Comparing Multiple Regression, Principal Componant Analysis, Partial Least Square Regression and Ridge Regression in Predicting Rangeland Biomass in the Semi Steppe Rangeland of Iran

Ali.Goharnejad [1]*, Azin.Zarei [1] and Pejman.Tahmasebi [2]

[1] *Ph.D. student of Range Management, Department of Rehabilitation of Arid and Mountainous Regions, University of Shahrekord. Tel: +989144465862.Fax:0381-4424423, Iran*

[2] *Department of Range and Watershed management, Faculty of Natural Resource and Earth Science , University of Shahrekord, Iran,*

Abstract

In this paper, the prediction of rangeland biomass using different methods including Multiple regression, Principal Component Analysis, Partial Least Square regression and Ridge regression were compared. For this goal, environmental factors such as elevation (m), rainfall (mm), slope (٪), caco3 (٪), Sand (٪), Sand (٪), Clay (٪), pH, EC (ds.m-1), Saturation (SP (٪)) were used to determine a relationship between environmental factors and the forage yield. The results showed that PLS, and ridge regressions were among the best models to predict rangeland biomass followed by multiple regressions and Principal component Analysis. PLS and ridge regression had a higher predicted accuracy than other calculation methods. It was shown that, the Partial Least Square regression values to R, RMSE and MAE were 0.83, 34.9 and 26.23, respectively.

Key words: Biomass/ Multiple regressions/ PCA ordination/ Ridge regression/ Partial Least Squares (PLS)/ NDVI index

1. Introduction

Plant biomass can be considered a diagnostic indicator for desertification assessment, and determine the health of rangeland for grazing capacity (Van Zee, and Havastad, 1998; Rubio and Bochet, 1998; Zhang and Chen, 2007). Plant biomass is an important factor in the study of functional plant biology and

*Corresponding author:
E-mail: Goharnejad.a@gmail.com

growth analysis, and is the basis for the calculation of net primary production and growth rate (Tackenberg, 2007). A reliable estimate of plant biomass is essential for ecological studies, but there is no universal method to estimate biomass validity across all plant communities and landscapes (Catchpole and Wheeler 1992). Two categories of methods are commonly used. The first one, which is mostly conducted when calibrating other indices, is to harvest, hand-sort and weigh samples of vegetation from a number of quadrates (Catchpole and Wheeler 1992). Direct harvesting techniques for estimating biomass are labor intensive and time consuming. Biomass of plant measurement via direction methods such as harvesting is destructive, expensive and time consuming (Reese et al. 1980). The second category of methods relies on indirect indices that are assumed to relate to biomass, if possible linearly. This includes several pasture plates or disk meters (Bransby et al. 1977, Schneider and Bell 1985), and methods such as point counting of various sorts: including the point-quadrat (Goodall 1952), canopy intercept (Frank and McNaughton 1990), point intercept (Jonasson 1988), point contact (Catchpole and Wheeler 1992)

and 3D quadrat methods (Said et al. 2005). These methods are non-destructive, but are often neither quick nor cheap. Alternatively, another indirect method has been developed based on the calibrated visual assessment of biomass (e.g. as used in the visual estimation component of the BOTANAL method, Tothill et al. 1992). The latter is typically quick (after preliminary training sessions) and has been proven to perform well in mountain grassland and tropical rangeland situations (Coser et al. 1991, Dall'Agnol et al. 2005, Lavorel et al. 2008, Lopez-Guerrero et al. 1999, Mellors 1991, t'Mannetje and Jones 2000). Fang et al. (2006) studied the relationship between above-ground biomass and RVI and suggested that linear regression models would better fit this relationship. However, due to compressing the rich information contained in the original data, the single vegetation index derived from remote sensing data causes the estimation model to be unstable and unviable, especially in those areas where topographical features and vegetation are heterogeneous (Murakami, 2004; Jiang et al., 2006;Liu et al., 2006; Cho, Skidmore, Corsi, van Wieren, and Sobhan, 2007).

Multivariate linear regression models are often used as alternative

methods to derive the relationship between multiple explanatory variables and the response variable. Multivariate statistics concerns understanding the different aims and background of each of the different forms of multivariate analysis, and how they relate to each other. The practical implementation of multivariate statistics to a particular problem may involve several types of univariate and multivariate analysis in order to understand the relationships between variables and their relevance to the actual problem being studied. In addition, multivariate statistics are concerned with multivariate probability distributions, this gives rise to the questions;

- how these can be used to represent the distributions of observed data ?
- how they can be used as part of statistical inference, particularly where several different quantities are of interest to the same analysis ? .

A multivariate the specific advantages of multivariate studies are such as, more data can be analyzed simultaneously, thereby providing greater statistical power and Regression models can give more insight into relationships between variables. also, The focus is on relationships among variables rather than on isolated individual factors.

Furthermore, ridge regression and partial least squares (PLS) regression approaches, due to their robustness to collinearity, which generally exist between multiple variables, are widely applied (Cai, Ci, Li, and Hong, 2003; Hansen and Schjoerring, 2003;Kooistra et al., 2003; Cho et al., 2007).Principal components analysis (PCA) creates a new set of orthogonal variables that contain the same information as the original set. It rotates the axes of variation to give a new set of orthogonal axes, ordered so that they summarize decreasing proportions of the variation.

In this study, we evaluated the precision of four models for production estimation, i.e., the PLS regression model, multilinear regression, PCA ordination and the ridge regression model, in Nazlou region with obtained images from Landsat2011 satellite.

2. Materials and Methods:

2.1 Study area

The area of study (Nazlou) is located in the north- west of the Iran in the west Azarbayjan province and the

remaining part of it, is located in Turkey (Figure 1) with an area of about 1480 km^2. The region lies within latitudes 37º28' and 37º59' N and within longitudes 44º 23' and 45º 08' E. The maximum elevation of the region is 3586

m and the minimum elevation is 1284 m. The climate of the region is semi arid according to Domarten method. Also, the maximum temperature is about 39 in July and minimum temperature -33 in January.
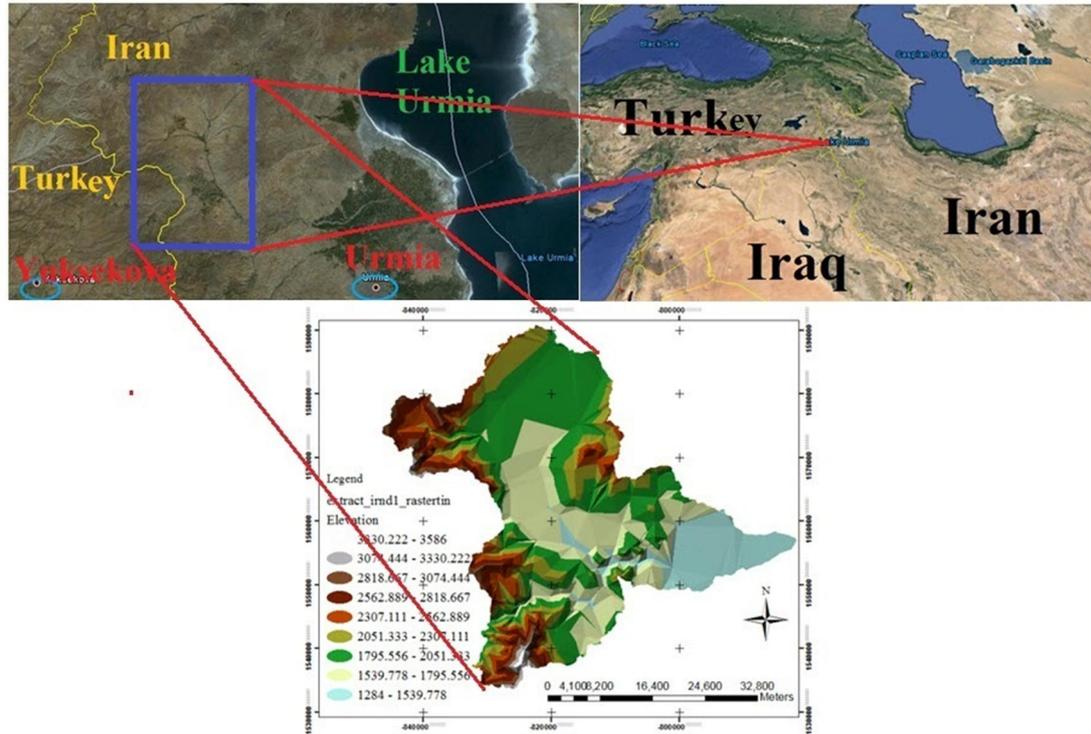

Figure 1: Location map of the study area

2.2 Measuring plant biomass and environmental factors

Forage yield is the actual amount of forage dry matter available in a pasture or rangeland on a per acre basis. In this research, production was determined by clipping and weighing method and NDVI method. Applied data from Land sat satellites in 2011, Also field data was

harvested in the same year. To integrate field data with satellite data, topographic map with 1:50000 scales were used for geo reference. The maps of land cover were prepared by use of satellite images, then the map was completed by field visits. The land cover map and assessment of primary production using remote sensing (NDVI index) and clipping and

weighing methods that were produced the following diagram shown in Figure 2. The geographic position of the site was obtained with GPS (global positioning system). To assess the biomass of plants, they were harvested in the field. Vegetation sampling was conducted in 75plots (Plot size was estimated based on a minimum of 1 square meter). In order to, after the parts are cut from a piece of grazing species in terms of the palatability degree, dry forage was calculated in each class level per kg/ha. Environmental factors such as elevation (m), Rainfall (mm), slope (⁒), Caco3 (⁒), sand (⁒), silt (⁒), clay (⁒), PH, EC (ds.m-1), Percent Saturation (SP(⁒)) were used to determine relation between environmental factors and forage yield. In order to determine the environmental factors, average rainfall data over 30 years from the stations in the basin and its margin were prepared the Map Curves Rain, then by interpolation, Average annual rainfall maps were determined for each pixel (30 x 30 m). To prepare the elevation and slope data was used on a topography map with 1:50000 scale and digital elevation model (DEM) in Arc GIS 9.3 software, Also, 35 soil profiles were taken to determine edaphically factor in each land unit.
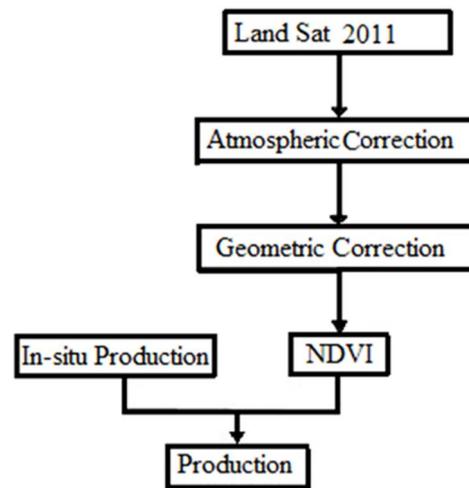

Figure 2: Flow chart of production mapping using remote-sensing data

In second method, the production was estimated by NDVI derived from satellite images in ENVI 4.2 software the satellite data include ETM Sensor of land Sat 2011 images. The NDVI index formula is (Eq. 1).

$$NDVI = \frac{(ETM4 - ETM3)}{(ETM4 + ETM3)} \qquad (1)$$

That NDVI reflect vegetation cover index:

ETM4: Band 4 in ETM sensor of Land sat satellite

ETM3: Band 3 in ETM sensor of Land sat satellite

ETM4 and ETM3 are spectral reflectance in the near-infrared band and reflectance in the red band, respectively. This is a less measure unit that their range

usually is between -1 to +1. The pixel value is closer to +1 density of vegetation is higher (Figure 3).
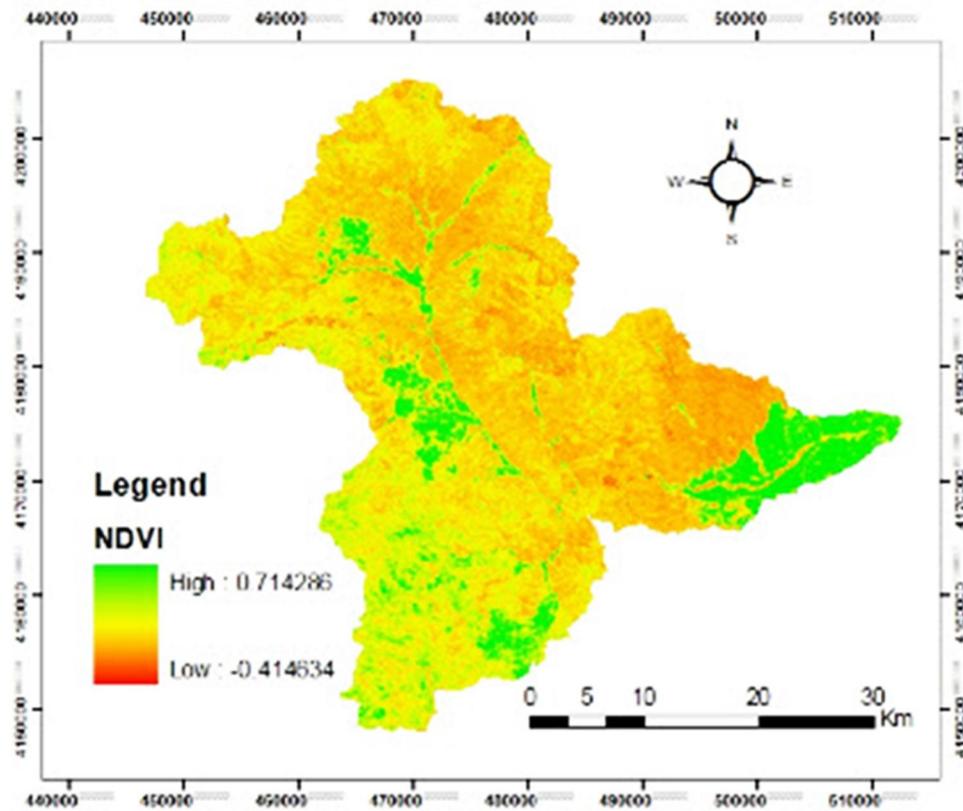

Figure 3: Vegetation cover NDVI index map of Nazlou

2.3 Research methods

In this paper, In order to determine the most appropriate method of plant biomass four methods of multiple regression (backward), ordination (PCA), Partial Least Square regression (PLS) and Ridge regression were used.

Multiple regressions are a statistical technique that allows us to predict someone's score on one variable on the basis of their scores on several other variables. Regression analysis is a statistical technique for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable (Y) and one or more independent variables (Xi). (Eq. 2).

$$Y = a + b_1X_1 + b_2X_2 + \cdots + b_kX_K \quad (2)$$

Multiple regression model of estimating forage determined was performed using SPSS statistical software

from multiple regression and backward method, so that, autocorrelation data lost by using backward method. In this method, biomass estimates based on vegetation indices of NDVI as a dependent variable (Figure 4) and other environmental parameter such as (silt, sand, clay, Lime and moisture content of saturated, PH, EC) and precipitation, elevation, percent slope in the region were considered as independent variables.
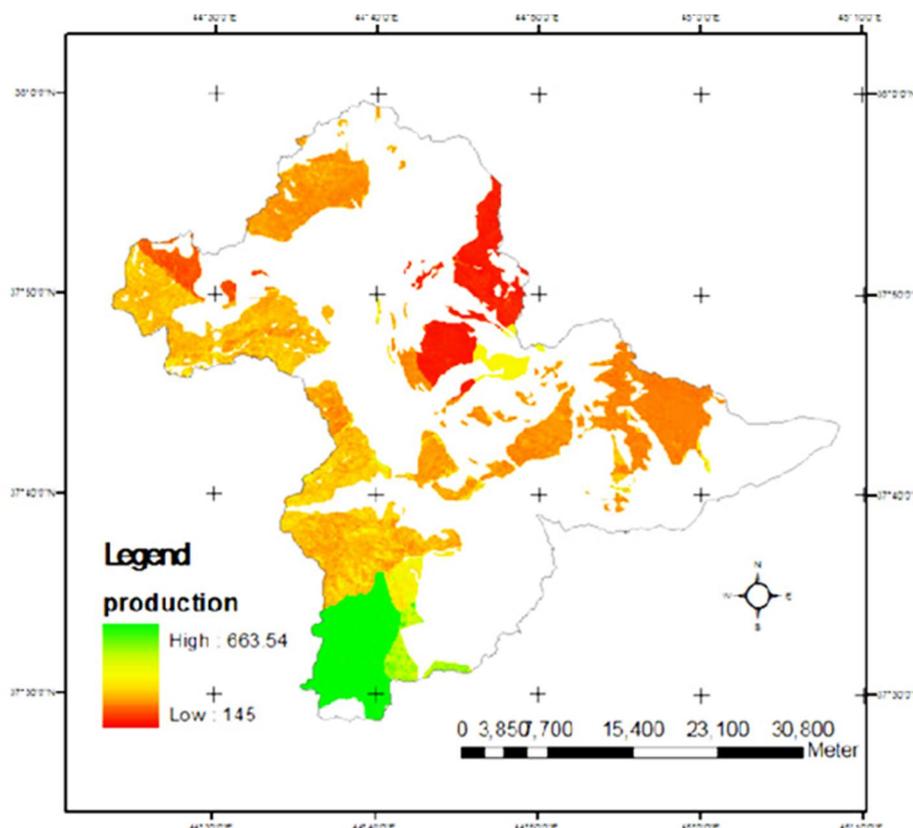

Figure 4: Yield map by plant cover index of NDVI nazlou

PCA ordination method was used to determine the model of rangeland forage in PC-Ord software. By this method studied variables have been reduced. It was useful when you have obtained data on a number of variables. So those, relation between variables and total relation of model were determined by using the eigenvalues of each parameters and sampling profiles. In this study, according to equation (10) was determined relation between parameters and then, was obtained correlation among 35 sampling units in the first vector of PCA with sampling of biomass. So that, biomass value can be calculated by using the (9) equation.

In this study, ridge regression method was used to determine model of rangeland forage in Matlab software. The ridge regression method was developed by Hoerl and Kennard in 1970 (Hoerl and Kennard, 1970). b = ridge(y, X, k) returns a vector b of coefficient estimates for a multi linear ridge regression of the responses in y on the predictors in X. X is an n-by-p matrix of p predictors at each of n observations. y is an n-by-1 vector of observed responses. k is a vector of ridge parameters. If k has m elements, b is p-by-m. By default, b is computed after centering and scaling the predictors to have mean 0 and standard deviation 1. The model does not include a constant term, and X should not contain a column of 1s.b = ridge (y, X, k, scaled) uses the {0,1}-valued flag scaled to determine if the coefficient estimates in b are restored to the scale of the original data. Ridge (y, X, k, 0) performs this additional transformation. In this case, b contains p+1 coefficient for each value of k, with the first row corresponding to a constant term in the model. Ridge (y, X, k, 1) is the same as ridge (y, X, k). In this case, b contains p coefficients, without a coefficient for a constant term.

The relationship between b0 = ridge (y,X,k,0) and b1 = ridge(y,X,k,1) is given by

$$m = mean(X);$$
$$s = STD(X,0,1)';$$
$$b1\_scaled = b1. /s;$$
$$b0 = [mean(y)-m*b1\_scaled;$$
b1_scaled]

This can be seen by replacing the x_i (i = 1... n) in the multilinear model y = b00 + b10x1 + ... + bn0xn with the z-scores $z_i = (x_i − \mu_i)/\sigma_i$, and replacing y with y − μ_y.

In general, b1 is more useful for producing plots in which the coefficients are to be displayed on the same scale, such as a ridge trace (a plot of the regression coefficients as a function of the ridge parameter). b0 is more useful for making predictions. Coefficient estimates for multiple linear regression models rely on the independence of the model terms. When terms are correlated and the columns of the design matrix X have an approximate linear dependence, the matrix $(X^TX)^{-1}$ becomes close to singular. $\widehat{B}$ is calculated according to (Eq. 3):

$$\widehat{B} = (X^TX)^{-1}X^Ty \qquad (3)$$

Where

$\widehat{B}$: is the least-squares

X: is an n-by-p independent variable matrix,

X^T: is the transpose of X,

Y: is an n-by-1 vector of observations (biomass),

As a result, the least-squares estimate becomes highly sensitive to random errors in the observed response y, producing a large variance. This situation of multi collinearity can arise, for example, when data are collected without an experimental design. Ridge regression addresses the problem by estimating regression coefficients use where k is the ridge parameter and I is the identity matrix. Small positive values of k improve the conditioning of the problem and reduce the variance of the estimates. While biased, the reduced variance of ridge estimates often result in a smaller mean square error when compared to least-squares estimates. $\widehat{B}$is calculated according to (Eq.4):

$$\widehat{B} = (X^T X + KI)^{-1} X^T y \qquad (4)$$

K: is a scalar constant (the ridge parameter).

I: is an n-by-n identity matrix

Hoerl and Kennard (1970) suggested Ridge Trace for determining k. Ridge trace is an easily applicaple method. It is obtained by plotting ' $\widehat{B}_R$ s versus k values which are usually taken in the interval of [0,1]. While k increases the values of B^_R's will stabilize. And k can be chosen as the smallest value which stabilizes all the $\widehat{B}_R$'s. Sinan1and Genç (2012) suggested that The basic method for determining ridge parameter is known as ridge trace. In this study, The best choice for k is 0.5.

The PLS regression method was developed by Wold et al. in 1983 (Wold et al., 1983).Since then, it has been widely applied and perfected in many fields (Hansenand Schjoerring, 2003; Næsset, Bollandsas, and Gobakken, 2005). It integrates fundamental functions of multiple linear regression, principal component analysis, and canonical correlation analysis together. Partial least squares have modeling advantages when using small samples and collinearity data. It extracts new synthetic variables from the original explanatory data and builds the regression equation by matching them for dependent variables. Then it is used to regress the equation using the original variables associated with the dependent. The number of new variables

(components) is determined by the cross-validation indicator (Wang, Wu, and Meng, 2006). The cross-validation indicator, defined as Qh, is calculated according to (Eq. 5):

$$Q_h^2 = 1 - \frac{S_{PRESS,h}}{S_{ss,h-1}} \qquad (5)$$

$S_{PRESS,h}$ and $S_{SS,h-1}$ calculated according to (Eq. 6, 7):

$$S_{PRESS,h} = \sum_{i=1}^{n}\left(y_i - \hat{y}_{h(-i)}\right)^2, \qquad (6)$$

$$S_{SS,h-1} = \sum_{i=1}^{n}\left(y_i - \hat{y}_{(h-1)i}\right)^2 \qquad (7)$$

n: is the number of sample plots,

y_i :is the biomass observation,

$\hat{y}_h(-i)$ is the ith sample simulated biomass value using the regression model that includes h components built with sample data, except the ith sample, $\hat{y}_{hi}$, is the ith sample simulated biomass value using the regression model, which includes $h - 1$ components built with all sample data. If Q_h^2 is not less than 0.0975, i.e., the ratio of $S_{PRESS,h}/S_{ss,h-1}$ is no more than 0.05, the hth component should be introduced in the regression model.

3. Results

3.1 Multiple regression (backward)

The correlation of studied variables with yield were determined, among the studied variables only rainfall and slope factors, respectively, had the greatest effect on the biomass (Figure 5, 6). Graphs show correlation of 0.67 between biomass and rainfall and 0.546 correlations between the biomass and slope.
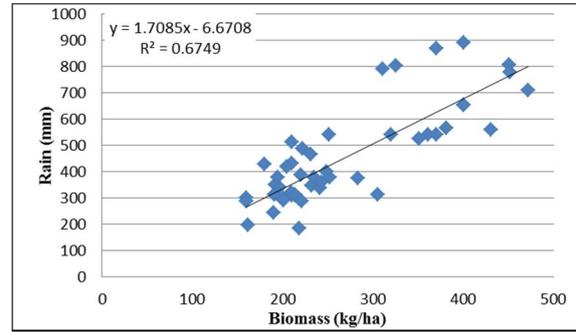
Figure 5: Relationship between rainfall and biomass data in Nazlou basin
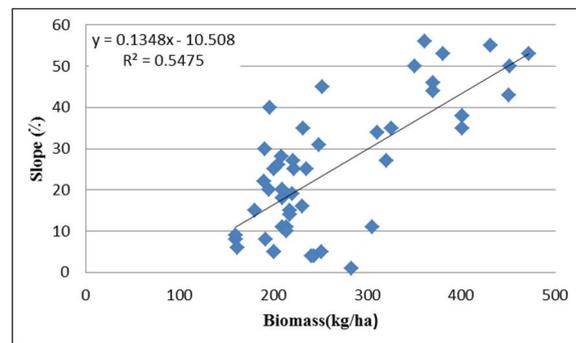
Figure 6: Relationship between percent slope and biomass data in Nazlou basin

Equation (8) was obtained the base on relation slope and rainfall with forage production in multiple regression method in SPSS software.

Biomass = 100.303 + 1.635Slope

$$+ 0.278Rain \qquad (8)$$

According to equation (8), among 10 studied parameters, rain and slope have the biggest effect on biomass.

3.2 Principal Componant Analysis (PCA)

In PCA model, Relation between studied variables has been shown in the Figure 5. Also, eigenvector of six axes for every factors has been indicated, which, slope, SP, EC, silt and sand have the most effective on biomass the based on eigenvalues of each parameter (Table1).
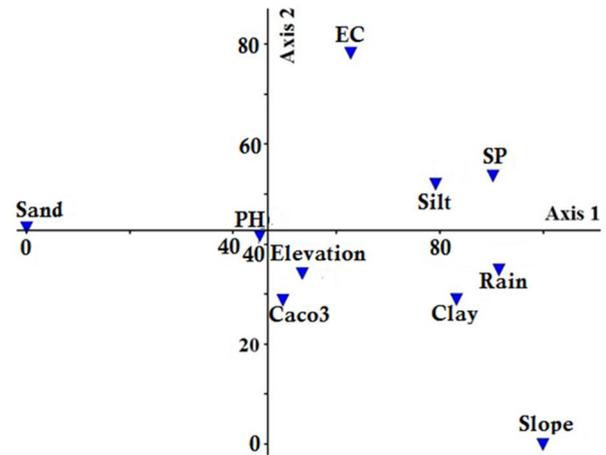

Figure 7: PCA ordination graph

Table 1:Eigenvector of six axes for every factor.

Factor	Eigenvector					
	1	2	3	4	5	6
Elevation (m)	0.169	-0.428	0.352	-0.33	0.3441	0.485
Slope (٪)	0.273	-0.424	0.299	0.371	-0.162	0.096
Rain (mm)	0.378	-0.128	0.243	0.467	0.067	-0.308
Sp (٪)	0.442	0.216	-0.02	0.24	0.042	0.046
PH	-0.259	-0.351	-0.368	0.308	-0.221	0.547
EC (ds.m-1)	0.129	0.556	0.241	0.026	0.324	0.484
Clay (٪)	0.363	-0.264	-0.167	-0.511	0.019	-0.184
Silt (٪)	0.384	0.213	-0.365	0.08	-0.275	0.272
Sand (٪)	-0.44	0.011	0.321	0.229	0.162	-0.07
Caco3 (٪)	0.016	-0.153	-0.517	0.242	0.769	-0.1

The equation for this model is as follows (Eq. 9, 10)

$$Biomass = 23.592X + 270 \qquad (9)$$

$$X = 0.169 + 0.378Rain + 0.44SP + 0.364Clay + 0.384Silt - 0.44Sand + 0.1293EC + 0.2736Slope - 0.259PH + 0.016Caco3 \qquad (10)$$

X: Relation between parameters in PCA method

3.3 Ridge regression

Figure 8 represents the ridge trace curves that offer an enlightening view over the stability of the regression estimators depending on parameter k which varies between 0 and 0.50 with step 0.05. It can be seen that while for the variables X1, X2… X9X10 the values of the regression coefficients estimators become stable for different values of k.
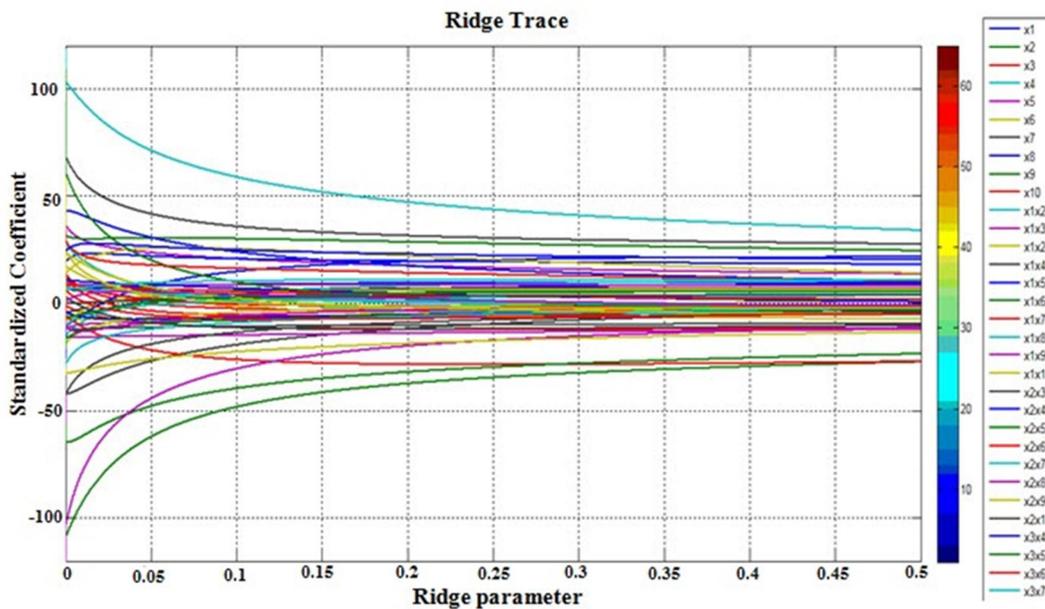

Figure 8: The values of the estimated regression coefficients plotted against k with using ridge trace method

The equation for this model is as follows (Eq.11).

$$Biomass = 8.78Rain(x1) + 5.79Slope(x2) - 11.69E(x3) - 2.4\ SP(x4) - 3.99Clay(x5) - 7.2Silt(x6) + 6.71Sand(x7) + 9.53EC(x8) + 3.82PH(x9) - 3.04Caco3(x10) + 61.88 \qquad (11)$$

3.4 Partial Least Square regression (PLS)

PLS shares characteristics with other regressions and feature transformation techniques. It is similar to ridge regression, in that it is used in situations

with correlated predictors. It is similar to stepwise regression (or more general feature selection techniques), in that it can be used to select a smaller set of model terms. PLS differs from these methods, however, by transforming the original predictor space into the new component space. Actually, PLS and ridge regression has a higher prediction accuracy than other calculation methods. Also, REMSE the PLS and ridge regression model gave lower values than other models. Figure 9 shows the PLS loading plot and PLS score plot.
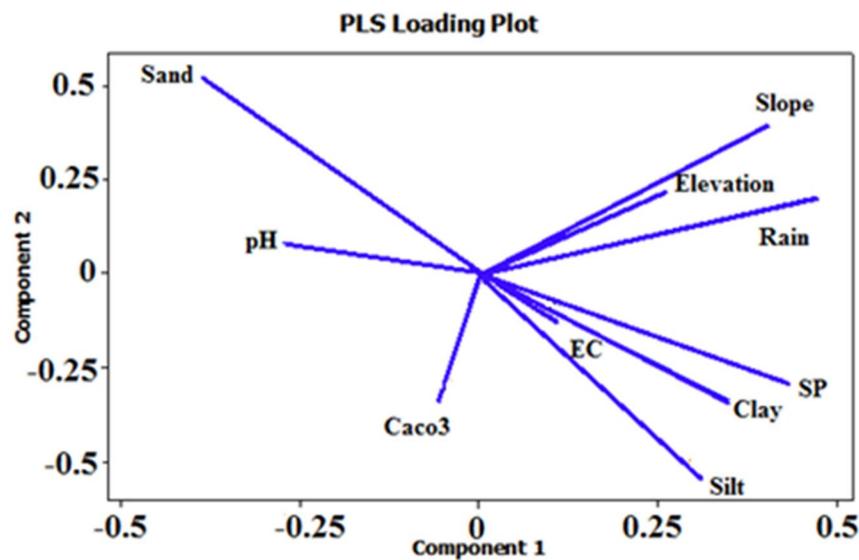

Figure 9: PLS loading plot and PLS score plot

The equation for this model is as follows (Eq. 12)

Biomass = - 91.14 - 0.069Elevation + 2.37slope + 0.3rain - 1.2SP + 39.02pH + 107.46 EC + 2.3clay - 0.79silt - 0.38sand - 1.5caco3   (12)

The accuracy assessment of models:

All models were validated by the predicting the biomass value of the same testing data set. For this goal, 25 test data were used to determine the accuracy of the model, so that the correlation between predicted and actual biomass (Clipping and weighing) were examined for each method. Consequently, the ridge regression and Partial least squares (PLS) work better than PCA ordination and Multi linear regression methods (Table2). Also, the correlation between predicted and actual biomass (Clipping and weighing) by multiple regression, ridge regression, PLS and PCA have been shown in the Figure 10-11.
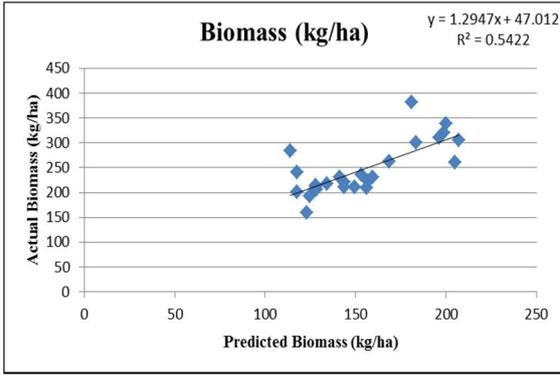
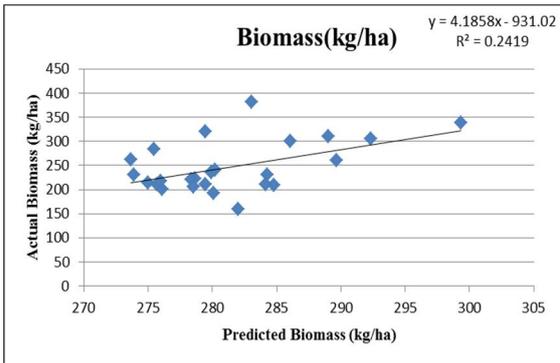**Figure 10**: Relationship between predicted and actual biomass (Clipping and weighing) by multiple regression

Figure 11: Relationship between predicted and actual biomass (Clipping and weighing) by PCA ordination
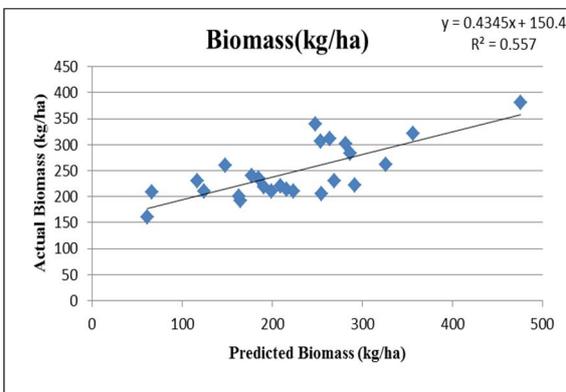

Figure 12: Relationship Correlation between predicted and actual biomass

(Clipping and weighing) by ridge regression
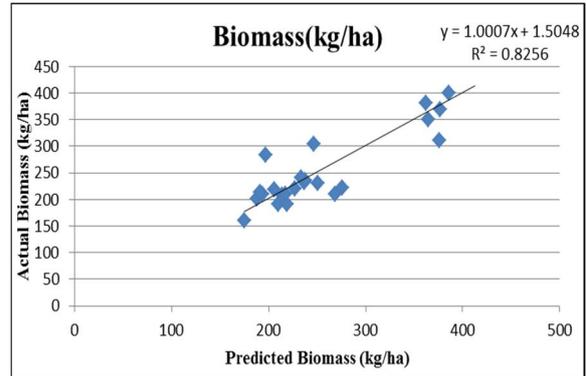
Figure 13: Correlation between predicted and actual biomass (Clipping and weighing) by PLS

in conclusion, there are different indices for the determination of prediction model functions, which is shown in table (2). In this research, the root mean squared error (RMSE), correlation coefficient(R) and mean absolute error (MAE) were used. The root mean squared error (RMSE) can be calculated by using (Eq. 13, 14).

$$RMSE = \left[\frac{\sum_{i=1}^{N}(S_i - O_i)^2}{N}\right]^{0.5} \qquad (13)$$

$$R = \frac{N\sum_{i=1}^{N} S_i O_i - \sum_{i=1}^{N} S_i \sum_{i=1}^{N} O_i}{\left\{\left[N\sum_{i=1}^{N} S_i^2 - \left(\sum_{i=1}^{N} S_i\right)^2\right]\left[N\sum_{i=1}^{N} O_i^2 - \left(\sum_{i=1}^{N} O_i\right)^2\right]\right\}^{0.5}} \qquad (14)$$

That, O_i is observed value in time step of i, S_i predicted value in time step of I, N is number of time step, $\overline{O_t}$ is mean of observed. Also, the mean absolute error (MAE) is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. The mean absolute error is given by (Eq. 15).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |f_i - y_i|$$

$$= \frac{1}{n}\sum_{i=1}^{n} |e_i| \qquad (15)$$

As the name suggests, the mean absolute error is an average of the absolute errors

$e_i = |f_i - y_i|$, where f_i is the prediction and y_i the true value. Note that alternative formulations may include relative frequencies as weight factors. The mean absolute error is a common measure of forecast error in time series analysis, where the terms "mean absolute deviation" is sometimes used in confusion with the more standard definition of mean absolute deviation. The same confusion exists more generally.

Table 2: R value for Multi linear regression, PCA ordination, Partial least squares (PLS) and Ridge regression

Methods	R	RMSE	MAE
Multi linear regression	0.54	66.29	54.5
PCA ordination	0.24	99.41	92.52
Partial least squares (PLS)	0.83	34.9	26.23
Ridge regression	0.56	62.2	52.28

5. Discussion

In general, the results suggest that effective selection of potential variables is essential before constructing the relationship between them, both for saving time and for improving the prediction accuracy of the model. On the other hand, these results also showed that selection of explanatory variables via each method is different. At the same time, the inherent uncertainty of the

remote sensing data also caused estimation error (Giacinto and Roli, 2001; Petit and Lambin, 2002). Hence, other mathematical methods or satellite platform data, or vegetation and land form classification, are worthy of further study in order to improve biomass and production estimate accuracy (Wittich and Hansing, 1995; Hirata et al., 2005). In conclusion, multiple regression analysis, which was used to relate the spectral data to biotic factors, i.e., above- biomass, and the PLS regression, as an effective method dealing with inter correlated explanatory variables, was a better alternative over the ridge regression for semi-arid Nazlou region in Iran. In addition, selected appropriate variables can also contribute to better biomass estimation.

4. Conclusions

Sustainable utilization of rangelands requires updated information on permanent vegetation parameters indices for quantitative estimation of measurement in the long term (Amiri, 2010).

In this paper, 4 methods of biomass estimate were used such as; PLS, ridge regression, Multi linear regression,

PCA and their accuracy was compared using a test data set (n=10). Predictive models based on PLS regression had the highest accuracy compared with those based on other methods. The result of the higher accuracy achieved by PLS regression is consistent with previous research, e.g., Hansen and Schjoerring (2003), Næsset et al. (2005), Cho et al. (2007), Cai et al (2003).There is no standard method to ascertain the ridge parameter (Fang et al.,1988; Wang et al., 2006). of all the methods studed in this paper, PLS, ridge regression, multi linear regression and PCA method have the highest accuracy, respectively. So that, RMSE and MAE values confirm it (Table 2). the PLS regression method improved the biomass estimation precision, the residual error that must be reduced, or even avoided, was still large , and the models need to be enhanced substantially. (Cai et al, 2003).

6. References

Amiri, F. 2008. **Modeling multiple use of rangeland by using GIS.** PH.D thesis, I. A. University Research and Science Branch, Tehran, Iran.

Bransby DI., Matches, AG., and Krause, GF. 1977. Disk meter for rapid

estimation of herbage yield in grazing trials. **Agron. J**. 69: 393–396.

Cai, T. J., L. J. Ci, C. G. Li, and J. Hong. 2003. Quantitative estimation of biomass based on RS and GIS in Mu Us sandland. **Science of Soil and Water Conservation** 1(2): 33–37.

Catchpole, W.R, and Wheeler ,C.J 1992. Estimating plant biomass areview of techniques. **Aust. J. Ecol**. 17: 121–131.

Cho, M. A., A. Skidmore, F. Corsi, S. E. van Wieren, and I. Sobhan. 2007. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. **International Journal of Applied Earth Observations and Geoin formation** 9: 414–424.

Coser AC, Donascimento D, Gomide JA, Dasilva JFC, Silva MDE, Garcia R, and Martins CE 1991. Comparison of Botanal to other methods of pasture evaluation in natural Grasslands. **PesquisaAgr opecuaria Brasileira** 26: 759–767.

Dall'Agnol M, Scheffer-Basso SM, do Nascimento JAL, SilveiraCAM, and Fischer RG 2005. Elephant grass forage yield under cold climate conditions. 2. Production and animal selectivity. **Revista Bra sileira De Zootecnia - Brazilian Journal of Animal Sci** 34: 425-432.

Das, D., K. Mishra, and N. Kalra. 1993. Assessing growth and yield of wheat using remotely-sensed canopytemperature and spectral indices. **Int. J. Remote Sens.** 14 (1993): 3081-3092.

De Soyza, A. G., W. G. Whitford, J. E. Herrick, J. W. Van Zee, and K. M. Havastad. 1998. Early warning indicators of desertification: Examples of tests in the Chihuahuan Desert. **Journal of Arid Environments** 39: 101–112.

Feng, X., C. Tong, L. Zhang, B. L. Miao, Y. Ding, and Y. M. Zhang. 2006. Assessment on grassland degradation at regional-scale in the Baiyinxile ranch, Inner Mongolia. **Journal of Nature Resources** 21(4): 575–583.

Frank DA and McNaughton SJ 1990. Aboveground biomass estimation with the canopy intercept method: a plant growth form caveat. **Oikos** 57: 57–60.

Giacinto, G., and F. Roli. 2001. Design of effective neural network ensembles for image classification purposes. **Image and Vision Computing** 19: 699–707.

Goodall, DW. 1952. Some considerations in the use of point quadrats for the analysis of vegetation. **Aust. J. Scientific Res Series B-Biol Sci** 5: 1–41.

Guevara, J.C., J.M. Gonnet, and O.R. Estevez. 2002. Biomass Estimation for native perennial grasses in the plain of Mendoza, Argentina. **Journal of Arid Environments** 50: 613-619.

Hansen, P. M., and J. K. Schjoerring. 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. **Remote Sensing of Environment** 86(4): 542–553.

Harmoney KR, Moore KJ, George JR, Brummer EC, and Russell JR 1997. Determination of pasture biomass using four indirect methods. **Agron. J.** 89: 665–672.

Hein, L., and N. D. Rodder. 2006. Desertification in Sahel: a reinterpretation. **Global Change Biology** 12: 751–758.

Hirata, M., N. Koga, H. Shinjo, H. Fujita, G. Gintzburger, J. Ishida, and A. Miyazaki. 2005. Measurement of above-ground plant biomass, forage availability and grazing impact by combining satellite image processing and field survey in a dry area of northeastern Syria. **Grass and Forage Science** 60: 25–33.

Hoerl, A. E and Kennard, R. W. 1970. Ridge regression: Biased estimation for nonorthogonal problems. **Thechnometrics** 12(1): 55-67.

Jiang, Z. Y., A. R. Huete, J. Chen, Y. H. Chen, J. Li, G. J. Yan, and X. Y. Zhang. 2006. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. **Remote Sensing of Environment** 101: 336–378.

Jonasson S 1988. Evaluation of the point intercept method for the estimation of plant biomass. **Oikos** 52: 101–106.

Kooistra, L., J. Wanders, G. F. Epema, R. S. E. W. Leuven, R. Wehrens, and L. M. C. Buydens. 2003. The potential of field spectroscopy for the assessment of sediment properties in river floodplains.

Analytica Chimica Acta 484: 189–200.

Lavorel S., Grigulis K., McIntyre S., Williams NSG., Garden D., Dorrough J., Berman S., Quetier F., Thebault A., and Bonis A. 2008. Assessing functional diversity in the field – methodology matters!. **Funct. Ecol.** 22: 134–147.

Liu, Z. Y., J. F. Huang, X. H. Wu, Y. P. Dong, F. M. Wang, and P. T. Liu. 2006. Hyperspectral remote sensing estimation models for the grassland biomass. **Transactions of the CSAE** 22(2): 111–115.

Lopez-Guerrero I, Fontenot JP, Garcia-Peniche TB 1999. Comparison of four biomass estimation methods in Tall Fescue pastures. **Revista Mexicana De Ciencias Pecuarias** 2: 209–220.

Mellors JE. 1991. An evaluation of rapid visual technique for estimating seagrass biomass. **Aquat. Bot**. 42: 67–73.

Moges, S., W.R. Raun, R. Mullen, K. Freeman, G.V. Johnson, and J. B. Solie. 2004. Evaluation of green, red, and near infrared bands for predicting winter wheat biomass, nitrogen uptake, and final grain yield. **Journal of Plant Nutrition** 27(8): 1431-1441.

Murakami, T. 2004. Seasonal variation in classification accuracy of forest-cover types examined by a single band or band combination. **Journal of Forest Research** 9: 211–215.

Naesset, E., O. M. Bollandsas, and T. Gobakken. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. **Remote Sensing of Environment** 94(4): 541–553.

Petit, C. C., and Lambin, E. F. 2002. Impact of data integration technique on historical land-use/land-cover change: Comparing historical maps with remote sensing data in the Belgian Ardennes. **Landscape Ecology** 17(2): 117–132.

Redjadj, C., Duparc, A., Lavorel, S., Grigulis, K., Bonenfant, C., Maillard, D., Saïd, S. and Loison, A. 2012. Estimating herbaceous plant biomass in mountain grasslands: a comparative study using three different methods. **Alpine Botany**. ISSN:1664-2201.

Reese, G.A., R.L. Bayn, and N.E. West. 1980. Evaluation of double-

sampling estimators of subalpine herbage production. **Journal of Range Management** 33: 300-306.

Rubio, J. L., and E. Bochet. 1998. Desertification indicators as diagnosis criteria for desertification risk assessment in Europe. **Journal of Arid Environments** 39:113–120.

Said S, Pellerin M, Guillon N, Debias F, Fritz H 2005. Assessment of forage availability in ecological studies. **Eur. J. Wildl Res.** 51: 242–247.

Schneider BH, Bell DT 1985. A simple, effective technique for rapid measurement of fuels in low shrub communities. **Aust. For Res.** 15(11): 79–84.

Serrano, L., I. Filella, and J. Penuelas. 2000. Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. **Crop Science** 40 (May-June): 723- 731.

Sinan, A., Genç, A. 2012. Comparing the most commonly used classical methods for determining the ridge parameter in ridge regression. **Journal of selçuk university natural and applied science** 2012 1(2).

t'Mannetje Ln and Jones RM. 2000. **Field and laboratory methods for grassland and animal production research.** CAB International, Wallingford.

Tackenberg, O. 2007. A new method for non-destructive measurement of biomass, growth rates, vertical biomass distribution and dry matter content based on digital image analysis. **Annals of Botany** 99: 777-783.

Tothill JC, Hargreaves JNG, Jones RM, and McDonald CK 1992. BOTANAL—a comprehensive sampling and computing procedure for estimating pasture yield and composition. 1. Field sampling. Division of tropical crops and pastures. CSIRO, St Lucia, Queensland.

Tucker, C. J. 1979. Red and Photographic infrared linear combinations for monitoring vegetation. Remote Sens. **Environ**. 8 (19791): 127-150.

Wang, H. W., Z. B. Wu, and J. Meng. 2006. **Partial least-squares regression—Linear and nonlinear methods.** National Defense Industry Press, Beijing.

Wittich, K. P., and O. Hansing. 1995. Area averaged vegetative cover fraction estimated from satellite data. **International Journal of Remote Sensing** 38: 209–215.

Wold, S., C. Albano, W. J. Dunn, K. Esbensen, S. Hellberg, E. Johansson, and M. Sjstrom. 1983. **Pattern recognition: finding and using regularities in multivariate** data, pp. 147–188, in H. Martens and H. Russworm, eds., Food research and data analysis. Applied Science Publishers, London.

Zhang, J. T., and T. G. Chen. 2007. Effects of mixed HippoPHae rhamnoides on community and soil in planted forests in the Eastern Loess Plateau, China. **Ecological Engineering** 31: 115–121.