

Estimation of daily average downward shortwave radiation from MODIS data using principal components regression method: Fars province case study

Razieh Barzin, Amin Shirvani*, and Hossein Lotfi

Department of Water Engineering, Oceanic and Atmospheric Research Center, College of Agriculture, Shiraz University, Shiraz, Iran

Received February 16, 2016; accepted November 29, 2016

A b s t r a c t. Downward shortwave radiation is a key quantity in the land-atmosphere interaction. Since the moderate resolution imaging spectroradiometer data has a coarse temporal resolution, which is not suitable for estimating daily average radiation, many efforts have been undertaken to estimate instantaneous solar radiation using moderate resolution imaging spectroradiometer data. In this study, the principal components analysis technique was applied to capture the information of moderate resolution imaging spectroradiometer bands, extraterrestrial radiation, aerosol optical depth, and atmospheric water vapour. A regression model based on the principal components was used to estimate daily average shortwave radiation for ten synoptic stations in the Fars province, Iran, for the period 2009-2012. The Durbin-Watson statistic and autocorrelation function of the residuals of the fitted principal components regression model indicated that the residuals were serially independent. The results indicated that the fitted principal components regression models accounted for about 86-96% of total variance of the observed shortwave radiation values and the root mean square error was about 0.9-2.04 MJ m⁻² d⁻¹. Also, the results indicated that the model accuracy decreased as the aerosol optical depth increased and extraterrestrial radiation was the most important predictor variable among all.

K e y w o r d s: daily average solar radiation, moderate resolution imaging spectroradiometer, principal components regression model, estimation, Fars province

INTRODUCTION

The information of the solar radiation received at the earth surface and global solar radiation distribution have extensive applications in the architecture, agriculture and climate investigations, and finding an optimal place for construction of renewable energy sites. Performance of solar energy systems requires knowledge of the spatial dis-

tribution of solar radiation at the earth surface and a reliable methodology to estimate global solar radiation based on the least available data, which is a challenge in renewable energy planning (Kaplanis and Kaplani, 2007). Monitoring the earth radiation budget (ERB) is essential for improving our understanding of the earth climate and potential climatic changes (Change, 2001; Houghton *et al.*, 2001). Furthermore, solar radiation can be a powerful player in local weather and climate. Point measurements of solar radiation are not sufficient for spatial land surface models. In addition, weather stations are generally located in flat areas so that the observations of daily average downward solar radiation (DADSR) tend to represent limited surrounding areas, which may exclude precipitous land surfaces or mountainous areas (Irmak and Allen, 2005; Samani *et al.*, 2007; Su *et al.*, 2005). The current solar radiation products have fine temporal and coarse spatial resolutions that are not suitable for evaluating land-atmospheric interaction and local weather models (Kim and Liang, 2010). Due to the requirement for expensive equipment and the difficulty in sensor calibration, it is impossible to obtain solar radiation from ground-based measurement at the high resolution spatial scale (Phakamas *et al.*, 2013). Thus, several studies have attempted to estimate solar radiation (both shortwave and longwave) components using remote sensing (RS) data based on radiative transfer models which need many ancillary data (such as CO₂, O₃, CH₄, temperature, cloud optical thickness *etc.*). Mueller *et al.* (2004) and Huang *et al.* (2012) estimated the daily average net surface shortwave radiation (R_{ns}) at 1 Km resolution accurately using a simplified radiation transfer algorithm. Kim and Liang

*Corresponding author e-mail: am_shirvani@hotmail.com,
ashirvani@shirazu.ac.ir

(2010) developed a hybrid method for estimating land surface shortwave net radiation from MODIS data. They used instantaneous net shortwave radiation data as an input for calculating daily average net shortwave radiation. Their findings showed that root mean square errors (RMSE) varied about 2.8-4.0 MJ m⁻² d⁻¹ (Kim and Liang, 2010). Qin *et al.* (2011) used MODIS and monthly products of tropical rainfall measuring mission (TRMM) and the artificial neural network (ANN) method to build a mathematical relationship between monthly mean daily global solar radiation (GSR) and satellite data. They reported that the ANN method can retrieve GSR with an acceptable accuracy. In addition, their results showed that Yang *et al.* (2001) method performed better than both the ANN-based method and Japanese aerospace exploration agency (JAXA) algorithm. Rahimikhoob considered both the Hargreaves and Samani equation and the ANN method to estimate global solar radiation in Ahvaz-Iran, using air temperature (T_{\min} , T_{\max}) and R_a as input parameters (Rahimikhoob, 2010) and showed that the ANN outperformed the Hargreaves and Samani equation in global solar radiation. Theoretical methods for estimating solar radiation can be expected to give higher accuracies than empirical methods, as the radiation attenuation process is considered comprehensively in the theoretical method (Huang *et al.*, 2012). On the other hand, many solar radiation based on remote sensing studies estimated daily average component using adjusted sinusoidal interpolation (Bisht *et al.*, 2005; Kim and Liang, 2010). Other studies have attempted to retrieve net surface shortwave radiation from MODIS data using only the solar zenith angle and the water vapour amount as the input parameters and mentioned that the MODIS sensor may not be good enough to provide daily average shortwave radiation since Terra and Aqua overpass the same site only two times each day (Tang *et al.*, 2006).

López and Batlles presented a new accurate parametric model to estimate global solar radiation under cloudless conditions (López and Batlles, 2014). Wang *et al.* (2015) reported that the temporal mean albedo is more important than the instantaneous albedo for calculation of the surface shortwave net radiation budget. Zhang *et al.* (2015) demonstrated that the radiation components exhibit very high spatial heterogeneity and are largely affected by different terrain orientations, slope, the cosine of the solar illumination angle, and atmospheric environments. Overall, the MODIS sensors on board the Terra and Aqua satellite that yield regular and frequent observation of biophysical variables of the entire globe have fine spatial and spectral resolution for studying energy balance over large areas. However, all of the polar orbit sensors have coarse temporal resolution.

The main objective in the present study is to apply the combination of principal components analysis (PCA) and regression models for estimating DADSR using MODIS data, which has not been reported in previous studies. In

this manner, the regression model based on the principal components is developed to estimate shortwave solar radiation over the Fars province in Iran using the information of spectral bands, aerosol optical depth (AOD), atmospheric water vapour, and extraterrestrial radiation (R_a).

MATERIAL AND METHODS

The satellite data were extracted from moderate resolution imaging spectroradiometer (MODIS) website (www.modis.gsfc.nasa.gov). In this study, the following MODIS data were used:

- the MOD02HKM Level1B data products provide an estimation of TOA radiance and reflectance, which include the 250 and 500 m spatial resolution bands. The first seven spectral bands cover from 620 to 2155 nm. The MOD02HKM product is calibrated Earth view data at 500 m resolution by the MODIS Characterization and Support Team (MCST) (Qu *et al.*, 2006).
- MOD04 (Level2) consists of aerosol optical depth (AOD). The standard MODIS aerosol land algorithm relies on finding dark targets; for this reason, it is unable to make retrievals over bright surfaces such as deserts (Levy *et al.*, 2013). Since the Fars province has been classified as an arid/semi-arid region, the AOD was not made with the previous algorithm over this region. The AOD product includes the ‘deep-blue’ algorithm recently developed (collection 6.0) to get AOD over bright land areas. This new AOD product was used to investigate the effect of AOD on DADSR variations.
- the near infrared algorithm of the MODIS total perceptible water product (MOD05_L2) consists of water-vapour column amounts (WV) over the clear area of the globe. The most variable gas that significantly affects downward solar radiation in the atmosphere is water vapour. Even daily fluctuation of water vapour from 0.4 to 4.12 g cm⁻² has been reported (Justice *et al.*, 1991). Therefore, daily transmittance of the atmosphere related to water vapour is considered in our study.

The extraterrestrial radiation (R_a) for each day of the year and for different latitudes can be estimated from the solar constant, the solar declination and the time of the year by the Angstrom equation (FAO, NO56) (Adeboye *et al.*, 2009):

$$R_a = \frac{24 \times 60}{\pi} G_{sc} d_r (\omega_s \sin \Phi \sin \delta + \cos \Phi \cos \delta \sin \omega_s), \quad (1)$$

where: R_a – extraterrestrial radiation (MJ m⁻² d⁻¹), G_{sc} – solar constant = 0.082 MJ m⁻² min⁻¹ = 1367 W m⁻²,

$$d_r = 1 + 0.033 \cos \left(\frac{2\pi}{365} J \right), \text{ inverse relative distance Earth-Sun,} \quad (2)$$

$$\omega_s = \arccos(-\tan \varphi \tan \delta) \text{ sunset hour angle (rad),} \quad (3)$$

φ = latitude (rad),

$$\delta = 0.409 \sin (J - 1.39), \text{ solar declination (rad),} \quad (4)$$

and J is Julian day.

The proposed methodology of estimating DADSR from MODIS data was applied over the Fars province, located in the south western of Iran, (approximately between 27.00°N and 32.00°N latitudes and between 51.50 and 55.50°E longitudes) with the elevation of 300 to 3500 m. The climate of the province is mainly arid and semi-arid with mean annual temperature from 10 to 24°C in the northern and southern parts of the province, respectively. Downward shortwave radiation for the period of 2009 to 2012 under clear sky conditions was collected for 10 synoptic stations from the Fars Meteorological Bureau. Ten-minute solar radiation values recorded by the data loggers were converted to daily average values.

The spatial distribution and geographical characteristics of the synoptic stations surveyed in this study are illustrated in Fig. 1 and Table 1.

The principal components analysis as an applied multivariate statistical technique is extensively used in the atmospheric sciences (Wilks, 2011) to reduce a data set containing a large number of variables to a data set containing fewer new variables. These new variables are linear combinations of the original ones, and these linear combinations are chosen to represent the maximum possible fraction of the variability contained in the original data. Data on multispectral bands (and atmospheric fields) are often highly correlated because of material spectral correlation and topography, and therefore analysis of all the

Table 1. Geographical characteristic of the stations used in this study

| No. | Station | Latitude N | Longitude E | Height (m a.s.l) |
|-----|------------|------------|-------------|------------------|
| 1. | Eghlid | 30.869 | 52.676 | 2300 |
| 2. | Bajgah | 29.726 | 52.585 | 1796 |
| 3. | Kazeroon | 29.607 | 51.665 | 838 |
| 4. | Shiraz | 29.544 | 52.603 | 1494 |
| 5. | Neyriz | 29.187 | 54.352 | 1627 |
| 6. | Fasa | 28.899 | 53.720 | 1310 |
| 7. | Firoozabad | 28.886 | 52.552 | 1363 |
| 8. | Darab | 28.790 | 54.296 | 1095 |
| 9. | Jahrom | 28.483 | 53.517 | 1080 |
| 10. | Lamerd | 27.364 | 53.203 | 403 |

original spectral bands is inefficient (Schowengerdt, 2006). The principal components (PC) is a linear transform with an image-specific matrix W_{PC} over an image spectral vector (DN):

$$PC = W_{PC} DN. \quad (5)$$

This transformation alters the covariance matrix as follows:

$$C_{PC} = W_{PC} C W_{PC}^T. \quad (6)$$

The PCA is optimum in the sense that, of all possible transformations, W_{PC} is the only one that diagonalizes the covariance matrix of the original multispectral image, so that:

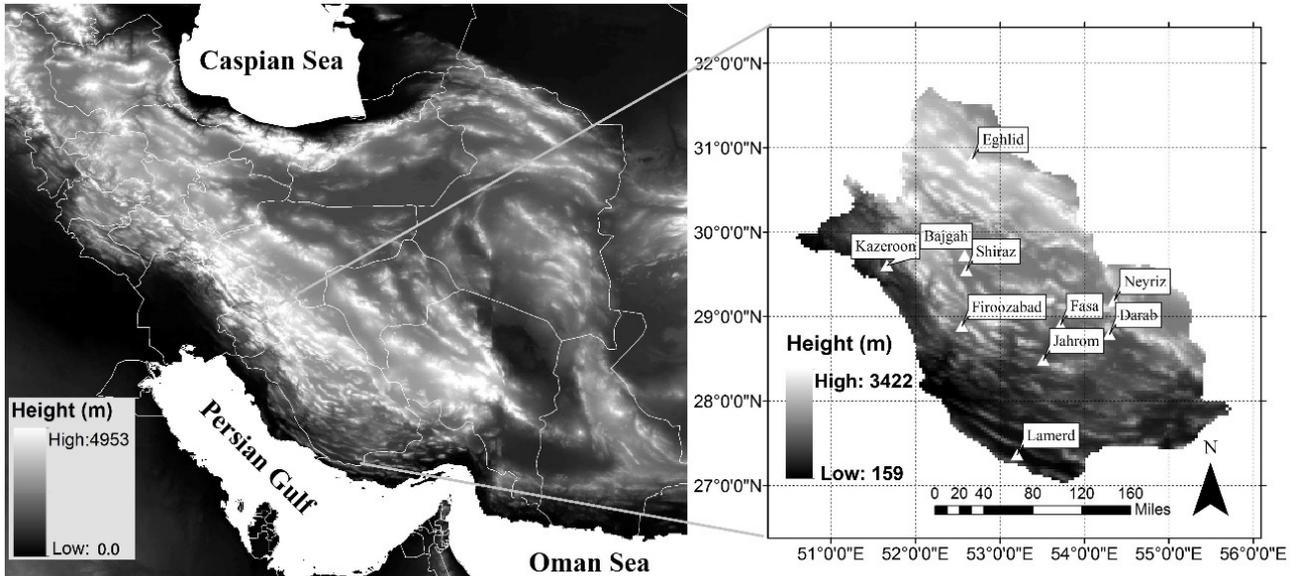


Fig. 1. Spatial distribution of the synoptic stations used in this study.

$$C_{PC} = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \lambda_K \end{bmatrix}. \quad (7)$$

The K – eigenvalues λ_K are found as the K roots of the characteristic equation:

$$|C - \lambda I| = 0, \quad (8)$$

where: C is the original data covariance matrix and I is the (diagonal) identity matrix. The PC coordinate axes are defined by the K eigenvectors, e_k , obtained from the vector-matrix equation for each eigenvalue λ_k :

$$(C - \lambda_k I) e_k = 0, \quad (9)$$

which form the rows of the transformation matrix W_{PC} ,

$$W_{PC} = \begin{bmatrix} e_1^t \\ \vdots \\ e_K^t \end{bmatrix} = \begin{bmatrix} e_{11} & \dots & e_{1K} \\ \vdots & e_{ij} & \vdots \\ e_{K1} & \dots & e_{KK} \end{bmatrix}, \quad (10)$$

where: e_{ij} is the j th element of the i th eigenvector. Each eigenvalue is equal to the variance of the respective PC image along the new coordinate axes, and the sum of all the eigenvalues must equal the sum of all the band variances of the original image, thus preserving the total variance in the data. Since C_{PC} is diagonal, the principal component images are uncorrelated and, by convention, are ordered by decreasing variance, such that PC_1 has the largest variance and PC_K has the lowest. The PC result is removal of any correlation present in the original K dimensional data (Schwengerdt, 2006). PC images are useful for reducing data dimensionality, condensing topographic and spectral information. Moreover, DADSR observations at ten synoptic stations are used to check whether the principal components regression (PCR) method can obtain satisfactory spatial variation

of solar radiation over the Fars province. The spatial mode (S-mode) of PCA was used to input variables (ten variables) in order to spatial modelling of DADSR. In this method, in the first step, the raster file of extraterrestrial solar radiation was obtained over the Fars province for each day and then, MODIS products (bands 1-7, AOD and WV) were cropped for Fars province boundaries. In the second step, all raster data were used as the input files for S-mode PCA computation. Finally, the extracted S-mode PCA values for all stations considered the calculating regression model. The spatial regression model was applied to estimate DADSR over the Fars province.

As the significant coefficient was estimated in the regression model, the model adequacy was examined by testing whether the residuals were serially independent. A useful test for serial correlation of regression residuals is the Durbin-Watson test. The null hypothesis of this test assumes that the residuals are serially independent. The alternative hypothesis assumes that the residuals are consistent with a first order autoregressive process. The Durbin-Watson statistic (d) is:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}, \quad (11)$$

which: e_i is the residual for case i .

The null hypothesis of this test is rejected if the Durbin-Watson statistic is less than the critical values for this test. Also, the autocorrelation function of the residual is plotted to reflect the statistical relationship between residuals for various lags. Normalized RMSE (NRMSE) and correlation coefficient (R^2) are used to assess the model skill.

The methodology of this study is schematically shown in Fig. 2.

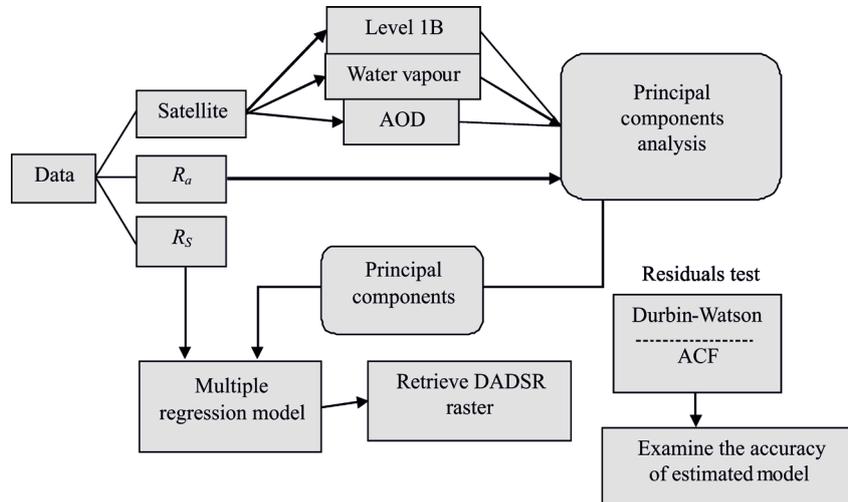


Fig. 2. Flowchart to estimate daily average downward shortwave radiation from the MODIS data for clear sky.

RESULTS AND DISCUSSION

Extraterrestrial radiation (R_a) and MODIS retrieved parameters, such as the aerosol optical depth, atmospheric water vapour, and level1B reflectance over bands 1-7, were used as the candidate predictor variables for estimating (R_s). The correlation matrices of these ten variables were separately calculated for each station. The applied methods were explained for Bajgah station, which is a meteorological station at the Agricultural Research Station of Shiraz University, and its meteorological parameters have been frequently used in previous works (Shirvani *et al.*, 2015). The applied methods for the other stations are similar to Bajgah station. The upper panel of Fig. 3 shows a correlation matrix of the ten variables in Bajgah station. This matrix indicates that the correlation coefficients between each two variables are high and statistically significant, and the highest correlation is between the reflectance in band 1 and 4. The lower panel of Fig. 3 is the scatter plot for these variables, arranged in the same pattern as the correlation matrix in the upper panel. The close relationship between the used bands can be seen in Fig. 3, while each band contains different information. To capture the information of the bands, PCA was carried out to extract uncorrelated principal components of the bands. Principal components of the

input predictor variables were computed based on the correlation matrix. Table 2 indicates the percentages of seven eigenvalues, or variance of each principal component, and the cumulative percentages of variances that accounted for the principal components for all stations. For example, the first four principal components for Bajgah station explain 97% of the total variance of the ten predictor variables.

The scree plot, which is the plot of the eigenvalues versus the corresponding principal component number for Bajgah station, is shown in Fig. 4. This Figure indicates that the magnitudes of eigenvalues of the last few principal components are close to zero and the behaviour of scree plot is like a straight line. Therefore, the first seven principal components, which carry almost all of the predictor information, were used in the regression models. For example, the applied regression revealed the significant (of 95% level) coefficients for the first five principal components for Bajgah station as Eq. (12):

$$R_s = 28.0 + 1.22_{(0.05)} PC1 - 0.827_{(0.14)} PC2 + 0.914_{(0.19)} PC3 + 3.14_{(0.21)} PC4 - 4.32_{(0.36)} PC5. \quad (12)$$

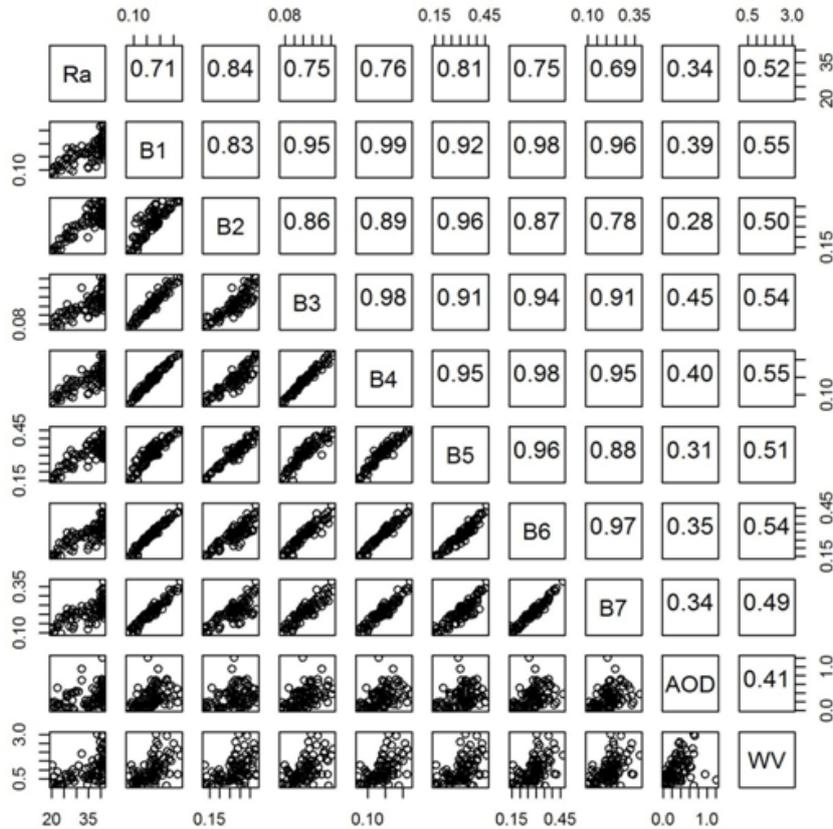


Fig. 3. Scatter plot matrix of the MODIS first seven spectral bands at the Bajgah station.

Table 2. Percentage and cumulative percentage from *PC1* to *PC7* variance in all

| Station | Variance | <i>PC1</i> | <i>PC2</i> | <i>PC3</i> | <i>PC4</i> | <i>PC5</i> | <i>PC6</i> | <i>PC7</i> |
|---------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 1. Eghlid | percentage | 69.0 | 15.8 | 9.4 | 2.4 | 2.1 | 1.0 | 0.2 |
| | cumulative | 69.0 | 84.8 | 94.2 | 96.6 | 98.7 | 99.7 | 99.9 |
| 2. Bajgah | percentage | 77.1 | 9.6 | 5.8 | 4.5 | 1.6 | 0.8 | 0.2 |
| | cumulative | 77.1 | 86.7 | 92.5 | 97 | 98.6 | 99.4 | 99.6 |
| 3. Kazeroon | percentage | 80.6 | 8.4 | 6.6 | 2.8 | 0.8 | 0.6 | 0.2 |
| | cumulative | 80.6 | 89.0 | 95.6 | 98.4 | 99.2 | 99.8 | 100.0 |
| 4. Shiraz | percentage | 82.0 | 9.1 | 4.7 | 1.7 | 1.5 | 0.6 | 0.3 |
| | cumulative | 82.0 | 91.1 | 95.8 | 97.5 | 99.0 | 99.6 | 99.9 |
| 5. Neyriz | percentage | 80.8 | 12.1 | 2.6 | 2.1 | 1.1 | 0.8 | 0.3 |
| | cumulative | 80.8 | 92.9 | 95.5 | 97.6 | 98.7 | 99.5 | 99.8 |
| 6. Fasa | percentage | 82.2 | 7.9 | 5.0 | 3.1 | 1.2 | 0.3 | 0.1 |
| | cumulative | 82.2 | 90.1 | 95.1 | 98.2 | 99.4 | 99.7 | 99.8 |
| 7. Firoozabad | percentage | 82.2 | 8.1 | 5.1 | 2.9 | 1.2 | 0.4 | 0.1 |
| | cumulative | 82.2 | 90.3 | 95.4 | 98.3 | 99.5 | 99.9 | 100.0 |
| 8. Darab | percentage | 80.1 | 10.2 | 4.2 | 3.3 | 1.3 | 0.5 | 0.3 |
| | cumulative | 80.1 | 90.3 | 94.5 | 97.8 | 99.1 | 99.6 | 99.9 |
| 9. Jahrom | percentage | 80.8 | 10.3 | 4.5 | 2.7 | 1.1 | 0.5 | 0.1 |
| | cumulative | 80.8 | 91.1 | 95.6 | 98.3 | 99.4 | 99.9 | 100.0 |
| 10. Lamerd | percentage | 82.6 | 7.3 | 6.6 | 1.9 | 0.9 | 0.5 | 0.1 |
| | cumulative | 82.6 | 89.9 | 96.5 | 98.4 | 99.3 | 99.8 | 99.9 |

The values given in the parenthesis signify the standard error of coefficient. These five coefficients were found to be statistically significant using the t-test. For example, the t-value of $\frac{1.22}{0.05} = 23$ is significantly higher than the quantiles of t-distribution with $\alpha = 0.05$ and 84 degree of freedom ($t_{0.025}(84) = 1.96$). The principal components 6 and 7 were not significant in Bajgah station. In the ANOVA (Analysis of variance) table, the value of the F-statistic was 181 for this station, which is statistically higher than the quantile

of F-distribution with five and 84 degree of freedom at 5 significance level ($F_{0.05}(5,84) = 1.96$). The corresponding P-value of F-statistic is close to zero, which indicates that the fitted principal components regression model is adequate for prediction of R_s . The computed Durbin-Watson statistic was 0.8, which is smaller than the critical value, indicating that the null hypothesis of Durbin-Watson test was accepted. Therefore, the residuals of the fitted model were serially independent. The autocorrelation function

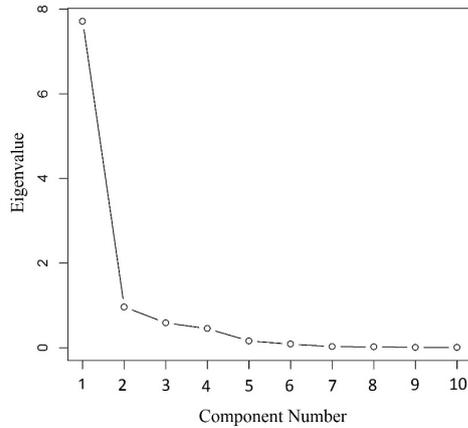


Fig. 4. Scree plot of Bajgah station components.

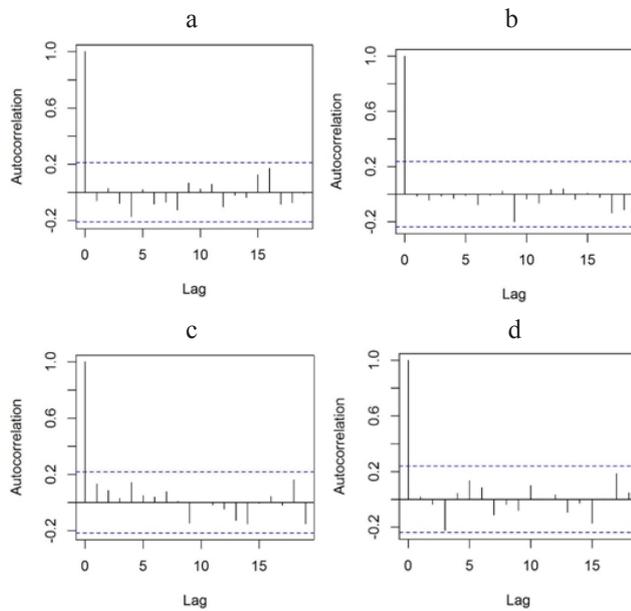


Fig. 5. Autocorrelation function of the residuals for: a – Bajgah, b – Darab, c – Fasa, and d – Shiraz.

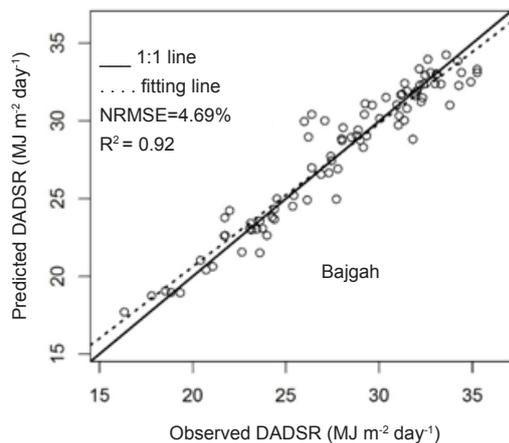


Fig. 6. Scatter plot of the observed versus predicted R_s in Bajgah station.

Table 3. Summary of statistical parameters of the stations used in this

| No. | Station | R^2 | RMSE ($\text{MJ m}^{-2} \text{d}^{-1}$) | NRMSE (%) | MBE ($\text{MJ m}^{-2} \text{d}^{-1}$) |
|-----|------------|-------|--|--------------|---|
| 1. | Darab | 0.96 | 0.99 | 3.7 | 0.00 |
| 2. | Eghlid | 0.95 | 1.14 | 4.46 | 0.00 |
| 3. | Fasa | 0.96 | 0.90 | 3.30 | 0.00 |
| 4. | Firoozabad | 0.91 | 2.04 | 6.16 | 0.00 |
| 5. | Jahrom | 0.93 | 1.00 | 4.71 | -0.045 |
| 6. | Kazeroon | 0.86 | 1.72 | 7.83 | 0.00 |
| 7. | Lamerd | 0.88 | 1.50 | 5.92 | 0.010 |
| 8. | Neyriz | 0.96 | 0.99 | 3.88 | 0.00 |
| 9. | Shiraz | 0.92 | 1.30 | 5.41 | 0.00 |
| 10. | Bajgah | 0.92 | 1.31 | 4.69 | 0.00 |

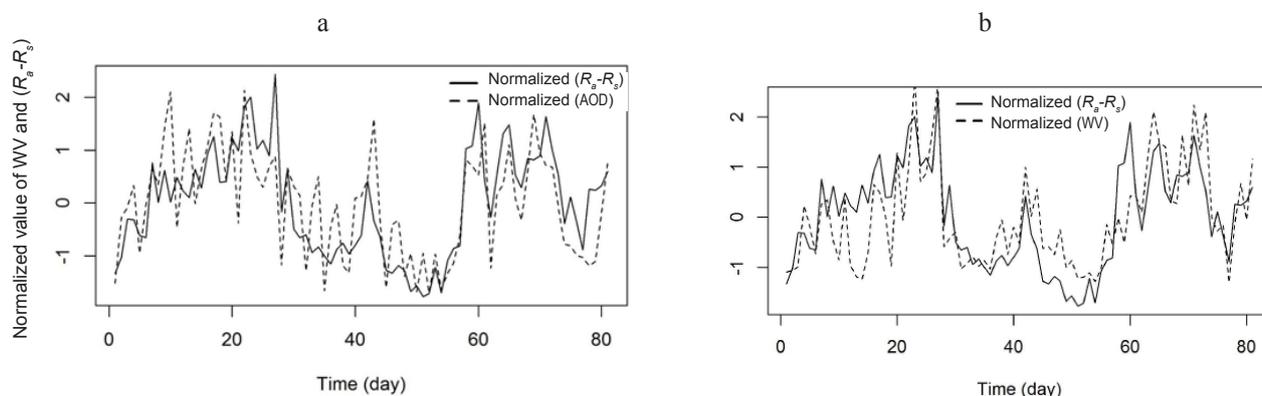
(ACF) of the residuals for Bajgah station (Fig. 5), does not show any patterns. Also, Fig. 5 indicates that the residuals are uncorrelated for various lags. Therefore, according to the given results, for Bajgah station, the fitted principal components regression model is considered to estimate R_s .

The scatter plot for observed and predicted R_s using the fitted model for Bajgah station is shown in Fig. 6. R^2 and NRMSE values of the fitted model are presented in Table 3.

Sensitivity analysis of the regression model to the input variables was performed to choose the best predictor among all, such that one variable was removed from ten predictors and then principal components regression was reconstructed using all remaining predictors. The results indicated that R_a is an important variable among all and the lowest R^2 and highest RMSE were produced when this predictor was removed, such that the estimated NRMSE was between 7 and 10.4. The results also indicated that the AOD and water vapour had the greatest influence on the calculated DADSR, such that the estimated NRMSE was between 3.5 and 7.8, when one of these predictors was removed from the candidate predictors. For example, in Eghlid station, inclusion of AOD and WV as inputs to the proposed model reduced the error to 4.5%, while the model performance showed an error equal to 7.6% without considering these two variables. Without applying *PCA*, the ten predictor variables were separately used in the stepwise multiple regression for each station. The results showed that the RMSE ($\sim 1.3 - 2 \text{ MJ m}^{-2} \text{d}^{-1}$) of this regression is higher than the principal component regression. For example, the R^2 and RMSE of the fitted multiple regression without applying *PCA* were 0.9 and $1.5 \text{ MJ m}^{-2} \text{d}^{-1}$, respectively, for Bajgah station. Therefore, the regression based on principal components predicts R_s better than regression based on original variables (Table 4).

Table 4. Statistical variables before and after applying *PCA* on regression model

| Stations | R^2 | | RMSE ($\text{MJ m}^{-2} \text{d}^{-1}$) | | NRMSE (%) | |
|----------|--------------------|------------|---|------------|--------------------|------------|
| | original variables | <i>PCA</i> | original variables | <i>PCA</i> | original variables | <i>PCA</i> |
| Bajgah | 0.90 | 0.92 | 1.50 | 1.30 | 5.36 | 4.65 |
| Darab | 0.96 | 0.97 | 1.30 | 0.99 | 3.99 | 3.66 |
| Fasa | 0.94 | 0.96 | 1.08 | 0.86 | 4.15 | 3.30 |
| Eghlid | 0.86 | 0.95 | 1.93 | 1.12 | 7.55 | 4.38 |
| Shiraz | 0.91 | 0.93 | 1.40 | 1.29 | 5.83 | 5.37 |
| Neyriz | 0.94 | 0.96 | 1.15 | 0.98 | 4.51 | 3.84 |

**Fig. 7.** Comparison between difference of the R_a and R_s (R_a-R_s) and: a – AOD and b – WV over the Bajgah.

The presented results of RMSE of principal components regression (PCR) is compatible with the results of Kim and Liang, 2010, who estimated daily shortwave net radiation with RMSE of 2.8-4 $\text{MJ m}^{-2} \text{d}^{-1}$, for the seven SURFRAD sites using the Tang *et al.* (2006) method (Kim and Liang, 2010; Tang *et al.*, 2006).

Figure 7 illustrates that the difference between R_a and global received radiation at the earth surface (R_s) increases when the atmospheric transparency decreases due to the effect of aerosols and water vapour and *vice versa*. Figure 8 shows the scatter plots of the estimated radiation versus ground observation using PCR for each station. Also, the time series of estimated DADSR versus observed data in Fasa station (low RMSE) and Firoozabad station (High RMSE) are shown in Fig. 9. The results show that the performance of the model is increasing when AOD with high variation was used within the model. For example, the mean and standard deviation of AOD in Lamerd (Firoozabad) station were 0.76 (0.71) and 0.48 (0.48), respectively.

Variations of the atmospheric transparency due to AOD variations can affect the performance of the solar radiation models.

To validate our spatial method, Fig. 10 shows comparisons of the DADSR estimated from MODIS observation and in situ measurements at 10 synoptic stations over the Fars province during the studied period. The RMSE between estimated and measured values is 2.3 $\text{MJ m}^{-2} \text{d}^{-1}$ in all stations. In this figure, we see that, generally, DADSR estimated from MODIS data is in good agreement with that measured in situ. The NRMSE at individual stations using the fitted model for all stations was estimated between 8.2 in Fasa station and 20 in Firoozabad station, indicating that the errors have increased in comparison with the given errors for the individual model in Table 3.

As shown in Table 3, the error of the model in Lamerd and Firoozabad station was large. On the other hand, the mean and standard deviation of AOD in Eghlid station were the smallest (0.16 and 0.20, respectively), and the estimated DADSR shows good agreement (in Lamerd station, the annual mean of AOD was 4.5 times that of Eghlid station). Figure 11 shows daily average global radiation over the Fars province in 4 clear sky days of separate months in 2011.

Developing and modelling spatial variations of daily average solar radiation via satellite data is one of the recent challenges in remote sensing studies. This study focused

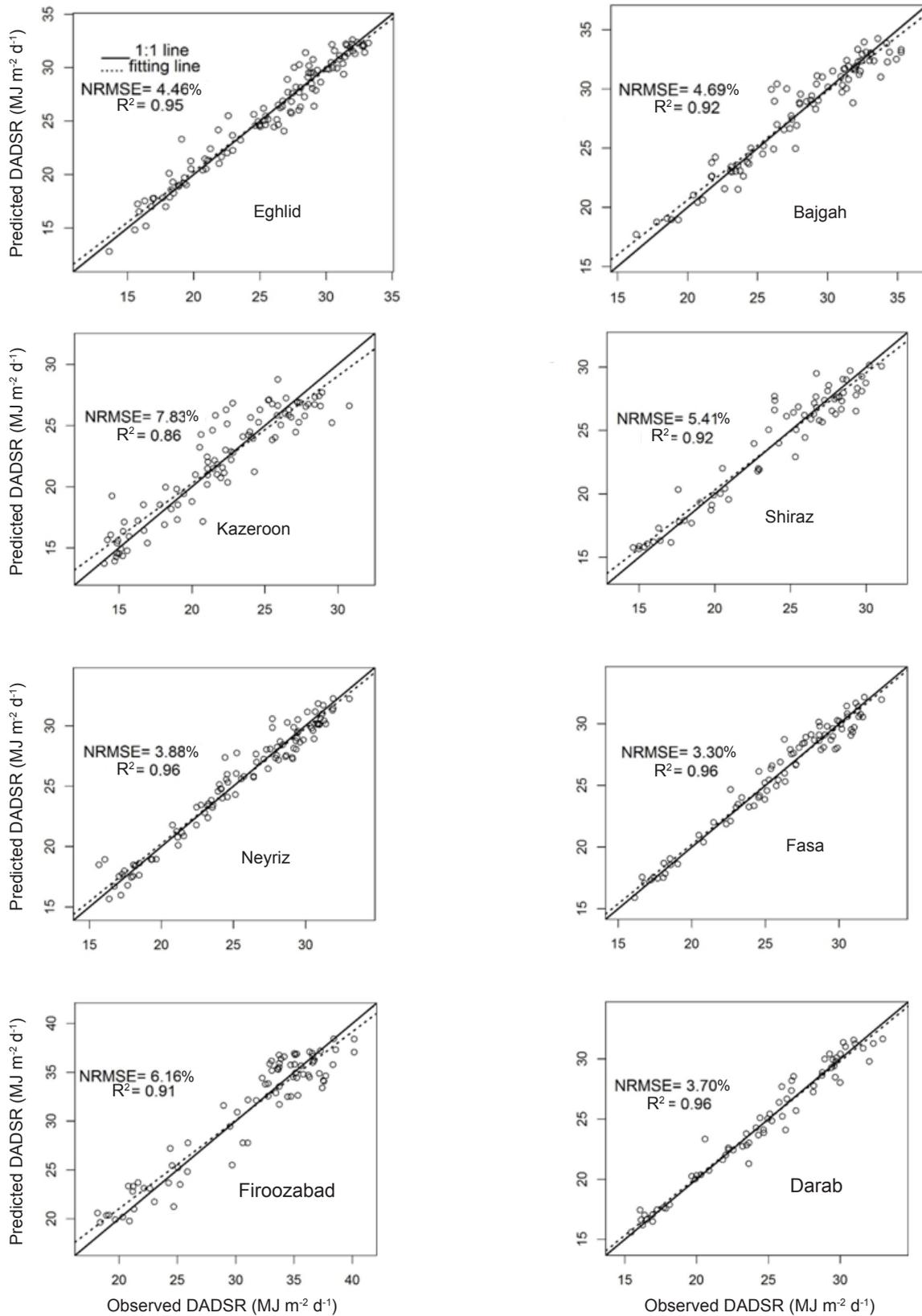


Fig. 8. Comparison of the estimated and observed daily average downward shortwave radiation ($\text{MJ m}^{-2} \text{d}^{-1}$) in the surveyed stations. Normalize RMS Error percent show the accuracy of the estimation.

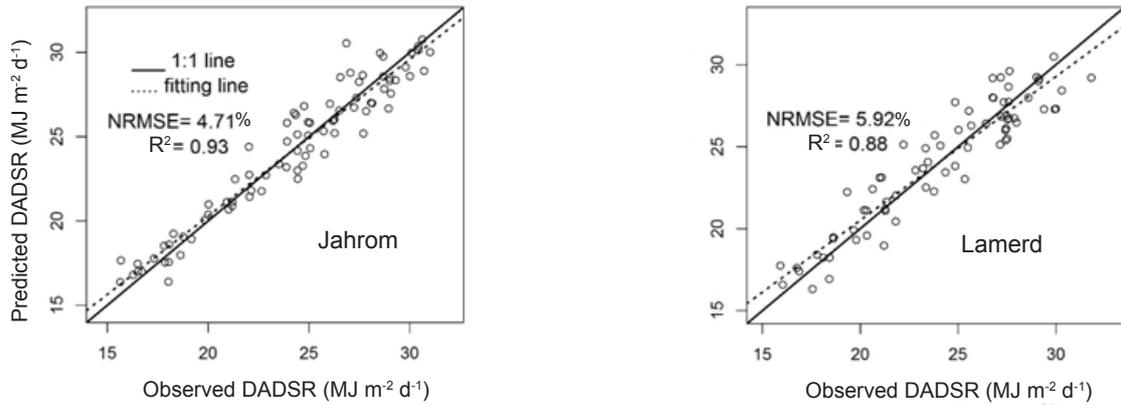


Fig. 8. Continuation.

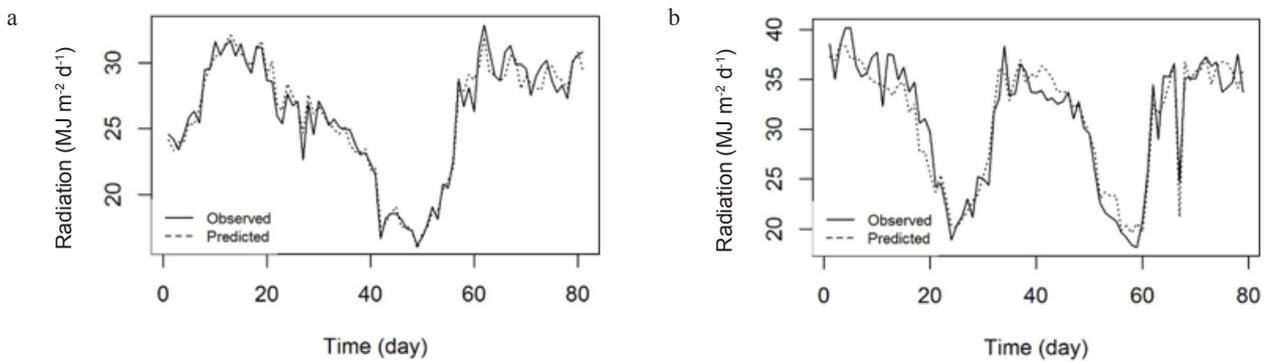


Fig. 9. Time series of the estimated versus observed of DADSR in: a – Fasa and b – Firoozabad stations, respectively during study period.

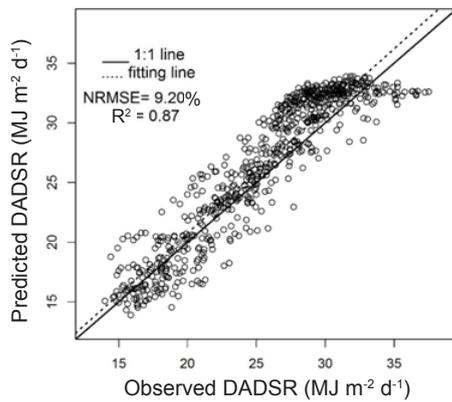


Fig. 10. Validation of the spatial estimation of DADSR (MJ m⁻² d⁻¹) using all station data.

on the estimation of daily average downward shortwave radiation using MODIS products as model input. Besides, the combination of the principal components analysis and regression models was used to estimate DADSR without any ancillary data. Principal components regression improved the prediction of R_s in comparison with the regression based on the original variables. Many current studies used a complicated algorithm for estimation the received surface solar radiation. The current study used

the multivariate statistical methods to estimate the daily average surface shortwave radiation adequately with considering the atmospheric effects on the solar spectrum. The used statistical methods could predict shortwave accurately with RMSE less than ~ 2 MJ m⁻² d⁻¹ which is less than previous attempts. We acknowledge that the proposed method does not involve any complex radiative transfer model.

CONCLUSIONS

1. Principal components analysis appeared to be useful technique to reduce a data set containing a large number of variables. Therefore, it can be widely used for modelling many multispectral sensor data especially moderate resolution imaging spectroradiometer. In addition, some techniques such as principal components analysis should be taken into account before applying regression models when the predictor variables are highly correlated.

2. Extraterrestrial radiation water vapour, and aerosol optical depth are important factors for prediction of received solar radiation at the earth surface and the seasonal variation of these parameters can considerably affect the performance of local radiation models.

3. Average downward solar radiation was skilfully estimated and the results indicated good agreement with field measurements at ten synoptic stations with NRMSE lower than 10% in all stations.

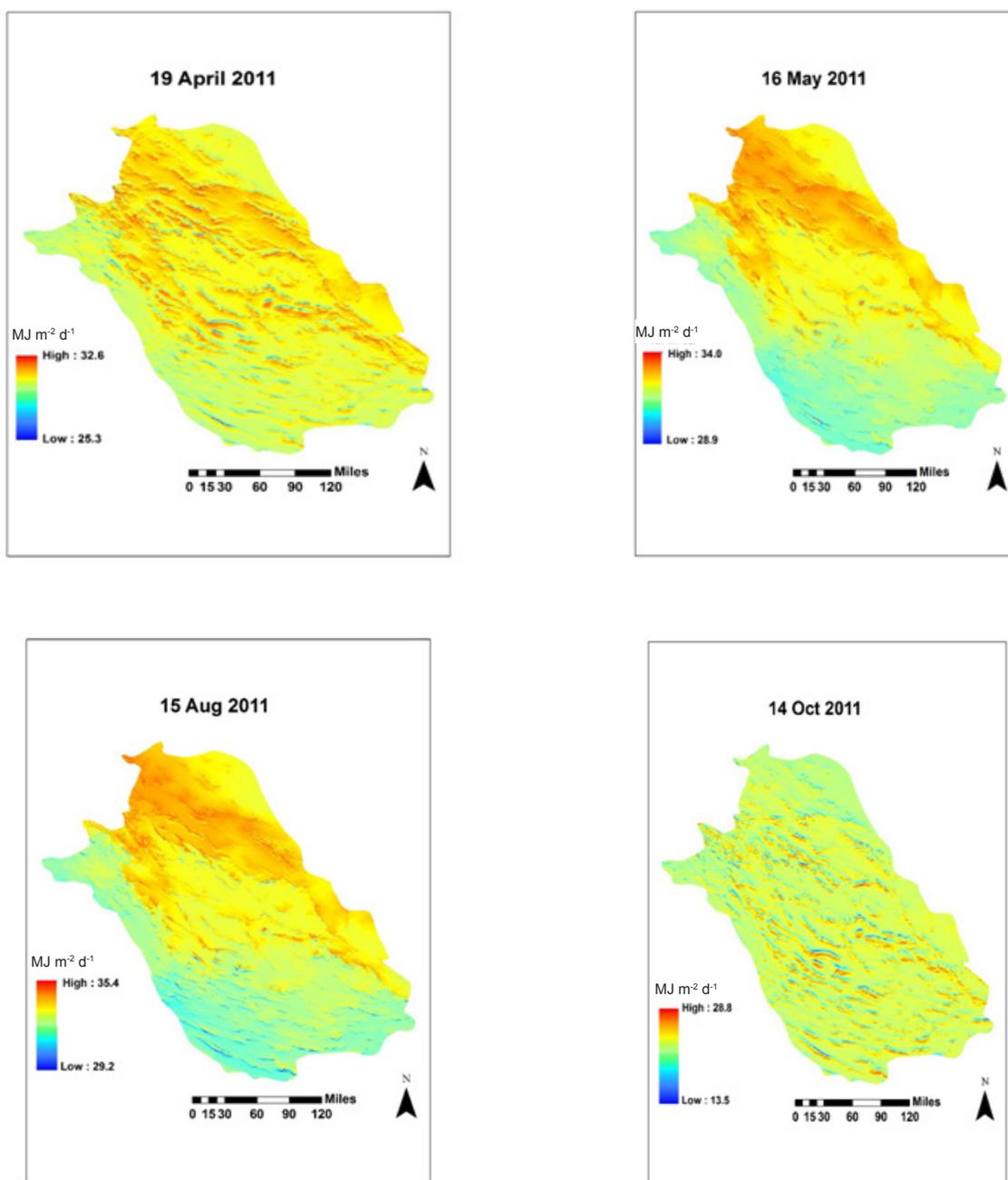


Fig. 11. Spatial distribution of the estimated ($\text{MJ m}^{-2} \text{d}^{-1}$) DADSR from principal components regression model over Fars province, Iran.

4. The daily average downward solar radiation variability was mainly dependent on the annual variation of aerosol optical depth. The annual mean of Aerosol optical depth in the southern parts of the Fars province was 4.5 times higher than that in the northern part.

Conflict of interest: The Authors declare no conflict of interest.

REFERENCES

- Adeboye O.B., Osunbitan J.A., Adekalu K.O., and Okunade D.A., 2009. Evaluation of FAO-56 Penman-Monteith and temperature based models in estimating reference evapotranspiration using complete and limited data, application to Nigeria. *Agric. Eng. Int.: CIGR J.*, Manuscript 1291, 11, October, 2009.
- Bisht G., Venturini V., Islam S., and Jiang L., 2005. Estimation

- of the net radiation using MODIS (Moderate Resolution Imaging Spectroradiometer) data for clear sky days. *Remote Sensing Environ.*, 97, 52-67.
- Change C., 2001.** The scientific basis, intergovernmental panel on climate change (Eds J.T. Houghton, Y. Ding, D.J. Griggs, *et al.*). Cambridge University Press, Cambridge, U.K.
- Houghton J.T., Ding Y., Griggs D.J., Noguer M., van der Linden P.J., Dai X., Maskell K., and Johnson C., 2001.** Climate change 2001: the scientific basis. Cambridge University Press, Cambridge, U.K.
- Huang G., Liu S., and Liang S., 2012.** Estimation of net surface shortwave radiation from MODIS data. *Int. J. Remote Sensing*, 33, 804-825.
- Irmak S. and Allen R.G., 2005.** Closure on Predicting daily net radiation using minimum climatological data (Eds S. Irmak, A. Irmak, J.W. Jones, T.A. Howell, J.M. Jacobs, R.G. Allen and G. Hoogenboom). *J. Irrigation Drainage Eng.*, 131, 389-390.
- Justice C., Eck T., Tanré D., and Holben B., 1991.** The effect of water vapour on the normalized difference vegetation index derived for the Sahelian region from NOAA AVHRR data. *Int. J. Remote Sensing*, 12, 1165-1187.
- Kaplanis S. and Kaplani E., 2007.** A model to predict expected mean and stochastic hourly global solar radiation I (h; nj) values. *Renewable Energy*, 32, 1414-1425.
- Kim H.-Y. and Liang S., 2010.** Development of a hybrid method for estimating land surface shortwave net radiation from MODIS data. *Remote Sensing Environ.*, 114, 2393-2402.
- Levy R., Mattoo S., Munchak L., Remer L., Sayer A., and Hsu N., 2013.** The collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech. Discuss.*, 6, 159-259.
- López G. and Batlles F.J., 2014.** Estimating Solar Radiation from MODIS Data. *Energy Procedia*, 49, 2362-2369.
- Mueller R., Dagestad K.-F., Ineichen P., Schroedter-Homscheidt M., Cros S., Dumortier D., Kuhlemann R., Olseth J., Piernavieja G., and Reise C., 2004.** Rethinking satellite-based solar irradiance modelling: The SOLIS clear-sky module. *Remote Sensing Environ.*, 91, 160-174.
- Phakamas N., Jintrawet A., Patanothai A., Sringam P., and Hoogenboom G., 2013.** Estimation of solar radiation based on air temperature and application with the DSSAT v4. 5 peanut and rice simulation models in Thailand. *Agric. Forest Meteorol.*, 180, 182-193.
- Qin J., Chen Z., Yang K., Liang S., and Tang W., 2011.** Estimation of monthly-mean daily global solar radiation based on MODIS and TRMM products. *Applied Energy*, 88, 2480-2489.
- Qu J.J., Gao W., Kafatos M., Murphy R.E., and Salomonson V.V., 2006.** Earth Science Satellite Remote Sensing. *Science Instruments*, 1, 408.
- Rahimikhoob A., 2010.** Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. *Renewable Energy*, 35, 2131-2135.
- Samani Z., Bawazir A.S., Bleiweiss M., Skaggs R., and Tran V.D., 2007.** Estimating daily net radiation over vegetation canopy through remote sensing and climatic data. *J. Irrigation Drainage Eng.*, 133(4), 291-297.
- Schowengerdt R.A., 2006.** Remote sensing: models and methods for image processing. Academic Press, San Diego, California.
- Shirvani A., Moradi F., and Moosavi A.A., 2015.** Time series modelling of increased soil temperature anomalies during long period. *Int. Agrophys.*, 29, 509-515.
- Su H., McCabe M., Wood E., Su Z., and Prueger J., 2005.** Modeling evapotranspiration during SMACEX: Comparing two approaches for local-and regional-scale prediction. *J. Hydrometeorology*, 6, 910-922.
- Tang B., Li Z.-L., and Zhang R., 2006.** A direct method for estimating net surface shortwave radiation from MODIS data. *Remote Sensing Environ.*, 103, 115-126.
- Wang D., Liang S., He T., Cao Y., and Jiang B., 2015.** Surface Shortwave Net Radiation Estimation from FengYun-3 MERSI Data. *Remote Sensing*, 7, 6224-6239.
- Wilks D.S., 2011.** Statistical methods in the atmospheric sciences. Academic Press, San Diego, California.
- Yang K., Huang G., and Tamai N., 2001.** A hybrid model for estimating global solar radiation. *Solar Energy*, 70, 13-22.
- Zhang Y., Li X., and Bai Y., 2015.** An integrated approach to estimate shortwave solar radiation on clear-sky days in rugged terrain using MODIS atmospheric products. *Solar Energy*, 113, 347-357.