

FROM DIVINING RODS TO RANDOM FORESTS USING PREDICTIVE ANALYTICS TO MAP WETLANDS IN HAWAII



<https://aeon.co/essays/why-dowsing-for-water-is-a-test-of-faith-and-of-science>

Mari K Reeves (USFWS) and Adonia Henry (Pacific Birds)



Mari Reeves, Spatial Ecologist with PIFWO, been here about a year,

Been working with Adonia Henry at Pacific Birds Joint Venture to address the Second of her data gaps since this past November.

Unlike this man who is dowsing, or waterwitching, or doodlebugging (which is how you look for oil) with his divining rod, we are using slightly more sophisticated tools, like the machine learning algorithm Random Forests to identify and map wetlands in Hawaii.

As you can see from the picture, Both methods come from trees...I'll explain more about the random forests aspect of that later.

It's possible both require Similar amounts of faith.

Going to walk you through the need, by showing you some of the existing data layers

Share with you how we developed the data set to train our model to identify water on the landscape.

Share the results of our preliminary models including the predictions they were able to make to landscapes they hadn't seen yet. .

WETLANDS IN HAWAII



- Statewide wetland loss at 15%, with spatial patterns
 - More loss on Oahu (65%)
 - More loss in coastal areas (44%)
 - Less loss at higher elevations (3%)
 - van Rees and Reed, 2014



Loss of wetlands enumerated at X% by van Reed

We estimate statewide wetland loss at 15 %, compared to 12 % from the 1990 estimate, ranging from 6 to 8 % loss on Maui, Moloka'i, Hawai'i, and Kaua'i to 65%loss on Oahu, the most developed of the islands. The majority of wetland losses occurred in coastal areas where 44 % of wetlands have been lost, while only 3 % were lost at higher elevations.
es and Reed 2014

So we're concerned about losing these habitats with high conservation value, especially to waterbirds, and a good, comprehensive wetland inventory for Hawaii is not currently available.

WHY MAP?



<http://legacy.lib.utexas.edu/maps/topo/250k/bu-pclmaps-topo-us-maui-1961.jpg>

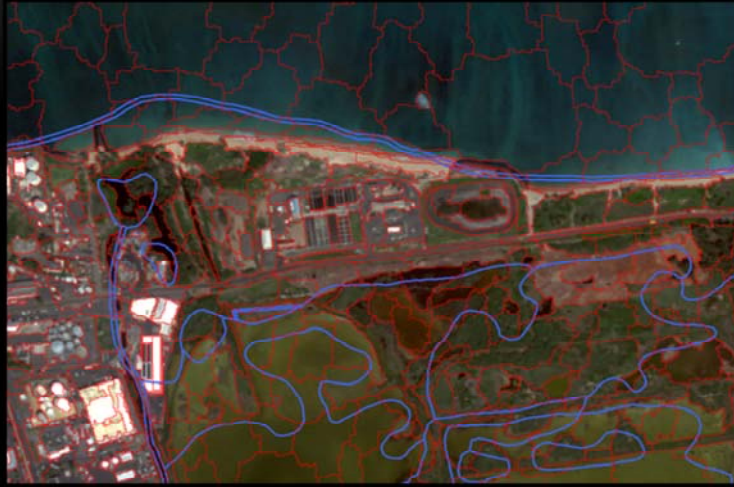


In addition to Adonia's points....

And that if you don't know where something was, it's hard to know when it's gone ...

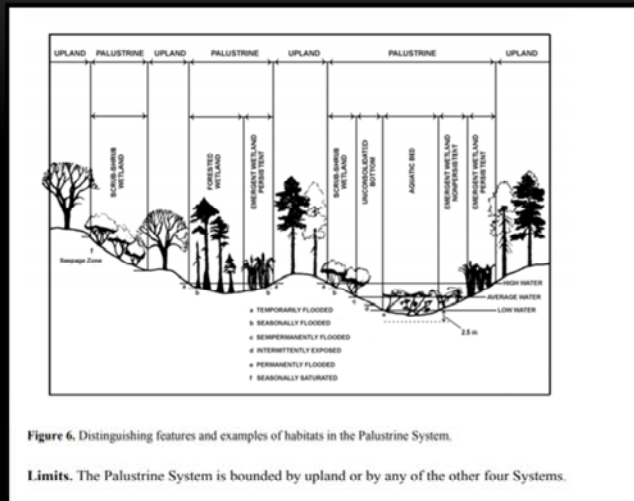
A lot of hawaii wetlands were mapped in early 1980s using topo maps. I've tried even 10 years ago in Alaska to navigate to wetlands using topo maps, and arrived to find grassy meadows there instead.

IMPROVED IMAGERY AND TECHNOLOGY



Moreover, the technology has improved. We have sub-meter scale satellite imagery and I'm showing you the NWI layers in marsh next to airport in Maui, blue is NWI overlaying the 1 m satellite imagery, and red is unsupervised image segmentation from a program called eCognition, which is not the way I did this analysis, but is one way that computer technology is enabling us to do land cover classification far more easily and efficiently that we used to do it.

NATIONAL WETLANDS INVENTORY

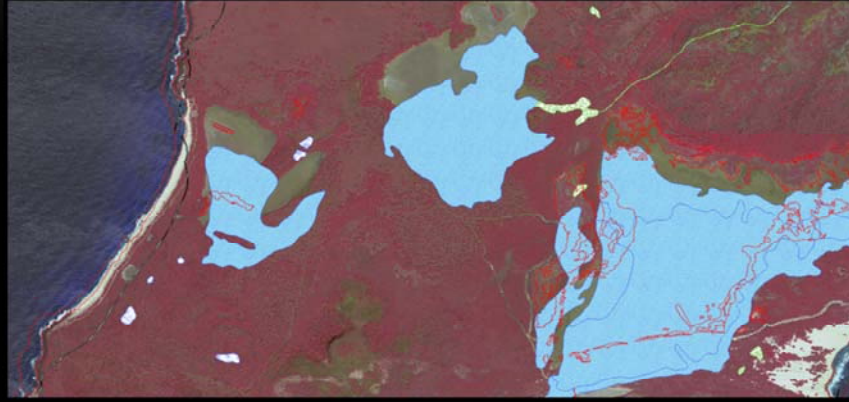


<https://www.fws.gov/wetlands/documents/Classification-of-Wetlands-and-Deepwater-Habitats-of-the-United-States-2013.pdf>



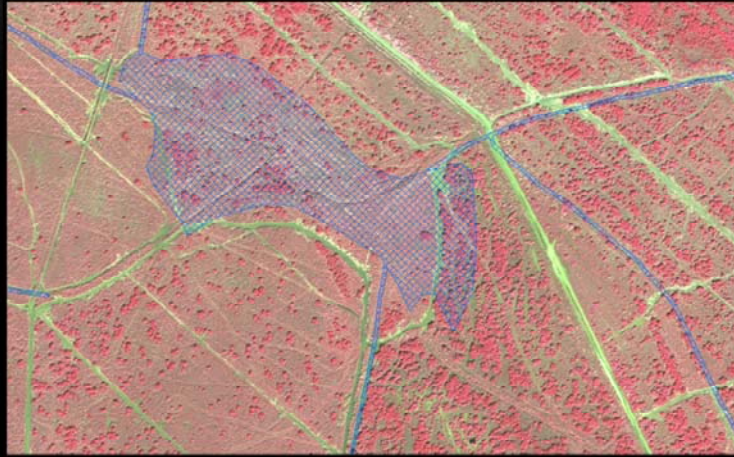
NWI is polygon based and has much greater detail about wetlands hydrology and vegetation - - It's not a complete landcover classification, it only contains wetlands. Unfortunately for Hawaii, it's horribly outdated.

NWI AND CCAP LAYERS IN NIIHAU



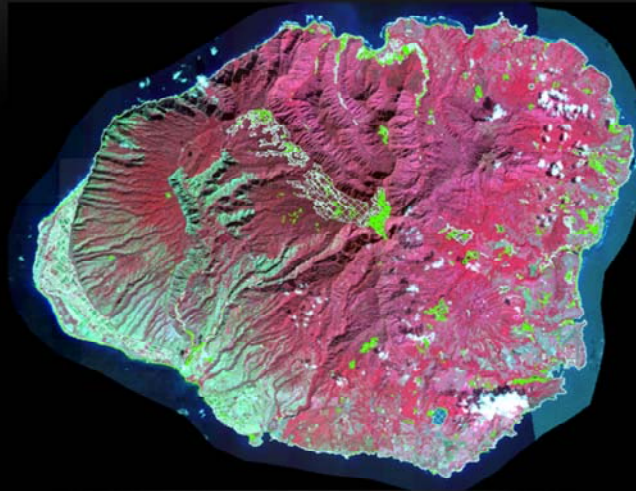
And in some places inaccurate. Here is Niihau. There are problems with even basic georeferencing and overlays on some islands (Niihau – Blue lines are NWI and are obviously off). NOAA has some better layers, the red lines on this air photo...but for reasons we'll discuss they are not as well suited to wetlands management as the NWI are.

LAND USE AND LAND COVER CHANGE



In addition there has been land use and land cover change since these maps were generated, where formerly wet areas have dried or been filled, moved, or drained. Lanai...wet, not wet.

KAUAI AND OAHU NWI IS MORE CURRENT



Using images from the mid 2000s, but still had drawn from the aerial photos, which were more like 5 m pixel scale.

NOAA CCAP LAYERS



Mad props to the NOAA coastal change analysis program, which has done a much more detailed land cover mapping of the Hawaiian Islands (and most of CNMI), and explicitly looks at land cover change.

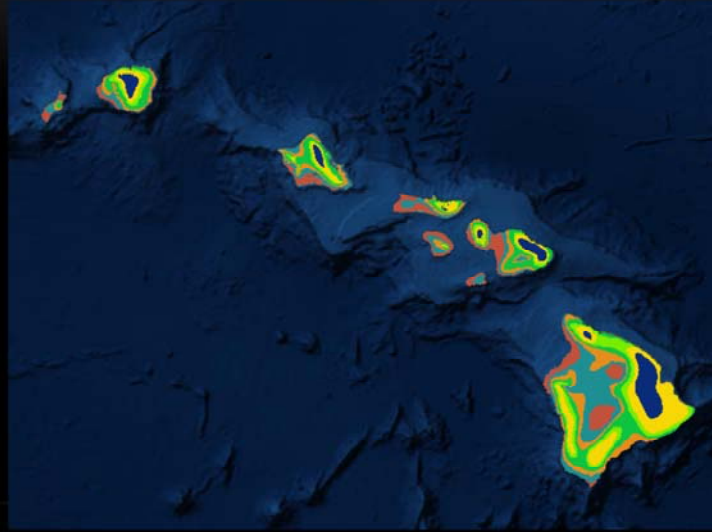
So they have layers from 2005 and 2011 for most islands, then have assessed change formally...we're using these layers heavily in the HSA efforts.

NOAA CCAP LAYERS



The NOAA layers are pretty detailed, where you can see trees on streets in Urban Honolulu, for example. This is good and bad from a wetlands mapping perspective, which we can discuss more if you want to.

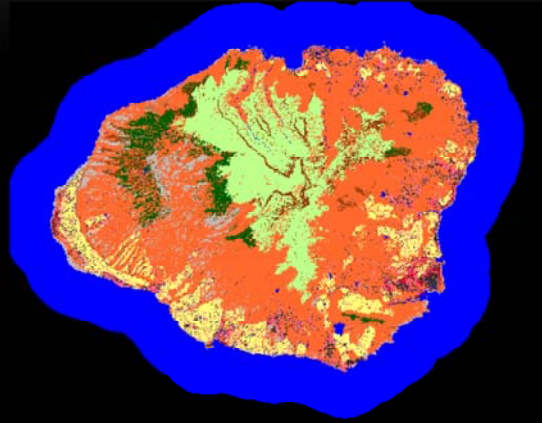
OTHER LAYERS – CLIMATE BASED MOISTURE



<https://pubs.usgs.gov/of/2012/1192/>



OTHER LAYERS – LANDFIRE VEGETATION TYPE



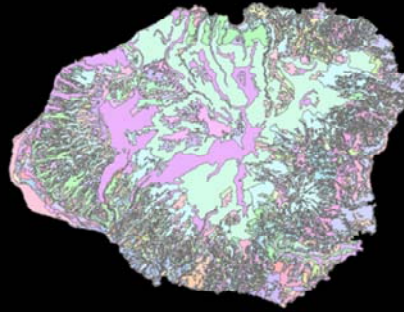
<https://www.landfire.gov/>



I b

OTHER LAYERS – SOILS AND GEOLOGY

- Soils

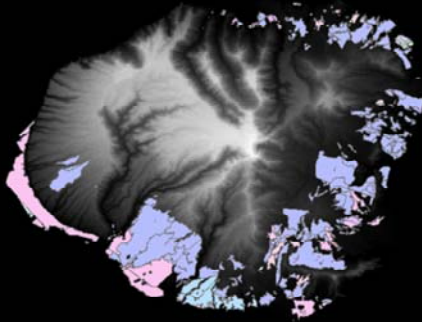


- Geology

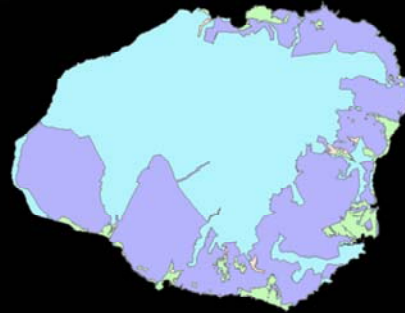


OTHER LAYERS – AGRICULTURE AND LAND USE

- Agriculture

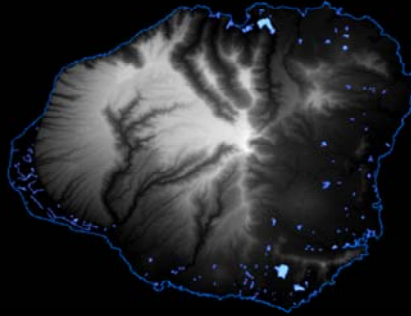


- Land Use Districts

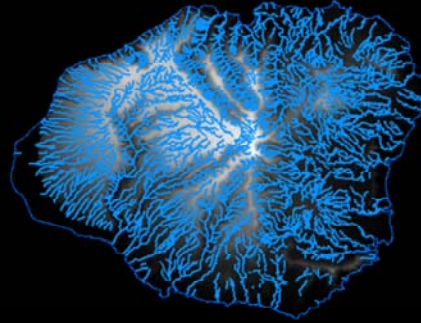


OTHER LAYERS – WATERBIRDS AND STREAMS

- Waterbird Survey Locations



- State Department of Aquatic Resources, Streams



BOTH CCAP AND NWI HAVE WETLAND INFORMATION...

- NOAA CCAP

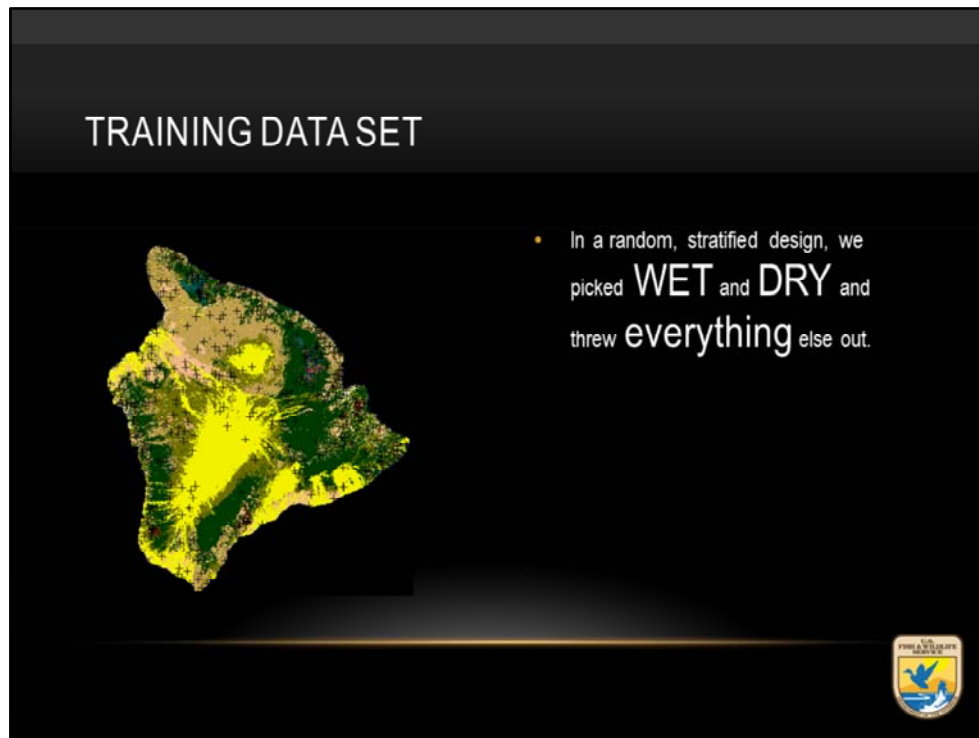
- USFWS NWI



There is a lot of overlap, but not perfect overlap, so who is right?

This is Kaneohe Bay

Neither layer is perfect for calculating areas of wetland habitat. But there is information about wetlands in both layers.



So because we had so much geospatial information that we could use to teach our models to waterwitch, we went with a point based approach for the training data,

Used NOAA layers to pull in the “not water” categories...omitting the more intermediate wetlands from analysis entirely.

Showed the models “definitely wet” and “definitely dry”, but nothing in between. The hope was it would use the information to identify those habitats in between.

Used spatially stratified design to randomly choose points from both the NWI and the CCAP layers

Used points so that we could drill down and gather data from other datasets as model input

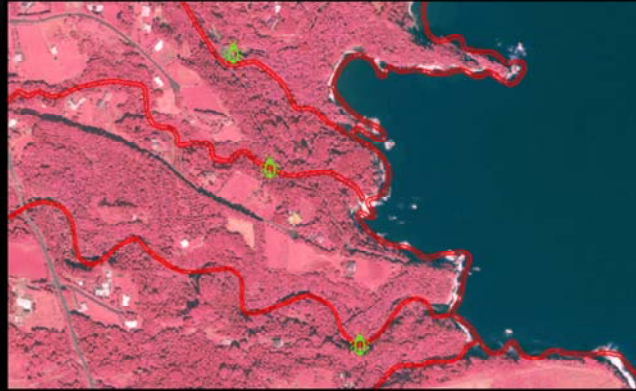
Picked sampling strata from NOAA and FWS layers that should be very clearly “water” and “not water”

Eliminated strata that were less obvious or we couldn't ground truth

~1300 points statewide (all Islands = more variation in training set)

Hope with this strategy was that the models would learn we were looking for water and identify more intermediate wetland habitats as having a moderate probability of being water

INCORPORATING T&E SPECIES



ONLY OPEN WATER WETLANDS



Makes you really appreciate what your eyes are doing when they look at an air photo image! Used only open water to make sure model knew what we were looking for.

TRAINING DATA FOR JAMES CAMPBELL NWR



Zoom in on James Campbell, to see the type of sampling intensity we had.

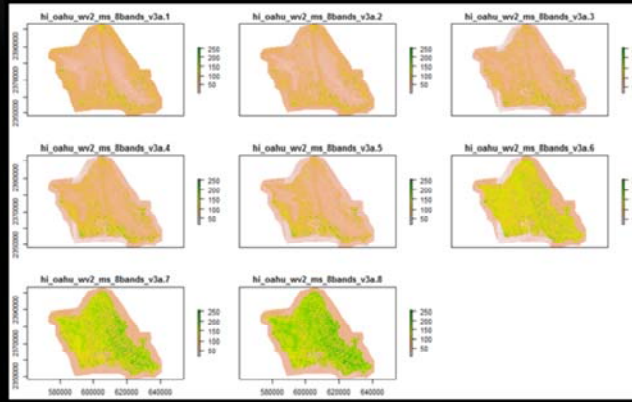
For each of these points (about 1300 of them, half from NWI and half from

SATELLITE IMAGERY



Mosaic'ed images sub meter scale, cloud free, and color enhanced thank you Tony Kimmet at NRCS.

MULTISPECTRAL IMAGERY



WE ONLY SEE THREE BANDS AT ONCE,
BUT THE SATELLITE SEES MORE...
AND SO DOES THE MODEL



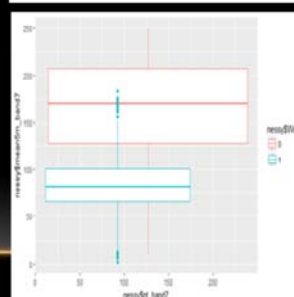
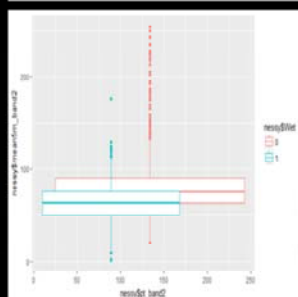
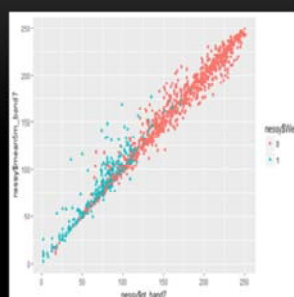
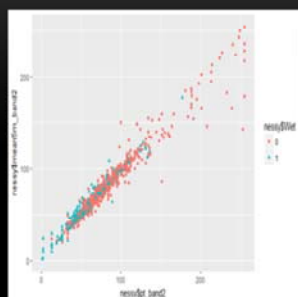
WHICH BANDS TO USE?



BAND 2

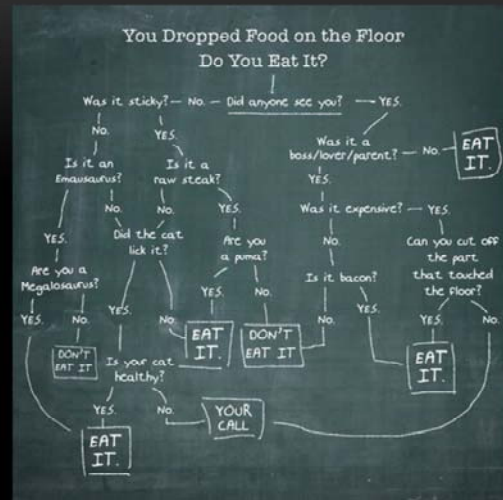
VS

BAND 7



MACHINE LEARNING HELPS US UNDERSTAND AND PREDICT

- Classification algorithms randomize the variables and split the data
 - Over and over and over again.
- In supervised classification, the analyst also classifies and splits the data
 - Training and Testing Datasets (75%/25%)
 - 10-fold cross validation



<https://laura.makes.org/>



Made a word play in my title about random forests, which are

LOCH NESS, MESSY NESSY, AND THE KRAKEN

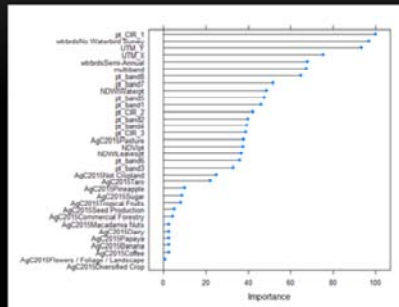


Random Forests
Boosted Regression
Partial Least Squares
Elastic Net Regression



Just talk about how so far these are preliminary models and need cleaning to finalize input data and outputs. Models help us understand and predict. These models are more run in the vein of predicting.

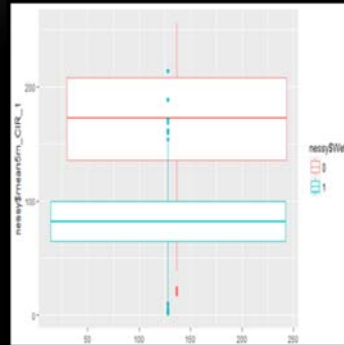
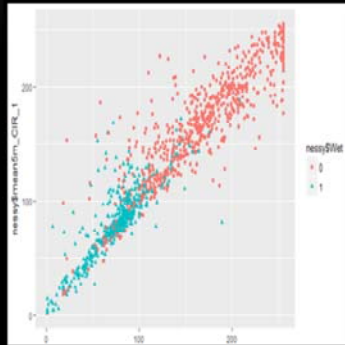
MODEL RESULTS...



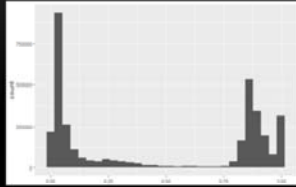
- Which variables are important
- How are they influencing the predictions?
- How well did the model work?



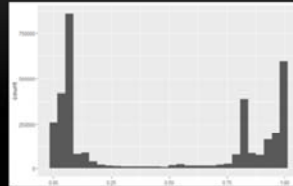
NRCS – TUNED COLOR INFRARED



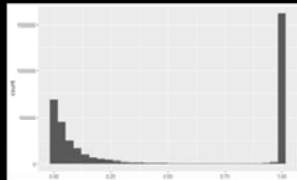
PREDICTIONS!!!!



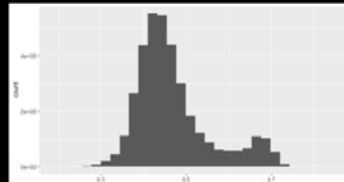
Random Forests (95% Accuracy)



Boosted Regression (93% Accuracy)

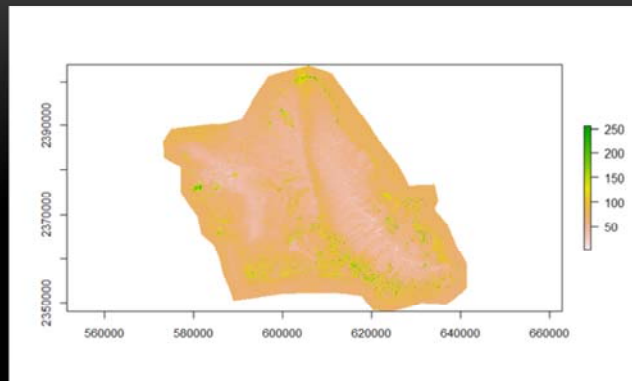


Elastic Net Regression (94% Accuracy)



Partial Least Squares Regression (91% Accuracy)

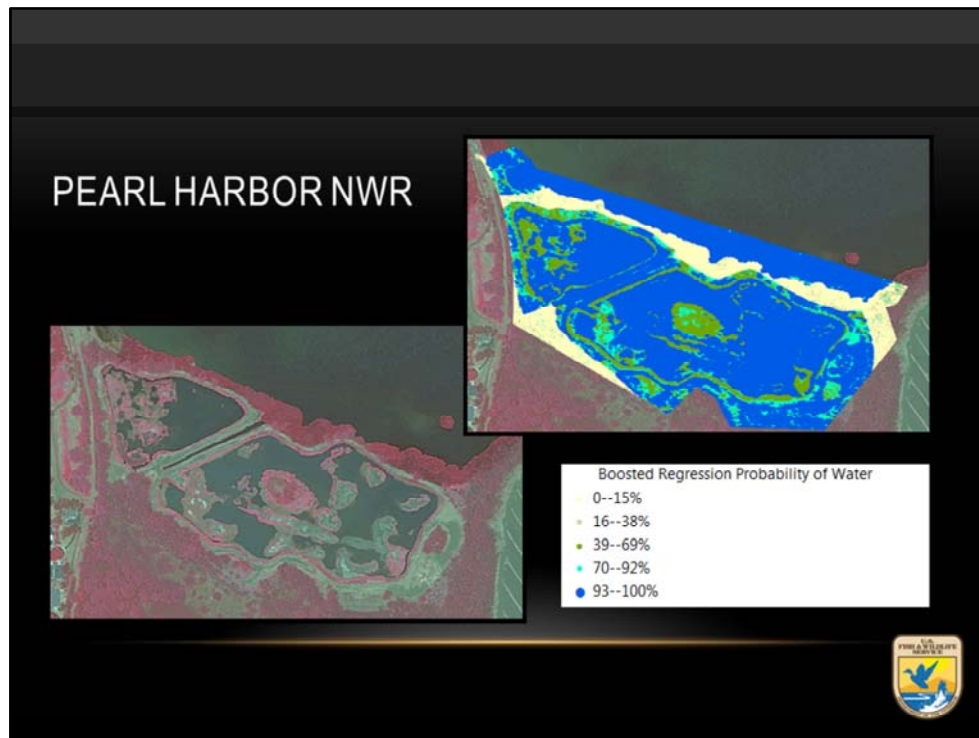




```
> system.time(plot(oahu8Band))
user system elapsed
10.36  0.69  21.05
```

```
> nrow(oahu8Band)
[1] 34786
> ncol(oahu8Band)
[1] 42698
1.48x109 cells
```





Add a legend -

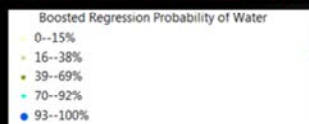
TRAINING DATA FOR JAMES CAMPBELL NWR



Zoom in on James Campbell, to see the type of sampling intensity we had.

For each of these points (about 1300 of them, half from NWI and half from

JAMES CAMPBELL NWR





Add Google



Insert Hawaii coast movie

SATELLITE POSITION MOVEMENT



2017-05-18

Source: © 2016 DigitalGlobe, GeoView License



Insert James Campbell movie.

ACKNOWLEDGEMENTS

- Pacific Birds, Adonia Henry
- Natural Resource Conservation Service, Tony Kimmet
- National Oceanic and Atmospheric Association, Ross Winans
- USFWS Strategic Habitat Conservation Team: Fred Amidon, Adam Vorsino, Stephen Miller



