

Comment on acp-2022-389

Corwin Wright (Referee)

Referee comment on "Intermittency of gravity wave potential energies and absolute momentum fluxes derived from infrared limb sounding satellite observations" by Manfred Ern et al., Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2022-389-RC1>, 2022

This manuscript by Ern et al uses data from the SABER and HIRDLS atmospheric limb sounders to quantify and characterise the intermittency of gravity wave momentum fluxes and potential energies in the terrestrial stratosphere and mesosphere.

The paper is well-written, clear and technically accurate, and I would have no objections to it being published in something very close to the current form. I have a few comments, but none of these are major criticisms and are intended to help make the paper stronger.

Main Comments

[1] I do have a probable answer to one question raised in the manuscript. In several places (e.g lines 532-543, lines 657-664, line 730) the authors highlight that their results appear quantitatively inconsistent with a previous study (W2013a, reference below). Specifically, the wave intermittencies measured in the current study are consistently higher than those seen in W2013a. However, I believe this arises from a key methodological difference. In W2013a, we used an approach (described by WG2013, reference below) which selects for multiple waves in a given HIRDLS measurement, and which typically (see W2015, reference below) identifies four discrete waves at any one measurement location. In contrast, the current study identifies at most a single wave at each point (L173). These 'additional' waves in W2013a tend to be lower-amplitude and have smaller momentum fluxes (WG2013, W2015) than the 'main' waves measured by the method used here, and will hence tend to strongly pull intermittency values down. I think this methodological difference is likely to explain most if not all of the differences between the current study and W2013a.

I emphasise that this difference is not an mistake in experimental design in the current

manuscript and I have no objections to the authors using a more cautious approach and identifying at most one wave as they do here - it is a perfectly valid choice. The WG2013 method has the advantage that it detects more smaller waves from the same data, but in particular for HIRDLS can be negatively affected by a known fault with the instrument which will introduce a small height-varying population of nonexistent waves into the data (W2015) and hence would have to be treated cautiously for a study of this type. We had not identified this problem at the time of writing W2013a, but it shouldn't affect the results presented there too much as the effect is very small at the low altitudes that study focuses on, and even at high altitudes is only a few percent of the measured waves - i.e. I suspect that the differences between the current study and W2013a will be almost entirely methodological rather than due to this data issue.

[W2013a] Wright, Osprey and Gille, JGR 2013: 10.1002/jgrd.50869

[WG2013] Wright and Gille, GRL 2013: 10.1002/grl.50378

[W2015] Wright, Osprey and Gille, ACP 2015: 10.5194/acp-15-8459-2015

[2] The methodological choice to normalise the distributions (L238-247) does make sense when the results are considered, but probably needs a little more justification. For QBO regions, where filtering varies strongly from year to year, the logic is clear and coherent, but I am less clear on the justification for doing so at extratropical latitudes (at least outside SSW periods). If this was a minor point I would be happy with the current presentation, but since this choice underpins most of the results presented I think it needs to be justified a bit more strongly.

[3] In section 5.1, I was still confused after several readings as to exactly how the gradient effect (line 444 onwards) was being compensated for - if the normalisation is taking place at the level of the bin, then how does this reduce spatial biasing due to gradients within the bin? I suspect I am misunderstanding something here, and as such would appreciate this section being made clearer.

[4] A minor concern I do have is that the manuscript feels very long and could probably benefit from trimming, but this is not a critical problem and the paper does contain a lot of data which does justify this. If the authors do choose to trim it, I think it would be better to do so by slimming down each section rather than removing some parts entirely, and by reducing repetition between sections.

Additional Comments

L131: note that the resolution of HIRDLS drops sharply above 60km, averaging ~2km above this level - see e.g. Figure 5.1.1 of the HIRDLS Data Quality Document

HIRDLS DQD: https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/HIRDLS/3.3_Product_Documentation/3.3.5_Product_Quality/HIRDLS-DQD_V7.pdf

L148: what kind of high-pass filtering is being applied here? Some types could introduce small waves, which may pull measured intermittencies down.

L201 and other places: while it's reasonable clear that by 'average' the authors mean 'mean', the mode, median and mean are all types of average - I would suggest switching from 'average' to 'mean' throughout to avoid any confusion.

L273: additionally, presumably SABER cannot access the smallest horizontal wavelengths due to the inter-profile spacing being about double that of HIRDLS. This is mentioned in s4.2.2 later, but this is the first mention where that effect is relevant so it might help to put it here instead.

L339 onwards: similar high intermittencies are seen in the open Southern Ocean using AIRS data in Hindley et al 2019 (their figure 9). While this has a very different observational filter, it may be relevant to the discussion here too.

Hindley et al, ACP 2019: doi:10.5194/acp-19-15377-2019

Section 4.2.4 - I like this!

L470: I understand here that you mean individual profiles (i.e. single profiles against altitude), but the text as-written could be read as referring to single levels *within* a profile (i.e. a single altitude level of a given profile) - it might help to clarify this.

L484: This normalisation depends on the averaging method used and more detail about this would help - for example, are the averaged values being stored at the grid-corner or grid-centre before being interpolated to the profile location?

L515 onwards: could the lack of shorter vertical wavelengths and/or different normalisation areas cause this higher Gini coefficient estimate?

L544: this is remarkable given the very different observational filters involved - do you think this is a real similarity and (e.g.) that the same intermittency effects are being

observed uniformly across the GW spectrum despite the very different physical scales, do you think the methods are actually observing the same waves, or do you think it's just a coincidence?