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
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
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
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

- Transformer (Vaswani et al. 2017)
- Natural language processing (NLP)
- Generative Pre-trained Transformer (GPT)
- Vision Transformer (ViT) (Dosovitskiy et al. 2020)
- Swin – Shifted windows Transformer (Liu et al. 2021)
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

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

Swin-Tunet – John Taylor
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

Correlation (Target vs Predicted SST anomaly)

Correlation Difference (Model - Persistence)

Lead = 3

Correlation (Target vs Predicted SST anomaly)

Correlation Difference (Model - Persistence)
Selected regions to evaluate predictions

Nino 3.4 – (170°W-120°W, 6°S-6°N)

(Eastern) IOD – (90°E-110°E, 10°S-0°)

Ningaloo Nino – (110°E-116°E, 32°S-22°S)

Great Barrier Reef – (146°E-156°E, 26°S-18°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