General comments

This paper presents a comparison of lidar signals and the derived aerosol parameters between 6 different lidar systems using a inter-comparison strategy recommended by EARLINET. The objective is to evaluate the performance of different lidar systems and quantify their consistency. Such work is important to the scientific community and has to be done for lidar networking. However, such comparison is not easy to organize, because of the transportation of lidars, arrangement of the campaign, the strategy of data comparison and so on. The subject fits very well the scope of <*Atmospheric Measurement and Technique*>. It is a pity that (at least it seems) the campaign did not last long, since I did not see measurements on multiple days and with different aerosol loading, clean atmosphere and cloudless conditions. The presentations of results are well done and results are well analyzed.
As to academic writing, I suggest the author and co-authors spending more time on polishing the manuscript before getting it published. Be careful with the following 3 points:

- Remember to not use long sentence if you are not sure you can handle the grammar and logic in complicated sentences.
- Avoid colloquial expressions. In particular, be careful with the words such as 'maybe', 'probably', 'might' and so on. There are quite a few expressions like this throughout the manuscript. It is OK to have open questions or unresolved issues in a scientific paper, but you should state clearly what the causes/considerations are. Otherwise, it gives the readers an impressive that the results or conclusions are vague and informal and cannot be trusted.
- Make it shorter. I would expect this paper to be concise and compact. The author should focus on the main messages that he/she wants to address to the readers.

Moreover, I see it not necessary to have a separate section 'Discussion' with only the comparison of dark measurement. The discussion part is supposed to analyze and interpret the results of the study, but it is not the case here.

Specific comments

P1L13-15: "In the end, the lidar systems were quality assured, of which the standard deviation of range-corrected signal can be controlled within 5 % at 500-2000 m while 10 % at 2000-5000 m." → One example for consideration "In the end, the lidar systems passed the quality control/assurance, ensuring that the standard deviation of range-corrected signal could be controlled within 5% at 500-2000 m and 10% at 2000-5000 m"
P2L29: '~~a~~ large ground-based lidar networks': remove 'a'
P2L45: '~~While~~ to achieve', remove 'while'
P2L51-52: "Since European lidar technology was independently developed at different stations from different countries,"-> "Since lidar systems in EARLINET were developed independently at stations in different countries.." .
This sentence is not clear-- "the devices and algorithms used are not the same (D'Amico et al., 2015)". EARLINET has its requirements and criteria for all the affiliated lidar stations, as well

as data processing algorithm, the Single Calculus Chain. I am not sure what devices and algorithms you are referring to.

P3L80: reference missing
P3L81: assessment of -> assess
P3L83: "At present, China is starting to build…" This sentence not clear, please revise it. Here is an example, for your reference-- "Currently, China is in the process of building a … network, which may lead to isolated and one-sided measurements from each observation station, therefore…spatial-temporal correlation…"

P3L86-87: "With the gradual shift from qualitative measurement to the qualitative application of atmospheric lidar, ... high, and direct mutual … level, China has" → As atmospheric lidars are shifting from qualitative to quantitative applications, …, increasingly high. Therefore, direct mutual…level. China has…

P4L95: in the other region → in other regions/countries.

P4L99: This sentence is too long and lacks of clarity, please revise it. "Based on the lidar inter-comparison observation campaign on September 2021 in the south of Beijing observatory, this paper introduces the lidar quality assessment strategy based on experience of EARLINET on self-calibration and inter-comparison methods for systematic improvement of lidar hardware, and evaluates the reliability of the 1064 nm channel of many sets of lidar systems, analyzes the deviation of the Mie-Rayleigh signal and its influence on the backscatter coefficient."

P4L100: ~~relatively~~, …., single wavelength -> single-wavelength
P4L105: ~~on the results~~ on the lidar signal….
P4L110: , which was -> . It was
P4L111: Cirrus was… and covered
P4L112: in most of the time
P4L113: feature -> featured
P4L114: Maybe it is more accurate to say "emission and reception modules" instead of "transmitting and receiving modules ".
P4L121: The 1064 nm light -> the backscattered 1064 nm light
P5L127: of which→ on which
P5L129: In this sentence "The first part is self-validation or calibration according…", what are the difference between (self-) validation and calibration? If these two terms are referring to the same thing, there is no need to use different terminology here, it could confuse the readers. The authors should also check if there are confusions of terminology elsewhere in the manuscript, such as self-test, self-validation, calibration, verification… and so on.

P5L131: CMA not defined
P5L131-134: "With the CMA's goal of promoting the use of lidar instruments and their data among the Chinese lidar network, the inter-comparison at the hardware level was made, in terms of range corrected lidar return signals inter-compared directly, and also the inter-comparison of aerosol backscatter coefficient at 1064 nm retrieved by each lidar system was

performed in this study." This sentence is not correct in grammar, please think of breaking long sentence into shorter ones for simplicity and clarity.

P6L137: "so that makes it …."-> This uniformity in data collection makes the inter-comparison easier.

P6L144: of each system. Again, the sentence is too long and contains fragments.
P6L147: due to the difference in each system efficiencies -> due to different lidar efficiencies/ transmissions

P7L150: Please specify what λr, λ0 and I are.
P7L151: please revise this sentence "It should be compared with simultaneous observation, and continuously collecting the original data for at least 180 minutes,
and selecting a period of no less than 30 minutes where the aerosol vertical distributions are stable."

P7L152: "The rang-square-correction signal" what is it? Range-corrected signal is commonly used in the lidar community.

P7L157: No. Ref lidar is confusing, better call it reference lidar for clarity.

P7L170: I do not see the link between signal saturation and fitting molecular signal to RCS. please explain.

P7L172: "…a good signal-to-noise ratio except No. L05 lidar adopted…". Break into 2 sentence "…a good signal-to-noise ratio. However, No. L05 lidar adopted….~~at such range~~"
P7L174: typo: a space missing 'than' and '5%'
P8L177--L179: Could you give more comment about the data quality of lidar No.05? Is it possible to correct this electronic noise? Could this mismatch with Rayleigh fit be caused by other reasons, for example, the divergence of laser beam? Have you ever got better Rayleigh fit from this lidar?

P8L180: Better to write "In this test, lidar profiles in every 30-minute time intervals were averaged."

P9 Figure 4: To me, raw lidar signal means it is without any correction, such as deadtime, bin shift, etc. In Figure 9, I guess corrections have been applied.

P10L193: This is a bit confusing, please rephrase it "which also means they can be used to observe the vertical distributions of aerosol and cloud without knowing the determining the amounts".

P10L194: quantitative analysis→ quantitatively analyze
P10L194: it is not proper to call it 'accuracy', better to use 'difference'

P10L195: overlap properties may cause misunderstanding, you can simply say 'overlap ranges'
P10L197: "In this investigation, single…..was selected"→ "Lidar observations between 18:30 and 19:00 CST from each lidar system were averaged for inter-comparison. "

P10L198: From the quicklooks, the readers can see the presence of aerosol layers in the range of 1500-2000, why did you normalize the lidar signal in the aerosol layer although you have mentioned in the previous sentence that the range 2000-5000 was expected to be clean? What if you normalize at higher altitude, should the results be better?

P11L202-L203: A few grammar mistakes in this sentence. Please rewrite it, here is one example "Due to the significant differences in the incomplete overlap region between different lidar systems, large relative deviations were observed within the 500 m range (Figure 6, b). As a result, a meaningful comparison cannot be made"

About P10 Figure 5:

- The ticks of Figure 5(c,d) are partly cut off, please replace with complete figures
- The structures within boundary layer looks similar in all quicklooks except for Figure 5(e), could you comment on that?
- Where do the repeating white stripes in Figure 5(b) come from? Is it a technical anomaly?
- In Figure 5(a) and 5(c), we see ripples in the transported aerosol layer below 3000 m from 19:00 to 23:00, can you comment on that?

P11, P12: Figure 6(c) and 7(c) do not carry so much information, therefore not necessary.
P13: Please write how you determine system noise and random noise and put references

The reviewer stops listing corrections of English writing from Page 11 and advices the author and coauthors to correct it by themselves.

Figure 8: (1) The scales of Y-axis in Figure 8(a) and Figure 8(b) are so different, making it difficult to conclude. (b) For Figure 8(b), did you take multiple dark measurements for L05 in order to check if such noises are stable or not?

P14L267: "we found … also such lidar system has a possibility to be miniaturization". It is not appropriate to draw such conclusion, because this paper does not talk about the miniaturization.