


Utilization of the Google Earth Engine for the evaluation of daily soil temperature derived from Global Land Data Assimilation System in two different depths over a semiarid region

Abolghasem Akbari^{1,2}  | Majid Rajabi Jaghargh³ | Azizan Abu Samah⁴ | Jonathan Peter Cox⁵ | Mojtaba Gholamzadeh⁶ | Alireza Araghi³ | Patricia M. Saco⁷ | Khabat Khosravi⁸

¹Department of Civil Engineering, Khavaran Institute of Higher Education, Mashhad, Iran

²Research & development unit, Sayyal Samen Co., Mashhad, Iran

³Department of Water Science and Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

⁴Institute of Ocean and Earth Sciences, University Malaya, Kuala Lumpur, Malaysia

⁵Caribbean Institute for Meteorology and Hydrology, Bridgetown, Barbados

⁶Islamic Azad University, Tehran, Iran

⁷School of Engineering and Centre for Water Security and Environmental Sustainability, The University of Newcastle, Callaghan, New South Wales, Australia

⁸School of climate change and adaptation, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada

Correspondence

Abolghasem Akbari, Department of Civil Engineering, Khavaran Institute of Higher Education, Mashhad, Iran.

Email: akbariinbox@yahoo.com

Abstract

The Google Earth Engine (GEE) was used to investigate the performance of the Global Land Data Assimilation System (GLDAS) soil temperature (ST) data against observed ST from 13 synoptic stations over a semiarid region in Iran. Three-hourly ST data were collected and analyzed in two depths (0–10 cm; 40–100 cm) and 5 years. In each depth, GLDAS-Noah ST data were evaluated for daily minimum, maximum, and average ST (i.e., T_{\min} , T_{\max} , and T_{avg}). Based on the correlation coefficient, Kling–Gupta Efficiency, and Nash–Sutcliffe Efficiency the overall performance of the GLDAS-Noah is 0.96, 0.66, and 0.79 for T_{\min} ; 0.97, 0.84, and 0.89 for T_{avg} ; and 0.95, 0.89, and 0.89 for T_{\max} , respectively in the first layer. Likewise, 0.97, 0.85, and 0.86 for T_{\min} ; 0.97, 0.77, and 0.80 for T_{avg} ; and 0.97, 0.69, and 0.69 for T_{\max} are obtained in the second layer. However, there is a significant negative bias which tends to underestimate ST in the two investigated layers, given by an average bias over all the stations analyzed of -24% , -12% , and -5% for T_{\min} , T_{avg} , and T_{\max} in the first layer, and average bias of -8% , -13% , and -17% for T_{\min} , T_{avg} , and T_{\max} in the second layer. This study reveals that GLDAS-Noah-derived ST can be used in arid regions where little or no observation data is available. Moreover, GEE performed as an advanced geospatial processing tool in regional scale analysis of ST in different layers.

KEYWORDS

GLDAS-Noah, Google Earth Engine, semiarid region, soil temperature, synoptic station

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

1 | INTRODUCTION

Soil temperature (ST) is essential in energy balance applications such as land surface modeling (LSM), numerical weather forecasting, agricultural activities, and long-term climate prediction. ST is also required for estimating most belowground ecosystem processes, including root growth and respiration, decomposition, and nitrogen mineralization. In addition, ST is an essential factor in agriculture and treatment of organic waste due to its close relationship to the growth of biological organisms (Ozgener et al., 2013). ST is similarly a key player in natural heating/cooling buildings systems, such as ground source heat pumps, earth-air heat exchangers for agricultural buildings as well as for agricultural greenhouse systems (Faridi et al., 2019; Mihalakakou, 2002). According to Cohen and Fielding (1979), another significant application of ST is to predict the frost depth of underground pipelines. The depth of ST measurement depends on the specific application (Holmes et al., 2008).

ST plays a key control on the transport and fate of contaminants in soils. However, unlike other hydro-meteorological parameters such as air temperature and precipitation, ST is rarely measured regularly. In many regions, historical ST databases have been compiled from meteorological networks but have limited spatial resolution. For instance, only 13 synoptic stations record the ST at standard depths in the Khorasan Razavi province. Given the large area of Khorasan Razavi (128,420 sq. km), the in situ measurements are inadequate to accurately depict the spatial variability of the minimum, maximum, and average ST. Similar issues regarding data scarcity have been reported by Li et al. (2020) in other regions. The main reasons for the lack of spatial observations are that the measurement of ST at different depths is expensive and often difficult (Napagoda & Tilakar-atne, 2012).

As soil temperature strongly depends on meteorological variables (i.e., solar radiation, air humidity, atmospheric pressure, temperature, precipitation, sunshine hours, and wind speed), several methods have been developed based on these input variables (Chow et al., 2011; Li et al., 2020; Ozgener et al., 2013), including empirical models (Bond-Lamberty et al., 2005), analytical models (Badache et al., 2016), numerical models (Gao et al., 2016), machine learning algorithms (Zeynoddin et al., 2020), and methods based on remote sensing technology. Empirical equations do not have reasonable prediction power due to their simple structure, while analytical and numerical models are challenging to use because of their complexity and vast data demand (Xing et al., 2018). Given the growing availability of earth observation technologies, the space-based remote sensing

methods have the potential to offer a very significant contribution to large-scale monitoring of ST (WMO, 2018). A simple, physically based model that can calculate relatively accurate near-surface (0–10 cm, 40–100 cm) ST profiles from a single observation can be a valuable tool for many environmental modeling applications. As reported by Holmes et al. (2008), earth observation technologies such as MODIS, AMSR-E, ENVISAT-AATSR, SSM/I, and TRMM-TMI provide LST products over most of the earth.

Despite the existence of many studies on ST prediction, the analysis of ST data derived from the Global Land Data Assimilation System (GLDAS) in Iran is still rare. GLDAS is a joint effort by the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), and the National Centers for Environmental Prediction (NCEP) (Rodell et al., 2004). It ingests satellite and ground-based observational data products using advanced land surface modeling (LSM) and data assimilation techniques to generate optimal fields of land surface states and fluxes (Rodell et al., 2004). The GLDAS drives multiple offline LSMs and integrates a vast quantity of observation-based data. Furthermore, it is implemented globally at high spatial resolutions (1 to 2.5 km), producing results in near-real-time (NASA, 2021). High-quality global land surface simulation is achieved using land surface models while minimizing errors by directing the models with observation-based data and satellite land data assimilation techniques. GLDAS is a joint effort by NASA, NOAA, and NCEP (Rodell et al., 2004). According to Rui and Beaudoin (2018), the GLDAS performs based on four LSM approaches, including the Community Land Model (Dai et al., 2003), Variable Infiltration Capacity (VIC) (Liang, 1994), Noah (Ek et al., 2003), and MOSAIC (Koster & Suarez, 1996).

In recent years, machine learning models have been applied by some researchers to predict soil moisture and ST at different depths. For instance, Afandi et al. (2022) applied Artificial Neural Network (ANN) for the prediction of temperature at 2-m depth, Li et al. (2022) used Random Forest (RF) and Support Vector Regression (SVR) for the prediction of soil moisture and ST, while Samadianfard, Ghorbani, and Mohammadi (2018) applied a hybrid ANN for forecasting soil temperature at multiple-depths and later applied Wavelet neural networks and gene expression programming models for prediction of short-term ST at different depths (Samadianfard, Asadi, et al., 2018). In the same line of research, Bayatvarkeshi et al. (2021) used complementary machine learning models for the prediction of ST from air temperature features in diverse climatic conditions. Datta and Faroughi (2023) combined a structural

equation model and an ANN for the estimation of surface soil moisture. Wang et al. (2023) constructed a structural equation model coupled with an artificial neural network for soil moisture inversion. Last but not least, Shamshirband et al. (2020) applied support vector machines and multilayer perceptron for predicting ST at different depths. It should be noted that machine learning models calibrated for a given study area cannot be used for other regions without recalibration, which can limit applicability.

Though previous work on applications and analysis of GLDAS-based ST has been carried out over parts of Asia (Wang, Cui, et al., 2016) and over South Asia (Yang et al., 2020; Yang & Zhang, 2018), none of the selected observational networks in these studies were located in the Middle East and West Asia. Therefore, this is the first study on the validation of GLDAS-based ST over the Middle East, with a specific focus on semiarid areas in Iran. Due to the unique properties of GLDAS data at the global scale, the accuracy of the data needs to be evaluated at local and regional scales before its potential use in future practical applications. In recent years, there has been an increasing interest in using cloud computing tools and reanalysis of data. In this study, we utilize the Google Earth Engine (GEE) platform to evaluate the accuracy of the GLDAS-based ST reanalysis dataset against the in situ measurements across a significant agricultural zone in northeastern Iran. The GEE combines a multi-petabyte catalog of satellite imagery and reanalyzed data with planetary-scale analysis capabilities. It has been used to detect changes, map trends, and quantify differences on the Earth's surface. This research benefits from the use of GEE because of its data coverage for this large region, extensive reanalyzed data such as GLDAS products generated from different LSMs, high speed, intensive processing capacity, advanced raster processing tools. GEE is now available free for academic and research use.

2 | STUDY AREA

The Khorasan Razavi province is located in the northeast of Iran in central Asia and extends within the 33–38° N and 56–62° E, covering a total area of 117,200 km². Agriculture is one of the main pillars of economic development in this region and considering the importance of ST on plant growth and photosynthesis process, it is worth using the GLDAS products to determine the temporal and spatial variation of ST in this region. The soil texture in this region is classified as loam, silt loam, silty clay loam, clay loam, and sandy loam. More than 90% of agricultural soils in Khorasan-Razavi province are classified in these five texture classes. The mean annual rainfall is

approximately 250 mm, and the average annual temperature is 14°C. Based on the Köppen climate classification, the climate of most of Khorasan Razavi is classified as semiarid or arid (Araghi et al., 2017). The province's highest elevation found at Mount Shirbad in the Binalud mountain range north of Neishabour, has an altitude of 3303 m, which contrasts to the lowest elevation of 235 m, in the Sarakhs plain close to the border of Iran and Turkmenistan, (Ashraf et al., 2011). The location of the study area and the corresponding synoptic stations are shown in Figure 1, and a step-by-step methodology and workflow is presented in Figure 2.

3 | MATERIALS AND METHODS

Three-hourly observed ST data for 5 years, from January 2008 to December 2012, were collected from 13 synoptic stations and analyzed statistically using time series analysis, specifically for the applicability of least squares regression methods. The validity of the linear regression model was carried out using analysis of variance (ANOVA) of observed data in all the stations based on the indicators of Durbin (1971), Durbin and Watson (1992), and tests of normality (Shapiro & Wilk, 1965) using the Statistical Package for the Social Sciences (SPSS) software. Due to the low significance level obtained (less than 5%), the assumption of normality was confirmed with a 95% probability.

Further investigation using residual statistics resulted in a constant variance that fulfils the other condition for acceptance of an ordinary linear regression model. Runs-test for residuals also confirmed that the time series are random and first-order autocorrelation does not exist.

3.1 | Ground-based observation soil temperature data

The ground-based data used in the current study were collected from synoptic stations in Khorasan Razavi province and managed by the Islamic Republic of Iran Meteorological Organization (IRIMO). The data from all the stations are standardized based on the World Meteorological Organization (WMO) guidelines, and ST is recorded at depths of 5, 10, 20, 50, and 100 cm below the ground surface. The ST were measured at 03:00, 09:00, and 15:00 Greenwich Mean Time (GMT) on a daily step. To evaluate the GLDAS-derived ST products, quality checks were conducted for time series data from the 13 synoptic stations to identify possible random or systematic errors. Among the 13 stations listed in Table 1, ST2 and ST3 did

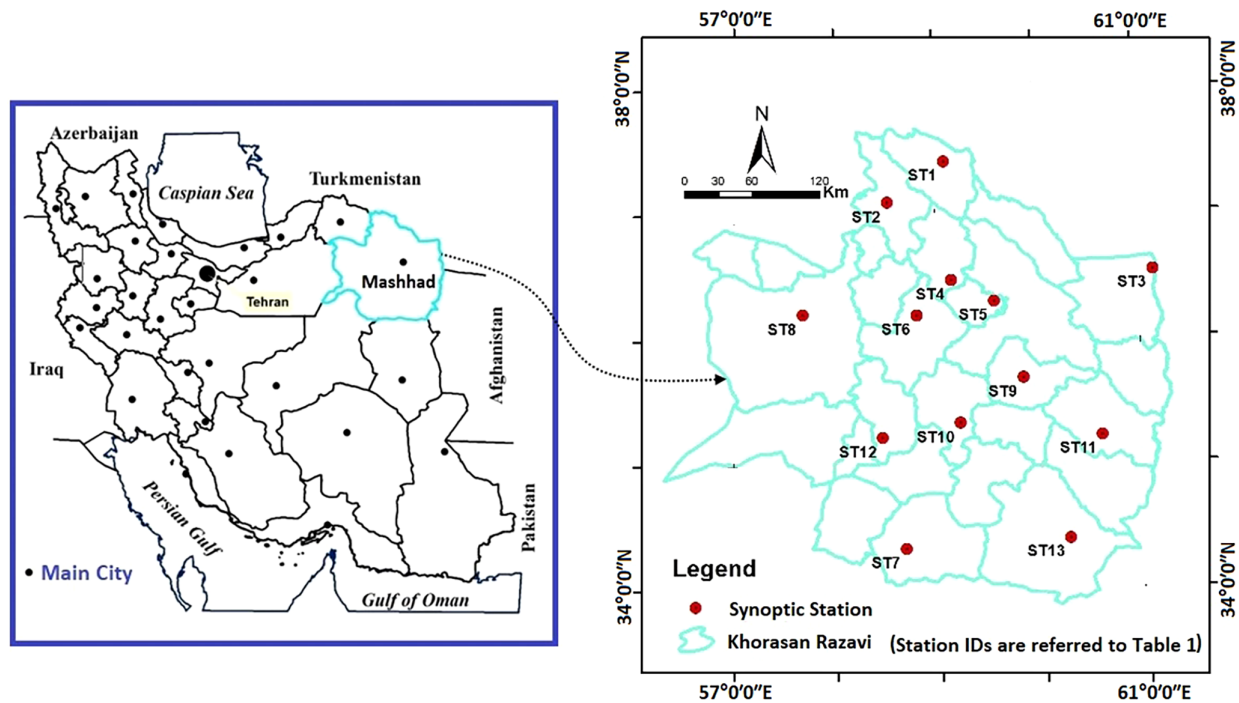


FIGURE 1 Location of the study area and synoptic stations employed for this research.

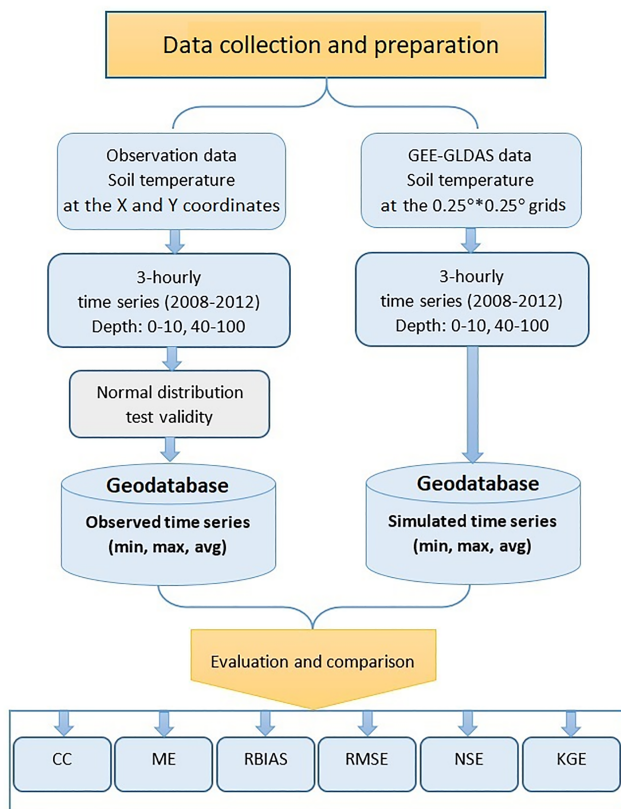


FIGURE 2 Methodology and workflow of this study.

not meet the quality check criteria due to missing records in the time series. Therefore, ST2 and ST3 were excluded from further analysis.

3.2 | Global Land Data Assimilation System-based ST time series

The GLDAS-2.1 dataset is the newest reprocessed dataset covering the period 2000–present (Rui & Beaudoin, 2019), and is archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC). The GLDAS-2.1 is simulated with the Noah model and forced using the global meteorological forcing dataset from Princeton University and observation-based datasets. The GLDAS-2.1 uses the MODIS-based land surface parameter datasets and includes initialization of soil moisture over deserts. GLDAS-2.1 products are provided with two resolutions of $1^{\circ} \times 1^{\circ}$ and $0.25^{\circ} \times 0.25^{\circ}$ and extend from the year 2000 to the present with approximately 1.5-month latency, and are updated monthly.

The ST data from GLDAS were collected for a period of 5 years, from January 2008 to December 2012, using the GEE. The GEE is increasingly recognized as a powerful worldwide platform for remote sensing data and a fast computation and visualization system for global-scale geospatial analysis with an extensive scientific data catalog (Gorelick et al., 2017; Kumar & Mutanga, 2018). The GEE provides tools for collecting, retrieving, and analyzing remotely sensed and processed data from different sources, including GLDAS. Essential characteristics of the GLDAS data are provided in Table 2.

The average layer ST is the depth-averaged temperature beneath the soil surface for that specified layer. The

TABLE 1 Synoptic stations used in this research for the evaluation of soil temperature.

^a ID	Name	Latitude (°N)	Longitude (°E)	Time periods	Elevation (m)
ST1	Dargaz	37.47	59.25	2008–2012	514
ST2	Quchan	37.07	58.50	2008–2010	1287
ST3	Sarakhs	36.53	61.17	2008–2011	235
ST4	Golmakan	36.48	59.28	2008–2012	1176
ST5	Neyshabour	36.27	58.80	2008–2012	1213
ST6	Mashhad	36.23	59.63	2008–2012	999
ST7	Gonabad	34.35	58.68	2008–2012	1056
ST8	Sabzevar	36.20	57.65	2008–2012	962
ST9	Fariman	35.58	59.83	2008–2012	1460
ST10	Torbat Heydarieh	35.33	59.22	2008–2012	1451
ST11	Torbat Jam	35.27	60.58	2008–2012	950
ST12	Kashmar	35.27	58.47	2008–2012	1110
ST13	Khaf	34.58	60.15	2008–2012	998

^aLocation of stations are shown by station ID in Figure 1.

TABLE 2 Basic characteristics of the Global Land Data Assimilation System data.

Contents	Water and energy budget components, forcing data
Latitude extent	−60° to 90° N
Longitude extent	−180° to 180° E
Spatial resolution	0.25°, 1.0°
Temporal resolution	3-hourly or monthly
Temporal coverage	2000 to present for the 0.25° data
Dimension	1440 (longitude) x 600 (latitude) for the 0.25° data
Origin (1st grid center)	(179.875° W, 59.875° S) for the 0.25° data
Land surface models	CLM 2.0 (1.0°) MOSAIC (1.0°) NOAH 2.1 (1.0°) VIC water balance (1.0°) NOAH 2.1 (0.25°)

number of vertical layers for ST in different GLDAS-LSMs is shown in Table 3 (NASA, 2021).

This study evaluated the accuracy of the GLDAS-2.1 Noah 3-hourly 0.25° × 0.25° ST product, named hereafter as GLDAS-Noah, using observed ST.

3.3 | Evaluation approach

The 3-hourly GLDAS-Noah soil temperature data are provided as averaged values for specific soil profile depths,

TABLE 3 Available land surface models and supporting vertical layers in Global Land Data Assimilation System.

Model	Vertical layers
CLM (10 layers)	0–0.018, 0.018–0.045, 0.045–0.091, 0.091–0.166, ..., 2.296–3.433 m.
Mosaic (3 layers)	0–0.02, 0.02–1.50, and 1.5–3.50 m.
Noah (4 layers)	0–0.1, 0.1–0.4, 0.4–1.0, and 1.0–2.0 m.
VIC (3 layers)	0–0.1, 0.1–1.6, and 1.6–1.9 m.

including 0–10, 10–40, 40–100, and 100–200 cm. For the evaluation of the product, the GLDAS-Noah ST values were obtained for each synoptic station using its geographic coordinates. To enable comparison with observations, and given that the GLDAS-Noah ST are averaged soil profile values (i.e., 0–10 cm), the observed ST corresponding to 5 and 10 cm depths, and also 50 and 100 cm, were averaged to generate new time series to compare with the GLDAS-Noah ST at 0–10 and 40–100 cm soil depths respectively. This is based on the approach proposed by Wang et al. (2018). The evaluation was undertaken separately for 03:00, 09:00, and 15:00 GMT, which can be assumed to represent the daily minimum, maximum, and average soil temperature values. Several statistical metrics were employed to evaluate the performance of GLDAS-Noah ST against the observational data.

3.4 | Statistical metrics

The statistical-based metrics used to evaluate the accuracy of the GLDAS ST dataset, include correlation

TABLE 4 Utilized metrics for statistical evaluation of Global Land Data Assimilation System-derived soil temperature in this study.

Statistic indices	Equations	Optimal value
R	$R = \frac{\frac{1}{N} \sum_{i=1}^n (T_{Sta})_i - \overline{(T_{GLDAS})} (\overline{(T_{Sta})})}{\sqrt{\frac{1}{N} \sum_{i=1}^n (T_{GLDAS})_i - \overline{(T_{GLDAS})}^2} \sqrt{\frac{1}{N} \sum_{i=1}^n (T_{Sta})_i - \overline{(T_{Sta})}^2}}$	1
ME	$ME = \frac{1}{N} \sum_{i=1}^n ((T_{GLDAS})_i - (T_{Sta})_i)$	0
RBIAS	$RBIAS = \frac{\sum_{i=1}^n ((T_{GLDAS})_i - (T_{Sta})_i)}{\sum_{i=1}^n (T_{Sta})_i}$	0
RMSE	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n ((T_{GLDAS})_i - (T_{Sta})_i)^2}$	0
KGE	$KGE = 1 - \sqrt{(CC - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$ $\beta = \frac{\overline{T_{GLDAS}}}{\overline{T_{Sta}}} \quad \gamma = \frac{CV_{GLDAS}}{CV_{Sta}}$	1
NSE	$NSE = 1 - \left(\frac{\sum_{i=1}^n ((T_{Sta})_i - (T_{GLDAS})_i)^2}{\sum_{i=1}^n ((T_{Sta})_i - \overline{T_{Sta}})^2} \right)$	1

Note: where T_{GLDAS} refers to predicted ST, T_{Sta} refers to observed ST, and n refers to the number of samples.

coefficient (R), mean error (ME), relative bias (RBIAS), root mean square error (RMSE), Kling–Gupta efficiency (KGE), and Nash–Sutcliffe efficiency (NSE), were used to evaluate the accuracy of the GLDAS ST dataset. The NSE and KGE are normalized statistics that determine the relative magnitude of the residual variance compared to the measured data variance of time series (Nash & Sutcliffe, 1970). They are commonly used to assess the predictive power of hydrological time series. However, it can also be used to quantitatively describe the accuracy of model outputs for other time series than discharge such as nutrient loadings and soil temperature (Tabari et al., 2015). The R is used to assess the strength and direction of the linear relationships between pairs of variables. A higher R value represents more minor differences between the observed data and the fitted values. The RMSE provides an estimate of how far the predicted values deviate, on average, from the actual values in the time series. RMSEs values of 0 indicate a perfect fit and the values less than half the standard deviation of the measured data indicate low reliability and unsatisfactory results (Singh et al., 2004). ME refers to the average of all the errors in a dataset and, hence, it reflects uncertainty in measurements or the difference between the predicted (in this case, the GLDAs values) and the observed values. Likewise, RBIAS also quantitatively estimates the systematic percentage error. From this performance indicator, the direction of uncertainties can be identified, such as whether the simulation is over or underestimated. NSE varies between $-\infty$ and 1. These indices have been

widely used in previous studies (Akbari et al., 2017; Nashwan et al., 2020; Shirmohammadi-Aliakbarkhani & Akbari, 2020; Thiemig et al., 2013). The equations and optimal values for each of the indicators are provided in Table 4. More details and information can be found in Nash and Sutcliffe (1970), Gupta et al. (1999), and Moriasi et al. (2007).

4 | RESULTS AND DISCUSSION

4.1 | GLDAS-Noah ST for the 0–10 cm layer

A linear relationship was obtained between the GLDAS-derived ST and ground observation time series for the two soil layers (0–10 cm, 40–100) and for the maximum (T_{max}), minimum (T_{min}), and average (T_{avg}) values. Scattergram was used to evaluate the accuracy of the GLDAS data visually and presented in Figure 3, in which “Sta” indicates observed ST values and GLDAS corresponds to GLDAS-Noah derived ST values. The number of data in the time series, the correlation coefficient, the linear equation, and the station ID are reported in each Scattergram. The values of the correlation coefficient reflect a robust linear relationship between the two datasets in the first layer (0–10 cm) and correspond to the ranges of 0.92–0.98, 0.90–0.98, and 0.94–0.98 for T_{min} , T_{max} , and T_{avg} , respectively. Results from the Scattergram indicate a high accuracy of the applied model for ST prediction.

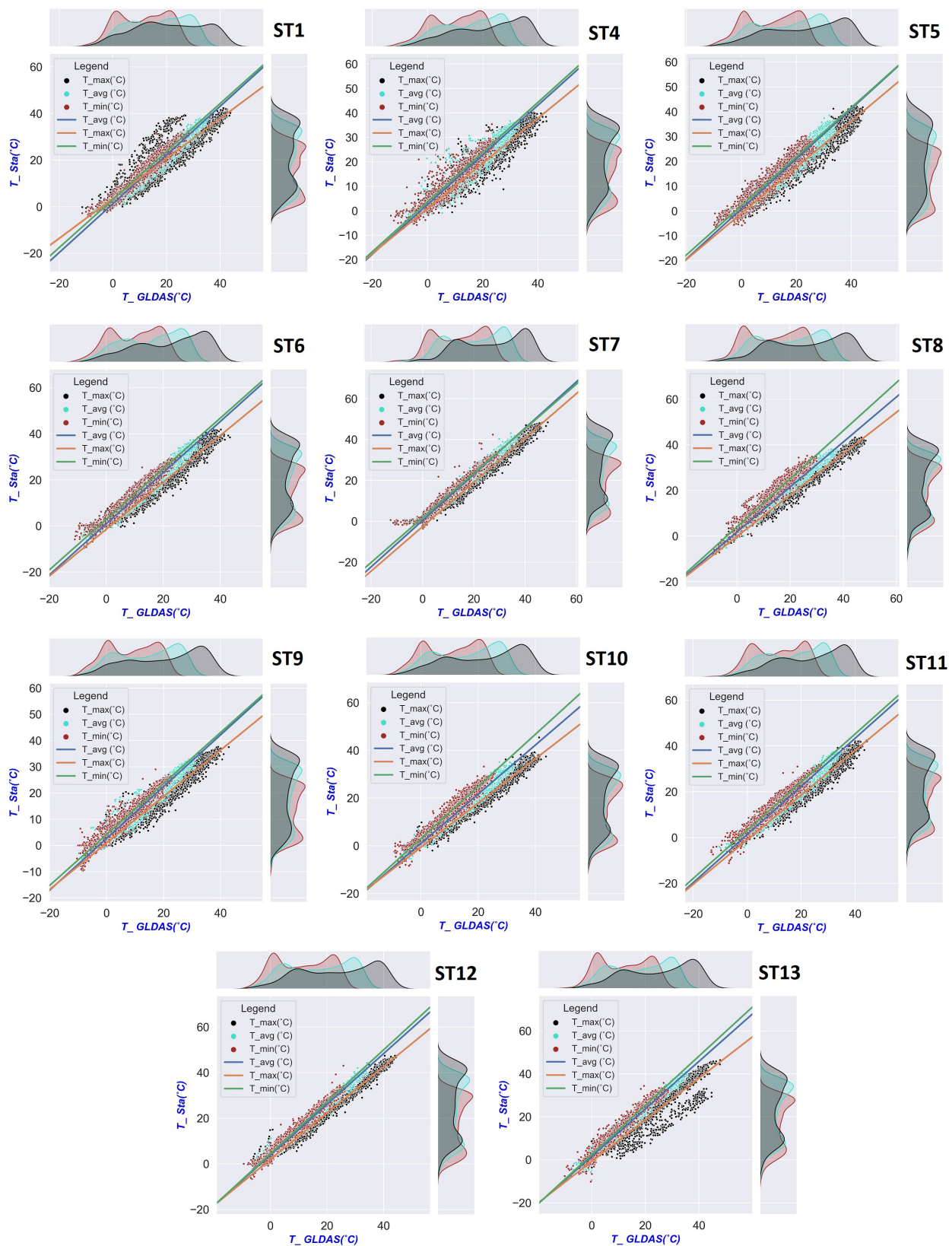


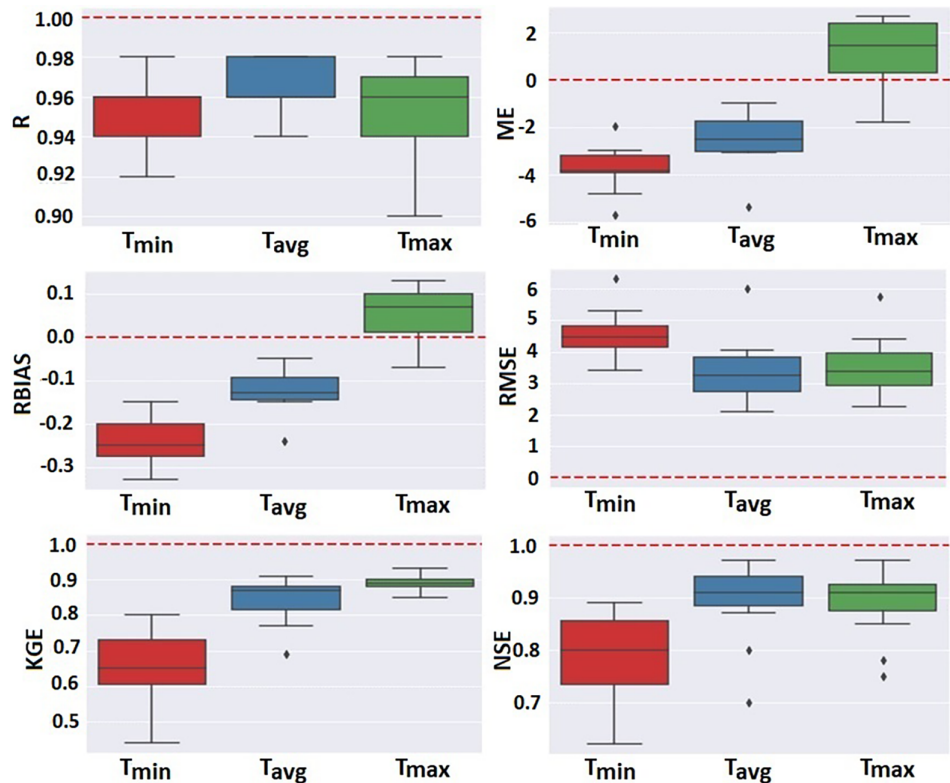
FIGURE 3 Scattergram visually displays the relationship between the observed and Global Land Data Assimilation System-derived soil temperature (T_{avg} , blue; T_{max} , black; and T_{min} brown) for the first layer (0–10 cm) at the stations ST1, ST4–ST6. Scattergram visually displays the relationship between the observed and GLDAS-derived ST (T_{avg} , blue; T_{max} , black; and T_{min} brown) for the first layer (0–10 cm) at the stations ST7–ST12. Scattergram visually displays the relationship between the observed and GLDAS-derived ST (T_{avg} , blue; T_{max} , black; and T_{min} brown) for the first layer (0–10 cm) at the station ST13.

TABLE 5 Statistical metrics used to evaluate the performance of predicted Global Land Data Assimilation System-Noah soil temperature (ST) against the observed ST at the synoptic stations in the first layer (0–10 cm).

No.	R			ME (°C)			RBIAS (%)			RMSE (°C)			KGE			NSE		
	T_{\min}	T_{avg}	T_{\max}	T_{\min}	T_{avg}	T_{\max}	T_{\min}	T_{avg}	T_{\max}	T_{\min}	T_{avg}	T_{\max}	T_{\min}	T_{avg}	T_{\max}	T_{\min}	T_{avg}	T_{\max}
ST1	0.96	0.96	0.96	-3.40	-2.01	-0.53	-0.23	-0.11	-0.02	3.96	2.85	5.74	0.67	0.87	0.91	0.83	0.94	0.78
ST4	0.92	0.94	0.94	-3.82	-2.80	0.79	-0.28	-0.15	0.04	4.77	4.05	3.79	0.59	0.80	0.93	0.75	0.87	0.91
ST5	0.94	0.96	0.94	-1.96	-0.96	2.69	-0.15	-0.05	0.13	3.40	2.60	4.10	0.75	0.91	0.86	0.89	0.95	0.90
ST6	0.96	0.96	0.96	-2.97	-3.01	0.30	-0.17	-0.13	0.01	4.40	3.49	2.78	0.80	0.87	0.90	0.88	0.91	0.97
ST7	0.96	0.98	0.98	-2.97	-3.01	0.30	-0.17	-0.13	0.01	3.61	3.67	2.25	0.80	0.87	0.90	0.88	0.91	0.97
ST8	0.96	0.98	0.98	-4.78	-1.51	2.66	-0.26	-0.07	0.11	5.30	2.11	3.40	0.64	0.89	0.89	0.74	0.97	0.92
ST9	0.94	0.96	0.94	-3.93	-2.49	1.46	-0.32	-0.15	0.08	4.48	3.25	3.39	0.44	0.77	0.90	0.73	0.90	0.91
ST10	0.96	0.98	0.96	-3.88	-1.31	2.57	-0.27	-0.08	0.13	4.55	2.18	3.14	0.62	0.89	0.87	0.80	0.94	0.90
ST11	0.94	0.96	0.94	-3.73	-2.18	1.55	-0.25	-0.11	0.07	4.33	3.03	2.85	0.65	0.87	0.89	0.80	0.92	0.93
ST12	0.98	0.98	0.98	-5.71	-5.35	-1.79	-0.33	-0.24	-0.07	6.34	6.00	3.01	0.54	0.69	0.89	0.62	0.70	0.85
ST13	0.98	0.98	0.90	-3.86	-3.06	2.17	-0.24	-0.14	0.09	4.86	3.98	4.40	0.71	0.83	0.85	0.73	0.80	0.75
Avg	0.96	0.97	0.95	-3.73	-2.52	1.11	-0.24	-0.12	0.05	4.55	3.38	3.53	0.66	0.84	0.89	0.79	0.89	0.89
Max	0.98	0.98	0.98	-1.96	-0.96	2.69	-0.15	-0.05	0.13	6.34	6.00	5.74	0.80	0.91	0.93	0.89	0.97	0.97
Min	0.92	0.94	0.90	-5.71	-5.35	-1.79	-0.33	-0.24	-0.07	3.40	2.11	2.25	0.44	0.69	0.85	0.62	0.70	0.75

Note: min, max, and avg represent the minimum, maximum, and average soil temperature, respectively.

FIGURE 4 Box plot of T_{\max} , T_{\min} , and T_{avg} and their position to the optimal values (red dashed line) based on CC, ME, PBIAS, RMSE, KGE, and NSE for GLDAS-derived ST in the first layer (0–10 cm).



To quantify GLDAS-Noah data accuracy with respect to the observation values, additional statistical metrics, including ME, RBIAS, RMSE, KGE, and NSE, were applied and are displayed in Table 5. GLDAS-Noah ST in the 0–10 cm profile showed an acceptable level of accuracy when compared to the observed ST data. In the studied stations, the values of ME varied from -5.71 to -1.96 , RMSE from RBIAS from -33% to -15% , KGE from 0.44 to 0.88, and NSE from 0.62 to 0.89 for T_{\min} . The negative RBIAS values indicated that the GLDAS-Noah soil temperature overestimated observed minimum ST in all stations.

Table 5 shows the results from the evaluation of GLDAS-Noah ST at 09:00 GMT (taken as a proxy for T_{\max}) in the 0–10 cm vertical layer. Considering the findings obtained for all the stations, ME ranges from -1.79 to 2.69°C , RBIAS from -7% to 13% , RMSE from 2.25 to 5.74°C , KGE from 0.85 to 0.93, and NSE from 0.75 to 0.97 for T_{\max} . Likewise, ME ranges from -5.71 to -1.96°C , RBIAS from -33% to -15% , RMSE from 3.4 to 6.34°C , KGE from 0.85 to 0.93, and NSE from 0.75 to 0.97 for T_{\min} . The RBIAS values showed that GLDAS-Noah T_{\max} was either underestimated or overestimated. However, the evaluation indices, specifically NSE and KGE, showed an enhanced performance of GLDAS-Noah ST for T_{\min} at depths of 0–10 cm for the investigated period from 2008 to 2012.

Likewise, the evaluation of ST at 15:00 GMT (proxy for T_{avg}) in the 0–10 cm profile is also presented in Table 5. Based on the findings obtained in the stations studied, ME varied from -5.35 to -0.96 , RBIAS from -24% to -5% , RMSE from 2.11 to 6.00°C , KGE from 0.44 to 0.80, and NSE from 0.62 to 0.89. The RBIAS values revealed that the GLDAS-based ST (T_{avg}) was overestimated in all stations. However, based on the KGE and NSE indices, T_{avg} is simulated with an acceptable level of accuracy by GLDAS-Noah in the 0–10 cm layer at 15:00 GMT.

As indicated in Table 5, the overall performance of GLDAS-Noah for prediction of T_{\min} over the study area is 0.96, -24% , 0.66, and 0.79 based on the correlation coefficient, RBIAS, KGE, and NSE, respectively. The values for T_{avg} are 0.97, -12% , 0.84, and 0.89; and for T_{\max} are 0.95, 5% , 0.89, and 0.89, correspondingly, which indicates a significant positive correlation between the observed and simulated time series with the lowest bias for T_{\max} and highest bias for T_{\min} . According to Maggioni et al. (2022), there is no fixed threshold limit for RMSE, but it is always better to obtain an RMSE value as low as possible. In this case, relatively high RMSE values are due to the retrieval methods, the Noah land surface model tends to underestimate downward longwave radiation from its forcing data, and also instrumental error which are associated to the observations data (Wang, Li, et al., 2016).

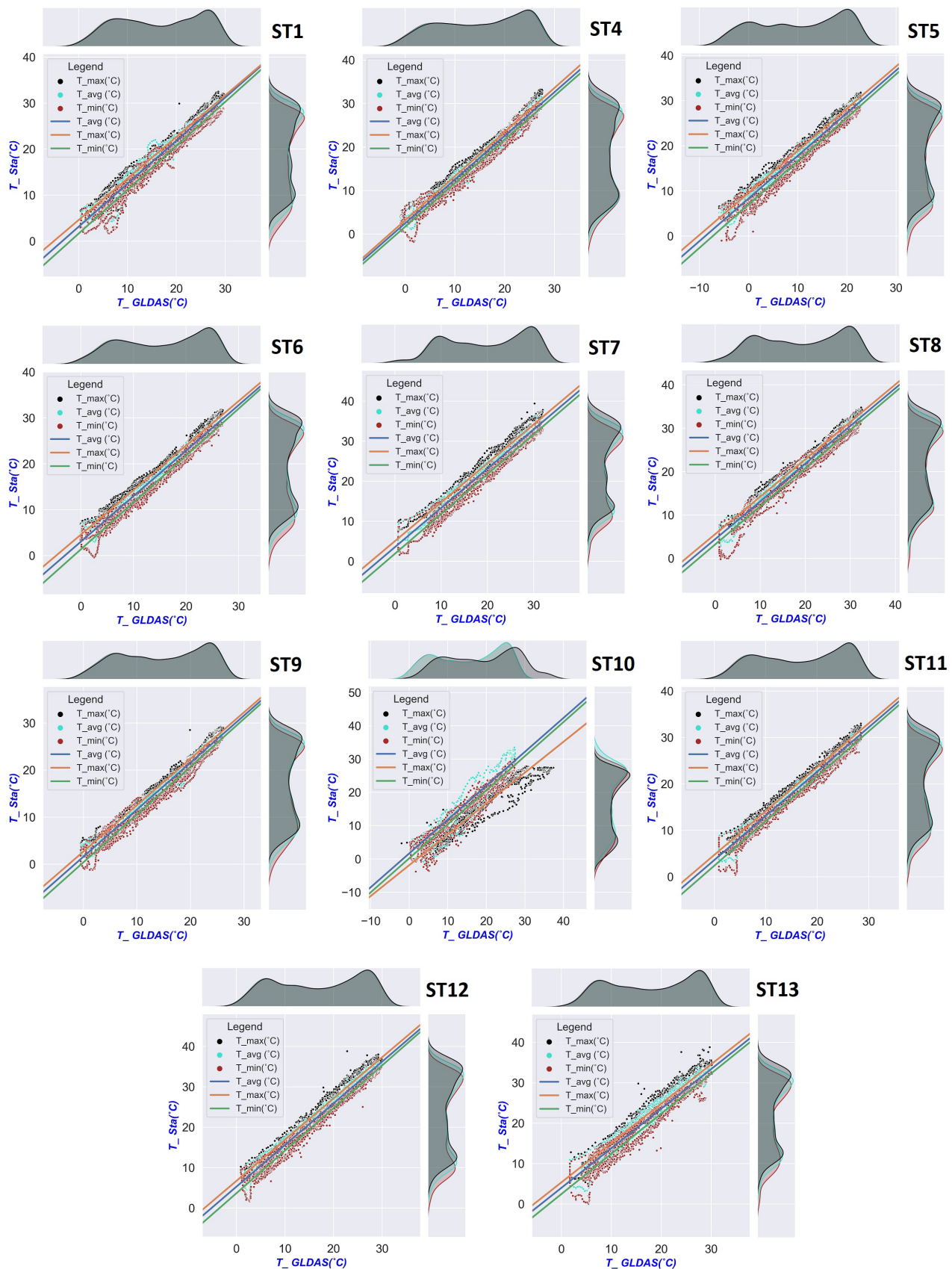


FIGURE 5 Scattergram visually displays the relationship between the observed and Global Land Data Assimilation System-derived soil temperature (T_{avg} , blue; T_{max} , black; and T_{min} brown) ST1, ST4-ST8 for the first layer (40–100 cm) at the stations. Scattergram visually displays the relationship between the observed and GLDAS derived ST (T_{avg} , blue; T_{max} , black; and T_{min} brown) for the second layer (40–100 cm) at the stations ST9-ST13.

TABLE 6 Statistical metrics used to evaluate the Global Land Data Assimilation System-Noah against the observed soil temperature at the synoptic stations in the second layer (40–100 cm).

No.	R			ME (C°)			RBIAS (%)			RMSE (C°)			KGE			NSE		
	T _{min}	T _{avg}	T _{max}	T _{min}	T _{avg}	T _{max}	T _{min}	T _{avg}	T _{max}	T _{min}	T _{avg}	T _{max}	T _{min}	T _{avg}	T _{max}	T _{min}	T _{avg}	T _{max}
ST1	0.96	0.98	0.96	-0.88	-2.06	-2.97	-0.05	-0.11	-0.15	1.74	2.46	3.37	0.89	0.78	0.68	0.95	0.90	0.81
ST4	0.96	0.96	0.96	-1.23	-2.06	-2.88	-0.07	-0.12	-0.15	2.21	2.87	3.74	0.85	0.79	0.74	0.82	0.78	0.72
ST5	0.96	0.98	0.98	-1.95	-3.09	-4.19	-0.10	-0.16	-0.13	6.72	7.81	8.93	0.71	0.85	0.75	0.82	0.85	0.78
ST6	0.98	0.98	0.98	-1.52	-2.72	-3.92	-0.09	-0.15	-0.2	2.05	3.08	4.30	0.87	0.75	0.64	0.91	0.83	0.69
ST7	0.98	0.98	0.98	-1.77	-3.13	-4.47	-0.08	-0.14	-0.18	2.17	3.27	4.63	0.88	0.78	0.69	0.94	0.85	0.70
ST8	0.98	0.98	0.98	-0.89	-1.92	-2.92	-0.04	-0.09	-0.13	1.93	2.44	3.32	0.82	0.73	0.65	0.94	0.91	0.83
ST9	0.96	0.98	0.98	0.10	-0.84	-1.73	0.01	-0.05	-0.1	1.45	1.94	2.80	0.95	0.88	0.80	0.74	0.71	0.65
ST10	0.96	0.96	0.94	-0.38	-1.75	-3.07	-0.02	-0.10	-0.16	1.97	2.71	4.03	0.94	0.85	0.76	0.92	0.88	0.79
ST11	0.98	0.98	0.98	-1.78	-2.79	-3.74	-0.1	-0.14	-0.18	2.07	2.93	3.90	0.83	0.74	0.66	0.93	0.86	0.72
ST12	0.98	0.98	0.98	-4.54	-5.8	-7.05	-0.21	-0.26	-0.29	4.75	5.88	7.13	0.71	0.61	0.52	0.74	0.59	0.38
ST13	0.96	0.96	0.96	-1.96	-3.22	-4.47	-0.1	-0.15	-0.19	3.01	3.92	5.14	0.85	0.74	0.66	0.76	0.67	0.53
Avg	0.97	0.97	0.97	-1.53	-2.67	-3.76	-0.08	-0.13	-0.17	2.73	3.57	4.66	0.85	0.77	0.69	0.86	0.80	0.69
Max	0.98	0.98	0.98	0.10	-0.84	-1.73	0.01	-0.05	-0.1	6.72	7.81	8.93	0.95	0.88	0.80	0.95	0.91	0.83
Min	0.96	0.96	0.94	-4.54	-5.8	-7.05	-0.21	-0.26	-0.29	1.45	1.94	2.80	0.71	0.61	0.52	0.74	0.59	0.38

Note: min, max, and avg represents the minimum, maximum, and average soil temperature, respectively.

Box and Whisker plots were used to illustrate the summary results of the correlation. Optimal values are shown with a red dashed line for each metric. The scatter plots for the statistical metrics in the first layer (0–10 cm) are presented in Figure 4.

Based on the box plots presented in Figure 4, predicted T_{\max} shows a closer tendency to the optimal values for NSE, KGE, RMES, and ME in the first soil layer (0–10 cm). In this case, minimum relative bias is obtained for T_{\max} .

4.2 | GLDAS-Noah ST for the 40–100 cm layer

The results for the GLDAS-Noah ST in 40–100 cm layer are presented in Figure 5 and Table 6. Similar to the first layer, a linear relationship was established between the GLDAS and the ground observation time series for T_{\max} , T_{\min} , and T_{avg} in the second layer. Surprisingly, a slightly better correlation was found between the two time series with values in the range of 0.96–0.98, 0.94–0.98, and 0.96–0.98 for T_{\min} , T_{\max} , and T_{avg} , respectively, for the second layer.

These findings demonstrated excellent performance of the GLDAS-Noah for the simulation of ST at 03:00 GMT (T_{\min}) in the 40–100 cm soil layer (Table 6). In the studied stations, ME ranged from -4.54 to 0.10 , RMSE from 1.45 to 6.72°C , RBIAS from -21% to -10% , KGE from 0.71 to 0.95 , and NSE from 0.74 to 0.95 . Similar to

the results from the first layer, the negative RBIAS values in all the stations indicated that the GLDAS-Noah overestimated ST in the second layer. Table 6 also reveals interesting results of GLDAS-Noah ST at 09:00 GMT (T_{\max}) for the 40–100 cm soil layer. According to the ME metric, simulated ST errors varied between -7.05 and -1.73 , RBIAS varied between -29% and -10% , RMSE varied between 2.80 and 8.93°C , KGE varied between 0.52 to 0.80 , and NSE varied between 0.38 to 0.83 . The RBIAS values showed that GLDAS-Noah T_{\max} was overestimated in the second layer in all stations. However, the primary evaluation indices, specifically NSE and KGE, revealed a good agreement between the observed and simulated ST in all stations except ST12, with the lowest NSE (0.38) for the investigated 2008–2012 period.

A final evaluation was conducted on the time series at 15:00 GMT (T_{avg}) in the 40–100 cm layer, with ME ranged from -5.80 to -0.84 , RBIAS from -26% to -5% , RMSE from 1.94 to 7.81°C , KGE from 0.61 to 0.88 , and NSE from 0.59 to 0.91 . RBIAS values revealed that the GLDAS-based ST (T_{avg}) was overestimated in all stations. However, based on the KGE and NSE indices, simulated T_{avg} values from GLDAS-Noah at 15:00 GMT in the 40–100 cm layer were found to have an acceptable level of accuracy.

The overall performance of GLDAS-Noah for simulation of T_{avg} over the study area is 0.97 , -13% , 0.77 , and 0.80 based on the correlation coefficient, RBIAS, KGE, and NSE, respectively, which indicates an accurate

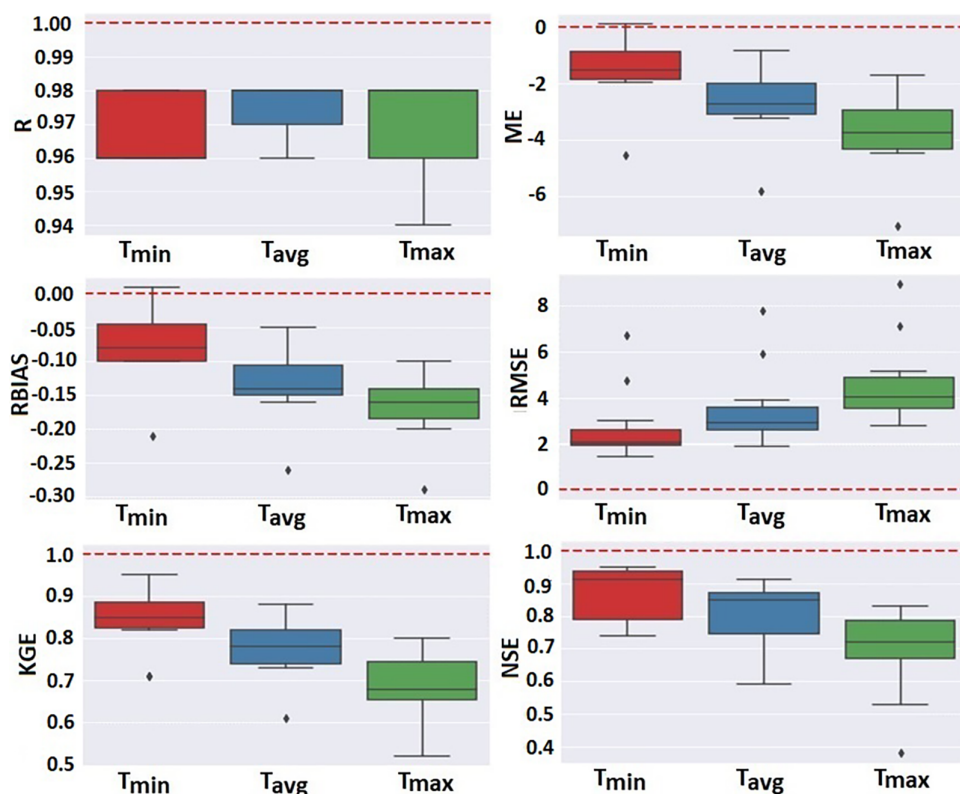
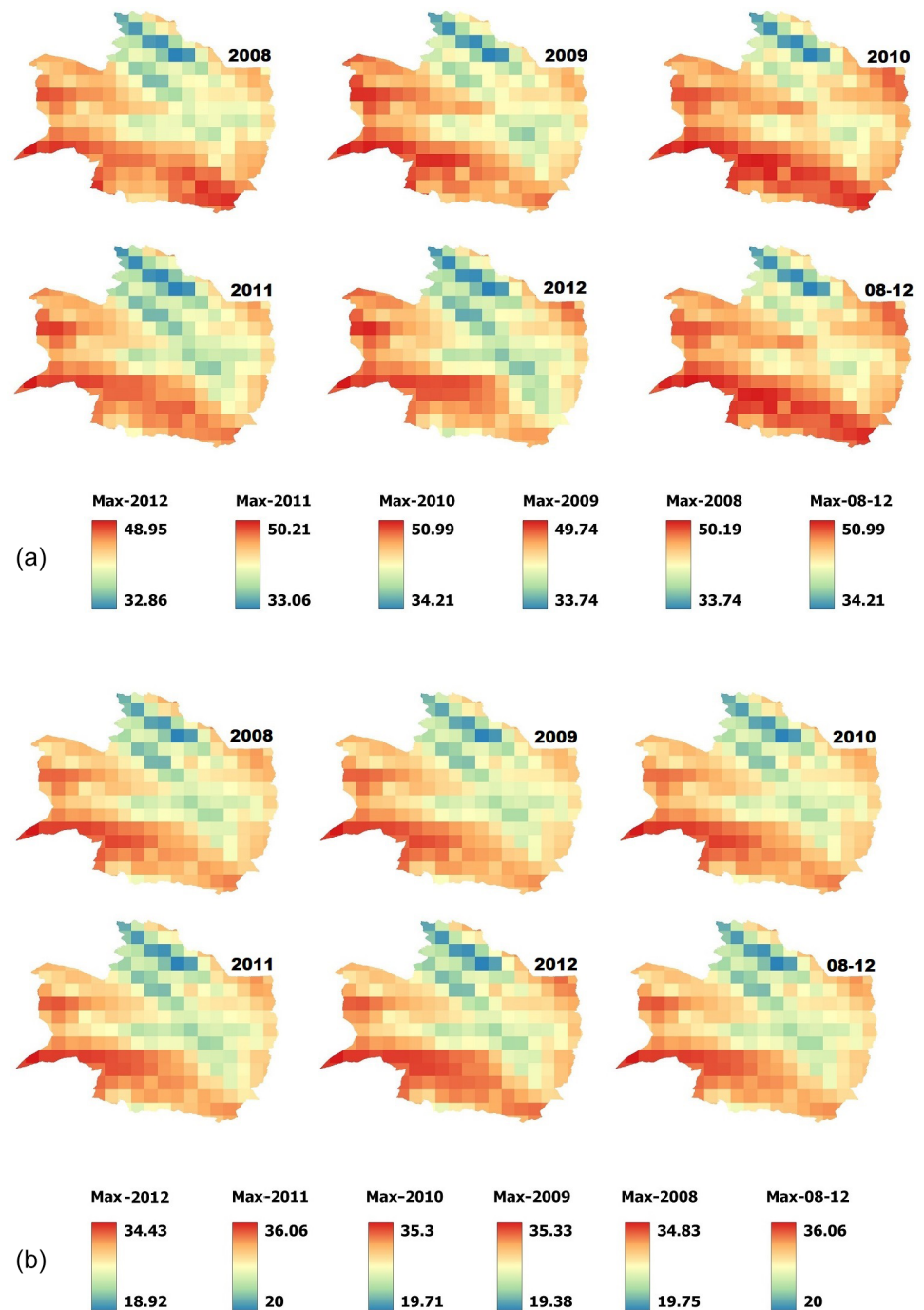


FIGURE 6 Box plot of T_{\max} , T_{\min} , and T_{avg} and their position to the optimal values (red dashed line) based on CC, ME, PBIAS, RMSE, KGE, and NSE for Global Land Data Assimilation System-derived soil temperature in the second layer (40–100 cm).

FIGURE 7 (a) Spatio-temporal variation of maximum soil temperature (ST) for the first layer depth 0–10 cm. (b) Spatio-temporal variation of maximum ST for the second layer depth 40–100 cm over the study area.



simulation of GLDAS-Noah for average daily ST in the second layers (40–100 cm).

As evident in Figure 6, predicted T_{\min} shows a closer proximity to the optimal values for NSE, KGE, RBIAS, RMES, and ME in the second soil layer (40–100 cm).

4.3 | Spatio-temporal variation of ST over the study area

Considering the strong correlation between the two time series, an interesting result emerges from the analysis of

the spatial-temporal variation of ST over the investigated area with the $0.25^{\circ} \times 0.25^{\circ}$ grid resolution. Using GEE and the Quantum Geographic Information System (QGIS) software, the ST map for the T_{\min} , T_{\max} , and T_{avg} was developed for the two investigated layers. The role of T_{\max} and T_{\min} is critically important in many applications, specifically in agriculture and civil engineering projects.

Figure 7 shows the spatial variation of annual T_{\max} for the 0–10 cm layer, over the Khorasan Razavi region during the investigated time period. (i.e., for the years 2008, 2009, 2010, 2011, 2012, and 2008–2012).

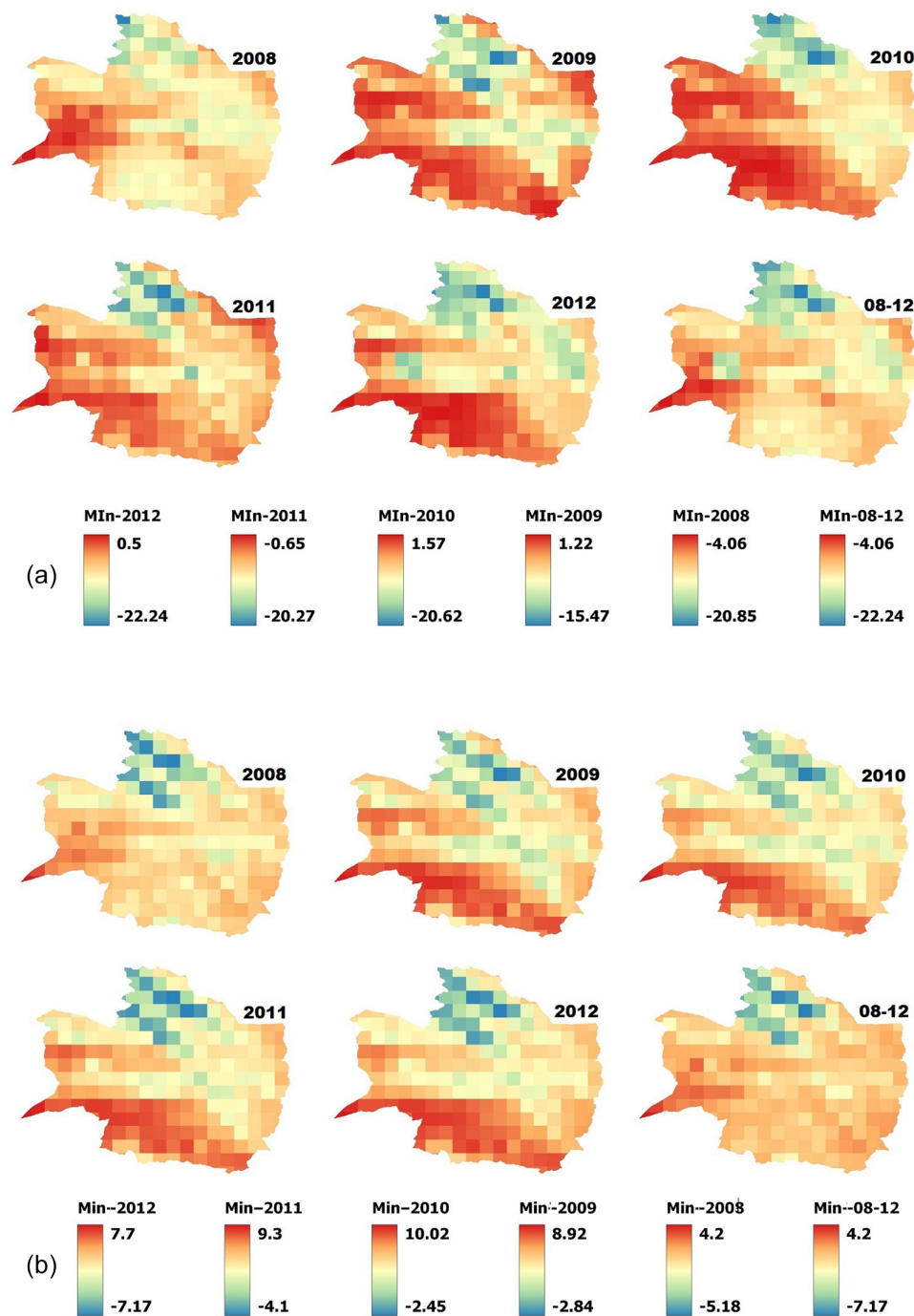


FIGURE 8 (a) Spatio-temporal variation of minimum soil temperature (ST) for the first layer depth 0–10 cm. (b) Spatio-temporal variation of minimum ST for the second layer depth 40–100 cm over the study area.

It is evident in Figure 7 that the maximum temperature in the south of the Khorasan Razavi region is much higher than in the northwest, with roughly a 16°C difference. T_{\max} tends to increase from the south to the north of Khorasan Razavi region. Figure 8 presents the spatial variation of annual T_{\max} in the second layer (40–100 cm), over Khorasan Razavi region during the investigated time period. As illustrated in Figure 8, the same trend is observed for the second layer with values in the range $18.92\text{--}36.06^{\circ}\text{C}$.

5 | CONCLUSION

We have evaluated the GLDAS-Noah soil temperature at the depths of 0–10 cm and 40–100 cm within the soil profile layers, comparing it with observed ST data at 03:00, 09:00, and 15:00 GMT in north-eastern Iran from January 2008 to December 2012 by utilization of GEE as a web-based data storage platform and data processing tools. The results demonstrate the promising utility of GLDAS-Noah in regions with limited observational data.

Notably, the study reveals that the accuracy of GLDAS-Noah ST is higher in shallower layers compared to deeper layers, where accuracy is lower, particularly for T_{\max} . Although accuracy is slightly better for T_{\min} , a significant negative bias leads to an overestimation of ST in the two examined soil layers.

The analysis indicates an average bias of -24% , -12% , and -5% for T_{\min} , T_{avg} , and T_{\max} , respectively, in the first layer, and an average bias of -8% , -13% , and -17% for T_{\min} , T_{avg} , and T_{\max} , respectively, in the second layer over the study period. Despite this bias, a close correlation is observed between GLDAS-Noah and observed ST, underscoring the algorithm's reliability for simulating ST in arid regions with limited observation data crucial for landscaping and construction, where soil temperature and moisture levels are pivotal. In addition, ST has significant effects in chemical and biological processes which take place in soils of study area. This includes seed germination, bugs and microbes, plant and animals that live in the environment. Furthermore, the method proves cost-effective and offers enhanced spatiotemporal coverage for the T_{\min} , T_{avg} , and T_{\max} over the whole study area.

The study also highlights the advantage of GEE, which significantly simplifies visualizing, manipulating, and editing, more than 30 years of historical imagery and reanalyzed datasets suitable for soil studies. It serves as a favorable platform for satellite-based soil research, particularly in evaluating GLDAS products. While this study focused on GLDAS-Noah LSM, it suggests future investigations to assess the performance of other LSMs, including CLM, Mosaic, and VIC, in arid regions.

AUTHOR CONTRIBUTIONS

Abolghasem Akbari: Conceptualization (equal); formal analysis (equal); methodology (equal); writing – original draft (equal). **Majid Rajabi Jaghargh:** Formal analysis (equal); visualization (equal). **Azizan Abu Samah:** Supervision (equal). **Jonathan Peter Cox:** Writing – review and editing (equal). **Mojtaba Gholamzadeh:** Investigation (equal). **Alireza Araghi:** Methodology (equal). **Patricia M. Saco:** Writing – review and editing (equal). **Khabat Khosravi:** Conceptualization (equal).

ACKNOWLEDGMENTS

We extend our sincere appreciation to NASA and the GLDAS team for their contribution in making soil temperature (ST) data publicly accessible. Similarly, we express gratitude to Google for the development of Google Earth Engine (GEE). Additionally, we wish to thank IRIMO for supplying the soil temperature data from the synoptic stations.

FUNDING INFORMATION

There is no funding for this research.

CONFLICT OF INTEREST STATEMENT

There is not any competing or conflict of interest among authors.

DATA AVAILABILITY STATEMENT

We do not have permission to share data from the synoptic stations, but GLDAS data related to this research are available free in the GEE repository.

ORCID

Abolghasem Akbari  <https://orcid.org/0000-0003-4394-1970>

REFERENCES

- Afandi, A., Lusi, N., Catrawedarma, I. & Rudiyanto, B. (2022) Prediction of temperature in 2 meters temperature probe survey in Blawan geothermal field using artificial neural network (ANN) method. *Case Studies in Thermal Engineering*, 38, 102309. Available from: <https://doi.org/10.1016/j.csite.2022.102309>
- Akbari, A., Daryabor, F., Samah, A.A. & Fanodi, M. (2017) Validation of TRMM 3B42 V6 for estimation of mean annual rainfall over ungauged area in semi-arid climate. *Environmental Earth Sciences*, 76, 1–10. Available from: <https://doi.org/10.1007/s12665-017-6867-3>
- Araghi, A., Mousavi-Baygi, M., Adamowski, J., Martinez, C. & van Der Ploeg, M. (2017) Forecasting soil temperature based on surface air temperature using a wavelet artificial neural network. *Meteorological Applications*, 24, 603–611. Available from: <https://doi.org/10.1002/met.1661>
- Ashraf, B., Mousavi Baygi, M., Kamali, G. & Davari, K. (2011) Prediction of seasonal variations of climatological parameters over next 20 years by using statistical downscaling method of HADCM3 data (case study: Khorasan Razavi province). *Water and Soil*, 25, 940–952. Available from: <https://doi.org/10.22067/jsw.v0i0.10267>
- Badache, M., Eslami-Nejad, P., Ouzzane, M., Aidoun, Z. & Lamarche, L. (2016) A new modeling approach for improved ground temperature profile determination. *Renewable Energy*, 85, 436–444. Available from: <https://doi.org/10.1016/j.renene.2015.06.020>
- Bayatvarkeshi, M., Bhagat, S.K., Mohammadi, K., Kisi, O., Farahani, M., Hasani, A. et al. (2021) Modeling soil temperature using air temperature features in diverse climatic conditions with complementary machine learning models. *Computers and Electronics in Agriculture*, 185, 106158. Available from: <https://doi.org/10.1016/j.compag.2021.106158>
- Bond-Lamberty, B., Wang, C. & Gower, S.T. (2005) Spatiotemporal measurement and modeling of stand-level boreal forest soil temperatures. *Agricultural and Forest Meteorology*, 131, 27–40. Available from: <https://doi.org/10.1016/j.agrformet.2005.04.008>
- Chow, T.T., Long, H., Mok, H. & Li, K. (2011) Estimation of soil temperature profile in Hong Kong from climatic variables. *Energy and Buildings*, 43, 3568–3575. Available from: <https://doi.org/10.1016/j.enbuild.2011.09.026>
- Cohen, A. & Fielding, M. (1979) Predicting frost depth; protecting underground pipelines. *Journal-American Water Works Association*, 71, 113–116. Available from: <https://www.jstor.org/stable/41270234>

- Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G. et al. (2003) The common land model. *Bulletin of the American Meteorological Society*, 84, 1013–1024. Available from: <https://doi.org/10.1175/BAMS-84-8-1013>
- Datta, P. & Faroughi, S.A. (2023) A multihead LSTM technique for prognostic prediction of soil moisture. *Geoderma*, 433, 116452. Available from: <https://doi.org/10.1016/j.geoderma.2023.116452>
- Durbin, J. (1971) Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov-Smirnov test. *Journal of Applied Probability*, 8, 431–453. Available from: <https://doi.org/10.2307/3212169>
- Durbin, J. & Watson, G.S. (1992) Testing for serial correlation in least squares regression. I. In: *Breakthroughs in Statistics*, Vol. 37. New York, NY: Springer, p. 409. Available from: <https://doi.org/10.2307/2332391>
- Ek, M., Mitchell, K., Lin, Y., Rogers, E., Grunmann, P., Koren, V. et al. (2003) Implementation of Noah land surface model advances in the National Centers for environmental prediction operational mesoscale eta model. *Journal of Geophysical Research: Atmospheres*, 108(D22), 8851. Available from: <https://doi.org/10.1029/2002JD003296>
- Faridi, H., Arabhosseini, A., Zarei, G. & Okos, M. (2019) Utilization of soil temperature modeling to check the possibility of earth-air heat exchanger for agricultural building. *Iranian (Iranica) Journal of Energy & Environment*, 10, 260–268. Available from: <https://doi.org/10.5829/ijee.2019.10.04.06>
- Gao, W., Whalley, W.R., Tian, Z., Liu, J. & Ren, T. (2016) A simple model to predict soil penetrometer resistance as a function of density, drying and depth in the field. *Soil and Tillage Research*, 155, 190–198. Available from: <https://doi.org/10.1016/j.still.2015.08.004>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. & Moore, R. (2017) Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. Available from: <https://doi.org/10.1016/j.rse.2017.06.031>
- Gupta, H.V., Sorooshian, S. & Yapo, P.O. (1999) Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*, 4, 135–143. Available from: [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135))
- Holmes, T., Owe, M., De Jeu, R. & Kooi, H. (2008) Estimating the soil temperature profile from a single depth observation: A simple empirical heatflow solution. *Water Resources Research*, 44. Available from: <https://doi.org/10.1029/2007WR005994>
- Koster, R.D. & Suarez, M.J. (1996) *Energy and water balance calculations in the Mosaic Lsm*. Greenbelt, Maryland: National Aeronautics and Space Administration, Goddard Space Flight Center. Available from: <https://gmao.gsfc.nasa.gov/pubs/docs/Koster130.pdf>
- Kumar, L. & Mutanga, O. (2018) Google earth engine applications since inception: usage, trends, and potential. *Remote Sensing*, 10, 1509. Available from: <https://doi.org/10.3390/rs10101509>
- Li, M., Wu, P. & Ma, Z. (2020) A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets. *International Journal of Climatology*, 40, 5744–5766. Available from: <https://doi.org/10.1002/joc.6549>
- Li, Q., Zhu, Y., Shangquan, W., Wang, X., Li, L. & Yu, F. (2022) An attention-aware LSTM model for soil moisture and soil temperature prediction. *Geoderma*, 409, 115651. Available from: <https://doi.org/10.1016/j.geoderma.2021.115651>
- Liang, X. (1994) *A two-layer variable infiltration capacity land surface representation for general circulation models*. USA: University of Washington. Available from: <https://ntrs.nasa.gov/citations/19980073396>
- Maggioni, V., Massari, C. & Kidd, C. (2022) Errors and uncertainties associated with quasiglobal satellite precipitation products. In: *Precipitation Science*. Netherlands: Elsevier. Available from: <https://doi.org/10.1016/B978-0-12-822973-6.00023-8>
- Mihalakakou, G. (2002) On estimating soil surface temperature profiles. *Energy and Buildings*, 34, 251–259. Available from: [https://doi.org/10.1016/S0378-7788\(01\)00089-5](https://doi.org/10.1016/S0378-7788(01)00089-5)
- Moriasi, D.N., Arnold, J.G., Liew, V.A.N., M.W., Bingner, R.L., Harmel, R.D. et al. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50, 885–900. Available from: <https://doi.org/10.13031/2013.23153>
- Napagoda, N. & Tilakaratne, C. (2012) Artificial neural network approach for modeling of soil temperature: a case study for Bathalagoda area. *Sri Lankan Journal of Applied Statistics*, 13, 39–59. Available from: <https://doi.org/10.4038/sljstats.v13i0.5123>
- NASA. (2021) *GLDAS: project goals*. USA: NASA. <https://ldas.gsfc.nasa.gov/gldas#:~:text=The%20goal%20of%20the%20Global,et%20al.%2C%202004a> [Accessed 22nd September 2021].
- Nash, J.E. & Sutcliffe, J.V. (1970) River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10, 282–290. Available from: [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Nashwan, M.S., Shahid, S., Dewan, A., Ismail, T. & Alias, N. (2020) Performance of five high resolution satellite-based precipitation products in arid region of Egypt: an evaluation. *Atmospheric Research*, 236, 104809. Available from: <https://doi.org/10.1016/j.atmosres.2019.104809>
- Ozgener, O., Ozgener, L. & Tester, J.W. (2013) A practical approach to predict soil temperature variations for geothermal (ground) heat exchangers applications. *International Journal of Heat and Mass Transfer*, 62, 473–480. Available from: <https://doi.org/10.1016/j.ijheatmasstransfer.2013.03.031>
- Rodell, M., Houser, P., Peters-Lidard, C., Kato, H., Kumar, S., Gottschalck, J. et al. (2004) NASA/NOAA's global land data assimilation system (GLDAS): Recent results and future plans. Proceedings of the ECMWF/ELDAS Workshop on Land Surface Assimilation, 61–68. <https://www.ecmwf.int/en/elibrary/76196-nasanoas-global-land-data-assimilation-system-gldas-recent-results-and-future>
- Rui, H. & Beaudoin, H. (2018) *README document for NASA GLDAS version 2 data products*. USA: NASA's Goddard Space Flight Center. https://data.mint.isi.edu/files/raw-data/GLDAS_NOAH025_M.2.0/doc/README_GLDAS2.pdf
- Rui, H. & Beaudoin, H. (2019) *README document for NASA GLDAS version 2 data products* [online]. Available from: https://data.mint.isi.edu/files/raw-data/GLDAS_NOAH025_M.2.0/doc/README_GLDAS2.pdf [Accessed 8th March 2022].
- Samadianfard, S., Asadi, E., Jarhan, S., Kazemi, H., Kheshtgar, S., Kisi, O. et al. (2018) Wavelet neural networks and gene expression programming models to predict short-term soil

- temperature at different depths. *Soil and Tillage Research*, 175, 37–50. Available from: <https://doi.org/10.1016/j.still.2017.08.012>
- Samadianfard, S., Ghorbani, M.A. & Mohammadi, B. (2018) Forecasting soil temperature at multiple-depth with a hybrid artificial neural network model coupled-hybrid firefly optimizer algorithm. *Information Processing in Agriculture*, 5, 465–476. Available from: <https://doi.org/10.1016/j.inpa.2018.06.005>
- Shamshirband, S., Esmailbeiki, F., Zarehaghi, D., Neyshabouri, M., Samadianfard, S., Ghorbani, M.A. et al. (2020) Comparative analysis of hybrid models of firefly optimization algorithm with support vector machines and multilayer perceptron for predicting soil temperature at different depths. *Engineering Applications of Computational Fluid Mechanics*, 14, 939–953. Available from: <https://doi.org/10.1080/19942060.2020.1788644>
- Shapiro, S.S. & Wilk, M.B. (1965) An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591–611. Available from: <https://doi.org/10.2307/2333709>
- Shirmohammadi-Aliakbarhiani, Z. & Akbari, A. (2020) Ground validation of diurnal TRMM 3B42 V7 and GPM precipitation products over the northeast of Iran. *Theoretical and Applied Climatology*, 142, 1413–1423. Available from: <https://doi.org/10.1007/s00704-020-03392-0>
- Singh, J., Knapp, H.V. & Demissie, M. (2004) Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. In: *Illinois state water survey contract report*. Illinois: Illinois Department of Natural Resources and the Illinois State Geological Survey. Available from: <http://www.isws.illinois.edu/pubdoc/CR/ISWSCR2004-08.pdf>
- Tabari, H., Hosseinzadeh Talaei, P. & Willems, P. (2015) Short-term forecasting of soil temperature using artificial neural network. *Meteorological Applications*, 22(3), 576–585. Available from: <https://doi.org/10.1002/met.1489>
- Thiemig, V., Rojas, R., Zambrano-Bigiarini, M. & De Roo, A. (2013) Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. *Journal of Hydrology*, 499, 324–338. Available from: <https://doi.org/10.1016/j.jhydrol.2013.07.012>
- Wang, L., Li, X., Chen, Y., Yang, K., Chen, D., Zhou, J. et al. (2016) Validation of the global land data assimilation system based on measurements of soil temperature profiles. *Agricultural and Forest Meteorology*, 218, 288–297. Available from: <https://doi.org/10.1016/j.agrformet.2016.01.003>
- Wang, S., Li, R., Wu, Y. & Wang, W. (2023) Estimation of surface soil moisture by combining a structural equation model and an artificial neural network (SEM-ANN). *Science of the Total Environment*, 876, 162558. Available from: <https://doi.org/10.1016/j.scitotenv.2023.162558>
- Wang, W., Cui, W., Wang, X. & Chen, X. (2016) Evaluation of GLDAS-1 and GLDAS-2 forcing data and Noah model simulations over China at the monthly scale. *Journal of Hydrometeorology*, 17, 2815–2833. Available from: <https://doi.org/10.1175/JHM-D-15-0191.1>
- Wang, X., Ding, Y., Zhao, C. & Wang, J. (2018) Validation of TRMM 3B42V7 rainfall product under complex topographic and climatic conditions over Hexi region in the northwest arid region of China. *Water*, 10, 1006. Available from: <https://doi.org/10.3390/w10081006>
- WMO. (2018) Guide to Instruments and Methods of Observation Volume 1–Measurement of Meteorological Variables. *Guide to Instruments and Methods of Observation*. <https://library.wmo.int/idurl/4/41650>
- Xing, L., Li, L., Gong, J., Ren, C., Liu, J. & Chen, H. (2018) Daily soil temperatures predictions for various climates in United States using data-driven model. *Energy*, 160, 430–440. Available from: <https://doi.org/10.1016/j.energy.2018.07.004>
- Yang, K. & Zhang, J. (2018) Evaluation of reanalysis datasets against observational soil temperature data over China. *Climate Dynamics*, 50, 317–337. Available from: <https://doi.org/10.1007/s00382-017-3610-4>
- Yang, S., Li, R., Wu, T., Hu, G., Xiao, Y., Du, Y. et al. (2020) Evaluation of reanalysis soil temperature and soil moisture products in permafrost regions on the Qinghai-Tibetan Plateau. *Geoderma*, 377, 114583. Available from: <https://doi.org/10.1016/j.geoderma.2020.114583>
- Zeynoddin, M., Ebtehaj, I. & Bonakdari, H. (2020) Development of a linear based stochastic model for daily soil temperature prediction: one step forward to sustainable agriculture. *Computers and Electronics in Agriculture*, 176, 105636. Available from: <https://doi.org/10.1016/j.scitotenv.2020.138015>

How to cite this article: Akbari, A., Rajabi Jaghargh, M., Abu Samah, A., Peter Cox, J., Gholamzadeh, M., Araghi, A., Saco, P. M., & Khosravi, K. (2024). Utilization of the Google Earth Engine for the evaluation of daily soil temperature derived from Global Land Data Assimilation System in two different depths over a semiarid region. *Meteorological Applications*, 31(4), e2221. <https://doi.org/10.1002/met.2221>