Journal of Biodiversity and Environmental Sciences (JBES)
ISSN: 2220-6663 (Print) 2222-3045 (Online)
Vol. 5, No. 1, p. 249-257, 2014
http://www.innspub.net

RESEARCH PAPER OPEN ACCESS

Prediction of shrub forms production by integrating statistical and remote sensing data (a case study: Chehelgazi watershed of Sanandaj-Iran)

Salahudin Zahedi[1*], Kaka Shahedi[2]

[1*]*Watershed Science of Sanru and Natural Resource Researcher of Agriculture and Natural Resources Research Center of Kurdistan, Kurdistan, Iran*

[2]*Watershed Sience Department, Faculty of Natural Resources, SANRU, Sari, Iran*

Article published on July 09, 2014

Key words: production estimation, semi-arid shrub lands, multivariate regression, TM sensor.

Abstract

Rangelands vegetation cover is the main resource production of protein in Iran. Inappropriate usage and misknowing of species combination, is the agent of decreasing of valuable species in rangelands. Annual production definition based on growth form or possibly could base on species is the major factor of accurate rangelands management specially grazing programing and natural or intentional fire prevention. This study applied double date TM imagery to estimate production of shrubberies forms of rangelands of Qeshlaq dam watershed. The images were processed by ERDAS IMAGINE software. The rangeland yield clipping and weighing system applied to measure green herbage biomass from ground truth sites by means of 300 medium plots ($5m^2$). Ground truth sites were selected to represent five rangeland types and four sites were sampled by systematic random method in each type to calibrate the relationship between satellite-derived Wavebands, vegetation indices and green yields. These yield data were compared with yields estimated by 6 main wavebands also 4 synthetic bands of 2 scenes of TM data in corresponding time and go through linear multivariate regression processing to make the model. Remote sensing yield estimation model were also analyzed for their precision and checked by actual 10 percent measured yields on four ground truth sites. Results showed that relationship between shrubberies growths is meaningful with, ND53 and TM5/TM3 bands. Resulted model have higher accuracy in estimating shrub forms growth production in order to permanent management of rangelands in comparison with traditional models.

***Corresponding Author:** Salahudin Zahedi ✉ zahedi51@gmail.com

Introduction

Rangelands as an important natural resource in Iran encountered with serious environmental challenges now, which most of them resulted from climate change. Climate change effects on agricultural crops, hydrological cycles and other agriculture systems are very obvious. Drought conditions in recent years, inappropriate utilization, widespread fire and unfamiliarity with plant communities' composition are agents of decrease worthy species which may severely effect rangelands and make them vulnerable about disease, pests, soil erosion and toxic and invader plants (Zahedi and ghasriani 2014).

Aboveground biomass and production can be used to monitor or assess the condition of rangelands. Biomass is used to approximate total carbon content while production is used to track new growth in the plant (Bradley., 2010). Estimating shrubs production often involves separating new plant growth, or "green" biomass, from wood, litter, and other biomass (Ren *et al.* 2012).

The most accurate method to estimate biomass involves clipping, oven-drying, and weighing the plant material (Bonham1989). This clip-and-weigh method is time and labor intensive, making it expensive and impractical for application across broad scales. Technological advances have brought promising new methods to remotely sense and estimate vegetation characteristics. Satellite imagery data are comprehensive responses of vegetation stand structure, vegetation density and vegetation species composition. Different plant structures have different reflectance and texture patterns in various wavelengths, and the relationships between biomass and remotely sensed data are different (Lu *et al*, 2002).

Lu *et al*, (2002) used TM images and field vegetation inventory data to analysis and estimation biomass of moist tropical areas in the Brazilian Amazon basin. Results of this study showed that TM spectral bands and vegetation indices alone are not sufficient to establish an effective model for biomass estimation. But multiple regression models improved biomass estimation performance. Jakubauskas *et al*, (2002) used time series NDVI of NOAA sensor and harmonic, or Fourier analysis (additive, amplitude, and phase terms) to develop an innovative technique for crop identification (corn, soybeans and alfalfa) based on temporal changes in NDVI values. Results showed that for crops that have a single distinct growing season and period of peak greenness, such as corn the majority of the variance was captured by the first and additive terms, while winter wheat exhibited a bimodal NDVI periodicity with the majority of the variance accounted for by the second harmonic term. Jin and Sader (2005) represented that SAVI, NDVI and PVI indices or even simple band ratios depend on shrub types and phonological stages were more sensitive than reflectance from green, red and near infrared bands. These indices had ability to discriminate various shrub species and separate shrub lands from grasslands.

Lawrence *et al*, (2006) used multi-temporal IKONOS imagery of the Missouri Coteau to mapping prairie pothole communities. The authors used classification tree analysis (CTA) with multitemporal imagery. The classification was carried out hierarchically, with three levels of discrimination with level 1 being the broadest classification and level 3 being several vegetation classes. Results of the study found that multi-temporal was important for distinguishing similar vegetation types with phenological variability.

Many studies showed that, through the use of simple regression or multiple regression analysis, correlations between NDVI and crop yield can be derived and used in yield estimation models for different vegetation types (corn, wheat, sugar beets, cotton, canola and grass) in various regions. It is found the suitability of NDVI for yield estimation varies depending upon the acquisition time of the remote sensing images (Ren *et al*, 2012).

Roy and Ravan (1996) tried to develop empirical models with satellite measured spectral response and biomass. The results indicate that there are significant relationships with spectral responses. These relationships have seasonal dependency in varying phonological conditions. The relationships are strongest in visible bands and middle infrared bands. However, spectral biomass models developed using middle infrared bands would be more reliable as compared to the visible bands as the later spectral regions are less sensitive to atmospheric changes. Multiple regression equations using brightness and wetness isolates have been used to predict biomass values. The model used has correlation coefficient of 0·77. Per cent error between observed and predicted biomass was 10·5%. The biomass estimated for the entire national park using stratified and spectral response modelling approaches were compared and showed similarity with the difference of only 4·69%. The results indicate that satellite remote sensing data provide capability of biomass estimation.

Many studies performed using different sensors such as TM, AVIRIS and SPOT in order to separation of plant spices area distribution and have gotten favorite results (Underwood *et al*, 2003).

Results of study by Arzani *et al* (1994) showed that there is no significant difference between estimated biomass and production by means of remote sensing and conventional methods. Results of a study by Duncan (1993) emphasized on efficiency of vegetation indices such as PVI, NDVI and SAVI in various bushes assessment and detachment of shrubbery land and grassland.

In the present study emphasis was on modelling of production estimation of shrubs and shrubs like spices based on field and integrated remote sensing data and statistical predicted method in rangelands which these plants are dominant in order to managing livestock grazing and calculation of flambable materials to fire control and prevention.

Materials and methods

Study area

Chehelgazi watershed with the area of 27 km² is located in northwestern Sanandaj. It is bounded by geo-coordinates UTM38 660447 to 680335 and 3920722 to 3944628 (fig. 1). Mean annual rainfall is approximately 470 mm. Rainfed agriculture (wheat, barley and chickpea) dominates hill foots and semi-flat landscapes and irrigated agriculture and orchard are clustered along the tributary and main river. In overall in the area topography is too rugged for agriculture and Rangelands dominates comprising over 80% of the region. Growth season is varied yearly and its duration is about late March to early August.

Sampling

Vegetation type and landuse maps prepared using two full scene of TM data from 25 Jun and 28 August 2010 false colour composites (FCC) of different bands, supervised and unsupervised classification with ERDAS imagine 9.2 softeware.1:25000 scale topographic map, digital elevation model (Dem) used as ancillary data. For estimating total above ground biomass and production of shrubs area of rangelands was divided into five different types as follows.

1) Ferula hausskenthi-Bromus tomentollus
2) Bromus tomentollus- Ferula hausskenthi
3) Astragalus sp-Bromus tomentollus
4) Astragalus sp-Psathyrostachys fragilis-Bromus tomentollus
5) Astragalus sp- Gundelia

Training sites selected at homogenous area of patches (pure shrub pixels) by systematic random method and pointed by GPS. Ground sampling was carried out by laying 5 m² sample plots in the center of a supposition quadrangular with 90 meter of their sides which contains 9 pixel of TM image. In the homogeneous vegetation area of each pixel the quantitative measurements of plant parameters in 9 plots with the distance of 5 meter measured and clipped.

For each sample, the percentage cover of herbaceous layer was estimated; current yearly production of all plant species were cut and weighed with a portable scale. Afterward the clipped samples dried at shadow and weighed again. Average value of 3×3 pixel counted as ground sample point. Vector digital maps of sampling points converted to raster and they were used with sensor bands and artificial bands such as vegetation indices to extract corresponding values of ground sampling plots.

Statistical analysis

In this study dependent variable is production (biomass) and independent variables are image data such as TM bands and vegetation indices. Multiple regression models are used to establish the relationships between production and remote sensing data.

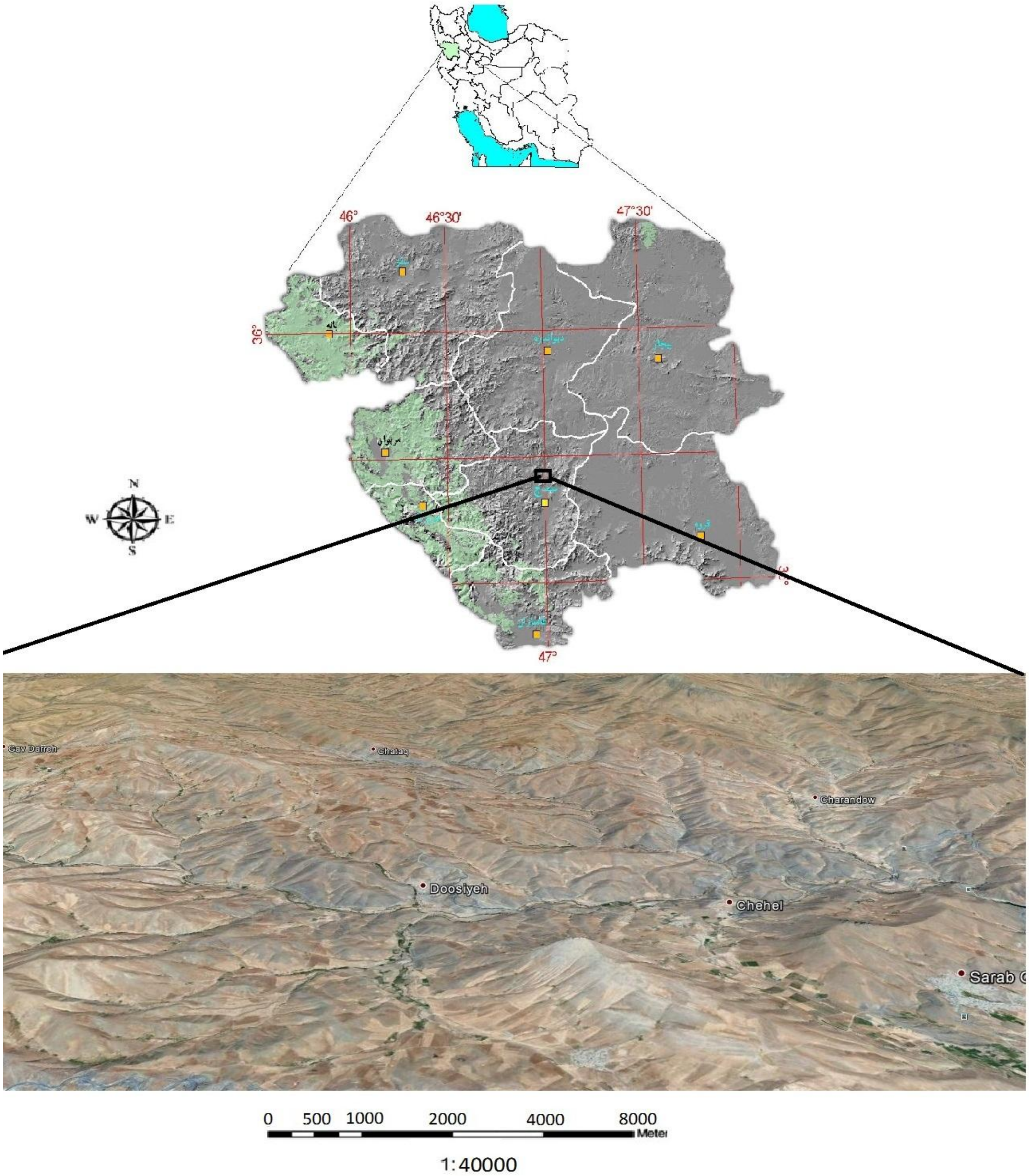

Fig. 1. Study area.

Results

Regression approach and correlation applied in order to determine relationship between variables. Correlation between the DN of TM different bands and artificial bands at the sampling points and production of different plant forms including grasses, forbs and shrubs computed and analyzed. Table 1 shows the correlation coefficients between production's variable of the shrub form and original TM bands and in table2 these values presented for artificial bands and production variable.

Table 1. Coefficients correlation between production parameter and original bands.

		Bush1	TM1	TM2	TM3	TM4	TM5	TM7
Bush1	Pearson Correlation	1	-.689**	-.654**	-.660**	.542**	-.066	-.464**
	Sig. (2-tailed)		.000	.000	.000	.000	.633	.000
	N	54	54	54	54	54	54	54
TM1	Pearson Correlation	-.689	1	.950**	.881**	-.135	.365**	.668**
	Sig. (2-tailed)	.000		.000	.000	.329	.007	.000
	N	54	54	54	54	54	54	54
TM2	Pearson Correlation	-.654	.950**	1	.895**	-.004	.483**	.774**
	Sig. (2-tailed)	.000	.000		.000	.975	.000	.000
	N	54	54	54	54	54	54	54
TM3	Pearson Correlation	.4160**	.881**	.895**	1	.050	.454**	.686**
	Sig. (2-tailed)	.000	.000	.000		.721	.001	.000
	N	54	54	54	54	54	54	54
TM4	Pearson Correlation	.542	-.135	-.004	.050	1	.437**	.135
	Sig. (2-tailed)	.271	.329	.975	.721		.001	.329
	N	54	54	54	54	54	54	54
TM5	Pearson Correlation	.66**	.365**	.483**	.454**	.437**	1	.837**
	Sig. (2-tailed)	.000	.007	.000	.001	.001		.000
	N	54	54	54	54	54	54	54
TM7	Pearson Correlation	.464**	.668**	.774**	.686**	.135	.837**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.329	.000	
	N	54	54	54	54	54	54	54

**. Correlation is significant at the 0.01 level (2-tailed).

Linear multivariate regression based on Equation (1) used to data analysis:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + ..b_n X_n + \varepsilon_i \quad \text{(Equation 1)}$$

Where Y is the predicted shrub production,

b_0 = constant value for model

b_1, b_2 and b_n are coefficients of independent variables

ε_i = Error of model

First the data base established and stepwise regression analysis was used carried out in SPSS software. Residuals analysis is an explicit and effective method in order to clarify defects of model in regression analysis. The residuals defined as:

$$e_i = Y_i - \hat{Y}_i \quad \text{(Equation 2)}$$

The difference between the observed value of the dependent variable (y_i) and the predicted value ($\hat{y}_i$) is called the residual (e_i). Each data point has one residual.

Residual = Observed value - Predicted value

Both the sum and the mean of the residuals are equal to zero. That is, Σ e= 0 and e= 0. Corresponding with each e_i, standardized residual (d_i) defined as:

$$d_i = e_i / \sqrt{MSD} \quad \text{(Equation 3)}$$

Where, MSD is mean squares residuals. Mean of standardized residual (d_i) is zero and its standard deviation is unit. Using normal probability graph and Kolmogorov-Smirnov test, we may peruse this hypotheses which the residuals have normal distribution. Calibration data set was randomly divided into a training subset consisting of 90% of the data and a validation subset consisting of the remaining 10% that was used to estimate the prediction error for variable and model complexity selection. Ground true production of samples compassion by estimated values of model separately. Acceptance criterion of the model is occurring sampling production at confine of predicted model. Confine of predicted model computed by SPSS software at 95% Level of significance. General form, 100 %(1- α) predicted range Y for supposition X_0 calculated using (Equation 4)

$$\hat{Y}_0 \pm t_{\alpha/2,} n-p-1\sqrt{MS_{res}}(1+\frac{1}{n}+X'.(X'X)^{-1}X_0)^{1/2}$$

(Equation4)

In proposed model for estimation of shrubs form production, predicted parameters are ND53 and TM5/TM3 and other bands and Indices because of low correlation exited from the model. Relationship between dependent variable (production) and independent variables (vegetation indices) is significant. Results of regression analysis presented at table3 and 4.

Table 2. Coefficients correlation between production parameter and artificial bands.

		Bush1	ND53	TM5/TM3	NDVI	RVI
Bush1	Pearson Correlation	1	.863**	.794**	.854**	-.859**
	Sig. (2-tailed)		.000	.000	.000	.000
	N	54	54	54	54	54
ND53	Pearson Correlation	.863**	1	.968**	.989**	-.990**
	Sig. (2-tailed)	.000		.000	.000	.000
	N	54	54	54	54	54
TM5/TM3	Pearson Correlation	.794**	.968**	1	.956**	-.954**
	Sig. (2-tailed)	.000	.000		.000	.000
	N	54	54	54	54	54
NDVI	Pearson Correlation	.854**	.989**	.956**	1	-.998**
	Sig. (2-tailed)	.000	.000	.000		.000
	N	54	54	54	54	54
RVI	Pearson Correlation	-.859**	-.990**	-.954**	-.998**	1
	Sig. (2-tailed)	.000	.000	.000	.000	
	N	54	54	54	54	54

**. Correlation is significant at the 0.01 level (2-tailed).

Table 3. Proposed model summary using original and artificial bands.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	F	Sig.
1	.863a	.745	.740	12.144758	151.976	.000a
2	.879b	.773	.764	11.569347	86.885	.000b

a. Predictors: (Constant), ND53
b. Predictors: (Constant), ND53, TM5/TM3

Table 4. Regression coefficients of production Proposed model and TM sensor bands.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17.217	2.530		6.805	.000
	ND53	391.571	31.763	.863	12.328	.000
3	(Constant)	16.069	2.453		6.550	.000
	ND53	686.882	121.473	1.514	5.655	.000
	TM5/TM3	-236.744	94.312	-.672	-2.510	.015

$$production-g/pic = b_0 + b_1 ND53 + b_2 TM5/TM3 + \alpha$$

a. Dependent Variable: Shrub

Corrected specific coefficient related to proposed model showed that 74% of variance of shrubs form production is related to ND53 index. Value of multiple correlation coefficients between shrub production and ND53 is 86%. So unfitted linear model hypothesis at regression model fitting rejected at 99% level of significance.

($P< 0.01$ $F=151.976, 86.885$)

Result of One-Sample K-S and also normal probability graph of standardized residuals emphasized on to be normal of errors distribution (Table5).

($P= 0.013$ and K-S $Z =1.583$)

By attention to results in table6, acceptance of 7 of 10 samples of control patches in model validation test is demonstrator of proposed model appropriate.

Table 5. Results of one- sample Kolmogorov-Smirnov test.

		Forb1
	N	54
Normal Parametersa,b	Mean	40.83433
	Std. Deviation	23.825426
Most Extreme Differences	Absolute	.184
	Positive	.109
	Negative	-.184
	Kolmogorov-Smirnov Z	1.350
	Asymp. Sig. (2-tailed)	.052

a. Test distribution is Normal.
b. Calculated from data.

Table 6. Proposed model validation test.

Ground kg/ truth pixel	Minimum kg/pixel	Maximum kg/pixel	result
0.242	0.117	0.231	reject
0.186	0.127	0.213	accept
0.524	0.481	0.632	accept
0.227	0.179	0.254	accept
0.191	0.175	0.232	accept
0.234	0.192	0.281	accept
0.99	0.111	0.158	reject
0.421	0.376	0.539	accept
0.589	0.501	0.623	accept
0.121	0.132	0.197	reject

Discussion

Results of this study showed that shrubs production at Chehelgazi watershed had the best correlation with two vegetation indices ND53 and TM5/TM3 rather than other bands of TM sensor. The spectral reflectance of shrubs is distinct compared to bare ground and senescent herbaceous cover in the imagery. Shrubs produce ephemeral leaves in early spring, which senesce during the growing season, and overwintering leaves later in the growing season, which senesce the following spring (Bilbrough and Richards 1993). So we used two TM image from 25 Jun and 28 August in order to better distinguish shrubs spectral reflectance from other plant forms and rainfed agriculture. Many pixels with low percent cover of shrubs were classified as having 0 percent and production. This pattern supports Okin *et al.*'s (2001) conclusion that spectral mixture analysis does not provide accurate estimates of vegetation cover, when vegetation cover within a pixel is less than 30 percent. Sepehri (2003) showed that because of the prevalence of spectral refection of the soil, estimation of plant cover (<40%) is difficult and there should be other data such as the type of the soil, color and leaf surface indexes to be included in the model for the estimation of plant cover less than 40%.

In this study, shrub cover of less than 20 percent considered as having 0 percent shrub cover and production. The production model based on multivariate linear regression resulted as statistically significant and explained 74.5% and 77.3% of the variability in shrub production for ND53 and total ND53 and TM5/TM3 as first and second predictor respectively. Many research used vegetation indices specially NDVI as a proxy of vegetation productivity instead of performing direct vegetation assessments (Santi *et al*, 2014). Arzani (1994, 1998) investigated on ability of some vegetation indices and had been proved the real ability indices that has been created based on middle infrared band. Farzad Mehr *et al.*, (2004) in a study performed in the similar region (Semirom county) estimated that the correlation between NDVI index and plant cover data was significant at P<0.5% error.

In this study, the relationships found between ND53 and TM/TM indices and shrub production and we developed a regression equation for predicting production of species common in a semi-arid region in Iran, mainly *Astragalus spp*. The regression models developed herein are suitable for predicting fuel biomass in similar shrub areas.

Analysis of dependent variable (shrubs form production), independent variables correlation matrix corresponding ground samples spectral values in original and artificial bands showed that in this area shrubs production has maximum correlation with to indices, ND53 and TM5/TM3. So we can resulted that production prediction of shrubs using these indices is possible in good accuracy. Similar results derived in previous researches (Farzadmehr *et al*, 2004, Arzani *et al*, 2009 and 1997).

Also by attention to livestock presence in the field simultaneous to sampling time and so decline of production and vegetation cover by grazing, we can result that vegetation indices show more sensitivity to little vegetation cover than other bands and fusions. Opposite to previous studies (Bing *et al*, 2003 and Ripple *et al*, 1991) mathematical transformation (logarithm, power, inverse and square root) on production variable and investigating correlations of new variables with independent variables are demonstrator of unchanged and in some cases decline of correlation coefficient. Based on proposed model and validation test results, Landsat TM images data in order to livestock grazing programing and determine fire risk or suppression has capability of production estimation of shrubbery forms in rangelands which these species are dominant and comparison with conventional methods needs to less time, cost and manpower.

Refrences

Arzani H, King G, Forster B. 1998. The application of the Landsat TM data to vegetation cover and yield measurement. Iranian Journal of Natural Resource. Vol. **1**, No. 50, pp. 3-21.

Arzani H, Noori S, Kaboli H, Moradi R, Ghelichnia H. 2009. Determination of suitable indices for vegetation cover assessment in summer rangelands in south of Mazandaran. Journal of the Iranian Natural Res., Vol. **61**, No. 4, pp. 997-1016.

Arzani H. 1994. Some aspects of estimating Short term and long-term rangeland carrying capacity. PhD Thesis. University of New South Wales. Australia.

Bilbrough CJ, Richards JH. 1993. Growth of sagebrush and bitterbrush following simulated winter browsing: mechanisms of tolerance. *Ecology* **74**, pp. 481-492.

Bing XU, Peng G, Ruliang P. 2003. Crown closure estimation of oak Savanah in a dry season with Landsat TM imagery: comparison of various indices through correlation analysis. International Journal of remote sensing. **24**(9). Pp.1811-1822.

Bradley BA. 2010. Assessing ecosystem threats from global and regional change: hierarchical modeling of risk to sagebrush ecosystems from climate change, landuse, and invasive species in Nevada, USA. *Ecography* **33**, 198-208. doi: 10.1111/j.1600-0587.2009.05684.x.

Bonham C D.1989. Measurements for Terrestrial Vegetation. New York, NY, USA: John Wiley and Sons. 352 p.

Lu D, Paul M, Eduardo B, Emilio M. 2002. Above-Ground Biomass Estimation og Successional and Mature Forests using TM Images in the Amazon Basin, 10th International Symposium on spatial Data Handling, pp. 183-196.

Duncan J, StowD, Franklin J, Hope A. 1993. Assessing the Relationship between spectral vegetation indices and shrub cover in the Yornada Basin, New Mexico. International Journal of Remote sensing, **14**, 3395-3416.

Farzadmehr DJ, Arzani H, Darvishsefat AA, Jafari M. 2004. Investigation in Estimating Vegetation Cover and Phytomass Production, Using Enhanced Landsat Data in A semi-arid Region, Journal of the Iranian Natural Res., Vol. **57**, No. 1, pp. 339-351.

Jin S, Sader SA. 2005. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. Remote sensing of environment. **94**, 364-372. DOI: 10.1016/j.rse.2004.10.012.

Lawrence R, Hurst R, Weaver T, Aspinall R. 2006. Mapping prairie pothole communities with multitemporal Ikonos satellite imagery, *Photogrammetric Engineering and Remote Sensing*. **72**(2), 169-174.

Jakubauskas ME, Legates DR, Kastens JH. 2002. Crop identification using harmonic analysis of time-series AVHRR NDVI data. Computers and Electronics in Agriculture. **37**, 127-139.

Okin GS, Roberts DA, Murray B, Okin WJ. 2001. Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment* 77, 212-225.

Ripple WJ, Wang S, Isaacson DL, Paine DP. 1991. A preliminary comparison of Landsat TM and Spot-1 HRV multispectral data for estimating coniferous forest volume, International Journal of remote sensing, **12**(9), 1971-1977.

Roy PS, Shirish AR. 1996. Biomass estimation using satellite remote sensing data—An investigation on possible approaches for natural forest, *Journal of Bioscience*. **21**(4), pp 535-561.

Santi E, Tarantino C, Amici V, Bacaro G, Blonda P, Borselli L, Rossi M, Tozzi S, Torri D. 2014. Fine-scale spatial distribution of biomass using satellite images. Journal of Ecology and the natural Environment. Vol. **6**(2), pp. 75-86. DOI: 10.5897/JENE2013.0416.

Seperhi A. 2003. Using Vegetation indices for estimate rangeland vegetation cover in Jahan nama refuge, Iranian Journal of Natural Resource. **55**(2), 20-31.

Underwood E, Ustin S, Dipietro D. 2003. Mapping nonnative plants using hyperspectral imagery. Remote Sensing of Environment. **86**, pp. 150–161.

Ren X, Wei X, Smith A. 2012. Remote Sensing, Crop Yield Estimation and Agricultural Vulnerability Assessment: a Case of Southern Alberta. *The Open Hydrology Journal,* 2012, ***6***, 68-77**.**

Zahedi S, Ghasriani F. 2014. Tolerence of Festuca ovina to different severities of simulated grazing effects in semi-arid Rangelands, Kurdistan, Iran. Journal of Biodiversity and Environmental Sciences (JBES). Vol.**4**. No.2, 369-377.