Statistical Analysis Based on Lake Michigan Fish Acoustic Data Using LASSO Method

Liming Xie

Department of Statistics, North Dakota State University, 1340 Administration Ave, Fargo, ND 58108, USA

limingxie2018@gmail.com

ABSTRACT

LASSO method is one of the most popular and more extensive regressions. It has been applied to many fields. However, it is rare seen to research with complicated big data in biology. This paper is to apply LASSO method to Lake Michigan Fish acoustic data. The main techniques include: Elastic Net selection, which tests validation from the average square error (ASE) to predict the error for the model by computing separately for each of these subsets; defaulting group LASSO to test multiple parameters by splitting a couple constituent parameters, such as successive intervals, multiple continuous depth layers, to estimate the Schwarz Bayesian information criterion (SBC) to find the lowest value for the model; The adaptive LASSO selection, which is applied to each of the parameters in constructing the LASSO constraint for weights, that is, the response y has mean zero and the regressor x are scaled to have mean zero and common standard deviation. The empirical results show that the fish density (Y) has strong relationships with area backscattering coefficient (PRC_ABC), secondly, significant interactions with PRC_ABC and Exclude below line depth mean), among PRC_ABC, fish density in the intervals and layers of acoustic survey transect of Lake Michigan.

KEY WORDS: Fish acoustic data, Least absolute shrinkage and selection operator (LASSO), area backscattering (PRC_ABC), Average square error (ASE), Schwarz Bayesian Information Criterion (SBC), The Elastic Net, Group LASSO, Adaptive LASSO.

INTRODUCTION

Today, a lot of complicated data need be analyzed by some statistical technical methods. Since 1996 LASSO (least absolute shrinkage and selection operator) was proposed by Robert Tibshirani, it has been very popular now and it has been more expensively used to almost every field. LASSO regression analysis is a shrinkage and variable selection methods for linear regression model. the goal is to receive the subset of predictors that minimizes prediction error for a quantitative response variable. The LASSO is mainly to impose a constraint on the model parameters that regression coefficients for some variables to shrink toward zero. After the shrinkage process is excluded from the model, variables with a regression coefficient is trend to zero. However, those coefficient variables that have non-zero regression are most associated with the response variable. Especially for data with large amounts of predictors, we want to test them that must use more powerful functions. For many years underwater acoustics is the study of the propagation of sound in water and the interaction of the mechanical wave. I would like to study underwater acoustics by using Lake Michigan Fish acoustic data 2011 to 2016. In the data, each line represents the acoustic information that we can estimate fish density for a single depth layer of water. It includes many variables: transects, which are divided horizontally into successive intervals; there are multiple continuous depth layers; area backscattering (ABC), mean acoustic size (sigma); fish density is reported for each unique transect-interval-layer from Lake Michigan in the year 2011-2016; area backscattering (PRC_ABC), mean acoustic size (sigma); the fish density in the intervals and layers of acoustic survey transects of Lake Michigan 2011 to 2016. The source was used by a stratified and systematic design that has transect locations randomized in stratum. Thus, each year investigators get different transect location.

I am to test those variables through using LASSO methods to estimate whether the relationships each other exist among them. Therefore, the model selections are very important.

This data collection comes from “USGS Science for a changing world”, title “Lake Michigan Fish Acoustic Data from 2011 to 2016”. It comes from the following linker: https://www.sciencebase.gov/catalog/item/57ed22cfe4b090825011d3f5 . In raw data, since I plan to make big data analysis by LASSO, there are some minor or nonsignificant variables to be shifted away so that SAS can optimize data. SAS generally requires numeric variables. For example, “transect”, “EV_filename”, and “program_version” are not sensitivity for LASSO regression analysis. Dataset includes 13147 observations that one dependent (total fish density, “fish/ha”) is chosen as Y, 18 independent variables are the following:

X_1= Successive intervals that reports the interval number of the cell being analyzed, X_2=Multiple successive depth layers that presents the layer number of the cell being analyzed, X_3=Mean acoustic size (sigma), X_4 =Depth Mean that denotes the mean depth of the domain being analyzed, X_5= Date of start in the first ping in the domain to be analyzed, X_6=Time of start in the first ping in the domain to be analyzed, X_7=Time of the end in the last ping in the domain to be analyzed, X_8=Time in the middle ping in the domain to be analyzed, X_9=Latitude of start in the first ping in the domain to be analyzed, X_{10} =Longitude of start first ping in the domain to be analyzed, X_{11} = Latitude of the end that represents the last ping in the domain to be analyzed, X_{12}= Longitude of the end that represents the last ping in the domain to be analyzed, X_{13}= Latitude represents the middle ping in the domain to be analyzed, X_{14}= Longitude represents the middle ping in the domain to be analyzed, X_{15}=Exclude below line depth mean X_{16}=Processing date that denotes analysis date, X_{17}=Area Backscattering (ABC) that denotes Area backscattering coefficient for the domain to be analyzed, X_{18}= Year of survey.

MATERIALS AND METHODS

The paper applies regular LASSO, group LASSO, adaptive LASSO, other techniques such as fitting generalized additive models with the GAML procedure, classification and regression tree models with the HPSPLIT procedure, etc. Because of the limit space of the space for journal, I will focus on analysis the dataset by regular LASSO, Group LASSO, adaptive LASSO. In general, GMESELECT procedure is used to analyze coefficient effects of predictors. It supports the mode selection methods: Forward selection, which get starts with no effects and adds effects; Backward elimination which has to be beginning with all effects and deletes effects; Stepwise selection that gets start with no effects that added and be able to be deleted; least angle regression that begins with no effects and adds effects and estimate βs by shrinking to zero; LASSO that is constraining sum of absolute βs; at least one β close to 0. Elastic net is a kind of constrains sums of absolute and squared βs, and at least one β set to 0. Adaptive LASSO is a kind of constraint sum of absolute weighted βs, and at least one β set to 0; Group LASSO is constraint sum of Euclidean norms of βs with effects and all βs for the same effects are set to 0 or βs are probably to be non-zero.

The variable selection is more important for the high-dimensional datasets like Lake Michigan Fish data 2011-2016. When feature selection process, the variables are probably having a non-zero coefficient if the shrinking process are selected to be part of the model. For using LASSO methods to analyze predictors, there is many advantages: it provides accurate prediction. After shrinking and removing the coefficients, variance is reduced without a substantial increase of the bias. Secondly, LASSO can help argument the model interpretability by eliminating irrelevant v variables without the response variables.

Group LASSO method was introduced by Yuan and Lin in 2006. It is mainly to let predefined groups of covariates to be selected into or out of single model. In a specific group all of members are entered or not. When levels of a categorical variable are coded as a collection of covariates are either included or excluded from single model.

Adaptive LASSO method is a particular LASSO technique. It generates consistent estimates of the parameters when retaining the convexity property of the LASSO. The objective is to favor predictors with univariate strength and avert spurious selection of noise predictors. Also, adaptive LASSO can give the correct model under milder conditions than regular LASSO.

The LASSO definition

Suppose that there are data (W^n, Q_n), n=1, 2, …, N, where W^n= $(W_{n1}, \dots, W_{nk})^T$ are the predictor variables and $Q_n$are the response. If there is the usual regression, the observations are independent or that the $Q_n s$ are conditionally independent given the W_{nj}s. If we assume there are W_{nj} for being standardized so that $\sum_i \frac{W_{nj}}{N} = 0, \sum_i \frac{W_{nj}^2}{N} = 1.$

Suppose that we have $(\hat{\theta} = \widehat{\theta_1}, \dots, \widehat{\theta_p})^T$, the LASSO estimate $(\hat{\sigma}, \hat{\theta})$ is regarded as

$$(\hat{\sigma}, \hat{\theta}) = \text{arg min } \{\sum_{i=1}^{N}(Q_i - \sigma - \sum_j \theta_j W_{nj})^2\} \quad \text{subject to } \sum_j |\theta_j| \leq q. \quad (1)$$

where it is q $\geq$ 0 a tuning parameter. $\forall q$, the solution for σ is $\hat{\sigma}$=$\bar{Q}$. We assume without loss of generality that $\bar{Q} = 0$ with omitting σ.

LASSO Selection

We suppose data W=($w_1, w_2, \dots, w_n$) express the matrix of covariates and let Q represents the response, where the $W_i s$ have been centered and scaled to have unit standard deviation and mean 0, and there is mean 0 for Q, so, if parameter t exist, the LASSO regression coefficients θ =$(\theta_1, \theta_2, \dots, \theta_n)$ are the solution to the constrained optimization problem. Thus, we have the following:

$$\text{minimize } || \mathrm{Q} - \mathrm{W}\,\theta||^{2} \text{ subject to } \sum_{j=1}^{n} |\,\theta_j| \le t \tag{2}$$

here, the parameter t should be enough small, which it is exactly zero for some regression coefficients. Therefore, LASSO as selecting a subject of the regression coefficients in which the nonzero coefficients are each step correspond to selected parameters.

The Elastic Net

Suppose a simple and extreme example that the coefficient for a variable Z_j with a specific value for τ is $\hat{\theta}_j$>0. When a data is augmented with an identical copy $W_{j\prime}$=W_j, the coefficient in infinitely a lot of ways could be shared: ∀ $\tilde{\theta}_j$ +$\tilde{\theta}_{j\prime}$=$\hat{\theta}_j$ with both pieces positive and the loss and $\ell_1$penalty is indifferent. Thus, the coefficients for this pair are not defined. Hence, a quadratic penalty could be divided $\hat{\theta}_j$ exactly equally between two twins. In fact, the elastic net takes a compromise between the ridge and the LASSO penalties:

$$\underset{(\theta_0\theta)\,\in\,\mathrm{R}\times\mathrm{R}^p)}{min} \;\{\, \tfrac{1}{2}\textstyle\sum_{i=1}^{N}(Q_i - \theta_0 - \sigma_i^T\theta)^2 + \gamma[\,\tfrac{1}{2}(1-\sigma)||{}_2^2 + \sigma||\,\theta||_1]\,\} \tag{3}$$

Here, we have $\sigma \in [0,1]$ for a parameter that is varied. By construction, the penalty applied to an individual coefficient (here the weight γ >0) should be given by

$$\tfrac{1}{2}(1-\sigma)\theta_j^2 + \sigma|\theta_j|. \tag{4}$$

Here, σ = 1. It is probably to be the ℓ_1-norm or LASSO penalty, and with σ = 0, it reduces to the squared ℓ_2-norm, which is corresponded to the ridge penalty.

The Group LASSO

If there is a linear regression model with X groups of covariates, which is x=1,2, …, X, the vector $V_j \epsilon \mathrm{R}^{pk}$ denotes the covariates in group x. To predict a real-valued response S ∈ R by the collection of covariates $(S_1, \dots, S_k)$, we suppose a linear model for the regression function C (Al S) given by $\theta_0 + \sum_{x=1}^{X} S_k^T\,\theta_k$, which is $\theta_j \in \mathrm{R}^{pk}$dented by a group of $p_j$regression coefficients.

Suppose a collection of T samples $\{(Q_i, s_{i1}, \ldots, s_i, K)\}_{i=1}^{N}$, so, we could obtain the following:

$$\min_{(\theta_0 \in \mathrm{R}, \theta_j \in, \mathrm{R}^{pj}} \{\frac{1}{2}\sum_{i=1}^{N}(Q_i - \theta_0 - \sum_{k=1}^{J} S_{ik}^{T}\, \theta_k)^2 + \gamma \sum_{k=1}^{K} || \,\theta_k ||_2\}, \tag{5}$$

here, $||\theta_k||_2$ is regarded as the Euclidean norm of the vector θ_k.

Adaptive LASSO

Adaptive LASSO is proposed by Zou in 2006. Since LASSO could not be an oracle procedure. He introduced asymptotic setup was somewhat. From LASSO definition and the selection, the coefficients to be equally penalized in the $\ell_1$penalty. If assigning a couple of weights to different coefficients. Suppose the weighted LASSO, we define the weighted LASSO (2):

$$\arg\min_{\beta} || \,y - \sum_{k=1}^{p} \sigma_k \theta_k \,||^2 + \gamma \sum_{k=1}^{p} Q_k |\theta_k| \,, \tag{6}$$

here, Q is a weights vector that the weights are data=dependent and cleverly chosen, it is probably for the weighted lasso to be the oracle properties. That is, if $\hat{\theta}$ is a root-n-consistent estimator to θ^*; here we could use $\hat{\theta}$ (ols). Picking a $\mu > 0$, the weight vector $\hat{Q}$=$\frac{1}{|\,\hat{\theta}|^{\mu}}$. Therefore, we got adaptive LASSO estimates $\hat{\theta}^{*(n)}$ are given by

$$\hat{\theta}^{*(n)} = \arg\min_{\beta} || \,y - \sum_{k=1}^{p} \sigma_k \theta_k \,||^2 + \gamma_n \sum_{k=1}^{p} \hat{Q}_k |\theta_k|. \tag{7}$$

STATISTICAL ANALYSIS

The Elastic Net Selection

The Lake Michigan fish acoustic data is big observations. I reserve observation for training, validation, and testing. A model fit on the training data is scored on the validation and test data. Each data is computed separated by the average square error (ASE) for the error sum of squares for observations in that role divided by the number of observations in the role.

The following table "LASSO Selection Summary" is the STEPS-sub-option of the

SELECTION=option specifies performs 120 steps of LASSO selection, but LASSO method terminates by step 42. Since the selected model is a perfect fit and the number of effects that can be selected by LASSO is bound by the number of training samples. In addition, the table showed that the step 42 is the minimum of the validation ASE. Thus, the model at this step is selected. The effected step is 27 (Table 1).

We can see the following figure that displays the standardized coefficients of all the effects selected at a couple of step of the LASSO method, which is plotted as a function of the step number (Figure 1).

For LASSO method, selecting the number of effects is upper-bonded by the number of training sample, the elastic net method is more worked without a restriction. It can incorporate an additional ridge regression penalty. The following Table "Elastic Net Selection Summary" shows 28 steps due to the final step is effect. Comparing with the regular LASSO method, the elastic net method was more variables and its effected numbers were restricted by the number of samples. The following figure is the standardized coefficients of all the effects that are selected at some step of the elastic net method, plotted as a function of the step number.

To construct a validation data set, the author applies the elastic net method selection. The ridge regression parameter is set to the value that achieves the minimum validation ASE. In the ridge regression parameter validation ASE is higher than the one of the elastic net selection. We can see the following the validation is lowered and optimized on the validation data. In addition, the k-fold cross validation is applied to this paper. It is easy to find that the elastic net method gets the smallest CVPRESS score at step 9 that generates 10 selected effects.

Figure 2 is to reflect that k-fold cross validation used a least square fit to compute the CVPRESS score. Thus, the criterion does not directly depend on the penalty regression used in the elastic net method, and the CVPRESS curve looks like OK. On the other hand, the author uses the elastic net method to compare the selection of both the selected variables in the model and the ridge regression parameter. The result displays that the curve of CVEXPRESS statistic as a function of the step number is smoother than the CVPRESS statistic. However, the CVPRESS statistic is based on an ordinary least squares model; CVEXPRESS statistic is according to a penalized model.

We also see that Table 2.b shows that the smallest CVEXPRESS score at step 26 with 27 selected effects.

Group Lasso selection

It is a variant of Lasso that specific linear model. in this case, there are 18 continuous effects, that is, X1-X18 and 3 CLASS variables, X1-X3. It has multiple degree of freedom. This paper was selected to stopped by the specified number of 20 steps. In Lasso selection, there are multiple parameters that can split into their constituent parameters. In this paper the spline effect has a couple of parameters by default and each of the three CLASS effects have four parameters. The mode had a total of 6727 parameters. In "LASSO Selection Summary" the standardized coefficients of all the effects selected as a couple of steps of the LASSO method, which plotted as a function since we specified CHOOSE=SBC to pick the best model, the SBC value for the model at every step is also shown in the following figure. The number of effects is Step 20. Thus, the model at this step is selected, it is possible that the resulting in 21 selected effects, which is noted that SBC= 2n ln (ACL (τ)) +p ln (n), ACL (τ), (Average check loss); ACL (τ)=D (τ) /n. In Table 3, the Elastic net selection is displayed: The smallest CVEXPRESS score.

In LASSO selection, we can choose the multiple parameters to split into them constitute parameters and the spline have a couple of parameters by defaulting. In this case, in the three

CLASS effects X3 has 13 parameters, three parameters for X1, one parameter for X2, as well as X15*X17. To build the best a group of the continuous effects to use a collection effect, the author applied GROUP LASSO procedure. Group Lasso selection is not split by default. In the following table contained 27 effects and 223 parameters effects. Also, it exhibited the standardized coefficients of all the effects selected at some step of the group Lasso method, plotted as a function of the step number. In this plot, the CHOOSE=SBC option is selected the model at step 16 is the minimum value of the SBC, the resulting model contains the 4 effects with all the true ones, and 89 parameters.

Figure 3 is a spline effect and classification effects. In Figure 4 a group of the three of the continuous effects are constructed. The RHO=0.8 option specifies the value of ρ for determining the regularization parameter $\lambda = \rho^{i}$ as the i th step of the group Lasso selection process. the figure displayed a finer coefficient progression. The group Lasso method add or drop more than one effect. Clearly in figure5, Step7, Step 13, and Step 18 each added two effects to the model. Simple selection breaks down since group Lasso does not accept a piecewise linear constant solution path for a regular Lasso.

In Figure 6. the interaction between X11 and X13; in Figure 7 cross validation details suggest that there are strong interactions between X10 and X11; in Figure 8 non-monotone is increasing, which is the sequence of entry p-values at each step and stopping when all candidate entering effects are not significant at the prespecified SLE value doe not guarantee that.

Adaptive LASSO

For GLMSELECT procedure, the algorithms that make them customized with specifying criteria of some effects are applied. This paper was used to significance level(SLS) for removal and effects not yet in the model whose addition was significant at the entry significance level(SLE) were candidates for addition to the model by the default SLE and SLS value of 0.15. Figure 9 is the stepwise selection process that stopped at an earlier step with using predicted residual sum of squares (PRESS) to assess the selected models as stepwise selection progresses other than Schwarz Bayesian information criterion (SBC). The following figure showed that X17*X2 was optimal value of criterion (Table 4). In addition, “stop=PRESS” statement is used to confirm the above results (Table 5).

Some researchers thought that Adaptive LASSO was better than regular LASSO methods. However, the author does not agree with view. We could compare both to see if there are difference place. We can see out that both selected the same set of predictor variables (X17, X2*X18, and X18*X17), even if the solution paths were little different. Also, the estimated coefficient values are closer. On the other hand, we see “Fit Statistics Tables” by using LASSO and Adaptive LASSO that the ASE of the test data for adaptive LASSO (145251) is slightly greater than LASSO (145073).

Of course, the main advantage of adaptive LASSO over regular LASSO should be its asymptotic consistency. It makes a difference for very large data sets like this paper. But, asymptotic consistency cannot generally automatically result in optimal prediction performance for finite samples. Some researchers thought regular LASSO can still be benefit from difficult prediction problems (Zou 2006).

Table 6 displays the parameter estimates and the fit statistics of the model that was selected by elastic net. Obviously, two parameters (X3*X17 and X2* X18) were so closed to 0. (-0.000613, -7.331039E11). The ASE of the test data was 167246. Also, for the mode evolves the selection process, the author used the QUANTSELECT procedure to help the coefficient plot, the average check loss plot, and a couple of criterion plots in either packed or unpacked forms by which the Lake Michigan Fish data was required by τ=0.1, STOP=AIC criterion, CHOOSE=SBC criterion, and SH=7 option (Not shown due to space limit). Table 7 suggests the LASSO and Adaptive LASSO. In Table 8 "Fit Statistics" shows the penalized log likelihood and the roughness penalty. Information criteria such as Akaike's Information criterion (AIC), Akaike's bias-corrected information criterion(AICC), etc. These criteria penalize the -2log likelihood for effective degree of freedom. Using the GCV criterion is to compare against other generalized additive models that are penalized. Obviously, AIC is the smallest one. In Table "Parameter Estimates" it shows that the regression parameter and dispersion parameter estimates. We can see that the intercept is the only regression parameter since all variables are characterized by spline terms and no parametric effects are presents and the constant effect was absorbed by the intercept. Also, maximizing the likelihood make estimate to the dispersion parameter.

RESULTS

The elastic net, group LASSO, and Adaptive LASSO are the best and accurate analysis and forecasting parameters and models in the LASSO techniques. I believe that the LASSO method is most comprehensive and effective way in the regression analysis to solve the real problem thoroughly. From statistical analysis, the author concludes that fish density is strong relationship with Area Backscattering (ABC). Also, there are significant interactions between "Exclude below line depth mean" and ABC, between ABC and "Multiple successive depth layers that presents the layer number of the cell being analyzed". On the other hand, there are interaction between latitude of the end that represents the last ping in the domain to be analyzed and latitude represents the middle ping in the domain to be analyzed, as well as relative relationship between multiple successive depth layers that presents the layer number of the cell being analyzed and year of survey for Lake Michigan.

DISCUSSION

For sparse underwater acoustic channel equalization, two researchers introduced a family of sparse group Lasso recursive least squares algorithms such as $\ell_1\ell_2, -norm, \ell_1\ell_\infty,$ etc. as the sparsity constraint in the penalty function to develop the sparsity of the underwater acoustic communication system. Their experimental result showed that a direct adaptive decision feedback equalizer receiver with the proposed family of sparse group Lasso RLS algorithms obtained good performance in convergence rate, mean square deviation and symbol error rate (Lu Liu et al, 2017). A couple of researchers used machine leaning techniques to predict the performance of an underwater acoustic network. They displayed a machine-learning model based on logical regression to capture the spatio-temporal variation in the performance of an underwater acoustic network and captured the effect of environmental factors such as wind speed, tide, current velocity (Kalaiarasu, V et al, 2017). One scholar used regression analysis to effective hydro cast in underwater environment. To get data he developed a simple regression model by ns2 simulator and tested without any autocorrelation between them with Durbin Watson analysis by statistical package (Anand, J. V., et al, 2014). A study group applied Logistic Regression Models to underwater cylindrical objective detection. the mathematical mode was based on the size of the cylinder used in the experiment. They obtain effective results for under the circumstance and a mathematical model without enough data (Seo, Y., et al, 2018). Two IEEE members developed Robust Regression for tracking underwater targets. They used gaussian model of noise to advance underwater target tracking. The Monte Carlo simulation showed the robustness of the proposed estimation procedure (Ferial et al, 1995). Some scholars thought that it is important for variable selection to analyze GWAS data with both the lasso and the elastic net and alternative tuning criterion to minimum MSE (Partrik et al, 2013). Researchers solved the questions with a small number of smooth nonzero patches with the latter of different degree of sparsity by using Elastic Net and Elitist Lasso models. They found mode interpretable neurophysiological patter (Deirel et al, 2013). In the supervised learning theory some professional scholars studied that the algorithm is an iterative procedure for the minimization of the regularized empirical error and they solved LASSO, elastic-net and Tikhonov regularization (Emeslo et al, 2011). A study group made the generalization of weight-fused elastic net to perform group variable selection with combining weight-fused LASSO and elastic net (Fu, G. H. et al 2014). To assist the mixed model selection, two researchers used the adaptive LASSO penalized term to propose a two-stage selection procedure to choose both the random and fixed effects. They achieved effective results (Pan, J., et al, 2018).

CONCLUSION

The author think that big data or big sample size data is a complicated and thorny problems. It is very difficult to solve them if using single or two regular LASSO method. Therefore, the author tries to test multiple and comprehensive LASSO methods to analyze and estimate 13147 observations and big dimensions (121 parameters or more) from Lake Michigan fish acoustic data during 2011-2016. The author believes that more research papers regarding multiple to comprehensive researches using LASSO methods will be increasing in the future.

ACKNOWLEDGEMENTS

I thank edits and three reviewers. They gave constructive suggestions and helps.

COMPETING INTERESTS

The author declares no competing or financial interest.

FUNDING

This work is no any funding from any grants.

REFERENCES

1. Liming Xie, Time Series Analysis and Prediction on Cancer Incidence Rates. J Med Discov (2017); 2(3): jmd17030; DOI:10.24262/jmd.2.3.17030
2. Robert Tibshirani, Regression Shrinkage and Selection vis the Lasso. Journal of the Royal Statistical Society (1996), pp.267-288
3. Hui Zou, The Adaptive Lasso and its Oracle Properties. Journal of the American Statistical Association (2006)
4. Lu Liu, D.J. Sun, A family of sparse group Lasso RLS algorithms with adaptive regularization parameters for adaptive decision feedback equalizer in the underwater acoustic communication system. Physical Communication (2017)
5. Vignesh Kalaiarasu, H. Vishnu, Predicting underwater acoustic network variability using machine learning techniques. Ocean-Anchorage (2017)
6. Anand, J V, Titus, S, Regression based analysis of effective hydro cast in underwater environment. TENCON. IEEE Region (2014)
7. Yoojeong Seo, B. On, Underwater Cylindrical Objective Detection Using the Spectra Features of Active Sonar Signals with Logistic Regression Models, Applied Sciences (2018)
8. Ferial El-Hawary, Yuyang Jing, Robust Regression-Based EKF for Track Underwater
9. Targets. IEEE Journal of Oceanic Engineering 20 (1995)
10. Partrik Waldmann, G. Meszaos, Evaluation of the lasso and the elastic net in genome-wide association studies, Frontiers in Genetics 04 (2013)
11. Deirel Paz-Linares, M. Vega-Hernandez, Spatio Temporal EEG Source Imaging with the Hierarchical Bayesian Elastic Net and Elitist Lasso Models, Frontier in Neuroscience, 16

(2017)

12. Emeslo De Vito, V. Umanita, A consistent algorithm to solve lasso, elastic-net and Tikhonov regularization. Journal of Complexity 27(2011)

13. Guang-hui Fu, W.M. Zhangm, Group Variable Selection with Oracle Property by Weight-Fused Adaptive Elastic Net Model for Strongly Correlated Data. Journal of Communications in Statistics-Simulation and Computation 43(2014)

14. Juming Pan, J.F. Shang, Adaptive LASSO for linear mixed model selection via profile log-likelihood, Journal of Communications in Statistics-Theory and Methods 47(2018)

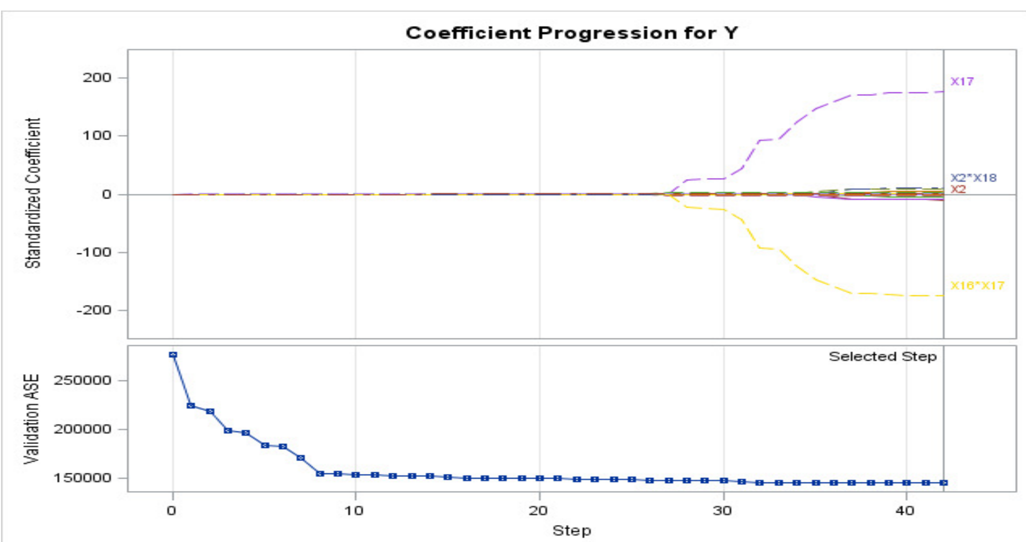
15. Yan Cui, X. Chen, Adaptive Lasso for generalized linear models with a diverging number of parameters. Journal of Communications in Statistics-Theory and Methods 46 (2017)

16. Caiya Zhang, Y.B. Xiang, On the oracle property of adaptive group Lasso in high-dimensional linear models. Statistical Papers, 57(2016)

17. Deborah Gefang (2013) Bayesian doubly adaptive elastic-net Lasso for VAR shrinkage. International Journal of Forecasting 30 (2013)

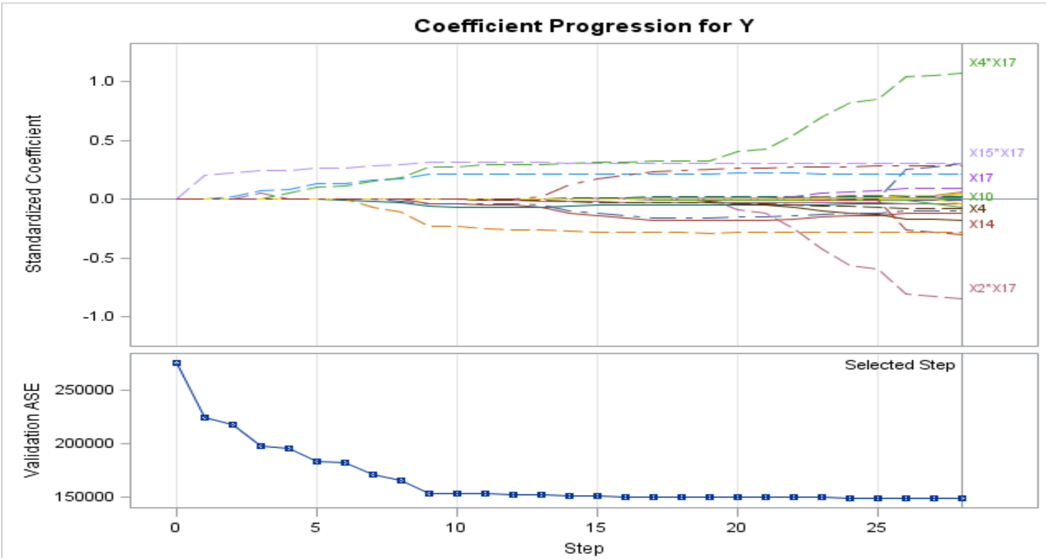
FIGURE LEGENDS

Figure 1.



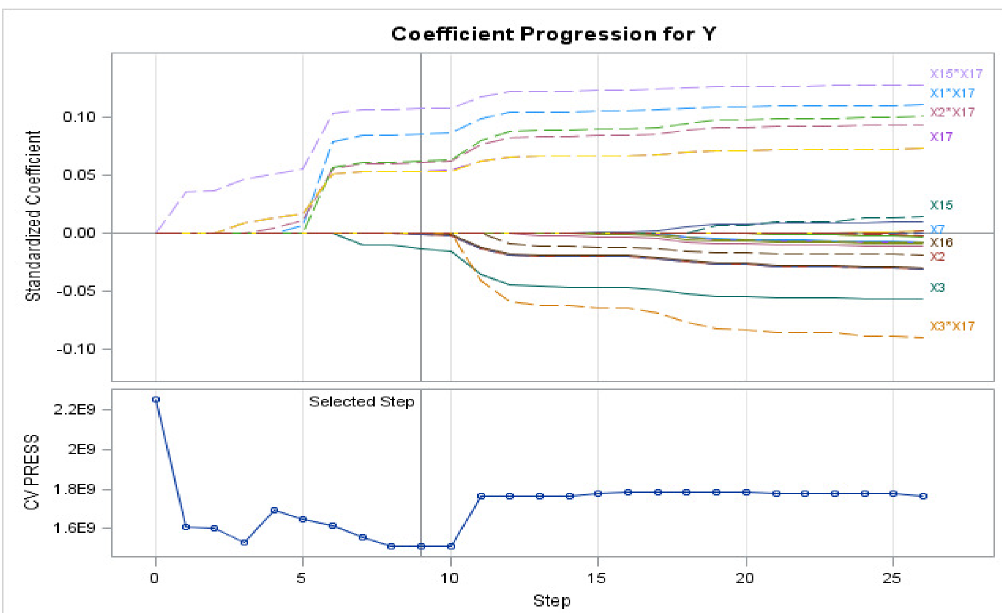
Standardized coefficients of all the effect selection.

Figure 2.



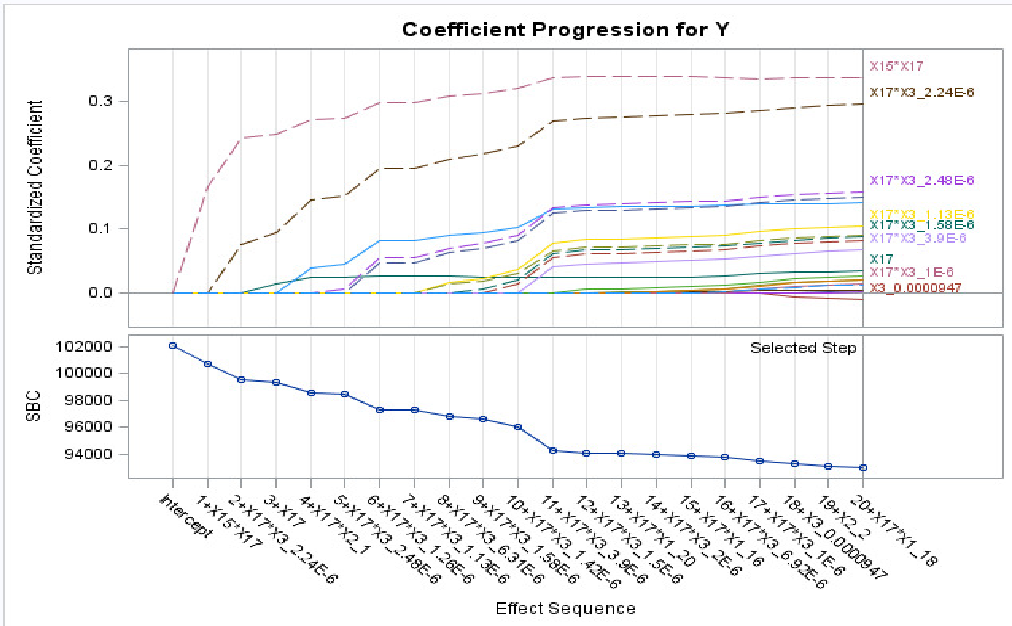
Cross validation using elastic net.

Figure 3



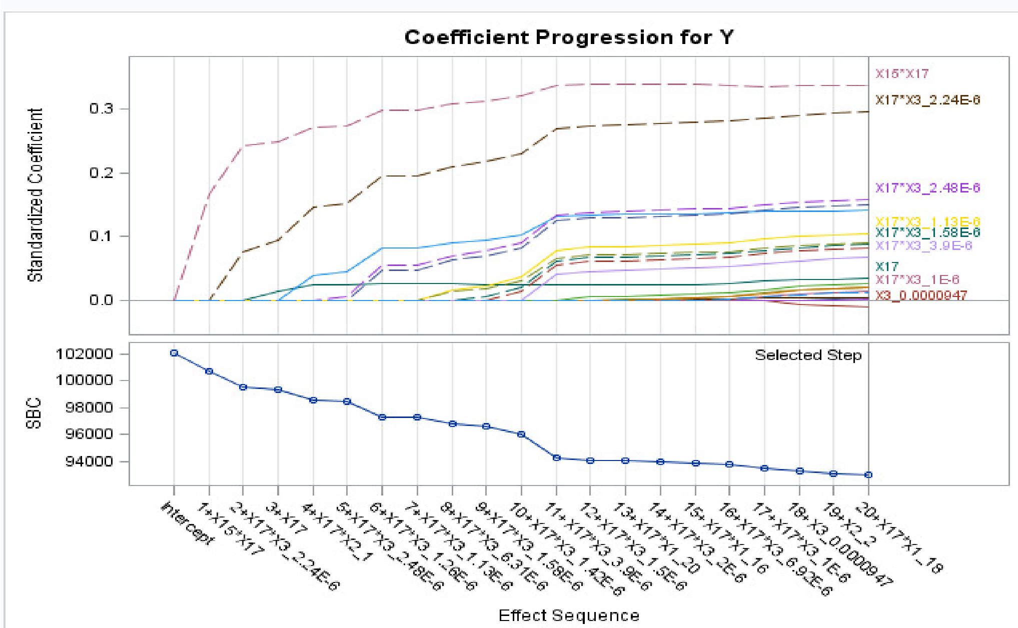
. k-fold cross validation with CVPRESS.

Figure 4



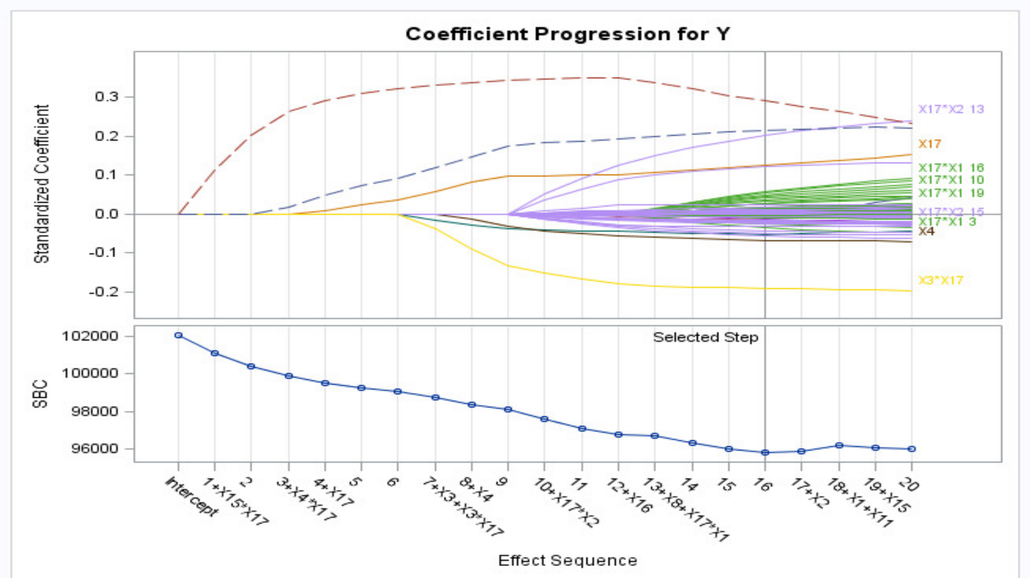
k-fold cross validation with CVEXPRESS.

Figure 5.



Group LASSO Standardized Coefficient: Standardized coefficient with effects Sigma.

Figure 6.



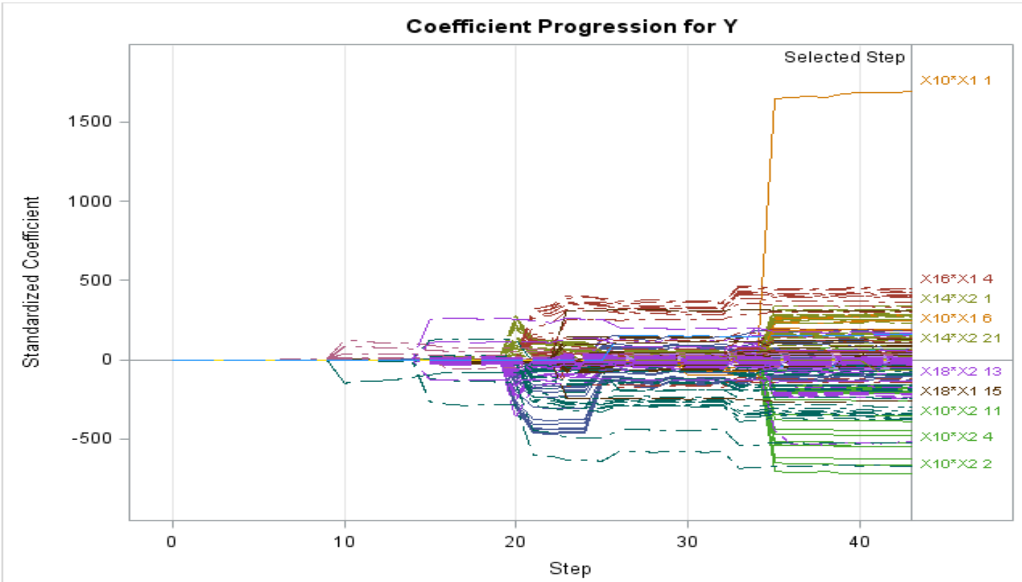
Group LASSO Standardized Coefficient: Standardized coefficient with a spline effect.

Figure 7.

Stop Details					
Candidate For	Effect	Candidate Significance		Compare Significance	
Entry	X11*X13	0.1518	>	0.1500	(SLE)
Removal	X13*X14	0.1290	<	0.1500	(SLS)

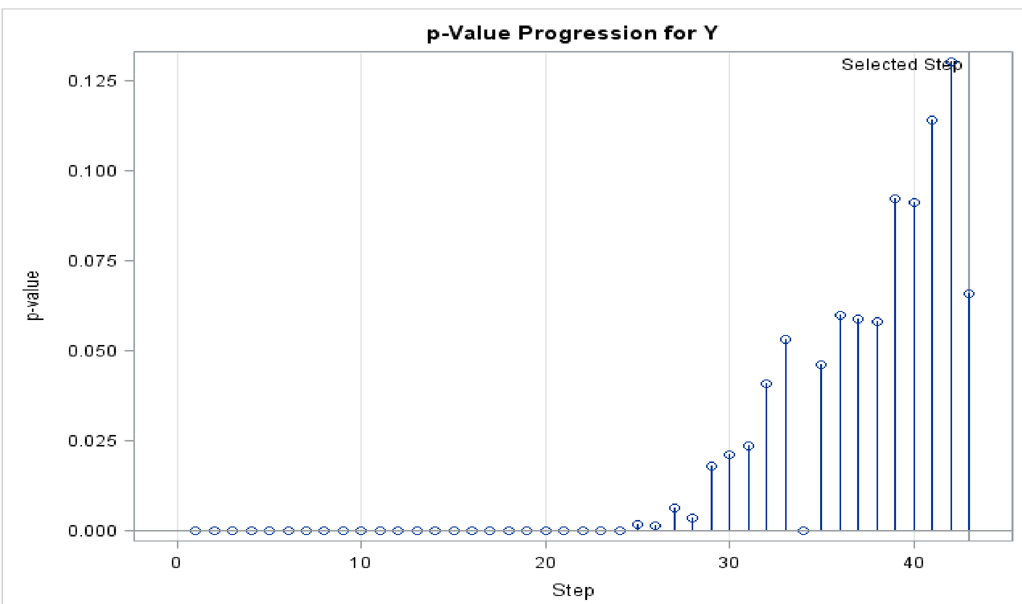
Stop Details statement.

Figure 8.



Cross Validation Details.

Figure 9.



The p-values at each step.

Table 1 LASSO Selection Summary

LASSO Selection Summary					
Step	Effect Entered	Effect Removed	Number Effects In	ASE	Validation ASE
0	Intercept		1	276210.576	276210.576
1	X15*X17		2	224493.993	224493.993
2	X1*X17		3	217588.104	217588.104
3	X2*X17		4	197889.557	197889.557
4	X4*X17		5	195789.419	195789.419
5		X2*X17	4	183012.332	183012.332
6	X3		5	182280.074	182280.074
7	X3*X17		6	170982.722	170982.722
8	X2*X18		7	153824.138	153824.138
9	X2		8	153785.191	153785.191
35	X5		24	145072.643	145072.643
36	X2		25	144999.794	144999.794
37	X9		26	144956.172	144956.172
38		X13	25	144955.885	144955.885
39	X11		26	144953.314	144953.314
40	X13		27	144953.229	144953.229
41		X14	26	144952.868	144952.868
42	X14		27	144952.670	144952.670*
* Optimal Value of Criterion					

Selection stopped because all effects are in the final model.

Table 2. Elastic net selections: K-fold cross validation with CVPRESS score

Elastic Net Selection Summary						
Step	Effect Entered	Effect Removed	Number Effects In	ASE	Test ASE	CV PRESS
0	Intercept		1	276210.576	276210.576	2251171488
1	X15*X17		2	265357.551	265357.551	1608147829
2	X17		3	265296.905	265296.905	1602466577
3	X16*X17		4	257474.397	257474.397	1529131613
4	X2*X17		5	253177.447	253177.447	1695127279
5	X1*X17		6	247113.853	247113.853	1651553662
6	X4*X17		7	196290.254	196290.254	1613818555
7	X3		8	193155.194	193155.194	1554937688
8	X2*X18		9	193138.912	193138.912	1512652347
9	X2		10	192216.612	192216.612	1512122143*
10	X4		11	191420.482	191420.482	1512735192
11	X3*X17		12	179639.952	179639.952	1764291908
12	X16		13	175179.708	175179.708	1766796706
13	X8		14	174332.869	174332.869	1765710780
14	X6		15	174213.318	174213.318	1766331058
15	X1		16	173807.262	173807.262	1777545580
16	X5		17	173770.278	173770.278	1783734492
17	X18		18	172729.821	172729.821	1784045572

Table 3. Elastic net selections: The smallest CVEXPRESS score

Elastic Net Selection Summary						
Step	Effect Entered	Effect Removed	Number Effects In	ASE	Test ASE	CVEX PRESS
0	Intercept		1	276210.576	276210.576	276345.602
1	X15*X17		2	261397.571	261397.571	260277.133
2	X17		3	261310.548	261310.548	260197.045
3	X16*X17		4	252633.467	252633.467	252456.365
24	X14		25	164089.702	164089.702	182231.785
25	X2*X4		26	163516.023	163516.023	181785.834
26	X13		27	163274.911	163274.911	181597.307*
* Optimal Value of Criterion						

Table 4. Group LASSO selection with 27 effects and 223 parameters

Group LASSO Selection Summary

Step	Effect Entered	Effect Removed	Number Effects In	Number Parms In	SBC
0	Intercept		1	1	102082.107
1	X15*X17		2	2	101102.085
2			2	2	100399.661
3	X4*X17		3	3	99891.158
4	X17		4	4	99507.962
5			4	4	99244.410
6			4	4	99071.153
7	X3 X3*X17		6	6	98736.492
8	X4		7	7	98380.322
9			7	7	98121.751
10	X17*X2		8	32	97613.169
11			8	32	97090.175
12	X16		9	33	96739.311
13	X8 X17*X1		11	85	96706.996
14			11	85	96279.327
15			11	85	95987.071
16			11	85	95790.434*
17	X2		12	110	95869.517
18	X1 X11		14	162	96178.614
19	X15		15	163	96054.619
20			15	163	95960.304
* Optimal Value of Criterion					

Table 5. Adaptive LASSO: GLMSELCT procedure with stay significant level and entry significant level.

Stop Details					
Candidate For	Effect	Candidate Significance		Compare Significance	
Entry	X2	0.1547	>	0.1500	(SLE)
Removal	X16	0.0346	<	0.1500	(SLS)

Stepwise Selection Summary							
Step	Effect Entered	Effect Removed	Number Effects In	Number Parms In	PRESS	F Value	Pr > F
0	Intercept		1	1	2250840088	0.00	1.0000
1	X17*X2		2	26	1566731423*	272.08	<.0001
2	X17*X1		3	76	3188205695	36.18	<.0001
3	X3*X17		4	77	3138509263	535.19	<.0001
4	X18*X2		5	102	3034393851	10.18	<.0001
5	X1		6	152	4300637383	4.90	<.0001
6	X15*X17		7	153	4156715546	101.31	<.0001
7	X4*X17		8	154	4263627241	130.14	<.0001
8	X16*X17		9	155	4043917287	105.38	<.0001
9	X3		10	156	6933476622	25.74	<.0001
10	X8		11	157	6753788984	20.92	<.0001
11	X15		12	158	6641603778	8.49	0.0036
12	X16		13	159	6537948806	4.46	0.0346
* Optimal Value of Criterion							

Table 6. Adaptive LASSO: Stepwise selection with stop=PRESS

Stepwise Selection Summary							
Step	Effect Entered	Effect Removed	Number Effects In	Number Parms In	PRESS	F Value	Pr > F
0	Intercept		1	1	2250840088	0.00	1.0000
1	X17*X2		2	26	1566731423*	272.08	<.0001
* Optimal Value of Criterion							

Selection stopped at a local minimum of the PRESS criterion.

Stop Details				
Candidate For	Effect	Candidate PRESS		Compare PRESS
Entry	X17*X1	3188205695	>	1566731423
Removal	X17*X2	2250840088	>	1566731423

Table 7. LASSO and Adaptive LASSO

LASSO

Root MSE	381.44629
Dependent Mean	138.45983
R-Square	0.4748
Adj R-Sq	0.4733
AIC	105024
AICC	105024
SBC	97043
ASE (Train)	145073
ASE (Test)	145073

Root MSE	409.23332
Dependent Mean	138.45983
R-Square	0.3945
Adj R-Sq	0.3938
AIC	106157
AICC	106157
SBC	98085
ASE (Train)	167246
ASE (Test)	167246
CVEX PRESS	187388

Adaptive LASSO

Root MSE	381.49305
Dependent Mean	138.45983
R-Square	0.4741
Adj R-Sq	0.4732
AIC	105018
AICC	105018
SBC	96981
ASE (Train)	145251
ASE (Test)	145251

Parameter Estimates		
Parameter	DF	Estimate
Intercept	1	83.049635
X2	1	-1.212169
X3	1	-183280
X17	1	7719788
X1*X17	1	12985420
X2*X17	1	25749336
X3*X17	1	-7.331039E11
X4*X17	1	4515462
X15*X17	1	2771735
X16*X17	1	0.349767
X2*X18	1	-0.000613

Table 8. Fit Statistics with penalized log likelihood and the roughness penalty.

Fit Statistics	
Penalized Log Likelihood	-1479600
Roughness Penalty	2837686
Effective Degrees of Freedom	46.29069
Effective Degrees of Freedom for Error	8097.75991
AIC (smaller is better)	121606
AICC (smaller is better)	121607
BIC (smaller is better)	121930
GCV (smaller is better)	177766

Parameter Estimates

Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	138.459825	4.686865	872.7356	<.0001
Dispersion	1	178963	257694		

Estimates for Smoothing Components

Component	Effective DF	Smoothing Parameter	Roughness Penalty	Number of Parameters	Rank of Penalty Matrix	Number of Knots
Spline(X1)	6.06617	642.2	381678	9	10	51
Spline(X2)	1.00000	7.0879E8	0.3639	9	10	25
Spline(X3)	8.71983	2.96E-13	748840	9	10	2000
Spline(X4)	7.75137	11460.8	326769	9	10	2000
Spline(X12)	6.50043	0.1463	472748	9	10	1197
Spline(X15)	5.82681	224561	702626	9	10	1274
Spline(X16)	6.47888	6.0693E8	166662	9	10	18
Spline(X17)	1.00000	0.9997	2.555E-7	9	10	2000
Spline(X18)	1.00000	5.78E129	153E-127	5	6	6

Tests for Smoothing Components

Component	Effective DF	Effective DF for Test	F Value	Pr > F
Spline(X1)	6.06617	7	38.56	<.0001
Spline(X2)	1.00000	1	5.21	0.0225
Spline(X3)	8.71983	9	380.31	<.0001
Spline(X4)	7.75137	8	39.94	<.0001
Spline(X12)	6.50043	7	126.39	<.0001
Spline(X15)	5.82681	7	73.06	<.0001
Spline(X16)	6.47888	7	152.87	<.0001
Spline(X17)	1.00000	1	2774.99	<.0001
Spline(X18)	1.00000	1	5.02	0.0250