Optimizing Seasonal-to-Decadal Analog Forecasts with a Learned Spatially-Weighted Mask

Jamin K. Rader[1], and Elizabeth A. Barnes[1]

[1]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA.

Key Points:

- An interpretable neural network provides a spatially-weighted mask for selecting optimal analogs
- Analogs selected with the weighted mask offer more skillful forecasts than traditional methods for selecting analogs
- The learned mask highlights precursor regions for predicting large-scale climate anomalies in a perfect model framework

Corresponding author: Jamin K. Rader, `jamin.rader@colostate.edu`

Abstract

Seasonal-to-decadal climate prediction is crucial for decision-making in a number of industries, but forecasts on these timescales have limited skill. Here, we develop a data-driven method for selecting optimal analogs for seasonal-to-decadal analog forecasting. Using an interpretable neural network, we learn a spatially-weighted mask that quantifies how important each grid point is for determining whether two climate states will evolve similarly. We show that analogs selected using this weighted mask provide more skillful forecasts than analogs that are selected using traditional spatially-uniform methods. This method is tested on two prediction problems within a perfect model framework using the Max Planck Institute for Meteorology Grand Ensemble: multi-year prediction of North Atlantic sea surface temperatures, and seasonal prediction of El Niño Southern Oscillation. This work demonstrates a methodical approach to selecting analogs that may be useful for improving seasonal-to-decadal forecasts and understanding their sources of skill.

Plain Language Summary

Understanding how the climate will look in one to ten years is useful for many industries, but this task is very difficult. One method for making forecasts on these timescales is called analog forecasting. In analog forecasting, a researcher finds past states in observations, or states in a climate model simulation, that look like the current state of the climate, and uses how those maps changed over time to predict how the climate will change over time. Some regions are more important for determining how a climate state will change over time, and we use a machine learning method called a neural network to identify these important regions. We find that if we only look at these important regions when determining if two climate states are similar or not, we can improve our analog forecasting skill.

1 Background

Forecasts on seasonal-to-decadal timescales are crucial for decision-makers in a number of industries, but forecasts on these timescales have limited skill (Kushnir et al., 2019; Merryfield et al., 2020; Towler et al., 2018). Analog forecasting, predicting what will happen based on previous states with similar initial conditions, is an intuitive method for seasonal-to-decadal prediction. It is built on the premise that similar geophysical states will evolve in similar ways (Lorenz, 1969). It follows that analogs—similar looking states to the initial state that is being forecast—can provide insight into how that initial state will continue to evolve. The analog forecasting approach is powerful for seasonal-to-decadal climate prediction (e.g., Ding et al., 2018, 2019; Menary et al., 2021; Delle Monache et al., 2013; Zhang et al., 2023) and can outperform general circulation models (GCMs) initialized with observations, which struggle with initialization shock and climate model drift (Merryfield et al., 2020; Mulholland et al., 2015).

A major hurdle in obtaining successful analog forecasts is that the climate system is noisy and chaotic, and thus small differences between two initial states can result in vast differences in their evolution (Lorenz, 1963). Thus, a successful analog forecast for a particular initial climate state, which we refer to as the state of interest (SOI), requires that the analogs and SOI are sufficiently similar such that their evolutions do not significantly diverge during the prediction timeframe. Sufficiently similar analogs can be difficult to find in the observational record since the number of independent observations we have on seasonal-to-decadal scales (e.g., fewer than 100 during the satellite era) is so much smaller than the number of degrees of freedom within a global geophysical field (e.g., Van den Dool, 1994). While observations are in short supply, there is a wealth of simulated climate data and many recent studies have used "model-analogs" (Ding et al., 2018) drawn from climate model output instead (e.g., Lou et al., 2023; Peng et al., 2021; Wu & Yan, 2023).

We refer to the library of climate model states that can be used for analog forecasting as "potential analogs." Once a potential analog has been identified to be sufficiently similar to the SOI we refer to it as an analog. Forecasts are made by taking the mean evolution of the top-N analogs, where N is chosen by the user. There are several ways to quantify the similarity between the potential analogs and the SOI. The most straightforward method is to compute the global correlation between each potential analog and the SOI (e.g., Mahmood et al., 2022). Using a global correlation assumes that the similarity between the maps at each grid point globally matters equally. A natural next step in complexity is to compute a correlation over a region that is known to be important for predictability of a given target, such as the North Pacific for predicting the Pacific Decadal Oscillation (e.g., Wu & Yan, 2023). While this approach removes some regions that may not be useful for determining the best analogs, it still assumes that each grid point within the region is equally important and the region must be known *a priori*.

In the following work, we train an interpretable neural network on a proxy task that is similar to the analog problem (Section 3). The network learns a weighted mask which is used for determining analogs. The forecasting skill of the analogs selected using the learned weighted mask is tested through a perfect model approach where climate model data substitutes observations and is used to predict future climate model data. We show in two examples, forecasting 5-year sea surface temperature (SST) anomalies in the North Atlantic (Section 4) and wintertime SST anomalies in the tropical Pacific (i.e. El Niño Southern Oscillation; Section 5), that analogs identified using the weighted mask provide more skillful forecasts than analogs that are identified in a way that is globally or regionally uniform. In addition, we show that these masks, once generated by a neural network, can be modified *post hoc* to further investigate the importance of each region for seasonal-to-decadal prediction (Section 5).

2 Data and Metrics

2.1 Climate Model Data

We use monthly SST from the historical run of the Max Planck Institute (MPI) for Meteorology Grand Ensemble (GE; Maher et al., 2019) at $2°$ latitude by $2°$ longitude resolution. This dataset contains 100 members and each simulates 156 years (1850-2005) of the Earth's climate with historical forcing. The MPI-GE uses the MPI Earth System Model version 1.1 (ESM1.1; Giorgetta et al., 2013). Each member is initialized using a different year of the preindustrial control simulation such that the differences between ensemble members are a product of internal variability.

2.2 Standardization and Selection

Subsets of the MPI-GE ensemble members are used for different purposes. Our library of potential analogs is made up of members 1-35. Members 36-50 are the SOIs for training the neural network, members 51-55 are the SOIs for the early stopping validation set (which is used to prevent overfitting to the training data), and members 56-60 are the SOIs for the tuning validation set (which is used to identify optimal hyperparameters for the neural network). Finally, members 96-100, which are withheld until the very end, are the test set for making and evaluating the analog forecasts. Details on the process of tuning and training the neural network, including selecting the hyperparameters, can be found in Section S1.

Each sample i or j, from the SOIs or the library of potential analogs, is composed of an input field ($I_{SOI,i}$ or $I_{analog,j}$) and a target ($T_{SOI,i}$ or $T_{analog,j}$). The input fields are one or more maps of global SST leading the targets over some earlier period (the "input period"). The targets are time- and area-mean SST anomalies over a certain region and forecast window.

We removed the forced signal from the climate model data by subtracting the ensemble mean of the library of potential analogs at each location and year from each set of data. After the forced signal was removed, the data was standardized by dividing by the standard deviation at each grid point across the library of potential analogs. By using the library of potential analogs to calculate the forced signal and internal variance we treat the SOIs as if they are truly unseen data as we would when forecasting.

2.3 Metrics

We measure forecasting skill with a mean absolute error (MAE) skill score. This skill score is calculated by comparing the MAE of the analog prediction for the SOIs in the test set with the MAE of climatology, as:

$$\text{Skill Score} = 1 - \frac{MAE_{pred}}{MAE_{climo}}$$

such that a perfect prediction has a score of one, and a climatology prediction has a score of zero. Climatology is the prediction by the mean state, which is zero for this standardized data. Analog forecasts made using the weighted mask are compared with the following additional baselines: a global analog forecast, a target region analog forecast, a mean target evolution forecast, and a random forecast. In the global analog forecast (target region analog forecast), the analogs are selected if the unweighted MSE over the entire globe (target region) is the smallest. The mean target evolution forecast is based on how the targets in the input period evolve on average and is detailed in Section S2. The random forecast is made by randomly selecting targets from the library of potential analogs and using them as the prediction.

3 Optimized Analog Forecasting Approach

Our goal is to find optimal analogs for forecasting a specific target. To do this, we train a neural network to identify a spatially-weighted mask. This weighted mask is then multiplied by the SOI and potential analogs and the mean-squared error (MSE) between the weighted maps is used to determine how similar they are (Figure 1). This weighted mask should contain large values where similarity between the analogs and the SOI is most important for predicting the target and near-zero values where similarity between the maps is not important. With this architecture, the MSE will be low if the maps agree where the mask weights are high, regardless of the differences between the maps where the mask weights are low. For the plots in this paper, the mask is normalized by dividing by the sum of the weights times the size of the input, such that the mean weight is one.

We generate the weighted mask by training a neural network on a proxy task that is tangential to our main goal. While our goal is to identify a weighted mask that is optimized for making an analog forecast, our proxy task is to predict the difference in $T_{SOI,i}$ and $T_{analog,j}$ given $I_{SOI,i}$ and $I_{analog,j}$. En route to making this prediction, the neural network must learn the weighted mask, multiply it by the two input maps, compute the MSE between these weighted maps, and finally convert the MSE into a predicted difference in the targets. This process is depicted in the red box of Figure 1.

Once the weighted mask has been learned, a neural network is no longer needed to make analog predictions. The weighted mask is multiplied by the SOI and each potential analog, the MSE is computed between the weighted SOI and the weighted potential analogs, and the potential analogs with the lowest MSE are used to make the analog forecast. While the proxy task is not identical to the analog problem, it provides a weighted mask that improves analog forecasting skill, as we will show in Sections 4 and 5.

4 Multi-year Prediction of North Atlantic Sea Surface Temperature

We first test our analog forecasting approach on a multi-year prediction of SSTs over the North Atlantic. North Atlantic SSTs exhibit clear variability on multi-annual timescales (Jackson et al., 2022) and exhibit potential for skillful decadal forecasts (Hawkins et al., 2011; Sutton & Allen, 1997). SST variability in the North Atlantic has been as-

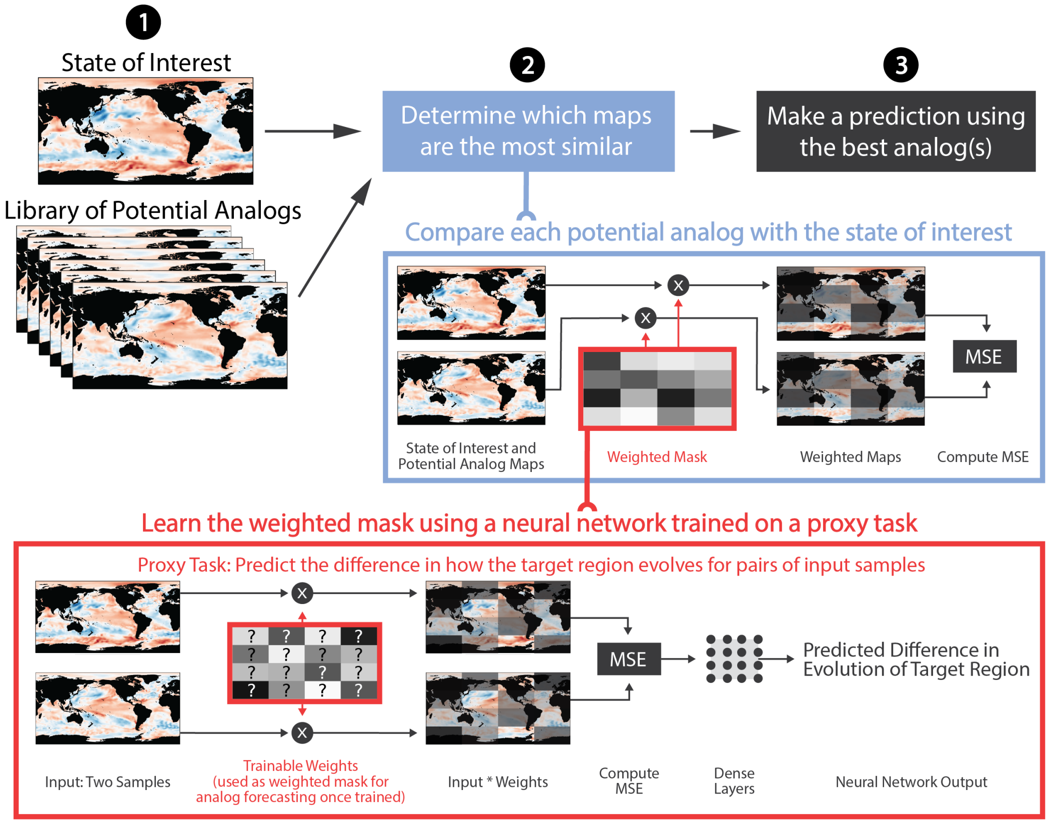

Figure 1. Optimized analog forecasting method and interpretable neural network architecture. The analog forecasting method can be described in three steps: 1) identify a state of interest and a library of potential analogs. 2) Determine which maps are the most similar. 3) Make a prediction using the best analog(s). In the blue box, we show our weighted-mask approach for determining the similarity of two maps. The weighted mask is multiplied by the state of interest and a potential analog before computing the mean squared error (MSE). In the red box, the interpretable neural network architecture is shown. Two input samples are multiplied by a matrix of trainable weights and the MSE is computed. This MSE is then converted to a predicted difference in the sample targets using a group of fully-connected dense layers. Note that the weighted mask has the same dimensions as the input field(s), despite the coarser resolution in this figure.

sociated with weather and climate anomalies globally, including Atlantic hurricane frequency and intensity (Goldenberg et al., 2001; Balaguru et al., 2018), northern hemisphere precipitation (Enfield et al., 2001; Si et al., 2023), and the strength of the Asian summer monsoon (Shekhar et al., 2022). In this prediction problem, we use global maps of SST, averaged over the previous five years, to predict the mean SST anomaly in the North Atlantic (40°-60°N, 10°-70°W) over the following five years.

The weighted mask learned by the neural network is shown in Figure 2a. The Greenland Sea and the gulf stream region in the western North Atlantic emerge as the most important regions for identifying analogs in the MPI-GE. Over the western North Atlantic, there is an area of zero weight between two areas of high weight. These may be where the boundaries of persistent SST anomalies vary, and the neural network has learned that the specific locations of these boundaries are not important for the prediction problem. Previous studies that have used an analog approach to assess North Atlantic decadal predictability selected the best analogs by taking a correlation over the whole globe (Mahmood et al., 2022) or the entire North Atlantic basin (Menary et al., 2021). As shown in Figure 2b-d, when using the weighted mask, the best analogs only have to look like the SOI in the highest weight regions. An example SOI is shown in Figure 2b and its best analog in Figure 2c. These two maps look similar in the North Atlantic, but are starkly different in the North Pacific and Indian Ocean, among other regions. Once the weighted mask has been applied to the SOI (Figure 2d) and its best analog (Figure 2e), the maps look nearly identical.

These results suggest that using uniform weights across the entire North Atlantic basin, or the whole globe, may lead to a selection of analogs that are not optimized for forecasting multi-year variability in the North Atlantic. Indeed, we see that this is true in the skill scores shown in Figure 3a. For $1 \leq \mathrm{N} \leq 50$, where the top-N analogs are averaged, our weighted mask analog forecast outperforms the global and target region analog forecasts, as well as the climatology, mean target evolution, and random baselines. The skill score is lowest when only the single best analog is used for forecasting, and subsequently improves for larger N. Given that the skill score maximizes around $\mathrm{N} = 10$, and the spread of the targets associated with the analogs (i.e. the uncertainty of the forecast) increases with N (Figure S1), we elect to focus on results for $\mathrm{N} = 10$ analogs. The prediction by the top-10 analogs, and the spread of the targets, are shown in Figure 3b for 200 years of SOIs. The analog predictions do a good job of capturing the variability of North Atlantic sea surface temperatures, though they do struggle to forecast the most extreme anomalies.

5 Seasonal Prediction of El Niño Southern Oscillation

In addition to improving multi-year forecasts of SST in the North Atlantic, the learned weighted mask improves forecasts of ENSO on seasonal timescales. ENSO is the leading mode of global annual SST variability (Hsiung & Newell, 1983) and has an extensive influence on global weather and climate (reviewed in Yeh et al., 2018). Analog forecasting has been applied to seasonal prediction of ENSO in several studies due to its potential to outperform initialized GCM forecasts (e.g., Ding et al., 2018, 2019). In the following example, we use wintertime (November-March) global SST anomalies to forecast SST anomalies in the Niño3.4 region (5°S-5°N, 120-170°W; Barnston et al., 1997; Hanley et al., 2003) the following winter.

The weighted mask for forecasting ENSO looks markedly different from that for forecasting North Atlantic multi-year variability (Figure 4a). While a few regions are assigned higher weights, the weights in Figure 4a are much more uniform across the globe than in Figure 2a. The four main regions that stand out in this weighted mask have also been identified as important precursors in previous literature: the western North Pacific (e.g., S.-Y. Wang et al., 2012), the Pacific Meridional Mode (e.g., Amaya, 2019), the Central Atlantic (e.g., Martín-Rey et al., 2015), and the tropical Pacific itself (e.g., Capotondi & Sardeshmukh, 2015). The skill score of the global analog forecast (Figure 4b) is similar to that of our weighted mask analog forecast (but always lower, see Figure S3), which is not surprising since the values of the weighted mask are near one for most areas of the globe.

Since the weighted mask can be manually updated *post hoc*, we use this to explore the sensitivity of the forecast skill to which regions are included in the weighted mask. Figure 5a shows the weighted mask for ENSO prediction (Figure 4a) but where the smallest 95 percent of the weights have been set to zero. Forecasts made with this "constrained" weighted mask have similar skill to the original weighted mask (as shown in Figure S4). From the constrained weighted mask, we identify four main precursor regions for ENSO: the West Pacific (ocean grid points bounded by 0°-40°N, 100°-170°E), the Tropical Pa-

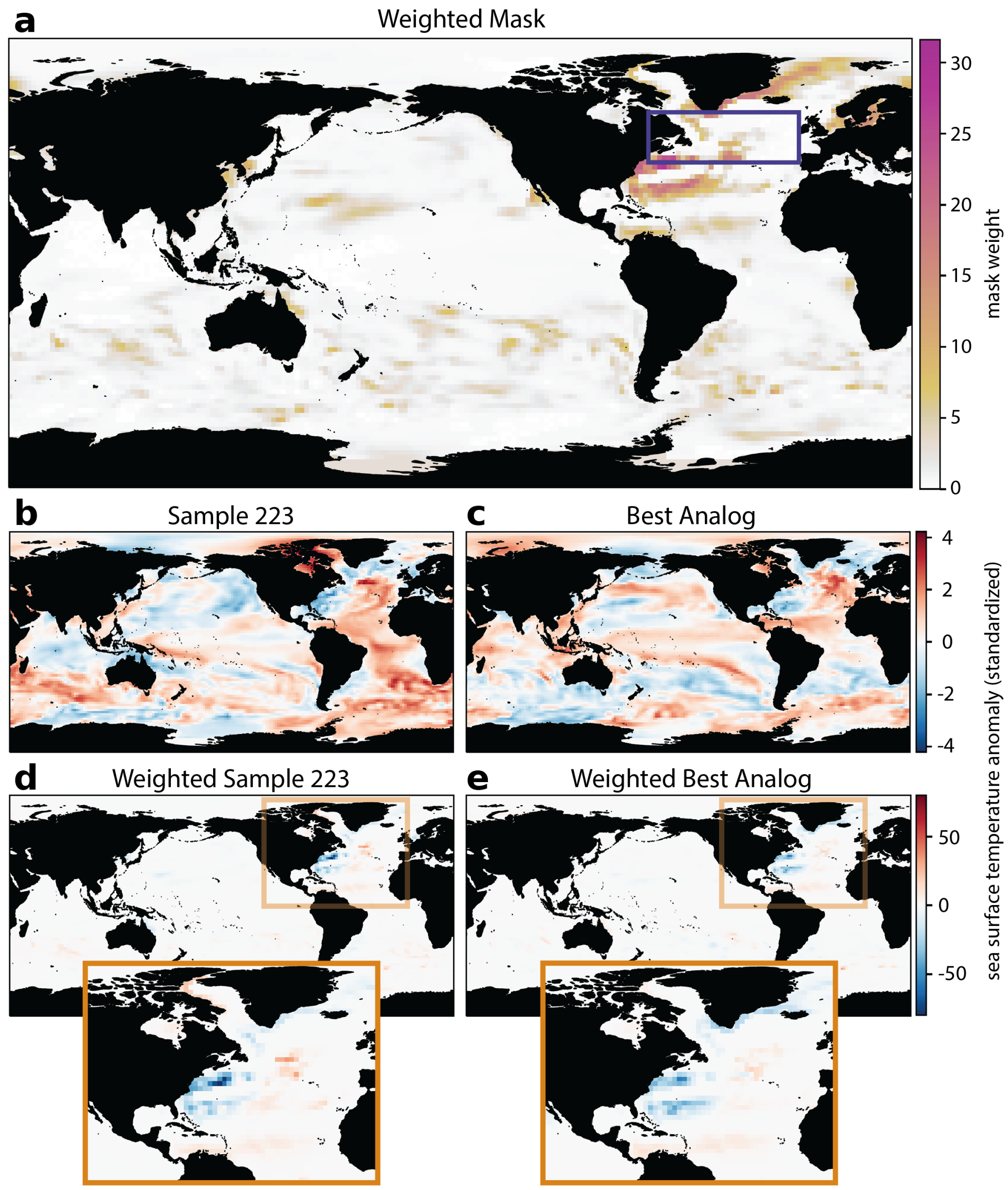

Figure 2. Weighted mask and example for multi-year predictions of North Atlantic SST. (a) Weighted mask, as learned by the interpretable neural network. (b) Standardized SST anomalies for a sample state of interest (SOI). (c) Standardized SST anomalies for the best analog associated with the SOI. (d) Weighted SOI. (e) Weighted best analog.

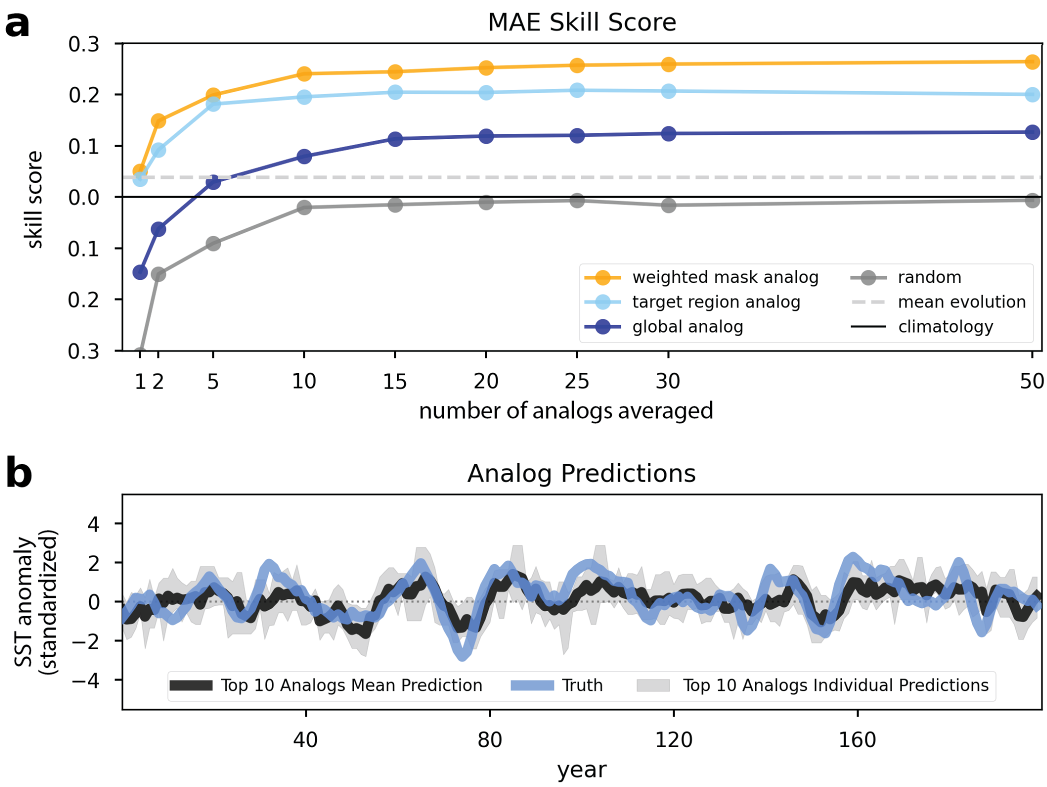

Figure 3. Analog forecasts of North Atlantic sea surface temperature. (a) Skill scores for our weighted mask analog forecast and other baselines. (b) Weighted mask analog forecasts for 200 years of MPI-GE simulations, including the mean prediction from the top-10 analogs, the spread of these predictions, and the truth values.

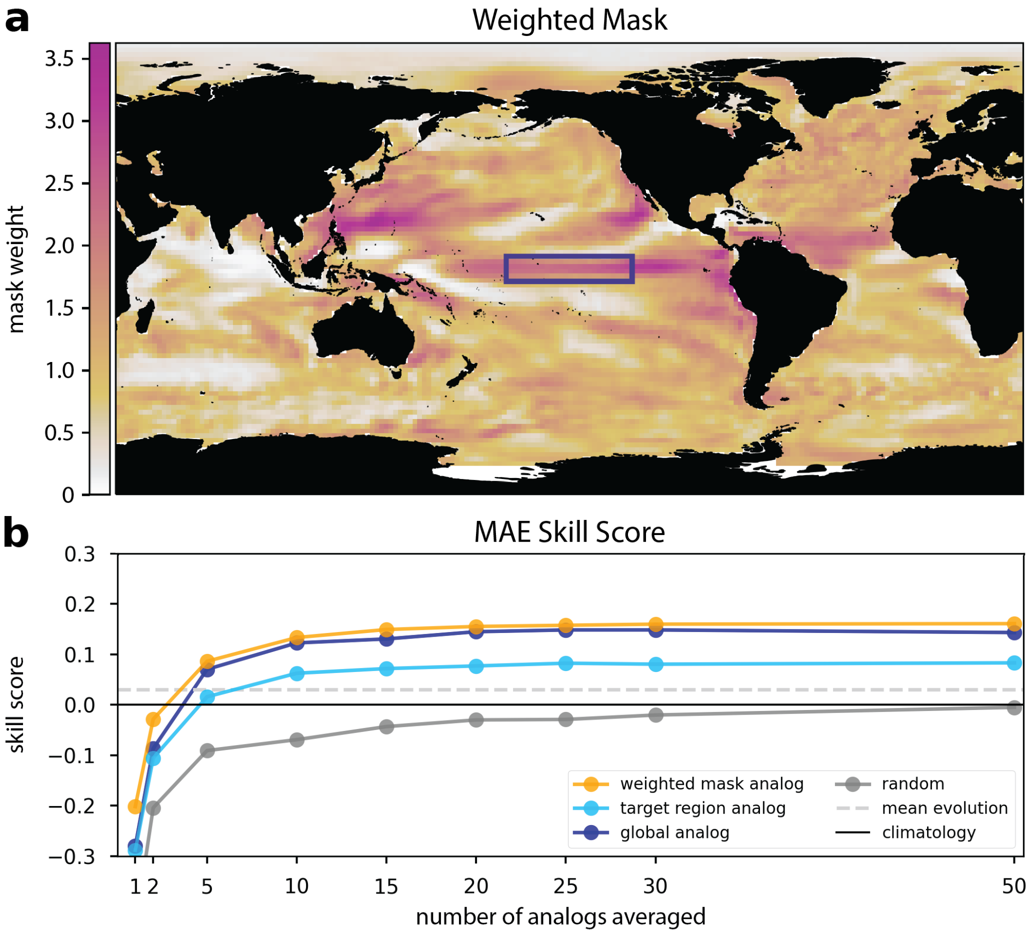

Figure 4. Weighted mask and skill scores for seasonal predictions of El Niño Southern Oscillation. (a) Weighted mask. (b) Skill scores for our weighted mask analog and other baselines.

cific (25°S-10°N, 170°E-65°W), the Baja Coast (10°N-40°N, 110°-140°W), and the Tropical Atlantic (0°-20°N, 20°-80°W).

We assess how important each precursor region is in two ways. In the first approach, we test the skill score of analog forecasting when each region is occluded from the constrained weighted mask (weights in that region are set to zero). When all four regions are included, the skill score is 0.146. Removing any of the four regions from the weighted mask results in a skill score decrease. Interestingly, removing the Tropical Atlantic results in the most drastic decrease in prediction skill. While the Tropical Atlantic has been connected to ENSO predictability (e.g., Martín-Rey et al., 2015), it is not considered a primary driver (C. Wang, 2018). In the second approach, we isolate each of the four regions (weights outside that region are set to zero). There is no improvement over climatology when just the Baja Coast or Tropical Atlantic is used to select analogs, and more skill when just the West Pacific or Tropical Pacific is used. However, no region alone provides anywhere near the skill that all four regions do together.

6 Discussion and Conclusions

We have shown how an interpretable neural network can be used to identify a weighted mask that improves the selection of analogs for seasonal-to-decadal forecasting. The precursors identified in the weighted masks are not necessarily causal, but they do provide the optimal predictors for the given input. In this work we have constrained the neural network to learn one mask that represents all pathways of predictability, however allowing the network to learn different masks for different SOIs could lead to better analog forecasts.

While we only used a single input map of SST to predict a future target SST in this work, this neural network architecture can be used for many other forecasting approaches. For example, a combination of multiple variables can be used as the predictors (such as SST and sea surface height, as in Ding et al., 2018) or other geophysical variables may be selected as the target (e.g., predicting precipitation over land). In addition, one may also include variables at multiple lead times to capture the time tendency of the climate system. We show results that include SST tendency as an input for the North Atlantic multi-year prediction example in Figure S5.

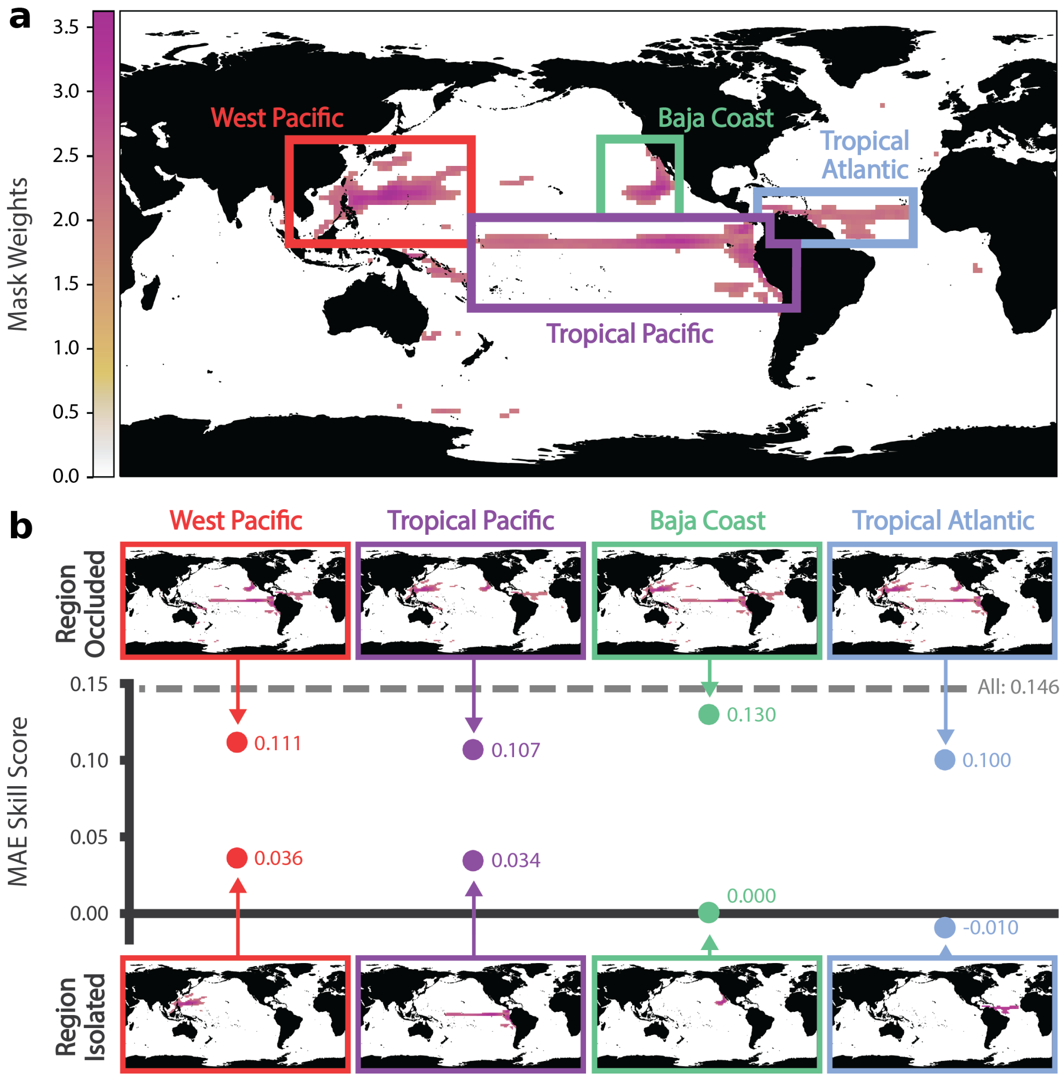

Figure 5. Analog forecasting skill of El Niño Southern Oscillation when various regions are occluded or isolated. (a) As in Figure 4a, but the lowest 95 percent of weights are set to zero. Four regions of focus are highlighted by the colored boxes. (b) Skill scores for analog forecasts when each region is occluded from the mask (top) and when the region is isolated to make a forecast (bottom).

We have explored this method through a perfect model setup. As such, the identified precursors are intrinsic to MPI-ESM1.1 and may not reflect patterns of predictability in the observed Earth system. Training the weighted mask on a multi-model ensemble may provide patterns that are more consistent with observations (e.g. Kirtman et al., 2014; Rader et al., 2022) and allow for enhanced analog predictions on real data. Additionally, we could train on models and observations at the same time to identify a weighted mask that is more representative of the true Earth System. We believe that this weighted mask approach will be influential to analog forecasting moving forward.

Open Research Section

The data used in this study, simulations from the Max Planck Institute for Meteorology Grand Ensemble, are publicly available at https://esgf-data.dkrz.de/projects/mpi-ge/. The weighted masks, and all python code used to generate the data and figures in this paper, can be found at XXX REVIEWERS, THIS CAN CURRENTLY BE FOUND AT https://github.com/jaminrader/WeightedMaskAnalogForecasting AND WILL BE UPLOADED TO ZENODO WHEN FINISHED XXX.

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0020347, and by grant AGS-2210068 from the National Science Foundation. The authors would like to thank colleagues at the NOAA Physical Sciences Laboratory for their helpful feedback.

References

Amaya, D. J. (2019, December). The pacific meridional mode and ENSO: a review. *Current Climate Change Reports*, *5*(4), 296–307. doi: 10.1007/s40641-019-00142-x

Balaguru, K., Foltz, G. R., & Leung, L. R. (2018, May). Increasing magnitude of hurricane rapid intensification in the central and eastern tropical atlantic. *Geophys. Res. Lett.*, *45*(9), 4238–4247. doi: 10.1029/2018gl077597

Barnston, A. G., Chelliah, M., & Goldenberg, S. B. (1997, September). Documentation of a highly ENSO-related sst region in the equatorial pacific: Research note. *Atmosphere-Ocean*, *35*(3), 367–383. doi: 10.1080/07055900.1997.9649597

Capotondi, A., & Sardeshmukh, P. D. (2015, November). Optimal precursors of different types of ENSO events. *Geophys. Res. Lett.*, *42*(22), 9952–9960. doi: 10.1002/2015gl066171

Delle Monache, L., Eckel, F. A., Rife, D. L., Nagarajan, B., & Searight, K. (2013, October). Probabilistic weather prediction with an analog ensemble. *Mon. Weather Rev.*, *141*(10), 3498–3516. doi: 10.1175/mwr-d-12-00281.1

Ding, H., Newman, M., Alexander, M. A., & Wittenberg, A. T. (2018, July). Skillful climate forecasts of the tropical Indo-Pacific ocean using Model-Analogs. *J. Clim.*, *31*(14), 5437–5459. doi: 10.1175/JCLI-D-17-0661.1

Ding, H., Newman, M., Alexander, M. A., & Wittenberg, A. T. (2019, February). Diagnosing secular variations in retrospective ENSO seasonal forecast skill using CMIP5 model-analogs. *Geophys. Res. Lett.*, *46*(3), 1721–1730. doi: 10.1029/2018gl080598

Enfield, D. B., Mestas-Nuñez, A. M., & Trimble, P. J. (2001, May). The atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.*, *28*(10), 2077–2080. doi: 10.1029/2000gl012745

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., … Stevens, B. (2013, July). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model intercomparison project phase 5. *J. Adv. Model. Earth Syst.*, *5*(3), 572–597. doi: 10.1002/jame.20038

Goldenberg, S. B., Landsea, C. W., Mestas-Nunez, A. M., & Gray, W. M. (2001, July). The recent increase in atlantic hurricane activity: causes and implications. *Science*, *293*(5529), 474–479. doi: 10.1126/science.1060040

Hanley, D. E., Bourassa, M. A., O'Brien, J. J., Smith, S. R., & Spade, E. R. (2003, April). A quantitative evaluation of ENSO indices. *J. Clim.*, *16*(8), 1249–1258. doi: 10.1175/1520-0442(2003)16⟨1249:AQEOEI⟩2.0.CO;2

Hawkins, E., Robson, J., Sutton, R., Smith, D., & Keenlyside, N. (2011, December). Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach. *Clim. Dyn.*, *37*(11-12), 2495–2509.

doi: 10.1007/s00382-011-1023-3

Hsiung, J., & Newell, R. E. (1983, October). The principal nonseasonal modes of variation of global sea surface temperature. *J. Phys. Oceanogr.*, *13*(10), 1957–1967. doi: 10.1175/1520-0485(1983)013⟨1957:TPNMOV⟩2.0.CO;2

Jackson, L. C., Biastoch, A., Buckley, M. W., Desbruyères, D. G., Frajka-Williams, E., Moat, B., & Robson, J. (2022, March). The evolution of the north atlantic meridional overturning circulation since 1980. *Nature Reviews Earth & Environment*, *3*(4), 241–254. doi: 10.1038/s43017-022-00263-2

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., ... Wood, E. F. (2014, April). The north american multimodel ensemble: Phase-1 Seasonal-to-Interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Am. Meteorol. Soc.*, *95*(4), 585–601. doi: 10.1175/BAMS-D-12-00050.1

Kushnir, Y., Scaife, A. A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., ... Wu, B. (2019, January). Towards operational predictions of the near-term climate. *Nat. Clim. Chang.*, *9*(2), 94–101. doi: 10.1038/s41558-018-0359-7

Lorenz, E. N. (1963, March). Deterministic nonperiodic flow. *J. Atmos. Sci.*, *20*(2), 130–141. doi: 10.1175/1520-0469(1963)020⟨0130:DNF⟩2.0.CO;2

Lorenz, E. N. (1969, July). Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, *26*(4), 636–646. doi: 10.1175/1520-0469(1969)26⟨636:APARBN⟩2.0.CO;2

Lou, Newman, M., & Hoell, A. (2023, February). *Multi-decadal variation of ENSO forecast skill since the late 1800s.* doi: 10.21203/rs.3.rs-2544766/v1

Maher, N., Milinski, S., Suarez-Gutierrez, L., & others. (2019). The max planck institute grand ensemble: enabling the exploration of climate system variability. *Journal of Advances*. doi: 10.1029/2019MS001639

Mahmood, R., Donat, M. G., Ortega, P., Doblas-Reyes, F. J., Delgado-Torres, C., Samsó, M., & Bretonnière, P.-A. (2022, October). Constraining low-frequency variability in climate projections to predict climate on decadal to multi-decadal timescales – a poor man's initialized prediction system. *Earth Syst. Dyn.*, *13*(4), 1437–1450. doi: 10.5194/esd-13-1437-2022

Martín-Rey, M., Rodríguez-Fonseca, B., & others. (2015). Atlantic opportunities for ENSO prediction. *Geophys. Res. Lett.*. doi: 10.1002/2015GL065062

Menary, M. B., Mignot, J., & Robson, J. (2021, June). Skilful decadal predictions of subpolar north atlantic SSTs using CMIP model-analogues. *Environ. Res. Lett.*, *16*(6), 064090. doi: 10.1088/1748-9326/ac06fb

Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A. S., … Yeager, S. (2020, September). Subseasonal to decadal prediction: Filling the Weather–Climate gap. *Bull. Am. Meteorol. Soc.*, *101*(9), 767–770. doi: 10.1175/BAMS-D-19-0037.A

Mulholland, D. P., Laloyaux, P., Haines, K., & Balmaseda, M. A. (2015, November). Origin and impact of initialization shocks in coupled Atmosphere–Ocean forecasts. *Mon. Weather Rev.*, *143*(11), 4631–4644. doi: 10.1175/MWR-D-15-0076.1

Peng, W., Chen, Q., Zhou, S., & Huang, P. (2021, March). CMIP6 model-based analog forecasting for the seasonal prediction of sea surface temperature in the offshore area of china. *Geoscience Letters*, *8*(1), 1–8. doi: 10.1186/s40562-021-00179-7

Rader, J. K., Barnes, E. A., Ebert-Uphoff, I., & others. (2022). Detection of forced change within combined climate fields using explainable neural networks. *Journal of Advances*.

Shekhar, M., Sharma, A., Dimri, A. P., & Tandon, S. K. (2022, July). Asian summer monsoon variability, global teleconnections, and dynamics during the last 1,000 years. *Earth-Sci. Rev.*, *230*, 104041. doi: 10.1016/j.earscirev.2022.104041

Si, D., Hu, A., Jiang, D., & Lang, X. (2023, February). Atmospheric teleconnection associated with the atlantic multidecadal variability in summer: assessment of the CESM1 model. *Clim. Dyn.*, *60*(3), 1043–1060. doi: 10.1007/s00382-022-06331-z

Sutton, R. T., & Allen, M. R. (1997, August). Decadal predictability of north atlantic sea surface temperature and climate. *Nature*, *388*(6642), 563–567. doi: 10.1038/41523

Towler, E., PaiMazumder, D., & Done, J. (2018, March). Toward the application of decadal climate predictions. *J. Appl. Meteorol. Climatol.*, *57*(3), 555–568. doi: 10.1175/JAMC-D-17-0113.1

Van den Dool, H. M. (1994, May). Searching for analogues, how long must we wait? *Tellus A*, *46*(3), 314–324. doi: 10.1034/j.1600-0870.1994.t01-2-00006.x

Wang, C. (2018, October). A review of ENSO theories. *Natl Sci Rev*, *5*(6), 813–825. doi: 10.1093/nsr/nwy104

Wang, S.-Y., L'Heureux, M., & Chia, H.-H. (2012, March). ENSO prediction one year in advance using western north pacific sea surface temperatures. *Geophys. Res. Lett.*, *39*(5). doi: 10.1029/2012GL050909

Wu, Y., & Yan, X. (2023, May). Evaluating changes in the multiyear predictability of the pacific decadal oscillation using model analogs since 1900. *J. Mar. Sci. Eng.*, *11*(5), 980. doi: 10.3390/jmse11050980

Yeh, S. W., Cai, W., Min, S. K., McPhaden, M. J., & others. (2018). ENSO atmospheric teleconnections and their response to greenhouse gas forcing. *Reviews of*. doi: 10.1002/2017RG000568

Zhang, L., Delworth, T. L., Yang, X., Morioka, Y., Zeng, F., & Lu, F. (2023, February). Skillful decadal prediction skill over the southern ocean based on GFDL SPEAR Model-Analogs. *Environ. Res. Commun.*, *5*(2), 021002. doi: 10.1088/2515-7620/acb90e