Editor

The revised paper now compares two waveguide diagnostics, and the reviewer is not happy with the discussion and conclusions drawn from the results.
Having gone over the revised paper, I tend to agree with the reviewer on this, that more explicit statements are needed.
Here are some of my thoughts that might help -
there are two main differences between the two diagnostics - one is the zonalization process- a low frequency filtering vs a wave-activity based zonalization of PV contours. The other difference is due to the difference between PV gradients and the stationary wavenumber which is essentially grad(PV)/U. In fact for a zonal mean the relation between the two should be:

Ks^2=-fo*grad(ln(|PV|))/U (Bukenberger et al 2023).

It might help for the comparison to actually calculate this quantity and compare to the standard Ks waveguide, to more directly see any effects of the zonalization method.

As for the direct comparison between using Ks^2 and PV gradients, the two in my mind show different things - Ks^2 shows, for a given zonal wavenumber, the meridional wavenumber that satisfies the dispersion relation, in regions of meridional wave propagation. The relation of the wave amplitude to this is complex. Strong PV gradients indicate regions in which meridional displacements will most directly create strong PV perturbations.

Theoretically I expect both are a proxy of where waves will propagate zonally and of the latitude where we expect most waves to occur at each longitude. The analysis in this paper can more closely and explicitly examine this in real data

Thank you for these comments and suggestions. Overall, we are very grateful to both yourself and all reviewers for comments that have led to substantial improvements to this manuscript. We agree that separating out the impacts of the background flow differences with the impacts of the waveguide definition differences is valuable, and so have added additional analysis on this. To keep the analysis zonally asymmetric, we decided to calculate the PV-waveguides using, instead of the rolling-zonalization method of smoothing, the same time- and zonal-filter that we use for the KS-waveguides. This analysis shows that it is predominantly the waveguide-definition that matters the most, although the background flow is also important, with the rolling-zonalization PV-waveguides showing the strongest correlations with QSWs.

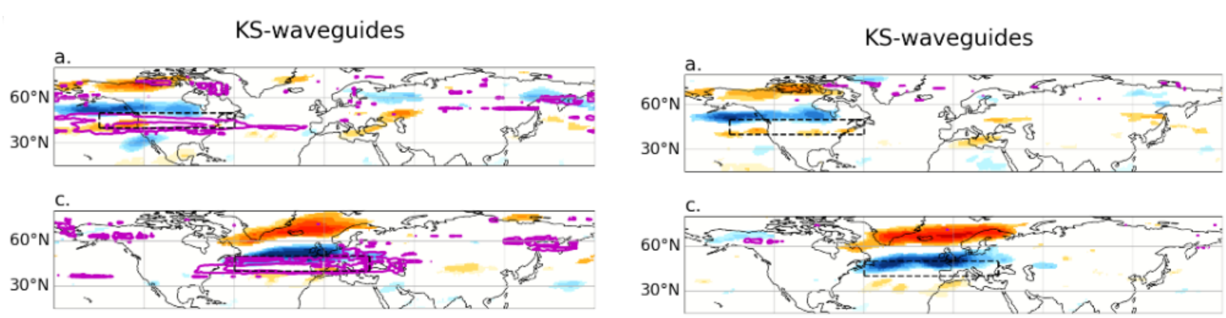
I agree that both methods theoretically tell us something about where we would expect most waves at each longitude, and this is a good phrasing – we have added sentences on this in the introduction, and in the discussion/conclusions.

Reviewer 1

The paper has been deeply modified since the first submission. The paper now contains a systematic comparison between two notions of waveguides: one based on the classical stationary wavenumber, which was present in the initial submission, and the other on potential vorticity gradient that has been recently introduced by Wirth and Polster papers and inserted in the revised version of the paper. I think this comparison made the paper much more attractive and I enjoyed reading it. Additionally, all my comments have been carefully considered by the authors. Therefore, I recommend publication of the paper once the following minor suggestions have been taken into account:

1) The waveguide strength is a new diagnostic on which many figures rely on but its definition is not entirely easy to find. Line 165, the maximum waveguide strength W_s=K_s-k is introduced. But in order to build composites from Figures 4 to 9, the text says (line 218) that the waveguide strength is the sum of all K_s-k for k between 4 and 9. Why did you choose the sum of K_s-k for the composites and not the maximum K_s-k ? Since all the composites rely on that parameter it would be good to introduce a mathematical notation for that parameter. This would help to clarify the captions of Figures 4 to 9. Different wordings are used to say the same thing: in Figures 4 to 7 "strongest waveguide presence" is used, in Figure 8 "strongest average waveguide strength", Figures B1-B5 "strongest ... average waveguides". Please keep the same wording or use a specific notation.

Thank you for this comment. The results are not sensitive to using the maximum W_s versus using the sum/average W_s across wavenumbers (shown below for a sample of the plots). Because waves can exist in a range of wavenumbers, and not necessarily in the wavenumber coinciding with the maximum strength, we choose to show the sum. We have now referred to this as total W_S, and noted that the results are not sensitive to this choice/definition. Captions have been updated to be consistent throughout.



Zonal wind composites for strong waveguide days for two regions (black boxes) based on total W_S (left column) and maximum W_S (right column)

2) Line 227: "in the appendix B"
Corrected, thank you.

3) Line 270: I do not understand the end of the sentence "however, these results .... background flow"

This has now been reworded to clarify: “It is possible that the higher latitude blocking is creating local KS-waveguide conditions - blocks are known to impact the jet, and thus the subsequent movement of smaller, more transient eddies (e.g. Shutts 1983). These results highlight that the KS-waveguide methodology is unable to effectively separate the blocking perturbation from the background flow, despite the blocks typically occurring on length scales smaller than k=2, the upper bound of the spatial Fourier filter.

4) There are two groups of sentences in the discussion (lines 418-422) and in the conclusion (lines 434-438) saying roughly the same thing. I would suggest to suppress one of the two.

These sentences have been substantially changed given the addition of the PV-waveguides using the Fourier-transform background flow, and we have removed the repetition.

Reviewer 2

We thank the reviewer for his continued critique of this paper, and whilst we may not agree on all points, we know that the revised paper is substantially better because of his reviews, and we appreciate the time he has spent on this. Since he chose to not be anonymous, we have added a statement to the acknowledgements to acknowledge his contribution.

Major comments:

A key result of the revised manuscript is the fact that the two different waveguide definitions are associated with rather strong differences in their climatological/ statistical properties. As I said, that’s interesting and worth a publication, given the fact that the Ks-waveguide has been used extensively in the past. However, I disagree with the author’s interpretation saying that “it is unclear which [of the two] is the more ‘accurate’ description [of reality]”; such or related statements appears several times in the text (occasionally somewhat indirectly). First of all, the Ks-definition is based on assumptions that are badly violated in reality (Wirth 2020). It does, therefore, not come as a surprise that the waveguidability diagnosed with the Ks-algorithm is unable to explain properties of true waveguidability (Wirth, 2020). Moreover, Wirth and Polster (2021) have shown that the Ks-method is prone to artefacts in the presence of large-amplitude waves. In fact, the current work corroborates the existence of such artifacts by showing that certain “double jet features” co-occur with blocking-like flow patterns, and the abstract of this manuscript says that the “Ks-methodology…. does not sufficiently separate waves from the background”. Given this body of earlier work and the authors own statement in the abstract, a more logical interpretation of the substantial differences between the two methods would be saying that the current work corroborates earlier indications that the Ks-definition may be inappropriate to diagnose waveguides.

To test the reviewers hypothesis, we have now completed additional analysis looking at the impacts of the background flow methodology, with conclusions that, in most cases, it is the KS- vs PV- definition, and not the method of calculating the background flow, that dominates in the differences between the KS-waveguides and the rolling-zonalization PV-waveguides. We therefore now agree with the reviewer that this works corroborates earlier indications that the KS-definition may be inappropriate to diagnose waveguides. This is reflected in several changes throughout the manuscript, examples given below:
"We recommend PV-waveguides using rolling-zonalized background flow for the study of zonally varying waveguides and their connections to waves. This study adds further caution against using KS-waveguides on time- and/or zonally-varying scales."
"Given the two waveguide definitions have large differences in the relative climatological waveguide frequency between different hemispheres and seasons, it is clear that the two waveguide methods cannot both be accurately quantifying the waveguidability of the atmosphere. Existing concerns over the KS-waveguide methodology suggest that the PV-waveguide climatology is likely a more accurate description of the atmosphere's waveguidability; in the following sub-sections we further explore the differences in the waveguide definitions."
"Given the large differences between the two waveguide datasets, the concerns around separation of waves from the waveguides for the KS-waveguides, and the stronger positive correlations between QSWs and PV-waveguides, we recommend that PV-waveguides on rolling-zonalized flow \citep{polster_new_2023} are used for detecting time-varying zonally asymmetric waveguides, particularly for studying the conditions conducive for quasi-stationary waves and related extreme weather events."

My second issue is related to the first issue. It concerns the authors "explanation" ("this can be understood…..") of the differences between the two waveguide detection algorithms, most notably the fact that the definition of Ks implies a division by the wind speed U, while the PV-gradient method does not. Obviously, the division by U explains some of the differences between the two statistics on an algorithmic level. However, the current text conveys the impression that both methods are equally valid, suggesting that the jury is out which of the two methods is better suited (see above, my first major issue). More explicitly, in her reply the author says: "We do not consider this [i.e., the differences in the seasonal and hemispheric behavior between the two waveguide definitions] necessarily to be a weakness of the Ks-waveguides, however, as the results are consistent with the mathematics that the zonal wind strength appears in the denominator of the Ks equation". I disagree, because this argument lacks logic. Of course, the differences in the statistics can be traced back to differences in the algorithms, but this exercise does not make any statement about the appropriateness of an algorithm to represent reality. For the reasons given above, I think that the Ks-algorithm is not a good waveguide diagnostic. A proof that an algorithm does what is it supposed to do cannot save it from being inappropriate.

I accept that the reviewer has already come to a conclusion that the KS-waveguides are, under all circumstances, not valid; however, I do not agree that the scientific community has agreed that KS-waveguides are not valid under any circumstances. I agree that the reviewer has shown in his earlier work that there are certain situations in which the background flow used for the KS-waveguides includes aspects of the waves, but, to my understanding, it has not been proven that these situations dominate, and therefore the KS-waveguide method should be considered invalid at all time. I therefore think the differences between the KS-waveguides and the PV-waveguides are interesting, and worthy of explanation. The end conclusion of the paper is indeed that the KS-waveguides seem less valid and less useful than the PV-waveguides, and this is made clear in the abstract and conclusions. The new results show that most of the differences between the PV-waveguides and KS-waveguides actually come from the different waveguide definitions, not the different methods of calculating the background flow, and thus I continue to believe that explaining how and why these two methods are different is useful to readers. We have strengthened statements about the inappropriateness of using the KS-waveguides (see response to above comment). But I continue to think that understanding the differences, rather than completely dismissing the KS-waveguides from the beginning, is useful.

Minor comments:

Line 5: what is an “objective algorithm”?
Fair comment – the word objective has been removed.

Lines 6,7: given that the concept of “waveguidability” is more appropriate than a binary decision between “waveguide” or “no waveguide”, such a number (“40% of all days”) is meaningless, because it sensitively depends on the threshold chosen in the definition of a waveguide.
Agreed – this sentence has been removed.

Line 34: does this argument apply only to quasi-stationary waves, or does it also apply to travelling waves?
Unsure – I suspect wave dissipation may play a larger role for quasi-stationary waves remaining at a strong amplitude than for travelling waves passing through at a strong amplitude? Perhaps we should discuss!

Line 44, “there are two main methods….”: this formulation suggests that both methods are well established and equally valid. I would argue that this is not the case: the Ks-method was shown in earlier papers to be fraught by a number of issues (see my first major issue)
Reworded to better highlight the concerns that have been raised about the KS-waveguides.

Line 61, “without this clear separation….”: I would argue that without a scale separation, the entire concept of a waveguide breaks down, which is much worse than what is suggested by the simple phrase “care must be taken”.

The new analysis shows that it is not the method of background flow separation that determines most of the differences between the two methods, but actually the waveguide definitions themselves. From my perspective a lack of ‘clear’ separation does not equal no separation, and thus the phrase ‘much care’ is reasonable.

Line 180, “however”: this sentence lacks logic (see my second major issue). The word “however” suggests that the argument that follows solves the riddle that was given in the previous sentence. The previous sentence refers to the fact that summer has generally weaker jets than winter, and that according to previous work weaker jets are weaker waveguides. The argument given in the following sentence CANNOT resolve this issue, as it only shows why the Ks theory yields a stronger waveguide in summer than in winter; one might just as well conclude that for this reason Ks-theory is inappropriate, because it contradicts the previous wisdom that weaker jets are weaker waveguides.
I argue that the KS-theory states that narrower jets are stronger waveguides, and that this was shown by Manola et al. (2013), and thus the however is reasonable: the basic idea that stronger jets = stronger waveguides is not necessarily true as the width of the jet also needs to be taken into account. However (!), I have reworded this paragraph now to remove this clause. The paragraph is simply trying to explain the results seen, not argue that therefore the KS-waveguide method is OK. I think it is still worthwhile to understand why a method shows the results that it does, particularly when they are different from other methods. I have added a short paragraph at the end of the waveguide frequency sub-section saying:
“Given the two waveguide definitions have large differences in the relative climatological waveguide frequency between different hemispheres and seasons, it it clear that the two waveguide methods cannot both be accurately quantifying the waveguidability of the atmosphere. Existing concerns over the KS-waveguide methodology suggest that the PV-waveguide climatology is potentially a more accurate description of the atmosphere's waveguidability; in the following sub-sections we further explore the differences in the waveguide definitions..”

Line 183, these idealized tests: I do not quite understand the result of these tests, because they contradict the following simple asymptotic argument. If you increase the strength of the jet and leave the width of the jet constant, I expect that Ks saturates to some finite value, but it should not decrease. Basically, in this limit one can neglect the first term on the right-hand side of (2), and Uyy divided by U essentially yields a meridional wavenumber that characterizes the meridional width of the jet.
Agreed, but I don’t think we can make the assumption that we are acting in the limit of being able to neglect the first term in (2), i.e. the planetary vorticity – and in that case, increasing U, whilst keeping the jet half width the same, does indeed decrease K_S. This has been now specified: “Idealised tests of calculating \$K_S\$ on linear multiples of different realistic \$U\$ profiles confirms that, unless the planetary vorticity (first term on the right hand side of Eq. \ref{eq:betaM}) is negligible, then if \$U\$ increases in strength but the latitudinal shape of the \$U\$ profile (i.e. the jet half-width) remains constant, \$K_S\$ decreases in magnitude.”

Line 213: I suggest to change “much more” to “more”
Changed.

Line 214, “U in the denominator…”: again, this only makes plausible that algorithmically the differences are due to the U in the denominator, but it does not make any statement about which of the two options is more realistic (see my second major issue). Similarly on line 345 and 409, “…. This can be understood…”: yes, algorithmically the difference can be “understood”, but the more important question would be which of the two definitions is more realistic. You simply write that “further study is required”; however, I would argue that the Ks-definition has already accumulated a significant amount of criticism, including the fact that the underlying assumptions are badly violated (see my first major issue). For that reason, one could conclude from these differences that the Ks-definition is less realistic. In your text you make the reader believe that the jury is out (see also lines 356, 357, or on line 411 “it is unclear which is the more ‘accurate’ description….”), but in my eyes is really isn’t.
As noted above, I still think it is helpful to understand why they are different, and so most of this text remains. To address the reviewers comment about which is more realistic, we have now added the short paragraph:
“Given the two waveguide definitions have large differences in the relative climatological waveguide frequency between different hemispheres and seasons, it it clear that the two waveguide methods cannot both be accurately quantifying the waveguidability of the atmosphere. Existing concerns over the KS-waveguide methodology suggest that the PV-waveguide climatology is potentially a more accurate description of the atmosphere's waveguidability; in the following sub-sections we further explore the differences in the waveguide definitions..”
Following the analysis on the background flow impacts, showing that much of the differences comes from the waveguide definition, and the stronger QSW correlations for the PV-waveguides, we have also removed the comments about it being unclear which is more realistic, and hopefully the text is now clearer that the PV-waveguides are preferred:
“This study adds further caution against using KS-waveguides on time- and/or zonally-varying scales, and we recommend using rolling-zonalized PV-waveguides for the study of time- and zonally-varying waveguides and their connections to quasi-stationary atmospheric waves.”
“The results in this paper highlight some strong differences between temporally and zonally varying KS-waveguides and PV-waveguides. We can understand many of these differences, such as seasonal and hemispheric variations, at least partially by the different formulations of the waveguide definition, with the strength of the zonal wind U appearing in the denominator of KS-waveguides, but not for PV-waveguides. \cite{wolf_quasistationary_2018} show that QSWs are stronger in NH winter than in NH summer, in line with the seasonal cycle of PV-waveguides. Given the stronger correlations found with QSWs for the PV-waveguides, we conclude that the PV-waveguides likely give a more accurate description of the seasonal cycle of waveguidability.”
“Given the large differences between the two waveguide datasets, the concerns around separation of waves from the waveguides for the KS-waveguides, and the stronger positive

correlations between QSWs and PV-waveguides, we recommend that PV-waveguides on rolling-zonalized flow \citep{polster_new_2023} are used for detecting time-varying zonally asymmetric waveguides, particularly for studying the conditions conducive for quasi-stationary waves and related extreme weather events."

Line 237, "…. Enhanced zonal wind inside the waveguide region": I cannot verify this statement on panels 4c, g, i

Text edited to note that it is not true for all regions for the KS-waveguides, highlighting panels c and g (I would argue it is true to some extent in panel i).

Line 248: could you point to those panels in Fig 4 where this "double jet feature" is clearly visible?!

Added direct references to panels a (North America) and e (Asia), where it is most clear. Removed the reference for Europe, as it is weaker (although clearer when the statistical significance masking is removed).

Line 303: suggest to change "much stronger" ◊ "stronger"

Amended as suggested.

Line 432,433: I disagree. I think here one can dare a statement about causality. One important advantage of the rolling zonalization approach is the fact that the so-obtained background state is (almost) independent of the waves that may be present on the background state (see Wirth and Polster 2021). For this reason, one may at least hypothesize that the strength of the waveguide has an impact on the strength of the waves, but not vice versa!

Agreed, although this study has not proven causality, which is what I meant to imply. Given that the positive correlations also exist (albeit slightly weaker) for the PV-waveguides with the zonal and time filtering, I am hesitant to claim causality any stronger than is done in the previous part of the sentence. However, I agree that it is unlikely in the PV-waveguide definition that causality is in the opposite direction. This has therefore been reworded to: "suggesting that strong PV-waveguide conditions make amplified QSWs more likely, although causality has not been established here"

Reviewer 3

The authors have significantly revised the manuscript by adding a new discussion on the waveguide defined using the gradient of PV, in addition to the original definition based on the gradient of absolute vorticity (i.e., refractive index). This addition offers a more comprehensive exploration of the waveguide and provides new insights into understanding the behavior of waveguides variations from different perspectives. The authors have fully addressed my previous concerns, and thus I recommend accepting the current version of the manuscript

We thank the reviewer for their time and comments that helped improve the manuscript.