STATE OF URBANIZATION IN SF BAY AREA / SILICON VALLEY

Alexander Wiegman | Dr. W. Takeuchi | *Remote Sensing Course*
17 November 2017

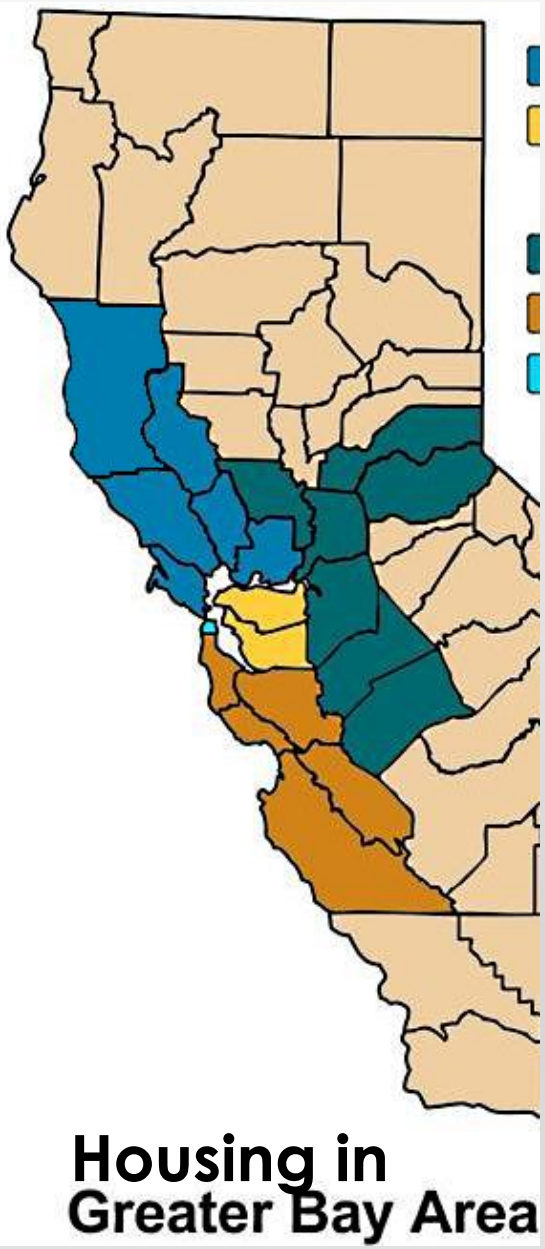
WHERE IS


SILICON VALLEY
San Francisco
Oakland
Berkeley
Sacramento
San Jose
Palo Alto
Mountain View
Sunnyvale
Santa Clara
Cupertino
Fremont


SILICON VALLEY
hooli
facebook


HIGHS TODAY
108 UKIAH
106 CLOVERDALE
104 LAKEPORT
106 SANTA ROSA
103 NAPA
104 FAIRFIELD
HOT
105 SAN RAFAEL
108 CONCORD
105 ANTIOCH
88 SAN FRANCISCO
97 OAKLAND
99 FREMONT
106 LIVERMORE
WARM
68 HALF MOON BAY
102 REDWOOD CITY
103 SAN JOSE
102 GILROY
80 SANTA CRUZ

WHERE IS THE BAY AREA?



Housing in Greater Bay Area

TIME-LAPSE LOOK

- Here, we can see a time-lapse of images taken from the LANDSAT satellites.
- The range is from 1984, 1987, 1997, 2007, and 2017. (LANDSAT 5, 7, and 8)
- Blue (Class 0) = Water
  Brown (Class 1) = Bare Ground
  Green (Class 2) = Greenery
  Yellow (Class 3) = Urban

METHODOLOGY

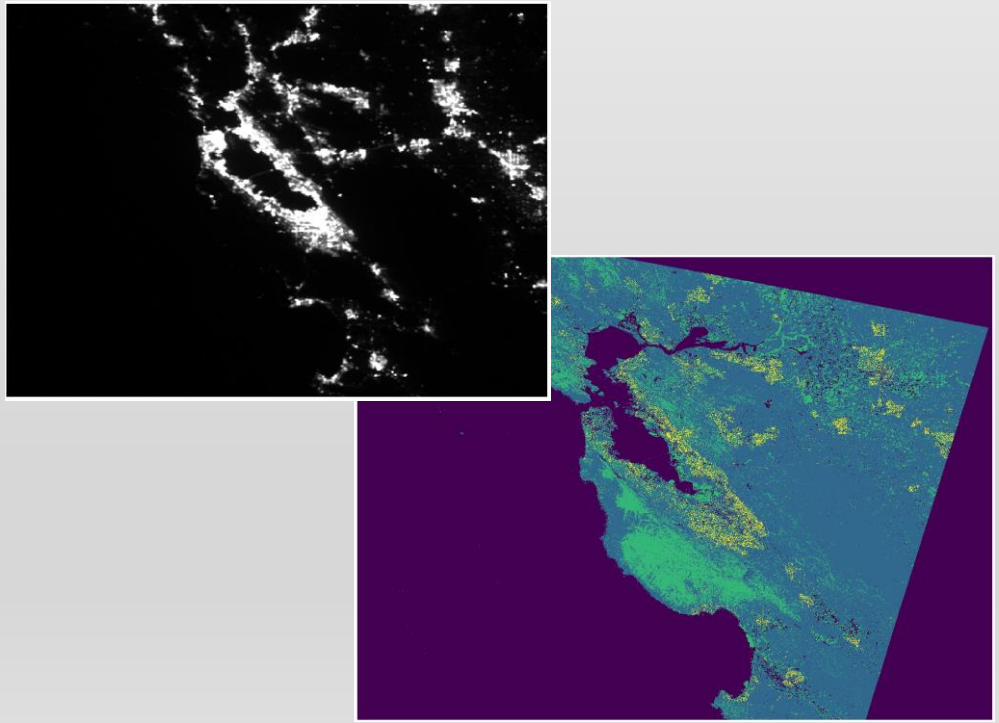
- Now that we have an initial prediction level (2017) and we have seen what it looks like in Google Earth Engine with the classification system, let's use our previous data and try to see if we can make an AI that will accurately predict the future land use distribution.
- We'll use a nighttime light-density image from early September 2017 as a benchmark for real-data population density.
- Against that, we shall use the land change maps (that you saw on the previous slides) for 1997 and 2007 as benchmarks. To check my results, I will check the difference between my 1984 and 2007 maps as well and do statistical validation.

ANALYSIS VIA MACHINE LEARNING

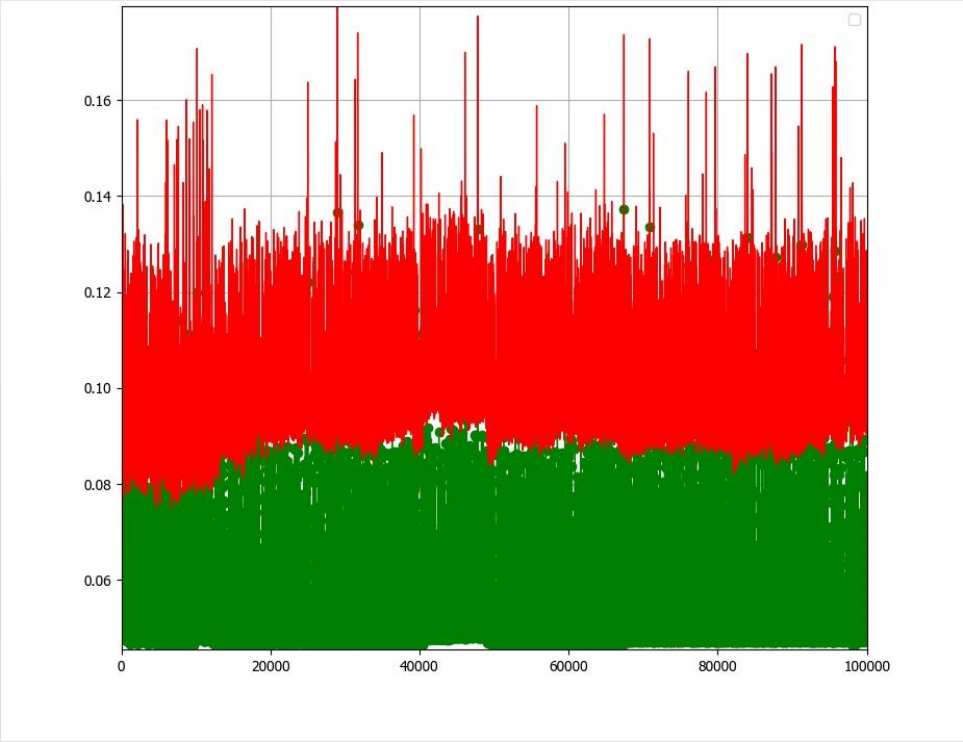
- We will use the MOLUSCE plugin of QGIS program.
- We start with inputting the data, then running a comparison (how much change there is in each kind of class), a neural network, a predictive mapping program, and a validation against real program.
- I will use a neural network of 100 000 (100 thousand, 10-man) nodes.
- The predictive mapping program will run with all default settings.
- The validation program will run default: logarithmic correction sequences with five iterations.

TRAINING DATA

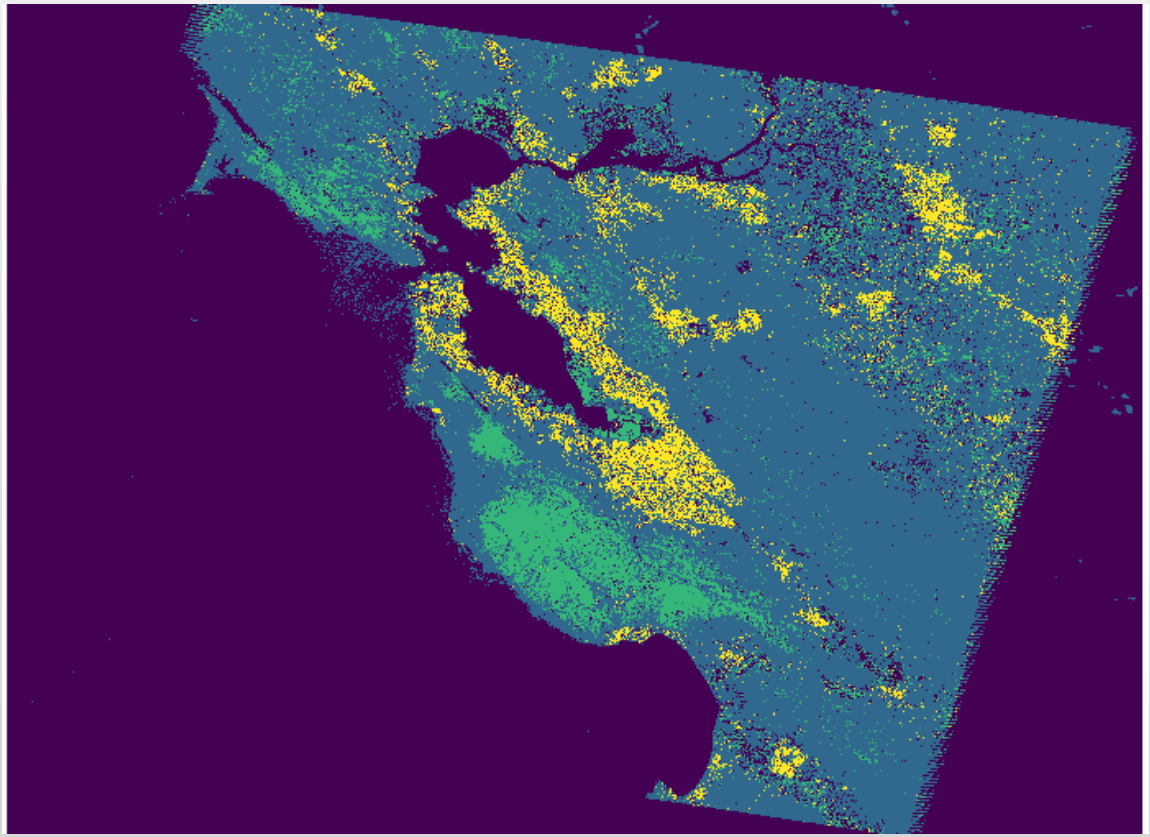
- We will run the program with TIFF maps exported (they look like the below).



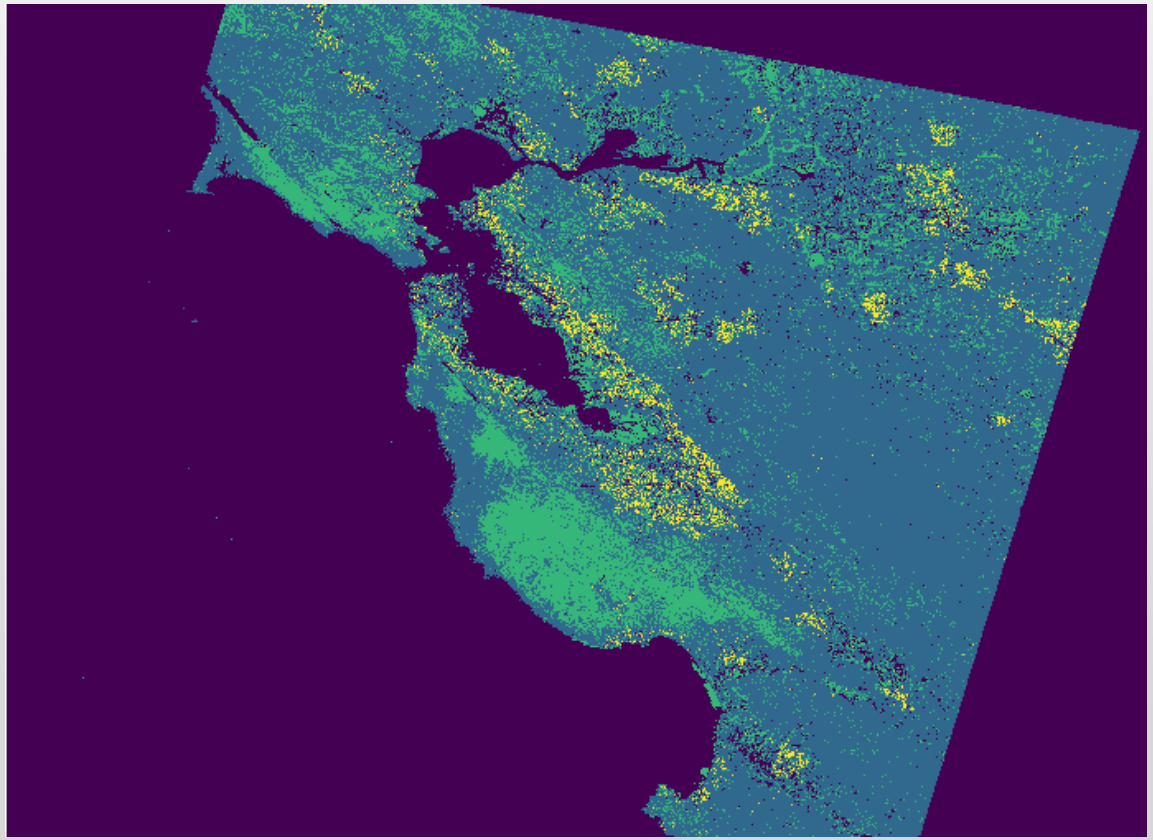
- The AI prediction and validation runs off neural network data (like below).



RESULTS OF ANALYSIS, 1997 + 2007 DATA

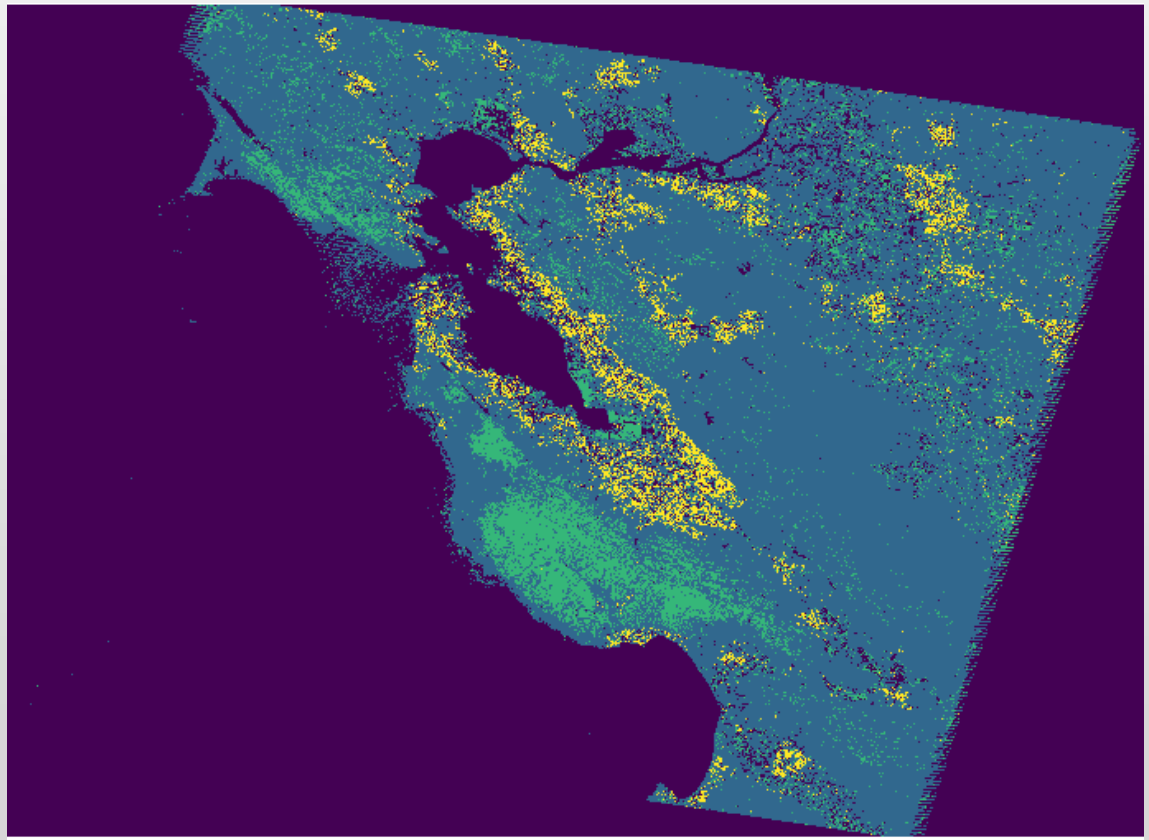


Simulation, 2017

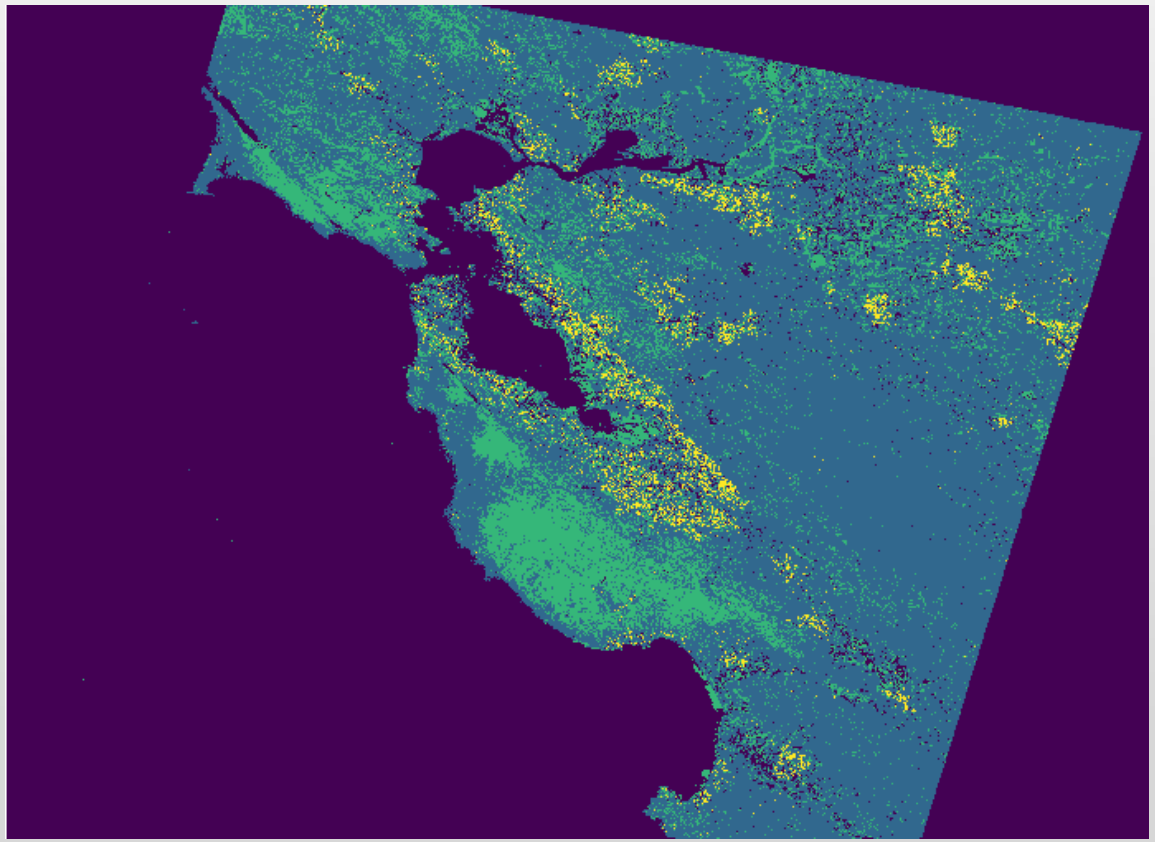


Actual data, 2017

RESULTS OF ANALYSIS, 1984 + 2007 DATA



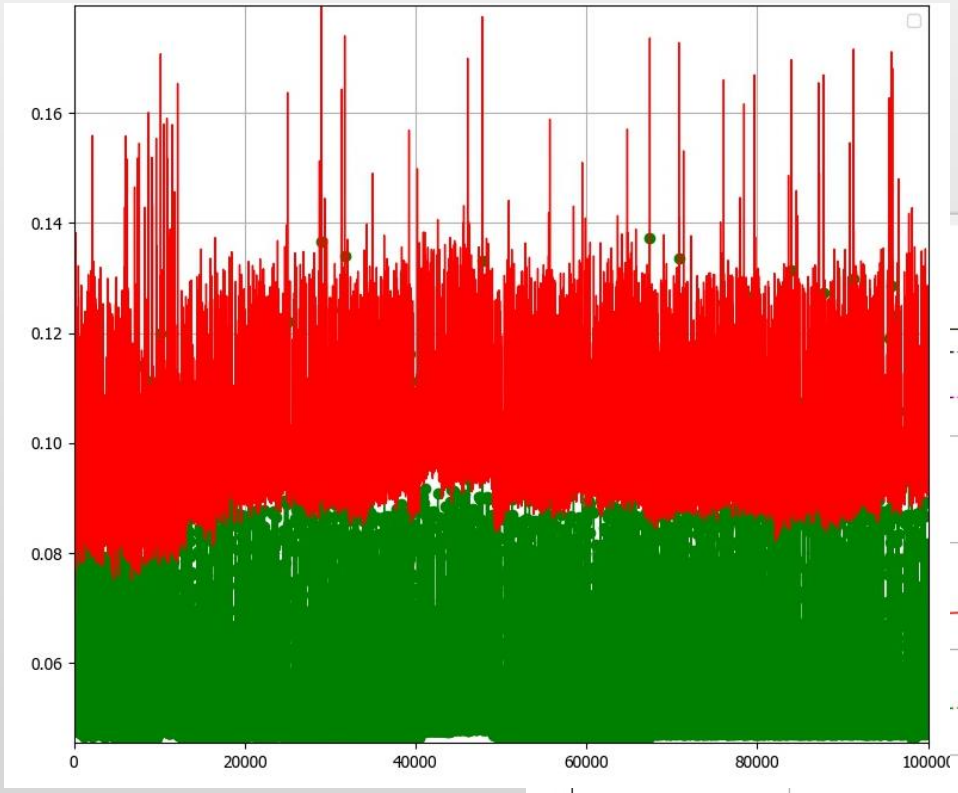
Simulation, 2017

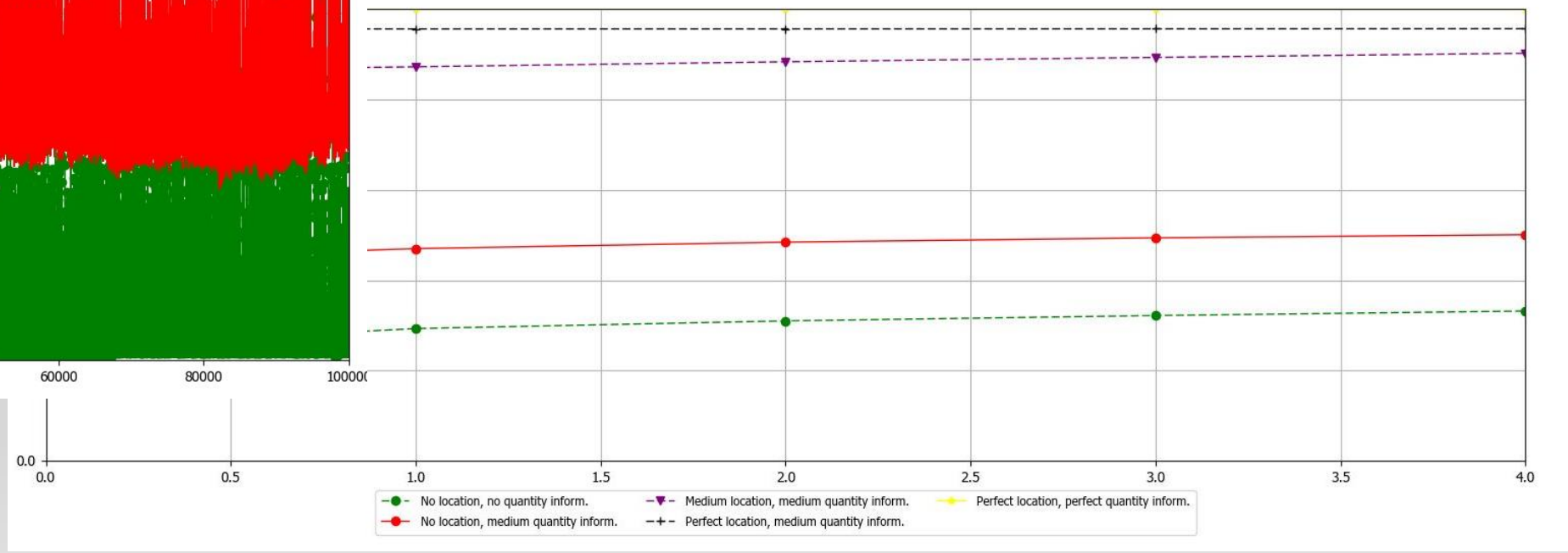


Actual data, 2017

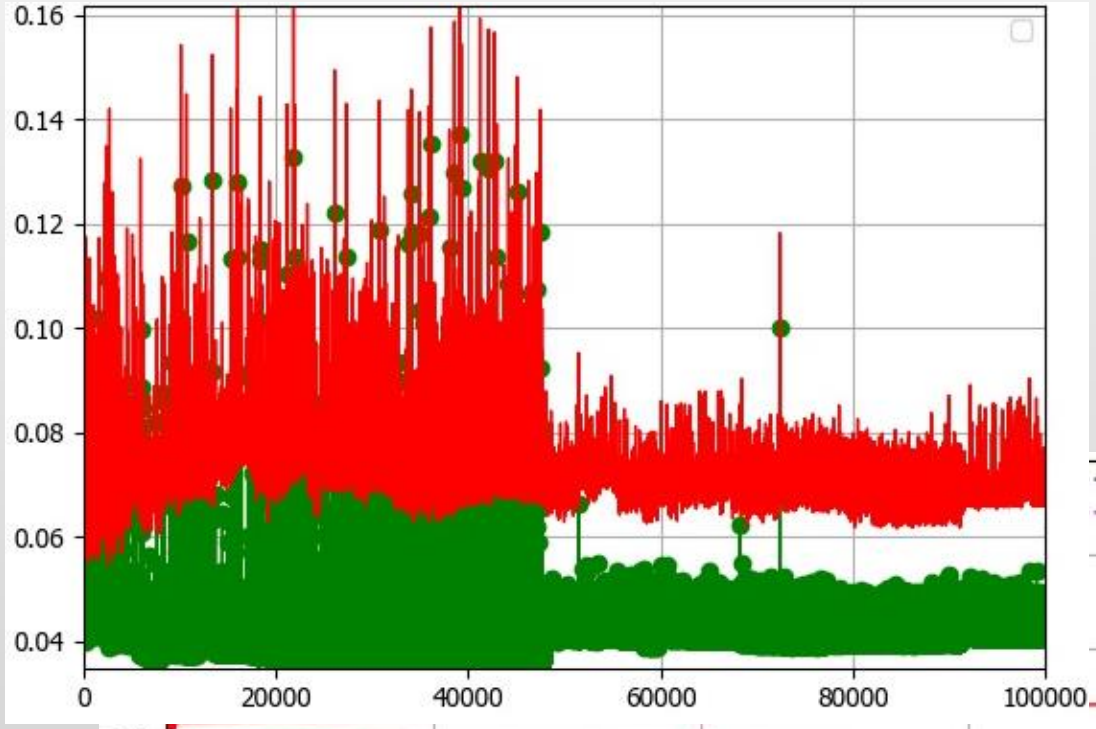
RESULTS OF ANALYSIS, 1997 + 2007 DATA

% of Correctness	86.26207
Kappa (overal)	0.75346
Kappa (histo)	0.92248
Kappa (loc)	0.81677

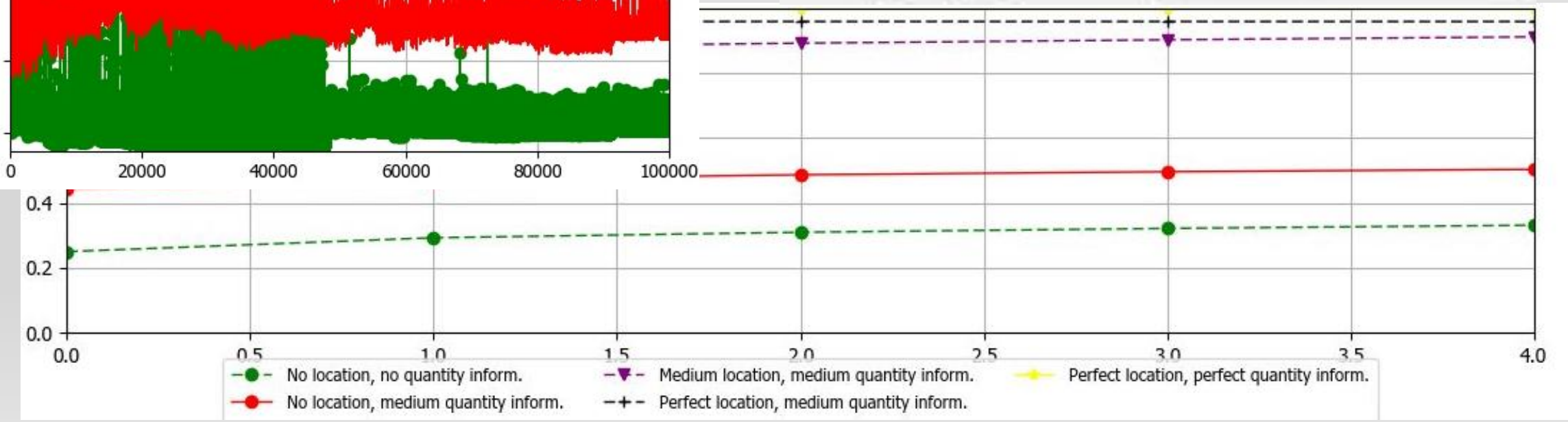




RESULTS OF ANALYSIS, 1984 + 2007 DATA



% of Correctness	86.81435
Kappa (overal)	0.76242
Kappa (histo)	0.92788
Kappa (loc)	0.82169



RESULTS OF ANALYSIS, STATISTICAL SIGNIFICANCE

- As we see, the validation data shows that my exported land use maps from the two cases are of similar values.
- If we interpret the *non*-histogram kappa values as t-test allowance values, then we see a 7-9% chance of random (false) positive. While not the best, it is acceptable.
- The delta of the kappa values (~0.01%) helps us see that this is (likely) similar enough to disregard type-1 error safely.
- It's an interesting issue to see that the neural network begins to flatten early in the 1984 comparison, but is always active throughout the 1997 comparison towards 2007.
- The lack of trials is a statistical *faux pas*.
- Of course, the null-hypothesis is never impossible. There may be no relation between the data no matter how similar the predictive validation rates.

ABOUT THOSE 100000 NODES AND ANALYSIS…

- So, how long and how strong of a computer would you need to run 1 million node prediction?
- Would there be a point (number of iterations) at which an AI prediction is unhelpful or even damaging relative the real result?

```
[Thu Nov 16 2017 23:30:41] Start logging
[Thu Nov 16 2017 23:30:49] Set intial layer to Bay Area 2007 50
[Thu Nov 16 2017 23:30:53] Set intial layer to Bay Area 1997
[Thu Nov 16 2017 23:30:55] Set final layer to Bay Area 2007 50
[Thu Nov 16 2017 23:30:59] Added factor layer nightmap_2017
[Thu Nov 16 2017 23:32:05] Class statistics and transition matrix are updated
[Thu Nov 16 2017 23:35:45] Init raster should be in PseudoColor mode. Style not applied.
[Thu Nov 16 2017 23:35:45] Change Map is created
[Thu Nov 16 2017 23:37:12] Init ANN model
[Thu Nov 16 2017 23:37:13] Set training data
[Thu Nov 16 2017 23:37:14] Start trainig ANN model
[Fri Nov 17 2017 12:44:56] ANN model trained
[Fri Nov 17 2017 12:47:39] Simulation process is started
[Fri Nov 17 2017 14:23:39] Simulation process is finished
[Fri Nov 17 2017 14:26:00] Validation process is started
[Fri Nov 17 2017 14:26:04] Validation process is started
[Fri Nov 17 2017 15:29:09] Validation process is finished
[Fri Nov 17 2017 15:40:33] Kappa validation process is started
[Fri Nov 17 2017 15:40:33] ERROR: An unknown error occurs during validation procedure
[Fri Nov 17 2017 15:41:25] Kappa validation process is finished
```

17 hours on an i7-6700HQ overclocked 7th generation CPU (@ 3.35GHz).

This will get slower as we increase node count so it is worth considering GPU use for even larger node counts.

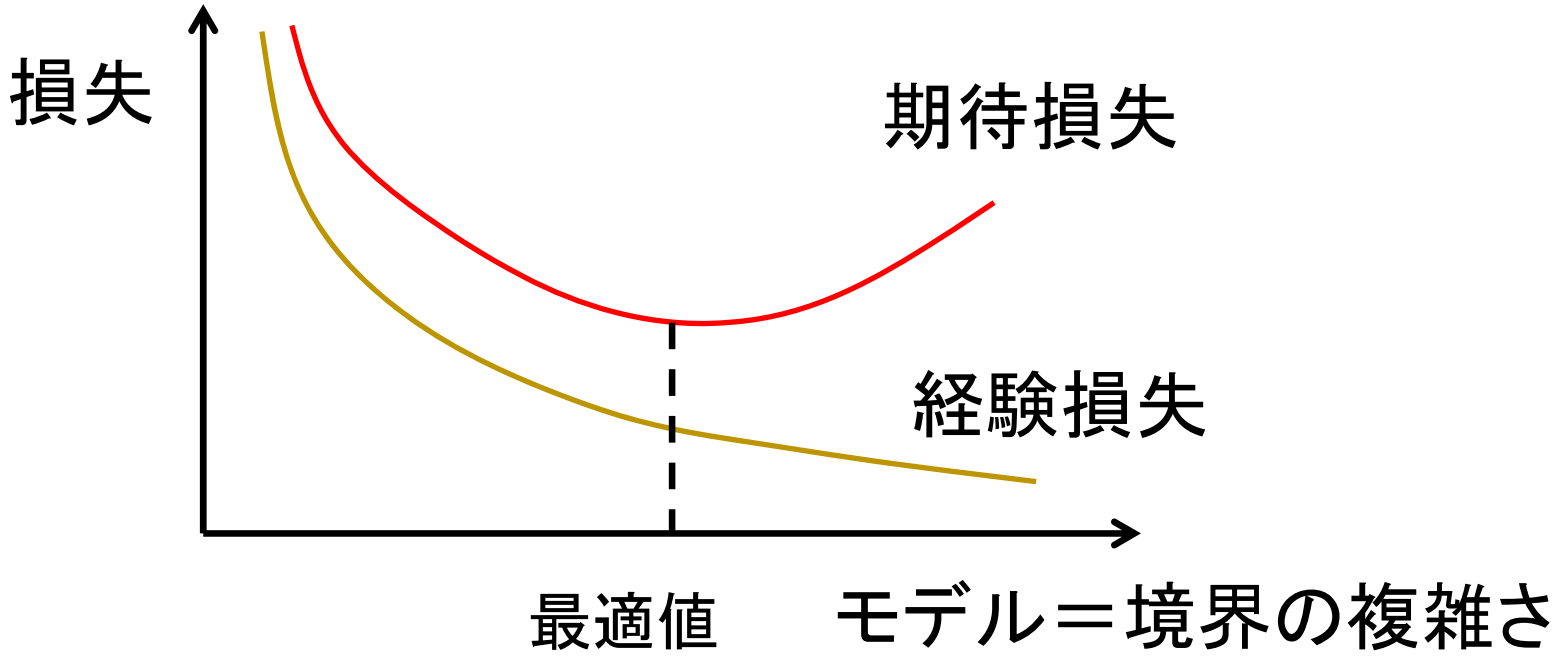
過学習（Over Learning/Training）

- 学習データに対する性能が良ければ
  評価データに対する性能も良いのか？

- そうでない場合がある
  - 学習データは確率分布からのサンプル
    ⇒　偏り（サンプルバイアス）が入っている
  - その偏りまで学習してはまずい
  - 学習用データに過剰に適応してしまうとまずい

経験損失と期待損失

- 経験損失＝学習データに対する損失
- 期待損失＝損失の期待値



ANY QUESTIONS?

For example:

- Why didn't you analyze the way waterways changed with the horrid climate in California and the landfill areas created to accommodate new housing and cities?
- Why did you give up on analyzing Tokyo?
- Is Berkeley communist?
- Thanks for listening…

SOURCES

- https://upload.wikimedia.org/wikipedia/en/3/33/Silicon_valley_title.png
- https://www.siliconmaps.com/2018/SV18-Preview.jpg
- http://cdn.abclocal.go.com/content/kgo/images/cms/211688 3_1280x720.jpg
- https://dogtrekker.com/user files/new_bay_area_map.gif
- https://i.imgur.com/AVDE4R0.jpg
- http://portal.mlslistings.com/help/files/2011/03/MLSRecipMap-4-16-10-web.png
- http://wtlab.iis.u-tokyo.ac.jp/~wataru/lecture/rs/
- https://www.youtube.com/watch?v=D7hoBTKhldE
- Presentation. Dr. Asoh, University of Tokyo. 2017 Nov 10. *Intensive Course on Artificial Intelligence, Part 2*. Excerpt.