Analysis of Two Acceptance Procedures for Aluminum Welds

RICARDO T. BARROS

ABSTRACT

Radiographic procedures are used to inspect and accept welds incorporated into aluminum overhead sign support structures. Radiography is expensive, however, so it is obviously desirable to inspect no more frequently than necessary to assure that defective welds are absent from accepted structures. The principles underlying the New Jersey Department of Transportation's present radiographic acceptance procedure are discussed and an alternative procedure is proposed that may save more than $10,000 annually without diminishing the present level of protection.

Aluminum welds are used by the New Jersey Department of Transportation (NJDOT) in the fabrication of aluminum overhead sign support structures. Every weld is first subjected to a visual inspection and then, in addition, a random sample from each lot of welds is subjected to radiographic inspection. Radiographic inspection is costly, currently $43 per radiograph, so there is an obvious incentive to reduce the number of radiographs taken to the minimum required. Those that are taken must still provide adequate protection against the acceptance of defective welds.

The existing radiographic acceptance procedure appears to provide the required protection. Analysis of the risks involved indicates that, in most cases, lots that contain an excessive number of flawed welds stand a small chance of passing undetected. This inference is based on established theory using the hypergeometric probability distribution.

However, a discrepancy between statistical theory and practical application of the acceptance procedure introduces a potential flaw into conventional statistical analyses, the impact of which was not previously known. As a practical expedient, a cluster sampling technique is used rather than pure random sampling. Risk analyses that assume one sampling procedure may be invalidated if another is used. Consequently, it was necessary to quantify the nature and magnitude of the potential bias introduced. If this step had not been taken, it could have been possible that the inferred protection was nonexistent but assumed to be present simply because the quality levels submitted to date have been exceptionally high.

Prerequisite information for any analysis of risk is the concept of acceptable and rejectable weld quality. Knowledge of specific quality levels is necessary to provide reference points at which the risk of not detecting flawed welds can be meaningfully compared. It was found that these specific definitions of quality, as such, had not been explicitly developed within the NJDOT or the American Welding Society (AWS) for aluminum welds. Instead, previous attention focused on engineering acceptance limits. Because of this lack of explicit quality definitions, it was first necessary to identify reasonable quality levels that it would be in the agency's interest to consistently accept and other quality levels that should be consistently rejected. This was done by evaluating historical NJDOT data in the context of operating characteristic (OC) curve analyses.

Given quality levels thus established, it was the objective of this study to critically evaluate the current NJDOT acceptance procedure and to propose an alternative that would afford either or both of two important benefits: a reduced exposure to risk or a reduced radiographic inspection cost. The alternative acceptance procedure subsequently developed was highly successful in both regards. Considerable cost savings ($10,000 per year or more) may be realized with risks not only stabilized near but, in some cases, substantially lowered from their present levels.

Initiating a change in an accepted practice is difficult, however, particularly when appreciation of the benefits to be obtained is not reinforced by a dissatisfaction with the procedure already in place. In addition, the relative merits of competing concerns may not be clear and thus favor making no change to the status quo. Consequently, many of the topics relevant to the selection of the best weld inspection strategy, which will enable management to make an appropriate, well-informed decision, are discussed in this paper.

EXISTING RADIOGRAPHIC ACCEPTANCE PROCEDURE

The NJDOT has, for several years, used a minimum sampling rate of 25 percent of the total number of welds in a lot. (Each structure is comprised of several lots.) If more than 10 percent of the radiographed welds are found to be defective, all of the remaining welds are subsequently radiographed. In any event, all defective welds found are repaired and the lot is eventually accepted.

The NJDOT pays for the cost of all radiographs except for those that reveal a defective weld. These are paid for by the fabricator at the current rate of $43 per radiograph. The total cost to the NJDOT of administering the existing radiographic inspection program ranges from roughly $40,000 to more than $200,000 per year, depending on the intensity of construction activity.

Attempts to determine the origin of this plan have been unsuccessful. Apparently it is not explicitly patterned after any existing standard but

---

New Jersey Department of Transportation, 1035 Parkway Avenue, CN 600, Trenton, N.J. 08625.

simply "evolved" many years ago. NJDOT engineers are under the impression that 100 percent radiographic inspection was originally used but later reduced because of (a) the generally satisfactory quality that was being received, (b) the relatively high cost of radiography, and (c) a belief that there is a sufficient amount of structural redundancy to preclude the sudden collapse of a sign support structure.

It is interesting that, in an isolated case, an existing and apparently serviceable aluminum sign support structure was dismantled from its field location and transported to the laboratory where it was subjected to 100 percent radiographic inspection. A subsequent analysis revealed that the tested structure would most likely have failed to pass the initial acceptance criteria. That is, it would have triggered the 100 percent inspection provision. Implications of this finding are discussed later in this paper.

QUALITY LEVELS AND ACCEPTANCE LIMITS

Meaningful risk analyses refer to quality levels, not acceptance limits. The distinction between these two terms is subtle but important. An acceptance limit represents the critical engineering tolerance that precipitates one of two actions--the acceptance or the rejection of a material--and hence expresses the policy that is to be followed in dealing with materials that may be submitted with differing levels of quality. Quality levels, on the other hand, are a more fundamental measure and reflect the degree to which a material could be expected to meet specified serviceability requirements if it were to be accepted.

In the present case, acceptable weld quality comprises those quality levels for which weld defects incorporated into a structure do not prevent the structure from providing adequate service during the structure's intended life. Certainly flawless welds meet this criteria, but so do welds that contain flaws not sufficiently large or frequent to diminish the serviceability rendered below some designated threshold. Historical observations strongly indicate that acceptable weld quality levels in a structure can, without a doubt, include some welds that are flawed.

A basic, well-established parameter used to represent the various levels of quality in analyses of this type is percent defective. This parameter simply quantifies the proportion of the welds in a structure that are flawed. (Weld flaws are defined by specific engineering tolerances for porosity, cracking, incomplete penetration, etc.) Structures with some low value of percent defective are judged to be of acceptable quality whereas others at higher levels of percent defective are not.

The percent defective parameter is intrinsically keyed to the acceptance limit defining weld flaws by the statistical acceptance procedure. Critical information relevant to the evaluation of an acceptance limit is the frequency with which welds defined as flawed by this limit are accepted. The net effect of a seemingly stringent limit in a (poorly designed) acceptance procedure may be that the average quality actually accepted is substantially worse than the stringent limit might suggest.

Acceptance limits must be considered to be primarily the expression of a policy decision, and the adequacy of this policy decision can be evaluated only relative to the assessment of the quality levels between which it is capable of distinguishing. This study seeks to identify specific quality levels through analyses of historical data and to assess the relative discriminating power of several alternative acceptance procedures in the recognition of these quality levels. Existing engineering tolerances that characterize welds as defective or not will remain unchanged. In so doing a basis will be established on which to comparatively evaluate the merits of the various acceptance procedures.

RISKS AND ENGINEERING DESIGN

The analyses performed in this study are of an admittedly empirical nature. An abstract parameter, percent defective, is used to gauge quality in structures observed to have provided specific levels of service. Percent defective considerations do not explicitly enter into the original design considerations, however, and a question may be raised about the relevance of this specific abstraction and whether another, more tangible procedure might not more meaningfully evaluate risks in a model with physically measurable dimensions. Indeed, an alternative analytical procedure does exist. Reliability analyses quantify the risk of specific structural elements failing in prescribed modes. These analyses require comprehensive knowledge of material properties and applied loads, however, which effectively places them beyond the scope of the present investigation. A brief digression may help justify the relevance of the empirical analyses performed and demonstrate why, with historical information available, risks may be addressed in a generalized manner.

The erection of a completed structure may be viewed as the end product of many, distinct engineering analyses. Three of these are the selection of an overall structural configuration as well as individual member dimensions, the specification of engineering limits on desirable material properties, and an analysis of the risks present in (materials) quality assurance. These analyses are obviously interrelated. Stronger material properties permit a more sparse structural configuration. Low confidence in the materials acceptance procedure would require a compensating degree of redundancy in structural support or surplus in material strength.

Given a design load, a structural element, and a specified material limit, the problem becomes one of assuring that the material limit is not exceeded. If this can be accomplished with a reliability comparable to that which has historically been achieved, then it can be inferred that the existing (and satisfactory) balance has been preserved. Thus the presence or absence of defects in a weld becomes the pertinent criterion and the operating characteristic curve the primary analytical tool.

USE OF THE HYPERGEOMETRIC PROBABILITY FUNCTION IN AN ACCEPTANCE PROCEDURE

Operating characteristic curves for lots that contain a discrete number of defects are calculated with the hypergeometric probability function (1). Probabilities are determined as a function of the lot size (N), the sample (n), the total number of defects in the lot (D), and the observed number of defects in the sample (d). A lot is accepted if the sample contains c or fewer defects, where c is the maximum number of allowable defects in a sample. Typically, lots with more than c defects are then subjected to 100 percent inspection and all detected defects are repaired or replaced. Sampled observations are implicitly assumed to meet the requirements of independent, random selection.

Operating characteristic curves developed with the hypergeometric probability function are, in a sense, more limited than curves developed with continuous functions because only discrete integers may be used. It is not possible to have 1.5 defects, for example. It sometimes occurs that an incremental change in one of the foregoing variables (i.e., N, n, D, d, or c) results in a noticeable jump in the incremental probability.

The existing acceptance procedure is especially subject to this type of fluctuation. For example, because the probabilities of acceptance are more sensitive to absolute sample sizes than sampling rates, two lots from which a 25 percent sample is taken will necessarily incur different risks if one contains 60 welds and the other 80 welds. Also, if the acceptance requirement is that 10 percent or less of the sample be defective, then additional imprecision is introduced because c may be set to 1 or 2 defects but never 1.5. And, finally, the existing acceptance procedure randomly selects weld nodes of various sizes until the cumulative number of sampled welds exceeds the minimum number required. (The actual sample size used is the cumulative number rather than the minimum required.) Thus, by chance, two lots of equal size may be represented by unequal sample sizes and incur different risks.

RQL, TOLERABLE RISKS, AND AVERAGE OUTGOING QUALITY

Established tolerable risks at specified quality levels have not been universally established in the existing state of the art. The American Welding Society does not recognize statistical risks in the acceptance procedure and favors strictly controlled fabrication conditions. The American Society for Testing and Materials, the American Society for Nondestructive Testing, and the American Society for Quality Control were also contacted but were unable to provide further guidance regarding the tolerable risk of accepting marginal quality.

It is fortunate that the NJDOT has developed historical information regarding weld quality, and this information gains in authority when other references remain silent. An aluminum sign support structure, scheduled for dismantlement after about 17 years of satisfactory service, was subjected to 100 percent radiographic inspection to provide additional quality and performance data. Scattered porosity was, by far, the most common defect found. A smaller number of cracks and tungsten inclusion were also observed, and even fewer incidences of lack of fusion were detected.

Although a great deal was learned about this particular structure, the conspicuous lack of other information necessitates the making of certain assumptions if specific inferences are to be generalized. The structure must be thought of as representative. Or, more specifically, it must be assumed that the relative frequencies of defect types found in the tested sections are not unusual, that the weld defects found at the time of inspection were present at the time of fabrication, and that the field loading exposure was not atypically light. NJDOT engineers queried on this matter considered this structure to be generally representative of others in use and thought that these assumptions were reasonably met.

A conservative rejectable quality level (RQL) value can then be derived. The observed quality levels for the five sections were found to range from 17 to 56 percent defective and the weighted average value for all trusses was 33 percent defective. Thus, for the purposes of this analysis, it is assumed that the RQL is no smaller than 33 percent defective.

Setting the RQL at 33 percent defective implies that trusses with no more than this amount of defective welds would serve at least as well as the tested sections. It also implies that trusses with more than RQL defects are not acceptable. Note that this latter implication is fairly conservative because two of the five tested sections were actually in excess of 50 percent defective. Note also that the percent defective parameter applies strictly to the degree of compliance with a specified engineering limit, such as a maximum limit for linear porosity. A change in this engineering limit may necessitate a corresponding change in the definition of the RQL.

The tolerable risk at the RQL is dependent on several considerations. These include the likelihood of structural failure should an RQL situation occur, the mode of the structural failure, and the potential consequences. (Recall that all welds are subjected to a visual inspection and that the risks discussed hereafter apply exclusively to those defects detectable only through radiographic inspection.)

On the basis of these observations, the likelihood of a structural failure exactly at the RQL appears to be extremely small. Should a failure occur, PennDOT sources personally contacted report that individual struts tend to disengage first and are visually detectable from the roadway. In their experience, this allowed sufficient time for the structure to be dismantled in a timely fashion. (PennDOT failures were generally attributable to incomplete penetration and lack of fusion and were not catastrophic. Also, PennDOT's structures were accepted on the basis of fabricator certification rather than a statistical acceptance procedure.) Thus the primary consequences of historical weld failures have been engineering costs. Should a catastrophic failure someday occur, it could have a human cost as well. Therefore the risk of incorporating RQL or worse quality welds into an overhead sign support structure should be kept reasonably small.

Not a single weld defect-related structural failure has occurred in New Jersey during the roughly 20 years during which the department has installed aluminum sign support structures. The present statistical acceptance procedure has been in effect for approximately 15 of these years. Therefore, for the purposes of this paper, it is assumed that the tolerable risk at the RQL is the risk historically borne over this period. Operating characteristic curve analyses indicate this risk has ranged from virtually 0.0 to more than 0.25, depending on the lot size, and the approximate median value of 0.05 is taken to be the tolerable risk at the RQL.

The corresponding risk of rejecting acceptable quality level (AQL) lots is not a significant concern in the present application because rejected lots are simply submitted to 100 percent inspection. Thus it is not necessary to define an AQL or to quantify the risk of rejection at the AQL. The cost of unnecessary inspection is a concern, however, and this cost is very much a function of the risk of rejection. The higher the risk of rejection, the greater the overall cost of inspection.

Of the 2,833 welds radiographed in 1984, 7 percent were found to be defective. If it can be assumed that 7 percent defective reasonably represents the construction quality of recent years, in which not a single aluminum weld-related structural failure has occurred, then the optimum sampling strategy can be determined.

The reasonableness of 7 percent defective as a representative value is supported by the average outgoing quality limit (AOQL). As shown in Figure 1, the average outgoing quality (AOQ) is dependent on the incoming quality level. When rejected lots are subjected to 100 percent inspection and all defects are repaired, an AOQL is established. This is the

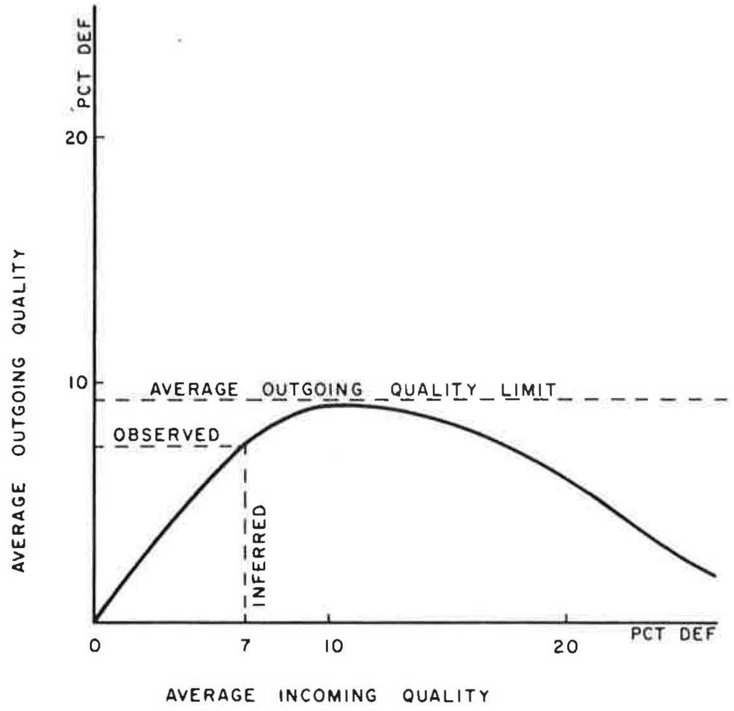

FIGURE 1 Average outgoing quality concepts.

maximum possible value for the AOQ. The AOQL for the present acceptance procedure is approximately 9 percent defective. This means that the worst the average outgoing quality could have been in the past is 9 percent defective, a value, reasonably close to the observed level.

To achieve the average outgoing quality of 7 percent defective observed in 1984, the average incoming quality level could have ranged anywhere from approximately 7 to 20 percent defective. Should the average incoming quality level have been toward the upper part of this range, however, the 100 percent inspection provision would have been triggered more frequently than actually observed. Thus it is reasonable to infer that the average incoming quality level of historical projects was truly in the vicinity of 7 percent defective, that this quality level adequately represents the quality of construction of existing structures in the field, and, based on empirical observation, that the existing quality levels in the field have been entirely satisfactory.

OPTIMUM SAMPLING STRATEGY

The optimum sampling strategy in the present application is determined by two criteria. First, for any lot size the probability of acceptance at the RQL should not be greater than 0.05. Second, the acceptance procedure should have the minimum average total inspection (ATI) at the 7 percent defective level. (The ATI is computed as the sum of two products: the probability of acceptance times the initial sample size plus the probability of rejection times the lot size. Identification of the optimum plan using these criteria is most conveniently accomplished through an iterative procedure with computer assistance.) If two plans have similar ATI values, then that plan with the smaller initial sample size is generally preferable because it must also have the lesser ATI value for smaller percent defective. Plans that meet these criteria will effectively provide protection comparable to what has historically been achieved at the minimum cost.

Two examples will illustrate the difference between the optimum sampling strategy and the existing acceptance procedure. Table 1 shows that, for relatively large lots, both the existing and the optimum sampling plans incur suitably small risks of accepting RQL lots. Further, both plans virtually never miss lots that are 40 percent defective or worse. For smaller percent defective values, however, inspection of the ATI columns reveals that the optimum plan requires fewer welds to be radiographed. Thus, although both plans afford comparable protection, the optimum plan is less expensive to operate.

TABLE 1 Comparison of Selected Characteristics of Two Acceptance Procedures for a Large Lot Size

	Plan 1, Existing Lot Size, N = 100 Sample Size, n = 25 Acceptance No., c = 3		Plan 2, Proposed Lot Size, N = 100 Sample Size, n = 17 Acceptance No., c = 2	
Percent Defective	P(accept)	Avg Total Inspection	P(accept)	Avg Total Inspection
0	1.0	25.0	1.0	17.0
2	1.0	25.0	1.0	17.0
7	0.94	29.8	0.91	24.7
33 (RQL)	0.01	100.0	0.03	97.2
40	0.00	100.0	0.00	100.0

Table 2 gives analogous information for a case in which the lot size is relatively small. The existing plan is grossly insensitive to the recognition of RQL lots, but the optimum plan maintains virtually the same risk as before. Of course, to achieve this protection the average total inspection of the optimum plan must be higher, and this is most noticeable when percent defective values are moderately large. Thus, in this case, it is the optimum sampling plan

TABLE 2 Comparison of Selected Characteristics of Two Acceptance Procedures for a Small Lot Size

	Plan 1, Existing Lot Size, N = 20 Sample Size, n = 5 Acceptance No., c = 1		Plan 2, Proposed Lot Size, N = 20 Sample Size, n = 7 Acceptance No., c = 0	
Percent Defective	P(accept)	Avg Total Inspection	P(accept)	Avg Total Inspection
0	1.0	5.0	1.0	7.0
2	1.0	5.0	1.0	7.0
7	1.0	5.0	0.65	11.6
33 (RQL)	0.41	13.9	0.02	19.7
40	0.31	15.4	0.01	19.9
60	0.06	19.1	0.0	20.0

that is more expensive to operate. It is fortunate that the increased inspection is negligible for very small lot sizes and that such small lot sizes are not very common. In any case, the increased ATI is simply the price to be paid if protection at the RQL is to be assured.

A complete set of optimum acceptance procedures for every lot size from N = 6 to N = 150 is presented in another report (2). The lot sizes that were observed in 1984, along with their frequency of occurrence, are given in Table 3. Also given are the acceptance criteria for the optimum and existing procedures as well as selected operating characteristics. Every one of the proposed acceptance procedures allows for that reasonably large acceptance number, c, which still restricts the risk of not detecting an RQL lot at 0.05 or smaller. The risk of not detecting RQL lots with the existing procedure is, of course, variable.

The product of the lot frequency and the ATI for 7 percent defective provides a reasonable estimate of the number of welds radiographed for each lot size, and the sum of these products estimates the number of welds radiographed in 1 year. Comparison of these two bottom line figures in Table 3 reveals that, on the average, the optimum plans require 334 (11 percent) fewer welds to be inspected per year than the existing acceptance procedure.

It is possible that, because of the sampling technique in which clusters of welds are selected, the actual sample size may be greater than the minimum required. Table 4 gives the same information as Table 3, except here every sample size has been increased by two welds. Under these conditions, and when the average percent defective value is 7 percent, the optimum sampling plans require 5 percent fewer welds to be radiographed annually.

As a rule, the optimum sampling plans require more welds to be radiographed than does the existing acceptance procedure when lot sizes are small. Small lot sizes occur less frequently, as inferred from 1984 data, so inspection savings for the larger lot sizes play the dominant role in determining which set of plans is most economical. Note that the net savings is expected to be greater still if percent defective values less than 7 percent are typically submitted for acceptance. Indeed, up to 20 percent savings would be realized if quality levels were to consistently approach 0 percent defective. And, finally, the optimum acceptance plans achieve this economy with a stabilized risk. Thus the optimum sampling plans appear to be clearly preferable to the plans currently in use.

COST CONSIDERATIONS

Reduced radiography rates do not translate directly into proportionately reduced costs. Many elements within the radiographic program represent fixed expenses. Travel, equipment, and film badge costs, for example, would remain virtually constant while labor and film costs might fluctuate.

In 1984 921 radiographs were shot at a total cost of approximately $42,000. Knowledgeable NJDOT personnel have indicated that this was a relatively light year and that up to five times this number of radiographs have been shot annually in the past.

TABLE 3 Summary of Operating Characteristics for Observed Lots, Minimum Sample Size

		Existing Plan					Optimum Plan				
Frequency	Lot Size	Sample Size	c	ATI (7%)	Weighted ATI	β (33%)	Sample Size	c	ATI (7%)	Weighted ATI	β (33%)
1	8	2	1	2.0	2.0	0.89	6	0	7.5	7.5	0.00
4	12	3	1	3.0	12.0	0.76	6	0	9.0	36.0	0.03
3	14	4	1	4.0	12.0	0.55	6	0	9.4	28.2	0.03
2	16	4	1	4.0	8.0	0.63	7	0	10.9	21.8	0.03
6	20	5	1	5.0	30.0	0.41	6	0	10.2	61.2	0.02
4	24	6	1	7.0	28.0	0.32	7	0	15.6	62.4	0.03
1	32	8	1	9.4	9.4	0.14	11	1	13.3	13.3	0.03
1	34	9	1	10.6	10.6	0.12	12	1	14.6	14.6	0.03
1	36	9	1	13.0	13.0	0.11	11	1	16.4	16.4	0.04
1	40	10	1	14.5	14.5	0.08	12	1	17.9	17.9	0.03
3	48	12	2	12.5	37.5	0.14	12	1	17.4	52.2	0.03
2	62	16	2	18.3	36.6	0.04	16	2	18.3	36.6	0.04
3	64	16	2	18.2	54.6	0.04	16	2	18.2	54.6	0.04
1	66	17	2	22.0	22.0	0.03	16	2	20.4	20.4	0.03
11	68	17	2	21.9	240.9	0.03	17	2	21.9	240.9	0.03
2	72	18	2	23.2	46.4	0.02	16	2	19.9	39.8	0.04
10	80	20	2	29.7	297.0	0.01	17	2	23.7	237.0	0.03
9	84	21	3	23.0	207.0	0.03	16	2	21.4	192.6	0.04
7	88	22	3	24.1	168.7	0.02	17	2	22.9	160.3	0.03
14	96	24	3	28.5	399.0	0.01	16	2	23.0	322.0	0.04
11	100	25	3	29.8	327.8	0.01	17	2	24.7	271.7	0.03
12	104	26	3	31.0	372.0	0.01	17	2	24.3	291.6	0.04
9	112	28	3	36.9	332.1	0.0	21	3	24.6	221.4	0.03
1	116	29	3	38.2	38.2	0.0	21	3	24.3	24.3	0.04
4	128	32	4	36.1	144.4	0.0	21	3	25.2	100.8	0.04
1	140	35	4	42.4	42.4	0.0	21	3	26.2	26.2	0.04
Weighted total					2,906.1					2,571.7	
Weighted average						0.10					0.03

TABLE 4 Summary of Operating Characteristics for Observed Lots, Minimum Sample Size Plus Two

		Existing Plan					Optimum Plan				
Frequency	Lot Size	Sample Size	c	ATI (7%)	Weighted ATI	β (33%)	Sample Size	c	ATI (7%)	Weighted ATI	β (33%)
1	8	4	1	4.0	4.0	0.50	8	-	8.0	8.0	0.0
4	12	5	1	5.0	20.0	0.42	8	0	10.7	42.8	0.0
3	14	6	1	6.0	18.0	0.24	8	0	11.4	34.2	0.0
2	16	6	1	6.0	12.0	0.35	9	0	12.9	25.8	0.0
6	20	7	1	7.0	42.0	0.18	8	0	12.8	96.8	0.0
4	24	8	1	9.6	38.4	0.14	9	0	18.3	73.2	0.01
1	32	10	1	12.0	12.0	0.06	13	1	16.0	16.0	0.01
1	34	11	2	11.0	11.0	0.21	14	1	17.2	17.2	0.01
1	36	11	2	11.6	11.6	0.19	13	1	19.7	19.7	0.02
1	40	12	2	12.6	12.6	0.15	14	1	21.2	21.2	0.01
3	48	14	2	14.7	44.1	0.07	14	1	20.8	62.4	0.01
2	62	18	2	21.1	42.2	0.02	18	2	21.1	42.2	0.02
3	64	18	2	20.9	62.7	0.02	18	2	20.9	62.7	0.02
1	66	19	2	25.5	25.5	0.01	18	2	23.8	23.8	0.02
11	68	19	2	25.3	278.3	0.01	19	2	25.3	278.3	0.01
2	72	20	2	26.6	53.2	0.01	18	2	23.2	46.4	0.02
10	80	22	3	24.6	246.0	0.02	19	2	27.7	277.0	0.02
9	84	23	3	25.8	232.2	0.01	18	2	25.2	226.8	0.02
7	88	24	3	26.9	188.3	0.01	19	2	26.7	186.9	0.01
14	96	26	3	31.8	445.2	0.0	18	2	27.3	382.2	0.02
11	100	27	3	33.1	364.1	0.0	19	2	29.0	319.0	0.02
12	104	28	3	34.3	411.6	0.0	19	2	28.6	343.2	0.02
9	112	30	3	40.8	367.2	0.0	23	3	27.8	250.2	0.02
1	116	31	4	33.6	33.6	0.0	23	3	27.5	27.5	0.02
4	128	34	4	39.2	156.8	0.0	23	3	28.7	114.8	0.02
1	140	37	4	46.1	46.1	0.01	23	3	30.0	30.0	0.02
Weighted total					3,178.7					3,008.3	
Weighted average						0.05					0.02

Thus the total cost of the aluminum weld inspection program could very well exceed $200,000 per year. Excluding the share paid by fabricators for defective welds found, the flexible cost associated with the 1984 construction season was approximately $32,000. This cost, which could reach the $150,000 mark in a busier year, is most conveniently evaluated as the cost rate per 1,000 radiographs shot.

Table 5 gives the annual flexible cost as a function of the radiographs shot and the anticipated savings resulting from a decrease in the sampling rate. A reasonable number of annual radiographs to consider may be the median value in Table 5, approximately 3,000 per year. For this value, an annual savings of from $5,000 to $20,000 could be realized with the implementation of the alternative sampling plan previously identified. The lower limit of this range would result if it were commonly necessary to inspect more welds than the minimum required, and the upper limit would result if, as a result of the alternative plan's implementation, fabricator quality were to be spurred to improvement. Perhaps the most reasonable value to expect is an annual savings of approximately $9,000 to $10,000. It is thought that most of this savings would result simply from the reduced sample sizes, but a small contribution from increased quality of production is also intuitively expected.

TABLE 5 NJDOT Savings per 1,000 Radiographs

No. of Radiographs Shot per Year	Annual Flexible Cost ($)	Savings ($) Resulting from % Reduction in Sampling Rate		
		5	10	20
1,000	31,870	1,594	3,187	6,374
2,000	63,740	3,187	6,374	12,748
3,000	95,610	4,781	9,561	19,122
4,000	127,480	6,374	12,748	25,496
5,000	159,350	7,968	15,935	31,870

CLUSTER SAMPLING

A discrepancy between the statistical theory assumed appropriate and practical application of weld radiography introduces a flaw into the preceding analysis. Fortunately this discrepancy was found to have a small impact in the present application, but its effects and implications represent a potential concern that could not go unaddressed.

The discrepancy arises from the known violation of a fundamental, underlying assumption. Contrary to theory, the welds inspected are not selected in accordance with standard procedures for obtaining independent, random samples. They are selected in fixed clusters as they naturally occur. Thus, after the first weld is randomly selected from all possible welds, adjacent welds are automatically inspected and nonadjacent welds may escape inspection altogether. If the weld fabrication environment is such that the defects produced tend to be correlated with one another, then the specter of clusters that are entirely defective and that may fail to be detected is raised. Conventional risk analyses are insensitive to this and, fooled by the large number of welds inspected, may substantially understate the incurred risk.

A computer simulation model was developed to investigate the nature and degree of bias introduced when the fundamental assumption of independent, random sampling is violated. Lots of varying size were generated in which the total number of defective welds was a controlled variable and in which the degree of association between two consecutively generated welds could be specified. (Serial correlation was specified within a continuous variable and converted to attribute-type data in the simulated structure. This is believed to realistically represent the manner in which defective welds would tend to be correlated.)

Each of the modeled structures was then sampled in two ways that simulate alternative acceptance procedures: cluster sampling (the current NJDOT practice) and true random sampling. It was possible to tabulate whether the acceptance procedure disposed of each structure properly because the simulated quality levels were known. The long-run average frequency with which each procedure correctly rejected defective lots and accepted nondefective lots could then be compared.

The impacts of the sampling technique and the degree of serial correlation (r_s) on the acceptance procedure are given in Table 6. It can be shown that when r_s = 0.0, application of random or cluster sampling procedures results in equivalent operating characteristics. When the degree of correlation is large, both the producer's risk and the buyer's risk are increased. When the correlation is large and cluster sampling procedures are used, these risks are increased to a still greater extent.

TABLE 6 Impact of Cluster Sampling and Serial Correlation

Sampling Procedure	r_s = 0, Independent Observations	r_s = Large, Associated Observations
Random	Reference datum	Seller's risk (α) and buyer's risk (β) slightly increased
Cluster	Same as reference datum	Seller's risk (α) and buyer's risk (β) increased to a greater extent

Serial correlation and cluster sampling have a disproportionate and increasingly larger effect on the acceptance procedure as the absolute value of r_s approaches 1.0. This effect is negligible for small percent defective values and increases as the percent defective value grows. This phenomenon is shown in Figure 2. Horizontal lines would have been produced if the probabilities of acceptance had been independent of the serial correlation. It may be observed that serial correlation introduces greater bias (a steeper slope in Figure 2) when the percent defective values are moderately large. Fortunately, the probability of acceptance (without triggering the 100 percent inspection provision) is relatively small in this region. Extremely large values of serial correlation may also affect the probabilities of acceptance even at small percent defective levels, although such high correlation values are quite improbable.

Simulation analyses indicate that the impact of cluster sampling is practically negligible in this application for low levels of serial correlation among weld defects and relatively low levels of percent defective. The degree of serial correlation would have to be rather large (e.g., r_s = 0.5) for its effect to be pronounced. At the r_s = 0.2 level of serial correlation, a value higher than actually observed in the few lots checked, probabilities of acceptance were increased by approximately 0.03 (or less) in the vicinity of the RQL. Near the 5 percent defective level, the opposite effect was observed with cluster sampling reducing the probability of accepting satisfactory lots by an even smaller amount. AOQL and ATI values were not greatly affected.

Serial correlation itself cannot be controlled, of course, so it is the manner in which it is treated by the acceptance procedure that must be considered. The computer simulation tests strongly suggest that cluster sampling is not a serious problem in the present application.

CONCLUSIONS

The influence of cluster sampling in a procedure in which random sampling is assumed has been determined.

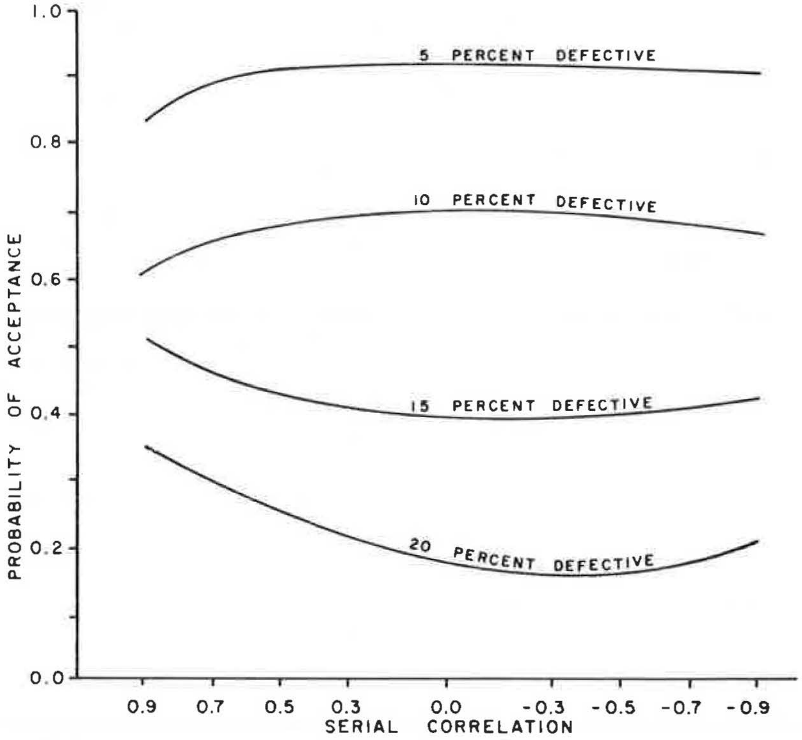

FIGURE 2 Effect of serial correlation on the probability of acceptance in clustered samples.

It is fortuitous that, in the current application, this influence was found to be negligible.

The existing aluminum weld radiographic acceptance procedure appears to provide adequate protection against the acceptance of defective welds, even if this degree of protection is not consistent. Should quality levels worse than 33 percent defective be submitted, these defects will usually be detected provided the lot size is reasonably large. The acceptance procedure becomes increasingly more lenient as the lot sizes are reduced, however, although the increased risks are somewhat offset by the relative scarcity of very small lots. The cost of administering this acceptance procedure is dependent on the level of construction activity in any given year. In general, this cost is expected to run between $40,000 and $200,000 annually.

An alternative acceptance strategy, which stabilizes the risk of failing to detect defective welds, has been identified. This risk is kept small regardless of the size of the lot. In comparison with the existing sampling strategy, small lots are inspected more thoroughly and large lots are inspected more efficiently. The net result is a reduction in the number of radiographs required to be shot. This reduction may range from 5 to 20 percent of the number presently required. Translated to dollars, one estimate of the associated savings is $10,000 per year. Reasonable lower and upper bounds on this savings might be $1,000 and $32,000, respectively, depending on the quality levels actually submitted, the level of construction activity, and the efficiency with which welds may be included on a radiograph.

Regardless of the acceptance strategy used, a risk always exists that defective welds may pass undetected. The proposed acceptance plans stabilize this risk near the existing minimum level, rendering these plans both more effective and more economical.

REFERENCES

1. E.L. Grant and R.S. Leavenworth. Statistical Quality Control. 5th ed. McGraw-Hill Book Company, New York, 1980, pp. 359-386.
2. R. Barros. Analysis of Two Aluminum Weld Acceptance Procedures. New Jersey Department of Transportation, Trenton, 1986.

Publication of this paper sponsored by Committee on Quality Assurance and Acceptance Procedures.