8-2016

Estimation of the instantaneous downward surface shortwave radiation using MODIS data in Lhasa for all-sky conditions

Xu Lu
International Development, Community and Environment, xulu2@clarku.edu

Estimation of the instantaneous downward surface shortwave radiation using MODIS data in Lhasa for all-sky conditions

Xu Lu

AUGUST 2016

A THESIS

Submitted to the faculty of Clark University, Worcester, Massachusetts, in partial fulfillment of the requirements for the degree of Master of Sciences in the department of IDCE

And accepted on the recommendation of

John Rogan, Chief Instructor

ABSTRACT

Estimation of the instantaneous downward surface shortwave radiation using MODIS data in Lhasa for all-sky conditions

Xu Lu

Measuring the solar irradiance with high accuracy is the basis of PV power forecasting. Although the downward surface shortwave radiation (DSSR) data derived from satellite images are widely used in the PV industry, the instantaneity and accuracy of these data are not suitable for PV power forecasting in a short-time period. In this study, an algorithm to calculate instantaneous DSSR for all-sky conditions was developed by combining clear-sky radiative transfer model and 3D radiative transfer model using MODIS products (MOD03-07, 09). The algorithm was evaluated by ground measurements from a station in Lhasa and a reference dataset from FLASHFlux. The results indicate that the errors of DSSR using combining model are less than FLASHFlux. The time consuming of running 3D radiative transfer model can be reduced by narrowing down the extent of input data to 8km.

John Rogan, Ph.D.
Associate Professor
Chief Instructor

ACADEMIC HISTORY

Name: Xu Lu   **Date: AUGUST 2016**

Baccalaureate Degree: Bachelor of Science

Source: Wuhan University, Hubei, China   **Date: JUNE 2014**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1   Introduction

1.1 Background

Energy is a fundamental requirement for all socio-economic activities of human and its well-being. Without adequate energy supply, people cannot enjoy the brilliant results of modern industrial civilization. Up to now, people still use fossil fuels as the major primary energy sources, which are not renewable sources, and the energy demand is constantly growing as the development of the world. In the International Energy Outlook 2013, U.S. Energy Information Administration (2013) predicted that the total energy consumption is projected to increase by 43.4% from 2015 to 2040 as shown in Figure 1. As shown in Figure 2, EIA predicted that fossil fuels would still account for the major parts of human energy consumption in the long run.
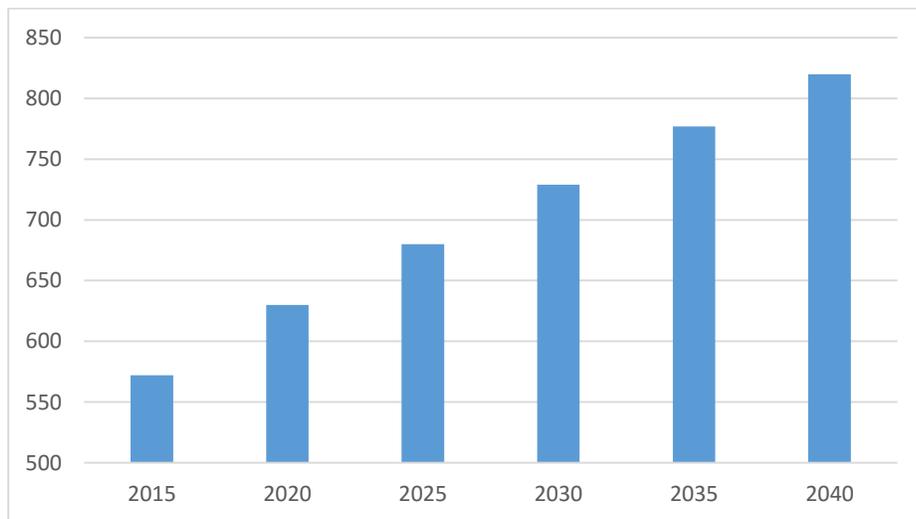
Figure 1 World total energy consumption (quadrillion Btu)
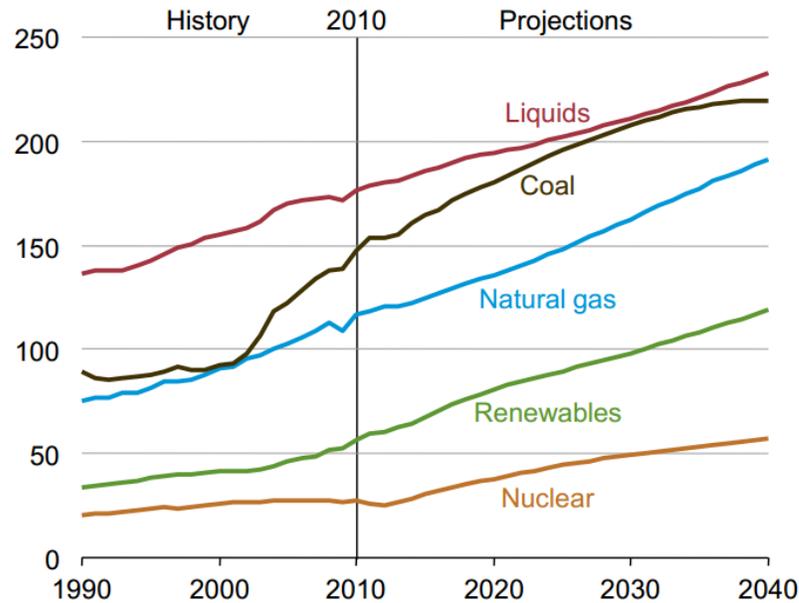
Figure 2 World energy consumption by fuel type from 1990 to 2040 (quadrillion Btu) (U.S. Energy Information Administration 2013)

Although fossil fuels would not be exhausted in the near future, the overuse of fossil fuels results in heavy pollution and high levels of CO_2 emissions since the industrial revolution. People have to face the side-effects of prosperity, such as environmental degradation and climate change. These problems do not only occur in the countries who overuse fossil fuels, but go beyond the boundaries and become a global nightmare. People used to only consider the interests of themselves, but overlook the consequences of their behaviors, which cause "the tragedy of commons" (Hardin 1968). Nowadays, environmental degradation and climate change are serious challenges to human survival and development. People cannot heavily use fossil fuels as before anymore and need to discover new energy

sources.

Renewable energy sources are widely known as alternative energy sources to fossil fuels, because renewable energy sources, like wind, solar, hydropower, and geothermal, can reduce local and global atmospheric emissions. Then, the pollution and climate change problems would be solved or relieved. In fact, the role of renewable energy is not limited to protect environment. First, the potential of renewable energy sources is abundant. The reserves of renewable energy are much more adequate than fossil fuels. Currently, the total solar power is close to 10000 times the rate of global energy consumption and sun can supply solar energy for billions of years. In addition, the solar energy usually becomes of the other kinds of renewable energy, such as wind. The kinetic energy of the wind amounts is more than 100 times global energy consumption. Thus, renewable energy can support the sustainable development of human. Compared to renewable energy, the reserves of fossil fuel are limited. Even if human maintains current energy consumption rate, fossil fuel would be exhausted in the next century, let alone considering the growing energy demand. Second, renewable energy resources are widely distributed over the world, unlike other energy sources, which are mainly concentrated in a few countries. For example, although the intensity of sunlight varies in different part of the earth due to a different latitude and climate, every place on the earth have enough solar energy to be exploited. Thus, even rural and remote areas and

developing countries, where fossil fuels are scarce, can exploit solar energy to develop economically.

Solar power is presently considered as one of the most promising sources of renewable energy and solar photovoltaic (PV) is a major method to exploit solar energy. The steady technological progress of PV industry reduces the cost of PV module and increase the efficiency, so that PV is close to compete with conventional thermal power plant in the near future. The globally cumulative installed PV capacity was 227 GW at the end of 2015, at least 10 times higher than in 2009 (International Energy Agency 2016). Although PV industry grows rapidly, the variability and uncertainty issues of PV result in the unreliable nature of PV system and impose some constraints on the development of high penetration rates of PV systems.

PV power is sometimes criticized for being intermittent and viewed as garbage power. Because solar energy is weather-dependent, a PV system's performance could change more than 50% in a few minutes due to a thunder shower. The intensive fluctuations of PV system will shorten the operation lives of transformers, which connect to the distribution network. Thus, the high penetrate of PV will bring detrimental effect to the power distribution network (Zhang et al. 2013) and the most severe constraint in practical projects is the risk of PV system that cause blackouts.

However, the problem is not insoluble and there are two practical strategies to cope with. The one is applying energy storage facilities to eliminate the intermittency and instability of renewable energy sources and to increase the reliability of the power grid (Evans, Strezov, and Evans 2012). However, bulk energy storage facilities are not feasible for most of PV system at the current technological level. The other one is forecasting the variability of PV and reduce the uncertainty of prediction results. Then, the grid operators manage this predefinition intermittent production input more efficiently to keep the balance between energy production and demand. Thus, forecasting and enhanced grid management techniques can increase the predictability and integration of solar energies for widespread penetration and to advance the development of PV generation (Paulescu et al. 2012).

1.2 Research purposes

With the advance of technology, satellite can both have very high temporal and fine spatial resolution with multiple spectral bands, like Gaofen-4 (Lin and Singer 2016). The physical model using satellite imagery might be the most promising way to predict solar irradiance, which is critical for PV forecasting, with high accuracy for both short-term and long-term. For short-term forecasting, the location of cloud cover can be predicted based on the fine cloud motion vectors, which can be extracted from satellite images with high spatial resolution. For the long-term forecasting, the weather measurements with higher spatial

resolution derived from satellite images are very critical inputs for Numerical Weather Prediction (NWP) models to get accurate forecasting results.

However, due to the low computational efficiency of 3D (three-dimensional) radiative transfer model, most researches for operational purposes are still using 1D (one-dimensional) radiative transfer model (Chen et al. 2012, Huang et al. 2013) or even empirical model (Gupta et al. 2001, Huang, Liu, and Liang 2012) to calculate cloud transmittance. In addition, the temporal resolution (hourly, daily) and spatial resolution (several kilometers) of the downward surface shortwave radiation (DSSR) products for all-sky conditions are usually very coarse. In fact, the time scale of the fluctuations on solar irradiance can be a minute or less (Tomson 2010), so that coarse temporal resolution is not suitable for forecasting the solar irradiance with high fluctuations in a very short time. Moreover, the coarse spatial resolution is also not suitable to predict the solar irradiance on a small area.

The primary objectives of this study are: 1) using a 3D radiative transfer model and MODIS products (MOD04-07) to estimate instantaneous solar irradiance on all-sky days and assessing the accuracy of solar irradiance estimates; 2) discussing the advantages and disadvantage of using 3D radiative transfer model and finding a solution to apply a 3D radiative transfer model in operational environment. Therefore, this research combines two existing radiative transfer model to estimate DSSR using MODIS (Moderate Resolution

Imaging Spectroradiometer) products (MOD03-07, 09). Yang's model (Yang, Koike, and Ye 2006) was adopted to calculate clear transmittance and I3RC (Intercomparison of 3D Radiation Codes) Monte Carlo 3D radiative transfer community model (Cahalan et al. 2005) was adopted to calculate cloud transmittance. In addition, a method to model the 3D cloud structure, which is an input of I3RC model, in Tibet were proposed. Finally, an overall assessment of the combined model and a suitable extent to apply 3D theory would be discussed by comparing the calculated results with DSSR measurements from a ground monitor station in Tibet and the downward shortwave surface flux dataset from FLASHFlux (Fast Longwave And SHortwave Radiative Fluxes) project.

Chapter 2   Review of Literature

The PV system's performance is decided by solar irradiance, the system factors that address the specific design of PV system, and ambient factors (like temperature, wind speed etc.). Although the latter two factors affect the efficiency of PV system, their influence on the performance during operation is typically less than 20% (Kleissl 2013). Thus, PV system's performance is fairly linear to the incoming solar irradiance, which is the most important type of factor compared to the other types of factors. Thus, PV forecasting can be divided into two parts: solar irradiance measurements and forecasting model.

2.1 Solar irradiance derived from satellite observation

Solar irradiance can be measured by ground monitor station or derived from satellite observation. Due to the scarcity of ground monitor stations and data, solar irradiance data derived from satellite are used more widely in PV industry.

Solar irradiance varies on a wide range of time scales from a fraction of a second to centuries. The variability of solar irradiance mainly consists of two parts: astronomical conditions and atmospheric conditions. The astronomical conditions result in the basic seasonal and diurnal variability. Because this kind of variability is very regular, so that the top-of-atmosphere (TOA) insolation can be predicted by mathematical formulae with high

accuracy (Gupta et al. 2001).

The atmospheric conditions' effects on solar irradiance at the horizontal ground are irregular, so that it is much more complicated to estimate. Besides the TOA insolation, the model to estimate DSSR typically consists of two distinct parts: clear-sky transmittance, and cloud transmittance. These two parts correspond to clear-sky and cloudy-sky conditions. All-sky refers to the combination of clear-sky and cloudy-sky.

The clear-sky transmittance is the result of a global and diffuse solar radiation model on a clear-sky day. The radiative transfer model can be viewed as a function of astronomical parameters, geographical parameters, meteorological parameters, and quantities related to atmospheric turbidity (Badescu et al. 2012). Astronomical parameters can be used to calculate the path length of radiation fluxes. Geographical parameters provide the elevation and location information for satellite. Meteorological parameters provide the composition of atmospheric gases, such as water vapor and ozone content, and environmental state. Quantities related to atmospheric turbidity are critical to building the absorption and scattering processes of aerosol. These four types of parameters can be measurements of ground, estimations, and satellite derived data. (Badescu et al. 2012) compared a comprehensive set of 54 clear-sky models, and there was no model to be ranked "the best" for all sets of input data. Simple models as well as more complex models may get the

accurate results. It is possible to predict the DSSR with good accuracy if various inputs are known; however, the models using more inputs usually perform better than the models using a few of inputs (Widén et al. 2015). In addition, these models can be classified as either broadband that estimates solar radiation in a single broadband or spectral that estimates solar radiation in different wavelength bands. In fact, PV industry mainly concerns about the broadband amount solar irradiance to predict electrical power output.

Compared with clear-sky transmittance, cloud transmittance has profound effects on DSSR, because cloud is the strongest modulator of the solar radiant energy absorbed by the earth–atmosphere system among all the atmospheric constituents (Chen et al. 2012). There are mainly two categories to consider the cloud transmittance: empirical models and physical models.

Empirical models do not consider the physical characteristics of cloud, but assume that cloud transmittance is a function of the satellite or ground measurements. Cano et al. (1986) proposed cloud (cover) index, which is defined as:

$$N = (C - C_{min})/(C_{max} - C_{min}), \tag{2.1}$$

where C, C_{min}, C_{max} are values of the current, minimum (usually a clear-sky background value) and maximum observed satellite counts (pixel's gray value without radiative

meaning). There are many operational models, such as Heliosat (Beyer, Costanzo, and Heinemann 1996), LPSA (Gupta et al. 2001), SUNY model (Perez et al. 2002), SolarGIS model (Cebecauer, Šúri, and Perez 2010), are functions of cloud index to calculate cloud transmittance. Similar to cloud index, cloud transmittance can be also assumed as a function of sunshine duration (Angstrom 1924). There are many evolved models based on sunshine duration (Gopinathan 1988, Yang, Koike, and Ye 2006). Although these models are very simple and easy to calculate, the effects of clouds are not physically based and the parameters of empirical methods are hard to predicts accurately from forecasting models, so that errors of all-sky conditions from empirical models are inevitable (Sun and Liu 2013).

Physical methods apply radiative transfer theory to estimate cloud transmittance from satellite retrievals. Considering the spatial dimensions of the atmosphere, most atmospheric radiative transfer models and algorithms can be classified into two broad categories: one-dimensional (1D) plane-parallel radiative transfer models and three-dimensional (3D) radiative transfer models (Chen et al. 2012).

1D radiative transfer theory is the most widely used radiative transfer models in atmospheric sciences. Due to the composition (ice and water), vertical structure, global, and spatial variability of cloud, it is hard to model radiative transfer processes in clouds very precisely. 1D theory is simplified by omitting the two horizontal spatial dimensions so that

the radiation field depends only on an angle and vertical location and assumes that there is

no net transfer of radiation in the horizontal (Pincus and Evans 2009). The most simple way

to consider cloud effects is Parallel Plane Approximation (PPA), which assumes an

atmosphere stratified in parallel planes in which all the physical properties are invariant over

a plane (Chandrasekhar 1960). When PPA calculates the cloud transmittance using satellite

retrievals, the optical and microphysical properties of cloud of each pixel are computed from

the mean values of satellite cloud retrievals around the pixel. A more advanced method is

called independent pixel approximation (IPA), which is used by most 1D algorithms, because

it considers that atmosphere is horizontally heterogeneous. IPA treats each pixel in a satellite

image independently and neglects the radiative transfer effects of neighboring columns

(pixels). Both approximations can calculate cloud transmittance using a simple cloud

parameterization model, such as (Stephens, Ackerman, and Smith 1984),or radiative transfer

codes, such as DISRORT (Stamnes et al. 1988), MODTRAN (Berk, Bernstein, and

Robertson 1987). However, the cloud transmittance can be problematic, because the

assumption of 1D theory is not always valid.

The main drawbacks of 1D theory are that clouds are not 1D and the horizontal exchange

of photons between different parts of a cloud or between clouds cannot be accounted for by

1D theory (Cahalan et al. 2005). The flaws of 1D theory can impact fluxes at the top and

bottom of the atmosphere, local heating rates, and especially the estimation of atmospheric

properties from remotely sensed radiation. 1D radiative transfer theory is only suitable to be

applied at medium and coarse scales (Diner and Martonchik 1984). Wyser et al. (2002)

pointed out that a shift of the apparent position of clouds and their shadows, and the diffusion

of radiation are the effects of spatially inhomogeneous cloud fields that need to be correct

when applying an IPA algorithm to process high-to-medium (up to 1km) resolution image.

The former two cloud effects can be called satellite parallax and solar geometry. Satellite

parallax is a displacement in the apparent position of cloud on the ground viewed along

sensor view angle and overhead. Solar geometry results in that the cloud shadows on the

ground is dislocated from the position of the cloud in the sky due to the zenith and azimuth

angle of solar. The third one is known as adjacency effect, which reduces apparent surface

contrast by decreasing the top of the atmosphere radiance over bright pixels and increasing

the brightness of the dark pixels, and it is caused by the scattering removes part of the

radiation from an atmospheric column and distributes it to neighboring columns (Lyapustin

and Kaufman 2001).

In 3D radiative transfer theory, the radiation field depends on three spatial dimensions

of atmosphere. There are mainly two broad classes of 3D radiative transfer model: explicit

method and Monte Carlo (MC) method (Cahalan et al. 2005). The essence of different 3D

radiative transfer models is to solve the radiative transfer equation. Explicit methods start with a first-guess solution and iteratively adjust the radiation field until it agrees with the radiative transfer equation to some specified accuracy; Monte Carlo methods solve the radiative transfer equation by computing a large number of discrete trajectories through the full radiation field (Pincus and Evans 2009). The most widely used explicit method is Spherical Harmonics Discrete Ordinate Method (SHDOM; (Evans 1998)), which is publicly available at http://nit.colorado.edu/~evans/shdom.html. For MC method, I3RC (Intercomparison of 3D Radiation Codes) Monte Carlo community model of 3D radiative transfer is the most widely used MC codes and it is publicly available at http://i3rc.gsfc.nasa.gov/.

Although the advantages of 3D radiative transfer model are widely acknowledged and the cloud transmittance can be calculated more accurately than 1D theory, there are two barriers to use it. First, describing the instantaneous distribution of extinction, single-scattering albedo, and scattering phase function in three spatial dimensions at very small scales is difficult; Second, solving the 3D radiative transfer is much more computationally expensive than solving 1D theory with the same purpose (Pincus and Evans 2009). For the first barrier, I3RC model use Mie theory and cloud structure with physical properties to deal with. Accurate cloud structure can be directly built using the optical sensors with multiple

viewing angles, such as Multi-angle Imaging SpectroRadiometer (MISR) on Terra, or radar sensors, such as the Cloud Profiling Radar (CPR) on CloudSat. The physical properties of a cloud, such as liquid water content, can be calculated based on corresponding algorithm and satellite retrievals. Although it is impossible to rebuild an accurate cloud structure just based on the normal optical sensor, such as MODIS, and many assumptions or accessory data can be helpful to model a cloud structure using MODIS data (Miller et al. 2014). In fact, the second barrier is the main reason that limits the application of 3D radiative transfer model in an operational environment.

2.2 Forecasting model

The accuracy of forecasting models depends on the time horizon and the spatial resolution, and different methods are suitable for different temporal and spatial domains (Widén et al. 2015). There are two broad categories of forecasting models, which are data analysis models and physical models, and three forecasting horizons, which are intra-hour, intra-day, and day(s) ahead (Kostylev and Pavlovski 2011).

Data analysis models are based on the analysis of historical data in order to make predictions. They are often used as short-term (intra-hour and intra-day) PV forecasting. Because it can predict the power output directly from historical data, such as PV outputs, solar irradiance, ambient factors, Data analysis models usually predict the performance of

PV system directly based on rather than predicting solar irradiance first. Sometimes, data analysis methods even make predictions without the inputs of solar irradiance.

Statistics is a conventional category of data analysis models. A typical model is autoregressive (AR) analysis of time series and it was used in PV forecasting as early as the 1980s (Chowdhury and Rahman 1987). Up to now, the AR methods are still widely used by many companies to do PV forecasting. AR is often combined with other methods, such as moving average (MA), integrative part (I), non-linearity (N) or exogenous inputs (X) to form more sophisticated forecast models (Widén et al. 2015). Among these extra methods, exogenous inputs are the most useful method to combine with AR, because it can effectively decrease the RMSE. Bacher, Madsen, and Nielsen (2009) applied an ARX day ahead model that takes global irradiance from a NWP model as exogenous inputs, and the RMSE of ARX is 12% lower than the result of AR. This result is consistent with a more recent study conducted by Li, Su, and Shu (2014). This study used ARMAX method, which takes temperature, precipitation amount, insolation duration, and humidity as exogenous inputs, and the ARMAX had 26.7% lower RMSE than an ARIMA model.

An emerging category of data analysis models is machine learning. Deep learning is one of the most popular branches of machine learning methods. It has two features: (1) models consisting of multiple layers or stages of nonlinear information processing; and (2) methods

for supervised or unsupervised learning of feature representation at successively higher, more abstract layers (Deng and Yu 2014). The multiple layers of deep learning method can minimize the error by tuning the weights of the input dataset. These features enable the deep learning methods to effectively exploit complex, compositional nonlinear functions, so that it has been popular in forecasting applications. For example, artificial neural network (ANN) is the one of the most successful algorithms. Paoli et al. (2010) showed that ANN method presents predictions similar to or even better than conventional statistical method, such as ARIMA techniques, or simple machine learning method, like k-Nearest-Neighbors method. Chen et al. (2011) built an ANN model to predict 24-hour ahead PV output by using the mean daily solar irradiance, relative humidity, air temperature and power output of the PV system as the model input. The testing days' mean absolute percentage error ranges from 6.36% to 54.44%. In the study of Chupong and Plangklang (2011), recurrent ANN achieved a 24-hour-ahead PV array power forecast with a mean absolute percentage error of 16.83% for Thanyburyi, Thailand, without using solar radiation measurement as input. In addition, one machine learning algorithm can be combined with the conventional statistical methods or different types of machine learning algorithms to improve the prediction accuracy (Chaouachi et al. 2009, Ji and Chee 2011).

Physical models can be used for both short-term and long-term solar irradiance

forecasting. There are mainly two types of physical model: cloud motion and numerical weather prediction (NWP). Cloud motion approaches extract cloud motion vectors from successive images and forecast the location of cloud cover in the short-term based on assuming persistence of cloud speed, size, and shape (Lorenz, Hammer, and Heinemann 2004). Using satellite images can detect the cloud patterns at a moderate spatial resolution for intra-day forecasting. Chow et al. (2011) said the frozen cloud advection derived from GOES satellite images can provide accurate forecasts up to 6 h ahead at a spatial resolution of 1km. For high temporal and spatial resolution of intra-hour forecasts, Chow et al. (2011), and Chow, Belongie, and Kleissl (2015) provided solutions to extract cloud motion derived from ground-based sky imagers. NWP builds mathematical models of the atmosphere and oceans to predict the evolution of weather based on initial weather conditions up to several days ahead (Diagne et al. 2013, Pelland et al. 2013). The spatial resolution of NWP varies from a few kilometers to tens of kilometers and spatial resolution varies from one hour to three hours (NOAA 2016, Northwest Modeling Consortium 2016). Compared to statistical methods, the solar irradiance estimates are critical for physical models, because it depends on modeling the physical processes that affect the output of a PV system. Thus, physical models need post-processing to estimate solar irradiance and PV output. The solar irradiance can be computed by the methods referred in the section 3.3 based on the output of a physical

forecasting model. After acquiring the solar irradiance, the performance of a PV system can be calculated by the physical model of PV system using the system factors (Paulescu et al. 2013). Similar to statistical methods, combining different types of physical model, such as NWP, can also achieve improvements in forecast accuracy (Perez et al. 2011).

There are both advantages and disadvantages of pure data analysis models and physical models. Data analysis models usually have better accuracy than physical models in the short-term, but worse for the long-term (Widén et al. 2015). Moreover, data analysis models always need huge historical datasets to build models and the overfitting of models would undermine their ability to perform well on unseen data. Physical models do not need historical data and outperform the accuracy of data analysis models in the long-term. To overcome the deficiencies of using one class of models, hybrid models that combine data analysis and physical models are desirable choice. Physical models can cope with the chaotic nature of the atmosphere and weather to control the systematic errors of forecasting at a moderate level for both short-term and long-term. At the same time, it can reduce the dependencies on historical datasets and be used at anywhere. The outputs of the physical models require the post-processing and can be improved with statistical method to minimize errors. Diagne et al. (2014) showed that the hybrid model can effectively reduce the errors of solar irradiance forecasting. In addition, any measure to improve the performance of either

model in a hybrid model would reduce errors.

Chapter 3　Methods and Materials

In this section, the datasets used in this study were introduced at first. Then, the procedures to preprocess these datasets were presented. Furthermore, the algorithms for instantaneous DSSR were described in detail. Last, the indices to assess the results of the algorithms were provided.

3.1 Datasets of ground observation and remote sensing

Three types of data were used in this study. The first type of data is in-situ ground observation data, which are viewed as real solar irradiance values. The second type of data is MODIS products, which are used as input data of the algorithms for instantaneous DSSR. The last type of data is an existing DSSR dataset of FLASHFlux data.

3.1.1　In situ data from Lhasa monitor station

In situ observation from the Lhasa station (29.6466°N, 91.1779°E, which is 3680m above mean sea level (AMSL) in Tibet University (New Campus). Lhasa is the capital city of the Tibet Autonomous Region of the People's Republic of China at an altitude of 3650m. Due to its very high elevation, Lhasa has a cool semi-arid climate with frosty winter and mild summer. Lhasa has an annual precipitation of 426 millimeters (16.8 in) with rain falling mainly in July, August and September. The driest month is January at 0.8 millimeters (0.03

in) and the wettest month is August, at 120.6 millimeters (4.75 in). The rains come mostly at night and Lhasa is still sunny during the daytime, so that Lhasa receives nearly 3,000 hours of sunlight annually. The specific location of the station is shown in Figure 3(a) and (b). The total solar irradiance data were measured by the instrument, which is called pyranometer, in the left side of Figure 3(c).
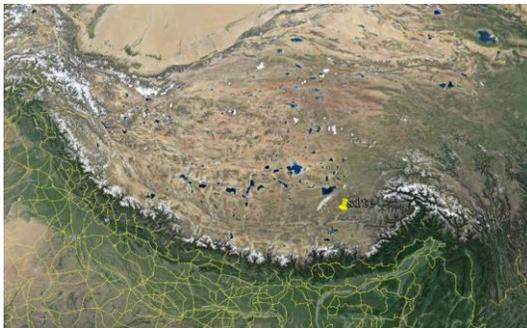
(a)   Station location in Tibet

(b)   Station location in Lhasa

(c)   The measuring system at Lhasa station (29.6466°N, 91.1779°E)

Figure 3 Overview of the Lhasa station

The CMP-6 broadband pyranometer was produced by Kipp & Zonen Company in Netherlands as shown in Figure 4. CMP-6 is a high quality radiometer designed for measuring shortwave irradiance on a plane surface. CMP-6 can accurately measure the intensity of solar radiation and the radiation response curve of the instrument is shown in

Figure 5. Its expected daily accuracy is ±5% and wavelength range of 310 nm to 2800 nm and the maximum radiation output signal intensity of 2000 W/m². Based on the extraterrestrial spectrum (Gueymard 2004), the total solar irradiance between 310 nm and 2800 nm is 1312.8 W/m².

Figure 4 The connection of CMP-6
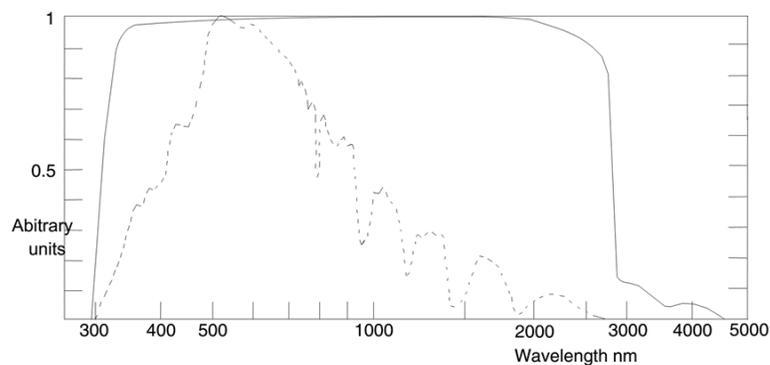
Figure 5 Solar irradiance spectrum at the Earth's surface and pyranometer response (CMP-6 user manual)

The instrument is operated and maintained by Tibet University. In this study, the total solar irradiance data (including direct insolation and diffuse insolation) from July 12th 2012 to October 31st 2013 were selected. The time interval of data sampling is one minute and the

units of total solar irradiance is W/m^2. For example, Figure 6 shows the minutely solar

irradiance scatter plot on July 13th, 2012.
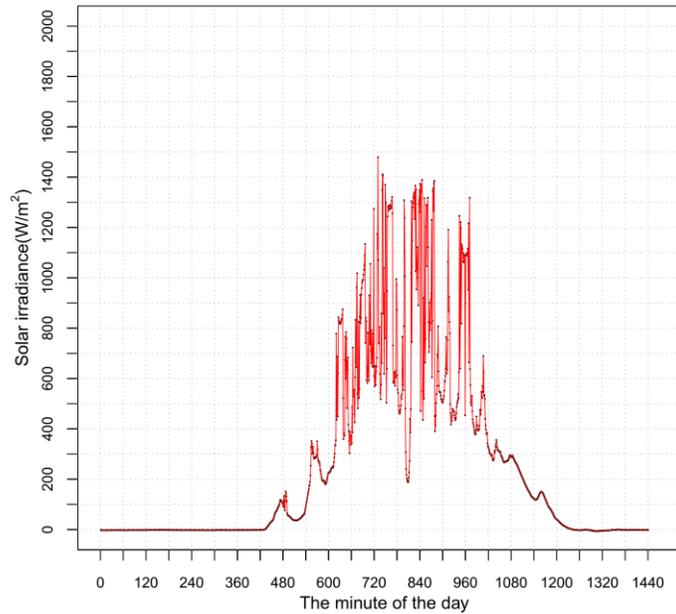
Figure 6 Minutely (GMT+8) solar irradiance scatterplot on July 13th, 2012

The variability of clouds will result in the variability of solar irradiance and affects the

accuracy of estimation. The variability of solar irradiance is low in a clear-sky or when the

clouds are uniform distributed. The variability is high in a cloudy day that the clouds are

dispersed (Figure 6). The variability metric of solar irradiance for single location is defined

as the standard deviation of the change in the clear-sky index of the location using the

specified time interval over the selected time period (Kleissl 2013). The variability metric

can be computed by:

$$\sigma(\Delta Kt^*_{\Delta t}) = \sqrt{VAR[\Delta Kt^*_{\Delta t}]}, \qquad\qquad (3.1)$$

where: $\Delta Kt^*_{\Delta t}$ is the change in clear-sky index Kt^* over a specific time interval Δt. Kt^* can describe the weather/clouds factors on solar irradiance and it is equal to cloud transmittance τ_C in this study. Δt decides the variability and usually the less Δt is, the higher variability is. In this study, Δt is 1 minute based on the in-situ data.

3.1.2   MODIS products data

MODIS (Moderate Resolution Imaging Spectroradiometer) is spaceflight scientific instrument for remote sensing (NASA 2016g). The instrument acquires data in 36 spectral bands ranging in wavelength from 0.4 μm to 14.4 μm and at three spatial resolutions (2 bands at 250 m, 5 bands at 500 m and 29 bands at 1 km) (NASA 2016h). A ±55-degree scanning pattern at the EOS (Earth Observing System) orbit of 705 km achieves a 2,330-km swath and provides 1 or 2 days repeat observations of the Earth. MODIS was loaded on two satellites: The one is Terra satellite, which was launched on December 18, 1999; The other one is Aqua satellite, which was launched on May 4, 2002. Thus, the pairs of satellites can image one place on the earth 1 or 2 times during daytime.

Standard MODIS data products are widely applied in geoscience. They are produced by the MODIS Adaptive Processing System (MODAPS), who converts the radiances received

by the instruments from both the Terra and Aqua platforms to geophysical quantities (NASA 2016f). The physical theory and the mathematical procedures and possible assumptions of the MODIS products can be found in their Algorithm Theoretical Basis Documents (ATBD) (NASA 2016a). LAADS (Level 1 and Atmosphere Archive and Distribution System) receives data processed by MODAPS and provides the free download service of Level 1 products (calibrated radiances and geolocation) and Level 2 and 3 Atmosphere and Land products for the public. In this study, data products (Collection 6) of geolocation, atmosphere and land were used. These products can be obtained from LAADS in hierarchical data format (HDF). There are two types MODIS product files: MOD, containing data collected from the Terra platform; and MYD, containing data collected from the Aqua platform. MXD refers to both MOD and MYD products. Although the data from the two different satellites, the differences of the two types of products are small. Thus, the MOD and MYD products were assumed to be identical in this study.

All datasets that are used as unknown variables in the three procedures in section 3.3 were shown from Table 1 to Table 3. Based on the time range of in-situ data, the MOD 03-07, and 09 data from July 12th 2012 to October 31st 2013 were downloaded.

Table 1 Required datasets of MODIS products for calculating TOA insolation

Variable name	symbol	MODIS products	Datasets name	Spatial resolution
Solar zenith angle	θ_{sol}	MOD03	SolarZenith	1 km

Table 2 Required datasets of MODIS products for calculating clear-sky transmittance

Variable name	symbol	MODIS products	Datasets name	Spatial resolution
Solar elevation angle	h	MOD03	SolarZenith	5 km
Total column precipitable Water Vapor	w	MOD05	Water_Vapor_Near_Infrared	1 km
Thickness of the ozone layer	l	MOD07	Total_Ozone	5 km
Aerosol optical depth and wavelength	$\delta(\lambda),\ \lambda$	MOD04	Deep_Blue_Spectral_Aerosol_Optical_Depth_Land	10 km
Ångström exponent	α	MOD04	Deep_Blue_Angstrom_Exponent_Land	10 km

Table 3 Required datasets of MODIS products for calculating cloud transmittance

Variable name	symbol	MODIS products	Datasets name	Spatial resolution
Solar zenith angle	θ_{sol}	MOD03	SolarZenith	1 km
Solar azimuth angle	γ_{sol}	MOD03	SolarAzimuth	1 km
Sensor zenith angle	θ_{sen}	MOD03	SensorZenith	1 km
Sensor azimuth angle	γ_{sen}	MOD03	SensorAzimuth	1 km
Cloud top height AMSL	CTH	MOD06	cloud_top_height_1km	1 km
Cloud top temperature	T_{top}	MOD06	cloud_top_temperature_1km	1 km
Cloud water path	CWP	MOD06	Cloud_Water_Path	1 km
Cloud effective radius	r_{eff}	MOD06	Cloud_Effective_Radius	1 km
Surface temperature	T_s	MOD06	Surface_Temperature	5km

3.1.2.1   Geolocation data

The MODIS geolocation product (MXD03) provides latitude, longitude, ground elevation, solar zenith angle, solar azimuth angle, satellite zenith angle and satellite azimuth angle for each MODIS 1 km pixel. The latitude and longitude datasets with 1 km spatial

resolution of MXD03 are indispensable to project higher level MODIS products, like

MOD04-07, into particular projection and can decide the actual pixel of each image that

contains the station. In addition, the transit time of the satellite can be obtained via MXD03.

3.1.2.2 Atmosphere data

The MODIS aerosol product (MXD04) provides aerosol optical depth (AOD) at four

wavelengths, 0.412, 0.47, 0.55, and 0.66 μm and Ångström exponent over land. The MXD04

Level 2 data are produced at the spatial resolution of a 10x10 1-km (at nadir)-pixel array

(NASA 2016b). The AOD over land is estimated using both dark target and deep blue

approaches (Levy et al. 2013).

The MODIS precipitable water product (MXD05) provides total column precipitable

water vapor at the spatial resolution of a 1-km (at nadir)-pixel array using the near-infrared

algorithm during the day (NASA 2016e).

The MODIS cloud product (MXD06) combines infrared and visible techniques to

determine both physical and radiative cloud properties (NASA 2016d). It provides cloud-top

temperature, cloud-top height, cloud water path, and effective cloud-particle radius at the 1-

km spatial resolution of the MODIS. In addition, MXD06 contains surface temperature at

the 5-km spatial resolution from ancillary data.

The MODIS atmospheric profile product (MXD07) provides total-ozone burden, which is an estimate of the total-column tropospheric and stratospheric ozone content (NASA 2016c). The MXD07 Level 2 data are produced at the spatial resolution of a 5x5 1-km (at nadir)-pixel array.

3.1.2.3   Land data

Surface Reflectance Product (MOD09) provides an estimate of the surface spectral reflectance for each band as it would have been measured at ground level if there were no atmospheric scattering or absorption. It is computed from the MODIS Level 1B land bands 1, 2, 3, 4, 5, 6, and 7 (centered at 648 nm, 858 nm, 470 nm, 555 nm, 1240 nm, 1640 nm, and 2130 nm, respectively). The albedo of the ground surface of the station is calculated by averaging cloud free pixels that contain the ground station of MOD09. In this study, the albedo for I3RC input is 0.15.

3.1.3   FLASHFlux data

The FLASHFlux (Fast Longwave And SHortwave Radiative Fluxes) project provides global near real-time surface and TOA radiative fluxes based on Terra and Aqua's CERES (Clouds and the Earth's Radiant Energy System) and MODIS TOA observations with fast radiation algorithms. The home page of the project is http://flashflux.larc.nasa.gov/ and the

data can be downloaded from https://eosweb.larc.nasa.gov/project/ceres/ssf_table.

FLASHFlux Level 2 product provides downward shortwave surface flux dataset on a near real-time, which using the Langley parameterized shortwave algorithm described in (Gupta et al. 2001). Because CERES and MODIS loaded on the same satellite, their data are captured at the same time and nearly every scene of MODIS has corresponding FLASHFlux Level 2 products. Thus, the downward shortwave surface flux dataset of FLASHFlux Level 2 products is very good references to evaluate the performance of instantaneous DSSR estimates using MODIS products.

FLASHFlux Level 2 product contains many datasets, which are introduced by Geier et al. (2003). There are two downward shortwave surface flux datasets, "CERES downward SW surface flux - Model A", and "CERES downward SW surface flux - Model B". Because the former one lacks data for cloudy-sky conditions, the latter one was selected as a reference data to compare with the DSSR estimates computed by the method in section 3.3. In addition, "Clear area percent coverage at subpixel resolution" (hereinafter referred to as clear percent) dataset was used as indicators to distinguish clear-sky and cloudy-sky. The dataset represent the clear area percent within CERES Field-of-View (FOV). The optical FOV at nadir is 16 × 32 km or 20 km resolution (Geier et al. 2003). Considering the time range of in-situ data, the FLASHFlux Level 2 data from July 12th, 2012 to October 31st, 2013 were used in this

study.

3.2 Data preprocessing

The incorrect, missing, or inconsistent data always appear in the raw dataset of ground observations and remote sensing data. The corrupt or inaccurate data can lead to false conclusions, so that identifying, modifying, or deleting these dirty data are required before further processing. In addition, remote sensing data usually need data preprocessing, such as clipping and projecting before extracting the parameter values.

3.2.1 Ground observations

The first step is to identify the incorrect, missing data and to delete them. The ground observations are minutely data, so that each total solar irradiance (W/m^2) record has a timestamp (Figure 7).

```
2012-08-02 16:31:54     966.046
2012-08-02 16:32:54     967.230
2012-08-02 16:33:54     963.068
2012-08-02 16:34:54     945.338
2012-08-02 16:35:54     979.600
2012-08-02 16:36:54     956.328
2012-08-02 16:37:54     942.910
2012-08-02 16:38:54    1011.411
```

Figure 7 The format of ground observations data

There are three types of missing records: 1) The solar irradiance values are missing; 2) The timestamps are missing; 3) Both solar irradiance values and timestamps are missing.

Moreover, the total solar irradiance values and timestamps can be duplicated. Due to the high variability of the solar irradiance data, imputation might add some new error in the original dataset. Thus, the records that have missing or duplicated values were deleted.

The second step is to identify the inconsistent data and to correct them. The inconsistent data are identified by visualizing the data using scatter plot. As shown in Figure 6, the number of minute of a day (timestamp) is the x axis and the solar irradiance value is the y axis. The missing data can be easily found by skimming the graphs of solar irradiance. The inconsistent data can be identified by examining the features of graphs on clear-sky days. For example, on a clear-sky day, the scatterplot is close to a sinusoidal curve (Figure 8(a)). The center line of the sinusoidal curve is around 840th minute of a day. The maximum values of the sinusoidal curve of two adjacent clear-sky days will be very close. Thus, the shape of curve, the center line of curve, and the maximum value of a scatterplot are the features of graphs.

If there is no gap in a scatterplot, like Figure 8(a), the data in this time period are assumed to be consistent. If the data have a big gap in a scatterplot, like Figure 8(b), the inconsistent data might emerge in this time period. If the feature of the graph on a clear-sky day before the gap was different from the feature on the following clear-sky day after the gap, the data on the two days are inconsistent. The inconsistent data are assumed to be

corrected by a line regression function, whose dependent variable is correct data and the independent variable data is inconsistent. The correct data and inconsistent data have the same date in different years.

| (a) A clear-sky day | (b) A day has a gap |
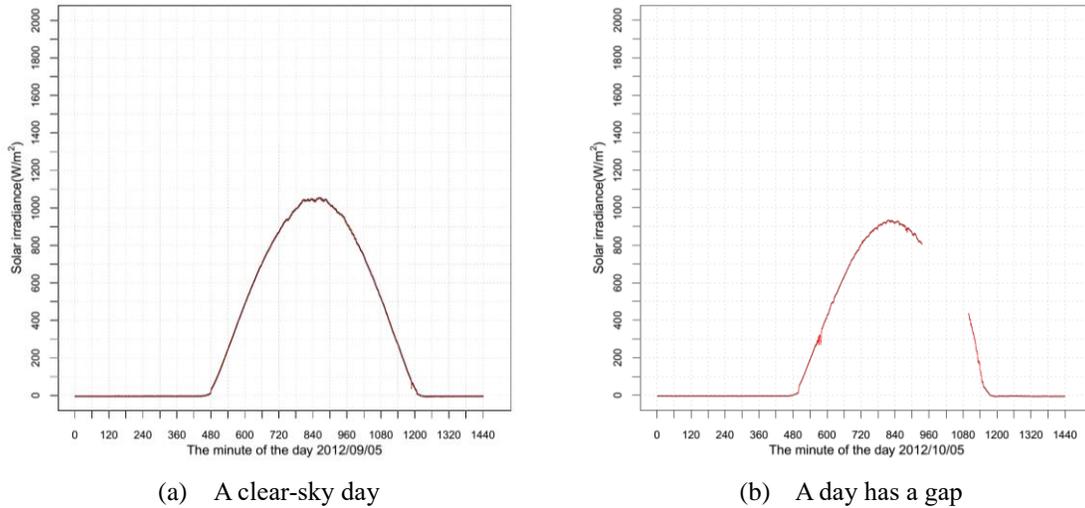
Figure 8 Typical scatterplots

3.2.2   Remote sensing data

For MODIS products, data preprocessing has four steps. First step is to remove the scenes that do not contain the station. Although users can download the MODIS products from LAADS based on area of interest, some retrievals still do not contain the area of interest. MOD03 provides Input Geometry (IGM) file, which contains the latitude and longitude information for each pixel. Thus, the scenes that do not contain the station can be deleted based on IGM file.

The Second step is to clip a square subset of MODIS products with 1km spatial

resolution around the station pixel and the side length of the square is set to 51 pixels. Because the effects of clouds that are more than 25 km from the station on the solar irradiance measurements are negligible at transit time of MODIS and a smaller image can reduce the computation time for the following steps, clipping a subset of MODIS products for each scene is necessary for the study.

The Third step is to georeference remote sensing data. Based on IGM file, the subsets of MODIS products can be projected to Universal Transverse Mercator (UTM) map projection. UTM is a conformal projection, so it preserves angles and approximates shape. Although UTM distorts distance and area, the latitude of the station is not high and the distortion of distance is no very big. Due to the longitude of the station, the UTM zone is north zone 46. If one MODIS product was corrupt that cannot get valid data after above-mentioned two steps, all MODIS products of the same scene would be ignored.

The fourth step is imputation. The imputation methods are depended on the types of MODIS products and the fraction of missing data. If the variables do not change a lot in a short time period, the missing values can be imputed from an average value of previous and next days. If the variables vary greatly, the missing values are replaced by a constant value based on prior knowledge.

For FLASHFlux data, removing the scenes that do not contain the station is necessary

based on IGM file. Then, extracting the downward shortwave surface flux values of the pixels that contains the station is enough for the study. It is unnecessary to do clipping, georeferencing, and imputation.

3.3 Algorithms for instantaneous DSSR

The instantaneous DSSR for all-sky conditions can be calculated from a broadband shortwave model (Gupta et al. 2001):

$$F_{GHI} = F_{TOA} \ \tau_A \ \tau_C, \tag{3.2}$$

where F_{TOA} is the solar irradiance on a horizontal surface at the TOA; τ_A is the transmittance of clear-sky; τ_C is the transmittance of clouds.

3.3.1   TOA insolation

F_{TOA} can be calculated by (Peixoto and Oort 1992):

$$F_{TOA} = S(d_m/d)^2 \cos\theta_{sol}, \tag{3.3}$$

where d is the actual distance and d_m is the mean distance between the sun and the earth; S is the solar constant, which is 1366.1 W/m² (Gueymard 2004); θ_{sol} is the solar zenith angle, which can be acquired from the ancillary data of MODIS products. However, S should be the total solar irradiance within the pyranometer response waveband (310-2800 nm) in

this study. The S is assigned as 1312.8 W/m^2.

The eccentricity correction factor, $(d_m/d)^2$, of the earth's orbit can be computed by (Spencer 1971):

$$(d_m/d)^2 = 1.00010 + 0.034221 \ \cos\Gamma + 0.001280 \ \sin\Gamma + \ 0.000719 \ \cos2\Gamma + 0.000077 \ \sin2\Gamma, \tag{3.4}$$

where Γ, in radians, is called the day angle, which can be represented by:

$$\Gamma = 2\pi(d_n - 1)/365, \tag{3.5}$$

where d_n is the day number of the year; for the leap year, 365 can be replaced by 366. Iqbal (1983) said the maximum error in $(d_m/d)^2$ computed with Equation (3.4) is 0.0001.

3.3.2   Clear-sky transmittance

For clear-sky conditions, the τ_A can be calculated by a global and diffuse solar radiation model. Yang's model (Yang, Koike, and Ye 2006) was widely used in some researches (Yang et al. 2008, Huang, Liu, and Liang 2012), which need to calculate beam and diffuse irradiance under clear skies. And the high performance of the model was verified by many studies (Madkour, El-Metwally, and Hamed 2006, Badescu et al. 2012). Thus, Yang's model was adopted in this study. The detailed transmittance functions of the all the atmospheric compositions considered in the model are described as:

$$\tau_A = \tau_{b,clear} + \tau_{d,clear}, \tag{3.6}$$

where: $\tau_{b,clear}$ is solar beam radiative transmittance and $\tau_{d,clear}$ is the solar diffuse radiative transmittance under clear skies.

$$\tau_{b,clear} \approx \max(0, \ \tau_{oz} \ \tau_w \ \tau_g \ \tau_r \ \tau_a - 0.013), \tag{3.7}$$

$$\tau_{d,clear} \approx \max\{0, 0.5[\ \tau_{oz} \ \tau_w \ \tau_g(1 - \ \tau_r \ \tau_a) + 0.013]\}, \tag{3.8}$$

$$\tau_{oz} = \exp[-0.0365(ml)^{0.7136}], \tag{3.9}$$

$$\tau_w = \min[1.0, 0.909 - 0.036ln(mw)], \tag{3.10}$$

$$\tau_g = \exp(-0.0117m_c^{0.3139}), \tag{3.11}$$

$$\tau_r = \exp[-0.008735m_c(0.547 + 0.014m_c - 0.00038m_c^2 + 4.6 \times 10^{-6}m_c^3)^{-4.08}], \tag{3.12}$$

$$\tau_a = \exp\{-m\beta[0.6777 + 0.1464(m\beta) - 0.00626(m\beta)^2]^{-1.3}\}, \tag{3.13}$$

$$m = 1/[\sinh + 0.15(57.296h + 3.885)^{-1.253}], \tag{3.14}$$

$$m_c = mp_s/p_0, \tag{3.15}$$

where: τ_{oz}, τ_w, τ_g, τ_r, and τ_a are the broadband radiative transmittances due to ozone absorption, water vapor absorption, permanent gas absorption, Rayleigh scattering, and aerosol extinction, respectively; m is the air mass; m_c is the pressure-corrected air

mass; h, in radians, is the solar elevation; p_0 is the standard air pressure, which is 1.013×10^5 Pa; p_s is the surface pressure in Pa; w is the precipitable water in cm; l is the thickness of the ozone layer in cm or 1000 Dobson Units; β is the Ångström turbidity coefficient.

When the surface pressure is not available, it can be estimated by:

$$p_s = p_0 \exp(-z/H_T), \tag{3.16}$$

where: z is the altitude in meter, and H_T is the scale height of an isothermal atmosphere, which is 8430 m.

The Ångström turbidity coefficient can be expressed by:

$$\delta(\lambda) = \beta \lambda^{-\alpha}, \tag{3.17}$$

where: λ (μm) is the wavelength; $\delta(\lambda)$ is the AOD value; α is the Ångström exponent.

All the unknown variables, like h, w, l, λ, $\delta(\lambda)$, and α, can be acquired from MODIS products and z is decided by the study area.

3.3.3   Cloud transmittance

3.3.3.1   I3RC Monte Carlo 3D radiative transfer community model

In this study, I3RC Monte Carlo 3D radiative transfer community model was used to calculate cloud transmittance. Compared to the other types of 3D radiative transfer model,

SHDOM, Monte Carlo model has been proved to be more efficient in computing flux divergence and downwelling flux in the stratocumulus scene, especially at higher accuracies (Pincus and Evans 2009). I3RC model mainly has three programs and the input files of these programs can be modified based on study purposes.

a) Make Mie table

This program creates a scattering phase function look-up table as a function of effective radius for gamma or lognormal size distribution of spherical particles using Mie theory. The look-up table decides the extinction and single scattering albedo properties of water or ice particles in the cloud during I3RC simulation. The input file contains many parameters. The name, brief description, and value of these parameters that used in this study are listed in Table 4.

b) Physical properties of clouds to domain

This program creates a domain file given input files describing the 3D distribution of cloud particle properties, including cloud water content (WC) and effective radius. The extinction and single scattering albedo properties of particle can be read from the scattering phase function look-up table from the last program. The extent of the 3D domain is defined by the parameters listed in Table 5. The 3D distribution of cloud particle is defined on a 3D grid. Besides the basic coordinate information, such as X, Y, Z axes and their cell sizes, the

temperatures of vertical boundaries are required to be defined. The spatial location of each

cell within the 3D domain is defined by its 3D coordinate and its properties are WC and

effective radius as shown in

Table 6. Both two cloud properties can be extracted from MOD06 product. The specific

structure of cloud derived from MODIS products will be discussed in section 3.3.3.2.

Table 4 The input parameters of making Mie table program

Parameter name	Value	Description
WAVELEN1	0.3	Starting wavelength (μm) of band
WAVELEN2	2.8	Ending wavelength (μm) of band
AVGFLAG	A	Average or center properties over wavelength (A, C)
DELTAWAVE	0.1	Wavelength interval if averaging (μm)
PARTYPE	W	Water, Ice, or Aerosol spherical particles (W, I, A)
DISTFLAG	G	Particle size distribution type: Gamma (G) or Lognormal (L)
ALPHA	7	The shape parameter α in Gamma distribution
NRETAB	-35	Number of tabulated effective radius (<0 for log-space effective radius)
SRETAB	1.0	Starting tabulated effective radius (μm)
ERETAB	30.0	Ending tabulated effective radius (μm)
MAXRADIUS	100.0	Maximum particle size (μm)

Table 5 The extent of the 3D domain

Parameter name	Description
nX	Number of X grid cells, east-west direction
nY	Number of Y grid cells, south-north direction
nZ	Number of Z grid cells, vertical direction, can be viewed as layer
deltaX deltaY	X and Y grid cell size in km
Zlevels(1:nZ+1)	increasing heights of cell boundaries in km
Temps(1:nZ+1)	temperatures of boundaries (K)

Table 6 The information of each cell

Parameters	Description
iX,iY,iZ	indices from 1 to nX,nY,nZ
LWC	cloud liquid water content (g/m³)
Reff	the effective radius (μm)

c) Monte Carlo driver

The last program implements Monte Carlo simulation for I3RC model. It computes pixel by pixel outgoing monochromatic solar fluxes at the bottom of the domain. The cloud transmittance is the ratio of the solar flux at the bottom of the domain to original solar flux. The input parameters of this program are listed in Table 7. The "solarFlux" is set to 1. The "solarMu", "solarAzimuth", "solarAzimuth" are extracted from MODIS geolocation data. The "surfaceAlbedo" are extracted from MODIS land data. The "numPhotonsPerBatch" is set to 100000 and the "numBatches" is set to 100. The other input parameters keep their default values.

Table 7 The input parameters of Monte Carlo driver

Parameters	symbol	Description
solarFlux		Normalized solar flux
solarMu	$cos\theta_{sol}$	Cosine of the solar zenith angle
solarAzimuth	γ_{sol}	Azimuthal direction of incoming photons
surfaceAlbedo	ρ	Albedo of the ground surface
numPhotonsPerBatch		The number of photons per batch in the MC calculation
numBatches		The number of batches in the MC calculation

3.3.3.2   3D Modeling Cloud structure

In this study, the 3D cloud structure is rebuilt based on MODIS products. Although it is impossible to model accurate cloud structure, like multi-angle data or radar data, a coarse model can be built to approximate the 3D structure and properties of clouds based on MODIS products.

For parallax effect, the oblique viewing angle results in that the apparent position of a single cloud layer as seen from the satellite will be off its true location on the ground (Figure 9). The apparent position of a cloud can be viewed as the projecting the real cloud at a cloud base height to the surface for the given viewing angle. Thus, the parallax effect can be corrected if the cloud top height and the viewing geometry are known.

The horizontal displacement of a single cloud layer on the surface can be computed as:

$$D_g = CTH\sin\theta_{sen}, \tag{3.18}$$

where: D_g is the displacement distance on the surface; CTH is the cloud top height. The transition on the X and Y axes can be described as:

$$D_x = D_g\sin\gamma_{sen}, \tag{3.19}$$

$$D_y = D_g\cos\gamma_{sen}, \tag{3.20}$$

where: D_x and D_y are the transition on the X and Y axes, respectively. If the values of the CTH were close within a scene, the shift at two axes of the cloud layer after the correction

can be calculated by the median value of all D_x and D_y.
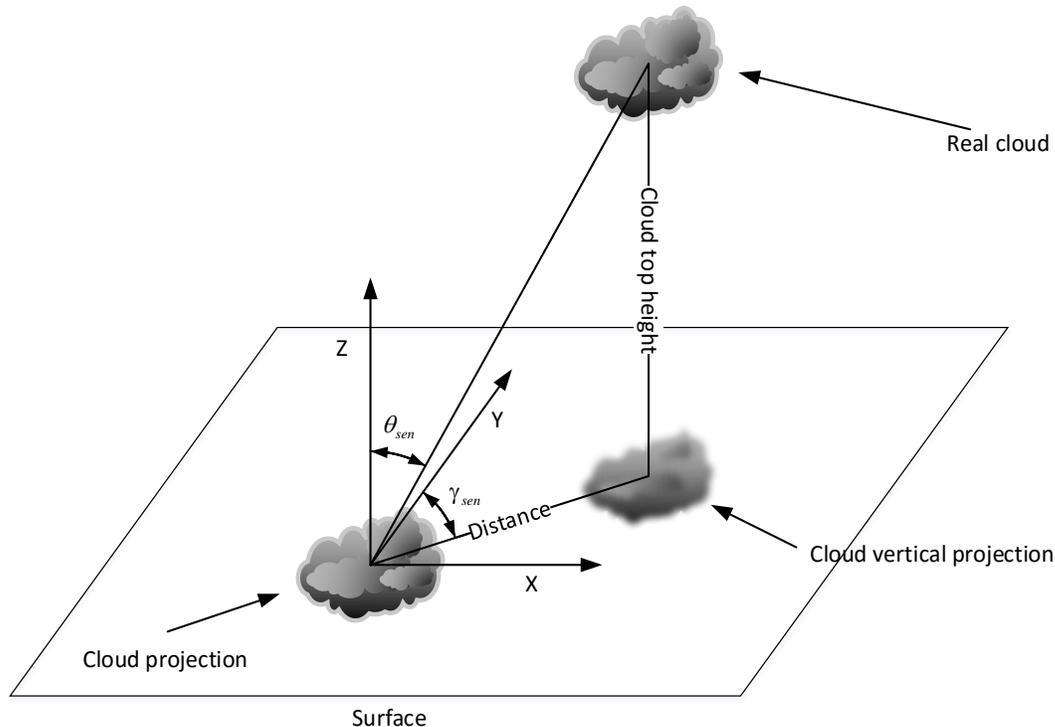
Figure 9 Principle of parallax correction. The horizontal displacement of the apparent position is a function of the cloud base height and satellite zenith and azimuth angle. A correction becomes necessary if the distance between the apparent and the true location exceeds the spatial resolution (Wyser et al. 2002).

In fact, clouds seldom only exist at a single layer with similar height, but have 3D structure, which includes different heights and thicknesses. The extent of cloud horizontal displacement can vary due to the different height. The pixel by pixel parallax correction for multiple layers of cloud structure may result in gaps (Figure 10). There are some gap pixels within cloud area after parallax correction (Figure 11(b)) compared with original image (Figure 11(a)). Lábó, Kerényi, and Putsay (2007) proposed an algorithm that can correct the

gaps. Based on Lábó's model, the flowchart of gap correction algorithm is shown in Figure

12 and example of filling gaps is shown in (Figure 11(c)).
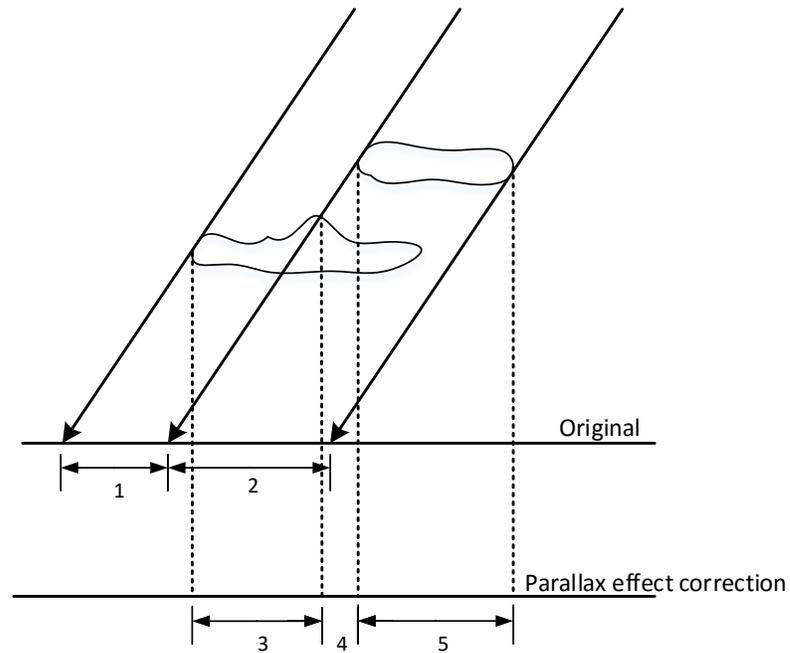
Figure 10 Possible gap results from parallax correction. No.1 and No.2 areas are the projection of the cloud in from a satellite viewing angle. No.3 and No.5 areas are the vertical projection of cloud after parallax correction. No.4 area is the gap.

| (a) Original cloud | (b) After correction | (c) Fill gap |
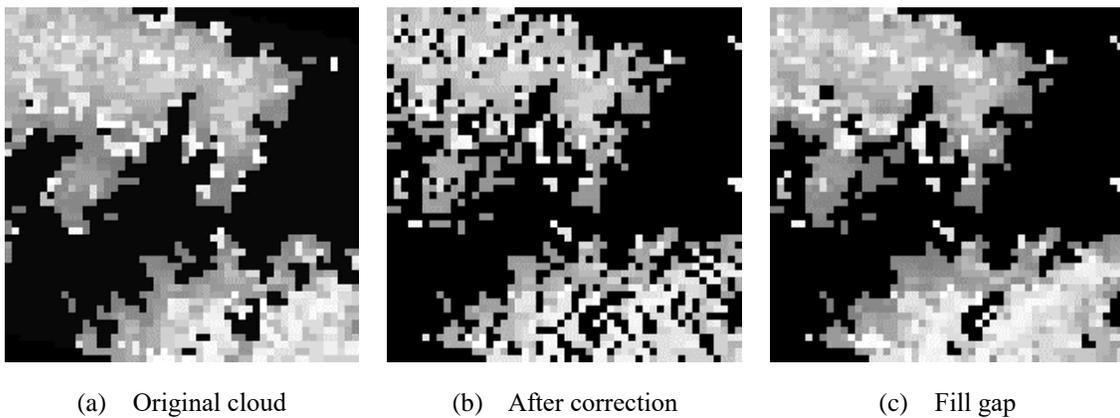
Figure 11 Gaps appearing on the image due to parallax correction
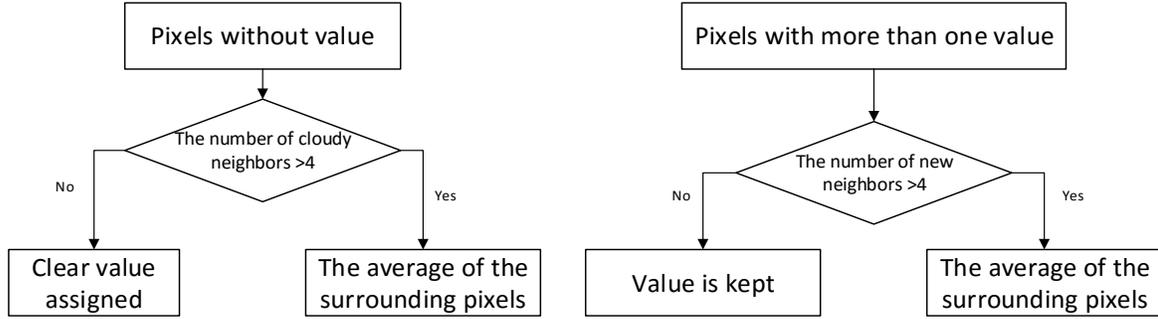
Figure 12 Flow chart of gap correction

Based on the cloud base height above ground level (AGL), cloud can be classified into three groups as shown in Table 8. The cloud base height can be computed by (Meerkötter and Bugliaro 2009):

$$CBH = CTH - CGT, \tag{3.21}$$

where: CBH is cloud top height; CGT is cloud geometrical thickness, which can be computed as (Sharma et al. 2015):

$$CGT = \frac{CWP}{WC}, \tag{3.22}$$

where: CWP is cloud water path, which can be acquired from MOD06, and the parameter's view angle dependence is not strong in MOD06 (Horváth, Seethala, and Deneke 2014); WC is water content, which can be liquid water content (LWC) or ice water content (IWC). According to (Hess, Koepke, and Schult 1998), the value of continental WC varies between different types of cloud and the reference values are shown in Table 9. However, cumulus

and cumulonimbus can grow vertically and be really thick, so that these two types of clouds

cannot be classified into any group based on cloud base height.

Table 8 Approximate Height of Cloud Bases AGL for Various locations (Ahrens 2016)

Cloud group	Tropical region	Mid-latitude region	Polar region
High clouds: Cirrus, Cirrostratus, Cirrocumulus	6000 to 18000 m	5000 to 13000 m	3000 to 8000m
Middle clouds: Altostratus, Altocumulus	2000 to 8000 m	2000 to 7000 m	2000 to 4000 m
Low clouds: Stratus, Stratocumulus, Nimbostratus	0 to 2000 m	0 to 2000 m	0 to 2000 m

Table 9 Microphysical properties of cloud models (Hess, Koepke, and Schult 1998)

Component	r_{eff} (μm)	WC (g/m^3)
Stratus	7.33	0.28
Cumulus (clean)	5.77	0.26
Cumulus (polluted)	4.00	0.30
Cirrus 1: −25°C	91.7	0.0260
Cirrus 2: −50°C	57.4	0.00193

According to a study of Liu et al. (2015), the clouds over the Tibetan Plateau (TP) can

be generally classified into the high clouds with cloud top above 6 km AGL, and middle to

low-level clouds with a top below 4 km AGL as shown in Table 10. The clouds over Lhasa

are assumed to have two layers, which are upper layer and lower layer. As shown in Figure

13, there are four cases of two cloud layers. Sometimes, such as cumulonimbus, the two

layers can be merged into a single layer.

However, it is impossible to get the real structure of two cloud layers from one viewing

angle, especially the overlay case. Thus, there are two assumptions of cloud structure in the

same scene: overlook the overlay cases and clouds are thin.

Table 10 Characteristics of different types of cloud over TP (Liu et al. 2015)

Cloud type	The range of AGL	Maximum cloud depth
Cumulus	2 km to 4 km	2 km
Altocumulus	6 km to 10 km	4 km
Altostratus	6 km to 11 km	5 km
Stratocumulus	Around 2 km	< 1 km
Nimbostratus	0 km to 8 km	8 km
Deep convective case	0 km to 16.5 km	16.5 km
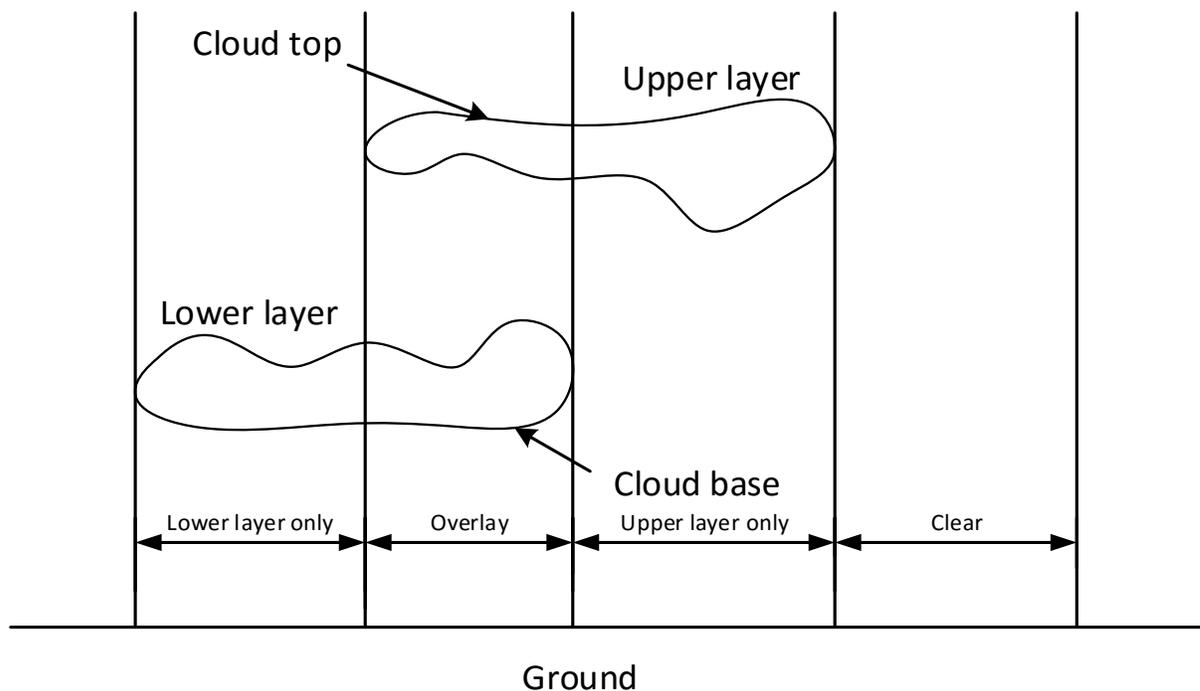
Figure 13 Four cases two cloud layers

The first assumption is to simplify the cloud structure model by overlooking the overlay

cases. For the second assumption, because the finest spatial resolution of MODIS products

used in the study is 1 km, the horizontal displacement caused by cloud thickness seldom exceeds the spatial resolution, except cumulus and cumulonimbus. The area of very thick cloud projection on the surface can be greater than the area of cloud vertical projection on the surface as shown in Figure 14. Thus, a rainy day with thick cloud might have different cloud projection from a cloudy-sky day. This case was not considered in this study and the clouds are assumed to be thin. But the accuracy of solar irradiance estimation in rainy days should be examined separately.
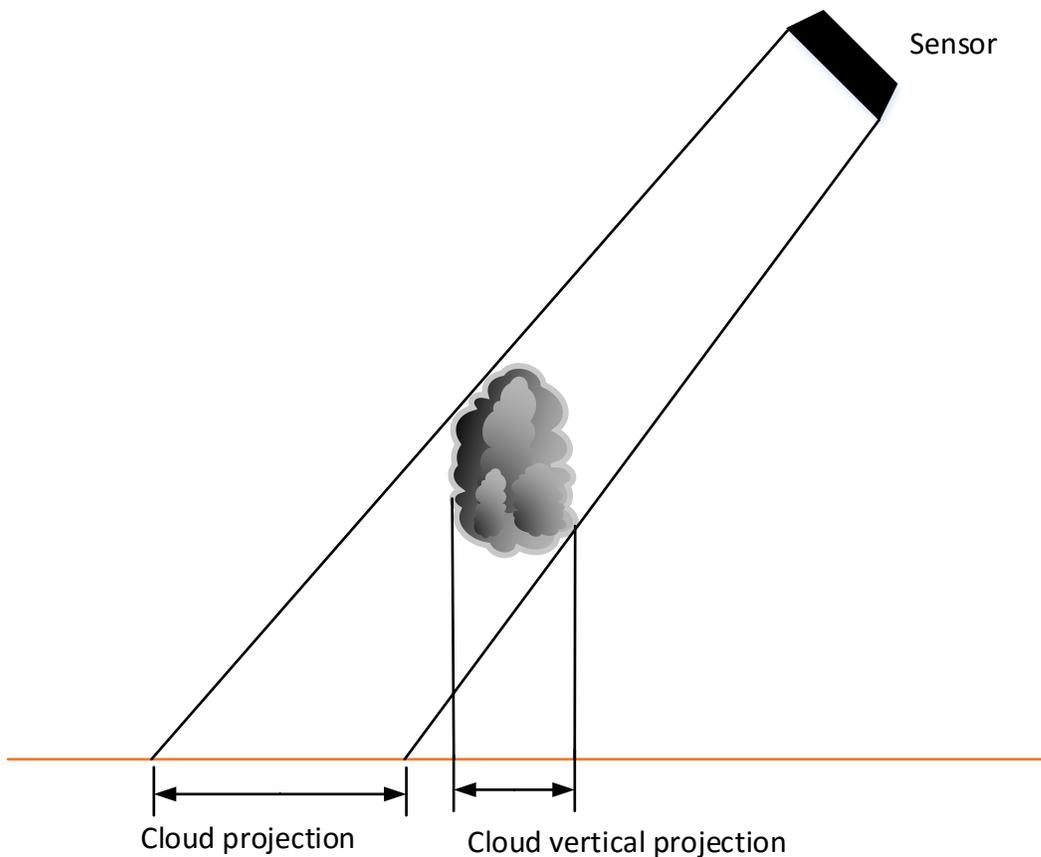
Figure 14 The projection of very thick cloud

Before building cloud structure, three types of cloud, which are water, ice, and mixed, should be determined. Figure 15 shows the flowchart to determine the types of cloud. The LWC of water cloud is assumed to be equal to stratus' LWC in Table 9. The minimum cloud base height is assumed to be 1km. If the LWP of a pixel is very large, the cloud will be assumed as cumulonimbus and its LWC would be the average of LWP over the cloud depth. The temperature of cloud was assumed to be a function of height and it can be calculated by:

$$T_{cloud} = T_{top} - \Gamma_e (CTH - H_{cloud}), \qquad (3.23)$$

where: T_{cloud} is the temperature of a cloud layer in °C; H_{cloud} is the height of the layer AGL; Γ_e is the environmental lapse rate, which is assigned as 6.5°C/km.

An threshold of cloud icing temperature is assigned as -20 °C and the IWC of cloud id divided by the temperature in two groups, -40 ~ -20 °C and -65 ~ -40 °C (Liou et al. 2008). The cloud above -20 °C is assumed as water and the cloud below that temperature is assumed as ice. The maximum ice cloud depth is assumed as 3 km (NASA 2011). If the ice cloud depth is greater than the threshold, the rest fraction of cloud is assumed to be water cloud, which can be either stratus or cumulonimbus. The effective radius of each cloud column is acquired from MOD06.
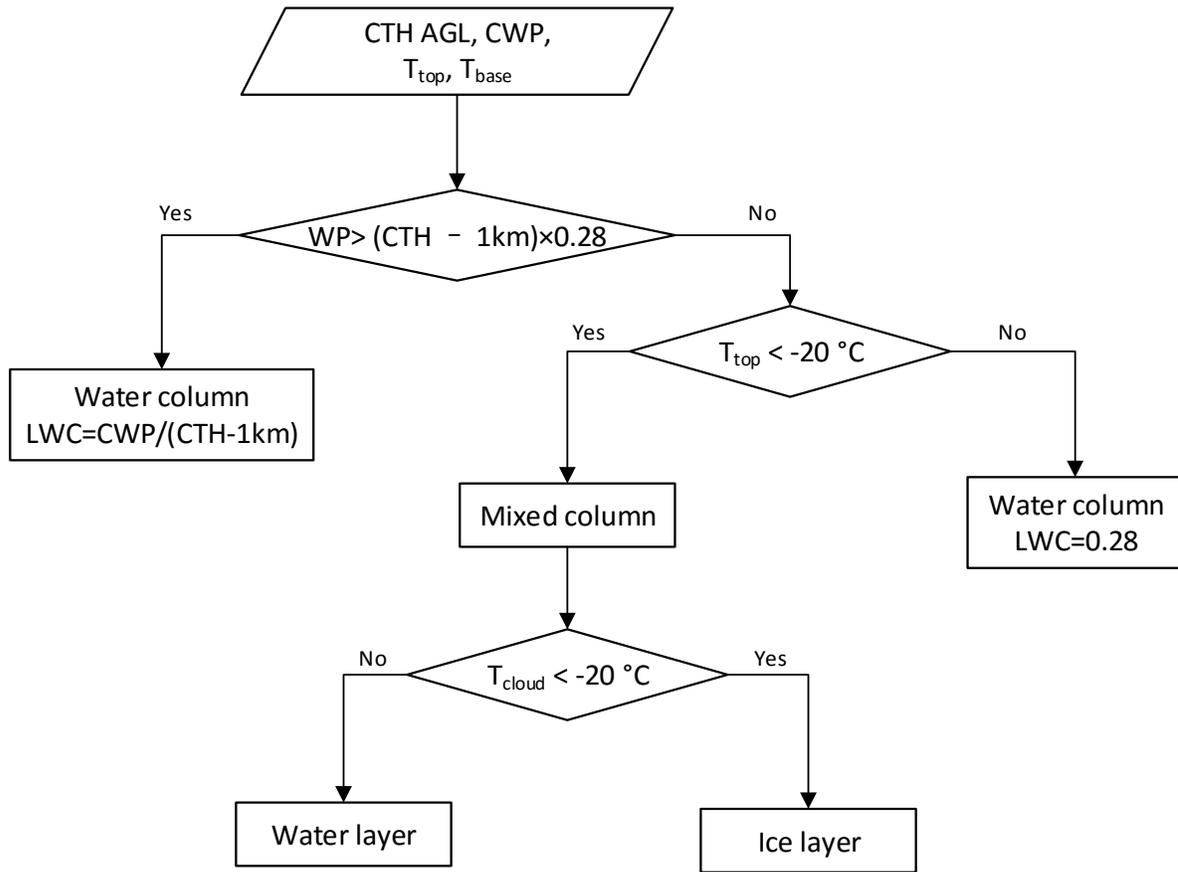
Figure 15 The flowchart to determine the types of cloud

However, some cloud pixels of MOD06 only contain the cloud top height and temperature and the CWP and effective radius data are missing. For example, the cloud areas in Figure 16(b) always have cloud top height value in Figure 16(a). However, the Figure 16(c) does not have CWP value in most cloud areas. In this case, a CWP value would be assigned to these pixels as shown in Figure 17. Their effective radiuses would be based on their cloud type in Table 9.

| (a) Cloud top height | (b) True color | (c) Water path |
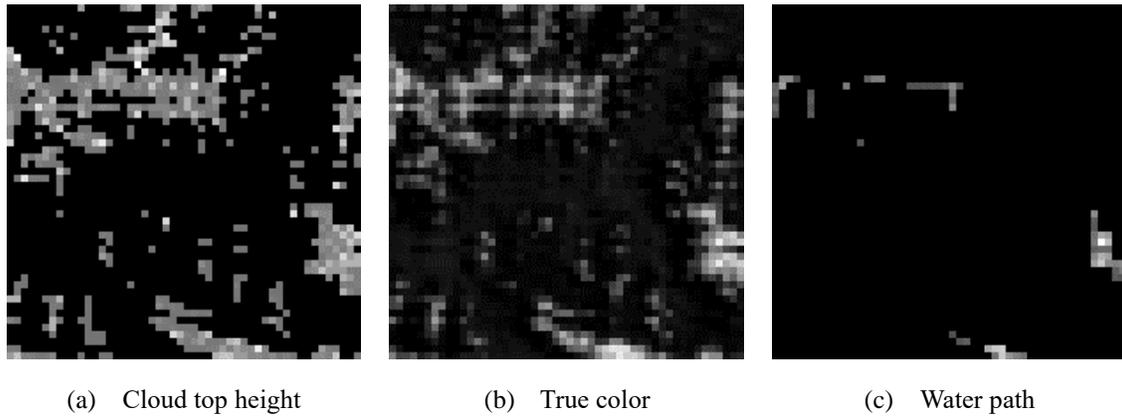
Figure 16 The data missing problem of MOD06. The three images were derived from MODIS image at 6:45 AM GMT the 271st day in 2012. The coordinates of the center pixel of the image is (1180, 32). The gray or white color of the pixel in the three figures means the pixel has value.
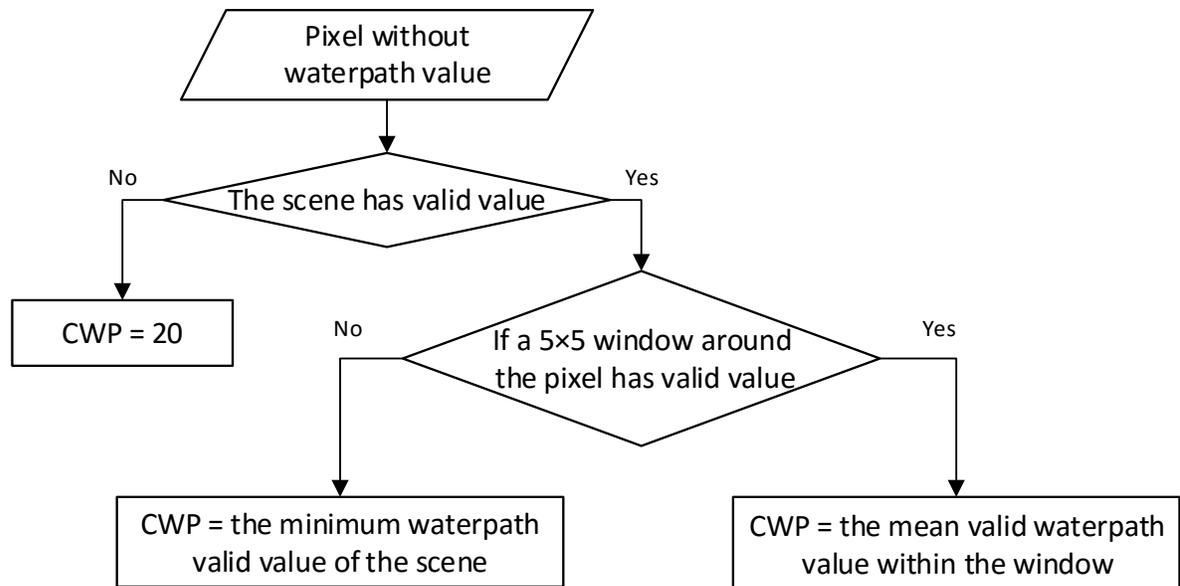
Figure 17 The flow chart to assign CWP value to the pixel without CWP value

3.4 Validation methods

By comparing the modeled instantaneous DSSR from satellite with the ground observations, the performance of models are usually quantitatively validated using statistical

indices, like Bias, root mean square error (RMSE), mean absolute percent error (MAPE), and coefficient of determination (R^2). The formulas for these statistics are listed in Table 11.

Table 11 Definitions of the statistical indices used to evaluate performance of the models (Huang, Liu, and Liang 2012)

Statistical indices	Description	Mathematical definition
$\bar{E}$ or $\bar{O}$	Mean of all n estimated DSSR or observed DSSR	$\bar{E} = \frac{1}{n}\sum_{i=1}^{n} E_i$ or $\bar{O} = \frac{1}{n}\sum_{i=1}^{n} O_i$
Bias	Bias error of $\bar{E}$ to $\bar{O}$	$\frac{1}{n}\sum_{i=1}^{n}(E_i - O_i)$
RMSE	Root mean square error	$\left[\frac{1}{n}\sum_{i=1}^{n}(E_i - O_i)^2\right]^{\frac{1}{2}}$
MAPE	Mean absolute percent error	$\frac{100}{n}\sum_{t=1}^{n}\left|\frac{E_i - O_i}{O_i}\right|$
R^2	Coefficient of determination	$\dfrac{\sum_{i=1}^{n}(E_i - \bar{E})(O_i - \bar{O})}{\left[\sum_{i=1}^{n}(E_i - \bar{E})^2 \sum_{i=1}^{n}(O_i - \bar{O})^2\right]^{\frac{1}{2}}}$

Chapter 4   Results

The description of the results after data preprocessing is presented at the beginning of the section. Then, the instantaneous DSSR estimates of the algorithms are evaluated by comparing them with the ground observation and FLASHFlux data. In addition, the effects of parallax correction, DSSR variability, and the distance to the closest cloud would be discussed in this section.

4.1 Data preprocessing

4.1.1   Ground observation

During the study time period from July 12th, 2012 to October 31st, 2013, there are 477 days and 680030 records that have both solar irradiance value and timestamp. Missing data and deleted data only account for 1.0% of the whole ground observation data. Most of missing data concentrate in a time period from October 2nd, 2012 to October 5th, 2012 and obvious inconsistent features happened on October 5th, 2012 (Figure 18). The maximum value on October 5th, 2012 (Figure 18(a)) is much lower than the maximum value on October 6th, 2012 (Figure 18(b)). The maximum values of clear-sky days with the same date in different years should be very close theoretically. However, the data from July 12th, 2012 to October 5th, 2013 are also much lower than the data from July 12th, 2013 to October 5th, 2012

and always greater than the theoretical maximum value of DSSR. Thus, the data after

October 5th, 2012 should be corrected.

(a) Minutely (GMT+8) solar irradiance scatter
plot on October 5th , 2012

(b) Minutely (GMT+8) solar irradiance scatter
plot on October 6th , 2012
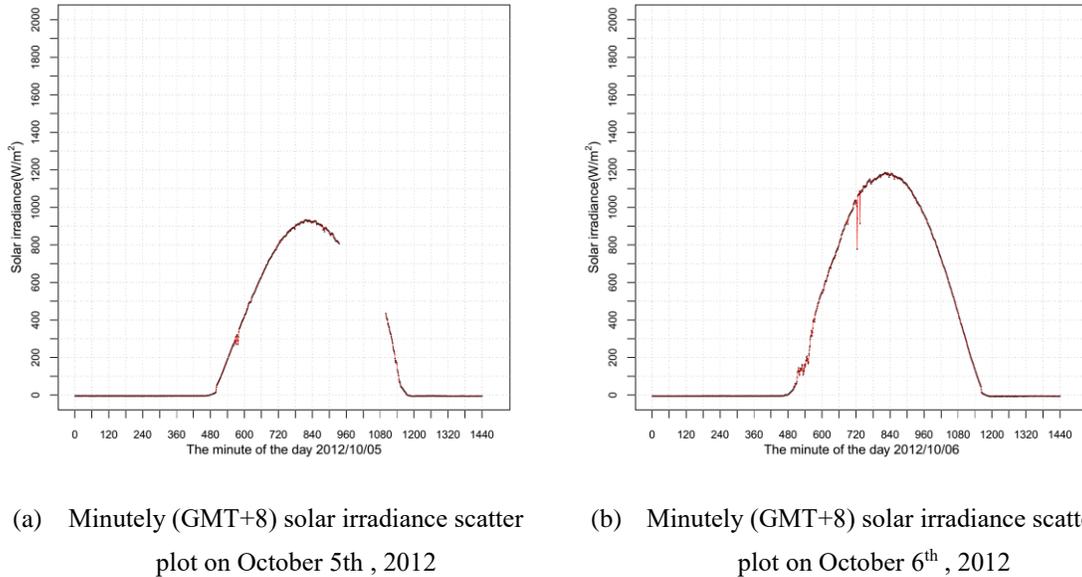
Figure 18 The inconsistent patterns of ground observation data

The days that under clear-sky conditions from July 12th to October 5th in different years

were used to build the correct function. The records in 2012 were dependent variable and

the records in 2013 were independent variable. However, from July to September are the

raining season of Lhasa, so that it is very hard to find clear-sky days in both 2012 and 2013,

which has perfect sinusoidal curves like Figure 18 (b). Thus, the time periods when the curve

is a fraction of the sinusoidal curve without great fluctuations are under clear-sky conditions.

For example, in Figure 19, the time period from the 570th minute to 720th minute can be

viewed as under clear-sky conditions in both 2012 and 2013. The time periods that both 2012

and 2013 records were under clear-sky conditions were listed in Table 12.

(a)   2012   (b)   2013
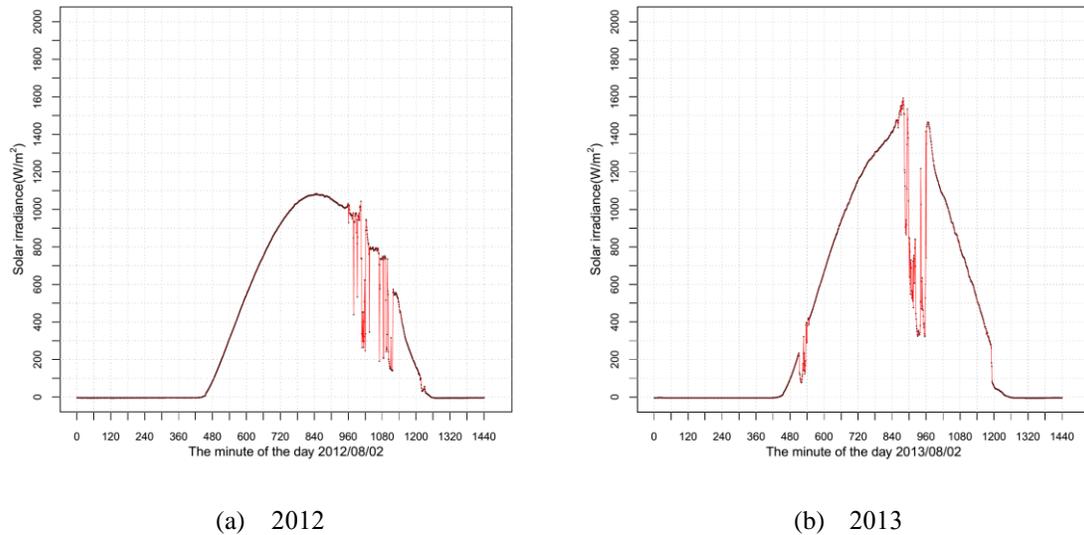
Figure 19 The DSSR curve on August 2nd

Table 12 The time periods that both 2012 and 2013 records were under clear-sky conditions

Date	Time period (the number of minute of day)
08-02	570-720
08-24	1130-1170
09-05	1110-1150
09-07	500-720
09-27	1080-1140
09-28	510-600; 1080-1140

The scatter plot of the independent variable and the dependent variable is shown in

Figure 20. The relationship between the independent variable and the dependent variable can

be modeled with a linear regression, which can be expressed by:

$$y = 0.7916x - 4.2252. \tag{4.1}$$

The R^2 of the regression analysis is 0.999, which means the linear relationship between independent and dependent variables is very strong. The slope and regression coefficients are both statistically significant. Thus, the factors that result in corrupted data from August 2nd, 2013 to September 28th,2013 can be simulated by a regression function. Because the corrupted data after October 5th, 2012 are sequential without gaps, the operational status of the pyranometer during that time was assumed to be the same. Therefore, the linear regression function was used to correct all the corrupted data after October 5th, 2012. However, the problematic ground observations reduce the reliability of calibrations in the following sections.
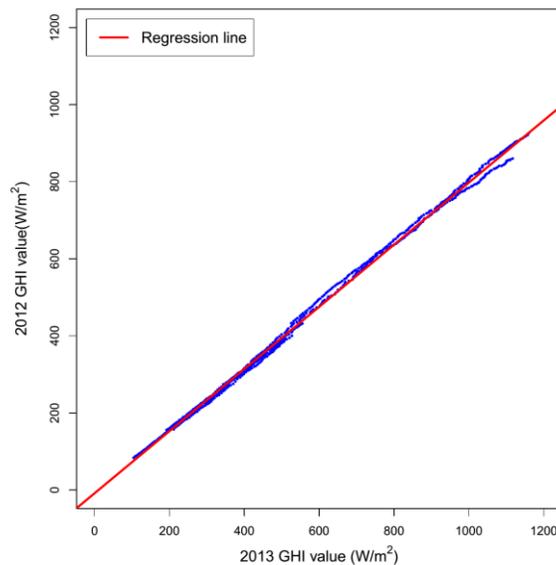
Figure 20 The scatter plot of the DSSR value in 2012 and 2013

56

4.1.2   Remote sensing data

4.1.2.1   MODIS products

From July 12th, 2012 to October 31st, 2013, there are 946 scenes that contain the station and have valid MODIS products from Terra or Aqua. Almost every day during the study period has two scenes. One scene is captured by Terra around 10:30 AM at local time and the other scene is captured by Aqua around 1:30 PM at local time. The transit time of all the scenes around two time points are relatively evenly distributed as shown in Figure 21.
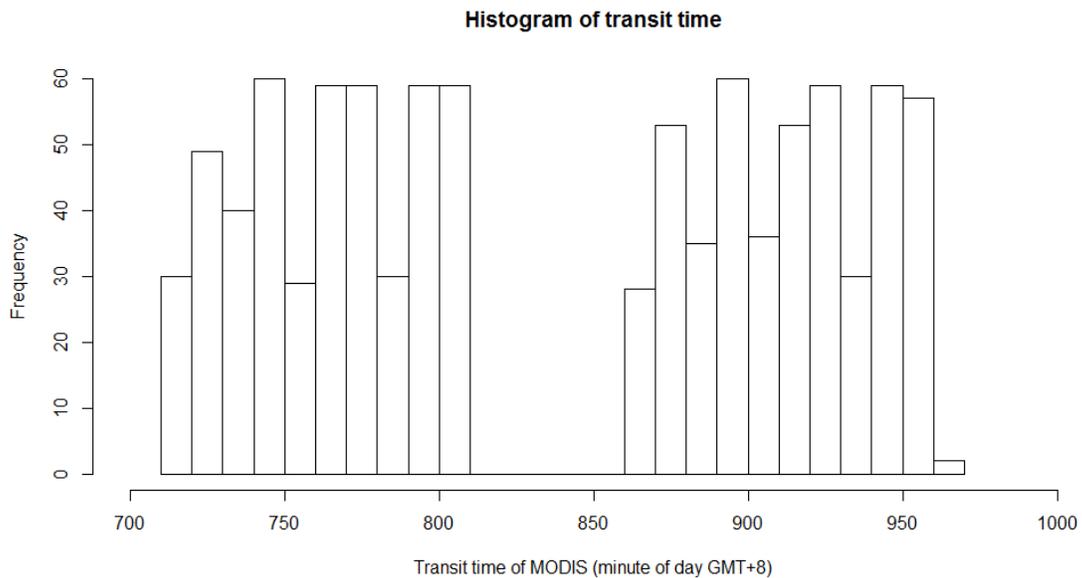
Figure 21 The histogram of transit time of all scenes of MODIS products

The date that only has one scene accounts for 0.85% of all valid scenes and all the dates are listed in Table 13. The distribution of the date is dispersed. Five scenes are captured by

Terra and three are captured by Aqua.

Table 13 The date only has one scene

Year	Day of year	Minute of day	MOD/MYD
2012	229	755	MOD
2012	307	767	MOD
2013	45	718	MOD
2013	94	954	MYD
2013	108	774	MOD
2013	124	774	MOD
2013	214	905	MYD
2013	215	948	MYD

Due to the algorithms of MODIS products, MODIS sensor, and weather, lots of pixels in MODIS products do not have value. Thus, when the value parameters that list in Table 1, Table 2, and Table 3 were extracted from the pixel contained the station during the study period, there are so much missing data (Table 14). The θ_{sol}, γ_{sol}, θ_{sen}, γ_{sen}, and T_s do not have missing data. w has a small fraction of missing data. l has a moderate fraction of missing data. CTH, T_{top}, CWP, r_{eff}, $\delta(\lambda)$, λ, and α have a large fraction of missing data. In fact, missing data do not mean failures to acquire the value variables and they are meaningful in some cases. For example, the missing value of CTH, T_{top}, CWP, and r_{eff}, mean the pixel is cloud free. The different percentages of CTH and CWP reflect the phenomenon in Figure 16. However, some variables are indispensable to estimate solar irradiance, such as w, l, $\delta(\lambda)$, λ, and α. Therefore, it is necessary to impute the missing data.

Table 14 The percentage of missing data of each parameter

Variable name	symbol	The percentage of missing data
Solar zenith angle	θ_{sol}	0%
Solar azimuth angle	γ_{sol}	0%
Sensor zenith angle	θ_{sen}	0%
Sensor azimuth angle	γ_{sen}	0%
Cloud top height AMSL	CTH	63.2%
Cloud top temperature	T_{top}	63.2%
Cloud water path	CWP	75.1%
Cloud effective radius	r_{eff}	75.1%
Surface temperature	T_s	0%
Total column precipitable water vapor	w	3.9%
Thickness of the ozone layer	l	34.1%
AOD and wavelength	$\delta(\lambda), \lambda$	77.4%
Ångström exponent	α	77.4%

The missing value of l was imputed from an average value of the closest former and latter valid records. The pixels without w value are always covered by very high cloud, so that the missing value of w was imputed by zero. $\delta(\lambda)$, λ, and α are not the inputs of Yang's model, but are inputs of Equation (3.17) to compute β. Because Tibet's air is always very clean, β should be very small in most scenes. The missing value of β was assigned a mode of existing β, which is 0.0256. The substituted β value is small, which means the air in Lhasa is very clean and the aerosol loading's effects on DSSR would be very small as well.

4.1.2.2   FLASHFlux data

From July 12th, 2012 to October 31st, 2013, there are 896 scenes that contain the station

and have valid downward shortwave surface flux value from Terra or Aqua. The number of the scenes is a little less than the number of MODIS products' scenes. The numbers of day of year of missing scenes are shown in Figure 22. The missing scenes are not uniformly distributed. The months from July to October have more missing scenes than other months.
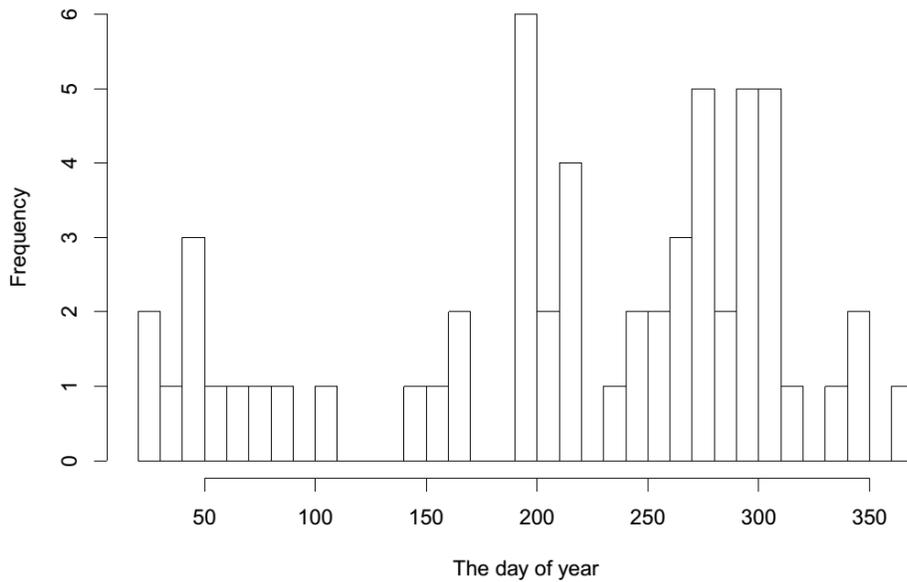
Figure 22 The histogram of missing scenes of MODIS products

4.2 Instantaneous DSSR estimates

The instantaneous DSSR on all-sky days were estimated using the algorithm discussed in section 3.3. The estimates are separated into two groups, clear-sky and cloudy-sky conditions. For clear-sky conditions, the only Yang's model was applied to estimate DSSR. For cloudy-sky conditions, both Yang's model and I3RC model were applied to estimate DSSR.

4.2.1 DSSR estimates using Yang's model

Three variables (w, l, β) are imputed with substituted values. Missing values account for a small fraction of w and the substituted values of it are close to the real values. l has a relatively small effect on DSSR and it does not vary large in a short period. Therefore, the imputation method would not lead to great errors to DSSR results. However, 77.4% of scenes do not have β at the station pixel in this study and aerosols can have great impacts on DSSR, so that the influences of the substituted values of β on DSSR should be examined.

4.2.1.1 The effects of substituted β on DSSR estimates

Because most of scenes do not have valid β, the scenes that can compute valid β are suitable to examine the influences of the substituted values of β on DSSR using Yang's model. The comparison between the ground measurements and DSSR estimates that β exists using Yang's model is shown in Figure 23(a). The comparison between the ground measurements and DSSR estimates that existing β are replaced by a constant using Yang's model is shown in Figure 23(b). The count, bias, RMSE, MAPE, and R^2 for the two cases is shown in Table 15. The case that calculates DSSR estimates using substituted β has much lower bias, RMSE, and MAPE values and higher R^2 value than the case that calculates DSSR estimates using original β. Generally, the points in Figure 23(b) are more close to the 1:1 line than the points in Figure 23(a). These evidences show that using substituted β to

compute DSSR estimates has higher accuracy than using original β.

Table 15 Count, bias, RMSE, MAPE, and R^2 for DSSR when β exists

	Count	Bias (W/m^2)	RMSE (W/m^2)	MAPE (%)	R^2
Original β	199	-66.908	152.321	14.780	0.366
Substituted β	199	-26.209	109.997	10.981	0.621

(a) β exists   (b) existing β are replaced by a constant
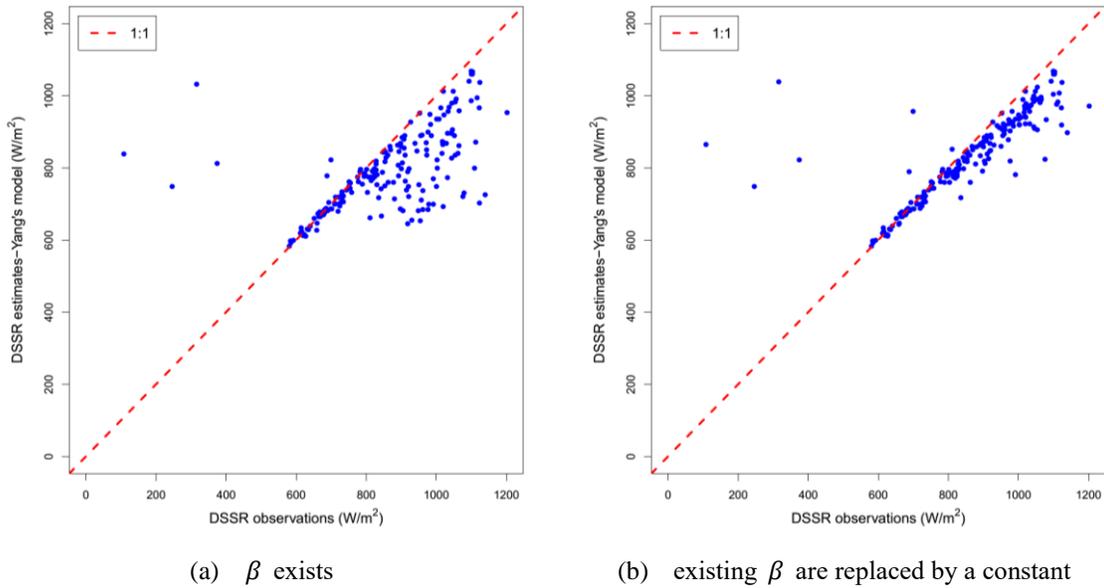
Figure 23 Scatter plots of ground measurements versus DSSR estimates using Yang's model

However, there are still many points are far away from the 1:1 line. Because the spatial

resolution of MOD04 is 10 km, existing β does not mean that the area around the station in

that scene is cloud free and original β can be greater than real value. Using Yang's model

with substituted β might reduce the clouds' effects on DSSR estimates, so that the case

using substituted β has better performance. To get rid of the effects of clouds, the clear

percent variable was taken into account. As shown in Figure 27 (b), when the clear percent

is greater than 90%, there is no point of FLASHFlux data that deviates far from the 1:1 line. Thus, it is assumed that clouds would not heavily affect the DSSR estimates using Yang's model when clear percent is greater 90%.

The comparisons between the ground measurements and DSSR estimates for the two cases using Yang's model when the clear percent is greater than 90% are shown in Figure 24. The count, bias, RMSE, MAPE, and R^2 for the two cases when clear percent is greater than 90% is shown in shown in Table 16.

(a) β exists

(b) existing β are replaced by a constant
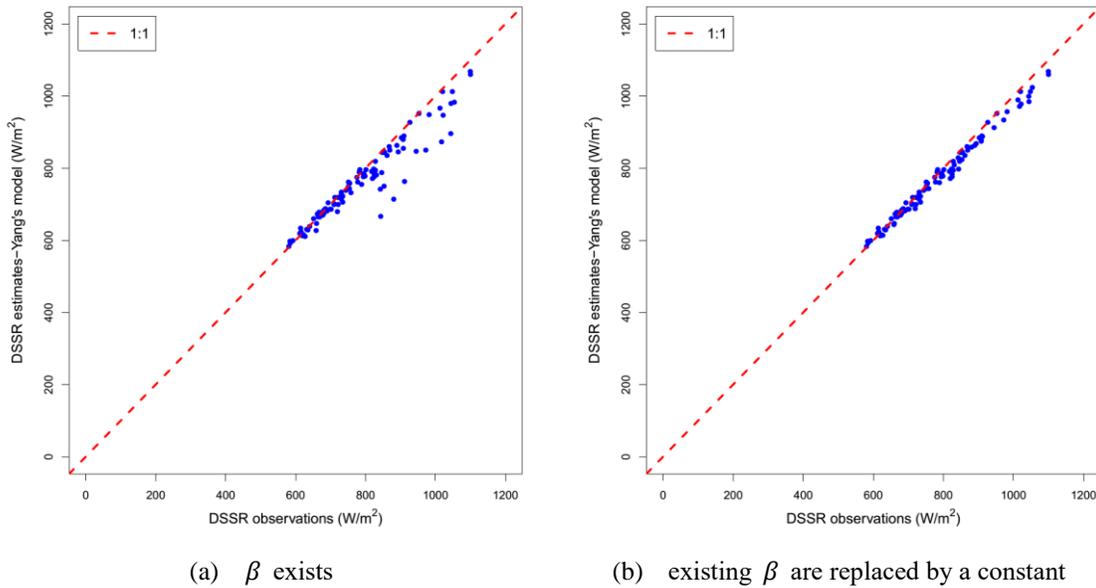
Figure 24 Scatter plots of ground measurements versus DSSR estimates using Yang's model when clear percent is greater than 90%

Table 16 Count, bias, RMSE, MAPE, and R^2 for DSSR estimates when β exists and clear percent is greater than 90%

	Count	Bias (W/m^2)	RMSE (W/m^2)	MAPE (%)	R^2
Original β	104	-23.876	46.494	3.177	0.914
Substituted β	104	-12.672	21.038	1.944	0.990

The results are similar to the cases without considering clear percent. The case that calculates DSSR estimates using substituted β still has much lower bias, RMSE, and MAPE values and higher R^2 value than the case using original β. Some points in Figure 24(a) obviously deviate from the 1:1 line. Most of the points in Figure 24(b) are close to 1:1 line and the linear pattern of all points is obvious. Thus, using substituted β to compute DSSR estimates for clear-sky conditions still has higher accuracy than using original β. The changes of DSSR estimates are obvious when large original β was replaced by mode as shown in Figure 25.
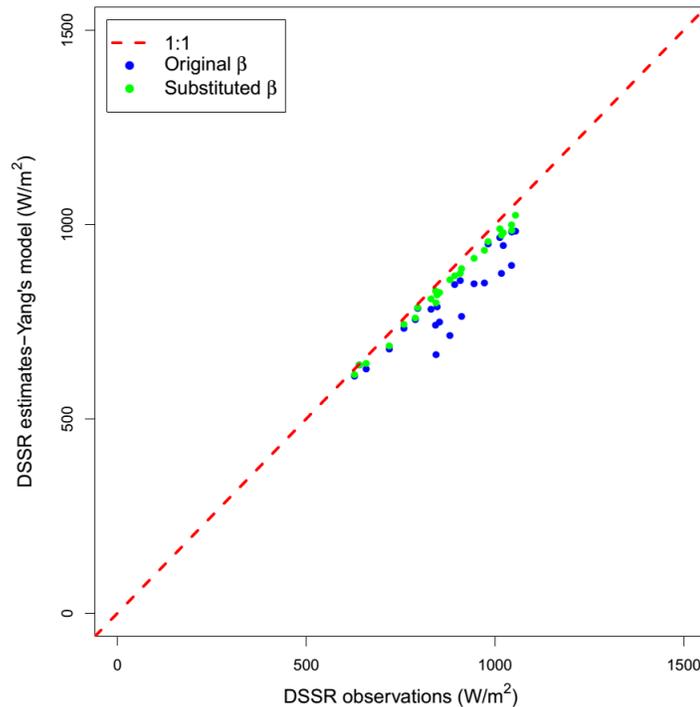

Figure 25 Scatter plots of ground measurements versus DSSR estimates using Yang's model when clear percent is greater than 90% and original β are greater than 0.0256

The histogram of the day of year of the scenes that original β are greater than 0.0256 is shown in Figure 26. The months from July to October have higher frequencies than other months. There are a few scenes that have large original β in the period from December to April.
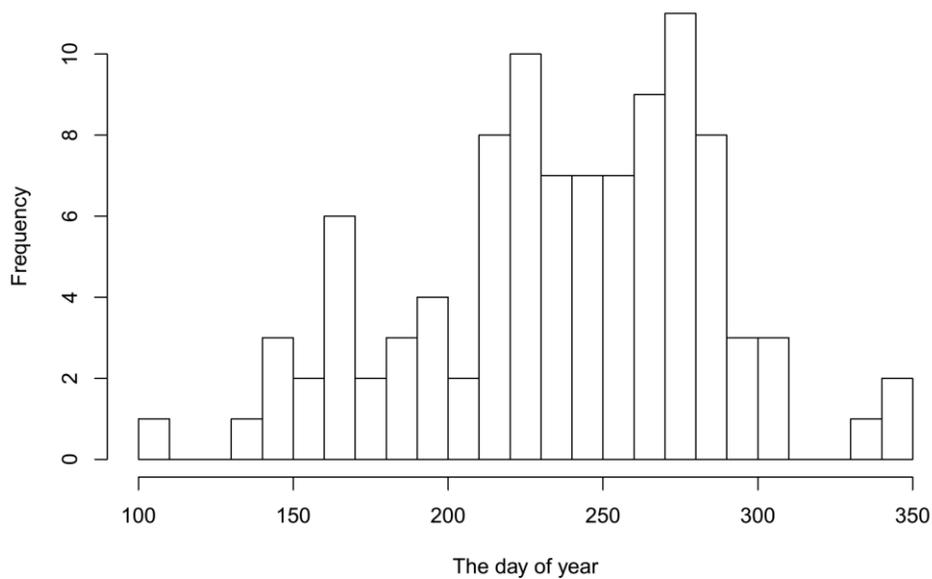
Figure 26 The histogram of the scenes that existing β is greater than 0.0256

4.2.1.2　Comparisons for clear-sky conditions

Yang's model can only calculate the clear-sky transmittance, so that only ground observations on clear-sky days are suitable to examine the performance of modeled instantaneous DSSR using Yang's model. Three different clear percent values, which are 90%, 95%, and 100%, are used for identifying the different levels of clear-sky conditions.

The "CERES downward SW surface flux - Model B" (DSWSF) dataset of FLASHFlux was used as a reference to evaluate the performance of Yang's model.

The comparisons between ground measurements and DSSR estimates of Yang's model or FLASHFlux with three clear percent values are shown in Figure 27. The clear percent, count, bias, RMSE, MAPE, and R^2 is shown in Table 17. All points in six subfigures of Figure 27 are very close to 1:1 line, which means the linear relationship between the independent variable and dependent variable is very strong. There is no obvious difference between the DSSR estimates of Yang's model and FLASHFlux at three clear percent levels. The only difference is that the blue bands in Figure 27(a), (c), and (e) are a little thinner than the blue bands in Figure 27(b), (d), and (f), separately. The DSSR estimates of Yang's model have lower bias, RMSE, and MAPE values and higher R^2 value than the DSSR estimates of FLASHFlux at three clear percent levels. Thus, Yang's model can estimate the DSSR for clear-sky conditions better than the model of FLASHFlux.

Table 17 Count, bias, RMSE, MAPE, and R^2 for DSSR estimates of Yang's model and FLASHFlux at conditions with three different three clear percent values

Model	Clear percent	Count	Bias (W/m^2)	RMSE (W/m^2)	MAPE (%)	R^2
Yang's model	>90	233	-14.232	20.725	2.159	0.986
Yang's model	>95	214	-12.662	18.992	2.003	0.987
Yang's model	=100	139	-10.155	15.620	1.700	0.989
FLASHFlux	>90	233	-31.993	36.134	4.288	0.980
FLASHFlux	>95	214	-30.446	34.121	4.118	0.983
FLASHFlux	=100	139	-28.106	31.135	3.915	0.985
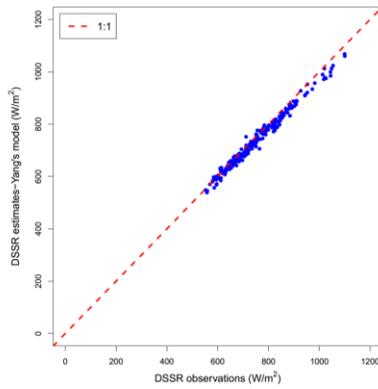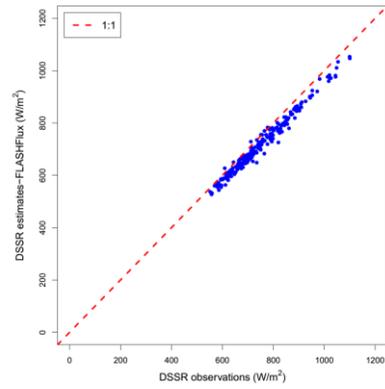
(a) Clear percent > 90%   (b) Clear percent > 90%
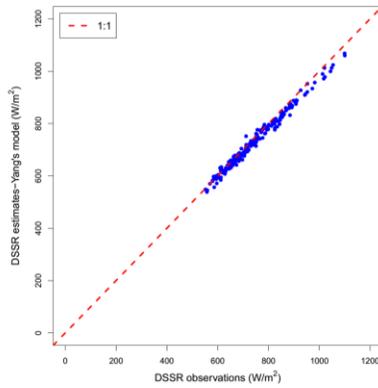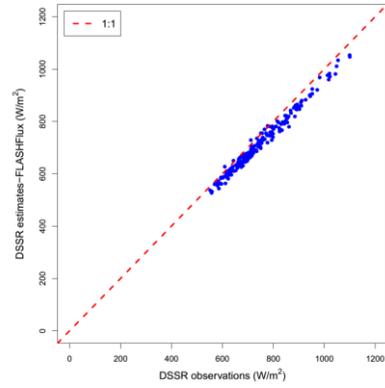
(c) Clear percent > 95%   (d) Clear percent > 95%
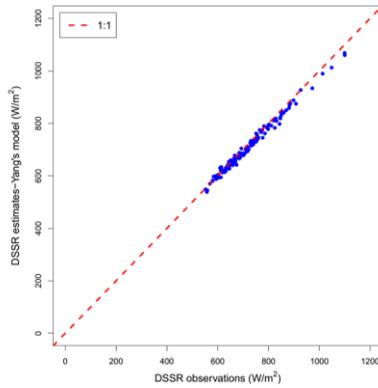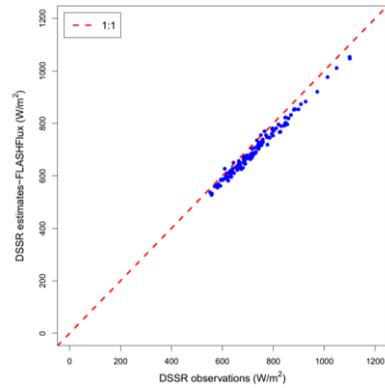
(e) Clear percent = 100%   (f) Clear percent = 100%

Figure 27 The comparisons between DSSR derived from Yang's model and FLASHFlux with three clear percent thresholds. (a), (c), and (e) are scatter plots of ground measurements versus DSSR estimates using Yang's model; (b), (d), and (f) are scatter plots of ground measurements versus DSSR estimates of FLASHFlux

67

4.2.2 DSSR estimates using I3RC model

After computing the clear-sky transmittance, the cloud transmittance was calculated by I3RC model. Then, the DSSR estimates for all-sky conditions can be computed based on Equation (3.2). In this section, the effects of parallax correction, DSSR variability, and the distance to the closest cloud on DSSR estimates would be discussed at first. Then, the DSSR estimates for all-sky conditions would be compared with the ground observations and FLASHFlux data.

4.2.2.1 Effects of parallax correction

The cloud structure decides the cloud transmittance using I3RC model. The parallax correction change the cloud structure compared with the cloud structure derived from original MOD04 dataset directly. Moreover, the different methods to correct the parallax effect would also change the cloud structure. In this study, the parallax effect was corrected by two methods: one is correct the parallax effect pixel by pixel; the other one is that the horizontal displacement of a scene is computed by the median value of horizontal displacement pixel by pixel. The cloud transmittance was computed by three different cloud structures, including the case without correction, the case that was corrected pixel by pixel, and the case that was corrected by a median value of the horizontal displacements within the scene. The three cases were used to examine the effect of parallax correction in the following

sections.

4.2.2.2   The DSSR estimates for all-sky conditions

The scatter plots scatter plots of ground measurements versus DSSR estimates on all-sky days for the three cases are shown in Figure 28. The three subfigures in Figure 28 are very similar to each other. A lot of points are located near the 1:1 line range from 500 to 1100 W/m^2. The other points are scattered within a symmetrical area looks like spindle. The extent of y axis in Figure 28(c) is a little greater than the extents of y axes in Figure 28(a) and (b). Count, bias, RMSE, MAPE, and R^2 for DSSR estimates on all-sky days for the three cases are shown in Table 18.

The pixel by pixel case has smaller bias, RMSE, and MAPE and greater R^2 than the case without correction. Except the bias, the statistical indices of the median case are superior to the case without correction. Thus, the parallax correction can reduce the errors of DSSR estimates. When the pixel by pixel case is compared with the median value case, the former has lower bias and MAPE than latter and latter has lower RMSE and high R^2 than former. However, the absolute differences among the statistical indices of three cases are very small.

(a) Origin

(b) Median

(c) Pixel by pixel
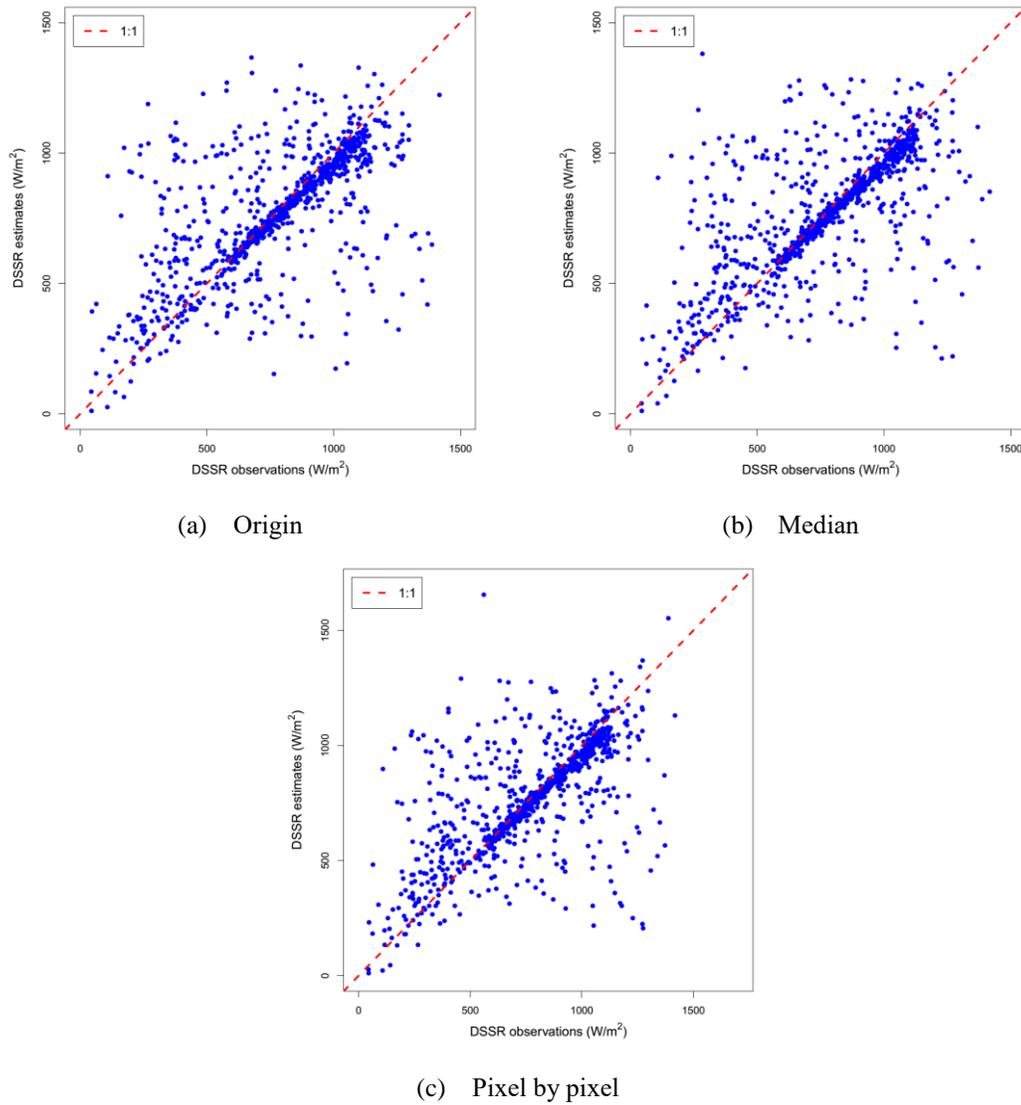
Figure 28 The scatter plots of ground measurements versus DSSR estimates for three different cloud structures. Parallax effect was not corrected in (a). Parallax effect was corrected by a median value of the horizontal displacements of the scene in (b). Parallax effect was corrected pixel by pixel in (c).

Table 18 Count, bias, RMSE, MAPE, and R^2 for DSSR estimates on all-sky days

Case	count	Bias (W/m^2)	RMSE (W/m^2)	MAPE (%)	R^2
Origin	934	1.992	220.733	25.885	0.407
Median	934	2.393	215.792	24.624	0.425
Pixel by pixel	934	1.608	218.051	24.384	0.420

4.2.2.3   The effects of DSSR variability

Because the spatial resolution of the DSSR estimates is 1 km, which is usually greater than a broken cloud, a small piece of cloud that can affect the DSSR around it might not be identified by MODIS products. The small cloud that passes the area around the station would result in high DSSR variability of the ground observations. In addition, when the cloud properties are not even distributed within an area, the ground observations could also have high DSSR variability in a short time period. Thus, the DSSR variability (Equation (3.1)) was used as an index, which indicates the stability of weather, to examine the performance of DSSR estimates. When the DSSR variability is small, the sky would be either clear or covered by evenly distributed clouds. If the DSSR variability is large, the sky would be covered by a few cloud or the properties of cloud vary large within an area.

The cumulative frequency function plot of DSSR variability is shown in Figure 29. The maximum DSSR variability is 37. The slope of the curve in Figure 29 reduce a lot when the DSSR variability is around 4. The scenes that DSSR variability is less than 4 account for 50% of available scenes.

The plot of DSSR variability versus bias for the three cases is show in Figure 30. The lines of three cases are very close to each other and their trends are similar. The three lines get their minimum value when the DSSR variability is from 1 to 2. With the increase of

DSSR variability from 2 to 10, the bias increase to 0. When the DSSR variability is greater

than 10, the lines of bias fluctuate around 0.
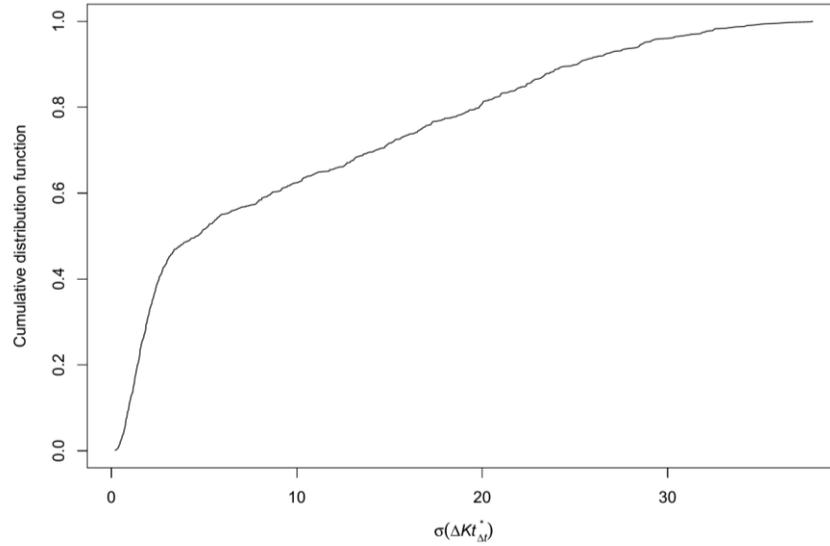
Figure 29 The cumulative frequency function of DSSR variability plot for all available scenes
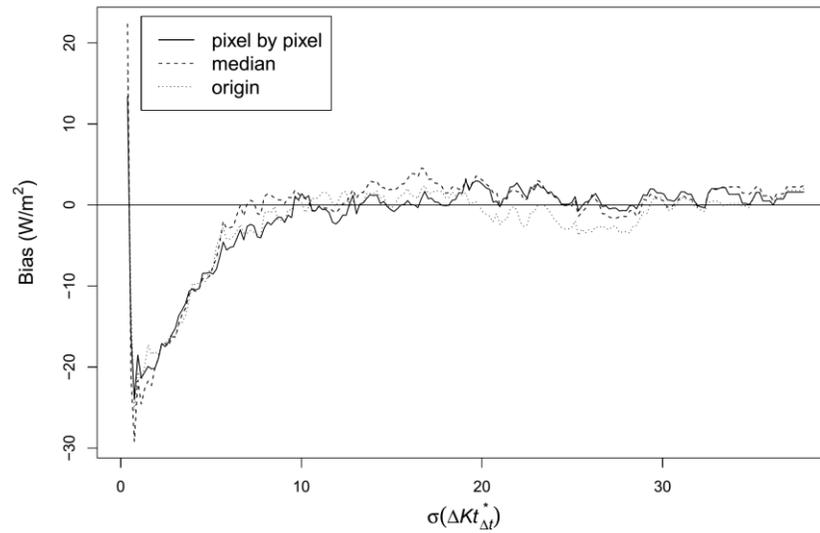
Figure 30 The plot of DSSR variability versus bias for the three cases to examine parallax effect.
The y axis value is the bias of scenes that DSSR variability is less than the x axis value.

The plot of DSSR variability versus RMSE for the three cases is show in Figure 31. The three lines are close to each other and have similar pattern. They get their minimum values from 1 to 2. Then, the RMSE of three cases increase as the rise of DSSR variability. The median line has lower RMSE than the other two lines for the range from 10 to 37.
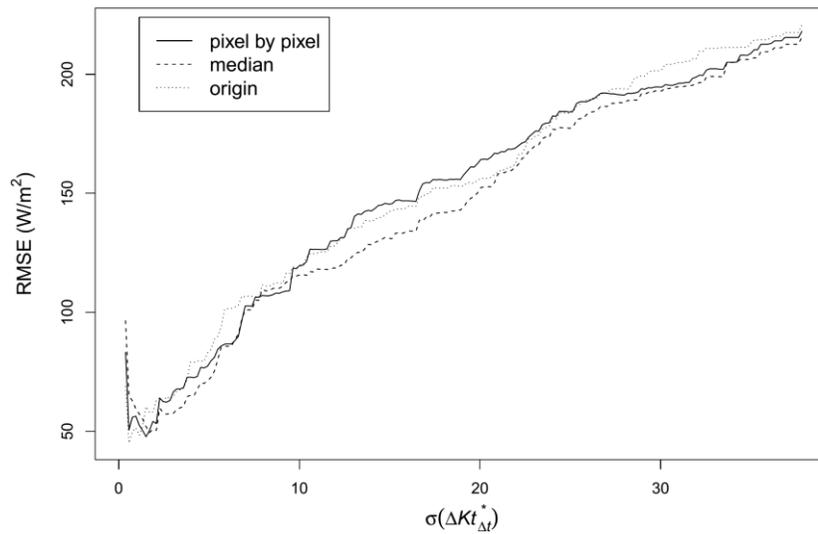
Figure 31 The plot of DSSR variability versus RMSE for the three cases to examine parallax effect. The y axis value is the RMSE of scenes that DSSR variability is less than the x axis value.

The plot of DSSR variability versus MAPE for the three cases is shown in Figure 32. The three lines are close to each other and have similar pattern. They get their minimum values from 2 to 3. Then, the MAPE of three cases increase as the rise of DSSR variability. The origin case has greater MAPE than the other cases for the range from 2 to 37.
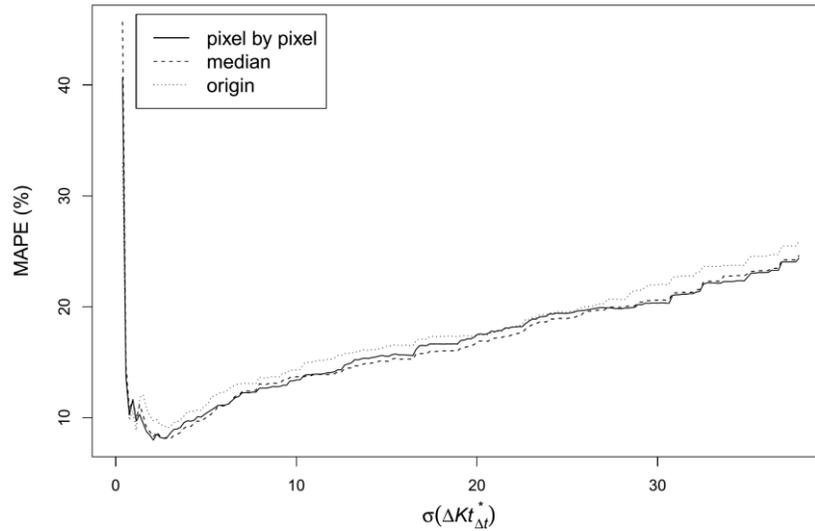
Figure 32 The plot of DSSR variability versus MAPE for the three cases to examine parallax effect. The y axis value is the MAPE of scenes that DSSR variability is less than the x axis value.

The plot of DSSR variability versus R^2 for the three cases is shown in Figure 33. The three lines are close to each other and have similar pattern. The R^2 of three cases decrease as the increase of DSSR variability. The median case has greater R^2 than the other cases for the range from 9 to 37.

When the DSSR variability is small (less than 3), the three cases have similar statistical indices. With the increase of DSSR variability, the case without parallax correction performs worse than the cases that apply parallax correction based on the figures of RMSE, MAPE, R^2. The case that was corrected by median value performs better than the case that was corrected pixel by pixel based on RMSE and R^2.
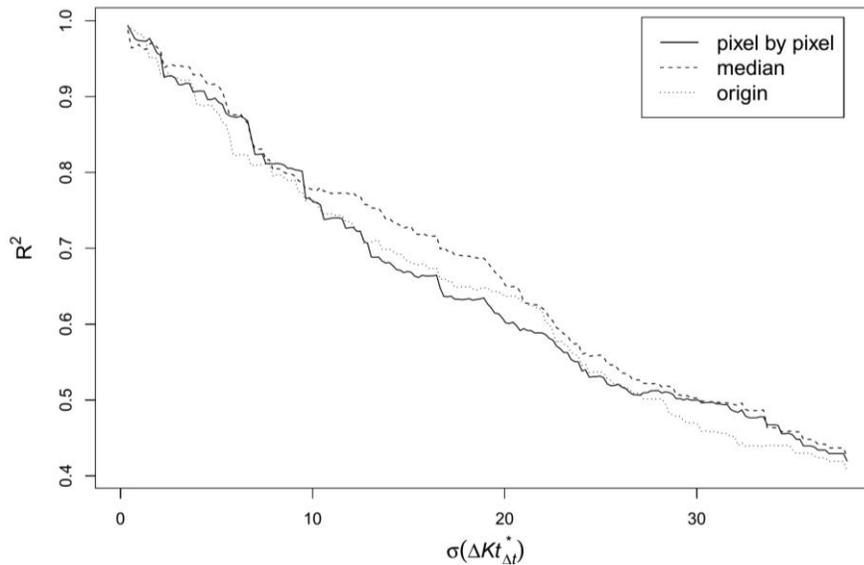
Figure 33 The plot of DSSR variability versus R^2 for the three cases to examine parallax effect. The y axis value is the R^2 of scenes that DSSR variability is less than the x axis value.

4.2.2.4　The effects of the distance between the station and the closest cloud

In fact, the DSSR variability cannot derive from satellite data. Thus, the closest distance from the station pixel to the cloud pixel can be an alternative index. The cumulative frequency function plot of the distance to the closest cloud (DCC) for case without parallax correction is shown in Figure 34. The maximum DCC is 40. The scenes that the DCC is zero account for 37%. And the scenes that the DCC is 40 account for 8%.

The plot of DCC versus bias for the three cases is show in Figure 35. The original line has higher bias than the other two line from 1 to 4. Then, the three lines are close to each others and have similar trends. The three lines get their minimum value when the DSSR variability is 6. When the DSSR variability is greater than 8, the three lines are overlapped

75

and the bias values are approaching -18 W/m^2.
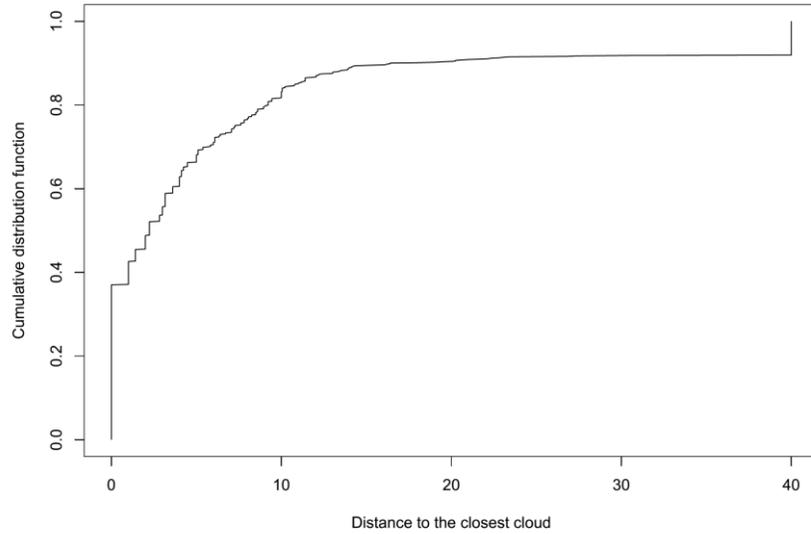
Figure 34 The cumulative frequency function plot of the closest distance to cloud for case without parallax correction
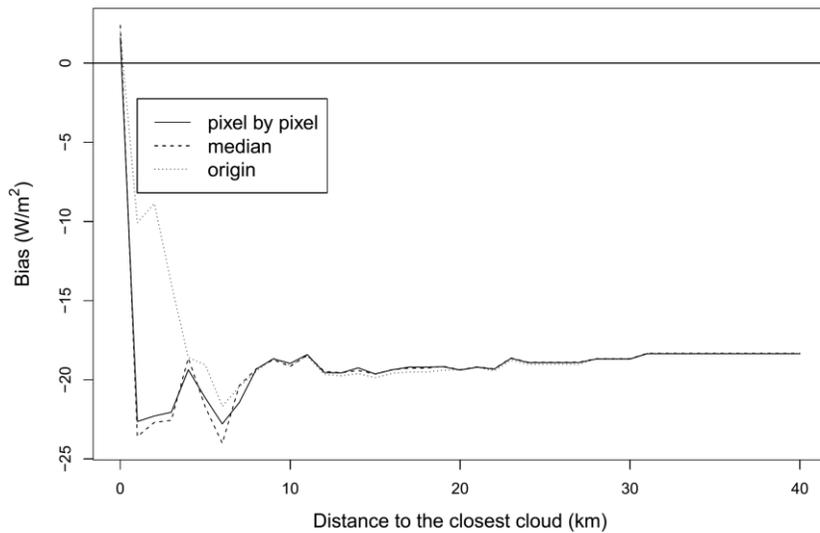
Figure 35 The plot of DCC versus bias for the three cases to examine parallax effect. The y axis value is the bias of scenes that DCC is greater than the x axis value.

The plot of DCC versus RMSE for the three cases is show in Figure 36. The three lines are overlapped and RMSE of three cases vary little when the DCC is greater than 8. The pixel by pixel line has lower RMSE that the other two lines from 3 to 6.
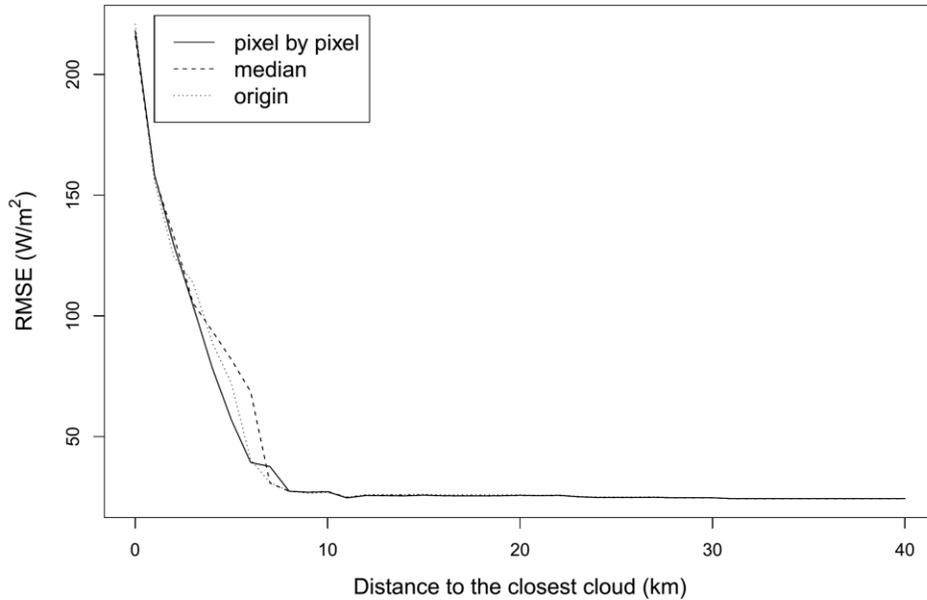

Figure 36 The plot of DCC versus RMSE for the three cases to examine parallax effect. The y axis value is the bias of scenes that DCC is greater than the x axis value.

The plot of DCC versus MAPE for the three cases is show in Figure 37. The three lines are overlapped and MAPE of three cases does not change a lot when the DCC is greater than 7. The pixel by pixel line has lower MAPE that the other two lines from 1 to 7.

The plot of DCC versus MAPE for the three cases is show in Figure 38. The three lines are overlapped and MAPE of three cases does not change a lot when the DCC is greater than 8. The pixel by pixel line has higher R^2 that the other two lines from 2 to 7.
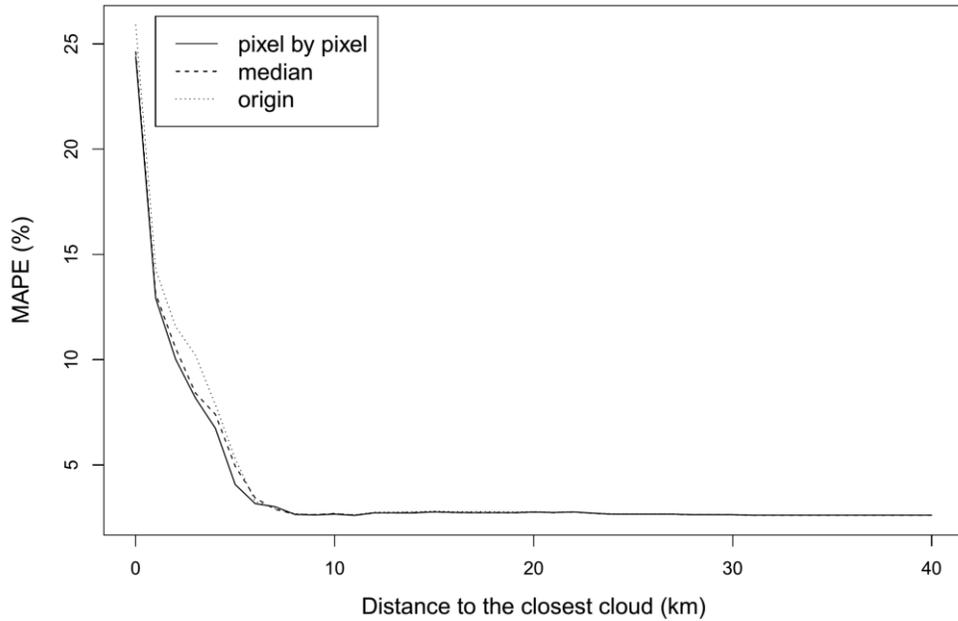
Figure 37 The plot of DCC versus MAPE for the three cases to examine parallax effect. The y axis value is the bias of scenes that DCC is greater than the x axis value.
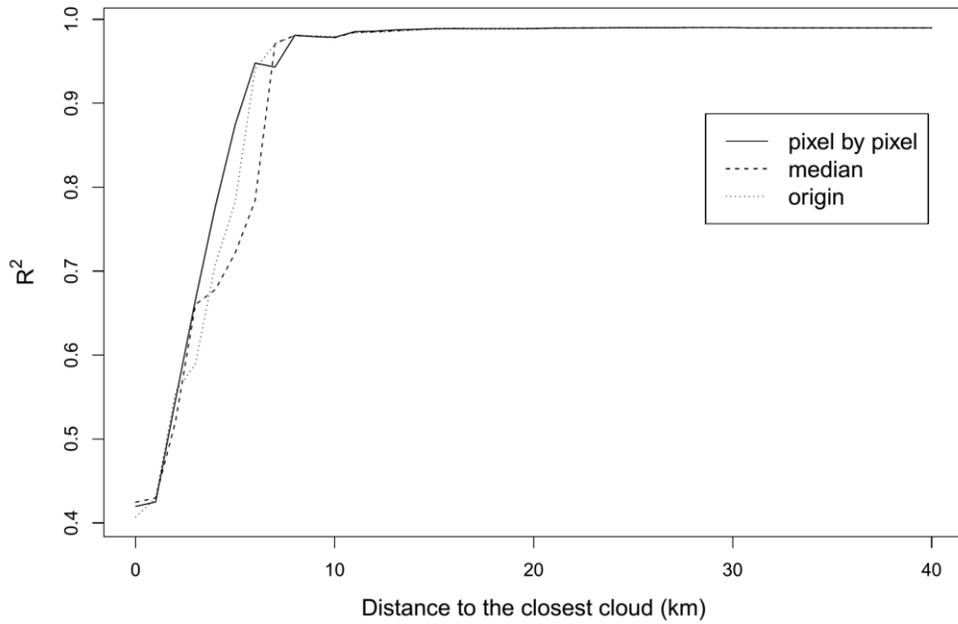
Figure 38 The plot of DCC versus R^2 for the three cases to examine parallax effect. The y axis value is the bias of scenes that DCC is greater than the x axis value.

When the DCC is large (greater than 8), the three cases have similar statistical indices. The case that was corrected pixel by pixel performs better than the other two cases when the DCC is less than 7 based on the plots of RMSE, MAPE, and R^2.

4.2.2.5 Comparisons with FLASHFlux data

The DSSR estimates that parallax effect was corrected by median case were compared with the DSSR estimates of FLASHFlux. The scenes that have both DSSR estimates of median case and FLASHFlux were selected to make comparisons. The scatter plots are shown in Figure 39.

(a) Median case   (b) FLASHFlux
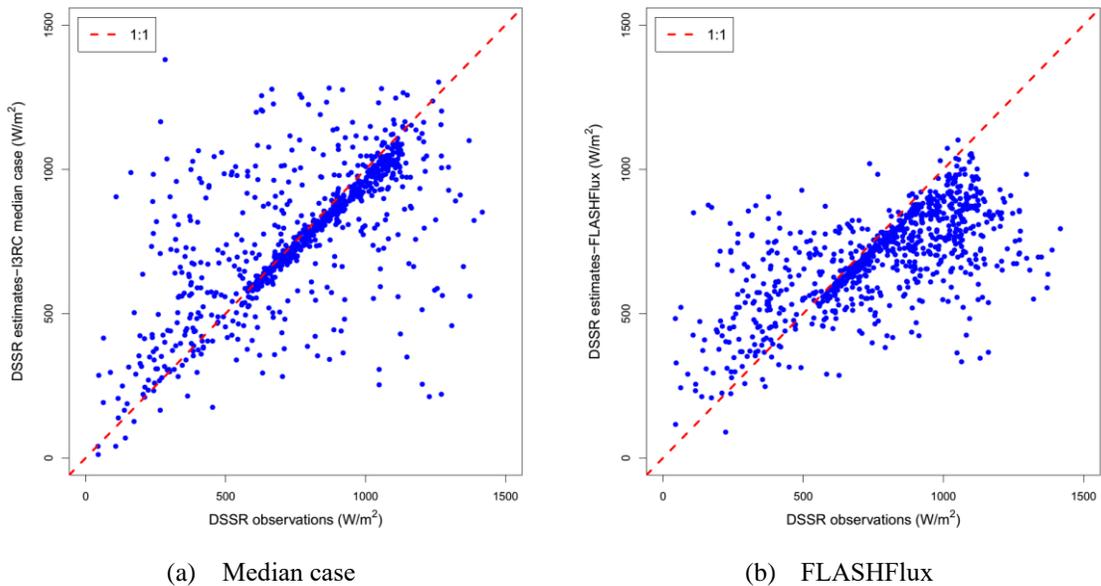
Figure 39 Scatter plots of ground measurements versus DSSR estimates of median case and FLASHFlux for all-sky conditions

The distribution of the points in the scatter plot of median case is symmetric based on

1:1 line. Some points gather at the area near the 1:1 line, and the other points are dispersedly distributed. The scatter plot of FLASHFlux is not symmetric and the points are more concentrated than median case's. The range of DSSR estimates of FLASHFlux is less than median case's.

The statistical indices are listed in Table 19. The DSSR estimates of median case have much better bias and MAPE results than FLASHFlux's and the RMSE and R^2 of former are also slightly better than latter. If the DSSR variability was considered, the performance of median case's DSSR estimates would be superior to FLASHflux's DSSR estimates (Figure 40). Thus, the DSSR estimates using median case have fewer errors than the DSSR estimates from FLASHFlux.

Table 19 Count, bias, RMSE, MAPE, and R^2 for DSSR estimates of Median case and FLASHFlux for all-sky conditions

Data	count	Bias (W/m^2)	RMSE (W/m^2)	MAPE (%)	R^2
Median case	889	3.026	215.017	24.588	0.431
FLASHFlux	889	-75.304	222.620	29.411	0.421
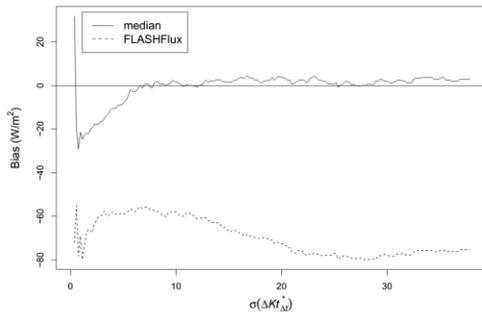
(a) Bias
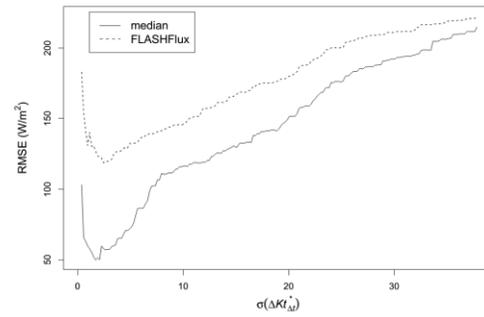
(b) RMSE
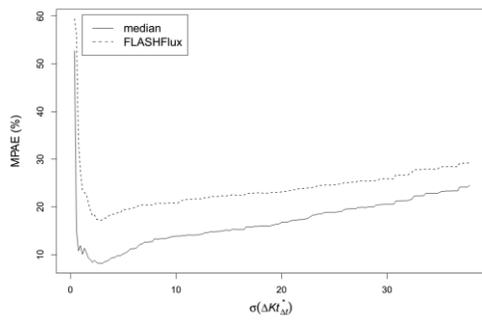
(c) MAPE
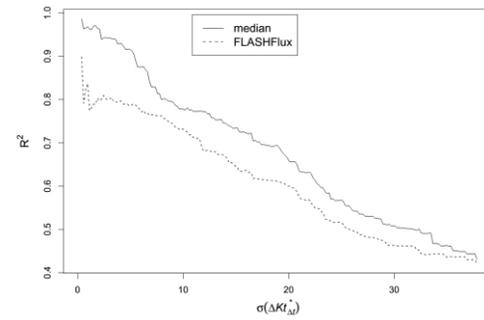
(d) R^2

Figure 40 The plot of DSSR variability versus four statistical indices for DSSR estimates of median case and FLASHFlux. The y axis value is the statistical indices of scenes that DSSR variability is less than the x axis value.

Chapter 5   Discussions

5.1 The effects of aerosol loading

Based on the results in section 4.2.1, using the valid β derived from MOD04 products would lead to larger errors than using a small value when DSSR estimates were calculated for clear-sky conditions. Even if the area around the station is cloud free, the DSSR estimates using the substituted β still have better performance than using original β, especially for the scenes that the original β are greater than substituted β.

In fact, due to the climate and location of the station, it is rarely to have very high optical aerosol thickness around the station. First, the height of aerosols is usually lower than 3 km (Hart et al. 2005) and the altitude of the station is 3680 m, so that it is rarely to exist thick aerosol layer at such a high altitude. The heavy particulate air pollutants from other areas, such as Indian and Mainland China, cannot reach Lhasa. Second, the aerosol sources around the station are very limited. Lhasa had a few factories that can emit huge aerosols in 2012 and 2013. The artificial aerosols can be particulate air pollutants that come from vehicle exhaust gas and firing biomass and coal, especially in the winter and spring. The natural aerosols can come from sandstorm in the winter and spring. If there were some days that have heavy aerosols, these days were more likely to appear in the winter and spring because of the rainfall in summer and fewer aerosol sources than winter and spring. This deduction

is similar to the result of Xia et al. (2011), that AOD at a Tibet station (30.773 °N, 90.962 °E, 4730 m AMSL) had an evident seasonal change, higher in spring and lower in other seasons. However, Figure 26 shows that the number of summer days with high β is much greater than the number of winter or spring days with high β. Thus, the β might be overestimated by the MODIS products at Lhasa. This inference means that aerosol loading is overestimated, so that the DSSR estimates using original β perform worse than using substituted β.

In addition, it is a common thing that satellite radiation datasets are very likely to have lower accuracy on the Tibetan Plateau (TP). Huang et al. (2013) indicated that Global Land Surface Satellite-downward surface shortwave radiation (GLASS-DSSR) products' accuracy at the stations in TP were evidently worse than other stations in mainland China. The similar results were also reported by many previous studies (Yang et al. 2006, Yang et al. 2008, Gui et al. 2010). But these study did not explain why the phenomenon happened.

Yang et al. (2012) suggested that the decline of solar radiation, which is particular outstanding in summer, on the TP over recent three decades is caused by the increase in water vapor amount and deep cloud cover rather than aerosol loading on the TP. The change and value of AOD over last three decades were small, so that the effects of AOD on DSSR could be neglect. The explanation can be applicable for the estimation of instantaneous solar radiation. The aerosol loading's effects on instantaneous DSSR estimates are neglect. If the

satellite derived AOD was overestimated, it might result in declines of DSSR estimates and introduce big errors. That might be the reason that satellite radiation datasets always have lower accuracy on the TP.

There are two reasons that can lead to the overestimation of AOD. The first one is that the coarse spatial resolution of AOD might exaggerate the AOD at subpixel level. The algorithm of MOD04 can only calculate valid AODs in a fraction of the pixels within the 10x10 1-km window. The AOD of each pixel In the MOD04 is the mean value of the valid AOD within each window. If there were some pixels within the window that valid AODs are large, the mean value would be much greater than the actual AOD at subpixel level. Second, the weather of TP in the summer is often cloudy and the cloud adjacency effect can increase the satellite derived AOD in the vicinity of clouds. Cloud adjacency effect refers to an enhancement in the illumination of the cloud-free column through the reflection of sunlight by nearby clouds (Marshak et al. 2008). Wen et al. (2007) reported bias error in one-dimensional (1D) aerosol optical thickness retrieval ranges from 50 to 140% depending on wavelength and the optical depth of nearby clouds, as well as aerosol optical thickness, so that the cloud adjacency effect should not be ignored. Redemann et al. (2009) indicated that the MODIS-derived AOD increases near cloud edges were much greater than the increases observed by airborne sun photometer. Beside the actual increase comes from the

humidification and consequent growth of aerosol particles in the moist cloud environment, the enhancement in the cloud-free column radiance comes from enhanced Rayleigh scattering can lead to the overestimated increase of AOD.

5.2 The advantages and disadvantages of applying I3RC model

The advantages of I3RC model are obvious. Applying 3D radiative transfer theory can compute the DSSR estimates with higher accuracy than 1D radiative transfer theory for all-sky conditions. When the variability of weather is lower (less than 4%), DSSR estimates of I3RC model has relatively low errors. However, the superiority of I3RC model is reduced with the increase of variability of weather.

The disadvantages of I3RC model are also obvious. It is very time consuming to apply I3RC model to calculate cloud transmittance as referred by many previous studies (Cahalan et al. 2005, Pincus and Evans 2009, Chen et al. 2012). In this study, the CPU of the computer to run I3RC model is Intel Core i7-4700HQ @2.40GHz, which has 4 cores and 8 threads and 6 threads were allocated to run I3RC model. Due to the values of the input parameters for Monte Carlo simulation (Table 7), the consuming time of each scene was at 10^3 seconds level per thread, so that the RMSE in normalized flux down is less than 2 percent (Pincus and Evans 2009). The theoretical RMSE is very small. If the weather was stable and the cloud structure was close to the real condition, the accuracy of DSSR estimates using I3RC

would be high and obviously outperform the DSSR dataset from FLASHFlux. The time consuming would be worth. However, if the cloud structure was not accurate when the weather is not stable, there would be no noticeable difference between the overall errors of the DSSR estimates using I3RC and FLASHFlux data. Using I3RC model would spend much more time than using traditional 1D radiative transfer model and get the same bad result as latter. The accurate cloud structure is the premise of the result using I3RC, otherwise it is unnecessary to apply I3RC model for radiative transfer model.

In addition, the dimensions of the input cloud structure in this study is around 51×51×50 pixels. The horizontal dimension is only 51×51, which is much smaller than the dimensions of a MOD03 product, which is 1354×2030. To get the DSSR-+ dataset of a MODIS scene with the same accuracy would spend much more time used in this study. Thus, it is not very practical to apply I3RC model to generate operational remote sensing products with large dimension at real time. However, if the dimension of the scene can be reduced to a small scope and the reduction would not heavily affect the result that people need, the time consuming feature of I3RC is acceptable because it can generate more accurate result than 1D theory in some cases.

5.3 The possible causes of the DSSR estimates with big errors

In the scatterplot of ground observations and DSSR estimates, lots of points are far away

from the 1:1 line. These points that absolute percentage error is greater than 10% are regarded as having big errors. The points with big errors accounts for 38% of all points in the median case. Thus, these errors should be systematic and the reason of these errors should be also carefully examined.

5.3.1   Cloud structure model

The horizontal location of cloud after parallax correction is not accurate enough, so that accurate 3D cloud structure cannot be rebuilt based on MODIS products. Theoretically, the case that parallax was corrected pixel by pixel should perform the best among the three cases in section 4.2.2.1. However, there is no big difference of statistical indices among the three cases for all-sky conditions. When the DSSR variability was used to examine the performance of the three cases, the case that uses median value even outperforms the pixel by pixel case at large DSSR variability. Although the pixel by pixel case seems to be the best when the DCC was applied to examine the errors of the three cases when the DCC is greater than 1 km, it only indicates that DSSR estimates of pixel by pixel case have lower errors when the variability of weather is not large.

As mentioned in the section 2.1, the optical sensors with multiple viewing angles and radar sensors can rebuild accurate 3D cloud structure. The former applies stereo pair photography to build the 3D model. The latter can measure the direction and distance

between sensor and objects based on backscatter to build the 3D model. Apparently, the optical sensor with single viewing angle as MODIS can not rebuild based on above-mentioned two principles. Although the section 3.3.3.2 refers to a simple method to correct parallax effect based on cloud top height and viewing angle of sensor and an algorithm to fill the gaps caused by the correction, the gaps can not be filled at some cases as shown in Figure 41. If the gaps were close to the station, the DSSR estimates would be introduced great errors.

(a)   Original   (b)   After parallax correction and gap filling
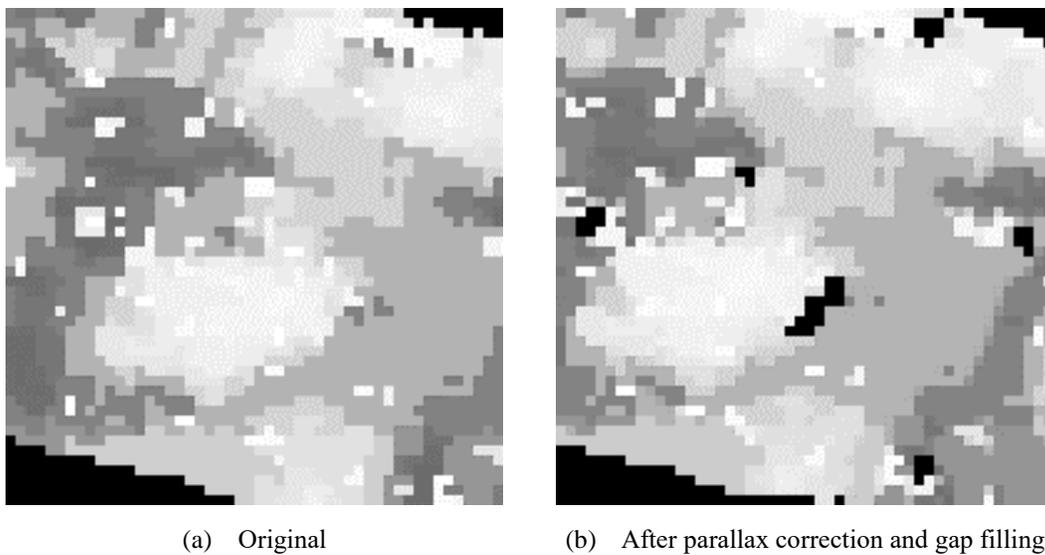
Figure 41 The failure of a scene after parallax correction and gap filling. The scene was captured by Terra on June 1st, 2013. The clouds in (a) have more than one layer without apparent gaps. However, after parallax correction and gap filling, many gaps emerge in (b).

The model of vertical cloud layer is over simplified. The number of cloud layer was assumed to be one without overlay in this study. In fact, the cloud can have two or even three layers and the clouds in lower layers can be overlaid by the upper clouds. If the cases were

not considered, sun light blocked by lower clouds would reach the surface (Figure 42) and DSSR would be overestimated.

Besides the spatial structure of cloud, that the properties of cloud were decided by assumptions can result in errors. Only two cloud types, stratocumulus and cirrus were taken into account and the two types were categorized by the cloud temperature. The other types of cloud were classified into the two types. In addition, the water content of cloud and effective radius were set to constant values based on the type of cloud. These assumptions make the modeled cloud properties to be very different from the real cloud properties. In fact, the properties of cloud have great influences on DSSR. For example, the cloud with smaller effective radius can reduce the DSSR more effectively the cloud with bigger effective radius, even the total water path is equal. The water content decides the cloud depth, which affect the radiation that pass through the cloud directly. The water content of different cloud types varies a lot. Even the same type cloud can also have different water contents. Thus, the inaccurate cloud properties of 3D cloud structure would introduce great errors of DSSR estimates.
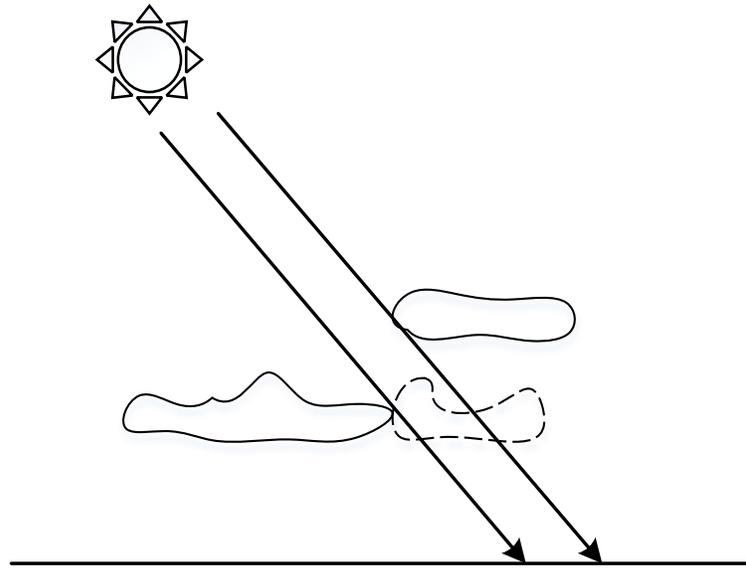
Figure 42 A conceptual diagram that DSSR might be overestimated by one layer cloud structure model. The cloud drawn by dash line is the cloud that cannot be rebuild by one layer model.

5.3.2 Spatial resolution and map projection

The subpixel cloud can have great effects on DSSR. The spatial resolution of MODIS cloud products is 1 km. A 1 km × 1 km area is too large to identify a small piece of cloud, so that the small subpixel cloud cannot be identified by coarser pixel. However, it can still affect the transmission of light, and it would result in overestimating the DSSR. Similarly, a small area of clear sky can also affect DSSR like subpixel cloud, but it would lead to underestimate the DSSR.

Map projection deformation would affect the relative geographic location between clouds and the station. It is possible that the deformation would result in small displacement

of clouds and affect the DSSR at the edge of clouds.

5.3.3 Data quality

The datasets from MODIS products to calculate clear-sky transmittance are precipitable water, ozone thickness, thickness of the ozone layer, and Ångström turbidity coefficient. The data quality of first three parameters have small effects on DSSR estimates in Lhasa and the errors that they introduce are negligible for clear-sky conditions. The missing data of them account for relatively small part of the whole dataset and can be imputed with reasonable value. However, the AOD retrievals from MOD04 have lots of missing data and the imputed value can be very different from the true value. In addition, the accuracy of data is lowered by coarse spatial resolution and cloudy weather on TP. Thus, the Ångström turbidity coefficient, which is derived from AOD of MOD04, cannot be very accurate.

Cloud top height, water path, and effective radius are three indispensable datasets from MOD06 to rebuild cloud structure. However, the quality of these data is not good. First, although the cloud top height data cover most cloud areas of each scene, lots of cloud areas do not have water path and effective radius values. To rebuild 3D cloud structure, the imputation is required, even if the imputed value is far from the real value. Second, the accuracy of these three datasets are not high. For, example, the cloud top height is decided by sea surface temperature, cloud top temperature, and environmental lapse rate (Menzel,

Frey, and Baum 2015). However, the cloud top temperature is not decided by height. If the weather was not stable, the cloud temperature could be very different from air temperature and lapse rate can change a lot as well. Thus, the absolute cloud top height differences between the of CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) and MODIS can be very large. The mean differences for single-layered cirrus is greater than 1.5 km. King et al. (1997) suggested that errors of ±25% can arise in the optical thickness for errors of ±50% in effective radius. Thus, the algorithms of these datasets cannot get very accurate retrievals for the three parameters.

5.3.4   Radiative transfer model

Yang's model and I3RC model both can introduce errors, but the errors could not be big. Gueymard (2003) pointed out that Yang's model can result in a MAPE of 1.1% in the direct normal irradiance compared with the Simple Model of the Atmospheric Radiative Transfer of Sunshine (SMARTS). The RMSE caused by I3RC can be controlled by computation time (Pincus and Evans 2009). Thus, there is no need to consider the errors, which result from radiative transfer model, before solving the above-mentioned problems.

Chapter 6   Conclusions

This study explored the accuracy of instantaneous DSSR on all-sky days estimated by a DSSR algorithm based on MODIS products (MOD03-07, 09). The DSSR algorithm consists of three parts: 1) TOA insolation model (Peixoto and Oort 1992); 2) Yang's model (Yang, Koike, and Ye 2006) was adopted to calculate clear transmittance; 3) I3RC Monte Carlo 3D radiative transfer community model (Cahalan et al. 2005) was adopted to calculate cloud transmittance. In addition, a model to build the 3D cloud structure, which is an input of I3RC model, in Tibet were proposed. After computing DSSR estimates, the effects of parallax correction, DSSR variability, and the distance to the closest cloud on DSSR estimates were discussed. Then, the DSSR estimates for all-sky conditions were compared with the ground observations and FLASHFlux data.

The main empirical findings are chapter specific and were included in the "Results" and "Discussions" section. The synthesized findings of this study are:

1) 3D radiative transfer model does generate DSSR with higher accuracy than the FLASHFlux dataset, which uses 1D radiative transfer model.

2) When the distance to the closest cloud is greater than 8 km, the cloud effects on DSSR are negligible.

Applying 3D radiative transfer model is a better choice to calculate DSSR than traditional 1D radiative transfer model. Although 3D radiative transfer model is time consuming, the computation time can be reduced by narrowing down the dimension of the input data. Based on the two findings, it can be inferred that the algorithm using 3D radiative transfer model may be suitable to estimate DSSR at a limited area, such as a PV system, to forecast DSSR with high accuracy in an operational environment.

However, the DSSR estimates of 38% of the scenes have big errors compared with ground observations. These big errors might result from inappropriate assumptions of the cloud structure model, coarse spatial resolution and map projection, and poor data quality of MODIS products. Moreover, only one station and the errors of ground observations also reduce the reliability of calibrations. In addition, MODIS on Terra/Aqua only revisit the same place for one or two times during the day time, the temporal resolution is not high enough for real time short-term forecasting.

For future research, there are four suggestions: 1) more ground stations for calibration to examine the practicality and reliability of the DSSR algorithm proposed in this study; 2) The satellite system with several geostationary satellites with high spatial and spectral resolution can build more accurate cloud structure in real time; 3) The algorithm to generate atmosphere products based on satellite data should be improved to enhance the data quality;

4) the historical cloud property data captured by radar sensor, such as Cloud Profiling Radar (CPR) on CloudSat, can be helpful to build more accurate cloud structure rather than assumptions (Miller et al. 2014).

Although a few of situations applies 3D radiative transfer model at present, the benefit of it to estimate DSSR with higher accuracy is obvious. It has the potential to improve the quality of PV power forecasting and to promote the development of the PV industry in the future.

Bibliography

Ahrens, C Donald. 2016. *Meteorology today: an introduction to weather, climate, and the environment*: Cengage Learning.

Angstrom, Anders. 1924. "Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation." *Quarterly Journal of the Royal Meteorological Society* 50 (210):121-126.

Bacher, Peder, Henrik Madsen, and Henrik Aalborg Nielsen. 2009. "Online short-term solar power forecasting." *Solar Energy* 83 (10):1772-1783.

Badescu, Viorel, Christian A Gueymard, Sorin Cheval, Cristian Oprea, Madalina Baciu, Alexandru Dumitrescu, Flavius Iacobescu, Ioan Milos, and Costel Rada. 2012. "Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models." *Renewable and Sustainable Energy Reviews* 16 (3):1636-1656.

Berk, Alexander, Lawrence S Bernstein, and David C Robertson. 1987. MODTRAN: A moderate resolution model for LOWTRAN. DTIC Document.

Beyer, Hans Georg, Claudio Costanzo, and Detlev Heinemann. 1996. "Modifications of the Heliosat procedure for irradiance estimates from satellite images." *Solar Energy* 56 (3):207-212.

Cahalan, Robert F, Lazaros Oreopoulos, Alexander Marshak, K Franklin Evans, Anthony B Davis, Robert Pincus, Ken H Yetzer, Bernhard Mayer, Roger Davies, and Thomas P Ackerman. 2005. "The I3RC-Bringing together the most advanced radiative transfer tools for cloudy atmospheres." *Bulletin of the American Meteorological Society* 86 (9):1275-1293.

Cano, Daniel, Jean-Marie Monget, Michel Albuisson, Hervé Guillard, Nathalie Regas, and Lucien Wald. 1986. "A method for the determination of the global solar radiation from meteorological satellite data." *Solar Energy* 37 (1):31-39.

Cebecauer, Tomáš, Marcel Šúri, and Richard Perez. 2010. "High performance MSG satellite model for operational solar energy applications." ASES National Solar Conference, Phoenix, USA.

Chandrasekhar, Subrahmanyan. 1960. *Radiative transfer*: Dover Publications.

Chaouachi, Aymen, Rashad M Kamel, R Ichikawa, H Hayashi, and K Nagasaka. 2009. "Neural network ensemble-based solar power generation short-term forecasting." *World Academy of Science, Engineering and Technology* 54:54-59.

Chen, Changsong, Shanxu Duan, Tao Cai, and Bangyin Liu. 2011. "Online 24-h solar power forecasting based on weather type classification using artificial neural network." *Solar Energy* 85 (11):2856-2870.

Chen, Ling, Guangjian Yan, Tianxing Wang, Huazhong Ren, Josep Calbó, Jing Zhao, and Richard McKenzie. 2012. "Estimation of surface shortwave radiation components under all sky conditions: Modeling and sensitivity analysis." *Remote Sensing of Environment* 123:457-469.

Chow, Chi Wai, Serge Belongie, and Jan Kleissl. 2015. "Cloud motion and stability estimation for intra-hour solar forecasting." *Solar Energy* 115:645-655.

Chow, Chi Wai, Bryan Urquhart, Matthew Lave, Anthony Dominguez, Jan Kleissl, Janet Shields, and Byron Washom. 2011. "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed." *Solar Energy* 85 (11):2881-2893.

Chowdhury, Badrul H, and Saifur Rahman. 1987. "Forecasting sub-hourly solar irradiance for prediction of photovoltaic output." 19th IEEE Photovoltaic Specialists Conference.

Chupong, Charnon, and Boonyang Plangklang. 2011. "Forecasting power output of PV grid connected system in Thailand without using solar radiation measurement." *Energy Procedia* 9:230-237.

Diagne, Maimouna, Mathieu David, John Boland, Nicolas Schmutz, and Philippe Lauret. 2014. "Post-processing of solar irradiance forecasts from WRF model at Reunion Island." *Solar Energy* 105:99-108.

Diagne, Maimouna, Mathieu David, Philippe Lauret, John Boland, and Nicolas Schmutz. 2013. "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids." *Renewable and Sustainable Energy Reviews* 27:65-76.

Diner, David J, and John V Martonchik. 1984. "Atmospheric transfer of radiation above an inhomogeneous non-Lambertian reflective ground—II. Computational considerations and results." *Journal of Quantitative Spectroscopy and Radiative Transfer* 32 (4):279-304.

Evans, Annette, Vladimir Strezov, and Tim J. Evans. 2012. "Assessment of utility energy

storage options for increased renewable energy penetration." *Renewable and Sustainable Energy Reviews* 16 (6):4141-4147. doi: 10.1016/j.rser.2012.03.048.

Evans, K Franklin. 1998. "The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer." *Journal of the Atmospheric Sciences* 55 (3):429-446.

Geier, EB, R Green, DP Kratz, P Minnis, WF Miller, SK Nolan, and CB Franklin. 2003. "CERES data management system: Single Satellite Footprint TOA/surface fluxes and clouds (SSF) collection document, release 2, version 1, 212 pp. and appendixes, NASA Langley Res. Cent., Hampton, Va." *Cent., Hampton, Va, Radiat. and Aerosol Branch, Atmos. Sci. Res., NASA Langley Res. Cent., Hampton, Va, available at:* *http://asd-www.larc.nasa.gov/ceres/collect_guide/SSF_CG.pdf*.

Gopinathan, KK. 1988. "A general formula for computing the coefficients of the correlation connecting global solar radiation to sunshine duration." *Solar energy* 41 (6):499-502.

Gueymard, Christian A. 2003. "Direct solar transmittance and irradiance predictions with broadband models. Part I: detailed theoretical performance assessment." *Solar Energy* 74 (5):355-379.

Gueymard, Christian A. 2004. "The sun's total and spectral irradiance for solar energy applications and solar radiation models." *Solar energy* 76 (4):423-453.

Gui, Sheng, Shunlin Liang, Kaicun Wang, Lin Li, and Xiaotong Zhang. 2010. "Assessment of three satellite-estimated land surface downwelling shortwave irradiance data sets." *Geoscience and Remote Sensing Letters, IEEE* 7 (4):776-780.

Gupta, Shashi K, David P Kratz, Paul W Stackhouse Jr, and Anne C Wilber. 2001. *The Langley parameterized shortwave algorithm (LPSA) for surface radiation budget studies*.

Hardin, Garrett. 1968. "The tragedy of the commons." *science* 162 (3859):1243-1248.

Hart, William D, James D Spinhirne, Steven P Palm, and Dennis L Hlavka. 2005. "Height distribution between cloud and aerosol layers from the GLAS spaceborne lidar in the Indian Ocean region." *Geophysical research letters* 32 (22).

Hess, M, P Koepke, and I Schult. 1998. "Optical properties of aerosols and clouds: The software package OPAC." *Bulletin of the American meteorological society* 79 (5):831-844.

Horváth, Ákos, Chellappan Seethala, and Hartwig Deneke. 2014. "View angle dependence of MODIS liquid water path retrievals in warm oceanic clouds." *Journal of Geophysical Research: Atmospheres* 119 (13):8304-8328.

Huang, Guanghui, Shaomin Liu, and Shunlin Liang. 2012. "Estimation of net surface shortwave radiation from MODIS data." *International journal of remote sensing* 33 (3):804-825.

Huang, Guanghui, Weizhen Wang, Xiaotong Zhang, Shunlin Liang, Shaomin Liu, Tianbao Zhao, Jinming Feng, and Zhuguo Ma. 2013. "Preliminary validation of GLASS-DSSR products using surface measurements collected in arid and semi-arid regions of China." *International Journal of Digital Earth* 6 (sup1):50-68. doi: 10.1080/17538947.2013.825655.

International Energy Agency. 2016. 2015 SNAPSHOT OF GLOBAL PHOTOVOLTAIC MARKETS. In *Photovoltaic power systems programme*.

Iqbal, M. 1983. *An introduction to solar radiation*.

Ji, Wu, and Keong Chan Chee. 2011. "Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN." *Solar Energy* 85 (5):808-817.

King, Michael D, Si-Chee Tsay, Steven E Platnick, Menghua Wang, and Kuo-Nan Liou. 1997. "Cloud retrieval algorithms for MODIS: Optical thickness, effective particle radius, and thermodynamic phase." *MODIS Algorithm Theoretical Basis Document* 1997.

Kleissl, Jan. 2013. *Solar energy forecasting and resource assessment*: Academic Press.

Kostylev, V, and A Pavlovski. 2011. "Solar power forecasting performance–towards industry standards." 1st International Workshop on the Integration of Solar Power into Power Systems Aarhus, Denmark.

Lábó, Eszter, Judit Kerényi, and Mária Putsay. 2007. "The parallax correction of MSG images on the basis of the SAFNWC cloud top height product." Proceedings EUMETSAT Meteorological Satellite Conf. and 15th Satellite Meteorology and Oceanography Conf. Amer. Meteor. Soc., Amsterdam, the Netherlands.

Levy, RC, S Mattoo, LA Munchak, LA Remer, AM Sayer, and NC Hsu. 2013. "The Collection 6 MODIS aerosol products over land and ocean." *Atmos. Meas. Tech* 6 (11):2989-3034.

Li, Yanting, Yan Su, and Lianjie Shu. 2014. "An ARMAX model for forecasting the power output of a grid connected photovoltaic system." *Renewable Energy* 66:78-89.

Lin, Jeffrey, and P.W. Singer. 2016. "GAOFEN 4, THE WORLD'S MOST POWERFUL GEO SPY SATELLITE, CONTINUES CHINA'S GREAT LEAP FORWARD INTO SPACE." Last Modified 01/08/2016 Accessed 02/08. http://www.popsci.com/gaofen-4-worlds-most-powerful-geo-spy-satellite-continues-chinas-great-leap-forward-into-space.

Liou, KN, Y Gu, Q Yue, and G McFarguhar. 2008. "On the correlation between ice water content and ice crystal size and its application to radiative transfer and general circulation models." *Geophysical Research Letters* 35 (13).

Liu, Liping, Jiafeng Zheng, Zheng Ruan, Zhehu Cui, Zhiqun Hu, Songhua Wu, Guangyao Dai, and Yahao Wu. 2015. "Comprehensive radar observations of clouds and precipitation over the Tibetan Plateau and preliminary analysis of cloud properties." *Journal of Meteorological Research* 29:546-561.

Lorenz, Elke, Annette Hammer, and Detlev Heinemann. 2004. "Short term forecasting of solar radiation based on satellite data." EUROSUN2004 (ISES Europe Solar Congress).

Lyapustin, A. I., and Y. J. Kaufman. 2001. "Role of adjacency effect in the remote sensing of aerosol." *Journal of Geophysical Research: Atmospheres* 106 (D11):11909-11916. doi: 10.1029/2000jd900647.

Madkour, MA, M El-Metwally, and AB Hamed. 2006. "Comparative study on different models for estimation of direct normal irradiance (DNI) over Egypt atmosphere." *Renewable Energy* 31 (3):361-382.

Marshak, Alexander, Guoyong Wen, James A Coakley, Lorraine A Remer, Norman G Loeb, and Robert F Cahalan. 2008. "A simple model for the cloud adjacency effect and the apparent bluing of aerosols near clouds." *Journal of Geophysical Research: Atmospheres (1984–2012)* 113 (D14).

Meerkötter, R, and L Bugliaro. 2009. "Diurnal evolution of cloud base heights in convective cloud fields from MSG/SEVIRI data." *Atmospheric Chemistry and Physics* 9 (5):1767-1778.

Menzel, W Paul, Richard A Frey, and Bryan A Baum. 2015. *Cloud top properties and cloud phase algorithm theoretical basis document*: University of Wisconsin--Madison.

Miller, Steven D, John M Forsythe, Philip T Partain, John M Haynes, Richard L Bankert, Manajit Sengupta, Cristian Mitrescu, Jeffrey D Hawkins, and Thomas H Vonder Haar. 2014. "Estimating three-dimensional cloud structure via statistically blended satellite observations." *Journal of Applied Meteorology and Climatology* 53 (2):437-455.

NASA. 2011. Joint Polar Satellite System (JPSS) VIIRS Cloud Base Height Algorithm Theoretical Basis Document (ATBD). Greenbelt, Maryland: Goddard Space Flight Center.

NASA. 2016a. "Algorithms." Accessed 02/12. http://modis.gsfc.nasa.gov/data/atbd/.

NASA. 2016b. "The introduction of MODIS aerosol product." Accessed 02/14. http://modis-atmos.gsfc.nasa.gov/MOD04_L2/index.html.

NASA. 2016c. "The introduction of MODIS atmospheric profile product." Accessed 02/14. http://modis-atmos.gsfc.nasa.gov/MOD07_L2/index.html.

NASA. 2016d. "The introduction of MODIS cloud product." Accessed 02/14. http://modis-atmos.gsfc.nasa.gov/MOD06_L2/index.html.

NASA. 2016e. "The introduction of MODIS water vapor product." Accessed 02/14. http://modis-atmos.gsfc.nasa.gov/MOD05_L2/index.html.

NASA. 2016f. "LEVEL 1 AND ATMOSPHERE ARCHIVE AND DISTRIBUTION SYSTEM (LAADS)." Accessed 02/12. https://earthdata.nasa.gov/about/daacs/daac-laads.

NASA. 2016g. "MODIS Component." Accessed 02/11. http://modis.gsfc.nasa.gov/about/components.php.

NASA. 2016h. "MODIS Design." Accessed 02/11. http://modis.gsfc.nasa.gov/about/design.php.

NOAA. 2016. "The High-Resolution Rapid Refresh (HRRR)." Last Modified 01/17/2016 Accessed 02/08. http://ruc.noaa.gov/hrrr/.

Northwest Modeling Consortium. 2016. "Pacific Northwest Environmental Forecasts and Observations ", Last Modified 02/07/2016 Accessed 02/07. http://www.atmos.washington.edu/mm5rt/.

Paoli, Christophe, Cyril Voyant, Marc Muselli, and Marie-Laure Nivet. 2010. "Forecasting of preprocessed daily solar radiation time series using neural networks." *Solar Energy* 84 (12):2146-2160.

Paulescu, Marius, Eugenia Paulescu, Paul Gravila, and Viorel Badescu. 2012. *Weather modeling and forecasting of PV systems operation*: Springer Science & Business Media.

Paulescu, Marius, Eugenia Paulescu, Paul Gravila, and Viorel Badescu. 2013. "Outdoor Operation of PV Systems." In *Weather Modeling and Forecasting of PV Systems Operation*, 271-324. Springer.

Peixoto, José P, and Abraham H Oort. 1992. *Physics of climate*.

Pelland, Sophie, Jan Remund, Jan Kleissl, Takashi Oozeki, and Karel De Brabandere. 2013. "Photovoltaic and solar forecasting: state of the art." *IEA PVPS, Task* 14.

Perez, Richard, Mark Beauharnois, K Hemker, Sergey Kivalov, Elke Lorenz, Sophie Pelland, Jim Schlemmer, and Glenn Van Knowe. 2011. "Evaluation of numerical weather prediction solar irradiance forecasts in the US." ASES Annual Conference. Raleigh, NC, USA.

Perez, Richard, Pierre Ineichen, Kathy Moore, Marek Kmiecik, Cyril Chain, Ray George, and Frank Vignola. 2002. "A new operational model for satellite-derived irradiances: description and validation." *Solar Energy* 73 (5):307-317. doi: http://dx.doi.org/10.1016/S0038-092X(02)00122-6.

Pincus, Robert, and K Franklin Evans. 2009. "Computational cost and accuracy in calculating three-dimensional radiative transfer: Results for new implementations of Monte Carlo and SHDOM." *Journal of the Atmospheric Sciences* 66 (10):3131-3146.

Redemann, Jens, Qin Zhang, Philip B Russell, John M Livingston, and Lorraine A Remer. 2009. "Case studies of aerosol remote sensing in the vicinity of clouds." *Journal of Geophysical Research: Atmospheres* 114 (D6).

Sharma, S, R Vaishnav, MV Shukla, P Kumar, PK Thapliyal, S Lal, and YB Acharya. 2015. "Evaluation of cloud base height measurements from ceilometer CL31 and MODIS satellite over Ahmedabad, India." *Evaluation* 8:11729-11752.

Spencer, JW. 1971. "Fourier series representation of the position of the sun." *Search* 2 (5):172-172.

Stamnes, Knut, S-Chee Tsay, Warren Wiscombe, and Kolf Jayaweera. 1988. "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media." *Applied optics* 27 (12):2502-2509.

Stephens, Graeme L, Steven Ackerman, and Eric A Smith. 1984. "A shortwave parameterization revised to improve cloud absorption." *Journal of the Atmospheric Sciences* 41 (4):687-690.

Sun, Zhian, and Aixia Liu. 2013. "Fast scheme for estimation of instantaneous direct solar irradiance at the Earth's surface." *Solar Energy* 98:125-137.

Tomson, Teolan. 2010. "Fast dynamic processes of solar radiation." *Solar Energy* 84 (2):318-323.

U.S. Energy Information Administration. 2013. International Energy Outlook. US Energy Information Administration, Washington, DC.

Wen, Guoyong, Alexander Marshak, Robert F Cahalan, Lorraine A Remer, and Richard G Kleidman. 2007. "3‐D aerosol‐cloud radiative interaction observed in collocated MODIS and ASTER images of cumulus cloud fields." *Journal of Geophysical Research: Atmospheres* 112 (D13).

Widén, Joakim, Nicole Carpman, Valeria Castellucci, David Lingfors, Jon Olauson, Flore Remouit, Mikael Bergkvist, Mårten Grabbe, and Rafael Waters. 2015. "Variability assessment and forecasting of renewables: A review for solar, wind, wave and tidal resources." *Renewable and Sustainable Energy Reviews* 44:356-375.

Wyser, Klaus, William O'Hirok, Catherine Gautier, and Charles Jones. 2002. "Remote sensing of surface solar irradiance with corrections for 3-D cloud effects." *Remote Sensing of Environment* 80 (2):272-284.

Xia, Xiangao, Xumei Zong, Zhiyuan Cong, Hongbin Chen, Shichang Kang, and Pucai Wang. 2011. "Baseline continental aerosol over the central Tibetan plateau and a case study of aerosol transport from South Asia." *Atmospheric environment* 45 (39):7370-7378.

Yang, Kun, Baohong Ding, Jun Qin, Wenjun Tang, Ning Lu, and Changgui Lin. 2012. "Can aerosol loading explain the solar dimming over the Tibetan Plateau?" *Geophysical Research Letters* 39 (20).

Yang, Kun, Toshio Koike, Paul Stackhouse, Colleen Mikovitz, and Stephen J Cox. 2006. "An assessment of satellite surface radiation products for highlands with Tibet instrumental data." *Geophysical research letters* 33 (22).

Yang, Kun, Toshio Koike, and Baisheng Ye. 2006. "Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets." *Agricultural and Forest Meteorology* 137 (1):43-55.

Yang, Kun, Rachel T Pinker, Yaoming Ma, Toshio Koike, Margaret M Wonsick, Stephen J Cox, Yuanchong Zhang, and Paul Stackhouse. 2008. "Evaluation of satellite estimates of downward shortwave radiation over the Tibetan Plateau." *Journal of Geophysical Research: Atmospheres* 113 (D17).

Zhang, Peng, Wenyuan Li, Sherwin Li, Yang Wang, and Weidong Xiao. 2013. "Reliability assessment of photovoltaic power systems: Review of current status and future perspectives." *Applied Energy* 104:822-833. doi: 10.1016/j.apenergy.2012.12.010.