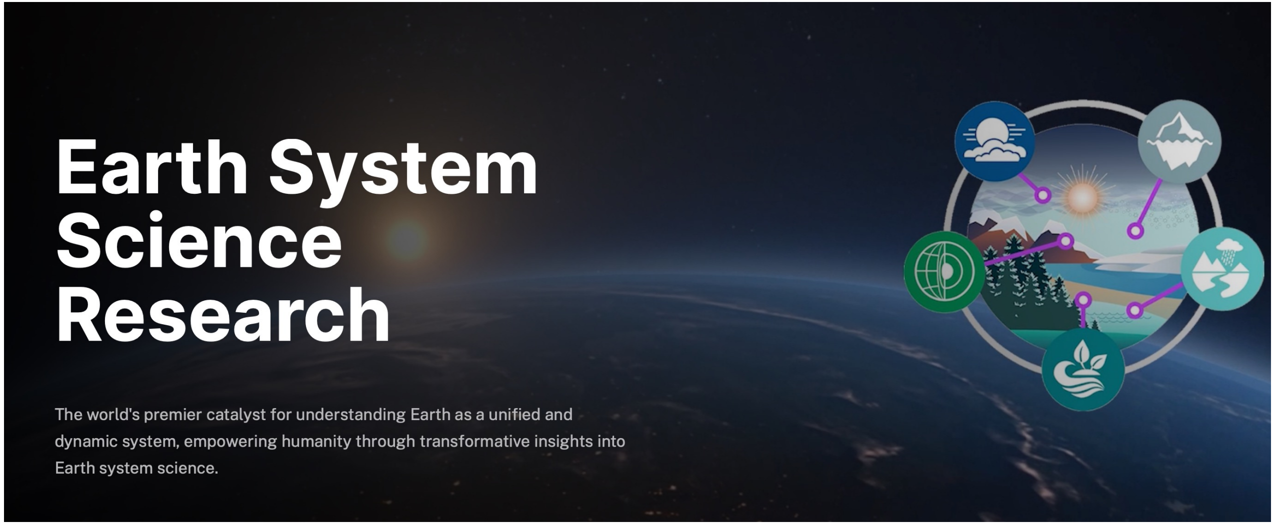
A Deep Neural Network Approach to Atmospheric Lidar Denoising and Impacts on Validation for Data Assimilation

Willem J. Marais, **Danica B. Fliss Ehmer**, Robert E. Holz,
Mark A. Vaughan, Charles R. Trepte,
Johnathan W. Hair, Richard A. Ferrare,
Chris A. Hostetler, Robert D. Nowak

Slides from Bob Holz:

Update on NASA Atmospheric SIPS

Thoughts informed by 2023 Continuity Workshop, recent AtmoSphere community meeting, and forthcoming Foundational Data Products project



Atmosphere

Atmosphere fosters research and innovation targeting the dynamics and thermodynamics of the atmosphere, its physical and chemical composition, and the interdependent impacts that these have on the Earth's radiative balance, air quality, and weather.

Thoughts informed by 2023 Continuity Workshop, recent AtmoSphere community meeting, and forthcoming Foundational Data Products project

FOUNDATIONAL DATA PRODUCTS (FDP)

The FDP project focuses on stewardship and continuity of Earth system essential observations and measurements. This project will evolve the processing and archival functions historically performed by the Science Investigator-led Processing Systems, a core function of the Earth sciences division and community partners, ensuring transparency, traceability, and sustained availability of authoritative data records. It will address mission transitions, algorithm development and evolution, and reprocessing needs, while positioning NASA to support future scientific and applied uses of Earth system data. The project leverages existing capabilities among NASA centers and elsewhere, while defining a forward-looking processing architecture that supports long-term data integrity, efficient delivery, and technical adaptability. The project will coordinate closely with the ESDIS project, following defined interfaces to ensure consistency, interoperability, and alignment across data production, user support, and data access services.

The Foundational Data Products (FDP) project ensures that NASA continues to advance high-quality standardized data necessary for models, research, and commercial applications. FDP also provides a sustainable path for innovative data products that is independent of a single data product/source. This is critical to ensure smooth and cost-effective user transitions between data from different missions as new missions launch, and to maximize value of Earth Science data as advanced instruments replace historical ones.

nasa-fiscal-year-2027-full-budget-request.pdf

Science team and supported products

Product Short Name	Product Description	ST Lead	Distribution
AERDB_L2_VIIRS_[SNPP|NOAA20|21] **AERDB_D3_VIIRS_[SNPP|NOAA20|21]** **AERDB_M3_VIIRS_[SNPP|NOAA20|21]**	**Deep Blue Aerosol (day only)** **Standard and NRT**	Jaehwa Lee (NASA GSFC)	LAADS (standard) LANCE (NRT)
AERDT_L2_VIIRS_[SNPP|NOAA20|21]	**Dark Target Aerosol (day only)** **Standard and NRT**	Robert Levy (NASA GSFC)	LAADS (standard) LANCE (NRT)
CLDMSK_L2_VIIRS_[SNPP|NOAA20|21] **CLDMSK_L2_MODIS_Aqua**	**Continuity Cloud Mask (day/night)** **Standard and NRT**	Bob Holz (SSEC UW)	LAADS (standard) LANCE (NRT)
CLDMSK_L2_VIIRS_[SNPP|NOAA20|21] CLDMSK_L2_MODIS_Aqua	Imager/Sounder Continuity Cloud Mask (day/night) Standard	**Bob Holz (SSEC UW)**	**LAADS**
CLDPROP_L2_VIIRS_[SNPP|NOAA20] CLDPROP_D3_VIIRS_[SNPP|NOAA20] CLDPROP_M3_VIIRS_[SNPP|NOAA20] CLDPROP_L2_MO CLDPROP_D3_MODIS_Aqua CLDPROP_M3_MODIS_Aqua	Continuity Cloud Properties (day/night) Standard	Kerry Meyer (NASA GSFC)	LAADS
CLDPROP_L2_VIIRS_[SNPP|NOAA20] CLDMSK_L2_MODIS_Aqua	Imager/Sounder Continuity Cloud Properties (day/night) Standard	**Kerry Meyer (NASA GSFC)**	**LAADS**
SNDRSNCrISL1BIMG (SNPP) SNDRJ1CrISL1BIMG (NOAA-20)	Collocated VIIRS Level 1 and cloud mask statistical summary((IMG)	Joe Tayler (SSEC UW)	GESDISC

Back to the talk

The UW Madison team

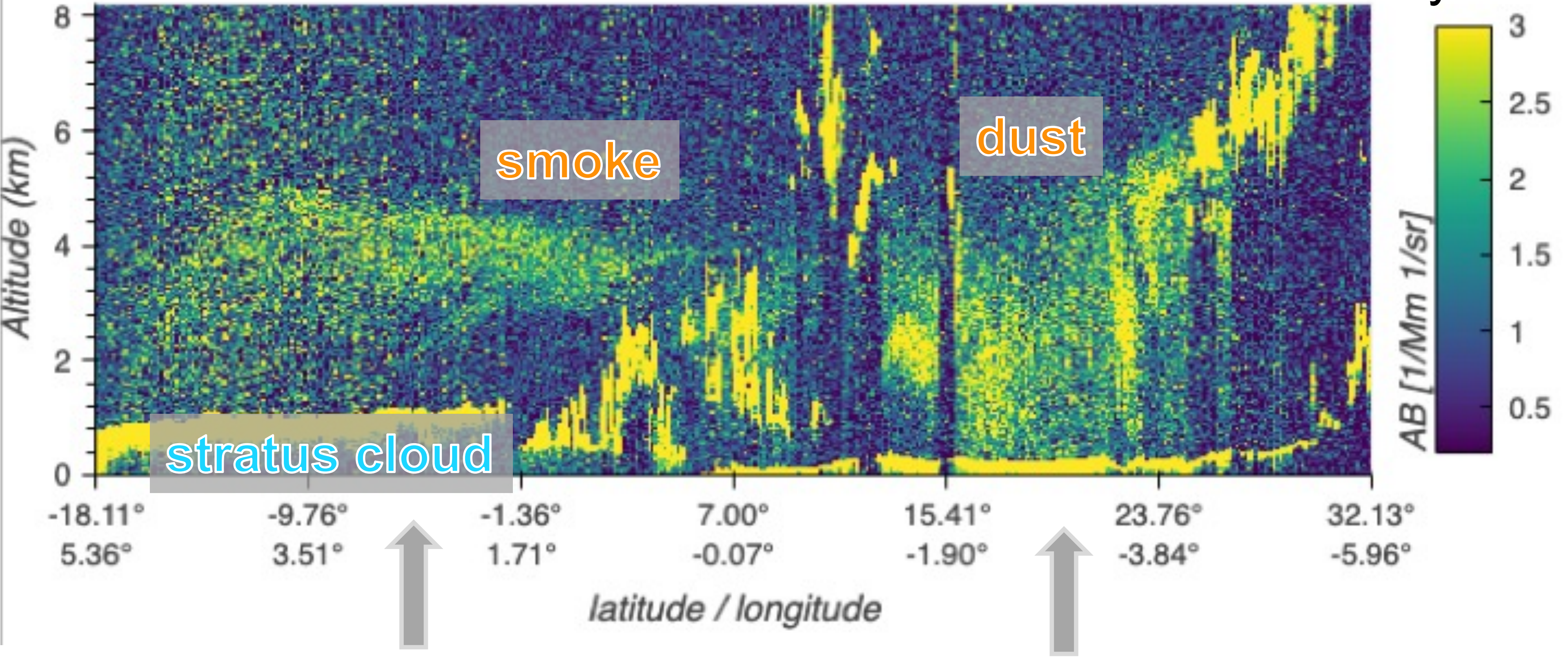
Danica Fliss Ehmer
who did all the work
(UW Madison ECE)
Graduating this Spring 2026

Robert Nowak
Danica's academic advisor
(UW Madison ECE)

Bob Holz
The visionary
(UW Madison SSEC)

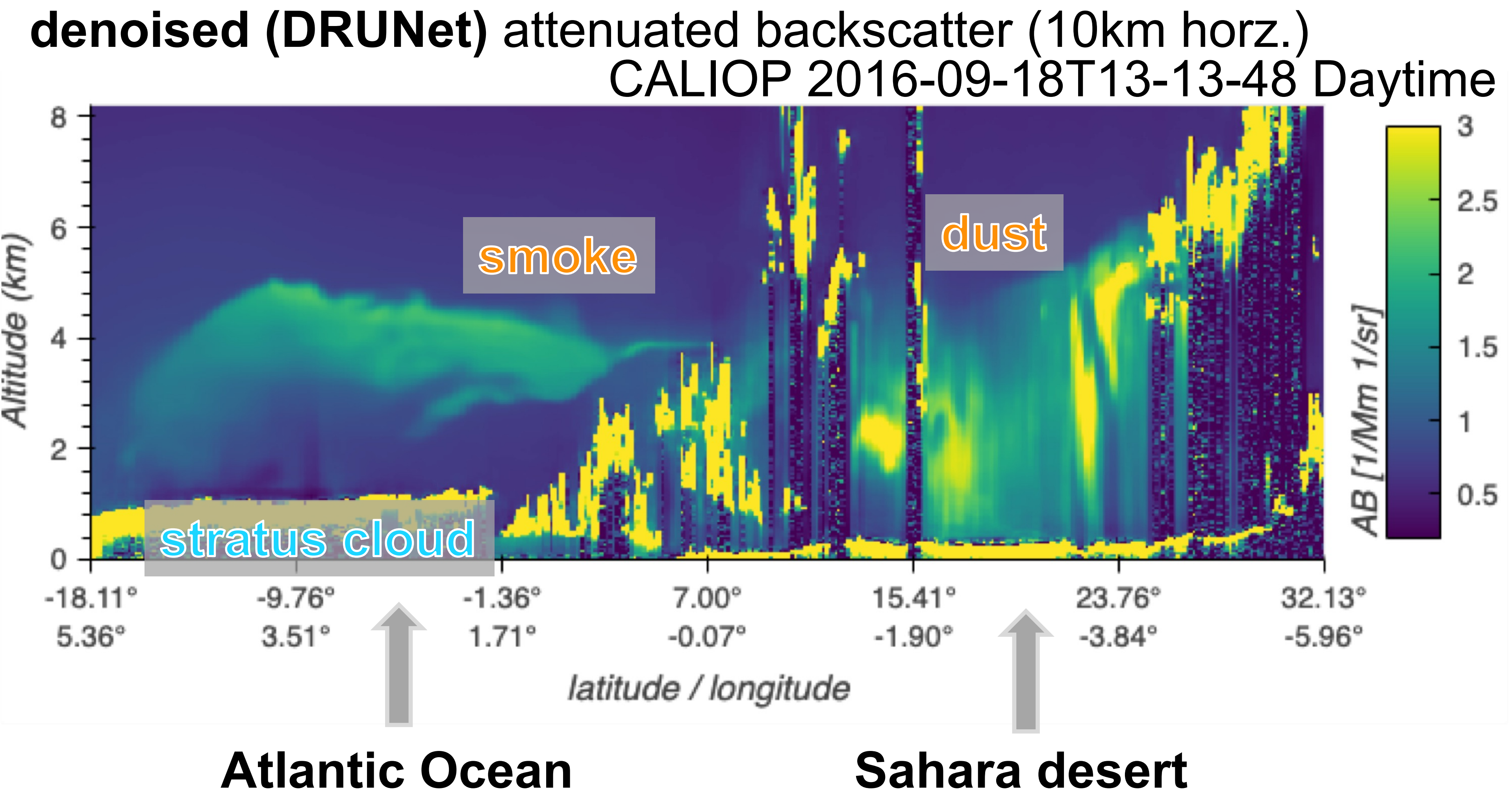
standard attenuated backscatter (10km horz.)

CALIOP 2016-09-18T13-13-48 Daytime

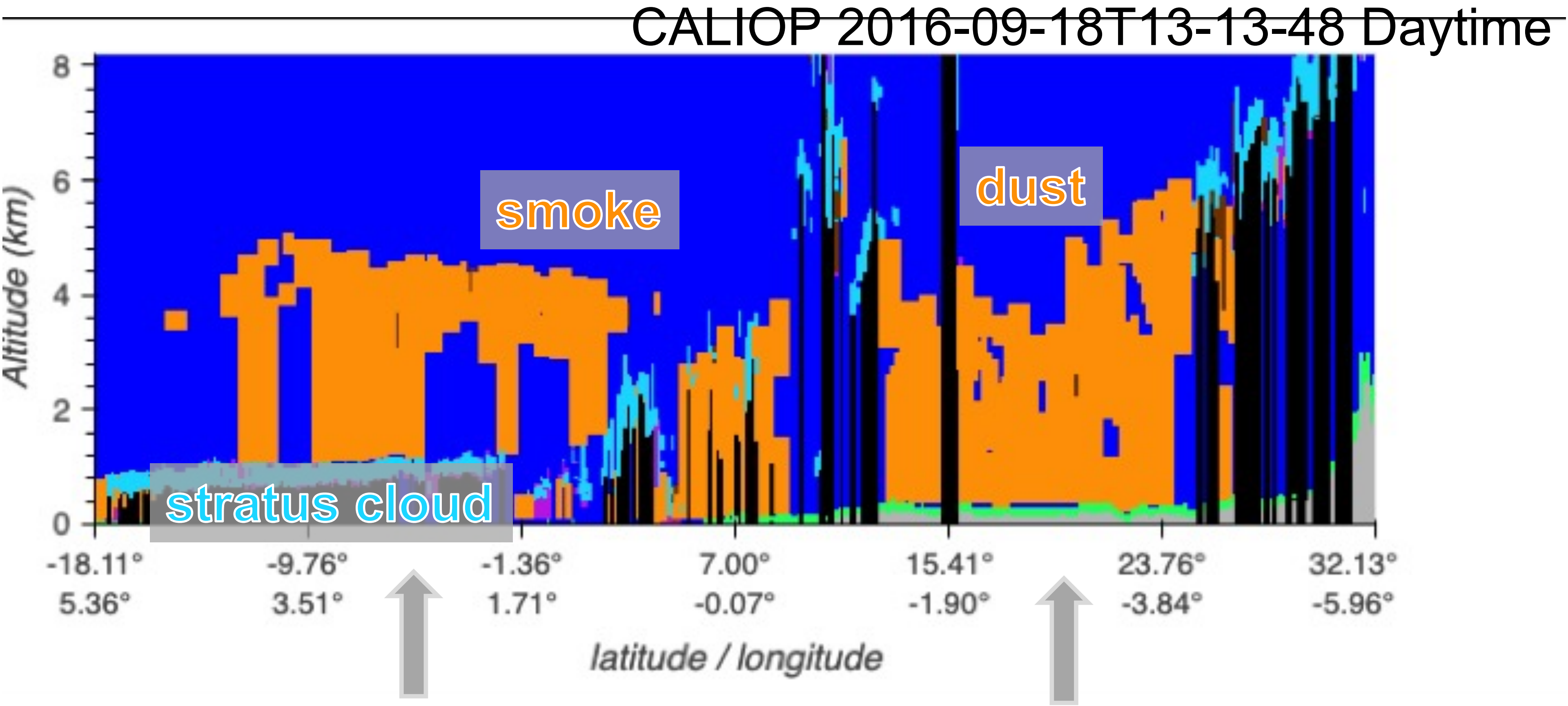


Atlantic Ocean

Sahara desert


denoised (DRUNet) attenuated backscatter (10km horz.)
CALIOP 2016-09-18T13-13-48 Daytime
Altitude (km)
0
2
4
6
8
smoke
dust
stratus cloud
AB [1/Mm 1/sr]
0.5
1
1.5
2
2.5
3
-18.11° 5.36°
-9.76° 3.51°
-1.36° 1.71°
7.00° -0.07°
15.41° -1.90°
23.76° -3.84°
32.13° -5.96°
latitude / longitude
Atlantic Ocean
Sahara desert

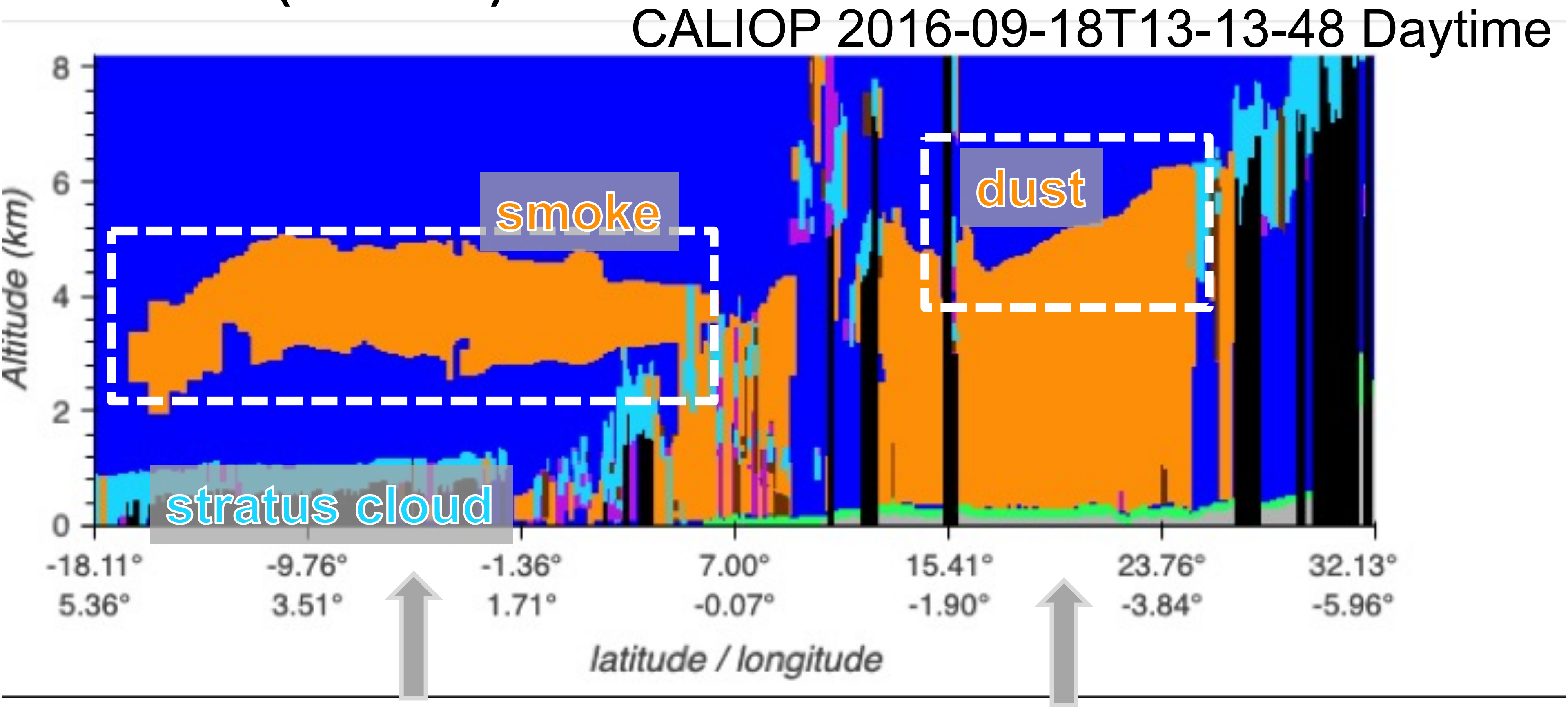
standard vertical feature mask



Atlantic Ocean

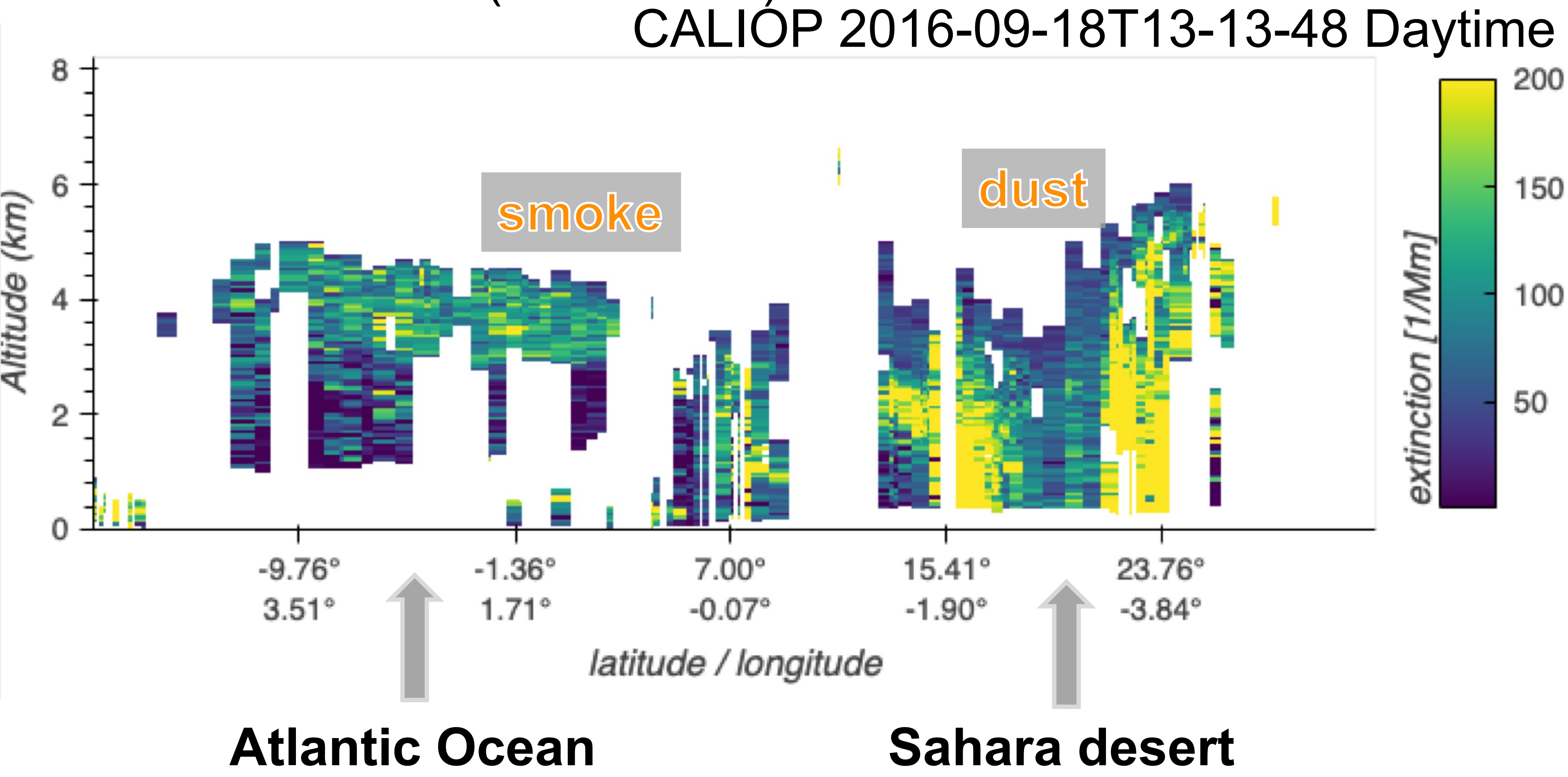
Sahara desert

denoised (DRUNet) vertical feature mask



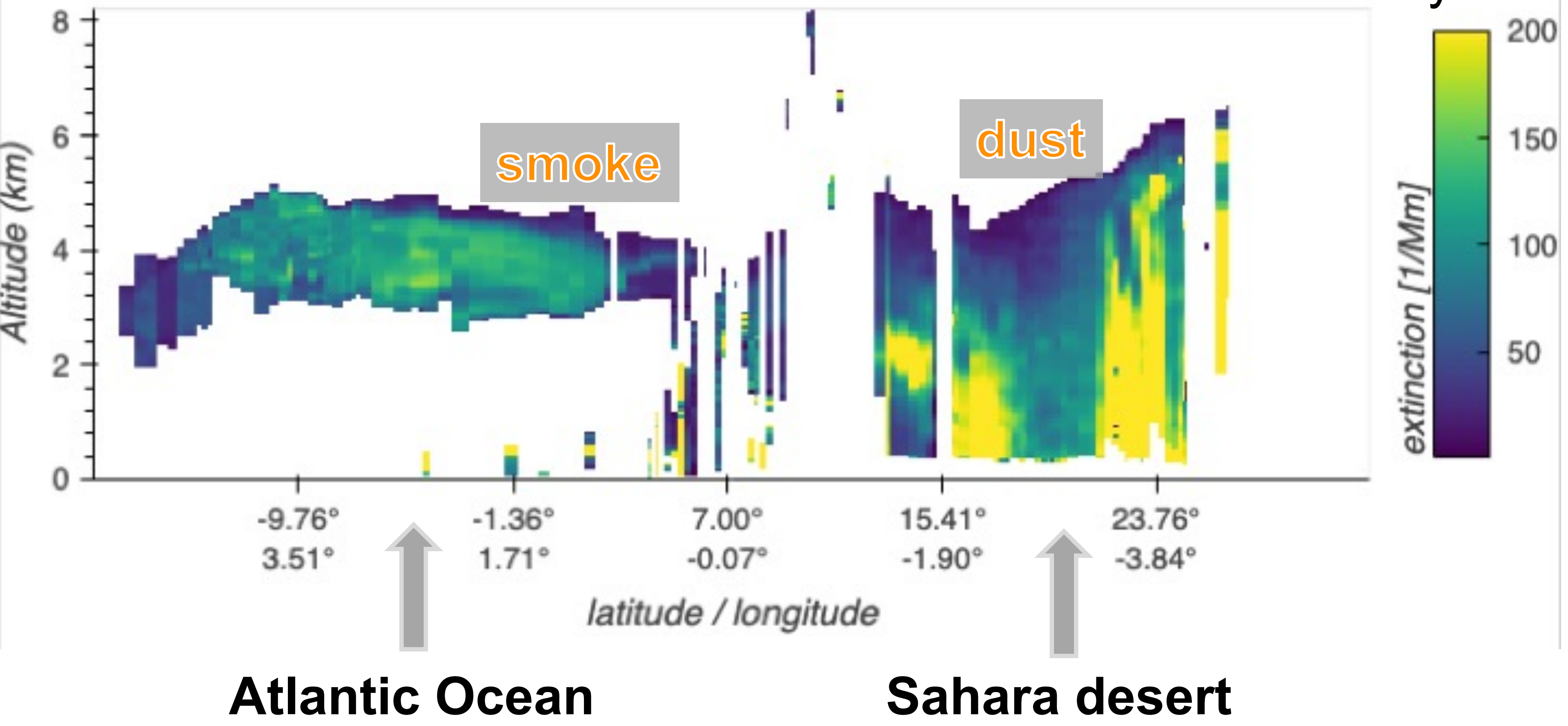
Atlantic Ocean **Sahara desert**

standard extinction (10km horz.)



denoised (DRUNet) extinction (10km horz.)

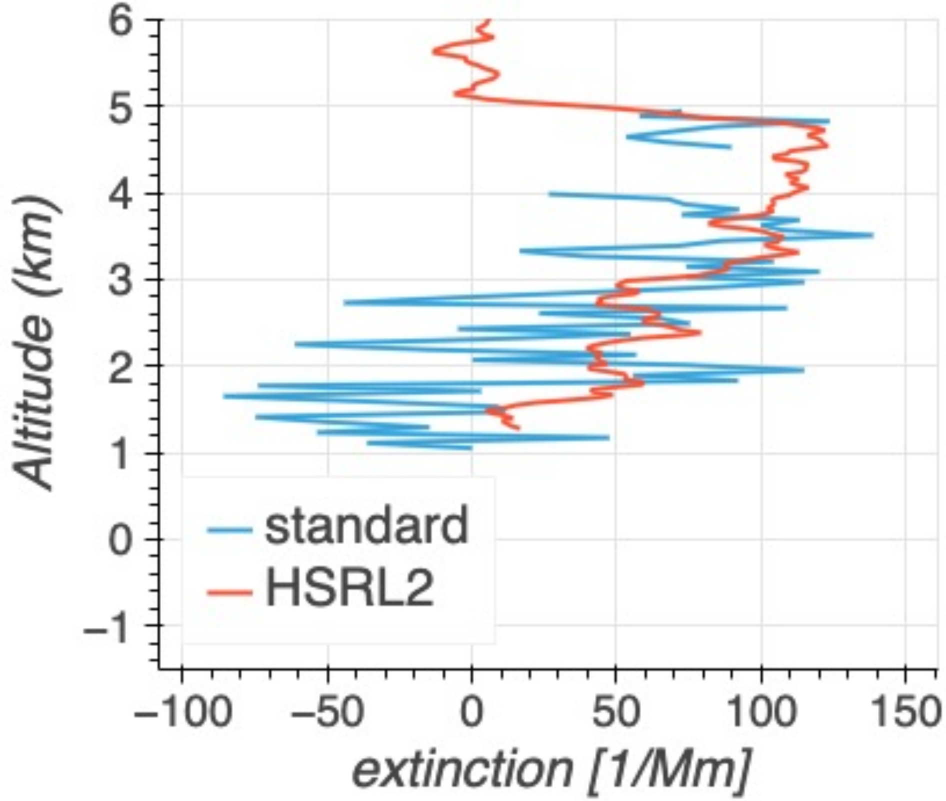
CALIOP 2016-09-18T13-13-48 Daytime



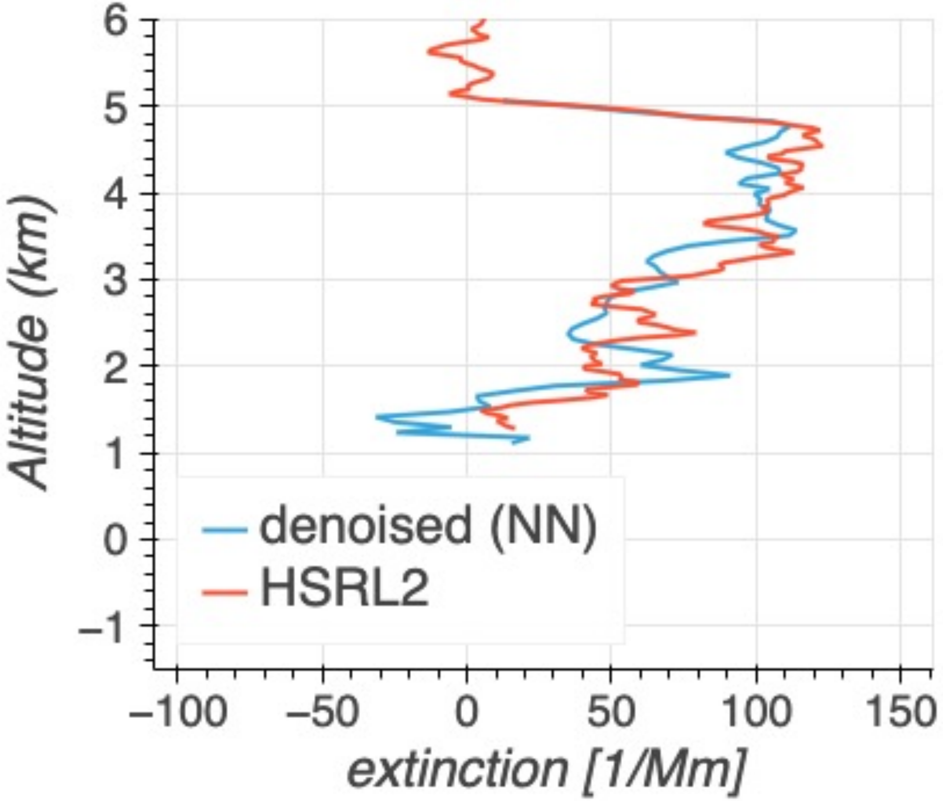
Profile plots at co-incident time

CALIOP 2016-09-18T13-13-48 Daytime

Extinction retrieval with no denoising at 10km horizontal resolution



Extinction retrieval after denoising via DRUNet at 10km horizontal resolution



Overview of the denoising framework

- The denoising NN method (DRUNet) we employ has not been trained on lidar imagery at all
- DRUNet is instead an off-the-shelve pretrained NN
  - that has been trained image web-images
  - to denoise Gaussian noisy images
  - and not photon limited images such as space-based lidar
- The challenge therefore is to transform the lidar images so that an off-the-shelve NN denoising method can be use

General Assumptions of Gaussian denoisers

1. **Gaussian Noise:** Assumes the distribution of the noise is additive Gaussian Noise, and an approximation of the noise standard deviation is known prior
2.
3.

General Assumptions of Gaussian denoisers

1. **Gaussian Noise:** Assumes the distribution of the noise is additive Gaussian Noise, and an approximation of the noise standard deviation is known prior
2. **Collective Info:** Uses information from neighboring pixels to promote continuity and smoothness
3.

General Assumptions of Gaussian denoisers

1. **Gaussian Noise:** Assumes the distribution of the noise is additive Gaussian Noise, and an approximation of the noise standard deviation is known prior
2. **Collective Info:** Uses information from neighboring pixels to promote continuity and smoothness
3. **Smooth:** Fine-tuned to natural images which have a defining characteristic of being mostly smooth with a few sharp edges

CALIOP Data

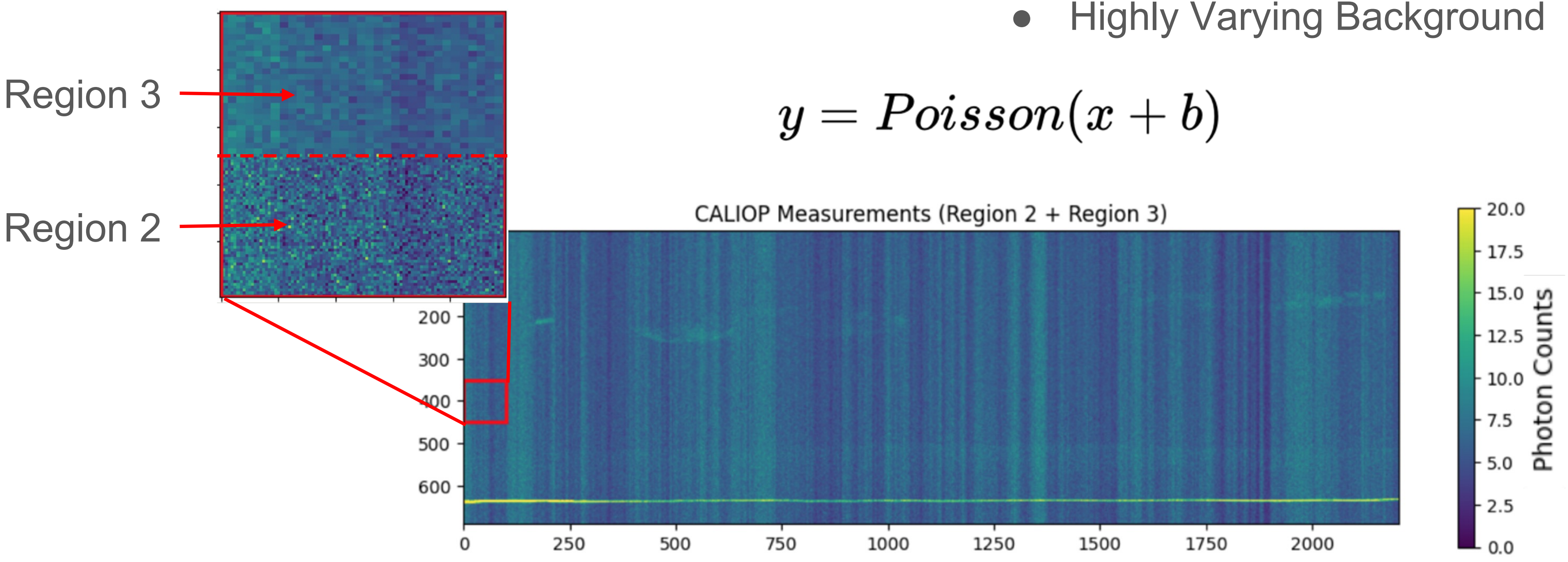
Gaussian denoisers assumptions

- ~~Gaussian Noise~~
- ~~Collective Info~~
- ~~Smooth~~

Assumptions violated by atmospheric lidar

- Poisson like Noise
- Resolution Change Between Altitude Regions
- Highly Varying Background

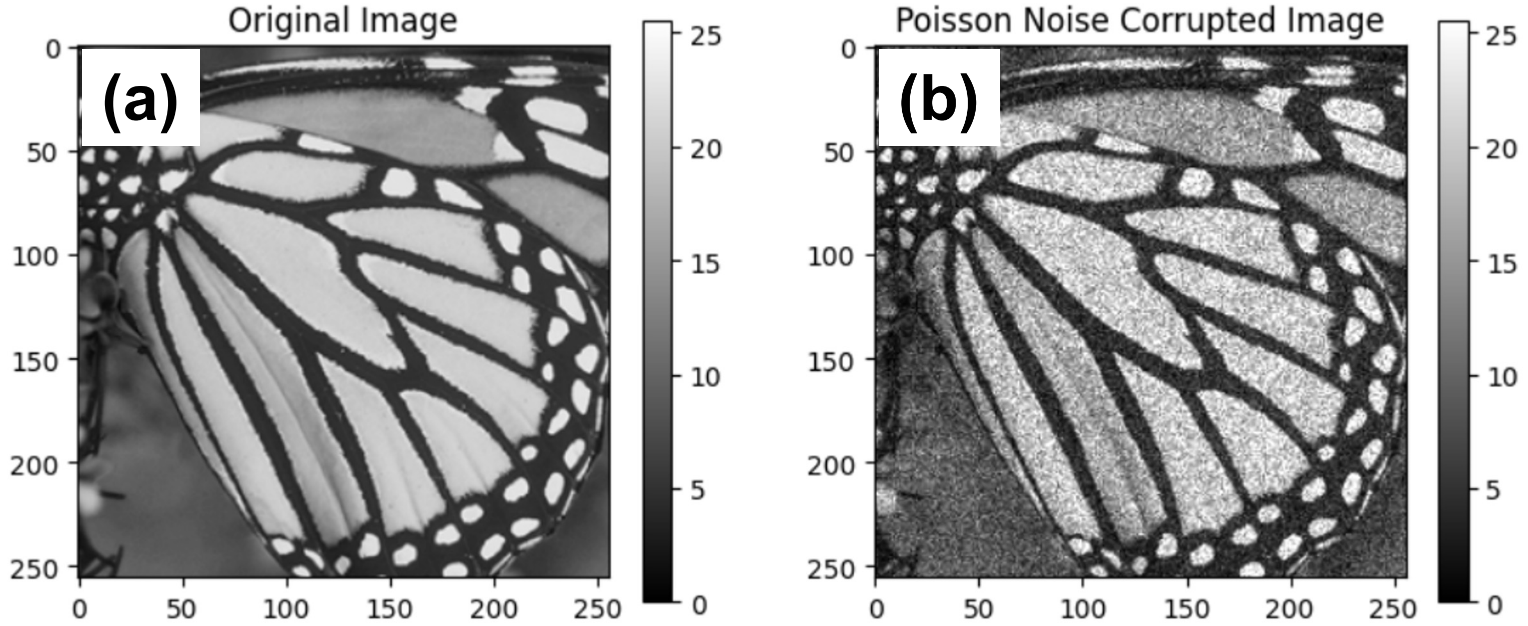
$$y = Poisson(x + b)$$



Challenge 1: Poisson like noise

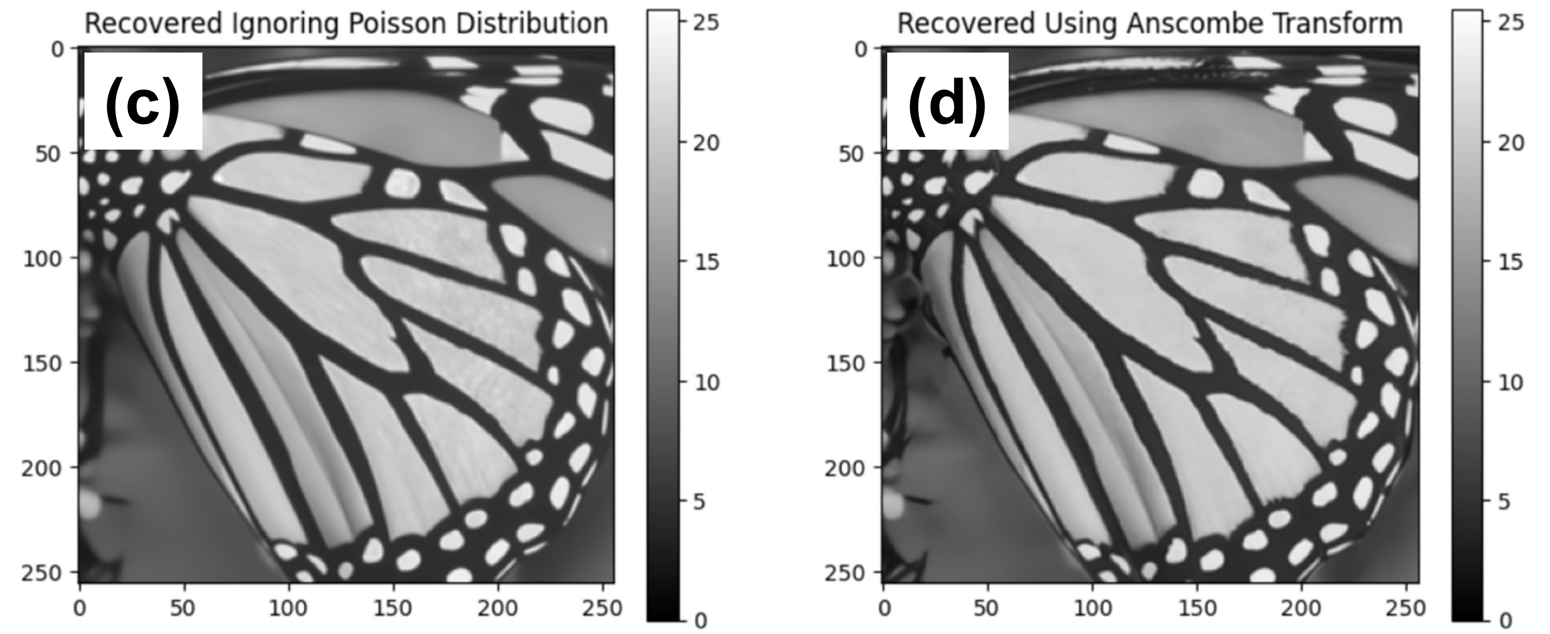
Poisson to Gaussian Noise

Comparing denoising with and without variance stabilization transformation



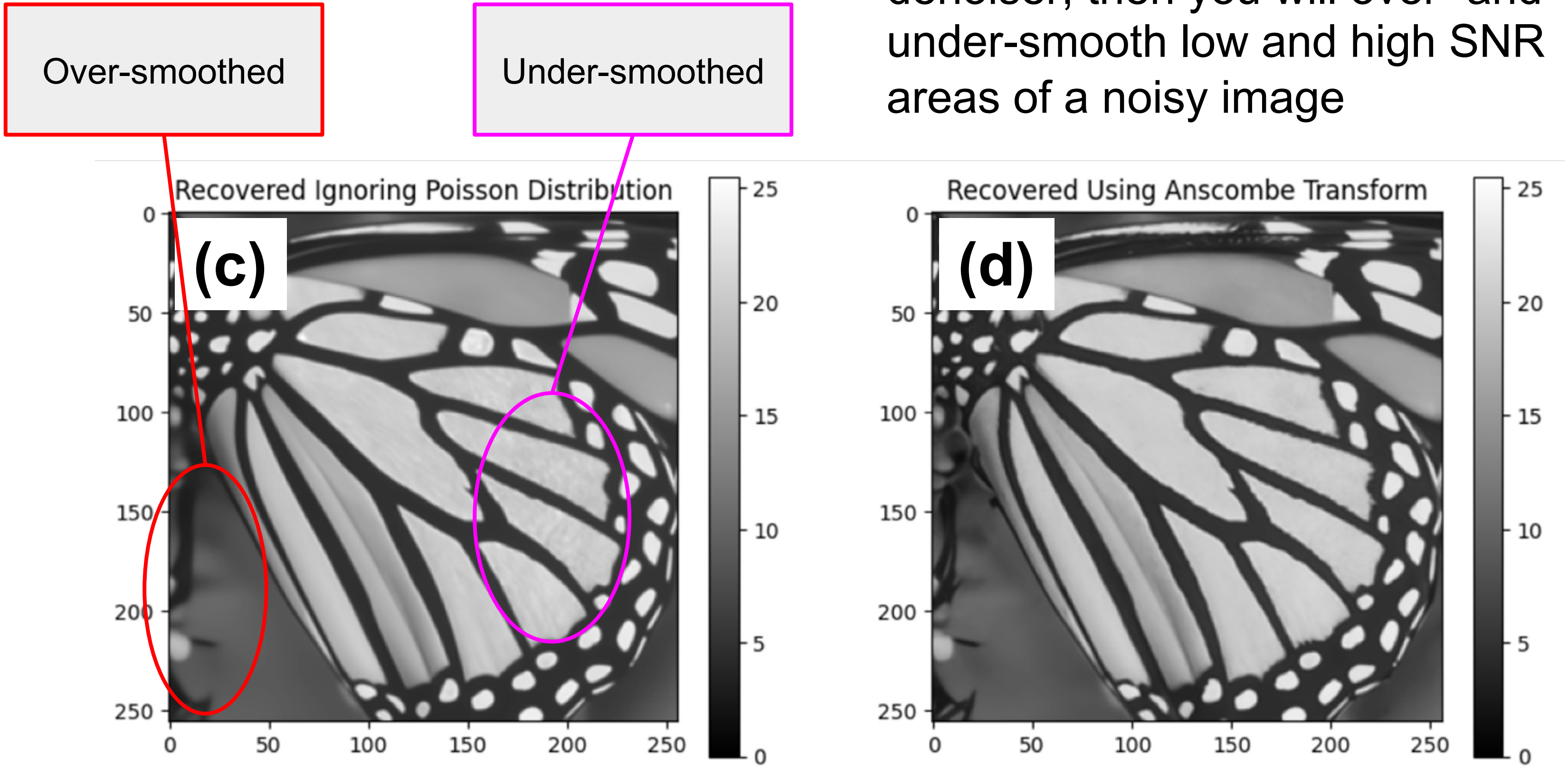
Poisson to Gaussian Noise

The variance stabilization transformation used here is known as Anscombe transform



Poisson to Gaussian Noise

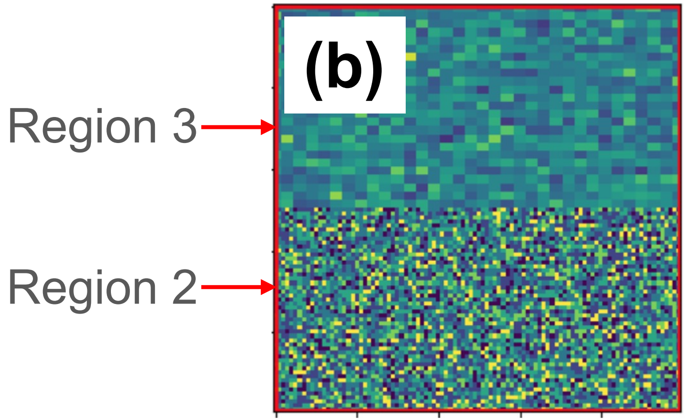
If you denoise Poisson noisy images with a Gaussian denoiser, then you will over- and under-smooth low and high SNR areas of a noisy image

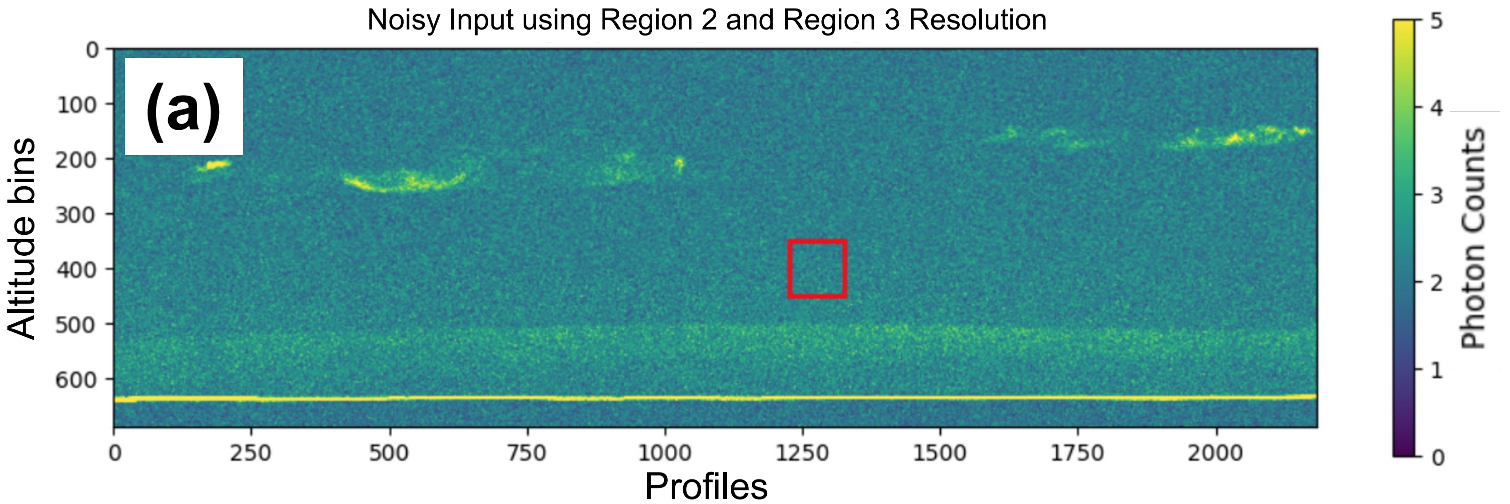


Challenge 2: CALIOP resolution boundaries

Resolution Boundaries

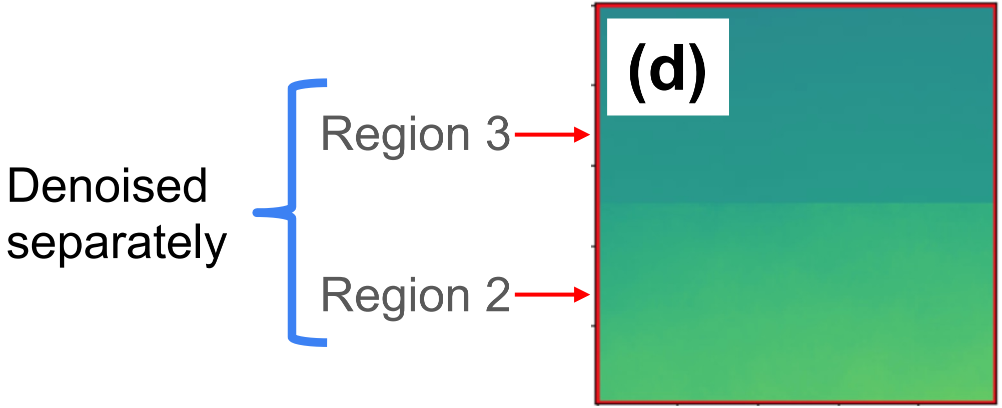
- Denoisers do not handle changes in resolution
- Denoising the regions separately causes artifacts at the boundaries

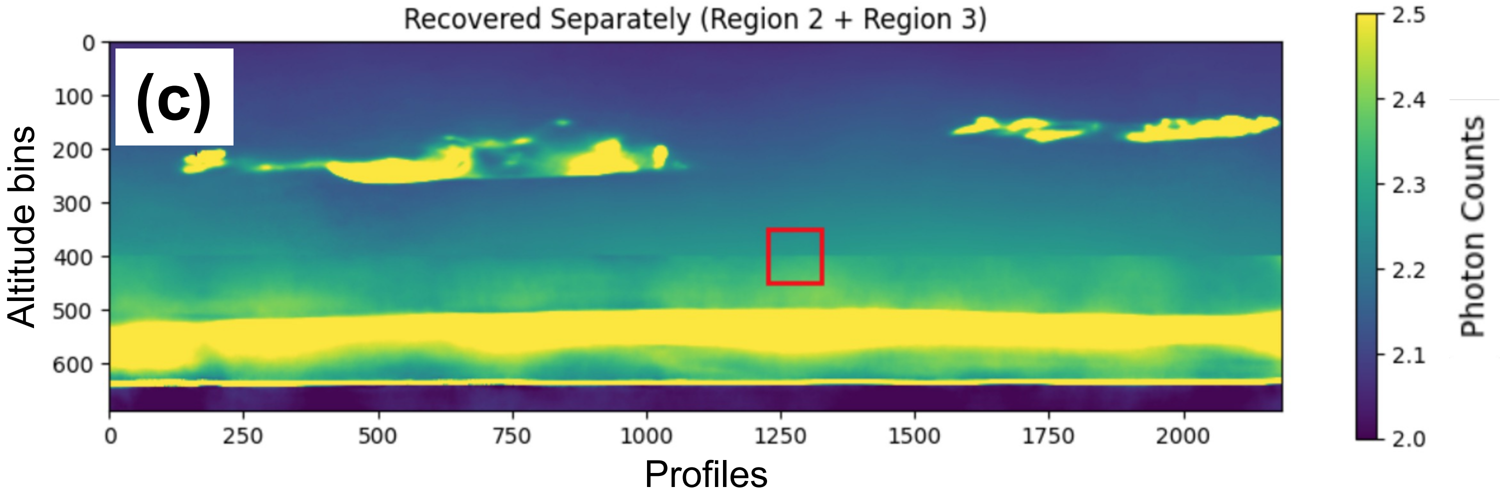




Resolution Boundaries

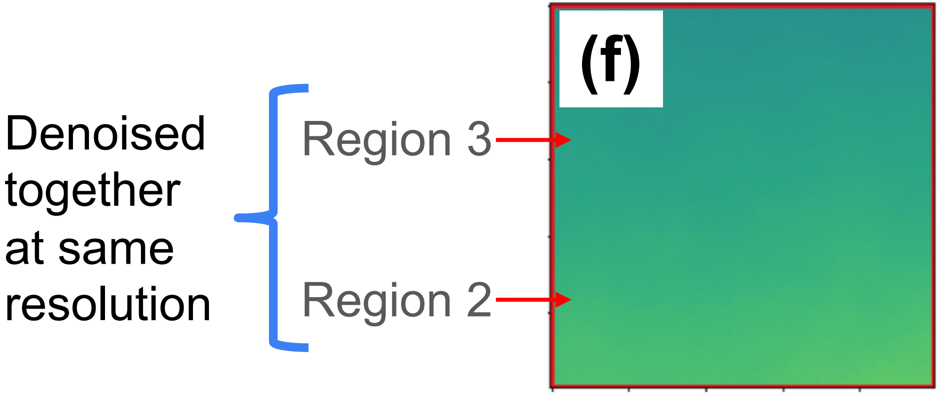
- Denoisers cannot promote “smoothness” across resolution boundaries

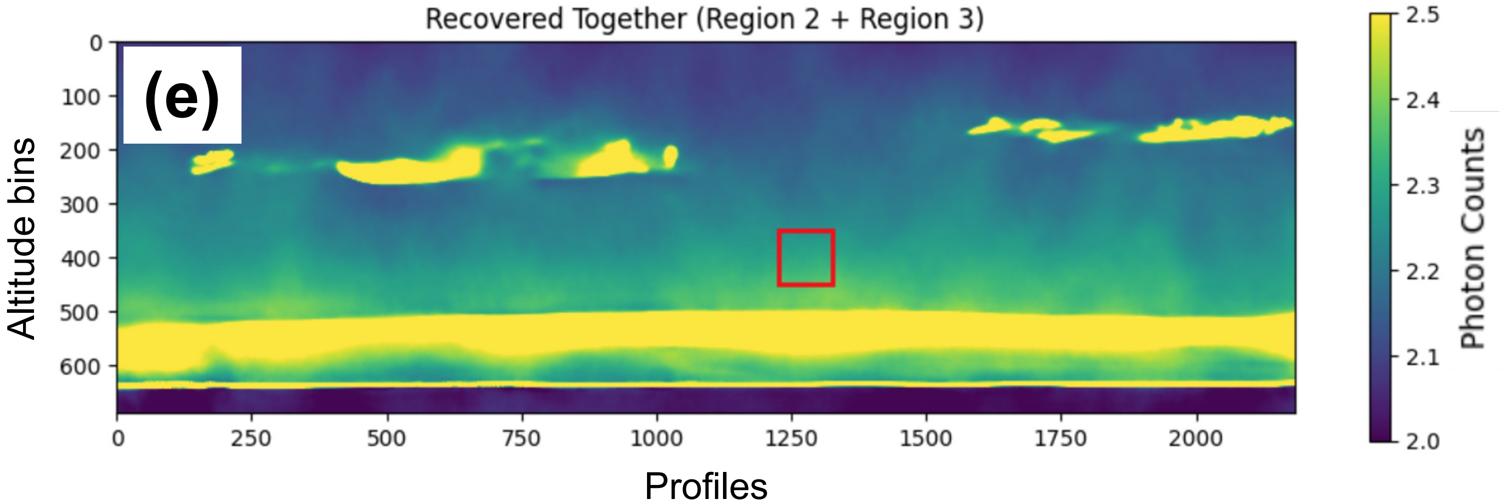




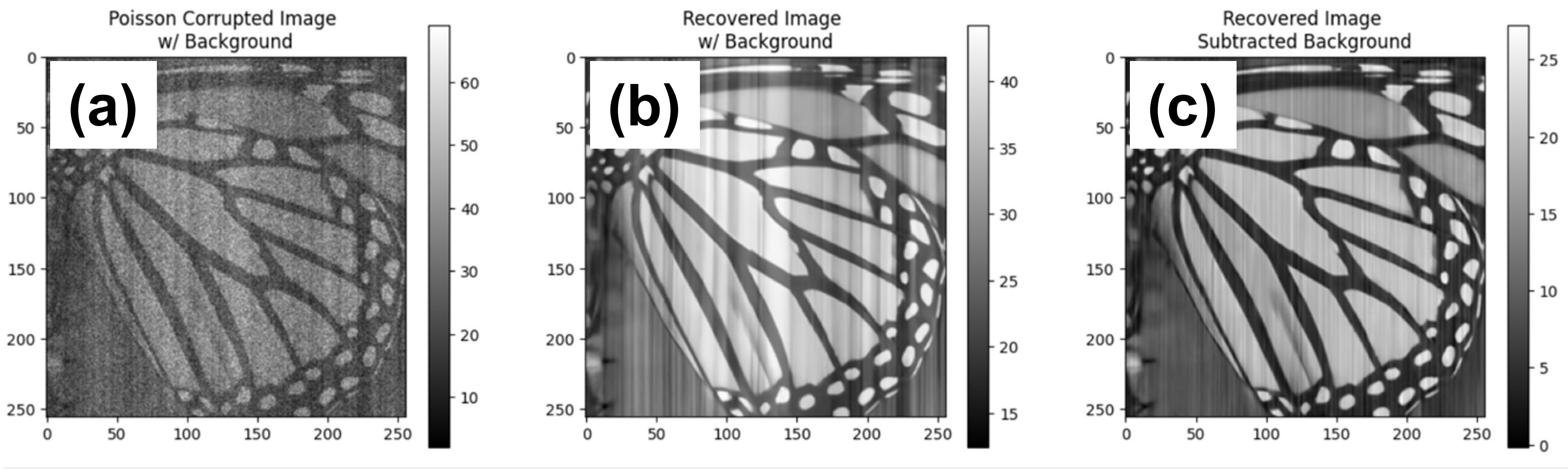
Single Resolution

- When the data is denoised together at a single resolution, the boundary artifacts are avoided

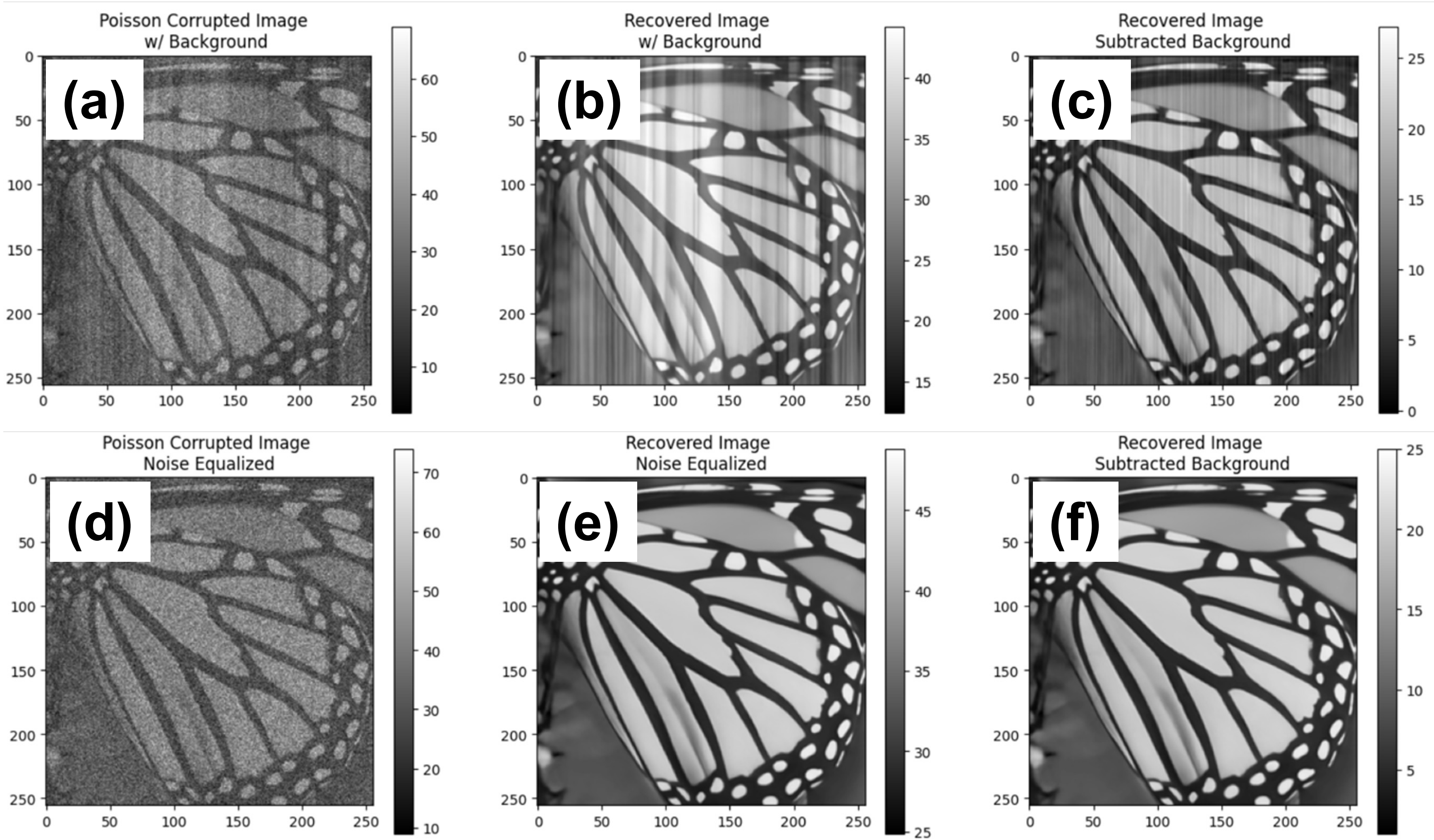




Challenge 3: Varying Background (Lacks Underlying Smoothness)


Poisson Corrupted Image
w/ Background
(a)
Recovered Image
w/ Background
(b)
Recovered Image
Subtracted Background
(c)

Increased the solar background to make the background noise uniform



Now, what about ATLID EarthCARE?

- The proposed denoising might work on ATLID EarthCARE data...
  - ...as long as the noise is approximately Poisson distributed
- However, there are caveats:
  - The ATLID backscatter [m^{-1} sr^{-1}] is a ratio between two channels
  - Denoising the two channels and taking the ratio of each might induce artifacts in the clear-air regions
  - The same issue applies to the depolarization ratio computation
- The denoising method is agnostic of the lidar physical model:
  - Hence, one should not be surprised to see non-physical extinction [m^{-1}] and lidar ratio [sr] retrievals in areas where there are low aerosol loading

Future work

- The denoising framework we presented on allows for any state-of-the-art Gaussian denoising method to be used to denoise CALIOP data
- Meaning, in the future when more advanced Gaussian denoising methods are developed we can easily adopt such methods
- Hence, the denoising framework we presented on can be transposed to other atmospheric lidar systems that have Poisson noisy like observations such as ICESat-2, EarthCARE ATLID, etc.
- The next challenge for the denoising framework is to make the denoising method aware of the lidar physical model
    - Specifically, the current denoising framework has no knowledge about the beer lambert law (signal attenuation) and molecular backscatter
    - We can expect that if the denoising framework is more aware of the lidar physical model, we can more accurately capture say the attenuation below clouds and aerosols

Extra slides

CALIOP simulation

Quantify denoising performance of both aerosol and cirrus cloud attenuated backscatter coefficients

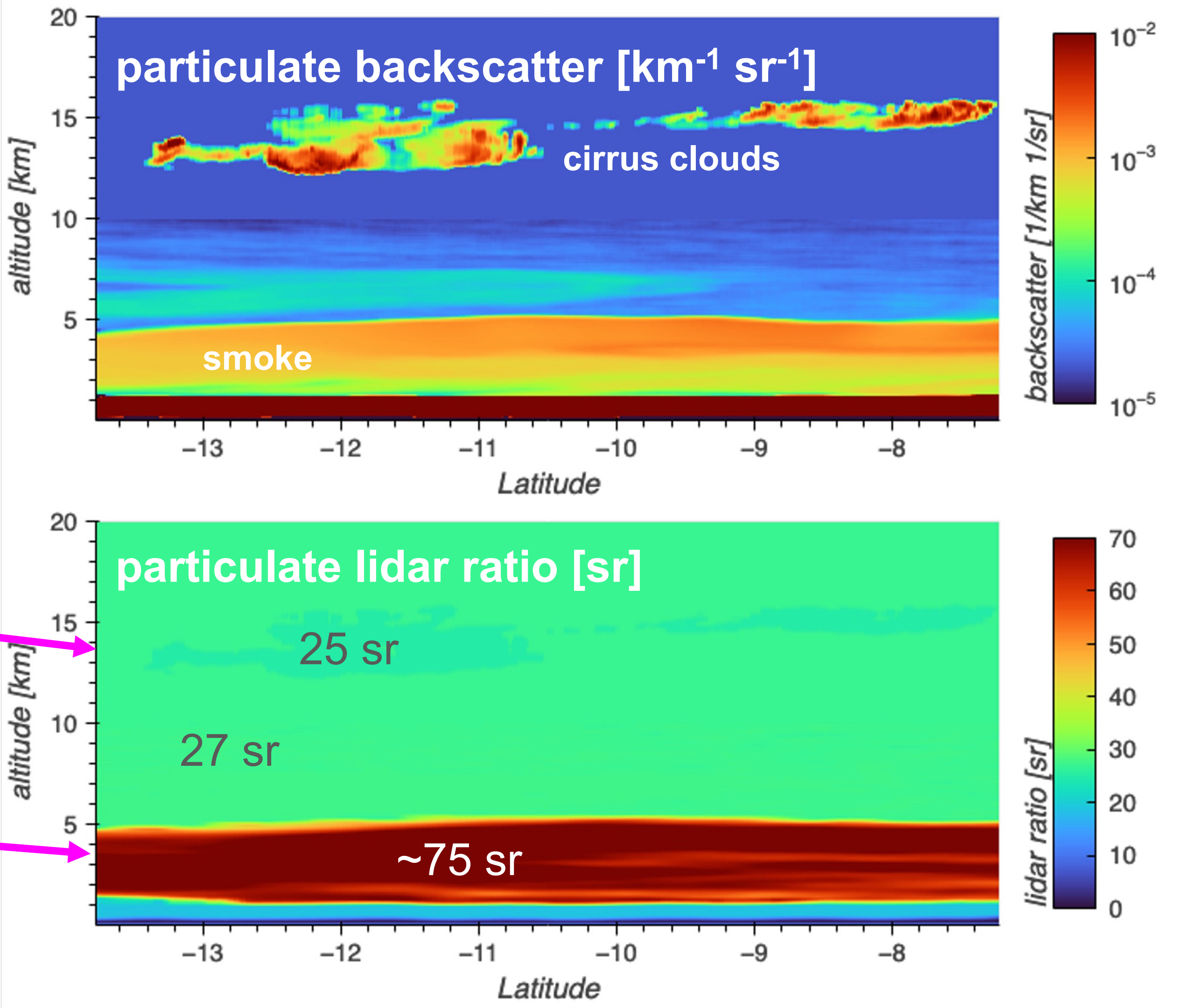
CALIOP simulations

- Though we can use CALIOP and HSRL2 under flight measurements to validate the denoised attenuated backscatter, there are no validation data for cirrus clouds attenuated backscatter retrievals for CALIOP (that Willem is aware of)
- Hence, we employ the ACCP/AOS lidar working group CALIOP simulator to simulate a daytime CALIOP scene with aerosols and cirrus clouds
- The following five slides show
  - the attenuated backscatter coefficient geophysical variables that were used as input to the CALIOP simulator
  - and the approximate raw observations of the CALIOP instrument
  - These slides specifically highlight the severity of the solar background radiation noise

The geophysical variables that are used in the CALIOP simulator

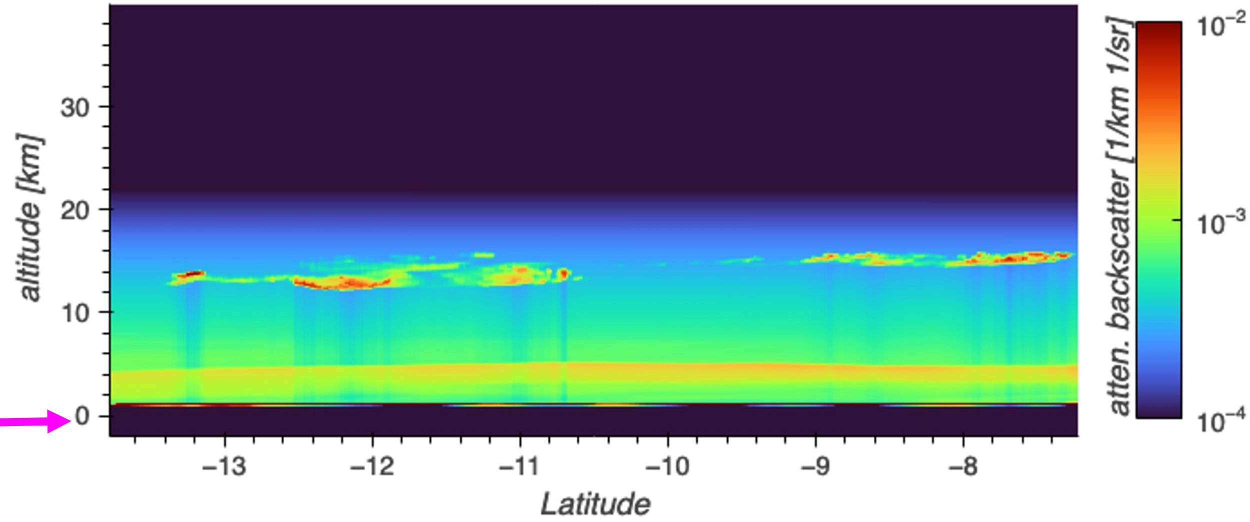
From UW HSRL
(CAMP2EX, Manilla)

From NASA HSRL2
(Oracles)

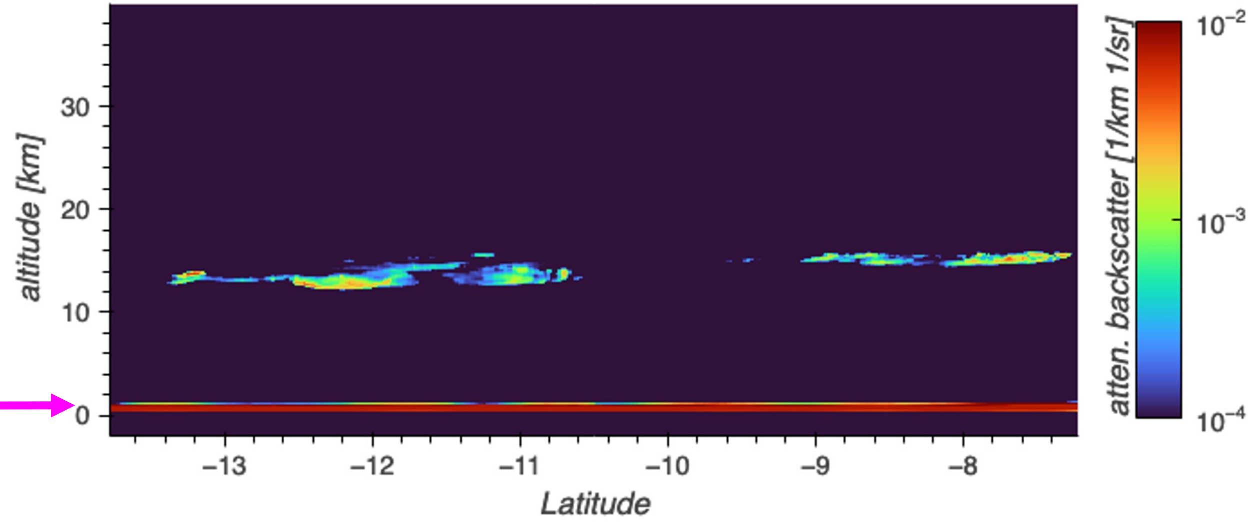


The simulated attenuated backscatter coefficients

parallel attenuated backscatter [km^{-1} sr^{-1}]

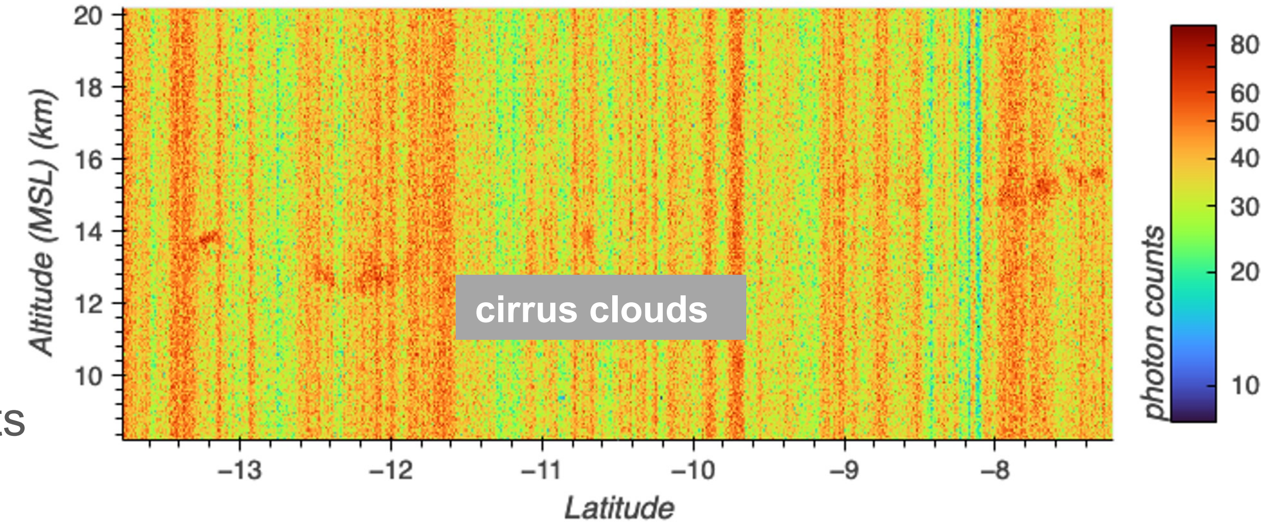


perpendicular attenuated backscatter [km^{-1} sr^{-1}]

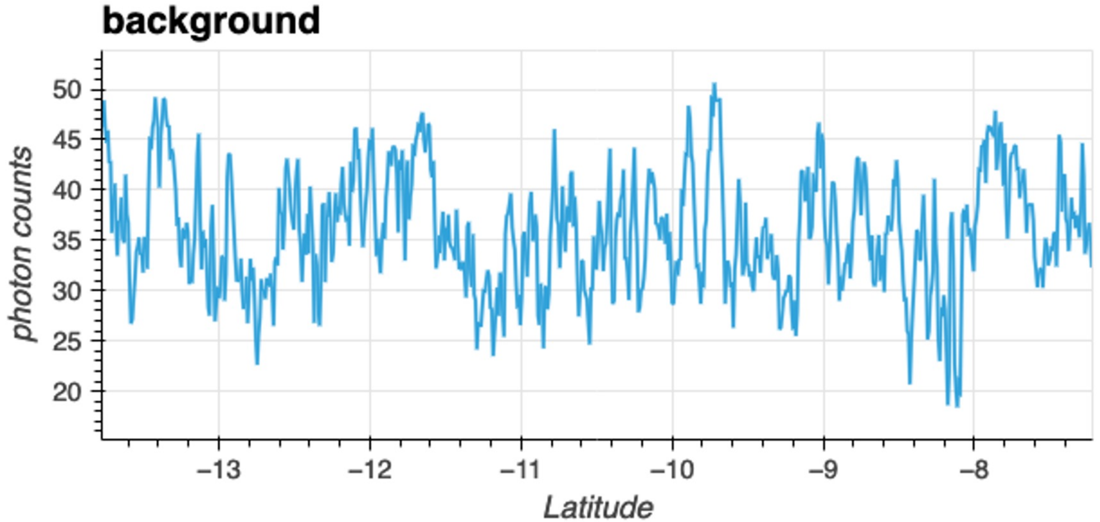


Simulated CALIOP noisy observations between 8.2km and 20km

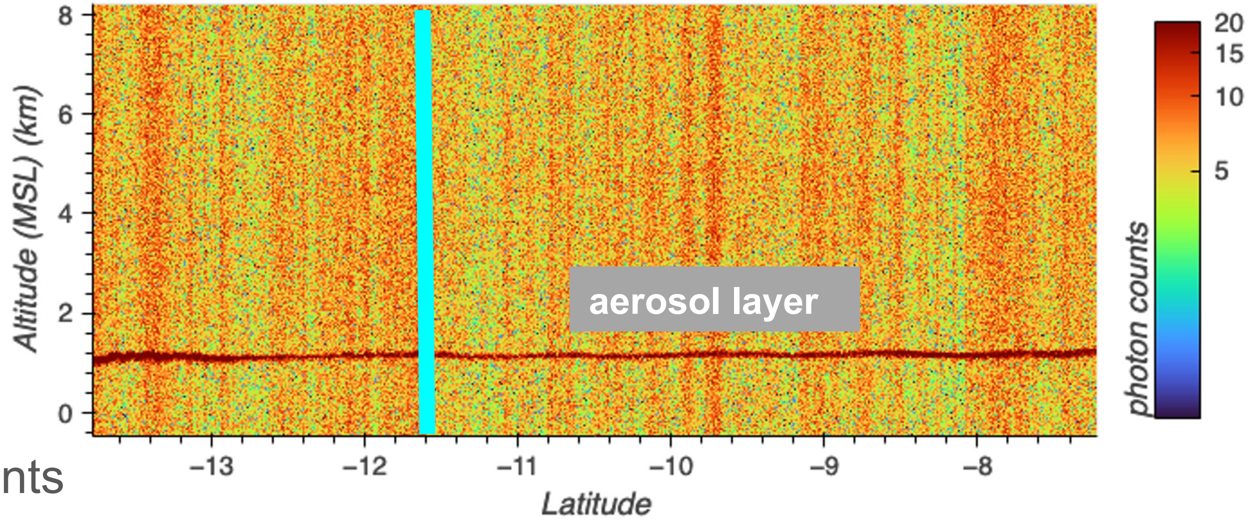
Region 3 parallel channel photon counts



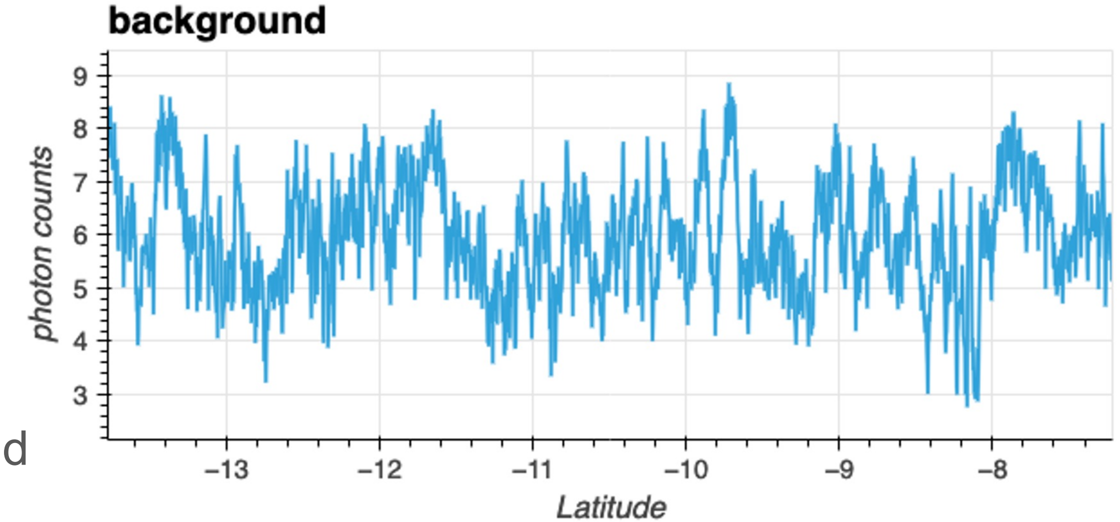
Region 3 parallel channel background



Simulated CALIOP noisy observations between -0.5km and 8.2 km



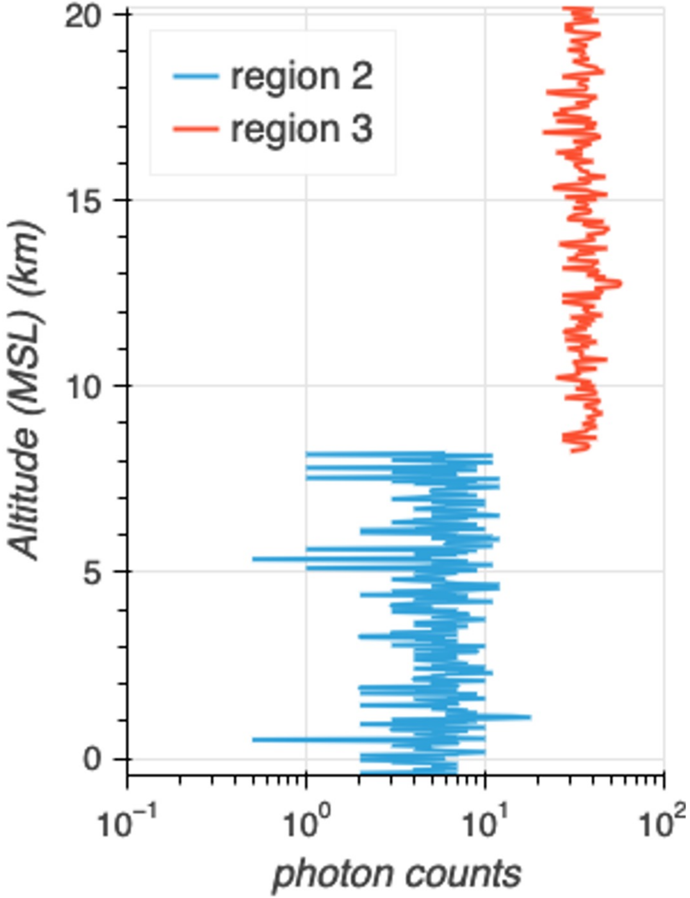
Region 2 parallel channel photon counts



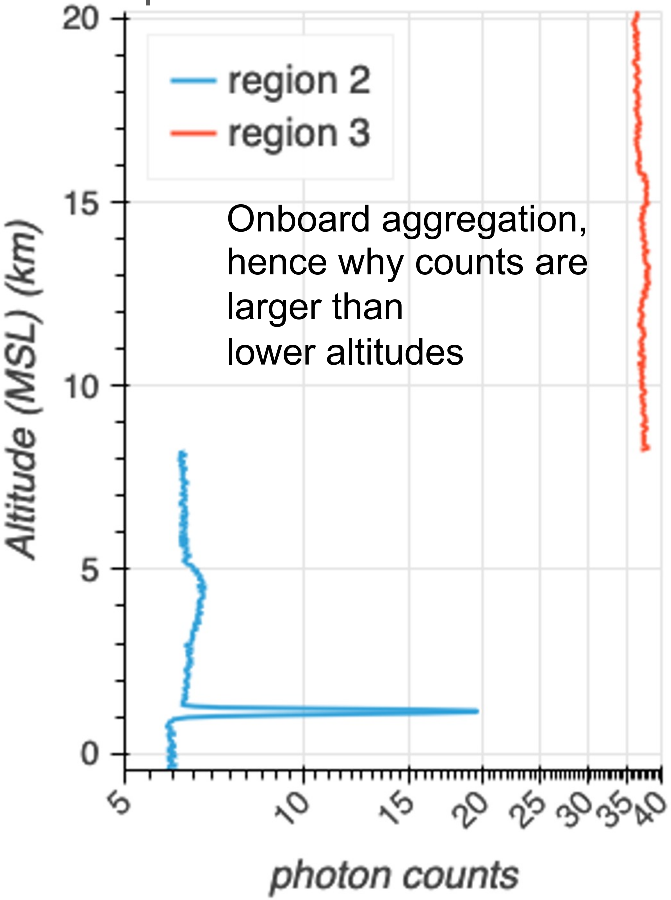
Region 2 parallel channel background

These vertical altitude plots give you a quantitative perception of the noise levels of CALIOP

Single profile photon counts at latitude -12.44°

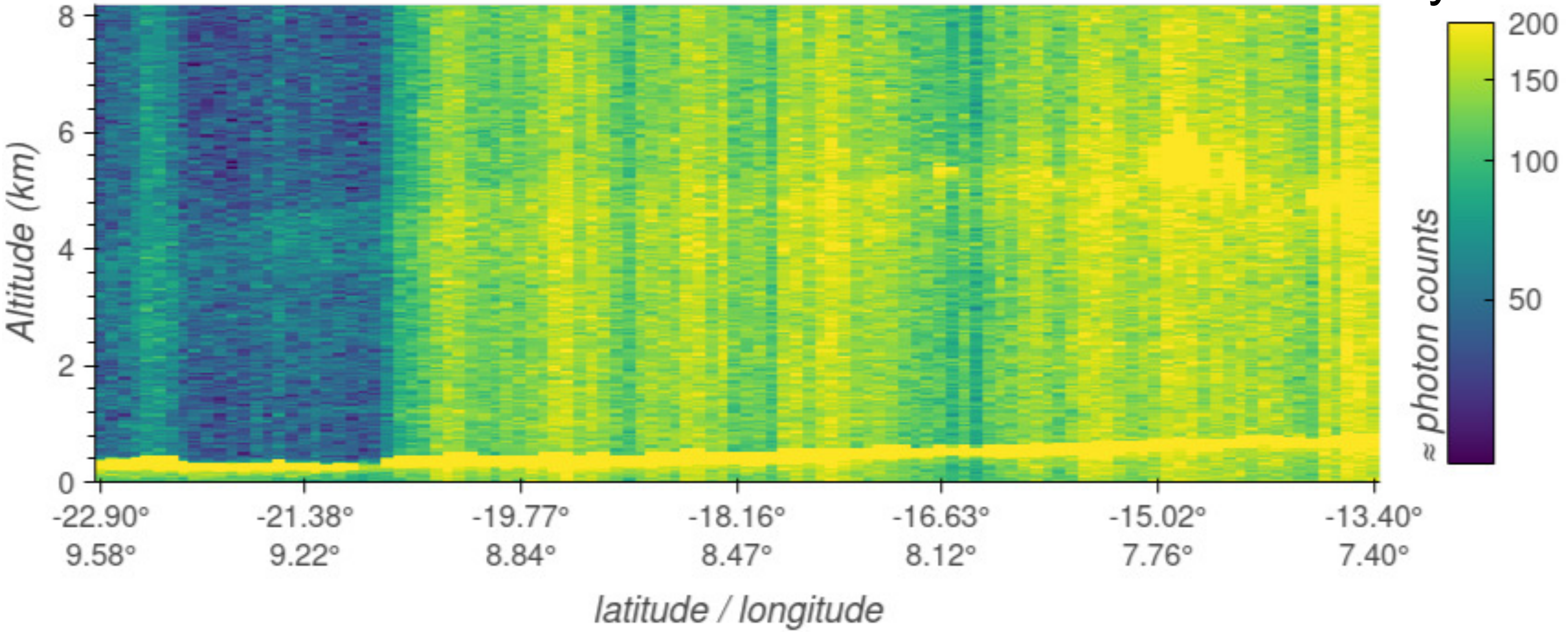


Average profile photon counts



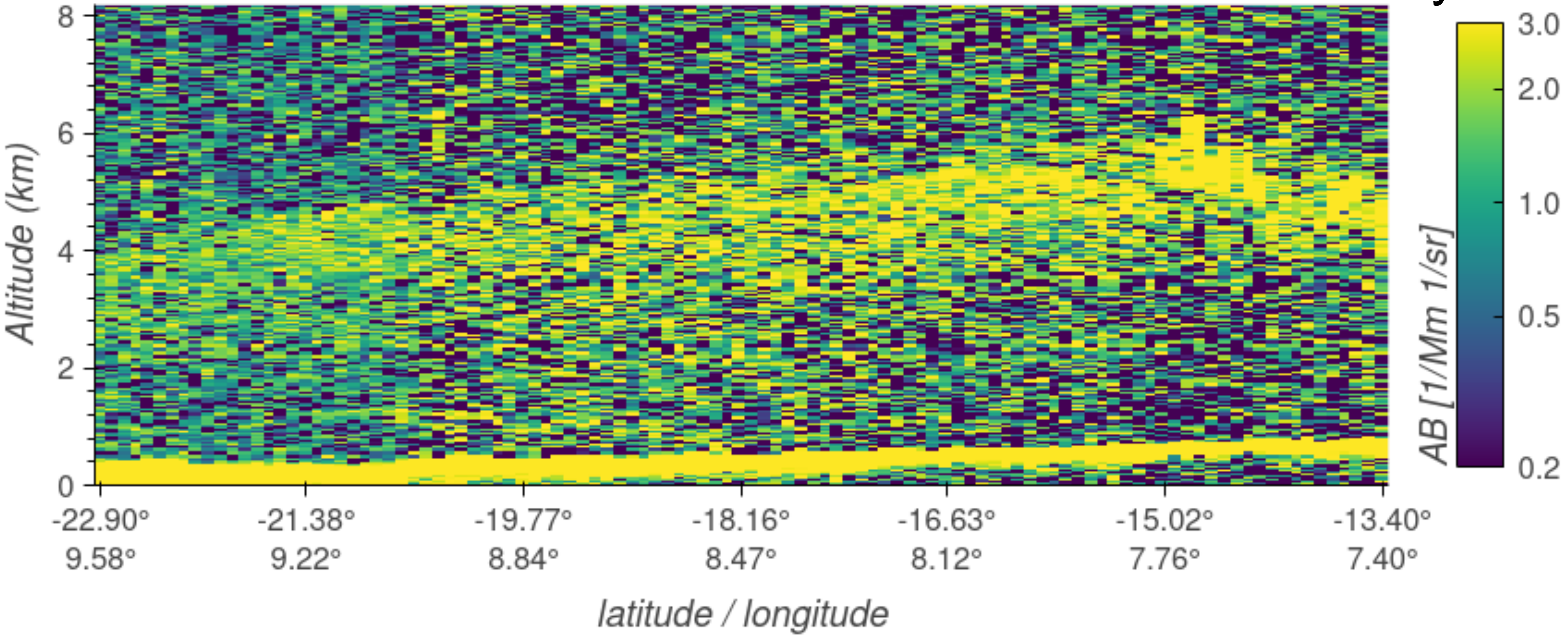
More denoising results

parallel 532nm **noisy counts** (10km horizontal resolution)
CALIOP 2016-09-20T13-01-09 Daytime

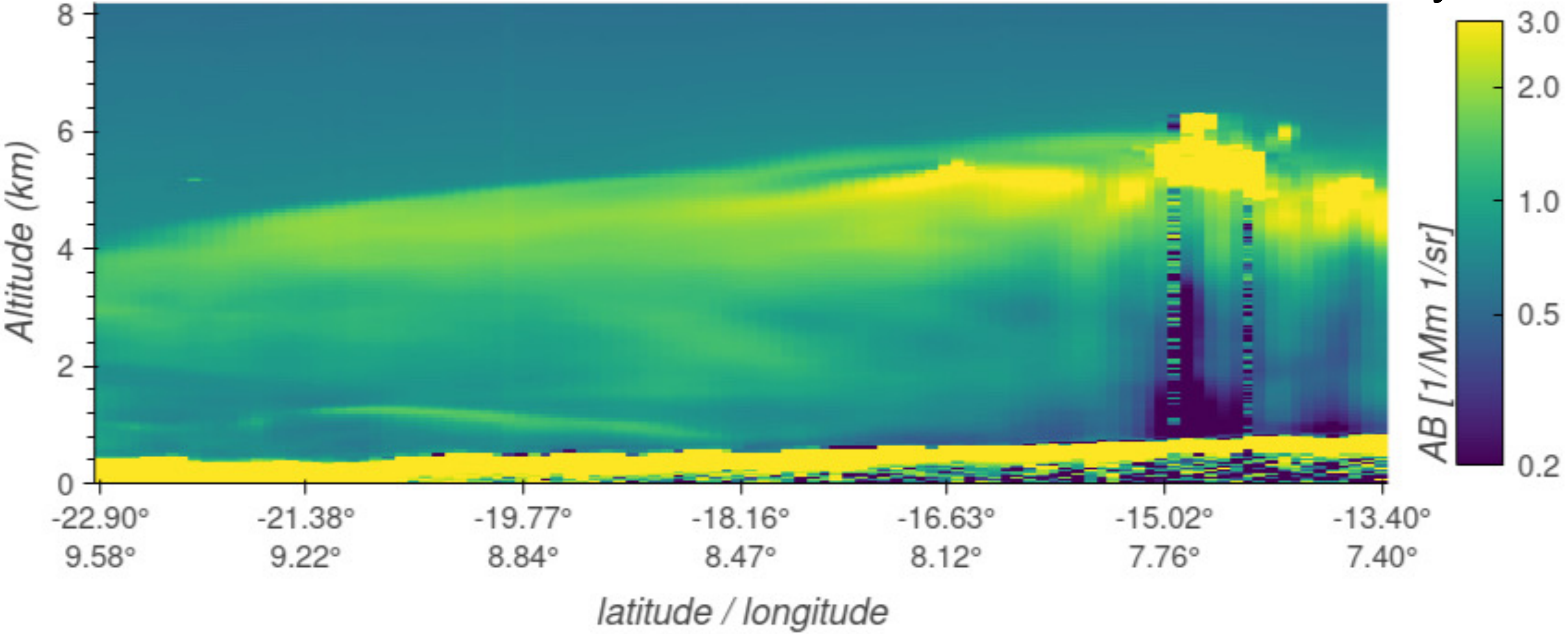


standard attenuated backscatter (10km horz.)

CALIOP 2016-09-20T13-01-09 Daytime

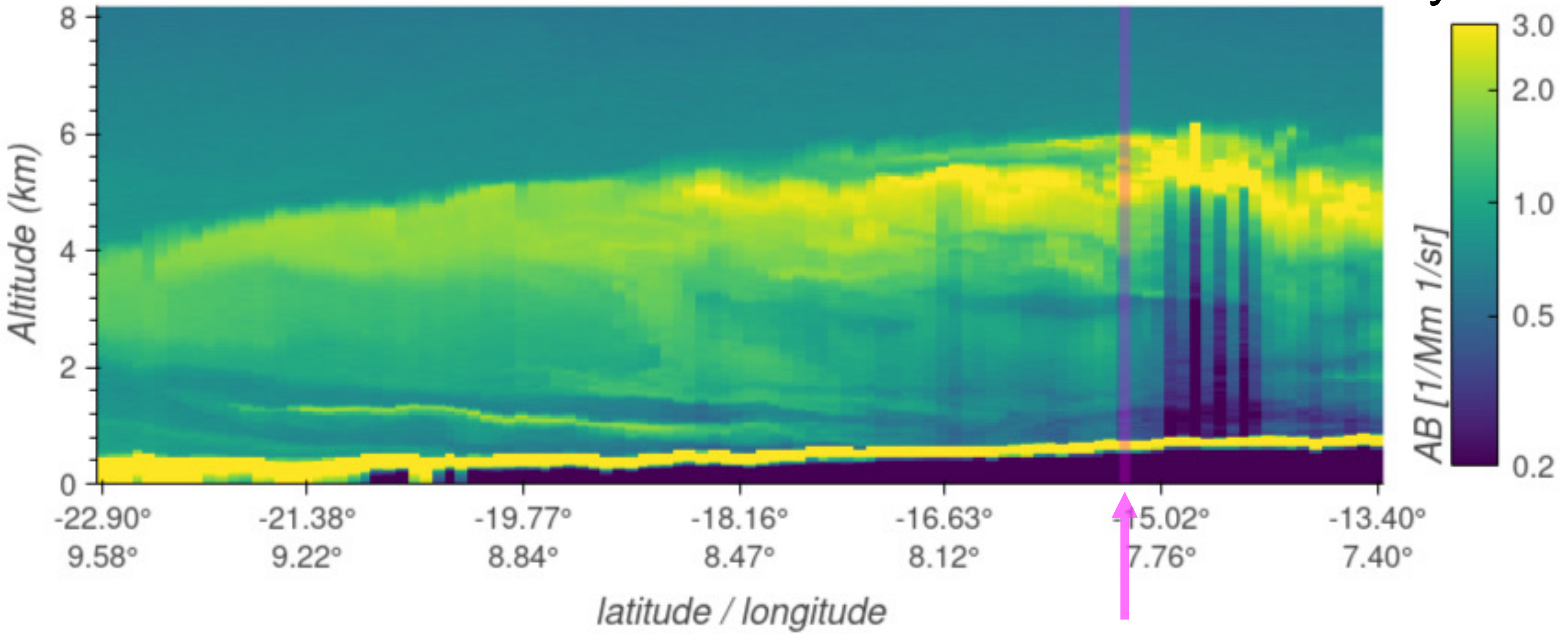


denoised (DRUNet) attenuated backscatter (10km horz.)
CALIOP 2016-09-20T13-01-09 Daytime



Proxy truth HSRL2 attenuated backscatter (10km horz.)

CALIOP 2016-09-20T13-01-09 Daytime

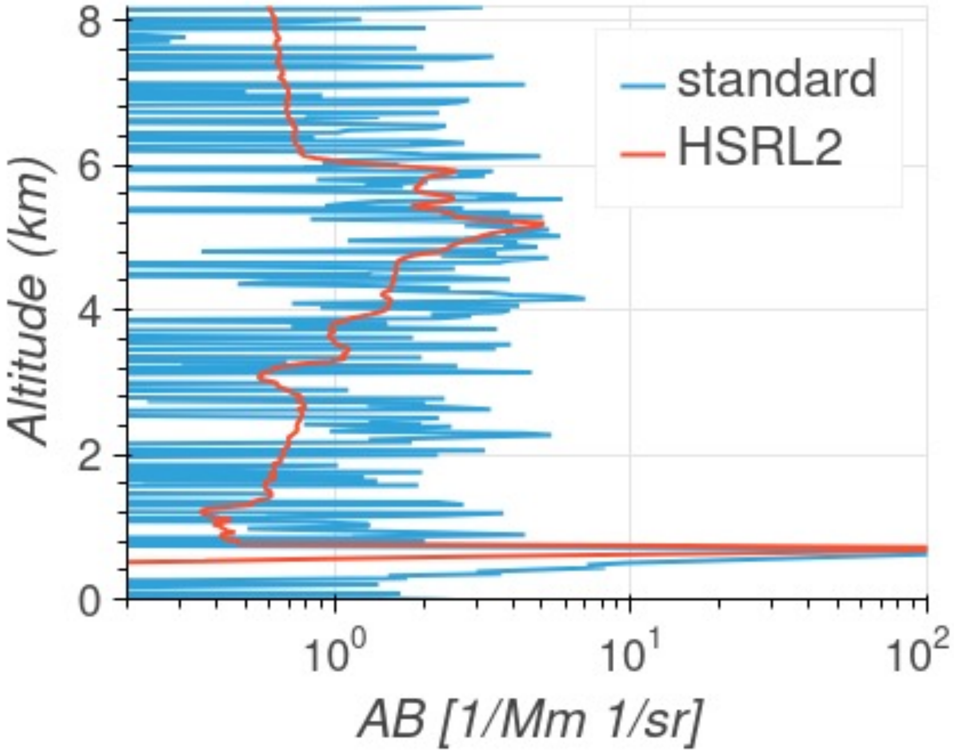


CALIOP / HSRL2
coincident time

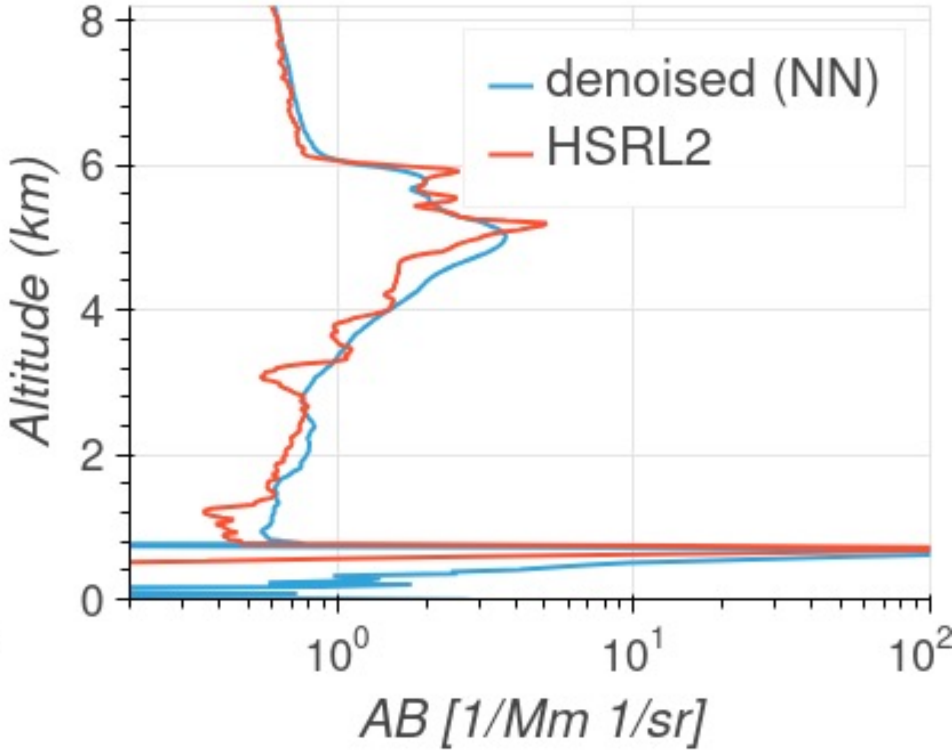
Profile plots at co-incident time

CALIOP 2016-09-20T13-01-09 Daytime

Attenuated backscatter (AB) retrieval with no denoising at 10km horizontal resolution



AB denoised via DRUNet at 10km horizontal resolution



Proposed denoising frame vs previous approaches

- With 1 GPU, it will take ~1 hour to denoise 1 channel of 1 L1B with a (old) 2018 GPU
  - Further optimization can be done to reduce processing time
- Proposed denoising framework does not use a forward model at all
  - Meaning, the framework has no knowledge about say the molecular backscatter
  - (Refer to last slide for future work)