Resilient Structures and Sustainable Construction
Edited by Pellicer, E., Adam, J. M., Yepes, V., Singh, A., and Yazdani, S.

ISBN: 978-0-9960437-4-8

SPATIAL REGRESSION ANALYSIS FOR MODELING THE SPATIAL VARIATION IN HIGHWAY CONSTRUCTION COSTS

MINSOO BAEK[1] and BAABAK ASHURI[2]

[1]*School of Building Construction, Georgia Institute of Technology, Atlanta, USA*
[2]*School of Building Construction and School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, USA*

Price volatility in wages, materials, and equipment has a significant impact on highway construction costs. As the construction market and economy have experienced dynamic changes in prices, the price volatility becomes less predictable. In addition, various levels of the price volatility in different market locations aggravate the prediction. Thus, in developing highway construction costs, transportation agencies should consider geographical location of construction projects and market conditions of the locations. Transportation agencies face significant uncertainties in price volatility across different geographical locations. This volatility may not be uniformly distributed across different geographical locations due to changes in the availability of local contractors, materials, equipment, and labor. The objective of this research is to develop statistical models that are capable to explain spatial variations in submitted unit prices for asphalt line items in highway projects considering local market condition factors. Historical bid data used in this research consist of resurfacing and widening projects let in the state of Georgia, the United States, between 2008 and 2015. The methodology of this research is a spatial regression analysis to explain the spatial variation in the submitted unit prices for asphalt line items. The findings of this research indicate that volatility in submitted bid prices is not uniformly distributed across different geographical locations within the same transportation agency. The contribution to the body of knowledge of this research is an improved understanding of the role of local construction market and macroeconomic conditions to explain geographic variability in construction costs.

Keywords: Spatial variation, Unit price, Uncertainty, Geographically weighted regression, Spatial analysis.

1 INTRODUCTION

Construction cost data collected from past projects are critical sources in developing cost estimates at different stages of the project's life cycle (AASTO 2011). Moreover, the historical cost data such as the contractors' submitted unit prices for the bid item contain valuable information for preparing bids and establishing risk-sharing strategies for a project. However, the construction costs are subjected to significant variations because of several factors such as the price volatility of materials, the geographical location of a project, market conditions, and the unique characteristics of a project (Ayed 1997). Significant variations in construction costs cause significant uncertainty in developing construction costs. The significant uncertainty impedes the transportation agencies to make accurate decisions for investing, programming, and budgeting construction projects. With the significant amount of uncertainty, the transportation agencies can overestimate construction costs which leads to the cancellation or delays by tying up funds that could be allocated to other projects. Thus, the transportation agencies should analyze the factors affecting variation in construction costs to extract valuable information from historical data in developing construction costs and establishing risk-sharing strategies for highway projects.

Several studies have been conducted to determine the factors affecting construction costs and the relationship between the factors and construction costs using historical data. Shrestha *et al*. (2014) analyzed the bid data of 151 road projects conducted in Clark County, Nevada, to examine the relationship between the unit price bids and the quantity of the unit item. This paper showed that an increase in the quantities lowers the unit price bids of the contract bid items by using a correlation analysis. Basavaraj (2011) also examined a relationship between the asphalt unit prices and the bid quantity. The author showed that as the bid quantity increases, the asphalt unit prices tend to decrease. Moreover, Hegazy and Ayed (1998) identified factors affecting the contract costs by using eighteen bids submitted by construction contractors. This study found that season, location, type of project, contract duration, and contract size significantly impacted on individual contract costs. Ilbeigi *et al*. (2011) used the historical unit prices to identify factors that explain the variation in the bid price. The authors suggested three critical factors, the quantity of the item, total bid price of the projects, and asphalt cement price index at the bid data. Shahandashti and Ashuri (2016) identified the leading indicators that affect the national highway construction cost index using the unit root tests. The authors identified the key leading indicators, crude oil price and average hourly earnings in the construction industry.

A comprehensive literature review confirmed that the construction costs are significantly affected by prices of construction resources, project characteristics, and market conditions. However, there is no research that empirically examines whether the impacts of these factors are uniformly distributed across different geographical locations of projects.

2 RESEARCH OBJECTIVE

The primary objective of this study was to explain spatial variations in the in submitted unit prices for asphalt line items in highway projects considering local market condition factors. To achieve this main objective, the sub objectives of this research were to: (1) identify potential factors that might be influential on the submitted unit prices for asphalt line items; (2) develop an ordinary least squares (OLS) regression model; (3) develop geographically weighted regression (GWR) model; and (4) compare the OLS regression and GWR models.

3 RESEARCH METHODOLOGY

An overview of the research methodology is presented in Fig. 1. In compiling data, we collected the submitted unit prices for asphalt line items. In order to explain spatial variations in the submitted unit prices, this study deployed the geographically weighted regression (GWR). To develop the GWR model, the study used the submitted unit prices by contractors for asphalt line items and several explanatory variables. The major tasks for this study are as follows:

1. Compiling data for analysis: including extraction of data corresponding the research problem, critical evolution of data sources, assessment of data quality, and compilation of data

2. Determining strength of relationship between variables: conducting Pearson Correlation analysis to study the linear relationship between the dependent variable and explanatory variables
3. Conducting exploratory regression: selecting a “best” model among competing models while diagnosing multicollinearity in the model
4. Developing the OLS regression model: with the identified variables from Task 3, estimating global (fixed) variable coefficients for all locations
5. Developing the GWR model: with the identified variables from Task 3, developing local variable coefficients for each location.
6. Finding the best fit model: determining the best model that has higher adjusted R-square and Akaike's Information Criterion (AICc).

With regard to the dependent Variable, this paper used the submitted unit prices for hot mix recycle asphaltic concrete. The submitted unit prices were collected on 1424 highway resurfacing and widening projects let in the entire state of Georgia between 2008 and 2015. Next, explanatory variables were classified into two groups: project related and construction market variables. The project-related variables contained the project characteristics such as the quantity of the bid item, total bid price, and project length. In addition, this study included several construction market variables: number of bidders, number of asphalt plants within 50 miles of the geographic locations of projects, housing market index (HMI), and Georgia asphaltic cement (AC) price index, job opening and labor turnover survey, and quarterly census of employment & wages. The input variables are briefly described in Table 1.

Table 1. Input variables for spatial analysis.

	Input Variables	Description/Source
Project related variables	Natural Logarithm of Quantity of Item	The quantity of asphalt line item (Ton)/ Bid-Tabs database of Oman Systems
	Total Bid Price	Total proposal bid price of the project ($)/Bid-Tabs database of Oman Systems
	Project Length	Miles to be constructed (Miles)/ Online database Bid Express
Construction market variables	Number of Bidders	Number of contractors participating in the bidding process for a project (No.)/ Bid-Tabs database of Oman Systems
	Number of Asphalt Plants Within 50 mi	Number of asphalt plants within 50 miles around a project location (No.)/Georgia DOT
	Housing Market Index	A gauge of the housing builders’ outlook (IDX)/ The National Association of Home Builders (NAHB)
	Georgia Asphalt Cement Price Index	The monthly price of asphalt cement in state of Georgia ($/Ton)/ Georgia DOT
	Job Opening and Labor Turnover Survey	The number of job openings in construction industry (In thousands)/ Bureau of Labor Statistics
	Quarterly Census of Employment & Wages	A count of employment and wages reported by employers (No.)/ Bureau of Labor Statistics

The research methodology includes the two major steps to develop a statistical model that explain spatial variations in the submitted unit prices. The first step is to identify influential factors on the submitted unit prices for asphalt line items. To find the potential factors, literature is reviewed and person correlation analysis is conducted as initial assessment of the potential factors. The second step is to develop a statistical model. Two major statistical techniques, exploratory regression and geographically weighted regression were deployed to develop a statistical model that is capable to explain spatial variation in the submitted unit prices.

4 SPATIAL ANALYSIS

The aim of the GWR analysis is to capture the spatial variation that a simple global model, such the OLS regression model, cannot explain (Brunsdon *et al*. 1996). This study utilized an exploratory regression tool for finding the best model by looking at all possible combination of candidate explanatory variables. The best model was identified based three criteria: (1) adjusted R-square; (2) the variance inflation factors (VIFs); and (3) Akaike's Information Criterion (AICc). The VIFs allow to identify the explanatory variables that are involved in the multicollinearity. The VIFs larger than 10 indicate problems with multicollinearity (Montgomery *et al*. 2015). The AICc allow to measure the performance of between two models. The difference of AICc values between two models more than 2 infer substantial difference in the performance of the two models (Nakaya *et al*. 2005).

Geographically weighted regression (GWR) allows local variations in rates of change so that the coefficients in the model rather than being global estimates are specific to a location. The GWR model is expressed in Eq. (1) (Brunsdon *et al*. 1996).

$$y_i = a_{i0} + \sum_{k=1,m} a_{ik} x_{ik} + \varepsilon_i \qquad (1)$$

Here, a_{ik} is the value of the kth parameter at point/location i. The approach of GWR estimates parameters using a weighting function based on distance so that locations closest to the estimation point have more influence on the estimate (Fotheringham *et al*. 2001).

5 RESULTS

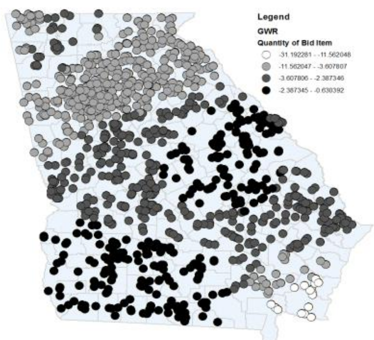
Through the exploratory regression tool, the best model was identified with five significant variables including the quantity of bid items, total bid price, number of asphalt plants within 50 miles, Georgia AC price index, and HMI. With these key variables, the OLS regression and the GWR models are developed. The estimation results for the OLS regression and the GWR models shows in Table 2. Table 2 presents summaries (i.e., minimum, maximum, 1st quartile, median, 3rd quartile) of local variable coefficients of the GWR model.

Table 2. Results of the OLS regression and GWR models.

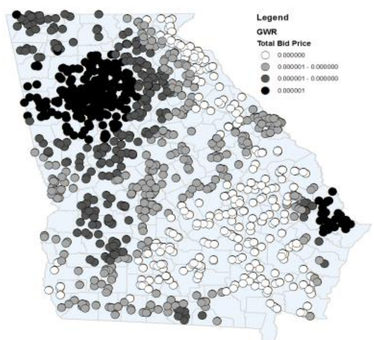
Regression Model	**OLS Regression Model**	**GWR Model**
	Coefficients	
Intercept	75.038	52.949, 259.569 (69.404, 71.921, 78.827)[a]
Natural Logarithm of Quantity of Item	-3.399	-31.1923, -0.630 (-3.833, -3.225, -2.332)
Total Bid Prices	0.00000012	-0.0000004, 0.0000009 (0.0000002, 0.0000003, 0.0000005)
Number of Asphalt Plants Within 50 mi	-0.153	-3.283, 0.361 (-0.412, -0.116, 0.056)
Georgia Asphalt Cement Price Index	0.037	-0.001, 0.086 (0.027, 0.033, 0.038)
Housing Market Index (HMI)	0.231	-0.061, 1.587 (0.129, 0.191, 0.250)
Akaike's Information Criterion (AICc)	10872.984	10100.672
Adjusted R-squared	0.280 (28%)	0.596 (59.6%)

Note: [a] Minimum, maximum (1st quartile, median, 3rd quartile) of the parameter estimates.

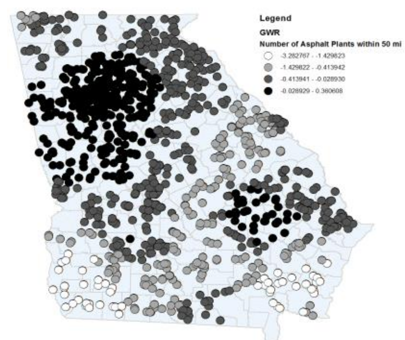
Based on the AICc, as one of criteria for selecting the best model, the difference of the AICc values between the OLS regression and GWR models are 772 which is more than 2. Since the GWR has the lower AICc value than the OLS regression model, the GWR model achieved the greater performance than the OLS regression model. Moreover, while the OLS regression model explained 28% of the variation in the submitted unit prices, the GWR model explained 59.6% of the variation in the submitted unit prices for asphalt line items. As the biggest benefit of GWR, the GWR model provided locally estimates coefficients for identified variables in the model.



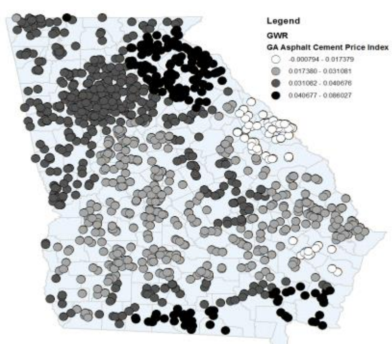
(a) Natural logarithm of quantities.



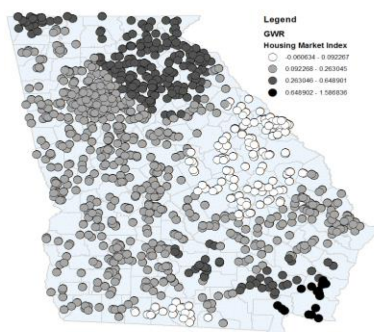
(b) Total bid price.



(c) Number of asphalt plant.



(d) Georgia asphalt cement price index.



(e) Housing market index.

Figure 1. GWR local variable coefficients for five significant variables.

In Figure 1, the local variable estimates are depicted for a better view on the spatial variation of the GWR parameter estimates. The findings of the GWR analysis indicate that:

- The quantity of asphalt line items has a negative relationship in both the OLS regression and the GWR models. The quantity of the item has a stronger negative influence in the southeastern Georgia and the northern Georgia than other areas.
- The total bid price generally shows a positive influence in the GWR model. With the darker shaded points/areas, the total bid price has a stronger influence in the northern Georgia than other areas.
- The number of asphalt plants within 50 miles of project locations generally has a negative relationship with the submitted unit prices. The result of the GWR analysis indicates that as the number of asphalt plants has a stronger negative impact in the southern Georgia, the brighter shaded points, than other areas.

- Georgia AC price index mostly has a positive relationship in both the OLS regression and the GWR models. The result of the GWR analysis indicates that the AC price index has a stronger negative impacts in the northeastern and the southern Georgia.
- HMI shows a positive relationship in both the OLS regression and the GWR models. With the darker shaded points, the HMI has a stronger influence in the northeastern and the southeastern Georgia.

6 CONCLUSIONS

Appropriate variation analysis for the historical bid data is essential for estimating more accurate construction costs and establishing risk-sharing strategies for highway projects. As a simple global model lacks capability to explain spatial variations in the submitted unit prices, this study included GWR analysis to develop a statistical model with the project-related and construction market variables. Compared to an OLS model, the proposed approach provides the higher performance in explaining the spatial variation with key identified factors including the quantity of asphalt line items, total bid price, number of asphalt plants within 50 miles, AC price index, and HMI. The primary result of this study indicates that the key identified factors are not uniformly distributed across different geographical locations. The contribution to the body of knowledge of this research is an improved understanding of the role of local construction market and macroeconomic conditions to explain geographic variability in construction costs.

References

AASTO, *AASHTO Practical Guide To Estimating*, 2011.

Ashuri, B., Shahandashti, S. M., and Lu, J., "Empirical tests for identifying leading indicators of ENR Construction Cost Index." *Construction Management and Economics*, 30(11), 917–927, 2012.

Ayed, A. S., "Parametric Cost Estimating of Highway Projects using Neural Networks", 1997.

Brunsdon, C., Fotheringham, A. S., and Charlton, M. E., "Geographically weighted regression: a method for exploring spatial nonstationarity." *Geographical Analysis*, 281–298, 1996.

Chua, D. K. H., and Li, D., "Key Factors in Bid Reasoning Model." *Journal of Construction Engineering and Management*, 126(5)(October), 349–357, 2000.

Fotheringham, A. S., Charlton, M. E., and Brunsdon, C., "Spatial Variations in School Performance: a Local Analysis Using Geographically Weighted Regression." *Geographical & Environmental Modelling*, 5(1), 43–66, 2001.

Hegazy, T., and Ayed, A., "Neural Network Model for Parametric Cost Estimation of Highway Projects." *Journal of Construction Engineering and Management*, 124(3), 210–218, 1998.

Ilbeigi, M., Ashuri, B., Asce, M., and Shayegh, S., "Price Adjustment Clauses and Submitted Bid Prices for Major Asphalt Line Items in Highway Projects", 2011.

Mekki Basavaraj, V., "Influence of Bid Quantity on Asphalt Unit Prices in The State of Oklahoma." Oklahoma State University, 2011.

Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to linear regression analysis*. John Wiley & Sons, 2015.

Nakaya, T., Fotheringham, A. S., Brunsdon, C., and Charlton, M., "Geographically weighted Poisson regression for disease association mapping." *Statistics in Medicine*, 24(17), 2695–2717, 2005.

Shahandashti, S. M., and Ashuri, B., "Highway Construction Cost Forecasting Using Vector Error Correction Models." *Journal of Management in Engineering*, 32(2), 1–8, 2016.

Shrestha, P. P., Pradhananga, N., and Mani, N., "Correlating the quantity and bid cost of unit price items for public road projects." *KSCE Journal of Civil Engineering*, 18(6), 1590–1598, 2014.

Wilmot, C. G., and Cheng, G., "Estimating Future Highway Construction Costs." *Journal of Construction Engineering and Management*, 129(June), 272–279, 2003.