Estimation of instantaneous air temperature using remote sensing data

Foroogh Golkar, Ali Akbar Sabzparvar, Reza Khanbilvardi, Mohammad Jafar Nazemosadat, Shahrokh Zand-Parsa & Yousef Rezaei


To link to this article: http://dx.doi.org/10.1080/01431161.2017.1382743

Published online: 01 Oct 2017.
Estimation of instantaneous air temperature using remote sensing data

Foroogh Golkar\textsuperscript{a,b}, Ali Akbar Sabziparvar\textsuperscript{c}, Reza Khanbilvardi\textsuperscript{d}, Mohammad Jafar Nazemosadat\textsuperscript{e}, Shahrokh Zand- Parsa\textsuperscript{f} and Yousef Rezaei\textsuperscript{g}

\textsuperscript{a}Agrometeorology, Bu-Ali Sina University, Hamedan, Iran; \textsuperscript{b}NOAA-CREST Institute, The City University of New York, New York, NY, USA; \textsuperscript{c}Meteorology, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran; \textsuperscript{d}Civil Engineering, NOAA-CREST Institute, The City University of New York, New York, NY, USA; \textsuperscript{e}Department of Oceanic and Atmospheric Research Centre and Water Eng., Faculty of Agriculture, Shiraz University, Shiraz, Iran; \textsuperscript{f}Department of Water Engineering, Faculty of Agriculture, Shiraz University, Shiraz, Iran; \textsuperscript{g}Department of Civil Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamedan, Iran

\textbf{ABSTRACT}

In arid areas, the variation of air temperature can be considerable, so instantaneous air temperature ($T_{ai}$) estimation is needed in different environmental researches. In this research, two different remote sensing data are used for estimating $T_{ai}$ for clear sky days in 2009 in Fars Province, Iran, including atmospheric temperature profile and land surface temperature (LST) data from Moderate Resolution Imaging Spectroradiometer. The $T_{ai}$ from a number of surface weather sites is used to judge the best $T_{ai}$ estimation. Stations’ elevation, latitude, and land cover type are considered to show their effect on $T_{ai}$ estimation. The estimated $T_{ai}$ evaluation focuses on daily and seasonal timescales in the daytime and night time separately. Both LST and vertical temperature profile data produced relatively high coefficient of determination values and small root mean square error value for $T_{ai}$ estimation, especially during the night time. Land cover and elevation vary the error values in $T_{ai}$ estimation more, when LST data is used. In comparison atmospheric temperature profile indicates a smaller error in $T_{ai}$ estimation in spring and summer and in urban land cover type, while using LST data presents a better result in fall and winter especially at night time.

\textbf{1. Introduction}

Air temperature is known as measured temperature at the shelter height 2 m above ground (hereafter $T_a$). Knowledge of instantaneous air temperatures (hereafter $T_{ai}$) is critical to understand the environmental conditions. Spatial and temporal information on instantaneous air temperature, which is considered in current study, is important for modelling regional evapotranspiration and estimating net radiation. Land surface evapotranspiration models routinely use solar radiation directly or, in combination with longwave radiation, to provide a measure of the net energy available to evaporate.
water. Variation of the downwelling components of net radiation is dominated by atmospheric variations, primarily air temperature and clouds (Prigent, Aires, and Rossow 2003). The air temperature has temporal and spatial variations. The density of the station network (number of station per unit area) is frequently insufficient to represent the spatial distribution of air temperature at detailed spatial scales due to lack of high-resolution data, especially in developing countries (Emamifar, Rahimikhoob, and Noroozi 2013). However, the development of satellite remote sensing technology provides an opportunity to obtain such high-resolution data.

Several authors have proposed methods to estimate air temperature using remote sensing data. In estimating air temperature, land surface temperature (LST) has been frequently used. Surface temperatures are governed by land–atmosphere interactions and the energy fluxes between both. LST is a key parameter in land surface processes due to its control of the upward terrestrial radiation, and consequently, the control of the surface sensible and latent heat flux exchange with the atmosphere (Arya 2001). Surface heat fluxes can induce local convection in the boundary layer, producing changes in air temperature, surface winds, cloudiness, and (potentially) precipitation (Arya 2001).

Although LST and air temperature are strongly correlated, both have different physical meanings, magnitudes, measurement techniques, response to atmospheric conditions, and diurnal phase (Jin and Dickinson 2010). Vancutsem et al. (2010) explained that the lapse rate between LST and instantaneous air temperature is controlled by a complex surface energy. The temperature lapse rate can vary greatly in a diurnal cycle, and patterns are influenced seasonally through variations in the day and night lengths. In this case, Cresswell et al. (1999) study showed that during daytime, surface temperature is generally higher than air temperature, and at night time, the opposite occurs. This point is used indirectly (due to lack of measured LST data in study area) to evaluate Moderate Resolution Imaging Spectroradiometer (MODIS) LST product in current study.

Several authors have mentioned many other factors that have impact on the LST–Ta lapse rate such as soil emissivity, moisture content and wind velocity, turbulence, cloud cover, water vapour content, elevation, topography, leaf area index, and vegetation (Huband and Monteith 1986; Jin and Dickinson 2010; Mildrexler, Zhao, and Running 2011). Benali et al. (2012) stated that several factors such as wind velocity and vegetation are important in energy balance in the land–atmosphere system, and consequently, influence the LST–instantaneous air temperature relation; however, accurate wind data is difficult to obtain in the studied area and the biophysical controls of vegetation in the energy balance are quite complex. Thus, the main goal of this work is to estimate instantaneous air temperature using easy remote sensing methods with minimum input data.

Most recent studies have focused on estimating the minimum, maximum, mean daily, or mean monthly air temperatures using LST data; few works have addressed estimating instantaneous air temperature. Determining Ta is an indispensable element to calculate hydrological variables, such as evapotranspiration (Allen et al. 2007) and net radiation on a regional scale, especially in arid climate. Instantaneous air temperature at the time of satellite overpass is needed when using remote sensing methods in solar net radiation estimation (Bisht et al. 2005). Nevertheless, despite the methods used, previous studies reported errors of about 2.0–3.0°C for a variety of target variables and both spatial and temporal resolutions (Zakšek and Schroedter-Homscheidt 2009).
Sun et al. (2005) suggested a method based on thermodynamics to retrieve regional $T_{ai}$ using MODIS LST and normalized difference vegetation index (NDVI) data. Crop Water Stress Index (CWSI) was used as the basis for the establishment of the relationship between air temperature and LST. An accuracy of within 3.0°C was achieved for more than 80% of the data processed in their study. The determination of CWSI and aerodynamic resistance are two parameters that were crucial in their method. Lack of input data in calculating the parameters such as net radiation as well as CWSI were the limitations in the use of Sun et al.’s (2005) method for estimating $T_{ai}$ based on satellite data and thermodynamics in our study.

Artificial intelligence methods are another approach which have been developed for estimating air temperature using remote sensing data, including artificial neural network (Zhao, Zhang, and Shijin 2008; Mao et al. 2007; Sahhin 2012) and the M5 model tree (Emamifar, Rahimikhoob, and Noroozi 2013). Zhao, Zhang, and Shijin (2008) developed an algorithm based on back propagation (BP) neural network for retrieval of near-surface daily mean, maximum, and minimum air temperature from remotely sensed data in southwestern China. Parameters used for the training of the BP neural network included: remotely sensed albedo, NDVI, layered meteorological data the station’s observed daily mean, maximum, and minimum temperature provided by geographical information system (GIS) as well as the digital elevation model (DEM) of the study site. They concluded that the BP neural network integration with surface meteorological observations could be a promising approach for retrieving near-surface air temperature with reliable accuracy.

Mao et al. (2007) used an algorithm based on the radiance transfer model (Moderate Resolution Atmospheric Transmission (MODTRAN4)) and a dynamic learning neural network for estimation of near-surface air temperature from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. The comparison of estimation results with ground measurement data at meteorological stations indicates that the RM-NN (radiative transfer model with neural network) can be used to estimate near surface air temperature with acceptable accuracy from ASTER data. Mean monthly air temperature was forecasted in 20 cities of Turkey based on remote sensing and artificial neural network (Sahhin 2012). In their research city, month, altitude, latitude, longitude, monthly mean LST were chosen as input for network. The results showed the accuracy of 1.0–1.3°C between the estimated and measured mean monthly air temperature. Emamifar, Rahimikhoob, and Noroozi (2013) used M5 model tree to estimate air temperature in the southwest of Iran. The input variables for the M5 model tree were the daytime and night time MODIS-Terra LST, extraterrestrial solar radiation, and Julian day. The results of their study showed that mean daily air temperature can be estimated with acceptable levels of the statistical indicators from MODIS data and from the two geographic parameters using the M5 model. These methods require observation data to provide training data for predictions using artificial intelligence. However, the goal of the current study is focusing on estimating instantaneous air temperature without the direct use of observation data.

In addition to the above-mentioned methods, Benali et al. (2012) explained that the regression models can be suitable for areas with complex landscape characteristics. Low density of the observed data, complex landscape characteristics, and lack of various parameters in the studied area motivated this study to estimate $T_{ai}$ based on simple
statistical analysis in remote sensing method using MODIS LST data as the only input data.

As another approach, MODIS atmospheric temperature data is used to estimate $T_{ai}$ as well. Some authors have proposed methods to estimate $T_{ai}$ using temperature profile data. The satellite temperature profile data includes vertical temperature profile data (Retrieved_Temperature_Profile). Rhee and Jungho (2014) improved a method suggested by Mendez (2004) to estimate air temperature based on a linear interpolation between the bottom profile level (1000 hPa) and ground level (2 m above ground level). He used the pressure difference between 1000 hPa and 620 hPa levels, considering the vertical properties of the atmosphere, and calculated the ground level air temperature as the temperature of the 1000 hPa plus the adiabatic lapse rate, which resulted in a strong agreement with the observed data ($R^2$ (coefficient of determination) = 0.8 and RMSE (root mean square error) = 1.5°C) for their study area. Tang and Li (2008) argued that approximating air and dew point temperatures at 1000 hPa as near-surface temperatures in estimating net longwave flux may be inappropriate due to variations caused by Earth’s terrain and suggested using the hydrostatic assumption in the atmosphere to estimate near surface temperatures. They assumed a hydrostatic atmosphere assumption to extrapolate $T_{ai}$ provided at the lowest vertical pressure level from the MODIS atmospheric profile product to estimate $T_{ai}$. Tang and Li (2008) and Bisht and Bras (2011) improved the instantaneous upwelling longwave radiation calculation using hydrostatic assumption in extrapolating $T_{ai}$ from MODIS temperature profile data. Air temperature derived from MODIS temperature profile data improved the RMSE by 1.0°C and 1.5°C for day and night, respectively (Bisht and Bras 2011).

The objective of this study is to estimate real-time air temperature ($T_{ai}$) in a simple way with minimum input data using remote sensing based methods. Two principal approaches have been used in this study to map $T_{ai}$ from remote sensing data: statistical analysis using MODIS LST data and extrapolating $T_{ai}$ from MODIS air temperature profile. Stations’ elevation, latitude, and land cover are consider to demonstrate their effect on $T_{ai}$ estimation in daily and seasonal timescale (at daytime and night time).

2. Materials and methods

2.1. Study area and data set

The study area for the present work is the Fars Province in southern Iran (surface area of 122, 608 km$^2$, one of the largest provinces in Iran), which is known for its arid and semi-arid climate and flourishing agriculture. The study area experiences four distinct seasons over the year and has complex topography with various land cover types. Such conditions allow the analysis of the effect of various factors, including seasonality, elevation, latitude, and land cover type on air temperature estimation, thus the spatial and temporal heterogeneity has been tested as well in the study area (in Section 2.2).

In Fars Province there are 22 Automatic Weather Stations (AWS). Among these stations, 16 reliable stations were selected based on continuous instantaneous air temperature records for each month in 2009 (Figure 1). Mao-Gui and Wang (2010) improved a computer optimization package based on mean of surface with nonhomogeneity theory with the purpose of improving accuracy in the global estimation of some spatial
properties, given a spatial sample distributed over a heterogeneous surface. They introduced and applied a spatial sampling optimization method using meteorological network. Wang et al. (2012) also reviewed spatial sampling in collecting data for heterogeneous areas such as the studied area. However, based on stratified heterogeneity study (Section 2.2), several subareas have been identified (Figure 1). Spatial sampling study could improve the accuracy of estimation, but a limited number of stations in our study area make the spatial sampling study impossible; thus, this study is based on the only 16 reliable AWS in Fars Province for 2009 and no station selection method has been tested.

AWS record near-surface air temperature (at 2 m height) and soil temperature ($T_{10}$ at 10 cm depth) every 10 min. Measured soil and air temperature data are used indirectly to evaluate LST data from MODIS. Measured air temperature data are also used to validate estimated $T_{air}$. The land cover type of the meteorological stations and the study area is assigned using the land cover map of Fars Province, prepared by Fars Agricultural Organization. Figure 1 shows the dominant land cover in the locality, where the stations are located. All data are from 2009 in this study.

MODIS sensor products from the satellites Terra and Aqua are used as remote sensing data sources. LST in MOD11 and MYD11, and temperature profile from MOD07 and MYD07 are used in the satellite-based data method. Aqua passes over the Fars Province at about 8:55–10:35 and 21:30–23:20 UTC (Coordinated Universal Time) and Terra passes at about 6:25–8:10 and 17:45–19:25 UTC.

This research is based on clear-sky days (with less than 25% cloud in sky) recorded in 2009 in Fars Province, southwest Iran, under the condition that no precipitation occurred in three previous days (to have the same soil moisture condition in the studied days). In this year, 238 days met the desired conditions. Instantaneous air temperature during the overpass of satellites Terra and Aqua is gathered from 16 automated weather stations.

Figure 1. Geographical distribution of AWS in Fars province, Iran; the stations are grouped into three elevation classes and three latitude classes (a), and four land cover types (b).
2.2. MODIS LST data

Using nearest neighbour method in space and time, the MODIS LST data were collected in the studied days. For evaluating MODIS LST data in the study area, measured air temperature and soil temperature at 10 cm depth are used indirectly. Lack of direct surface measurements of LST in the study area means that we can only check the qualitative consistency of the satellite values by comparison with the stations’ air temperature and soil temperature at 10-cm depth. Uncertainty in LST estimates increases when significant variations of temperature occur. The accuracy of MODIS LST in arid and semi-arid areas has been reported to be lower due to higher overestimations of surface emissivity (Hulley and Hook 2009). Moreover, errors in LST retrieval may be larger in bare soil and highly heterogeneous areas due to large uncertainties in surface emissivity, and when the column water vapour content is high (Hulley and Hook 2009). The use of LST products in $T_{ai}$ estimation is strongly influenced by errors on the LST retrievals. In the present work, the accuracy of MODIS LST products on board Terra and Aqua spacecraft were examined indirectly, based on expected behaviour of LST in comparison to measured $T_{ai}$ and $T_{10}$ data.

The main LST and $T_{ai}$ and $T_{10}$ difference lies in their diurnal cycles. The expectation is that during daytime, especially in summer, the value of LST will generally be larger than both $T_{ai}$ and $T_{10}$ and during nighttime, especially in winter, the value of LST will generally be smaller than $T_{ai}$ and $T_{10}$ (Arya 2001). Solar radiation is the driving factor influencing this difference. The surface responds more rapidly to a changing solar force. The surface will warm up more quickly in the morning with the rising sun, inducing a positive LST–$T_{ai}$ and LST–$T_{10}$ difference during daytime, whereas it will cool more rapidly at night, generating a negative LST–$T_{ai}$ and LST–$T_{10}$ difference. The fraction of absorbed solar radiation at the surface is determined by the surface albedo that depends on the vegetation type and fractional cover, the albedo usually decreasing with increasing vegetation density. However, the dominant vegetation impact is associated with its effect on evaporative cooling, denser vegetation is usually associated with more underlying soil moisture and with less restricted transpiration (Prigent, Aires, and Rossow 2003). In the absence of in situ measurements to validate the retrieved LST, the following analysis will attempt to examine the variations of the difference of LST–$T_{ai}$ and LST–$T_{10}$ with the factors that are expected to affect it. Given that all these factors are interconnected, it is difficult to isolate the influence of a single parameter. However, we will confirm that the expected behaviour is observed, and, if that is the case, the errors on the used LST data cannot be large.

The daytime average difference of LST–$T_{10}$ is between 11.0°C and 12.3°C; at night the average difference of LST–$T_{10}$ is between −2.7°C and −4.7°C. The average difference of LST–$T_{ai}$ is 9.7–10.3°C and −5.3 to −5.9°C at daytime and nighttime, respectively. The daytime and nighttime histograms of all differences are plotted for each season to show the distribution of negative and positive difference values. For both MODIS, Terra and Aqua LST data, the LST–$T_{10}$ and LST–$T_{ai}$ value showed a dominant positive result during daytime and negative values during nighttime in all seasons. Thus, the MODIS values of the LST meet the expectations. Figure 2 shows the winter and summer daytime and nighttime LST–$T_{ai}$ value.
In this study, the correlation between LST and observed $T_{ai}$ for daytime and nighttime (at the time of satellite overpass) is examined in several ways, considering stations’ elevation, latitude, and land-use type. These spatial factors were examined to show their effect on the accuracy of $T_{ai}$ estimation. For this purpose, spatial stratified heterogeneity should be tested at the early stage of spatial data analysis (Wang, Zhang, and Bo-Jie 2016). Spatial variation of attributes or uneven distribution of events or their relationship across a region is known as spatial heterogeneity. Stratification of heterogeneity means that the observations are homogeneous within each stratum but not between strata. A stratified heterogeneity is mostly significant if the values within the strata are homogeneous or the variance within the strata is zero; a stratification of heterogeneity vanishes when there is no difference between the strata (Wang, Zhang, and Bo-Jie 2016). The $q$-statistic has been proposed by Wang, Zhang, and Bo-Jie (2016) to measure a spatial stratified heterogeneity. The value of the $q$-statistic varies between 0 and 1. As a larger $q$ value indicates a stronger stratified heterogeneity effect (Wang, Zhang, and Bo-Jie 2016), we used $q$-statistic to show and measure the degree of stratified heterogeneity in our study. The $p$-value test is used to test the hypotheses. If $p < \alpha$, then there is stratified heterogeneity at the significant level $\alpha$.

**Figure 2.** MODIS histograms of LST–$T_{ai}$ in winter daytime (a), winter nighttime (b), summer daytime (c), and summer nighttime (d).
software (Wang, Xin-Hu, and Christakos 2010), is used for computing $q$-statistic and $p$-value for the stratification in the current study. The results are presented in Table 1. As it is shown, all spatial factors including elevation, latitude, and land-use type show significant heterogeneity across the study area. The strongest heterogeneity was seen in land-use type strata and the weakest was seen in latitude strata (it should be mentioned that the elevation and latitude stratification are based on prior knowledge of the study area). These results emphasize on the fact that these spatial factors affect the $T_{ai}$ estimation, and thus the spatial stratification improves the accuracy of estimation. Table 2 compares the average of $T_{ai}$ in each stratum based on ‘Risk Detector’ worksheet of Geodetector software. As it is shown, the average difference in land-use cover strata is more than elevation strata and the average difference in latitude strata is the least. As well as spatial heterogeneity, temporal heterogeneity is critical in current study, and the same test was run for diurnal (day and night) and seasonal temporal factors (Table 1). As expected, for both diurnal and seasonal strata, the heterogeneity was significant (which was stronger for day and night). These results indicate that considering the effect of diurnal and seasonal factors in $T_{ai}$ estimation is critical as it was proved for spatial factors. Table 2 shows the average of $T_{ai}$ in each temporal stratum. In addition to spatial factors, temporal factors show a significant difference in average, which highlights the need for considering both spatial and temporal heterogeneities in current study.

Correlation between LST and observed $T_{ai}$ in all 16 AWS was calculated for daytime and nighttime separately (hereafter, $A_{day}$ and $A_{night}$). Since the study province has a wide range of elevation and changes in elevation cause temperature oscillations, the stations are grouped into three elevation classes: stations with elevation from 400 to 1000 m above sea level (hereafter, ASL), stations with elevation from 1000 to 1600 m ASL, and stations with elevation from 1600 to 2300 m ASL (Figure 1). The MODIS LST and observed $T_{ai}$ correlation were examined in these three groups for daytime and nighttime as well (hereafter, $E_{day}$ and $E_{night}$ methods). Stations are also grouped based on latitude distribution which conveniently groups stations according to climate characteristics.

| Table 1. $q$-statistic test for spatial and temporal stratification heterogeneity. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Statistic       | Land use | Elevation | Latitude | Day and night | Season |
| $q$-statistic   | 0.58     | 0.42       | 0.33     | 0.65           | 0.41   |
| $p$-value       | 0.0000   | 0.0000     | 0.000002 | 0.0000         | 0.0000 |
| Number of strata| 4        | 3          | 3        | 2              | 4      |

| Table 2. $T_{ai}$ average in each stratum computed by Geodetector software in ‘Risk Detector’ worksheet (°C). |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Factor          | Agriculture | Low forest | Bare land and poor pasture | Urban |
| Land-use type   | 19.5 (°C)    | 23.3 (°C)   | 26.8 (°C)         | 24.4 (°C) |
| Elevation (m)   | 400–1000      | 1000–1600    | 1600–2300         |       |
|                 | 28.1 (°C)    | 24.4 (°C)   | 19.4 (°C)         |       |
|                 | 19.4 (°C)    | 25.6 (°C)   | 27.2 (°C)         |       |
| Latitude        | North       | Centre      | South             |       |
| Day and Night   | Day         | Night       |                   |       |
|                 | 35.8 (°C)    | 11.0 (°C)   |                   |       |
| Season          | Spring      | Summer      | Fall              | Winter |
|                 | 26.6 (°C)    | 32.4 (°C)   | 16.3 (°C)         | 9.4 (°C) |
These classes are named the North, Centre, and South class (Figure 1). Within the latter grouping, stations with severe arid climate type are placed in the South class (in Southern Fars). The North class (in Northern Fars) consists of stations with semi-arid climate type; the Centre (in Central Fars) class consists of stations with arid climate type. In this classification, the MODIS LST data and observed $T_{ai}$ correlation were examined based on latitude for daytime and nighttime (hereafter, $L_{day}$ and $L_{night}$ methods). Also, based on land cover type, stations are grouped into four land cover classes: agriculture, low forest, urban, and bare land and poor pasture (Figure 1). Fars Province includes 21.2%, 7.2%, 18.6%, 51.3%, and 1.7% agriculture, low forest, urban, bare land and poor pasture, and other (salt land and rock), respectively. There is no station located in salt land or rock land cover type. The MODIS LST and observed $T_{ai}$ correlation in this classification are named as $C_{day}$ and $C_{night}$ method in this study.

The scatter plots showed the linear regression as the most appropriate one in all LST–$T_{ai}$ correlations ($T_{ai}$ day = $aLST$ day + $b$ and $T_{ai}$ night = $aLST$ night + $b$).

The estimated $T_{ai}$ in different classes was evaluated in daily as well as seasonal timescale for daytime and nighttime separately to show the effect of diurnal and seasonal differences on $T_{ai}$ estimation.

Figure 3 shows the best fitted regression between LST and observed $T_{ai}$ data for all 16 automated weather stations for daytime ($A_{day}$). Figure 4 shows the best daytime fitted regressions between LST and observed $T_{ai}$ data in three elevation classes ($E_{day}$).

### 2.3. MODIS temperature profile data

The MODIS Atmospheric Profile data include temperature profile data (Retrieved_Temperature_Profile) for 20 geopotential heights (Retrieved_Height_Profile), from 5 to 1000 hPa (Sobrino et al. 2014), evaluated the MODIS MOD07 daytime and nighttime products (from Terra) over the Iberian Peninsula during the decade from 2000 to 2010 using nine radiosonde stations. The validation provided satisfactory results, with bias around 1 K on average and standard deviation between 2 and 3 K for air temperature at different pressure levels.

![Figure 3. Regression plot between all stations’ LST data at the time of satellite over pass and observed $T_{ai}$ in daytime.](image-url)
Radiosonde data at 00:00 UTC was gathered from the only available station in Fars Province in Shiraz station to investigate the accuracy of the satellites (Aqua and Terra) temperature profile data at nighttime from 1000 to 620 hpa pressure level (the layers which are important in this study). The results show large errors in comparison with Sobrino et al.’s 2014 study for Terra (which was expected). Greater level of accuracy could be obtained in the validation process for Terra if the measured data used in the validation process were taken at the instantaneous time at the Terra pass over the Fars Province (not with differences in time). Aqua shows much better accuracy with RMSE around 1.4–2.3 K, especially when the satellite overpass was later than 23:00 UTC. Seasonal average errors for Aqua satellite data show under estimate of air temperature in higher pressure levels (1000, 900, 850 hpa) in wintertime, while the lowest errors have been seen in summertime. This evaluation is not reliable for Terra satellite data, due to the long-time difference between measured radiosonde data and satellite overpass. So overall, despite the limitations of the measured data in the evaluation process, it can be concluded that MODIS atmospheric temperature profile is reliable from both Terra (Sobrino et al. 2014) and Aqua (current study investigate) satellites.

The adiabatic lapse rate phenomena occur between 1000 and 620 hPa (Rhee and Jungho 2014). In this study air temperature data were obtained based on the linear interpolation/extrapolation, using the geopotential height and air temperature data of 620 hPa and 1000 hPa and the elevation (Rhee and Jungho 2014) for daytime and nighttime, separately. The linear relationship between the geopotential height and air temperature in the lower levels is assumed in Equation (1), where $z$ is geopotential

![Figure 4. Regression plots between LST data and observed $T_{ai}$ data in three classes based on stations’ elevation for daytime ((a) < 1000 m, (b) 1000–1600 m, and (c) 1600–2300 m ASL).](image-url)
height. Slope $a$ and intercept $b$ may be obtained using data at 620 hPa and 1000 hPa levels (Equations (2) and (3)), where $T_{a1000}$ hPa, $T_{a620}$ hPa, $z_{1000}$ hPa, and $z_{620}$ hPa are instantaneous air temperature at 1000 hPa level, instantaneous air temperature at 620 hPa level, geopotential height at 1000 hPa level, and geopotential height at 620 hPa level, respectively. The $T_{ai}$ can then be calculated using the height above mean sea level (AMSL), which is elevation AMSL plus 2 m (Jocik Mendez 2004). This method is named as $P_{1\text{day}}$ and $P_{1\text{night}}$ in this study for daytime and nighttime $T_{ai}$ estimation, respectively:

$$T = a \cdot z + b,$$

$$a = \frac{(T_{1000} - T_{620})/z_{1000} - z_{620}}{z_{1000} - z_{620}},$$

$$b = T_{1000} - z_{1000} \cdot a.$$  

Moreover, a hydrostatic atmosphere was assumed to extrapolate $T_{ai}$ using MODIS MOD07 and MYD07 products. The hydrostatic atmospheric assumption can be expressed as Equation (4), as indicated by (Tang and Li 2008)

$$\frac{dP}{dz} = \frac{(P_S - P_L)}{\Delta z} = -\rho g,$$

where $dP/dz$ is the changes in pressure while the height changes, $g$ is the standard gravity, $\rho$ is the density, and $P_L$ is the lowest pressure level of the MODIS atmospheric profile measurement, while $P_S$ is the surface pressure level obtained from the MODIS data. The ambient lapse rate is assumed to be $-6.5 \text{ Kkm}^{-1}$ (Cosgrove et al. 2003) and can be used to relate instantaneous temperature at the lowest pressure level, $T_{aL}$, and instantaneous near-surface temperature, $T_{aS}$ (Equation 5). Combining Equation (4) and Equation (5) and rearranging the terms, instantaneous near-surface air temperature can be estimated as Equation (6). Bisht and Bras (2011) developed this method for estimating instantaneous $T_{ai}$ in solar net radiation calculation at the time of satellite overpass. This method is named as $P_{2\text{day}}$ and $P_{2\text{night}}$ in this article for daytime and nighttime $T_{ai}$ estimation, respectively.

$$T_{aL} - T_{aS} \Delta z = -6.5\text{ Kkm}^{-1},$$

$$T_{aS} = T_{aL} - 6.5\text{ Kkm}^{-1} - \rho g (P_S - P_L).$$

If the lowest 1000 hPa layer was not accessible for the study area, the nearest lower layer was used in its place in both methods. Since the studied area is located in a high elevation area, the 1000 hPa pressure level is not available except in some areas in southern Fars. For missing 1000 hPa level data, 950 hPa level or less pressure level data were used instead. Elevation, latitude, and land cover classifications in Section 2.2 are also mentioned in $P_{1}$ and $P_{2}$ methods validation (Section 3).

### 3. Results and discussions

Measured data from the only 16 AWS are used to validate the estimated air temperature from remote sensing method using MODIS LST and temperature profile data.

Validation processes are performed through statistical criteria, such as the coefficient of determination, root mean square error, and mean bias error (MBE). The
estimated air temperature at the time the satellite passes was compared to data from the meteorological stations. In this case, the error values, such as MBE and RMSE, become very important criteria. The validations are performed in daily and seasonal timescale.

3.1. Daily validation

Considering $R^2$, RMSE, and MBE value in Table 3, $P_1$ and $P_2$ methods show a better agreement with observed data ($P_2$ method seems to be slightly more accurate than $P_1$ method) than all LST methods. Generally, temperature profile data and LST data explain 81–86% and 68–78% variability, respectively. This difference is possibly due to the larger temperature variation of daily LST data than the atmospheric temperature profile (Rhee and Jungho 2014). Both LST and atmospheric temperature profile methods show a relatively better $T_{ai}$ agreement with observed data at nighttime rather than daytime. Vancutsem et al. (2010) explained that higher stability of atmospheric temperature profile and especially LST products at nighttime due to lack of solar radiation and its effect on the thermal infrared signal for LST is the reason of this difference. Moreover, angular anisotropy has a larger effect on LST data accuracy during daytime, due to mixed sunlit and shaded areas within a pixel, and its impact depends on the land cover structure (Prigent, Aires, and Rosswow 2003). Mostovoy et al. (2006) study showed that the difference between LST and estimated $T_{ai}$ increases with view angle from 0° to ±65°. The error increased by 1.0–2.0° with increased view angle in Mostovoy et al. (2006) study.

Table 3 shows that in using LST in $T_{ai}$ estimation, the errors vary by land cover and elevation classification, while only slight variation is noted by latitude classification. Land cover has a potential impact on $T_{ai}$ estimation due to its effect on land emissivity. Rhee and Jungho (2014) explained that during the daytime lower LST have been seen in agricultural land cover due to the cooling effect of vegetation on LST. Larger error value is shown in Table 3 for daytime agricultural regions. Mountain regions (north class and some centre regions) also showed a larger error for both daytime and nighttime using LST data, unlike temperature profile method. The higher accuracy of $P_1$ and $P_2$ methods in altitude regions can be explained by the fact that in north and centre classes the 1000 or even 950 hPa layers are rarely possible where the air temperature values are more

<table>
<thead>
<tr>
<th>Method</th>
<th>$R^2$</th>
<th>MBE (°C)</th>
<th>RMSE (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{day}$</td>
<td>0.71</td>
<td>4.7</td>
<td>4.9</td>
</tr>
<tr>
<td>$A_{night}$</td>
<td>0.73</td>
<td>4.5</td>
<td>4.7</td>
</tr>
<tr>
<td>$E_{day}$ (150–1000, 1000–1600, 1600–2300)</td>
<td>0.80; 0.77; 0.74</td>
<td>3.5; 3.6; 4.0</td>
<td>3.8; 3.9; 4.3</td>
</tr>
<tr>
<td>$E_{night}$ (150–1000, 1000–1600, 1600–2300)</td>
<td>0.79; 0.78; 0.76</td>
<td>3.2; 3.4; 3.8</td>
<td>3.5; 3.6; 3.9</td>
</tr>
<tr>
<td>$L_{day}$ (North, Centre, South)</td>
<td>0.70; 0.77; 0.71</td>
<td>4.2; 3.7; 3.9</td>
<td>4.6; 3.8; 3.8</td>
</tr>
<tr>
<td>$L_{night}$ (North, Centre, South)</td>
<td>0.71; 0.76; 0.73</td>
<td>4.1; 3.5; 3.6</td>
<td>4.4; 3.7; 3.6</td>
</tr>
<tr>
<td>$C_{day}$ (agriculture, low forest, bare land and urban)</td>
<td>0.71; 0.73; 0.76; 0.67</td>
<td>4.0; 3.6; 3.1; 4.2</td>
<td>3.9; 3.8; 3.3; 4.6</td>
</tr>
<tr>
<td>$C_{night}$ (agriculture, low forest, bare land and urban)</td>
<td>0.75; 0.74; 0.78; 0.68</td>
<td>3.5; 3.4; 3.0; 4.1</td>
<td>3.5; 3.7; 3.2; 4.4</td>
</tr>
<tr>
<td>$P_1_{day}$</td>
<td>0.83</td>
<td>2.8</td>
<td>3.3</td>
</tr>
<tr>
<td>$P_1_{night}$</td>
<td>0.86</td>
<td>2.6</td>
<td>3.2</td>
</tr>
<tr>
<td>$P_2_{day}$</td>
<td>0.81</td>
<td>2.5</td>
<td>3.1</td>
</tr>
<tr>
<td>$P_2_{night}$</td>
<td>0.85</td>
<td>2.4</td>
<td>2.9</td>
</tr>
</tbody>
</table>
unstable. The measured surface pressure in high-elevation stations showed the range of 735–890 hPa in 2009.

The results show a larger error in the north of Fars Province when using LST data, due to large urban areas (land-cover type) and mountainous regions (1600–2300 m class in elevation classification) in comparison to centre and south; however the uncertainty in the south and southeast of the study area is more variable due to lower station density. Rhee and Jungho (2014) explained that the larger error of LST in urban areas for estimating $T_{ai}$ is related to high percentage of impervious areas, which exacerbate the discrepancy between the surface and air temperature, especially during the daytime. In this case, the error values in $T_{ai}$ estimation are smaller in $P1$ and $P2$ methods than the LST method. LST product seems to be more sensitive to land cover and the time of satellite overpass, especially during daytime, however, estimated $T_{ai}$ using atmospheric temperature profile seems to be sensitive to elevation due to surface pressure variations and atmosphere stability in higher layers.

Overall, hydrostatic assumption in estimating $T_{ai}$ using MODIS vertical temperature profile ($P2$) seems to be more accurate than linear extrapolation ($P1$), and both these methods seem to be more suitable than LST method, especially during daytime for the study area.

### 3.2. Seasonal validation

The error estimation in seasonal timescale shows that using atmospheric temperature profile data for $T_{ai}$ estimation has a better agreement with observation data in spring and summer (Table 4). This can have particular relevance on air temperature inversion in

<table>
<thead>
<tr>
<th>Method</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{day}$</td>
<td>5.5</td>
<td>5.3</td>
<td>4.5</td>
<td>4.3</td>
</tr>
<tr>
<td>$A_{night}$</td>
<td>5.0</td>
<td>5.1</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>$E_{day}$</td>
<td>150–1000 (m)</td>
<td>3.5</td>
<td>4.2</td>
<td>3.1</td>
</tr>
<tr>
<td>$E_{night}$</td>
<td>1000–1600 (m)</td>
<td>3.8</td>
<td>4.3</td>
<td>3.1</td>
</tr>
<tr>
<td>$L_{day}$</td>
<td>1600–2300 (m)</td>
<td>4.3</td>
<td>4.7</td>
<td>3.3</td>
</tr>
<tr>
<td>$L_{night}$</td>
<td>150–1000 (m)</td>
<td>3.6</td>
<td>3.8</td>
<td>3.0</td>
</tr>
<tr>
<td>$C_{day}$</td>
<td>1000–1600 (m)</td>
<td>3.8</td>
<td>4.0</td>
<td>2.9</td>
</tr>
<tr>
<td>$C_{night}$</td>
<td>1600–2300 (m)</td>
<td>4.0</td>
<td>4.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Agriculture</td>
<td>4.5</td>
<td>4.9</td>
<td>4.1</td>
<td>3.7</td>
</tr>
<tr>
<td>Low forest</td>
<td>4.4</td>
<td>4.2</td>
<td>3.8</td>
<td>3.5</td>
</tr>
<tr>
<td>Bare land</td>
<td>3.9</td>
<td>4.1</td>
<td>3.6</td>
<td>3.5</td>
</tr>
<tr>
<td>Urban</td>
<td>4.0</td>
<td>4.5</td>
<td>3.8</td>
<td>3.5</td>
</tr>
<tr>
<td>$P1_{day}$</td>
<td>3.6</td>
<td>4.0</td>
<td>3.6</td>
<td>3.4</td>
</tr>
<tr>
<td>$P1_{night}$</td>
<td>3.8</td>
<td>4.2</td>
<td>3.7</td>
<td>3.4</td>
</tr>
<tr>
<td>$P2_{day}$</td>
<td>4.2</td>
<td>4.4</td>
<td>3.9</td>
<td>3.5</td>
</tr>
<tr>
<td>$P2_{night}$</td>
<td>3.8</td>
<td>4.2</td>
<td>3.7</td>
<td>3.4</td>
</tr>
</tbody>
</table>

Table 4. RMSE values in $T_{ai}$ estimation in seasonal timescale (°C).
fall and winter and more uncertainty in extrapolating \((\text{P1 method})\) and hydrostatic assumption methods \((\text{P2 method})\) for estimating \(T_{ai}\) in these seasons. The measured temperature profile data from Shiraz station (at 00:00 UTC) shows that throughout the winter there are frequent surface temperature inversions (only six days did not show this). These inversions extend from the surface to maximum heights of 225, 270, and 203 m in January, February, and March, respectively. The evaluation process in Section 2.3 showed the same results for cold seasons. However, using LST data shows better results than \(\text{P1 and P2 methods}\) in fall and winter. This result is explained by the high variation of LST in warm seasons in LST method as mentioned in Section 3.1 Fabiola Flores and Mario Lillo (2010) reported that coastal areas raise the value of RMSE on a regional level in the warmest months using linear extrapolation method, while these areas show the lowest errors for the coldest months. They related this result to higher cloud fraction and higher relative humidity in the warmest months in Chile which caused difficulties in extracting data from MODIS images. However, unlike Chile, the cloud fraction and relative humidity are higher in the coldest months in Fars.

Benali et al.’s (2012) study showed higher accuracy of daily average air temperature, using LST data in summer. They explained that the higher accuracy of daily average air temperature estimations in the summer is due to the higher accuracy of the LST product, when clear-sky days are more common, probably. Their result does not agree with the current study and Rhee and Jungho (2014) research. That can refer to the dry climate type of the study area (with rare cloudy days during the year, especially in recent decades) which is different from Benali et al.’s (2012) study area. They suggested that cloud cover had an inverse relationship with model performance. Cloud contamination decreases the LST value leading to lower LST–\(T_{ai}\) differences.

Estimation errors of air temperature from \(\text{P1, P2, and especially LST}\) can vary by elevation, latitude, and land cover classes. Differences are observed between land cover classes for LST method during spring, summer, and fall (Table 4) and \(\text{P1 and P2 methods}\) during summer and fall (not shown). The station-average MBE and RMSE in different land cover classes show that the error value in LST method is larger in stations with urban land cover class, especially in summer and spring. It is likely that the error value in \(T_{ai}\) estimation is smaller in urban area using \(\text{P1 and P2 methods}\). This result can be explained by the same reason as mentioned in Section 3.1. Lee et al. (2011) finding showed that land cover can also be especially important for dry regions (like the study area), where winter cereals and pastures typically have a strong contrast in vegetation cover between wet and dry seasons. Table 4 shows that during spring, summer and fall the cooling effect of vegetation on LST causes larger error values in \(T_{ai}\) estimation during daytime.

For elevation classes no increasing and decreasing error were observed in seasonal scale for LST method, except in summertime in high elevation classes. The latitude classes (like elevation classes) show no increasing and decreasing error in seasonal scale, but a larger error value in the north of Fars was due to higher elevation and larger urban area in this class. The inter-annual pattern varies from cold to mild winter and mild to warm summer from north to south in three latitude classes. In temperature profile method, elevation classes show fewer errors in summer and higher elevation classes. The sensitivity of \(\text{P1 and P2 method}\) to land cover type does not seem to be significant as it is shown in this research.

Comparatively, as a sample Figure 5 shows the results of daily \(T_{ai}\) estimation methods in Shiraz station in daytime. Shiraz station is classified in urban land cover class in centre
latitude class with 1484 m elevation ASL. Considering station land cover, atmosphere temperature profile data in estimating \( T_a \) for the studied area shows the smallest error.

### 3.3. Sources of uncertainty

The accuracy of the LST and temperature profile has a significant effect on estimated \( T_{ai} \). Uncertainty sources in LST retrieval can be addressed as radiometric noise, surface emissivity, atmospheric contribution, and sensor view angle. Emissivity is one of the largest uncertainty sources in LST production. Benali et al. (2012) reported that the difference of 0.01 in emissivity can cause a 2.0°C error in the LST retrieval. The MODIS emissivity is calculated using the MODIS land cover. Wan (1997) showed that emissivity is a function of soil and vegetation conditions.

View angle is another possible uncertainty source for LST retrieval. However, the view-angle effect may be ignored when the angle is less than 45° (Wan 1997). Similarly they reported that view angle effect can be reduced when averaged over several pixels. Prigent, Aires, and Rossow (2003) showed that angular anisotropy has a larger effect during daytime, due to mixed sunlit and shaded areas within a pixel, and its impact depends on the land cover structure. During the day, errors of MODIS LST estimations increase with increasing satellite viewing angles (Prigent, Aires, and Rossow 2003). This can be another reason for the higher accuracy of \( T_{ai} \) estimation at nighttime using LST data.

Several sources of errors must be addressed in MODIS temperature profile inaccuracy. Surface uncertainties such as surface elevation, emissivity, and skin temperature as well as atmospheric transmittance calculation error and cloud detection are some sources of error in this product (Seemann et al. 2003). Further work to improve MODIS temperature profile and MODIS LST product will include improving surface emissivity and land cover. A more accurate skin temperature and cloud mask product should also be investigated in improving the MODIS temperature profile.

![Figure 5. The daytime comparison of estimated \( T_{ai} \) values against observed data.](image)
4. Conclusion

Air temperature data can be used for many environmental and agricultural studies, such as surface net radiation estimation. In this study, LST and atmospheric temperature profile data from MODIS were used in estimating high spatial resolution instantaneous air temperature data for Fars province (Iran) with limited *in situ* data.

LST and vertical temperature profile data showed acceptable accuracy in estimating daily and seasonal air temperature (with temperature profile method superior to LST method). They showed quite reasonable coefficient of determination values and the results were within the accuracies reported in the literature considering the methodologies for $T_{\text{air}}$ estimation based on remote sensing data. Daily and seasonal $P1$ and $P2$ outperformed LST in terms of error values, especially in urban regions and high elevation regions, but LST showed higher accuracy during fall and winter, especially at nighttime. However, the relation between LST and $T_{\text{air}}$ varies with local conditions, because the LST is influenced by land cover, elevation, and vegetation.

Further developments will consider mainly two topics including cloudy days and soil moisture effect on air temperature. These factors were eliminated in this study due to lack of data and, thus, clear sky days with no precipitation in three previous days were selected as studied days.

In conclusion, the relatively low spatial resolution of measured $T_{\text{air}}$ data can be improved using remote sensing methods. The current study results can be extended to areas with limited *in situ* data due to complex topography and a wide range of temperature.

Acknowledgements

This study was a part of the PhD student sabbatical work done in CUNY/CREST Institute and Bu-Ali Sina University, Hamedan, Iran. The authors also acknowledge the kind support of NOAA/CREST at City University of New York.

Disclosure statement

No potential conflict of interest was reported by the authors.

References


