3DeepCT: Learning Volumetric Scattering Tomography of Clouds, Supplementary Material

Yael Sde-Chen, Yoav Y. Schechner, Vadim Holodovsky

Viterbi Faculty of Electrical and Computer Engineering,

Technion - Israel Institute of Technology, Haifa, Israel

yael.sde.chen@gmail.com, yoav@ee.technion.ac.il, vholod@ef.technion.ac.il

Eshkol Eytan

Department of Earth and Planetary Science,

The Weizmann Institute of Science, Rehovot, Israel

eshkol.eytan@weizmann.ac.il

Abstract

This is a supplementary document to the main manuscript. Here we provide additional numerical results.

1. Introduction

This supplementary material contains eight sections. Section 2 describes the BOMEX and CASS setups for the cloud fields simulations. Section 3 describes the noise model we use in our simulated database. Section 4 presents a comparison between 3DeepCT and other initialization schemes. Section 5 describes the **10 Viewpoints** geometry. Section 6 presents some of the ablation studies. Section 8 provides additional data and results of the *subset of seven* clouds, which were not included in the main manuscript due to space limits. Section 7 AirMSPI real data cross validation results.

2. Datasets

The cloud field was simulated using the BOMEX [5] and CASS setups [6]. The BOMEX case study is based on surface fluxes, large-scale tendencies, and profiles of wind, humidity, and temperature in trade wind cumulus cloud fields over the Atlantic ocean near Barbados.

The CASS simulation of terrestrial clouds is based on a composite case called the Continental Active Surface-Forced Shallow Cumulus (CASS) [6]. The setup was constructed from measurements of the ARM project (Atmospheric Radiation Measurements), during 1997-2009 in the Southern Great Plains of the USA. A composite of the diurnal cycle of surface fluxes, large-scale forcing, and profiles of wind, humidity, and temperature was established over all days of shallow convection with non-precipitating cumulus clouds and was used to initiate the LES model.

The simulated domain for both setups is 12.82 km × 12.82 km wide, with cyclic horizontal boundary conditions. The time duration of the BOMEX simulation is 8 hours (including 2 hours of spin-up time), while CASS simulates 12 hours with varying conditions according to the diurnal cycle. From each simulation, we use a snapshot every 2 minutes, to produce the database of cloud fields. The simulation evolved in 1 second increments, each yielding a different 3D spatial field which includes dozens of clouds. The data created takes 1.2 TB of memory. It took $\approx$ 5 days to generate, on an Intel Xeon Gold 5115 with 256 cores.

3. Noise Model

Section 2.1 in the main manuscript explains how a forward model derives a theoretic radiance field $I(\boldsymbol{x}, \boldsymbol{\omega})$. However, real-world radiance is in the form of a random photon flux, which obeys a Poissonian distribution. The photons are converted to discrete electric charges at the sensor. Furthermore, the sensor introduces noise due to various causes, according to its specifications. Let i^{e} be the expected photo-electron count of a pixel. At darkness and infinitesimal exposure time, readout noise has standard deviation ρ_{read}, in electrons. At temperature $\mathcal{T}$, the sensor dark current in electrons/sec is $D_{\mathcal{T}}$. The exposure time is Δt. The standard deviation of quantization noise in electrons is $\rho_{\mathrm{digit}} = g^{\mathrm{e}}/\sqrt{12}$, where g^{e} is the number of photo-electrons required to change a unit gray level [4]. Overall, in a pixel readout, in units of electrons, the noise has variance of approximately

$$V = i^{\mathrm{e}} + D_{\mathcal{T}}\Delta t + \rho_{\mathrm{read}}^2 + \rho_{\mathrm{digit}}^2 \ . \qquad (1)$$
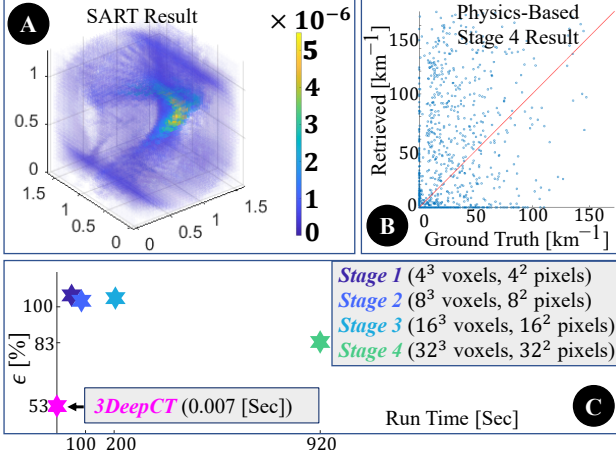
Figure 1. Recovering the cloud of Figs.3 of the main paper. (a) SART yields negligible initialization. (b) Scatter plots following coarse-to-fine analysis. (c) The error ϵ and run-time clock after each coarse-to-fine stage.

Our simulations of a perspective camera have a noise model based on the CMV4000 sensor [1]. The pixel size is $5.5 \times 5.5 \text{micron}^2$, $\rho_{\text{read}} = 13$ electrons, $D_{\mathcal{T}} = 125$ electrons/sec at 25°C, full well of a pixel is 13,500 electrons. It uses 10bit quantization, thus $g^{\text{e}} = 13,500/2^{10}$. The exposure time is adjustable according to the radiance that reached the camera such that the camera sensor reaches the full well.

4. Initialization Schemes

3DeepCT provides good initialization to the physics-based inverse rendering. We compare it with initializations used in prior art. As expected, applying SART yields negligible values, 6-8 orders of magnitude lower than the true β (See Fig. 1a above). Initialization by SART for cloud scattering-CT is as initialization by a null cloud. Fig. 1b,c here shows progression from $\times 8$ coarser grids (in each spatial coordinate) to full resolution by a coarse-to-fine [2, 3] process. It does not seem to save much time here. It is far slower and less effective than 3DeepCT.

5. 10 Viewpoints Geometry

The **10 Viewpoints** geometry, presented in Sec. 4.2 of the main manuscript is visualized in Fig. 2 herein. Recall from the main manuscript that this geometry uses 10 satellites which orbit at altitude of 600km. Nearest-neighbor satellites are 100km apart. They view the same field in off-nadir angles -46°, $\pm 34^{\circ}$, $\pm 26^{\circ}$, $\pm 18^{\circ}$, $\pm 9^{\circ}$ and 0°. Each carries a perspective camera. The field of view of each camera is 0.22°, corresponding to a ground footprint at the nadir of $1.6\text{km} \times 1.6\text{km}$, at 50 m ground resolution. Solar azimuth and zenith angles are 45° and 30°, respectively.
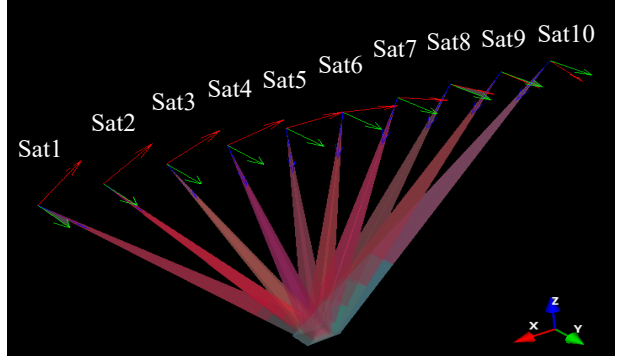

Figure 2. **10 Viewpoints** geometry; orbit at 600km altitude, nearest-neighbor satellites are 100km apart, remotely sensing clouds in the atmosphere from space.

Model	ϵ	δ	Time [millisec]
U-Net	91±12%	38±16%	7 ± 1
ResNet	94.5±12%	43±20%	5 ± 0.7

Table 1. Summary of test results: Mean $\pm$ standard deviation (std) of U-Net and ResNet models, discussed in Sec. 3.3 of the main manuscript. Equations for ϵ and δ are in Eq. (8) of the main manuscript.

6. Ablation Studies

The architecture of 3DeepCT, its components and hyperparameters were carefully examined and tuned by ablation studies. For example, Fig. 3 herein shows how the NN depth and batch normalization affect convergence, and Fig. 4 herein shows the training and evaluation loss of 3DeepCT compared to U-Net and ResNet architectures. Table 1 herein shows the summary results of these two architectures.

Space carving ablation studies were also examined. Without space-carving, we obtain $\epsilon = 85 \pm 14\%$ and $\delta = 9.5 \pm 18\%$ when using the Model 1 described in Table 2 in the main manuscript. Space-carving improved the results, as seen in Fig. 5 herein.

7. AirMSPI Cross Validation Results

Fig. 6 shows the perspective nadir view which was left out of the 3DeepCT input. 3DeepCT recovered the volumetric cloud without using the nadir view of the cloud. We rendered the recovered result in the nadir direction and perform cross validation between this rendered view and the original left-out image.

8. Simulated Inference Results

Fig. 3 in the main manuscript shows the 3D reconstructions of cloud extinction. Fig. 4 of the main manuscript shows the scatter plots of recovery results related to methods A and B, described in Sec. 5 of the main manuscript.
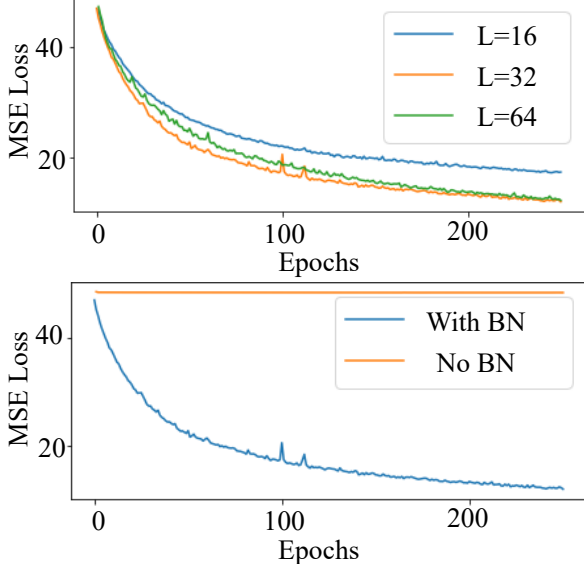
Figure 3. Hyper-parameters and ablation experiment: [First row] Tuning experiment of hyper-parameter L ; [Second row] Ablation study of the batch normalization (BN) module of the NN.
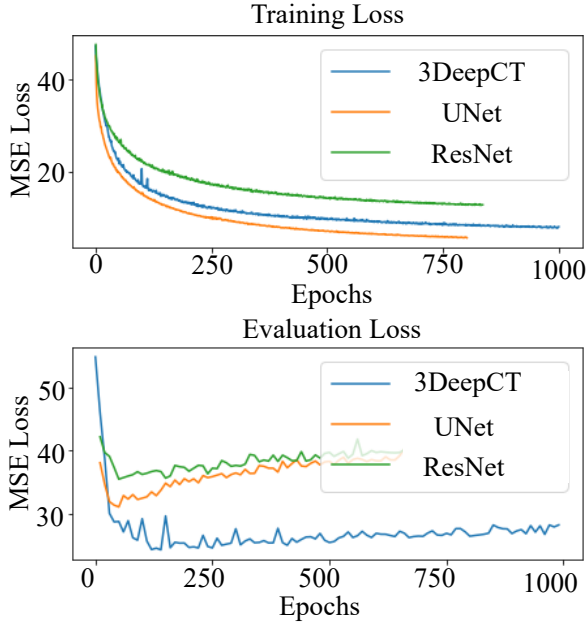

Figure 4. Training and evaluation loss of 3DeepCT, compared to those of U-Net and ResNet architectures.

Fig. 7 herein relates to the additional methods C and D. Fig. 8 herein shows the *subset of seven* clouds. In the main manuscript, inference results of our Model 1 are demonstrated on one cloud out of the *subset of seven* clouds. This cloud is shown in Fig. 8(a) herein. Three additional clouds are shown in Figures 8(b)-(d) herein. Figures 9(a)-(c) herein show their respective recovery, using the four approaches
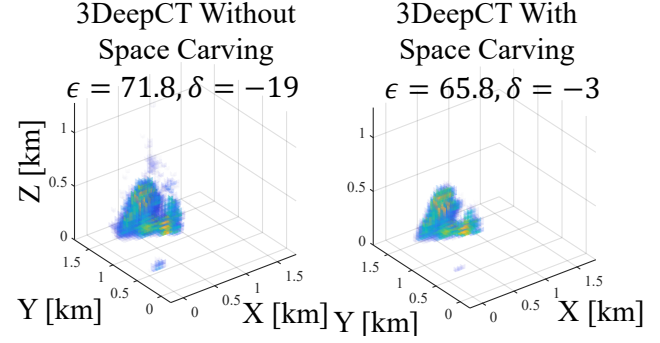

Figure 5. Recovery results: [Left] 3DeepCT raw recovery result without space-carving, [Right] 3DeepCT recovery results using the space-carving mask. The results show the improvement when using space-carving.

described in Sec. 5 of the main manuscript.

References

[1] CMV4000 sensor. https://ams.com/cmv4000. 2

[2] Ioannis Gkioulekas, Anat Levin, and Todd Zickler. An evaluation of computational imaging techniques for heterogeneous inverse scattering. In *ECCV*, pages 685–701. Springer, 2016. 2

[3] Tamar Loeub, Aviad Levis, Vadim Holodovsky, and Yoav Y Schechner. Monotonicity prior for cloud tomography. In *Proceedings of ECCV*, pages 24–29, 2020. 2

[4] Yoav Y. Schechner, Shree Nayar, and Peter N. Belhumeur. Multiplexing for optimal lighting. *PAMI*, 29:1339–1354, 2007. 1

[5] A Pier Siebesma, Christopher S Bretherton, Andrew Brown, Andreas Chlond, Joan Cuxart, Peter G Duynkerke, Hongli Jiang, Marat Khairoutdinov, David Lewellen, Chin-Hoh Moeng, et al. A large eddy simulation intercomparison study of shallow cumulus convection. *Journal of the Atmospheric Sciences*, 60(10):1201–1219, 2003. 1

[6] Yunyan Zhang, Stephen A Klein, Jiwen Fan, Arunchandra S Chandra, Pavlos Kollias, Shaocheng Xie, and Shuaiqi Tang. Large-eddy simulation of shallow cumulus over land: A composite case based on ARM long-term observations at its Southern Great Plains site. *Journal of the Atmospheric Sciences*, 74(10):3229–3251, 2017. 1
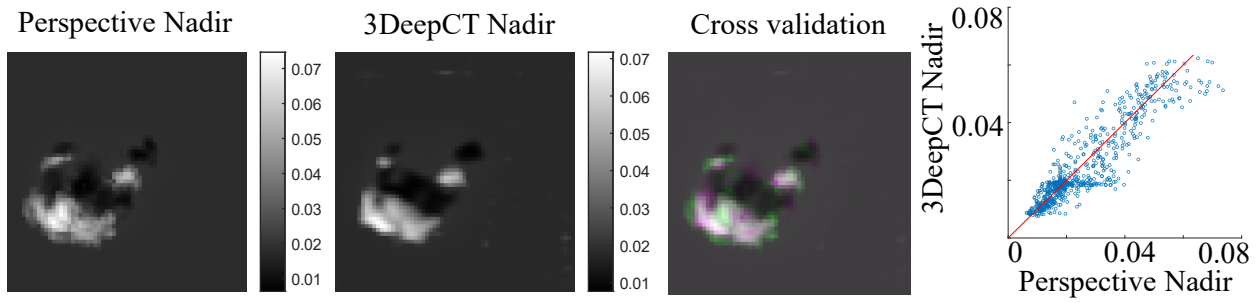
Figure 6. From left to right: Left-out perspective nadir view; Nadir view rendered from the results of 3DeepCT; Cross validation of the nadir view - the physics-based inverse scattering left-out image is displayed in the green channel; the 3DeepCT result is displayed in magenta; scatter-plot of the cross-validation pixel values of radiance.
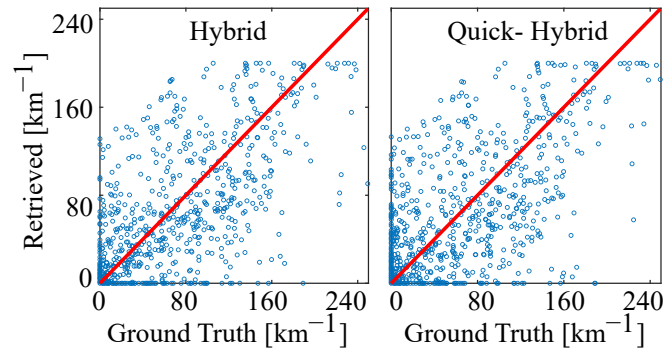

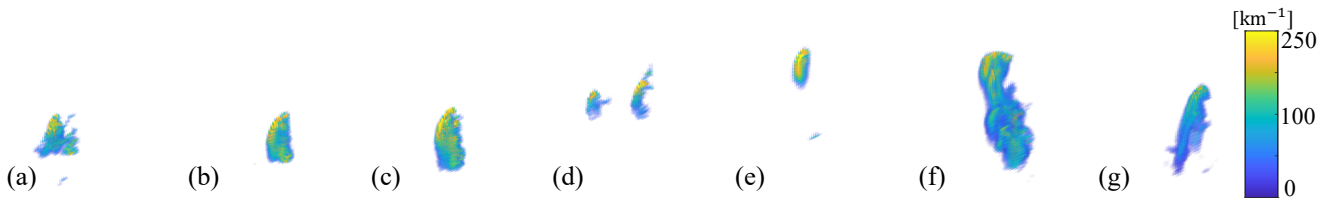Figure 7. Recovering the cloud of Fig. 3 of the main manuscript.

Figure 8. The *Subset of seven* clouds.

(a) 3D extinction recovery of cloud (b) from Fig. 8 herein.

(b) 3D extinction recovery of cloud (c) from Fig. 8 herein.

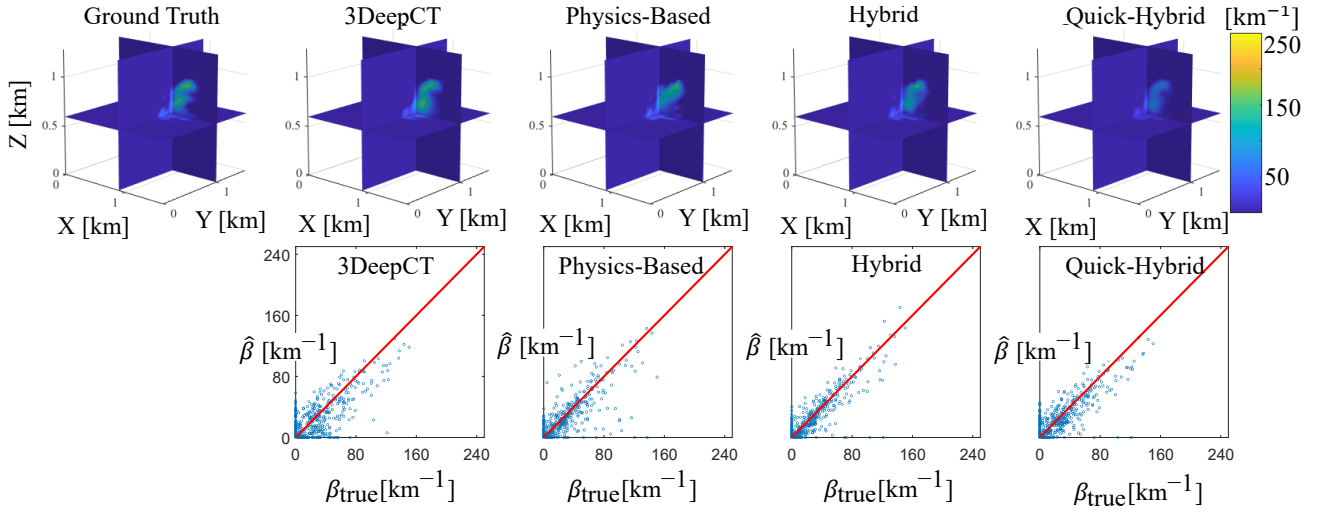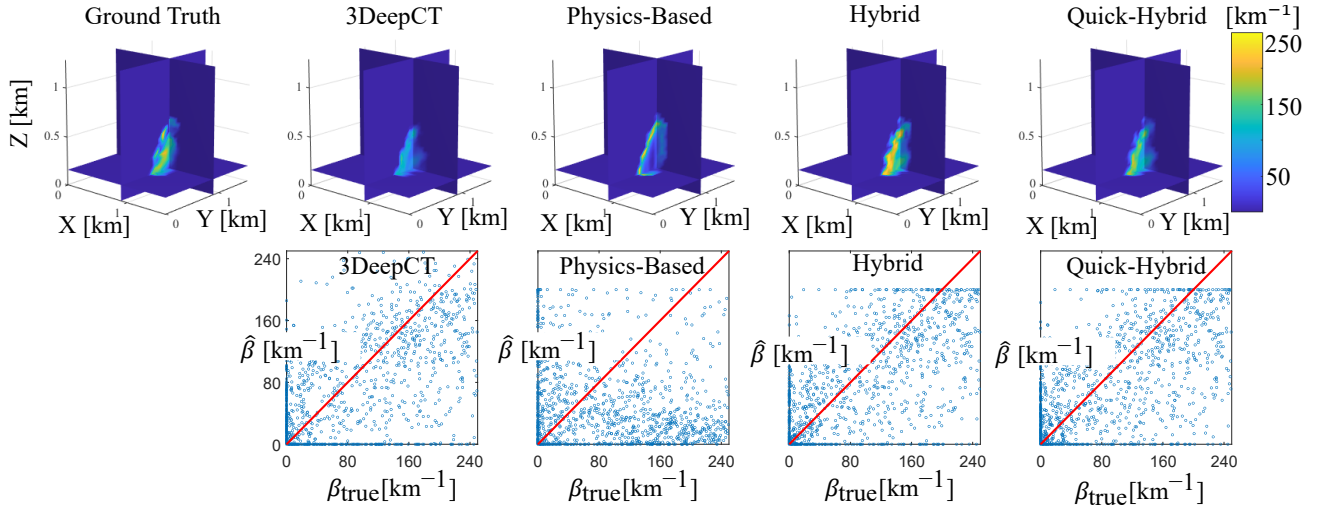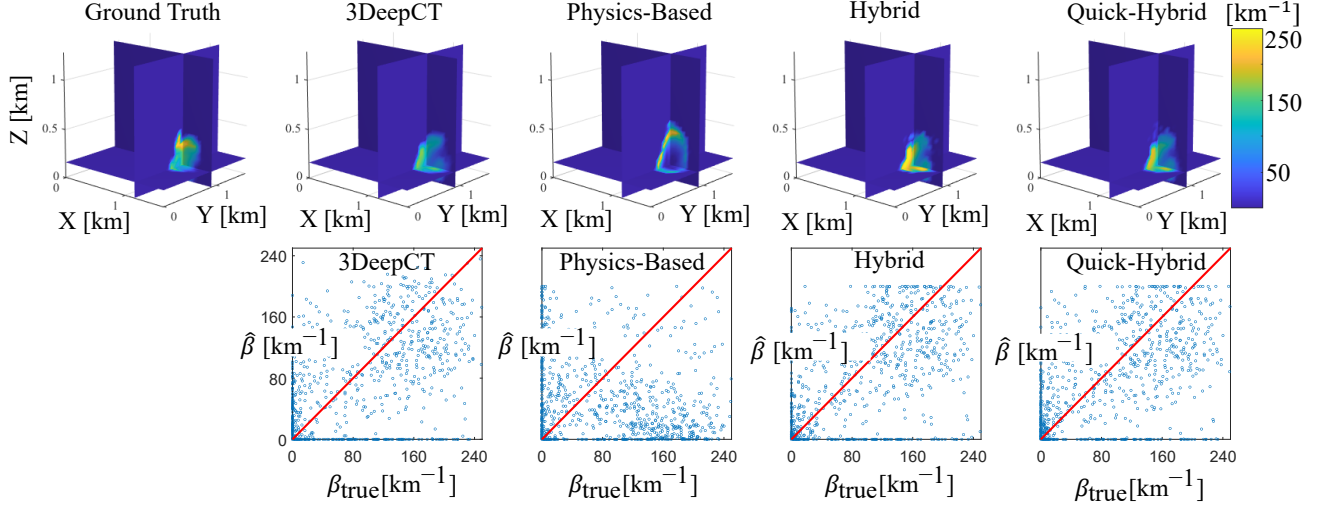(c) 3D extinction recovery of cloud (d) from Fig. 8 herein.

Figure 9. [First row] From left to right: 3D ground-truth extinction of the cloud; 3D reconstructed extinction using the four methods mentioned in Sec. 5 of the main manuscript. [Second row] Scatter plots of the recovery results. The plots relate to the four methods at the top row. The reconstructed 3D extinction is $\hat{\beta}$. The red line represents ideal reconstruction, where $\hat{\beta} = \beta_{\text{true}}$.