



# Geostatistical Methods for Filling Gaps in Level-3 Monthly-Mean Aerosol Optical Depth Data from Multi-Angle Imaging Spectroradiometer

Manoj K. Singh<sup>1\*†</sup>, Parvatham Venkatachalam<sup>1</sup>, Ritesh Gautam<sup>1,2\*†</sup>

<sup>1</sup> Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai- 400076, India

<sup>2</sup> Interdisciplinary Program in Climate Studies, Indian Institute of Technology Bombay, Mumbai- 400076, India

## ABSTRACT

The Aerosol Optical Depth (AOD) retrieved from satellite remote sensing measurements such as from MISR and MODIS, both onboard the Terra platform, are widely used for studying regional and global patterns of aerosol loading. Aerosol products from these sensors are also used for analyzing feedbacks and relationship between aerosols and climatic variables including clouds, precipitation, and radiation fluxes. Several statistical techniques leading to the understanding of such relationships, including empirical orthogonal function and temporal trend extraction methods, require spatially complete AOD data records. Inherent to remote sensing of aerosols, cloud cover significantly affects aerosol retrievals and results in missing data across the AOD products. This paper demonstrates widely-used geostatistical techniques, such as Co-Kriging (CK) and Regression Kriging (RK), for spatially-filling missing data in the MISR AOD product for the period 2001–2013. Among the unique characteristics of this data-filling algorithm is that it utilizes additional AOD information obtained from MODIS. The mean accuracy of the predicted MISR AOD using CK method is estimated to be 0.05, globally. The gap-filled MISR AOD data are also compared with 131 ground-based Aerosol Robotic Network (AERONET) stations, located around the world. It is found that Root Mean Squared Error of the gap-filled AOD dataset and the original MISR AOD product with respect to AERONET data are 0.143. The gap-filled AOD dataset can be used in applications where the presence of missing values is undesirable such as for global/regional aerosol variability and trend analysis.

**Keywords:** Gap-filling; Geostatistics; Kriging; MISR; MODIS.

## INTRODUCTION

Atmospheric aerosols are of great relevance to environmental and climate change studies (Ramanathan *et al.*, 2001; Rosenfeld *et al.*, 2001), as they play a vital role in recent global and regional changes in air quality and climatic patterns (de Meij *et al.*, 2012; Hsu *et al.*, 2012; Murphy, 2013). A growing body of evidence suggests that aerosols significantly impact the Earth's radiation budget through their direct radiative effects due to scattering and absorption of solar radiation (Ramanathan *et al.*, 2001). In addition, aerosols also indirectly alter microphysical properties of clouds, and thus can induce changes in precipitation patterns

(Rosenfeld *et al.*, 2001). Columnar Aerosol Optical Depth (AOD) is a measure of the amount of aerosol loading in the atmosphere. Two of the widely-used satellite datasets for AOD retrieval are available from NASA's MODerate resolution Imaging Spectro-radiometer (MODIS) and Multi-angle Imaging SpectroRadiometer (MISR) observations (Kahn *et al.*, 2010; Levy *et al.*, 2013). The AOD products from these spaceborne sensors are mainly available in two levels, i.e., swath (Level-2) and gridded (Level-3) formats. Several studies involving aerosol characterization, aerosol-climate effects and aerosol trend analysis utilize the monthly-mean gridded AOD products from MISR and MODIS (Level-3 data) (Shrestha and Barros, 2010; Zhang and Reid, 2010; Li *et al.*, 2013). The Level-3 MODIS AOD dataset includes aerosol retrievals over ocean as well as Dark-Target retrievals over land (Levy *et al.*, 2013). In addition, AOD retrievals over bright surfaces such as arid and desert regions, from the Deep Blue algorithm, are also available from both Terra and Aqua platforms (Hsu *et al.*, 2013).

Inherent to remote sensing of aerosols, cloud cover significantly hampers clear-sky retrievals, and results in missing data across various aerosol products. Here, by missing data we mean the pixel with invalid AOD retrieval. Furthermore, sampling biases resulting from the variable

<sup>†</sup> Now at Environmental Defense Fund, Washington DC 20009, USA

<sup>‡</sup> Now at University of Petroleum and Energy Studies, Dehradun 248007, India

\* Corresponding author.

E-mail address: rgautam@edf.org (R. Gautam); mkumar@ddn.upes.ac.in (M.K. Singh)

revisit orbital frequency can also impact the statistics of the monthly-mean aerosol products. For example, the fact that MISR views the same region of the Earth once in 7–9 days, leads to only ~4 daily retrievals averaged into the monthly mean AOD product, representing about 13% sampling. On a regional basis, where cloud cover is persistent, e.g., Amazon basin, parts of Central and western Africa and South-East Asia, the missing data in MISR AOD varies from 28% to 50% (Table 1). Whereas, the MODIS has a revisit frequency of close to once-a-day, yielding a superior sampling in terms of the monthly-mean AOD data.

Studies involving spatio-temporal analysis of aerosols on monthly, annual and decadal time scales require the time series to be of reliable sample size at a pixel-level. If the time series consists of a number of missing values above a certain threshold, the resultant trend obtained from the time series may not be a reliable or accurate representation of the change over time. Empirical Orthogonal Functions (EOFs) are another technique used widely in the seasonal decomposition of aerosols (or other environmental/climatic variables) based on satellite datasets (Shrestha and Barros, 2010; Li *et al.*, 2013), where central to the EOF analysis is the variance-covariance matrix. The elements of this matrix are computed using the covariance between two time series. By elimination of missing data, an estimate of the variance-covariance matrix can be made. However, given the significant amount of missing data, this matrix may not be positive definite, which is an essential property of the variance-covariance matrix.

In order to address the challenges related to the assessment of spatio-temporal variability of aerosols and associated climatic effects, this paper demonstrates the application of geostatistical methods to produce a gap-filled satellite aerosol data record. Algorithms to fill data gaps in satellite-retrieved geophysical parameters have evolved over the years. One of the oldest methods to fill data gaps is optimal interpolation (Gandin and Hardin, 1965), which is a reduced version of Kalman filtering. Bayesian models (e.g., dynamic linear models) have also been used to fill missing data (Kaplan *et al.*, 1997). Techniques such as optimal interpolation and Kalman filtering require apriori knowledge of the error covariance matrix. However, in the presence of missing data, the estimate of covariance matrix can be incorrect, as elements of the matrix do not reflect true variance or covariance in the presence of missing data in the AOD time

series. Additionally, other widely-used geostatistical based methods have also been used to fill missing data in different datasets. Buytaert *et al.* (2006) used Universal Kriging (UK) to interpolate a rainfall dataset and compared the results with spatial interpolation using Thiessen polygons. They found that the UK (mean error of cross validation: 0.03) performed better than the Thiessen polygons (mean error of cross validation: -0.127). Carrera-Hernandez and Gaskin (2007) used various Kriging methods to interpolate rainfall and temperature data such as ordinary Kriging, Kriging with external drift, block Kriging with external drift, and found that their results improved by taking elevation as a co-variable. Furthermore, Zhang *et al.* (2009) used Co-Kriging to fill gaps in cloudy pixels of multispectral remotely sensed imagery, and Ruiz-Arias *et al.* (2013) used ordinary Kriging to fill data gaps in MODIS daily Level-3 datasets.

Merging of multiple satellite-based datasets is another approach towards producing spatially and spatio-temporally complete AOD dataset (Nguyen *et al.*, 2012, 2014; Xu *et al.*, 2015). Xu *et al.* (2015) used a maximum likelihood method, where the weights are derived using Root Mean Squared Error (RMSE) from the ground-based Aerosol Robotic Network (AERONET) to create an AOD dataset, based on combined satellite observations over mainland China. Nguyen *et al.* (2012) used a spatial-statistical method to merge Level-2 AOD datasets, retrieved from MISR and MODIS, and found that their resultant dataset improved over a dataset obtained using a Bayesian melding method (Fuentes and Raftery, 2005), and another obtained using fixed rank Kriging (Cressie and Johannesson, 2008). Additionally, Djuric *et al.* (2016), Chatterjee *et al.* (2010), and Kinne (2009) utilized AERONET data alongside satellite-based datasets to produce merged AOD products. Chatterjee *et al.* (2010) interpolated the AERONET data using a space-time UK method with the AOD retrieved from MISR and MODIS as co-variables. However, the merging and gap-filling methods/objectives have an inherent difference. In general, the goal of a merging algorithm is to produce a new product whose accuracy is better than the input datasets, whereas in spatial gap-filling, the objective is to make a dataset complete while retaining its basic properties.

In this paper, we apply the Co-Kriging (CK) and Regression Kriging (RK) methods to produce a spatially-complete and gap-filled monthly-mean MISR Level-3

**Table 1.** Fraction of missing data in MISR and fraction of pixels with valid retrievals of MODIS and missing data from MISR for regions shown in Fig. 2. The second column represents the total number of pixels used for the period 2001–2013, third column shows the percent of missing pixels of MISR AOD data and the fourth column represents percent of missing pixels of MISR AOD where MODIS AOD has valid data.

Regions	Total pixels ( $\times 10^5$ )	MISR missing pixels (%)	MISR missing pixels with presence of MODIS (%)
Global	179.7	9.94	96.77
Region 1	5.5	43.31	93.01
Region 2	1.4	28.26	87.27
Region 3	2.1	34.40	96.08
Region 4	3.6	36.80	91.87
Region 5	1.6	49.69	98.86
Region 6	2.3	43.12	98.90

AOD dataset, taking advantage of the presence of MODIS AOD data. Here, the latter has larger sample size in terms of the number of pixels used to create the Level-3 monthly-mean data. The CK method uses the spatial structure of MISR as well as MODIS AOD data to fill gaps in the MISR AOD product. This method can predict AOD at any pixel irrespective of whether the MODIS data are present or not. The RK method uses only the spatial structure of MISR AOD and predicts at those pixels only where the MODIS data are present, by also taking into account the correlation of MISR and MODIS AOD data. The accuracy of the predicted MISR AOD data is estimated using correlation coefficients and RMSE with the ground-based aerosol dataset from the AERONET. In addition, the accuracy of the gap-filled MISR AOD product is also estimated by using a cross-validation approach. Here, the prediction algorithm is used to estimate the filled value at pixels with valid AOD retrievals. The predicted AODs are then compared with the operational MISR AOD dataset.

## SATELLITE AND AERONET AEROSOL DATASETS

Our study area is bounded by the global region between 40°N and 40°S latitudes. The MODIS Level-3 monthly-mean AOD data from Terra (Collection 6) is used in this study, which are available at a spatial resolution of  $1^\circ \times 1^\circ$ . MODIS has a swath width of 2330 km; with such a large swath, it has a global coverage of almost once a day. It is typically easier to detect aerosols over dark ocean surfaces than is to detect aerosols over land. This is particularly true over bright surfaces such as arid/desert regions where the contribution of the aerosols to the top of atmosphere signal observed by a satellite instrument is relatively smaller as compared to the surface reflectance. The Deep Blue or DB (Hsu *et al.*, 2013) algorithm is especially designed for retrieving AOD from MODIS over bright surfaces. Over dark surfaces, such as vegetated land cover, the Dark-Target or DT (Levy *et al.*, 2013) algorithm retrieves AOD from MODIS data. The overall uncertainty of AOD retrievals of DT and DB products is reported to be  $0.05 \pm 0.15\tau$  (Levy *et al.*, 2013) and  $0.03 \pm 0.2\tau$  (Sayer *et al.*, 2014), respectively. In the Collection MODIS aerosol products, the DT and DB datasets are separately provided as well as are a combined product (Levy *et al.*, 2013; Sayer *et al.*, 2014), using a monthly Normalized Difference Vegetation Index (NDVI) climatology. Over bright surfaces ( $\text{NDVI} < 0.2$ ) the DB product is selected, over dark surfaces ( $\text{NDVI} > 0.3$ ) the DT product is selected, while for surfaces with  $0.2 \leq \text{NDVI} \leq 0.3$ , the average of the respective AODs from the DT and DB products are taken in the combined dataset. Our study period is limited from January 2001 to December 2013, for a total of 156 months.

The MISR Level-3 monthly AOD data used in this study has a spatial resolution of  $0.5^\circ \times 0.5^\circ$ . The MISR has a swath width of 380km and views the entire Earth once in 7–9 days. With such a significant difference in swath size, the MISR AOD product has fewer samples to aggregate in the computation of the monthly-mean compared to MODIS.

Therefore, in general, the fraction of missing data in MISR is greater compared to MODIS. Here, the fraction of missing data represents the ratio of the number of pixels with invalid AOD retrievals and the total pixels for the study period (2001–2013). In our study period (2001–2013), missing data in the MISR AOD product are almost twice as frequent compared to MODIS. MISR AOD retrievals are reported at four wavelengths; we selected AOD product in the mid-visible wavelength ( $0.55 \mu\text{m}$ ). The uncertainty of the MISR AOD product has been reported to be maximum of 0.05 and  $0.2\tau$  (Kahn *et al.*, 2010), where  $\tau$  represents the AOD.

As previously noted in section 1, MODIS AODs are used to fill gaps in the MISR AOD product using CK and RK methods. The RK method can be used to predict a missing MISR AOD value only at the spatial coordinates where a co-variable, in our case MODIS AOD, is present. Thus in the RK method, the spatial coordinates of MISR and MODIS pixels must be the same. In other words, if we assume the center of a pixel as a spatial coordinate, the spatial resolution must be the same for the RK method for filling data gaps in MISR using MODIS AOD. Hence, to use the RK method for data filling, we have downscaled the MODIS AOD to  $0.5^\circ$  degree grids by assuming that all the four  $0.5^\circ$  grid cells contain the same value as their parent grid.

As a representative of ground truth, the AERONET data are used for the inter-comparison of our gap-filled-AOD data with ground-based sites located globally. More information about AERONET can be found in Holben *et al.* (1998). Since the overpass of the Terra satellite is at  $\sim 10:30\text{AM}$  local-time (which includes both MODIS and MISR data), sunphotometer measurements as part of AERONET, made between 10:00AM and 11:00AM local-time only are considered for our analysis.

## METHODOLOGY

In this section, our methodology to fill missing data in the MISR AOD product is described. For our study region (bounded by 40°S–40°N) and time period (2001–2013), there are a total of 156 (months)  $\times$  160 (latitudes)  $\times$  720 (longitudes) =  $1.8 \times 10^7$  data points. Among these,  $1.8 \times 10^6$  instances of the MISR AOD data have missing data. In order to fill these data points with suitable interpolated values, Kriging methods (CK and RK) are used. Out of the  $1.8 \times 10^6$  instances of the missing data from MISR, the MODIS aerosol retrievals are present at  $1.7 \times 10^6$  instances. This implies that for a MISR AOD missing pixel, the probability that the MODIS AOD will have valid data is 96.77%. Therefore, the MODIS AOD dataset is suitable for filling gaps in MISR AOD dataset. We have used CK and RK methods to predict MISR AOD data, both of which take advantage of the presence of MODIS data. The RK method can predict AOD for only those pixels, where MODIS has valid data. However, CK method can predict AOD at any pixel irrespective of whether the MODIS data are present or not. The algorithms used in our study for filling MISR Level-3 monthly-mean AOD data are based on two steps:

1. The MISR AOD is predicted at a pixel with missing

data using CK, which is an interpolation method involving co-variables. Co-variables are the variables present along with the target variable. Co-variables must be denser than the target variable, so that they can be useful in data filling of the target variable. In our case, the target variable is MISR AOD data, which has higher fraction of missing data compared to co-variable (MODIS AOD). Here, a spatial structure (semivariogram) is prepared using both the variable and co-variable. Semivariogram is a function which indicates spatial correlation in observations for a region. More details about semivariogram are discussed in section S1 of the supplementary material. This structure is used to prepare a Kriging system of equations which can then be used to predict missing data of the variable under study. Solving Kriging system of equations is computationally expensive as it requires inverting a square matrix of large size. The region of our study is thus divided into 72-smaller non-overlapping regions of size  $20^\circ \times 20^\circ$  to make the invertible matrix smaller in size. To form smaller regions, latitudes between  $40^\circ\text{N}$ – $40^\circ\text{S}$  are divided into 4 equal regions at an interval of  $20^\circ$  i.e.,  $40^\circ\text{S}$ – $20^\circ\text{S}$ ,  $20^\circ\text{S}$ – $0^\circ$ ,  $0^\circ$ – $20^\circ\text{N}$  and  $20^\circ\text{N}$ – $40^\circ\text{N}$ . Similarly, longitudes between  $180^\circ\text{W}$ – $180^\circ\text{W}$  are divided into 18 equal regions at an interval of  $20^\circ$ . These 4 latitude intervals and 18 longitude intervals are used for the formation of 72 small regions. Since each region is bounded, the estimation of AOD using Kriging suffers from the boundary problem including edge effects, where relationship across bordering pixels are ignored. To reduce edge effect, the smaller regions of size  $20^\circ \times 20^\circ$  are padded by adding outer pixels adjacent to the boundary of the region. Thus, the Kriging method is used for filling gaps at each of the 72 regions of size  $22^\circ \times 22^\circ$  (after padding). After filling missing pixels for each region, the final dataset is produced by combining

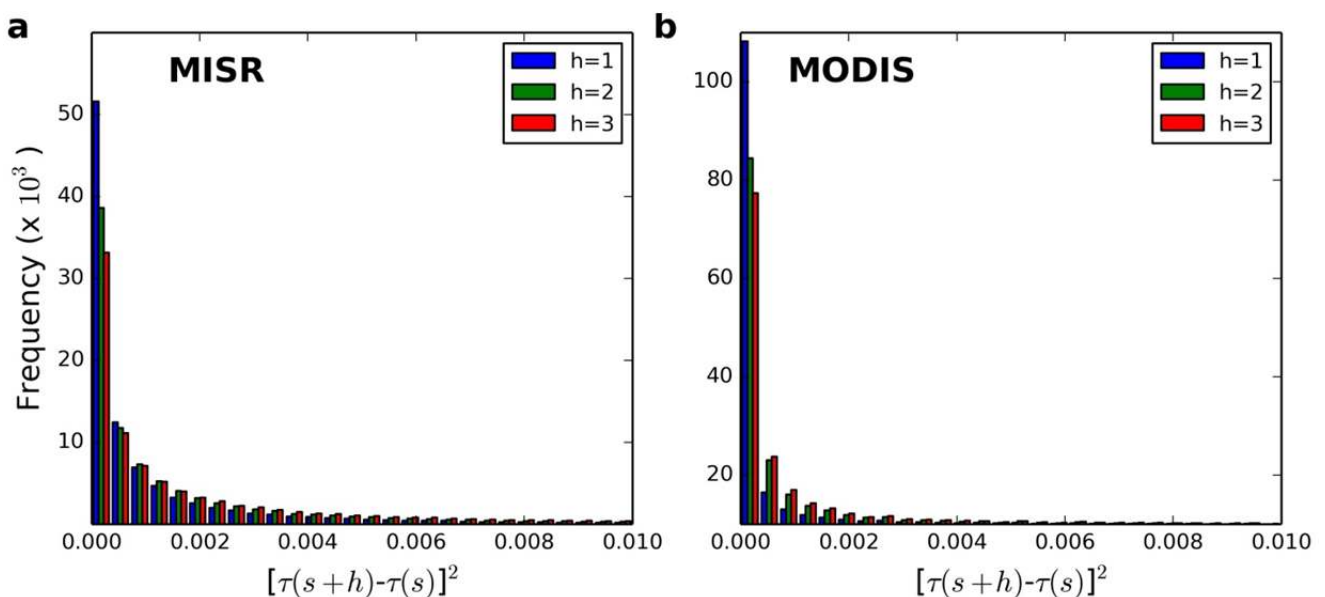
the 72 regions and by discarding the padding pixels.

- To inter-compare the AOD predicted using CK, another method- RK is also implemented to estimate AOD. This method also takes advantage of the presence of co-variable, where a spatial structure is prepared using the target variable only. The RK method can predict missing data only at pixels with presence of co-variable. Further, the ordinary Kriging method is used for making estimate of AOD at a pixel where the co-variable (MODIS) is not present. We have used R- package *gstat* (Pebesma, 2004) for estimating AOD using CK and RK.

### Kriging

Co-Kriging and Regression Kriging methods are widely-used techniques for interpolation in spatial statistics, applicable to the fields of environment science, hydrology, climate science (Goovaerts, 1997; Aalto et al., 2013). One of the important assumptions is the stationarity of the spatial field. For CK and RK, a weaker form of stationarity is required. It is sufficient that the  $E[\tau(s+h) - \tau(s)]^2$  is constant, where  $\tau(s)$  is the AOD at a pixel  $s$  and  $\tau(s+h)$  is the AOD at a pixel with distance  $h$  from  $s$ . Fig. 1 shows the histogram of  $[\tau(s+h) - \tau(s)]^2$  for the pixels separated by 1, 2 and 3 pixels for MISR (Fig. 1(a)) and MODIS (Fig. 1(b)). The figure shows that the distribution is biased towards a single value. Therefore,  $E[\tau(s+h) - \tau(s)]^2$  can be assumed to be constant and hence MISR and MODIS data can be assumed to be stationary.

Here, we briefly give an overview of Kriging concepts. More details can be found in the literature for spatial statistics methods (e.g., Cressie, 1993; Chiles and Delfiner, 2009). The MISR AOD data are available at pixels with centers  $(x_i, y_i)$ , where  $x_i \in [-179.75, -179.25, \dots, +179.75]$  and  $y_i \in [-39.75, -39.25, \dots, +39.75]$  ( $x_i$  stands for longitudes and  $y_i$  stands for latitudes). In the CK method, the interpolation at a pixel  $s_0(x_0, y_0)$  is made using Eq. (1).



**Fig. 1.** Histogram of  $[\tau(s+h) - \tau(s)]^2$  for pixels separated by 1, 2 and 3 pixel units for (a) MISR and (b) MODIS pixels from 2001–2013, globally.

$$\hat{Y}(s_0) = \omega_1'Y + \omega_2'Z \quad (1)$$

where  $\omega = (\omega_1, \omega_2)$  is the weight vector associated with available MISR data vector  $Y$  and MODIS data vector  $Z$ . Weight vector  $\omega$  can be found by solving the system of equations given in Eq. (2) (Ver Hoef and Cressie, 1993).

$$\begin{bmatrix} \Sigma & X' \\ X & A \end{bmatrix} \times \begin{bmatrix} \omega \\ -\phi \end{bmatrix} = \begin{bmatrix} \sigma \\ a \end{bmatrix} \quad (2)$$

where  $X$  is a two column matrix whose first  $n$  rows are  $[1 \ 0]$  and next  $m$  rows are  $[0 \ 1]$ , with  $n$  as the number of available MISR pixels and  $m$  as the number of available MODIS pixels.  $A$  is a  $2 \times 2$  matrix of zeros,  $\phi$  is a vector with two elements to be solved along with  $\omega$  (where  $a$  is the vector  $[1 \ 0]'$ ).  $\sigma$  is a vector with element as covariance between  $Y(s_0)$  and available MODIS and MISR pixels.  $\Sigma$  is the (cross-) covariance matrix between MODIS pixels and MISR pixels. Form of  $\Sigma$  is a block matrix as shown in Eq. (3).

$$\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2' & \Sigma_3 \end{bmatrix} \quad (3)$$

where  $\Sigma_1$ ,  $\Sigma_2$  and  $\Sigma_3$  are submatrices of size  $m \times m$ ,  $m \times n$  and  $n \times n$ , respectively. Elements of  $\Sigma_1$  and  $\Sigma_3$  are filled with the help of a semivariogram function in Eq. (S3) whereas elements of  $\Sigma_2$  are filled with the help of a cross semivariogram function defined in Eq. (S5). Using Eq. (1), an estimate of AOD at those pixels can be made where MISR data are not available.

### Regression Kriging

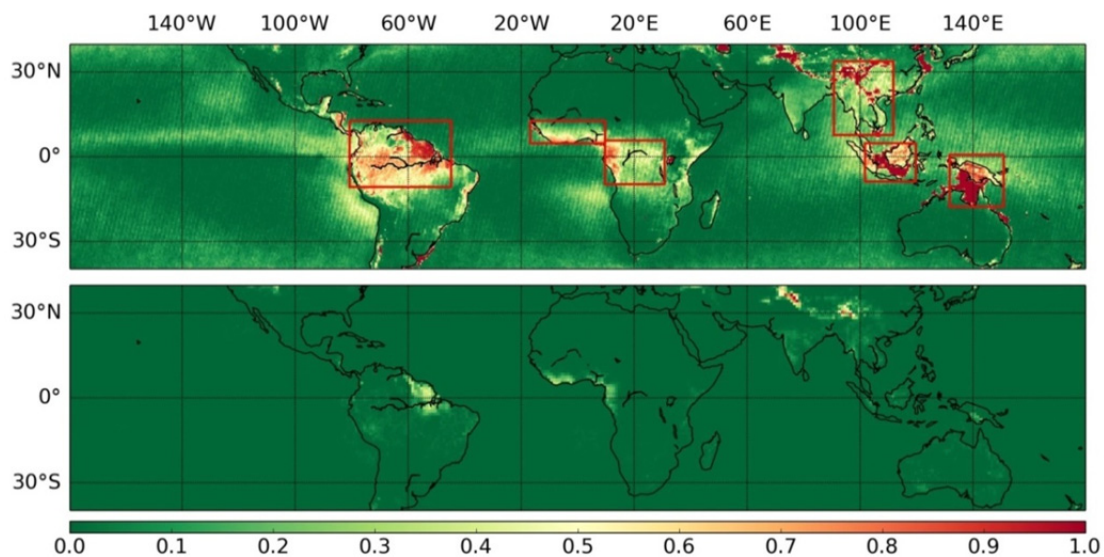
Similar to CK, the RK method is also used for spatial

interpolation of a spatial process. The covariates are used in RK in a similar way as in multivariable regression. Let  $n$  be the number of pixels with known MISR AOD data and  $m$  be the number of pixels, where the MISR AOD needs to be interpolated. The interpolated values will be given by Eq. (4).

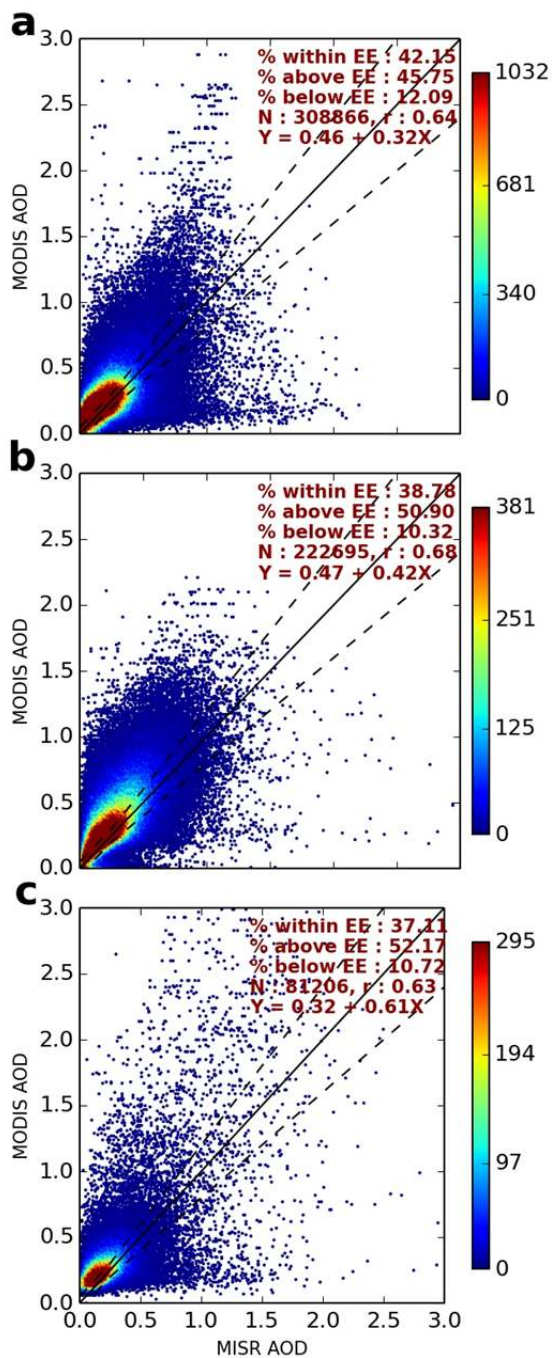
$$Y_m = B_m\beta + \eta \quad (4)$$

where  $\beta$  is a vector of coefficients corresponding to 2 covariables, unity and MODIS AOD.  $B_m$  is the matrix of size  $m \times 2$  where the first column of the matrix is filled with 1s and second column represents MODIS AOD data for  $n$  locations. The first column of the matrix  $B_m$  is associated with the constant term in the linear regression model of Eq. (4). The last term  $\eta$  in Eq. (4) represents the residual term, which represents the difference between true AOD and interpolated AOD. The least squares estimate of the expected value of predicted AOD gives rise to Regression Kriging system of equations. By solving this system of equations, Kriging weights (relative to the available MISR pixels and regression coefficients relative to available MODIS pixels) can be computed. Using these weights, the AOD can be predicted at a pixel with unknown AOD data (in our case MISR AOD is the predicted AOD).

Before proceeding with the algorithm, we have performed an initial check of the relationship between the MISR and MODIS AODs. For the analysis of gap-filled MISR AOD dataset, we have selected six regions, shown in the Fig. 2 with high fraction of invalid MISR pixels. Fig. 3 shows the comparison of the MISR and MODIS AODs for region 1 (Amazon), region 4 (central Asia) and region 5 (Indonesia/Borneo). We found that the percent of MODIS AOD falling within the expected Error Envelope (EE) of MISR for regions



**Fig. 2.** Fraction of missing AOD data of MISR (top) and MODIS (bottom) for the period 2001–2013. Areas bounded by rectangular boundaries show regions with high missing values, in terms of pixel fractions. The rectangular areas are labeled as Region 1 to Region 6 from leftmost to rightmost. Region 1 is northern part of South America (the Amazon River basin), region 2 is over western Africa, region 3 covers central Africa, region 4 is located over south-east Asia, region 5 is the area surrounding Indonesia/Borneo and region 6 is the area surrounding New Guinea.



**Fig. 3.** Inter-comparison between MISR and MODIS AODs for the period 2001–2013 over (a) the Amazon (81°W–45°W; 13°N–11°S), (b) central Asia (91°E–112°E; 8°N–34°N) and (c) Indonesia/Borneo (102°E–120°E; 9°S–5°N). Dashed lines are the Envelope of Expected Error in terms of maximum of 0.05 or 20% × AERONET AOD (minimum of –0.05 or –20% × AERONET AOD).

1, 4 and 5 are 42%, 39% and 37%, respectively. It has been also noted by Kahn *et al.* (2011) that the AODs from the MISR and MODIS fall outside the EE when compared to the AERONET data. The percent of MODIS AOD falling above the EE of MISR for regions 1, 4 and 5 are 46%, 51% and 52%, respectively, implying that the MODIS

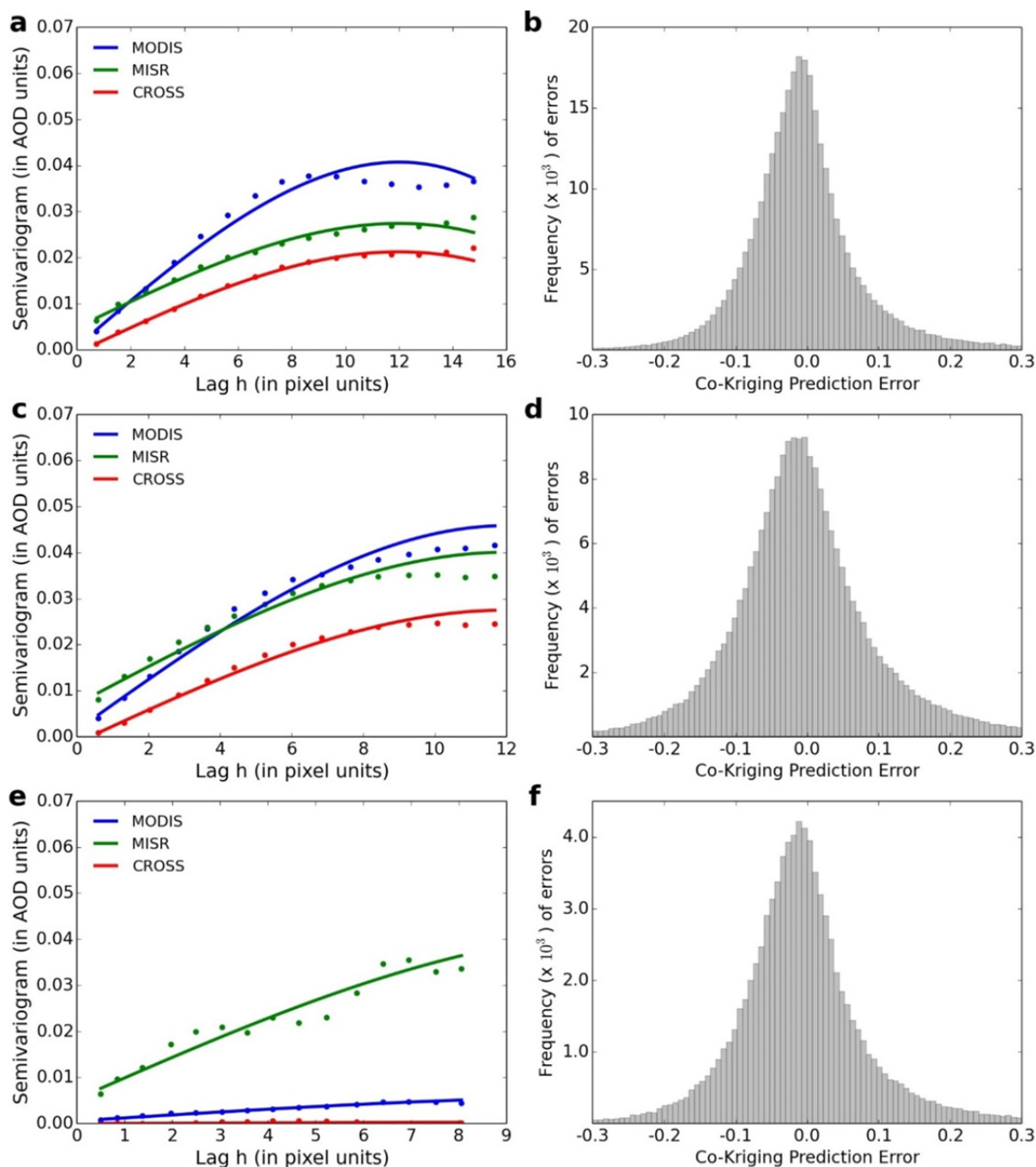
AODs are higher than the MISR AODs. However, we note that the correlation coefficient ( $r$ ) between the MISR and MODIS AODs in these regions are 0.64, 0.68 and 0.63, respectively. Therefore, there may be a linear relationship between MISR and MODIS AODs.

The Regression Kriging uses linear relationship between MISR and MODIS AODs. Higher correlation between MISR and MODIS is an indicator that the MODIS data can be used as a co-variable in Regression Kriging. In addition, there is a spatial relationship in the MISR and MODIS AOD products. Fig. 4 shows the semivariogram of the MISR and MODIS AODs for the three regions. A semivariogram represents the spatial relationship in terms of the difference of AOD at two pixels. The three parameters which determine the fitted semivariogram are nugget, sill and range. Over Amazon, the nugget, sill and range for MODIS semivariogram are 0.0007, 0.04 and 12.0, respectively. For MISR, the nugget, sill and range are, 0.004, 0.02 and 12.0, respectively. As found from the Fig. 4(a), the sill associated with MODIS data is larger than that of MISR. The range for both the semivariogram is same. The expected value of the difference follows an increasing pattern which can be fitted using a spherical function (Eq. (S4)). The CK and RK algorithms use the variance and co-variance between the AODs at different pixels. The variance and co-variance are computed by using the semivariogram model. Since there is a good fit of semivariogram model to the experimental semivariogram, the CK and RK methods can be used for filling gaps in the MISR AOD dataset.

## RESULTS AND DISCUSSION

### Cross Validation

The prediction accuracy of the CK and RK methods are evaluated by cross validating MISR AOD on a regional basis. For a given region and time, all the valid pixels are randomly divided into ten groups. The AOD values at all the pixels in a group are then computed using the Kriging method, taking the valid data of other nine groups as input. Similarly, pixels in the other groups are also interpolated. We have used 72 regions for a total of 156 months for the spatial interpolation. The results of cross validation are shown in the Table 2, where the RMSE in cross validation gives the measure of accuracy of the CK and RK predictors. The RMSE of CK is 0.05. To verify the strength of the CK prediction, six regions shown in Fig. 2 are selected, where the fraction of missing data is high. Table 1 shows the percentage of missing values in MISR and MODIS data. It is observed that all regions have significantly larger fraction of missing AOD data from MISR compared to MODIS. Therefore, fewer samples are available for the estimation of semivariogram and prediction of missing values at a pixel. For this reason, high predicted RMSE, in these regions are found. On the other hand, Region 6 (around New Guinea) has smaller RMSE compared to others regions. It is noted here that the fraction of missing MISR data in this region is high, however the presence of MODIS data is also high compared to other regions. The accuracy of prediction using Kriging methods depends on the accuracy of the fitted



**Fig. 4.** (Left plot) Semivariogram model and experimental semivariogram plot of MISR and MODIS AOD product for March 2001, (a) for the Amazon region, (c) central Asia region and (e) for Indonesia/Borneo. The semivariogram plots also show cross semivariogram (red color) between MISR and MODIS AOD, where symbols denote the measure of experimental semivariogram at different lag shown in X-axis. The fitted lines to the symbols are estimated using spherical semivariogram model. (Right plot) Histogram of cross validated prediction error in MISR using Ordinary Co-Kriging with MODIS as co-variable, (b) at the Amazon, (d) central Asia and (f) Indonesia/Borneo.

theoretical semivariogram. This accuracy can be computed by the mean squared difference between the selected theoretical semivariogram model and experimental semivariogram. Table 3 shows the accuracy of the fitted semivariogram models (spherical semivariogram in this study). Column two and three in Table 3 represent the accuracy of the fitted spherical semivariogram model for MODIS and

MISR AOD data. In the fourth column, the accuracy of the fitted cross semivariogram is also provided in terms of mean squared difference between MODIS and MISR AOD data. Smaller values in the mean squared difference represent higher accuracy and higher values represent lower accuracy of the fitted semivariogram model. We note that region 6 (around New Guinea) has highest accuracy of the MODIS,

**Table 2.** Cross validation parameters of CK and RK predictors for regions shown in Fig. 2. The bias is given by the expected value of the difference of predicted AOD data and measured AOD data. The RMSE is root mean square error of the predicted AOD with respect to measured AOD data.  $r$  is the correlation coefficient between predicted AOD and measured AOD. Globally, cross validation is performed for the year 2001 only.

	Total Pixels	CK			Total Pixels	RK		
		Bias	RMSE	$r$		Bias	RMSE	$r$
Global	$4.1 \times 10^5$	$5.0 \times 10^{-6}$	0.05	0.91	$4.1 \times 10^5$	$6.5 \times 10^{-6}$	0.05	0.91
Region 1	$4.1 \times 10^5$	$7.4 \times 10^{-5}$	0.10	0.72	$3.1 \times 10^5$	$7.3 \times 10^{-4}$	0.08	0.83
Region 2	$1.0 \times 10^5$	$4.7 \times 10^{-4}$	0.12	0.82	$1.0 \times 10^5$	$1.3 \times 10^{-5}$	0.09	0.91
Region 3	$1.5 \times 10^5$	$-2.5 \times 10^{-4}$	0.14	0.76	$1.4 \times 10^5$	$1.4 \times 10^{-4}$	0.12	0.83
Region 4	$2.3 \times 10^5$	$5.9 \times 10^{-5}$	0.12	0.76	$2.2 \times 10^5$	$1.0 \times 10^{-4}$	0.09	0.87
Region 5	$8.8 \times 10^4$	$-1.7 \times 10^{-4}$	0.13	0.71	$8.1 \times 10^4$	$7.9 \times 10^{-4}$	0.11	0.81
Region 6	$1.4 \times 10^5$	$-2.9 \times 10^{-5}$	0.07	0.59	$1.3 \times 10^5$	$2.0 \times 10^{-4}$	0.06	0.72

**Table 3.** Semivariogram fitting and AOD prediction error for regions showed in Fig. 2. Second and third columns show expected error in the fitting of spherical semivariogram model to experimental semivariograms of MODIS and MISR AOD data for the time period of January 2001 to December 2013. The fourth column shows the semivariogram fitting error of cross experimental semivariogram between MODIS and MISR. The fifth column shows the expected value of prediction error using CK of MISR AOD at a pixel.

Regions	MODIS AOD semivariogram error	MISR AOD semivariogram error	Cross variogram error	Prediction error
Global	NA	NA	NA	0.05
Region 1	$5.7 \times 10^{-3}$	$1.8 \times 10^{-3}$	$2.8 \times 10^{-3}$	0.10
Region 2	$2.9 \times 10^{-3}$	$3.7 \times 10^{-3}$	$1.8 \times 10^{-3}$	0.12
Region 3	$3.0 \times 10^{-3}$	$2.7 \times 10^{-3}$	$1.7 \times 10^{-3}$	0.14
Region 4	$4.0 \times 10^{-3}$	$3.8 \times 10^{-3}$	$2.6 \times 10^{-3}$	0.12
Region 5	$1.4 \times 10^{-2}$	$3.9 \times 10^{-3}$	$4.6 \times 10^{-3}$	0.13
Region 6	$4.4 \times 10^{-4}$	$1.3 \times 10^{-3}$	$2.2 \times 10^{-4}$	0.07

MISR and cross semivariogram models, which is reflected in the lower prediction error of MISR AOD. On the other hand, the prediction error in region 3 (central Africa) is the largest among six regions under this study, associated with the large error of the fitted semivariogram models of MODIS AOD, MISR AOD and the cross semivariogram between MODIS and MISR. For region 2 (western Africa), the percent of missing MISR AOD data, for which MODIS has valid AOD data (87.27%), is less compared to region 5- Indonesia/Borneo (98.86%). Therefore, the prediction error (0.12) in region 2 is high, albeit the error in fitted semivariogram models in region 2 being smaller than region 5. In region 2 and region 4 (southeast Asia), the errors in prediction and fitted semivariogram are similar. Therefore, we can conclude that the prediction error depends on two factors, i.e., accuracy of the fitted semivariogram model and percent of missing pixels of MISR and MODIS AOD.

The regional and global behavior of the prediction error are also discussed temporally, based on results from 156 months of data (2001–2013). The prediction error varies for different periods for a given region, which depends on the fitting of semivariogram model. The prediction procedure for March 2001 in region 1 (Amazon basin) is discussed here in more detail as the Amazon region has highest percentage of missing MISR AOD data, among all selected regions. In addition, the regional analysis of the gap-filled MISR AOD product is also performed over region 4 (central Asia) and region 5 (Indonesia/Borneo). In March 2001, 46% of the total

area in region 1 has missing data while region 4 and region 5 have 20% and 64% missing data, respectively. Fig. 4 shows the experimental and fitted spherical semivariograms of MODIS and MISR, together with the cross semivariogram between MODIS and MISR over region 1 (Fig. 4(a)), region 4 (Fig. 4(c)) and region 5 (Fig. 4(e)). The error in the fitted cross semivariogram between MODIS and MISR AOD is  $7.7 \times 10^{-4}$ , whereas the errors in the fitted semivariograms of MODIS and MISR are  $2.6 \times 10^{-3}$  and  $1.1 \times 10^{-3}$ , respectively, over the Amazon. Thus, the errors in MODIS and MISR AOD semivariogram fitting for March 2001 are 2 times smaller than that for the entire study period (2001–2013). However, the percent of invalid MISR pixels with valid MODIS pixels is (85.7%) less than that for the entire study period (93.01). The expected value of the prediction error over the Amazon region is therefore large, estimated to be 0.12.

Over the region 4, the error in the fitted cross semivariogram between MODIS and MISR AOD is  $1.3 \times 10^{-3}$ , whereas the errors in the fitted semivariograms of MODIS and MISR are  $2.5 \times 10^{-3}$ . Thus, the errors in cross semivariogram and MISR AOD semivariogram fitting for March 2001 are 3 and 1.5 times, respectively, smaller than that for the entire study period (2001–2013). The expected value of the prediction error over region 4 is therefore smaller, estimated to be 0.09. Over region 5, the percent of MISR pixels is higher (64%) compared to the entire period (50%). The expected value of the prediction error over



region 4 is therefore higher, estimated to be 0.14.

We also calculated the bias  $E(Y - \hat{Y})$  present in the prediction using the residuals of the cross validation data, where the residual is defined as the difference between measured AOD and interpolated AOD. The expected value of the residual is called the bias present in the predictor. Fig. 4(b) shows the distribution of residuals, which indicates a normally distributed variable with mean close to zero and standard deviation (SD) 0.076 over the Amazon. Therefore, the bias present in the interpolated data using CK will be  $\sim 0 \pm 0.076$ . Over regions 4 and 5, the residuals follow normal distribution with mean close to zero and SD 0.093 and 0.085, respectively. The correlation coefficient ( $r$ ) between predicted AOD and measured MISR AOD is also computed, which is found to be greater than 0.59, as shown in Table 2.

The filled-MISR data along with the original MISR and MODIS AOD data are also compared with AERONET AOD data, globally and for the 8 regions (different from Fig. 2), as shown in Table 4. Table 4 shows the RMSE of the four aerosol products with AERONET data observations between 10:00AM and 11:00AM, i.e., MODIS, MISR, CK-filled MISR and RK-filled MISR datasets. Globally, the RMSE of MISR and filled-MISR data are higher than that of MODIS. As indicated in Table 4, over northern India, MISR RMSE is significantly greater than that of MODIS. The northern parts of India are largely covered by clouds for close to four months, during the monsoon season. The persistent cloud cover together with smaller swath width results in significantly fewer samples of MISR retrievals towards the generation of Level-3 monthly-mean AOD product, resulting in larger fraction of missing data in MISR Level-3 product than MODIS data. Globally, the accuracy of the gap-filled MISR data is increased by a very small margin of 0.0004 (0.25%). Over eastern Asia and northern India, the accuracy of the gap-filled MISR was found to increase by 2.87% and 5.19%, respectively.

#### Comparison with Regression Kriging

We have also used the RK method to fill missing data in MISR AOD where MODIS AOD data are present. Table 2 shows the results related to the cross validation of RK and CK methods. The RMSE of the prediction using RK is

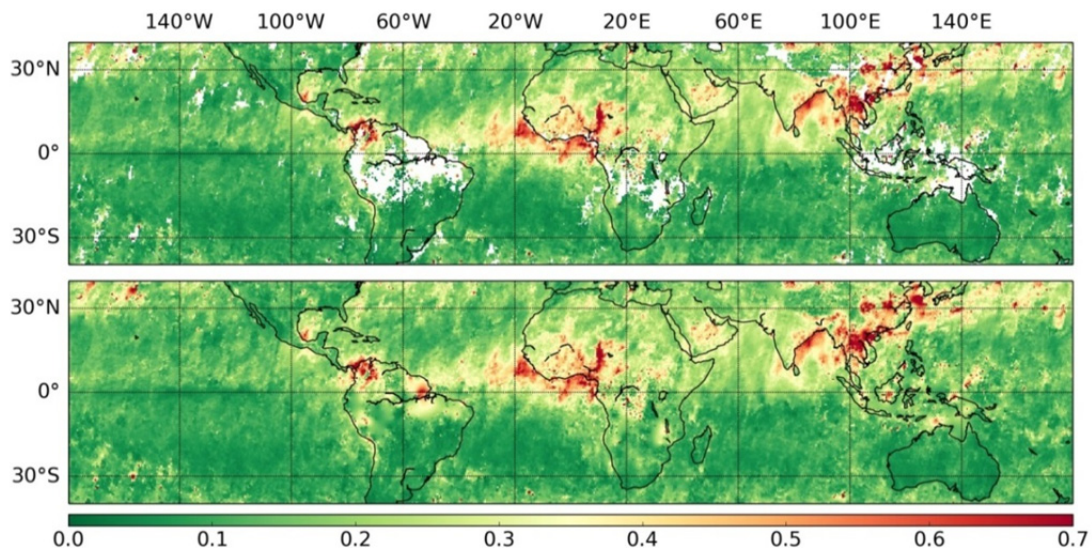
0.05, similar to the RMSE of the CK method. The residuals of RK also follow normal distribution with mean close to zero and standard deviation close to 0.05. Since the expected value of the residuals show bias of the predictor, globally, both RK and CK predictors have positive bias close to zero (for the year 2001). However, at regional scale, where MISR has higher density of missing data, we find that the RK predictor has larger bias as compared to the CK predictor bias. Additionally,  $r$  between the predicted and original MISR AOD data is low for CK, as compared to RK method. The computations related to the RK method are limited to pixels with valid MODIS AOD.

Fig. 5 shows the spatial distribution of MISR AOD data together with gap-filled MISR AOD for March 2001. The missing pixels are mainly concentrated over the Amazon, southern Africa, and the regions surrounding Indonesia/Borneo and New Guinea. The gap-filled AOD map does not have any gaps and apparent discontinuity in comparison to the original MISR AOD. In addition, those pixels with valid data in the original spatial distribution of AOD are preserved in the new gap-filled dataset with missing values filled using CK. The CK prediction method has a small positive bias and a prediction error of 0.051 for March 2001, which is similar to the global prediction error (for the year 2001).

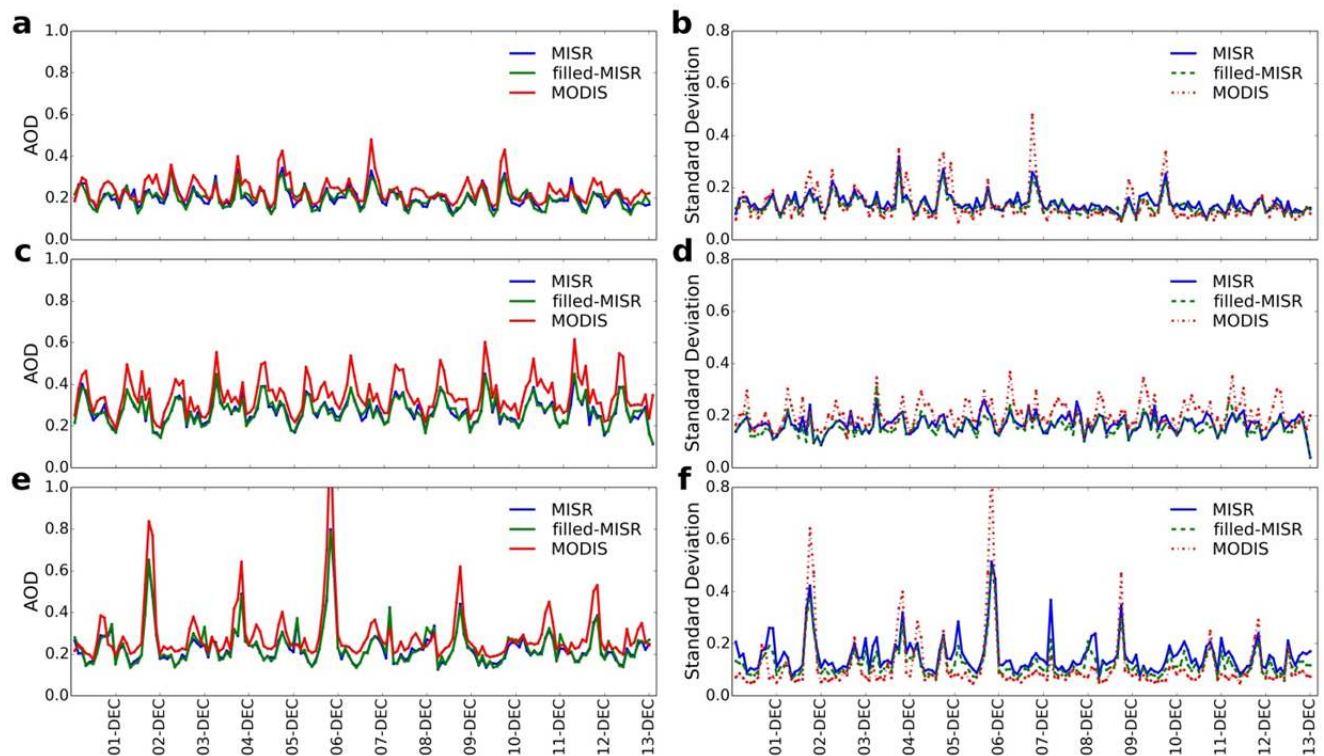
The interannual variability time series of MODIS, MISR and filled-MISR AOD (using CK) datasets for the Amazon region, central Asia and Indonesia/Borneo are also shown (Figs. 6(a), 6(c) and 6(e)) for the period 2001–2013. The interannual variability of filled-MISR data follows very similar pattern compared to that of MISR AOD data. The  $r$  of the two time series in the region 1, region 4 and region 5 are 0.91, 0.97 and 0.98, respectively, which are slightly less than the global  $r$  of 0.99. Over these regions, the fraction of missing data is larger than the global fraction of missing MISR AOD data; and the filling of larger number of missing data leads to greater departure from the regional mean of AOD data. Therefore,  $r$  between MISR AOD and filled-MISR AOD in the Amazon is less than global  $r$ . We find that for some periods, the absolute difference between these two products is on the higher side, i.e., above 0.02, where the missing pixels count in MISR AOD product is larger

**Table 4.** The RMSE of different satellite AOD datasets with respect to AERONET. Monthly time series of a satellite AOD product at the pixel containing the AERONET station is used to compute RMSE with the monthly time series of the AERONET station AOD. AOD measurements between 10.00AM to 11.000AM local-time at an AERONET station are only considered to calculate the monthly time series. n represents total number of samples used for computation. Total number of samples for CK and RK method are equal.

Region	MODIS (n)	MISR (n)	MISR filled (n) (CK)	MISR filled (n) (RK)
Global	0.128 (41375)	0.143 (38372)	0.143 (41496)	0.161
Northern India	0.210 (2614)	0.317 (2065)	0.300 (2652)	0.336
North America	0.072 (15719)	0.065 (15056)	0.065 (15756)	0.085
South America	0.073 (1248)	0.065 (1204)	0.065 (1248)	0.167
South Africa	0.082 (624)	0.052 (591)	0.052 (624)	0.104
Western Africa	0.140 (2025)	0.159 (1978)	0.158 (2028)	0.155
Arabian Peninsula	0.127 (3120)	0.101 (3112)	0.100 (3120)	0.118
Eastern Asia	0.218 (7009)	0.223 (5914)	0.217 (7020)	0.230
Australia	0.029 (1091)	0.039 (1081)	0.039 (1092)	0.100



**Fig. 5.** Aerosol Optical Depth of original MISR (top) and filled-MISR dataset (using Ordinary Co-kriging) (bottom) for March 2001. White shading in the top image shows locations where MISR AOD data not present.



**Fig. 6.** (Left) Time series of MODIS AOD, MISR AOD and filled-MISR AOD with CK. (Right) Time series of standard deviation of MODIS AOD, MISR AOD and filled-MISR AOD dataset, with CK for the period 2001–2013 (a, b) over the Amazon region, (c, d) central Asia and (e, f) Indonesia/Borneo.

than 30%. For instance, in January 2001, the difference in the regional mean of filled-AOD and MISR AOD is 0.031, with 50% pixels having missing MISR AOD data over Amazon. Since it is difficult to estimate in advance whether a missing pixel has high or low AOD, the difference of the regional mean of filled AOD product and original MISR AOD product is equally difficult to assess. The difference in regional mean of filled-AOD and MISR AOD can be

explained by the bias obtained from CK method for each month. For example, the regional-mean filled AOD (0.22) for January 2001 is higher than the original MISR AOD (0.19) data, where the expected CK prediction bias at a pixel for this period is  $-0.06$ . This indicates that the expected predicted value is larger by  $AOD = 0.06$  compared to the observed AOD. Therefore, for the time series shown in Fig. 6(a), the higher regional-mean filled AOD (for January

2001) may be due to a negative bias of CK predictor. Similarly for May 2002, the lower value in regional mean filled AOD product is likely due to a positive bias of 0.12.

We also computed regional standard deviation (SD) of the AOD for the Amazon basin, central Asia and Indonesia/Borneo regions. It is observed that the expected value for the SD of MISR AOD dataset over regions 1, 4 and 5 are 0.14, 0.17 and 0.15, respectively, i.e., higher than the expected SD of the gap-filled MISR data, which are 0.13, 0.16 and 0.13, respectively. Figs. 6(b), 6(d) and 6(f) show the time series of SD for the three regions where it is evident that the SD of filled-MISR AOD datasets is lower than the original MISR AOD data. With the reduction in spatial variability of MISR-AOD dataset, we believe that the analysis related to the time series of AOD dataset (including trend analysis over the Amazon) will be better representative of the region.

## CONCLUSIONS

In this paper, an application of geostatistical methods is demonstrated to produce a gap-filled aerosol dataset for Level-3 monthly-mean MISR AOD at  $0.5^\circ \times 0.5^\circ$  spatial resolution. The spatially complete AOD dataset can be especially used in studies where the presence of missing data are undesirable, such as for global and regional aerosol spatio-temporal variability and trend analysis. Two Kriging techniques (Co-Kriging and Regression Kriging) are used to fill MISR AOD data, with valid AOD from MODIS data. To assess the quality of the spatially-filled data set, cross validation is performed. The cross validation results show that the prediction error follows a normal distribution with mean close to zero. The RMSE of the predicted data at pixels with valid AOD retrievals (in cross validation, we predict AOD at pixels with valid retrievals) using both CK and RK is approximately 0.05. Across the six selected regions around the globe, where large fraction of missing AOD data from MISR is widespread, biases from both CK and RK have small values. However, at regional scale, predicted AOD using the RK method has larger biases as compared to the predicted AOD using CK. In general, at regional scale, the  $r$  between the observed and predicted MISR AOD using RK is higher than that using CK. However, the RK method can be used only at pixels with valid MODIS AOD. Furthermore, the prediction using Kriging method on regional scale shows that as the fraction of missing MISR AOD data grows (combined with the reduced presence of MODIS data), the accuracy of the Kriging predictors decreases. Additionally, the comparison of gap-filled MISR and MISR AOD with AERONET data shows that the deviation of filled-MISR data from AERONET data is smaller than the original MISR data. These results suggest that the CK method is an effective interpolation algorithm for filling data gaps in MISR AOD data, using co-existing MODIS data as a co-variable.

## ACKNOWLEDGMENTS

Authors thank MISR and MODIS teams for the

provision of aerosol datasets used in this paper.

We also gratefully acknowledge the principal investigators of AERONET and their team for establishing and maintaining the ground-based aerosol dataset. The gap-filled MISR AOD datasets described in the paper can be presently obtained by requesting the authors. We thank three anonymous reviewers for their insightful comments which helped improve an earlier version of the manuscript.

## SUPPLEMENTARY MATERIAL

Supplementary data associated with this article can be found in the online version at <http://www.aaqr.org>.

## REFERENCES

- Aalto, J., Pirinen, P., Heikkinen, J. and Veninen, A. (2013). Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models. *Theor. Appl. Climatol.* 112: 99–111.
- Buytaert, W., Celleri, R., Willems, P., Bivre, B.D. and Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *J. Hydrol.* 329: 413–421.
- Carrera-Hernandez, J. and Gaskin, S. (2007). Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. *J. Hydrol.* 336: 231–249.
- Chatterjee, A., Michalak, A.M., Kahn, R.A., Paradise, S.R., Braverman, A.J. and Miller, C.E. (2010). A geostatistical data fusion technique for merging remote sensing and ground-based observations of aerosol optical thickness. *J. Geophys. Res.* 115: D20207.
- Chiles, J.P. and Delfiner, P. (2009). *Geostatistics: Modeling Spatial Uncertainty*, Vol. 497, John Wiley and Sons.
- Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *J. R. Stat. Soc. B* 70: 209–226.
- Cressie, N.A.C. (1993). *Statistics for Spatial Data*, Revised Edition Edn, Wiley.
- de Meij, A., Pozzer, A. and Lelieveld, J. (2012). Trend analysis in aerosol optical depths and pollutant emission estimates between 2000 and 2009. *Atmos. Environ.* 51: 75–85.
- Djuric, N., Kansakar, L. and Vucetic, S. (2016). Semi-supervised combination of experts for aerosol optical depth estimation. *Artif. Intell.* 230: 1–13.
- Fuentes, M. and Raftery, A.E. (2005). Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics* 61: 36–45.
- Gandin, L.S. and Hardin, R. (1965). *Objective Analysis of Meteorological Fields*, Vol. 242, Israel Program for Scientific Translations Jerusalem.
- Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*, Oxford University Press.
- Holben, B., Eck, T., Slutsker, I., Tanr, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenue, F., Jankowiak, I. and Smirnov, A. (1998).

- AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.* 66: 1–16.
- Hsu, N.C., Gautam, R., Sayer, A.M., Bettenhausen, C., Li, C., Jeong, M.J., Tsay, S.C. and Holben, B.N. (2012). Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010. *Atmos. Chem. Phys.* 12: 8037–8053.
- Hsu, N.C., Jeong, M.J., Bettenhausen, C., Sayer, A.M., Hansell, R., Seftor, C.S., Huang, J. and Tsay, S.C. (2013). Enhanced Deep Blue aerosol retrieval algorithm: The second generation. *J. Geophys. Res.* 118: 9296–9315.
- Kahn, R.A., Gaitley, B.J., Garay, M.J., Diner, D.J., Eck, T.F., Smirnov, A. and Holben, B.N. (2010). Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. *J. Geophys. Res.* 115: D23209.
- Kahn, R.A., Garay, M.J., Nelson, D.L., Levy, R.C., Bull, M.A., Diner, D.J., Martonchik, J.V., Hansen, E.G., Remer, L.A. and Tanré, D. (2011). Response to “Toward unified satellite climatology of aerosol properties. 3. MODIS versus MISR versus AERONET”. *J. Quant. Spectrosc. Radiat. Transfer* 112: 901–909.
- Kaplan, A., Kushnir, Y., Cane, M.A. and Blumenthal, M.B. (1997). Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *J. Geophys. Res.* 102: 27835–27860.
- Kinne, S. (2009). Remote sensing data combinations: Superior global maps for aerosol optical depth, In *Satellite Aerosol Remote Sensing over Land*, Kokhanovsky, A. and de Leeuw, G. (Eds), Springer Praxis Books, Springer Berlin Heidelberg, pp. 361–381.
- Levy, R.C., Mattoo, S., Munchak, L.A., Remer, L.A., Sayer, A.M., Patadia, F. and Hsu, N.C. (2013). The collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech.* 6: 2989–3034.
- Li, J., Carlson, B.E. and Lacis, A.A. (2013). Application of spectral analysis techniques in the intercomparison of aerosol data: 1. an EOF approach to analyze the spatial-temporal variability of aerosol optical depth using multiple remote sensing data sets. *J. Geophys. Res.* 118: 8640–8648.
- Murphy, D.M. (2013). Little net clear-sky radiative forcing from recent regional redistribution of aerosols. *Nat. Geosci.* 6: 258–262.
- Nguyen, H., Cressie, N. and Braverman, A. (2012). Spatial statistical data fusion for remote sensing applications. *JASA* 107: 1004–1018.
- Nguyen, H., Katzfuss, M., Cressie, N. and Braverman, A. (2014). Spatio-temporal data fusion for very large remote sensing datasets. *Technometrics* 56: 174–185.
- Pebesma, E.J. (2004). Multivariable geostatistics in S: The gstat package. *Comput. Geosci.* 30: 683–691.
- Ramanathan, V., Crutzen, P.J., Kiehl, J.T. and Rosenfeld, D. (2001). Aerosols, climate, and the hydrological cycle. *Science* 294: 2119–2124.
- Rosenfeld, D., Lohmann, U., Raga, G.B., O’Dowd, C.D., Kulmala, M., Fuzzi, S., Reissell, A. and Andreae, M.O. (2008). Flood or drought: How do aerosols affect precipitation? *Science* 321: 1309–1313.
- Ruiz-Arias, J., Dudhia, J., Lara-Fanego, V. and Pozo-Vzquez, D. (2013). A geostatistical approach for producing daily Level-3 MODIS aerosol optical depth analyses. *Atmos. Environ.* 79: 395–405.
- Sayer, A.M., Munchak, L.A., Hsu, N.C., Levy, R.C., Bettenhausen, C. and 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.* 119: 13965–13989.
- Shrestha, P. and Barros, A.P. (2010). Joint spatial variability of aerosol, clouds and rainfall in the Himalayas from satellite data. *Atmos. Chem. Phys.* 10: 8305–8317.
- Ver Hoef, J. and Cressie, N. (1993). Multivariable spatial prediction. *Math. Geol.* 25: 219–240.
- Xu, H., Guang, J., Xue, Y., de Leeuw, G., Che, Y., Guo, J., He, X. and Wang, T. (2015). A consistent aerosol optical depth (AOD) dataset over mainland China by integration of several AOD products. *Atmos. Environ.* 114: 48–56.
- Zhang, C., Li, W. and Travis, D.J. (2009). Restoration of clouded pixels in multispectral remotely sensed imagery with cokriging. *Int. J. Remote Sens.* 30: 2173–2195.
- Zhang, J. and Reid, J.S. (2010). A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products. *Atmos. Chem. Phys.* 10: 10949–10963.

Received for review, February 20, 2016

Revised, October 15, 2016

Accepted, March 28, 2017