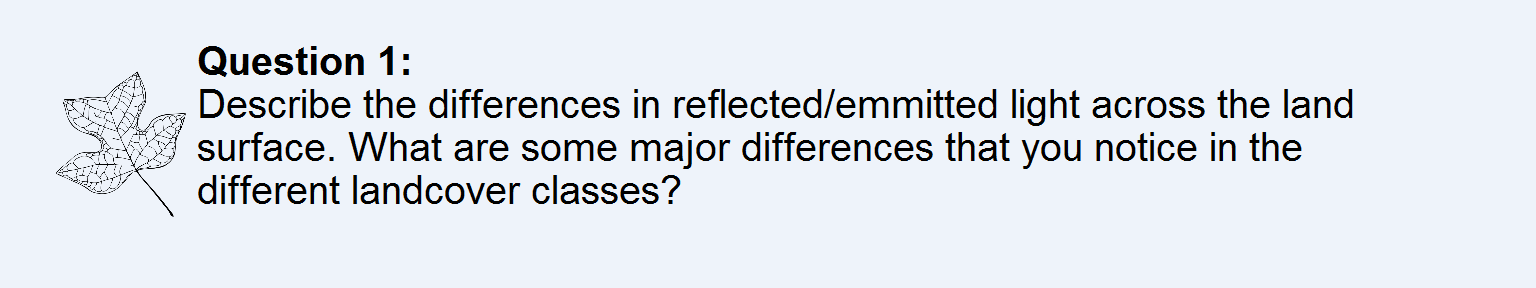
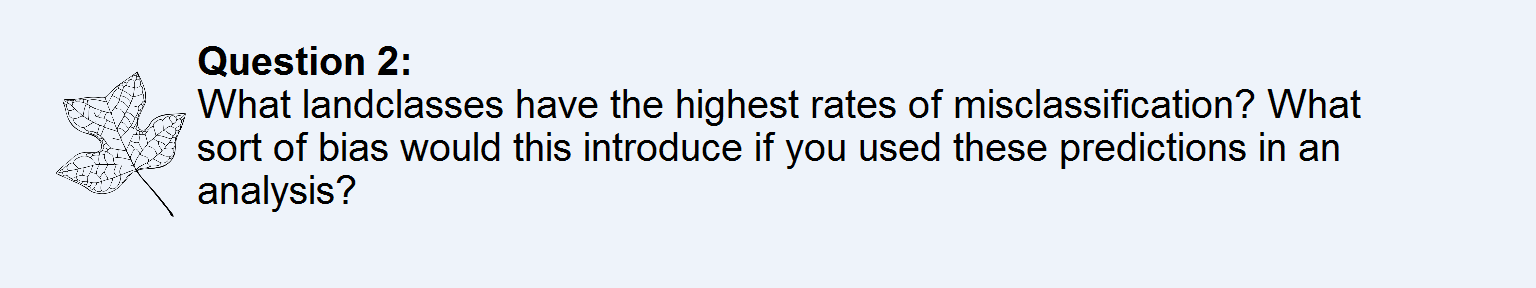
Charlie Huemmler

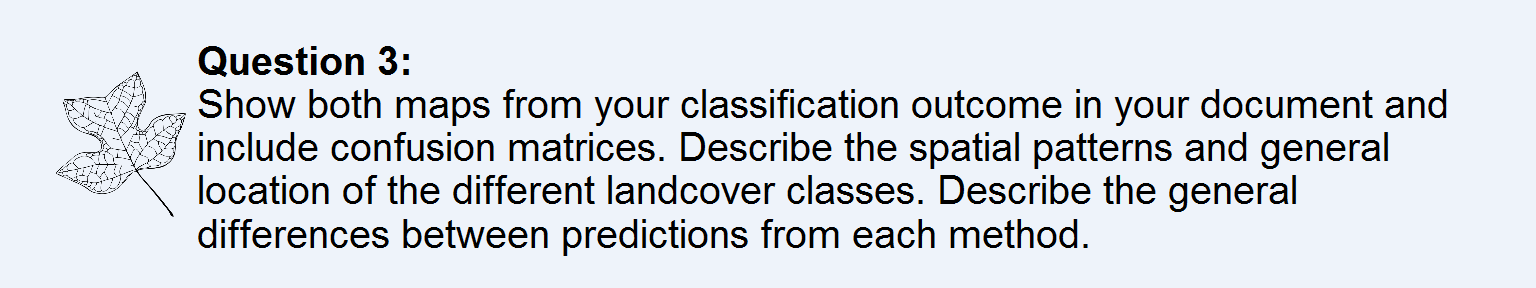
EDS Activity 7

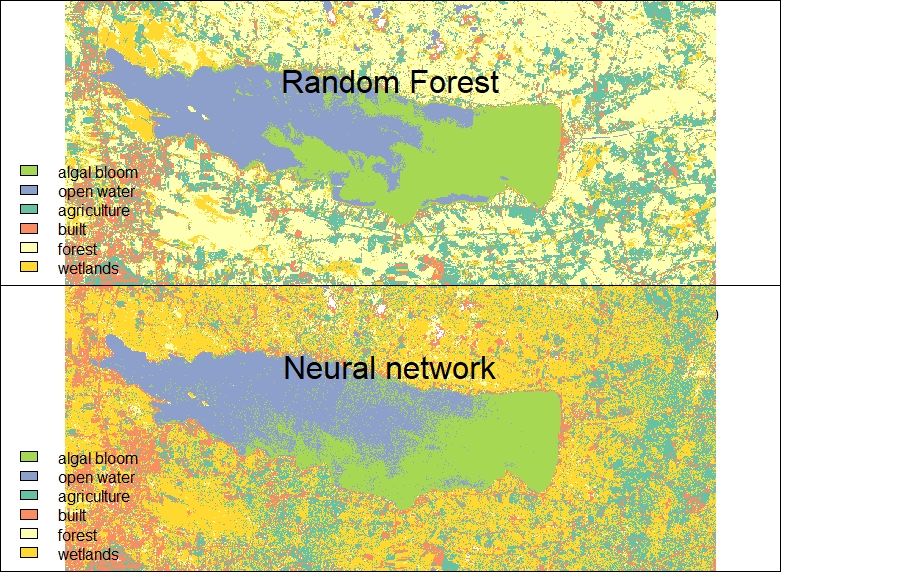


Different areas of land reflect different spectrums of light for our sensors to read. You can tell which areas of land are wooded, farms, or urban based on the bands that are most present in those regions. In the false color image, we can see the urban areas come out as gray, forests as purple, and agriculture as pink.



Both ‘Built’ and ‘Agriculture’ have high rates of misclassification, frequently being mistaken for each other. This introduces a bias for more of these land class to appear than is actually true.





Random Forest:

Reference

Prediction algal bloom open water agriculture built forest wetlands

algal bloom 60 0 0 0 0 0

open water 0 60 0 0 0 0

agriculture 0 0 54 5 0 0

built 0 0 4 53 0 0

forest 0 0 2 0 59 1

wetlands 0 0 0 0 1 59

Neural Network:

Reference

Prediction algal bloom open water agriculture built forest wetlands

algal bloom 58 1 0 0 0 0

open water 2 59 0 0 0 0

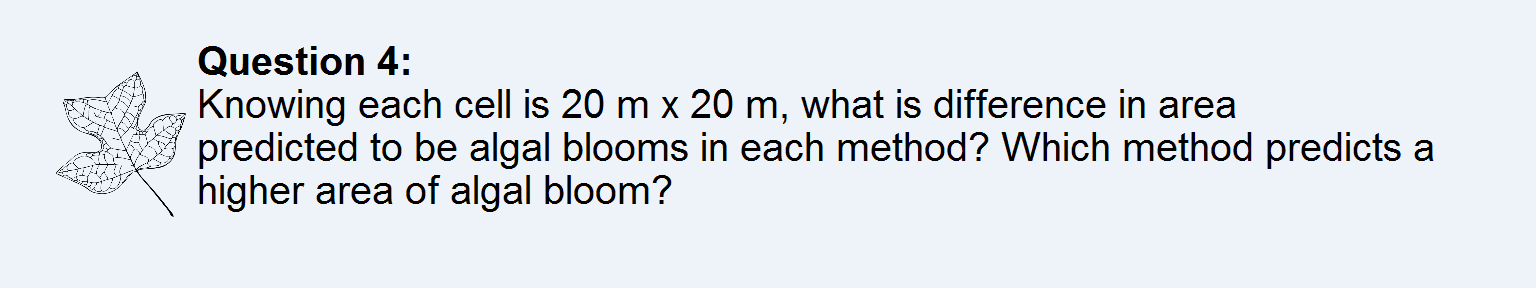
agriculture 0 0 45 2 9 3

built 0 0 7 56 0 1

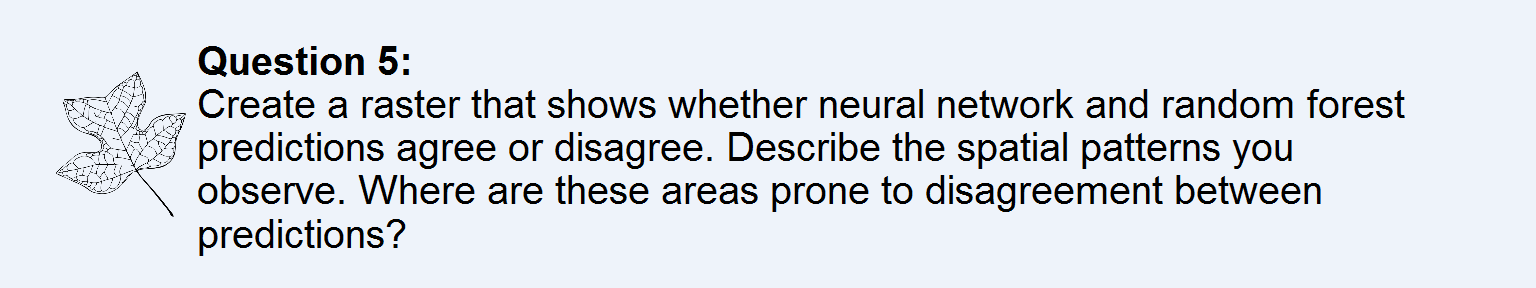
forest 0 0 0 0 2 5

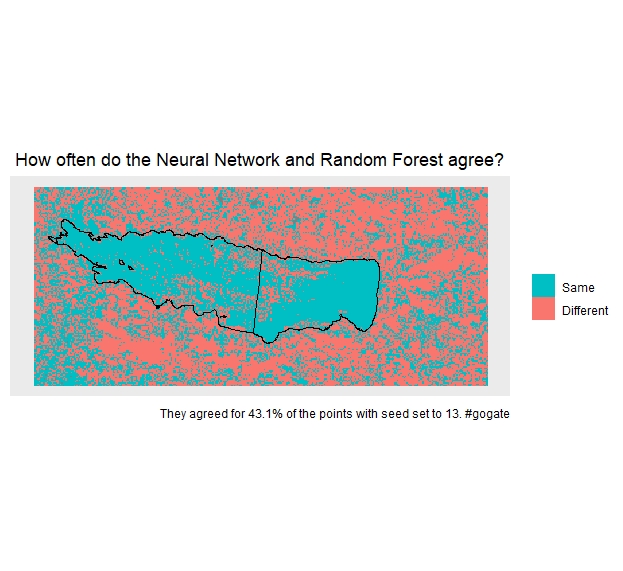
wetlands 0 0 8 0 49 51

Algal blooms appear in the eastern half of the lake, with open water to the west. Syracuse, a ‘built’ area, is to the south west of like Oneida. Agriculture is scatter around the map. Across the board, the neural network sees wetlands where the random forest sees forest. The landclasses are much more dotted throughout the neural network, while more clumped with the random forest.

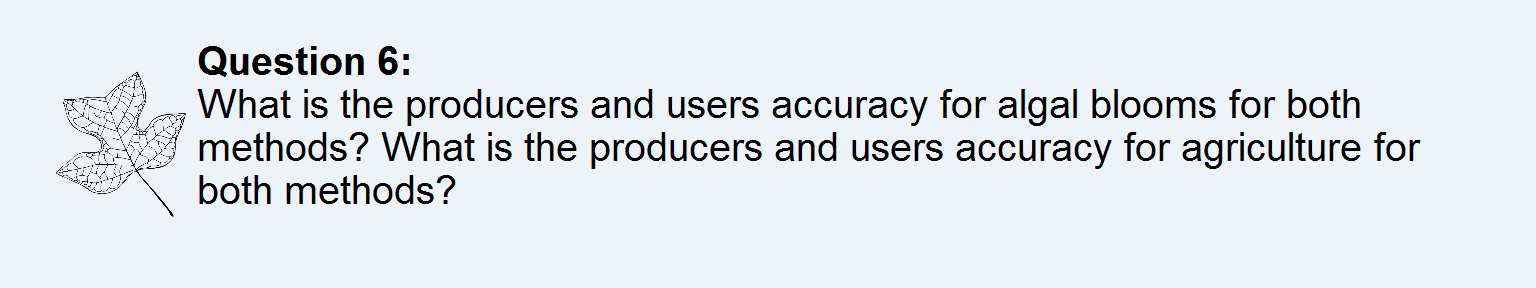


The random forest predicts 7,974,800 more square meters of algal bloom than the neural network model. This is 3.85% of the entire size of lake Oneida. (79.8 square miles)

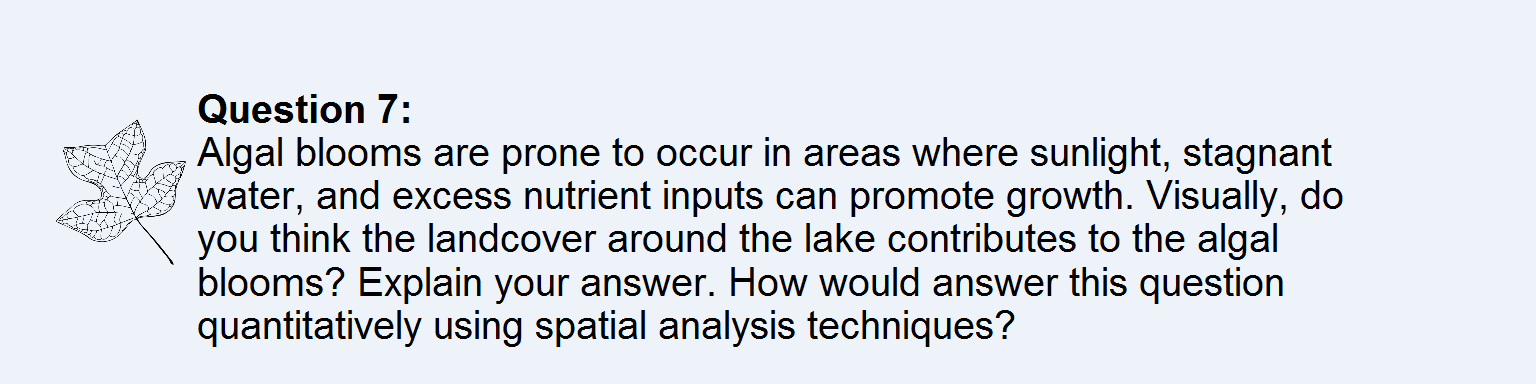




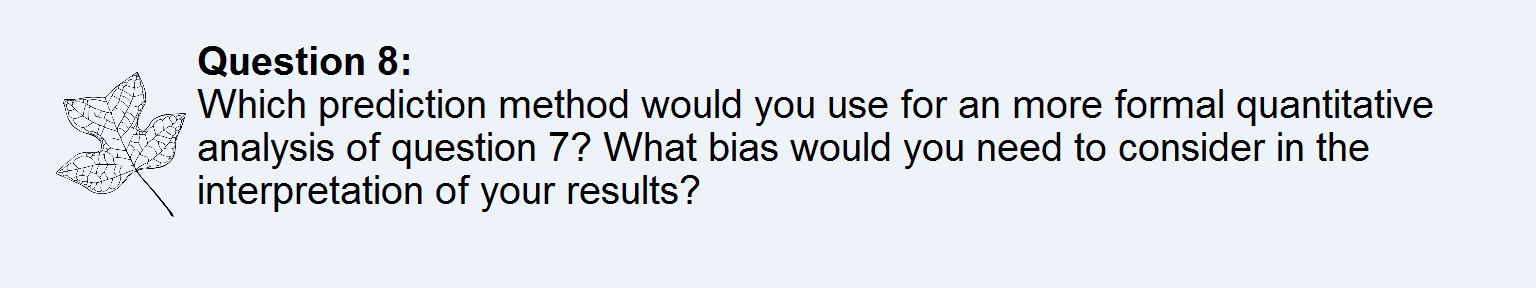
The models disagree mostly between the wetlands and forest, as the neural network repeatedly mis-predicted the wetland class. Furthermore, there are some discrepancy on the edge of the algal bloom.



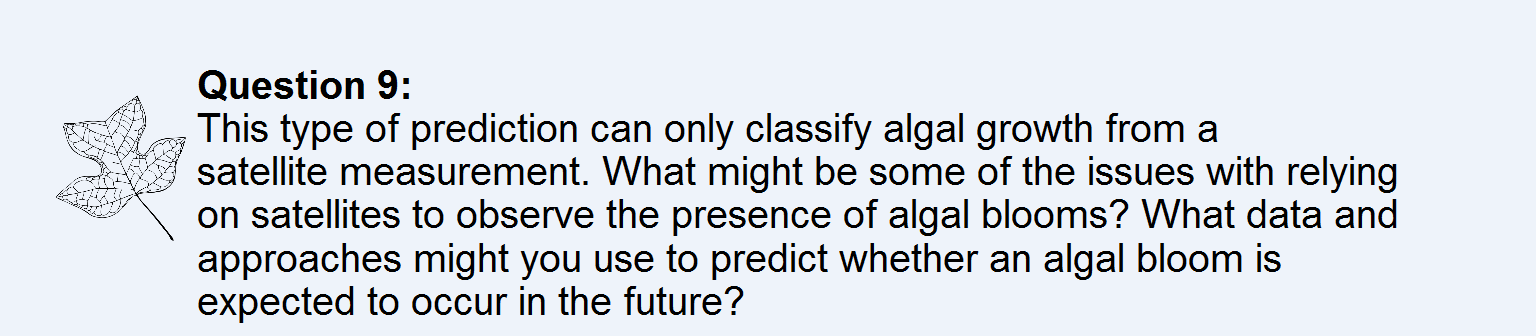
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RF User | RF Prod | NN User | NN Prod |
| Algal | 0% | 0% | 98.3% | 96.6% |
| Agriculture | 91.5% | 90% | 76.2% | 75% |



I do believe the landcover around the lake affects the algal blooms. The large amounts of agricultural land to the east of the lake contributes to the algal blooms, while the more built area to the west (Syracuse) does not shed fertilizer into the lake. This could be analyzed quantitatively by introducing watershed data to the equation. We could see the amount of agricultural land, and thus fertilizer, enters the lake from each water shed and how the currents and algal blooms map around that. The coastline would be weighted by how much runoff enters the lake from each unit of coastline.



I would use the random forest model as it was more accurate than the neural network model in the previous predictions. I would need to consider the bias of mis-predictions between agricultural land and built land in the interpretation of my results.



One concern I have is how deep the algal blooms reach. Perhaps some sections are much deeper, unobserved from satellite imagery. If it is predictions we are after, then incorporating weather data may be useful. Furthermore, if we could obtain data on fertilizer usage of nearby farms that may prove predictive as well.

