Response to Reviewer #1

The authors have made substantial modifications to the modeling configuration and substantial revisions that have significantly improved the manuscript and addressed my previous comments. I have included additional minor comments on specific sections of the text below. Overall, this manuscript makes a valuable contribution to the community's understanding of bacteria dynamics and to the modeling of the WAP ecosystem and I look forward to seeing the final manuscript published after these final issues are addressed.

Thank you for your positive feedback. We have revised the manuscript to address each of your specific comments below. We also should note that the related paper on the modelling and data assimilation framework, which we shared with the previous round of revisions, has now been accepted in *Geoscientific Model Development*:

> Kim, H. H., Luo, Y.-W., Ducklow, H. W., Schofield, O. M., Steinberg, D. K., and Doney, S. C.: WAP-1D-VAR v1.0: Development and Evaluation of a One-Dimensional Variational Data Assimilation Model for the Marine Ecosystem Along the West Antarctic Peninsula, Geosci. Model Dev., https://doi.org/10.5194/gmd-2020-375, Accepted, 2021

Minor comments / technical corrections

Line 252 "with relatively high correlations" and line 255 "with lower correlations": Specify if the correlations are significant.

Included (line 294).

Lines 261, 264, 371, caption of Table 3: The references to the tables in the text and caption appear incorrect.

All fixed.

Consider changing "Tables 1-2" to "Tables 2-3" in line 261, and "Table 2" to "Table 3" in lines 264 and 371.

All fixed.

In the caption of Table 3, consider changing "Table 1" in "Table 2".

Fixed.

Line 264: Uniform the portability index between Table 3 and the text for 2011-2012.

Fixed.

Lines 266-267: Change "Figure S6" to "Figure S8" Figure S8: The time evolution for February and March is missing. The time series should cover the whole growth season, i.e. until March 31

as in Figure 4. The multiplication factor is hidden under the title for 2 sub-figures and parentheses are missing in titles.

We updated Figure S8 and the figure caption.

Line 333: Change “HHA” to “HNA”

Fixed.

Line 373: Change “wth similar” to “with similar”

Fixed.

Lines 383-384: “Assimilating each bacterial group’s biomass allows for the partitioning of [...] that were never measured in this study.” The sentence is unclear. Do you mean “in this study area”?

Yes, and we fixed the sentence to “… never measured for each bacterial group in this study area” because the bulk BP has been measured but not the group-specific BP (line 424).

Lines 399-400: “The WAP typically exhibits strong interannual variability (Ducklow et al. 2007)” Could you be more specific? Do you mean the WAP meteorological or hydrodynamic conditions, or ecosystem?

All of them; so we have listed those now (line 448-449).

Line 476: Change “export” to “sinking flux”

Fixed.

Figure 5: Add arrows for flows 8 and 9 “nutrient uptake’ and “regeneration” or remove arrows for the other nutrient flows on the 4 sub-figures.

We have removed the arrows for the nutrient flows to only present C stocks and flows and updated the figure captions accordingly for both Figure 5 and Figure S9.

Response to Reviewer #2

For the revised ms the authors expanded their model from 0D to 1D (3 layers). The ms has greatly improved, mostly by providing the model equations (although apparently incomplete, see below), whose omission had made it impossible for me to understand the model structure in the previous round. The parameter estimation has much improved and its description has become OK now. But even after the long time it has taken the authors to prepare the revised ms, it still leaves a strong impression of sloppiness. Several sentences are simply incomprehensible and little attention seems to have been paid to the readability, correctness, and design of some of the figures. Only some of the changes to the previous ms are highlighted. The model description in the main text is still very much unclear and this applies also to the mode concept. Nevertheless, having seen the equations, the study seems to be much better than I had feared based on the original ms. After another major revision or two, I now think it could become a useful contribution.

Thank you for your time thoroughly reviewing our manuscript again. Please see below our response specific to each of your comment. We also should note that the related paper on the modelling and data assimilation framework, which we shared with the previous round of revisions, has now been accepted in *Geoscientific Model Development*:

> Kim, H. H., Luo, Y.-W., Ducklow, H. W., Schofield, O. M., Steinberg, D. K., and Doney, S. C.: WAP-1D-VAR v1.0: Development and Evaluation of a One-Dimensional Variational Data Assimilation Model for the Marine Ecosystem Along the West Antarctic Peninsula, Geosci. Model Dev., https://doi.org/10.5194/gmd-2020-375, Accepted, 2021

One of the remaining problems is the confusion of assumptions and results. The authors mention in the response letter (introductory para, point 4) that the "larger cell-specific BP and SDOC uptake rates of HNA cells than those of LNA cells" indicate the robustness of their analysis. This finding is also referred to in the results and discussion sections (lines 276, 384–385). But this is a model assumption, not a result: "maximum bacterial growth rate of the HNA group (μHNA, d-1) was ensured to be optimized to be higher than that of the LNA group (μLNA, d-1)" (lines 168–169).

While it is true that the maximum HNA growth rate was kept higher than the maximum LNA growth rate over the course of simulations, there are additional model stock and flow variables that determine bacterial SDOC uptake, respiration, and production. One of them is the bacterial stock (C_{BAC}, here $_{BAC}$ can be applied to HNA and LNA both; Eq. A.4.53) whose time rate of change is determined prognostically in the model by multiple source (total DOC uptake) and sink terms (DOC excretion, respiration, grazing, and viral mortality). Bacterial SDOC uptake is proportional to the bacterial stock (Eq. A.4.13), and bacterial respiration is affected by the bacterial stock and total DOC uptake both (Eq. A.4.25). By definition, cell-specific BP is the biomass-normalized difference between total DOC uptake and respiration (i.e., cell-specific BP = $(G^{C}_{BAC,DOC} - R^{C}_{BAC})/ C_{BAC}$, Eq. A.4.14, A.4.25, A.4.53). Because of these intertwined processes whose magnitudes are continuously adjusted and determined during optimization, keeping the high

maximum HNA growth rate does not simply guarantee higher SDOC uptake and cell-specific BP of HNA than those of LNA. One simple counterexample case would be: despite higher maximum growth rate of HNA and that of LNA ($\mu_{HNA} > \mu_{LNA}$) assigned initially and kept during optimization, if LNA biomass is larger than HNA biomass ($C_{HNA} < C_{LNA}$, due to the relative magnitude of sink and source terms for each bacterial group, Eq. A.4.53), this could potentially lead to larger SDOC uptake rate of LNA than that of HNA (Eq. A.4.13). Thus, these findings are the result of optimization in conjunction with the model assumption about class growth rates. We have added this explanation in Section 4.2 (line 428-435).

The authors have now clarified that their model considers flexible (Chl:)C:N:P stoichiometry, but this is mentioned only in the equations and the supplement. This information must be provided in the main text, e.g., under Sections 2.1 or 2.2 or a new 2.x section, as this information is quite crucial for understanding the model design. The statement that the model has 12 state variables (line 81) is simply wrong (I counted 32). This misinformation had led me to conclude that the model was based on a fixed-stoichiometry approach in my previous review. Fig. 1 has been amended regarding the flows of inorganic nutrients to phytoplankton. But it still remains a source of confusion. Fig. 1 shows two compartments, "Higher level" and "RDOM", which do not have corresponding differential equations, so the authors should either add the missing equations or modify Fig. 1 to clarify what these are (this applies also to Fig. 5).

We have expanded on Section 2.1 to detail the model's flexible stoichiometry and referred the rest information to Text S1 (line 91-102).

By 12 prognostic model state variables, we are referring to the subset of carbon stocks of biological compartments and dissolved inorganic nutrient compartments analyzed and presented in the study. These include diatoms (Eq. A.2.41), cryptophytes (Eq. A.3.37), HNA bacteria (Eq. A.4.53), LNA bacteria (Eq. A.4.53), microzooplankton (Eq. A.5.24), krill (Eq. A.6.27), detritus (Eq. A.7.4), LDOC (Eq. A.8.4), SDOC (Eq. A.8.7), NH_4 (Eq. A.9.2), NO_3 (Eq. A.9.3), and PO_4 (Eq. A.9.4). We have made this point clear in Section 2.1 (line 81-85).

In the previous version we mentioned that both higher levels and RDOM are implicitly represented as model closure terms in Section 2.1. Higher levels and RDOM play a role as source or sink terms of other explicit model state variables (i.e., krill removal by higher level, detrital production by higher level, SDOM production by higher level in Eq. A.6.24, bacterial RDOM excretion in Eq. A.4.26, krill RDOM excretion in Eq. A.6.21, SDOM to RDOM conversion in Eq. A.8.2) but the model does not calculate time derivatives of their concentrations. As suggested, we have modified both Figure 1 and Figure 5 captions to clarify that these two compartments are implicit as well as included the explanation above in Section 2.1 (line 85-90).

The sentence "Total (bulk) bacterial production (BP; BP = BPHNA + BPLNA) was constrained by observations, and therefore, the group-specific production (BPHNA and BPLNA, mmol C m-3 d-1) was determined during optimization:" (lines 99–100) is unclear. Does this mean that Eqs.

(3) and (4) apply only during the optimization? How do you calculate BP_HNA and BP_LNA when not optimizing?

Our apologies for the confusion. To answer your question first, Eq. (3) and (4) apply both before and during optimization. Before optimization the assigned initial parameter values are used to calculate LDOC uptake ($G^C_{HNA,LDOC}$, Eq. A.4.12), SDOC uptake ($G^C_{HNA,SDOC}$, Eq. A.4.13), and respiration (R^C_{HNA}, Eq. A.4.25) and the resulting BP_{HNA}. These parameter values are adjusted during the optimization resulting in new updated BP_{HNA} values (based on optimized parameters) presented throughout our manuscript (the same applies to LNA). We have added this explanation in Section 2.1 (line 118-123).

On line 104, you state that "The modelling framework consisted of a dynamic (mechanistic) part and a data-driven part (Figure 2)" but Fig. 2 is about the data assimilation scheme and does not show or mention dynamic and data-driven parts.

We have modified Figure 2 to show the mechanistic and the data-driven parts. Also, by its design, the data assimilation methodology is a fusion of model dynamics and data constraints.

On lines 109–110, you introduce fmodes as functional modes, but even after reading the whole ms several times, it remains unclear what these are, e.g., which functions the fmodes describe. Since the fmodes are used later on in the statistical analysis, they should be explained clearly.

The fmode constructs are described on line 133-134: "functional modes (*f*modes hereafter) were derived from predicted community metabolic structure." For further details we refer the reader to Bowman et al. (2017). In brief the functional modes are derived in exactly the same way as the taxonomic modes, except that the SOM is trained on the abundance of predicted metabolic pathways rather than taxa.

On lines 310–311, "These results suggest a clear link between the modelled ecosystem functions and observed bacterial taxonomic (modes) and physiological (fHNA) traits observations." This, together with the absence of any significant relations for fmodes, seems to indicate that the functions (not decribed in the ms) of the functional modes were chosen inappropriately.

We appreciate the comment but disagree with the interpretation. In general, we find that functional modes are poorer predictors of ecological processes than taxonomic modes. This falls in part from the very different distribution of the underlying data in the marine environment. Taxa are much more sensitive to ecological processes (as drivers and responders) whereas many metabolic pathways – which may be widely distributed across taxa – are not.

The authors moved from a 0D to a 1D setup for the model but do not provide any information about the 1D setup except the depth levels. No indication about vertical mixing is given in the equations either, so they must be considered incomplete. Also, I am not convinced that, given the shallow model domain (20 m), a 1D design provides a significant advantage over 0D. But again, essential information is missing to allow a firm judgement, e.g., the depth of the mixed layer and

its seasonal variations. If the mixed layer is usually deeper than 20 m at the modelled site, then a 1D model offers no advantage over a 0D model. Also, no information is provided regarding the vertical geometry (are the three layers of the same height?) or the mixing scheme (implicit, explicit, positive definite, etc.). The authors should also indicate how the mixing coefficients were obtained or calculated. The reference to Kim et al. (2021) is insufficient, as this has not been published. The authors should just add a short section describing the vertical configuration and modify the equations accordingly.

Please note that in the previous version we detailed physical forcings of the model that takes the vertical structure and configuration into account (Figure S1, Text S2). The model implements the mixing scheme where vertical advection and detrital sinking are demonstrated with a third-order direct space-time upwind-biased scheme (Hundsorfer & Trompert 1994) and the Sweby flux limiter (Sweby 1984) but simplified to work for 1-D vertical advection only. Vertical diffusion is applied using a Crank-Nicholson vertically variable diffusion operation (Press et al. 1986), with a closed upper boundary and an open bottom boundary. We have added this information on the mixing scheme in Text S2 (line 304-311). We modeled 3 layers, 0, 10, and 20 m, which has 2, 16, and 4 m layer thickness, respectively, so that the center of each layer corresponded to the depths (surface, 10, and 20 m) from observations as closely as possible (added in Section 2.2 line 150-151). More importantly, the original model article (Kim et al. 2021) has been accepted and will be published soon in *Geoscientific Model Development* with which we hope to navigate potential readers for more details about the model framework. A preprint of the paper is now publicly available from https://gmd.copernicus.org/preprints/gmd-2020-375/

In the previous version we switched to the 1D framework to simulate seasonally-varying MLD and light impacts on the model stocks and flows with more realistic, observation-based boundary conditions at the base of the layer of interest (10 m, full description in Text S3 line 354-360) that were previously fixed to constant concentrations over time in the 0D setup. These constant vertical fluxes of the variables at 10 m are not appropriate for representing a seasonal cycle of the system and especially troublesome when MLD is deeper than 10 m. In the previously added Figure S1, we show that MLD is mostly deeper than 10 m but frequently shallower than 20 m over the seasonal cycle, so the 1D setup the model now correctly introduces the vertical fluxes from the base of the 10 m by mixing. On a minor note, despite the advantage of simulating the full water-column layers, we judged that it would be best to exclude depth levels without bacterial traits observations, yet to include an adequate number of depth levels for seasonal MLD and light impacts, and ultimately chose to model 0, 10, and 20 m. We have added this explanation in Section 2.2 (line 153-155).

The concept of the bacterial modes remains rather confusing. Since this is one of the main foundations of the present study, this must be clarified. My main problem with Fig. 6 is still that the explanation of the modes (also in the authors' response letter) does not seem to match what is shown in the panels. For example, Candidatus Pelagibacter ubique is supposed to dominate mode 6 (Fig. 6c) and C. Thioglobus singularis should dominate mode 1 (Fig. 6e). However, the relative

abundance of C. T. singularis in mode 1 never exceeds 0.25, whereas C. P. ubique has relative abundances between 0.25 and 0.35 in mode 1, according to panel c, so it appears that both modes 1 and 6 are dominated by C. P. ubique. I did go through Bowman et al. (2017) but could not find an explanation there either.

Thank you for noting the discrepancy between the figure and the caption. We have modified the example to *Dokdonia* sp. MED134 as that is a clearer example (line 1178-1180).

The description of the data assimilation and parameter optimisation has become much more accessible by the added explanations. Still, several points remain unclear. On lines 166–167, you write "... group-specific bacterial model parameters were optimized in the direction to properly represent the dynamics associated with each group ..." I do not understand what this means, even with the explanation in the next sentence, which describes a constraint imposed in the maximum bacterial growth rates.

We have significantly elaborated and rearranged elements in Section 2.3 to demonstrate details on the parameter optimization process (line 183-185, 193-209). Please see if the added paragraphs make sense and help you understand the optimization process better.

On lines 187–188, "When converting Chl to phytoplankton C (N) biomass, the maximum Chl to N ratio was used along with other reference ratios ..." it remains unclear why you use the maximum (rather than, e.g., an average) Chl:N ratio and what the other reference ratios are. This must be clarified. Also, throughout the ms, you mostly refer to C biomass, so it is unclear where, when, or why you convert to N biomass here.

We needed the Chl:C ratio to convert Chl to C biomass to calculate the fractional contribution of phytoplankton to the total (observed) POC, which also includes zooplankton and detritus carbon. However, the Chl:C ratio of phytoplankton has not been measured at the study site. The maximum Chl/N ratio (Θ, g Chl *a* (mol N)$^{-1}$, Table S2-6) is the parameter that the model directly requires to calculate phytoplankton C growth and Chl production, the value of which is available from other data assimilation studies (Luo et al. 2010), so we used it and multiplied Θ by the Redfield C/N ratio of 0.15 to obtain phytoplankton C biomass. We acknowledge that by using the maximum Chl/C ratio we applied a minimum estimate of C and N biomass converted from Chl that the model needs to match. We also have updated the percentage of total POC and PON that living biomass account for that was mistakenly written in the previous version (Section 2.4, line 224-227).

On lines 198–199, you refer to "... normalized costs of individual data types (J'm) ..." The J'm seem to be indicated in Table 2 but they are never defined.

Fixed (line 237).

On lines 205–213, you refer to the depth of the mixed layer as affecting the calculation of the target error. Besides the lack of information on the mixed-layer depth, it is also unclear whether the

mixed layer was always deeper than 20 m, or whether you always applied the same CV throughout the whole model domain. Please explain clearly.

Apologies for the insufficient explanation about target errors in the previous version. In Kim et al. (2021) we calculated the climatological mean and standard deviation of each variable in the mixed layer per observation (vertical profile) over an extended year period (2002-03 to 2011-12) to get a more generalized picture with large sample size. We then used the climatological CV (from the same climatological mean and standard deviation) for target errors of the most data types and the same climatological standard deviation for target errors of the log-transformed data types. In other words, each data type was assigned with its own but non-time varying CV.

In the present study though, MLD is mostly deeper than 10 m but frequently shallower than 20 m (Figure S1), and following the methodology as in Kim et al. (2021) would throw out most vertical profiles, decrease the sample size, and make inadequate cases for representing the overall observational errors and seasonal-interannual variations (e.g., for all data types the depth levels measured typically span surface, 10 m, 20 m, etc, so shallow MLD would leave vertical profiles with only one or at most two data points within MLD). As stated in the previous version, we instead calculated the climatological standard deviation, and CV in the upper 20 m (i.e., 0, 10, and 20 m) per profile over the four study years of the present study, which we adjusted to similar values in Kim et al. (2021) as they were, of course, higher than those in Kim et al. (2021), largely due to the inclusion of the observations at 20 m when MLD < 20 m. For adjustment, we derived the ratios of the climatological CV (for most data types) and standard deviation (for the log-converted variables) between our study and Kim et al. (2021), averaged the ratio for the same categorical data types (e.g., nutrients (NO_3 and PO_4), phytoplankton (diatoms, cryptophytes, and primary production), bacteria (HNA and LNA biomass and production)), and multiplied this ratio to what we calculated for the present study to reduce to the level in the "mixed layer" to avoid an overestimated target error of each data type. Though complicated, we chose to do this way of combining target errors in Kim et al. (2021) and error adjustment hoping that it would more realistically represent the dynamics at Palmer Station B and for the 4 study years in the present study because target errors in Kim et al. (2021) were calculated for the 11-year period of the data from slightly offshore Palmer Station E. We have created a new Text S4 section (line 361-381) to include this explanation (added also to refer in main text, line 245-249).

On lines 240–241, "... 5-7 constrained parameters and 3-6 optimized parameters ..." Please clearly explain how you define and determine constrained and optimized parameters and how they differ. Also, explain CS and OP in Tables S2–S6.

We have made this part clear in Section 2.3 (line 194-199).

On lines 254–155 "However, the model skill for HNA biomass slightly degraded in the climatological model (Figure 3b), with lower correlations and normalized standard deviation and higher RMSD than the four years together (Figure 3a)." Since the individual years have been optimized individually, this result was to be expected. What could be more informative is a comparison with simulations for the individual years but with the same parameter set, e.g., the

most portable one. This could provide insight into the influence of parameter differences compared to that of different boundary conditions and forcings between the different years.

Thanks for your suggestion. Please note that the rationale for the climatological model skill assessment in Taylor diagrams is because we use the climatological model parameters for the climate change simulations, not the model parameters from any specific year. The portability of the most portable model (PI = 0.76 ± 0.11, 2012-13) is still quite similar with the next portable year's model (PI = 0.76 ± 0.11 for 2013-14, Table 2) thereby making it hard to choose one specific year's model parameter set over others to represent the overall "mean" ecosystem required for climate change simulations. We have updated Section 3.1 (line 301-304) and highlighted again the parts on model portability relevant to your comment (line 354-357, line 416-420).

These sentences are incomprehensible and must be corrected. I could not figure out what you wanted to say here: Lines 273–275: "C stocks and flows averaged over the growth (Figure 5) and normalized by NPP (normalized by NPP in 1-day for C stocks; Figure S9) season for each year summarized an annual snapshot of the group-specific bacterial dynamics."

Fixed (line 313-315).

Lines 283–284: "NO3, POC, and SDOC in unassimilated years were modelled to values comparable to those in other assimilated years (Figure 5)."

Fixed (line 323-324).

Line 399: "Modelled nutrient stocks were above detect limits and indicated the lack of macronutrient limitations."

Fixed (line 447-448).

Fig. 3: The caption for panel a (2010 - 2013) is confusing. The simulations also cover 2014 and this caption gives the impression that the simulations went through 2010–2013 continuously, which is not what you did. You should come up with a better caption. Also, as mentioned above, a third panel showing results for different years with a single parameter set could be useful. The last sentence of the caption seems to make no sense, since the x-axes of both panels are the same.

Thanks. We have deleted the captions from both (a) and (b) in Figure 3 and instead explained those in the caption. Please see our response above about using the single year model parameter set in Taylor diagrams. We also have made both Taylor diagrams look less busy by labeling each data point number rather than text.

Fig. 4: The numbers and letters are very hard to read and often overlap. Maybe rearrange in 4 columns (2 for the means and 2 for the CVs)? The units in (b) are wrong, the CV is dimensionless.

Please see our updated Figure 4 if this addresses your concerns.

Fig. 5: The caption says that the panels show C stocks and flows in units of mmol C m-2 and mmol

C m-2 d-1 but the panels also show NH4, NO3, and PO4, so the associated numbers must have different units. The caption explains the numbers in the first rows and the numbers in parentheses but not the numbers in the second rows next to the arrows. Then it says that "N and P flows, as well as the flows smaller than 0.01 mmol C m-3 d-1, are omitted." but the panels show arrows from and to NH4 and to the inorganic nutrient compartments.

We have fixed the Figure 5 and Figure S9 captions as suggested.

Fig. 8 suffers from the same problems as Fig. 4 (% numbers are also dimensionless). In addition, the first rows in (b) should be left out as they are always 0 by definition.

We have fixed Figure 8 the same as Figure 4, but have kept the first row in both panels (a) and (b) to explicitly represent the baseline state and for consistency between the two panels. We also have corrected Figures S10-11.