The manuscript "The colonization of the oceans by calcifying pelagic algae" by B. Sucheras-Marx et al. describes colonization of the oceans by coccolithophorids since the last 200 M. This well written manuscript is based on the compilation of nannoplankton accumulation rates in sediments, brought in context with previously published species richness, coccolith size as well as atmospheric CO2. Results indicate a colonization of the oceans in distinct phases, shaped by the reproduction strategy, interactions with other planktonic organisms and the physical environment.

The research is original and provides interesting findings to the community. The data set compilation seems to have been carried out with great care, even though, sadly, the available data is confined largely to the Atlantic, therefore I would suggest to maybe rephrase the main conclusions of the manuscript from "World Oceans" to "Atlantic". The manuscript is concisely written, however, could benefit from a re-organization of the Discussion paragraph in my opinion, so that each phase is discussed in its own paragraph, instead of discussing the colonization twice in 4.1 and 4.2.

I have some reservations regarding the smoothing of the NAR and the seemingly arbitrary reference to sometimes the smoothed trend and sometimes the underlying raw data. The authors should carefully re-examine each time the NAR is discussed and elaborate on when which datatype is discussed (see major comments below).

I would recommend publication of this manuscript after minor revisions have been carried out. I wish the authors good luck with the revisions and remain available for further feedback and discussions.

Please see my comments below (p=page, l=line):

Major comments:

NAR calculation: Since the majority of the manuscript hinges on the NAR, it would be great if the authors could provide an propagation of error for the NAR values, as they are calculated from 3 other variables. Additionally the NAR in Fig. 2 has a high variability of several orders of magnitude, can the authors elaborate on this a bit, e.g. is this caused by pooling different ocean locations, where changes could have happened at a different point in time?

Smoothed curve versus raw data: Currently, in some time periods smoothed NAR values are discussed and sometimes the raw data. Please state each time, which data is taken (raw data or smoothed trend). Please be careful in not mixing the two.

e.g. p9 l29 " a steady production for the rest of the epoch" seem to be rather subjective, as there seems to be rather a huge variability in observed NAR post K-Pg until the end of the Paleocene, just the chosen smoothing factor results in a steady NAR. How have the authors assessed "stable phases" in NAR versus "changing phases" of NAR? Only by visual observation of the smoothed trend?

By just looking at the smoothed curve, variability in the NAR data is lost. While I agree that in some time points a SF of 0.1 is influenced by the sampling resolution, however, in other

time points variability and trends are lost by a higher smoothing factor (e.g. the increase in NAR since the middle Paleogene, which is "smoothed away" otherwise).

Furthermore (p9 l27) here the average NAR shows no change during the K.Pg event, but NAR clearly changes, which is also discussed.

Layout Figure 2: please mark the individual colonization phases in a way, that they are easy to be put into context with the NAR record. Currently, the phases are indicated on the far right and the NAR record is on the far left, making it hard to see the exact phase changes. I would suggest shading of the background. Please also indicate the Torcian and Valangian. And add a line for the K-Pg event, as some of the statements (e.g. p9 l28 " ..the NAR recovered to pre-extinction levels within less than 4 Myr") are hard to follow with the current Figure layout.

Minor Comments:

p2 l2: represent (without s)

p2 l6-13: also refer to the Kuenen Event in the discussion or remove from Introduction

p3 l17 mm^2

p4 Fig. 1 caption: type of outcrop: rephrase outcrop; deep sea drilling is not an outcrop

p5 l5: SI= suppl. inform. (define).

p7 l 6: I would structure the paragraph according to the different phases, e.g. add a break in the middle of l. 6.

p7 l 14: regarding the versatile readership of BGD, I would refrain from using too many specific terms such as *Cope-Deperets* rule, which are not explained in the Introduction, same for *Margalefs mandala* in p9 l12, also explain briefly K and r strategists (for readers from a more geological background).

p7 l 17: ecospace or ecological niche?

p7 l 24: dominance: rephrase, as modern oceans are not dominated by *Ehux*, but it is the dominant cocco

p8 Fig 3: please add also a time stamp to panel c (Valanginian?)

p9: when the term *species* is used, calc. nannoplankton species is meant? or

coccolithophorids?

p 9: I find the terms *R-pole* and *K-pole* confusing, are these commonly used terms? Or do they just hint towards the respective areas in Margalefs mandala?

p9 l21: the maximum occurs much later, this need to be rephrased

p9 l24: please explain "roughly stable"

p9 l. 32: where is the "ecological specialization" seen in the data?

p10 l10: What are "red lineage algae"? Those belonging to the Red Queen Model?

p10 l 18 - 20: please add citations

p 10 l 28: where is the "post crisis Invasion period" in Fig 2?

p10 l 31: "smaller sized species than in the Mesozoic": to me it looks like the average coccolith size is relatively the same between this period and the Jurassic portion of the Mesozoicum

p11 l 1 The "decrease in pCO2" during the Neogene is not visible in Fig2, maybe another dataset would be more suitable? Also, how do the authors then explain the stable coccolith mean size and increasing NAR during the Jurassic, where pCO2 showed the largest drop?