

General comments (initial paragraph evaluating the overall quality of the preprint):

This manuscript is a relevant piece of work paving the way for an effective design and usage of LSR in future lidar missions. It contributes to the expansion of the database for the research field of surface reflection properties from only passive to also active instruments. The results on the excellent sensitivity of Aeolus LSR's to white surfaces (snow and ice) are very promising for future studies. The manuscript addresses relevant scientific questions and presents a novel concept and data resulting in substantial and reasonable conclusions. The objectives of the paper are well described at the end of chapter 1. The method is described to a sufficient extent, so that reproduction by fellow scientists seems feasible. The structure of the overall presentation is reasonable, but the language might need some revision in certain paragraphs. Own new contributions are indicated and proper credit to related work is provided including a good number of references of high quality. The amount and quality of the supplementary material seems appropriate.

I recommend the acceptance of this manuscript after the necessary revisions. Specific minor comments and unfortunately a large number of technical corrections are suggested below.

Specific comments (addressing individual scientific questions/issues):

1. Line 1: Why is "Phase-F" mentioned so prominently in to paper title? Isn't "new" enough information? As "Phase-F" appears only one more time in the whole manuscript, I recommend to delete this term from the title.
2. Line 10: The term "incidence angle" might need additional reference information (incident onto what?). Maybe consider using the term "off-nadir" angle. (Further cases in lines 67, 97, ...)
3. Line 22: Why are the reference orbits only from the period where Aeolus used its laser A? What about data from laser B with L1B 7.12? Please mention this including the used baseline version. → See also line 107 and Table 1. Are different results or performances expected from FM-B data?
4. Line 36: It might be worth noting shortly here that Calipso measured at a different wavelength than Aeolus.
5. Line 67: What do you refer to here specifically with "unique Aeolus setup"? In which way does it constitute a challenge to retrieve robust LSR. → Presumably you mean the coarse vertical resolution.
6. Line 120: Table 1. As you introduce the "reprocessing product" in line 107, it would be good if you add the Baseline to the column "Version" at least for the L1B.
7. Line 139: Could you provide a reference that shows how real surfaces can behave in terms of angular dependent reflection?
8. Line 149: Why do you perform a resampling from 0.25° to a coarser resolution of 2.5°?
9. Line 165: Why can't NDVI become negative (see equation in line 164)? Is the absolute value of NIR always equal to or larger than that of VIS?
10. Line 187: To my knowledge the Aeolus mission uses the ACE-2 DEM: <https://sedac.ciesin.columbia.edu/data/set/dedc-ace-v2>.

11. Line 189 ff.: To my knowledge the Aeolus ground detection looks for signal drops going upwards (in terms of altitude) first and then looks for signal drops going downwards.
12. Line 200: Northern or southern hemisphere winter in general, or both? Or do you refer to cases b and c?
13. Line 226: Assuming that the ground signal (disregarding the atmospheric contribution to the signal of the bin) emanates from a more or less infinitesimally thin layer, why shall it make sense to multiply here with the thickness of the whole bins (apart from getting a useful unit)? Doesn't the variation of the range-bin-thickness (0.25 km – 2 km, line 97) inadequately affect the LSR'?
14. Line 365: Can sea ice already be expected close to these latitude thresholds, e.g. around +/-40°?
15. Line 367: What should be the reason for a low amount of sea ice right after the southern hemisphere winter?
16. Line 429: Do you have any idea what could be the reason for the ocean returns being place exclusively in such regions? Algae (might explain the seasonality), white caps (maybe rather the cyan values around -60° latitude?), microplastics, ...? It might be worth putting the info from line 474/476 already here?
17. Line 440: Figure 6, top left panel: Could you comment on the apparent increase of valid data in Africa (Sahel zone, region south of African rainforest, Mozambique), South America (north and south of the rainforest) and South East Asia from September to November? Is it related to the general unavailability of data, or could it have a meteorological (or another) reason, such as the movement of the ITCZ or cloud climatology? Or are lines 453-457 sufficient for explanation?
18. Line 460: Did you encounter a decrease in the number valid LSR data over the FM-A period correlated with the decrease of the laser energy of Aeolus?
19. Line 466/467: To me this doesn't look localized, but rather like a clear trend for the whole arctic region for these three months, as also Greenland shows this behaviour. But can a potential "wetting" of the ice/snow with warmer temperatures be consistent with increasing LSR? By melting snow leaving behind the below ice surface?
20. Line 490: It would be good to have this rough classification and allocation of LSR values to surface types already before Fig. 6.
21. Line 498/499: Why only the Southern Ocean? I see a lack of valid signals over all ocean areas between +/-70° lat.
22. Line 504: I could well imagine here a 4-by-4 plot with maps of the number of valid Aeolus observations per grid point, in order to get an idea of the distribution and significance of the comparison for certain regions.
23. Line 524: What's the reason for showing the y-axes in log scale, unlike for Fig. 11+12?
24. Line 814: I would recommend to also explicitly mention the excellent sensitivity of Aeolus LSR to white surfaces (snow/ice conditions) in your abstract.

Technical corrections (typing errors, etc.):

1. Line 12: Although well known, the first appearance of the abbreviation "UV" could be written out.

2. Line 24: Assure consistency for LERG and LERT naming with (l. 137 ff, l. 518 ff., 540 ff., 572, 604 ff.) and without subscripts.
3. Line 33: The “we demonstrated” seems to be either a left over or misplaced (parentheses?) to me.
4. Line 49: ... at the lidar ...
5. Line 100: I assume the either “both” or “two” is redundant.
6. Line 111: As you explain in line 114 that FM stands for feature mask, you might also want to state what PRO stands for.
7. Line 119: described
8. Line 150/151: no comma needed between “estimates” and “reflected”.
9. Line 152/153: Because of the use of “although” this seems to be an unfinished sentence.
10. Line 176: ... where the DEM ...
11. Line 179: ... from the ground ...
12. Line 182: Comma after “Aeolus bin” really necessary?
13. Line 188: Is “sought” correctly used here? Alternatively: “searched”, “looked for” or “sought for” (also line 176).
14. Line 189: ... below the height of the DEM ...
15. Line 189: Proposal: “The algorithms searches for the highest bin that contributes to the ground signal, starting from the first bin with non-negative valid useful signal.”
16. Line 191: Proposal: “... in upward direction.”
17. Line 191: lowermost
18. Line 199: What shall be expressed by the “ - “ between “potentially” and “noise”?
19. Line 222: ... and the highest ...
20. Line 226: Eq.1 should most probably start with “LSR’ ” instead of “y,”. Otherwise introduce y’ in the text here, as done with y before Eq.4.
21. Line 226: Eq.1 – missing explanation of z_i
22. Line 244: ... the optical depth we ...
23. Line 250: overestimation
24. Line 250: probably two spaces after “situated”
25. Line 255: ... the given ??? ...
26. Line 263: aerosol
27. Line 266: two commata after “(HSRL)”
28. Line 270: RayOD \rightarrow OD_{Ray}
29. Line 271: deleted the second “both” in the same sentence
30. Line 280: ... of the LSR ...
31. Line 281: Please explain/write out AB already here than in line 329.
32. Line 295: probably two spaces after “echo”
33. Line 318: ... calculating the scientific ...
34. Line 320: probably two spaces after “year”
35. Line 324: Fig.2 contains “L1A” in the top box, although this has not yet been mentioned. Perhaps you mean L1B.
36. Line 330: orbits
37. Line 330: Proposal: “... with colored indexed for the selected observations placed in each subplot.”
38. Line 334: ... or weakened the LSR ...
39. Line 336: qflag

40. Line 340: ... having the highest ...
41. Line 348: ... ground bins ...
42. Line 352: Fig.3 could have a more expressive design, e.g. with filled, colored and semi-transparent background boxes.
43. Line 353: (ATB) should be deleted here, because it defines a second acronym for a quantity that has already been defined before.
44. Line 354: According to the text above case 1 (red frame) has a $qflag = 0\%$.
45. Line 357: Proposal: "Note that the term *index* above each subplot stands for the respective orbit number of the Aeolus mission."
46. Line 363: This is not a full sentence.
47. Line 383: Would be nice to have all four plots with the same x-axis range.
48. Line 384: latitude
49. Line 370/371: There is no error propagation mentioned in the caption of Figure 4.
50. Line 374: Within the parentheses the values for water seem to be missing.
51. Line 374: ... land are explained ...
52. Line 377: ... are clipped for water surface from ...
53. Line 388: covers
54. Line 392: probably two spaces after "hypotheses" and before "strong"
55. Line 398: probably two spaces after "we"
56. Line 399: probably two spaces after "Several"
57. Line 440: The caption mentions panels a-f, but the letters do not appear anywhere.
58. Line 429 (same for line 460+477+503): The title textboxes of the lower four plot seem to overlap with their above plots. Resolution of the figure could be improved. What about the information on the top left? Lat/Lon info for the y/x axes would be helpful.
59. Line 490: Proposal: ... representing mostly sea ice ... (red colour appears also elsewhere)
60. Line 533: Is this a second caption to the same table? Please unify the information.
61. Line 540: LER instead of LERs
62. Line 542 ff.: Please insert the values for the correlation coefficients also in Figures 11 and 12, as done for Fig. 13.
63. Line 547: ... along the y-axis ...
64. Line 548: Proposal: "various snow types and conditions" instead of just "snow"
65. Line 548: ... by a strongly ...
66. Line 553: ... such an experimental ...
67. Line 556: ... is increased to $r =$...
68. Line 564: Fig. 11: Not a good resolution. Labels of y-axes overlap with colour scale of neighboured plots. Caption: LER should be LER_G or LER_G or LER_G. Please make consistent with x-axis labels.
69. Line 573: probably two spaces after "the"
70. Line 575: delete "comparisons"
71. Line 576: Does a second "agreement" make sense here?
72. Line 579: Fig. 12: Not a good resolution. Caption: LER should be LER_T or LER_T or LER_T. Please make consistent with x-axis labels.
73. Line 586: Fig. 13: Panel a and b are mentioned in the caption but are not noted anywhere in the figure. → use left and right?
74. Line 598: ... from an ecosystem ...
75. Line 599: ... and a correlation ...

76. Line 601: probably two spaces after “was”
77. Line 605: Fig 14c refers to the Sahara region, not to snow covered Scandinavia (a).
78. Line 607: Where do you see the ~0.3 of LER(G/T?) in Fig. 14a?
79. Line 608: “Northern Canada” ... not shown here
80. Line 610: Where are these values for Central Greenland in Fig. 15? →
81. Line 611ff: Do these values match the colours in Fig. 15, or ist the colour scale wrong?
82. Line 614: probably two spaces after “changes”
83. Line 619/620: Grammatically incorrect first part of the sentence.
84. Line 624: very god agreement only too LERT but obviously not to LERG
85. Line 655: Fig. 14. You show LERT and LERG, but mention only LERG in the labels of the y-axes.
What about the availability of error bars for LERT?
86. Line 625: Fig. 14a refers to Scandinavia but not to Sahara
87. Line 625: sensitive → sensitivity
88. Line 634: Fig. 14a refers to Scandinavia but not to Sahara
89. Line 638: ... of the southern ...
90. Line 643: 15 d → 14 d
91. Line 648: Refer to Figure 14 e.
92. Line 650: 15 f → 14 f
93. Line 714: Fig. 16 caption: ... during the entire ...
94. Line 714: Fig. 17 is clipped at the top. Caption: ... during the entire ... colorbar ...
95. Line 695: This is not a fully sentence in a grammatical sense. Also “moreover” doesn’t make sense here as you already referred to the pattern in the previous sentence.
96. Line 699: Fig 16 – 17 – 18: Bad resolution → hardly readable label of y-axes. Please use capital letters for “LSR” in the labels.
97. Line 703: Please explain the value 0.05 here. What does it mean and where do you get it from? Is it 5% of the area covered by snow?
98. Line 705: Fig. E4 → Fig. S8
99. Line 706: No need to refer to the same figure again.
100. Line 734: Fig. 18. Not well arranged with partly mutual clipping of the subplots. Does “snow cover cases” represent a count/number here? If so, it needs a second y-axis. In Figure 16 it is “snow cover” only. Otherwise formulate the sentence more clearly. Think of a more appropriate label for the x-axes or add a second one for SNW. Are the grey bars error bars?
101. Line 727: ... in our previous paper ...
102. Line 729: level
103. Line 731: ... in the discussion chapter in ...
104. Line 740: What do you mean with “during post-commissioning phase”? Doesn’t the commissioning phase of a satellite comprise only the first few weeks to months? If so, this would be long ago for Aeolus, but you seemingly want to express that your LSR product is to be placed into the Level 2A product in the near future, which is already after the end of the Aeolus mission.
105. Line 751: ... in the FM-A ...
106. Line 760: ... different magnitudes of return signal ... effects in the returns ...
107. Line 768: Decide for one consistent representation of the ranges, either “ $x - y$ ” or “ $x < LSR < y$ ”
108. Line 791: ... to a low ...
109. SUPPLEMENT Line 19: What does “grind” mean in the x-axis labels?