February 13, 2024

Responses to comments of "Assessment of object-based indices to identify convective organization" egusphere-2023-1985 to GMD

We deeply thank the referee for the very insightful and helpful comments which show that he did understand very well our work. The suggestions offered have resulted in revisions to the manuscript, which enhances its overall quality.

In the following, we have made a point-by-point response to the comments and revised the manuscript accordingly. The referee's comments are presented alongside our responses (in blue) and the textual modifications (in red). Both authors have thoroughly reviewed the revised manuscript and unanimously agreed to its submission in this improved form.

1 Conceptual

Comment

1. I_org assessment
It is interesting that the I_org metric is so sensitive in all categories of robustness criteria. I wonder to what extent the sensitivity relates to the number of objects considered. Is I_org robust when considering a large number of objects? How large fraction of the scenes considered in this study have less than 35 objects (where the I_org metric is no longer reliable)?

The I_org metric has been identified as unreliable for a small number of convective centroids (<20) before:
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019GL086927
Not in a systematic way as in the present study, but perhaps it should still be mentioned.

Answer

We thank the referee for suggesting the work of Semie and Bony (2020) that has been cited in the manuscript.

In some tests, the sensitivity strongly depends on the number of objects, for example, conditions 1 and 2. For this reason, the sensitivity has been shown as a function of the number of objects in Fig.5. On the other hand, the results shown in Fig.4 and Table 1 include all events therefore they are averaging events where $N < 35$ and where $N > 35$.

In some other tests, the sensitivity does not depend very much on the number of objects (see later), and for those tests I_{org} can be robust also where $N < 35$. Therefore, we did not show any results as a function of the number of objects similar to Fig.5.

The fraction of scene as a function of the number of object can be roughly seen in Fig.1. Scenes where $N < 35$ ($N > 35$) are 93% (7%) of the total.

Changes in Manuscript (line 231)

The index I_{org} is very sensible to noise at a low number of objects, and it becomes noise-safe when more than 35 objects are present. This result is in agreement with Semie and Bony (2020) who raised a similar statement of robustness for I_{org}.

Comment

In several studies, I_org is used to assess aggregation from local minima in brightness temperature, which includes multiple convective centroids (convective cores) in a large convective object. With this approach the area of convective objects is implicitly included (as a large convective object introduce several closely connected convective cores), and the issue with a small number of convective centroids is addressed. Perhaps assessing I_org from the convective cores approach can be insightful to better understand the utility of this metric. Otherwise, perhaps just presenting the alternative approach and clarifying that the statement in the present study relates to the method applied to convective objects.
In this study:
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019AV000155
the authors mention that a similar result is obtain from using the object-based approach and the convective cores approach. Albeit, in that study the domain is the whole tropics, so the number of convective centroids is likely sufficiently large regardless.

Answer

We thank the referee for raising such an important point that was not mentioned in the manuscript.

In this work, we have used convective object reconstruction. The conditions 1-7 are designed to be satisfied for objects with a real size (e.g. condition 2), and all the presented results relate to this approach. This is the only approach that can be used to directly compare so many indices of organization. On the other hand, I_{org} and L_{org} can also be applied to local minima, more related to convective cores. However, such a method supposes point-like objects, and therefore it precludes the possibility to compute all the other indices.

Following the referee's remark, I_{org} and L_{org} can be applied to both local minima and convective objects. In this case, they can have different behaviors. Such a clarification has been added in the manuscript in the conclusions when all the results are presented for each index.

Lastly, the referee's question is part of a wider question, that is how to define the objects under study. There are many possibilities: mesoscale convective systems defined by brightness temperature, deep convective cores defined by strong precipitation, convective cores defined by local minima in brightness temperature, and many others. This question would lead to a new assessment, which focuses only on object definition.

Changes in Manuscript (line 396)

The above-mentioned behaviors of I_{org} are obtained using convective object reconstruction. However, since I_{org} does not consider the size of the objects, some studies (e.g. Semie and Bony, 2020; Bony et al., 2020) applied it to local minima in brightness temperature which may be seen as a proxy for the convective core positions. Such an approach may modify the behavior of this index. Identical considerations apply to L_{org}.

Changes in Manuscript (line 401)

Equivalently to I_{org}, the present results hold only when convective objects are reconstructed.

Comment

2. ROME assessment
From working with the ROME metric to assess the tropical domain with DOC, I know that the metric is highly correlated with mean area of the domain. It was interesting to see that the impact of changes to the proximity of convective regions was so small. I suspect changes to the proximity of convective objects has a greater relative effect on the metric if the scenes are sub-sampled to scenes with similar energetic constraints (similar mean convective area, similar vertical velocity, similar mean precipitation etc.). Perhaps the distance component of ROME is also more significant when very large distances are considered (where the squared edge distance is a larger number of multiples of the smaller pair object). Further, considering that the proximity scaling applies to every object pair, a larger number of objects all moving together may also highlight the proximity scaling. I reserve the possibility, that the metric simply is unable to factor in the proximity of objects as it occurs in realistic settings, past the change in proximity which results in joining two objects. However, it would be interesting to test some of these considerations to highlight the limitations / utility of this metric.

Answer

Indeed it would be interesting to subsample the events and test the indices behaviors for each sample. We did not fully compute such a study, however, we checked a few cases and we report here a few examples. Fig.R1 displays the sensitivity test of condition 3 on four different subsamples of the dataset.

The upper panels show the proximity test for two bins of the number of objects. Despite the large difference in the number of objects, the results are similar. The lower panels show the proximity test for two bins of the total area of convection ($\sum_i A_i$). The bins of area of convection are about 4% and 40% of the domain under consideration. Of course, the results depend on the number of objects and the total area of convection, however, the dependency is small if compared with conditions 1 and 2.

Comment

3. ABCOP assessment
In the conclusion, it sounds like it is recommended to use this metric. While the metric captures

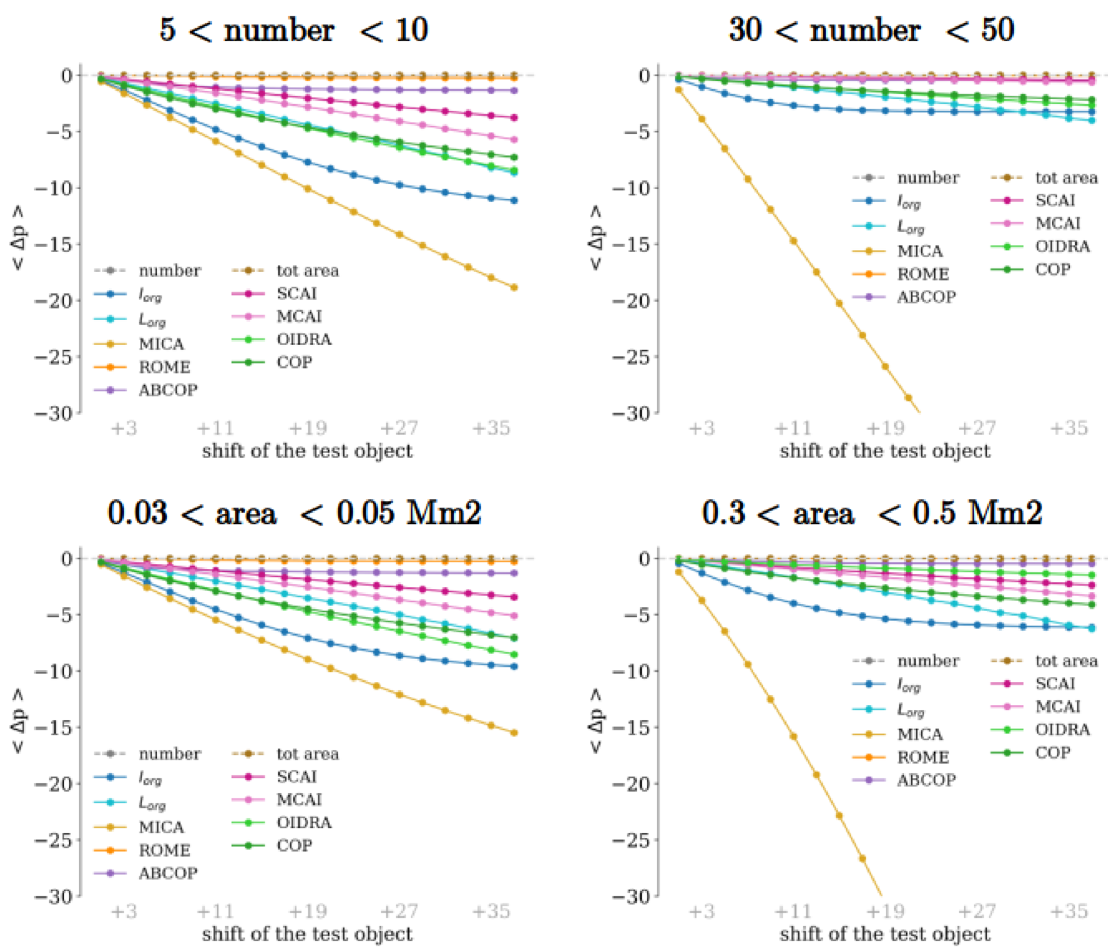

Figure R1: Following the referee's comments, the figure presents the sensitivity test to proximity (condition 3 of the manuscript) for 4 different subsamples of the data. The subsamples are: events with a number of objects within 10 and 15 (top left), events with a number of objects within 30 and 50 (top right), events with a total area of convective objects within 0.03 and 0.05 Mm^2 (bottom left), events with a total area of convective objects within 0.3 and 0.5 Mm^2 (bottom right).

changes in proximity, and is robust in most criteria, from what I understand, the metric does not correctly capture fundamental changes in aggregation; adding a random single convective gridbox increases aggregation, and merging objects decrease aggregation, which are the opposite signs of change from the conceptual interpretation. Perhaps it can be highlighted that these features make the metric unsound in this regard.

Answer

The referee does understand well the results, and he is right. ABCOP is well behaving in most criteria, but the two features he mentioned make it a bad candidate to spot convective organization.

We changed the conclusion to highlight more of those behaviors.

Changes in Manuscript (line 407)

- ABCOP: The index ABCOP has a large correlation with the total objects' area $\sum_i A_i$. Therefore, it mostly reflects the total objects' area. This relationship uniquely comes from the larger objects in the images, while the smaller objects do not play a significant role in the value of ABCOP. This behavior comes from the $\max_{j\neq i}()$ function in the ABCOP definition (equation A8), which gives great importance to large objects. As a consequence, ABCOP is not very sensitive to noise. However, it strongly depends on the number of objects because it is defined as a sum over each object instead of as a mean like COP. This characteristic is visible in two of the studied conditions and it negatively influences ABCOP response to noise. First, when one convective grid box is added randomly in the domain, ABCOP incorrectly increases instead of decreasing because the number of objects is increasing. Second, when two close objects are merged, ABCOP incorrectly decreases instead of increasing because the number of objects decreases. In addition, ABCOP follows the behavior the number of objects even under a change of horizontal resolution. ABCOP proves to be robust under changes in resolution, shifts in time, and shifts of the considered domain. Last but not least, the index ABCOP increases with the proximity of the objects but slower than the majority of indices.

2 Technical

Comment

When testing condition 4 (changing the size of one object) - from the schematic in figure 6, it appears that the edge of the test object effectively move closer to the other objects when the area of the test object increases. Consequently, there will be a proximity component affecting metrics that depend on the edge distance between objects. To avoid a proximity influence when testing condition 4 on ABCOP and ROME, the test object could be uniformly extended Eastward in these cases.

Answer

The referee raised a good point. The perturbation of condition 4 modifies also the distances between objects for 4 indices: ABCOP, MCAI, ROME, and OIDRA. ABCOP and MCAI are affected because they use the effective distance $d_{eff} = d(i,j) - r_i - r_j$, where r_i is the equivalent radii of the object. ROME and OIDRA directly use the distances between edges via the python package "shapely". For these four indices, the influence of proximity should be removed as suggested by the referee. However, if we implement such a correction, the results of the original test on I_{org}, L_{org}, COP, MICA, and SCAI cannot be compared with the results of the new test on ABCOP, MCAI, ROME, and OIDRA. Therefore, we decided to perform one single test on all the indices.

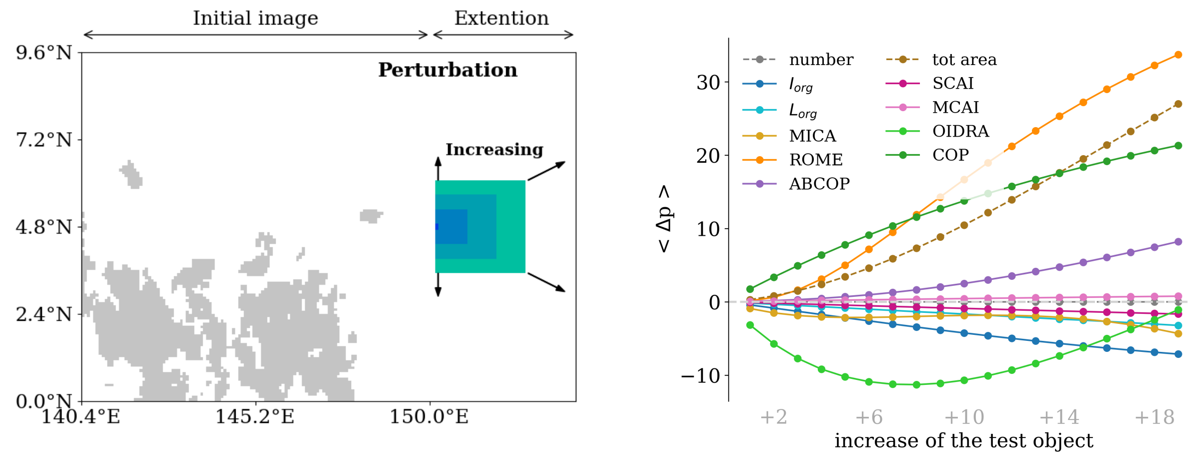

Figure R2: Sketch of the perturbation suggested by the referee (left) and the associated $< \Delta p >$ for all the indices (right).

To provide a complete answer to the referee, we have run the test the referee has suggested, and the results are shown in Fig.R2. For each index, the result can be compared with condition 4 shown in Fig.6 and Fig.7 of the manuscript.

The sensitivity to this new perturbation is similar to the one of condition 4. The main difference is that I_{org}, L_{org}, MICA, and SCAI have negative trends because are affected by the proximity component.

Comment

In the methodology section, it was mentioned that scenes with one or no objects were removed. What fraction of scenes contain only one object, and are they significant for describing the degree of organization? Further, are they small objects or very large objects spanning most of the domain.

Answer

Only 4.1% of the scenes contain no objects. Only 2.9% of scenes contain one object. Therefore we did not consider these 7% of the scenes.

The scenes with a small number of objects usually have very small objects. It is partially shown in Fig.1 of the manuscript, where the total coverage of convective objects is plotted as a function of the number.

Changes in Manuscript (line 56)

Then, images with less than two objects are rejected (4.1% and 2.9% of the events with no object and one object, respectively).

Comment

Is it important for a metric to be able to handle a singular large object for the 10x10 degree domain?

Answer

The referee asked a very meaningful question that does not have a well-defined answer. To date, there is no consensus on how a metric should behave when one singular large object spans the domain, and the authors do not have better insights on the topic than other experts. Such a question should be discussed in the scientific community.

Comment

In the introduction in line 27 there is a statement: 'Such studies have been so far performed only for example cases'. It could be nice with a reference to this. In the ROME metric paper, for example, there is an evaluation of different organization metrics for example cases, so maybe that paper can be referenced here:
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019JD031801

Answer

We thank the referee for the suggestion. The right references have been added to the text.

Changes in Manuscript (line 26)

Such studies have been so far performed only for example cases (Retsch et al., 2020; Jin et al., 2022).

Sincerely,

Giulio Mandorli,
Claudia Stubenrauch