The mixed ridge is evaluated using wavelet analysis as an alternative to conventional penalized approaches (simulation study)

Dler H. Kadir[1,2], Alan G. Rahim[1], Dlshad M. Saleh1,3 and Azhin M. Khudhur[1]

[1]Department of Statistics and Informatics, College of Administration and Economics, Salahaddin University-Erbil, Kurdistan Region - F.R. Iraq

[2]Department of Business Administration, Cihan University-Erbil, Kurdistan Region, Iraq

[3]Department of Accounting and Financial, College of Administration and Economics, Lebanese French University, Erbil, Iraq

***Abstract*—In this research, we use threshold wavelet approaches to estimate tuning parameters in the presence of the contamination problem (noise) for calculated parameters in multiple linear regression. To estimate tuning parameters, the suggested technique uses wavelet shrinkage (for wavelets family, Symlets and Fejer-Korovkin and threshold, Universal, SURE, and Minimax). It is also utilized in penalized methods for variable selection and linear model coefficient estimation. Using simulation experiments for contamination levels of 10% and 25% as well as actual data, in contrast to the conventional penalized method, this study presents a straightforward estimate for parameter adjustment based on the penalized method's wavelet shrinkage (Ridge). The results of using the statistical criteria (MAE and MSE) to compare the suggested technique to a traditional penalized method. The penalized technique's wavelet shrinkage produces the best results and is more precise than the conventional approach in all simulations, according to the MAE and MSE criteria.**

***Keywords*—Ridge regression, Wavelet Shrinkage, Threshold, Wavelet Ridge, and the penalized technique.**

I. INTRODUCTION

Penalized regression approaches for linear regression have emerged in recent decades to solve prediction accuracy difficulties with regular least squares regression (Van der Kooij, 2007).

The residual squared error is decreased to obtain ordinary least squares (OLS) estimations. The data analyst frequently rejects OLS estimates for a couple of parameters. The first is prediction accuracy: OLS estimates often have high variation but low bias. In certain circumstances, prediction precision may be improved by decreasing or zeroing out specific components. Ordinary least squares (OLS) estimations are obtained by minimizing the residual squared error. For two factors, the OLS estimates usually leave the data analyst unsatisfied. Lowering or setting some coefficients to 0 can improve prediction precision in some situations. The first is prediction accuracy: OLS estimates often have a high variance but little bias. By doing this, we give up some bias to lower the projected values' variance, which could increase the prediction. Interpretation is the second reason. We frequently want to identify a smaller subset of predictors with the biggest impacts when there are many (Tibshirani, 1996). Multiple regression is commonly used to build a model for predicting future responses or to study the relationship between the variable that responded and the predictor variables. The model's prediction precision is crucial for the first purpose, while its complexity is more interesting for the second To address the shortcomings of ordinary least squares (OLS) regression, which frequently performs inadequately in regards to model complexity and precision of prediction, several established regression techniques have been developed in recent decades, beginning with Ridge regression (Hoerl & Kennard, 1970).

However, the precision of the tuning parameter utilized in the penalty functions is required for this system to work well. There are several options for calculating the tuning parameter. They are ascertained by applying an appropriate criterion. These criteria might be lowered in proportion to the adjustment parameter to get the desired selection. Currently, the majority of widely used methods based on data are generalized cross-validation (GCV) and cross-validation (CV) (Fan & Li, 2001).

Donoho and Johnstone (1995) Developed the wavelet threshold technique, which uses thresholding coefficients to reassemble signals. The wavelet threshold technique's denoising effect is threshold-dependent. A threshold set excessively will remove some important information, but a low enough threshold will keep some noise. Many authors researched threshold determination strategies to overcome this issue. Donoho (1993) evaluated a normal Gaussian noise model and suggested a universal threshold. These systems have the drawback of frequently establishing a universal threshold. These systems have the drawback of typically having an excessively high universal threshold, which may result in an excess of relevant data.

However, the distribution could not be applied to a single signal, all these strategies rely on a particular coefficient distribution. Donoho and Johnstone (1994)developed a new threshold approach based on the minimax criteria. However, this strategy requires previous information on the signal, which is difficult to get in real-world situations. The

DOI: http://doi.org/10.24086/icafs2025/paper.1764

generalized cross-validation (GCV) criterion, Stein's unbiased risk assessment (SURE) criteria, and the concept of parameter estimates were presented (Jansen & Bultheel, 1999). Wavelet shrinking for ridge is suggested in this paper as a practical solution to these problems. The efficacy of the proposed methodologies is investigated through simulated tests and real-world data applications.

II. MATERIALS AND METHODOLOGY

The most significant regularization methodology for coefficient regression was examined by the methodologies used in the simulation study-based estimate procedures This was demonstrated via Wavelet shrinkage for penalized techniques and traditional penalized algorithms. The mean squared and mean absolute errors are used to determine relative efficiency.

A. Penalized Methods

Penal approaches have grown in favor of statisticians in recent years due to their use in performing simultaneous parameter estimation and variable selection. Consequently, a variety of penalty techniques have been put forth, incorporating a penalty constraint into the regression models (Tutz & Ulbricht, 2009). By imposing limits on transactions that require certain transactions to have a value of zero, the penalty constraint aims to regulate the model's complexity and offer a criterion for the selection of variables (Helwig, 2017).

The chosen model's variance and bias are balanced by the penalty constraint quantity. In contrast, a large penalty amount results in the selection of a few explanatory variables with a large bias but a lower variance. More explanatory variables with a small bias but a large variance are chosen when the penalty amount is small. Consequently, a well-chosen penalty amount improves the model's predictability as well as its readability and interpretation (Li & Sillanpää, 2012).

It is commonly referred to as Penalized Linear Regression (PLR).

In this manner

$$PLR(\beta;\lambda) = (Y - X\beta)^T(Y - X\beta) + \lambda \sum_{j=1}^{p} P_\lambda(|\beta_j|) \tag{1}$$

Because $(\lambda \geq 0)$, the tuning parameter is denoted by (λ). The penalty limit relies on the value of (λ), which governs the amount of parameter shrinkage. The penalty term, $P_\lambda(|\beta_j|)$, is a function of coefficients. When $\lambda = 0$, the Ordinary Least Squares approach (OLS) estimates the value. On the other hand, if the value of (λ) rises, more variables will be left out of the model (Wood, 2006).

The following equation is used in partial linear regression to find estimates of the model parameters:

$$\hat{\beta}_{PLR} = argmin(Y - X\beta)^T(Y - X\beta) + \lambda \sum_{j=1}^{p} P_\lambda(|\beta_j|) \tag{2}$$

When the variable is unbiased for significant real factors, it is argued that an appropriate penalty term should provide an estimator with three properties. The fundamental characteristic is impartiality. Second, simplicity causes small estimators to be exactly zero. As a result, the computed continuity is (continuous) in the data, helping avoid model prediction inaccuracy (Fan & Li, 2001).

The features of several penalized techniques have been developed and analyzed, including Ridge, Elastic-Net, Bridge, Least Absolute Shrinkage and Selection Operator (LASSO), and others.

B. Ridge Regression (RR)

Selecting variables and estimating parameters for regression modeling using related explanatory variables is a complex issue. This is because when multicollinearity arises, there is insufficient information in the data matrix to discriminate between the impacts of a correlated variable and an associated variable (Hoerl & Kennard, 1970). Because the penalty quantity decreases the regression coefficients, the Ridge Regression method reduces variance in coefficient estimations through the addition of a penalty quantity resulting from the rule (L2-norm) to the sum of the squares of the residuals that remain (Birdawod et al., 2024; Saleh et al., 2023). The term ridge is used to characterize penalized linear regression which is as below:

$$PLR(\beta;\lambda)^{RR} = (Y - X\beta)^T(Y - X\beta) + \lambda \sum_{j=1}^{k} \beta_j^2 \tag{3}$$

$$\hat{\beta}_{PLR}^{RR} = (X^TX + \lambda I)^{-1}X^TY \tag{4}$$

C. Wavelets Shrinkage

A quick nonlinear technique for image denoising that preserves discontinuity is wavelet shrinkage. The basic idea behind it is to break down an image using a wavelet basis, shrink all of the coefficients by a modest amount, and then use the shrunken coefficients to reconstruct the filtered image. The success of this process depends on the supposition that a comparatively limited number of wavelet coefficients with large magnitude may accurately reflect the original image, but moderate Gaussian noise impacts all coefficients, albeit less severely (Antoniadis, 2007).

Since sines and cosines are the foundation of Fourier analysis, researchers have been searching for more suitable functions to model unpredictable signals. According to Donoho and Johnstone (1995), these functions are non-local and have an infinite range. As a result, they approximate sharp spikes very poorly. However, wavelet analysis allows us to use condensed approximation functions in finite domains. An

DOI: http://doi.org/10.24086/icafs2025/paper.1764

effective method for approximating data with abrupt discontinuities is to use wavelets. The wavelet prototypical function, also identified as an analyzing wavelet or mother wavelet, is utilized to begin wavelet analysis (Kadir et al., 2024).

Symlet Wavelets

As updates to the db family, Daubechies presented the Symlets, which are almost symmetrical wavelets (Ali & Saleh, 2021).

Fejer-Korovkin (fk)

This wavelet filter can be considered more symmetrical than the Daubechies filters, but less soft. That kind of filter consists variety of uses in the theory of approximation and provides an adequate response to frequency as support increases (Varanis & Pederiva, 2017).

Thresholding

The simplest basic approach for non-linear wavelet denoising is thresholding, which divides the wavelet coefficient into two brunches, one expressing the signal while the other expressing the noise (Hamad, 2010).

Universal Threshold

David L Donoho and Johnstone (1994) suggested the universal threshold, it is supplied as follows:

$$\eta^{U} = \tilde{\sigma}_{(MAD)}\sqrt{2logN} \tag{5}$$

SURE Threshold

Donoho and Johnstone (1994) proposed a criterion that relied on reducing Stein's unbiased risk estimate. The soft threshold estimator assigns a threshold value of η_j to every level j of the wavelet coefficients in the sure threshold.

$$SURE(\eta, W) = N - 2 \neq \{j: |W_j| \leq \eta\} - \sum_{j=0}^{d} min\,(|W_j|, \eta) \tag{6}$$

Where $\{W_j: j = 1,2, \ldots, d\}$ be a wavelet coefficients in the jth level, and

Then, select η^S that minimizes SURE (η, W).

$$\eta^S = arg\ arg\ SURE\ (\eta, W)$$

The SURE threshold is a hybrid thresholding strategy that combines the SURE and universal thresholds, claim Donoho and Johnstone (1995). If the collection of coefficients is sparsely displayed, the universal threshold is utilized for choosing a threshold level; alternatively, the SURE threshold is applied.

The level j is considered scant if:

$$W_{SS}(\eta) \leq 1 + \frac{\left(log\ log\ N_j\right)^{\frac{3}{2}}}{\sqrt{N_j}} \tag{7}$$

$$W_{SS}(\eta) = \sum W_{j,t}^2 \tag{8}$$

Minimax Threshold

An estimator $\tilde{f}$ that achieves the minimax risk is the foundation of the ideal minimax threshold, which David L Donoho and Johnstone (1995) suggested as an enhancement to the universal threshold.

$$\tilde{R}(F) = in_{\tilde{f}} f sup_{f \in \tilde{R}(F)} R(\tilde{f}, f) \tag{9}$$

Where

$$R(\tilde{f}, f) = \frac{1}{N}\sum_{i=1}^{N} E[\tilde{f} - f]^2 \tag{10}$$

III. PROPOSED METHOD

When the parameters have heavy-tailed distributions and de-noising values, the proposed approach employs wavelet shrinkage to estimate the tuning parameter in penalized linear regression. This technique employs a threshold criterion before relying on the small wave filter. To calculate a multiple linear regression model, compute (MSE and MAE), and compare it with the classical ones, the data is adjusted for Wavelet shrinkage for Penalized methods (Wavelet Ridge) after DWT and denoise data are found using the outputs.

The inverse of DWT is applied to the decreased set of coefficients after the detail coefficients have been trimmed.

$$\widehat{\beta}_{PLR}^{WURR} = (X^T X + \quad {}^{U} I)^{-1} X^T y \tag{11}$$

SURE threshold $W_{SS}(\eta) = \lambda$

$$\widehat{\beta}_{PLR}^{WSRR} = (X^T X + W_{SS}(\eta) I)^{-1} X^T y \tag{12}$$

Each level will have its threshold for wavelet shrinkage. After assessing in (5) and substituting the Constant form threshold (also known as the Universal threshold) method with the tuning parameter from (4), the tuning parameter from (4) is changed for SURE and Minimax in (8) and (9) in the following manner.

Estimating tuning parameters by:

Universal threshold ${}^{U} =$

Minimax threshold $\tilde{R}(F) = \lambda$

DOI: http://doi.org/10.24086/icafs2025/paper.1764

$$\hat{\beta}_{PLR}^{WMRR} = \left(X^T X + \tilde{R}(F)I\right)^{-1} X^T y \tag{13}$$

As demonstrated in (Sym3) and (fk6), the values of (observations of the processed response variable) that might be used in conjunction with the independent variable are extracted via the wavelet matrix to estimate the parameters of the multiple linear regression model.

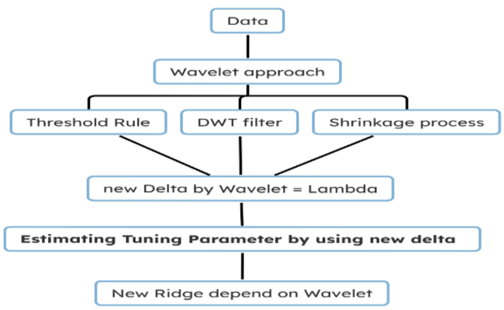

Fig. 1. Suggested Approach (Wavelet shrinkage for penalized techniques).

Finally, the techniques utilized to estimate and analyze the validity of penalized linear regression in Wavelet shrinkage for penalized methods (wavelet Ridge) are presented, as depicted in Fig. 1.

IV. APPLICATION PART

This section compares the procedures used in the estimation process that use Wavelet shrinkage for penalized approaches to traditional penalized methods. The most essential formalization technique for coefficient regression was revealed to be relative efficiency, as measured by the mean squared error and mean absolute error.

Simulation Study

There were several instances in the simulation experiment since two sample sizes (100) and (200) were employed when the parameter numbers (P) ranged between (11 and 41). We also contaminated the (ei) vector without changing the explanatory variables, which could result in outliers. Various levels of the following factors were used to implement the simulation experiments. In this scenario, the original (ei) values are taken from a regular normal distribution with a zero mean and standard deviations of 1 and 5.

The Laplace distribution with location = 2 and Hawkins (1980) used the formula $f(x) = (1-p) * f_1(x) + p * f_2(x)$. With a mean of zero and a standard deviation of one, the explanatory variables do not follow a normal distribution. As soon as the quantity of parameters (P) equals (3 -5 0 0 -0.5 0 0 0.5 5 0 0) the non-zero coefficient numbers where q=5, and the second case (P) equal to (4 8 0 -3 0 3 0 1 0 0.5 0 -8 5 0 3 -0.5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -5) where q=11 is the coefficient numbers of that are not zero. It is now possible to define parameters for the default model by comparing the approaches to the estimating process that each method represents. The wavelet shrinkage was produced for penalized techniques (Wavelet Ridge) and the traditional penalized techniques (Ridge) the assumed regression model for the frequency of 1000 iterations, as well as the instances shown in tables (I, II, III, IV). The MSE and MAE were used to calculate the relative efficiency.

TABLE I
The Average (MAE and MSE) Values for the Suggested and Traditional Approaches. Where (σ = 1) and (P=11)

Wavelet	Criteria	n=100					
		New Method					
		Data Contamination: 10%			Data Contamination: 25%		
		Universal rule	SURE rule	Minimax rule	Universal rule	SURE rule	Minimax rule
Sym3	MAE	0.948	0.997	0.996	1.109	1.167	1.163
	MSE	1.692	1.854	1.857	2.294	2.515	2.520
	q	3.998			3.9980		
fk6	MAE	**0.942**	0.998	0.994	**1.108**	1.166	1.166
	MSE	**1.665**	1.862	1.843	**2.274**	2.516	2.514
	q	3.998			3.998		
Classical Ridge Approach							
MAE		1.191			1.406		
MSE		2.617			3.136		
q		5			5		
n=200							
Sym3	MAE	0.926	0.970	0.963	1.095	1.139	1.136
	MSE	1.539	1.682	1.654	2.112	2.286	2.275
	q	5			5		
fk6	MAE	**0.925**	0.970	0.969	**1.093**	1.135	1.137
	MSE	**1.466**	1.676	1.679	**2.105**	2.268	2.275
	q	3.999			3.999		
Classical Ridge Approach							
MAE		1.147			1.193		
MSE		1.962			2.697		
q		5			5		

DOI: http://doi.org/10.24086/icafs2025/paper.1764

TABLE II
The Mean (MAE and MSE) Values for the Proposed and Existing Techniques Where (σ = 5) and (P=11)

Wavelet	Criteria	(n=100) New Method Data Contamination: 10% Universal rule	Sure rule	Minimax rule	Data Contamination: 25% Universal rule	Sure rule	Minimax rule
Sym3	MAE	3.832	3.841	3.822	3.872	3.882	3.869
	MSE	25.985	26.106	25.787	26.517	26.710	26.433
	q	4			4		
fk6	MAE	**3.818**	3.830	3.837	**3.859**	3.878	3.883
	MSE	**25.778**	25.911	25.984	**26.325**	26.567	26.668
	q	4			4		
Classical Ridge Approach							
MAE		4.396			4.450		
MSE		26.742			27.639		
q		5			5		
n=200							
Sym3	MAE	3.920	3.927	3.931	3.973	3.986	3.991
	MSE	25.565	25.661	25.691	26.277	26.429	26.490
	q	3.992			3.9920		
fk6	MAE	**3.917**	3.931	3.936	**3.955**	3.985	3.986
	MSE	**25.530**	25.680	25.758	**26.021**	26.462	26.434
	q	3.992			3.992		
Classical Ridge Approach							
MAE		4.618			4.582		
MSE		26.394			27.733		
q		5			5		

TABLE III
The Average Values (MAE and MSE) for the Proposed and Existing Techniques. Where (σ = 1) and (P = 41)

Wavelet	Criteria	(n=100) New Method Data Contamination: 10% Universal rule	Surerule	Minimaxrule	Data Contamination: 25% Universal rule	Sure rule	Minimax rule
Sym3	MAE	1.091	1.115	1.117	1.197	1.221	1.302
	MSE	3.158	4.029	4.051	3.807	4.838	4.861
	q	9.998			9.9980		
fk6	MAE	**1.088**	1.118	1.113	**0.925**	1.223	1.312
	MSE	**3.133**	4.057	4.010	**2.741**	4.849	4.874
	q	9.896			9.896		
Classical Ridge Approach							
MAE		3.553			3.607		
MSE		32.097			34.094		
q		11			11		
(n=200)							
Sym3	MAE	1.148	1.212	1.203	1.279	1.344	1.338
	MSE	2.656	3.159	3.110	2.508	3.885	3.841
	q	10			10		
fk6	MAE	**0.862**	0.865	1.147	**1.158**	1.343	1.341
	MSE	**1.505**	1.516	2.644	**2.282**	3.875	3.860
	q	10			10		
Classical Ridge Approach							
MAE		2.337			2.396		
MSE		11.426			12.187		
q		11			11		

DOI: http://doi.org/10.24086/icafs2025/paper.1764

TABLE IV
The Average (MAE and MSE) Values for the Suggested and Traditional Approaches. Where (σ = 5) and (P=41)

Wavelet	Criteria	**(n=100)**					
		New Method					
		Data Contamination: 10%			Data Contamination:25%		
		Universal rule	Sure rule	Minimax rule	Universal rule	Sure rule	Minimax rule
Sym3	MAE	3.210	3.228	3.231	3.239	3.261	3.271
	MSE	27.185	33.485	33.603	27.562	34.157	34.357
	q	9.989			9.989		
fk6	MAE	**3.209**	3.232	3.222	**3.238**	3.249	3.264
	MSE	**27.1798**	33.585	33.349	**27.534**	34.009	34.361
	q	9.989			9.989		
Classical Ridge Approach							
MAE		4.762			4.794		
MSE		72.435			73.437		
q		11			11		
(n=200)							
Sym3	MAE	3.691	3.700	3.688	3.734	3.743	3.737
	MSE	26.714	28.885	28.719	27.391	29.524	29.477
	q	9.999			9.9990		
fk6	MAE	**3.681**	3.698	3.697	**3.729**	3.740	3.754
	MSE	**26.638**	28.588	28.862	**27.343**	29.479	29.723
	q	9.799			9.799		
Classical Ridge Approach							
MAE		4.220			4.236		
MSE		37.386			37.945		
q		11			11		

CONCLUSION

1. Based on the criteria of MAE, MSE, and q, the suggested approach for wavelet kinds (Sym3 and fk6) is superior to the traditional method in all cases which 10% and 25% of contaminants where (σ = 1 and 5) and (P = 11 and 41) for sample sizes 100 and 200.

2. Due to the criterion of (MAE) and (MSE) less than (SURE and Minimax) of wavelet type (Sym3), the proposed threshold approach (Universal) for wavelet type (fk6) is shown by the majority of contamination instances (10% and 25%, respectively).

3. The average (MAE and MSE) indicates that the wavelet type (fk6) is superior to the wavelet type (Sym3) in every scenario and for every sample size.

DOI: http://doi.org/10.24086/icafs2025/paper.1764

REFERENCE

Ali, T. H., & Saleh, D. M. (2021). COMPARISON BETWEEN WAVELET BAYESIAN AND BAYESIAN ESTIMATORSTOREMEDYCONTAMINATION IN LINEAR REGRESSION MODEL. *PalArch's Journal of Archaeology of Egypt/Egyptology*, *18*(10), 3388-3409.

Antoniadis, A. (2007). Wavelet methods in statistics: some recent developments and their applications. *Statistics Surveys*, *1*(none), 16-55, 40. https://doi.org/ 10.1214/07-SS014

Birdawod, H. Q., Khudhur, A. M., Kadir, D. H., & Saleh, D. M. (2024). A Wavelet Shrinkage Mixed with A Single-level 2D Discrete Wavelet Transform for Image Denoising. *Kurdistan Journal of Applied Research*, *9*(2), 1-12. https://eprints.cihanuniversity.edu.iq/id/eprint/2747

Birdawod, H. Q. (2022). Using factor analysis to determine the most important factors affecting student absenteeism at Cihan University-Erbil. *Cihan University-Erbil Scientific Journal*, *6*(2), 1-8.

Donoho, D. L. (1993). Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and nosiy data. *Proc. Symposia in Applied Mathematics*, *47*.

Donoho, D. L., & Johnstone, I. M. (1994). Ideal Denoising in an orthonormal basis chosen from a library of bases. *Comptes rendus de l'Académie des sciences. Série I, Mathématique*, *319*(12), 1317-1322.

Donoho, D. L., & Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the american statistical association*, *90*(432), 1200-1224. https://doi.org/https://doi.org/10.1080/01621459.1995.10476626

Fan, J., & Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the american statistical association*, *96*(456), 1348-1360. https://doi.org/https://doi.org/10.1198/016214501753382273

Hamad, A. S. (2010). *Using Some Thresholding Rules in Wavelet Shrinkage to Denoise Signals for Simple Regression with Application in Rezgary Hospital–Erbil* University of Sulaimania, Iraq].

Hawkins, D. (1980). *Identification of Outliers* (1 ed.). Springer Dordrecht. https://doi.org/https://doi.org/10.1007/978-94-015-3994-4

Helwig, N. E. (2017). Adding bias to reduce variance in psychological results: Atutorial on penalized regression. *The Quantitative Methods for Psychology*, *13*(1), 1-19.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1), 55-67.

Jansen, M., & Bultheel, A. (1999). Multiple wavelet threshold estimation by generalized cross validation for images with correlated noise. *IEEE transactions on image processing*, *8*(7), 947-953. https://doi.org/10.1109/83.772237

Kadir, D. H., Yahya, R. O., Khudhur, A. M., & Ali, T. H. (2024). Data De-Noise for Multivariate T2 and S-Charts Using Multivariate Wavelets. *Zanco Journal of Human Sciences*, *28*(6), 343-360. https://doi.org/https://doi.org/10.21271/zjhs.28.6.19

Li, Z., & Sillanpää, M. J. (2012). Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. *Theoretical and applied genetics*, *125*(3), 419-435. https://doi.org/https://doi.org/10.1007/s00122-012-1892-9

Massoudi, A. H., Birdawod, H. Q., & Raewf, M. B. (2023). Personal digital marketing influence on successful marketing campaign in today's digital age. *Cihan University-Erbil Journal of Humanities and Social Sciences*, *7*(1), 158-165.

Qader, H. M., Birdawod, H. Q., Qader, H. M., & Sedeeq, B. S. (2025). Capability Process to Optimize Specification by Impact of Wavelet Analysis. *Cihan University-Erbil Journal of Humanities and Social Sciences*, *9*(1), 53-60.

Saalih Kahwachi, W. T., & Birdawod, H. Q. (2023). A new hybridization of bilateral and wavelet filters for noisy De-noisy images. *Eurasian Journal of Science and Engineering*, *9*(1), 99-115.

Saleh, D. M., Kadir, D. H., & Jamil, D. I. (2023). A Comparison between Some Penalized Methods for Estimating Parameters. *Qalaai Zanist Journal*, *8*(1), 1122-1134. https://doi.org/https://doi.org/10.25212/lfu.qzj.8.1.44

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *58*(1), 267-288. https://doi.org/https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tutz, G., & Ulbricht, J. (2009). Penalized regression with correlation-based penalty. *Statistics and Computing*, *19*, 239-253. https://doi.org/https://doi.org/10.1007/s11222-008-9088-5

Van der Kooij, A. J. (2007). *Prediction accuracy and stability of regression with optimal scaling transformations*. Leiden University.

Varanis, M., & Pederiva, R. (2017). The influence of the wavelet filter in the parameters extraction for signal classification: An experimental study. *Proceeding series of the Brazilian society of computational and applied mathematics*, *5*(1).

Wood, S. N. (2006). On confidence intervals for generalized additive models based on penalized regression splines. *Australian & New Zealand Journal of Statistics*, *48*(4), 445-464.