Article

A Kalman Filter-Based Method for Reconstructing GMS-5 Global Solar Radiation by Introduction of *In Situ* Data

Jingying Fu [1,2], Dong Jiang [1,*], Yaohuan Huang [1], Dafang Zhuang [1] and Yong Wang [1]

[1] State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A, Datun Road, Chaoyang District, Beijing 100101, China; E-Mails: fujy@lreis.ac.cn (J.F.); huangyh@lreis.ac.cn (Y.H.); zhuangdf@igsnrr.ac.cn (D.Z.); wangy@lreis.ac.cn (Y.W.)

[2] University of Chinese Academy of Sciences, Beijing 100049, China

* Author to whom correspondence should be addressed; E-Mail: jiangd@igsnrr.ac.cn; Tel. +86-10-64889433; Fax: +86-10-64855049.

Abstract: Solar radiation is an important input for various land-surface energy balance models. Global solar radiation data retrieved from the Japanese Geostationary Meteorological Satellite 5 (GMS-5)/Visible and Infrared Spin Scan Radiometer (VISSR) has been widely used in recent years. However, due to the impact of clouds, aerosols, solar elevation angle and bidirectional reflection, spatial or temporal deficiencies often exist in solar radiation datasets that are derived from satellite remote sensing, which can seriously affect the accuracy of application models of land-surface energy balance. The goal of reconstructing radiation data is to simulate the seasonal variation patterns of solar radiation, using various statistical and numerical analysis methods to interpolate the missing observations and optimize the whole time-series dataset. In the current study, a reconstruction method based on data assimilation is proposed. Using a Kalman filter as the assimilation algorithm, the retrieved radiation values are corrected through the continuous introduction of local *in-situ* global solar radiation (GSR) provided by the China Meteorological Data Sharing Service System (Daily radiation dataset_Version 3) which were collected from 122 radiation data collection stations over China. A complete and optimal set of time-series data is ultimately obtained. This method is applied and verified in China's northern agricultural areas (humid regions, semi-humid regions and semi-arid regions in a warm temperate zone). The results show that the mean value and standard deviation of the reconstructed solar radiation data series are significantly improved, with

greater consistency with ground-based observations than the series before reconstruction. The method implemented in this study provides a new solution for the time-series reconstruction of surface energy parameters, which can provide more reliable data for scientific research and regional renewable-energy planning.

Keywords: Kalman filter; solar radiation; time series; remote sensing; reconstruction

---

1. Introduction

The parameters of surface energy balances are important inputs for research on global climate change, crop-yield assessment and ecological environment evaluation. At present, remote sensing-derived radiation, temperature and other data sets on a global scale have already been used as standardized products in research and applications. Satellite remotely sensed information is convenient and easily accessed over a large area. However, due to the impact of clouds, aerosols, solar elevation angle and bidirectional reflection, the surface energy parameters retrieved from remote-sensing data are often interrupted on the spatial and temporal scales. A complete dataset for a large region is always difficult to obtain [1,2]. In addition, an indirect retrieval method and the instantaneous features of monitoring have varying degrees of impact on the accuracy of surface-parameter retrieval [3,4]. To reduce such impacts, a time-composite method is generally adopted. For example, the hourly parameters retrieved from geostationary meteorological satellite data are combined into daily average data, and 8-day Moderate Resolution Imaging Spectroradiometer (MODIS) radiation products are combined from daily data. Despite these interpolation methods, the problems of missing data or unstable data quality are still very serious at the regional scale [5], and this limitation will eventually affect the accuracy of land surface-energy balance analyses and simulations [6].

Fitting surface energy parameter time series is a fundamental task for further study [7]. For surface energy parameters, the time-series fitting and noise-removal methods commonly used include the mean diurnal variation and nonlinear regression methods [8–10]. Moffat *et al.* [11] reviewed 15 techniques for estimating the missing values of net ecosystem CO_2 exchange in eddy covariance time series and evaluated their performance for different artificial gap scenarios. Different reconstruction methods have their respective advantages and disadvantages. The nonlinear regression method, look-up table method, dynamic linear regression method and artificial neural networking require *in situ* meteorological data [4,12]. Meteorological data are usually not required for the mean diurnal variation method, therefore, this method can be used even when meteorological data are missing. However, under significant changes of environmental conditions, the modeled results often cannot represent the actual situation [3,13].

In recent years, data-assimilation methods have been adopted in the reconstruction of time-series data. Gu *et al.* [14] designed a simplified NDVI time-series reconstruction method based on data assimilation, and the results showed that the quality and efficiency of long NDVI time-series data processed using this method were superior to those processed using the threshold removal and nonlinear fitting methods. Zhou *et al.* [15] verified the effects of using an ensemble Kalman filter algorithm to calculate the soil moisture and depth of the frozen soil during soil freezing and thawing in

an improved model. Assimilation results based on the data at the testing points showed that an ensemble Kalman filter algorithm improved the calculations of energy and moisture variables using the model. Alavi *et al.* [16] applied the data-assimilation method to reconstruct missing latent heat-flux data using the Penman-Monteith equation as the model operator for the assimilation algorithm and an ensemble Kalman filter as the reconstruction method. A data assimilation experiment of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) land surface temperature (LST) into the Joint UK Land Environment Simulator (JULES) land-surface model via an Ensemble Kalman Filter shows an improvement in the modeled LST, soil moisture and latent and sensible heat fluxes. The results indicate that data assimilation can indeed prove to be a consistent and reliable method of constraining the simulations of complex land surface models [17]. Kawai *et al.* [18] did some research on validating and improving satellite-derived downward surface shortwave radiation which is calculated from GMS-5/VISSR data in 0.05° × 0.05° grids over the northwestern Pacific Ocean using abundant *in situ* data. The cloud attenuation coefficient that was used in the study was affected by satellite-measured albedo, solar zenith angle and even the latitude. In addition, this comparison was only performed at the point scale and could not be used in the regional scale. In this paper, GSR data primarily derived from GMS-5/VISSR are used. The major objective is to establish a time-series reconstruction method based on an ensemble Kalman filter. The method was tested with *in situ* solar radiation measurements. Using climate-zoning information, the method could be applied at a large regional scale to produce a refined solar radiation time-series dataset.

2. Study Area and Data Acquisition

2.1. Data Source

In this study, daily solar radiation data retrieved from a geostationary meteorological satellite (the Japanese Geostationary Meteorological Satellite 5, GMS-5) were used as the main data source. The GMS-5 satellite was launched in March, 1995 into geostationary orbit, located at nominal geodetic coordinates of 0° latitude and 140° E longitude. Full disk images are relayed to ground receiving stations approximately 25–28 times per day, providing up to half-hourly temporal coverage during some periods of the day. The full-width scan angle subtended at the satellite is about 19.6°, which includes some space view. Earth-view data are obtained over a full-scan angle width of 17.4°. The GMS-5 Visible and Infrared Spin Scan Radiometer (VISSR) has four channels: one visible, two thermal infrared channels and a channel sensing thermal radiation emitted by the strong water vapor rotation band at 6.3 μm. Some characteristics of the VISSR are listed in Table 1 [19].

With the support of the China-Netherlands cooperative project *Establishment of a Chinese Energy and Water Balance Monitoring System for Desertification and Food Security Applications*, the China National Satellite Meteorological Centre (CNSMC) and the Institute of Geographic Science and Natural Resources Research (IGSNRR) jointly established the China Energy and Water Balance Monitor System (CEWBMS). The system has now been updated to Version 2.0 (with additional data-processing capacity from China's FengYun-series Weather Satellite), operationally running at IGSNRR, and offering standard satellite-derived GSR data products and services. The GSR data of the research area in 2002 were selected for further study. However, the design life of GMS-5 was five

years, and VISSR had a declining performance. The parameters and models used for the calibration and atmospheric correction have the unavoidable problem of uncertainty, which will affect the data accuracy. In the CEWBMS, two calibration coefficients were introduced in the radiation transfer model to decrease errors mention above. Detailed description can be found in [20].
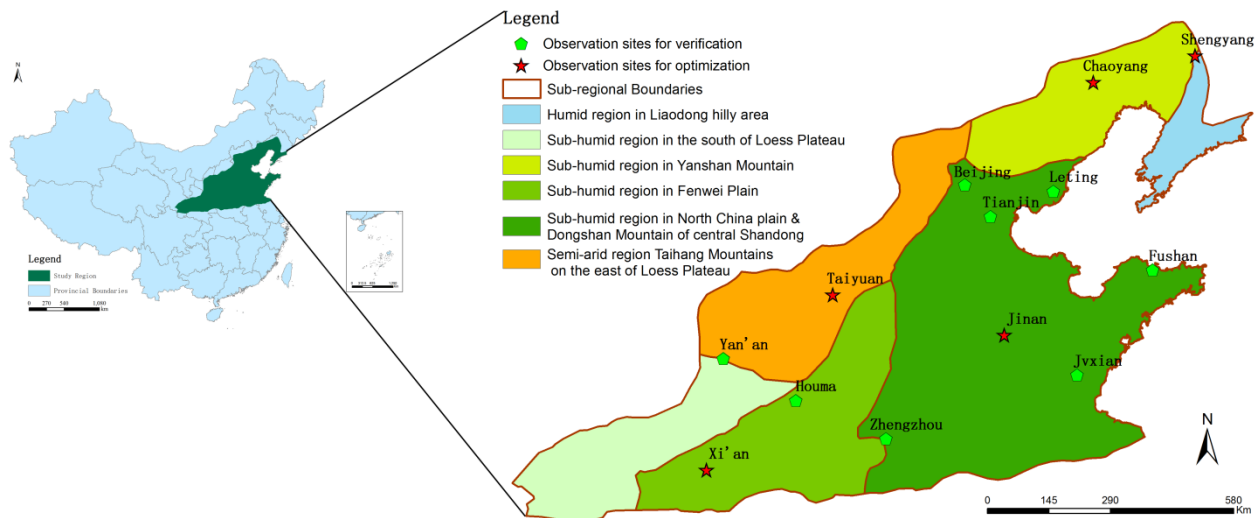
Table 1. Characteristics of the VISSR on GMS-5.

Function	Visble channels	Infrared channels
Number	4 (+4 redundant)	3 (+3 redundant)
IFOV	35×31 μrad	140×140 μrad
Band	0.55–0.90 μm	10.5–11.5 μm (IR1)
		11.5–12.5 μm (IR2)
		6.5–7.5 μm (IR3)
Resolution	1.25 km	5 km
Noise performance	S/N $\geq$ 84 (albedo = 100%) S/N $\leq$ 6.5 (albedo = 25%)	NEΔT (300 K) NEΔT (220 K)
		IR1 $\leq$ 0.35K $\leq$ 1.00K
		IR2 $\leq$ 0.35K $\leq$ 0.90K
		IR3 $\leq$ 0.22K $\leq$ 1.50K
Digitization	6-bit	6-bit
Calibration	Preflight	On board blackbody and space view

The *in situ* GSR data which were measured at each meteorological station were provided by the China Meteorological Data Sharing Service System. There are 122 radiation data collection stations over China where the GSR has been measured since 1957. The daily radiation dataset_(Version 3) was used in this paper which was set up at 28 June 2005. The geographic extent is between latitudes 4.00° and 53.52° N, and longitudes 73.60° and 135.08° E. The *in situ* data were obtained using the surface induction automatic telemetering radiometer and the original data have a relative error of 0.5. The dataset used has been modified by the quality control method. The data type was integer with units of 0.01 MJ/m^2.

2.2. Study Area

The study area, which is located in North China, is one of China's important agricultural areas, with humid, semi-humid and semi-arid climates in a warm temperate zone. Six climatic sub-regions are included in the research area: a sub-humid region in a warm temperate zone in the North China plain and Dongshan Mountain of central Shandong, a sub-humid region in a warm temperate zone in Fenwei Plain, a sub-humid region in a warm temperate zone of the Yanshan mountains, a humid region in a warm temperate zone in the Liaodong hilly area, a semi-arid region in a warm temperate zone in the Taihang mountains on the east of the Loess Plateau, and a sub-humid region in a warm temperate zone in the southern part of the Loess Plateau (Figure 1).

Figure 1. Study area and distribution of observation sites.

3. Method

A time-series reconstruction of the satellite-derived GSR data was conducted as follows:

1. Time-series reconstruction based on a single observation site (point scale):

    ✓ The *in situ* GSR data at meteorological stations in the research area were processed;
    ✓ Daily GSR retrieved from GMS-5 (VISSR) were extracted corresponding to the location information of the meteorological stations ;
    ✓ The time series of GSR retrieved from GMS-5(VISSR) were reconstructed using an ensemble Kalman filter.

2. Verification of the time-series reconstruction based on single site-pixels.
3. Application of the method to the whole research area: taking the climatic sub-regions as the basic unit, the measured radiation data at representative sites in the sub-regions were assimilated using the methods noted above for each pixel in the research area.
4. Regional verification: using the measured radiation values at the sites of (3) as "true values", the application results of our method were verified over the whole research area.

3.1. Kalman Filter-Based Reconstruction Algorithm

A Kalman filter was used as the reconstruction algorithm in this study. In 1960, Kalman *et al.* [21] proposed the concept of a Kalman filter for the state estimation of stochastic processes. The method has been modified and widely applied in many fields of research and has been adopted as one of the most traditional data assimilation algorithms [22–24]. The Kalman filter addresses the general problem of trying to estimate the state X_k of a discrete-time-controlled process that is governed by the linear stochastic-difference equation [25,26]:

$$X_k = \Phi_{k,k-1} X_{k-1} + \Gamma_{k,k-1} W_{k-1} \tag{1}$$

with a measurement Z_k that is defined as follows:

$$Z_k = H_k X_k + V_k \tag{2}$$

The random variables W_{k-1} and V_k represent the process and measurement noise, respectively. These factors are assumed to be independent (of each other), composed of white noise, and with normal probability distributions. In practice, the process noise-covariance and measurement noise-covariance matrices ($\Phi_{k,k-1}$ and $\Gamma_{k,k-1}$) might change with each time step or measurement; however, here, we assume that they are constant. The matrix H_k in the measurement Equation (2) relates the state to the measurement. In practice, H_k might change with each time step or measurement, but here, we assume it is constant, $H = 1$.

We assume the following relations:

$$\begin{cases} E[W_k] = 0 \;, E[W_k W_j^T] = Q_k \delta_{kj} \\ E[V_k] = 0 \;, E[V_k V_j^T] = R_k \delta_{kj} \\ E[WV_j^T] = 0 \end{cases} \tag{3}$$

where Q_k is the nonnegative covariance matrix of W_k and R_k is the positive covariance matrix of V_k; δ_{kj} is the function of *Kronec*ker$-\delta$.

We define N as the number of days of measurement, then:

$$Q_k \;=\; 0.5 * \text{cov}(\text{randn}(1, N)) \tag{4}$$

$$R_k \;=\; 0.5 * \text{cov}(\text{randn}(1, N)) \tag{5}$$

We define $\hat{X}_{k-1}$ to be our *a priori* state estimate at step k, given knowledge of the process prior to step k, and $\hat{X}_k$ is our *a posteriori* state estimate at step k, given measurement Z_k. We can then define the *a priori* state as follows:

$$\hat{X}_{k,k-1} = \Phi_{k,k-1} \, \hat{X}_{k-1} \tag{6}$$

and the *a posteriori* state as follows:

$$\hat{X}_k = \hat{X}_{k,k-1} + K_k [Z_k - H_k \, \hat{X}_{k,k-1}] \tag{7}$$

The *a priori* estimate of the error covariance is as follows:

$$P_{k,k-1} = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + \Gamma_{k,k-1} Q_{k-1} \Gamma_{k,k-1}^T \tag{8}$$

The *a posteriori* estimated error covariance is as follows:

$$P_k = [I - K_k H_k] P_{k,k-1} [I - K_k H_k]^T + K_k R_k K_k^T \tag{9}$$

The matrix *I* is the unit matrix. The matrix *K* in Equation (5) is chosen to be the gain or blending factor that minimizes the *a posteriori* estimated error covariance. One form of the resulting *K* that minimizes (1.6) is given by the following relations:
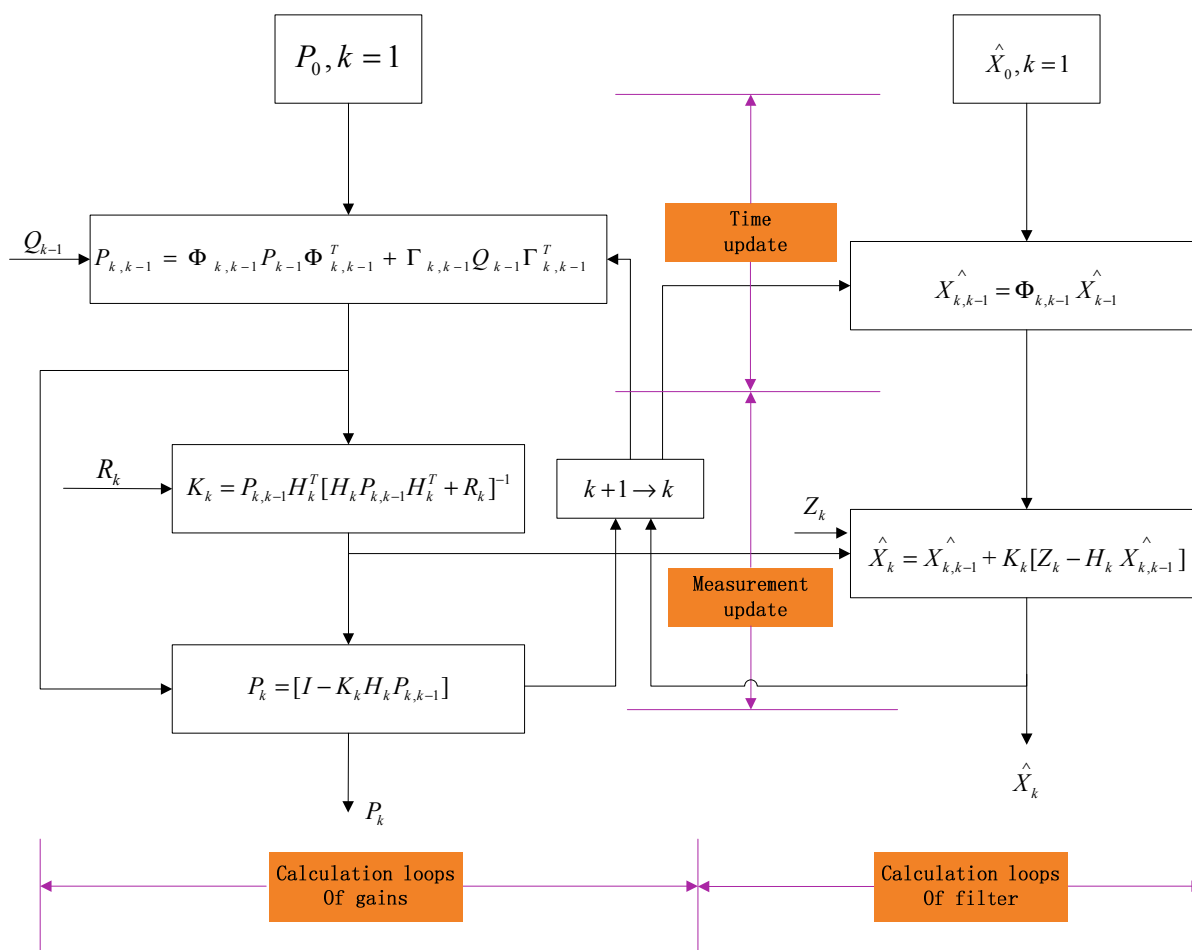
$$K_k = P_{k,k-1} H_k^T [H_k P_{k,k-1} H_k^T + R_k]^{-1} \quad \text{or} \quad K_k = P_k H_k^T R_k^{-1} \tag{10}$$

$$P_k = [I - K_k H_k P_{k,k-1}] \quad \text{or} \quad P_k^{-1} = P_{k,k-1}^{-1} + H_k^T R_k^{-1} H_k \tag{11}$$

The calculation process of a Kalman filter is a constant "forecast-correction" process. Within a filter cycle, in view of the sequence of using systematic and observed information, a Kalman filter has two distinct information-updating processes: a time-update process and an observation-update process. In the time-update process, the predicted values at the current state of the present moment are generated according to the state of the preceding moment. In the observation-update process, the *in situ* GSR is introduced, and the state is re-analyzed by estimating the minimum variance; through several iterations, the global optimum is reached. We can draw the Kalman filter process as presented in the diagram of Figure 1 using Equations (4–9).

In Figure 2, $\hat{X}_0$ is the initial estimated state of the satellite GSR. Z_k is the *in situ* GSR, where k is one day of the year of 2002 which has the measured GSR. Then the gain matrix K_k can be calculated. The *a posteriori* estimated error covariance P_k can be obtained later. The time k and the *in situ* GSR Z_k will update during the processes of "Time update" and "Measurement update", meanwhile, the previous optimal GSR has been estimated.

Figure 2. Algorithm Diagram of the Kalman Filter.

3.2. From Sites to the Regional Scale

After the time-series reconstruction based on a single observation site, the algorithm is applied to the whole research area. First, it is assumed that the altitude, temperature, rainfall, aridity, topography and other conditions are basically the same in each climatic sub-region [27]. In addition, the solar

radiation values retrieved from remote-sensing data represent the unit values for approximately 25 km^2 of land area. Therefore, the solar radiation values corresponding to pixels of the entire research area can be reconstructed using the *in situ* GSR at the representative sites in each climatic sub-region to obtain a spatial distribution map of the solar radiation at different dates after reconstruction over the entire research area.

4. Result and Analyses

4.1. Results and Analysis of Reconstruction on the Single Site-Pixel Scale

At each radiation-observation site, which is located near the center of the corresponding climatic sub-region, the *in situ* GSR can be obtained and applied to reconstruct the daily GSR retrieved from the GMS-5 (VISSR) in 2002 by using the Kalman filter-based algorithm mentioned in Section 3. The GMS-5 (VISSR) derived GSR, the mean square error (MSE) and the differences of each pair (retrieved GSR minus *in situ* GSR, reconstructed GSR minus *in situ* GSR and reconstructed GSR minus retrieved GSR) are shown in Figure 3.

Figure 3. Reconstructed GSR, *in situ* GSR, retrieved GSR, differences and the MSE of the three datasets at (**a**) Taiyuan; (**b**) Chaoyang; (**c**) Xi'an; (**d**) Jinan; and (**e**) Shenyang meteorological stations in 2002.

(**a**1)   (**a**2)   (**a**3)

(**b**1)   (**b**2)   (**b**3)

Figure 3. *Cont.*

(**c**1)　　　　　　　　(**c**2)　　　　　　　　(**c**3)

(**d**1)　　　　　　　　(**d**2)　　　　　　　　(**d**3)

(**e**1)　　　　　　　　(**e**2)　　　　　　　　(**e**3)
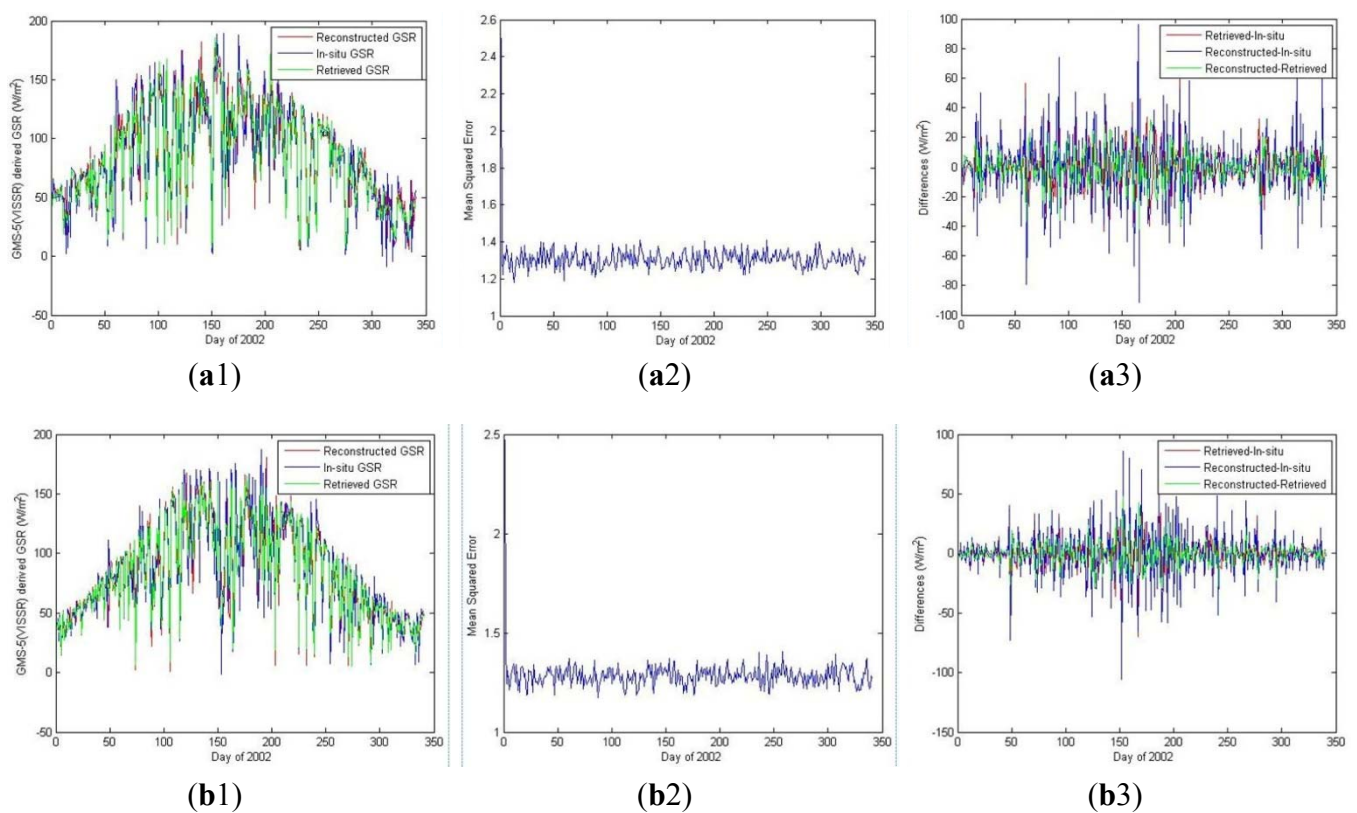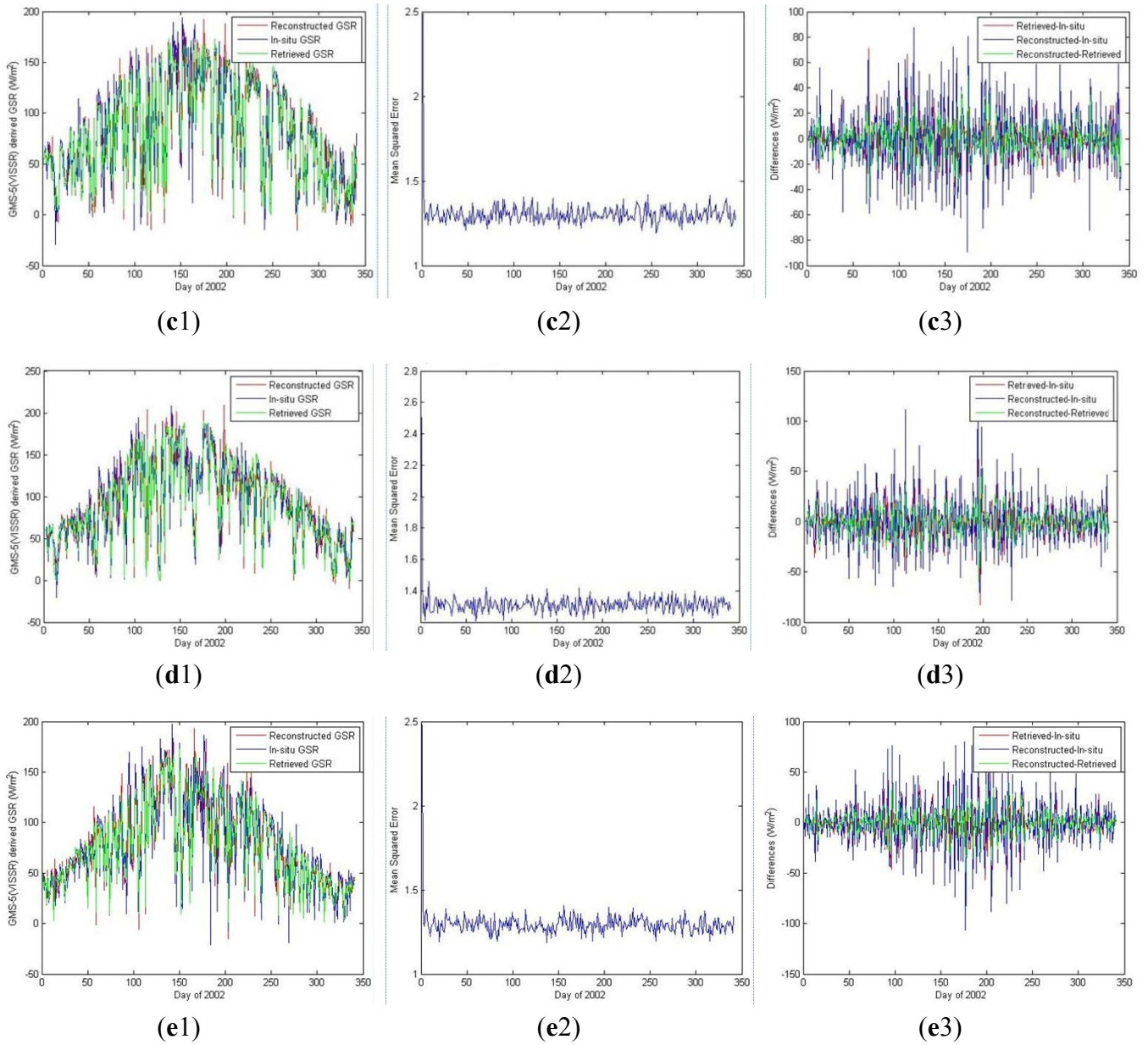
Panels A to E in Figure 3 presented the meteorological stations of Taiyuan, Chaoyang, Xi'an, Jinan and Shenyang, respectively, which were used for reconstruction and their precise positions were shown in Figure 1. The first column is the time series of retrieved, *in situ* and reconstructed GSR. The second column is the MSE which is an estimator to determine the extent to which the model fits the data. In this paper, MSE is used to quantify the difference between the reconstructed GSR and the *in situ* GSR [Equation (12)]. It is about 2.5 at the beginning of reconstruction of GSR by Kalman filter and decreased to less than 1.5 and stayed stable, which means the variance of the two dataset became stabilized:

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(insitu_k - reconstruct_k)^2 \tag{12}$$

The third column is the time series of retrieved GSR minus in–situ GSR (the red line), reconstructed GSR minus *in situ* GSR (the blue line) and reconstructed GSR minus retrieved GSR (the green line). From this column we can found that most of the differences are near the line of zero. Most of the differences in the winter (the first fifty days and the last fifty days in 2002) are lower than 10 W/m^2. The differences between reconstructed GSR and *in situ* GSR are higher than the other two, and are less than 20 W/m^2.

The statistics in Table 2 show that the standard deviations of the solar radiation data for each site are significantly improved after the reconstruction using the Kalman filter. Meanwhile, the means (mean before reconstruction, mean of *in situ* data, and mean after reconstruction) of the reconstructed dataset show more consistent variation with those of *in situ* dataset (Table 1).

Table 2. Comparison of GMS-5(VISSR) derived GSR, *in situ* GSR and the reconstructed GSR.

Meteorological stations	Mean of GMS-5 (VISSR) derived GSR (w/m^2)	Mean of *in situ* GSR (w/m^2)	Mean after reconstruction (w/m^2)	Standard deviations of GMS-5 (VISSR) derived GSR	Standard deviations of *in situ* GSR	Standard deviations after reconstruction
Chaoyang	88.13	84.52	84.50	44.89	39.17	42.77
Taiyuan	87.65	86.09	86.16	44.37	41.20	42.00
Xi'an	86.50	82.50	82.45	51.35	48.31	49.88
Shenyang	84.20	78.56	78.57	45.73	40.70	44.94
Jinan	93.37	93.17	93.14	44.08	47.44	47.99

The spatial variation of the solar radiation data was relatively stable (on scales of pixels of dozens of meters or a few kilometers), and the temporal variation was significant. Such a variation pattern of the time series is obviously seasonal, which can be used as the basis for precision analyses and the reconstruction of time series of radiation data products retrieved from remote sensing data.

4.2. Reconstruction on the Research-area Scale

The reconstruction results of the solar radiation values retrieved from the remote-sensing data in each climatic sub-region on the representative dates in the four seasons of 2002 are shown as follows (Figure 4).

Figure 4. The distribution of global solar radiation on (**a**) 15 February; (**b**) 15 May; (**c**) 15 August; and (**d**) 15 November, 2002.

(**a**1)   (**a**2)   (**a**3)

Figure 4. *Cont.*

(**b**1)   (**b**2)   (**b**3)

(**c**1)   (**c**2)   (**c**3)

(**d**1)   (**d**2)   (**d**3)
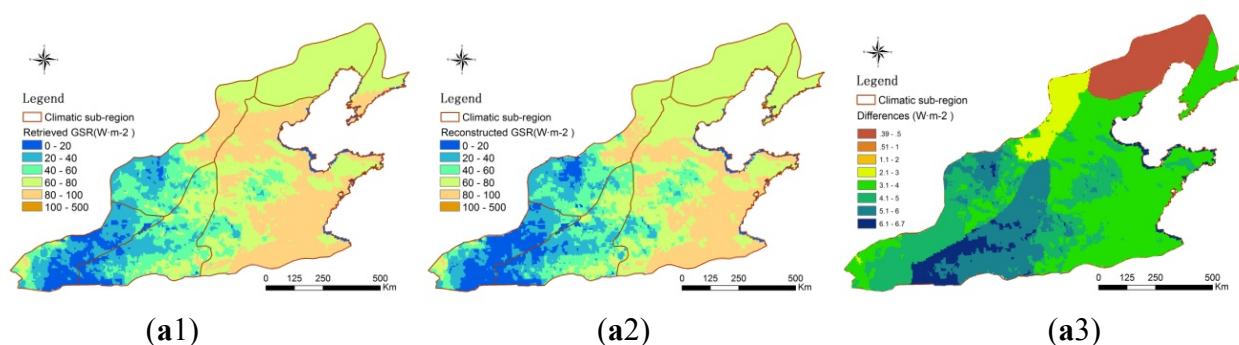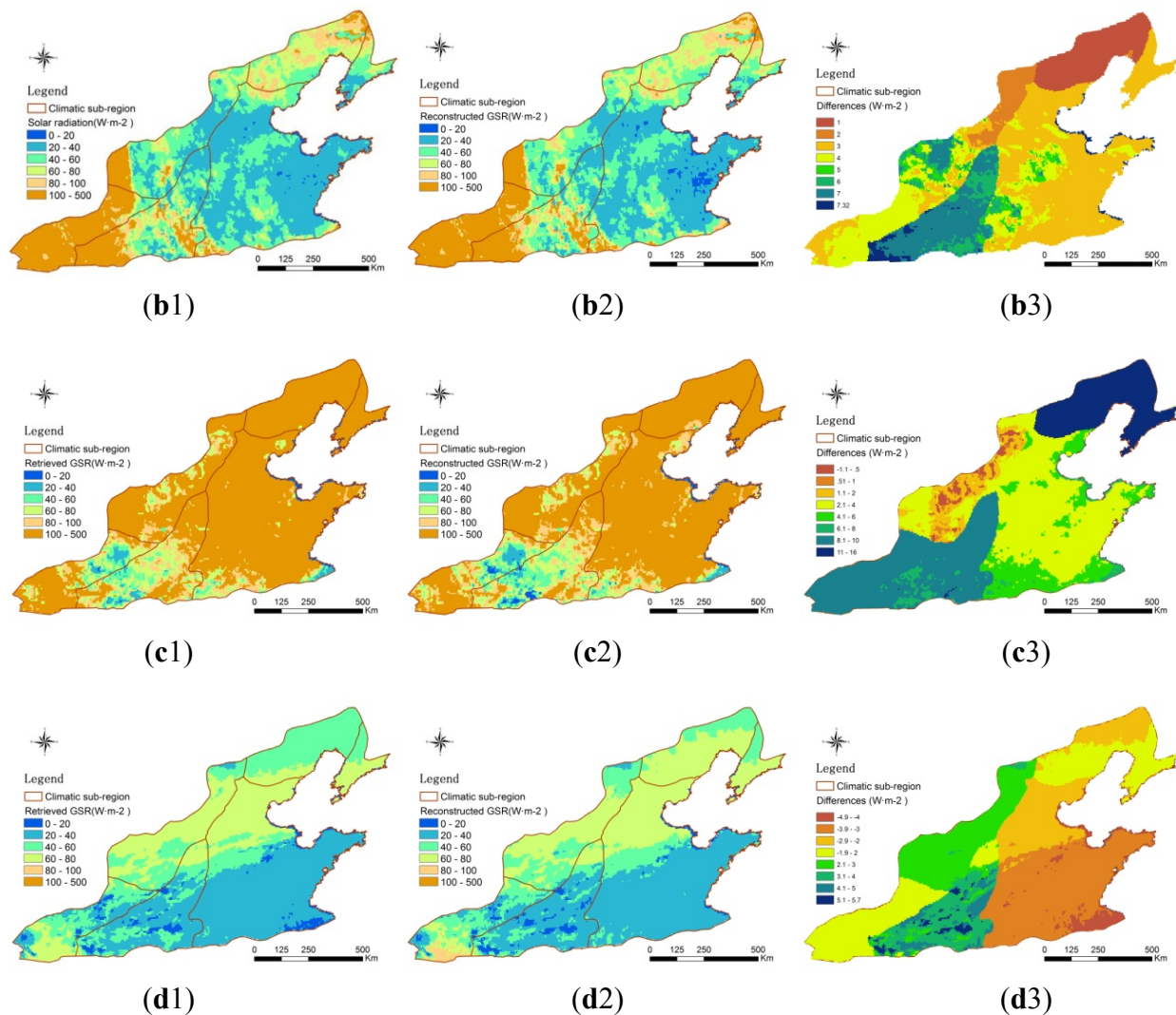
In Figure 4, columns 1–3 represent the distributions of the retrieved GSR, reconstructed GSR and the differences between them, while rows a–d show four different days (15 February, 15 May, 15 August and 15 November, 2002) corresponding to four seasons in China. From Figure 4, we can see that in the four seasons, two dates show significant variation of the solar radiation values before and after reconstruction: 15 February and 15 November. On 15 February, after the reconstruction of the spatial distribution of the solar radiation over the research area, except the semi-humid regions in the warm temperate zone of the Yanshan mountains, there is a considerable reduction in the radiation values in all of the climatic sub-regions. On 15 November, the radiation values show an increasing trend in the semi-humid regions in the warm temperate zone of the Yanshan mountains and in the semi-arid regions of the warm temperate zone of the Taihang mountains on the eastern Loess Plateau. On 15 May and 15 August, the high solar radiation values show less variation, both in degree and scope. In the c column, the differences for dates 1–4 is 0.39~6.7, 1–7.32, −1.1–16 and −4.9–5.7 w/m², respectively. The biggest difference of 16 w/m² occurred on 15 August in the humid region.

Except for the data at the radiation observation sites that were used to establish the fitting equation, the data at the remaining observation sites were used to verify the reconstructed results over the research area. Data at the observation sites in Yan'an, Houma, Beijing, Tianjin, Leting, Zhengzhou, Ju

County and Fushan were selected for verification. Statistical analysis was conducted for the solar radiation values of 2002 retrieved from the VISSR data (before reconstruction) at these sites, the *in situ* GSR and the solar radiation values after reconstruction (Table 3).

Table 3. Comparison of the remotely sensed solar radiation data before and after reconstruction.

Meteorologi cal stations	Mean of GMS-5 (VISSR) derived GSR (W/m^2)	Mean of *in situ* GSR (W/m^2)	Mean after reconstruction (w/m^2)	Standard deviations of GMS-5 (VISSR) derived GSR	Standard deviations of *in situ* GSR	Standard deviations after reconstruction
Yan'an	86.17	84.16	84.64	47.22	42.15	47.70
Houma	88.04	82.41	83.95	47.92	43.21	47.34
Beijing	87.94	84.69	87.49	44.84	43.13	43.45
Tianjin	90.09	87.28	89.42	45.84	43.20	44.42
Leting	88.56	84.84	88.13	46.93	42.84	45.47
Zhengzhou	88.90	88.00	88.48	49.18	46.96	47.65
Jvxian	93.81	91.27	93.55	46.93	42.75	45.55
Fushan	89.47	87.28	89.07	47.04	41.79	45.58

From Table 3, we can see that the means of the reconstructed solar radiation values are close to those of the measured values before the reconstruction. Meanwhile, the standard deviations of the solar radiation values at the majority of the radiation-observation sites are smaller than those before the reconstruction, which means the scatter of the data set has been reduced. In the verification points of Yan'an, the standard deviations of the data after reconstruction are slightly larger than those before reconstruction, but the numerical change is small. The possible causes of error are analyzed in the following section.

4.3. Error Analysis

Based on the analysis of the reconstruction results using the Kalman filter in Section 3.2 and the verification of the reconstruction results over the whole research area in Section 4.2, our method for time-series reconstructions of short-term solar radiation data achieves good results on the whole. However, at some sites, the accuracy fails to be significantly improved. The possible sources of error are analyzed as follows:

1. Errors from the satellite data source. The global solar radiation data for 2002 came from Japan's GMS-5(VISRR), launched in 1995. The satellite had been running beyond its intended service life even by 2000, and the sensor performance has been declining. The parameters and models used during the calibration and atmospheric correction have the unavoidable problem of uncertainty, which affects the data accuracy. However, the products we used in this paper have been modified by two calibration coefficients.

2. The retrieval algorithm is an indirect method. The radiation signals received by the sensor are the results of the entire process of atmospheric reflection, absorption, transmission to the ground and reflection back to the sensor, which is very complex. The satellite remote-sensing retrieval involves many unknown parameters, and empirical or semi-empirical formulas are often adopted. Therefore, a precise quantification is almost impossible.

3. Errors may inevitably arise from the reconstruction method established in this study in the following two aspects: during the reconstruction process using the Kalman filter, the setting of the initial values is uncertain and requires experience. The accuracy of the initial value setting will impact the reconstruction results. We used the first *in situ* GSR as the initial estimated state in this paper, and the MSE of the first day of 2002 was the largest. Moreover, there are scale differences between the ground-observation sites and the satellite pixels, which is another source of error. The *in situ* GSR obtained at the point scale while the GMS-5(VISRR) derived GSR were retrieved at the resolution of 5 km. We used MSE in this study to quantify the difference between the reconstructed GSR and the *in situ* GSR. It is about 2.5 at the beginning of reconstruction of GSR by Kalman filter and decreased to less than 1.5 to the end.

5. Conclusions

In this study, based on daily solar radiation retrieved from remote-sensing data, a complete set of time-series reconstruction methods for solar radiation values was proposed. The following main conclusions are reached:

1. The spatial variation of the GMS-5 solar radiation data was relatively stable on scale of pixels of 25 km^2 in this paper, and the temporal variation was significant. Such a variation pattern of the time series is obviously seasonal, which can be used as the basis for precision analyses and the reconstruction of time series of radiation data products retrieved from remote-sensing data.

2. On the site-pixel scale, the accuracy and consistency of the entire time series was effectively improved by reconstruction using a Kalman filter. The MSE which can quantify the difference between the reconstructed GSR and the *in situ* GSR is about 2.5 at the beginning of reconstruction of GSR by Kalman filter and decreased to less than 1.5 later to the end. Some representative sites in the research area were selected, and the measured solar radiation values were constantly introduced as "true values" to reconstruct the retrieved data using a Kalman filter. The means and standard deviations of the solar radiation data after reconstruction were significantly improved which can be seen from Table 2; From the third column of Figure 3, most of the differences of the pairs were near zero: lower than 10 W/m^2 in winter and lower than 20 W/m^2 in other days. This finding indicates that the time-series reconstruction method established in this study for solar radiation data based on Kalman filtering is effective for applied research.

3. Apart from the Kalman filter algorithm and the improved algorithm, other new data-assimilation methods can also be used in further studies. Through comparative analysis of the application effects of various methods, we can obtain the desired reconstruction results. With respect to the conversion method for data on the point scale to data on the regional scale, data-assimilation techniques using a land surface-process model as the model operator and a Kalman filter as the assimilation algorithm have been rapidly developed. However, there have been few reports on the application of data-assimilation techniques in time-series reconstruction for surface parameters. Therefore, this technology has great application potential and research value in this respect.

Acknowledgments

Conflict of Interest

The authors declare no conflict of interest.

References

1. Kondrashov, D.; Ghil, M. Spatio-temporal filling of missing points in geophysical data sets. *Nonlinear Process. Geophys.* **2006**, *13*, 151–159.
2. Borak, J.S.; Jasinski, M.F. Effective interpolation of incomplete satellite-derived leaf-area index time series for the continental United States. *Agric. For. Meteorol.* **2009**, *149*, 320–332.
3. Neteler, M. Estimating daily Land Surface Temperatures in mountainous environments by reconstructed MODIS LST data. *Remote Sens.* **2010**, *2*, 333–351.
4. Chen, J.; Jonsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay Filter. *Remote Sens. Environ.* **2004**, *91*, 332–344.
5. Huang, Y.H.; Jiang, D.; Zhuang, D.F.; Ren, H.Y.; Yao, Z. Filling gaps in vegetation index measurements for crop growth monitoring. *Afr. J. Agric. Res.* **2011**, *6*, 2920–2930.
6. Li, R.; Zhang, X.; Liu, B.; Zhang, B. Review on methods of remote sensing time-series data reconstruction. *J. Remote Sens.* **2009**, 335–341.
7. Lovell, J.L.; Graetz, R.D. Filtering pathfinder AVHRR land NDVI data for Australia. *Int. J. Remote Sens.* **2001**, *22*, 2649–2654.
8. Schafer, J.L.; Graham, J.W. Missing data: Our view of the state of the art. *Psychol. Methods* **2002**, *7*, 147–177.
9. Viovy, N.; Arino, O.; Belward, A.S. The best index slope extraction(BISE): A method for reducing noise in NDVI time-series. *Int. J. Remote Sens.* **1992**, *13*, 1585–1590.
10. Wang, D.; Liang, S. Singular Spectrum Analysis for Filling Gaps and Reducing Uncertainties of MODIS Land Products. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008.
11. Moffat, A.M.; Papale, D.; Reichstein, M.; Hollinger, D.Y.; Richardson, A.D.; Barr, A.G.; Beckstein, C.; Braswell, B.H.; Churkina, G.; Desai, A.R. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agric. For. Meteorol.* **2007**, *147*, 209–232.
12. Huang, Y.H.; Wang, J.H.; Jiang, D.; Zhou, Q. Reconstruction of MODIS-EVI time-series data with S-G filter. *Geomat. Inf. Sci. Wuhan Univ.* **2009**, *34*, 1440–1443.
13. Xu, Z.; Liu, S.; Xu, T.; Wang, J. Comparison of the gap filling M ethods of evapotranspiration measured by eddy covariance system. *Adv. Earth Sci.* **2009**, 372–382.
14. Gu, J.; Li, X.; Huang, C.; Okin, G.S. A simplified data assimilation method for reconstructing time-series MODIS NDVI data. *Adv. Space Res.* **2009**, *44*, 501–509.

15. Zhou, J.; Wang, G.X.; Li, X.; Yang, Y.M.; Pan, X.D. Data assimilation algorithm apply to energy-water balance analysis of the high cold ecosystem at qinghai-tibet plain, Northwest China. *Adv. Earth Sci.* **2008**, *23*, 965–973.

16. McLaughlin, D.; Zhou, Y.H.; Entekhabi, D. Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Mon. Weather Rev.* **2006**, *134*, 2128–2142.

17. Ghent, D.;Kaduk, J.; Remedios, J.; Ardo, J.; Balzter, H. Assimilation of land-surface temperature into the land surface model JULES with an ensemble Kalman filter. *J. Geophys. Res.* **2010**, *115*, doi:10.1029/2010JD014392.

18. Kawai, Y.; Kawamura, H. Validation and improvement of satellite-derived surface solar radiation over the northwestern Pacific ocean. *J. Oceanogr.* **2005**, *61*, 79–89.

19. Prata, A.; Cechet, R. An assessment of the accuracy of land surface temperature determination from the GMS-5 VISSR. *Remote Sens. Environ.* **1999**, *67*, 1–14.

20. Kalman, R.E. A new approach to linear filtering and prediction theory. *Trans. ASME-J. Basic Eng.* **1960**, 34–45.

21. Rosema, A. Using METEOSAT for operational evapotranspiration and biomass monitoring in the Sahel region. *Remote Sens. Environ.* **1993**, *45*, 1–25.

22. Huang, C.; Li, X.; Lu, L. Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. *Remote Sens. Environ.* **2008**, *112*, 1320–1336.

23. Sakov, P.; Evensen, G.; Bertino, L. Asynchronous data assimilation with the EnKF. *Tellus Ser. A* **2010**, *62*, 24–29.

24. Aravéquia, J.A.; Szunyogh, I.; Fertig, E.J.; Kalnay, E.; Kuhl, D.; Kostelich, E.J. Evaluation of a strategy for the assimilation of satellite radiance observations with the local ensemble transform kalman filter. *Mon. Weather Rev.* **2011**, *139*, 1932–1951.

25. Welch, G.; Bishop, G. *An Introduction to the Kalman Filter*; University of North Carolina at Chapel Hill: Chapel Hill, NC, USA, 1995; Volume 7.

26. Fu, M.Y.; Deng, Z.H.; Zhang, J.W. *Kalman Filtering Theory and Its Application in Navigation Systems*; Science Press: Beijing, China, 2003; pp. 17–19.

27. Zheng, J.Y.; Yin, Y.H.; Li, B.H. A new scheme for climate regionalization in China. *Acta Geogr. Sin.* **2010**, *65*, 3–12.