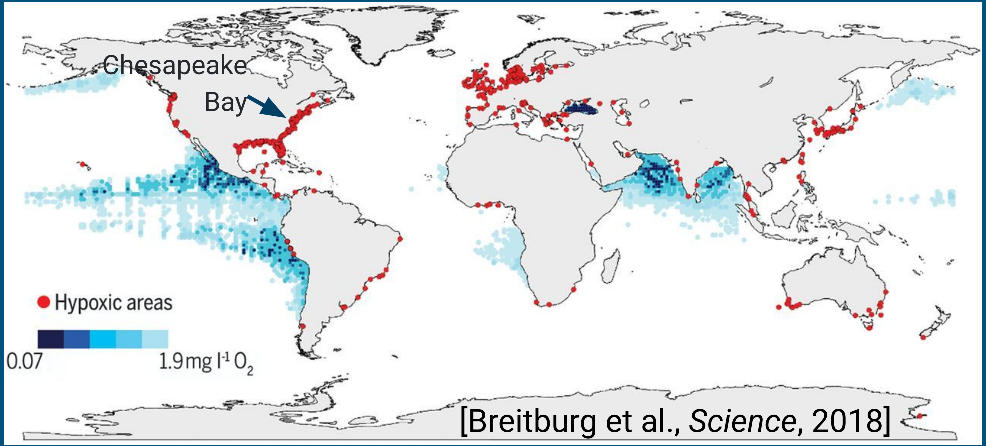
HypoxAI: A Model to Forecast Chesapeake Bay Hypoxia Using Convolutional and Recurrent Neural Networks

Guangming Zheng [1,2] and Stephanie Schollaert Uz [3], Marjorie Friedrichs [4], Pierre St-Laurent [4], Amita Mehta [5], Paul M. DiGiacomo [1]

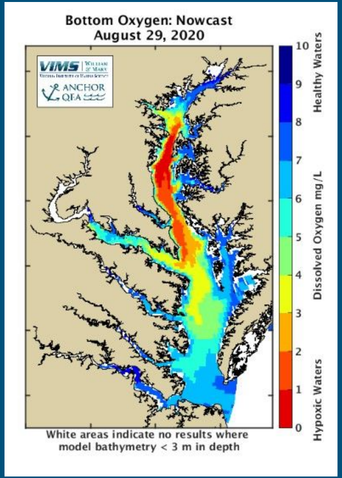
[1] NOAA/NESDIS/Center for Satellite Applications and Research, 5830 University Research Court, College Park, MD 20740
[2] Cooperative Institute for Satellite Earth System Studies, Earth System Science Interdisciplinary Center, University of Maryland College Park, 5825 University Research Court, College Park, MD 20740
[3] Earth Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771
[4] Virginia Institute of Marine Science, William & Mary, Gloucester Point, VA 23062
[5] NASA Goddard Earth Sciences Technology and Research, University of Maryland Baltimore County, Baltimore, MD 21250

Motivation

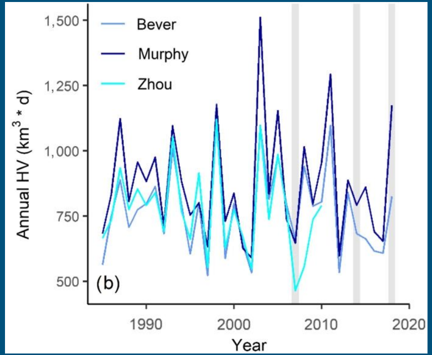
- Oxygen has been declining in both the open ocean and coastal waters during the past half-century
- In many open ocean and coastal waters oxygen levels dropped below hypoxic level (<2 mg L^{-1})
- Expanding hypoxia is one of three primary ocean consequences of rising atmospheric CO_2
- Hypoxia affects many processes ranging from biogeochemistry to food security



[Breitburg et al., *Science*, 2018]



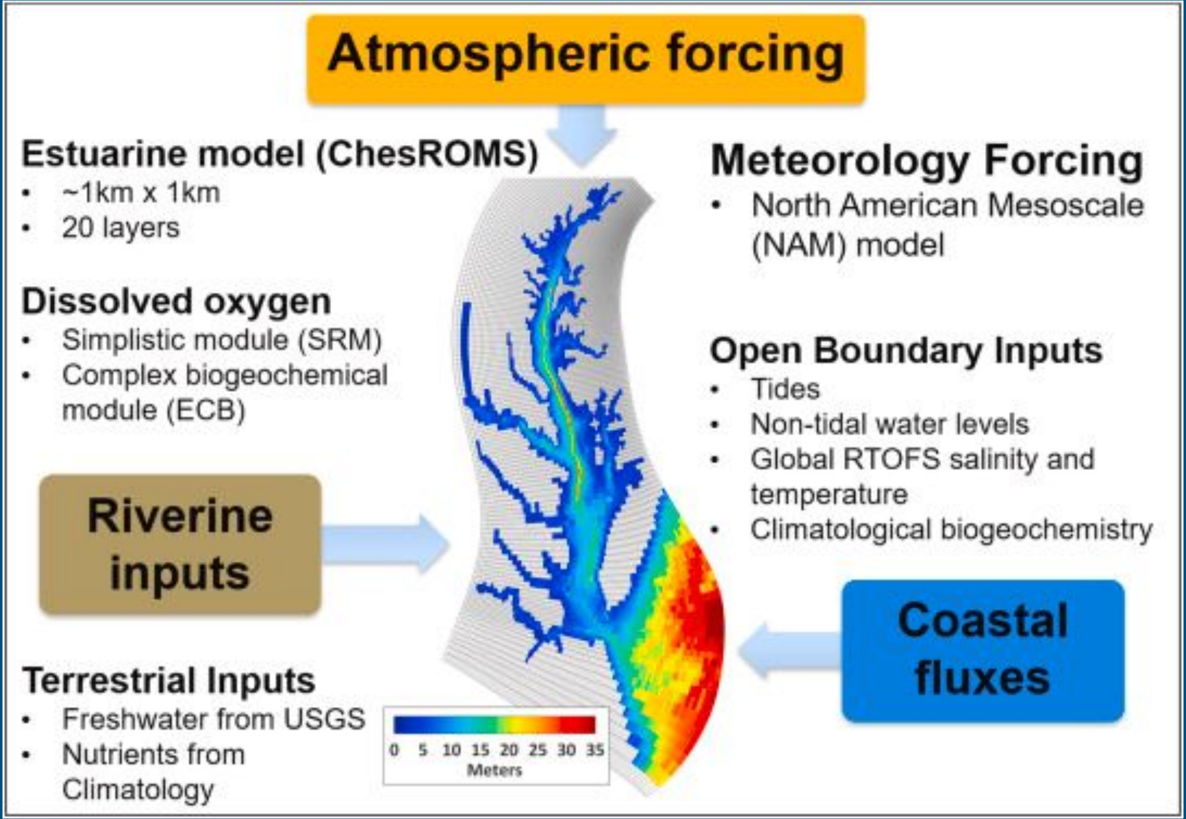
[Credit: M. Friedrichs]



[Scavia et al., *Ecol. Appl.*, 2021]

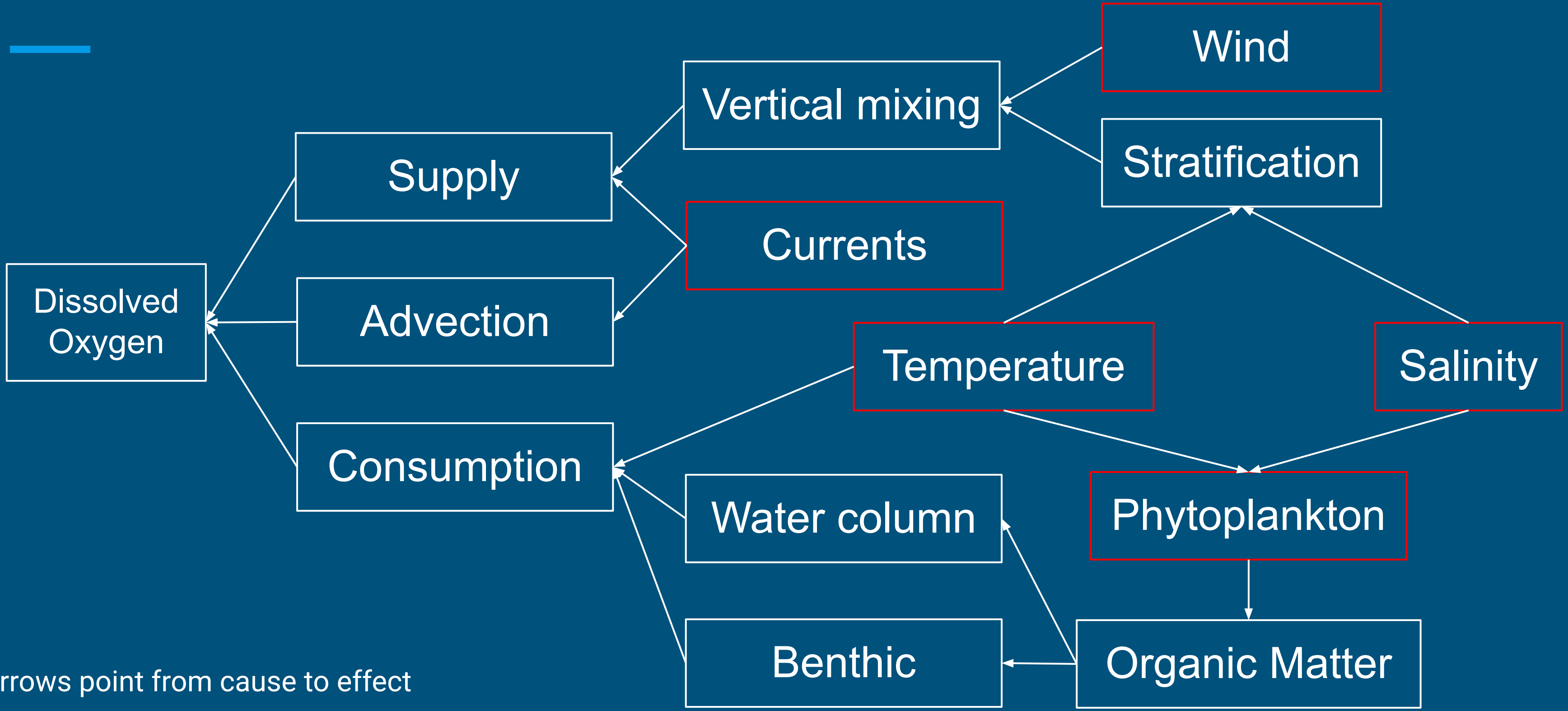
Traditional Models for Hypoxia Forecast

Chesapeake Bay Environmental Forecast System (CBEFS)



[Bever et al., *Environ. Model. Soft.*, 2021]

Root Cause Analysis of Chesapeake Bay Hypoxia



Arrows point from cause to effect

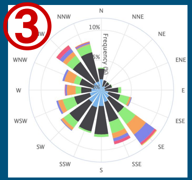
Root causes are highlighted in red boxes

Data and Preprocessing

① **MODIS-Aqua Satellite**:
- Rayleigh-corrected reflectance at 12 bands
- Weekly minimum (1st percentile)
- 2002-2020

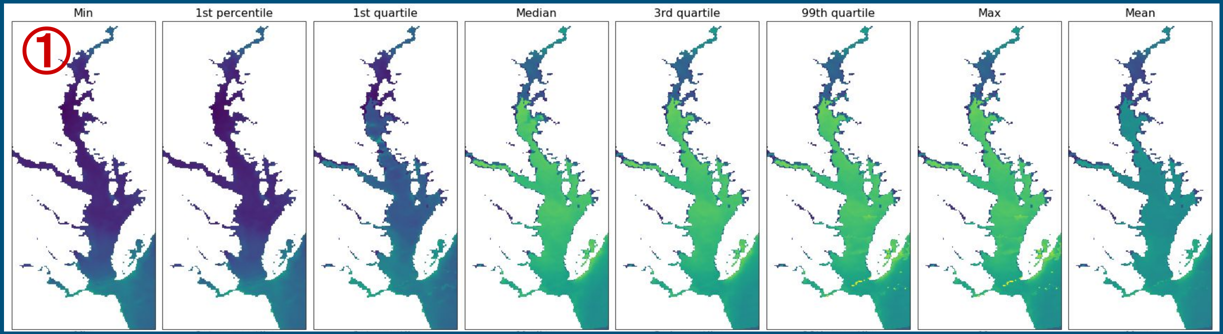
② **Virginia Institute of Marine Science (VIMS)**:
- Chesapeake Bay Environmental Forecast System (**CBEFS**)
- Currents (u, v, & w)
- Water temperature (T), Salinity(S)
- 3-D reprojected to fixed grids.
- Weekly mean

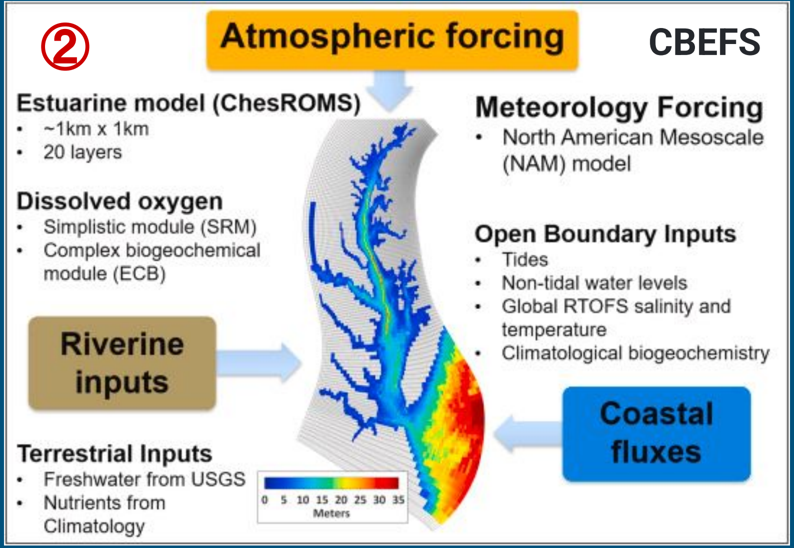
③ **ECMWF ERA5**:
- Wind (u & v)
- Hourly data
- Wind rose tabular data



④ **Chesapeake Bay Program (CBP)**:
- Field-measured Dissolved Oxygen concentration
- Mainstem stations only

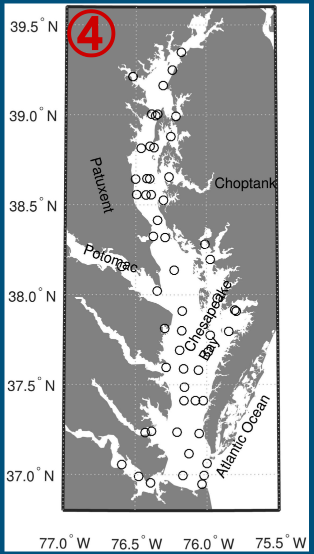
MODIS weekly statistics





[Bever et al., *Environ. Model. Soft.*, 2021]

CBP Stations



Features (x) and Target (y)

Features:

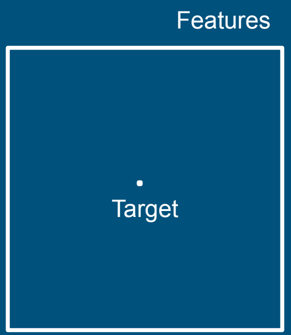
- Weekly time series (8 weeks):

  In a 10*t* by 6*t* km box around target:

  - MODIS Rayleigh-corrected reflectance at 16 bands
    (*t, lat, lon, λ*)
  - Currents at 3 depths, <5, 5-10, >10 m
    (*t, lat, lon, z*)
  - Temperature/Salinity at 3 depths
    (*t, lat, lon, z*)

  At target location:

  - Wind statistics (probability distribution)
    (*t, w_spd, w_dir*)
- Ancillary information:
  - Sampling depth (*non-dim*)
  - Bottom depth (*non-dim*)



Target:

DO concentration at sampling depth
1 day after last timestamp in features

Symbols:

t, time steps before the DO observation, e.g., 1 week before …

lat, latitude

lon, longitude

λ, light wavelength

z, depth

w_spd, wind speed

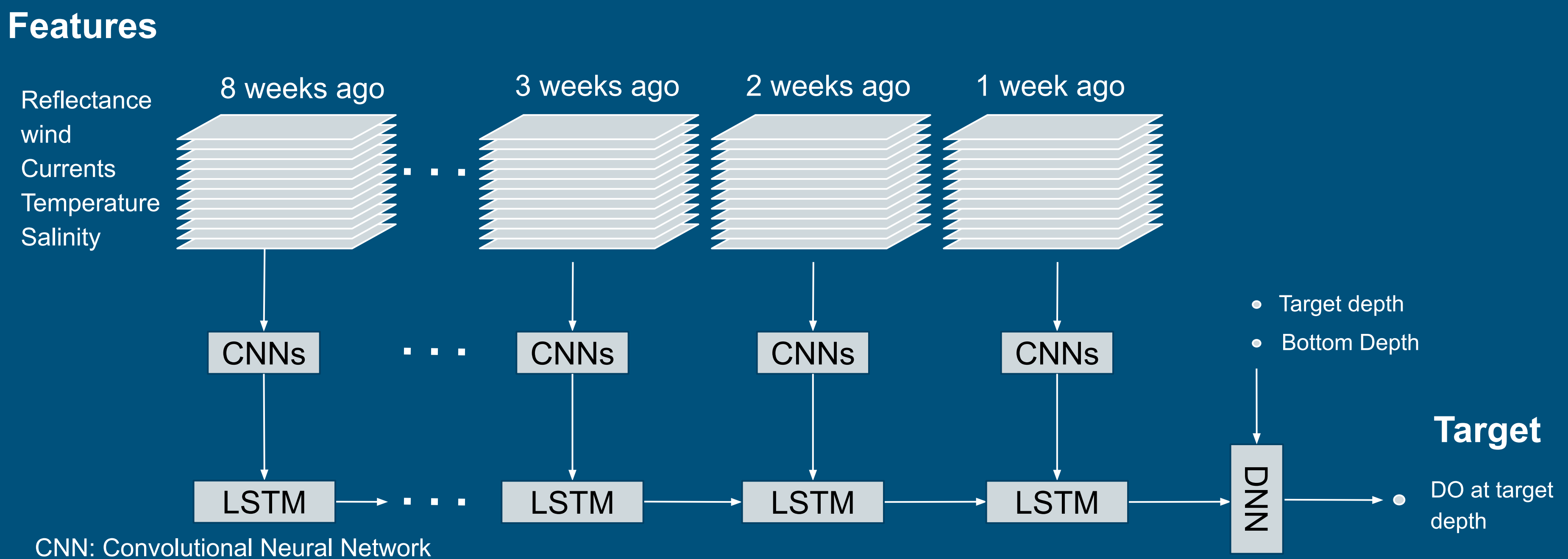
w_dir, wind direction

Training, Validation, and Test datasets

Split by year

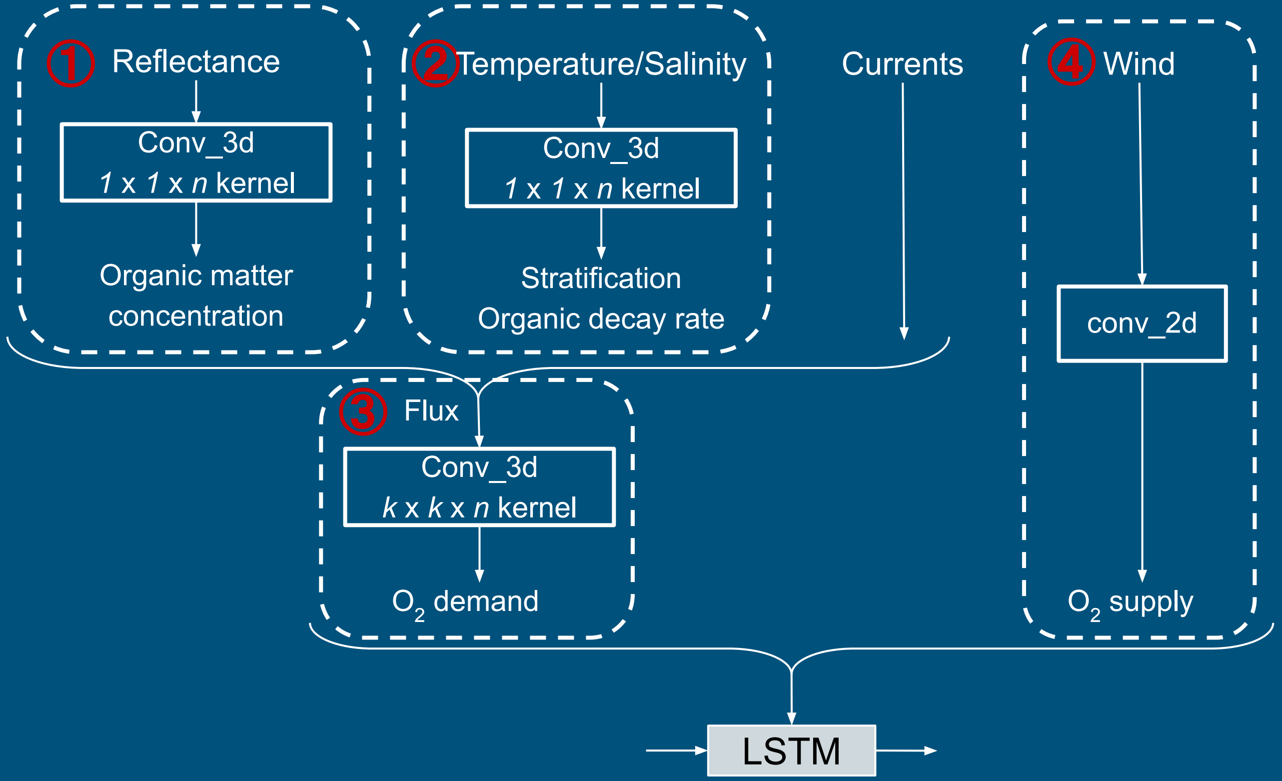
Splits	Year	Size	%
Training	2002-2017	140,636	86.7
Validation	2018	10,020	6.2
Test	2019-2020	11,618	7.2

Model Architecture - General Conceptual Flow



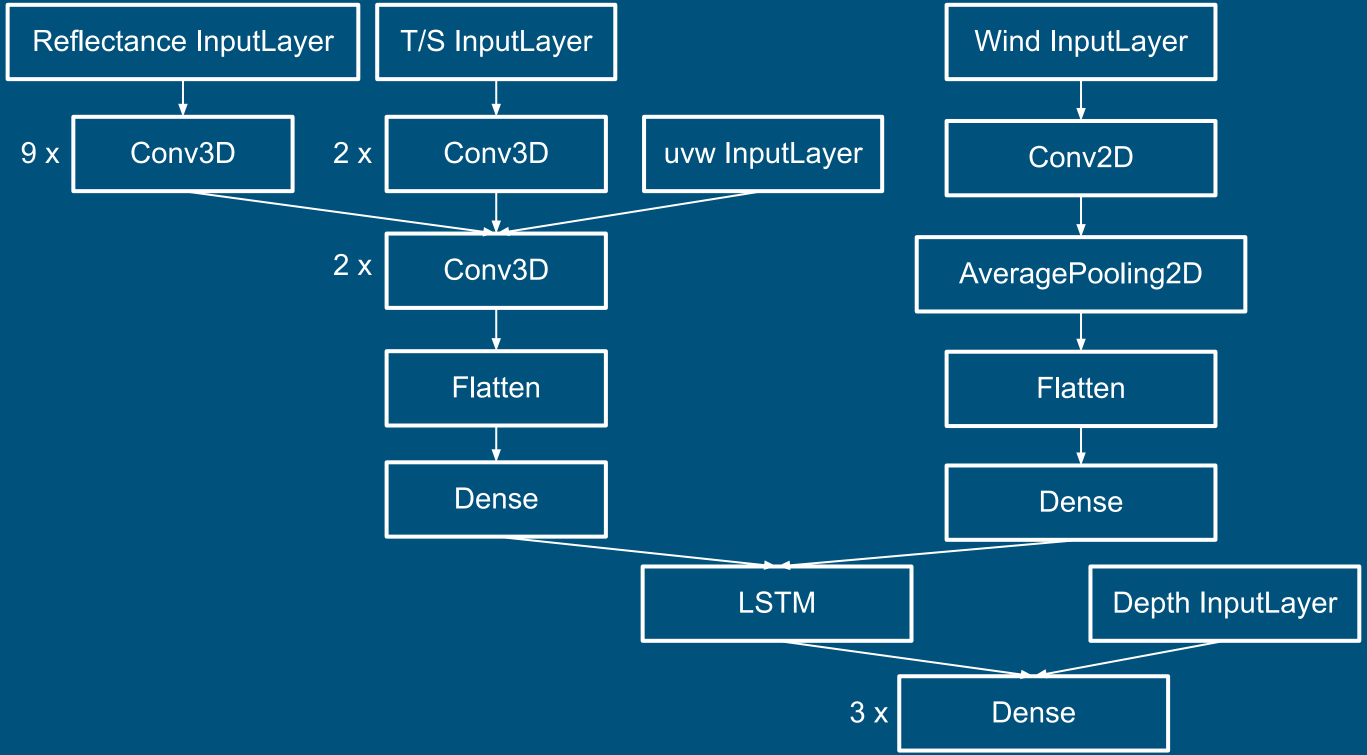
CNN: Convolutional Neural Network
LSTM: Long and Short Term Memory
DNN: Deep Neural Network

Model Architecture - Conceptual Flow at Each Time Step



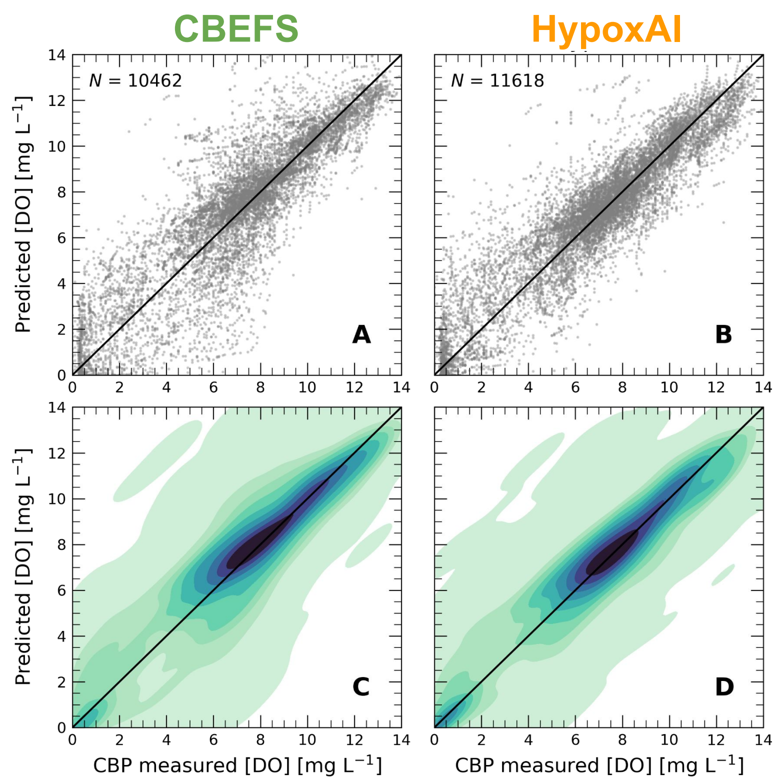
Module ①, Extract surface organic concentration
Module ②, Extract stratification and organic decomposition speed
Module ③, Combine organic flux and decomposition information to estimate oxygen demand
Module ④, Estimate oxygen supply from wind-mixing

Model Architecture - Final Full Model



- Hyperparameter optimization within a search space
- Total trainable params:
  22,234

Model Evaluation on Testset (2019-2020)

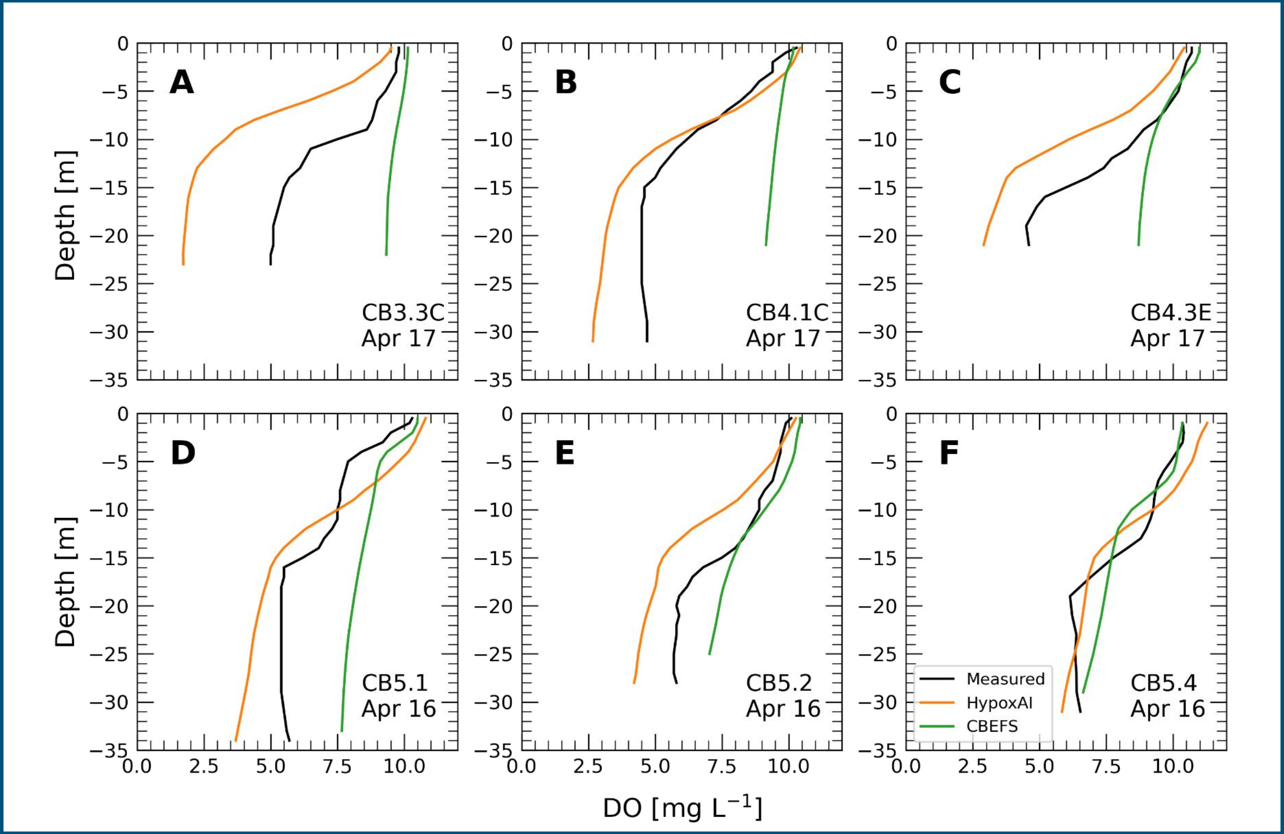


Metrics	RMSE	MAE	R^2	Slope	Bias	*N*
HypoxAI	1.255	0.905	0.849	0.868	0.119	11618
CBEFS	1.538	1.098	0.763	0.892	0.155	10462

- Significant improvement compared with CBEFS model

Comparison of DO Profiles in Spring

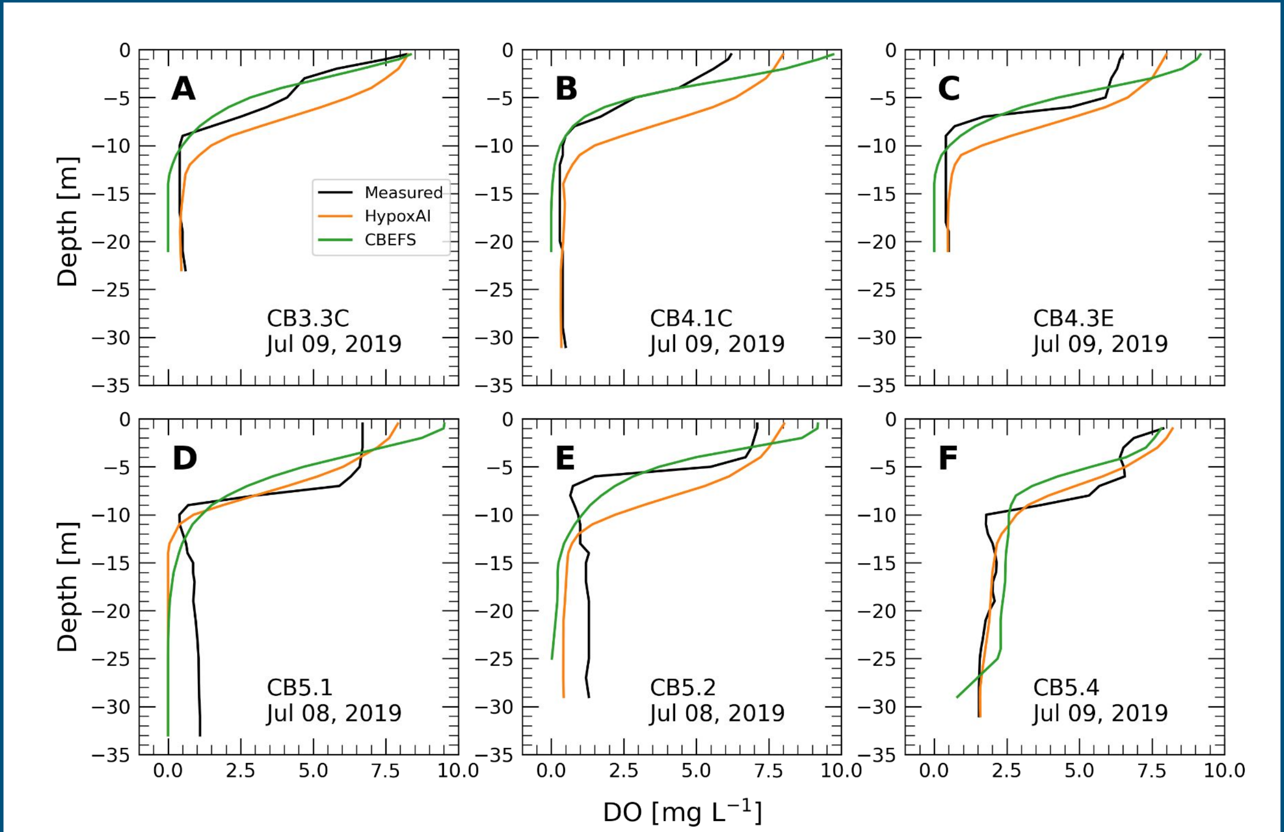
- HypoxAI-predicted profiles exhibit variable vertical shape and good agreement with measurements
- Measurement show abrupt changes in DO
- Predicted profiles are smooth and do not reproduce fine structures



All data from 2019

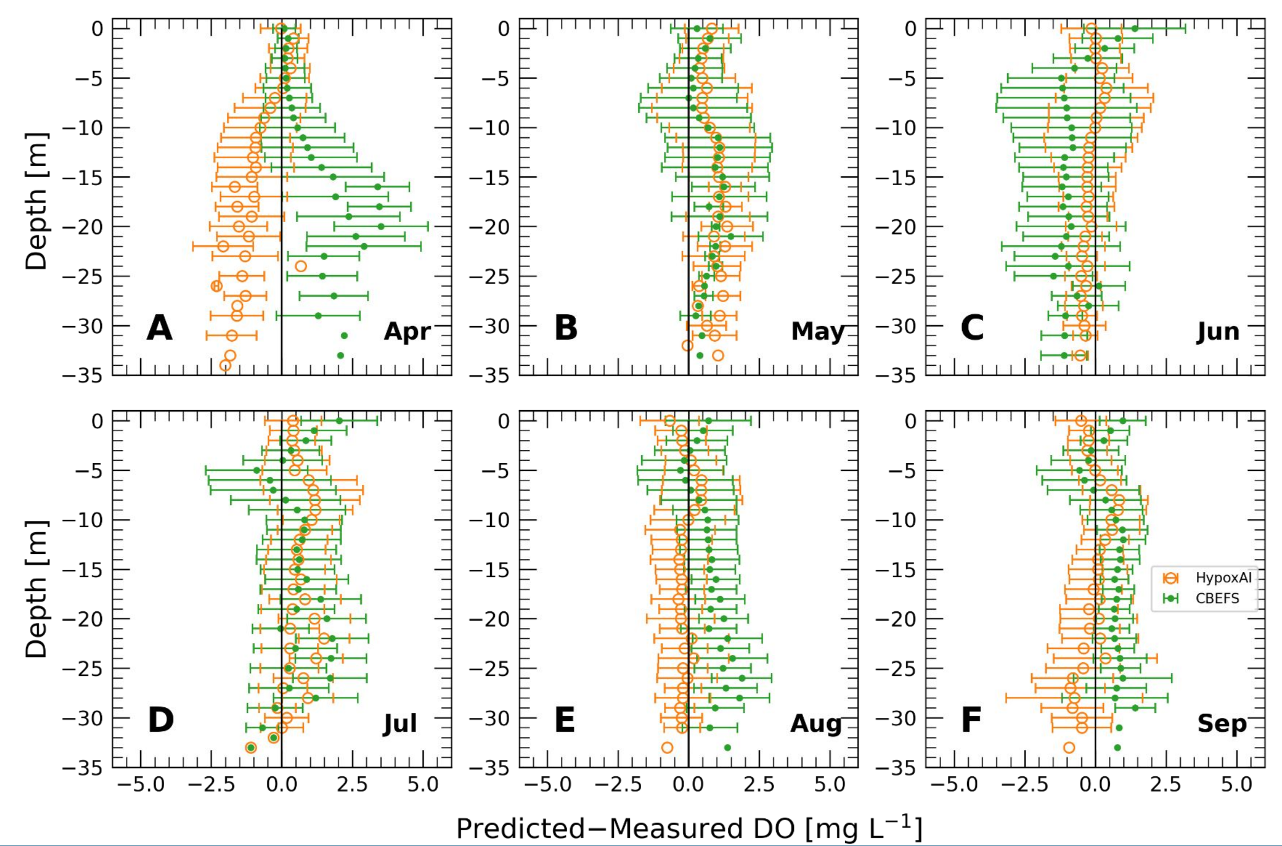
Comparison of DO Profiles in Summer

- HypoxAI does reproduce some stratification
- Measured profiles also show abrupt changes in DO which is hard for model to reproduce
- HypoxAI gives overall realistic DO profile

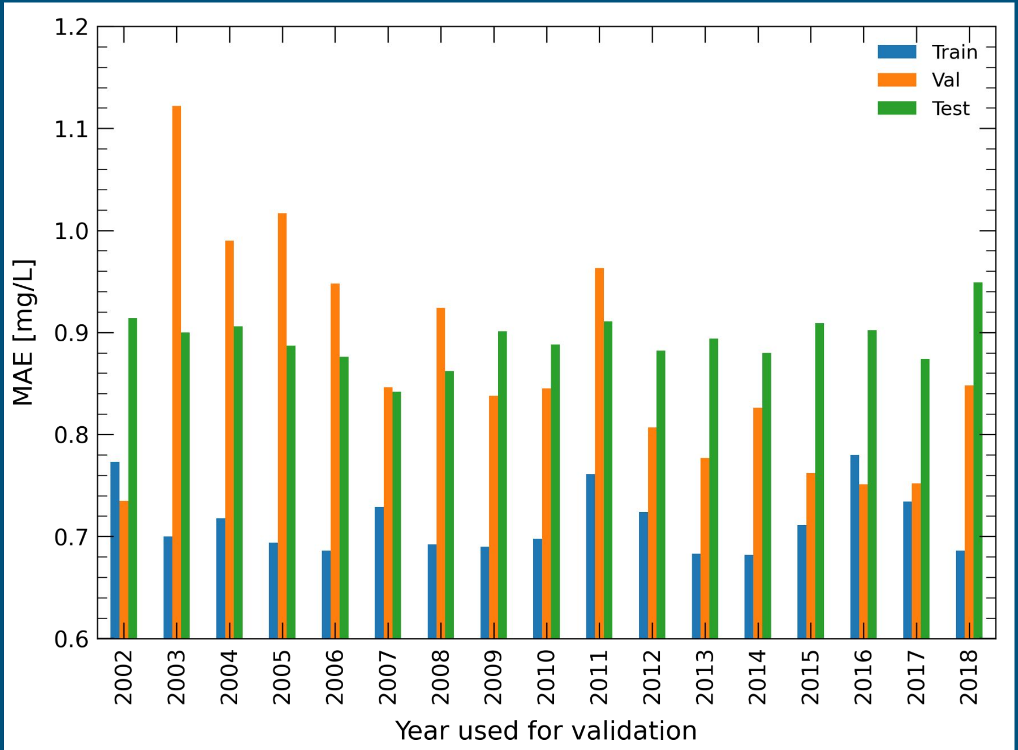


Monthly Difference between Predicted and Measured

- Accuracy varies with month
- Best overall performance in June and August
- Increased error at 5-10 m due to inability to predict sharp DO changes



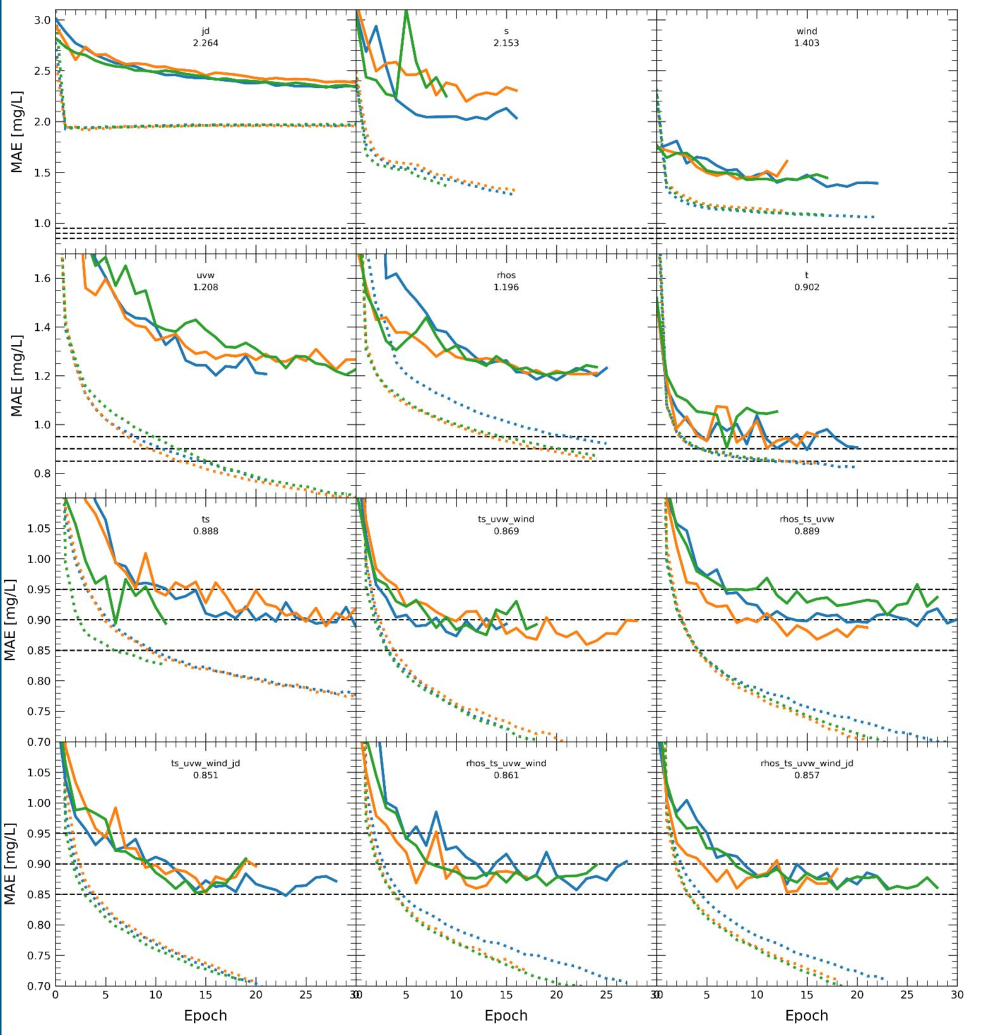
Cross validation using different year as validation



Ablation study

Component	Symbol	MAE
Baseline*		2.457
Julian day	jd	2.264
Salinity	s	2.153
Wind	wind	1.403
Currents	uvw	1.208
Ocean color	rhos	1.196
Temperature	t	0.902

* Sample and bottom depths only



Summary

- Developed a short-term hypoxia forecast model using deep learning which is capable of taking in features with disparate data structure
- HypoxAI achieves good overall agreement with measurements; produces reasonable vertical oxygen profiles; learned multiple types of vertical profiles
- Temperature is the strongest predictor of oxygen, followed by satellite-derived Rayleigh-correction reflectance, currents, wind, and etc.
- Future:
  - Use only satellite-derived sea surface temperature and ocean color as predictors
  - Improved satellite chlorophyll products

Acknowledgements

NASA Advanced Information Systems Technology Program (Grant AIST-18-0007)

Chesapeake Bay Program

Troy J. Ames (NASA GSFC)