Earth Syst. Sci. Data Discuss., 7, C110–C127, 2014
www.earth-syst-sci-data-discuss.net/7/C110/2014/

Open Access Earth System Science Data Discussions

Interactive comment on "Global ocean biomes: mean and temporal variability" *by* A. R. Fay and G. A. McKinley

A. R. Fay and G. A. McKinley

arfay@wisc.edu

Received and published: 18 June 2014

Author Response to Review #2 Journal: Earth System Science Data (ESSD) Title: Global ocean biomes: mean and temporal variability Author(s): A. R. Fay and G. A. McKinley MS No.: essd-2014-7 MS Type: Review Article

General comments: Fay and McKinley present maps showing delineation of "biomes" obtained by applying limits to time-variable interpolated gridded datasets. Using these definitions, they provide 13 annual mean maps of 17 biomes between 1998 and 2010 and show that the area of some biomes changes in time. They also isolate clusters of grid points whose biome assignment does not change in time to highlight regions where biome definitions are temporally stable. Overall, I think this paper suffers many flaws in

logic and organization. There is also no estimate of error and no sensitivity analysis is presented to assess how the results from their classification scheme depend on their choices of parameter threshold. Lastly, since the authors re-use existing and already published data sets and only provide limited analysis and interpretation, I don't think that the material presented in this manuscript is sufficient to warrant publication.

In summary, I think the main problems are:

1) No error/sensitivity analysis

>We have included discussion of the uncertainties in the input products (Sections 2.2-2.4) and of sensitivity of the results to alternative input datasets (GlobColour, Reynolds SST, others) in the Discussion section.

2) No justification for the parameter thresholds used

>We have improved our justification at the start of Section 2.

3) Limited interpretation and no mention of the possible role of the seasonal cycle. The interplay between the seasonal cycle of MLD and Chl a seems particularly relevant, and so is the seasonal cycle of sea ice.

>We have enhanced our interpretation of the results. Other authors have directly addressed biome seasonality (Reygondeau et al., 2013, Landschützer et al. 2013). We include the amplitude of annual seasonal variation as part of our biome definitions (maxMLD, spring/summer Chl). Our goal is to consider year-to-year variability, which has not been accomplished by these previous studies.

4) Not clear what is gained (aside from the temporal aspect) from using the limited number of biomes described here as opposed to other published biomes and studying the temporal variability of these existing definitions. It feels like a step back.

>As discussed in more detail below, the community of scholars considering large-scale variability and trends in the carbon cycle has not taken up the use of previously defined

biomes (e.g. Longhurst 1995; 2007). This indicates that the global ocean biogeochemistry community has a need for a consistent, simply defined set of biomes that are reasonable for large-scale aggregation of data and analysis; and that also offer the potential to be used in analysis of numerical models.

5) Organization of the text: Section 2.1 for example, describes some biomes, but there is no justification why these are the focus of the manuscript or how these are defined.

>We justify that these are open-ocean regions of interest to the global carbon cycle and other studies of global scale biogeochemical cycles.

6) No explanation is provided why the parameters used here (SST, Chla, Ice, MLD) are accurate descriptor of biogeography. There is also no definition of biogeography.

>We directly define biogeography in the Introduction and include discussion of why these are the appropriate variable for use in Section 2.

7) They claim the community will want to use these biomes to study carbon fluxes, yet, they do not allow the equatorial regions used to define equatorial biomes, home of El Nino (one of the main driver of inter-annual variability), to vary.

>We have changed the equatorial regions to be defined by the same variables as the other biomes.

Specific comments:

The list below provides specific comments related to the text.

p108/L20: perhaps “latitudinally-defined “

>Thank you for this suggestion. We have made this correction.

p108/L25: It seems to me that heterogeneities on land are much greater than in the ocean. Be more specific about what you mean by biogeography and can that definition apply similarly on land and in the ocean?

>We concur that spatial decorrelation length scales are smaller on land than in the ocean, but nevertheless, the land biosphere is frequently grouped into very large biomes (e.g. tundra, boreal forest, etc.), particularly in studies of the global carbon cycle (e.g. McGuire, A. D. et al. An assessment of the carbon balance of Arctic tundra: comparisons among observations, process models, and atmospheric inversions. Biogeosciences 9, 3185–3204 (2012)). The ocean is vast and diverse, and sampling is difficult and expensive, and thus data are sparse and heterogeneously distributed. Yet researchers have a clear need to use these data to their best ability to identify and track change in ocean ecosystems and biogeochemistry. We have clarified these issues in this section and more clearly defined biogeography for the reader.

P109/L9-10: Why do you focus on open ocean only? Provide a rationale for that choice. Define what you mean by open-ocean. Why not study the variability of the Reygondeau et al (2013) or Longhurst (1995) biomes instead?

>Coastal oceans are an important and diverse region of study as they constitute the transition zone between terrestrial and open-ocean regions. These regions receive and transport large amounts of organic matter, making them a very biogeochemically dynamic area. However, different factors affect biogeochemical processes in the coastal oceans, including runoff rates, composition, and tides. Therefore, coastal oceans deserve their own focus and these areas cannot be grouped with open-ocean regions, a conclusion supported by Longhurst’s provinces. Independent groups that focus solely on coastal issues exist for various regions of the world. The North American effort is led by the Coastal Carbon Synthesis (CCARS) group (http://coastalcarbon.pbworks.com/w/ page/15143273/FrontPage) and a European effort is included in the new Blue Carbon Initiative (http://thebluecarboninitiative.org/). Some studies use bathymetric conditions to eliminate coastal data while others use salinity cutoffs. With our biomes, we did not specify any specific constraints for excluding coastal regions however the biomes are defined using criteria based on open-ocean values so coastal regions which are often much higher in SST (due to shallower

depths) and Chl (due to runoff or upwelling) are excluded from the surrounding biomes. Additionally, coastal areas need to be studied with a finer spatial resolution than used here (1°x1°). The biomes in this study have been defined from the start based on a small number of datasets (SST, CHL, MLD, and ice fraction) and with latitudinal restrictions in a few cases. The work of Reygondeau et al. (2013) takes the 56 provinces that Longhurst (1995, 2007) defined based on an array of climatological satellite chlorophyll and hydrographic data, and fits more recent data within these defined provinces in the interest of studying seasonal and interannual variability. Our effort is a different, and complementary approach to their effort. We rely on a smaller set of input data and a simpler approach for consideration of interannual variability. Our approach, being simpler, should also be easier for researchers using global biogeochemical models to adopt for their studies of large-scale biogeochemical functioning. It is important to note that, despite their availability for more than two decades, the 56 provinces of Longhurst (1995, 2007) have frequently not been adopted for use in large-scale biogeochemistry and carbon cycle studies and, instead, the use of regions defined solely on latitude and longitude has become standard (Gurney et al. 2008, Gruber et al. 2009, Schuster et al. 2013, Ishii et al. 2014). Thus, we have been motivated in this work by the apparent need of the global open-ocean community to have a concise set of biomes that have been clearly and simply defined based on relevant biogeochemical and physical datasets. Finally, we also note that we have added and discussed a figure comparing our 17 biomes to the 56 biomes of Longhurst as well as the regions used in the RECCAP project, so that the similarities and differences are evident (Figure 6).

P109/L21-23: Explain explicitly why you think these biomes would be better to use than other existing definitions. Also, if the biogeochemistry of a model were to differ from the observations, why would one want to analyze fluxes or other parameters using these biomes instead of biomes defined specifically for the particular model? In addition, why use biomes at all and why not try to compare the models with the actual measurements where measurements exist? Why interpolate sparse measurements into biomes at all?

>In P109, Line 20 we do explicitly say why we think these biomes are better to use than existing definitions: "We argue that these biomes are preferable because they are defined by relevant environmental parameters instead of by lines of latitude." The Longhurst provinces do have a significant component of their definition that is based on broad latitude/longitude strokes. Also, as we state in P109 L10, we are interested in providing biomes that will address differences at the scale of ocean gyres and in those inter-gyre regions. Longhurst's, and subsequently Reygondeau's provinces split the gyres into multiple provinces, which may limit the ability to aggregate observations or fluxes at the large-scale. We emphasize, as noted above, that there has been a lack of adoption of the Longhurst provinces for biogeochemical and carbon studies that aggregate at the scale of gyres. Models are not perfect at replicating the ocean state. For example, while they will incorporate an El Nino event, it may not happen at the exact same time that the El Niño actually occurred in the oceans. Additionally, limitations exist in the physics of ocean models, especially in the Southern Ocean, and this impacts the biogeochemistry. Models also may be output at coarser or finer resolution than the biomes defined here. All of these reasons are rational for why a modeler would want to create model-specific biomes using the criteria thresholds listed in Table 1. Model output certainly can be compared to measurements directly, but it can also be of interest to study mechanisms and patterns of change at the larger scale (Schuster et al. 2013). Further, global climate model output such as that available in the CMIP5 archive, offer output only at monthly timescales, which complicates direct point-by-point comparison and can lead to a desire to bin data to larger scales for comparison. These biomes offer one possible approach for that larger-scale binning. When studying changing biogeochemistry on the scales of ocean gyres, we unfortunately do not have observations uniformly in space and time. For example, surface ocean pCO2 which ourselves and others have put substantial effort into interpolating the available measurements to the larger scale (e.g. Landschützer et al. 2013). Our previous studies (Fay and McKinley 2013, McKinley et al. 2011) have shown that in most ocean regions and for most time scales, the available data does offer reasonable information about behavior of these

larger regions. Sarmiento et al. (2004) made a similar effort for study of surface ocean productivity. With these encouraging results already in the literature, our goal is to support further use of large-scale biomes.

P110/L5: Biomes are described before being defined (section 2.1). Why are these biomes the ones that matter?

>Thank you for your feedback on the organization of the biome presentation in the manuscript. We have added a sentence to the beginning of Section 2 to define the biomes before going into their descriptions.

P110/L10-11: Why use 50%? Why not 20% or any other value? That seems very arbitrary.

>Please see response to P111/L16/17 below.

P110/L12-23: If wind stress curl is such an important factor, as described here, why is it not used in the definition of the biomes?

>We have chosen not to include wind stress curl in our biome definition as it is poorly observed globally, and thus would have to be taken from reanalyses of numerical models for weather prediction (e.g. NCAR Reanalysis, ECMWF). The inclusion of these uncertainties in the biomes' definitions is not warranted given the other observed oceanographic variables that are available and that do allow us to define biomes quite reasonably.

P110/L17: Surface temperature is largely a function of latitude. I am not sure I understand your statement that the Pacific is warmer? Be more specific.

>You are correct- surface temperature is generally a function of latitude. The text and biome criteria have been clarified to remove the "warmer" statement. Now the only difference between the Atlantic and Pacific criteria is the chlorophyll levels allowed, due to the HNLC attributes of the Pacific versus the Atlantic.

P110/L24-p111/L6: Why fix the equatorial domain with latitude/longitude lines. I thought the point of that paper (as stated in the abstract/introduction) is to allow for flexible boundaries? Given the role of El Nino on inter-annual variability, I would imagine that equatorial variability should be an important component of this analysis.

>We have incorporated biogeochemically defined equatorial regions and allow them to shift from year to year. Thank you for this suggestion.

P111/L7-9: Why exclude these regions? Variability in these regions that are ignored is large, so if the purpose is to define how variable the biomes are, ignoring these seems very limiting. Why is the ICE biome not defined to be coastal and thus ignored?

>The biome definitions are designed to incorporate as much of the global open-ocean regions as possible. The regions that do not fall within the biome criteria are anomalous regions in the gyres or on their borders. We do consider how the biomes borders shift with time for years 1998 through 2010, but we do not address variability within the biome in this manuscript. The ICE biome is defined using ice fraction criteria, and the ICE biomes do not extend all the way to the coast in most regions because grid-cells containing land are excluded. In other regions, it is the typically warmer SSTs and higher CHL that causes coastal regions to be undefined. With respect to the ICE biome being all coastal as postulated by the reviewer, these biomes directly illustrate that there are large regions of the open ocean that have significant ice cover for at least 50% of the year.

P111/L13-15: Please provide a discussion of the errors associated with this data product.

>We have added discussion of the uncertainty in the input datasets in the text in sections 2.2-2.4. We have selected these products due to their strengths in reconstructing even high-latitude regions based on their thorough comparisons to available in-situ observations. It is reasonable to interpret the difference between the core biomes and the mean biomes as our best measure of the uncertainty in the mean biomes due to

interannual variability. The areal differences range from 10-30% in most biomes, to a maximum of 57% in the west equatorial Pacific and equatorial Atlantic biomes (Figure 4). In the text, we discuss that the core biomes would the most conservative choice for biomes because they avoid boundary regions of substantial interannual variability in biome definition. Though we acknowledge that this is not a complete measure of uncertainty, we do also note that it is an improvement over previous biome schemes in which no effort to quantify uncertainties were made (Longhurst 1995; 2007; Reygondeau et al., 2013; Sarmiento et al., 2004).

P111/L16/17: Why use that 0.5 threshold? What is the effect of that choice on your final product? Why not follow Sarmiento et al. (2004)?

>We use a more conservative criterion than Sarmiento et al. (2004). We elected to make that threshold 50% ice coverage as it encompasses regions with a majority of their area being ice covered for the year. These areas are therefore more ice-influenced. If we were to choose a smaller ice percentage threshold, our ICE biome would have extended only slightly further into the subpolar region. Below is a figure showing the maximum ice fraction for the year 2009 (other years are very similar). Note the small fraction of the colored region that reports a max ice fraction of less than 1. This map illustrates that the choice of 0.5 ice fractions will not have a significant influence on the extent of the ICE biome. We have added note of this limited impact in the text in section 2.2.

P112/L3: What is SMIGEN?

>SMIGEN is a program provided by NASA scientists that reads in a Level-3 space-binned or time-binned file and generates a Standard Mapped Image (SMI) HDF product containing one of 5 possible statistical measures. Rather than downloading the provided 9kmx9km mapped images and reprocessing ourselves, we opted to use SMIGEN to ensure that our reprocessing to a coarser resolution was consistent with the NASA SeaWiFS protocol. Thank you for this question. We have updated the text to

clarify this point (Section 2.3).

P112/L7: I understand there are data limitations in winter, but variability in the seasonal cycle is a very important factor governing mean inter-annual variability. This choice of using mean spring/summer values must be justified and the effect of this assumption on the final product must be evaluated.

>We do not deny the importance of seasonal variability of chlorophyll. As stated in the manuscript on P112, L6-8, we use mean spring/summer chlorophyll in order to avoid biases that would be caused by cloud cover during the winter months. We also considered using peak chlorophyll values or the amplitude of annual chlorophyll change. The summer/spring months have the most consistent coverage (over 75% of the northern (southern) hemisphere have at least 5 (3) of the spring/summer months with coverage) and express the most about the strength of the biological pump during each year. The selected chlorophyll criteria thresholds reflect this choice by through higher values than annual mean chlorophyll criteria since chlorophyll peaks during spring/summer. This justification is now noted in the text.

P112/L12: Artificially filling in 2008 values with the climatological mean and then interpreting inter-annual variability for that year is not a very robust approach. If you do that, provide an estimate of how this choice affects the final product.

>We admit that filling in this year with the climatological mean is not optimal, however there are no other viable options available given the lack of data. Two further points: (1) More than just chlorophyll goes into defining the biomes, and (2) Chlorophyll is not used as a criteria for all of the biomes, so this exception doesn't affect all biomes in 2008. As stated in the manuscript, this only affects the southern hemisphere 2008 values and chlorophyll is only used to define the STSS and STPS biomes. From Table 1 we see that the criteria for STSS either includes chlorophyll OR maxMLD conditions to be met for this biome classification. This optional criteria reduces the impact of the missing 2008 chlorophyll coverage. Comparison of the biome maps for 2007, 2008,

and 2009, shows many fluctuations between years as expected. We alert the reader to this "condition" of 2008 so that they may proceed carefully if conducting analysis for that year specifically.

P112/L13-14: Does this have an influence on the final boundaries set for STPS?

>This question refers to the use of the southern hemisphere summer chlorophyll (Dec-March) in defining the Indian STPS biome despite its extension into the northern hemisphere. Chlorophyll in these oligotrophic regions does not vary much annually and so this choice does not have a significant influence on the boundaries. However, we have elected to change this and to allow the northern portion (that north of 10°S) to utilize north hemisphere summer chlorophyll definitions and the southern portion of the biome to use southern hemisphere chlorophyll to address this concern. There is no discernable change in the extent of the STPS biome with this change. We have edited the manuscript and maps to reflect this change and clarification. Thank you for this question.

P112/L19: MLD does not vary annually. >Thank you for this suggestion.

P113/L5-8: If filling missing MLD grid cell has no influence on the biome assignment, why discuss it? It would be more helpful to provide a rationale why you are using threshold values or 125 and 150 m.

>We include this discussion of MLD, as it is a common concern among scientists that MLD climatologies have large gaps. We admit these gaps and share our filling procedure, but also note the relatively small impact of the choice. If others chose to use a MLD climatology with missing values in the extratropics or subpolar regions, they would find more impact on the biome classifications. We have elected to retain this clarification in the text, but thank you for your comment. The choice of 125m and 150m depth was based off of previous work, especially that by Sarmiento et al. (2004) who chose 150m as their maxMLD cutoff. This depth is commonly deep enough to access the nutrient rich waters of the thermocline. We selected slightly shallower criteria for the

northern hemisphere maxMLD criteria in order to capture the high chlorophyll areas of the STSS biomes.

P113/L10-11: I don't understand the motivation for this smoothing step. Why not defining the core-biomes without this? This smoothing appears to artificially limit variability, yet the purpose of this paper seems to provide an estimate of variability. This seems counterproductive.

>This smoothing step only fills in small (1 gridcell) gaps in the biomes if it is surrounded on 3 sides by another biome classification. As noted above by the reviewer, there is uncertainty in the input datasets and this smoothing is also motivated in part by these uncertainties at the 1°x1° scale. As explained in the manuscript, this does not change a large number of gridcells in the maps and mostly acts to fill in missing gridcells in the southern hemisphere.

P113/L14: What is the rationale for this choice of 300-400 cell?

>This was not a choice, but rather a result of the iterative smoothing procedure. This is the number of gridcells that changed in the automated, iterative smoothing process as described in the text. We quote this number so as to verify that the smoothing process did not make significant changes to the biome maps nor did it significantly influence one year more than another. With our updated biomes the smoothing process now changes between 350-467 gridcells per year with the increase due to the time-varying equatorial biome and other updates.

P113/L17: More than the relative number of cell that changes in each biome, it is important to comment on the location of these cells that change assignment. Do these cells form coherent clusters?

>Thank you for this comment. The location of the "smoothed" gridcells is indeed important and we have included a comment on this in the manuscript (P113 L19-21): "These smoothed areas occur predominantly in the intergyre regions, the hardest region to de-

fine due to strong interannual variability." Most of the changed gridcells are located in the southern hemisphere STPS biomes. Random gridcells with deep maxMLD climatology are excluded from the STPS classification. However, these gridcells can have low chlorophyll, making them fit the attributes of the subtropical gyres. The smoothing process reassigns these gridcells. Additionally, gridcells in the northern hemisphere STSS biomes are smoothed if they are surrounded on all sides by STPS biome gridcells (and vice versa). As stated in the manuscript, it is not a large number of gridcells that are changed, nor is it a large fraction of gridcells in any one biome that are altered in this smoothing process.

P113/L18-19: the previous sentence says 15%, not 14%? If more that 300-400 cells were allowed to change, would more cells change?

>The 15% is referring to the maximum percentage of gridcells that change in the time-varying biomes (the 13 individual year maps). The 14% is referring to the percentage of gridcells that change in the MEAN biome map. As mentioned above, we do not limit the number of gridcells that change to 300-400. We do iterative smoothing until no more gridcells meet the criteria to be changed (no more rogue/isolated gridcells exist on the maps) and this number never exceeds 400 (now 467). This takes anywhere from 8-12 rounds of smoothing to accomplish depending on the map (for the mean and each of the individual years). With the update to the biomes, including varying equatorial biomes, these numbers have increased. Now the number of gridcells that change in the time-varying biome maps ranges from 350-467.

P114/L23: Show the maps for each year in the paper directly ... this is after all the point of the paper.

>Thank you for this comment. We will include the maps of the time-varying biomes in the manuscript (Figure 2) as well as in the downloadable files.

P115/L5-P116/L7: Provide a deeper analysis of why grid cells change assignment? What variable controls this change? Is variability in the defining parameter correlated

in these cells? Is variability in the defining parameters (e.g. SST and Chl a) correlated anywhere? If there is correlation, why use fixed biome definitions? The definitions should take into account covariance between defining parameters. To what extent are these annual mean changes governed by changes in seasonality?

>The question refers to the core biomes and how/why biomes in each year shift. The variables that control these changes are chlorophyll concentration and sea surface temperature. Seasonality is not considered in these biomes. Other efforts (Reygondeau et al. 2013; Landschützer et al. 2013) have looked at how provinces change with season. In this manuscript we focus on interannual changes of the biomes' extent. We have not looked at correlations or covariance between parameters and the impact on interannual variability. As mentioned above, the goal of this paper is to describe the biome definitions and present the data product. There is still much analysis that could be done with these biomes and it is a goal of our future work to look into such relationships.

P116/L22: "there are NO significant trends"

>Thank you for this correction.

P116/L24: is that trend significant?

>The stated trend (-2.31 $\pm$ 1.07 105 km/yr) shows that it is statistically significant given that with the stated uncertainty it is still a negative trend. Other biomes do report a small trend but none are statistically significant given the 95% confidence interval.

P117/... : What is significance of these trends? What does the "error" reported in each case represent?

>The trends represented here are linear trends showing the change in area of a biome over the 13 years included in our analysis (y = a + bx). The errors are a 95% confidence interval on the trend (b) as computed using statistical techniques (Wilkes, 2006): CI = +/-t * RMSE * sqrt(1/(sum((Xi – mean(X))ˆ2)))

where t is the two-tailed t-statistic for 95% confidence for N-2 degrees of freedom.

These details are now referenced in the text.

P117/L11: These represent just 2 years shift, hardly a stringent test. Also, none of the mean area presented come with an estimate of error such that the errors presented here are greatly underestimating the true errors.

>We concur with the reviewer that these trends of the North Atlantic ICE biome shrinking as the SPSS biome expands are not stringently tested here. Our goal is to describe the results as presented in this paper, and future studies could delve into this issue more deeply. In our previous work, we have found significant sensitivity of short-term trends to the selection of start and end years, so we do want to point out that this particular trend is not an artifact of this selection.

P117/L17-19: Is the variability reported in this analysis important to a degree that would lead us to doubt previous conclusions provided by previous papers by Fay and McKinley? If so, discuss this in detail.

>The shifting of the biomes has been recognized as an undeniable fact but previous studies have not allowed dynamic biome boundaries for their trend analysis. As discussed in this manuscript, many previous studies have not taken into account boundaries defined by anything other than latitude and longitude (Gruber et al., 2009; Schuster et al., 2013; Lenton et al., 2013). This is beginning to change and groups are now considering how biomes are shifting with time (Landschützer et al., 2013). We look forward to completing such analysis as an update to our previous work however we do not foresee the results using these dynamic biomes contradicting previous results from Fay and McKinley (2013). However, limitations of years available for dynamic biomes (1998-2010) will remain a challenge to utilizing time-varying biomes, and thus we also offer the core biomes as a conservative choice for studies extending beyond this timeframe.

P117/L23: Fay and McKinley.... provide explicit references.

>We have amended the reference. Thank you for this comment.

P117/L24: are the same MLD criterion used in all the studies?

>A similar maxMLD criteria is used in previous studies (150m) however we have changed the maxMLD limit to 125m for the north hemisphere biomes in this manuscript thanks to improvements in the MLD climatology used. By shoaling the requirement (125m) we are now able to capture most of the intergyre regions in the north Atlantic and Pacific.

P117/L24: What aspect of MLD is improved? Also, variability in MLD is not considered. Is it possible that variability in MLD would counteract variability in the other variables such that this would bias the biome assignment of some grid cells? Would that have an effect on the trends discussed in Figures 3 and 4 and in section 4?

>The improved maxMLD climatology (as discussed in Section 2.4) uses an algorithm to calculate maxMLD rather than only a density threshold as in previous MLD climatologies. Unfortunately, MLD is not a variable that is measured often enough and with sufficient spatial coverage to define interannual variations. Improvements provided by the ARGO program will hopefully enable this in the future. The maxMLD criteria is used to designate the STSS and STPS biomes as the STSS regions have higher summer chlorophyll concentrations caused by deeper mixing that supplies nutrients from below the thermocline. It is for this nutrient supply process that the maxMLD acts as a proxy. It would be ideal for the maxMLD to vary with time, but the inclusions of the chlorophyll criteria does partially compensate for its lack. The maxMLD and chlorophyll criteria should not counteract each other in terms of interannual variability as one would expect deeper maxMLD and higher chlorophyll to go together.

P118/L9: I don't think that statement is true: "...full coverage of the open ocean...". Looking at the maps you provide, there are clearly grid cells that many would consider "open ocean" that are not covered. "open ocean" is not defined.

>Thank you for this comment. We indeed do not get full coverage of the ocean. There are areas near coastal bays or in shallow plateau regions (Falkland Plateau, Arabian Basin, Bay of Bengal) that are not captured by our biomes. The physics and biogeochemical processes in these regions are difficult to study and interpret and are very different from open ocean processes as circulation pattern differ and interactions with the shallow plateau regions can have strong impacts on nutrient supply. These differences make them important regions to study, but are not considered open-ocean as their anomalous attributes exclude them from any open ocean biome. We purposely have strict criteria for the biomes so that they can be used as homogenous regions where data can be assumed to represent the entire region with similar circulation patterns and biological Thank you also for noting the need to clarify our definition of "open-ocean". In our case, we define open-ocean as the regions of the major ocean basins that do not include coastal regions or smaller bodies of saltwater including bays, seas, gulfs, and straights.

P118/L19-21: "If as widely...". Rewrite sentence.

>Thank you for this comment.

Table/Figures:

Table 1: No rationale is provided for the choice of variable and thresholds listed in Table 1.

>The choice of variables included in our biome creation is explained in the text (Section 2). The threshold selection is based on values put forth by previous studies (Sarmiento et al. 2004, Longhurst 2007) but final selections are tuned to the values presented here as they provide the most complete coverage of the open-ocean. All selections are based on biogeochemical and physical ocean properties for the biomes.

Figures 3 and 4: Color-coordinate the lines in figures 3 and 4 with the colors of the regions on the maps.

>Thank you for this suggestion.

Interactive comment on Earth Syst. Sci. Data Discuss., 7, 107, 2014.