We sincerely thank Reviewer#2 for his/her comments, which gave us the opportunity to deepen some points of our paper. We indicate here our reply in blue colour and some corrections we propose to implement on the text of the manuscript in italic red. The full version of our reply, if accepted, will include all the corrections declared here.

The manuscript "Characterisation of extreme events waves in marine ecosystems: the case of Mediterranean Sea" describes a new method to characterize extreme events bases upon simulated chlorophyll concentrations for the period 1994-2012. Using a cluster analysis applied to a set of indices that define the occurrence of extreme events waves, different ecosystem regimes were defined. In my opinion, the manuscript is very interesting and deserves publication in Biogeosciences after minor to moderate revisions. In detail:

1) The language should be checked by a native English speaking person. There are several typos that should be removed. For instance in line 270 there is a reference to Fig. 9 that does not exist.

Thank you for the suggestion. We will carefully revise the text references to Figures and Tables and we will send the new version of the manuscript to an English Editing Service.

2) The method uses surfaces chlorophyll concentrations averaged over the uppermost 10 m and does not consider vertical profiles. Please discuss why subsurface blooms do not play a role.

We considered surface chlorophyll particularly suitable to show the functioning of the method, since surface chlorophyll has been widely investigated in literature and it is comparable also with remote sensing measurements (as done in Sect. 3.1, Fig. 4). However, also subsurface processes (e.g., associated to the Deep Chlorophyll Maximum feature) can be analysed by our method, using the same model implementation (which provides 3D chlorophyll concentration fields). For sure, it can be a further interesting application of our method and we will specify it in the new version of the manuscript, in the Discussion section.
In particular, we propose to substitute the two sentences at lines 360-364 with:
We have applied the method to the surface chlorophyll, as one of the more representative and investigated variables of the marine ecosystem, identifying maxima of chlorophyll as events which potentially influence the ecosystem function (e.g., food web and carbon fluxes). However, our method can be applied also to integrated chlorophyll, to account for subsurface growth of phytoplankton, or to other variables whose impacts on the ecosystem can be relevant, such as HAB-like phytoplankton groups (Vila and Masó, 2005), provided the availability of a continuous dataset in time and space.
The method may also be applied to other specific subsurface quantities, as for example the surface identifying key biogeochemical properties as the Deep Chlorophyll Maximum (e.g. as defined in Salon et al., 2019), the oxygen minimum or the oxygen deficiency (OSPAR, 2013; Ciavatta et al., 2016).
Anyway, as replied also to Reviewer#1, we decided to avoid the reference to "bloom" processes in the new version of the manuscript, and to restrict our argumentation to extreme events (i.e., maxima of surface chlorophyll distribution).

3) It is discussed but I am still worried about the stability of the identified regimes using different thresholds. You showed that the clustering of the mean values of the indices do not change much. However, are there changes in the spatial distribution of the regimes shown by Figure 6?

Thank you for this comment, which gives us the opportunity to deepen some aspects.

We applied the fuzzy k-means analysis also in case of 98th and 99.5th percentile thresholds and we report the results in Figs. R2.1 and R2.2, respectively.

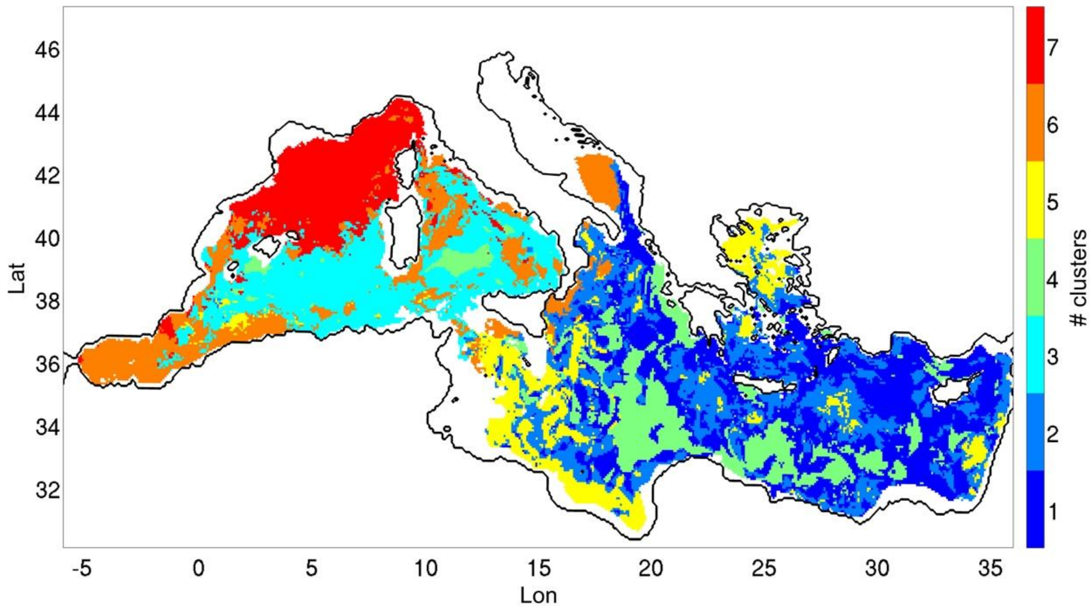

Figure R2.1 Fuzzy clusters with maximum membership, in case of 98th percentile threshold.

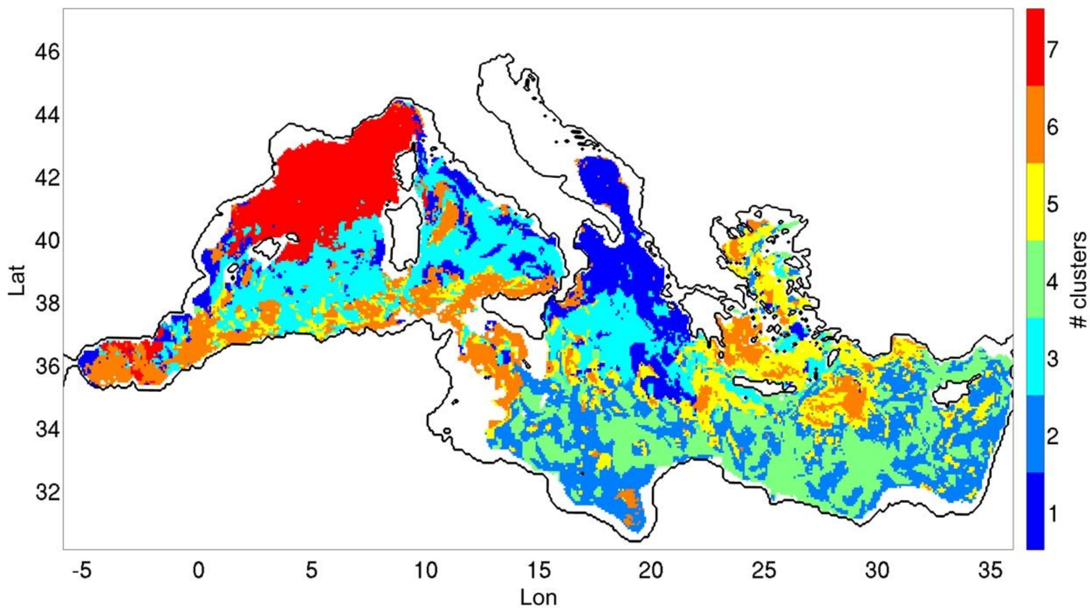

Figure R2.2 Fuzzy clusters with maximum membership, in case of 99.5th percentile threshold.

From a general point of view, the clusters distribution for 99.5th percentile threshold (Fig. R2.2) is very similar to the clusters distribution for 99th percentile (Fig. 6 of the manuscript), while the clusters distribution for 98th percentile (Fig. R2.1) differs mainly in Ionian and Levantine Sea.

More in details, the spatial distribution of clusters #3, #6, #7 in the western basin does not considerably change with respect to Fig. 6 of the manuscript, as well as the spatial distribution of cluster #4 in the eastern basin.
Clusters #1, #2, #3 #5 and #6 display instead differences, mainly in the clusters distribution for 98th percentile, in the eastern basin, i.e., southern Adriatic Sea, central Ionian Sea, Aegean Sea, Libyan coast and eastern Levantine Sea.

Since the identification of clusters depends on four indexes (i.e., mean severity, uniformity, duration, anomaly) which do not necessarily scale in the same way with the local threshold, the fact that the spatial distribution resulting from the fuzzy k-means analysis (as "combination" of the four indexes) for different thresholds can differ from Fig. 6 is a reasonable result.
Moreover, it should be highlighted that Figs. 6, R2.1, R2.2 show the clusters referred to the maximum membership.
In addition to maximum membership, we considered also the "confusion index" (Burrough et al., 1997), which, applied to our case, quantifies how well each point of the Mediterranean domain has been classified. High values of the confusion index (CI) index are related to higher sensitivity of some areas in the cluster classification with respect to variations in the local threshold (i.e., to differences of Figs. R2.1, R2.2 with respect to Fig. 6).
We estimated the confusion index (CI) as:
$CI=1-(MF_{max}-MF_{max2})$,
where MF_{max} denotes the dominant membership value and MF_{max2} is the subdominant membership value for each point, and we computed it in case of 99th percentile threshold. Figure R2.3 shows values of CI greater than 0.7 (i.e., "high values" of CI) as black dotted points.
We can observe that most of the areas displaying differences in the spatial distribution of the clusters with respect to variations in the threshold (Figs. R2.1, R2.2) correspond to high CI in the reference case (Fig. R2.3) and that the identification of the clusters generally appears to be consistent with the other two clusterizations.

We recognise that this point deserves a revision in the manuscript.
Therefore, we propose to substitute Fig. 6 with Fig. R2.3 in the revised manuscript, adding the expression:
with black points indicating a confusion index higher than 0.7
at the end of the figure caption and the definition of the confusion index in the text. Accordingly, our statement about the robustness of the classification (old lines 387-390) will be reformulated and referred to this new version of Fig. 6.

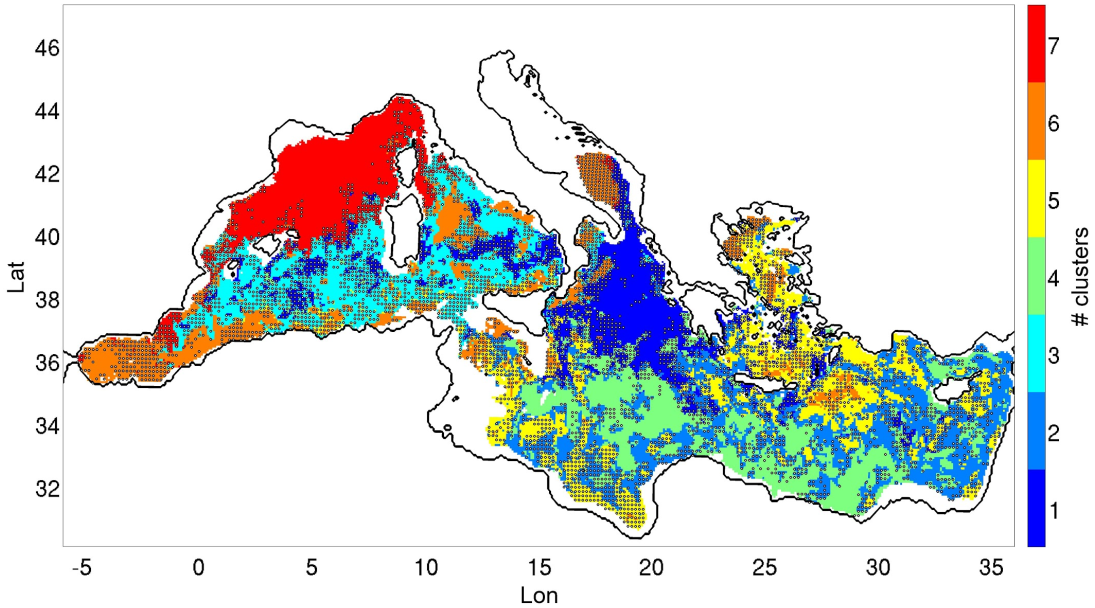

Figure R2.3 Fuzzy clusters with maximum membership, in case of 99th percentile threshold, with black points indicating a confusion index higher than 0.7.

4) In Section 4 and the Appendix, ecosystem dynamics characterizing some of the regimes are discussed. However, some clusters lack any dynamical explanation and might be rather artificial. It would increase the scientific value of the manuscript if you could discuss these clusters more in detail as you have done it for NWM in the Appendix.

Section 3 and Appendix of the first version of the manuscript present in detail one of the Extreme Event Waves (EEWs) identified in the 1994-2012 dataset, to show how the method works. In particular, Appendix highlighted that the method catches all and only the relevant information of the event. We selected the EEW associated to the highest value of mean severity of the whole dataset (and occurred within the area covered by cluster #7) as a particularly meaningful case.
Since the main focus of the paper is to propose a method to identify and classify EEWs, rather than to analyse in detail surface chlorophyll dynamics, we did not include other examples of EEWs (e.g., associated to areas covered by other clusters, as suggested) in the previous version of the manuscript, avoiding to enlarge too much the length of the manuscript and to shift the attention more on the specific application than on the method. Thus, we would prefer not to add new figures and comments in Appendix.
However, if the Reviewer#2 still suggests to do it, we are open to include other examples of EEWs occurring in areas covered by other clusters.

5) In the abstract you mentioned that "There is a growing interest about events that can affect ecosystem functions and services in a changing climate". However, is the method suitable for following the temporal shifts in the regimes without any discontinuities? I suggest that you split the period in half and that you use the cluster analysis for both periods. By this, the impact of trends in some of the indices on the spatial distribution of the regimes might be investigated and compared with observed changes.

Thank you for the comment.
Yes, our method is suitable for following the temporal changes in the characteristics of the EEWs, provided that the time series is long enough for a robust trend or regime shift evaluation on a multi-

decadal scale. Unfortunately, our time series is not very long (i.e., 19 years). Thus, we think that the approach to split the chlorophyll dataset in two parts, i.e. 9-10 years for each part, would not guarantee statistical robustness to the analysis.
Indeed, we conducted a non parametric analysis of the trend slope (Theil-Sen method) and showed the results in Tab.2 of the manuscript, by reporting the cases in which the ratio between slope and mean is higher than 1% (red cells) and lower than -1% (blue cells). For sake of clarity, we report here the annual time series (and trend slope) of each cluster for the duration, uniformity, mean severity and anomaly indexes (Figs. R2.4-R2.7). As it can be intuitively grasped, the length of the time series does not allow further speculations on temporal changes of the EEWs characterisation and classification.

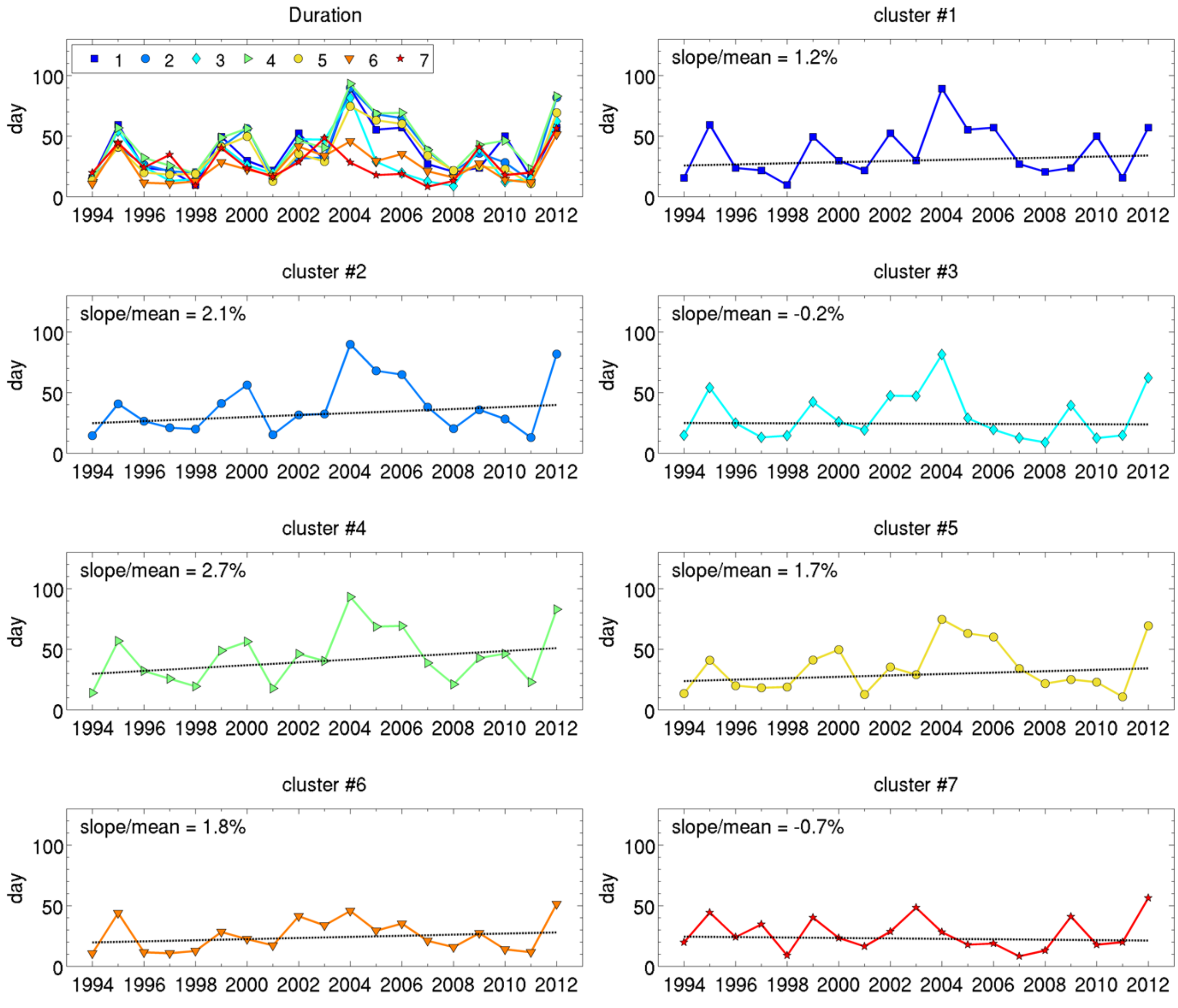

Figure R2.4 Trend evaluation (as slope over mean) on the clusters in Fig.6, referred to duration index.

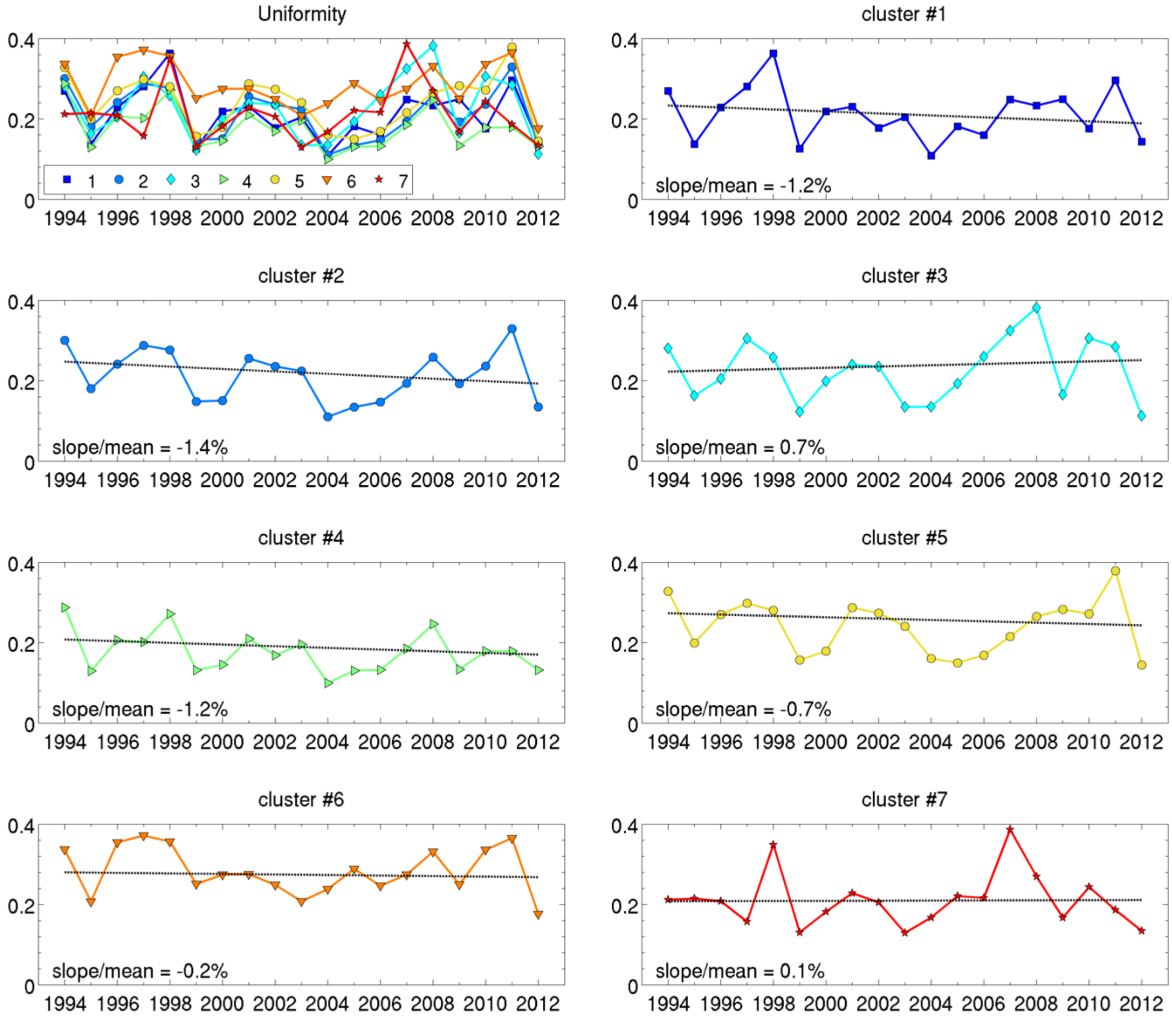

Figure R2.5 As Fig. R2.4, but referred to uniformity index.

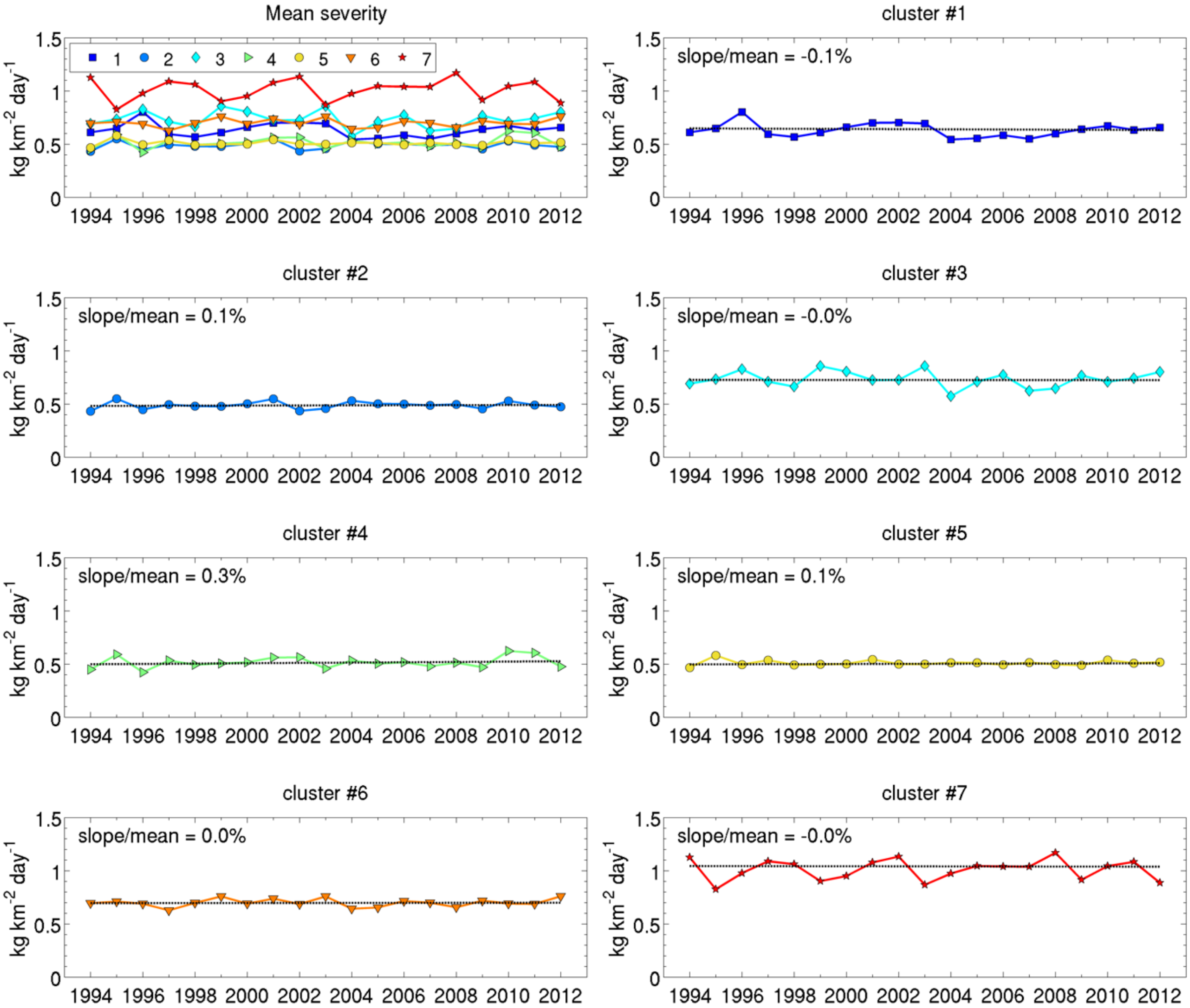

Figure R2.6 As Fig. R2.4, but referred to mean severity index.

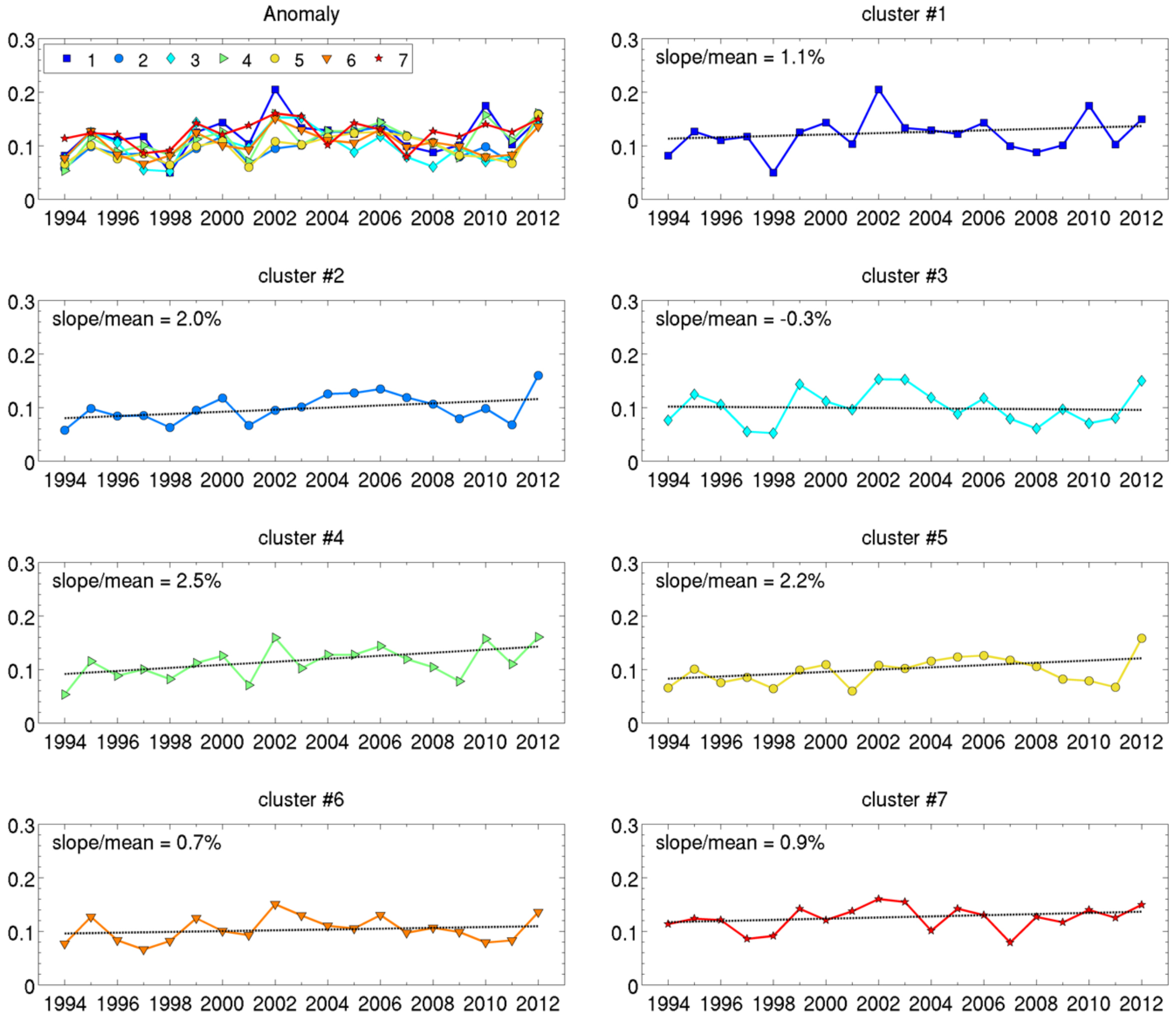

Figure R2.7 As Fig. R2.4, but referred to anomaly index.

Bibliography:

Burrough, P. A., van Gaans, P. F. M. and Hootsmans, R.: Continuous classification in soil survey: spatial correlation, confusion and boundaries, Geoderma, 77(2), pp. 115–135. doi: https://doi.org/10.1016/S0016-7061(97)00018-9, 1997.

Ciavatta, S., Kay, S., Saux-Picart, S., Butenschön, M., and Allen J. I.: Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem, J. Geophys. Res.-Oceans, 121, 1824–1845, https://doi.org/10.1002/2015JC011496, 2016.

OSPAR: Common procedure for the identification of the eutrophication status of the OSPAR maritime area, Tech. Rep. 2013-8, London, UK, available at: https://www.ospar.org/documents?d532957 (last access: 28 October 2019), 2013.

Salon, S., Cossarini, G., Bolzon, G., Feudale, L., Lazzari, P., Teruzzi, A., ... and Crise, A.: Novel metrics based on Biogeochemical Argo data to improve the model uncertainty evaluation of the

CMEMS Mediterranean marine ecosystem forecasts. Ocean Sci., 15(4), 997-1022 https://doi.org/10.5194/os-15-997-2019, 2019.