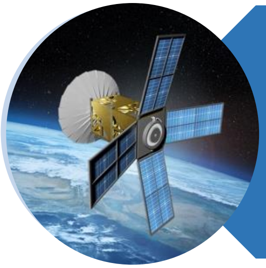
Applications of ML for Hydrodynamics and water quality Simulations in Chesapeake Bay

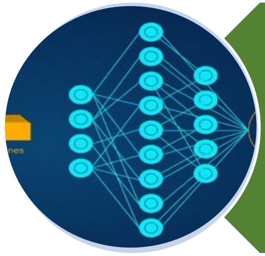
Jian Shen

Virginia Institute of Marien Sciences

Outlines



Background

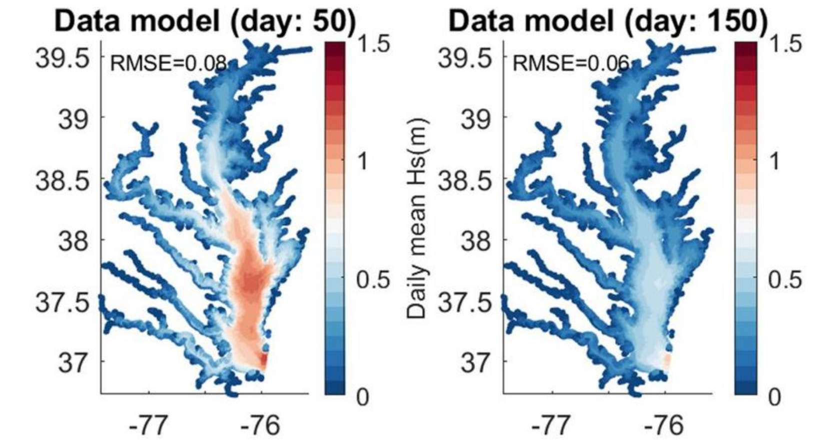


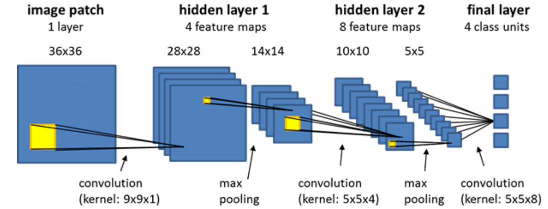
ML Example

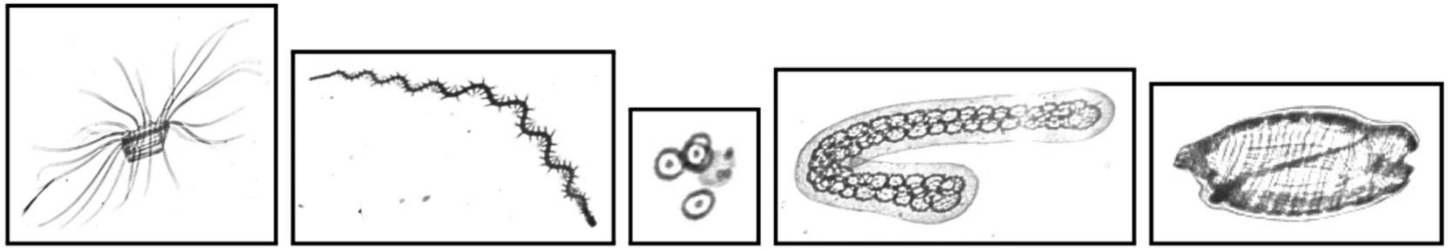
Future development

Background

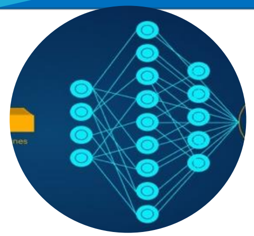
- AL/ML has been applied in the Chesapeake Bay
  - Water Quality & Phytoplankton (Ocean Color / Satellite Imagery)
  - Water Quality Detection for Aquaculture
  - Harmful algal species
  - Wetland Mapping with Deep Learning
  - Shoreline evolution/shoreline armoring
  - Classify submerged vegetation from aerial imagery
  - Plankton image recognition
- VIMS estuarine and coastal modeling group have focused on applications of ML for predictions and management for both hydrodynamics and water quality





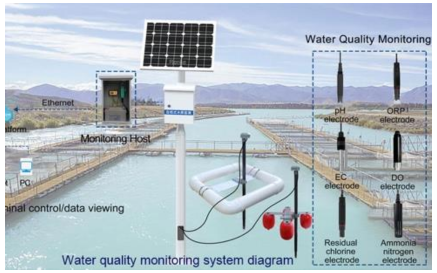


Applications in Chesapeake Bay



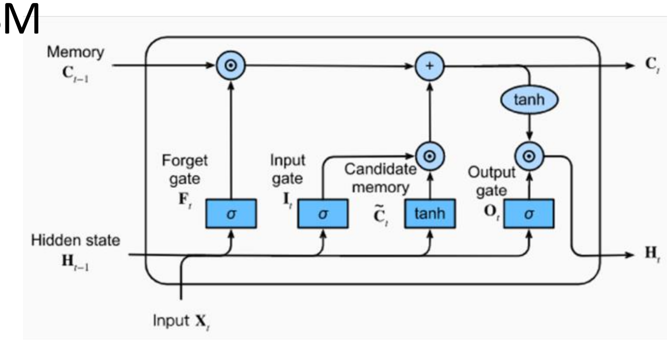
Computer model

Observation data

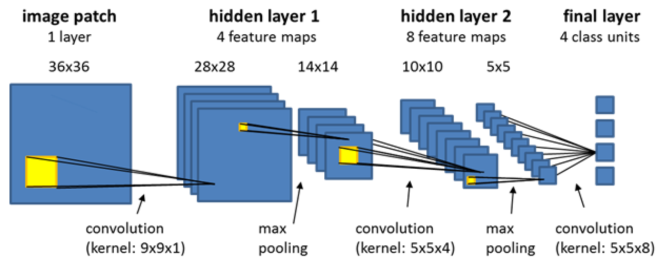


- Forecasting model
    - Storm surge, Flooding
    - Tide, Wave
    - Salinity, Temperature
- Saltwater intrusion
- Ecosystems and Water quality model
    - Primary production
    - Dissolved oxygen/hypoxia volume
    - Harmful algal bloom
    - Phytoplankton model
    - Water quality assessment
- Management model
    - Phytoplankton vs nutrient reduction

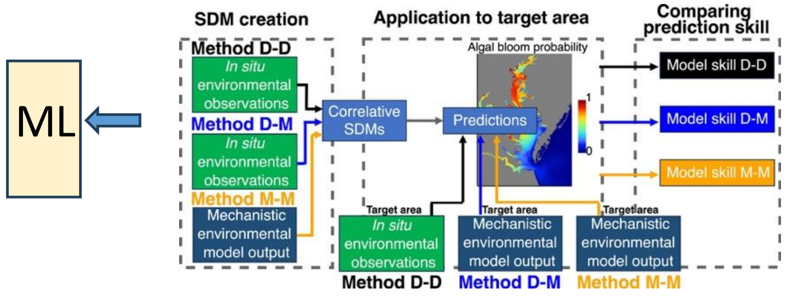
LTSM



CNN/Transformer with attention



Combine numerical model and ML



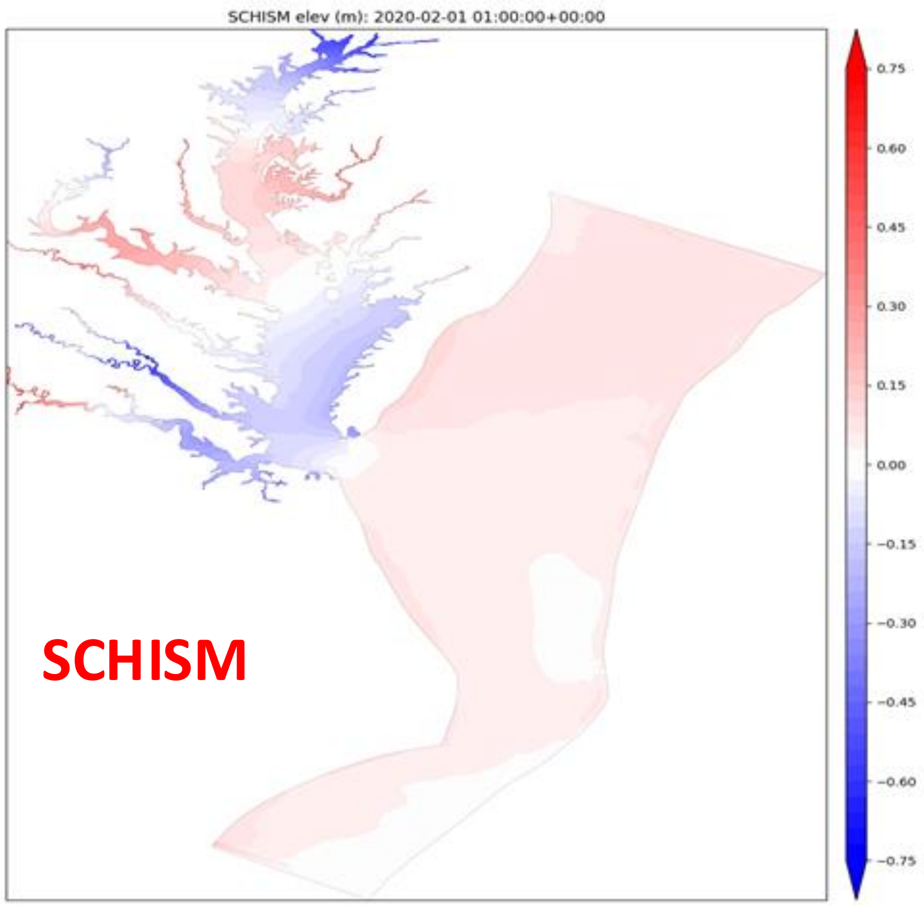
Available Data

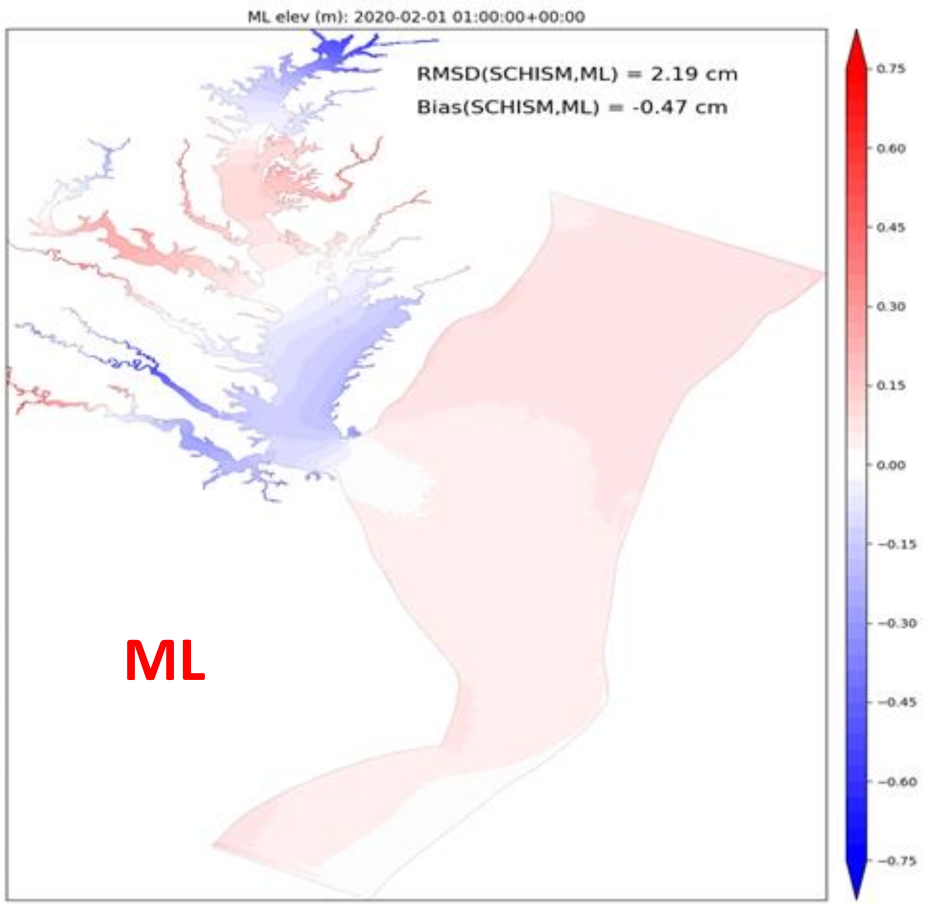
- Observations
  - CBP Long-term water quality observations (more than 40 years)
  - NOAA long-term monitoring tide, wind, temperature), and buoy data
  - NOAA shallow water monitoring data (tide, salinity, Chl a, temperature etc.)
- Numerical model results
  - Reliable numerical models have been aviation in the Bay and generated long-term model results (ROMs, SCHISM, SWWN, FVCOM etc.)
- Satellite data
- These data can be used for training ML
  - Provide fast and effective predictions
  - Use limited forcing data for for casting to the future and reconstruction of historical evolution
  - Provide uncertainty anabasis
  - Provide alternative tools for management

Hydrodynamic modeling

VIMS WILLIAM & MARY
VIRGINIA INSTITUTE OF MARINE SCIENCE

- ML simulation of hourly tide in high resolution (Wang, Zhang, et al.)
  - Use hourly 3D model simulations (20-year) of tide to train ML model
  - ML forced by selected tide, salt, wind at boundary
  - The ML model can simulate hourly effectively with high model skill.

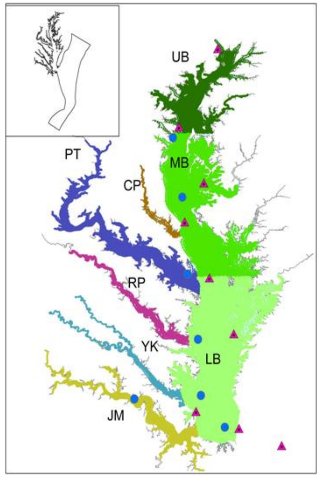




Wave Simulations

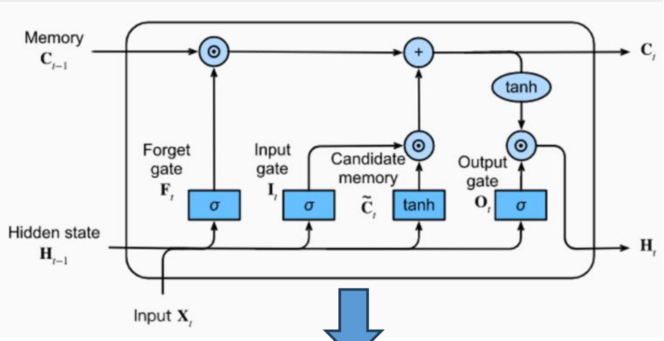
VIMS WILLIAM & MARY
VIRGINIA INSTITUTE OF MARINE SCIENCE

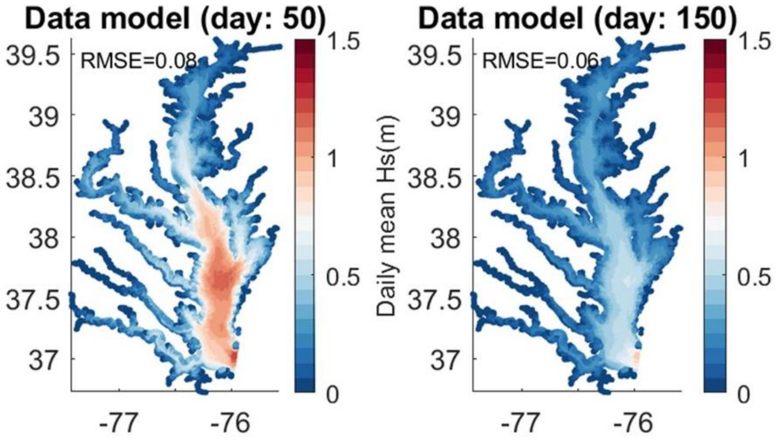
- Use numerical wave model data to train ML model, and verified the model using NOAA observations
- Predict daily mean/maximum significant wave based on daily wind data at 9 stations.



Daily wind data at 9 stations

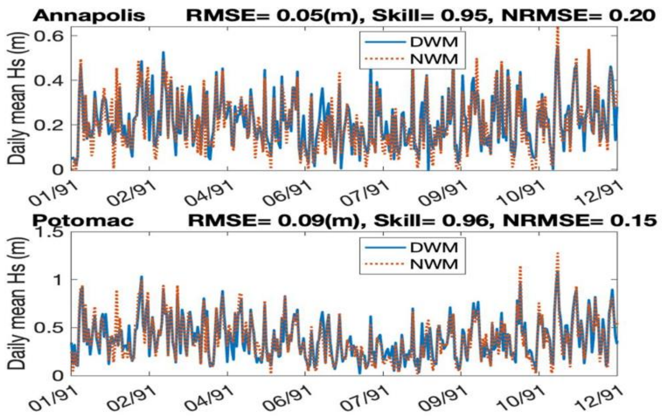
LSTM model



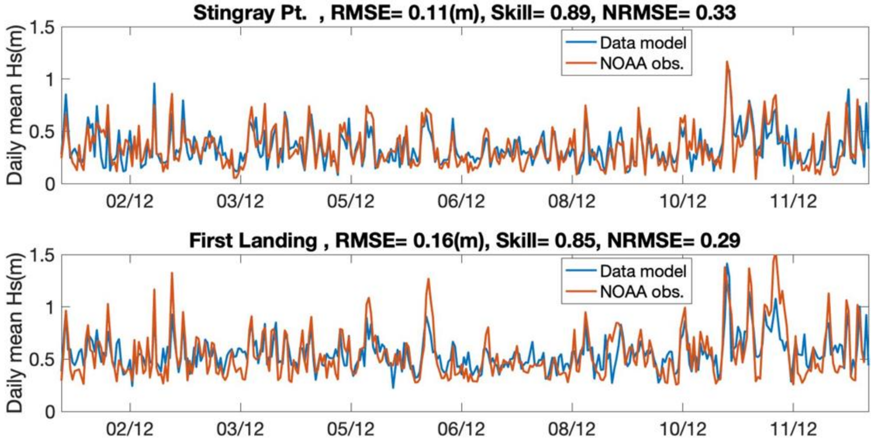


(Shen et al. 2024)

Test



Verification NOAA data



Advantage

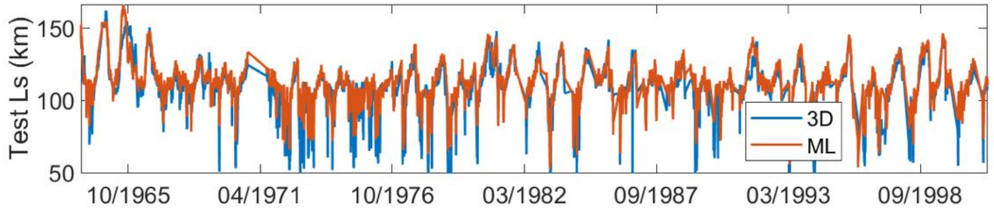
- Only needs wind data
- It can conveniently do forward and backward simulation of wave
- Easy to access climate change

Salinity Intrusion in Chesapeake Bay

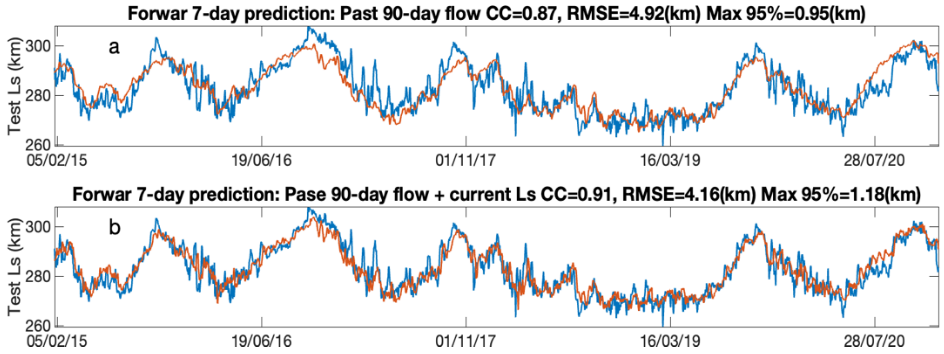
VIMS | WILLIAM & MARY
VIRGINIA INSTITUTE OF MARINE SCIENCE

- Simulate saltwater intrusion trained using 3D model simulation (CNN + ensemble)
- Forward forecast using preceding 90-day flow data
- Reconstruction history saltwater intrusion
- Conduct management scenarios

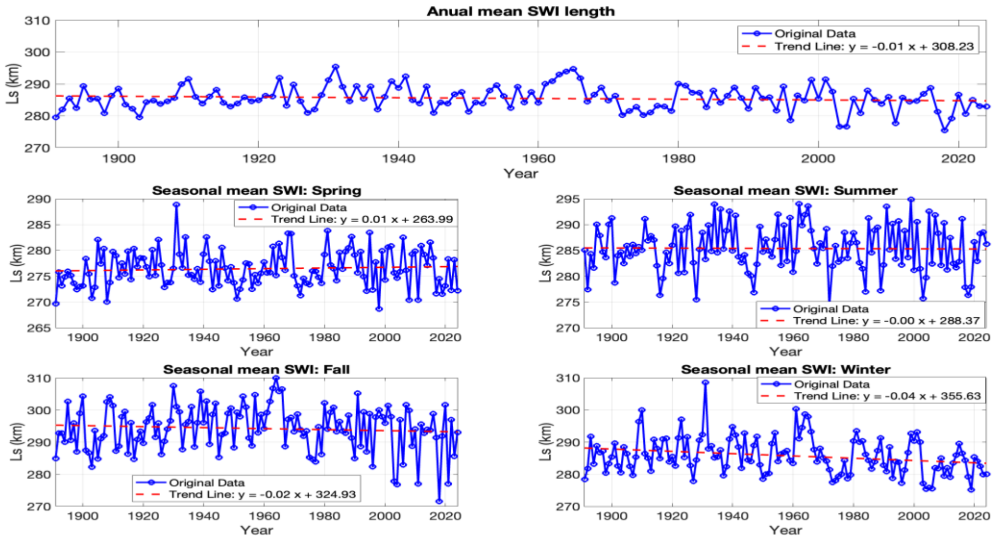
Verified using Delaware Bay data



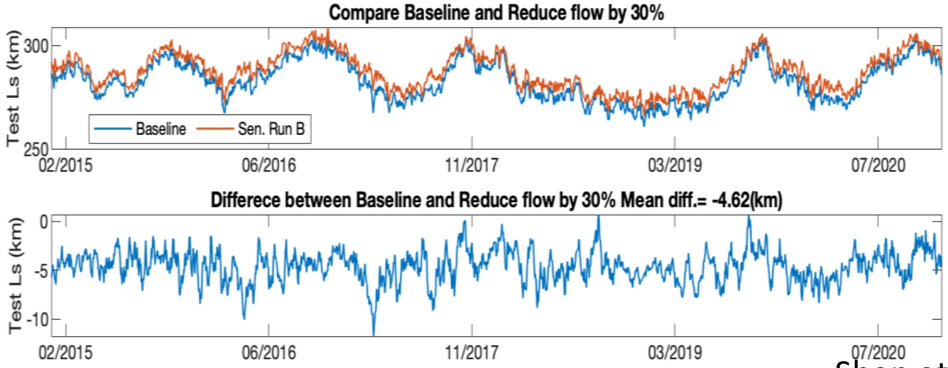
Forward prediction (7 to 14 days)



Analysis of historical vairones



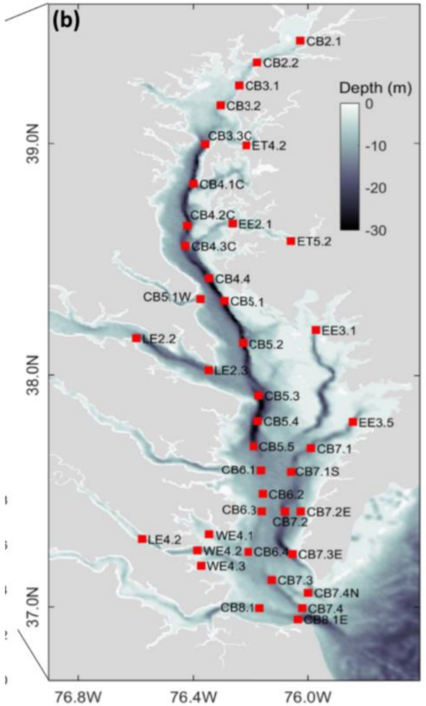
Flow reduction scenarios

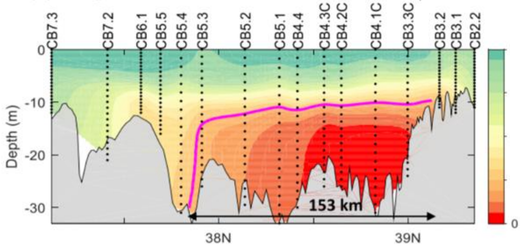


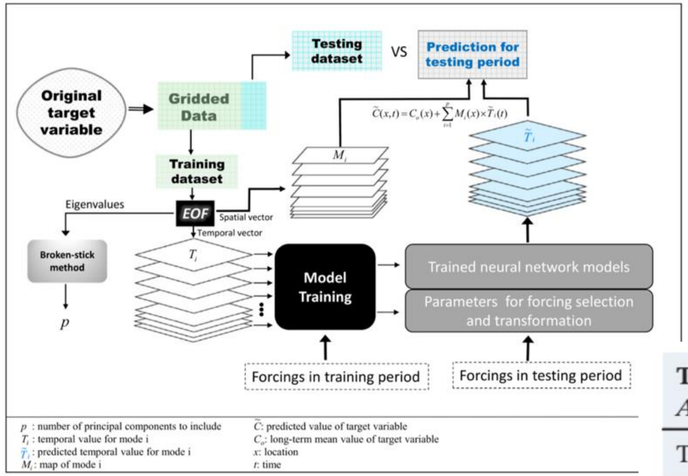
Shen et al., 2026

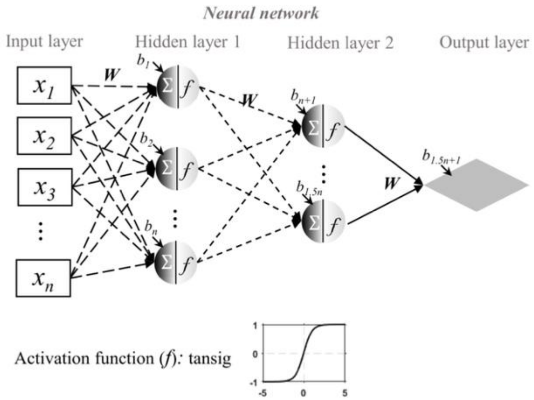
DO Simulaiton

- Prediction of DO









- Use EOF to separate spatial and temporal variation.
- Spatial vector preserved vertical structure
- Drive by all external forcing: flow, N loading, wind, temperature (similar to numerical model)
- Use parameter transformation
- Be able to use for management

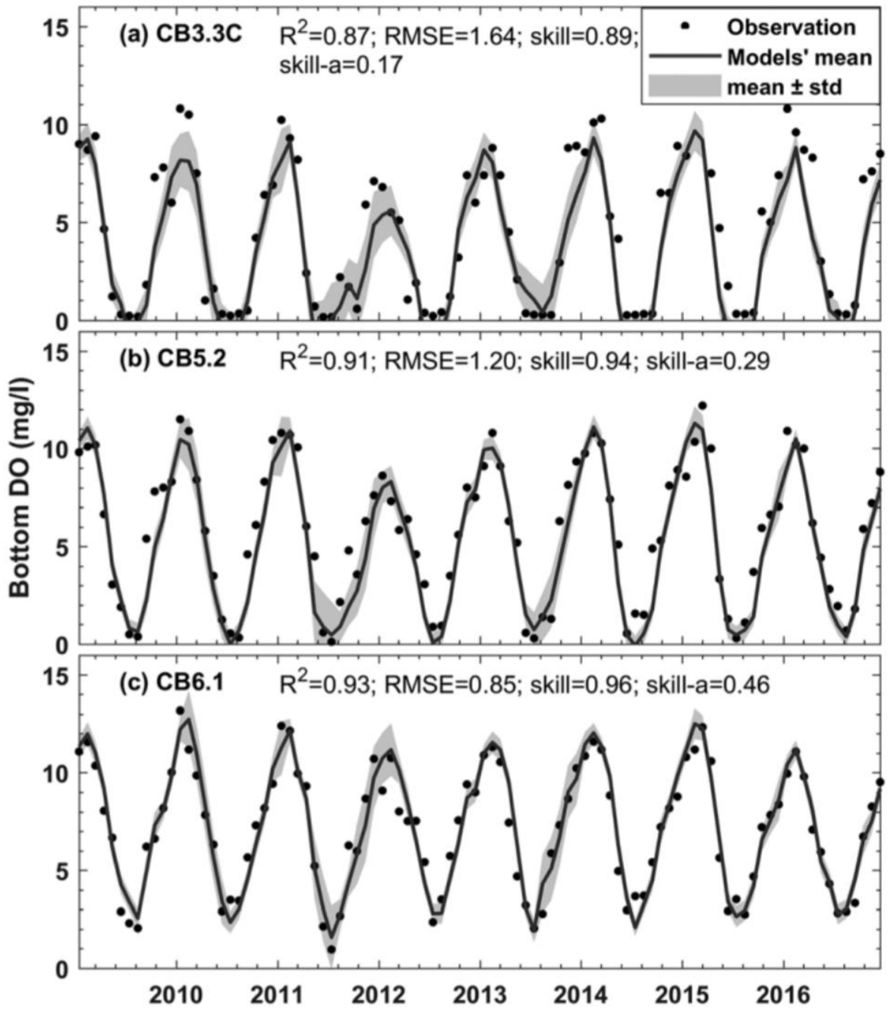
Table 1
A List of the Transformation Options in the Data-Driven Model

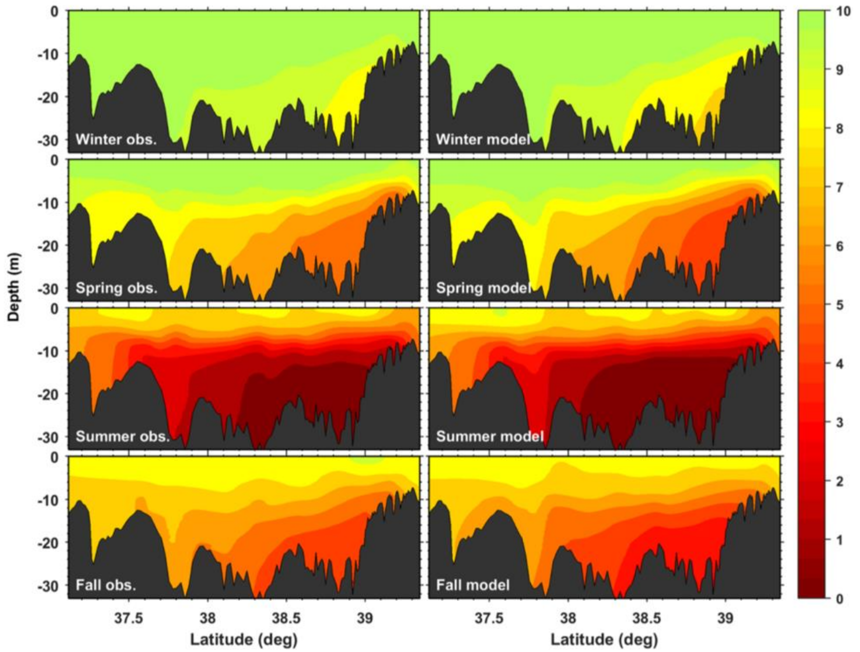
Transformation	Subtypes	Formula
Time-lag transformation	1–7	ϕ(t) = x(t-lag), with lag ranging within 0, 10, ... 60 days
Accumulative transformation	1–13	ϕ(t) = mean(x(τ)), where $\tau \in [t1 - acc, t2]$, with acc ranging from 0 to 120. *t1* and *t2* are the beginning and end of each month; t1 = t − 15 and t2 = t + 15
Regular transformation	1	ϕ = x
	2	ϕ = log(x)
	3	ϕ = 1/x
	4	ϕ = exp((x-mean(x))/std(x))
	5	ϕ = x/(p50 + x), also known as Monod-type filter
	6	ϕ = x/(p75 + x)
	7	ϕ = x/(p25 + x)
	8	ϕ = (x-mean(x))/std(x)

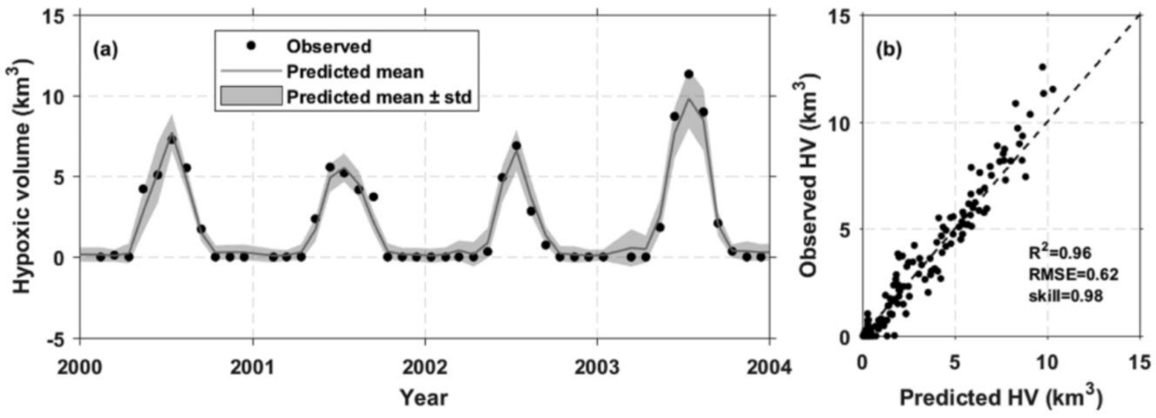
Note. x = forcing variable; ϕ = transformed forcing variable; t = time; std(x) = the standard deviation of x; mean(x) = the mean value of x; P25, P50, P75 = the 25, 50, and 75 percentile of x.

(Xu et. 2020, Water Resource Research)

2. Applications: DO

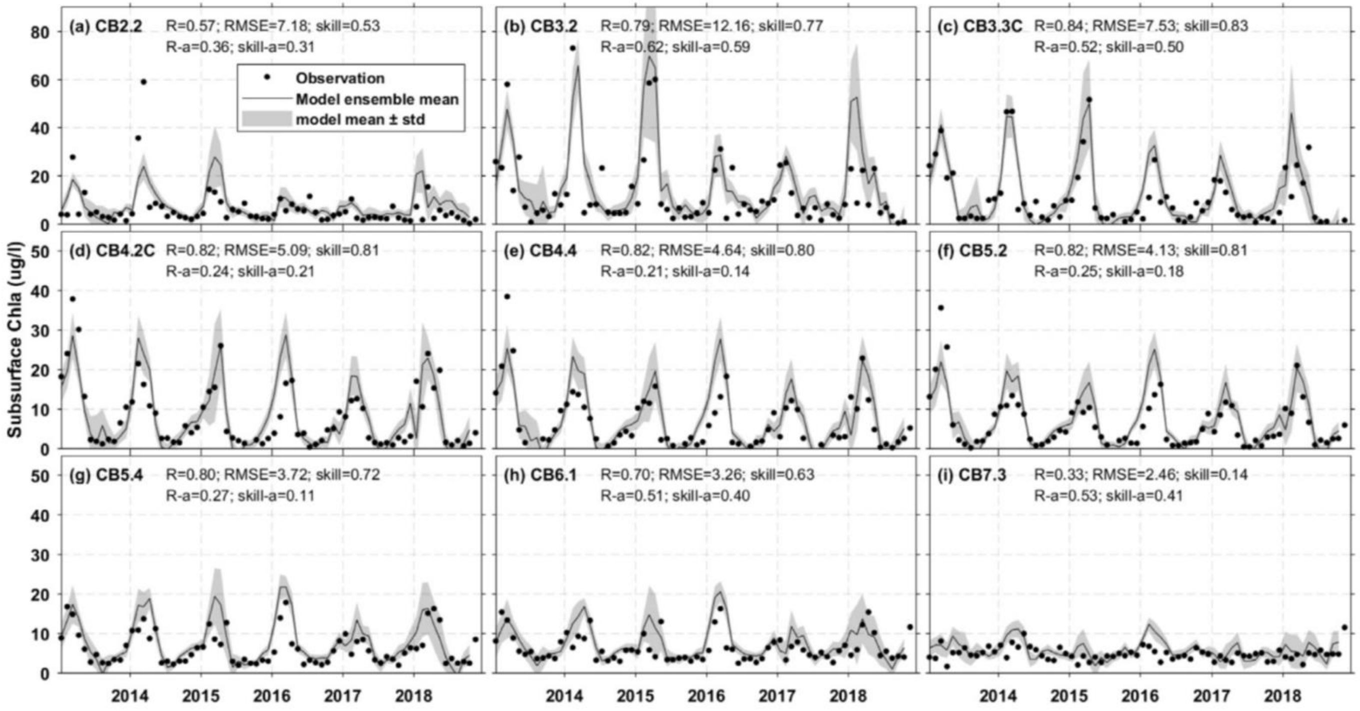


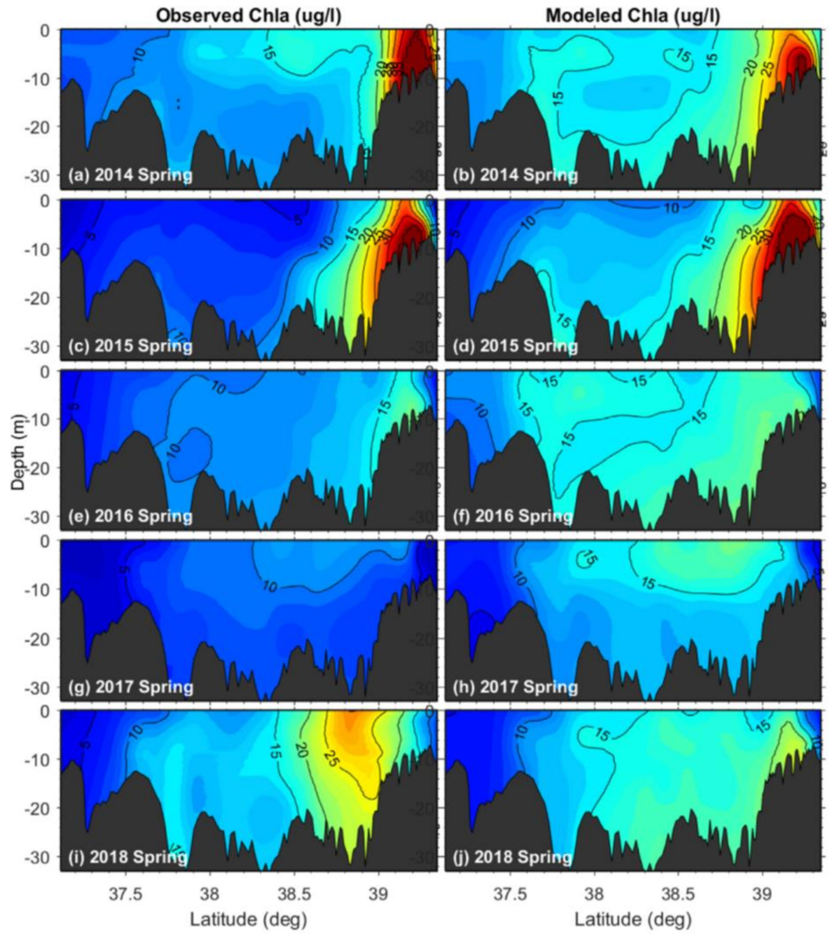




Phytoplankton Simulation

- Predict phytoplankton

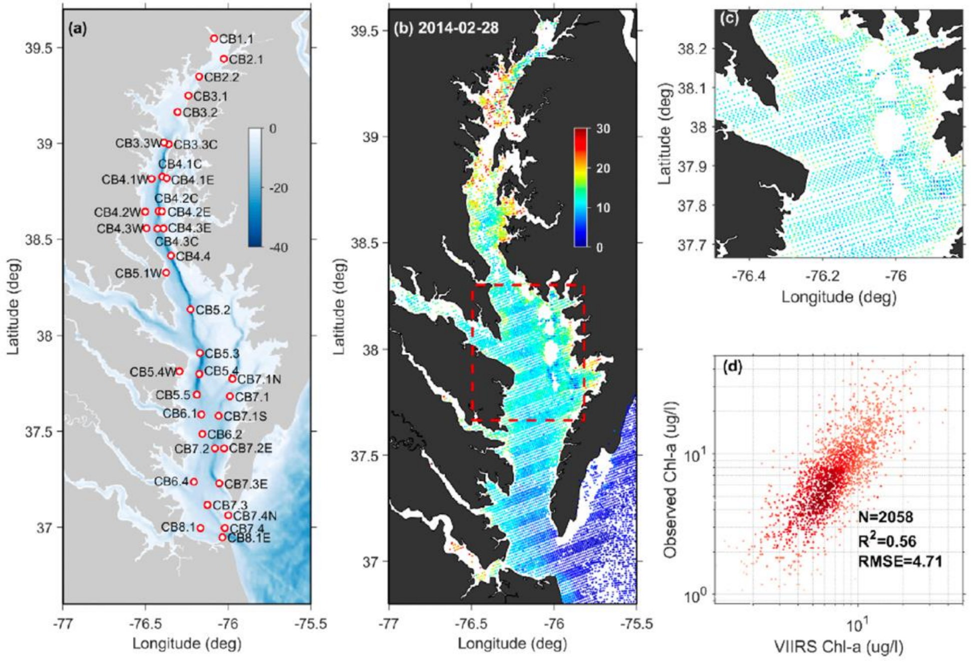


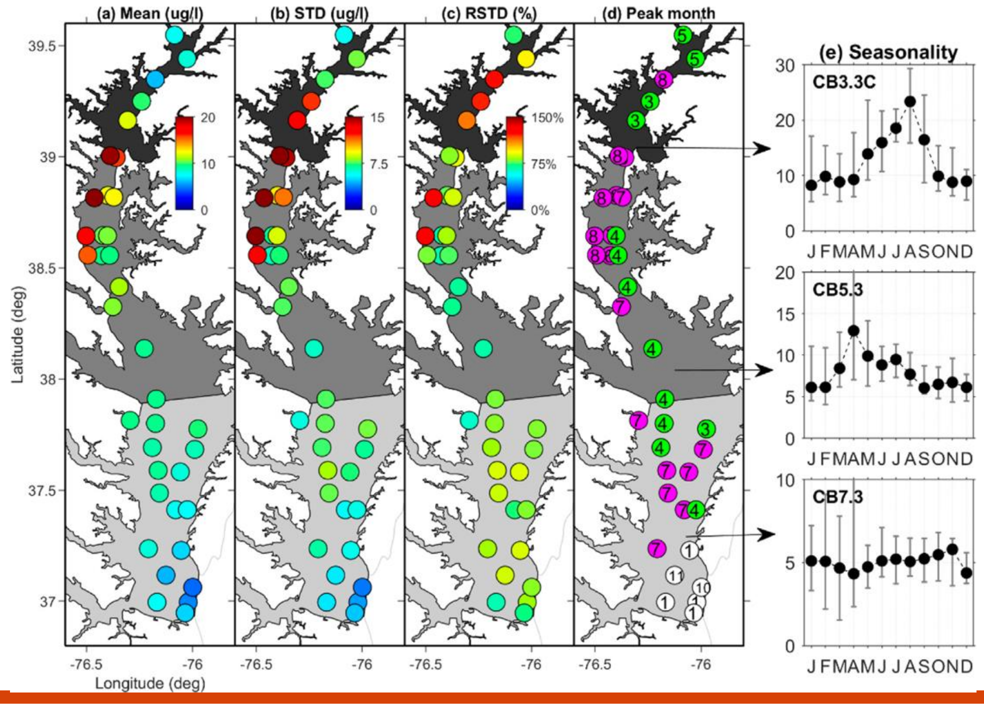


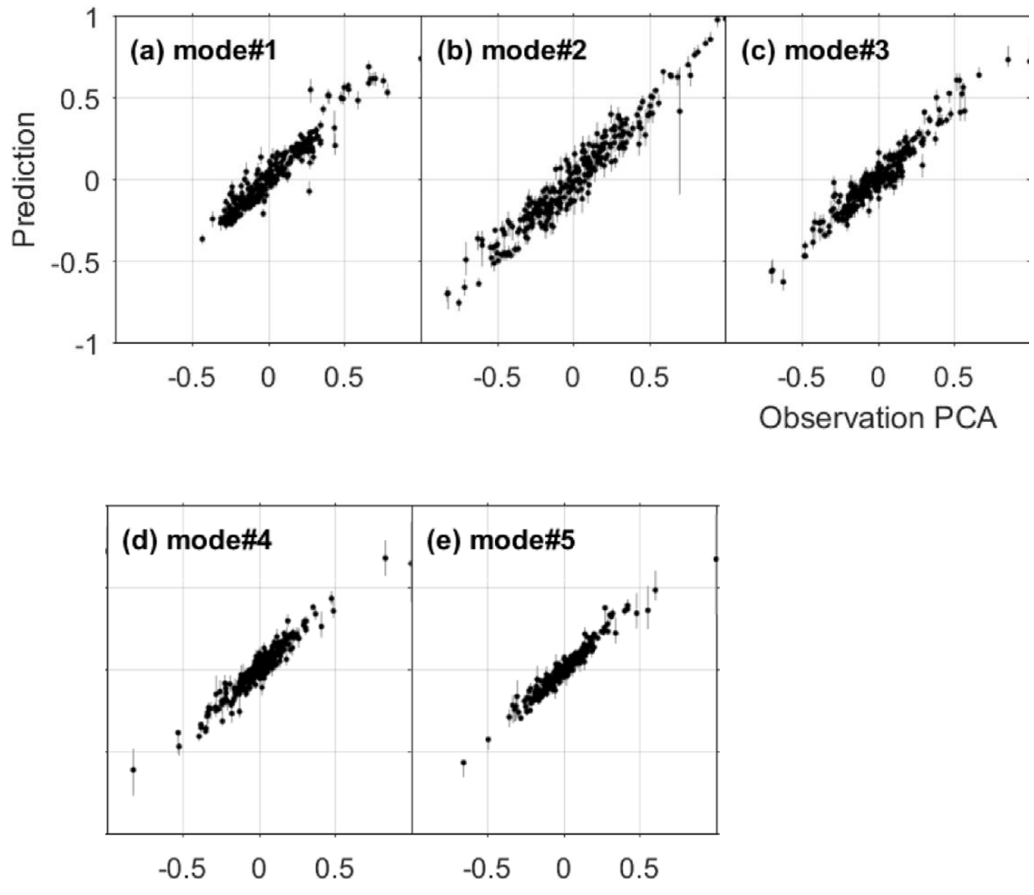
(Xu and Shen, 2021. Ocean Modeling)

Phytoplankton Simulations

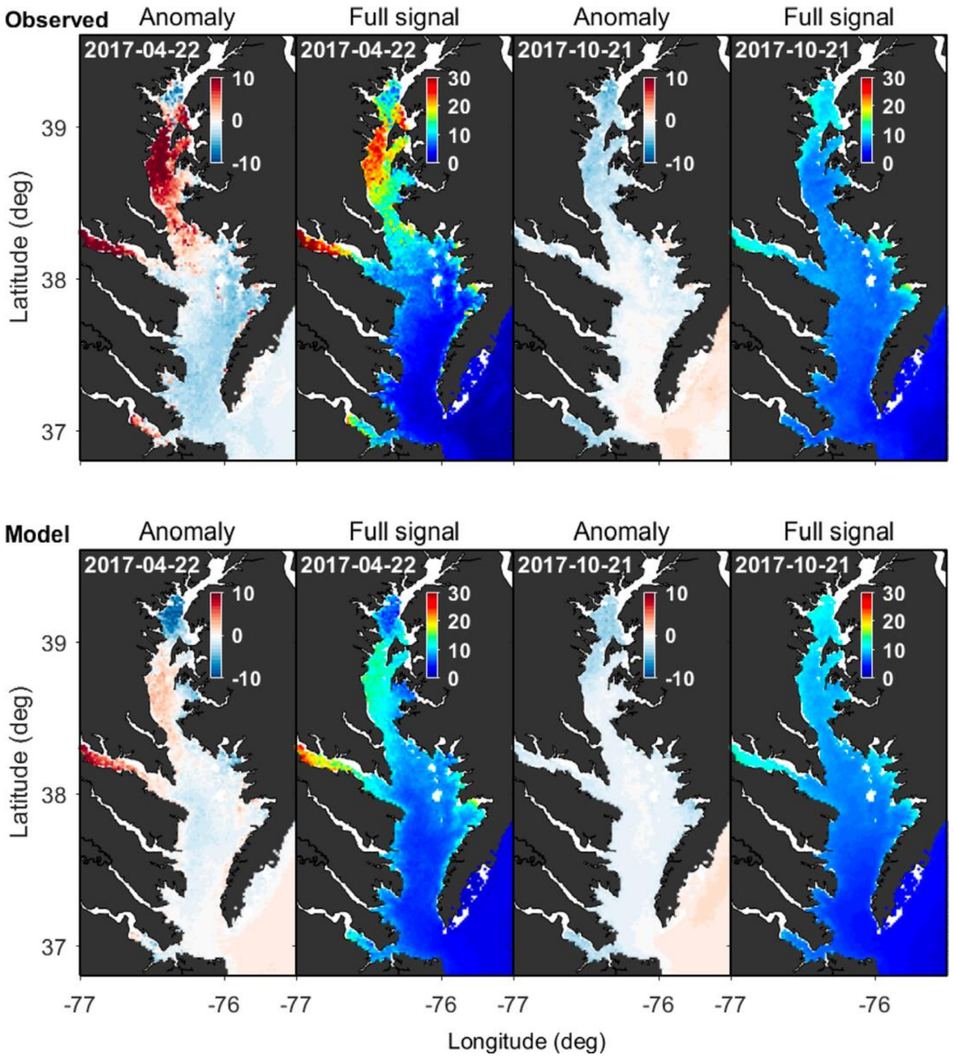
- Use daily *VIIRS satellite data of Chl-a*
- DINEOF is used to efficiently estimate the missing records (7-day mean)
- CNN to train model using external forcing (Flow, TN,TP loading, temperature, solar radiation)
- Parameter transform were applied to find the best parameters







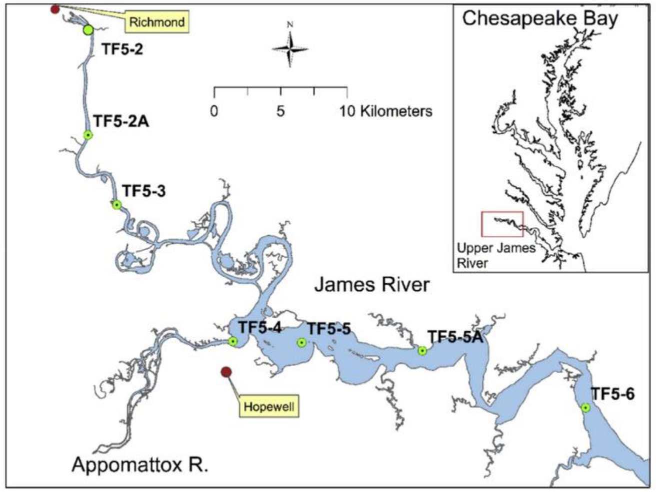
(a) mode#1
(b) mode#2
(c) mode#3
(d) mode#4
(e) mode#5
Prediction
Observation PCA
1
0.5
0
-0.5
-1
-0.5
0
0.5

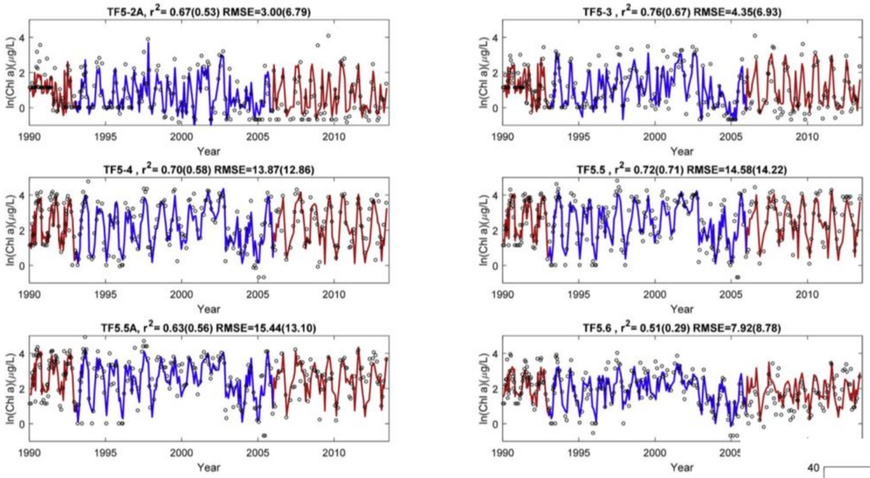


Observed
Model
Anomaly
Full signal
2017-04-22
2017-10-21
10
0
-10
30
20
10
0
Latitude (deg)
39
38
37
Longitude (deg)
-77
-76

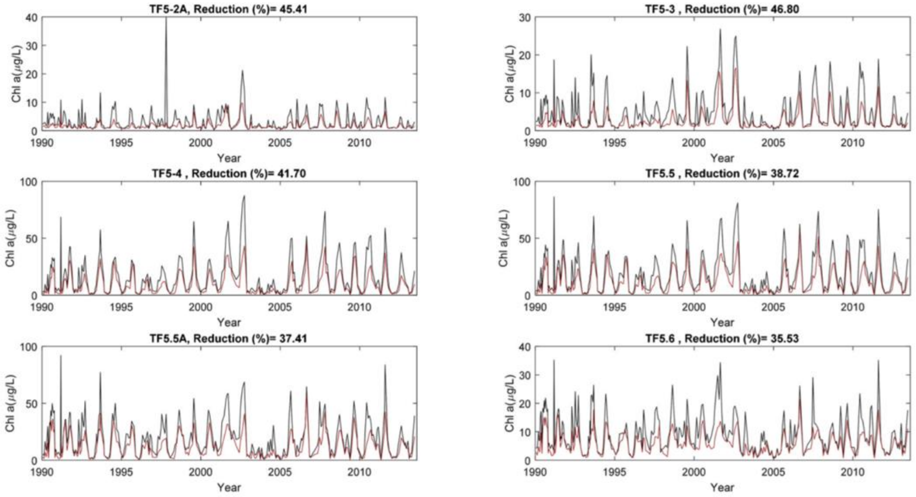
Phytoplankton Prediction in James

- Can we use ML model for management
- Predicate phytoplankton (Shen et al., 2019, Ecological modeling





Management scenario:
Reduce nutriment by 50%



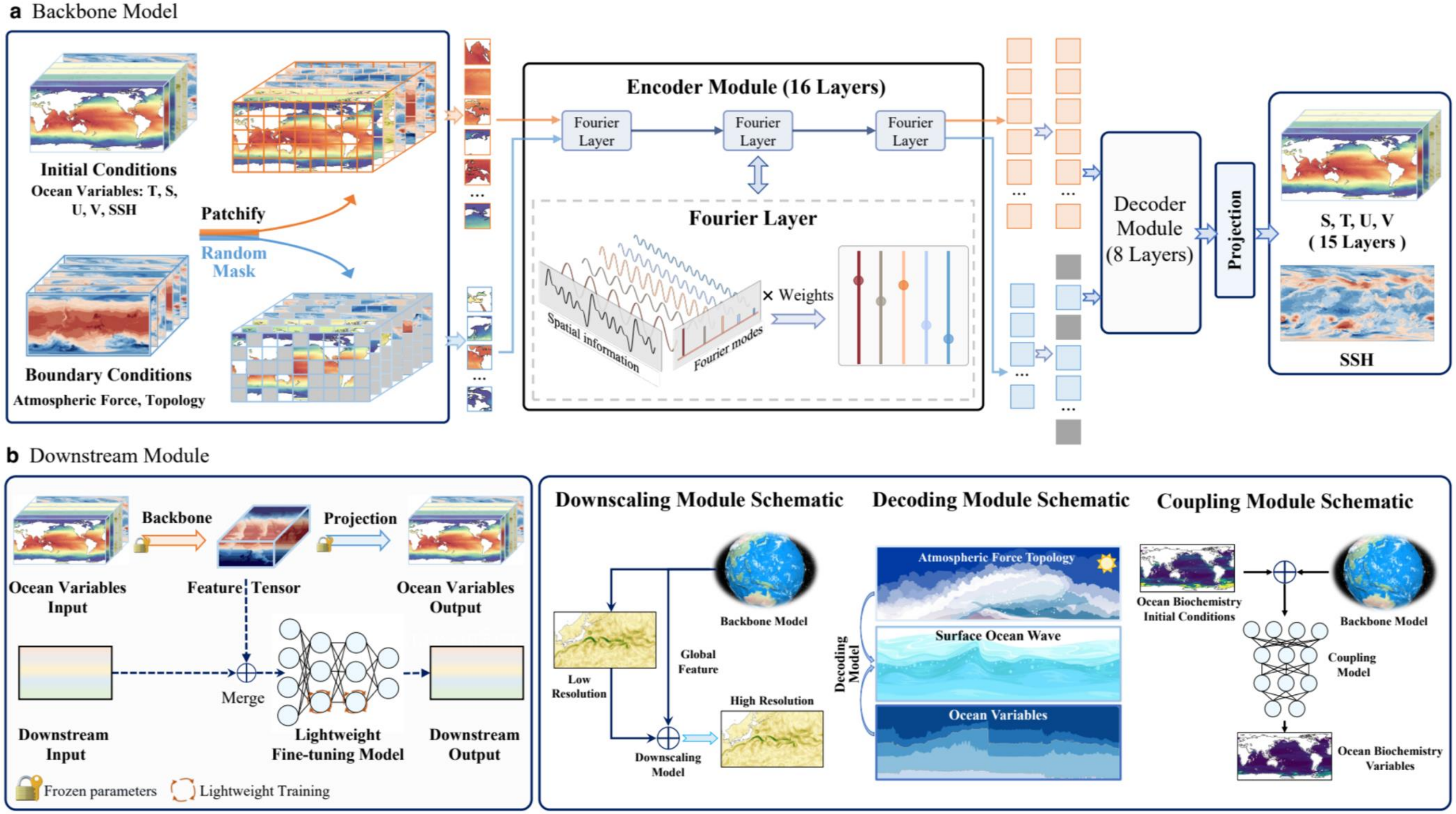
- Input parameters: watershed model outputs (flow, nutrients, temperature)
- Support vector machine LS-SVM (project to high dimension)
- Parameter transformation
- Without use temperature as an independent variable

$$TN_{new} = \frac{TN}{H_{TN} + TN}\theta^{T-20}$$

(Shen et al., 2019, Ecological modeling)

3. Future development: Digital twin

- AI-GOMS model (Xiong et al., 2023, arXiv)



Conclusions

- Estuarine and coastal modeling group of VIMS have engaged in ML to develop prediction models for hydrocyanic and water quality simulations in the Chesapeake bay
- There are large amounts of observational data and many numerical models, which can be used for training ML model
- ML has been applied to the Bay for various applications
- ML has the potential to be applied for forecasting and management