Gaussian Conditional Random Fields for Aggregation of Operational Aerosol Retrievals

Nemanja Djuric, Vladan Radosavljevic, Zoran Obradovic, Slobodan Vucetic

***Abstract*—We present a Gaussian Conditional Random Field model for aggregation of Aerosol Optical Depth (AOD) retrievals from multiple satellite instruments into a joint retrieval. The model provides aggregated retrievals with higher accuracy and coverage than any of the individual instruments, while also providing an estimation of retrieval uncertainty. The proposed model finds an optimal, temporally-smoothed combination of individual retrievals that minimizes Root Mean Squared Error of AOD retrieval. We evaluated the model on five years (2006 - 2010) of satellite data over North America from 5 instruments (Aqua and Terra MODIS, MISR, SeaWiFS, and OMI), collocated with ground-based AERONET ground-truth AOD readings, clearly showing that aggregation of different sources leads to improvements in accuracy and coverage of AOD retrievals.**

***Index Terms*—Remote sensing, Aerosol Optical Depth, Gaussian Conditional Random Fields, Data aggregation.**

I. Introduction

Aerosols have been recognized among the most important quantities in understanding the Earth's climate [1]. Consequently, accurate retrieval of Aerosol Optical Depth (AOD or τ), a measure of an extinction of Solar radiation by scattering and absorption between the top of the atmosphere and the surface, is of great importance to characterize their effect on Earth's radiation budget [2]. Currently, a number of satellite-borne sensors monitor the Earth's atmosphere and report their AOD measurements on a daily basis, such as Moderate Resolution Imaging Spectroradiometer (MODIS) aboard Terra and Aqua satellites [3], Multi-angle Imaging SpectroRadiometer (MISR) aboard Terra [4], Ozone Monitoring Instrument (OMI) aboard Aura [5], or Sea-viewing Wide Field-of-view Sensor (SeaWiFS) aboard SeaStar [6].

The coverage and quality of AOD retrievals from different instruments can vary for a number of reasons. For example, swath of MODIS is $2{,}330km$, allowing MODIS to cover the entire Earth's surface every day, as opposed to $360km$ swath of MISR which results in global coverage only every 9 days. Quality of AOD estimates from different instruments also varies with atmospheric and surface conditions [7]. In addition to satellite sensors, AOD is measured by ground-based, highly accurate sensors from AErosol RObotic NETwork (AERONET) [8]. AERONET instruments are placed at several hundred unevenly distributed locations across the globe, and their measurements are considered a ground-truth. However, AERONET cannot provide global estimation of AOD required for climate models due to limited spatial coverage.

Authors are with Dept. of Computer and Information Sciences, Temple University, USA. e-mail: {nemanja, vladan, zoran, vucetic}@temple.edu

Manuscript received March XX, 2014; revised XX XX, 2014.

Different spatial and temporal coverage, design, and specific mission objectives of the instruments mean that they observe and measure different, possibly complementary, aspects of the same phenomenon. Instead of considering various data sources in isolation, combining retrievals from different sources into a unique, aggregated AOD retrieval might be the best path towards obtaining a higher-quality AOD data product. This was observed in [7], where simple average of collocated Terra MODIS and MISR retrievals led to an improved accuracy of AOD retrieval. This result indicates that further improvements might be possible if more powerful schemes were used. Equally important issue in remote sensing of aerosols, in addition to obtaining point-estimate of AOD, is estimation of retrieval uncertainty. Since AOD retrievals are used as inputs to complex climate models [9], adequate knowledge about the uncertainty and quality of AOD retrievals from satellite instruments is of extreme importance for climate studies.

Several issues need to be considered for the task of aggregation of AOD retrievals. Namely, AOD distribution is characterized by strong temporal and spatial correlation, which could be used to improve the accuracy of aggregated retrievals. Furthermore, due to a number of reasons (e.g., limited coverage of sensors, sensor maintenance, sunglint), it is common that satellite or ground-based retrievals are missing. In this paper we propose an aggregation approach that handles these issues. The method is based on Gaussian Conditional Random Field (GCRF) [10], which assumes that AOD follows Gaussian distribution conditionally dependent on satellite retrievals. The GCRF model can utilize correlations in the values of AOD, while allowing learning and inference in a presence of missing retrievals. Finally, the approach provides an easy-to-calculate point estimate of AOD, as well as estimation uncertainty.

II. Gaussian Conditional Random Field (GCRF)

GCRF is a special type of Conditional Random Field (CRF) that provides a probabilistic framework for incorporation of various aspects of complex data into a single model. Let us denote by $\mathbf{x}$ a vector of covariates, and by $\mathbf{y} = [y_1, \ldots, y_N]^{\mathrm{T}}$ an N-dimensional vector of real-valued output variables. For example, y_i can be an actual AOD at a particular time and place, while $\mathbf{x}$ are all available measurements related to AOD at different times and locations. The conditional distribution $\mathbb{P}(\mathbf{y}|\mathbf{x})$ for CRF can be represented in a convenient form as

$$\mathbb{P}(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta})} \exp\big(-\sum_{i=1}^{N} A(\boldsymbol{\alpha}, y_i, \mathbf{x}) - \sum_{i \sim j} I(\boldsymbol{\beta}, y_i, y_j, \mathbf{x})\big), \quad (1)$$

where $A(\boldsymbol{\alpha}, y_i, \mathbf{x})$ is an *association potential* with weights $\boldsymbol{\alpha}$, $I(\boldsymbol{\beta}, y_i, y_j, \mathbf{x})$ is an *interaction potential* with weights $\boldsymbol{\beta}$, $i \sim j$ denotes that y_i and y_j are assumed correlated (referred to as neighbors), and $Z(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a normalization function.

In general, both learning and inference with model defined in (1) can be difficult due to integration over real-valued $\mathbf{y}$ in $Z(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta})$. However, the potentials could be designed in a way which allows efficient learning and inference. First, let us define the association potential as follows,

$$A(\boldsymbol{\alpha}, y_i, \mathbf{x}) = \sum_{m=1}^{M} \alpha_m \big(y_i - \theta_m(\mathbf{x}_i^m)\big)^2, \tag{2}$$

where $\theta_m(\cdot)$ is the m^{th} baseline predictor, α_m is weight of the m^{th} predictor, $\mathbf{x}_i^m$ is a vector of covariates used by the m^{th} predictor to predict y_i, M is the number of prediction models, and $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_M]^{\mathrm{T}}$. Baseline predictor $\theta_m(\cdot)$ can be any predictor of y_i (e.g., operational aerosol retrieval algorithm for a particular instrument). Quadratic function is easy to interpret: value of y_i close to $\theta_m(\mathbf{x}_i^m)$ is more likely by the model in (1). We can introduce an arbitrary number of baseline predictors and their relevance will be determined during training: relevant predictors will be given bigger α weights, whereas irrelevant ones will get weights close to 0, thus reducing their influence.

Further, let us define the interaction potential as follows,

$$I(\boldsymbol{\beta}, y_i, y_j, \mathbf{x}) = \sum_{l=1}^{L} \beta_l \delta_{ij}^l \cdot (y_i - y_j)^2, \tag{3}$$

where L is the number of interaction (or, neighborhood) definitions, each assigned a different weight β_l, δ_{ij}^l is a $0/1$ indicator function describing whether the i^{th} and the j^{th} outputs are connected according to the l^{th} neighborhood definition, and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_L]^{\mathrm{T}}$. If two outputs are neighbors (e.g., AOD at the same location for two consecutive days), the interaction potential (3) will force them to have similar values.

When the potentials are defined as in (2) and (3), it can be shown that the resulting CRF model corresponds to a multivariate Gaussian distribution $\mathcal{N}\big(\boldsymbol{\mu}(\mathbf{x}), \boldsymbol{\Sigma}(\mathbf{x})\big)$ [10]. For this reason, we call the resulting model the Gaussian CRF. It is important to observe that both mean and covariance matrix are not constant and that they depend on $\mathbf{x}$. However, for the simplicity of notation, we will use $\boldsymbol{\Sigma} \equiv \boldsymbol{\Sigma}(\mathbf{x})$ and $\boldsymbol{\mu} \equiv \boldsymbol{\mu}(\mathbf{x})$.

To obtain explicit expressions for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, let us first define an N-dimensional vector $\mathbf{b} = [b_1, \ldots, b_N]^{\mathrm{T}}$ with elements

$$b_i = 2 \sum_{m=1}^{M} \alpha_m \theta_m(\mathbf{x}_i^m), \tag{4}$$

and $N \times N$ matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ with elements

$$\begin{aligned} \mathbf{Q}_{1ij} &= \begin{cases} \sum_{m=1}^{M} \alpha_m, & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases} \\ \mathbf{Q}_{2ij} &= \begin{cases} \sum_{n=1}^{N} \sum_{l=1}^{L} \beta_l \delta_{in}^l, & \text{if } i = j, \\ -\sum_{l=1}^{L} \beta_l \delta_{ij}^l, & \text{if } i \neq j. \end{cases} \end{aligned} \tag{5}$$

As shown in [10], the inverse of the covariance matrix (i.e., precision matrix) of $\mathbb{P}(\mathbf{y}|\mathbf{x})$ can be calculated as

$$\boldsymbol{\Sigma}^{-1} = 2(\mathbf{Q}_1 + \mathbf{Q}_2), \tag{6}$$

and the mean of $\mathbb{P}(\mathbf{y}|\mathbf{x})$ as $\boldsymbol{\mu} = \boldsymbol{\Sigma}\mathbf{b}$.

A. Training and inference in the GCRF model

Given the model from (1) and a training set $\mathcal{D} = (\mathbf{x}, \mathbf{y}) = \{(\mathbf{x}_i, y_i)\}_{i=1,\ldots,N}$, the training task is to find $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that the conditional log-likelihood $\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is maximized,

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}) = \arg\max_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}), \text{ where } \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \log \mathbb{P}(\mathbf{y}|\mathbf{x}), \tag{7}$$

solved using gradient descent [10]. On the other hand, given a trained GCRF model with parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the inference task is to find the point estimate $\hat{\mathbf{y}}$ of outputs $\mathbf{y}$ for given inputs $\mathbf{x}$. We select $\hat{\mathbf{y}}$ that maximizes $\mathbb{P}(\mathbf{y}|\mathbf{x})$, equal to the expected value $\boldsymbol{\mu}$,

$$\hat{\mathbf{y}} = \boldsymbol{\mu} = \boldsymbol{\Sigma}\mathbf{b}. \tag{8}$$

An important property of the GCRF model is that the uncertainty of the calculated point estimates is easily computed. More specifically, 95%-confidence intervals of outputs $\hat{\mathbf{y}}$ are estimated from the mean and the covariance matrix as

$$\mathbb{P}\big(\hat{\mathbf{y}} - 1.96 \cdot \mathrm{diag}(\boldsymbol{\Sigma}) \leq \mathbf{y} \leq \hat{\mathbf{y}} + 1.96 \cdot \mathrm{diag}(\boldsymbol{\Sigma})\big) = 0.95, \tag{9}$$

where $\mathrm{diag}(\boldsymbol{\Sigma})$ denotes the main diagonal of $\boldsymbol{\Sigma}$ matrix.

B. Handling the missing predictors

In many real-world applications, it is often the case that some baseline predictors $\theta_m(\mathbf{x}_i)$ might not be available. To address this issue, we define the association potential as

$$A(\boldsymbol{\alpha}, y_i, \mathbf{x}) = \sum_{m=1}^{M} \alpha_m \delta_i^m \big(y_i - \theta_m(\mathbf{x}_i^m)\big)^2, \tag{10}$$

where we introduced $0/1$ functions δ_i^m equal to 1 if the m^{th} baseline predictor provided prediction for the i^{th} output, and 0 otherwise. This results in slightly modified expressions (4) and (5) for $\mathbf{b}$ and $\mathbf{Q}_1$, where α_m is replaced by $\alpha_m \delta_i^m$.

III. GCRF FOR AGGREGATION OF AOD RETRIEVALS

We address the problem of aggregation of satellite AOD retrievals. More formally, we assume we are given training set $\mathcal{D} = \{\theta_i^{aqua}, \theta_i^{omi}, \theta_i^{sw}, \theta_i^{terra}, \theta_i^{misr}, y_i\}_{i=1,\ldots,N}$, where N is size of the data set, index i corresponds to AOD at particular time and location, y_i is an AERONET retrieval taken as a ground-truth due to instrument's high accuracy, and, to make the notation more intuitive, instead of covariates $\mathbf{x}$ and baseline predictors $\theta_m(\mathbf{x}_i^m)$ introduced in (2), $m = 1, \ldots, 5$, we used θ_i^{aqua}, θ_i^{omi}, θ_i^{sw}, θ_i^{terra}, θ_i^{misr}, respectively, which denote operational AOD retrievals from Aqua MODIS, OMI, SeaWiFS, Terra MODIS, and MISR, respectively. Furthermore, training points with available AERONET retrieval are referred to as labeled, otherwise they are referred to as unlabeled points.

Due to the limitations of the sensors, such as limited coverage and sensitivity to atmospheric conditions, it is common that $\mathcal{D}$ has large number of missing satellite and ground-based retrievals. As an example, let us consider a typical spatial coverage of Terra and Aqua MODIS, MISR, and AERONET instruments in the USA during a single day, given in Figure 1. We can see that different areas may have very different coverage. For most of labeled data points (e.g., data with AERONET retrieval from locations A, B, and E) and

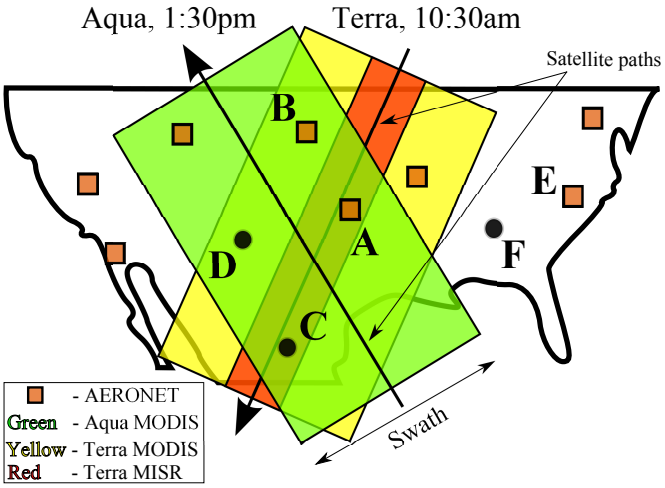


Fig. 1: Coverage of instruments over the USA

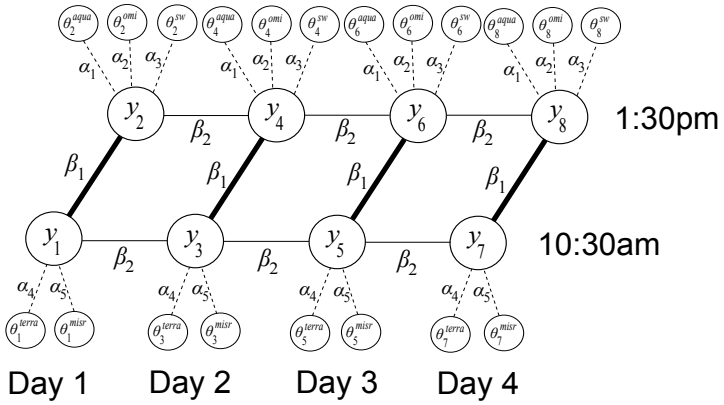


Fig. 2: Graphical representation of GCRF for retrieval aggregation (thick line: within-day interaction; thin line: between-day interaction; dashed line: association between inputs and outputs)

unlabeled data points (e.g., data without AERONET retrieval from locations C, D, and F), AOD retrievals from some of the satellite instruments are missing. Locations A and C have retrievals from all 3 satellite sensors, while B and D are just outside MISR's swath and do not have its AOD retrieval. Locations E and F are outside the swaths of all instruments, and do not have any satellite retrieval available. Moreover, we note that even when a location is covered by an instrument, availability of retrievals also depends on sensitivity of the instruments to technical, atmospheric, and surface conditions, further exacerbating the problem of limited coverage.

Satellite sensors considered in this study can be divided into two groups, one that collocates with AERONET sites around 10:30*am* local time (comprising Terra MODIS and MISR instruments), and another that does so around 1:30*pm* local time (comprising Aqua MODIS, OMI, and SeaWiFS instruments). We choose to provide aggregated AOD retrievals at these two discrete time-points every day at every location. Thus, each location for each day contributes two data points to $\mathcal{D}$, one with AOD retrievals from two satellite instruments with morning overpass, and the other data point with AOD retrievals from three satellite instruments with afternoon overpass.

A. GCRF aggregation model

The graphical representation of the GCRF model for AOD retrieval aggregation, derived from the GCRF model from Section II-B, is shown in Figure 2. We did not consider spatial correlations due to sparse distribution of AERONET sites in our data set, and we set all interaction weights β between i^{th} and j^{th} outputs to zero if i and j correspond to different locations. As a result, different AERONET sites are independent, and Figure 2 corresponds to a single location. We note that the same $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ parameters are used for all locations. As we consider retrievals from 5 satellite sensors, we set $M = 5$ in equation (2) for the association potential, and represent influence of satellite retrievals on outputs with a dashed line in Figure 2. For the interaction potential, we assume AOD values are temporally correlated. To encode this assumption, we linked within-day outputs at 10:30*am* and 1:30*pm*, and associate weight β_1 with these links (thick lines in Figure 2). We also linked 10:30*am* outputs from two consecutive days, as well as 1:30*pm* outputs from two consecutive days, and associate weight β_2 with these day-to-day links (thin lines in Figure 2). The resulting potentials are

$$\begin{aligned} A(\boldsymbol{\alpha}, y_i, \mathbf{x}) &= \alpha_1 \delta_i^1 (y_i - \theta_i^{aqua})^2 + \alpha_2 \delta_i^2 (y_i - \theta_i^{omi})^2 \\ &\quad + \alpha_3 \delta_i^3 (y_i - \theta_i^{sw})^2 + \alpha_4 \delta_i^4 (y_i - \theta_i^{terra})^2 \\ &\quad + \alpha_5 \delta_i^5 (y_i - \theta_i^{misr})^2, \\ I(\boldsymbol{\beta}, y_i, y_j, \mathbf{x}) &= \beta_1 \delta_{ij}^1 \cdot (y_i - y_j)^2 + \beta_2 \delta_{ij}^2 \cdot (y_i - y_j)^2, \end{aligned} \tag{11}$$

where δ_i^m, $m \in \{1, \ldots, 5\}$ are $0/1$ indicator functions returning 1 if, for the i^{th} output, there is an available retrieval from Aqua MODIS, OMI, SeaWiFS, Terra MODIS, and MISR, respectively, and 0 otherwise, while δ_{ij}^1 and δ_{ij}^2 are $0/1$ indicator functions returning 1 if y_i and y_j are within-day neighbors or between-day neighbors, respectively, and 0 otherwise. We note that indicator functions δ_i^1, δ_i^2, and δ_i^3 always return 0 for odd-indexed y_i, while δ_i^4 and δ_i^5 always return 0 for even-indexed outputs.

GCRF aggregation model has very useful properties. In particular, it can aggregate baseline retrievals even when an arbitrary number of baseline predictors are unavailable for some output y_i. GCRF model learns the importance of different baselines by assigning higher α weight to more accurate ones, and can also utilize temporal correlations in AOD values, where strength of temporal correlations is learned and quantified through β weights. Moreover, GCRF readily provides uncertainty estimates of output retrievals. Lastly, assuming we use indexing of data points shown in Figure 2, the resulting matrix $\boldsymbol{\Sigma}^{-1}$ in (6) is pentadiagonal. Consequently, inverse of $\boldsymbol{\Sigma}^{-1}$, required in the gradient-based optimization, can be found in very favorable $\mathcal{O}(N)$ time [11].

B. GCRF training with missing AERONET AOD retrievals

As discussed previously, it is possible to have unlabeled data points in training data. In this case the aggregation of satellite retrievals becomes a semi-supervised task. In particular, let us denote the joint probability of labeled and unlabeled points as

TABLE II: RMSE of instruments and GCRF (upper part shows $1{:}30pm$, middle part $10{:}30am$, bottom part labeled data results)

Coverage	No. of points	Individual sensors	Only β_1	Only β_2	All $\boldsymbol{\beta}$
All MODIS Aqua	7,246	0.0872	0.0822	0.0825	**0.0756**
All OMI	10,142	0.2390	0.1930	0.1482	**0.0934**
All SeaWiFS	2,205	0.0739	0.0630	0.0642	**0.0607**
MODIS Aqua alone	2,102	0.0889	0.0850	0.0773	**0.0741**
OMI alone	4,528	0.2934	0.2744	0.2010	**0.1114**
SeaWiFS alone	237	0.0800	0.0771	0.0784	**0.0753**
MODIS + OMI	3,868	0.0893 | 0.2123	0.0886	0.0925	**0.0827**
MODIS + SeaWiFS	222	0.0982 | 0.0747	0.0665	0.0655	**0.0634**
OMI + SeaWiFS	692	0.1011 | 0.0837	0.0758	0.0760	**0.0691**
MODIS + OMI + SeaWiFS	1,054	0.0717 | 0.0715 | 0.0650	**0.0475**	0.0506	0.0495
All MODIS Terra	8,725	0.0905	0.0826	0.0819	**0.0800**
All MISR	2,165	**0.0652**	0.0664	0.0735	0.0725
MODIS Terra alone	7,552	0.0863	0.0845	0.0822	**0.0806**
MISR alone	992	**0.0618**	0.0636	0.0657	0.0677
MODIS + MISR	1,173	0.1142 | **0.0680**	0.0686	0.0795	0.0763
All labeled	44,445	–	0.1634	0.1328	**0.1080**
All labeled with any satellite	22,420	–	0.1430	0.1158	**0.0852**
All labeled without satellites	20,025	–	0.1817	0.1481	**0.1271**

TABLE I: Coverage of baseline predictors in the 5-year data

Instrument	No. of days	Coverage
Terra MODIS	20,286	20.37%
MISR	4,906	4.93%
Aqua MODIS	16,711	16.78%
OMI	23,332	23.43%
SeaWiFS	4,567	4.59%

$\mathbb{P}(\mathbf{y}_L, \mathbf{y}_U|\mathbf{x}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and let us separate prediction vector $\boldsymbol{\mu}$ and precision matrix $\boldsymbol{\Sigma}^{-1}$ into labeled and unlabeled parts as $\boldsymbol{\mu} = [\boldsymbol{\mu}_L^{\mathrm{T}}, \boldsymbol{\mu}_U^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{\Sigma}^{-1} = [\mathbf{Q}_{LL}, \mathbf{Q}_{LU}; \mathbf{Q}_{UL}, \mathbf{Q}_{UU}]$. Then, the marginal distribution of the labeled part is equal to $\mathbb{P}(\mathbf{y}_L|\mathbf{x}) \sim \mathcal{N}(\boldsymbol{\mu}_L, (\mathbf{Q}_{LL} - \mathbf{Q}_{LU}\mathbf{Q}_{UU}^{-1}\mathbf{Q}_{UL})^{-1})$. To obtain a semi-supervised training procedure, we redefine the likelihood as $\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \log \mathbb{P}(\mathbf{y}_L|\mathbf{x})$ and maximize modified (7).

IV. Experiments

We used ground-based AERONET data [8] and data from 5 satellite instruments from 2006 to 2010, collected over North America, and considered AOD at $550nm$ wavelength. If a source did not provide AOD retrievals at this wavelength, we performed a linear interpolation or extrapolation in the log-scale of retrievals at two closest wavelengths to $550nm$ [3]. There were, on average, 56 working AERONET sites each year, and we estimated AOD twice a day at every AERONET location, resulting in a total of 199,134 data points in the collected data set. In all experiments we report results after leave-one-year-out cross-validation. After 5 repetitions, test data were pooled together and Root Mean Squared Error (RMSE) on labeled test points was calculated and reported.

For the AERONET sensors we downloaded data for all sites from AERONET website[1], and used the highest-quality Level 2.0 AOD data. For the satellite sensors, we downloaded data from Multi-sensor Aerosol Products Sampling System (MAPSS) website[2]. We only used research-quality AOD data products: 1) **MODIS** - Daily Level 2 aerosol data product, MODIS collection 5.1 (MOD04_L2 and MYD04_L2, for Terra and Aqua, respectively); 2) **MISR** - MIL2ASAE data product, a MISR Level 2 aerosol product; 3) **OMI** - OMAERUV, a Level-2 near-UV aerosol absorption and extinction optical depth and single scattering albedo OMI data product; 4) **SeaWiFS** - SWDB_L2, Deep Blue Aerosol Optical Depth Daily Level 2 data product. Coverage of operational retrievals is given in Table I. We see that MISR and SeaWiFS instruments cover less than 5% of days during the 5-year period, while OMI has the highest coverage of nearly 24%.

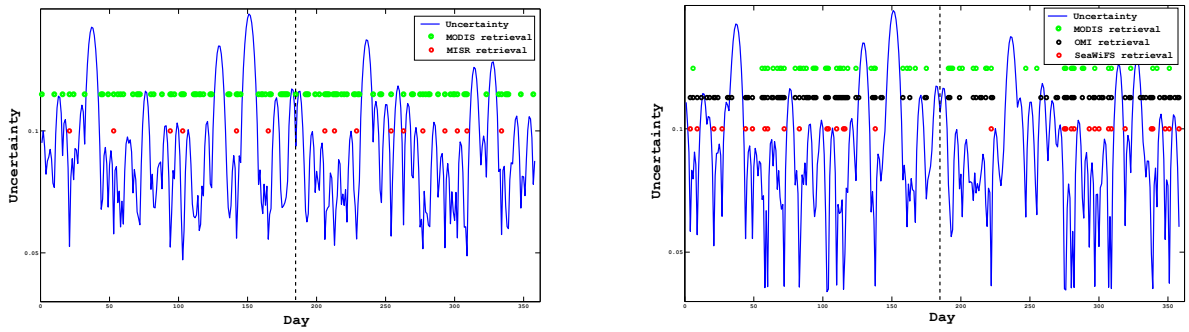
We report RMSE only on subsets of labeled points, summarized in Table II. For example, row "*MODIS + OMI*" reports results on a subset of labeled points that have both Aqua MODIS and OMI retrieval, but not SeaWiFS retrieval. In the bottom part we do not list RMSE when a model did not have a complete coverage on test data.

Let us first discuss results of baseline predictors shown in the "*Individual sensors*" column. The results indicate that the performance of different instruments varied significantly, both in accuracy and coverage. Regarding $1{:}30pm$ results, we see that SeaWiFS retrievals overall were more accurate than both Aqua MODIS and OMI retrievals. It is interesting to observe that OMI accuracy was particularly low when MODIS and SeaWiFS retrievals were not available. Whenever OMI provided AOD retrieval along with some other instrument, its accuracy improved significantly, which indicates there could be certain issues with the quality checks of OMI retrievals. We can also see that SeaWiFS consistently outperformed Aqua MODIS, but that it had around 3 times smaller coverage. Regarding $10{:}30am$ results, MISR achieves 38% lower RMSE than Terra MODIS, while having 4 times smaller coverage. Interestingly, the accuracy of MODIS in the absence of MISR was much higher than its accuracy when MISR was available, explained by larger MODIS sensitivity to sunglint [4].

Further, we considered GCRF models with increasing levels of complexity, and investigated how the introduction of β parameters influences performance. For that purpose, we trained GCRF with different combinations of β parameters: (1) model without interactions ($\beta_1 = \beta_2 = 0$); (2) model with diurnal interaction ($\beta_1 \neq 0$, $\beta_2 = 0$); (3) model with day-to-day interaction ($\beta_1 = 0$, $\beta_2 \neq 0$); (4) model with both interactions ($\beta_1 \neq 0$, $\beta_2 \neq 0$). In the bottom three rows we see that the overall RMSE on all labeled points dropped significantly when we included more temporal interactions in the model. Overall

[1] aeronet.gsfc.nasa.gov/cgi-bin/combined_data_access_new, March 2014

[2] disc.sci.gsfc.nasa.gov/aerosols/services/mapss/mapssdoc, March 2014



(a) Uncertainty at 10:30am

(b) Uncertainty at 1:30pm

Fig. 3: GCRF uncertainty estimate for AERONET site located at Maryland Science Center, Baltimore, for year 2007 (colored dots denote that satellite AOD retrieval is available for that particular day; dashed vertical line denotes day 185)

RMSE on labeled points with at least one satellite retrieval dropped from 0.1430 to 0.0852 when both day-to-day and diurnal interactions were included, an improvement of 41%.

Compared to the RMSE of individual instruments, GCRF accuracies for 1:30pm have improved. In the case of 10:30am results, RMSE on data points with MISR retrievals increased due to temporal averaging enforced by β_1 and β_2. Thus, it seems that we should use retrievals from very accurate MISR retrievals when available, and use the power of temporal smoothing when MISR retrieval is missing. However, it is important to note that only 4.76% of data points had high-quality MISR or SeaWiFS retrievals. Moreover, results shown in upper and middle parts of Table II cover only 22,420 labeled points that had retrievals from at least one of the instruments. This amounts to 11.2% coverage of the whole data set, while GCRF model provides aggregated retrieval for all 199,134 points due to its ability to leverage temporal correlation.

In Figure 3 we give an example of daily uncertainty estimates for the AERONET site at Maryland Science Center in Baltimore, USA. Colored circles represent availability of individual satellite retrievals. We see that uncertainty gradually increases with distance from the nearest satellite retrieval, an example of the influence of the β_2 parameter modeling the daily interaction between retrievals. It is interesting to observe a drop in uncertainty at day 185 in Figure 3b, although no instrument retrieved AOD at 1:30pm on that or any of the neighboring days. This can be explained by the Terra MODIS retrieval observed at 10:30am that day, as seen in Figure 3a. This is an example of the influence of the β_1 parameter, which models the interaction between 10:30am and 1:30pm retrievals. Influence of $\boldsymbol{\alpha}$ parameters is also visible in Figure 3; the availability of MISR or SeaWiFS retrievals always results in sharp drops in uncertainty, as these instruments were assigned the largest α parameters (omitted for lack of space).

V. Conclusion

We presented a GCRF model for fusion of AOD retrievals from multiple instruments. Ease of modeling interaction between outputs, ability to handle missing data, high-quality of the aggregated AOD retrievals, as well as the interpretability of its outputs, strongly suggest that the GCRF model can represent an important tool in remote sensing applications.

Acknowledgment

This work was funded by NSF grant, no. NSF-IIS-1117433.

References

[1] R. J. Charlson, S. E. Schwartz, J. M. Hales, R. D. Cess, J. A. Coakley, J. E. Hansen, and D. J. Hofmann, "Climate Forcing by Anthropogenic Aerosols," *Science*, vol. 255, no. 5043, pp. 423–430, 1992.

[2] R. C. Levy, L. A. Remer, and O. Dubovik, "Global aerosol optical properties and application to Moderate Resolution Imaging SpectroRadiometer aerosol retrieval over land," *Journal of Geophysical Research*, vol. 112, pp. 13,210–13,224, 2007.

[3] L. A. Remer, Y. J. Kaufman, D. Tanré, S. Mattoo, D. A. Chu, J. V. Martins, R.-R. Li, C. Ichoku, R. C. Levy, R. G. Kleidman, T. F. Eck, E. Vermote, and B. N. Holben, "The MODIS Aerosol Algorithm, Products, and Validation," *Journal of Atmospheric Science*, vol. 62, pp. 947–973, 2005.

[4] R. A. Kahn, D. L. Nelson, M. J. Garay, R. C. Levy, M. A. Bull, D. J. Diner, J. V. Martonchik, S. R. Paradise, E. G. Hansen, and L. A. Remer, "MISR Aerosol Product Attributes and Statistical Comparisons With MODIS," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 12, pp. 4095–4114, dec. 2009.

[5] O. Torres, R. Decae, J. P. Veefkind, and G. de Leeuw, "OMI aerosol retrieval algorithm," in *OMI Algorithm Theoretical Basis Document, Volume III, Clouds, Aerosols, and Surface UV Irradiance*, P. Stammes, Ed., 2002. [Online]. Available: http://eospso.gsfc.nasa.gov/eos_homepage/for_scientists/atbd/

[6] M. Wang, S. Bailey, and C. R. Mcclain, "SeaWiFS Provides Unique Global Aerosol Optical Property Data," *EOS*, vol. 81, no. 18, 2000.

[7] M. I. Mishchenko, L. Liu, I. V. Geogdzhayev, L. D.Travis, B. Cairns, and A. A. Lacis, "Toward unified satellite climatology of aerosol properties. 3. MODIS versus MISR versus AERONET," *Journal of Quantitative Spectroscopy & Radiative Transfer*, vol. 111, pp. 540–552, 2010.

[8] B. N. Holben, T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, "AERONET - A Federated Instrument Network and Data Archive for Aerosol Characterization," *Remote Sensing of Environment*, vol. 66, no. 1, pp. 1–16, 1998.

[9] K. Zhang, D. O'Donnell, J. Kazil, P. Stier, S. Kinne, U. Lohmann, S. Ferrachat, B. Croft, J. Quaas, H. Wan, S. Rast, and J. Feichter, "The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations," *Atmospheric Chemistry and Physics*, vol. 12, no. 19, pp. 8911–8949, 2012.

[10] V. Radosavljevic, S. Vucetic, and Z. Obradovic, "Continuous Conditional Random Fields for Regression in Remote Sensing," in *Proceedings of European Conference on Artificial Intelligence (ECAI-10)*, 2010.

[11] A. D. A. Hadj and M. Elouafi, "A fast numerical algorithm for the inverse of a tridiagonal and pentadiagonal matrix," *Applied Mathematics and Computation*, vol. 202, no. 2, pp. 441–445, 2008.