At the same time, there is a tendency to avoid establishing precise design criteria in the Green Book. Instead, ranges of design values are provided, which affords the design engineer greater flexibility in selecting the design features of a roadway. This design flexibility may be viewed as a two-edged sword. It allows the design engineer to be innovative and provides freedom to exercise discretion. However, the plaintiff in a tort lawsuit can present alternative designs that are claimed to prevent the tort-related accident. Both designs could satisfy the design criteria and guidelines of the Green Book. The jury faces a dilemma in trying to determine whether the original design was inadequate and therefore hazardous.

The best defense for a public agency is to document the decision-making process when selecting the design for new or reconstructed roadways. If the design does not comply with the Green Book, it is imperative that the reasons for noncompliance be explained and documented for use as potential evidence. Because a multitude of potential designs can be developed in accordance with the Green Book, it is important that the discretionary decisions made by the design engineer also be documented. This documentation could provide the primary evidence necessary to successfully defend a future tort lawsuit.

REFERENCES


Publication of this paper sponsored by Committee on Geometric Design.

New Approach to Geometric Design of Highways

MOSHE HIRSH, JOSEPH N. PRASHKER, AND MOSHE BEN-AKIVA

A basic deficiency in the current practice of geometric highway design is a lack of sensitivity to traffic volume, traffic composition, and construction and user cost factors. Current practice is based on a deterministic approach, whereas the factors involved in the geometric design process (e.g., speed, friction, reaction time) are stochastic in nature and vary among road users. The current approach employs only a single value to represent each factor. Criteria that are used to generate these representative values are not made explicit. An alternative approach to geometric design of highways is presented in which sensitivity to the stochastic nature of the various factors involved in the design process and utilization of their distribution are used in calculating design values. The proposed approach also attempts to achieve a cost-effective design by taking into account all the cost elements associated with the highway. An empirical example of a horizontal curve demonstrates the advantages of the probabilistic approach.

This paper is concerned with the concepts used in the geometric design of highways. A modified approach is proposed that would achieve more meaningful and cost-effective designs.

Current geometric design practice is based heavily on design standards and the following basic design process is used. First the highway section to be designed is classified into one of the several functional classes (e.g., freeway, arterial, local). Then a design speed is selected for the highway on the basis of its classification and local conditions. After highway classification and design speed have been specified, design values for the various highway elements are selected from a set of predefined design standards (1–3).

This design practice has two major advantages. First, the
design concepts are transparent. This enables highway engineers to be trained easily and quickly. Second, this practice supports consistent design. For example, geometric elements of freeways in different locations designed for 100 km/hr will have the same design values.

The foregoing practice, however, has been subjected recently to increased criticism. It has become evident that such a practice is rigid and does not allow the designer to use his own judgment in special cases in which deviations from the standards are clearly justified (4). The practice is not always sensitive to important factors such as traffic volume, construction cost, and traffic composition. That is, once the highway class and the design speed have been selected, the minimum design value of a horizontal curve, for example, is fixed, regardless of the traffic volume that will use the road and the costs associated with implementing the design standard. Thus, a horizontal curve on an arterial road designed for 80 km/hr and serving a very low traffic volume in mountainous terrain will have the same minimum design value as a horizontal curve on a highway that serves a high traffic volume in level terrain.

Another criticism is that inflexible design standards tend to be based mainly on safety considerations, which results in excessively high design standards in many situations (5). For example, in the design of a vertical curve, the relevant inputs are the driver’s reaction-perception time, the speed, the friction factor, and the driver’s eye height. The design standards specify safe values for all these factors. For example, a reaction-perception time of 2.5 sec is used because it is valid for a large percentage of the population. Thus, the sight distance based on these values would result in a costly design.

Still another criticism addresses the concept of the design speed, which is the basis of the current design practice. The design speed is defined as the speed of the 85th-percentile driver in the speed distribution. However, it is not always clear which distribution is applicable, because there is a tangent speed distribution, a curve speed distribution, and speed distributions for cars and trucks. A recent change in design practice has been the replacement of the design speed concept with the concept of a consistent design (6, 7), which is also based on the 85th-percentile driver. In this approach it is desired to limit the maximum value of speed changes that the 85th-percentile driver experiences. However, it has been shown (8) that the 85th-percentile driver in the speed distribution found for a tangent highway section is not the same as that in a curve speed distribution. Moreover, it is theoretically possible for the speed distributions on a tangent section and on a following curve to be identical, which implies a consistent design, yet all drivers may experience significant speed changes.

The use of a design speed has also been criticized in other studies that claim that it is irrelevant in specific cases, especially in the case of a horizontal curve (9, 10). In the latter case, the determination of the design speed value for a curve with a given radius and superelevation rate is based on the value of the maximum superelevation rate. Thus, the same horizontal curve (i.e., the same radius and superelevation rate) may have various design speed values based on various values of the maximum superelevation rate.

The foregoing deficiencies reflect the deterministic approach adopted in these design processes. Although all the factors involved in the geometric design process (e.g., speed, friction, reaction time) are stochastic in nature and are distributed among road users, the current approach is based on a single, arbitrarily chosen value to represent each factor. The failure to account for the stochastic nature of these factors is likely to lead under some circumstances to poor designs. That is, in some cases the combination of those deterministic values that are chosen out of the distributions may not be representative of the road user population, and hence may result in an unsuitable design of a highway section.

In summary the current design process is not sufficiently sensitive to traffic volume and composition and does not explicitly consider the variation of cost factors among different locations. Also, because of its deterministic nature, it may lead to a poor design. These problems have been recognized by highway engineers for some time and several solutions have been proposed. In order to overcome the excessively high construction costs that often result from adherence to general design standards, alternative design standards have been developed for special cases such as low-volume roads (11), low-cost roads, and roads in developing countries (12). These efforts justify the use of standards that are lower than the usual ones but continue to employ the deterministic design process that is based almost entirely on road classification and design speed.

A modified approach to geometric design of highways is presented here. This approach is fully sensitive to the specific conditions of the design problem at hand, that is, to traffic volume and composition and to construction, maintenance, and user costs. It is also based on the stochastic nature of the various factors involved in the design process. The proposed approach is an attempt to obtain an optimal or a cost-effective design that takes into account all the cost elements associated with the highway. In this approach the designer is made aware of the economic and safety implications of alternative designs.

The proposed approach is presented along with an empirical example. The differences between the current and the modified approach are demonstrated, and the potential of a future development in the proposed direction is discussed.

THE PROPOSED APPROACH

The first stage in the proposed approach is to obtain information about traffic volume, traffic composition, driver performance, and vehicle characteristics. The relevant information includes reaction-perception times, speed distributions at various highway locations, and vehicle dimensions and characteristics such as friction factors. It is recognized that precise information for a specific site may not be available. However, empirical studies may be used to obtain reasonable distributions for the relevant parameters. It may be noted that in the current design approach the designer uses empirical values determined in other studies to design new highways.

The major difference, however, between the current approach and the proposed approach is that the current approach requires a single deterministic value for each parameter and the proposed approach utilizes the full distribution of parameter values. For example, in order to determine the required design value (i.e., the standard) of sight distance for a design speed (V) of 80 km/hr, the current approach assumes...
that the reaction-perception time \( t \) is 2.5 sec, the friction-factor value \( f \) is 0.2, and thus the required sight distance \( S \) is 80 m, which results in the following expression for stopping distance:

\[
S = Vt + \frac{V^2}{(2gf)}
\]  

(1)

where \( g \) is the gravity acceleration constant. The proposed approach assumes that the distributions of the travel speed, the reaction-perception time, and the friction factor are known. On the basis of many empirical studies, it is possible to assume these distributions when observations are not available. It may be noted that these parameters are not independently distributed. The relevant factors are often correlated with the travel speed. For example, the friction-factor value decreases with increasing travel speed, and the perception-reaction time may also.

The second stage of the proposed approach is the determination of the relevant physical or behavioral relationships, or both, from which the design value may be calculated. In the current approach, these are the various design equations that relate the design value to the design speed, as in Equation 1. As was mentioned earlier, however, more than one relationship may determine the design value of a specific element. For example, one may design a horizontal curve to satisfy the dynamic forces acting on the vehicle, and thus use the following relationship:

\[
R = \frac{V^2}{g(e + f)}
\]  

(2)

where \( R \) is the curve and \( e \) is the superelevation rate. On the other hand, one may wish to have a consistent design such that

\[
\text{Max}(V_t - V_e) = 15
\]  

(3)

where \( V_t \) is the travel speed on the tangent section and \( V_e \) is the speed on the curve.

The variables on the right-hand side of the various relationships will be referred to here as the input parameters. Thus, the purpose of the first stage in the proposed approach is to determine the distributions of the input parameters. Once the input parameters are available, these distributions can be incorporated into the design relationships to get an output distribution for the parameters on the left-hand side of the various relationships. In the current approach, the resultant parameter on the left-hand side is the desired design value, or the standard. In the proposed approach, however, a full distribution of possible design values is obtained, out of which only one may be selected. Thus, the variables on the left-hand side in the proposed approach are referred to here as the intermediate variables. The analytical determination of the derived distribution of an intermediate variable is not a simple task. Consider, for example, the sight-distance relationship in Equation 1. Even if all the input parameters are independent and normally distributed, the distribution of the stopping distance is not normal. If the intermediate variables were a linear function of the input parameters, the distribution would also be normal. [An example is the application to the design of a climbing lane by Ben-Akiva et al. (13).] In other situations it is always possible to apply a numerical approach to determine the derived distribution.

To demonstrate the ideas presented so far and the feasibility of applying a numerical approach, the following empirical example for the design of a horizontal curve may be considered. The relevant physical relationship that governs the curve design was given in Equation 2.

In this example, the maximum superelevation rate used is 6 percent. It is assumed that the speed distribution can be approximated by the normal distribution with a mean value of 50 km/hr, and two values for the standard deviation are compared: 25 km/hr (e.g., heterogeneous traffic, which may include slow trucks and fast cars) and 10 km/hr (e.g., more homogeneous traffic). By using these two speed distributions, a sample of 1,000 vehicle speeds was randomly drawn from each distribution, and the resultant two distributions are shown in Figure 1.

The side friction factor has been shown in many studies to be highly correlated with the travel speed (14). It is assumed that the side friction factor is normally distributed; that is,

\[
f \sim N(\mu_f, \sigma_f^2)
\]  

(4)

On the basis of an empirical study reported recently by Lamm (14), it is assumed that

\[
\mu_f = 0.37(0.0000214 V^2 - 0.0064 V + 0.77)
\]  

(5)

where \( V \) is the travel speed in kilometers per hour. The standard deviation of the distribution is assumed to be 0.0555. For each vehicle in the sample drawn from the speed distribution a friction-factor value based on its speed is now determined. A value was randomly drawn from a normal distribution with the mean value given by Equation 4 and a standard deviation of 0.0555. The resultant friction-factor distributions for the two speed distributions are presented in Figure 2.

It may be noted that the percentile values in Figure 2 indicate the percentage of the population for which the given friction factor or less is applicable. However, the geometric design process considers the percentile of the population for which at least a given value of the friction factor is applicable. Thus, a friction-factor value that covers 90 percent of the population corresponds to the 10th percentile in Figure 2.

Each vehicle in the sample has now been assigned both a travel speed and a friction factor. By using Equation 2 the minimum horizontal curve radius required by that vehicle may be calculated. Because each vehicle has a different speed and side friction factor, each vehicle also requires a different minimum curve radius. Figure 3 presents the curve radius distribution for the two samples for the two speed distributions. As may be expected, the wider speed distribution results in a wider radius distribution.

Figure 3 shows, for example, that 90 percent of the drivers for the relatively homogeneous speed distribution may be satisfied with a curve radius of 150 m, whereas a curve of 265 m is required to satisfy 90 percent of the drivers under the heterogeneous speed distribution. The corresponding radii for 85 percent of the population are 135 and 220 m, respectively.

At this point it may be useful to compare the radius distribution in Figure 3 with the design standard that may result from
the current approach. As was mentioned earlier, the design speed is usually the speed of the 85th-percentile driver. Thus, the narrow speed distribution results in a design speed ($V_1$) of 60.2 km/hr and the wide speed distribution results in a 73.4-km/hr design speed ($V_2$). In order to calculate the value of the minimal radius, a friction-factor value must be assigned to each design speed. For that purpose the relationship in Equation 4 is used, which gives the mean value of the friction factor. These values for Equation 4 are 0.17 and 0.15 for $V_1$ and $V_2$, respectively. However, for design, safe values of the friction factor are necessary rather than mean values. Because the standard deviation of the friction factor distribution is assumed to be 0.0555, the 85th percentile is approximately the mean value minus one standard deviation, and the 90th percentile is approximately equal to the mean minus 1.3 times the standard deviation. Thus, the 85th-percentile friction factors are 0.115 and 0.097 for $V_1$ and $V_2$, respectively, and the corresponding 90th-percentile values are 0.099 and 0.081 for $V_1$ and $V_2$, respectively. The 90th-percentile friction factors yield radii of 179 and 301 m for $V_1$ and $V_2$, respectively. Incorporating the design speed value and the 85th-percentile friction-factor value into Equation 2 yields curve radii of 162 and 269 m for $V_1$ and $V_2$, respectively.

With Figure 3 it can now be determined what percentile of
the road user population is satisfied by these design values. In the first case, where the 85th-percentile values of the friction factor are used, the corresponding percentiles of the population that are covered by the design standards are 93 and 90 percent for \( V_1 \) and \( V_2 \), respectively. In the second case, where the 90th-percentile values of the friction factor are used, the percentiles covered by the standards are 95 and 93 percent, respectively.

Thus, a design speed of 73 km/hr and a corresponding 85th-percentile friction factor result in a standard that covers 90 percent of the population, whereas a design speed of 60 km/hr with an 85th-percentile friction factor result in a standard that covers 93 percent of the population. The results show that using the 85th-percentile values of the speed and friction-factor distributions may allow the derivation of a design standard that covers 93 percent of the population, which may be viewed as an overdesign, because it was intended to satisfy only 85 percent of all drivers.

Once the output-value distribution, as shown in Figure 3, has been established, there is a need to select only one design value from the distribution. To do so, a percentile criterion may be used that states that the single design value selected needs to satisfy the requirements of at least a certain percentage of the drivers. As will be restated later, there is no need to state a priori the value of the minimum percentile. Examination of the shape of the output-value distribution may lead to a reasonable selection of a design value.

The percentile criterion has an intuitive safety implication in that the higher the percentile value is, the safer the design is. However, it is obvious that there is no need to design the highway for the 100th-percentile driver. On the basis of the output-value distribution, the designer may determine the implication of a specific design alternative. For example, if the curve radius is constrained to be only 100 m, the designer may determine that such a design will satisfy the requirements of only 60 percent of the drivers.

Besides the intuitive meaning, another advantage of the percentile criterion is its sensitivity to the traffic composition. As may be seen from Figures 1 and 3, more homogeneous traffic volumes (fewer vehicle types and homogeneous driver performance) result in lower design values.

The disadvantage of the percentile criterion is the lack of sensitivity to the volume of traffic and cost considerations. Thus, another design criterion, called the minimum-cost criterion, is suggested here. An objective function is defined that expresses an expected total cost as a function of the design value, and the value that brings this function to a minimum is selected. The cost function has two components: the road user cost and the construction and maintenance cost. Each possible design value results in a different road user cost and a different construction and maintenance cost, and hence a different total cost. In general, a safer standard will imply savings in user cost and increased construction cost, and a chosen design value reflects a specific trade-off between these two cost components.

A cost-minimization approach has been used in past studies (15) to help in the selection of a cost-effective design among various discrete alternatives. The uniqueness of the criterion proposed here is that it accounts for the full range of all the possible design values and is linked with the distributions of the intermediate values. The importance of the second aspect will be explained shortly.

The user cost function includes accident costs, vehicle operating costs, and the value of the driver's time. These cost elements are dependent on the selected design value and on the distributions of the intermediate values. Accident costs, for example, may be dependent on the difference between the design value required by a driver and the selected design value. The same applies to the fuel consumption cost, which is partly dependent on the level of the speed changes that are imposed on the drivers by the chosen design value. To calculate the total cost, it is necessary to sum these values for a heterogeneous driver population.

The other component of the total cost is the construction and
maintenance cost. This component depends only on the selected design value and not on the input-value distributions.

Formally, the cost criterion may be defined as follows. Let $X$ be the design value of interest and denote by $x$ the values of the various input and intermediate variables. Let $f(x)$ be the probability density function of $x$, $C(X,x)$ the road user cost function, and $I(X)$ the construction and maintenance cost. The objective function may be written as

$$\min \int C(X,x)f(x)dx + I(X)$$

If desired, the objective function may be subjected to various design constraints.

To demonstrate the concept behind this criterion, the problem of selecting an optimal design value for a horizontal curve $R$ is reconsidered. Let $r$ be an intermediate variable; that is, $r$ is the curve radius required by each driver. For simplicity, the perfectly homogeneous case, in which all road users are identical, is used and $r$ therefore takes a single value. It may be noted that this assumption is the one used in the current design practice; that is, all road users require the same design value. In other words, in this example $r$ may be defined as the design standard. The road user cost may be defined as

$$C(R,r) = \begin{cases} N \cdot b_1 (r - R)^2 & \text{for } R < r \\ 0 & \text{for } R \geq r \end{cases}$$

in which $b_1$ is the accident parameter cost and $N$ is the number of vehicles using the road. For simplicity $B_1$ is defined as equal to $N\cdot b_1$. In this cost function, the only relevant element is assumed to be the accident cost. Thus, when the selected design value $(R)$ is smaller than the intermediate value $(r)$, which is needed by road users, accidents may occur. If the selected design value, however, is equal to or greater than $r$, the accident costs are assumed to be zero.

The construction cost function may be approximated by

$$I(R) = b_2 R^2$$

Here the construction cost is also represented by a nonlinear function with a parameter $b_2$, and the cost is considered to be mainly right-of-way costs. The total cost function is the sum of Equations 7 and 8. It is evident that for the case of $R \geq r$, the optimal design value is $R = r$. Thus, the optimal design value is in the range $R \leq r$ and the following objective function is derived:

$$\min [B_1 (r - R)^2 + b_2 R^2]$$

The optimal solution is then given by

$$R_{opt} = r/[1 + (b_2/B_1)]$$

The conditions in Equation 10 indicate that the optimal design value is related to the intermediate variable $r$ by the ratio of the parameters of road user cost to construction cost. Unless $b_2$ equals 0 or $B_1$ is infinite, the optimal design value $(R)$ will always be less than the intermediate value $r$. It may be recalled that $r$ is in fact the current design standard. Thus, implicitly, the current design practice ignores the trade-off between construction and accident costs. This may be a reasonable assumption to make in level terrain, but not in mountainous terrain where the cost of earth movement is a very significant factor. However, current design practice evidently does not completely ignore this trade-off with construction cost. The deterministic values selected are high (or safe) percentile values, but do not cover the 100th-percentile driver. Thus, the selection of a specific percentile value less than 100 is an implicit recognition of this trade-off. The optimal cost criterion, however, makes the trade-off explicit.

The optimal cost criterion may also be applied numerically. In this case, it was assumed that the 1,000 vehicles sampled for the heterogeneous speed distribution make up the average daily traffic volume that traverses the curve. For the cost function whose components are given in Equations 7 and 8, the following parameter values were assumed: $b_1$ lies in the range of $0.0001/(day \cdot m^2)$ to $0.0004/(day \cdot m^2)$ and $b_2$ has a value of $0.05/(day \cdot m^2)$, which corresponds to a construction cost of $200/m. The total daily cost for various design values in the range of 50 to 300 m was calculated. Figure 4 presents the total cost as a function of the design value for different $b_1$-values. It may be seen that as the value of $b_1$ increases—that is, as the importance of the accident costs increases—the value of the optimal design value also increases. It may be noted, however, that the optimal value for $b_1 = 0.0002$, for example, is only 140 m, which meets the requirements of only 70 percent of road users. These results show that it may be cost effective to reduce the standard values.

DISCUSSION OF RESULTS

A new approach to geometric design of highways has been presented that utilizes the full distribution of input parameters and attempts to achieve a cost-effective design.

However, there are several problems associated with the implementation of the approach. First, to be sensitive to local conditions, the approach needs the appropriate input-value distributions. In some cases this may call for an extensive data collection effort. Because many parameters may be correlated (e.g., speed and reaction time), such a data collection effort is not a simple task. Second, once the appropriate input-value distributions have been established, the relevant output-value distribution must be derived. Analytical derivation of the output-value distribution is always the preferred approach. However, for many highway design problems, the analytical derivation is a very complicated task and a numerical approach is suggested instead.

Another problem is the construction of an appropriate cost function. First, there is a need to identify all the relevant cost components. The conventional road user cost components are accident costs, vehicle operating costs, and the value of travel time. In some instances, some of these elements may not be relevant (e.g., the value of time). Also, even when the total cost components are known, there is still a need to assign a monetary value to such elements as accidents and value of time,
which are difficult to quantify. Because highways are designed to serve for several years, there is a need to forecast the value of the various parameters associated with the design process, such as future volumes and costs. The proposed approach involves many factors, and this may introduce additional uncertainty into the design process. As a result, the selected single design value may not be optimal or cost effective. A possible solution to this problem is to perform a sensitivity analysis by varying the values of the various parameters. An example of such an analysis has been given elsewhere (13) for the case of climbing-lane design.

Another methodological problem is the interdependency that may exist between the input variables and the selected design value. In the case of a horizontal curve, for example, it was found that drivers adjust their speed and the level of the accepted lateral acceleration rate to the curve radius (16). The design process should be able to account for this phenomenon as well. A related issue is the three-dimensional aspect of the design. The methodology presented in this paper demonstrates the design process for a single highway element. However, highway elements are interrelated in a three-dimensional system. An optimal design should take into account all the highway elements in a given section.

The last problem discussed here is the incorporation of the new approach into practice. On a day-to-day basis, it is not practical to conduct an extensive study each time there is a need to design a highway section. Rather, it is desired to be able to upgrade the current design standards to include the features of the new approach. A simple improvement would be to introduce additional factors into the specification of a design standard. For example, the travel speed may be represented by two parameters—mean and standard deviation—and a design value may be calculated for a range of percentile values. The volume of traffic may also be introduced and a range of optimal design values may be presented for each combination of speed distribution and traffic volume.

REFERENCES


*Publication of this paper sponsored by Committee on Geometric Design.*