A Simple and Universal Aerosol Retrieval Algorithm for Landsat Series Images Over Complex Surfaces

Jing Wei, Bo Huang, Lin Sun, Zhaoyang Zhang, Lunche Wang, and Muhammad Bilal

Key Points:
- An improved 30 m resolution Landsat AOD retrieval algorithm is proposed.
- AOD retrievals agree well with AERONET ground measurements.
- AOD retrievals show better accuracy and less bias than MOD04 AOD products.

Abstract
Operational aerosol optical depth (AOD) products are available at coarse spatial resolutions from several to tens of kilometers. These resolutions limit the application of these products for monitoring atmospheric pollutants at the city level. Therefore, a simple, universal, and high-resolution (30 m) Landsat aerosol retrieval algorithm over complex urban surfaces is developed. The surface reflectance is estimated from a combination of top of atmosphere reflectance at short-wave infrared (2.22 μm) and Landsat 4–7 surface reflectance climate data records over densely vegetated areas and bright areas. The aerosol type is determined using the historical aerosol optical properties derived from the local urban Aerosol Robotic Network (AERONET) site (Beijing). AERONET ground-based sun photometer AOD measurements from five sites located in urban and rural areas are obtained to validate the AOD retrievals. Terra MODerate resolution Imaging Spectrometer Collection (C) 6 AOD products (MOD04) including the dark target (DT), the deep blue (DB), and the combined DT and DB (DT&DB) retrievals at 10 km spatial resolution are obtained for comparison purposes. Validation results show that the Landsat AOD retrievals at a 30 m resolution are well correlated with the AERONET AOD measurements ($R^2 = 0.932$) and that approximately 77.46% of the retrievals fall within the expected error with a low mean absolute error of 0.090 and a root-mean-square error of 0.126. Comparison results show that Landsat AOD retrievals are overall better and less biased than MOD04 AOD products, indicating that the new algorithm is robust and performs well in AOD retrieval over complex surfaces. The new algorithm can provide continuous and detailed spatial distributions of AOD during both low and high aerosol loadings.

1. Introduction
Atmospheric aerosols are liquid or solid tiny suspended particles in the atmosphere with diameters ranging from $10^{-3}$ μm to $10^2$ μm; they show a variety of characteristics and are caused by natural and anthropogenic activities (Kauffman, Tanre, et al., 1997). Aerosols play important roles in the radiation budget, climate change, air quality, and visibility via scattering and absorption of the incoming solar light, and they have serious effects on human health (Kauffman et al., 2002; Li et al., 2011; Ramanathan et al., 2001). To fully understand the impacts of aerosol particles and develop climate and air quality control strategies, the retrieval of aerosol optical properties is required. Aerosol optical depth (AOD) is an important quantitative parameter used to estimate the total aerosol amount in the atmosphere (Clarke et al., 2001; Holben et al., 2001). AOD is the sum extinction of both scattering and absorption of solar light due to aerosol particles (Van de Hulst, 1974). Currently, remote sensing technology has been an effective method for detecting the optical properties and spatial distribution of aerosols from local to global scales.

In the mid-1970s, satellite remote sensing of aerosol was initiated to retrieve the amount of dust particles over the ocean using Landsat and NOAA-POES/GOES data sets (Fraser, 1976; Norton et al., 1980). Aerosol optical properties have been retrieved using the NOAA advanced very high resolution radiometer (Hauser et al., 2005; Li et al., 2013; Mei et al., 2013), Total Ozone Mapping Spectrometer (TOMS; Herman et al., 1997; Torres et al., 2002), Polarization and Directionality of the Earth’s Reflectances (POLDER; Leroy et al., 1997), Multi-angle Imaging Spectro-Radiometer (Kahn et al., 2005), Sea-Viewing Wide Field-of-View Sensor (Sayer et al., 2012), the MODerate resolution Imaging Spectrometer (MODIS; Hsu et al., 2004; Kaufman, Tanre, et al., 1997, Kaufman, Wald, et al., 1997; Levy, Remer, & Dubovik, 2007, Levy, Remer, et al., 2007), and the Visible/Infrared Imager Radiometer Suite (VIIRS; Jackson et al., 2013; Zhang et al., 2016) over both land
and ocean. The main objective of satellite aerosol remote sensing is to retrieve optical properties between the top of atmosphere (TOA) and the ground surface as signals received from satellite sensors, which is the amalgamation of the atmospheric path reflectance and surface reflectance. Therefore, estimation of surface reflectance is an important step in developing an aerosol retrieval algorithm. Satellite remote sensing techniques have been used to estimate the surface reflectance over both dark targets (Kaufman, Tanre, et al., 1997; Kaufman, Wald, et al., 1997; Levy, Remer, & Dubovik, 2007; Levy et al., 2010) and bright surfaces (Hsu et al., 2004, 2006).

China has been experiencing serious air pollution problems (i.e., haze, smoke, and dust), mainly caused by economic development and the resultant release of pollutants from agricultural and industrial activities, urban sprawl, and dust transportation from Mongolian regions. Atmospheric particulate pollution is becoming more serious, and urban aerosol has been the primary pollutant and major source in city and local areas (Huang et al., 2014; Sun, Wei, Duan, et al., 2016; Tang et al., 2016). Operational MODIS AOD products (MOD04) provide daily global AOD distributions at 10 km spatial resolution from 2001 and are most suitable for global aerosol spatial distribution studies. Due to their coarse spatial resolutions, MOD04 AOD products are very limited in their application to air quality monitoring and spatial variations of atmospheric pollutants at the city level or other small scales (Bilal et al., 2013; Bilal et al., 2014; Li et al., 2005; Sun et al., 2017).

Many researchers are studying aerosol and atmospheric particulates at the city scale, and many high-resolution aerosol retrieval algorithms have been developed for MODIS (Bilal et al., 2013; Li et al., 2005; Li et al., 2014; Lyapustin et al., 2011; Oo et al., 2010), VIIRS (Wang et al., 2017; Zhang et al., 2016), Landsat (Luo et al., 2015; Sun et al., 2015), China land resources satellite series (i.e., HJ-1 and GF-1; Li et al., 2012; Sun et al., 2010, 2017; Zhong et al., 2017), and other satellites. The spatial resolutions of distributions of aerosol have gradually increased from 10 km to 3 km (Oo et al., 2010; Remer et al., 2013), 1 km (Li et al., 2005; Lyapustin et al., 2011; Oo et al., 2010), 500 m (Bilal et al., 2013; Li et al., 2014), 30 m (Luo et al., 2015; Zhong et al., 2017), and higher resolutions (Luo et al., 2015) for AOD studies. The Landsat program has launched several sensors and provided long-term data records at 30 m resolution every 16 days over the last 40 years. Due to the high spatial resolution, long-time observations and high image quality, Landsat can play an important role in monitoring urban air quality. Previous studies have shown that outside of special conditions such as fire smoke, air conditions typically do not vary considerably within 30 m, which allows aerosol retrieval at high resolution (Chudnovsky et al., 2014; Hsu et al., 2004; Richter, 1997). Therefore, the main purpose of this paper is to explore a high-resolution aerosol retrieval method for Landsat satellites.

Beijing is the capital of China and is located in Eastern China (39–42°N, 115–118°E, Figure 1). The region features complex surfaces and consists of various land use types. Vegetated areas including grassland and forest with dense vegetation coverage are mainly distributed in the west, northeast, and northwest of the region, whereas the southeastern areas are mainly covered agricultural farmlands. Bright areas refer to land surfaces covered by little or sparse vegetation with high surface reflectance (i.e., urban, bare land, arid/semiarid areas, or deserts). Central Beijing is a typically bright urban area with a dense population and extensive industrial activity. However, irregular changes in common land use types in the region create challenges in the estimation of land surface reflectance (LSR), which has been the key factor affecting AOD retrieval from remote sensing data. Previous studies found that 0.01 estimation errors in the surface reflectance can lead to −0.1 errors in AOD retrieval (Kaufman, Tanre, et al., 1997). Therefore, to obtain high-quality AOD retrievals, the accurate estimation of surface reflectance is essential for aerosol retrieval over such complex land surfaces.

To solve the above problems, two new approaches for improving the estimation of LSR over two customized extinct categories, densely vegetated areas and bright areas, are developed in this paper. The surface reflectance of these areas in visible channels is determined through modified relationships between visible and short-wave infrared (SWIR) channels and a precalculated LSR database constructed with existing high-quality Landsat 4–7 surface reflectance products. Then, a simple, universal, and high-resolution (30 m) AOD retrieval algorithm for Landsat sensors (i.e., TM, ETM+, and OLI (Operational Land Imager)) is developed based on the above new approaches for the estimation of surface reflectance as well as for improving cloud mask, assumptions of aerosol type, and other features over complex urban areas. Landsat TM/ETM+/OLI images covering the period 2008–2016 are collected and used to perform aerosol retrieval experiments with the new proposed algorithm. Then, the aerosol retrievals are validated against the Aerosol Robotic Network (AERONET)
ground-based AOD measurements and compared with current MODIS standard aerosol products at 10 km spatial resolution.

2. Remote Sensing and Ground-Based Data Sets

2.1. Landsat Satellite Data

The Landsat program has launched eight satellites since 23 July 1972, and the Landsat series of satellites provides the longest temporal record of multispectral geophysical data of the Earth’s surface on a global basis, spanning the last 40 years. Landsat 4 and 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI)/Thermal Infrared Sensor are three main Landsat sensors and were launched in 1 March 1984, 15 April 1999, and 11 February 2013, respectively. Continuing the Landsat program, Landsat 8 retains and includes all spectral channels settings of TM and ETM+ and offers improvements including (i) the addition of two channels: a coastal channel for coastal zone detection and a short-wave infrared channel for cloud detection, (ii) avoidance of the water vapor absorption effect at 0.825 μm, and (iii) narrowing of the spectral range in the panchromatic channel to better distinguish vegetation and nonvegetation areas. Landsat series images have been collected continuously and provide a unique resource to assist a broad range of specialists in managing the world’s agriculture, water, forests, and other natural resources and disaster mitigation. Table 1 lists the detailed parameters of different Landsat series satellites.

Landsat 4–7 surface reflectance climate products were generated from the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS). Water vapor, air pressure, air temperature, ozone, geopotential height, AOD, and digital elevation along with Landsat TM or ETM+ images are inputs into the Second Simulation of a Satellite Signal in the Solar Spectrum (6S) radiative transfer model to perform the atmospheric correction and generate the surface reflectance (Vermote et al., 1997). The water vapor, pressure, and air temperature data are obtained from the National Center for Environmental Prediction Reanalysis Grid. The digital elevation data are provided by the Global Climate Model DEM, and the ozone data are obtained from the
OMI/TOMS ozone observations. The AOD is obtained from the correlation between chlorophyll absorption and bound water absorption of the scene (Kalnay et al., 1996; Masek et al., 2006).

Most Landsat 4–7 images can be processed with LEDAPS surface reflectance products; images that cannot be processed include those with missing auxiliary ozone information and those affected by atmospheric correction in following conditions: (1) hyperarid or snow-covered regions, (2) low Sun angle conditions, (3) coastal regions where the land area is small relative to adjacent water, and (4) areas with extensive cloud contamination. The atmospheric correction accuracy of LEDAPS surface reflectance is typically 0.05 compared with AERONET-derived reflectance and 0.07 compared with MODIS surface reflectance under favorable conditions. No significant differences in the performance have been found between Landsat TM and ETM+ surface reflectance products (Claverie et al., 2015; Ju et al., 2012; Maiersperger et al., 2013). LEDAPS LSR product has been updated numerous times since December 2012 with further adjustments to spectral indices, altered by conversion, subset, and reprojection, and the newest version, 7.1, published in December 2016, is selected for this study.

2.2. AERONET Ground-Based Measurements

AERONET is a federation of ground-based remote sensing aerosol networks and has provided a long-term, continuous, and accessible public domain database of globally distributed observations of spectral AOD. AERONET contains aerosol optical properties including size distribution, single scattering albedo, asymmetry parameter, phase functions, and the complex index of refraction in diverse aerosol regimes (Holben et al., 2001; Smirnov et al., 2000). The AOD observations cover a wide range of visible to near-infrared channels (0.34–1.02 μm) with a high temporal resolution of 15 min and a low uncertainty of 0.01–0.02.

The processing algorithms have evolved from version (V) 1.0 to V 2.0 (fully released in July 2006) and now are available from the official AERONET websites. The AOD measurements are computed as three data quality levels (L): L1.0 (unscreened), L1.5 (cloud screened), and L2.0 (cloud screened and quality assured). The AERONET V3.0 L1.0 and L1.5 near-real-time (NRT) databases became available on 22 June 2016. Compared with the V2.0 products, the V3.0 L1.0 product applies new temperature characterizations and NO2 OMI L3 climate climatology for the period 2004–2013. However, the L1.5 product is calculated from the L1.0 product with new automatic quality controls. For cloud screening, the L1.5 product improves the removal of optically thin cirrus contamination and preserves more highly variable smoke, comparing well with V2.0 L2.0 AOD measurements. For data quality control, it removes sensor temperature artifacts and aerosols with poor spectral dependence and those affected by solar eclipses and window obstructions (http://aeronet.gsfc.nasa.gov/new_web/Documents/AERONET_V3_AOD.pdf).

To help monitor the air quality over Beijing, five ground-based air quality AERONET sites have been established. Four of these sites are typical urban sites: Beijing (39.977°N, 116.381°E), Beijing_RADI (40.005°N, 116.379°E), Beijing_CAMS (39.933°N, 116.317°E), and Beijing_PKU (39.992°N, 116.310°E). These sites are located in the center of bright urban areas and have been operated since 2001, 2010, 2012, and 2016 (Figure 1). XiangHe (39.754°N, 116.962°E) is located in the less bright suburbs 70–80 km from the Beijing site and has operated since 2001. The AERONET V3.0 L1.5 aerosol measurements are selected for validation purposes in this study.
3. Landsat Aerosol Retrieval Algorithm

A simple, universal, and high-resolution (30 m) Landsat (SUHL) aerosol retrieval algorithm is developed with the simplified look-up tables (LUT) constructed from the 6S model. The SUHL algorithm requires Landsat TM, ETM+, and OLI images along with metadata files to calculate the TOA reflectance. The LSR is estimated from the combination of the relationships among TOA reflectance at visible and SWIR channels and Landsat 4–7 surface reflectance climate data records over densely vegetated areas and bright areas. The aerosol properties, including the single scattering albedo and the asymmetric factor, are obtained from the historical AERONET ground-based AOD measurements at the Beijing site.

Landsat series images provide satellite-received signals from visible to midinfrared channels and are recorded using Digital Numbers (DN). The radiometric calibration must be performed first to convert the DN values into TOA reflectance using reflectance rescaling coefficients provided in the product metadata file (Chander et al., 2009; Sun et al., 2015). The principle of aerosol retrieval is to separate the surface reflectance and TOA reflectance from the total satellite signals. Based on the atmospheric radiative transfer model, the satellite-received TOA reflectance is a function of atmospheric path reflectance (scattering of solar radiation within the atmosphere) and surface function (reflection of solar radiation from the surface that is directly transmitted to TOA reflectance) and can be estimated as follows:

$$\rho^*_{i,\theta,\phi} = \rho_{\text{Ray}}(i,\theta,\phi) + \rho_{\text{Aer}}(i,\theta,\phi) + \frac{T(\theta_i)T(\phi_i)\rho_{\text{Aer}}(i,\theta,\phi)}{1 - \rho_{\text{Ray}}(i,\theta,\phi)S_i}$$  \hspace{1cm} (1)

where $$\rho^*_{i,\theta,\phi}$$ is the TOA reflectance, $$\rho_{\text{Ray}}(i,\theta,\phi)$$ is the aerosol reflectance resulting from multiple scattering in the absence of molecules, $$\rho_{\text{Aer}}(i,\theta,\phi)$$ is the Rayleigh reflectance resulting from multiple scattering in the absence of aerosols, $$T(\theta_i)$$ and $$T(\phi_i)$$ are the transmission of the atmosphere along the sun surface path and the surface sensor path, respectively, $$\rho_{\text{Aer}}(i,\theta,\phi)$$ is the surface reflectance, $$S$$ is the atmospheric backscattering ratio, and $$\theta$$ and $$\phi$$ are the view zenith angle and relative azimuth angle, respectively.

The Rayleigh scattering has significant impacts on the visible channels, especially in the blue bands with short wavelengths. Rayleigh scattering correction depends on the determination of the Rayleigh phase function and the Rayleigh optical depth (ROD), and the ROD can be calculated using equation (2) (Bodhaine et al., 1999; Mishchenko & Travis, 1997):

$$\tau_{\text{Ray}}(i, Z) = 0.00877\lambda^{-4.05} \exp(-Z/8.5)$$  \hspace{1cm} (2)

where $$\tau_{\text{Ray}}$$ is the ROD, $$\lambda$$ is the wavelength (µm), 8.5 km is the exponential scale height of the atmosphere, and $$Z$$ is the elevation of the surface target obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (GDEM) at 30 m spatial resolution.

3.1. Surface Reflectance Estimation

To improve the estimation of surface reflectance over complex surfaces, extensive efforts have been made, and two typical extinction schemes are proposed for different land surface types according to their variation characteristics. Land surfaces are divided into two different categories: (1) Densely vegetated areas with high vegetation coverage, where the LSR can be estimated more readily, and (2) bright areas, which mainly include urban, desert, bare land, and arid/semiarid areas with sparse or little vegetation coverage; reflectance in these regions is more difficult to estimate accurately.

3.1.1. Densely Vegetated Areas

For densely vegetated areas, the surface reflectance shows obvious seasonal changes; considerable variation could be observed in the growing season, whereas long periods of stability could be observed in winter. Previous studies showed that the DT algorithm is accurate in densely vegetated areas, with nearly fixed ratios of surface reflectance between the visible and SWIR channels (Kaufman, Wald, et al., 1997). Such relationships have been further explored and updated to improve the estimation of LSR for MODIS data at a global scale (Levy, Remer, & Dubovik, 2007; Levy et al., 2013, 2010). Therefore, the DT algorithm is selected here for AOD retrieval in densely vegetated areas. Considering the differences in spectral channels and spectral response functions among different sensors, a new modified relationship between visible and SWIR channels for Landsat images is developed to estimate LSR over such areas.
In the second MODIS DT algorithm, the normalized difference vegetation index (NDVI) calculated from the SWIR wavelengths (NDVI_{SWIR}) of 1.24 and 2.12 μm is used to divide the land surface into three types of vegetated area. Pixels with values of NDVI_{SWIR} > 0.6 are considered dense vegetation covered areas, whereas NDVI_{SWIR} < 0.2 is representative of dormant or sparse vegetation (Levy et al., 2013). However, Landsat sensors do not include 1.24 nm; therefore, the regular NDVI, which is highly correlated with NDVI_{SWIR} in aerosol-free conditions (Levy et al., 2013), is selected instead. To improve the estimation of surface reflectance over land surfaces, the second enhanced DB algorithm divides the pixels into three NDVI classes, 0.10 < NDVI ≤ 0.20, 0.20 < NDVI ≤ 0.55, and NDVI > 0.55, representing areas with different vegetation coverage and then explores the relationships between the surface reflectance of visible channels and Rayleigh-corrected reflectance at 0.865 μm (Hsu et al., 2013). Based on previous studies, we define pixels with NDVI values greater than 0.55 (NDVI > 0.55) as densely vegetated areas. These can be extracted from natural vegetated land surfaces along with the Landsat Global Land cover mapping at 30 m spatial resolution (GlobeLand30) in 2010 provided by the National Geographic Information Center (Chen et al., 2014).

Considering the spectral differences among sensors, the relationships between the surface reflectance of visible and SWIR channels need to be reexplored to improve the aerosol retrieval accuracy over densely vegetated areas. Thus, Landsat series images from 2008 to 2016 that cover different AERONET sites are collected, and TOA reflectance from visible to SWIR channels and the corresponding AERONET ground-based AOD measurements are extracted. Then, these parameters are input into the 6S model to perform accurate atmospheric correction, and the actual surface reflectance from visible to SWIR channels is obtained (Vermote et al., 1997). Here a total of 6,985 effective points of surface reflectance between visible and SWIR channels is collected over densely vegetated areas.

Figure 2 plots the relationships of atmospherically corrected surface reflectance of blue (0.48 μm) and red (0.66 μm) channels versus that in the SWIR (2.22 μm) and their regression lines. It is apparent that there are overall low dispersions between them and that the surface reflectance of the visible channels exhibits strong linear regression relationships with that of the SWIR channel. For the blue channel, its surface reflectance agrees well with that of SWIR channel, with a high correlation of 0.847 (P < 0.001) and a small intercept of 0.002, and it is approximately 0.260 times smaller than that of the SWIR channel (ρ_{0.48} = 0.260 × \rho_{2.22} + 0.004); this result is close to that (0.259 times) of the MODIS second generated DT algorithm (Levy et al., 2013). A similar result is obtained for the red channel; its surface reflectance is approximately 0.531 times smaller than that of the SWIR channel (ρ_{0.66} = 0.531 × \rho_{2.22} - 0.002, P < 0.001), which is slightly smaller than that of the MODIS (0.55 times; Levy et al., 2013) second generated DT algorithm. The new relationships for the Landsat series data are updated in the DT algorithm and used to estimate the LSR for visible channels for aerosol retrieval over densely vegetated areas.

3.1.2. Bright Areas

The LSR is quite high and does not exhibit stable relationships between visible and SWIR channels over bright areas resulting in large difficulties in LSR estimation (Hsu et al., 2004; Sun et al., 2015). Aerosol retrieval over bright areas has been an issue of much interest and represents an important problem to be solved in this paper. However, we noticed that the LSR of bright surfaces does not significantly vary over time and that the effect of bidirectional reflectance distribution function is weaker for bright surfaces than for vegetated surfaces (Hsu et al., 2013).

Based on the assumption that the LSR of bright areas with sparse or little vegetation coverage changes little over a given period (Levy, Remer, & Dubovik, 2007; Sun et al., 2015; Sun, Wei, Wang, et al., 2016; Wei & Sun, 2017), a new approach for LSR estimation is proposed based on a precalculated monthly LSR database with Landsat 4–7 surface reflectance climate data records. The Landsat 4–7 LSR products with a low cloud amount of less than 30% and encompassing the entire period 2008–2016 are collected to construct the prior LSR database. To further remove the influence of cloud contamination, cloud, or mountain shadow in remote
sensing images, the second lowest surface reflectance for each pixel of all the images in a month is chosen to be the pixel for the 1 month series (equation (3))

\[
I(i,j) = \text{Second}_\text{Min}(I_1(i,j), I_2(i,j), ..., I_k(i,j))
\]  

(3)

where \( I \) represents the synthetic image; \( I_1, I_2, ..., I_k \) represent different Landsat 4–7 surface reflectance images in a given month; and \( i \) and \( j \) represent the row and column, respectively, in an image.

The LSR database is constructed based on the above method and contains 12 LSR images for each month at 30 m spatial resolution for the visible channels. Figure 3 shows four LSR images for the blue channel over the study area in January, April, July, and October. It is clear that the LSR images show overall high data quality with little cloud contamination and can better reflect the LSR variations over different underlying surfaces. It is apparent that LSR is overall higher in winter (January) than in summer (July), which is mainly due to the influence of decreasing vegetation and soil moisture as well as increasing snow cover during the winter (Franch et al., 2017; Wei & Sun, 2017; Zhu et al., 2015). LSR is obviously higher over the center of urban areas than over the surrounding areas that are covered by vegetation. However, the LSR values are generally low in most areas, with more than 99% of the pixels having LSR values less than 0.15; exceptions are a few pixels over extremely bare land. Thus, the LSR database is selected to provide accurate surface reflectance for AOD retrieval over bright areas.

Figure 3. Land surface reflectance images in (a) January, (b) April, (c) July, and (d) October over Beijing and surrounding areas.
3.2. Aerosol-Type Assumption

Aerosol type describes the chemical composition and aerosol optical properties in a certain region and is another key parameter that affected the accuracy of AOD retrieval. Due to the complex sources, the composition of aerosol models is constantly changing with time, position, and season in different areas. Therefore, obtaining a full understanding of the aerosol particles and the proper assumption of aerosol type can improve the accuracy of AOD retrieval from remote sensing data.

In this paper, we assume that the aerosol type does not vary spatially over the study area. The monthly averages of two main optical properties, single scattering albedo (SSA), and the asymmetry parameter (g), during 2008–2016, are collected from the ground-based Beijing site. Single scattering albedo is the sum of extinction caused by scattering and is used to characterize the absorption and scattering of electromagnetic radiation. The asymmetry parameter indicates the relative dominance of forward or backward scattering. Figure 4 shows the long-term temporal monthly variations of SSA and g at 550 nm, which is interpolated from ground-based measurements at 440, 500, and 675 nm using the Ångström exponent algorithm.

The results show that the SSA exhibits a significant increasing trend, increasing by 0.04% per year \( (P < 0.001) \). This finding indicates the continuous and decreasing absorption characteristics of aerosol type in the study area. Moreover, it exhibits the maximum value in the autumn and the minimum value in the spring, displaying obvious seasonal variation. However, the slope of SSA variation is small, less than 0.001 during the period 2008–2016, and the differences are small overall within the same season among years. The asymmetry parameter \( (g) \) shows an overall decreasing trend, decreasing by \( -0.01\% \) per year \( (P = 0.163) \), and it exhibits similar but smaller seasonal variation than does single scattering albedo. Thus, it is observed that the aerosol optical properties of aerosol type show overall small changes during a single season, and we assumed that the dominant aerosol type in the study area is a function of season and comprises fine mode aerosols with weak and moderate absorption (Levy et al., 2013, 2010). The seasonal averages of the single scattering albedos and asymmetry parameters are calculated from monthly average AERONET ground measurements from 2008 to 2016. Table 2 lists the aerosol optical properties of aerosol models in different seasons used in the AOD retrieval in this paper.

3.3. Pixel Selection and Aerosol Retrieval

The success of MODIS aerosol retrieval depends on its ability to throw out unsuitable pixels, including those of clouds, snow, and inland water bodies. Failure to fully remove these unsuitable pixels can lead to large errors and insufficient aerosol coverage (Levy et al., 2013). The recent proposed Fmask (Function of mask) for cloud and cloud shadow detection in Landsat imagery is selected for cloud, cloud shadow, snow, and inland water screening (Zhu, Wang, & Woodcock, 2015; Zhu & Woodcock, 2012). Based on Landsat TOA reflectance, brightness temperatures and cloud physical properties, Fmask first uses rules to separate Potential Cloud Pixels (PCPs) and clear sky pixels, and second, a normalized temperature probability, spectral variability probability and brightness probability are combined to generate the probability

![Figure 4](image)

**Figure 4.** Long-term variations of aerosol properties at 550 nm obtained from the monthly AERONET measurements at the Beijing site during the period 2008–2016.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Optical Properties of Different Aerosol Models Used in AOD Retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Season</td>
<td>Aerosol type</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring</td>
<td>Weak absorption</td>
</tr>
<tr>
<td>Summer</td>
<td>Weak absorption</td>
</tr>
<tr>
<td>Autumn</td>
<td>Moderate absorption</td>
</tr>
<tr>
<td>Winter</td>
<td>Moderate absorption</td>
</tr>
</tbody>
</table>
mask for clouds over land and water separately. Then, the potential cloud mask is derived using the PCPs and the cloud probability mask, and the cloud shadow is derived using the flood-fill transformation technology. The Fmask algorithm has made several improvements and expansions to cloud, cloud shadow and snow detection for Landsat 4–8 images, allowing more accurate results to be generated. The Fmask result is recorded as five values ranging from 0 to 5, which represent clear, cloud, cloud shadow, snow and water, respectively (Zhu, Wang, & Woodcock, 2015; Zhu & Woodcock, 2012). In this study, only clear pixels (value = 0) are retained for AOD retrieval.

In this paper, the preconstructed LUT is used to perform and improve the efficiency of AOD retrieval for blue and red channels using their own spectral response functions for Landsat TM, ETM+, and OLI satellites. LUT is built using the 6S model, which carefully considers aerosol scattering and gaseous (including H2O, O3, O2, and CO2) absorption, among other factors (Vermote et al., 1997). The 6S model has been widely used for LUT calculation for AOD retrieval from many sensors. Four custom aerosol types along with parameters such as solar and view zenith angles, relative azimuth angle, AOD, and surface reflectance are used as inputs in the 6S model to construct the precompiled LUTs for visible channels for Landsat series satellites (Sun et al., 2015). Figure 5 shows the flowchart of the SUHL aerosol retrieval algorithm for Landsat images.

### 3.4. Evaluation Source and Approaches

AERONET AOD measurements are selected to quantitatively evaluate the reliability of the AOD retrievals. AERONET does not provide AOD ground observations at 550 nm; thus, the AODs at 550 nm are calculated...
using the Ångström exponent algorithm with the measured AODs at 440, 500, and 675 nm wavelengths from AERONET to make direct comparisons with the AOD retrievals. AOD retrievals within a common sampling window of 5 × 5 pixels around the AERONET sites are obtained, and the average of effective values (\( \text{AOD} > 0 \)) is taken as the AOD retrieval at this site. Then, the average of at least two AERONET AOD measurements at each site within ±30 min of the Landsat satellites overpass times is calculated (Hsu et al., 2004, 2006; Kaufman, Tanre, et al., 1997; Levy, Remer, & Dubovik, 2007; Levy, Remer, et al., 2007).

The linear regression equation, mean absolute error (MAE), root-mean-square error (RMSE), error (equation (4)), relative mean bias (RMB, equation (5)), and MOD04 expected error (EE, equation (6)) are selected for evaluation purposes. The linear regression equation is selected to estimate the slope and intercept, and the correlation coefficient (\( R^2 \)) is a good indicator of the extent of agreement between the AOD retrievals and AERONET measurements. MAE and error are the most widely used natural measures of mean error magnitude. RMSE is selected to measure the differences between AOD retrievals and measurements and is sensitive to both systematic and random errors. RMB is used to analyze the average estimation uncertainty of AOD retrievals, where RMB > 1.0 or <1.0 indicates the overestimation or underestimation uncertainty, respectively, of the AOD retrievals. The EE for the DB product over land is selected to evaluate the accuracy (within EE, %) for different AOD retrievals (Hsu et al., 2013).

\[
\text{Error} = \text{AOD}_{\text{Retrieval}} - \text{AOD}_{\text{AERONET}} \quad (4)
\]

\[
\text{RMB} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\text{AOD}_{\text{Retrieval}}}{\text{AOD}_{\text{AERONET}}} \right| \quad (5)
\]

\[
\text{EE} = \pm (0.05 + 0.2 \times \text{AOD}_{\text{AERONET}}) \quad (6)
\]

Figure 6. Color composite images (red-blue-green) and corresponding spatial distribution of Landsat AOD retrievals (30 m) at 550 nm in the study area.
4. Results and Discussion

4.1. Spatial Distribution of AOD Retrievals

One hundred thirty-six images including Landsat 5 TM images from 2008 to 2012, Landsat 7 ETM+ images from 2010 to 2016, and Landsat 8 OLI images from 2013 to 2016 with a small cloud amount of less than 30% are collected to perform the AOD retrieval experiments. However, due to sustained variations of abundant and detailed surface information, noise is unavoidable in the retrievals. Therefore, before considering the spatial distribution of aerosols, a median filter with a kernel size of five is applied to reduce the effects of noise and generate more continuous spatial distributions of aerosols. Figure 6 shows the color composite images (Red-Blue-Green: 4-2-3) to show the degree of air pollution and AOD spatial distributions at 550 nm on 29 April 2014 and 19 August 2014 for Landsat 8 OLI at 30 m spatial resolution with different AOD loadings in the study area. It is observed that the new algorithm can achieve continuous AOD distributions in the study area over both dark and bright areas, even in the centers of urban areas with buildings and roads, which exhibit high surface reflectance and complex patterns. Moreover, the surrounding areas with vegetation coverage show overall low AOD values, whereas the urban areas have more severe air pollution and high AOD values. The SUHL algorithm can describe the aerosol distribution and variability in great detail at a high spatial resolution of 30 m under both low and high aerosol loadings.

4.2. Validation With AERONET AOD Measurements

The high-resolution (30 m) AOD retrievals from Landsat series images were first validated with AERONET AOD measurements from Beijing, Beijing_CAMS, Beijing_RADI, Beijing_PKU, and XiangHe sites. Figure 7 shows the validation of AOD retrievals against AERONET AODs from these five sites. The dotted lines represent the MOD04 DB EE lines, the black dashed lines are the 1:1 lines, and the red solid lines represent the regression lines. A total of 284 AOD collections is obtained, comprising 93 collections for the Beijing site, 90 collections for the XiangHe site, 51 collections for the Beijing_CAMS site, 44 collections for the Beijing_RADI site, and 6 collections for the Beijing_PKU site. It is apparent that the Landsat AOD retrievals agree well with AERONET AOD measurements, exhibiting a high correlation coefficient of 0.939 and small MAE and RMSE errors of 0.097 and 0.134 and entailing only 2.9% overestimation uncertainty (RMB = 1.029) at the Beijing site. Similar good agreement of the AOD retrievals is found for the Beijing_CAMS and Beijing_RADI sites. AOD retrievals for the Beijing_CAMS site show a high correlation (R² = 0.927) with AERONET AOD measurements and exhibit small MAE and RMSE errors of 0.094 and 0.120, with an average underestimation uncertainty of 14.1% (RMB = 0.859). For the Beijing_RADI site, there is high consistency between the AOD retrievals and AERONET AOD measurements (R² = 0.932) and low MAE and RMSE errors of 0.095 and 0.135, leading to approximately 1.3% overestimation uncertainty (RMB = 1.013). Due to short-term ground-based observations, only six pairs are collected for the Beijing_PKU site, which are highly consistent with the AERONET AODs (R² = 0.997, MAE = 0.098, and RMSE = 0.129). In general, 70.97%, 72.55%, 79.55%, and 100% of the AOD retrievals fall within the EE for the Beijing, Beijing_CAMS, Beijing_RADI, and Beijing_PKU sites, respectively, indicating that the SUHL algorithm can accurately retrieve AOD over bright urban surfaces. Moreover, the suburban XiangHe site shows the highest accuracy, with 84.44% of the AOD collections falling within the EE, and the AOD retrievals are well correlated with AERONET AOD measurements (R² = 0.940), with low MAE and RMSE errors of approximately 0.076 and 0.116, respectively. Figure 7f provides a scatterplot of all of the Landsat AOD collections and AERONET AOD measurements from all five sites over the Beijing region. The Landsat AOD retrievals agree well with the AERONET AOD measurements (R² = 0.932), and a total of 77.46% of them meet the acquirement of the EE, with overall low MAE and RMSE errors of 0.090 and 0.126, respectively. These results demonstrate that the SUHL algorithm is robust and can accurately retrieve AOD over both bright urban and dark vegetated surfaces. Moreover, 20.43%, 24.44%, 17.65%, 20.45%, and 33.33% of the collections had AOD values greater than 0.8 at the Beijing, XiangHe, Beijing_CAMS, Beijing_RADI, and Beijing_PKU sites, respectively. The results show that 93.33% of the AOD retrievals fall within the EE with low MAE and RMSE values of 0.0942 and 0.1410, indicating that the SUHL algorithm performs well overall on heavily polluted days.

4.3. Applicability for Different Landsat Sensors

To investigate the efficiency and stability of the SUHL aerosol retrieval algorithm for different Landsat satellites, AOD retrievals of Landsat TM, ETM+, and OLI images are extracted from all of the sites and validated.
against AERONET AOD measurements. Forty-four collections for Landsat TM, 155 collections for Landsat ETM+, and 85 collections for Landsat OLI from five AERONET sites are collected for the period 2008–2016.

Figure 8 shows the validation of AOD retrievals against AERONET AOD measurements and their errors (AODretrieval - AODAERONET) from all five sites for Landsat TM, ETM+, and OLI images.

The Landsat TM AOD retrievals agree well with AERONET AODs, with a high correlation coefficient of 0.956, and approximately 79.55% of the AOD collections fall within the EE, and with low overall MAE and RMSE.
values of 0.077 and 0.103, respectively, resulting in less than 5% estimation uncertainty (RMB = 0.955). Landsat ETM+ AOD retrievals show high consistency with AERONET AOD measurements (\(R^2 = 0.925\)), and 74.84% meet the acquirement of the EE, with average low MAE and RMSE errors of 0.095 and 0.128, respectively, causing approximately 5% underestimation uncertainty (RMB = 0.950). Moreover, Landsat OLI AOD retrievals are highly correlated to AERONET AOD measurements (\(R^2 = 0.948\)), and 81.18% of the collections fall within the EE, with an MAE of 0.085 and RMSE of 0.132, respectively, causing approximately 4.9% of the overestimation uncertainty. The error plot shows different discrepancies in retrieved AODs, varying from \(-0.256\) to 0.281, \(-0.226\) to 0.360, and \(-0.180\) to 0.498 for Landsat TM, ETM+, and OLI sensors, respectively, with average low errors of \(0.019\), \(0.001\), and \(0.044\), respectively. Moreover, approximately 77.27%, 64.52%, and 75.29% of the AOD retrievals are highly correlated with AERONET AOD measurements at very low discrepancies, with errors of less than 10%. The validation results show that the new proposed algorithm is robust and performs well overall for different Landsat series sensors.

### 4.4. Comparison With MODIS AOD Products

For monitoring of atmospheric particle pollution, NASA provides long time series of daily global aerosol products (MOD04) at a coarse spatial resolution of 10 km generated from both the dark target (DT) and deep blue (DB) aerosol retrieval algorithms. MOD04 products provide three typical data sets: DT retrievals, DB retrievals, and merged DT and DB (DT&DB) retrievals. A data set named quality assurance (QA) is provided and represents the quality of AOD retrievals with QA values ranging from 0 to 3 and indicating low to high accuracy (Levy et al., 2010, 2013).

Considering the similar overpass time between MODIS and Landsat satellites (AM 10:30 ± 30 min local solar time), the MODIS Terra C6 AOD retrievals are selected here for comparison purposes. Only highest-quality common retrievals for MOD04 DT retrievals (QA = 3), DB retrievals (QA ≥ 2), and DT&DB retrievals (QA = 3) at 550 nm as Landsat images are selected during 2008–2016. Figure 9 shows the comparison of SUHL AOD retrievals (30 m) and MOD04 AOD products (10 km) against AERONET AOD measurements.

A total of 140 collections for Landsat AODs and C6 DT AODs, 270 collections for Landsat AODs and C6 DB AODs, and 249 collections for Landsat AODs and C6 DT&DB AODs is obtained and validated with AERONET AOD measurements from five AERONET sites. It is found that MOD04 AOD retrievals show overall high
agreement with AERONET AOD measurements ($R^2 = 0.730$), yet only 52.86% of the collections fall within the 
EE with MAE of 0.185 and RMSE of 0.240, which led to seriously overestimation of the AOD loadings in the 
urban areas ($RMB = 1.161$). Compared with the MOD04 DT retrievals, the Landsat AOD retrievals achieve a 
higher correlation with AERONET AODs ($R^2 = 0.928$), a stronger slope of 1.090, a smaller positive offset of 
0.042, and lower MAE and RMSE values of 0.099 and 0.123, respectively, which significantly reduce the 
estimation uncertainty ($RMB = 1.006$). In general, 73.57% of the AOD collocations fall within the EE, which 
is approximately 1.4 times more than the corresponding percentage of MOD04 DT retrievals, mainly 
because the DT algorithm is unable to retrieve accurate AODs over urban surfaces.

The MOD04 DB algorithm has the ability to retrieve AOD over both dark and bright surfaces, thus providing a 
large number of collocations ($N = 270$), approximately 1.93 times more the number obtained with the DT

---

**Figure 9.** (a, c, and e) Comparisons of the AOD retrievals at 30 m resolution with (b) MOD04 C6 DT, (d) DB, and (f) DT&DB AOD products (10 km) with AERONET AODs in the study area from 2008 to 2016. The dotted lines represent EE lines, the black solid lines represent 1:1 lines, and the red solid lines represent regression lines.
algorithm in urban areas. The DB retrievals agree well with AERONET AOD measurements ($R^2 = 0.793$) and show higher accuracy, with 61.11% of the collections falling within the EE, and overall lower MAE and RMSE values of 0.129 and 0.216, respectively, and lower uncertainty than that observed with DT retrievals. However, the Landsat AOD retrievals agree more overall with AERONET AOD measurements ($R^2 = 0.934$) than do the DB retrievals, with 78.15% of the retrievals falling within the EE and lower MAE and RMSE values of 0.087 and 0.121, respectively. However, the Landsat AOD retrievals showed a similar and low underestimation uncertainty of approximately less than 3% in the study area.

The MOD04 DT&DB algorithm combines the advantages of the DT and DB algorithms. It provides more retrievals than does the DT algorithm and improves the overall accuracy above that of both the DT and DB algorithms. There is high consistency ($R^2 = 0.825$) between the DT&DB AODs and AERONET AOD measurements, and a total of 65.86% of the collections fall within the EE, with a mean MAE of 0.135 and an RMSE of 0.216. However, it still yields large overestimation uncertainty over the urban surfaces. Compared with the DT&DB retrievals, the Landsat AOD retrievals show a higher consistency with AERONET AOD measurements ($R^2 = 0.935$), and approximately 77.51% of them fall within the EE, with a lower MAE of 0.087 and an RMSE of 0.122. These comparison results show that overall, the SUHL algorithm performs better than the routine MOD04 AOD products over complex surfaces, mainly due to improvements in the estimation of surface reflectance and in the assumptions of aerosol models.

5. Conclusion

In this paper, a simple, universal, and high-resolution aerosol retrieval algorithm for Landsat series satellites (SUHL) is proposed. It presents several main improvements to cloud screening, the estimation of surface reflectance, and the selection of aerosol type. The Fmask (Function of mask) is selected to mask all of the unsuitable pixels including cloud, snow, shadow, and inland water. The land surface reflectance (LSR) is estimated using two typical surface reflectance schemes, densely vegetated and bright areas, where the LSR of visible channels is estimated via new explored relationships with the top of atmosphere (TOA) reflectance of the shortwave infrared (SWIR) channel and a precalculated LSR database constructed from Landsat 4–7 surface reflectance climate data records. The aerosol type is determined via the historical aerosol optical properties derived from the local Aerosol Robotic Network (AERONET) site. The SUHL algorithm has three assumptions: (1) the surfaces are Lambertian, (2) the LSR of most features remains unchanged in a month over bright surfaces, and (3) the aerosol types vary little spatially and temporally in a season over the study area. The Landsat AOD retrievals are validated against AERONET AOD measurements and compared with the current MODIS Collection (C) 6 aerosol products (MOD04). Validation and comparison results show the following:

1. Landsat AOD retrievals can provide continuous and detailed AOD distributions at 30 m spatial resolution over Beijing, which features complex patterns with high land surface reflectance.
2. Landsat AOD retrievals agree well with AERONET AOD measurements ($R^2 = 0.932$), and 77.46% of them fall within the EEs, with low MAE and RMSE errors of 0.090 and 0.126, indicating high stability. The SUHL algorithm can accurately retrieve AODs over both bright urban and dark vegetated surfaces, even under high-pollution episodes.
3. In total, 79.55%, 74.84%, and 81.18% of the Landsat TM, ETM+, and OLI AOD retrievals, respectively, fall within the EE, with more than 60% of the retrievals showing errors $(\text{AOD}_{\text{retrieval}} - \text{AOD}_{\text{AERONET}}) < 10\%$, indicating that the new algorithm performs well overall for aerosol retrieval for different Landsat series sensors.
4. Comparison with routine MOD04 aerosol products showed that Landsat AOD retrievals at 30 m spatial resolution are more consistent with AERONET AOD measurements and exhibit less bias than do the MOD04 DT, DB, and DT&DB AOD products, mainly due to improvements in the estimation of surface reflectance and the assumption of aerosol types.

However, although the AOD retrievals obtained from Landsat series images have good accuracy overall, noise is unavoidable in the spatial distribution due to constant changes in land surfaces. Therefore, the signal-to-noise ratio (SNR) is a nonignorable factor affecting aerosol retrieval, and we will focus on investigating and discussing the influence of SNR on aerosols in our future work.
Acknowledgments
Landsat TM, ETM+, OLI images and Landsat 4–7 surface reflectance climate data records were obtained from https://earthexplorer.usgs.gov/. MODIS AOD products are available from the Goddard Space Flight Center Level 1 Atmosphere Archive and Distribution System (https://ladsweb.nascom.nasa.gov/). AERONET measurements are available from the Earth Resources Observation and Science Center (http://glovis.usgs.gov/). The work was supported by the Hong Kong Research Grants Council General Research Fund (14652016), National Key R & D Program of China (2017YFB0503605), National Natural Science Foundation of Shandong Province (2017YFF0307901), and National Natural Science Foundation of China (41271213, 41573087, 41675029). We thank all authors who helped review the paper. All authors made substantial contributions to this work. JING WEI designed the experiments and wrote the paper, and Bo Huang, Lin Sun, Zhaoyang Zhang, Lunche Wang, and Mohammad Bilal helped review the paper. All authors declare that no conflicts of interest exist.

References


