WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

GIS-Based Traffic Volume Estimation Using Spatial Interpolation Techniques: A Review of Methods, Applications, Validation, and Research Gaps

Mr. Pravej Akhtar[1], and Dr. Arun Kumar Patel[1]

Department of Civil Engineering, Vedica Institute of Technology, RKDF University, Bhopal, Madhya Pradesh, India,
Email: arunpatel123@gmail.com, Phone: +91-9407144228

ABSTRACT :

Traffic volume is a fundamental parameter in transportation planning, road design, traffic management, and road-safety assessment. However, traffic counts are generally available at selected monitoring stations, while many road segments remain unmonitored because continuous traffic data collection requires substantial equipment, time, and financial resources. Geographic Information Systems (GIS) provide a spatial framework for integrating traffic counts with road-network characteristics and estimating traffic volumes at locations where direct observations are unavailable. This review examines the application of spatial interpolation techniques for traffic volume estimation, focusing on inverse distance weighting (IDW), ordinary kriging, universal kriging, natural neighbour interpolation, geographically weighted regression (GWR), network-based interpolation, and hybrid geostatistical approaches. Previous studies have demonstrated the potential of these methods for estimating annual average daily traffic and filling gaps in traffic-count datasets. Selby and Kockelman (2013) investigated universal kriging and GWR for predicting traffic levels at unmeasured locations, while Lowry (2014) used network centrality to improve traffic-count estimation. Yang et al. (2018) demonstrated the value of incorporating both spatial and temporal relationships for traffic-volume data imputation. More recent work has explored network regression kriging and machine-learning-based approaches. The review highlights that interpolation accuracy depends on the distribution and representativeness of observed counts, road-network connectivity, temporal consistency, explanatory variables, and validation strategy. Conventional Euclidean-distance interpolation may not adequately represent the structure of a road network. Therefore, network-aware and hybrid methods may be preferable where reliable road-network data and supporting variables are available. The review recommends a GIS-based framework combining traffic-count surveys, road-network attributes, comparative interpolation, independent validation, and uncertainty assessment to produce more reliable traffic-volume estimates for urban road networks.

Keywords: Traffic volume; Geographic Information System; spatial interpolation; inverse distance weighting; kriging; road network; annual average daily traffic; traffic estimation.

1. Introduction

Traffic volume is one of the most important indicators used to assess the operation and demand of a road network. It is commonly expressed as the number of vehicles passing a specified point during a defined period. Depending on the purpose of the study, traffic volume may be reported as hourly traffic, peak-hour traffic, daily traffic, or annual average daily traffic (AADT). Reliable traffic-volume information supports road capacity assessment, pavement and geometric design, intersection planning, public transport development, environmental assessment, and road-safety analysis.

Obtaining comprehensive traffic-volume data across an entire urban road network is challenging. Automatic traffic counters and permanent monitoring stations provide continuous observations at selected locations, while manual traffic counts are often conducted at a limited number of sites and during specific periods. Consequently, traffic volumes may be unavailable for many local roads and secondary links. This limitation can affect transportation models and spatial assessments that require traffic information for the complete road network.

GIS offers a practical environment for integrating observed traffic counts with road-centreline data, functional road classification, number of lanes, speed limits, intersection characteristics, land use, and network connectivity. Spatial interpolation can then be used to estimate traffic volumes at unsurveyed locations based on measured values and their spatial relationships.

Wang and Kockelman (2009) investigated kriging-based methods for estimating AADT at locations without traffic detectors using Texas highway-count data. Selby and Kockelman (2013) subsequently compared universal kriging and geographically weighted regression for predicting traffic levels at unmeasured locations. Their work considered both Euclidean and road-network distances and incorporated road characteristics into the estimation process.

Lowry (2014) proposed a traffic-count interpolation approach based on origin–destination centrality, demonstrating the potential of road-network structure to explain traffic variation. Wu et al. (2022) investigated a hybrid geostatistical approach using network regression kriging for citywide traffic-volume estimation in Edmonton, Canada. These studies illustrate the progression from conventional spatial interpolation toward approaches that explicitly account for road-network structure and explanatory variables.

This review examines the principles, applications, strengths, and limitations of GIS-based traffic-volume interpolation. It also identifies methodological issues concerning spatial dependence, temporal variation, network connectivity, and validation, with the objective of proposing a framework suitable for urban road-network studies.

2. Review Scope and Approach

This paper adopts a narrative review approach to examine research on traffic-volume estimation using GIS and spatial interpolation. The literature covers conventional interpolation methods, geostatistical techniques, network-based estimation, regression approaches, and spatiotemporal imputation of missing traffic observations.

The review addresses five principal questions:

1. How can GIS support traffic-volume estimation at unsurveyed road locations?
2. Which spatial interpolation techniques have been applied to traffic-count estimation?
3. How does road-network connectivity influence interpolation accuracy?
4. How should interpolated traffic volumes be validated?
5. What research gaps remain for urban traffic-volume mapping?

The paper is a narrative review rather than a systematic review. A manuscript claiming systematic coverage should additionally document database sources, search strings, search dates, eligibility criteria, screening stages, and the number of studies included.

3. Role of GIS in Traffic Volume Determination

GIS provides the spatial database and analytical environment needed to integrate traffic-count observations with the road network. Each traffic observation can be associated with a monitoring point, road segment, survey date, observation period, and vehicle classification.

The major applications of GIS in traffic-volume determination include:

3.1 Traffic-count mapping

Observed traffic counts can be displayed as graduated symbols or road-segment attributes. This allows researchers to distinguish high-volume corridors from lower-volume streets and examine the spatial distribution of traffic demand.

3.2 Road-network integration

Traffic-count points can be linked to road-centreline features using spatial matching or network referencing. Road attributes such as functional classification, lane count, speed limit, road length, and intersection connectivity can be added to the dataset.

3.3 Spatial interpolation

Interpolation estimates traffic volume at locations without direct measurements. The resulting estimates can be associated with unsurveyed road segments and visualised as continuous surfaces or network-based traffic maps.

3.4 Traffic data gap filling

Interpolation can be used to estimate missing observations in a traffic dataset. When observations are missing over time as well as space, spatiotemporal methods may provide better estimates than methods that use spatial proximity alone (Yang et al., 2018).

3.5 Transportation planning and road safety

Estimated traffic volumes can support network-level analysis where direct observations are sparse. For example, traffic volume can be integrated with accident records to examine whether high accident counts coincide with high traffic exposure. However, interpolated traffic data should be clearly distinguished from observed traffic counts when used in subsequent analyses.

4. Spatial Interpolation Techniques for Traffic Volume Estimation

4.1 Inverse Distance Weighting

Inverse Distance Weighting (IDW) estimates an unknown value using a weighted average of observations at surrounding locations. Nearby observations generally receive higher weights than distant observations.

The estimated traffic volume at location (s_0) can be expressed as:

where:

$$\bar{u}(x) = \frac{\sum_{j=1}^{n} w_j v_j}{\sum_{j=1}^{n} w_j}$$

Where the weight w_j for each known point is inversely proportional to the distance d_j raised to a power (p):

$$w_j = \frac{1}{d_j^p}$$

Expanding this into single equation

$$\bar{u}(x) = \frac{\sum_{j=1}^{n} \left(\frac{v_i}{d_j^p} \right)}{\sum_{j=1}^{n} \left(\frac{1}{d_j^p} \right)}$$

$\bar{u}(x)$= is the estimated traffic volume at the target location, v_i is the observed traffic volume at location (j); d_j^p = is the distance between the target and observed locations; (p) is the distance-power parameter; • A higher power (p = 2 or more) gives more weight to the closest points, creating a localized, peaky surface. and (n) is the number of neighbouring observations used.

IDW is relatively simple to implement in GIS and does not require the estimation of a variogram. However, it assumes that nearby observations provide useful information about the target location. This assumption may be inappropriate when neighbouring locations lie on disconnected roads or roads with substantially different traffic functions.

For road-network applications, network distance can be used instead of Euclidean distance. The results also depend on the distance-power parameter,

neighbourhood size, and distribution of the traffic-count stations.

4.2 Ordinary Kriging

Ordinary kriging is a geostatistical interpolation method that estimates unknown values using spatial dependence represented by a variogram or semivariogram. It assigns weights to observed values according to their spatial relationships and the fitted spatial model.

A general estimator is: $\bar{z}(s_0)$at an unsampled location s_0 is given by:

$$\bar{z}(s_0) = \sum_{j=1}^{n} \lambda_i Z(s_i)$$

Where $Z(s_i)$ = the known measured value at sample location s_i; λ_i= the kriging weight assigned to sample I ; n = the total number of nearby sample points used in the estimation

Kriging can provide estimates and an associated measure of estimation uncertainty under the assumptions of the fitted model. Its performance depends on the spatial distribution of observations, the adequacy of the variogram model, and the degree to which the assumptions of spatial stationarity are satisfied.

Wang and Kockelman (2009) developed kriging-based methods for estimating AADT at locations without traffic detectors. Their study highlighted the practical value of using available traffic-count data to estimate traffic volumes across a wider highway network, while also reporting difficulties associated with low-volume roads and extreme observations.

Kriging should not automatically be assumed to outperform simpler methods. Its performance should be assessed through independent validation using observed traffic counts withheld from model fitting.

4.3 Universal Kriging

Universal kriging extends ordinary kriging by allowing a spatially varying trend to be modelled using explanatory variables. In traffic-volume applications, the trend may be related to road classification, number of lanes, speed limit, or other characteristics.

Selby and Kockelman (2013) compared universal kriging and geographically weighted regression for estimating average daily traffic at unmeasured locations across the Texas road network. Their study considered both Euclidean and network distances and incorporated roadway characteristics. The authors reported that the spatial methods reduced errors compared with non-spatial regression approaches, although estimation errors remained substantial at some locations, particularly where traffic counts were sparse.

Universal kriging may be appropriate when traffic volumes vary systematically with road characteristics in addition to exhibiting spatial dependence. Its use requires suitable explanatory variables and a defensible model of the remaining spatial structure.

4.4 Natural Neighbour Interpolation

Natural neighbour interpolation estimates values from nearby observations using the spatial geometry of the surrounding points. It is often used to create continuous surfaces from irregularly distributed observations.

In traffic-volume applications, natural neighbour interpolation may be useful for exploratory mapping where traffic-count points are distributed across an urban area. However, the method is based on spatial proximity and does not inherently represent traffic movement along connected road segments.

Consequently, a continuous natural-neighbour surface should not automatically be interpreted as traffic volume on every road. A network-aware procedure is needed if the final output is intended to represent estimated traffic counts on specific road links.

4.5 Geographically Weighted Regression

Geographically Weighted Regression (GWR) estimates spatially varying relationships between a response variable and explanatory variables. Unlike conventional interpolation, GWR models how traffic volume relates to road characteristics at different locations.

Potential explanatory variables include road class, number of lanes, speed limit, intersection density, network accessibility, and surrounding land use, where these data are available and appropriately measured.

Selby and Kockelman (2013) compared GWR with universal kriging for traffic-level prediction at unmeasured locations. Their findings indicated that the spatial methods could improve upon non-spatial regression, while the comparative performance depended on the data and validation setting.

GWR is not a pure interpolation technique. It is a local regression model that can be used to estimate traffic volumes at unsurveyed locations when suitable predictor variables are available. Its reliability depends on bandwidth selection, the distribution of observations, and the stability of local regression estimates.

4.6 Network-Based Interpolation

Conventional spatial interpolation typically uses straight-line distance between observations. In road networks, however, traffic movement follows connected road segments rather than unrestricted geographic space. Network-based interpolation incorporates the structure of the road system and may use road-network distance or connectivity-based measures.

Shiode (2011) investigated network-based extensions of IDW and ordinary kriging for street-level spatial interpolation. The study is relevant to traffic-volume research because it demonstrates how interpolation can be adapted to network-constrained data.

Lowry (2014) proposed an approach for estimating AADT using origin–destination centrality, a measure of the structural importance of a road link within a network. Rather than relying exclusively on proximity to observed counts, the method uses network structure to explain traffic variation. The study reported an out-of-sample (R^2) of 0.95 in its case study; this result is specific to the study dataset and should not be interpreted as a general performance guarantee.

Network-based interpolation is particularly relevant where roads are separated by barriers, where parallel routes carry different traffic levels, or where the target output is an estimate for individual road segments.

4.7 Hybrid Geostatistical Methods

Hybrid approaches combine spatial interpolation with regression or network information. Network regression kriging, for example, models systematic variation using explanatory variables and then estimates the remaining spatially structured component.

Wu et al. (2022) developed a hybrid geostatistical method for estimating citywide traffic volumes in Edmonton, Canada. Their network regression kriging approach incorporated variables such as road type, speed limit, and network accessibility. The study reported improved accuracy compared with standard kriging in the evaluated network, with benefits from treating road types separately where the stationarity assumption did not hold across the entire network.

Hybrid approaches may be useful when traffic volume depends on both network structure and local road characteristics. However, the additional complexity requires adequate training data, appropriate model assumptions, and independent validation.

5. Comparison of Interpolation Methods

Method	Main principle	Strength	Main limitation
IDW	Distance-weighted average of observed counts	Simple and easy to implement	Does not model spatial dependence explicitly
Ordinary kriging	Spatial dependence represented through a	Provides estimates and kriging	Sensitive to variogram assumptions

	variogram	uncertainty	and data distribution
Universal kriging	Kriging with a spatial trend or explanatory variables	Accounts for trend and spatial residuals	Requires suitable covariates and model specification
Natural neighbour	Uses local point geometry	Useful for exploratory surface mapping	Does not inherently respect road connectivity
GWR	Locally varying regression relationships	Models spatial variation in predictor effects	Sensitive to bandwidth and local sample size
Network IDW/kriging	Uses road-network distances or topology	Better suited to network-constrained data	Requires a reliable, connected road network
Network regression kriging	Combines regression, network information, and spatial residuals	Integrates road characteristics and spatial dependence	More demanding data and modelling requirements
Spatiotemporal kriging	Uses spatial and temporal relationships	Can fill missing traffic observations across time and space	Requires consistent timestamps and temporal data

The literature indicates that there is no universally optimal interpolation method. The choice depends on whether the target is a continuous geographic surface, an estimate for an individual road link, a missing sensor observation, or AADT at an unsurveyed location.

6. Temporal Considerations in Traffic Volume Estimation

Traffic volume varies by time of day, day of the week, season, and year. Therefore, spatial interpolation should be conducted using traffic observations that represent comparable time periods.

For example, an hourly count measured during the morning peak should not be directly interpolated with an annual average daily traffic value unless the data have been converted to a consistent basis. Similarly, counts recorded on weekdays and weekends may represent different traffic patterns.

Yang et al. (2018) proposed a kriging-based spatiotemporal method for imputing missing traffic-volume observations. Their approach incorporated both spatial relationships between traffic monitoring locations and temporal relationships between observations. The authors compared the method with historical-average and K-nearest-neighbour approaches under different missing-data scenarios.

Their study highlights the importance of temporal information when the objective is to fill missing sensor data. However, spatiotemporal imputation of missing observations at monitored sites is not identical to spatial estimation of long-term AADT on an unmonitored road segment. The target variable and intended application should be clearly defined before selecting the method.

For urban traffic studies, a suitable workflow should distinguish between:

- Short-duration observed traffic counts;
- Peak-hour traffic volume;
- Average daily traffic;
- Annual average daily traffic; and
- Time-specific traffic observations from sensors.

Where short-duration surveys are used to estimate AADT, appropriate expansion factors and procedures should be documented. Spatial interpolation should not be used as a substitute for temporal adjustment.

7. Validation of Interpolated Traffic Volumes

Validation is essential because a visually smooth traffic map does not demonstrate that the estimated values are accurate.

7.1 Independent validation data

Observed traffic counts should be divided into training and validation sets. The interpolation method is fitted using the training observations, and its estimates are compared with actual counts at locations withheld from model fitting.

Where observations are spatially clustered, random splitting may overestimate performance because training and validation sites may be close to each other. Spatially separated validation or road-segment-based holdout procedures may provide a more demanding assessment of performance at unobserved locations.

7.2 Evaluation metrics

Commonly used metrics include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_i'|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_i'|^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}|Y_i - Y_i'|^2}{\sum_{i=1}^{n}|Y_i - Y'|^2}$$

MAE represents average absolute error, while RMSE penalises larger errors more strongly. (R^2) measures the degree of agreement in variation between observed and estimated values, but it should not be used alone to evaluate interpolation performance.

Traffic-count datasets often contain both low-volume roads and a small number of very-high-volume links. Researchers should therefore report errors by road class or traffic-volume range where possible, because aggregate metrics may conceal poor performance for particular groups.

7.3 Model comparison

A fair comparison requires the same validation sites, consistent target units, and comparable input information across methods. Results should include a simple baseline, such as nearest-neighbour estimation or a road-class average, alongside more complex approaches.

Kriging variance or other model-derived uncertainty measures can also be reported where supported by the selected method. However, model-based uncertainty should not be confused with a guarantee that the actual prediction error will fall within a specified range.

8. Applications in Road-Safety Assessment

Interpolated traffic volume can support road-safety analysis by providing estimates for road segments where direct traffic counts are unavailable. Accident counts and traffic volumes can then be examined together to better understand spatial patterns.

For example, a road segment with a high accident count may also carry a high volume of traffic. Without exposure information, it is difficult to determine whether the number of crashes is unusually high relative to the amount of traffic using the road.

Interpolated traffic volumes can therefore support exploratory analysis of accident hotspots, traffic exposure, and road-network prioritisation. Nevertheless, traffic-volume estimates introduce additional uncertainty. A safety model should distinguish observed from interpolated traffic counts and, where feasible, evaluate whether conclusions remain stable under alternative interpolation methods.

For a GIS–AHP road-sign requirement study, traffic volume can be one criterion in a composite index. However, estimated traffic volume should be standardised consistently with the other criteria, and its uncertainty should be considered. A high traffic-volume estimate alone does not prove that a road segment requires more signs; the assessment should also consider road geometry, junction complexity, crash history, sign condition, visibility, and applicable road-sign standards.

9. Research Gaps and Future Directions

The reviewed studies indicate several opportunities for further research.

First, network-aware interpolation needs wider comparative evaluation. Conventional IDW and kriging should be compared with network-based alternatives using the same road network and validation observations.

Second, traffic volume should be modelled with relevant road characteristics. Road classification, number of lanes, speed limit, network accessibility, intersection characteristics, and land use may improve estimation where these variables are available and reliable.

Third, temporal consistency needs greater attention. Researchers should distinguish short-duration counts from daily and annual traffic measures and avoid combining incompatible observation periods.

Fourth, validation should reflect spatial transferability. Independent spatial holdouts can provide a more realistic assessment of performance on road segments far from existing monitoring stations.

Fifth, uncertainty should be communicated explicitly. Traffic estimates should be accompanied by error metrics, validation results, and, where supported, prediction uncertainty. This is important when estimated traffic volumes are used in road-safety or infrastructure-prioritisation decisions.

Sixth, more urban case studies are needed in data-scarce settings. Many cities have incomplete traffic-count coverage, making GIS-based interpolation attractive. However, models developed in one region should be locally tested before being transferred to cities with different network structures, traffic patterns, and road classifications.

10. Proposed GIS-Based Framework

A practical framework for estimating traffic volume on an urban road network can be organised into the following stages:

1. Collect traffic counts at representative road locations and document survey dates, duration, vehicle classes, and counting procedures.
2. Prepare a topologically consistent road network and assign each count station to the appropriate road segment.
3. Standardise the target variable, distinguishing hourly traffic, peak-hour traffic, daily traffic, and AADT.
4. Integrate relevant road attributes, such as road class, lane count, speed limit, intersection connectivity, and network accessibility.
5. Implement baseline and alternative methods, including IDW, ordinary kriging, and a suitable regression or network-aware method.
6. Select interpolation parameters using training data and appropriate diagnostics.
7. Validate the models using independent observations and consistent error metrics.
8. Compare model performance across road classes, traffic-volume ranges, and spatially separated validation locations.
9. Generate a GIS road-segment layer containing estimated traffic volume, method used, validation information, and uncertainty where available.
10. Integrate the estimates into transportation planning or road-safety analysis while clearly distinguishing observed and estimated values.

This workflow provides a reproducible structure for traffic-volume estimation and supports transparent interpretation of the resulting traffic maps.

11. Conclusion

GIS-based spatial interpolation offers a practical way to estimate traffic volumes at locations where direct traffic counts are unavailable. IDW and natural neighbour methods provide relatively simple alternatives, while kriging and regression-based methods can incorporate spatial dependence and road characteristics. Network-based interpolation is particularly relevant when the intended output is traffic volume on individual road links, because road connectivity and network structure may not be adequately represented by straight-line distance.

Previous research by Wang and Kockelman (2009), Selby and Kockelman (2013), Lowry (2014), Yang et al. (2018), and Wu et al. (2022) demonstrates the range of approaches available for estimating traffic levels and imputing missing observations. Their findings also highlight the importance of data

coverage, explanatory variables, temporal consistency, and validation.

No single interpolation method is best for every road network. A reliable GIS workflow should compare suitable methods using independent traffic counts, report estimation errors, consider road-network structure, and distinguish measured values from modelled estimates. These practices are essential when interpolated traffic volume is used for road design, traffic management, accident analysis, or road-sign requirement prioritisation.

REFERENCES

- Katayama, H., Yasuda, S., & Fuse, T. (2022). Comparative validation of spatial interpolation methods for traffic density for data-driven travel-time prediction. *International Journal of Intelligent Transportation Systems Research, 20*, 830–837. https://doi.org/10.1007/s13177-022-00326-0
- Lowry, M. (2014). Spatial interpolation of traffic counts based on origin–destination centrality. *Journal of Transport Geography, 36*, 98–105. https://doi.org/10.1016/j.jtrangeo.2014.03.007
- Selby, B., & Kockelman, K. M. (2013). Spatial prediction of traffic levels in unmeasured locations: Applications of universal kriging and geographically weighted regression. *Journal of Transport Geography, 29*, 24–32. https://doi.org/10.1016/j.jtrangeo.2012.12.009
- Shamo, B., Asa, E., & Membah, J. (2015). Linear spatial interpolation and analysis of annual average daily traffic data. *Journal of Computing in Civil Engineering, 29*(1). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000281
- Shiode, N. (2011). Street-level spatial interpolation using network-based IDW and ordinary kriging. *Transactions in GIS, 15*(4), 457–477. https://doi.org/10.1111/j.1467-9671.2011.01278.x
- Wang, X., & Kockelman, K. M. (2009). Forecasting network data: Spatial interpolation of traffic counts from Texas data. *Transportation Research Record: Journal of the Transportation Research Board, 2105*(1). https://doi.org/10.3141/2105-13
- Wu, M., Kwon, T. J., & El-Basyouny, K. (2022). A hybrid geostatistical method for estimating citywide traffic volumes—A case study of Edmonton, Canada. *Journal of Geographical Research, 5*(2). https://doi.org/10.30564/jgr.v5i2.4513
- Yang, H., Yang, J., Han, L. D., Liu, X., Pu, L., Chin, S.-M., & Hwang, H. (2018). A kriging based spatiotemporal approach for traffic volume data imputation. *PLOS ONE, 13*(4), e0195957. https://doi.org/10.1371/journal.pone.0195957