ALiDAn: Spatiotemporal and Multiwavelength Atmospheric Lidar Data Augmentation

Adi Vainiger, *Graduate Student Member, IEEE*, Omer Shubi, Yoav Y. Schechner, *Member, IEEE*, Zhenping Yin, Holger Baars, Birgit Heese, and Dietrich Althausen

***Abstract*— Methods based on statistical learning have become prevalent in various signal processing disciplines and have recently gained traction in atmospheric lidar studies. Nonetheless, such methods often require large quantities of annotated or resolved data. Such data are rare and require effort, especially when exploring evolving phenomena. Existing simulators and databases primarily focus on atmospheric vertical profiles. We propose the Atmospheric Lidar Data Augmentation (ALiDAn) framework to fill this gap. ALiDAn serves as an end-to-end generation and augmentation framework of spatiotemporal and multiwavelength resolved lidar simulated data. ALiDAn employs a hybrid approach of physical models, data statistics, and sampling processes. In addition, it takes into account geographical and seasonal characteristics of aerosols and meteorological conditions along with short- and long-term phenomena that affect lidar measurements. This approach can provide diversified data and robust benchmarks to assist in developing and validating new lidar processing algorithms. We demonstrate simulations compatible with a pulsed time-of-flight lidar. Our approach leverages a broader use of existing databases and can inspire similar data augmentation to other types of lidars and active sensors.**

***Index Terms*— Atmospheric lidar, data augmentation, data-driven models, lidar simulations and databases, statistical learning.**

I. Introduction

STATISTICAL learning tools have a significant potential to improve retrieval quality and reduce run-time in atmospheric lidar research. A ground-based aerosol lidar is a key tool in atmospheric sensing. A lidar, however, is affected by changes caused by external factors and internal wear, such as in-door temperature fluctuations affecting optical alignment and laser energy. Thus, the analysis should address the dynamic phenomena of both the atmosphere and the lidar system. Lidar analysis is often done in synergy with other systems [1], [2], [3]. Thus, it requires many resources, including expert supervision to obtain well-validated continuous retrievals.

A lidar essentially samples vertical profiles in the atmosphere. Over the course of a day, consecutive vertical profiles create a two-dimensional (2D) spatiotemporal map. Advancing spatiotemporal lidar research can be achieved by adapting powerful deep-learning (DL) methods that had become highly developed in image processing. Recent image processing methods, such as segmentation and detection [4], [5], denoising [6], [7], and classification [8], have shown that spatiotemporal analysis better overcomes challenges posed by sparse or low-SNR signals. The use of sequential methods has also gained attention in recent boundary layer studies [9], [10], [11].

DL uses inputs that follow the statistical distribution of real measurements. However, this often requires large databases to facilitate the training. We suggest building a ground-truth *spatiotemporal and multiwavelength* (STMW) lidar database to support such approaches. According to our hypothesis, let us define the requirements for a desirable database.

1) The database should contain resolved or annotated STMW lidar data. This allows developing and implementing supervised or semisupervised learning algorithms.
2) The database should account for typical variability factors affecting ground-based lidar measurements, including the spatiotemporal variation of aerosol type and density; diurnal and environmental signal effects along with system-related effects; and seasonal- and geographical-related effects. This allows developing robust algorithms by expanding their scope to natural lidar signals.

Currently, there are limited continuously validated lidar databases [12], [13]. Inspired by the recent contribution of synthesized databases to learning-based methods in various disciplines [14], [15], [16], we suggest augmenting the desired data with simulations. Potentially, we can achieve this using

Manuscript received 16 January 2022; revised 26 June 2022; accepted 27 July 2022. Date of publication 30 August 2022; date of current version 1 December 2022. This work was supported in part by the European Research Council (ERC) through the European Union's (EU's) Horizon 2020 Research and Innovation Program under Grant Agreement 810370: CloudCT and Grant 654109: ACTRIS-2, in part by the EU Seventh Framework Programme (FP7) 2007–2013 under Grant 262 254: ACTRIS, and in part by the German-Israeli Science Foundation (GIF) under Grant I-1262-401.10/2014. The work of Yoav Y. Schechner was supported by the Taub Foundation. *(Corresponding author: Adi Vainiger.)*

Adi Vainiger and Yoav Y. Schechner are with the Viterbi Faculty of Electrical and Computer Engineering, Technion—Israel Institute of Technology, Haifa 3200003, Israel (e-mail: addalin@campus.technion.ac.il).

Omer Shubi is with the Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa 3200003, Israel.

Zhenping Yin is with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China.

Holger Baars, Birgit Heese, and Dietrich Althausen are with the Leibniz Institute for Tropospheric Research, 04318 Leipzig, Germany.

This article has supplementary downloadable material available at https://doi.org/10.1109/TGRS.2022.3201436, provided by the authors.

Digital Object Identifier 10.1109/TGRS.2022.3201436

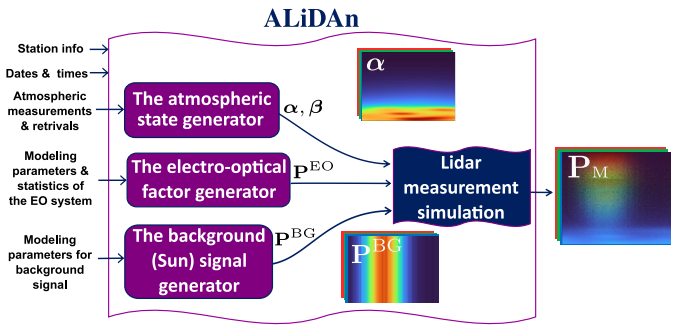


Fig. 1. Diagram of the ALiDAn framework.

existing lidar databases and simulators. We elaborate on these tools and their limitations concerning our goal in Section II. Following this review, we conclude that each tool separately has not yet met the requirements above.

We present the Atmospheric Lidar Data Augmentation (ALiDAn) framework.[1] ALiDAn is a model- and data-driven generation and augmentation approach of STMW resolved atmospheric lidar measurements. In addition to spatial changes, we also include short- and long-term temporal changes. The proposed framework is adaptable to varying geographical locations and times.

To achieve diversified simulated lidar data, we define and employ a sampling space of typical generation parameters of each measurement component. Generation parameters are derived and augmented from an ensemble of physical models, measurements, and past analysis [1], [2]. The simulations demonstrated in this work are compatible with a pulsed time-of-flight [17], [18], [19] lidar. We derive simulation parameters from the literature and measurements collected in a field campaign during 2016–2019. In this campaign, a Polly^{XT} [20] system was assembled on the rooftop of the Meyer Building (Electrical Engineering (EE) Faculty), Technion, Haifa, Israel [21], [22], [23], [24], [25]. To generate simulation parameters adapted to other scenarios, we suggest alternative sources throughout this work.

The suggested framework of ALiDAn consists of three generation modules of measurement factors, as shown in Fig. 1: the atmosphere (aerosols and air) presented in Section V; the *electro-optical* (EO) system factor presented in Section VI; and the *background* (BG) signal (scattered sunlight) presented in Section VII. We provide an additional simulation model that uses the augmented parameters to generate a new STMW simulated lidar measurement, as presented in Section IV. We provide simulated products of ALiDAn modules throughout Sections V–VII. In Section VIII, we demonstrate a fully STMW simulated lidar measurement and, accordingly, a real one for qualitative comparison. The outline of this work is presented in Table I.

Using the ALiDAn framework, we demonstrate the concept of resolved atmospheric lidar database augmentation. As a result, augmented databases can better accommodate real-measurements statistics. Thereby, one can apply future data-driven algorithms to real data following minimal modification. Such a framework may assist various lidar processing tasks, such as inversion, aerosol/cloud classification, and aerosol typing. Its modular nature could benefit various types of lidar or sensors along with future extensions or modifications.

[1]The project repository is available at https://github.com/Addalin/pyALiDAn

TABLE I
OUTLINE OF ALiDAn

Section	Topic
III	Atmospheric lidar modeling
IV	Lidar measurement simulation modeling
V	The atmospheric state generator
VI	The electro-optical factor generator
VII	The background signal generator
VIII	Demonstration of simulated lidar measurements

II. PRIOR DATABASES AND SIMULATORS

Here, we review and examine possible ways of using existing tools to obtain the desired lidar data, as discussed previously in Section I. The European Aerosol Research Lidar Network (EARLINET) provides one-dimensional (1D) synthetic vertical lidar profiles for typical aerosol conditions [26], [27]; the Lidar Climatology of Vertical Aerosol Structure (LIVAS) database [28] provides aerosol and cloud properties in multiple wavelengths that are designed to be used in space-based lidar simulations. However, both do not address the time-varying atmospheric state or measurement-related phenomena that affect ground-based lidars. The Optical Properties of Aerosols and Clouds (OPAC) database [29], [30] offers a few predefined 1D atmospheric profiles or a user interface for manual specifications [31]; however, OPAC has not addressed temporal atmospheric variations.

A possible way to create simulated STMW validated lidar data is using existing simulators. Radiative transfer simulators [32], [33], [34], [35] usually require predefined ingredients, such as aerosol profiles commonly obtained from OPAC. The radiative transfer library (LIBRADTRAN) [36] offers a lidar model. In addition to aerosol profiles, LIBRADTRAN requires predefined lidar system parameters. However, some parameters, such as the overlap between the laser beam and the receiver field-of-view, are not addressed.

The Earth Clouds, Aerosols, and Radiation Explorer (EarthCARE) [37] simulator (ECSIM) [38], [39], [40] refers to a space-based multisensor platform, which also includes a lidar [41]. ECSIM is a comprehensive simulator that addresses a setup different than that of this article. The lidar measurements of EarthCARE are affected differently by external and internal factors. Therefore, adapting ECSIM to ground-based lidars is not straightforward.

The polarimetric lidar simulator (PLS) [42] provides a tool for a thorough simulation and assessment of the instrument components. Nonetheless, it does not include the statistics of the environment inducing the backscatter and the background signals.

The existing simulators and databases mentioned above focus on vertical profiles and do not address the dynamic

phenomena of both the system and the atmosphere. To emulate such data with current simulators, one should manually provide the system and atmospheric state variations. Since the system and the atmosphere change over time, their respective simulations should also vary according to real statistics. Here, we suggest a framework to accomplish such simulations practically.

The value of using statistical databases in lidar study has already been demonstrated in aerosol typing study [43]. The ground-based lidar networks, EARLINET [44], MPLNET [45], and PollyNet [20], [46], have accumulated data globally and over time. Recently, the *PollyNet Processing Chain* [2], [47] created a vast database of lidar-derived aerosol optical property profiles from past years for several stations around the globe. These data, together with AERONET inversion products [48], yield rich statistics. Here, we demonstrate how to leverage these databases to achieve lidar data augmentation.

III. Atmospheric Lidar Modeling

A. Lidar Model

An outdoor scene has downwelling radiance induced by natural lighting at wavelength λ[nm]. Denote by $p^{\mathrm{BG}}(\lambda)$[photons] the resulting BG number of photons detected by a detector on the ground during exposure time δt[sec]. An atmospheric lidar [49] is mounted on the ground, pointing upward. The lidar operates in a pulsed *time-of-flight mode*, emitting $p_0(\lambda)$ [photons] per pulse. The temporal pulsewidth δt of the lidar and the *speed-of-light* c [km/sec] define the lidar's axial resolution $\Delta r = (c\delta t/2)$ [km]. The lidar radiation interacts with the atmosphere and backscatters downward to the detector. This creates a *backscattered* (BS) signal $p^{\mathrm{BS}}(r, \lambda)$[photons]. The number of photons per pulse measured at λ, corresponding to a volume element at height r [km] above the lidar, is

$$p(r, \lambda) = p^{\mathrm{BS}}(r, \lambda) + p^{\mathrm{BG}}(\lambda) \text{ [photons]}. \tag{1}$$

Now, we derive $p^{\mathrm{BS}}(r, \lambda)$. Let the backscatter and extinction coefficients of particles at r be $\beta(r, \lambda)$ [1/(km · sr)] and $\alpha(r, \lambda)$ [1/km], respectively. The unitless optical depth corresponding to distance r is

$$\tau(r, \lambda) = \int_0^r \alpha(z, \lambda) dz. \tag{2}$$

The *attenuated backscatter* coefficient for elastic interactions is

$$\beta_{\mathrm{ATTN}}(r, \lambda) = \beta(r, \lambda) \exp[-2\tau(r, \lambda)] \left[\frac{1}{\mathrm{km} \cdot \mathrm{sr}}\right]. \tag{3}$$

The lidar detector has area A [km^2]. With a unitless system efficiency η, the *lidar constant* (LC) [1], [50] (sometimes referred to as the lidar factor) is

$$p^{\mathrm{LC}}(\lambda) = p_0(\lambda) \Delta r A \eta \text{ [photons} \cdot \mathrm{km}^3\text{]}. \tag{4}$$

In practice, the lidar laser beam does not fully coincide with the *field-of-view* of the detector. The portion of the laser beam that can geometrically be detected varies continuously with height. This portion is expressed by the unitless *overlap function* $O(r)$. We define an EO system factor

$$p^{\mathrm{EO}}(r, \lambda) \triangleq p^{\mathrm{LC}}(\lambda) \cdot O(r) \text{ [photons} \cdot \mathrm{km}^3\text{]}. \tag{5}$$

Then,

$$p^{\mathrm{BS}}(r, \lambda) = \frac{p^{\mathrm{EO}}(r, \lambda) \cdot \beta_{\mathrm{ATTN}}(r, \lambda)}{r^2} \text{ [photons]}. \tag{6}$$

An atmospheric lidar samples heights above r_1 [km]

$$r_i = r_1 + \Delta r \cdot (i - 1) \text{ [km]} \quad \forall\, i \in [1, m],\ m \in \mathbb{Z}. \tag{7}$$

A lidar typically averages the signals during a period of Δt[sec] that includes a number of laser pulses n_{pulse} [20]. When acquisition starts at time t_1, the time measurement samples are

$$t_j = t_1 + \Delta t \cdot (j - 1) \quad \forall\, j \in [1, n],\ n \in \mathbb{Z}. \tag{8}$$

Each measurement is done simultaneously at $w \in \mathbb{Z}$ wavelength channels. Hence, a discrete time-dependent representation of (1) at $\lambda_k \in (\lambda_1, \ldots, \lambda_w)$ is

$$\begin{aligned} p_{i,j,k} &\triangleq \frac{1}{n_{\mathrm{pulse}}} \sum_{t=t_{j-1}}^{t_j} \left[p^{\mathrm{BS}}(r_i, t, \lambda_k) + p^{\mathrm{BG}}(t, \lambda_k)\right] \\ &= p^{\mathrm{BS}}_{i,j,k} + p^{\mathrm{BG}}_{j,k} \text{ [photons]}. \end{aligned} \tag{9}$$

B. Optical Properties of Scattering Particles

Air molecules are a major source of the *Rayleigh scattering*, denoted as the *molecular signal* (mol). Both the molecular *number density* $\rho^{\mathrm{mol}}_{\#}(r)$ [1/km] and the *scattering cross section* $\sigma^{\mathrm{mol}}(r, \lambda)$[km^2] depend on the air pressure $\mathrm{P}_{\mathcal{A}}(r)$ [hPa], temperature $\mathrm{T}_{\mathcal{A}}(r)$ [K], and relative humidity $\mathrm{H}_{\mathcal{A}}(r)$ [%] at height r. The number density $\rho^{\mathrm{mol}}_{\#}(r)$ is derived from the U.S. Standard Atmosphere models [51], and $\sigma^{\mathrm{mol}}(r, \lambda)$ is derived using the modeled Rayleigh scattering cross section [52], [53]. Then, the *extinction coefficient* of atmospheric gases is $\alpha^{\mathrm{mol}}(r, \lambda) = \sigma^{\mathrm{mol}}(r, \lambda) \cdot \rho^{\mathrm{mol}}_{\#}(r)$ [1/km].

Aerosols often affect light by *Mie scattering*, denoted as *aerosol signal* (aer). Denote $\rho^{\mathrm{aer}}_{\#}(r)$ [1/km^3] as the number density and $\sigma^{\mathrm{aer}}(r, \lambda)$[km^2] as the effective extinction cross section. They set the aerosols extinction coefficient $\alpha^{\mathrm{aer}}(r, \lambda) = \sigma^{\mathrm{aer}}(r, \lambda) \cdot \rho^{\mathrm{aer}}_{\#}(r)$ [1/km]. The total extinction coefficient is

$$\alpha(r, \lambda) = \alpha^{\mathrm{mol}}(r, \lambda) + \alpha^{\mathrm{aer}}(r, \lambda) \left[\frac{1}{\mathrm{km}}\right]. \tag{10}$$

Particles have a unitless *single scattering albedo* $\varpi(r, \lambda)$ and a *scattering phase function*, whose value at backscatter angle π is $\mathcal{P}(\pi, \lambda)$ [1/sr]. The *lidar ratio* (LR) expresses the relation of extinction to the backscatter coefficients defined as [54], [55], [56]

$$\mathrm{LR(r, \lambda)} \triangleq \frac{\alpha(r, \lambda)}{\beta(\mathrm{r}, \lambda)} = \frac{1}{\varpi(r, \lambda) \cdot \mathcal{P}(\pi, \lambda)} \text{ [sr]}. \tag{11}$$

For Rayleigh scattering, $\mathrm{LR}^{\mathrm{mol}} \approx (8\pi/3)$ [sr] [57]. Let the LR of aerosols be $\mathrm{LR}^{\mathrm{aer}}$(r, λ)[sr]. Then, using (10) and (11), the total backscatter coefficient is

$$\begin{aligned} \beta(r, \lambda) &= \beta^{\mathrm{mol}}(r, \lambda) + \beta^{\mathrm{aer}}(r, \lambda) \\ &= \frac{\alpha^{\mathrm{mol}}(r, \lambda)}{\mathrm{LR}^{\mathrm{mol}}} + \frac{\alpha^{\mathrm{aer}}(r, \lambda)}{\mathrm{LR}^{\mathrm{aer}}(\lambda)} \left[\frac{1}{\mathrm{km} \cdot \mathrm{sr}}\right]. \end{aligned} \tag{12}$$

An additional wavelength-dependent property of an aerosol is the Å*ngström exponent* Å [37], [43], [58], [59], [60]. Let us denote a temporal spectral relation using Å

$$\mathring{\Lambda}_{\tilde{k}\to k} \triangleq \frac{\tau(r,\lambda_k)}{\tau(r,\lambda_{\tilde{k}})} = \begin{cases} \left(\frac{\lambda_k}{\lambda_{\tilde{k}}}\right)^{-\mathring{\Lambda}_{k,\tilde{k}}}, & \lambda_k < \lambda_{\tilde{k}} \\ \left(\frac{\lambda_{\tilde{k}}}{\lambda_k}\right)^{\mathring{\Lambda}_{\tilde{k},k}}, & \lambda_{\tilde{k}} < \lambda_k. \end{cases} \tag{13}$$

IV. Lidar Measurement Simulation Modeling

We present an array-based formalism for measurements. Let the temporal and the altitude bin measurements be $i \in [1,\ldots,m]$ and $j \in [1,\ldots,n]$, respectively. Let the wavelength channel bins be $k \in [1,\ldots,w]$. These create an STMW array $\mathbb{Z}_+^{m\times n\times w}$. Then, from (9), the expected measurements can be arranged in an STMW array

$$\begin{aligned} \mathbf{P} &\triangleq \mathbf{P}^{\mathrm{BS}} + \mathbf{P}^{\mathrm{BG}} = \left(p^{\mathrm{BS}}_{i,j,k}\right) + \left(p^{\mathrm{BG}}_{j,k}\right) \\ &\equiv \left(p_{i,j,k}\right) \in \mathbb{R}_+^{m\times n\times w}\ [\text{photons}]. \end{aligned} \tag{14}$$

The ALiDAn framework is built on top of the XARRAY platform, designed to handle multichannel spatiotemporal gridded data [61]. The array formalism in (14) is elaborated on in Section IV-A. To induce natural statistics of photoelectron signals, measurements are simulated using a Poissonian distribution [62]

$$\mathbf{P}_{\mathrm{M}} \sim \mathrm{Poiss}(\mathbf{P}) \in \mathbb{Z}_+^{m\times n\times w}\ [\text{photons}]. \tag{15}$$

We present simulations of the multiwavelength polarization Raman lidar (Polly$^{\mathrm{XT}}$) by TROPOS [20], [63]. The system operates continuously with $\Delta t = 30$ [s]. We simulated the far-field elastic backscatter channels $\lambda_k \in (355, 532, 1064)$ [nm]. The far-field channels measure up to an altitude of $\sim$48 [km], with $\delta t = 50$ [ns] and $\Delta r \approx 7.5$ [m]. The partial overlap zone is up to an altitude of $\sim$500 [m]. Here, we set a period of 24 h with time bins of $n = 2880$ and altitude bins of $m = 3000$, covering heights up to $\sim$22.5 [km].

A. Array Formalism

Each element of $\mathbf{P}$ in (14) satisfies (9). Similar to (14), we denote the sampled optical coefficients in arrays $\boldsymbol{\alpha} \in \mathbb{R}_+^{m\times n\times w}$ [1/km] and $\boldsymbol{\beta} \in \mathbb{R}_+^{m\times n\times w}$ [1/(km · sr)]. Let $\boldsymbol{d}_i^\top$ be the ith row of a lower left triangular matrix $\mathbf{D} \in \mathbb{R}_+^{m\times m}$. The elements of the matrix $\mathbf{D}$ are

$$\mathbf{D} \triangleq \begin{pmatrix} r_1 & & & \\ r_1 & \Delta r & & \\ \vdots & \vdots & \ddots & \\ r_1 & \Delta r & \ldots & \Delta r \end{pmatrix} \equiv \begin{pmatrix} — & \boldsymbol{d}_1^\top & — \\ — & \boldsymbol{d}_2^\top & — \\ & \vdots & \\ — & \boldsymbol{d}_m^\top & — \end{pmatrix} [\text{km}]. \tag{16}$$

Let $\boldsymbol{\alpha}_{j,k}$ be the jth column at the kth channel of array $\boldsymbol{\alpha}$. Then, the discrete form of (2) is

$$\begin{aligned} \tau_{i,j,k} &\approx \sum_{\imath=1}^{i} \alpha_{\imath,j,k} \cdot dz_\imath \\ &= r_1 \cdot \alpha_{1,j,k} + \Delta r \sum_{\substack{\imath=2 \\ \forall i>1}}^{i} \alpha_{\imath,j,k} \triangleq \boldsymbol{d}_i^\top \boldsymbol{\alpha}_{j,k}. \end{aligned} \tag{17}$$

Using (16) and (17), the sampled array of optical depths is

$$\boldsymbol{\tau} \triangleq \mathbf{D}\boldsymbol{\alpha} = \left(\tau_{i,j,k}\right) \in \mathbb{R}_+^{m\times n\times w}. \tag{18}$$

Denote by $\odot$ the elementwise *Hadamard product*. Use elementwise exponent $\mathbf{exp}(-2\boldsymbol{\tau})$ with (3) to define an array for the attenuated backscatter

$$\begin{aligned} \boldsymbol{\beta}_{\mathrm{ATTN}} \triangleq \boldsymbol{\beta} \odot \mathbf{exp}(-2\boldsymbol{\tau}) &= \left(\beta_{\mathrm{ATTN}i,j,k}\right) \\ &\in \mathbb{R}_+^{m\times n\times w} \left[\frac{1}{\text{km}\cdot\text{sr}}\right]. \end{aligned} \tag{19}$$

Temporal variations in the EO system factor are mainly caused by the LC variations. Except for the pulse duration, all of the EO components are prone to disturbances, distortions, and wear. Hence, a temporal form of (4) is $p^{\mathrm{LC}}(t,\lambda) = p_0(t,\lambda)\,\Delta r\,A(t)\,\eta(t)$[photons · km^3]. The samples of p^{LC} at channel k create a temporal vector $\boldsymbol{p}_k^{\mathrm{LC}} = (p_{1,k}^{\mathrm{LC}},\ldots,p_{n,k}^{\mathrm{LC}}) \in \mathbb{R}_+^{1\times n}$. Overlap changes are expected to be slower than those of the LC. Hence, the height-dependent overlap function remains constant within n time samples $\boldsymbol{O} \in \mathbb{R}_+^{m\times 1}$.

An array *broadcast* operation [61] shares on values of an array across additional dimensions. Using broadcasting, we share $\boldsymbol{p}_k^{\mathrm{LC}}$ across all heights into $\mathbf{P}_k^{\mathrm{LC}} \in \mathbb{R}_+^{m\times n}$. A sampled LC array is $\mathbf{P}^{\mathrm{LC}} \triangleq \{\mathbf{P}_k^{\mathrm{LC}}\}_{\forall k} \in \mathbb{R}_+^{m\times n\times w}$ [photons · km^3]. Similarly, broadcasting the vector $\boldsymbol{O}$ twice, across all wavelength and times, creates $\mathbf{O} \in \mathbb{R}_+^{m\times n\times w}$. Then, from (5), the EO factor is

$$\mathbf{P}^{\mathrm{EO}} = \mathbf{P}^{\mathrm{LC}} \odot \mathbf{O} \in \mathbb{R}_+^{m\times n\times w}\ [\text{photons}\cdot\text{km}^3]. \tag{20}$$

A time vector $\boldsymbol{p}_k^{\mathrm{BG}} = (p_{1,k}^{\mathrm{BG}},\ldots,p_{n,k}^{\mathrm{BG}}) \in \mathbb{R}_+^{1\times n}$ includes the samples of the BG signal of channel k. Broadcasting $\boldsymbol{p}_k^{\mathrm{BG}}$ through height in all channels constructs $\mathbf{P}^{\mathrm{BG}} \in \mathbb{R}_+^{m\times n\times w}$ [photons]. Let us denote a diagonal matrix of inverse square altitudes $\mathbf{R}^{-2} \triangleq \mathrm{diag}(r_1^{-2},\ldots,r_m^{-2}) \in \mathbb{R}_+^{m\times m}$ [km^{-2}]. Finally, the formation model of a simulated lidar signal presented in (14) is

$$\begin{aligned} \mathbf{P} \triangleq \mathbf{P}^{\mathrm{BS}} + \mathbf{P}^{\mathrm{BG}} &= \mathbf{R}^{-2}\mathbf{P}^{\mathrm{EO}} \odot \boldsymbol{\beta}_{\mathrm{ATTN}} + \mathbf{P}^{\mathrm{BG}} \\ &\in \mathbb{R}_+^{m\times n\times w}\ [\text{photons}]. \end{aligned} \tag{21}$$

V. Atmospheric State Generator

This section describes the generation process of atmospheric optical properties. The process consists of: 1) generation of the molecular optical coefficients $\boldsymbol{\alpha}^{\mathrm{mol}} \in \mathbb{R}_+^{m\times n\times w}$ [1/km] and $\boldsymbol{\beta}^{\mathrm{mol}} \in \mathbb{R}_+^{m\times n\times w}$ [1/(km · sr)] in Section V-A and 2) generation of the aerosol optical coefficients $\boldsymbol{\alpha}^{\mathrm{aer}} \in \mathbb{R}_+^{m\times n\times w}$ [1/km] and $\boldsymbol{\beta}^{\mathrm{aer}} \in \mathbb{R}_+^{m\times n\times w}$ [1/(km · sr)] in Section V-B. Then, (10) and (12) yield $\boldsymbol{\alpha} \in \mathbb{R}_+^{m\times n\times w}$ [1/km] and $\boldsymbol{\beta} \in \mathbb{R}_+^{m\times n\times w}$ [1/km · sr]. Equation (19) yields $\boldsymbol{\beta}_{\mathrm{ATTN}} \in \mathbb{R}_+^{m\times n\times w}$ [1/km · sr]. The atmospheric state generator is outlined in Fig. 2.

A. Molecular Optical Properties

For a given location (longitude, latitude) and time, one can generate an atmospheric state using meteorological measurement of $\mathbf{P}_{\mathcal{A}} \in \mathbb{R}_+^{\hat{m}\times\hat{n}}$, $\mathbf{T}_{\mathcal{A}} \in \mathbb{R}_+^{\hat{m}\times\hat{n}}$, and $\mathbf{H}_{\mathcal{A}} \in \mathbb{R}_+^{\hat{m}\times\hat{n}}$ [64], [65], [66], [67]. Here, a state is defined on $\mathbb{Z}_+^{\hat{m}\times\hat{n}}$, a sparse

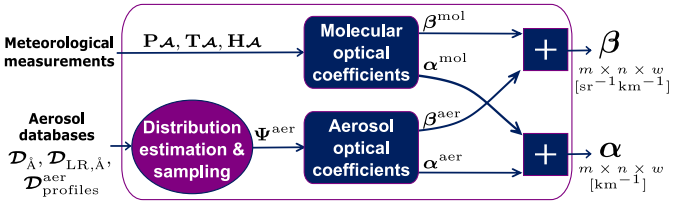


Fig. 2. Diagram of the atmospheric state generator.

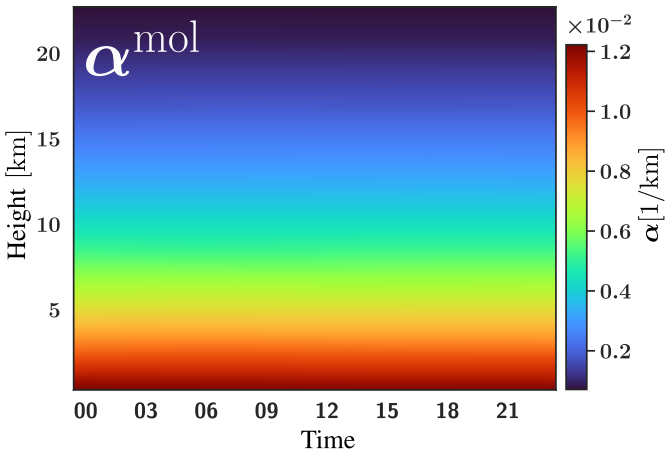


Fig. 3. Daily generated molecular extinction coefficient $\boldsymbol{\alpha}^{\text{mol}}$[1/km] at 532 [nm] for September 1, 2017, Haifa.

subarray of $\mathbb{Z}_{+}^{m\times n}$. Assume that the spatiotemporal distribution of air molecules does not change significantly between adjacent measurements. Hence, a spatiotemporal interpolation is performed and broadcast to an STMW array. According to Section III-B, we calculate elementwise $\boldsymbol{\alpha}^{\text{mol}} \in \mathbb{R}_{+}^{m\times n\times w}$ [1/km] using [68]. Then, with LR^{mol} and (11), we get $\boldsymbol{\beta}^{\text{mol}} \in \mathbb{R}_{+}^{m\times n\times w}$ [1/(km · sr)].

To calculate daily optical coefficients of air gases, we use meteorological measurement profiles by GDAS, available globally every 3 h on a 360 × 181 (1° × 1°) latitude–longitude grid [64]. Fig. 3 shows the generated extinction coefficient by air molecules at $\lambda = 532$ [nm] $\boldsymbol{\alpha}^{\text{mol}}$ for September 1, 2017, in Haifa.

B. Aerosol Optical Properties

To generate aerosol optical coefficients, our model addresses the following four questions: 1) where are the aerosols found? 2) is there a typical 2D texture associated with their spatiotemporal densities? 3) what are the typical optical coefficients of particles per wavelength? and 4) how can one simulate measurements typical to a geographic site or a season?

In the troposphere, generally, aerosol concentration decreases with altitude. Beyond a *reference height* r_{ref} [km] [69], [70], [71], [72], the aerosol optical coefficients become negligible in comparison to air. Aerosols tend to have a smooth spatiotemporal distribution. Therefore, we assume that their density can be approximated or represented by smooth base functions, specifically spatiotemporal Gaussians. Here, we further leverage the aerosol conversion features noted in LIVAS [28] to generate typical random spatiotemporal variations of aerosols.

We denote $\boldsymbol{\Psi}^{\text{aer}}$ as the aerosol optical parameter space used to statistically generate an STMW aerosol optical density. The parameter space $\boldsymbol{\Psi}^{\text{aer}}$ consist of the following: (I) A reference height r_{ref}[km]. (II) A typical maximum extinction coefficient $\alpha_{\tilde{k}}^{\max} \in \mathbb{R}_{+}[1/\text{km}]$ at $\lambda_{\tilde{k}}$. (III) A set of spectral conversion matrices $\{\boldsymbol{\Gamma}_k \in \mathbb{R}_{+}^{n\times n}\}_{\forall k}$. (IV) A set of LR conversion matrices $\{\boldsymbol{\Gamma}_k \in \mathbb{R}_{+}^{n\times n}\}_{\forall k}$. Here, each matrix $\mathring{\boldsymbol{\Lambda}}_{\tilde{k}\to k}$ and $\boldsymbol{\Gamma}_k$ is diagonal, containing respectively the temporal values of spectral relations $(\mathring{\Lambda}_{1,\tilde{k}\to k}, \cdots, \mathring{\Lambda}_{n,\tilde{k}\to k})$ and LR relations $(\text{LR}^{\text{aer}}_{1,k}, \cdots, \text{LR}^{\text{aer}}_{n,k})$.

To simulate various atmospheric states, we initially randomly generate $\boldsymbol{\Psi}^{\text{aer}}$, as illustrated in 2. Section V-B1 describes how to derive these parameters. Then $\boldsymbol{\Psi}^{\text{aer}}$ yields the optical coefficients as elaborated in Section V-B2. Simulating an atmosphere state with a mixture of aerosols is discussed in Section V-B3. Section V-B4 shows a few diurnal simulations of aerosol optical coefficients.

1) Aerosol Statistical Model: Each parameter of $\boldsymbol{\Psi}^{\text{aer}}$ is sampled from an empirical continuous probability density function (pdf), which is related to an aerosol optical property at a given location and time of year. Let $\boldsymbol{f}(\mathcal{D})$ be a *truncated multivariate Gaussian mixture* (GM) pdf [73] fit to dataset $\mathcal{D}$.

For data given prior to statistical assessment as raw measurements, we used *kernel density estimation* (KDE) to calculate $\boldsymbol{f}(\mathcal{D})$ [73]. If data were provided after statistical analysis, we used the mean and uncertainty of each data point to generate a single Gaussian in a weighted mixture to calculate $\boldsymbol{f}(\mathcal{D})$.

Denote by $\mathcal{D}^{\text{aer}}_{\text{profiles}}$ a dataset of previously retrieved aerosol optical coefficients profiles [2]. From $\mathcal{D}^{\text{aer}}_{\text{profiles}}$, we derive $\hat{f}(\alpha_{\tilde{k}}^{\max}, r_{\text{ref}})$. This pdf empirically represents the typical maximum value of the extinction coefficient at a reference wavelength $\lambda_{\tilde{k}}$, e.g., $\alpha_{352}^{\max}$, as a function of the aerosols' reference height.

Denote by $\mathcal{D}_{\text{Å}}$ a dataset of Ångström measurements at two pairs of wavelengths, e.g., $\text{pair}_1 = \{355, 532\}$ [nm] and $\text{pair}_2 = \{532, 1064\}$ [nm]. These are usually derived from a sunphotometer as that of AERONET [48]. From $\mathcal{D}_{\text{Å}}$, we calculate $\hat{f}(\text{Å}_{\text{pair}_1}, \text{Å}_{\text{pair}_2})$ to sample the spectral dependency for three different wavelengths. The LR is sampled from $\hat{f}(\text{LR}_k, \text{Å})$, given the presampled Å values. This pdf can be generated by using aerosol data derived from previous field campaigns [1], [56], [74], [75], denoted by $\mathcal{D}_{\text{LR,Å}}$.

The parameters are mainly sampled from monthly statistics, corresponding to the location and time of the synthesized measurements. The values of r_{ref} and $\alpha_{\tilde{k}}^{\max}$ are sampled once in a period of n temporal measurements. With these values, we produce varying spatiotemporal optical coefficients, as discussed in Section V-B2. To simulate dynamic changes in aerosol composition within such a period, we sample values of Å and LR at $n_{\text{s}} < n$ times e.g., three to five times.

Let $\mathcal{I}_n(\{y_j, t_j\}_{n_{\text{s}}})$ be a spline interpolation from a set of n_{s} values y_j at times t_j to n times [76], e.g., the Bèzier interpolation. Using such an operator, we produce continuous changes of Å and LR based on their corresponding sampled values. The sampling process is summarized in Algorithm 1.

From the PollyNet Processing Chain, the automated analysis software of Polly$^{\text{XT}}$ [2], [47], we derived $\mathcal{D}^{\text{aer}}_{\text{profiles}}$. The

Algorithm 1 Modeling and Augmenting Aerosol Optical Properties

Input: $\mathcal{D}^{\text{aer}}_{\text{profiles}}, \mathcal{D}^{\text{aer}}_{\text{Å}}, \mathcal{D}^{\text{aer}}_{\text{LR,Å}}$
Output: $\boldsymbol{\Psi}^{\text{aer}}$

// Calculate empirical aerosol PDFs
1 $\hat{f}\left(\alpha^{\max}_{\tilde{k}}, r^{\text{ref}}\right) \leftarrow \boldsymbol{f}(\mathcal{D}^{\text{aer}}_{\text{profiles}})$
2 $\hat{f}\left(\text{Å}_{\text{pair1}}, \text{Å}_{\text{pair2}}\right) \leftarrow \boldsymbol{f}(\mathcal{D}^{\text{aer}}_{\text{Å}})$
3 $\hat{f}\left(\text{LR}_k, \text{Å}\right) \leftarrow \boldsymbol{f}(\mathcal{D}_{\text{LR,Å}})$
// Sample parameters
4 $\left(\alpha^{\max}_{\tilde{k}}, r_{\text{ref}}\right) \sim \hat{f}\left(\alpha^{\max}_{\tilde{k}}, r_{\text{ref}}\right)$
5 **for** $s \in [1, \cdots, n_s]$ **do**
6 $\quad \left(\text{Å}_{\text{pair1}}, \text{Å}_{\text{pair2}}\right)_s \sim \hat{f}\left(\text{Å}_{\text{pair1}}, \text{Å}_{\text{pair2}}\right)$
7 $\quad \text{LR}_{k,s} \sim \hat{f}\left(\text{LR}_k | \text{Å} = \text{Å}_s\right)$, s.t. Å_s is one of Å_{pair1} or Å_{pair2} samples
8 $\quad$ Set t_s, s.t. t_s is a unique time in $[t_1, \ldots, t_n]$.
// Produce continuous changes of Å and LR
9 $\mathring{\boldsymbol{\Lambda}}_{\tilde{k}\to k} \overset{\forall k\neq k}{\leftarrow} \text{diag}\left(\mathcal{I}_n\left(\{\mathring{\Lambda}_{j,\tilde{k}\to k}, t_j\}_{n_s}\right)\right)$
10 $\boldsymbol{\Gamma}_k \overset{\forall k}{\leftarrow} \text{diag}\left(\mathcal{I}_n\left(\{\text{LR}_{j,k}, t_j\}_{n_s}\right)\right)$
11 **return** $\boldsymbol{\Psi}^{\text{aer}} \leftarrow \left\{r_{\text{ref}}, \alpha^{\max}_{\tilde{k}}, \left\{\mathring{\boldsymbol{\Lambda}}_{\tilde{k}\to k}\right\}_{\forall k\neq\tilde{k}}, \left\{\boldsymbol{\Gamma}_k\right\}_{\forall k}\right\}$

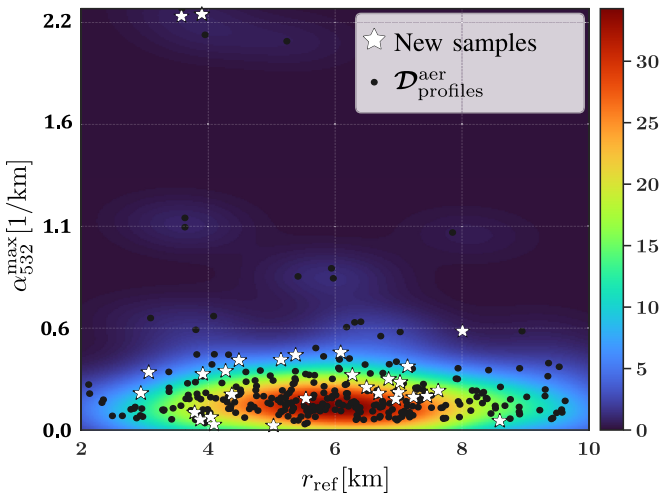


Fig. 4. 2D heatmap of the empirical pdf $\hat{f}(\alpha^{\max}_{532}, r_{\text{ref}})$ for September 2017, Haifa. Values for estimating the pdf, shown as black dots, are extracted from the aerosol profile dataset $\mathcal{D}^{\text{aer}}_{\text{profiles}}$, provided by the PollyNet Processing Chain. White stars indicate samples for generating new aerosol densities. Refer to Lines 1 and 4 in Algorithm 1 for the estimation and the re-sampling processes, respectively.

empirical pdf $\hat{f}(\alpha^{\max}_{\tilde{k}}, r_{\text{ref}})$ at $\lambda_{\tilde{k}} = 532$ [nm] for September 2017 is presented in Fig. 4. During the campaign, an AERONET sunphotometer was operated near the lidar. Using AERONET cloudless inversion products (level > 1.5), we created $\mathcal{D}_{\text{Å}}$. We first calculated optical depths, and using (13), the values of $\text{Å}_{355,532}$ and $\text{Å}_{532,1064}$ were calculated for each AERONET measurement in a month. The optical depths of wavelengths not measured by the sunphotometer were interpolated from neighboring measured wavelengths. Empirical distributions $\hat{f}(\text{Å}_{355,532}, \text{Å}_{532,1064})$ for April and September 2017 are presented in Fig. 5.

2) Generating the Optical Coefficients of Aerosols: We suggest representing the aerosol spatiotemporal distribution using a 2D GM in the spatiotemporal domain. In Appendix A, we compare this assumption with data derived from measurements. Accordingly, this section presents the generation process of the predefined sampled parameters $\boldsymbol{\Psi}^{\text{aer}}$. We generate random spatiotemporal Gaussians up to a given reference height r_{ref} [km]. A randomly weighted summation of these Gaussians creates $\boldsymbol{\rho} \in \mathbb{R}^{m\times n}_{+}$ such that $\|\boldsymbol{\rho}\|_1 = 1$. We now denote a few definitions in (22)–(24). These definitions are used in the process of creating spatiotemporal optical coefficients.

Let us denote the global min–max normalization scaling operator of an arbitrary array $\mathbf{X}$ to be

$$\boldsymbol{\mathcal{S}}_{\text{GN}}(\mathbf{X}) = \frac{\mathbf{X} - \min(\mathbf{X})}{x^{\max}} \tag{22}$$

where the global normalization factor is $x^{\max} = \max(\mathbf{X}) - \min(\mathbf{X}) \in \mathbb{R}_{+}$. A temporal min–max normalization scaling operator of a spatiotemporal array $\mathbf{X} \in \mathbb{R}^{m\times n}_{+}$ is

$$\boldsymbol{\mathcal{S}}_{\text{TN}}(\mathbf{X}) = \begin{pmatrix} | & & | \\ \boldsymbol{\mathcal{S}}_{\text{GN}}(\boldsymbol{x}_1) & \cdots & \boldsymbol{\mathcal{S}}_{\text{GN}}(\boldsymbol{x}_n) \\ | & & | \end{pmatrix} \in \mathbb{R}^{m\times n}_{+} \tag{23}$$

such that $\boldsymbol{x}_j$ is a column of $\mathbf{X}$ at time t_j. The normalization factors of columns $\boldsymbol{x}_j \, \forall j \in [1, \ldots, n]$ are arranged in a diagonal matrix

$$\mathbf{X}^{\max} = \begin{pmatrix} x^{\max}_1 & & \\ & \ddots & \\ & & x^{\max}_n \end{pmatrix} \in \mathbb{R}^{n\times n}_{+}. \tag{24}$$

Using (22), we set a unitless global normalized density $\boldsymbol{\rho}_{\text{GN}} = \boldsymbol{\mathcal{S}}_{\text{GN}}(\boldsymbol{\rho}) \in \mathbb{R}^{m\times n}_{+}$. For given $\alpha^{\max}_{\tilde{k}}$ and $\boldsymbol{\rho}_{\text{GN}}$, one can generate a new spatiotemporal array of extinction coefficients at $\lambda_{\tilde{k}}$

$$\boldsymbol{\alpha}^{\text{aer}}_{\tilde{k}} = \boldsymbol{\rho}_{\text{GN}} \boldsymbol{\alpha}^{\max}_{\tilde{k}} \in \mathbb{R}^{m\times n}_{+} \left[\frac{1}{\text{km}}\right]. \tag{25}$$

Using (23), we get $\boldsymbol{\rho}_{\text{TN}} = \boldsymbol{\mathcal{S}}_{\text{TN}}(\boldsymbol{\alpha}^{\text{aer}}_{\tilde{k}}) \in \mathbb{R}^{m\times n}_{+}$, the density normalized per time t_j. Using (24), we set $\boldsymbol{\alpha}^{\max}_{\tilde{k}} \in \mathbb{R}^{n\times n}$ as the normalization factors of extinction coefficients at $\lambda_{\tilde{k}}$. Then, the extinction coefficient at $\lambda_{\tilde{k}}$ can be rewritten as

$$\boldsymbol{\alpha}^{\text{aer}}_{\tilde{k}} = \boldsymbol{\rho}_{\text{TN}} \boldsymbol{\alpha}^{\max}_{\tilde{k}} \in \mathbb{R}^{m\times n}_{+} \left[\frac{1}{\text{km}}\right]. \tag{26}$$

For September 1, 2017, the generated unitless $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_{\text{TN}}$ are presented in Fig. 6(a) and (b), respectively.

Let the aerosol particle type be vertically uniform at each time t_j. Then, from (18) and (26), the optical depths at $\lambda_{\tilde{k}}$ are

$$\boldsymbol{\tau}^{\text{aer}}_{\tilde{k}} = \mathbf{D}\boldsymbol{\rho}_{\text{TN}} \boldsymbol{\alpha}^{\max}_{\tilde{k}} \in \mathbb{R}^{m\times n}_{+}. \tag{27}$$

According to (13), the spectral relations set the temporal optical depths at λ_k

$$\boldsymbol{\tau}^{\text{aer}}_{k} = \boldsymbol{\tau}^{\text{aer}}_{\tilde{k}} \mathring{\boldsymbol{\Lambda}}_{\tilde{k}\to k} \in \mathbb{R}^{m\times n}_{+} \quad \forall \lambda_k \neq \lambda_{\tilde{k}}. \tag{28}$$

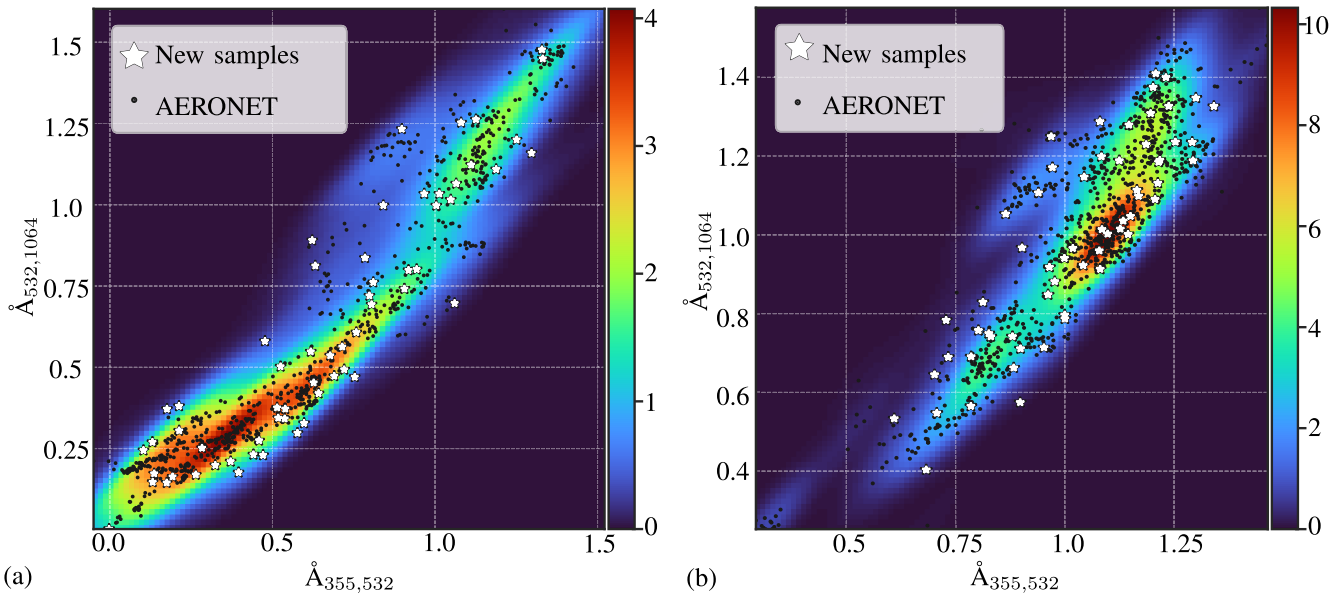


Fig. 5. Monthly empirical pdfs $\hat{f}(\mathring{A}_{355,532}, \mathring{A}_{532,1064})$ calculated based on AERONET inversion products and their corresponding sample values for (a) April and (b) September 2017, Haifa. Black dots indicate the values used for the pdf estimation, while the white stars indicate new samples employed in the aerosol generation process. Refer to Lines 2 and 6 in Algorithm 1 for the estimation and the resampling processes, respectively.

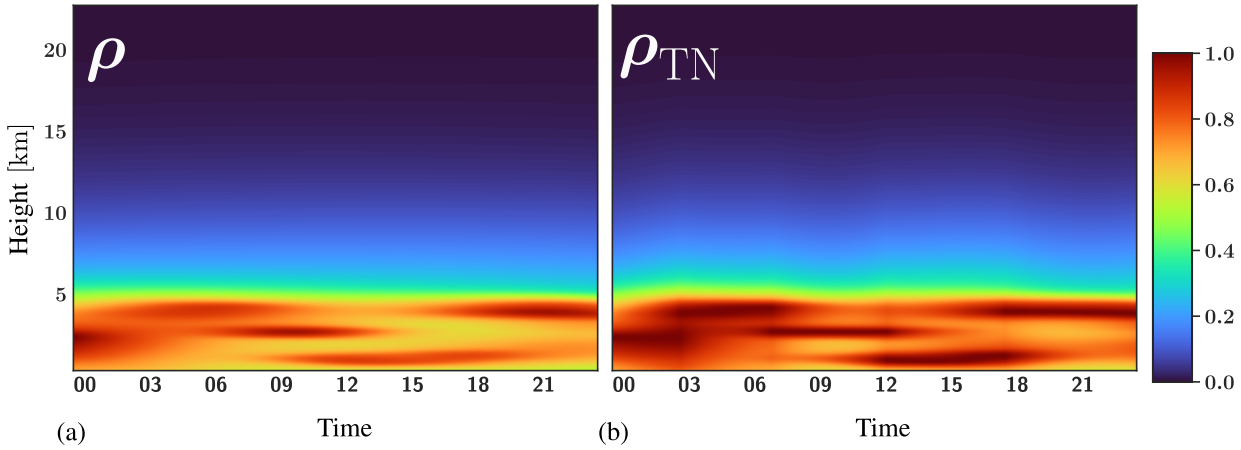


Fig. 6. Simulated densities for September 1, 2017. (a) Unitless density $\boldsymbol{\rho}$ generated from random spatiotemporal Gaussians. (b) Unitless temporally normalized density $\boldsymbol{\rho}_{\mathrm{TN}}$.

Then, by plugging (27) into (28), and multiplying from the left by $\mathbf{D}^{-1}$, the spatiotemporal array of the extinction coefficient at λ_k is

$$\boldsymbol{\alpha}_k^{\mathrm{aer}} \approx \boldsymbol{\rho}_{\mathrm{TN}}\, \boldsymbol{\alpha}_{\tilde{k}}^{\max}\, \mathring{\boldsymbol{\Lambda}}_{\tilde{k}\to k} \in \mathbb{R}_+^{m\times n} \left[\frac{1}{\mathrm{km}}\right] \quad \forall\, \lambda_{\mathrm{k}} \neq \lambda_{\tilde{\mathrm{k}}}. \tag{29}$$

From (12), the backscatter coefficient at any λ_k is

$$\boldsymbol{\beta}_k^{\mathrm{aer}} = \boldsymbol{\alpha}_k^{\mathrm{aer}} \boldsymbol{\Gamma}_k \in \mathbb{R}_+^{m\times n} \left[\frac{1}{\mathrm{km}\cdot\mathrm{sr}}\right]. \tag{30}$$

3) Aerosol Types: Consider a scene having several aerosol types at time t_j e.g., marine minerals, smoke, dust, and soot [23], [25], [76]. Here, we suggest two methods for simulating such a scene.

The first approach considers an atmosphere with an effective mixed aerosol type. It yields a mixed distribution of $\hat{f}(\mathrm{LR}_k, \mathring{\mathrm{A}})$, affecting $\boldsymbol{\Psi}^{\mathrm{aer}}$. Denote $\boldsymbol{\chi} = (\chi_0, \ldots, \chi_v)$ to be a vector of weights of v aerosol types s.t. : $\|\boldsymbol{\chi}\| = 1$. For example, a set of weights may indicate the presence of different aerosol types typical of a specific location [1] and/or a season [79], [80]. Then,

$$\hat{f}(\mathrm{LR}_k, \mathring{\mathrm{A}}) = \sum_{l=1}^{v} \chi_l \cdot \hat{f}_l(\mathrm{LR}_k, \mathring{\mathrm{A}}). \tag{31}$$

The second approach addresses an atmosphere state containing several aerosol layers. Here, optical coefficients are generated per each aerosol type, as elaborated on in Section V-B2. Then, one can set $\chi_{i,j,l}$ to represent the percentage share of aerosol type l at height r_i and time t_j. This creates a weighting map $\boldsymbol{\chi} \in \mathbb{R}_+^{m\times n}$ for each aerosol type l. Fig. 7(b) illustrates an example of such a map. Then, the optical coefficients of a

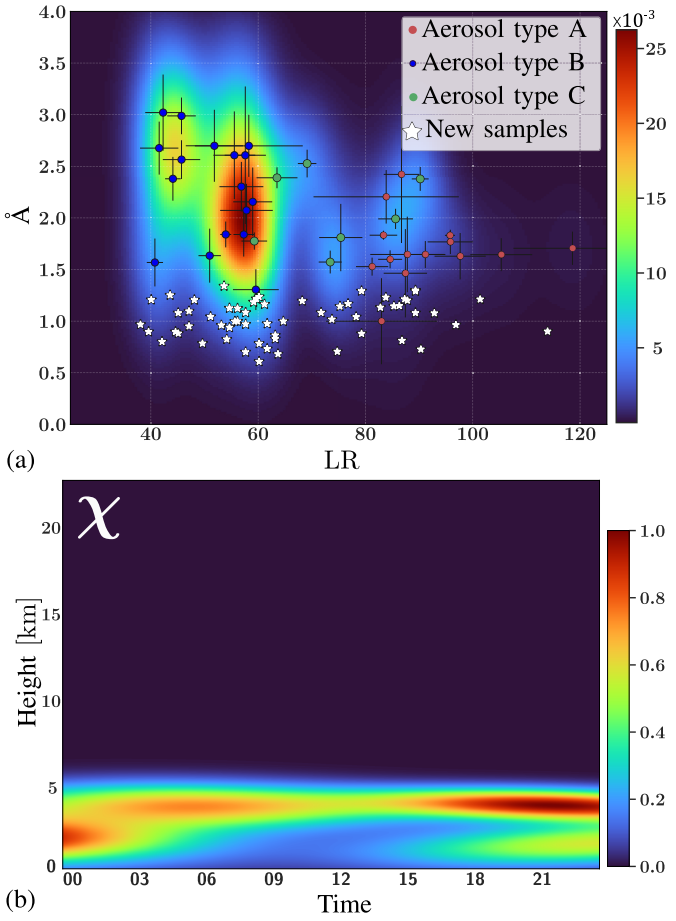


Fig. 7. Two approaches for generating various aerosol types. (a) Weighted joint pdf $\hat{f}(\mathrm{LR}, \mathring{\mathrm{A}})$ based on [78], demonstrating a single effective aerosol type at any time. Here, $\boldsymbol{\chi} = (0.05, 0.75, 0.2)$ for types A (biomass burning), B (urban/industrial), and C (a mixture of desert dust with biomass burning), respectively. White star markers illustrate new samples from $\hat{f}(\mathrm{LR}|\mathring{\mathrm{A}} = \mathring{\mathrm{A}}_s)$, where $\mathring{\mathrm{A}}_s$ are samples from $\hat{f}(\mathring{\mathrm{A}}_{355,532}, \mathring{\mathrm{A}}_{532,1064})$ for September 1, 2017. Refer to Lines 3 and 7 in Algorithm 1 for the estimation and the resampling processes, respectively. (b) Aerosol type weighting map $\boldsymbol{\chi} \in \mathbb{R}_+^{m\times n}$, demonstrating a mixture of aerosol layers.

mixture of aerosols at wavelength λ_k is a simple 2D blending

$$\boldsymbol{\alpha}_k^{\mathrm{aer}} = \sum_{l=1}^{v} \boldsymbol{\alpha}_{k,l}^{\mathrm{aer}} \odot \boldsymbol{\chi}_l \in \mathbb{R}_+^{\mathrm{m}\times\mathrm{n}} \left[\frac{1}{\mathrm{km}}\right] \tag{32a}$$

$$\boldsymbol{\beta}_k^{\mathrm{aer}} = \sum_{l=1}^{v} \boldsymbol{\beta}_{k,l}^{\mathrm{aer}} \odot \boldsymbol{\chi}_l \in \mathbb{R}_+^{\mathrm{m}\times\mathrm{n}} \left[\frac{1}{\mathrm{km}\cdot\mathrm{sr}}\right]. \tag{32b}$$

4) Examples of Diurnal Simulated Aerosol Optical Coefficients: Here, we demonstrate diurnal simulations of aerosol optical coefficients for April 4, 2017, May 16, 2017, and September 1, 2017. We derived $\hat{f}(\mathrm{LR}_k, \mathring{\mathrm{A}})$ from [78]. With the first approach proposed in Section V-B3, we calculate a mixed $\hat{f}(\mathrm{LR}, \mathring{\mathrm{A}})$. As such, the weighting values for aerosols of types A (biomass burning), B (urban/ industrial), and C (a mixture of desert dust with biomass burning) are $\boldsymbol{\chi} = (0.05, 0.75, 0.2)$, respectively. Such pdf is presented in Fig. 7(a).

For each day, we sample the Ångström values from $\hat{f}(\mathring{\mathrm{A}}_{355,532}, \mathring{\mathrm{A}}_{532,1064})$ of the corresponding months. Based on sampled values of the Ångström exponent, we sample $\hat{f}(\mathrm{LR}|\mathring{\mathrm{A}})$. The respective samples for September 2017 are marked with white stars in Fig. 7(a). The original pdf from [78] is given at $\lambda = 355$ [nm]. Here, we applied the LR values to be similar at all wavelengths. Diurnal Ångström and LR values are shown in Fig. 8.

Then, with $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_{\mathrm{TN}}$, as earlier presented in Fig. 6, we calculated the aerosol optical coefficients in other wavelengths according to Section V-B2. Fig. 9 presents accordingly the simulated diurnal aerosol extinction coefficients $\boldsymbol{\alpha}^{\mathrm{aer}}[1/\mathrm{km}]$ for these days.

VI. Electro-Optical Factor Generator

A. LC Generation

The LC factor is affected by variations in the EO setup, including laser power drift, optical transmittance changes at the detector due to environmental exposure disturbances, and so on [50]. On a scale of days, observations show a monotonic decline in LC values. The expected descending trend resembles an exponential decay at rate t_{decay}. Maintenance in every two to three months resets the LC to a higher range of values. We use these trends to generate simulated LC factors varying over time.

Denote a *prediction band* $\Delta_{\mathrm{LC}}(t, \lambda)$ [81] of LC values at any time t. For any time between maintenance interrupts, fluctuations are introduced to the source power. We achieve this by setting $\Delta_{\mathrm{LC}}(t, \lambda)$, within which randomized fluctuations deviate from the average parametric curve. After maintenance at time t_{LC_0}, the value of $\Delta_{\mathrm{LC}}(t, \lambda)$ is narrowest, and then, it expands over time. As an example, the uncertainty ranges from 5% to 20% by the end of the period. Hence, the generation model for the time-varying LC factor is

$$p^{\mathrm{LC}}(t, \lambda) = \mathrm{LC}_0(\lambda) \cdot \exp\left[-\frac{t - t_{\mathrm{LC}_0}}{t_{\mathrm{decay}}}\right] + \mathcal{N}\Big(0, \Delta_{\mathrm{LC}}(t, \lambda)\Big)\,[\mathrm{photons}\cdot\mathrm{km}^3]. \tag{33}$$

The parameters in (33) are set based on typical past LC retrievals for each λ. Except for operational maintenance, the power is expected to be smooth and continuous over time; thus, samples are temporally interpolated into $\boldsymbol{p}_k^{\mathrm{LC}} = (p_{1,k}^{\mathrm{LC}}, \ldots, p_{n,k}^{\mathrm{LC}}) \in \mathbb{R}_+^{1\times n}\,[\mathrm{photons}\cdot\mathrm{km}^3]\,\forall k$. We generate LC values for an extended period, e.g., two months. For each arbitrary period of n time bins within this period, we use the corresponding temporal section of LC values for measurement generation, as illustrated in the diagram in Fig. 10.

Fig. 11(a) shows the retrieved LC values from September to October 2017 at Haifa. The newly simulated values of such a period are presented in Fig. 11(b). Here, we used typical values such as $\mathbf{LC} = (15, 45, 35)\,[10^3\mathrm{photons}\cdot\mathrm{km}^3]$. The value of t_{LC_0} was randomly selected within a period of $t_{\mathrm{decay}} = 70$ days.

B. Overlap Generation

The optical setup directly affects the overlap function, e.g., when the laser beam is slanted. Usually, changes in the overlap function are caused by maintenance interrupts or temperature fluctuations in the lidar's cabin. Various overlap functions can be realistically simulated by adjusting a modeled or estimated

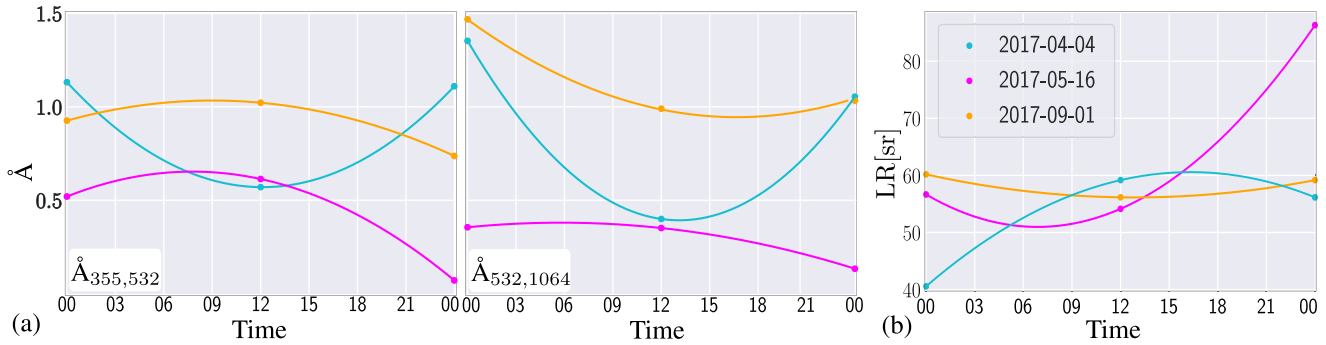


Fig. 8. Generated diurnal LR and Å for April 4, May 16, and September 1, 2017. (a) (Left) $Å_{355,532}$ and (Middle) $Å_{532,1064}$. (b) LR values derived from samples of the empirical distribution $\hat{f}(\text{LR}|Å)$.

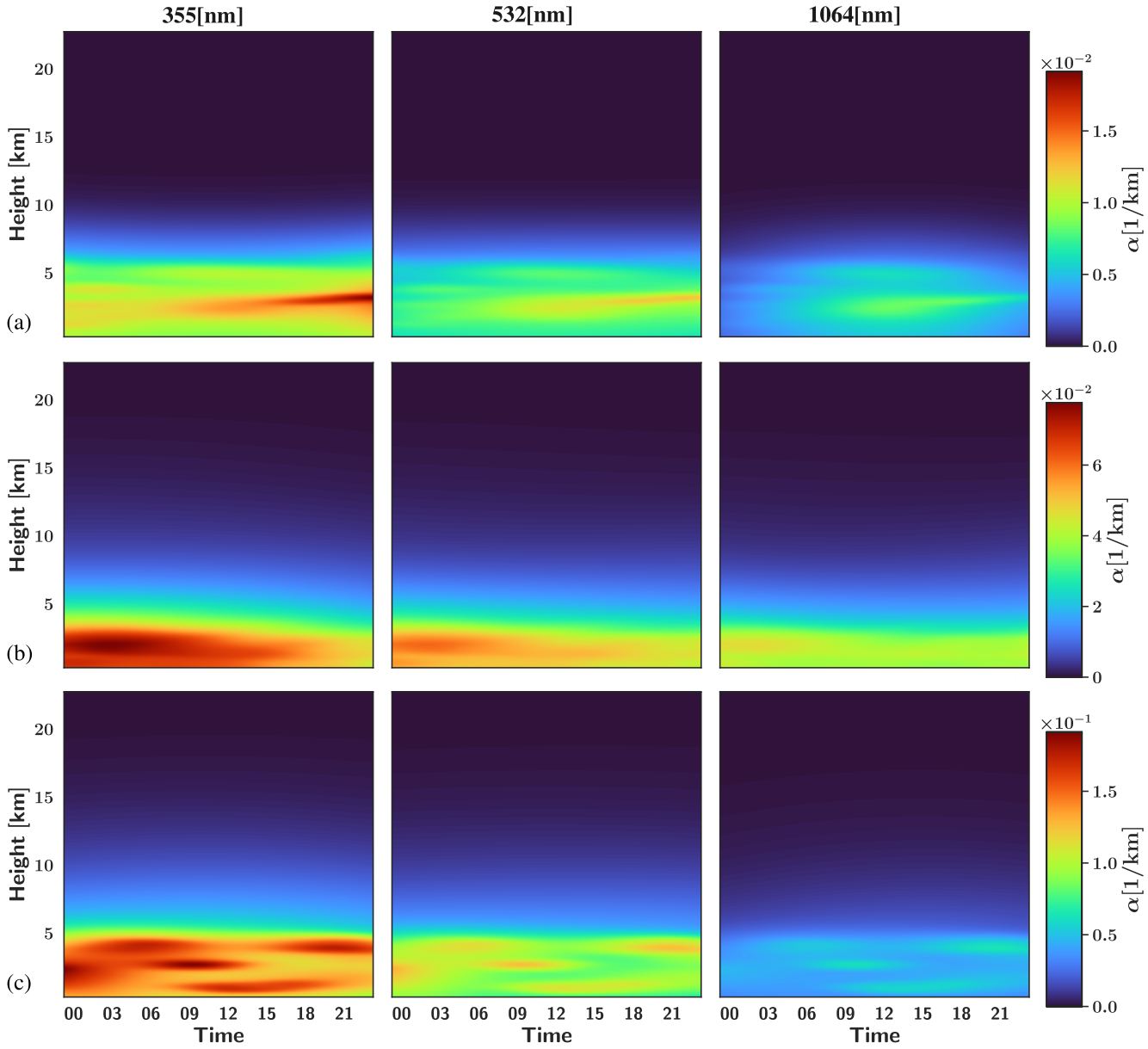


Fig. 9. Generated diurnal aerosol extinction coefficients $\boldsymbol{\alpha}^{\text{aer}}[1/\text{km}]$ for (a) April 4, (b) May 16, and (c) September 1, 2017.

overlap function [82], [83]. We generate overlap functions based on a typical set of past overlap retrievals. Denote r_{MO}, d, g, and s as parameters in $\mathbb{R}_+$. The overlap function resembles an "S" curve; therefore, we model it as a generalized logistic function

$$O(r) = \frac{1}{(1 + d\exp[-g(r - r_{\text{MO}})])^s}. \tag{34}$$

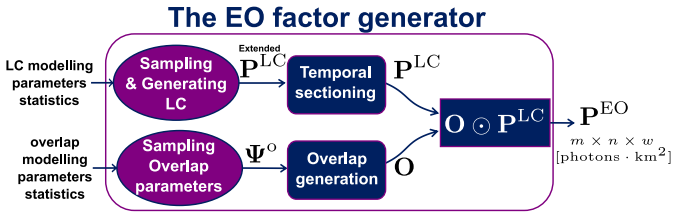


Fig. 10. Diagram of the lidar EO system factor generator.

Previously retrieved overlap functions are fit to (34). From these, we derive normalized distributions for the parameters above. To simulate new overlap functions, the parameters are sampled from their respective distributions. The overlap sampled parameters are denoted by $\mathbf{\Psi}^{\mathrm{o}}$ in Fig. 10. The overlap function tends to be more stable than other system factors, so it can be generated infrequently, e.g., once per several months. For the simulations presented in this work, we assumed a similar overlap function across all wavelengths. Fig. 11(c) shows several examples of newly generated overlap functions.

VII. Background Signal Generator

The Sun's position in the sky affects p^{BG}, with a diurnal symmetry. A lidar samples the temporal average of p^{BG} before every laser trigger event. This enables creating a statistical model. Denote a_0, a, $t_{\mathrm{peak}} \geq 0$, and $u > 0$ as Gaussian curve parameters. Then, the expected daily BG signal model per wavelength λ is

$$\left\langle p^{\mathrm{BG}}(t,\lambda)\right\rangle = a_0(\lambda) + a(\lambda)\exp\left[-\frac{[t - t_{\mathrm{peak}}(\lambda)]^2}{2u^2(\lambda)}\right] \text{[photons]}. \quad (35)$$

Now, we elaborate on how each of the parameters is tailored to a particular day of the year and geographic location. Initially, an actual lidar measurement from a clear day, denoted as (orig), was fit to (35).

Let θ be the solar elevation at any time [84], [85]. Let t_{noon} be the noon of *true solar time* [86]. At t_{noon}, the Sun is at its highest angle θ_{noon}. However, the BG peaks on $t_{\mathrm{peak}}(\lambda)$, which varies around t_{noon}, depending on the wavelength. For example, the signal at $\lambda = 1064$ [nm] peaks slightly later than the shorter wavelengths. We set $t_{\mathrm{peak}}(\lambda)$ of an arbitrary day in proportion t_{noon} using the ratio $t^{\mathrm{orig}}_{\mathrm{peak}}(\lambda)/t^{\mathrm{orig}}_{\mathrm{noon}}$.

We use a clear sky model of global downwelling solar irradiance from [87] with air mass turbidity conditions. We utilize its trend by calculating a fit to a function with parameters a_θ, b_θ, c_θ, and c_θ expressed by $\mathcal{E}(\theta) = a_\theta \cos(b_\theta\theta + c_\theta) + d_\theta$. Then, values of $a_0(\lambda)$ and $a(\lambda)$ for an arbitrary day are set to be proportional to $\mathcal{E}(\theta_{\mathrm{noon}})/\mathcal{E}(\theta^{\mathrm{orig}}_{\mathrm{noon}})$. This is illustrated in Fig. 12(a).

To set $u(\lambda)$, we use the twilight times. Let $\Delta t_{\mathrm{daylight}}$[sec] be the time difference between dusk and dawn. Denote the number of photons per δt reaching the detector due to sunlight at twilight as $p_{\mathrm{twilight}}(\lambda) = p^{\mathrm{BG}}(t_{\mathrm{twilight}}, \lambda)$[photons]. During twilight, we assume similar clear sky lighting conditions, i.e., $p_{\mathrm{twilight}} \equiv p^{\mathrm{orig}}_{\mathrm{twilight}}(\lambda) > a_0(\lambda)$, for any arbitrary day at a given location. From the symmetry of a Gaussian curve, we set

$$u(\lambda) = \frac{\Delta t_{\mathrm{daylight}}}{2\sqrt{2\ln\left(\frac{1}{q_{\mathrm{twilight}}(\lambda)}\right)}} \quad (36)$$

where

$$q_{\mathrm{twilight}}(\lambda) = \frac{p_{\mathrm{twilight}}(\lambda) - a_0(\lambda)}{a(\lambda)}, \quad a(\lambda) > 0. \quad (37)$$

In addition to θ, the BG signal is influenced by environmental conditions [87]. We, thus, introduce temporal fluctuations to the expected signal within a diurnal prediction band $\Delta_{\mathrm{BG}}(t,\lambda)$. For example, one can set the prediction band to vary with θ. Then, the sampled BG signal is

$$p^{\mathrm{BG}}_{j,k} = \left\langle p^{\mathrm{BG}}_{j,k}\right\rangle + \mathcal{N}(0, \Delta_{\mathrm{BG}\,j,k}) \text{ [photons]}. \quad (38)$$

A BG signal was created by taking an original BG measurement from April 4, 2017. We adapted the BG signal for each day of 2017. A generated p^{BG} for the original day and the one adapted to December 21, 2017, are presented in Fig. 12(b) and (c), respectively. An annual time series of p^{BG} is shown in Fig. 12(d).

VIII. Demonstration of Simulated Lidar Measurements

For September 1, 2017, we calculate the backscatter coefficient $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{mol}} + \boldsymbol{\beta}^{\mathrm{aer}}$ [1/(km · sr)] and the optical depth $\boldsymbol{\tau} = \boldsymbol{\tau}^{\mathrm{mol}} + \boldsymbol{\tau}^{\mathrm{aer}}$. Using (19), we calculate the total $\boldsymbol{\beta}_{\mathrm{ATTN}}$. Extinction coefficients at all wavelengths are shown in Fig. 9(c), and $\boldsymbol{\beta}_{\mathrm{ATTN}}$ is, respectively, shown in Fig. 13(a).

Then, we derive $\mathbf{P}$[photons] using (14). The lidar range corrected signal (RCS) $\mathbf{R}^2\mathbf{P}$[photons·km^2] is shown in Fig. 13(b). Following a Poisson process [see (15)], we get a simulated diurnal lidar measurement $\mathbf{P}_{\mathrm{M}}$[photons]. In Fig. 13(c), we present the lidar RCS measurement $\mathbf{R}^2\mathbf{P}_{\mathrm{M}}$[photons · km^2].

For comparison, Fig. 14 shows a real example of a Polly$^{\mathrm{XT}}$ lidar RCS for a cloudless day during September 2017 at Haifa. Simulated measurements by ALiDAn in Fig. 13(c) possess similar characteristics to those of the Polly$^{\mathrm{XT}}$ lidar presented in Fig. 14. One can observe resemblance in the dynamic range of values per wavelength channel, in the diurnal atmospheric dynamic variability, and in the diurnal change due to the Sun BG signal. In this work, the aerosol optical properties, as shown in Fig. 7(a), are derived from a modified distribution of [78]. Hence, we expect only a partial statistical resemblance.

IX. Discussion

This work presents ALiDAn, a generation and augmentation framework of STMW lidar data based on actual signal statistics. ALiDAn incorporates effects related to atmospheric compounds, diurnal and seasonal phenomena, and EO effects. In addition, ALiDAn can be tailored to different geographical locations and times and, thus, can assist in simulations of new stations.

Generation via ALiDAn promotes broader use of well-established databases. ALiDAn can help generate traceable signals at any stage of their creation, eventually yielding

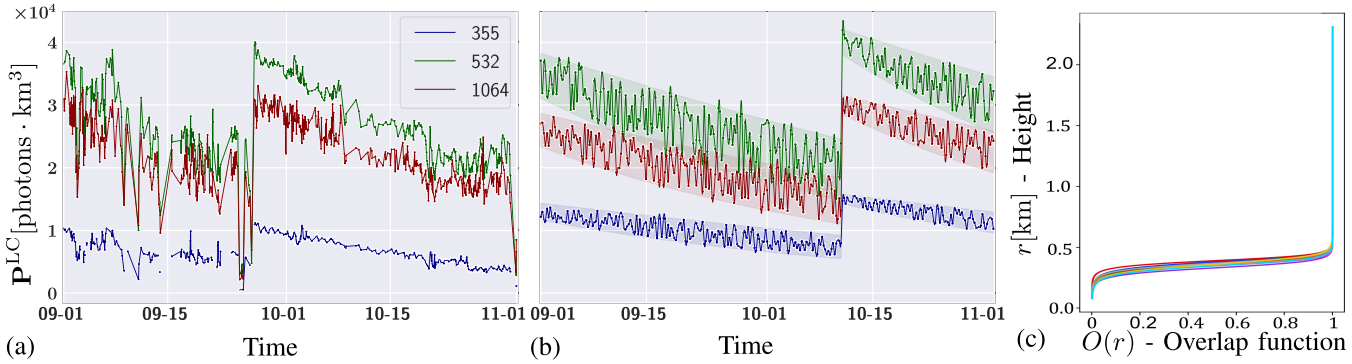


Fig. 11. EO factor generation. LC varying factor values for September–October 2017: (a) estimated by the PollyNet Processing Chain and (b) new randomly generated. LC values within shaded areas illustrate the prediction bands $\Delta_{\mathrm{LC}}(t, \lambda)$. (c) Several examples of generated overlap functions.

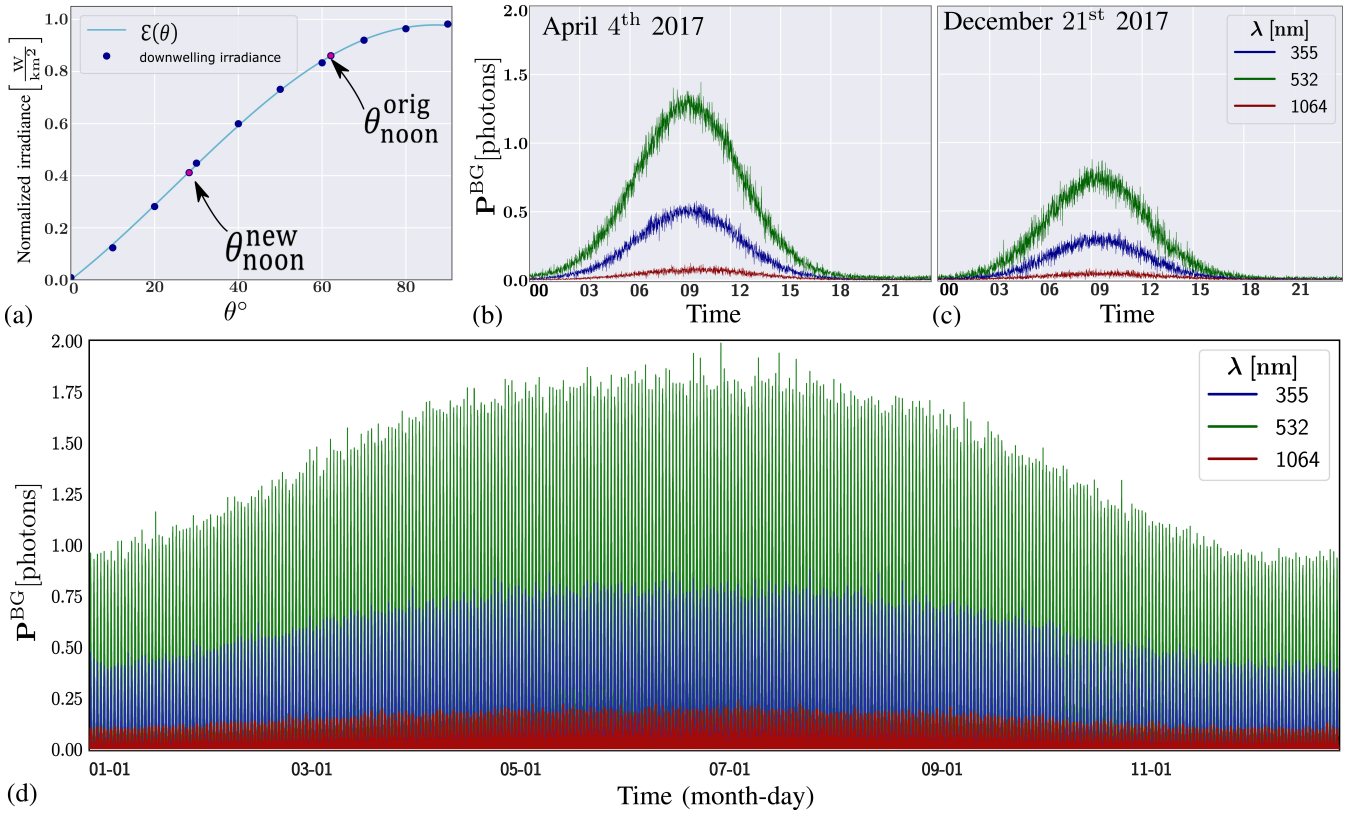


Fig. 12. (a) Normalized irradiance. (b) Diurnal simulated BG signals of April 4, 2017. (c) Diurnal simulated BG signals of December 21, 2017. (d) Annual simulated BG signal of 2017.

annotated databases. The generated data can assist in developing and testing lidar algorithms. We demonstrate this for supervised learning of lidar calibration in [88]. Future extensions or modifications can apply to other types of lidar and active sensors. For example, adding modules to ALiDAn can support polarized or inelastic Raman channels; this may address aerosol typing and clouds or water vapor analysis [89], [90], [91].

Two significant aspects set the strengths, and the limitations of ALiDAn are given as follows.

1) The first aspect relates to the source data statistics that our model relies on. Though being derived from well-established databases, the source data might be biased by various causes, such as temporal averaging of measurements, erroneous lidar calibration, and incorrect aerosol estimation arising from faulty assumptions or initialization in state-of-the-art retrieval algorithms, e.g., the LR assumption [1], [92]. Mitigating such issues is possible as lidar measurement technologies and analysis improve.
2) The second aspect relates to ALiDAn parametric model. This relates to a tradeoff between accuracy and practicality. The more complex the model (i.e., with more parameters it has), there is potentially better accuracy of its products. However, in practice, a more complex model requires resources, including training data and computations. We believe that, as time progresses, more data and better resources will facilitate better simulations by ALiDAn.

Modifications related to source statistics may enhance desirable characteristics and enrich simulation by ALiDAn, e.g.,

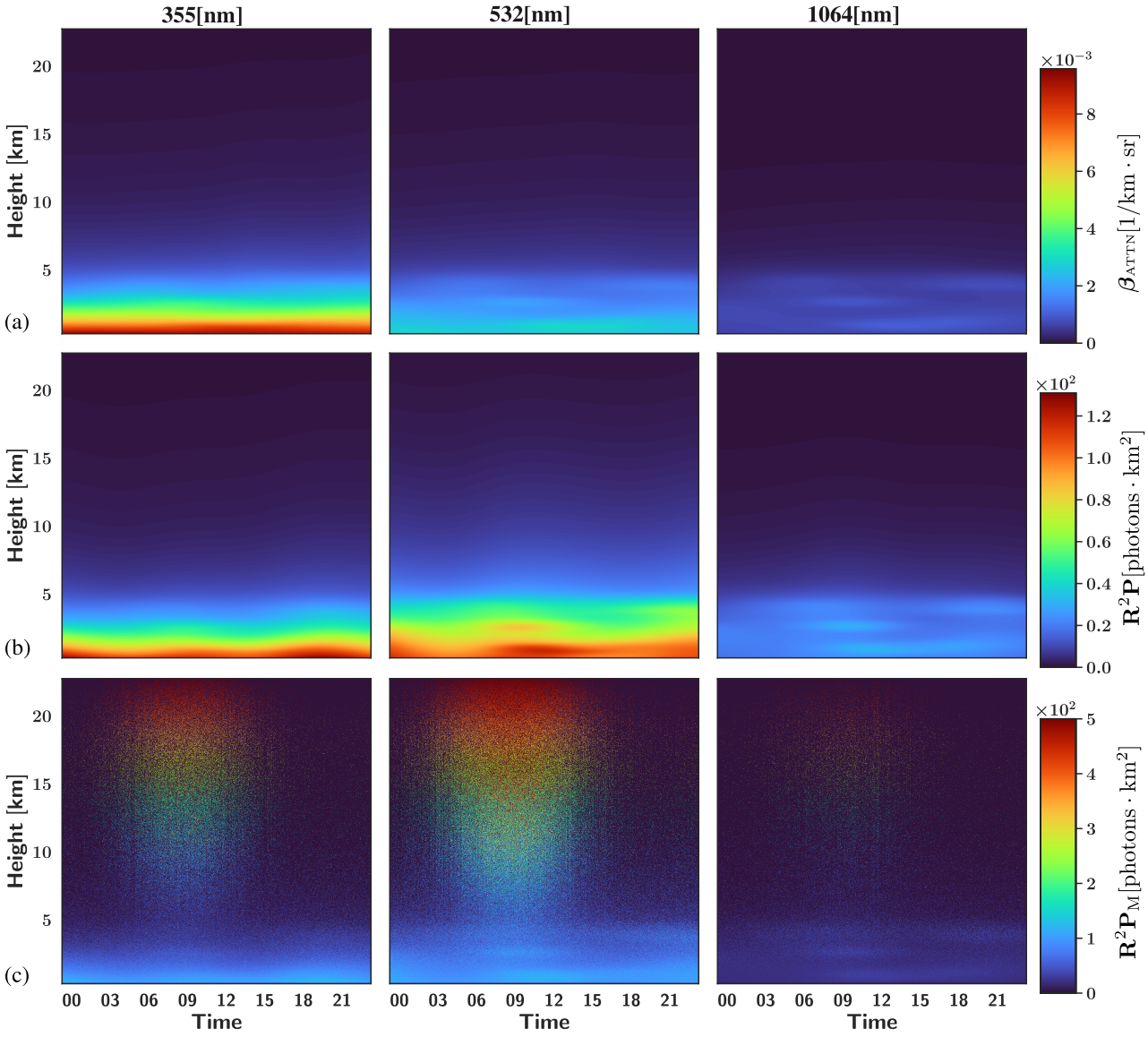


Fig. 13. Diurnal simulations for September 1, 2017. (a) Total attenuated backscatter coefficient $\boldsymbol{\beta}_{\text{ATTN}}[1/(\text{km} \cdot \text{sr})]$. (b) Lidar RCS $\mathbf{R}^2\mathbf{P}[\text{photons} \cdot \text{km}^2]$. (c) RCS of lidar measurement $\mathbf{R}^2\mathbf{P}_{\text{M}}[\text{photons} \cdot \text{km}^2]$.

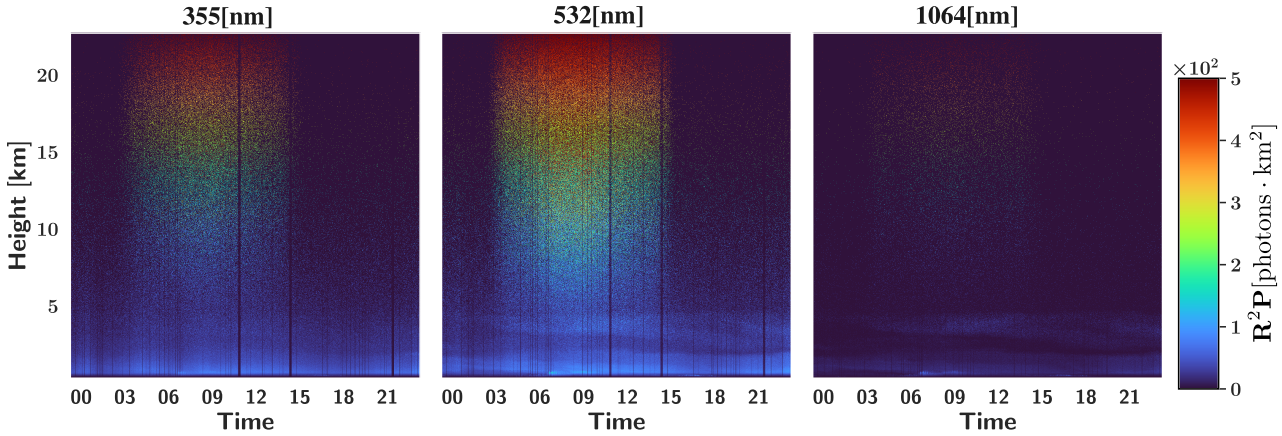


Fig. 14. An example of a lidar RCS $\mathbf{R}^2\mathbf{P}[\text{photons} \cdot \text{km}^2]$ measured on a cloudless day from September 12, 2017, at Haifa.

incorporation of meteorological data by ERA5 [66] for the molecular model (see Appendix B). The aerosol model can benefit from additional databases [26], [28], [67], [93], [94], [95], from which one can generate $\boldsymbol{\mathcal{D}}^{\text{aer}}_{\text{profiles}}$, $\boldsymbol{\mathcal{D}}_{\text{Å}}$, and $\boldsymbol{\mathcal{D}}_{\text{LR,Å}}$. Such databases may also be useful for locations without a sun-photometer. Wavelength-related statistics of $\boldsymbol{\mathcal{D}}_{\text{LR,Å}}$ obtained

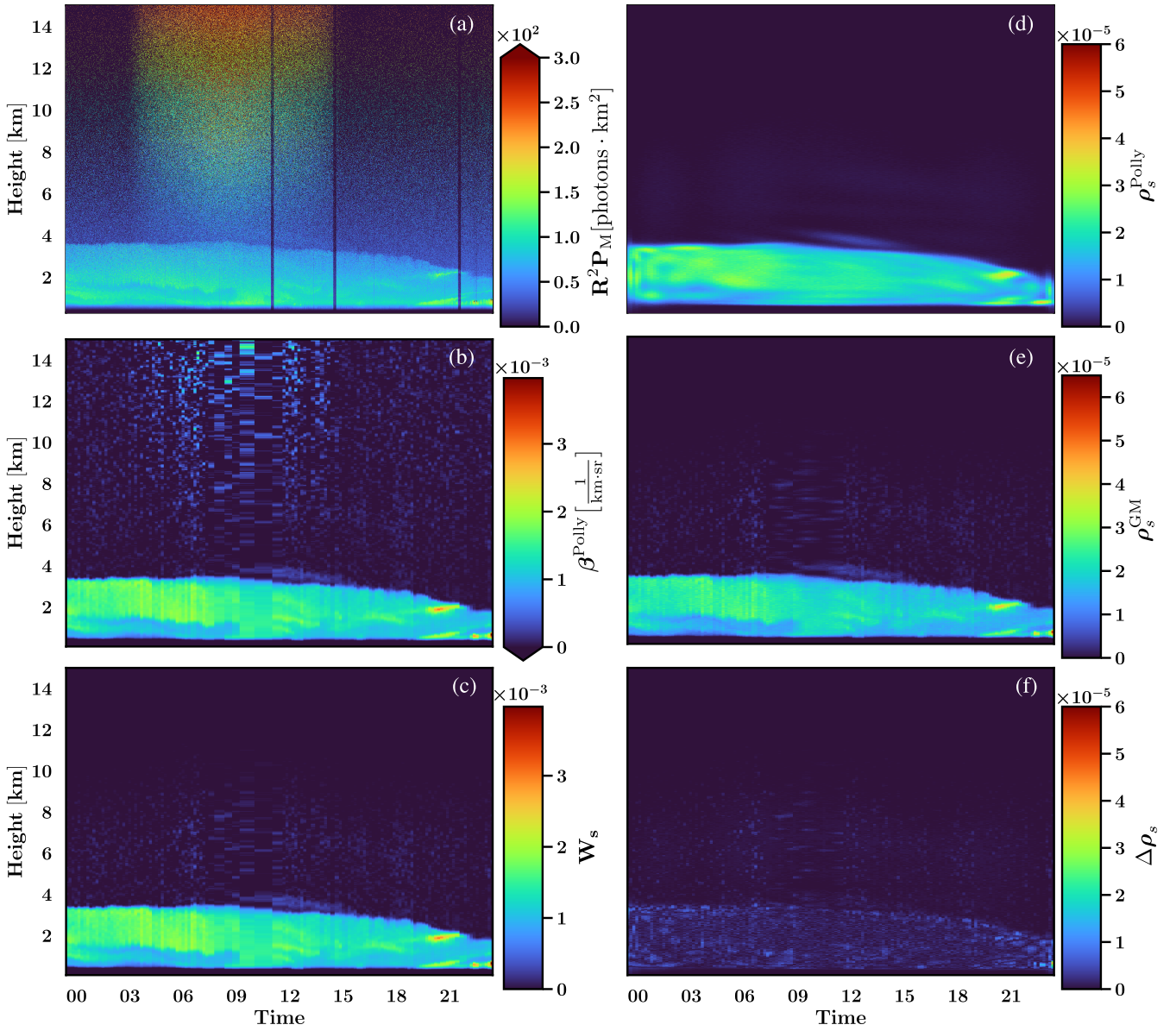


Fig. 15. Representation of empirical spatiotemporal aerosol data by a 2D GM distribution. (a) (Left) Lidar RCS $\mathbf{R}^2\mathbf{P}$[photons · km^2] at $\lambda = 532$ [nm] measured on October 2, 2017, at Haifa. (b) Daily estimated aerosol backscatter by PollyNet Processing Chain $\boldsymbol{\beta}^\mathrm{Polly}$[1/(km · sr)]. (c) Sample weights $\mathbf{W_s}$. (Right) Samples' densities: (d) $\boldsymbol{\rho}_s^\mathrm{Polly}$—derived from weights of $\boldsymbol{\beta}^\mathrm{Polly}$, and (e) $\boldsymbol{\rho}_s^\mathrm{GM}$—derived from the trained GM model. (f) Difference of the densities $\Delta\boldsymbol{\rho}_s = |\boldsymbol{\rho}_s^\mathrm{Polly} - \boldsymbol{\rho}_s^\mathrm{GM}|$.

over long periods [78], [96], [97] can expand data on aerosols, in particular, advanced analysis at 1064 nm [98].

ALiDAn can be adapted to handle aerosol microphysical properties, i.e., $\{\varpi^\mathrm{aer}, \mathcal{P}^\mathrm{aer}, \sigma^\mathrm{aer}, \rho_\#^\mathrm{aer}\}$, in addition to aerosol optical properties. It may be useful when having statistical particle compositions instead of optical properties. Data of this kind can be available via reanalysis databases, such as CAMS [95] or MERRA-2 [67]. Enabling such an extension requires an *a priori* calculation of aerosol optical properties. Such a calculation is available via OPAC [30] or a modeling tool of aerosol optical properties [99].

Modifications related to the parametric model can focus on aerosols. The aerosols' variability is related to changes in their density or type. The variability depends on atmospheric dynamics. Dynamics can be simulated using physical simulators or can be learned from measurements. Representation of aerosol type variability may be enhanced via conditional sampling of LR and Å values from typical aerosol type variations. Typical aerosol type variations may be characterized by learning joint change patterns of LR and Å.

Sequential lidar measurements are often represented by a 2D spatiotemporal map. Statistical image-based methods [100], [101] may yield spatiotemporal density representations to enrich aerosol generation. To support effective learning of natural aerosol variability, these approaches may require additional computational efforts and sufficient spatiotemporal aerosol data.

The overlap function depends on the mechanical structure stability of the lidar system [102]. Diurnal or abrupt variations of the overlap function may occur due to variations in ambient temperature. The overlap function variability can be represented by learning joint change patterns of additional indicators, such as the lidar's power and housing temperature.

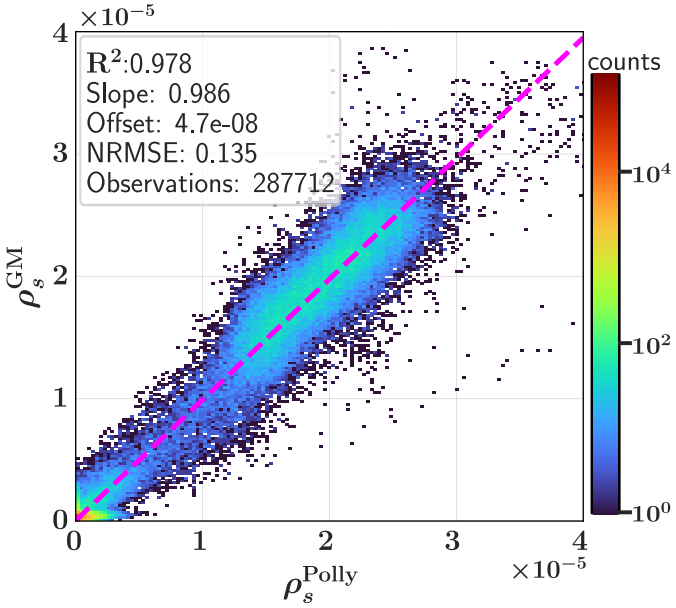


Fig. 16. Scatter plot comparing $\boldsymbol{\rho}_s^{\mathrm{GM}}$ and $\boldsymbol{\rho}_s^{\mathrm{Polly}}$ and the corresponding linear regression.

APPENDIX A

An aerosol distribution varies vertically and over time. To model aerosols' spatiotemporal distribution, we suggest using a 2D GM distribution[2] in Section V-B2. Here, we demonstrate a representation of empirical aerosol data in the spatiotemporal domain using a 2D GM distribution. We employ a *scikit-learn* [108], [109] implementation for fitting a GM distribution [110], [111].

Fig. 15(a) presents the lidar RCS measured at Haifa on October 2 at $\lambda = 532$. The aerosol backscatter profiles are estimated by the PollyNet Processing Chain [47], using the Klett–Fernald algorithm [92], for an average reference height $\langle r_{\mathrm{ref}} \rangle = 4.65$ [km] and an assumed LR $= 50$ [sr]. Each profile is retrieved from a 10 [min] period of consecutive lidar samples. Fig. 15(b) presents $\boldsymbol{\beta}^{\mathrm{Polly}}$, a set of 122 aerosol backscatter coefficient profiles.

We apply time interpolation to $\boldsymbol{\beta}^{\mathrm{Polly}}$; this allows using data from $n_s = 144$ profiles appearing every 10 [min] in the day. Each profile has $m_s = 1998$ height bins, corresponding to measurements up to an altitude of 15 [km]. Additional processing steps applied to $\boldsymbol{\beta}^{\mathrm{Polly}}$ include trimming of negative values, removing of noisy values above the aerosol layer using a height-dependent sigmoid function, and applying min–max scaling using (22) to set weights $\mathbf{W_s} \in \mathbb{R}^{m_s \times n_s}$, as presented in Fig. 15(c).

The weights $\mathbf{W_s}$ are used for deriving samples that feed the GM fitting process; Fig. 15(d) presents accordingly the samples density $\boldsymbol{\rho}_s^{\mathrm{Polly}} \in \mathbb{R}^{m_s \times n_s}$. Fig. 15(e) presents $\boldsymbol{\rho}_s^{\mathrm{GM}} \in \mathbb{R}^{m_s \times n_s}$, the density of samples derived from a continuous fit GM distribution by 32 2D Gaussians. Fig. 15(f) presents the density difference. The normalized root mean square error is NRMSE $= (\|\boldsymbol{\rho}_s^{\mathrm{Polly}} - \boldsymbol{\rho}_s^{\mathrm{GM}}\|_2 / \|\boldsymbol{\rho}_s^{\mathrm{Polly}}\|_2) = 0.135$. Fig. 16 presents a linear regression for a scatter plot of $\boldsymbol{\rho}_s^{\mathrm{GM}}$ versus $\boldsymbol{\rho}_s^{\mathrm{Polly}}$ at corresponding grid locations.

[2]Representations of vertical aerosol profiles via Gaussian basis functions have been used in aerosol optical depth estimations [103], [104], [105], [106] and profiles inversion [107] from satellites.

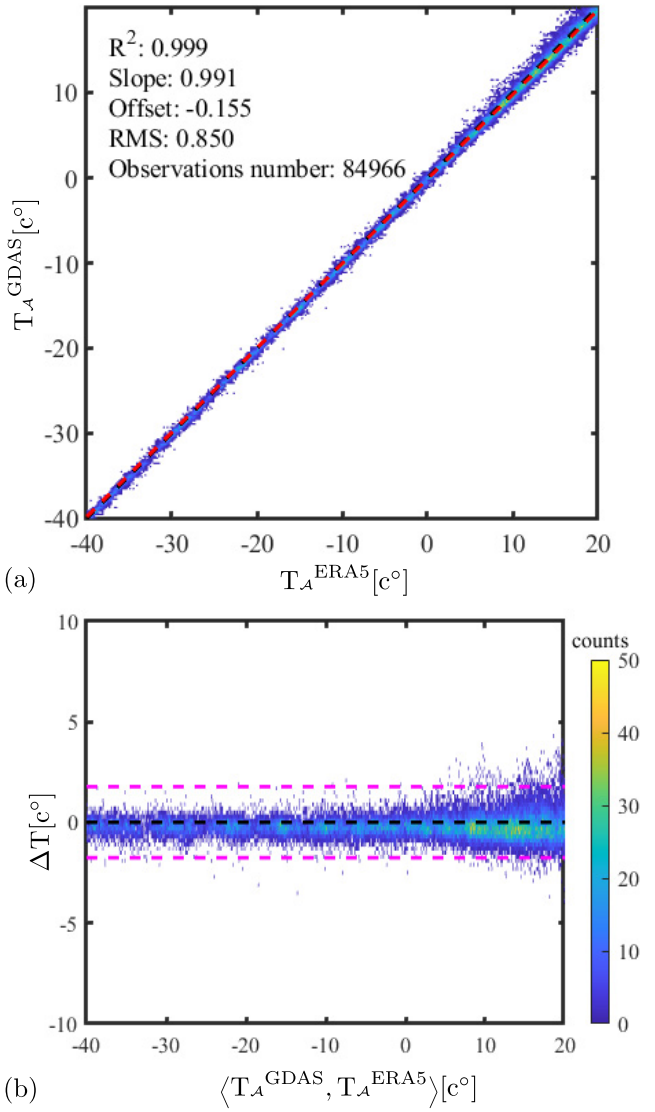


Fig. 17. Comparison between GDAS and ERA5 of temperature measurements for Haifa during April, May, September, and October 2017. (a) Linear fit expressing the relationship between GDAS and ERA5 temperatures. (b) 2D histogram of temperature differences between GDAS and ERA5, $\Delta\mathrm{T[c°]} \triangleq \mathrm{T}_{\mathcal{A}}^{\mathrm{GDAS}} - \mathrm{T}_{\mathcal{A}}^{\mathrm{ERA5}}$, for varying average temperature levels. The corresponding mean and standard deviation are $\mu_{\Delta\mathrm{T}} = 0.014°$ and $\sigma_{\Delta\mathrm{T}} = 0.899°$, respectively. The magenta dashed lines show a confidence level of 95% of the data.

APPENDIX B

As mentioned in Section V-A, the molecular distribution is calculated from meteorological measurements. For the molecular model, we have used GDAS [64] to be consistent with the PollyNet Processing Chain [47]. ERA5 [66] provides similar data, available every hour at four times higher spatial resolution. Although ERA5 provides a higher resolution of temperature and wind speed/direction, we have not observed significant differences for data related to molecular optical coefficients. This conclusion is supported by a temperature comparison between GDAS and ERA5 at corresponding pressure levels. The comparison was made on measurements of Haifa during April, May, September, and October 2017, corresponding to months of data presented in this work. Fig. 17(a) and (b), respectively, shows linear regression of GDAS and ERA5 temperatures and their residuals. The results show a

high correlation between GDAS and ERA5 measured temperatures and pressures.

ACKNOWLEDGMENT

The authors thank I. Czerninski, Y. Sde Chen, M. Tzabari, Y. Bertschy, M. Fisher, J. Hofer, and A. Floutsi for their advice and R. Hengst, I. Talmon, and D. Yagodin for technical support. Yoav Y. Schechner's work was conducted at the Ollendorff Minerva Center; Minvera is funded through the BMBF.

REFERENCES

[1] H. Baars *et al.*, "An overview of the first decade of Polly^{NET}: An emerging network of automated Raman-polarization LiDARs for continuous aerosol profiling," *Atmos. Chem. Phys.*, vol. 16, no. 8, pp. 5111–5137, 2016.

[2] Z. Yin, H. Baars, P. Seifert, and R. Engelmann, "Automatic LiDAR calibration and processing program for multiwavelength Raman polarization LiDAR," in *Proc. EPJ Web Conf.*, vol. 237, 2020, p. 8007.

[3] M. Wendisch *et al.*, "The Arctic cloud puzzle: Using ACLOUD/PASCAL multiplatform observations to unravel the role of clouds and aerosol particles in Arctic amplification," *Bull. Amer. Meteorolog. Soc.*, vol. 100, pp. 841–871, May 2019.

[4] R. A. Weekley, R. K. Goodrich, and L. B. Cornman, "Aerosol plume detection algorithm based on image segmentation of scanning atmospheric LiDAR data," *J. Atmos. Ocean. Technol.*, vol. 33, no. 4, pp. 697–712, Apr. 2016.

[5] E. Cromwell and D. Flynn, "LiDAR cloud detection with fully convolutional networks," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 619–627.

[6] W. J. Marais, R. E. Holz, Y. H. Hu, R. E. Kuehn, E. E. Eloranta, and R. M. Willett, "Approach to simultaneously denoise and invert backscatter and extinction from photon-limited atmospheric LiDAR observations," *Appl. Opt.*, vol. 55, no. 29, pp. 8316–8334, 2016.

[7] Z. Zheng, W. Chen, Y. Zhang, S. Chen, and D. Liu, "Denoising the space-borne high-spectral-resolution LiDAR signal with block-matching and 3D filtering," *Appl. Opt.*, vol. 59, no. 9, pp. 2820–2828, 2020.

[8] S. Zeng *et al.*, "Identifying aerosol subtypes from CALIPSO LiDAR profiles using deep machine learning," *Atmosphere*, vol. 12, no. 1, p. 10, Dec. 2020.

[9] D. R. Vivas, E. Sánchez, and J. H. Reina, "Deep learning the atmospheric boundary layer height," 2020, *arXiv:2004.04353*.

[10] N. Kumar, K. Soni, and R. Agarwal, "Prediction of temporal atmospheric boundary layer height using long short-term memory network," *Tellus A, Dyn. Meteorol. Oceanogr.*, vol. 73, no. 1, pp. 1–14, 2021.

[11] D. Ziaei, J. Sleeman, M. Halem, V. Caicedo, R. M. Delgado, and B. Demoz, "Convolutional LSTM for planetary boundary layer height (PBLH) prediction," *UMBC Joint Center Earth Syst. Technol.*, vol. 2021, pp. 1–6, Mar. 2021.

[12] U. Wandinger, A. Hiebsch, I. Mattis, G. Pappalardo, L. Mona, and F. Madonna, "Aerosols and clouds: Long-term database from spaceborne lidar measurements," ESTEC Contract 21487, Final Rep., 2011. Accessed: Nov. 14, 2022. [Online]. Available: https://www.tropos.de/fileadmin/user_upload/Institut/Abteilungen/Fernerkundung/Daten_PDF/Wandinger-ESA_2011.pdf

[13] G. Pappalardo *et al.*, "Four-dimensional distribution of the 2010 Eyjafjallajökull volcanic cloud over Europe observed by EARLINET," *Atmos. Chem. Phys.*, vol. 13, no. 8, pp. 4429–4450, 2013.

[14] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, Dec. 2018.

[15] J. Tremblay *et al.*, "Training deep networks with synthetic data: Bridging the reality gap by domain randomization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 969–977.

[16] S. I. Nikolenko, *Synthetic Data for Deep Learning*. Switzerland: Springer, 2021, doi: 10.1007/978-3-030-75178-4.

[17] R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE J. Quantum Electron.*, vol. 37, no. 3, pp. 390–397, Mar. 2001.

[18] G. Berkovic and E. Shafir, "Optical methods for distance and displacement measurements," *Adv. Opt. Photon.*, vol. 4, no. 4, pp. 441–471, 2012.

[19] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin, "Phasor imaging: A generalization of correlation-based time-of-flight imaging," *ACM Trans. Graph.*, vol. 34, no. 5, pp. 1–18, 2015.

[20] R. Engelmann *et al.*, "The automated multiwavelength Raman polarization and water-vapor LiDAR Polly^{XT}: The neXT generation," *Atmos. Meas. Techn.*, vol. 9, no. 4, pp. 1767–1784, 2016.

[21] A. Aides, Y. Y. Schechner, V. Holodovsky, A. Levis, and D. Althausen, "Measuring atmospheric scattering in 3D," in *Proc. Propag. Characterization Atmos. Ocean. Phenomena*, 2018, pp. 1–2.

[22] B. Heese, J. Hofer, H. Baars, R. Engelmann, D. Althausen, and Y. Y. Schechner, "Wild fire aerosol optical properties measured by LiDAR at Haifa, Israel," in *Proc. EPJ Web Conf.*, vol. 176, 2018, p. 05049.

[23] D. Althausen *et al.*, "Vertical profiles of dust and other aerosol types above a coastal site," in *Proc. E3S Web Conf.*, vol. 99, 2019, p. 02005.

[24] A. Aides, A. Levis, V. Holodovsky, Y. Y. Schechner, D. Althausen, and A. Vainiger, "Distributed sky imaging radiometry and tomography," in *Proc. IEEE Int. Conf. Comput. Photogr. (ICCP)*, Apr. 2020, pp. 1–12.

[25] B. Heese *et al.*, "The vertical aerosol type distribution above Israel–2 years of LiDAR observations at the coastal city of Haifa," *Atmos. Chem. Phys.*, vol. 22, no. 3, pp. 1633–1648, Feb. 2022.

[26] L. Mona *et al.*, "EARLINET database: New design and new products for a wider use of aerosol LiDAR data," in *Proc. EPJ Web Conf.*, vol. 176, 2018, p. 9016.

[27] C. Böckmann *et al.*, "Aerosol LiDAR intercomparison in the framework of the EARLINET project. 2. Aerosol backscatter algorithms," *Appl. Opt.*, vol. 43, no. 4, pp. 977–989, 2004.

[28] V. Amiridis *et al.*, "LIVAS: A 3-D multi-wavelength aerosol/cloud database based on CALIPSO and EARLINET," *Atmos. Chem. Phys.*, vol. 15, no. 13, pp. 7127–7153, Jul. 2015.

[29] M. Hess, P. Koepke, and I. Schult, "Optical properties of aerosols and clouds: The software package OPAC," *Bull. Amer. Meteorol. Soc.*, vol. 79, no. 5, pp. 831–844, 1998.

[30] P. Koepke, J. Gasteiger, and M. Hess, "Technical note: Optical properties of desert aerosol with non-spherical mineral particles: Data incorporated to OPAC," *Atmos. Chem. Phys.*, vol. 15, no. 10, pp. 5947–5956, May 2015.

[31] R. Pedrós *et al.*, "AEROgui: A graphical user interface for the optical properties of aerosols," *Bull. Amer. Meteorolog. Soc.*, vol. 95, no. 12, pp. 1863–1871, Dec. 2014.

[32] J. Vicent *et al.*, "Comparative analysis of atmospheric radiative transfer models using the atmospheric look-up table generator (ALG) toolbox (version 2.0)," *Geosci. Model Develop.*, vol. 13, no. 4, pp. 1945–1957, 2020.

[33] D. P. Donovan, H. K. Baltink, J. S. Henzing, S. R. de Roode, and A. P. Siebesma, "A depolarisation LiDAR-based method for the determination of liquid-cloud microphysical properties," *Atmos. Meas. Techn.*, vol. 8, no. 1, pp. 237–266, Jan. 2015.

[34] H. W. Barker, R. K. Goldstein, and D. E. Stevens, "Monte Carlo simulation of solar reflectances for cloudy atmospheres," *J. Atmos. Sci.*, vol. 60, no. 16, pp. 1881–1894, Aug. 2003.

[35] K. Stamnes, S.-C. Tsay, W. Wiscombe, and K. Jayaweera, "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media," *Appl. Opt.*, vol. 27, no. 12, pp. 2502–2509, 1988.

[36] B. Mayer and A. Kylling, "Technical note: The libRadtran software package for radiative transfer calculations—Description and examples of use," *Atmos. Chem. Phys.*, vol. 5, no. 7, pp. 1855–1877, Jul. 2005.

[37] A. J. Illingworth *et al.*, "The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation," *Bull. Amer. Meteorol. Soc.*, vol. 96, no. 8, pp. 1311–1332, 2015.

[38] R. Voors *et al.*, "ECSIM: The simulator framework for EarthCARE," *Proc. SPIE*, vol. 6744, Oct. 2007, Art. no. 67441Y.

[39] M. Reverdy *et al.*, "An EarthCARE/ATLID simulator to evaluate cloud description in climate models," *J. Geophys. Res., Atmos.*, vol. 120, no. 21, pp. 11–90, 2015.

[40] D. Donovan *et al.*, "ECSIM model and algorithms document," *ESTEC Contract*, vol. 20003, no. 6, pp. 1–207, 2008.

[41] J. P. do Carmo, G. de Villele, A. Helière, K. Wallace, A. Lefebvre, and F. Chassat, "ATLID, ESA atmospheric backscatter LIDAR for the ESA EarthCARE mission," *CEAS Space J.*, vol. 11, no. 4, pp. 423–435, Dec. 2019.

[42] J. A. Bravo-Aranda *et al.*, "Assessment of LiDAR depolarization uncertainty by means of a polarimetric LiDAR simulator," *Atmos. Meas. Techn.*, vol. 9, no. 10, pp. 4935–4953, Oct. 2016.

[43] D. Nicolae *et al.*, "A neural network aerosol-typing algorithm based on LiDAR data," *Atmos. Chem. Phys.*, vol. 18, no. 19, pp. 14511–14537, Oct. 2018.

[44] G. Pappalardo *et al.*, "EARLINET: Towards an advanced sustainable European aerosol LiDAR network," *Atmos. Meas. Tech.*, vol. 7, no. 8, pp. 2389–2409, 2014.

[45] E. J. Welton, J. R. Campbell, J. D. Spinhirne, and V. S. Scott, "Global monitoring of clouds and aerosols using a network of micropulse LiDAR systems," *Proc. SPIE*, vol. 4153, Feb. 2001, pp. 151–158.

[46] H. Baars *et al.*, "Polly^{NET}—An emerging network of automated Raman-polarizarion LiDARs for continuous aerosolprofiling," in *Proc. EPJ Web Conf.*, vol. 176, 2018, p. 9013.

[47] Z. Yin and H. Baars, "PollyNET/pollynet_processing_chain: Version 2.1," Zenodo, 2021. [Online]. Available: https://github.com/PollyNET/Pollynet_Processing_Chain, doi: 10.5281/zenodo.3774689.

[48] B. N. Holben *et al.*, "AERONET—A federated instrument network and data archive for aerosol characterization," *Remote Sens. Environ.*, vol. 66, no. 1, pp. 1–16, Oct. 1998.

[49] C. Weitkamp, *LiDAR: Range-Resolved Optical Remote Sensing of the Atmosphere*, vol. 102. New York, NY, USA: Springer, 2006.

[50] H. Baars, P. Seifert, R. Engelmann, and U. Wandinger, "Target categorization of aerosol and clouds by continuous multiwavelength-polarization LiDAR measurements," *Atmos. Meas. Techn.*, vol. 10, no. 9, pp. 3175–3201, Sep. 2017.

[51] *US Standard Atmosphere*, Washington, DC, USA: Nat. Ocean. Atmos. Admin., 1976.

[52] A. Bucholtz, "Rayleigh-scattering calculations for the terrestrial atmosphere," *Appl. Opt.*, vol. 34, no. 15, pp. 2765–2773, Aug. 1995.

[53] R. B. Miles, W. R. Lempert, and J. N. Forkey, "Laser Rayleigh scattering," *Meas. Sci. Technol.*, vol. 12, no. 5, p. R33, 2001.

[54] J. Ackermann, "The extinction-to-backscatter ratio of tropospheric aerosol: A numerical study," *J. Atmos. Ocean. Technol.*, vol. 15, no. 4, pp. 1043–1050, Aug. 1998.

[55] O. Dubovik *et al.*, "Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust," *J. Geophys. Res.*, vol. 111, no. D11, pp. 1–34, 2006.

[56] D. Müller *et al.*, "Aerosol-type-dependent LiDAR ratios observed with Raman LiDAR," *J. Geophys. Res.*, vol. 112, no. D16, pp. 1–11, 2007.

[57] A. Marshak and A. Davis, *3D Radiative Transfer in Cloudy Atmospheres*. Berlin, Germany: Springer, 2005.

[58] A. Ångström, "The parameters of atmospheric turbidity," *Tellus*, vol. 16, no. 1, pp. 64–75, 1964.

[59] Y. Song *et al.*, "Correlation between the LiDAR ratio and the Ångström exponent of various aerosol types," *Particuology*, vol. 40, pp. 62–69, Oct. 2018.

[60] T. Eck *et al.*, "Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols," *J. Geophys. Res.*, vol. 104, no. 24, pp. 31333–31349, Dec. 1999.

[61] S. Hoyer and J. Hamman, "xarray: ND labeled arrays and datasets in Python," *J. Open Res. Softw.*, vol. 5, no. 1, p. 10, Apr. 2017.

[62] D. P. Kroese, T. Taimre, and Z. I. Botev, *Handbook of Monte Carlo Methods*, vol. 706. Hoboken, NJ, USA: Wiley, 2013.

[63] D. Althausen *et al.*, "Portable Raman LiDAR Polly^{XT} for automated profiling of aerosol backscatter, extinction, and depolarization," *J. Atmos. Ocean. Technol.*, vol. 26, no. 11, pp. 2366–2378, 2009.

[64] NOAA, "Global Data Assimilation System (GDAS1) archive information," 2014. Accessed: Nov. 14, 2022. [Online]. Available: https://www.ready.noaa.gov/gdas1.php

[65] L. Oolman, "Atmospheric soundings measurements," Univ. Wyoming. Accessed: Sep. 1, 2022. [Online]. Available: http://weather.uwyo.edu/upperair/sounding.html

[66] H. Hersbach *et al.*, "ERA5 hourly data on single levels from 1979 to present," Copernicus Climate Change Service (C3S) Climate Data Store (CDS), ECMWF, Reading, U.K., 2018, vol. 10, doi: 10.24381/cds.adbb2d47.

[67] R. Gelaro *et al.*, "The modern-era retrospective analysis for research and applications, version 2 (MERRA-2)," *J. Climate*, vol. 30, pp. 5419–5454, Jul. 2017.

[68] I. Binietoglou. (2015). *LiDAR Molecular*. [Online]. Available: https://gitlab.com/ioannis_binietoglou/lidar_molecular/

[69] F. G. Fernald, B. M. Herman, and J. A. Reagan, "Determination of aerosol height distributions by LiDAR," *J. Appl. Meteorol.*, vol. 11, no. 3, pp. 482–489, Apr. 1972.

[70] A. Ansmann, U. Wandinger, M. Riebesell, C. Weitkamp, and W. Michaelis, "Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter LiDAR," *Appl. Opt.*, vol. 31, no. 33, pp. 7113–7131, 1992.

[71] V. Freudenthaler, H. Linné, A. Chaikovski, D. Rabus, and S. Groß, "EARLINET lidar quality assurance tools," *Atmos. Meas. Techn. Discuss.*, pp. 1–35, 2018. Accessed: Nov. 14, 2022. [Online]. Available: https://amt.copernicus.org/preprints/amt-2017-395/, doi: 10.5194/amt-2017-395.

[72] V. Freudenthaler, "Lidar Rayleigh–fit criteria," in *EARLINET-ASOS 7th Workshop*, 2009. Accessed: Nov. 14, 2022. [Online]. Available: https://epub.ub.uni-muenchen.de/12970/, doi: 10.5282/ubm/epub.12970.

[73] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. Hoboken, NJ, USA: Wiley, 2015.

[74] N. Papagiannopoulos *et al.*, "CALIPSO climatological products: Evaluation and suggestions from EARLINET," *Atmos. Chem. Phys.*, vol. 16, no. 4, pp. 2341–2357, Feb. 2016.

[75] A. M. de Oliveira *et al.*, "Analysis of atmospheric aerosol optical properties in the Northeast Brazilian atmosphere with remote sensing data from MODIS and CALIOP/CALIPSO satellites, AERONET photometers and a ground-based LiDAR," *Atmosphere*, vol. 10, no. 10, p. 594, Oct. 2019.

[76] M. Pharr, W. Jakob, and G. Humphreys, *Physically Based Rendering: From Theory to Implementation*. San Mateo, CA, USA: Morgan Kaufmann, 2016.

[77] I. Rogozovsky *et al.*, "Impact of aerosol layering, complex aerosol mixing, and cloud coverage on high-resolution MAIAC aerosol optical depth measurements: Fusion of LiDAR, AERONET, satellite, and ground-based measurements," *Atmos. Environ.*, vol. 247, Feb. 2021, Art. no. 118163.

[78] E. Giannakaki, P. G. van Zyl, D. Müller, D. Balis, and M. Komppula, "Optical and microphysical characterization of aerosol layers over South Africa by means of multi-wavelength depolarization and Raman LiDAR measurements," *Atmos. Chem. Phys.*, vol. 16, no. 13, pp. 8109–8123, Jul. 2016.

[79] A. Hänel, H. Baars, D. Althausen, A. Ansmann, R. Engelmann, and J. Y. Sun, "One-year aerosol profiling with EUCAARI Raman LiDAR at Shangdianzi GAW station: Beijing plume and seasonal variations," *J. Geophys. Res., Atmos.*, vol. 117, no. D13, pp. 1–11, Jul. 2012.

[80] Z. Yin *et al.*, "Long-term variations of aerosol optical properties over Wuhan with polarization LiDAR," *Atmos. Environ.*, vol. 259, Aug. 2021, Art. no. 118508.

[81] D. P. Kroese, Z. I. Botev, T. Taimre, and R. Vaisman, *Data Science and Machine Learning: Mathematical and Statistical Methods*. London, U.K.: Chapman & Hall, 2019.

[82] U. Wandinger and A. Ansmann, "Experimental determination of the LiDAR overlap profile with Raman LiDAR," *Appl. Opt.*, vol. 41, no. 3, pp. 511–514, Jan. 2002.

[83] X. Shen *et al.*, "Development of ZJU high-spectral-resolution LiDAR for aerosol and cloud: Calibration of overlap function," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 257, Dec. 2020, Art. no. 107338.

[84] NOAA. (2022). *Solar Positions Calculations*. [Online]. Available: https://gml.noaa.gov/grad/solcalc/calcdetails.html

[85] S. Kennedy. (2009). *Astral, Version: 2.2*. [Online]. Available: https://pypi.org/project/astral/

[86] B. Guinot, "Solar time, legal time, time in use," *Metrologia*, vol. 48, no. 4, pp. S181–S185, Aug. 2011.

[87] C. Rigollier, O. Bauer, and L. Wald, "On the clear sky model of the ESRA—European solar radiation atlas—With respect to the Heliosat method," *Sol. Energy*, vol. 68, no. 1, pp. 33–48, Jan. 2000.

[88] A. Vainiger *et al.*, "Supervised learning calibration of an atmospheric LiDAR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2022, pp. 7317–7320, doi: 10.1109/IGARSS46834.2022.9883078.

[89] S. Qi *et al.*, "Classification of atmospheric aerosols and clouds by use of dual-polarization LiDAR measurements," *Opt. Exp.*, vol. 29, no. 15, pp. 23461–23476, 2021.

[90] C. Jimenez *et al.*, "The dual-field-of-view polarization LiDAR technique: A new concept in monitoring aerosol effects in liquid-water clouds–case studies," *Atmos. Chem. Phys.*, vol. 20, no. 23, pp. 15265–15284, 2020.

[91] A. Fix, F. Steinebach, M. Wirth, A. Schäfler, and G. Ehret, "Development and application of an airborne differential absorption LiDAR for the simultaneous measurement of ozone and water vapor profiles in the tropopause region," *Appl. Opt.*, vol. 58, no. 22, pp. 5892–5900, 2019.

[92] F. G. Fernald, "Analysis of atmospheric LiDAR observations: Some comments," *Appl. Opt.*, vol. 23, no. 5, pp. 652–653, 1984.

[93] D. M. Winker, J. L. Tackett, B. J. Getzewich, Z. Liu, M. A. Vaughan, and R. R. Rogers, "The global 3-D distribution of tropospheric aerosols as characterized by CALIOP," *Atmos. Chem. Phys.*, vol. 13, no. 6, pp. 3345–3361, Mar. 2013.

[94] (Mar. 2019). *EARLINET Database*. [Online]. Available: https://www.earlinet.org/index.php?id=125

[95] A. Inness *et al.*, "The cams reanalysis of atmospheric composition," *Atmos. Chem. Phys.*, vol. 19, no. 6, pp. 3515–3556, 2019.

[96] C. Kaduk, "Characterization of the optical properties of complex aerosol mixtures observed with a multiwavelength Raman polarization LiDAR during the 6-weeks BACCHUS campaign in Cyprus in spring 2015," M.S. thesis, Fac. Phys. Geosci., Leipzig Univ., Leipzig, Germany, 2017.

[97] A. A. Floutsi *et al.*, "Advection of biomass burning aerosols towards the southern hemispheric mid-latitude station of Punta Arenas as observed with multiwavelength polarization Raman LiDAR," *Remote Sens.*, vol. 13, no. 1, p. 138, Jan. 2021.
[98] M. Haarig *et al.*, "First triple-wavelength LiDAR observations of depolarization and extinction-to-backscatter ratios of Saharan dust," *Atmos. Chem. Phys.*, vol. 22, no. 1, pp. 355–369, Jan. 2022.
[99] J. Gasteiger and M. Wiegner, "MOPSMAP V1.0: A versatile tool for the modeling of aerosol optical properties," *Geosci. Model Develop.*, vol. 11, no. 7, pp. 2739–2762, Jul. 2018.
[100] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*.
[101] I. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–9.
[102] H. Baars, "Aerosol profiling with lidar in the Amazon basin during the wet and dry season 2008," Ph.D. dissertation, Leipzig Univ., 2012. Accessed: Nov. 14, 2022. [Online]. Available: https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa-98757
[103] N. C. Hsu, S.-C. Tsay, M. D. King, and J. R. Herman, "Aerosol properties over bright-reflecting source regions," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 557–569, Mar. 2004.
[104] Y. Wu, M. de Graaf, and M. Menenti, "The impact of aerosol vertical distribution on aerosol optical depth retrieval using CALIPSO and MODIS data: Case study over dust and smoke regions," *J. Geophys. Res., Atmos.*, vol. 122, no. 16, pp. 8801–8815, Aug. 2017.
[105] C. Li, J. Li, O. Dubovik, Z.-C. Zeng, and Y. L. Yung, "Impact of aerosol vertical distribution on aerosol optical depth retrieval from passive satellite sensors," *Remote Sens.*, vol. 12, no. 9, p. 1524, May 2020.
[106] Z. Song *et al.*, "Effect of the vertical distribution of absorbing aerosols on the atmospheric correction for satellite ocean color remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, 2022.
[107] O. Dubovik *et al.*, "Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations," *Atmos. Meas. Techn.*, vol. 4, pp. 975–1018, May 2011.
[108] F. Pedregosa *et al.*, "Scikit-Learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 10, pp. 2825–2830, Jul. 2017.
[109] Scikit-learn, "A Bayesian Gaussian Mixture class," Accessed: Sep. 1, 2022. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.mixture.BayesianGaussianMixture.html
[110] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131–146, Nov. 2008.
[111] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," *J. Amer. Stat. Assoc.*, vol. 112, no. 518, pp. 859–877, 2017.

Adi Vainiger (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, Israel, in 2012 and 2019, respectively, where she is currently pursuing the Ph.D. degree with Viterbi Faculty of Electrical and Computer Engineering.

She worked at Intel and Mobileye as a Software and Algorithms Engineer. Her research interests include computer-vision algorithms for 3D retrievals, remote sensing, computational photography, and statistical learning.

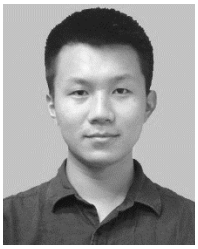
Omer Shubi received the B.Sc. degree in data science and engineering from the Technion—Israel Institute of Technology, Haifa, Israel, in 2021, where he is currently pursuing the M.Sc. degree in data science with the Faculty of Industrial Engineering and Management.

His research focuses on the intersection between natural language processing, psycholinguistics, and cognition.

Yoav Y. Schechner (Member, IEEE) received the B.A. and M.Sc. degrees in physics and the Ph.D. degree in electrical engineering (EE) from the Technion—Israel Institute of Technology, Haifa, Israel, in 1990, 1994, and 2000, respectively.

He was a Research Scientist with Columbia University, New York, NY, USA. Since 2002, he has been a Faculty Member with the Viterbi Faculty of Electrical Engineering, Technion. From 2010 to 2011, he was a Visiting Scientist with the California Institute of Technology (Caltech), Pasadena, CA, USA, and the NASA's Jet Propulsion Laboratory, Pasadena. He is currently the Diane and Mark Seiden Chair in Science at Technion. He is also a Principal Investigator and a Coordinator of the CloudCT project, funded by the European Research Council (ERC). His research interests include outdoor phenomena and all aspects of imaging.

Dr. Schechner was a recipient of the Technion Distinguished Lecturer Award in 2020, the Best Student Paper Award at CVPR in 2017, the Best Paper Awards at ICCP in 2013 and 2018, the Ray and Miriam Klein Research Award, the Henry Taub Prize for Academic Excellence, the Otto Schwarz Foundation Excellence Award, and the Landau Fellowship. He was also the Program Chair of ICCP 2021.



Zhenping Yin received the Ph.D. degree in space physics from Wuhan University, Wuhan, China, in 2021.

He started his post-doctoral research at the School of Remote Sensing and Information Engineering, Wuhan University, in August 2021. His research interests include lidar-radar synergistics and polarization lidar techniques.

Holger Baars received the Diploma degree in meteorology from Leipzig University, Leipzig, Germany, in 2007, and the Ph.D. degree in 2012.

Since then, he has been a Scientist with the Leibniz Institute for Tropospheric Research, Leipzig. He is a data analysis specialist and responsible for the PollyNET—a network of small, compact research lidars for profiling aerosol and cloud optical properties. He is also involved in spaceborne lidar missions, such as Aeolus and EarthCARE. His research interests include active remote sensing of aerosol and clouds.

Birgit Heese received the Diploma degree in meteorology from the University of Kiel, Kiel, Germany, in 1991, and the Ph.D. degree from the University of Bremen, Bremen, Germany, in 1995. The research for her diploma thesis was done at the Nansen Remote Sensing Center, Bergen, Norway, from 1990 to 1991, on a comparison between satellite- and ground-based stratospheric ozone measurements. Her Ph.D. thesis was based on ozone lidar measurements from Spitsbergen, Norway, which were conducted at the Alfred Wegener Institute for Marine and Polar Research (AWI), Bremerhaven, Germany.

From 1994 to 1998, she was an Associate Professor with The University Centre in Svalbard (UNIS), Spitsbergen, lecturing on the middle polar atmosphere. From 1998 to 2000, she held a post-doctoral position at the Service d'Aeronomie du CNRS, Paris, France, working with airborne ozone lidar measurements. From 2000 to 2006, she was a Scientist with the Lidar Group, Meteorological Institute, Ludwig Maximilian University of Munich, Munich, Germany, now measuring aerosol particles by lidar in the troposphere. Since 2007, she has been employed at the Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany, where she is currently a member of the Polly XT Lidar Group and continues with aerosol observations by lidar.

Dietrich Althausen studied physics from 1977 to 1982 at Leipzig University, Leipzig, Germany. He received the Ph.D. degree in 1992.

Since then, he has been a Senior Scientist with the Leibniz Institute for Tropospheric Research, Leipzig. Within these 30 years, the focus of his work is the development and application of multiwavelength polarization Raman lidars for the vertically resolved observations of particles, clouds, and water vapor in the atmosphere. In 2002, he started the development of small and compact aerosol lidar systems, the "Polly" family. Meanwhile, these systems are applied worldwide in the frame of ACTRIS. An important result of this research is the observation and characterization of aerosol layers that are transported intercontinentally at different heights—in the troposphere and even in the stratosphere.