



*Supplement of*

## **Synergistic Fusion of Aerosol Optical Depth over India from multi-sensor satellite retrievals with ground-based measurements**

**Shiba Shankar Gouda et al.**

*Correspondence to:* Mukunda M. Gogoi (dr\_mukunda@vssc.gov.in)

The copyright of individual parts of the supplement might differ from the article licence.

## Supplementary Document

**Table S1:** Error statistic of SVR models; (i) MODIS and Ground and (ii) MISR and Ground.

<b>Sensor</b>	<b>ML-function</b>	<b>R<sup>2</sup></b>	<b>R</b>	<b>RMSE</b>	<b>MAE</b>
<b>MODIS</b>	Train	0.48	0.7	0.22	0.15
	Test	0.72	0.85	0.15	0.12
<b>MISR</b>	Train	0.59	0.76	0.20	0.15
	Test	0.60	0.84	0.13	0.12

**Table S2:** Spatio-temporal sensitivity test statistics regarding the instantaneous satellite overpass (for specific domains) against ground observations (for specific time window) in the month of Nov-2021.

Spatio-temporal domain	R Pearson Correlation coefficient	RMSE	MAE	N (Colocation Points)	$\Delta$ RMSE (RMSE <sub>current cell</sub> -RMSE <sub>previous cell</sub> )	$\Delta$ MAE (MAE <sub>current cell</sub> -MAE <sub>previous cell</sub> )	$\Delta$ R (R <sub>current cell</sub> -R <sub>previous cell</sub> )
<b>MISR</b>							
0.2deg-30min	0.94	0.20	0.12	24			
0.2deg-60min	0.93	0.20	0.12	25	-0.0004	0.0003	-0.0008
0.2deg-120min	0.94	0.19	0.12	26	-0.0035	-0.0026	0.0060
<b>0.5deg-30min</b>	<b>0.93</b>	<b>0.18</b>	<b>0.10</b>	<b>31</b>			
0.5deg-60min	0.92	0.18	0.10	33	-0.0022	-0.0038	-0.0116
0.5deg-120min	0.91	0.17	0.10	36	-0.0034	-0.0008	-0.0052
1deg-30min	0.84	0.22	0.13	38			
1deg-60min	0.82	0.22	0.13	41	0.0032	0.001	-0.0232
1deg-120min	0.83	0.21	0.13	46	-0.0158	-0.0053	0.0159
2 deg-30min	0.83	0.24	0.15	48			
2 deg-60min	0.82	0.24	0.15	52	-0.0005	0.0025	-0.0164
2 deg-120min	0.79	0.22	0.14	59	-0.0150	-0.0070	-0.0244
<b>MODIS</b>							
0.2deg-30min	0.93	0.14	0.10	110			
0.2deg-60min	0.93	0.14	0.10	114	-0.0003	0.000931	-0.0004
0.2deg-120min	0.93	0.14	0.10	118	-0.0023	0.001017	0.0020
<b>0.5deg-30min</b>	<b>0.93</b>	<b>0.13</b>	<b>0.10</b>	<b>148</b>			
0.5deg-60min	0.93	0.13	0.10	157	-0.0028	0.000615	0.0003
0.5deg-120min	0.93	0.13	0.09	165	-0.0045	-0.00233	0.0035
1deg-30min	0.93	0.16	0.10	175			
1deg-60min	0.93	0.16	0.10	189	-0.0059	-0.00222	0.0010
1deg-120min	0.93	0.16	0.10	202	-0.0033	-0.00158	0.0002
2 deg-30min	0.89	0.20	0.12	187			
2 deg-60min	0.89	0.19	0.12	204	-0.0105	-0.00415	0.0026
2 deg-120min	0.90	0.19	0.11	219	-0.0084	-0.00565	0.0033

**Table S3:** Variogram used in sensitivity of **Fig. S24** analysis

<b>Sensor</b>	<b>Model</b>	<b>Nugget</b>	<b>sill</b>	<b>Range(in km)</b>
<b>MODIS</b>	Matheron	0	0.030028	548.35
<b>MISR</b>	Exponential	0.000022	0.014509	638.76

**Table S4:** MODIS, MISR and fused AOD at different regions of India during Winter (January).

	MODIS	MISR	FUSED	MODIS	MISR	FUSED	MODIS	MISR	FUSED
	2012			2016			2021		
<b>NE</b>	0.38±0.21	0.25±0.15	0.42±0.19	0.35±0.21	0.25±0.16	0.27±0.19	0.29±0.18	0.18±0.09	0.28±0.19
<b>IGP</b>	0.70±0.18	0.48 ±0.12	0.53±0.13	0.99±0.31	0.60±0.14	0.87±0.16	0.83±0.24	0.49±0.12	0.74±0.23
<b>Peninsular</b>	0.27±0.1	0.26 ±0.05	0.32±0.05	0.38±0.10	0.35±0.08	0.52±0.13	0.50±0.21	0.46±0.16	0.74±0.19
<b>Central</b>	0.41 ±0.08	0.31 ±0.06	0.39 ± 0.07	0.40±0.12	0.31±0.12	0.52±0.11	0.57±0.14	0.38±0.16	0.59±0.14
<b>NW</b>	0.29±0.11	0.29 ±0.07	0.38±0.07	0.32±0.10	0.29±0.07	0.48±0.08	0.35±0.11	0.30±0.10	0.40±0.10

**Table S5:** MODIS, MISR and fused AOD at different regions of India during Pre-monsoon (May).

	<b>MODIS</b>	<b>MISR</b>	<b>FUSED</b>	<b>MODIS</b>	<b>MISR</b>	<b>FUSED</b>	<b>MODIS</b>	<b>MISR</b>	<b>FUSED</b>
	<b>2012</b>			<b>2016</b>			<b>2021</b>		
<b>NE</b>	0.41±0.16	0.31±0.13	0.32±0.13	0.44±0.19	0.30±0.12	0.27±0.11	0.51±0.20	0.36±0.13	0.46±0.22
<b>IGP</b>	0.65±0.15	0.53±0.09	0.58±0.14	0.80±0.19	0.61±0.13	0.59±0.15	0.72±0.18	0.53±0.10	0.74±0.22
<b>Peninsular</b>	0.40±0.10	0.42 ±0.09	0.43±0.12	0.30 ±0.07	0.35±0.06	0.46±0.07	0.42±0.16	0.39±0.14	0.51±0.22
<b>Central</b>	0.36 ±0.09	0.44 ±0.06	0.39±0.08	0.35±0.11	0.47±0.06	0.57±0.07	0.38±0.09	0.40±0.07	0.49±0.13
<b>NW</b>	0.34 ±0.08	0.44 ±0.06	0.45±0.08	0.33 ±0.14	0.44 ±0.06	0.52±0.05	0.29±0.16	0.49±0.11	0.54±0.15

**Table S6:** MODIS, MISR and fused AOD at different regions of India during Post-monsoon (November).

	<b>MODIS</b>	<b>MISR</b>	<b>FUSED</b>	<b>MODIS</b>	<b>MISR</b>	<b>FUSED</b>	<b>MODIS</b>	<b>MISR</b>	<b>FUSED</b>
	<b>2012</b>			<b>2016</b>			<b>2021</b>		
<b>NE</b>	0.15±0.11	0.12±0.09	0.22±0.11	0.09±0.05	0.09±0.07	0.11±0.04	0.15±0.09	0.13±0.07	0.08±0.06
<b>IGP</b>	0.72±0.18	0.47±0.08	0.61±0.20	0.69±0.19	0.45±0.11	0.60±0.23	0.74±0.18	0.51±0.10	0.80±0.25
<b>Peninsular</b>	0.41±0.11	0.36±0.08	0.45±0.17	0.41±0.10	0.34±0.09	0.47±0.14	0.43±0.14	0.36±0.13	0.50±0.22
<b>Central</b>	0.44±0.09	0.37±0.07	0.40±0.09	0.42±0.11	0.34±0.08	0.36±0.09	0.55±0.09	0.42±0.08	0.59±0.16
<b>NW</b>	0.35±0.11	0.33±0.09	0.31±0.10	0.32±0.09	0.29±0.05	0.43±0.10	0.41±0.10	0.34±0.07	0.40±0.12

**Table S7:** List of ARFINET stations and respective Latitudes and Longitudes.

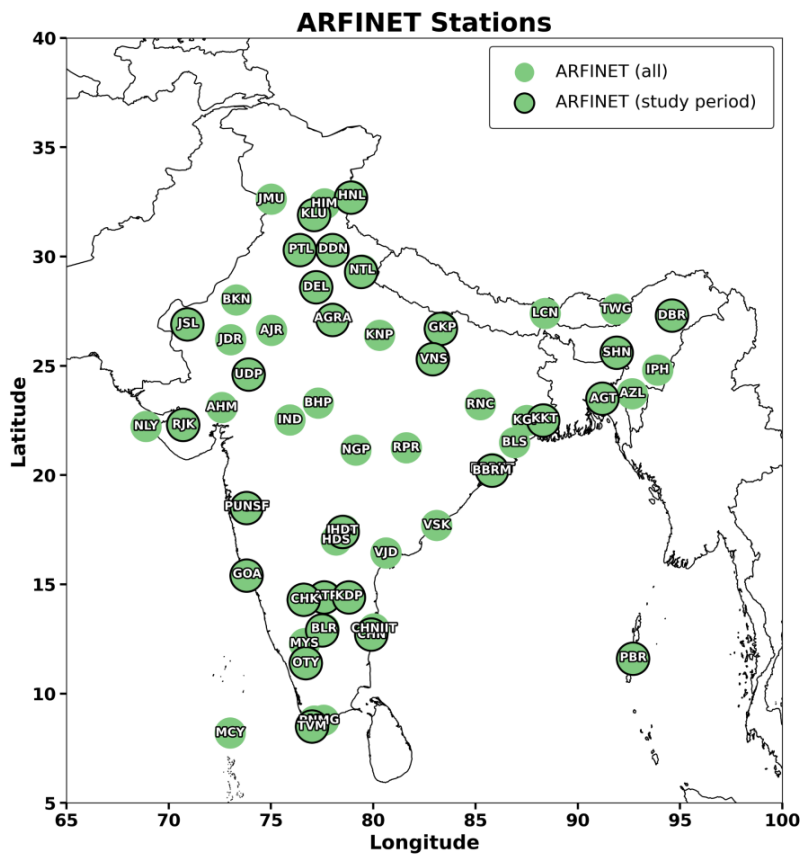
<b>Station Name</b>	<b>Station ID</b>	<b>Longitude</b>	<b>Latitude</b>
Agartala	AGT	91.2	23.5
Agra	AGR	78.0	27.1
Anantapur	ATP	77.6	14.4
Bangalore	BLR	77.5	12.9
Bhubaneswar	BBR	85.8	20.2
Challekere	CHK	76.6	14.3
Chennai	CHN	79.9	12.7
Dehradun	DDN	78.0	30.3
Delhi	DEL	77.2	28.6
Dibrugarh	DBR	94.6	27.3
Goa	GOA	73.8	15.4
Gorakhpur	GKP	83.3	26.7
Hanle	HNL	78.9	32.7
Hyderabad	HDT	78.5	17.4
Jaisalmer	JSL	70.9	26.9
Kadappa	KDP	78.8	14.4
Kolkata	KKT	88.3	22.5
Kullu	KLU	77.1	31.9
Nainital	NTL	79.4	29.3
Ooty	OTY	76.7	11.4
Patiala	PTL	76.4	30.3
Port-Blair	PBR	92.7	11.6
Pune	PUN	73.8	18.5
Rajkot	RJK	70.7	22.3
Shillong	SHN	91.9	25.6
Trivandrum	TVM	77.0	08.5
Udaipur	UDP	73.9	24.6
Varanasi	VNS	82.9	25.3

**Table S8:** List of AERONET stations and respective Latitudes and Longitudes.

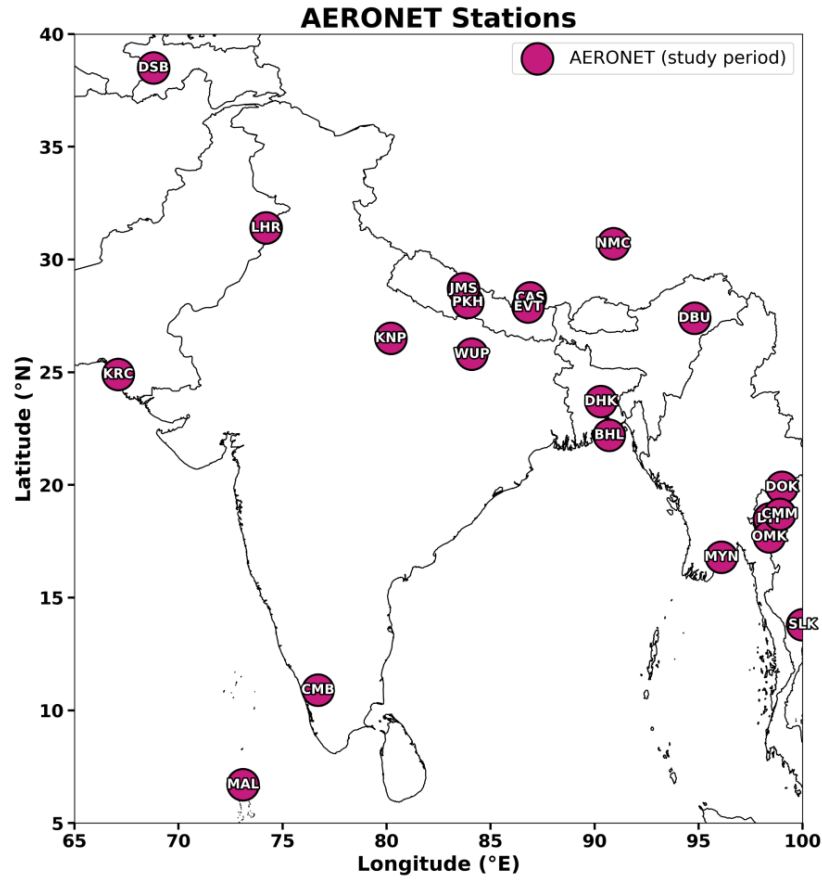
<b>Station Name</b>	<b>Station ID</b>	<b>Longitude</b>	<b>Latitude</b>
Bhola (Bangladesh)	BHL	90.7	22.2
Dhaka university	DHK	90.3	23.7
Doi_Ang_Khang (Thailand)	DOK	99.0	19.9
Dushanbe	DSB	68.8	38.5
Gandhi college (West UP)	WUP	84.1	25.8
Kanpur (India)	KNP	80.2	26.5
Karachi (Pakistan)	KRC	67.1	24.9
NAM_CO (Tibet)	NMC	90.9	30.7
Pokhara (Nepal)	PKH	83.9	28.1
QOMS-CAS (China)	CAS	86.9	28.3
Doi_Inthanon (Thailand)	DIT	98.4	18.5
Karunya university (Coimbatore)	CMB	76.7	10.9
MCO_Hanimaadho (Maldives)	MAL	73.1	6.7
Omkai (Thailand)	OMK	98.4	17.7
Silpakorn university (Thailand)	SLK	100.0	13.8
Chiang_Mai_Met_Sta (Thailand)	CMM	98.9	18.7
Jomsom	JMS	83.7	28.7
Lahore (Pakistan)	LHR	74.2	31.4
EVK2-CNR	EVT	86.8	27.9
Myanmar	MYN	96.1	16.8
Dibrugarh	DBU	94.8	27.4

**Table S9:** Observed and modified values of AOD used in the sensitivity analysis for UK and RK-ML method.

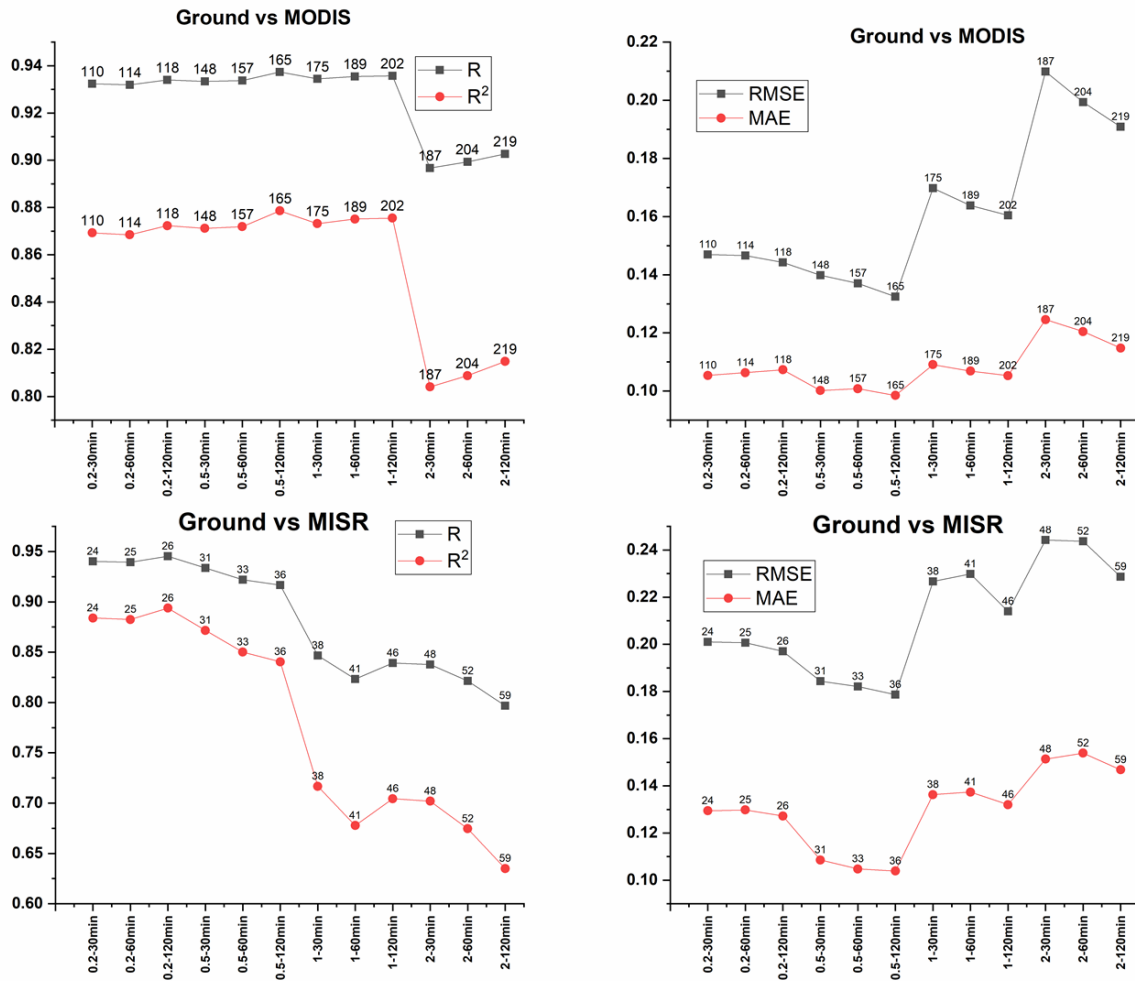
Station ID	Longitude	Latitude	AOD (observed)	AOD (modified)
DDN	78.04	30.34	0.50	0.70
JSL	70.95	26.92	0.36	0.10
KLU	77.10	31.90	0.38	0.20
PBR	92.71	11.64	0.22	0.90
JMS	83.71	28.77	0.04	1.00
LHR	74.26	31.47	0.77	1.30



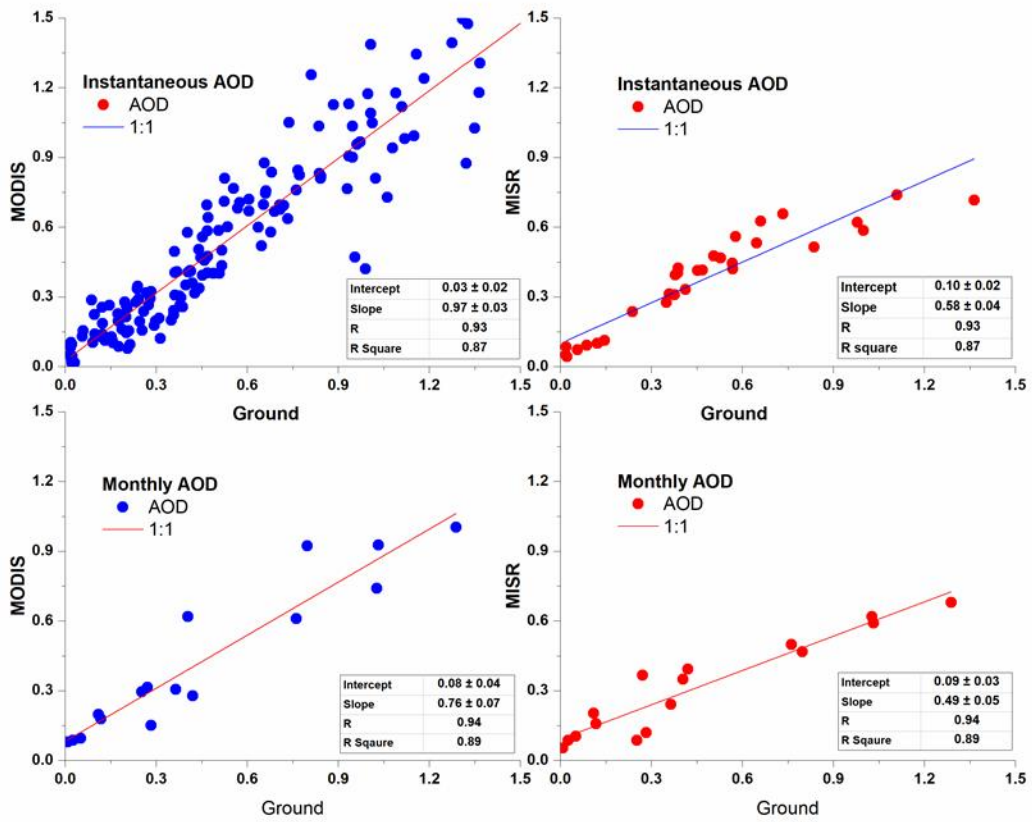
**Figure S1:** ARFINET stations (green circles) over different parts of India; stations indicated by circles with black outline are used in the present study. Refer **Table S7** for station names.



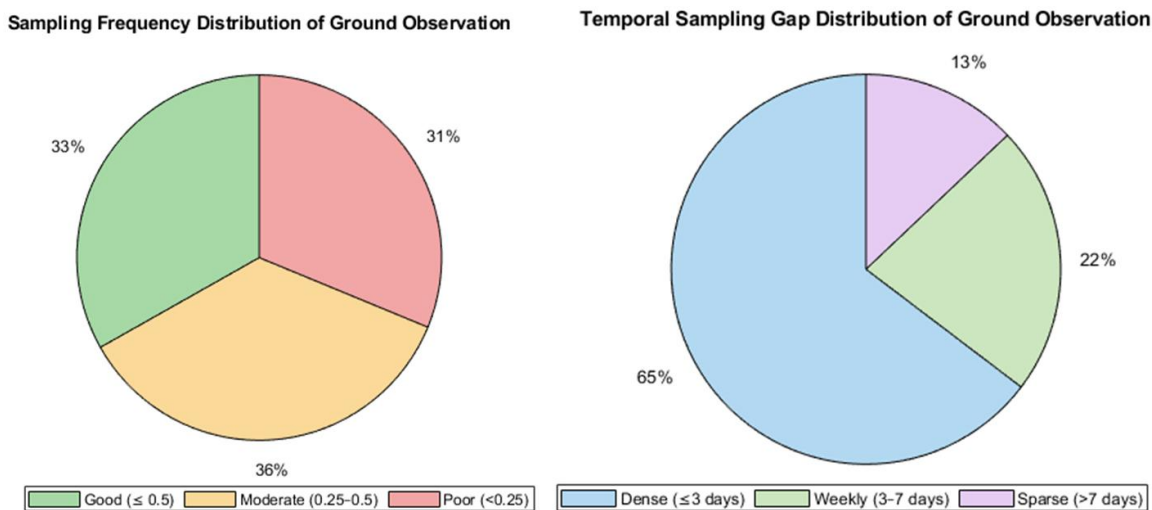
**Figure S2:** AERONET stations over different parts of South-Asian regions used in the current study. Refer **Table S8** for station names.



**Figure S3:** Spatio-temporal association between collocated satellite and ground AOD for Nov-2021. (a) R and R<sup>2</sup> and (b) RMSE and MAE for MODIS vs ground; (c) R and R<sup>2</sup> and (d) RMSE and MAE for MISR vs. ground. The value on the x-axis shows the format of spatial-resolution followed by temporal window of averaging (e.g., 0.2-30min; 0.2 is spatial resolution, 30 min is temporal window). The numerical values above the lines represent the number of collocated points for different spatio-temporal windows. (Refer **Table S2**)



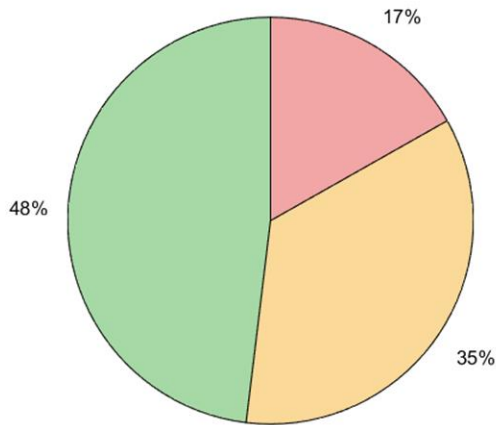
**Figure S4:** Scatter plot between collocated points between MODIS and ground (top left: Instantaneous AOD; bottom left: monthly mean); and MISR and ground (top right: Instantaneous AOD; bottom right: Monthly mean)



**Figure S5:** (Left panel) Pie-charts of sampling frequency per month of ground observations. Green represents the good number of sampling days per month (15 days or  $\geq 50\%$  observations), orange being a moderate number of sampling days per month (08-15 days; or 25-50% of observations) and red representing low sampling days per month ( $< 08$  days; or  $< 25\%$  of observations). (Right panel) Pie-charts of temporal sampling gap between consecutive days in ground observations. Sky-blue color representing dense data with minimal gap ( $\leq 3$ days), light-green represents data with 3-7days gap and purple showing data with sparse data ( $> 7$  days).

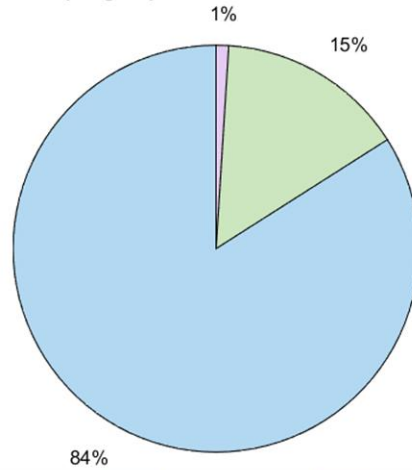
The compiled ground-based dataset consists of 2507 observation days across 22 stations on average per month during January, May, and November for the years 2012, 2016, and 2021. This corresponds to approximately 41% of the maximum possible observations, which is consistent with typical data availability in aerosol observations primarily influenced by cloud screening and instrument downtime. The average temporal sampling gap is  $\sim 3-4$  days, indicating adequate coverage for reliable monthly statistics.

Sampling Frequency Distribution of MODIS Observation



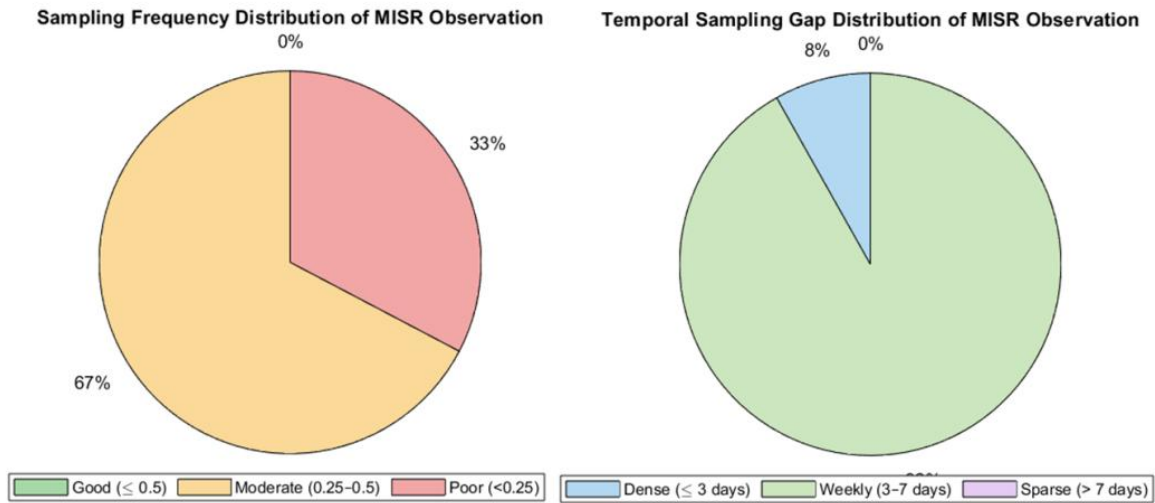
Good ( $\leq 0.5$ ) Moderate (0.25-0.5) Poor ( $< 0.25$ )

Temporal Sampling Gap Distribution of MODIS Observation

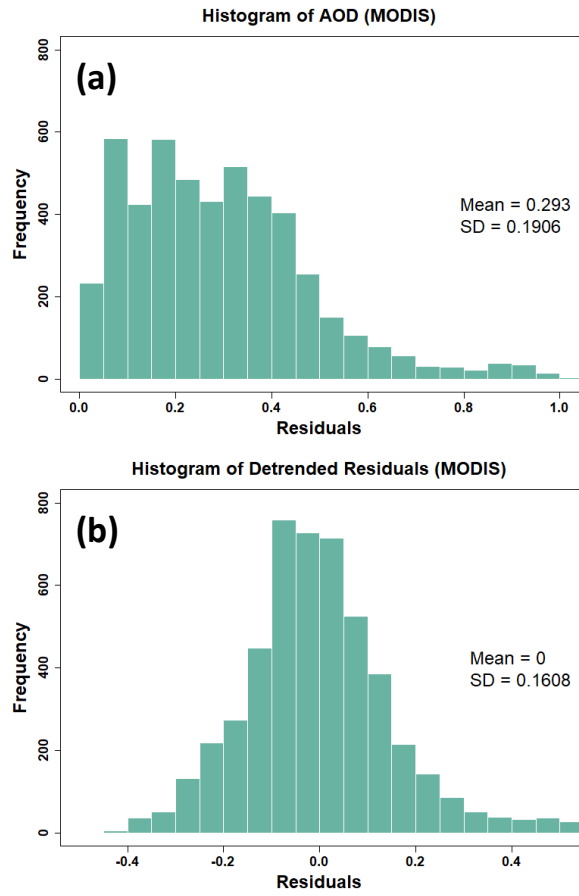


Dense ( $\leq 3$  days) Weekly (3-7 days) Sparse ( $> 7$  days)

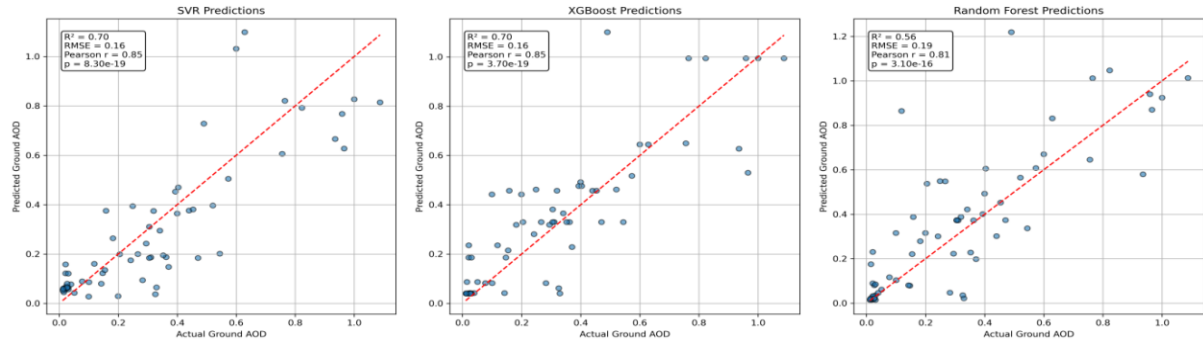
**Figure S6:** Same as **Fig. S5** but for MODIS observations. Hence, similar inference can be made for MODIS. MODIS provides approximately 49% of observations (days) in a month.



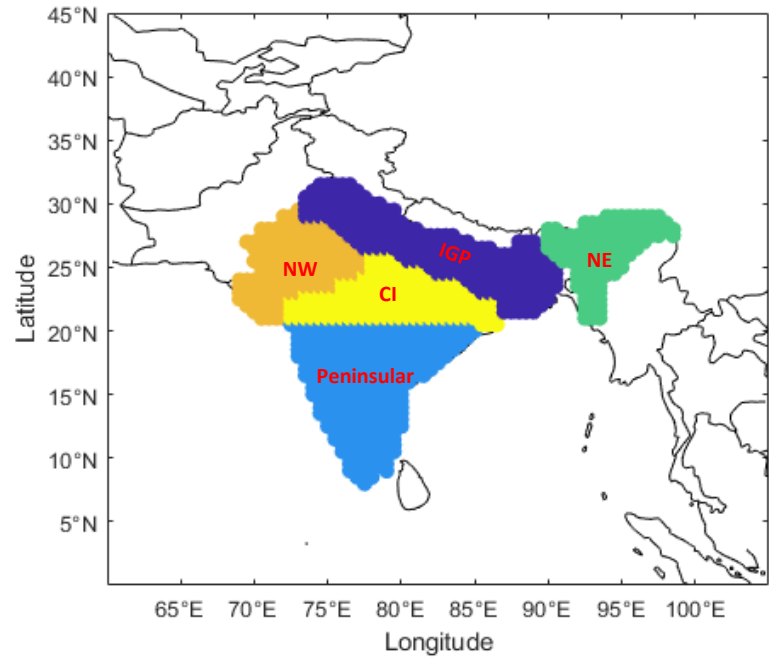
**Figure S7:** Same as **Fig. S5** but for MISR observations. MISR observations occur on average 25.6% of the days in a month, corresponding to a mean revisit interval of approximately 4 days.



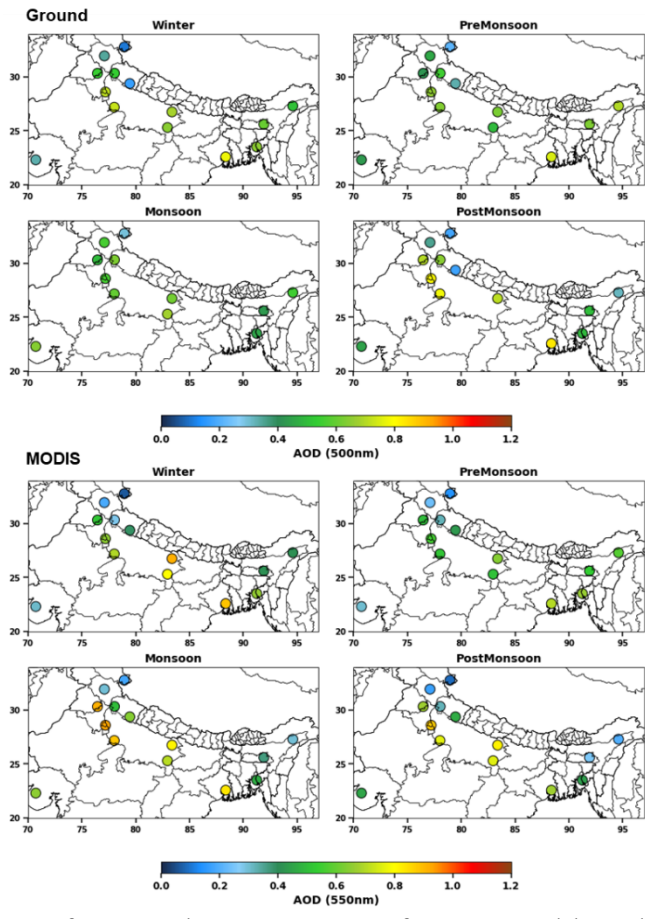
**Figure S8:** Histogram of (a) MODIS AOD, and (b) Detrended (Spatial detrending) residuals from MODIS AOD with respect to latitude, longitude, elevations.



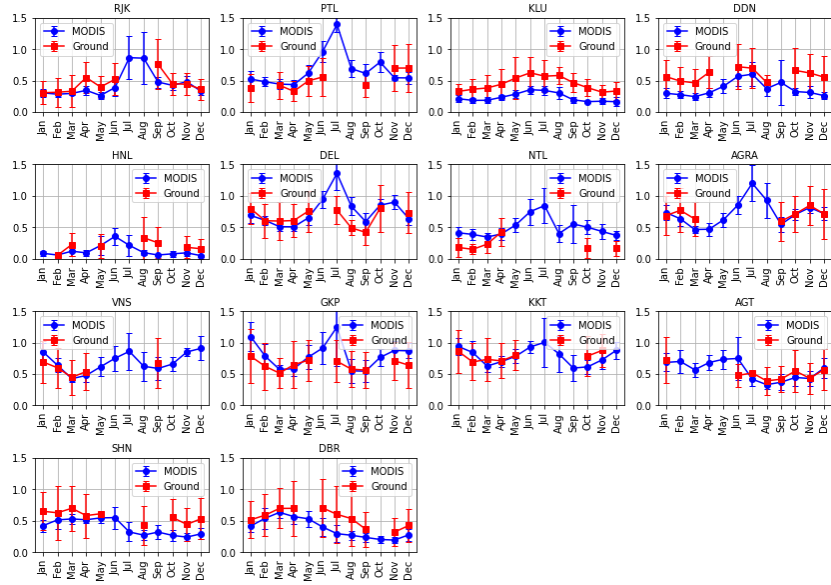
**Figure S9:** Comparison plots showing the performances of SVR, XGBoost and Random Forest.



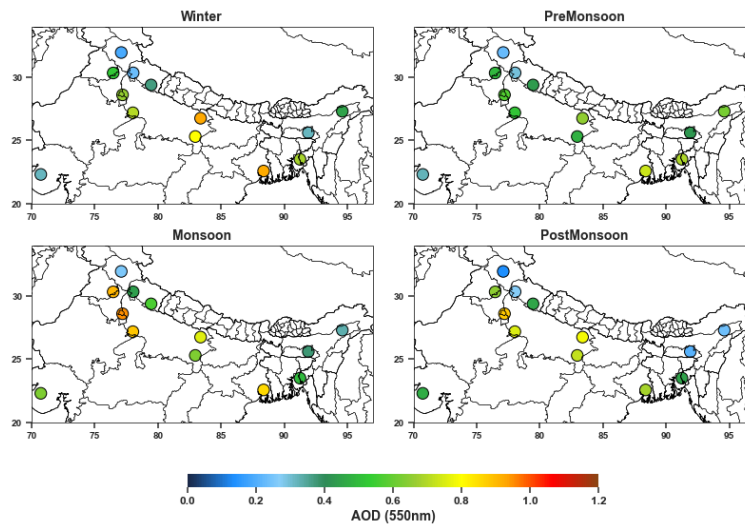
**Figure S10:** Classification of different regions of India and their boundaries used in this study.



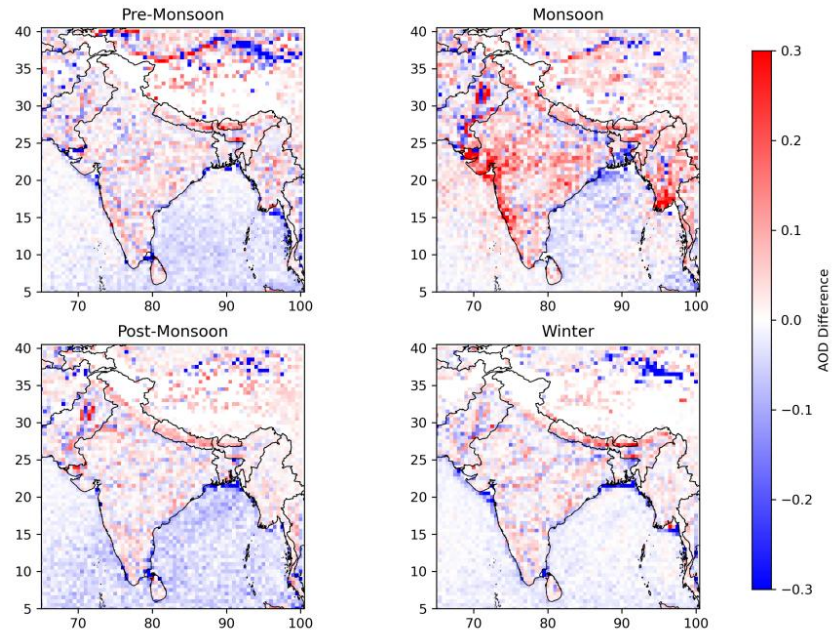
**Figure S11:** Comparison of seasonal average AOD from ground-based and satellite (MODIS) observations over different ground locations of ARFINET.



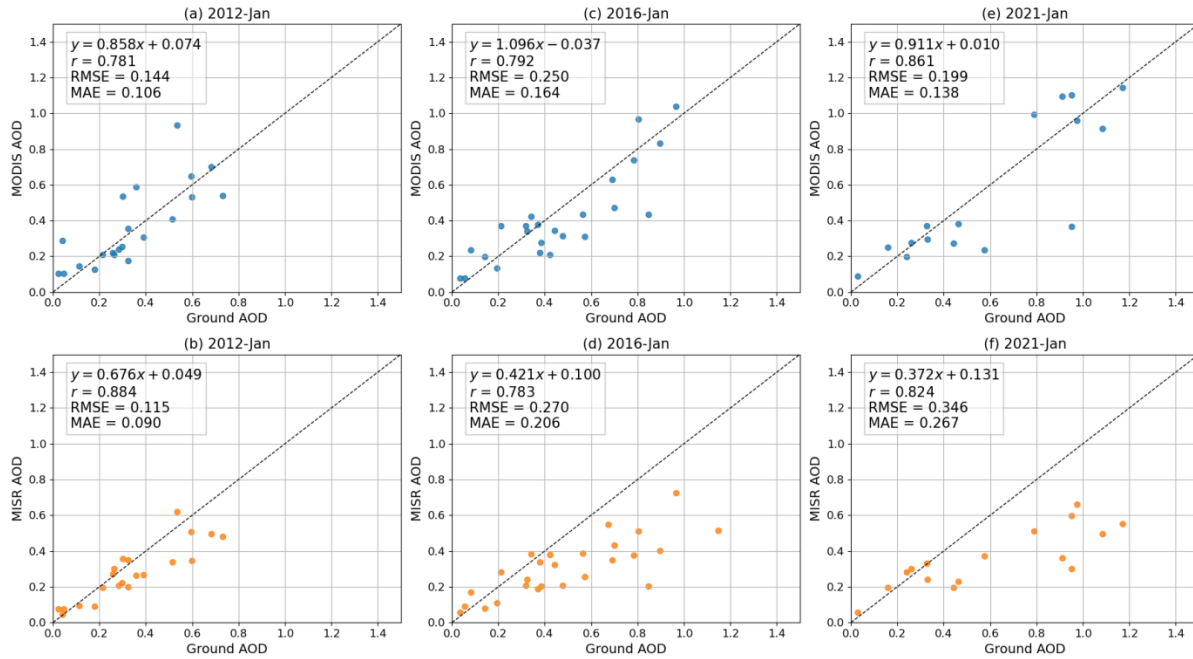
**Figure S12:** Time series plots of long-term ground-based AOD (monthly mean with standard deviation) and satellite AOD (MODIS; monthly mean with standard deviation) over selected locations of ARFINET.



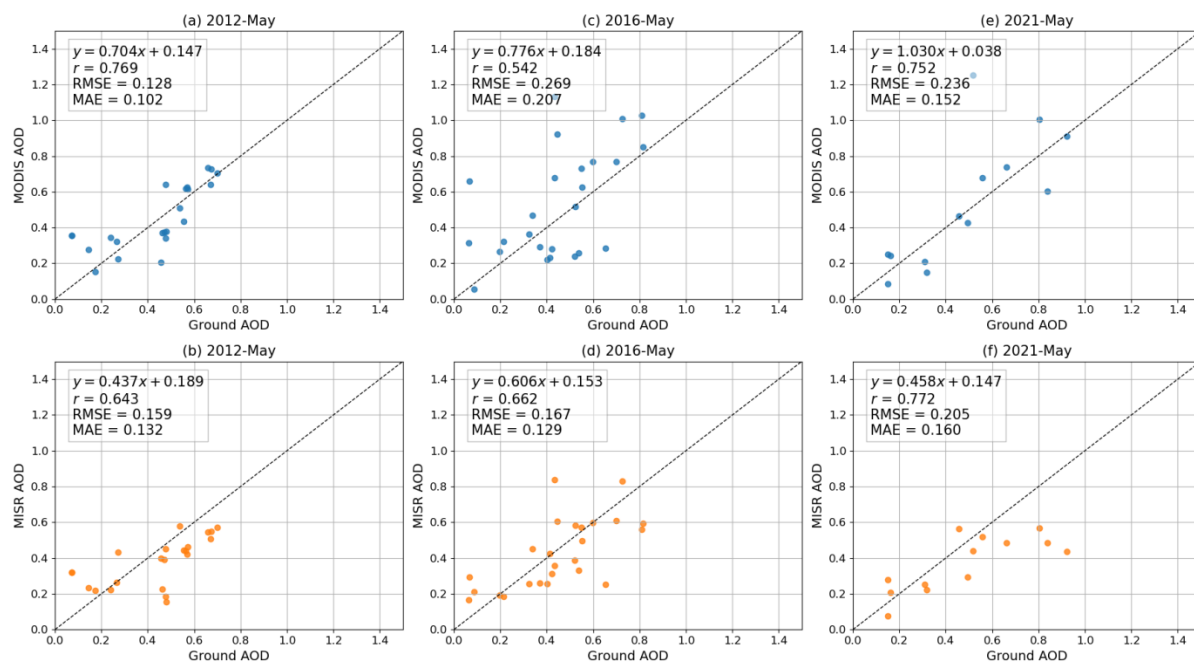
**Figure S13:** Long-term seasonal mean AOD over ground monitoring locations derived from MODIS AOD after removing pixels for haze/cloud-discrimination criteria.



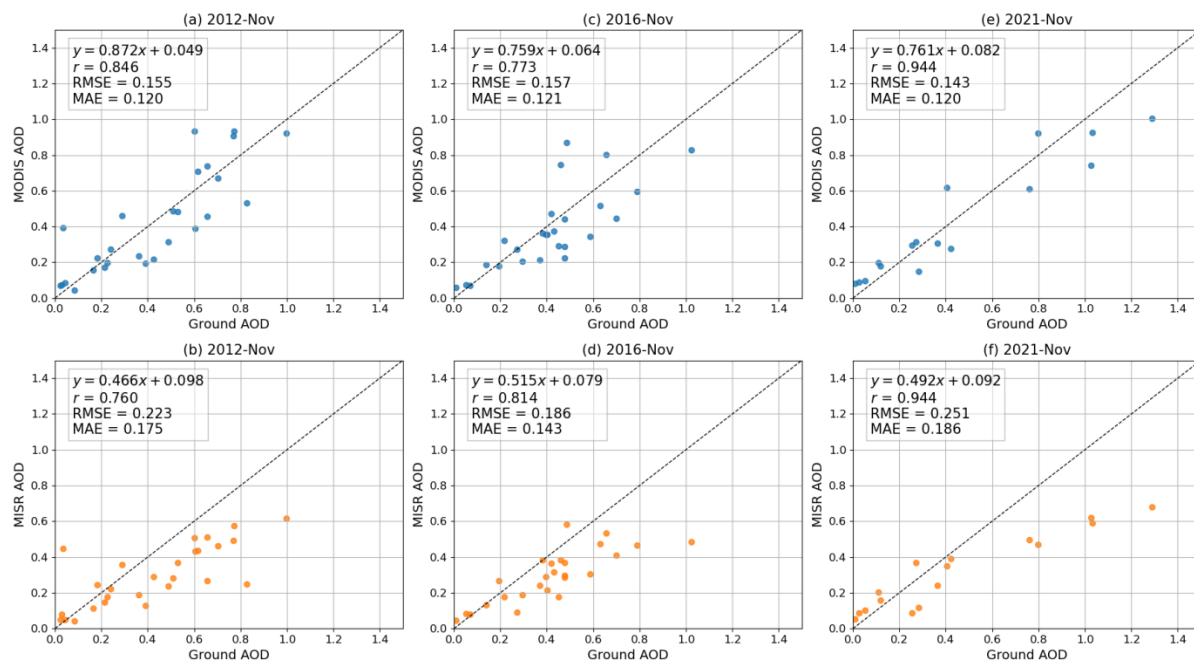
**Figure S14:** Difference between long-term seasonal mean AOD derived from MODIS AOD before and after removing pixels using haze/cloud-discrimination criteria.



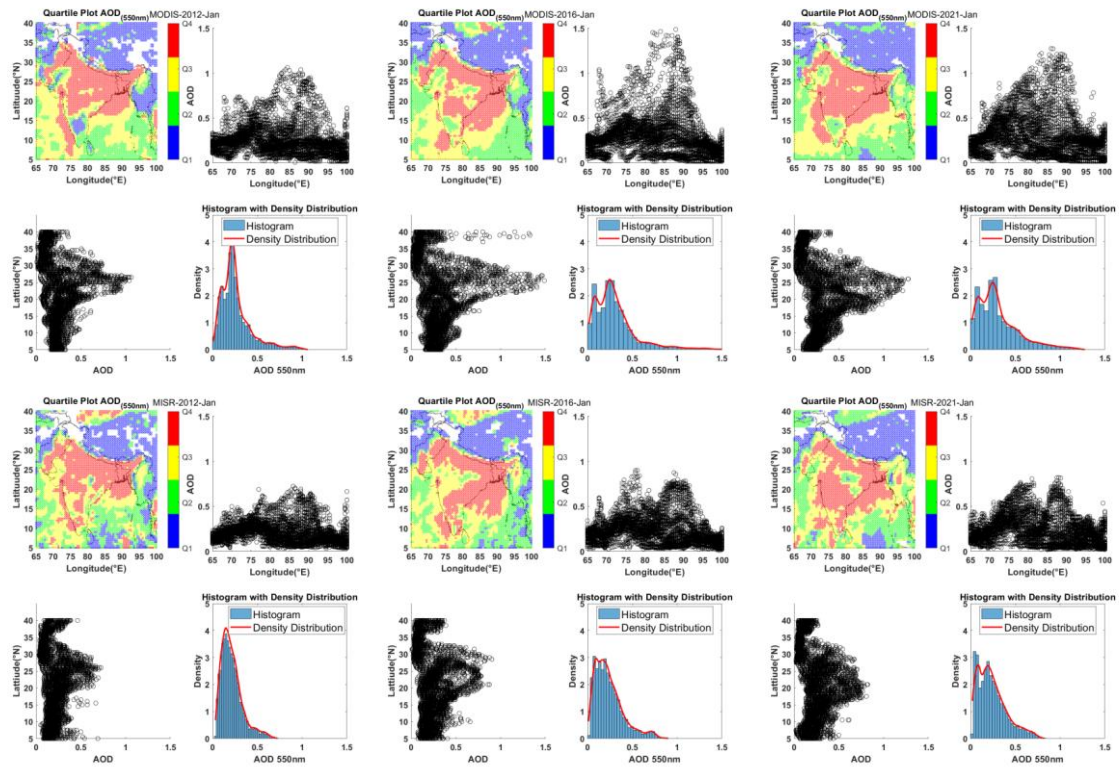
**Figure S15:** Comparison between MODIS (top) and MISR (bottom) versus ground-based AOD for the month of January 2012 (N=21), January 2016 (N=26), and January 2021(N=16), respectively.



**Figure S16:** Comparison between MODIS (top) and MISR (bottom) versus ground-based AOD for the month of May 2012 (N=22), May 2016 (N=25), and May 2021(N=13), respectively.



**Figure S17:** Comparison between MODIS (top) and MISR (bottom) versus ground-based AOD for the month of November 2012 (N=27), November 2016 (N=26), and November 2021 (N=16), respectively.



**Figure S18:** Quartile plots of MODIS and MISR AOD at 550 nm for January 2012, 2016, and 2021 over India followed by AOD distribution over longitude and latitude and its density distribution.

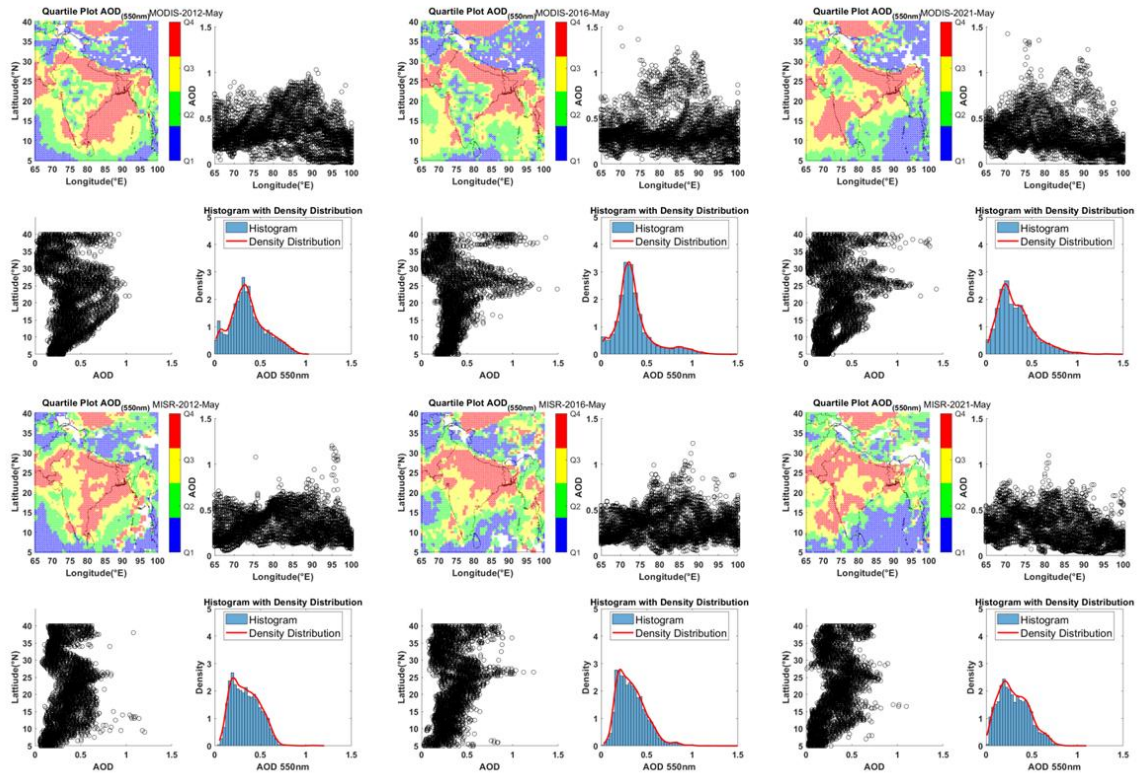


Figure S19: Similar analysis as in Fig. S18 but for May month..

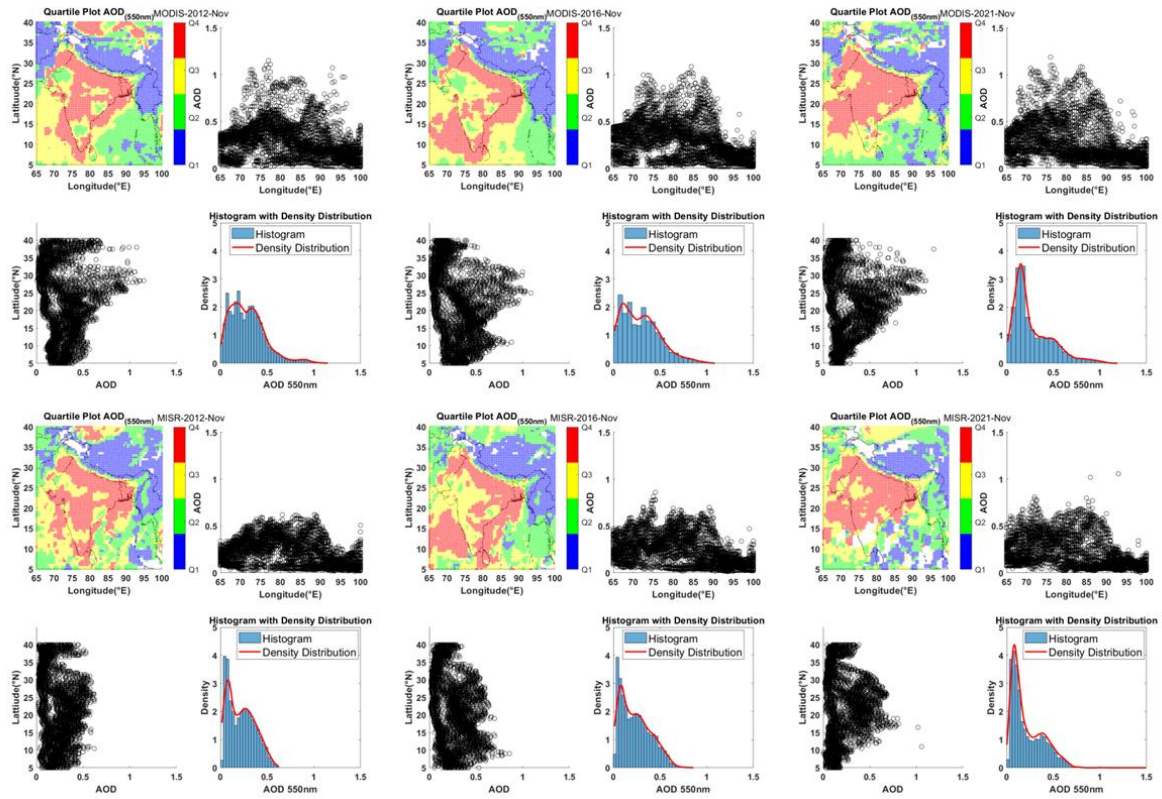
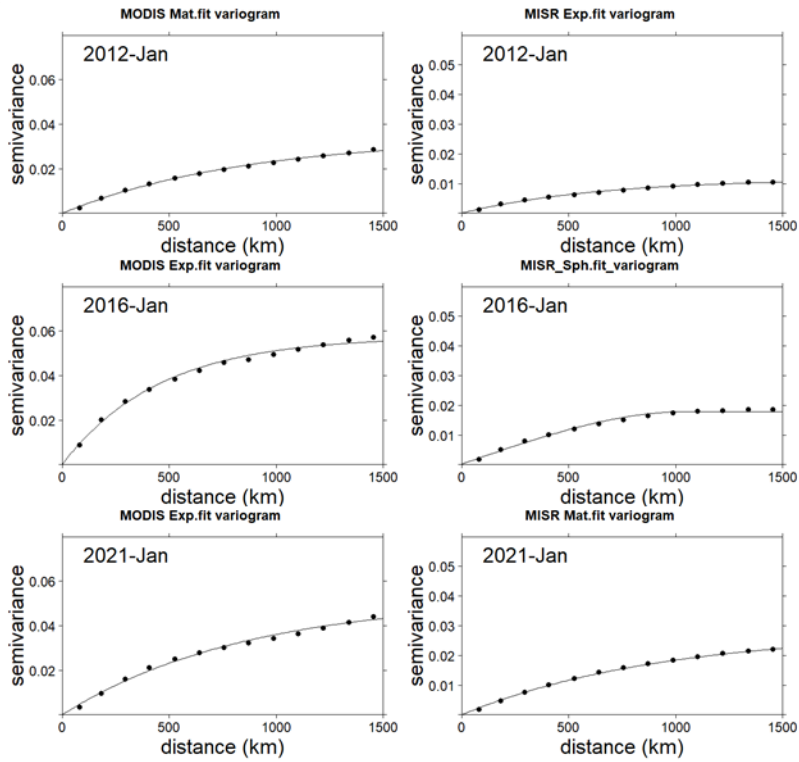
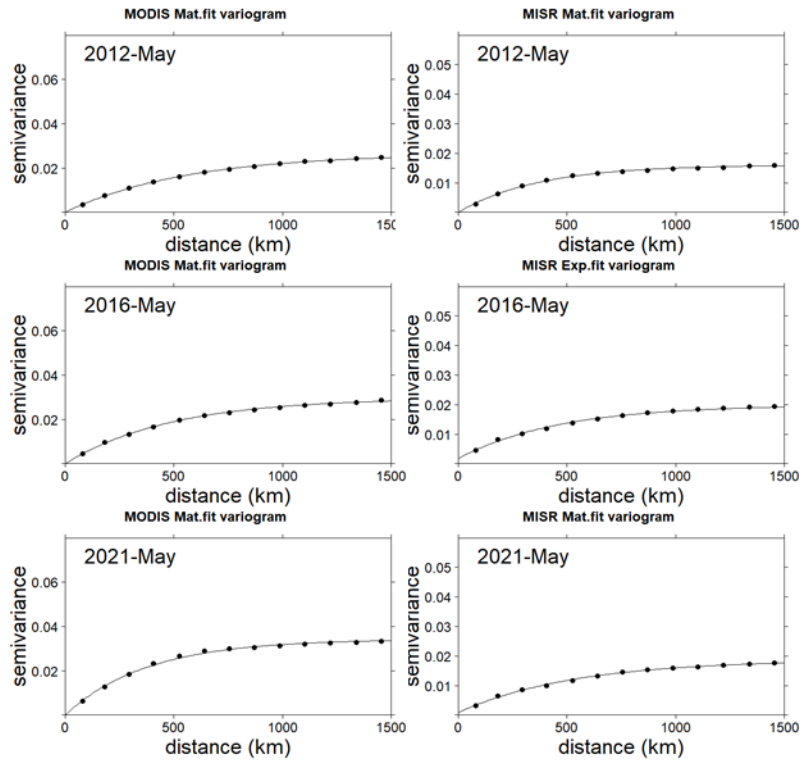


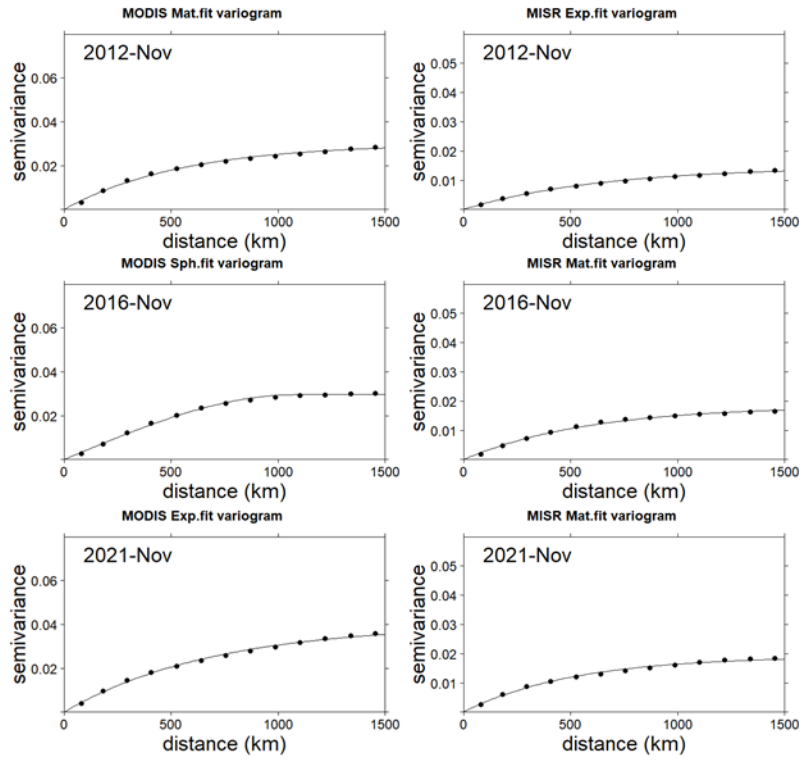
Figure S20: Similar analysis as in Fig. S18 but for November month.



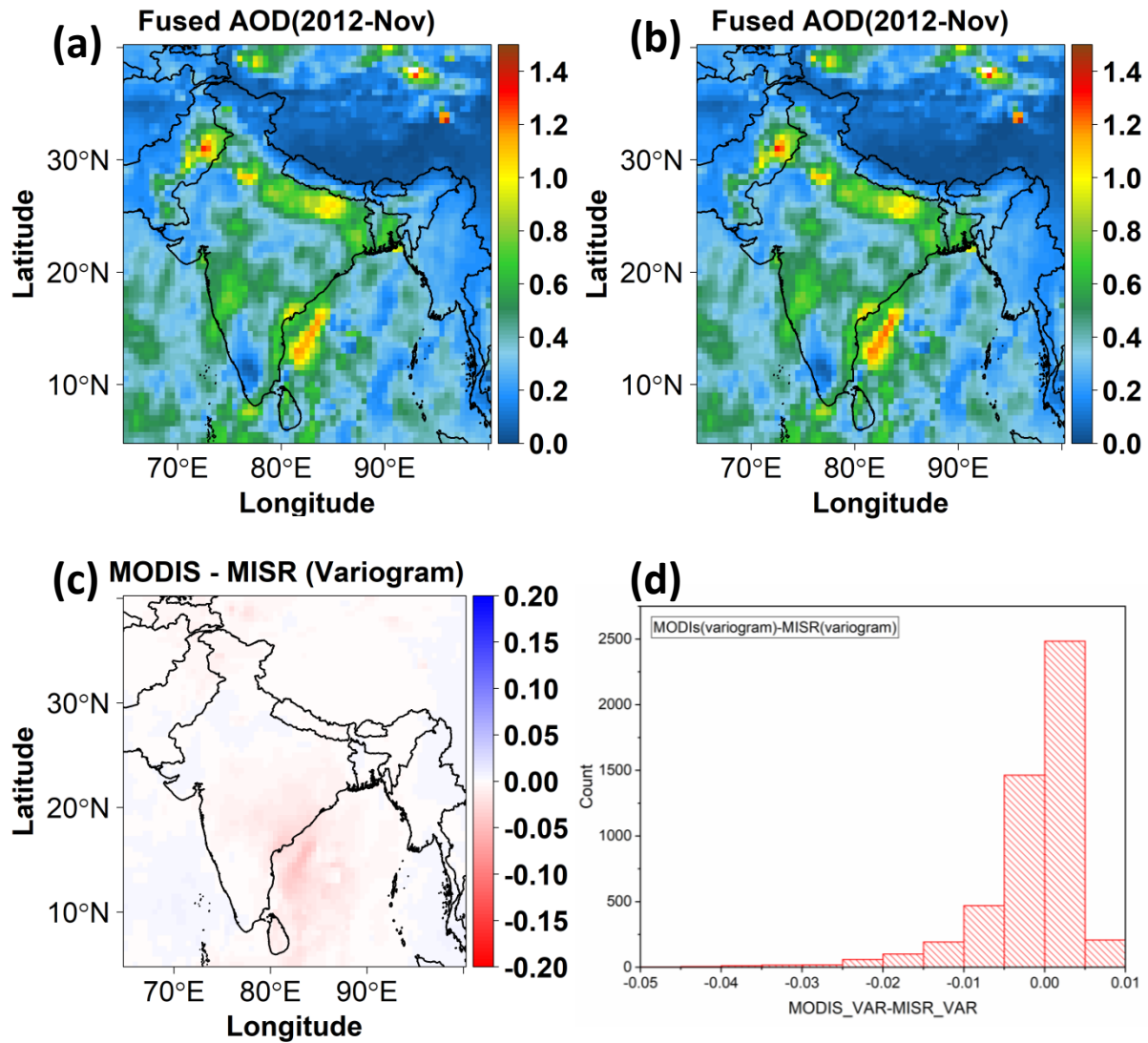
**Figure S21:** Experimental variogram of MODIS (left) and MISR (right) AOD for the month of January 2012, 2016, and 2021.



**Figure S22:** Similar analysis as in **Fig. S21** but for May month.

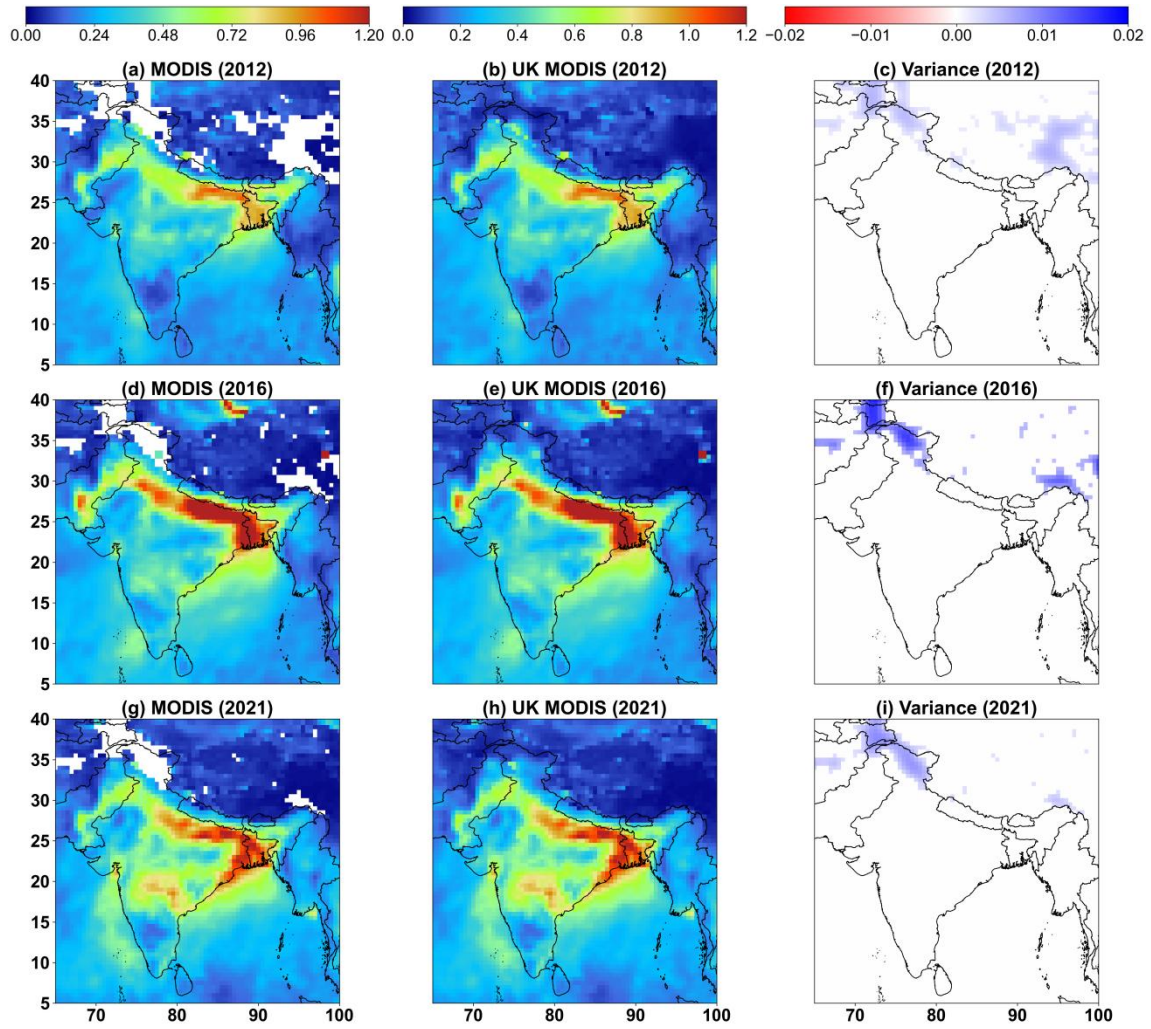


**Figure S23:** Similar analysis as in **Fig. S21** but for November month.

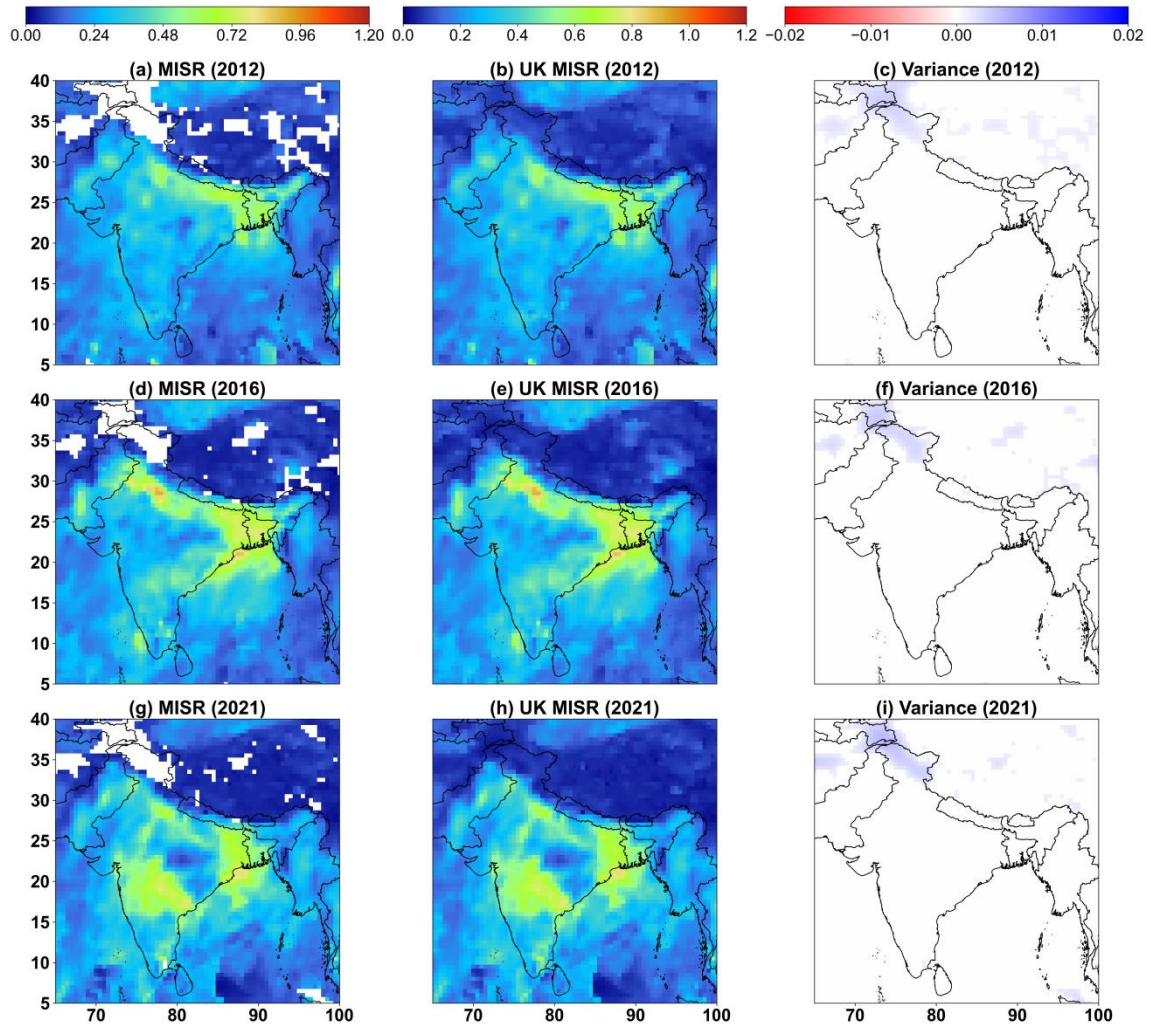


**Figure S24:** Fused AOD generated from (a) MODIS variogram and (b) MISR variogram (c) difference between fused AOD obtained using MODIS and MISR variogram, (d) Histogram of values shown in (c). Variogram details are in (Table S3).

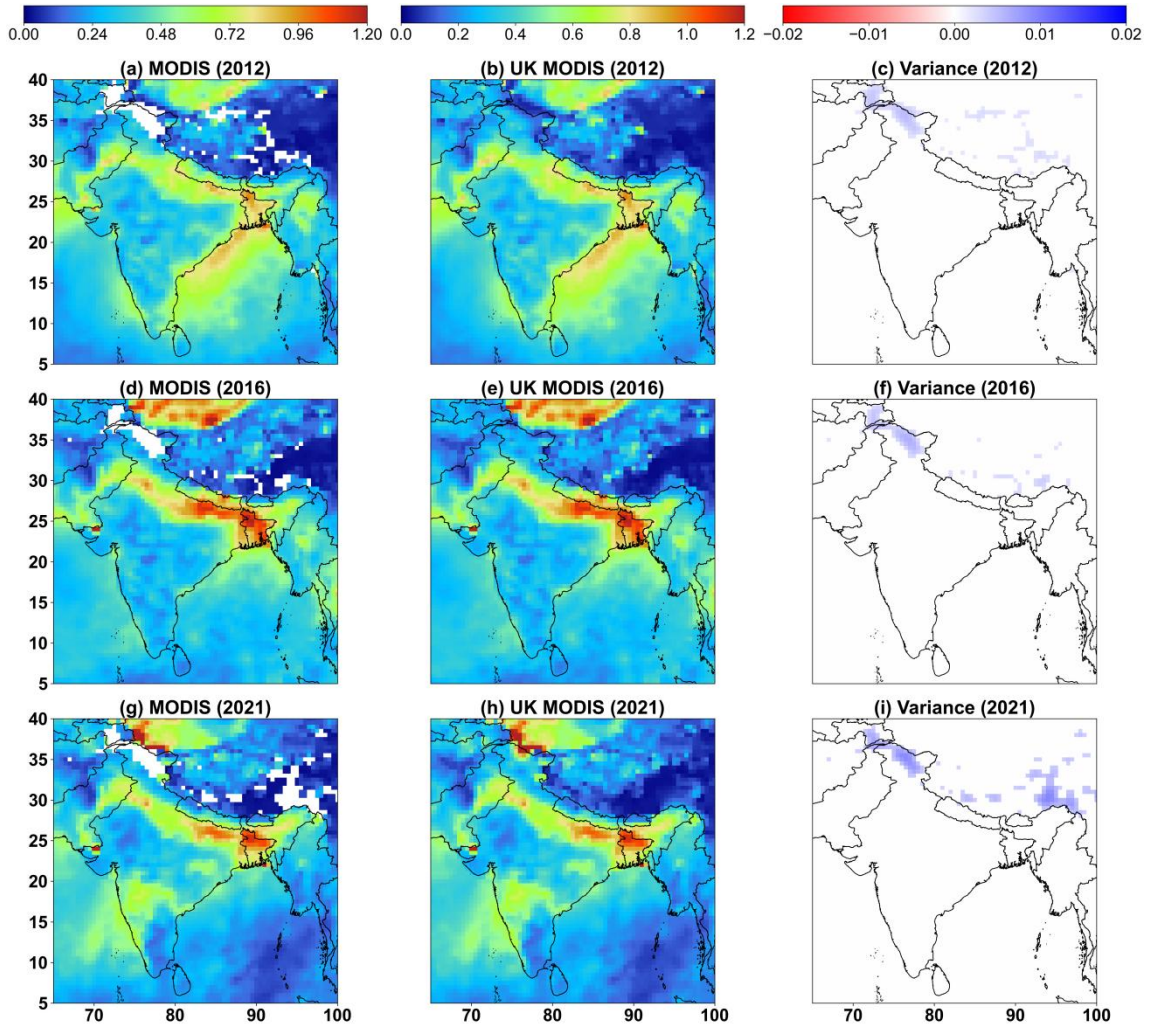
The difference values are  $\sim -0.01$  to  $0.01$  while over mainland it is close to  $\sim 0$  than oceanic regions.



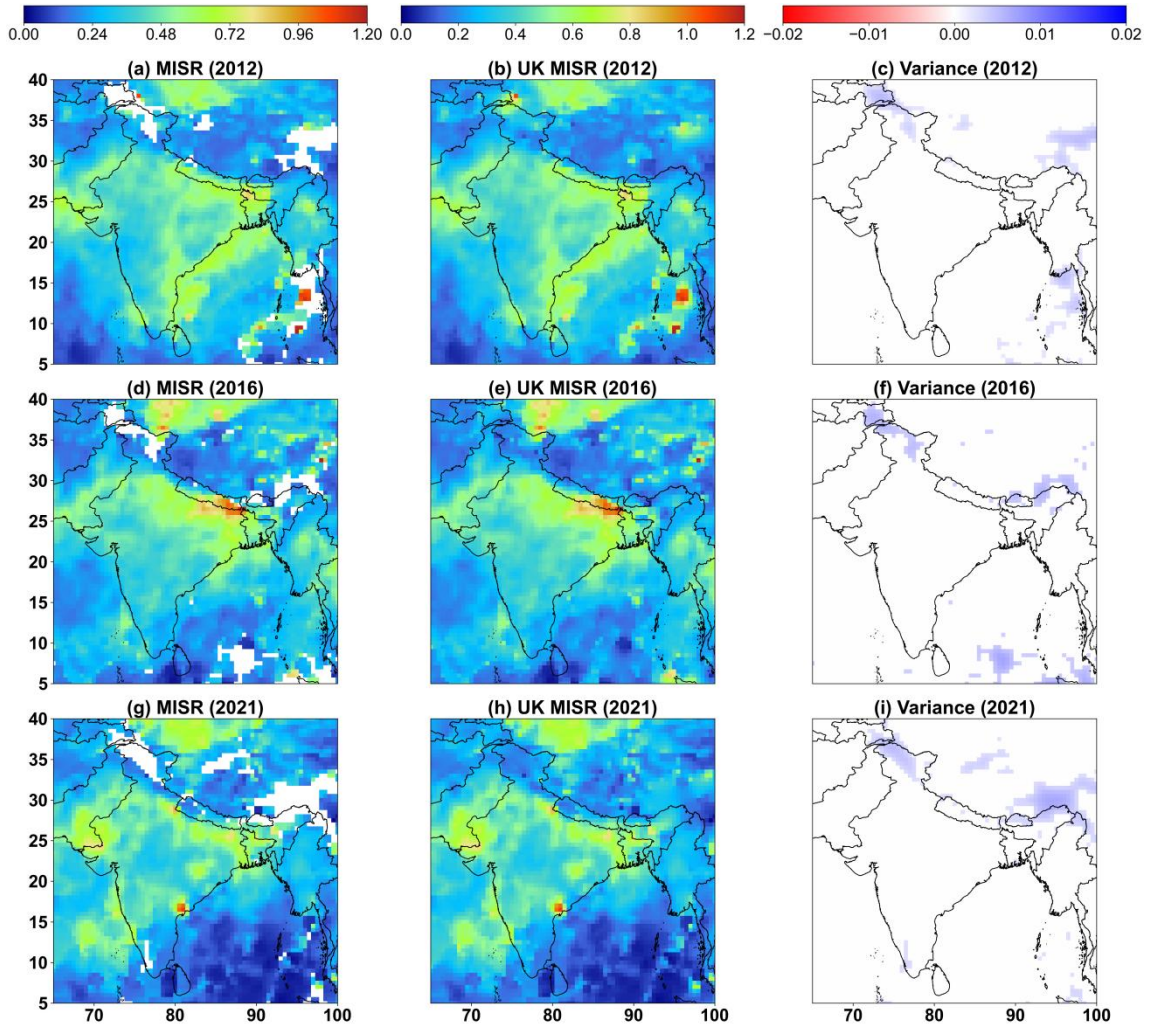
**Figure S25:** Monthly mean  $0.5^\circ \times 0.5^\circ$  gridded MODIS AOD for January 2012, 2016, and 2021; MODIS AOD [(a), (d) and (g)], spatial prediction using UK [(b), (e) and (h)], and variance of predictions [(c), (f) and (i)].



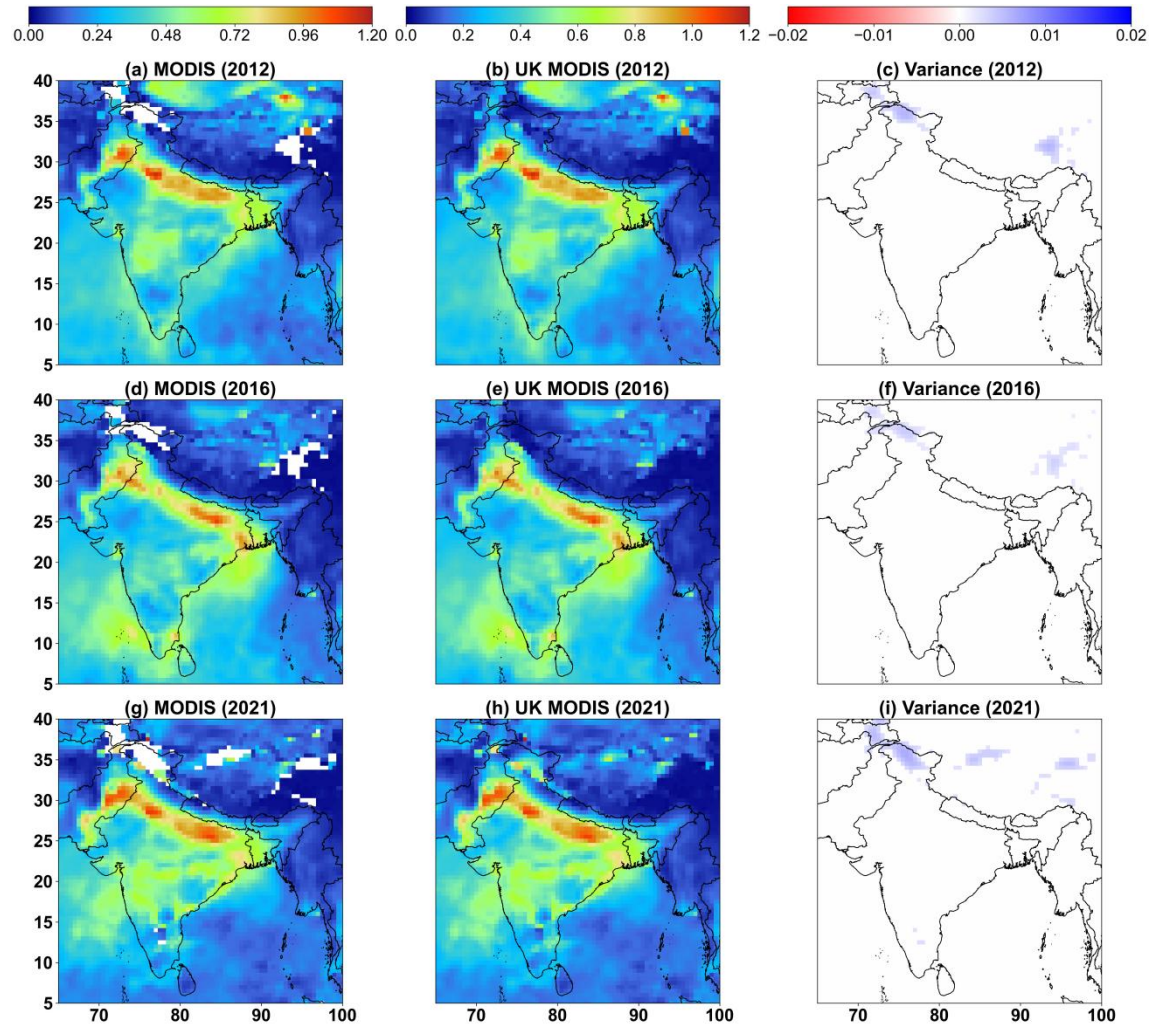
**Figure S26:** Monthly mean  $0.5^\circ \times 0.5^\circ$  gridded MISR AOD for January 2012, 2016, and 2021; MODIS AOD [(a), (d) and (g)], spatial prediction using UK [(b), (e) and (h)], and variance of predictions [(c), (f) and (i)].



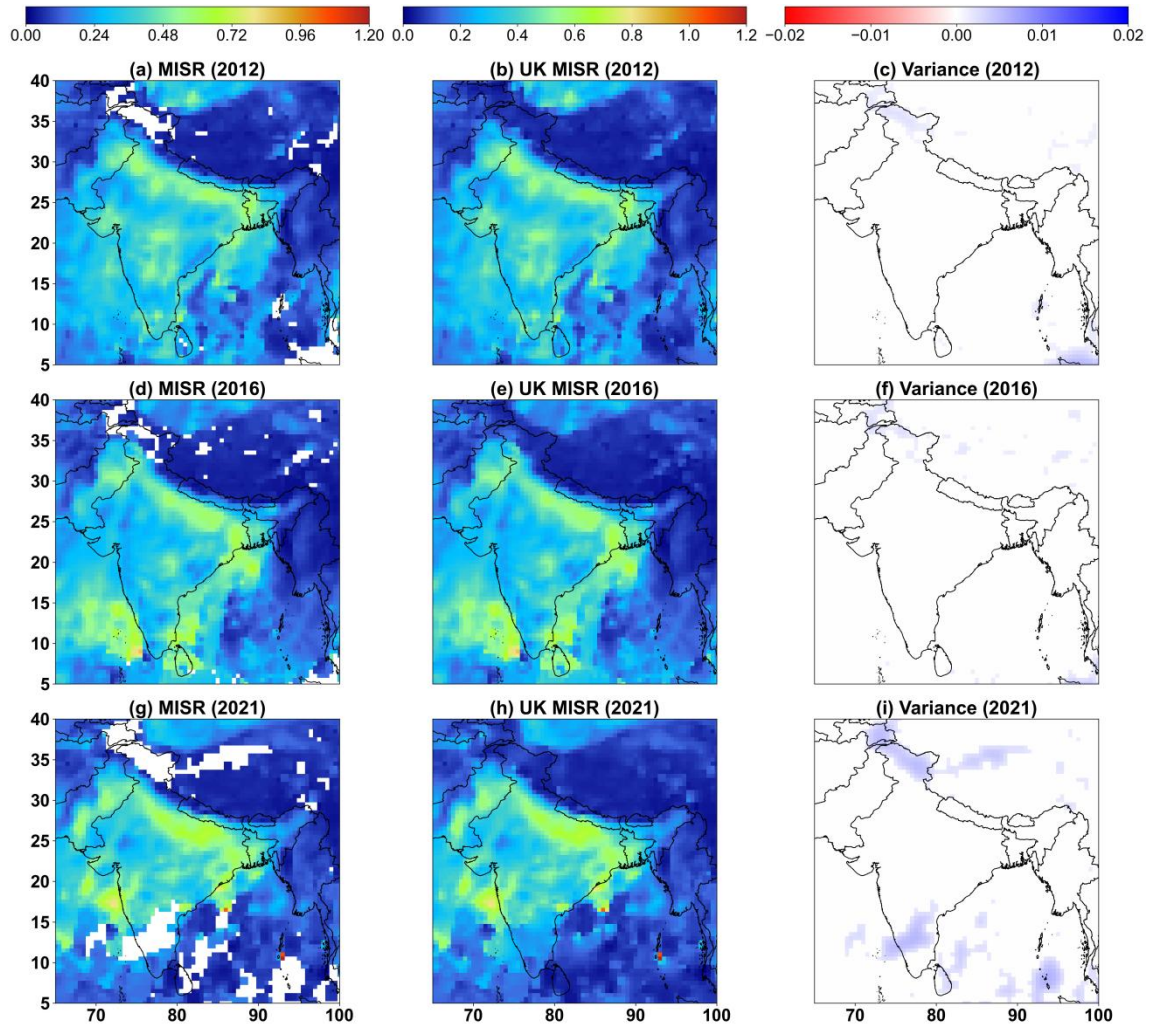
**Figure S27:** Similar analysis as in Fig. S25 but for May month.



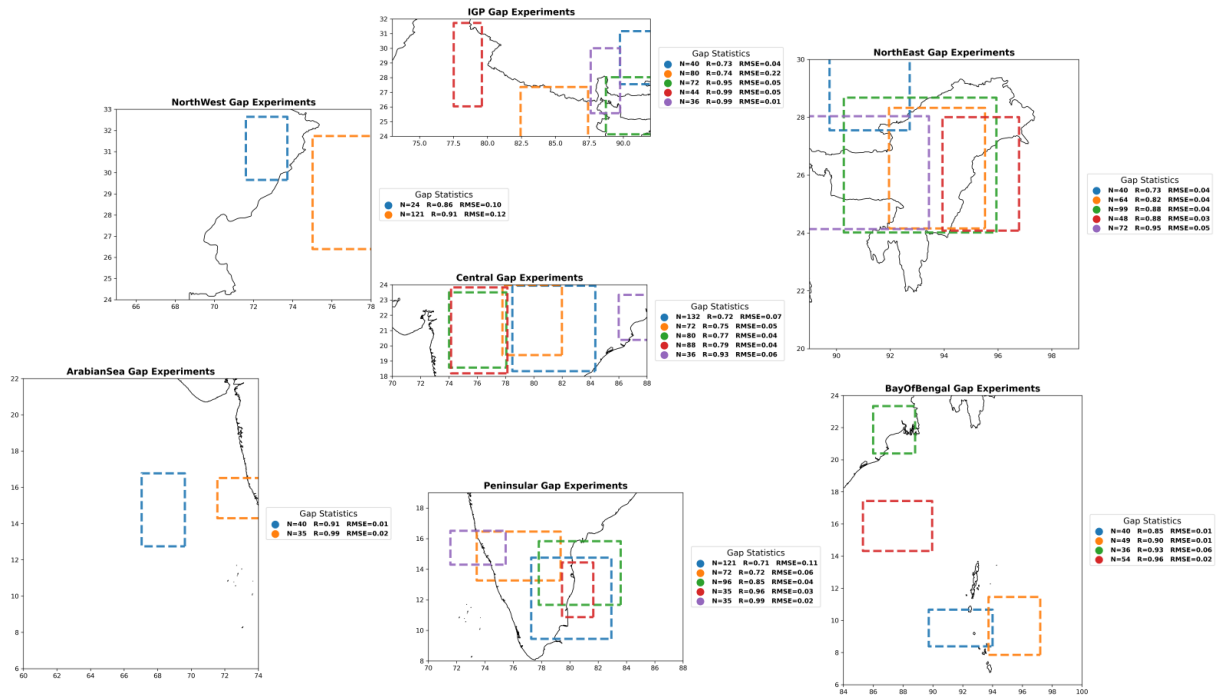
**Figure S28:** Similar analysis as in **Fig. S26** but for May month.



**Figure S29:** Similar analysis as in Fig. S25 but for November month.

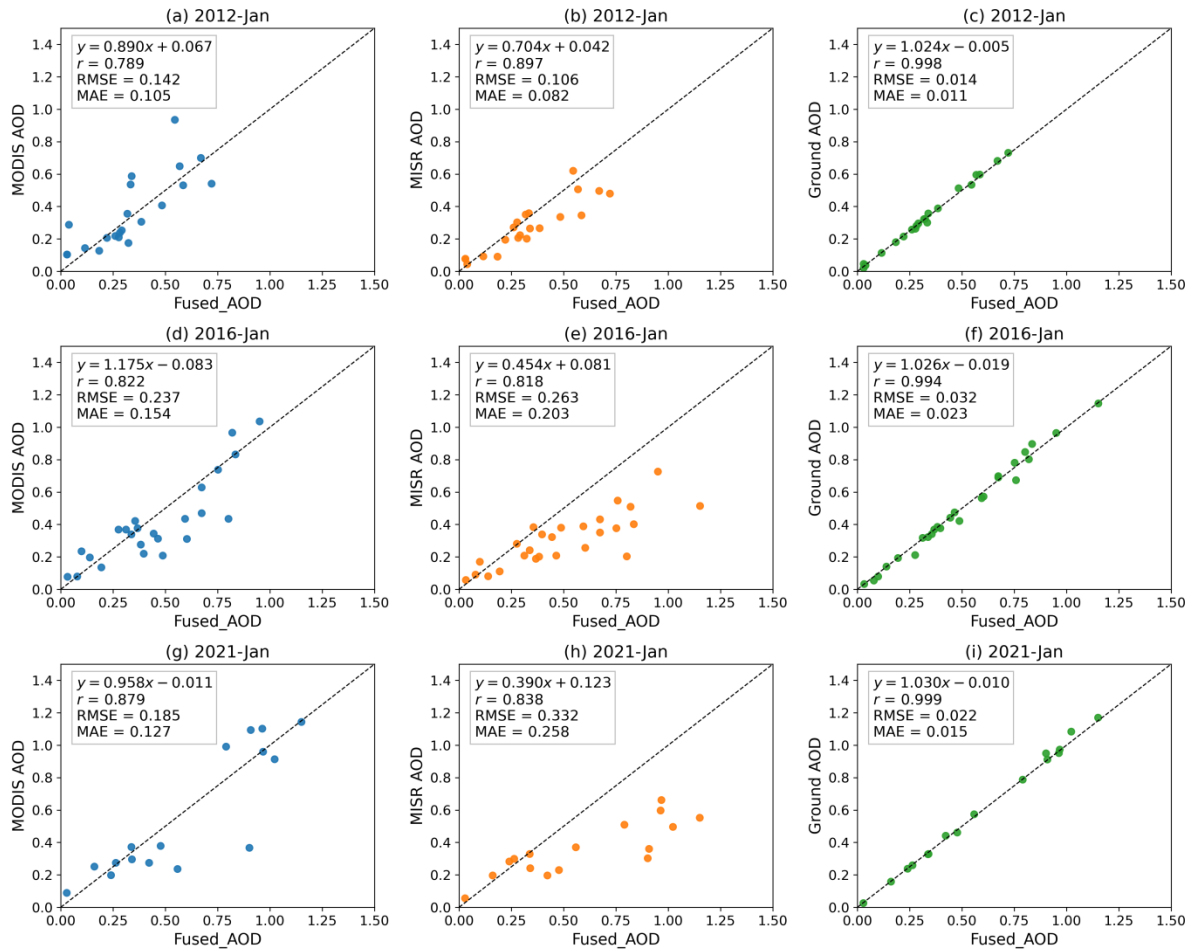


**Figure S30:** Similar analysis as in **Fig. S26** but for November month.

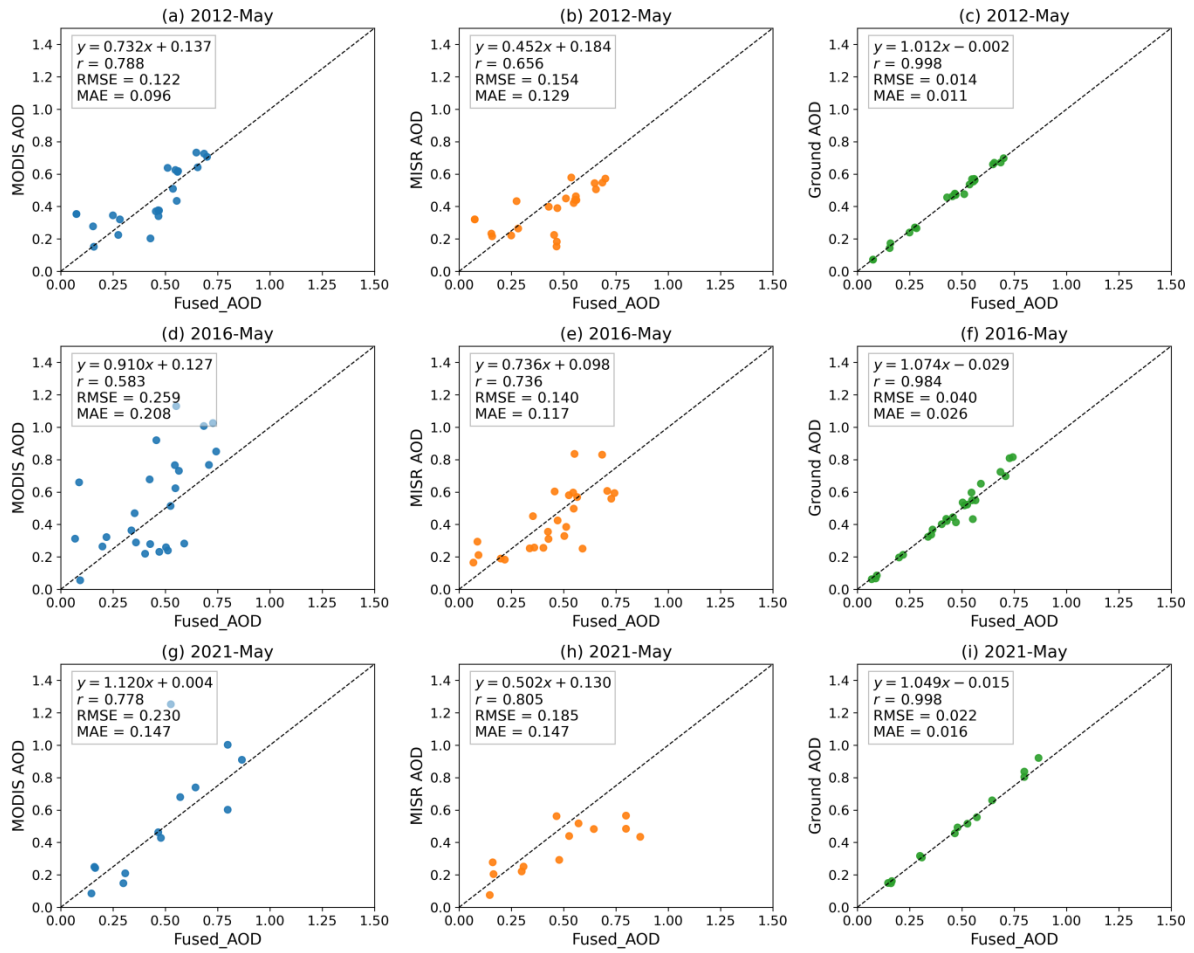


**Figure S31:** Different areas across the study region, where data were removed for validation of spatial interpolation using the Universal Kriging approach, are shown using dotted square boxes. The number of data points removed (N), correlation between removed data and predicted (R) and error between the original removed and predicted points (RMSE) between the actual AOD in the removed regions and the spatially predicted AOD over those same regions are represented by corresponding colors.

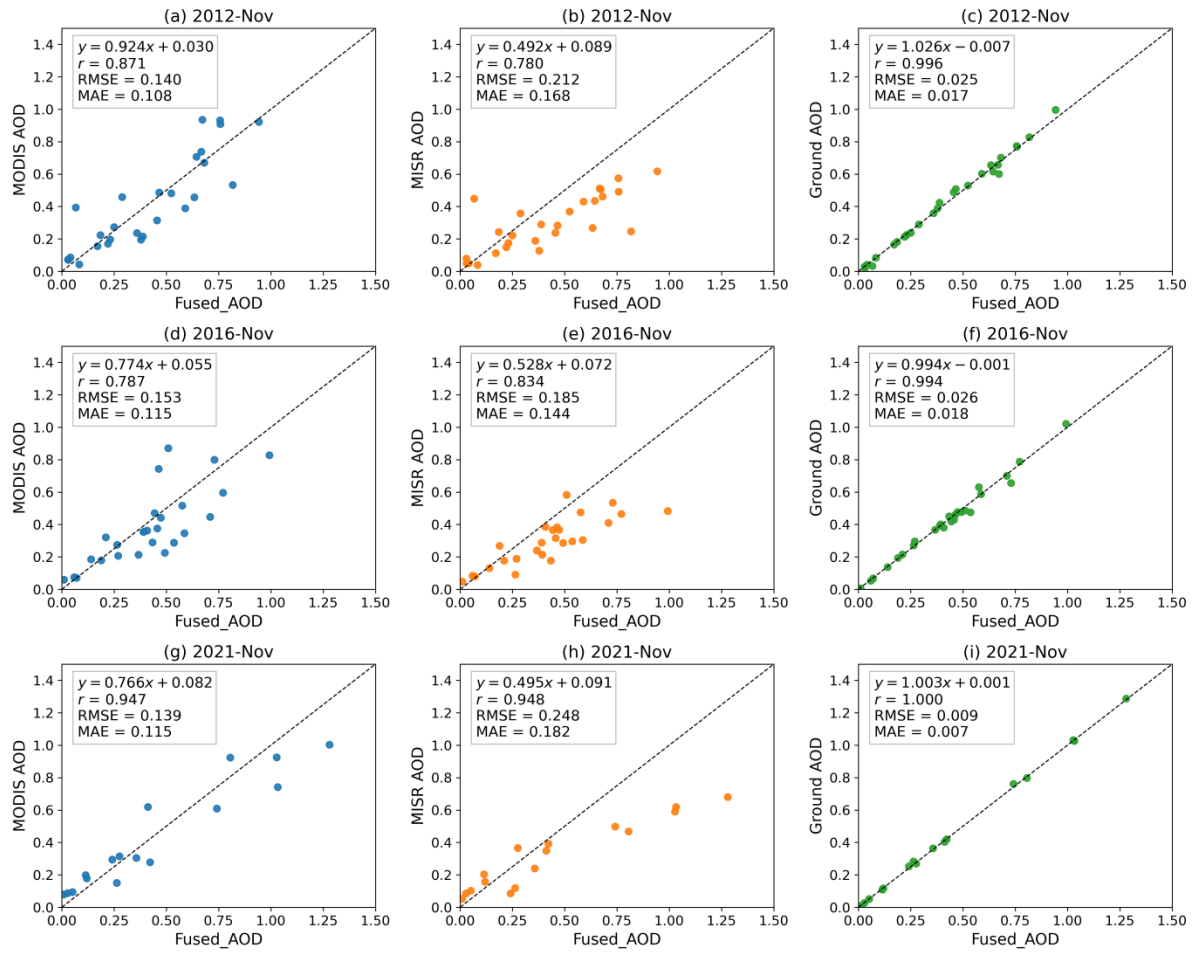
The sensitivities of predictability of gap filling methods showing both best and worst cases of validations. This highlights correlation  $R \sim 0.99 - 0.71$  for different cases while RMSE varies from 0.01 to 0.12 which conveys the accuracy of gap filled method using UK, however the predicted AOD depends on the nearby observed data points. Hence, the larger area gaps over heterogeneous landmasses show relative lower correlations.



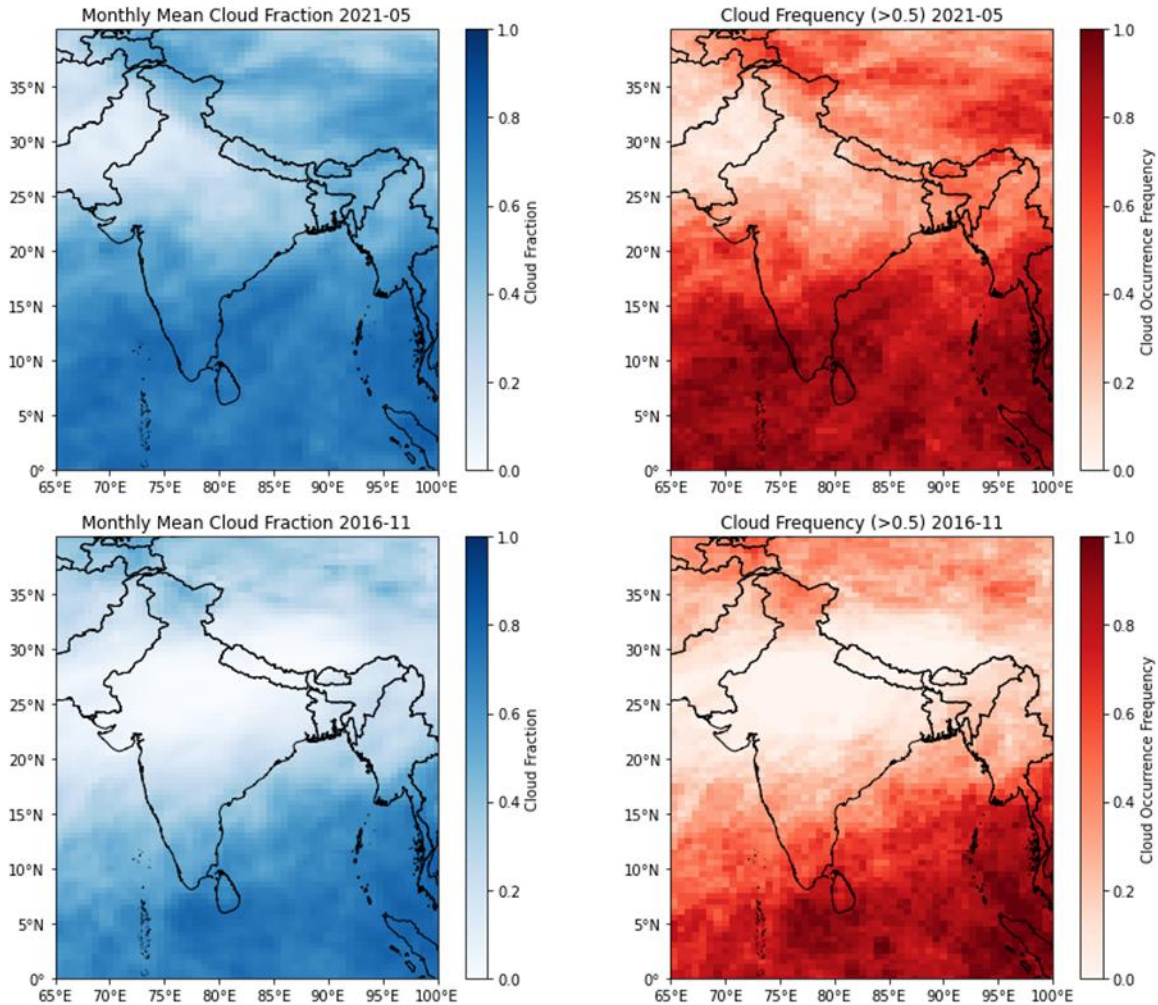
**Figure S32:** Comparison of MODIS [(a), (d) and (g)], MISR [(b), (e) and (h)], and ground observations [(c), (f) and (i)] with fused AOD for the month of January 2012, 2016, and 2021.



**Figure S33:** Similar analysis as in Fig. S32 but for May.

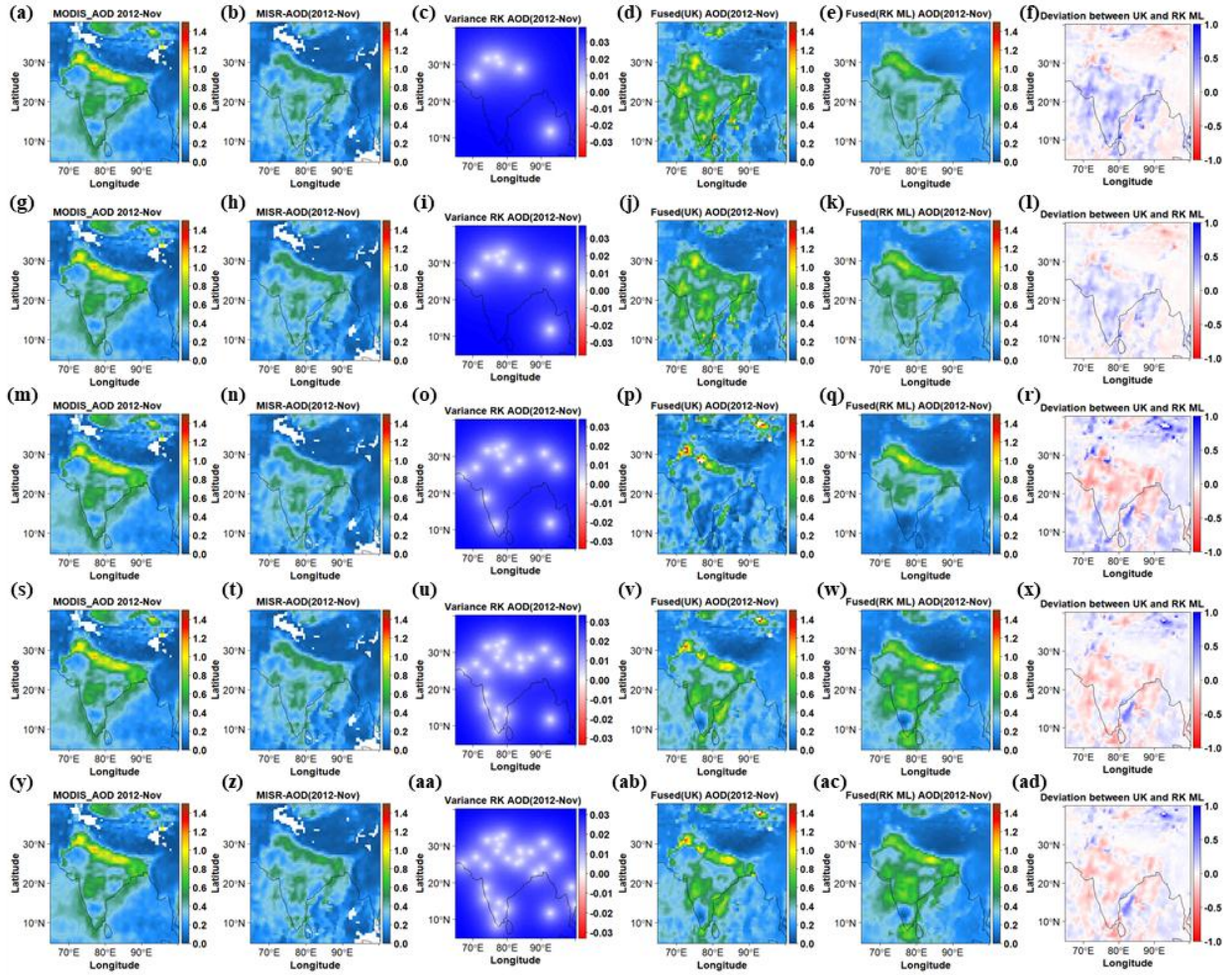


**Figure S34:** Similar analysis as in Fig. S32 but for November.

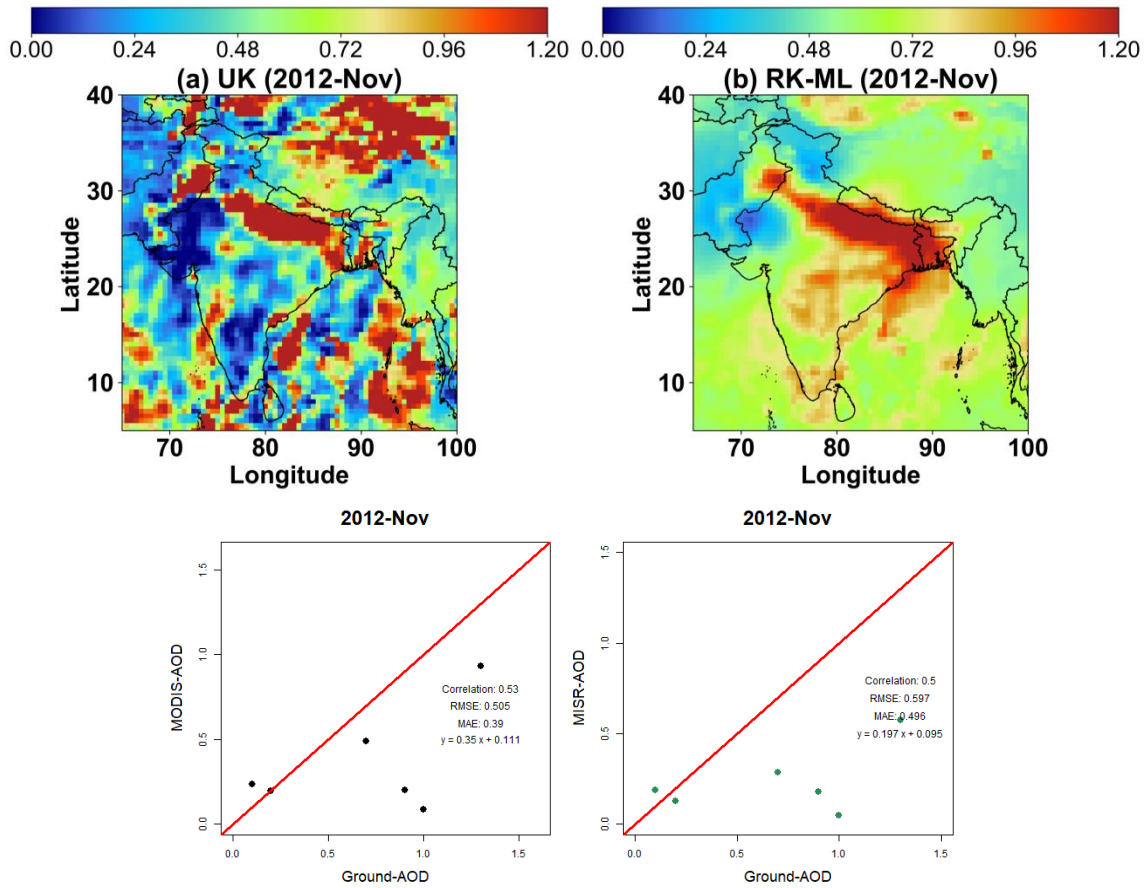


**Figure S35:** Monthly mean cloud fractions (Left panels) and monthly cloud occurrence frequency for the cloud fractions  $>0.5$  (Right panels). For the May 2021 (top panels) and November 2016 (bottom panels). The data shown here is for 10:00 AM – 11:00 AM IST.

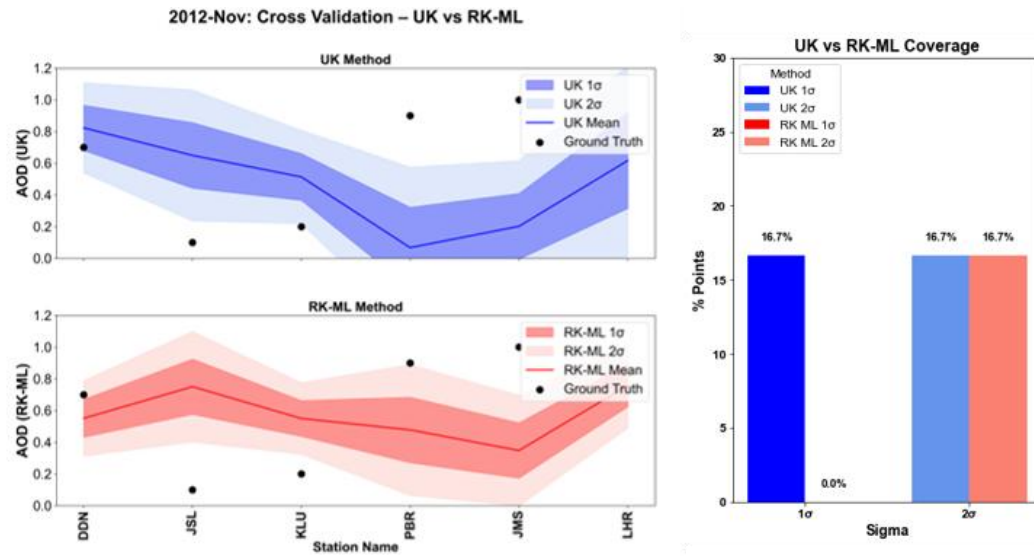
Cloud fraction map for Indian region from MERRA2 cloud fraction indicates the number of cloudy pixels during (10:00 AM – 11:00 AM IST) time, cloud frequency shows how many days were cloudy in that month cloud frequency for cloud fraction  $>0.5$ .



**Figure S36:** Sensitivity study of AOD from two methods (UK and RK-ML) in Nov 2012. The ground points are taken as 6 [(a) – (f)], 8 [(g) – (l)], 13[(m) – (r)], 22[(s) – (x)], 27[(y) – (ad)] respectively. The impact of using less number of ground data points are observed to change the spatial signature from respective satellite sensors in the UK method, while RK-ML is robust to such changes.



**Figure S37:** Sensitivity study of modified AOD from two methods (UK and RK-ML) in Nov 2012. The impact of using less number of ground data (**Table S9**) are seen when erroneous points present.



**Figure S38:** Line plots of LOOCV results from UK method (blue line) and RK-ML method (red line), covering the ground AOD (black dots) for the modified AOD values in Nov 2012 within  $\pm 1\sigma$  (dark shade) and  $\pm 2\sigma$  (light shade). LOOCV comparison between two methods (UK and RK-ML). Both scores are very low and hence, LOOCV scores (right panel) can be treated as one of the criteria to accept the fused estimates using both the methods.