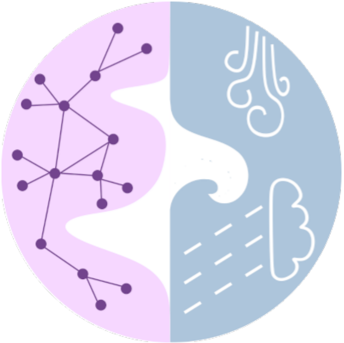
A multi-modal representation of El Nino Southern Oscillation Diversity



Jakob Schlör, Bedartha Goswami

Cluster of Excellence - Machine Learning for Science, University of Tübingen

Climate data

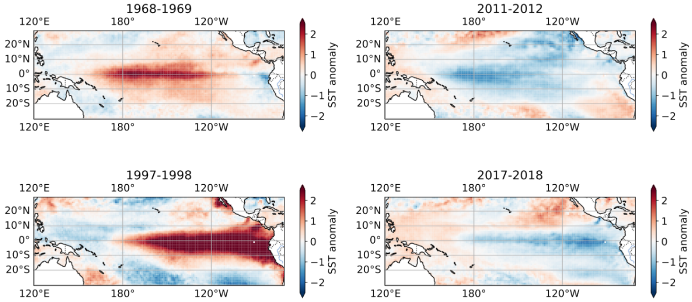


Figure: Sea surface temperature anomalies from ERA5

- Observational climate data are high-dimensional
- Goal: Reduction of information to essential features

Climate data

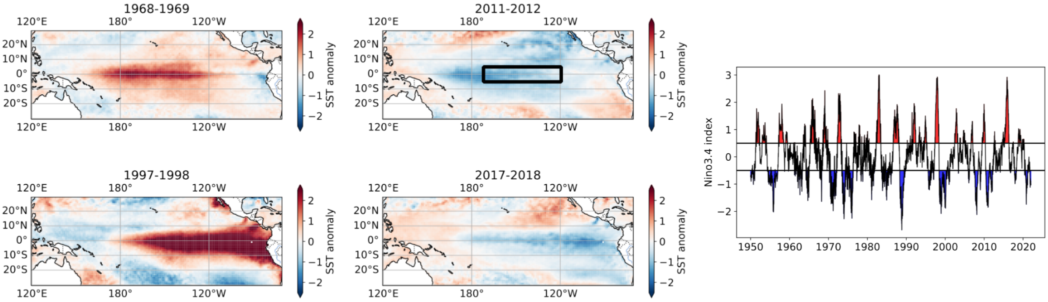
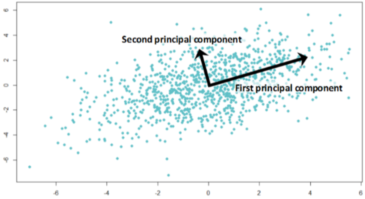


Figure: Sea surface temperature anomalies

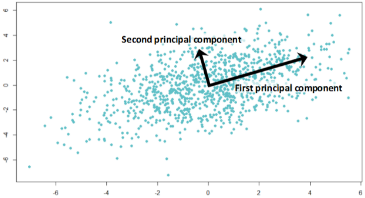
- Observational climate data are high-dimensional
- Goal: Reduction of information to essential features

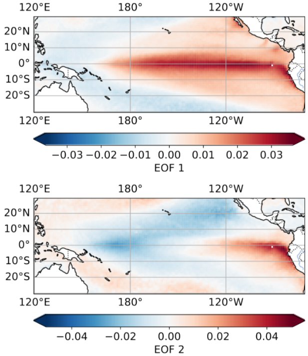
Recap: Empirical orthogonal function analysis

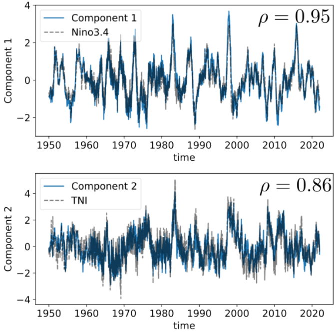


- Linear decomposition into components that explain most variance
- SSTA data reduction from 10 000 to 2 dimensions

Recap: Empirical orthogonal function analysis

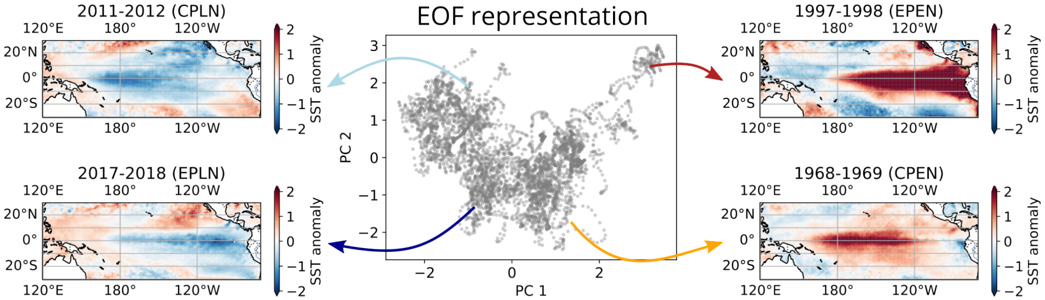






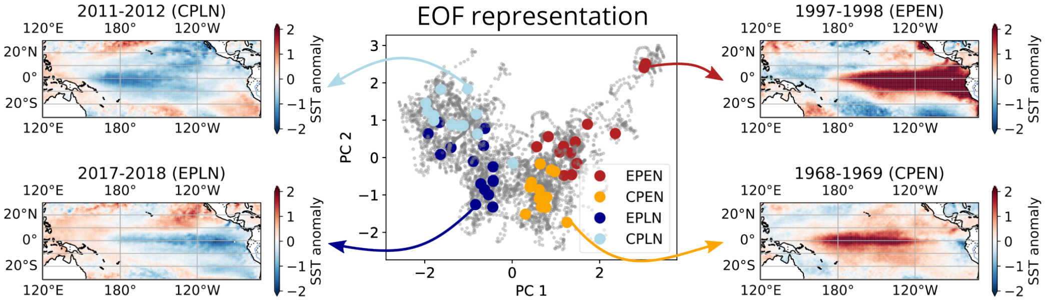
- Linear decomposition into components that explain most variance
- SSTA data reduction from 10 000 to 2 dimensions explaining $\approx$ **32%** of variance

ENSO Diversity in EOF space



ENSO Diversity in EOF space

EBERHARD KARLS UNIVERSITÄT TÜBINGEN



Spatial diversity of ENSO events are classified into [1]

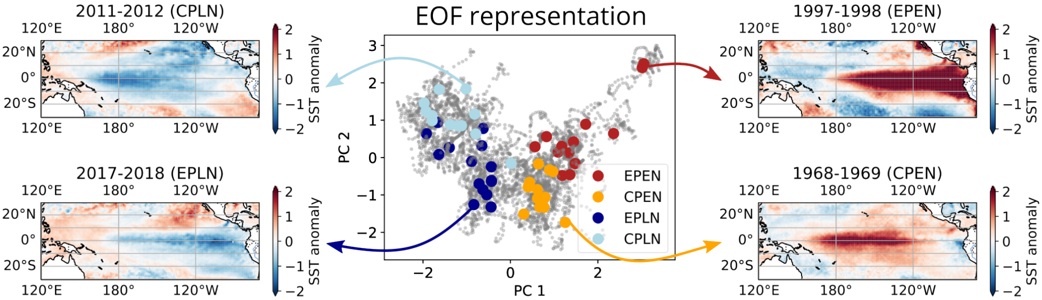
EP: Eastern Pacific Niño3 > Niño4

CP: Central Pacific Niño3 < Niño4

[1] Ashok et al. 2007, Kao et al. 2009

ENSO Diversity in EOF space

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

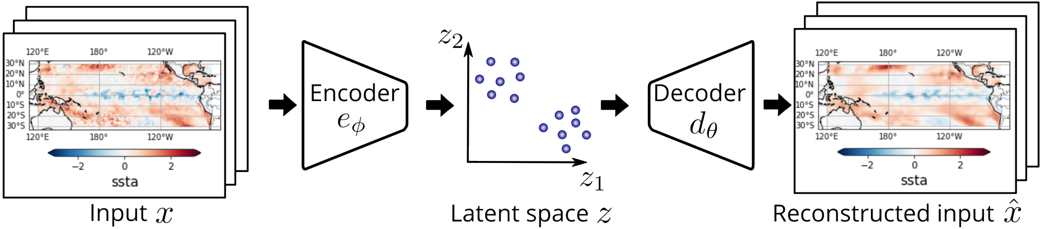


Is the classification a good description of the spatial variability of ENSO?

Dimensionality reduction

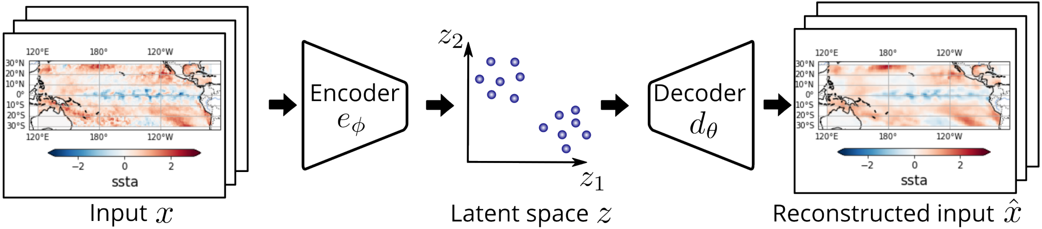
A machine learning perspective

Autoencoder (AE):



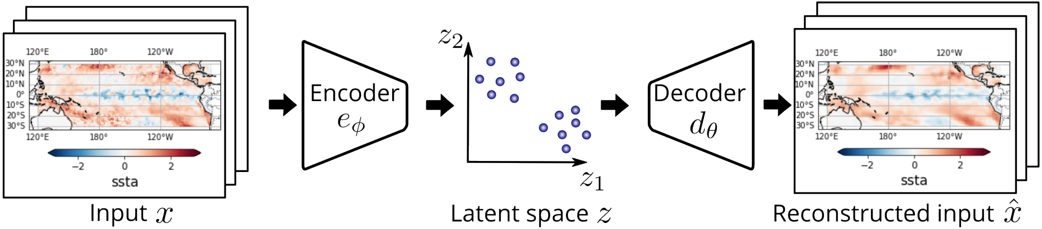
- Non linear dimensionality reduction

Autoencoder (AE):



- Non linear dimensionality reduction
- The encoder and decoder functions are Neural Networks

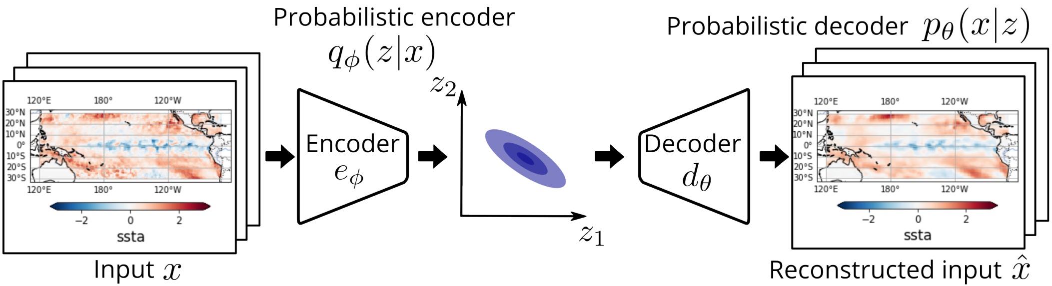
Autoencoder (AE):



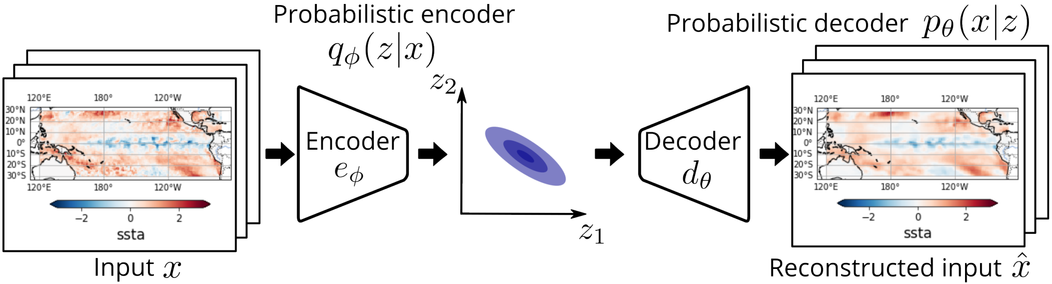
- Non linear dimensionality reduction
- The encoder and decoder functions are Neural Networks
- The network parameters θ and ϕ try to minimize the reconstruction error:

$$L_{AE}(\theta, \phi) = \frac{1}{n} \sum_{i=1}^{n} \left(\mathbf{x}^{(i)} - d_\theta \left(e_\phi \left(\mathbf{x}^{(i)} \right) \right) \right)^2$$

Variational Autoencoder (VAE):



Variational Autoencoder (VAE):



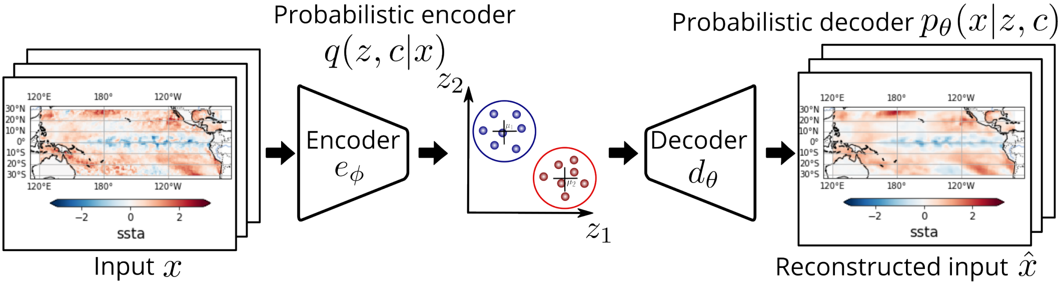
- Approximate the true data distribution:

$$x \sim p(x) = \int dz p(x|z)p(z)$$

- Continous and smooth latent space

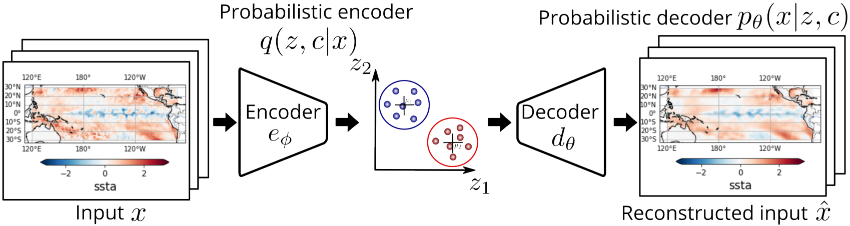
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Gaussian mixture variational Autoencoder (GMVAE):



- Assume: Data are generated from a mixture of Gaussians
- Unsupervised "soft" clustering in the latent space

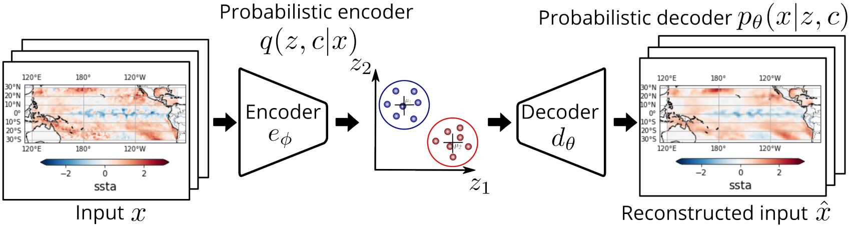
GMVAE on Pacific SSTA



Data:

- ERA5 (1950-2021), detrended and subtracted from daily climatology
- Tropical Pacific **[120**°E, **90**°W**]** and **[30**°S, **30**°N**]** on a 1°x 1° resolution
- Select boreal winters (DJF) where average Nino3.4 $\leq -\mathbf{0.5}$ K or Nino3.4 $\geq \mathbf{0.5}$ K

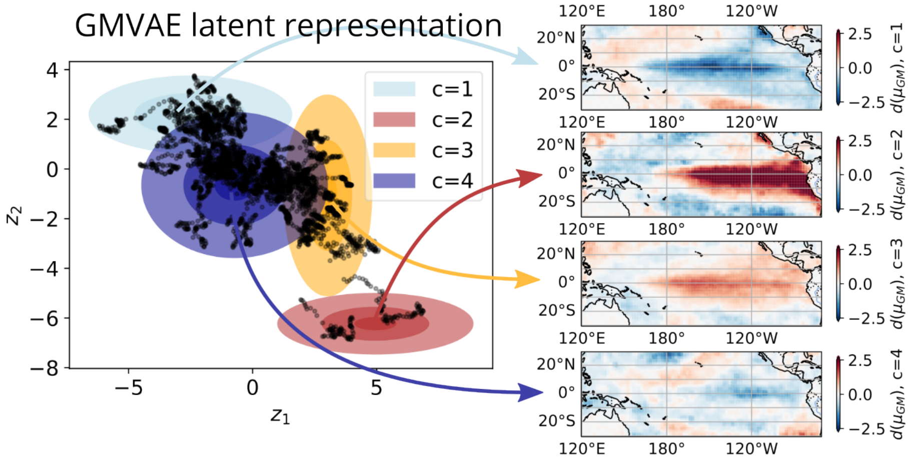
GMVAE on Pacific SSTA



Computation:

- 3-layer convolutional NN for encoder and decoder architecture
- 4 Gaussian mixtures randomly initialized

GMVAE on Pacific SSTA

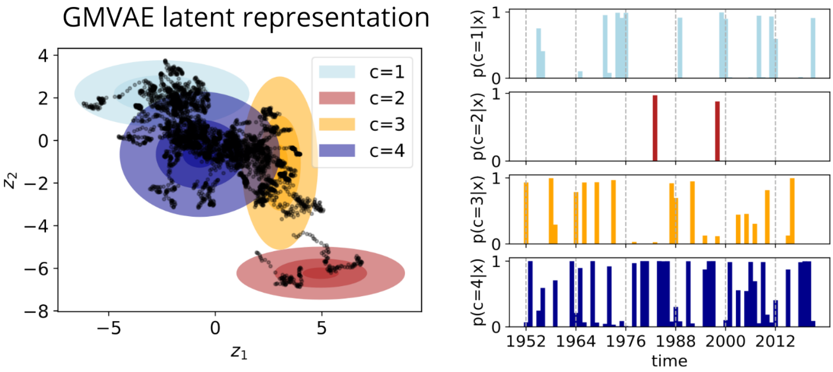


- Cluster mean capture known extreme SSTA patterns
- El Niños are rather separated by their intensity

GMVAE allows soft-clustering

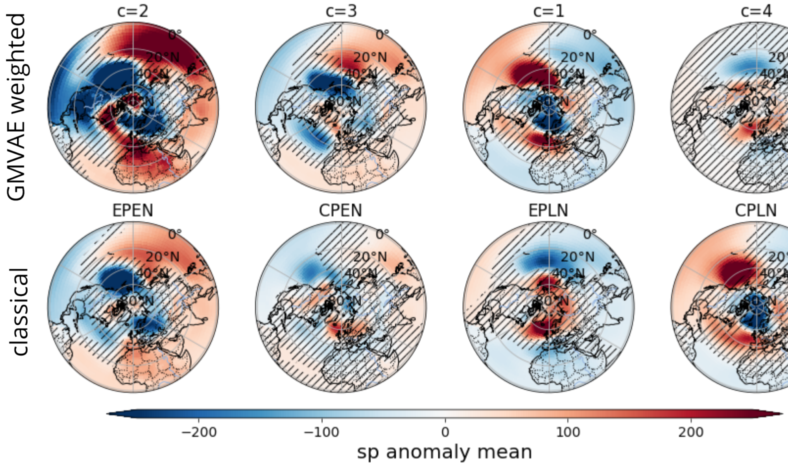
Assign a probability to each day $p(c_k|x_t)$:

⇒ **Soft-clustering**



GMVAE composites

Probabilities $p(c_k|x_t)$ allows to compute composites:



Summary

Conclusion:

- Soft clustering represents continuity of ENSO events
- Non-linear decomposition does not separate EP and CP
- Probabilistic weighting allow more expressive composites

Summary

Conclusion:

- Soft clustering represents continuity of ENSO events
- Non-linear decomposition does not separate EP and CP
- Probabilistic weighting allow more expressive composites

Outlook:

- Include other variables, e.g. thermocline depth
- Time-evolution of events
- How exactly do composites change?

Thank you!

Figure: Felix Strnad, Bedartha Goswami

https://machineclimate.de/
@MachineClimate

GMVAE: Some maths

The generative process:

$$p(\mathbf{x}, \mathbf{z}, c) = p(\mathbf{x} \mid \mathbf{z})p(\mathbf{z} \mid c)p(c)$$

Maximizing the Evidence lower bound (ELBO):

$$\mathcal{L}_{\text{ELBO}}(\theta, \phi, \psi) = \langle \log p_\theta(\mathbf{x} \mid \mathbf{z}) + \log p(\mathbf{z} \mid c) + \log p(c) - \log q_\phi(\mathbf{z} \mid \mathbf{x}) - \log p(c \mid \mathbf{z}) \rangle_{q(\mathbf{z}, c \mid \mathbf{x})}$$