

LETTER • OPEN ACCESS

Combining climate models and observations to predict the time remaining until regional warming thresholds are reached

To cite this article: Elizabeth A Barnes *et al* 2025 *Environ. Res. Lett.* **20** 014008

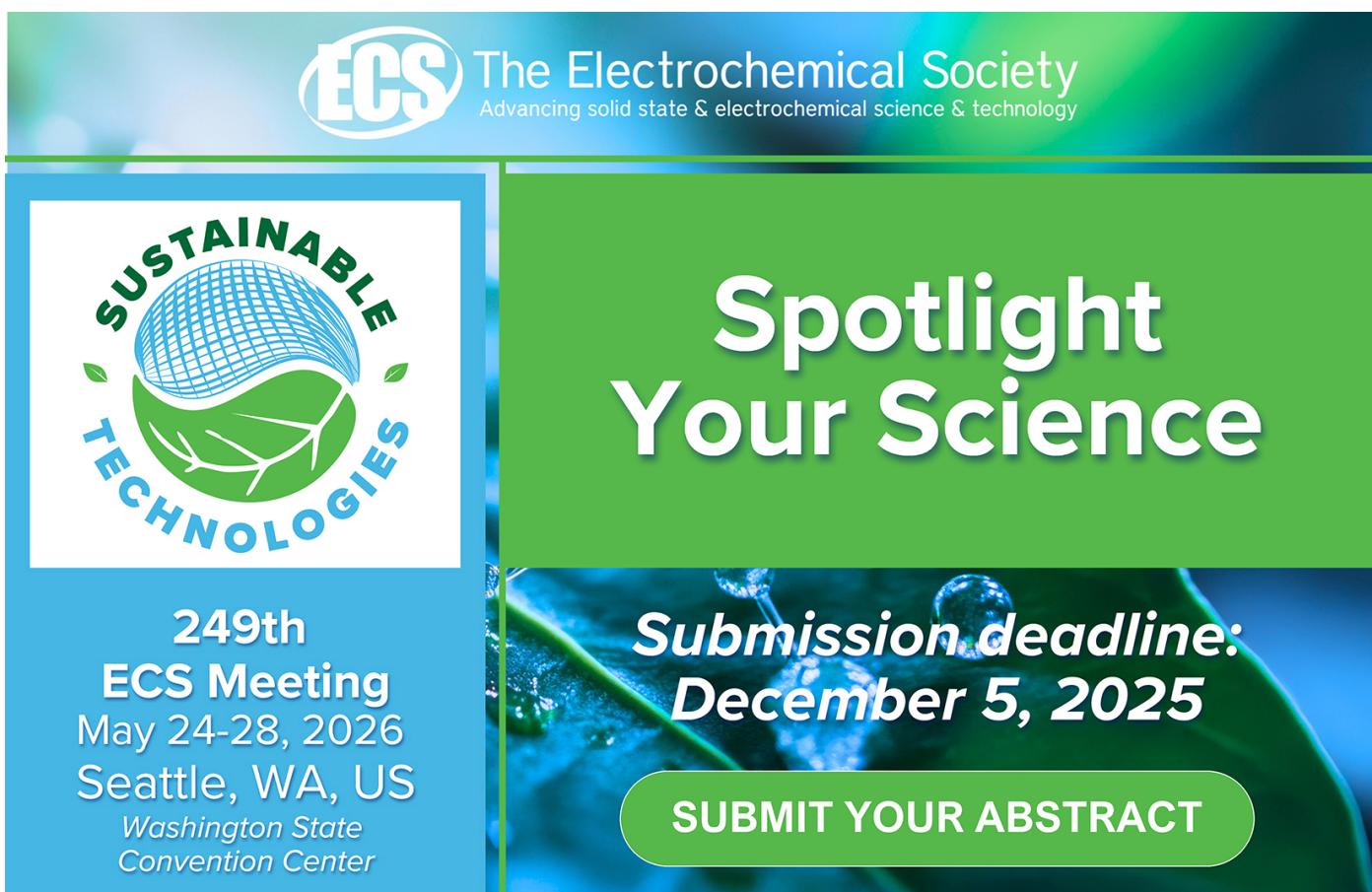
View the [article online](#) for updates and enhancements.

You may also like

- [Multi-messenger Observations of a Binary Neutron Star Merger](#)
B. P. Abbott, R. Abbott, T. D. Abbott et al.

- [An ensemble approach to quantify global mean sea-level rise over the 20th century from tide gauge reconstructions](#)
M D Palmer, C M Domingues, A B A Slanger et al.

- [Effects of Impact and Target Parameters on the Results of a Kinetic Impactor: Predictions for the Double Asteroid Redirection Test \(DART\) Mission](#)
Angela M. Stickle, Mallory E. DeCoster, Christoph Burger et al.



The advertisement features the ECS logo and the text "The Electrochemical Society Advancing solid state & electrochemical science & technology". On the left, there's a circular graphic for "SUSTAINABLE TECHNOLOGIES" featuring a stylized globe and leaves. Below it, the text reads: "249th ECS Meeting May 24-28, 2026 Seattle, WA, US Washington State Convention Center". On the right, the text "Spotlight Your Science" is prominently displayed, along with a "Submission deadline: December 5, 2025" and a "SUBMIT YOUR ABSTRACT" button.

ENVIRONMENTAL RESEARCH LETTERS



OPEN ACCESS

RECEIVED
14 March 2024

REVISED
25 October 2024

ACCEPTED FOR PUBLICATION
12 November 2024

PUBLISHED
10 December 2024

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



LETTER

Combining climate models and observations to predict the time remaining until regional warming thresholds are reached

Elizabeth A Barnes^{1,*} , Noah S Diffenbaugh² and Sonia I Seneviratne³

¹ Department of Atmospheric Science, Colorado State University, Fort Collins, CO, United States of America

² Doerr School of Sustainability, Stanford University, Stanford, CA, United States of America

³ Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

* Author to whom any correspondence should be addressed.

E-mail: eabarnes@colostate.edu

Keywords: climate change, machine learning, climate models, regional warming, transfer learning, regional climate change

Supplementary material for this article is available [online](#)

Abstract

The importance of climate change for driving adverse climate impacts has motivated substantial effort to understand the rate and magnitude of regional climate change in different parts of the world. However, despite decades of research, there is substantial uncertainty in the time remaining until specific regional temperature thresholds are reached, with climate models often disagreeing both on the warming that has occurred to-date, as well as the warming that might be experienced in the next few decades. Here, we adapt a recent machine learning approach to train a convolutional neural network to predict the time (and its uncertainty) until different regional warming thresholds are reached based on the current state of the climate system. In addition to predicting regional rather than global warming thresholds, we include a transfer learning step in which the climate-model-trained network is fine-tuned with limited observations, which further improves predictions of the real world. Using observed 2023 temperature anomalies to define the current climate state, our method yields a central estimate of 2040 or earlier for reaching the 1.5 °C threshold for all regions where transfer learning is possible, and a central estimate of 2040 or earlier for reaching the 2.0 °C threshold for 31 out of 34 regions. For 3.0 °C, 26 °C out of 34 regions are predicted to reach the threshold by 2060. Our results highlight the power of transfer learning as a tool to combine a suite of climate model projections with observations to produce constrained predictions of future temperatures based on the current climate.

1. Introduction

It has been known for decades that the impacts of climate change will be determined not only by the rate and magnitude of global warming, but also by the associated climate change that occurs at the local and regional scales where people and ecosystems exist (e.g. Cohen 1990, Schleussner *et al* 2016, Seneviratne *et al* 2016). Thus, while the time remaining until critical global warming thresholds are reached has been the subject of intense international policy and scientific focus (IPCC 2018, Diffenbaugh and Barnes 2023) the timing of regional warming thresholds is critical for understanding the vulnerability of natural and human systems to climate change in the coming decades. The fact that some areas are warming faster

than typically simulated in climate model projections (Rantanen *et al* 2022, Vautard *et al* 2023, Schumacher *et al* 2024) elevates the need to quantify the time until critical regional warming thresholds are likely to be reached.

The importance of regional climate change for climate impacts has motivated substantial effort to understand the processes that govern the rate and magnitude of regional climate change in different parts of the world (e.g. Cohen 1990, Giorgi and Mearns 1991, Hewitson *et al* 2014, James *et al* 2017, Seneviratne *et al* 2018). Despite this deep body of research, quantifying the time to regional warming thresholds poses persistent challenges. First, even for a given climate forcing pathway, there is uncertainty in the time until specific global warming thresholds

will be reached (e.g. Diffenbaugh and Barnes 2023, Lee *et al* 2023). Second, while it is possible to calculate the regional forced response per unit of global warming (e.g. Lee *et al* 2021 figure 4.32; see also Seneviratne *et al* 2016, Seneviratne and Hauser 2020), uncertainty in both large-scale processes (such as the spatial details of the response of the large-scale atmospheric circulation (e.g. Vautard *et al* 2021) and regional-scale processes (such as feedbacks associated with snow albedo (e.g. Hall and Qu 2006) and soil moisture (e.g. Diffenbaugh and Ashfaq 2010; Vogel *et al* 2017)) contribute to uncertainty in the pattern of the regional forced response. Third, even for a robust regional forced response, internal climate variability creates uncertainty in regional warming that is generally greater than the uncertainty in global-scale warming (e.g. Hawkins and Sutton 2009, Deser *et al* 2012, Diffenbaugh *et al* 2017). As a result, even if the future climate forcing pathway were known, there is substantial uncertainty in the time until a given regional warming threshold will be reached, with the magnitude of that uncertainty varying by region (e.g. Diffenbaugh and Scherer 2011, Hawkins *et al* 2014, Diffenbaugh and Charland 2016, King and Harrington 2018).

Much of this uncertainty is measured within single- or multi-model ensembles of climate simulations. Efforts to narrow that uncertainty through the use of historical observations to determine which climate simulations provide the most accurate future projections have been underway for decades (e.g. Giorgi and Mearns 1991, 2002, Diffenbaugh and Scherer 2013, Xie *et al* 2015, Ribes *et al* 2022, Dai *et al* 2024). One recent approach is to train a neural network on historical and future climate model simulations, and then use historical observations as input to the trained network to make out-of-sample predictions based on the current state of the climate system. This method has recently been applied to make predictions of the time until global warming thresholds are reached based on the observed global map of annual temperature anomalies (Diffenbaugh and Barnes 2023).

In this study, we adapt the approach of Diffenbaugh and Barnes (2023) to predict the time until a wide range of regional warming thresholds are reached (see Methods). We train a separate convolutional neural network (CNN) for each region using the subset of Phase 6 of the Coupled Model Intercomparison Project (CMIP6) climate models that have archived at least 10 realizations in the SSP3-7.0 scenario. Thus, each CNN is trained to iteratively predict the time series of the regional forced temperature response, using only the global map of annual temperature anomalies as input. In this study, in addition to predicting regional rather than global warming thresholds, we add a transfer learning step in which the CMIP-trained CNN is fine-tuned with

limited observations to further improve predictions of the real world. Once the CNNs have been trained, we use the global map of annual temperature anomalies in 2023 to predict the time until the forced temperature response reaches 1.5 °C, 2.0 °C and 3.0 °C above the pre-industrial baseline in each region.

2. Methods

2.1. Datasets

We predict the time remaining until various regional warming thresholds are reached ('time-to-threshold') within each of 43 IPCC regions (Chen *et al* 2021) using CNNs that are trained on annual-mean temperature anomalies simulated by multiple realizations of multiple global climate models (CMIP variable 'tas'). We use climate model data from CMIP6 (Eyring *et al* 2016). Given the availability of different climate models with large numbers of realizations, we focus on the predicted time-to-threshold under the SSP3-7.0 future climate forcing scenario; however, our methodology can be applied to any future climate scenario assuming there is ample data available. To include sampling of internal climate variability, we include the seven CMIP6 climate models that have archived at least 10 realizations, and use only 10 from each climate model (even though some have archived more than 10). Seven of the ten members are used for training the CNNs, two for validation, and one for testing. We compare our results with the threshold exceedance years of an additional 36 CMIP6 climate model simulations under the SSP3-7.0 scenario, where we use a single ensemble member from each climate model (i.e. the widely-used 'one-model-one-vote' approach, also referred to hereafter as 'single-member CMIP6 dataset').

We use observed maps of annual temperature from the Berkeley Earth Surface Temperature ('Berkeley') dataset (Rohde and Hausfather 2020) to make predictions of the time-to-threshold for the real world. For both the climate models and Berkeley, we regrid the annual-mean anomalies to a common 2.5° by 2.5° grid, and calculate the annual-mean anomalies relative to the 1951–1980 climatology at each grid point. Following the IPCC (Lee *et al* 2023), calculations of the time-to-threshold for the regionally-averaged temperature use a preindustrial baseline of 1850–1900. CNN training is performed using data from years 1970–2100.

2.2. Regional warming thresholds

There are many possible definitions of a regional temperature 'threshold', including the time at which the ensemble mean exceeds a temperature value (as was used in Diffenbaugh and Barnes 2023). For the CMIP6 climate models with 10 members used for training, validation and testing of the CNNs, we follow this approach and define the exceedance year

by when the 10-member ensemble-mean exceeds the temperature value. To facilitate transfer learning with the observations and to analyze the single-member CMIP6 dataset (in both cases where we do not have an ensemble mean), we instead define the time-to-threshold for each ensemble member separately. To do so, we apply a backward 5 year rolling mean to the 1970–2100 annual temperatures and then fit a quadratic for each region (for Berkeley we fit from 1970–2023). This procedure is displayed visually in Supplementary figure S4. The year the temperature threshold is reached is defined as the year the quadratic fit exceeds the threshold. These time-to-thresholds serve as the labels for our CNN prediction task.

As described below, a single CNN is trained to make predictions for a range of set temperature thresholds. When creating the climate model training data, temperature thresholds that are not reached prior to 2100 by an ensemble member are not included for that member. Although the CNNs are trained using five specific temperature thresholds (1.0°C , 1.5°C , 2.0°C , 2.5°C , 3.0°C), we use the CNNs to interpolate between these thresholds when making predictions. We focus our main analysis on temperature thresholds of 1.5°C , 2.0°C and 3.0°C above the preindustrial baseline.

2.3. CNNs

We train a CNN to ingest both maps of annual-mean temperature anomalies and the desired temperature threshold to predict the time-to-threshold and its uncertainty for each region. While one could attempt to train a single CNN for all regions simultaneously, here we choose to train a separate CNN for each region, both so that each region is equally weighted during training and so that the network has a focused task which eases our interpretation of the transfer learning and explainable AI (XAI) analyzes. The input to the network is an input map of 72 latitudes by 144 longitude grid points, as well as a single value which represents the temperature threshold for the desired prediction. As depicted in figure 1, the input map is fed into a padding layer (to account for the discontinuity in longitude at 360°E) and the padded map is then fed into three convolution-max pooling layers, after which, it is flattened. At this point, we concatenate the flattened output with the input temperature threshold, and this is then fed into a series of dense layers (10 units each). We separate the predictions of the mu and its uncertainty (σ). We again concatenate the input temperature threshold with the output of the dense layers, feed this through another dense layer (10 units) until finally, the network outputs predictions of the mu and its uncertainty. Following Difffenbaugh and Barnes (2023), these two output units represent the mean and standard deviation of a conditional Gaussian distribution,

and we train the network to minimize the negative log-likelihood. Supplementary figure S1 depicts an example prediction with its corresponding inputs and outputs. Supplementary figure S2 shows the climate model testing error for each region and demonstrates that most time-to-threshold predictions are within 3–4 years of the correct answer. Supplementary figure S3 shows the probability integral transform of the validation and testing predictions for the CMIP6 models for six representative regions and provides evidence that the CNN's estimates of uncertainty are reliable (e.g. Heinrich 2021, Haynes *et al* 2023).

We train the CNNs via back propagation using an Adam optimizer with a learning rate of 0.00005 and a batch size of 32 and apply early stopping on the validation loss using a patience of 10 epochs. ReLu activations are applied throughout the network except in the final dense layer where we use a hyperbolic tangent activation, and in the final output layer which is linear. The model is coded and trained using Tensorflow 2.13.0, and Tensorflow-Probability 0.21.0.

We conduct a number of tests to probe the sensitivity of our results to network choices. The specific networks shown in the paper are those that use a random seed of 66, which sets both the ensemble member train/validation/test split, as well as the initialization of the untrained network. Supplementary figure S7 shows IPCC region West-Central Europe (WCE) predictions for a range of network architectures, learning rates, and additional random seeds (44 and 55). As expected, the predictions vary slightly for different architectures and hyperparameter choices (as some of these choices are not-optimal), but the metrics and results are not overly sensitive, supporting confidence in our results.

2.4. Transfer learning

We further refine the CNNs trained on climate model data (referred to as the 'Base-CNN') through transfer learning on the Berkeley observations to produce new CNNs that generally better agree with the observations (termed 'Transfer-CNNs'). Specifically, we create the Berkeley training set using maps of annual-mean temperature anomalies from 1970–2023 and fine-tune the base-CNN using the three most recently exceeded temperature thresholds out of the following options: 0.8°C , 1.0°C , 1.2°C , 1.5°C , 2.0°C , 2.5°C , and 3.0°C . If only 0.8°C and 1.0°C have been exceeded thus far in the observations, we still perform transfer learning. If only 0.8°C has been exceeded, transfer learning is not performed (e.g., NZ = New Zealand). We test the sensitivity of choosing the three most recent exceedances by repeating our transfer learning training using the fixed thresholds of 0.8°C , 1.0°C , 1.2°C for every region (Supplementary figure S6) and we find that our results are not overly sensitive to this choice. Transfer learning is performed on only the final four dense layers of the mean prediction

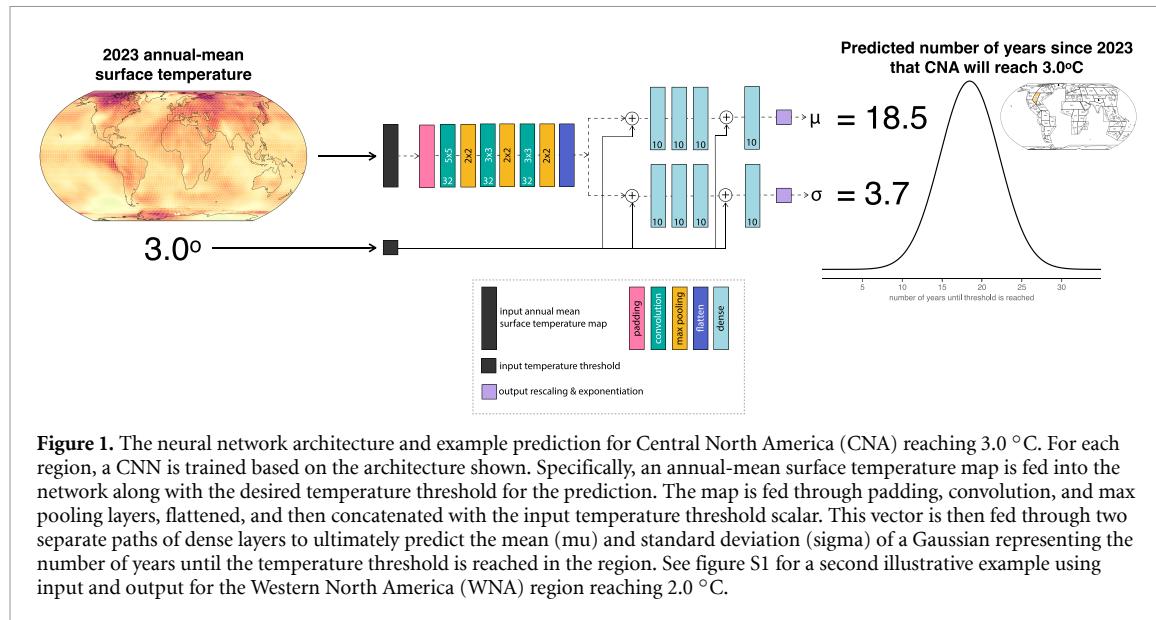


Figure 1. The neural network architecture and example prediction for Central North America (CNA) reaching 3.0°C . For each region, a CNN is trained based on the architecture shown. Specifically, an annual-mean surface temperature map is fed into the network along with the desired temperature threshold for the prediction. The map is fed through padding, convolution, and max pooling layers, flattened, and then concatenated with the input temperature threshold scalar. This vector is then fed through two separate paths of dense layers to ultimately predict the mean (μ) and standard deviation (σ) of a Gaussian representing the number of years until the temperature threshold is reached in the region. See figure S1 for a second illustrative example using input and output for the Western North America (WNA) region reaching 2.0°C .

branch (μ ; figure 1), and so all of the dense layers associated with the uncertainty (σ) and convolutional layers are frozen. Training is performed using a learning rate five times smaller than that of the main network training and uses early stopping based on when the training data no longer reduces its mean absolute error by 0.25°C after 50 epochs.

2.5. Predictions initialized with observations

Once the base-CNNs and transfer-CNNs are trained, we make predictions based on the Berkeley observations by feeding the global map of 2023 Berkeley annual-mean temperature anomalies into the CNNs along with our desired temperature threshold. For our main analysis in figure 2, we use the observed global maps to predict the uncertainty in the time until regional warming of 1.5°C , 2.0°C and 3.0°C . To create the teal and orange timeseries in figure 3, we make predictions for temperature thresholds between 0.5°C and 5.0°C in increments of 0.1° . Note that the network learns to predict negative time-to-thresholds if the threshold has already been exceeded by the year 2023.

2.6. Using XAI to understand the spatial patterns most relevant for predictions

We apply the XAI method ‘deep-SHAP’ (SHapley Additive exPlanations) to quantify the importance of individual grid points in the final network prediction (Lundberg and Lee 2017). We use a ‘baseline of zeros’ for the input maps and set the baseline temperature threshold to be the same as the prediction we are analyzing. In this way, deep-SHAP provides a heatmap of the regions of the input map that made the network modify its prediction from that if the input map had been all zeros (i.e. zero anomalies). Deep-SHAP heatmaps for the observations are computed using the base-CNN’s predictions

over the years 2000–2023 and then averaged together for visualization in figure 4(A). We additionally analyze the transfer-CNNs with deep-SHAP in the same way and show the difference between the two heatmaps in figure 4(B). An important technical consideration is that for this difference to be valid, the base-CNN and the transfer-CNN must produce similar predictions when fed the ‘baseline of zeros’ maps. We have confirmed that this is the case.

3. Results

The pace of regional warming in the SSP3-7.0 scenario varies substantially across the CMIP6 ‘one model, one vote’ ensemble (figure 2; gray bars). For the 1.5°C regional warming threshold, the range in time-to-threshold across the CMIP6 ensemble exceeds two decades for most regions (gray bars in figure 2(A)). Some regions exhibit even larger ranges in time-to-threshold, including Southern South America (NSA; >6 decades) and Northern Europe (NEU; >6 decades). The ranges in time-to-threshold across the CMIP6 ensemble are similarly large for the 2.0°C (gray bars in figure 2(B)) and 3.0°C (gray bars in figure 2(C)) regional warming thresholds.

Because our training climate models are a subset of the CMIP6 ensemble, our training ensemble (teal triangles in figure 2) generally exhibits a smaller range in time-to-threshold than the full ‘one-model, one-vote’ CMIP6 ensemble. However, because our training ensemble consists of 10 realizations of each training climate model, in some cases the mean value of an individual training climate model falls outside of the values in the CMIP6 ‘one model, one vote’ ensemble (e.g. NES = Northeastern South America, for the 1.5°C and 2.0°C thresholds; figure 2(A)). The ranges

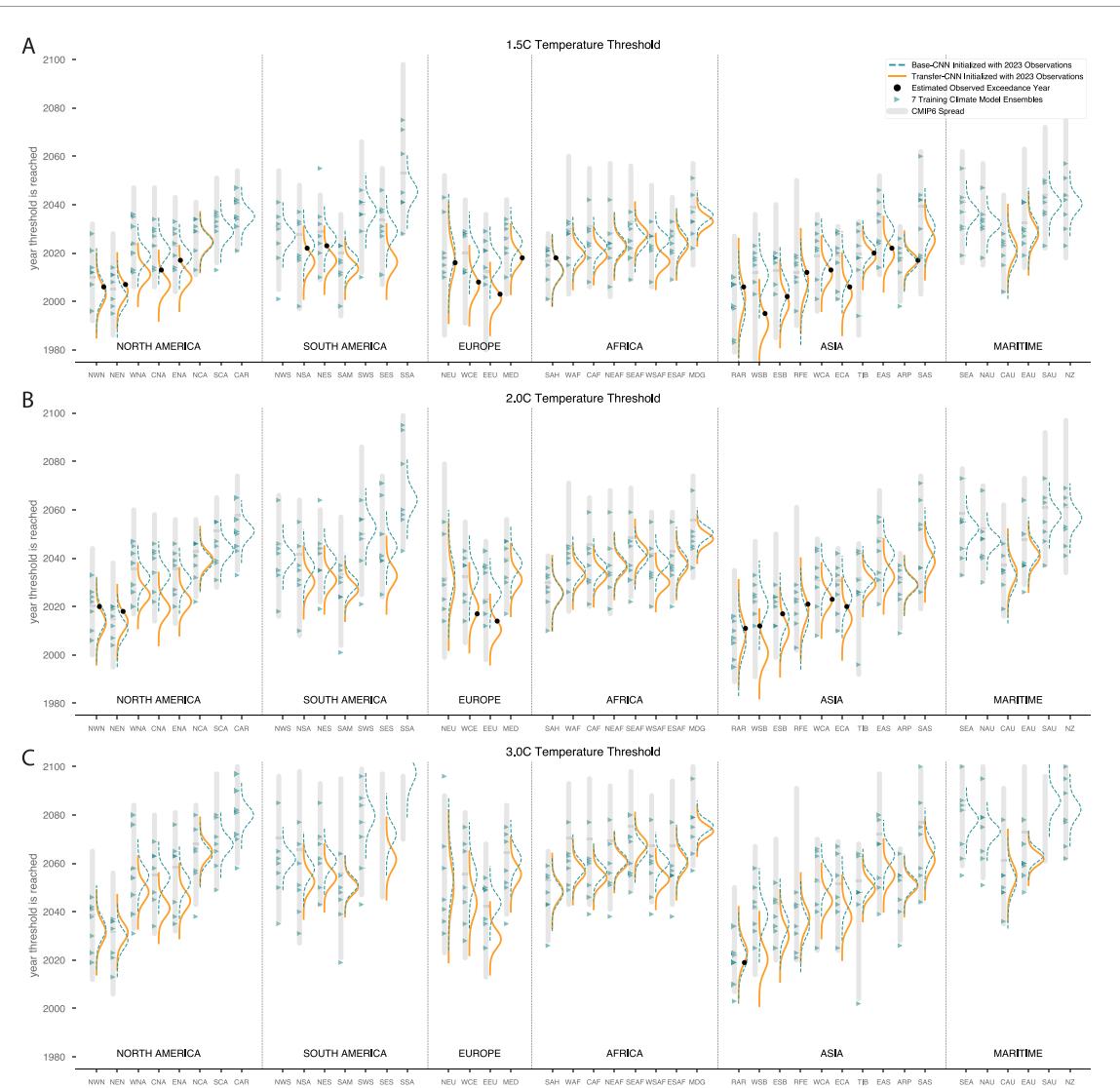


Figure 2. Predictions of the exceedance year 1.5°C , 2.0°C and 3.0°C thresholds for each of the IPCC regions. Gray bars denote the full spread across the CMIP6 models and the teal triangles denote the years for each of the CMIP6 GCMs used for training the network. Dashed teal Gaussians denote the base-CNN predictions initialized with 2023 Observations while the orange Gaussians denote the transfer-CNN predictions. Black dots denote the estimated observed exceedance year if it has already occurred.

in regional time-to-threshold predicted by the CNNs (PDF curves in figure 2) are substantially smaller than the ranges of both the CMIP6 ensemble and our training ensemble. This reduction in uncertainty reflects the refinement of the climate model uncertainty as the network learns the best indicators of the regional temperature response, and also reflects the recent state of those indicators in the temperature observations provided as input for the network prediction.

Our primary prediction (transfer-CNN; orange PDF curves in figure 2) uses a transfer-learned CNN initialized with the observed global map of annual temperature anomalies in 2023 (see Methods). For the 1.5°C threshold, that primary prediction yields a central estimate of 2040 or earlier—and a very high likelihood of 2060 or earlier—for all regions where transfer learning is possible (figure 2(A)). Our primary prediction also yields a central estimate of

2060 or earlier for the 2.0°C threshold for all regions where the transfer learning is possible, and 2040 or earlier for 31 out of 34 regions where the transfer learning is possible (figure 2(B)). Further, our primary prediction yields a central estimate of 2070 or earlier for the 3.0°C threshold for all but one region where the transfer learning is possible, 2060 or earlier for 26 out of 34 regions where transfer learning is possible, and 2040 or earlier for 8 out of 34 regions where transfer learning is possible (figure 2(C)).

We can use the regions that have already reached the predicted temperature threshold (black circles in figure 2) to evaluate the accuracy of the CNN predictions. While the observed time-to-threshold falls within the CMIP6 range for all applicable regions, the CMIP6 ranges are so large that in most cases the CNN-based predictions represent a substantial

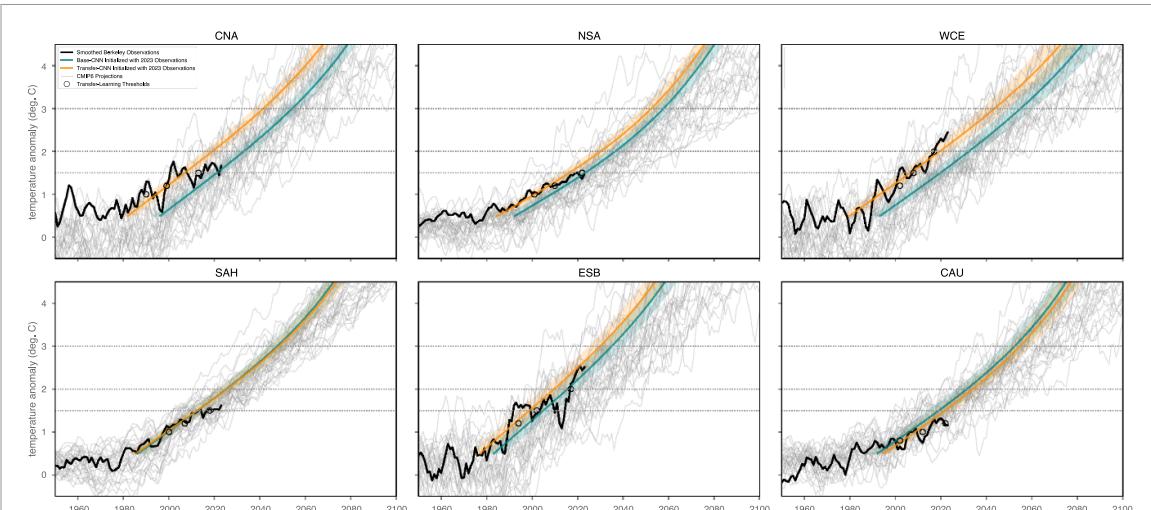


Figure 3. Predicted exceedance years and uncertainties output by the regional networks initialized with 2023 observations. The teal curves denote predictions with the base-CNNs while the orange curves denote predictions with the transfer-CNNs. Shading denotes \pm one sigma as predicted by the CNN. The black lines denote the observed temperature time series, and the open circles denote the years and temperatures used for transfer learning. Light gray lines show the range of regional temperature anomalies across the CMIP6 models.

refinement (both with and without transfer learning). The fact that the observed predictions are a substantial refinement to the CMIP6 range is particularly notable for the results without transfer learning since CMIP6 biases could have caused the CNN to produce wildly inconsistent predictions based on observed inputs. However, this was not the case. In addition, for the majority of regions where the transfer learning substantially shifts the distribution of predicted time-to-threshold, the distribution is earlier for the CNN with transfer learning (orange PDF curves in figure 2) than without (teal PDF curves in figure 2). For a number of the regions that have already reached the predicted warming threshold, the agreement between the predicted and observed time-to-threshold is improved for the CNN with transfer learning. This is seen most clearly in Europe, where the observed time-to-threshold has substantially higher likelihood in the distribution predicted by the transfer-CNN for 5 of the 6 instances where the regional temperature threshold has already been reached (figures 2(A) and (B)).

The magnitude of difference in predicted time-to-threshold between the CNNs with and without transfer learning varies by region (figure 2). This can be clearly seen in the predicted trajectory of regional warming over the 21st century in the SSP3-7.0 scenario (figure 3). For example, in some regions (such as SAH = Sahara and CAU = Central Australia) the predicted trajectory is quite similar, with the uncertainty in the regional temperature prediction overlapping throughout the 21st century. In contrast, in other regions (such as CNA = Central North America and WCE = West-Central Europe) the predicted trajectory is shifted substantially earlier after transfer learning.

The shape of the predicted warming trajectory over the 21st century also varies across regions (figure 3). The difference in the shape of predicted warming confirms that the different CNNs are learning different relationships between the global map of temperature anomalies and the temperature response of each region. This heterogeneity is reflected in the different areas of the globe that the network has learned are the best sources of information for each region (figure 4(A)), as well as in the changes in those areas after transfer learning (figure 4(B)). Interestingly, for some regions (such as NSA = Northern South America, SAH = Sahara, and CAU = Central Australia) the temperature within the region is a strong indicator of the time-to-threshold, but for others (such as CNA = Central North America, WCE = West-Central Europe and ESB = Eastern Siberia) the temperature within the region is not as strong an indicator of the time-to-threshold as the temperature in other regions (figure 4(A)). Likewise, the change with transfer learning is not typically weighted towards the temperature within the region, with the largest changes occurring elsewhere in the world (figure 4(B)).

There are a number of areas of the globe where the annual temperature seems to be a consistent indicator of the regional temperature response, even for geographically disparate regions (figure 4(A)). For example, despite a range of warming rates across the illustrative regions (figure 3), annual temperatures in areas of the Southern Ocean (including south of the Horn of Africa), the high elevations of central Asia, the Arctic coast of North America, and northwestern Africa all seem to be indicators of the regional temperature response (figure 4(A)). While the areas of the globe that change the most after transfer learning

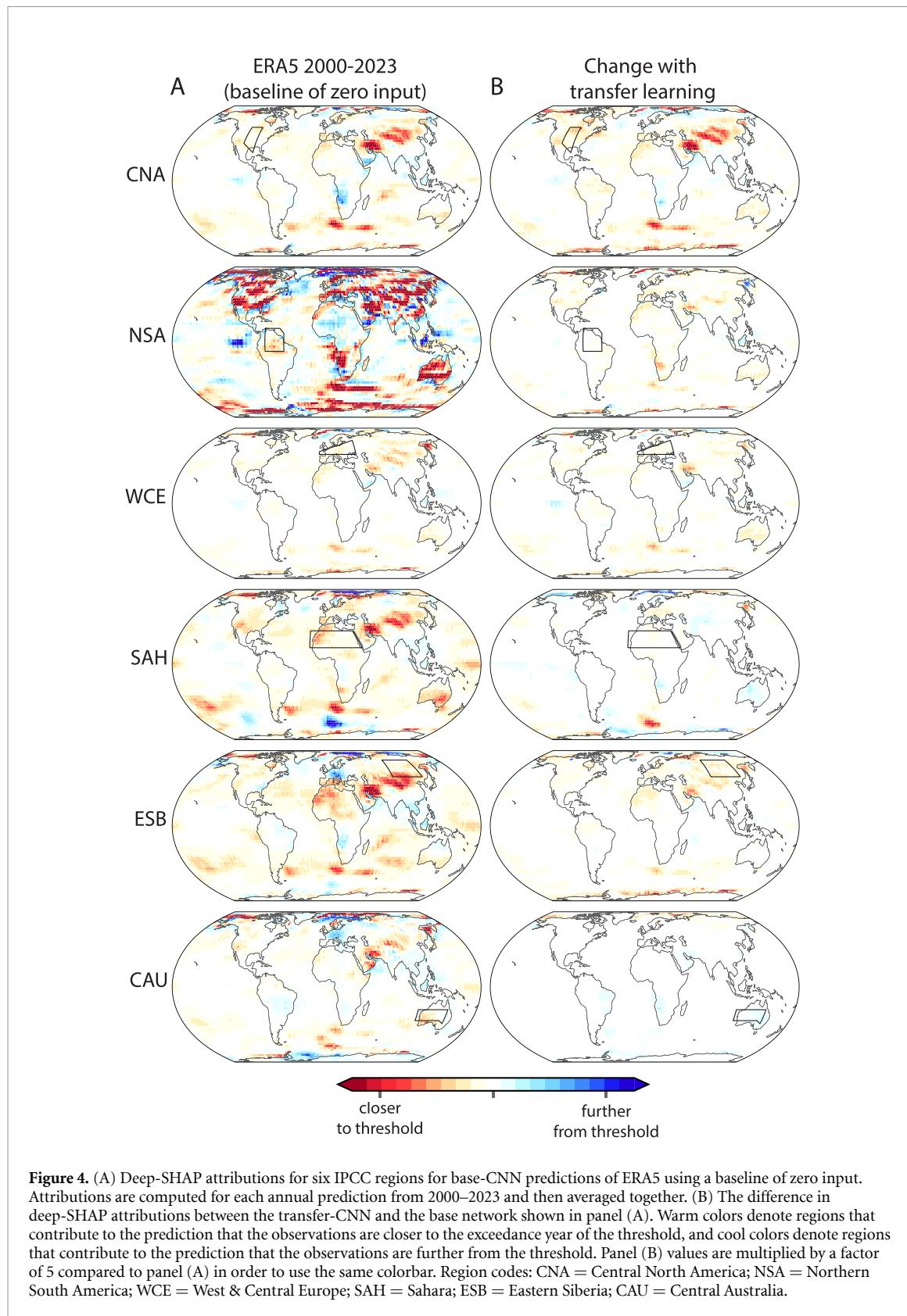


Figure 4. (A) Deep-SHAP attributions for six IPCC regions for base-CNN predictions of ERA5 using a baseline of zero input. Attributions are computed for each annual prediction from 2000–2023 and then averaged together. (B) The difference in deep-SHAP attributions between the transfer-CNN and the base network shown in panel (A). Warm colors denote regions that contribute to the prediction that the observations are closer to the exceedance year of the threshold, and cool colors denote regions that contribute to the prediction that the observations are further from the threshold. Panel (B) values are multiplied by a factor of 5 compared to panel (A) in order to use the same colorbar. Region codes: CNA = Central North America; NSA = Northern South America; WCE = West & Central Europe; SAH = Sahara; ESB = Eastern Siberia; CAU = Central Australia.

are not uniform across the illustrative regions, these same areas (such as the Southern Ocean south of the Horn of Africa and the high elevations of central Asia) do tend to increase in importance after transfer learning.

4. Discussion and conclusions

We use observations of annual global temperature anomalies as input for neural networks trained on GCMs from the CMIP6 ensemble in order to refine

multi-model ensemble projections of regional warming. We find that the addition of transfer learning leads to predictions of the regional time-to-threshold that are generally earlier than the corresponding predictions without transfer learning, often improving the agreement with the observed time until the regional warming threshold (figure 2). The shift towards earlier regional warming seems to be particularly consistent in regions of strong soil moisture limitation on evapotranspiration and consequent strong soil moisture–temperature feedback (Seneviratne *et al* 2010, 2013), such as CNA; Koster *et al* (2004)), WCE; (Seneviratne *et al* 2006), the Mediterranean; Seneviratne *et al* 2006, Diffenbaugh *et al* (2007), South Asia; (Koster *et al* 2004), East Asia; (Koster *et al* 2004), NSA; (Koster *et al* 2004) and Central Africa; (Koster *et al* 2004). This highlights that potential issues with the representation of regional soil moisture–temperature feedbacks in Earth System Models could explain some of the discrepancies identified with our analysis.

Our results highlight the power of transfer learning as a tool to combine a suite of climate model projections with observations to produce constrained predictions of future temperatures based on the current climate (e.g. Ham *et al* 2019; see also Immorlano *et al* (2023) for a different but related transfer learning approach to reduce uncertainty in climate projections). Deep learning requires a large amount of data and this approach allows for the majority of learning to be performed on climate model output. However, given that climate models have systematic biases (e.g. Bock *et al* 2020), transfer learning refines the network weights based on a limited amount of observations with the aim to improve predictions of the real world. The fact that transfer learning both improves the network predictions for temperature thresholds that have already been exceeded in observations, and also shifts the exceedance years earlier for those that have not, suggests that many regions may reach 2.0 °C and 3.0 °C earlier than the majority of CMIP6 climate models suggest.

Although the areas of the globe that the deep-SHAP XAI analysis identifies as being important for predicting the regional time-to-threshold are not identical between the different illustrative regions, there are a number of common features (figure 4). The fact that there are so many common features for regions that are geographically dispersed and climatically diverse (such as, for example, CNA, WCE, the SAH, ESB, and CAU), and that the strongest indicators are generally outside a given region, suggests that these common features are indicators of the global time-to-threshold. As the global warming level is a leading indicator of regional warming, areas of the globe that are leading indicators of the time until the global time-to-threshold are also likely to be strong indicators of the regional time-to-threshold.

The prediction of the regional time-to-threshold does not seem to be heavily influenced by disparities in the climate sensitivity of the training climate models. In particular, one concern about training on the available CMIP6 climate models is that some have been identified as ‘hot models’ due to their high global climate sensitivity (Hausfather *et al* 2022, Diffenbaugh and Barnes 2023). To test whether training on these climate models contribute to the network’s predictions generally falling in the earlier part of the CMIP6 distribution, we repeat our analysis after removing the high-sensitivity climate models (as identified by Meehl *et al* 2020) from the training set. We find that the predicted regional time-to-thresholds are strikingly similar even after the so-called ‘hot models’ are removed (figure S5). This robustness suggests that the indicators of time-to-threshold learned by the network are not highly dependent on the climate models’ global climate sensitivity.

Uncertainty in the timing of regional warming has been an area of persistent scientific and policy concern. We find that machine learning with transfer learning offers the potential to substantially refine predictions of the time until specific regional warming thresholds are reached. For many regions, this refinement yields predictions that are weighted towards the early part of the CMIP6 distribution, with transfer learning generally shifting the prediction earlier than without transfer learning. These results suggest a very high likelihood of 2.0 °C of regional warming by 2040 for the majority of regions, along with a likelihood of 3 °C by mid-century. Given the impacts on extreme events, water resources, food security, human health, livelihoods and ecosystems that are already emerging at the current levels of regional warming (Figure SPM.2 in IPCC 2022), the likelihood that these higher levels of regional warming are reached in the next 2–3 decades poses substantial risks. Our data-driven predictions using observed maps of annual global temperature as input into neural networks to constrain multi-model GCM uncertainty suggest that these risks are likely to emerge earlier than suggested by the full climate model ensemble over many regions of the world.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://aims2.llnl.gov/search/cmip6/>.

Acknowledgment

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output,

the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. EAB was supported, in part, by the Regional and Global Model Analysis program area of the US Department of Energy's Office of Biological and Environmental Research as part of the Program for Climate Model Diagnosis and Intercomparison project. NSD acknowledges support from Stanford University, and computational support from the Stanford Doerr School Center for Computation and the Stanford Research Computing Center. SIS acknowledges partial support from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 101003469 (XAIDA project), from the European Union's Horizon Europe research and innovation actions under Grant Agreement No. 101137682 (AI4PEX project), from the Swiss State Secretariat for Education and Innovation (SERI), and from the Stanford Woods Institute for the Environment (sabbatical stay from April-July 2024).

Data, materials, and software availability

All code for processing the data and performing the analysis is available on GitHub at https://github.com/eabarnes1010/paperRegional_thresholds_erl2024 and is archived on Zenodo with permanent DOI <https://doi.org/10.5281/zenodo.14045995> (Barnes 2024).

ORCID iDs

Elizabeth A Barnes  <https://orcid.org/0000-0003-4284-9320>

Noah S Difffenbaugh  <https://orcid.org/0000-0002-8856-4964>

References

- Bock L, Lauer A, Schlund M, Barreiro M, Bellouin N, Jones C, Meehl G A, Predoi V, Roberts M J and Eyring V 2020 Quantifying progress across different CMIP phases with the ESMValTool *J. Geophys. Res.* **125** e2019JD032321
- Chen D *et al* 2021 Framing, context and methods *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed V Masson-Delmotte *et al* (Cambridge University Press) pp 147–286
- Cohen S J 1990 Bringing the global warming issue closer to home: the challenge of regional impact studies *Bull. Am. Meteor. Soc.* **71** 520–26
- Dai P, Nie J, Yu Y and Wu R 2024 Constraints on regional projections of mean and extreme precipitation under warming *Proc. Natl Acad. Sci. USA* **121** e2312400121
- Deser C, Knutti R, Solomon S and Phillips A S 2012 Communication of the role of natural variability in future North American climate *Nat. Clim. Change* **2** 775
- Difffenbaugh N S *et al* 2017 Quantifying the influence of global warming on unprecedented extreme climate events *Proc. Natl Acad. Sci. USA* **114** 4881–6
- Difffenbaugh N S and Ashfaq M 2010 Intensification of hot extremes in the United States *Geophys. Res. Lett.* **37** 15
- Difffenbaugh N S and Barnes E A 2023 Data-driven predictions of the time remaining until critical global warming thresholds are reached *Proc. Natl Acad. Sci.* **120** e2207183120
- Difffenbaugh N S and Charland A 2016 Probability of emergence of novel temperature regimes at different levels of cumulative carbon emissions *Front. Ecol. Environ.* **14** 418–23
- Difffenbaugh N S, Pal J S, Giorgi F and Gao X 2007 Heat stress intensification in the Mediterranean climate change hotspot *Geophys. Res. Lett.* **34** L11706
- Difffenbaugh N S and Scherer M 2011 Observational and model evidence of global emergence of permanent, unprecedented heat in the 20(th) and 21(st) centuries *Clim. Change* **107** 615–24
- Difffenbaugh N S and Scherer M 2013 Using climate impacts indicators to evaluate climate model ensembles: temperature suitability of premium winegrape cultivation in the United States *Clim. Dyn.* **40** 709–29
- Elizabeth Barnes 2024 eabarnes1010/paperRegional_thresholds_erl2024: Publication (1.0) (Zenodo) (<https://doi.org/10.5281/zenodo.14045995>)
- Eyring V, Bony S, Meehl G A, Senior C A, Stevens B, Stouffer R J and Taylor K E 2016 Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization *Geosci. Model Dev.* **9** 1937–58
- Giorgi F and Mearns L 1991 Approaches to the simulation of regional climate change: a review *Rev. Geophys.* **29** 191–216
- Giorgi F and Mearns L 2002 Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the ‘reliability ensemble averaging’ (REA) method *J. Clim.* **15** 1141–58
- Hall A and Qu X 2006 Using the current seasonal cycle to constrain snow albedo feedback in future climate change *Geophys. Res. Lett.* **33** L03502
- Ham Y-G, Kim J-H and Luo J-J 2019 Deep learning for multi-year ENSO forecasts *Nature* **573** 568–72
- Hausfather Z, Marvel K, Schmidt G A, Nielsen-Gammon J W and Zelinka M 2022 Climate simulations: recognize the ‘hot model’ problem *Nature* **605** 26–29
- Hawkins E *et al* 2014 Uncertainties in the timing of unprecedented climates *Nature* **511** E3–5
- Hawkins E and Sutton R 2009 The potential to narrow uncertainty in regional climate predictions *Bull. Am. Meteorol. Soc.* **90** 1095–107
- Haynes K, Lagerquist R, McGraw M, Musgrave K and Ebert-Uphoff I 2023 Creating and evaluating uncertainty estimates with neural networks for environmental-science applications *Artif. Int. Earth Syst.* **2** 1–58
- Heinrich C 2021 On the number of bins in a rank histogram *Q. J. R. Meteorol. Soc.* **147** 544–56
- Hewitson B C, Janetos A C, Carter T R, Giorgi F, Jones R G, Kwon W T, Mearns L O, Schipper E L F and van Aalst M 2014 Regional context *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed V R Barros, C B Field, D J Dokken, M D Mastrandrea, K J Mach, T E Bilir and M Chatterjee (Cambridge University Press) pp 1133–97
- Immorlano F, Eyring V, le Monnier de Gouville T, Accarino G, Elia D, Aloisio G and Gentile P 2023 Transferring climate change knowledge (arXiv:2309.14780)
- IPCC 2022 Summary for policymakers *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed H O Pörtner, D C Roberts, M Tignor, E S Poloczanska, K Mintenbeck, A Alegría and M Craig (Cambridge University Press)
- IPCC 2018 Summary for policymakers. In: global warming of 1.5 °C *An IPCC Special Report on the Impacts of Global Warming of 1.5 °C above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate*

- Change, Sustainable Development, and Efforts to Eradicate Poverty* ed V Masson-Delmotte *et al* (Cambridge University Press) pp 3–24
- James R, Washington R, Schleussner C F, Rogelj J and Conway D 2017 Characterizing half-a-degree difference: a review of methods for identifying regional climate responses to global warming targets *Wiley Interdiscip. Rev. Clim. Change* **8** e457
- King A D and Harrington L J 2018 The inequality of climate change from 1.5 to 2 °C of global warming *Geophys. Res. Lett.* **45** 5030–3
- Koster R D *et al* 2004 Regions of strong coupling between soil moisture and precipitation *Science* **305** 1138–40
- Lee J-Y, Marotzke J, Bala G, Cao L, Corti S, Dunne J P and Engelbrecht F 2023 future global climate: scenario-based projections and near-term information *Climate Change 2021—The Physical Science Basis* ed V Masson-Delmotte, P Zhai, A Pirani, S L Connors, C Péan, S Berger and N Caud (Cambridge University Press) pp 553–672
- Lee J-Y *et al* 2021 *Future Global Climate: Scenario-Based Projections and Near-Term Information. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed V P Zhai *et al* (Cambridge University Press) pp 553–672
- Lundberg S M and Lee S-I 2017 A unified approach to interpreting model predictions *Advances in Neural Information Processing Systems* 30 ed I Guyon, U V Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan and R Garnett pp 4765–74 (Curran Associates, Inc) (available at: <http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>)
- Meehl G A, Senior C A, Eyring V, Flato G, Lamarque J F, Stouffer R J, Taylor K E and Schlund M 2020 Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models *Sci. Adv.* **6** eaab1981
- Rantanen M, Yu A, Karpecho A L, Nordling K, Hyvärinen O, Ruosteenoja K, Vihma T and Laaksonen A 2022 The Arctic has warmed nearly four times faster than the globe since 1979 *Commun. Earth Environ.* **3** 168
- Ribes A, Boé J, Qasmi S, Dubuisson B, Douville H and Terray L 2022 An updated assessment of past and future warming over France based on a regional observational constraint *Earth Syst. Dyn.* **13** 1397–415
- Rohde R A and Hausfather Z 2020 The Berkeley Earth land/ocean temperature record *Earth Syst. Sci. Data* **12** 3469–79
- Schleussner C-F *et al* 2016 Differential climate impacts for policy-relevant limits to global warming: the case of 1.5 °C and 2 °C *Earth Syst. Dyn.* **7** 327–51
- Schumacher D, Singh J, Hauser M, Fischer E and Seneviratne S I 2024 Why climate models underestimate the exacerbated warming in Western Europe (available at: www.researchsquare.com/article/rs-3314992/v1)
- Seneviratne S I *et al* 2013 Impact of soil moisture-climate feedbacks on CMIP5 projections: first results from the GLACE-CMIP5 experiment *Geophys. Res. Lett.* **40** 5212–7
- Seneviratne S I *et al* 2018 The many possible climates from the Paris Agreement's aim of 1.5 °C warming *Nature* **558** 41–49
- Seneviratne S I, Corti T, Davin E L, Hirschi M, Jaeger E B, Lehner I, Orlowsky B and Teuling A J 2010 Investigating soil moisture-climate interactions in a changing climate: a review *Earth Sci. Rev.* **99** 125–61
- Seneviratne S I, Donat M G, Pitman A J, Knutti R and Wilby R L 2016 Allowable CO₂ emissions based on regional and impact-related climate targets *Nature* **529** 477–83
- Seneviratne S I and Hauser M 2020 Regional climate sensitivity of climate extremes in CMIP6 vs CMIP5 multi-model ensembles *Earth's Future* **8** e2019EF001474
- Seneviratne S I, Lüthi D, Litschi M and Schär C 2006 Land-atmosphere coupling and climate change in Europe *Nature* **443** 205–9
- Vautard R *et al* 2023 Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends *Nat. Commun.* **14** 6803
- Vautard R, Kadygrov N, Iles C, Boberg F, Buonomo E, Bülow K and Coppola E 2021 Evaluation of the large EURO-CORDEX regional climate model ensemble *J. Geophys. Res.* **126** e2019JD032344
- Vogel M M, Orth R, Cheruy F, Hagemann S, Lorenz R, van den Hurk B J J M and Seneviratne S I 2017 Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks: soil moisture and hot regional extremes *Geophys. Res. Lett.* **44** 1511–9
- Xie S-P *et al* 2015 Towards predictive understanding of regional climate change *Nat. Clim. Change* **5** 921–30