Development of a Crash Reduction Model for Horizontal Curves

by

William C. Taylor
Cyrus Safdari

Department of Civil and Environmental Engineering
Michigan State University
October 1997

MICHIGAN STATE UNIVERSITY
1855

COLLEGE OF ENGINEERING
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Reciepient's Catalog No.
4. Title and Subtitle Development of a Crash Reduction Model for Horizontal Curves.		5. Report Date October 1997
		6. Performing Organization Code
7. Author(s) William C. Taylor and Cyrus Safdari		8. Performing Organization Report No.
9. Performing Organization Name and Address Michigan State University Civil & Envir Engineering Department East Lansing, MI 48824-1226		10. Work Unit No.
		11. Contract or Grant No. MDOT-94-1521-B
12. Sponsoring Agency Name and Address Michigan Department of Transportation 425 W. Ottawa Lansing, MI 48933		13. Type of Report and Period Covered FINAL REPORT
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract In Michigan over 25 percent of fatal traffic crashes take place on non-freeway trunkline highways. Research has consistently demonstrated that crash rates on horizontal curves are many times higher than that of the tangent sections on the same road, and most studies have found the degree of curvature to be the most significant single factor related to curve crashes. However, other roadway features, such as superelevation and skid resistance of the pavement surface, traffic control elements, driving environment and human factors, individually or in combination are major contributors as well.

The purpose of this study was to analyze horizontal curve crashes experienced on two-lane trunkline roads in the State of Michigan, and to devise procedures to identify curved road segment grouping attributes that correspond to the crash rate on curves. A second goal was to identify curves that exhibited crash frequencies significantly higher than the mean for their group, or which potentially may exhibit such crash frequencies.

Two sets of simple regressions, one for the curve crashes and the other for the difference between the curve crashes and tangent crashes, versus the independent variables were performed. The regression lines, and the coefficients of regression all indicate that simple regression models are poor predictors of crashes. The use of multiple regressions models improved the predictive capability of the models, but these models still explained only a small percentage of the variation in the crash rate on curves.

Discriminant analysis was used to determine the variables which can be used to distinguish between high and low crash rate curves.

The curve length, the presence of a turn or curve warning sign, the radius of the curve and the tangent crash rate are the discriminating variables identified. Using these variables 79.1% of the curves were correctly classified.

Discriminant analysis provides information useful in meeting the objectives of this study. Specifically, it can be used to identify those characteristics of low crash rate curves which distinguish them from high crash rate curves. Having done this, it can be used to identify those curves with a high crash rate that should (based on their characteristics) have a low crash rate. These curves are the ones that should be studied for possible countermeasure implementation.

Cluster analysis was used to identify the variables with a strong association with the crash rate. One cluster identified the variables associated with curves that have a low crash rate, a second cluster was formed around curves with an intermediate crash rate, and the third around high crash rate curves.

The clustering of high, medium and low crash rate curves with other variables is clear, with cluster one having a crash rate of 3.08, cluster two a crash rate of 7.78 while the third cluster has a crash rate of 18.05.

The same variables identified in the discriminant analysis were important in the cluster analysis. The ADT, curve radius and length, and the presence of traffic control devices (arrow and chevron) are all important in defining the clusters.

17. Key Words		18. Distribution Statement	
Horizontal Curves, Discriminant Analysis, Cluster Analysis, Crash Reduction			
19. Security Classif. (of this report) Unclassified	20. Security Classif (of this page) Unclassified	21. No. of Pages 123	22. Price

Table of contents

List of figures

List of tables

EXECUTIVE SUMMARY

Introduction

In Michigan over 25 percent of fatal traffic crashes take place on non-freeway trunkline highways. Research has consistently demonstrated that crash rates on horizontal curves are many times higher than that of the tangent sections on the same road, and most studies have found the degree of curvature to be the most significant single factor related to curve crashes. However, other roadway features, such as superelevation and skid resistance of the pavement surface, traffic control elements, driving environment and human factors, individually or in combination are major contributors as well.

Several models, most notably the Glennon Model and the Zegeer Model, have been developed to explain the relationship between curve features and curve crashes. However, when applied to Michigan data, their results are not sufficiently reliable to establish corrective or preventative programs.

The purpose of this study was to analyze horizontal curve crashes experienced on two-lane trunkline roads in the State of Michigan, and to devise procedures to identify curved road segment grouping attributes that correspond to the crash rate on curves. A second goal was to identify curves that exhibited crash frequencies significantly higher than the mean for their group, or which potentially may exhibit such crash frequencies.

Regression analysis results

To accomplish the objective of this study, a multi-step approach was utilized. Step one was to acquire geometric data for all the rural, two-way, two-lane trunkline highways in Michigan. Based on the selection criteria shown in Table 1, the candidate curves were selected and the control section (reference system used by MDOT for trunklines) and the mile points of the beginning and ending of the curves were noted.

Table 1: Selection criteria

Rural two-lane, two-way.

No taper, no extra lanes.

No curb, no parking.

No median, and no intersections.

At least 306 meters (0.19 mile, about 1000 feet) of tangent at each end of each curve.

At least 611 meters (0.38 mile) of tangent between the two curves.

In addition to the data from the geometric file and the crash data file, data was obtained from the photo logs and the curve superelevation and pavement friction were obtained in the field.

For each of the 220 curves, all the crashes corresponding to the mile points from 306 meters (0.19 mile) before the start of the curve to 306 meters (0.19 mile) after the end of the curve were extracted from the MDOT crash files for the six year period of 1989 to 1994, yielding 3107 total crashes.

The crash report forms for these crashes were obtained and processed to verify the location of the individual crash as being on the curve or on the tangent.

The Geometric data included 44 variables such as degree of curvature, curve length, average lane width, total shoulder width (right and left), etc. The crash data consisted of 120 variables such as mile point of crash, highway area type, highway area code, etc.

The photo log data were used for variables such as the presence of traffic signs (arrow, chevron, etc.), the mile point at which the curve was first observed, etc. The data also included a subjective measure of the roadside clearance/hazard, on a scale of one to seven. The data acquisition was performed twice, once for each direction of the traffic flow.

The field data collection was performed to obtain only two variables; a measure of the superelevation of the road, and a measure of the skid resistance of the pavement surface.

For the analyses used in this project, only the Curve Related crashes consisting of the following types of crashes were considered:

Table2: Curve related crashes

CODE	DESCRIPTION
000	Miscellaneous 1 Vehicle
010	Overturn
060	Fixed Object
070	Other Object
141	Head-on
543	Side-Swipe Opposite

Selection of the curve related crashes yielded 994 crashes corresponding to the 178 roadway segments which had at least one related crash. Not all of the selected roadway segments had crashes in both the tangent and curve portion of the roadway segment.

In addition to analyzing all related crashes, crashes occurring under different road surface conditions, weather conditions and lighting conditions were also analyzed.

A sub-set of curves consisting of only those with the field data were analyzed separately.

All analyses were based on the assumption that the non-measurable, non-quantifiable environmental and traffic conditions along the entire length of each curve can be considered to be the same as that of the average of the tangents at each end. The basic unit of tangent length at each end of the curves was 306 meters (0.19 mile). To compare the curve crashes with the tangent crashes, 611 meters (0.38 mile) was used as a unit length and the curve crash rate was adjusted for this length. The resulting variables were called Cper380 for curve crashes and Tper380 for tangent crashes. Another variable, C-T was defined to represent the difference between curve and tangent crashes. This variable has a value equal to Cper380-Tper380.

Two sets of simple regressions, one for the curve crashes (Cper380), and the other for the difference between the curve crashes and tangent crashes C-T, versus the independent variables were performed. The regression lines, and the coefficients of regression all indicate that simple regression models are poor predictors of crashes. The use of multiple regressions models improved the predictive capability of the models, but these models still explained only a small percentage of the variation in the crash rate on curves.

The difference between the design speed and the advisory speed or posted speed limit was calculated and linear regression models for the Cper380 and C-T values were developed. These models were also found to indicate a weak correlation.

The curve crash data versus their predicted value from the Glennon and Zegeer models were then calculated. While both models appear to show the correct trends, neither model explains the curve crash variation in Michigan data, as shown in Figure 1 and Figure 2.

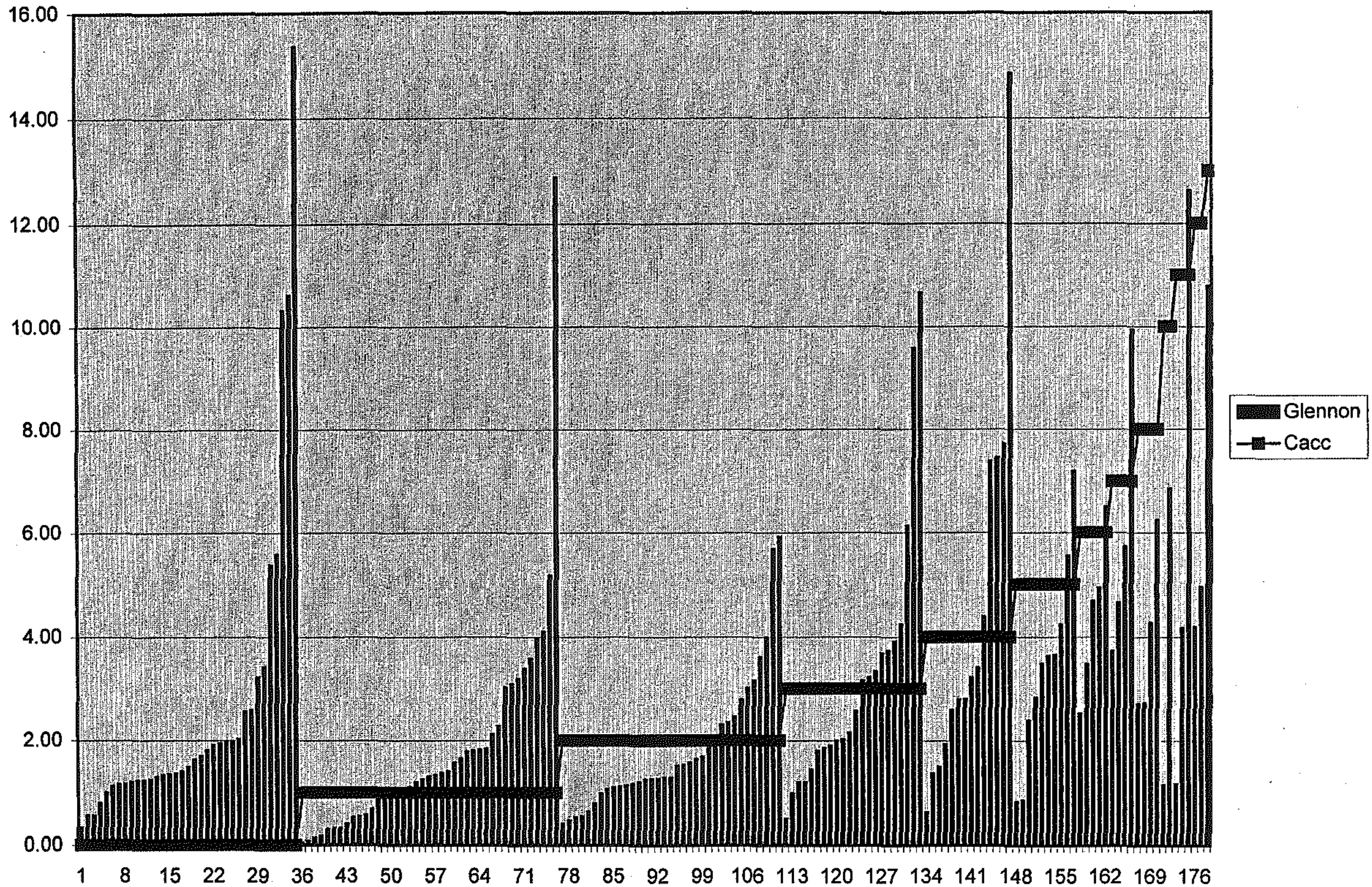


Figure 1 Comparison of the predicted number of curve crashes using Glennon's model (Glennon), and the actual number of curve crashes (Cacc)

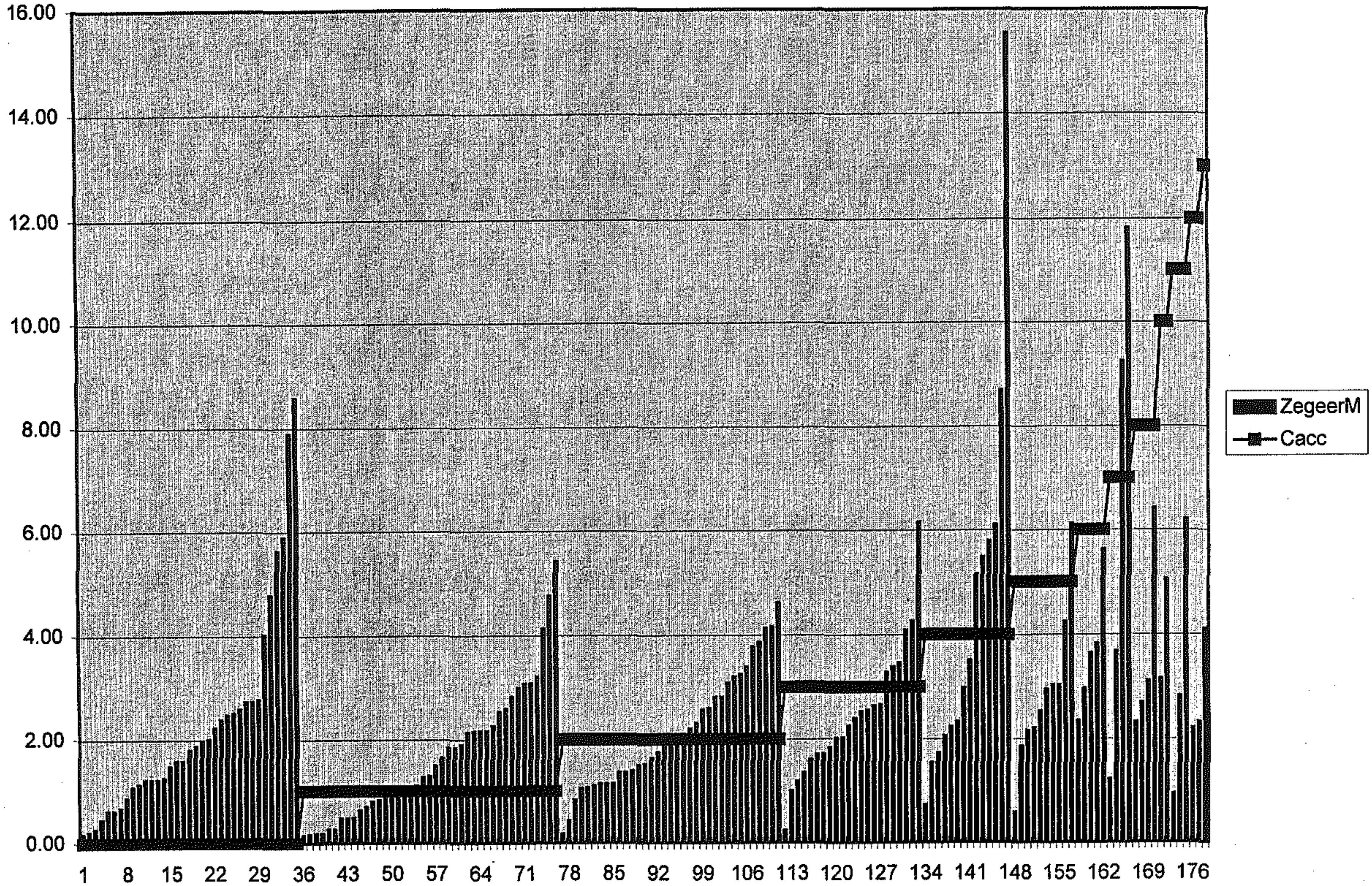


Figure 2 Comparison of the predicted number of curve crashes using Zegeer's model without spiral (ZegeerM), compared with the actual number of curve crashes (Cacc)

Discriminant analysis and cluster analysis models

The conclusion from these analyses was that neither simple linear regression nor multiple linear regression are powerful enough tools to depict the large variations in the curve crash rate, or to be useful in establishing crash reduction policies for the Department.

Having determined that the variation in crash frequency found on Michigan curves can not be satisfactorily explained by models based on simple linear regression, simple non-linear regression, multiple linear regression or multiple non-linear regression, alternative statistical techniques were tested to determine which techniques could satisfactorily explain the data variation. Discriminant analysis and cluster analysis techniques were found to accomplish the task.

DISCRIMINANT ANALYSIS

Discriminant analysis is a multivariate technique used to distinguish between two or more groups of cases and for studying the overlap between groups, or divergence of one group from the others.

The variables with a high contribution toward explaining membership in each group, generally not all the original variables, are considered the predictor variables or the discriminating variables.

For this study, discriminant analysis was used to determine the variables which can be used to distinguish between high and low crash rate curves. Table 3 shows the results of the analysis using the curve crash rate as the grouping variable.

		GRPCLT5	Predicted Group Membership 1.00	2.00	Total
Original	Count	1.00	64	24	88
		2.00	26	64	90
	%	1.00	72.7	27.3	100.0
		2.00	28.9	71.1	100.0

a. 71.9% of original grouped cases correctly classified.

Variables	Tolerance	Sig. of F to Remove	Wilks' Lambda
HCLFT	.866	.001	.827
HCRFT	.848	.002	.822
ADT	.978	.002	.821

Table 3 Results of the discriminant analysis for curve crash rate (Cper380)

Group one represents curves with the expected value of the crash rate is lower than 5.0 crashes per 306 meters (1000 ft), and group 2 represents curves with an expected value greater than 7 crashes per 306 meters.

The curve length, the presence of a turn or curve warning sign, the radius of the curve and the tangent crash rate are the discriminating variables identified in this case. Using these variables 79.1% of the curves were correctly classified. A second analysis was conducted using the difference between the curve crash rate (Cper380) and the tangent crash rate (Tper380) as the grouping measure.

As shown in Table 4, the curve radius, curve length and the presence of a warning sign are the three most important discriminating variables. For this analysis, 75.6% of the curves were correctly classified using these three variables. Using this model, 90.7% of the high crash rate curves were correctly identified.

Discriminant analysis provides information useful in meeting the objectives of this study. Specifically, it can be used to identify those characteristics of low crash rate curves which distinguish them from high crash rate curves. Having done this, it can be used to identify those curves with a high crash rate that should (based on their characteristics) have a low crash rate. These curves are the ones that should be studied for possible countermeasure implementation.

		LOCMNST	Predicted Group Membership 1.00	2.00	Total
Original	Count	1.00	5	23	28
		2.00	10	97	107
	%	1.00	17.9	82.1	100.0
		2.00	9.3	90.7	100.0

75.6% of original grouped cases correctly classified.

Step	Variables	Tolerance	Sig. of F to Remove	Wilks' Lambda
1	HCRFT	1.000	.000	
2	HCRFT	.987	.003	.917
	CTSIGN	.987	.004	.912

Table 4 Results of the discriminant analysis for modified curve minus tangent crash rate (ModC-T)

Using the discriminant analysis results from the modified Cper380 analysis, sixteen curves fell in this category. The crash rate on these curves ranged from 7.13 to 21.71 when they should have fallen in the group with a crash rate below 5.0. These curves are shown in Table 5, along with the value of some of the variables used in the analysis.

The significant characteristics of these curves include:

Most do not have curve signs, target arrows and delineators

There are no chevrons

The observed sight distance is usually short

The radius is relatively large

The tangent crash rate is low

CRVno	CS	BMP	CTsign	CHEVRON	ARROW	DLNTR	OBSDSTW	HCLFT	HCRFT	Tper380	Cper380
136	45012	5540	0	0	1	1	10	845	1042	0.00	7.13
14	5051	7280	0	0	0	0	40	264	2865	1.00	7.60
72	24011	4377	1	0	0	0	23	1056	2292	3.00	7.60
200	73131	0	0	0	0	0	0	264	2865	2.00	7.60
3	2021	15020	0	0	0	0	40	739	1910	1.00	8.14
39	12021	490	0	0	0	0	70	739	2292	3.00	8.14
33	10011	5620	1	0	1	0	33	475	2865	0.00	8.44
82	28052	5530	0	0	1	0	40	475	1910	1.00	8.44
81	28052	4790	1	0	0	0	50	634	2865	2.00	9.50
94	31013	5810	0	0	0	1	30	370	1910	3.00	10.86
117	38071	7490	1	0	1	0	10	1214	2865	8.00	13.22
87	30062	1640	1	0	0	1	10	1478	2456	0.00	13.57
156	51011	50	0	0	0	0	50	581	1146	1.58	13.82
19	8011	8990	0	0	0	1	80	211	1763	1.00	19.00
193	67011	2130	0	0	1	1	40	475	1637	4.00	21.11
172	58032	4150	0	0	0	0	80	370	2644	4.00	21.71

Table 5 Curves with a high curve crash rate (Cper380) from the discriminant analysis

CLUSTER ANALYSIS

Cluster Analysis is a systematic technique to look for regularities in a data set. Once the regularities are depicted, this procedure groups the data based on these regularities and their interpretations. Unlike Discriminate Analysis, which requires prior knowledge of the group membership for the data cases, cluster analysis does not require such knowledge. Cluster analysis was used to identify the variables with a strong association with the crash rate. While any number of clusters can be created, three clusters were used in this study. One cluster identified the variables associated with curves that have a low crash rate, a second cluster was formed around curves with an intermediate crash rate, and the third around high crash rate curves.

Utilizing cluster analysis produced results which proved to be useful for the objectives of this study. Table 6 shows the output for a three cluster case in which Modified Cper380, as discussed previously, was used to define the number of curves included in the analysis.

The clustering of high, medium and low crash rate curves with other variables is clear, with cluster one having a crash rate of 3.08, cluster two a crash rate of 7.78 while the third cluster has a crash rate of 18.05. Some variables, such as curve length and radius, show great variations between at least two of the three clusters. This is an indication of an important variable in the prediction model. The important variables are shown in Table 7. The same variables identified in the discriminant analysis were important in the cluster analysis. The ADT, curve radius and length, and the presence of traffic control devices (arrow and chevron)are all important in defining the clusters. Interestingly, the high crash

	Cluster		
	1	2	3
ADT	472.72	536.05	549.14
ALW	11.31	11.19	11.06
ARROW	.21	.09	.29
CHEVRON	.03	.03	.13
CLRNCW	3.69	3.66	4.09
CTSIGN	.34	.44	.56
DLNTR	.31	.19	.27
EDGLN	1.00	.98	1.00
GRAIL	.21	.13	.23
HCLFT	1704	590	520
HCRFT	2471	2383	963
MODCPER	3.08	7.78	18.05
MPHS	.10	.09	.30
NPZC	.90	1.06	1.96
OBSDSTW	45.24	44.27	38.37
PSL	54.66	54.53	53.29
PSW	10.79	6.56	7.03
SCT	1.66	1.53	1.60
TPER380	2.52	3.44	2.98
TSW	19.45	18.72	18.56

Table 6 **The numerical values of all variables in defining the clusters grouped by the modified curve crash rate (ModCper)**

	Cluster		
	1	2	3
ADT	472.72	536.05	549.14
ALW			
ARROW	.21	.09	.29
CHEVRON	.03	.03	.13
CLRNCW			
CTSIGN			
DLNTR			
EDGLN			
GRAIL			
HCLFT	1704	590	520
HCRFT	2471	2383	963
MODCPER	3.08	7.78	18.05
MPHS			
NPZC			
OBSDSTW			
PSL			
PSW			
SCT			
TPER380			
TSW			

Table 7 **The numerical values of the important variables in defining the clusters grouped by the modified curve crash rate (ModCper)**

rate curves are associated with the highest probability of having chevrons and target arrows deployed. However, this is explained by the fact that this cluster contains the short radius curves, where these devices tend to be deployed. Perhaps the most interesting cluster is the third one, which clusters moderately high crash rate curves with curves of large radius but short length. These tend to not have traffic control devices deployed because of their large radius and subsequently their high design speed.

Similar results were found when C-T was used as the grouping variable. This is consistent with the results above, since most of the misclassified curves had a low value of Tper380, they would fall in the high range of C-T values.

The results of the cluster analysis are consistent with prior studies, but they also provide additional information that may be useful in reducing traffic crashes. Low crash rates are clustered with curves with a large radius and long length. The average radius for curves in this group (based on modified Cper380) is 398 meters (1305 ft). The average length for the same curves is 274 meters (900 ft). These curves tend to have target arrows but no chevrons.

High crash rates are clustered with short, sharp curves as expected. These curves tend to have both chevrons and target arrows in place, but still tend to experience crashes because of their geometry.

The third cluster is the most difficult to explain, and possibly the group of curves where countermeasures may be most effective. These curves have a crash rate over twice as high as the low crash rate curves, even though they have approximately the same radius. The primary geometric difference is that they are very short curves, averaging 95 meters (312 ft). These curves generally do not have chevrons or target arrows in place.

Chevrons and target arrows are not intended for these types of curves according to the Michigan Manual of Uniform of Traffic Control Devices (MMUTCD), since they do not constitute a sharp change in alignment. However, based on the analysis, it may be appropriate to consider the use of these signs to increase the visibility of the curves.

This same clustering of curves into these groups are observed whether the crash rate variable was Cper380, Modified Cper380, C-T, or modified C-T. There were approximately 70 curves that belong to this cluster. Table 8 lists the curves for which both the Cper380 and C-T were significantly higher than the average for this cluster.

CRVno	CS	BMP	CTsign	CHEVRO	ARROW	DLNTR	BSDST	HCLFT	HCRFT	Tper380	Cper380	CmnsT
39	12021	490	0	0	0	0	70	739	2292	**3.00**	**8.14**	5.14
200	73131	0	0	0	0	0	0	264	2865	**2.00**	**7.60**	5.60
68	23051	2220	1	1	0	0	20	845	2083	**6.00**	**11.88**	5.88
177	61012	4910	0	0	0	0	48	327	2292	**12.00**	**18.39**	6.39
14	5051	7280	0	0	0	0	40	264	2865	**1.00**	**7.60**	6.60
4	2021	23640	0	0	0	1	70	581	2865	**0.00**	**6.91**	6.91
3	2021	15020	0	0	0	0	40	739	1910	**1.00**	**8.14**	7.14
82	28052	5530	0	0	1	0	40	475	1910	**1.00**	**8.44**	7.44
81	28052	4790	1	0	0	0	50	634	2865	**2.00**	**9.50**	7.50
94	31013	5810	0	0	0	1	30	370	1910	**3.00**	**10.86**	7.86
33	10011	5620	1	0	1	0	33	475	2865	**0.00**	**8.44**	8.44
12	5031	3900	1	0	0	0	60	370	2292	**2.00**	**10.86**	8.86
214	81031	750	1	0	0	0	10	317	2292	**10.00**	**19.00**	9.00
100	31051	9143	1	0	0	0	13	338	1910	**1.00**	**11.88**	10.88
172	58032	4150	0	0	0	0	80	370	2644	**4.00**	**21.71**	17.71
88	30062	2900	1	0	0	1	30	581	1719	**3.00**	**20.73**	17.73
19	8011	8990	0	0	0	1	80	211	1763	**1.00**	**19.00**	18.00
101	32011	3050	1	0	0	0	30	370	2292	**4.00**	**27.14**	23.14
62	22021	499	0	0	0	0	49	306	1879	**16.00**	**45.86**	29.86
140	45013	11700	1	0	1	0	60	634	1910	**2.00**	**34.83**	32.83
215	81031	1370	1	1	0	0	70	370	2989	**6.00**	**43.43**	37.43
71	23111	3670	1	0	0	0	30	211	1910	**3.00**	**47.50**	44.50

Table 8 Curves with both a high curve crash rate (Cper380) and a high curve minus tangent crash rate (C-T)

CONCLUSIONS

Based on the analyses conducted in this study, the following conclusions were reached.

1. The variation in the crash frequency or rate between horizontal curves with similar geometry is too large to be explained by regression techniques. The only studies that report high correlation coefficients are those that aggregate curves into groups with similar characteristics and then conduct the regression analysis on the group means. This type of analysis may be useful in the design of new highways, but it is not useful in meeting the objectives of this study.
2. The predicted crash rate using existing models (Zegeer and Glennon) does not accurately depict the actual crash rates on Michigan two-way, two-lane rural trunklines. These models can not be used to identify curves locations where countermeasures could successfully be deployed to reduce crashes.
3. The distance on the approach at which the curve first becomes visible to the motorist is not highly correlated with the crash rates as a single variable, but it was found to be a contributor to some of the models that use multiple variables.
4. The addition of data on superelevation and the drag factor contributed little to the prediction capability of the models.
5. Discriminant analysis techniques, using the variables collected for this study, can successfully distinguish the high crash rate curves from the low crash rate curves. This technique can be used to identify outliers in each of the two categories (high and low) for both the absolute crash rate on curves (Cper380) or the difference in the crash rate between the curve and the tangent roadway segments (C-T).
6. Cluster analysis identified three distinct groups of curves. The group with a high crash rate (Cper380) is characterized by short radii and short curve lengths. These curves generally are marked with a curve sign, advisory speed panels and chevrons or delineators. The high crash rate on the first group of curves is probably related to

constraints the geometry imposes on driver ability to negotiate the curve at their approach speed.

The group with a low crash rate are characterized by large radii and long curve length where the curve is obvious, and little or no driver input is required.

The third group, with an intermediate crash rate, are characterized by large radii but short curve lengths. The intermediate crash rate curves appear to be the group of curves where the benefits of low cost traffic engineering measures may be most effective. The crashes on these curves may be related to the driver perception (or lack of perception) of the presence of a curve. Thus, even though the curve geometry does not require extraordinary driver input to negotiate safely, the presence of the curve is not being effectively communicated to the driver.

RECOMMENDATIONS

1. The curves identified in Table 5 from the discriminant analysis results should be targeted for analysis and potential countermeasures implementation. These sixteen curves have the characteristics of low crash rate curves, but are experiencing a high rate of crashes.
2. The curves identified in Table 9 from the cluster analysis results should be targeted for analysis and potential countermeasure implementation. These curves have been identified as experiencing a crash rate at least twice that of the average crash rate for curves in their cluster.
3. Curves characterized by a large radius and short length should be analyzed to determine if there are inexpensive countermeasures that could be applied at these curves to reduce the crash rate. These curves have been identified from the cluster analysis as having a higher crash rate than that explained by the curve geometry. The curves from this group with both a high crash rate and a large difference in the curve crash rate compared to the tangent crash rate are shown in Table 8.
4. Discriminant analysis and cluster analysis techniques should be used to analyze other sets of curves on state trunkline highways. These techniques have been useful in identifying specific curves that are candidates for countermeasures. It should be determined whether these techniques are equally valid for curves that are not screened for approach tangents and intersections. The techniques may also be useful to identify high crash rate curves on four-lane cross sections.
5. If recommendations 1, 2, and 3 are adopted, a careful before and after study should be designed to document any change in the crash rate resulting from implementation of the selected countermeasures.
6. If resources are available in the Department of Transportation, these analyses could be conducted internally. Alternatively, these analyses could form the basis of a study for the Michigan State University Center of Excellence.

CRVno	CS	BMP	CTsign	CHEVRON	ARROW	DLNTR	OBSDSTW	HCLFT	HCRFT	Tper380	Cper380	C>2Mn
23	8031	2990	1	0	1	1	50	1267	1763	**6.00**	**9.50**	2.85
35	11052	14040	1	0	0	0	10	1320	2865	**12.00**	**10.64**	3.99
117	38071	7490	1	0	1	0	10	1214	2865	**8.00**	**13.22**	6.56
87	30062	1640	1	0	0	1	10	1478	2456	**0.00**	**13.57**	6.92
92	31012	4227	0	0	0	0	17	343	477	**4.00**	**35.08**	0.87
28	10011	7470	1	0	1	1	30	158	521	**2.00**	**38.00**	3.79
152	47041	21730	1	1	0	1	60	158	286	**2.00**	**38.00**	3.79
181	62031	3160	0	0	0	0	10	264	820	**7.00**	**38.00**	3.79
211	79081	8450	0	0	1	1	18	539	1008	**7.00**	**44.71**	10.50
18	8011	7100	1	0	0	0	30	211	229	**4.00**	**47.50**	13.29
196	72051	7673	0	0	0	0	10	143	1146	**1.00**	**56.30**	22.09
85	29042	6270	0	0	0	0	20	106	1146	**4.00**	**57.00**	22.79
168	56032	8814	1	0	0	0	34	380	1146	**4.00**	**58.06**	23.85
199	73061	3930	0	1	0	1	10	370	727	**6.00**	**65.14**	30.94
29	10011	8920	1	0	1	0	30	317	215	**3.00**	**82.33**	48.13
151	47041	19440	1	1	0	0	30	211	744	**8.00**	**95.00**	60.79
177	61012	4910	0	0	0	0	48	327	2292	**12.00**	**18.39**	3.14
19	8011	8990	0	0	0	1	80	211	1763	**1.00**	**19.00**	3.75
214	81031	750	1	0	0	0	10	317	2292	**10.00**	**19.00**	3.75
88	30062	2900	1	0	0	1	30	581	1719	**3.00**	**20.73**	5.48
172	58032	4150	0	0	0	0	80	370	2644	**4.00**	**21.71**	6.47
101	32011	3050	1	0	0	0	30	370	2292	**4.00**	**27.14**	11.90
140	45013	11700	1	0	1	0	60	634	1910	**2.00**	**34.83**	19.59
215	81031	1370	1	1	0	0	70	370	2989	**6.00**	**43.43**	28.18
62	22021	499	0	0	0	0	49	306	1879	**16.00**	**45.86**	30.62
71	23111	3670	1	0	0	0	30	211	1910	**3.00**	**47.50**	32.25

Table 9 **Curves with a crash rate (Cper380) greater than twice the average for their cluster**

INTRODUCTION:

In Michigan over 25 percent of fatal traffic crashes take place on non-freeway trunkline highways. These highways typically have all the elements associated with a high number of serious crashes. Lack of a separation buffer from the opposing traffic, combined with rather high speeds, lack of, or at times adverse lighting conditions sets the stage for such crashes.

Research has consistently demonstrated that crash rates on horizontal curves are many times higher than that of the tangent sections on the same road, and most studies have found the degree of curvature to be the most significant single factor related to curve crashes. However, other roadway features, such as superelevation and skid resistance of the pavement surface, traffic control elements, driving environment and human factors, individually or in combination are major contributors as well.

Several models, most notably the Glennon Model and the Zegeer Model, have been developed to explain curve crashes. However, when applied to Michigan data, the results are not sufficiently reliable for establishing corrective or preventative programs.

OBJECTIVES:

The purpose of this study was to analyse horizontal curve crashes experienced on two-lane trunkline roads in the State of Michigan, and to devise procedures to identify curved road segment grouping attributes that correspond to the crash rate on these curves. A second goal was to identify curves that exhibited crash frequencies significantly higher than the mean for their group, or which potentially may exhibit such crash frequencies. The specific objectives were to:

1) Identify the factors influential in horizontal curve crashes based on Michigan's crash data.

2) Prepare guidelines as to where and to what extent improvement of horizontal curves is warranted.

LITERATURE REVIEW:

Modeling of Crashes on Horizontal Curves:

Prior to 1985, modeling of crash frequencies or rates on horizontal curves was normally based on a single variable. For example, Jorgenson (1) in 1978 reported a linear relationship between crashes and the degree of curvature.

In 1985, Glennon et. al. (2) published a report titled " Safety and Operational Consideration for Design of Rural Highway Curves". The research was performed to study the safety and operational characteristics of two-lane, rural highway curves. A series of independent research methodologies were employed, including (a) multivariate crash analyses; (b) simulation of vehicle/driver operations using Highway Vehicle Operation Simulation Model (HVOSM); (c) field studies of vehicle behavior of highway curves; and (d) analytical studies of specific problems involving highway curve operations.

The crash studies indicate that, in general, the Jorgenson model is correct; as curve radius decreases, crash rate increases. However, radius of curve is not the only geometric element affecting safety. The crash and field studies showed that the design of highway curves must consider a series of trade-offs among the basic elements of a curve-radius, superelevation, and curve length.

The study also found that either very sharp or very long highway curves tend to produce more crashes. Larger angles (i.e., greater than 45 degree) require either sharp curvature, or a long curve length and should be avoided when possible.

Studies of crashes on highway curves showed single-vehicle run-off-road crashes to be of paramount concern. Roadside treatment countermeasures were found to offer the greatest potential for mitigating the frequency and severity of crashes on rural highway curves. Studies involving a single factor have generally reached the following conclusions:

Lane Width

The crash studies did not conclusively establish a meaningful effect of lane width on crash rates at highway curves. This lack of sensitivity probably resulted because very few roads less than 20 feet wide were observed in the crash study data base.

Shoulder Width

As shoulder width increases, the probability that the highway curve will be a high crash location decreases.

Roadside Character

The crash studies indicate that roadside character (roadside slope, clear zone width, and coverage of fixed-objects) is the most dominant contributor to the probability that a highway curve is a high-crash location.

Pavement Surface

As pavement skid resistance decreases, the probability that a highway curve will be a high-crash location increases.

Stopping Sight Distance

Limited sight distance increases the probability that a curve will be a high crash location. Two special considerations of stopping sight distance are important:

(a) the increased friction demand of a vehicle that is both cornering and braking; and

(b) the loss of the eye height advantage for truck drivers on highway curves when the horizontal sight restriction is either a row of trees, a wall, or vertical rock cut.

Approach Conditions

The crash studies did not indicate a measurable effect of approach conditions (such as approach sight distance, preceding vertical or horizontal alignment, etc.) on the crash experience of highway curves.

MODELING EFFORTS:

Based on these analyses, a crash model, namely Glennon model, was developed and presented in the Transportation Research Board's Special Report 214.

$A = ARs\,(L)(V) + 0.0336\,(D)(V)$ for $L >= Lc$

where,

A = Total number of crashes on the roadway segment.

ARs=Crash rate on comparable straight roadway segments in crashes per million vehicle miles.

L=Length of roadway segment in miles

V=Traffic volume in millions of vehicles

D=Curvature in degrees

Lc=Length of curved component in miles

In the development of this model, cross-tabulations and data analysis supported the following findings:

1). Lane width may have a minor effect on reported crash rates (not in the model).

2).Volumes appear to have a small effect as well.

3). The data showed no consistent and pronounced relationship between crash rate and either curve length or curve central angle.

As noted in Special Report 214, the accuracy of this horizontal curve model "may be diminished for curves sharper than about 15 degrees, the approximate limit recorded in the data base from which the model was calibrated". This model does not consider the following factors and curve design parameters: Curve length, Superelevation and superelevation run-off, Spiral transitions, Cross-slope break, Roadside, Geometric design consistency.

In 1986, Zegeer et al (3) reported the result of their study "Safety Effects of Cross-section Design for Two-lane Roads, Volume I". In this study, they quantified the effects of lane width, shoulder width, and shoulder type on highway crash experience on extended sections of roadways based on an analysis of data for nearly 5,000 miles of two-lane highway from seven states. The following crash prediction model resulted from that study:

$$AO/M/Y = 0.0019\,(ADT)^{0.8824}\,(0.8786)^{W}\,(0.9192)^{PA}\,(0.9316)^{UP}\,(1.2356)^{H}\,(0.8822)^{TER1}\,(1.3221)^{TER2}$$

where:

AO/M/Y = related crashes (i.e., single-vehicle plus head-on plus opposite direction sideswipe plus same direction sideswipe crashes) per mile per year.

ADT = average daily traffic

W = lane width in feet.

PA = average paved shoulder width in feet.

UP = average unpaved shoulder width in feet.

H = roadside hazard rating, a subjective measure with values of 1 to 7 (least to most hazardous), based on a visual assessment.

TER1 = 1 if terrain is flat, otherwise 0.

TER2 = 1 if terrain is mountainous, otherwise 0.

The model is applicable only to:

- two-lane, two-way paved rural highways of state primary and secondary systems.
- lane widths of 8 to 12 feet.
- shoulder widths of 0 to 10 feet.

- ADT's less than 10,000 vpd.

- homogenous roadway sections.

The model does not include the intersection related crashes or those within the horizontal curve that are not expressly stated on the previous page. The model did not explain the variance in crash experience on horizontal curves, as it does not consider the effects of horizontal or vertical alignment or the frequency of horizontal curves, the frequency of sight-restricted vertical crest curves, etc.

In 1991 Zegeer et al (5) formulated a model for predicting crashes on horizontal curves:

$A=[1.552(L)(V)+0.014(D)(V)-0.012(S)(V)](0.978)^{(W-30)}$

where:

A=number of total crashes on the curve in a 5-year period.

L=length of curve in miles (or fraction of a mile)

V=volume of vehicles in million vehicles in a 5-year period passing through the curve (both directions)

D=degree of curve

S=presence of spiral, S=0 if no spiral exists and S=1 if there is a spiral.

W=width of the roadway on the curve in feet.

The purpose of this study was to determine the horizontal curve features which affect safety and operations and to quantify the effects on crashes of various curve-related improvements. The primary data base developed and analyzed consisted of 10,900 horizontal curves in Washington State. Three existing federal data bases on curves were also analyzed. These data bases included the cross-section data base of nearly

5,000 miles of roadway from seven states, a surrogate data base of vehicle operations on 78 curves in New York state, and 3,277 curve roadway segments from four states.

Based on statistical analyses and model development, variables found to have a significant effect on crashes include degree of curve, roadway width, curve length, ADT, presence of a spiral, superelevation, and roadside condition.

In a comprehensive review of design features related to highway safety, McGee et al (6) concluded that the Zegeer and Glennon models were the best models available for predicting crashes on horizontal curves. They reported that:

"The Zegeer model relating crashes to horizontal alignment appears to represent the best available relationship to estimate the number of crashes on individual horizontal curves on two-lane rural roads, although it does have limitations. While the model explicitly considers curve length, degree of curvature, roadway width, and presence of a spiral transition, it does not explicitly consider roadside parameters or the effect of upstream or downstream alignment. The fact that it does not consider roadside or even some surrogate rating for roadside is a major limitation, especially since crash research has shown that roadside design is a determinant of horizontal curve safety.

The model does not consider the effect of vertical alignment or the consistency with respect to the design of all curves within the highway section (e.g., geometric design consistency). The model also does not consider the frequency of horizontal curves greater than three degrees within the section, the frequency of sight-restricted vertical crest curves, or the percent grade. The average operating speeds or design speeds are also not considered explicitly. The model does not consider the influence of access points, driveways or intersections that may be in close proximity to the subject curve."

In 1992, Kach and Benac (7) used the Zegeer and Glennon models and Michigan Trunkline data, and found a poor fit between the predicted and actual crash frequency, as shown in Figures 1, 2 and 3.

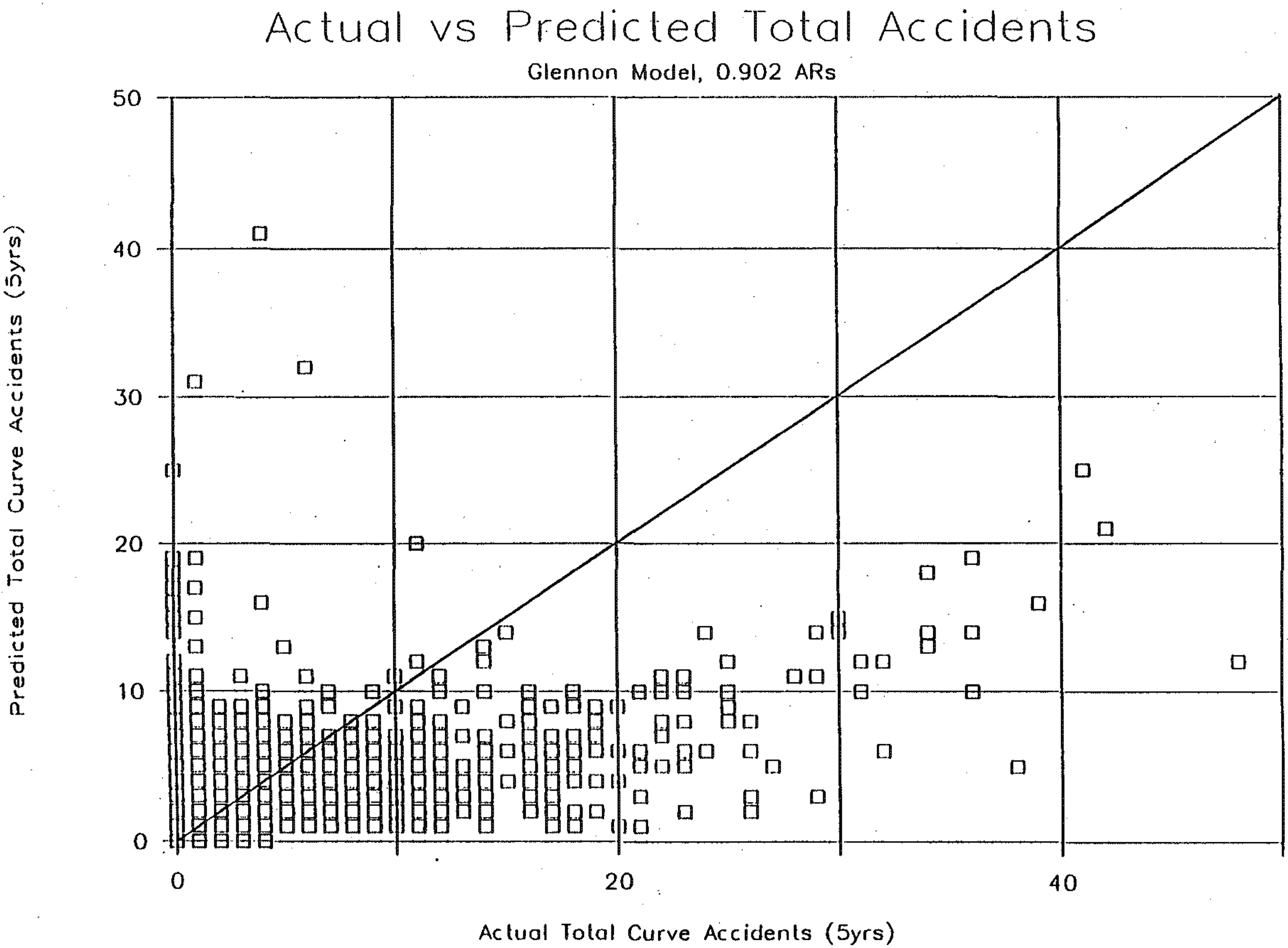
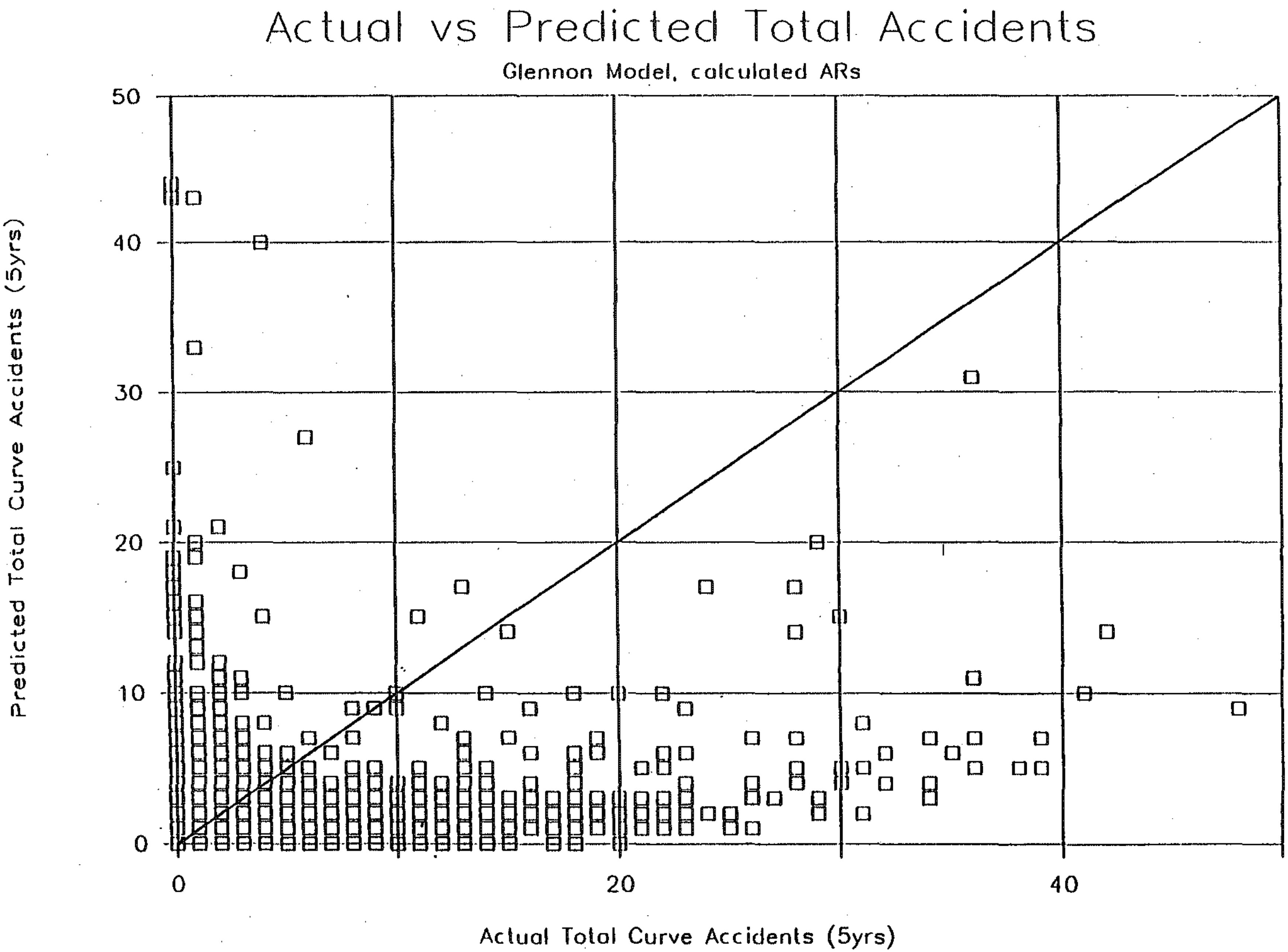
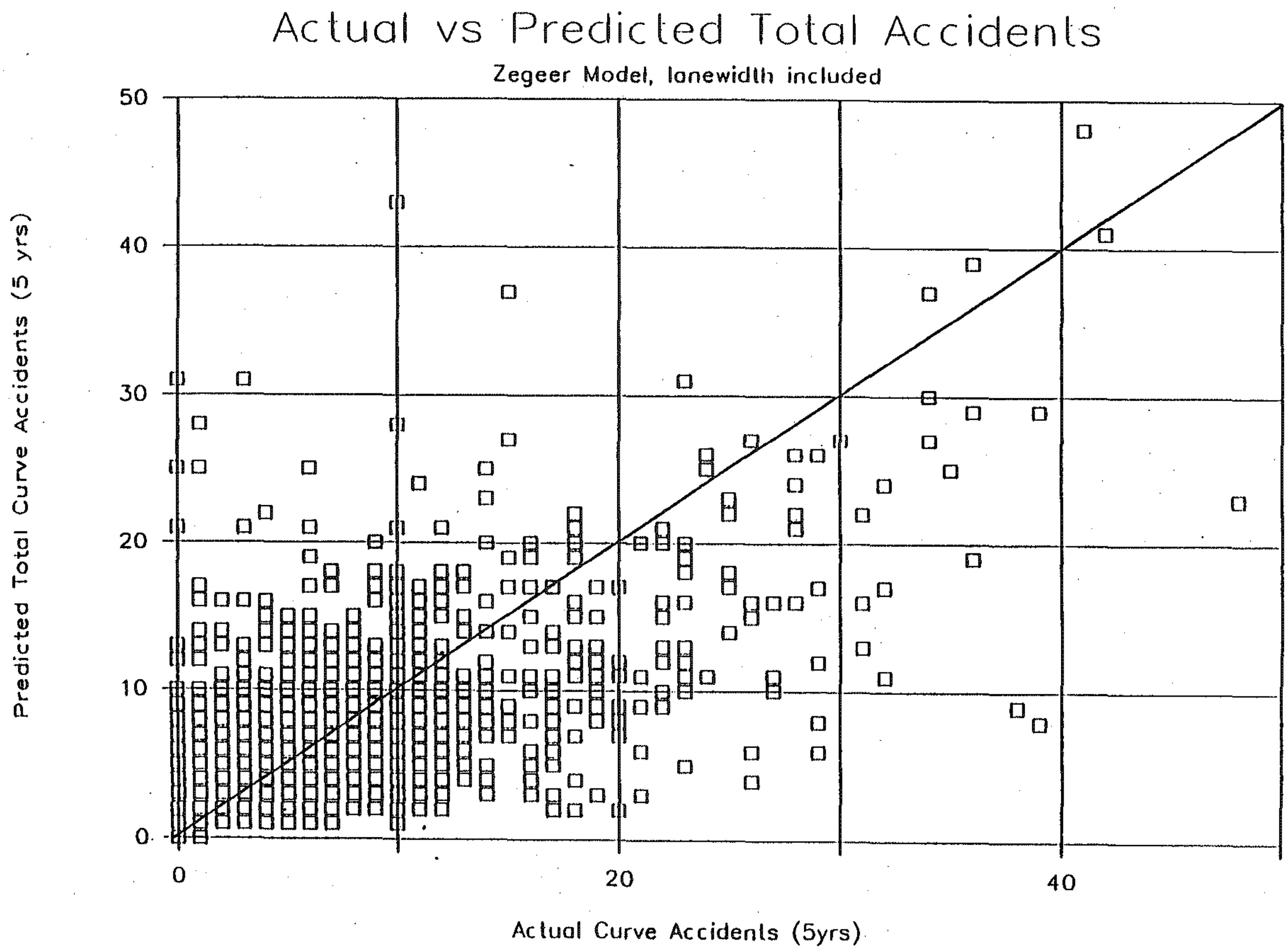

Actual vs Predicted Total Accidents
Glennon Model, 0.902 ARs
Predicted Total Curve Accidents (5yrs)
Actual Total Curve Accidents (5yrs)
0
10
20
30
40
50
0
20
40

Figure 1



Actual vs Predicted Total Accidents
Glennon Model, calculated ARs
Predicted Total Curve Accidents (5yrs)
Actual Total Curve Accidents (5yrs)
0
10
20
30
40
50
0
20
40

Figure 2



Actual vs Predicted Total Accidents
Zegeer Model, lanewidth included
Predicted Total Curve Accidents (5 yrs)
Actual Curve Accidents (5yrs)
0
10
20
30
40
50
0
20
40

Figure 3

After reviewing the models developed by Glennon and Zegeer, they identified the following weaknesses of these models:

1. Total crashes are predicted, instead of "curve related" crash types:

   A. Fixed-Object

   B. Overturn

   C. Head-On

   D. Sideswipe-Opposite

2. Models do not recognize an "influence zone" for curves.

3. Models do not adequately address the actual variability in the crash experience for all the curves with a given length and degree of curvature.

In 1995, Fink and Krammes (8) reported on a study of the effect of tangent length and sight distance on crashes at horizontal curves. This study included a review of previous models.

Their report concluded that most models for evaluating operating-speed consistency on two-lane rural highways estimate operating-speed profiles based upon tangent length and degree of horizontal curvature. Some models also consider the effect of sight distance to horizontal curves. To add insight on the effects of these variables on safety and operations at horizontal curves, a base relationship between crash rates at horizontal curves and degree of curvature was established, and the effects of approach tangent length and approach sight distance on this relationship were examined.

The results confirm that degree of curvature is a good predictor of crash rates on horizontal curves. Although the effects of approach tangent length and sight distance were not as clear, the results suggest that the adverse safety effects of long approach tangent length and short approach sight distance become more pronounced on sharp curves.

Four other studies considered tangent length among a set of candidate predictors of crash rates at horizontal curves (10-13). Their findings with respect to tangent length were mixed. Datta et al. (10) found tangent length to be a significant predictor of outside-lane crash rates for one subset of 25 curve sites in Michigan. Terhune and Parker (11) evaluated tangent length (among other variables) using data bases of 78 curves in New York, 40 curves in Ohio, and 41 curves in Alabama, and concluded that tangent length was not significant. Matthews and Barnes (12) studied 4,666 curves on the rural two-lane portion of State highways in New Zealand.

They found a significant relationship that involved tangent length in combination with other variables and concluded that crash risk was particularly high on short radius curves at the end of long tangents, on steep down grades, and on relatively straight sections of roads.

Zegeer et al. (13) evaluated the significance of the minimum and maximum distance to the adjacent curve; although neither variable was significant, they observed, "there appears to be evidence that tangents above a certain length may result in some increase in crashes on the curve ahead."

Glennon et al. (14) concluded that approach sight distance was not a significant variable in a discriminate analysis of curve sites with high and low crash rates. Fambro et al. (15) concluded that available stopping sight distance is not a good indicator of crashes, with the exception that "when there are intersections within limited sight distance portions of crest vertical curves, there is a marked increase in crashes.".

The study by Fink and Krammes (8) developed two models:

1) A regression model for predicting mean crashes per million vehicle kilometers versus mean degree of curvature:

mean crash rate = 0.05 + 0.23 mean degree of curvature

The model has an r^2 value of 0.94. The r^2 is much higher than typically observed in crash analyses, because the unit of observation is a grouping of curve sites into nine degree-of-curvature categories which eliminates much of the variability among individual sites.

2) A regression model for predicting the crash rate based on the approach tangent length. Three categories were defined representing the shortest 25 percent (<=107 m [350 ft]), middle 50 percent (107 m[350 ft] to 427 m [(1400 ft]), and longest 25 percent (>427 m [1400 ft]) of tangent lengths in the database. The regression models were as follows:

* Shortest 25%:

mean crash rate = 0.35 + 0.16 mean degree of curvature

* Middle 50%:

mean crash rate = -0.30 +0.32 mean degree of curvature

* Longest 25%:

mean crash rate = 0.52 + 0.20 mean degree of curvature

The results indicate that the slope and intercept for the middle 50 percent of tangent lengths are significantly different from the slope and intercept for the shortest and the longest 25 percent. (See Figure 4)

These models, like those of Zegeer and Glennon, fail to explain the variation in crash rate experienced at different curves with the same degree of curvature or the same approach tangent length.

While all of the models found in the literature may have some value when considering design alternatives, none are suitable for identifying hazardous curves.

They also provided no assistance in determining countermeasures once a location is identified as being hazardous.

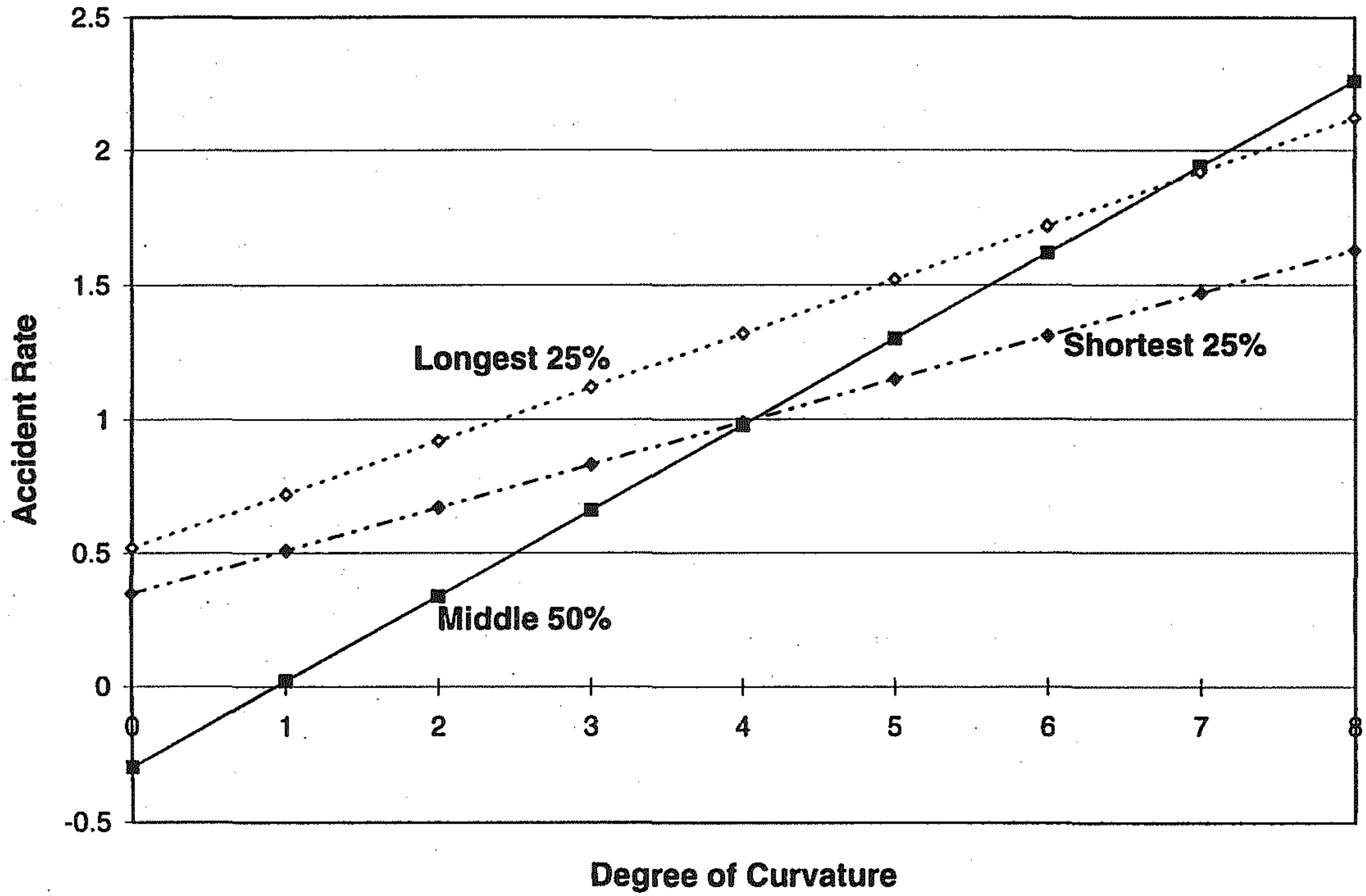


Figure 4 **Accident rate versus degree of curvature**

METHODOLOGY:

To accomplish the objective of this study, a multi-step approach was utilized. Step one was to acquire geometric data for all the rural, two-way, two-lane trunkline highways in Michigan from the Michigan Department of Transportation (MDOT). Based on the selection criteria, (Table 1) the candidate curves were selected and the control section (reference system used by MDOT for trunklines) and the mile points of the beginning and ending of the curves were noted.

The next step consisted of obtaining additional data from the Photo log. In addition to data acquisition, data verification was also performed and locations which, based on this observation, did not meet the selection criteria were removed from the database.

While this step was in progress, field data collection was being performed to obtain the curve superelevation and pavement friction. Field data collection further rendered some of the curves invalid. After this step 220 curves were left for the final analysis. For each of the 220 curves, all the crashes corresponding to the mile points from 306 meters (0.19 mile) before the start of the curve to 306 meters (0.19 mile) after the end of the curve were extracted from the MDOT crash files. This procedure was performed six times for the six year period of 1989 to 1994, yielding 3107 total crashes (Table 6).

The crash report forms for all these crashes were obtained and processed to locate the individual crash as being on the curve or on the tangent. After this step, various analyses were performed, including comparison of the actual curve crashes and those predicted by the models.

THE DATA:

Data for the project consists of the four following sets:

1) Geometric data provided by the MDOT

2) Six years of crash data for the years 1989 through 1994

3) Data obtained from the photo log for all 220 segments

4) Field data for 81 segments (see page 95)

The Geometric data consisted of 44 variables such as Control Section, Beginning Mile point, Ending Mile Point, Average Lane Width, Total Shoulder Width (Right and Left), etc. The variables selected from this file for use in this study are shown in Table 3.

The crash data are from the Michigan State Police "State of Michigan crash Master File". This file contains information on up to three vehicles involved in a crash, but the data for the second and third vehicles were not used in the study. The original source of the data is the "State of Michigan Traffic Crash Report" (Form UD-10). The data consisted of 120 variables such as District, Control Section, Mile point of Crash, Highway area Type, Highway Area Code, etc. The data were for the crashes for both traffic directions combined. The variables selected from this file for use in this study are shown in Table 4.

The photo log data were used for dichotomous variables such as the presence of traffic signs (Arrow, Chevron, etc.) and other variables such as the mile point at which the curve was first observed, etc. The data also included a subjective measure of the roadside clearance/hazard, on a scale of one to seven. One being "Clear" (least hazardous) and seven being "Not Clear" (most hazardous). The data acquisition was performed twice, once for each direction of the traffic flow.

The field data collection was performed to obtain only two variables; a measure of the superelevation of the road, and a measure of the skid resistance of the pavement surface. The superelevation was obtained by use of an ordinary 48 inch long level. The difficulty with superelevation is the fact that unlike some other variables, an average value will not substitute for the lowest value and the highest value. If there is an optimal value, any deviation from it, positive or negative, could result in lower safety. However, since there was no procedure available to record continuous values of superelevation, representative locations on the curve were selected and the average value for each lane was coded. Occasionally the superelevations were in the opposite direction, i.e., banking towards the outside of the curve. In these cases the superelevation is coded with a negative sign.

The friction factor was obtained and calculated by dragging a piece of tire filled with concrete to weigh 22.7 kilograms (50 lbs) (16). The horizontal force required to pull it over the pavement (divided by its weight), would have been the friction factor, had the tire been smooth. However, the reading corresponded to a value higher than the actual friction factor because the treads of the tire and the gravel particles on the road would "engage" and to some extent act like teeth gears. Occasionally the required horizontal force exceeded 22.7 kilograms, yielding friction factors higher than one. Since this variable was for comparison across the curves and not for the absolute values, the resulting values were used for the study. However, to avoid confusion it was referred to as the Drag Factor rather than the Friction Factor.

TABLE 1

CURVE SELECTION CRITERIA:

a) Rural two-lane, two-way.

b) No taper, no extra lanes.

c) No curb, no parking.

d) No median, and preferably no intersections.

e) At least 306 meters (0.19 mile, about 1000 feet) of tangent at each end of each curve.

f) Preferably at least 611 meters (0.38 mile) of tangent between the two curves.

g) Degree of curvature greater than one.

This geometric selection criteria yielded a total of 285 roadway segments, each consisting of a curve and two tangents. Based on the photo log observation, 50 of the selected roadway segments did not fit the specified criteria and were eliminated from the study. Fifteen more were eliminated from the list based on the field observation. Examples of such cases are listed in Table 2. The final data set consisted of 220 valid roadway segments.

Table 2

Examples of the Disqualified Roadway segments:

(based on the photo log/field observations)

Control Section	Listed BMP*	Listed EMP*	Length km	Actual Length km	Comments
23111	3670	3710	0.06	0.21	Intersection Corner.
32092	60	190	0.21	-	Intersection widening (M-52/M-36)
38073	9810	9920	0.18	0.26	Curve not found.
38073	14350	14500	0.24	-	Curve not found.
46011	5770	5900	0.21	0.10	Three Lanes (intersection with left turn lane)
46012	110	300	0.31	-	Three Lanes (intersection with left turn lane)
46051	380	490	0.18	0.27	Not found. Two curves near listed location.
46074	20	130	0.14	0.18	Intersection (with median and right turn lanes).

* (coded in 0.001 mile with implied decimal point)

Table 3

Geometric data variables coded for the study and their names:

(Where specific cases were selected the condition is listed under "SELECTED IF:" and no variable name is listed for them since the item is no longer a variable.)

VARIABLE DESCRIPTION	SELECTED IF:	VARIABLE NAME
District		DNO
Control Section		CS
Beginning Mile Point of Roadway segment (MALI)		BMP
Ending Mile Point of Roadway segment (MALI)		EMP
Roadway Area Type Flag	Midblock	
Number of Basic Lanes	Two	
Roadway Type	Two-Way	
Miscellaneous Extra Lanes (Right)	None	
Miscellaneous Extra Lanes (Left)	None	
On-Street Parking (Right)	No	
On-Street Parking (Left)	No	
Average Lane Width		ALW
Total Shoulder Width (Right)		TSWR
Shoulder or Curb Type (Right)	No Curb	

Table 3 (continued)

Paved Shoulder Width (Right)		PSWR
Total Shoulder Width (Left)		TSWL
Shoulder or Curb Type (Left)	No Curb	
Paved Shoulder Width (Left)		PSWL
No Passing Zone Code		NPZC
Roadside Development Code	Rural	
Posted Speed Limit		PSL
Degree of Curvature, Number of Degrees		HCD
Degree of Curvature, Number of Minutes		HCM
Roadway segment File Record Number		SFRN
Intersection File Record Number		IFRN
Average Daily Traffic (Divided by 10)		ADT

Using BMP, EMP, HCD and HCM, four more variables were calculated as follows:

Degree of Curvature in decimal degrees	HCDD
Curve Length in feet	HCLFT
Curve Radius in feet	HCRFT
Central Angle in decimal degrees	CANG

Additionally four more variables related to the design speed were calculated as described on pages 44 and 45.

Table 4

Crash data used in the study:

District	Driver 1 Violation
Control Section	Contrib. Circumst., Vehicle 1
Crash Mile Point	Visual Obstruction, Vehicle 1
Highway Area Type	Direction of Travel, Vehicle 1
Highway Area Code	Alcohol/Drug use, Vehicle 1
Hour of Occurrence	Object Hit, Vehicle 1
Route Class	Situation, Vehicle 1
Weather Condition	Vehicle Size, Vehicle 1
Lighting	Impact Code, Vehicle 1
Road Surface Condition	Vehicle Condition, Vehicle 1
"A" Injuries	Trailer, Vehicle 1
"B" Injuries	Road Type, Vehicle 1
"C" Injuries	Number of Lanes
Road Alignment	Average Daily Traffic
Traffic Control	Number of Persons Killed
Crash Type	Number of Persons Injured
Distance From Crossroads	Number of Occupants
Direction From Crossroads	Crash Location
Intersecting Street name	Crash Route Number

TABLE 4 (Continued)

Number of Persons Uninjured	Original Prime Street Name
Vehicle 1 Type	Operator Number, Vehicle 1
Vehicle 1 Make	Year Of Crash
Age of Driver 1	Film Reel Number
Residence of Driver 1	Film Frame Number
Sex of Driver 1	PR Number
Degree of Injury to Driver 1	PR Mile Point
Driver 1 Intent	

PHOTO LOG and FIELD DATA:

For these variables two values were obtained, one for each direction of traffic, denoted with prefix P for plus direction and M for minus direction. The plus direction is the direction of increasing mileage in the control section.

The "Mile Point When Curve Observed" was converted to the "Distance from curve when it was observed" (coded in 0.001 mile with implied decimal point). The variables obtained from these two sources are listed in Table 5.

Table 5

Variables obtained from the photo log and field observations.

VARIABLE DESCRIPTION	VARIABLE NAME
Curve Sign	CURVES
Turn Sign	TURNS
Advisory Speed Sign	MPHS
Guard Rail	GRAIL
Chevron	CHEVRON
Arrow Sign	ARROW
Delineator	DLNTR
Edge Line	EDGLN
Mile point when Curve Observed	OBSDSTW
Roadside Clearance/hazard	CLRNCW
Superelevation	SPRELVN
Drag Factor	DRGFCTR

VARIABLE MODIFICATION:

Since the crash data were for both directions, variables with two values, one for each traffic direction, were reduced to a single value. These included all the photo log data, some geometric data and the two field data variables.

For the following variables, if for either direction of traffic the variable had a value of YES, the variable was coded as 1. If neither direction had a value of YES, it was coded as 0 (zero). Variables in this category consisted of: Curve Sign, Turn Sign, Guard Rail, Chevron, Arrow Sign and Delineator.

The variable "Mile Point Where Curve Observed", was converted to a distance and the lower of the two was used. For the subjective value of the "Roadside Clearance", the higher of the two values was used.

From the geometric data, Total Shoulder Width Right and Left were combined into one value, the sum of the two. Similarly the Paved Shoulder Width Right and Left was replaced by the sum of the two values. The Shoulder or Curb Type Right and Shoulder or Curb Type Left, each with a value of 1 or 2 were collapsed into one value. If both values were the same that value was used. If one value was 1 and the other 2, a value of 2 was used.

The drag factor and superelevation also had two values, one for each side of the road. For the drag factor the lower of the two was used. For the superelevation, the lower of the two was used for one analysis, and then the analysis was repeated using the higher value.

CRASH TYPES:

For the analyses used in this project, several types of crashes were eliminated from the crash data. Only the "Curve Related" crashes consisting of the following types of crashes were considered:

CODE	DISCRIPTION
000	Miscellaneous 1 Vehicle
010	Overturn
060	Fixed Object
070	Other Object
141	Head-on
543	Side-Swipe Opposite

Selection of the "Related" crashes yielded 994 crashes corresponding to the 178 roadway segments which had "Related" crashes. Not all selected roadway segments had crashes in both the tangent and curve portion of the roadway segment.

CRASH CASES:

In addition to analyzing all crashes, crashes occurring under different road surface conditions, weather conditions and lighting conditions were also analyzed.

A sub-set of curves consisting of only those with the field data (superelevation and drag factor) were analyzed separately. Similar analyses for the sub-set of crash cases based on weather, surface or lighting were not performed due to the fact that the two field variables were not found to be significant in predicting curve crashes.

DATA CATEGORIES:

The assumption was that the non-measurable, non-quantifiable environmental and traffic conditions along the entire length of each curve can be considered to be the same as that of the average of the tangents at each end. The basic unit of tangent length at each end of the curves was 306 meters (0.19 mile). As such, to compare the curve crashes with the tangent crashes, 611 meters (0.38 mile) was used as a unit length and the curve crash rate was adjusted for the length of 611 meters (0.38 mile). The resulting variables were called Cper380 for curve crashes and Tper380 for tangent crashes. Another variable, C-T was defined to represent the difference between curve and tangent crashes. This variable has a value equal to: Cper380-Tper380.

CRASH LOCATION MILE POINTS:

The location of each crash along its control section is indicated by a mile point. Based on the mile point of the crash location compared with the mile points of the two ends of a curve, one could presumably determine if the crash was on the curve or tangent. However, it was evident that locating the crash in the field was not very accurate. A plot of crashes showed that the crashes tend to accumulate at tenths or quarters of a mile from the nearest intersection.

To remedy this problem the UD-10s for all crashes were manually checked. If the crash was drawn on a curve, it was assigned to the curve, even if based on the mile point it would fall on the tangent. The UD-10 forms also provide a check box for the road alignment and if the box for curve was checked, the crash was assigned to the curve. The reason being that it was unlikely that an investigator would draw a tangent section of a road showing curve, however they may draw the curve section as a tangent but check the curve box and use the code for curve.

SPECIAL DATA CONSIDERATIONS:

Even though typically each roadway segment consists of two tangents of 306 meters each, and the curve itself, there were exceptions. In 14 cases the control section number changed within the 306 meters of tangent section of the roadway segment, of which only 10 contained "Related" crashes. In these cases the 306 meters of tangents existed for both ends of the curve, however, the mileage of tangents within the same control section were less than 306 meters. Pro-rated values were used to determine the tangent crashes for 612 meters (Tper380) of these 10 cases. There were no such cases of different control section numbers within a curve, among the 220 curves.

In another 8 cases even though there were 306 meters of tangents at each end of the curves, the distance between the end of one curve and start of another was less than 612 meters. In other words there was an overlap between the two tangents. In only two cases were there crashes in the overlap section of the two tangents, of which only one case contained "Related" crashes. The crashes corresponding to this overlapping section of tangents, (3 crashes), where appropriate, were counted twice, once for one tangent and again for the other.

Table 6

FOLLOWING IS THE NUMBER OF ALL CRASHES:

1994	519
1993	491
1992	503
1991	532
1990	503
1989	559
TOTAL	**3,107**

Out of the 3107 total crashes, 991 were in curves and 2116 in the tangents. The total number of "Related" crashes were 994 of which 463 were in the curves and 531 in the tangents.

NOTE: 13 of the 220 roadway segments did not have any crashes in the curve or the two tangent sections. For the "Related" crashes only 178 roadway segments had crashes in either curve or tangent sections.

DATA PRESENTATION:

The crash data described in the preceding pages is presented in graphical form in Figures 5 through 10.

Figure 5 shows the Cper380 for the 178 roadway segments which had "related" crashes in their tangent sections or their curved section. The Cper380 values are sorted in ascending order including the roadway segments which did not have any crashes in their curved section.

Figure 6 shows the Tper380 for the same 178 roadway segments, some with no crashes in their tangent sections. Similarly, the Tper380 values are sorted in ascending order.

Figure 7 is the Tper380 values when sorted by ascending values of Cper380.

Figure 8 is the superimposed graph of Figure 5 and Figure 7.

Similarly, Figure 9 shows the values of C-T, when sorted in ascending order and Figure 10 is the C-T values sorted by ascending values of Cper380.

From the Figures 7, 8 and 10 it is clear that the crash rate on the tangent section approaching the curve is not a reliable predictor of the curve crash rate. This is evidenced by the fact that the values of Tper380 and C-T do not display a consistent pattern when compared with the sorted values of Cper380.

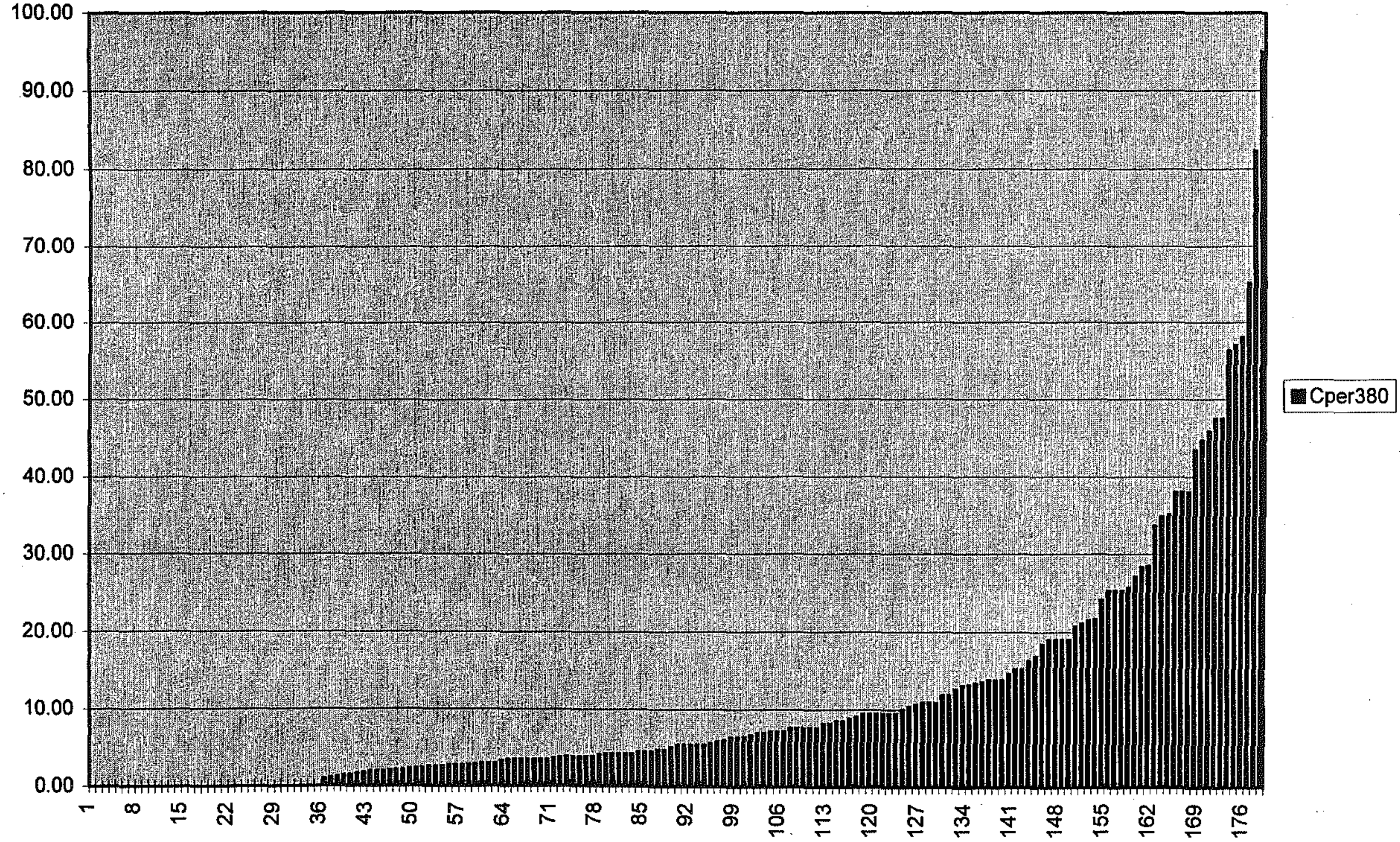


Figure 5 **Curve crash rate (Cper380), arranged in ascending order**

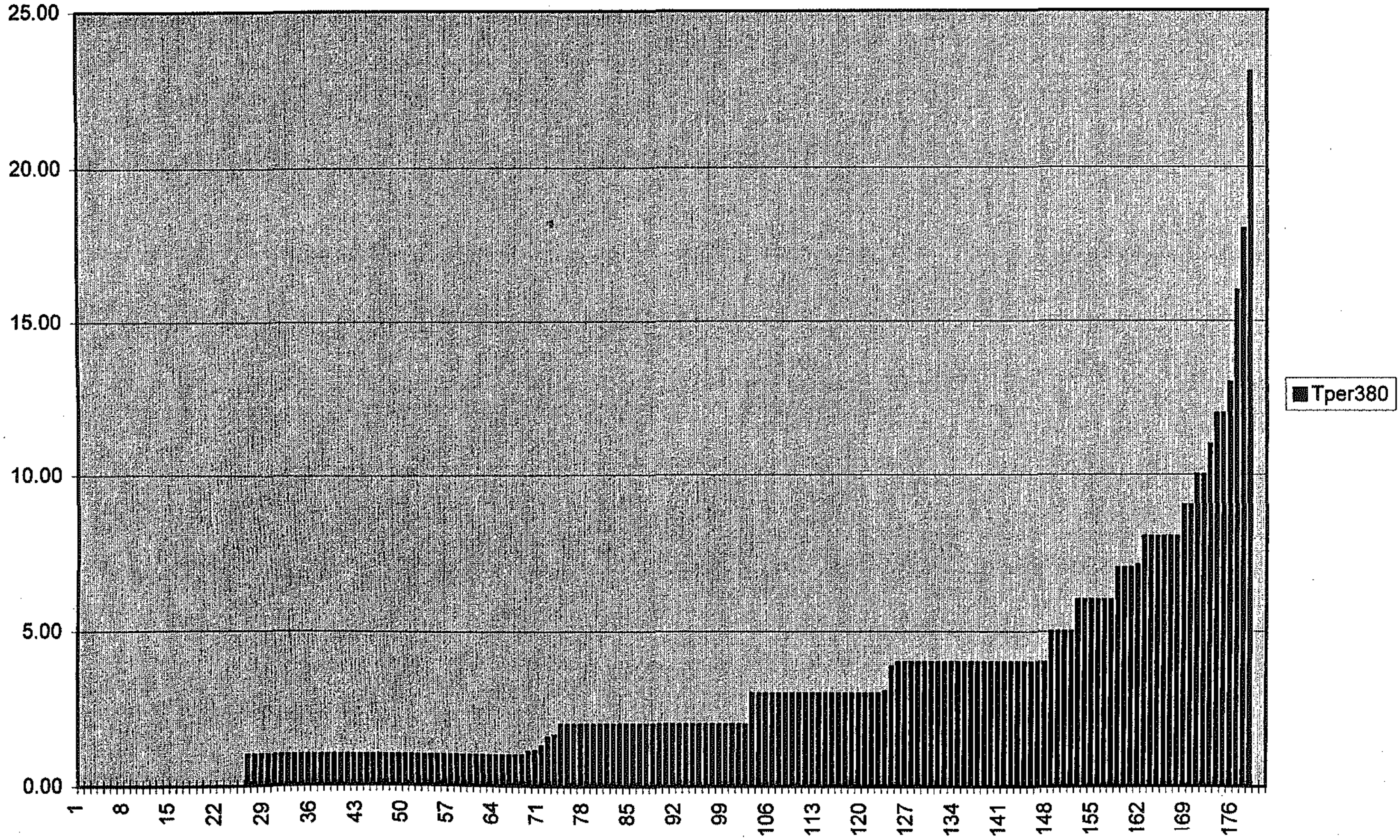


Figure 6 Tangent crash rate (Tper380), arranged in ascending order

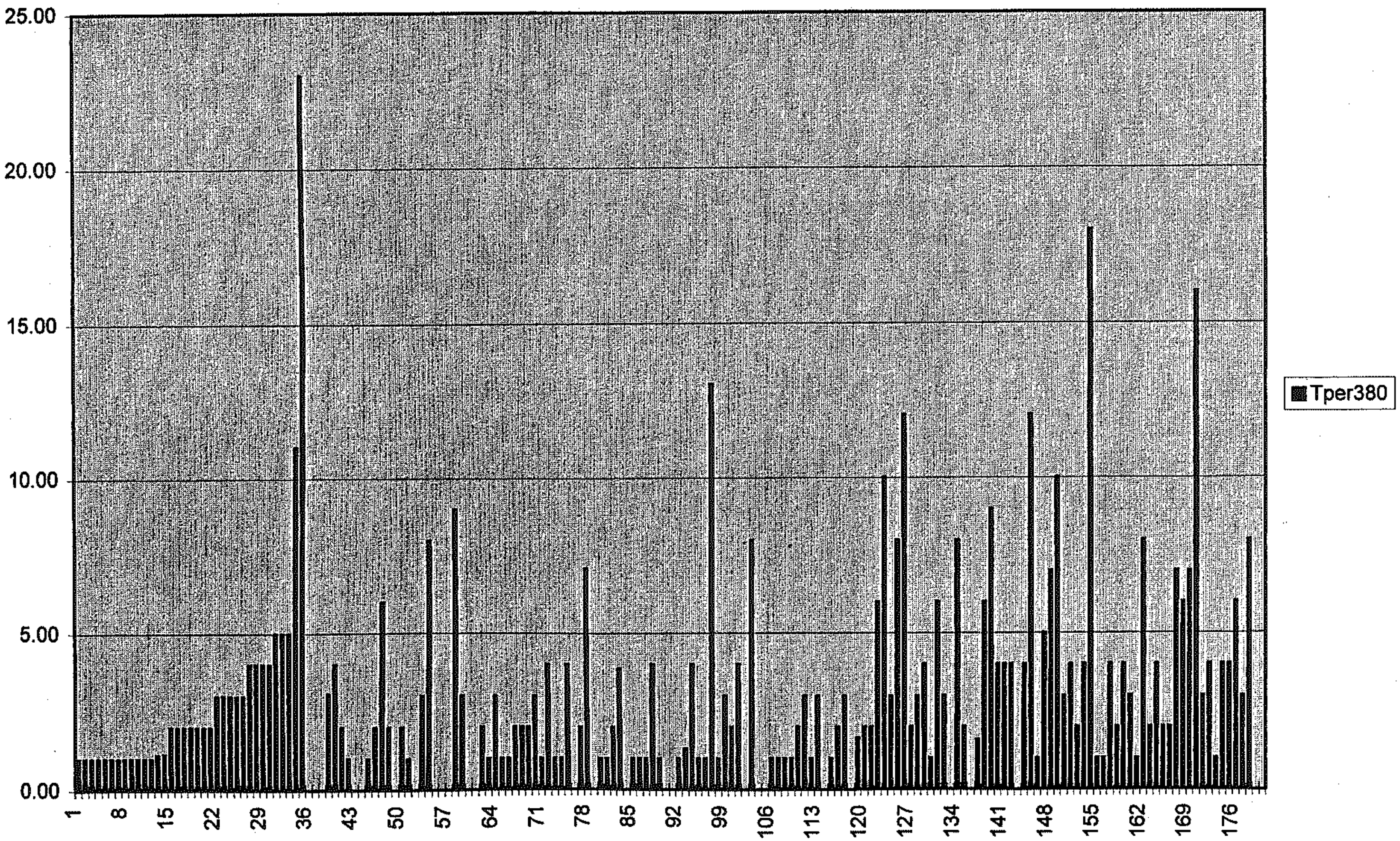


Figure 7 Tangent crash rate (Tper380), arranged in ascending order of curve crash rate (Cper380)

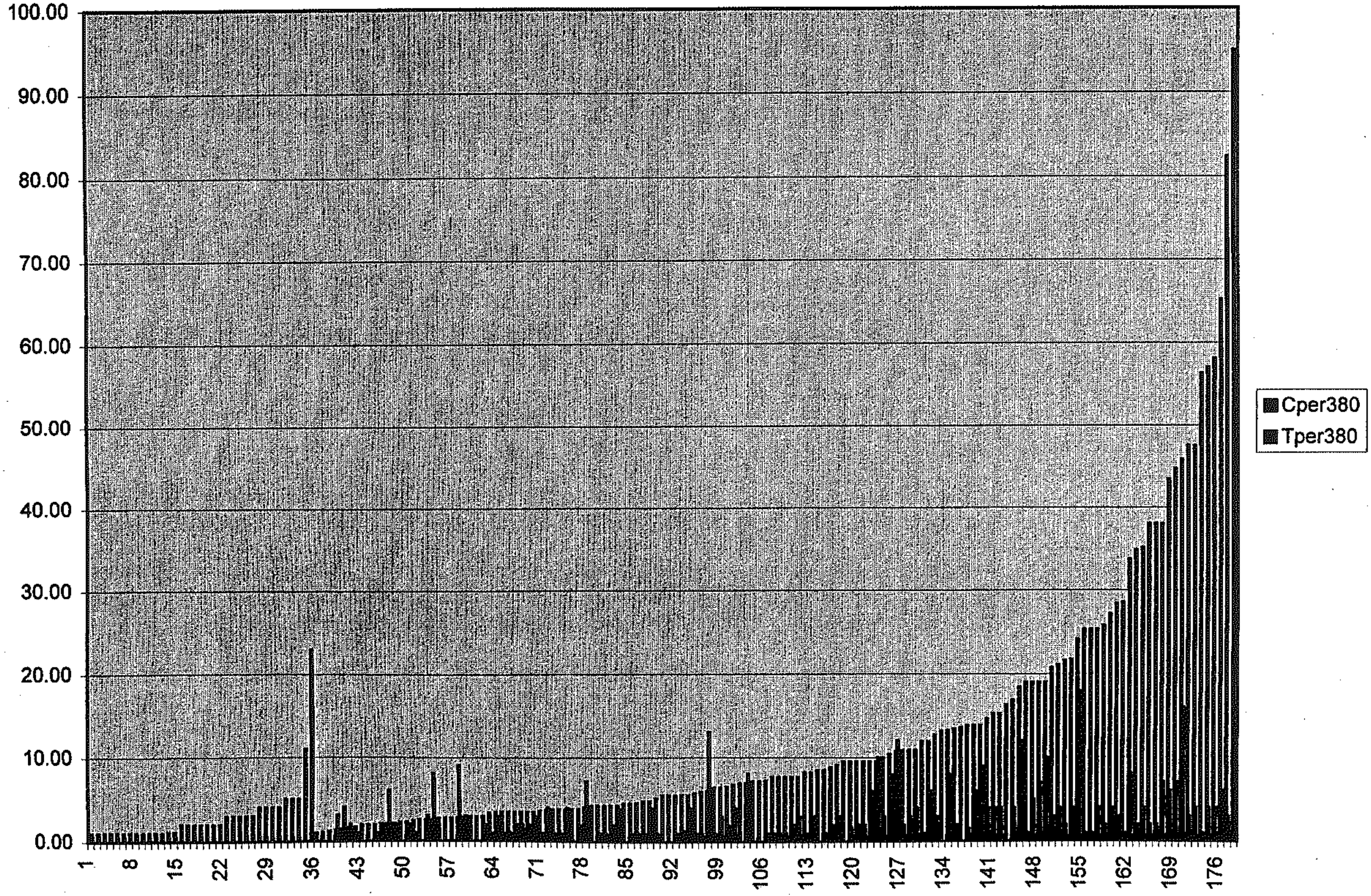


Figure 8 Curve crash rate (Cper380), and tangent crash rate (Tper380), arranged in ascending order of Cper380

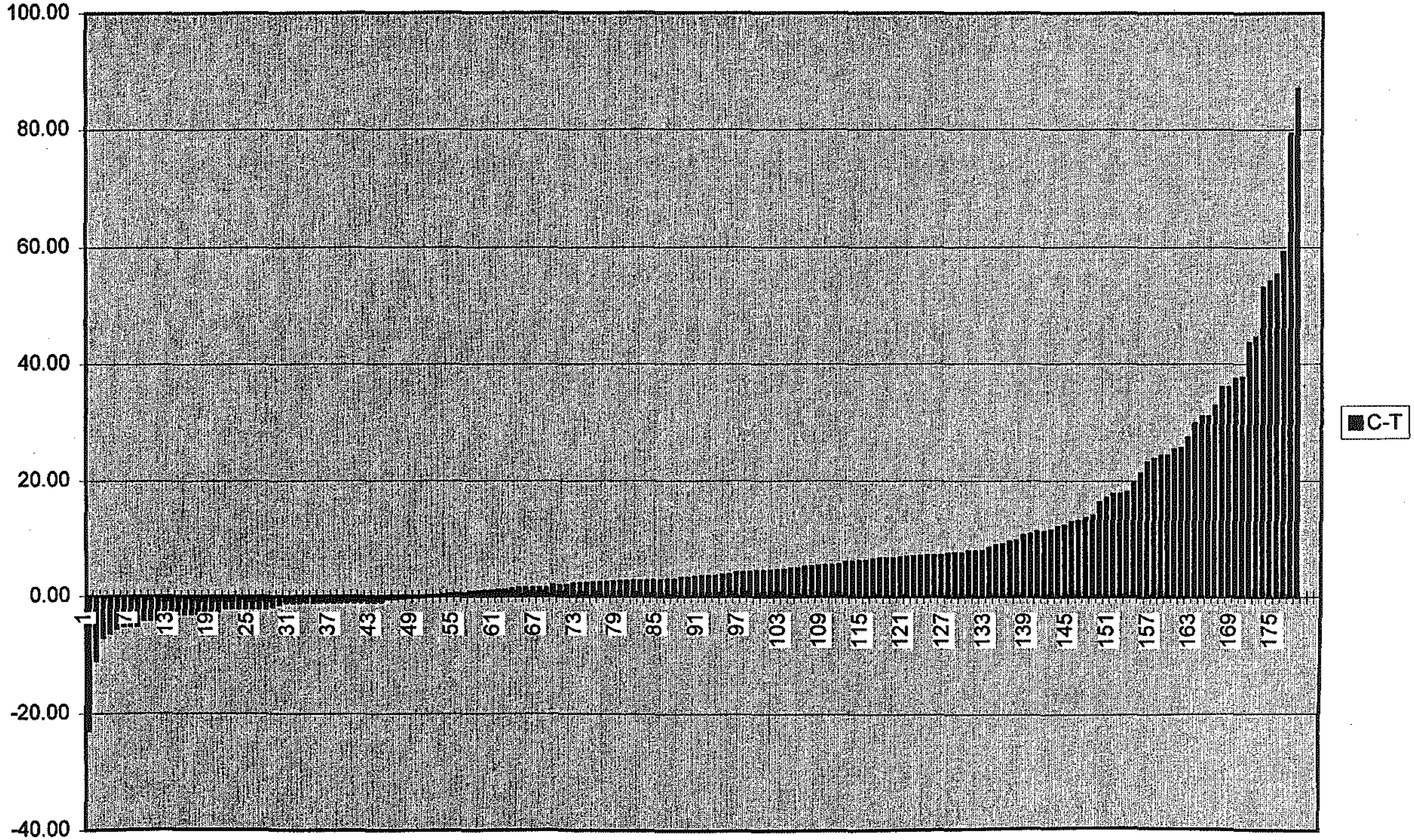


Figure 9 **Curve crash rate minus tangent crash rate (C-T), arranged in ascending order**

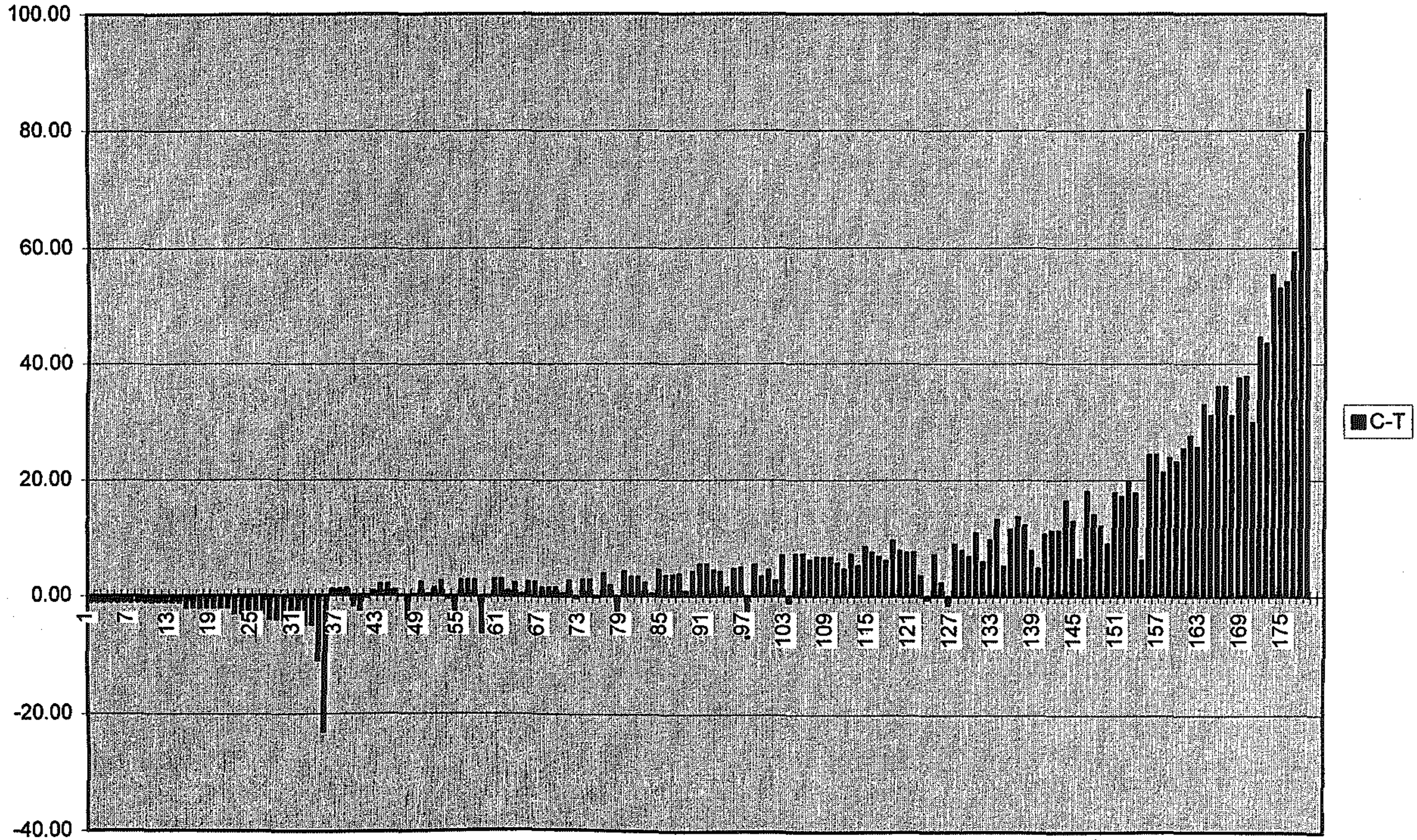


Figure 10 Curve crash rate minus tangent crash rate (C-T), arranged in ascending order of curve crash rate (Cper380)

DATA ANALYSIS:

As a first step in the analysis, two sets of simple regressions, one for the curve crashes (Cper380), and the other for the difference between the curve crashes and tangent crashes C-T, versus the independent variables ADT, Tper380, HCLFT, HCRFT, CLRNCW and OBSDSTW were constructed. The results are shown in Figures 11 through 26. The scatter plots, the regression lines, and the coefficients of regression all indicate that simple regression models are poor predictors of crashes.

For the variable HCLFT (curve length), it appeared that there might be a nonlinear relationship. However the quadratic and cubic regression lines showed little improvement over the linear model, as shown in Figures 23 and 24.

Many different multiple regressions models were analyzed but with unsatisfactory results. Table 7 shows the results of one such model. In this model, the variables HCRFT, Tper380, HCLFT and MPHS best explain the "related" curve crashes. These linear multiple regression models also produced low coefficients of regression, which is consistent with previous research results.

Further more, four new variables were defined and computed . These variables were calculated based on the "Design Speed" and field measurements of the superelevation. The design speeds were calculated from the equation: $R=V^2 / 15(e+f)$ where R is the curve radius in feet, V is the design speed in MPH, e is the superelevation and f is the wet friction factor for which a value of 0.19 was substituted. Two sets of design speeds were computed. The one for the lower value of the superelevation of the two sides of the road was named “DsgnSpdL” and the one for the higher value was named “DsgnSpdH”.

The difference between the design speed and the advisory speed was calculated and named “DiffSpdL” and “DiffSpdH” corresponding to the lower and higher values of the superelevation as described before. Where an advisory speed was not posted, 55 MPH was used as the posted speed limit.

The linear regression models for the Cper380 values and these four variables were analyzed and found to indicate weak correlation.

Figures 25 and 26 show two such regression plots for Cper380 versus DsgnSpdL and DiffSpdL

The conclusion from these analyses was that neither simple linear regression nor multiple linear regression are powerful enough tools to depict the large variations in the curve crash rate.

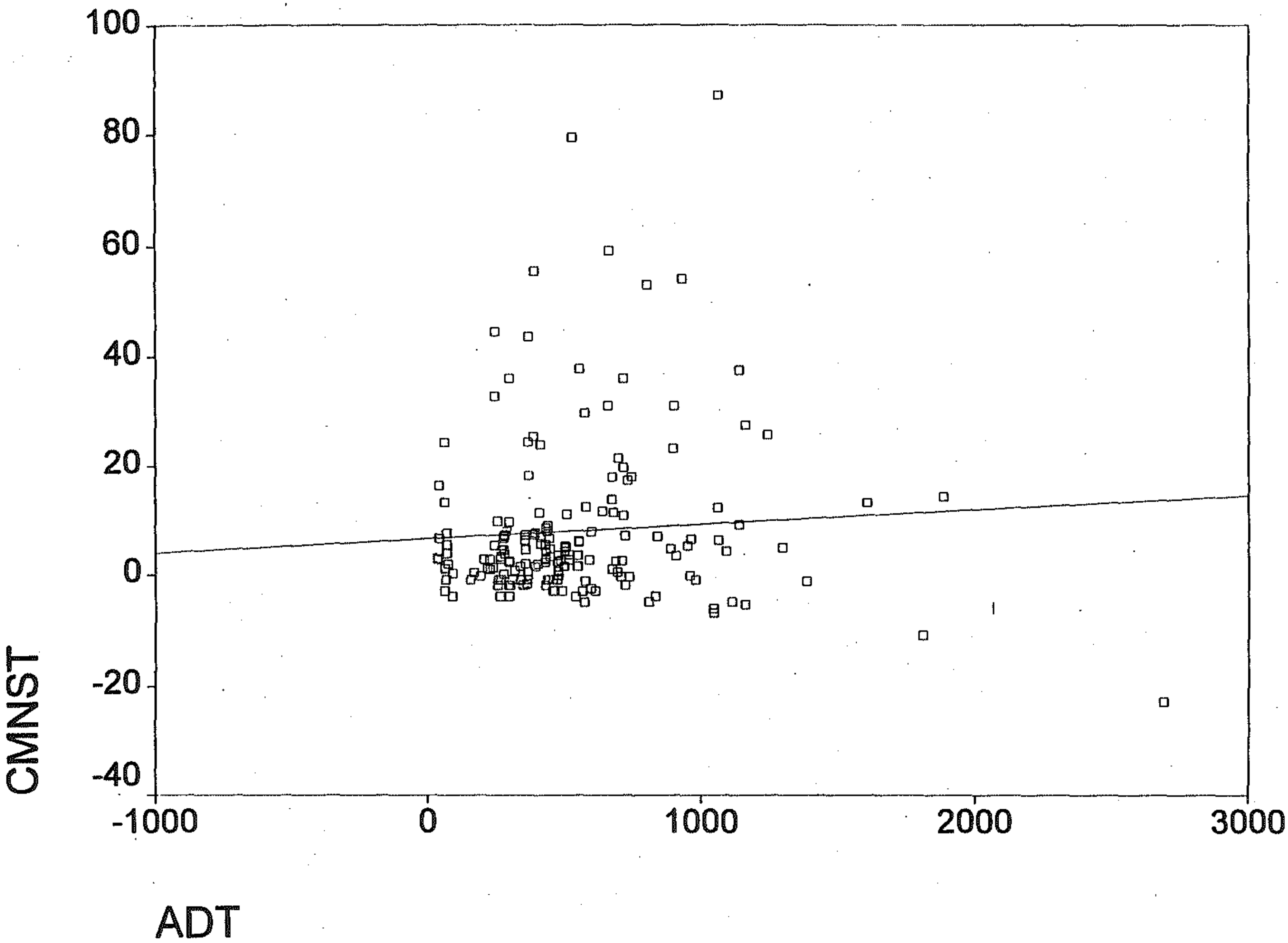


Figure 11 Curve crash rate minus tangent crash rate (C-T), regression line for various values of average daily traffic (ADT)

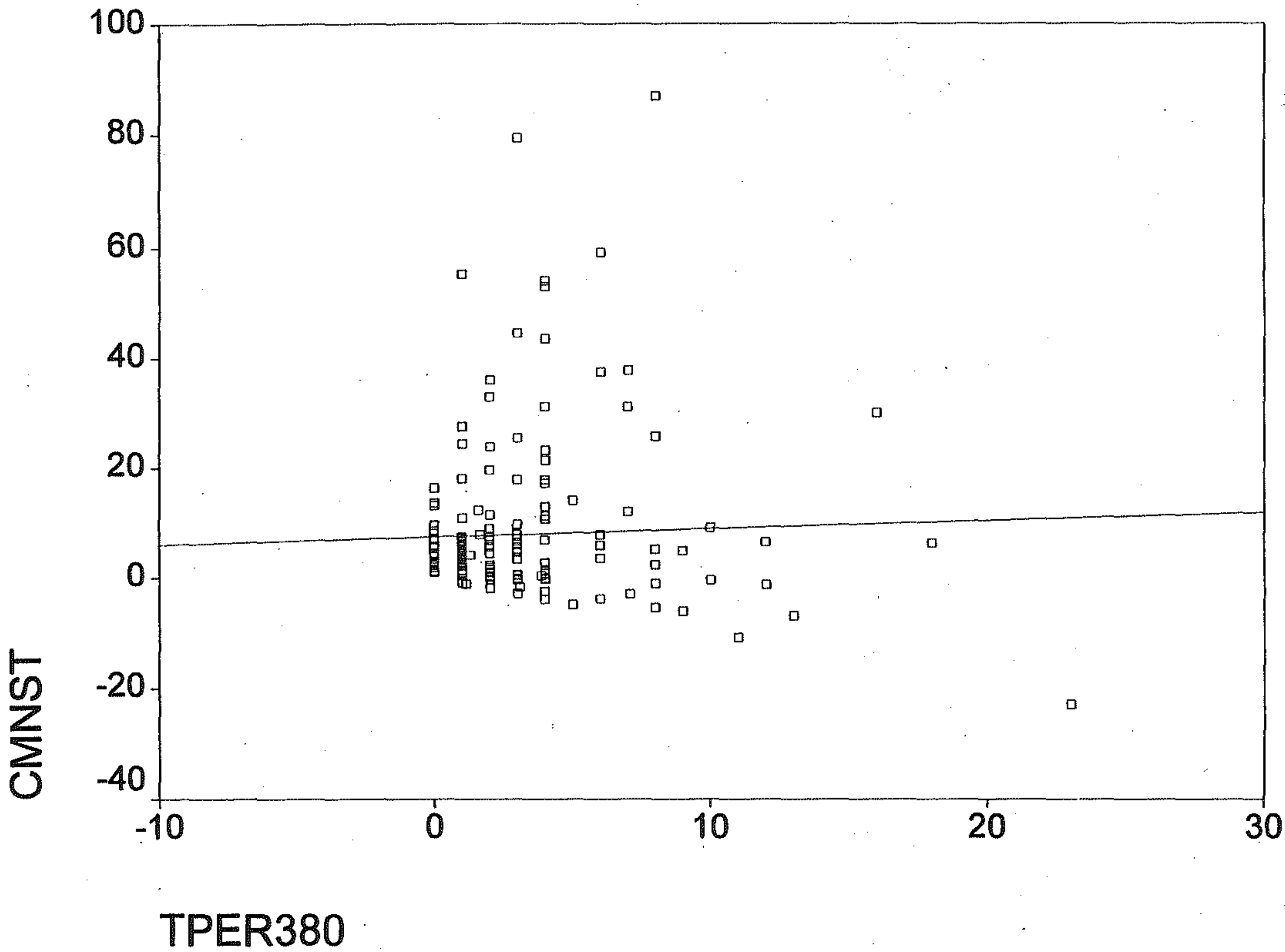


Figure 12 **Curve crash rate minus tangent crash rate (C-T), regression line for various values of tangent crash rate (Tper380)**

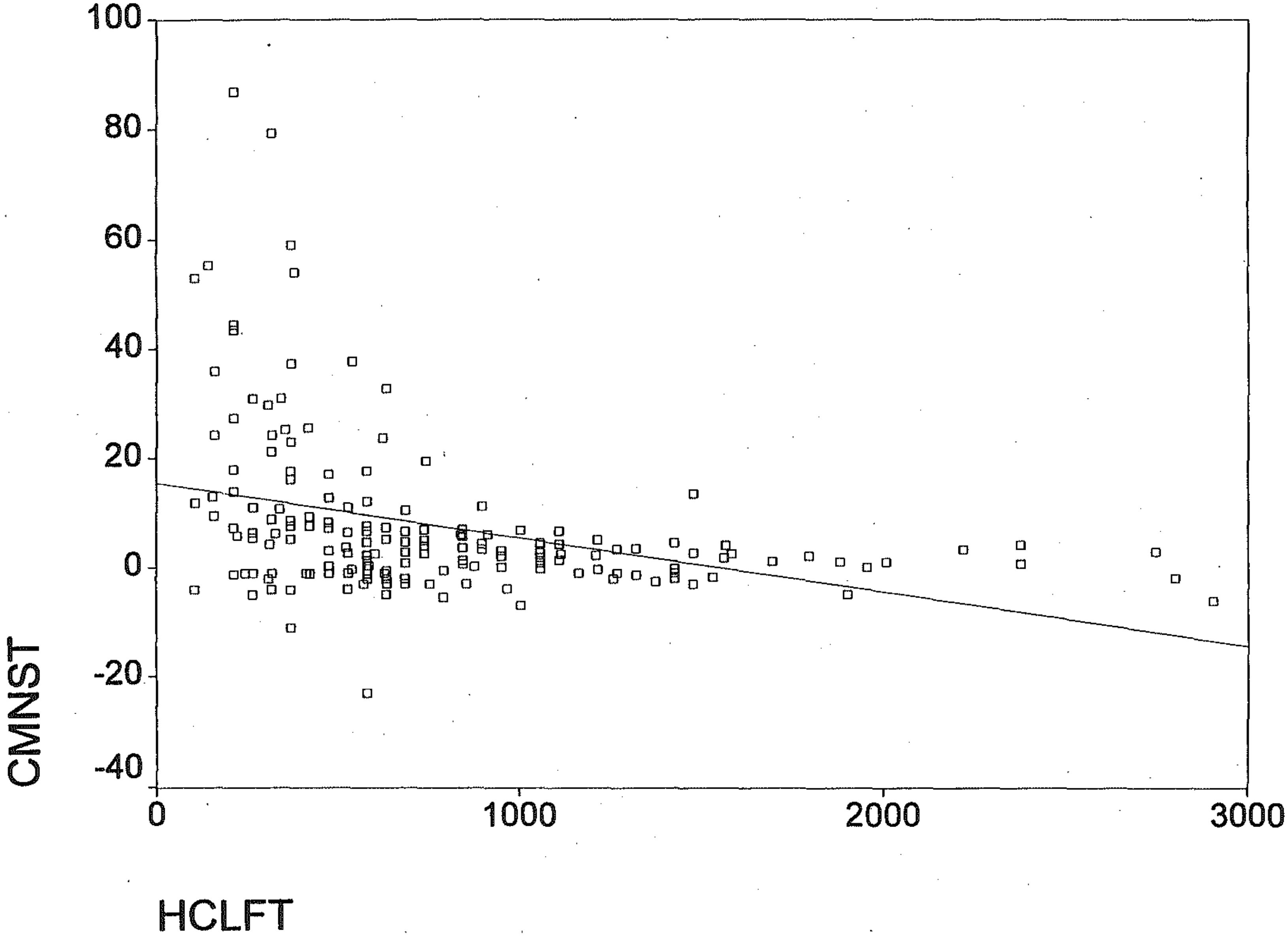


Figure 13 **Curve crash rate minus tangent crash rate (C-T), regression line for various values of curve length in feet (HCLFT)**

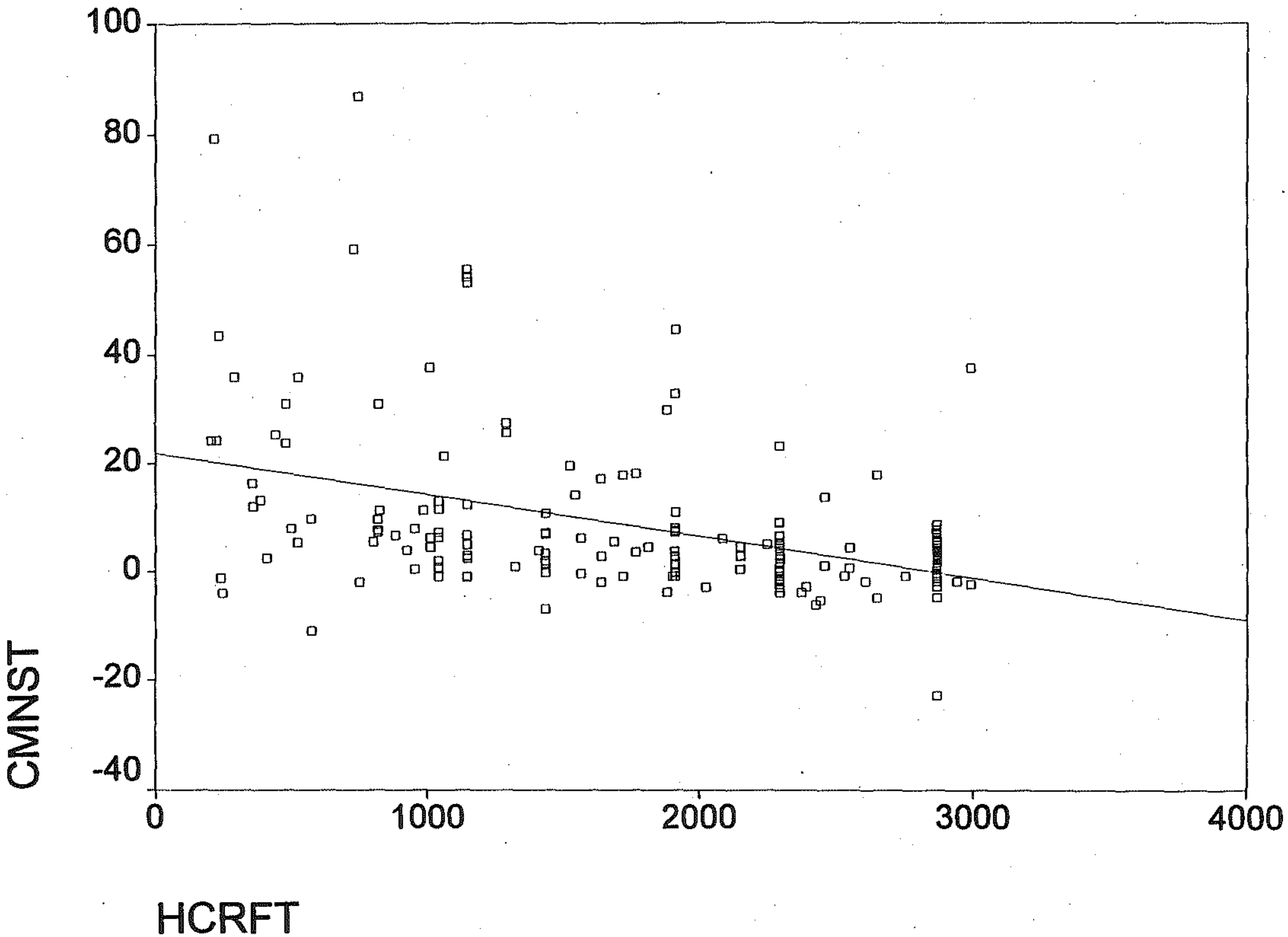


Figure 14 **Curve crash Rate minus tangent crash rate (C-T), regression line for various values of curve radius in feet (HCRFT)**

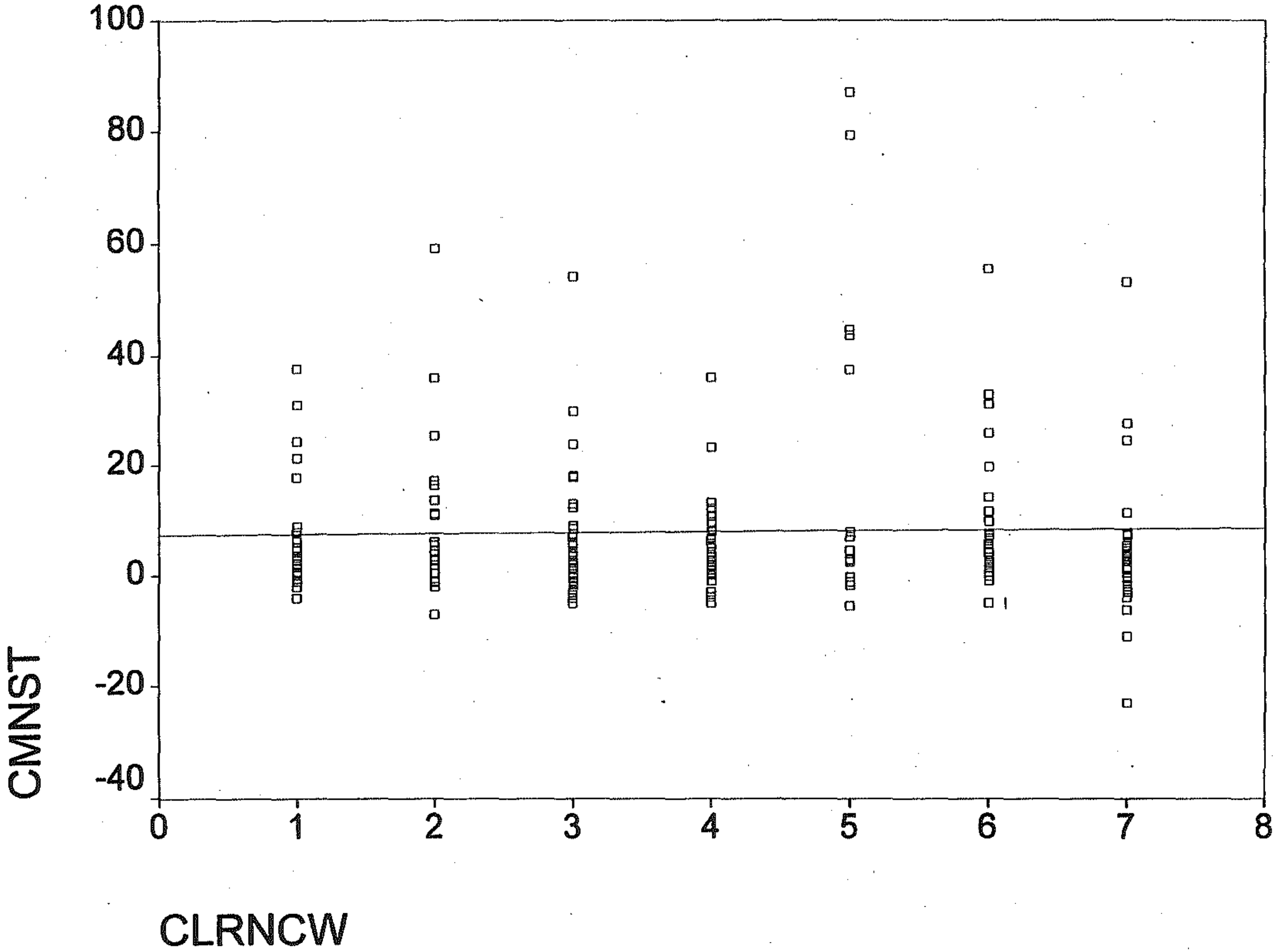


Figure 15 **Curve crash rate minus tangent crash rate (C-T), regression line for various values of roadside clearance (CLRNCW)**

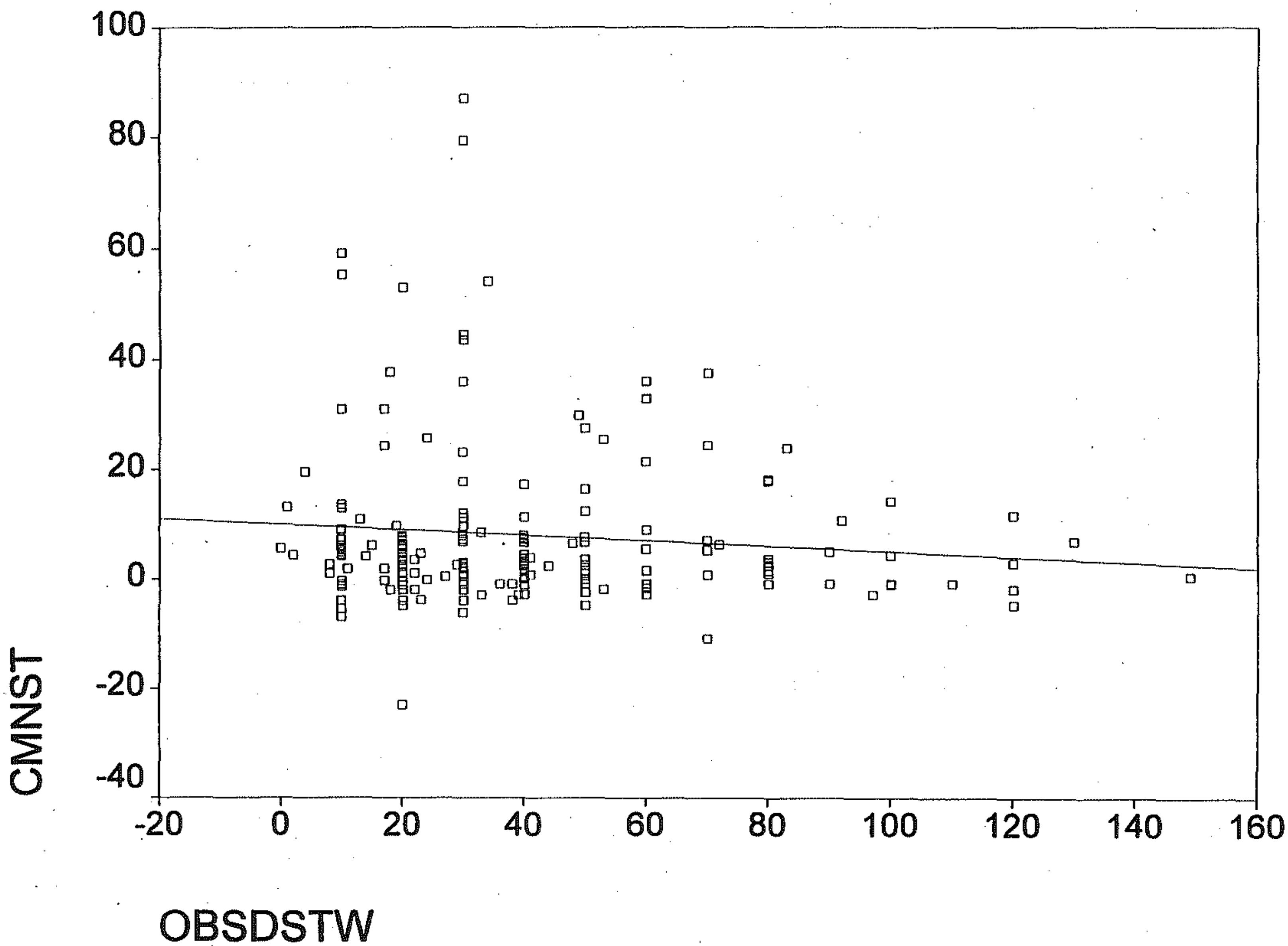


Figure 16 **Curve crash Rate minus tangent crash rate (C-T), regression line for various values of sight distance to the beginning of curve (OBSDSTW)**

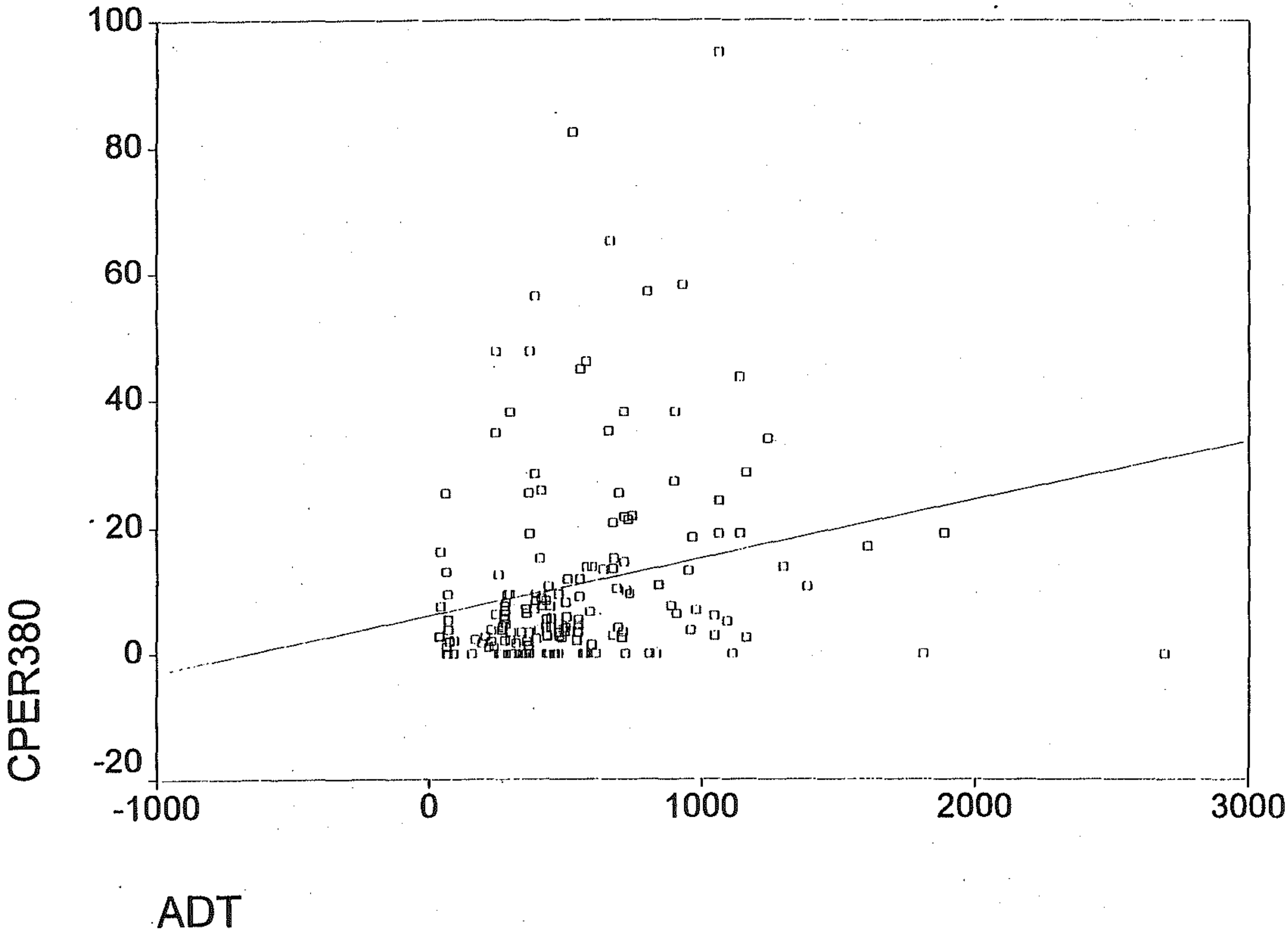


Figure 17 Curve crash rate (Cper380), regression line for various values of average daily traffic (ADT)

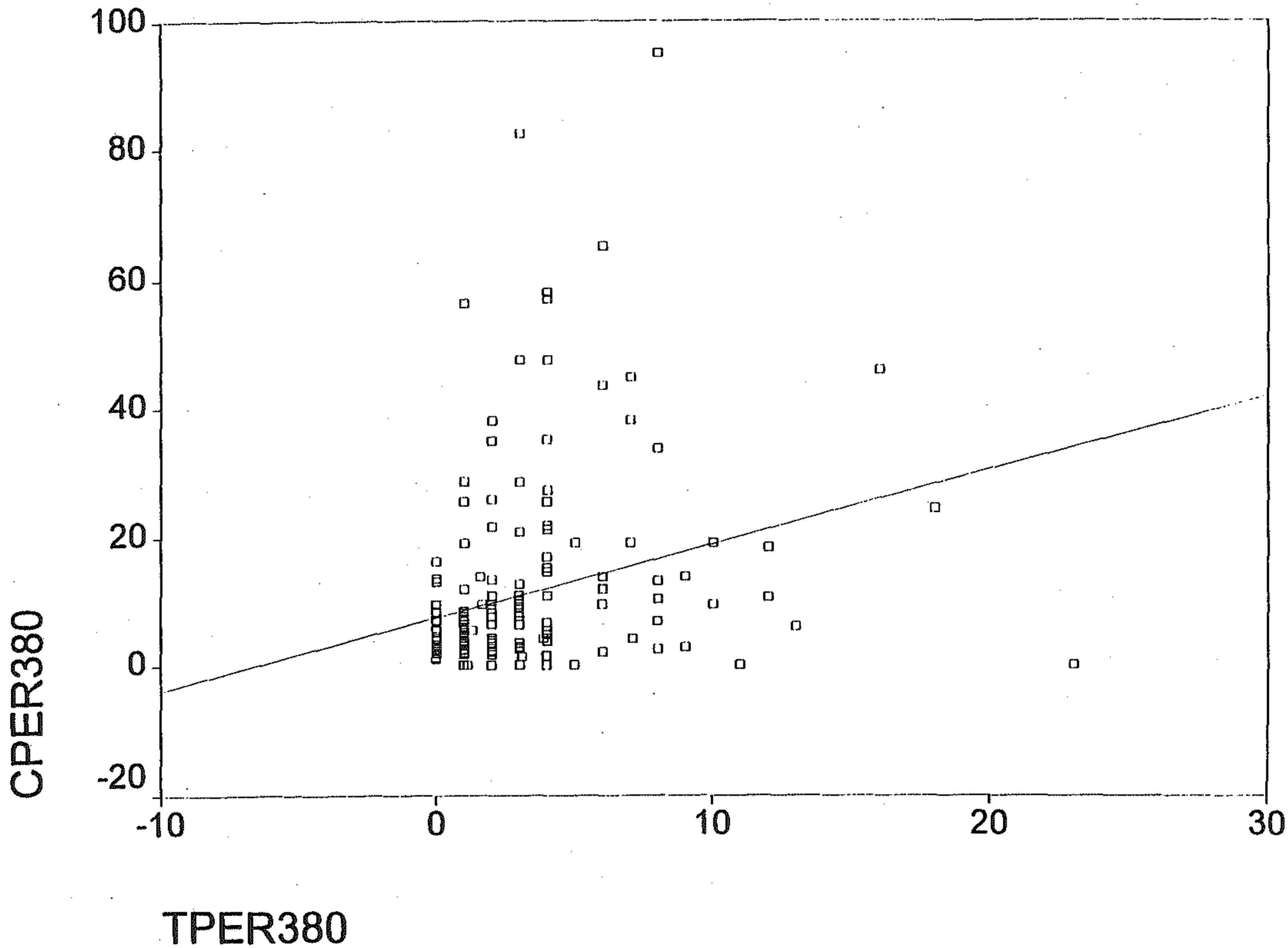


Figure 18 Curve crash rate (Cper380), regression line for various values of tangent crash rate (Tper380)

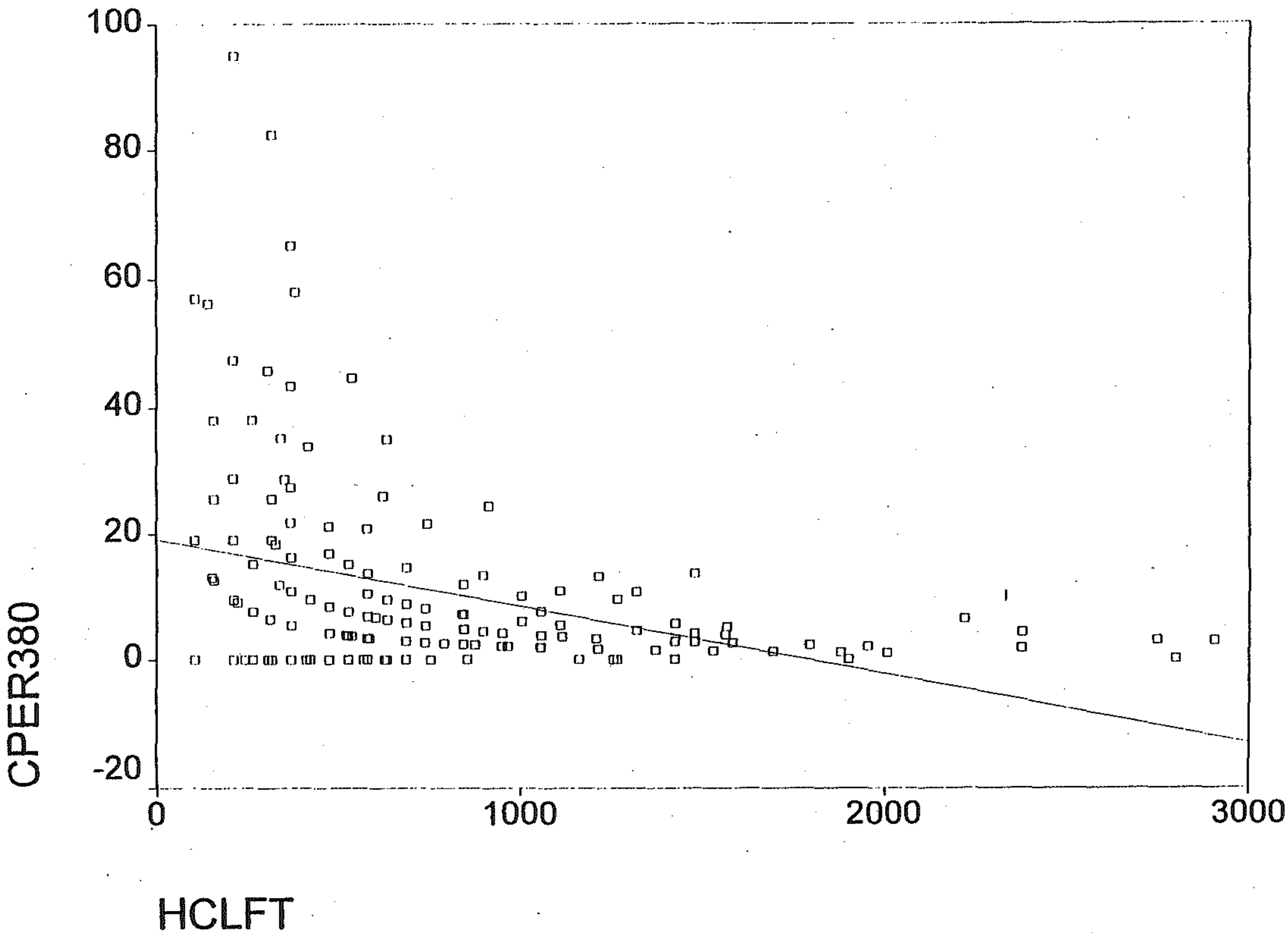


Figure 19 Curve crash rate (Cper380), regression line for various values of curve length in feet (HCLFT)

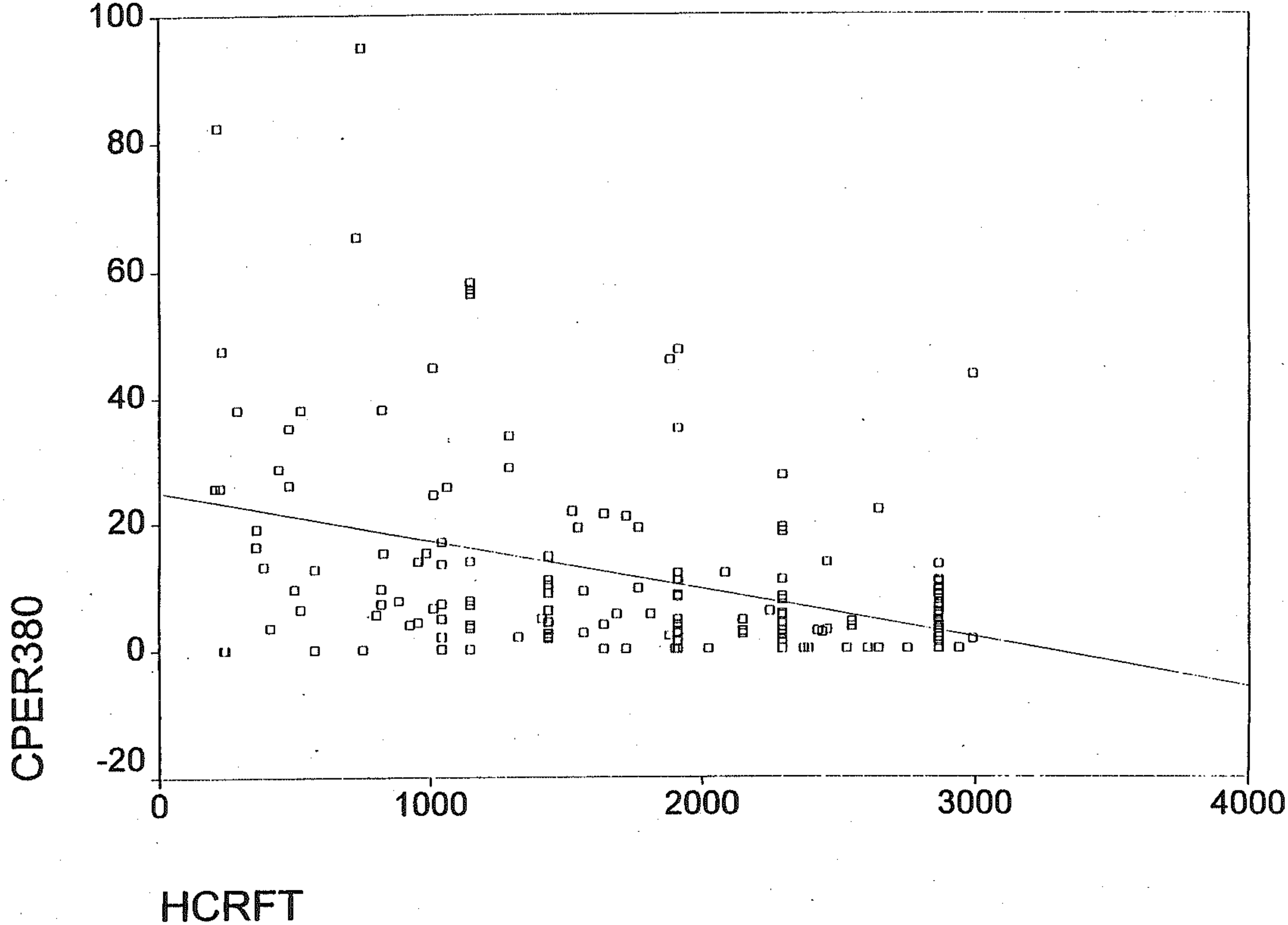


Figure 20 **Curve crash rate (Cper380), regression line for various values of curve radius in feet (HCRFT)**

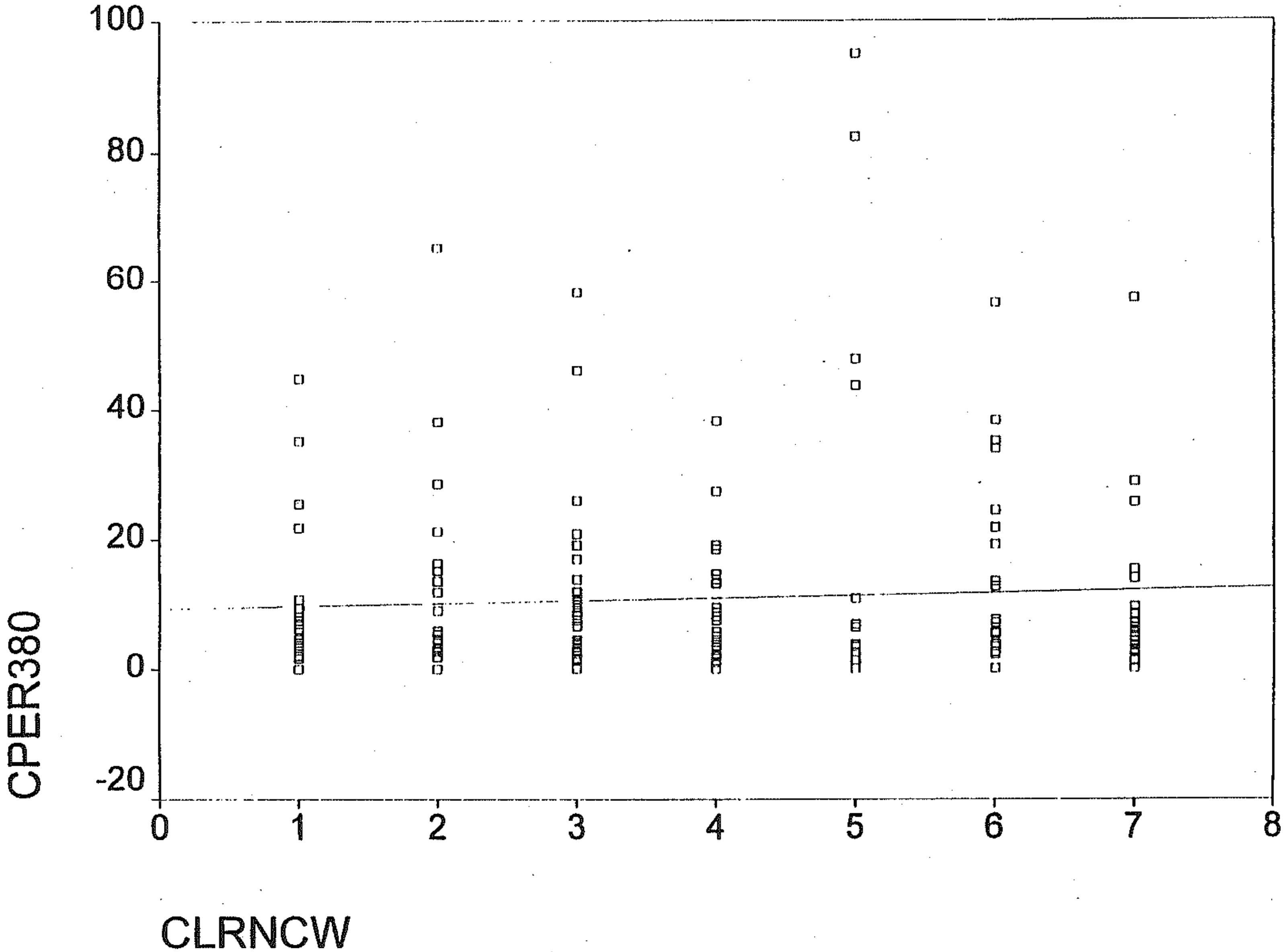


Figure 21 Curve crash rate (Cper380), regression line for various values of roadside clearance (CLRNCW)

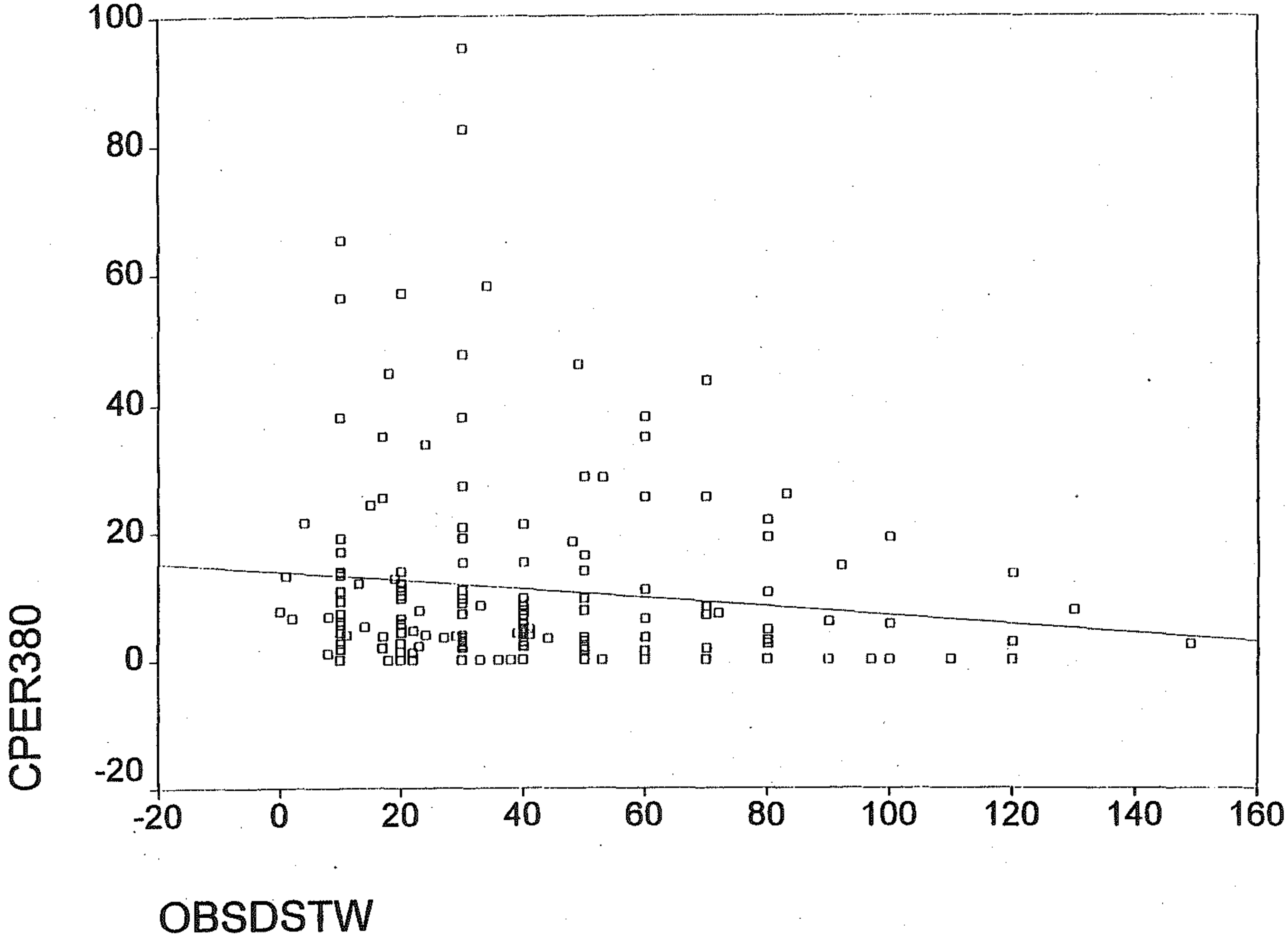


Figure 22 **Curve crash rate (Cper380), regression line for various values of sight distance to the beginning of curve (OBSDSTW)**

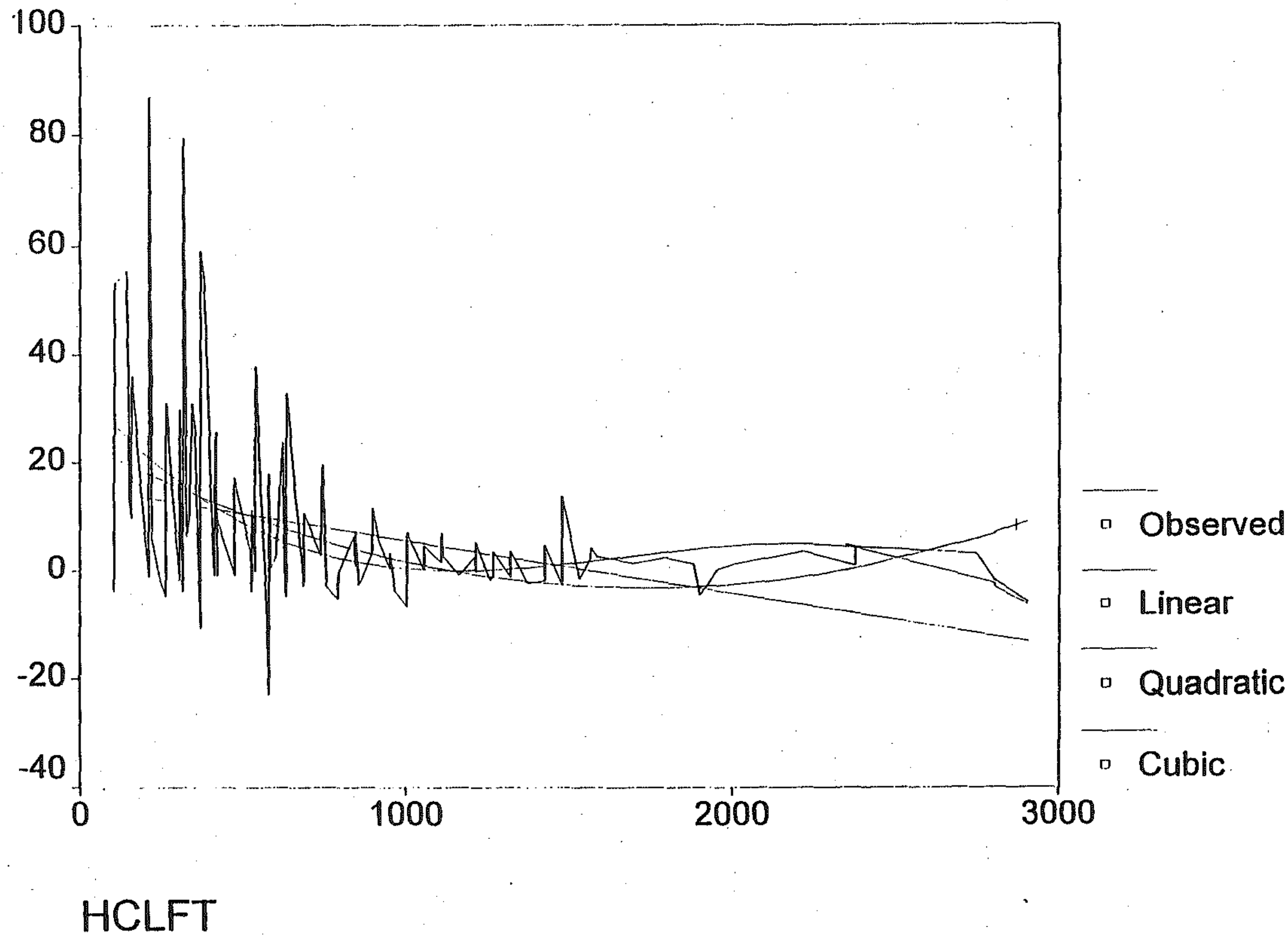


Figure 23 **Curve crash Rate minus tangent crash rate (C-T), regression lines for various values of curve length in feet (HCLFT)**

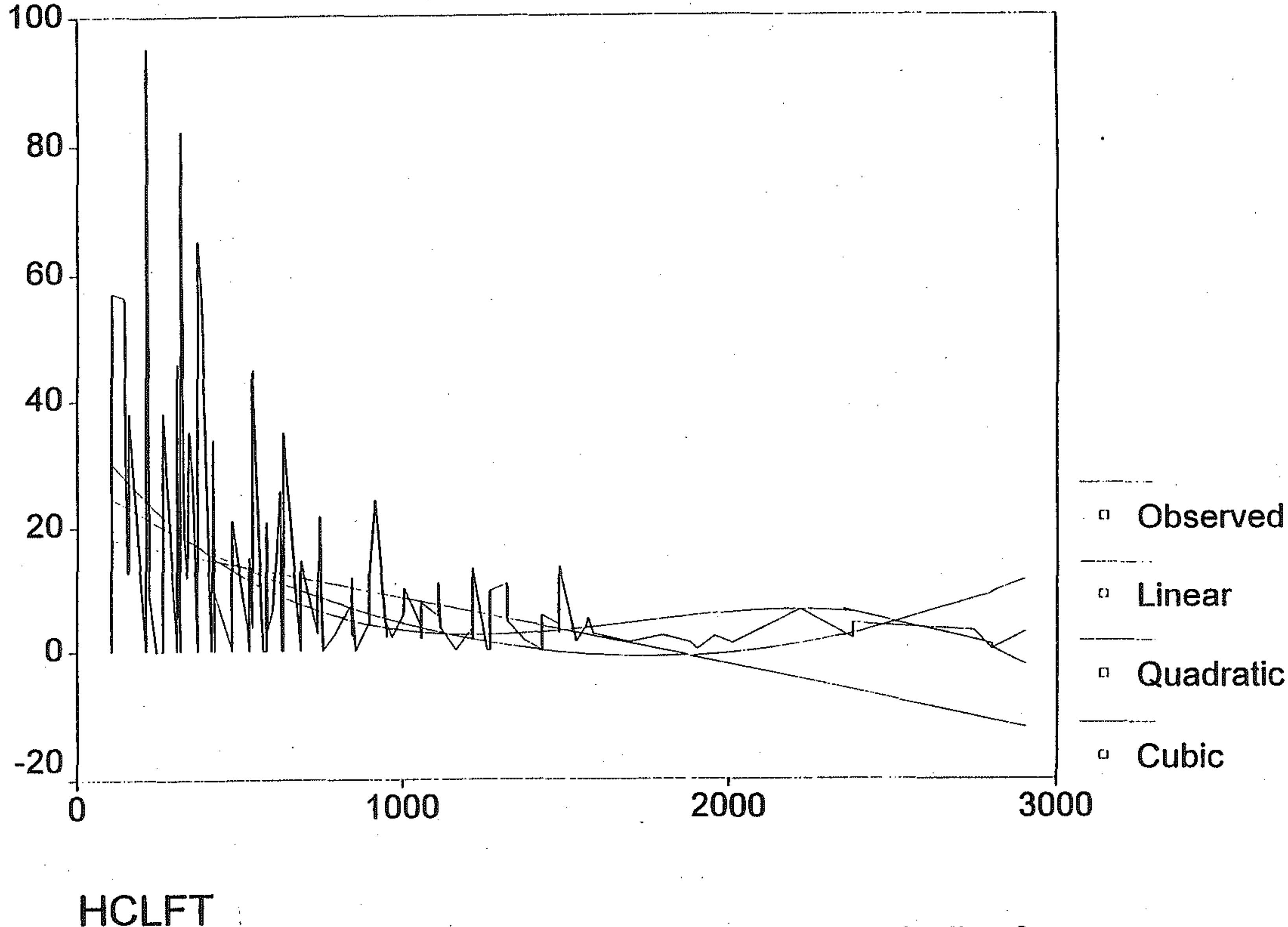


Figure 24 **Curve crash rate (Cper380), regression lines for various values of curve length in feet (HCLFT)**

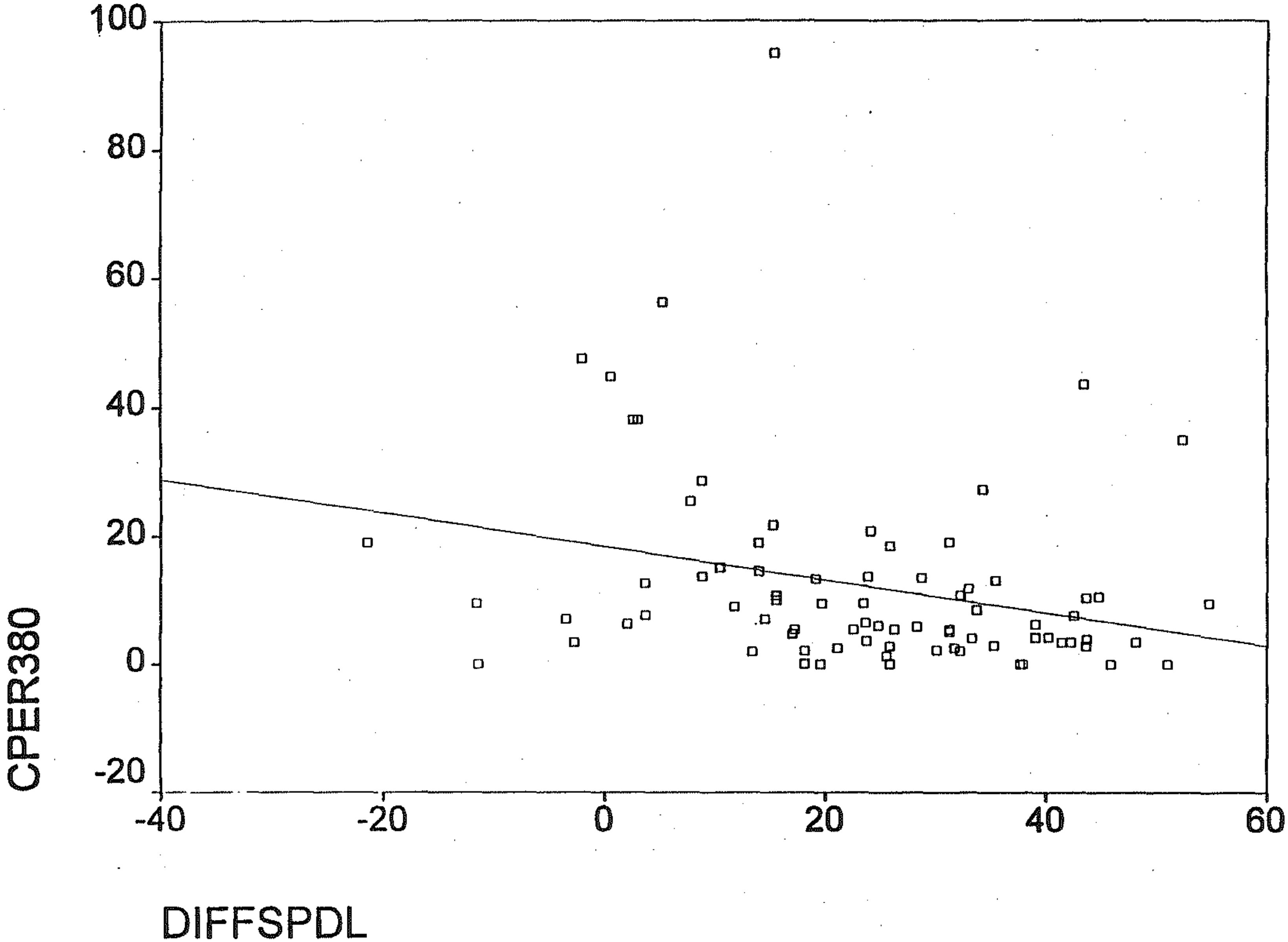


Figure 25 **Curve Crash rate (Cper380), regression line for various values of speed difference (DIFFSPDL)**

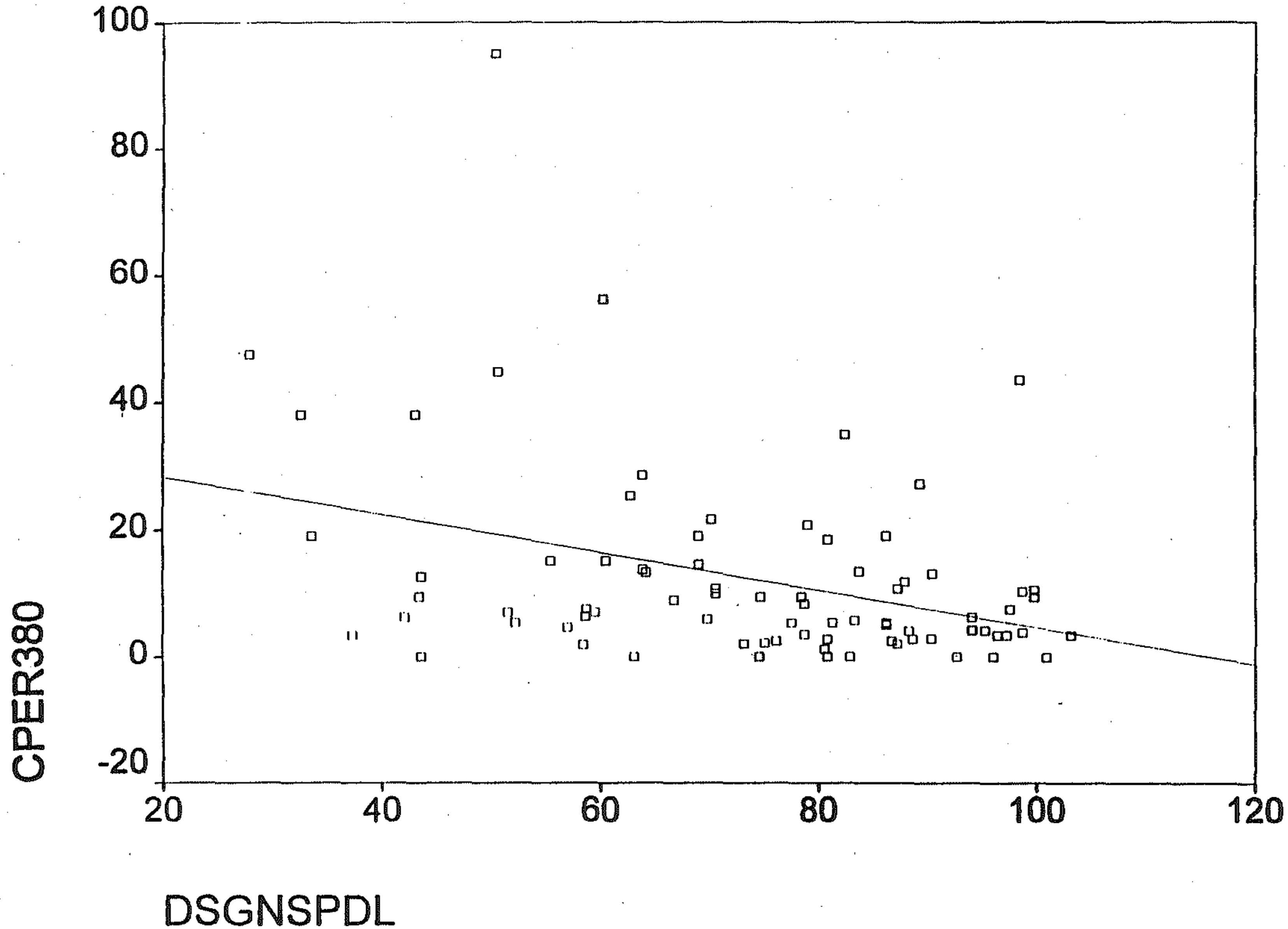


Figure 26 **Curve crash rate (Cper380), regression line for various values of design speed difference (DSGNSPDL)**

Regression Equation:

CPER380=7.35 MPHS -6.32 HCLKFT+.936 TPER380 -4.73 HCRKFT+20.034

Model	R	R Square
1	.408[a]	.166
2	.484[b]	.234
3	.519[c]	.269
4	.547[d]	.299

Coefficients[a]

Model 4		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
	(Constant)	20.034	2.802		7.150	.000	14.504	25.564
	HCRFT	-4.73E-03	.001	-.252	-3.328	.001	-.008	-.002
	TPER380	.936	.297	.205	3.147	.002	.349	1.523
	HCLFT	-6.32E-03	.002	-.223	-3.099	.002	-.010	-.002
	MPHS	7.350	2.691	.187	2.731	.007	2.039	12.661

a. Dependent Variable: CPER380

Table 7 **Results of the multiple linear regression analysis for curve crash rate (Cper380)**

TEST OF EXISTING MODELS:

The next step was to compare the curve crash data versus their predicted value from the Glennon and Zegeer models identified in the literature review.

The Glennon Model

$A = ARs (L)(V) + 0.0336 (D)(V)$ for $L >= Lc$

where,

A = Total number of crashes on the roadway segment.

ARs = Crash rate on comparable straight roadway segments in crashes per million vehicle miles.

L = Length of highway roadway segment in miles

V = Traffic volume in millions of vehicles

D = Curvature in degrees

Lc = Length of curved component in miles

For AR_s the value of Tper380 was used. This value was converted to appropriate units for the comparison.

The Zegeer Model

$$A=[1.552(L)(V)+0.014(D)(V)-0.012\,(S)(V)](0.978)^{(W-30)}$$

where:

A=number of total crashes on the curve in a 5-year period.

L=length of curve in miles (or fraction of a mile)

V=volume of vehicles in million vehicles in a 5-year period passing through the curve (both directions)

D=degree of curve

S=presence of spiral, S=0 if no spiral exists and S=1 if there is a spiral.

W=width of the roadway on the curve in feet.

For the Zegeer model the predicted values were obtained for both the with spiral, ZegeerS, and without spiral, ZegeerM, assumptions.

The plots of the predicted values of curve crashes versus actual values of curve crashes, (Cacc), are shown in Figures 27-31. This analysis considered only "related" curve crashes with the model adjusted for the length of the individual curves, not for the 612 meters. While both the Zegeer model and the Glennon model appear to show the correct trend, neither model explains the variation in "related" crash rates observed in the Michigan data. Thus it does not appear that these models are beneficial in identifying curves that should be reviewed for possible safety improvements.

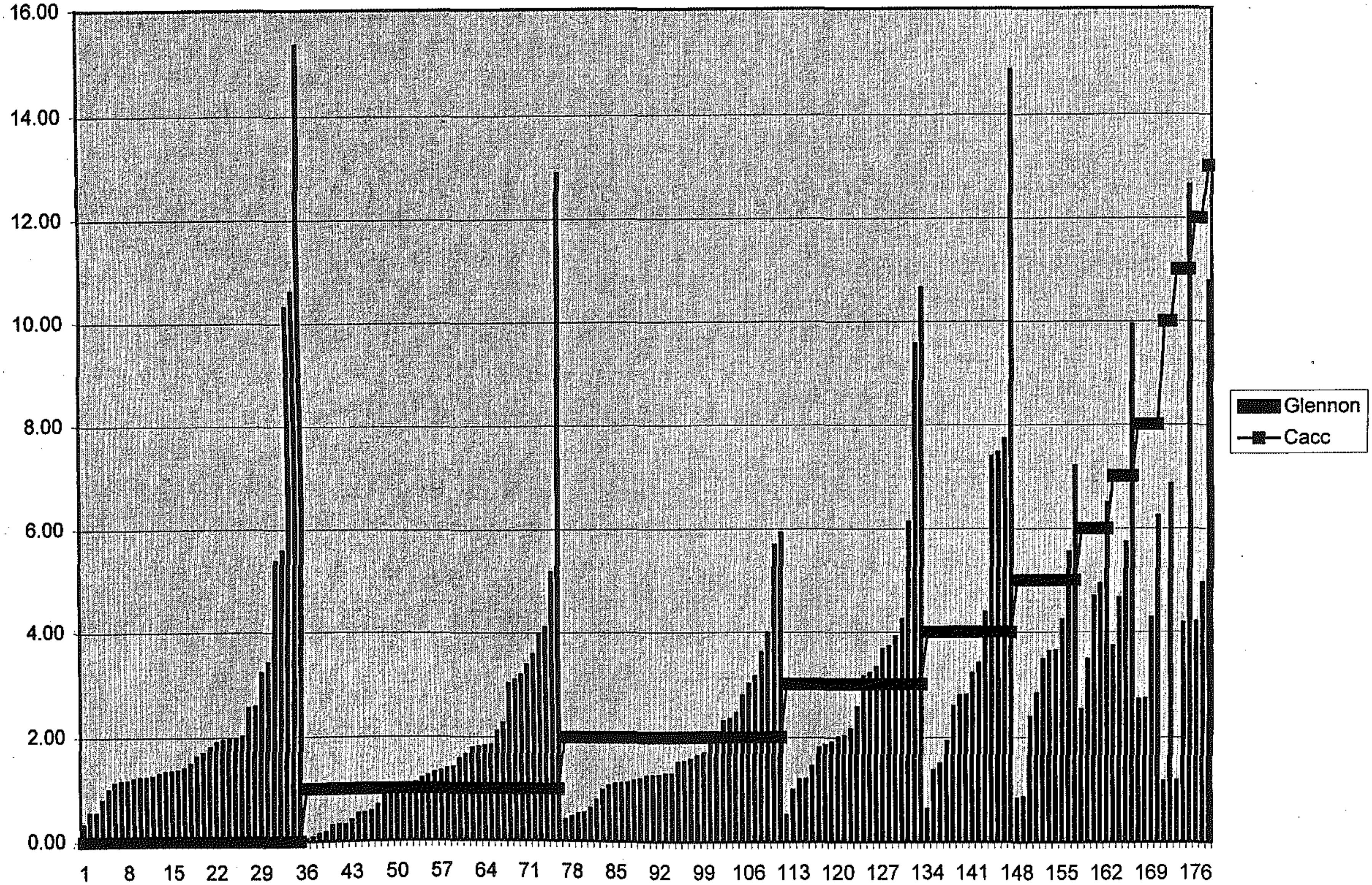


Figure 27 Comparison of the predicted number of curve crashes using Glennon's model (Glennon), and the actual number of curve crashes (Cacc)

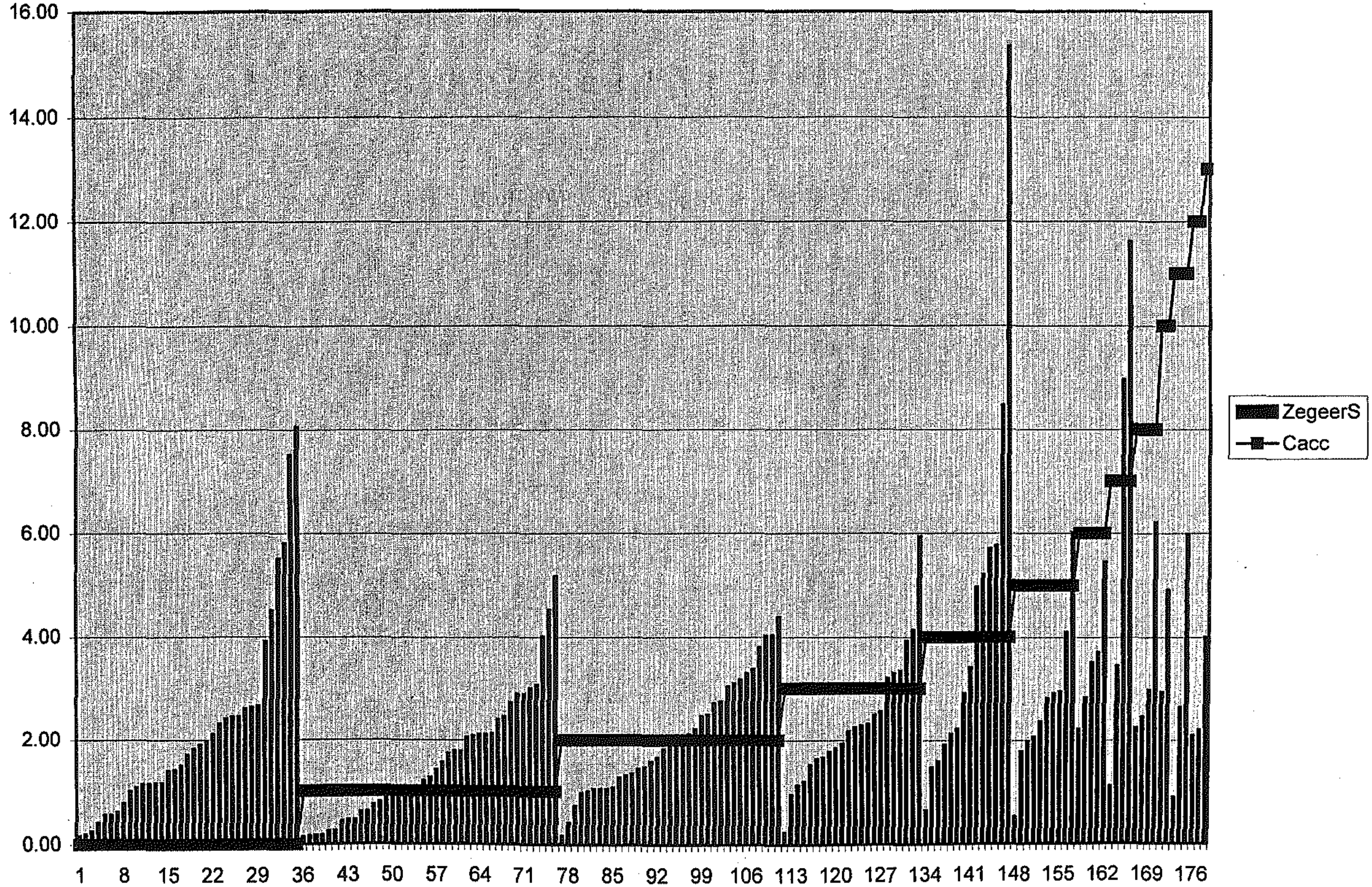


Figure 28 Comparison of the predicted number of curve crashes using Zegeer's model with spiral (ZegeerS), and the actual number of curve crashes (Cacc)

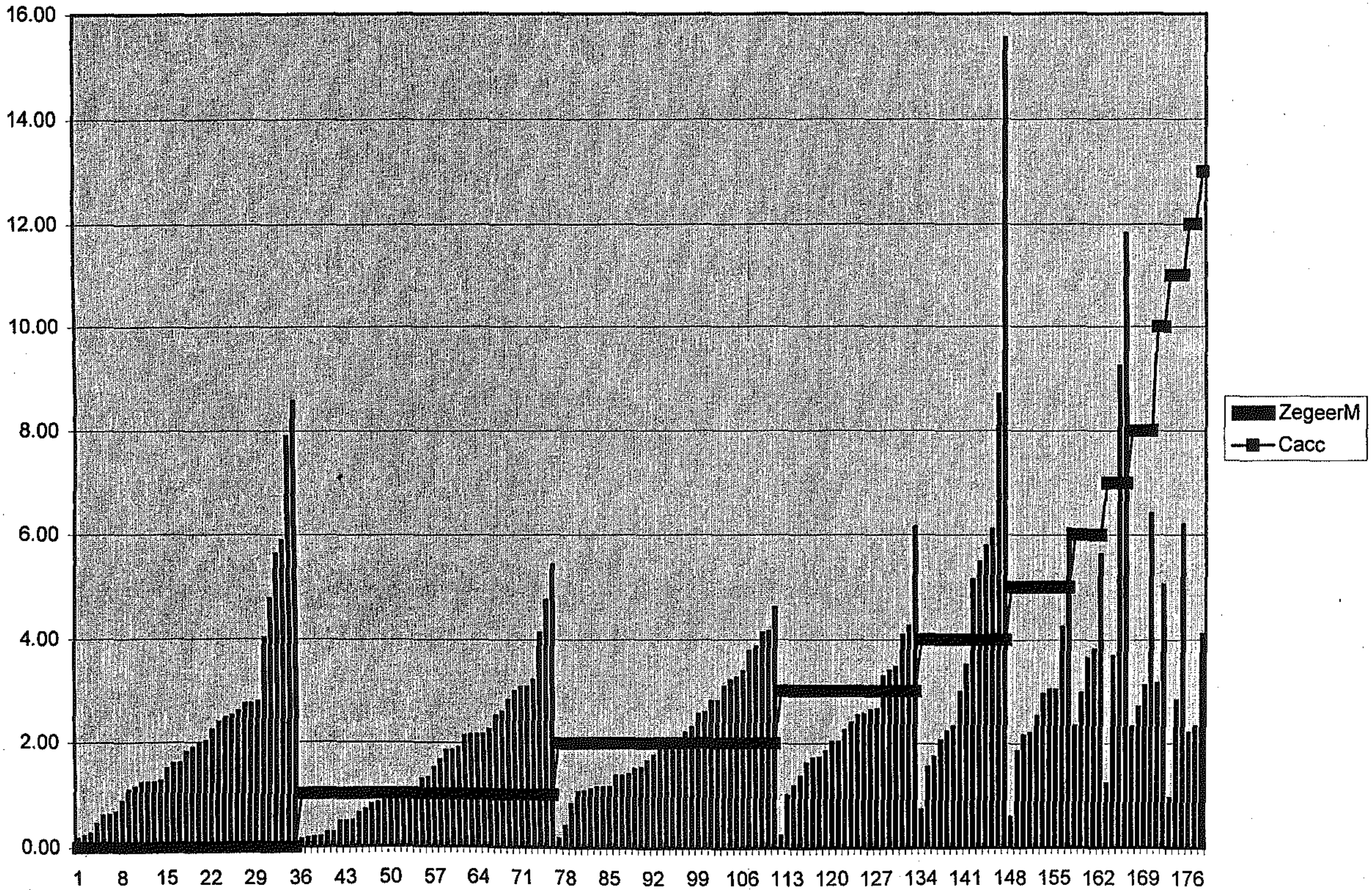


Figure 29 Comparison of the predicted number of curve crashes using Zegeer's model without spiral (ZegeerM), compared with the actual number of curve crashes (Cacc)

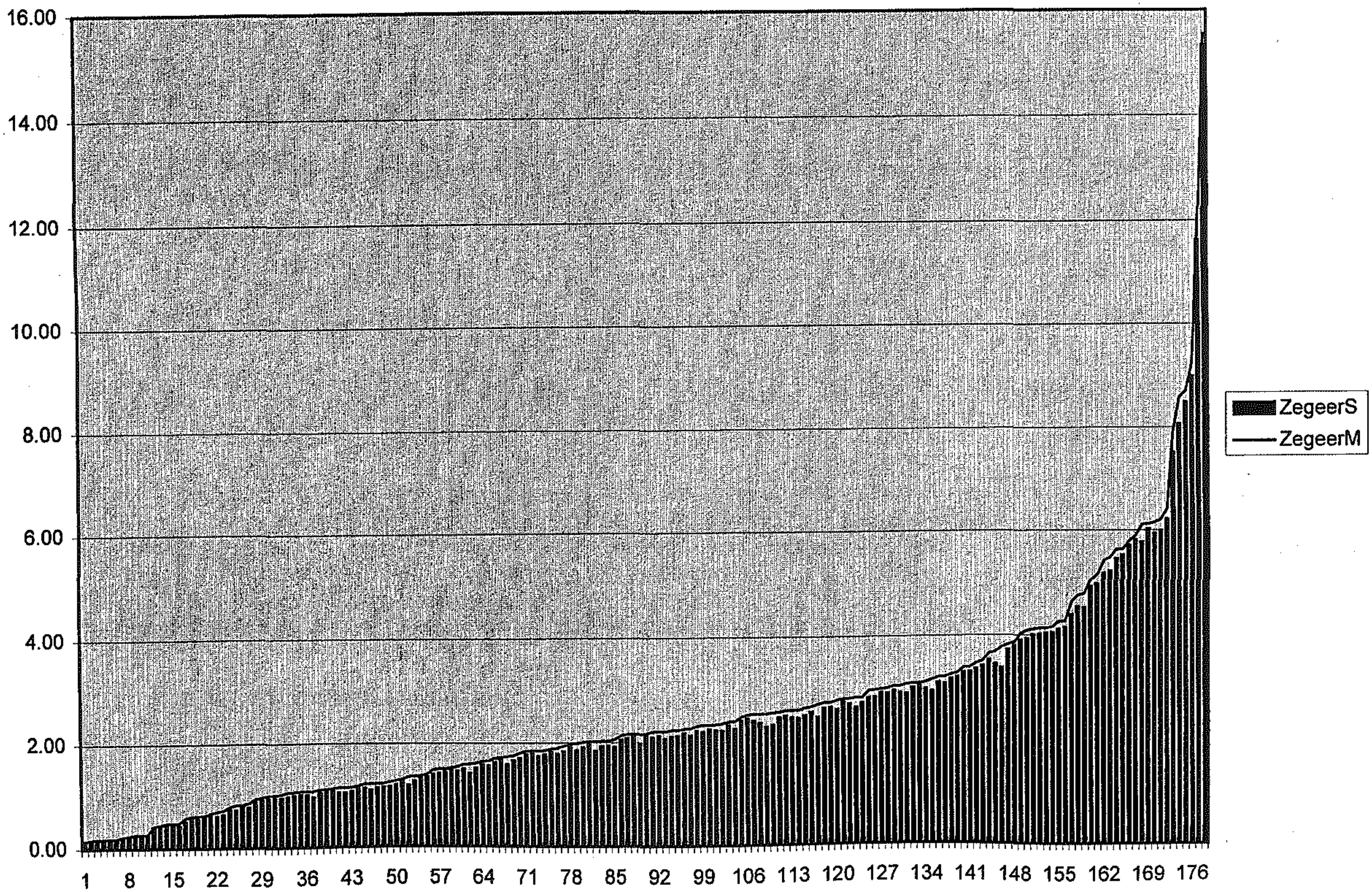


Figure 30 Comparison of the predicted number of curve crashes using Zegeer's model without spiral (ZegeerM), and that of the model with spiral (ZegeerS)

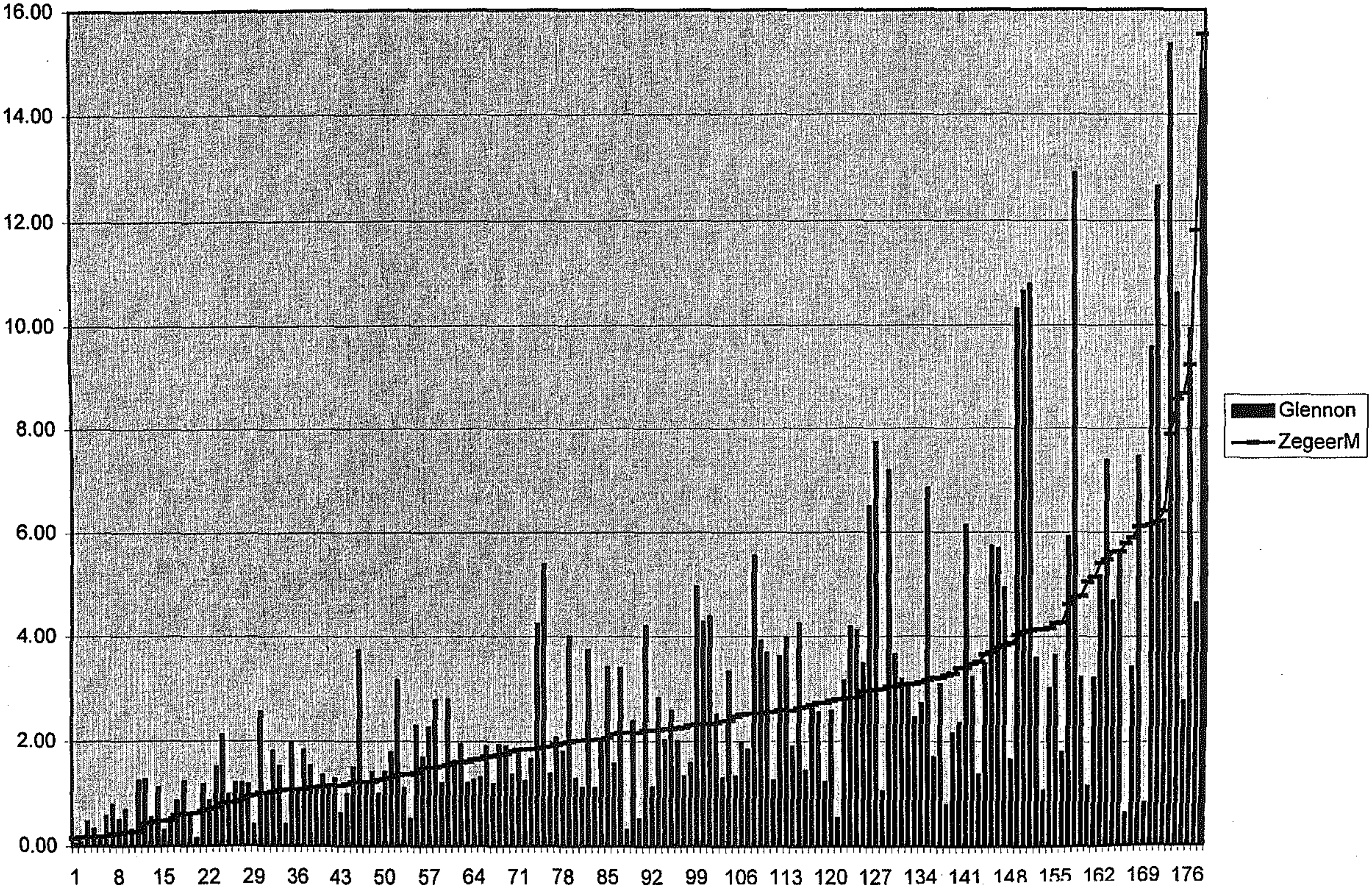


Figure 31 Predicted curve crashes using the Glennon's model (Glennon), arranged in ascending order of predicted curve crashes by Zegeer's model without spiral (ZegeerM)

Alternative Model Structures:

Having determined that the variation in crash frequency found on Michigan curves can not be satisfactorily explained by models based on simple linear regression, simple non-linear regression, multiple linear regression or multiple non-linear regression, alternative statistical techniques were tested to determine if these techniques could satisfactorily "explain" the data variation.

Discriminant analysis, cluster analysis and factor analysis techniques (as described in the following sections),were utilized.

DISCRIMINANT ANALYSIS

DESCRIPTION:

Discriminant analysis is a multivariate technique used to distinguish between two or more groups of cases and for studying the overlap between groups, or divergence of one group from the others. Statistically the objective is to define discriminating functions by weighting and linearly combining the variables such that the groups become associated with variables as distinctly as possible.

The variables with a high contribution toward explaining membership in each group, generally not all the original variables, are considered the predictor variables or the discriminating variables. It is then possible to predict group membership by their association with these discriminating variables.

The discriminant functions can be thought of as the axis of a geometric space in which each group centroid is a point. The weighting coefficients then can be interpreted as the contribution of a variable along the respective dimension of such space.

For this study, discriminant analysis was used to determine the variables which can be used to distinguish between high and low crash rate curves. The analysis was conducted with the definition of high and low crash rates based on Cper380 and then again with some of the curves removed from the sample as explained on page 72.

Analysis and Results:

All of the variables included in the database were used to conduct the first discriminant analysis. For this study, the analysis was used to define membership in one of two groups, either a high crash group or a low crash group.

A value of Cper380 equals 5 resulted in approximately half of the curves being defined as belonging to the high crash group and the other half being classified as the low crash group and it was selected as the defining value between high and low crash rates.

The results of this analysis are shown in Table 8. The curve length and the curve radius were the two most important discriminating variables followed by ADT. Using only these variables 71.9% of all cases were correctly classified. There were 26 curves that were placed in the low category that had a Cper380 value of greater than 5, and 24 curves that were misclassified in the other direction.

Since our primary interest is determining whether it was possible to distinguish between high crash locations and low crash locations (rather than some intermediate group), the data set was reduced to eliminate the curves with a value of Cper380 approximately equal to five. A new variable called Modified Cper380 (ModCper) was defined. This variable is the same as Cper380 but 15 curves with a Cper380 value near the average for all curves were excluded from the analysis.

Table 9 shows the results of the analysis using the modified Cper380 as the grouping variable. Group 2 being for ModCper > 7 crashes and group 1 for ModCper < 5 crashes. In this analysis the curve sign and turn sign were replaced by a single variable called

CTsign. If either sign were present, CTsign was assigned the value of 1 otherwise 0 (zero).

The curve length, the presence of a turn or curve warning sign, the radius of the curve and Tper380 are the discriminating variables identified in this case. Using these variables 79.1% of the curves were correctly classified. As expected , removing the marginal cases improved the predictive capability of the model. With this modification, only 16 curves were misplaced as low and 18 curves were misplaced as high.

For the next analysis the difference between the curve crash rate (Cper380) and the tangent crash rate (Tper380) is used as a grouping measure. This variable, (C-T), was also modified to more clearly distinguish the curves with higher crash rates relative to their tangent crash rates. The cases with curve crashes nearly equal to the tangent crashes were eliminated. A total of 43 curves with C-T=-1.36 to C-T=1.90 were eliminated from the analysis.

As shown in Table 10, the variables Curve Radius, Curve Length and the presence of a warning sign are the three most important discriminating variables. For this analysis, 75.6% of the curves were correctly classified using these three variables. Using this model, 90.7% of the high crash rate curves were correctly identified, with only 10 curves being misclassified in this direction. The problem with this model is that too many low crash rate curves, (23) were placed in the high crash category.

Variables in the Analysis

Step		Tolerance	Sig. of F to Remove	Wilks' Lambda
3	HCLFT	.866	.001	.827
	HCRFT	.848	.002	.822
	ADT	.978	.002	.821

Classification Results[a]

			Predicted Group Membership		
		GRPCLT5	1.00	2.00	Total
Original	Count	1.00	64	24	88
		2.00	26	64	90
	%	1.00	72.7	27.3	100.0
		2.00	28.9	71.1	100.0

a. 71.9% of original grouped cases correctly classified.

Table 8 Results of the discriminant analysis for curve crash rate (Cper380)

Variables in the Analysis

Step		Tolerance	Sig. of F to Remove	Wilks' Lambda
4	HCLFT	.882	.000	.725
	CTSIGN	.990	.000	.718
	HCRFT	.866	.001	.706
	TPER380	.971	.004	.697

Classification Results

			Predicted Group Membership		
		GRPNT5T7	1.00	2.00	Total
Original	Count	1.00	70	18	88
		2.00	16	59	75
	%	1.00	79.5	20.5	100.0
		2.00	21.3	78.7	100.0

a. 79.1% of original grouped cases correctly classified.

Table 9 Results of the discriminant analysis for modified curve crash rate (ModCper)

Variables in the Analysis

Step		Tolerance	Sig. of F to Remove	Wilks' Lambda
1	HCRFT	1.000	.000	
2	HCRFT	.987	.003	.917
	CTSIGN	.987	.004	.912

Classification Results

		LOCMNST	Predicted Group Membership 1.00	Predicted Group Membership 2.00	Total
Original	Count	1.00	5	23	28
		2.00	10	97	107
	%	1.00	17.9	82.1	100.0
		2.00	9.3	90.7	100.0

75.6% of original grouped cases correctly classified.

Table 10 Results of the discriminant analysis for modified curve minus tangent crash rate (ModC-T)

The next group of analyses was performed using Cluster Analysis.

CLUSTER ANALYSIS

DESCRIPTION:

Cluster Analysis is a systematic technique to look for regularities in a data set. Once the regularities are depicted, this procedure groups the data based on these regularities and their interpretations. Unlike Discriminate Analysis, which requires prior knowledge of the group membership for the data cases, cluster analysis does not require such knowledge.

Cluster analysis uses the concept of "distance" and "similarity" in generating new clusters or collapsing them into a lesser number of clusters. There are many methods of calculating "distance" and the analyst must use interpretative judgment and inspection in addition to the quantitative analysis.

Cluster analysis was used to identify the variables with a strong association with the crash rate. While any number of clusters can be created, three clusters were used in this study. One cluster identified the variables associated with curves that have a low crash rate, a second cluster was formed around curves with an intermediate crash rate, and the third around high crash rate curves.

Analysis and Results:

Utilizing cluster analysis produced results which proved to be useful for the objectives of this study. Table 11 shows the output for a three cluster case in which Modified Cper380, as discussed previously, was used to define the number of curves included in the analysis.

The clustering of high, medium and low crash rate curves with other variables is clear, with cluster one having a crash rate of 3.08, cluster two a crash rate of 7.78 while the third cluster has a crash rate of 18.05. Variables such as Lane Width (ALW), that show little variance between the three clusters indicate that either this variable is unimportant in predicting the curve crashes, or that there is little variance in the variable across all curves.. For this variable the latter is true. Other variables, such as Curve Length and Radius, show great variations between at least two of the three clusters. This is an indication of an important variable in the prediction model. The important variables are shown in Table 12.

The same variables identified in the discriminant analysis were important in the cluster analysis. The ADT curve radius and length, and the presence of traffic control devices (arrow and chevron)are all important in defining the clusters. Interestingly, the high crash rate curves are associated with the highest probability of having chevrons and target arrows deployed. However, this is explained by the fact that this cluster contains the short radius curves, where these devices tend to be deployed.

An analysis using Cper380 instead of ModCper shows similar results (Table 13). Most notably, the clustering of high crash rates with short curves and low radii while the low crash rate curves are clustered with long curves with large radii. This finding is consistent with prior research. Using this measure of the crash rate, ADT was replaced by the presence of an advisory speed plate and the paved shoulder width as explanitory variables. Perhaps the most interesting cluster is the third one, which clusters moderately high crash rate curves with curves of large radius but short length. These tend to not have traffic control devices deployed because of their large radius and subsequently their high design speed.

Tables 14 and 15 show two more cluster analysis results. These results are also in agreement with the previous findings. In Table 14 the difference between the curve crash and the tangent crash (C-T), is used as the curve crash rate variable, while in Table 15 the variable, ModC-T, as described before, was used.

It was hypothesized that the variation in crash rates within each cluster would be lower than that of all curves combined, and thus regression analysis techniques might show better results.

To test this hypothesis, simple and multiple regression were applied to each of the three clusters obtained from the cluster analysis. However, regression failed to depict even mild correlation. As examples Figures 32 through 37 show the regression plots of Cper380 with HCRFT and HCLFT for each of the referenced three clusters.

Final Cluster Centers

	Cluster		
	1	2	3
ADT	472.72	536.05	549.14
ALW	11.31	11.19	11.06
ARROW	.21	.09	.29
CHEVRON	.03	.03	.13
CLRNCW	3.69	3.66	4.09
CTSIGN	.34	.44	.56
DLNTR	.31	.19	.27
EDGLN	1.00	.98	1.00
GRAIL	.21	.13	.23
HCLFT	1704	590	520
HCRFT	2471	2383	963
MODCPER	3.08	7.78	18.05
MPHS	.10	.09	.30
NPZC	.90	1.06	1.96
OBSDSTW	45.24	44.27	38.37
PSL	54.66	54.53	53.29
PSW	10.79	6.56	7.03
SCT	1.66	1.53	1.60
TPER380	2.52	3.44	2.98
TSW	19.45	18.72	18.56

Number of Cases in each Cluster

Cluster	1	29.000
	2	64.000
	3	70.000
Valid		163.000
Missing		15.000

Table 11 The numerical values of all variables in defining the clusters grouped by the modified curve crash rate (ModCper)

Final Cluster Centers

	Cluster		
	1	2	3
ADT	472.72	536.05	549.14
ALW			
ARROW	.21	.09	.29
CHEVRON	.03	.03	.13
CLRNCW			
CTSIGN			
DLNTR			
EDGLN			
GRAIL			
HCLFT	1704	590	520
HCRFT	2471	2383	963
MODCPER	3.08	7.78	18.05
MPHS			
NPZC			
OBSDSTW			
PSL			
PSW			
SCT			
TPER380			
TSW			

Number of Cases in each Cluster

Cluster	1	29.000
	2	64.000
	3	70.000
Valid		163.000
Missing		15.000

Table 12 The numerical values of the important variables in defining the clusters grouped by the modified curve crash rate (ModCper)

Final Cluster Centers

	Cluster		
	1	2	3
ADT			
ALW			
ARROW	.19	.30	.10
CHEVRON	.03	.14	.03
CLRNCW			
CPER380	3.33	17.10	7.62
CURVES			
DLNTR			
EDGLN			
GRAIL			
HCLFT	1707	522	608
HCRFT	2490	974	2392
MPHS	.09	.32	.10
OBSDST			
PSL			
PSW	11.09	7.26	6.53
SCT			
TPER380			
TSW			
TURNS			

Number of Cases in each Cluster

Cluster	1	32.000
	2	76.000
	3	70.000
Valid		178.000
Missing		.000

Table 13 **The numerical values of the important variables in defining the clusters grouped by the curve crash rate (Cper380)**

Final Cluster Centers

	Cluster		
	1	2	3
ADT			
ALW			
ARROW	.10	.30	.19
CHEVRON	.03	.14	.03
CLRNCW			
CMNST	4.32	14.05	.88
CTSIGN			
DLNTR			
EDGLN			
GRAIL			
HCLFT	608	522	1707
HCRFT	2392	974	2490
MPHS	.10	.32	.09
OBSDSTW			
PSL			
PSW	6.53	7.26	11.09
SCT			
TSW			

Table 14 The numerical values of the important variables in defining the clusters grouped by the curve minus tangent crash rate (C-T)

Final Cluster Centers

	Cluster		
	1	2	3
ADT			
ALW			
ARROW	.11	.33	.25
CHEVRON	.04	.18	
CLRNCW			
CTSIGN			
DLNTR			
EDGLN			
GRAIL	.07	.25	.20
HCLFT	607	471	1757
HCRFT	2351	902	2481
MODCMNST	5.59	17.67	1.40
MPHS	.09	.37	.15
PSL			
PSW	6.78	7.15	12.90
SCT			
TSW			
OBSDSTW			

Table 15 The numerical values of the important variables in defining the clusters grouped by the modified curve minus tangent crash rate (ModC-T)

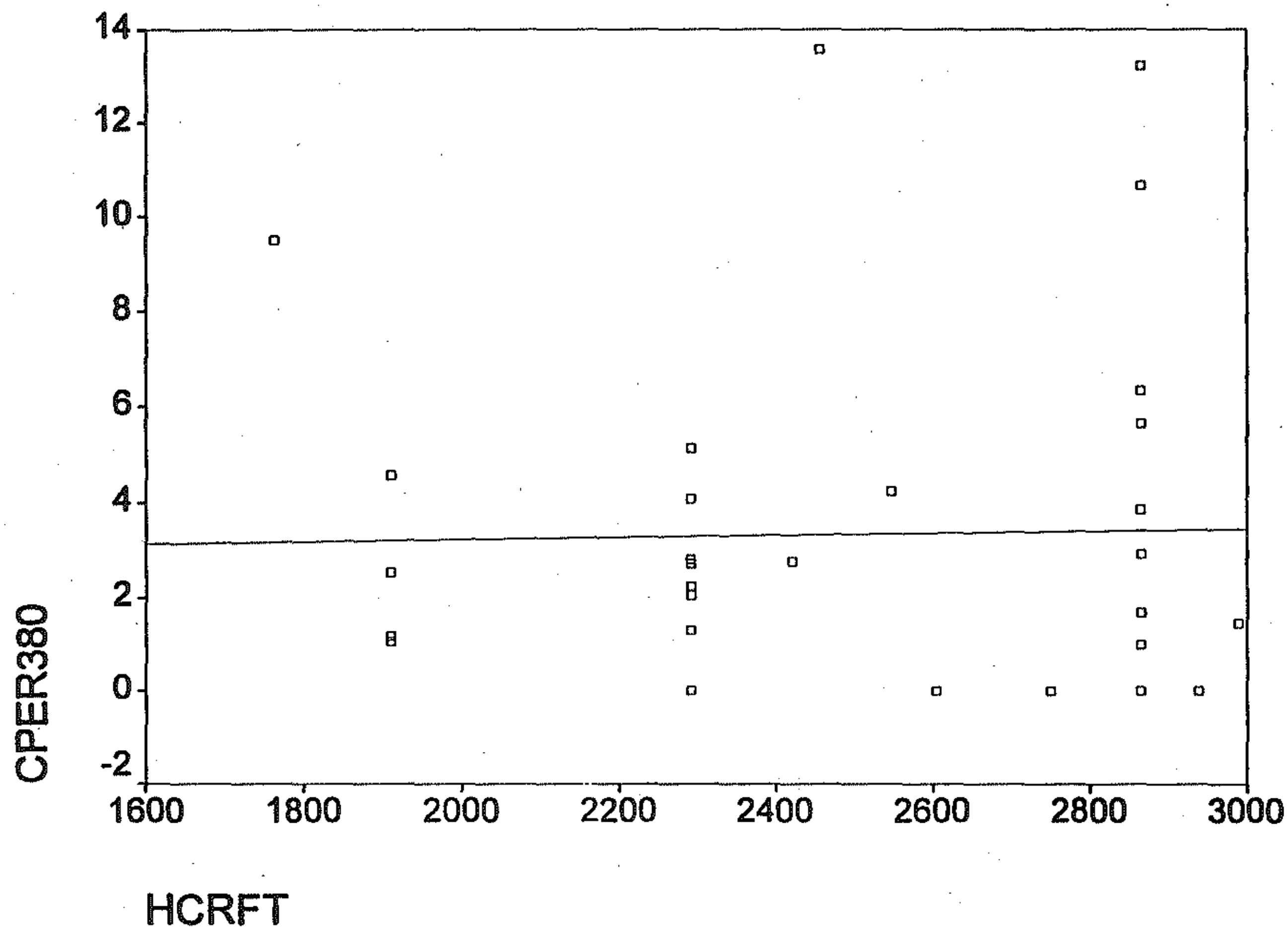


Figure 32 Curve crash rate (Cper380), regression line for various values of curve radius in feet, cluster 1

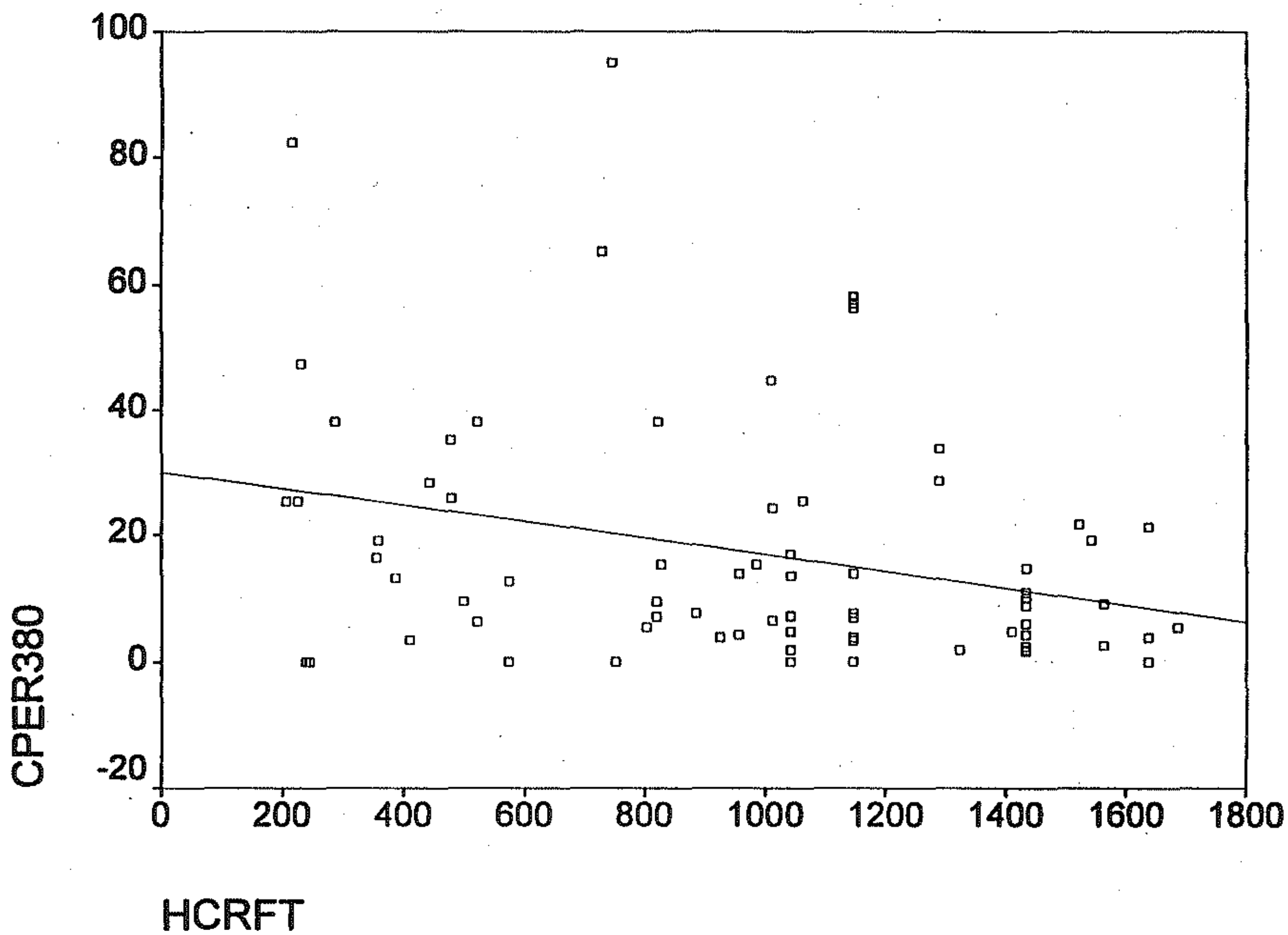


Figure 33 Curve crash rate (Cper380), regression line for various values of curve radius in feet, cluster 2

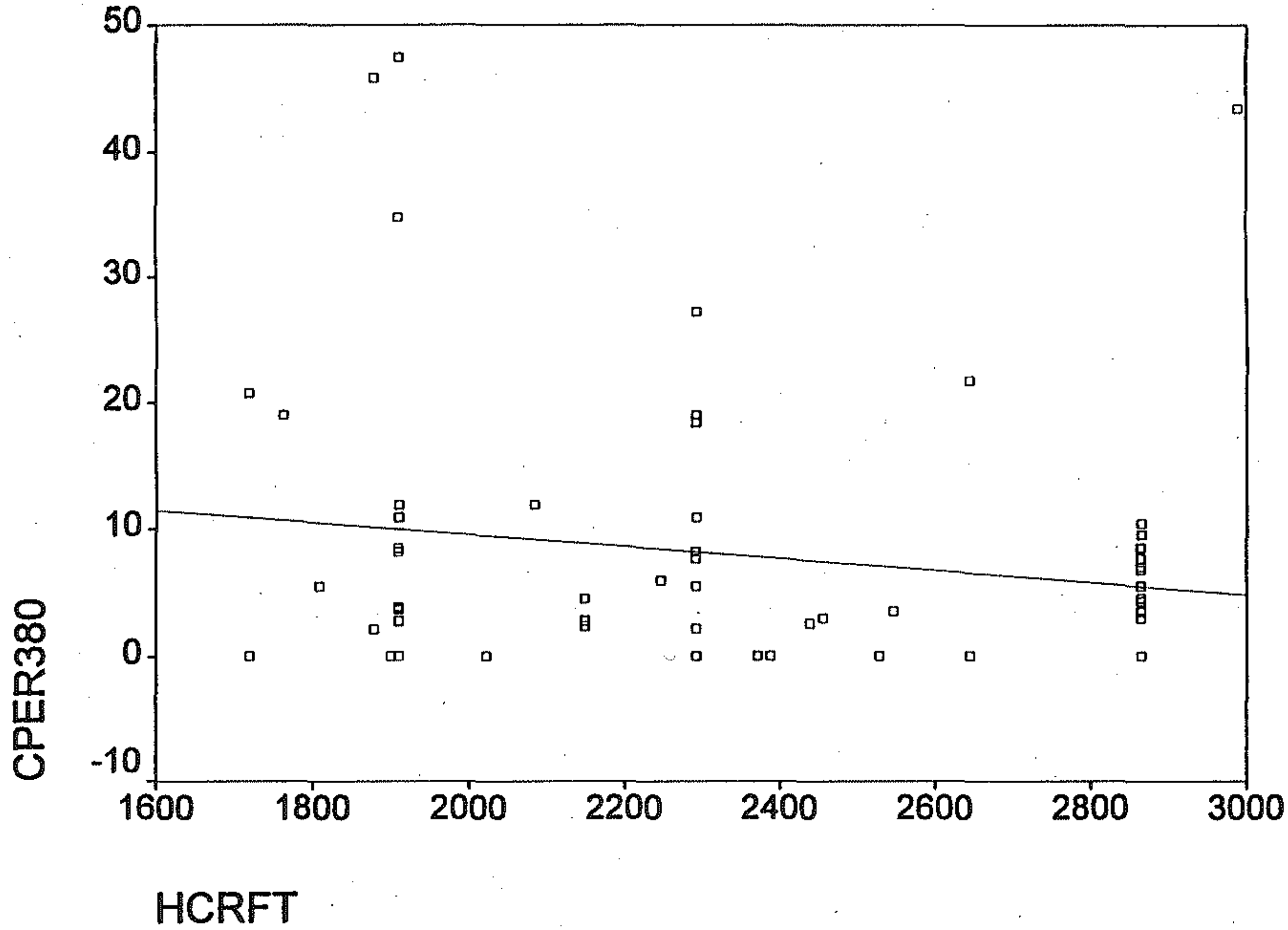


Figure 34 **Curve crash rate (Cper380), regression line for various values of curve radius in feet, cluster 3**

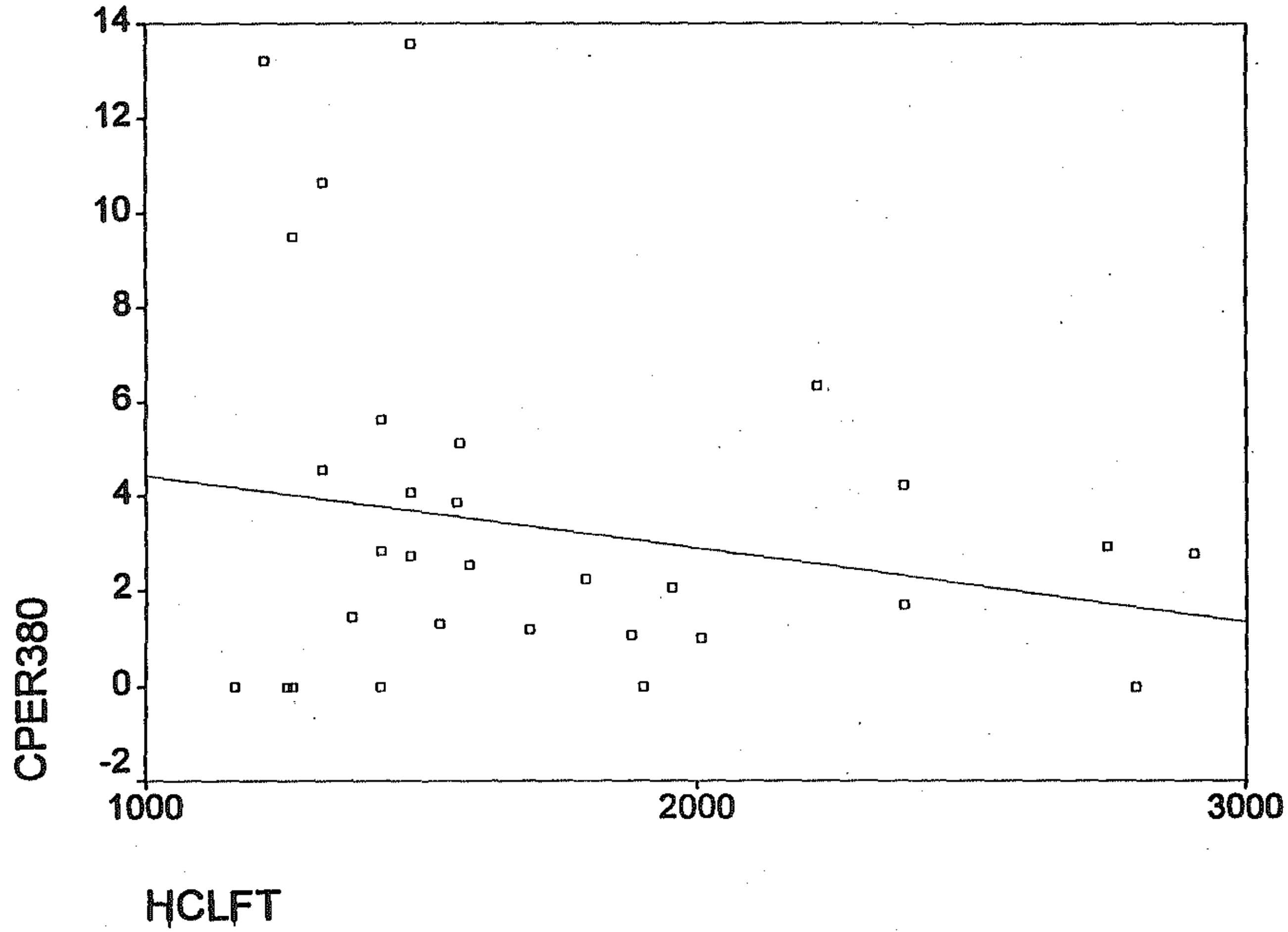


Figure 35 **Curve crash rate (Cper380), regression line for various values of curve length in feet, cluster 1**

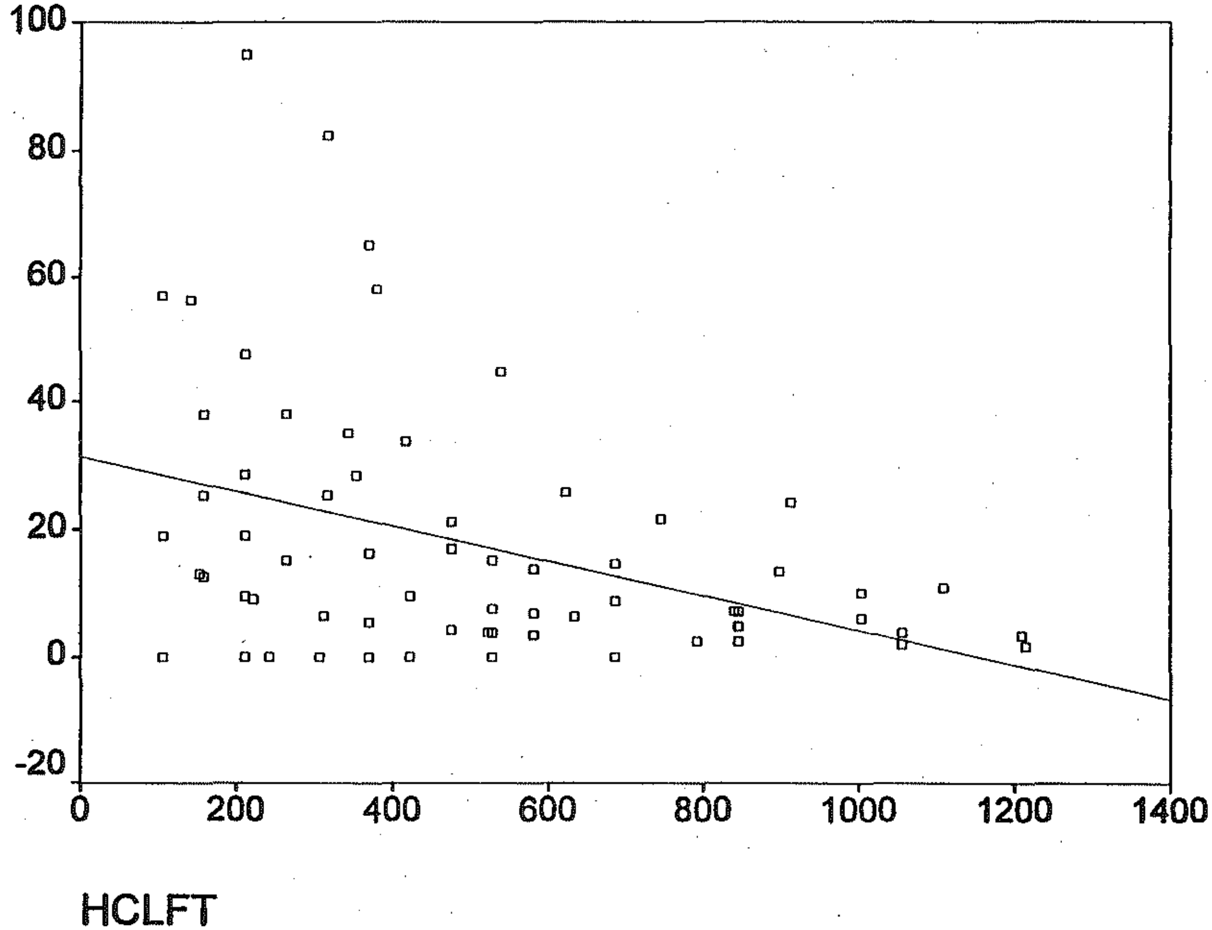


Figure 36 Curve crash rate (Cper380), regression line for various values of curve length in feet, cluster 2

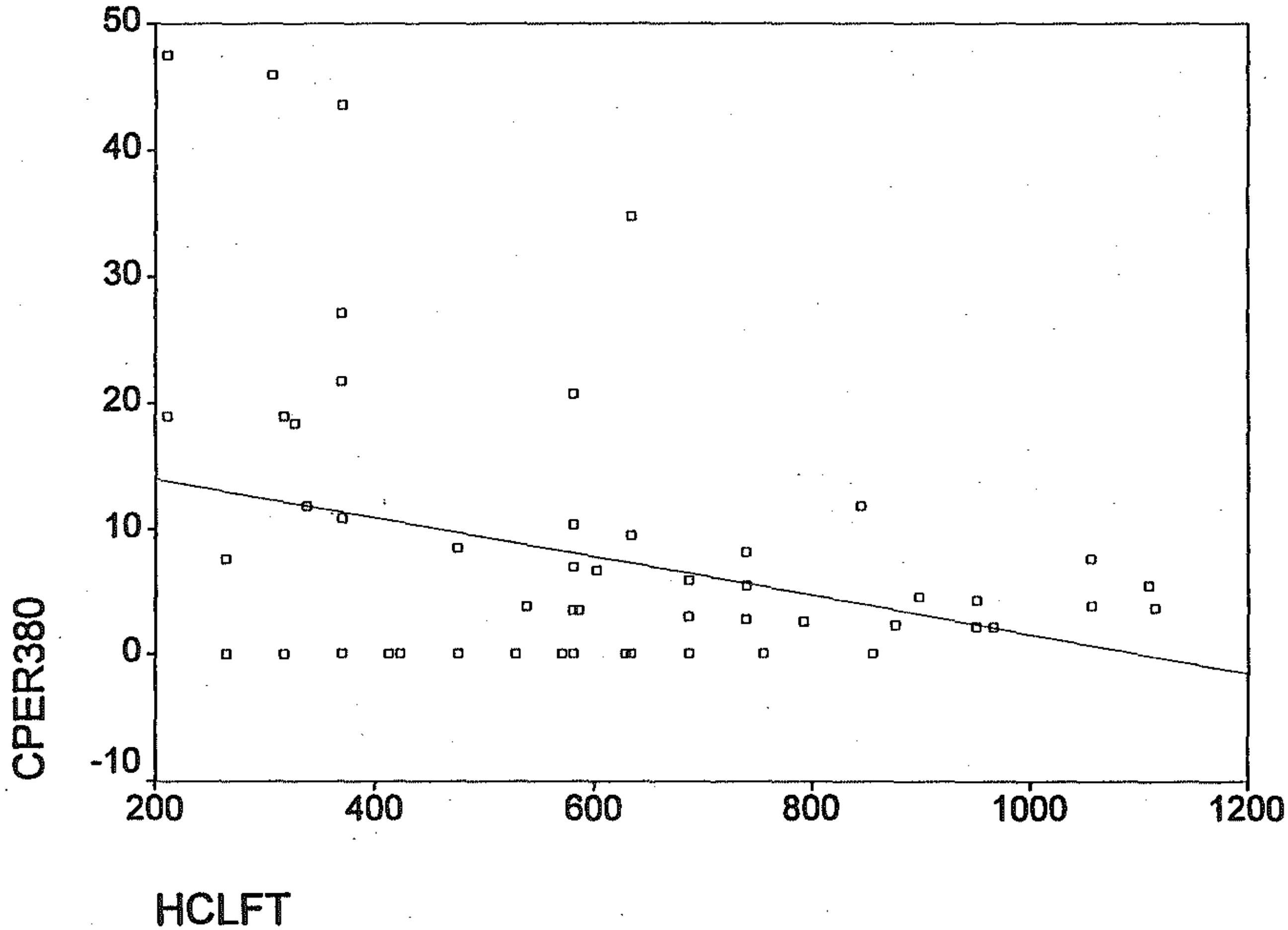


Figure 37 **Curve crash rate, (Cper380), regression line for various values of curve Length in Feet, cluster 3**

FACTOR ANALYSIS

DESCRIPTION:

While variables are the common method of describing statistical values, other concepts which are readily understood by individuals, (such as aggressiveness) may not be describable by variables. Often, the number of variables required to describe such a concept are numerous, with interdependencies and interrelations; and the variables included may even be seemingly contradictory.

Factor analysis is a technique used to reduce many variables into a smaller set of factors. Each factor describes a "concept". Ideally the concept will be readily understood by individuals and there may even be an existing name for the concept. If not, the analyst can often understand the concept and give it an appropriate name.

Factor analysis starts with a set of variables, or better stated, the scores related to a set of variables. Next, a set of new variables is constructed based on the interrelations exhibited in the data. The first factor is defined as the best linear combination of variables explaining the variance in the data as a whole. The other factors are similarly defined as the best linear combination of variables which explains the variance remaining in the data as a whole. As such, the first factor is more important than the second one and so on. The first few factors usually explain most of the variance in the data.

Analysis and Results:

Factor analysis was conducted for many cases of differing variables, factoring criteria, rotation method and number of extracted factors. However, the use of this technique did not add significanly to an understanding of the relationships among the variables and crash rates.

Table 16 shows the results of one factor analysis with the first three factors extracted. The variables that contribute the most to the factor score coefficients for the three factors are those shown in Table 17. Only one of the three factors includes the crash rate (Cper380). Factor 1 includes Cper380 and the presence of certain traffic control devices (chevron and advisory speed panels), curve length, radius, and roadside clearance (inversely). All of these variables, with the exception of the roadside clearance variable were also included in the discriminant analysis and cluster analysis results.

Factors 2 describes curves with high ADT and safe roadside, while Factor 3 describes curves with more hazardous roadside conditions and a lower ADT. This can be interpreted to indicate that the high volume State Trunkline roads have a safer roadside than do those trunkline highways with lower volumes. However, nothing is revealed about the difference in crash rates between these two combinations of variables.

Factor Score Coefficient Matrix

	Factor		
	1	2	3
ADT	.014	.507	.277
ALW	-.032	.066	-.089
ARROW	.033	-.109	.053
CHEVRON	.128	.032	.009
CLRNCW	-.176	-.289	.611
CPER380	.321	.018	-.019
CURVES	.018	-.004	.049
DLNTR	.012	-.022	.006
EDGLN	.013	.018	.012
GRAIL	-.012	-.013	.101
HCLFT	-.124	.041	.024
HCRFT	-.347	.105	-.038
MPHS	.218	.039	.071
OBSDST	-.004	.006	.006
PSL	.002	-.033	-.018
PSW	-.029	.097	-.031
SCT	.013	.181	-.024
TPER380	.035	.142	.117
TSW	-.004	.124	-.044
TURNS	.103	-.062	-.063

Factor Score Covariance Matrix

Factor	1	2	3
1	.736	1.850E-03	4.219E-02
2	1.850E-03	.766	4.699E-02
3	4.219E-02	4.699E-02	.768

Table 16 Factor score coefficient matrix for all the variables

Factor Score Coefficient Matrix

	Factor		
	1	2	3
ADT		.507	.277
ALW			
ARROW			
CHEVRON	.128		
CLRNCW	-.176	-.289	.611
CPER380	.321		
CURVES			
DLNTR			
EDGLN			
GRAIL			
HCLFT	-.124		
HCRFT	-.347		
MPHS	.218		
OBSDST			
PSL			
PSW			
SCT			
TPER380			
TSW			
TURNS			

Table 17 **Factor score coefficient matrix of relatively high values**

ANALYSES INCLUDING FIELD DATA:

The next set of analysis was performed using the subset of curves for which the field data, superelevation and drag factor, were collected. A total of 81 roadway segments containing 531 crashes, (279 in tangents and 252 in curves), were among those with the field data. Only 71 of the 81 roadway segments had crashes on their curved section. The values of these variables for the 81 roadway segments are shown in Figures 38-40.

Analyses similar to those performed previously for all the roadway segments, were conducted for only the roadway segments with the field data. The analyses were conducted with the addition of the two field data variables, superelevation and drag factor, for each direction of traffic individually and combined. The analysis was done twice, once for the higher values of the superelevation for the two directions, SPRELVN and again for their lower value, SELELO.

Figures 41 and 42 show graphs sorted by ascending value of Cper380 for those curves with the field data. Figures 43-48 show the simple linear regression results of Cper380 and C-T with these variables.

Neither the drag factor nor the superelevation, individually or in combination, showed any significance in explaining the curve crashes or assisting in the identification of curves to be modified.

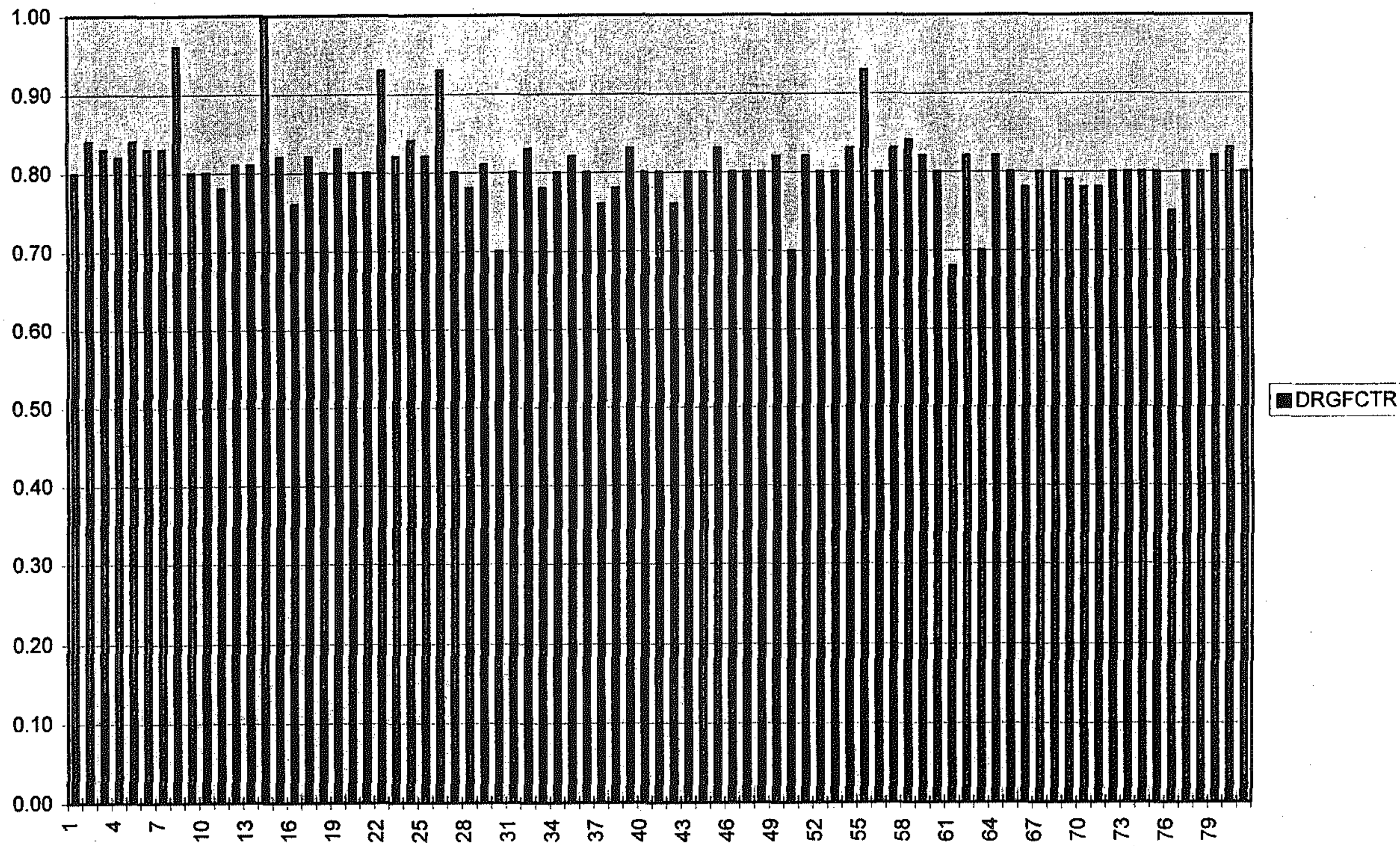


Figure 38 Drag factor (DRGFCTR), arranged in ascending order of curve crash rate (Cper380)

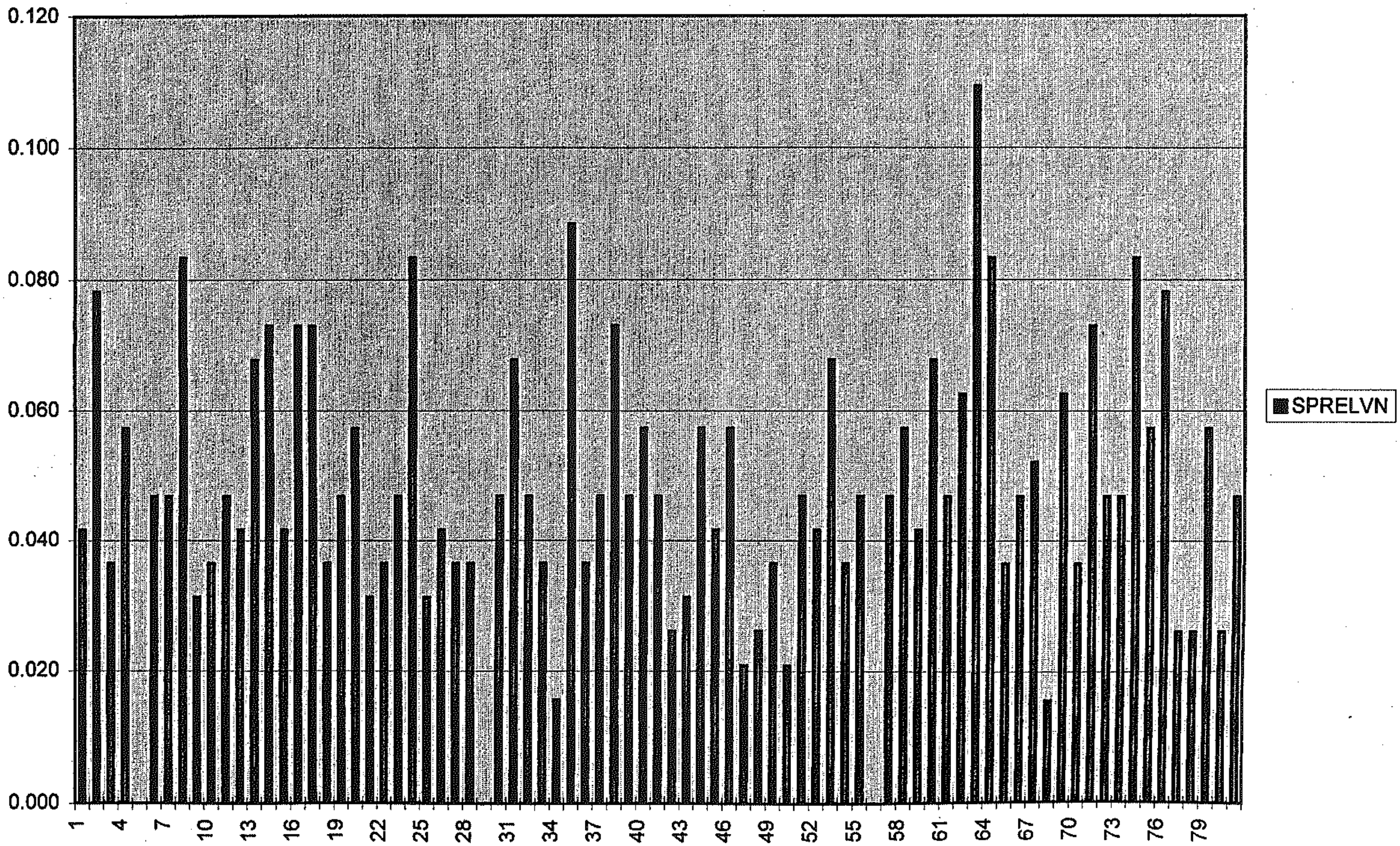


Figure 39 Superelevation high values (SPRELVN), arranged in ascending order of curve crash rate (Cper380)

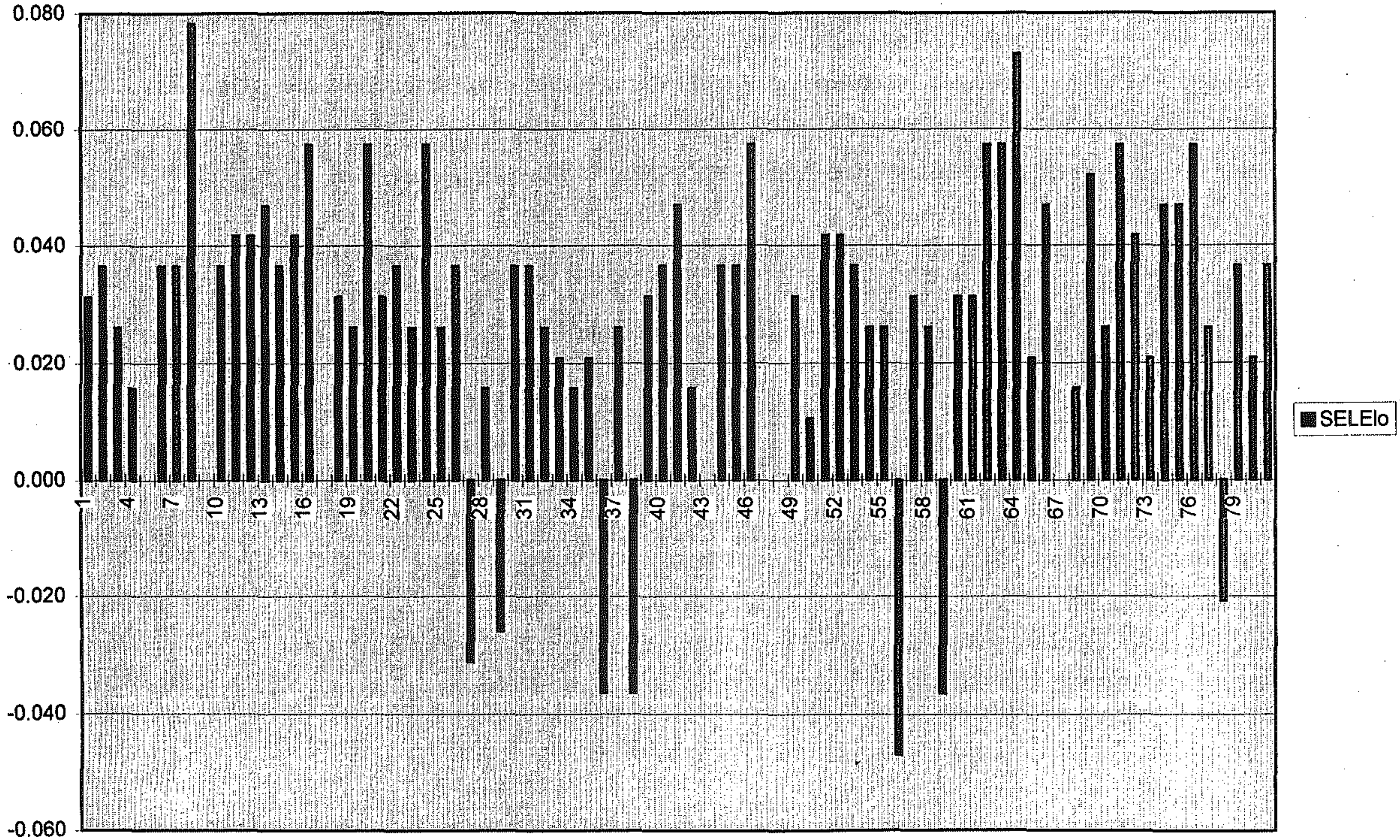


Figure 40 **Superelevation low values, (SELElo), arranged in ascending order of curve crash rate (Cper380)**

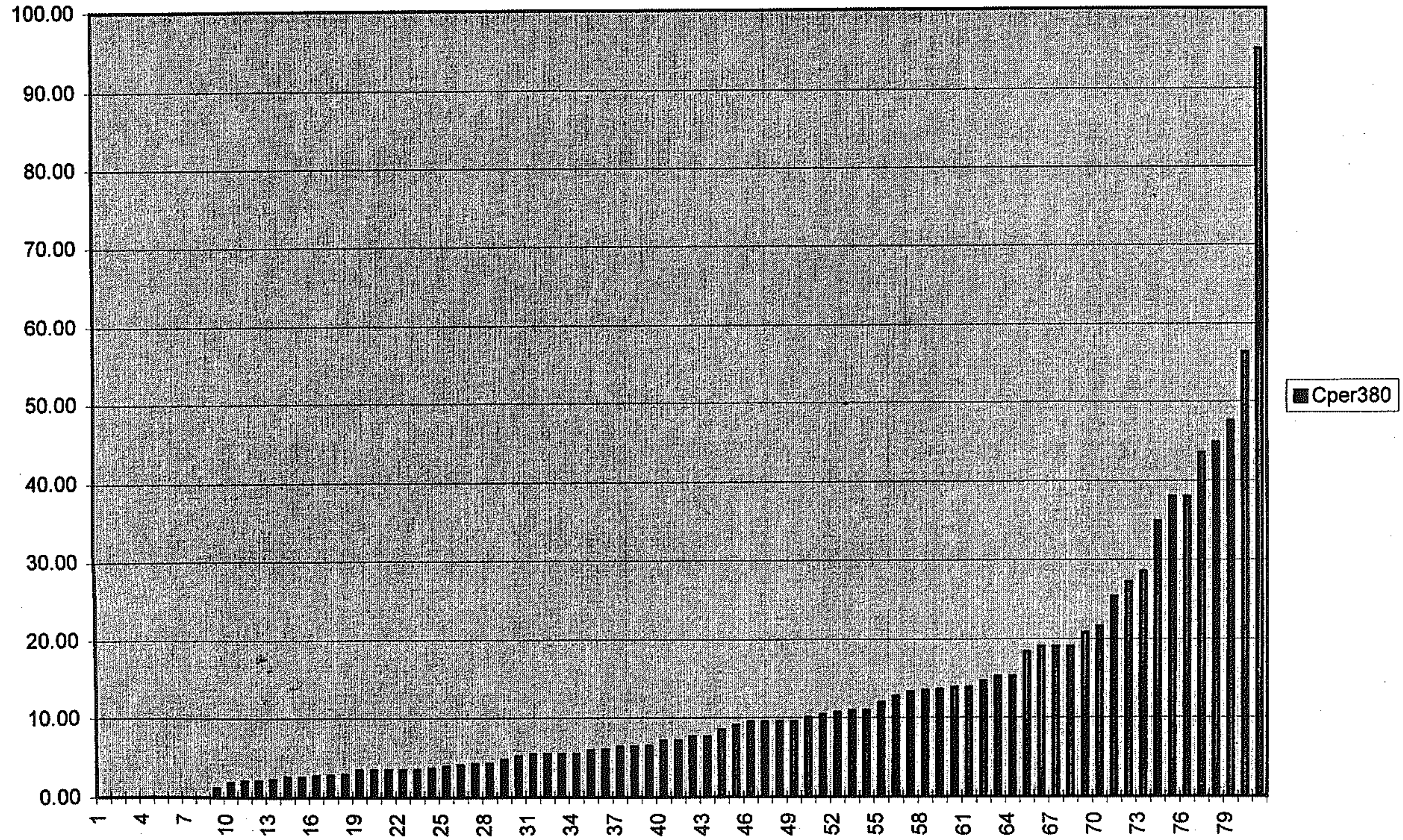


Figure 41 Curve crash rate (Cper380), arranged in ascending order

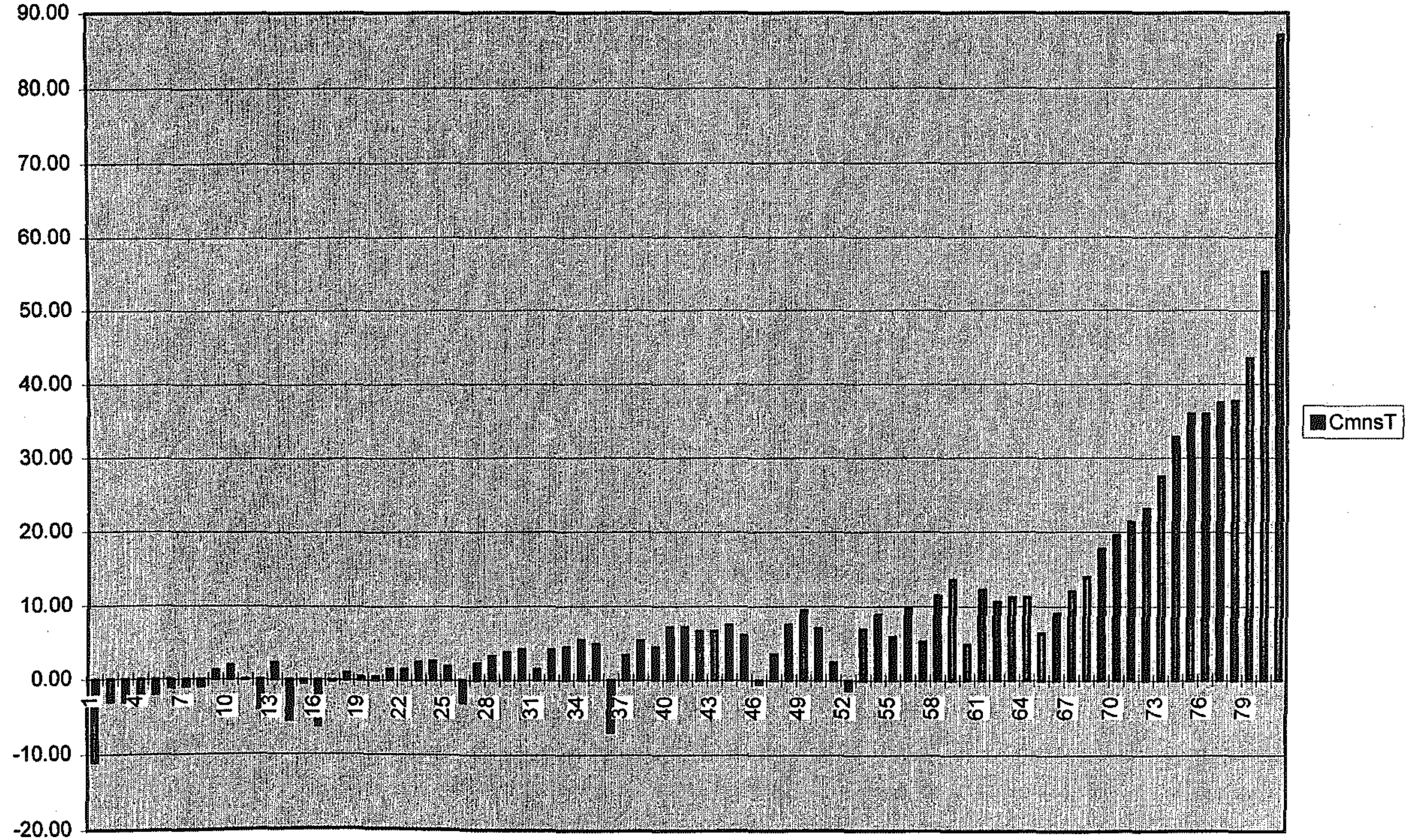


Figure 42 **Curve Crash Rate minus Tangent Crash Rate (C-T), arranged in ascending order of curve crash rate (Cper380)**

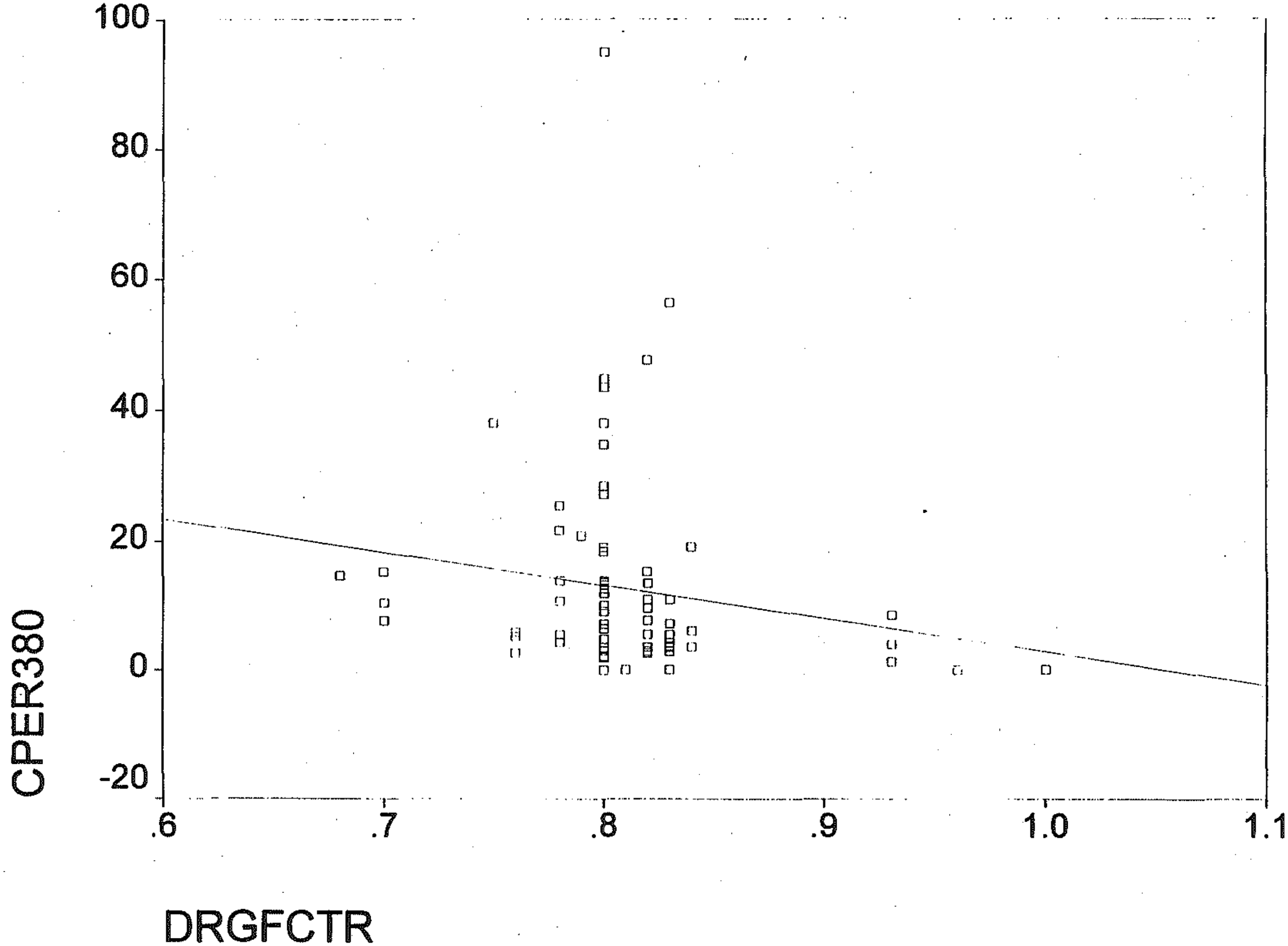


Figure 43 **Curve crash rate (Cper380), regression line for various values of drag factor (DRGFCTR)**

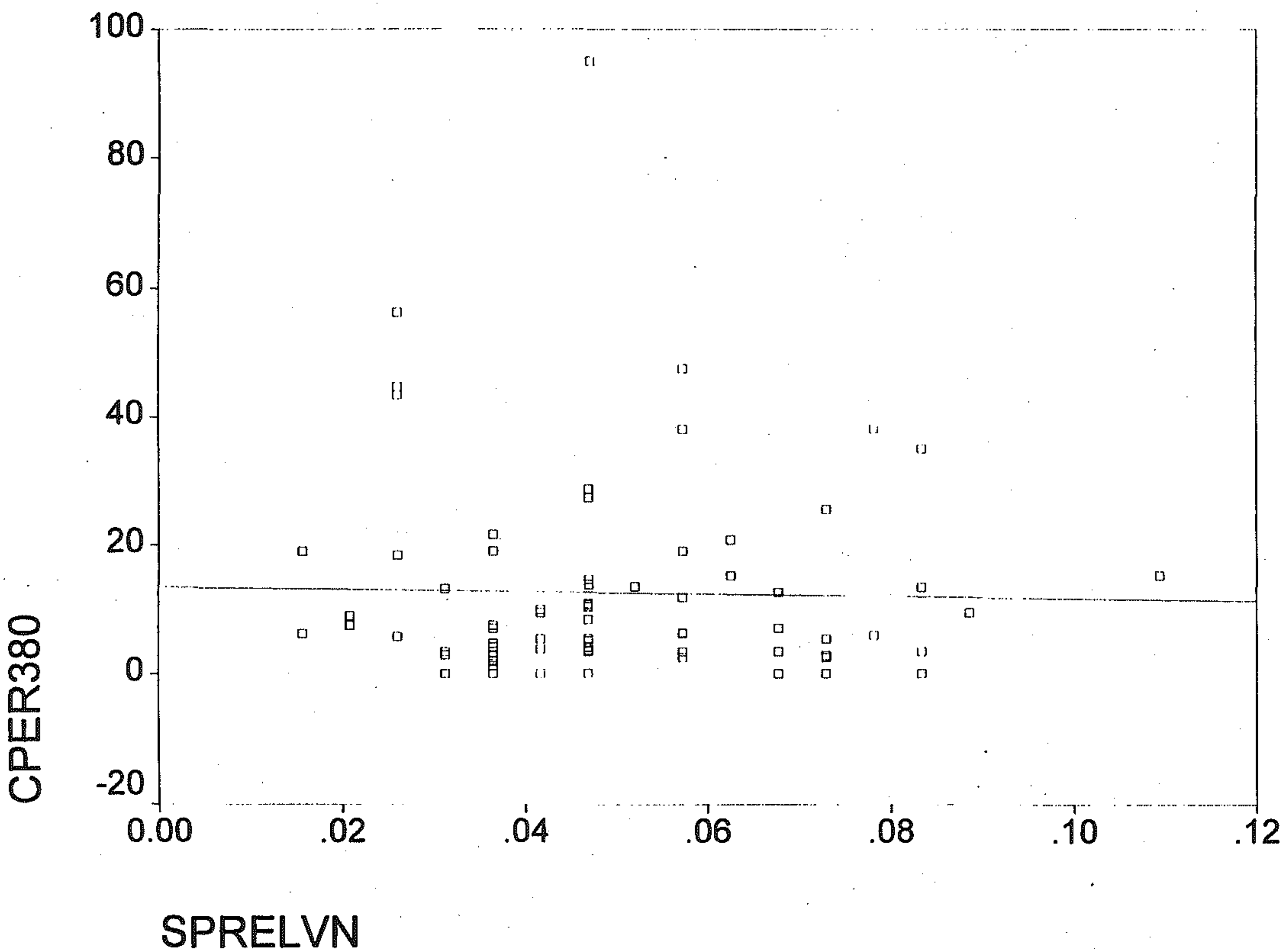


Figure 44 **Curve crash rate (Cper380), regression line for various values of superelevation high values (SPRELEV)**

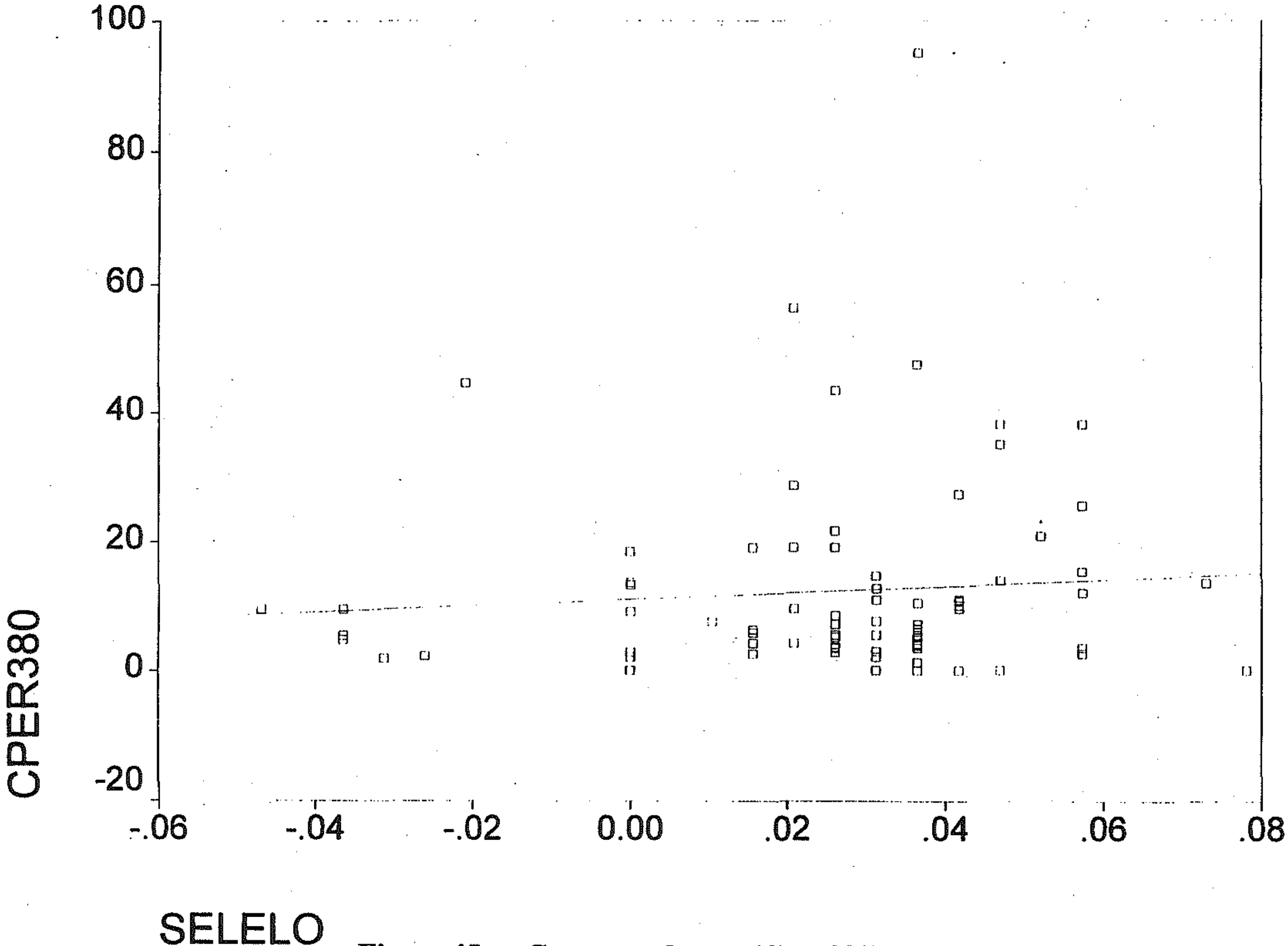


Figure 45 **Curve crash rate (Cper380), regression line for various values of superelevation low values (SELELO)**

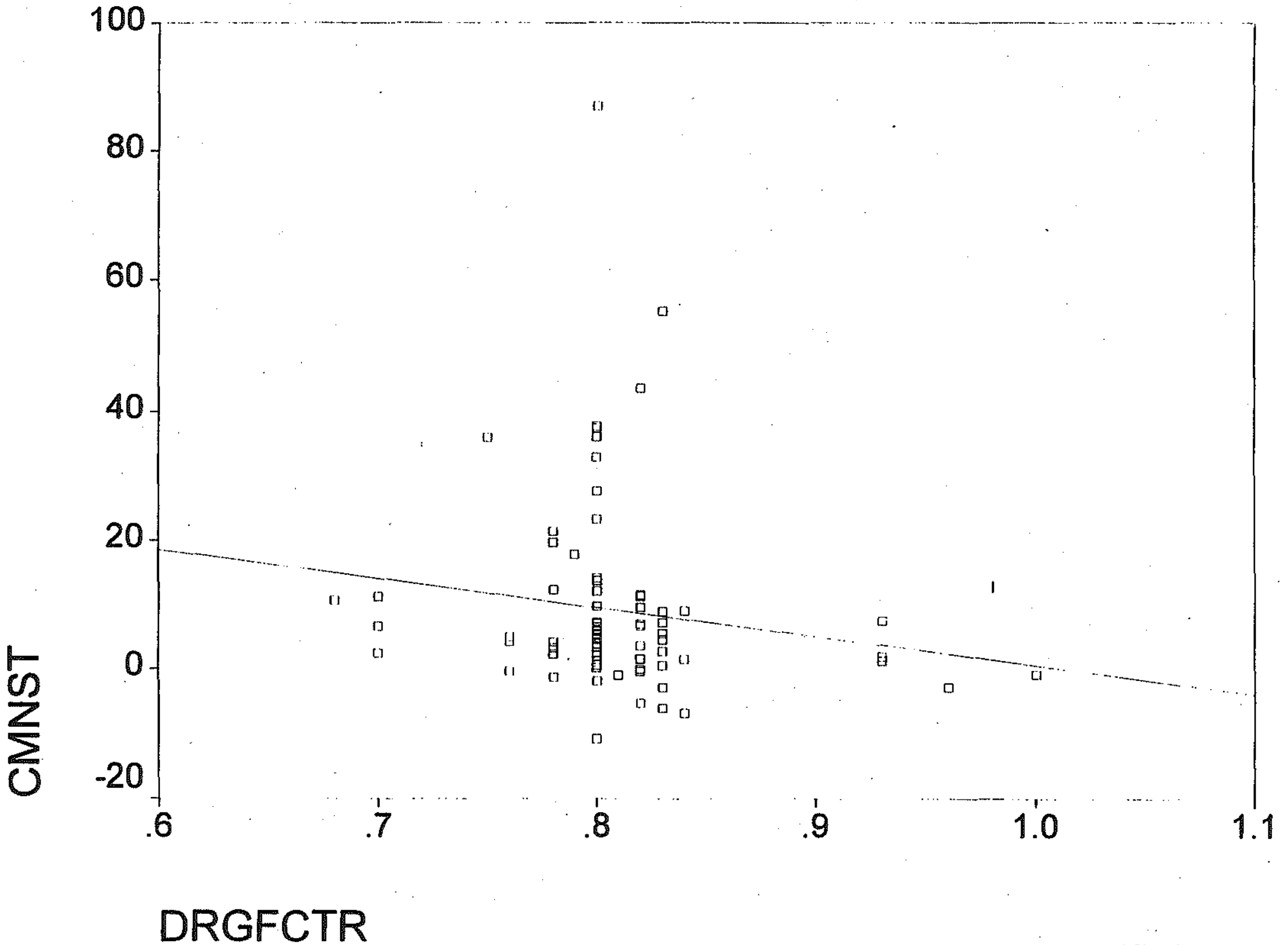


Figure 46 **Curve crash rate minus tangent crash rate (C-T), regression line for various values of drag factor (DRGFCTR)**

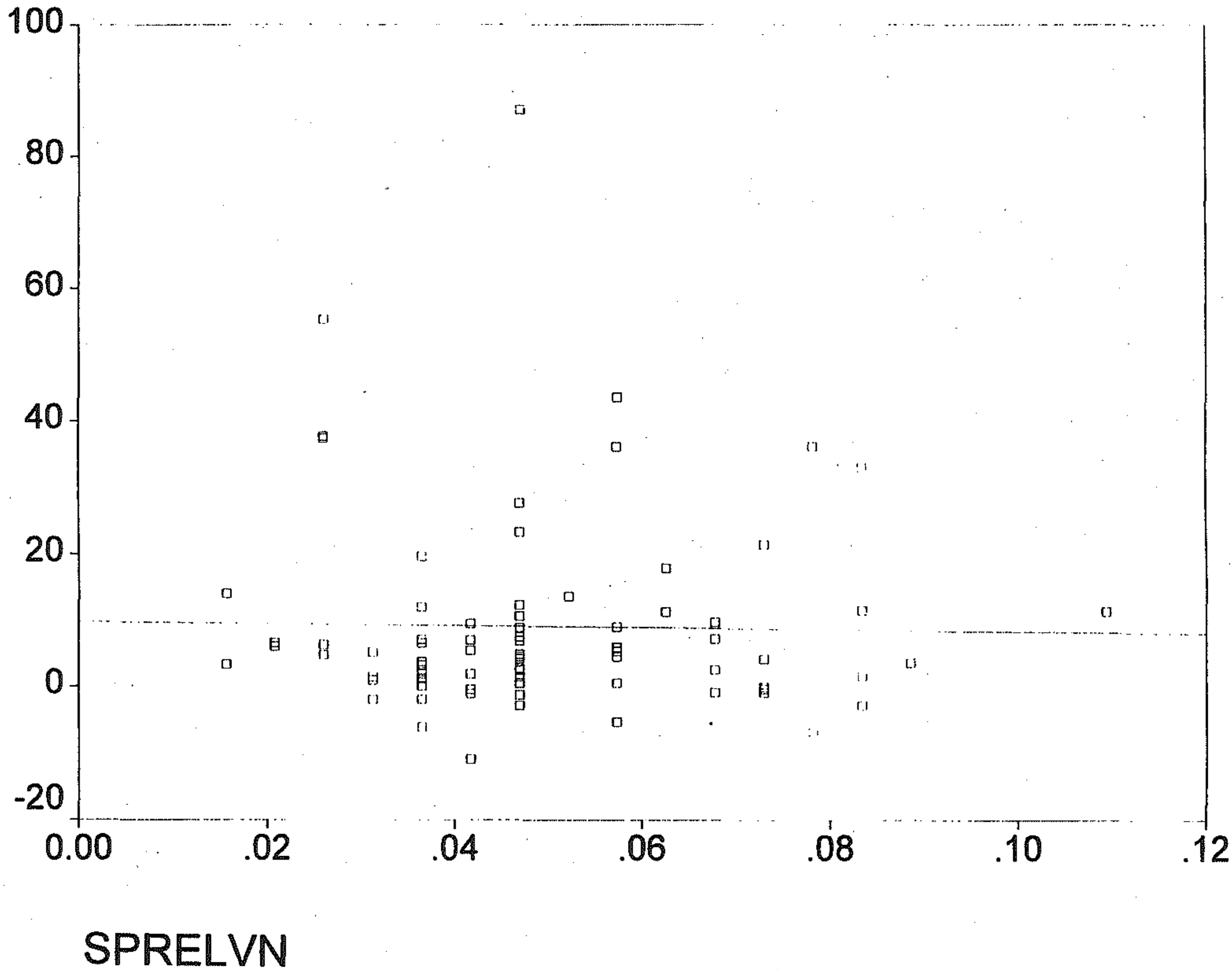


Figure 47 **Curve crash rate minus rangent crash rate (C-T), regression line for various values of superelevation high values (SPRELEV)**

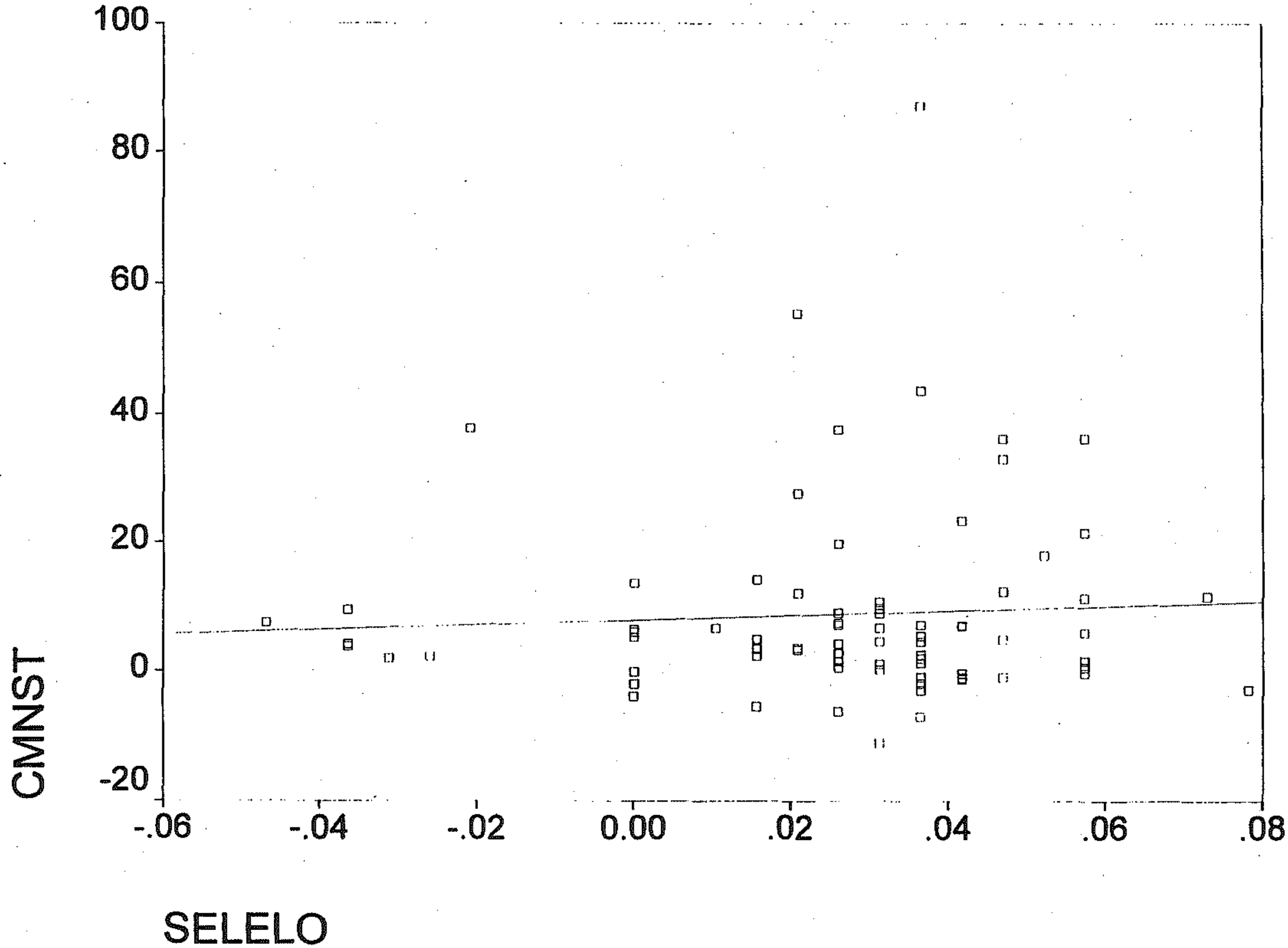


Figure 48 **Curve crash rate minus tangent crash rate (C-T), regression line for various values of superelevation low values (SELELO)**

DISCUSSION OF THE RESULTS:

Discriminant analysis provides information useful in meeting the objectives of this study. Specifically, it can be used to identify those characteristics of low crash rate curves which distinguish them from high crash rate curves. Having done this, it can be used to identify those curves with a high crash rate that should (based on their characteristics) have a low crash rate. These curves are the ones that should be studied for possible countermeasure implementation.

Using the discriminant analysis results from the modified Cper380 analysis, sixteen curves fell in this category. The crash rate on these curves ranged from 7.13 to 21.71 when they should have fallen in the group with a crash rate below 5.0. These curves are shown in Table 18, along with the value of some of the variables used in the analysis.

The significant characteristics of these curves include:

- Most do not have curve signs, target arrows and delineators
- There are no chevrons
- The observed sight distance is usually short
- The radius is relatively large
- The tangent crash rate is low

CRVno	CS	BMP	CTsign	CHEVRON	ARROW	DLNTR	OBSDSTW	HCLFT	HCRFT	Tper380	Cper380
136	45012	5540	0	0	1	1	10	845	1042	**0.00**	**7.13**
14	5051	7280	0	0	0	0	40	264	2865	**1.00**	**7.60**
72	24011	4377	1	0	0	0	23	1056	2292	**3.00**	**7.60**
200	73131	0	0	0	0	0	0	264	2865	**2.00**	**7.60**
3	2021	15020	0	0	0	0	40	739	1910	**1.00**	**8.14**
39	12021	490	0	0	0	0	70	739	2292	**3.00**	**8.14**
33	10011	5620	1	0	1	0	33	475	2865	**0.00**	**8.44**
82	28052	5530	0	0	1	0	40	475	1910	**1.00**	**8.44**
81	28052	4790	1	0	0	0	50	634	2865	**2.00**	**9.50**
94	31013	5810	0	0	0	1	30	370	1910	**3.00**	**10.86**
117	38071	7490	1	0	1	0	10	1214	2865	**8.00**	**13.22**
87	30062	1640	1	0	0	1	10	1478	2456	**0.00**	**13.57**
156	51011	50	0	0	0	0	50	581	1146	**1.58**	**13.82**
19	8011	8990	0	0	0	1	80	211	1763	**1.00**	**19.00**
193	67011	2130	0	0	1	1	40	475	1637	**4.00**	**21.11**
172	58032	4150	0	0	0	0	80	370	2644	**4.00**	**21.71**

Table 18 Curves with a high curve crash rate (Cper380)

Similar results were found when C-T was used as the grouping variable. This is consistent with the results above, since most of the misclassified curves had a low value of Tper380, they would fall in the high range of C-T values.

The results of the cluster analysis are consistent with prior studies, but they also provide additional information that may be useful in reducing traffic crashes. Low crash rates are clustered with curves with a large radius and long length. The average radius for curves in this group (based on modified Cper380) is 398 meters (1305 ft). The average length for the same curves is 274 meters (900 ft). These curves tend to have target arrows but no chevrons.

High crash rates are clustered with short, sharp curves as expected. These curves tend to have both chevrons and target arrows in place, but still tend to experience crashes because of their geometry.

The third cluster is the most difficult to explain, and possibly the group of curves where countermeasures may be most effective. These curves have a crash rate over twice as high as the low crash rate curves, even though they have approximately the same radius. The primary geometric difference is that they are very short curves, averaging 95 meters (312 ft). These curves generally do not have chevrons or target arrows in place.

Chevrons and target arrows are not intended for these types of curves according to the Michigan Manual of Uniform of Traffic Control Devices (MMUTCD), since they do not constitute a sharp change in alignment. However, based on the analysis, it may be appropriate to consider the use of these signs to increase the visibility of the curves.

This same clustering of curves into these groups are observed whether the crash rate variable was Cper380, Modified Cper380, C-T, or modified C-T. There were approximately 70 curves that belong to this cluster. Table 19 lists the curves for which both the Cper380 and C-T were significantly higher than the average for this cluster.

CRVno	CS	BMP	CTsign	CHEVRO	ARROW	DLNTR	BSDST	HCLFT	HCRFT	Tper380	Cper380	CmnsT
39	12021	490	0	0	0	0	70	739	2292	**3.00**	**8.14**	5.14
200	73131	0	0	0	0	0	0	264	2865	**2.00**	**7.60**	5.60
68	23051	2220	1	1	0	0	20	845	2083	**6.00**	**11.88**	5.88
177	61012	4910	0	0	0	0	48	327	2292	**12.00**	**18.39**	6.39
14	5051	7280	0	0	0	0	40	264	2865	**1.00**	**7.60**	6.60
4	2021	23640	0	0	0	1	70	581	2865	**0.00**	**6.91**	6.91
3	2021	15020	0	0	0	0	40	739	1910	**1.00**	**8.14**	7.14
82	28052	5530	0	0	1	0	40	475	1910	**1.00**	**8.44**	7.44
81	28052	4790	1	0	0	0	50	634	2865	**2.00**	**9.50**	7.50
94	31013	5810	0	0	0	1	30	370	1910	**3.00**	**10.86**	7.86
33	10011	5620	1	0	1	0	33	475	2865	**0.00**	**8.44**	8.44
12	5031	3900	1	0	0	0	60	370	2292	**2.00**	**10.86**	8.86
214	81031	750	1	0	0	0	10	317	2292	**10.00**	**19.00**	9.00
100	31051	9143	1	0	0	0	13	338	1910	**1.00**	**11.88**	10.88
172	58032	4150	0	0	0	0	80	370	2644	**4.00**	**21.71**	17.71
88	30062	2900	1	0	0	1	30	581	1719	**3.00**	**20.73**	17.73
19	8011	8990	0	0	0	1	80	211	1763	**1.00**	**19.00**	18.00
101	32011	3050	1	0	0	0	30	370	2292	**4.00**	**27.14**	23.14
62	22021	499	0	0	0	0	49	306	1879	**16.00**	**45.86**	29.86
140	45013	11700	1	0	1	0	60	634	1910	**2.00**	**34.83**	32.83
215	81031	1370	1	1	0	0	70	370	2989	**6.00**	**43.43**	37.43
71	23111	3670	1	0	0	0	30	211	1910	**3.00**	**47.50**	44.50

Table 19 Curves with a high curve minus tangent crash rate (C-T)

The curves categorized in each of the three clusters were then plotted in ascending order of the value of Cper380 to identify the outliers within each group. Figure 49 shows these values. It is clear that even within a cluster there is a significant range of values for the crash rate. These high crash rate curves are the ones that should be studied for possible countermeasure implementation. Table 20 lists these curves which have a crash rate equal to or greater than twice the average value of the cluster.

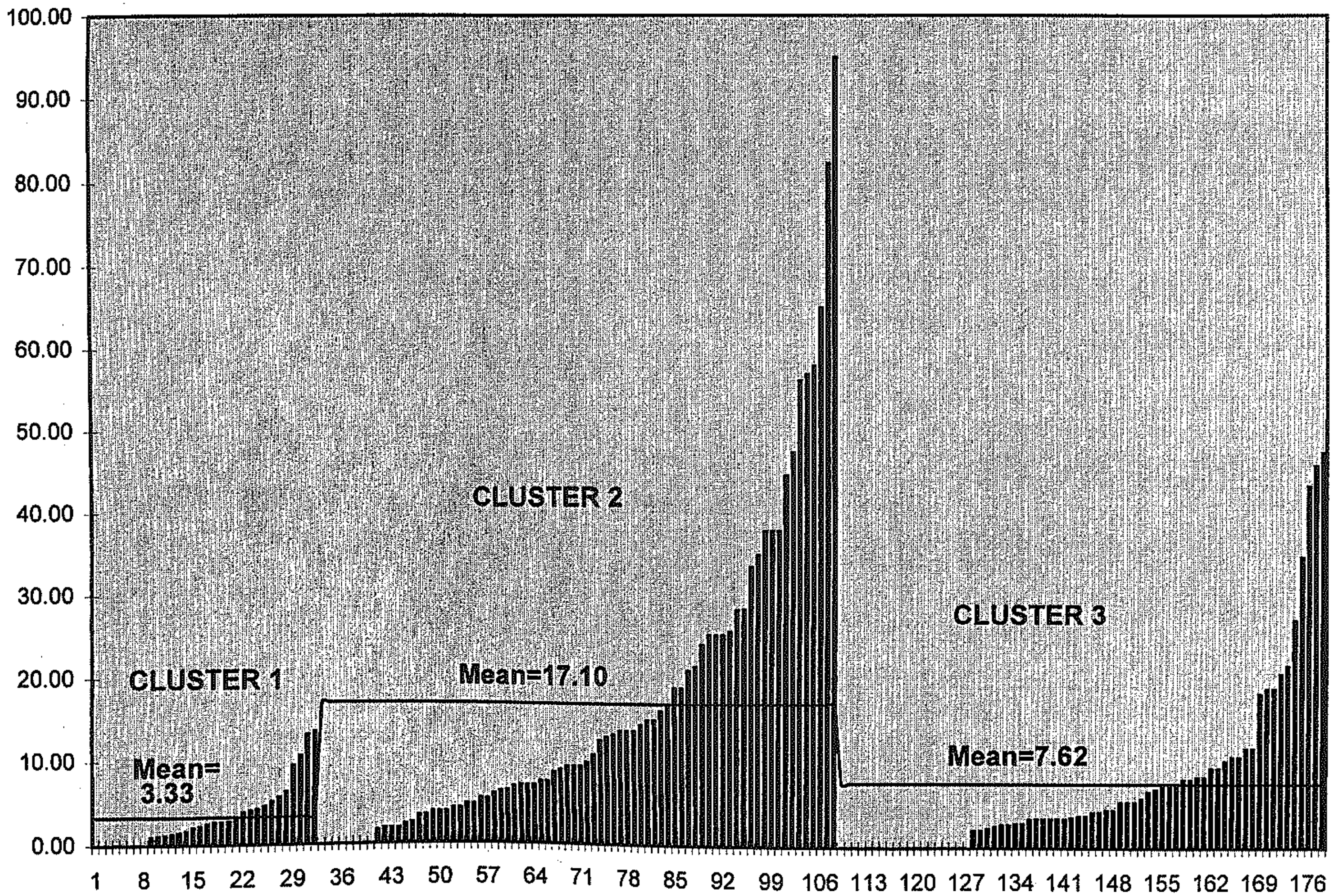


Figure 49 Curve crash rate (Cper380) for the three clusters, arranged in ascending order of Cper380 within each cluster

CRVno	CS	BMP	CTsign	CHEVRON	ARROW	DLNTR	OBSDSTW	HCLFT	HCRFT	Tper380	Cper380	C>2Mn
23	8031	2990	1	0	1	1	50	1267	1763	**6.00**	**9.50**	2.85
35	11052	14040	1	0	0	0	10	1320	2865	**12.00**	**10.64**	3.99
117	38071	7490	1	0	1	0	10	1214	2865	**8.00**	**13.22**	6.56
87	30062	1640	1	0	0	1	10	1478	2456	**0.00**	**13.57**	6.92
92	31012	4227	0	0	0	0	17	343	477	**4.00**	**35.08**	0.87
28	10011	7470	1	0	1	1	30	158	521	**2.00**	**38.00**	3.79
152	47041	21730	1	1	0	1	60	158	286	**2.00**	**38.00**	3.79
181	62031	3160	0	0	0	0	10	264	820	**7.00**	**38.00**	3.79
211	79081	8450	0	0	1	1	18	539	1008	**7.00**	**44.71**	10.50
18	8011	7100	1	0	0	0	30	211	229	**4.00**	**47.50**	13.29
196	72051	7673	0	0	0	0	10	143	1146	**1.00**	**56.30**	22.09
85	29042	6270	0	0	0	0	20	106	1146	**4.00**	**57.00**	22.79
168	56032	8814	1	0	0	0	34	380	1146	**4.00**	**58.06**	23.85
199	73061	3930	0	1	0	1	10	370	727	**6.00**	**65.14**	30.94
29	10011	8920	1	0	1	0	30	317	215	**3.00**	**82.33**	48.13
151	47041	19440	1	1	0	0	30	211	744	**8.00**	**95.00**	60.79
177	61012	4910	0	0	0	0	48	327	2292	**12.00**	**18.39**	3.14
19	8011	8990	0	0	0	1	80	211	1763	**1.00**	**19.00**	3.75
214	81031	750	1	0	0	0	10	317	2292	**10.00**	**19.00**	3.75
88	30062	2900	1	0	0	1	30	581	1719	**3.00**	**20.73**	5.48
172	58032	4150	0	0	0	0	80	370	2644	**4.00**	**21.71**	6.47
101	32011	3050	1	0	0	0	30	370	2292	**4.00**	**27.14**	11.90
140	45013	11700	1	0	1	0	60	634	1910	**2.00**	**34.83**	19.59
215	81031	1370	1	1	0	0	70	370	2989	**6.00**	**43.43**	28.18
62	22021	499	0	0	0	0	49	306	1879	**16.00**	**45.86**	30.62
71	23111	3670	1	0	0	0	30	211	1910	**3.00**	**47.50**	32.25

Table 20 Curves with a curve crash rate larger than twice the mean for their cluster

CONCLUSIONS

Based on the analyses conducted in this study, the following conclusions were reached.

1) The variation in the crash frequency or rate between horizontal curves with similar geometry is too large to be explained by regression techniques. The only studies that report high correlation coefficients are those that aggregate curves into groups with similar characteristics and then conduct the regression analysis on the group means. This information may be useful in the design of new highways, but it is not useful in meeting the objectives of this study.
2) The predicted crash rate using existing models (Zegeer and Glennon) does not accurately depict the actual crash rates on Michigan two-way, two-lane rural trunklines. These models can not be used to identify curves locations where countermeasures could successfully be deployed to reduce crashes.
3) The addition of data on the distance on the approach at which the curve first becomes visible to the motorist is not highly correlated with the crash rates as a single variable, but it was found to be a contributor to some of the models that use multiple variables.
4) The addition of data on superelevation and the drag factor also showed a low simple correlation with the crash rate and contributed little to multiple variable analyses.
5) Discriminant analysis techniques, using the variables collected for this study, can successfully distinguish the high crash rate curves from the low crash rate curves. This technique can be used to identify outliers in each of the two categories (high and

low) for both the absolute crash rate on curves (Cper380) or the difference in the crash rate between the curve and the tangent roadway segments (C-Tper380).

6) Cluster analysis identified three distinct groups of curves. The group with a high crash rate (Cper380) is characterized by short radii and short curve lengths. These curves generally are marked with curve sign, advisory speed panels and chevrons or delineators.

   The group with a low crash rate are characterized by large radii and long curve lengths.

   The third group, with an intermediate crash rate, are characterized by large radii but short curve lengths. These results are shown in Figure 50 and 51.

   The high crash rate on the first group of curves is probably related to constraint the geometry imposes on the driver ability to negotiate the curve at their approach speed. The intermediate crash rate curves may be related to the driver perception (or lack of perception) of the presence of a curve that does not require extraordinary driver input to negotiate safely.

7) The factor analysis results are more difficult to interpret, but do support the cluster analyses results. In general, the variables significant in defining the factor groups are the same as those used to distinguish the groups membership in cluster analysis.

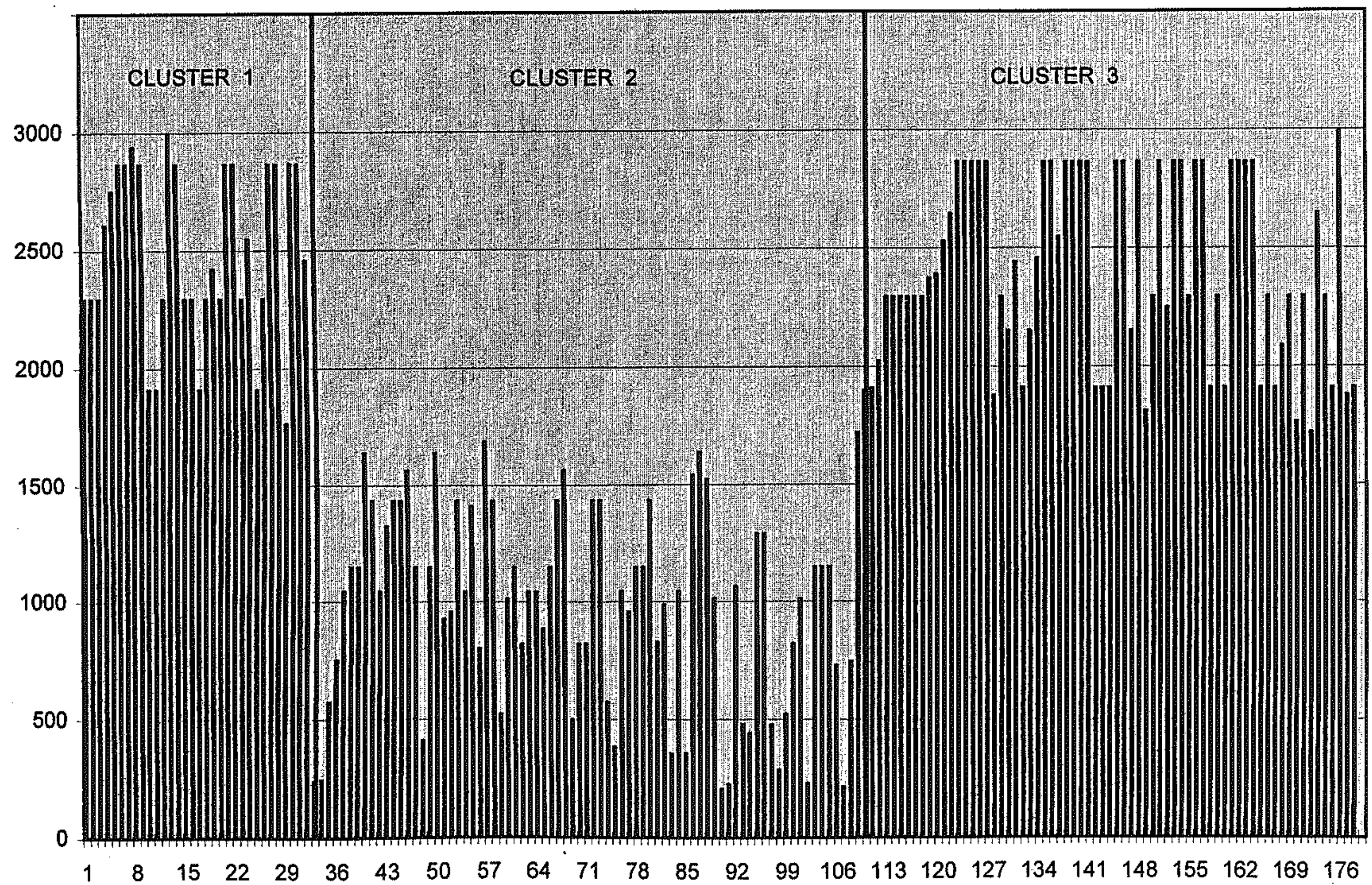


Figure 50 Curve radius in feet (HCRFT) for the three clusters, arranged in ascending order of Cper380 within each cluster

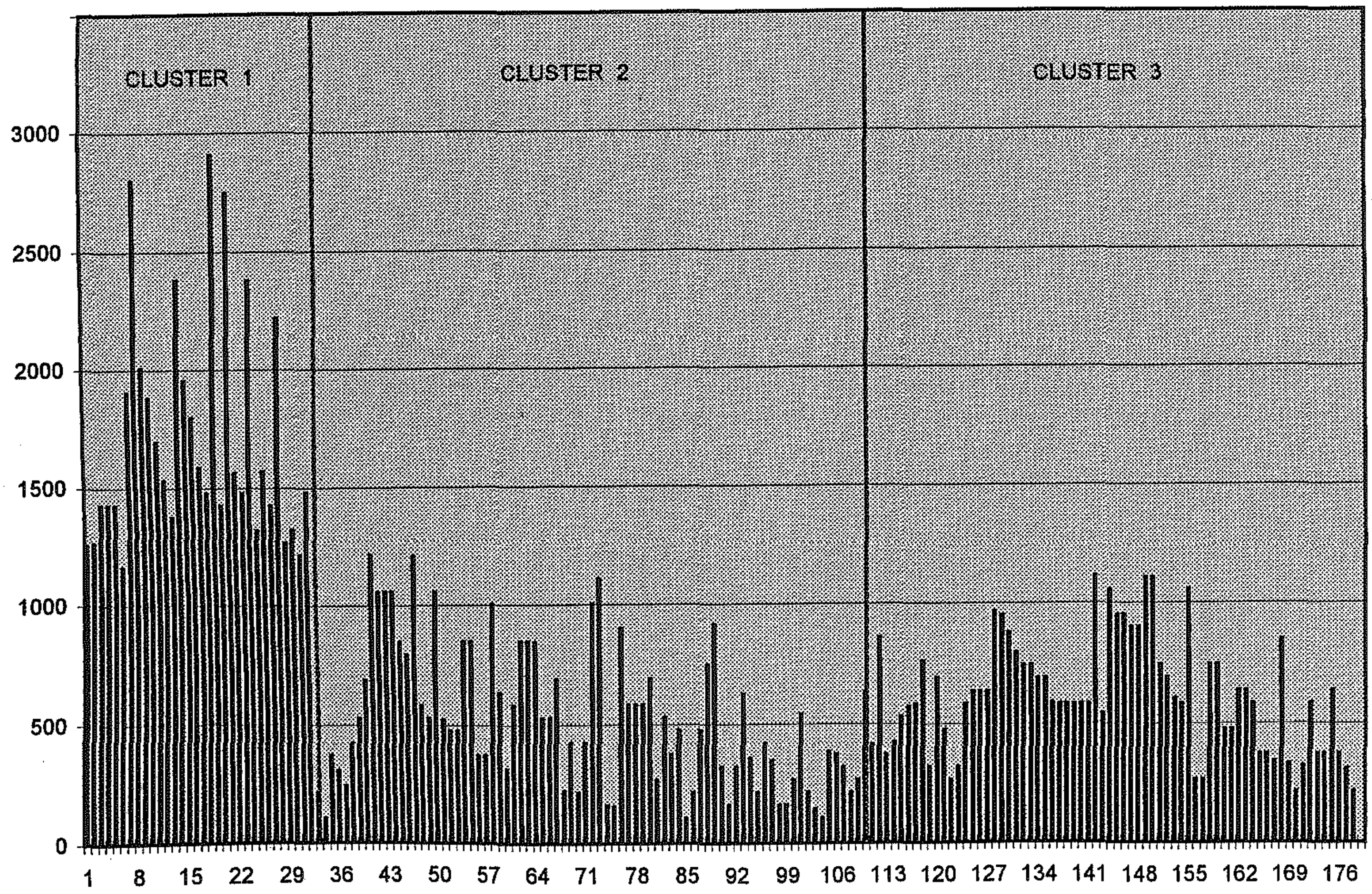


Figure 51 Curve length in feet (HCLFT) for the three clusters, arranged in ascending order of Cper380 within each cluster

RECOMMENDATIONS

1) The curves identified in Table 18 from the discriminant analysis results should be targeted for analysis and potential countermeasures implementation. These sixteen curves have the characteristics of low crash rate curves, but are experiencing a high rate of crashes.

2) The curves identified in Table 20 from the cluster analysis results should be targeted for analysis and potential countermeasure implementation. These curves have been identified as experiencing a crash rate at least twice that of the average crash rate for curves in their cluster.

3) Curves characterized by a large radius and short curve length should be analyzed to determine if there are inexpensive countermeasures that could be applied at these curves to reduce the crash rate. These curves have been identified from the cluster analysis as having an intermediate crash rate which is not explained by the curve geometry. The curves from this group with both a high crash rate and a large difference in the curve crash rate compared to the tangent crash rate are shown in Table 19.

4) Discriminant analysis and cluster analysis techniques should be used to analyze other sets of curves on state trunkline highways. These techniques have been useful in identifying specific curves that are candidates for countermeasures. It should be determined whether these techniques are equally valid for:

a) curves that are not screened for approach tangents and intersections.

b) curves on four-lane cross sections.

5) If recommendations 1, 2, and 3 are adopted, a careful before and after study should be designed to document any change in the crash rate resulting from implementation of the selected countermeasures.

6) If resources are available in the Department of Transportation, these analyses could be conducted internally. Alternatively, these analyses could form the basis of a study for the Michigan State University's Center of Excellence.

REFERENCES:

1. Roy Jorgenson Associates, Inc., MD, "Cost and Safety Effectiveness of Highway Design Elements", NCHRP Report 197, 1978.

2. Glennon, J.C., Neuman, T.r., Leisch, J.E. "Safety and Operational Considerations For Design of Rural Highway Curve", Jack E. Leisch & Associates, Evanston, IL, FHWA-RD-86/035, 1985.

3. Zegeer, C.V., Hummer, J., Reinfurt, D., Herf, L. and Hunter, W., "Safety ffects of cross-section Design for two-lane Roads, Volume I," Report No. FHWA/RD-87/008, FHWA, Washington DC, December 1986.

4. Zegeer, C.V., and Deacon, J.A., "Effects of Lane Width, Shoulder Width, and Shoulder Type on Highway Safety".

5. Neuman, T.R., Zegeer, C., and Slack, K.L. "Design Risk Analysis Volume II - User's Guide", FHWA Contract No. DTFH 71-88-C-00011, U.S. Department of Transportation, April 1991.

6. McGee, H.W., Hughes, W.E., Daily, K., Bellomo-McGee, Inc., "Effect of Highway Standards on Safety", Final Report, prepared for National Cooperative Highway Research Program, Transportation Research Board, National Research Council, Vienna, Virginia.

7. Kach, Brett, Benac, Jack D., "Horizontal Curve Crash Model Assessment", Traffic and Safety Division, Michigan Department of Transportation, June 1992.

8. Fink, K.F., Krammes, R.A., "Tangent Length and Sight Distance Effects of Crash

Rates at Horizontal Curves on Rural two-lane Highways", The 74th Annual Meeting, Transportation Research Board, January 1995

9. Zegeer, C.V., Twomey, J.M., Heckman, M.L. and Heyward, J.C., "Safety Effectiveness of Highway Design Features", Volume II, Alignment. Report No. FHWA-RD-91-045. FHWA,U.S. Department of Transportation, 1992

10. Datta T.K. et al., "Crash Surrogates for Use in Analyzing Highway Safety Hazards". Report No. FHWA/RD-82/103. Washington, D.C., 1983.

11. Terhune, K. W. and Parker, M.R., "An Evaluation of Crash Surrogates for Safety Analysis of Rural Highways". Volume 2. Technical Report. Report No. FHWA/RD 86/128. Federal Highway Administration, Washington, D.C., 1986.

12. Matthews, L.r. and Barnes, J.W., "Relation Between Road Environment and Curve Crashes". Proceedings, 14th Australian Road Research Board Conference, Vol. 14, Part 4, Victoria, Australia, 1988.

13. Zegeer, C., Stewart, R., Reinfurt, D. et al., "Cost-Effective Geometric Improvements for Safety Upgrading of Horizontal Curves". Report No. FHWA-RD-90-021. Federal Highway Administration, Washington, D. D., 1991.

14. Glennon, J.C., Neuman T.R. and Leisch, J.P. "Safety and Operational Considerations for Design of Rural Highway Curves". Report No, FHWA/RD-86/035. FHWA, U.S. Department of Transportation, 1985.

15. Fambro, D.B., Urbanik, T. II, Hinshaw, W.M., Hanks Jr., J.W.,
Ross, M.S., Tan, C.H. and Pretorius, C.J. "Stopping Sight Distance Considerations at Crest Vertical Curves on Rural two-lane Highways in Texas", (TX-90/1125-1F). Texas Transportation Institute, College Station, Texas, 1989.

16. Baker, Stannard J., Northwestern University Traffic Institute, The Traffic-Accident Investigation Manual, At-Scene Investigation and Technical Follow-Up, Evanston, IL

17. Berry, William D., Feldmam, Stanley, University of Kentucky. Multiple Regression in Practice, Sage Publications, Newberry Park London New Delhi, 1990

18. Bhatti, M. Ishaq, School of International Business Relations, Testing Regression Models Based on Sample Survey Data, Avebury, Aldershot Brookfield USA Hong Kong Singapore Sydney, 1995

19. Hand, D.J., Biometrics Unit, London University Institute of Psychiatry, Discrimination and Classification, John Wiley and sons, Cichester New York Brisbane Toronto, 1981

20. NATO Advanced Study Institute, Edited by T. Cacoullos, University of Athens, Athens, Greece. Discriminant Analysis and Applications, Academic Press, New York and London, 1973.

21. Jambu, M., Centre National d' Etudes des Te'le'communications, Issy-les-Molineaux, France and Lebeaux, M-O., Centre National de la Recherche Scientifique, Centre d' Etudes Sosiologiques, Paris, France. Cluster Analysis and Data Analysis, North-Holland Publishing Company. Amsterdam New York Oxford, 1983

22. Tryon, Robert Choate, Associate Professor of Psychology, University of California. Cluster Analysis, correlation profile and Orthometric (Factor) Analysis. Edward Brothers, Inc., Ann Arbor, Michigan, 1939.