nature portfolio

Peer Review File

Decadal changes in atmospheric circulation detected in cloud motion vectors

Corresponding Author: Professor Larry Di Girolamo

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
"Decadal changes in atmospheric circulation detected in cloud motion vectors" is timely. Actually, it's overdue, and I think it's a great fit for Nature.

Some of the most readily validated modeled or reanalyzed metrics of circulation strength and or latitude are the surface wind maxes and zero crossings. But even these are challenging to validate, because the wind speeds are overwhelmingly measured over land. Wind speeds aloft would be very nice to be able to analyze, but they are, of course, very sparse. The use of MISR cloud motion vectors increases the amount of wind measurements orders of magnitude, with the caveat that these wind measurements are taken exclusively at cloud tops.

This is a huge caveat, and I have mixed feelings about how the authors handle it. They do mention it explicitly a couple of times (and I point out one in the minor comments below), but they also make a couple of claims that I don't think are justified (because of the caveat) without additional work. Specifically, the article has language in it that makes it sound like MISR data can be used to constrain zonal mean modeled circulation shifts, and that the small-to-absent circulation shifts over the SH in MISR data alone cast doubt on larger modeled shifts. I agree that MISR data seem likely to be a valuable source of data for constraining modeled circulation change, I don't think there's enough in the paper to support the claim that they can be used to constrain zonal mean circulation change without substantial additional work. I think they can be used to validate conditional change – like the kind of change shown in the second column of Figure 1 and the first four extended data figures. However, even subsampling the reanalysis data to the wind data point locations sampled by MISR introduces a bit of an error source. If it is wind data that you are trying to validate or refute in ERA5, and if the occurrence of clouds depends on that wind data, the location of the clouds cannot be assumed to match up between MISR and ERA5. So while the methodology applied to get the second column is good, it is not complete, and it won't work to create an apples-to-apples comparison between MISR and model data. A complimentary approach – and one that I think will work with ERA5 or CMIP data – is to simulate a MISR orbit and sample winds at locations in the reanalysis or model that have simulated clouds. That said, I think that sort of analysis is more than can reasonably be expected of the current paper. So I think that the writing needs to be revised to walk back the claims just a tad and make sure that the claims made are those that can be backed up by the current work, and others are stated as speculation.

In lieu of that, I think at least a strong empirical relationship between the MISR-sampled data and a modeled quantity – along with a mechanistic explanation – would need to be demonstrated. Poleward, cloudy branches of the midlatitude circulation and surface winds are indeed a great place to look, and the discussion of the results and their consistency with other work strongest when that is the emphasis.

I also think more discussion is also needed about why the authors are so confident that the differences between the MISR and MISR-sampled ERA5 points are due to biases in ERA5 and not to shortcomings in the retrieval. Granted, the retrieval looks to be based on simple optics and geometry, so I can imagine them being traceable to standards. But I'm actually not familiar enough with the MISR cloud motion vectors to say for sure. And with the point I brought up earlier about clouds potentially being in the wrong locations in ERA5, I think there is room for ERA5 biases affecting the measured ERA5-MS product in counter-intuitive ways. Just treating the ERA5-MS wind values as measures of the same circulation features as those sampled by MISR is risky. In any case, I don't think that the work in the paper alone is enough to back up the claim that the biases seen are biases in ERA5's cloudy circulation.

Another area I think needs to be discussed a bit more clearly is the potential impact of changes in cloudiness to calculated trends in wind. In humidity increased (decreased), then cloudiness would likely increase (decrease) even in the absence of changes in the wind field. At a given location, then, increasing moisture could conceivably be misconstrued as a decrease in wind speed. So I think this should be addressed in the text.

Finally, the paper mentions that the time period analyzed by MISR is short, but the article discusses MISR as something that can be used to constrain modeled trends, some of which are too large. But the largest trends are some of the oldest, and included very little of the time period analyzed by MISR. So (given additional work like I mentioned above) MISR will probably be useful for constraining cloudy circulations over the most recent couple of decades, it still may not be enough to directly constrain longer-term trends.

This is all to say that I think the authors need to be a bit more circumspect about their claims and the stated (or implied) scope of the current work. This work is a good step towards constraining models with MISR cloud motion vectors, but it is not enough to demonstrate it yet.

Minor comments and edits

Line 62: how about “consistent” instead of “relevant”?

Line 82: a word on their vertical resolution would be interesting.

Figure 1 (and others): ERA5_MS and ERA5_AW are used in the main text, but ERA5-MS and ERA5-AW are used in the figures. I recommend consistency. (I searched “ERA5-AW” to go up to a definition I had initially read over, and couldn’t find it.)

Line 132: The “Hadley cell in the SH” is not the same as the “SH Hadley cell.” The wintertime Hadley circulation extends into the summer hemisphere a bit. So this change could be caused by a change in either the SH cell strengthening or the NH cell weakening, or a mix of both. So “and” isn’t quite the right word, here. That said, the meridional wind trends you see in the deep tropics seem pretty consistent between seasons. So I think what you say is true, but it requires the seasonal breakdown in the extended data figures to verify. I would modify your wording or cite the extended figures here.

Line 154: Really what they indicate is that there is a difference. You attribute this difference to a bias from ERA5, rather than in the wind retrieval algorithm. And while you have good reason to be right, these findings alone do not indicate that ERA5 biases are to blame. You’re drawing on other knowledge (which you briefly mentioned in the intro), and I think some discussion of the possible errors of ERA5 and the retrieval would be valuable. Or maybe some discussion of how this apparent ERA5 wind bias compares with other work and what we know about the behavior of the large scale dynamical and cloud fields in ERA5 and models. (Is this wind bias something we would suspect from models? Or is there no way for us to know this without something like MISR, since ERA5 runs on a model backbone?)

Line 157: I think the bigger impact here is that you’re sampling the poleward-V phase of the waves. Even if you weren’t sampling the wavier component of the flow, you’d still get the largest difference – the difference in V – from this fact alone. You may also get the difference in U from this fact alone, too, since the warm conveyor belt curves cyclonically towards the west as it extends poleward and upward. (At least that’s how things seem to work if I picture warm conveyor belts and cyclones.) It makes me think that much of the angular momentum transport from the tropics to the extra tropics is happening above the clouds.

Line 172: You mentioned the consistency (and the underestimation by ERA5) in the midlatitudes, but you didn’t say anything about the deep tropics, where ERA5-MS depicts a westerly trend no meridional wind trend.

Also, I’m frustrated by a lack of discussion of the third panel. Some discussion of the third panel early on could be a good way to give context for the changes seen in columns 1 or 2 of Figure 1. Granted, it’s been seen before, but the difference between the MISR-sampled winds (or ERA5-MS) and ERA5-AW is interesting! ERA5 shows little V trend in the deep tropics, and it only matters a bit whether it’s -AW or -MS.

There are some tropical biases (double ITCZ, etc.) that will likely make it so that clouds are in the wrong place. Also, I think that (in the deep tropics), you’re more likely to have high clouds at longitudes where you have a Hadley cell-like overturning, rather than a reversed, monsoon-like overturning. In other words, I think sampling clouds favors sampling a Hadley regime. Based on that, I’m not surprised to see only weak hints of the deep tropical change in panel g showing up in panel i. And I would be willing to entertain the idea that differences between panels d and e (or g and h) are in part because of modeled ITCZ biases.

Line 232: This is one of those times you bring up a caveat well. Thanks.

Line 243: While this is true in general, I think it’s problematic here. This is one of the statements that bugs me in this paper. It sounds like you’re saying that models that show significant SH widening should be treated with greater skepticism because the MISR sampled rates were low and not significant. In the previous paragraph, you mentioned that the widening rates in the MISR data are not directly comparable to widening rates overall—that means in models, too, but it feels like that statement was forgotten by the time you said this.

Line 255: Same comment as before; these results alone don't suggest a problem with ERA5; you're suggesting that based on outside info.

Line 269: I think you should delete the comma between longest and climate-quality.

Line 312: Nice to see you chose nearest neighbor here. Good call, I think.

Referee #2

(Remarks to the Author)
This is an excellent manuscript topic. After Aeolus, not many height-resolved satellite winds appear for accurate model validation. MISR indeed has also value in trend observation, which Aeolus has not. Over the oceans and at the surface wind scatterometers can do a good job too in terms of dynamical trends.

Having said that, several aspects of CMVs may be trending and which in my view need more attention before publication. Rain increases substantially and it falls from the clouds observed by MISR. Does more rain imply more energy and momentum flux in clouds, deeper clouds, faster changing clouds? If the sampling target changes, does it change the winds it measures, even if the mean circulation does not change? I'd guess these topics would interest many readers of the manuscript.

Therefore, the implied topic of cloud changes and how these could affect CMVs needs to be clarified and understood before presenting results on wind trends. This discussion is now very weak in the manuscript. Are the trends presented cloud trends or circulation trends?

A second related point is that it is not clear how many "all Wind" (AW) points there are as compared to MISR sampled (MS) cloud points? Following Aeolus experience, there would be about 20% MS and 80% AW points, which explains some of the large differences shown I guess.

Finally, satellite sampling affects the results for MS very much indeed. Obtaining the same sampling data set from ERA5 however does not guarantee fully physical representative winds to MISR as ERA5 clouds will appear in different places generally than real clouds. This also explains some of the differences presumably.

I provide a commented/edited version of the Word file for further details and noted typos.

Cheers,

Ad Stoffelen

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
In my previous review for this manuscript, I suggested the authors either (1) perform an additional analysis (basically perform a OSSE with MISR) to eliminate the impact of differences in cloud location in MISR and ERA5, or (2) soften their conclusions. I considered my proposal for the first option to "...be more than can reasonably be expected of the current paper."

In the authors' rebuttal, they argue well for the accuracy of MISR, and describe how they took an alternative approach to eliminating the conditionality problem: they just performed an additional comparison, using only the 71.2% of the locations and times in which there were MISR retrievals and clouds in ERA5. My suggestion was overkill; but the authors took a better approach that resolved the same concern.

At the end of their rebuttals, the authors said, regarding our concerns, "We hope that changes we've made to the manuscript have addressed them." And my own response is that they have. Handily. My major and minor critiques have all been addressed. I agree with the changes – particularly in the methods and extended data, but also in the main text. I happily recommend publication, and look forward to the discussion that will follow.

Referee #2

(Remarks to the Author)
To be brief, may I congratulate the authors with their improved, original and significant manuscript. Most of the reviewer's suggestions are taken up well.

Just to highlight a few remaining points:
- The additional conditioning that ERA5 also be a cloud-top at the times and locations of the MISR CMVs had negligible

impact on your analysis and it didn't matter what the cloudy or clear state was for the ERA5_MS samples, their comparison to MISR CMVs led to the same result. Given results on tropical moist convection with scatterometers this may not be surprising. King et al. (2022, https://doi.org/10.3390/rs14051147) show that extreme convergence and divergence are associated with heavy rain, representing resp. updrafts and downdrafts. Collocated ECMWF ERA5 winds do not show these extreme divergences and convergences and hence lack some of the dynamics seen with real rain clouds. One could say, it does not matter much whether moist convection exists or not in ERA5, while in the real world it does matter. Hence, could it be that MISR results are still affected by climate changes in moist convection, while it is hard to prove as models do not show real moist convection dynamics?
- Grand biases in tropical (tropopauze) closure in the ECMWF model are shown against radiosondes at the locations of the radiosondes in Houchi et al. (2010, https://doi.org/10.1029/2009JD013196), in particular for vertical wind shear. Belmonte Rivas and Stoffelen (2019, https://doi.org/10.5194/os-15-831-2019) show rather large tropical errors in ocean surface winds, while Zagar et al. (2025, https://essopenarchive.org/doi/pdf/10.22541/essoar.173776264.46417622/v1) highlight tropopause shear errors. These publications appear to link well with the MISR work presented and warrant discussion.
- Unfortunately, as may be understood from the previous points, radiosondes are insufficient to control the model dynamical biases in the tropics or elsewhere. Rather, moist convection processes and associated 4D dynamics are too fast (30 min. time scale) for a global NWP model to track with the observing system at hand. The poor representation of the associated dynamics is the main reason for dynamics that are independent of the presence of cloud, not data assimilation of radiosondes (cf. Houchi et al.).

It would be sufficient to consider these in a minor revision.

Referees' comments:

We are extremely grateful for the detailed reviews provided by Reviewers 1 and 2. We believe that many other readers would have brought up the same questions and concerns. We've undertaken additional analyses (see below) to address the reviewers' comments and have made revisions needed to ensure their concerns are clearly addressed in this version of the manuscript.

Fortunately, these changes have not led to changes in the figures or any of our conclusions in Main; in fact, they are now even better supported.

Clarifying statements and edits were made to Main. However, we added substantial analysis and discussion in the Extended Data Analysis and Discussion section to address the reviewers' main concerns on (1) confounding effects in the interpretation of our reported trends (Lines 429-483), (2) ERA5 not placing clouds at the right height as a factor in its comparison to MISR CMVs (Lines 486-523), and (3) a comparison to other works to support our statements on ERA5 wind biases in the upper troposphere (Lines 526-584). We have included three additional Extended Data Figures (Extended Data Figures 1 to 3) and 14 additional citations to address these concerns. We have also added additional analysis in Methods to demonstrate the long-term stability of MISR stereo measurements for climate research (Lines 335-352).

All Line Numbers that the reviewers point to are from the original manuscript. All Line Numbers in our response are for the revised manuscript. The original comments from the reviewers are in black. Our response are in blue.

Referee #1 (Remarks to the Author):

"Decadal changes in atmospheric circulation detected in cloud motion vectors" is timely. Actually, it's overdue, and I think it's a great fit for Nature.

Thanks!

Some of the most readily validated modeled or reanalyzed metrics of circulation strength and or latitude are the surface wind maxes and zero crossings. But even these are challenging to validate, because the wind speeds are overwhelmingly measured over land. Wind speeds aloft would be very nice to be able to analyze, but they are, of course, very sparse. The use of MISR cloud motion vectors increases the amount of wind measurements orders of magnitude, with the caveat that these wind measurements are taken exclusively at cloud tops.

This is a huge caveat, and I have mixed feelings about how the authors handle it. They do mention it explicitly a couple of times (and I point out one in the minor comments below), but

they also make a couple of claims that I don't think are justified (because of the caveat) without additional work.

We have added additional analyses and discussions in Extended Data Analysis and Discussion. The results from our analyses now clearly justifies our approach, and supports our analyses and conclusions presented in Main. More on this below.

Specifically, the article has language in it that makes it sound like MISR data can be used to constrain zonal mean modeled circulation shifts, and that the small-to-absent circulation shifts over the SH in MISR data alone cast doubt on larger modeled shifts.

Perhaps "constrain" has more meaning than intended, so we replaced "constrain" with "benchmark" (Lines 34 and 306). We demonstrated that MISR CMVs can be used as a benchmark for zonal mean cloud-top-conditioned circulation shifts for the 2000-2020 period. We used the ERA5 reanalysis data to demonstrate this rather than climate models. Benchmarking climate models has the added complication of needing to contend with internal climate variability – doable[8,9,35], but not done here.

In terms of "small-to-absent circulation shifts over the SH from MISR data alone", we take it that this is with reference to Figure 2b, where SH edge of the Hadley cell does not exhibit a statistically significant ($p < 0.05$) shift during the 2000-2020 period. But this is also true for ERA5_MS and ERA5_AW. So MISR is not alone.

In terms of casting "doubt on larger modeled shifts", we believe we traced this to a poorly worded sentence in Discussion and Conclusion of Main. It has been removed. Thanks for catching that. It was not our intent to cast doubt on larger climate model and reanalysis shifts found in other studies for other periods. Our study is specific to the 2000-2020 period.

These points are now addressed in Discussion and Conclusion, Lines 285-292.

I agree that MISR data seem likely to be a valuable source of data for constraining modeled circulation change, I don't think there's enough in the paper to support the claim that they can be used to constrain zonal mean circulation change without substantial additional work. I think they can be used to validate conditional change – like the kind of change shown in the second column of Figure 1 and the first four extended data figures. However, even subsampling the reanalysis data to the wind data point locations sampled by MISR introduces a bit of an error source. If it is wind data that you are trying to validate or refute in ERA5, and if the occurrence of clouds depends on that wind data, the location of the clouds cannot be assumed to match up between MISR and ERA5. So while the methodology applied to get the second column is good, it is not complete, and it won't work to create an apples-to-apples comparison between MISR and model data. A complimentary approach – and one that I think will work with ERA5 or CMIP data – is to simulate a MISR orbit and sample winds at locations in the reanalysis or model that have simulated clouds. That said, I think that sort of analysis is more than can

reasonably be expected of the current paper. So I think that the writing needs to be revised to walk back the claims just a tad and make sure that the claims made are those that can be backed up by the current work, and others are stated as speculation.

In lieu of that, I think at least a strong empirical relationship between the MISR-sampled data and a modeled quantity – along with a mechanistic explanation – would need to be demonstrated. Poleward, cloudy branches of the midlatitude circulation and surface winds are indeed a great place to look, and the discussion of the results and their consistency with other work strongest when that is the emphasis.

We have undertaken the analysis suggested by the reviewer. The results are shown and discussed in Extended Data Analyses and Discussion (EDAD) under the subheading "An ERA5 Cloud Conditional Analysis", Lines 486-523. It is a wonderful addition to the manuscript, so we thank the reviewer for suggesting it. As we show, the additional conditioning that ERA5 also be a cloud-top at the times and locations of the MISR CMVs had negligible impact on our analysis in Main. In fact, it didn't matter what the cloudy or clear state was for the ERA5_MS samples, their comparison to MISR CMVs led to the same result. We provide a discussion as to why we think we get the same results – it may be due to ERA winds being strongly constrained by the assimilation of vast amounts of observational data (e.g., rawinsondes) that are independent of the presence of cloud. We also discuss that our original approach was indeed an apples-to-apples comparison for identifying wind errors in ERA5, and that this additional conditioning that ERA5_MS also be a cloud-top is only a diagnostic, but a valuable one. See discussion for details.

Edits to Main includes (1) a pointer to the EDAD discussion in the first paragraph of "Consistency with Atmospheric Reanalyses", Lines 156-159, and (2) a highlight of the key finding from EDAD in "Discussion and Conclusion", Lines 278-283.

I also think more discussion is also needed about why the authors are so confident that the differences between the MISR and MISR-sampled ERA5 points are due to biases in ERA5 and not to shortcomings in the retrieval. Granted, the retrieval looks to be based on simple optics and geometry, so I can imagine them being traceable to standards. But I'm actually not familiar enough with the MISR cloud motion vectors to say for sure. And with the point I brought up earlier about clouds potentially being in the wrong locations in ERA5, I think there is room for ERA5 biases affecting the measured ERA5-MS product in counter-intuitive ways. Just treating the ERA5-MS wind values as measures of the same circulation features as those sampled by MISR is risky. In any case, I don't think that the work in the paper alone is enough to back up the claim that the biases seen are biases in ERA5's cloudy circulation.

More discussion was added as follows:

1) The differences that we discuss in the manuscript are only the differences that are much larger (i.e., 95% CI or larger) than the uncertainty in the MISR data. We now make this clear

throughout the manuscript. (Lines 172-173, 192-193). We also added an additional figure, now Extended Data Figure 1 that better clarifies the statistical differences in trends between MISR and ERA5_MS.

2) We have added additional details about the MISR stereo algorithm to further emphasize why it is a climate quality record and why the behavior in the differences between MISR and ERA5_MS cannot be easily explained as a behavior in error in MISR. We also add additional analyses in Methods that shows the long-term stability of the MISR CMV record for climate research (Lines 320 to 352). Here's an additional discussion below for the reviewer to consider:

The MISR stereo product has been extensively validated, with validation results summarized in Methods with full citation. As described in Main and in Methods, the errors are also fully traceable and self-contained (i.e., there's no ancillary data, such as temperature profiles, needed). The algorithm is also agnostic to what MISR is looking at, be it high clouds, low cloud or land surfaces. It's also agnostic to where and when (e.g., high latitude summer vs low latitude winter) the data is collected. The MISR instrument and algorithm are also incredibly stable, which we now show with results of additional analyses added to Methods. There is no evidence that exists that the MISR stereo errors summarized in Methods are in error.

Now consider reanalysis data, such as the ERA5. The second paragraph in Main describes several sources of uncertainties for ERA5. These uncertainties are not completely characterized, leading to calls (e.g., Staten et al. 2018) to validate reanalysis data with high-quality, independent data. Here, we're using MISR to answer this call.

3) We've added a section to Extended Data Analyses and Discussion (EDAD) under the subheading "Comparison to other works", Lines 526-584. This is in response to another comment by Reviewer 1 below with reference to "Line 154", but fits in well here too. In this section, we examined the quality of ERA5 winds against other independent wind measurements (e.g., scatterometers, doppler wind lidars). We also carried out additional analyses with MISR for comparison with these other works. We conclude that the differences of those observations with ERA5 are very similar to the differences reported here between ERA5 and MISR in the very limited regions of the troposphere examined by these other works (basically surface winds and one detailed study of vertical profile over Singapore). For most of the rest of the troposphere, validation of ERA5 winds against independent data doesn't exist – except now from MISR. The fact that both MISR and these other independent data show similar difference with ERA5 supports that the notion that the differences are due to ERA5 wind errors.

In addition, as noted in the new EDAD section, the positive impact that the assimilation of Aeolus winds had on NWP model forecasts, including ECMWF (e.g., Rennie et al. 2021; Zuo et al. 2024), is further evidence that modeled winds still have room for improvements, particularly in the upper troposphere (i.e., where MISR CMV show the largest disagreement with ERA5_MS in Figure 1 and Extended Data Figure 1). We also point to studies that have shown issues in reanalysis data properly capturing upper tropospheric dynamics (e.g., Flannaghan and

Fueglistaler 2014). Again, these studies support the notion of ERA5 wind errors in the upper troposphere.

So, based on the above discussions given in (2) and (3), there's good reason to expect that the differences between MISR CMV and ERA5_MS that are well beyond MISR's uncertainties are due to errors in ERA5_MS. Our examination following the reviewer's suggestions has not identified evidence or reason to say otherwise. Still, we have added words such as "may" and "likely" to play it safe.

4) The reviewer may be suggesting with this comment that we may be a bit too negative on ERA5. That was not our intention. We think the agreement between ERA5_MS and MISR are very good. We have gone through the manuscript to soften the text here and there, while highlighting the very good agreement between the two datasets (e.g., Line 163-168; Line 228; Line 274; Lines 285-290).

5) The reviewer's point on clouds being in the wrong place in ERA5 and potentially impacting our analysis is addressed in Extended Data Analyses and Discussion (EDAD) under the subheading "An ERA5 Cloud Conditional Analysis", as described in response to an earlier comment above.

Another area I think needs to be discussed a bit more clearly is the potential impact of changes in cloudiness to calculated trends in wind. In humidity increased (decreased), then cloudiness would likely increase (decrease) even in the absence of changes in the wind field. At a given location, then, increasing moisture could conceivably be misconstrued as a decrease in wind speed. So I think this should be addressed in the text.

We have added an extensive quantitative analyses and discussion on confounding factors to the Extended Data Analysis and Discussion with the subheading of "Examination of Confounding Factors", Lines 429-483, along with the following statement at the beginning of the second paragraph in Main under the subheading "Decadal Trends in CMVs", Lines 109-114:

"Interpreting trends in MISR CMVs as changes in cloud-top-conditioned circulations requires careful consideration of several potential confounding factors, as described and analyzed in Extended Data Analyses and Discussion (EDAD). Our EDAD shows that the impacts of these confounding factors on the observed trends are small, supporting the conclusion that much of the observed MISR CMV changes (Figure 1) are due to changes in the cloud-top-conditioned atmospheric circulation."

Please see EDAD for the full analyses and discussion.

Finally, the paper mentions that the time period analyzed by MISR is short, but the article discusses MISR as something that can be used to constrain modeled trends, some of which are

too large. But the largest trends are some of the oldest, and included very little of the time period analyzed by MISR. So (given additional work like I mentioned above) MISR will probably be useful for constraining cloudy circulations over the most recent couple of decades, it still may not be enough to directly constrain longer-term trends.

We agree that the record length of MISR is still relatively short and have emphasized limitations related to the record length in the revised manuscript. This is a similar comment as the second major comment above, which we addressed. Please see that response, which includes clarifying edits to the manuscript.

This is all to say that I think the authors need to be a bit more circumspect about their claims and the stated (or implied) scope of the current work. This work is a good step towards constraining models with MISR cloud motion vectors, but it is not enough to demonstrate it yet.

We hope that extensive analyses and discussions that we added to EDAD have addressed the reviewer's concerns. The reviewer's comments gave us the opportunity to provide more evidence and discussion to further support and clarify the findings in our original manuscript. For this, we are grateful to the reviewer.

Minor comments and edits

Line 62: how about "consistent" instead of "relevant"?

Changed.

Line 82: a word on their vertical resolution would be interesting.

We add "…, binned to 1 km vertical bins, …" a few lines earlier.

Figure 1 (and others): ERA5_MS and ERA5_AW are used in the main text, but ERA5-MS and ERA5-AW are used in the figures. I recommend consistency. (I searched "ERA5-AW" to go up to a definition I had initially read over, and couldn't find it.)

Fixed. We now use ERA5_MS and EAR5_AW throughout.

Line 132: The "Hadley cell in the SH" is not the same as the "SH Hadley cell." The wintertime Hadley circulation extends into the summer hemisphere a bit. So this change could be caused by a change in either the SH cell strengthening or the NH cell weakening, or a mix of both. So "and" isn't quite the right word, here. That said, the meridional wind trends you see in the deep tropics seem pretty consistent between seasons. So I think what you say is true, but it requires

the seasonal breakdown in the extended data figures to verify. I would modify your wording or cite the extended figures here.

We agree that the observed changes can arise from seasonally varying changes. We modified the wording and removed the reference to the Hadley cells. We now describe the circulation changes as "an intensification of the cross-equator circulation that contributes to the low-troposphere convergence and the upper-troposphere divergence near 6°N" (See Lines 144-146)

Line 154: Really what they indicate is that there is a difference. You attribute this difference to a bias from ERA5, rather than in the wind retrieval algorithm. And while you have good reason to be right, these findings alone do not indicate that ERA5 biases are to blame. You're drawing on other knowledge (which you briefly mentioned in the intro), and I think some discussion of the possible errors of ERA5 and the retrieval would be valuable. Or maybe some discussion of how this apparent ERA5 wind bias compares with other work and what we know about the behavior of the large scale dynamical and cloud fields in ERA5 and models. (Is this wind bias something we would suspect from models? Or is there no way for us to know this without something like MISR, since ERA5 runs on a model backbone?)

We addressed this comment above in response to a related comment made by the reviewer. Please see the above response for details, which includes added analyses and discussions to Methods and Extended Data Analysis and Discussion (EDAD). Line 154 has been modified to read as:

"Since these differences are much larger than the uncertainty in MISR CMVs (Methods and EDAD), these findings indicate that further improvements to ERA5 are needed to reduce these upper tropospheric biases – a finding supported by other evidence (see EDAD)." (see Lines 172-175)

Line 157: I think the bigger impact here is that you're sampling the poleward-V phase of the waves. Even if you weren't sampling the wavier component of the flow, you'd still get the largest difference – the difference in V – from this fact alone. You may also get the difference in U from this fact alone, too, since the warm conveyor belt curves cyclonically towards the west as it extends poleward and upward. (At least that's how things seem to work if I picture warm conveyor belts and cyclones.) It makes me think that much of the angular momentum transport from the tropics to the extra tropics is happening above the clouds.

We agree that the differences between ERA5_MS and ERA5_AW are likely related to the conditional sampling. We have modified the wording to make this reasoning clearer.

Line 172: You mentioned the consistency (and the underestimation by ERA5) in the midlatitudes, but you didn't say anything about the deep tropics, where ERA5-MS depicts a westerly trend no meridional wind trend.

Thanks for taking note of this. We've rewritten the bottom half of this paragraph, and added what is now called Extended Figure 1, to focus on and highlight the statistical differences in the trends between ERA5_MS and MISR CMV, which includes the tropical upper troposphere. See Lines 185-195.

Also, I'm frustrated by a lack of discussion of the third panel. Some discussion of the third panel early on could be a good way to give context for the changes seen in columns 1 or 2 of Figure 1. Granted, it's been seen before, but the difference between the MISR-sampled winds (or ERA5-MS) and ERA5-AW is interesting! ERA5 shows little V trend in the deep tropics, and it only matters a bit whether it's -AW or -MS.

We too think the differences are interesting! We do come back to these differences in the Discussion and Conclusion section as they are important. It was discussed in an earlier draft around Line 172, but we removed it since it was a duplicate of what comes later, and we were trying to trim the text to make it fit into Nature's word limit. We've modified the text around Line 172 (now it's own paragraph on AW) to bring some of it back. See Lines 197-203.

There are some tropical biases (double ITCZ, etc.) that will likely make it so that clouds are in the wrong place. Also, I think that (in the deep tropics), you're more likely to have high clouds at longitudes where you have a Hadley cell-like overturning, rather than a reversed, monsoon-like overturning. In other words, I think sampling clouds favors sampling a Hadley regime. Based on that, I'm not surprised to see only weak hints of the deep tropical change in panel g showing up in panel i. And I would be willing to entertain the idea that differences between panels d and e (or g and h) are in part because of modeled ITCZ biases.

The issue of ERA5 placing clouds in the wrong place was addressed in response to an earlier comment, which led to a new section in Extended Data Analysis and Discussion (EDAD) subtitled "An ERA5 Cloud Conditional Analysis" and the conclusion the ERA5 placing clouds in the wrong spot was not the issue. But we agree with the reviewer that differences in the tropical upper troposphere between d and e (or g and h) are likely model biases. Indeed, the biases that show up in the tropical upper troposphere are also observed in other studies and linked to known model weaknesses in the dynamics in this region of the atmosphere. This is reviewed in a new EDAD section subtitled "Comparison to other works". (Lines 526-584)

We also agree with the reviewer that differences in trends between panel g and panel i are partly due to the conditional sampling of MISR CMV to cloud-top, unlike ERA5 (which we note throughout), and that the trends in meridional winds may be stronger in regions with a deep cross-equatorial overturning circulation, such as the western Pacific, than in regions with shallow overturning circulation, such as the eastern Pacific. But if so, then this difference would show up between panel h and i as well, given our EDAD discussion. We speculate instead on the existence of trends in the input data assimilated in into ERA5 that would impact both

ERA5_MS and ERA5_AW in the tropical upper troposphere, which we have added to the text (Lines 192-195; Lines 202-203)..

Line 232: This is one of those times you bring up a caveat well. Thanks.

Thanks.

Line 243: While this is true in general, I think it's problematic here. This is one of the statements that bugs me in this paper. It sounds like you're saying that models that show significant SH widening should be treated with greater skepticism because the MISR sampled rates were low and not significant. In the previous paragraph, you mentioned that the widening rates in the MISR data are not directly comparable to widening rates overall—that means in models, too, but it feels like that statement was forgotten by the time you said this.

This comment is part of an earlier comment that we responded to above. The sentence was poorly worded and removed to keep our study focused on the 2000-2020 period. Thanks for catching this.

Line 255: Same comment as before; these results alone don't suggest a problem with ERA5; you're suggesting that based on outside info.

We responded to this comment above.

Line 269: I think you should delete the comma between longest and climate-quality.

Done.

Line 312: Nice to see you chose nearest neighbor here. Good call, I think.

Thanks!

We are extremely grateful for the detailed review provided by Reviewer 1. We hope that changes we've made to the manuscript have addressed them.

If the reviewer has additional questions or concerns, they may contact us directly so that we can discuss further. Thanks!

Referee #2 (Remarks to the Author):

This is an excellent manuscript topic. After Aeolus, not many height-resolved satellite winds appear for accurate model validation. MISR indeed has also value in trend observation, which Aeolus has not. Over the oceans and at the surface wind scatterometers can do a good job too in terms of dynamical trends.

Thanks. Aeolus was an incredible mission. I hope a longer mission comes out of it. In our response to Reviewer 1, we turned to several Aeolus and scatterometer studies for supporting evidence on the quality of MISR CMVs and it's differences with ERA5. This is now captured in a new section in Extended Data Analysis and Discussion, subtitled "Comparison to other works," Lines 526-584.

Having said that, several aspects of CMVs may be trending and which in my view need more attention before publication. Rain increases substantially and it falls from the clouds observed by MISR. Does more rain imply more energy and momentum flux in clouds, deeper clouds, faster changing clouds? If the sampling target changes, does it change the winds it measures, even if the mean circulation does not change? I'd guess these topics would interest many readers of the manuscript.

Therefore, the implied topic of cloud changes and how these could affect CMVs needs to be clarified and understood before presenting results on wind trends. This discussion is now very weak in the manuscript. Are the trends presented cloud trends or circulation trends?

Reviewer 1 raised the same concern. We duplicate that response here:

We have added an extensive quantitative analyses and discussion on confounding factors to the Extended Data Analysis and Discussion with the subheading of "Examination of Confounding Factors", Lines 429-483, along with the following statement at the beginning of the second paragraph in Main under the subheading "Decadal Trends in CMVs", Lines 109-114:

"Interpreting trends in MISR CMVs as changes in cloud-top-conditioned circulations requires careful consideration of several potential confounding factors, as described and analyzed in Extended Data Analyses and Discussion (EDAD). Our EDAD shows that the impacts of these confounding factors on the observed trends are small, supporting the conclusion that much of the observed MISR CMV changes (Figure 1) are due to changes in the cloud-top-conditioned atmospheric circulation."

Please see EDAD for the full analyses and discussion.

As to physical reasoning driving the observed CMV trends in the upper troposphere, we brought up several arguments (Lines 255-258), as summarized in the first paragraph of

Discussion and Conclusion. But these are hypotheses at this point. The rain hypothesis in the reviewer's comment fits nicely into the group that we mentioned. Of course, the hypotheses will need to be tested one day with data.

A second related point is that it is not clear how many "all Wind" (AW) points there are as compared to MISR sampled (MS) cloud points? Following Aeolus experience, there would be about 20% MS and 80% AW points, which explains some of the large differences shown I guess.

ERA5_MS and ERA5_AW shouldn't add up to 100%. This is because AW are always present and MS is a sub-sample of them. We can estimate the fraction of ERA_MS to ERA_AW as follow: we get one MS at a cloud top, which can only occupy one 1-km altitude bin at any instance. In contrast, AW occupies all 1-km altitude bins at this same instance. There are about 11 1-km altitude bins in the troposphere when averaged over latitude. About 85% of 17.6km x 17.6 km areas (the resolution of a CMV retrieval) contain some cloud that can trigger a retrieval. Therefore, (1/11)x0.85 = 0.077; that is, the fraction of AW samples (at MISR sampling times and areas) over all tropospheric altitudes and latitudes that contain a MS sample is about 8%. The distribution of this 8% and how they are trending is shown in Extended Data Figure 2. Since the number of AW samples is a constant for all altitudes, latitudes and time, the trends shown in Extended Data Figure 2 can only be attributed to trends in MS samples. This is now included in the EDAD discussion with the subheading of "Examination of Confounding Factors" (Lines 429-483). As noted above, the confounding factors are small, leading to the conclusion that much of the observed MISR CMV changes shown in Figure 1 are due to changes in the cloud-top-conditioned atmospheric circulation.

Finally, satellite sampling affects the results for MS very much indeed. Obtaining the same sampling data set from ERA5 however does not guarantee fully physical representative winds to MISR as ERA5 clouds will appear in different places generally than real clouds. This also explains some of the differences presumably.

Reviewer 1 also made the same comment. We repeat our response here:

We have undertaken the analysis suggested by the reviewer. The results are shown and discussed in Extended Data Analyses and Discussion (EDAD) under the subheading "An ERA5 Cloud Conditional Analysis", Lines 486-523. It is a wonderful addition to the manuscript, so we thank the reviewer for suggesting it. As we show, the additional conditioning that ERA5 also be a cloud-top at the times and locations of the MISR CMVs had negligible impact on our analysis in Main. In fact, it didn't matter what the cloudy or clear state was for the ERA5_MS samples, their comparison to MISR CMVs led to the same result. We provide a discussion as to why we think we get the same results – it may be due to ERA winds being strongly constrained by the assimilation of vast amounts of observational data (e.g., rawinsondes) that are independent of the presence of cloud. We also discuss that our original approach was indeed an apples-to-

apples comparison for identifying wind errors in ERA5, and that this additional conditioning that ERA5_MS also be a cloud-top is only a diagnostic, but a valuable one. See discussion for details.

Edits to Main includes (1) a pointer to the EDAD discussion in the first paragraph of "Consistency with Atmospheric Reanalyses", Lines 156-159, and (2) a highlight of the key finding from EDAD in "Discussion and Conclusion", Lines 278-283.

I provide a commented/edited version of the Word file for further details and noted typos.

Thanks! We've taken into consideration all edits and minor clarifying-comments that the reviewer suggested in the edited version of the manuscript that they provided. The reviewer grabbed the major comments in the manuscript and summarized them above, which we've addressed above. There are a few other comments that we lifted from the Word file and address here:

Commented [AS2]: We addressed the confounding factors above with the new analyses presented in Extended Data Analysis and Discussion (EDAD). The includes observed changes in the number of samples shown in Extended Data Figure 2b. We bring this up again here because the reviewer points to one paper of many that describe small changes in cloud fraction. In our discussion of Extended Data Figure 2b, we caution readers not to equate cloud fraction changes with changes in CMV sampling, for reasons discussed in EDAD.

Commented [AS4]: Yes. But the reviewer's comment about increased moisture due to warming is also on point. So we have it to added to the text (Lines 130-132). Thanks!

Commented [AS10]: Sorry, but we didn't understand this comment.

Commented [AS12]: Thanks for asking! This gave us the opportunity to add an additional analysis that shows just how stable the MISR CMVs are. The results are shown in Methods (Lines 335-352). This also addresses the follow-on comment [AS13] and whether changes in the geometry between MISR views has changes. The remainder of [AS13] is on confounding effects, which we addressed above.

Commented [AS14]: We now cite several Aeolus papers. We acknowledge that the assimilation of Aeolus had a substantial impact on model forecast, and that they could be used for further ERA5 model validation as done in one study over Singapore. We discuss these papers in EDAD section "Comparison to other works" (Lines 526-584). But we didn't place the suggested edit, since Aeolus data hasn't yet been used to quantify biases in ERA5 throughout the troposphere – it was done in Banyard et al. (2024) over Singapore only. We have edited this sentence to clarify our unique contribution (Lines 269-270).

Commented [AS15]: Sampling trends are now included and discussed in responses to other comments above. Our statement here is just speculation. Its verification can only be had with a detailed analysis of the ERA5 data, which is well beyond the scope of our study.

Commented [AS16]: This paragraph discusses trends in upper tropospheric winds. Scatterometers would not be useful since they only retrieve surface winds.

Commented [AS18]: Both pieces of information that the reviewer points to here are important and used in EDAD in quantifying confounding effects and justifying our choice for a 1 km vertical bin width (Lines 453-466).

Commented [AS19]: Agreed, geostationary CMVs are not all that accurate, in part because of their poor cloud-top heights. In the studies cited, they were careful to address this point by restricting their analyses to cloud-top heights where geostationary and MISR agreed to within some height tolerance. We now point this out and we are careful to state that these are relative biases (Lines 333-334). We believe that the smaller MISR CMV wind biases relative to ground targets are truer because of its traceability. Still, we report the full range of MISR CMV error assessments so that we are complete. We use the more pessimistic estimate of MISR CVM wind biases from geostationary when interpreting wind differences with ERA5, just to be safe when we ascribe possible bias to ERA5.

Cheers,

Ad Stoffelen

We are extremely grateful for the detailed review provided by Reviewer 2. We hope that changes we've made to the manuscript have addressed them.

If the reviewer has additional questions or concerns, they may contact us directly so that we can discuss further. Thanks!

Re: Revised manuscript "Decadal changes in atmospheric circulation detected in cloud motion vectors" (MS # 2024-10-22368A-Z)

Point-by-point response to Reviewer 2's remaining minor comments

All Line Numbers in the response below are with reference to the revised manuscript. The reviewer's comments are in blue italic fonts and our responses are in black regular fonts.

To be brief, may I congratulate the authors with their improved, original and significant manuscript. Most of the reviewer's suggestions are taken up well.
Just to highlight a few remaining points:

We thank the reviewer for their encouraging remarks and appreciation of the improvements made to our manuscript. Their constructive comments were instrumental in strengthening the clarity and significance of our work. We address their highlighted remaining points individually below.

- The additional conditioning that ERA5 also be a cloud-top at the times and locations of the MISR CMVs had negligible impact on your analysis and it didn't matter what the cloudy or clear state was for the ERA5_MS samples, their comparison to MISR CMVs led to the same result. Given results on tropical moist convection with scatterometers this may not be surprising. King et al. (2022, https://doi.org/10.3390/rs14051147) show that extreme convergence and divergence are associated with heavy rain, representing resp. updrafts and downdrafts. Collocated ECMWF ERA5 winds do not show these extreme divergences and convergences and hence lack some of the dynamics seen with real rain clouds. One could say, it does not matter much whether moist convection exists or not in ERA5, while in the real world it does matter. Hence, could it be that MISR results are still affected by climate changes in moist convection, while it is hard to prove as models do not show real moist convection dynamics?

We appreciate the reviewer's insightful comment highlighting the limitations of ERA5 in representing moist convection dynamics. And we agree. We have updated the discussion in the Extended Data Analysis and Discussion section (EDAD) in Methods, Lines 625 – 633, with the points being made here by the reviewer.

- Grand biases in tropical (tropopauze) closure in the ECMWF model are shown against radiosondes at the locations of the radiosondes in Houchi et al. (2010, https://doi.org/10.1029/2009JD013196), in particular for vertical wind shear. Belmonte Rivas and Stoffelen (2019, https://doi.org/10.5194/os-15-831-2019) show rather large tropical

errors in ocean surface winds, while Zagar et al. (2025, https://essopenarchive.org/doi/pdf/10.22541/essoar.173776264.46417622/v1) highlight tropopause shear errors. These publications appear to link well with the MISR work presented and warrant discussion.

We thank the reviewer for highlighting these relevant studies. The publications suggested by the reviewer have now been explicitly discussed and cited in EDAD, specifically within the "Comparison to other works", Lines 693 – 698. These references indeed provide valuable context and further support our MISR-based observations regarding biases in ECMWF and ERA5 tropical and tropopause-level winds.

- Unfortunately, as may be understood from the previous points, radiosondes are insufficient to control the model dynamical biases in the tropics or elsewhere. Rather, moist convection processes and associated 4D dynamics are too fast (30 min. time scale) for a global NWP model to track with the observing system at hand. The poor representation of the associated dynamics is the main reason for dynamics that are independent of the presence of cloud, not data assimilation of radiosondes (cf. Houchi et al.).

We agree with the reviewer's point regarding the limitations of radiosondes and the importance of properly representing moist convection dynamics. In line with this comment and the reviewer's first comment, we have revised our earlier statement about the impact of data assimilation (radiosondes being only one example of the data being assimilated) and incorporated the reviewer's perspective. As in the first comment, this revision appears in Lines 625 – 633. Also, the original statement on the data assimilation point in Main Discussion and Conclusion has been removed; it now points the reader to EDAD for this discussion. See Lines 263 – 264.