Journal of Computer Science and Electrical Engineering
Print ISSN: 2663-1938
Online ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3165

FUZZY SEMIPARAMETRIC PARTIAL LINEAR MODELS BASED ON RIDGE ESTIMATION

XiaoXu Liu*, MingJuan Ma, DaMing Chen
Aviation University of Air Force, Changchun 130000, Jilin, China.
**Corresponding Author: XiaoXu Liu*

Abstract: Multicollinearity may lead to unstable least-squares coefficients and introduce errors in the smooth part of a fuzzy semiparametric partial linear model. This paper introduces a ridge estimator for five linear meteorological effects and models wind speed nonparametrically under LR fuzzy responses. A Diamond-distance residual criterion is combined with a quadratic penalty, and a two-stage procedure is applied to 21 daily fuzzy carbon-monoxide observations formed from hourly measurements in Rome. The six inputs are standard-centred; the reported fit uses a Gaussian kernel with bandwidth 0.08 and a ridge parameter of 0.1. Ridge fitting reduces MSEe from 0.0763 to 0.0508 and MAPEe from 2.4003% to 1.6370% compared with ordinary two-step estimation. It also improves on the reported fuzzy linear comparator for these measures, although that model uses a different predictor subset and preprocessing. The case shows how ridge shrinkage can stabilise the parametric part without affecting the smooth term.
Keywords: Fuzzy semiparametric partial linear model; Ridge estimation; LR fuzzy number; Multicollinearity; Kernel smoothing

1 INTRODUCTION

Fuzzy ridge regression has been extended to imprecise inputs and outputs [1]. Recently, fuzzy nonparametric methods have employed kernel smoothing for fuzzy responses [2], and robust kernel formulations address outliers in fuzzy regression [3].

Recently, some alternative biased estimators for collinearity and ridge behaviour in factor-augmented designs have been proposed [4-5]. Kernel ridge-type shrinkage has also been used in partially linear models with correlated errors [6]. Recent studies of multicollinearity have revisited VIF- and condition-based diagnostics [7], while fuzzy ridge and quantile-based research have investigated penalty selection and fitting with fuzzy data [8-9]. Although these approaches are motivating, they do not provide an estimator for LR-fuzzy responses in the AFFSPLM structure.

This paper considers a scenario where relevant meteorological predictors are retained despite collinearity and the response is an asymmetric LR triangular fuzzy number. Five predictors are in the linear part, and wind speed is in the kernel-smoothed component. A fuzzy residual criterion based on the Diamond distance is extended by a quadratic ridge penalty; then, the estimated linear coefficients are used to obtain a second-stage estimate of the smooth effect. The procedure is tested on 21 daily fuzzy carbon-monoxide observations that have been aggregated from hourly monitoring in Rome between 5 and 25 October 1992. All six inputs have been standardised before collinearity diagnosis and model fitting.

Empirically, we compare ordinary two-step estimation with a published fuzzy linear model. Record the Gaussian-kernel bandwidth and selected penalty; evaluate fit using the reported error and R-squared measures; retain observed-versus-fitted plots for centres and spreads. Ridge fit has a lower reported MSEe and MAPEe than the ordinary two-step fit; however, comparison with the fuzzy linear model is not a controlled estimator-only test because that model uses a different predictor subset and preprocessing. The difference is that although regularisation can improve stability, it introduces bias; therefore, optimisation must consider the actual data and loss function specified. Therefore, ridge regularisation is tested here, but it does not have generalised estimator superiority. The ridge penalty follows the bias-variance trade-off in recent estimators for collinearity [4-5].

2 MODEL FORMULATION AND RIDGE ESTIMATION

2.1 LR Fuzzy Responses and Distance

An LR fuzzy number is characterised by a central value and left-right shape parameters. Its shape functions show how the membership decreases on either side of the center; a triangular case can be obtained with symmetric triangular shapes. This way, both the central tendency and the spread of the response are retained without reducing each observation to a crisp value. Recent studies have covered fuzzy ridge fitting [1], fuzzy-response kernel smoothing [2], robust kernel regression [3], and quantile-based fuzzy regression [9].

The LR representation and weighted Diamond distance for residual comparison are as follows:

$$\widetilde{A} = (a, l_a, r_a) \tag{1}$$

$$d_E^2(\widetilde{A}, \widetilde{B}) = \omega_m(a - b)^2 + \omega_l(l_a - l_b)^2 + \omega_r(r_a - r_b)^2 \tag{2}$$

© By the Author(s) 2026, under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

The three positive weights apply to differences in centers and left/right shape parameters. The distance is used to compare fuzzy residuals of the same response representation. Recently, some research has introduced a fitting criterion for uncertain responses [1], and fuzzy nonparametric methods have also been developed using kernel-based procedures for smooth components [2].

2.2 Fuzzy Semiparametric Partial Linear Model

AFFSPLM is a combination of a linear part and an unknown smooth component. Based on the above design, five meteorological variables are included in the parametric section and wind speed enters the non-parametric part. The model is given by:

$$\tilde{y}_i = \sum_{\ell=1}^{5} (\tilde{\beta}_\ell \otimes x_{i\ell}) \oplus \tilde{f}(x_{i6}) \oplus \tilde{\varepsilon}_i, \quad i = 1, \ldots, n \tag{3}$$

The response, coefficients, smooth function and error here are all fuzzy quantities; the six meteorological inputs are precise. The first five predictors make up the linear part, and the sixth is a smooth function. First, estimate the fuzzy linear coefficients in the two-stage procedure, and then smooth the residual centre and shape components at selected design points.

2.3 Ridge Criterion and Closed-Form Estimate

Recently, regularised estimators for collinearity have been introduced that reduce coefficient bias to lower variance [4-5]. The fuzzy model also works in the same way; a square Diamond-distance residual criterion is added to a quadratic penalty on the five fuzzy coefficients. When there is no penalty, it is equivalent to ordinary least squares; otherwise, as the penalty increases, so does parameter shrinkage. The goal is:

$$PRSS(\beta; \lambda) = \sum_{i=1}^{n} d_E^2(\tilde{y}_i, \hat{\tilde{y}}_i) + \lambda \sum_{j=1}^{5} (\tilde{\beta}_j \otimes \tilde{\beta}_j) \tag{4}$$

Minimising the above criterion gives the fuzzy ridge coefficient estimate:

$$\hat{\tilde{\beta}}_r(\lambda) = \Phi^{-1} S \tilde{Y} \tag{5}$$

The matrices in the closed-form expression are formed by the design matrix, the fuzzy-distance transformation and the identity penalty. Their components in the notation of derivation are:

$$\Phi = t^T t - t^T D\, t - t^T D^T t + t^T D^T D\, t + \lambda I_p \tag{6}$$
$$S = t^T - t^T D^T - t^T D + t^T D^T D$$

The order of the identity matrix is equal to the number of linear coefficients. The second-stage smooth term is obtained by kernel matrix estimation from the residual center and shape series. The Gaussian kernel used in the application is:

$$K(u) = (2\pi)^{-1/2} \exp(-u^2/2) \tag{7}$$

Bandwidth determines the extent of local smoothing, and a ridge parameter introduces a constraint on the size of the linear coefficients. Both are presented with the application results; no further tuning rules are added to the selection described in the study.

3 EMPIRICAL DESIGN

3.1 Data Construction and Predictors

The empirical series is the Rome air-quality application that Coppi and others have reported [10]. Carbon monoxide concentration was continuously recorded every hour for 21 days between October 5th and 25th, 1992. Temperature, relative humidity, pressure, rainfall, solar radiation and wind speed were recorded once daily during this time. Daily hourly CO observations are summarised as asymmetric LR triangular fuzzy responses: the centre is the daily mean, and the two sides show deviations below and above this mean. The obtained fuzzy response retains the variation within a day but keeps the predictors exact. Six predictors in Table 1, their units and assigned model components are listed.

Table 1 Meteorological Predictors and Their Model Roles

No.	Predictor	Unit	Model component
1	Temperature	°C	Linear
2	Relative humidity	%	Linear
3	Atmospheric pressure	mbar	Linear
4	Rainfall	mm	Linear
5	Solar radiation	mW/cm²	Linear
6	Wind speed	m/s	Kernel-smoothed

The predictors have different units and scales, so all six were standard-centred before estimation. The parametric component therefore has five standardised inputs, and wind speed is used as the only smoothing variable. The ordinary two-step and ridge versions use the same AFFSPLM structure for within-study comparison.

3.2 Collinearity Assessment and Tuning

Collinearity was tested by pairwise correlation, the design-matrix condition number and variance inflation factors according to the above linear regression diagnostics [7]. The strongest reported pairwise relationship is that of temperature and solar radiation, with a correlation coefficient of 0.833. The condition number is 24.848, which is within the range for mild collinearity as indicated by the source study; VIF analysis also shows some correlations among predictors, but individual values are not listed. Based on the above analyses, there is only a small amount of multicollinearity in this sample.

$$k = \frac{\lambda_1}{\lambda_p} = 24.848 \tag{8}$$

The reported application uses a Gaussian kernel with a bandwidth of 0.08 and selects a ridge parameter of 0.1. Table 2 presents the five coefficient estimates as LR triples, ordered by centre, left component, and right component.

4 RESULTS AND COMPARISON

4.1 Coefficient Estimates

Table 2 Fuzzy Coefficient Estimates (Center, Left, and Right Components)

Coefficient	Ordinary two-step estimate	Ridge estimate
1	(1.0399, 0.6613, 1.1835)	(1.1059, 0.7121, 1.2625)
2	(0.3813, 0.1342, 0.5322)	(0.4033, 0.1409, 0.5662)
3	(0.5618, 0.3743, 0.9358)	(0.6199, 0.4156, 1.0416)
4	(0.0062, -0.0404, 0.3106)	(0.0119, -0.0353, 0.3408)
5	(-0.7401, -0.8057, -0.3452)	(-0.7878, -0.8565, -0.3755)

Ridge Shrinkage modifies the estimated centres and both fuzzy-side components but keeps all five linear predictors. The estimated smooth term is then obtained from the residual part. Figure 1 shows the ordinary two-step fitted response against the observed fuzzy response; the centre and two sides track the broad temporal pattern, with localised differences around the mid-series decline.

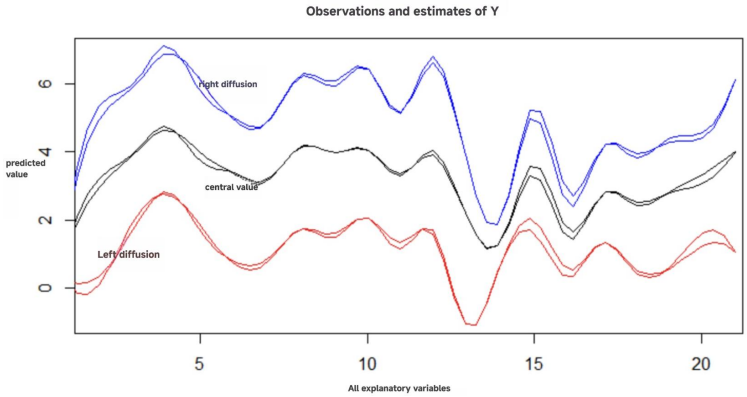


Figure 1 Ordinary Two-Step Fit and Observed Fuzzy CO Response

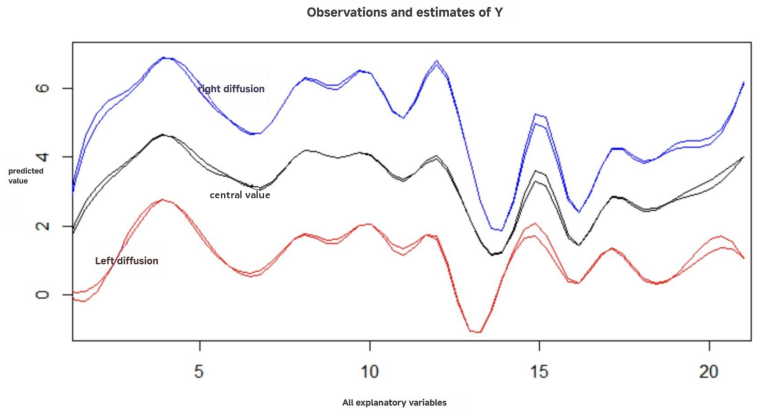


Figure 2 Ridge-Estimated Fit and Observed Fuzzy CO Response

Figure 2 is the ridge fit. Its predicted center and side curves stay close to the observed series for all 21 days, including a noticeable drop in the middle of the observation period. The above visual comparison is in line with the smaller error values shown by the ridge fit.

4.2 Fit Measures and Comparator

Table 3 shows the six fit measures of all three types: the ordinary two-step estimator, the ridge estimator, and the Coppi fuzzy linear model. Compared with the ordinary two-step fit, the ridge estimate reduces MSEe from 0.0763 to 0.0508 and MAPEe from 2.4003% to 1.6370%; the two reported R-squared values increase from 0.9521 and 0.9799 to 0.9844 and 0.9866. Both the negative MPEe and SMAPEe have moved closer to zero. The Coppi model has an MSEe of 1.8953 and a MAPEe of 5.4236%.

Table 3 Reported Overall Fit Measures

Measure	Ordinary two-step	Ridge AFFSPLM	Coppi fuzzy linear
MSEe	0.0763	0.0508	1.8953
MPEe	-0.1512	-0.0831	-0.4287
MAPEe	2.4003%	1.6370%	5.4236%
SMAPEe	-0.2857	-0.1048	-0.4762
$R1^2$	0.9521	0.9844	0.5148
$R2^2$	0.9799	0.9866	0.5000

Figure 3 shows the fitted fuzzy response of the comparison model. The general trend is the same, but the magnitude of deviation and fit index are relatively large. This comparison is informative but not an estimator-only experiment: the comparator retains humidity, rainfall and wind speed, whereas the AFFSPLM uses all six meteorological predictors; the comparator also does not standard-center its predictors. Therefore, predictor selection and scale handling differ with the estimation method.

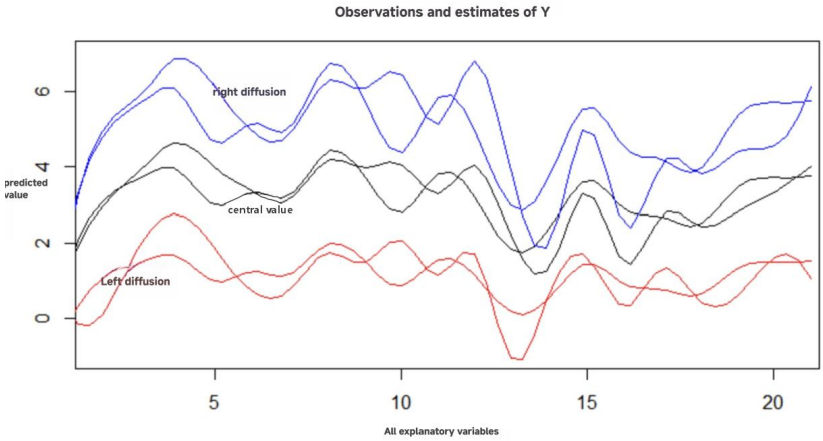


Figure 3 Coppi Fuzzy Linear Model Fit and Observed Response

5 CONCLUSION

A ridge estimator was proposed for a fuzzy semiparametric partial linear model with five linear effects and a kernel-smoothed wind-speed effect. LR fuzzy responses retain the centre and side information, and the Diamond-distance residual criterion and quadratic penalty provide a basis for parameter estimation. In the 21-day Rome CO application, standard-centering was followed by collinearity diagnostics, Gaussian kernel smoothing, and ridge tuning. The reported ridge fit has improved over the ordinary two-step fit in all the main error measures and the two reported R-squared values. It also shows some advantages over the Coppi fuzzy linear model based on the reported indicators, although these differences may be due to different predictors and preprocessing methods. Based on the above results, ridge regularisation is suitable for this fuzzy-response application; however, it does not exhibit general superiority or out-of-sample performance. Further tests using more data sets and a fixed predictor set will help determine whether the improvement is due to shrinkage alone.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Behdani Z, Darehmiraki M. Ridge regression with intuitionistic fuzzy input and output: a parametric approach. Fuzzy Optimization and Modeling Journal, 2024, 5(2): 19-31. DOI: 10.71808/FOMJ.2024.1002660.

[2] Hesamian G, Torkian F, Johannssen A, et al. A fuzzy nonparametric regression model based on an extended center and range method. Journal of Computational and Applied Mathematics, 2024, 436: 115377. DOI: 10.1016/j.cam.2023.115377.

[3] Kong L, Song C. Fuzzy robust regression based on exponential-type kernel functions. Journal of Computational and Applied Mathematics, 2025, 457: 116295. DOI: 10.1016/j.cam.2024.116295.

[4] Saray K F, Alheety M I N. A new convex estimator for linear regression model. Iraqi Journal of Science, 2025, 66(12): 5563-5583. DOI: 10.24996/ijs.2025.66.12.24.
[5] He Y. Ridge regression under dense factor augmented models. Journal of the American Statistical Association, 2024, 119(546): 1566-1578. DOI: 10.1080/01621459.2023.2206082.
[6] Ahmed S E, Yilmaz E, Aydin D. Kernel ridge-type shrinkage estimators in partially linear regression models with correlated errors. Mathematics, 2025, 13(12): 1959. DOI: 10.3390/math13121959.
[7] Osman E G A. Entropy based solutions for detecting and treating multicollinearity in econometrics. Discover Analytics, 2025, 3: 18. DOI: 10.1007/s44257-025-00047-0.
[8] Homaida A, Pekalp M H, Ebegil M. Efficient k value computation for enhanced fuzzy ridge regression. Journal of Computational and Applied Mathematics, 2026, 472: 116813. DOI: 10.1016/j.cam.2025.116813.
[9] Arefi M, Khammar A H. Nonlinear prediction of fuzzy regression model based on quantile loss function. Soft Computing, 2024, 28(6): 4861-4871. DOI: 10.1007/s00500-023-09190-w.
[10] Coppi R, D'Urso P, Giordani P, et al. Least squares estimation of a linear regression model with LR fuzzy response. Computational Statistics & Data Analysis, 2006, 51(1): 267-286. DOI: 10.1016/j.csda.2006.04.036.