IMPROVEMENT OF ESTIMATION METHOD FOR GRAMINEOUS CROP PRODUCTIVITY USING NORMALIZATION OF HYPERSPECTRAL DATA

Shinya Odagawa[1], Akihiro Nakazawa[2] and Tomomi Takeda[3]
[1]Asia Air Survey Co., Ltd., 1-2-2, Manpukuji, Asao-ku, Kawasaki, Kanagawa, 215-0004, Japan,
Email: sh.odagawa@ajiko.co.jp
[2]Asia Air Survey Co., Ltd., 1-8-30, Tenmanbashi, Kita-ku, Osaka, 530-6029, Japan,
Email: aki.nakazawa@ajiko.co.jp
[3]Japan Space Systems, 3-5-8, Shibakoen, Minato-ku, Tokyo, 105-0011, Japan,
Email: Takeda-Tomomi@jspacesystems.or.jp

KEY WORDS: Hyperpsectral data application, Lasso regression, Machine Learning, Rice quality estimation, Pasture yield estimation

ABSTRACT: This paper describes the normalization of hyperspectral data in order to construct a high accuracy estimation method for gramineous crop productivity. Generally, the estimation accuracy of crop is affected by the shadow caused by the topography and the object structure. Some vegetation indices of multispectral sensor are expected to reduce the influence of shadow, therefore previous studies used vegetation indices for estimation of crop productivity. Recently, needs for remote sensing trend with higher accuracy, so that hyperspectral data having high potential estimation capability have been employed in remote sensing applications. However, the technique of reducing shadow influence included in hyperspectral data is still not carefully reviewed. This study suggests the unit vectorized reflectance (UVR), which is one of the normalization for spectral data. The normalization is expected to reduce the shadow influence included in hyperspectral data. Our results show that UVR reduced the effect of shadow and improved the determination coefficient of dry matter pasture from 0.58 to 0.83, as well as the protein content rate of rice from 0.76 to 0.84. While, in comparison with different sun elevation, UVR improves of estimation accuracy under low sun elevation, which increases the influence of shadow. This study shows that the suggested method achieves to reduce shadow influence of hyperspectral data, as well as UVR is effective for improving estimation accuracy for gramineous crop productivity.

1. INTRODUCTION

Remote sensing can be utilized as an efficient monitoring system for food security and quality control of crop (Asaka and Shiga, 2003). However, enough precision is not provided in some cases (Akiyama et al., 2007). This is because observation range of multispectral sensor is not suited for information extraction of observed objects. Recently, hyperspectral sensor with the high wavelength resolution and the continuous spectral observation ability is being utilized for remote sensing application. The analysis method with generalization capability based on machine learning algorithm can produce high accuracy estimation model using hyperspectral data (Odagawa et al., 2012). However, high precision was not provided in some cases. This study suggests that influence of the shadow caused by the topography and the object structure is the main cause of low precision. In particular, gramineous crop is more susceptible to shadow due to the perpendicular structure of leaves. Previous studies attempt to estimate the quality of the object using vegetation index having the effect of shadow reduction (Asaka et al., 2006). To date, there are still few attempts to improve the estimation accuracy by reducing the shadow influence included in hyperspectral data. This study suggests the unit vectorized reflectance (UVR), which is one of the normalization. UVR is expected to reduce the shadow influence included in hyperspectral data. This study attempts to show the general versatility of the suggested method using two different gramineous crops, which are pasture and rice.

2. UNIT VECTORIZED REFLECTANCE

Unit vectorization was used to compare easily spectrums of mass spectrometer (Alfassi, 2004). In remote sensing application, UVR was used to reduce the shadow influence caused by topographic effect (Odagawa et al., 2012). UVR reduces volatility without relative spectral shape change because UVR holds a numerical change from zero to one. UVR can be expressed as equation (1).

$$Ruv_i = \frac{Ra_i}{\sqrt{\sum_{k=1}^{m} Ra_k^2}} \quad \dots \quad (1)$$

Ruv_i is UVR of band i, Ra_i and Ra_k are ground surface reflectance of band i and band k, m is the number of bands. Even if reflectance in shadow changes by topographic effect, UVR is able to restore the original spectrum. However, in the area under strong influence of scattering light with wavelength dependency, this method cannot restore the original spectrum. This study supposes that crop field on flat land is not subject to the influence of scattering light.

3. METHOD

3.1 Field Survey

Study area of pasture is Motonopporo farm owned by Rakuno Gakuen University in Ebetsu, Hokkaido, northern Japan (shown in figure 1). Timothy (TY), Orchardgrass (OG), Reed Canarygrass (RCG) and weed of oxeye were growing. Field survey conducted in the growing phase on 29 July to 1 August 2008. Forty-two sample points were set up. Dominant grass species were visually confirmed in each point. Above-ground part in fifty square centimeters was taken for laboratory measurement of the dry matter yield.

Study area of rice is three commercially available fields in Sakata, Yamagata prefecture, north-east Japan (shown in figure 1). An area of field, which is shown as a white rectangle in figure 1, is fifty meter in east-west, one hundred meter in north-south. Forty-eight sample points were set up in three fields. There are two rice cultivars, which are "Haenuki" and "Koshihikari". "Haenuki" is planted in AR-3 and AR-5, "Koshihikari" in AR-4. Rice samples were taken for laboratory measurement in late September 2008. Protein content rate of rice sample was measured by GS-1000J manufactured by Shizuoka Seiki Co., Ltd.

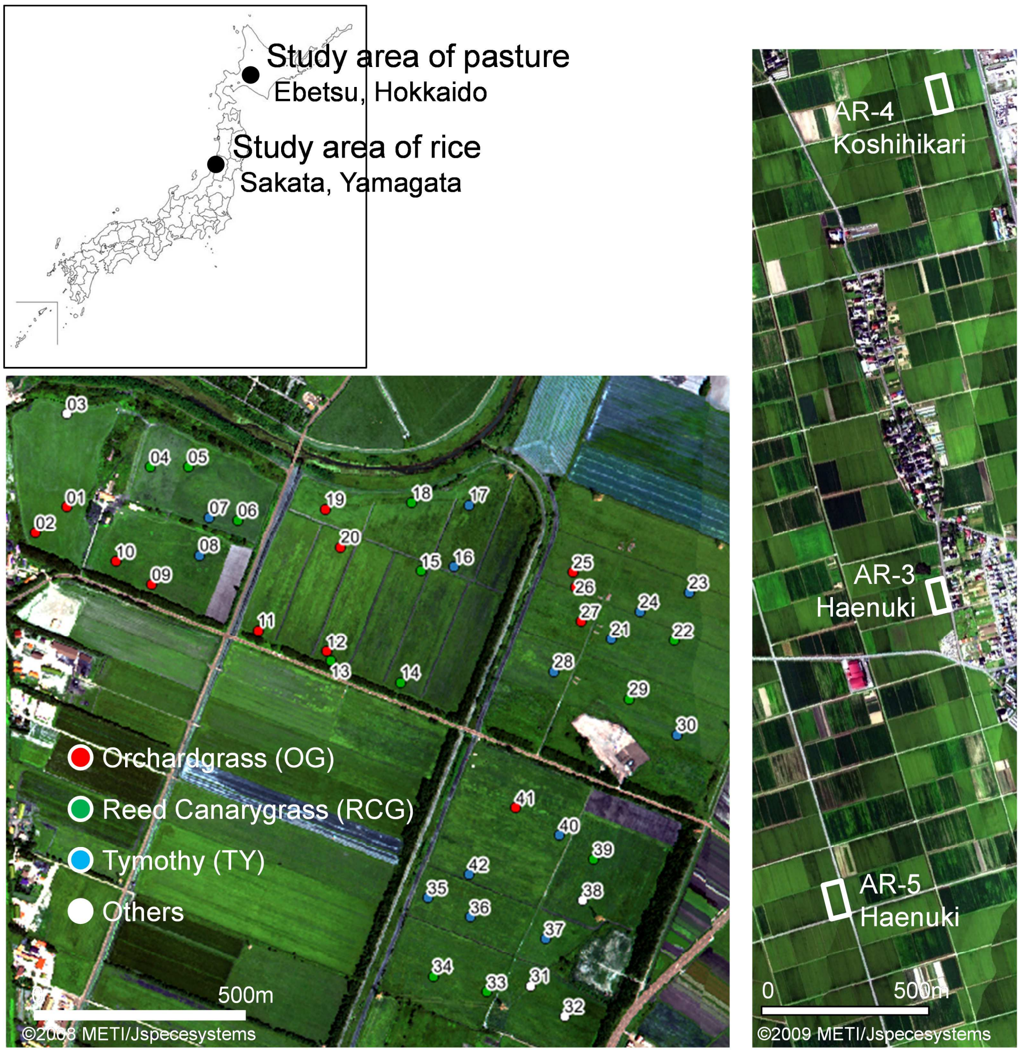

Figure 1. Study area of pasture (left) and rice (right).

3.2 Airborne Hyperspectral Sensor Observation

Airborne hyperspectral sensor for pasture is CASI-3, which is spectral range from 400 nm to 1060 nm with 10 nm spectral interval. The number of bands is 68. Spatial resolution was 1.5 m. The imagery was acquired at 10:40AM on 30 July 2008 (Japan Standard Time). The weather was clear and sunny.

Airborne hyperspectral sensor for rice is AISA, which is spectral range from 400 nm to 2450 nm with 9 nm spectral interval in visible-near infrared region and 11 nm spectral interval in short wavelength infrared region. The number of band in visible-near infrared region is 68, in short wavelength infrared region is 127. Spatial resolution was 1.5 m. The imagery was acquired at 14:45PM on 8 August 2008 (Japan Standard Time). The weather was clear and sunny.

After wavelength calibration correcting the shift of center wavelength (Odagawa et al., 2011), hyperspectral data was converted to ground surface reflectance using atmospheric correction function FLAASH in the image analysis software ENVI. The atmospheric model is Middle Latitude Summer, while the aerosol model is Rural.

3.3 Lasso Regression

This study used Lasso (Least Absolute Shrinkage and Selection Operator) regression for building the estimation model. Lasso regression is multiple regression analysis with regularization. Lasso selects the appropriate band for the estimation model (Tibshirani, 1996, Freedman et al., 2010). The decision criterion of Lasso can be expressed equation (2).

$$E = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - (\beta_0 + x_i^T\boldsymbol{\beta})\right)^2 + \lambda P(\boldsymbol{\beta}) \quad \ldots \quad (2)$$

N is the number of sample points, y is objective variable which means the result of field survey, x is explanatory variable, which is hyperspectral data, λ is regularization parameter, β is vector of regression coefficient for each band and $p(\beta)$ is sum of absolute value of regression coefficient. First term is mean square error (MSE) of regression, second term is regularization. If λ is very large, regression coefficients of all bands become zero. In this case, any regression model will not be constructed. $p(\beta)$ increases with a small number of λ, then the regression model is built. Because some regression coefficients are given a value, learning of regression coefficient in each regularization parameter achieves to adopt adequately bands and coefficients. Previous studies used MSE for discriminant, this study used the discriminant with Akaike Information Criteria (AIC). This reason is that AIC tends to build a simple model which uses a few bands (Odagawa et al., 2012). Simple model facilitates interpretation between object characteristics and hyperspectral data.

Reflectance of sample point of pasture is average of 5 by 5 pixel, rice is 3 by 3 pixel. UVR was calculated from this average reflectance. Validation method for pasture is three folds cross validation, rice is Leave-One-Out cross validation. At last, determination coefficient between predicted and measured value was acquired. This study used the package “glmnet” available in the statistical analysis software “R” for Lasso regression (Freedman et al., 2010).

4. RESULTS AND DISCUSSION

Table 1 shows field survey for pasture. The number of sample points for dominant species of TY, OG and RCG is 14, 12 and 12, respectively. The determination coefficients using reflectance and UVR are 0.32 and 0.47, respectively for all points. The estimation model was not built. However, only TY model was built in the analysis of each species. Figure 2 shows the predicted and measured value for TY. The determination coefficients using reflectance and UVR are 0.58 and 0.83, which shows that UVR improves estimation accuracy of pasture.

Table 2 shows field survey results for rice. The protein content rate for “Haenuki” ranges from 7.2 to 8.6 %, while for “Koshihikari” ranges from 6.7 to 7.6 %. Figure 3 shows the predicted and measured value of all points. The determination coefficients of reflectance and UVR are 0.76 and 0.84, which shows that UVR also improve estimation accuracy of rice. The result does not depend on the difference of rice cultivar. Taking into consideration, the results of the two different crops, our results show that UVR is effective for improving estimation accuracy.

Table 1. Dry matter of pasture

Loc. No.	TY	OC	RCG	Others	Total	Loc. No.	TY	OC	RCG	Others	Total	Loc. No.	TY	OC	RCG	Others	Total
1	0.00	38.13	0.00	8.00	46.13	15	0.00	101.87	0.00	0.00	101.87	29	0.00	0.00	60.93	0.00	60.93
2	0.00	22.41	0.00	1.00	23.41	16	89.39	0.00	0.00	0.00	89.39	30	19.00	0.00	0.00	3.00	22.00
3	0.00	0.00	0.00	33.00	33.00	17	83.87	0.00	0.00	4.00	87.87	31	0.00	0.00	45.39	1.00	46.39
4	0.00	2.00	25.00	1.00	28.00	18	0.00	0.00	88.71	1.00	89.71	32	27.02	0.00	0.00	19.00	46.02
5	0.00	0.00	65.34	1.00	66.34	19	0.00	47.36	0.00	4.00	51.36	33	0.00	0.00	96.72	2.00	98.72
6	0.00	0.00	30.69	0.00	30.69	20	2.00	63.20	0.00	0.00	65.20	34	0.00	0.00	19.00	8.00	27.00
7	19.00	0.00	0.00	0.00	19.00	21	18.00	0.00	0.00	4.00	22.00	35	56.71	0.00	2.00	1.00	59.71
8	27.95	0.00	0.00	3.00	30.95	22	0.00	0.00	56.52	0.00	56.52	36	54.58	0.00	0.00	0.00	54.58
9	0.00	38.75	0.00	1.00	39.75	23	43.19	3.00	0.00	0.00	46.19	37	29.68	1.00	0.00	1.00	31.68
10	0.00	32.98	0.00	1.00	33.98	24	41.80	0.00	0.00	2.00	43.80	38	23.00	0.00	0.00	43.09	66.09
11	2.00	29.94	0.00	1.00	32.94	25	1.00	73.04	0.00	2.00	76.04	39	1.00	0.00	48.35	8.00	57.35
12	0.00	39.96	0.00	0.00	39.96	26	0.00	32.00	0.00	0.00	32.00	40	76.46	0.00	0.00	4.00	80.46
13	0.00	0.00	43.82	2.00	45.82	27	0.00	0.00	104.97	0.00	104.97	41	5.00	64.94	0.00	1.00	70.94
14	0.00	0.00	48.65	2.00	50.65	28	27.22	0.00	0.00	0.00	27.22	42	41.20	0.00	0.00	0.00	41.20

Unit: $g/0.25m^2$

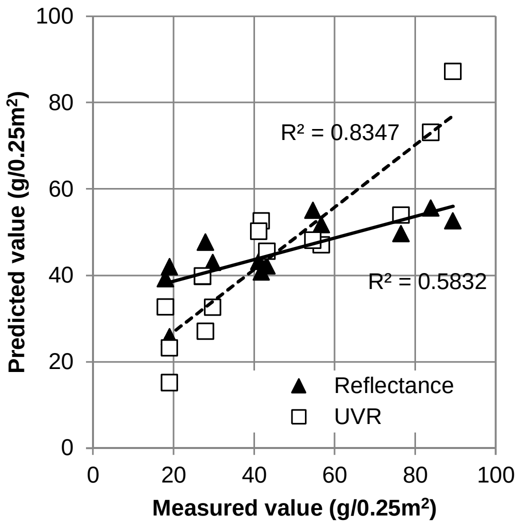

Figure 2. Correlation between predicted and measured value of dry matter of TY.

Table 2. Protein content rate of rice

Loc. No.	Protein (%) "Haenuki"	Loc. No.	Protein (%) "Koshihikari"	Loc. No.	Protein (%) "Haenuki"
AR-3-01	7.9	AR-4-01	7.1	AR-5-01	7.8
AR-3-02	7.6	AR-4-02	7.2	AR-5-02	7.9
AR-3-03	7.5	AR-4-03	7.4	AR-5-03	8.1
AR-3-04	7.7	AR-4-04	7.2	AR-5-04	8.2
AR-3-05	7.9	AR-4-05	7.4	AR-5-05	8.1
AR-3-06	8	AR-4-06	7.6	AR-5-06	8.1
AR-3-07	7.9	AR-4-07	7.1	AR-5-07	7.7
AR-3-08	8.2	AR-4-08	7.2	AR-5-08	7.4
AR-3-09	7.9	AR-4-09	6.9	AR-5-09	7.7
AR-3-10	7.6	AR-4-10	7.2	AR-5-10	7.9
AR-3-11	7.5	AR-4-11	-	AR-5-11	8
AR-3-12	7.8	AR-4-12	7.2	AR-5-12	8.6
AR-3-13	7.4	AR-4-13	6.7	AR-5-13	8.3
AR-3-14	7.5	AR-4-14	6.9	AR-5-14	7.7
AR-3-15	7.4	AR-4-15	6.8	AR-5-15	7.9
AR-3-16	7.2	AR-4-16	6.9	AR-5-16	7.3

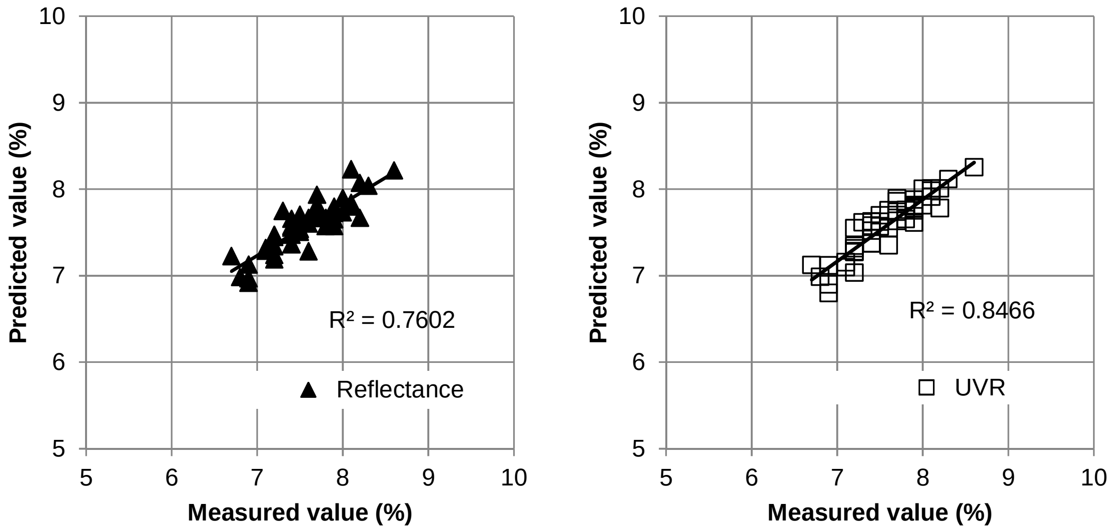

Figure 3. Correlation between predicted and measured value of protein content of rice observed under low sun elevation (45 degree). The left is result of reflectance, the right is UVR.

Another hyperspectral data for rice was measured under high sun elevation (60 degree). The imagery was acquired at 13:15PM on 8 August 2008. The influence of shadow was reduced under high sun elevation. Figure 4 shows the predicted and measured value for rice protein content under high sun elevation. The determination coefficient of reflectance and UVR are 0.76 and 0.72, which indicates that UVR cannot improve the estimation accuracy under high sun elevation. However, results under lower sun elevation (45 degree) shows that UVR is more effective in improving estimation accuracy using hyperspectral data including the influence of shadow.

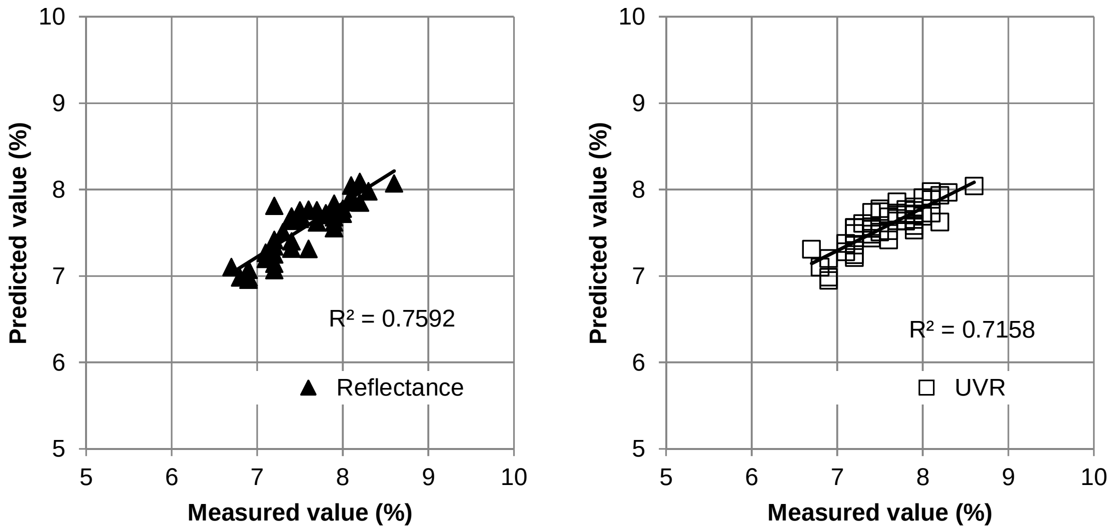

Figure 4. Correlation between predicted and measured value of protein content of rice observed under high sun elevation (60 degree). The left is result of reflectance, the right is UVR.

6. CONCLUSION

Unit vectorization achieved the improvement of estimation accuracy for different crop species. This result shows that UVR reduces the influence of shadow and is a versatile method for improving better estimation accuracy. In comparison with different sun elevation, UVR improves of estimation accuracy under low sun elevation, which includes more influence of shadow. This result shows that improvement by UVR increases under the influence of shadow. This result also indicates that UVR is superior for satellite sensor observation with more constraint condition. Therefore, UVR is an effective and versatile method for improving estimation accuracy.

This study is a part of "Research and Development Project for the Next-generation Earth Observation Satellite Utilization Technology" conducted by Japan Space Systems entrusted by Ministry of Economy, Trade and Industry.

Acknowledgment

The authors would like to thank Prof. M. Kaneko and Prof. B. Hoshino (Rakuno Gakuen University, Japan) for accommodation for field survey of pasture. Also, the authors would like to thank K. Oda (Yamagata prefecture, Japan) and assistant Prof. Y. Sasaki (Yamagata University, Japan) for accommodation for field survey of rice. Especially, the authors would like to thank Prof. Y. Kosugi, extraordinary Prof. G. Saito and assistant Prof. K. Uto (Tokyo Institute of Technology, Japan) for mentorship of this study.

References

Akiyama, T., Ishizuka, N., Ogawa, S., Okamoto, K., Saito, G. and Uchida, S., 2007. Agriculture Remote Sensing Handbook, The Japan Agricultural System Society, Tokyo, pp. II30-II32.

Alfassi, Z. B., 2004. On the Normalization of a Mass Spectrum for Comparison of Two Spectra, American Society for Mass Spectrometry, 15, pp. 385-387.

Asaka, D. and Shiga H., 2003. Estimating Rice Grain Protein Contents with SPOT/HRV Data Acquired at Maturing Stage, Journal of The Remote Sensing Society of Japan, 23(5), pp. 451-457.

Asaka, D., Shiga, H. and Hayashi, T., 2006. The map of grain protein content of winter wheat using satellite observed NDVI at maturing stage, The Japan Agricultural System Society, 22(2), pp. 89-98.

Odagwa, S., Kosugi, U., Uto, K., Saito, G., Sasaki, Y., Oda, K. and Kato, M., 2012. Development of an estimation method for the rice growth status, yield and protein content using airborne hyperspectral data, Journal of the Japan Society of Photogrammetry and Remote Sensing, 51(5), pp. 270-284.

Odagawa, S., Sanga, T. and Kato, M., 2012. A Proposal for the Reduction Method of Vegetation Influence from Mixed Spectrum Using Unit Vectorized Reflectance and NDVI Function in Vegetated Area, Journal of The Remote Sensing Society of Japan, 32(5), pp. 320-329.

Odagawa, S., Suhama, T., Kato, M., Saito, G. and Kosugi, Y., 2011. Artifact reduction by wavelength calibration for hyperspectral imagers, Journal of the Japan Society of Photogrammetry and Remote Sensing, 50(6), pp. 347-360.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso, J. Roy. Statist. Soc. Ser. B, 58, pp. 267-288.

Freedman, J., Hastie, T. and Tibshirani, R., 2010. Regularization Paths for Generalized Linear Model via Coordinate Descent, Journal of Statistical Software, 33(1), pp. 1-22.