Youngstown Temperature Forecast

David Gessler and Nicole Zimmerman
Advised by: Dr. Moon Nguyen and Dr. Thomas Wakefield
Youngstown State University, Youngstown, OH

Abstract

We predicted temperatures in Youngstown over a 30- year period (1987-2017). We used several regression models to fit the data and determined the error between the values predicted and the actual values to ultimately choose the best model. The results indicated that mean temperature sine model is the best model, among the four chosen models, for predicting weather.

Introduction

Predict temperature in Youngstown, using 30-year data, 1987-2017:

- 25-years for training, 5-years for testing
- Four months represent the four seasons in a year: March, June, September, December
- Data frequency: annual

Data sources:

- The Actuaries Climate Index (ACI)
- Weather Underground

Models: Four regression models

Mathematical Models

Linear: y=ax+b

Exponential: $y=ae^{bx}$

Sine: y=asin(bx+c)+d

Mean Temperature Sine: y=asin(bx+c)+dx+e

- For our variables:
  - a,b,c,d,e are constants and are our model parameters
  - x represents the number of years from 1987, starting at 1987 when x=0
  - y is our resulting temperature in degrees Fahrenheit

Model Evaluation

Absolute Percentage Error (APE) = $\frac{1}{\bar{r}}\sum_{i=1}^{N}\frac{|r_i-\tilde{r}_i|}{N}$

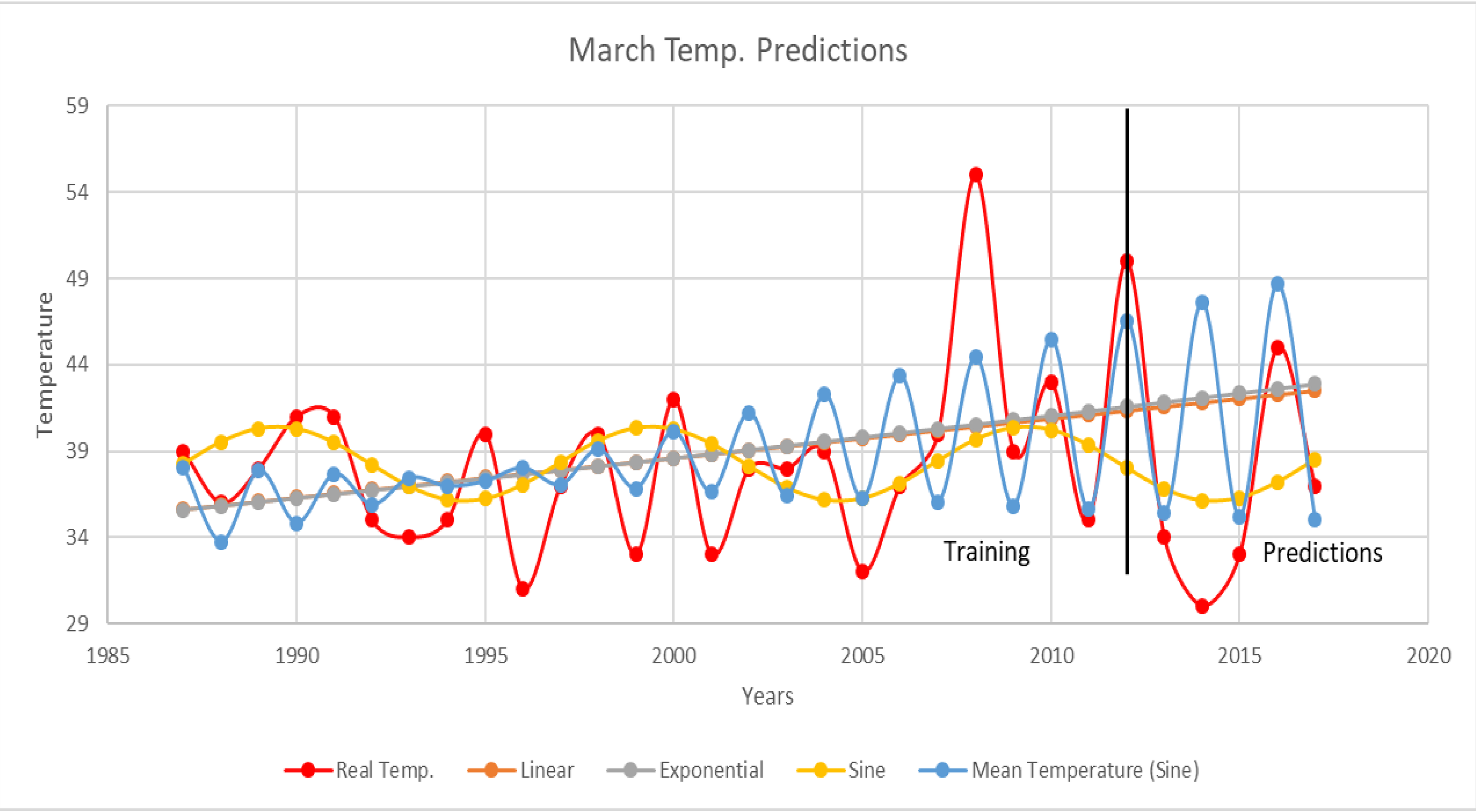
Average Absolute Error (AAE) = $\sum_{i=1}^{N}\frac{|r_i-\tilde{r}_i|}{N}$

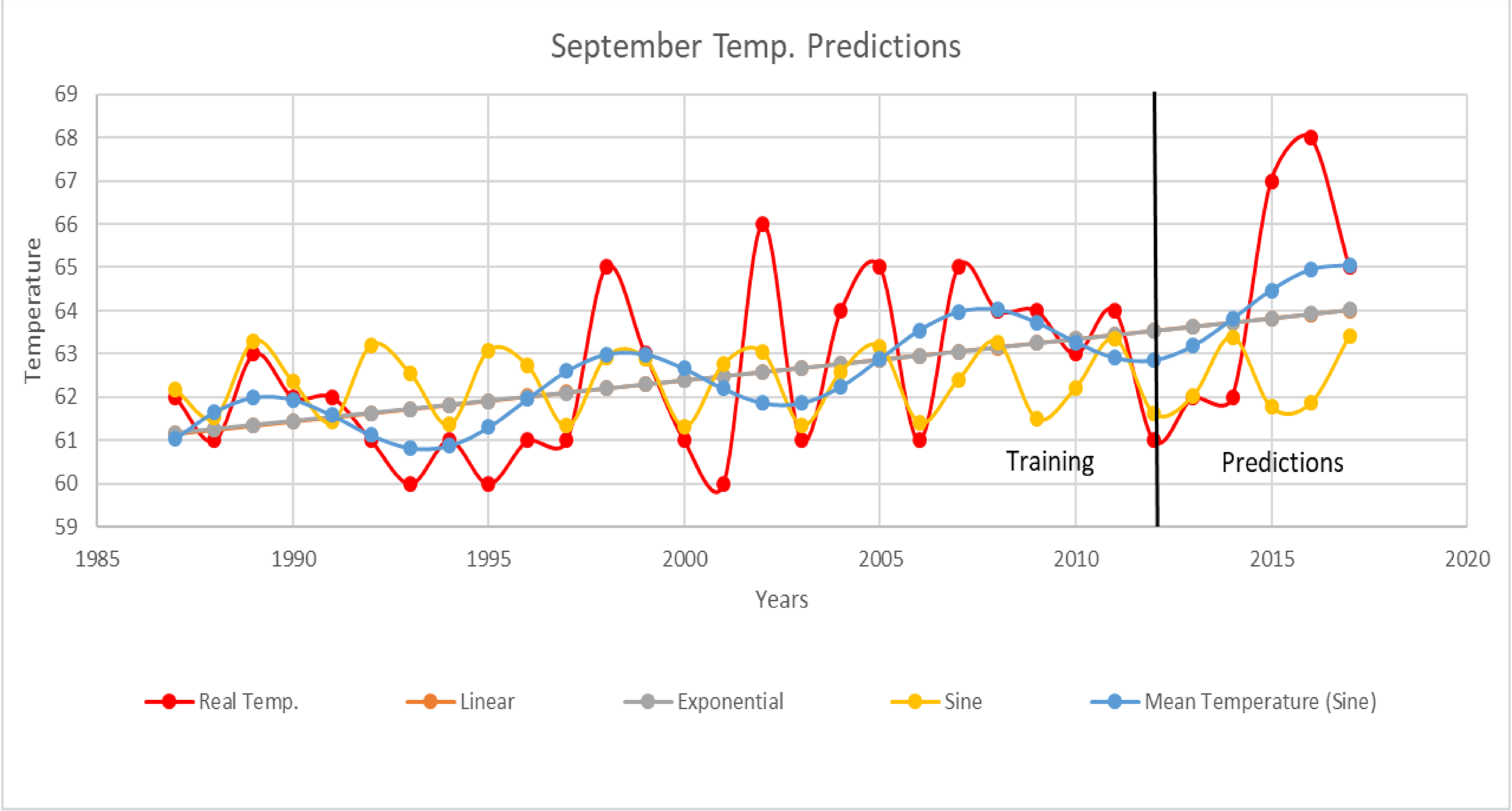
Average Relative Percentage Error (ARPE) = $\frac{1}{N}\sum_{i=1}^{N}\frac{|r_i-\tilde{r}_i|}{N}$

Root-Mean-Square Error (RMSE) = $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_i-\tilde{r}_i)^2}$

- For our variables:
  - "N" equals the number of data points
  - "r" is the actual temperature
  - "$\tilde{r}$" is the predicted temperature
  - "$\bar{r}$" is the average temperature determined from the actual temperatures

Results





Numerical Results

Error Estimator	AAE	RMSE	ARPE	APE
Linear	0.7646	4.4406	0.1027	0.0193
Exponential	0.7725	4.4927	0.1026	0.0195
Sine	0.6853	4.3557	0.1021	0.0115
MT (Sine)	0.6639	4.3542	0.1004	0.0170

Conclusions

- **Mean Temperature Sine** is the best among the four chosen models: y=0.85sin(13.25x + 6.42) + 0.11x + 60.93
- Limitations
  - Used fixed month data: March, June, September, December
  - Less accuracy due to using annual data
- Future work:
  - Not using fixed month data
  - Using more frequency data such as monthly, weekly, or daily data
  - Add a seasoning parameter to the models

Acknowledgement : We would like to thank the Choose Ohio First Program at YSU for supporting throughout our entire research project. We would also like to thank our advisors for their help, we could not have accomplished any of this without them.

References:

Actuaries Climate Index At a Glance. (n.d.). Retrieved February 20, 2018.

Alaton, P., Djehiche, B., & Stillberger, D. (n.d.). *On Modelling and Pricing Weather Derivatives*. Retrieved February 20, 2018

Nguyen, N., Nguyen, D., & Wakefield, T. P. (2017, March 7). *Using the Hidden Markov Model to Improve the Hull-White Model for Short Rate*. Retrieved February 20, 2018

Weather History for KYNG. (n.d.). Retrieved February 20, 2018