

Peer-reviewed research

Carbon Prices Forecasting Using Group Information

Xiaohang Ren¹, Kang Yuan¹, Lizhu Tao^{2a}, Cheng Yan³¹ School of Business, Central South University, China, ² College of Mathematics, Sichuan University, China, ³ Essex Business School, University of Essex, UK

Keywords: Carbon return predictability, High-dimensional models, Group structure, JEL C52 C53 Q43

<https://doi.org/10.46557/001c.36615>

Energy RESEARCH LETTERS

Vol. 4, Issue 4, 2023

We select 44 macroeconomic variables as predictors and employ multiple statistical models to forecast EU carbon futures price returns. The predictors in this study are high-dimensional and have the group structure, and we find that, in this case, the accuracy of the high-dimensional models for forecasting carbon prices are higher than traditional time series models. In addition, the introduction of group structure variables into the high-dimensional model improves forecasting performance.

I. Introduction

The issue of greenhouse gas emissions is now receiving widespread global attention. According to estimates from relevant studies, the harm caused by each ton of CO₂ is worth approximately US\$50 (Revesz et al., 2017). In order to alleviate the worsening environmental problems, the Kyoto Protocol, established in Kyoto, Japan, in 1997, brought the carbon emission rights market to reality. As a derivative market of the carbon spot market, the carbon futures market theoretically has the function of price forecasting and risk hedging, and has a decisive impact on the rational allocation of carbon resources, risk reduction and mitigation of climate change. The European Union Emissions Trading Scheme (EU ETS) is the earliest and largest carbon emission reduction market. This paper examines the international carbon futures market by taking the EU carbon allowance futures, the main carbon contract futures under the EU ETS, as the object of study.

Accurate and stable carbon price forecasts are important for policy making by government departments and decision making by market participants (Ren et al., 2022; Xiong et al., 2019). However, there are still relatively few studies on carbon price forecasting. For example, Benz and Truck (2009) use Markov switching and AR-GARCH models to forecast the EUA price, based on historical data. Chevallier (2009) considers seven macroeconomic variables as predictors of carbon prices and uses a variety of GARCH-type models for forecasting. Byun and Cho (2013) employ GARCH-type models to study the forecasting of carbon price volatility and include several energy futures prices as predictors. Hong et al. (2017) construct a multiple linear regression forecasting model to predict carbon prices using seven variables from the stock and energy markets.

The main contributions of this paper include the following: Firstly, we introduce a large number of macroeconomic variables as predictors of the carbon price (EU carbon allowance futures), whereas existing studies have typically included only a small number of exogenous predictor variables (Byun & Cho, 2013; Chevallier, 2009; Hong et al., 2017). These macroeconomic variables, 44 in total, include information on energy commodities, financial markets and economic activity in different countries. In this respect, our paper extends the study on carbon price predictors. Secondly, we use six different statistical models to forecast carbon prices: the ARMAX, GARCHX, LASSO, Adaptive LASSO, the Group SCAD, and the Group LASSO models. Benz and Truck (2009) and Byun and Cho (2013) compare the performance of several traditional statistical models for predicting carbon prices, but do not introduce high-dimensional statistical models. Comparing the prediction accuracy between high-dimensional and traditional statistical models, we find that the high-dimensional statistical model is more suitable for carbon price prediction with high-dimensional predictor variable inputs. Finally, by analyzing and comparing the prediction performance of different high-dimensional models, we find that carbon price prediction using group high-dimensional models is better when the input high-dimensional data has a distinct group structure.

The rest of the paper is structured as follows: Section 2 presents our use of methods and data; Section 3 reports our empirical results. Section 4 concludes.

^a Corresponding author:
lizhutao@scu.edu.cn

II. Methodology and data

A. The candidate models

We introduce exogenous inputs to extend the ARMA model (Box et al., 2015; Whittle, 1953) to the ARMAX model, which is constructed as follows:

$$Y_t = c + \varepsilon_t + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{i=1}^q \beta_i \varepsilon_{t-i} + \sum_{i=1}^g \gamma_i X_{t-i}, \quad (1)$$

where α_i , β_i and γ_i are the parameters of the observation Y_{t-i} , white noise error terms ε_{t-i} and the exogenous covariate X_{t-i} , respectively. c is a constant.

In this paper, we follow the ARMAX (1,1,1) models to understand and predict our carbon price return data

$$p_t = c + \varepsilon_t + \alpha p_{t-1} + \beta \varepsilon_{t-1} + \gamma X_{t-1}, \quad (2)$$

As with the ARMAX model, we construct the GARCHX model as follows:

$$y_t = \mu_t + \varepsilon_t, \varepsilon_t | \Psi_{t-1} \sim N(0, \sigma_t^2), \quad (3)$$

$$\sigma^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{i=1}^g \gamma_i X_{t-i}, \quad (4)$$

where Ψ_{t-1} represents a collection of information until time $t-1$. ε_t is the error term at time t . The expectation value of y_t at time t is μ_t . σ_t^2 is the conditional variance, X is the exogenous covariate. ω , α_i , β_i and γ_i are used as corresponding parameters.

In this paper, we employ GARCHX (1,1,1) to predict the carbon price. The models are as follows:

$$p_t = \mu_t + \varepsilon_t, \varepsilon_t | \Psi_{t-1} \sim N(0, \sigma_t^2) \quad (5)$$

$$\sigma^2 = \omega + \alpha p_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma X_{t-1} \quad (6)$$

Following Tibshirani (1996), the LASSO estimator is constructed as follows:

$$\hat{\beta}^{LASSO} = \arg \min_{\beta} \left\{ \frac{1}{N} \|Y - X\beta\|_2^2 + \lambda \sum_{i=1}^N |\beta_i| \right\}, \quad (7)$$

where λ is the regularization parameter, and for sparsity the penalty $\sum_{i=1}^N |\beta_i|$ is adopted.

The carbon price prediction model with LASSO is:

$$\hat{p}_{t+1} = \hat{\beta}_0^{LASSO} + \sum_{i=1}^N \hat{\beta}_0^{LASSO} x_{i,t} \quad (8)$$

Here

$$\begin{aligned} \hat{\beta}^{LASSO} = \arg \min_{\beta} & \left\{ \frac{1}{t-1} \sum_{i=1}^{t-1} \left(p_{t+1} - \beta_0 - \sum_{i=1}^n \beta_{i,x_{i,l}} \right)^2 \right. \\ & \left. + \lambda_{cv} \sum_{i=1}^N |\beta_i| \right\}, \end{aligned} \quad (9)$$

In reference to the approach of Zou (2006), Adaptive LASSO is given by

$$\begin{aligned} \hat{\beta}^{adapt} = \arg \min_{\beta} & \left\{ \frac{1}{N} \|Y - X\beta\|_2^2 + \lambda \sum_{i=1}^n \frac{|\beta_i|}{|\hat{\beta}_{init,i}|} \right\}. \end{aligned} \quad (10)$$

The carbon price prediction model with Adaptive LASSO is

$$\hat{p}_{t+1} = \hat{\beta}_0^{adapt} + \sum_{i=1}^N \hat{\beta}_0^{adapt} x_{i,t}. \quad (11)$$

Here

$$\begin{aligned} \hat{\beta}^{LASSO} = \arg \min_{\beta} & \left\{ \frac{1}{t-1} \sum_{i=1}^{t-1} \left(p_{t+1} - \beta_0 - \sum_{i=1}^n \beta_{i,x_{i,l}} \right)^2 \right. \\ & \left. + \lambda_{cv} \sum_{i=1}^N \frac{|\beta_i|}{|\hat{\beta}_{init,t}|} \right\}, \end{aligned} \quad (12)$$

B. Group method

In practical terms, the parametric vector in a high-dimensional regression model has a group structure $\{g_1, g_2, \dots, g_q\}$ and an index number $\{1, 2, \dots, p\}$. That is, $\bigcup_{j=1}^q g_j = \{1, 2, \dots, p\}$ and $g_i \cap g_k = \emptyset$. We combine the group structure with two methods, LASSO and SCAD, respectively, and evaluate their performance in predicting carbon prices.

Following Yuan and Lin (2006), the group LASSO estimator is

$$\begin{aligned} \hat{\beta}^{gLASSO} = \arg \min_{\beta} & \left\{ \frac{1}{N} \|Y - X\beta\|_2^2 + \lambda \sum_{j=1}^N m_j \|\beta_{gj}\|_2 \right\}, \end{aligned} \quad (13)$$

where standard Euclidean norm is $\|\beta_{gj}\|_2$ and $\|\beta_{gj}\|_2 = \left(\sum_{l=1}^k \beta_{gj,l}^2 \right)^{\frac{1}{2}}$. The multiplier m_j is designed to balance the size of the different groups and is usually set to

$$m_j = \sqrt{T_j}, \quad (14)$$

where T_j indicates the number of parameters in j th group.

The carbon price prediction model with Group LASSO is

$$\hat{p}_{t+1} = \hat{\beta}_0 + \sum_{i=1}^N \hat{\beta}_0 x_{i,t}. \quad (15)$$

Here

$$\begin{aligned} \hat{\beta} = \arg \min_{\beta} & \left\{ \frac{1}{t-1} \sum_{i=1}^{t-1} \left(p_{t+1} - \beta_0 - \sum_{i=1}^n \beta_{i,x_{i,l}} \right)^2 \right. \\ & \left. + \lambda_{cv} \sum_{j=1}^N m_j \|\beta_{gj}\| \right\}. \end{aligned} \quad (16)$$

The group SCAD estimator is

$$\begin{aligned} \hat{\beta}^{gSCAD} = \arg \min_{\beta} & \left\{ \frac{1}{N} \|Y - X\beta\|_2^2 + \lambda \sum_{j=1}^N \|\beta_{gj}\|_2 \right\}, \end{aligned} \quad (17)$$

where β_{gj} are the parameters in g_j th group, and the penalty $P_{\lambda}(\cdot)$ is

$$P_{\lambda}(|x|) = \begin{cases} \lambda|x|, & \text{if } |x| \leq \lambda. \\ -\frac{(|x|^2 - 2a\lambda|x| + \lambda^2)}{2(a-1)}, & \text{if } \lambda < |x| < a\lambda. \\ \frac{(a+1)\lambda^2}{2}, & \text{if } |x| > a\lambda. \end{cases} \quad (18)$$

The carbon price prediction model with Group SCAD is

$$\hat{p}_{t+1} = \hat{\beta}_0 + \sum_{i=1}^N \hat{\beta}_0 x_{i,t}. \quad (19)$$

Here

$$\hat{\beta} = \arg \min_{\beta} \left\{ \frac{1}{t-1} \sum_{i=1}^{t-1} \left(p_{t+1} - \beta_0 - \sum_{i=1}^n \beta_{i,x_i,t} \right)^2 + \lambda_{cv} \sum_{j=1}^N m_j \|\beta_{gj}\| \right\}. \quad (20)$$

C. Out-of-sample Comparisons

Using the full dataset usually leads to over-fitting, hence we use out-of-sample performance tests to objectively evaluate the predictive power of statistical models. Specifically, we split the sample into an in-sample dataset and an out-of-sample dataset. The statistical model is trained on the in-sample dataset and the predictions are then compared to the out-of-sample dataset.

Referring to Campbell and Thompson (2008) and similar approaches in the extensive literature that exists, we use the mean-squared prediction error ($MSPE$), the mean absolute prediction error ($MAPE$), the R^2 statistic of mean-squared prediction error (R^2_{MSPE}) and the R^2 statistic of the absolute value of prediction error (R^2_{MAPE}) to evaluate the out-of-sample performance of the prediction models.

R^2_{MSPE} is defined as follows:

$$R^2_{MSPE} = 1 - \frac{MSPE_C}{MSPE_B}, \quad (21)$$

and

$$MSPE_B = \frac{1}{q} \sum_{i=1}^q (r_{m+i} - \hat{r}_{m+i}^B)^2, \quad (22)$$

$$MSPE_C = \frac{1}{q} \sum_{i=1}^q (r_{m+i} - \hat{r}_{m+i}^C)^2, \quad (23)$$

where $MSPE_B$ and $MSPE_C$ are $MSPE$ of the benchmark and candidate models respectively. r_{m+i} denotes the true value in dataset at the time $m+i$ while $\hat{r}_{m+i}^B$ and $\hat{r}_{m+i}^C$ represents the out-of-sample predicted value of the benchmark and the candidate model at the time $m+i$, respectively.

Likewise, R^2_{MAPE} is defined as follows

$$R^2_{MAPE} = 1 - \frac{MAPE_C}{MAPE_B}, \quad (24)$$

and

$$MAPE_B = \frac{1}{q} \sum_{i=1}^q |r_{m+i} - \hat{r}_{m+i}^B|, \quad (25)$$

$$MAPE_C = \frac{1}{q} \sum_{i=1}^q |r_{m+i} - \hat{r}_{m+i}^C|, \quad (26)$$

where $MAPE_B$ and $MAPE_C$ are the $MAPE$ of the benchmark and candidate models respectively. r_{m+i} denotes the true value at the time $m+i$ while $\hat{r}_{m+i}^B$ and $\hat{r}_{m+i}^C$ represents the out-of-sample predicted value of the benchmark and the candidate models at the time $m+i$, respectively.

D. Data

The EU carbon futures price is the target of our attention. The monthly closing price of the ICE ECX EUA futures continuous contract are selected as a sample for our study.

We make a logarithmic difference between the month-end closing prices of carbon futures to obtain carbon futures returns as the core dependent variable in this study. Our sample period is from January 2011 to December 2020. Data is obtained from DataStream. Descriptive statistics for carbon futures returns are shown in [Table 1](#).

Positive mean monthly price returns on carbon futures imply an upward trend in carbon prices over the sample period. The negative skewness means that the carbon price return series is left-skewed and the kurtosis value indicates that the return series is platykurtic. The results of the ADF and J-B tests indicate that the carbon futures return data are stationary and non-normal.

We adopt 44 macroeconomic variables as predictor variables, reflecting the situation of energy markets, financial markets, and economic fundamentals. These variables have the potential to predict carbon price. The data source is DataStream. Furthermore, we group the 44 macroeconomic variables into energy groups, energy price groups, stock index groups, monetary policy groups and economic information groups. Each group consists of identical indicators from different regions. For data stationarity, log returns are calculated for all macroeconomic variables.

III. Empirical results

A. Baseline Out-of-sample Forecasting Results

[Table 2](#) shows the performance of each statistical model for out-of-sample predictions after training on the in-sample dataset, including two predictive evaluation indicators, $MSPE$ and $MAPE$. By definition [of $MSPE$ and $MAPE$], a smaller value of $MSPE$ and $MAPE$ indicates a more accurate prediction of the statistical model, while the opposite indicates a greater bias in the prediction. Six-month and twelve-month forecasts are reported in [Table 2\(a\)](#) and [Table 2\(b\)](#) respectively.

From the results in [Table 2](#), we find that among the candidate models, the LASSO and Adaptive LASSO models for high-dimensional variables generally outperform the traditional time-series statistical models (the ARMAX (1,1,1) and the GARCHX (1,1,1)) in the six-month prediction test. However, in the twelve-month test the high-dimensional prediction models outperform the ARMAX (1,1,1) model but are weaker than the GARCHX (1,1,1) model.

For the four high-dimensional forecasting models, the Group high-dimensional forecasting model had significantly smaller values for two of the forecasting indicators than it did for the other models. The only exception is that the Group SCAD model had slightly higher $MAPE$ values than the LASSO model in the six-month forecasting test while the empirical results are fully consistent with our expectations in the twelve-month forecasting test.

B. Comparison of the predictive performance between different models

From [Table 2](#), we can briefly conclude that high-dimensional models (LASSO, Adaptive LASSO, Group SCAD and Group LASSO) exhibit better forecasting performance than traditional time-series models (the ARMAX (1,1,1) and the

Table 1. Descriptive statistics for the carbon price return data

Mean	Stdev	Skewness	kurtosis	ADF	Jarque-Bera
0.0065	0.1373	-0.6408	1.5527	-4.3550***	0.0000***

Notes: This Table reports summary statistic for the response variable —the monthly carbon price returns, and the sample period runs from Jan 2011 to December 2020. The ADF shows the value of the Augmented Dickey-Fuller (ADF) test with the null hypothesis of non-stationarity. Jarque-Bera shows the p-values of the Jarque-Bera test with the null hypothesis of normality. ***, **, and * denote, respectively, statistically significance at 1%, 5% level, and 10% levels.

Table 2. The MSPE and MAPE in the six-month forecast and twelve-month forecast

(a) six-month forecast			(b) twelve-month forecast		
forecasting model	MSPE	MAPE	forecasting model	MSPE	MAPE
ARMAX (1.1.1)	0.1057	0.2692	ARMAX (1.1.1)	0.1701	0.3057
GARCHX (1.1.1)	0.0321	0.1358	GARCHX (1.1.1)	0.0264	0.1333
LASSO	0.0122	0.0941	LASSO	0.0468	0.1476
Adaptive LASSO	0.0222	0.1484	Adaptive LASSO	0.0444	0.146
Group SCAD	0.0115	0.0942	Group SCAD	0.0153	0.1086
Group LASSO	0.0089	0.0816	Group LASSO	0.0157	0.1079

Notes: Table 2 reports MSPE and MAPE values of the six-month and twelve-month forecasts. The smaller MSPE and MAPE values indicate higher prediction accuracy of the forecast model.

GARCH (1,1,1) for carbon price forecasts that take into account a large number of other variables. This result inspired us to make a specific comparison of the predictive performance of different high-dimensional models. Specifically, we use the traditional time series model as the baseline model and the remaining high-dimensional model as the candidate model. The predictive performance of the candidate models is evaluated by calculating the R^2_{MSPE} and R^2_{MAPE} values of the candidate models relative to the benchmark model; larger values of R^2_{MSPE} and R^2_{MAPE} indicate stronger predictive performance of the candidate models compared to the benchmark model, with positive values indicating that the candidate models have better predictive power than the benchmark model, and inversely, that the candidate models have weaker predictive performance than the benchmark model.

Table 3 displays the results of using ARMAX (1, 1, 1) and GARCHX (1, 1, 1) models as the baseline model to obtain values for R^2_{MSPE} and R^2_{MAPE} . Table 3 (a) represents the result for ARMAX (1, 1, 1) as the baseline model and Table 3 (b) represents the result for GARCHX (1, 1, 1) as the baseline model. Panel A and Panel B in Table 3 are the six-month prediction test and the twelve-month prediction test respectively. The first column of each table indicates which candidate model has been used.

Table 3 shows that the R^2_{MSPE} and R^2_{MAPE} values of the high-dimensional model with group structure are positive in both the six-month test and the twelve-month test when the baseline model is the ARMAX (1,1,1) and GARCH (1,1,1) models, indicating that the accuracy of the group high-dimensional model in predicting carbon prices out-of-sample is better than that of the traditional time series model in both tests.

On the other hand, we also consider the performance of LASSO and Adaptive LASSO models. When the ARMAX (1,1,1) is used as the benchmark model, their R^2_{MSPE} and

R^2_{MAPE} values are positive in both tests. When the GARCHX (1,1,1) is used as the benchmark model, their R^2_{MSPE} and R^2_{MAPE} values are positive in the six-month tests. It is, however, worth noting that both LASSO and Adaptive LASSO have negative R^2_{MSPE} and R^2_{MAPE} in the twelve-month forecast test, implying that high-dimensional models do not outperform traditional time-series models in carbon price forecasting in all cases.

Finally, we compare the R^2_{MSPE} and R^2_{MAPE} values of the group high-dimensional model with other high-dimensional models. In most cases, the R^2_{MSPE} and R^2_{MAPE} values of the group high-dimensional model are larger than those of the other models, except for a few cases where the R^2_{MAPE} values of the Group SCAD model are close to those of the GARCH model. In particular, comparing the LASSO, Adaptive LASSO and Group LASSO models, the R^2_{MSPE} and R^2_{MAPE} values of the Group LASSO model are larger in any case, indicating that Group LASSO has better carbon price forecasting performance under our forecasting conditions. This implies that the use of group structure variables in our study to optimize the forecasting model is positive, a result that is also evidenced in Table 2.

IV. Conclusion

This study uses data from 2011 to 2020, including EU carbon allowance futures returns and 44 macroeconomic indicators, to investigate the prediction of carbon prices using six statistical models. In particular, we propose two-group high-dimensional models for carbon price forecasting. The results of the study show that the high-dimensional model has better predictive power for carbon prices than the traditional linear model, suggesting the usefulness of incorporating a variety of macroeconomic factors into the prediction of carbon prices. In addition, in the group high-dimensional models, the high-dimensional

Table 3. The MSPE and MAPE in the six-month forecast and twelve-month forecast

(a) ARMAX (1,1,1) as the benchmark model			(b) GARCHX (1,1,1) as the benchmark model		
R^2_{MSPE}	R^2_{MAPE}	forecasting model	R^2_{MSPE}	R^2_{MAPE}	
Panel A: Six-month forecast			Panel A: Six-month forecast		
0.8845	0.6504	LASSO	0.7248	0.5171	
0.7899	0.4487	Adaptive LASSO	0.7389	0.5224	
0.8912	0.6500	Group SCAD	0.9100	0.6447	
0.9157	0.6968	Group LASSO	0.9077	0.6470	
Panel B: Twelve-month forecast			Panel B: Twelve-month forecast		
0.6199	0.3070	LASSO	-0.7727	-0.1072	
0.3084	-0.0927	Adaptive LASSO	-0.6818	-0.0952	
0.6417	0.3063	Group SCAD	0.4204	0.1852	
0.7227	0.3991	Group LASSO	0.4053	0.1905	

Notes: Table 3 reports the results of R^2_{MSPE} and R^2_{MAPE} in the six-month and twelve-month forecasts. Table 3(a) and 3(b) show the results when the benchmark models are the ARMAX (1,1,1) model and the GARCHX (1,1,1) model respectively. The larger positive value of R^2_{MSPE} and R^2_{MAPE} , the better prediction ability of the candidate forecast model compared with the corresponding benchmark model.

models with cluster structure have better predictive performance than other models.

Our findings have important implications for policy makers, investors, and other stakeholders, as they reveal the important influence of macroeconomic factors on carbon price dynamics. Policymakers can enhance the effec-

tiveness of intervention policies based on important predictors and expectations of the price of carbon. Effective forecasting of carbon prices also helps investors diversify their portfolios and provide reasonable risk diversification.



References

- Benz, E., & Truck, S. (2009). Modeling the price dynamics of CO₂ emission allowances. *Energy Economics*, 31(1), 4–15. <https://doi.org/10.1016/j.eneco.2008.07.003>
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.
- Byun, S. J., & Cho, H. (2013). Forecasting carbon futures volatility using GARCH models with energy volatilities. *Energy Economics*, 40, 207–221. <https://doi.org/10.1016/j.eneco.2013.06.017>
- Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4), 1509–1531. <https://doi.org/10.1093/rfs/hhm055>
- Chevallier, J. (2009). Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics*, 31(4), 614–625. <https://doi.org/10.1016/j.eneco.2009.02.008>
- Hong, K., Jung, H., & Park, M. (2017). Predicting European carbon emission price movements. *Carbon Management*, 8(1), 33–44. <https://doi.org/10.1080/17583004.2016.1275813>
- Ren, X., Duan, K., Tao, L., Shi, Y., & Yan, C. (2022). Carbon prices forecasting in quantiles. *Energy Economics*, 108, 105862. <https://doi.org/10.1016/j.eneco.2022.105862>
- Revesz, R., Greenstone, M., Hanemann, M., Livermore, M., Sterner, T., Grab, D., Howard, P., & Schwartz, J. (2017). Best cost estimate of greenhouse gases. *Science*, 357(6352), 655. <https://doi.org/10.1126/science.aao4322>
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Whittle, P. (1953). The analysis of multiple stationary time series. *Journal of the Royal Statistical Society: Series B*, 15(1), 125–139. https://doi.org/10.1007/978-1-4612-0667-5_7
- Xiong, S. H., Wang, C. F., Fang, Z. M., & Ma, D. (2019). Multi-Step-Ahead Carbon Price Forecasting Based on Variational Mode Decomposition and Fast Multi-Output Relevance Vector Regression Optimized by the Multi-Objective Whale Optimization Algorithm. *Energies*, 12(1). <https://doi.org/10.3390/en12010147>
- Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 68, 49–67. <https://doi.org/10.1111/j.1467-9868.2005.00532.x>
- Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476), 1418–1429. <https://doi.org/10.1198/016214506000000735>