Replies to editor comments

Dear authors,

The two reviewers are satisfied with the way the first revision has been made. But both reviewers have still some suggestions and comments you will have to consider. Like one of the reviewer, I would appreciate a better writing of the new paragraph in the introduction describing the choice of the two variables. Maybe in this paragraph or in the section method (subsection 2.1) it would be worth mentioning the link between the two variables. The temperature horizontal gradient intervenes in the Q-vector definition and in that sense the EGR is involved in the Q-vector definition. For a given upper-level disturbance, the omega-forcing might be stronger depending on the intensity of the baroclinicity. Of course the presence of a synoptic upper-level disturbance renders the two diagnostics different but maybe some sentences relating the two variables could be useful for future readers. I look forward to the new version of the paper.

Best regards,
Gwendal Rivière

Dear Gwendal

We polished the paragraph in the introduction (Line. 60) and added in Section 2.1 a comment on the link between the two variables (via the horizontal temperature gradient) and that they are not fully independent. In our study the EGR is computed at lower levels, while the Q-vector is computed at upper levels. In our case it is therefore not the same temperature gradient that enters the computation of both variables. But in general, we agree that this would be the case if the variables were computed on the same vertical level.

Best regards
The authors

Replies to comments by Reviewer 2

The authors have responded to almost all of my previous comments in a very satisfactory manner. There are two exceptions: details of the ΔSLP diagnostic (#3 below) and the effect of the bin size (#4 below) which in my opinion still require further consideration and revisions to the manuscript. Furthermore, when reading the revised manuscript, I noticed a few minor issues that lacked clarity (points 1, 2, 5, 6) and thus a few additional revisions are still warranted to improve both the clarity of the manuscript.

1. Line 60. This new paragraph in the introduction does not flow well with the rest of the text and it is not clear whether the purpose of this paragraph is the motivate / justify the choice of the two variables or explain what they are. I think this paragraph should be primarily motivation and this needs to be written more clearly. e.g “The main motivation for selecting these two variables is...”
The paragraph starting a Line 60 was polished and now contains our main motivation for the choice of the two variables. Our motivation results from the classical picture of baroclinic

cyclone development. As suggested by the reviewer we now start the paragraph explicitly with *"The main motivation for the selection of these two variables results from the classical picture of the extra-tropical cyclone development."* We also streamlined the remainder of the paragraph to make the choice of our forcing factors, hopefully, very clear to the readers.

2. Line 99. This sentence is confusing "The static stability is not constant in the domain, but a 1d vertical profile is used instead" as it includes reality (static stability varies in the horizontal and vertical) and what is done to compute to QG omega (static stability is assumed constant in the horizontal and only varies in the vertical). A few more words are needed in this sentence to make it clearer.
We added a few more sentences in Line 103. The need to use a 1d vertical profile arises from numerical problems in the inversion, which might occur in situations of near neutral or negative static stability. Instead of using the full 3d static stability, a horizontal average is used to compute a 1d reference profile at each time step. This pragmatic choice has been made already in previous studies and does not substantially affect the final outcome.

3. Figure 1. ΔSLP. The issue I previously raised (see reviewer 2, minor point 3) has not been resolved. It is still unclear when reading the manuscript exactly how this diagnostic is calculated, for example if the real time between time of genesis and time of minimum MSLP is 48 hours, are there four 12-hr values of deepening rate calculated (-48 to -36, -36 to -24, -24 to -12 and -12 to 0hr) or is a sliding window used or is only the maximum value used? Secondly, a sentence needs to be added to the manuscript explaining why there are positive values (weakening cyclones) in the distributions of figure 1 even when only the normalised times from -1 to 0 (the intensification phase) are considered.
We clarified this in Section 2.3 near L. 174. We do not use a sliding window or the maximum value, instead we use all time intervals between genesis and minimum MSLP. In your example of a 48 hour growth period there are four values that contribute to the statistic. For this reason positive values occur when the deepening phase is not a simple linear deepening and are accepted because we consider the full growth period.

4. Previously, I asked how the size of the bins and the differing number of points per bin may affect the results. Thank you for providing additional analysis on this matter. However, having seen Figure 6 in your replies, I do not think your conclusions (first bullet point in conclusions section and also text in section 3.2) as currently written are fully supported by your analysis - I think it is more complex than you state and there are a few subtle points that you should stress more clearly.
Firstly, Figure 6 in your reply, bottom right panel (zoomed in part) shows that, for this limited part of the parameter space, that the strongest deepening rates occur for high EGR but low QG omega. This is not consistent with Figure 3a in the manuscript and needs explaining. Potentially the mean values in this part of the parameter space are not statistically different though and considering the distributions may clarify this point.
Yes indeed. We must not expect a linear increase in the deepening rates with linearly increasing EGR and QG Omega. The standard deviation was added to Figure S6 of our manuscript for each bin to clarify this issue. Because the standard deviation in each bin is larger than the difference between two consecutive mean values between two bins, we must expect that due to a certain skewness of the distribution that very high values occur in a bin that has however a lower mean value compared to its neighbouring bin.

Secondly, when Figure 6 in your reply is considering together with the 2D histogram shownin Figure 3a in the manuscript, I think the correct interpretation of these figures / analysis is that the while the strongest deepening rates occurs for high EGR and high Q, strong deepening rates can also occur for high EGR and moderate values of QG omega. This is somewhat written in the first bullet point of the conclusions but I feel it is a result which should be stressed more and better explained. Related to this, Figure 3a strongly suggests that EGR has a stronger influence on deepening rate than QG omega – this is already somewhat touched on (but rather indirectly and briefly) by the authors when discussing the asymmetry between the bottom right and top left corners. This subtle result is quite interesting and the manuscript would benefit if this was highlighted more clearly and physically explained. One hypothesis to consider is can the instability (high EGR) be effectively released as long as there is a reasonable amount of upper level forcing? (moderate to high Q)? In summary, the authors should carefully revisit and revised section 3.2 and the first bullet point in the conclusions, potentially even splitting this conclusion into two.

This is a very interesting comment and suggestion. Thank you! A detailed analysis of the asymmetry between EGR and QGω forcing would be interesting, but we also think that it asks for a systematic and detailed analysis, and hence goes beyond the scope of this study. Still, we are happy to include the reviewer's comment into the manuscript. More specifically, we have added the following text to section 3.2 (closely following the reviewer's suggestion):

> *The asymmetry between the two opposite corners could also point to a (potentiallly) subtle difference in EGR and QGω forcing. It indicates that EGR has a stronger influence on the deepening rates than QGω, and that baroclinic instability might be released as long as there is a reasonable (even moderate) amount of upper-level forcing. In short, moderate upper-level forcing (QGω) might be sufficient to trigger substantial deepening rates if EGR is high. In contrast, if EGR is low, weaker deepening rates result even if substantial upper-level forcing is discernible.*

Additionally, as suggested by the reviewer, we discuss this specific point in an extra bullet point in the conclusions.

An interesting asymmetry is discernible between conditions with high EGR and low QGω forcings and, conversely, low EGR and high QGω forcing: larger deepening rates result for the former than the latter. This indicates that baroclinic instability and substantial deepening rates can result even in situations with moderate upper-level (QGω) forcing as long as the EGR is high enough. The opposite is, however, not true: even substantial upper-level forcing results only in weak deepening rates if EGR is low.

5. Line 160. Suggest you revise “49 bins” to “49 linearly distributed bins”.
Corrected.

6. Line 276-277. This sentence could be clearer – currently it sounds rather alarming (that the results are almost meaningless). I think the authors intend to say that the mean evolution shown in Figure 4 is not representative of any one individual cyclone lifecycle. Please revise this.
The sentence was revised accordingly (now L. 286).

7. Figure 4. While this is a nice addition to the manuscript, showing only the mean values clearly hides the large amount of variability as the values on the y-axes cover much smaller ranges than what is shown on the x- and y-axis of Figure 3. Is it possible to add some additional lines to this figure e.g the 25th and 7th percentile values?
The Figure has been updated and includes the mean values + 0.5 STD in either direction.

Replies to comments by Reviewer 1

The authors have satisfactorily addressed all my comments and concerns and I think the paper is suitable for publications after some minor revision. I only have a few comments:

1. I still think it's worth adding the separate Q and E maps (Fig. 1 in the response) to the SI, since it does provide additional information. It shows for each one of the cases Q(up,down) E(up,down) where the main contribution is coming from (from either Q,E, or from both), which is interesting.

The separate Q and E maps have been added to the Supplement.

2.Fig.2a in the paper- is it possible that there is some mistake in the calculation here? Why is the density distribution almost three times larger than in the other cases? Especially after seeing Fig.1 in the response, this is not clear to me, since the region of Q_up and E_down do not overlap much. For example, based on Fig.1 in the response, I would expect Q_up and E_up (whose regions overlap) to have higher density than the former.

It seems odd that the Q_up E_down density is so much higher than the other ones. As you stated, it feels counterintuitive when considering the separate Q and E distributions. However, one thing to note is that when calculating the density for the separate distributions, we use a different data range. For instance: For E_down, we use only the lower range of EGR but use the entire range of QGomega (in the Forcing Histogram, this corresponds to the first two columns and all rows). However, for Q_up E_down, we only use the first two columns and first two rows (upper left corner of the histogram). This difference in data range may account for discrepancies between the combined forcing distributions and the separate distributions for E and Q.

3. Can this study be useful for understanding why the midlatitude storm tracks are too zonal in global simulations (e.g., in historical CMIP simulations)?
Yes, thank you for pointing to this potential application of the method. We added the following final statement at the end of the paper.

> *Finally, the combined EGR and QGω perspective could be used to identify biases in the representation of cyclone dynamics and storm tracks in present-day and historic CMIP6 simulations (Priestley et al., 2020). For instance, midlatitude storm tracks are known to be too zonal in historical CMIP6 simulations, and the robust EGR/QGω diagnostic could readily be applied to temporally coarser dataset than the six-hourly ones used in this study.*