Paper No. 03-3680

Multivariate Statistical Model for Predicting Occurrence and Location of Broken Rails

C. Tyler Dick, Christopher P. L. Barkan, Edward R. Chapman, and Mark P. Stehly

Broken rails are the leading cause of major accidents on U.S. railroads and frequently cause delays. A multivariate statistical model was developed to improve the prediction of broken-rail incidences (i.e., service failures). Improving the prediction of conditions that cause broken rails can assist railroads in allocating inspection, detection, and preventive resources more efficiently, to enhance safety, reduce the risk of hazardous materials transportation, improve service quality, and maximize rail assets. The service failure prediction model (SFPM) uses a combination of engineering and traffic data commonly recorded by major railroads. A Burlington Northern Santa Fe Railway database was developed in which the locations of approximately 1,800 service failures over 2 years were recorded. The data on each location were supplemented with information on other engineering and traffic volume parameters. A complementary database with the same parameters was developed for a randomly selected set of locations at which service failures had not occurred. The combined databases were analyzed using multivariate statistical methods to identify the variables and their combinations most strongly correlated with service failures. SFPM accuracy in predicting service failures at specific locations exceeded 85%. Although further validation is necessary, SFPM is promising in the quantitative prediction of broken rails, thereby improving a railroad's ability to manage its assets and risks.

Derailments from broken rails have been a safety concern for more than a century (*1, 2*). Improvements in rail manufacturing, inspection, and rail defect detection have greatly reduced the incidence of broken rails. However, broken rails frequently cause service interruptions and are a leading cause of derailments. Improving the prediction of locations where broken rails are likely to occur has economic and safety benefits, enabling more effective allocation of resources to detect and prevent broken rails (*3–5*). Previous work has focused on the development of fracture-mechanic approaches in combination with empirical testing (*6–9*) and single-variable probabilistic methods using Weibull analysis (*10*). The first studies have assisted in understanding the underlying mechanisms of rail defect occurrence; the latter studies have assisted in predicting the useful life of rail, given basic information on traffic volume and loads. A statistical approach based on more variables can potentially improve the prediction of rail life and may also provide more insight into associated mechanisms (*3*). Over the past decade, railroads have expanded their use of information technology to include extensive geographic and engineering data systems. Large, multivariate databases that extensively detail key parameters likely to affect the occurrence of broken rails have been developed, thereby making such an approach feasible.

C. T. Dick and C. P. L. Barkan, Railroad Engineering Program, University of Illinois at Urbana–Champaign, 1201 Newmark Civil Engineering Laboratory, 205 North Mathews Avenue, Urbana, IL 61801. E. R. Chapman and M. P. Stehly, Hazardous Materials, Burlington Northern Santa Fe Railway, 2600 Lou Menk Drive, Fort Worth, TX 76131. Current affiliation for C. T. Dick: HDR Engineering, Inc., 512 Main Street, Suite 600, Fort Worth, TX 76102.

DEFINITION OF SEVERE DERAILMENTS

This study focused on identifying derailment causes most likely to lead to a severe accident in which many cars derail at speed. Generally such accidents will have the greatest potential for harm to people, property, equipment, and track. Further, analyses of FRA accident data have shown that accidents with these characteristics strongly correlate with the release of hazardous materials, if present, in the vicinity of the train derailment (*11, 12*). Consequently, for both safety and economic reasons, information on such types of derailments was of particular interest.

Derailment Severity–Frequency Analysis

To determine the causes of accidents most likely to lead to severe derailments, a simple risk analysis was conducted using FRA data for 3,504 mainline derailments that occurred during the 1994–1998 period (*13*). The FRA reporting system requires the identification of a primary cause (and other contributing causes if applicable). FRA groups data on accident causes hierarchically. Data at the FRA "subcause" level (the second-highest level of aggregation) were used in this study. The average number of cars derailed in accidents attributed to each subcause was calculated and plotted against the frequency of derailments with the same subcause (Figure 1).

Figure 1 is divided into four quadrants by vertical and horizontal lines that represent the average value of the two variables with respect to the x and y axes, respectively. The vertical line represents the average frequency of accidents for all recorded causes combined, and the horizontal line is the average number of cars derailed due to each cause. Causes above or below these lines are, by definition, above or below average for the respective axis.

The causes in the upper right quadrant are most interesting and pose the greatest risk—they are more frequent and more severe than average. It is clear that the most frequent cause of high-consequence accidents related to the FRA cause code for rail and joint bars. More detailed analyses revealed that most of these accidents were

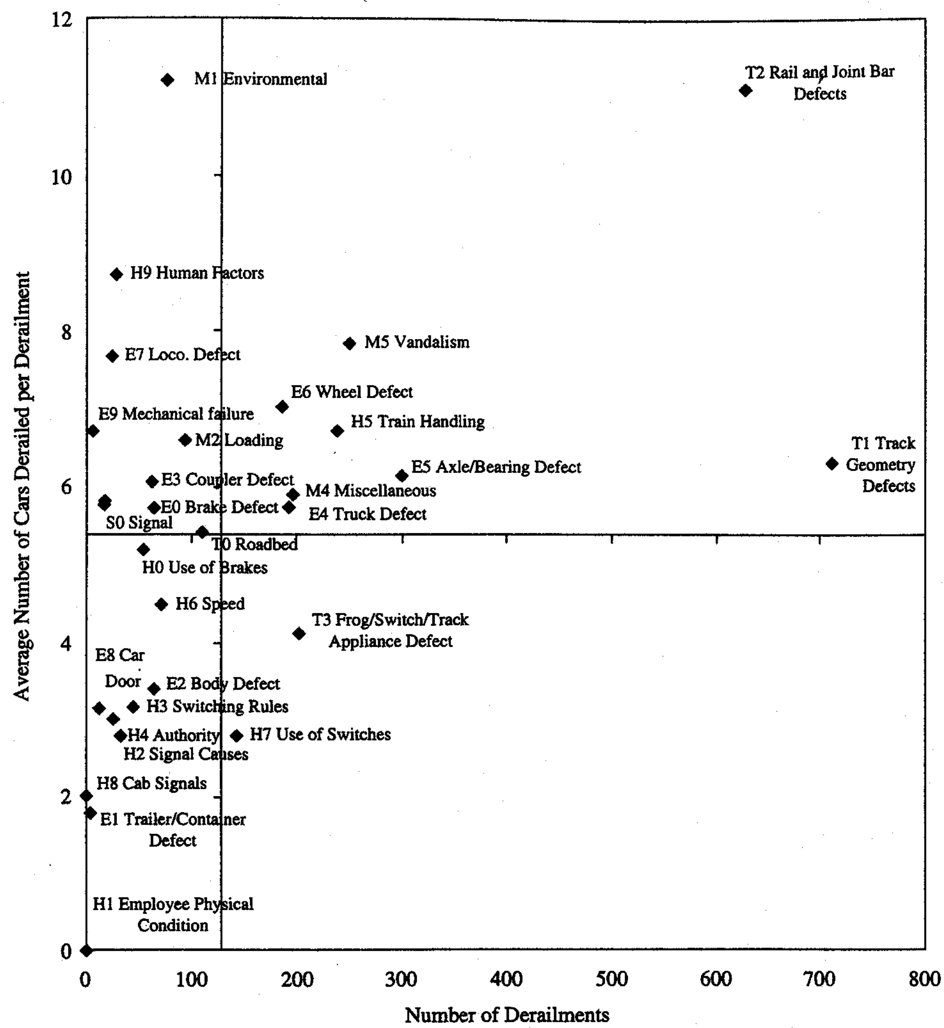

FIGURE 1 Frequency–severity graph of mainline derailments, 1994–1998.

the result of broken rails. On the basis of these results, a more detailed analysis was made of factors contributing to the occurrence of broken rails (*11*). Several recent hazardous materials accidents have underscored the importance of this particular aspect of risk from broken rails.

Broken Rails

Most broken rails do not result in derailments. Instead the break is detected, usually by the track circuit system or by track inspectors, and repaired (on several North American railroads, these detected broken rails are referred to as service failures). Broken rail derailments appear to correlate with the occurrence of service failures (*11*). Therefore, predicting the occurrence of service failures has a potential safety benefit—enabling railroads to allocate broken rail prevention measures, detection technology, and inspection efforts more effectively (*3–5, 7*). Further, understanding the factors correlated with service failure occurrence could help identify contributing causal factors, thereby enabling better preventive measures. The objective of this research was to develop a probabilistic model to predict the circumstances most likely to lead to the occurrence of a service failure.

Model Form and Data Set

Ideally, the model developed would enable the user to input values for the relevant parameters at a specific location on the railroad and determine a measure of the probability of a service failure there. The output of the model is an index value between 0 and 1, with 0 indicating the lowest probability of service failure and 1 representing the highest. Because a probability is the desired output and there are only two possible outcomes—service failure or no service failure at each location—the model can be constructed as a discrete choice model.

A discrete choice model, such as the logit model, fits an appropriate equation to the data and uses this equation to score each location relative to a threshold value, above which failure is predicted to occur (*14*). The logit model then uses a logistic distribution to consider the uncertainty and error in the estimated score and threshold value and determine the probability that the score is above the threshold value. The calculated probability is then used as an estimate of the service failure probability at that particular location.

To fit a discrete choice logit model, two sets of data were required—one to characterize locations where service failures occurred and another to characterize locations where service failures had not occurred. Data development began with Burlington Northern Santa Fe detailed information on the date, location, and type of 1,903 service failures that occurred over 2 years. The data were supplemented with engineering and operational data on each service failure location. A new dependent variable was created and assigned a value of 1 for each of these records signifying that a service failure had occurred.

The second set of data was created with records for locations where no service failure occurred during the same interval. This data set, of about the same size as the first, was developed by selecting a random sample of locations from the railroad and assembling the same information as for the service failure locations. The dependent variable for these records was assigned a value of 0.

Ultimately, a test database was developed that contained 3,676 records with complete service failure and descriptive parameter information. On the basis of a univariate analysis of the service failure data and a literature review on the circumstances of rail defect growth and broken rail occurrence (*8, 15, 16*), track structure and dynamic effects (*17–19*), and rail fracture mechanics (*6, 20*), the following parameters were selected for inclusion in the multivariate service failure model:

- Rail age,
- Rail weight,
- Degree of curve,
- Speed,
- Average tons per car,
- Average dynamic tons per car,
- Percent grade,
- Annual gross tonnage,
- Annual wheel passes,
- Insulated joints, and
- Mainline turnouts.

All of the parameters are continuous variables except the last two, insulated joints and mainline turnouts, which are both discrete. The parameters were assigned a value of 1 if present at a location, and 0 if not.

Model Development

The service failure probability model was developed using the statistical analysis system (SAS) and the LOGISTIC procedure. The LOGISTIC procedure fits a discrete choice logit model to the test database. Stepwise regression was used to determine the most relevant parameters and combinations of parameters (two-factor interaction terms) for inclusion in the model (*21*). The stepwise regression procedure uses an iterative process to select variables on the basis of their ability to explain the variance in the input data. The model conducts a goodness-of-fit test for each step and adds or subtracts variables or combinations of variables until the addition of another parameter does not significantly improve the fit. At this point, the last version of the model is considered the best and the resultant parameters, coefficients, and functional relationships comprise the final model.

Retrospective and Prospective Models

The service failure model was developed in two steps. First, the model was fit to the test database. Approximately half of the 3,676 locations in this database had a service failure during a 2-year period, and the other half were a random sample of locations that did not. Because the model makes predictions about broken rails that have already occurred, it was termed a "retrospective model." This version of the model is used primarily to assess the accuracy of the model's predictions relative to the test database.

The second step of the process developed a "prospective model." This model is modified from the retrospective model by adjusting a constant term to reflect the actual average service failure probability over a specific portion of the railroad system. After this adjustment, the prospective model can be used to calculate the annual probability of a service failure at particular locations, or along any portion of track.

Retrospective Service Failure Model

The retrospective service failure probability model was developed using the LOGISTIC procedure as follows:

$$P_{SF2} = \frac{e^U}{(1 + e^U)}$$

where

P_{SF2} = probability that a service failure occurred at a particular point during the study period;
$U = Z + Y$;
$Z = -4.569$, model-specific constant;
$Y = 0.059A + 0.025AC - 0.00008A^2C^2 + 5.101T/S + 217.9W/S - 3861.6W^2/S^2 + 0.897(2N - 1) - 1.108P/S$;
A = rail age (years);
C = degree of curvature (= 0 for tangent);
T = annual traffic [million gross tons (MGT)];
S = rail weight (pounds);
$W = 4T/L$ = annual number of wheel passes (millions);
$P = L(1 + V/100)$ = estimated average dynamic wheel load;
N = 1 if at turnout, 0 if not at turnout;
L = tons per car; and
V = track speed.

The fitted model includes a model-specific constant or intercept term, Z, that is related to the average service failure probability. The retrospective model is fit to a data set in which approximately half of the records are for locations with service failures. The average service failure probability on an actual system would be far lower, so this term would be adjusted accordingly.

Interpretation of Model Terms

The service failure probability model has terms that describe different effects and relationships among service failure probability, infrastructure characteristics, and traffic characteristics.

The first term in the model, $0.059A$, reflects the effect of rail age. As rail age increases, service failure probability increases. This result is consistent with extensive industry experience (*9, 10*). Older rail is likely to have carried more tonnage, experienced more thermal stress cycles, and may have been manufactured using processes that produced more flaws in the rail. A recent study of rail failures on Railtrack in Great Britain supports the importance of this parameter (*22*).

The second and third terms in the model, $0.025AC - 0.00008A^2C^2$, reflect the interaction between rail age and degree of curve. As either rail age or degree of curve increases, service failure probability is predicted to increase. Because the interaction between rail age and curvature is multiplicative, the model indicates that in terms of service failure probability, higher degree (sharper) curves are more sensitive to the effects of rail age and vice versa.

The fourth term in the model, $5.101T/S$, reflects the effect of annual traffic (MGT) normalized by rail weight. As annual gross tonnage increases, service failure probability increases. However, the increase in service failure probability associated with a unit increase in annual traffic is greater on segments of track with relatively light rail.

The fifth and sixth terms in the model, $217.9W/S - 3861.6W^2/S^2$, describe the effect of annual wheel passes or load cycles normalized by rail weight. Service failure probability increases as the number of wheel passes or load cycles increases. However, just as with gross tonnage, the increase in service failure probability associated with a unit increase in the annual number of wheel passes is greater on segments of track with relatively light rail. This situation probably reflects the greater stress of lighter rail under a given load than heavier rail. Thus, the amount of crack growth per fatigue cycle is greater in lighter rail than heavier rail.

The model includes terms that describe annual traffic relative to gross tonnage and number of wheel passes. The relationship between annual traffic and service failure probability is a function of both the total amount of load applied to a section of rail and the number of times the load is applied. This relationship is consistent with fracture mechanics models of fatigue crack growth in rails that depend on both the applied stress and the number of load cycles (*10, 20*).

The seventh term in the model, $0.897(2N - 1)$, describes the effect of mainline turnouts. The model indicates that proximity to a turnout increases the probability of a service failure. Several possible explanations relate to inferences about rail stress. Turnouts may tend to anchor the track structure, thereby causing greater thermal stress cycling as the nearby rail expands and contracts. Also, to the extent that turnouts tend to be associated with locations where trains are braking or accelerating, rails in these locations may tend to experience more traction-induced stresses.

The final term in the model, $-1.108P/S$, describes the effect of estimated average dynamic load on service failure probability. This variable for dynamic load was not directly measured by wheel impact load detectors. Such data would have been preferable but were unavailable for most of the locations where broken rails were recorded. Instead, the value was calculated using the average gross rail load data and track speed at each location using the formula in the American Railway Engineering and Maintenance-of-Way Association manual (*19*, p. 16-10-9). The values ranged from 30,000 to 55,000 lb and thus do not represent the full spectrum of dynamic loads, particularly the most damaging ones (*23*).

The final term is negative, indicating that as average dynamic load increases, service failure probability decreases. This is an unexpected result and the opposite of what was suggested by a single variable analysis conducted before developing the multivariate model (*11*). However, the relative effect of this term is weak. For example, at an annual tonnage level of 50 MGT, on 136-lb rail, in tangent track, varying the annual wheel passes between the highest and lowest possible values changes P_{SF2} by approximately 0.17. Under the same conditions, varying the dynamic load term between its extreme values changes P_{SF2} by only 0.03. In the stepwise regression, this final term added to the model has the least predictive ability of the other terms (as indicated by the low chi-square value) (Table 1). The artificial nature of the computed value for this term, combined with the way the model handled it, suggests that it does not represent a real physical relationship.

Table 1 also indicates that during the stepwise regression process, an interaction term between rail age and annual gross tonnage was initially included in the model. By multiplying rail age by annual gross tonnage, this term provided an estimate of cumulative tonnage; however, it was not a direct measure of this important variable. Although this estimated cumulative tonnage term was initially significant, as more detailed terms describing the effects of rail age, rail size, annual tonnage, turnouts, curvature, etc., were added to the model, the cumulative tonnage term became less significant and was finally removed. Thus, the variance in service failure probability that was initially explained by the estimated cumulative tonnage term in a model with two terms could be better explained by a model with more terms and a combination of effects involving other variables.

This result should not be interpreted as meaning that cumulative tonnage is not an important factor in predicting the occurrence of broken rails. There are several reasons for this. First, if a direct measure of accumulated tonnage was available for the analysis, a term based on it might not have been removed. Such a variable would have been preferable, but it was not consistently available systemwide. Second, the two elements of the cumulative tonnage term—age and annual tonnage—appear in several other terms, indicating that these factors are important. Third, the calculated cumulative tonnage term was a strong predictor in the absence of other variables. This term was removed only when most of the other terms were added. This finding is consistent with industry experience that cumulative tonnage is a good predictor of broken-rail frequency. Part of the point of the multivariate statistical approach is that it reveals other variables that have subtler or perhaps interactive effects.

It is also interesting which parameters did not appear in the final model. The effects of grade, speed, average wheel load, and insulated joints were tested and not found to significantly improve the predictive ability of the model and were not included in the final model. Conversely, other variables would have been useful relative to physical factors that cause broken rails, but the requisite data were unavailable. In addition to the dynamic load and cumulative tonnage variables, other

TABLE 1 Model Term Selection Order

Step	Term Added	Term Removed	Chi-Squared
1	Wheel Passes / Rail Weight	--	155
2	Annual Gross Tonnage x Rail Age	--	--
3	(Wheel Passes / Rail Weight)^2	--	202
4	Annual Gross Tonnage / Rail Weight	--	63
5	Rail Age	--	204
6	Turnout	--	41
7	Degree of Curve x Rail Age	--	47
8	(Degree of Curve x Rail Age)^2	--	33
9	Dynamic Load / Rail Weight	--	8
10	--	Annual Gross Tonnage x Rail Age	--

variables that would have been useful were rail steel type, rail surface roughness, neutral temperature, and temperature at the time of the break. One benefit of model development is that it can highlight the importance of certain types of data and the potential benefit from expanded database development. It would be advantageous to include these variables in the future development of comparable databases.

Retrospective Service Failure Model Performance

Two methods were used to evaluate the ability of the retrospective model to predict locations where service failures occurred. The first method calculated a goodness-of-fit statistic for the model on the basis of P_{SF2} computed for each record in the input data. If the model completely accounted for all of the sources of variance, $P_{SF2} = 1$ would be expected at all of the service failure locations and $P_{SF2} = 0$ at all of the locations where service failures did not occur. In this case, the summation of P_{SF2} over all service failure locations should equal the total number of service failures, and the summation of $1 - P_{SF2}$ over all locations where service failures did not occur should equal the total number of locations where they did not occur. All sources of variance are unlikely to be accounted for by any statistical model. Therefore, when the summations are computed for actual values of P_{SF2}, they will correctly account for only a percentage of the total. This percentage reflects the goodness of fit or the amount of variance explained by the retrospective model (*14*). Using this approach, the goodness-of-fit statistic is calculated using the following expression, where n_{sf} is the actual number of locations where service failures occurred, and n_{nosf} is the number of locations where they did not:

$$\text{Goodness of fit} = \frac{\left(\sum_{n_{sf}} P_{SF2} + \sum_{n_{nosf}} 1 - P_{SF2}\right)}{(n_{sf} + n_{nosf})}$$

$$= \frac{(1,507 + 1,462)}{(1,861 + 1,815)}$$

$$= 0.808$$

On the basis of this analysis, the retrospective model accounted for 80.8% percent of the variance in the service failure data.

The second method to evaluate the performance of the model was to compare the value of P_{SF2} to the event that actually occurred at a location. The decision criterion, or threshold value, for service failure prediction was $P_{SF2} = 0.5$. If $P_{SF2} < 0.5$, it was classified as predicting no failure and if > 0.5, it was classified as predicting a service failure. Of these predictions, 87.4% were correct (Table 2). Of the incorrect predictions, there were twice as many false positives than missed service failures. This finding indicates that the model is somewhat conservative because it is more likely to provide a false positive than miss a service failure. The decision criterion of 0.5 could be adjusted by users of the model to make the results more or less conservative (*11*). Further work in which additional variables are incorporated might reduce the error rate.

These two evaluations indicate that the model had a reasonably high level of accuracy in predicting the occurrence of service failures in the database from which it was developed. The next steps in assessing the model's accuracy would be to test it using data from another time period or another railroad, or both.

Prospective Service Failure Model

To use the model to predict the annual probability of a service failure at a particular location, the retrospective model must be transformed into a prospective model. This transformation is accomplished by adjusting the value of the model specific constant, Z, to reflect the average service failure probability across the entire system of interest. There were 1,861 service failures in the test database over the 2-year period for which complete records were available. The probability that one of these service failures falls into any given segment of track is a function of the length of the segment. To capture as much detail as possible, and to avoid the use of average values over a segment that may introduce additional variance, the segments should be kept relatively short. The maximum resolution in the data available for most of the parameters of interest was 0.01 mi (52.8 ft). The total system represented by the database was approximately 23,750 mi of mainline. Thus, there were 2, 375,000 segments, each 0.01 mi in length. Given this value, the average probability that a service failure is found in any one of those segments over a 2-year period is approximately 0.00078. This probability can be converted into a new model-specific constant, Z^*, through the use of the log-odds operator (*21*):

$$Z^* = Z + \ln\left[\frac{P_{SFavg}}{(1 - P_{SFavg})}\right]$$

$$= -4.569 + \ln\left[\frac{0.00078}{(1 - 0.00078)}\right]$$

$$= -11.763$$

This new model-specific constant, Z^*, adjusts the scale of the probability calculated by the prospective service failure model so that the model predicts service failures at a rate comparable to the observed rate.

The retrospective model calculated the probability of a service failure for a 2-year period. This probability can be converted to an annual probability simply by dividing by 2 when transforming the U score into a probability. After these two adjustments are made, the annual service failure probability for any 0.01-mi segment can be calculated with the prospective service failure model. The prospective service failure probability model has the following form:

$$P_{SF} = \frac{e^U}{[2(1 + e^U)]}$$

TABLE 2 Results of Goodness-of-Fit Test for a Threshold Value of $P_{SF2} = 0.5$

Model Prediction	Actual Event	Events	Percent of Total	Outcome
Service Failure ($P_{SF2} > 0.5$)	Service Failure	1,700	87.4	Correct Prediction
No Failure ($P_{SF2} < 0.5$)	No Failure	1,513		
Service Failure ($P_{SF2} > 0.5$)	No Failure	302	8.2	False Positive
No Failure ($P_{SF2} < 0.5$)	Service Failure	161	4.4	Missed Failure

where

P_{SF} = annual probability of a service failure in the 0.01-mi segment of interest,
$U = Z^* + Y$, and
$Z^* = -11.763$, prospective-model–specific constant.

Service Failure Probability and Expected Service Failures per Mile

A cursory review of the annual service failure probabilities calculated by the prospective model might suggest that they are too low. However, the probability is based on a segment of track that is only 0.01 mi long. The calculated probability is approximately equal to the expected number of service failures per year in that 0.01-mi segment. Annual service failures per mile is a metric more typically used by North American railroads, so it is useful to calculate a per-mile rate by multiplying P_{SF} by 100.

$$\text{SF/MI/YR} = \frac{100\,e^U}{[2\,(1 + e^U)]}$$

where SF/MI/YR is the expected service failure rate on segment of interest (service failures per mile per year).

This rate can be applied to a segment of track of any length as long as the values of the parameters in the service failure model remain constant along that section of track. A service failure rate of 2 SF/MI/YR indicates that for every mile of track for which the rate applies, two service failures are expected to occur. If the track section to which this rate applied is 0.5 mi long, one service failure is expected along that length; if the section is 2 mi, four service failures are expected along that length. In all three cases, the service failure rate, 2 SF/MI/YR, is the same. The number of service failures expected in a section of track in which the service failure rate is constant is a linear function of the length of the section.

Example of Service Failure Model Application

The following example illustrates how SFPM can be used to obtain a measure of service failure probability and rate. A hypothetical 1.5-mi, single-track portion of a railroad mainline is illustrated in Figure 2, and the relevant parameters are presented in Table 3. The segment has been divided into several subsegments over which the input parameters are constant.

Some of the rail is 47 years old and weighs 132 lb/yd. The remaining rail is 5 years old and weighs 136/lb/yd. Mainline turnouts are located at Mile 0 and also at Mile 0.7, where another mainline connects to the line. A 1° curve is located between Mile 0.25 and Mile 0.45. Track speed on the segment is 50 mph. The annual traffic is 80 MGT between Mile 0.0 and Mile 0.7. At Mile 0.7, 40 MGT are routed on the connecting mainline, with the remaining 40 MGT being routed on the segment under consideration between Mile 0.7 and Mile 1.5. The average gross rail load is 100 tons eastward and 80 westward, with the higher value of 100 tons used in the calculations. The dynamic load computes to 150 tons per car, and the annual traffic of 80 MGT and a 100-ton average per car results in an estimated 3.2 million wheel passes.

Because this is the prospective model, $Z^* = -11.763$ was used to calculate the U score for each portion of the segment of interest and then transformed into an estimate of service failure rate. The estimated service failure rate (service failures per mile per year) for each subsegment is summarized in Table 4 and presented graphically in Figure 3. Multiplying each subsegment's calculated service failure rate by its length provides an estimate of the expected number of service failures per mile per year in that subsegment. Summing all of the subsegment values provides an estimate of the expected number of service failures per year on the 1.5 mi of the segment of interest. In this case, the expected number of service failures for the segment is 0.316.

The service failure profile in Figure 3 highlights how interactions between the various parameters affect service failure rate. Between Mile 0.0 and Mile 0.1, the rail is relatively old and a turnout is present. The combination of these two factors results in a relatively high service failure rate prediction. At Mile 0.1, the service failure rate drops as the rail is no longer close enough to the turnout to be subject to its effects. Between Mile 0.1 and Mile 0.25, the track is tangent but the old rail produces a higher service failure rate than on the segment between Mile 0.45 and Mile 0.6, where the track is tangent but the rail is relatively new. This difference in service failure rate illustrates the importance of rail age. Under the traffic conditions in this example, the age difference of 42 years results in a service failure rate that is 16 times higher on the older section of rail. At Mile 0.25, the track transitions from tangent to a 1° curve and the service failure rate increases approximately threefold. Compared with Mile 0.45, where the new rail transitions from curve to tangent and the service failure rate increases by a factor of only 1.5, the increase in service failure rate at Mile 0.25 is large. This increase results from the interaction of rail age and curvature that makes the old rail on this subsection of track sensitive to curvature. At Mile 0.3, the rail on the 1° curve changes from rail that is 47 years old to rail that is 5 years old. The model suggests that newer rail is less sensitive to curvature, so the service failure rate drops from 0.86 to 0.03 service failures per mile per year. Because there is one half the traffic between Mile 0.7 and Mile 1.5 than there is between Mile 0.0 and Mile 0.7, the service failure rate is also correspondingly lower.

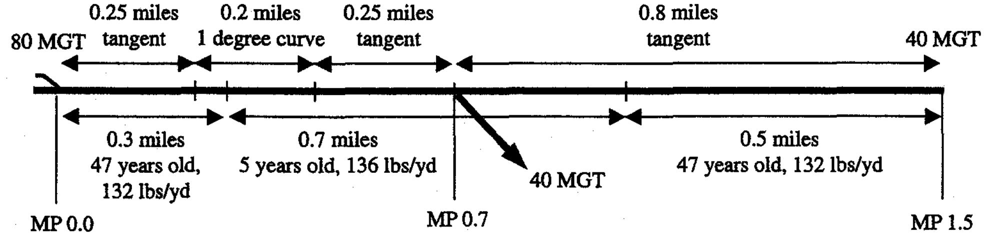

FIGURE 2 A hypothetical section of mainline track.

TABLE 3 Input Parameters for a Hypothetical Section of Mainline Track

Start MP	End MP	Z	A (age)	C (degree)	T (MGT)	S (pounds)	W (million)	P (tons)	N (turnout)	U
0.00	0.10	-11.763	47	0	80	132	3.2	150	1	-3.25
0.10	0.25	-11.763	47	0	80	132	3.2	150	0	-5.04
0.25	0.30	-11.763	47	1	80	132	3.2	150	0	-4.04
0.30	0.45	-11.763	5	1	80	136	3.2	150	0	-7.47
0.45	0.60	-11.763	5	0	80	136	3.2	150	0	-7.60
0.60	0.70	-11.763	5	0	80	136	3.2	150	1	-5.80
0.70	0.80	-11.763	5	0	40	136	1.6	150	1	-8.26
0.80	1.00	-11.763	5	0	40	136	1.6	150	0	-10.05
1.00	1.50	-11.763	47	0	40	132	1.6	150	0	-7.53

NOTE: MP = milepost.

CONCLUSIONS

A simple risk analysis showed that broken rails are the leading cause of severe accidents as measured by the number of cars derailing. Improved detection and prevention of broken rails potentially has important safety and economic benefits. Further, service quality and reliability benefits can accrue if the incidence of broken rails can be reduced. Improving the ability to predict the conditions that can lead to broken rails can help railroads allocate inspection, detection, and preventive resources more efficiently, thereby enhancing safety and reducing service interruptions due to broken rails.

A statistical model was developed that provides probabilistic estimates of the likelihood of service failure occurrence on the basis of engineering and operational input parameters. Although further validation needs to be conducted, the service failure prediction model shows promise in improving the ability to predict the occurrence of broken rails. If the requisite data for a railway system can be systematically developed in a consistent, easily accessed, electronic format, the model can be applied to any portion of a system to generate location-specific estimates of service failure probability. If the data include appropriate geographical information, the service failure model could be incorporated into a geographic information system that would generate service failure and broken rail derailment profiles automatically from railway databases.

Previous models have been based on a combination of fatigue and fracture principles and a limited number of parameters available for in-service rail. The information technology and computer revolution has resulted in large, comprehensive databases and made practical the use of powerful statistical tools. The present research would not have been feasible 10 years ago. The results of these analyses, coupled with the graphical output capabilities typical of current PCs, can improve managers' access to information and enhance the quality and pace of decision making. The potential benefit of the approach is greater precision in predicting the occurrence of broken rails, along with wider availability and enhanced interpretation of the results. This capability is important as railroads strive to improve safety and at the same time more efficiently use their resources and extract more value from assets such as rail.

ACKNOWLEDGMENTS

Douglas Simpson and Todd Treichel provided helpful assistance and review of the statistical methods used. Thanks also to Hank Lees, Tom Wright, and Scott Staples, who assisted in obtaining the data needed for the analysis. Frederick Lawrence patiently shared his insight regarding the fracture mechanics of rail, and Kevin Sawley and David Davis also provided helpful discussion. The first two

TABLE 4 Service Failure Probabilities Along a Hypothetical Track Section

Start MP	End MP	Length	U	SF/MI/YR	Expected SF
0.00	0.10	0.10	-3.25	1.866	0.187
0.10	0.25	0.15	-5.04	0.322	0.048
0.25	0.30	0.05	-4.04	0.865	0.043
0.30	0.45	0.15	-7.47	0.028	0.004
0.45	0.60	0.15	-7.60	0.025	0.004
0.60	0.70	0.10	-5.80	0.151	0.015
0.70	0.80	0.10	-8.26	0.013	0.001
0.80	1.00	0.20	-10.05	0.002	0.000
1.00	1.50	0.50	-7.53	0.027	0.014
Total (0.0 to 1.5)				0.211	0.316

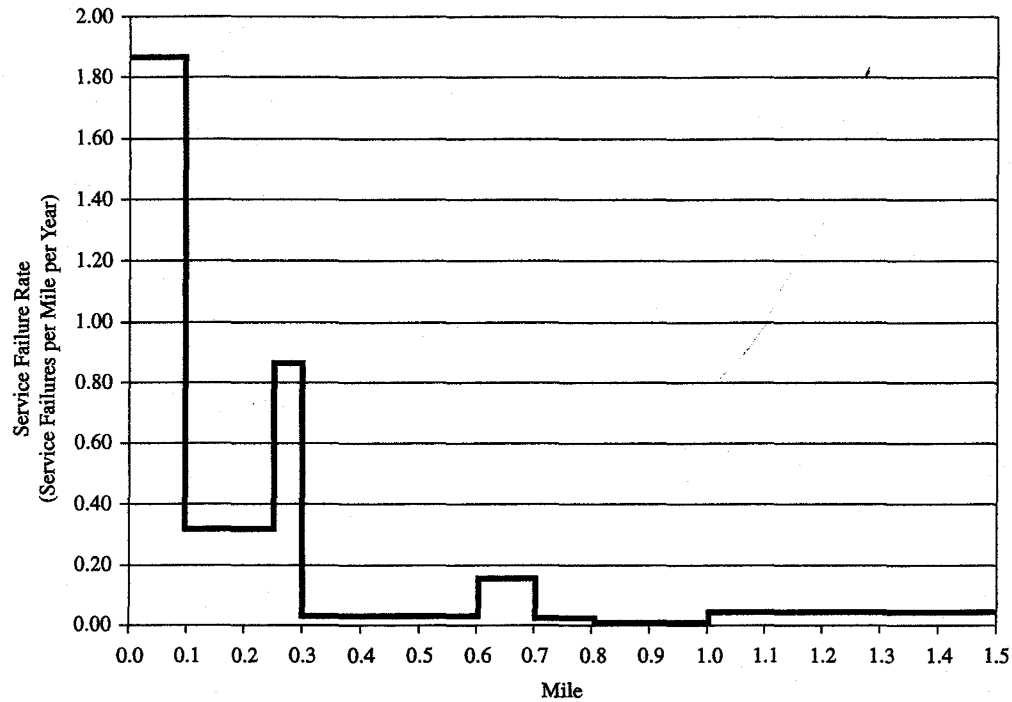

FIGURE 3 Service failure probability along a hypothetical track segment.

authors would like to express their gratitude to the Burlington Northern Santa Fe Railway for its support of this research.

REFERENCES

1. Thompson, A. W. Service Performance of Railroad Rails. In *Symposium on Railroad History, Vol. 2* (A. W. Thompson, ed.), A. C. Kalmbach Library, Chattanooga, Tenn., 1992.
2. Aldrich, M. The Peril of the Broken Rail: Carriers, Steel Companies, and Rail Technology, 1900–1945. *Technology and Culture,* Vol. 40, 1999, pp. 263–291.
3. Shyr, F.-Y., and M. Ben-Akiva. Modeling Rail Fatigue Behavior with Multiple Hazards. *Journal of Infrastructure Systems,* Vol. 2, No. 2, June 1996, pp. 73–82.
4. Palese, J. W., and T. W. Wright. Risk-Based Ultrasonic Rail Test Scheduling on Burlington Northern Santa Fe. *Proc., AREMA 2000 Annual Conference* (CD-ROM), Chicago, Ill., American Railway Engineering and Maintenance-of-Way Association, Landover, Md., 2000.
5. Palese, J. W., and A. M. Zarembski. BNSF Tests Risk-Based Ultrasonic Detection, *Railway Track and Structures.* Vol. 97, No. 2, 2001, pp. 17–21.
6. Orringer, O., and M. W. Bush. Applying Modern Fracture Mechanics to Improve the Control of Rail Fatigue Defects in Track. *American Railway Engineering Association Bulletin,* Bulletin 689, Vol. 83, 1982, pp. 19–53.
7. Orringer, O., Y. H. Tang, D. Y. Jeong, and A. B. Perlman. *Risk-Benefit Assessment of Delayed Action Concept for Rail Inspection.* Office of Research and Development, FRA, U.S. Department of Transportation, 1999.
8. Lawrence, F. V., C. P. L. Barkan, and H. M. Reis. *Summary of a Workshop on New Technologies for the Detection of Broken Rail.* Railroad Engineering Program, University of Illinois, Urbana–Champaign, 2000.
9. Jeong, D. Y. Progress in Rail Integrity Research. *Proc., AREMA 2001 Annual Conference* (CD-ROM), Chicago, Ill., American Railway Engineering and Maintenance-of-Way Association, Landover, Md., 2001.
10. Roney, M. D., and W. Ebersohn. Maintaining Optimal Wheel and Rail Performance. In *Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues.* International Heavy Haul Association, Virginia Beach, Va., May 2001.
11. Dick, C. T. *Factors Affecting the Frequency and Location of Broken Railroad Rails and Broken Rail Derailments.* M.S. thesis. University of Illinois, Urbana–Champaign, 2001.
12. Barkan, C. P. L., C. T. Dick, and R. Anderson. Railroad Derailment Factors Affecting Hazardous Materials Transportation Risk. In *Transportation Research Record: Journal of the Transportation Research Board,* No. 1825, TRB, National Research Council, Washington, D.C., 2003, pp. 64–74.
13. FRA. Accident Incident Database, 1994–2000. www.safetydata.fra.dot.gov/officeofsafety/. Accessed May 2000.
14. Ben-Akiva, M., and S. R. Lerman. *Discrete Choice Analysis.* MIT Press, Cambridge, Mass., 1985.
15. Reiff, R. P. *Proc., Rail Defect and Broken Rail Defects Expanded Workshop, July 1997.* Transportation Technology Center, Association of American Railroads, Pueblo, Colo., 1997.
16. Clark, R. Defective and Broken Rails—How They Occur. In *Summary of a Workshop on New Technologies for the Detection of Broken Rail* (F. V. Lawrence, C. P. L. Barkan, and H. M. Reis, eds.), Railroad Engineering Program, University of Illinois, Urbana–Champaign, 2000.
17. Hay, W. W. *Railroad Engineering,* 2nd ed. John Wiley and Sons, New York, 1982.
18. Selig, E. T., and J. M. Waters. *Track Geotechnology and Substructure Management.* Thomas Telford Publications, London, 1994.
19. *Manual for Railway Engineering* (CD-ROM). American Railway Engineering and Maintenance-of-Way Association, Landover, Md., 2000.
20. Lawrence, F. V. An Analytical Model for Broken Rail. In *Summary of a Workshop on New Technologies for the Detection of Broken Rail* (F. V. Lawrence, C. P. L. Barkan, and H. M. Reis, eds.), Railroad Engineering Program, University of Illinois, Urbana–Champaign, 2000.
21. McCullagh, P., and J. A. Nelder. *Generalized Linear Models,* 2nd ed., Chapman and Hall, New York, 1989.
22. Sawley, K., and R. Reiff. *Rail Failure Assessment for the Office of the Rail Regulator.* Transportation Technology Center, Association of American Railroads, Pueblo, Colo., 2000.
23. Kalay, S., and A. Tajaddini. *Condemning Wheels Due to Impact Loads: Preliminary Survey—Six Railroads' Experience.* Report R-754. Chicago Technical Center, Association of American Railroads, Chicago, Ill., 1990.

Publication of this paper sponsored by Committee on Railway Maintenance.