Predicting Non-Working Days Due to Adverse Weather Conditions for Highway Construction Projects in Brazos County, Texas Using 2nd Order Linear Regression

Wonkee Lee, MSCM candidate and Paul Woods, Doctor of Environmental Design, RA and Wei Jue, MSCM and Michael Speed, PhD
Texas A&M University
College Station, Texas

The objective of this research was to develop and evaluate a statistical model to predict the number of non-working days for highway construction projects in Brazos County, Texas, based on the historical weather data and Daily Work Reports. A 2nd order, linear regression model was developed that can be used by contractors to estimate the number of non-working days for any highway construction project in Brazos, County. The variables considered in the model were monthly average precipitation, temperature, wind speed, and number of weekend days and holidays per month. Interactions between the weather variables were also considered in an attempt to improve model effectiveness. With an accuracy of about 95% and an adjusted R-square near 80%, this model provides a relatively easy, valid and powerful method to predict project duration. Additionally, all the required weather data to calculate the number of non-working days are conveniently available on the web.

Key Words: Non-Working Days, Regression, Temperature, Precipitation, Highway Construction

Introduction

Background

Due to the continuous exposure to an outdoor environment, highway construction projects are greatly affected by weather conditions. When adverse weather conditions, which cause non-working days, are not adequately considered in highway project scheduling, delay in the construction project is unavoidable (Kenner, et al., 1998). The delay leads to request for contract time extension. Texas Department of Transportation (TxDOT) specifies contractors are required to decide project completion dates based on their own risks. This delivery method does not allow for contractors to request time extension due to the adverse weather conditions except for some cataclysmic cases (Kenner, et al., 1998). As a result, the number of disputes between the owner and contractor are increasing regarding the reasonable number of non-working days (Woods, et al., 2006). Therefore, an efficient and effective method for calculating the expected number of non-workdays due to the adverse weather conditions is necessary (Xi, Balaji & Molenaar, 2005).

Although several models have been developed to predict the weather related non-working days in a construction project, those methods are not easy to apply and/or are not handy to use for contractors. This study focused on developing an effective and efficient regression model for calculating non-working days that could help in estimating the number of non-working days for any future construction project in Brazos County. Weather data and Daily Work Reports (DWR) for a two-year highway construction project in Brazos County area were used to develop the regression model.

Statement of the problem

The objective of this research was to develop and evaluate a second-order, multiple linear regression model predicting the number of non-working days due to adverse weather conditions for highway construction projects in Brazos County, Texas.

Research Objectives and Study Procedure

The general objectives of this study and study procedure were as follows:

1. To identify weather factors that affect to the number of non-working days:
   1) Review literature related to weather factors and non-working days.
   2) Identify the weather variables to construct a regression model.
   3) Obtain the Local Weather Data from National Climatic Data Center (NCDC).

2. To develop a statistical model to predict the number of non-working days for a highway construction project in Brazos County, Texas:
   1) Analyze the DWR (daily work record) of the highway construction project, and calculate the actual number of weather-related non-working days.
   2) Construct a regression model between the dependent variable, number of non-working days, and independent variables, weather related factors.
   3) Run statistical programs to test hypothesis and to obtain the beta coefficient values.
   4) Analyze the statistical result and finalize the model.

3. To evaluate the effectiveness and efficiency of the statistical model:
   1) Estimate monthly predicted number of working days based on the model developed.
   2) Compare the predicted number of working days with the actual number of working days of the project.

Delimitations / Limitations

1. The model developed in this research was based on the two-year (24 month) sample of a highway construction project in Brazos County, Texas.
2. The data applied in this study were monthly average values, thus monthly average number of non-working days were main concern of the study.
3. The Wind Speed values were the average values of the hourly data obtained from NCDC. Other independent variables, such as precipitation and temperature, were the average values of the daily data obtained from NCDC.

Assumptions

1. The daily work reports obtained through TxDOT were complete and accurate.
2. The Local Weather Data obtained from NCDC used in this research accurately reflects the weather conditions at the construction project site.
3. The sample project was representative of the population.

Review of Literature

Development of Working Day Weather Chart for Transportation Construction in South Dakota

Kenner, et al. (1998) provided a method to estimate the number of non-working days due to adverse weather condition for highway construction projects in the state of South Dakota. Temperature and precipitation were used as the weather variables affecting the number of working days. Through an expert panel review, threshold values were used in determining the adverse weather days. These are; a precipitation threshold of 7.62 mm for a one-day delay, 19.5 mm of rain or more for an additional day delay, and daily maximum temperature threshold of less than 0° C for a day delay. A sample of 54 projects and historical weather data were used to obtain the number of non-working days. Charts were provided for the total number of available working day weather for six geographical locations (zones) and two construction types (grading, surfacing and structural) for the benefit of the contractors. The working day weather charts developed in the study were ready-to-use for contractors of the state of South Dakota. However, the method was applicable to the state only. Also, a validation method comparing the working days obtained from the model to the actual working days was not provided.

Quantify Construction Delays Due to Weather

Xi, Balaji & Molenaar (2005) developed a model to estimate the number of weather related non-working days using synthetic daily weather data, based on different types of weather conditions and construction processes. Three weather variables; precipitation, temperature, and wind speed were used as the weather variables affecting the number of non-working days. The authors emphasized that a mathematical model, such as Markov chain and two-step k-NN weather generator presented in the study, would increase the efficiency of estimating the number of delay days accurately. The model is applicable to other states, but it is a somewhat complex mathematical model, making it difficult for contractors to understand and use.

Predicting Non-Workdays for Highway Projects in Brazos County, Texas from Historical Weather Records and Daily Project Work Reports

Woods, et al. (2006) developed a first-order, linear regression model to predict the number of non-working days in a construction project based on the historical weather data and DWR. The three significant weather factors were precipitation, temperature, and the number of weekend days and holidays per month. The results of the statistical tests showed the effectiveness of the model is reasonably high with an adjusted R-square value of 0.704. When applied to the sample project, the model only produced a 3% deviation from the actual number of project non-work days. The results showed that a first-order, linear regression model can be applied in predicting the number of non-working days. The model developed in this study is efficient because the input data are relatively easy to obtain and its application to the model does not require any complex mathematical algorithm.

Research Methods

Data Collection: Non-working Days

This study used the DWR data analyzed by Woods, et al. (2006). The DWR, recorded for the phase 2 construction of highway FM 158 in Brazos County area, was obtained from TxDOT, (Texas Department of Transportation). The project lasted for 24 months (from August 2002 to July 2004). By examining the DWR closely, the actual number of non-working days and working days were calculated for each month during the 24 month period (Woods, et al., 2006). The actual number of non-working days per month was used as the dependent variable in the multiple linear regression equation.

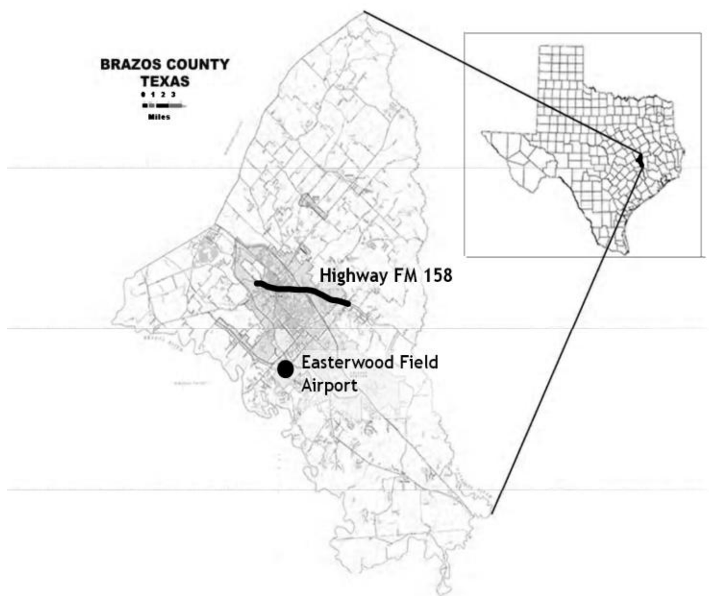

Figure 1: Easterwood Field Airport and FM 158 in Brazos County, TX

Data Collection: Weather Variables

Through the literature review, three weather variables were chosen for the study. Those were; precipitation (inches), temperature (F), and wind speed (mph). Easterwood Field Airport weather station, College Station, TX, was chosen as a source of the weather data. The Easterwood Field Airport was the closest weather station from the project site. All of the local weather data was downloaded from NCDC, National Oceanic and Atmospheric Administration (NOAA). The geographic relationship between the weather station and the project site is shown in Figure 1.

Analysis

Variables Considered

Seven independent variables were considered for the regression analysis. The first independent variable considered was the number of weekend days and holidays per month - 8. The other three were monthly weather factors; precipitation (inches), temperature (F), and wind speed (mph), and the last three were interactions between the weather factors; precipitation*temperature (P*T), precipitation*wind speed (P*WS), and wind speed* temperature (WS*T). The reason for subtracting eight from the number of weekend days and holidays per month was to meet the statistical assumptions for a multiple regression, a simple method of data transformation. It can be thought of logically as accounting for the fact that there is no month that has fewer than a total of eight weekend days and holidays. Interactions between the variables were the product of the two associated variables. As the dependent variable, the actual number of non-working days per month was used. The variables were obtained from the DWR of the sample project.

Model Considered

The following regression equation expresses the initial model that was built to carry out the objectives of the study. This model includes all the variables hypothesized to have a significant influence on the number of non-working days.

Non-Working Days per Month = β0 + β1 (Weekend Days + Holidays – 8) + β2 (Precipitation) + β3 (Temperature) + β4 (Wind Speed) + β5 (Precipitation * Temperature) + β6 (Precipitation * Wind Speed) + β7 (Temperature* Wind Speed) + ε

Selection Method

A number of selection procedures can be applied to select the best regression model. This study used the backward elimination procedure (Ott & Longnecker, 2001). The backward elimination procedure starts with the complete model containing all the independent variables and eliminates the least significant predictor (i.e. eliminates the variable with smallest partial correlation with the dependent variable) and fits a new model with the rest of the predictors. The process repeats until a reasonable regression model is found.

Results

Predictive Statistics

Tables 1, 2, and 3 show the results of the regression analysis. Analysis of variance (ANOVA) showed that the number of weekend days and holidays -8, average monthly precipitation, and the interaction between the average monthly precipitation and temperature (Precipitation * Temperature) were all significant at p-value <= 0.05. This means that all of the three independent variables have significant effects on the number of non-working days. The adjusted R-Square was 0.792. This means that almost 80% of the variability in the variable of interest, number of non-working days, can be explained by the relationship with the three independent variables. The final regression model is;

> ***Non-working days per month*** = 8.361 + 0.824 (Weekend Days + Holidays –8) + 1.974 (Precipitation) - 0.022 (Precipitation * Temperature)

Even though the sample size (n=24) in this study was relatively small, all the statistical diagnostics were positive. The residuals were randomly scattered around zero. This not only indicates a linear relationship between the independent and dependent variables, but also shows that the residuals are relatively evenly distributed.

Because of the small sample size, the degree of freedom less than 50, Shapiro-Wilks test was used to test the normality of the residual. From the Shapiro-Wilks test, since the P-value equals 0.586 (> 0.05), we failed to reject the null hypothesis that the residual is normally distributed.

Since the data is normal, we then could use the Breusch-Pagan test to test if we have a constant variance. Given H_o: $Var(\varepsilon_i) = \sigma_\varepsilon$ for all *i*, the Breusch-Pagan test gave out a P-value equals 0.311 (greater than 0.05). Again, we failed to reject the null hypothesis that we have constant variance.

Table 1

Model Summary

Model	R	R Square	Adjusted R Square	Standard. Error of the Estimate
1	.905(e)	.819	.792	1.25616

Table 2

ANOVA

Model		Sum of Squares	df	Mean Square	F	Significance
1	Regression	142.899	3	47.633	30.187	.000
	Residual	31.559	20	1.578		
	Total	174.458	23			

Table 3

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Significance
		B	Standard. Error	Beta		
1	(Constant)	8.361	.557		15.002	.000
	Weekend Days + Holidays – 8	.824	.182	.449	4.534	.000
	Precipitation	1.974	.310	2.184	6.366	.000
	T*P	-.022	.004	-1.787	-5.192	.000

Note. Excluded variables: Temperature, Wind Speed, T*WS, P*WS.

Validation / Application

The regression model developed in this study can be validated by comparing it with the actual number of working days of a construction project. Table 4 shows the comparison with the predicted number of working days and the actual number of working days. It showed that only 3.5% deviation from the actual number of non-work days of a sample project. Figure 2 shows the differences between the predicted working days and actual working days.

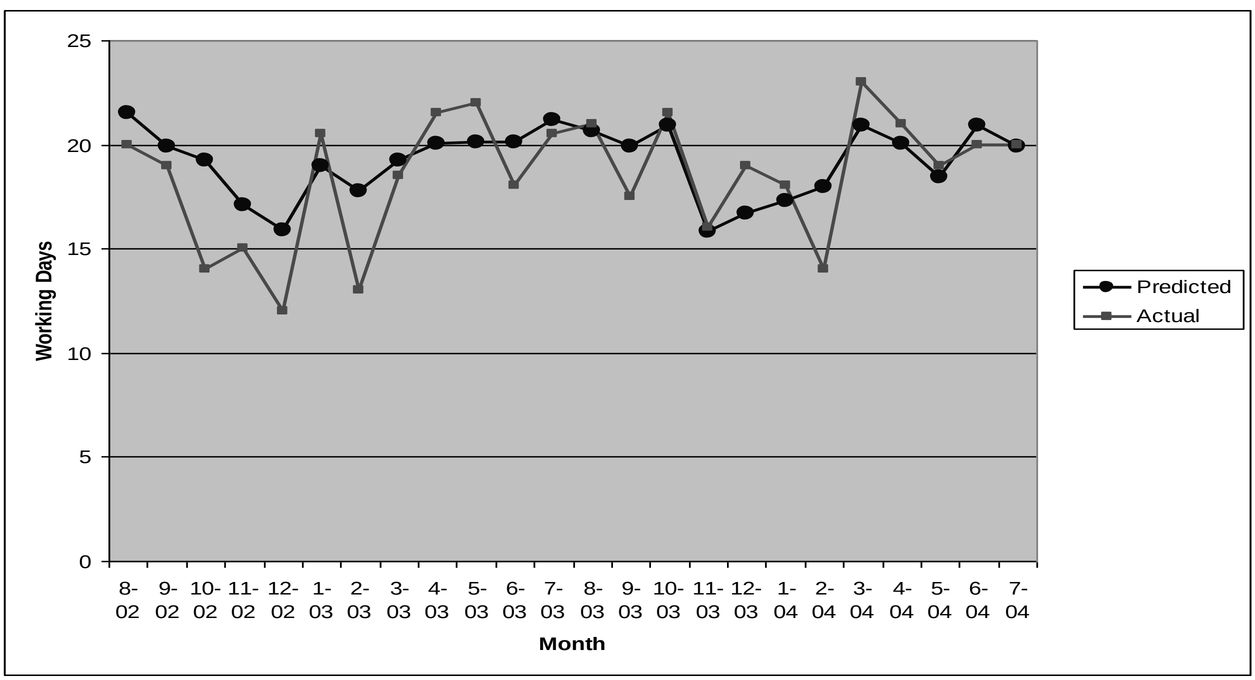

Figure 2: Working Days (Predicted .vs. Actual)

The model is applicable to any future construction project only if long-term weather data is available. Using the regression model developed in this study, ***Non-working days per month*** = 8.361 + 0.824(Weekend Days + Holidays –8) + 1.974(Precipitation) - 0.022(Precipitation * Temperature). Predicted number of working days can then be calculated from the non-working days of the month. For example, for the month of August 2002, non-working days = 8.361 + 0.824 (1) + 1.974 (2.62) - 0.022 (220.55) = 9.51 days. Therefore, the predicted number of working days is; 31 – 9.51 = 21.49 days, as shown in Table 4.

Table 4

Predicted number of working days and the actual working days

			1971 – 2000		Non-Working Days		Working Days		
Month	**Days**	**Weekend & Holidays –8**	**Precipitation (inch)**	**T*P**	**Predicted**	**Actual**	**Predicted**	**Actual**	**Deviation**
8-02	31	1	2.62	220.55	9.51	11.0	21.49	20.0	-1.49
9-02	30	1	3.77	298.10	10.08	11.0	19.92	19.0	-0.92
10-02	31	2	4.01	280.37	11.75	17.0	19.25	14.0	-5.25
11-02	30	3	3.16	187.81	12.93	15.0	17.07	15.0	-2.07
12-02	31	5	3.17	163.52	15.14	19.0	15.86	12.0	-3.86
1-03	31	1	3.24	159.79	12.06	10.5	18.94	20.5	1.56
2-03	28	0	2.33	125.08	10.22	15.0	17.78	13.0	-4.78
3-03	31	2	2.80	171.43	11.77	12.5	19.23	18.5	-0.73
4-03	30	0	3.01	196.82	9.98	8.5	20.02	21.5	1.48
5-03	31	1	5.28	395.41	10.90	9.0	20.10	22.0	1.90
6-03	30	1	3.89	315.86	9.92	12.0	20.08	18.0	-2.08
7-03	31	1.5	2.01	168.50	9.85	10.5	21.15	20.5	-0.65
8-03	31	2	2.62	220.55	10.33	10.0	20.67	21.0	0.33
9-03	30	1	3.77	298.10	10.08	12.5	19.92	17.5	-2.42
10-03	31	0	4.01	280.37	10.11	9.5	20.89	21.5	0.61
11-03	30	4.5	3.16	187.81	14.17	14.0	15.83	16.0	0.17
12-03	31	4	3.17	163.52	14.32	12.0	16.68	19.0	2.32
1-04	31	3	3.24	159.79	13.71	13.0	17.29	18.0	0.71
2-04	29	1	2.33	125.08	11.04	15.0	17.96	14.0	-3.96
3-04	31	0	2.80	171.43	10.12	8.0	20.88	23.0	2.12
4-04	30	0	3.01	196.82	9.98	9.0	20.02	21.0	0.98
5-04	31	3	5.28	395.41	12.55	12.0	18.45	19.0	0.55
6-04	30	0	3.89	315.86	9.09	10.0	20.91	20.0	-0.91
7-04	31	3	2.01	168.50	11.09	11.0	19.91	20.0	0.09
Annual	365.5				135.35	143.50	230.15	222.00	-8.15
					100%	106%	100%	96%	-3.5%

Conclusions

This study provides a second order, linear regression model to predict the number of non-work days in a highway construction project. The results showed that a multiple linear regression model can successfully predict the number of non-working days for a construction project. The three most significant weather factors were; the number of weekend days and holidays per month, precipitation, and temperature. The interaction between the precipitation and the temperature (precipitation* temperature) was considered in order to better explain the variability of the dependent variable. In the regression model, an increased adjusted R-square value was obtained from this research when compared to the first order model developed earlier by Woods et al. This increases the effectiveness of the model in terms of predicting the number of expected non-working days. The model developed in this study is efficient because the input data are relatively easy to obtain and the application of the data to the model does not require any complex mathematical algorithm.

References

Kenner, S., Johnson, L. R., Matt, A. S., Miller, R. J., & Salmen, A. J. (1998). *Development of working day weather chart for transportation construction in South Dakota* [WWW document]. URL http://ntl.bts.gov/lib/9000/9000/9083/SD97-07final.pdf

National Oceanic and Atmospheric Administration (2004). *Climatography of the United States No. 20 1971-2000, College Station Easterwood Airport*. National Climatic Data Center, Asheville, North Carolina, URL http://www.ncdc.noaa.gov/oa/ncdc.html

Ott, L. R., & Longnecker, M. T. (2001). *An Introduction to Statistical Methods and Data Analysis* (5th ed.). Australia ; Pacific Grove, Calif.: Duxbury.

Static GIS Overlay Maps, Brazos County Maps [WWW document]. URL http://www.co.brazos.tx.us/maps/

Woods, P. K., Jue, W., Speed, M., & Burt, R. (2006). *Predicting Non-Workdays for Highway Projects in Brazos County, Texas from Historical Weather Records and Daily Project Work Reports.* (Final Report to Technology Study).

Xi, Y., Balaji, R., & Molenaar, K. (2005). *Quantify construction delays due to weather* [WWW document]. URL http://forecast.colorado.edu/~apipatta/delay/