THE PENNSYLVANIA DOT TRAINED OBSERVER SURVEY: DESIGN AND PRELIMINARY RESULTS

Theodore H. Poister, Pennsylvania State University
William R. Moyer, Pennsylvania Department of Transportion

This paper discusses the design and initial findings of a trained observer survey developed for the Pennsylvania Department of Transportation. With a shift in priorities from construction to maintenance it became clear that some systematic, objective basis for assessing the condition of the state's 45,000 miles of highway was needed for analyzing needs and monitoring program performance over time. The trained observer approach was chosen because it could be tailored to fit the varied concerns of the maintenance program, and the instrument was designed to provide an intensive examination of surface, foundation, shoulders, drainage and appurtenances. Preliminary findings show that there is widespread variation in conditions among road sections across the state with high percentages of deficient roads on many items. To some extent this variation is attributable to systematic differences by Maintenance Functional Code, pavement type, and district and county. Sample reliability with the initial 3 percent sample is fairly weak and sample size will have to be expanded on subsequent cycles of data collection; however, the survey's reportable condition variables appear to be more reliable than PSI measures in discriminating good roads from bad roads.

This paper discusses the development of a Trained Observer Survey of road conditions for the Pennsylvania Department of Transportation and presents some initial findings from the first round of data collection. The survey was designed by university researchers, working in conjunction with staff from the Bureau of Maintenance and Operations Review Group, as part of a more comprehensive effort to develop systems for monitoring the performance of the Department's highway programs.

As a result of a worsening fiscal crisis which made it impossible to carry on a large new construction program and adequately maintain the existing system in the face of a growing backlog of resurfacing projects, the Department's priorities were shifted away from construction to maintenance of the 45,000 mile system. (1) Concurrently, a new administration came into office determined to economize where possible, use resources more effectively and upgrade a highway system which had become a shambles. (2) Part of the approach of this new management-control oriented administration rested on the concept of developing performance indicators to monitor trends in activities, outputs and effects on a continuing basis. With the heavy priority on maintenance, monitoring the efficiency of maintenance activities and the results of these efforts was essential.

Early on in the performance indicator project, it became apparent that this system had to include a means of periodically assessing the condition of state owned roads across the state. While the preferred approach was to utilize existing reporting systems as much as possible--such as using the Highway Maintenance Management System for efficiency analysis--the Department simply did not have any systematic, comprehensive road condition survey in use. Thus, a major part of the project has been devoted to developing and testing such a survey, to be conducted on an ongoing basis. The trained observer approach was chosen because it could be tailored to fit the varied concerns of the maintenance program and because it could be geared up fairly quickly with no major capital outlay. This paper describes how the survey works and illustrates the kinds of analysis for which it can be used.

Trained Observer Survey

The trained observer survey entails the use of trained professionals to physically inspect a number of conditions on a sample of highway sections in each of the state's 67 counties on a periodic basis. Each sample section is also run with a Mays Meter to obtain an indication of roughness. As with the system currently in use in Ohio, the number of "reportable conditions" of deficiencies in surface, foundation, drainage, shoulders, and other safety features which are observed can be used to obtain an overall rating of the "as built" maintenance condition of the roads in each county and district. (3,4)

The data generated by this system are intended to serve two purposes relating to needs assessment and performance monitoring. First, analysis of the numbers of reportable conditions and overall maintenance ratings should provide an improved basis for allocating funds for maintenance activities. The results should show where the greatest needs are for maintenance efforts and permit a better allocation of resources among counties or districts and a more efficient targeting of funds within counties or districts by type of highway and by maintenance cost function.

Secondly, the trained observer system will provide the Department with a means of tracking the effectiveness of the maintenance program. Since highways in each county will be sampled for observation on a periodic basis, this system will accumulate time series data which can be examined to determine whether the number of certain types of reportable conditions is being decreased or the overall maintenance ratings improved over time. Such measures of change over time can also be correlated with the amount of selected maintenance activities conducted and their associated costs to examine the effectiveness and efficiency of these activities.

Instrument Design

In determining what kinds of deficiencies would be included in the trained observer survey and defining the reportable conditions to be noted, three criteria were taken into account:

1. The system should be comprehensive in terms of highway maintenance concerns and should include measures representing conditions of major structural characteristics and major safety features.
2. The list of reportable conditions should include those which relate to the Department's high cost maintenance activities (excluding winter maintenance) and should only include deficiencies to which the Department responds. For example, the condition of surface flushing would not be included because the Department does not take action intended to directly counteract this problem.
3. The measuring instrument itself must be reliable, based on observable, tangible characteristics or conditions rather than impressions. To assure inter-observer reliability, the system must be based on clearly defined reportable conditions which would be observed and counted the same by different individuals.

Offsetting these criteria for valid measurement is the need for an instrument which is not too cumbersome to use. The system needs to be workable in the field, and an unduly complicated set of measures will make the observers' work too tedious and time-consuming. During pre-testing, the instrument was simplified and streablined to make the fieldwork smoother and quicker.

Figure 1 shows the set of deficiencies included in the survey. They reflect the major structural and safety concerns of highway maintenance as grouped in the general categories of roadways (surface and foundation), shoulders, drainage and appurtenances. Figure 1 also shows the individual reportable conditions which are counted as indicating given types of deficiencies. For example, surface deterioration may be indicated by dust layering, slopes of greater than ½ inch per foot, depressions, minor cracking or "mapcracking", and gaps in transverse or longitudinal joints. Specific definitions of each type of deficiency have been established as illustrated in Figure 2 and the observers have been drilled and trained in applying them before beginning actual fieldwork.

An attempt has been made to minimize problems of inter-observer reliability by establishing limits on the amount of each type of deficiency which counts as a reportable condition. The prevalence of each type of reportable condition is to be measured by the frequency with which the observers see that condition along the sample stretch of highway they are inspecting. The definition of each type of reportable condition specifies a minimum--how much of that condition must be observed in order to be counted. For instance, minor cracking must cover at least one square foot in order to count.

In addition, a unit count is specified for each specific type of reportable condition. If a given deficiency extends for more than this unit count, it counts as more than one reportable condition. For example, the unit count for minor cracking is every 25 lineal feet. Thus, if the observers encounter mapcracking of one square foot or more which extends 12 lineal feet, this counts as one reportable condition. However, if the mapcracking were more extensive and found to cover 60 lineal feet, this would be counted as three reportable conditions. With this system, the observers have less responsibility for determining the severity of the problems they come across.

It should also be noted that some degree of continuity among the different reportable conditions is built into the system. For instance, small depressions or holes less than 2 inches deep count as surface deterioration while potholes 2 inches or more in depth count as surface obstructions as long as they are not bigger in area than 2 square yards. Large potholes or depressions covering 2 square yards or more count as areas that are broken up or have mud surfacing through units of foundation failure.

In addition to visually inspecting the sample highway segments and recording the observed reportable conditions, crews are riding each segment in a vehicle equipped with a May's Ride Meter to obtain an indication of roughness. The resulting roughness is converted into a Present Serviceability Index (PSI) which can be thought of as an indicator of rideability and service as opposed to a structural or safety indicator. The Department has been using May's Meters on selected roadways and relying primarily on PSI measures to determine the need for resurfacing. Rideability is only one aspect of road condition, however, and the trained observer counts on a variety of conditions should greatly improve the ability to discriminate good roads from bad roads.

Sample Selection

The trained observer survey is to be conducted twice each year, once in the spring and once in the fall, with a "floating" sample; in other words a new random sample of road segments will be selected for observation each time. While the sampling fraction must be low--given time and cost considerations and the fact that the total system contains some 45,000 miles--over time data will build up on most of the system. The use of floating samples is advisable in order to obtain more complete coverage over the span of a few observation periods and to avoid possible problems arising from the targeting of maintenance activities on those few segments known to be included in a stationary sample.

Given the purposes of the trained observer system as aiding in the allocation of maintenance funds and monitoring the progress of the maintenance program across time, the development of a sample selection procedure was based on the following criteria:

1. The overall sample should be approximately representative of the statewide highway network in

Figure 1. Reportable conditions - trained observer survey.

Components	Deficiencies	Reportable Conditions	Unit Counts
ROADWAYS	Surface Deterioration	Dust Layering	25 LF
		Slope - ½"/ft.	25 LF
		Depressions	25 LF
		Minor Cracking	25 LF
		Joints	25 LF
	Surface Obstructions	Potholes	Each
		Foreign Objects	Each
		Blowups	Each
		Virginia Joints	Each
	Foundation Failure	Soft Spots	25 LF
		Major Cracking	25 LF
		Broken up or Mud	25 LF
		Bituminous Patch	25 LF
SHOULDERS	Deterioration	Slope - ½"/ft.	25 LF
		Depressions	25 LF
		Minor Cracking	25 LF
		Raveling	25 LF
		Buildup	25 LF
	Obstructions	Potholes	Each
		Foreign Objects	Each
		Washouts or Slides	Each
		Bad Drives	Each
	Failure	Major Cracking	25 LF
		Rutted	25 LF
		Broken up or Mud	25 LF
	Drop Off	Edge Pavement 2"	100 LF
		Edge Pavement 4"	
DRAINAGE	Obstruction	Non-functional Ditch	100 LF
		Non-functional Inlet	Each
		Pipe ½ Inlet	Each
	Failure	Bad Pipe	Each
		Broken Inlets	Each
		Non-functional Endwalls	Each
APPURTENANCES	Guard Rails and Median Barriers	Bad Stripping	500 LF
		Rotted Posts	Each
		Non-functional Elems.	100 LF
		Median Barrier	100 LF
	Signs	Regular Signs	Each
		Delineators	Each
		L.R. or Station Markers	Each
	Litter	Litter	25 LF

terms of the distribution across the 67 counties and the 11 Engineering Districts.

2. Within individual counties the samples should provide for fairly precise estimates of the number of reportable conditions for the county as a whole.

3. While the samples will not be very reliable for a given type of road for any given county, the overall sample should provide reliability for a given maintenance functional code on a district basis.

Given these criteria it was decided that the most appropriate type of sample would be a random sample of the approximately 2,600 observable segments stratified by county and Maintenance Functional Code (MFC). Thus a separate sample is selected from each county, and within each county care is taken to include sufficient examples of each road type. Since there are relatively fewer miles of the higher classes of highways, they have to be sampled more heavily in order to have good sample reliability on a district basis. Therefore, while collectors and local roads were sampled at 2% for cycle 1 of the survey, the sampling fraction for Interstates was 15%.

On the average, the length of the highway sections that are inspected is approximately one half mile. To generate the sample, the highway sections entered in the road log were combined or subdivided where necessary to form segments from .3 to .7 miles in length. In combining small sections care was taken not to mix very dissimilar roads in the segments which would be used to draw the sample. Thus, entries in the road log could be combined only if they were part of the same legislative route, had the same urban/rural status, were classified the same according to the maintenance functional code, and fell within the same general range of average daily traffic (ADT). The ranges of ADT which were established by Bureau of Maintenance personnel were 1 - 1,499 vehicles, 1,500 - 2,999 vehicles, 3,000 - 4,999 vehicles, and 5,000 vehicles and above. Highway sections with ADT's of more than 5,000 vehicles

Figure 2. Sample definitions of reportable conditions.

Rigid Base Roads - Include all highways with a base course of plain or reinforced cement concrete and a wearing course of bituminous concrete, brick, or block.

Surface Deterioration

Minor cracking - The presence of irregular cracks, less than 1/8 inches wide, often referred to as map-cracking, covering at least one square foot of area. Unit count - every 25 lineal feet.

Depression - Any depression hole or corrugation which is greater than ½ inch and less than 2 inches deep and at least one square foot in area. Unit count - every 25 lineal feet.

Joints - Any gap in a transverse or longitudinal joint at least one foot long which needs to be filled. Also, any crack greater than ¼ inch wide and more than one foot long, which need to be filled. Unit count - every 25 lineal feet.

could be combined only if the difference in ADT did not exceed 10,000 vehicles.

From the resulting file of these .3 to .7 mile segments, then, roads were randomly selected for a sample which is stratified proportionally by county and disproportionally by MFC. The total mileage and number of sections sampled for each MFC is shown in Figure 3.

Figure 3. Sample selection statistics.

MFC	Total Mileage	Sampling Fraction	Sections Observed
Interstate	1,060	15%	339
Principal Arterials	4,060	8%	633
Minor Arterials	8,460	2.5%	399
Collectors	17,700	2.0%	728
Local	12,780	2.0%	482
Total	44,060	3.0%	2,581

This overall 3% sample clearly contains a substantial number of highway segments to be observed; yet if there is wide variation in the rate of reportable conditions, a larger sample may be needed to draw valid conclusions on a district by MFC basis. On the basis of the results of Cycle 1 of the survey it may well be determined that the sample size must be expanded and/or reallocated among MFC classes for subsequent cycles.

Conduct of the Field Work

The definitions of reportable conditions were field tested by piloting the program in Lycoming County with two Engineers from Operations Review Group. Lycoming County is a rural County in North Central Pennsylvania with one major city, Williamsport, and 869 miles of State highway. It was felt that this was a representative County; even though it has no Interstate it has several major four lane roads. On two lane roads only one direction was walked and observations extended only to the center line. On three and four lane roads, both directions were walked and the data combined for one section. Upon completion of the pilot representatives from the Bureau of Highway Maintenance, Penn State, and the Operations Review Group reviewed the results and made numerous changes. First, several definitions were revised to reflect actual conditions and make reporting easier, and unit counts were standardized for ease in observing and reporting. Directions of walk were also set to alternate to provide greater randomness and actual representation.

A two-man team procedure for performing observations and recording data was established for the following reasons: (1) Use of a two-man team with two vehicles and parking one at each end of the section saved considerable time lost by walking back to the vehicle at the original starting point. (2) Many of the reportable conditions are based on frequencies or occurrences within a 25 or 100 foot length. Through experimentation the easiest method of tracking the distances was found to be with the use of a measuring wheel. This requires one member of the team to use the wheel and note distances and conditions while the other member records the counts on the chart and is observing the conditions. A set of standard packages was developed for each county for use by the trained Observers including: (1) Straight line diagrams for each section of roadway to be examined to help team members in locating roads in counties lacking station and/or legislative route markers, (2) County maps with all sections plotted in color to identify approximate locations of sections as an aid to the team in planning a days work with the minimal amount of travel time between sections, and (3) Copy of tabulation of sections by MFC, starting and stopping station, length in feet and miles. Finally, it was decided to inspect bridges up to a 20 foot span length because this was the actual cut-off length of the Bridge Inspection Team.

It was anticipated that a minimum of five two-man teams would be needed to complete the work before the Winter. New employees worked with the two engineers who did the pretesting until they were proficient and then as we hired new employees they were placed with experienced employees. One of the engineers who performed the pilot in Lycoming County worked with the program as the coordinator during the entire cycle. He or his supervisor was available to answer any questions concerning interpretations, definitions, etc., via phone and periodically visited crews in the Counties.

An attempt was made to overcome weather problems by starting with the Counties in the Northern and Western part of the State and working generally toward the Southeast corner. Whenever possible team members were rotated to eliminate the same two people working together constantly. This helped to ensure uniform interpretation and application. At the completion of the cycle and prior to release of the temporaries, a two-day seminar was held to discuss each of the definitions and conditions and identify any potential problem areas that required attention before the next cycle.

The cost for the first cycle of field operations was approximately $200,000. For the second cycle efficiency has been increased by plotting all counties prior to starting field work and dividing the State into five geographical areas. At the start of the cycle each team was provided all the packages for their assigned counties. This permits them to reduce wasted time and perform isolated sections with adjacent counties or perform a complete days work on Interstates. Furthermore, data processing has been streamlined. The tally sheets for the current and subsequent cycle have been computerized to reduce handwork and transpose information. All headings and pertinent information is now printed when the computer selects the sample.

Results

The preliminary findings from the trained observer survey presented here concern ratings of existing road conditions as well as estimates of total needs for maintenance work. The results pertain to statewide conditions and also show comparisons across districts and counties. While most of the data are purely descriptive, some attempt is made here to apply statistical tests and evaluate sample reliability.

Overall Conditions

Figure 4 shows aggregate statewide frequency distributions for four illustrative reportable conditions: surface depressions, minor cracking on the surface, nonfunctional ditches and nonfunctioning guardrail elements. The incidence of all conditions is reported on a count per mile basis. The most striking feature of these distributions is that they reflect widespread variation and are highly skewed to the right. While some of the 2,581 road sections that were inspected had no surface depressions, for example, some had a moderate-to-heavy incidence of depressions and a few had the maximum possible, 211-220. On these highways depressions were encountered in every 25 foot stretch that was observed.

That these reportable conditions are highly variable--and this characterizes almost all the items included in the survey--is further evidenced by the fact that the standard deviations exceed their respective mean averages by a considerable margin. The high degree of dispersion weakens sample reliability and makes it more difficult to compare means of obtaining precise interval estimates. This is a strong indication that sample sizes should be expanded to future cycles of the survey.

The one-sidedness of these distributions along with the skews toward extremely high values also makes the mean average less reliable as an indicator of central tendency. Median averages are preferable as a more accurate indication of what values the more typical roads take on for various conditions. It is also the case that on many of the indicators a significant number of roads have zero counts. Thus, the percentage of road sections with some counts, the percent deficient roads, is also a measure which is worthwhile to look at. To concentrate on the extremely high values on the basis that in some respects the worst roads in a county or district are of greatest concern, the eightieth or ninetieth percentile might also be useful.

As can be seen in Figure 4, then, surface depressions and minor cracking have similar distributions ranging from zero to the maximum, although a substantially higher percentage of all roads evidenced depressions (small potholes) than minor cracking. Comparing the two medians also indicate that depressions were observed more than minor cracking. Figure 4 shows that only 18.4% of the highway sections that were inspected had nonfunctional guardrail elements and that the median was only slightly above .1 observed count per mile. The drainage condition of a pipe being half-full or more was encountered in nearly 40% of all the sections, and the median was .33 counts per mile. The overall range of this condition is much less than many others because along many stretches of highway there are simply no pipes to observe.

The wide variability in the condition of Pennsylvania's road network is not surprising considering the inadequacy of funding of the maintenance program over the past decade and the wide range of roads included in the system. Unlike most states, whose responsibility covers only trunk-line highway systems, the Pennsylvania DOT system also incorporates many lower order collectors and farm-to-market and small borough "local use" roads which are not likely to stack up as well in terms of many reportable conditions. Given this context, the distributions seen in Figure 4 are what would be expected as a profile of the condition of the state's roads. Now that this profile is known, it raises questions about the preferred strategy of an improved maintenance program: should the Department seek to steepen the skew by upgrading the good and moderately good roads to zero deficiency roads or should it concentrate on trying to eliminate roads in the worst possible condition?

Maintenance Classification Comparisons

Since programming and work planning are geared in part to functional class of highway, it is worthwhile to make initial comparisons of reportable conditions by MFC. Figure 5 shows the percentage of deficient roads on selected surface conditions as well as median averages. On Interstates and some other higher order roads the percent deficient will be overstated relative to lower MFC classes because lane highways were walked in both directions. There is little systematic variation in percent deficient in depressions except that local roads have more deficient roads than the higher categories. However, looking at medians it is apparent that the typical Interstate has fewer depressions per mile than other roads and that this indicator increases with the lower order MFC's. This pattern holds for "all potholes" which includes depressions, potholes, and broken up areas with mud surfacing through. Minor cracking and major cracking are more prevalent on the lower order roads than on interstates and principal arterials as would be expected given maintenance priorities. However, the incidence of joint deterioration varies in the other direction, from 44 counts per mile on the interstates and 33 per mile on principal arterials to almost nonexistent on collectors and 6% on local roads, largely because joints are only observed on rigid base and rigid pavement roads. Yet the 92% deficient on interstates and 73% on principal arterials would have to be considered unacceptable.

As seen in Figure 6, insufficient shoulder slopes are mainly a problem on collectors and local roads, largely because this condition occurs much less frequently with paved shoulders. Shoulder buildup varies in the opposite direction, perhaps reflecting the greater use of antiskid material on the higher order roads. It is interesting to note that although some shoulder buildup is encountered on a fairly high percentage of roads, the problem is not severe in that on the average there is substantially less than one count per mile. Surprisingly, washouts have their lowest incidence on local roads and Interstates and are more prevalent on arterials although overall the counts are low. Edge of pavement dropoffs greater than 4 inches have the lowest incidence on interstates, as expected, and do not vary substantially by MFC otherwise. The percent of roads with major cracking on the shoulders is substantially higher for Interstates--at 6 counts per mile--than for other roads which have fewer concrete and bituminous shoulders.

With the exception of interstates, the percent of roads with nonfunctioning ditches is greater with the lower order roads, as shown in Figure 7. On the other hand, the percent of roads with nonfunctioning inlets as well as median averages tend to decrease with the lower order roads, as they have fewer pipes.

Figure 4. Frequency distribution.

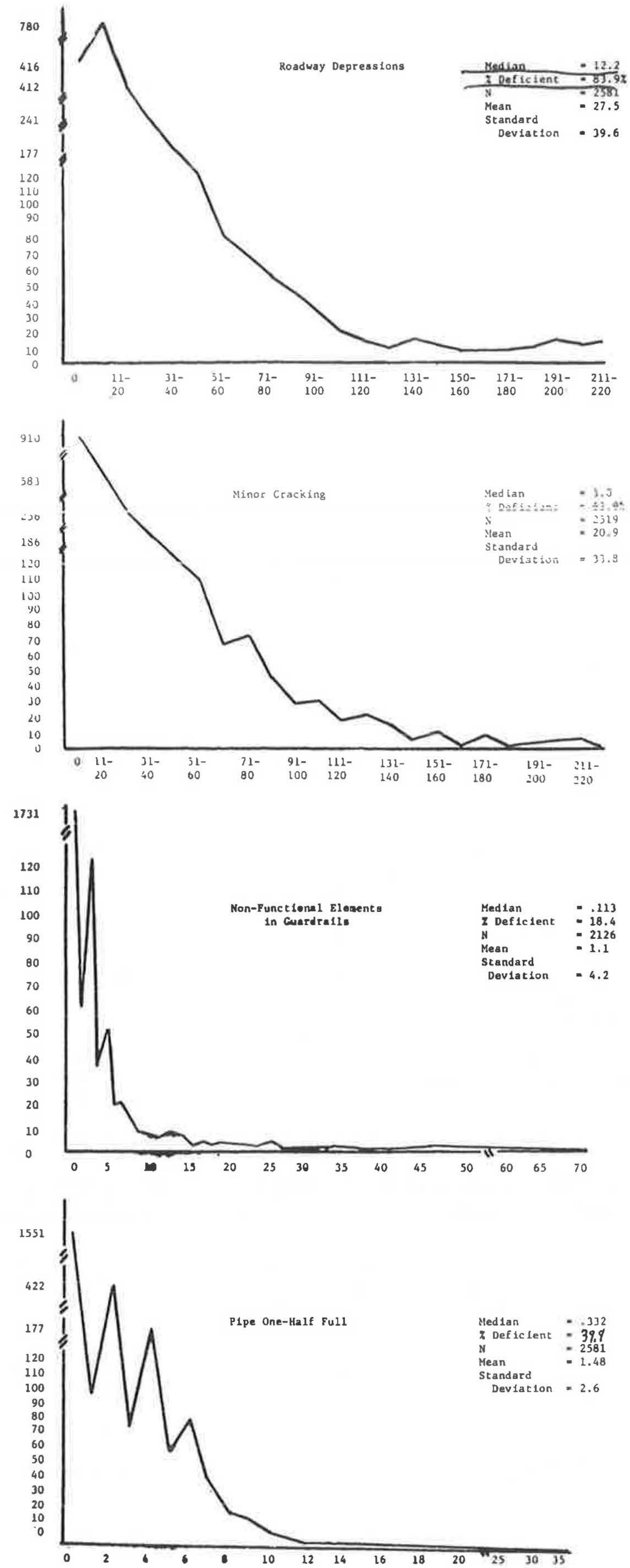


Figure 5. Median average counts and percentage of sections with reportable conditions--selected surface items.

MFC	Surface Depressions	Minor Cracking	Joint Deterioration	Major Cracking	All Potholes
Interstate	5.1 83.5%	0.5 49.8%	44.0 92.0%	1.2 57.8%	5.7 85.2%
Principal Arterial	10.6 81.8%	2.5 59.7%	33.1 73.2%	2.8 59.4%	11.7 82.7%
Minor Arterial	8.3 78.4%	7.7 71.4%	0.5 49.6%	10.9 69.2%	9.4 81.4%
Collectors	16.7 83.6%	9.2 66.2%	0.1 16.9%	8.5 64.0%	19.0 85.9%
Local	27.4 91.7%	13.0 70.1%	0.0 5.6%	12.1 64.7%	35.6 93.2%

Figure 6. Median average counts and percentage of sections with reportable conditions--selected shoulder items.

MFC	Shoulder Slope	Shoulder Buildup	Washouts	Major Cracking	Edge Pavement 4"
Interstate	0.1 8.3%	0.5 48.5%	0.2 26.2%	6.3 60.5%	0.0 8.9%
Principle Arterial	0.3 34.2%	0.3 38.8%	0.3 33.5%	0.4 41.3%	0.1 16.6%
Minor Arterial	0.5 48.9%	0.3 34.7%	0.2 31.5%	0.4 45.1%	0.1 17.3%
Collectors	14.9 65.1%	0.2 26.6%	0.2 24.3%	0.3 37.3%	0.1 20.2%
Local	22.2 69.7%	0.2 27.3%	0.1 20.4%	0.5 59.3%	0.1 16.6%

Figure 7. Median average counts and percentage of sections with reportable conditions--selected drainage items.

MFC	Nonfunctional Ditches	Nonfunctional Inlet	Pipe 1/2 Full
Interstate	0.1 22.4%	0.1 21.2%	0.1 20.9%
Principal Arterial	0.1 15.3%	0.2 23.7%	0.3 40.9%
Minor Arterial	0.1 19.5%	0.1 12.0%	0.3 37.8%
Collectors	0.2 25.1%	0.1 8.6%	0.4 45.3%
Local	0.2 25.9%	0.0 5.8%	0.4 45.4%

Principal arterials are at the same level as interstates on this indicator, reflecting the fact that they have as many inlets but perhaps receive a little less maintenance effort. The prevalence of roads with cross-pipes which are one-half full or more is lowest for Interstates, higher for arterials, and still higher for collectors and local roads, even though they have fewer pipes. This would indicate less maintenance attention to pipes on these roads. Because the incidence of drainage problems is so heavily dependent upon the actual numbers of ditches and structures, in Cycle 2 these totals will also be tabulated and the deficiencies will be analyzed on a percentage basis.

Figure 8 shows median averages the percent deficient on selected appurtenance conditions by MFC.

Striping is good on the interstates and evidences only a minor problem on lower order roads, except local roads. While many local roads do not have striping, the level of maintenance is less on those that do as compared with higher MFC's as would be expected. The reportable conditions concerning both nonfunctioning guardrail elements and missing or damaged regular signs are more frequent on the higher order roads, largely because such roads have more of these appurtenances in place. Missing or damaged delineators are the greatest problem on local roads and, surprisingly, on the Interstates; the lowest count is on principal arterials. The percentage of roads with some litter count is almost uniformly high except for local roads, but the median count decreases dramatically with the lower

Figure 8. Median average counts and percentage of sections with reportable conditions--selected appurtenances.

MFC	Bad Striping	Nonfunctional Guardrail Elements	Regular Signs	Missing Delineators	Litter	(N)
Interstate	0.0	0.1	0.1	2.9	81.5	
	0.0%	21.4%	22.4%	76.9%	100.0%	339
Principal Arterials	0.0	0.2	0.1	0.4	64.7	
	4.4%	25.3%	19.6%	46.9%	97.8%	633
Minor Arterials	0.0	0.1	0.1	1.1	45.4	
	4.4%	21.8%	13.3%	59.9%	95.5%	399
Collectors	0.0	0.1	0.1	2.3	32.0	
	7.1%	16.5%	13.3%	63.9%	95.9%	728
Local	0.2	0.0	0.1	2.5	14.8	
	2.4%	7.7%	9.7%	67.6%	89.2%	482

MFC's mainly a function of their lower levels of usage. This reportable condition may need to be redefined in order to provide a more discriminating indicator.

Because some of the variation by MFC presented above seems to be explicable in part by pavement type, it will be interesting to make some comparisons on this basis as illustrated by Figure 9. Median values and the percentage of roads exhibiting some potholes are highest for unpaved roads and lowest for rigid pavement roads, as might be expected given their respective design standards and maintenance priorities. Yet the most stark impression conveyed here is that the overall incidence of potholes is so high. Cracking--major and minor cracking combined--vary in the same direction although in a lower order of magnitude. While cracking is not applicable to unpaved roads, it was found in 64% of the flexible base roads observed and in only 24% of the rigid pavement roads.

Figure 9. Median average and percent roads deficient by pavement type--potholes and cracking.

Pavement Type	All Potholes Md	All Potholes %	All Cracking Md	All Cracking %	N
Unpaved	33.2	95.2	-	-	63
Flexible Base	20.1	87.2	36.4	63.7	1,419
Rigid Base	11.6	83.3	35.4	60.7	493
Rigid Pavement	6.8	80.0	3.2	24.4	601
					2,576

District and County Comparisons

Figure 10 shows the percentage of road sections with deficient shoulder slopes broken down by Engineering district and MFC. In general, the pattern of shoulder slope problems increasing with lower order MFC's is replicated in the 11 districts with some exceptions. In districts 4 and 5 there is less incidence of this problem on minor arterials than on Interstates and major arterials. Overall, the percentage deficient ranges from 38% in District 5 to 61% in District 10. From a macro level management perspective, this level of information should be of primary importance in evaluating maintenance efforts, particularly over time. District engineers can then be called upon to explain poor performance and to deal with problems in problem counties internally.

Following a hierarchical approach to data display, district engineers can be given results in the format of Figure 11, showing the incidence of minor cracking by district and counties within districts. The districts range from 53% to 75% roads with some deficiency and from a median of 1.8 to 13.2 counts per mile. While there is some degree of homogeneity within districts on this indicator, there is clearly substantial variation within. District 3, for example, is in relatively good shape in terms of minor cracking with the exception of Northumberland county which pulls the overall median up.

Figures 12 and 13 show geographic distributions of the percent deficient roads with major cracking and nonfunctional guardrail elements respectively. Major cracking problems do not coincide with district boundaries to any substantial degree. District 10 including Jefferson and Armstrong counties and District 4 including Luzerne, Bradford and Wayne counties both contain the upper and lower extremes. Most of the other districts include some near zero deficient counties as well as counties with a high percentage of deficient roads.

By contrast, the variation in the percentage of roads with nonfunctional guardrail elements coincides with district boundaries to a much greater degree. Districts 8 and 9 in the middle of the southern tier counties along with District 3 to the north have almost uniformly low percentages of deficient roads. Districts with concentrations of counties in the highest ranges of percent roads with nonfunctional guardrail elements include District 6 around Philadelphia, District 4 in the northeast corner, and District 12 in the southwest. This kind of clustering facilitates delegating responsibilities to correct problems to the District level.

Needs Assessment

A second major purpose of the trained observer survey is to project maintenance needs on the basis of the estimated total amount of each deficiency in a district or county or particular MFC. This is done by applying the overall mean count per mile and the percentage of deficient roads to the total mileage of a particular MFC in a district or county. While the ratings discussed above are based only on roads that are applicable--for a certain condition--counts of nonfunctioning guardrail elements are recorded only for road segments that have guardrails, for example--for the purposes of estimating the total magnitude of a condition rates and percentages from the complete subsample (including "not applicables")

Figure 10. Percentage of roads with deficient shoulder slope by district and MFC.

District	Interstate	Principal Arterial	Minor Arterial	Collectors	Local Roads	Row Total
1	17%	48%	54%	72%	67%	53%
2	0%	10%	52%	57%	83%	40%
3	0%	10%	35%	59%	62%	43%
4	11%	57%	30%	60%	68%	45%
5	6%	33%	24%	57%	63%	38%
6	0%	50%	70%	52%	45%	49%
8	8%	23%	38%	62%	66%	45%
9	0%	31%	32%	68%	70%	52%
10	8%	27%	64%	77%	90%	61%
11	7%	45%	74%	88%	82%	59%
12	14%	37%	77%	73%	76%	59%
	9%	34%	49%	65%	70%	49%

Figure 11. Incidence of Minor Cracking by district and county.

	Md	%
District 1	2.1	56%
Crawford	4.4	68
Erie	12.2	77
Forest	56.0	78
Mercer	.19	27
Venango	.32	39
Warren	2.0	58
Lawrence	7.0	60
District 2	6.8	69%
Centre	26.0	84
Clearfield	4.3	64
Clinton	10.7	67
Cameron	13.5	100
McKean	6.8	73
Potter	9.5	72
Mifflin	4.0	65
Elk	.23	31
District 3	13.2	75%
Columbia	18.0	90
Lycoming	3.7	64
Montour	.25	33
Northumberland	44.0	94
Snyder	1.0	50
Sullivan	24.0	90
Tioga	17.0	70
Union	14.0	94
District 4	4.9	61%
Bradford	16.5	79
Lackawanna	.31	38
Luzerne	.41	45
Pike	.37	42
Susquehanna	27.0	86
Wayne	21.0	85
Wyoming	.50	50
District 5	9.7	72%
Berks	1.50	53
Carbon	15.5	79
Lehigh	28.0	86
Monroe	2.5	58
Northampton	44.0	94
Schuylkill	8.0	81

	Md	%
District 6	1.9	57%
Bucks	.45	47
Chester	3.0	63
Delaware	2.3	62
Montgomery	12.0	72
Philadelphia	.21	30
District 8	7.9	73%
Cumberland	1.67	55
Franklin	2.1	62
York	22.0	74
Dauphin	8.5	84
Juniata	25.0	63
Lancaster	23.0	79
Lebanon	13.2	86
Perry	10.5	81
District 9	2.0	54%
Bedford	12.0	79
Blair	12.2	83
Cambria	.25	33
Fulton	.15	13
Huntingdon	.36	42
Somerset	4.5	60
District 10	3.9	62%
Armstrong	26.5	82
Butler	.46	48
Clarion	8.5	81
Indiana	4.0	65
Jefferson	.25	33
District 11	12.0	70%
Allegheny	16.0	77
Beaver	2.0	54
District 12	1.8	53%
Fayette	13.0	65
Greene	10.0	67
Washington	.21	29
Westmoreland	3.7	64

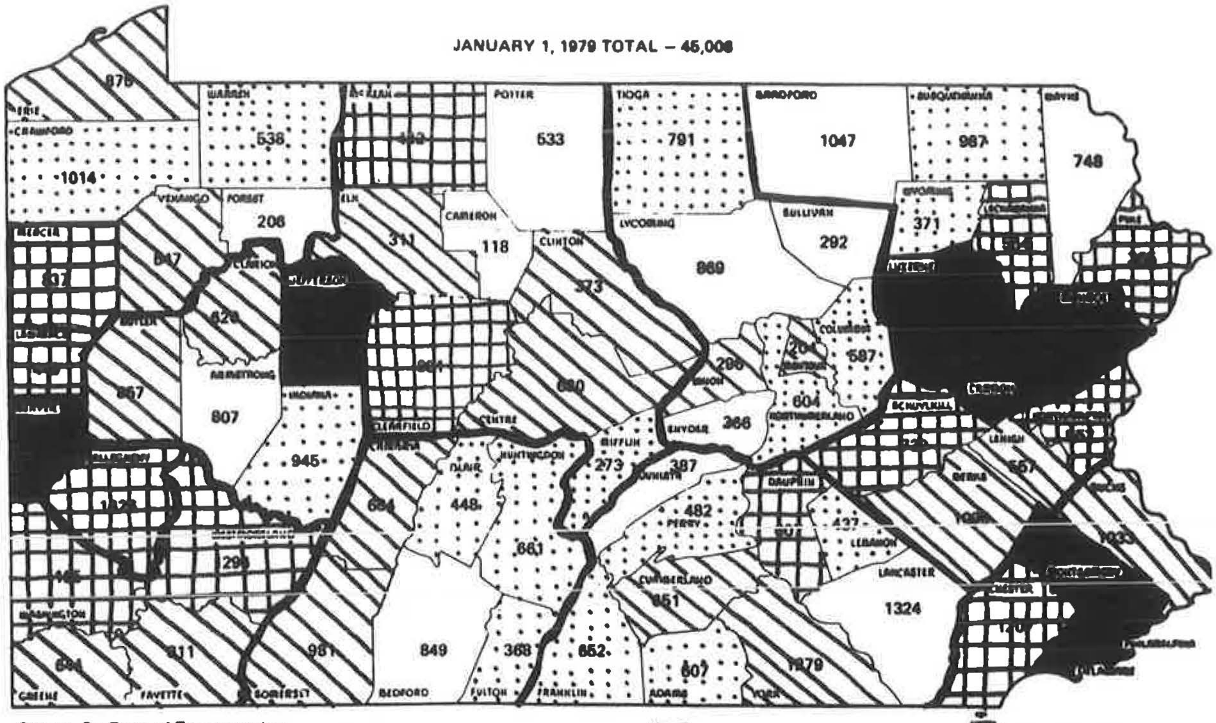
are appropriate. Thus the issue for a given district is, how many counts of nonfunctioning guardrail elements on principal arterials are there, taking into account the fact that many road segments, have no guardrail?

Figure 14 illustrates these findings with respect to all potholes (depressions + potholes + broken up areas) by district and MFC. Looking at the totals it shows that the overall sample is being expanded to 44,209 miles by applying the original sampling fractions. Of this total mileage it is estimated that 38,324 miles have at least one pothole per mile. In addition, the number of 25 foot sections with one or more potholes is projected to be 1,808,640; this would be the total estimated number of pothole clusters in the state system. Stated another way, if these 25 foot sections containing potholes were placed end to end, they would amount to 8,563 miles of highway, roughly 20% of the entire state system.

This kind of output should be most useful for programming because instead of showing the average condition on roads with certain design characteristics, it indicates the magnitude of deficiency across all highways in a given district. Depending on the size of the network of highways in different districts, their relative magnitudes of needs may or may not conform to their ranking in terms of average condition. For instance, District 4 has a lower rate of potholes per mile than does District 12, but the total projected pothole count is nevertheless higher in District 12. Districts 11 and 12 have the same rate of potholes per mile while District 12 has twice the total magnitude of deficiency. Similarly, Districts 5 and 8 have roughly equivalent ratings on potholes per mile, but District 8 has a much greater total need.

The output in Figure 14 can also be examined to make comparisons across MFC's. Although District 3 has a somewhat larger highway system than District 2, for example, it has substantially lower projected total pothole counts on Interstates, principal arterials, minor arterials and collectors, reflecting in general a higher level of maintenance on arterials and collectors. Yet, District 3 has a much higher projected pothole count on local roads than does District 2, due to the fact that there are many more local roads in District 3 than District 2. When this kind of analysis is conducted for all reportable

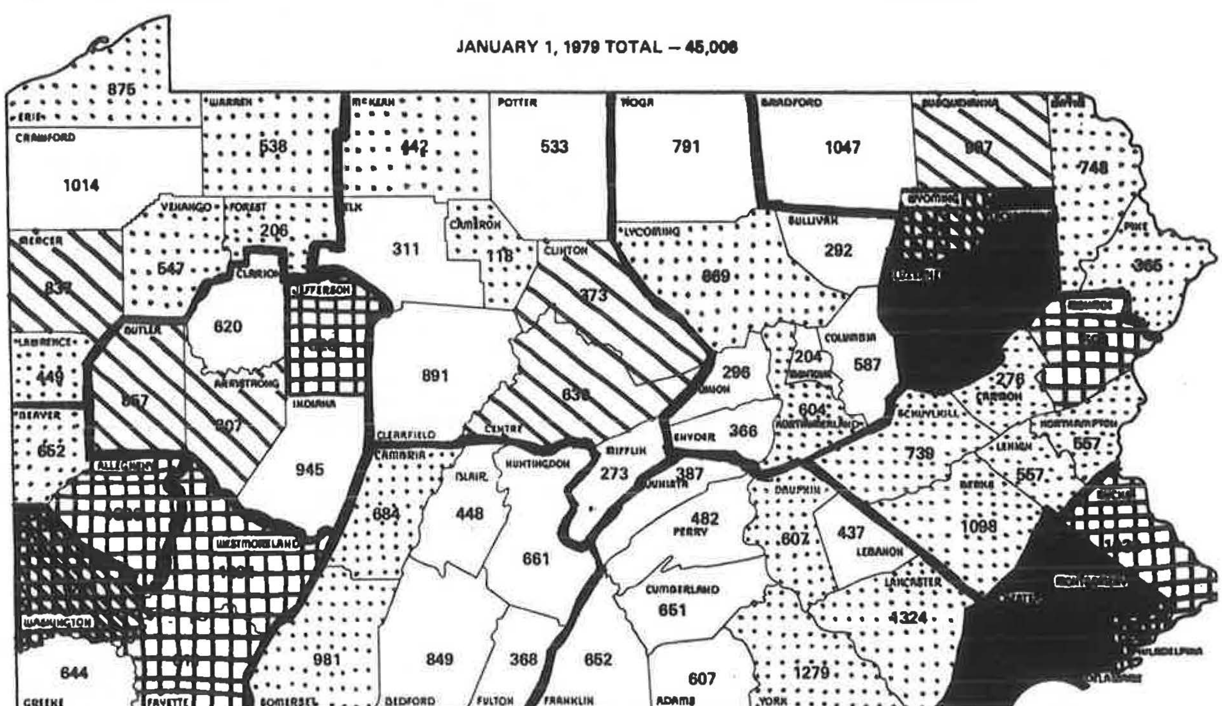
Figure 12. Percentage of roadway segments with major cracking on roadway surface.



Source: Pa. Dept. of Transportation
Bureau of Transportation Planning Statistics

81 - 100.00
61 - 80.99
41 - 60.99
21 - 40.99
0 - 20.99

Figure 13. Percentage of roadway segments with nonfunctional elements in guardrails.



Source: Pa. Dept. of Transportation
Bureau of Transportation Planning Statistics

LEGEND

51 - 100.00%
41 - 50.99%
31 - 40.99%
21 - 30.99%
11 - 20.99%
0 - 10.99%

Figure 14. Projected total pothole counts by district and MFC.

District		Interstate	Principal Arterial	Minor Arterial	Collectors	Local Roads	Total
1	Total Mileage	166	191	1,027	1,980	1,057	4,421
	Deficient Mileage	149	154.3	859.2	1,809	1,031	4,003
	Total Count	1,433	4,613	33,828	107,433	55,351	202,659
2	Total Mileate	99	279	870	1,545	737	3,530
	Deficient Mileage	92.6	163.3	692.1	1,283	737	2,968
	Total Count	984	3,845	19,100	55,542	38,029	117,500
3	Total Mileage	52	223	629	1,695	1,369	3,968
	Deficient Mileage	48.9	128.4	426.8	1,464	1,269	3,337
	Total Count	471	2,710	11,906	47,203	70,167	132,458
4	Total Mileage	178	161	860	1,877	1,897	4,973
	Deficient Mileage	155.7	140	650	1,642	1,872	4,460
	Total Count	2,352	5,943	16,613	68,276	134,238	227,421
5	Total Mileage	163	415	648	1,607	872	3,705
	Deficient Mileage	138.7	381	543.5	1,428	797	3,289
	Total Count	1,581	9,491	14,340	38,900	29,249	93,561
6	Total Mileage	63	1,053	736	1,327	629	3,803
	Deficient Mileage	51.8	962	715.5	1,106	524	3,359
	Total Count	1,201	29,641	17,337	40,891	36,010	125,080
8	Total Mileage	161	593	998	2,748	1,783	6,283
	Deficient Mileage	120	452	825.3	2,146	1,612	5,155
	Total Count	1,744	9,003	15,181	73,071	70,563	169,561
9	Total Mileage	24	387	556	1,649	1,342	3,958
	Deficient Mileage	20.6	317.9	402.6	1,374	1,165	3,280
	Total Count	391	8,818	21,646	48,146	73,202	152,203
10	Total Mileage	81	197	699	1,649	1,177	3,803
	Deficient Mileage	81	163	555.1	1,446	1,126	3,371
	Total Count	864	3,397	17,167	136,842	115,704	273,973
11	Total Mileage	51	453	463	664	268	1,899
	Deficient Mileage	37.4	411	442.9	638	268	1,797
	Total Count	346	16,145	17,675	42,726	26,776	103,667
12	Total Mileage	101	349	673	1,790	948	3,861
	Deficient Mileage	69.7	280.6	525.8	1,563	848	3,287
	Total Count	1,073	6,583	21,578	79,264	101,186	209,684
Total State							
	Total Mileage	1,139	4,301	8,159	18,531	12,079	44,209
	Total Deficient	968	3,553	6,646	15,909	11,252	38,324
	Total Count	12,411	100,051	206,429	740,527	749,224	1,808,640

conditions, such comparisons should facilitate a more objective basis for programming maintenance activities. Staff should be cognizant of differences between districts on total projects count of a given deficiency broken down by MFC, as well as differences in the projected magnitudes of various kinds of reportable conditions across districts.

Sensitivity Analysis

One concern with this trained observer survey is making comparisons between districts and counties which are statistically significant, concluding from the comparison of three percent samples that the incidence of some reportable condition in one area's entire road system is higher or lower than that of another area. Figure 15 shows the results of two-tailed difference of means tests (2 sample t tests) across the range of deficiencies for selected pairs of counties. These findings indicate a few important points regarding sample reliability for drawing conclusions about differences between counties. First, among the trained observer deficiencies, the surface foundation and shoulder items tend more to show significant differences than do the drainage and appurtenance items. Secondly, the former set of items also tend to indicate more significant differences than do the PSI measures. Although the PSI measures have compact frequency distributions with low standard deviations, the trained observer variables tend to discriminate better and worse counties than does PSI.

A final point about these comparisons highlights a problem. Although many of the pairs of counties shown in Figure 15 are seen to have statistically significant mean average counts, they represent some of the extreme pairings of counties. Many other pairs of counties with sample differences that are of practical significance do not turn out to have statistically significant differences. Even looking at Figure 15, shoulder deterioration is not statistically significant between Erie and Adams counties although the mean count for Erie is more than double that for Adams, a difference of more than 50 points. This is due to the extreme variation and resulting large standard deviations in most of these measures.

This indicates that larger sample sizes are required for future cycles of the survey. Using one-tailed tests with the present sampling fraction, differences in foundation failure means with a difference of 28 points are significant at the .05 level.

Figure 15. Two tailed difference in means test results for selected pairs of counties.

Variable	Susquehanna		Lycoming	Lycoming		Erie	Erie		Adams	Adams		York
Surface Deterioration	156.0	*	32.3	32.3	*	80.7	80.7	*	50.1	50.1	*	112.4
Surface Obstruction	3.7	*	0.4	0.4		3.9	3.9		0.6	0.6	*	1.6
Foundation Failure	72.0	*	17.1	17.1	*	81.8	81.8	*	7.6	7.6	*	40.7
Shoulder Deterioration	132.8	*	28.2	28.2		99.9	99.9		46.4	46.4	*	111.8
Shoulder Obstruction	9.2		2.5	2.5	*	8.1	8.1	*	2.4	2.4	*	12.7
Shoulder Failure	10.6		5.8	5.8	*	30.6	30.6		14.5	14.5		15.5
Shoulder Dropoff	16.5	*	3.2	3.2		7.0	7.0	*	0.9	0.9	*	12.8
Drainage Obstruction	3.8		2.1	2.1		3.6	3.6		1.3	1.3	*	3.9
Drainage Failure	0.9		0.2	0.2		0.3	0.3		0.1	0.1		0.2
Bad Striping	0.0		0.0	0.0		0.3	0.3		0.0	0.0	*	0.8
Guardrail Deterioration	1.4	*	0.5	0.5		1.0	1.0	*	0.0	0.0	*	2.7
Median Barrier Deterioration	0.0		0.0	0.0		2.0	2.0		0.0	0.0		0.0
Sign Deterioration	6.4		4.0	3.0		4.6	4.6	*	1.1	1.1	*	3.0
Litter	55.8		37.6	37.6		53.6	53.6		39.9	39.9		43.6
PSI LOW	2.5		2.6	2.6		2.9	2.9		2.7	2.7	*	2.3
PSI AVE	2.8		2.9	2.9		3.2	3.2		2.9	2.9		2.6
N =	47		41	41		57	57		32	32		69

* t test results statistically significant at the .05 level.

For districts the required difference is cut down to roughly 12 counts. Confidence interval estimates for the mean counts of foundation failure at the 95% level would run from 2.6 to 12.6, from 59.8 to 112.3 and from 30.3 to 50.6 for Adams, Erie and York counties respectively. Sample reliability seems a little more useful when looking at differences in the percentage of deficient roads. Initial tests conducted on surface depressions, minor cracking, major cracking and all potholes, for instance, indicates that differences of 60 to 10 percentage points between districts are statistically significant.

Further Analysis

The kind of descriptive analysis discussed in this paper, refined and conducted in a thorough fashion, will be beneficial to the Department in terms of both programming and monitoring performance over time. Beyond this, however, there are a number of subsequent types of analysis which can further improve the usefulness of the survey. These include:

1. The development of a weighting scheme for computing indexes of roadways, shoulders, drainage and appurtenances for ranking counties in the aggregate and inclusion in an allocation formula for maintenance funds.
2. Incorporation of ADT to build in a use factor in analyzing variation in condition and projecting total needs.
3. Conducting sampling experiments with data collected on 100% samples of major and minor arterials in Berks County to obtain a feel for increasing sample reliability.
4. Correlating the trained observer roadway items with PSI data to determine the degree of complementarity or redundancy.
5. Analyzing the road condition data in conjunction with maintenance expenditure data to assess the responsiveness of the program and, over time, sensitivity of road condition to maintenance efforts.

References

1. T. H. Poister, T. D. Larson, and S. Rao. "Fiscal Planning and Highway Programming: The Pennsylvania Response to a Changing Environment," Transportation Research Record, No. 654, Washington, DC: Transportation Research Board, 1977, pp. 16-22.

2. Secretary Thomas D. Larson's Address Before the 68th General Assembly of Boroughs, Hershey Convention Center, Tuesday, June 26, 1979.

3. E. L. Miller. "A Method of Measuring the Quality of Highway Maintenance". Transportation Research Record, No. 506, Washington, D.C.: Transportation Research Board, 1976, pp. 1-14.

4. Tallamy, Byrd, Tallamy and MacDonald. "A Study of Maintenance Quality Levels in Ohio." Prepared for the Ohio Department of Highways. Falls Church, Virginia. December 31, 1970.