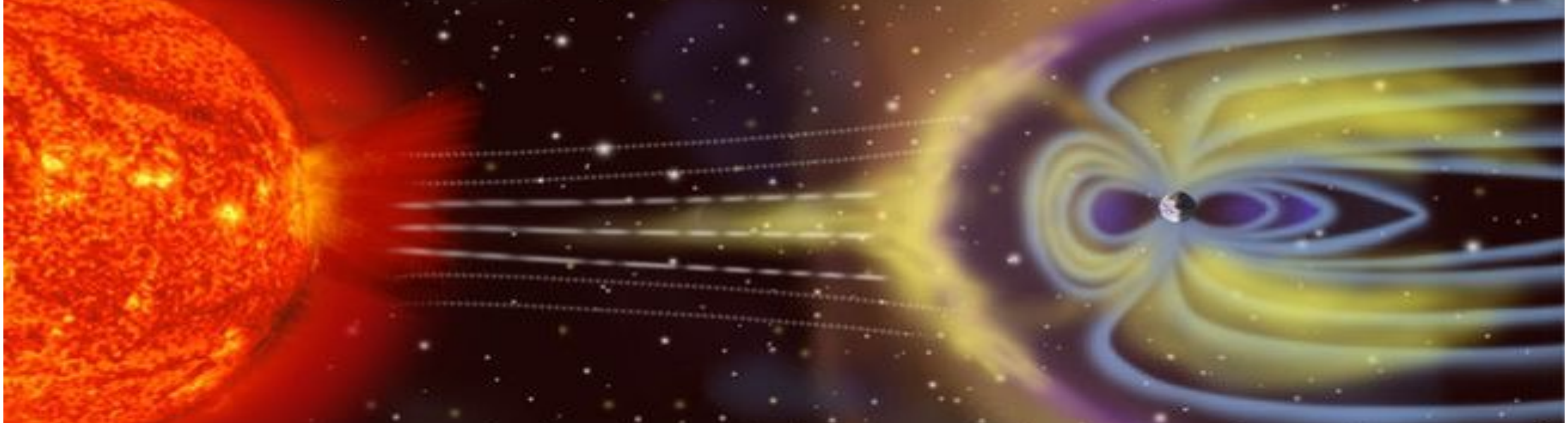


Statistical Modeling and Machine Learning for Space Physics

Steven K. Morley

Space Science and Applications (ISR-1), Los Alamos National Laboratory

What ground will I cover?

Background and focus topics

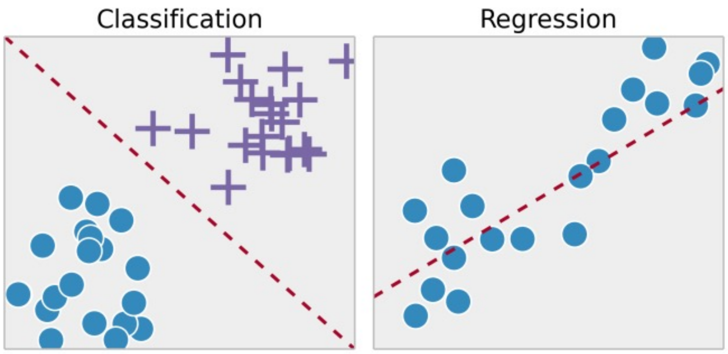
- A few different types of problem
- Using different techniques to tackle a simple problem
- Example 1: Plasmaspheric number density
- Example 2: Kp index
- Example 3: GNSS time-difference-of-arrival
- Final Thoughts

Problem Classes

Fitting statistical models

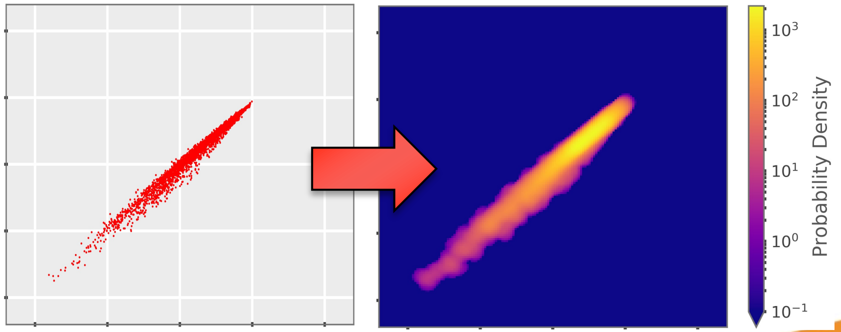
Basic classes of problem include:

- Regression
  - Predict y = f(x)
  - OLS, NNs, superposed epoch analysis, GLMs
- Classification
  - Predict label of data given (x, y)
  - Algorithm selection, decision trees,
- Density Estimation
  - Estimate P(x, y) and P(y|x)
  - E.g., Distribution fitting, mixture models, kernel density estimation



scorecardstreet.wordpress.com

Density Estimation



What's the difference between applied statistics and machine learning?

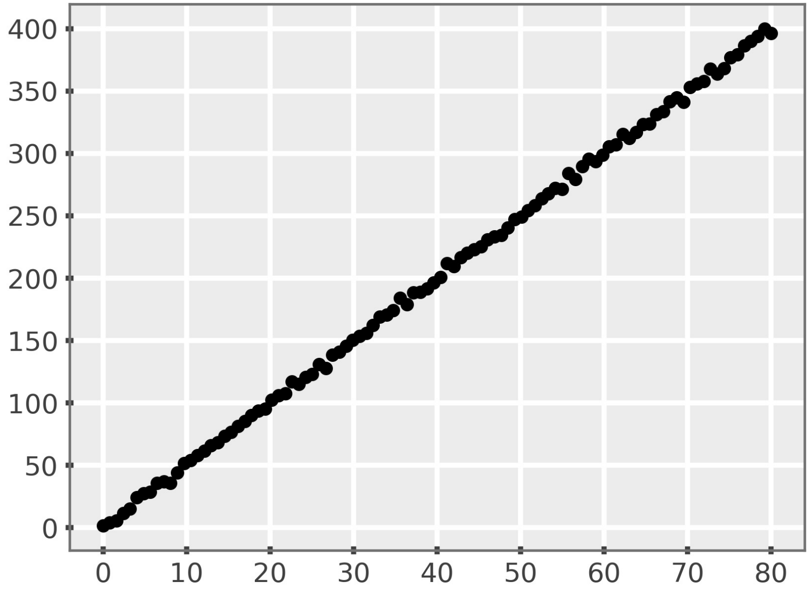
What's the difference between applied statistics and machine learning?

"Machine learning is where only the machine learns something"

- Paul O'Brien

Regression

Fitting functions: Linear regression example

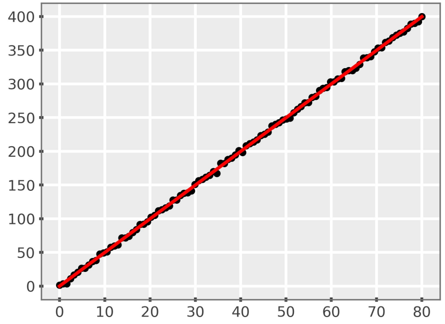


Starting with a simple linear model

- Generate training data
- $y = 5x + \varepsilon$
- $\varepsilon = \mathcal{N}(0, 2.5)$

Solving the same problem in different ways

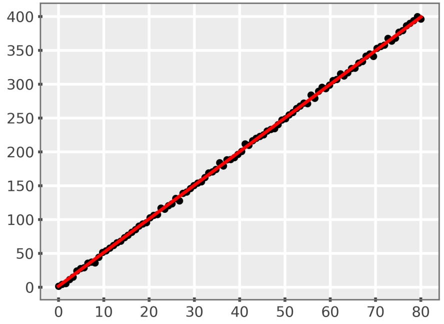
Statistical modeling or machine learning?



OLS:

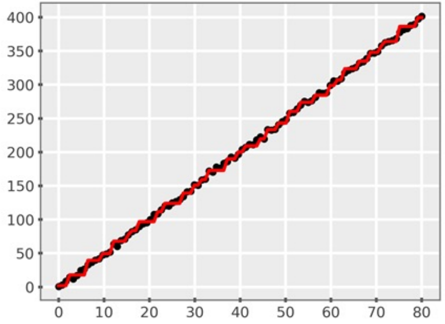
$$m = \frac{Cov(x,y)}{Var(x)}$$

$$c = \bar{y} - m\bar{x}$$



Neural network:

- Single layer, linear activation, mean squared error penalty
- OLS the hard way



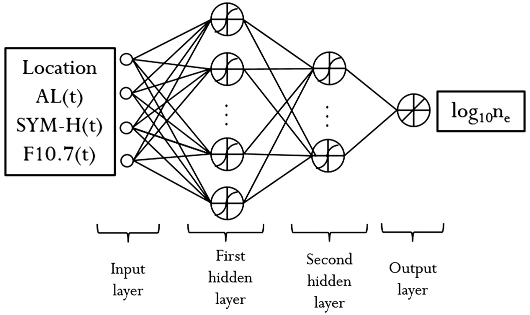
Decision Tree:

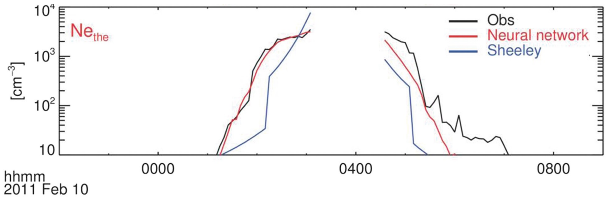
- 5 layers
- No concept of slope or intercept

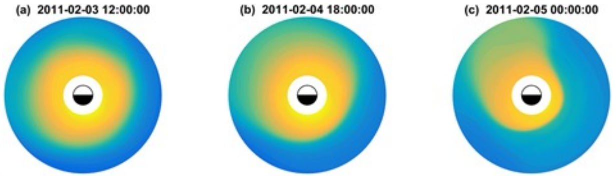
Neural Networks Fit Functions

Example Application 1: Plasmasphere number density

Chu et al., JGR-Space, 2017a,b







- A neural network with a linear activation function is equivalent to linear regression, it's just solved "the hard way"
- NNs provide a flexible framework for finding nonlinear mappings from inputs to outputs
- Chu et al. (2017a,b) model has a 180-element input vector including time histories, 2 hidden layers
- Predictions outperform older empirical model and show physical behavior like erosion and refilling

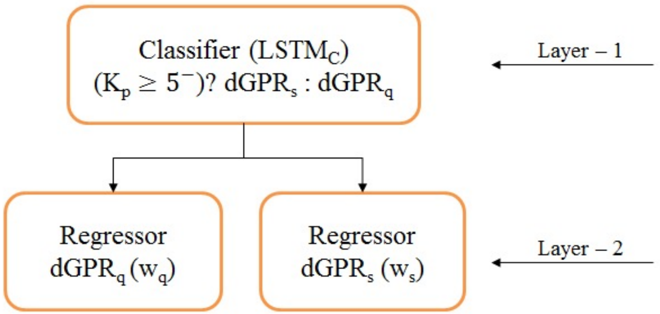
Example Application 2:

Probabilistic prediction of the Kp index

Shibaji Chakraborty and Steve Morley

Probabilistic prediction of Kp index

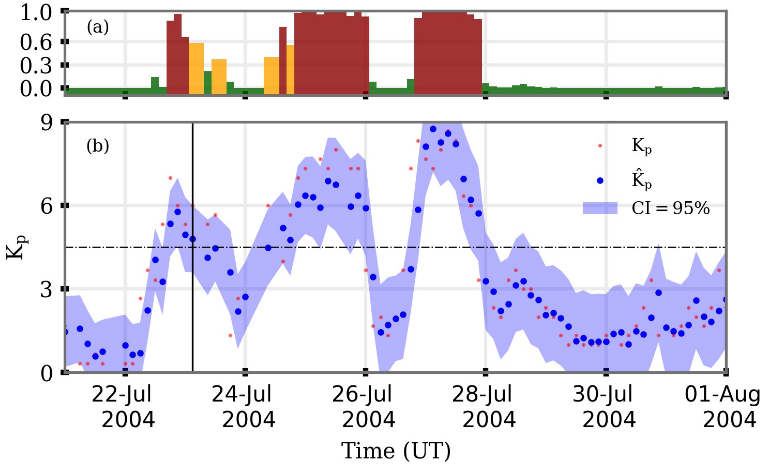
Hybrid architecture



- Convection (and hence Kp) response is different during storms
- Use a recurrent neural network to classify sequence of input vector
  - *LSTM: Long short term memory*
- Use a deep Gaussian process for probabilistic regression
  - A Gaussian process models a distribution of functions
  - A deep GP uses *nested* GPs

Probabilistic prediction of Kp index

Hybrid architecture



- Gaussian Process gives output distribution
- Uncertainty bounds let us calculate probability of exceeding a given Kp
- *Including solar X-ray flux in the input vector improves storm-time prediction*

Example Application 3:

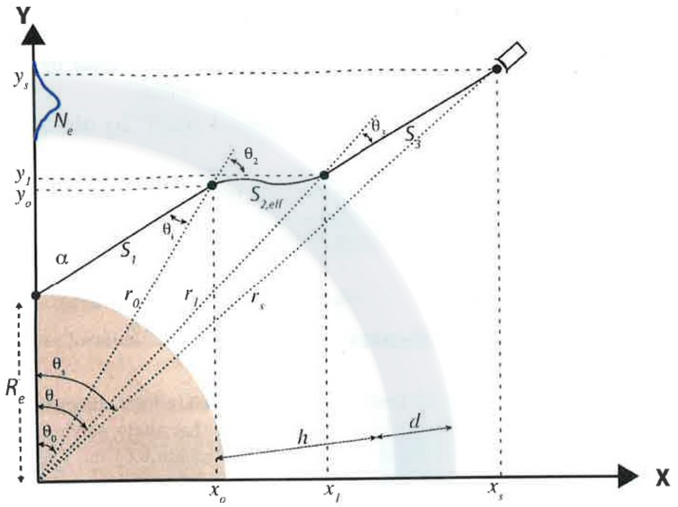
Predicting model error for radio propagation

Steve Morley, Erin Lay, et al.

Model error in time-difference-of-arrival for radio

Use density estimation methods

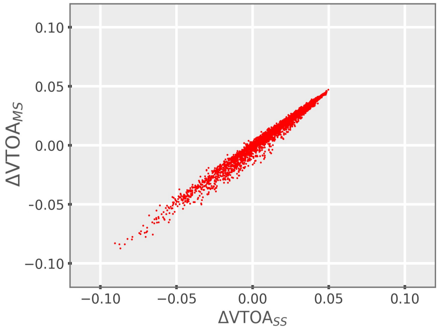
Lay et al., LA-UR-18-20440, 2018

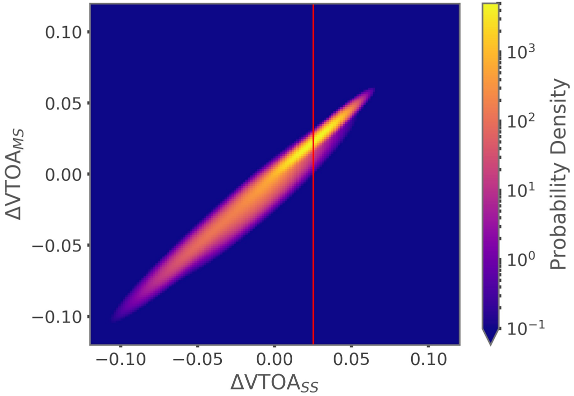


- Time-difference of arrival (TDOA) is used for satellite-based augmentation of GNSS, geolocation
- Single-shell ionosphere models are often used
- For given TDOA from single shell-model, estimate what we would have seen with a *ground truth* (multishell) model

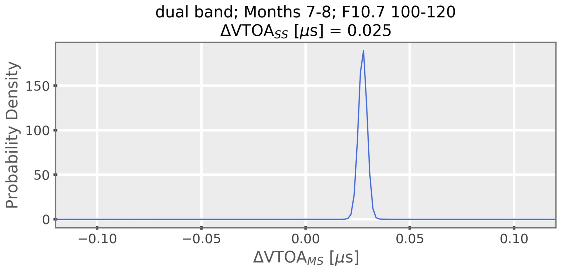
Model error in time-difference-of-arrival for radio

Bayesian Gaussian Mixture Model





- Uncertainty in arrival time of radio signal propagating through ionosphere
  - Compare multi-shell to single shell ionospheric raytracing model
- Use density estimation to find P(x,y) and P(y|x)



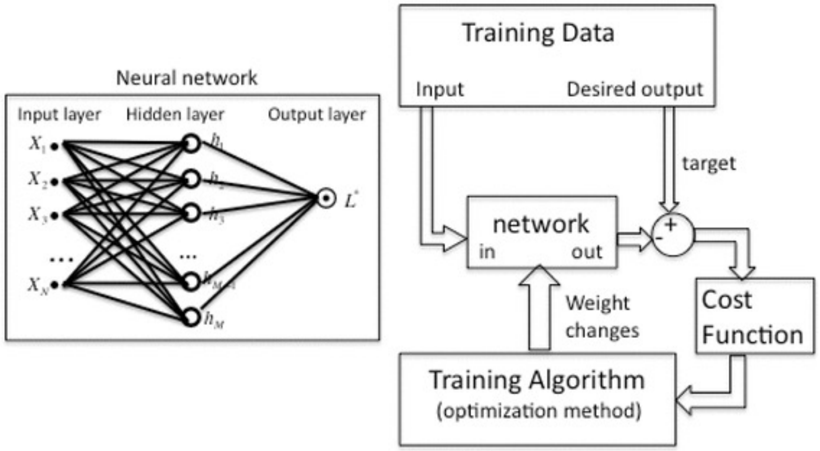
“Statisticians have to become opportunistic.” – *J. Rice*

“…faced with a problem, they must find a reasonable solution by whatever method works.” – *L. Breiman*

“Numerical experimentation by itself, unguided by theory, is prone to faddish wandering:
Rule 1. New methods always look better than old ones.
…
Rule 2. Complicated methods are harder to criticize than simple ones.
…” – *B. Efron*

Can the human learn something too?

ML outcomes can bring insight; SM methods can bring insight



Yu et al., Space Weather, 2012

- ML can be thought of as *application* of flexible *algorithmic* methods with the aim of *prediction*
- Statistical modeling has general aims of *describing* and *understanding* a system
- *“The whole point of science is to open up black boxes, understand their insides, and build better boxes for the purposes of mankind.”* – B. Efron

Machine Learning and Statistical Modeling

Approach and Resources

Key questions:

- What problem do I need to solve?
- Does my method need to be interpretable?
- How should I judge the performance of my model?
- Which methods are appropriate for my data/problem?

- Software: Scikit-learn; Tensorflow; Torch; Edward; PyMC3; …
- Literature:
  - Statistical Methods in the Atmospheric Sciences. D.S. Wilks
  - Forecast Verification: A practitioner's guide in atmospheric science. Ed. Jolliffe & Stephenson
  - Machine Learning Techniques for Space Weather. Ed. Camporeale, Wing, & Johnson
  - Snakes on a Spaceship—An Overview of Python in Heliophysics, Burrell et al., JGR-Space, 2018
  - Statistical Modeling: The two cultures, Breiman, Statistical Science, 2001

Los Alamos NATIONAL LABORATORY EST. 1943