

# Safety Effects of Widening Shoulders on Rural Multilane Roads

## Developing Crash Modification Functions with Multivariate Adaptive Regression Splines

Juneyoung Park and Mohamed Abdel-Aty

The presented study used crash modification (CM) factors and CM functions to assess the safety effectiveness of widening shoulders. The study first evaluated CM factors for widening shoulders on rural multilane roadways with the empirical Bayes (EB) observational before–after method to check the overall safety effects. Second, the CM factors were calculated for each aggregated break point according to roadway characteristics such as the original shoulder widths of treated sites in the before period and the actual increased widths. Finally, CM functions were developed with multiple linear regression and multivariate adaptive regression splines (MARS) models to determine the variation of CM factors. The results indicated that widening the shoulders resulted in a 12% crash reduction for total crashes (all severities), 18% for injury crashes, and 21% for severe crashes on rural multilane roadways for all types of crashes. For run-off-roadway crashes, the overall crash reductions of the treatment were 25% for total crashes, 28% for injury crashes, and 31% for severe crashes. Moreover, the results showed that the safety effects varied across sites, depending on roadway characteristics. In particular, the original shoulder widths of treated sites in the before period and the actual increased width were significant parameters that affected the variation of CM factors. Moreover, it was found that the MARS models showed more reliable estimates than normal regression if the variation of CM factors with specific parameters had nonlinear relationships and interaction effects. The proposed CM functions provide general insights into roadway design and selection of sites for increasing shoulder width for reducing crashes.

Part D of the *Highway Safety Manual* (HSM) presents a variety of crash modification (CM) factors for safety treatments on roadway segments and at intersections (1). A CM factor can be used to compute the expected number of crashes after a given countermeasure is implemented at a specific site. CM factors have been estimated with observational before–after studies and with cross-sectional methods. The before–after method evaluates the safety effects of treatments by comparing crash frequencies during the periods before and after implementation of the treatments (2). The cross-sectional method is used when the before–after study cannot be used

because of the following conditions: (a) the date of the treatment installation is unknown, (b) data for the period before the treatment installation are not available, and (c) the effects of other factors on crash frequency must be controlled for with a crash modification (CM) function (3, 4). Cross-sectional methods are also known as safety performance functions (SPFs) or crash prediction models.

Although the HSM provides many estimates of safety effects (i.e., CM factors) of roadway treatments, the effect is given as a fixed single factor. Because this single value represents an average safety effect of the treatment for all treated sites, the heterogeneous effects of roadway characteristics on CM factors at treated sites are ignored. To overcome this limitation, CM functions have been developed to predict the variation in CM factors according to the site characteristics. Elvik proposed an approach for estimating CM functions (5) for the same or similar treatments by means of meta-regression analysis (6). By using this framework, Elvik estimated CM functions for the installation of a bypass and converting signalized intersections to roundabouts according to population changes (7). The results showed that CM factors increase with population for both treatments. However, to develop good CM functions, large amounts of data are needed. Elvik also identified the variation of safety effectiveness of horizontal curves with the lengths of their radii (8). He estimated the summary CM function to assess the international transferability of national CM functions from 10 countries. The results showed that the estimated CM function can be a summary of these national functions. It was found that the safety effects decrease as the radius of the curve decreases and that the variations in safety performance appear to be the same for all 10 countries.

To reflect the nonlinear relationship between CM factors and roadway characteristics, Elvik applied five nonlinear functions to develop CM functions for speed enforcement (9). The CM function illustrates the effect of speed enforcement on injury accidents as a function of the relative change in the level of speed enforcement. It was found that higher levels of enforcement are associated with a reduction in crash frequencies. The nonlinear logarithmic function best fitted the data points from 13 previous studies, but the inverse function also fitted the data well. Similarly to this study, Park et al. applied one linear and four nonlinear functions to develop CM functions for adding shoulder rumble strips, widening shoulders, and a combination to check the relationship between CM factors and original shoulder width of treated sites (10). It was found that the exponential nonlinear regression was the best fitted function to develop CM functions. However, other traffic and roadway characteristics [e.g., annual average daily traffic (AADT), actual increased width]

Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, FL 32816-2450. Corresponding author: J. Park, jypark@knights.ucf.edu.

*Transportation Research Record: Journal of the Transportation Research Board*, No. 2583, Transportation Research Board, Washington, D.C., 2016, pp. 34–41. DOI: 10.3141/2583-05

were not considered in developing CM functions. To consider the variation of CM factors over time, Park et al. used nonlinearizing link functions in developing CM functions (11). Similarly, Sacchi et al. developed CM functions to incorporate the effectiveness of a safety treatment (signal head upgrade program) with changes over time by using Poisson lognormal linear intervention and nonlinear intervention models (12). Sacchi and Sayed estimated CM functions that accounted for AADT changes among treated sites and time trends by using the same data for evaluation of the safety effectiveness of the signal head upgrade program (13). Although they identified the variation of safety effects with both AADT changes and time trends, only the single roadway characteristic was used to develop the CM function. Park et al. applied one linear and four nonlinear functions to develop CM functions for adding a bike lane on urban arterials (14). The results of CM functions show that inverse, quadratic, and exponential nonlinear regression models were the best fitted functions for various single roadway characteristics. It was found that CM functions with multiple parameters show better model fit than simple models. The study also showed that the multiple regression models with backward and stepwise subset selections were the best fitted for various multiple roadway characteristics. However, there have been few studies on variation in the safety effects of widening shoulders among treated sites with different original shoulder width and actual increased width.

Thus, the objective of this study was to identify the relationship between the safety effects of widening shoulders and roadway characteristics through (a) evaluation of CM factors with the use of the empirical Bayes (EB) observational before–after method and (b) development of CM functions based on site characteristics and actual increased width to reflect variation of the safety effects. Crash severities were categorized according to the KABCO scale as follows: K = fatal, A = incapacitating injury, B = nonincapacitating injury, C = possible injury, and O = property damage only.

## SAFETY EFFECTS OF SHOULDER WIDTH

Roadside elements are one of the most important cross-sectional roadway characteristics for safety. In particular, widening shoulders is one of the roadside safety countermeasures known to reduce crashes. Hadi et al. developed negative binomial (NB) models to analyze safety effects of shoulder width on rural multilane highways (15). The study showed that a small reduction in crashes (1% to 3%) could be attained if the unpaved shoulder is widened by 1 ft. Jovanis and Gross estimated safety effects of shoulder width by using case control and cohort methods (16). The results of the two methods showed that crashes decrease as the shoulder width increases. Moreover, Harkey et al. developed CM factors for shoulder width on rural multilane roadways with more than 2,000 vehicles per day (17). They found that increasing shoulder width reduces crashes. Zeng and Schrock evaluated the safety effects of 10 shoulder design types in winter and nonwinter periods (18). They developed CM factors by using cross-sectional methods. The results showed that wider and upgraded shoulders have significantly less impact on safety in winter than in nonwinter periods. Park et al. assessed the safety effects of shoulder rumble strips and widening shoulders on rural multilane roadways by using the observational before–after method with comparison group and EB and a cross-sectional study (10). The study showed that two single treatments and combination are generally safety effective for total and single-vehicle run-off-road (ROR) crashes. The results also showed that for roadway segments with shoulders 9 ft or wider, a

single treatment can produce better safety effects than two treatments. The results for total crashes (KABCO) showed that shoulder rumble strips are more effective in reducing crashes for roadway segments with shoulder widths of less than 7 ft, whereas increasing shoulder width is more effective for roadway segments with shoulder widths of 7 ft or more. Similar results were found for the safety effects of shoulder rumble strips and widening shoulders on rural two-lane roadways (19).

However, according to Stamatiadis et al., wider shoulders may encourage higher operating speeds because they may communicate to the driver the presence of wider space for correcting errors (20). Some studies have explained associated interaction effects between lane width and shoulder width according to differences in local conditions. Gross et al. reported that the effects of lane width on crash frequency were neither consistently positive nor consistently negative because of the variation in shoulder widths (21). Thus, they suggested that determination of CM factors consider the interaction between lane width and shoulder width. Potts et al. also recommended that narrowing lane width be used as a treatment according to local roadway characteristics such as shoulder width because the effect of lane width varies with location (22). However, the variation in the safety effects of shoulder width based on the original shoulder width and the actual increased width has not been investigated.

## DATA PREPARATION

In this study, more detailed roadway information and additional treated locations were obtained in addition to a previously used data set (10). Three sets of data maintained by the Florida Department of Transportation were used in this study: roadway characteristic inventory (RCI) data for 8 years (2004 to 2011), financial project information, and the crash analysis reporting system (CARS) database. The RCI database provides current and historical roadway characteristics data and reflects features of specific segments for the selected dates. Treated sites were identified from the financial project information and the RCI data set. All segments that were treated between the end of 2006 and the beginning of 2009 were selected for analysis to ensure sufficient sample size. There were no other geometric changes between 2006 and 2009 for the identified treated sites except shoulder width and AADT, respectively. Crash records were collected from CARS for 2 years (2004 and 2005) for the before period and 2 years (2010 and 2011) for the after period. Crash records for 2006 and 2009 were not included in the analysis to account for several data issues (e.g., initial period to prepare roadway construction, finalizing period of construction, stable time for drivers to get used to the new roadway conditions). A total of 241 treated roadway segments, 185.822 mi long, and 1,796 reference sites, 881.882 mi long, were identified. Distributions of each variable among these treated segments are summarized in Tables 1 and 2.

## METHODOLOGY

### Safety Performance Functions

Generally, an SPF relates the crash frequency to traffic and roadway characteristics. The NB model is most commonly used to develop an SPF since the function can account for overdispersion. Two types of SPFs, the full SPF and the simple SPF, have been used in the literature. The full SPF relates the frequency of crashes to both traffic and

TABLE 1 Descriptive Statistics of Treated Segments

Variable	Crash Frequency in Before Period				Crash Frequency in After Period			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
All KABCO crashes	4.037	6.773	0	57	3.249	5.148	0	33
All KABC crashes	2.398	3.850	0	24	1.680	2.750	0	19
All KAB crashes	1.506	2.467	0	13	0.942	1.687	0	11
ROR KABCO crashes	0.950	2.041	0	22	0.622	1.487	0	12
ROR KABC crashes	0.577	1.253	0	10	0.344	0.881	0	7
ROR KAB crashes	0.407	0.909	0	6	0.203	0.581	0	5

roadway characteristics, whereas the simple SPF considers a traffic parameter only, such as AADT, as an explanatory variable. The HSM provides CM factors based on the simple SPF only. However, the simple SPF is an oversimplified function for reflecting the relationship between crash frequency and roadway characteristics since crash frequency is affected by not only traffic volume (14). Thus, in this study, the full SPF was used for calculating CM factors in the EB before–after method. The functional form of SPF for fitting the NB regression models is shown in Equation 1:

$$N_{\text{predicted},i} = \exp(\beta_0 + \beta_1 \ln(\text{AADT}_i) + \dots + \beta_k (X_{ki})) \quad (1)$$

where

$$\begin{aligned} N_{\text{predicted},i} &= \text{predicted crash frequency on segment } i, \\ \beta_k &= \text{coefficients for variable } k, \\ \text{AADT}_i &= \text{AADT of segment } i \text{ (vehicles per day), and} \\ X_{ki} &= \text{roadway characteristic } k \text{ of segment } i \text{ (} k > 2 \text{)}. \end{aligned}$$

In this study, six full SPFs were developed with the NB model for combinations of crash type and severity levels using 2-year before and 2-year after crash data. The SPFs were developed for reference sites of rural multilane roadways in Florida shown in Table 3. In general, the results of the six full SPFs show that crash frequency is higher for the roadway segments that have higher AADT and are longer. The results also show that the crash frequency is lower for roadways with wider medians and lower speed limits. For all (KABCO) crashes, the results indicate that an increase in lane width

can increase crash frequency. It has been known that wider lanes are effective for reducing crashes (23–25). However, some studies found an opposite effect (4, 19, 21, 26). In particular, Hauer suggested that an increase in separation of vehicles on wider lanes tends to increase vehicle speeds and reduce spacing between vehicles (27). Consequently, an increase in lane width may instead increase crash frequency. To account for the trend of crash frequency based on time changes, a binary variable (i.e., before period) was included to represent the 2-year before period. The model with a categorical variable for each year was assessed, but it was not statistically significant. The results indicate that the crash frequency in the after period is lower than in the before period for both all crashes and ROR crashes, and this trend is consistent with the declining trend of traffic crashes in an 8-year period (2004 to 2011) in the United States (28). Since this decline trend on crashes could affect the evaluation of safety effects of treatment, it is better to capture the time changes in the SPFs to account for the trend of crash frequency in the EB analysis.

### Before and After with EB Method

In the EB before–after method, the expected crash frequencies at the treatment sites in the after period had the countermeasures not been implemented are estimated more precisely through use of data from the crash history of a treated site, as well as what is known about the safety of reference sites with a similar yearly traffic trend, physical characteristics, and land use. The EB before–after method, which is based on research by Hauer (29), is a well-accepted approach to evaluating safety effects of treatments because of its statistical strength. The safety effectiveness of a treatment is calculated by comparing the observed number of crashes and the expected number of crashes in the after period. One of the main advantages of the before–after study with EB is that it accurately accounts for changes in crash frequencies in the before and after periods at the treatment sites that may be related to regression-to-the-mean bias. The detailed procedure of the EB method is available in other work (29, 30). In the EB before–after method, the expected number of crashes without treatment ( $N_{\text{expected},B}$ ) can be estimated with Equation 2:

$$N_{\text{expected},B} = w(N_{\text{predicted},B}) + (1-w)(N_{\text{observed},B}) \quad (2)$$

where

$$\begin{aligned} N_{\text{expected},B} &= \text{expected crash frequency in the before period,} \\ N_{\text{predicted},B} &= \text{predicted crash frequency estimated with the SPF in the before period,} \\ N_{\text{observed},B} &= \text{observed crash frequency in the before period, and} \end{aligned}$$

TABLE 2 Variables Related to Traffic and Roadway Geometric Characteristics

Variable	Mean	SD	Min.	Max.
AADT (vpd) in before period	20,548.02	13,491.79	4,200	60,500
AADT (vpd) in after period	20,272.82	12,987.71	4,100	51,500
Length (mi)	0.771	1.000	0.1	4.634
Lane width (ft)	11.975	0.156	11	12
Median width (ft)	46.232	18.718	10	130
Max. speed limit (mph)	59.274	9.519	40	70

NOTE: vpd = vehicles per day. Number of lanes: 4 lanes = 226 sites; 6 lanes = 17 sites. Original shoulder width: 2–4 ft = 8 sites; 5–6 ft = 9 sites; 7–8 ft = 39 sites; 9–10 ft = 75 sites; 11–12 ft = 110 sites. Actual increased width: 1 ft = 50 sites; 2 ft = 32 sites; 3 ft = 35 sites; 4 ft = 15 sites; 5 ft = 20 sites; 6 ft = 69 sites; 7–8 ft = 15 sites; 9–10 ft = 5 sites.

TABLE 3 Florida Specific Calibrated SPFs for Rural Multilane Roadways by Crash Type and Severity Level

Crash Type	Estimated Coefficient ( <i>p</i> -value)									
	Constant	ln (AADT)	Length	Before Period (2004–2005)	Max. Speed Limit	Median Width	Lane Width	Dispersion	Deviance	AIC
All (KABCO)	-13.9082 (<.0001)	1.3072 (<.0001)	1.0244 (<.0001)	0.0718 (.1445)	NS	-0.0047 (.0011)	0.0953 (.0535)	1.4801	3,507.5	13,191.2
All (KABC)	-14.2983 (<.0001)	1.3374 (<.0001)	1.0163 (<.0001)	0.1122 (.0344)	0.0125 (.0029)	-0.0053 (.0038)	NS	1.3581	3,166.6	10,000.7
All (KAB)	-13.3037 (<.0001)	1.1501 (<.0001)	1.0093 (<.0001)	0.1755 (.0027)	0.0184 (<.0001)	-0.0058 (.0054)	NS	1.1965	2,802.8	7,443.2
ROR (KABCO)	-11.8034 (<.0001)	0.8311 (<.0001)	0.8701 (<.0001)	0.1459 (.0888)	0.0299 (<.0001)	NS	NS	1.5529	1,857.8	3,952.5
ROR (KABC)	-12.2116 (<.0001)	0.7835 (<.0001)	0.8644 (<.0001)	0.1734 (.0992)	0.0357 (<.0001)	NS	NS	1.3286	1,431.5	2,681.4
ROR (KAB)	-11.6202 (<.0001)	0.6718 (<.0001)	0.8292 (<.0001)	0.2513 (.0428)	0.0419 (<.0001)	-0.0079 (.0937)	NS	1.0601	1,167.6	1,988.2

NOTE: NS = not significant; AIC = Akaike information criterion.

$w$  = weight factor estimated with overdispersion parameter from NB model (SPF) and predicted crash frequency in the before period for the treated site, as shown in Equation 3.

$$w = \frac{1}{1 + k \times N_{\text{predicted},B}} \quad (3)$$

where  $k$  is an overdispersion parameter.

The expected crash frequency in the after period can be calculated as a product of the expected crash frequency in the before period and the ratio of the predicted crash frequency in the after period to the predicted crash frequency in the before period. According to Persaud and Lyon (31) and Gross et al. (2), the CMF can be estimated as shown in Equation 4.

$$\text{CMF} = \frac{\left[ \frac{N_{\text{observed},A}}{N_{\text{expected},A}} \right]}{\left[ 1 + \frac{\text{var}(N_{\text{expected},A})}{N_{\text{expected},A}^2} \right]} \quad (4)$$

$$\text{var}(N_{\text{expected},A}) = N_{\text{expected},A} \times \frac{N_{\text{predicted},A}}{N_{\text{predicted},B}} \times (1 - w)$$

where

$N_{\text{expected},A}$  = expected crash frequency in the after period,  
 $N_{\text{observed},A}$  = observed crash frequency in the after period,  
 $N_{\text{predicted},A}$  = predicted crash frequency estimated with the SPF in the after period, and

$\text{var}(N_{\text{expected},A})$  = variance of expected crash frequency in the after period.

The standard error of the overall safety effectiveness can be calculated by taking the square root of the variance of CMF. The variance of CMF can be calculated with Equation 5.

$$\text{var}(\text{CMF}) = \frac{\text{CMF}^2 \left[ \left( \frac{1}{N_{\text{observed},A}} \right) + \left( \frac{\text{var}(N_{\text{expected},A})}{N_{\text{expected},B}^2} \right) \right]}{\left[ 1 + \frac{\text{var}(N_{\text{expected},A})}{N_{\text{expected},A}^2} \right]^2} \quad (5)$$

### Multivariate Adaptive Regression Splines

The multivariate adaptive regression splines (MARS) analysis can be used to model complex relationships with a series of basis functions (32). According to Abraham et al., MARS is a multivariate piecewise regression technique, and the splines can represent the space of predictors broken into number of regions (33). Piecewise regression, also known as segmented regression, is a useful method when the independent variables, clustered into groups, exhibit different relationships between the variables in these groups (34). The independent variable is partitioned into intervals, and a separate line segment is fit to each interval. MARS divides the space of predictors into multiple knots (i.e., the boundary between regions) and then fits a spline functions between these knots (32). The MARS model is defined as shown in Equation 6 (35). A log form of the MARS model was fitted to develop CM functions in this study.

$$\hat{y} = \exp \left( b_0 + \sum_{m=1}^M b_m B_m(x) \right) \quad (6)$$

where

$\hat{y}$  = predicted response variable,  
 $b_0$  = coefficient of the constant basis function,  
 $b_m$  = coefficient of the  $m$ th basis function,  
 $M$  = number of nonconstant basis functions, and  
 $B_m(x)$  =  $m$ th basis function.

There are three main steps to fit a MARS model (35–37). The first step is a constructive phase, in which basis functions are introduced in several regions of the predictors through use of a forward stepwise selection procedure. The predictor and the knot location that contribute significantly to the model are searched and selected in an iterative way in this step. Also, the introduction of an interaction is checked to improve the model at each iteration. The second step (pruning phase) performs a backward deletion procedure to eliminate the least-contributing basis functions. A generalized cross-validation (GCV) criterion is used in this pruning step to find the best model. The criterion can be estimated with Equation 7. The last step, which is the selection phase, selects the optimum MARS model from a group of recommended models according to the fitting results of each.

$$GCV(M) = \frac{1}{n} \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\left(1 - \frac{C(M)}{n}\right)^2}$$

$$C(M) = M + dM \quad (7)$$

where

$y_i$  = response for observation  $i$ ,

$n$  = number of observations,

$C(M)$  = complexity penalty function, and

$d$  = defined cost for each basis function optimization.

## RESULTS AND DISCUSSION

### Estimation of CM factors with EB Method

Table 4 presents the CM factors estimated with the EB observational before–after analysis method. In general, the safety effects of widening the shoulders were positive for both all crashes and ROR crashes. CM factors for ROR crashes were lower than CM factors for all crashes. These results indicate that widening the shoulder is more effective in reducing ROR crashes than all crashes. Moreover, it was found that safety effects are greater for crashes that are more severe.

To identify changes of CM factors on the basis of site characteristics, the safety effects of widening the shoulder were calculated for the treated sites with different original shoulder widths and actual widened widths. The results show that the safety effects are greater for roadway segments with a narrow original shoulder width (i.e., 2 to 8 ft) for both all crashes and ROR crashes. The results also show that the safety effects of widening the shoulder were greater as actual width increased. Thus, it can be concluded that the safety effects vary depending on the original shoulder widths and actual widened widths among treated sites. Some CM factors were not significant at a 90% confidence level. Although these CM factors may not represent reliable safety effects of treatments statistically, insignificant CM factors could be used to check the general impact of treatments with relatively large variation. The HSM suggests that a standard error of 0.1 or less indicates that the CMF value is sufficiently accurate, precise, and stable. For treatments that have CM factors with a standard error of 0.1 or less, related CM factors with standard errors of 0.2 to 0.3 may be included and considered to account for the effects of the same treatment on other facilities, other crash types, or other severities (*1*).

### Development of CM Functions

The CM functions were developed to determine the variation of CM factors with different site characteristics among treated segments, as shown in Tables 5 and 6. Because of the low frequency of all KABCO crashes and ROR crashes, the CM functions were evaluated for all KABCO crashes and all KABC crashes only. A total of 241 roadway segments with the same roadway characteristics and roadway ID were grouped into 24 data points according to the most important variables (different original shoulder width, actual increased width, AADT, median width, and maximum speed limit) so there would be a reasonable number of individual segments for each data point (average 10 segments for each point). Moreover, as suggested by Sacchi and Sayed (*13*) and Park and Abdel-Aty (*19*), the log form of the models was used to ensure that the CMF value from CM function cannot be a negative estimate. The CM functions were developed with multiple linear regression and MARS models. In this study, the ADAPTIVEREG procedure in the SAS program was used to fit a MARS model (*38*). It was found that there is little difference between selecting the default condition (two-way maximum interactions) and increasing the maximum number of interactions (e.g., three-way or four-way) in the analysis. Although increasing model complexity by adding more interactions could improve the predictive power for highly structured data, the applicability of the model could be decreased. Thus, two-way maximum order of interactions was used consistently for the various crash severities in this study. Moreover, the basis functions were constructed for each severity level since the rate of changes can vary within the range for different severities. Park and Abdel-Aty recommended use of a MARS model to examine the nonlinearity and interaction impacts between variables (*37*).

Overall, the results show that the CM factors increase as original shoulder width increases for both all KABCO and all KABC crashes. In other words, widening of shoulders has greater safety effects for roadways with a narrow shoulder width. To evaluate more reliable estimates, the variables for actual increased width and median width were transformed as binary variables. The results show that widening the shoulder has lower CM factors for the roadways with narrower median width, possible because the safety treatments generally are more safety effective when they are implemented for hazardous roadway conditions (e.g., narrower shoulder and median widths, higher traffic volumes in each lane, more roadside obstacles). As found from the developed SPFs shown in Table 2, roadways with

TABLE 4 Estimated CM factors for Original and Increased Shoulder Widths

Crash Type (severity)	Overall Safety Effects		Different Original Shoulder Width				Different Actual Increased Width			
			2–8 ft		9–12 ft		1–4 ft		5–10 ft	
	CMF	SE	CMF	SE	CMF	SE	CMF	SE	CMF	SE
All (KABCO)	0.88**	0.04	0.72**	0.07	0.94	0.05	0.94	0.07	0.85**	0.05
All (KABC)	0.82**	0.05	0.73**	0.09	0.84**	0.06	0.85*	0.09	0.80**	0.06
All (KAB)	0.79**	0.06	0.69**	0.12	0.82**	0.08	0.84	0.12	0.77**	0.08
ROR (KABCO)	0.75*	0.08	0.66**	0.15	0.77**	0.09	0.77*	0.14	0.74**	0.09
ROR (KABC)	0.72*	0.10	0.62**	0.18	0.74**	0.11	0.73	0.17	0.71**	0.12
ROR (KAB)	0.69**	0.11	0.57**	0.19	0.73*	0.14	0.71	0.21	0.68**	0.13

NOTE: SE = standard error.

\* = significant at a 90% confidence level; \*\* = significant at a 95% confidence level.

**TABLE 5 Estimated CM Functions for Widening Shoulders, Regression Model**

Parameter	All KABCO			All KABC		
	Estimate	SE	p-Value	Estimate	SE	p-Value
Constant	-0.5170	0.0486	<.0001	-0.5394	0.0867	<.0001
Original shoulder width in before period (ft)	0.0258	0.0041	<.0001	0.0246	0.0072	.0028
Actual increased shoulder width indicator (1 = sites with 1~4 ft shoulder width increased; 0 = sites with 5~10 ft shoulder width widened)	0.1648	0.0205	<.0001	0.1729	0.0365	.0001
Median width indicator (1 = sites with less than 40-ft median width, 0 = sites with 40-ft or more than 40-ft median width)	-0.0599	0.0250	.0265	-0.0653	0.0446	.1587

NOTE: MSE = mean square error. All KABCO: MSE = 0.0024;  $R^2 = .8826$ ; adjusted  $R^2 = .8649$ . All KABC: MSE = 0.0077;  $R^2 = .7084$ ; and adjusted  $R^2 = .6647$ .

a wide median have fewer crashes, which indicates that a narrower median represents a hazardous roadway condition. Therefore, it could be more safety effective to widen the right shoulder on roadways with a narrower median than on roadways with a wide median. The treatment is still effective for reducing crashes in general. Also, it was found that CM factors decrease as actual widened shoulder width increases.

In the MARS models, the estimated parameters of basis functions were statistically significant at a 90% confidence level. The basis functions were constructed with truncated power functions based on knot values (39). The knots are automatically chosen in the ADAPTIVEREG procedure. In the MARS model for total crashes, the first basis function, BF0, is the intercept. The second basis function, BF1, is “10–original shoulder width” when original shoulder width is lower than 10 and is 0 otherwise (where the knot value is 10). Other basis functions are constructed similarly with

different knot values. Various interaction impacts among variables for ranges based on knot values were found from MARS, whereas no interaction impact was found in the linear regression models. Moreover, two variables (AADT and maximum speed limit) that were not captured in the regression model were found to be significant in MARS. The results also show that the MARS models generally provide better model fits than the regression models, possible because MARS can account for both nonlinear effects and interaction impacts between variables.

To check the capability of selected break points for the basis functions in MARS models, regression models with dummy variables based on the break points from MARS were developed, as presented in Table 7. Although the developed models show better model fitness than the normal regression models in Table 4, MARS models still produce more reliable estimates because they can account for interaction impacts between multiple variables for various ranges.

**TABLE 6 Estimated CM Functions for Widening Shoulders, MARS Model**

Basis Function	Basis Function Information	Estimate	SE	p-Value
MARS Model for All KABCO Crashes				
BF0	Constant	-0.2257	0.0163	<.0001
BF1	max (10 – original shoulder width, 0)	-0.0151	0.0083	.0874
BF2	max (original shoulder width – 10, 0)	—	—	—
BF3	Actual increased shoulder width indicator (1 = sites with 1~4 ft shoulder width increased, 0 = sites with 5~10 ft shoulder width increased)	0.1726	0.0174	<.0001
BF4	Median width indicator (1 = sites with less than 40-ft median width, 0 = sites with 40-ft or more than 40-ft median width)	-0.1720	0.0479	.0021
BF5	BF2 × max (10.02127 – ln AADT, 0)	-0.0371	0.0170	.0426
BF6	BF4 × max (original shoulder width – 6, 0)	0.0247	0.0101	.0252
MARS Model for All KABC Crashes				
BF0	Constant	-0.5535	0.0502	<.0001
BF1	max (original shoulder width – 4, 0)	0.1001	0.0318	.0055
BF2	Actual increased shoulder width indicator (1 = sites with 1~4 ft shoulder width increased, 0 = sites with 5~10 ft shoulder width increased)	0.1765	0.0324	<.0001
BF3	max (original shoulder width – 6, 0)	-0.0888	0.0390	.0354
BF4	Median width indicator (1 = sites with less than 40-ft median width, 0 = sites with 40-ft or more than 40-ft median width)	—	—	—
BF5	BF4 × max (maximum speed limit – 65, 0)	-0.0439	0.0149	.0086
BF6	BF4 × max (10.16585 – ln AADT, 0)	-0.0565	0.0502	.1027

NOTE: — = not applicable; For MARS model for all KABCO crashes: MSE = 0.0014;  $R^2 = .9385$ ; adjusted  $R^2 = .9215$ . For MARS model for all KABC crashes: MSE = 0.0049;  $R^2 = .8329$ ; and adjusted  $R^2 = .7865$ .

**TABLE 7** Estimated CM Functions for Widening Shoulders, Regression Model with Dummy Variables Based on Break Points from MARS

Parameter		Estimate	SE	p-Value
<b>All KABCO</b>				
Constant		-0.2525	0.0155	<.0001
Original shoulder width in before period (base: original shoulder width $\geq$ 10 ft)	Original shoulder width < 6 ft 6 ft $\leq$ original shoulder width < 10 ft	-0.2332 -0.1305	0.0256 0.0229	<.0001 <.0001
Actual increased shoulder width indicator (1 = sites with 1~4 ft shoulder width increased, 0 = sites with 5~10 ft shoulder width increased)		0.1552	0.0195	<.0001
Median width indicator (1 = sites with less than 40-ft median width, 0 = sites with 40 ft or more than 40-ft median width)		NS	—	—
log AADT indicator (1 = sites with less than 10.02127, 0 = sites with 10.02127 or more than 10.02127)		NS	—	—
<b>All KABC</b>				
Constant		-0.2770	0.0333	<.0001
Original shoulder width in before period (base: original shoulder width $\geq$ 10 ft)	Original shoulder width < 6 ft 6 ft $\leq$ original shoulder width < 10 ft	-0.1969 -0.2055	0.0848 0.0507	.0315 .0007
Actual increased shoulder width indicator (1 = sites with 1~4 ft shoulder width increased, 0 = sites with 5~10 ft shoulder width increased)		0.1928	0.0339	<.0001
Median width indicator (1 = sites with less than 40-ft median width; 0 = sites with 40 ft or more than 40-ft median width)		-0.0873	0.0371	.0296
log AADT indicator (1 = sites with less than 10.02127; 0 = sites with 10.02127 or more than 10.02127)		NS	—	—
Maximum speed limit indicator (1 = sites with less than 65 mph; 0 = sites with 65 mph or more than 65 mph)		NS	—	—

NOTE: NS = not significant; MSE = mean square error; KABCO em dashes represent the following: MSE = 0.0020;  $R^2 = .9021$ ; adjusted  $R^2 = .8874$ ; KABC em dashes represent the following: MSE = 0.0063;  $R^2 = .7759$ ; adjusted  $R^2 = .7287$ .

## CONCLUSIONS

This study assessed safety effectiveness of widening shoulders on rural multilane roadways considering the variation of CM factors with various site characteristics. To determine this variation, CM functions were developed with various statistical approaches. In particular, the MARS modeling approach was applied to quantify the changes of CM factors according to varying influential factors because of its ability to account for nonlinearity and interaction impacts between variables.

The results of estimated CM factors indicate that widening shoulders will reduce crash frequencies. In particular, the estimated CM factors show greater safety effects for severe crashes. Moreover, the CM factors for ROR crashes are lower than the CM factors for all crashes. The CM factors were also estimated for various ranges of original shoulder width and actual increased width. It was found that CM factors estimated separately for ranges of original shoulder width and actual increased width can better capture the effects of interactions between safety effects and site characteristics.

The CM functions were derived from this observed relationship. The results of CM functions show that the CM factors increase as original shoulder width increases for both all KABCO and all KABC crashes. Moreover, it was found that the CM factors decrease as actual shoulder width increases. The results also show that widening shoulders has greater safety effects for roadways with narrower medians. The study demonstrated that the CM functions developed with the MARS model can better reflect variations in safety effects of widening shoulders than can CM functions that use multiple linear regression.

Although the study provided empirical evidence of the variation of safety effects depending on site characteristics, there are some limitations in this study. MARS may not clearly consider specific nonlinear trend because the rate of change is assumed to be fixed within a given range of a variable, although the rate can vary within the range. Generalized nonlinear models possibly could be adopted to overcome this limitation. A more general relationship between safety effects and site characteristics could not be observed because of a lack of samples in this study. Thus, the developed CM functions could be used to check the overall impact of widening shoulders with relatively large variation. Also, inclusion of multiple target areas (e.g., more states) in the analysis could produce more generalized conclusions.

## ACKNOWLEDGMENT

The authors thank the Florida Department of Transportation for providing the data used in this study and for funding this research.

## REFERENCES

1. *Highway Safety Manual*. AASHTO, Washington, D.C., 2010.
2. Gross, F., B. Persaud, and C. Lyon. *A Guide to Developing Quality Crash Modification Factors*. FHWA-SA-10-032. FHWA, U.S. Department of Transportation, 2010.
3. Abdel-Aty, M., C. Lee, J. Park, J. Wang, M. Abuzwidah, and S. Al-Arifi. *Validation and Application of Highway Safety Manual (Part D) in Florida*. Florida Department of Transportation, Tallahassee, 2014.

4. Lee, C., M. Abdel-Aty, J. Park, and J. Wang. Development of Crash Modification Factors for Changing Lane Width on Roadway Segments Using Generalized Nonlinear Models. *Accident Analysis and Prevention*, Vol. 75, 2015, pp. 83–91.
5. Elvik, R. Introductory Guide to Systematic Reviews and Meta-Analysis. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1908*, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 230–235.
6. Van Houwelingen, H. C., R. A. Lidia, and T. Stijnen. Advanced Methods in Meta-Analysis: Multivariate Approach and Meta-Regression. *Statistics in Medicine*, Vol. 21, No. 4, 2002, pp. 589–624.
7. Elvik, R. Developing Accident Modification Functions. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2103*, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 18–24.
8. Elvik, R. International Transferability of Accident Modification Functions for Horizontal Curves. *Accident Analysis and Prevention*, Vol. 59, 2013, pp. 487–496.
9. Elvik, R. Developing an Accident Modification Function for Speed Enforcement. *Safety Science*, Vol. 49, 2011, pp. 920–925.
10. Park, J., M. Abdel-Aty, and C. Lee. Exploration and Comparison of Crash Modification Factors for Multiple Treatments on Rural Multilane Roadways. *Accident Analysis and Prevention*, Vol. 70, 2014, pp. 167–177.
11. Park, J., M. Abdel-Aty, J. Wang, and C. Lee. Assessment of Safety Effects for Widening Urban Roadways in Developing Crash Modification Functions Using Nonlinearizing Link Functions. *Accident Analysis and Prevention*, Vol. 79, 2015, pp. 80–87.
12. Sacchi, E., T. Sayed, and K. El-Basyouny. Collision Modification Functions: Incorporating Changes over Time. *Accident Analysis and Prevention*, Vol. 70, 2014, pp. 46–54.
13. Sacchi, E., and T. Sayed. Accounting for Heterogeneity Among Treatment Sites and Time Trends in Developing Crash Modification Functions. *Accident Analysis and Prevention*, Vol. 72, 2014, pp. 116–126.
14. Park, J., M. Abdel-Aty, J. Lee, and C. Lee. Developing Crash Modification Functions to Assess Safety Effects of Adding Bike Lanes for Urban Arterials with Different Roadway and Socio-Economic Characteristics. *Accident Analysis and Prevention*, Vol. 74, 2015, pp. 179–191.
15. Hadi, M. A., J. Aruldas, L. Chow, and J. Wattleworth. Estimating Safety Effects of Cross-Section Design for Various Highway Types Using Negative Binomial Regression. In *Transportation Research Record 1500*, TRB, National Research Council, Washington, D.C., 1995, pp. 169–177.
16. Jovanis, P., and F. Gross. Estimation of Safety Effectiveness of Changes in Shoulder Width with Case Control and Cohort Methods. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2019*, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 237–245.
17. Harkey, D. L., R. Srinivasan, J. Baek, F. M. Council, K. Eccles, N. Lefler, F. Gross, B. Persaud, C. Lyon, E. Hauer, and J. A. Bonneson. *NCHRP Report 617: Accident Modification Factors for Traffic Engineering and ITS Improvements*. Transportation Research Board of the National Academies, Washington, D.C., 2008.
18. Zeng, H., and S. Schrock. Safety Effectiveness of Various Types of Shoulders on Rural Two-Lane Roads in Winter and Non-Winter Periods. Presented at 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2013.
19. Park, J., and M. Abdel-Aty. Development of Adjustment Functions to Assess Combined Safety Effects of Multiple Treatments on Rural Two-lane Roadways. *Accident Analysis and Prevention*, Vol. 75, 2015, pp. 310–319.
20. Stamatiadis, N., J. Pigman, J. Sacksteder, W. Ruff, and D. Lord. *NCHRP Report 633: Impact of Shoulder Width and Median Width on Safety*. Transportation Research Board of the National Academies, Washington, D.C., 2009.
21. Gross, F., P. P. Jovanis, and K. Eccles. Safety Effectiveness of Lane and Shoulder Width Combinations on Rural, Two-Lane, Undivided Roads. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2103*, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 42–49.
22. Potts, I. G., D. W. Harwood, and K. R. Richard. Relationship of Lane Width to Safety for Urban and Suburban Arterials. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2023*, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 63–82.
23. Lord, D., and J. A. Bonneson. Development of Accident Modification Factors for Rural Frontage Road Segments in Texas. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2023*, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 20–27.
24. Yanmaz-Tuzel, O., and K. A. Ozbay. Comparative Full Bayesian Before-and-After Analysis and Application to Urban Road Safety Countermeasures in New Jersey. *Accident Analysis and Prevention*, Vol. 42, 2010, pp. 2099–2107.
25. Labi, S. Efficacies of Roadway Safety Improvements Across Functional Subclasses of Rural Two-Lane Highways. *Journal of Safety Research*, Vol. 42, 2011, pp. 231–239.
26. Mehta, G., and Y. Lu. Calibration and Development of Safety Performance Functions for Alabama: Two-Lane, Two-Way Rural Roads and Four-Lane Divided Highways. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2398*, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 75–82.
27. Hauer, E. *Lane Width and Safety*. 2000. <http://ezrahauer.files.wordpress.com/2012/08/lane-width-and-safety.pdf>. Accessed April 2014.
28. *2012 Motor Vehicle Crashes: Overview*. NHTSA, U.S. Department of Transportation, 2013.
29. Hauer, E. *Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*. Pergamon Press, Oxford, United Kingdom, 1997.
30. Hauer, E., D. Harwood, F. Council, and M. Griffith. Estimating Safety by the Empirical Bayes Method: A Tutorial. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1784*, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 126–131.
31. Persaud, B., and C. Lyon. Empirical Bayes Before-After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions. *Accident Analysis and Prevention*, Vol. 39, 2007, pp. 546–555.
32. Friedman, J. Multivariate Adaptive Regression Splines. *Annals of Statistics*, Vol. 19, 1991, pp. 1–67.
33. Abraham, A., D. Steinberg, and N. Philip. Rainfall Forecasting Using Soft Computing Models and Multivariate Adaptive Regression Splines. In *IEEE SMC Transactions: Special Issue on Fusion of Soft Computing and Hard Computing in Industrial Applications*, IEEE, New York, 2001.
34. Snedecor, G. W., and W. G. Cochran. *Statistical Methods* (7th ed.), Iowa State University Press, Ames, 1980.
35. Put, R., Q. Xu, D. Massart, and Y. Heyden. Multivariate Adaptive Regression Splines (MARS) in Chromatographic Quantitative Structure-Retention Relationship Studies. *Journal of Chromatography A*, Vol. 1055, 2004, pp. 11–19.
36. Haleem, K., A. Gan, and J. Lu. Using Multivariate Adaptive Regression Splines (MARS) to Develop Crash Modification Factors for Urban Freeway Interchange Influence Areas. *Accident Analysis and Prevention*, Vol. 55, 2013, pp. 12–21.
37. Park, J., and M. Abdel-Aty. Assessing the Safety Effects of Multiple Roadside Treatments Using Parametric and Nonparametric Approaches. *Accident Analysis and Prevention*, Vol. 83, 2015, pp. 203–213.
38. *The ADAPTIVEREG Procedure*. SAS Institute, Cary, N.C., 2012.
39. Kuhfeld, W. F., and W. Cai. Introducing the New ADAPTIVEREG Procedure for Adaptive Regression. Presented at SAS Global Forum: Statistics and Data Analysis, 2013.

---

*The opinions, findings, and conclusions in this paper are those of the authors.*

*The Standing Committee on Highway Safety Performance peer-reviewed this paper.*