Assessing Tropical Pacific-induced Predictability of Southern California Precipitation Using a Novel Multi-input Multi-output Autoencoder

Linsey S. Passarella[1], **Salil Mahajan**[1]

[1]Computational Earth Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Key Points:

- We design a novel MIMO-AE to capture the non-linear relationships between tropical Pacific SSTs and Southern California precipitation.
- We use long-short term memory models of a MIMO-AE derived index to assess predictability of Southern California precipitation.
- MIMO-AE offers statistically significant improvement in predictive skill of Southern California precipitation on sub-seasonal scales.

Corresponding author: Linsey S Passarella, `passarellals@ornl.gov`

Abstract

We construct a novel Multi-Input Multi-Output Autoencoder-decoder (MIMO-AE) to capture the non-linear relationship of Southern California precipitation (SC-PRECIP) and tropical Pacific Ocean sea surface temperature (TP-SST). The MIMO-AE is trained on both monthly TP-SST and SC-PRECIP anomalies simultaneously. The co-variability of the two fields in the MIMO-AE shared nonlinear latent space can be condensed into an index, termed the MIMO-AE index. We use a transfer learning approach to train a MIMO-AE on the combined dataset of 100 years of output from a historical simulation with the Energy Exascale Earth Systems Model version 1 (E3SMv1) and a segment of observational data. We further use Long Short-Term Memory (LSTM) networks to assess sub-seasonal predictability of SC-PRECIP using the MIMO-AE index. We find that the MIMO-AE index provides enhanced predictability of SC-PRECIP for a lead-time of up-to four months as compared to Niño 3.4 index and the El Niño Southern Oscillation Longitudinal Index.

Plain Language Summary

Traditional El Niño Southern Oscillation indices, like the Niño 3.4 index, although well-predicted themselves, fail to offer reliable sub-seasonal to seasonal predictions of Western US precipitation. Here, we use a machine learning approach called a multi-input, multi-output autoencoder to capture the relationship between tropical Pacific and Southern California precipitation and project it onto a new index, which we call MIMO-AE index. Using machine learning based time-series predictions, we find that MIMO-AE index offers enhanced predictability of Southern California precipitation up-to a lead time of four months as compared to other ENSO indices.

1 Introduction

Extracting sub-seasonal and seasonal predictability of precipitation over California from the leading mode of climate variability - the El Niño-Southern Oscillation (ENSO) - remains a challenge, even during strong ENSO episodes (e.g. L'Heureux et al., 2021; Pan et al., 2019; S. Wang et al., 2017). This was apparent during the 2015-16 El Niño event, when California received just above average precipitation despite strong Niño 3.4 index anomalies. And, in contrast to the predicted (Hoell et al., 2016) (also see CPC precipitation outlooks: https://www.cpc.ncep.noaa.gov/products/archives/long_lead/) high likelihood of heavy precipitation, which occurred there during the 1982-83 and 1997-98 strong El Niño events (e.g. Cohen et al., 2017; Lee et al., 2018; L'Heureux et al., 2017). Studies suggest that atmospheric intrinsic variability (e.g. Kumar et al., 2016; Cheng et al., 2021; L'Heureux et al., 2021), such as that of the jet-stream, and remote influences like Arctic oscillation (Cohen et al., 2017) could have dampened California precipitation response in 2015-2016. However, some studies also suggest that the contrast in spatial pattern of tropical Pacific SST anomalies between the Central Pacific (or Modoki) 2015-16 El Niño event and other canonical (Eastern Pacific) events also contributed to the differing Western US precipitation response (e.g. Patricola et al., 2020).

ENSO indices, like the static area-averaged Niño3.4 index, do not capture this ENSO diversity in the spatial pattern of SST anomalies (e.g. Trenberth & Stepaniak, 2001; Williams & Patricola, 2018). Linear statistical modeling suggests that Niño3.4 index can only explain about 9-25% of variability of California precipitation on seasonal to inter-annual time-scales, with the largest skill over Southern California (e.g. Jong et al., 2016; X. Huang & Ullrich, 2017; G. Wang et al., 2021; Cheng et al., 2021). Also, ENSO teleconnections exhibit clear non-linear behavior with stronger response for El Niño events than La Niña events and non-linear response to extreme ENSO events (Hoerling et al., 1997; Frauen et al., 2014). A recent study suggests that the predictive skill of precipitation over the Western US from ENSO can be enhanced by using a non-linear index that accounts for the spatial diversity of SST anomalies, and hence convective activity and extra-tropical wave-train response, during ENSO events (e.g. Patricola et al., 2020).

To enhance regional seasonal predictability afforded by the low frequency variability modes, it is thus important to account for the complexity of the modes as well as their teleconnections more fully. Here, in a novel approach, we adapt a multitask learning autoencoder technique to capture non-linear co-variability patterns of tropical Pacific SSTs and Southern California precipitation. Autoencoders are artificial neural networks, with the same input and output layers, that regenerate the original data from efficient representations (encodings) of the data like principal component analysis (PCA). Autoencoders, however, transform data to non-linear latent spaces via non-linear activation functions, thus imparting the additional capability of capturing the underlying non-linear relationships within the data (e.g. Y. Wang et al., 2016; Charte et al., 2018; Masti & Bemporad, 2018). Studies show that autoencoders can better detect dominant variability patterns over other techniques, like the PCA (e.g. Y. Wang et al., 2016; Zamparo & Zhang, 2015; Fournier & Aloise, 2019). Some studies (e.g. Tang & Hsieh, 2003; He & Eastman, 2020) have also demonstrated the use of autoencoders to effectively identify modes of climate variability, including those related to ENSO. Multitask learning, on the other hand, solves multiple learning tasks at the same time while exploiting commonalities and differences across tasks (e.g. Caruana, 1997). It has been applied to many problems including natural language processing (e.g. S. Chen et al., 2021), speech recognition (e.g. Toshniwal et al., 2017) and computer vision (e.g. Kendall et al., 2018) to improve prediction accuracy and learning efficiency of task-specific models. Ghifary et al. (2015) implemented multitask learning with a multi-output autoencoder, for related cross-domain object recognition. It has a single input layer and multiple output layers corresponding to different domains - where a domain refers to transformation of the object image, like rotation of the viewing angle.

Here, we expand on this approach and design a new network, termed as the multi-input multi-output autoencoder (MIMO-AE), to effectively extract the most prominent shared features between monthly TP-SSTs and SC-PRECIP anomalies and capture their underlying non-linear relationships, using an Earth System Model (ESM) simulation data and observational data. A similar design, albeit for convolutional neural networks, was proposed by Raza et al. (2017) for cell features segmentation in biomedical microscopy images. Our architecture of the MIMO-AE yields a temporal index of the co-variability of the two variables. We further use Long Short-term Memory (LSTM) models to predict this monthly index, which we decode to generate predicted SC-PRECIP, and evaluate its predictive skill. We show that MIMO-AE is a powerful tool to isolate important teleconnections and wield enhanced sub-seasonal regional predictability.

1.1 Model Simulations and Data

We use a 165 years-long historical simulation of the Energy Exascale Earth System Model version 1 (E3SMv1) (E3SM Project, 2018), and utilize the first 100 years of the simulation for training the MIMO-AE network. E3SMv1 is found to effectively capture temporal variability of ENSO and reproduce ENSO associated spatial SST patterns when compared to observational datasets (Golaz et al., 2019), although with a larger westward extent of SST anomalies during El Niño events. It also simulates the teleconnections of ENSO to US winter season precipitation well (Mahajan et al., 2021). We use observed precipitation data from NOAA's PRECipitation REConstruction over Land (PREC/L) at 1° resolution (M. Chen et al., 2002). PREC/L is a global analysis of interpolated rain gauge observations from 1948 to 2020. We use observed SSTs for the same period from the Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST 1.1) dataset available at a 1° resolution (Rayner et al., 2003).

2 Methodology

2.1 Autoencoder

An autoencoder is an unsupervised neural network that is trained to learn an identity function, a function that returns the same value as its input. It aims to efficiently compress and encode data by minimizing the reconstruction error. A simple autoencoder, shown in figure 1a, contains a hidden layer h that describes a representation of important attributes of the input (e.g. Goodfellow et al., 2016). The general autoencoder consists of two parts: an encoder and a decoder. The encoder maps input x to h by a chosen activation function $f()$,

$$h = f(x \cdot w_e) \tag{1}$$

where w_e are the encoder weights. The decoder then maps h to the reconstruction of x, represented by x':

$$x' = f(h \cdot w_d) \tag{2}$$

where w_d are the decoder weights.

By using a linear activation function, the single hidden layer autoencoder behaves similarly to a PCA (e.g. Bourlard & Kamp, 1988; Plaut, 2018). The number of hidden layers can also be increased to create a deep autoencoder, with the middle layer often referred to as the bottleneck layer. Tang and Hsieh (2003) used a simple autoencoder to extract the leading nonlinear mode of interannual variability of upper ocean heat content over the tropical Pacific, with a single node bottleneck hidden layer, to reveal an asymmetry in the spatial pattern between characteristic El Niño and La Niña episodes. For spatio-temporal data, the temporal vectors at the bottleneck nodes are analogous to principal components (PC) of PCA. The value of a temporal vector at a given time t results from passing the spatial data at t through the network. The non-linear activation functions imply that the spatial pattern derived from reconstructing the data using the decoder varies with the magnitude of the temporal vector at t, unlike PCs which yield a standing spatial pattern (e.g. Tang & Hsieh, 2003).

2.1.1 MIMO-AE

Figure 1b illustrates our MIMO-AE architecture designed to extract the non-linear relationship between TP-SSTs and SC-PRECIP on monthly timescales. The encoder consists of two separate input temporal vectors (TP-SST and SC-PRECIP) that are passed through two hidden layers before concatenating and passing through a single hidden node. The input (and output) vectors represent SST anomalies at each grid box within the boxed domain over tropical Pacific (20°N to 20°S, 120°E to 70°W) and precipitation anomalies over each grid box in the boxed domain over Southern California (32°N to 35°N, 120°W to 114°W) (Figure 1b). The first hidden layer, consisting of 50 nodes each for the two variables, can be thought of as feature extraction of the original data. The next hidden layer then shrinks the data to 10 hidden nodes, again separately for the two variables, in order to reduce the computational complexity of data. This data is then passed to a single hidden node that is shared by the two input variables. This hidden node represents the shared non-linear latent structure of both the SST and precipitation vectors. The vectors are then split back into two from the shared hidden node and passed through the decoder, which is identical in structure (with different weights) to the encoder, to reconstruct back to the original shape in the output layer. We use the ”tanh”, or hyperbolic tangent, activation function for all the hidden layers.

We performed several iterations of the network design with different number of hidden layers, neurons and activation functions and chose the MIMO-AE architecture (described above) that exhibited a low value of the training loss function as well as explaining a large fraction of the variability of SC-PRECIP. The loss is calculated by using a mean squared error (MSE) using the following equation:

$$MSE = \frac{1}{N}\sum_{i}(P_i - T_i)^2 \tag{3}$$

where P_i is the predicted value of the reconstructed data at point i and T_i is the true value of the data at point i, which here is the original input data. The input variables are scaled using a min-max scaler before training is performed. MIMO-AE was trained on first 100 years of the E3SM simulation data for 100 epochs with an AdaGrad loss optimizer using tensorflow on one CPU node on the National Energy Research Scientific Computing Center's (NERSC) Cori super computer. The training loss for the scaled TP-SST reconstruction (orange) and SC-PRECIP (blue) are shown in figure 1c. We refer to this MIMO-AE network as MIMO-AE-E3SM, hereafter.

Figure 1d and e show the R^2 values (fraction of variance explained) between the reconstructions from the MIMO-AE-E3SM and the original data for the 100 years of training data for TP-SSTs and SC-PRECIP respectively. MIMO-AE explains more than 80% of the variability of Southern California for most grid points and about 20% of the variability of TP-SSTs over most of the domain. The relatively weaker explained variability of MIMO-AE over tropical Pacific is an artefact of our network design preference. We ad-hocly chose a network that explained a larger fraction of the variability of SC-PRECIP, while also capturing the teleconnections to TP-SST, since our goal was largely to assess predictability of SC-PRECIP here. In the future, we plan to make the network design more systematic, for example, by adding a penalizing term for explained variability of each field in the loss function.

We refer to the temporal vector of the single node bottleneck layer that represents the dominant non-linear mode of co-variability of TP-SSTs and SC-PRECIP as the MIMO-AE index, hereafter. We apply the MIMO-AE-E3SM trained on 100 years of E3SM historical simulation on the latter 65 years of the run. As a form of transfer learning, we combine the first 100 years of the E3SM simulation with 32 years of observational data (1948-1979) to train another MIMO-AE network for application to remaining observational data (1980-2020), termed MIMO-AE-OBS. Although, we find that using MIMO-AE-E3SM on observational data imparted similar predictability skills (Results section) as MIMO-AE-OBS for observational data. Ham et al. (2019) also used a transfer learning approach, whereby, to predict the Niño 3.4 index they train a convolutional neural network (CNN) with observational SST and heat content data but with the network weights initialized from a network trained on historical simulations of 21 CMIP5 models. While not investigated in our exploratory study of MIMO-AE here, we plan to apply this and other transfer learning methods to MIMO-AE in the future.

2.2 LSTM

To study predictability, we also train long-short term memory (LSTM) recurrent neural networks as our time series prediction models. LSTMs are a special kind of recurrent neural network that learn long term dependencies (Hochreiter & Schmidhuber, 1997). They have recently been shown to perform better at time series prediction of Niño 3.4 index over linear models (A. Huang et al., 2019; Mu et al., 2020; Broni-Bedaiko et al., 2019; Gupta et al., 2020).

LSTM models are constructed individually for each of the time series of MIMO-AE index, Niño 3.4 index, ELI , SVD and regionally averaged SC-PRECIP anomalies using the first 100 years of the E3SM data. We train separate LSTMs for the above listed time series using the first 32 year segment (1948-1979) of observational dataset used. Given a predicted value of MIMO-AE index, predicted SC-PRECIP (and TP-SSTs) can be constructed by passing the index through the decoder of MIMO-AE. We optimize the LSTM architecture by choosing the number of hidden nodes that maintains a low training loss for all indices, found to be 100 nodes. We train separate LSTMs for each of the forecast lead times ranging from 1 to 12 months and evaluate their predictive skill on the remaining 65 years of E3SM data and the 41 years of observational data.

3 Results

3.1 MIMO-AE

Figure 2a shows the the three-month moving average of the standardized MIMO-AE index time-series for a segment (last 40 years, 1974-2013) of the 65 years of the E3SM testing data. MIMO-AE index was generated by passing the TP-SST and SC-PRECIP data through the MIMO-AE network trained on the prior 100 years of simulation. As a linear approach, we also conduct a singular value decomposition (SVD) of the covariance matrix of grid point data anomalies of SST over TP-SST and precipitation over the SC-PRECIP region using the training data. SVD is a commonly used method of extracting linear co-variability patterns in multivariate climate data (e.g. Bretherton et al., 1992; Chang et al., 1997; Mahajan et al., 2010). Figure 2a also shows the time series of the expansion coefficients of SST over the TP-SST domain (SVD1-SST) corresponding to the first leading mode of the SVD (SVD1) for the same segment of the testing period. SVD1 explains 28% of the squared covariance of TP-SST and SC-PRECIP. Also shown are the Niño 3.4 index, ELI and domain averaged SC-PRECIP. The correlations of each time-series against domain averaged SC-PRECIP is also listed for the smoothed data. Fig. 2b shows the same but for a segment of the observational data (1980-2019) and using MIMO-AE-OBS. To reiterate, MIMO-AE index for observations is computed by passing the observed data through the MIMO-AE-OBS network. SVD1-SST time series for the segment is also computed by projecting the TP-SST testing data on the SVD1 computed from a combination of E3SM and observational training data.

For both E3SM and observations, the correlation between SC-PRECIP and MIMO-AE index is higher than that between SC-PRECIP and Niño 3.4 index or ELI, given that precipitation data is fed in the generation of MIMO-AE index and MIMO-AE explains a large fraction of the SC-PRECIP variability. The time series of the expansion coefficient of precipitation over SC-PRECIP domain also shows a high correlation with SC-PRECIP for E3SM (0.98) and observations (0.98). The correlation between SVD1-SST time-series and SC-PRECIP is weaker for both data (about 0.17) pointing to the limitations of linear methods. The correlation between MIMO-AE and both Niño 3.4 index and ELI is weak both for E3SM (0.35 and 0.27 respectively) and observational data (0.43 and 0.39 respectively). However, the correlation between MIMO-AE and Niño 3.4 is higher than the correlation between SC-PRECIP and Niño 3.4. The above correlations are indicative of the shared variability captured by MIMO-AE. The correlation of SVD1-SST time-series and Niño3.4 index (0.94) is higher than that of MIMO-AE and Niño3.4 indices suggesting that linear methods like SVDs are dominated (and limited) by the variability over static regions when capturing relationships with the tropical Pacific. Further, in observational data, all indices spike during the 1982-83 and 1996-97 El Niño events, but only the Niño 3.4 index and SVD1-SST time-series peak during the 2015-16 El Niño event. MIMO-AE also categorizes the 2015/2016 event weaker than the Niño 3.4, similar to the ELI (Williams & Patricola, 2018).

Fig. 2d and e show the probability density functions of the Niño 3.4 index, ELI, SVD1-SST time-series, MIMO-AE index and domain averaged SC-PRECIP for E3SM testing data and observations. While the Niño 3.4 index and SVD1-SST time-series tend to be symmetric, the ELI is skewed towards the left (westwards), both for E3SM data and observations as noted by Williams and Patricola (2018), with a thicker right tail (eastwards). MIMO-AE which represents the shared co-variability between the TP-SSTs and SC-PRECIP also shows a leftwards skewed distribution with a larger number of strong positive events than strong negative events. The leftwards skewness may follow from the density function of precipitation that is naturally skewed leftwards, even for monthly average data (e.g. Mahajan et al., 2012). But, it could also be reflective of the skewed relationship between TP-SSTs and SC-PRECIP, with some events over the tropical Pacific triggering extreme positive anomalous events in SC-PRECIP. While, co-variability between the two remains weaker during strong negative SC-PRECIP anomalous events. The skewness of MIMO-AE index is noted to be stronger in E3SM than in observations.

Fig. S1a show the composite of reconstructions of TP-SST and SC-PRECIP during the strongest 10 positive and negative monthly MIMO-AE index values for the E3SM testing data. Strong positive (negative) SC-PRECIP anomalies during those events are associated with strong positive (negative) anomalies in central, north-central and northeast tropical Pacific, weak positive (negative) anomalies in the eastern tropical Pacific and negative (positive) anomalies over the northwestern Pacific. Similar patterns are noted for the strongest positive and negative MIMO-index values for reconstructions of TP-SST when observation data is passed through the MIMO-AE network (Fig. S1b).

Fig. 2c shows the December to February average reconstructions for the three strongest El-Niño events (1981-82, 1997-98, 2015-16) in observations. It is apparent that the spatial patterns of these reconstructions are not standing - more clearly here than the composite plots (Fig. S1) - with varying contour patterns of SST anomalies for each of the three events. The 2015-16 El Niño events is associated with weak positive anomalies in the MIMO-AE latent space for SC-PRECIP and TP-SST over much of tropical Pacific. In contrast, the 1981-82 and 1997-98 events are associated with strong positive anomalies both in SC-PRECIP and TP-SST, with the stronger 1997-98 SC-PRECIP anomalies associated with stronger TP-SST anomalies and varied contour delineations. When a separate MIMO-AE network is trained on all of the observation data (1948-2020) and with no E3SM data, the spatial pattern of the TP-SST during 1981-82 and 1997-98 exhibits a narrow band of strong anomalies over equatorial Pacific including coastal Eastern Pacific (not shown), illustrating the influence of E3SM model bias in MIMO-AE-OBS. In the future, we plan to explore ways to reduce the influence of model bias in MIMO-AE-OBS, for example by appropriately weighing the observational training data.

3.2 Predictability of MIMO-AE Index

We use LSTM models to predict the MIMO-AE index for lead times of 1 to 12 months. Figure 3a shows the predictive skill of the LSTMs to predict the MIMO-AE index for the 65 years of E3SM testing data. The predictive skill is computed as the correlation between the LSTM predicted value and the true value of the MIMO-AE index when data is passed through the network. The predictive skill of Niño 3.4 index, ELI index and SVD1-SST using LSTMs are also shown. One standard deviation spread, computed using the Fisher transformation, are shown as color shadings. The MIMO-AE index exhibits a lower predictive skill than Niño 3.4 index, ELI index and SVD1-SST at all lead times longer than one month. This is likely due to the dominance of noisy precipitation data in MIMO-AE, which demonstrates poor temporal auto-correlations on these time scales (e.g. Mahajan et al., 2012), offering little predictive skill.

This is evident in the Figure 3a, which also shows the predictive skill of domain averaged SC-PRECIP using LSTMs, and serves as a baseline for evaluation of predictive skill. Precipitation shows a high skill at a lead time of one month like the other indices, but offers poor predictive skill at longer lead times. MIMO-AE index provides more predictive skill than precipitation itself for two and three month lead times, likely due to the inclusion of TP-SSTs, which have higher predictive skill due to the thermal inertia of the oceanic mixed layer. But, MIMO-AE index provides poor skill for longer lead times. Figure 3b shows the LSTM skills as a function of the calendar month when the prediction is initialized for all indices and generally reflect Fig. 3a, while also showing the well-known spring predictability barrier associated with Niño3.4 and ELI.

The above results hold for the observational data too, with the the MIMO-AE index exhibiting poorer predictive skill when compared to Niño 3.4 index, ELI and SVD1-SST on these monthly time scales. Similar to E3SM data, MIMO-AE index demonstrates weaker skill at 2-months lead times and longer, while precipitation time series exhibits no skill at lead times longer than one month irrespective of the initial month of predictions (Figure 3d). Although, the skill of predicting MIMO-AE index is substantially higher than that of predicting SC-PRECIP.

3.3 SC-PRECIP Predictability from MIMO-AE Index

To evaluate the predictability of SC-PRECIP using the MIMO-AE index, we pass the predicted MIMO-AE index values through the decoder part of the MIMO-AE to construct spatio-temporal predictions of SC-PRECIP anomalies. Figure 3e shows the skill of predicted SC-PRECIP. The predicted spatial pattern of the SC-PRECIP constructed by the MIMO-AE decoder is domain averaged to compute predictive skill. For the Niño 3.4 index, ELI, and SVD1-SST we predict domain average SC-PRECIP from LSTM predicted values of the indices by using linear regression models (Fig. 3e). The linear regression models were constructed using the training data for E3SM and observations separately. Domain-averaged SC-PRECIP computed from reconstructions of SC-PRECIP from LSTM of SVD1-PRECIP predictions yield similar skill (not shown) as SC-PRECIP index, given their high correlation. MIMO-AE generated predicted precipitation exhibits stronger skill than other indices for lead times of up to 3-months.

However, MIMO-AE index's skills at lead times of one to three months are statistically indistinguishable from that of SC-PRECIP index at the 95% confidence level based on a two-tailed Student's t-test of the Fisher transformations of the correlations. To account for the auto-correlation in the time-series', we use effective sample size for the null hypothesis tests. We calculate this effective sample size using the following equation $N_{effective} = \frac{N}{1+2\sum_i^N \gamma_i^2}$ where γ_i is the auto-correlation of our SC-PRECIP time series at lag i and N is our total number of samples (Livezey & Chen, 1983). Although, the improved skill is a significant improvement over that of Niño 3.4 index, ELI and SVD1-SST. MIMO-AE skills are weaker and also indistinguishable from that of SC-PRECIP for longer lags, and become statistically indifferent from zero at a lead time of 6-months and longer. The skill of Niño 3.4 and ELI is statistically insignificant at all lead times on these monthly scales. A two-channel LSTM method, commonly used in areas such as text classification (e.g. Liang et al., 2021), that uses both Niño3.4 index and SC-PRECIP as inputs to predict both the indices and captures non-linear shared variability of the two indices also performs poorly than the MIMO-AE index (Fig. S2). This further supports a role of regions of the tropical Pacific, other than regions like Niño3.4, in enhancing the predictability of MIMO-AE. The enhanced predictive skill of precipitation from MIMO-AE up to a lead time of 3 months is noted for almost all initialization calendar months of the year as compared to other indices (Figure 3f).

Enhanced predictive skill of MIMO-AE of SC-PRECIP is also noted for the 41 years of observation testing data (Figure 3g). The improvement in MIMO-AE skill as compared to other indices is statistically significant at two to four months lead times at the 95% confidence level. The high skill at 1-month lead time is statistically indifferent from that of SC-PRECIP. And, the skills are statistically zero for 6-months lead time and longer. Also, the enhanced skill of MIMO-AE is noted for almost all initialization calendar months of the year (Figure 3h). Similar to E3SM, the Niño 3.4 index, ELI index and SVD1-SST demonstrate weak skill at all lead month lengths on monthly scales, although they are statistically different from zero for 1-month and 2-month lead times.

4 Summary and Discussion

In a novel approach, we apply MIMO-AE to extract the non-linear relationships between TP-SST and SC-PRECIP on monthly scales. Using LSTMs, we find MIMO-AE to be a powerful tool to enhance sub-seasonal to seasonal regional predictability that offers statistically significant improvements in predictive skill of SC-PRECIP up to a lead time of four months, as compared to that imparted by Niño 3.4 index, ELI index and SVD1-SST. The poor skill imparted by these indices is consistent with other studies (e.g. L'Heureux et al., 2021; Pan et al., 2019; S. Wang et al., 2017) that find poor skill from ENSO on noisier sub-seasonal time scales over Western US - largely due to atmospheric noise - in spite of significant correlations between SC-PRECIP and Niño 3.4 index on smoother seasonal and inter-annual time scales in observational data (e.g. Jong et al., 2016; X. Huang & Ullrich, 2017; G. Wang et al., 2021; Cheng et al., 2021, also Fig. 2b). Using the number of rainy days, with a threshold-based definition of rainy days, as a predictor of total precipitation, Cheng et al. (2021) found that ENSO can explain 20% (correlation skill of about 0.45) of the variability of winter season precipitation over Southern California. This is lower than the average skill of MIMO-AE upto a lead time of 3 months (about 0.6).

The underlying non-linear relationships, captured by MIMO-AE, are missed when variability over the tropical Pacific is represented by broad indices based on area-averages like Niño3.4 and linear transformations like SVDs. We find that a distinctive combination of SST anomalies over central, north-central, northeastern, eastern and northwestern tropical Pacific are associated with SC-PRECIP anomalies as illustrated by the MIMO-AE reconstructions during strong MIMO-AE defined events (Fig. 2c, Fig. S1). SC-PRECIP anomalies are driven by extra-tropical storm systems, jet streams as well as atmospheric rivers that are influenced by wave-trains emanating from the tropical Pacific forced by TP-SSTs. Fig. 2f shows the linear regression of vertically integrated moisture transport anomalies (IVT) - a proxy for processes impacting precipitation - against the standardized Niño3.4 index, ELI, SVD1-SST as well as against SC-PRECIP index for the 100 years of E3SM testing data. The weak IVT response associated with these indices indicate that they fail to adequately represent the processes associated with SC-PRECIP (Fig. 2f), and thus are poor predictors. Since, the spatial pattern of SST anomalies associated with MIMO-AE is not standing (Fig 2e), we use the principal component (PC1) of the leading EOF of MIMO-AE monthly SST reconstructions over the tropical Pacific as an index to represent MIMO-AE-derived SST anomaly pattern. The leading EOF explains almost all of the variability of the MIMO-AE SST reconstructions. A regression of IVT against the standardized PC1 reveals a strong IVT response over Southern California - about 70% of the moisture transport response associated with SC-PRECIP anomalies (Fig. 2f). This indicates that MIMO-AE learnt SST anomaly patterns, unlike those associated with linear approaches like SVD, strongly influence processes that result in SC-PRECIP anomalies, allowing MIMO-AE to provide enhanced predictive skill there.

Studies (e.g. L'Heureux et al., 2021; S. Wang et al., 2017; G. Wang et al., 2021; Cheng et al., 2021) suggest enhanced sub-seasonal and seasonal predictability of Western US precipitation from atmospheric variables; like geopotential heights, upper level zonal winds, moisture transport, etc.; as well as Northern Pacific SSTs. We plan to explore the benefits of including more variables in MIMO-AE framework. We also plan to explore techniques, like de-noising autoencoders, that systematically account for system noise - often associated with poor predictability of the climate system. Our MIMO-AE approach can also be applied to assess the predictability of regional climate across the globe, both where linear correlations are known to exist and where the signal to noise ratio is low. Expanding the precipitation domain from Southern California to all of the conterminous US, we find that MIMO-AE can explain a significant fraction of variability over other regions like the Southeastern US (Fig. S3). Additionally, our results demonstrate the promise of multi-task learning to enhance predictability afforded by remote teleconnections, supporting a focused exploration of other pertinent multi-task and multi-modal methods, like multi-task CNNs for such purposes, using the wealth of multi-model simulation ensembles.

5 Acknowledgements

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research.

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government

5.1 Open Research

The E3SMv1 data used in this study is freely available through the Earth System Grid Federation (ESGF) distributed archives via `https://doi.org/10.1029/2018MS001603` and is available through the ESGF interface `https://esgf-node.llnl.gov/projects/e3sm/` (E3SM Project, 2018).

Observational SST data from the HadISST 1.1 dataset (Rayner et al., 2003) can be downloaded from the web at `https://www.metoffice.gov.uk/hadobs/hadisst/`. Observed precipitation data from NOAA's PREC/L (M. Chen et al., 2002) can also be found open access at `https://psl.noaa.gov/data/gridded/data.precl.html`.

References

Bourlard, H., & Kamp, Y. (1988). Auto-association by multilayer perceptrons and singular value decomposition. *Biological cybernetics*, *59*(4), 291–294.

Bretherton, C. S., Smith, C., & Wallace, J. M. (1992). An intercomparison of methods for finding coupled patterns in climate data. *Journal of climate*, *5*(6), 541–560.

Broni-Bedaiko, C., Katsriku, F. A., Unemi, T., Atsumi, M., Abdulai, J.-D., Shinomiya, N., & Owusu, E. (2019). El niño-southern oscillation forecasting using complex networks analysis of lstm neural networks. *Artificial Life and Robotics*, *24*(4), 445–451.

Caruana, R. (1997). Multitask learning. *Machine learning*, *28*(1), 41–75.

Chang, P., Ji, L., & Li, H. (1997). A decadal climate variation in the tropical atlantic ocean from thermodynamic air-sea interactions. *Nature*, *385*(6616), 516–518.

Charte, D., Charte, F., García, S., del Jesus, M. J., & Herrera, F. (2018). A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines. *Information Fusion*, *44*, 78–96.

Chen, M., Xie, P., Janowiak, J. E., & Arkin, P. A. (2002). Global land precipitation: A 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology*, *3*(3), 249–266.

Chen, S., Zhang, Y., & Yang, Q. (2021). Multi-task learning in natural language processing: An overview. *arXiv preprint arXiv:2109.09138*.

Cheng, R., Novak, L., & Schneider, T. (2021). Predicting the interannual variability of california's total annual precipitation. *Geophysical Research Letters*, *48*(7), e2020GL091465.

Cohen, J., Pfeiffer, K., & Francis, J. (2017). Winter 2015/16: A turning point in enso-based seasonal forecasts. *Oceanography*. Retrieved from `https://doi.org/10.5670/oceanog.2017.115`

E3SM Project. (2018, April). *Energy Exascale Earth System Model (E3SM).* [Computer Software] `https://dx.doi.org/10.11578/E3SM/dc.20180418.36`. Retrieved from `https://dx.doi.org/10.11578/E3SM/dc.20180418.36` doi: 10.11578/E3SM/dc.20180418.36

Fournier, Q., & Aloise, D. (2019). Empirical comparison between autoencoders and traditional dimensionality reduction methods. In *2019 ieee second international conference on artificial intelligence and knowledge engineering (aike)* (pp. 211–214).

Frauen, C., Dommenget, D., Tyrrell, N., Rezny, M., & Wales, S. (2014). Analysis of the nonlinearity of el niño–southern oscillation teleconnections. *Journal of Climate*, *27*(16), 6225 - 6244. Retrieved from `https://journals.ametsoc.org/view/journals/clim/27/16/jcli-d-13-00757.1.xml` doi: 10.1175/JCLI-D-13-00757.1

Ghifary, M., Kleijn, W. B., Zhang, M., & Balduzzi, D. (2015). Domain generalization for object recognition with multi-task autoencoders. In *Proceedings of the ieee international conference on computer vision* (pp. 2551–2559).

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., … others (2019). The doe e3sm coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2089–2129.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1) (No. 2). MIT press Cambridge.

Gupta, M., Kodamana, H., & Sandeep, S. (2020). Prediction of enso beyond spring predictability barrier using deep convolutional lstm networks. *IEEE Geoscience and Remote Sensing Letters*.

Ham, Y.-G., Kim, J.-H., & Luo, J.-J. (2019). Deep learning for multi-year enso forecasts. *Nature*, *573*(7775), 568–572.

He, J., & Eastman, J. R. (2020). A sequential autoencoder for teleconnection analysis. *Remote Sensing*, *12*(5), 851.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.

Hoell, A., Hoerling, M., Eischeid, J., Wolter, K., Dole, R., Perlwitz, J., ... Cheng, L. (2016). Does el niño intensity matter for california precipitation? *Geophysical Research Letters*, *43*(2), 819–825.

Hoerling, M. P., Kumar, A., & Zhong, M. (1997). El niño, la niña, and the nonlinearity of their teleconnections. *Journal of Climate*, *10*(8), 1769–1786.

Huang, A., Vega-Westhoff, B., & Sriver, R. L. (2019). Analyzing el niño–southern oscillation predictability using long-short-term-memory models. *Earth and Space Science*, *6*(2), 212–221.

Huang, X., & Ullrich, P. A. (2017, 08). The Changing Character of Twenty-First-Century Precipitation over the Western United States in the Variable-Resolution CESM. *Journal of Climate*, *30*(18), 7555-7575. Retrieved from `https://doi.org/10.1175/JCLI-D-16-0673.1` doi: 10.1175/JCLI-D-16-0673.1

Jong, B.-T., Ting, M., & Seager, R. (2016). El niño's impact on california precipitation: Seasonality, regionality, and el niño intensity. *Environmental Research Letters*, *11*(5), 054021.

Kendall, A., Gal, Y., & Cipolla, R. (2018). Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 7482–7491).

Kumar, K. N., Ouarda, T. B., Sandeep, S., & Ajayamohan, R. S. (2016). Wintertime precipitation variability over the arabian peninsula and its relationship with enso in the cam4 simulations. *Climate dynamics*, *47*(7), 2443–2454.

Lee, S.-K., Lopez, H., Chung, E.-S., DiNezio, P., Yeh, S.-W., & Wittenberg, A. T. (2018). On the fragile relationship between el niño and california rainfall. *Geophysical Research Letters*, *45*(2), 907-915. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL076197` doi: https://doi.org/10.1002/2017GL076197

L'Heureux, M. L., Takahashi, K., Watkins, A. B., Barnston, A. G., Becker, E. J., Di Liberto, T. E., ... others (2017). Observing and predicting the 2015/16 el niño. *Bulletin of the American Meteorological Society*, *98*(7), 1363–1382.

L'Heureux, M. L., Tippett, M. K., & Becker, E. J. (2021). Sources of subseasonal skill and predictability in wintertime california precipitation forecasts. *Weather and Forecasting*, *36*(5), 1815 - 1826. Retrieved from `https://journals.ametsoc.org/view/journals/wefo/36/5/WAF-D-21-0061.1.xml` doi: 10.1175/WAF-D-21-0061.1

Liang, S., Chen, X., Ma, J., Du, W., & Ma, H. (2021). An improved double channel long short-term memory model for medical text classification. *Journal of healthcare engineering*, *2021*.

Livezey, R. E., & Chen, W. (1983). Statistical field significance and its determination by monte carlo techniques. *Mon. Wea. Rev*, *111*(1), 46–59.

Mahajan, S., North, G. R., Saravanan, R., & Genton, M. G. (2012). Statistical significance of trends in monthly heavy precipitation over the us. *Climate dynamics*, *38*(7), 1375–1387.

Mahajan, S., Saravanan, R., & Chang, P. (2010). Free and forced variability of the tropical atlantic ocean: Role of the wind–evaporation–sea surface temperature feedback. *Journal of Climate*, *23*(22), 5958 - 5977. Retrieved from `https://journals.ametsoc.org/view/journals/clim/23/22/2010jcli3304.1.xml` doi: 10.1175/2010JCLI3304.1

Mahajan, S., Tang, Q., Keen, N. D., Golaz, J.-C., & van Roekel, L. P. (2021). Simulation of enso teleconnections to precipitation extremes over the us in the high resolution version of e3sm. *Journal of Climate*, 1 - 62. Retrieved from `https://journals.ametsoc.org/view/journals/clim/aop/JCLI-D-20-1011.1/JCLI-D-20-1011.1.xml` doi: 10.1175/JCLI-D-20-1011.1

Masti, D., & Bemporad, A. (2018). Learning nonlinear state-space models using deep autoencoders. In *2018 ieee conference on decision and control (cdc)* (pp.

3862–3867).

Mu, B., Ma, S., Yuan, S., & Xu, H. (2020). Applying convolutional lstm network to predict el niño events: Transfer learning from the data of dynamical model and observation. In *2020 ieee 10th international conference on electronics information and emergency communication (iceiec)* (pp. 215–219).

Pan, B., Hsu, K., AghaKouchak, A., Sorooshian, S., & Higgins, W. (2019). Precipitation prediction skill for the west coast united states: From short to extended range. *Journal of Climate*, *32*(1), 161 - 182. Retrieved from `https://journals.ametsoc.org/view/journals/clim/32/1/jcli-d-18-0355.1.xml` doi: 10.1175/JCLI-D-18-0355.1

Patricola, C. M., O'Brien, J. P., Risser, M. D., Rhoades, A. M., O'Brien, T. A., Ullrich, P. A., ... Collins, W. D. (2020). Maximizing enso as a source of western us hydroclimate predictability. *Climate Dynamics*, *54*(1), 351–372.

Plaut, E. (2018). From principal subspaces to principal components with linear autoencoders. *arXiv preprint arXiv:1804.10253*.

Rayner, N., Parker, D., Folland, C., Horton, E., Alexander, L., & Rowell, D. (2003). The global sea-ice and sea surface temperature (hadisst) data sets. *J. Geophys. Res.*

Raza, S. E. A., Cheung, L., Epstein, D., Pelengaris, S., Khan, M., & Rajpoot, N. M. (2017). Mimo-net: A multi-input multi-output convolutional neural network for cell segmentation in fluorescence microscopy images. In *2017 ieee 14th international symposium on biomedical imaging (isbi 2017)* (pp. 337–340).

Tang, Y., & Hsieh, W. W. (2003). Nonlinear modes of decadal and interannual variability of the subsurface thermal structure in the pacific ocean. *Journal of Geophysical Research: Oceans*, *108*(C3). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JC001236` doi: https://doi.org/10.1029/2001JC001236

Toshniwal, S., Tang, H., Lu, L., & Livescu, K. (2017). Multitask learning with low-level auxiliary tasks for encoder-decoder based speech recognition. *arXiv preprint arXiv:1704.01631*.

Trenberth, K. E., & Stepaniak, D. P. (2001). Indices of el niño evolution. *Journal of climate*, *14*(8), 1697–1701.

Wang, G., Zhuang, Y., Fu, R., Zhao, S., & Wang, H. (2021). Improving seasonal prediction of california winter precipitation using canonical correlation analysis. *Journal of Geophysical Research: Atmospheres*, *126*(17), e2021JD034848.

Wang, S., Anichowski, A., Tippett, M. K., & Sobel, A. H. (2017). Seasonal noise versus subseasonal signal: Forecasts of california precipitation during the unusual winters of 2015–2016 and 2016–2017. *Geophysical Research Letters*, *44*(18), 9513–9520.

Wang, Y., Yao, H., & Zhao, S. (2016). Auto-encoder based dimensionality reduction. *Neurocomputing*, *184*, 232–242.

Williams, I. N., & Patricola, C. M. (2018). Diversity of enso events unified by convective threshold sea surface temperature: a nonlinear enso index. *Geophysical Research Letters*, *45*(17), 9236–9244.

Zamparo, L., & Zhang, Z. (2015). Deep autoencoders for dimensionality reduction of high-content screening data. *arXiv preprint arXiv:1501.01348*.

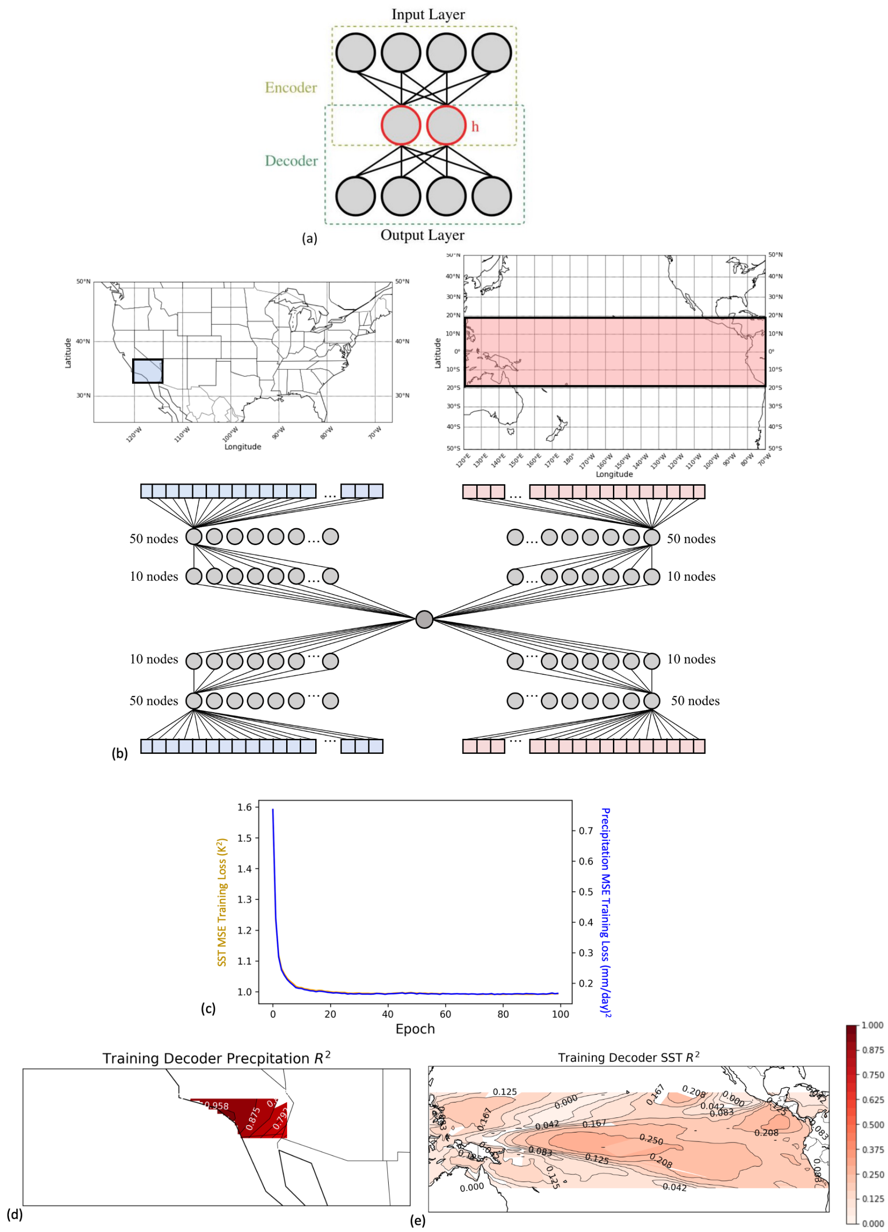


Figure 1. Network architecture. A simple autoencoder architecture (a), MIMO-AE architecture (b), training losses for MIMO-AE over 100 epochs using scaled data (c), and the average R^2 between the MIMO-AE reconstructed and original input data for Southern California precipitation (d) and Tropical Pacific SST (e). We note that the connections between neurons in (b) are shown selectively for clarity and in actuality all neurons are connected with every neuron in the adjacent layers.

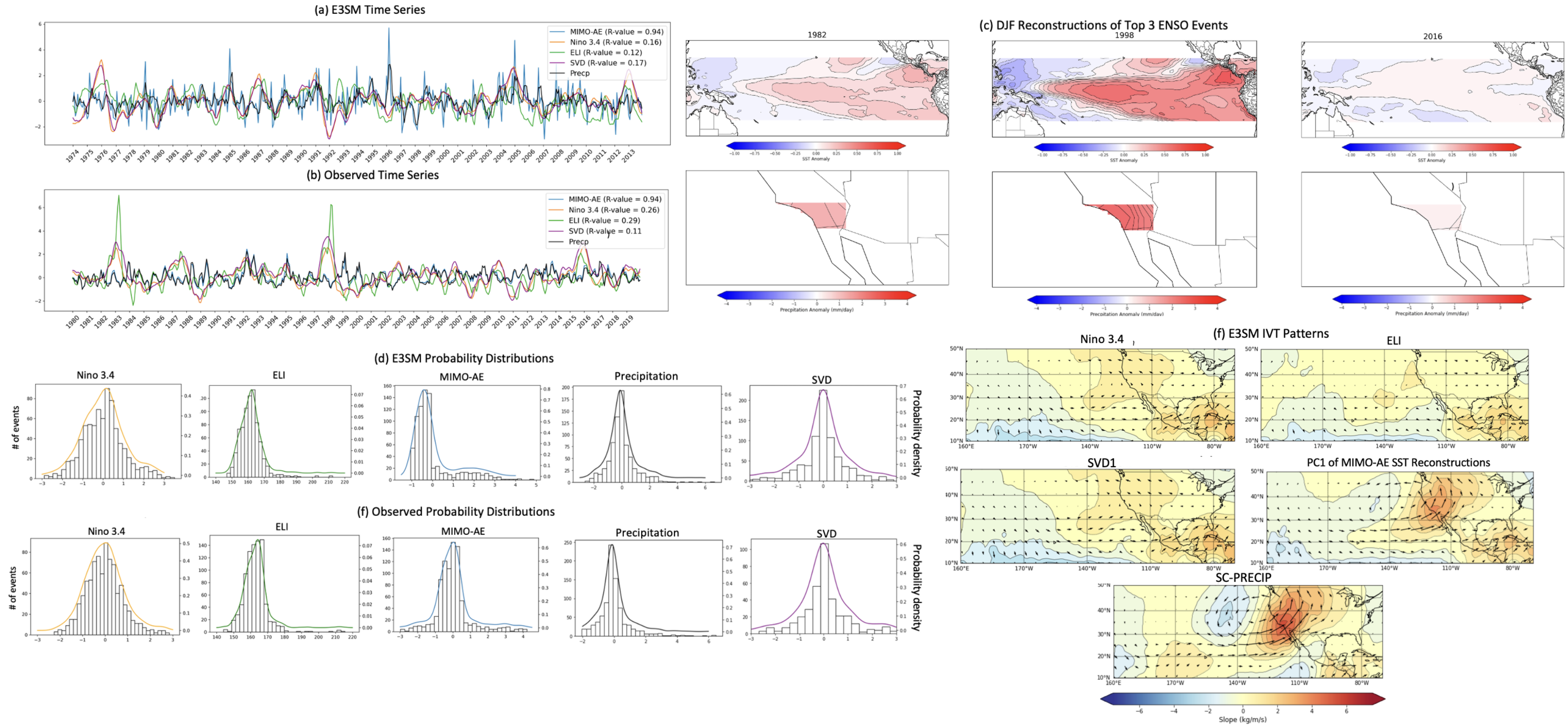


Figure 2. Three-month moving average of the standardized time series of MIMO-AE index (blue), Niño 3.4 index (orange), ELI index (green), expansion coefficients of the first leading mode of SVD (SVD1) and domain averaged SC-PRECIP (black) for a segment (last 40 years, 1974-2013) of the 65 years of testing data for the E3SM historical simulation (a); and a segment (last 40 years, 1980-2019) of the observed testing data (b). Correlations of each time-series against domain averaged SC-PRECIP is also listed. The December to February average reconstructions for the three strong El-Niño events of 1981-82, 1997-98, and 2015-16 in observations (c). The probability density distributions for the Niño 3.4 index (orange), ELI (green), MIMO-AE index (blue), SVD1 time-series (purple) and SC-PRECIP (black) for E3SM data (d); and observational data (e). Regression of the vertically integrated moisture transport (IVT) vectors and magnitudes against standardized Niño 3.4 index, ELI, SVD1 time-series and MIMO-AE index in E3SM testing data (f).

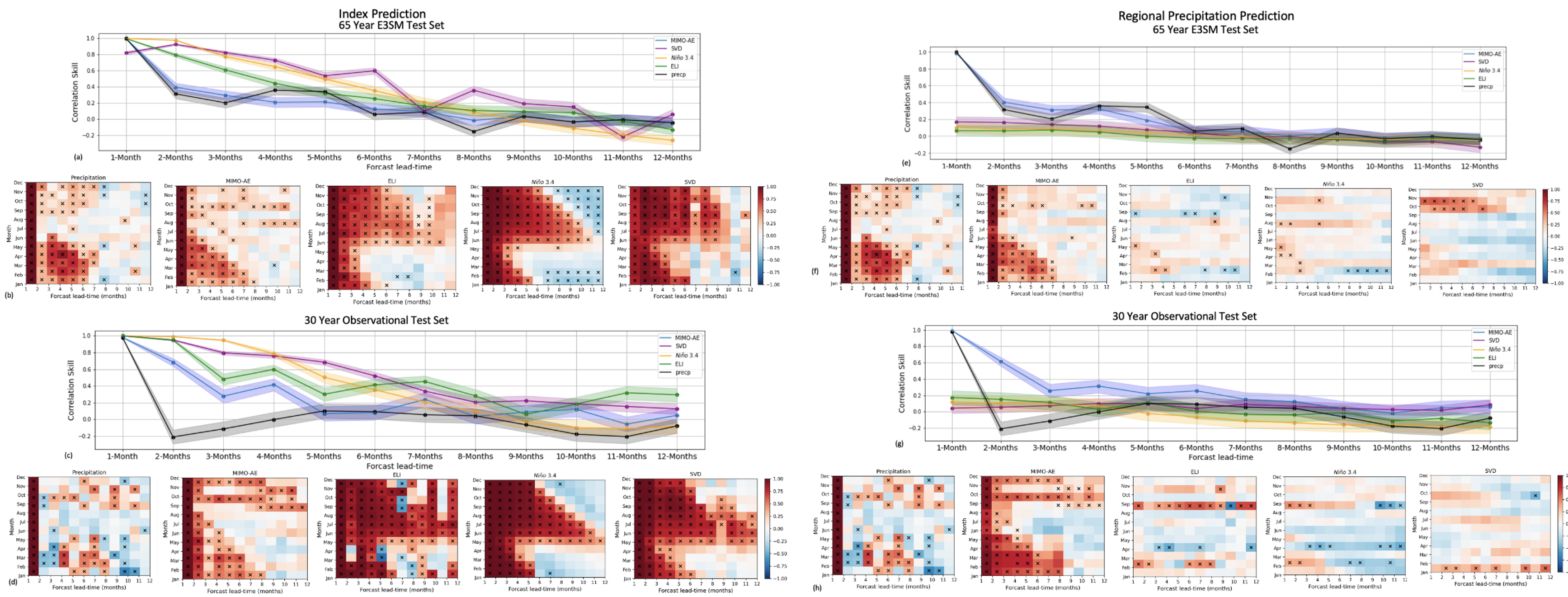


Figure 3. Predictability of MIMO-AE index. Predictive skill of LSTMs of MIMO-AE index (blue), Niño 3.4 index (orange), ELI (green), SVD1 (purple) and domain averaged SC-PRECIP (black) at forecast lead times of 1 to 12 months for E3SM testing data (a); and observational testing data (c). Shading represents one standard deviation of the correlation coefficients. Predictive skill of LSTMs as a function of initialization calendar month and forecast lead time from domain average SC-PRECIP, MIMO-AE index, ELI, SVD1 and Niño 3.4 index for E3SM testing data (b); and observational testing data (d). Predictability of SC-PRECIP using MIMO-AE. Predictive skill of MIMO-AE index (blue), Niño 3.4 index (orange), SVD1 (purple) and ELI (green) at predicting domain averaged SC-PRECIP at forecast lead times of 1 to 12 months for E3SM testing data (e); and observational testing data (g). Shading represents one standard deviation of the correlation coefficients. Predictive skill of domain average SC-PRECIP as a function of initialization calendar month and forecast lead time from domain average SC-PRECIP, MIMO-AE index, ELI, SVD1 and Niño 3.4 index for E3SM testing data (f); and observational testing data (h). Cross markings in b,d,f and h indicate values that are statistically different from zero at the 95% confidence level.

Figure 1.

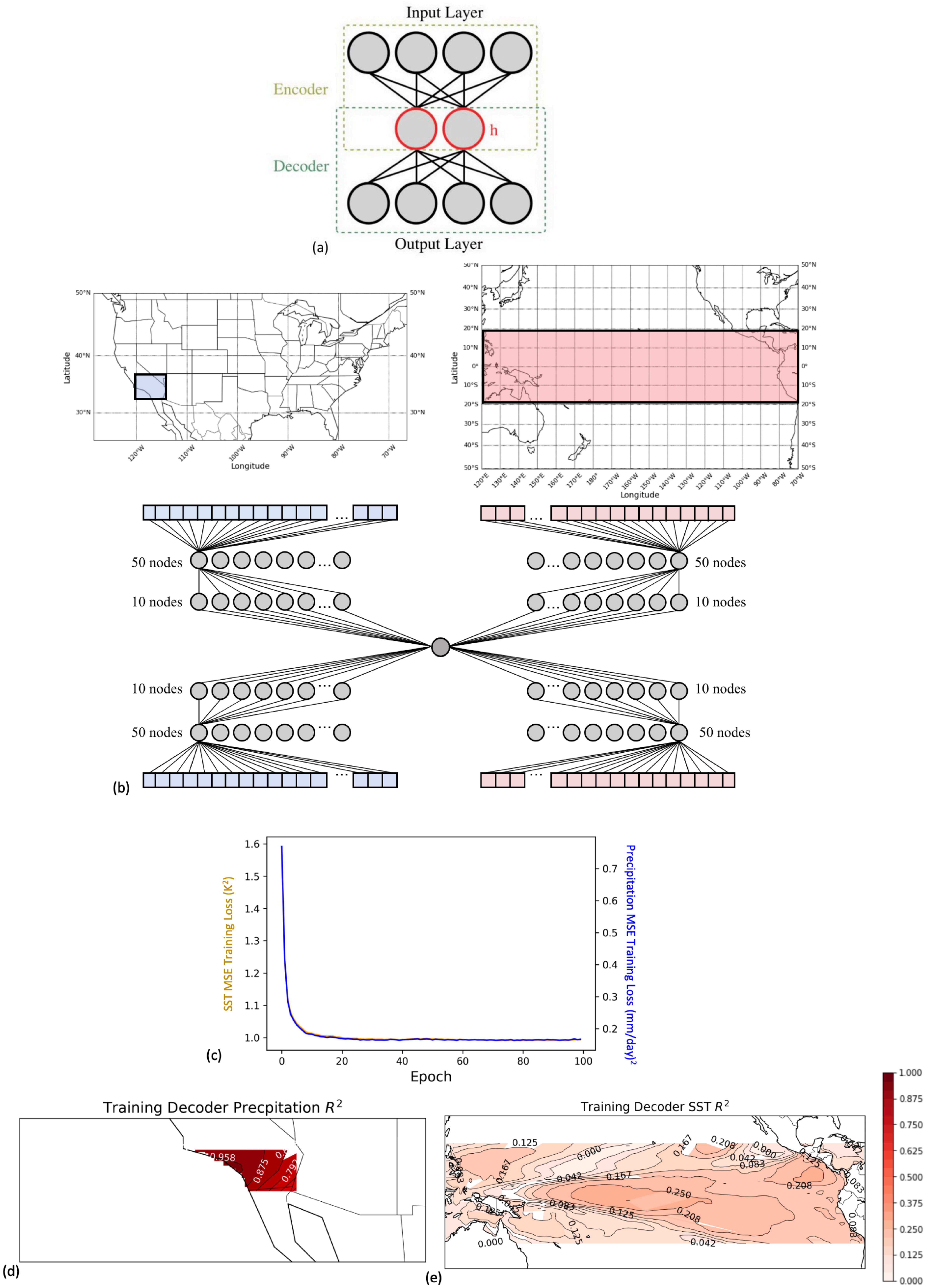
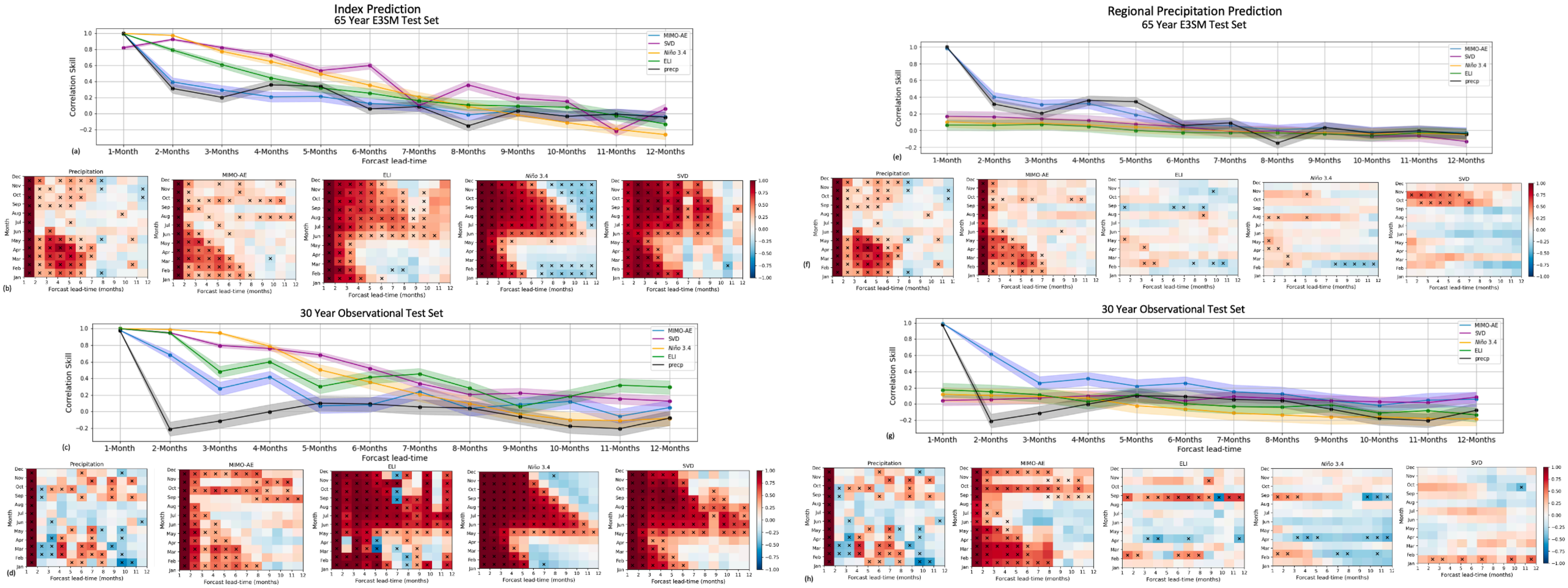
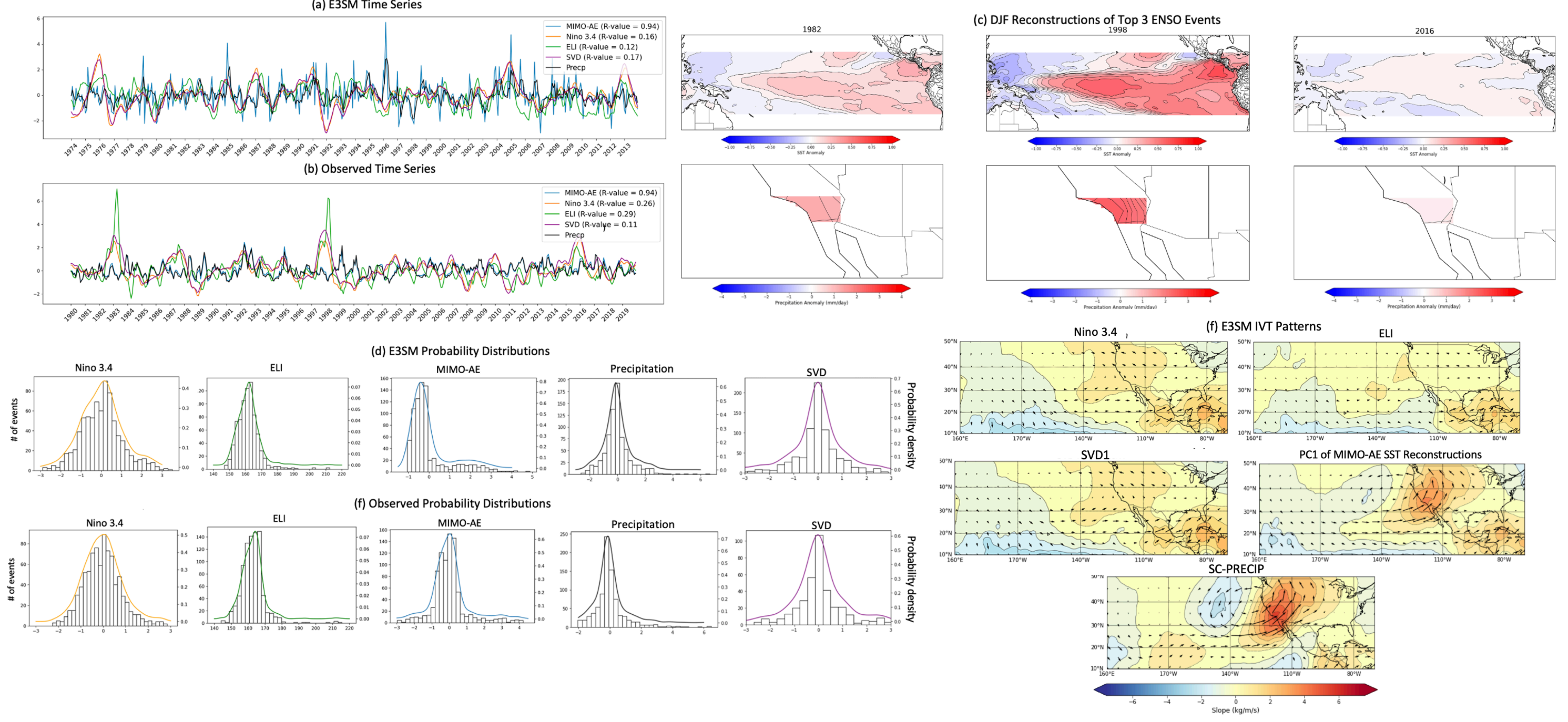

Input Layer
Encoder
h
Decoder
Output Layer
(a)
Latitude
Longitude
50 nodes
10 nodes
10 nodes
50 nodes
(b)
SST MSE Training Loss (K²)
Precipitation MSE Training Loss (mm/day)²
Epoch
(c)
Training Decoder Precpitation R^2
(d)
Training Decoder SST R^2
(e)

Figure 2.



Index Prediction
65 Year E3SM Test Set
30 Year Observational Test Set
Regional Precipitation Prediction
65 Year E3SM Test Set
30 Year Observational Test Set
Correlation Skill
Forcast lead-time
Forcast lead-time (months)
MIMO-AE
SVD
Niño 3.4
ELI
precp
Precipitation
Month
1-Month
2-Months
3-Months
4-Months
5-Months
6-Months
7-Months
8-Months
9-Months
10-Months
11-Months
12-Months
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Figure 3.


(a) E3SM Time Series
MIMO-AE (R-value = 0.94)
Nino 3.4 (R-value = 0.16)
ELI (R-value = 0.12)
SVD (R-value = 0.17)
Precp
(b) Observed Time Series
MIMO-AE (R-value = 0.94)
Nino 3.4 (R-value = 0.26)
ELI (R-value = 0.29)
SVD (R-value = 0.11)
Precp
(c) DJF Reconstructions of Top 3 ENSO Events
1982
1998
2016
SST Anomaly
Precipitation Anomaly (mm/day)
(d) E3SM Probability Distributions
Nino 3.4
ELI
MIMO-AE
Precipitation
SVD
# of events
Probability density
(f) Observed Probability Distributions
Nino 3.4
ELI
MIMO-AE
Precipitation
SVD
# of events
Probability density
(f) E3SM IVT Patterns
Nino 3.4
ELI
SVD1
PC1 of MIMO-AE SST Reconstructions
SC-PRECIP
Slope (kg/m/s)