Predicting Surf Conditions With Neural Networks: Using h2o to Predict H^2O

by Aaron Politsky
December 11, 2015

Introduction

Forecasting good surfing conditions requires constant innovation. Surfers have always been looking to get the most out of their time in the water, and having an edge on knowing when and where to surf is as much a part of the sport as is jockeying for position in a crowded lineup. Whether you're planning a trip weeks in advance, or deciding whether to wake up early for dawn patrol, forecasts are paramount. The current state of the art is to visit websites such as Surfline or Magicseaweed--the weathermen of the surf world--and digest their surf forecasts. These sites employ experts who combine weather, wind, and buoy data, develop and tune nearshore physics models, and incorporate reports from local reporters and cameras. While this may seem like a strong incumbent presence, it relies on significant human expertise. We hypothesize that machine learning may be able to augment, compete with, or even improve upon these forecasting methods, all without the domain-specific knowledge of complicated weather expertise or nearshore models of the ocean's floor and shoreline.

Current surf forecasting methods draw on data from, among other sources, hundreds of buoys deployed for scientific purposes throughout our oceans, whose data is made available by the National Oceanographic and Atmospheric Administration's (NOAA) National Data Buoy Center. These buoys float in the water above fixed geographical locations and produce readings of wind speed and direction, wave height, direction, and period, tide, air pressure, and a host of other relevant local conditions.

The goal of this project is to show that neural networks can forecast conditions at a nearshore buoy based on observations from buoys hundreds or even thousands of miles away. For instance, winter storms in the North Pacific ocean produce strong, sustained gusts of wind which send groups of waves, known as swells, across the Pacific until their energy dissipates or until they reach land. These storms cause many of the surfable waves that days later reach California and Hawaii. In theory, patterns of buoy observations in the North Pacific should be able to predict conditions of buoys near the shores of California and Hawaii. Over time, surfers and forecasters have come to learn what nearshore buoy conditions--such as wave height, direction, and period--work well for local surf locations. It stands to reason that one could train a model to predict a defined condition, e.g. waves between 6 and 12 feet arriving from the northwest every 14 to 18 seconds, days in advance, based on buoy activity in the North Pacific.

While we do not expect a neural network model to alone compete with the location resolution of Surfline's or Magicseaweed's methods, in the optimistic case, our models may be able to forecast regional surf conditions fairly well without expertise in meteorology, physics, or location specifics. Our results can serve as a proof of concept to make the practical case for incorporating machine learning into surf forecasting.

The paper is organized as follows. Part I provides background information on surf prediction, introducing common terminology and domain-specific knowledge, and how one can think about the problem. Part II describes the data and its challenges. Part III discusses our methods, including data cleaning and modeling decisions. Part IV discusses results. Part V provides conclusions and discussion of potential future improvements.

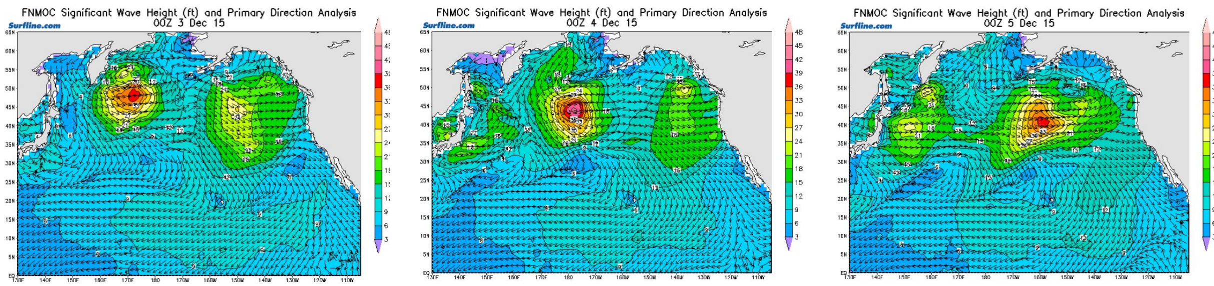
Part I: Background on the Surf Prediction Problem

Surf Conditions

Forecasting surf conditions is an integral part of the sport requiring constant innovation. The most important surf conditions are wave height, wave period (seconds between waves), swell direction, tide, and wind. Tide and wind are mainly a product of time of day, with wind conditions generally being favorable in the early morning or evening. Tides affect conditions in ways specific to a surf spot, based on the structure of the ocean's floor. We choose not to predict tide or wind. Tide is highly predictable. Wind is less predictable, but we leave this to the meteorologists. We focus on predicting swell conditions of wave height, period, and direction as consumers of our predictions could easily monitor tide and wind in conjunction with our predictions.

The Formation of Waves

Ocean storms produce strong, sustained winds which transfer their energy to the water in the form of waves. Waves radiate outward from their source, forming groups called swells, traveling largely unencumbered until landfall or until their energy decays. Swell travel time is an important consideration in forecasting because it allows us to use distant weather observations to forecast nearshore surf. Swells travel in the open ocean at around 30 nautical mph, but this varies with the energy and period of the wave with longer period waves traveling faster than shorter period waves. The further the source, the more dispersed in time the strong and weak waves will be at landfall. A forecaster might therefore observe a wide window of time in predicting surf conditions for a given time. Offshore ocean buoys give us indications of weather, while nearshore buoys indicate surf conditions. We know that by observing North Pacific weather, we can predict swell in Hawaii. The images below show how a storm's swells travel over a period of three days.



Nearshore Wave Behavior

As deepwater waves approach shore, their lower reaches begin to drag against the rising ocean floor, pushing them upwards and slowing their lower portions relative to their upper portions. This causes them to grow in height, steepness, and instability. When they reach a depth of about 1.3 times their height, they topple over and "break." Wave period actually affects wave height such that longer period waves break taller than shorter period waves. Longer period waves, with periods over 14 seconds, are called "groundswell," meaning they were formed by faraway storms, as opposed to "windswell," which were formed by local wind conditions. Groundswell is more desirable than windswell, so we incorporate period into our defined "good surf" condition. Since we are using only buoy data, we leave to the experts the modeling of interaction of waves and the nearshore ocean floor.

Part II: Data

Noaa provides about 8000 buoy-years of historical data.[1] The buoys of interest report on wave data, including average period, dominant period, dominant direction; wind data, both average and gust; and weather data, including barometric pressure, and air and water temperature.[2] The data is numeric, and typically reported hourly or half-hourly with a timestamp.

We chose to train on data from one entire year and test on another, because wanted to use periods of time having both weather activity and inactivity. Most North Pacific weather activity happens in the winter months, but we wanted to train on an entire year to get a sense of patterns occurring in different times of the year. Future efforts might extend this to train on multiple years, as years could vary in type and amount of activity (e.g. el nino vs. la nina years).

The data was not without its practical challenges. We found that many buoys are not online over the entire duration of a given year. Some are missing whole months or seasons, while others are generally online all year, but have fairly regular, short dropouts throughout the year. Some buoys report some predictors (e.g. wind speed) in one year and not another, which poses a problem since training and testing require that predictors be present in both years. In short, finding buoys in the desired geographical locations, with rich data across years posed a

[1] http://www.ndbc.noaa.gov/data/historical/stdmet/
[2] Data description http://www.ndbc.noaa.gov/measdes.shtml

practical challenge. We discuss our choice of buoys and methods of working around these challenges in the next section.

Part III: Methods

Data Arrangement

We employed the following general method to arrange our data. We chose a set of one or more source buoys and a destination buoy provided that each had rich data in at least two years. For our destination buoy, we created a factor variable called "go.surf." We defined our good surf condition to be if dominant wave height was between 2 and 5 meters, and wave period was between 12 and 20 seconds. One could easily define this differently for different surf locations and train accordingly.

For each source buoy, we formed a time series for each of its wind, wave, and weather readings. For each of these columns, we created a series of lags representing an offset window of time, with window width and offset depending on how far our source buoys are from our destination buoy.

Merging the source and destination data, we formed rows consisting of a window of time for each of our predictors associated with a binary surf condition at our destination. We could now treat our rows separately, even though a row contained the lagged data of its predecessors.

Why We Chose Neural Networks

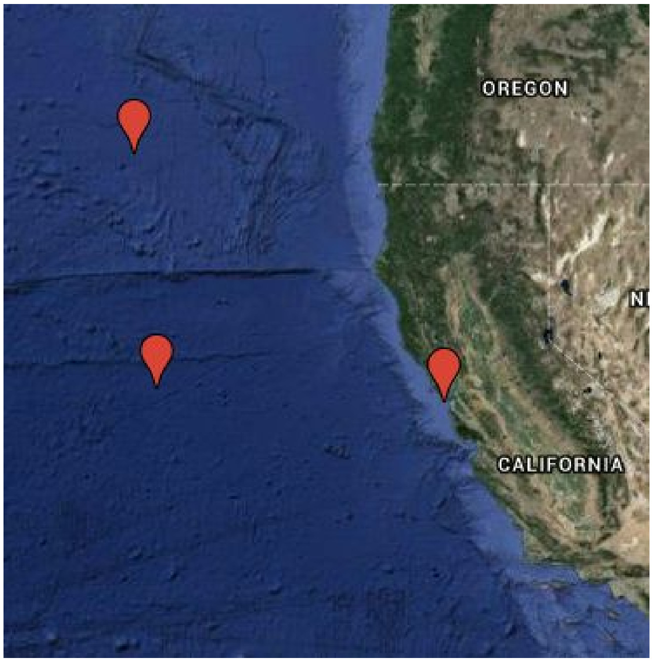
We chose to use neural networks primarily because this problem seems analogous to two of the neural network applications discussed in class and on Piazza: the LeNet5[3] MNIST handwritten character recognizer, and MarI/O[4], the automated video game player. Both methods sweep across input images, and both are highly predictive. One can consider sequential time windows of buoy data to be analogous to images, so we hypothesized that a neural net will work well in our application. Furthermore, we hypothesize that deep neural networks will be able to find patterns in our source data to identify factors causing patterns in our destination data.

Trial Experiments

Before attempting our main experiment of predicting Hawaiian surf from Alaskan buoys, we first ran some trial experiments using more proximal source and destination buoys. For our initial trial, we chose a destination buoy just off the shore of San Francisco's Ocean Beach, a popular surf spot that can handle waves between 6 and 15 feet. We chose a buoy 357 Nautical miles West of San Francisco as our source. For a second trial, we added a buoy 257 Nm West of Oregon as a second source. 2008 would serve as our training year, and 2011 our test.

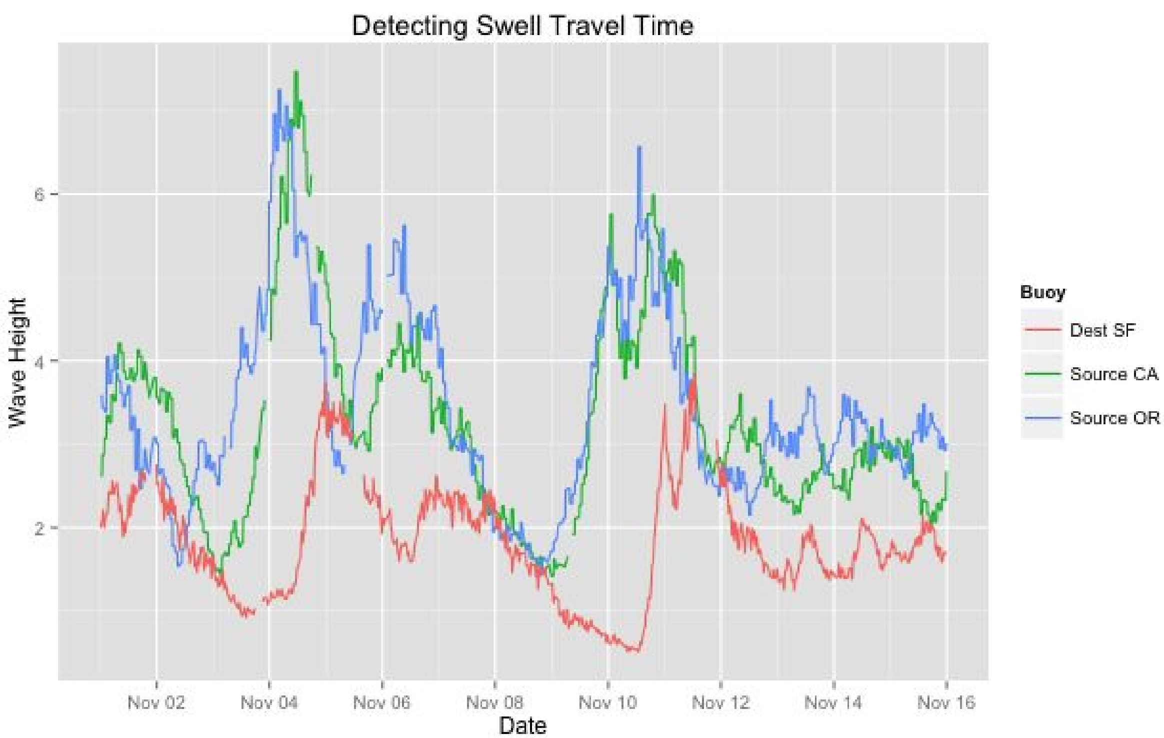
---

[3] http://yann.lecun.com/exdb/lenet/
[4] MarI/O video and explanation: https://www.youtube.com/watch?v=qv6UVOQ0F44



Based on the average deepwater swell speed of 30Nmph, we figured the most direct swell's travel time would be about 12 hours from the CA buoy and longer from the OR buoy. Depending on their direction, swells would arrive at our buoys with different delays. Rather than offset each buoy separately, for simplicity we decided to create a wide enough window to capture this variation and let our neural net determine direction, but our chosen offset could not be so long that we would be predicting yesterday's surf. This is a practical problem of choosing source buoys so near destination buoys.

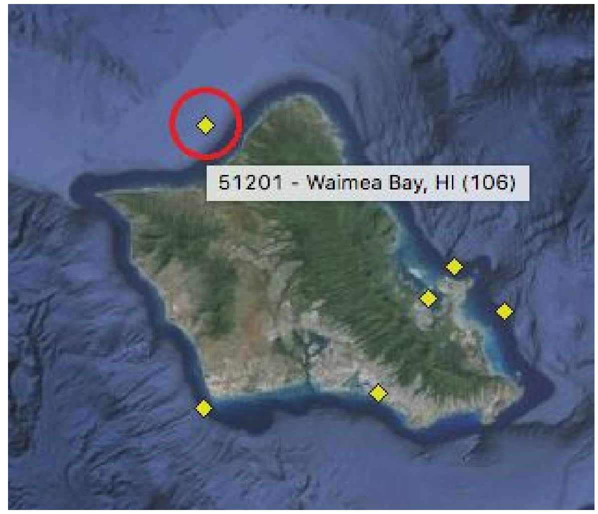
We examined plots to see if we could detect travel time, and it seemed that our destination was about 10-12 hours behind sources. We chose a 12-hour window delayed by 10 hours. In the plot below, the November 10 swell arrives at the OR just before the CA buoy, and about 10hrs before the SF buoy. It is likely coming from the WNW.



For these reasons we decided to use two input buoys in our main experiment. By using two buoys that are somewhat geographically spaced, we allow our model to better triangulate the origin and direction of weather activity. We discuss our trial results in Part IV.

Main Experiment:

Buoys several hundred miles off the West coast should be able to predict nearshore buoy activity quite well, but at the cost of a shorter forecast window. This served as our initial trial of our modeling hypothesis (see results in Part IV).



Alaskan buoys should indicate the sort of storms that produce swells that reach Hawaii. For our main experiment, we chose a buoy on the North shore of Oahu for a particular reason. California gets swells from storms originating in the the North, West, and South pacific. There are few buoys in the Southern hemisphere of the Pacific to indicate weather activity, so South swells would make our predictions much more noisy. The Waimea Bay buoy on the North shore of Oahu is shielded from Southern swells by the island, so it will only pick up swells from storms in the North Pacific (primarily in Winter).

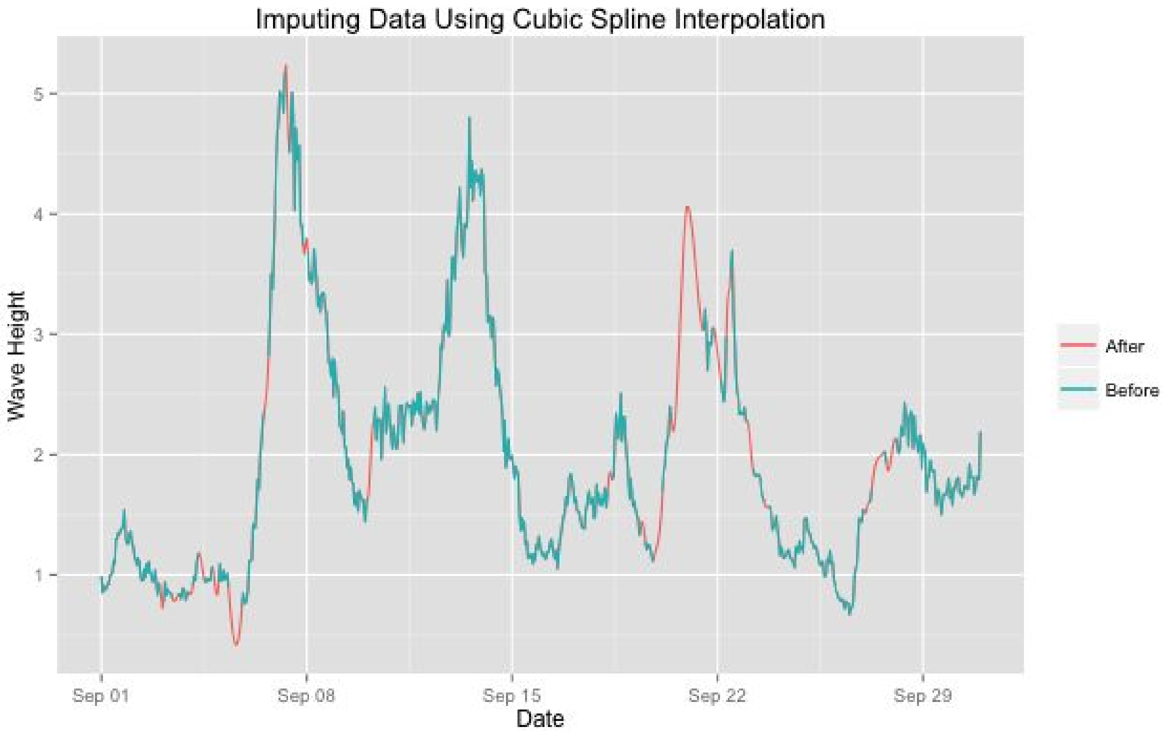
Data Challenges and Cleaning

Our main tasks with the data were to identify buoys having rich enough data over several years, arrange it on a standardized hourly time grid, and impute reasonably sized gaps in the data.

We were able to identify two Alaskan buoys which had rich data in 2007 and somewhat more sparse data in 2014. 2007 would serve as our training year and 2014, while more sparse, our test. The HI buoy has nearly complete data in many years, so it did not pose a problem. As for columns, we could not find a pair of years each having wind data, so we excluded it. Wave height and period, air pressure, and air and water temperature were to be our predictors.

A closer look at the previous plot (Detecting Swell Travel Time) reveals discontinuities in the data. When lagging data, any missing value would render many lagged rows incomplete, so we needed to impute small gaps in the data. A naive way to fill small gaps was to repeat the previous value. This worked for small gaps, but our Alaskan test data had many gaps, some longer than a few hours. Furthermore, we used much longer lag windows in our main experiment due to swell travel dispersion, and so needed a better solution. Our chosen lag window used 96 two-hourly lags. This would reduce our complete rows from around 77% to 9%. (Further experiments might vary the lag width, resolution, and offset, to better identify actual swell travel time).

Our solution was to use cubic spline interpolation for gaps of up to 36 hours.



This brought our test set completeness to 87%, with missing values entirely at ends of the year. We were now ready to train and test models.

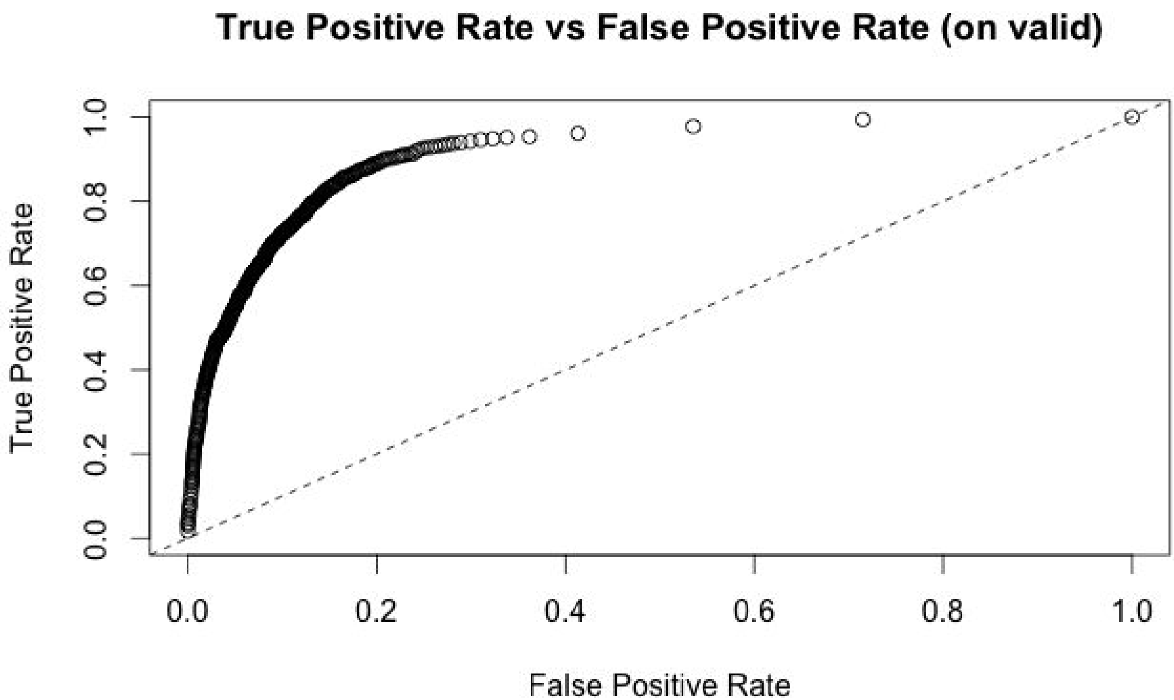
Model Selection

We used the h2o package to train deep learning neural network models. We initially experimented over a grid varying the number of neurons and hidden layers, activation function, input dropout ratios, l1 regularization values, and epochs. Initial grid results showed that models having 1024 neurons, whether they had hidden layers or not, required the longest training runtime but offered no clear out of sample performance advantage over our smaller models. Since having more input neurons helps with nonlinearities in the data, this suggests that our data does not require such flexibility. Eventually, we settled on grid consisting of 64, (64,64), (64,64,64), and 256 neurons, also varying Tanh activation with and without dropout, input dropout ratio, and l1 regularization. (See appendix for grid.)

Part IV: Results and Discussion

Trial Experiment Results

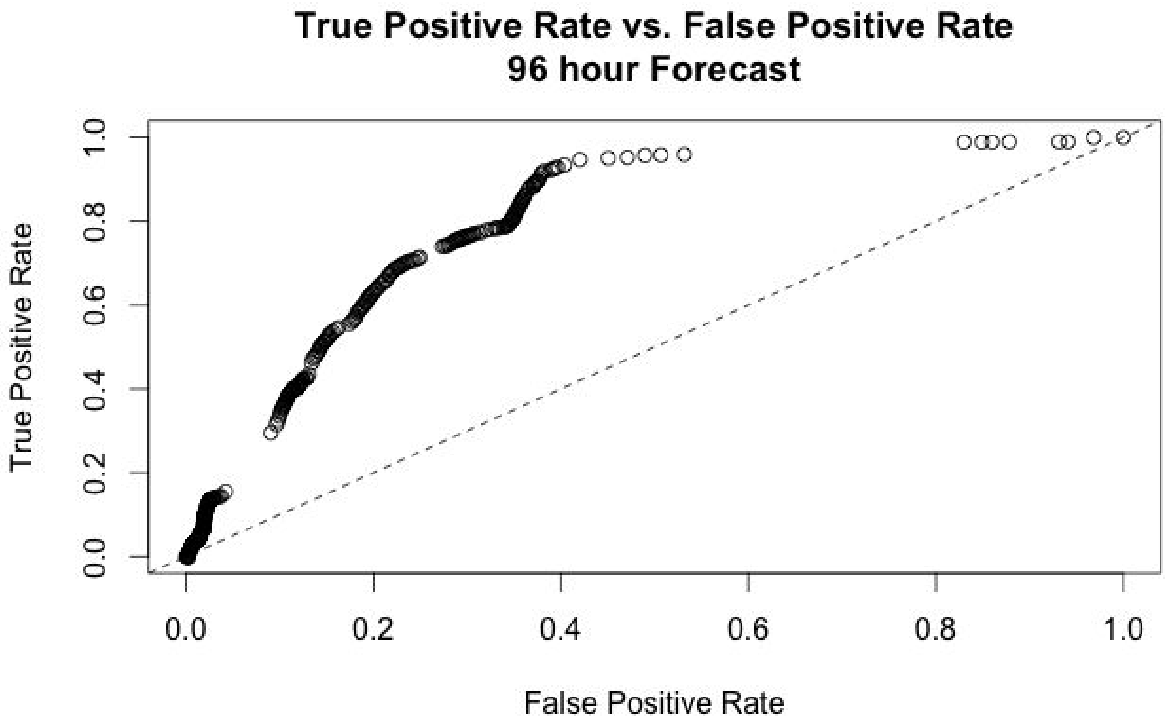
Our trial experiment to predict San Francisco surf from offshore California and Oregon buoys yielded promising results. Our best model's error rate was 14.5% at a threshold of 0.123. It employed a single layer of 64 neurons, 2 epochs, Tanh activation with Dropout, and l1 regularization. Its ROC curve is shown below.



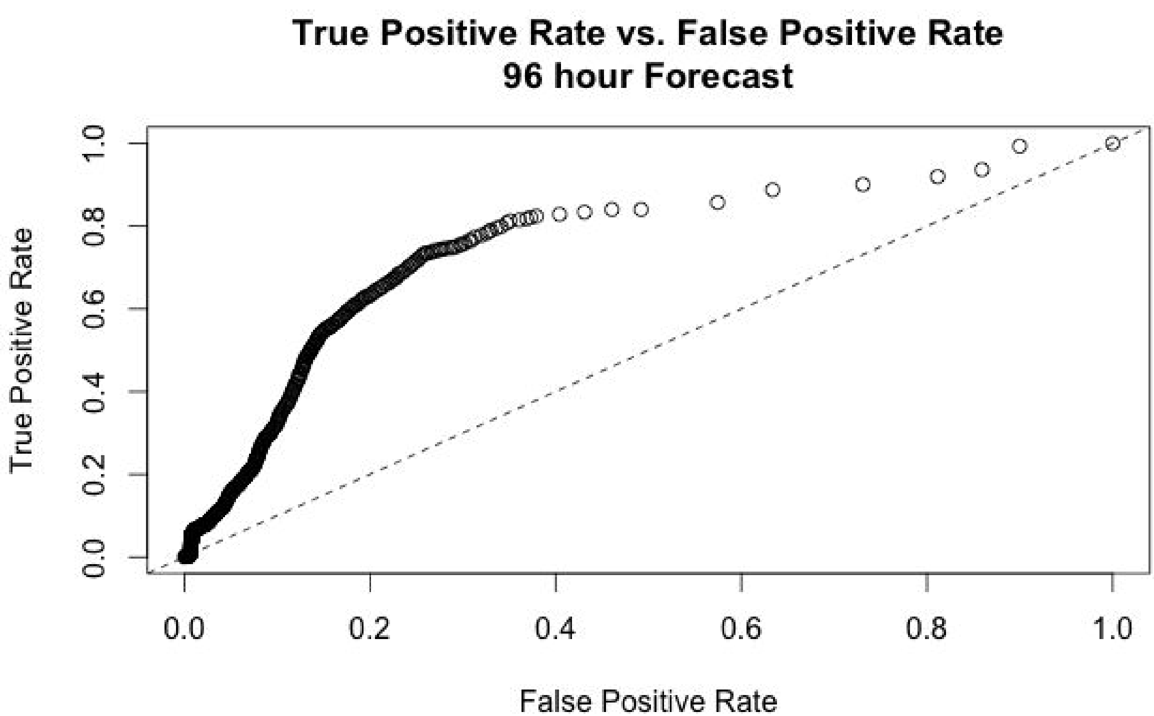
This suggested that our neural net methodology could work. While we were pleased with this result, we remained skeptical that our approach would be valid when predicting over much longer distances. Predicting using buoys so near to shore was perhaps too easy of a problem.

Main Experiment Results

Our main experiment was to predict 2014 Hawaii surf based on Alaska buoy data using a model trained on 2007 data. Our best model exhibited a 24% error rate at a threshold of 0.0856. It employed 64 neurons in its input layer, a single 64-neuron layer, Tanh activation, l1 regularization of 0.2, and 2 epochs. Its ROC curve is shown below.



This ROC curve is a bit irregular, which may be difficult to interpret. Another model had an error rate of 26% at a threshold of 0.0283, using only a single layer of 64 neurons, Tanh activation with 20% input dropout and 0.2 l2 activation. Its ROC curve is below:



Part V: Conclusions and Future Improvements

Our goal was to show that neural networks can forecast surf conditions at a nearshore buoy based on observations from buoys hundreds or even thousands of miles away. Overall, our results seem to validate this hypothesis as a proof of concept. Neural networks can indeed use buoy data alone to predict surf conditions well in advance of landfall without a drop of expertise.

Future research may include experimenting with other prediction methods (such as Hidden Markov Models, or tree methods), using more than two source buoys, and determining the optimal relationship between source-destination buoys, window width, and offset time. Ideally, one could move on from a classification problem to a predicting continuous measures of wave height, wave period, and direction, rather than a defined condition. Future researchers may also look to incorporate other, non-buoy data if available, such as satellite weather or wind data.

Appendix A

Final Hyper Params Grid:

	epochs	hidden	activation	input_dropout_ratio	l1
1	2	64, 64, 64	Tanh	0.0	1e-05
2	2	64, 64	Tanh	0.0	1e-05
3	2	256	Tanh	0.0	1e-05
4	2	64	Tanh	0.0	1e-05
5	2	64, 64, 64	TanhWithDropout	0.0	1e-05
6	2	64, 64	TanhWithDropout	0.0	1e-05
7	2	256	TanhWithDropout	0.0	1e-05
8	2	64	TanhWithDropout	0.0	1e-05
9	2	64, 64, 64	Tanh	0.2	1e-05
10	2	64, 64	Tanh	0.2	1e-05
11	2	256	Tanh	0.2	1e-05
12	2	64	Tanh	0.2	1e-05
13	2	64, 64, 64	TanhWithDropout	0.2	1e-05
14	2	64, 64	TanhWithDropout	0.2	1e-05
15	2	256	TanhWithDropout	0.2	1e-05
16	2	64	TanhWithDropout	0.2	1e-05
17	2	64, 64, 64	Tanh	0.0	2e-01
18	2	64, 64	Tanh	0.0	2e-01
19	2	256	Tanh	0.0	2e-01
20	2	64	Tanh	0.0	2e-01
21	2	64, 64, 64	TanhWithDropout	0.0	2e-01
22	2	64, 64	TanhWithDropout	0.0	2e-01
23	2	256	TanhWithDropout	0.0	2e-01
24	2	64	TanhWithDropout	0.0	2e-01
25	2	64, 64, 64	Tanh	0.2	2e-01
26	2	64, 64	Tanh	0.2	2e-01
27	2	256	Tanh	0.2	2e-01
28	2	64	Tanh	0.2	2e-01
29	2	64, 64, 64	TanhWithDropout	0.2	2e-01
30	2	64, 64	TanhWithDropout	0.2	2e-01
31	2	256	TanhWithDropout	0.2	2e-01
32	2	64	TanhWithDropout	0.2	2e-01

Code in Appendix from RMarkdown.