

Full Length Research Paper

Establishment of rice yield prediction model using canopy reflectance

K. W. Chang

Department of Leisure and Recreation Studies, Aletheia University, Tainan, 721, Taiwan, Republic of China.
E-mail: ckw550320@gmail.com. Tel: 886-6-5703100 ext. 7420. Fax: 886-6-5704222.

Accepted 4 September, 2012

The major objectives of this study were to identify spectral characteristics associated with rice yield and to establish their quantitative relationships. Field experiments were conducted at Shi-Ko experimental farm of TARI's Chiayi Station, during 2001 to 2005. Rice cultivar Tainung 67 (*Oryza sativa* L.), the major cultivar grown in Taiwan, was used in the study. Various levels of rice yield were obtained via nitrogen application treatments. Canopy reflectance spectra were measured during entire growth period and dynamic changes of characteristic spectrum were analyzed. Relationships among rice yields and characteristic spectrum were studied to establish yield estimation models suitable for remote sensing purposes. Spectrum analysis indicated that the changes of canopy reflectance spectrum were least during booting stages. Therefore, the canopy reflectance spectra during this period were selected for model development. Two multiple regression models, constituting of band ratios (NIR/RED and NIR/GRN) were then constructed to estimate rice yields for first and second crops separately. Results of the validation experiments indicated that the derived regression equations successfully predicted rice yield using canopy reflectance measured at booting stage unless other severe stresses occurred afterward.

Key words: Rice yield, canopy reflectance, prediction model.

INTRODUCTION

Yield maps are the basis of making precision management decisions. Through accumulated yield maps during past seasons, maps for field management can be produced. Regions always having higher or lower yields can be easily delineated, which can be very useful for diagnosing the causes responsible for low yield. Proper management strategies can then be applied. Where available, remotely sensed images showing spatial and spectral variations resulting from soil and crop characteristics are important source of data for making yield maps (NRC, 1997). A remote sensed yield map would not be affected by the inaccuracies (problems connected with grain flow dynamics and accurate logging of geographic position) associated with combine yield monitors, as suggested by Lark et al. (1997) and Arslan and Colvin (1999). However, difficulty results from a lack of valid regression models to convert imagery spectral information to a yield map.

Rice is one of the world's major staple foods, and paddy rice fields account for approximately 15% of the world's

arable land (IRRI, 1993). Remote sensing data can acquire temporal, large spatial and vast spectrum data and also track the past data. Remote sensing, which quantitatively measures the light reflected from the surface of the earth, is a powerful tool for studying regional-scale ecosystem dynamics. So, remote sensing techniques have the potential to provide information on agricultural crops quantitatively, instantaneously and non-destructively over large areas. Ability to estimate rice yields within fields from remote sensing images is not only fundamental to applications of precision agriculture, but can also be very useful to governmental administrators for food provisions management. Though many researches have been devoted to rice planting acreage estimation (Leblon et al., 1991; Prince, 1991; Bouman, 1992; Wiegand et al., 1992; Field et al., 1995; Clevers and Leeuwen, 1996; Bach, 1998; Moulin et al., 1998; Reynolds et al., 2000; Serrano et al., 2000), few studies have been conducted attempting to relate canopy reflectance spectra measurements to grain yields. Basic

Table 1. Analysis of variance for rice yield at four nitrogen application rates (0, 45, 90, 180 kg N ha⁻¹) and five years (2001 to 2005).

Sources of variation	df	Mean squares	Pr
Year (Y)	4	4.19026276	<0.0001
Nitrogen rate (N)	3	7.00343400	<0.0001
Y × N	12	0.58496987	0.2310

studies regarding the timing of reflectance measurements and regression models for rice yield prediction/estimation are still very lacking. Our objectives were thus to collect the reflectance spectrum of rice canopies, to identify spectral characteristics associated with rice yield, and to establish their quantitative relationships.

MATERIALS AND METHODS

Study site and experimental treatments

The study was conducted at Shi-Ko experimental farm of TARI's Chiayi Station (23°35'4" N, 120°24' E) during 2001 to 2005 growing seasons. The soil at this site is a silt clay loam (mixed hyperthermic Haplaquepts). Rice cultivar Tainung 67 (*O. sativa* L.), the major cultivar grown in Taiwan, was used in the study. Different nitrogen levels, 0, 45, 90, 180 kg N ha⁻¹, were used in each year to affect rice yield. Field plots were shifted every year to other well-fertilized production fields to avoid any residue fertilizer effect from the previous year. The plots in each year were all arranged in randomized complete block design with three replications. Individual plot dimensions were at least 10 m × 10 m.

The rice was grown under conventional two-season cropping system. Three-leaf old rice seedlings were transplanted in early February for first season crop, and in early July for second season crop. Transplanting density was 0.15 m by 0.25 m with three plants per hill. Other than nitrogen, all plots were fertilized with 150 kg P₂O₅ ha⁻¹ and 150 kg K₂O ha⁻¹ at the time of transplanting as the basal dose. The nitrogen fertilizer, fractionated as three quotas, were applied as basal at transplanting, and top-dressed at active tillering and panicle initiation stages, respectively. In-season weed and pest controls were practiced according to regional recommendations.

Canopy reflectance measurement

Canopy reflectance spectra were measured during entire growth period of each cropping season using a portable spectroradiometer (LI-1800, LICOR) with remote cosine receptor (LI-1800-02, LICOR) attached to a 1.5 m extension arm. The arm was held 1 m above the canopy; at this height, a target area of 1 m-radius may have occupied 80% of the view. The man holding the extension arm always wore dark clothes and stands sideways to reduce measurement error. All the measurements were made near midday, within 2 h ± solar noon. Incident and reflected solar radiations were measured by facing the remote cosine receptor upward and downward, respectively. The measurements were taken over the wavelength range from 400 to 1100 nm at a scanning interval of 10 nm and executed consecutively three times per subplot to reduce the possible effect of changing sky conditions. The reflectance of canopies was then calculated from the mean of three repetitions.

Incremental values of spectral reflectance were averaged within

0.45 to 0.52 μm, 0.52 to 0.60 μm, 0.63 to 0.69 μm, and 0.76 to 0.90 μm to give, respectively, values of blue (BLU), green (GRN), red (RED), and near-infrared (NIR) bands of reflectance. Two normalized difference vegetation indices (NDVI and GNDVI) and two ratio vegetation indices (NIR/GRN and NIR/RED) were then calculated. The dynamic changes of these four vegetation indices during growing season were analyzed for each treatment at different year and crop season. The indices and timing best for yield prediction model development were then selected accordingly.

Grain yield estimation

At maturity, grain yield was estimated by hand harvesting three 1 m × 1 m area of each subplot. After threshing, fully filled grains sieved by wind selection were sun dried. After drying, grain yield per subplot was weighted and adjusted to a constant moisture basis of 13.5% water. Grain yield, vegetation index, year, and season were analyzed via ANOVA with a mixed model (Statistica Ver. 6, StatSoft, 2001).

Establishment and validation of rice yield prediction models

Rice yield prediction models were established by regression analysis of relationships between grain yields, spectral reflectance and vegetation indices using the General Linear Models procedure. Independent data of canopy reflectance at booting stage and grain yield from different nitrogen rate experiments conducted also at Shi-Ko experimental farm on the two crop growing seasons during 2002 to 2005. The experimental design of the nitrogen experiments, with levels of 0, 45, 90, 180 kg N ha⁻¹, was similar to those described earlier.

In this study, validation of rice yields prediction models was carried out using cross validation method. We judged the model fitting abilities with coefficients of determination (R²) and root mean squared error of cross validation (RMSECV). Finally, we will conduct the model diagnostics with scatter plots and residual plots (Neter et al., 1999).

RESULTS AND DISCUSSION

The ground truth data of rice yield

The main effects of N rate and the interaction between N rate and year were statistically significant for yield (Table 1). Yields were also significantly different among the years and crop seasons. These results indicated that rice yield variability was not only affected by the amount of applied N fertilizers but also by the differences in climatic conditions between the years and crop seasons.

The features of rice canopy reflectance

Typical temporal changes of rice canopy reflectance spectra during first crop season were shown in Figure 1. Three weeks after transplanting, rice seedlings were just recovering from damages induced by transplanting. At this time, percent ground cover was less than 15%. Values of reflectance in the visible region were only slightly lower than those in the near-infrared region

Table 2. The results of RMSECV and R^2 for rice yield prediction models using different band of spectral reflectance.

First order linear models	RMSECV	R^2
BLUE	1.073	0.059
GRN	1.103	0.002
RED	1.023	0.130
NIR	0.886	0.336
BLUE GRN	1.077	0.062
BLUE RED	0.982	0.199
BLUE NIR	0.791	0.474
GRN RED	0.974	0.208
GRN NIR	0.684	0.605
RED NIR	0.742	0.535
BLUE GRN RED	0.945	0.261
BLUE GRN NIR	0.698	0.591
BLUE RED NIR	0.742	0.538
GRN RED NIR	0.695	0.592
BLUE GRN RED NIR	0.702	0.586

Table 3. The results of RMSECV and R^2 for rice yield prediction models by vegetation index.

First order linear models	RMSECV	R^2
NDVI	0.702	0.582
GNDVI	0.679	0.610
SR	0.710	0.573
NIR/GRN	0.705	0.580
NDVI GNDVI	0.688	0.601
NDVI SR	0.709	0.575
NDVI NIR/GRN	0.702	0.586
GNDVI SR	0.687	0.602
GNDVI NIR/GRN	0.718	0.570
SR NIR/GRN	0.712	0.573
NDVI GNDVI SR	0.703	0.585
NDVI GNDVI NIR/GRN	0.725	0.564
NDVI SR NIR/GRN	0.708	0.580
GNDVI SR NIR/GRN	0.727	0.562
NDVI GNDVI SR NIR/GRN	0.738	0.551

because underlying water and soil contributed most to the measured canopy reflectance spectrum. As rice plants grew, contribution from the plants gradually increased. At active tillering stage, six weeks and nine weeks after transplanting, tiller number and leaf area increased much more rapidly at rate about 1.3 tillers per day. Accordingly, reflectance in near-infrared region increased rapidly as a result of increased light scattering by leaves and stems. However, the reflectance in the visible region decreased due to absorption by pigments, chlorophyll in particular. About 12 weeks after transplanting, reflectance in the near-infrared region reached the highest value of the

season while reflectance in visible portion reached lowest value. At later stages, yellowing and wilting of rice plants gradually appeared. Therefore, reflectance in visible region increased as a result of decreasing chlorophyll concentration but reflectance in near infrared region decreased due to wilting, the exposing of soil background.

The rice yield prediction models

In this study, the rice yield prediction models were made using rice yields as dependent variable and different bands of spectral reflectance, band ratios and vegetation indexes as independent variables. After establishment of rice yield prediction models, we used RMSECV and R^2 values as criticism to judge the fitting abilities of models through cross validation.

Table 2 is the results of RMSECV and R^2 for rice yield prediction models using different band of spectral reflectance. From Table 2, we find that the simple regression models with BLUE, GRN, RED and NIR as independent variable have higher RMSECV values and lower R^2 values. So, the fitting abilities of these models are poor and not suitable for predicting rice yields. But the multiple regression models with GRN and NIR as independent variable have lower RMSECV values and higher R^2 values which fitting ability are well and suitable for predicting rice yields. Other multiple regression models with three or four independent variables have lower RMSECV values and higher R^2 values, their fitting abilities are most well and suitable for predicting rice yields.

Table 3 is the results of RMSECV and R^2 for rice yield prediction models by vegetation index (NDVI, GNDVI, SRVI and GRVI). From Table 3, we can clearly find that the simple regression models with vegetation indexes as independent variable have lower RMSECV and higher R^2 values than those of the simple regression models with different bands of spectral reflectance as independent variable. So, the fitting abilities of these models with vegetation indexes as independent variable are better and suitable for predicting rice yields than those models with different bands of spectral reflectance as independent variable. In all first order linear models, the model with GNDVI as independent variable has the lowest RMSECV and the highest R^2 values than others which is the best rice yield prediction model.

Diagnostics of rice yield prediction models

When a regression model is considered for an application, we can usually not be certain in advance that the model is appropriate for that application. Any one, or several, of the features of the model, such as linearity of the regression function or normality of the error terms, may not be appropriate for the particular data at hand. Hence, it is important to examine the aptness of the model for the data before inferences based on that model are

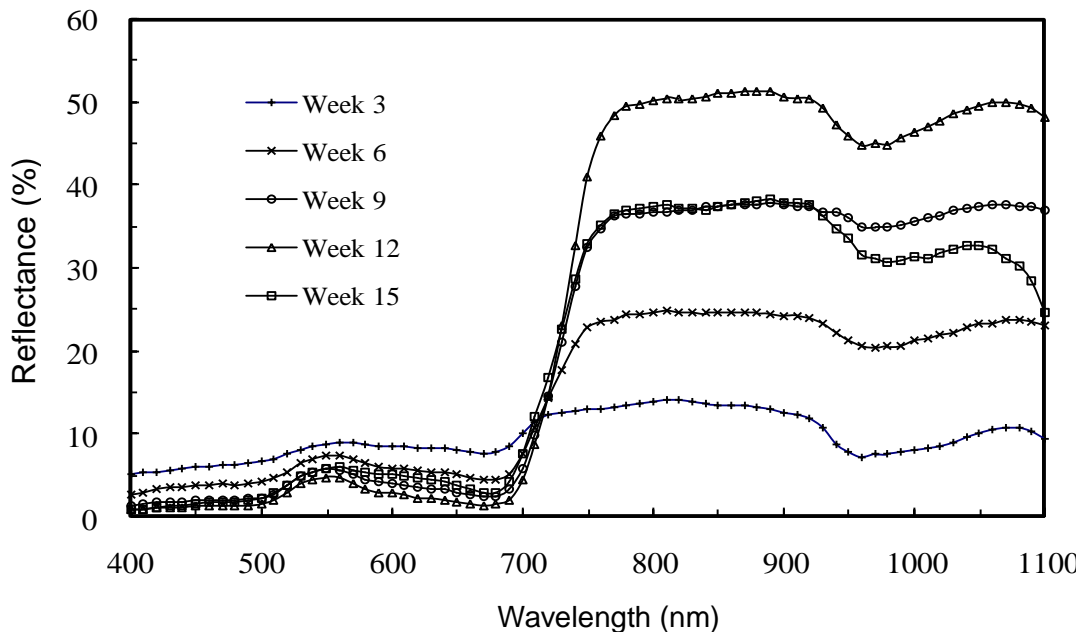


Figure 1. Typical temporal changes of rice canopy reflectance spectra.

undertaken.

Figure 2 is the plots of actual and predicted values of yield in first order linear models using BLUE, GRN, RED, and NIR as independent variable. From Figure 2, it is shown that the simple regression models with BLUE, GRN, and RED as independent variable appeared over or under estimation, but the model with NIR as independent variable has better performance. For multiple regression models, those performances are better than simple regression models. In general, multiple regression models with different bands of spectral reflectance as independent variables do not appear over or under estimation situation except those with BLUE and GRN, BLUE and RED, GRN and RED as independent variables.

Figure 3 is the plots of actual and predicted values of yield in linear models using vegetation indexes (NDVI, GNDVI, SRVI, GRVI, NDVI and GNDVI, NDVI and SRVI) as independent variables. From Figure 3, it is shown that the rice yield prediction models with vegetation indexes as independent variables do not appear serious over or under estimation situation.

Direct diagnostic plots for the response variable Y are ordinarily not too useful in regression analysis because the values of the observations on the response variable are a function of the level of the predictor variable. Instead, diagnostics for the response variable are usually carried out indirectly through an examination of the residuals. When a linear regression model is appropriate, the residuals fall within a horizontal band centered on 0, displaying no systematic tendencies to be positive and negative.

Figure 4 is the scatter plots of residual and actual value of yield in first order linear models using BLUE, GRN,

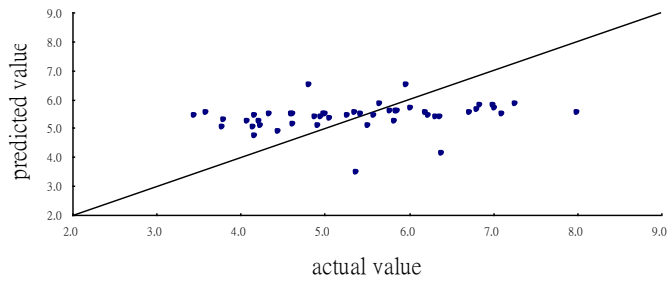
RED, and NIR band values as independent variables. The simple regression models with BLUE, GRN, and RED as independent variable show declination situation of different levels between residual and actual value of yield from Figure 4, but the model with NIR as independent variable has better performance. For multiple regression models, those performances are better than simple regression models. In general, multiple regression models with three band values as independent variable do not show serious declination situations between residual and actual value of yield except that with BLUE, GRN and RED as independent variables.

Figure 5 is the scatter plots of residual and actual value of yield in linear models with vegetation indexes as independent variables. The simple regression models with NDVI, SRVI and NIR/GRN (GRVI) as independent variable showed declination situation of different levels between residual and actual value of yield from Figure 5, but the model with GNDVI as independent variable has better performance. The multiple regression models with two vegetation indexes as independent variables also showed declination situation of different levels between residual and actual value of yield, which means the fitting ability of these models are not so good. But the multiple regression models with three vegetation indexes as independent variables do not show serious declination situations between residual and actual value of yield; those models possess better fitting abilities.

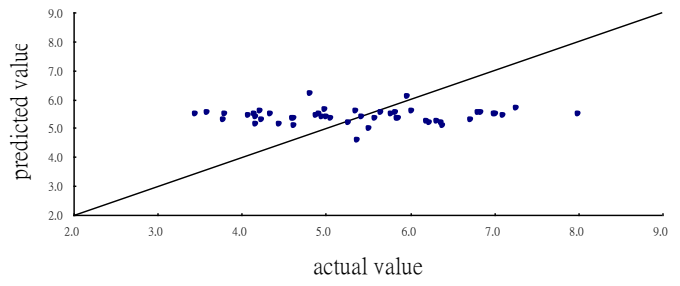
Conclusion

Ability to estimate rice yields within fields from remote

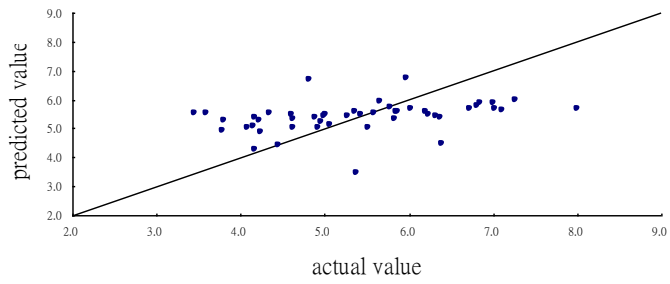
(a) BLUE



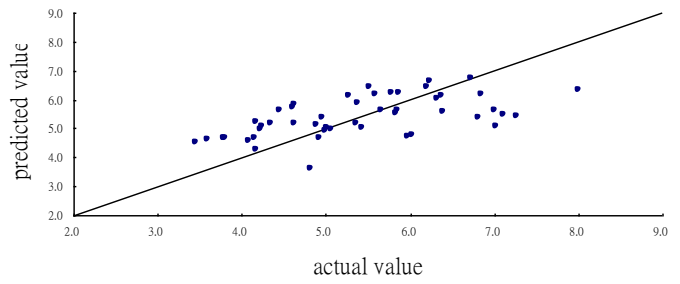
(b) GRN



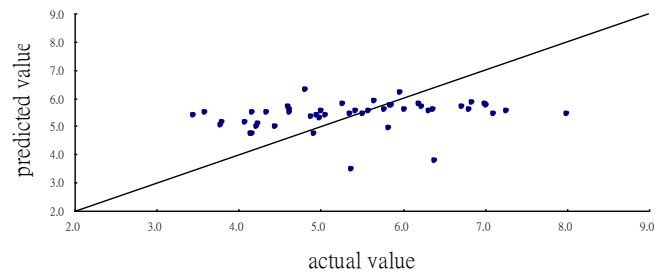
(c) RED



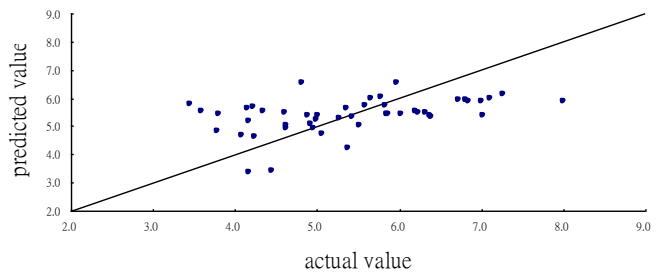
(d) NIR



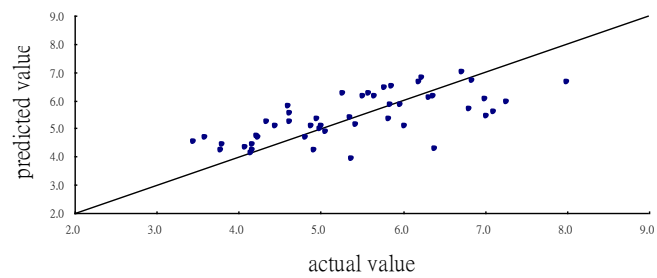
(e) BLUE GRN



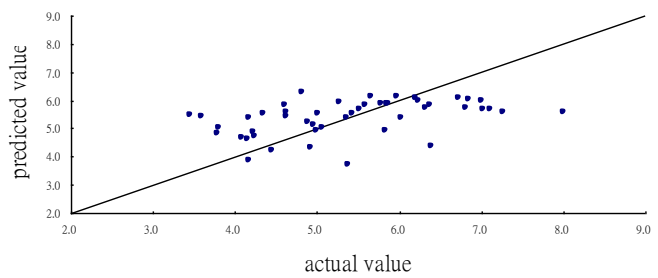
(f) BLUE RED



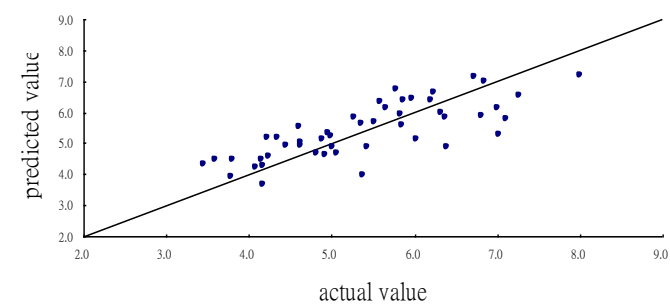
(g) BLUE NIR



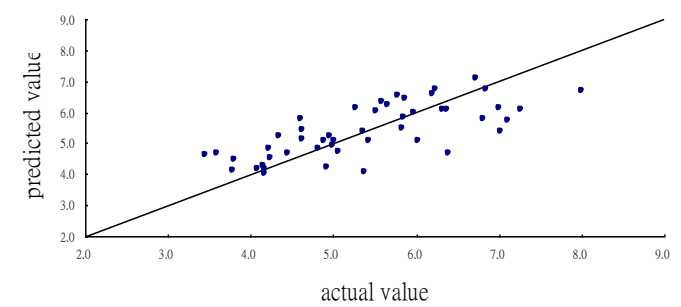
(h) GRN RED



(i) GRN NIR



(j) RED NIR



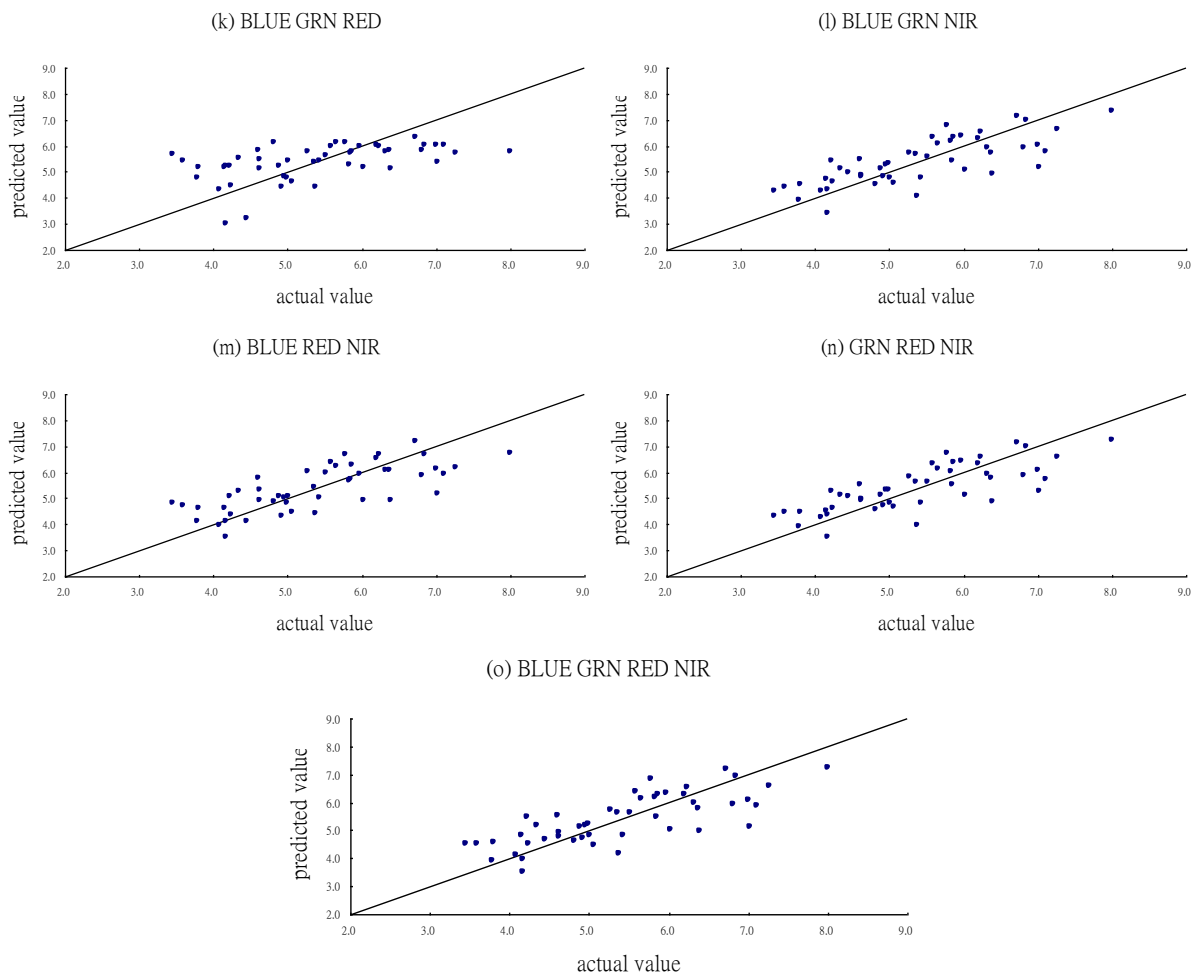
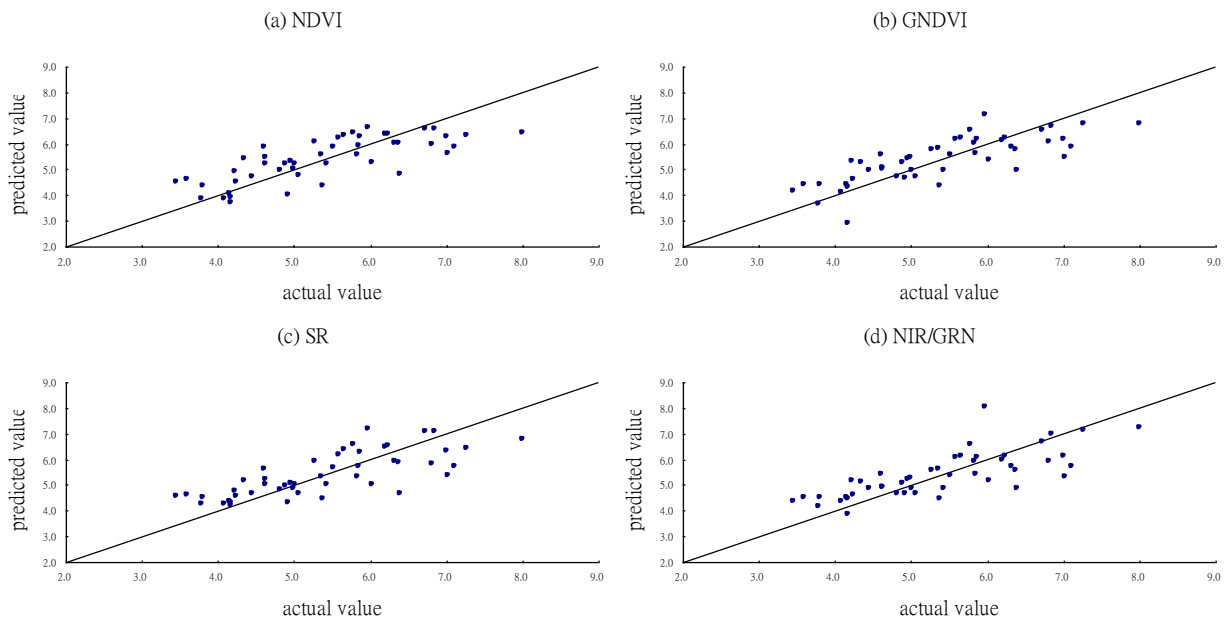
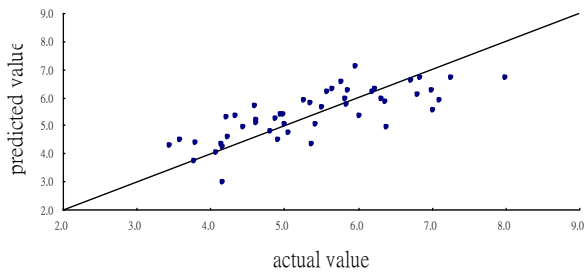


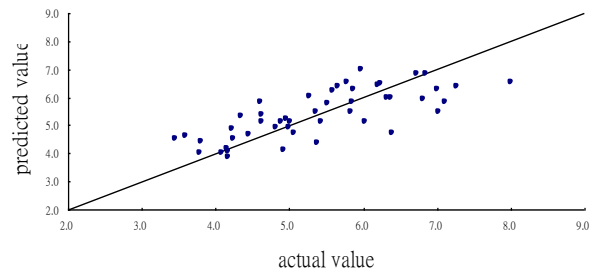
Figure 2. The plots of actual and predicted value of yield in first order linear models using (a) BLUE, (b) GRN, (c) RED, (d) NIR, (e) BLUE and GRN, (f) BLUE and RED, (g) BLUE and NIR, (h) GRM and RED, (i) GRN and NIR, (j) RED and NIR, (k) BLUE, GRN and RED, (l) BLUE, GRN and NIR, (m) BLUE, RED and NIR, (n) GRN, RED and NIR, (o) BLUE, GRN, RED and NIR.



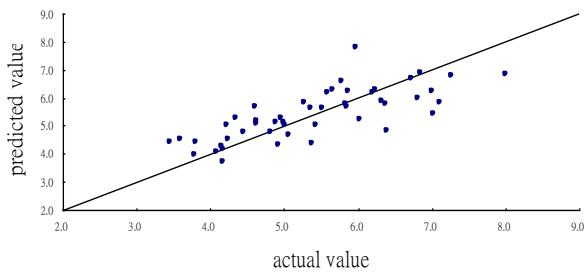
(e) NDVI GNDVI



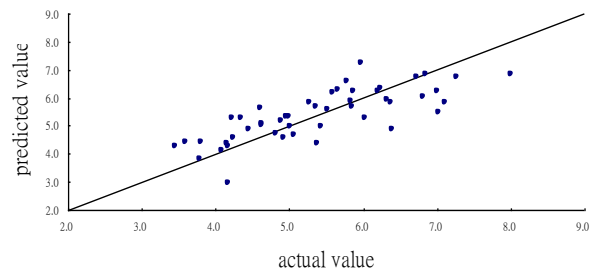
(f) NDVI SR



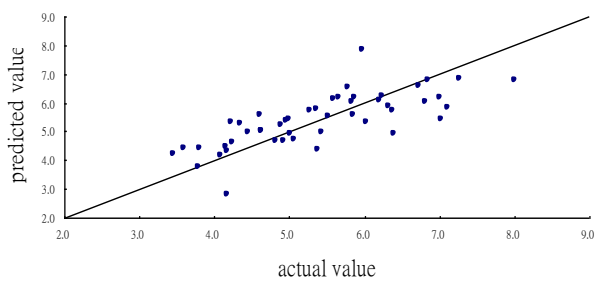
(g) NDVI NIR/GRN



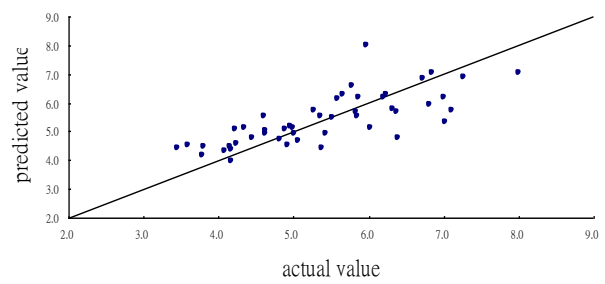
(h) GNDVI SR



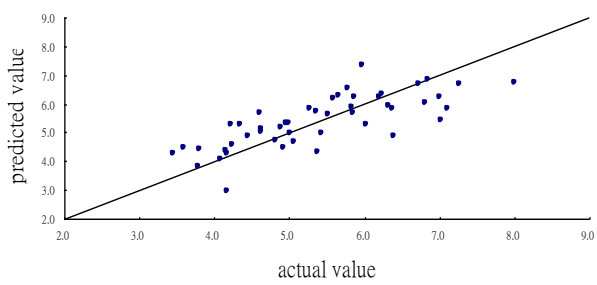
(i) GNDVI NIR/GRN



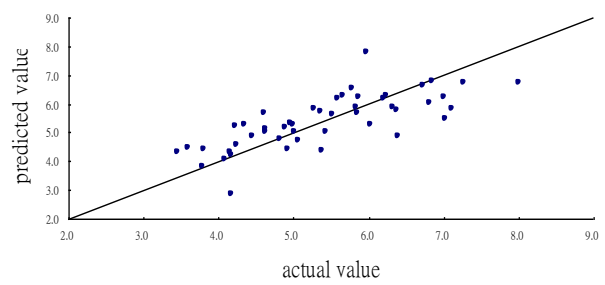
(j) SR NIR/GRN



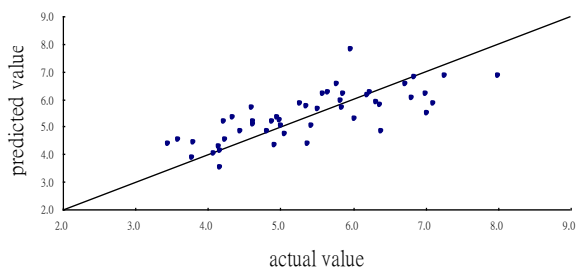
(k) NDVI GNDVI SR



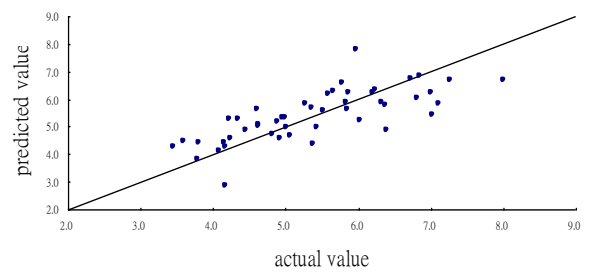
(l) NDVI GNDVI NIR/GRN



(m) NDVI SR NIR/GRN



(n) GNDVI SR NIR/GRN



(o) NDVI GNDVI SR NIR/GRN

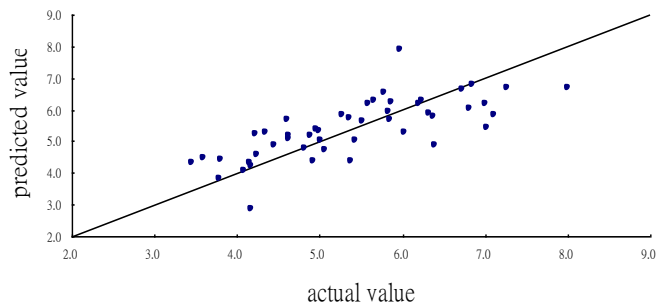
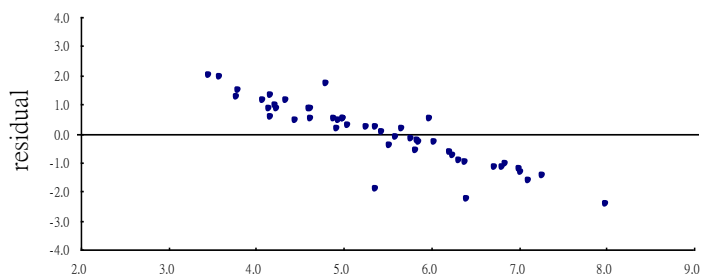
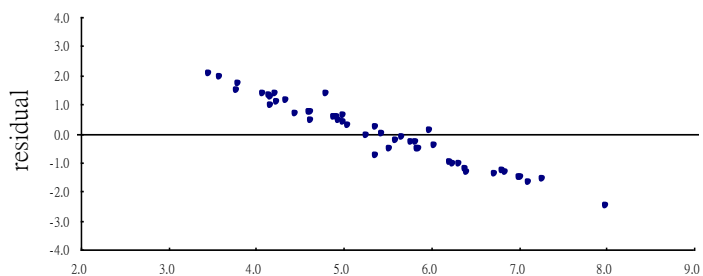


Figure 3. The plots of actual and predicted value of yield in linear models using (a) NDVI, (b) GNDVI, (c) SR, (d) NIR/GRN, (e) NDVI and GNDVI, (f) NDVI and SR, (g) NDVI and NIR/GRN, (h) GNDVI and SR, (i) GNDVI and NIR/GRN, (j) SR and NIR/GRN, (k) NDVI, GNDVI and SR, (l) NDVI, GNDVI and NIR/GRN, (m) NDVI, SR and NIR/GRN, (n) GNDVI, SR and NIR/GRN, and (o) NDVI, GNDVI, SR and NIR/GRN.

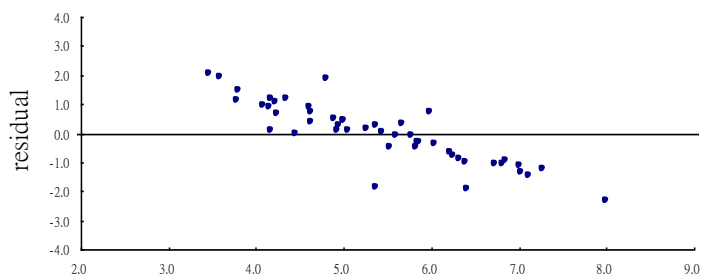
(a) BLUE



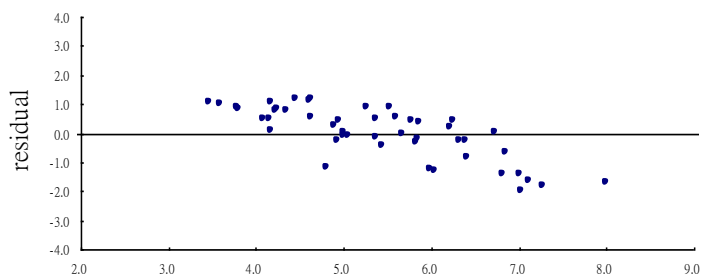
(b) GRN



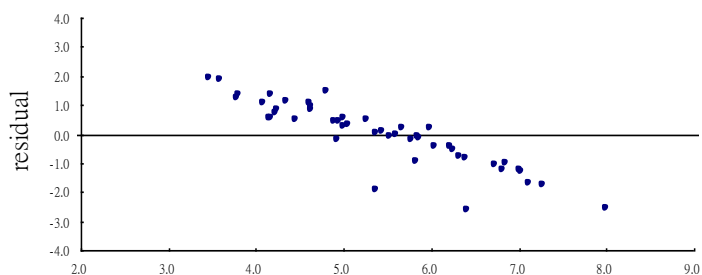
(c) RED



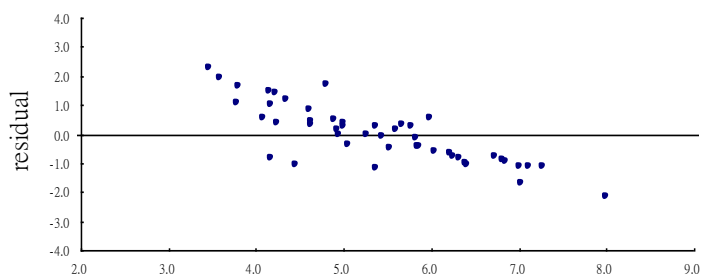
(d) NIR



(e) BLUE GRN



(f) BLUE RED



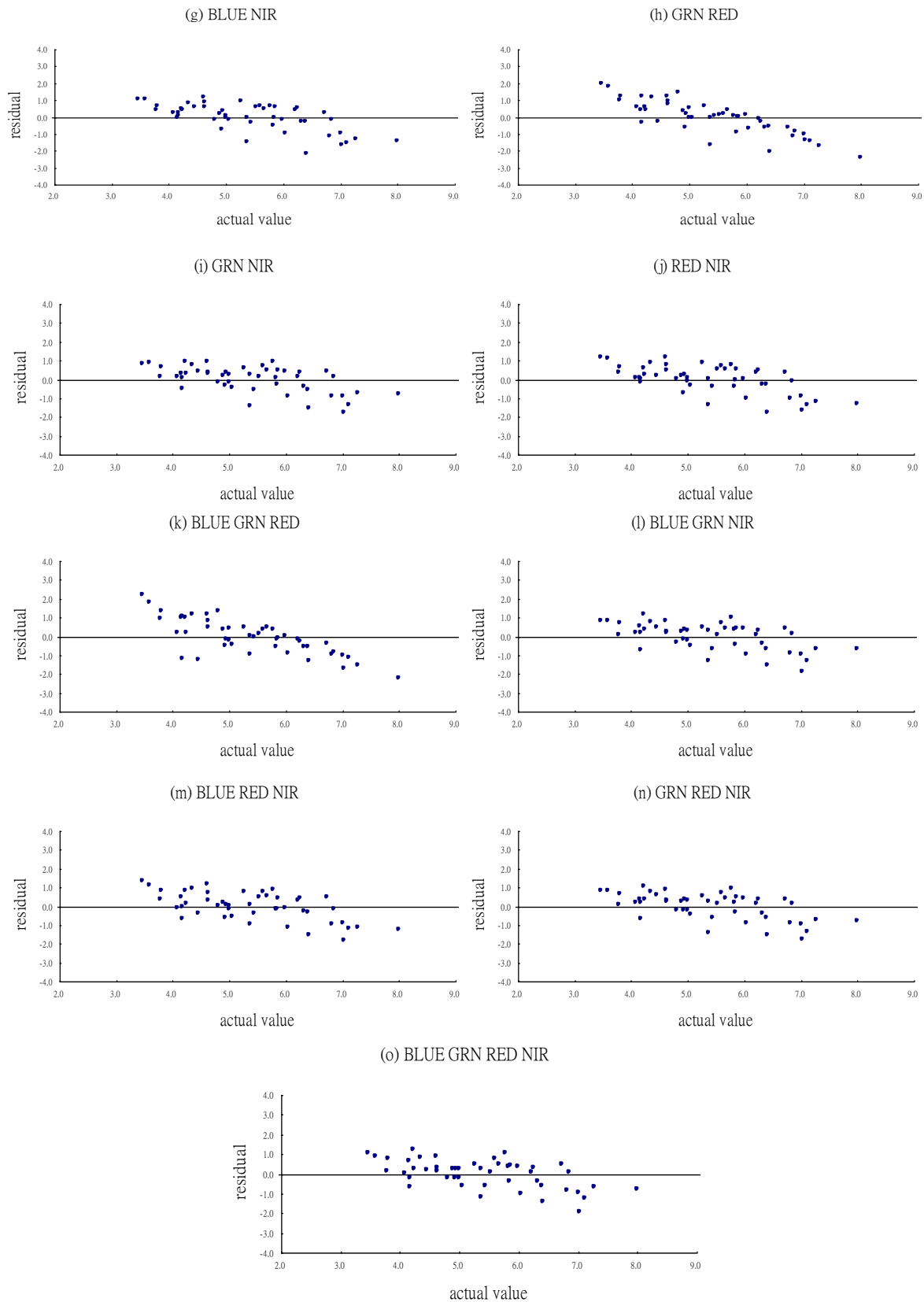
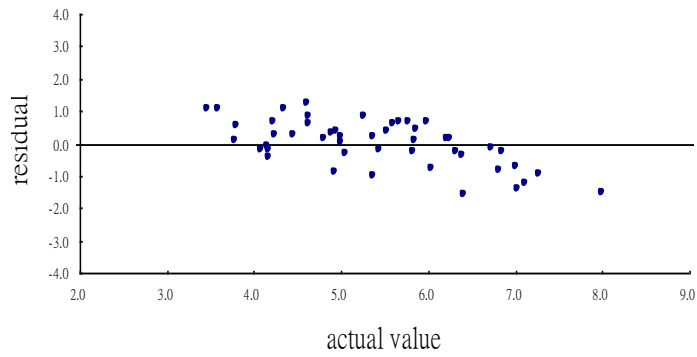
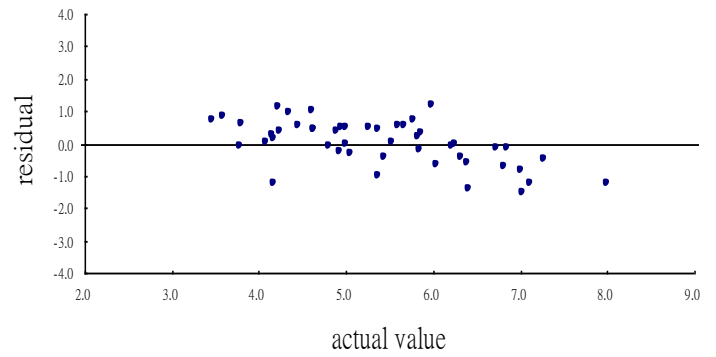


Figure 4. The scatter plots of residual and actual value of yield in first order linear models using (a) BLUE, (b) GRN, (c) RED, (d) NIR, (e) BLUE and GRN, (f) BLUE and RED, (g) BLUE and NIR, (h) GRM and RED, (i) GRN and NIR, (j) RED and NIR, (k) BLUE, GRN and RED, (l) BLUE, GRN and NIR, (m) BLUE, RED and NIR, (n) GRN, RED and NIR, (o) BLUE, GRN, RED and NIR.

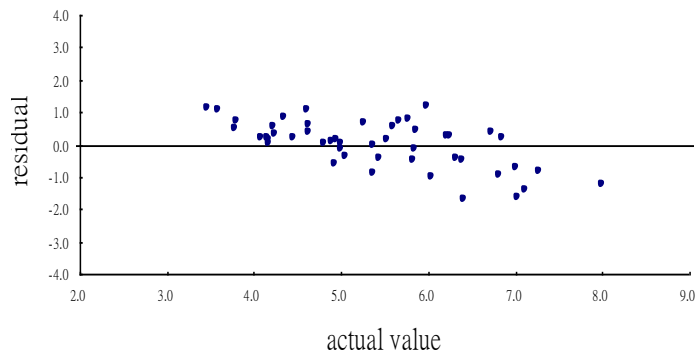
(a) NDVI



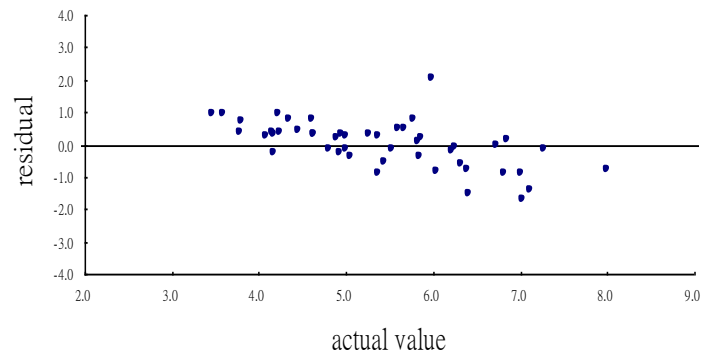
(b) GNDVI



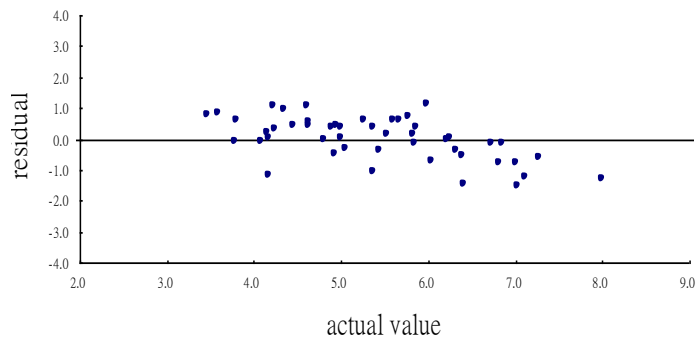
(c) SR



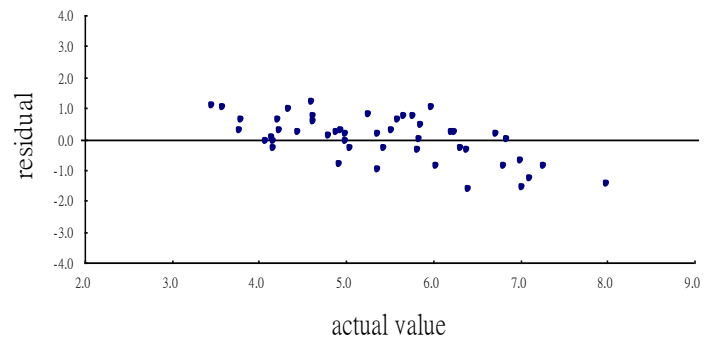
(d) NIR/GRN



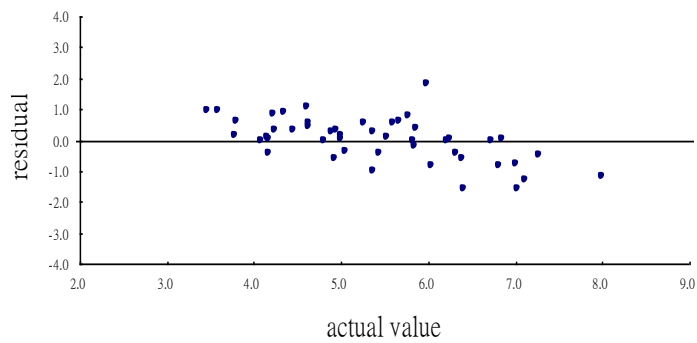
(e) NDVI GNDVI



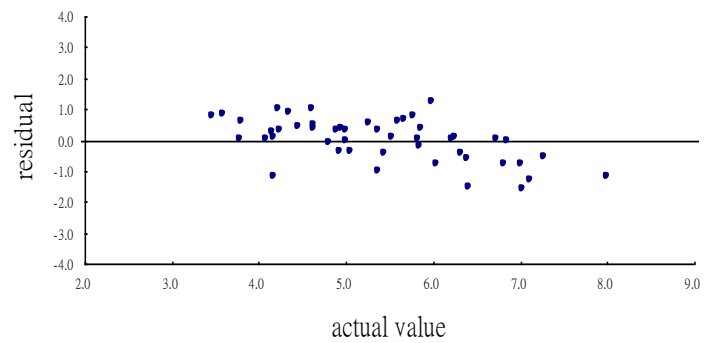
(f) NDVI SR



(g) NDVI NIR/GRN



(h) GNDVI SR



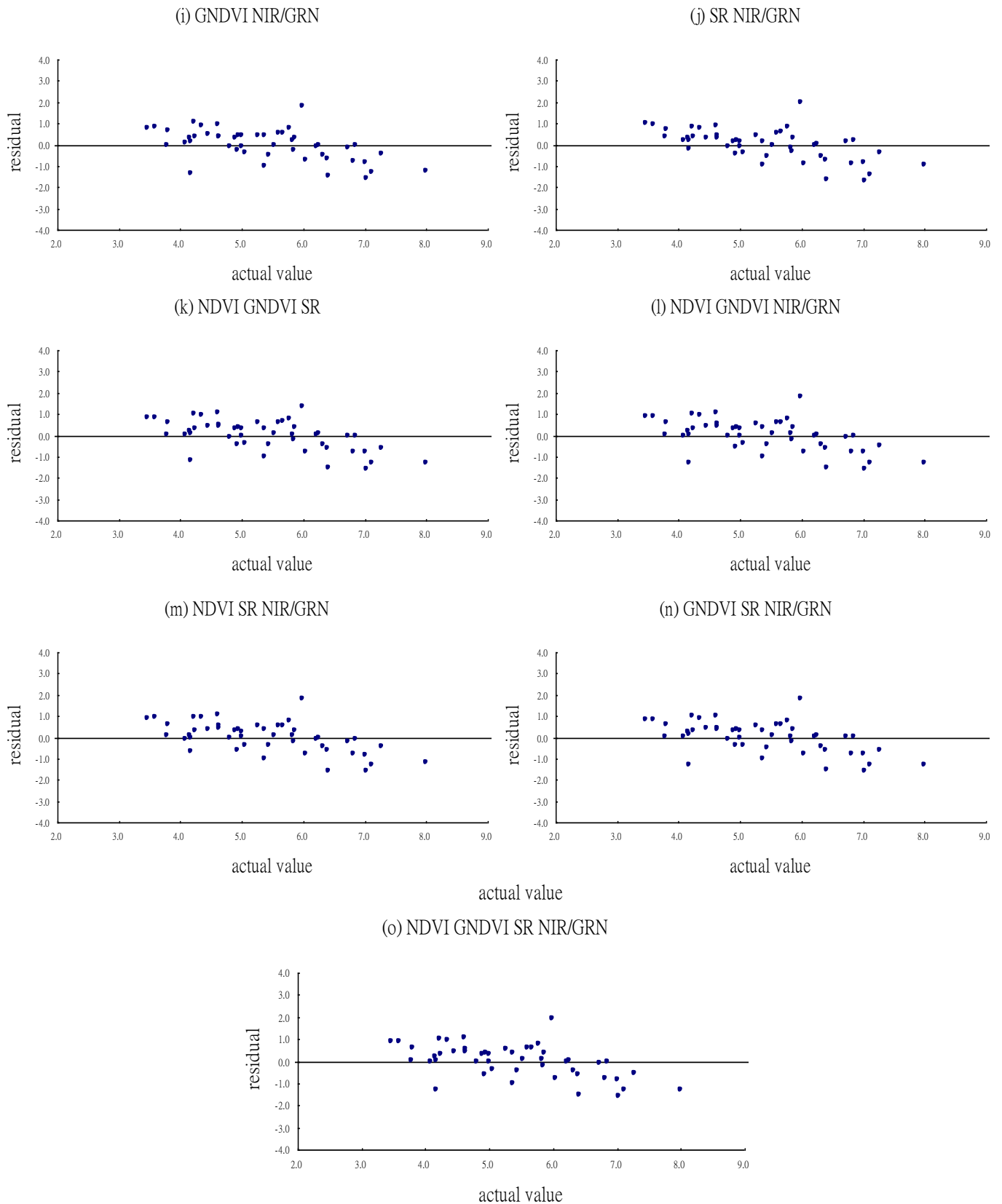


Figure 5. The scatter plots of residual and actual value of yield in linear models using (a) NDVI, (b) GNDVI, (c) SR, (d) NIR/GRN, (e) NDVI and GNDVI, (f) NDVI and SR, (g) NDVI and NIR/GRN, (h) GNDVI and SR, (i) GNDVI and NIR/GRN, (j) SR and NIR/GRN, (k) NDVI, GNDVI and SR, (l) NDVI, GNDVI and NIR/GRN, (m) NDVI, SR and NIR/GRN, (n) GNDVI, SR and NIR/GRN, (o) NDVI, GNDVI, SR and NIR/GRN.

sensing images is not only fundamental to applications of precision agriculture, but can also be very useful to food provisions management.

ACKNOWLEDGEMENT

Financial support for this research came from project (NSC 99-2632-H-156 -001 -MY3) of National Science Council, Taiwan, ROC.

REFERENCES

- Arslan S, Colvin TS (1999). Laboratory performance of a yield monitor. *Appl. Eng. Agric.* 15:189-195.
- Bach H (1998). Yield estimation of corn based on multitemporal Landsat-TM data as input for an agrometeorological model. *Pure Appl. Optics* 7:809-825.
- Bouman BAM (1992). Linking physical remote sensing models with crop growth simulation models, applied for sugar beet. *Int. J. Remote Sens.* 13:2565-2581.
- Clevers JGPW, van Leeuwen HJC (1996). Combined use of optical and microwave remote sensing data for crop growth monitoring. *Remote Sens. Environ.* 56:42-51.
- Field CB, Randerson JT, Malmstrom CM (1995). Global net primary production: combining ecology and remote sensing. *Remote Sens. Environ.* 51:74-88.
- Lark RM, Stafford JV, Bolam HC (1997). Limitations on the spatial resolution of yield mapping for combinable crops. *J. Agric. Eng. Res.* 66:183-193.
- Leblon B, Guerif M, Baret F (1991). The use of remotely sensed data in estimation of PAR use efficiency and biomass production of flooded rice. *Remote Sens. Environ.* 38:147-158.
- Moulin S, Bondeau A, Delecolle R (1998). Combining agricultural crop models and satellite observations: from field to regional scales. *Int. J. Remote Sens.* 19:1021-1036.
- National Research council. [NRC] (1997). Precision agriculture in the 21st century: Geospatial and information technologies in crop management. Rep. 59-0700-4-139. NRC, Washington, DC.
- Neter J, Kutner MH, Nachtsheim CJ, Wasserman W (1999). Applied linear statistical models. McGraw-Hill. pp. 54-56.
- Prince SD (1991). A model of regional primary production for use with coarse-resolution satellite data. *Int. J. Remote Sens.* 12:1313-1330.
- Reynolds CA, Yitayew M, Slack DC, Hutchinson CF, Huete A, Petersen MS (2000). Estimating crop yields and production by integrating the FAO crop specific water balance model with real-time satellite data and ground-based ancillary data. *Int. J. Remote Sens.* 21:3487-3508.
- Serrano L, Fillela I, Penuelas J (2000). Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Sci.* 40:723-731.
- Wiegand CL, Maas SJJ, Aase K, Hatfield JL, Pinter Jr PJ, Jackson RD, Kanemasu ET, Lapitan RL (1992). Multisite analyses of spectral-biophysical data for wheat. *Remote Sens. Environ.* 42:1-21.