Supplementary Information (SI) for Analytical Methods.

Identification of Dyes on Fabric Exposed to Lake and Ocean Waters Using Near-Infrared Excitation Raman Spectroscopy

Claire Sasaki[1], Shannon Bober[1], Aidan P. Holman[1] and Dmitry Kurouski*[1]

1. Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843, United States

Supporting Information

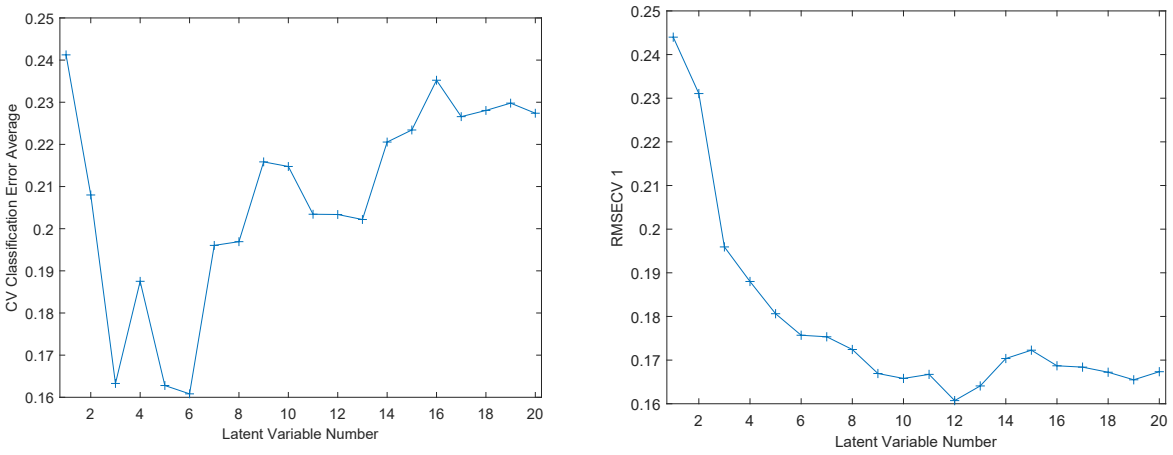


Figure S1. Cross-Validation (CV) Error (left) and Root Mean Square Error of Cross-Validation (RMSECV) Plots (right) of PLS-DA model used to predict faring stages of magenta-colored cotton exposed to ocean water (Figure 3, A). Based on these plots, 4 latent variables (LVs) were chosen.

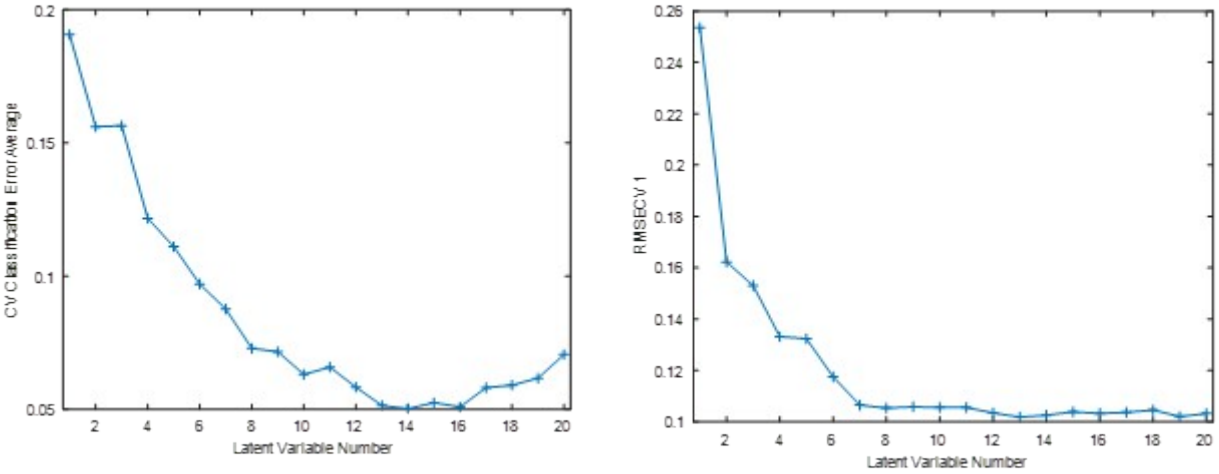


Figure S2. Cross-Validation (CV) Error (left) and Root Mean Square Error of Cross-Validation (RMSECV) Plots (right) of PLS-DA model used to predict faring stages of magenta-colored cotton exposed to lake water (Figure 3, B). Based on these plots, 2 latent variables (LVs) were chosen.

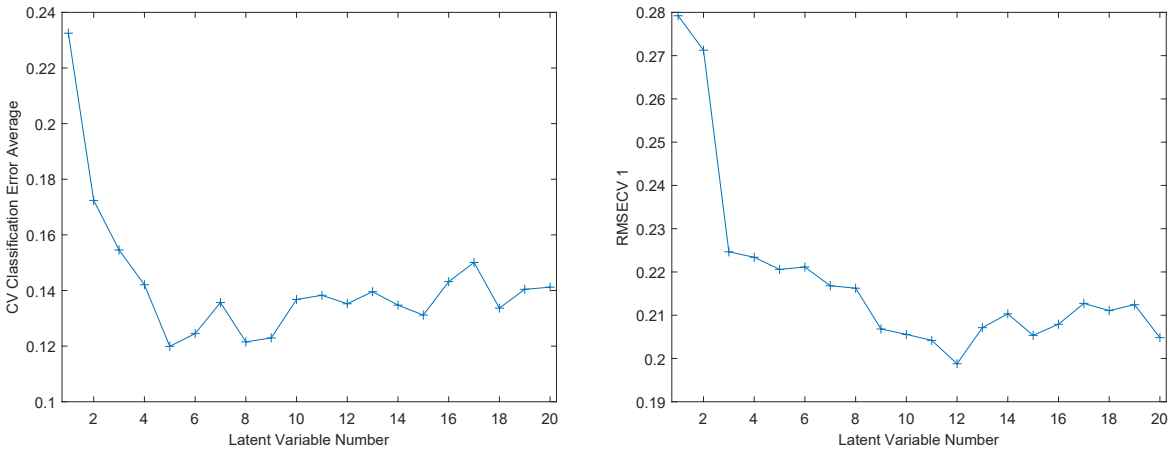


Figure S3. Cross-Validation (CV) Error (left) and Root Mean Square Error of Cross-Validation (RMSECV) Plots (right) of PLS-DA model used to predict faring stages of blue-colored cotton exposed to ocean water (Figure 5, A). Based on these plots, 5 latent variables (LVs) were chosen.

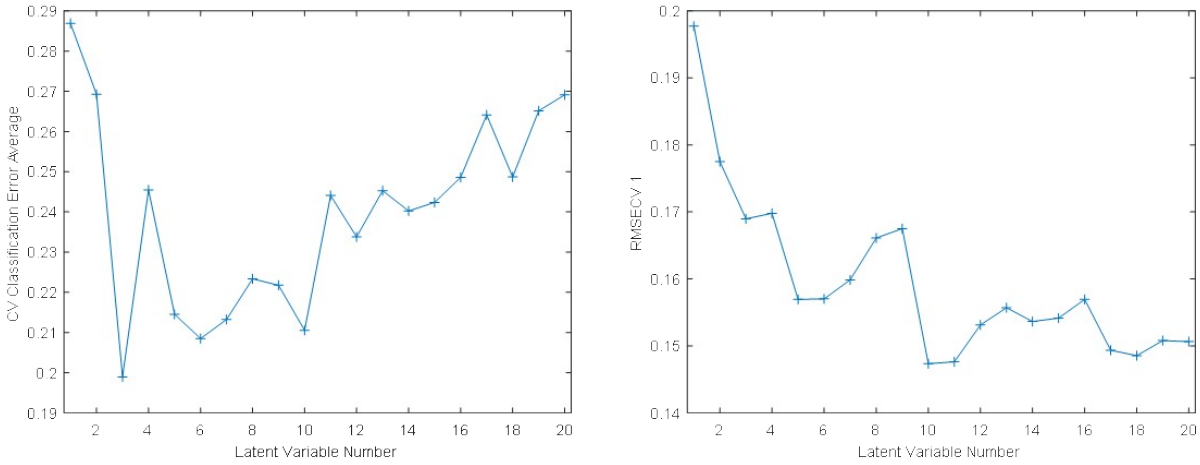


Figure S4. Cross-Validation (CV) Error (left) and Root Mean Square Error of Cross-Validation (RMSECV) Plots (right) of PLS-DA model used to predict faring stages of blue-colored cotton exposed to ocean lake (Figure 5, B). Based on these plots, 5 latent variables (LVs) were chosen.

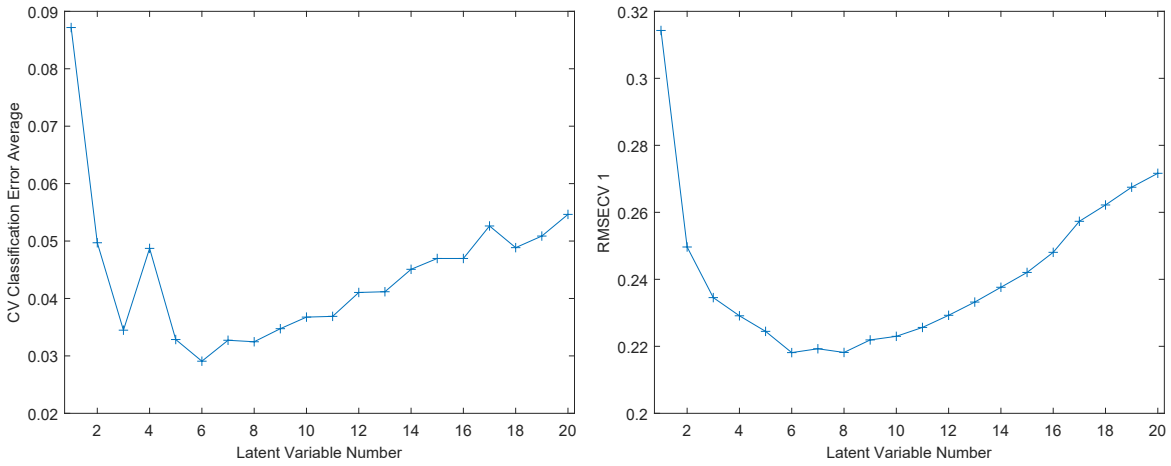


Figure S5. Cross-Validation (CV) Error (left) and Root Mean Square Error of Cross-Validation (RMSECV) Plots (right) of PLS-DA model used to predict dyes on colored cotton exposed to ocean and lake water (Table 2). Based on these plots, 6 latent variables (LVs) were chosen.

Figure S6. Magenta-colored fabric at week 0 (left) and week 14 (right) after exposure to ocean water.

Figure S7. Magenta-colored fabric at week 0 (left) and week 14 (right) after exposure to lake water.

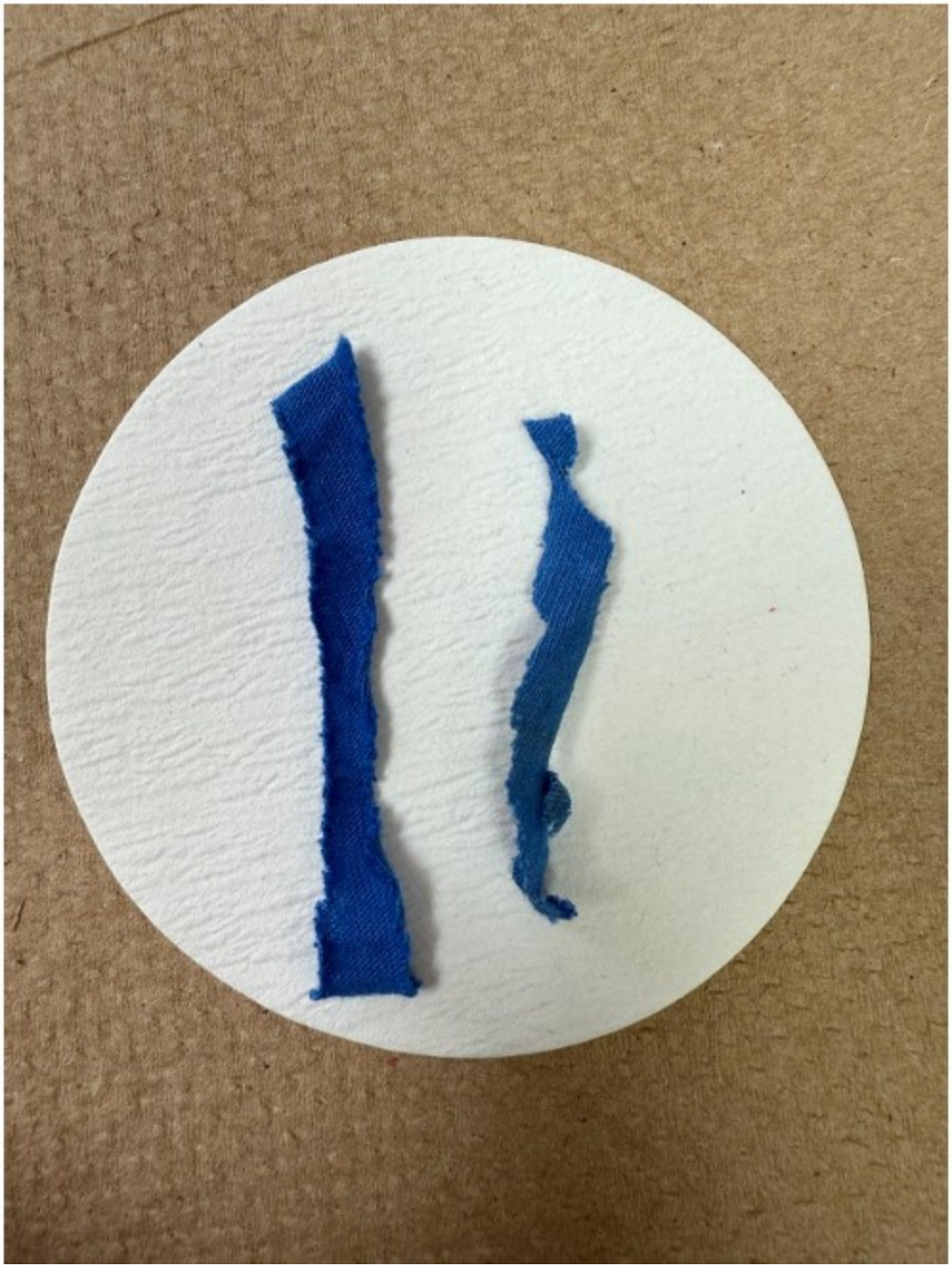
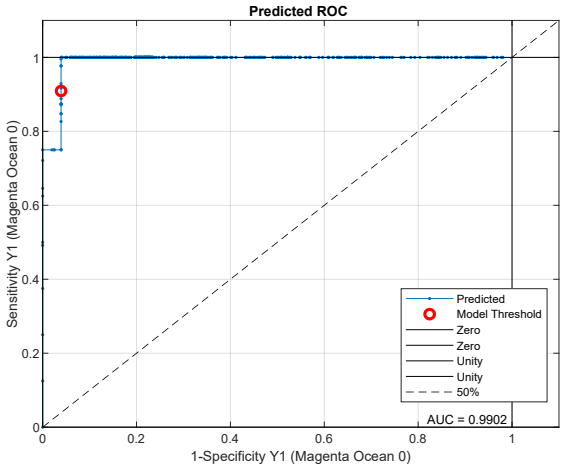
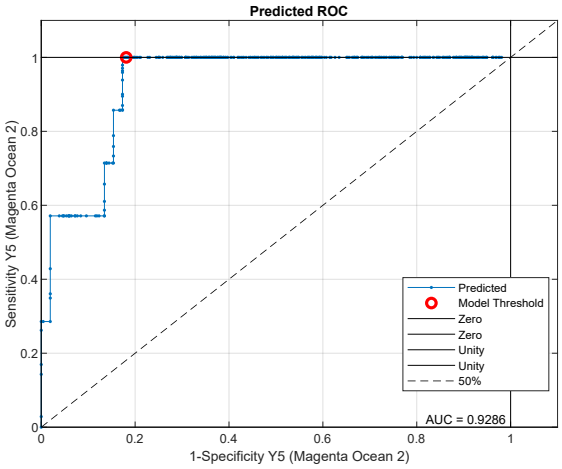
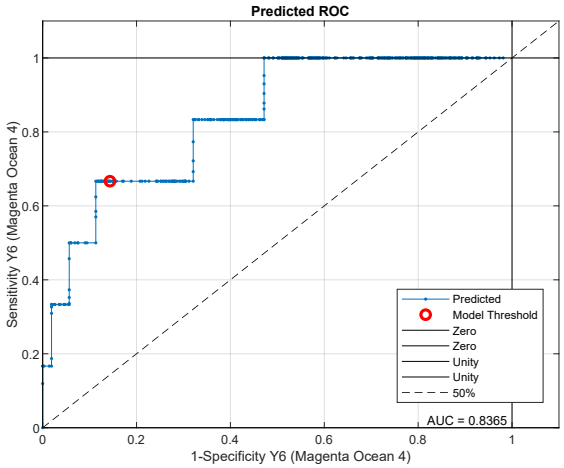


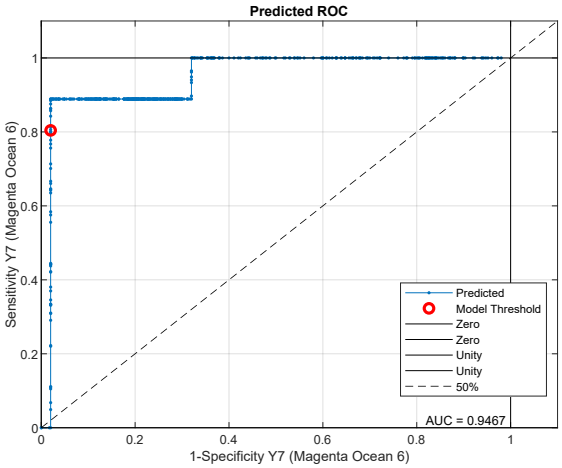
Figure S8. Blue-colored fabric at week 0 (left) and week 14 (right) after exposure to ocean water.

Figure S9. Blue-colored fabric at week 0 (left) and week 14 (right) after exposure to lake water.


Predicted ROC
Sensitivity Y1 (Magenta Ocean 0)
1-Specificity Y1 (Magenta Ocean 0)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.9902


Predicted ROC
Sensitivity Y5 (Magenta Ocean 2)
1-Specificity Y5 (Magenta Ocean 2)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.9286


Predicted ROC
Sensitivity Y6 (Magenta Ocean 4)
1-Specificity Y6 (Magenta Ocean 4)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.8365


Predicted ROC
Sensitivity Y7 (Magenta Ocean 6)
1-Specificity Y7 (Magenta Ocean 6)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.9467

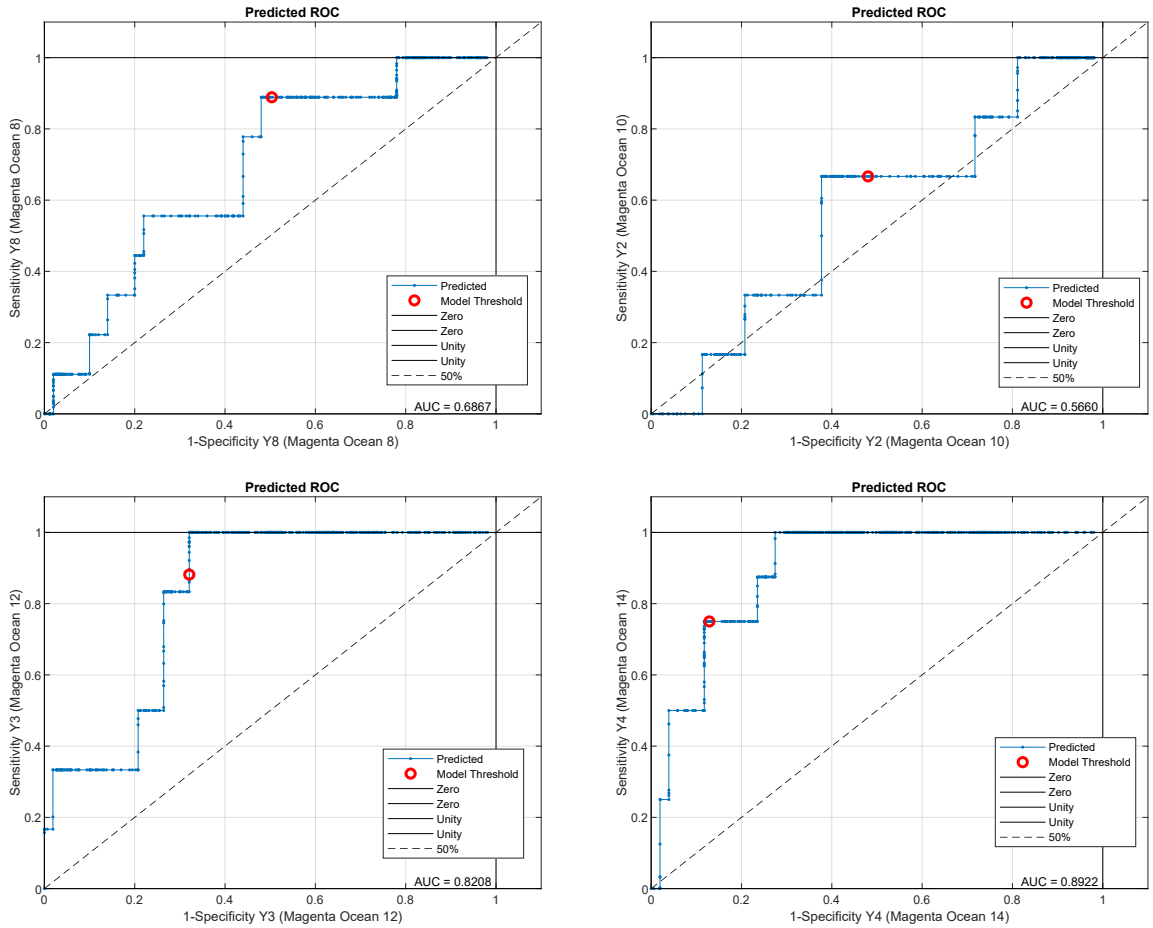
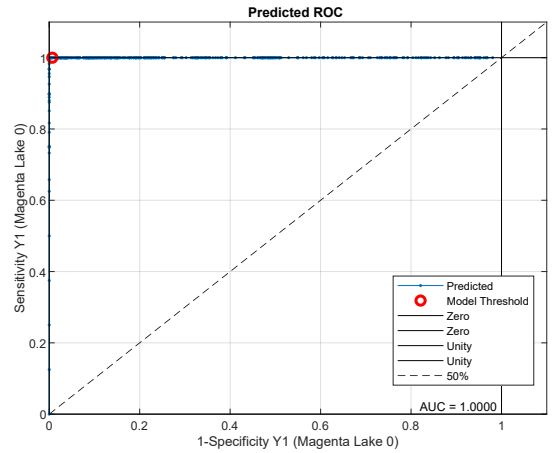
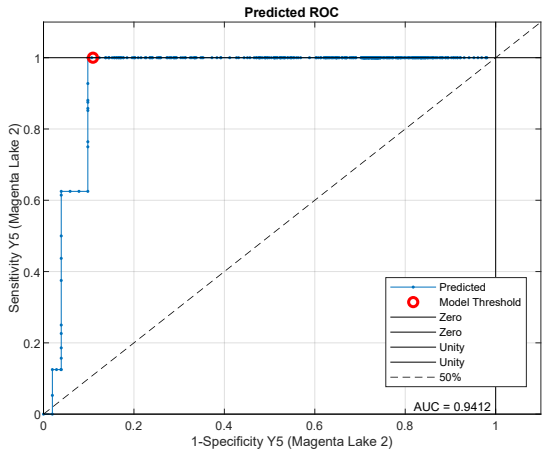
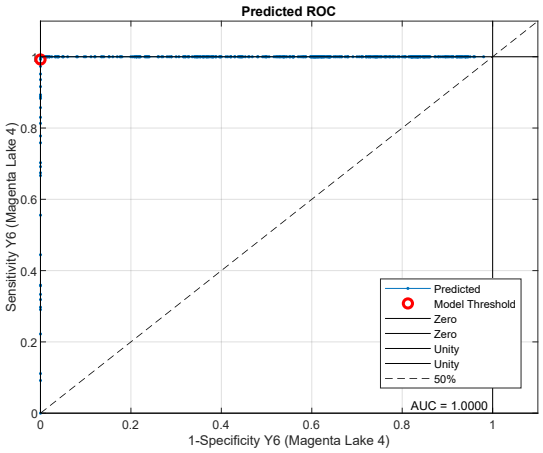
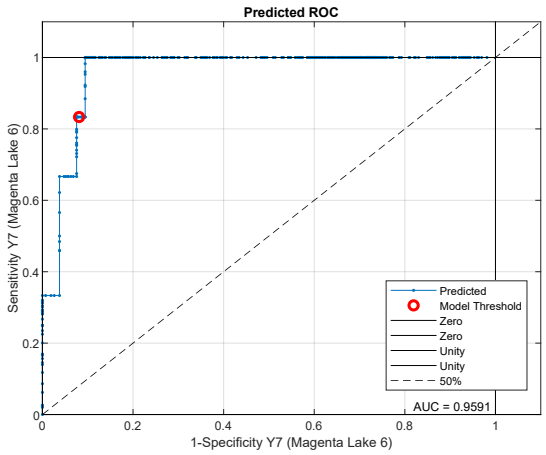


Figure S10. ROC plots for PLS-DA model used to predict faring stages of magenta-colored cotton exposed to ocean water (Figure 3, A).


Predicted ROC
Sensitivity Y1 (Magenta Lake 0)
1-Specificity Y1 (Magenta Lake 0)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 1.0000


Predicted ROC
Sensitivity Y5 (Magenta Lake 2)
1-Specificity Y5 (Magenta Lake 2)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.9412


Predicted ROC
Sensitivity Y6 (Magenta Lake 4)
1-Specificity Y6 (Magenta Lake 4)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 1.0000


Predicted ROC
Sensitivity Y7 (Magenta Lake 6)
1-Specificity Y7 (Magenta Lake 6)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.9591

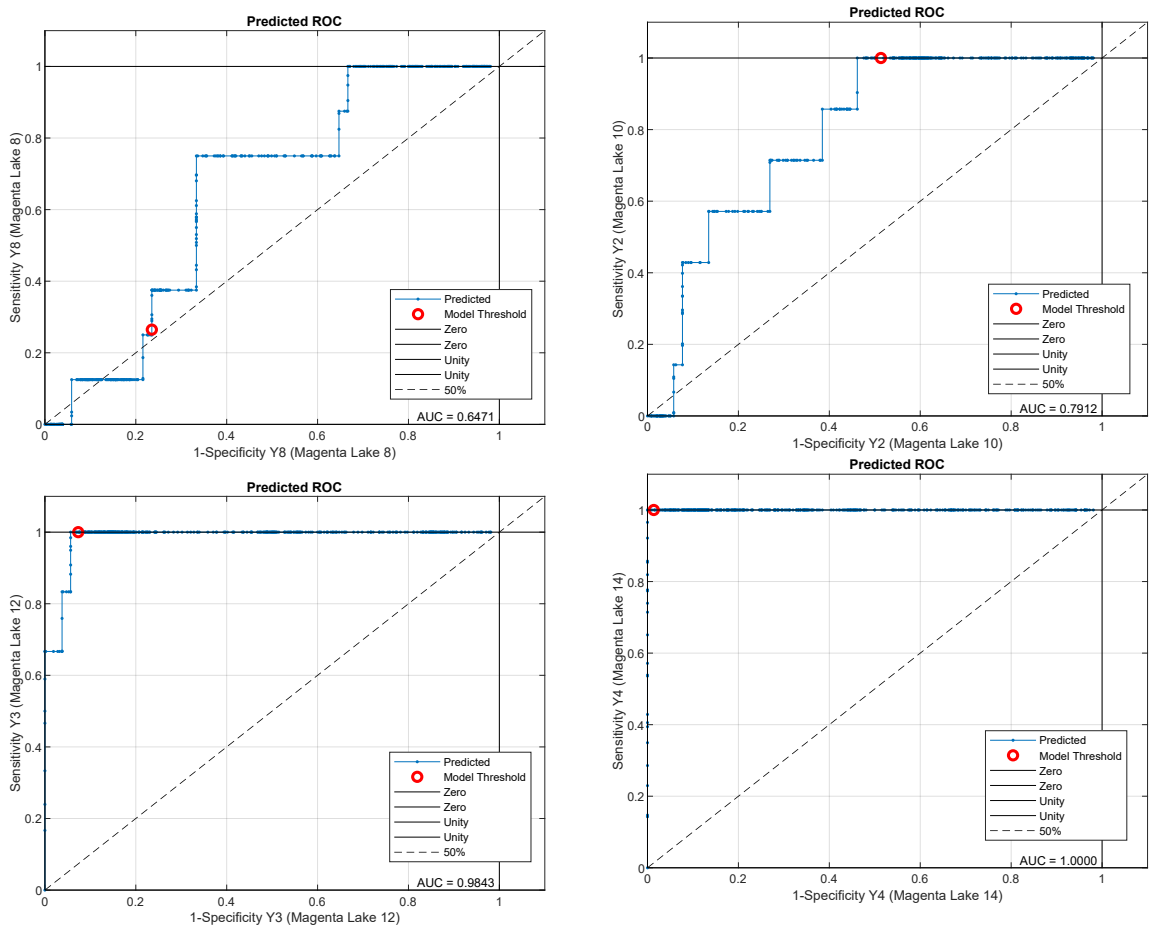
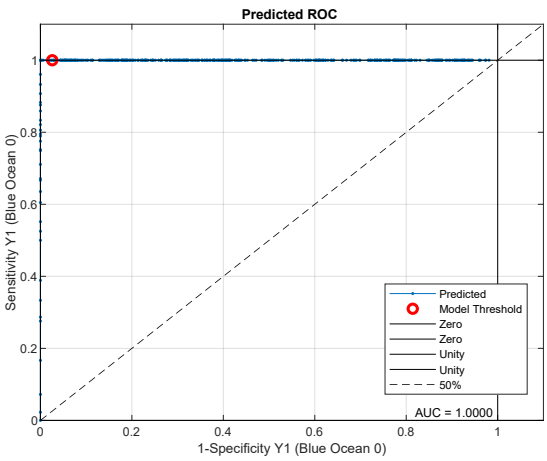
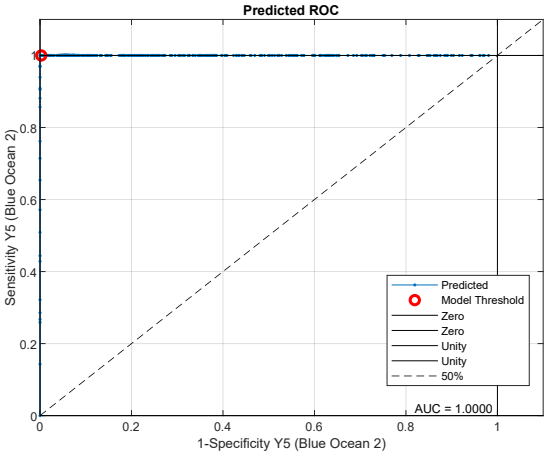
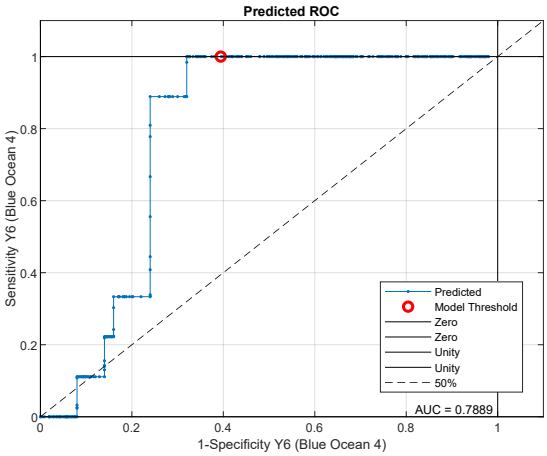
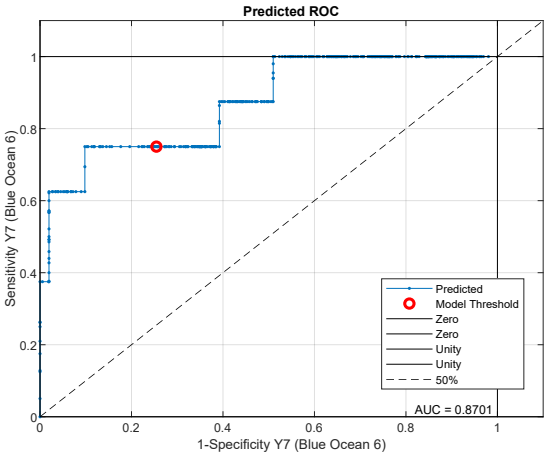


Figure S11. ROC plots for PLS-DA model used to predict faring stages of magenta-colored cotton exposed to lake water (Figure 3, A).


Predicted ROC
Sensitivity Y1 (Blue Ocean 0)
1-Specificity Y1 (Blue Ocean 0)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 1.0000


Predicted ROC
Sensitivity Y5 (Blue Ocean 2)
1-Specificity Y5 (Blue Ocean 2)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 1.0000


Predicted ROC
Sensitivity Y6 (Blue Ocean 4)
1-Specificity Y6 (Blue Ocean 4)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.7889


Predicted ROC
Sensitivity Y7 (Blue Ocean 6)
1-Specificity Y7 (Blue Ocean 6)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.8701

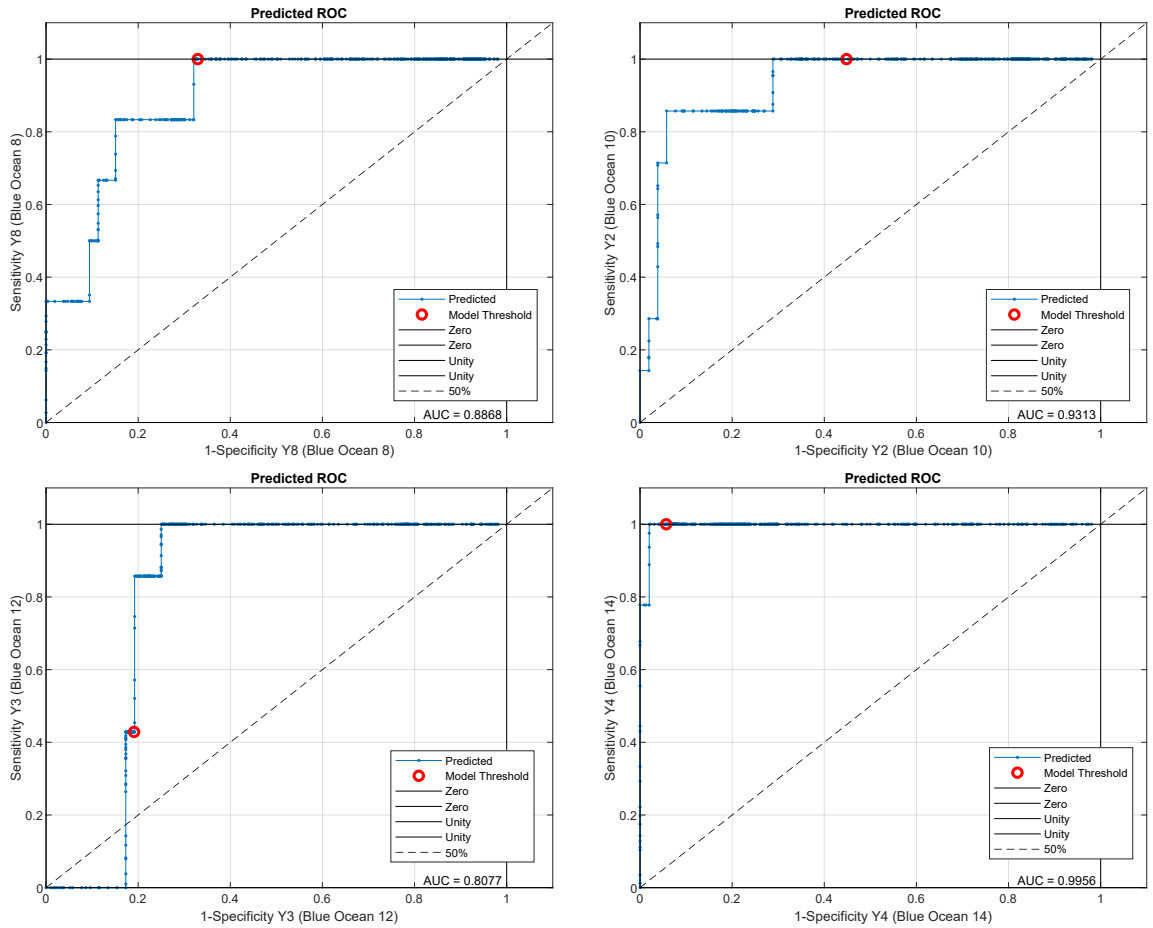
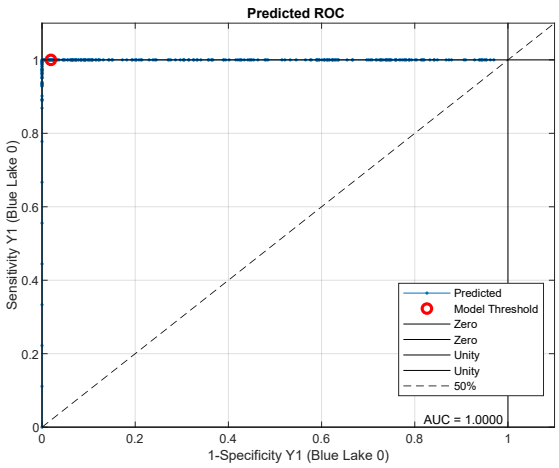
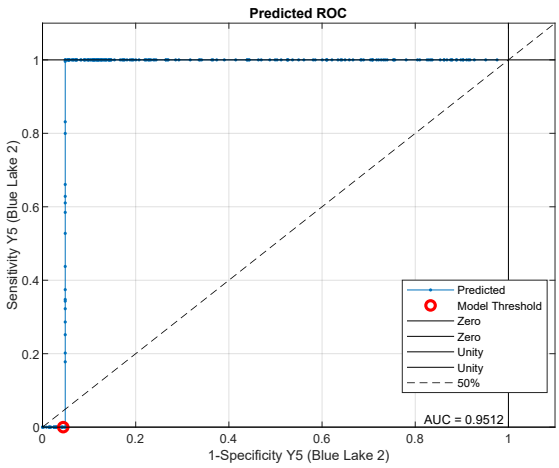
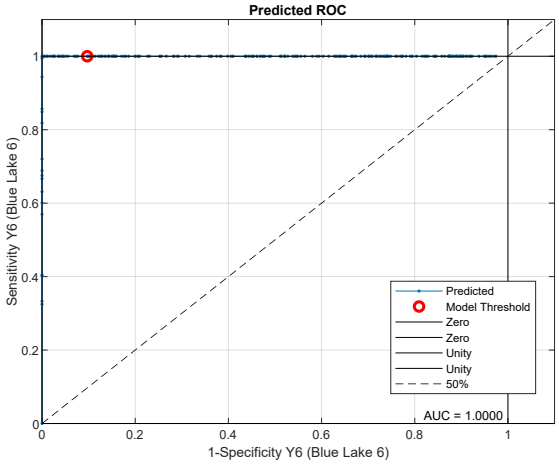
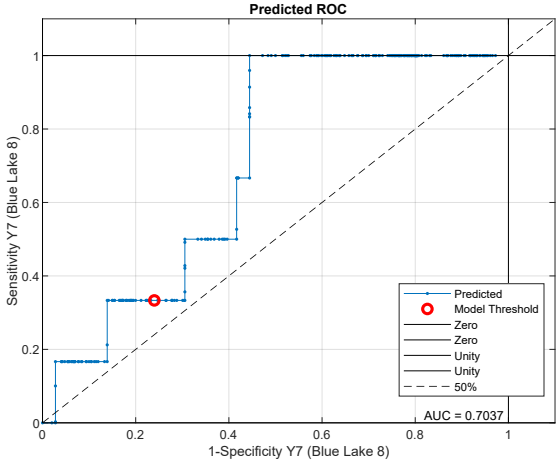


Figure S12. ROC plots for PLS-DA model used to predict faring stages of blue-colored cotton exposed to ocean water (Figure 5, A).


Predicted ROC
Sensitivity Y1 (Blue Lake 0)
1-Specificity Y1 (Blue Lake 0)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 1.0000


Predicted ROC
Sensitivity Y5 (Blue Lake 2)
1-Specificity Y5 (Blue Lake 2)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.9512


Predicted ROC
Sensitivity Y6 (Blue Lake 6)
1-Specificity Y6 (Blue Lake 6)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 1.0000


Predicted ROC
Sensitivity Y7 (Blue Lake 8)
1-Specificity Y7 (Blue Lake 8)
Predicted
Model Threshold
Zero
Zero
Unity
Unity
50%
AUC = 0.7037

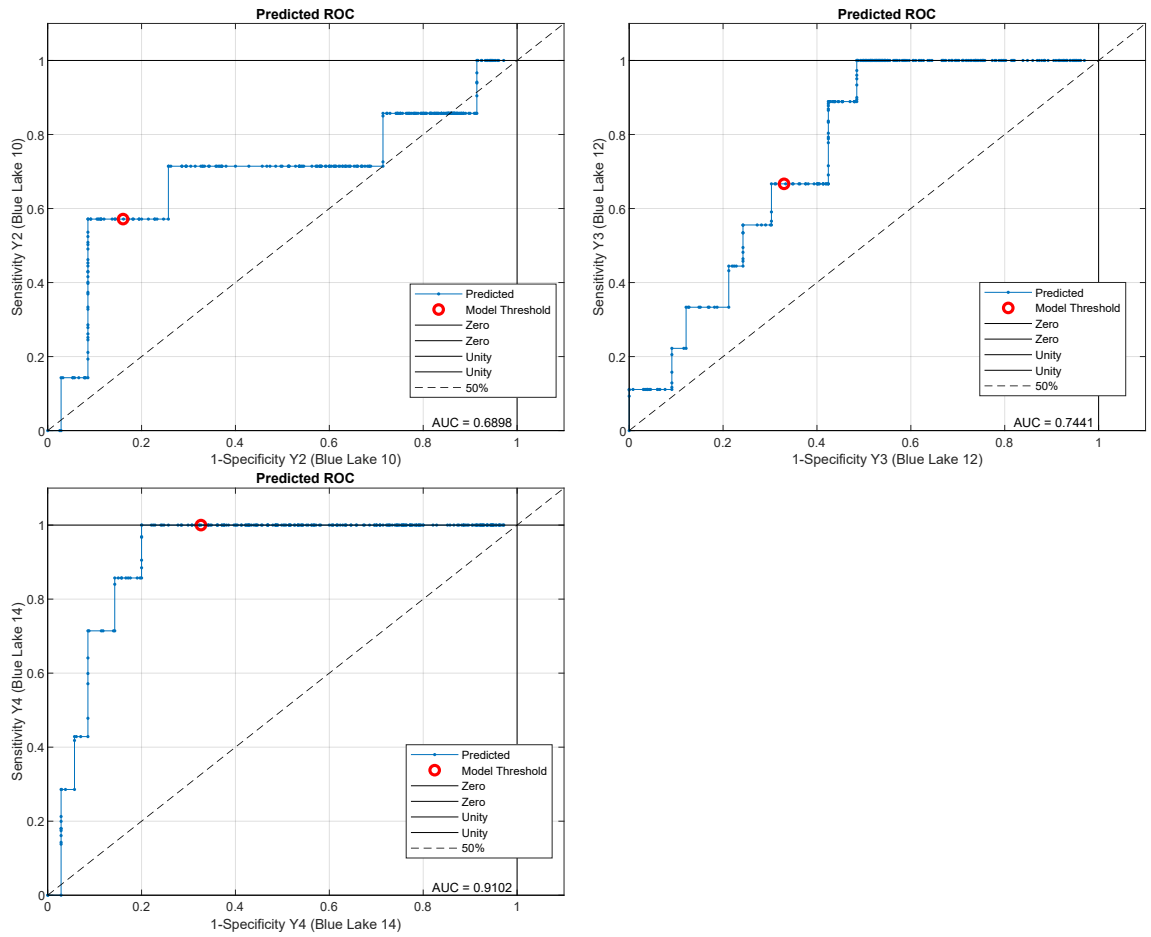


Figure S13. ROC plots for PLS-DA model used to predict faring stages of blue-colored cotton exposed to lake water (Figure 5, A).

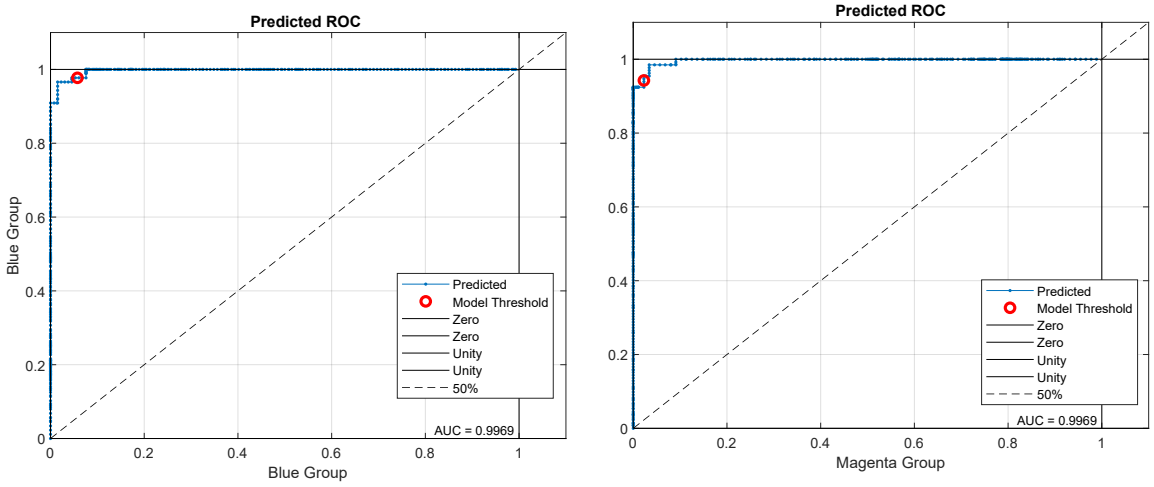


Figure S14. ROC plots for PLS-DA model used to predict faring stages of blue-colored cotton exposed to ocean and lake water (Table 2).

Magenta Ocean

```
Statistics for each y-block column:
Modeled Class: 1   2   3   4   5   6   7   8
Sensitivity (Cal):  1.000  0.920  0.920  0.920  1.000  0.960  0.720  0.920
Specificity (Cal):  0.971  0.379  0.707  0.828  0.839  0.782  0.925  0.305
Sensitivity (CV):  1.000  0.760  0.840  0.920  0.920  0.880  0.640  0.840
Specificity (CV):  0.971  0.431  0.713  0.822  0.839  0.776  0.914  0.305
Class. Err (Cal): 0.0142857 0.350345 0.186552 0.126207 0.0804598 0.129195 0.177356 0.387701
Class. Err (CV): 0.0142857 0.404483 0.223678 0.12908 0.12046 0.172069 0.223103 0.427701
RMSEC: 0.195708 0.32873 0.305846 0.290856 0.302851 0.301699 0.270861 0.328642
RMSECV: 0.200368 0.330892 0.310557 0.297958 0.304747 0.305894 0.28296 0.330443
Bias: 0.00139602 -0.00211777 0.000903964 0.00393892 0.0017842 -0.000288663 -0.00129316 -0.00432351
CV Bias: -0.000588994 -0.00161474 -0.000153678 0.00416464 0.00336524 0.00211714 -0.00314917 -0.00414044
R^2 Cal: 0.638881 0.0162684 0.148433 0.229994 0.16505 0.171364 0.332116 0.0169227
R^2 CV: 0.622237 0.00545213 0.123259 0.194702 0.154755 0.148564 0.273923 0.00787354
```

```
Key:
mag o*1   = mag ocean 0
mag o*2   = mag ocean 10
mag o*3   = mag ocean 12
mag o*4   = mag ocean 14
mag o*5   = mag ocean 2
mag o*6   = mag ocean 4
mag o*7   = mag ocean 6
mag o*8   = mag ocean 8
```

Magenta Lake

```
Statistics for each y-block column:
Modeled Class: 1   2   3   4   5   6   7   8
Sensitivity (Cal):  1.000  0.960  1.000  1.000  0.960  1.000  0.958  1.000
Specificity (Cal):  1.000  0.491  0.838  0.855  0.751  0.613  0.764  0.480
Sensitivity (CV):  0.958  0.960  1.000  1.000  0.960  1.000  0.917  0.920
Specificity (CV):  1.000  0.497  0.838  0.850  0.757  0.613  0.787  0.445
Class. Err (Cal):     0 0.274335 0.0809249 0.0722543 0.144277 0.193642 0.138649 0.260116
Class. Err (CV): 0.0208333 0.271445 0.0809249 0.0751445 0.141387 0.193642 0.147989 0.317457
RMSEC: 0.15953 0.321001 0.274437 0.267586 0.306687 0.316642 0.292955 0.324222
RMSECV: 0.162275 0.321889 0.278007 0.269458 0.307014 0.317569 0.294298 0.325489
Bias:     0     0     0     0     0     0     0     0
CV Bias: -0.00122963 0.000617691 -0.000705469 -0.000730119 0.000545723 0.000778805 0.000257905 0.000465093
R^2 Cal: 0.761079 0.0659765 0.3173 0.350958 0.14742 0.091174 0.194303 0.0471412
R^2 CV: 0.752808 0.0608168 0.299451 0.341869 0.145628 0.0858487 0.186937 0.0399882
```

```
mag lake 0
mag lake 10
mag lake 12
mag lake 14
mag lake 2
mag lake 4
mag lake 6
mag lake 8
```

Blue Ocean

```
Statistics for each y-block column:
Modeled Class: 1   2   3   4   5   6   7   8
Sensitivity (Cal):  1.000  0.920  0.920  1.000  1.000  0.960  1.000  0.920
Specificity (Cal):  1.000  0.920  0.800  0.949  0.994  0.794  0.903  0.857
Sensitivity (CV):  0.960  0.840  0.840  0.920  1.000  0.960  0.960  0.880
Specificity (CV):  1.000  0.931  0.829  0.937  0.977  0.760  0.914  0.874
Class. Err (Cal):     0  0.08  0.14 0.0257143 0.00285714 0.122857 0.0485714 0.111429
Class. Err (CV):  0.02 0.114286 0.165714 0.0714286 0.0114286  0.14 0.0628571 0.122857
RMSEC: 0.224482 0.258479 0.285814 0.203819 0.129599 0.314181 0.280101 0.262073
RMSECV: 0.230464 0.279739 0.304798 0.237199 0.147488 0.316755 0.289052 0.27459
Bias: 0.000938531 0.00269348 -0.000180958 -0.00950192 -0.00301156 0.000269148 0.0076495 0.00114377
CV Bias: 0.00170313 0.0123003 0.00418162 -0.0126185 -0.00851548 0.00209743 0.00286431 -0.00201284
R^2 Cal: 0.539281 0.389217 0.253121 0.621013 0.846521 0.0975094 0.283218 0.372062
R^2 CV: 0.514439 0.293084 0.165281 0.487362 0.807615 0.0833102 0.237898 0.311438
```

```
Key:
blue *1   = blue ocean 0
blue *2   = blue ocean 10
blue *3   = blue ocean 12
blue *4   = blue ocean 14
blue *5   = blue ocean 2
blue *6   = blue ocean 4
blue *7   = blue ocean 6
blue *8   = blue ocean 8
```

Blue Lake

```
Statistics for each y-block column:
Modeled Class: 1   2   3   4   5   6   7
Sensitivity (Cal):  1.000  0.750  0.958  0.958  1.000  1.000  0.000
Specificity (Cal):  1.000  0.496  0.717  0.743  0.985  0.911  0.973
Sensitivity (CV):  1.000  0.792  0.958  0.958  0.667  1.000  0.080
Specificity (CV):  1.000  0.478  0.726  0.735  0.978  0.935  0.946
Class. Err (Cal):     0 0.377212 0.162426 0.149152 0.00746269 0.0443548 0.513393
Class. Err (CV):     0 0.365229 0.158001 0.153577 0.177861 0.0322581 0.486786
RMSEC: 0.143474 0.379268 0.332936 0.337316 0.102933 0.234915 0.347794
RMSECV: 0.149066 0.382102 0.338513 0.3423 0.121271 0.247518 0.363809
Bias: 0.00250803 0.0124428 -0.00426954 -0.0147492 -0.00662865 0.0145327 -0.00383616
CV Bias: 0.00589129 0.0134488 -0.00388278 -0.0172416 -0.0109332 0.0101845 0.00253293
R^2 Cal: 0.857582 0.00589988 0.232989 0.214063 0.507377 0.359937 0.189273
R^2 CV: 0.846459 0.00131968 0.207074 0.191264 0.323459 0.288487 0.122221
```

```
Key:
blue *1   = blue lake 0
blue *2   = blue lake 10
blue *3   = blue lake 12
blue *4   = blue lake 14
blue *5   = blue lake 2
blue *6   = blue lake 6
blue *7   = blue lake 8
```

Figure S15. CV classification error and RMSE values for each model.

Table S1. P-values for ANOVA of magenta-colored fabric exposed to ocean water.

Comparison	p
0-10	< 0.001

0-12	< 0.001
0-14	< 0.001
0-4	0.006
0-6	< 0.001
0-8	< 0.001
10-12	0.005
10-14	< 0.001
10-2	< 0.001
12-2	< 0.001
12-4	< 0.001
12-8	0.049
14-2	< 0.001
14-4	< 0.001
14-8	0.012
2-4	0.034
2-6	< 0.001
2-8	< 0.001
4-6	0.002
4-8	.013

Table S2. P-values for ANOVA of magenta-colored fabric exposed to lake water.

Comparison	p
0-10	< 0.001
0-12	< 0.001
0-14	< 0.001
0-4	0.015
0-6	< 0.001
0-8	< 0.001
10-12	< 0.001
10-14	< 0.001
10-2	0.012
10-6	0.042
10-8	0.007
12-2	< 0.001
12-4	< 0.001
12-6	0.006
12-8	0.030
14-2	< 0.001

14-4	< 0.001
14-6	< 0.001
14-8	< 0.001
2-6	< 0.001
2-8	< 0.001
4-6	0.001
4-8	< 0.001

Table S3. P-values for ANOVA of blue-colored fabric exposed to ocean water.

Comparison	p
0-10	< 0.001
0-12	< 0.001
0-14	< 0.001
0-2	0.004
0-4	0.003
0-6	< 0.001
0-8	< 0.001
10-12	< 0.001
10-14	< 0.001
10-2	0.024
10-4	0.031
10-8	0.008
12-2	< 0.001
12-4	< 0.001
12-6	< 0.001

14-2	< 0.001
14-4	< 0.001
14-6	< 0.001
14-8	0.013
2-8	< 0.001
4-8	< 0.001
6-8	< 0.001

Table S4. P-values for ANOVA of blue-colored fabric exposed to lake water.

Comparison	p
0-10	0.006
0-12	< 0.001
0-14	< 0.001
0-2	< 0.001
0-6	0.007
0-8	< 0.001
10-12	< 0.001
10-14	< 0.001

10-2	0.015
10-8	0.006
12-6	< 0.001
12-8	0.002
14-6	< 0.001
14-8	0.002
2-6	0.034