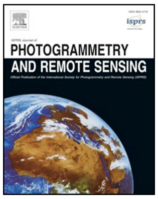
ISPRS Journal of Photogrammetry and Remote Sensing 157 (2019) 1–12

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs



Large-scale MODIS AOD products recovery: Spatial-temporal hybrid fusion considering aerosol variation mitigation

Yuan Wang[a], Qiangqiang Yuan[a,e,f,*], Tongwen Li[b,*], Huanfeng Shen[b,d,f], Li Zheng[a], Liangpei Zhang[c,f]

[a] School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei 430079, China
[b] School of Resource and Environmental Sciences, Wuhan University, Wuhan, Hubei 430079, China
[c] The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei 430079, China
[d] The Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, Wuhan, Hubei 430079, China
[e] The Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan University, Wuhan, Hubei 430079, China
[f] The Collaborative Innovation Center for Geospatial Technology, Wuhan, Hubei 430079, China

ARTICLE INFO

Keywords:
Large scale
MODIS
AOD recovery
Spatial-temporal hybrid fusion
Aerosol variation mitigation

ABSTRACT

Aerosol optical depth (AOD) is a pivotal parameter to reflect aerosol properties, such as aerosol radiative forcing and atmospheric corrections of the aerosol effect. Unfortunately, the valid pixels of moderate resolution imaging spectroradiometer (MODIS) AOD products are scarce, which has attracted great attention from scholars. In recent years, numerous AOD recovering algorithms have been proposed and the algorithms merely employing a single temporal AOD image are regarded as the most convenient and flexible for large-scale practical applications. However, current algorithms face the challenge of insufficiently considering the impacts of aerosol variation resulted from the temporal difference. Meanwhile, the improvement of AOD valid pixels is also poor due to the scarce excavation of complementary information. In order to address these issues, a novel algorithm of spatial-temporal hybrid fusion considering aerosol variation mitigation (ST-AVM) is developed to fill the missing pixels in Aqua AOD products with a single Terra AOD image in large scale. The results show that the total recovered AOD products nearly maintain the original accuracy of MODIS. Meanwhile, the AOD coverage is significantly improved in the study areas and the degrees of improvements regionally vary. Overall, the AOD coverage over land is increased by 123.9% (from 20.5% to 45.9%) after the recovery. Besides, the spatial distribution of recovered monthly AOD products remains fairly consistent as the original Aqua. Also, the recovered annual AOD spatial distribution shows more coherent, which indicates the reliability of ST-AVM algorithm.

1. Introduction

Atmospheric aerosols are liquid and solid particles suspending in the air (Li et al., 2014), which can affect the climate directly and indirectly (Che et al., 2018). The aerosols pollution that originates from natural and anthropogenic sources has attracted great attention from researchers (Emetere et al., 2019; Hou et al., 2019; Ma et al., 2019; Volkamer et al., 2006). Generally, one of the key parameters, i.e., aerosol optical depth (AOD), which is defined as the vertical integral of light extinction by aerosol in the atmospheric column (Van Donkelaar et al., 2010) is employed to depict aerosol optical properties (Della Ceca et al., 2018). Considering that ground-based measurements are incapable of providing a global perspective, different sensors on onboard remote satellites have been used to retrieve AOD products, such as MODIS (Kaufman et al., 2005), MISR (Kahn et al., 2010), OMI (Ahn et al., 2014), and so on (Sayer et al., 2012; Zhang et al., 2019). With 36 spectral channels, a temporal resolution of 1–2 days, and spatial resolutions of 250 m, 500 m, and 1000 m (Bisht et al., 2005; Justice et al., 1998; King et al., 1992), moderate resolution imaging spectroradiometer (MODIS) can generate daily AOD products based on the dark target (DT) algorithm (Gupta et al., 2016; Levy et al., 2013) and the deep blue (DB) algorithm (Hsu et al., 2013; Wang et al., 2019). AOD products from MODIS are the most widely used among scholars and performs well at both local and global scales. While intrinsic drawbacks still exist in the MODIS AOD algorithms and cause the data missing of AOD products. For example, the DT algorithm fails to retrieve AOD

---

* Corresponding authors at: School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei 430079, China (Q. Yuan). School of Resource and Environmental Sciences, Wuhan University, Wuhan, Hubei 430079, China (T. Li).

E-mail addresses: whuwy@yahoo.com (Y. Wang), yqiang86@gmail.com (Q. Yuan), litw@whu.edu.cn (T. Li), shenhf@whu.edu.cn (H. Shen), lzheng@sgg.whu.edu.cn (L. Zheng), zlp62@whu.edu.cn (L. Zhang).

https://doi.org/10.1016/j.isprsjprs.2019.08.017
Received 14 May 2019; Received in revised form 25 July 2019; Accepted 23 August 2019

values over bright areas where the surface reflectance is high (Gupta et al., 2016; Levy et al., 2013); The DB algorithm is unable to acquire AOD values in the regions contaminated with cloud, snow, and ice (Hsu et al., 2004, 2013; Wang et al., 2019). Meanwhile, the width of MODIS swath (∼2330 km) is not enough to completely cover the areas with low latitude (Sayer et al., 2015), which results in the scanning gaps of AOD products between two adjacent satellite orbits. It has been manifested that the annual completeness of AOD coverage in DT and DB algorithm is poor (Jinnagara Puttaswamy et al., 2014; Tang et al., 2016; Xu et al., 2015; Yang and Hu, 2018; Zhang et al., 2017), indicating that the MODIS AOD products are not available for most days of the year. Therefore, it is necessary and meaningful to recover the missing pixels of MODIS AOD products to enhance the availability of data and obtain spatially continuous AOD products.

To date, the algorithms of AOD recovery have been greatly developed. For instance, Zhang et al. (2018b), Zhao et al. (2019) utilized statistic and machine learning algorithms to train the models estimating AOD values, which seek the relationships among AOD retrievals from satellites and other factors (e.g., meteorological parameters). Nevertheless, the quality of statistical based algorithms is usually unsatisfactory due to the scarce accuracy of under-fitting or over-fitting models. Meanwhile, the universality of such end-to-end models is also poor. Except for the above algorithms, various merging algorithms have been proposed and most of them depend on the auxiliary AOD products from datasets of other satellites (e.g., SeaWiFS and MISR) (Jinnagara Puttaswamy et al., 2014; Tang et al., 2016; Xu et al., 2015) for a decade. However, these merging algorithms usually need multiple auxiliary data and the uncertainties of different AOD products are difficult to be unified. As a result, a few new algorithms, which consider the contiguous temporal knowledge of the original AOD products themselves, have been developed, including weighted linear regression with normalized difference vegetation index (NDVI) (Zhang et al., 2017) and spatiotemporal kriging (Yang and Hu, 2018). Generally, the recovering algorithms only based on a single temporal AOD image are regarded as the most convenient and flexible for practical applications (e.g., particulate matter estimation and air quality monitoring). For MODIS, the interval of overpass time between Terra (∼10:30) and Aqua (∼13:30) is approximately 3 h (Sayer et al., 2015), suggesting a good correlation. Filling the missing pixels in AOD products from Aqua with the help of AOD from Terra on the same day is reasonable. However, the variation of aerosol particles is non-ignorable even in a short time (see red marks in Fig. 1), especially for large areas in which the diverse aerosol models distribute. Although the algorithms based on temporal information of a single auxiliary AOD image possess great potential, yet they face the challenge of insufficiently considering the impacts of aerosol variation. Meanwhile, the improvement of AOD coverage is poor due to the scarce excavation of complementary information in a single temporal AOD image. Besides, the amount of these algorithms is still small and linear function is commonly used in large-scale applications at present (He et al., 2019; Li et al., 2017a, 2017b). Hence, it is required to propose a brand new algorithm to address the issues of how to accurately and maximally fill the missing pixels with scarce complementary information on a large scale.

The purpose of this paper aims at developing a novel algorithm of spatial-temporal hybrid fusion considering aerosol variation mitigation (ST-AVM) to fill the missing pixels in Aqua AOD products with a single Terra AOD image in large scale. To be specific, the proposed algorithm is pixel-based on the active windows of Terra and Aqua in the iterative framework. A total of three approaches, i.e., spatial, temporal, and spatiotemporal, are jointly employed to fill the missing AOD data as much as possible under the most suitable circumstances. Besides, aerosol variation mitigation is designed to reduce the effects of aerosol variation resulted from the temporal difference. NDVI is also considered in our algorithm to provide land cover information due to the association with AOD values (Guo et al., 2012).

The rest of this paper is structured as follows. Section 2 introduces the study areas (53–137°E, 3–57°N) and datasets (from satellites and ground-based sites) in our study. Section 3 presents the methodology. Spatial, temporal, and spatiotemporal approaches are organically merged in the iterative framework. The experiment results and discussion are given in Section 4, which are divided into three parts: assessment of AOD accuracy, coverage, and spatial distribution. Finally, Section 5 provides the conclusion.

2. Study areas and datasets

2.1. Study areas

In our study, an area with total coverage of ∼453600 sq·km. located in Asia (53–137°E, 3–57°N) is selected, including China, India, Kazakhstan, Mongolia and so on. The study areas contain some typically populated regions (e.g., China, India, and Pakistan) and numerous related researches in the past worked on them (Guo et al., 2012; He et al., 2019; Wang et al., 2019). Hence, spatiotemporally continuous and accurate AOD products are required in these areas. Meanwhile, the land cover types of the study areas are complex and diverse (see the legends in Fig. 2), which may help examine the universality of ST-AVM algorithm.

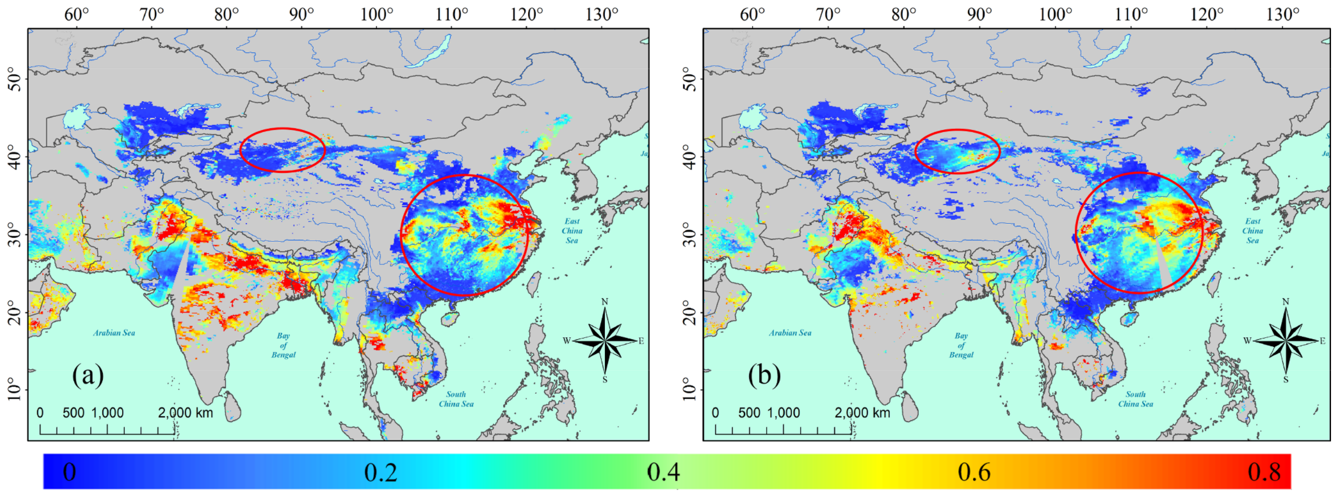


Fig. 1. Maps of daily AOD distribution for (a) original Terra AOD and (b) original Aqua AOD on 29th February 2016. The color bar represents the AOD values. The red marks highlight some regions with different AOD values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

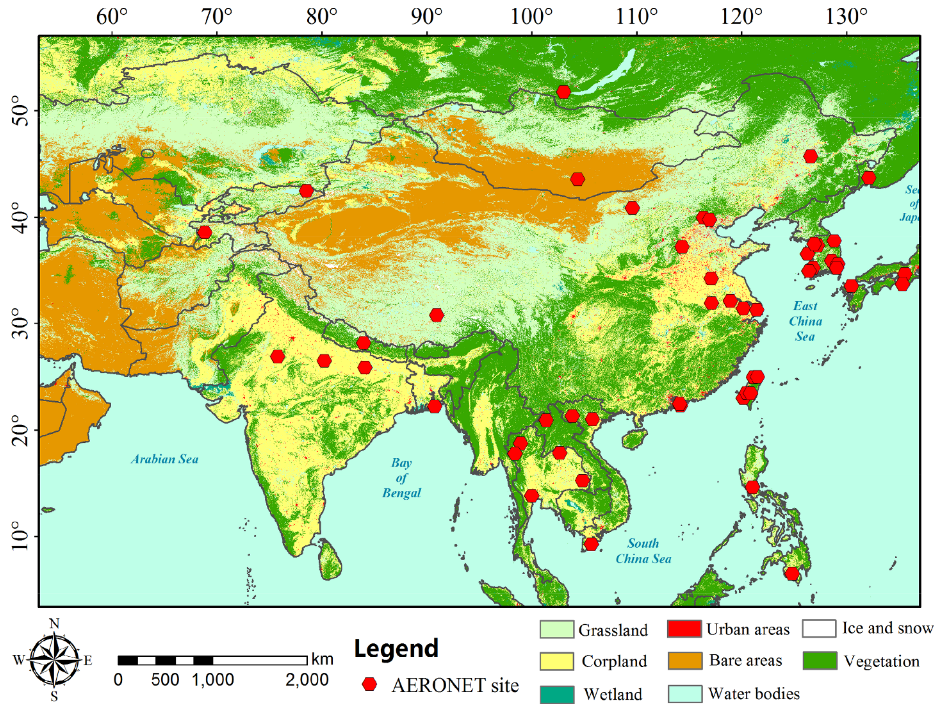


Fig. 2. Distribution of the AERONET sites considered in this paper. The detailed information of each site can be found in Table 1.

2.2. Datasets

In this paper, DB products with a 10-km spatial resolution in collection 6.1 from Terra and Aqua are utilized. In addition, the NDVI auxiliary products (MOD13A1) are also employed in our algorithm. The recovered AOD products are validated using Version 3.0 Level 2.0 AEROsol Robotic NETwork (AERONET) AOD products from 51 ground sites in Asia.

2.2.1. AERONET AOD products

The AERONET is a global ground-based network of sun photometers, which provides spectral AOD measurements with a high temporal resolution (~15 min) in the bands of 0.34–1.06 μm (https://aeronet.gsfc.nasa.gov). Datasets with three quality levels: Level 1.0 (unscreened), Level 1.5 (cloud-screened), and Level 2.0 (cloud-screened and quality-assured) are available at the AERONET website. The Version 3.0 AERONET AOD datasets of Level 1.5 and Level 2.0 with low uncertainty (Giles et al., 2019) are seen as the ground truth to validate the recovered AOD products in our study. A total of 51 AERONET sites in Asia with multiple land cover types are listed in Table 1.

2.2.2. MODIS DB AOD products

At first, the DB algorithm was devised to retrieve aerosol properties over bright areas, where the surface reflectance is usually high in the red and near-infrared bands. On the contrary, surface reflectance in these places is much darker in the band at 0.5 μm (Hsu et al., 2004). Specifically, DB retrieves AOD values with 1-km spatial resolution employing the surface reflectance databases at three bands (0.412, 0.47, and 0.65 μm), and then aggregates pixels to 10 km. After ten years, an enhanced algorithm which integrates the precalculated databases and NDVI to obtain the surface reflectance has been developed (Hsu et al., 2013). In conclusion, the AOD coverage from the collection 6 DB algorithm has been expanded to more land areas except for the regions of snow and ice. In the latest collection 6.1, a total of three new improvements have been introduced in the DB algorithm (Wang et al., 2019). In our study, the record named "Deep_Blue_Aerosol_Optical_Depth_550_Land_Best_Estimate" from

Table 1
Detailed information about the AERONET sites considered in this paper. Lat: latitude; Lon: longitude.

AEROENT	Lat/Lon (°)	AEROENT	Lat/Lon (°)	AEROENT	Lat/Lon (°)
Anmyon	36.54/126.33	Jaipur	26.91/75.81	Taipei_CWB	25.01/121.54
Bac_Lieu	9.28/105.73	Kanpur	26.51/80.23	Ubon_Ratchathani	15.25/104.87
Beijing	39.98/116.38	KORUS_Kyungpook_NU	35.89/128.61	Ussuriysk	43.70/132.16
Bhola	22.23/90.76	KORUS_Mokpo_NU	34.91/126.44	XiangHe	39.75/116.96
Chen-Kung_Univ	22.99/120.20	KORUS_UNIST_Ulsan	35.58/129.19	ND_Marbel_Univ	6.50/124.84
Chiang_Mai_Met_Sta	18.77/98.97	Luang_Namtha	20.93/101.42	Lulin	23.47/120.87
Chiayi	23.50/120.50	NGHIA_DO	21.05/105.80	Manila_Observatory	14.64/121.08
Dalanzadgad	43.58/104.42	Nong_Khai	17.88/102.72	Hong_Kong_PolyU	22.30/114.18
Dushanbe	38.55/68.86	Omkoi	17.80/98.43	Hong_Kong_Sheung	22.48/114.12
EPA-NCU	24.97/121.19	Osaka	34.65/135.59	AOE_Baotou	40.85/109.63
Fukuoka	33.52/130.48	Pokhara	28.19/83.98	SONET_Shanghai	31.28/121.48
Gandhi_College	25.87/84.13	Pusan_NU	35.24/129.08	SONET_Nanjing	32.12/118.96
Gangneung_WNU	37.77/128.87	Seoul_SNU	37.46/126.95	SONET_Hefei	31.91/117.16
Gwangju_GIST	35.23/126.84	Shirahama	33.69/135.36	SONET_Harbin	45.71/126.61
Hankuk_UFS	37.34/127.27	Silpakorn_Univ	13.82/100.04	XuZhou-CUMT	34.22/117.14
IAOCA-KRSU	42.46/78.53	Son_La	21.33/103.91	SONET_Xingtai	37.18/114.36
Irkutsk	51.80/103.09	Taihu	31.42/120.22	NAM_CO	30.77/90.96

Table 2
Detailed information about some concepts used in the proposed ST-AVM algorithm.

Name	Description
P_A	The missing pixel in Aqua
P_T	The pixel in Terra, which owns the same grid coordinates of the missing pixel in Aqua
active window in Aqua (AW_A)	The pixel-based active window centered on P_A of Aqua, which represents the surrounding AOD spatial structure around P_A
matching window in Terra (MW)	The pixel-based active window in the matching range of Terra, which is utilized to calculate the similarity with AW_A
active window in Terra (AW_T)	The pixel-based active window centered on the proper pixel of Terra, which owns the best similarity with AW_A

daily AOD products (MOD04_L2 and MYD04_L2) during 2016.01.01–2016.12.31 is considered as the appropriate data (Hsu et al., 2013).

2.2.3. MODIS NDVI products

MOD13A1 products can provide every 16 days at a 500-m spatial resolution as a gridded global NDVI in the Sinusoidal projection. They are computed from atmospherically corrected bidirectional surface reflectance, which has been masked for cloud, cloud shadow, heavy aerosol, and water. The NDVI products are usually used for monitoring of vegetation conditions, displaying the surface properties and land cover changes. In our study, we select the record: "500m_16_days_NDVI" during 2016.01.01–2016.12.31 as the auxiliary data. Meanwhile, the auxiliary data is aggregated to the 10-km spatial resolution by averaging the NDVI in the windows of 20 * 20 pixels.

3. Methodology

Some concepts employed in ST-AVM algorithm are listed in Table 2. The proposed algorithm is pixel-based on the active windows (AW_A, M_W, and AW_T), which are centered on each missing pixel of Aqua and on the proper pixel of Terra. The dates of the input AOD products from Terra and Aqua are in the one-to-one correspondence. As depicted in the flowchart (Fig. 3), the iterative framework is introduced in our algorithm. Firstly, the data processing, i.e., AW_A selection, is performed to acquire the active window on P_A, characterizing the robust spatial information of AOD round missing pixel. Next, temporal approaches, which depend on the temporal information from Terra, are utilized on the condition that AW_A doesn't exist. Otherwise, aerosol variation mitigation is adopted to search the proper MW (AW_T) on P_T compared with AW_A, reducing the effects of aerosol variation resulted from the temporal difference. If the temporal information from AW_T of Terra is still unreliable, only spatial approaches leaning on spatial information from Aqua could be used for recovery. On the contrary, spatiotemporal approaches (simultaneously merging the spatial information from Aqua and temporal information from Terra) is implemented to attain more dependable recovered result provided that the similarity between AW_A and AW_T is strong. A total of three approaches are jointly included in our algorithm in order to ensure the missing AOD data is filled as much as possible. In the end, the increment of AOD coverage is regarded as the criteria (less than 5%) during each iteration. The specific procedures of ST-AVM algorithm are described as follows.

3.1. Data processing (AW_A selection)

The spatial adjacent knowledge around P_A is the basis for recovering processing. The strategy of active windows to represent spatial properties was widely considered in the reconstruction task for remote sensing image processing (Xu et al., 2019; Zeng et al., 2018; Zhang et al., 2018a). Therefore, the selection of active window centered on P_A, i.e., AW_A, is adopted and includes two steps: "self-adaptive window selection" and "edge window removal" in our study.

3.1.1. Self-adaptive window selection

The valid AOD pixels in the active windows used to reflect the surrounding AOD spatial structure of the missing pixel is important. Usually, the windows with the predetermined length were widely considered, which include the valid AOD pixels of a fixed number. However, this way appears insufficient flexibility and ignores the spatial characteristics of AOD distribution. For each P_A, less AOD difference among the surrounding valid pixels is desired, which contains robust spatial information. Generally, the standard deviation is a measurement of the degree for statistical distribution, as shown in Eq. (1).

$$\sigma(D) = \sqrt{\frac{1}{N(D)-1}\sum_{i=1}^{N(D)-1}(\tau_{Aqua}(D_i) - \bar{\tau}_{Aqua}(D))^2} \tag{1}$$

where σ represents the AOD standard deviation of the active window; N represents the number of valid AOD pixels of the active window in Aqua; D represents the radius of the active window in Aqua; τ_{Aqua} represents the valid AOD values of the active window in Aqua; The overbar represents the averaged valid AOD values of the active window in Aqua. Therefore, the windows based on the local minimum AOD standard deviation are adopted as the most proper active windows of Aqua. In our algorithm, the input radiuses of active windows vary from 2 to 7 with the increment of 1, which is seen as a process of the self-adaptive selection.

3.1.2. Edge window removal

If the AOD pixels of the spatial reference are illustrated as edge windows (see Fig. 4(a)), where the P_A distributes at the edge of valid pixels, the recovered result tends to be unreliable due to the significant lack of AOD spatial continuous information. On the contrary, when the AOD pixels distribution is similar to Fig. 4(b), where the valid pixels mostly surround the P_A, the reliability of the recovered result would be much more increased. In conclusion, the edge windows ought to be removed and the P_A of them would not be filled. Considering the fact that the process of edge window removal is complicated, we have appended the detailed flowchart and several examples in the supplementary materials to make it easier for understanding.

3.2. Temporal approaches (NDVI-based local regression)

AW_A is able to reflect the adjacent AOD spatial structure around P_A. When the spatial information of AW_A is unstable (i.e., The number of valid AOD pixels around P_A is very small or zero in AW_A; AW_A is seen as an edge window), spatial approaches are inadvisable. Meanwhile, we also fail to trace the variation of aerosol particles due to lack of the valid spatial properties around P_A, suggesting that only the temporal information around P_T could be acquired. In other words, if none of AW_A could be discovered on P_A, NDVI-based local regression is utilized as temporal approaches for recovery in our study.

The linear function is suitable for practical applications due to the short time consumption compared to complex models. Generally, the accuracy of linear function may not meet the requirements of large-scale AOD recovery, only considering the holistic AOD information instead of local distinctions. On the contrary, the linear function performed on a local scale would improve the reliability of recovered AOD results. Besides, the NDVI values, which could display the surface properties, are associated with AOD values (He et al., 2019; Li et al., 2014). The direct introducing of NDVI in the regression model would

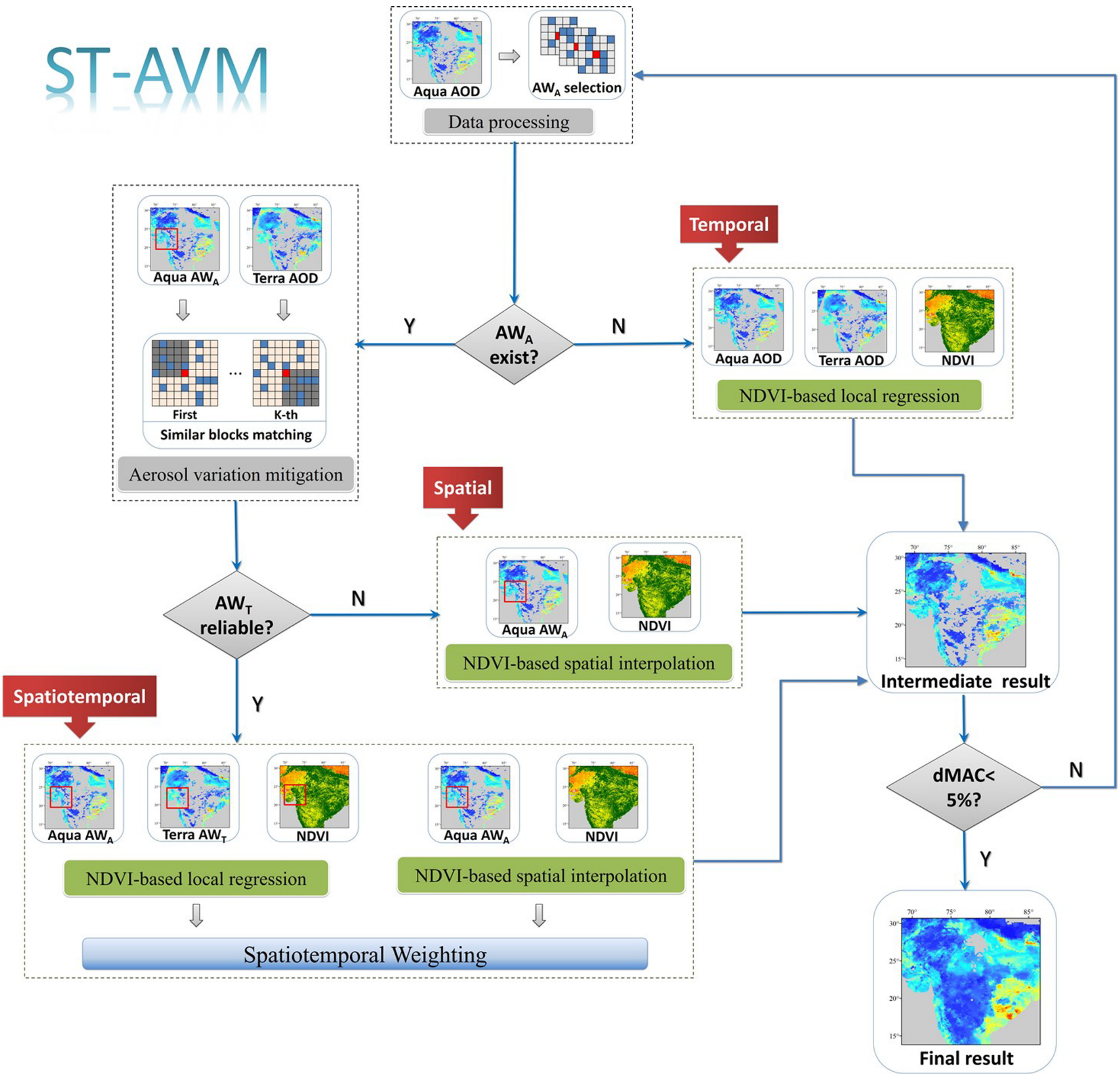


Fig. 3. Flowchart of the proposed ST-AVM algorithm. The red squares represent the active windows in Aqua and Terra. dMAC represents the percentage of AOD coverage increment after each iteration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

enhance the robustness for diverse land cover types. Specifically, as depicted in Fig. 5, the input window (centered on P_A) and the output window (centered on P_T) are primarily fixed, which include enough valid AOD pixels. Next, the active window in NDVI with the same coordinates of the output window is selected. Eventually, the values of the jointly valid pixels in the three windows are employed to fit the relationship with the binary linear regression, as shown in Eq. (2).

$$\tau_{Aqua} = \mathrm{a}\cdot\tau_{Terra} + \mathrm{b}\cdot NDVI_{Aqua} + \mathrm{c} \tag{2}$$

where τ_{Aqua}, τ_{Terra}, and $NDVI_{Aqua}$ denote the jointly valid values in the active windows of Aqua, Terra, and NDVI, respectively; a, b, and c are coefficients. Therefore, the fitted functions are different everywhere with the grid coordinate varying, sufficiently including the local information of AOD distribution and surface properties. It's worth noting that the AOD value of the center pixel in the input window will be filled with spatial interpolation provided that the AOD value on P_T is invalid.

3.3. Aerosol variation mitigation

As mentioned above, the variation of aerosol particles is non-ignorable and consequently it's required to reduce the effects of aerosol variation resulted from the temporal difference. In other words, the AOD value of P_A may be derived from surrounding pixels in the temporal data, instead of the same grid coordinate. Hence, aerosol

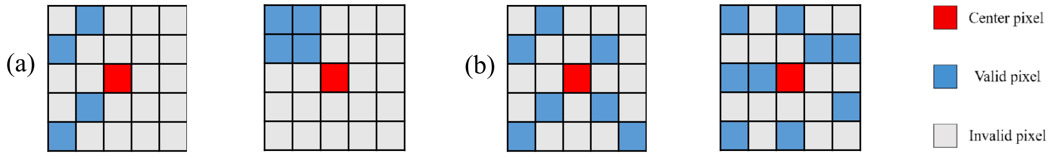


Fig. 4. Examples of (a) edge windows and (b) desired windows.

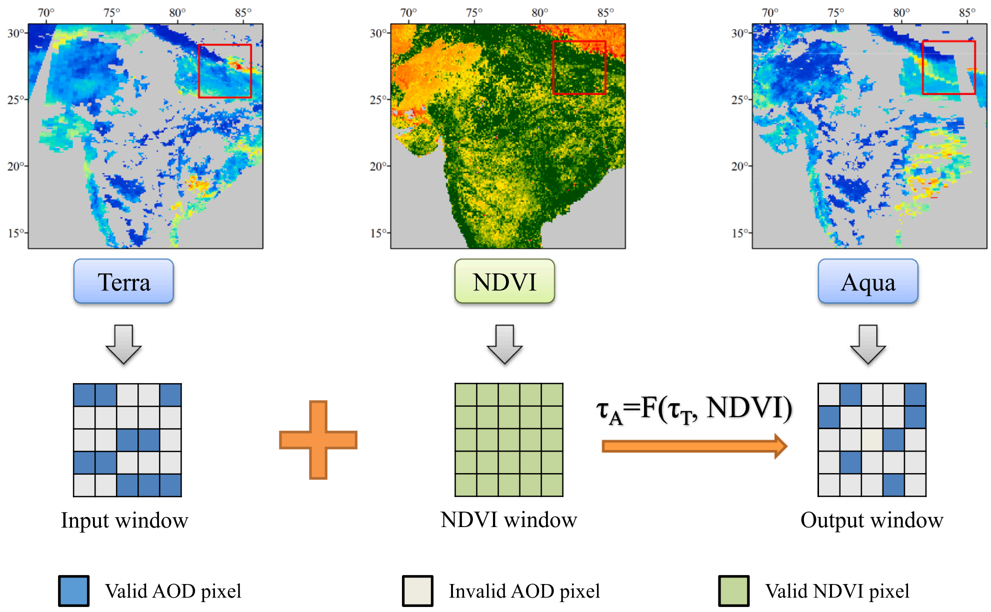


Fig. 5. Schematic diagram of NDVI-based local regression. The red squares represent the input window, NDVI window, and output window. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

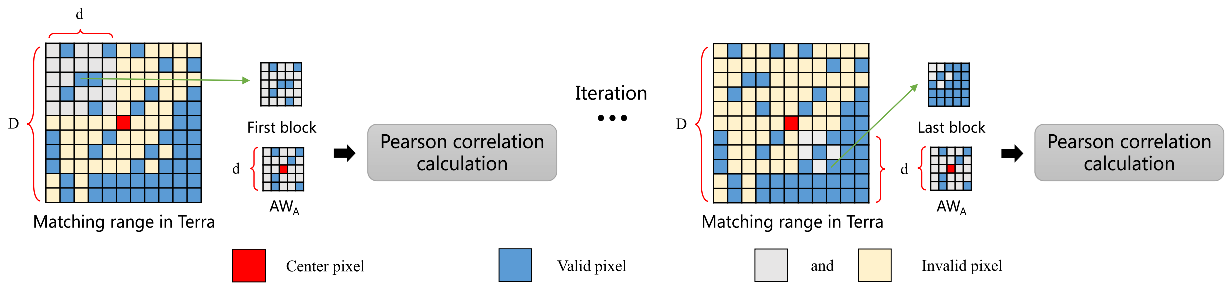


Fig. 6. Schematic diagram of aerosol variation mitigation.

variation mitigation is performed to determine the most similar MW (AW_T) on the condition that AW_A exists on P_A. In our algorithm, similar blocks matching is regarded as the approach for variation mitigation. To be specific, as illustrated in Fig. 6, the matching range centered on P_T in Terra with a diameter of D is fixed. Pearson correlation coefficient is a statistical parameter used to reflect the degree of the relationship between two variables, as shown in Eq. (3).

$$r(\tau_{Aqua}, \tau_{Terra}) = \frac{\sum_{j=1}^{n} (\tau_{Aqua\,j} - \bar{\tau}_{Aqua})(\tau_{Terra\,j} - \bar{\tau}_{Terra})}{\sqrt{\sum_{j=1}^{n} (\tau_{Aqua\,j} - \bar{\tau}_{Aqua})^2}\sqrt{\sum_{j=1}^{n} (\tau_{Terra\,j} - \bar{\tau}_{Terra})^2}} \tag{3}$$

where r represents the Pearson correlation coefficient; n represents the number of jointly valid AOD pixels in the active windows of Terra and Aqua; τ_{Aqua} and τ_{Terra} represent the jointly valid AOD values in the active windows of Aqua and Terra, respectively; The overbar represents the averaged valid AOD values in the active windows of Terra and Aqua. For each pixel in the matching range, we select the block of the same size as AW_A and calculate the Pearson correlation coefficient between the two windows. Eventually, the block with the highest correlation coefficient is considered as AW_T of Terra.

3.4. Spatial approaches (NDVI-based spatial interpolation)

After mitigation of the effect from aerosol variation, AW_A and AW_T are supposed to show the best similarity. Nevertheless, the temporal information could be invalid in case of significant aerosol variation in the matching range. Under the circumstances, spatial approaches will be used for recovery on the condition that the similarity between AW_A and AW_T is poor (Pearson correlation coefficient less than 0.8).

Inverse distance weighted (IDW) is a simple spatial interpolation approach, which takes the square distance between target points and sample points as the weight. However, the distribution of aerosol particles is complicated and consequently the square distance may be not enough to reflect the local AOD variation. Similarly, the knowledge of land surface from NDVI products could be employed to enhance the reliability of spatial interpolation. As a result, the NDVI-based IDW is adopted for spatial interpolation in our algorithm. The weight is defined as Eq. (4).

$$\begin{aligned} a(i, j, c, k) &= ((x_i - x_c)^2 + (y_i - y_c)^2 + \varepsilon_1)\cdot\sqrt{|NDVI_{i,j}^2 - NDVI_{c,k}^2| + \varepsilon_2} \\ b(i, j) &= \sum_{c=i-r}^{i+r} \sum_{k=j-r}^{j+r} {}^{1}\!/_{a(i,j,c,k)} \\ W(c, k) &= \frac{{}^{1}\!/_{a(i,j,c,k)}}{b(i,j)} \\ \varepsilon_1 &= 10^{-5},\ \varepsilon_2 = 10^{-3} \end{aligned} \tag{4}$$

where (x_i, y_i) and $NDVI_{i,j}$ represent the grid coordinate and NDVI of the missing pixel, respectively; (x_c, y_k) and $NDVI_{c,k}$ represent the grid coordinate and NDVI of valid pixels in the interpolation window, respectively; W represents the weight of valid pixels in the interpolation window; r represents the radius of the interpolation window. ε_1 and ε_2 are small values that prevent $a(i, j, c, k)$ from equaling zero, which are empirically determined. As a result, $((x_i - x_c)^2 + (y_i - y_c)^2 + \varepsilon_1)$ describes the spatial distinction and $\sqrt{|NDVI_{i,j}^2 - NDVI_{c,k}^2| + \varepsilon_2}$ reflects the

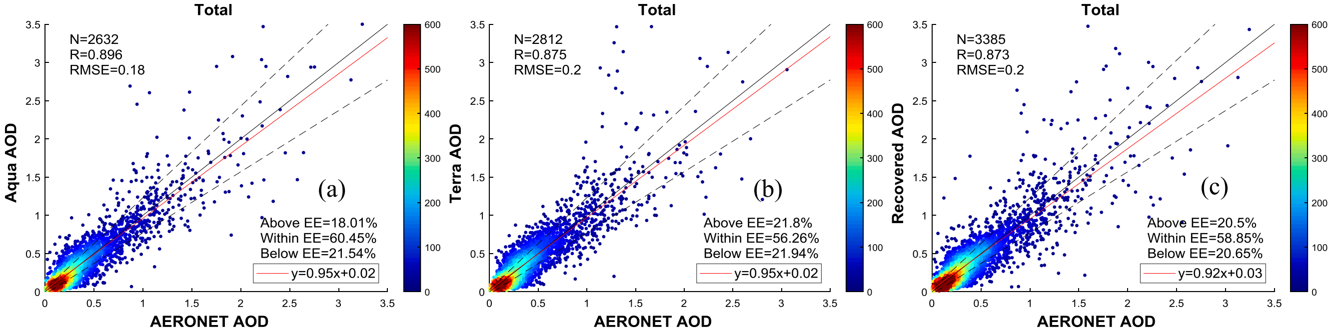


Fig. 7. Validation results of total products for (a) original Aqua AOD, (b) Terra AOD, and (c) recovered AOD. The color bars represent the counts of points. The red solid line represents the regression line, the dashed lines are the EE lines, and the black solid line is the 1:1 line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

difference of land cover types.

3.5. Spatiotemporal approaches (spatiotemporal weighting)

Although the qualities of spatially and temporally recovered results have been stringently controlled, yet some unexpected situation may still exist, which could lead to unstable biases in final results. Generally, spatial (or temporal) information is able to provide complementary knowledge for temporal (or spatial) information (Li et al., 2019; Yuan et al., 2018). In order to acquire more robust recovered results, spatiotemporal approaches are introduced in our algorithm provided that the reliability of spatially recovered result and temporally recovered result meet the requirement. To be specific, the spatially recovered result is obtained with NDVI-based spatial interpolation. Meanwhile, the temporally recovered result is attained with NDVI-based local regression (AW_T as input window and AW_A as output window).

In our algorithm, the spatiotemporal weighting is employed to take spatial and temporal properties into account at the same time. The weight is defined as Eq. (5) after the spatially and temporally recovered results are acquired.

$$w_1 = \frac{N_s}{l_s^2},\ w_2 = \frac{N_t r}{l_t^2}$$
$$w = w_1 + w_2$$
$$\tau_{st} = \frac{w_1\tau_s + w_2\tau_t}{w} \qquad (5)$$

where N_s and N_t represent the number of valid pixels in AW_A and AW_T, respectively; r represents the correlation coefficient between AW_A and AW_T; l_s and l_t represent the minimal distance between missing pixel and valid pixels in AW_A and AW_T, respectively (if the center pixel in AW_T exists, ls and lt are set to 1); τ_s and τ_t represent the spatially and temporally recovered results, respectively; τ_{st} represent the spatiotemporally recovered result. As a consequence, N_s and N_t reflect the abundance of valid information in AW_A and AW_T, respectively. l_s describes the reliability of spatially recovered results. l_t and r show the reliability of temporally recovered results.

4. Experiment results and assessment

In our study, three indicators, i.e., accuracy, coverage and spatial distributions, are considered to evaluate the quality of DB AOD recovery.

4.1. Assessment of AOD accuracy

4.1.1. Validation methodology

The total recovered AOD products (including the original and filled) are compared to the original AOD from Aqua and Terra. Meanwhile, the recovered missing AOD products (only including the filled) are compared with two common methods: (1) Replacement. The missing data in Aqua AOD is directly replaced with Terra AOD on the same day; (2) Linear function. For each day, AOD products of Terra and Aqua are integrally used to fit a linear regression, which could predict the missing data in Aqua. Since ground-based sites offer reduplicative measurements of a point, while satellites provide AOD retrievals of a certain region at a single moment. Hence, AERONET measurements and satellites AOD retrievals must be matched in space and time. In our study, matchups with the average of the AERONET AOD measurements at the MODIS overpass time (± 30 min) and pixels within a radius of 0.25° centered on the AERONET site are adopted (Tang et al., 2016). AERONET network collects AOD data in multiple wavelengths, many of which are slightly different from the MODIS (550 nm). Therefore, the ground-based AOD at 550 nm is interpolated using the Ångström exponent (α), which is defined as shown in Eq. (6).

$$\alpha = -\frac{\ln\left({}^{\tau_1}/_{\tau_2}\right)}{\ln\left({}^{\lambda_1}/_{\lambda_2}\right)} \qquad (6)$$

where τ_1, τ_2 represent the AOD at wavelengths λ_1, λ_2. The results are validated using the expected error (EE), as shown in Eq. (7) (Sayer et al., 2014); the root-mean-square error (RMSE), as shown in Eq. (8).

$$EE = \pm(0.03 + 0.2 \times \tau_{AERO}) \qquad (7)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\tau_{(MOD)i} - \tau_{(AERO)i})^2} \qquad (8)$$

where τ_{AERO} and τ_{MOD} represent the AOD products from satellites and ground-based sites, respectively.

4.1.2. Validation results

As shown in Fig. 7(a)–(c), the total recovered AOD performs well for a count of 3385 matched points with the correlation coefficient (R) reaching 0.873. Other metrics slightly decrease compared to Aqua, which likely results from the quality of Terra. Therefore, the total recovered AOD products nearly maintain the original accuracy of MODIS. As for missing data depicted in Fig. 8(a)–(c), great overestimation and high RMSE (0.25) could be observed in the replaced AOD, suggesting the variation of AOD values in Terra and Aqua. In addition, the R of the linear function (0.826) is even less than that of the replaced AOD (0.839), which indicates that the holistic relationship is unsuitable for large areas recovery. In contrast, our algorithm achieves the best performance, especially for the fraction within EE (57.25%) and R (0.866). Besides, the number of matched points is much more than other methods. However, the larger number of samples occasionally results in a higher correlation. In order to acquire a more reasonable comparison, the matched points for the three methods are adjusted to the same in Table 3. As can be seen, the quality of the recovered AOD still prevails over others regardless of the matched points counts, suggesting that our algorithm is reliable and robust.

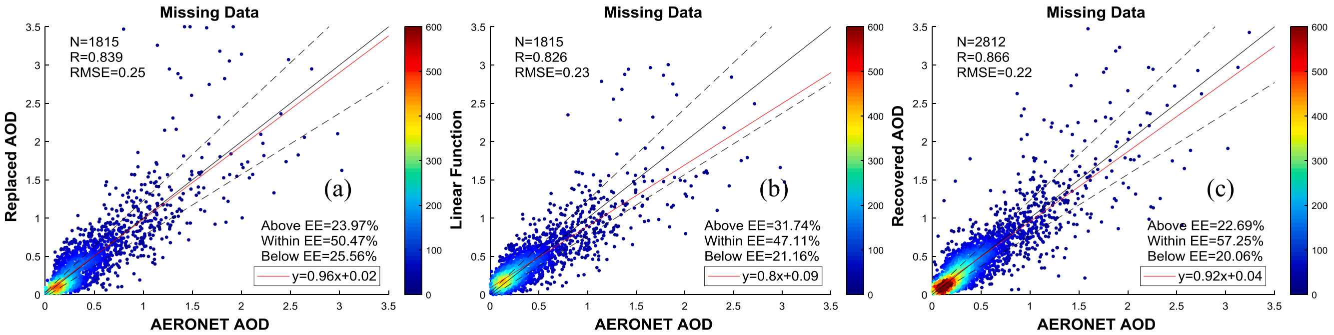


Fig. 8. Validation results of missing data for (a) replaced AOD, (b) linear function, and (c) recovered AOD. The color bars represent the counts of points. The red solid line represents the regression line, the dashed lines are the EE lines, and the black solid line is the 1:1 line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Validation results of missing data with the same matched points.

Name	N	R	RMSE	Above EE	Within EE	Below EE
Replaced AOD	1737	0.840	0.25	23.2%	50.83%	25.96%
Linear function	1737	0.829	0.24	30.86%	47.67%	21.47%
Recovered AOD	1737	0.870	0.21	21.01%	57.86%	21.13%

4.2. Assessment of AOD coverage

Primarily, we quantitatively evaluate the daily coverage of original Aqua AOD products, linear function/replaced AOD products, and recovered AOD products. The completeness was presented using the percentage, which is calculated with the number of valid AOD pixels over land in the study areas. The time series plot of daily AOD coverage in 2016 is shown in Fig. 9. As can be seen, the averaged original AOD coverage of Aqua is only 20.5%. Meanwhile, the improvement of linear function (or replaced) AOD shows slight as 7.8%, suggesting that the complementary information of a single temporal AOD image is insufficient. In contrast, the averaged AOD coverage of proposed algorithm reaches 45.9%, which is increased by 123.9% compared to the original Aqua AOD.

Next, the annual coverage of the original Aqua AOD and the recovered AOD is mapped for comparison. As demonstrated in Fig. 10, the AOD coverage is significantly improved in the study areas, especially for India, Pakistan, Afghanistan, South of Mongolia, and North of China. Meanwhile, the degrees of improvements regionally vary, depending on the local spatial and temporal properties. It can be observed that the recovered AOD emerges in the snow contaminated areas of Tibet, where the original Aqua AOD is absent all through the year. Unfortunately, the recovered AOD is still void in a small region contaminated with snow (see the black mark in Fig. 10(b)), requiring further researches. At last, the study areas are stratified with annual AOD coverage to clearly show the coverage variation in different regions. As depicted in Fig. 11, the valid regions where annual AOD coverage excels 20% mainly distribute in some areas with the bright surface for the original Aqua. However, the regions almost expand to the whole study areas after the recovery. As the stratifying threshold rises, the percentage of valid regions for original Aqua AOD products rapidly declines. Similarly, areas in these regions of the recovered AOD products still greatly exceed the original Aqua, particularly for the stratifying threshold of 80%.

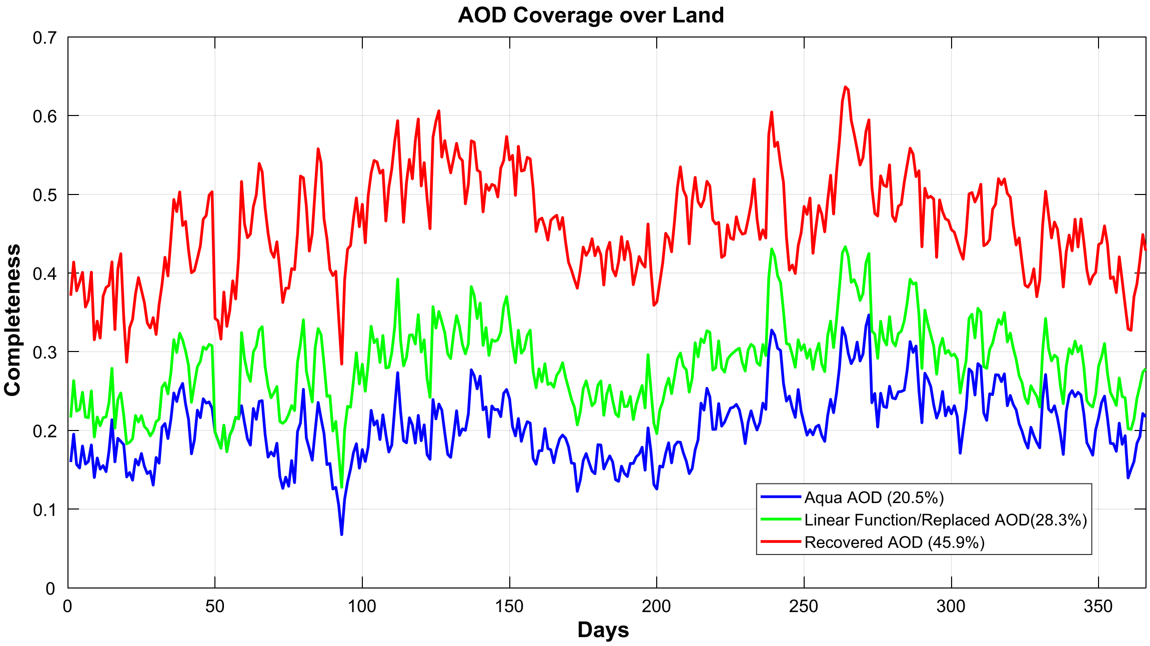


Fig. 9. Time series plot of daily AOD coverage over land in 2016 for original Aqua AOD, linear function/replaced AOD, and recovered AOD. The numbers in parentheses represent the averaged AOD coverage.

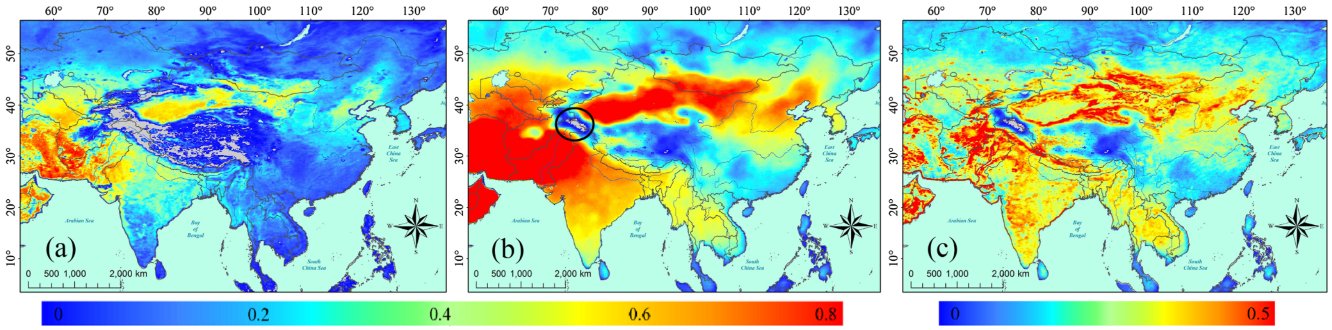


Fig. 10. Maps of annual AOD coverage in 2016 for (a) original Aqua AOD, (b) recovered AOD and (c) recovered AOD–original Aqua AOD. The color bars represent the completeness for (a), (b) and completeness difference for (c). The black mark highlights the small region with the void recovered AOD value.

4.3. Assessment of AOD spatial distribution

At first, the monthly recovered AOD products are selected in order to evaluate the spatial distributions of our algorithm in diverse aerosol modes and land cover types. As can be seen from Fig. 12, the AOD coverage has been expanded after the recovery for each month, which is conducive to temporally continuous analysis. Meanwhile, the distribution of recovered AOD products remains fairly consistent as the original Aqua with the changes of aerosol particles and surface types in different months. It’s worth noting that the details are more abundant for the recovered AOD, which could tell the visible difference of spatial structure. Next, the annual spatial distinctions in 2016 of the original Aqua AOD and the recovered AOD are presented. As illustrated in Fig. 13, the annual AOD shows more coherent in spatial distribution after the recovery. It’s observed that the boundaries among areas with various AOD values are more significant. Besides, some “extreme cases” in the red mark, which are resulted from the scarce number of valid pixels throughout the year, are well mitigated.

5. Conclusion and future work

MODIS AOD products are scarce for most days of the year and consequently it is meaningful to recover the missing pixels. In our study, a novel algorithm considering the effects of aerosol variation, i.e., ST-AVM, is developed to recovery the missing pixels in Aqua AOD products with a single Terra AOD image in large scale. The proposed algorithm is pixel-based on the active windows of Terra and Aqua in the iterative framework. A total of three approaches, i.e., spatial, temporal, and spatiotemporal, are jointly employed to fill the missing AOD data as much as possible under the most suitable circumstances. The results show that the total recovered AOD performs well for a count of 3385 matched points. Other metrics slightly decrease compared to Aqua, which likely results from the quality of Terra, suggesting that total recovered AOD products nearly maintain the original accuracy of MODIS. Meanwhile, the AOD coverage is significantly improved in the study areas and the degrees of improvements regionally vary, depending on the local spatial and temporal properties. Overall, the AOD coverage over land is increased from 20.5% to 45.9% (increment of 123.9%). For recovered AOD products, the valid regions where annual AOD coverage excels 20% almost expand to the whole study areas. Besides, the spatial distribution of monthly recovered AOD products remains fairly consistent as the original Aqua and the annual AOD spatial distribution also shows more coherent after the recovery.

In our study, we only employ DB AOD products to demonstrate the advantages of ST-AVM algorithm. Meanwhile, it is obvious that the proposed algorithm is suitable for other AOD products, such as DT and DT&DB merged. Besides, the recovered AOD products could offer more spatially continuous information of aerosol for practical applications,

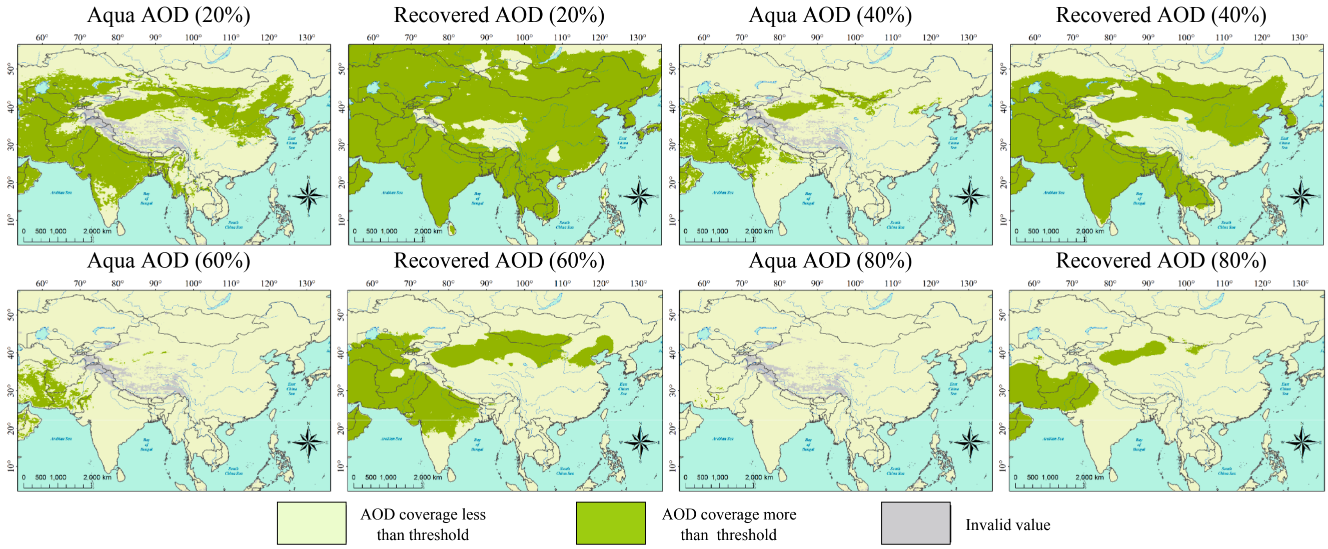


Fig. 11. Maps of valid regions stratified with annual AOD coverage in 2016 for original Aqua AOD and recovered AOD. The numbers in parentheses represent the stratifying threshold.

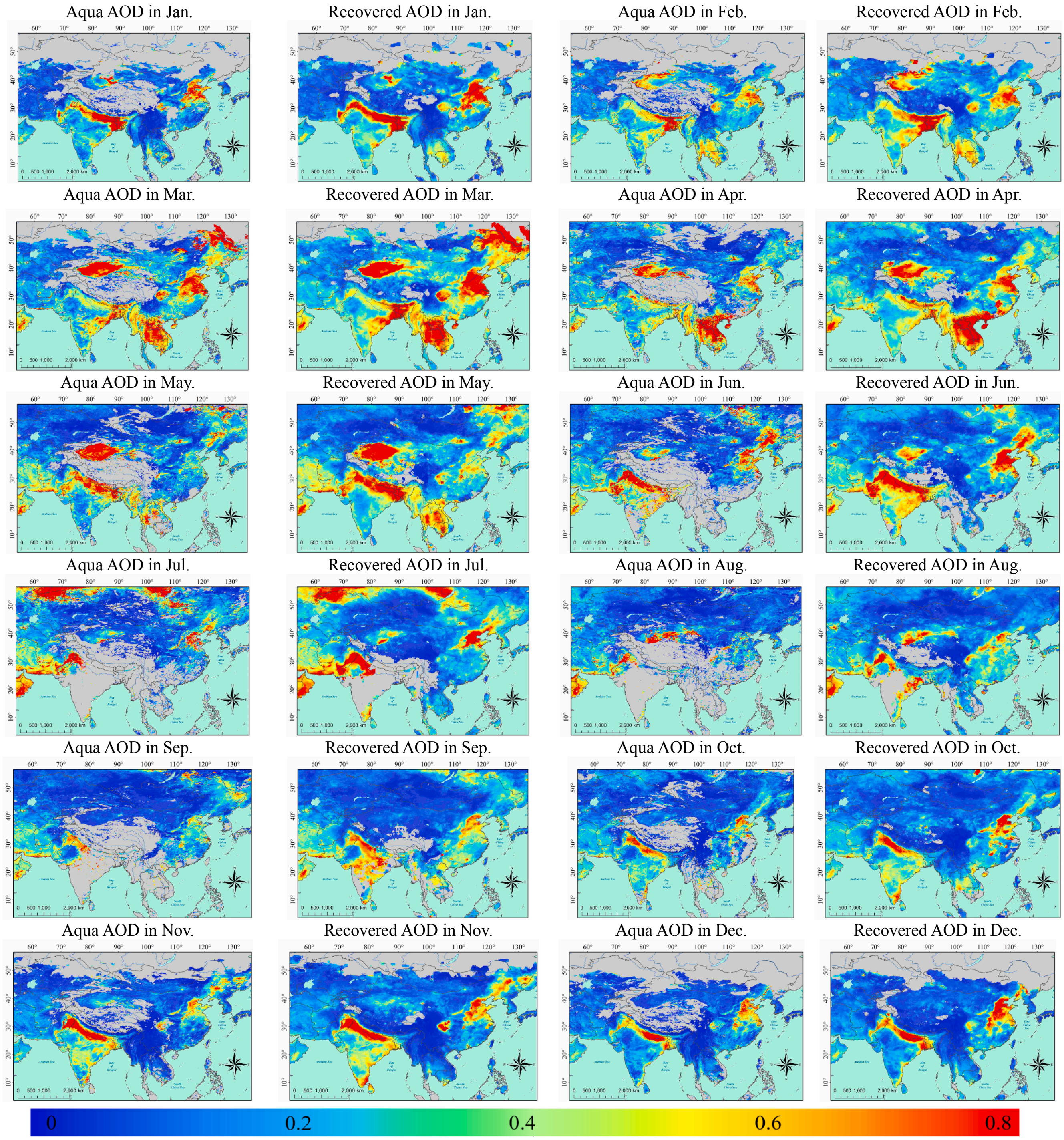


Fig. 12. Maps of monthly AOD distribution in 2016 for original Aqua AOD and recovered AOD. The color bar represents the AOD value.

such as the estimation (Li et al., 2017a, 2017b; Yao et al., 2019) and the analysis (Yang et al., 2017, 2019) of particulate matter. Although ST-AVM algorithm achieves well performance, yet the improvement of AOD coverage is still deficient due to the inherent defect. Multi-temporal AOD images can provide more effective knowledge compared to a single temporal AOD image. However, the distinction between multitemporal AOD images also greatly rises, which enlarge the difficulty to reasonably employ the valid information. There is no doubt that the accuracy of such recovered AOD products will decline without correctly considering the significant AOD variation. Furthermore, the increment of auxiliary AOD products could lead to more time consumption and less convenience. Hence, it requires more researches and we will spare no effort working on it in the future.

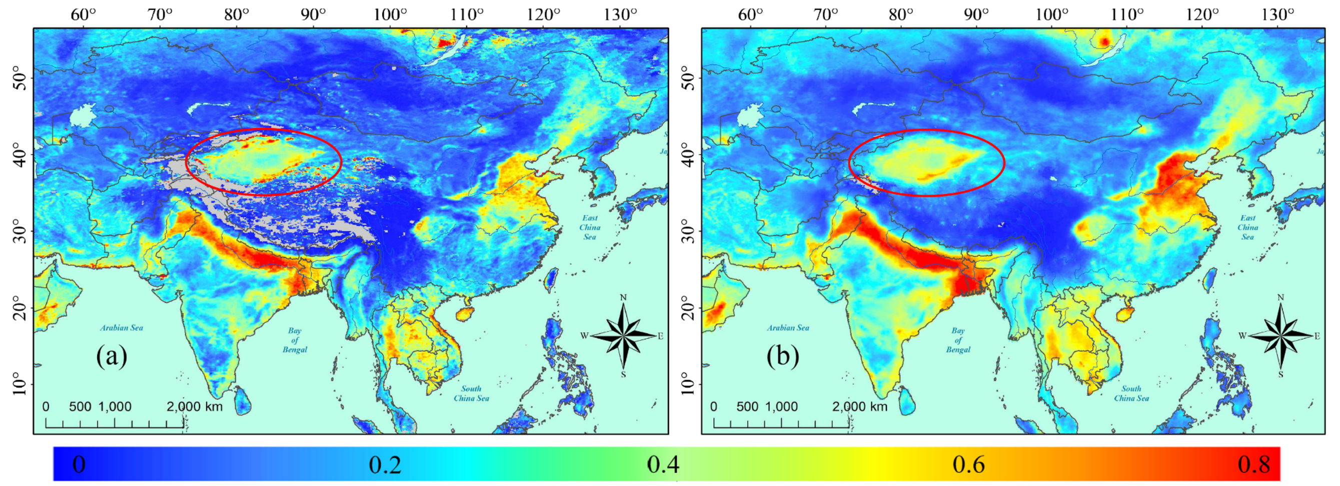


Fig. 13. Maps of annual AOD distribution in 2016 for (a) original Aqua AOD and (b) recovered AOD. The color bar represents the AOD value. The red mark highlights some “noises” caused by the absence of valid pixels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Acknowledgments

This work was supported by the National Key R & D Program of China (No. 2016YFC0200900), the National Natural Science Foundation of China (No. 41922008), and the Science and Technology Major Project of Hubei Province (No. 2019AAA046). The authors would like to express our gratitude to the Atmosphere Archive and Distribution System (LAADS) for providing the MODIS AOD and NDVI products, and the Principle Investigators for establishing and maintaining the AERONET sites.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2019.08.017.

References

Ahn, C., Torres, O., Jethva, H., 2014. Assessment of OMI near-UV aerosol optical depth over land. J. Geophys. Res. Atmosph. 119 (5), 2457–2473.

Bisht, G., Venturini, V., Islam, S., Jiang, L.E., 2005. Estimation of the net radiation using MODIS (Moderate Resolution Imaging Spectroradiometer) data for clear sky days. Remote Sens. Environ. 97 (1), 52–67.

Che, Y., Xue, Y., Guang, J., She, L., Guo, J., 2018. Evaluation of the AVHRR DeepBlue aerosol optical depth dataset over mainland China. ISPRS J. Photogramm. Remote Sens. 146, 74–90.

Della Ceca, L.S., Ferreyra, M.F.G., Lyapustin, A., Chudnovsky, A., Otero, L., Carreras, H., Barnaba, F., 2018. Satellite-based view of the aerosol spatial and temporal variability in the Córdoba region (Argentina) using over ten years of high-resolution data. ISPRS J. Photogramm. Remote Sens. 145, 250–267.

Emetere, M., Sanni, S.E., Okoro, E.E., 2019. Analysing the dangerous trends of air pollution over Bo-Sierra Leone using fourteen years of aerosol optical depth data. Nat. Environ. Pollut. Technol. 18 (2), 149–358.

Giles, D.M., Sinyuk, A., Sorokin, M.G., Schafer, J.S., Smirnov, A., Slutsker, I., Welton, E.J., 2019. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database–automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmos. Meas. Tech. 12 (1), 169–209.

Guo, Y., Hong, S., Feng, N., Zhuang, Y., Zhang, L., 2012. Spatial distributions and temporal variations of atmospheric aerosols and the affecting factors: a case study for a region in central China. Int. J. Remote Sens. 33 (12), 3672–3692.

Gupta, P., Levy, R.C., Mattoo, S., Remer, L.A., Munchak, L.A., 2016. A surface reflectance scheme for retrieving aerosol optical depth over urban surfaces in MODIS Dark Target retrieval algorithm. Atmos. Meas. Tech. 9 (7), 3293–3308.

He, Q., Gu, Y., Zhang, M., 2019. Spatiotemporal patterns of aerosol optical depth throughout China from 2003 to 2016. Sci. Total Environ. 653, 23–35.

Hou, X., Zhu, B., Kumar, K.R., Lu, W., 2019. Inter-annual variability in fine particulate matter pollution over China during 2013–2018: role of meteorology. Atmos. Environ. 116842.

Hsu, N.C., Tsay, S.C., King, M.D., Herman, J.R., 2004. Aerosol properties over bright-reflecting source regions. IEEE Trans. Geosci. Remote Sens. 42 (3), 557–569.

Hsu, N.C., Jeong, M.J., Bettenhausen, C., Sayer, A.M., Hansell, R., Seftor, C.S., Tsay, S.C., 2013. Enhanced deep blue aerosol retrieval algorithm: the second generation. J. Geophys. Res. Atmosph. 118 (16), 9296–9315.

Jinnagara Puttaswamy, S., Nguyen, H.M., Braverman, A., Hu, X., Liu, Y., 2014. Statistical data fusion of multi-sensor AOD over the continental United States. Geocar. Int. 29 (1), 48–64.

Justice, C.O., Vermote, E., Townshend, J.R., Defries, R., Roy, D.P., Hall, D.K., Lucht, W., 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. IEEE Trans. Geosci. Remote Sens. 36 (4), 1228–1249.

Kahn, R.A., Gaitley, B.J., Garay, M.J., Diner, D.J., Eck, T.F., Smirnov, A., Holben, B.N., 2010. Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. J. Geophys. Res. Atmosph. 115 (D23).

Kaufman, Y.J., Koren, I., Remer, L.A., Tanré, D., Ginoux, P., Fan, S., 2005. Dust transport and deposition observed from the Terra-Moderate Resolution Imaging Spectroradiometer (MODIS) spacecraft over the Atlantic Ocean. J. Geophys. Res. Atmosph. 110 (D10).

King, M.D., Kaufman, Y.J., Menzel, W.P., Tanre, D., 1992. Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). IEEE Trans. Geosci. Remote Sens. 30 (1), 2–27.

Levy, R.C., Mattoo, S., Munchak, L.A., Remer, L.A., Sayer, A.M., Patadia, F., Hsu, N.C., 2013. The Collection 6 MODIS aerosol products over land and ocean. Atmos. Meas. Tech. 6 (11).

Li, T., Shen, H., Zeng, C., Yuan, Q., Zhang, L., 2017a. Point-surface fusion of station measurements and satellite observations for mapping PM2.5 distribution in China: methods and assessment. Atmos. Environ. 152, 477–489.

Li, T., Shen, H., Yuan, Q., Zhang, X., Zhang, L., 2017b. Estimating ground-level PM2.5 by fusing satellite and station observations: a geo-intelligent deep learning approach. Geophys. Res. Lett. 44 (23).

Li, L., Yang, J., Wang, Y., 2014. An improved dark object method to retrieve 500 m-resolution AOT (Aerosol Optical Thickness) image from MODIS data: a case study in the Pearl River Delta area, China. ISPRS J. Photogramm. Remote Sens. 89, 1–12.

Li, X., Wang, L., Cheng, Q., Wu, P., Gan, W., Fang, L., 2019. Cloud removal in remote sensing images using nonnegative matrix factorization and error correction. ISPRS J. Photogramm. Remote Sens. 148, 103–113.

Ma, Z., Liu, R., Liu, Y., Bi, J., 2019. Effects of air pollution control policies on PM 2.5 pollution improvement in China from 2005 to 2017: a satellite-based perspective. Atmos. Chem. Phys. 19 (10), 6861–6877.

Sayer, A.M., Hsu, N.C., Bettenhausen, C., Ahmad, Z., Holben, B.N., Smirnov, A., Zhang, J., 2012. SeaWiFS Ocean Aerosol Retrieval (SOAR): algorithm, validation, and comparison with other data sets. J. Geophys. Res. Atmosph. 117 (D3).

Sayer, A.M., Hsu, N.C., Bettenhausen, C., Jeong, M.J., Meister, G., 2015. Effect of MODIS Terra radiometric calibration improvements on Collection 6 Deep Blue aerosol products: validation and Terra/Aqua consistency. J. Geophys. Res. Atmosph. 120 (23), 12–157.

Sayer, A.M., Munchak, L.A., Hsu, N.C., Levy, R.C., Bettenhausen, C., Jeong, M.J., 2014. MODIS Collection 6 aerosol products: comparison between Aqua's e-Deep Blue, Dark Target, and “merged” data sets, and usage recommendations. J. Geophys. Res. Atmosph. 119 (24), 13–965.

Tang, Q., Bo, Y., Zhu, Y., 2016. Spatiotemporal fusion of multiple-satellite aerosol optical depth (AOD) products using Bayesian maximum entropy method. J. Geophys. Res. Atmosph. 121 (8), 4034–4048.

Van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., Villeneuve, P.J., 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. Environ. Health Perspect. 118 (6), 847–855.

Volkamer, R., Jimenez, J.L., San Martini, F., Dzepina, K., Zhang, Q., Salcedo, D., Molina, M.J., 2006. Secondary organic aerosol formation from anthropogenic air pollution:

rapid and higher than expected. Geophys. Res. Lett. 33 (17).

Wang, Y., Yuan, Q., Li, T., Shen, H., Zheng, L., Zhang, L., 2019. Evaluation and comparison of MODIS Collection 6.1 aerosol optical depth against AERONET over regions in China with multifarious underlying surfaces. Atmos. Environ. 200, 280–301.

Xu, H., Guang, J., Xue, Y., de Leeuw, G., Che, Y.H., Guo, J., Wang, T.K., 2015. A consistent aerosol optical depth (AOD) dataset over mainland China by integration of several AOD products. Atmos. Environ. 114, 48–56.

Xu, M., Jia, X., Pickering, M., Jia, S., 2019. Thin cloud removal from optical remote sensing images using the noise-adjusted principal components transform. ISPRS J. Photogramm. Remote Sens. 149, 215–225.

Yang, J., Hu, M., 2018. Filling the missing data gaps of daily MODIS AOD using spatiotemporal interpolation. Sci. Total Environ. 633, 677–683.

Yang, Q., Yuan, Q., Li, T., Shen, H., Zhang, L., 2017. The relationships between PM2.5 and meteorological factors in China: seasonal and regional variations. Int. J. Environ. Res. Public Health 14 (12), 1510.

Yang, Q., Yuan, Q., Yue, L., Li, T., Shen, H., Zhang, L., 2019. The relationships between PM2.5 and aerosol optical depth (AOD) in mainland China: about and behind the spatio-temporal variations. Environ. Pollut. 248, 526–535.

Yao, F., Wu, J., Li, W., Peng, J., 2019. A spatially structured adaptive two-stage model for retrieving ground-level PM2. 5 concentrations from VIIRS AOD in China. ISPRS J. Photogramm. Remote Sens. 151, 263–276.

Yuan, Q., Zhang, Q., Li, J., Shen, H., Zhang, L., 2018. Hyperspectral image denoising employing a spatial-spectral deep residual convolutional neural network. IEEE Trans. Geosci. Remote Sens. 57 (2).

Zeng, C., Long, D., Shen, H., Wu, P., Cui, Y., Hong, Y., 2018. A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud. ISPRS J. Photogramm. Remote Sens. 141, 30–45.

Zhang, Q., Yuan, Q., Zeng, C., Li, X., Wei, Y., 2018a. Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network. IEEE Trans. Geosci. Remote Sens. 56 (8), 4274–4288.

Zhang, R., Di, B., Luo, Y., Deng, X., Grieneisen, M.L., Wang, Z., Zhan, Y., 2018b. A nonparametric approach to filling gaps in satellite-retrieved aerosol optical depth for estimating ambient PM2.5 levels. Environ. Pollut. 243, 998–1007.

Zhang, T., Zeng, C., Gong, W., Wang, L., Sun, K., Shen, H., Zhu, Z., 2017. Improving spatial coverage for aqua MODIS AOD using NDVI-based multi-temporal regression analysis. Remote Sens. 9 (4), 340.

Zhang, Z., Wu, W., Fan, M., Tao, M., Wei, J., Jin, J., Wang, Q., 2019. Validation of Himawari-8 aerosol optical depth retrievals over China. Atmos. Environ. 199, 32–44.

Zhao, C., Liu, Z., Wang, Q., Ban, J., Chen, N.X., Li, T., 2019. High-resolution daily AOD estimated to full coverage using the random forest model approach in the Beijing-Tianjin-Hebei region. Atmos. Environ. 203, 70–78.