Title : Forecast of green tide distribution in the Yellow Sea

Shiying Lin and Liang zhao

Laboratory of Marine Environment and Numerical Simulation, College of Marine and Environmental Sciences, Tianjin University of Science and Technology, Tianjin,300000

1. Aim: This study aimed 1) to forecast the distribution of green tide in the Yellow Sea from 2006 to 2017; 2) to reveal the main environmental factors affecting the distribution of green tide.

2. Methods: The distribution data of green tide were collected from the Moderate Resolution Imaging Spectroradiometer (MODIS) provided by the National Aeronautics and Space Administration (NASA) from 2006 to 2017. In this study, the latitude range of study region was approximately 31.50°N-37.50°N and the longitude range was 119.00°E-124.50°E (Fig. 1). To avoid the effect of spatial autocorrelation on the accuracy of the model, we randomly removed the points whose distance between two points was less than 0.02°.

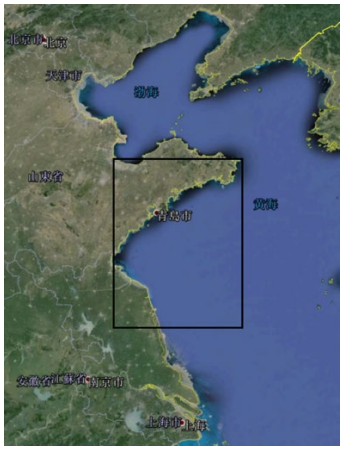

Fig .1 The range of study region

The environmental factors in this study are derived from the Regional Ocean Model System (ROMS) and the 5th major atmospheric reanalysis (ERA5) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). The mean, maximum and minimum values and standard deviation were calculated as environmental factors, respectively.

To avoid the influence of multicollinearity and the MaxEnt model overfitting of these variables, Pearson correlation analysis was performed on these variables in Statistical Product and Service Solutions (SPSS). When the correlation coefficient of two

variables exceeded 0.80, the variable with higher significance was reserved. In total, 12 environmental variables were maintained in the MaxEnt model (Table 1) for distribution of green tide.

Table .1 Environmental variables

Code	Name	Unite
Gt	Green tide position	
V10min	Min 10m northward velocity	m/s
Po_4mean	Mean phosphate	$mmol/m^3$
V10max	Max 10m northward velocity	m/s
U10min	Min 10m eastward velocity	m/s
U10min	Max 10m eastward velocity	m/s
sstmean	Mean surface temprature	K
No_3mean	Mean nitrate	$mmol/m^3$
ssrdmean	Min surface solar radiation downwards	J/m^2
dep	depth	m
tpmean	Mean temperature	K
saltmean	Mean salinity	

In this study, we intend to use data from 2012 and 2016 for validation and other years for model learning. To forecast the dispersion of the green tide, MaxEnt 3.4.1 was utilized. 75% of the occurrence data are chosen for training and 25% are chosen for testing in MaxEnt once the green tide distribution data and the filtered environmental variable data have been imported. The MaxEnt model then creates a continuous map with an occurrence probability that ranges from 0 to 1. The jackknife test was used to determine the contribution rate and significance of parameters. Model performance was assessed using the receiver operating characteristic curve (ROC) and the area under the ROC curve (AUC), with the AUC value ranging from 0 to 1. More accurate prediction outcomes are indicated by an AUC value that is nearer to 1. The model's performance was rated as failing (0.5-0.6), bad (0.6-0.7), fair (0.7-0.8), good (0.8-0.9), and exceptional (0.9-1.0).

3. Result: Maxent performed well in generating the potential distribution of green tide, with all mean AUC greater than 0.850 (Table 2). The top three environmental factor included in the prediction of the MaxEnt model that impacted green tide distribution were green tide position in the first half of the month, 10m eastward velocity, and salinity. The parameter with the highest gain when applied in isolation was green tide position. Therefore, green tide position appeared to have the most useful information by itself. In

summary, the key parameters that limited the selection of green tide distribution were green tide position in the first half of the month, 10m eastward velocity, and salinity.

Table 2. AUC of model

Year	AUC
2008	0.940
2009	0.852
2010	0.959
2011	0.957
2012	0.991
2013	0.972
2014	0.964
2015	0.931
2016	0.879

The training results of the model are shown in Fig. 2. The distribution of the trained green tide matches well with the actual distribution, indicating that the model is well trained and can be used for forecasting. The distribution of green tide in the first half of July 2008 was mainly north of the mouth of the Yangtze River, near Rizhao and Qingdao, and the green tide occupied a medium-sized area. The green tide in 2009 was widely distributed, with the northern border of Weihai, the southern border of northern Yancheng, and the eastern border extending to the middle of the Yellow Sea, and the distribution area of the green tide was the largest in this decade. The green tide in 2011 and 2013 were less frequent and concentrated north of the Yangtze River mouth, near Rizhao and Qingdao. The distribution of green tide in 2014 showed a small portion of green tides in the south compared to previous years, and green tides in the north remained widely distributed. The 2015 green tide was also present in the south, while the northern green tide remained widespread and close to land.

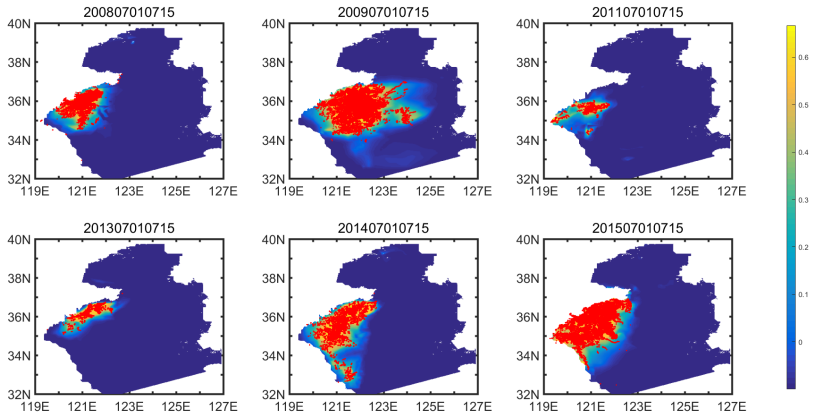

Fig .2 Green tide distribution in the first half of July, the color filled graphs are model predictions, red dots are measured values

We looked at how the environment affected green tide locations in previous years and forecast where they would go in 2012, as shown in Fig 3. The forecasts for 2014 and 2015 are more in line with the current circumstances. The possibility of a distribution in 2008 is low, and a more southerly distribution of the green tide is predicted. Predictions from 2009 and 2011 don't match up with reality. Although there is a slight match between 2013 projections and the actual dispersion of the green tide, the overall result is still inconsistent.

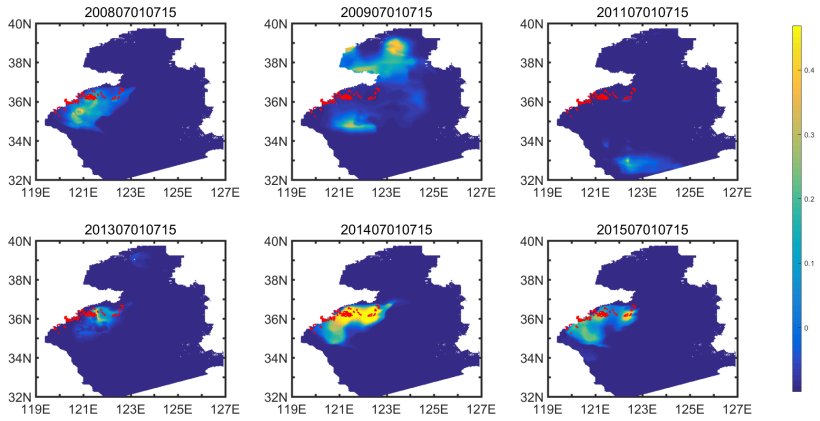

Fig .3 Predicted distribution of green tide in the first half of July 2012, the color filled graph is the predicted value, the red dots are the actual distribution in 2012

We also investigated the connection between environment and the position of green tides in prior years, and we forecast the location of the green tide in 2016 as shown in Fig 4. When used to forecast the location of green tides in 2016, the model did not do well. The spread of the 2016 green tide was less anticipated by the 2008 and 2013 training settings, which did not overlap. The Bohai Sea and the Yellow Sea along the coast make up the majority of the 2016 green tide distribution anticipated by the training environment in 2009. Predicted locations in the Yellow Sea section agree with the actual distribution, although the chance of the green tide is low. The majority of the southern portion of the study area was where the distribution of the 2016 green tide was predicted by the training environment in 2011, 2014, and 2015. Although the predicted distribution of the green tide was very large, it did not correspond to the actual distribution of the green tide,

making the prediction unsatisfactory.

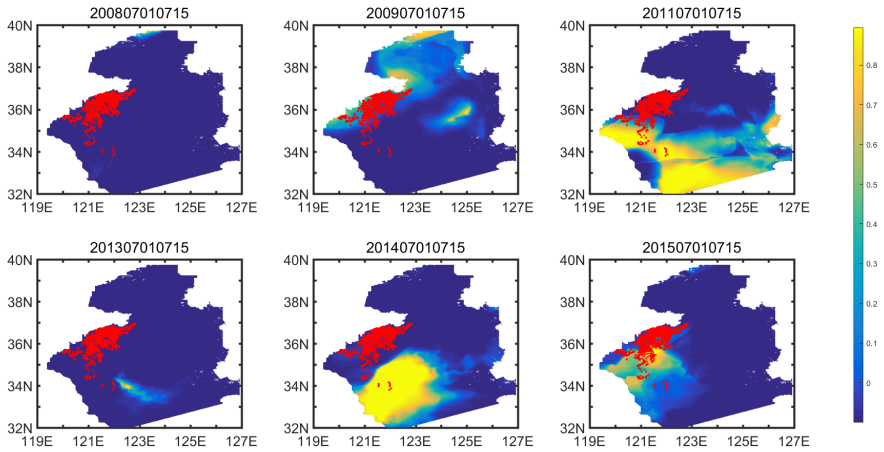

Fig. 4 Predicted distribution of green tide in the first half of July 2016, the color filled graph is the predicted value, the red dots are the actual distribution in 2016

4. Discussion: While the MaxEnt model did a good job of simulating current distribution, the findings for the prediction of green tide were less successful, which may have been due to environmental conditions. Green tide location in the first half of the month, 10m eastward velocity, and salinity were the key determinants of green tide dispersion. More than 75% of the distribution of green tide may be attributed to green tide position in the first half of the month; thus, it can be argued that this position will dominate future green tide distribution, with salinity and 10m eastward velocity having less of an impact. The MaxEnt model prediction is too dependent on green tide position in the first half of the month, which is detrimental to the prediction effect.

4. Conclusion: Green tide is mostly found north of the Yangtze River mouth, close to Rizhao and Qingdao, with the worst green tide episodes occurring in 2009, 2014, and 2015. The green tide cannot be accurately predicted by the existing model.

5 Perspectives in future: In order to enhance the accuracy of model, we will reevaluate and choose more representative environmental elements.