DOI: 10.1111/tgis.12607

RESEARCH ARTICLE

Transactions in GIS WILEY

Incorporating dynamic factors for improving a GIS-based solar radiation model

Mingxi Zhang[1,2,3] | Bin Wang[3] | De Li Liu[3,4] | Jiandong Liu[5] | Hong Zhang[2,3] | Puyu Feng[2,3] | Dongdong Kong[6] | James Cleverly[2] | Xihua Yang[1,2,7] | Qiang Yu[1,2,8]

[1]State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling, China

[2]School of Life Sciences, Faculty of Science, University of Technology Sydney, Sydney, Australia

[3]NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, Australia

[4]Climate Change Research Centre and ARC Centre of Excellence for Climate Extremes, University of New South Wales, Sydney, Australia

[5]State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China

[6]Department of Water Resources and Environment, Sun Yat-sen University, Guangzhou, China

[7]NSW Office of Environment and Heritage, Sydney, Australia

[8]College of Resources and Environment, University of Chinese Academy of Science, Beijing, China

Correspondence
Qiang Yu, State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling 712100, Shaanxi, China.
Email: yuq@nwafu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 41730645

Abstract

Solar radiation has been a major input to agricultural, hydrological, and ecological modeling. However, solar radiation is usually influenced by three groups of dynamic factors: sun–earth position, terrain, and atmospheric effects. Therefore, an integrated approach to accurately consider the impacts of those dynamic factors on solar radiation is essential to estimate solar radiation over rugged terrain. In this study, a spatial and temporal gap-filling algorithm was proposed to obtain a seamless daily MODIS albedo dataset. A 1 km-resolution digital elevation model was used to model the impact of local topography and shading by surrounding terrain on solar radiation. A sunshine-based model was adopted to simulate radiation under the influence of clouds. A GIS-based solar radiation model that incorporates albedo, shading by surrounding terrain, and variations in cloudiness was used to address the spatial variability of these factors in mountainous terrain. Compared with other independent solar radiation products, our model generated a more reliable solar radiation product over rugged terrain, with an R^2 of 0.88 and an RMSE of 2.55 MJ m^{-2} day^{-1}. The improved solar radiation products and open source app can be used further in practice or scientific research.

1 | INTRODUCTION

Solar radiation is the primary driving force for earth system processes, and its supply is a major input to agricultural, hydrological, and ecological models (Aguilar, Herrero, & Polo, 2010; Brock, 1981; Fu & Rich, 2002). Therefore, knowledge of the spatial and temporal variability of incoming solar radiation is critical for understanding these processes. Additionally, fine spatial and temporal mapping and monitoring of solar radiation components are essential for the design of solar energy systems.

The spatial and temporal heterogeneity of solar radiation over rugged terrain is determined by three groups of dynamic factors: sun–earth position, terrain, and atmospheric effects (Pintor et al., 2015). Based on the sun–earth geometry formulation, the first group can be precisely calculated. For the other two groups, the effects of terrain (shadowing, absorption, and reflection) and atmosphere are difficult to model due to their dynamic nature. Particularly, the albedo of the underlying surface modulates the amount of solar radiation absorbed and reflected by that surface and directly controls the distribution of solar radiation between the surface and the atmosphere. Additionally, shadows cast by complex topography due to different incident angles of the rays determine the fraction of direct and diffuse radiation in global solar radiation. Furthermore, clouds play a major role in the atmospheric attenuation of incoming solar radiation, but modeling of the radiative effects of clouds is challenging due to their variability in time, location, and condition. Hence, quantitative modeling of the impacts of those dynamic factors on solar radiation is essential to accurately estimate solar radiation over rugged terrain.

Three major methods have been used for solar radiation modeling over the past few decades, namely traditional interpolation methods, GIS-based solar radiation models, and satellite-derived solar radiation estimates (Hofierka & Suŕi, 2002; Qin et al., 2015; Ruiz-Arias, Tovar-Pescador, Pozo-Vázquez, & Alsamamra, 2009; Zhang, Li, & Bai, 2015). In spatial interpolation methods, unknown values are reliably predicted from ground-based measurements and external complementary information. However, the reliability of such methods strongly depends on sample size and the complexity of the topography (Alsamamra, Ruiz-Arias, Pozo-Vázquez, & Tovar-Pescador, 2009). By contrast, GIS-based solar radiation models (Table 1)—such as Solar Analyst (SA) (Fu & Rich, 2002), SRAD (Wilson & Gallant, 2000), Solei-32 (Mészároš & Miklánek, 2006), and r.SUN (Hofierka & Suŕi, 2002)—have been developed to calculate the incoming solar radiation for each cell of a digital elevation model (DEM) during recent decades. These GIS-based models are technologically interoperable and scientifically rigorous, but they use different algorithms (either physically based or empirically based), thus their results show large differences in estimating solar radiation (Ruiz-Arias et al., 2009). Two limitations of these GIS-based solar radiation models are that they are computationally demanding and that they have difficulty incorporating dynamic factors that contribute to solar radiation estimates (Freitas, Catita, Redweik, & Brito, 2015). In particular, SA is a GIS-based sun–earth geometric model, but it ignores reflected radiation from nearby surfaces. However, accounting for reflected radiation

TABLE 1 A comparison between the proposed tool and related tools

DEM-based model	Environment	Computing capacity	Ground parameter	Sky view factor
Our tool	GEE	Cloud-based unlimited	Dynamic albedo	Surrounding terrain
r.SUN	GRASS	Multi-processor limited	Static albedo	Slope itself
SRAD	ArcGIS	Multi-processor limited	Static albedo	Slope itself
Solar Analyst	ArcGIS	Single processor limited	Not included	Slope itself
Solei-32	DOS	Single processor limited	Static albedo	Slope itself

is vital at locations with high albedo due to snow cover, because any variation in snow-cover albedo can have a great impact on solar radiation (He, Liang, & Song, 2014). Unlike Solar Analyst, SRAD estimates reflected radiation, but its reliability declines when monthly average cloudiness and sunshine hours are used to adjust daily shortwave radiation. Furthermore, processing of large-scale DEMs is not appropriate using Solei-32, Solar Analyst, or SRAD, all of which suffer from heavy computational demand with very large datasets (Tabik, Villegas, Zapata, & Romero, 2012). Additionally, both Solei-32 and r.SUN require appropriate parameters for estimating the atmospheric attenuation of incoming solar radiation, such as atmospheric transmissivity, the circumsolar coefficient, and atmospheric turbidity. However, vertical profiles of many atmospheric parameters are rarely available, especially in mountainous areas. Even when atmospheric parameters are available for these GIS-based tools, they consider only shelter effects due to the slope, whereas effects of the surrounding topography should be taken into account (Wang, Qiu, Wang, Wang, & Liu, 2014). Recent studies have found that satellite-based solar radiation estimates provide reasonable values and large spatial coverage. One weakness of satellite-based estimates resides in cloud detection, where even a small cloud can make solar radiation estimates less accurate. In addition, the accuracy of satellite-based solar radiation estimates for complex topography is still limited (Romano et al., 2018; Roupioz, Jia, Nerry, & Menenti, 2016; Tang et al., 2016; Yeom, Seo, Kim, & Han, 2016).

Much effort is therefore needed to build a computationally economical, next-generation GIS-based solar radiation model, which could explain influential impacts from albedo, surrounding terrain, and cloud. However, a GIS-based solar radiation model that allows for the treatment of high spatial and temporal variability in sun–earth position, terrain, and atmospheric effects has not yet been developed for monitoring daily solar radiation. In recent years, the advanced cloud-based geospatial computing platform, Google Earth Engine (Gorelick et al., 2017), has given researchers the opportunity to use big data for planetary-scale environmental data analysis. The present study covers this gap by complementing existing solar radiation studies with a dynamic spatial perspective, by incorporating the spatial heterogeneity of factors into a model and by applying cloud-based geospatial computing techniques to the problem.

In this study, DEM and land-surface albedo data were used to determine whether each point in the landscape was shaded by the surrounding terrain. A generic spatial and temporal gap-filling algorithm was then developed to retrieve seamless albedo datasets from the raw MODIS product (see Section 2). It is worth noting that other remote-sensing indices with missing values can also be gap-filled using this algorithm. A sunshine-based submodel was used in this study as a module to address actual radiation under the influence of clouds. An assessment of overall model accuracy was made by comparing our modeling results with ground-observed data and existing solar radiation products. The GIS-based model developed in this study has been released to the research community in a publicly available online platform, the spatial and temporal mountainous solar radiation (STMSR) model, after comparison with current GIS-based solar radiation modeling software. This online mountainous solar radiation model can be extended to other locations with around-the-world complex terrain.

2 | MATERIALS AND METHODS

2.1 | Study area and observed solar radiation data

The Loess Plateau is a 64 m ha, semi-arid region located in north-central China (33°43′–41°16′N, 100°54′–114°33′E) (Lü et al., 2012). The Loess Plateau has irregular topography with varying elevation between 422 and 3,390 m above mean sea level (Figure 1). Studying the topographic impact on solar radiation is of major importance on the Loess Plateau because of its distinct variation in topography. The Loess Plateau's extensive landscape is diverse. At the local scale, the terrain in the Loess Plateau includes eroded gully, near-vertical slopes, varying terraces, shoulders, and summits. For macro landforms, the diverse topography contains high mountains, rough hills, broken tablelands, and low plains. This region has played an increasingly important role in China's ecological

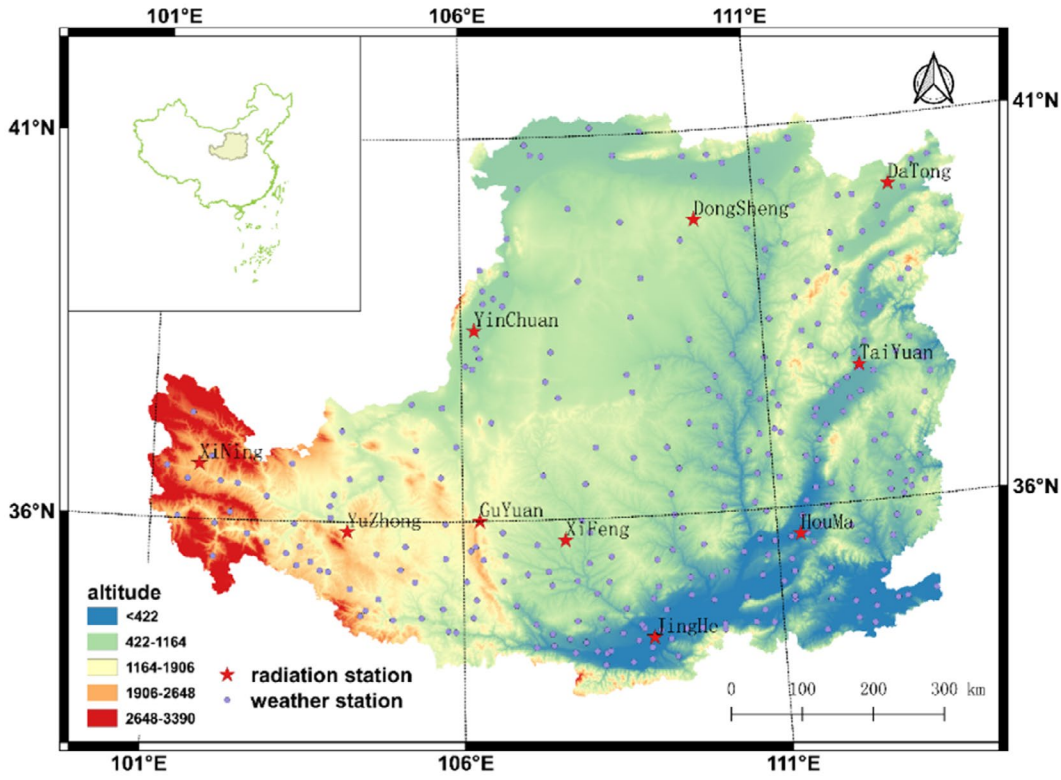

FIGURE 1 Study area showing the Loess Plateau located in north-central China along with 10 radiation stations and 301 weather stations

TABLE 2 Data sources for calculating and comparing solar radiation

Sequence number	Data name	Data time span	Data source
1	Sunshine data	2005–2014	China Meteorological Administration
2	DEM	2010	USGS/SRTMGL1_003
3	MODIS albedo	2000–2017	NASA LP DAAC at the USGS EROS Center
4	Surface solar radiation	2007–2014	Third Pole Environment Database
5	GLDAS2.1	1979–2018	NASA

security and natural resources supply (Zhao, Mu, Wen, Wang, & Gao, 2013). Since the ecological restoration projects such as "Natural Forest Protection" were implemented in this area, sloping cropland was converted to orchard land, and forest land has increased significantly. Simultaneously, there has been accelerated warming in the southwest region of the Loess Plateau (Sun, Miao, Duan, & Wang, 2015).

We acquired data from 301 Loess Plateau weather stations, carefully examined the data for quality, removed null values, and ingested the data into cloud storage. Other relevant data sources were DEM data and MODIS surface albedo from the cloud data catalog (Table 2) (i.e., SRTM Digital Elevation Data 90 m; Farr et al., 2007). During ingestion, DEM data were stored at various levels of resolution, from native resolution (90 m) to increasingly coarse levels. This was done by aggregating data in a pyramid structure such that pixel values of an upper level are the mean of pixels at the next lower level. The resolution of the DEM used for calculation was the closest scale equal to or less than the scale of the data source, with the coarsest native resolution in our analysis. The coarsest

resolution in our study was obtained from the daily MCD43A3.006 product, which provides 1 km resolution for black-sky and white-sky albedo across each of the MODIS surface reflectance bands (from band 1 to band 7), as well as three broad-spectrum bands (Schaaf & Wang, 2015).

GLDAS assimilates satellite and ground-based observational data products (Rodell et al., 2004) to generate land-surface parameters. This dataset supports agricultural and meteorological modeling. The GLDAS dataset started on January 1, 1948 and continues to the present time. The temporal and spatial resolution is 3 hr and 0.25°, respectively.

Land process research requires high spatial and temporal forcing data of surface solar radiation (SSR), which was derived by the fusion method of MODIS and MTSAT (Tang et al., 2016). MTSAT data includes MTSAT-1R and MTSAT-2, obtained from the Japanese Meteorological Agency. The temporal resolution of a MTSAT image is 30 min. A MTSAT image has five channels, and the spatial resolution for the visible sensor at nadir is 1 km, and for the other infrared sensors is 4 km. SSR was estimated by combining signals of polar-orbit (MODIS) and geo-stationary satellites (MTSAT).

2.2 | Schematic of the modeling

The GIS-based solar radiation model developed in this study (STMSR) can also be seen as a DEM-based model that integrates with a geospatial cloud-based computation platform to simulate the dynamics of solar radiation. The required inputs for the model include a DEM, MODIS albedo, in-situ observational data, and empirical coefficients. The entire process of modeling solar radiation in a mountainous terrain includes the three steps shown in Figure 2. The first step was to estimate extraterrestrial solar radiation and sky view factor on slopes in a high-performance computing (HPC) environment by using parallel raster image processing before uploading those image datasets to the cloud data catalog. The second step was to retrieve horizontal solar radiation data, including global solar

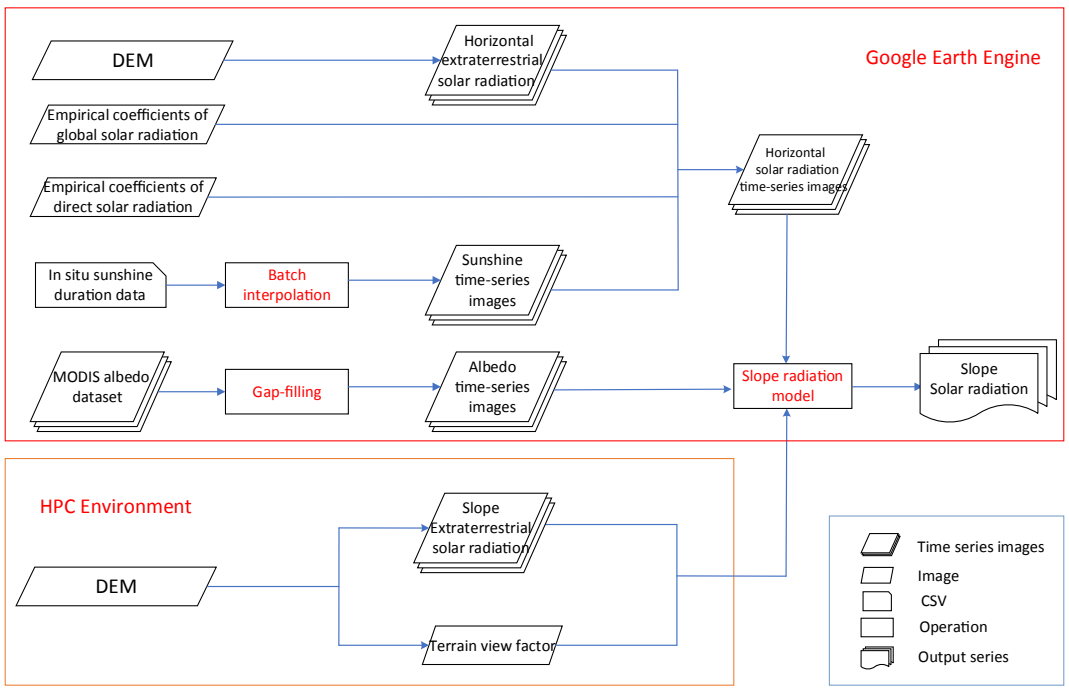

FIGURE 2 Flowchart of steps for calculation of solar radiation in mountainous terrain

radiation, direct solar radiation, diffuse solar radiation, and gap-filled MODIS albedo. The third step was to build a spatial and temporal mountain solar radiation model with those input parameters and create an online spatial and temporal mountain solar radiation modeling app.

2.3 | Distributed global solar radiation (GSR) model for rugged terrain

2.3.1 | Radiation on the horizontal surface

As reflected radiation on a horizontal surface is negligible, the radiation on a horizontal surface is partitioned into two parts, the beam and diffuse radiation, which are usually estimated by statistical regression of observed data (Liu et al., 2009):

$$K_b = B_h / Q_h \tag{1}$$

$$K_d = D_h / Q_h \tag{2}$$

$$K_b + K_d = 1 \tag{3}$$

where the direct radiation fraction K_b is called the direct radiation transmittance, and K_d is named the diffuse radiation fraction.

Since clouds are dynamic and site-specific, much observational data is required to parameterize cloud effects. Observations of routine meteorological variables such as sunshine and temperature do not require complicated instruments. A sunshine-based submodel is used in this study, because it produces better solar radiation estimates than cloud-based or temperature-based models (Iziomon & Mayer, 2001; Podestá, Núñez, Villanueva, & Skansi, 2004; Trnka, Žalud, Eitzinger, & Dubrovský, 2005). For example, the major limitation of cloud-based models is that they show systematically larger differences between measured and modeled values as cloud cover increases (Trnka et al., 2005).K_d is derived as a polynomial function of sunshine duration (Iqbal, 1983). Q_h is often derived from sunshine duration percentage using the Ångström formula (Angstrom, 1927). B_h is a polynomial function of relative sunshine duration (Louche, Notton, Poggi, & Simonnot, 1991). A further step is that Zeng, Qiu, Liu, and Jiang (2005) established an exponential function of direct radiation and global horizontal radiation:

$$Q_h = (a_h + b_h s)\, Q_{sh} \tag{4}$$

$$B_h = (1-a)\left(1 - e^{\frac{-bs^c}{(1-s)}}\right) Q_h \tag{5}$$

where Q_h is the horizontal solar radiation (MJ m^{-2} day^{-1}), Q_{sh} is the horizontal extraterrestrial radiation (MJ m^{-2} day^{-1}), s is the relative sunshine duration (i.e., the ratio of daily bright sunshine duration to the maximum possible duration of sunshine in daylight hours), a_h, b_h, a, b, and c are regression coefficients. The coefficients a_h and b_h in the Ångström equation were calibrated individually for each station in China using monthly observations. The calibration of direct radiation coefficients was achieved using least-squares linear regression of Q_{sh}/Q_h against s, where Q_{sh}, Q_h, and s are monthly mean global solar radiation (MJ m^{-2} day^{-1}), monthly mean extraterrestrial solar radiation (MJ m^{-2} day^{-1}), and monthly mean relative sunshine duration, respectively. Similarly, the coefficients of a, b, and c in the horizontal diffuse solar radiation model were determined from each month's observation (i.e., January, February, etc.), generating a set of coefficients for each month. The IDW interpolation method was then used to derive the spatial distribution of calibrated coefficients.

2.3.2 | Radiation on the inclined surface

Global solar radiation on an inclined surface was calculated as the sum of direct, diffuse, and reflected radiation from all sectors. This process was repeated for each grid cell in the DEM, thus producing an insolation map. Global solar radiation based on a DEM can be expressed as:

$$Q_{\beta w} = B_{\beta w} + D_{\beta w} + R_{\beta w} \tag{6}$$

where $Q_{\beta w}$ is total solar radiation for rugged terrain. The direct, diffuse, and reflected solar radiation components within rugged terrain are $B_{\beta w}$, $D_{\beta w}$, and $R_{\beta w}$, respectively.

Similar to the clear-sky conditions on a horizontal surface, direct transmittance K_b was used to solve the atmospheric attenuation of direct radiation on a rough surface (Liu et al., 2012). Direct irradiance on the inclined surface can be expressed as:

$$B_{\beta w} = \frac{Q_{sw}}{Q_{sh}} B_h \tag{7}$$

where Q_{sw} is slope extraterrestrial solar radiation.

In general, diffuse radiation coming from the sky is anisotropic. However, the calculation of anisotropy on a slope is complex and challenging (Dubayah & Rich, 1995). To simplify the calculation, diffuse radiation is divided into two parts: (1) from solar illumination direction; (2) from isotropic modeling. The diffuse radiation is given by Zeng, Qiu, He, and Liu (2008) as:

$$D_{\beta w} = D_h \left[K_b Q_{sw} / Q_{sh} + V \left(1 - K_b\right) \right] \tag{8}$$

When $k_b \to 0$, the sky is overcast and radiation is calculated from the isotropic model; when $k_b \to 1$, radiation is primarily from direct beam radiation. V is the sky view factor, which is associated with each grid cell. The detailed calculation process of V is illustrated in the Supporting Information.

Radiation that is reflected from nearby surfaces (e.g., mountains) is a function of albedo, the sky view factor, and horizontal solar radiation. The sky view factor is defined by the proportion of unobstructed sky over a horizontal surface such that $V = 0$ if the view of the sky is completely obstructed at a given location (Fu & Rich, 2002). Reflected radiation from nearby surfaces is calculated as

$$R_{\beta w} = Q_h \rho \left(1 - V\right) \tag{9}$$

where $R_{\beta w}$ is radiation reflected by surrounding cells, ρ is surface albedo, in which ρ was determined as the ratio of reflected to incident solar radiation at the surface.

The algorithm for solar radiation over rugged terrain is calculated per pixel using iterative calculations for sunshine duration and sky view factor. Currently, the front-end JavaScript programming and back-end GIS functions are not powerful enough in cloud computing platforms to implement intensive and iterative algorithms (Gorelick et al., 2017). To quickly obtain daily extraterrestrial solar radiation data over the vast area of the Loess Plateau (*c.* 1×10^6 km^2), a parallel extraterrestrial solar radiation algorithm on a local HPC environment was developed using Python multiprocessing and the GDAL package for parallel processing. First, the multi-band image was split into tiles equaling 90% of the HPC cores. After running the algorithm, the Mosaic tool was used to combine the resulting tiles into complete and seamless time series of extraterrestrial solar radiation images of the Loess Plateau and sky view factor images.

2.4 | Spatial and temporal MODIS albedo gap filling

Albedo is composed of direct and scattered radiation components. Therefore, actual clear-sky albedo can be calculated by a linear combination of direct and scattered albedo:

$$\partial = (1 - S)\,\partial_h + S\partial_b \tag{10}$$

where ∂ is land surface broadband albedo, ∂_h is broadband black-sky albedo, ∂_b is broad white-sky albedo, and S is a conversion coefficient.

The spatial resolution of MODIS albedo is 1,000 m, which is much finer than the surface albedo data resolution of 8 km obtained by NOAA/AVHRR data. As seen in Figure 3, the gap-filled product showed a similar spatial distribution for high and low extremes of surface albedo. Albedo of various land types exhibited differences (Figure 4). For example, very high albedo was observed in desert lands (up to 0.7; Figure 4) during periods with snow cover. Outside of snow-covered periods, cropland, desert, and grassland land types showed a relatively stable albedo, with the highest albedo for desert landscapes (0.41 ± 0.23), intermediate albedo for crops (0.14 ± 0.03) and grasslands (0.13 ± 0.03), and lowest for forests (0.10 ± 0.04).

The main problem encountered in applying GIS-based methods is missing values of MODIS albedo. We overcame this problem using spatio-temporal correlation of ungapped MODIS albedo data, interpolating across gaps using the surrounding data and producing a seamless dataset. This gap-filling method comprised five steps: (1) retrieve albedo data from the original MODIS albedo product; (2) create a mask of missing albedo data; (3) use a spatial neighborhood interpolation method to fill the missing data; (4) stack the yearly albedo image collection into time series; (5) apply the Whittaker algorithm to smooth the time series. This method was used to fill missing values without modifying existing values.

The Whittaker algorithm is based on penalized least squares, proposed by Whittaker 100 years ago (cited in Eilers, 2003). The Whittaker smoother has many advantages: it is extremely fast, much faster than the Savitzky–Golay filter in preliminary tests; it handles missing values efficiently; and it allows for full control over smoothness parameters (Eilers, 2003):

$$\left(W + \lambda D_d' D_d\right) z = Wy \tag{11}$$

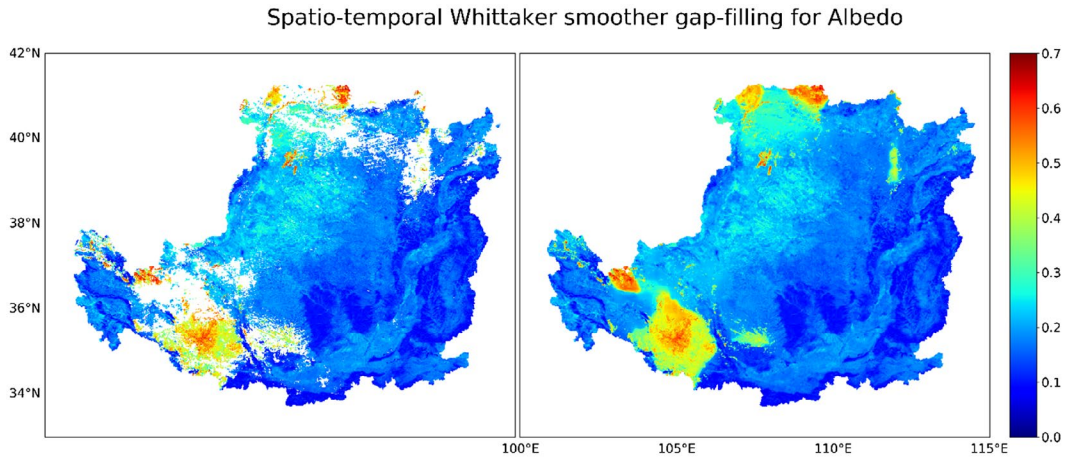

FIGURE 3 Albedo map of Loess Plateau at January 1, 2011 shown as an example of gap filling. Left panel: missing values (white) in the northern and western regions of the plateau. Right panel: Whittaker smoother gap-filled albedo map

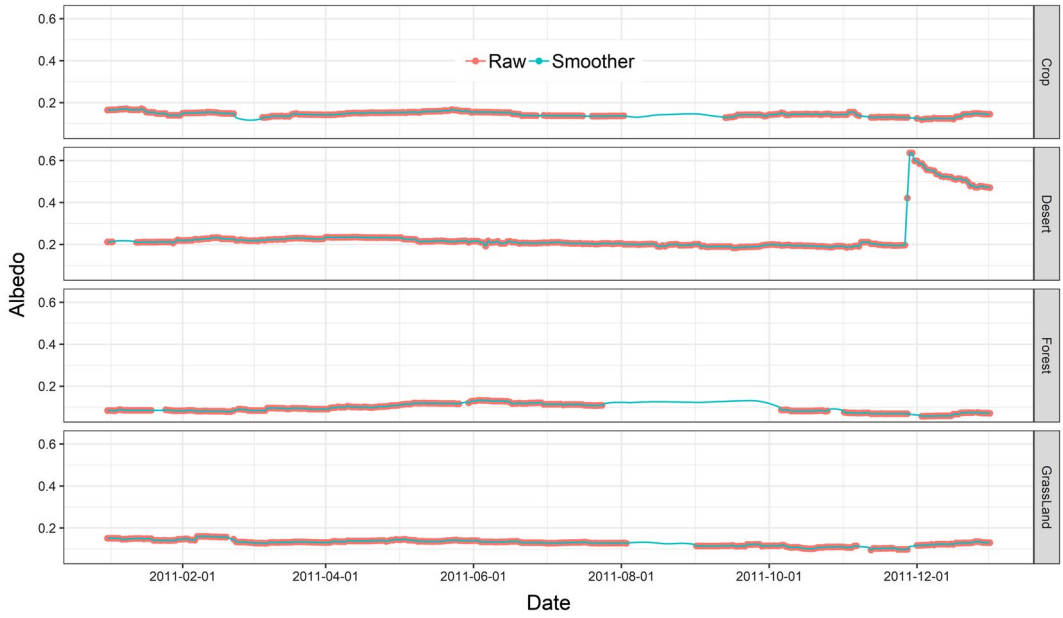

FIGURE 4 Variation of daily albedo for different land types. Missing values in raw albedo images were filled by a spatio-temporal gap-filling method. Those gap values in the curves were fitted by the Whittaker smoother method, with λ = 20, iterative = 3

The missing elements of *y* are set to zero, and the diagonal elements of the weights matrix *W* are set to zero for missing data and one for other values. At each missing point where *y* is zero, *z* was smoothed using Equation (11). *D* is a matrix such that $\mathrm{D}z = \Delta z$, and the subscript *d* represents the order of differences.

2.5 | Model evaluation

To evaluate the accuracy of the gap-fill predictions, the random knockout validation method (Gerber, de Jong, Schaepman, Schaepman-Strub, & Furrer, 2018) was applied during one year (2011) to a rectangle that had fewer than 50% missing observations in the original data. 10 points were selected, representing 50, 60, 70, 80, and 90% of the available original observations. 10 locations were randomly chosen from a validation area (Figure 5), and 10 temporal observations were randomly removed from each of these 10 time series (one time series per location). These missing values that were removed from the data were then filled using the spatial and temporal gap-filling algorithm described above, and gap-filled values were compared to the originally removed observations. Figure 5 shows that gap-filled values at the 10 randomly chosen observation sites were significantly correlated with observed data, resulting in R^2 = 0.962 and RMSE = 0.005 at daily timescales.

Validating and assessing the overall accuracy of STMSR based on only a few stations is inadequate for such a large region as the Loess Plateau. A comparison of STMSR estimates against those of other independent solar radiation products over China can provide an alternative approach for a regional evaluation of model performance. The comparisons include three key steps. First, daily SSR values were integrated from hourly values. Second, 1,000 locations were randomly created across the extent of the Loess Plateau for comparison. Third, daily SSR values of the points in different datasets were retrieved from those 1,000 locations using the point sampling method. This procedure was performed at the same 1,000 locations for STMSR and SSR, where the RMSE between the two models was evaluated across the entire Loess Plateau.

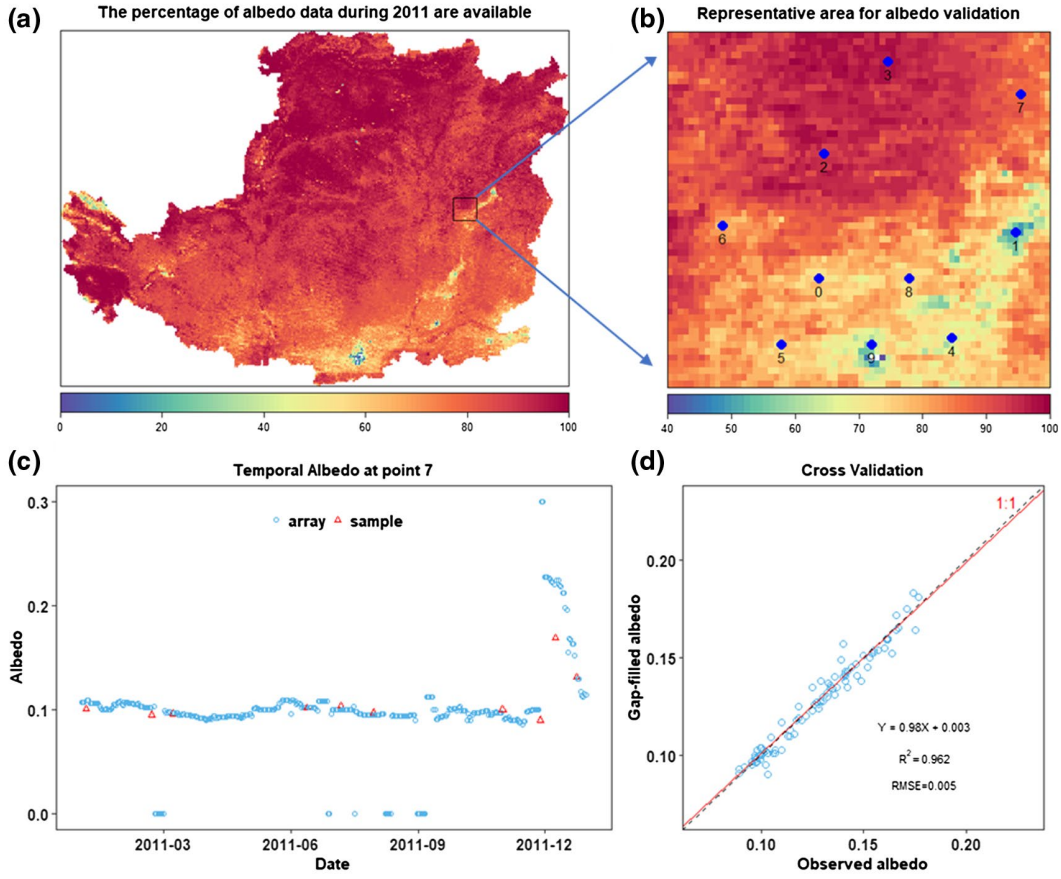

FIGURE 5 The percentage of albedo data during 2011 for the whole Loess Plateau (a), a representative validation area with 10 points (b), the temporal variation of daily albedo at point 7 with 10 randomly observed albedo (c), and cross validation for 100 samples during 2011 (d)

The performance of STMSR was also compared with two other GIS tools—SA in ArcGIS and r.SUN in GRASS. In SA, the Points Solar Radiation tool was used to calculate time series of global solar radiation simulations at the Yuzhong site because it contained continuous global direct and diffuse solar radiation datasets. The diffuse proportion of the radiation parameter was set to 0.4 under clear-sky conditions, and the transmittivity parameter was set to a default value of 0.5. The r.SUN program in GRASS cannot be used to simulate point radiation, but it is able to input parameters in raster format. Thus, a small patch of DEM around the Yuzhong site (400 km^2) was clipped to simulate daily global solar radiation. Albedo was set to the default value of 0.2, and the linke turbidity coefficient was set to an annual average value of 1.9.

R^2 is a statistic that describes the goodness-of-fit for a model, while RMSE is used to measure the difference between values predicted by a model and those which were actually observed. We used these two statistical criteria to validate our model. The two validation measurements were calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_i - O_i\right)^2} \tag{12}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n} \left(O_i - \overline{O}\right)\left(P_i - \overline{P}\right)}{\sqrt{\sum_{i=1}^{n} \left(O_i - \overline{O}\right)^2} \sqrt{\sum_{i=1}^{n} \left(P_i - \overline{P}\right)^2}} \right)^2 \quad (13)$$

where P_i and O_i are the predicted and observed daily SSR, respectively, $\overline{O}$ and $\overline{P}$ are the mean daily SSR, i is the ith sample, and n is the number of samples.

3 | RESULTS

3.1 | Model validation in the Loess Plateau

Different timescales have Ångström coefficients (a_h and b_h) of varying magnitude. Previous studies have shown that a better fit between n/N and R_s/R_a can be obtained using monthly data than yearly data. In this study, coefficients used for horizontal global and direct-beam radiation models were obtained from monthly in-situ radiation data, which were obtained from a previous study (Zeng et al., 2005).

To validate the performance of our model, observed data from 10 solar radiation stations were used (Figure 1), of which one station (the Yuzhong station) included measurements of both direct and diffuse solar radiation. Model performance for simulation of monthly solar radiation was evaluated for the period 2005–2009. In terms of global solar radiation, simulated monthly global solar radiation matched well with observations (Figure 6). Figure 6 shows that: (1) our model simulations at the 10 observation sites were significantly correlated with observed global radiation, resulting in high R^2 ($R^2 \geq 0.9$) and low RMSE (RMSE ≤ 45 MJ m^{-2} $month^{-1}$); and (2) the slopes were within ±10% of the 1:1 line across all study locations. At these 10 observation sites, our model performed very well in the Loess Plateau and can be used further to generate highly accurate solar radiation estimates for mountainous locations with local calibration/validation data.

3.2 | Comparison with other SSR and GSR products

Figure 7 illustrates the annual mountain solar radiation spatial map from STMSR, the SSR product, and the GLDAS net shortwave radiation product. In comparison to SSR, STMSR produced higher estimates of solar radiation in the drylands of the northwest Loess Plateau and lower estimates in the mountains to the south (cf. Figures 7a and b). GLDAS net shortwave radiation values (lower left panel) showed little consistency with spatial patterns in STMSR or SSR and little association with topography (Figure 7). Maximum radiation was highest in STMSR (*c.* 7,000 MJ m^{-2} day^{-1}), intermediate in SSR (*c.* 6,500 MJ m^{-2} day^{-1}), and lowest for GLDAS (*c.* 5,500 MJ m^{-2} day^{-1}; Figure 7). STMSR similarly produced the lowest minimum radiation values (*c.* 3,500 MJ m^{-2} day^{-1}). Figure 8 shows performance comparisons between STMSR, SSR, and GLDAS on a daily timescale, illustrating that solar radiation estimates from the current study were better than those from the other products. R^2 for STMSR (0.88) was better than that of the other two products (0.76–0.84), although R^2 for all was quite good (Figure 8). However, only two of the products (STMSR and SSR) showed a 1:1 response against observations (Figure 8). Overall, radiation estimates simulated by STMSR were slightly improved relative to SSR and greatly improved relative to GLDAS. We observed that STMSR–SSR RMSE (i.e., RMSE between two derived products, not observations) increased from north to south, indicating an increasing discrepancy between radiation products towards the south (Figure 9). Following this trend, RMSE was slightly higher in the Guanzhong Plain (in the southeastern Loess Plateau) than in the mountains extending to the north along the eastern boundary of the Loess Plateau (cf. Figures 1 and 9). RMSE was also large in the westernmost region of the Loess Plateau, which has the highest elevations and steepest

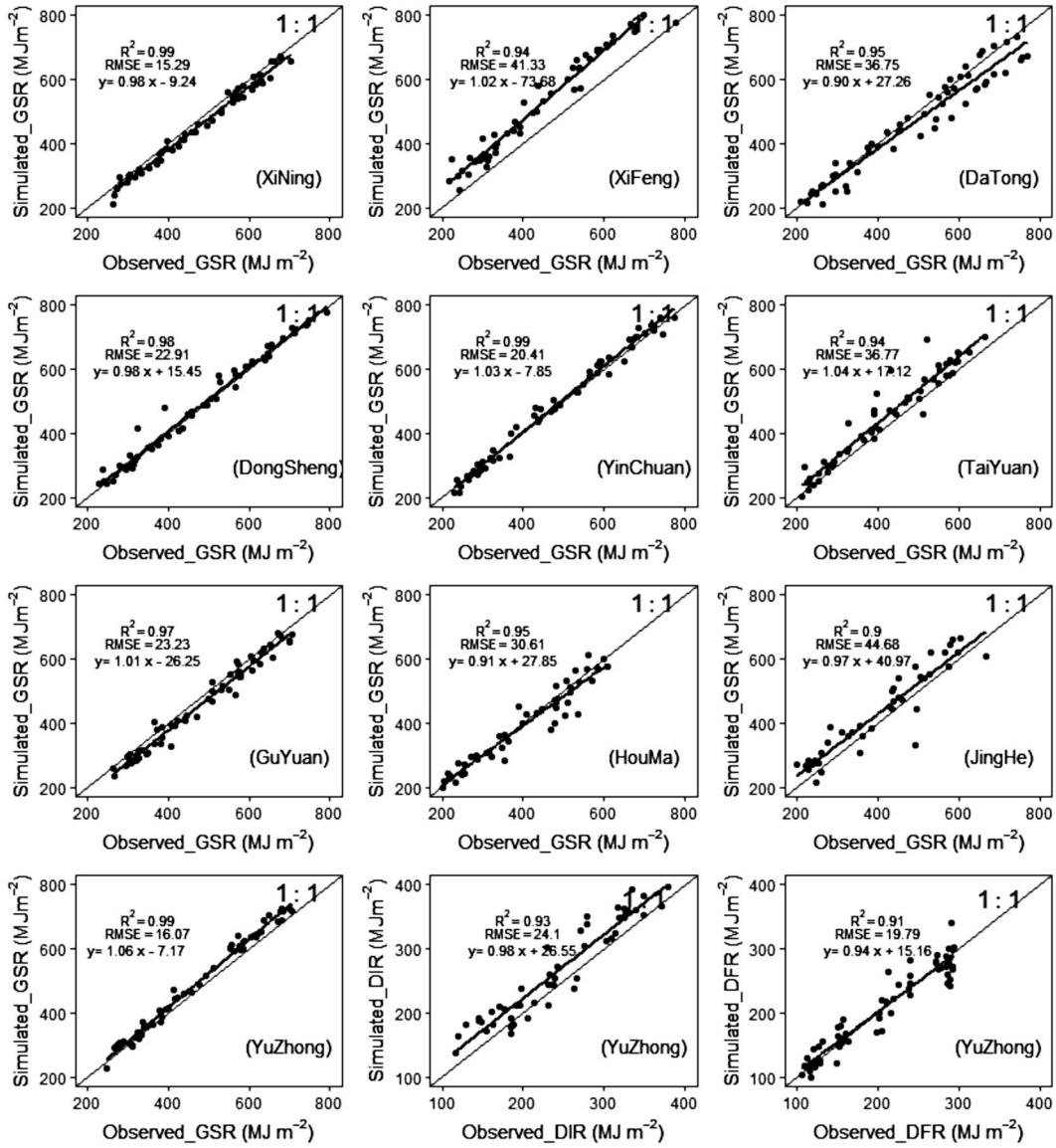

FIGURE 6 Comparison of annual observed and estimated (by mountain solar radiation model) monthly global solar radiation (GSR) for 10 radiation sites on the Loess Plateau, China, during 2005 to 2009. Comparisons for direct radiation (DIR) and diffuse radiation (DFR) are shown only for YuZhong

slopes. Small discrepancies between STMSR and SSR in the central Loess Plateau suggest that both products produce reasonable radiation estimates. By contrast, larger discrepancies in the Guanzhong Plain and the western mountains suggest an improvement on radiation estimates by STMSR in these areas; thus STMSR can be used further to generate realistic solar radiation maps for mountain and valley regions.

Figure 10 shows a comparison of three algorithms against observed values at Yuzhong in 2009. The SA and r.SUN algorithms clearly overestimated radiation in the middle of the year (March–October), and SA underestimated radiation in the winter months (November–February). In contrast, our STMSR model slightly overestimated observed radiation in March–October, but closely predicted monthly global solar radiation across the remaining months. It should be noted that finer tuning of input parameters, such as direct transmittance and

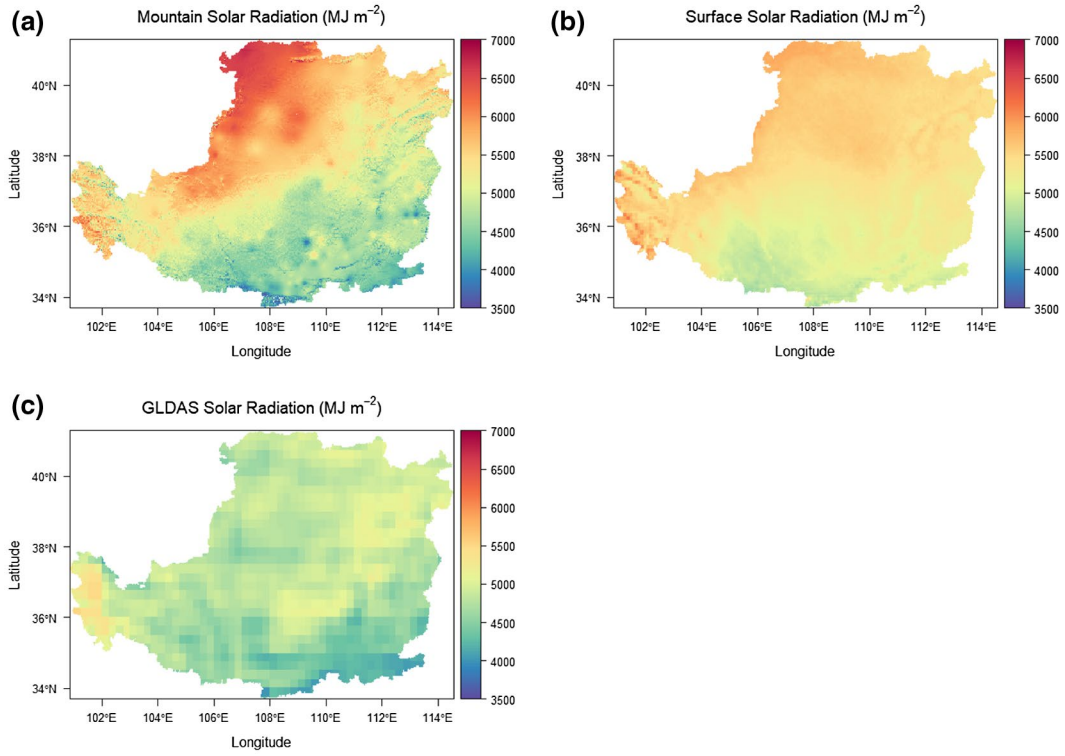

FIGURE 7 Spatial distributions of yearly solar radiation on the Loess Plateau in 2011 by mountain solar radiation produced by STMSR model, surface solar radiation, and GLDAS

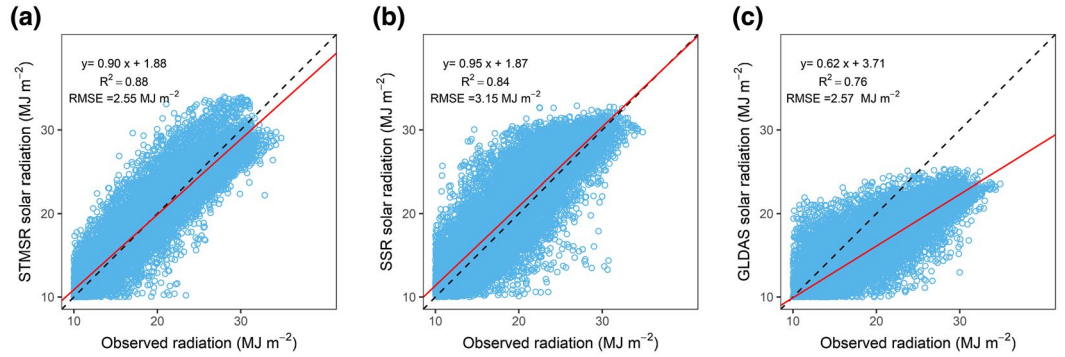

FIGURE 8 Summary statistics for estimated daily solar radiation produced by the STMSR model (a), the SSR model (b), and the GLDAS model (c) compared with observed data across 10 solar radiation stations in 2007–2013

diffuse proportion in SA or default atmospheric parameters in r.SUN, might result in improved estimates from those products.

4 | DISCUSSION

The online tool (STMSR) we provide as part of this study is available for use in complex terrain globally (https://geogismx.users.earthengine.app/view/stmsr) (Figure 11). In the left panel, users can define a time period along

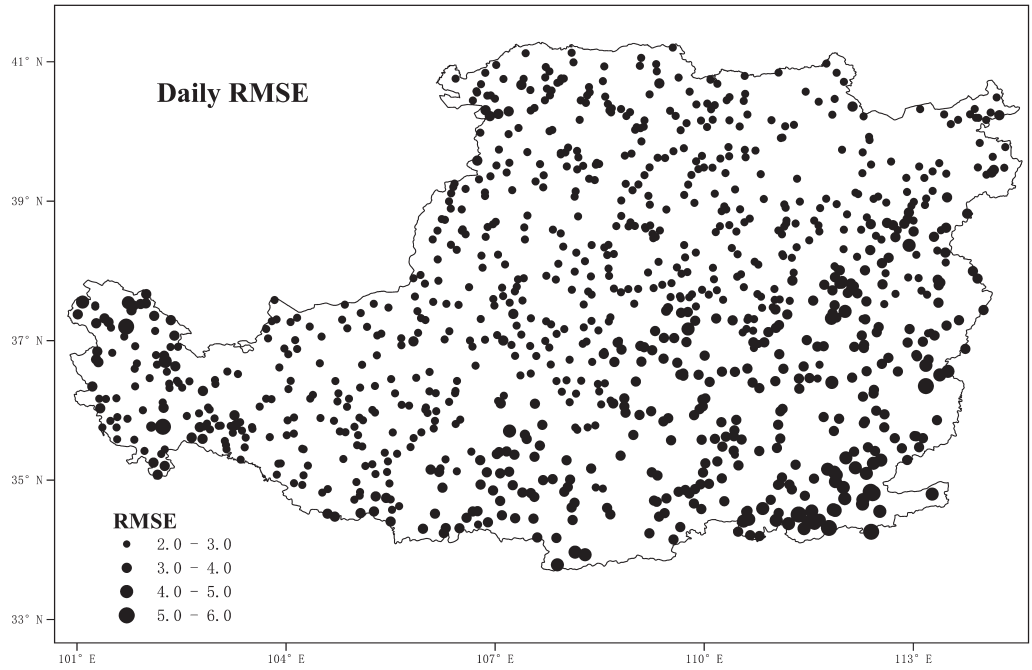

FIGURE 9 Spatial distributions of RMSE calculated between the daily solar radiation of the STMSR model and SSR at 1,000 randomly selected points in 2011 over the Losses Plateau. Circle diameters correspond to the size of RMSE. RMSE units MJ m^{-2}

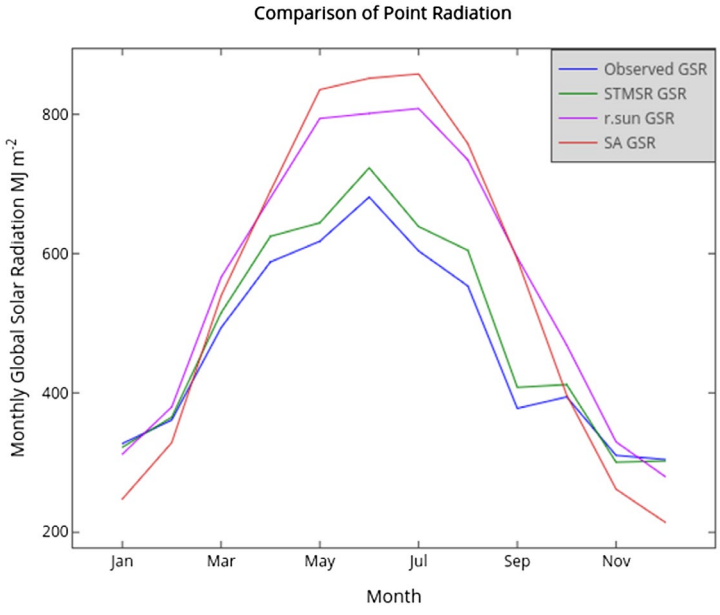

FIGURE 10 Comparison of different GSR products with in-situ observations at YuZhong in 2009. STMSR: spatio-temporal mountain solar radiation, SA: Solar Analyst, r.SUN: radiation integrated in GRASS

with a location by designating latitude and longitude, or by clicking on the map, and then the right panel will show a time series of point solar radiation, along with the three components of global solar radiation (direct, diffuse, and reflected). STMSR can also export composited images of astronomical solar radiation (i.e., the radiation which

would be incident at the planet's surface in the absence of an atmosphere) and global solar radiation components, as shown in Figure 12.

The solar radiation estimated by our model performed better in the Loess Plateau than other SSR products and GLDAS, as quantified by R^2 and RMSE. This may be attributed to differences in satellite data sources, methods, and the scales of prediction. Bias in satellite-based models depends on the clear-sky index and solar zenith angle, together with atmospheric parameters such as aerosol, ozone, and precipitable water (PW). Polar-orbiting satellites such as MODIS only measure instantaneous values, which are then extrapolated to daily solar radiation values using a sinusoidal function. This approximation is likely to incur a larger error than that provided by geostationary satellites like MTSAT (Qin et al., 2015; Roupioz et al., 2016). Even if MODIS and MTSAT measurements were integrated together to improve satellite-based solar radiation estimates, parameterization of the topographic correction remains overly simple. However, the time delay between atmospheric parameters derived from MODIS and actual cloud variation can lead to significant errors. The spatial resolution of our mountain solar radiation estimates (about 1 km) is much finer than that of GLDAS net shortwave radiation data (0.25°) and SSR (5 km). The coarse spatial resolution of these two radiation products resulted in larger mean error relative to in-situ measurements than for STMSR.

Our open source GIS-based model STMSR has both advantages and disadvantages. As described, one bottleneck was that GIS functions on the back end of cloud computing are not powerful enough to implement iterative algorithms. In this study, some of those calculations were performed beforehand (e.g., duration of possible sunshine on slopes, the sky view factor). Then, they were uploaded to cloud storage for users to access widely and for incorporation in a cloud-based library of solar radiation models to decrease processing times dramatically. Another potential difficulty was that our model could be set up only when the coefficients of the sunshine-based model were available. However, this is not a problem because many solar radiation sites worldwide make available calibrated and accurate local direct and diffuse coefficients for sunshine-based models (Liu et al., 2009; Trnka et al., 2005). In such a case, the empirical parameterization scheme used in our model proved to be an economical and practical method for estimating actual solar radiation from sunshine hours under the influence of cloud cover.

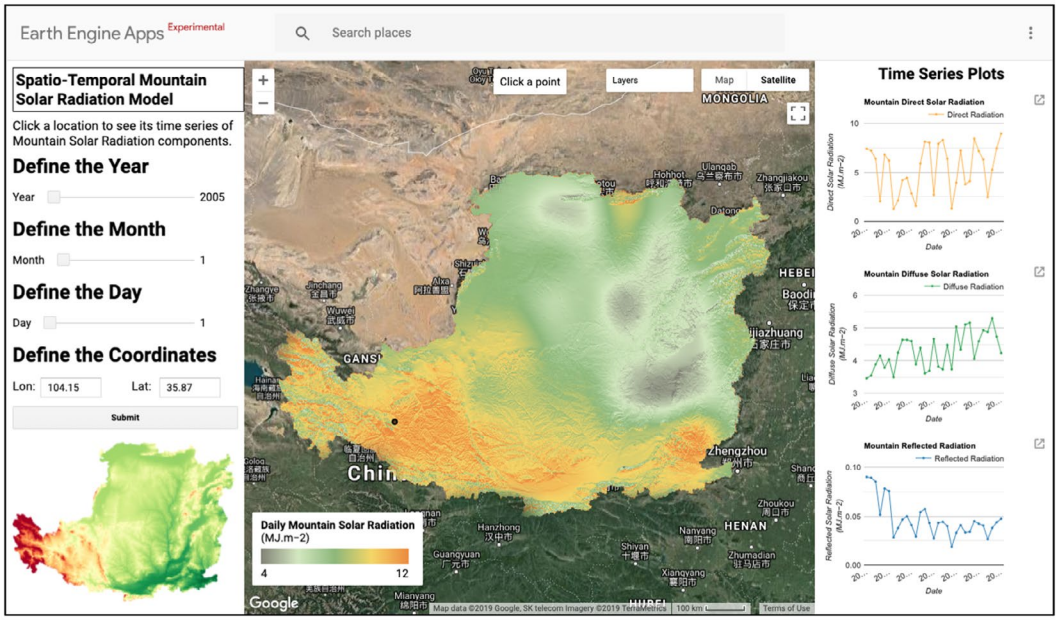

FIGURE 11 Application interface for the mountain solar radiation model on the Google Earth Engine APP platform

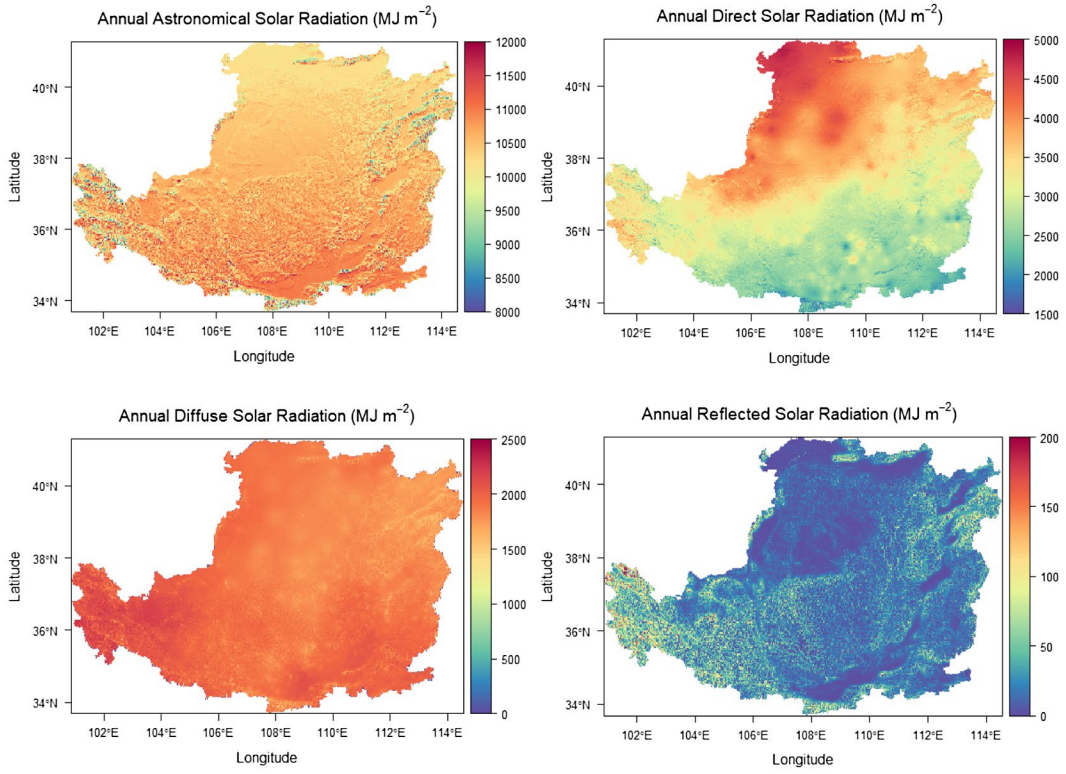

FIGURE 12 Spatial distribution of annual astronomical solar radiation, direct solar radiation, diffuse solar radiation, and reflected solar radiation in 2011 over the Loess Plateau

By contrast, satellite-based methods provide an advantage for retrieving atmospheric parameters from ungauged areas. Zhang et al. (2015) used two atmosphere products from MODIS—aerosol optical depth and PW—as input parameters for solar radiation modeling to decrease atmospheric estimation errors. Concerning the rapidly rising array of satellite products available, integration with more atmospheric products would be an important asset for future research.

As shown in Figure 3, the available daily MODIS albedo can be an issue in some regions for GIS-based solar radiation models due to data scarcity. Roupioz et al. (2016) chose to use the 8-day composite MODIS albedo product for the daily solar radiation modeling. However, this 8-day resolution is too coarse for investigating rapid changes in albedo over the Tibetan Plateau. We overcame this problem by developing a spatial and temporal gap-filling algorithm to provide a seamless daily albedo dataset for estimating variations in solar radiation. This seamless dataset made possible the quick estimation of albedo over snowy landscapes, also providing further capabilities such as smoothing of other ecological indices and extracting phenological characteristics from data types such as NDVI, EVI, or LAI (Pan, Hu, & Cao, 2017). However, care must be taken when selecting the smoothing parameter (λ) in the Whittaker algorithm, which is very sensitive to this parameter. In this study, λ was determined by trial and error to be equal to 20, but further research is needed to evaluate the relationship between λ and kurtosis, mean, and variance.

We found divergence between STMSR and SSR in the southern and western Loess Plateau (Figure 9). Large uncertainty in the southern portion of the Loess Plateau could be due to complicated cloud distribution, which reduces the accuracy of cloud parameter estimates, potentially leading to substantial errors in SSR estimation (Tang et al., 2016). By contrast, uncertainty in the western Loess Plateau was likely related to shading and surface

inclination effects, both of the surface itself and in the adjacent terrain (Liu et al., 2012). Furthermore, spatial interpolation of regression coefficients across cloudy or mountainous regions can still be problematic (Liu et al., 2017), even though interpolation can be valid across some regions where the atmospheric turbidity is similar (e.g., across the central Loess Plateau; Figure 9). To overcome uncertainty due to interpolation errors, improved spatial and temporal interpolation of complex calibrated coefficients and sunshine hours in future studies could be achieved through integration of a geographical and temporal weighted regression (Fotheringham, Crespo, & Yao, 2015). Many factors can affect RMSE between radiation products, including interpolation of Ångström model coefficients, spatial variability in elevation, water vapor content, and other climate characteristics (Liu et al., 2017), reflective features of the surface, cloud contamination, aerosols, and atmospheric water vapor (Stocker, 2014).

Since the "Grain-for-Green" program has been implemented, large areas of re-vegetated land are now present in southern and eastern parts of the Loess Plateau (Zhang et al., 2018). Based on SSR theory, land use/land cover (LULC) can change outgoing/reflected shortwave radiation and absorbed shortwave radiation by changing land surface albedo. However, it remains unclear whether LULC can change the incoming shortwave radiation reaching the land surface. LULC changes can be estimated by mean vegetation cover during the growing season, where vegetation cover is estimated by the normalized difference vegetation index (NDVI). Future research could focus on changes in vegetation cover during the growing season, for exploring the impact of LULC changes on solar radiation.

5 | CONCLUSIONS

We developed an improved GIS-based solar radiation model (STMSR) that allows for treatment of high spatial and temporal variations in albedo, surrouding terrain shading, and cloud cover for monitoring daily solar radiation at large scale. By comparison with other well-known GIS-based solar radiation models such as SA in ArcGIS and r.SUN in GRASS, our STMSR model showed better performance. The resulting estimates of global, direct, and diffuse solar radiation were validated with high estimation accuracy against the measured solar radiation data from 10 observation stations across the Loess Plateau. Compared with other high-resolution solar radiation datasets, the global solar radiation presented in this article has higher accuracy of daily solar radiation estimates over the Loess Plateau than other methods, generating higher R^2 and RMSE. Our STMSR model also has the potential to be applied globally for distributed modeling applications across a variety of landscapes.

ACKNOWLEDGMENTS

This research was supported by the Natural Science Foundation of China (Grant No. 41730645). The first author acknowledges the Chinese Scholarship Council for providing a PhD scholarship. We also acknowledge the Data Assimilation and Modeling Center for Tibetan Multi-spheres for providing the high-resolution surface solar radiation dataset. Facilities for conducting this study were provided by the New South Wales Department of Primary Industries and the University of Technology, Sydney.

CODE AVAILABILITY

The code to generate the final solar radiation products is available at https://code.earthengine.google.com/?accept_repo=users/geogismx/mountain_solar and https://github.com/geogismx/mountain-solar-radiation-map

ORCID

Mingxi Zhang https://orcid.org/0000-0002-5431-9274

REFERENCES

Aguilar, C., Herrero, J., & Polo, M. J. (2010). Topographic effects on solar radiation distribution in mountainous watersheds and their influence on reference evapotranspiration estimates at watershed scale. *Hydrology & Earth System Sciences*, *14*(12), 2479–2494.

Alsamamra, H., Ruiz-Arias, J. A., Pozo-Vázquez, D., & Tovar-Pescador, J. (2009). A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain. *Agricultural & Forest Meteorology*, *149*(8), 1343–1357.

Angstrom, A. (1927). Solar and terrestrial radiation. *Quarterly Journal of the Royal Meteorological Society*, *50*, 121–126.

Brock, T. D. (1981). Calculating solar radiation for ecological studies. *Ecological Modeling*, *14*(1–2), 1–19.

Dubayah, R., & Rich, P. M. (1995). Topographic solar radiation models for GIS. *International Journal of Geographical Information Science*, *9*(4), 405–419.

Eilers, P. H. (2003). A perfect smoother. *Analytical Chemistry*, *75*(14), 3631–3636.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., … Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews in Geophysics*, *45*(2), RG2004.

Fotheringham, A. S., Crespo, R., & Yao, J. (2015). Geographical and temporal weighted regression (GTWR). *Geographical Analysis*, *47*(4), 431–452.

Freitas, S., Catita, C., Redweik, P., & Brito, M. C. (2015). Modelling solar potential in the urban environment: State-of-the-art review. *Renewable & Sustainable Energy Reviews*, *41*, 915–931.

Fu, P., & Rich, P. M. (2002). A geometric solar radiation model with applications in agriculture and forestry. *Computers & Electronics in Agriculture*, *37*(1–3), 25–35.

Gerber, F., de Jong, R., Schaepman, M. E., Schaepman-Strub, G., & Furrer, R. (2018). Predicting missing values in spatio-temporal remote sensing data. *IEEE Transactions on Geoscience & Remote Sensing*, *56*, 2841–2853.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thaub, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing for Environment*, *202*, 18–27.

He, T., Liang, S., & Song, D. X. (2014). Analysis of global land surface albedo climatology and spatial–temporal variation during 1981–2010 from multiple satellite products. *Journal of Geophysical Research D: Atmospheres*, *119*(17), 10281–10298.

Hofierka, J., & Suŕi, M. (2002). The solar radiation model for open source GIS: Implementation and applications. In *Proceedings of the Open Source GIS-GRASS Users Conference*, Trento, Italy.

Iqbal, M. (1983). *An introduction to solar radiation*. New York, NY: Academic Press.

Iziomon, M. G., & Mayer, H. (2001). Performance of solar radiation models: A case study. *Agricultural & Forest Meteorology*, *110*(1), 1–11.

Liu, J. D., Pan, T., Chen, D. L., Zhou, X. J., Yu, Q., Flerchinger, G., … Shen, Y. (2017). An improved Ångström-type model for estimating solar radiation over the Tibetan Plateau. *Energies*, *10*(7), 892.

Liu, M., Bárdossy, A., Li, J., & Jiang, Y. (2012). GIS-based modeling of topography-induced solar radiation variability in complex terrain for data sparse region. *International Journal of Geographic Information Science*, *26*(7), 1281–1308.

Liu, X., Mei, X., Li, Y., Zhang, Y., Wang, Q., Jensen, J. R., & Porter, J. R. (2009). Calibration of the Ångström-Prescott coefficients (*a*, *b*) under different time scales and their impacts in estimating global solar radiation in the Yellow River basin. *Agricultural & Forest Meteorology*, *149*(3–4), 697–710.

Louche, A., Notton, G., Poggi, P., & Simonnot, G. (1991). Correlations for direct normal and global horizontal irradiation on a French Mediterranean site. *Solar Energy*, *46*(4), 261–266.

Lü, Y., Fu, B., Feng, X., Zeng, Y., Liu, Y., Chang, R., …Wu, B. (2012). A policy-driven large scale ecological restoration: Quantifying ecosystem services changes in the Loess Plateau of China. *PLoS ONE*, *7*(2), e31782.

Mészároš, I., & Miklánek, P. (2006). Calculation of potential evapotranspiration based on solar radiation income modeling in mountainous areas. *Biologia*, *61*(Suppl. 19), S284–S288.

Pan, Z., Hu, Y., & Cao, B. (2017). Construction of smooth daily remote sensing time series data: A higher spatiotemporal resolution perspective. *Open Geospatial Data, Software & Standards*, *2*, 25.

Pintor, B. H., Sola, E. F., Teves, J., Inocencio, L. C., Ang, M., & Concepcion, R. (2015). Solar energy resource assessment using r.SUN in GRASS GIS and site suitability analysis using AHP for groundmounted solar photovoltaic (PV) farm in the Central Luzon Region (Region 3), Philippines. In *Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings*, Philippines, Vol. 15, Article 3.

Podestá, G. P., Núñez, L., Villanueva, C. A., & Skansi, M. A. (2004). Estimating daily solar radiation in the Argentine Pampas. *Agricultural & Forest Meteorology*, *123*(1–2), 41–53.

Qin, J., Tang, W., Yang, K., Lu, N., Niu, X., & Liang, S. (2015). An efficient physically based parameterization to derive surface solar irradiance based on satellite atmospheric products. *Journal of Geophysics Research D: Atmospheres*, *120*(10), 4975–4988.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., … Toll, D. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381–394.

Romano, F., Cimini, D., Cersosimo, A., Di Paola, F., Gallucci, D., Gentile, S., … Viggiano, M. (2018). Improvement in surface solar irradiance estimation using HRV/MSG data. *Remote Sensing*, *10*, 1288.

Roupioz, L., Jia, L., Nerry, F., & Menenti, M. (2016). Estimation of daily solar radiation budget at kilometer resolution over the Tibetan Plateau by integrating MODIS data products and a DEM. *Remote Sensing*, *8*(6), 504.

Ruiz-Arias, J. A., Tovar-Pescador, J., Pozo-Vázquez, D., & Alsamamra, H. (2009). A comparative analysis of DEM-based models to estimate the solar radiation in mountainous terrain. *International Journal of Geographical Information Science*, *23*(8), 1049–1076.

Schaaf, C., & Wang, Z. (2015). *MCD43A3: MODIS/Terra and Aqua Albedo Daily L3 Global 500 m SIN Grid V006 distributed by NASA EOSDIS Land Processes DAAC*. https://doi.org/10.5067/MODIS/MCD43A3.006

Stocker, T. (Ed.). (2014). *Climate Change 2013: The physical science basis: Working Group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.

Sun, Q., Miao, C., Duan, Q., & Wang, Y. (2015). Temperature and precipitation changes over the Loess Plateau between 1961 and 2011, based on high-density gauge observations. *Global & Planetary Change*, *132*, 1–10.

Tabik, S., Villegas, A., Zapata, E. L., & Romero, L. F. (2012). A fast GIS-tool to compute the maximum solar energy on very large terrains. *Procedia Computer Science*, *9*, 364–372.

Tang, W., Qin, J., Yang, K., Liu, S., Lu, N., & Niu, X. (2016). Retrieving high-resolution surface solar radiation with cloud parameters derived by combining MODIS and MTSAT data. *Atmospheric Chemistry & Physics*, *16*(4), 2543–2557.

Trnka, M., Žalud, Z., Eitzinger, J., & Dubrovský, M. (2005). Global solar radiation in Central European lowlands estimated by various empirical formulae. *Agricultural & Forest Meteorology*, *131*(1–2), 54–76.

Wang, L., Qiu, X., Wang, P., Wang, X., & Liu, A. (2014). Influence of complex topography on global solar radiation in the Yangtze River Basin. *Journal of Geographical Sciences*, *24*(6), 980–992.

Wilson, J. P., & J. C. Gallant (Eds.). (2000). *Terrain analysis: Principles and applications*. New York, NY: Wiley.

Yeom, J.-M., Seo, Y.-K., Kim, D.-S., & Han, K.-S. (2016). Solar radiation received by slopes using COMS imagery, a physically based radiation model, and GLOBE. *Journal of Sensors*, *2016*, 1–15.

Zeng, Y., Qiu, X., He, Y., & Liu, C. (2008). Distributed modeling of diffuse solar radiation over rugged terrain of the Yellow River basin. *Chinese Journal of Geophysics*, *51*(4), 700–708.

Zeng, Y., Qiu, X., Liu, C., & Jiang, A. (2005). Distributed modeling of direct solar radiation of rugged terrain over the Yellow River basin. *Acta Geographica Sinica*, *60*(4), 680–688.

Zhang, S., Yang, D., Yang, Y., Piao, S., Yang, H., Lei, H., & Fu, B. (2018). Excessive afforestation and soil drying on China's Loess Plateau. *Journal of Geophysical Research G: Biogeosciences*, *123*, 923–935.

Zhang, Y., Li, X., & Bai, Y. (2015). An integrated approach to estimate shortwave solar radiation on clear-sky days in rugged terrain using MODIS atmospheric products. *Solar Energy*, *113*, 347–357.

Zhao, G., Mu, X., Wen, Z., Wang, F., & Gao, P. (2013). Soil erosion, conservation, and eco environment changes in the Loess Plateau of China. *Land Degradation & Development*, *24*(5), 499–510.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Zhang M, Wang B, Liu DL, et al. Incorporating dynamic factors for improving a GIS-based solar radiation model. *Transactions in GIS*. 2020;24:423–441. https://doi.org/10.1111/tgis.12607