

# PREDICTING SURFACE TEMPERATURES OF ROADS:

Utilizing a Decaying Average in Forecasting

## Student Authors



**Peter Boyd** is a senior studying applied statistics, actuarial science, and mathematics, and he enjoys conducting research projects in earth, atmospheric, and planetary sciences, mathematics, and forestry and natural resources. Boyd hopes to continue his education by

pursuing a master's degree in environmental statistics after graduation. Boyd's involvement on campus includes Purdue Student Government, Honors College organizations, and Mortar Board.



**Weston Phillips** is a senior majoring in psychological science while minoring in forensic science, law and society, and statistics at Purdue University. Phillips' interests include law, criminal behavior, law enforcement, politics, and various sports. His research

experience includes working with Dr. Michael Baldwin in the Department of Earth, Atmospheric, and Planetary Sciences as well as working with other professors in the Departments of Statistics and Psychological Sciences. Phillips will graduate in spring 2017 and plans to attend graduate school afterward to pursue a master's degree in forensic psychology.

## Mentor



**Michael Baldwin** is an associate professor in the Department of Earth, Atmospheric, and Planetary Sciences at Purdue University. Baldwin came to Purdue after a 15-year career supporting the Environmental Modeling Center (NCEP) as well as the National

Severe Storms Laboratory (NOAA), working to improve operational weather forecasting. Since starting at Purdue in 2006, his research has focused mainly on improving the understanding and prediction of high-impact weather events. Through Baldwin's research, a "feature-specific" analysis and forecast process has been developed that used automated procedures to identify and track specific weather events in time and space. Baldwin has also worked to improve understanding of the effects of global climate change on high-impact weather systems and is recognized as one of the leaders in the field of forecast evaluation. Recently, he has worked on research related to the analysis and prediction of winter road weather conditions.

## Abstract

While being aware of the air temperature during the winter months is very important, many overlook the importance of the road temperature. Knowing the temperature of the road helps transportation departments decide whether or not salt will need to be distributed onto the road, as well as what type of salt. This research was conducted in order to better a forecast model on surface temperature predictions that was shown to be inaccurate. Data was used from the 2013–2014 winter from three cities across Indiana. The data included variables such as air and surface temperatures, precipitation, wind speed, and other variables that could affect the road temperature. These variables were recorded hourly for approximately 5 months. The data was run through both Python and R Studio in order to better visualize and compare the predictions to the observed data. Variables were weighted in different ways to find the variables that contributed most to temperature discrepancies in the forecast. After many tests, the results showed that adding a decaying average over the last 14 days to the predicted temperature yielded the strongest correlation in comparison to the other options available. These results permit additional degrees to be added to our prediction model that will ultimately lead to more accurate predictions, allowing transportation departments to use the predictions to implement in their daily tasks.

Boyd, P., & Phillips, W. (2016). Predicting surface temperatures of roads: Utilizing a decaying average in forecasting. *Journal of Purdue Undergraduate Research*, 6, 9–15. <http://dx.doi.org/10.5703/1288284316160>

## Keywords

time series, road temperature, decaying average, ARIMA, seasonality, autocorrelated errors, forecasting, data analysis

## INTRODUCTION

To prepare for severe weather conditions during the winter months, the Indiana Department of Transportation (INDOT) utilizes weather forecasts and surface temperature predictions for the roads yielded from weather models at Purdue University. These temperature predictions are one of the most vital components of the decision-making process for INDOT. Because snow and ice can lead to hazardous driving conditions, INDOT must be prepared to counteract severe weather with appropriate control methods, such as different salt components.

When the surface temperature of the road falls within a certain range, materials with stronger melting capabilities are preferred to create a safer driving experience. Stronger materials, however, are higher in price, so the materials must be selected carefully to ensure both affordability and effectiveness. If INDOT is able to know, with a strong degree of accuracy, the temperature of the roads, the department will be able to safely monitor the roads in a more cost-effective manner.

The basis of INDOT's weather-related decisions are weather predictions modeled by Dr. Michael Baldwin at Purdue University in West Lafayette, Indiana. The predictions created at Purdue University result from models that are very similar to models used by the National Weather Service. Because these modeling approaches are quite commonplace, the topics discussed in this paper will be useful to a large audience of individuals and organizations. By utilizing the models created by the National Weather Service, one can have access to weather predictions at essentially no cost. Using these models and simply making corrections to eliminate bias is then a cost-effective and productive manner of forecasting.

The model mentioned above does not directly predict the surface temperature of the roads; it uses temperatures of the ground in the areas surrounding the road site. These lands include vegetation, farmland, and soil. Because these ground types have different components from road surfaces, the model has significant room for improvement.

One such option for improving forecasts and diminishing bias is a decaying average bias correction. To improve predicted temperatures, weather forecasters have begun to utilize this decaying average approach by analyzing previous days' errors in the forecast and using their findings to improve upon the forecast for the current day. In this paper, we will analyze the effect of this approach to



improve the predictions for surface temperatures of roads in several Indiana towns.

## DATA

The forecasts in this report were generated by a model created by Dr. Baldwin that follows common practices used by the National Weather Service. The model forecasts many variables, but this paper will focus on its predictions for surface temperatures of Indiana roads in the wintertime, November through March.

The data was collected in six towns, mostly in the central part of the state. Westfield, Frankfort, and Sullivan were selected for use in this study because the data collected at these sites was recorded consistently and accurately. Frankfort and Westfield are 30 miles apart, but Sullivan is 120 miles from Westfield. These distances offer a good spread of data and representation of weather patterns in the state.

The observations were recorded by a device that was located close to the highway near the respective towns. Several variables were recorded, including air temperature, surface temperature, wind speed, and the number of minutes after minute 0 of each hour that the observation was recorded.

The data was contained in spreadsheets that were separated by town, containing the forecasted temperatures and the observations collected for the road temperatures.

The times used in this analysis were in Coordinated Universal Time (UTC). The towns studied in this report were located in the Eastern time zone, which falls 5 hours behind UTC. All temperatures in this report are expressed in degrees Celsius.

## PRELIMINARY ANALYSES

We began our study by analyzing the existing model that was used for the road temperature predictions to learn its forecasting ability and the accuracy of its predictions. Because the data was recorded hourly, all analyses were done on an hourly basis. We first looked at the correlations between the observed and predicted surface temperatures of the three sites. The aforementioned correlations at Westfield, Frankfort, and Sullivan were .866, .886, and .897, respectively (Figure 1). Although these correlations were not poor, by increasing the correlation at all, we could increase our confidence in forecasting surface temperatures.

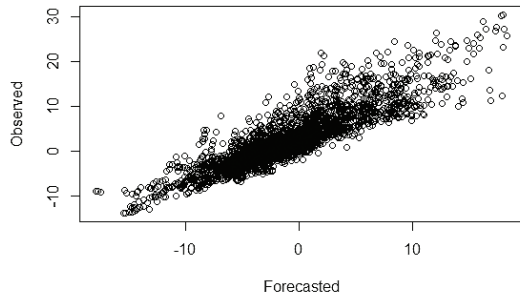
To test the forecasting ability of other parts of the model, we analyzed its accuracy in predicting the air temperature at these three sites. We found that the forecasts for air temperatures have a high degree of accuracy; the median errors for Westfield, Frankfort, and Sullivan are  $-.56$ ,  $-.69$ , and  $-.50$ , respectively. As seen in Figure 2, the correlations between predicted and observed air temperatures are quite strong, all above .930.

As we continued our study, it became apparent that there were several strong relationships among the data. The model seemed to be consistently predicting temperatures that were lower than the observed value. In our analyses, all errors were calculated by subtracting the observed temperature from the forecasted temperature. In the analysis of all surface temperatures, the median errors for Westfield, Frankfort, and Sullivan are  $-2.88$ ,  $-2.56$ , and  $-2.0$ , respectively (Figure 3). The plots display a large spread of errors, but the vast majority are negative, with a surprisingly small number of positive errors. This observation portrays the fact that the model is consistently producing forecasted temperatures that are lower than the true observed surface temperature. The median error in this analysis is  $-2.81$ , suggesting that the model is consistently producing estimates that are too low. We then analyzed the histogram and qq plot of the errors (Figures 4 and 5). The plots suggest that errors are not normally distributed, with a heavy skewness to the left.

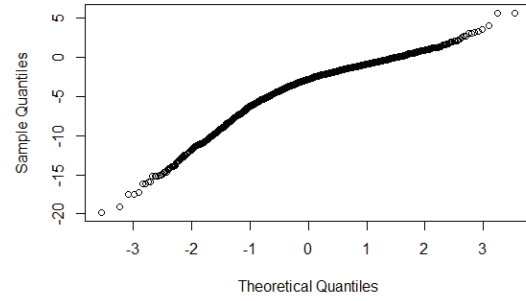
Because the errors are not distributed evenly, or about 0, we realized that the model is not predicting surface temperature in the most effective manner.

In order to proceed, we performed time series analytics to capture the essence of the problem. Figure 6 displays the time series plot of observed temperatures, which shows the temperature decreasing with the onset of winter. We conclude that the expected value of the surface temperature does indeed depend on time.

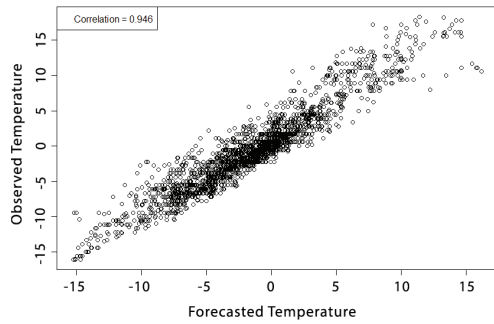
To further investigate how the data is time-dependent, we needed to look into the autocorrelation of the errors in the model. Autocorrelation of errors indicates that observations are not independent, which produces a need for time series analysis. As seen in the previous diagnostic plots, it is quite likely that the errors of the dataset are related sequentially. To further our understanding of this relationship, we plotted the sample autocorrelation function (ACF) and the partial autocorrelation function (PACF) (Figures 7 and 8). The two plots indicate that an



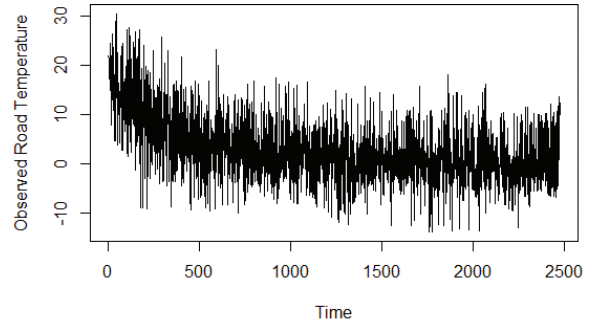
■ **Figure 1.** Forecasted vs. observed surface temperature.



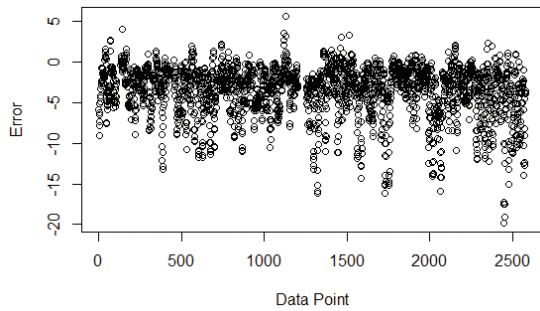
■ **Figure 5.** QQ plot of errors.



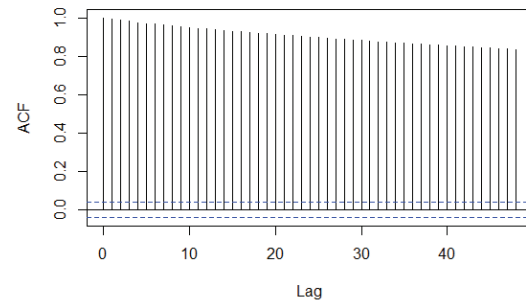
■ **Figure 2.** Westfield air temperature.



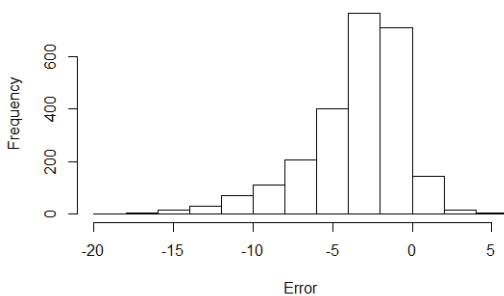
■ **Figure 6.** Time series plot of observed temperatures.



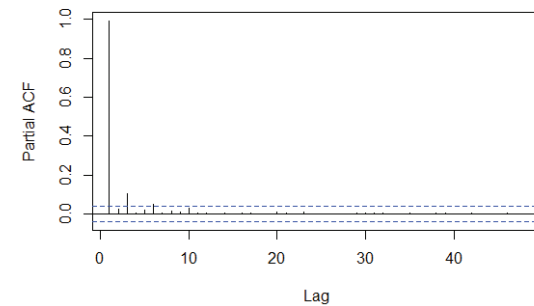
■ **Figure 3.** Westfield error plot.



■ **Figure 7.** ACF of errors.



■ **Figure 4.** Histogram of errors.



■ **Figure 8.** PACF of errors.

autoregressive process with lag 1, or an AR(1) model, may fit the distribution.

However, we have reason to believe that differencing the data may be of value; by differencing each error observation with the same hour of the previous day, all observations become more centralized. The PACF of differenced data is seen in Figure 9.

By looking at the ACF and PACF of the observed temperature, we see that the plot “cuts off,” but peaks at intervals of 24 lags (Figure 10). This realization suggests that the data has a seasonality aspect, meaning that each 24th value is related, or a certain hour in one day is related to the same hour in the previous days.

To understand the way in which these hourly deviations occur, we performed analyses and created plots to clearly show changes on an hourly basis. Plotting the mean error by hour, Figure 11 shows that the mean error for Westfield at hour 0 is roughly  $-4.5$ . The mean error’s magnitude by hour decreases (temperature rises) until hour 15, when it peaks at  $-2.07$ . The plot then shows a quick decrease until hour 21 at  $-6.57$ , which is followed by a steady increase.

In order to continue this analysis, we separated the data by month and replotted the mean error. We see that as the months progress from December (Westfield data started in December) until March, the mean errors become larger compared to previous months. This trend seems to indicate that, as the sun rises in the sky and the road surface is exposed to more heat for a longer duration, the model’s accuracy weakens.

As previously mentioned, the model is fairly accurate in its predictions. However, an increase in forecasting accuracy of any measure is beneficial. Considering the previously mentioned trends, we needed to apply an adjustment to the model to compensate for the error pattern seen by hour and by time of the year. We turned to the decaying average bias correction method to improve the model.

## TIME SERIES ANALYSIS

The idea to analyze the forecast by utilizing a decaying average of errors was based on methods discussed in “Determining an Optimal Decay Factor for Bias-Correcting MOS Temperature and Dewpoint Forecasts” (Glahn, 2014). In our study, we looked at a bias correction method for surface temperatures of the road, rather than dewpoint

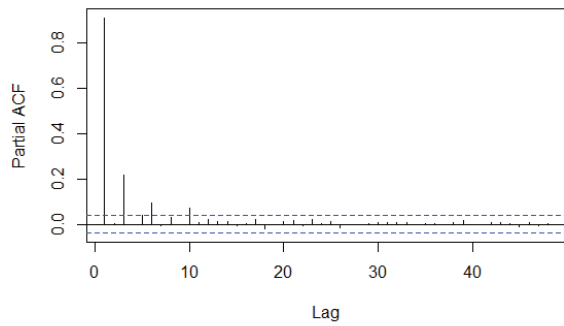
forecasts. A similar approach, however, is utilized in our analysis. As seen in the PACF plot, there are noticeable links between points 24 hours apart. We speculate that by considering the error made on a certain day at a specific hour, we can improve the predictive model by including this error in a future day’s hourly prediction. This improvement can be achieved through an autoregressive integrated moving average (ARIMA) process. This process will allow us to add a series of error terms from previous observations that help correct for consistent mistakes in the forecast.

In order to conduct an analysis involving a decaying average, we calculated the error for each data point, every hour from November (December for Westfield) to March. The next step was averaging the errors recorded on previous days at the same hour. For example, if we were analyzing January 8 at 12:00 p.m. by using a 7-day decay model, we averaged the errors recorded at 12:00 p.m. from January 1–7. We experimented with two variables: number of days to average and weights assigned to the days. We began with a pure average of equal weights for 3-, 4-, 5-, 6-, and 7-day error means. In addition, we utilized a decaying average approach for the same number of days. We then continued the decay analysis with 8, 10, 12, and 14 days. This mean error value was then subtracted from the forecast in order to create a new forecast for the surface temperature. The purpose of this analysis was to determine if the new predicted value is more strongly related to the observed value than to the original forecast.

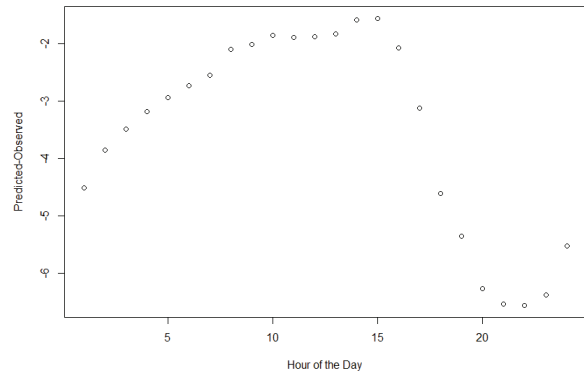
## RESULTS

The bias correction utilizing an average of previous days’ errors proved to be a successful addition to the forecast. Table 1 displays the correlations of observed and predicted temperatures using different lengths of time and weights assigned to the days. As the length of the analysis increases, the correlation generally increases. The tables for Westfield and Sullivan show that the pure average of 14 days’ errors yields the strongest correlation, although it is very similar to the weighted average used in the decay approach. For Frankfort, however, the resultant correlation for the decay method is noticeably stronger than that of the pure average.

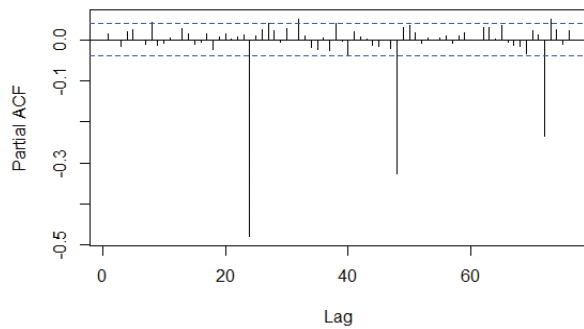
Experimenting with different weight assignments in the average calculation was also an important part of the analysis. The weights were most dependent on the number of days included in the average calculation. For the shorter analyses, the weight assigned to the first day could be as much as 40%. As the duration



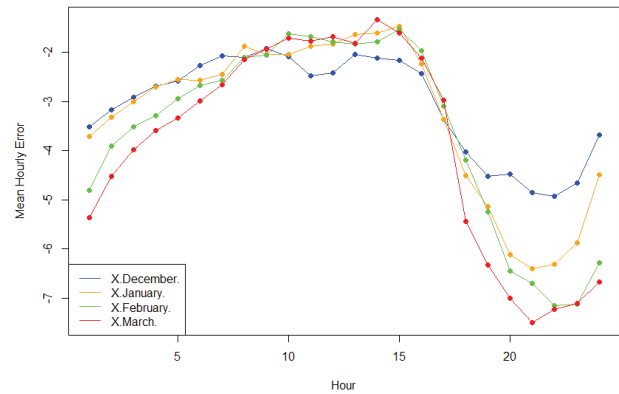
■ **Figure 9.** PACF of 24 lag differenced errors.



■ **Figure 11.** Westfield average error by hour.



■ **Figure 10.** PACF of observations.



■ **Figure 12.** Hourly errors by month (Westfield).

Westfield (.882 initial)	3 Days	4 Days	5 Days	6 Days	7 Days	8 Days	10 Days	12 Days	14 Days
Pure Average	.907	.909	.910	.913	.915	.913	.913	.918	.919
Method 1	.904	.892	.908	.908	.911	.913	.913	.915	.915
Method 2									.917

Frankfort (.886 initial)	3 Days	4 Days	5 Days	6 Days	7 Days	8 Days	10 Days	12 Days	14 Days
Pure Average	.921	.923	.923	.924	.926	.925	.924	.928	.929
Method 1	.920	.910	.923	.923	.925	.940	.927	.941	.942
Method 2									.941

Sullivan (.897 initial)	3 Days	4 Days	5 Days	6 Days	7 Days	8 Days	10 Days	12 Days	14 Days
Pure Average	.923	.924	.927	.929	.929	.927	.927	.928	.929
Method 1	.918	.913	.923	.923	.925	.926	.926	.928	.927
Method 2									.928

■ **Table 1.** The tables above indicate the correlation between forecasted and observed temperatures utilizing different decaying average lengths for 3 different locations.

of the calculation increased, however, the assigned weight was usually closer to 20% for the first day. The weights dropped a few percentage points each day, and depending on the number of days included, the final day in the formula could be 1%.

The addition of the bias correction through decay modeling also significantly lowered the median error for surface temperatures. The median errors for the three sites were lowered to .101 (initially  $-2.88$ ),  $-.034$  (initially  $-2.56$ ), and  $-.044$  (initially  $-2.0$ ). By making the median error closer to 0, we feel confident that the decay model we utilized greatly improves the forecasting ability of the model for surface temperatures.

## CONCLUSIONS

By adding an average of previous days' errors to the predicted surface temperature of the road, the forecast's accuracy is improved, limiting the error between the predicted and observed temperatures. After many trials of differing weights and durations of time, we found that a decaying average of the previous 14 days' errors yields the strongest correlation when added to the forecasted surface temperature.

After examining many different variables for correlations with surface temperature, we found very little indication of an approach that could be used to improve the forecasting ability of the model. We began to utilize the decaying average bias correction method and began to realize an improvement in the model. We experimented by assigning different

weights to the days in the calculation of the average. After finding an improvement in the accuracy of the forecast when adding the bias correction to the model for short periods of time, we lengthened our study, finishing with a 14-day decaying average, which proved to yield the strongest correlation to observed temperatures.

Although we were able to considerably increase the accuracy of the forecast, the capacity to improve always exists. As previously discussed, there was a distinct oscillating pattern of errors recorded by hour as the day progressed. After improving the model by incorporating the 14-day decaying average bias correction, we replotted the mean hourly errors and found that the pattern still existed (Figure 12). This trend can likely be attributed to the result of net radiation and how much sunlight is absorbed by the road, creating additional heat later in the day. Our analysis did not account for this factor, but further research may be able to account for this variable and lead to continued improvements in the forecast for road temperatures.

## ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1246818.

## REFERENCE

- Glahn, B. (2014). Determining an optimal decay factor for bias-correcting MOS temperature and dewpoint forecasts. *American Meteorological Society*. Retrieved from <http://journals.ametsoc.org/doi/abs/10.1175/WAF-D-13-00123.1?journalCode=wefo>