A multi-modal representation of ENSO Diversity

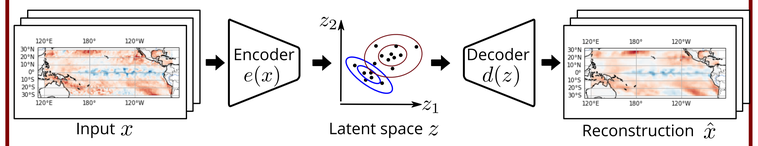
Jakob Schlör[1], Felix Strnad[1], Antonietta Capotondi[2], Bedartha Goswami[1]

[1]Cluster of Excellence - ML for Sciences, University of Tübingen, Germany
[2]Physical Sciences Laboratory, NOAA, Boulder, Colorado, USA

1. Motivation

- El-Niño Southern Oscillation (ENSO) shows strong event-to-event variability, known as **ENSO diversity**[1]
- El Niño and La Niña events are typically separated into Eastern Pacific (EP) and Central Pacific (CP) type
- The low-dimensional representation of SSTA suggests that events rather **form a continuum with regions of higher density**.
- We suggest a probabilistic unsupervised clustering to approximate ENSO diversity by a **multi-modal Gaussian distribution**.

2. Method



a. Low-dimensional representation

Linear ⟷ Emprical Orthogonal Functions (EOF)

Nonlinear ⟷ Autoencoder (AE) Neural Network[2]

b. Unsupervised "soft" clustering

Gaussian mixture model (GMM)[3] approximates the low-dimensional distribution:

$$p(\mathbf{Z}) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\mathbf{z} \mid \boldsymbol{\mu}_k, \Sigma_k\right)$$

$\boldsymbol{\mu}_k$: mean
Σ_k : covariance
π_k : probability of Gaussian

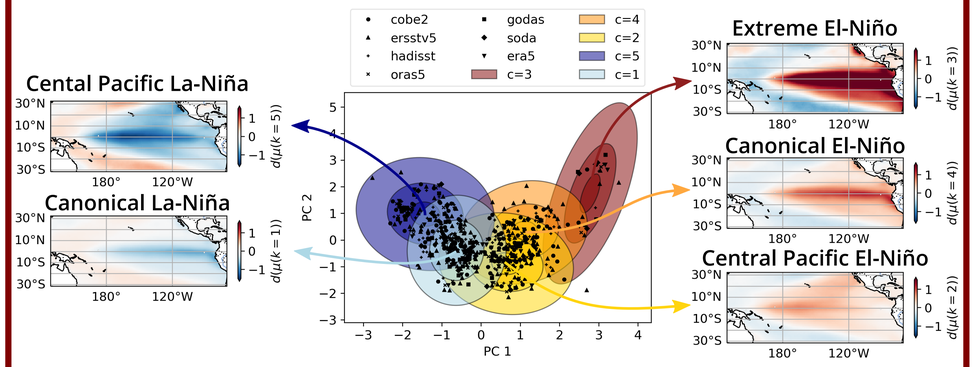
The optimal number of clusters is obtained by minimizing the Bayesian Information Criterion (BIC).

3. Data

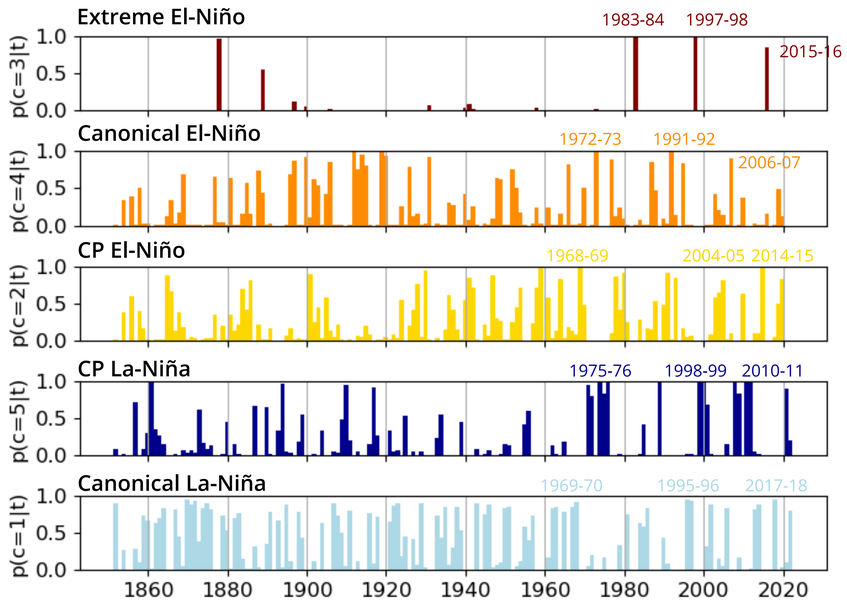
- Monthly Sea Surface Temperature Anomalies (SSTA) from 7 reanalysis datasets : (ERSSTv5, ERA5, COBE-SST, HadISST, ORAS5, SODA3.12, GODAS).
- Pacific region (130◦E - 70 ◦W, 30◦S - 30◦N) interpolated on a 1°×1° resolution
- Select El-Niño/La-Niña by SSTA in the Nino3.4 region when DJF average is larger/smaller than 0.5 K/-0.5 K.

4. Results

4.1. Linear decomposition

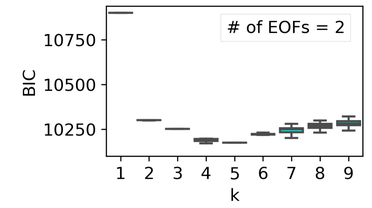


- Latent space is optimally described by five Gaussians
- El-Niño events by three Gaussians and La-Niña events by two
- Gaussians overlap which describes the continuity of ENSO Diversity

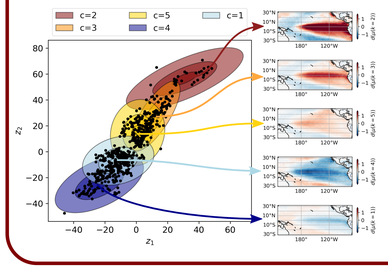


Conditional probability of each event to belong to each cluster.

4.3. Robustness of GMM

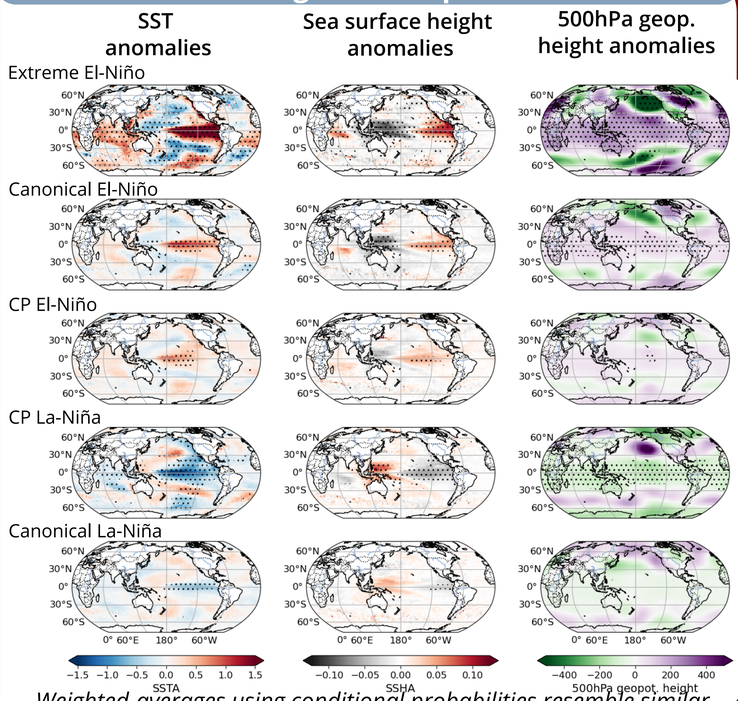


The optimal number of clusters, k, are obtained by minimizing BIC of the EOF-GMM.



Nonlinear decomposition, AE-GMM, yield qualitatively similar clusters than the linear decomposition.

4.2. Probabilistic weighted composites



Weighted-averages using conditional probabilities resemble similar patterns than averages over EP and CP events which are classified by the classical indices.

References:
[1] Capotondi et al., ENSO Book (2020)
[2] Kramer, AIChE J. (1991)
[3] McLachlan, Peel; Finite Mixture Models (2000)

jakob.schloer@uni-tuebingen.de
@schloer_jakob