Reply to Zender (Referee)

We are grateful to the referee for his constructive and thorough criticism and suggestions to our manuscript. Please find below a detailed point-by-point reply (*referee's comment in italic*).

General Comments

This manuscript presents a new lossy compression algorithm called "Digit Rounding" (DR), and evaluates its performance against and with other lossy and lossless compression algorithms on idealized and remote sensing datasets. The manuscript addresses the growing need to archive meaningful data rather than noise, and to do so reliably and quickly. The study presents an original advance in lossy compression whose implementation unfortunately hampers its utility. The study is understandable yet poorly written. This potentially useful study of lossy compression techniques needs a thorough overhaul before publication.

We will improve the writing of the manuscript by contacting a native English speaker/writer. As suggested by the Anonymous Referee #1, the paper will be reworked to highlight the main contributions of our work and focus more on the application to CFOSAT and SWOT datasets.

Specific Comments

Originality: DR is an improvement on "Bit Grooming" (BG) which I invented as an improvement on "Bit Shaving". In that sense I am qualified to comment on its originality. The heart of DR is essentially a continuous version of BG: Whereas BG fixes the number of bits masked for each specified precision, and masks these bits for every value, DR recomputes the number of bits masked for each quantized value to achieve the same precision. BG did not implement the continuous method because I thought that computing the logarithm of each value would be expensive, inelegant, and yield only marginally more compression. However, DR cleverly uses the exponent field instead of computing logarithms, and so deciphers the correct number of bits to mask while avoiding expensive floating point math. This results in significantly more compressibility that (apparently) incurs no significant speed penalty (possibly because it compresses better and thus the lossless step is faster?). Hence DR appears to be a significant algorithmic advance and I congratulate the authors for their insight.

Thank you for your congratulations. They are much appreciated. Indeed, the speed penalty of DR is compensated by the fact that the lossless step is faster.

The manuscript stumbles in places due to low quality English, and cries out for more fluent editing. Not only is the word choice often awkward, but the manuscript is like a continuously choppy sea of standalone sentences with few well developed paragraphs that swell with meaning then yield gently to the next idea. GMD readers deserve and expect better.

We will improve the grammar and writing (see first answer). We will also modify the paragraph structures to avoid too small paragraphs and better take care of the transitions.

Does DR guarantee that it will never create a relative error greater than half the value of the least significant digit? BG chooses the number of digits to mask conservatively, so it can and does guarantee that it always preserves the specified precision. Equations (1)-(7) imply that DR can make the same claim, but this claim is never explicitly tested or made. The absence of this guarantee is puzzling because it would strengthen the confidence of users in the algorithm. However, the guarantee must be explicitly tested, because it undergirds the premise that the comparison between DR and BG is fair. In any case, clearly state whether DR ever violates the desired precision, even if that happens only rarely.

Equations (1)-(7) imply that DR guarantees that it always preserves the specified precision. We will explicitly add that claim in the text and show that DR always provides the desired precision on the number Pi with nsd varying from 1 to 8. We will also provide the maximum absolute error on artificial data of 1 000 000 values spanning [1.0, 2.0) in equal-increment steps of 1e-6.

p. 16 L13: "Code and data availability: The Digit Rounding software source code and the data are currently only available upon request to Xavier Delaunay (xavier.delaunay@thalesgroup.com) or to Flavien Gouillon (Flavien.Gouillon@cnes.fr)." The GMD policy on code and data is here: https://www.geoscientific-model-development.net/about/code_and_data_policy.html. This manuscript provides no code access nor explanation, and no dataset access, and thus appears to violate GMD policy in these areas.

The code and the datasets will be made publicly available on the CNES gitlab.

Common comparisons would help build confidence in your results. It would have been more synergistic to evaluate the algorithms on at least one of the same datasets as Zender (2016), which are all publicly available. I am glad the authors used the publicly available NCO executables. Why not release the DR software in the same spirit so that the geoscience community can use (and possibly improve) it?

Comparisons with BG will be provided on the same MERRA dataset used in Zender (2016). The DR software will be released under MIT-style open source license.

The lossless and lossy compression algorithms analyzed seem like a fairly balanced collection of those most relevant to GMD readers. Most methods that were omitted are, to my knowledge, either non-competitive (e.g., Packing) or not user-friendly, e.g., research grade but not widely available (e.g., Layer Packing) and too hard to independently implement.

Table 6 on p. 19 shows the maximum absolute error (MAE) of BG is quite similar to DR, as I would expect. However, Table 7 on p. 20 shows the maximum absolute error (MAE) of BG is nearly 10x less than DR. Why are the MAEs similar for dataset s1 and significantly different for dataset s3D? I expect DR has a greater mean error (and lower SNR) than BG due to the algorithms, yet the difference in MAEs surprises me. Zender (2016) Table 3 shows that BG is tuned to have an MAE just shy of violating the precision guarantee. An MAE that is nearly 10x larger seems like it might violate the precision guarantee.

These results show that BG can sometimes be too conservative. As shown in Table 1 on the value Pi, BG sometimes preserves more bits in the mantissa than what is strictly necessary to achieve the required precision. This is what happens on the dataset s3D. On the contrary, DR adapts the quantization step to each value of the input dataset. Doing so, it can achieve the required precision while preserving less mantissa bits than DR does. This results both in a higher mean absolute error and in a higher MAE than BG. This explanation will be added to the text.

The preceding comment is a request to more carefully analyze the underlying cause of the behaviors reported in the data. The next two comments are to report more results to deepen the analyses and explain the behavior of DR more robustly.

Please include the maximum absolute error or maximum absolute relative error (which normalizes the error by the original value) to Tables 5–10.

MeanAE is an important statistic that is complementary to MaxAE. MeanAE is the average absolute (no compensation between positive and negative) bias in the dataset, and is more familiar and relevant than SNR to at least some geophysicists. Please consider including MeanAE in Tables 5–10.

As suggested, the maximum absolute error and the mean absolute error (MeanAE) will be added to the tables allowing deeper analysis of DR behavior.

Zender (2016) and Silver and Zender (2017) consider four primary criteria to evaluate compression algorithms: Compression Ratio, Accuracy, Speed, and User-friendliness. This manuscript neglects explicit consideration of the last, though usability seems (in addition to performance) seems to be an implicit reason why they recommend BG not DR for the "real world" use cases in Sections 5.1 and 5.2. The manuscript would benefit from a more explicit consideration of usability throughout. Examples include software availability, flexibility, and complexity of invocation, as well as transparency (will users have all the necessary software required to read the compressed data?), and instructions to mitigate these issues for DR.

As for BG, there is no "decompression" associated to DR. DR does not require any software to read the rounded data. This argument will be added into the text. The reason why BG is recommended rather than DR for the compression of CFOSAT dataset in section 5.1 is that this dataset is compressed in absolute error bounded compression mode. DR only works for relative error bounded compression mode. Nevertheless, some results using DR on this dataset will be provided for completeness. In section 5.2, BG (in the absolute error bounded compression mode) is recommended rather than DR for the compression SWOT L2 pixel cloud product. This recommendation is based on the compression ratio obtained. We will add the maximum absolute error and the mean absolute error (MeanAE) to Tables 8 to 10 for fairer comparisons. Moreover, we will provide a supplement to the article with the commands and datasets necessary to reproduce the results.

Tables 1 and 3 follow Tables 1 and 2 of Zender (2016). This should be noted in the text and/or caption of the tables.

The reference to Zender (2016) will be added in the caption of Tables 1 and 3.

It seems like Table 2, the algorithm description, should be a figure rather than a table.

This will be corrected as suggested.

The manuscript is awkward in that it introduces a demonstrably superior lossy compression algorithm but recommends a different algorithm (BG) for "real world" cases (Section 5), partly because DR is unavailable in software that potential users have easy access to, and its implementation appears to be too inflexible to use on generic datasets. The recommendation of BG not DR does attest to the objectivity of the study, yet it seems to be an unsatisfying conclusion to what was clearly a time-consuming study. In this sense the manuscript seems premature, since if DR were "ready for primetime" then the authors could have recommended it rather than BG in Section 5. Perhaps the authors should re-evaluate whether the manuscript is premature, i.e., whether it should both introduce a new lossy algorithm before it is ready to use in optimized workflows for generic geoscientific data compression.

As previously answered, the manuscript will be reworked to highlight the main contributions of our work and focus on the applications to the CFOSAT and the SWOT datasets. The maximum absolute error and the mean absolute error (MeanAE) will be added to Tables 5 to 10 for fairer comparisons that will allow mitigating the previous conclusions that were based on the compression ratio only. Moreover, some results using DR on CFOSAT dataset will be provided for completeness of the manuscript.

Response: we thank you for these suggestions that will help us to improve the manuscript.