JAXA

esa

EUROPEAN SPACE AGENCY

EarthCARE 2026 Science and Validation workshop

8–12 June 2026 | Rhodes House | Oxford, UK

Will validation succeed? Practical criteria for comparing spaceborne and ground-based lidars: application to ATLID/EarthCare and IPRAL

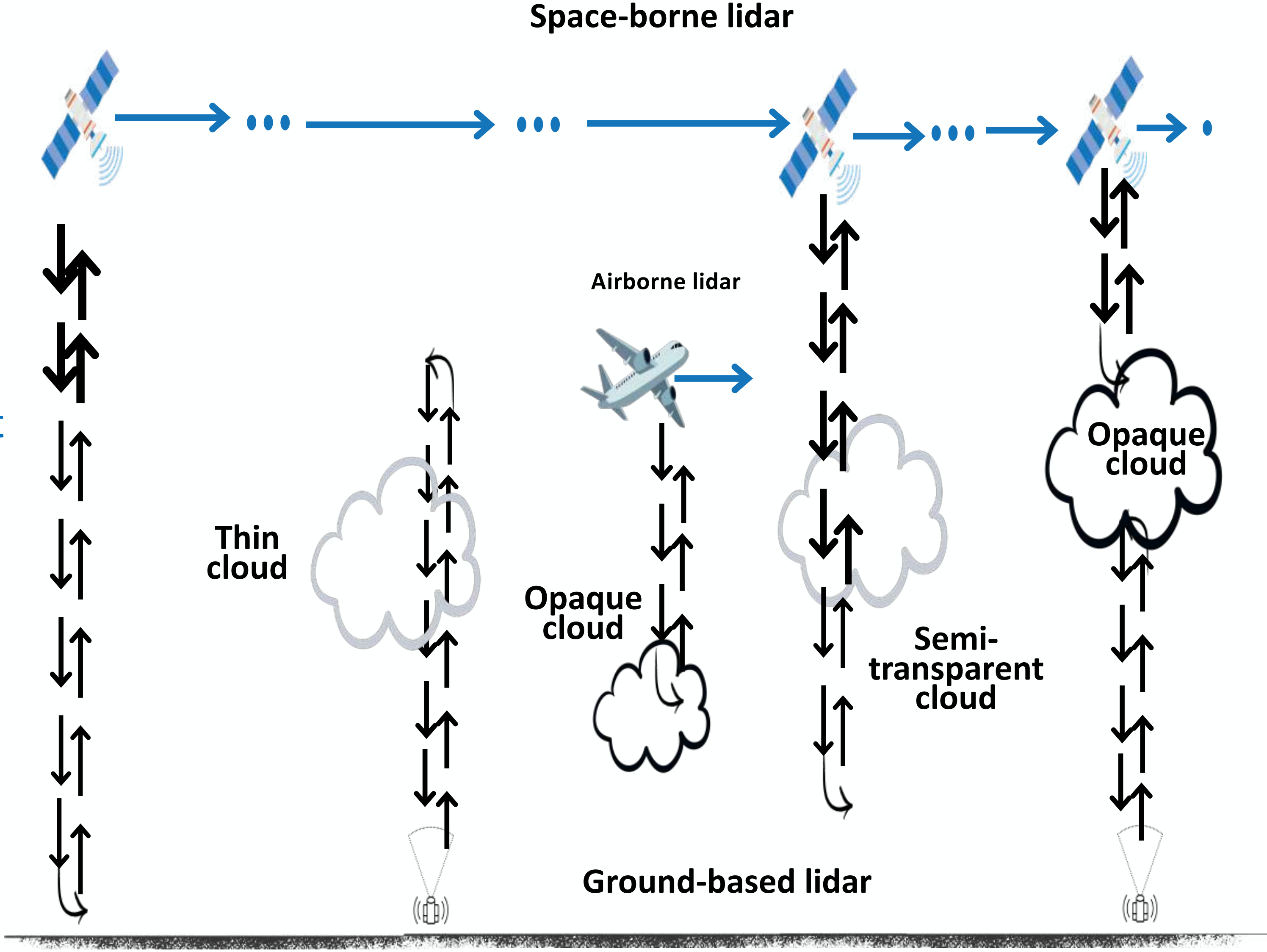
Feofilov A.[1] , Haddouche A.[1] , Chepfer H.[1] , Khaykin S.[2] , Lapouge F.[3] , Delville P.[4] , Drouin M.[4] ,
Haeffelin M.[5] , Dupont J.[6] , Pietras C.[4]

1 - Sorbonne University / LMD / CNRS / Ecole Polytechnique, Paris, France; 2 - LATMOS, UVSQ, Sorbonne Université, CNRS, IPSL
3 - Ecole Polytechnique / LMD; 4 - CNRS / LMD; 5 - CNRS / IPSL / LMD; 6 - UVSQ / IPSL / LMD

Formulation of the problem

- Spaceborne lidar observations need validation
- Airborne lidars are expensive
- Ground-based lidars provide controllable reference data
- But, there are two obstacles for reliable collocated data analysis:
  (1) Data comparability issues linked to different viewing geometries
  (2) Spatiotemporal variability

- Data used:
  ATLID L1/L2, 2024-now
- IPRAL 355nm lidar: two telescopes, 0-40km, co-polar/cross-polar, analog/photocounting, research L1 product and L2 retrievals.



What do we compare and how?

$$ATB(\lambda,z) = (\beta_{mol}(\lambda,z) + \boxed{\beta_{part}(\lambda,z)}) \times e^{-2\int_{z_{lidar}}^{z}(\alpha_{mol}(\lambda,z\prime)+\eta\alpha_{part}(\lambda,z\prime))dz\prime}$$

Total Attenuated Backscatter, L1

$$AMB(\lambda,z) = \beta_{mol}(\lambda,z) \times e^{-2\int_{z_{lidar}}^{z}(\alpha_{mol}(\lambda,z\prime)+\eta\alpha_{part}(\lambda,z\prime))dz\prime}$$

Attenuated Molecular Backscatter, L1

$$ATB_{mol}(\lambda,z) = \beta_{mol}(\lambda,z) \times e^{-2\int_{z_{lidar}}^{z}\alpha_{mol}(\lambda,z\prime)dz\prime}$$

Attenuated Molecular Backscatter in clear sky conditions

$$SR1(\lambda,z) = \frac{ATB(\lambda,z)}{ATB_{mol}(\lambda,z)}$$

Scattering Ratio for non-HSRL lidars (theoretical ATB_{mol})

$$SR2(\lambda,z) = \frac{ATB(\lambda,z)}{AMB(\lambda,z)}$$

Scattering Ratio for HSRL lidars (measured AMB)

Comparability score (CS) for two vectors, a and b: 0 is bad, 1 is good

$$CS = 0 \;\; if \;\; \vec{a} \approx 0 \; or \; \vec{b} \approx 0$$

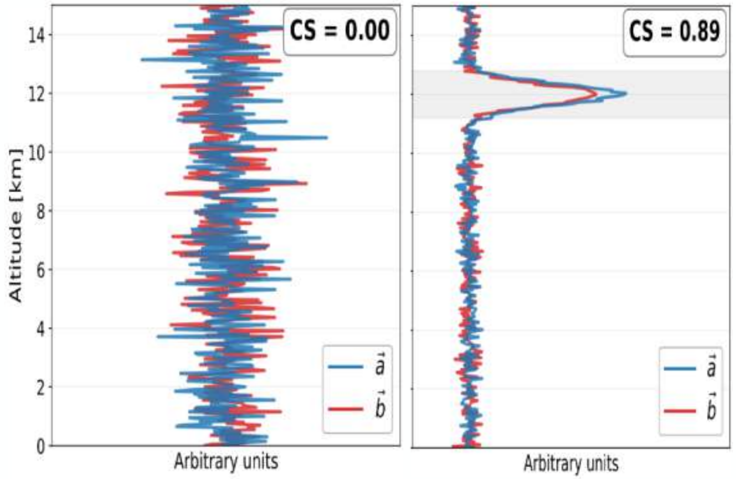
Excluding no-signal profiles

$$CS = \frac{N_{joint}}{N_{tot}} \times \{(1 - w_{shape}) \times S_{corr} + w_{shape} \times S_{shape}\}$$

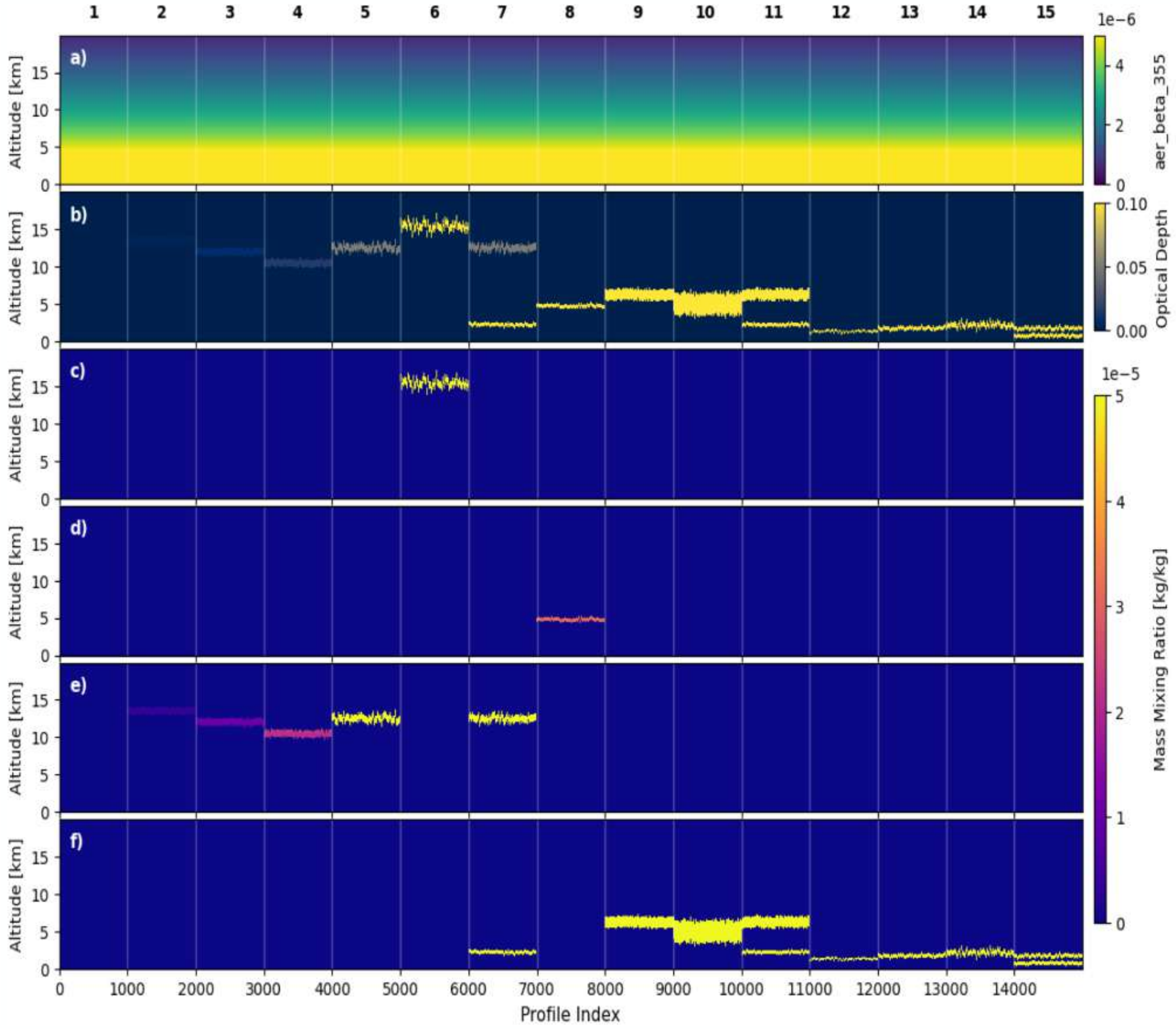
Three contributors

$$S_{corr} = \frac{\rho(\vec{a},\vec{b})+1}{2}; \qquad S_{shape} = e^{-2\times\frac{|a-b|}{|\vec{a}|+|\vec{b}|}}$$

Pearson's correlation and scaled difference terms



Atmospheric scenarios used in simulations



N	Scene name	Altitude range(s) (km)	COT	Particle type	R_{eff} (µm)	Lidar ratio	P-t-p variab (m)	Corr. length (m)
1	Clear Sky	0–30	0	N/A	N/A	N/A	N/A	N/A
2	Very thin subvisible ***Ci***	13.0–14.0	0.05	Ice	20	18.3	80	200
3	Subvisible ***Ci***	11.5–12.5	0.15	Ice	24	19.2	90	300
4	Thin ***Ci***	10.0–11.0	0.3	Ice	30	20.8	100	400
5	Semi-transp. ***Ci***	12.0–13.0	1.0	Ice	37	23.3	200	600
6	Opaque ***Ci***	15.0–16.0	3.0	Ice	45	26.3	350	800
7	***Ci*** above thick low	12.0–13.0 upper)	1.0	Ice	37	23.3	200	600
		2.0–2.5 (lower)	10.0	Liquid	37	23.3	100	600
8	Thin ***Ac***	4.5–5.0	1.0	Liquid	12	19.3	100	150
9	Semi-transp. ***As***	5.5–7	3.0	Liquid	14	19.2	100	500
10	Opaque ***As***	3.5–6.5	10.0	Liquid	16	18.9	150	600
11	Semi-transp. ***As*** above thick low	5.5–7.0 (upper)	3.0	Liquid	14	19.2	100	500
		2–2.5 (lower)	10.0	Liquid	14	19.2	100	500
12	Thin ***Sc***	1.5–2.5	3.0	Liquid	9	19.3	100	150
13	Semi-transp. ***Sc***	1.5–2.5	5.0	Liquid	12	19.3	120	200
14	Opaque ***Sc***	1.8–2.5	20.0	Liquid	14	19.2	200	250
15	Semi-transp. ***Sc*** above thick low	1.5–2 (upper)	5.0	Liquid	12	19.3	120	200
		0.5–1 (lower)	10.0	Liquid	12	19.3	100	200

Simulation flowchart

WACCM6 CESM output (P, T,) [1.0x1.25° hourly], (Gettelman, 2019)

Clear sky + 14 cloud scenarios and their variability (this work), [1000 pts/scene, H=300m, V=50m]

- Extraction of data
- Interpolation on the grid corresponding to horizontal resolution of the lidar

Model outputs: vertical profiles for the pseudo-tracks (P, T, clouds, aerosols), [H=300m, V=50m]

Running COSPv2 (Swales et al., 2018) lidar observables simulator for nadir and zenith

Noise-free ATB(z), APB(z), AMB(z) zenith and nadir profiles at 355nm and 532nm for the pseudo-tracks, [H=300m, V=100m]

Adding instrumental noise to space-borne and ground-based lidars

Noisy nighttime and daytime ATB(z), APB(z), AMB(z) zenith and nadir profiles at 355nm and 532nm for the pseudo-tracks, [H=300m, V=100m]

Calculating SR(z) from L1 data with/without averaging

Noisy SR(z) zenith/nadir profiles 355/532nm [H = 300, 3000, 30000m, V=100m]

Optical properties retrieval from L1 data with/without averaging

Noisy $\alpha(z)$ and $\beta(z)$ profiles 355/532nm [H = 300, 3000, 30000m, V=100m]

Calculating SR(z) from L2 data

Noisy SR(z) zenith/nadir profiles 355/532nm [H = 300, 3000, 30000m, V=100m]

Calculating Comparability Score (CS) values for pairs of zenith/nadir profiles

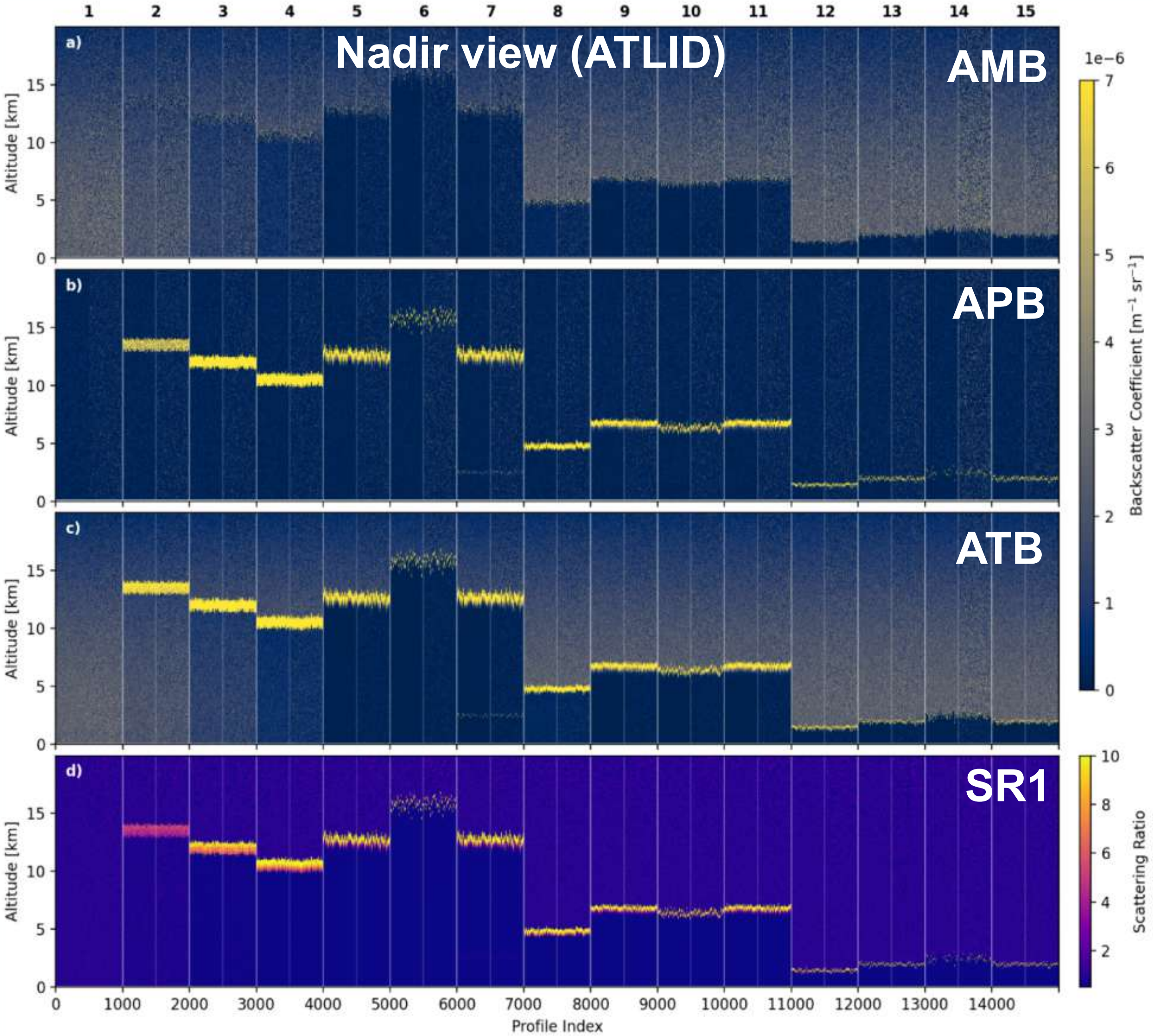
L1 comparability: average CS values for 15 scenes at 355/532nm, day/night, [H = 300, 3000, 30000m] for perfect overlaps and 150km distance

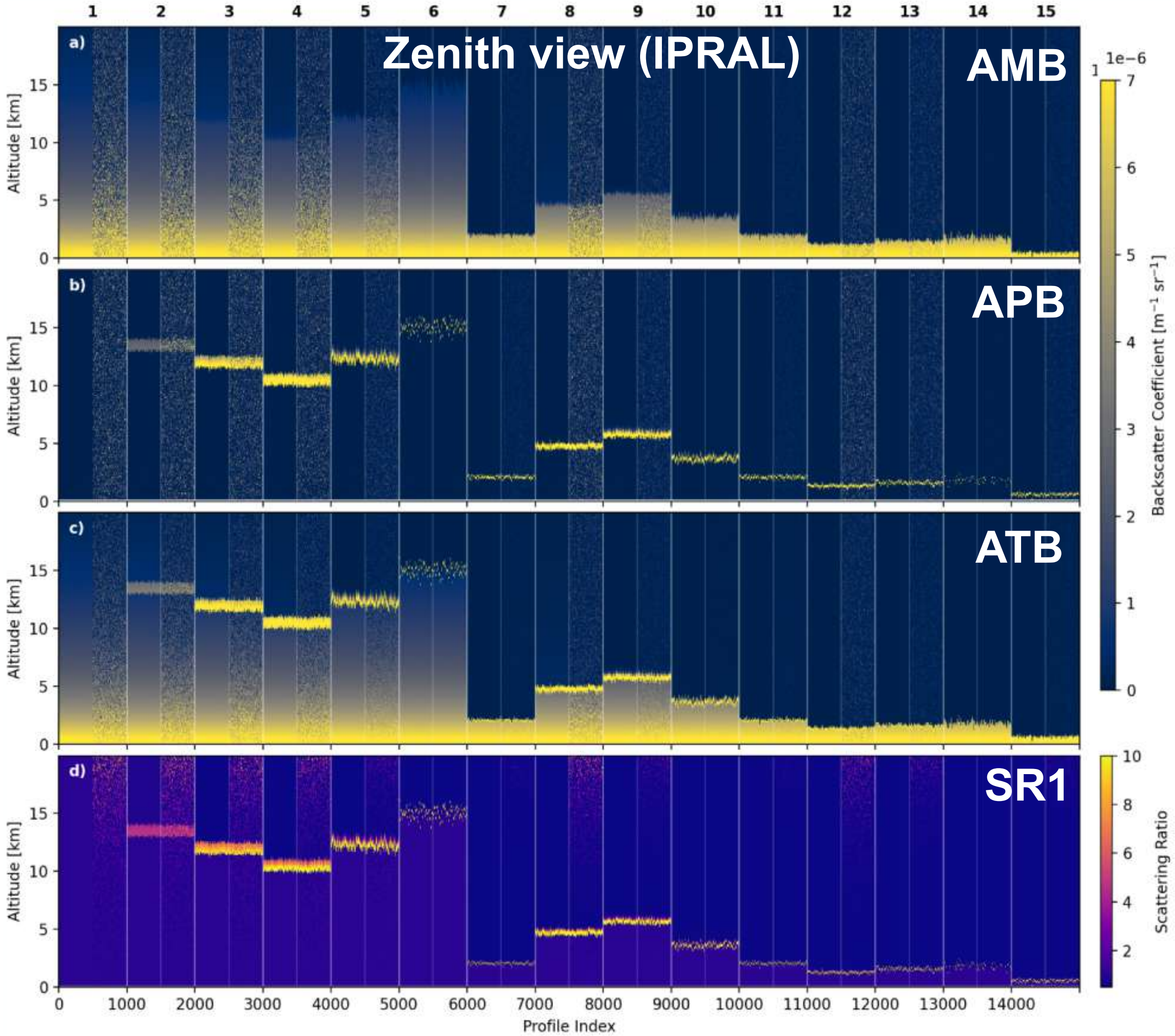
L1 comparability: average CS values for 15 scenes at 355/532nm, day/night, [H = 300, 3000, 30000m] for perfect overlaps and 150km distance

Key features:

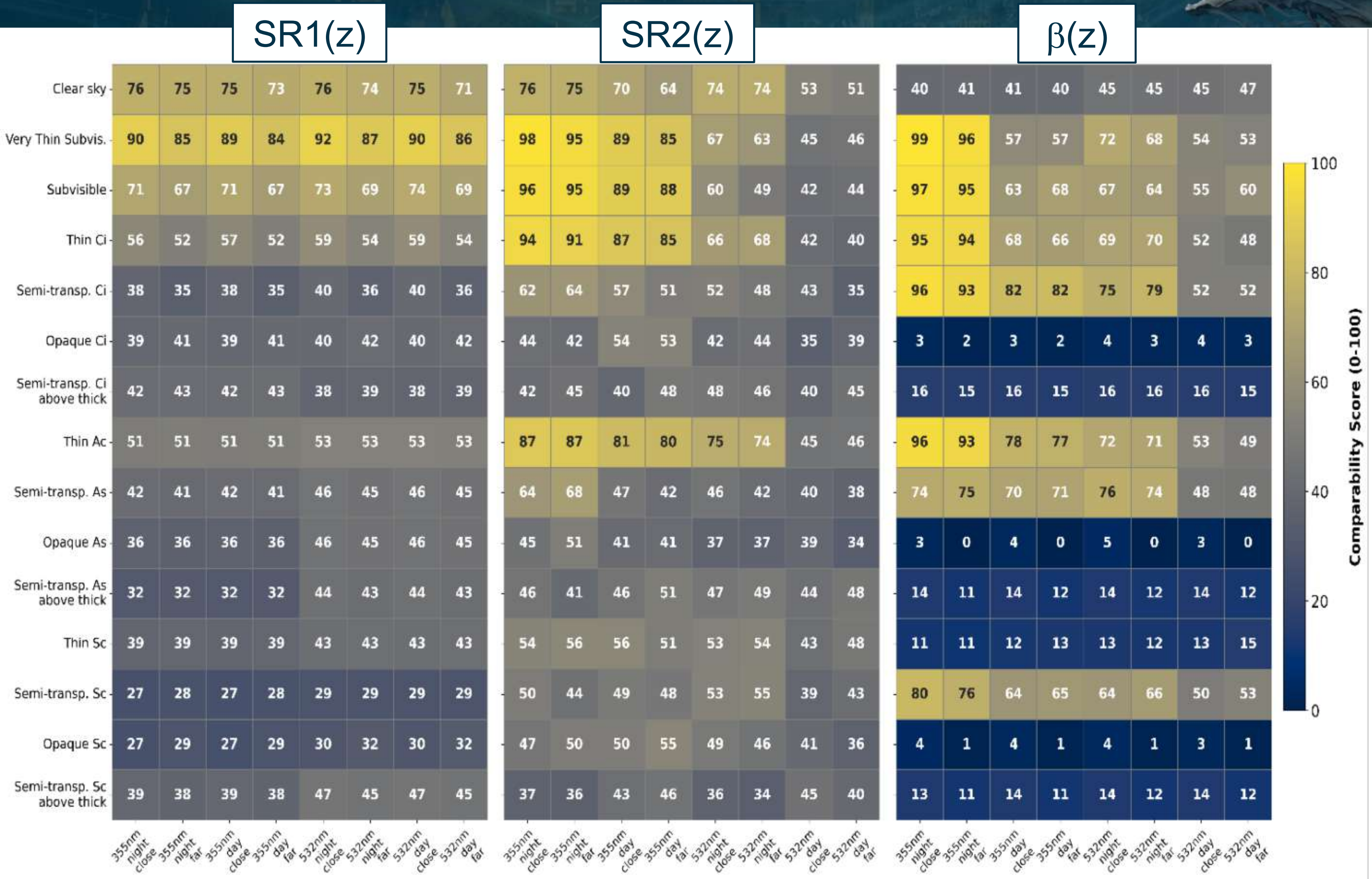
- Uses a combination of COSP lidar simulator (Swales et al., 2018) and instrument-specific noise.
- Approach tested for cirrus and stratocumulus cloud observations by ATLID (Feofilov et al., AMT, 2023).
- Resolution: 300m(H) x 50m(V)
- 30000 measurements simulated in nadir and zenith

Nadir and zenith simulations: 15 scenes, day/night





Comparability scores (CS) for different scenes and variables



SR1(z)

Scene	355nm night close	355nm night far	355nm day close	355nm day far	532nm night close	532nm night far	532nm day close	532nm day far
Clear sky	76	75	75	73	76	74	75	71
Very Thin Subvis.	90	85	89	84	92	87	90	86
Subvisible	71	67	71	67	73	69	74	69
Thin Ci	56	52	57	52	59	54	59	54
Semi-transp. Ci	38	35	38	35	40	36	40	36
Opaque Ci	39	41	39	41	40	42	40	42
Semi-transp. Ci above thick	42	43	42	43	38	39	38	39
Thin Ac	51	51	51	51	53	53	53	53
Semi-transp. As	42	41	42	41	46	45	46	45
Opaque As	36	36	36	36	46	45	46	45
Semi-transp. As above thick	32	32	32	32	44	43	44	43
Thin Sc	39	39	39	39	43	43	43	43
Semi-transp. Sc	27	28	27	28	29	29	29	29
Opaque Sc	27	29	27	29	30	32	30	32
Semi-transp. Sc above thick	39	38	39	38	47	45	47	45

SR2(z)

Scene	355nm night close	355nm night far	355nm day close	355nm day far	532nm night close	532nm night far	532nm day close	532nm day far
Clear sky	76	75	70	64	74	74	53	51
Very Thin Subvis.	98	95	89	85	67	63	45	46
Subvisible	96	95	89	88	60	49	42	44
Thin Ci	94	91	87	85	66	68	42	40
Semi-transp. Ci	62	64	57	51	52	48	43	35
Opaque Ci	44	42	54	53	42	44	35	39
Semi-transp. Ci above thick	42	45	40	48	48	46	40	45
Thin Ac	87	87	81	80	75	74	45	46
Semi-transp. As	64	68	47	42	46	42	40	38
Opaque As	45	51	41	41	37	37	39	34
Semi-transp. As above thick	46	41	46	51	47	49	44	48
Thin Sc	54	56	56	51	53	54	43	48
Semi-transp. Sc	50	44	49	48	53	55	39	43
Opaque Sc	47	50	50	55	49	46	41	36
Semi-transp. Sc above thick	37	36	43	46	36	34	45	40

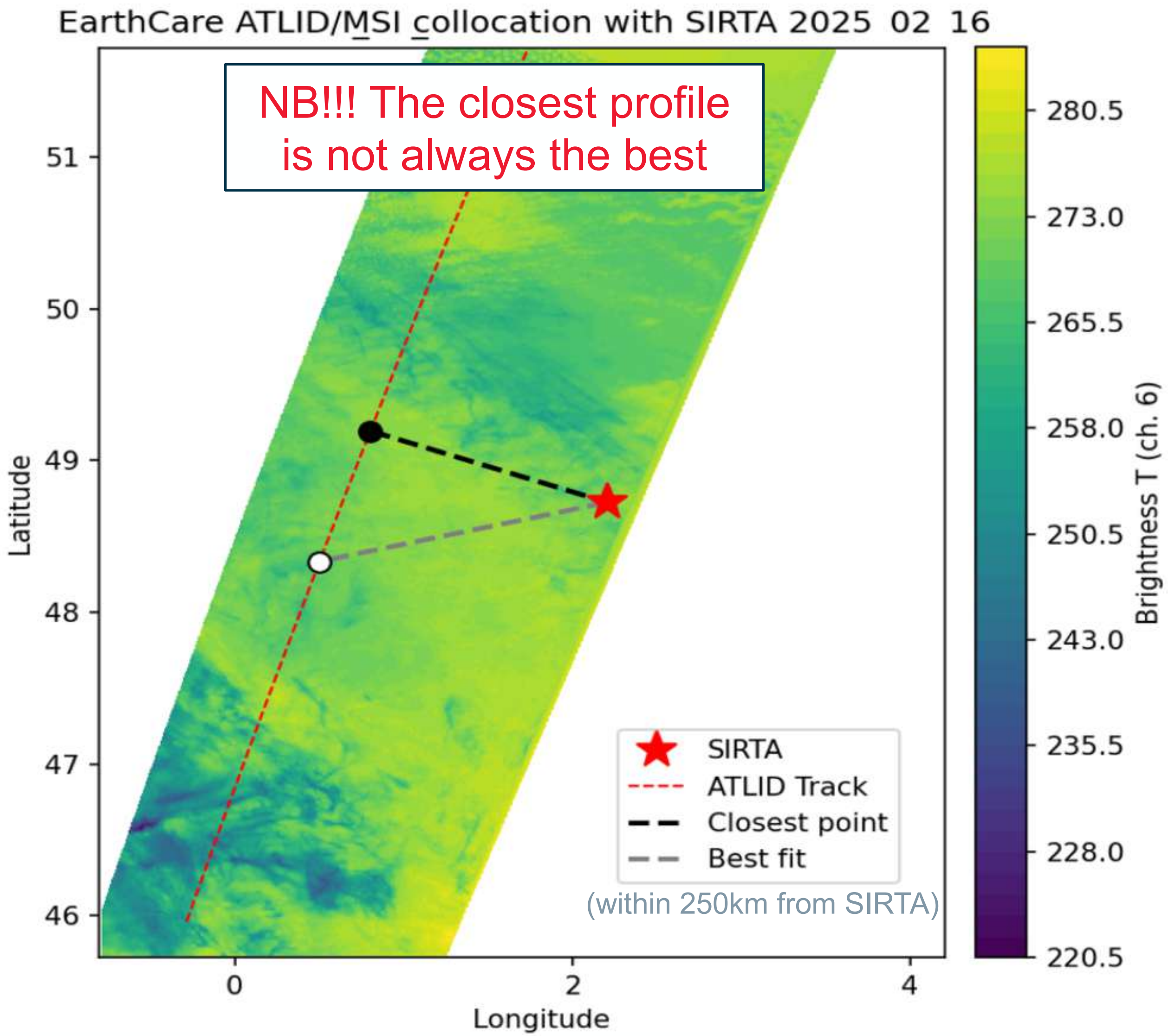
β(z)

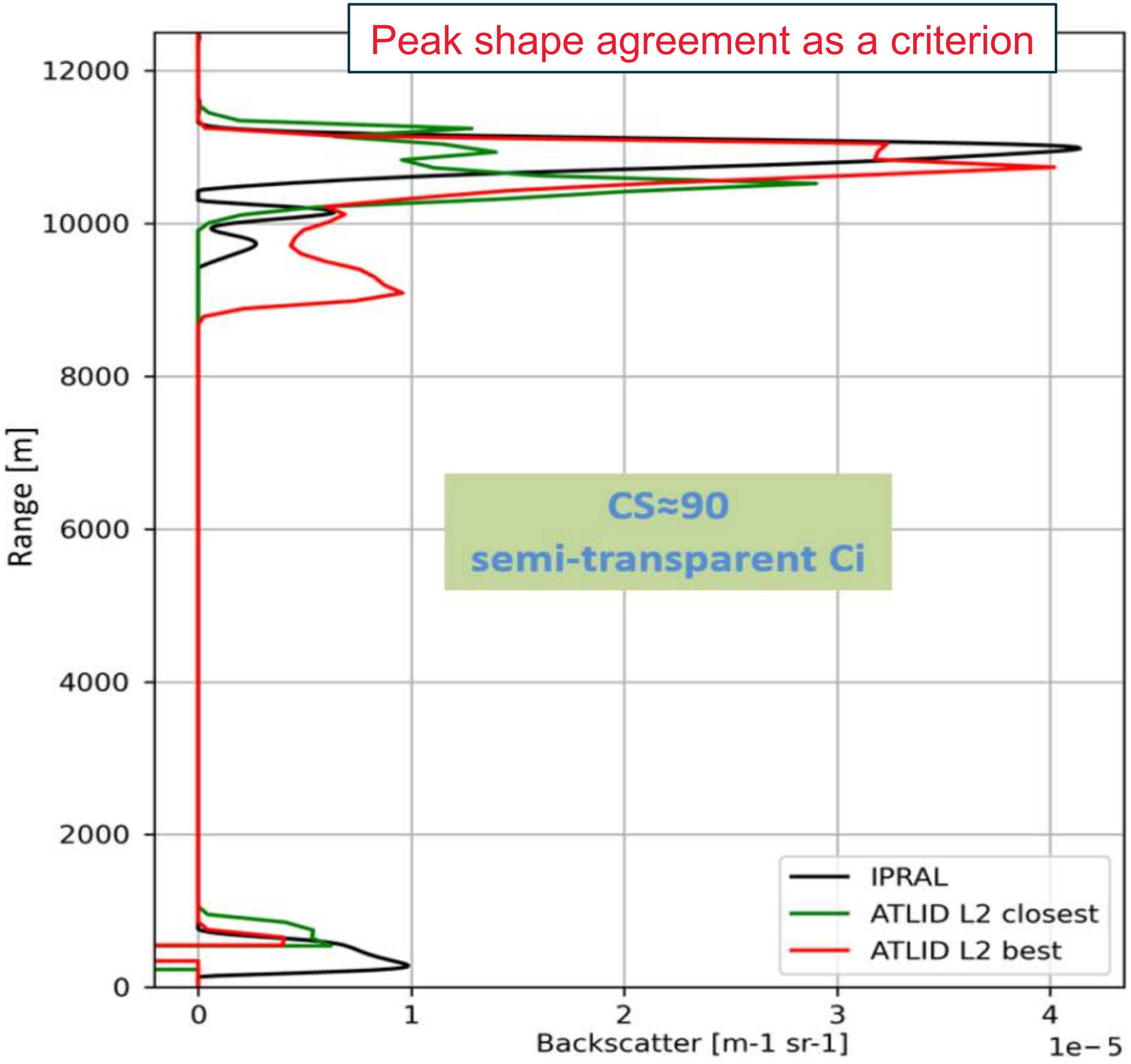
Scene	355nm night close	355nm night far	355nm day close	355nm day far	532nm night close	532nm night far	532nm day close	532nm day far
Clear sky	40	41	41	40	45	45	45	47
Very Thin Subvis.	99	96	57	57	72	68	54	53
Subvisible	97	95	63	68	67	64	55	60
Thin Ci	95	94	68	66	69	70	52	48
Semi-transp. Ci	96	93	82	82	75	79	52	52
Opaque Ci	3	2	3	2	4	3	4	3
Semi-transp. Ci above thick	16	15	16	15	16	16	16	15
Thin Ac	96	93	78	77	72	71	53	49
Semi-transp. As	74	75	70	71	76	74	48	48
Opaque As	3	0	4	0	5	0	3	0
Semi-transp. As above thick	14	11	14	12	14	12	14	12
Thin Sc	11	11	12	13	13	12	13	15
Semi-transp. Sc	80	76	64	65	64	66	50	53
Opaque Sc	4	1	4	1	4	1	3	1
Semi-transp. Sc above thick	13	11	14	11	14	12	14	12

Takeaway messages:

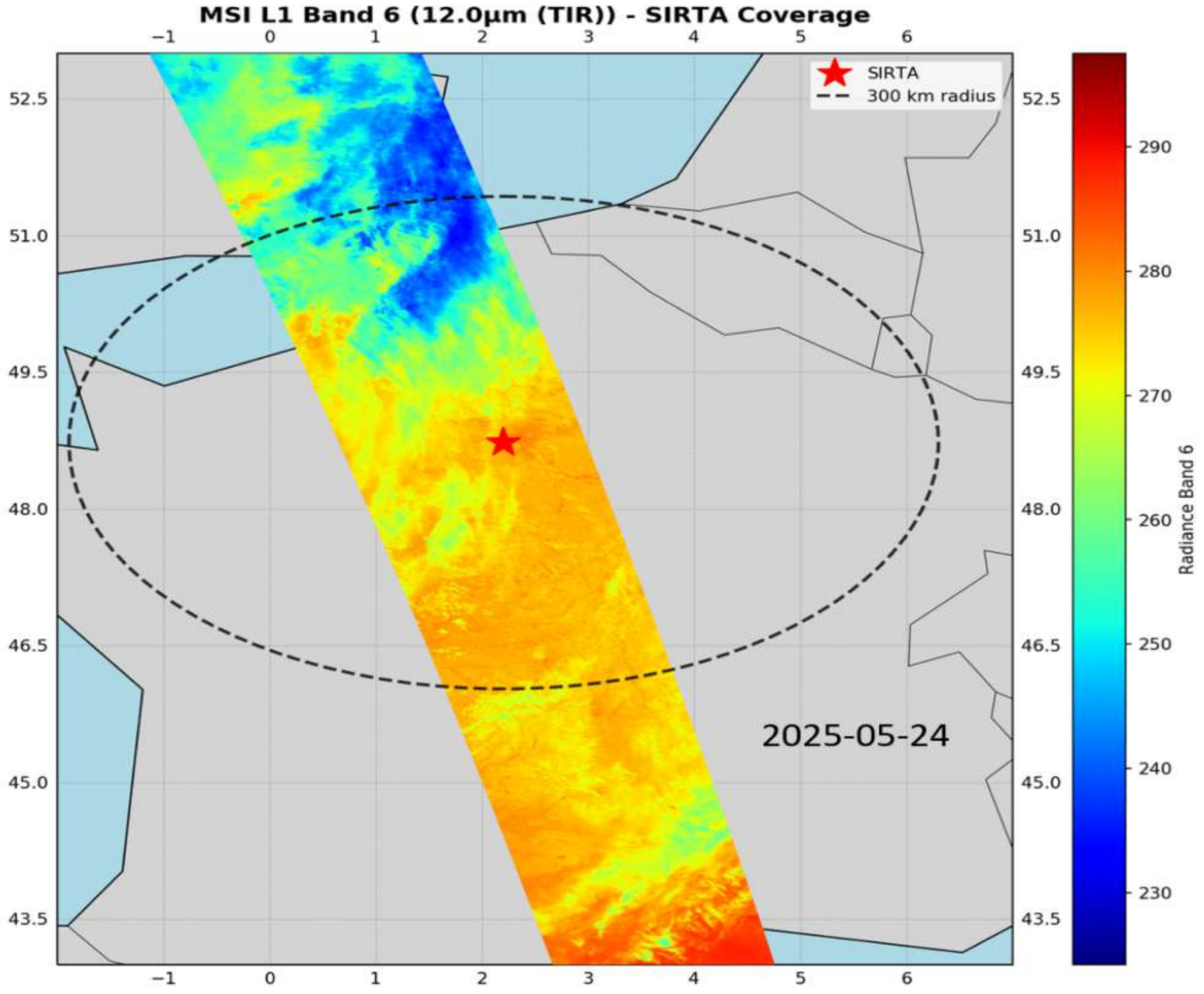
- Optically thin cases work better for all compared profiles
- SR2-based comparisons yield higher CSs than SR1-based ones
- Retrieved profiles β(z) yield higher CSs than SR2-based ones
- Daytime CSs are lower than the nighttime ones

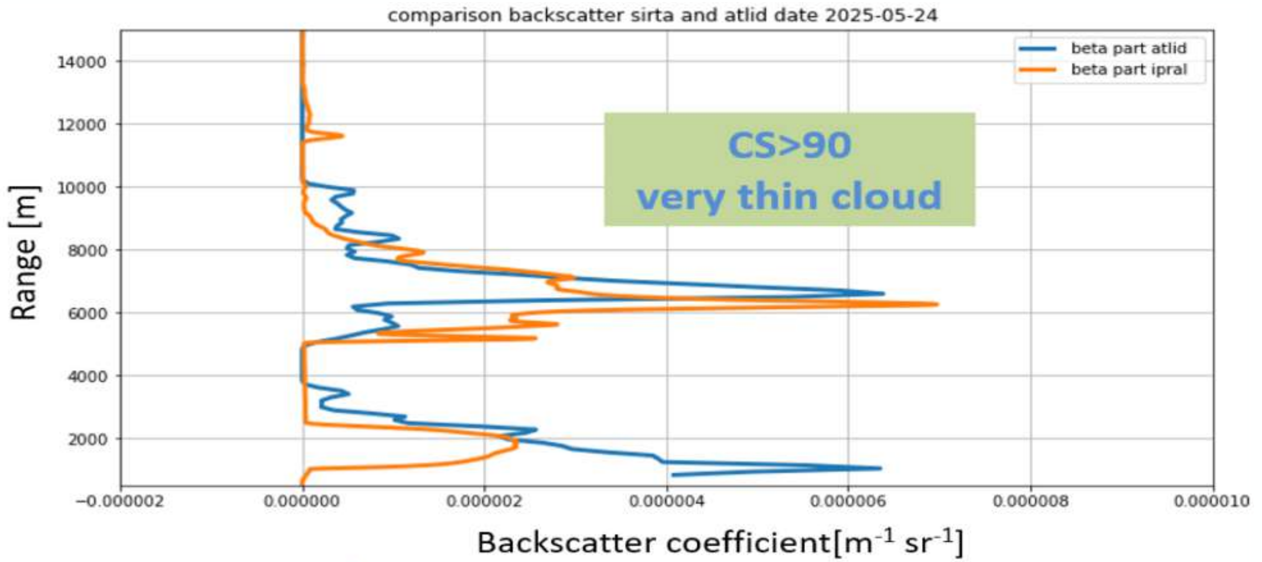
Spatiotemporal variability affecting the comparison

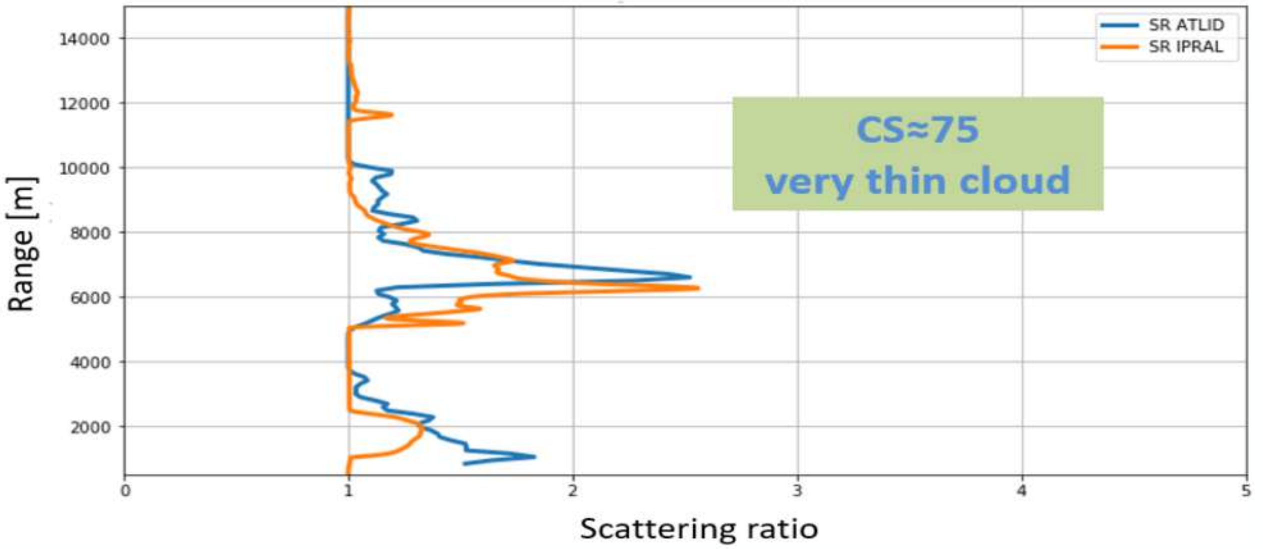




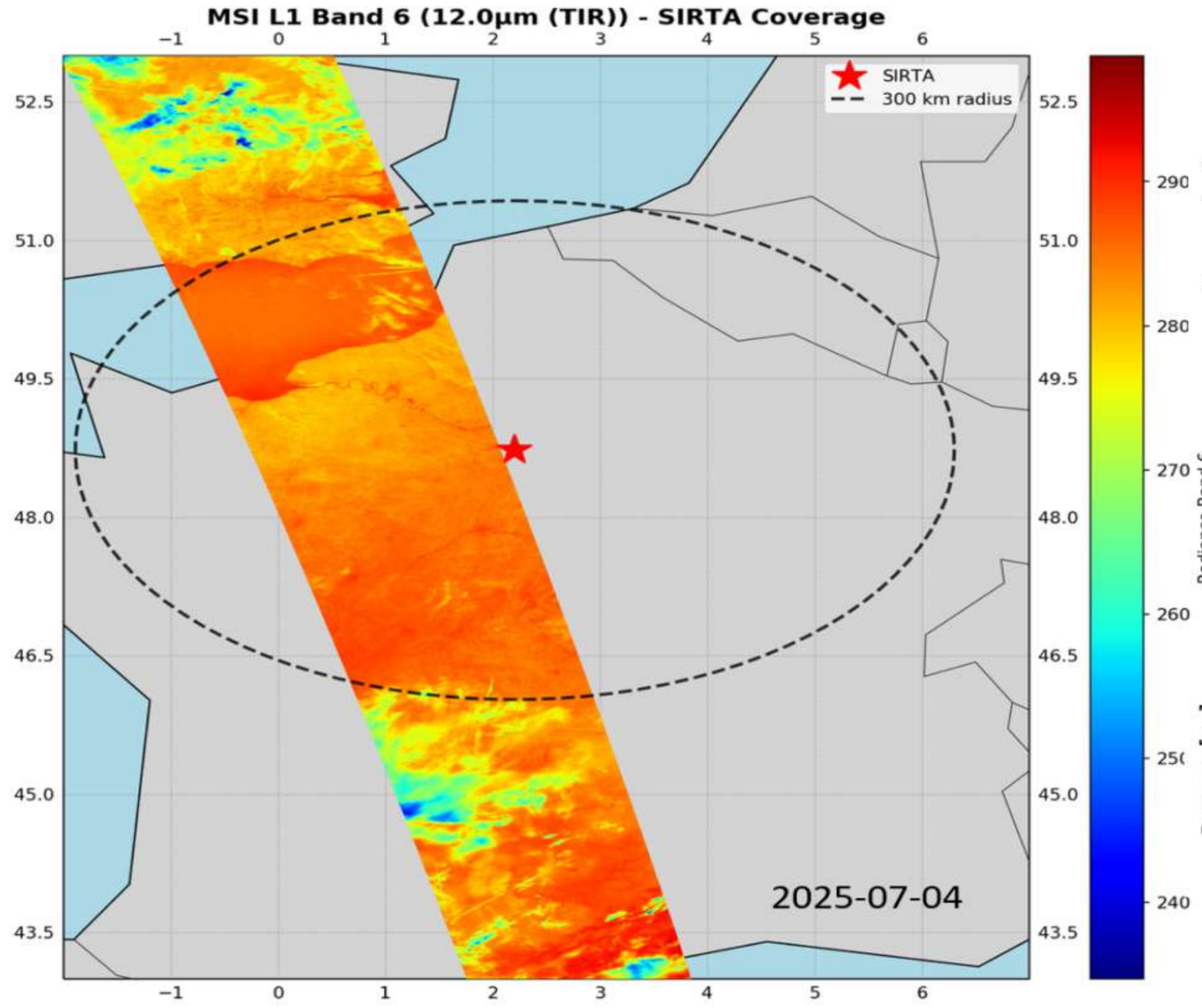
2025-05-24, good CS example

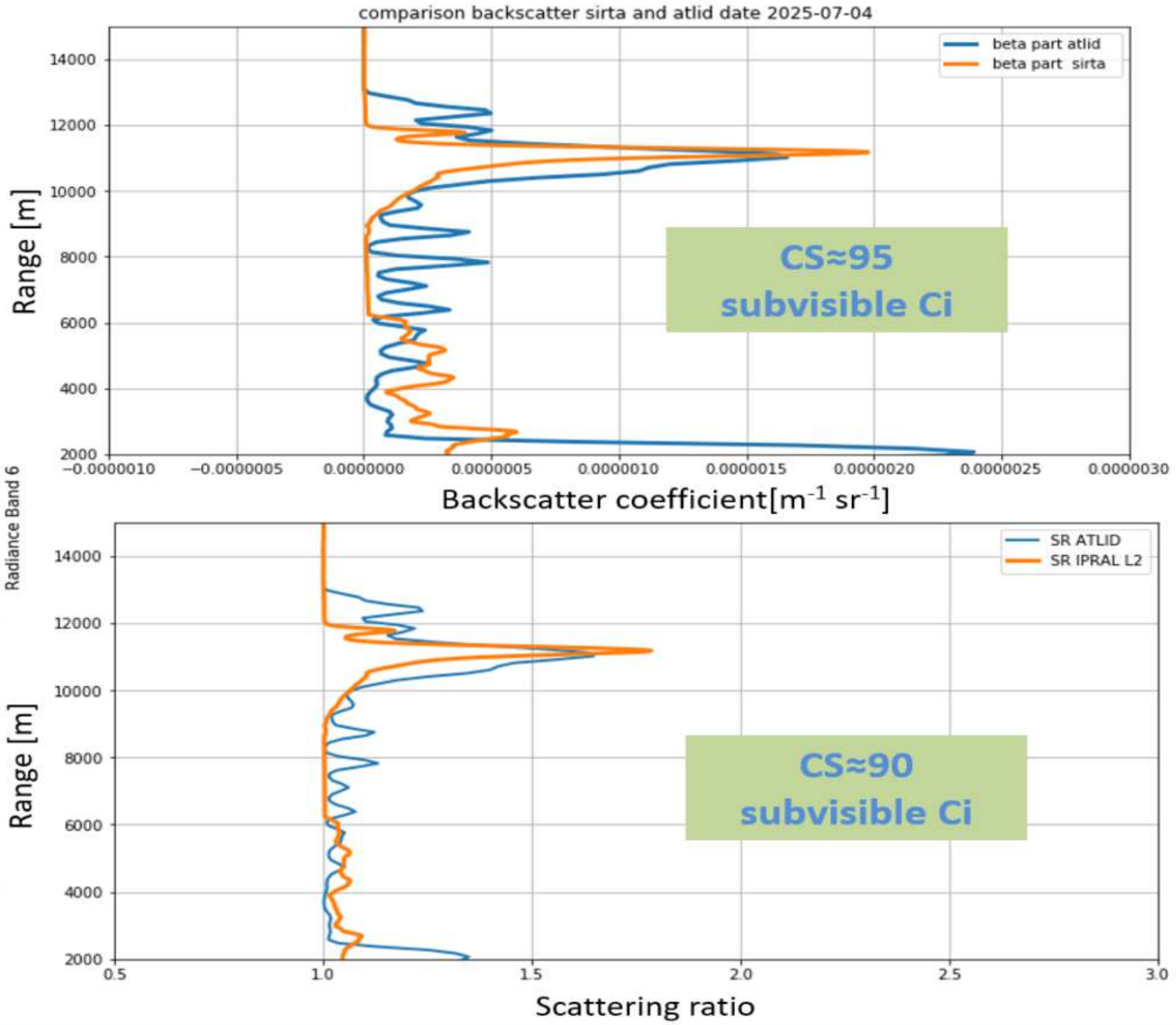




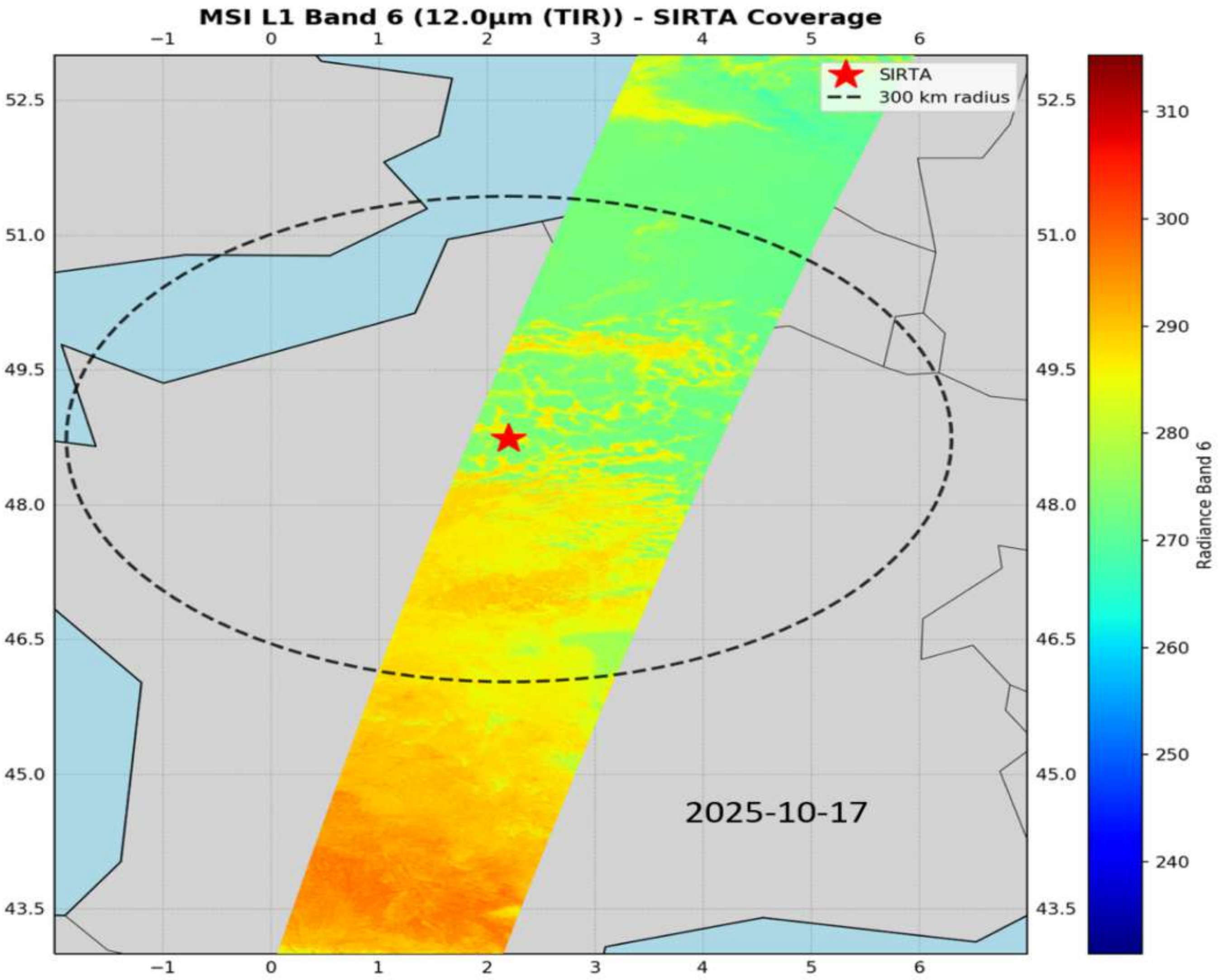


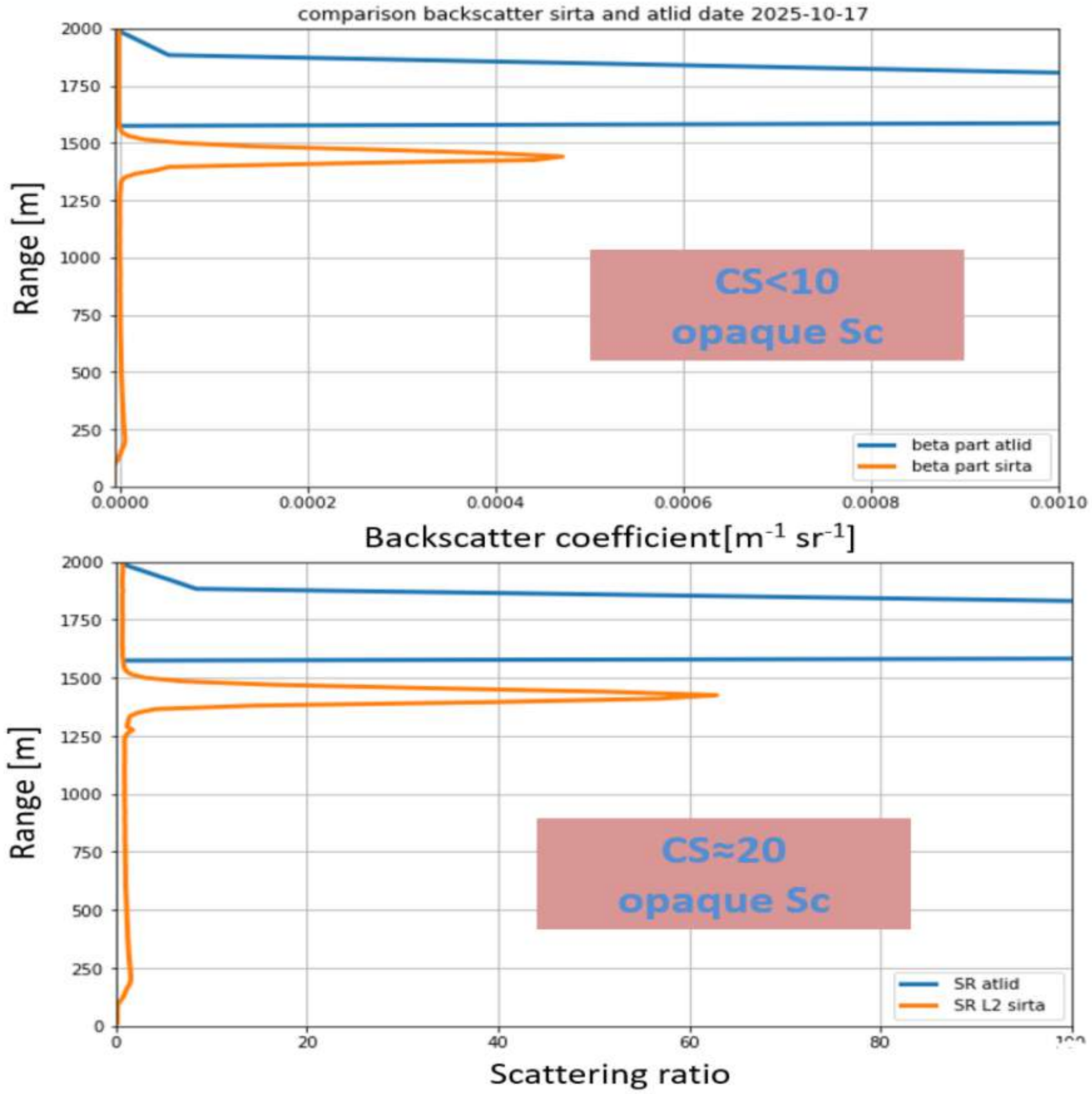
2025-07-04, good CS example





2025-10-17, bad CS example





So, will the validation succeed?

- Colocalizations with low CS are not comparable (70% of cases, including clear sky)
- If the cloud is opaque or multi-layer – NO
- Colocalizations with high cloud inhomogeneity (15%) - are not directly comparable
- The closest profile is not always the best profile for colocalized data
- Geostationary data analysis will help tracking the comparable volumes
- Colocalizations with CS > 60 (subvisible to semi-transparent) – YES.

  For IPRAL/ATLID colocations:

  - for β, RMS of the peak = ±10%

  - for SR1, RMS of the peak = ±8%

Adds on yesterday's discussion on cloud phase obs & validation

G. Cesana, H. Chepfer, D. Winker, B. Getzewich, X. Cai, O. Jourdan, G. Mioche, H. Okamoto, Y. Hagihara, V. Noel, and M. Reverdy (2016), Using in situ airborne measurements to evaluate three cloud phase products derived from CALIPSO, J. Geophys. Res. Atmos., 121, 5788–5808, doi:10.1002/2015JD024334.

Abstract: « …. main differences between the CALIPSO cloud phase products stem from the cloud detection (horizontal averaging, fully attenuated pixels) rather than the cloud phase determination procedures. »

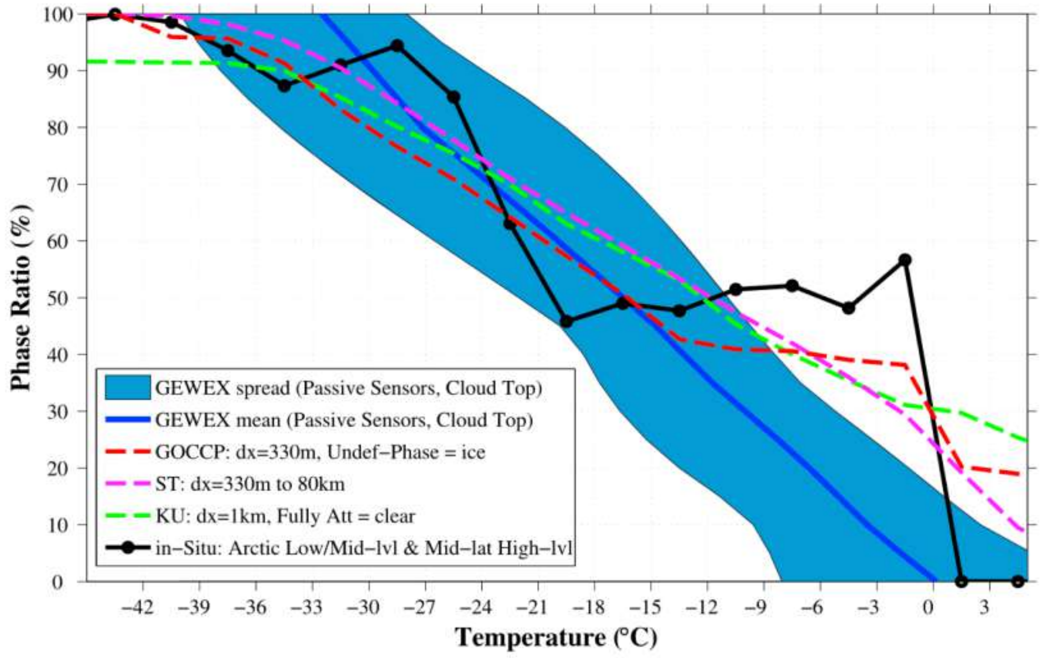


Figure 6. Phase Ratio (%) as a function of the temperature (°C). Red, green and magenta dashed lines correspond to GOCCP, KU and ST relationships, respectively, using 1 year of monthly data (2008 for GOCCP/KU and September 2006 to August 2007 for ST). The blue line and area are the mean relationship using five passive sensor satellites from the GEWEX-CA experiment and its spread. The black line with circle corresponds to the relation using measurements of nine in situ flights.