

Automated Unit Price Visualization Using ArcPy Site Package in ArcGIS

K.J. Shrestha^a and H.D. Jeong^b

^aDepartment of Engineering, Engineering Technology, and Surveying, East Tennessee State University, USA

^bDepartment of Construction Science, Texas A&M University, USA

E-mail: shresthak@etsu.edu, djeong@arch.tamu.edu

Abstract –

State Departments of Transportation (DOTs) in the U.S. have an increasing amount of digital data from various sources. One such set of data is structured unit price data collected from bid lettings. Such data contain unit prices of thousands of bid items from hundreds of projects every year. While state DOTs have such data from over a decade-long period, utilizing such data has been challenging because of the lack of automated analytical and visualization methodologies and tools to generate meaningful and actionable insights. This study develops an automated methodology to quickly and accurately generate color-coded visualization maps representing unit price variation across a geographical region. It uses Inverse Distance Weighted (IDW) technique that is based on the Tobler's First Law of Geography. The law states that points closer together in space are more likely to have a similar value than points that are farther away. The methodology is automated using ArcPy site package in ArcGIS. It imports unit price data from preformatted spreadsheets and boundary maps from existing ArcGIS shape files to generate unit price maps. The tool and the visualizations are expected to aid state DOTs in generating and communicating meaningful insights for making data-driven decisions. It can be used to investigate areas with higher unit prices for various items which can aid state DOTs in identifying potential causes of higher unit prices such as lack of competition and lack of sources of materials (e.g. quarry) in nearby locations.

Keywords –

Geographic-Information-System; Unit-Price-Visualization; Spatial-Analysis; Bid-Price-Estimating; Construction-Market

1 Introduction and Background

Computing accurate cost estimates for highway

construction projects is a never ending challenge faced by state Department of Transportation (DOT) engineers. Estimating accurate construction costs is especially challenging in early phases of the project when a limited amount of information is available. Some of the major factors affecting construction costs and its accuracies include project location, project size, project type, and market inflation. State DOTs collect a large amount of data that can be used to improve the construction estimation accuracies. While project size, type, and market inflation can be represented easily using numerical or categorical variables to account for construction cost variations, accounting for the location factor is more challenging. The spatial information collected by state DOTs coupled with computational tools such as ArcGIS can be used to account for the spatial variation of construction costs. A methodology for such process is described in a previous paper by the authors [1]. Although such methodology has been developed, adaptation of such methodology is likely to be hindered by complications in its implementation. Thus, there is a need to develop a tool to automate such spatial analysis to ease its implementation. This paper discusses a tool that is developed to automatically generate and import relevant data and create a unit price map that can be used as a visual communication tool about the construction market across a geographical region and to generate unit price estimates that account for the spatial variation.

2 Prior Studies

The location of a construction project is one of the most important factor affecting the project's productivity, and costs [2,3]. Those effects are a result of the cost for hauling materials, mobilizing equipment, and other site conditions. Other potential reasons that affect the construction costs across a geographical region include the level of competition in the construction market and local regulations such as minimum wage rates. Highway contractors tend to limit their projects in a certain region to remain competitive.

Similarly, the availability of limited space for construction and state DOTs' desire to complete projects faster in populated areas can significantly increase the construction costs in urban areas. Thus, accounting for the spatial factor is a must for accurate cost estimating. However, current practices have been very limited to partial accounting for such spatial effects on construction costs. For example, Iowa DOT uses an average unit prices from across the state to generate estimates for new construction projects [4]. Using the statewide average unit prices basically ignores the effect of spatial variations in the construction costs. In other cases, experienced engineers tend to account for the location factor based on their experience with previous projects. Several organizations have developed location adjustment factors to address this issue. This includes ENR city indexes and RS Means location factors [5,6]. However, those indexes provide data for a limited number of cities and are not necessarily focused on heavy civil construction projects and hence might be of little use for state DOTs' estimation purposes. A previous study by the authors developed a new methodology to account for spatial variation of construction costs across a geographical region by utilizing historical bid data and ArcGIS [1]. Similarly, another study by Zhang [7] also developed GIS based method to adjust construction cost estimates by project location. Another study utilized spatial interpolation method to understand housing and commercial establishment market [8]. However, while the methodologies are powerful, the implementation of such methodologies is challenging for state DOTs' day to day use because of the lack of automation and complexity in manually utilizing such methodologies. Thus, there is a need to develop a tool to automate the methodology to develop unit price maps of bid items which can be used to communicate and account for the spatial variation of the unit prices across a geographical region.

3 Limitations of Prior Studies

While some studies have attempted to solve the issue, the biggest challenge to communicate and account for spatial variation of construction costs in actual practice by state DOT estimators is the lack of automation. While researchers may be able to utilize advanced data analytics to prove the applicability of innovative techniques, such methods are rarely used in actual practice by state DOTs. State DOT employees have limited time and are not necessarily familiar with statistical and visualization theories and computer applications. This study tackles this challenge by developing an automated script that utilizes Python-based ArcPy Site Package developed by Environmental Systems Research Institute (ESRI) for ArcGIS.

4 ArcGIS and ArcPy Site Package

The ArcGIS is a powerful mapping and analytical platform developed by ESRI [9]. While powerful, it can also be complicated and time consuming to manually conduct spatial analysis using ArcGIS – especially when such analysis includes multiple steps and are to be performed by those who are not very familiar with the software. Thus, ESRI has developed an ArcPy Site Package to automate spatial data analysis, data conversion, data management, and map development [10]. The site package is based on Python programming language that is widely used for statistical analysis, text mining, as well as general purpose programming. This study utilizes the ArcPy Site Package to automate the unit price map development.

5 Framework for Automated Unit Price Visualization

The automated unit price visualization framework can be divided into four components (Figure 1): a) database preparation, b) base map preparation, c) unit price map development, and d) unit price estimation.

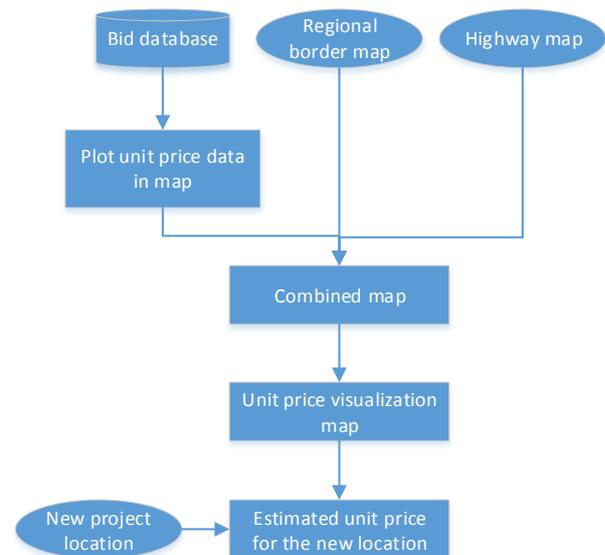


Figure 1. Automated unit price visualization methodology

5.1 Database Preparation

State DOTs have collected digital structured bid data for years. This database consists of historical project information such as the project location and bid data that include the unit prices of various bid items. The project information database can include data such as project ID, contract ID, project description, winning bid

(or contractor), and project location coordinates. The bid database can consist of the project ID that the bid data belongs to, item IDs of all bid items for a given project, and quantities and unit prices of various bid items.

5.2 Base Map Preparation

The geographical maps of the state border lines and highway maps are the two mapping components required for the unit map preparation. The border map becomes a boundary within which the color-coded visualization map will be developed based on the unit prices of items at various locations. The highway maps make it easier to comprehend the visual data points indicating the locations of the projects in respect to the highways.

5.3 Unit Price Map Development

In this step, first, the project data and bid data are combined to generate a list of unit prices of a bid item of interest and coordinates of the project locations where the bid item was used. The data are then plotted in an ArcGIS map to indicate exact project locations and unit prices of the bid item at those locations. Then, the unit price data are interpolated using the Inverse Distance Weighted (IDW) method to generate a map showing the estimated unit prices of the bid item of interest in any location within a boundary map.

5.3.1 Inverse Distance Weighted (IDW) Method

The IDW method is based on the Tobler's First Law of Geography. The law states that points closer together in space are more likely to have a similar value than points that are farther away [11]. Mathematically, the law can be expressed as:

$$p_j = \frac{\sum_{i=0}^n \frac{p_i}{d_{ij}}}{\sum_{i=0}^n \frac{1}{d_{ij}}} \quad (1)$$

In this equation (1), p_j is the estimated unit price of the bid item in a new project (j) location, p_i indicates a unit price of the same bid item from a past project (i), and d_{ij} indicates the distance between the past project location and the new project location. Thus, given the locations of the past and new projects and the unit prices of the item of interest in all the past projects, the unit price of the item for the new project location can be estimated.

5.4 Unit Price Estimation

Once the unit price visualization map is developed, estimators can input the location of a new project and obtain an estimated unit price for the location based on the IDW technique.

6 Results

The researchers collected historical bid data from the Montana Department of Transportation (MDT) in Excel format. The most important datasets relevant to this study were the project information and bid item datasets for all the projects awarded from 2010 to 2014. These datasets were imported in MS Access database program to enable generating customized data quickly for only the bid item of interest for select year or range of years by combining the two datasets. The data generated from the query was exported and saved as a comma separated values. Table 1 shows a sample unit price data generated for the bid item 4020368 (emulsified asphalt CRS-2P). This item is one of the largest items in terms of total dollar value accounting for \$120,350,573 in the projects included in the database and is also one of the most frequently used items.

This data is read by a python script and a new shape file capable of storing location and unit prices of the bid items is programmatically developed. This shape file is combined with a Montana state boundary map and a highway map obtained from MDT's GIS division. This intermediate map is shown in Figure 2.

Table 1. Sample unit price data for item 4020368 (emulsified asphalt CRS-2P)

Latitude	Longitude	Unit Price	Project ID
48.43666667	-104.4969444	654.26	1196
48.02916667	-115.3302778	658	1257
45.88305556	-111.87	463	791
45.88305556	-111.87	463	790
46.32666667	-112.745	510	834
47.48805556	-111.1813889	925	463
46.36472222	-113.3725	465	845
46.36472222	-113.3725	465	844

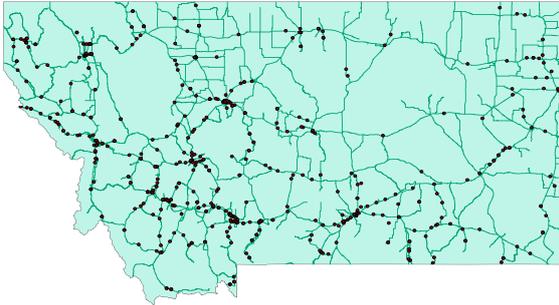


Figure 2. Montana state border map and highway map with project locations.

Finally, a color-coded interpolation map indicating unit prices of the bid items throughout the state is developed as a new shape file and combined with the maps. Figure 3 shows a unit price map generated for the bid item 402020368: emulsified asphalt CRS-2P per ton. The unit prices of the item vary from as low as \$125 to as high as \$3,500. The red color indicates higher unit prices while the blue color indicates lower unit prices.

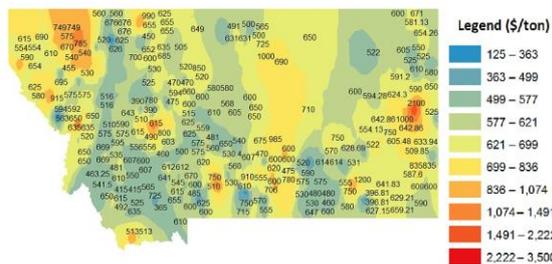


Figure 3. Unit price map for emulsified asphalt CRS-2P per ton

7 Discussion

The automated unit price visualization tool developed in this study enables estimators to easily develop unit price maps of desired bid items. This tool can be used to get insights about the market condition across a geographical region aiding state DOTs to identify areas with higher construction costs. Such insights can direct state DOT estimators to further investigate the reasons behind the higher construction costs and develop strategic plans to potentially reduce construction costs. Further, this highly detailed cost information about the regional construction market enables estimators to determine more accurate construction cost estimates by accounting for the spatial effect that is otherwise challenging to address. This tool also offers a higher granularity than any other location

factors that are currently available to state DOT estimators.

8 Future Research

Currently this tool has two separate components: a) generating the required data from the master database containing all the project level and bid level information and b) unit price map generation based on the data generated from the previous component. While both steps are automated, those steps are isolated and hence they need to be executed separately. A future tool should combine those two components into a single coherent system. This can potentially be achieved by using the database such as ArcSDE where the required data are extracted systematically using programmatically generated dynamic Structured Query Language (SQL) queries. Further, the current Python based tool does not have a Graphical User Interface (GUI). Adding a user friendly GUI will make it even easier for the implementation of the tool for unit price map generation as well as estimation purposes.

Current tool is more focused on automating the visual unit map development. Current Python scripts enable estimators to develop a unit price map of one bid item at a time. This function can be further enhanced to develop a more sophisticated tool that possibly takes all bid items of a new project as input to generate unit price maps for all bid items at once. Thus, such tool can be used to effectively generate the total construction cost estimate of a new project.

Finally, the current tool serves as a powerful visualization tool. However, the tool only accounts for the spatial factor and ignores other factors such as project type. Such factors may be incorporated in the future tool by using pre-filtered data only from the projects of the same type to generate a more precise unit price map.

9 Conclusion

This study developed a tool to automate the unit price map generation process using the bid data collected by state DOTs. This tool enables estimators to utilize their own data for construction market analysis, visually communicating the market information, and develop bid item price estimates that can be used to develop construction cost estimates. The automation and ease of utilizing the tool will aid in implementation of the powerful unit price development methodology that might otherwise be a complicated process for DOT estimators.

10 References

- [1] Jain D, Shrestha J, Jeong D. Spatiotemporal Visualization of Major Cost Items in Highway Construction in Iowa 2014.
- [2] Erickson R, White K. Description of Federal Highway Administration's National Highway Construction Cost Index, 2011.
- [3] Ghosh A, Lynn R. DOD Area Cost Factors (ACF) 2014.
- [4] Iowa Department of Transportation (IADOT). Cost Estimating Database Details 2012.
- [5] Grogan T. What Drives ENR's Cost Indexes. Engineering News-Record 2008.
- [6] RSMeans Engineering Staff. RSMeans Building Construction Cost Data 2016. 74th edition. RS Means; 2015.
- [7] Zhang S. Validation of Geographically Based Surface Interpolation Methods for Adjusting Construction Cost Estimates by Project Location 2014:1–13. doi:10.1061/(ASCE)CO.1943-7862.0000850.
- [8] Montero JM, Larraz B. Interpolation methods for geographical data: Housing and commercial establishment markets. Journal of Real Estate Research 2011;33:233–244.
- [9] Environmental Systems Research Institute (ESRI). About ArcGIS | Mapping & Analytics Platform n.d. <https://www.esri.com/en-us/arcgis/about-arcgis/overview> (accessed December 19, 2018).
- [10] Environmental Systems Research Institute (ESRI). What is ArcPy?—Help | ArcGIS for Desktop n.d. <http://desktop.arcgis.com/en/arcmap/10.3/analyze/arcpy/what-is-arcpy-.htm> (accessed December 13, 2018).
- [11] Eberly S, Swall J, Holland D, Cox B, Baldrige E. Developing Spatially Interpolated Surfaces and Estimating Uncertainty. 2004.