Role of Spatial Dimension in Infrastructure Condition Assessment and Deterioration Modeling

RABI G. MISHALANI AND HARIS N. KOUTSOPOULOS

The treatment of the spatial dimension in assessing infrastructure condition, modeling its deterioration, and, consequently, maintenance decision making are achieved through the development of a spatial distress model that provides the necessary structure for such a treatment. The relationship of the spatial model developed to the temporal deterioration modeling literature is presented. The spatial model developed recognizes two deterioration mechanisms: the *environmental* mechanism, which describes deterioration as a consequence of causal factors and exhibits both macroscopic and microscopic scales, and the *interactive* mechanism, which describes deterioration as a result of distress at a location influencing the deterioration of neighboring locations and exhibits a microscopic scale. The results of the application of the model for the identification of uniformly behaving regions appropriate for condition assessment emphasize the importance of the explicit recognition of the spatial dimension within the infrastructure management process.

Information on facility condition is essential to infrastructure management. Infrastructure facilities are spatially extensive in nature. This paper is concerned with the treatment of the spatial dimension in assessing infrastructure condition, modeling its deterioration, and, consequently, facilitating maintenance decision making.

The Infrastructure management process can be characterized by the following three components (*1*):

- Data collection and analysis;
- Condition assessment and forecasting;
- Strategy selection for inspection and maintenance.

The first component involves gathering and analyzing the relevant data for decision making. This includes data on use, such as current traffic volume and mix, and data on the surrounding environment, such as soil condition, temperature fluctuations, and precipitation. Both use and the surrounding environment play a major role in the degradation process. Distress data representing this degradation are also important to collect. They include information on the location and magnitude of the different distress types exhibited. The most common types of distress for which data are collected are those appearing on the surface of facilities.

The second component of the management process entails assessing the current condition and forecasting the future condition using the data collected in the context of the first component along with deterioration models. Many studies have been conducted in the area of deterioration modeling—see Ramaswamy (*2*).

R. G. Mishalani, Center for Transportation Studies, Massachusetts Institute of Technology, 3 Cambridge Center, Room 208, Cambridge, Mass. 02142. H.N. Koutsopoulos, Department of Civil and Environmental Engineering, Carnegie-Mellon University, Pittsburgh, Pa. 15213.

Both the current assessment and the prediction of condition provide the necessary inputs to the third component of the management process, namely, selecting the strategy for inspection and maintenance over time. The function of this component is to select maintenance activities and inspection strategies that minimize total user and maintenance costs. Many studies have been conducted in the area of strategy selection for infrastructure management—see Madanat (*3*).

As the presentation above indicates, the literature has focused predominantly on the temporal aspects of the management process. The importance of the temporal dimension should not be underestimated. However, there is a clear lack of explicit treatment of the spatial dimension. This paper defines the role of the spatial dimension within the management process and presents a new understanding of the spatial behavior of infrastructure distress. A methodology for identifying this behavior is, in turn, developed. Finally, the importance of the new spatial understanding in achieving effective maintenance decisions is indicated through an application.

ROLE OF SPATIAL DATA WITHIN THE MANAGEMENT PROCESS

The spatial dimension plays a role in the following functions (relating to all three components of the infrastructure management process):

- Condition assessment and temporal behavior modeling;
- Inspection strategy selection with regard to spatial coverage; and
- Identification of systematic type measurement errors resulting from unexpected exogenous factors.

Temporal deterioration at each location may be strongly related to the deterioration of its neighboring locations. Hence, measures of condition are, in general, associated with a specific region. Typical measures include percentage of the area cracked and average variation in deformation per unit length. Such measures can only be quantified in the context of well-defined regions. Therefore, to use the available data for infrastructure management purposes, the behavior of distress over space must be understood. This understanding is also necessary for condition assessment and temporal behavior modeling. The critical consideration is to quantify condition based on regions that will behave uniformly over time. Otherwise, the prediction of the behavior of such stand-alone entities ceases to have any meaning. Therefore, the purpose of modeling behavior over space is to provide the necessary structure based on which current condition can be quantified and deterioration modeled.

The role of the spatial dimension is also relevant at the strategy selection level. Because inspection is conducted over the length of the facility, it is necessary to specify not only *when* to inspect but also *where*. Understanding the spatial behavior allows such spatial decisions. Finally, the role of the spatial dimension is also significant for the data collection and analysis component. This role relates to measurement errors. Understanding the spatial behavior can potentially enhance the identification and quantification of systematic type errors resulting from unexpected exogenous factors.

LITERATURE REVIEW ON SPATIAL AGGREGATION

The literature on the partitioning of infrastructure space into regions within which distress data can be aggregated for condition assessment is very limited. The Agency Method is based on the argument that since deterioration is a function of causal variables, space is partitioned into the largest possible regions such that each of the causal variables are "constant" within each region (*4*). Common causal variables include

- Daily traffic volume and mix (percentage of trucks);
- Structural design;
- Quality of construction;
- Environmental factors;
- Maintenance history.

The Agency Method identifies the locations of "change" in any of the causal variables. These locations are the boundaries of the regions used for condition assessment. Each causal variable is assumed to have a "constant" magnitude within each region. In reality, however, none of the causal variables have a *constant* magnitude even within short lengths along the facility. Therefore, given that many of the important variables are continuously changing along the facility, the critical requirement for the method to produce good results is the appropriate definition of "change" and, consequently, its detection. Unfortunately, the data on causal variables are unavailable at the desirable level of detail primarily because of the difficulties in collecting such data despite recent advances in remote sensing technologies (*5*). Therefore, agencies rely on historic data, which are spatially aggregate. In the very nature of such aggregate causal data is the presence of locations of change. Therefore, the definition of regions along the facility reduces to superpositioning the locations of change in the aggregate data.

Furthermore, the aggregate historic data on causal variables are unreliable due to the lack of knowledge about the original circumstances surrounding their construction and maintenance. The variation of the causal variables within regions thought to be uniform, therefore, could be significant. The thickness of the different pavement layers, for example, have been found to vary significantly compared with data obtained from original design plans and maintenance records (*6*). According to AASHTO (*4*), one of the most difficult variables to assess with regard to "changes" along the facility is the quality of the subgrade or foundation as characterized by the soil type.

In response to the uncertainty in the use of causal variables, AASHTO has proposed a method for delineating "homogenous" regions using either distress measurements or condition indexes (*4*). The method is based on a response function that represents either a measure of distress or the value of a distress index (which is a composite of measures of several distress types) along the highway. It is assumed that the function is piecewise constant. The locations where the response function intersects its mean are adopted as the boundaries of uniform regions. However, it is not necessary that the mean intersects the response function at *all* locations of abrupt change. Only when there is a single abrupt change, does this method guarantee the desired solution. Moreover, as AASHTO has indicated (*4*), the response function is not piecewise constant in nature. Therefore, under realistic situations of high small-scale variability, the method has the potential to identify many unnecessary small regions, resulting in misleading condition assessments. For a more detailed discussion on this method see Mishalani (*7*).

Another study that addressed the spatial dimension relates to the determination of maintenance activity regions that take into account maintenance implementation constraints (*8*). Such considerations require as input uniformly behaving regions.

METHODOLOGY

In this section, pertinent deterioration models are first reviewed. Based on the review, a spatial distress model is developed. Subsequently, a methodology for identifying regions of uniform behavior using the developed spatial model is presented.

Mechanisms of Deterioration

Once an infrastructure facility is constructed, it is subjected to both traffic loading and weathering. As a result, the facility undergoes a natural degradation over time. In addition to these exogenous factors, the durability of the facility (as captured by the design of the facility and the quality of its construction) and the material on which it is founded (as captured by soil conditions) play a significant role in its behavior over time. All these factors influencing the deterioration process are referred to as environmental variables in this paper. The variables associated with the design, construction, and original soil properties are static in nature, while traffic and weather vary over time with a cumulative effect on the facility. Given that facilities extend through a vast amount of space, one expects *both* the static and dynamic exogenous variables discussed above to *vary* over space. Hence, understanding the mechanisms of deterioration from a spatial perspective is fundamentally important in understanding the temporal behavior of a facility.

The focus of most studies on the mechanisms of deterioration is the temporal dimension. Despite time and space being closely related, very little effort has been directed toward investigating the spatial dimension. Therefore, the following review of the deterioration literature focuses primarily on exploring the spatial knowledge implicitly or explicitly assumed in temporal modeling.

Theoretical models (as opposed to empirical) capture the effects of traffic loading and weathering on the mechanical characteristics of the material of which the facility consists and, consequently, determine the distress that would result from that effect. Moavenzadeh and Brademeyer (*9*) and Markow and Brademeyer (*10*) developed model systems that use basic mechanical principles to explain the behavior of infrastructure facilities over time.

From a spatial perspective, Moavenzadeh and Brademeyer (*9*) explain surface deformation as a consequence of *microscopic* (or small-scale) variation in the material properties over space. The use of a spatial correlation function is proposed for capturing such variation. This is an important component of the overall spatial variation of interest in this study and, therefore, should be taken into account in building the spatial model that describes the behavior of distress along extensive facilities.

Markow and Brademeyer (*10*) adopt a dynamic structure where the change in condition is modeled. It is assumed that the condition of a facility at time $t + \Delta t$, denoted by $C(t + \Delta t)$, is a function of both the condition at time t, $C(t)$, and explanatory variables such as design and traffic. Measures of condition are, in general, associated with a specific region. Therefore, the dynamic nature of the model assumes that the condition of a region at a particular point in time is a fundamental determinant of the condition of that same region at a subsequent point in time. This implies that the space within the region considered evolves in the same manner over time. This is expected to be the case when the magnitude of the causal variables are "fairly" constant within the region.

Due to the complexity of the mechanisms involved and the high degree of variability in the factors affecting the deterioration process, it is difficult to develop a realistic mechanistic model that accurately explains the behavior of distress over time. Nevertheless, such models provide the necessary foundations for building empirical models. Empirical models have proven to be more successful and are widely used by facility management agencies. They are founded on direct observations of surface distress, and correlations of such observations with the explanatory variables and maintenance actions of interest.

Paterson (*11*) adopted such an approach. Condition is represented in a disaggregate manner by a vector whose elements are associated with the different distress types. The behavior of each element is modeled separately over time. The dynamic dimension is microscopic compared to Markow and Brademeyer's model. For the distress types that occur discretely in space—such as cracking, raveling, and potholes—the temporal model is characterized by two distinct phases: initiation and progression. The initiation model predicts the failure time, which is the time at which the first distress appears. The progression model is conditional on initiation having taken place and measures the change in condition from one point in time to the next.

The representation of the dynamic dimension at this microscopic level implicitly assumes that the distress at a location influences the deterioration of neighboring locations. Therefore, the mechanism of deterioration captured by these models relates to distress *interaction* in space. Once a crack occurs in a particular region, the temporal behavior *switches* from one of initiation to one of progression, implying that the *first* crack induces initiation and progression of other cracks in that region. Such regions are referred to as sections, and are defined as "nominally homogeneous." It is suggested that a "convenient" section has the width of a lane and a length of 320 m. The use of the width of the lane is appropriate due to the general confinement of traffic to lanes. The length of the section should capture the spatial extent within which distress interaction takes place. Although the scale associated with such interaction is expected to be *microscopic*, the section length of 320 m is not quantified as such but rather chosen based on engineering judgment.

In summary, three types of spatial behaviors are implicitly or explicitly assumed by the deterioration models examined above:

- Macroscopic environmental behavior;
- Microscopic environmental behavior; and
- Microscopic interactive behavior.

Spatial Model

The process that results in the occurrence of surface distress can be broken down into two mechanistic processes that occur simultaneously. The environmental process and the interactive process. The *environmental* process describes deterioration as a consequence of a multitude of environmental factors: subsurface conditions (such as soil conditions, design standards, and construction quality); and external conditions (such as traffic, weather, and drainage). These factors are not uniform over space, and different combinations will result in different deterioration propensities. The realized distress will mirror the environmental variability. For example, regions of both poor soil condition and high traffic volumes have a much higher likelihood of exhibiting high distress levels than regions of good soil conditions and low traffic volumes. It is this process that motivates the use of causal variables by agencies managing infrastructure facilities. In relation to the model developed by Paterson, it is the environmental factors that influence the time at which initiation takes place.

The *interactive* process describes deterioration as a result of distress at a location influencing the deterioration of neighboring locations. For example, the likelihood that additional cracks will initiate near other cracks is greater than of their doing so in a region exhibiting no cracks. Moreover, the likelihood that two neighboring cracks will propagate and connect is very strong. This phenomenon is known as crack coalescence in the material science literature (*12*). In terms of Paterson's initiation and progression model, it is both the interactive process and the environmental process that are responsible for the progression stage.

The environmental process exhibits a spatially extensive scale where the level of distress is "fairly" constant for long stretches along facilities. This is a direct consequence of the nature of the underlying causal variables. The variables are not expected to vary substantially within relatively long stretches and when a change occurs, it is expected to be in the form of an abrupt shift. For example, traffic volumes are expected to be "fairly" constant. Changes in volume occur at exits and entrances; therefore, any significant change in volume will most likely be abrupt in nature. Since design standards are usually based on traffic volumes, a similar pattern in the design is expected. Moreover, since construction quality is a function of the source of materials and the contractor, this variable will also be constant for relatively long stretches, and any changes will most likely be abrupt. Therefore, the scale of the environmental process is expected to be macroscopic with an order of magnitude of several kilometers. In terms of the model developed by Markow and Brademeyer, the uniformity of the deterioration of a region over time is consistent with this environmental process.

Although the causal variables in most cases remain relatively constant for some "long" stretches, they still exhibit *small-scale variations* as captured by the model developed by Moavenzadeh and Brademeyer. Moreover, the interactive process is expected to exhibit a spatially "local" scale compared to the macroscopic scale associated with the environmental process. Since interaction takes place primarily as a result of the weakness distress induces on its surroundings, from a mechanistic perspective the interactive process is governed by the environmental mechanism in the sense that the environment affects the magnitude of the weakness a distress induces and, consequently, the strength of the interaction. For example, in situations of high design standards, the strength of the interaction between two cracks is lower than in the case of poor design standards. Therefore, the interactive process contributes to the variability within the regions of "fairly" constant environmental variables.

The hypothesis that emerges from this discussion is a spatial pattern with regions of well-defined boundaries. [Within each region, distress fluctuates around a region-specific constant level.] Such regions are referred to as *fields* in this study. The developed spatial hypothesis is represented by a stochastic process. For the sake of simplicity, the case of a single distress type is presented. Let X_s be the measure of distress at location s of the facility. The magnitude of the distress at location s consists of two components:

- A systematic component, constant within each field, which captures the macroscopic behavior; and
- A stochastic component, with zero mean, which captures the inherent variability of the microscopic behavior resulting from both the interactive process and the small-scale environmental variations.

The stochastic component also captures the random nature of measurement errors. It is assumed that the systematic nature of measurement errors (i.e., the bias) has been corrected—see Humplick (13).

Mathematically, this model is represented as follows:

$$X_s = \mu_s + \varepsilon_s \text{ and } \mu_s = \begin{cases} \mu_1 & \text{if } s \in \mathrm{F}_1 \\ \bullet & \bullet \\ \mu_m & \text{if } s \in \mathrm{F}_m \\ \bullet & \bullet \\ \mu_M & \text{if } s \in \mathrm{F}_M \end{cases} \quad (1)$$

where

X_s = distress at location s;
s = locations along the facility;
m = index representing the fields ($m = 1, \ldots, M$);
F_m = set of locations contained within field m;
μ_s = systematic mean distress;
μ_m = systematic mean distress within field m; and
ε_s = random variable (of zero mean) representing the deviation of the actual observed distress from the mean μ_s.

Since the deterioration at the microscopic level occurs within the context of the systematic deterioration at the macroscopic level, the fields are expected to behave uniformly over time (with respect to both the mean distress level and its rate of change). The microscopic environmental process and the interactive process will result in spatial variation in distress within each field. That is, since the underlying dominant force governing deterioration is the environmental process, such local variations will take place conditional on the environment within which they occur. Since the local variation within such fields is a variation around a constant level of deterioration, any aggregation within a region fully contained within such a field is an estimate of the constant level of deterioration. Therefore, the larger the region, the better the estimate (i.e., the lower the variance of the estimate) as long as the region is contained within the field. On the other hand, if the region size exceeds the size of the field within which it was originally defined, more than one level of deterioration will be introduced within the same region, and, therefore, the estimate would not be an accurate representation of either level (i.e., the estimate is biased). Consequently, this results in an erroneous condition assessment.

Therefore, the best aggregation scheme is a configuration in which each macroscopic field is a region. The problem at this point lies in the lack of knowledge on the locations of the boundaries defining the fields of interest. Moreover, since the environmental variables at the level of detail of interest are either unavailable or unreliable from a measurement point of view, the most useful indicators of the boundaries defining the fields are the distress measures over space. From a methodological point of view, the problem reduces to the identification of the locations along the facility where the *mean* of the stochastic process undergoes an abrupt change.

Field Identification

Due to the insignificance of the lane width with respect to the longitudinal scale of interest, the two-dimensional nature of a highway lane is approximated as one-dimensional. Furthermore, since, in general, the longitudinal extent of discrete distress types such as cracking is small [averaging between 2 and 10 m in the cases examined—see Mishalani (*7*)] with respect to the scale of interest, for purposes of this analysis distress types are characterized by:

- Point events in continuous one-dimensional space (cracking, potholes); and
- Continuous variables in continuous one-dimensional space (rutting).

The point representation naturally lends itself to an analysis using spatial point processes. Point processes are a type of stochastic processes in which the events of interest are points occurring randomly in continuous space. Using point processes provides a convenient means for representing point data by a distress intensity function in space. The intensity function, $x(s)$, is the number of events per unit length for each location in continuous space and, therefore, represents the propensity of a particular location to exhibit a point event. Using the original point events, $x(s)$ is estimated using a nonparametric kernel-based estimator (*14*) that minimizes the mean square error. The estimator is given by:

$$\hat{x}(s) = \sum_{j=1}^{C} \frac{1}{w}\,\delta\!\left(\frac{S - S_j}{w}\right) \quad (2)$$

where

$\delta(\cdot)$ = kernel function that specifies the relative strength by which the existence of an event at s_j contributes to the estimate of the intensity at s;
s_j = location of event j;
C = total number of events; and
w = a parameter representing half the window width within which events contribute to the estimation of $x(s)$.

The statistical properties of the estimator $\hat{x}(s)$ are dependent mostly on the parameter w. If w is large, more events are used in the estimation but the role of each as an indicator is less significant. On the other hand, if w is small, the more significant events are used but there are fewer of them. Therefore, the effect of w on the estimates relates to efficiency and bias. Large values of w are associated with both high efficiency (i.e., low variance) and large bias, whereas small values of w are associated with low efficiency and small bias.

The functional form of the kernel is not important from a statistical point of view. Hence, for mathematical convenience, the uniform kernel is used:

$$\delta(y) = \begin{cases} 1/2 & \text{for } |y| \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Using the function above, an optimal parameter w^*—which minimizes the mean square error of the estimator $\hat{x}(s)$—can be determined. The determination of the optimal window width, $2w^*$, also allows for conveniently representing rutting over space. See Mishalani (*7*) and Mishalani and Koutsopoulos (*15*) for more detail on representing distress over space.

Once distress functions (crack and rut intensity functions) are quantified, the identification of the boundaries defining the fields can proceed. The probability density function of X_s is unknown. Therefore, a nonparametric solution approach for the detection of the boundaries is adopted. The field identification problem is formulated within a cluster analytic framework with the additional constraint that all observations belonging to the same cluster should be spatially contiguous. The boundaries of the clusters represent locations where the systematic mean function, μ_s, undergoes an abrupt shift in value.

Formulating the optimal clustering problem in a manner where the number of fields and their boundaries are determined simultaneously is mathematically intractable. Therefore, the following heuristic is proposed:

Step 1: Set the number of fields, M, to 2.

Step 2: Determine the locations of the optimal boundaries of the M fields.

Step 3: Examine the stopping criterion (this step is applicable for $M \geq 3$). If the criterion is satisfied, the optimal solution with $(M - 2)$ fields is the final solution. Otherwise, set M to $M + 1$ and go to Step 2.

The main idea behind this iterative approach is to reduce the complexity of the problem by making a tentative assumption on the number of fields. This allows for an optimal determination of the location of the boundaries. This process is repeated in an iterative manner until the stopping criterion is satisfied. The purpose of the stopping criterion is to determine which of the various sets of solutions best captures the spatial behavior.

The decision variables of the optimization problem of Step 2 are the locations of the boundaries. The objective is to minimize the *total within cluster variation*. Therefore, the boundaries will optimize a measure of similarity within the fields (i.e., observations within each cluster are as close as possible to their cluster mean). The additive structure of the objective function allows for the use of a dynamic programming algorithm (*16*) that guarantees a globally optimal solution.

The stopping criterion should indicate the number of fields that best characterize the spatial behavior. It consists of examining the incremental contribution of every additional boundary to the reduction in the overall variation throughout the facility. Let $Z^*(M)$ be the value of the objective function when all the observations over space are optimally partitioned into M clusters. $Z^*(M)$ is monotonically decreasing with M. $Z^*(1)$ captures the total variation of the intensity function with respect to the overall mean. The incremental contribution of the Mth additional field to explaining the total variation (or equivalently to the reduction in the objective function) is measured by the following ratio:

$$r(M) = \frac{Z^*(M-1) - Z^*(M)}{Z^*(1)} \tag{4}$$

The stopping criterion is satisfied when the change in the incremental contribution becomes insignificant.

APPLICATION AND EMPIRICAL FINDINGS

The developed methodology has been applied to three independent flexible pavement highway facilities, each of which is approximately 15 km long. All three applications reveal similar results. One of the three applications is presented in this paper. The facility is an undivided two-lane (one lane in each direction) highway in Mississippi. Its Federal Functional Classification is a rural major collector. It was constructed between 1947 and 1948 and was resurfaced once in 1984–1985. The current average daily traffic is 1,040 vehicles, 10 percent of which is truck traffic. The original design has a total pavement thickness of 17.8 cm to 43.2 cm. The resurfacing resulted in an additional 2.5 cm. Detailed surface distress data were collected by PaveTech (Oklahoma) on one of the two lanes during the summer of 1991 using a van-mounted state-of-the-art video camera. The video images were interpreted by experts but the distress measurements, such as crack lengths and areas, were quantified automatically using image-processing techniques.

The mean length (in the longitudinal direction) of the cracks is 2.53 m. Moreover, almost 99 percent of the cracks have lengths less than or equal to 30 m. Therefore, in the context of the expected macroscopic nature of the environmental process (where fields are expected to be several kilometers long), the assumed point representation of cracks is, in general, realistic. The optimal window width, $2w^*$, is 50 m. In this particular application, the crack intensity function reveals all the fields. Therefore, results focusing on cracking are presented. See Mishalani and Koutsopoulos (*15*) for a presentation of the results relating to rutting.

Figure 1 depicts the contribution ratio as a function of the number of fields M. It is clear that for $M \geq 7$, the incremental contribution ratio is consistently low suggesting that all fields beyond the first six overfit the spatial model. Therefore, there are six significant fields that explain the spatial behavior. This conclusion is further confirmed by the examination of the location of the boundaries associated with the six fields. The optimal boundaries along with the crack intensity function are indicated in Figure 2. The horizontal lines below the plot indicate the solutions where a vertical bar represents the location of a field boundary. Two solutions are given: the optimal solution and the solution of the Agency Method. Notice that the optimal boundaries (first line in Figure 2) reveal the piecewise constant nature of the mean of the crack intensity function. With six fields, the average field length is 2.5 km. This is consistent with the engineering expectation that the macroscopic environmental process (which is captured by the fields) exhibits such a scale.

Another important conclusion that can be drawn relates to the spatial extent of the interactive mechanism along with the microscopic environmental mechanism. This is achieved by examining the spatial correlation structure within each field m. The spatial correlation function is defined by $\rho_{mh} = \mathrm{Cov}[X_{ms}, X_{m(s+h)}]/\mathrm{Var}[X_{ms}]$. The sample estimate of the correlation function, $\hat{\rho}_{mh}$, indicates the structure of the linear dependence exhibited by the data. The positive spatial correlation exhibited within each field confirms the presence of the microscopic interactive and environmental mechanisms. The extent of this correlation is captured by the largest separation distance at which the spatial correlation within each field significantly

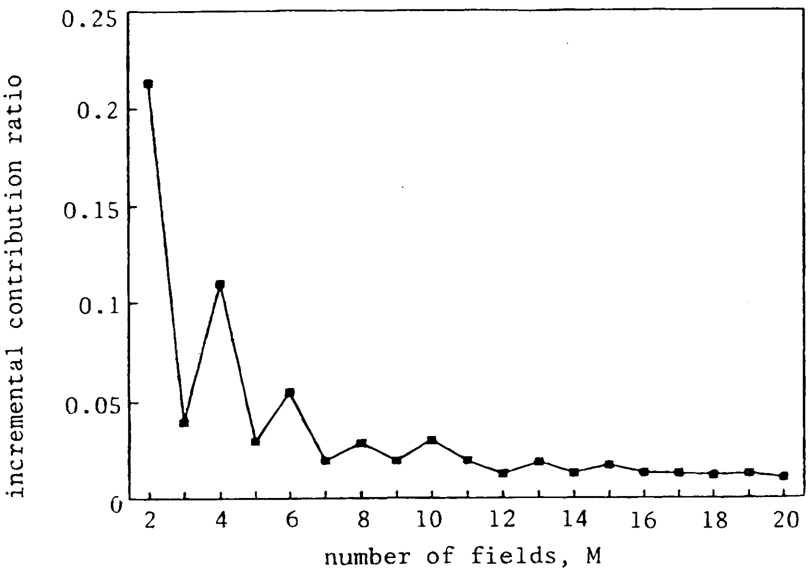

FIGURE 1 Incremental contribution ratio function.

differs from zero. That distance for all the fields of all three facilities examined varies between 20 m and 180 m considering both cracking and rutting. Figure 3 presents the estimate within Field 6 (i.e., the region from 11.53 km to 14.48 km). Notice that the correlation becomes insignificant at a separation greater than 50 m. Recall the range of interaction assumed by Paterson (*11*) at 320 m based on engineering judgment. The range of interaction identified based on this empirical study is of the same order of magnitude. However, the empirical analysis reveals a high degree of variability in the exhibited ranges and, in general, smaller magnitudes.

Having demonstrated empirically the validity of both the developed spatial model and the corresponding field identification methodology, it is worth examining the performance of the state-of-the-art methods in light of the findings of this study. Figure 2 indi-

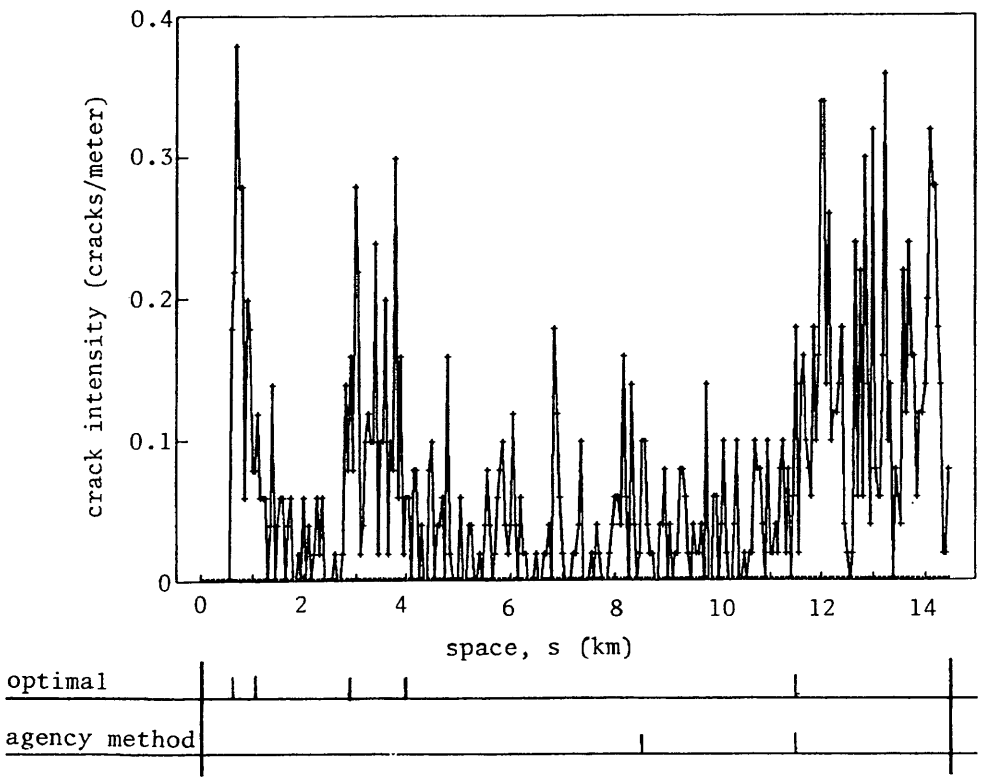

FIGURE 2 Field boundary solutions.

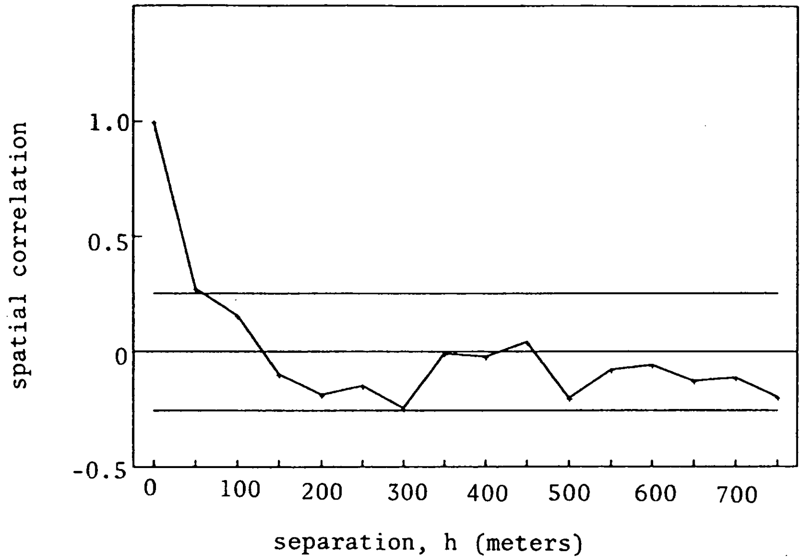

FIGURE 3 Estimate of Spatial Correlation Function Within Field 6 and 95 percent confidence region.

cates the solution of the Agency Method (on the second horizontal line below the plot). The boundary at 8.43 km is due to differences in the original design and the boundary at 11.79 km is due to a 1-year difference in the timing of the resurfacing activity. The Agency Method was able to identify only one of the boundaries detected by the developed field identification methodology. In addition, however, the Agency Method has identified a boundary at 8.43 km that the field identification methodology did not capture. One potential explanation follows. The resurfacing that was applied only a year before the data were collected is concealing the change in the distress manifestation that one expects to see. Even though the boundary was not evident at the time the distress data were collected, once further deterioration is allowed to take place the boundary is expected to become apparent.

Since one of the fundamental objectives behind this study is to provide the appropriate inputs for condition assessment, deterioration modeling, and subsequently decision making, one of the more important such inputs is computed using both the optimal fields and the Agency Method regions. The percentage area cracked is indicated in Table 1 using both sets of boundaries. The percentage difference between the Agency Method estimates and the optimal estimates (using the fields as basis) is also shown. One can see that the Agency Method appreciably underestimates the percentage area cracked in the region of Fields 2 and 4 and appreciably overestimates it in Fields 1 and 5. The Agency Method regions result in a contamination of the observations from one field with observations from other fields of different mean levels. The Agency Method overestimates or underestimates the percentage area cracked by at least 20 percent in regions comprising 65 percent of the total length. Such differences could potentially result in not maintaining regions that do need maintenance, thus resulting in a rapid deterioration that would require more costly maintenance in the future. It could also result in overspending on unnecessary maintenance work. In either case, the life cycle costs associated with the facilities would be suboptimal.

As for the AASHTO method, it is evident that under the highly stochastic nature of the spatial distress process revealed in this paper, it would result in too many small regions that would substantially over- or underestimate the mean distress level. Moreover, assuming for the sake of this argument that the small-scale vari-

TABLE 1 Comparing Percentage Area Cracked of Optimal Fields with Those of Agency Method

Optimal field number	1	2	3	4	5		6
From (km)	0	0.625	1.025	2.825	3.975		11.525
To (km)	0.625	1.025	2.825	3.975	11.525		14.48
% area cracking	0.69	26.02	5.57	7.94	2.83		11.83
% difference[a]	581.16	-81.94	-15.62	-40.81	66.08	34.63	0
Agency region	1					2	3
From (km)	0					8.43	11.525
To (km)	8.43					11.525	14.48
% area cracking	4.70					3.81	11.83

[a]Using the field measures as a reference.

ability is insignificant with respect to the piecewise constant mean function (which is definitely not the case in reality), as already discussed the AASHTO method does not guarantee the identification of all the boundaries of abrupt change. Hence, the usefulness of the AASHTO method is limited.

CONCLUSION

In this study, a model that describes the spatial behavior of infrastructure distress is established. The model captures both the macroscopic and microscopic scales of behavior. The macroscopic scale is associated with the environmental deterioration mechanism, and the microscopic scale is associated with both the environmental and interactive deterioration mechanisms. Based on the spatial distress model, a methodology that identifies the uniformly behaving spatial fields is developed. The spatial model developed is validated using detailed data along 15-km-long facilities (one of which is presented in this paper). Such an empirical analysis is the first of its nature in the context of infrastructure research.

The spatial model developed plays an important role not only in condition assessment and deterioration modeling but also in addressing a host of other issues within the infrastructure management process. For example, the past decade has witnessed the adoption of automated technologies by infrastructure agencies resulting in significant productivity improvements in relation to manual surface distress data collection processes. This, in turn, provides the opportunity to collect detailed data across the vast lengths of the facilities. This large amount of data poses a potential problem to agencies since their level of detail is not compatible with the scale of interest for maintenance application. The spatial model developed provides the necessary structure for aggregating the detailed distress data in a meaningful manner without any loss of information. Moreover, the spatial model allows for determining the optimal sampling schemes. This results in a more representative data collection by the new technologies. Finally, the spatial model can be used to identify systematic measurement errors resulting from unexpected exogenous factors. However, further research is required to fully understand the use of the spatial model in the two latter applications.

ACKNOWLEDGMENT

Support was provided by the Army Research Office through the Center for Construction Research and Education at Massachusetts Institute of Technology and the United Parcel Service Foundation through the Center for Transportation Studies at Massachusetts Institute of Technology.

REFERENCES

1. Ben-Akiva, M., F. Humplick, S. Madanat, and R. Ramaswamy. Infrastructure Management under Uncertainty: Latent Performance Approach. *Journal of Transportation Engineering*, Vol. 119, No. 1, 1993, pp. 43–58.
2. Ramaswamy, R. *Estimation of Latent Pavement Performance from Damage Measurements*. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1989.
3. Madanat, S. Optimal Infrastructure Management Decisions Under Uncertainty. *Transportation Research C*, Vol. 1C, No. 1, 1993, pp. 77–88.
4. *AASHTO Guide for Design of Pavement Structures*. American Association of State Highway and Transportation Officials, Washington D.C., 1986.
5. Maser, K., B. Brademeyer, and R. Littlefield. Pavement Condition Diagnosis Based on Multisensor Data. In *Transportation Research Record 1196*, TRB, National Research Council, Washington D.C., 1988, pp. 62–71.
6. Roddis, W. M. K., K. Maser, and A. Gist. Radar Pavement Thickness Evaluations for Varying Base Conditions. In *Transportation Research Record 1355*, TRB, National Research Council, Washington D.C., 1992, pp. 90–98.
7. Mishalani, R. G. *Extracting Spatial Information on Infrastructure Distress Using a Stochastic Modeling Approach: Application to Highway Maintenance Decision Making*. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1993.
8. Alfelor, R., and S. McNeil. Heuristic Algorithms for Aggregating Rail Surface Defect Data. *Journal of Transportation Engineering*, Vol. 120, No. 2, 1994, pp. 295–311.
9. Moavenzadeh, F., and B. Brademeyer. A Stochastic Model for Pavement Performance and Management. *4th International Conference on Structural Design of Asphalt Pavement*, Ann Arbor, Mich., *Proc.*, Vol. 1, 1977, p. 876.
10. Markow, M., and B. Brademeyer. *Modification of the System EAROMAR: Final Technical Report*. FWHA-RD-82-086, CTM Inc., 1984.
11. Paterson, W. D. O. *Road Deterioration and Maintenance Effects: Models for Planning and Management*. Johns Hopkins University Press, Baltimore, Md., 1987.
12. Herrmann, H. J., and S. Roux. *Statistical Models for the Fracture of Disordered Media*. Elsevier Science, Amsterdam, The Netherlands, 1990.
13. Humplick, F. Highway Pavement Distress Evaluation: Modeling Measurement Error. *Transportation Research B*, Vol. 26B, No. 2, 1992, pp. 135–154.
14. Diggle, P. A Kernel Method for Smoothing Point Process Data. *Applied Statistics*, Vol. 34, No. 2, 1985, pp. 138–147.
15. Mishalani, R. G., and H. N. Koutsopoulos. Infrastructure Condition Assessment: Definition and Identification of Uniformly Behaving Spatial Fields. *Journal of Infrastructure Systems*, Vol. 1, No. 1, 1995, pp. 44–55.
16. Hartigan, J. A. *Clustering Algorithms*. John Wiley and Sons, Inc., New York, 1975.