The Access and Development Prototype Project

MORTON SCHNEIDER, Creighton, Hamburg Inc.

•THIS PAPER is divided into two main parts. The first is a summary sketch of a project (1) based on a paper (2) presented at the Conference on Urban Development Models held at Dartmouth College in June 1967. (The conference was sponsored by the U.S. Bureau of Public Roads, the U.S. Department of Housing and Urban Development, and the Automotive Safety Foundation and conducted by the Highway Research Board.) The second elucidates certain aspects of the original paper, adding new theoretical material. The two parts do not have a great deal to do with each other.

THE PROTOTYPE PROJECT

The paper, Access and Land Development, attempts to show some general and computable relationships among travel, land development, and the time and cost characteristics of transportation. The purpose of this project was to develop a computer program capable of performing the calculations proposed there, partly to see if they could be performed, and partly to see if, having been performed, they appeared to mean anything. The computer program developed in the project is a prototype, which is to say that it does not try to accommodate the full range of possibilities that might be found in the real world. Its design—and prototype character—is due mostly to the decision to avoid, at this stage, the distracting, difficult, and expensive problems associated with realistically complicated transportation networks.

Review of Theory

The general sense of the theory can be seen in what it does, in the information on which it operates, and the results it produces. For a region of interest, the basic items of given information are (a) the total amount of floor area to occur in the region; (b) the transportation network, that is, the travel times and travel costs of all means of transportation that connect places with each other; (c) a measure of the inherent attractiveness of land—essentially just proportional to area, but also distinguishing, for example, fetid swampland from firm, salubrious acreage; and (d) any constraints, planned or expected to operate on individual sites, limiting or compelling their development—park land reservations, perhaps, or the continued existence of structures already in place.

The principal quantities calculated from these are (a) the amount of floor area that can be expected naturally to arise at each site in response to the access there, consistent with development at all other sites and with any constraints; (b) the number of trips that can be expected to originate at each site; and (c) the number of trips that can be expected to use each element (of all modes) of the transportation system.

The size of the region of interest, the fineness of partitioning into "sites" or "zones," the inclusiveness and detailing of the transportation system are all simply matters of precision and the mechanics of implementation as far as the theory is concerned. There is no innate specialization either to urban areas or to great regions.

The reasoning behind the theory proceeds in three more or less distinct phases. The first works with quite general considerations to show that trip generation at any site is related to the environment in which the site is embedded. Here the concept of "attractive stuff" is introduced as a necessary attribute of sites; it is called R because it has to be called something, but the physical meaning is left undetermined.

Paper sponsored by Committee on Land Use Evaluation.

Using only the assumption that travel can be described by a symmetric (but otherwise unspecified) trip-distribution function, the result is derived that the number of trips originating at any site is proportional to the amount of R at the site and to a quantity that can very naturally be thought of as the accessibility of the site, or the access integral around it (2, Eq. 4, p. 65).

The second phase takes up the ancillary question of the exact form of the trip-distribution function and, especially, the problem of mode of travel. It proposes that all trips do not have the same sensitivity to travel time and cost and that, because of this, there can be more than one effective "minimum path" between two points. Compounded into a mathematical treatment, this gives a definite all-mode trip distribution function (2, Eq. 11, p. 70).

The third phase identifies attractiveness, R, as the sum of two components, one due to land area and the other to floor area, where floor area is used as a convenient surrogate for capital improvement of all kinds. It then argues that floor area appears at a site in order to accommodate activity there, but generates new activity by its appearance. However, if total activity is governed by the trip generation relationship already given, there will be some definite amount of floor area just right for the land area (attractiveness) and accessibility of the site (2, Eq. 20, p. 173).

The Prototype Computer Program

The inputs and outputs of the prototype program are pretty much those listed above except that there is no link-by-link traffic volume output, but, instead, some trip-distribution tables. The transportation network is limited to a highly stylized three-moded system, designed around computational convenience. Each mode may be assigned any speed and any cost per mile, and any initial time and cost penalty. The modes compete but do not interconnect. They are assumed to operate along grid lines, but a grid cell need not be connected to all three modes.

In other respects, the prototype is a fairly usable model. It is organized on a regular grid system, in which a zone is simply a grid square. It will handle up to 2,500 such zones in any arrangement that fits within a 50 by 50 square. The scale of the grid system is arbitrary, so that a zone can be one mile on a side, or one-tenth of a mile, or ten miles—or any other size—with the dimensions of the whole region correspondingly expanded or contracted. Any zone can be constrained to have any given amount of floor area development, or less than, or more than any given amount.

The program is in FORTRAN IV. It has been operated successfully on both a Control Data 3800 computer and an IBM 360/50. The running time is roughly proportional to the square of the number of zones (on the CDC 3800, something like 40 seconds for a 230-zone case).

An output format worth mentioning is a supplemental maplike printout intended to give a quick perception of the calculated floor-area development pattern. It does.

Plan of Tests

If the theory were perfectly credible (or the world perfectly credulous)—that is, if there were no question of its validity—there would still be several steps necessary to make a going thing of it besides writing a computer program. These steps establish a very rough framework for exercising the prototype program to achieve some feeling for the behavior and validity of the whole idea.

The theory leaves three basic quantities undetermined, subject to empirical measurement. Two of these are the constants "a" and "b" in the trip distribution function. These constants may be thought of as the average values of time and cost respectively in trip-making. As "a" grows smaller, trips become willing to take more time, and as "a" grows larger, less. In just the same way, trips tend to spend more money when "b" is small and less when "b" is large. The magnitudes of "a" and "b" together determine average trip length and the distribution of trip lengths in a given system. But the relationship of "a" and "b" to each other determines the distribution of trips among competing paths, or modes. So it should be possible to estimate "a" and "b" through these two properties of length and mode distribution by comparing calculations with actual trip lengths and modal splits.

The third unfixed quantity is the relative scale of attractiveness due to land area, R_a, and that due to floor area, R_f. The units of R can be arbitrarily chosen, just as a foot is an arbitrary unit of length, but how much R resides in a square foot of land area and how much in a square foot of floor area is up to empirical measurement. If, for convenience, one unit of R is defined to be the amount of R produced by one square mile of lând, then the empirical problem becomes that of finding the rate of R for floor area. This is a rather more subtle problem than that of estimating "a" and "b," but it can be solved by making use of the not too obvious property of the theory that the pattern of floor-area development changes with a change in total R_F. In general, concentration effects in a given system increase with total development (as indeed seems to be the case in the real world). The procedure for scaling R_f would be to calculate the distribution of R_f's for some real region in successive trials, using different values of total R_F, until the shape of the R_f distribution agreed with the shape of the actual floor-area distribution—that is, until the ratios to each other of R_f's around the region were the same as the corresponding floor-area ratios, or, the same thing, the ratio of R_f to the floor area it was trying to represent was the same everywhere. The latter ratio, of course, is the sought-for conversion factor.

One further complication is that it can be difficult, though not strictly impossible, to calculate well trip distributions in order to estimate "a" and "b" without knowing how to measure R, and it is impossible to calculate floor-area distributions well, in order to measure R, without good values of "a" and "b." A certain amount of successive refinement is involved.

These are the things that would be done if things could be done that way. In practice the prototype tests, although they have this general scheme in mind, suffer from an exiguity of appropriate data and from their own ineptitude at representing complicated reality, and tend to limp and stumble from case to case. It is hard to know how to report the results of this kind of project in capsule form. The program works. Sometimes it is even fun. The mathematics make computational sense and look as though they might be reflecting the real world. Calculations exhibit agreement with data, but there are few points of contact. The approach to modal split seems to work well, but the trip-distribution function may be having some trouble in its ambitious effort to comprehend all trip behavior from the very short to the very long in one function. Interestingly enough, the scale of R_f seems to turn out to be about the same as that of R_a; a square mile of floor area apparently contains about one unit of R, just as does a square mile of land area.

On the whole, the idea looks promising to the intuition, and not too much more than that can be expected at this point. Perhaps its most engaging feature continues to be theoretical richness, the fact that it makes predictions, as opposed to statistical recapitulations, and focuses the viewpoint (distorting it at the same time, possibly). Two very general predictions deserve pointing out: if the quality and extensiveness of transportation does not change much, centralization will occur as the world accrues more and more fixtures of development; if transportation becomes more extensive rapidly enough, decentralization will occur.

It might be mentioned in passing that the theory does not particularly require every square foot of floor area to have the same R value as every other square foot, or every square foot of land area to be the same as every other. But it would be far beyond the resolving power of this project to distinguish differences of that order.

ELUCIDATIONS

The expression for floor area at a site is

$$R_f = \frac{R_a I R_F}{J - I R_F} \qquad (20)$$

where I is the accessibility integral around the site and J is the integral of these accessibilities throughout the region.

Regional "Temperature"

An examination of Eq. 20 reveals that the ratio J/R_F is a special quantity. Development distributions calculated separately for different regions (excluding constrained areas) have to give the same J/R_F in order to be consistent with the distribution obtained by lumping the regions together in a single superregion. The ratio can be formed for regions of any size at all, large or small, and will be the same for every place entering into a distribution. Presumably regions could have different J/R_F ratios if the interaction between them were weak, or if there were barriers or hidden constraints due to policies on travel, commerce, or immigration, but presumably also these differences would tend to diminish with time.

Indeed, this whole idea of land development can be restated in the form of a principle: the world acts to keep J/R_F the same everywhere within well-communicating regions. This has something of the ring of an entropy law to it, with J/R_F vaguely analogous to temperature. It says nothing about the dynamics of the process, only its goal. Carrying out the analogy, borders, oceans, and geographic and political barriers can be thought of as insulating walls slowing down or impeding the entropic process. The analogy is an attractive one, and may even be valid as far as it goes. But other important factors—things such as access to goods and resources, productivity, various kinds of costs—are not explicitly taken into account (although there seems to be theoretical elbow room for them to work their way in).

Continuing with the idea of regional "temperature," it is easy to let T stand for J/R_F and to rewrite Eq. 20 as

$$R_f = \frac{R_a I}{T - I} \tag{20a}$$

This form clearly suggests that if T is known, the distribution of development can be calculated without knowing the total development in advance: development becomes a matter of interaction between access at each point and the state parameter, T, instead of the distribution in detail of an exogenous total. Here T begins to function in a rather lofty role, trying to dictate the shape of the world. Presumably T changes with history, rising as the extensiveness and quality of transportation increases, though capable also of falling in a period of transportation stagnation combined with growth in total development. Almost certainly, T has been rising in the modern era (perhaps it has always been rising) and, from Eq. 20a, a site needs more access to get developed, or to stay developed, than it used to. Moreover, T increases especially fast when places of more than average desirability become accessible, another long-term tendency that works some hardship on fixed-access sites and on those that are less congenial.

The notion that T is some kind of fundamental system measure has considerable appeal to vagrant speculation. Perhaps deeper social, economic, and historical meanings can be found for it. Perhaps the dynamics of the development process can be understood in terms of flows of development governed by T differences from place to place. Perhaps T represents something to which a value can be attached in human events, so that a place of higher T may be said to be better off than one of lower T. But in a much more practical vein, the concept of T serves to make a connection among regions that could be treated together only in a system too big to handle comfortably: calculations done for areas all over the country can be made consistent with each other without having to stuff the whole continent into a computer at once.

The Dense Branch: Fiat Development

Given a region containing a transportation system but no development, there will be some value of J that may be called J*. If development is added uniformly everywhere, with the transportation system remaining fixed, it can be seen from definitions that J becomes, for any total development, R_F, and with total land, R_A:

$$J = (J^* - Z)(1 + R_F/R_A)^2 + Z(1 + R_F/R_A)$$

Here Z is a contribution to J* due to access from within the region to the world outside. (In a real, nonuniform case, J would increase more rapidly with R_F than this because of centralization effects.)

And T becomes:

$$T = \frac{1}{R_F}(J^* - Z)(1 + R_F/R_A)^2 + \frac{Z}{R_F}(1 + R_F/R_A)$$

This is a two-valued function, approaching infinity when R_F is either very small or very large, with a minimum at

$$R_F = R_A \sqrt{[J^*/(J^* - Z)]}$$

So starting with an empty system and knowledge of the world value of T, it appears that two quite different development patterns can be calculated, one at a lower total development and one at a higher. The lower density solution apparently is the proper one, representing normal incremental growth, and the higher solution can be interpreted as a kind of critical mass effect in which the greater development generates enough self-access to be stable. As Z, the contribution of the outside world, becomes large compared to J*, the minimum point separating the two branches of the T curve moves toward R_F, implying that the dense-branch solution for a small area well-connected to the rest of the world probably entails unreasonably high density. But for a substantially isolated region, the existence of the dense-branch solution offers the possibility that unexpectedly large self-sustaining development can occur, whether by fiat, special inducement, historical quirk, or otherwise.

Even assuming the general validity of Eq. 20a, the question of dense-branch development is an open one. It may very well be nothing but a whim of the mathematics. On the other hand, it may have real significance, especially when considering parts of the world, ancient or modern, having low values of T, which would allow smaller dense-branch development. Presumably dense-branch development could be accomplished suddenly by fiat through deliberate acts of political or economic power. It could also occur in more gradual evolution if T were to fall, go around the minimum, and rise again, dragging a city or two along with it. Conceivably the world might be made up of both low-branch and dense-branch development, with the distinction constituting a rather elegant definition of urbanization.

Everything else staying the same, the lower the low-branch solution the higher the dense-branch solution, and the greater the difference between them. The difference is mathematically capable of being small, but it looks as though it is commonly very large, possibly too large to be interesting, and growing all the time. It is intriguing to note that history records many examples of fiat development, most of which, like Karakorum, dwindled away soon after the fiat was removed. The two notable exceptions that come to mind are Constantinople and Alexandria—both enormously impressive in scale, even by modern standards. Perhaps it takes that kind of effort to make the great leap from low branch to dense branch, from normal village to hypernormal city.

Of course, both of these cities were well situated in the first place, unlike Karakorum, and their situations were improved in the course of founding by introduction of streets, roads, and shipping. This is a two-edged observation. On the one hand, it may mean that their low-branch solutions were fairly high and their dense-branch solutions fairly low, making them suitable candidates for bridging the gap. On the other hand, it opens the interpretation that no gap-jumping was involved at all, but that they are merely unusually dramatic instances of ordinary development.

Whether or not the dense branch has anything to do with fiat development, the founding of cities (or anything else) cannot really be decided without more and better numbers. In the absence of quantitative demonstration, it does not seem urgently needed to explain anything, yet there are things that it might very nicely explain.

The Cubic-Equation Form

From the theoretical point of view, which prefers to think of a site as indefinitely small and contributing virtually nothing to its own accessibility integral, Eq. 20 is the natural and lucid way to write the expression for floor area. There is another form, though, not so natural but still lucid in its own devious way, that turns out to be in general better behaved, and in some cases quite necessary for practical computing and numerical analysis.

Equation 20 is not a pure expression for R_f when dealing with sites or zones of finite size. R_f does not really stand isolated on the left of the equation, but occurs on the right side as well through its contributions to I and J. From the definition of I, it is easy to see that

$$I = I_0 + f_0R_f$$

where I_0 is I with the contribution due to R_f extracted, and f_0 is the intrazonal value of the trip-distribution function, its average value within the origin site itself; from the definition of J, it is possible (ignoring contraints, which complicate things somewhat) to work out that

$$J = J_0 + 2R_fI_0 + f_0R_f^2$$

where J_0 is J with the contribution due to R_f extracted. Putting these expansions of I and J back into Eq. 20, and picking through the notational shards for R_f, results in the cubic equation

$$f_0R_f^3 + R^2(2I_0 - f_0R_f) + R_f(J_0 - I_0R_F - R_af_0R_F) - I_0R_aR_F = 0$$

This actually represents a large set of simultaneous equations, of course, one for every site, with every R_f adding something to every I_0 and J_0. There is not much to be gained from trying to write out the solutions of such equations—each one can be solved numerically and a set of consistent solutions can be found by iteration. A certain amount of analyzing can be done, however, and, without pretending to say anything rigorously general or exhaustive about any large set of simultaneous cubic equations, the following properties appear to hold.

Each equation always has at least one entertainable solution, a root that is real, positive, and consistent with real and positive solutions for every other equation; often there can be two such roots, and sometimes three.

If only the smallest entertainable solutions are accepted for all equations that have more than one, the result is a unique set of solutions for the system—development at every site is uniquely determined. This appears to correspond to all real cases and is the policy adopted in the prototype program.

If an upper solution is used, there can be more than one development pattern depending on which site is chosen for the upper solution; these upper solutions entail concentrations of development at densities well above the point at which congestion effects and construction costs would decisively supervene in the present world. The physical meaning would seem to be that various stable development patterns involving extreme concentration could be achieved, ignoring congestion, if someone had the power to deliberately rearrange things that way. These upper solutions should not be confused with the dense-branch solutions discussed earlier. The contexts are entirely different; no doubt there is some kind of formal indirect relation, but no likely one worth pursuing.

In a finite region, with the transportation system remaining fixed, as more and more total development is added the lowest solution for the most central place will rise while the next higher solution will fall toward it. At some point these two solutions will meet, and then, if a little more development is thrown into the total, they will both abruptly become nonreal, leaving only a third, uppermost solution. Evidently a highly saturated region can become suddenly supersaturated, requiring the most central place to grow cataclysmically by sucking in development from all other places. Again, these effects occur at unfeasibly high densities. The physical meaning is problematical, but it is certainly a warning of some kind to the world.

It should be pointed out that solutions other than those in the "lowest entertainable" set are always of clouded significance, though it is not impossible that a meaning can sometimes be found for them. The treatment here is essentially numerical, making the assumption that intrazonal structuring does not exist and need not be considered. Obviously, the smaller the zone size, the better this assumption. But if Eq. 20 is taken to be a point function applying to infinitely small areas of infinitely small R_a, the cubic equation cannot even be developed for a point of finite R_f/R_a density, because such a point makes no appreciable contribution to I or J. In general, however, the three values of R_f as R_a approaches zero are the "proper" value taken directly from Eq. 20 (the limiting case of the lowest entertainable solution), and two other values, not necessarily real, positive, or meaningful, given by

$$\frac{1}{2f_0}\left[\left(f_0R_f - 2I\right) \pm \sqrt{4I^2 - 4f_0J + f_0^2R_F^2}\right]$$

The Quadratic Form

If Eq. 20a is subjected to the same kind of treatment that led to the cubic-equation form of Eq. 20 in the preceding section, the result is a quadratic equation,

$$f_0R_f^2 + R_f(I_0 + f_0R_a) + I_0R_a = 0$$

with solutions

$$R_f = \frac{1}{2f_0}\left(M \pm \sqrt{M^2 - 4f_0I_0R_a}\right)$$

where

$$M = T - (I_0 + f_0R_a)$$

Both of these solutions are always real and positive (unless the given value of T is too small for the system, in which case both are meaningless). The lower solution, the one using the minus sign, is the proper, well-behaved value. For an individual zone, the upper solution here corresponds to the dense branch of the T curve, but it is not clear that a consistent set of such solutions can be formed in a many-zoned region, or that that is the way to go about finding the dense-branch development pattern. It seems more likely that the dense-branch pattern would be found by nucleating the region with development and then working with the lower quadratic solutions—the dense-branch and the upper quadratic solution can be firmly associated with each other only when the entire region is treated as a single zone. If a large region were treated as a single zone, the lower and upper quadratic solutions would give the total development for the lower and dense branches respectively, but because they ignore internal structuring they would not necessarily be very accurate. The upper solution for the dense branch especially would suffer from this. In general, the lower solution would be too low and the upper solution would be too high.

Closing Note

With the exception of the cubic-equation form, which occurs as an algorithm in the prototype program, none of the above ideas has been computationally implemented or explored to any great extent. Most likely they will be given more attention, perhaps along with other extensions that may turn up, under a further contract with the U.S.

Bureau of Public Roads, the primary intent of which is to develop a computer model with more realistic abilities. The author is grateful to the Bureau, not only for its support, but for its patience and consideration during the course of this project.

REFERENCES

1. Transportation and Land Development. Creighton-Hamburg, Inc., final report of project under U.S. Bureau of Public Roads contract FH-11-6792, 1969.
2. Schneider, M. S. Access and Land Development. HRB Spec. Rept. 97, pp. 164-177, 1962.

through these two properties of length and mode distribution by comparing calculations with actual trip lengths and modal splits.

The third unfixed quantity is the relative scale of attractiveness due to land area, R_a, and that due to floor area, R_f. The units of R can be arbitrarily chosen, just as a foot is an arbitrary unit of length, but how much R resides in a square foot of land area and how much in a square foot of floor area is up to empirical measurement. If, for convenience, one unit of R is defined to be the amount of R produced by one square mile of land, then the empirical problem becomes that of finding the rate of R for floor area. This is a rather more subtle problem than that of estimating "a" and "b," but it can be solved by making use of the not too obvious property of the theory that the pattern of floor-area development changes with a change in total R_F. In general, concentration effects in a given system increase with total development (as indeed seems to be the case in the real world). The procedure for scaling R_f would be to calculate the distribution of R_f's for some real region in successive trials, using different values of total R_F, until the shape of the R_f distribution agreed with the shape of the actual floor-area distribution—that is, until the ratios to each other of R_f's around the region were the same as the corresponding floor-area ratios, or, the same thing, the ratio of R_f to the floor area it was trying to represent was the same everywhere. The latter ratio, of course, is the sought-for conversion factor.

One further complication is that it can be difficult, though not strictly impossible, to calculate well trip distributions in order to estimate "a" and "b" without knowing how to measure R, and it is impossible to calculate floor-area distributions well, in order to measure R, without good values of "a" and "b." A certain amount of successive refinement is involved.

These are the things that would be done if things could be done that way. In practice the prototype tests, although they have this general scheme in mind, suffer from an exiguity of appropriate data and from their own ineptitude at representing complicated reality, and tend to limp and stumble from case to case. It is hard to know how to report the results of this kind of project in capsule form. The program works. Sometimes it is even fun. The mathematics make computational sense and look as though they might be reflecting the real world. Calculations exhibit agreement with data, but there are few points of contact. The approach to modal split seems to work well, but the trip-distribution function may be having some trouble in its ambitious effort to comprehend all trip behavior from the very short to the very long in one function. Interestingly enough, the scale of R_f seems to turn out to be about the same as that of R_a; a square mile of floor area apparently contains about one unit of R, just as does a square mile of land area.

On the whole, the idea looks promising to the intuition, and not too much more than that can be expected at this point. Perhaps its most engaging feature continues to be theoretical richness, the fact that it makes predictions, as opposed to statistical recapitulations, and focuses the viewpoint (distorting it at the same time, possibly). Two very general predictions deserve pointing out: if the quality and extensiveness of transportation does not change much, centralization will occur as the world accrues more and more fixtures of development; if transportation becomes more extensive rapidly enough, decentralization will occur.

It might be mentioned in passing that the theory does not particularly require every square foot of floor area to have the same R value as every other square foot, or every square foot of land area to be the same as every other. But it would be far beyond the resolving power of this project to distinguish differences of that order.

ELUCIDATIONS

The expression for floor area at a site is

$$R_f = \frac{R_a I R_F}{J - I R_F} \tag{20}$$

where I is the accessibility integral around the site and J is the integral of these accessibilities throughout the region.

Regional "Temperature"

An examination of Eq. 20 reveals that the ratio J/R_F is a special quantity. Development distributions calculated separately for different regions (excluding constrained areas) have to give the same J/R_F in order to be consistent with the distribution obtained by lumping the regions together in a single superregion. The ratio can be formed for regions of any size at all, large or small, and will be the same for every place entering into a distribution. Presumably regions could have different J/R_F ratios if the interaction between them were weak, or if there were barriers or hidden constraints due to policies on travel, commerce, or immigration, but presumably also these differences would tend to diminish with time.

Indeed, this whole idea of land development can be restated in the form of a principle: the world acts to keep J/R_F the same everywhere within well-communicating regions. This has something of the ring of an entropy law to it, with J/R_F vaguely analogous to temperature. It says nothing about the dynamics of the process, only its goal. Carrying out the analogy, borders, oceans, and geographic and political barriers can be thought of as insulating walls slowing down or impeding the entropic process. The analogy is an attractive one, and may even be valid as far as it goes. But other important factors—things such as access to goods and resources, productivity, various kinds of costs—are not explicitly taken into account (although there seems to be theoretical elbow room for them to work their way in).

Continuing with the idea of regional "temperature," it is easy to let T stand for J/R_F and to rewrite Eq. 20 as

$$R_f = \frac{R_a I}{T - I} \qquad (20a)$$

This form clearly suggests that if T is known, the distribution of development can be calculated without knowing the total development in advance: development becomes a matter of interaction between access at each point and the state parameter, T, instead of the distribution in detail of an exogenous total. Here T begins to function in a rather lofty role, trying to dictate the shape of the world. Presumably T changes with history, rising as the extensiveness and quality of transportation increases, though capable also of falling in a period of transportation stagnation combined with growth in total development. Almost certainly, T has been rising in the modern era (perhaps it has always been rising) and, from Eq. 20a, a site needs more access to get developed, or to stay developed, than it used to. Moreover, T increases especially fast when places of more than average desirability become accessible, another long-term tendency that works some hardship on fixed-access sites and on those that are less congenial.

The notion that T is some kind of fundamental system measure has considerable appeal to vagrant speculation. Perhaps deeper social, economic, and historical meanings can be found for it. Perhaps the dynamics of the development process can be understood in terms of flows of development governed by T differences from place to place. Perhaps T represents something to which a value can be attached in human events, so that a place of higher T may be said to be better off than one of lower T. But in a much more practical vein, the concept of T serves to make a connection among regions that could be treated together only in a system too big to handle comfortably: calculations done for areas all over the country can be made consistent with each other without having to stuff the whole continent into a computer at once.

The Dense Branch: Fiat Development

Given a region containing a transportation system but no development, there will be some value of J that may be called J*. If development is added uniformly everywhere, with the transportation system remaining fixed, it can be seen from definitions that J becomes, for any total development, R_F, and with total land, R_A:

$$J = (J^* - Z)\ (1 + R_F/R_A)^2 + Z\ (1 + R_F/R_A)$$

Here Z is a contribution to J* due to access from within the region to the world outside. (In a real, nonuniform case, J would increase more rapidly with R_F than this because of centralization effects.)

And T becomes:

$$T = \frac{1}{R_F}(J^* - Z)\ (1 + R_F/R_A)^2 + \frac{Z}{R_F}\ (1 + R_F/R_A)$$

This is a two-valued function, approaching infinity when R_F is either very small or very large, with a minimum at

$$R_F = R_A \sqrt{[J^*/(J^* - Z)]}$$

So starting with an empty system and knowledge of the world value of T, it appears that two quite different development patterns can be calculated, one at a lower total development and one at a higher. The lower density solution apparently is the proper one, representing normal incremental growth, and the higher solution can be interpreted as a kind of critical mass effect in which the greater development generates enough self-access to be stable. As Z, the contribution of the outside world, becomes large compared to J*, the minimum point separating the two branches of the T curve moves toward R_F, implying that the dense-branch solution for a small area well-connected to the rest of the world probably entails unreasonably high density. But for a substantially isolated region, the existence of the dense-branch solution offers the possibility that unexpectedly large self-sustaining development can occur, whether by fiat, special inducement, historical quirk, or otherwise.

Even assuming the general validity of Eq. 20a, the question of dense-branch development is an open one. It may very well be nothing but a whim of the mathematics. On the other hand, it may have real significance, especially when considering parts of the world, ancient or modern, having low values of T, which would allow smaller dense-branch development. Presumably dense-branch development could be accomplished suddenly by fiat through deliberate acts of political or economic power. It could also occur in more gradual evolution if T were to fall, go around the minimum, and rise again, dragging a city or two along with it. Conceivably the world might be made up of both low-branch and dense-branch development, with the distinction constituting a rather elegant definition of urbanization.

Everything else staying the same, the lower the low-branch solution the higher the dense-branch solution, and the greater the difference between them. The difference is mathematically capable of being small, but it looks as though it is commonly very large, possibly too large to be interesting, and growing all the time. It is intriguing to note that history records many examples of fiat development, most of which, like Karakorum, dwindled away soon after the fiat was removed. The two notable exceptions that come to mind are Constantinople and Alexandria—both enormously impressive in scale, even by modern standards. Perhaps it takes that kind of effort to make the great leap from low branch to dense branch, from normal village to hypernormal city.

Of course, both of these cities were well situated in the first place, unlike Karakorum, and their situations were improved in the course of founding by introduction of streets, roads, and shipping. This is a two-edged observation. On the one hand, it may mean that their low-branch solutions were fairly high and their dense-branch solutions fairly low, making them suitable candidates for bridging the gap. On the other hand, it opens the interpretation that no gap-jumping was involved at all, but that they are merely unusually dramatic instances of ordinary development.

Whether or not the dense branch has anything to do with fiat development, the founding of cities (or anything else) cannot really be decided without more and better numbers. In the absence of quantitative demonstration, it does not seem urgently needed to explain anything, yet there are things that it might very nicely explain.

The Cubic-Equation Form

From the theoretical point of view, which prefers to think of a site as indefinitely small and contributing virtually nothing to its own accessibility integral, Eq. 20 is the natural and lucid way to write the expression for floor area. There is another form, though, not so natural but still lucid in its own devious way, that turns out to be in general better behaved, and in some cases quite necessary for practical computing and numerical analysis.

Equation 20 is not a pure expression for R_f when dealing with sites or zones of finite size. R_f does not really stand isolated on the left of the equation, but occurs on the right side as well through its contributions to I and J. From the definition of I, it is easy to see that

$$I = I_0 + f_0 R_f$$

where I_0 is I with the contribution due to R_f extracted, and f_0 is the intrazonal value of the trip-distribution function, its average value within the origin site itself; from the definition of J, it is possible (ignoring contraints, which complicate things somewhat) to work out that

$$J = J_0 + 2R_f I_0 + f_0 R_f^2$$

where J_0 is J with the contribution due to R_f extracted. Putting these expansions of I and J back into Eq. 20, and picking through the notational shards for R_f, results in the cubic equation

$$f_0 R_f^3 + R^2(2I_0 - f_0 R_f) + R_f(J_0 - I_0 R_F - R_a f_0 R_F) - I_0 R_a R_F = 0$$

This actually represents a large set of simultaneous equations, of course, one for every site, with every R_f adding something to every I_0 and J_0. There is not much to be gained from trying to write out the solutions of such equations—each one can be solved numerically and a set of consistent solutions can be found by iteration. A certain amount of analyzing can be done, however, and, without pretending to say anything rigorously general or exhaustive about any large set of simultaneous cubic equations, the following properties appear to hold.

Each equation always has at least one entertainable solution, a root that is real, positive, and consistent with real and positive solutions for every other equation; often there can be two such roots, and sometimes three.

If only the smallest entertainable solutions are accepted for all equations that have more than one, the result is a unique set of solutions for the system—development at every site is uniquely determined. This appears to correspond to all real cases and is the policy adopted in the prototype program.

If an upper solution is used, there can be more than one development pattern depending on which site is chosen for the upper solution; these upper solutions entail concentrations of development at densities well above the point at which congestion effects and construction costs would decisively supervene in the present world. The physical meaning would seem to be that various stable development patterns involving extreme concentration could be achieved, ignoring congestion, if someone had the power to deliberately rearrange things that way. These upper solutions should not be confused with the dense-branch solutions discussed earlier. The contexts are entirely different; no doubt there is some kind of formal indirect relation, but no likely one worth pursuing.

In a finite region, with the transportation system remaining fixed, as more and more total development is added the lowest solution for the most central place will rise while the next higher solution will fall toward it. At some point these two solutions will meet, and then, if a little more development is thrown into the total, they will both abruptly become nonreal, leaving only a third, uppermost solution. Evidently a highly saturated region can become suddenly supersaturated, requiring the most central place to grow cataclysmically by sucking in development from all other places. Again, these effects occur at unfeasibly high densities. The physical meaning is problematical, but it is certainly a warning of some kind to the world.

It should be pointed out that solutions other than those in the "lowest entertainable" set are always of clouded significance, though it is not impossible that a meaning can sometimes be found for them. The treatment here is essentially numerical, making the assumption that intrazonal structuring does not exist and need not be considered. Obviously, the smaller the zone size, the better this assumption. But if Eq. 20 is taken to be a point function applying to infinitely small areas of infinitely small R_a, the cubic equation cannot even be developed for a point of finite R_f/R_a density, because such a point makes no appreciable contribution to I or J. In general, however, the three values of R_f as R_a approaches zero are the "proper" value taken directly from Eq. 20 (the limiting case of the lowest entertainable solution), and two other values, not necessarily real, positive, or meaningful, given by

$$\frac{1}{2f_0}\left[\left(f_0R_f - 2I\right) \pm \sqrt{4I^2 - 4f_0J + f_0^2R_F^2}\right]$$

The Quadratic Form

If Eq. 20a is subjected to the same kind of treatment that led to the cubic-equation form of Eq. 20 in the preceding section, the result is a quadratic equation,

$$f_0R_f^2 + R_f(I_0 + f_0R_a) + I_0R_a = 0$$

with solutions

$$R_f = \frac{1}{2f_0}\left(M \pm \sqrt{M^2 - 4f_0I_0R_a}\right)$$

where

$$M = T - (I_0 + f_0R_a)$$

Both of these solutions are always real and positive (unless the given value of T is too small for the system, in which case both are meaningless). The lower solution, the one using the minus sign, is the proper, well-behaved value. For an individual zone, the upper solution here corresponds to the dense branch of the T curve, but it is not clear that a consistent set of such solutions can be formed in a many-zoned region, or that that is the way to go about finding the dense-branch development pattern. It seems more likely that the dense-branch pattern would be found by nucleating the region with development and then working with the lower quadratic solutions—the dense-branch and the upper quadratic solution can be firmly associated with each other only when the entire region is treated as a single zone. If a large region were treated as a single zone, the lower and upper quadratic solutions would give the total development for the lower and dense branches respectively, but because they ignore internal structuring they would not necessarily be very accurate. The upper solution for the dense branch especially would suffer from this. In general, the lower solution would be too low and the upper solution would be too high.

Closing Note

With the exception of the cubic-equation form, which occurs as an algorithm in the prototype program, none of the above ideas has been computationally implemented or explored to any great extent. Most likely they will be given more attention, perhaps along with other extensions that may turn up, under a further contract with the U. S.

Bureau of Public Roads, the primary intent of which is to develop a computer model with more realistic abilities. The author is grateful to the Bureau, not only for its support, but for its patience and consideration during the course of this project.

REFERENCES

1. Transportation and Land Development. Creighton-Hamburg, Inc., final report of project under U.S. Bureau of Public Roads contract FH-11-6792, 1969.
2. Schneider, M.S. Access and Land Development. HRB Spec. Rept. 97, pp. 164-177, 1962.

Although the scatter of the data from this simplified relation is great, the equation agrees reasonably well with the observation that many of the test road pavements reached a PSI rating of 1.5 when the mean rut depth reached 0.6 to 0.7 in. (virtual structural failure).

Causes of Pavement Deterioration and Failure

The subjective PSR gives no clue as to the cause of the deterioration of a pavement from the high value it presumably possessed when it was built. The empirical PSI indicates the major and minor factors in the loss of the initial serviceability, but does not define the mechanism by which they develop. In addition, the scatter of the data suggests that there are additional factors in the PSI.

Although deterioration or failure of the pavement to perform its function may be reflected in the surface condition, the seat of the trouble can be in any of the layers which make up the flexible pavement system: the surface course, the base course, the subbase (if any), and the subgrade or embankment. Furthermore, the initial failure of one may lead to a failure, often in a different form, in another. For example, cracking of the asphaltic surface may let surface water into the subgrade and cause its softening and eventual shear.

The mechanisms for pavement deterioration are suggested in Table 1. Some are primarily related to traffic load, whereas others are either independent of the traffic load or are related to the load only in that the failure is intensified or aggravated by the load rather than caused by it. Deterioration and failure in the surface were not within the scope of this investigation. The causes are listed because they had to be considered in diagnosing the mechanism of deterioration of existing pavements and deciding which failures were the result of inadequate pavement thickness. Deterioration and failure of the base and subbase were similarly beyond the scope of this study, except when the base was so similar to the subgrade in its properties that the base and subgrade had to be considered as a unit, as with topsoil and sand bases. The failure of the higher types such as sand-asphalt and soil-cement was not investigated.

A major, and possibly the most important, function of the pavement is to distribute the concentrated wheel load so that the stress does not exceed the supporting capabilities of the subgrade. Of course, the deformation and failure of the subgrade are reflected in the surface condition and thereby in the PSR or PSI. The elastic deformation, consolidation (densification) and shear (bearing capacity) failure of the subgrade are directly related to the wheel load and the resulting stress distributed through the surface and base courses.

TABLE 1

MECHANISMS FOR PAVEMENT DETERIORATION

Surface	Base and Subbase	Subgrade
Elastic deformation—rebound[a]	Elastic deformation[a]	Elastic deformation[a]
Densification (consolidation)[a]	Densification (consolidation)[a]	Densification (consolidation)[a]
Thermal expansion and contraction	Shear failure (bearing capacity)[a]	Shear failure (bearing capacity)[a]
Longitudinal shear failure (shoving)[a]	Deterioration of aggregate	Swell-shrink
Curvilinear shear (bearing failure)[a]	Deterioration of cementing agent	Pumping
Deterioration of bitumen	Swell-shrink	Settlement of deep strata
Separation of courses	Pumping	Mass shear failure (landslide)
Bleeding		Local mass shear (due to weak culverts, trenches)

[a]Primarily related to traffic loads.

The remaining mechanisms are not directly caused by the wheel loads. Swelling and shrinking of the subgrade depend on the moisture changes as well as the mineralogy of the soil. The effect may be bumps and hollows at irregular intervals not related to the traffic or load pattern. Swelling may have a secondary effect in that softening or weakening of the subgrade can lead to deflection or shear failure that is load related. Similarly, shrinkage has a secondary effect in producing tension and shear cracks in the pavement courses above. Although these are not necessarily load related, they may be aggravated by the load. Therefore, it is difficult to isolate the effects of deterioration due to swelling and shrinking, although the basic mechanism is different from the others.

Pumping is a complex phenomenon indirectly related to load; it arises from the effect of free available moisture on a susceptible subgrade or base. The load of the moving wheel causes the pavement components to deflect. After the load passes, the components rebound. If the upper layers rebound faster or more than the lower layers (which is likely because in the typical flexible pavement system, the upper layers are more rigid and possibly more nearly elastic), a temporary void is formed between the layers. If free water is available, it is sucked into the void, only to be expelled at the next loading. If the base or subgrade is easily softened or eroded, the pumping of water in and out creates an erosion cavity and eventually a structural failure.

Settlement of the roadway (ordinarily an embankment), because of consolidation of deeper strata, landslides and localized shear failures caused by weak culverts or improperly compacted backfills behind bridge abutments or in trenches, can cause disruption of the pavement surface and a loss of serviceability. None of these, however, are directly related to the design or adequacy of the pavement. Furthermore, the traffic loads are often not major factors in these phenomena because they may be small compared to the weight of the soil mass that is involved. Pavement deterioration due to these phenomena, therefore, must be discounted in evaluating observed pavement conditions for the purpose of developing a pavement design.

The major subgrade mechanisms that contribute to pavement deterioration are deformation and shear failure.

Subgrade Deflection.—The deflection of the subgrade under traffic load results from stresses particularly vertical, transmitted through the pavement system. Both theory and stress measurements show that vertical stresses become smaller with increasing depth below the pavement surface and with increasing horizontal distance from the center of the line of load application, depending on the elastic characteristics of the subgrade and base course (4).

These stresses have a two-fold effect on the subgrade (and on the other pavement layers). First, they produce a downward deflection of the subgrade surface due to the deformation of the soil without appreciable volume change. This can be visualized as the shortening and lateral building of the column of soil immediately below the load similar to the shortening of any axially loaded structural member. If it is assumed that the subgrade is a semi-infinite isotropic homogeneous elastic mass with a modulus of elasticity of E and is momentarily incompressible and that a uniform pressure of q is applied to a square area of width b, the deflection ρ due to deformation will be

$$\rho = \frac{0.6\,qb}{E} \tag{2}$$

that is, the deflection is the same as for a free-standing column of soil whose height is 0.6 times the width of the column. Of course, neither the distribution of the load nor the shape of the loaded area of the subgrade is as simple as the conditions assumed in this equation. More accurate, and more elaborate, mathematical representations of the deformation deflection of a subgrade are available. All are of the same general form as Eq. 2; therefore, this suffices as a model for illustrating the effects of some of the different factors involved. The deflection in any case is directly proportional to the pressure and the size of the loaded area and inversely proportional to the modulus of elasticity.

The second deflection mechanism is the consolidation or densification of the subgrade. Although the theories of soil settlement due to reduction in the volume of the voids have been primarily applied to foundations of structures, they apply also to the consolidation of the subgrade. The relation between void ratio change and stress increase is more complex than that for elastic deformation and, therefore, a simple expression for consolidation settlement is not available even for homogeneous soils. However, consolidation settlement does increase with increasing stress, not in direct proportion but more nearly in proportion to the log of the increase compared to the original stress due to the soil weight.

Under repeated loadings, progressive consolidation occurs. With each successive cycle of load and unload, the reduction in voids rapidly becomes less. Settlement appears to continue indefinitely, but at an ever decreasing rate. Subgrade deformation and consolidation cause an elongated depression in the wheelpath that is entirely below the original surface level. The deformation deflection is temporary and is recovered after the wheel passes. The major effect is an "alligator" cracking of the surface course if the deflection is sufficiently great. The estimated limiting deflection, based on the U. S. Navy airfield design, is 0.2 in., although some highway departments have suggested limiting deflections of 0.05 in. for major highways. Consolidation deflection causes a permanent rut entirely below the original surface. The rut may be accompanied by longitudinal and possibly transverse cracks. In addition, long longitudinal waves in the rut may be observed where there is severe consolidation.

Shear Failure.--Shear failure of the subgrade, similar to the bearing capacity failure of a foundation, can result if the stresses transmitted to the subgrade through the base and surface courses exceed the strength in a sufficiently large zone. If it is assumed that the subgrade is homogeneous and its properties can be described by the unit weight γ, the cohesion c, and the angle of internal friction φ, and if it is assumed that the pressure transmitted to the subgrade is vertical and uniform over an area of width b, the pressure at which the soil will shear q_o is defined by

$$q_o = \frac{\gamma b}{2} N_\gamma + c N_c + q' N_q \qquad (3)$$

In this expression N_γ, N_c, and N_q are dimensionless functions of the foundation shape and angle of internal friction and q' is the weight of the pavement and base above the subgrade.

Many variations of this expression, originally proposed by Terzaghi (5), have been published. The differences are in the mode of loading and assumed character of the zone of shear failure and they are manifested in differences in the values of the N-factors. So far no analysis has been developed for a nonuniform loading of indefinite width such as that transmitted through the pavement to the subgrade. However, it is to be expected that the general form of the equation will be little changed; instead, the values of N will reflect the nonuniform loading. Therefore, bearing capacities for subgrades computed by Eq. 3 and utilizing the N-values for one of the existing methods of analysis should be approximately proportional to the true bearing capacities. Or, conversely, the safety factor with respect to shear failure computed by Eq. 3 and utilizing certain existing N-factors and the average stresses transmitted to the subgrade through the pavement system should have some reasonably constant relation to the true safety factors.

Whereas the strength parameters c and φ reflect complete soil failure, they may not indicate the development of limited but accumulating shear under repeated loads that are not great enough to produce complete failure. Although little is known about the effects of repeated loading on progressive shear, the indications are that the magnitude of progressive failure increases with the increasing ratio of the actual stress to the failure stress. That is, progressive failure increases with a decrease in safety factor.

Shear in the subgrade is accompanied by a broad deep depression or rut in the wheelpath with the upheaval occurring beyond it. Longitudinal cracking may be severe and eventually leads to transverse cracking which forms a blocky pattern (6). Shear along the pavement edge may be accompanied by curved cracking and outward movement of the base and surface, and sometimes by severe outward tilting.

Apparent Safety Factor

The stresses computed by any of the elastic theories apply only to the state of elastic equilibrium on which that theory was based. If the elastic state is altered by nonlinear strain or by failure, the stress distribution may be altered. If failure develops suddenly from an elastic state, however, the stresses just before reaching failure are probably not greatly different from those of elastic equilibrium. If it further can be assumed that the pressure, q_o, required for complete failure and that required to initiate failure are approximately the same or proportional, then it is possible to compute an apparent safety factor against failure by

$$SF_a = \frac{\sigma_a}{q_o} \qquad (4)$$

In this expression, σ_a is the average vertical stress transmitted to the soil surface by the pavement system, as computed by an appropriate elastic theory, and q_o is the ultimate bearing capacity computed by Eq. 3, utilizing appropriate factors. The apparent safety factor is not the true safety factor (i.e., failure does not necessarily occur at a safety factor of 1), but it is reasonable to assume that both safety factors are proportional.

Summary

Pavement deterioration and failure is the result of a series of complex processes, none of which are clearly understood and only part of which are directly related to the loads supported. Although exact methods of analyzing the mechanical processes of subgrade deflection and shear failure are not available, approximations can be made that point out the relative importance of the different factors involved and also indicate the relative magnitude of possible deformation and the safety against shear failure.

The greatest unknown factors are those which involve the environment: temperature, frost action, groundwater, surface water infiltration, and other moisture changes. These profoundly influence the deformation and shear failure characteristics of all pavement components but particularly those of the subgrade. At the present time, little is known about the direct effects of the environment on the soil and too few facts are available to permit valid empirical correlations to be made.

SURVEY OF GEORGIA PAVEMENTS

A survey of Georgia pavements was undertaken in 1961 to locate typical areas in all four of the geologic regions (Coastal Plain, Piedmont, Blue Ridge, and Appalachian Ridge—Valley) in which comparable pavements had both exhibited good performance and deteriorated badly. An inquiry was sent to each of the Georgia State Highway Department field divisions asking for their suggested locations for study. From these a list of 84 was compiled for examination and testing.

A field examination was made of each location in the late summer, fall, and early winter of 1961. The pavement was examined visually and data on the roadway environment and pavement condition were obtained. A survey of the traffic was made during the period of pavement examination in which the total number of vehicles and the number of heavy trucks in the lane under study was counted and the percentage of heavy trucks estimated. Although such a short count is not a valid indication of the total traffic, it does give some picture of the character of the traffic on pavements for which no accurate information was available.

The typical depth of rutting was measured using a 4-ft straight-edge placed over the wheelpaths. The segment so measured was then photographed and a sketch made of the pattern of cracking (if any). The dates of construction and repair (if any) were obtained from the field division engineer. He also provided information on the design and construction of the pavement, where it was available. The pavement was marked at the location where samples were to be made, usually in the zone of the failure but not where the failure itself might have disrupted the soil. Finally, a serviceability rating was assigned utilizing the criteria described by Carey (3) and based on the visual observations of the surface condition and its riding qualities.

Sampling

Samples were secured in most of the locations by the Georgia State Highway Department Division of Materials and Tests. The sampling program was necessarily interspersed with the routine drilling and sampling work for new construction, and thus was spread out over several months. Practically all sampling was done in the late winter and spring of 1962 when the soil moisture conditions were likely to be at the highest.

The bituminous pavement was cored where possible and its thickness measured. The thickness of each deeper pavement course was measured and the materials were described visually. Undisturbed samples were secured of each base course layer that contained no gravel and of the top 2 to 3 ft of the subgrade, utilizing 3-in. O.D. thin-wall sample tubes. The samples were sealed in the field with plastic end caps and brought to the Georgia Institute of Technology Soil Engineering Laboratory.

Laboratory Tests

The samples were cut into 6-in. sections using either a high-speed abrasive saw or a metal-cutting band saw. Unfortunately, some of the samples were unsatisfactory because of gravel which caught under the edge of the tube and disturbed the soil or because of faulty sealing. Most, however, were suitable for testing.

Because of the limited amounts of sample available, only one form of test could be utilized. Considered the most representative of field conditions was the undrained triaxial test, utilizing the full sample diameter (approximately 2.8 in.) and no changes in moisture. Where possible, three confining pressures, 10, 20, and 40 psi, were employed; however, in some cases only 10 and 20 psi were used when the amount of sample was limited. The samples, each about 6 in. long, were loaded axially at a controlled strain rate of 0.8 to 1 percent/min. The test data were analyzed on a computer and the results plotted in the form of stress-strain curves from which the initial tangent modulus of elasticity for a confining pressure of 10 psi was found.

AASHO SUBGRADE TESTS

The AASHO test road was constructed to provide as uniform a subgrade as possible, so that initial subgrade variability would not influence the pavement performance. Therefore, the materials utilized were as nearly uniform as possible in composition, and the construction was controlled so that the moisture contents and densities could be kept within narrow limits.

Tests of the subgrade (embankment) base and surface materials were summarized in the AASHO Road Test reports and in other published data on the road (7). These included control tests for quality and physical tests by the U. S. Bureau of Public Roads to determine the structural properties of the compacted materials. In addition, samples were furnished to many state highway departments to be tested by the procedures commonly employed for their own design work. The results of these tests have also been published (8). Limited tests were made of soils in certain of the road embankments which had been removed from test routinely at the programmed end of testing earlier because of deterioration. These included moisture, density, CBR and K-factor tests (1).

Sampling

None of the published data included strength tests of samples of the subgrades as constructed. Four undisturbed samples were secured by the Illinois Division of Highways on about May 1, 1963, well after the spring thaw. These were of the subgrade, commencing 3 in. below the subbase, and were made with 24-in. long, 2-in. O.D. thin-walled tubes. They were sealed at the site and shipped to the Georgia Institute of Technology Soil Engineering Laboratory.

Laboratory Tests

Triaxial tests were made of all four samples utilizing the same method and pressures as for the tests of Georgia subgrades. The results were expressed in stress-strain curves and Mohr diagrams, and are summarized in Table 2. As can be seen, the results are not uniform. There is considerable variation in the densities, moistures, and strengths. With the exception of the samples from Sta. 60 + 00, where gravel made a full program of tests impossible, the samples exhibited comparable angles of internal friction of slightly more than 20° and cohesions between 6.2 and 20 psi. A composite plot of all data shows the weaker materials exhibit an average cohesion of 9.5 psi and an angle of internal friction of 20°. These values were used in subsequent analyses.

For comparison, the BPR tests of subgrade samples, laboratory-compacted to 95 percent of AASHO T99-49 maximum (the specified value), gave c and φ values of 11 psi and 31°, respectively, for Borrow Pit 1 and 8.9 psi and 21° for Pit 2. The corresponding CBR values were 2.7 and 2.5. The cooperative test (8) results were comparable in those cases where the soil was tested under similar conditions of compaction and moisture content.

ANALYSIS OF GEORGIA PAVEMENT PERFORMANCE

The performance of the Georgia pavements was analyzed utilizing the pavement descriptions and serviceability rating. The theoretical bearing capacity and deflection of the subgrade were computed by the methods described. These were correlated to form a semirational basis for pavement design evaluation.

Depth-Stress-Width

A previous paper (4) presented data on the vertical stresses at different depths beneath different pavement systems utilized in Georgia. These tests all indicated that the vertical stress was greatest immediately under the tire and became rapidly less with increasing horizontal distance and increasing vertical depth below the ground surface. For the purpose of analysis, it was assumed that the significant vertical stresses at any depth were those equal to or greater than one-half the maximum vertical stresses at that depth.

The average significant vertical pressures for a 9-kip dual-wheel load (4) were increased by ten-ninths to give the average significant vertical pressures for a 10-kip

TABLE 2

TRIAXIAL TEST DATA AASHO TEST ROAD

Station	Position (ft)	γ_d (pcf)	w (%)	c (psi)	φ (deg)
60 + 00	13.5 ft R of center WB	125	11	51	0
149 + 00	13.5 ft L of center EB	114	14	7.6	24
229 + 00	13.5 ft L of center EB	82	28	6.2	22
360 + 00	13.5 ft R of center WB	113	13	21.5	20
Composite[a]		-	-	9.5	20

[a]Weaker materials.

dual wheel (the present Georgia load limit of 20 kips per axle). A plot of these stresses (Fig. 1) shows the significant vertical stresses beneath different bases as a function of depth beneath the pavement surface. For most Georgia pavement systems, the curve for the 3-in. asphaltic surface and 8-in. topsoil-soil-macadam, or silt base applies. This is almost identical to the stress distribution computed by the Boussinesq theory and should apply reasonably well to all but soil-cement and sand-asphalt bases. The curves for the latter are shown in Figure 1.

The width of the zone of significant vertical stresses was also found from the stress distribution curves (Fig. 2). The curve can be approximated by the straight line whose equation is

$$b = 15 + 0.72 z \qquad (4)$$

where b is the equivalent width in inches and z is the depth below the pavement surface in inches.

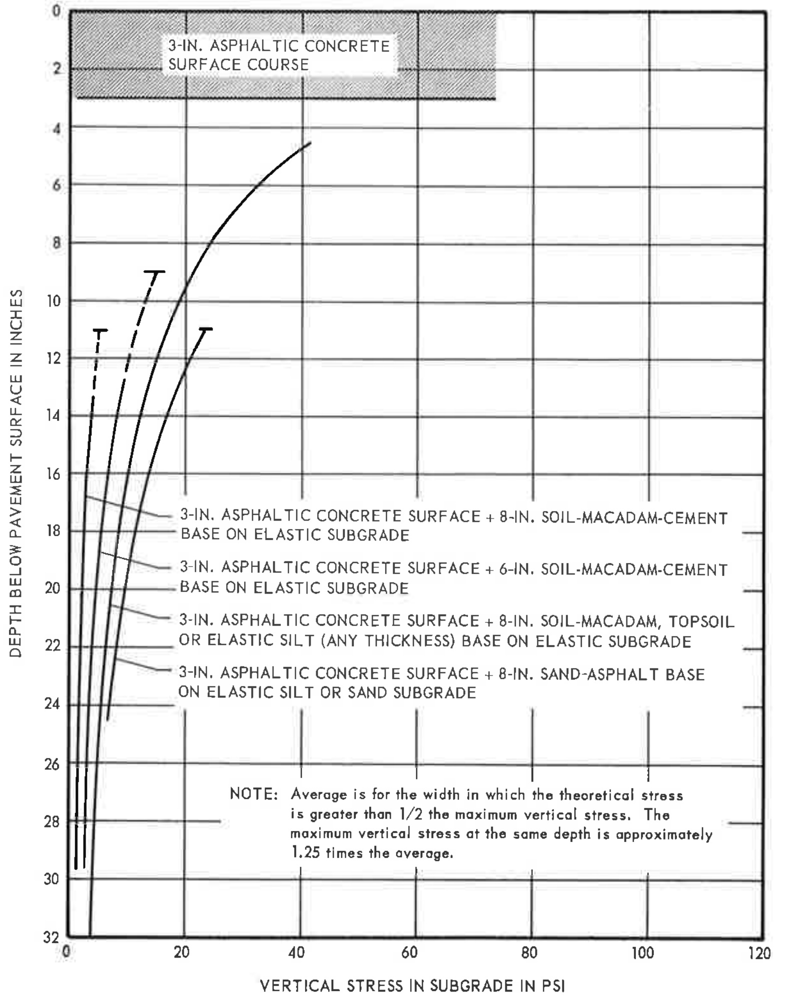

Figure 1. Average significant vertical stress in subgrade for different Georgia base courses and 20-kip axle loads on dual tires.

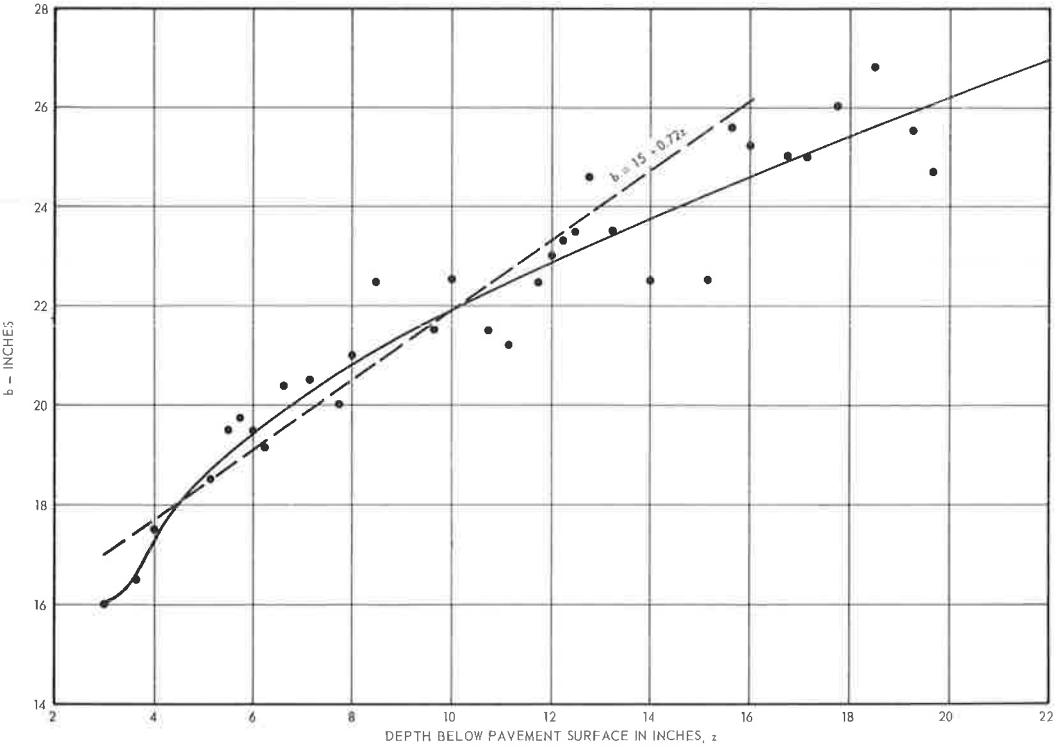

Figure 2. Equivalent width of area of significant vertical pressure in Georgia subgrade.

Deflection—Bearing Capacity

The elastic deformation of each subgrade under a 20-kip axle load was computed from Eq. 2 and the modulus of elasticity for the subgrade at a containing pressure of 10 psi. (The width, b, utilized in the computation was shown in Figure 2. The stress was shown for the depth of the top of the subgrade by Figure 1.) Of course, it would be false to conclude that this represents the true base deflection of the subgrade. However, it should be proportional to the true deflection if the modulus of elasticity determined by the laboratory tests is correct.

The bearing capacity of the subgrade, and in some cases the bearing of each different subgrade layer where the test data differed greatly, was computed from Eq. 3. The c and φ values were those of the soil tests and the b was found from Figure 2. The values of the bearing capacity factors were those computed from the simple Bell-Terzaghi equations as modified by the author (9). For use in these computations the relation was simplified slightly, based on the observation that the total thickness of pavement and base for Georgia is ordinarily 10 to 12 in. In such cases constants can be introduced in the terms involving b, d, and γ with little sacrifice in accuracy (considering the greater error involved in utilizing this or any other existing bearing capacity expression in analyzing pavement capacity).

The vertical stress exerted on the subgrade by the 20-kip axle was found from Figure 1. The ratio of the computed bearing capacity to the stress is the apparent safety factor which is probably not the true safety factor, but should be proportional to it. Further, it would be reasonable to conclude that the lower the safety factor, the greater the possibility of shear failure of the subgrade and the greater the amount of progressive shear.

Traffic

A short-term traffic count was made of each sample section. Data on estimated daily total traffic were obtained from the Georgia Division of Highway Planning. Most estimates were based on actual traffic counting at the regular stations in the area. However, some estimates, particularly for the secondary roads in remote areas, were based largely on experience. In no case was the pavement failure close enough to a point of long-term traffic study that the count can be considered accurate. Both the short-term count at the sample section and the Georgia State Highway Department estimate were utilized in determining the number of trucks per day (other than pickups) on the lane under study. This was converted to an equivalent number of 20-kip axles, utilizing a relationship established by the Alabama State Highway Department in their Loadometer studies (10). The total number of trucks multiplied by the weighting factor gives the equivalent 20-kip axles. The values of the factors used were 0.43 for Interstate and primary roads and 0.32 for secondary roads. The Alabama Loadometer studies were for an 18-kip load. The distribution factors of equivalent 20-kip loads in Georgia would probably be slightly smaller, but in the absence of data, the Alabama figures were used. The daily equivalent 20-kip axle-load figure multiplied by the number of days the pavement was in service gives the total axle loads at the time of the evaluation.

Considering the amount of estimating used to establish this traffic figure, it is likely that it may differ from the true value by 50 percent. An even greater variation is likely on the secondary roads with light traffic where even a moderate use by pulpwood trucks or other local highly specialized vehicles represents the major loading of the pavement.

Serviceability-Safety-Traffic

The serviceability for each pavement area was checked by photographs, measured rut depths, and crack patterns. Greatest weight was given to those factors which reflect the subgrade behavior. For example, although the overall serviceability of a

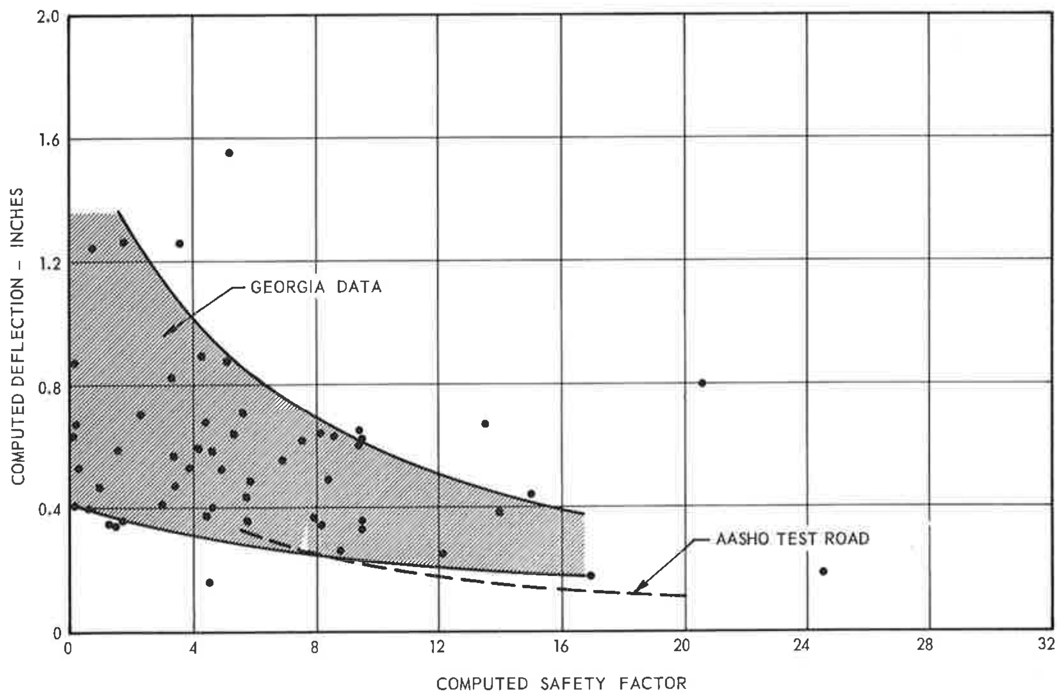

Figure 3. Computed elastic deflection vs computed safety factor of subgrades of Georgia pavements with 20-kip axle loads.

pavement suffering from the peeling of an overlay due to bad bond with the old pavement might be low, the serviceability of the pavement considering the rut depth and longitudinal profile might be high. Because this investigation is concerned with the design of a pavement to fit the subgrade, the subgrade behavior was given greatest weight.

Plots of the serviceability as a function of computed bearing capacity, apparent safety factor, and traffic were made to determine which of these factors was most significant in determining the behavior of the Georgia pavements. A plot of computed deflection vs computed safety factor is shown in Figure 3. Although there is considerable scatter, the relation shows that those pavements having the greatest safety factor against shear failure also exhibit the least elastic deflection; i. e., those soils having the greatest strength are also likely to be the most rigid. This relation also suggests that either deflection or bearing capacity alone might be a satisfactory criterion for design in that one reflects the other to some degree. Because of the limited time available for study and the many factors in both deflection and bearing capacity for which no data were available, no attempt was made to analyze the cause of the scatter.

The plot of serviceability vs apparent safety factor (Fig. 4) also exhibits considerable scatter. However, a general trend is apparent with serviceability decreasing with decreasing safety factor. If the traffic is considered, the trend becomes fairly well defined, with the lighter traffic requiring smaller safety factors than the heavier. Curves were drawn reflecting the largest safety factors required to maintain a given serviceability, for different levels of traffic. In reality, therefore, each curve represents an envelope. There are a few points that do not fit these relations. Some of these with high serviceabilities undoubtedly represent different qualities of initial construction, rather than any deterioration of the pavement. A few exhibit lower serviceabilities because of deterioration other than of the subgrade.

A major unknown factor in the scatter is the fact that the soil test data may not reflect the environmental conditions representative of the greatest degree of deterioration

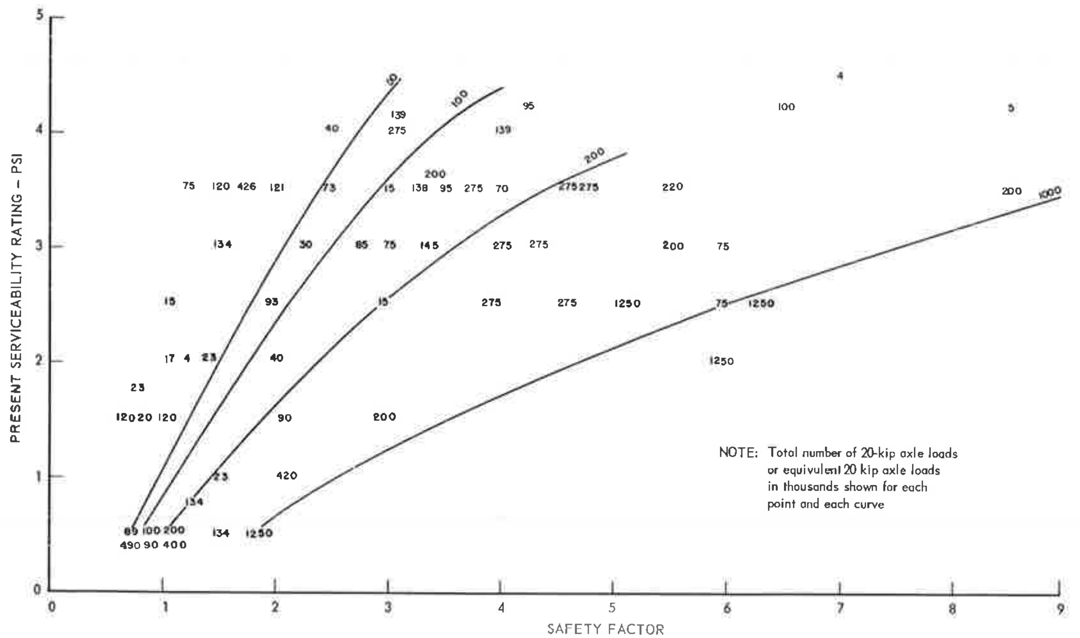

Figure 4. Required safety factors for different present serviceability indexes at end of service period for Georgia pavements and 20-kip axle loads.

and failure. For example, if the subgrade moisture increases in the winter and spring and if failure is most rapid during this period, then the tests should be made of samples obtained during this critical period. Although this was done, it is not known whether the soil at each location was sampled at its worst condition. Moreover, this might not be fair if the traffic during this period of soil weakening was materially less than average. Considering the variable factors which could not be evaluated in this investigation, the degree of correlation shown in Figure 4 is surprising.

AASHO DATA ANALYSIS

AASHO Flexible Pavement Evaluation

The evaluation of the AASHO flexible pavement tests is given in detail in Report 5 (1). A brief review of the program, however, is necessary to provide the background for this analysis.

The entire flexible pavement test program utilized a single subgrade soil, a silty clay classified as A-6 by the AASHO system. This was compacted under close control to densities between 95 and 100 percent of AASHO T99-49 maximum so as to provide as uniform a subgrade as possible and to eliminate the factor of variable subgrade support. The controlled variables were pavement component thickness and traffic. Although a few different base course materials were tested in limited sections, the major emphasis was on the effects of different combinations of surface, base, and subbase thickness under axle loads ranging from 2,000 to 48,000 lb, and with nearly continuous traffic. The serviceability of each pavement section was measured from time to time and a plot of serviceability as a function of the total number of axle loads was made for each pavement section. The results were analyzed statistically to develop empirical relations between axle load, number of axles, pavement design, and performance. The tests effectively demonstrated that serviceability decreased with increasing load and numbers of loads, and decreasing pavement thickness. Curves showing these relationships were developed by assuming a mathematical form and by finding the best fit for the assumed curve by statistical methods.

The method of analysis employed in the AASHO studies does not take into consideration the mechanisms contributing to deterioration or the relative contribution of each. The effect of possible variable subgrade support is not considered. The effect of environment, particularly moisture variation, is also ignored in the primary analysis. Therefore, the AASHO test results cannot be directly applied to the design of Georgia Highways (1, p. 3). Instead, the AASHO data for the 18-kip axle loads (which are nearly equal to the present Georgia legal limit of 20 kips) were analyzed in the same manner as the Georgia data in this report.

Deflection-Bearing Capacity-Traffic

The elastic deflection and bearing capacity of the AASHO subgrade were computed in the same way as for the Georgia subgrades. A single value was utilized for c, φ, and E in all segments, corresponding to the poorer soils tested (the composite on Table 2).

A plot of computed elastic deflection vs computed safety factor (Fig. 5) is well defined, as might be expected, because the only variable involved is pavement thickness. This does, however, suggest the validity of using a single index, either bearing capacity or deflection, as a basis for evaluating subgrade support. A comparison of Figure 5 with Figure 3 is of interest: the AASHO curve in Figure 3 approximately coincides with the lower limit of the Georgia data, suggesting that many of the Georgia soils are more elastic than those of the AASHO subgrade.

A plot of the safety factor of the AASHO pavements vs number of 18-kip axle loads required to reduce the serviceability to 1.5 is shown in Figure 6. A well-defined trend is evident, showing that as the safety factor increases, so does the number of axle loads required to reduce the serviceability to 1.5. Conversely, if a serviceability of 1.5 is demanded at the end of the service life of a pavement, the required safety factor must increase with increasing traffic.

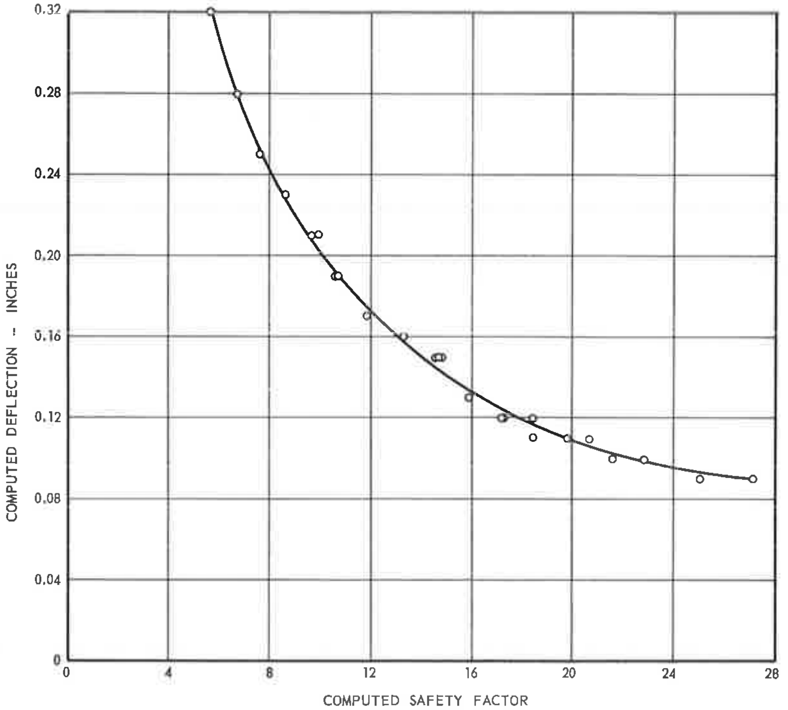

Figure 5. Computed elastic deflection vs computed safety factor of subgrades of AASHO Road Test and 18-kip axle loads.

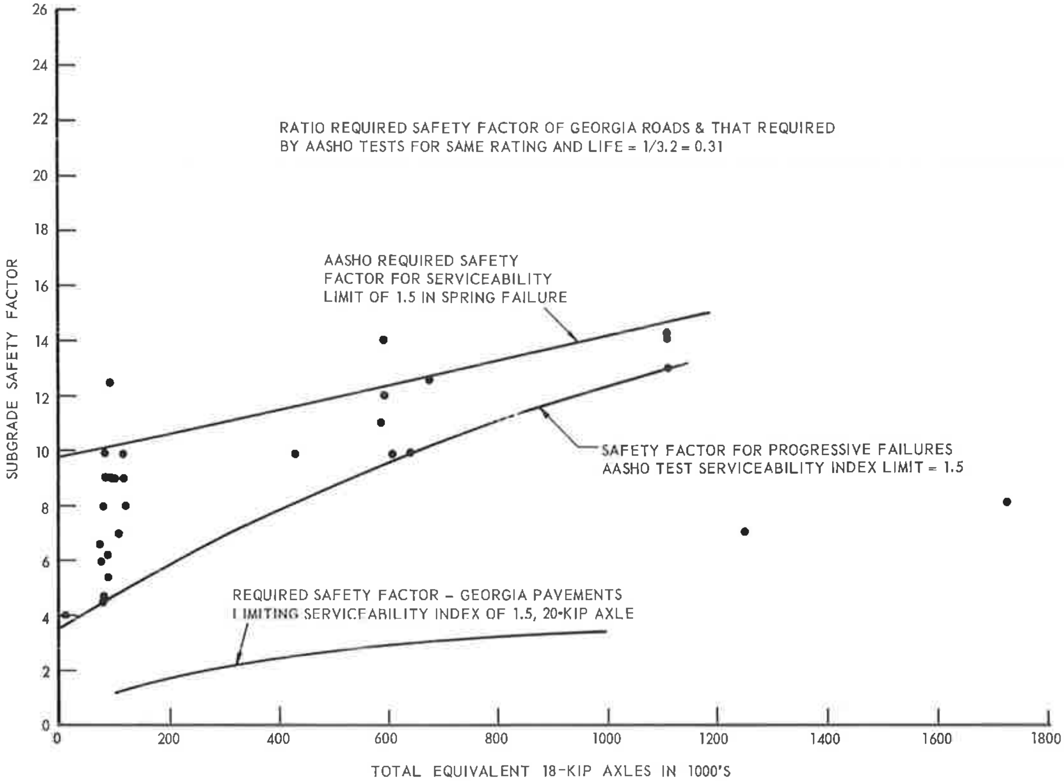

Figure 6. Required safety factor of AASHO pavement to provide serviceability index of 1.5 after different numbers of 18-kip axle loads.

The relationship exhibits considerable scatter, particularly at the lower numbers of axle loads. A study of the individual points shows that those exhibiting the higher safety factors failed suddenly in the spring of 1959 immediately after the thaw period. Two interpretations may be placed on this: (a) the failures were not the result of progressive failure or repeated load; or (b) the soil strength at this time was less than that indicated by the tests of samples made in the late spring of 1963. Those pavements which survived the spring breakup of 1959 exhibited a much better correlation between safety factor and traffic. Of these, however, those points above the lower line represent, for the most part, rather sudden failures corresponding to the spring breakup. The lower curve would seem to represent the more valid relationship between safety factor and traffic. If strength data were available for each test section for the period in which failure developed, the scatter would probably have been much less.

PAVEMENT DESIGN

AASHO Pavement Design

A tentative pavement design method was derived from the AASHO Road Test results by the AASHO committee on design and a draft was presented by Liddle (2). The basis for this development is outlined by Langsner, Huff, and Liddle (11).

The major Road Test correlation is pavement serviceability deterioration (from the initial constructed value) as a function of the pavement design, the axle load, and the number of axle loads. Thus, for a given initial serviceability and a desired serviceability level at the end of the pavement life, and for a required axle load and total traffic, the required pavement design can be found. The correlation is entirely empirical, based on curve fitting, and does not necessarily reflect any consideration of the mechanisms that contribute to failure. The correlation is valid only for the test road subgrade and only if the subgrade properties are uniform and unchanging. An attempt was made to include the variation of the subgrade with the season by assigning a greater weight to the number of load applications occurring during seasons of more rapid deterioration than to those during seasons of less rapid deterioration. The method of determining the factor (1) apparently was purely empirical; the weighting factors were adjusted until the serviceability and total load application data fit the assumed mathematical model with the least variation. This procedure does not indicate the mechanism by which the deterioration is accelerated. In fact, it applies the correction to the traffic rather than to the pavement support factors to which it more logically should apply. Therefore, although it may improve the fit of the AASHO data to an assumed mathematical curve, there is no reason to believe that it might be valid elsewhere.

The pavement design in the main load-performance-traffic-design relationship is expressed in terms of the structural number $\overline{SN}$ or equivalent thickness D; both definitions and symbols are used for the same thing, the first in Liddle's paper (2) and other design memoranda, and the second in the AASHO Report (1). This is related to the actual pavement components by

$$D = \overline{SN} = a_1D_1 + A_2D_2 + A_3D_3 \qquad (5)$$

where D_1, and D_2, and D_3 are the thicknesses in inches of the surface course, the base course and the subbase, respectively. The coefficients a_1, a_2, and a_3 are assumed to be indexes of the relative load-spreading or supporting qualities of each corresponding pavement course. The values found for the AASHO Road Test components varied with the traffic, load and the component thicknesses. The values for the 18-kip axle-load section are given in Table 3.

TABLE 3

VALUES OF COEFFICIENTS FOR 18-KIP AXLE-LOAD SECTION

Course	Coeff.	Unweighted	Weighted
Asphaltic concrete surface	a_1	0.39	0.44
Crushed stone base	a_2	0.15	0.14
Sand, gravel subbase	a_3	0.12	0.11

TABLE 4

COMPARISON OF PAVEMENT COEFFICIENTS

Course	Thickness (in.)	Stress in Course (psi)		Stress Reduction in Layer		
		Top	Bottom	Total (psi)	Per In. (psi)	Comparative a[a]
Surface	4	90	44	46	11.5	0.39
Base	3	44	27	17	5.7	0.19
Subbase	4	27	17	10	2.5	0.08

[a]Assuming surface = 0.39.

Although the AASHO Road Test report (1, p. 36) states that the weighted values indicate that an inch of surfacing ($a_1 = 0.44$) is about three times as effective as an inch of base ($a_2 = 0.14$) or four times as effective as an inch of subbase ($a_3 = 0.11$), this does not necessarily mean that these materials have support qualities or load-spreading abilities in the same ratio. For example, one design of the AASHO Loop 4, where the 18-kip axle load was employed, consisted of the layers shown in Table 4. If the vertical stresses are computed at the top and bottom of each layer using the Boussinesq equation (which applies to a semi-infinite homogeneous isotropic elastic mass), they will be seen to be less at the bottom of each successive course, as shown in the table. The stress reduction in each layer and stress reduction per inch of layer are also shown. The comparison indicates that the first layer is 2 times more effective than the second and 4.6 times more effective than the third. The effectiveness of each layer in terms of the pavement coefficient is tabulated. The resulting values are remarkably similar to the Road Test values for a. Therefore, it must be concluded that the relative values of a_1, a_2, and a_3 do not only reflect the load-spreading or supporting qualities of the pavement materials but also their relative position with respect to the pavement surface.

The Road Test correlation does not include any terms reflecting the subgrade soil support because it was assumed that this was constant and uniform. However, a possible "second" subgrade soil value was inferred from the pavement section with such thick crushed stone bases that the base, in effect, was the subgrade; however, this inference was not checked by any rational procedure. Arbitrary soil support values, S, were assigned to the subgrade and the thick stone base of 3 and 10, respectively. Of course, these values do not necessarily reflect relative support but instead are points of reference.

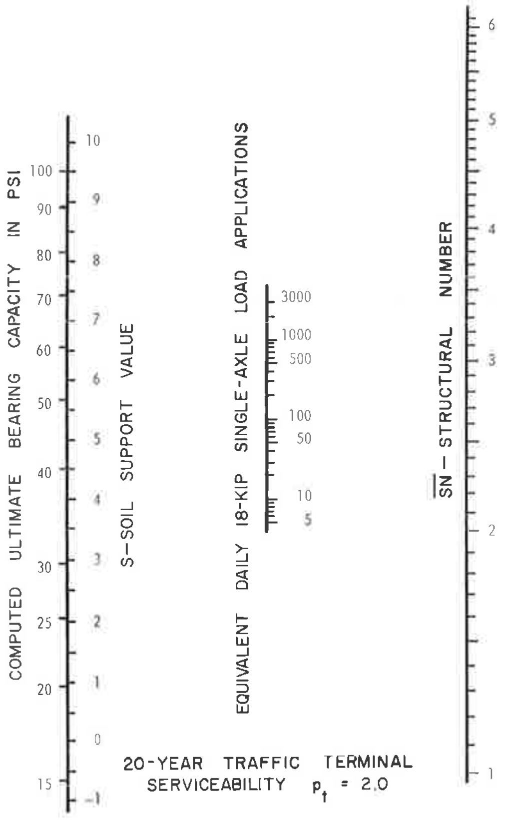

Figure 7. AASHO interim guide to design of flexible pavement (2) adapted to Georgia subgrades.

Nomographic design charts were constructed for the Road Test correlation utilizing an axle load of 18 kips and for terminal serviceabilities of 2 and 2.5. The former is reproduced in Figure 7, from the AASHO Interim Guide for the Design of Flexible Pavements. The same chart has been used by Liddle (2, 11).

Design Constants, Georgia Pavements

The use of the AASHO design charts requires calibration of the soil support scale in terms of some quantitative index or measure of the appropriate property of the subgrade soil for which the pavement is designed. Whereas the AASHO test results do not directly point to the mechanism of pavement deterioration and failure, clues are given by the results of the trenching program. Trenches were cut into pavement sections that had deteriorated to the point of removal from test in 1959. An extensive trench program was undertaken in 1960 when 39 pavement sections were investigated. In each, the transverse profile of the boundary between each of the pavement components and the densities of each layer in and beyond the wheelpath were obtained accurately. These tests indicated that about 25 percent of the thickness change of the pavement layers could be attributed to densification or consolidation of the layers. The remaining change, therefore, must be shear displacement. Such shear displacements can be clearly seen in the transverse profiles of the subgrade surface. Therefore, because it is shown that the major part of the subgrade's contribution to the deterioration of the pavement surface is shear failure, it appears reasonable to presume that the subgrade bearing capacity (its resistance to shear displacement) would be a valid index to the subgrade support of S. On this basis, the AASHO support value would represent an ultimate bearing capacity of 99 psi, based on tests of the samples secured in 1963 some time after the critical period of spring softening. This value probably does not represent the bearing capacity during the periods of most rapid deterioration. This is confirmed by the plot of safety factor vs traffic for the AASHO Road Test (Fig. 6). The lowest curve, which represents the more valid traffic-related deterioration, shows a safety factor of 3.6 required under conditions of very little traffic. The corresponding Georgia data gave a safety factor of about 1 for the same low traffic. Therefore, it is concluded that the actual bearing capacity of the AASHO subgrade was appreciably less than 99 psi during the critical periods of the Road Test.

A plot of the required Georgia subgrade safety factors for the same level of serviceability (1.5), based on Figure 4, is shown for comparison in Figure 6. The Georgia values everywhere are 1/3.2 or 31 percent of the indicated AASHO values. On this basis, the Georgia ultimate bearing capacity equivalent to the Road Test subgrade bearing capacity would be 0.31×99 or 31 psi. This value is recommended for use of pavement design in Georgia as the equivalent of the subgrade support value of 3 (2, 11).

Other values on the subgrade support value scale were established from this key bearing capacity utilizing the AASHO pavement thickness relation for an 18-kip axle load and 100 equivalent axle loads per day. The required safety factors for Georgia pavements for different amounts of traffic and different serviceabilities at the end of the pavement life were found from Figure 4 and plotted in Figure 8. One hundred axles per day for 20 years is a total of 730,000 axle loads. For a serviceability limit of 2.0, the required Georgia safety factor is 4. The safe limit of stress for the correct design would be $\frac{1}{4} \times 31$ or 7.75 psi. The total pavement thickness (3-in. surface plus soil-macadam base) required to maintain the stress at this level, from Figure 1, is 19.5 in. From the AASHO chart for a serviceability of 2, an S of 3 and 100 axles per day require a pavement $\overline{SN}$ of 3.58. The weighted average a for the Georgia pavement, therefore, must be 3.58/19.5 or 0.184.

A second point on the support scale can be found by utilizing a different assumed Georgia ultimate bearing capacity and the computed weighted average a for the Georgia pavement. For example, if the subgrade has an ultimate bearing capacity of 50 psi, the actual stress on the subgrade would be limited to $50/4 = 12.5$ psi. This corresponds in Figure 1 to a total thickness of 13.7 in. Utilizing the previously computed weighted average a, the $\overline{SN}$ would be $13.7 \times 0.184 = 2.52$. From the AASHO chart, the soil support number corresponding to $\overline{SN} = 1.52$ and 100 axle loads daily would be 5.7. Therefore, the S value corresponding to a bearing capacity of 50 psi would be 5.7.

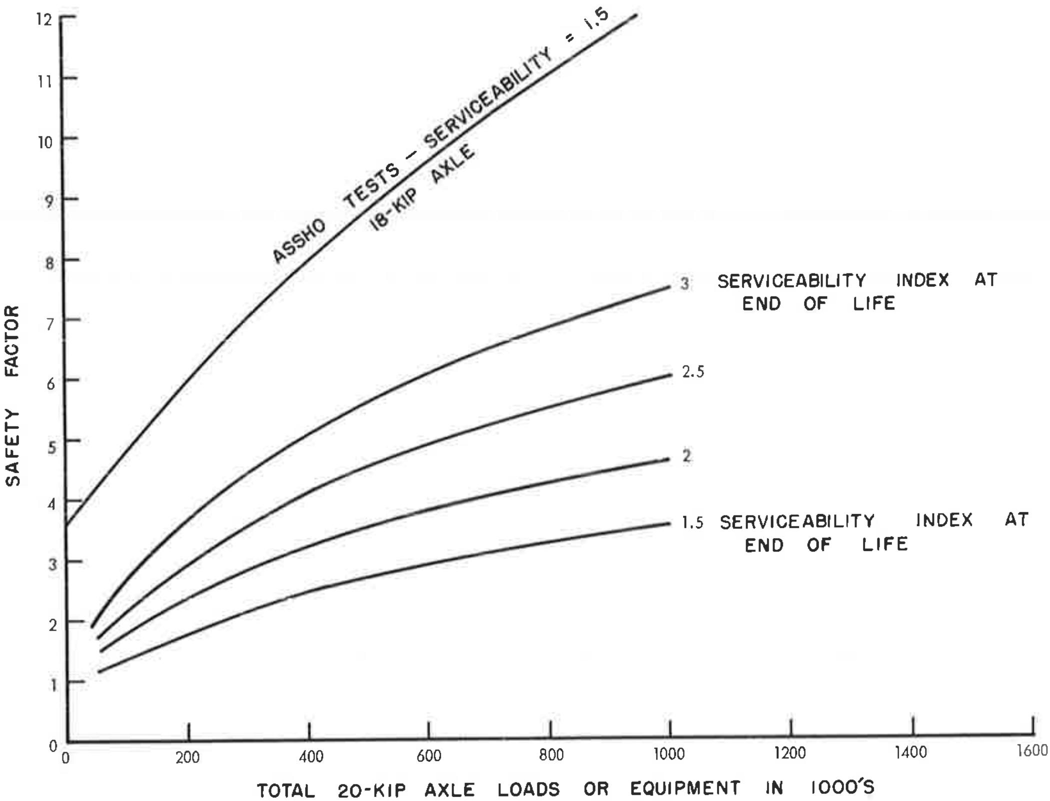

Figure 8. Required safety factors for Georgia subgrades for different numbers of 20-kip axle loads.

By this process the bearing capacities corresponding to various S values were found and are given in Table 5. These values apply to the 100 axle loads daily and the performance rating of 2.0 at the end of the pavement life. However, their applicability to other conditions is probably as valid as the other assumptions made in developing the design method.

In utilizing these values for design, consideration must be given to the test method on which the Georgia evaluations were based. The samples were secured in actual subgrades during the winter and the time of greatest soil moisture and lowest strength. Until data are available on the variations in soil moisture with the season, it is safe only to assume that the limit of capillary saturation is the limiting moisture corresponding to the Georgia test data. It is recommended that the bearing capacity be found from c and φ values determined as follows:

1. Compact two specimens of the subgrade to the lowest density and highest moisture permitted by the construction specifications.

2. Confine each in a triaxial chamber, one at a confinement of 10 psi and the other at 20 psi.

3. Subject each to a head of 1 ft of water from the bottom and allow to saturate until

TABLE 5

RELATION OF S TO COMPUTED ULTIMATE BEARING CAPACITY[a]

Bearing Capacity (psi)	S	Bearing Capacity (psi)	S
15	- 0.7	60	6.6
20	+ 0.9	70	7.4
30	2.9	80	8.2
40	4.5	90	8.9
50	5.7	100	9.5

[a]For using AASHO tentative design chart with subgrade bearing capacities computed from undrained triaxial shear tests of Georgia subgrade soils computed and inundated.

no more water is absorbed (period of saturation to be found by experiment).

4. Load axially until failure occurs, without further change in moisture. For design utilize a seasonal weighting factor of 1 throughout, as in Figure 7.

Determination of the a coefficients for use in the AASHO design method is more difficult because there is little on which to base a correlation. The AASHO values are 0.44 for the asphaltic concrete surface and 0.14 for the crushed stone base. The weighted a for a typical Georgia pavement of 3 in. surface and 8 in. soil-bound macadam would be $3/11 \times 0.44 + 8/11 \times 0.14 = 0.22$. This compares reasonably well with 0.18, indirectly computed from the equivalent thicknesses utilizing the AASHO design chart as previously described.

Based on the subgrade stress studies of the author, the value of 0.44 for the surface appears large. Considering the stress-spreading value of the layers alone, a value of 0.35 is suggested for the asphaltic surface and 0.14 for the stone base. The weighted average of these is 0.197, which is closer to that of the value computed from the bearing study. These values have a ratio of 2.5 to 1, which is close to that found on the basis of the Boussinesq distribution to be applicable to a flexible pavement system employing a granular base.

The a value of the soil-macadam-cement base can be found indirectly from Figure 1. This shows that an 8-in. soil-macadam-cement base (5 psi on subgrade) is equivalent to 22 in. of soil-bound macadam, etc., or would have an a of $22/8 \times 0.14 = 0.38$. The 6-in. soil-macadam-cement base (15 psi) is as effective as 9 in. of soil-bound macadam; it has an a of $9/6 \times 0.14 = 0.21$. At first glance the different a values for the same material might appear contradictory. However, stress theory indicates that the load-spreading ability of a layer capable of supporting tension is a nonlinear function of the layer thickness as well as the material rigidity. For an all-over design value, an a of 0.25 to 0.30 for a soil-macadam-cement base would appear reasonable. This is not greatly different from the value of 0.23 estimated for the AASHO pavements.

The 8-in. thick sand-asphalt base stressed the subgrade to 23 psi which is equivalent to a 5.5-in. thick soil-bound macadam base. The equivalent a value for the sand-asphalt, therefore, would be $5.5/8 \times 0.14 = 0.10$. This is considerably less than the 0.25 estimated from the AASHO test results. (Of course, the AASHO tests did not include a sand-asphalt base; the value was only a guess.) The great difference between the AASHO a values and the values inferred from the Georgia stress tests is possibly the result of the higher Georgia temperatures and resulting lower rigidity, as well as the slower rate of loading.

Alternate Georgia Design Method

An alternate design procedure can be evolved from the Georgia pavement evaluation data:

1. Determine the c and φ of the soil as outlined previously.
2. Compute the ultimate bearing capacity, using an assumed tentative pavement thickness D.
3. Find the appropriate safety factor from Figure 8.
4. Compute the safe bearing capacity by dividing the ultimate bearing capacity, Step 3.
5. Find the total pavement thickness from Figure 1 utilizing the appropriate curve for the type of base course to be employed.

This procedure is no more complicated than that of the AASHO interim guide. It makes use of the AASHO serviceability concept and the traffic-serviceability decline principle. It is based on Georgia performance and on the stress spreading ability of the Georgia base courses. Finally, it is a more nearly rational approach to design than is the AASHO method.

Recommendations for Further Study

Test sections of pavement should be constructed specifically for serviceability-performance studies. These should be a part of the highway system so as to reflect the use and traffic patterns of real highways. They should be placed on typical subgrades in each geologic region and should be constructed with varying pavement thicknesses and Georgia bases. They should be accompanied by a traffic count station where periodic Loadometer studies can be made to determine the distribution of the heavier axle loads. The soil moisture variation should be measured periodically, and the bearing capacity determined by laboratory tests of samples secured so as to reflect the typical range of moistures, particularly the highest. Pavement serviceability should be determined accurately by profile studies.

Subgrade moisture studies should be undertaken to define the range in moisture content changes for the typical subgrades in each geologic region and each different drainage regime.

Triaxial tests should be made on typical subgrade materials utilizing the procedure outlined in this paper for "saturating" the soils, or when more realistic subgrade moisture data become available, by making the tests at those moistures. The bearing capacities and deflections of these materials should be computed from the appropriate theories and correlated with the geology, soil classification, and region for use in preliminary design.

Finally more realistic theories should be developed for the bearing capacity and deflection of the subgrade and each of the pavement components.

CONCLUSIONS

1. There are numerous causes or factors involved in the failure of a pavement to perform adequately. Of the Georgia pavements studies, however, most were load related.

2. The study of the performance of Georgia pavements disclosed a correlation among the serviceability rating or index, traffic, and computed deflection and bearing capacity of the subgrade.

3. The study of the AASHO data disclosed a similar correlation among these factors.

4. The AASHO subgrade had an ultimate bearing capacity of 99 psi. For comparable load, traffic and performance, Georgia roads required an ultimate bearing capacity of 31 psi. The difference is probably the result of differences in environment. The 31-psi required bearing capacity for Georgia subgrades corresponds to the Soil Support Number 3 of the AASHO design.

5. Georgia pavement thicknesses can also be designed by the use of triaxial tests on subgrade soils tested under field moisture conditions. The required safety factor against a bearing capacity (shear) failure is the required thickness to reduce the stress to that necessary to provide the safe bearing. This factor can be found from graphs of the test data and stress distribution below a pavement.

ACKNOWLEDGMENTS

This work has been sponsored by the Georgia State Highway Department and the U. S. Bureau of Public Roads, through the Engineering Experiment Station of the Georgia Institute of Technology. The author is indebted to T. S. Wallace, T. K. Brasher and D. Wheeless, graduate research assistants; to Dr. A. Schwartz who directed the field studies of the pavements and the laboratory testing; to W. F. Abercrombie, Engineer of Materials and Tests and T. Moreland, Soil Engineer, the Georgia State Highway Department; and to the late W. E. Chastain, Sr., Engineer of Physical Research, Illinois Division of Highways.

REFERENCES

1. The AASHO Road Test, Report 5: Pavement Research. Highway Research Board Spec. Rept. 61E, 1962.
2. Liddle, W. J. Application of AASHO Road Test Results to the Design of Flexible Pavements. Preprint Vol. Supple., Int. Conf. on the Structural Design of Asphalt Pavements, Univ. of Mich., Ann Arbor, 1962.

3. Carey, W. N., Jr., and Irick, P. E. The Pavement Serviceability-Performance Concept. Highway Research Board Bull. 250, pp. 40-58, 1960.
4. Sowers, G. F., and Vesic, A. B. Vertical Stresses in Subgrades Beneath Statically Loaded Flexible Pavements. Highway Research Board Bull. 342, pp. 90-123, 1962.
5. Terzaghi, K. Theoretical Soil Mechanics. New York, John Wiley and Sons, 1941.
6. Hveem, F. N. Types and Causes of Pavement Distress. Highway Research Board Bull. 187, pp. 1-52, 1958.
7. The AASHO Road Test, Report 2: Materials and Construction. Highway Research Board Spec. Rept. 61B, 1962.
8. Shook, J. F., and Fang, A. Y. Cooperative Materials Testing Program at the AASHO Road Test. Highway Research Board Spec. Rept. 66, 1961.
9. Sowers, G. F., and Sowers, G. B. Introductory Soil Mechanics and Foundations. New York, Macmillan Co., 2nd Ed. 1961.
10. Eiland, E. Alabama Highway Departments Use of the AASHO Data. Proc. 6th Alabama Joint Highway Conf., Auburn, April 1963.
11. Langsner, G., Huff, T. S., and Liddle, W. J. Use of Road Test Findings by the AASHO Design Committe. Highway Research Board Spec. Rept. 73, pp. 399-414, 1962.