

Characterization and Estimation of Dates Palm Trees in an Urban Area Using GIS-Based Least-Squares Model and Minimum Noise Fraction Images

Muntadher Aidi Shareef^{1*}, Sumaya Falih Hasan¹

¹ Department of Surveying Engineering Techniques, Technical College of Kirkuk, Northern Technical University, Baghdad Road, 36001 Kirkuk, Iraq

* Corresponding author e-mail: muntadher.a.shareef@ntu.edu.iq

ABSTRACT

Date palm is the major food source and possesses an important role in the economic aspects, environmental parts, and society. These crops were subjected to degradation due to the financial and numerous military conflicts. Because of the expensive cost of monitoring and managing date palm in field measurements, and limited studies using satellite images, the authors proposed a method to estimate and map date palm using the Landsat-8 satellite images. The authors applied the least-squares multiple regression and GIS techniques to find suitable predictors from the set of variables such as original bands of Landsat-8, Minimum Noise Fraction (MNF) transformation, tasseled cap component transformation, and spectral index. In order to validate the proposed method, the field measurement data were utilized to assess the estimated date palm from the Landsat-8 images. A linear combination of MNF Landsat-8 band 4 (red, 0.636–0.673 μm), Normalized Difference Moisture Index (NDMI) and Enhanced Vegetation Index (EVI) were the best date palm predictor ($R^2_{\text{adj}} = 0.988$, root-mean-squared error (RMSE) = 0.013). The results demonstrate that the MNF Landsat-8 images in the least square regression help improve the date palm estimation and mapping for the practical use in the study area with high accuracy.

Keywords: dates palm, Landsat-8, minimum noise fraction, least-square model, GIS

INTRODUCTION

Date palm is one of the Arecaceae date group and a sign of existence in the desert (Khierallah et al. 2015). Date palm considered a native plant of the desert cultivated in severe local environments. The date is an ancient crop which has been grown in many regions and countries of the Middle East and North Africa for over 5000 years (Johnson 2010). Furthermore, Iraq could be considered the nation of the date palm, which can be found in different regions in Iraq (Chao and Krueger 2007). The number of date palm trees is to be around 22 million, covering more than 120 000 hectares (Khierallah et al. 2015). The date also represents a significant food source for the local population in the Middle East countries and North Africa. Furthermore, it has an essential role in the environment and society (Chao

and Krueger 2007). Date trees provide protection against high-speed wind as well as prevent soil desertification and degradation (Khierallah et al. 2015). Contrarily, in the year 2000, the number of trees in Iraq gradually reduced to 12 million, which cover about 50 000 hectares. The reduction in this important crop tree counts was due to the fighting activities that occurred in date palm regions since 1980 (Al-Khayri et al. 2015).

Additionally, planting of the date palm is affected by the biotic and abiotic factors (Al Shidi et al. 2018). Moreover, the date palm diseases play an essential part in degrading the palms by increasing the salinity of the soil, extreme water table, borers spreading, and aging of palm (Chao and Krueger 2007, El-Juhany 2010). The Arecaceae family, includes the well-known species such as date palm, coconut (scientific name is *Cocos nucifera*), and oil palm (or scientifically

named as *Elaeisguineensis*), which have a high profile compared others species (Sayan 2001). The remote sensing techniques applied widely for studying various types of vegetation with limiting of date palm (Alhammadi and Glenn 2008). The majority of optical and microwave remote sensing data focused on the plant, above-ground biomass, and palm oil rather than the date palm (Chong et al. 2017). The applications of date palm covered the surface analysis, mapping, and palm boundary extraction, counting of the tree, monitoring the change to distinguishing the variations in real resources, palm age and blight estimation, infection and disease detection, as well as carbon estimation (Santoso et al. 2016, Lazecky et al. 2018). Tan et al. (Li et al. 2016) studied the oil palm trees age. It is proven that the texture features and a shadow fraction are helpful in estimating the oil palm age. The estimation and evaluation of oil palm plantation and aboveground biomass were examined using multispectral Landsat images and microwave data (Morel et al. 2012).

Moreover, the machine learning-based methods were applied for the palm tree detection by using a scale-invariant feature transform (SIFT, and unmanned aerial vehicle (UAV) images (Malek et al. 2014) The circular autocorrelation of the polar shape matrix and a linear support vector machine were used to reduce the feature dimensions for detecting date palm (Manandhar et al. 2016). As follows, remote sensing is a valuable tool for implementing early discovery and constant observation of many difficulties (Chong et al. 2017). Several studies in remote sensing concentrated on the environmental factors and the multispectral properties or radar properties of the elements to extract multiple parameters through satellite images or data (Shi et al. 1997, Shareef 2019).

Various models were applied to study different parameters of the Earth's surface. The empirical models are considered the most straightforward and directly applied models; hence, they have been implemented to analyze uncertain determinants or to investigate the relationships between different features (Shareef et al. 2014a). Nevertheless, the remote sensing application in numerous kinds of the palm is confined in limited species and not focused directly on the date palm trees (Hansen et al. 2015). The objective of this study was to (1) produce a different approach that allows identifying the trees of date palm in urban employing the Landsat -8 images and a statistical

model, (2) assess the use of tasseled cap and spectral indices to extract the date palm.

DATA AND STUDY REGION

The Landsat 8 satellite officially started service in February 2013, being the eighth satellite of the Landsat group (Oishi et al. 2018). The Landsat-8 data herein were obtained on 09/11/2017 as well as pre-processed radiometric-ally and atmospherically using the ENVI5 software. The study was performed in the Baghdad region, the capital of Iraq, which is extended to the Al Dora city situated in the south of the area and to the Arabs Ejbura along with a distance of 12 km. Two hundred sixty-one samples of palm trees were obtained in October/2017 and separated within 211 samples as training data and 50 samples as the reference data. The references data employed to build and verify the producing models. The area of study can be shown in Figure 1.

Spectral models

Numerous spectral models applied to estimate or analyze the surface properties. Normalized Difference Vegetation Index (NDVI) is a commonly used model to detect vegetation (Lyk-hovyd 2020), and it was produced by Rouse in 1973. The NDVI depends on concentrating on the near-infrared and red band, which are more sensitive to the plant (Rouse Jr et al. 1974). On the other hand, the Enhanced Vegetation Index (EVI) was employed to enhance the response of the biomass region and to obtain the best signal of the crops or vegetation. It is supported to improve (1) the signal transmitted from the canopy, and (2) lessen the atmospheric influences (Liao et al. 2015). Moreover, the Normalized Difference Moisture Index (NDMI) was employed to obtain moisture using near-infrared (NIR) and shortwave infrared (SWIR), which represent the fifth and sixth bands in the Landsat-8 data. (Hardisky et al. 1983)

Figure 2 shows the general process followed in this paper is to assess and determine date palm by using the MNF Landsat-8 data. This methodology included the stages of passing through correction, transformation, producing the models, and mapping the date palm. The statistical regression analysis between the extracted parameters and fieldwork was conducted to create the date palm models

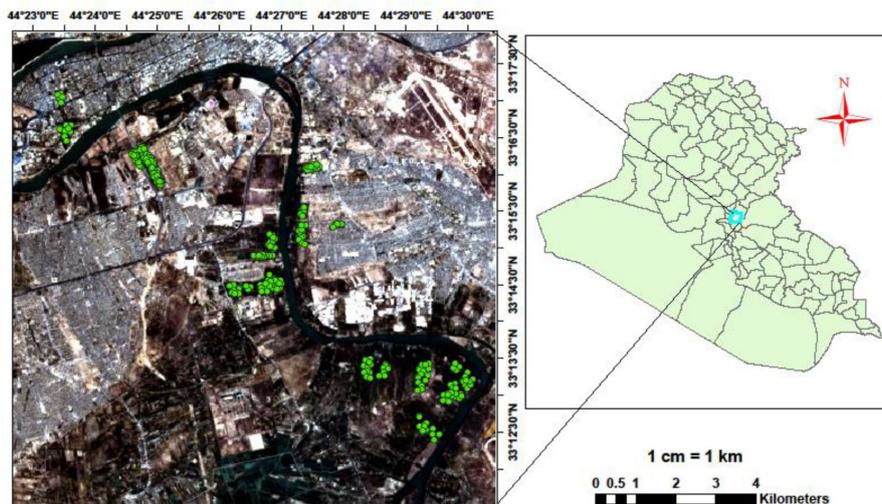


Fig. 1. Study region and data collection

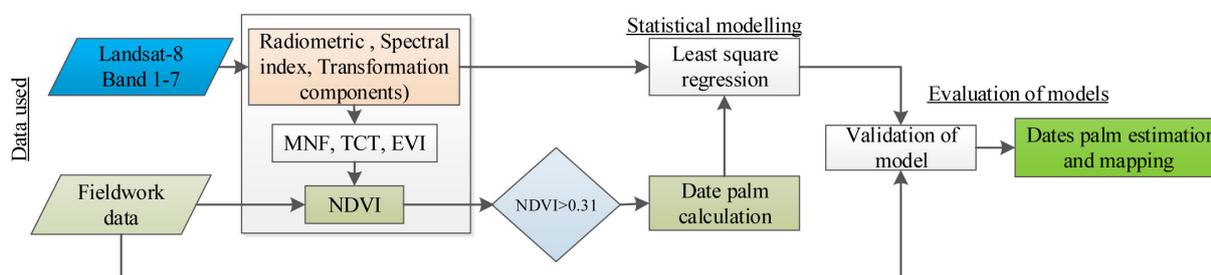


Fig. 2. The procedure to estimate date palms

Least square method

The least squares method is a statistical method that aims to estimate a regression line that reduces the sum of the primary deviations or errors in the observed points in the regression line. It minimizes the sum of squares of differences between actual and calculated values. The general model of this method is given in (Eq.1):

$$y = X\beta + \varepsilon \quad (1)$$

where: y is a vector of response X is a variables matrix,

β are the entered vector parameters, and ε is the error vector (Shareef et al. 2014b).

Transformation of data

Atmospheric correction was applied to the Landsat-8 data. The thermal infrared images were not used in this study. In order to determine the inherent dimensionality of the pre-processed image data acquired by Landsat-8, a Minimum Noise Fraction (MNF) transformation applied to pure images of the Landsat-8. MNF was performed for

isolating Landsat-8 noise and for diminishing the calculation necessities for succeeding the process and for reducing the multicollinearity among various features (Joseph 1994).

RESULT AND DISCUSSION

The Landsat-8 data used in this study includes the information concerning the reflectance (spectra) of different features in the region of interest. The MNF transformation was used herein to reduce the image noise in initial data and the produced data were applied to extract Tasselled. The spectral models have well presented the plants at a high brightness level, which herein represents the date palm trees. The wetness component shows the wetland zones by the high brightness, but date palm was mixed land covers in this component. Vegetation and water area reflected the interference or mixed pixels using the intensity and fourth components of TCT. Nevertheless, the fifth and sixth components present the plants in black color, including the date palm trees (Fig 3).

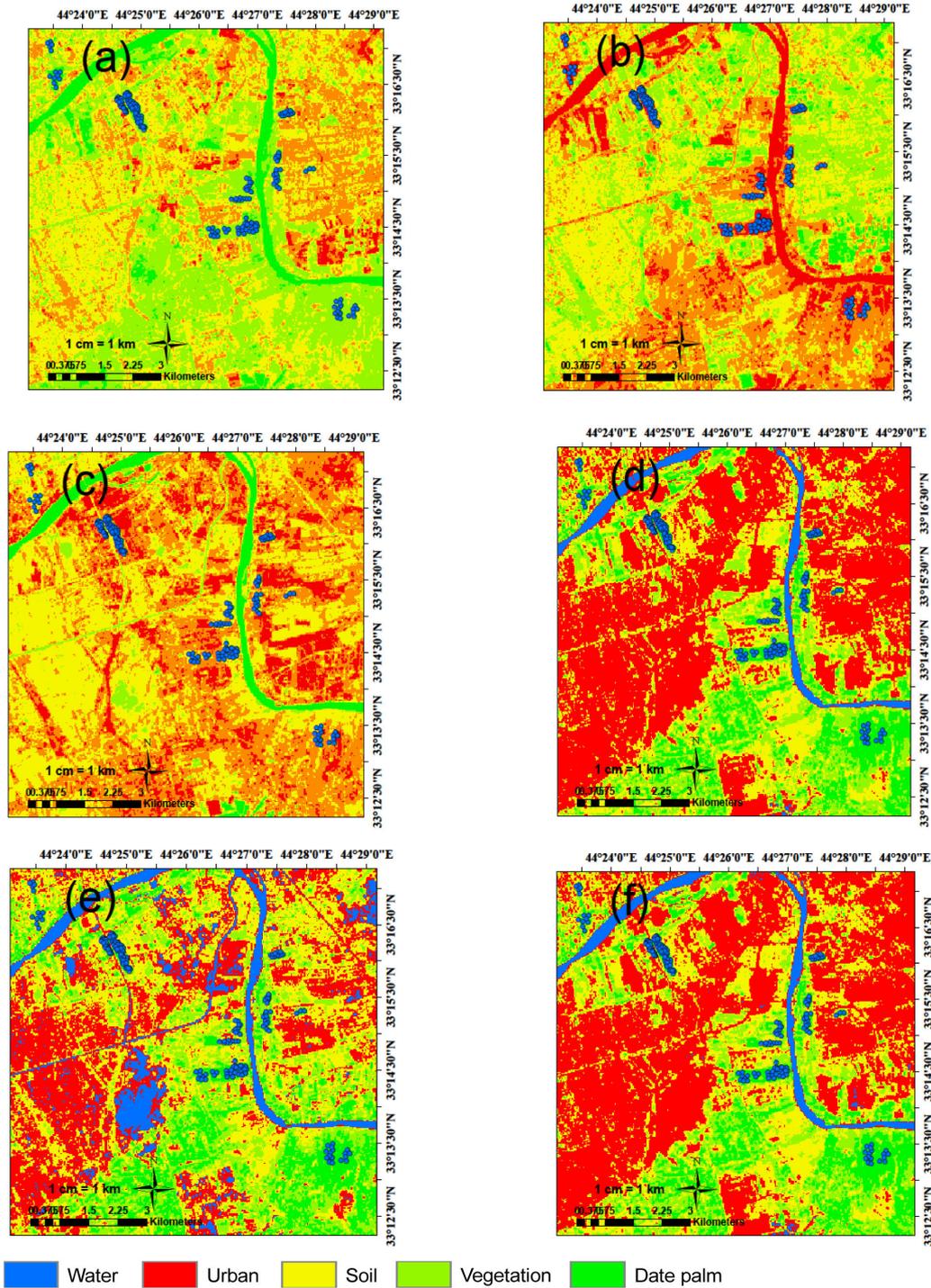


Fig. 3. (a, b, c) TCT components (greenness, brightness, wetness) and (d, e, f) spectral indices (NDVI,NDMI,EVI)

The autocorrelation analysis was used to find a correlation between our databases, a p-value of statistical significance was under 0.05, which indicated a tendency towards dispersion. The importance of F is to designate the goodness of the predicted model. Table 1 and Table 2 illustrate the correlation values of the date palm derived by NDVI correlated with various bands

of the Landsat-8, indices, and TCT components. The date palm effectively related to the 4th component of TCT (0.976), while EVI is low correlated (0.190) with date palm. On the other hand, the correlation value of the multispectral MNF bands and date palm reflects the strong correlation of 0.94 in band 7, which correspond to the SWIR band. The 5th and 6th components have

Table 1. Correlation of different indices, and components of TCT with NDVI

Spectral indices		Tasseled Cap components					
NDMI	EVI	Wetness component	Brightness component	Greenness component	Fourth component	Fifth component	Sixth component
0.805	0.190	0.468	-0.880	0.785	0.976	-0.933	-0.974

Table 2. Correlation of the MNF Landsat-8 data with NDVI

Multispectral MNF bands of Landsat-8						
Band1	Band2	Band3	Band4	Band5	Band6	Band7
-0.864	-0.880	-0.887	-0.915	-0.362	-0.888	-0.935

inverse correlation with the date palm, which was -0.93 and -0.97, respectively. Brightness has a higher value of correlation with date palm than Greenness.

In this study, three analyses were applied to produce the date palm model. The first analysis included a single variable correlated with date palm (Table 3), while the second analysis included two variables to produce the model (Table 4). The third analysis using the least square regression model was to three entered variables relative to date palm trees, as illustrated in Table 5.

The minimum correlation value using MNF data was noticed using B5. The lowest R^2_{adj} value

(0.216) noticed in the Wetness component, and on the contrary, it had the highest R^2_{adj} value (0.952) in Fourth. Moreover, the spectral NDMI is a low correlation with data palm, while EVI had a moderate value of R^2_{adj} (0.032 and 0.647), respectively. Next, two variables computed as independent variables were examined to produce the models using multi-regression least square. Table 5 shows the achieved results. The combination of spectral indices (NDMI and EVI) slightly increased the R^2_{adj} value (0.649) as opposed to using the same model with just one spectral index, an independent variable. Conversely, the combination of tasseled cap components of wetness and

Table 3. Evaluation of least square parameters required for producing models (single variable correlated relative to date palm)

	Variables	a	b	R	R^2_{adj}	R^2_{adj}	RMSE
Spectral indices	NDMI	0.60	0.99	0.81	0.65	0.65	0.07
	EVI	0.22	0.3	0.19	0.04	0.03	0.12
Tasseled cap components	Wetness component	0.78	2.04	0.47	0.22	0.22	0.11
	Brightness component	0.67	-1.20	0.88	0.78	0.77	0.06
	Greenness component	0.76	6.03	0.79	0.62	0.62	0.08
	Fourth component	0.30	4.15	0.98	0.95	0.95	0.03
	Fifth component	0.02	-15.77	0.93	0.87	0.87	0.04
	Sixth component	0.47	-15.37	0.97	0.95	0.95	0.03
	B7	0.60	-1.92	0.94	0.88	0.88	0.04

Table 4. Least square parameters required for producing models (two variables correlated relative to the date palm)

Independent variables	Variables	a	b	c	R^2_{adj}	R^2_{adj}	RMSE
Spectral indices	NDMI, EVI	0.579	0.977	0.095	0.652	0.649	0.072
Tasseled cap components	Wetness, Brightness	0.307	-2.113	-1.704	0.874	0.873	0.043
Spectral bands	B4, B5	0.157	-2.799	2.099	0.967	0.967	0.022
	B5,B6	0.244	1.662	-2.429	0.966	0.966	0.022

a, b are coefficients of least-square; R^2 : coefficient of determination; and R^2_{adj} : adjusted determination coefficient Table 3 illustrates main parameters that have the best value of R^2_{adj} when the MNF bands (B4 and B7) used as independent variables in the least-square model, R^2_{adj} the value was increased in the case of 0.84 to 0.88, respectively. Otherwise, the bands of Landsat-8, ranging from B1 to B3, had a moderate

Table 5. Least square parameters required for producing models (three variables correlated relative to the date palm)

Independent variables	a	b	c	d	R ²	R ² _{adj}
NDMI,EVI,B4	0.702	0.559	-0.030	-1.527	0.988	0.988
B3,B4,B5	0.147	3.060	-5.063	1.906	0.978	0.978
B4,B5,B6	0.203	-1.563	2.375	-1.455	0.978	0.977

brightness are generally increased the R^2_{adj} value (0.873) compared with those who have a single part. Thus, using the combination of the MNF Landsat-8 images improved the predicted models with high R^2_{adj} . The final step included three independent variables (MNF Landsat-8 images, spectral indices, and TCT component) used in the regression model (Table 5). This step classified combination into six groups according to the used parameters. The first group includes the combination of spectral indices and tasseled cap. In this group, R^2_{adj} is ranging between 0.903 (NDNI, EVI, and TC-Greenness) and 0.983 (NDNI, EVI, and TC-Sixth, which is considered the best in this group). On the other hand, MNF Landsat-8 has a high R^2_{adj} (0.988), mainly when the NDNI, EVI, and B4 indices are used. The third combination employs spectral indices, tasseled cap components, and MNF spectral bands; thus, the best R^2_{adj} for this combination is 0.97. The only mixture has obtained in this optimization was the combination of a spectral index, spectral bands, and tasseled cap, which is represented by EVI, TC-W, and B7. The combination of tasseled cap components has a good range between 0.928 (using TC-B, TC-G, and TC-FI) and 0.978 (TC-W, TC-B, and TC-S). The combination of the tasseled cap component with MNF Landsat-8 slightly improved, compared to the application of single tasseled

cap components images. The next combination, which represents the final group, uses the MNF Landsat-8 images to generate a date palm trees model. The best indication of the producing date palm was obtained using two variables, such in the following:

$$DP_s = 0.244 + 1.66B_5 - 2.429B_6 \quad (2)$$

where: DP refers to the date palm; s refers to the spectral band based MNF;

B_5, B_6 are the fifth and sixth band of Landsat-8.

Because the correlation between two spectral indices and the bands for of MNF Landsat-8 is the highest correlation of the produced model to estimate the amount of date palm given as:

$$DP_{ns} = 0.702 + 0.559 NDMI - 0.030 EVI - 1.527 B_4 \quad (3)$$

where: ns refers to the combination of indices and MNF band.

The validation method was applied on the produced models to measure their goodness. This method was used the polynomial linear fitting to estimate the degree of the confidence of the models between the estimated and field measurement values, as shown in Figure 4. The validation of

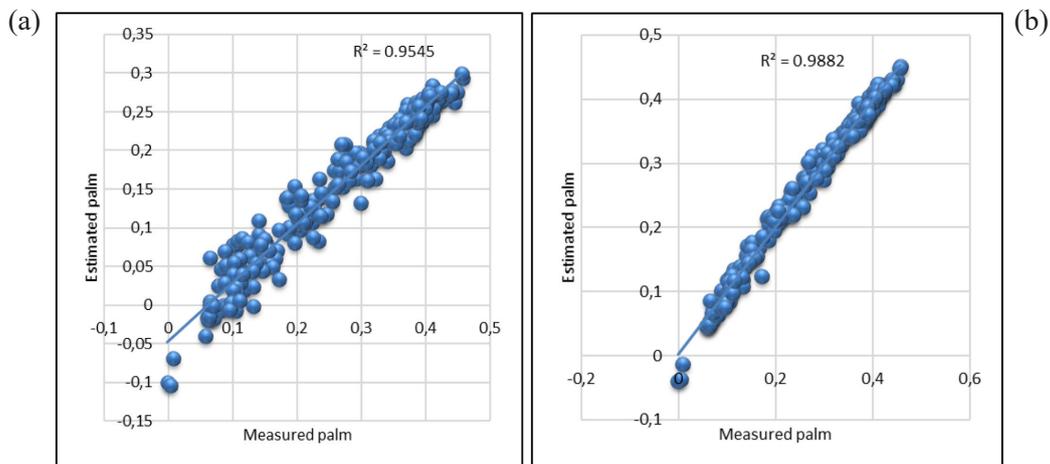


Fig. 4. Estimated (a) and calculated (b) date palm for two models

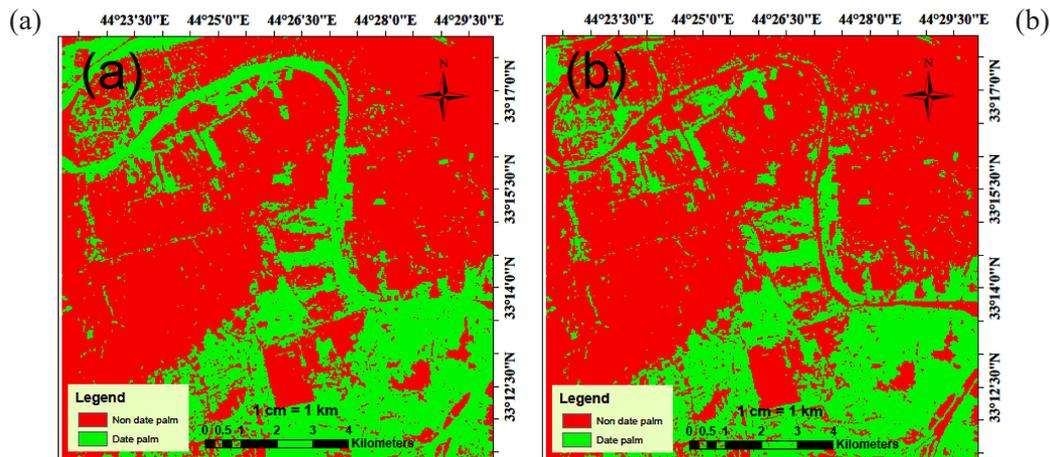


Fig. 5. Suitability maps of date palm based on DP_s (a) and DP_{ns} (b) models

the models gives legitimacy to use them within a specified accuracy. The produced accuracy was 0.95 and 0.988 for the first (DP_s) and the second model (DP_{ns}), respectively. The producing model with entering data enables to map the date palm in the spatial pattern. Figure 5 is established depending on the least square method. The produced models rely on the parameters extracted from the original Landsat-8 image. The final image provided by the models was a single image with a high brightness of the date palm. The maps were classified into two classes representing the non-date palm and date palm class. Thus, the representation of the date palm depends on the accuracy of the model used.

CONCLUSIONS

This study used different types of data extracted from the Landsat-8 images to estimate and map date palm using the least squares models in a GIS system. The relationship between the user parameter was examined to comprehend the influence of each applied variable relative to the other parameters and date palm.

Some parameters exhibited a low correlation with the vegetation index that it is necessary to generate the models to estimate date palm. Other parameters present high correlation values with the vegetation index needed to develop the estimation model. Overall, the obtained results demonstrate a strong effect of TCT, spectral index and MNF Landsat-8 of estimating and studying the date of palm trees with a useful finding pertaining to the use of single and multiple variables. Despite the many experiments in

this study, we have placed the strongest relationships between the variables that can identify the date palm with high accuracy. The produced result enables us to generate different types of estimating models depending on the data used and the accuracy required. Moreover, the distribution of estimated date palm trees in terms of the spatial map may help to guide the separations of the density of the different types of vegetation to make the decisions in the future to protect and exploit the date palm trees.

Acknowledgment

We thank the ministry of higher education for their support to achieve and complete this work.

REFERENCES

1. Al-Khayri J.M., S.M. Jain, and D.V. Johnson. 2015. Date palm genetic resources and utilization. Africa and the Americas 1.
2. Al Shidi R.H., L. Kumar, S.A. Al-Khatiri, M.M. Al-bahri, and M.S. Alaufi. 2018. Relationship of Date Palm Tree Density to Dubas Bug *Ommatissus lybicus* Infestation in Omani Orchards. Agriculture 8, 64.
3. Alhammadi M. and E. Glenn. 2008. Detecting date palm trees health and vegetation greenness change on the eastern coast of the United Arab Emirates using SAVI. International Journal of Remote Sensing 29, 1745–1765.
4. Chao C.T. and R.R. Krueger. 2007. The date palm (*Phoenix dactylifera* L.): overview of biology, uses, and cultivation. HortScience 42, 1077–1082.
5. Chong K.L., K.D. Kanniah, C. Pohl, and K.P. Tan. 2017. A review of remote sensing applications for

- oil palm studies. *Geo-spatial Information Science* 20, 184–200.
6. El-Juhany L.I. 2010. Degradation of date palm trees and date production in Arab countries: causes and potential rehabilitation. *Australian Journal of Basic and Applied Sciences* 4, 3998–4010.
 7. Hansen S.B., R. Padfield, K. Syayuti, S. Evers, Z. Zakariah, and S. Mastura. 2015. Trends in global palm oil sustainability research. *Journal of Cleaner Production* 100, 140–149.
 8. Hardisky M., V. Klemas, and M. Smart. 1983. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of *Spartina alterniflora* 49, 77–83.
 9. Johnson D. 2010. Worldwide dispersal of the date palm from its homeland. Pages 369–375 in *IV International Date Palm Conference* 882.
 10. Joseph W. 1994. Automated spectral analysis: A geologic example using AVIRIS data, north Grapevine Mountains, Nevada. In: *Proc. Tenth Thematic Conference on Geologic Remote Sensing*, Environmental Research Institute of Michigan, 1407–1418.
 11. Khierallah H.S., S.M. Bader, K.M. Ibrahim and I.J. Al-Jboory. 2015. Date palm status and perspective in Iraq. Pages 97–152 *Date palm genetic resources and utilization*. Springer.
 12. Lazecky M., S. Lhota, T. Penaz, and D. Klushina. 2018. Application of Sentinel-1 satellite to identify oil palm plantations in Balikpapan Bay. Page 012064 in *IOP Conference Series: Earth and Environmental Science*. IOP Publishing.
 13. Li W., H. Fu, L. Yu, and A. Cracknell. 2016. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sensing* 9, 22.
 14. Liao Z., B. He, and X. Quan. 2015. Modified enhanced vegetation index for reducing topographic effects. *Journal of Applied Remote Sensing* 9, 096068.
 15. Lykhovyd P.V. 2020. Sweet corn yield simulation using normalized difference vegetation index and leaf area index. *Journal of Ecological Engineering*, 21(3), 228–236.
 16. Malek S., Y. Bazi, N. Alajlan, H. AlHichri, and F. Melgani. 2014. Efficient framework for palm tree detection in UAV images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, 4692–4703.
 17. Manandhar A., L. Hoegner, and U. Stilla. 2016. Palm Tree Detection Using Circular Autocorrelation of Polar Shape Matrix. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 3, 465.
 18. Morel A.C., J.B. Fisher, and Y. Malhi. 2012. Evaluating the potential to monitor aboveground biomass in forest and oil palm in Sabah, Malaysia, for 2000–2008 with Landsat ETM+ and ALOS-PALSAR. *International Journal of Remote Sensing* 33, 3614–3639.
 19. Oishi Y., H. Ishida, and R. Nakamura. 2018. A new Landsat 8 cloud discrimination algorithm using thresholding tests. *International Journal of Remote Sensing* 39, 9113–9133.
 20. Rouse Jr, J. W., R. Haas, J. Schell, and D. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS.
 21. Santoso, H., H. Tani, and X. Wang. 2016. A simple method for detection and counting of oil palm trees using high-resolution multispectral satellite imagery. *International Journal of Remote Sensing* 37, 5122–5134.
 22. Sayan M.S. 2001. Landscaping with palms in the Mediterranean. *Palms-Lawrence* 45, 171–176.
 23. Shareef, H., Hasan, Noori. 2019. Integrating of GIS and fuzzy multi-criteria method to evaluate land degradation and their impact on the urban growth of Kirkuk city, Iraq. *International Journal of Advanced Science and Technology* 28, 800–815.
 24. Shareef M.A., A. Toumi, and A. Khenchaf. 2014a. Estimation of water quality parameters using the regression model with fuzzy k-means clustering. *international journal of advanced computer science and applications* 5, xx.
 25. Shareef M.A., A. Toumi, and A. Khenchaf. 2014b. Prediction of water quality parameters from SAR images by using multivariate and texture analysis models. Page 924319 in *SAR Image Analysis, Modeling, and Techniques XIV*. International Society for Optics and Photonics.
 26. Shi J., J. Wang, A.Y. Hsu, P.E. O’Neill, and E.T. Engman. 1997. Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data. *IEEE Transactions on Geoscience and Remote Sensing* 35, 1254–1266.