Agriculture Project

Christian Davis and Matthew Oehler

Brigham Young University

March 5, 2018

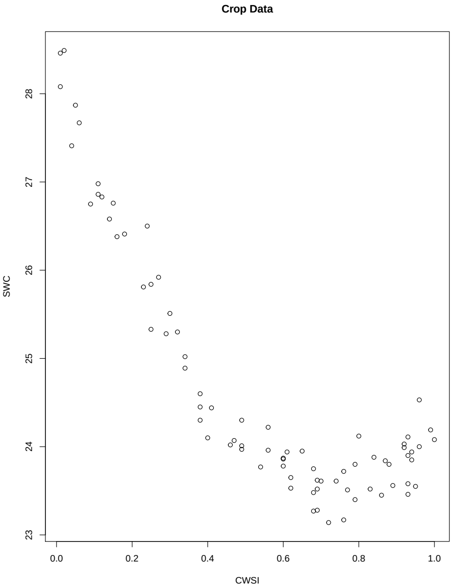
Problem Intro

Background

- Water scarcity is a problem for many regions of the United States
- Crop water stress index (CWSI) is widely used indicator to estimate the crop water status
- Soil water content (SWC) helps indicate how much water to add
- CWSI is much easier to measure

Goal: Use the CWSI to predict the SWC

Agriculture Data

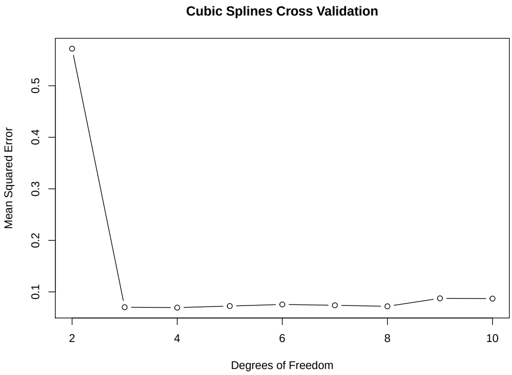


Problems/Issues

- Data is not linear

Solutions:

- Polynomial Regression
- Basis Function Expansion: B-splines
- Kernel Smoothing



Model

Polynomial Regression

$$\mathbf{y} = \beta_{\mathbf{0}} + \beta_{\mathbf{1}}\mathbf{x} + \beta_{\mathbf{2}}\mathbf{x}^2 + \epsilon$$

$$\text{where } \epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$\begin{aligned}
\mathbf{y} &= \text{response vector} \quad (nx1) \\
\mathbf{x} &= \text{CWSI values} \quad (1xp) \\
\boldsymbol{\beta} &= \text{model coefficients} \quad (px1) \\
\boldsymbol{\epsilon} &= \text{errors} \quad (nx1)
\end{aligned}$$

Model Assumptions

$$\mathbf{y} = \beta_{\mathbf{0}} + \beta_{\mathbf{1}}\mathbf{x} + \beta_{\mathbf{2}}\mathbf{x}^2 + \epsilon$$

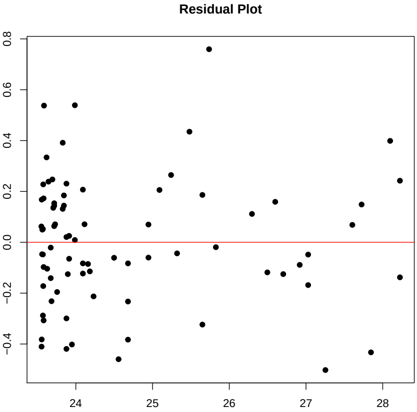
Still just a linear model – linear in the βs

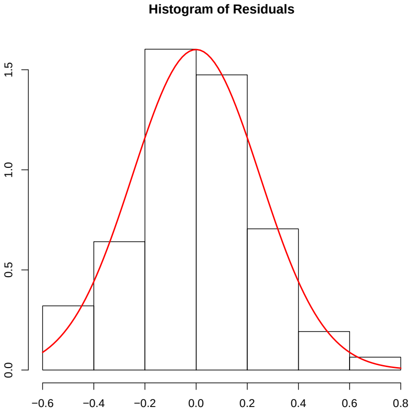
- Independence
- Normality of the Standardized Residuals
- Equal Variance

How Model Achieves Goals

- Easily fit non-linear data
- Can predict SWC from CWSI simply

Model Assumptions





Model Justification

Degree of Polynomial Regression

- x^3 term is not significant
- Performed cross validation to select the degree of polynomial regression

Comparison of Cubic and Quadratic Models:

	Cubic	Quadratic
Bias	8e-04	-6e-04
RPMSE	0.2591	0.2526
Coverage	0.9457	0.949
PIW	1.0285	1.0204
R^2	0.9689	0.9689

Results

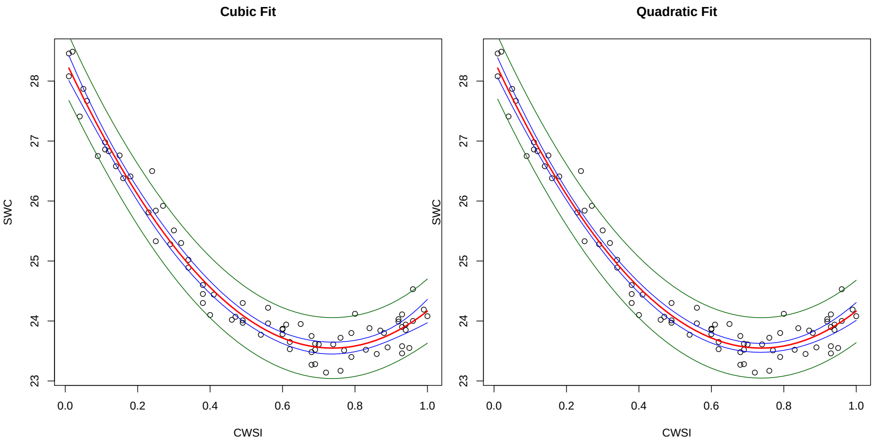
Table of Significant Quadratic Model Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	28.3500	0.0900	328.87	0.0000
poly(cwsi, 2, raw = T)1	-13.0200	0.3900	-33.57	0.0000
poly(cwsi, 2, raw = T)2	8.8300	0.3700	24.03	0.0000

Note: Practical interpretability of coefficients is lost when using polymonial regression.

	Quadratic
Bias	-6e-04
RPMSE	0.2526
Coverage	0.949
PIW	1.0204
R^2	0.9689

Results


Cubic Fit
Quadratic Fit
SWC
CWSI
23
24
25
26
27
28
0.0
0.2
0.4
0.6
0.8
1.0

Model Application

How much water to add for a CWSI of 0.4?

$$28.35 + -13.02 \times \mathbf{0.4} + 8.83 \times \mathbf{0.4}^2 = 24.56$$

Conclusions

Conclusions:

- Goals:
  - We fit a model that can be used by farmers to optimize their water usage
- Shortcomings:
  - Loss of coefficient interpretability
- Next steps:
  - Gather more data to be explore the relationship of SWC and volume of water sprinkled.

Distribution of Work

Problem Statement and Understanding	Christian
Describe the method/model(s) that are used . . .	Christian
Model Justification and Performance Evaluation	Matt
Results	Matt
Conclusions	Joint
Code	Joint