UF Herbert Wertheim College of Engineering
UNIVERSITY of FLORIDA

POWERING THE NEW ENGINEER TO TRANSFORM THE FUTURE

Fusing Remote Sensing Data with Spatiotemporal *In Situ* Samples for Red Tide Detection

Ronald Fick

2/22/2022

Problem Statement

- Can we improve remote sensing detection of red tide by considering recent spatiotemporal *in situ* samples?

In Situ Data

- Red tide concentration data from FWC
- Almost 90k samples from 2000 to 2020
  - This sampling goes back to 1954, but MODIS-Aqua has only been collecting since 2002
  - Some sampling is in response to blooms, but since 1998 it has been more consistent
- Data from the Southwest coast of Florida

FWC Red Tide Samples from 2000-2020

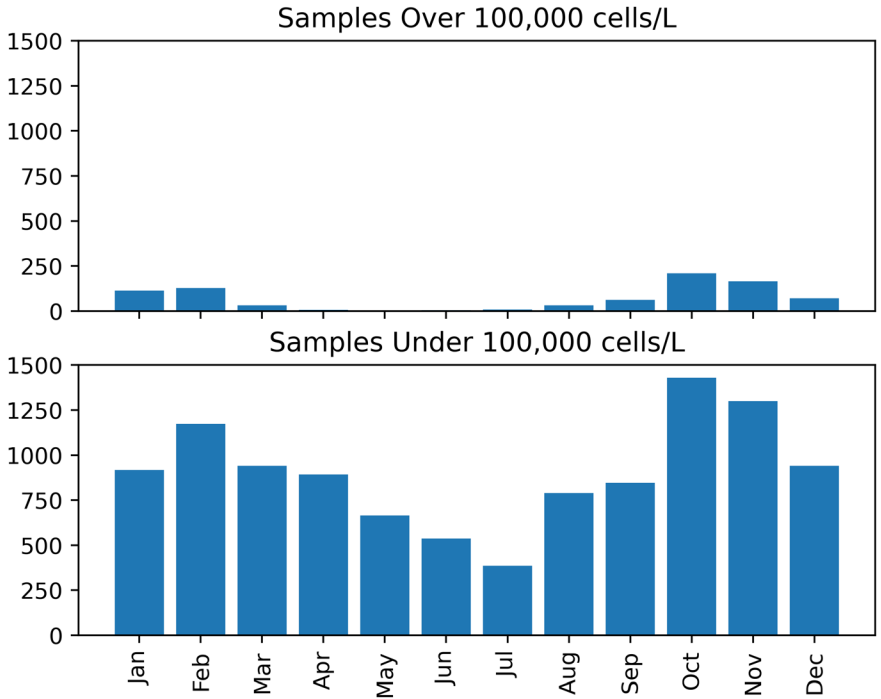
Remote Sensing Data

- We use MODIS-Aqua level 2 data for prediction
- The MODIS-Aqua sensor was chosen as it has been flying since 2002 so we have a long time series of data to consider for learning
- The MODIS-Aqua products have a 1 km spatial pixel resolution

Pairing of In Situ and Remote Sensing

- For each sample in the *in situ* dataset, it was paired with a remote sensing pixel if:
  - There was an image available from the same date as the sample
  - And the nearest pixel to the sample was within 1km and was not masked out due to cloud cover
- This gives a dataset of 11655 pixels with *in situ* red tide measurement.

Pairing of In Situ and Remote Sensing



Feature Selection

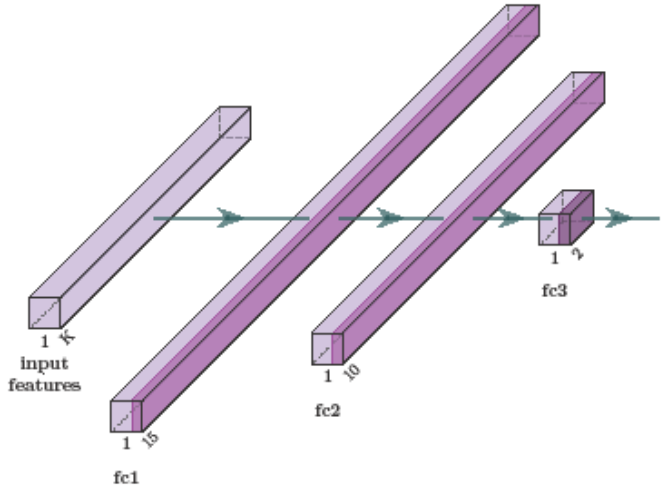
- Features were selected using recursive feature elimination (RFE)
- This method iteratively trains a classifier and removes the poorest performing feature according to some metric
- RFE was done using random forest classification with the gini coefficient being used to score features

Selected Features
par
Kd_{490}
$Chlor_a$
Rrs_{443}
Rrs_{469}
Rrs_{488}
nFLH

Features selected according to RFE

Neural Net Architecture

- The network architecture is shown on the right, it's a fully connected net with ReLU activations on the first 2 layers and softmax activation on the final layer
- This relatively shallow network was chosen as a result of the training set size



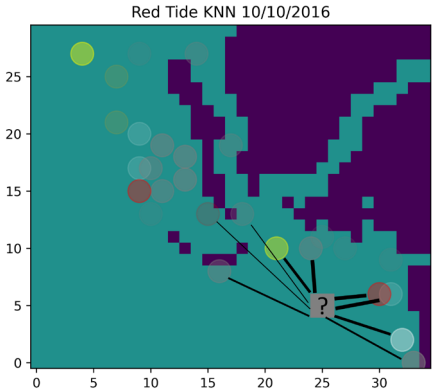
Neural Network Architecture

Machine Learning

- Neural networks were trained using the paired remote sensing/in situ data.
- The problem was set up as a two class classification/detection problem, with one class for samples with 100,000 cells/L or less (background, very low, and low in the FWC definitions), and the other class for samples over 100,000 cells/L (medium and high in the FWC definitions)
- All experiments were cross-validated by dividing the data into different train/test splits according to year

Weighted K-Nearest Neighbor Predictions

- In addition to the remote sensing information, we could consider information from recent ground samples
- Weighted KNN predicts the red tide concentration value at a test location and time according to a weighted summation of the red tide concentrations from the recent past/local area



Weighted-KNN Prediction with Recent Ground Truth

Weighted KNN Predictions

$$physical\ dist = \sqrt{(lat - lat_i\ )^2 + (lon - lon_i\ )^2}$$

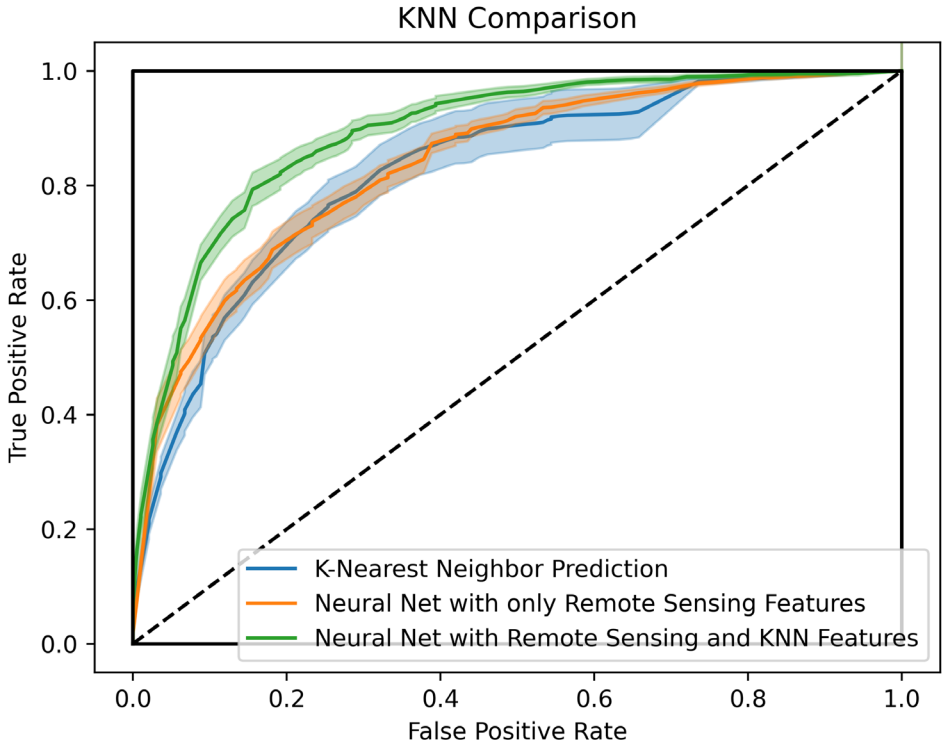
$$spatiotemporal\ dist = physical\ dist + \beta * (date - date_i)$$

$$inverse\ dist = \frac{1}{spatiotemporal\ dist}$$

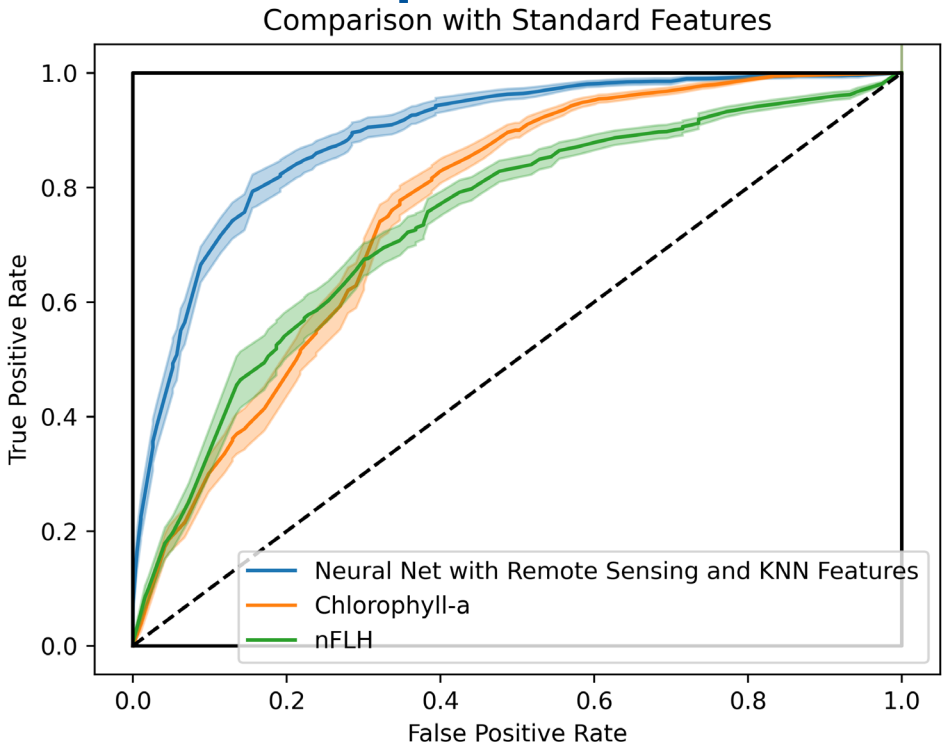
$$weight_i = \frac{inverse\ dist_i}{\sum_i inverse\ dist_i}$$

$$estimated\ concentrations = \sum_i weight_i * logConc_i$$

ROC Results – KNN Estimation



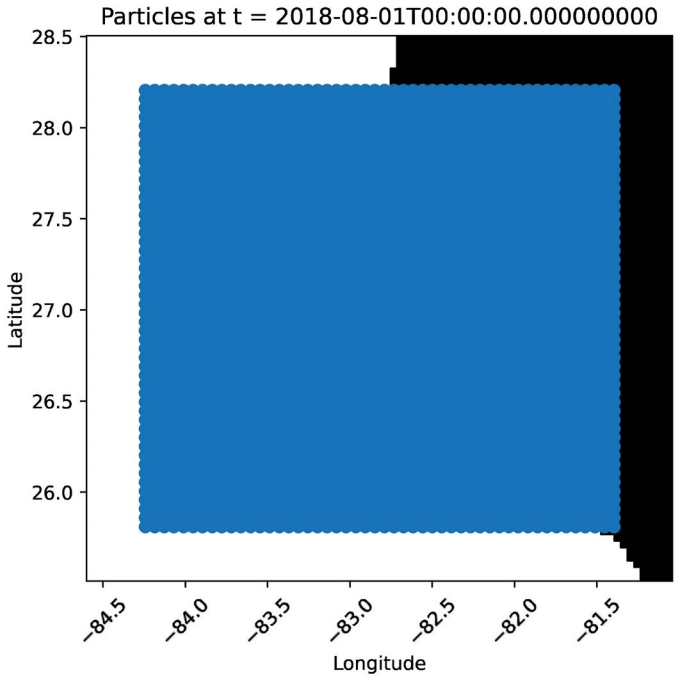
ROC Results – Comparison



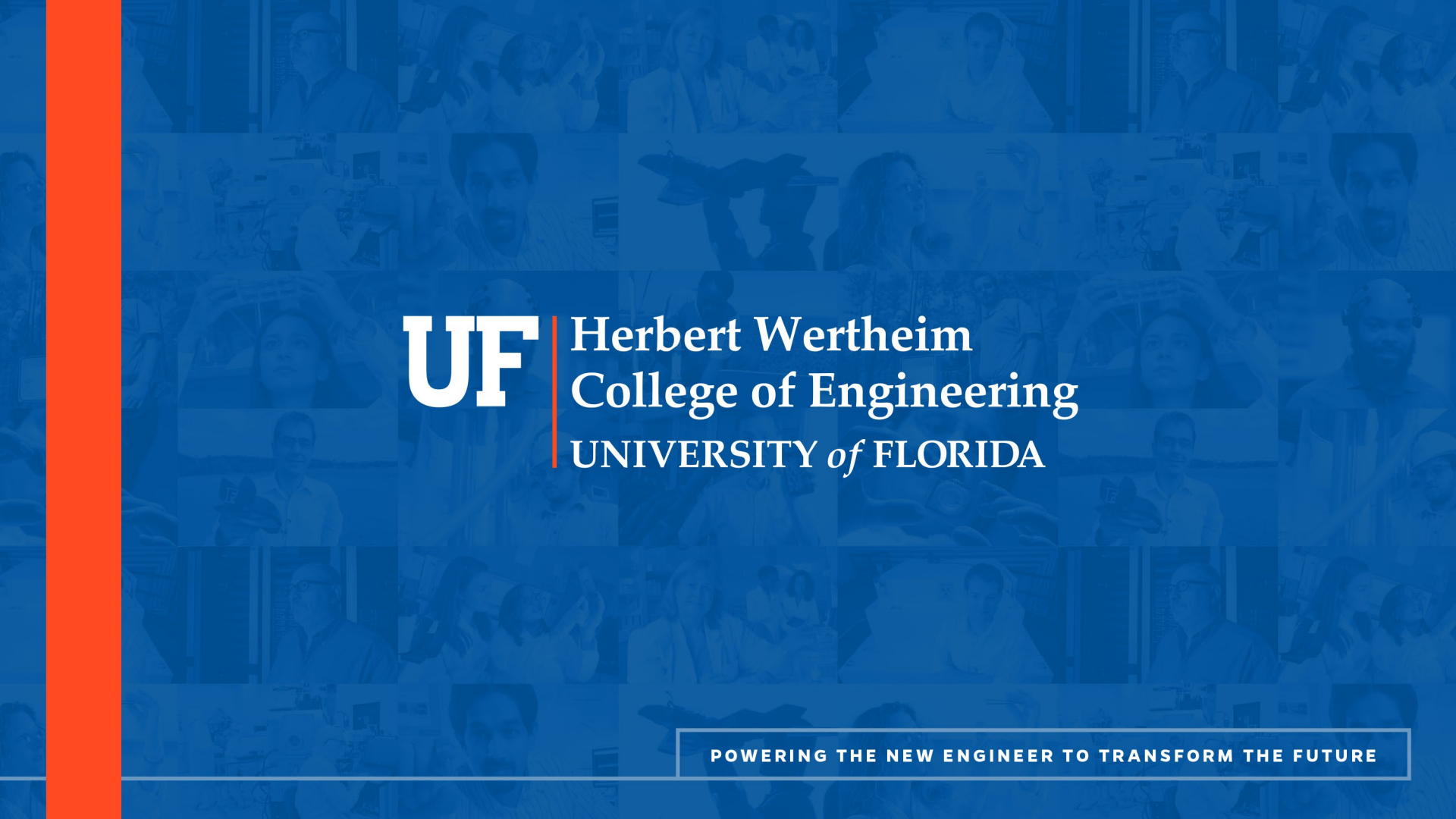
COMPUTING FOR LIFE
THE MACHINE LEARNING AND SENSING LABORATORY
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING UNIVERSITY OF FLORIDA

Future Directions

- The distance measure used in the KNN approach is a crude estimate of the relationship between locations in time and space
- We are considering building in current information to better understand the influence of past samples on prediction



Particle Tracking Simulation for 08/2018


UF
Herbert Wertheim
College of Engineering
UNIVERSITY of FLORIDA
POWERING THE NEW ENGINEER TO TRANSFORM THE FUTURE