Australia's National Science Agency

A deep learning model for forecasting global monthly mean sea surface temperature anomalies

Ming Feng | 29 November 2023

John Taylor[§], **Zeya Li**, Foivos Diakogiannis[§]

CSIRO Environment

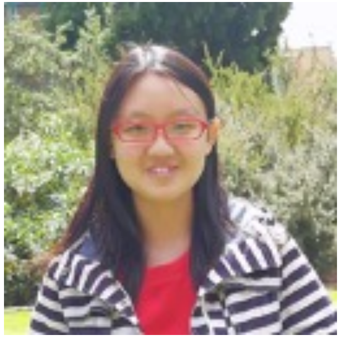
[§]Data61, CSIRO

Forum for Operational Oceanography

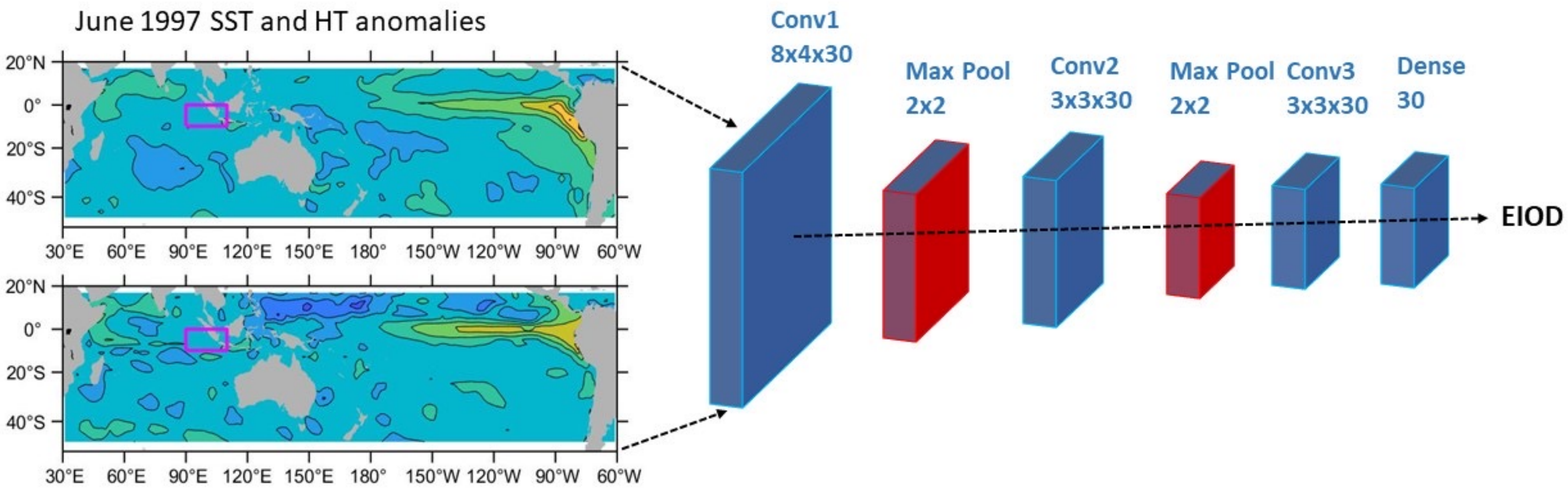
Outline of the talk

- Our previous machine-learning approaches
  - Convolutional neural network model
  - A Unet model for global SST predictions

- Zeya Li's postdoc project – Unlocking the predictability of marine heatwaves using AI techniques
  - Transformer based machine learning



Convolutional Neural Network (CNN) model



3-month SST and heat content anomaly maps

Issues:
Need long training data ~1000s years
Need to construct models for different target regions

Feng et al. 2022
Boschetti et al. 2023

ORIGINAL RESEARCH article

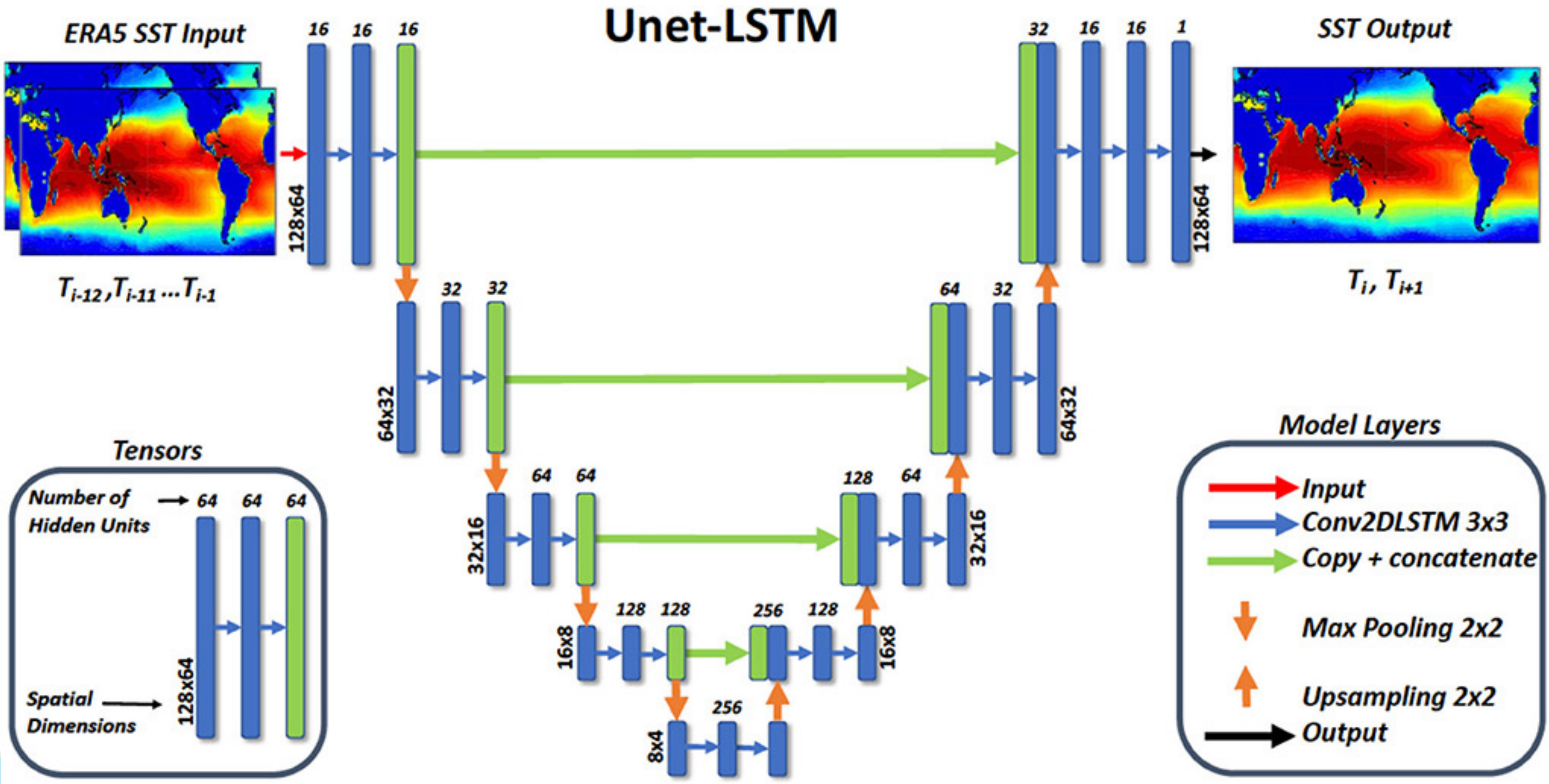
Front. Clim., 28 September 2022
Sec. Climate, Ecology and People
https://doi.org/10.3389/fclim.2022.932932

This article is part of the Research Topic
Advances in Marine Heatwave Interactions
View all 16 Articles >

A deep learning model for forecasting global monthly mean sea surface temperature anomalies

 John Taylor[1*†] and  Ming Feng[2,3†]

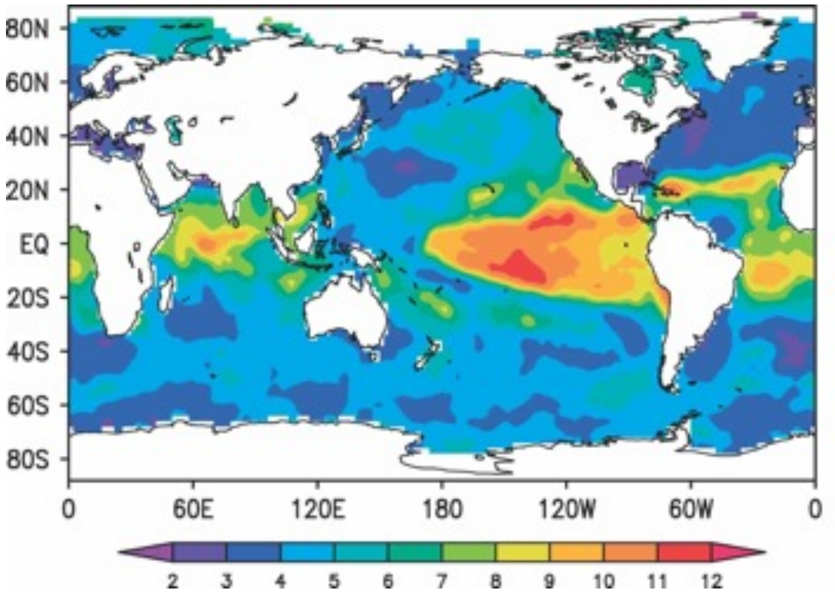
[1] Data61, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra, ACT, Australia
[2] Oceans and Atmosphere, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Perth, WA, Australia
[3] Centre for Southern Hemisphere Oceans Research, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Hobart, TAS, Australia


Unet-LSTM
ERA5 SST Input
$T_{i-12}, T_{i-11} \dots T_{i-1}$
SST Output
T_i, T_{i+1}
16 16 16
128x64
32 32
64x32
64 64
32x16
128 128
16x8
256
8x4
256 128
128 64
64 32
32 16 16 1
Tensors
Number of Hidden Units
64 64 64
Spatial Dimensions
128x64
Model Layers
Input
Conv2DLSTM 3x3
Copy + concatenate
Max Pooling 2x2
Upsampling 2x2
Output

Input/output of the Unet-LSTM model

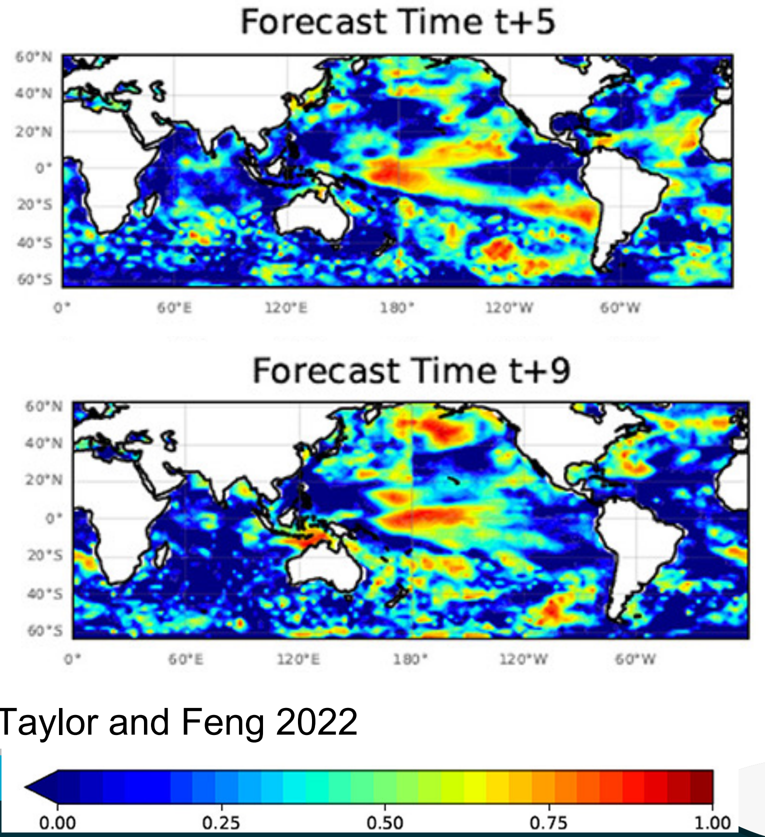
- ERA5 SST and air temperature data (from 1958)
- Map SST (Ta_2m) onto **[64,128] grid,** (64°S to 62°N in 2° increments; 180°W to 180°E in 2.8125° increments)
- Input data **uses 12 monthly steps**
- The model is trained to make **2-month predictions**, and recursively to predict the next 2 months

Predictability limit of monthly sea SST temperature in the global oceans



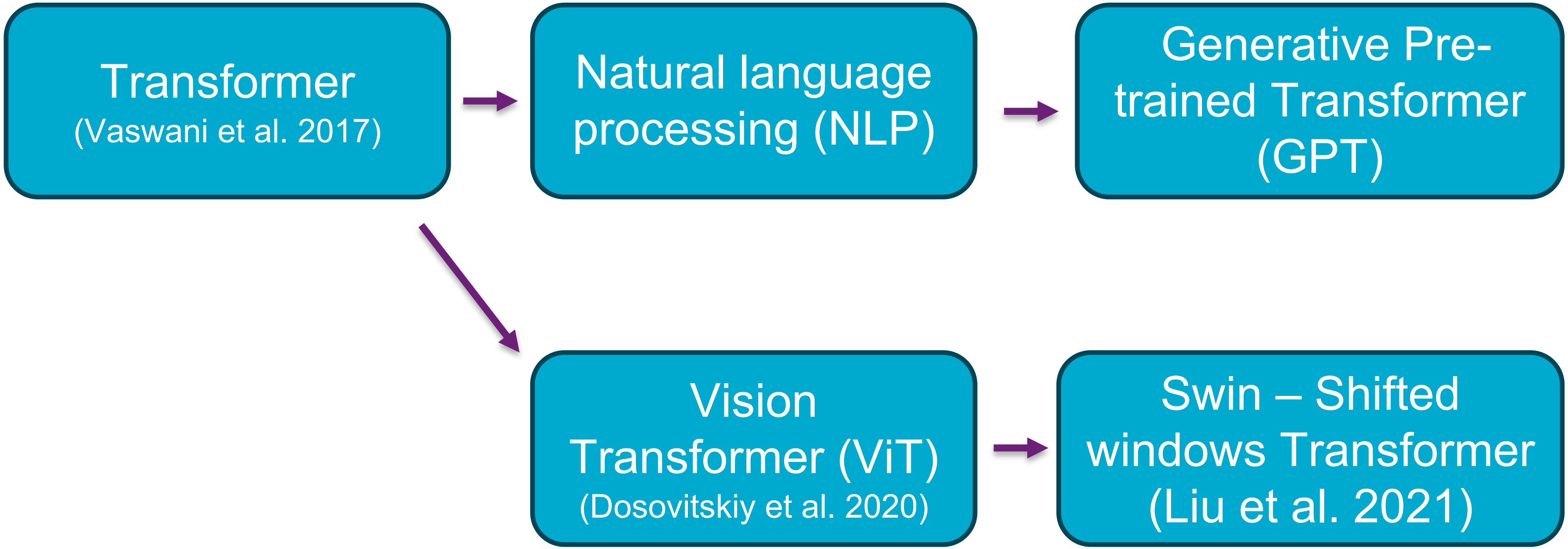
Li and Ding 2012

Nonlinear local Lyapunov exponent



Taylor and Feng 2022

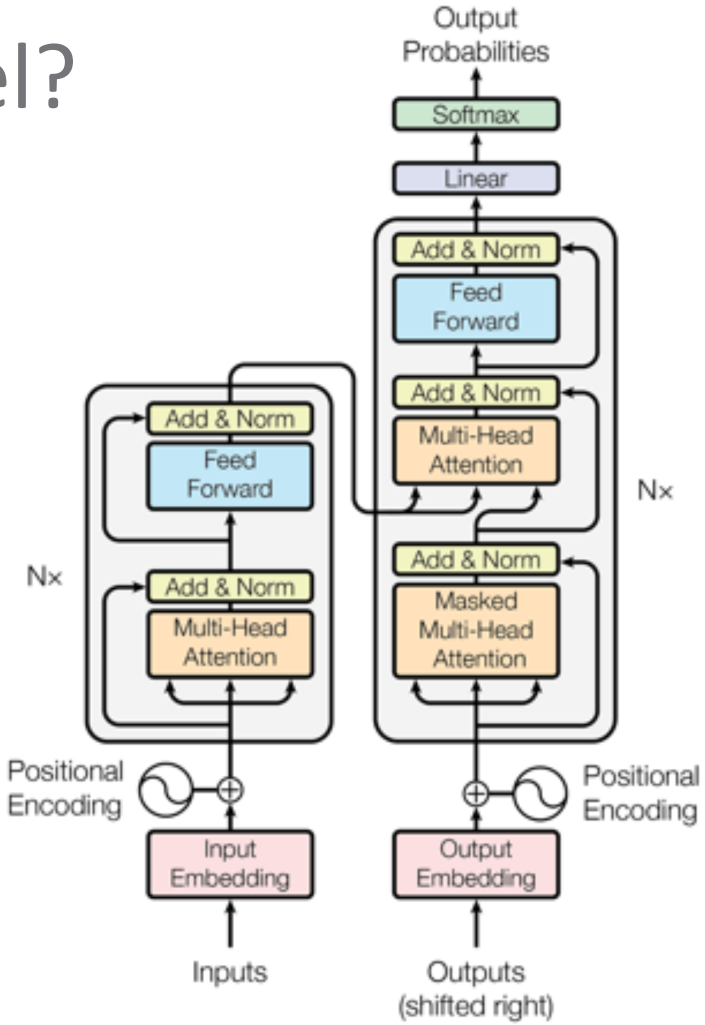
Transformer Model Architecture



What is a transformer model?

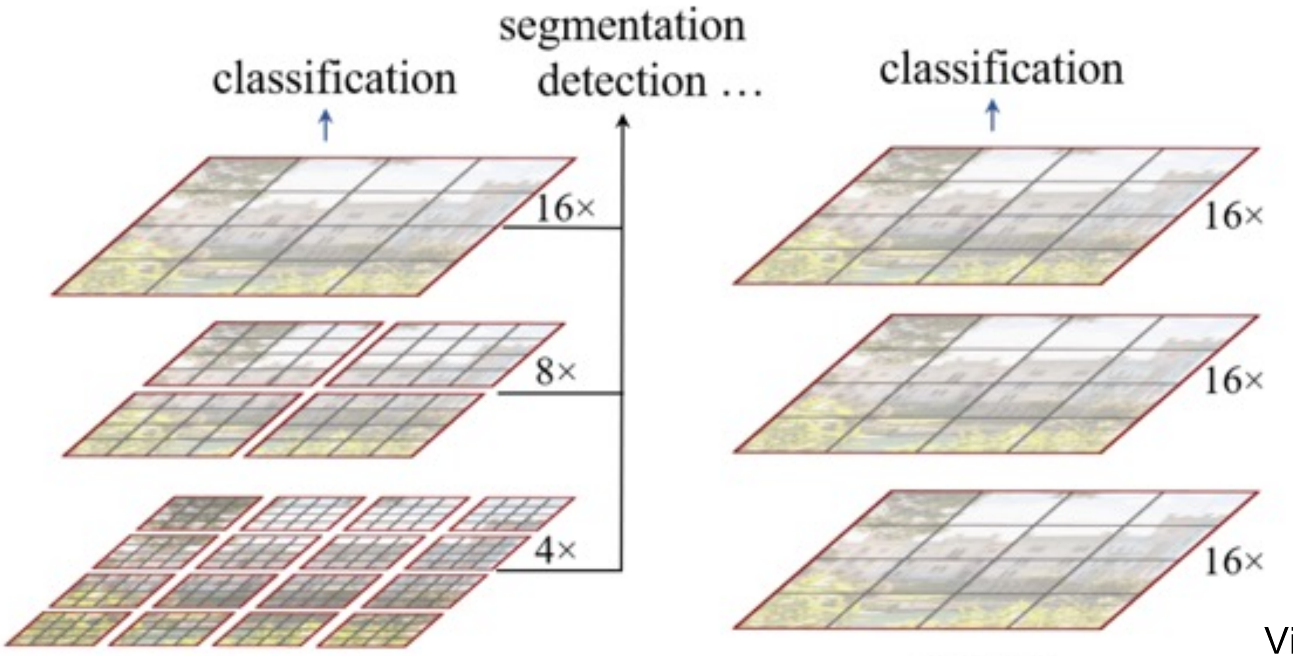
A neural network that learns context and thus meaning by tracking relationships in sequential data (like words in a sentence)

- Self-attention mechanism
- Establish multivariable relationships in parallel regardless of their spatial and temporal distances



Swin (shifted window) Transformer

Swin-Tunet – John Taylor



ViT – Vision Transformer
Dosovitskiy et al. 2021

Limit self-attention computation to non-overlapping local windows while also allowing for cross-window connection.

Liu et al. 2021
Microsoft Asia

Transformer model set up for FOO (Zeya)

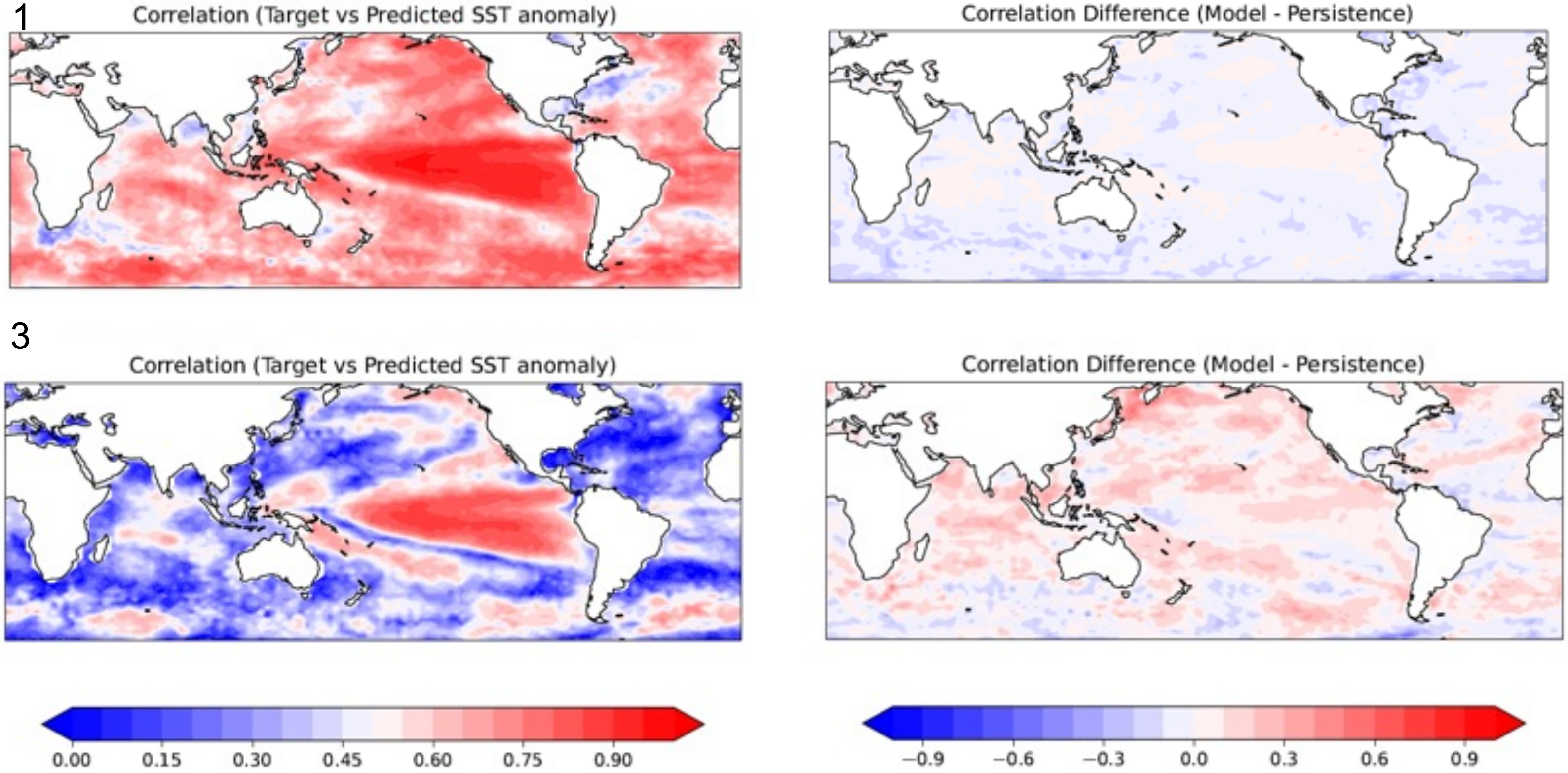
- [64,128] “global” grid
- Input variables: ERA5 SST and Ta 2m, EN4 upper ocean heat content from 1940
  - Input data’s time span: 3 months or 6 months
- Output variable: SST (surface temperature)
  - Model prediction: 3 months

Results

Improvement against persistence

Lead = 1

Lead = 3



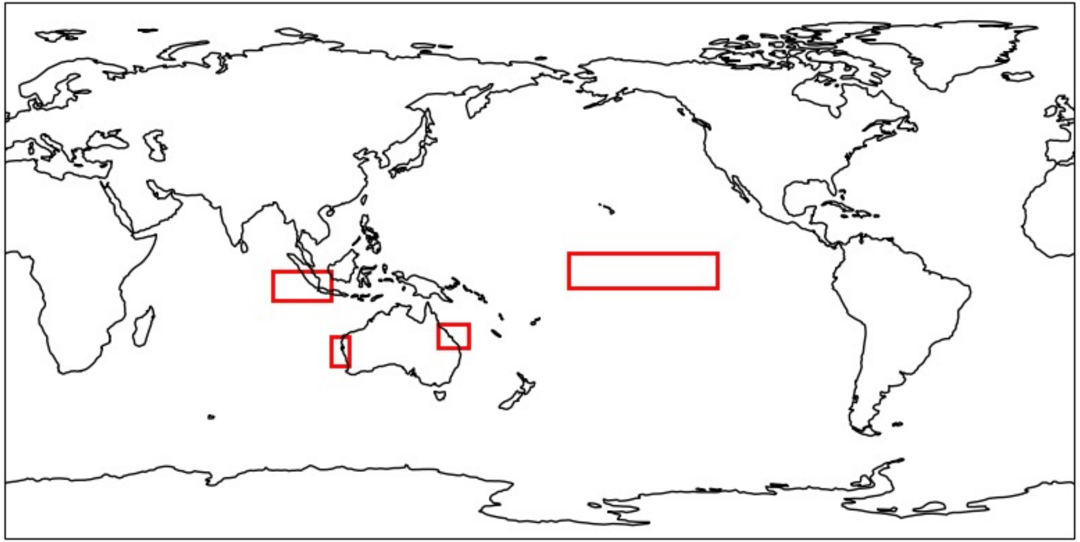
Selected regions to evaluate predictions

Nino 3.4 – (170^{o}W-120^{o}W, 6^{o}S-6^{o}N)

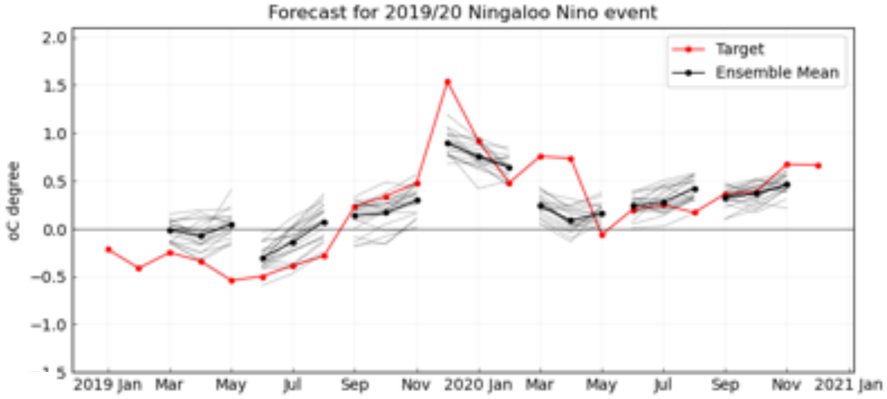
(Eastern) IOD – (90^{o}E-110^{o}E, 10^{o}S-0^{o})

Ningaloo Nino – (110^{o}E-116^{o}E, 32^{o}S-22^{o}S)

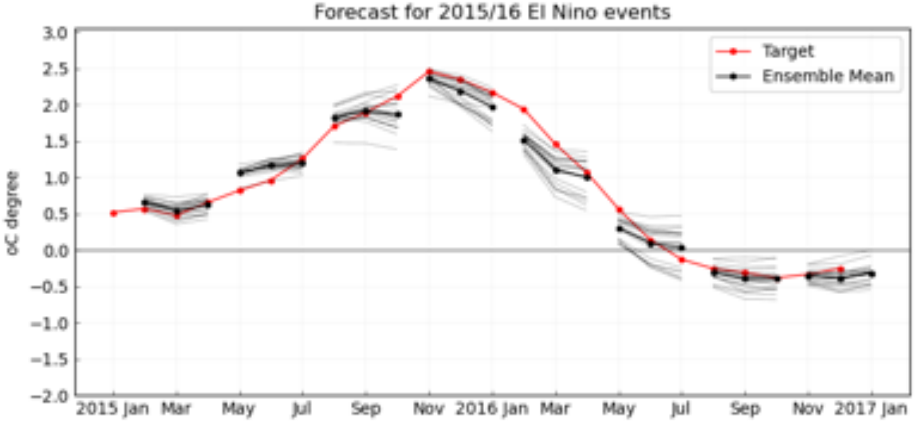
Great Barrier Reef – (146^{o}E-156^{o}E, 26^{o}S-18^{o}S)



3-month lead prediction for Nino3.4 and WA

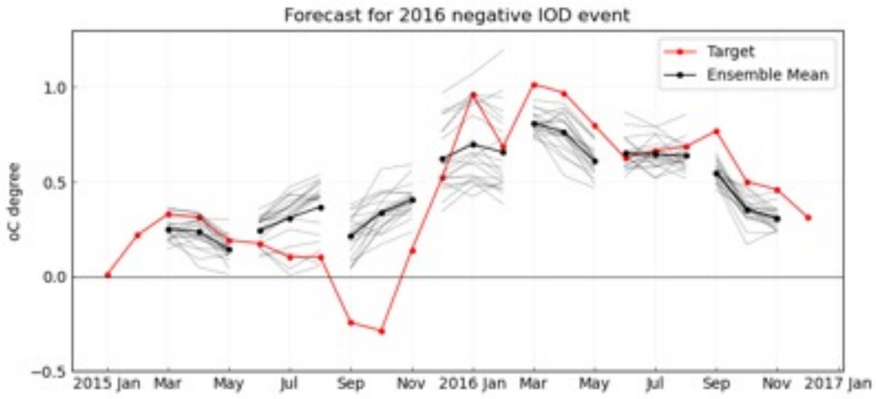


2015-16 El Nino



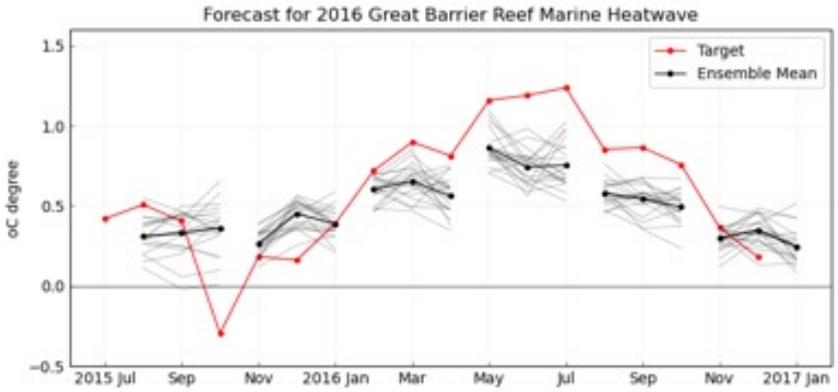
2019-20 Ningaloo Nino

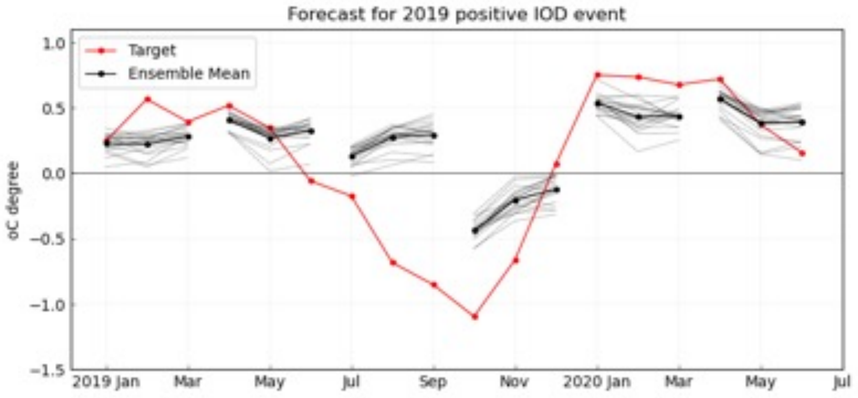
Predictions for eastern IOD and GBR



2016 GBR marine heatwave

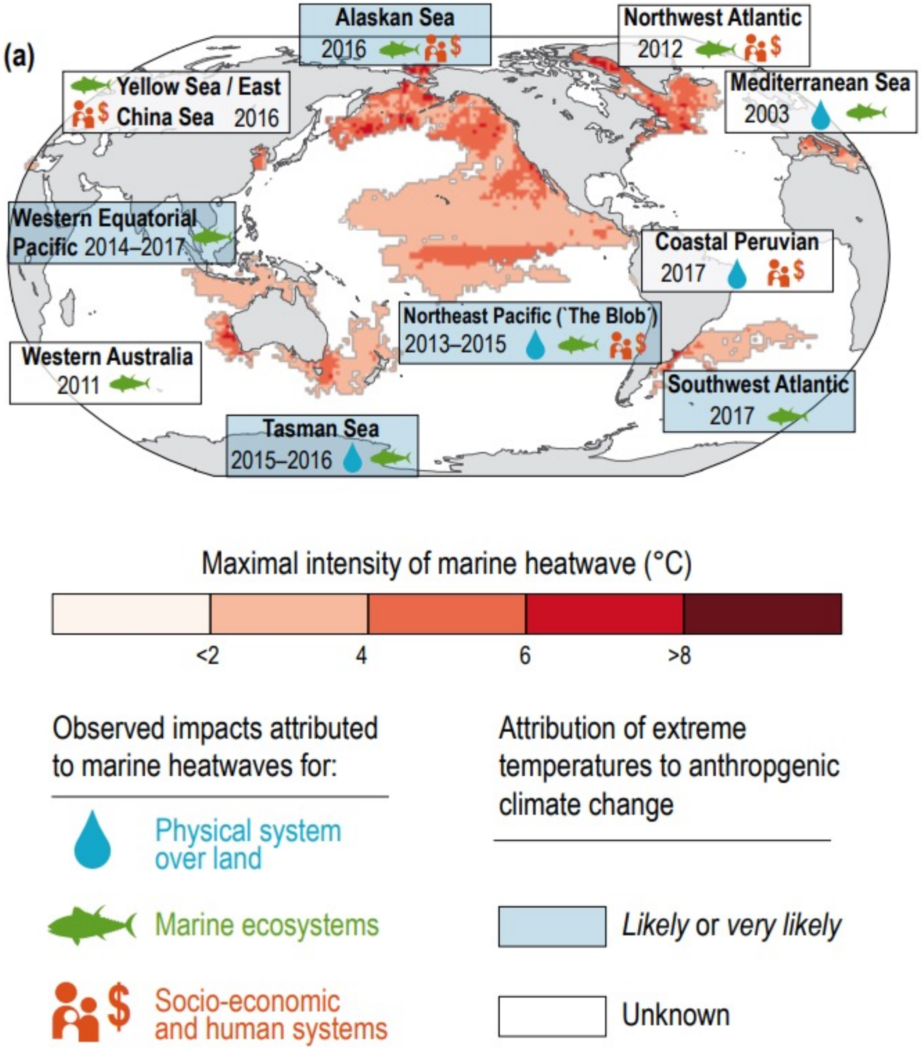
2016 negative IOD





2019 IOD

Prominent marine heatwave events in the recent decade



Summary and future works

- "Generations of ML models": CNN → Unet → Transformer
- The complex models may require less training data and show prediction skills at short-term lead (compared to persistence)

- Training data – adding other variables such as surface winds
- Assess longer lead predictions
- High spatial resolution (regional)
- Higher temporal resolution (daily) to predict marine heatwaves

Thank you

CSIRO Environment
Ming Feng
Senior Principal Research Scientist

+61 8 9333 6512
ming.feng@csiro.au
www.csiro.au