Supplementary Information

Supplementary information for the paper “Automated VIIRS boat detection based on machine learning and its application to fisheries monitoring in the East China Sea”

S1. Supplementary methods

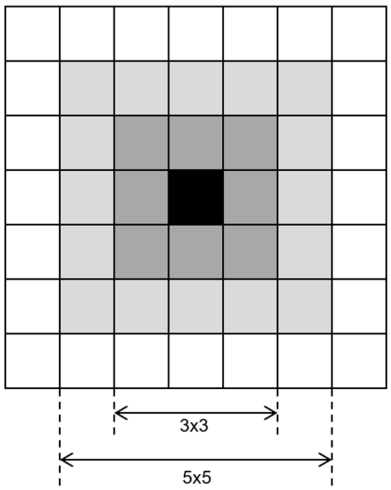

Figure S1: Schematic representation of the surrounding pixels. The black pixel indicates the target pixel and dark gray and pale gray pixels indicate the 3x3 and 5x5 range, i.e., $s = 3$ and $s = 5$, respectively. The larger surrounding pixels include a smaller range of pixels as their pixels, i.e., 5x5 surrounding pixels is a combination of pale gray, dark gray and black pixels. While not shown in the illustration, the maximum 9x9 surrounding pixels are considered in the algorithm.

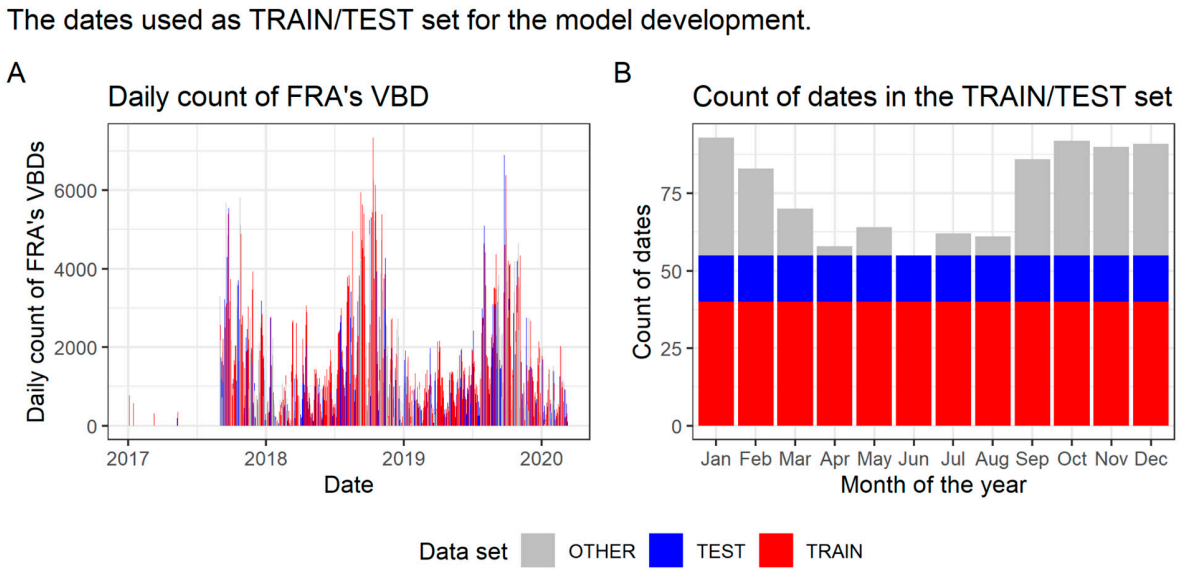

Figure S2: The dates used as TRAIN/TEST data set for developing the model. A: Daily count of FRA's VBD used in this study. The red indicates the dates used for model training and the blue indicates dates for testing the model performance and the dates shown in gray were not used for training and testing.

S2. Supplementary results

Variable importance

The result of the variable importance (permutation feature importance) for the production model (Figure S3) suggests that the radiance of the pixel (log10_rad_nano) is the most important variable to identify vessels from the background followed by the spike median index with broader surrounding areas (SMI9, SMI7, SMI5) as well as variables indicating moon reflection (luz and moon_illum_frac). The cloud masks are determined to be the least important features.

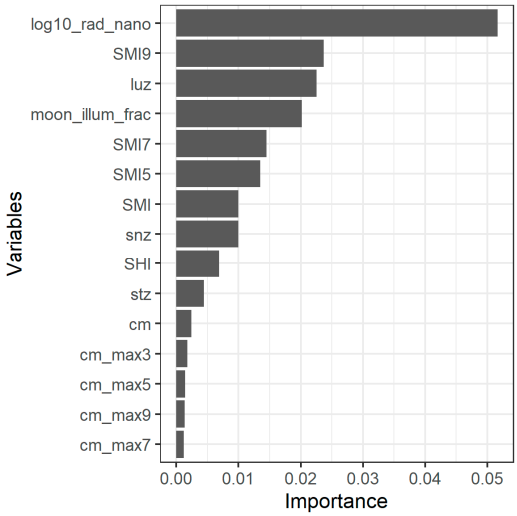

Figure S3: Variable importance of the production model.

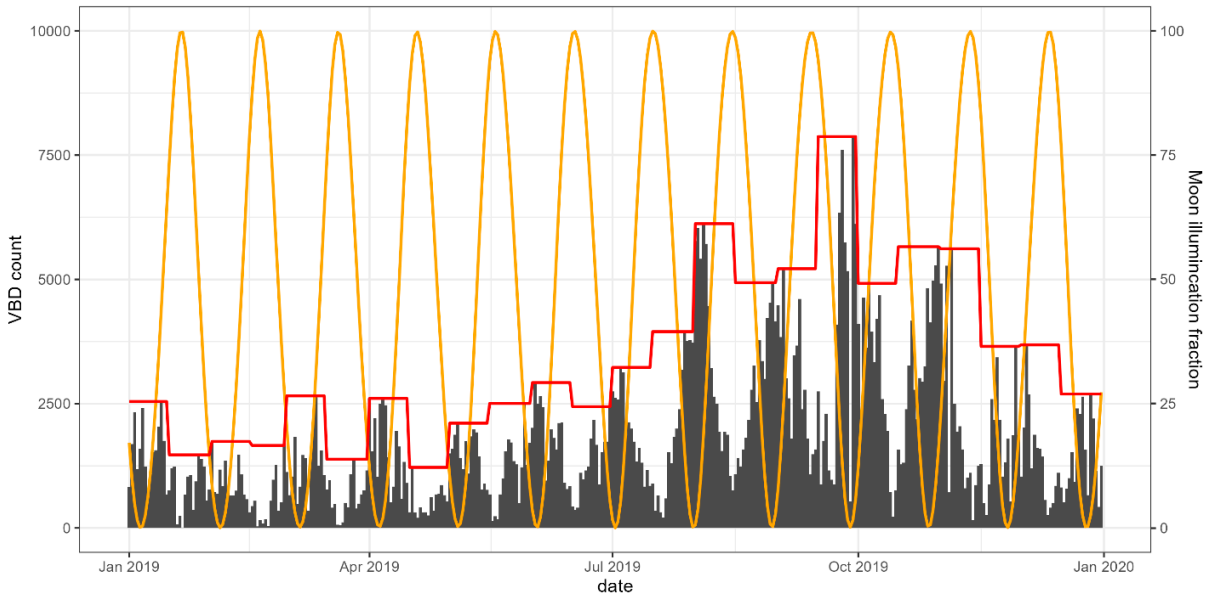

Figure S4: The relationship between the number of VBD count and moon phase. Black bar indicates the daily count of the present VBD and the red line indicates the maximum daily count of VBD for each half-month period. Orange line indicates moon phase (100 means full moon and 0 means new moon). Although only the result of the present VBD algorithm (NEW) is shown here, the other VBD algorithms (FRA, EOG) also have the same tendency.

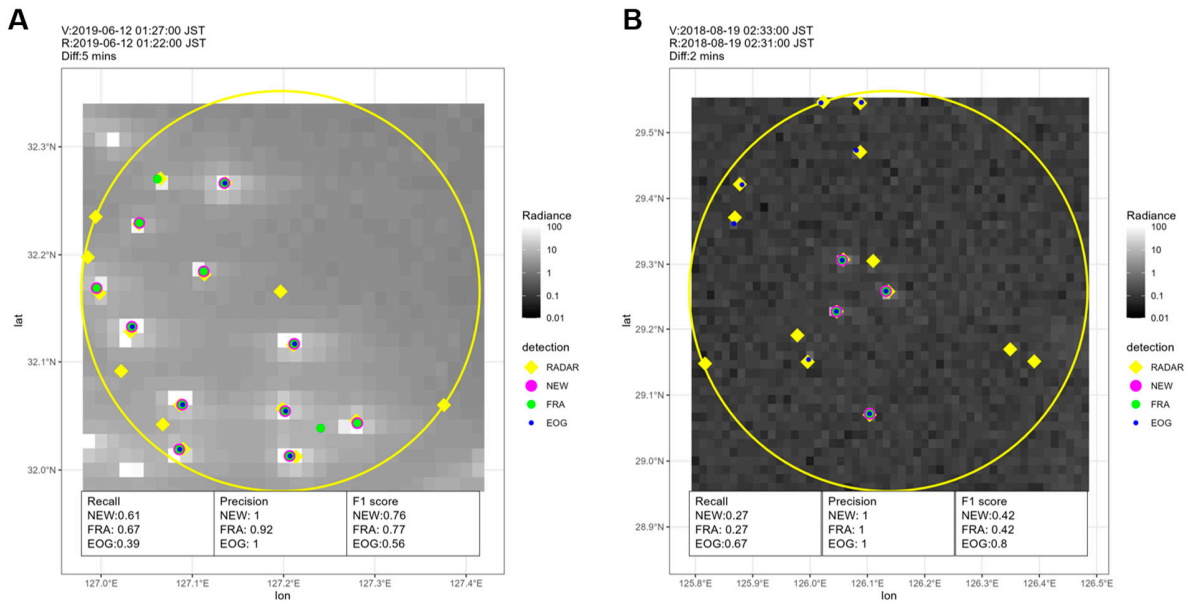

Figure S5: Example images of on-ship radar detection and VBDs. A: The example of the case when the NEW and FRA detect more than EOG under cloudy conditions. **B**: The example of the case when the EOG detect more than NEW and FRA under clear and no-moon conditions. The datetime shown in the title part of the image indicated by V and R is for VIIRS detection and on-ship radar detection respectively, and it's time differences in minutes are also shown. Yellow diamonds indicates on-ship radar detection. The points indicated by magenta, green and blue represents VIIRS detection from our's (NEW), FRA's and EOG's algorithm, respectively. The value of performance metrics (Precision, Recall and F1 score) for each VBD algorithms against on-ship radar data on this night is also shown in the text box. DNB radiance shown in the background is averaged for each 0.01 degree grid for the ease of visualization.

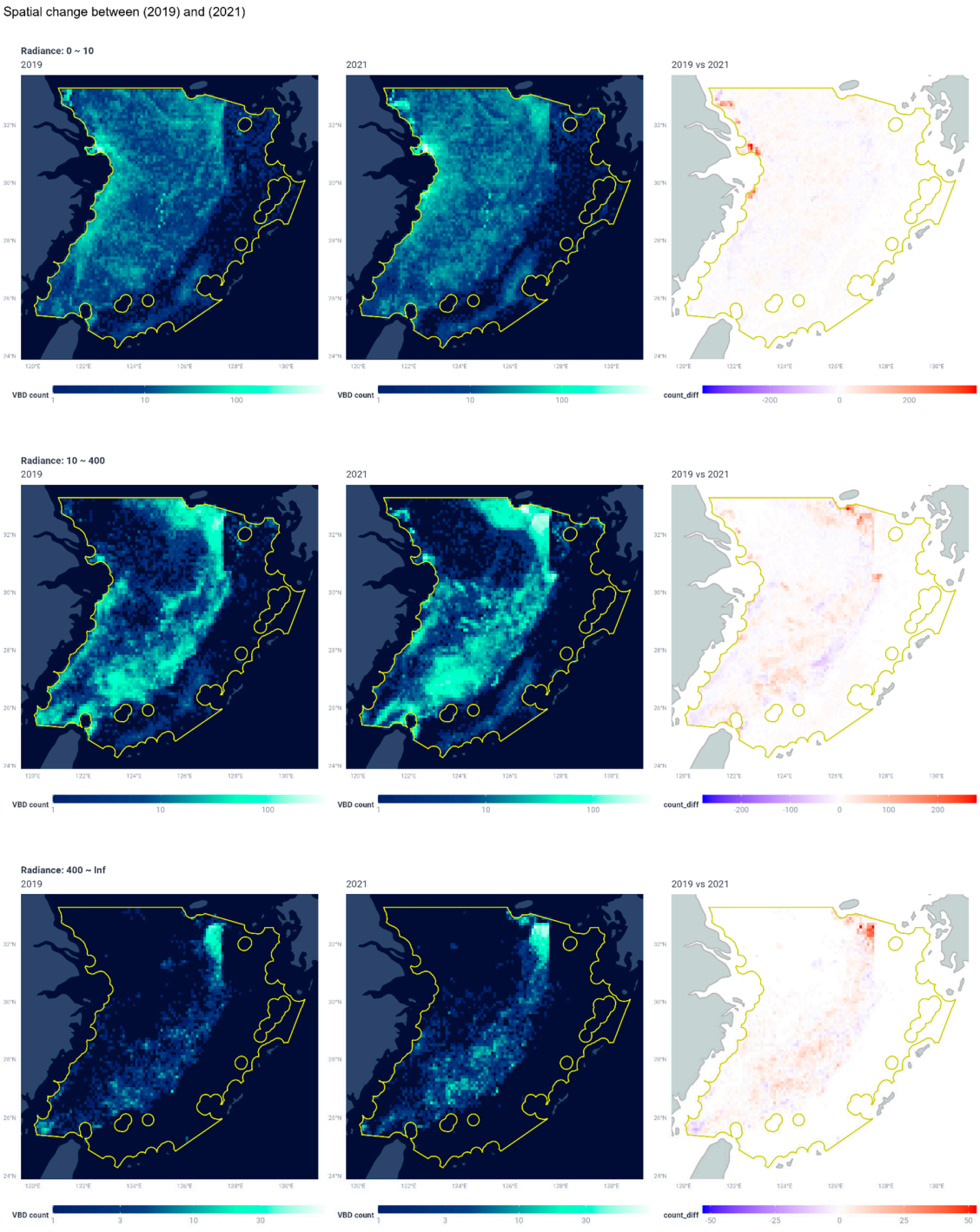

Figure S6: Comparison of spatial distributions of VBDs between 2019 and 2021. The spatial changes of VBD counts between 2019 and 2021. Each horizontal row of panels represents different radiance classes. Left most column of panel indicates spatial distribution of VBD in 2019 and the middle column indicates those in 2021. Right most column indicates the

differences of counts between 2019 and 2021. Red color indicates increase and blue indicates decrease in the VBD count.

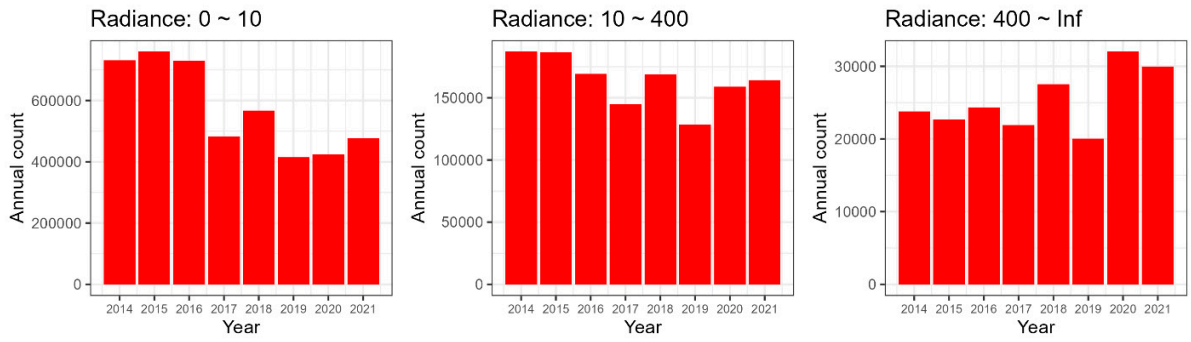

Figure S7: The annual aggregate of the EOG's VBD count in the ECS for each radiance class. The red bars indicate the annual summation of the maximum number of daily count of VBD for each half-month period.

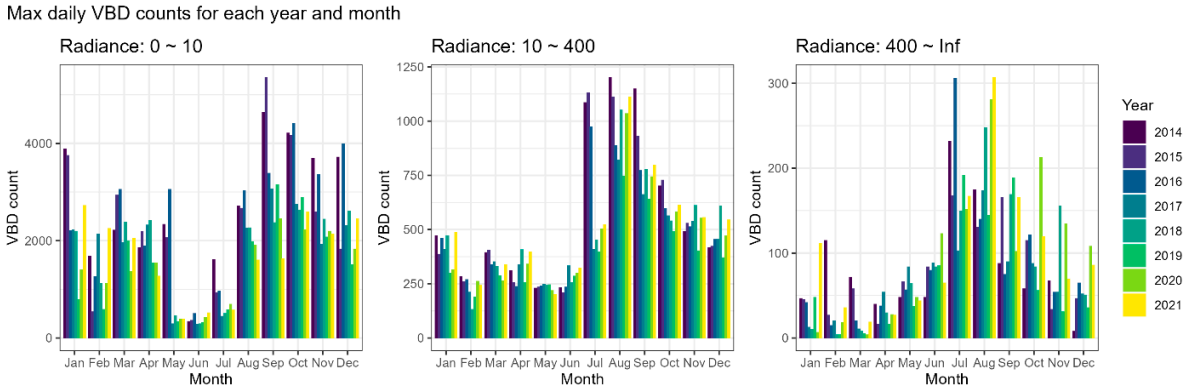

Figure S8: The maximum daily count of EOG's VBDs for each year and month in the study area. Each panel indicates the different radiance classes and colors indicates the different year, the horizontal axis indicates the months of the year, and the vertical axis indicates the maximum number of VBD observed in each year and month.