

TITLE: Geostatistical Approach for Operating Speed Modelling on Italian Roads

Topic Area Responding to “Human factor in geometric design”

Authors:

Alessandro Mazzella (PhD Student – Environmental Engineer)
Department of Geoengineering and Environmental Technologies, University of Cagliari
Piazza d’Armi 19 - 09123 Cagliari (ITALY)
Phone-number: +39-070-6755511, Fax: +39-070-6755523
e-mail: amazze@unica.it

Francesco Pinna (Assistant Professor – Civil Engineer)
Department of Land Engineering, University of Cagliari
Piazza d’Armi 19 - 09123 Cagliari (ITALY)
Phone-number: +39-070-6755256, Fax: +39-070-6755266
e-mail: fpinna@unica.it

Claudia Piras (PhD – Civil Engineer)
Department of Land Engineering, University of Cagliari
Piazza d’Armi 19 - 09123 Cagliari (ITALY)
Phone-number: +39-070-6755256, Fax +39-070-6755266
e-mail: cpiras@unica.it

Word count

text: 3889

figure: 250 x 12 = 3000

table: 250 x 0 = 0

total: 6889

ABSTRACT

Injuries are one of the main causes of death according to W.H.O. For this reason the attention of road safety researchers especially regards the study of the relationship between driver and road environment.

Several research works show that operating speed is an excellent driver behavior parameter. This article describes a different approach to the classical definition of prediction models for operating speed on horizontal curves. In this paper, the fundamental theories, the applied operating procedures and the first results obtained with the application of Geostatistics are discussed. The mathematical models expressing operating speed in function of horizontal curves characteristics found in International scientific literature, have mainly been built on the basis of Classical Statistics. For this reason, it needs to be pointed out that the interpolative techniques found in Classical Statistics are based upon the use of canonical forms (linear or polynomial regressions) 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.

The models described in this article have instead been created with the geostatistical interpolation technique (i.e. Kriging). This technique allows to obtain the "*best*" estimates possible because it considers the true correlation law between the measured data.

The applied methods are then described along with the results obtained in the field of road safety by applying Geostatistics which, for several years, have been used, with positive results, in all scientific and engineering fields dealing with empirical data analysis and processing.

INTRODUCTION

Numerous studies have been developed in the last decades in the field of road safety in order to forecast, already during the project phase, the relationship between real driver behavior and rural road features (especially on horizontal curve, which are more dangerous than tangent). Many of these studies focus on creating analytical models expressing driver behavior in relation to the characteristics of the two lane rural roads. In these models, driver behavior is expressed by the operating speed (V_{85} , that represents the 85th-percentile speed), while the road characteristics are generally represented by the Curvature Change Rate of the single curve (CCRs) or the horizontal curve radius (R) (1).

To create such models, various research groups have carried out a lot of survey campaigns to collect a numerous mass of inherent data on the characteristics of the road and the corresponding operating speed, to elaborate with statistical techniques (2).

The present article arises from these considerations with the objective of introducing the experiences accomplished by our research team in the developing new analytical models to predict V_{85} . These new predict models have been created by Geostatistics analysis and interpolation. This technique among the various advantages offered, allows to obtain the interpolation that minimizes the estimated variance value: to determine unknown values, it considers only the experimental values that exercise a real influence and weighs them with appropriate coefficients obtained by studying the correlation laws between experimental values themselves. The strong point of this approach is to admit that the collected data do not represent a sample of random values but they are a punctual manifestation of a phenomenological reality regulated by a precise law. This is a strong point of the Geostatistics approach in respect to the Statistics one.

The article is divided into four main parts. The first part briefly summarizes the main fields of application of Geostatistics and the advantages of its use, reporting a schematic description of the fundamental steps making up the geostatistical analysis. The second part describes the data collected to calibrate the proposed model. The third part describes the development of our proposed predict model using geostatistical techniques. The final part summarizes the main conclusions of the study and briefly indicates its future developments.

GEOSTATISTICAL ANALYSIS OF EXPERIMENTAL DATA

In all scientific fields, experimental data analysis and processing assume a considerable importance. In fact, in the majority of cases, starting from a limited set of collected information it is necessary to reconstruct a phenomenon of interest within its existing domain. In order to do that, the enlargement of the sampling mesh does not represent the best solution: investigation time and costs or, in some cases, the inaccessibility of points of measurement can hinder the feasibility of this choice. Therefore, the target is, in practice, to maintain low (as possible) the number of observations and to proceed to create representative models of the phenomenon in study.

Among the available alternatives, Geostatistics analytical and interpolation techniques have assumed a high importance today in all those branches of engineering science requiring experimental data process instruments. These techniques, now implemented in all most common software (for example, Surfer and ArcGIS), are based on correlation studies of collected observations and use of that law for the interpolation.

This is because the experimental data gathered may not represent a random sample of values but, in interpreting the precise expression of the same phenomenon, must be objectively related to each other. Recognition of the correlation law between the observations allows to carry out an interpolation based on the real law of phenomenon.

Geostatistics was born last century in the mining field, based on studies conducted by Georges Matheron, Danie Krige and Herbert Sichel, as science for providing tools for the evaluation of mineral deposits. Since then a lot of progresses has been made in the development of Geostatistics techniques

of analysis and interpolation. For this reason Geostatistics has become an extremely powerful tool for studying and modeling of space/time phenomena.

Therefore Geostatistics supplies a collection of techniques addressed to the study of spatial correlation between experimental values of a specific variable, representing a phenomenon in study in order to determine the value of the unknown points within the existing dominion (3) (4) (5).

In the research work proposed in this paper, the classical Geostatistics tools are used in an innovative way for studying the correlation law between the operating speed and road parameters (as CCRs or R) that are not a space and/or time coordinates but only design features.

More simply, a geostatistical analysis consists of the following main steps:

1. estimate of spatial correlation between measured observations. Such analysis is carried out by construction of an experimental curve (Figure 1) expressing the spatial and/or time correlation between collected data. This relationship, or more precisely, the variability of the size studied is represented by the function of the experimental semivariogram (Figure 1), defined by the equation:

$$\gamma(h) = \frac{1}{2} \text{Var} [Z(x) - Z(x+h)] \quad (1)$$

$\gamma(h)$ represents the semi-variance between collected pairs of observations separated by a specific distance (measured along the horizontal axis) or, rather, the error which would be made if the unknown point was replaced with the collected data at the definite distance (4).

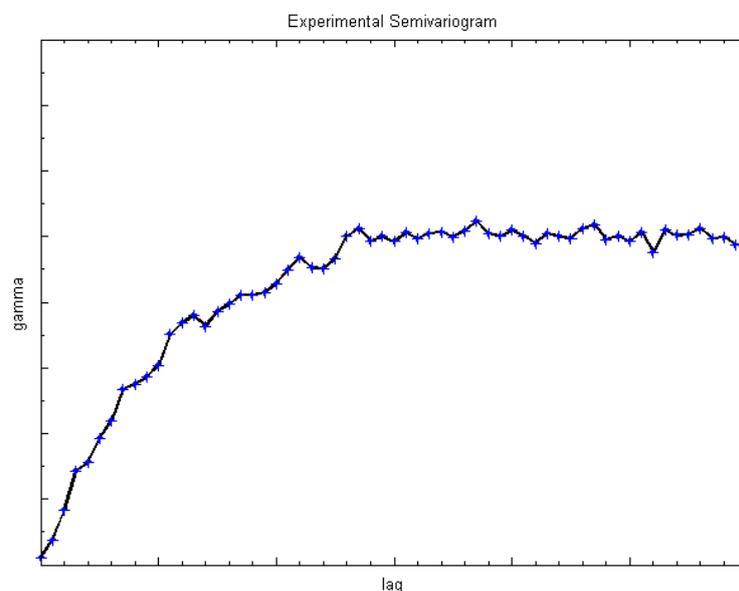


FIGURE 1 An experimental semivariogram example

2. modeling the experimental semivariogram using a theoretic model, of a well-known equation, that fits the trend of experimental semivariogram which, as already stated, represents the phenomenon law. Calculating and modeling the experimental semivariogram offers a series of important information of the phenomenon in examination, but it also represents the need of a basis to define the values of the unknown points.

In the Figures 2 some of the main theoretic models are shown.

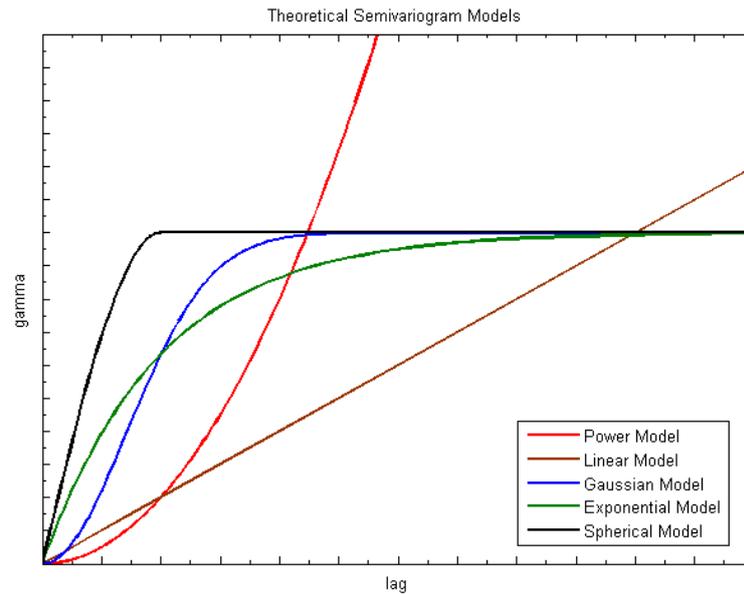


FIGURE 2 Theoretical models for experimental semivariogram interpretation

3. **definition of the variable parameters of the phenomenon in study.** This step is necessary to identify uniquely the phenomenon law in order to obtain the best fit between the experimental and theoretical semivariogram. Referring to the Figure 3, these parameters are:

- the **nugget effect**, which represents the value of the semivariogram for pairs of observations separated from a 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 in respect to the unknown point, should not be considered during the interpolation;
- the **sill** corresponds to the value of the semivariogram which is reached at the distance equal to the range.

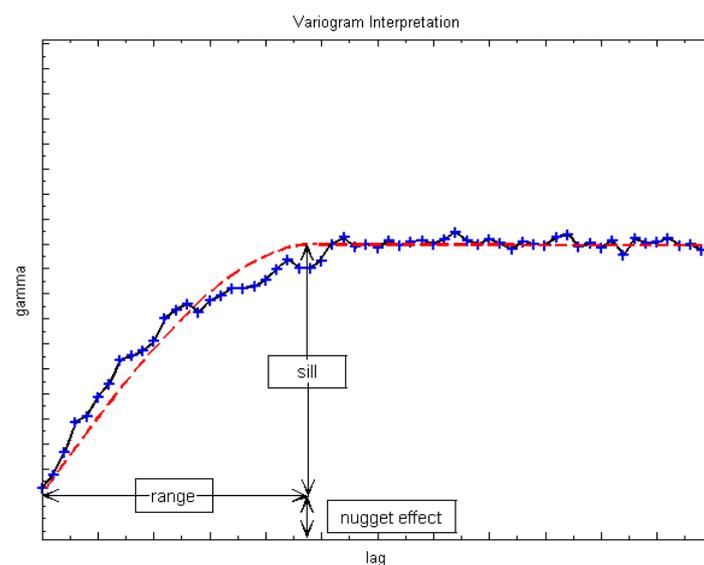


FIGURE 3 Variographic parameters of experimental semivariogram

4. **estimate of the unknown values and error calculation associated** with the geostatistical linear interpolation, named “Kriging” (4). The Kriging is able to obtain a goodness of interpolation otherwise difficult to achieve with other techniques.

Regarding classical interpolation techniques, Kriging advantages are mainly:

- a. Kriging estimation procedure is based on the real correlation law between the data (represented by the experimental semivariogram);
- b. Kriging interpolation leads to unique solutions;
- c. Kriging interpolates the data exactly (i.e. it returns the experimental value in the case of measured point estimation) and also it is able to give the estimated variance for each calculated value (in this way is possible to build both the phenomenon map and the estimation error map at once);
- d. Kriging results only depend from the model of semivariogram, from the distance between sampled points and unknown point to estimate, from sampling geometry and not from measured values.

For these reasons the kriging is defined as “best linear unbiased estimator” (B.L.U.EST.).

From a mathematical point of view Kriging estimation technique is defined by the equation 2:

$$\hat{Z}_x = \sum_{i=1}^n \lambda_i Z_i \quad (2)$$

where \hat{Z}_x represents the unknown point value, Z_i represents each measured value which will be considered for the interpolation and λ_i are appropriate weights which depend on the position and distance of each these point from unknown point and which satisfies the condition 3:

$$\sum_{i=1}^n \lambda_i = 1 \quad (3)$$

5. **results graphical restitution** (i.e. contour maps, block maps etc.) depending on the type of data and use of results. This procedure must be carried out after processing the data since the Kriging results are represented by a matrix of numerical values which is difficult to read by those who do not possess a specific know-how.

DATA COLLECTION

This paragraph describes the phase of the experimental data collection and the construction of a datafile to develop the investigation. Our analysis only concentrated on horizontal curves in order to construct a prediction model where operating speed is expressed as a function of CCRs, R and both R and L (where L is the length of the horizontal).

In order to make the results of our work comparable with those proposed in international bibliography, the survey campaigns were carried out in correspondence to several Italian two lane rural road sections with the following characteristics (1) (2):

- curves found in two lane rural roads;
- lack of elements which could influence driver (such as working zones signals);
- longitudinal grade below 5%;
- low traffic flow;

seconds) were considered.

At this point the operative speed associated to each section of recording was calculated: the 85th percentile speed was calculated for probability of distribution for each curve. A data file was then constructed containing design values for each section of the curves (CCRs, R, length, DC) and operating speeds.

GEOSTATISTICAL ANALYSIS OF COLLECTED SPEED DATA

The preliminary geostatistical study and successive application of the Kriging technique were conducted on imaginary grids composed of:

- a linear grid of points with coordinates (CCRs [gon/km], V_{85} [km/h]), for the construction of the model expressing CCRs vs V_{85} ;
- a linear grid of points with coordinates (R [m], V_{85} [km/h]), for the construction of the model expressing R vs V_{85} ;
- a 2D grid of points with coordinates (R [m], L [m], V_{85} [km/h]), for the construction of the model expressing R and L vs V_{85} .

Model $V_{85}=f(\text{CCRs})$

The calculation of the data collected in the experimental semivariogram gave us the result in the Figure 5.

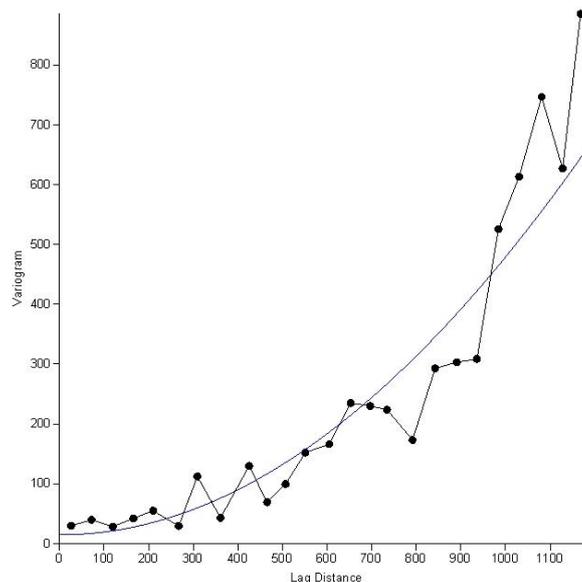


FIGURE 5 Experimental semivariogram for V_{85} vs CCRs and its interpretation

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

The Figure 6 represents the Kriging result of the data carried out using parameters of variographic calculated models.

The trend of the model seems almost linear, like the other models suggested by international literature, but it has not a perfect linear trend. This demonstrates that the *a priori* choose of the regression type produces errors in all the phenomenon evaluation.

That is underlined also in the Figure 7, which shows the comparison between bibliographic

international models and our one.

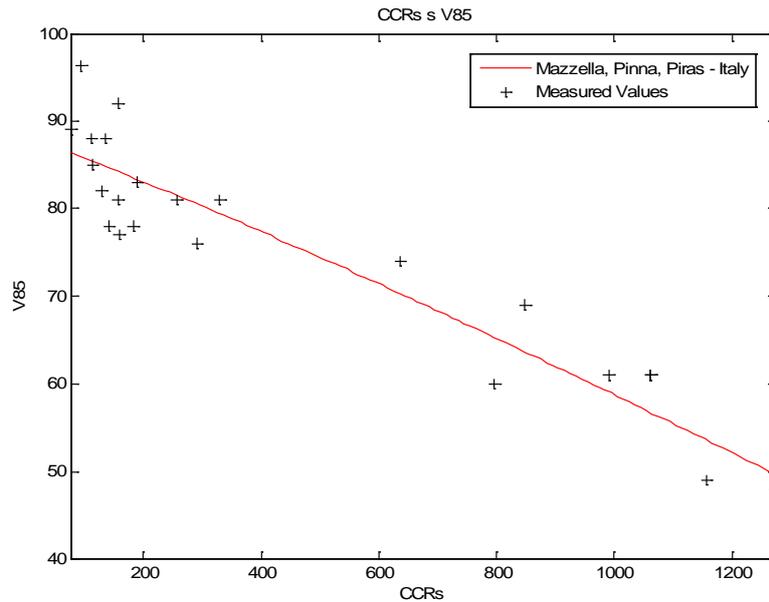


FIGURE 6 Kriging interpolation result for the V_{85} vs CCRs model

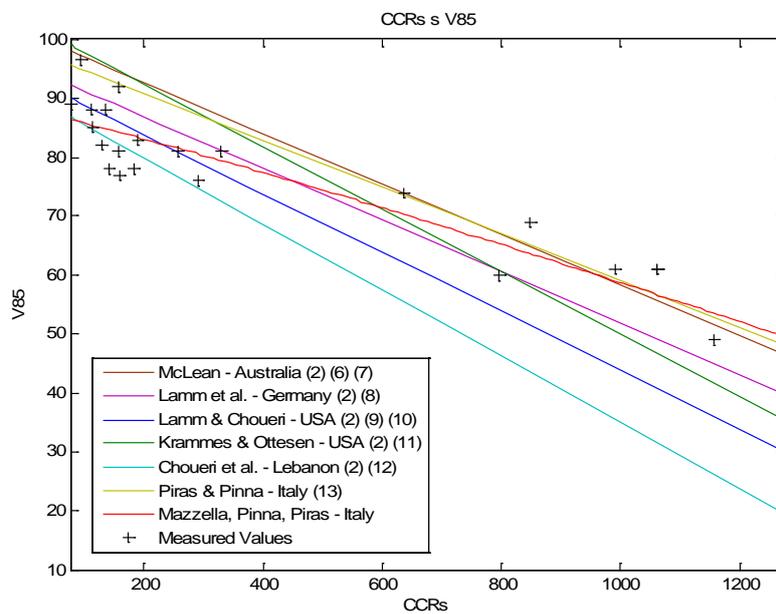


FIGURE 7 Comparison between international models and interpolation result for the V_{85} vs CCRs

The Figure 7 shows the evolution of speed on curves with a radius of 50 to 850 m. It shows a similar trend between the proposed model and most of those in use in various countries, even if the trend of our model, cutting the other ones, fits better the measured data.

Model $V_{85}=f(R)$

The calculation of the data collected in the experimental semivariogram gave us the result in the Figure 8.

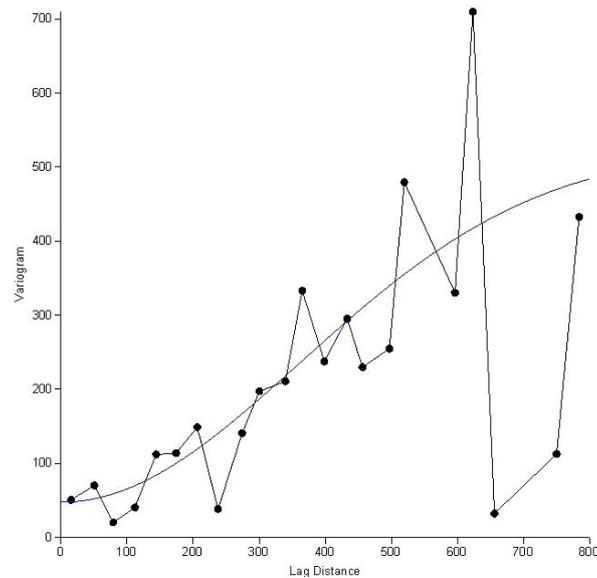


FIGURE 8 Experimental semivariogram for V_{85} vs R and its interpretation

The modeling was conducted by selecting the best fitting theoretical model to describe the general trend and to further estimate variographic parameters using least square criterion. The Figure 9 represents the Kriging results carried out using parameters of variographic calculated models.

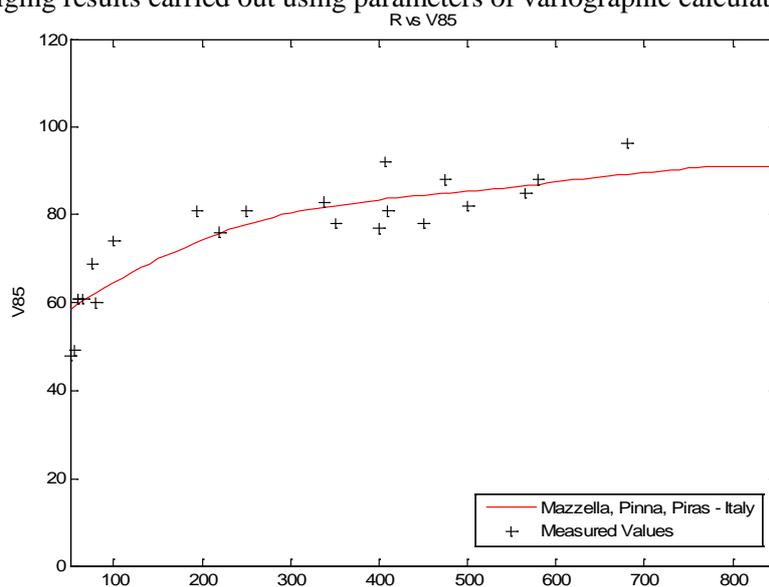


FIGURE 9 Kriging interpolation result for the V_{85} vs R model

The parabolic trend of the proposed model fits better than other models the distribution of collected data. Also in this case this demonstrates that the *a priori* choice of the regression type produces serious errors on the phenomenon evaluation.

Geostatistical curve built on measured data shows the existence of three different behaviors of the phenomenon: first for radii between 50 m and 300 m, second for radii between 300 m and 750 m, third for radii greater than 750 m (Figure 9). On this topic our research team is still working. Like the previous case, the Figure 10 shows some the comparison between bibliographic international models and our model.

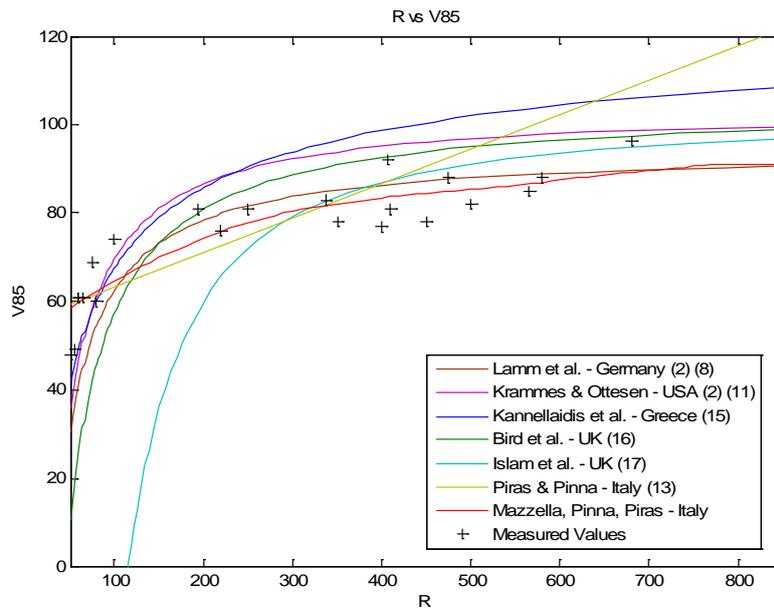


FIGURE 10 Comparison between international models and interpolation result for the V_{85} vs R

The Figure 10 shows a similar trend between the proposed model and most of those in use in various countries, especially for values of the radii between 150 and 700 m. Besides the difference between our model (with the lowest) and the Kannellaidis one (with the greatest values) is almost equal to 20 km/h.

Model $V_{85}=f(R, L)$

The calculation of the experimental semivariogram of the data collected provided the results shown in the Figure 11.

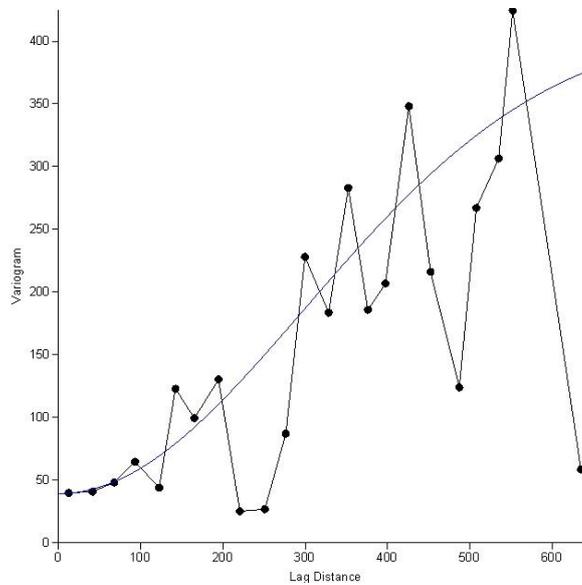


FIGURE 11 Experimental semivariogram for V_{85} vs R and L and its interpretation

The modeling was conducted by selecting the best fitting theoretical model to describe the general trend and to further estimate variographic parameters using least square criterion. The Figure 12 represents the Kriging result of the data carried out using parameters of variographic calculated models.

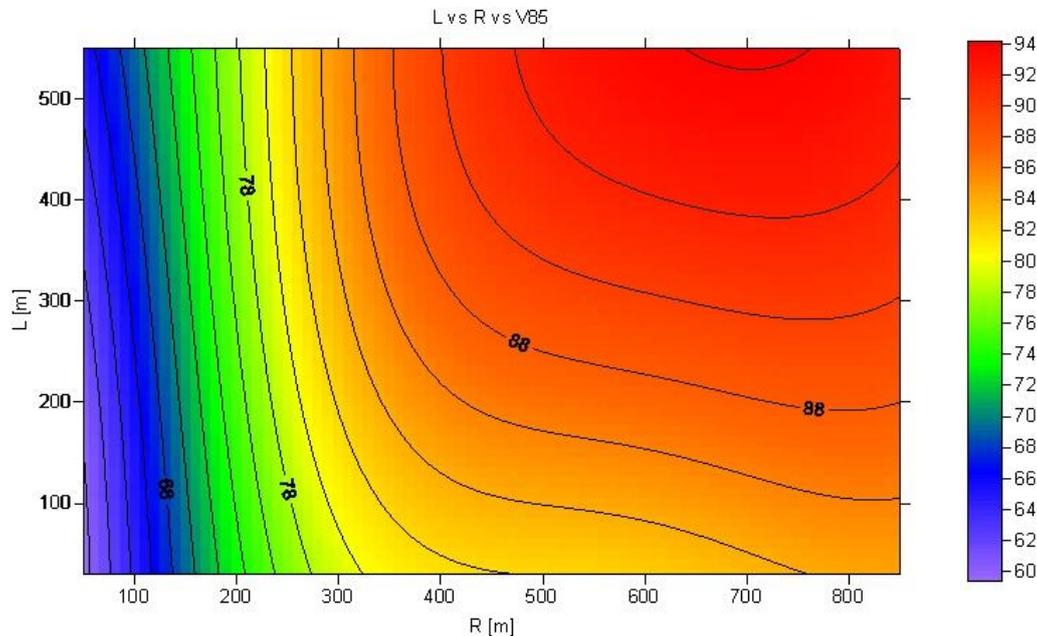


FIGURE 12 Kriging interpolation result for the V_{85} vs R and L model

The graph clearly shows the general V_{85} trend in variation to the radius R values and the length of the circular curves. In the image it is clearly shown the link between V_{85} , R and L.

CONCLUSIONS

There are numerous models in literature to calculate V_{85} in function of the rural road geometric characteristics, such as CCRs or horizontal curve radius. Almost all of these models were constructed using techniques own of Classical Statistics. No research group has still used Geostatistical analysis techniques even though they are largely used in many other fields of research inherent to Engineering sciences.

Some similarities can be seen when confronting the general trend of proposed models (using CCRs or R as independent variables) with those of respective models in use on an international level (created with classical Statistics techniques) (Figure 7 and 10). This clearly indicates a high implication of the validity of our proposed models, especially considering the fact that they are the first results using the geostatistical approach. In order to improve the models proposed in this paper we are still working on data collection on horizontal curves with radii greater than 700 m.

Using the correct study and modeling of such correlation, interpolations can be created which are based on the true variable of the phenomenon in exam where precise observations are obviously represented by exact demonstrations within the dominion of the study.

Geostatistics analytical instruments used correctly and in combination with the potentiality of

the Kriging algorithm, can create interpolations where the estimated variance is minimal (19).

In the current literature we find that many studies created models linking operating speed on curves to the horizontal curve features, or more precisely CCRs, radius of the horizontal curve, DC, etc. Even though it has been demonstrated that longitudinal grade or road alignment (tangent or curve) preceding the curve influence the speed in curves, there have been few studies that analyze such linking factors (2). For this reason, our research is leading towards the analysis of the relationship between driver behavior and other parameters, such as longitudinal grade, with the aid of geostatistical techniques. We are also in the process of defining a provisional 2D model where V_{85} is calculated by geostatistical analysis, contemporarily in function of R or CCRs and the longitudinal grade.

Furthermore, Geostatistics offers several analytical instruments that could be able to reveal the influence of the speed trend of vehicles going in the same direction for each section of study, meaning the element preceding the circular curve.

On the whole, the proposed work describes a new approach to study driver behavior in relation to the characteristics of the horizontal alignment, going beyond the limitations of classical statistics that:

1. chooses *a priori* regression type;
2. for the interpolation does not consider the true correlation between the sampled data.

REFERENCES

1. American Association of State Highway and Transportation Officials (AASTHO). *A Policy on Geometric Design of Highways and Streets*. AASTHO, Washington, D.C., 2001.
2. Lamm R., Psarianos B., Mailaender T. *Highway Design and Traffic Safety Engineering Handbook*. McGraw-Hill, New York, U.S.A., 1999.
3. Matheron G. *Principles of Geostatistics.*, Econ. Geol. 58, 1963 pp.246-1266.
4. Krige D.G. *Geostatistics and the definition of uncertainty.*, Inst. Min. Met. Trans., Vol. 93, Sect. A, pp. A41-47., 1984.
5. Journel A.G., Huijbregts C.J., *Mining Geostatistics*, Academic Press, 1978, pp. 600.
6. Lamm R., Choueiri E.M., Hayward J.C., Paluri A. *Possible Design Procedure to Promote Design Consistency in Highway Geometric Design on Two-Lane Rural Roads*. Transportation Research Record No. 1195, 1988.
7. McLean J.R., Morrall J.F. *Changes in Horizontal Alignment Design Standards in Australia and Canada*. International Symposium on Highway Geometric Design Practices, Transportation Research Board, Boston, Massachusetts, U.S.A, 1995.
8. Lamm R., Hiersche E.U., Mailaender T. *Examination of the Existing Operating Speed Background of the German Guidelines for the Design of Roads*. Institute for Highway and Railroad Engineering, University of Karlsruhe, Germany, 1993.
9. Lamm R., Choueiri E. M. *A Design Procedure to Determinate Critical Dissimilarities in Horizontal Alignment and Enhance Traffic Safety by Appropriate Low-Cost or High-Cost Projects*. Report to the National Science Foundation, Washington, D.C., U.S.A., 1987.
10. Lamm R., Choueiri E. M., *Rural Roads Speed Inconsistencies Design Methods*, Research Report for the State University of New York, research Foundation, Parts I and II, Albany, N.Y., U.S.A, 1987.
11. Ottesen J.L., Krammes R.A. *Speed Profile Model for U.S. Operating-Speed-Based Consistency Evaluation Procedure*. 73rd Annual Meeting of Transportation Research Board, Washington, D.C., U.S.A., 1994.
12. Choueiri E.M., Lamm R., Choueiri B.M., Choueiri G.M., *An International Investigation of Road Traffic Accidents*, Proceedings of the Conference Road Safety in Europe and Strategic Highway (SHRP), 1995.

13. Piras C. and Pinna F., *Influence of the plano-altimetric road characteristics on driver behavior: The operating speed.*, PhD Thesis, Department of Land Engineering, University of Cagliari, Cagliari, Italy, 2008.
14. Ng J., Sayed T. *Quantifying the Relationship Between Geometric Design Consistency and Road Safety*. Canadian Journal of Civil Engineering No. 31, 2004.
15. Gibreel G.M., Easa S.M., El-Dimeery I.A. *Prediction of Operating Speed on Three-Dimensional Highway Alignments*. Journal of Transportation Engineering No. 127, 2001.
16. Bird R.N., Hashim I.H., *Operating Speed and Geometry Relationships for Rural Single Carriageways in the UK*. 3rd International Symposium on Highway Geometric Design, Chicago, U.S.A., 2005.
17. Islam N.M., Seneviratne P.N. *Selection of Highway Design Parameters in the Presence of Uncertainty*, Transportation Research Record ISSN 0361-1981, Washington, D.C., 2007.
18. Bonneson J. A. and Pratt M. P., *A Model for Predicting Speed along Horizontal Curves on Two-Lane Highways*, Transportation Research Board, 88th Annual Meeting, Washington, D.C, January 2009.
19. Mazzella Al., Valera P., Mazzella An. *Kriging Assistant: an innovative and automatic procedure for geostatistical analysis of environmental data.*, European Congress on Regional Geoscientific Cartography and Information Systems 2009, Vol. I, pp.85-88.