Use of Kriging Technique to Study Roundabout Performance

Alessandro Mazzella, Claudia Piras, and Francesco Pinna

Road intersections are dangerous places because of the many conflicting points between motorized and nonmotorized vehicles. In the case of defined traffic volume, several research groups have proved that roundabouts reduced the number of injuries and fatal accident cases. In recent years, many countries have adopted roundabouts as a standard design solution for both urban and rural roads. Several recent studies have investigated the performance of roundabouts, including some with models that calculated the entering flow (Q_e) as a function of the circulating flow (Q_c). Most existing models have been constructed with the use of linear or exponential statistical regression. The interpolative techniques in classical statistics are based on the use of canonical forms (linear or polynomial) that completely ignore the correlation law between collected data. As such, the determined interpolation stems from the assumption that the data represent a random sample. In the research reported in this paper, a geostatistical approach was considered: the relationship Q_e versus Q_c is supposed to be a regionalized phenomenon. According to that supposition, collected data do not represent a random sample of values but are supposed to be related to each other with a defined law. This recognition allows the realization of interpolation on the basis of the real law of the phenomenon. This paper discusses the fundamental theories, the applied operating procedures, and the first results obtained in modeling the Q_e versus Q_c relationship with the application of geostatistics.

The number of roundabouts on roadways has been increasing. Italy, for example, has seen a significant increase in the use of roundabouts (as have many other European countries) in recent years. This increase in roundabouts is attributable mainly to their exceptional performance. If properly designed, roundabouts provide more safety under certain traffic flow conditions and roadway geometries than do other, conventional intersections. For that reason, many intersections have been converted into roundabouts, most of which consist of a single lane or two. Roundabouts appear in both rural and urban areas.

Sufficient literature is available to calculate the performance of roundabouts, and a substantial number of models have been built. Many of the existing models allow the entering flow value (Q_e) to be obtained as a function of the circulating flow value (Q_c) (*1–6*). Each of these models expresses the relationship between Q_e and Q_c with a specific trend (i.e., linear, quadratic, or exponential), which is the same for every Q_c value within the range of 0 ÷ 1,600 [pcu/h].

A. Mazzella, Department of Geoengineering and Environmental Technologies, and C. Piras and F. Pinna, Department of Land Engineering, University of Cagliari, Piazza d'Armi 19-09123, Cagliari, Italy. Corresponding author: C. Piras, cpiras@unica.it.

Transportation Research Record: Journal of the Transportation Research Board, No. 2241, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 78–86.
DOI: 10.3141/2241-09

This paper introduces the results of research to develop new models to predict Q_e versus Q_c on standard and nonstandard legs with the use of geostatistics.

In this research a geostatistical approach was considered: the relationship Q_e versus Q_c is supposed as a regionalized phenomenon (*7*). The strength of this approach is that the collected data do not represent a sample of random values. Rather, they are a punctual manifestation of a phenomenon regulated by a specific law (expressed by the correlation between collected data). This recognition allows the realization of interpolation on the basis of the real law of the phenomenon (*8*). First, available Q_c values are identified and, once the correlation law between the available data has been defined, the interpolation technique allows each unknown Q_e value to be predicted by considering only the experimental values that exercise a real influence.

This paper is divided into four parts. The first part briefly summarizes some of the international models with Q_e versus Q_c. The second part describes the data collected to calibrate the proposed model. The third part describes the development of the proposed predictive model obtained by geostatistics. The final part summarizes the main conclusions of the study and briefly discusses its future development.

EXISTING MODELS

Several models in the literature calculate the relationship between Q_e and Q_c for conventional roundabouts. Some of them are considered in this paper for purposes of comparison with the results reported here.

NCHRP Report 572 (*3*) proposed an exponential regression model, which is actually a gap-acceptance model, for single-lane and two-lane roundabouts. Entry capacity at single-lane roundabouts is expressed by the following:

$$Q_e = 1130 \cdot e^{(-0.001 \cdot Q_c)} \tag{1}$$

Entry capacity at two-lane roundabouts is shown by the following formula (*1*):

$$Q_e = 1130 \cdot e^{(-0.0007 \cdot Q_c)} \tag{2}$$

where Q_e is the entry capacity on every leg in passenger car units/h (pcu/h) and Q_c is the conflicting flow (pcu/h).

A procedure in the *Highway Capacity Manual* indicates that the capacity of a roundabout can be estimated by using gap acceptance techniques with the basic parameters of the critical gap and follow-up time. The procedure underlines the general assumption that the performance of each leg of a roundabout can be independently analyzed from the other legs; consequently, most techniques tend to use the information from one leg only. In the *Highway Capacity*

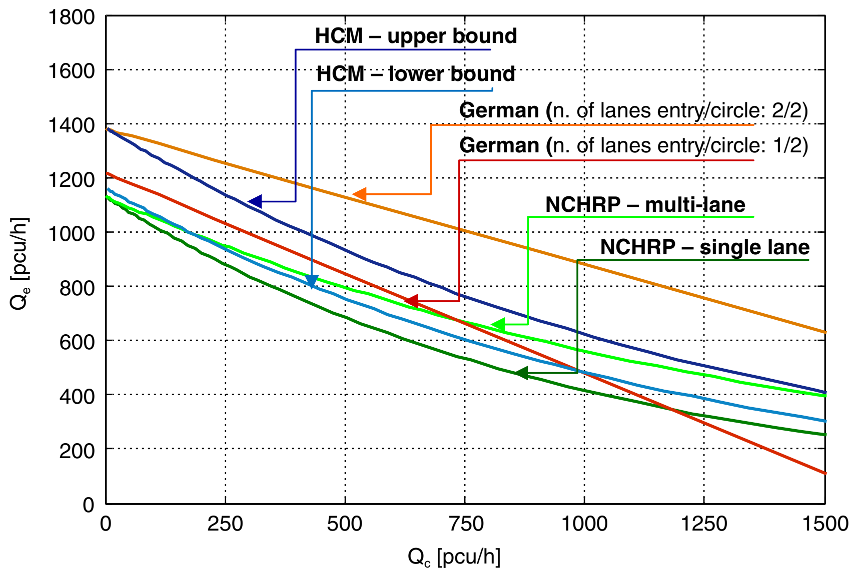


FIGURE 1 Comparison of existing roundabout models.

Manual (*2*), the approach capacity is estimated by using the following formula:

$$Q_e = \frac{Q_c e^{-Q_c t_c / 3{,}600}}{1 - Q_c e^{-Q_c t_f / 3{,}600}} \quad (3)$$

where

Q_e = approach capacity (pcu/h),
Q_c = conflicting flow (pcu/h),
t_c = critical gap (s), and
t_f = follow-up time (s).

The recommended ranges of values of the critical gap and follow-up time are as follows:

- Upper bound: critical gap = 4.1 s and follow-up time = 2.6 s and
- Lower bound: critical gap = 4.6 s and follow-up time = 3.1 s.

One of the German methodologies that calculates Q_e as a function of Q_c is expressed by a linear equation (*3*), which is

$$Q_e = C + DQ_c \quad (4)$$

where Q_e is the entry flow capacity (pcu/h) and Q_c is the circulating flow rate (pcu/h).

C and *D* depend on the number of entry and circle lanes. For this study, it was assumed that

- $C = 1{,}218$ and $D = -0.74$ (number of lanes entry/circle → 1/2) and
- $C = 1{,}380$ and $D = -0.50$ (number of lanes entry/circle → 2/2).

Similar models exist elsewhere [e.g., in Switzerland, France, and the United Kingdom (*4–6*)]. Some of the models, such as the British models, exhibit a trend that is much higher than other models and which is outside the data set range explored in the research described in this paper. To keep to a reasonable length, this paper does not describe such models further.

All of the above-mentioned models, which are built by a statistical regression among locally collected data (Figure 1), express the Q_e as a function of Q_c. The six models are quite similar for low Q_c values (the minor differences can probably be attributed to different driver behavior in different countries), in contrast to high Q_c rates for which the differences are more evident.

DATA COLLECTION

For the study reported here, an extensive data collection (which amassed more than 20,000 observations) was carried out to acquire data on the variation of the circulating flow with entering flow both in standard and nonstandard legs. Nonstandard legs were defined as those without deflection in the entries section (Figure 2).

Traffic volume data were collected for each leg of several roundabouts. Only urban roundabouts (in Cagliari, Italy) were taken into account. Generally, the chosen roundabouts had the following geometric features:

- Outer diameter between 26 m and 40 m,
- One single lane or two lanes at each of the entries,
- One single lane or two lanes at each of the exits, and
- Circle lane width of more than 7 m (with one or two marked lanes).

The number of entering vehicles and circulating vehicles was collected every 5 min during rush hour (7:30–9:30 a.m.). To take

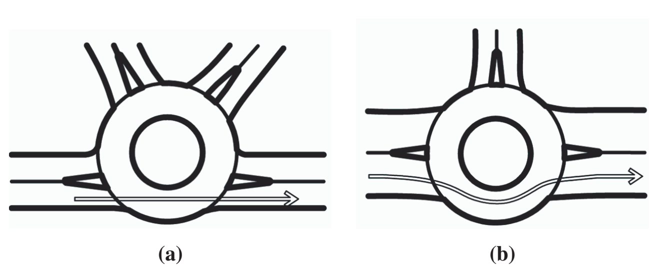


FIGURE 2 Examples of a nonstandard and a standard leg (*9*).

into account the influence of heavy vehicles, a procedure was adopted to estimate flow, which consisted of converting all traffic flows (for each time gap) into pcus.

STATISTICAL TREATMENT OF DATA

Data analysis and statistical regression (both linear and exponential) were performed to determine whether the data used for the geostatistical modelization were comparable to those of the existing international models. Figure 3 shows the relationship Q_e versus Q_c for standard legs. The exponential model to express that relationship is

$$Q_e = 1{,}366.49 \cdot e^{(-0.001 \cdot Q_c)} \quad (5)$$

The collected data were fitted by using a linear regression, too, and the equation is

$$Q_e = 1{,}212.19 - 0.73 \cdot Q_c \quad (6)$$

where Q_e is the entry capacity (pcu/h) and Q_c is the conflicting flow (pcu/h).

Figure 3 shows the trends of linear (light blue dotted line) and exponential (dark blue dotted line) models that are built by using classical statistical techniques; a comparison also is possible between international models and the statistical model described here. The standard roundabout legs in this statistical model show a relationship between Q_c and Q_e that is similar to the relationship in the existing models. The slight differences are related to the behavior of drivers in different countries.

Figure 4 shows the relationship Q_e versus Q_c for nonstandard legs. The models (respectively linear and exponential) are

$$Q_e = 764.35 + 0.014 \cdot Q_c \quad (7)$$

$$Q_e = 7713.37 \cdot e^{(-0.00003 \cdot Q_c)} \quad (8)$$

where Q_e is the entering flow (pcu/h), and Q_c is the conflicting flow (pcu/h).

Unlike standard legs or the international models, Figure 4 shows that Q_e stays nearly constant for each Q_c value. In fact, for low and high volumes of Q_c, nonstandard legs produced low-constant Q_e values. In the case of tangent legs, the Q_e value should have been higher than the value for the standard legs. That was not the case. It appears that geometric features do not influence the Q_e values, which seem to be linked only to local characteristics.

Q_e values also were constant for high Q_c values because of the forced entering maneuvers of drivers. In fact, nonstandard geometry (i.e., legs without deflections) allows drivers that are coming from a tangent leg to force their entry into the roundabout. Thus, Q_c stops to give way to Q_e.

GEOSTATISTICAL TREATMENT OF EXPERIMENTAL DATA

In all scientific fields, experimental data analyses and processing are important. In fact, in most cases it is necessary to start from a limited set of collected data in order to reconstruct the phenomenon of interest within its existing domain (not necessarily coincident with the whole data domain). To do that, the enlargement of the sampling mesh may not represent the best solution: investigation time and costs, or in some cases, the inaccessibility of measurement points can hinder the feasibility of research choice. In practice, therefore, researchers try to maintain as low as possible the number of observations, and proceed to create representative models of the phenomenon in their studies.

Among the available alternatives, geostatistical analytical and interpolation techniques have assumed high importance in all branches of engineering science that require experimental data processing instruments. These techniques, now implemented in all of the most common software (e.g., Surfer and ArcGIS) are developed on the basis of correlation studies of collected observations and on the use of related law to make the interpolation.

From a geostatistical point of view, collected data may not represent a random sample of values but, in the interpretation of the precise expression of the same phenomenon, must be objectively related to each other. Recognition of the correlation law between the

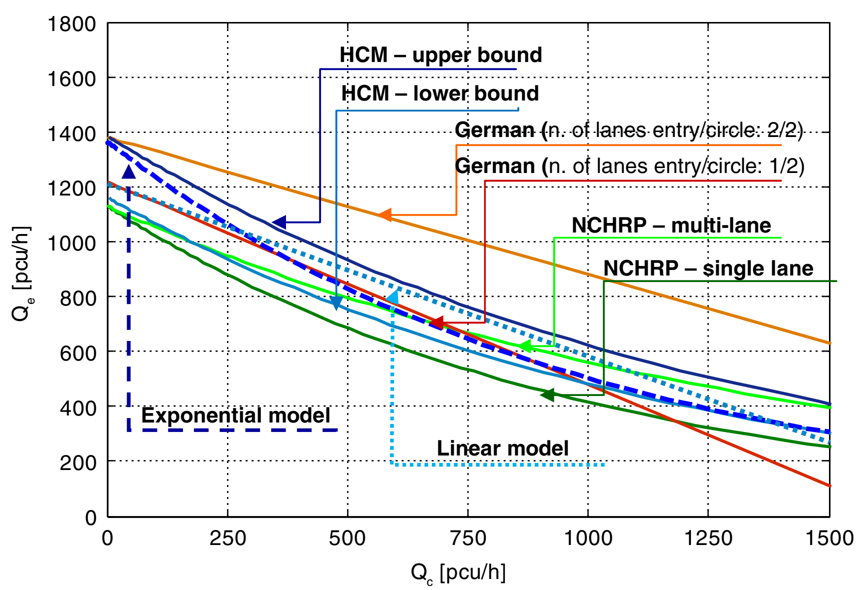


FIGURE 3 Results for Q_c versus Q_e (standard legs) modeled with statistical regression.

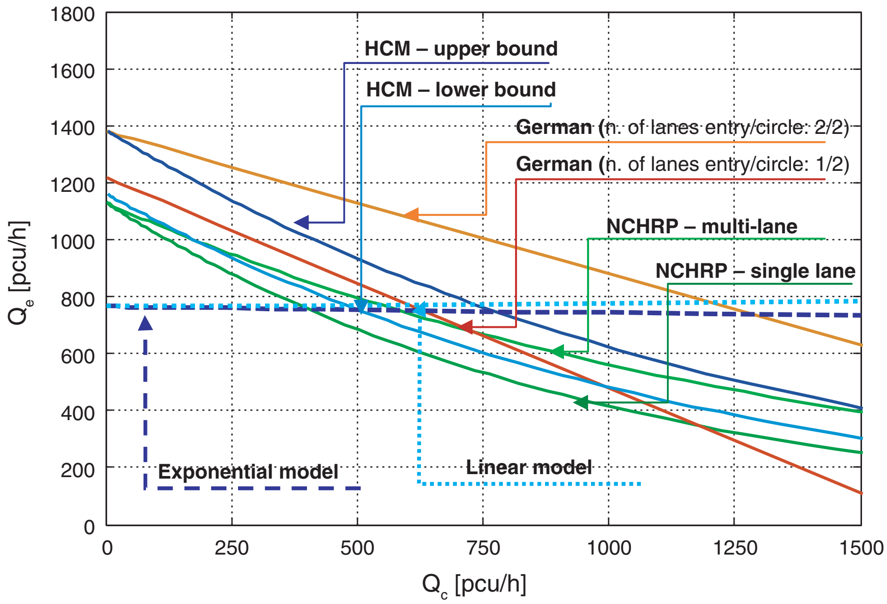


FIGURE 4 Results for Q_c versus Q_e (nonstandard legs) modeled with statistical regression.

observations allows an interpolation to be carried out on the basis of the real law of the phenomenon.

Geostatistics was born last century in the mining field, on the basis of the theories of Georges Matheron, Danie Krige, and Herbert Sichel as science for providing tools for the evaluation of mineral deposits. Since then, much progress has been made in the development of geostatistical techniques of analysis and interpolation. For this reason, geostatistics has become an extremely powerful tool for studying and modeling spatial and temporal phenomena.

Geostatistics supplies a collection of techniques that addresses the study of a spatial correlation between the experimental values of a specific variable, to estimate values in unknown points within the phenomenon's existing domain (*10–12*).

In the research described in this paper, classical geostatistical tools were used in an innovative way to study the correlation law, first between Q_e and Q_c (which are not spatial or temporal coordinate values) and then to use the law to interpolate available data.

A geostatistical analysis consists primarily of the following steps:

1. Estimate spatial correlation between measured observations. Such analysis is carried out by construction of an experimental curve (Figure 5) that expresses the correlation between collected data.

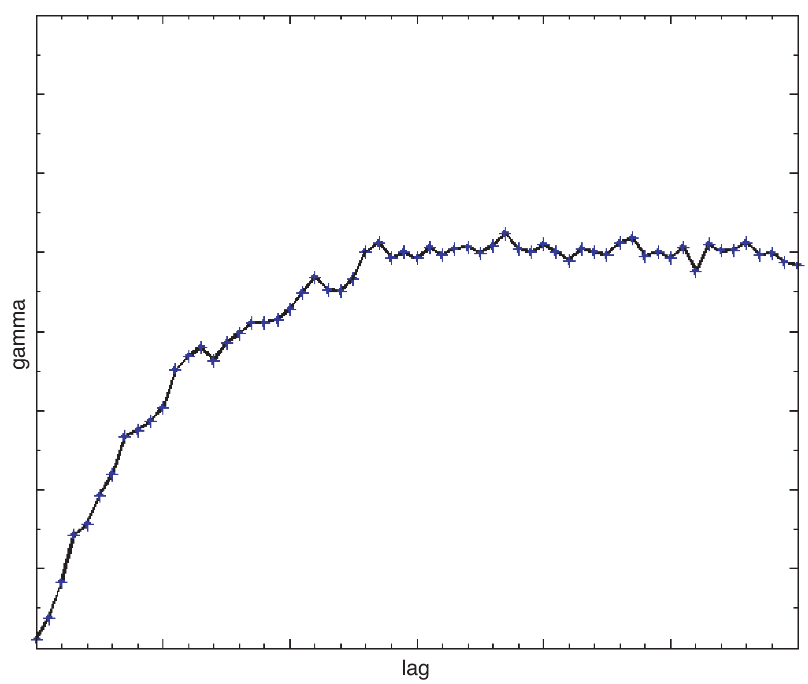


FIGURE 5 Sample experimental semivariogram.

This relationship (or more precisely, the variability of the size studied) is represented by the function of the experimental semivariogram (Figure 5) defined by the equation

$$\gamma(h) = \frac{1}{2}\mathrm{Var}[Z(x+h) - Z(x)]^2 \qquad (9)$$

where $\gamma(h)$ represents the error that would be made if the unknown point was replaced with the collected data at the definite distance (*11*).

2. Model the experimental semivariogram by using one of the theoretic models (Figure 6) of a well-known equation that fits the trend of experimental semivariogram, which, as already stated, represents the phenomenon law. Calculate and model the experimental semivariogram to obtain important information about the phenomenon under examination and also so that it can serve as the basis by which to define the values of the unknown points.

3. Define the variable parameters of the phenomenon that is undergoing study. This step is necessary to identify uniquely the phenomenon law and to obtain the best fit between the experimental and theoretical semivariograms. As shown in Figure 7, these parameters are the following:

– The nugget effect, which represents the value of the semivariogram for pairs of observations separated from distances close to zero;

– The range, which represents the distance, measured along the horizontal axis, where the curve of the semivariogram tends to flatten, or rather, beyond which the data become totally independent. For this reason, all values beyond this limit with respect to the unknown point are not considered during interpolation; and

– The sill, which corresponds to the value of the semivariogram reached at a distance equal to the range.

4. Estimate the unknown values and error calculation associated with the geostatistical linear interpolation known as "kriging" (*11*). The kriging technique is able to obtain a goodness of interpolation otherwise difficult to achieve with other techniques. With regard to classical interpolation techniques, the kriging advantages are mainly the following:

– Estimation procedure is on the basis of the real correlation law between the data (represented by the experimental semivariogram);

– Interpolation leads to unique solutions;

– Interpolation of data is exact (i.e., interpolation returns the experimental value in the case of measured point estimation) and is also able to give the estimated variance for each calculated value (in this way, it is possible to build both the phenomenon map and the estimation error map at once); and

– Results depend only on the model of the semivariogram (from the distance between sampled points and the unknown point to estimate) from sampling geometry and not from measured values.

For these reasons the kriging is defined as the best linear unbiased estimator. Once the variographic parameters are defined, interpolation of available data is conducted with the kriging method. From a mathematical point of view, the kriging estimation technique is defined by Equation 10

$$\hat{Z}_x = \sum_{i=1}^{n} \lambda_i(x) \cdot Z(x_i) \qquad (10)$$

where

$\hat{Z}_x$ = the unknown point value,
$Z(x_i)$ = each measured value to be considered for the interpolation, and
$\lambda_i(x)$ = appropriate weights, which depend on the position and distance of each of these points from the unknown point, and which satisfy the following condition

$$\sum_{i=1}^{n} \lambda_i(x) = 1 \qquad (11)$$

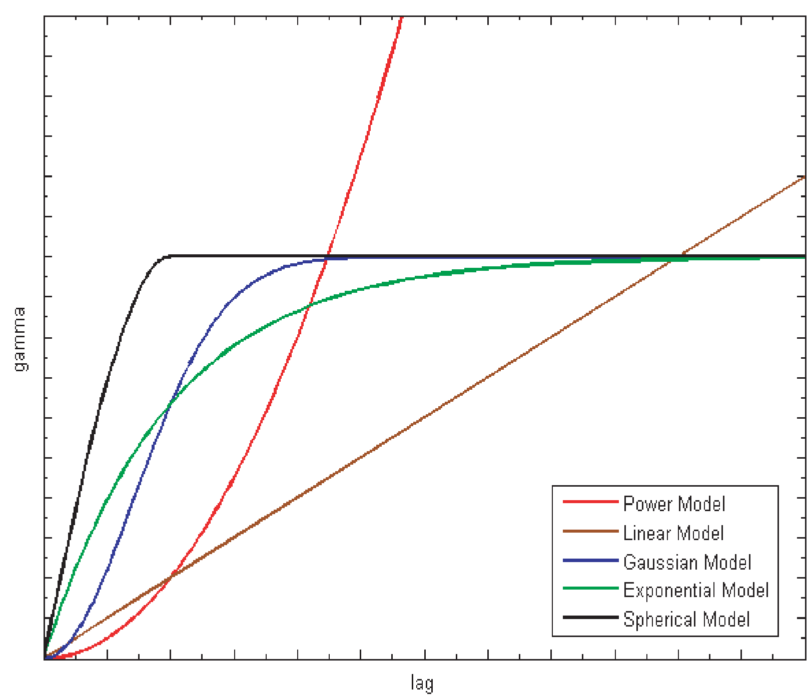


FIGURE 6 Theoretical models for experimental semivariogram interpretation.

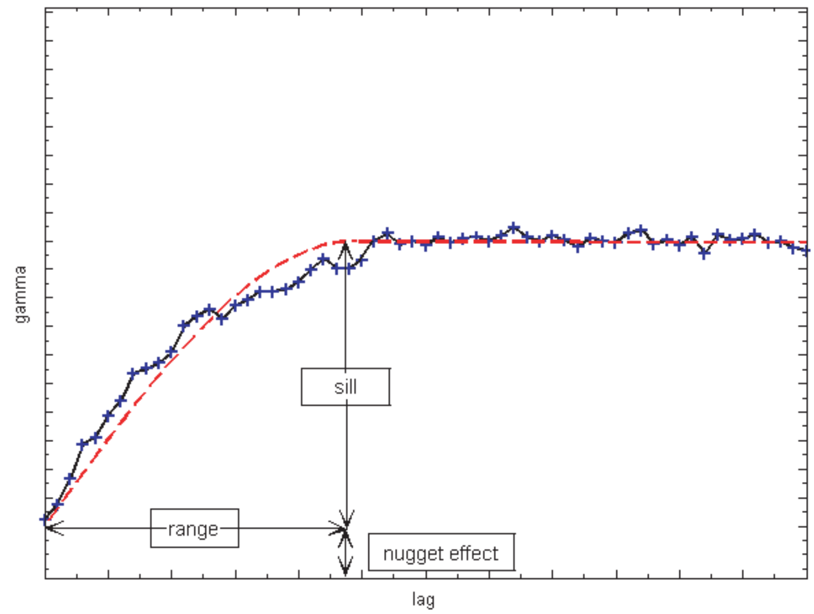


FIGURE 7 Variographic parameters of experimental semivariogram.

Figure 8 illustrates the kriging concept and provides insight into its potential. In summary, the kriging technique allows the enlargement of a cloud of data by using the information provided by the data that were actually collected.

5. Provide representations (e.g., contour maps, block maps) of the results, which depend on the type of data and the use of results. This procedure must be carried out after processing the data, because the kriging results are represented by a matrix of numerical values that is difficult to read by those who do not possess the specific know-how.

GEOSTATISTICAL ANALYSIS RESULTS

The preliminary geostatistical study and successive application of the kriging technique were conducted on imaginary grids composed of linear grids of points with the following coordinates:

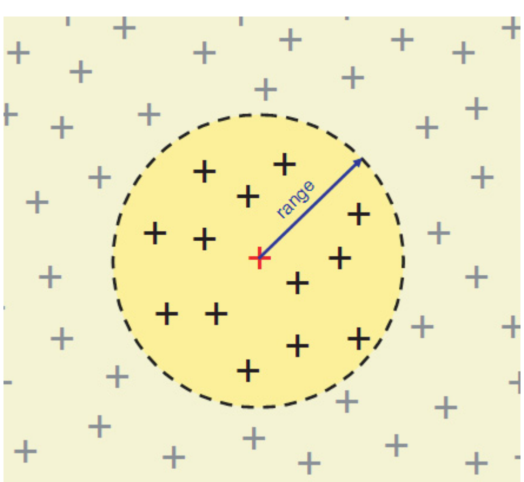


FIGURE 8 Representation of kriging techniques.

- Q_c (pcu/h), Q_e (pcu/h) for the construction of the model that expressed Q_e versus Q_c on standard legs and
- Q_c (pcu/h), Q_e (pcu/h) for the construction of the model that expressed Q_e versus Q_c on nonstandard legs.

Model Q_e = f(Q_c) on Standard Legs

As shown on the left side of Figure 9, the experimental semivariogram modelization was conducted with the following parameters:

- Theoretical model: Gaussian model and
- Variability parameters:
  - – Nugget effect = 89,000,
  - – Sill = 73,300, and
  - – Range = 423.

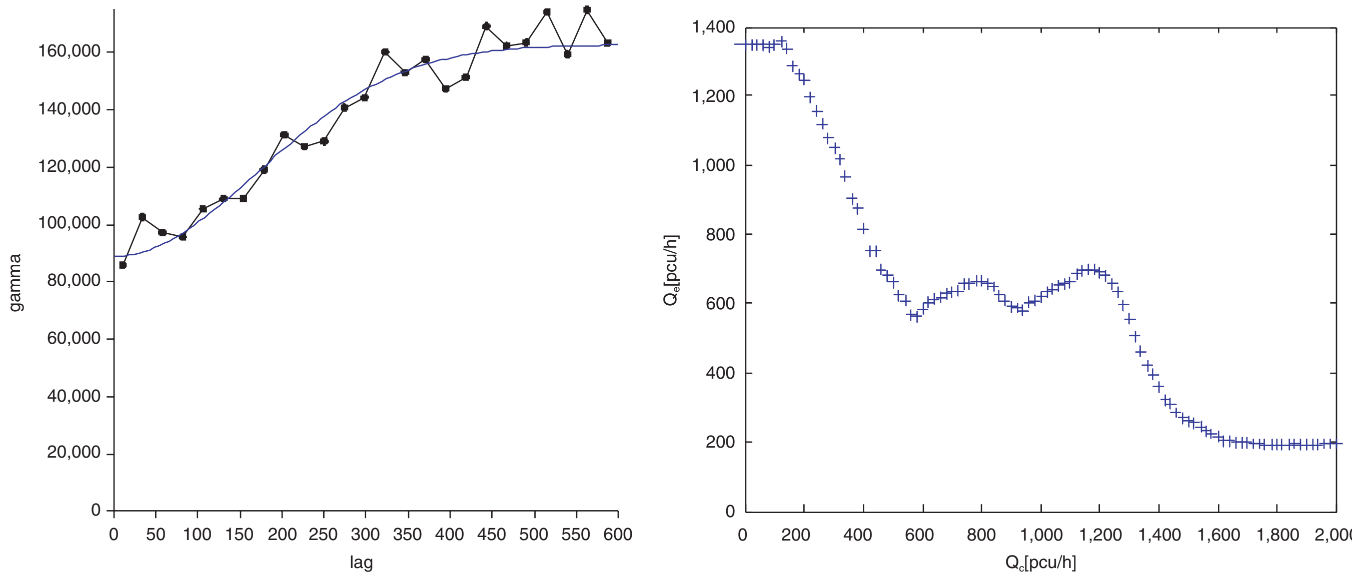
The modeling was conducted by choosing the best-fitting theoretical model to describe the general trend of the experimental semivariogram and to further estimate variographic parameters by using the least square criterion.

Figure 9 (right part) represents the result of a kriging interpolation carried out by using the previously defined parameters. In observing the trend, some considerations were taken into account.

The trend of a model on a geostatistical basis depends on the Q_c values, whereas other models suggested by the international literature have only one trend for every flow capacity value. This demonstrates that the a priori choice of regression type (linear, quadratic, or cubic) could produce dangerous errors in the evaluation.

In Figure 10, it is evident that the Q_e versus Q_c trend can be divided into three parts as follows:

1. When Q_c values are lower than 200 pcu/h, any further increases in Q_c imply that Q_e values remain constant. In this range, no specific



(a) (b)

FIGURE 9 Experimental semivariogram for Q_e versus Q_c: (*a*) interpretation and (*b*) representation of kriging result for Q_e versus Q_c modeling.

correlation exists between Q_c and Q_e. In fact, it is often emphasized that a model with Q_e versus Q_c values should be built for Q_c values that are higher than 200 pcu/h.

2. When Q_c values are greater than 1,500 pcu/h, the Q_e does not change if the Q_c values increase. In fact, when the Q_c is high, the entering flow probably forces entrance maneuvers. In this case, the roundabout often works like a normal intersection, in which drivers that are coming from another leg can force their way in, and thus cause the circulating flow to stop and give way to the entering vehicles. This research confirms the findings of other studies in which a specific model often has validity, between a given minimum and maximum Q_c independently by country (*1, 2, 5, 6*).

3. When Q_c values are 200 pcu/h ≤ Q_c ≤1,500 pcu/h, the Q_e values vary strictly according to the Q_c values, with at least another three behaviors.

As evident in Figure 10, it is not easy to define, without further studies, the exact limit between one range of variation and another within 200 pcu/h ≤ Q_c ≤ 1,500 pcu/h. Further studies are needed and, in fact, such research is in progress.

A first analysis was conducted on the basis of analytical geometry. The trends that have resulted from this study are significant: at least three variation fields of the modelization of Q_c versus Q_e were found within 200 pcu/h ≤ Q_c ≤ 1,500 pcu/h.

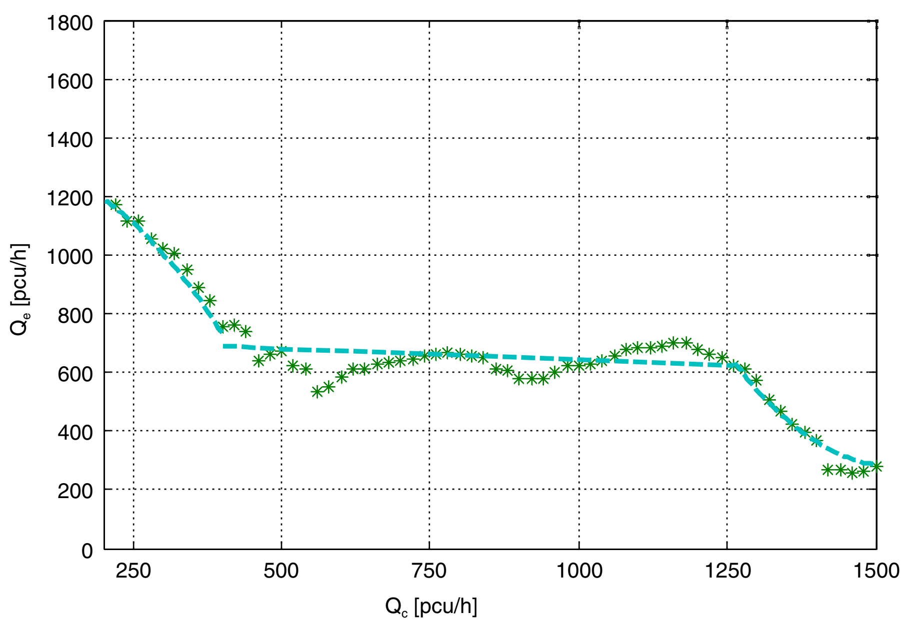


FIGURE 10 Interpolation results for Q_c versus Q_e modeling.

As shown in Figure 10, on the basis of the data in the range 200 ÷ 1,500 pcu/h, the geostatistical study underscored the existence of three subdomains.

1. For 200 pcu/h ≤ Q_c ≤ 400 pcu/h, the trend was decreasing.

$$Q_{e1} = -0.0044 \cdot Q_c^2 + 0.34 \cdot Q_c + 1{,}300 \quad (12)$$

2. For 400 pcu/h ≤ Q_c ≤ 1,250 pcu/h, the trend seemed constant, or subject to a quasi-increase, but other studies are needed to validate that

$$Q_{e2} = -0.077 \cdot Q_c + 720 \quad (13)$$

3. For 1,250 pcu/h ≤ Q_c ≤ 1,500 pcu/h, the trend was decreasing with a slope similar to the one for the 200 pcu/h ≤ Q_c ≤ 400 model.

$$Q_{e3} = 0.0049 \cdot Q_c^2 - 15 \cdot Q_c + 11{,}760 \quad (14)$$

The enormous potential of data analysis through geostatistical techniques has led to the discovery that the relationship between Q_c and Q_e within 200 pcu/h ≤ Q_c ≤ 1,500 pcu/h varies with the value given by the circulating flow. This is probably related to the fact that the magnitude of change causes variations in the flow, which affect the behavior of drivers in roundabouts. For strong circulating and entering flows, for example, the last one forces the entrance maneuvers.

These conclusions would not be underscored with the use of statistical regression techniques that do not take into account the changes in trends by using a single regression model (whether it is linear or exponential) and use only an a priori choice of regression type, which can cause a dangerous mistake in assessment.

Model $Q_e = f(Q_c)$ on Nonstandard Legs

As shown in Figure 11, experimental semivariogram modelization was conducted with the following parameters:

- Theoretical model: Gaussian model and
- Variability parameters:
  - – Nugget effect = 31,100,
  - – Sill = 35,100, and
  - – Range = 229.3.

The results were abnormal: the environmental conditions appeared to influence the relationship between circulating and entering flows.

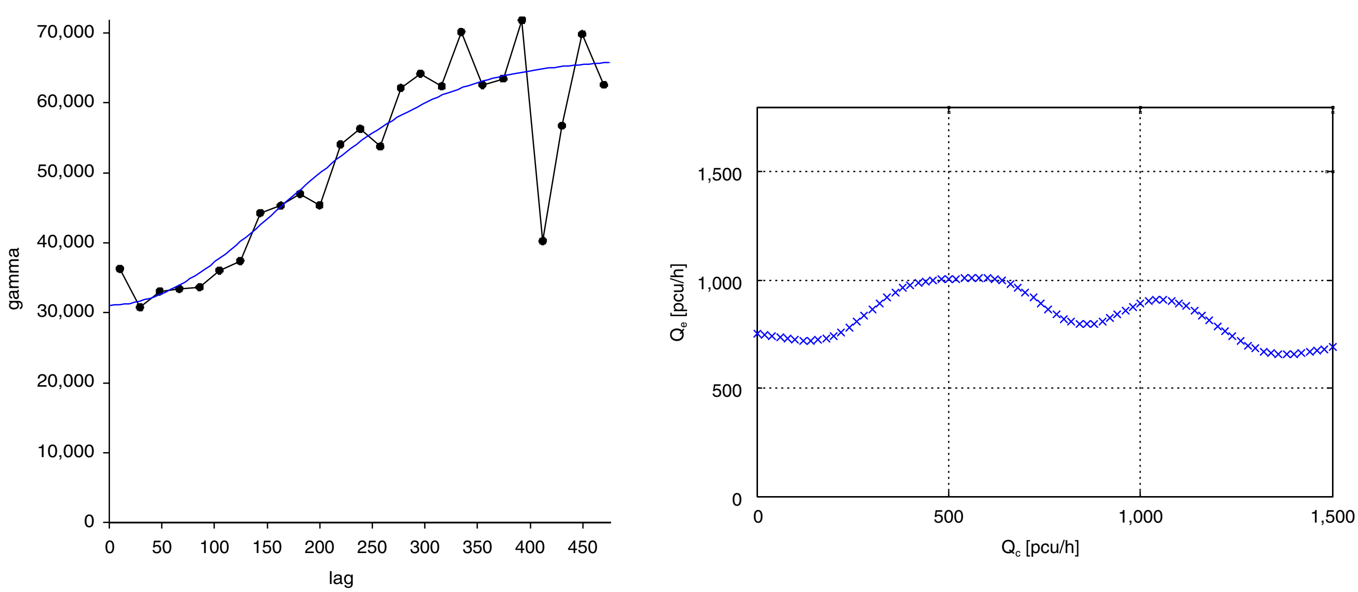
Unlike standard legs, or those of the international models, geometric characteristics did not influence the flow values; this result seemed to be linked to local features. Nonstandard features (in this case, tangent legs) allow drivers that are coming from a tangent leg to force their way into the roundabout, which causes the circulating flow of traffic to stop and give way to the entering flow.

CONCLUSIONS

Intersections are dangerous places in road systems because of the high number of conflicting points between motorized and nonmotorized vehicles. In the case of defined traffic volume, several research groups have proven that roundabouts reduce the number of injuries and fatal accident cases. In recent years, many countries have adopted roundabouts as a standard design solution for both urban and rural roads.

Researchers have investigated the performance of roundabouts with different models that can calculate the entering flow (Q_e) as a function of the circulating flow (Q_c) of traffic. All of these models have been constructed by using statistical regression techniques. Currently, no research group has used geostatistical analysis techniques, even though the techniques are largely used in many other fields of research in the engineering sciences.

An applied geostatistical approach starts from the consideration that, in this field of research, the collected data cannot represent a random sample of values. Rather, the data items must objectively



(a) (b)

FIGURE 11 Experimental semivariogram for Q_c versus Q_e: (*a*) interpretation and (*b*) kriging interpolation result for Q_c versus Q_e modeling.

relate to each other under a precise law: the phenomenon law. The study of the law between observations allows the realization of a interpolation based on kriging among collected data and the definition of models Q_e versus Q_c for both standard and nonstandard legs.

For standard legs, the kriging application resulted in the following findings:

- Confirmed that Q_e versus Q_c does not exist within [0, 200] *e* [1,500, 2000] which supports the existing international literature, and
- Underscored that the relationship between Q_c and Q_e within 200 pcu/h ≤ Q_c ≤1,500 pcu/h cannot be expressed by one trend only (such as that in the international models) but by two or three trends.

For nonstandard legs, the kriging application resulted in the following findings:

- Showed that geometric features do not influence the flow values, which seem to be linked to other local features, and
- Showed that nonstandard features (i.e., tangent legs) allow drivers that approach from a tangent leg to force their way into the roundabout.

The implications of this research have led to its continuation. Work is still in progress to understand the phenomena that influence the Q_e versus Q_c relationship and to validate the obtained results.

REFERENCES

1. Matheron, G. *The Theory of Regionalized Variables and Its Applications.* Paris School of Mines, Paris, 1971.
2. Mazzella, A., and A. Mazzella. Kriging Assistant: A Geostatistical Analysis and Evaluation Tool. American Society of Civil Engineers Conference Proceedings. Geoenvironmental Engineering and Geotechnics: Progress in Modeling and Applications (GSP 204). *Proc., GeoShanghai International Conference, Shanghai, China*, 2010, pp. 188–193.
3. Rodegerdts, L., M. Blogg, E. Wemple, E. Myers, M. Kyte, M. P. Dixon, G. F. List, A. Flannery, R. Troutbeck, W. Brilon, N. Wu, B. N. Persaud, C. Lyon, D. L. Harkey, and D. Carter. *NCHRP Report 572: Roundabouts in the United States.* Transportation Research Board of the National Academies, Washington, D.C., 2007.
4. *Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 2000.
5. Brilon, W. *Roundabouts: A State of Art in Germany.* Presented at National Roundabout Conference, Vail, Colo., 2005.
6. Akcelik, R. Lane-by-Lane Modeling of Unequal Lane Use and Flares at Roundabouts and Signalized Intersections: the SIDRA Solution. *Traffic Engineering and Control,* Vol. 38, Nos. 7–8, 1997 pp. 388–399.
7. Akcelik, R. *SIDRA Intersection User Guide.* Greythorn, Victoria, Australia, 2006.
8. Akcelik, R. *Evaluating Roundabout Capacity, Level of Service and Performance.* Presented at Institute of Transportation Engineers Annual Meeting, San Antonio, Tex., 2009.
9. Annunziata, F., M. Coni, F. Maltinti, F. Pinna, and S. Portas. *Le Infrastrutture Viarie: Dalla Ricerca Al Progetto.* Dario Flaccovio, Palermo, Italy, 2007.
10. Matheron, G. Principles of Geostatistics. *Economic Geology,* Vol. 58, 1963, pp. 1246–1266.
11. Krige, D. G. Geostatistics and the Definition of Uncertainty. *Inst. Min. Met. Trans.,* Vol. 93, 1984, pp. 41–47.
12. Journel, A. G., and C. J. Huijbregts. *Mining Geostatistics.* Academic Press, London, 1978.

The Statistical Methods Committee peer-reviewed this paper.