for SSP5-3.4OS and SSP5-3.4OS_10ye, respectively. The dashed black vertical line indicates the start of mitigation for SSP5-3.4OS (year 2040), and the dashed gray vertical line indicates the start of mitigation for SSP5-3.4OS_10ye (year 2031). (b) As in (a), but for global mean precipitation anomalies.

174 global temperatures rising until 2049 for SSP5-3.4OS_10ye and 2059 for SSP5-3.4OS. In
175 Figure S2 the time series of greenhouse gas concentrations show a corresponding peak
176 in carbon dioxide levels of about 515 ppm for SSP5-3.4OS_10ye and 571 ppm for SSP5-
177 3.4OS, which are nearly concurrent with the timing of the greatest global warming re-
178 sponse before the reversal of the upward trend. This contrasts with the continuing rise
179 of carbon dioxide under SSP5-8.5 that reaches 1135 ppm by 2100; that said, recent work
180 has shown that this climate scenario is becoming an implausible upper bound (e.g., Pielke
181 et al., 2022). The overshoot scenario results are broadly consistent with recent studies
182 (e.g., MacDougall et al., 2020) finding little warming after net zero emissions, but note
183 that these scenarios also include a drawdown of greenhouse gases. Strikingly, by 2100,
184 the difference in the ensemble-mean global mean surface temperature for SSP5-3.4OS_10ye
185 and SSP5-3.4OS is 0.53°C (Figure 1a). Even more revealing is that the ensemble spreads
186 do not overlap despite rapid mitigation efforts only starting a decade earlier in SSP5-
187 3.4OS_10ye. Comparing temperature trends over 2071 to 2100 also reveals widespread
188 cooling in both SSP5-3.4OS and SSP5-3.4OS_10ye, which is particularly amplified in higher
189 latitude regions of the Northern Hemisphere (Figure S4a-b). There are also hemispheric
190 differences in precipitation, including a southward shift in the annual mean climatology
191 of the Intertropical Convergence Zone. This could be related to the weakening of the At-
192 lantic Meridional Overturning Circulation (AMOC) as simulated by SPEAR (Delworth
193 et al., 2022) and will be investigated in future work.

194 Globally, precipitation increases in response to larger radiative forcing in SSP2-4.5
195 and even more so for SSP5-8.5 (Figure 1b). In contrast to global temperature, the re-
196 versal of the ensemble mean upward precipitation trend does not occur until about 10-
197 15 years later for both the SSP5-3.4OS_10ye and SSP5-3.4OS scenarios. Internal vari-
198 ability also contributes to overlapping ensemble member spreads in precipitation between
199 SSP1-1.9 and SSP2-4.5 along with the two overshoot scenarios, but this global mean re-
200 sponse continues to remain separate and distinct from the natural forcing scenario.

201 **2.2 Explainable Neural Network Approach**

202 Figure S1 summarizes our framework for using neural networks to detect which cli-
203 mate scenario is associated with maps of different climate variables. First, a classifica-
204 tion ANN is trained on annual mean global maps of temperature (or precipitation) from
205 SPEAR large ensembles simulated under either historical forcing from 1929 to 2014 or

206 under natural forcing, SSP1-1.9, SSP2-4.5, and SSP5-8.5 for the future years from 2015
207 to 2100. The aim of the ANN is to learn to associate individual inputs (the climate maps)
208 with the correct climate scenario (i.e., 5 possible classes/predictions). Figures S5-S6 show
209 sensitivity of the ANN performance to different choices in architecture, but overall we
210 find relatively similar mean skill across these networks. The ANN configuration that is
211 ultimately selected from this hyperparameter sweep is based on balancing median val-
212 idation accuracy and overall interpretability, which is further described in Text S1. Af-
213 ter training, validating, and testing is complete, we then input data from the 30 ensem-
214 ble members simulated under SSP5-3.4OS or SSP5-3.4OS_10ye into the ANN to see which
215 climate scenario class is predicted for every year from 2015 to 2100 during these miti-
216 gation scenarios. This is effectively out-of-sample data that the ANN has never seen be-
217 fore, and the ANN can again classify each year as either natural forcing, historical forc-
218 ing, SSP1-1.9, SSP2-4.5, or SSP5-8.5. For ease of interpretation in our results, we con-
219 catenate years from 2015 to 2030 using SSP5-3.4OS to complete the time series for SSP5-
220 3.4OS_10ye, which by itself does not diverge until 2031. In other words, the machine learn-
221 ing classifications for the years of 2015 to 2030 are the same between SSP5-3.4OS and
222 SSP5-3.4OS_10ye, so that they equally cover the same 2015-2100 period (86 years).

223 As discussed further below, we discover that there are jumps in the classifications
224 from one climate scenario to the next for the time evolution of the overshoot scenarios
225 (e.g., ANN consistently predicting SSP5-8.5 followed by an abrupt transition to consis-
226 tent SSP2-4.5 predictions as time progresses). To investigate these transitions in climate
227 scenario predictions more closely, we also train and test two binary classification ANNs,
228 which can predict either SSP5-8.5 versus SSP2-4.5 (Figure S1b) or SSP2-4.5 versus SSP1-
229 1.9 (Figure S1c). We again feed the out-of-sample data from the SSP5-3.4OS and SSP5-
230 3.4OS_10ye SPEAR large ensembles into the binary ANNs after their original training
231 is complete. The purpose of these additional ANNs is primarily for interpreting our ex-
232 plainable machine learning results, which is described in detail within Section 3.2. The
233 skill metrics for variations in the architecture of the binary ANNs are also provided in
234 Figure S7 and S9 for temperature and Figures S8 and S10 for precipitation.

235 For understanding which climate patterns are important for the ANNs to distin-
236 guish one scenario from another, we use a form of XAI called Integrated Gradients (Sundararajan
237 et al., 2017), which is an ad hoc feature attribution method that is used to describe the
238 contribution of each input pixel (e.g., an individual grid cell on a global map) to the over-

239 all prediction output (Baehrens et al., 2010). Integrated Gradients is similar to the method
240 of Input*Gradient (Shrikumar et al., 2016, 2017), but is designed to address potential
241 nonlinearities. Recent work, such as Mamalakis et al. (2022b), has shown that explana-
242 tions from Integrated Gradients have performed well compared to other XAI methods
243 on climate datasets with similar characteristics as ours. We also found close XAI results
244 after applying methods using different layer-wise relevance propagation rules (Bach et
245 al., 2015) (not shown). In this study, highly positive areas of relevance on the XAI heatmaps
246 can be interpreted as regions that pushed the ANN toward its predicted climate scenario
247 class, whereas negative areas of relevance are vice versa. While XAI is not itself a method
248 for proving causality, it can still help to aid in building user trust and insight into the
249 decision-making process of the machine learning black box (McGovern et al., 2019; Toms
250 et al., 2020; Jacovi et al., 2021; Mamalakis et al., 2022a; Bostrom et al., 2023). Here, our
251 XAI heatmaps provide a tool in identifying the relevant climate regions that were used
252 by the ANN to make its classifications (e.g., Labe & Barnes, 2022), especially for reveal-
253 ing the important time-evolving climate patterns after rapid mitigation efforts in the two
254 overshoot scenarios.

255 In summary, we use ANNs to take inputs of global temperature or precipitation
256 data from SPEAR and task the network to classify which climate scenario is associated
257 with each yearly map. Additional details regarding the choice and design of the ANNs
258 can be found in Text S1, and the final hyperparameter specifications that are uniquely
259 selected for each climate variable and classification task are listed in Table S2.

260 **3 Results**

261 **3.1 Classification of Climate Scenarios**

262 In Figure 2, we begin to evaluate the skill of our detection method on the 2 test-
263 ing ensemble members associated with the 5-class ANN and then show composites of the
264 relevance heatmaps for each predicted climate scenario class using the Integrated Gra-
265 dients method of XAI. We find higher accuracy for inputs of temperature maps (91%)
266 compared to precipitation (86%), which is likely due to their greater separation between
267 individual future projections (Figure 1a) and higher regional signal-to-noise ratio (Hawkins
268 & Sutton, 2011). Although our classes are balanced, we still show the metrics of recall,
269 precision, and F1 score for each climate scenario. Skill is generally similar for each cli-

270 mate scenario, except for the natural forcing ensemble members which have better per-
 271 formance for temperature and precipitation (Figure 2b,g).

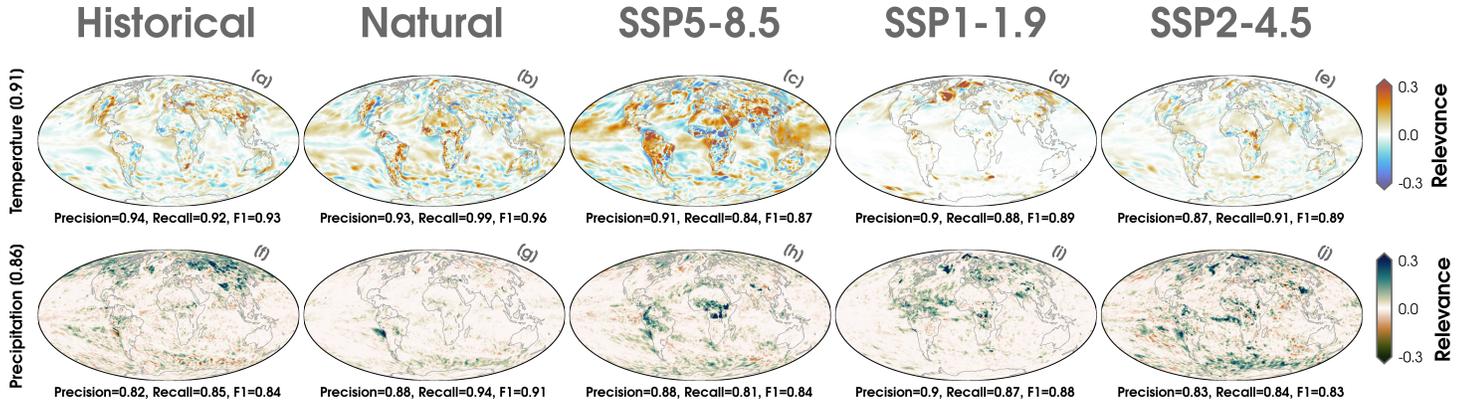


Figure 2. (a-e) Explainability maps using the Integrated Gradients method that are composed separately for each predicted climate scenario class using the testing ensemble members and global maps of temperature. The total accuracy is denoted in the far left label. The local precision, recall, and F1 scores for individual classes are denoted below each climate scenario composite. Relevance values are normalized by the absolute maximum relevance in each composite. (f-j) As in (a-e), but for maps of precipitation.

272 For inputs of temperature, we find several spatially-coherent regions of positive and
 273 negative areas of relevance in common across the climate scenarios. This indicates that
 274 these particular regions are important locations for the ANN to decide which scenario
 275 is associated with a given map. One of these regions is across eastern South America,
 276 where temperature anomalies in this region can therefore be interpreted as an impor-
 277 tant characteristic (or indicator) for correctly identifying a temperature map from the
 278 natural-forcing scenario (positive relevance; Figure 2b), but on the other hand, this re-
 279 gion also tends to confuse the ANN when given historical-forcing maps as it tries to push
 280 the network toward another class prediction (negative relevance; Figure 2a). Another
 281 important indicator region overlaid with areas of positive and negative relevance depend-
 282 ing on the specific climate scenario is found across Central Africa. Again, this suggests
 283 that temperatures in this region are a unique indicator for the ANN to identify the in-
 284 dividual climate scenario. Locations with highly positive and negative relevance values
 285 in close proximity are also found in some areas near higher topography and over the South-

ern Ocean, which is likely related to sharper temperature gradient features or simply insignificant, noisy XAI attributions. There are distinctive relevance patterns for individual scenarios too, such as the North Atlantic being most important for predicting SSP1-1.9 (Figure 2d) and a temperature signal across the tropical west-central Pacific that is important for predicting SSP5-8.5 (Figure 2c). This is similar to previous work that has found a contribution of scenario uncertainty to the evolution of the North Atlantic warming hole region, but even larger uncertainties exist if comparing across other GCMs (Park & Yeh, 2024).

Looking at the relevance maps for precipitation (Figure 2f-j), we find that features across the high latitude regions of the Arctic and the Southern Ocean are important for the ANN to make its scenario classifications. The locations of these positive relevance areas align with earlier work showing stronger signal-to-noise ratios from radiative forcing (e.g., H. Zhang & Delworth, 2018; Hawkins et al., 2020). We again find that the North Atlantic and Central Africa are associated with higher relevance, but one notably different relevance region is over the tropical Atlantic that is especially used for predicting either SSP1-1.9 (Figure 2i) or SSP2-4.5 (Figure 2j). Based on these XAI results, we mainly find that the ANN is focused on patterns of polar precipitation and the response of the Intertropical Convergence Zone in order to distinguish between different climate scenario classes.

Even though we have now shown that there are specific regions of temperature and precipitation information that the ANN is weighting together for discerning individual climate scenarios, it is still possible the network is simply learning to distinguish the climate scenarios by the differences in their mean of each map. To address this prospect, we set up a logistic regression model by inputting only the value of the global mean temperature or precipitation to attempt to predict the five scenarios. For this problem, we find that the logistic regression skill is highly variable due from a sensitivity related to different combinations of training ensemble members; nonetheless, it still only reaches a maximum accuracy up to 60% for temperature and precipitation for its best model (not shown). This baseline comparison provides further support to show that the ANN is learning important spatial information to connect the yearly maps with individual climate scenarios. This result is also not too surprising given that there is substantial overlap in the global means across scenarios when evaluating the data without considering their time evolution (Figure 1). For example, there are at least a few ensemble members in

319 the SSP1-1.9, SSP2-4.5, and SSP5-8.5 scenarios that at some point all observe a global
320 mean temperature anomaly of 1.5°C (Figure 1a), and even more overlaps in the ensem-
321 ble spreads are found for precipitation (Figure 1b).

322 **3.2 Identifying Indicators of Regional Change After Rapid Mitigation**

323 After finding that our data-driven framework can skillfully learn to associate maps
324 of temperature and precipitation with different climate scenarios, we now feed in data
325 from two overshoot simulations that were not used as part of the original training pro-
326 cess. To recall from earlier, these experiments are associated with aggressive climate mit-
327 igation that starts in 2040 (SSP5-3.4OS) or about a decade earlier in 2031 (SSP5-3.4OS_10ye)
328 after branching from a trajectory that mirrors SSP5-8.5 radiative forcing. The effects
329 of starting mitigation 10 years apart on the time-evolution of the predicted climate sce-
330 narios are displayed in Figure 3 using the 5-class ANN framework. These classifications
331 are sorted by the selected scenario for each of the 30 ensemble members for SSP5-3.4OS
332 and SSP5-3.4OS_10ye using annual-mean global maps of temperature (Figure 3a,c) and
333 precipitation (Figure 3b,d). Greater uncertainty across the individual ensemble class pre-
334 dictions is found prior to around 2030, which likely reflects the overlap in SSP projec-
335 tions as shown in Figure 1. In other words, there are fewer distinctive novel patterns that
336 the ANN can learn to connect with each unique climate scenario during this period of
337 time.

338 Looking at the yearly progression of predictions for SSP5-3.4OS, we find that SSP5-
339 8.5 is predicted by the majority of the ensemble members from the mid-2020s to about
340 2060 for inputs of temperature and precipitation (Figure 3a-b). Thereafter, the major-
341 ity of ensemble members are classified as the SSP2-4.5 scenario through 2100. In fact,
342 the highest agreement across ensemble members is found for these future SSP2-4.5 clas-
343 sifications, particularly for the temperature maps. This result is also consistent with the
344 high value of ensemble mean ANN confidence, as exhibited in Figure S11, for the yearly
345 evolution of the climate scenario classifications after the middle of the 21st century. In-
346 terestingly, however, we do find a reduction in mean ANN confidence for SSP2-4.5 and
347 a corresponding increase in confidence toward the SSP1-1.9 class for maps of precipita-
348 tion by the 2090s under SSP5-3.4OS (Figure S11b).

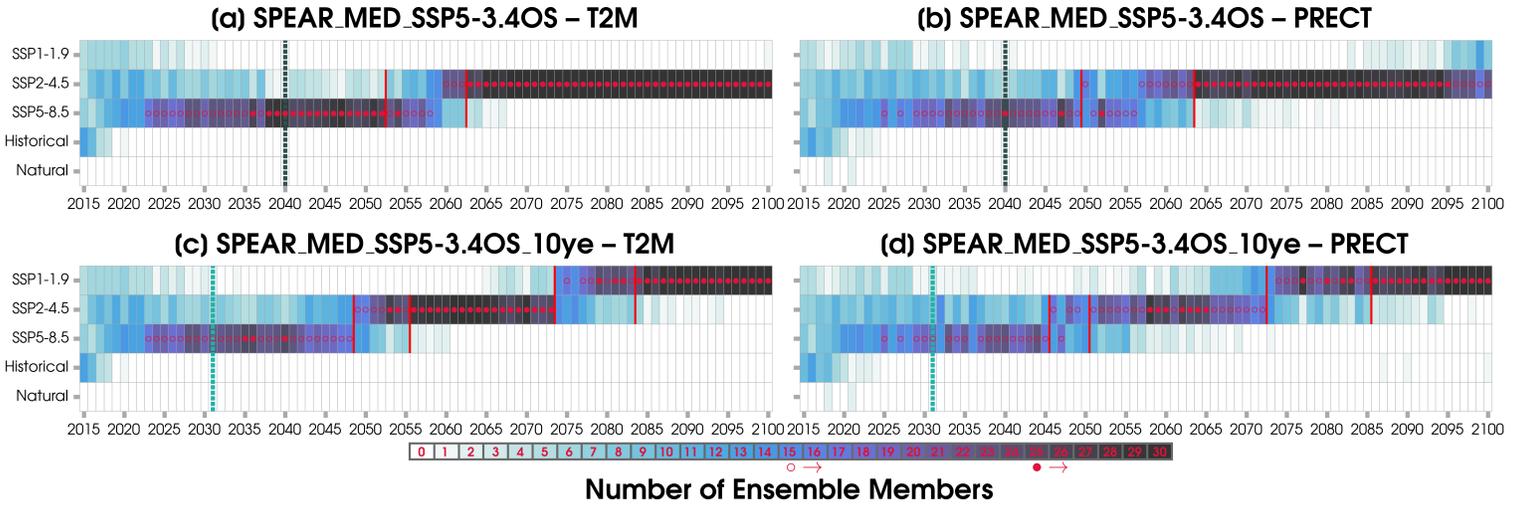


Figure 3. (a) Heatmap showing the number of ensemble members for each individual classification of SSP5-3.4OS temperature maps from 2015 to 2100. The dashed dark green line indicates the start of mitigation in 2040. The vertical red lines indicate the start and end of the transition in consistent predictions of the climate scenario classes from SSP5-8.5 to SSP2-4.5. See text for details. Open red dots denote that more than 15 ensemble members predicted that individual climate scenario, and filled red dots indicate that at least 25 ensemble members predicted that scenario. (b) As in (a), but for maps of precipitation. (c-d) As in (a-b), but for individual classification predictions of SSP5-3.4OS_{10ye}. The vertical red lines indicate the start and end of the transitions in consistent predictions of the climate scenario classes from SSP5-8.5 to SSP2-4.5 or from SSP2-4.5 to SSP1-1.9. The dashed bright green line indicates the start of mitigation in 2031.

349 For the ensemble of simulations following SSP5-3.4OS_10ye radiative forcing, we
350 find a different evolution of climate scenario classifications, as revealed in Figure 3c-d.
351 These predictions show a transition from mainly predicting SSP5-8.5 to SSP2-4.5 that
352 occurs earlier at around 2050 for maps of temperature and precipitation. Another changeover
353 then starts in the mid-2070s when the ANN begins to predict the SSP1-1.9 scenario, which
354 persists until the end of the century. Again, we find high agreement in these future cli-
355 mate scenario predictions across individual ensemble members. This suggests that the
356 ANN is learning robust patterns of regional climate indicators unique to each scenario
357 despite the background noise of internal variability. Another surprising result here is the
358 striking consistency in the timing of shifts between the consecutive climate scenario pre-
359 dictions found for both variables.

360 To more thoroughly evaluate these transitions in scenario classifications that are
361 selected for the overshoot experiments, we now turn to our two binary ANNs. Specif-
362 ically, we focus on compositing the differences in their relevance maps before and after
363 these transition periods (Figure 4), which are associated with lower model confidence (Fig-
364 ure S11-12) and greater variability in the predicted scenarios when looking across indi-
365 vidual ensemble members (Figure 3). Since the ANN can only predict one of two pos-
366 sible climate scenarios, we can more directly interpret these explainability maps. This
367 is unlike the earlier 5-class ANN, where their relevance maps cannot be compared directly
368 between one climate scenario and another (e.g., Figure 2), as this ANN must instead learn
369 to identify climate patterns that are unique to each of the five classes (Labe & Barnes,
370 2022).

371 We first consider the broader shift in classifying the SSP5-8.5 scenario to mostly
372 the SSP2-4.5 scenario for SSP5-3.4OS and SSP5-3.4OS_10ye maps of temperature (Fig-
373 ure 4a-b) and precipitation (Figure 4d-e). Note that this binary ANN (SSP5-8.5 or SSP2-
374 4.5) has an overall accuracy of 92% and average F1 score of 92% when evaluated on the
375 SPEAR testing ensemble members for temperature and returns an accuracy of 89% and
376 average F1 score of 89% for precipitation.

377 Next, we use another binary ANN that classifies a temperature or precipitation map
378 but this time as either SSP1-1.9 or SSP2-4.5 (testing data accuracy = 93% and average
379 F1 score = 93% for temperature; testing data accuracy = 91% and average F1 score =
380 91% for precipitation). This shift in climate scenario classification only occurs for data

381 from SSP5-3.4OS_10ye (Figure 3c-d), and therefore we only evaluate these difference in
382 relevance maps for the experiment where climate mitigation begins in 2031 (Figure 4c,f).

383 Since our XAI method returns a relevance heatmap for every year fed into the ANN,
384 we can therefore assemble these composites that show the difference in the relevance maps
385 around these transition periods for SSP5-3.4OS and SSP5-3.4OS_10ye. These XAI dif-
386 ferences are shown in Figure 4 and are calculated by taking the ensemble mean of the
387 five years after each transition period minus the five years before each transition period.
388 We can then interpret positive areas of relevance as locations that pushed the ANN to
389 select the later climate scenario class. For example, positive areas of relevance in Fig-
390 ure 4a are temperature features that made the ANN more likely to predict SSP2-4.5, and
391 negative relevance can then be interpreted as the opposite. These overall transition pe-
392 riods are outlined by the red lines in Figure 3 by considering whether the climate sce-
393 nario is predicted by at least 50% or 80% of the 30 ensemble members. Note that the
394 specific years and the raw data for the temperature and precipitation differences are dis-
395 played in a corresponding Figure S13. Although we acknowledge that these thresholds
396 are somewhat arbitrary, the purpose of this analysis is just to gain some broader insight
397 on how XAI tools could be used to investigate why there are robust and rapid switches
398 in climate scenario classifications associated with the aggressive mitigation runs. A closer
399 examination of these overshoot simulations is left for future work.

400 In general, we find that the North Atlantic is an important regional indicator dur-
401 ing these mean shifts in climate scenario classifications after the onset of climate mit-
402 igation for both inputs of temperature and precipitation (Figure 4). This relevance fea-
403 ture is consistent with a pattern of North Atlantic temperature anomalies that can be
404 influenced by the strength of AMOC (R. Zhang et al., 2019; Delworth et al., 2022), which
405 can have substantial implications for the magnitude of the global climate response (Bellomo
406 et al., 2021). Central Africa is another region of larger differences in relevance around
407 transition periods, which aligns closely with looking at the raw data differences shown
408 in Figure S13. For instance, the reduced precipitation over Central Africa in the late 21st
409 century under SSP5-3.4OS_10ye forcing (Figure S13f) is an important regional change
410 for pushing the ANN to begin predicting SSP1-1.9 instead of SSP2-4.5 (Figure 4f). Other
411 prominent features include the notable contrast in relevance between hemispheres for the
412 transition around predicting SSP2-4.5 to SSP1-1.9 with temperature (Figure 4c). This
413 is likely related to the larger cooling signal observed by the simulation with the SSP5-

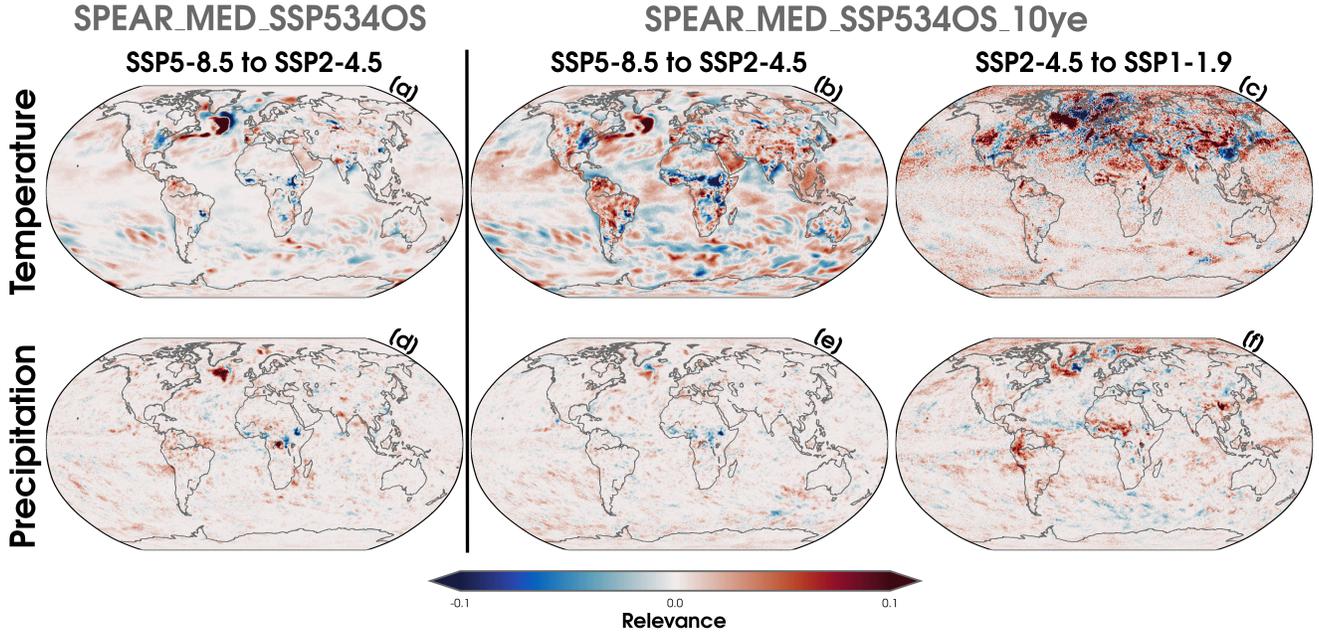


Figure 4. (a) Difference in the explainability spatial heatmaps for the ensemble mean of SSP5-3.4OS temperature predictions for the five years after the transition period in classifications from SSP5-8.5 to SSP2-4.5 minus the five years before the transition period. This transition period is designated by the vertical red lines outlined in Figure 3a. (b) As in (a), but for the ensemble mean of predictions using SSP5-3.4OS_10ye. This transition period is designated by the vertical red lines outlined in Figure 3c. (c) As in (b), but for the five years after the transition period in classifications from SSP2-4.5 to SSP1-1.9 subtracted by the five years before this transition period. The coarser appearance of this specific relevance composite for temperature inputs is due to the smaller ridge parameter selected for this binary ANN (Table S2). (d-f) As in (a-c), but for maps of precipitation using the transition periods outlined in Figure 3c,d.

414 3.4OS_10ye radiative forcing (Figure S13c), particularly over land. Regarding the pre-
415 cipitation XAI maps, we find that signals in the tropics are important for the ANN to
416 identify switches in the climate scenario classifications, but this appears less important
417 over the eastern Pacific Ocean and Indian Ocean basins (Figure 4d-f).

418 Lastly, we also highlight differences in the XAI heatmaps when compositing the
419 SSP5-3.4OS and SSP5-3.4OS_10ye simulations by their respective scenario predicted us-
420 ing the 5-class ANN for temperature and precipitation (Figures S13). Having said that,
421 we observe that the historical- and natural-forcing scenarios are rarely predicted for the
422 overshoot simulations, so the sample sizes of the mean relevance plots vary substantially
423 (Figure S14-S15). These relevance fields closely mirror the ones from the testing ensem-
424 ble members in Figure 2 and support our conclusion that the ANNs are learning to spatially-
425 weight distinctive temperature and precipitation features.

426 **4 Summary and Conclusions**

427 In our new detection method, we find that an ANN can skillfully identify a global
428 map with its associated radiative forcing scenario, even for a lower signal-to-noise vari-
429 able like precipitation (Hegerl et al., 2004; King et al., 2015; H. Zhang & Delworth, 2018;
430 Hawkins et al., 2020). By weighting spatial information, such as fingerprint patterns of
431 localized climate change, we find that this framework can identify between different ra-
432 diative forcing scenarios despite large internal variability and at times which share over-
433 lapping global mean characteristics. Then, by applying this framework to two overshoot
434 simulations, we show how this methodology can be used to reveal a difference in the av-
435 erage climate scenario impacts predicted over the 21st century after mitigation. In this
436 example, when aggressive climate mitigation efforts starts in 2031, we find that SSP1-
437 1.9 is predominately predicted by the 2070s for both temperature and precipitation. In
438 contrast, when climate mitigation instead begins in 2040, we find that SSP2-4.5 is clas-
439 sified for this same decadal period through the end of the run in 2100. This result in-
440 dicates that starting rapid mitigation in as little as a decade earlier can reduce the ex-
441 pected climate impacts that are typically associated with a more moderate emission sce-
442 nario (SSP2-4.5) compared to the lowest emission scenario (SSP1-1.9). Although we started
443 using XAI to explore the key regions of change associated with the climate scenario clas-
444 sifications, a deeper investigation into the physical responses associated with the tim-
445 ing of mitigation is crucial for assessing future climate risks, especially at the local level

446 (Diffenbaugh et al., 2023). While there is some spread in the specific classifications be-
447 tween the individual ensemble members due to internal variability in the earlier part of
448 the 21st century, we find that the majority of predictions are consistent by the mid 2020s.

449 More broadly speaking, this study highlights the benefit of this machine learning
450 approach for identifying time-evolving climate patterns and anomalies unique to differ-
451 ent radiative forcing scenarios, even in a single ensemble member with one realization
452 of internal variability. Large ensembles of additional radiative forcing simulations may
453 therefore not be needed when evaluating the ANNs after the training process. Given the
454 sensitivity of this neural network framework to learning crucial local spatial information,
455 it is conceivable that this architecture could also be extended to compare observations
456 with other climate modeling systems such as those that differ by examining new param-
457 eterization schemes, coupled model components, or sensitivities to different external forc-
458 ings. Alternatively, future work could investigate using spatial maps from multiple vari-
459 ables simultaneously, which might elucidate unique fingerprint patterns for compound
460 climate extremes across local scales.

461 The utility for near real-time monitoring of observations is a natural next exten-
462 sion of this work. Nevertheless, there are several remaining challenges. First, the ANNs
463 here are only trained on large ensemble experiments using a single GCM, and therefore
464 it is likely the ANN has learned any inherent biases associated with the SPEAR model
465 itself. Second, a key foundation of this work is on the availability of a large number of
466 ensemble members for training the ANN to learn each climate change scenario, which
467 allows the ANN to learn to distinguish the forced response from internal variability (Milinski
468 et al., 2020; Jain et al., 2023). This data availability is currently limited for other pub-
469 licly available initial-condition large ensembles, but it could be possible for a limited num-
470 ber of models such as MIROC6-LE (Shiogama et al., 2023) and SMHI-LENS (Wyser et
471 al., 2021). Third, and possibly the largest caveat to this work, is related to the constraints
472 of the classification scheme itself. In other words, the training here is limited to the pre-
473 diction of only a few pre-selected radiative forcing scenarios. In reality, the evolution of
474 greenhouse gases will not perfectly follow any of these scenario boundaries, and there-
475 fore how scientists reframe the development of new climate model scenarios for CMIP7
476 and beyond (e.g., Meinshausen et al., 2023; Nature, 2023; Sanderson et al., 2023) will
477 play a key role in how this detection method can be expanded in the future, particularly
478 as it pertains to more relevant regional applications for the climate services community.

479 **Open Research Section**

480 SPEAR_MED is described in Delworth et al. (2020), and our computational soft-
 481 ware is documented in Text S2. Data for the historical and SSP5-8.5 scenarios are avail-
 482 able from the SPEAR large ensemble data portal at GFDL (2023). Data for the other
 483 scenarios can be retrieved at Labe, Delworth, et al. (2023).

484 **Conflict of Interest**

485 The Authors declare no conflicts of interest for this study.

486 **Acknowledgments**

487 We appreciate the thoughtful reviews by Dr. Qinxue Gu and Dr. Tsung-Lin Hsieh on
 488 an earlier version of this work. This study was prepared by Zachary M. Labe under award
 489 NA18OAR4320123 from NOAA, U.S. Department of Commerce and through support
 490 from the NOAA Climate Program Office. High-performance computing was through base
 491 funding of GFDL provided by NOAA. The statements, findings, conclusions, and rec-
 492 ommendations are those of the authors and do not necessarily reflect the views of NOAA,
 493 or the U.S. Department of Commerce.

494 **References**

- 495 Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., ...
 496 Zhang, R. (2019, 10). The gfdl global ocean and sea ice model om4.0: Model
 497 description and simulation features. *Journal of Advances in Modeling Earth*
 498 *Systems*, *11*, 3167-3211. Retrieved from [https://agupubs.onlinelibrary](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS001726)
 499 [.wiley.com/doi/10.1029/2019MS001726](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS001726) doi: 10.1029/2019MS001726
- 500 Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., & Samek, W.
 501 (2015, 7). On pixel-wise explanations for non-linear classifier decisions by
 502 layer-wise relevance propagation. *PLoS ONE*, *10*, e0130140. Retrieved from
 503 <http://www.hfsp.org/>, doi: 10.1371/journal.pone.0130140
- 504 Baehrens, D., Schroeter, T., Harmeling, S., Kawanabe, M., Hansen, K., & Müller,
 505 K. R. (2010). How to explain individual classification decisions. *Journal of*
 506 *Machine Learning Research*, *11*.
- 507 Barnes, E. A., Toms, B., Hurrell, J. W., Ebert-Uphoff, I., Anderson, C., & An-
 508 derson, D. (2020, 9). Indicator patterns of forced change learned by an

- 509 artificial neural network. *Journal of Advances in Modeling Earth Systems*,
 510 *12*. Retrieved from [https://onlinelibrary.wiley.com/doi/10.1029/](https://onlinelibrary.wiley.com/doi/10.1029/2020MS002195)
 511 [2020MS002195](https://onlinelibrary.wiley.com/doi/10.1029/2020MS002195) doi: 10.1029/2020MS002195
- 512 Bellomo, K., Angeloni, M., Corti, S., & von Hardenberg, J. (2021). Future
 513 climate change shaped by inter-model differences in atlantic meridional
 514 overturning circulation response. *Nature Communications*, *12*. doi:
 515 [10.1038/s41467-021-24015-w](https://doi.org/10.1038/s41467-021-24015-w)
- 516 Bostrom, A., Demuth, J. L., Wirz, C. D., Cains, M. G., Schumacher, A., Mad-
 517 lambayan, D., ... Williams, J. K. (2023). Trust and trustworthy artificial
 518 intelligence: A research agenda for ai in the environmental sciences. *Risk Anal-*
 519 *ysis*. Retrieved from [https://onlinelibrary.wiley.com/doi/10.1111/](https://onlinelibrary.wiley.com/doi/10.1111/risa.14245)
 520 [risa.14245](https://onlinelibrary.wiley.com/doi/10.1111/risa.14245) doi: 10.1111/RISA.14245
- 521 Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti,
 522 R. (2020, 11). Reduced global warming from cmip6 projections when weight-
 523 ing models by performance and independence. *Earth System Dynamics*, *11*,
 524 995-1012. doi: [10.5194/ESD-11-995-2020](https://doi.org/10.5194/ESD-11-995-2020)
- 525 Brunner, L., & Sippel, S. (2023). Identifying climate models based on their daily
 526 output using machine learning. *Environmental Data Science*, *2*. doi: [10.1017/](https://doi.org/10.1017/eds.2023.23)
 527 [eds.2023.23](https://doi.org/10.1017/eds.2023.23)
- 528 Burgess, M. G., Ritchie, J., Shapland, J., & Pielke, R. (2020, 12). Ipccl baseline sce-
 529 narios have over-projected co2 emissions and economic growth. *Environmental*
 530 *Research Letters*, *16*, 014016. Retrieved from [https://iopscience.iop.org/](https://iopscience.iop.org/article/10.1088/1748-9326/abcd2)
 531 [article/10.1088/1748-9326/abcd2](https://iopscience.iop.org/article/10.1088/1748-9326/abcd2) doi: [10.1088/1748-9326/ABCDD2](https://doi.org/10.1088/1748-9326/ABCDD2)
- 532 Bône, C., Gastineau, G., Thiria, S., Gallinari, P., & Mejia, C. (2023, 10). Detec-
 533 tion and attribution of climate change using a neural network. *Journal of*
 534 *Advances in Modeling Earth Systems*, *15*, e2022MS003475. Retrieved from
 535 <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022MS003475>
 536 doi: [10.1029/2022MS003475](https://doi.org/10.1029/2022MS003475)
- 537 Davis, S. J., Lewis, N. S., Shaner, M., Aggarwal, S., Arent, D., Azevedo, I. L., ...
 538 Caldeira, K. (2018, 6). Net-zero emissions energy systems. *Science*, *360*.
 539 Retrieved from <https://www.science.org/doi/10.1126/science.aas9793>
 540 doi: [10.1126/SCIENCE.AAS9793/SUPPL_FILE/AAS9793-DAVIS-SM.PDF](https://doi.org/10.1126/SCIENCE.AAS9793/SUPPL_FILE/AAS9793-DAVIS-SM.PDF)
- 541 de Burgh-Day, C. O., & Leeuwenburg, T. (2023, 11). Machine learning for numeri-

- 542 cal weather and climate modelling: a review. *Geoscientific Model Development*,
 543 16, 6433-6477. Retrieved from [https://gmd.copernicus.org/articles/16/](https://gmd.copernicus.org/articles/16/6433/2023/)
 544 6433/2023/ doi: 10.5194/GMD-16-6433-2023
- 545 de Kleijne, K., Hanssen, S. V., van Dinteren, L., Huijbregts, M. A., van Zelm, R.,
 546 & de Coninck, H. (2022, 2). Limits to paris compatibility of co2 capture and
 547 utilization. *One Earth*, 5, 168-185. doi: 10.1016/J.ONEEAR.2022.01.006
- 548 Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J.-H., Dunne,
 549 K. A., ... Zhao, M. (2020, 3). Spear: The next generation gfdl modeling
 550 system for seasonal to multidecadal prediction and projection. *Journal of*
 551 *Advances in Modeling Earth Systems*, 12, e2019MS001895. Retrieved from
 552 <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS001895>
 553 doi: 10.1029/2019MS001895
- 554 Delworth, T. L., Cooke, W. F., Naik, V., Paynter, D., & Zhang, L. (2022, 8). A
 555 weakened amoc may prolong greenhouse gas-induced mediterranean dry-
 556 ing even with significant and rapid climate change mitigation. *Proceed-*
 557 *ings of the National Academy of Sciences of the United States of Amer-*
 558 *ica*, 119, e2116655119. Retrieved from [https://www.pnas.org/doi/abs/](https://www.pnas.org/doi/abs/10.1073/pnas.2116655119)
 559 [10.1073/pnas.2116655119](https://www.pnas.org/doi/abs/10.1073/pnas.2116655119) doi: 10.1073/PNAS.2116655119/SUPPL_FILE/
 560 PNAS.2116655119.SAPP.PDF
- 561 Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N.,
 562 ... Ting, M. (2020, 3). Insights from earth system model initial-condition
 563 large ensembles and future prospects. *Nature Climate Change*, 1-10. Re-
 564 trieved from <http://www.nature.com/articles/s41558-020-0731-2> doi:
 565 10.1038/s41558-020-0731-2
- 566 Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012, 2). Uncertainty in climate
 567 change projections: the role of internal variability. *Climate Dynamics*, 38, 527-
 568 546. Retrieved from [http://link.springer.com/10.1007/s00382-010-0977](http://link.springer.com/10.1007/s00382-010-0977-x)
 569 [-x](http://link.springer.com/10.1007/s00382-010-0977-x) doi: 10.1007/s00382-010-0977-x
- 570 Diffenbaugh, N. S., Barnes, E. A., & Keys, P. W. (2023). Probability of continued
 571 local-scale warming and extreme events during and after decarbonization. *En-*
 572 *vironmental Research: Climate*, 2. doi: 10.1088/2752-5295/accf2f
- 573 Dong, L., McPhaden, M. J., Deng, X., Huang, Y., Qin, Z., Fan, X., ... Xing, C.
 574 (2020, 10). Global surface air temperatures in cmip6: historical performance

- 575 and future changes. *Environmental Research Letters*, *15*, 104056. Retrieved
576 from [https://iopscience.iop.org/article/10.1088/1748-9326/abb051/](https://iopscience.iop.org/article/10.1088/1748-9326/abb051/meta)
577 [meta doi: 10.1088/1748-9326/ABB051](https://doi.org/10.1088/1748-9326/ABB051)
- 578 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E.,
579 Hauck, J., ... Zheng, B. (2023, 12). Global carbon budget 2023. *Earth System*
580 *Science Data*, *15*, 5301-5369. Retrieved from [https://essd.copernicus.org/](https://essd.copernicus.org/articles/15/5301/2023/)
581 [articles/15/5301/2023/](https://doi.org/10.5194/ESSD-15-5301-2023) doi: 10.5194/ESSD-15-5301-2023
- 582 Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck,
583 J., ... Zheng, B. (2022). Global carbon budget 2022. *Earth System Science*
584 *Data*, *14*. doi: 10.5194/essd-14-4811-2022
- 585 Fuss, S., Lamb, W. F., Callaghan, M. W., Hilaire, J., Creutzig, F., Amann, T., ...
586 Minx, J. C. (2018). *Negative emissions - part 2: Costs, potentials and side*
587 *effects* (Vol. 13). doi: 10.1088/1748-9326/aabf9f
- 588 GFDL. (2023). *Gfdl cmip6/spear (version 1) [dataset]*. Retrieved from [https://www](https://www.gfdl.noaa.gov/spear_large_ensembles/)
589 [.gfdl.noaa.gov/spear_large_ensembles/](https://www.gfdl.noaa.gov/spear_large_ensembles/)
- 590 Gordon, E. M., Barnes, E. A., & Davenport, F. V. (2023, 10). Separating
591 internal and forced contributions to near term sst predictability in the
592 cesm2-le. *Environmental Research Letters*, *18*, 104047. Retrieved from
593 <https://iopscience.iop.org/article/10.1088/1748-9326/acfdbc> doi:
594 10.1088/1748-9326/ACFDDBC
- 595 Hausfather, Z., Drake, H. F., Abbott, T., & Schmidt, G. A. (2020). Evaluating the
596 performance of past climate model projections. *Geophysical Research Letters*,
597 *47*. doi: 10.1029/2019GL085378
- 598 Hausfather, Z., & Peters, G. P. (2020). *Rcp8.5 is a problematic scenario for near-*
599 *term emissions* (Vol. 117). doi: 10.1073/pnas.2017124117
- 600 Hawkins, E., Frame, D., Harrington, L., Joshi, M., King, A., Rojas, M., & Sutton,
601 R. (2020, 3). Observed emergence of the climate change signal: From the
602 familiar to the unknown. *Geophysical Research Letters*, *47*. Retrieved from
603 <https://onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086259> doi:
604 10.1029/2019GL086259
- 605 Hawkins, E., & Sutton, R. (2009, 8). The potential to narrow uncertainty in regional
606 climate predictions. *Bulletin of the American Meteorological Society*, *90*, 1095-
607 1107. doi: 10.1175/2009BAMS2607.1

- 608 Hawkins, E., & Sutton, R. (2011). The potential to narrow uncertainty in projec-
 609 tions of regional precipitation change. *Climate Dynamics*, *37*. doi: 10.1007/
 610 s00382-010-0810-6
- 611 Hegerl, G. C., Zwiers, F. W., Stott, P. A., & Kharin, V. V. (2004). Detectabil-
 612 ity of anthropogenic changes in annual temperature and precipitation ex-
 613 tremes. *Journal of Climate*, *17*. doi: 10.1175/1520-0442(2004)017<3683:
 614 DOACIA>2.0.CO;2
- 615 Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., ...
 616 Zadeh, N. (2019, 11). Structure and performance of gfdl's cm4.0 climate
 617 model. *Journal of Advances in Modeling Earth Systems*, *11*, 3691-3727. Re-
 618 trieved from [https://agupubs.onlinelibrary.wiley.com/doi/10.1029/
 619 2019MS001829](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS001829) doi: 10.1029/2019MS001829
- 620 Huard, D., Fyke, J., Capellán-Pérez, I., Matthews, H. D., & Partanen, A. I. (2022,
 621 10). Estimating the likelihood of ghg concentration scenarios from probabilistic
 622 integrated assessment model simulations. *Earth's Future*, *10*, e2022EF002715.
 623 Retrieved from [https://agupubs.onlinelibrary.wiley.com/doi/10.1029/
 624 2022EF002715](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022EF002715) doi: 10.1029/2022EF002715
- 625 Huntingford, C., Jeffers, E. S., Bonsall, M. B., Christensen, H. M., Lees, T., &
 626 Yang, H. (2019). Machine learning and artificial intelligence to aid climate
 627 change research and preparedness. *Environmental Research Letters*, *14*. doi:
 628 10.1088/1748-9326/ab4e55
- 629 Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., ...
 630 Zhang, X. (2020). Harmonization of global land use change and management
 631 for the period 850-2100 (luh2) for cmip6. *Geoscientific Model Development*,
 632 *13*. doi: 10.5194/gmd-13-5425-2020
- 633 IPCC. (2022). *Ipc climate change 2022: Mitigation of climate change*. doi: 10.1017/
 634 9781009157926
- 635 Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., &
 636 Saynisch-Wagner, J. (2021, 8). Towards neural earth system modelling by
 637 integrating artificial intelligence in earth system science. *Nature Machine In-
 638 telligence*, *3*, 667-674. Retrieved from [https://www.nature.com/articles/
 639 s42256-021-00374-3](https://www.nature.com/articles/s42256-021-00374-3) doi: 10.1038/s42256-021-00374-3
- 640 Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021). Formalizing trust in

- 641 artificial intelligence: Prerequisites, causes and goals of human trust in ai.. doi:
642 10.1145/3442188.3445923
- 643 Jain, S., Scaife, A. A., Shepherd, T. G., Deser, C., Dunstone, N., Schmidt, G. A.,
644 ... Turkington, T. (2023). Importance of internal variability for cli-
645 mate model assessment. *npj Climate and Atmospheric Science*, 6. doi:
646 10.1038/s41612-023-00389-0
- 647 King, A. D., Donat, M. G., Fischer, E. M., Hawkins, E., Alexander, L. V., Karoly,
648 D. J., ... Perkins, S. E. (2015, 9). The timing of anthropogenic emergence
649 in simulated climate extremes. *Environmental Research Letters*, 10, 094015.
650 Retrieved from [https://iopscience.iop.org/article/10.1088/1748-9326/](https://iopscience.iop.org/article/10.1088/1748-9326/10/9/094015/meta)
651 [10/9/094015/meta](https://iopscience.iop.org/article/10.1088/1748-9326/10/9/094015/meta) doi: 10.1088/1748-9326/10/9/094015
- 652 Kriegler, E., Bauer, N., Popp, A., Humpenöder, F., Leimbach, M., Streffer, J., ...
653 Edenhofer, O. (2017, 1). Fossil-fueled development (ssp5): An energy and
654 resource intensive scenario for the 21st century. *Global Environmental Change*,
655 42, 297-315. doi: 10.1016/J.GLOENVCHA.2016.05.015
- 656 Labe, Z. M., & Barnes, E. A. (2022, 7). Comparison of climate model large en-
657 sembles with observations in the arctic using simple neural networks. *Earth*
658 *and Space Science*, 9, e2022EA002348. Retrieved from [https://doi.org/](https://doi.org/10.1029/2022EA002348)
659 [10.1029/2022EA002348](https://doi.org/10.1029/2022EA002348) doi: 10.1029/2022EA002348
- 660 Labe, Z. M., Barnes, E. A., & Hurrell, J. W. (2023, 3). Identifying the regional
661 emergence of climate patterns in the arise-sai-1.5 simulations. *Environmental*
662 *Research Letters*, 18, 1-12. Retrieved from [https://iopscience.iop.org/](https://iopscience.iop.org/article/10.1088/1748-9326/acc81a)
663 [article/10.1088/1748-9326/acc81a](https://iopscience.iop.org/article/10.1088/1748-9326/acc81a) doi: 10.1088/1748-9326/ACC81A
- 664 Labe, Z. M., Delworth, T. L., Johnson, N. C., & Cooke, W. F. (2023). *Spear cli-*
665 *mate scenario projections (version 1) [dataset]*. Retrieved from [https://](https://zenodo.org/doi/10.5281/zenodo.10083256)
666 zenodo.org/doi/10.5281/zenodo.10083256
- 667 Labe, Z. M., Johnson, N. C., & Delworth, T. L. (2024, 2). Changes in united states
668 summer temperatures revealed by explainable neural networks. *Earth's Future*,
669 12, e2023EF003981. Retrieved from [https://onlinelibrary.wiley.com/](https://onlinelibrary.wiley.com/doi/abs/10.1029/2023EF003981)
670 [doi/abs/10.1029/2023EF003981](https://onlinelibrary.wiley.com/doi/abs/10.1029/2023EF003981) doi: 10.1029/2023EF003981
- 671 Lehner, F., & Deser, C. (2023, 4). Origin, importance, and predictive limits of in-
672 ternal climate variability. *Environmental Research: Climate*. Retrieved from
673 <https://iopscience.iop.org/article/10.1088/2752-5295/accf30/meta>

- 674 doi: 10.1088/2752-5295/ACCF30
- 675 Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E., Brunner, L., ...
676 Hawkins, E. (2020). Partitioning climate projection uncertainty with mul-
677 tiple large ensembles and cmip5/6. *Earth System Dynamics Discussions*, 1-28.
678 doi: 10.5194/esd-2019-93
- 679 Liang, Y., Gillett, N. P., & Monahan, A. H. (2020). Climate model projections of
680 21st century global warming constrained using the observed warming trend.
681 *Geophysical Research Letters*, 47. doi: 10.1029/2019GL086757
- 682 Liu, Z., Deng, Z., Davis, S., & Ciais, P. (2023). *Monitoring global carbon emissions*
683 *in 2022* (Vol. 4). doi: 10.1038/s43017-023-00406-z
- 684 MacDougall, A. H., Frölicher, T. L., Jones, C. D., Rogelj, J., DamonMatthews, H.,
685 Zickfeld, K., ... Ziehn, T. (2020). Is there warming in the pipeline? a multi-
686 model analysis of the zero emissions commitment from co2. *Biogeosciences*,
687 17. doi: 10.5194/bg-17-2987-2020
- 688 Maher, N., Lehner, F., & Marotzke, J. (2020, 5). Quantifying the role of internal
689 variability in the temperature we expect to observe in the coming decades. *En-*
690 *vironmental Research Letters*, 15, 054014. Retrieved from [https://doi.org/](https://doi.org/10.1088/1748-9326/ab7d02)
691 10.1088/1748-9326/ab7d02 doi: 10.1088/1748-9326/ab7d02
- 692 Mamalakis, A., Ebert-Uphoff, I., & Barnes, E. A. (2022a). Explainable arti-
693 ficial intelligence in meteorology and climate science: Model fine-tuning,
694 calibrating trust and learning new science. In (Vol. 13200 LNAI, p. 315-
695 339). Springer Science and Business Media Deutschland GmbH. doi:
696 10.1007/978-3-031-04083-2_16
- 697 Mamalakis, A., Ebert-Uphoff, I., & Barnes, E. A. (2022b, 6). Neural net-
698 work attribution methods for problems in geoscience: A novel synthetic
699 benchmark dataset. *Environmental Data Science*, 1, e8. Retrieved
700 from [https://www.cambridge.org/core/journals/environmental](https://www.cambridge.org/core/journals/environmental-data-science/article/neural-network-attribution-methods-for-problems-in-geoscience-a-novel-synthetic-benchmark-dataset/DDA562FC7B9A2B30710582861920860E)
701 [-data-science/article/neural-network-attribution-methods-for](https://www.cambridge.org/core/journals/environmental-data-science/article/neural-network-attribution-methods-for-problems-in-geoscience-a-novel-synthetic-benchmark-dataset/DDA562FC7B9A2B30710582861920860E)
702 [-problems-in-geoscience-a-novel-synthetic-benchmark-dataset/](https://www.cambridge.org/core/journals/environmental-data-science/article/neural-network-attribution-methods-for-problems-in-geoscience-a-novel-synthetic-benchmark-dataset/DDA562FC7B9A2B30710582861920860E)
703 DDA562FC7B9A2B30710582861920860E doi: 10.1017/EDS.2022.7
- 704 Marotzke, J. (2019). Quantifying the irreducible uncertainty in near-term climate
705 projections. *Wiley Interdisciplinary Reviews: Climate Change*, 10. doi: 10
706 .1002/wcc.563

- 707 McGovern, A., Lagerquist, R., Gagne, D. J., Jergensen, G. E., Elmore, K. L., Home-
 708 yer, C. R., & Smith, T. (2019, 11). Making the black box more transpar-
 709 ent: Understanding the physical implications of machine learning. *Bulletin*
 710 *of the American Meteorological Society*, *100*, 2175-2199. Retrieved from
 711 [http://journals.ametsoc.org/bams/article-pdf/100/11/2175/4876688/](http://journals.ametsoc.org/bams/article-pdf/100/11/2175/4876688/bams-d-18-0195_1.pdf)
 712 [bams-d-18-0195_1.pdf](http://journals.ametsoc.org/bams/article-pdf/100/11/2175/4876688/bams-d-18-0195_1.pdf) doi: 10.1175/BAMS-D-18-0195.1
- 713 Medhaug, I., Stolpe, M. B., Fischer, E. M., & Knutti, R. (2017, 5). Reconcil-
 714 ing controversies about the ‘global warming hiatus’. *Nature*, *545*, 41-47.
 715 Retrieved from <https://www.nature.com/articles/nature22315> doi:
 716 10.1038/nature22315
- 717 Meinshausen, M., Nicholls, Z. R., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., ...
 718 Wang, R. H. (2020). The shared socio-economic pathway (ssp) greenhouse gas
 719 concentrations and their extensions to 2500. *Geoscientific Model Development*,
 720 *13*. doi: 10.5194/gmd-13-3571-2020
- 721 Meinshausen, M., Schleussner, C.-F., Beyer, K., Bodeker, G., Boucher, O., Canadell,
 722 J. G., ... Nicholls, Z. (2023). A perspective on the next generation of earth
 723 system model scenarios: towards representative emission pathways (reps).
 724 *Geoscientific Model Development Discussions*.
- 725 Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge,
 726 D. M., ... Weiss, R. (2017). Historical greenhouse gas concentrations for
 727 climate modelling (cmip6). *Geoscientific Model Development*, *10*. doi:
 728 10.5194/gmd-10-2057-2017
- 729 Melnikova, I., Boucher, O., Cadule, P., Tanaka, K., Gasser, T., Hajima, T., ...
 730 Ciais, P. (2022). Impact of bioenergy crop expansion on climate-carbon
 731 cycle feedbacks in overshoot scenarios. *Earth System Dynamics*, *13*. doi:
 732 10.5194/esd-13-779-2022
- 733 Milinski, S., Maher, N., & Olonscheck, D. (2020, 10). How large does a large en-
 734 semble need to be? *Earth System Dynamics*, *11*, 885-901. Retrieved from
 735 <https://esd.copernicus.org/articles/11/885/2020/> doi: 10.5194/esd-11
 736 -885-2020
- 737 Minx, J. C., Lamb, W. F., Callaghan, M. W., Fuss, S., Hilaire, J., Creutzig, F.,
 738 ... Dominguez, M. D. M. Z. (2018). *Negative emissions - part 1: Research*
 739 *landscape and synthesis* (Vol. 13). doi: 10.1088/1748-9326/aabf9b

- 740 Nature. (2023, 6). Reversing climate overshoot. *Nature Geoscience* 2023 16:6,
741 16, 467-467. Retrieved from [https://www.nature.com/articles/s41561-023](https://www.nature.com/articles/s41561-023-01213-3)
742 -01213-3 doi: 10.1038/s41561-023-01213-3
- 743 NCAR. (2020). *Us clivar multi-model le archive*. Retrieved from [https://www.cesm](https://www.cesm.ucar.edu/projects/community-projects/MMLEA/)
744 [.ucar.edu/projects/community-projects/MMLEA/](https://www.cesm.ucar.edu/projects/community-projects/MMLEA/)
- 745 O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., ...
746 van Vuuren, D. P. (2014). A new scenario framework for climate change re-
747 search: The concept of shared socioeconomic pathways. *Climatic Change*, 122.
748 doi: 10.1007/s10584-013-0905-2
- 749 O'Neill, B. C., Tebaldi, C., Vuuren, D. P. V., Eyring, V., Friedlingstein, P., Hurtt,
750 G., ... Sanderson, B. M. (2016, 9). The scenario model intercomparison
751 project (scenariomip) for cmip6. *Geoscientific Model Development*, 9, 3461-
752 3482. doi: 10.5194/GMD-9-3461-2016
- 753 Park, I. H., & Yeh, S. W. (2024, 12). Projections of the north atlantic warm-
754 ing hole can be constrained using ocean surface density as an emergent
755 constraint. *Communications Earth and Environment*, 5. doi: 10.1038/
756 s43247-024-01269-y
- 757 Pedersen, J. S. T., Santos, F. D., van Vuuren, D., Gupta, J., Coelho, R. E.,
758 Aparício, B. A., & Swart, R. (2021). An assessment of the performance of
759 scenarios against historical global emissions for ipcc reports. *Global Environ-*
760 *mental Change*, 66. doi: 10.1016/j.gloenvcha.2020.102199
- 761 Peters, G. P., & Hausfather, Z. (2020). Emissions - the 'business as usual' story is
762 misleading. *Nature*, 577.
- 763 Pielke, R., Burgess, M. G., & Ritchie, J. (2022, 2). Plausible 2005–2050 emissions
764 scenarios project between 2c and 3c of warming by 2100. *Environmental Re-*
765 *search Letters*, 17, 024027. Retrieved from [https://iopscience.iop.org/](https://iopscience.iop.org/article/10.1088/1748-9326/ac4ebf/meta)
766 [article/10.1088/1748-9326/ac4ebf/meta](https://iopscience.iop.org/article/10.1088/1748-9326/ac4ebf/meta) doi: 10.1088/1748-9326/
767 AC4EBF
- 768 Po-Chedley, S., Fasullo, J. T., Siler, N., Labe, Z. M., Barnes, E. A., eline W Bon-
769 fils, C. J., ... by Dennis Hartmann, E. (2022, 11). Internal variability and
770 forcing influence model–satellite differences in the rate of tropical tropospheric
771 warming. *Proceedings of the National Academy of Sciences*, 119, e2209431119.
772 Retrieved from <https://www.pnas.org/doi/abs/10.1073/pnas.2209431119>

- 773 doi: 10.1073/PNAS.2209431119
- 774 Rader, J. K., Barnes, E. A., Ebert-Uphoff, I., & Anderson, C. (2022, 7). Detection
775 of forced change within combined climate fields using explainable neural net-
776 works. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002941.
777 Retrieved from [https://onlinelibrary.wiley.com/doi/full/10.1029/
778 2021MS002941](https://onlinelibrary.wiley.com/doi/full/10.1029/2021MS002941) doi: 10.1029/2021MS002941
- 779 Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fuji-
780 mori, S., ... Tavoni, M. (2017, 1). The shared socioeconomic path-
781 ways and their energy, land use, and greenhouse gas emissions implica-
782 tions: An overview. *Global Environmental Change*, *42*, 153-168. doi:
783 10.1016/J.GLOENVCHA.2016.05.009
- 784 Ribes, A., Qasmi, S., & Gillett, N. P. (2021). Making climate projections conditional
785 on historical observations. *Science Advances*, *7*. doi: 10.1126/sciadv.abc0671
- 786 Ritchie, J., & Dowlatabadi, H. (2017). Why do climate change scenarios return to
787 coal? *Energy*, *140*. doi: 10.1016/j.energy.2017.08.083
- 788 Rolnick, D., Kaack, L. H., School, H., KOCHANSKI, E. Z. K., Jaques, N., Brain,
789 G., ... Creutzig, F. (2022). Tackling climate change with machine learning.
790 *ACM Computing Surveys*, *55*, 96. Retrieved from [https://doi.org/10.1145/
791 3485128](https://doi.org/10.1145/3485128) doi: 10.1145/3485128
- 792 Samset, B. H., Fuglestedt, J. S., & Lund, M. T. (2020). Delayed emergence of
793 a global temperature response after emission mitigation. *Nature Communica-*
794 *tions*, *11*. doi: 10.1038/s41467-020-17001-1
- 795 Sanderson, B. M., Booth, B. B., Dunne, J., Eyring, V., Fisher, R. A., Friedlingstein,
796 P., ... hereon, H.-Z. (2023). The need for carbon emissions-driven climate
797 projections in cmip7. *EGUsphere [preprint]*, 1-51. Retrieved from [https://
798 doi.org/10.5194/egusphere-2023-2127](https://doi.org/10.5194/egusphere-2023-2127) doi: 10.5194/egusphere-2023-2127
- 799 Saunio, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson,
800 R. B., ... Zhuang, Q. (2020). The global methane budget 2000-2017. *Earth
801 System Science Data*, *12*. doi: 10.5194/essd-12-1561-2020
- 802 Shiogama, H., Tatebe, H., Hayashi, M., Abe, M., Arai, M., Koyama, H., ... Watan-
803 abe, M. (2023, 11). Miroc6 large ensemble (miroc6-le): experimental de-
804 sign and initial analyses. *Earth System Dynamics*, *14*, 1107-1124. doi:
805 10.5194/ESD-14-1107-2023

- 806 Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important features
807 through propagating activation differences. In (Vol. 7).
- 808 Shrikumar, A., Greenside, P., Shcherbina, A., & Kundaje, A. (2016). Not just a
809 black box: Learning important features through propagating activation dif-
810 ferences. *34th International Conference on Machine Learning, ICML 2017*,
811 7.
- 812 Sippel, S., Meinshausen, N., Fischer, E. M., Székely, E., & Knutti, R. (2020, 1).
813 *Climate change now detectable from any single day of weather at global scale*
814 (Vol. 10). Nature Research. doi: 10.1038/s41558-019-0666-7
- 815 Sippel, S., Meinshausen, N., Székely, E., Fischer, E., Pendergrass, A. G., Lehner,
816 F., & Knutti, R. (2021). Robust detection of forced warming in the pres-
817 ence of potentially large climate variability. *Science Advances*, 7. doi:
818 10.1126/sciadv.abh4429
- 819 Sognaes, I., Gambhir, A., van de Ven, D. J., Nikas, A., Anger-Kraavi, A., Bui, H.,
820 ... Peters, G. P. (2021). A multi-model analysis of long-term emissions and
821 warming implications of current mitigation efforts. *Nature Climate Change*,
822 11. doi: 10.1038/s41558-021-01206-3
- 823 Sonnewald, M., Lguensat, R., Jones, D. C., Dueben, P. D., Brajard, J., & Balaji,
824 V. (2021, 7). Bridging observations, theory and numerical simulation of the
825 ocean using machine learning. *Environmental Research Letters*, 16, 073008.
826 Retrieved from [https://iopscience.iop.org/article/10.1088/1748-9326/
827 ac0eb0/meta](https://iopscience.iop.org/article/10.1088/1748-9326/ac0eb0/meta) doi: 10.1088/1748-9326/AC0EB0
- 828 Stott, P., Good, P., Jones, G., Gillett, N., & Hawkins, E. (2013). The upper end of
829 climate model temperature projections is inconsistent with past warming. *En-
830 vironmental Research Letters*, 8. doi: 10.1088/1748-9326/8/1/014024
- 831 Sundararajan, M., Taly, A., & Yan, Q. (2017, 3). Axiomatic attribution for deep
832 networks. *34th International Conference on Machine Learning, ICML 2017*, 7,
833 5109-5118. Retrieved from <https://arxiv.org/abs/1703.01365v2> doi: 10
834 .48550/arxiv.1703.01365
- 835 Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., ...
836 Ziehn, T. (2021). Climate model projections from the scenario model inter-
837 comparison project (scenariomip) of cmip6. *Earth System Dynamics*, 12. doi:
838 10.5194/esd-12-253-2021

- 839 Tebaldi, C., & Friedlingstein, P. (2013). Delayed detection of climate miti-
 840 gation benefits due to climate inertia and variability. *Proceedings of the*
 841 *National Academy of Sciences of the United States of America*, 110. doi:
 842 10.1073/pnas.1300005110
- 843 Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F.,
 844 & Knutti, R. (2020). Past warming trend constrains future warming in cmip6
 845 models. *Science Advances*, 6. doi: 10.1126/sciadv.aaz9549
- 846 Toms, B. A., Barnes, E. A., & Ebert-Uphoff, I. (2020, 9). Physically interpretable
 847 neural networks for the geosciences: Applications to earth system variabil-
 848 ity. *Journal of Advances in Modeling Earth Systems*, 12. Retrieved from
 849 <https://onlinelibrary.wiley.com/doi/10.1029/2019MS002002> doi:
 850 10.1029/2019MS002002
- 851 Wills, R., Sippel, S., & Barnes, E. A. (2020). Separating forced and unforced
 852 components of climate change: The utility of pattern recognition methods
 853 in large ensembles and observations. *US CLIVAR Variations*, 18. doi:
 854 10.5065/0DSY-WH17
- 855 Wyser, K., Koenigk, T., Fladrich, U., Fuentes-Franco, R., Karami, M. P., &
 856 Kruschke, T. (2021, 7). The smhi large ensemble (smhi-lens) with
 857 ec-earth3.3.1. *Geoscientific Model Development*, 14, 4781-4796. doi:
 858 10.5194/GMD-14-4781-2021
- 859 Zhang, H., & Delworth, T. L. (2018). Robustness of anthropogenically forced
 860 decadal precipitation changes projected for the 21st century. *Nature Communi-*
 861 *cations*, 9. doi: 10.1038/s41467-018-03611-3
- 862 Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y. O., Marsh, R., Yeager, S. G.,
 863 ... Little, C. M. (2019). *A review of the role of the atlantic meridional over-*
 864 *turning circulation in atlantic multidecadal variability and associated climate*
 865 *impacts* (Vol. 57). doi: 10.1029/2019RG000644
- 866 Zhang, S., Zhou, Z., Peng, P., & Xu, C. (2023, 10). A new framework for esti-
 867 mating and decomposing the uncertainty of climate projections. *Journal*
 868 *of Climate*, -1. Retrieved from [https://journals.ametsoc.org/view/](https://journals.ametsoc.org/view/journals/clim/aop/JCLI-D-23-0064.1/JCLI-D-23-0064.1.xml)
 869 [journals/clim/aop/JCLI-D-23-0064.1/JCLI-D-23-0064.1.xml](https://journals.ametsoc.org/view/journals/clim/aop/JCLI-D-23-0064.1/JCLI-D-23-0064.1.xml) doi:
 870 10.1175/JCLI-D-23-0064.1
- 871 Zhao, M., Golaz, J. C., Held, I. M., Guo, H., Balaji, V., Benson, R., ... Xiang, B.

872 (2018a, 3). The gfdl global atmosphere and land model am4.0/lm4.0: 1. sim-
873 ulation characteristics with prescribed ssts. *Journal of Advances in Modeling*
874 *Earth Systems*, 10, 691-734. Retrieved from <https://agupubs.onlinelibrary>
875 [.wiley.com/doi/10.1002/2017MS001208](https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017MS001208) doi: 10.1002/2017MS001208

876 Zhao, M., Golaz, J. C., Held, I. M., Guo, H., Balaji, V., Benson, R., ... Xiang,
877 B. (2018b, 3). The gfdl global atmosphere and land model am4.0/lm4.0:
878 2. model description, sensitivity studies, and tuning strategies. *Jour-*
879 *nal of Advances in Modeling Earth Systems*, 10, 735-769. Retrieved from
880 <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017MS001209>
881 doi: 10.1002/2017MS001209