The authors would like to thank the Reviewer for their careful assessment of our manuscript. Please find Reviewer comments reproduced below in blue font, with our responses provided in black font. Changes to the manuscript are provided in red font both below and in the accompanying manuscript.

## Main comments and recommendation

-----

The manuscript by Kvale at al. presents some new extensions to the ocean biogeochemistry in the widely used UVic Earth System Model, of which the most important one is a description of silicon cycling in the ocean. Together with the presentation of the new model, an evaluation of the models present-day biogeochemistry, and a future model scenario run that extends to the year 2300 are shown.

The future run is interesting, as it predicts a strong shift in phytoplankton species dominance after the year 2100, especially in the Southern Ocean, from diatoms to CaCO\$\_3\$-producing phytoplankton, which would have strong implications both for marine biology and biogeochemistry. I am, however, sceptical about the robustness of this result, given that the model description of the competitive advantages of the different phytoplankton functional groups in this model (as in most models) is overly simplistic and has been strongly tuned to agree with our (also still limited) knowledge about their present-day distribution. I think it is fair to say that at the present state of modelling, one can put some confidence in the modelled biogeochemistry, but not in the ecosystem functioning. The abstract of the manuscript almost entirely focusses on this, in my view still at least questionable, part of the results.

The authors agree with the assessment of the Reviewer that the future run is interesting and has implications for future biogeochemistry and biology, but that readers should be sceptical of the result, which is produced by a single parameter set of a single model. This is one reason why the last sentence of the abstract states that "These results are meant to serve as a baseline for sensitivity assessments to be undertaken with this model in the future"- it is a disclaimer against taking any results from this early version of the model too seriously.

We have re-arranged the abstract to place more emphasis on the historical simulations and model description itself, and to reiterate the RCP 8.5 projection is meant as a demonstration of the model function.

L 4:

This new model combines previously published parameterisations of a diatom functional type, opal production and export with a novel, temperature-dependent dissolution scheme. Model biogeochemical rates, carbon and nutrient distributions are similar to those found in previous model versions. We assess the fully-coupled model against modern ocean observations and the historical record since 1960...

L 10:

The model simulates a global decline in net primary production (NPP) of 1.8% having occurred since the 1960s, with the strongest declines in the tropics, northern mid-latitudes, and Southern Ocean. Based on a single parameter set tuned to observations, we also

# perform a first projection of potential biogeochemical and ecological changes under a business-as-usual atmospheric CO2 forcing to the year 2300.

I think this neither fits to the overall scope of the manuscript, nor to the journal this is submitted to. If this part of the manuscript has indeed such importance, its uncertainties should have been discussed much more in the main text of the manuscript, and the manuscript should probably been submitted to a different type of journal.

We appreciate the ecological trends in the transient simulations may be considered to fall in a grey zone with respect to the journal's scope and might be interpreted by readers as "science" (in which case, a thorough discussion of model uncertainties is absolutely necessary and a different journal would be more appropriate). However, biogeochemical models are commonly used to make projections of future ecosystem functioning, so reporting first results from transient experiments is useful for possible applications of the model, or other models, in the future. The Copernicus website lists at point 5 of the scope of Model description papers, "Examples of model output should be provided, with evaluation against standard benchmarks, observations, and/or other model output included as appropriate". We consider these transient simulations, and their ecological responses, to be baseline examples of model output using a standardised forcing scenario (extended RCP 8.5) that are comparable to other model output (UVic ESCM ecosystem simulations are typically run to year 2300). Section 4.2 includes discussion of how the model's transient responses compare with other, previously-published "science" runs (Schmittner et al., 2008, Kvale et al., 2015a, Nickelsen et al, 2015, Moore et al., 2018, Fu et al., 2016). Encouragingly, the model shows several similar long-term responses to other, even unrelated, models.

We plan to conduct "science" experiments using this model, in which ecosystem responses are explored, but using a more rigorous approach that includes parameter optimisation techniques and uncertainty quantification- not a single parameter set, but suites of parameter values. This kind of experimental setup is outside the scope of GMD.

# I will limit my review therefore to the description of the model and the evaluation of its present-day state, with a focus on the silicon cycle.

First to the description of the model. The two main new aspects of the silicon submodel that is presented here are a) the description of the dissolution of opal in the water column and b) a simple benthic transfer function for opal that thinks into the sediment. Other aspects, like the description of diatom growth and Si uptake are more standard (which is not a criticism). The second of the new aspects, the benthic transfer function, is fairly simple (30\% of the opal sinking into the sediment is permanently buried if the flux is above a certain flux threshold, and 5% below), but the description is clear and the parameterization is a reasonable approximation to our current understanding of global sedimentary Si fluxes.

### Thank you.

The first new aspect, however, is first of all neither mathematically nor verbally described well enough. Opal sinking is not modelled explicitly, and only the divergence of its vertical flux (equivalent to the release of dissolved silicic acid) is represented. The textual

explanation of this in the single sentence "We approximate (What?) (by) an exponential flux function and apply our e-folding temperature parameterisation to represent microbially enhanced dissolution" states this in a rather unclear way.

Firstly, we apologize for Equation 32, which was presented incorrectly. This parameterisation of opal vertical dissolution takes a similar approach as the default UVic ESCM parameterisation of CaCO3 (Schmittner et al., 2008). We have rephrased the text (P 9, L 227):

Dissolution of opal in the water column is calculated by assuming instantaneous sinking of the vertically integrated production, where the flux of opal is distributed down the water column using the e-folding temperature parameterisation (unitless), scaled by a dissolution rate constant ( $\lambda$ Opal, in day-1) which is multiplied by the depth (z, in meters), divided by a sinking rate (wOpal, in meters day-1)

The following equation 32 is not better: first, the boundaries in the vertical integral are not stated. Well, we can assume that it is over the whole water column. But then follows a flux divergence term (which is not stated as such) d/dz(something), but it is completely unclear which of the terms that this derivative is applied to actually depends on z. In the form as it stands here, the only potentially variable term is ocean temperature, as both \$\lambda\$ and \$w\_D\$ are constants. That cannot be correct: if temperature was constant there would be no dissolution? I am also worried that by scaling the dissolution with the water column integral of production (and not the integral obove the depth where the dissolution is calculated), one could potentially dissolve more than is produced in the upper model layer, and given the very fast dissolution in warm waters used here I actually suspect that this happens in the model.

Again, we apologise for the incorrect formulation of Eqn 32, which did not reflect what is actually coded into the model. The equation is now corrected in the text (P 9, L 231).

$$\mathrm{Di}(\mathrm{Opal}) = \int_{0} \mathrm{Pr}(\mathrm{Opal}) dz \cdot \frac{d}{dz} \; (\frac{\lambda_{\mathrm{Opal}} \times z}{w_{\mathrm{Opal}}} \; \exp^{\frac{T}{T_{b}}}).$$

We have also added boundaries to the integral. Lambda is in units of per day, while w\_opal is in units of m/day. If temperature is constant then exp^(T/T\_b) is non-zero. A higher temperature produces a sharper gradient in dissolution profile between the surface and deep ocean. The model code constrains the formulations to ensure that there cannot be more dissolution than production, and this is now explicitly stated in the text (P9 L232): Dissolution is presented in units of mol Si m\$^{-3}\$. The code constrains total water column dissolution to be no greater than total water column production.

Besides being described in a clearer way, the new parameterization should also be justified better. The existing justification (line 225-234) is rather superficial. Making the dissolution of opal as strongly temperature dependent as the breakdown of organic matter is justified here by the requirement to strip away the organic coating of diatom frustules before dissolution can set in. But that argument (which is incorrectly ascribed to Sarmiento and Gruber, 2006 here; it is in fact from Bidle et al. Science, 2002) has originally been proposed to explain just the opposite, namely that very little dissolution happens within the euphotic zone, as the bacteria first need to break down the coating before dissolution can set in, while the frustules have by then already sunk down into the ocean's interiour. So this justification simply does not hold, and the only remaining justification is that it improves the fit to the WOA dissolved silicon distribution. That is ok, given the contradictig information from the many experimental studies and the widely varying temperature dependencies in models. But it should then be acknowledged that this parameterization may compensate for physical deficiencies in the model, e.g. an overly small vertical mixing due to the low vertical resolution.

We thank the reviewer for pointing out this mistake in our reasoning. We have removed discussion of microbial activity and instead re-phrase the paragraph as follows (L 234): We find this parameterisation offers improved model fit to World Ocean Atlas silica distributions relative to other parameterisations that we tested, e.g. the temperature-dependent parameterisation of Gnanadesikan (1999) or the temperature and oxygen-dependent parameterisation of Enright et al. (2014). The Gnanadesikan (1999) parameterisation yields lower dissolution rates at low temperatures than the Enright et al. (2014) parameterisation, which is similarly formulated but which includes an additional oxygen scaling. The Enright et al. (2014) oxygen scaling is not justified in their model description, but it has the effect of increasing Si dissolution rates in the deep ocean (exacerbating the overestimation of Si dissolution in this region by the Gnanadesikan (1999) scaling described in Ridgwell et al. (2002) and decreasing Si dissolution rates (to a lesser extent) in the near-surface. Our temperature scaling has the effect of raising dissolution rates at the surface may be necessary to compensate for the low vertical resolution of the model.

One aspect of the model evaluation, the modelled present-day distribution of dissolved silicate, has been already discussed by reviewer 2, and I have nothing to add here. But another aspect of the evaluation has left me completely confused, namely the comparison of the global Si fluxes with the published data-based estimates from Treguer et al. (2021) and Treguer and De La Rocha (2012). I will try to explain the inconsistencies in this comparison, as I perceived them, acknowledging that I got confused in several places. The comparison is done in lines 327-340 and in Table 6, and I have reverted the numbers given there to the more commonly used unit of Tmol Si/year.

- The authors state that "Diagnosed surface opal production is within the range of a recent estimate (Tréguer et al., 2021)". This seems to be the case. From my unit conversion I get a production of 270

#### Tmol/yr, compared with 255 in Treguer.

- But the Si export out of the euphotic zone is about half of the estimate given in Tréguer et al., 2021: 57, compared to 112 Tmol/yr. Assuming that most of the production is taking place in the upper 130 m of the water column, I thus calculate that the dissolution of opal within the euphotic zone alone is 218 Tmol/yr.

- But that contradicts with the statement of opal dissolution in Table 6, which is 138 Tmol/yr. They compare this number to a number of Treguer et al, which they state as 170 Tmol/yr. In Treguer et al., I find a dissolution within the euphotic zone of 143 Tmol/yr, and in the water column below the euphotic zone of 28 Tmol/yr, so I assume that they mean their 138 Tmol/yr as the total dissolution both within the euphotic zone and in the water column below. But that cannot be correct, see above.

- The dissolution of 138 Tmol/yr also cannot be the dissolution only below the euphotic zone, as we have learned above that the total export of opal is only 57 Tmol/yr. So, what is the dissolution number actually?

- in Treguer, the majority of the exported Si lands in the sediment (84 of the 112) and is mostly dissolved there, only 9 Tmol/yr gets buried. From Table 8 I see that this 9 Tmol is compared to a model value of 47 (they call this net seafloor Si flux). And from the model description I get that this is returned to the ocean with rivers. Is that so? On the other hand they say that they underestimate the riverine influx. This does not fit, maybe I misunderstand something serious here.

We again thank the reviewer for pointing out a serious mistake in our Table 6. This table had not been correctly updated in our last revision. The 2 km opal flux was from the Honjo et al. (2008) dataset, but we have now changed this to the Treguer et al. (2021) estimate. Also, we have converted our fluxes to Tmol Si/y to be consistent with convention. And, the seafloor flux in our model includes opal flux above 2 km depth, so there is more seafloor flux than flux below 2 km. Since our model does not estimate sedimentary processes, our seafloor flux is not a good comparison to the Treguer et al. (2021) burial estimate (and has been removed to avoid confusion). We have also updated the text to reflect these changes (throughout Section 3, see L 337-347 especially).

It is true that silica parameterised to be lost to the sediments in our model is replaced via the river fluxes. In our model, currently the estimated river input is 1.52 Tmol Si/y, which is below 8.1 Tmol Si/y estimated by Treguer et al. (2021). This is discussed at L 343: The calculated river flux is 1.52 Tmol Si y\$^{-1}\$; lower than the \citet{Treguer etal 2021} estimate of 8.1 Tmol Si y\$^{-1}\$...

We have now made it more clear that we compare the total water column dissolution (euphotic zone plus below) to the Treguer et al. (2021) estimate, and have added a discussion of production to dissolution ratios (L 343):

...but the ratio of total water column dissolution to biogenic production is more encouraging; 0.5 in our model compared to 0.67 calculated from Tréguer et al., 2021).

Overall, what we now show are silica flux and dissolution rates that are only about ¼ of the Treguer et al. (2021) estimates. We have also added some discussion of these results in the context of inverse modelling of silica cycling, which estimates rates based on a different approach than extrapolation of direct measurements, such as Treguer et al. (2021). L 347:

However, it is worth noting that previous inverse estimates of global silica fluxes (Holzer et al., 2014; Pasquier and Holzer, 2017) estimate global export production (166 and 171 Tmol Si y–1, respectively) substantially higher than the 112 Tmol Si y–1 estimated by Tréguer et al. (2021), which suggests poor agreement still exists across estimates of the global silica budget

My overall recommendation is therefore that both the description and justification of the new temperature-dependent opal dissolution, and the description of the present-day global Si flux balance still need major revisions for the manuscript to become a useful addition to the literature.

We hope our responses to the above points adequately address the Reviewer's concerns.

Minor comments

Line 97: Maybe it would be good here to mention in what unit the phytoplankton biomasses are calculated. This information is added (L 99)

Line 111-112: It has become common to call the Michalis-Menten uptake limitation term "iron availability" but this term creates the possibility to confuse it with "bioavailability of iron" used in the biological lierature, which is related to iron speciation (and actually has the unit of a concentration, not dimensionless, as the MM-term).

The original language was modified to "availability" at the request of an earlier Reviewer. We have changed it back to "uptake". (L 115)

Line 124, Eq. 8, line 125: First of all I find it a typographical crime to state a number \$8 \cdot 10^{-4}\$ as 8e-4, like in matlab code. Second, the numbers in the equation are stated without unit, which is false.

### The formatting is changed and units are added. (L 128/129)

Equations 9, 10, 11: Making the ChI:C quota just dependent on iron is a gross simplification, as it takes away the much larger dependency of ChI:C on irradiance in acclimated cells. For the purpose here it still may be o.k., but that should be stated. Secondly, the two factors \$\alpha\$ and \$\theta\$ that are both made linearly dependent on Fe, appear only in the combination of \$\alpha \theta\$ in Eq. 9, so in effect the relationship is quadratic. Is there any justification for that?

None of these equations are new to this version of the model. Please note that Chl is not an explicit tracer. This is a one-off gross calculation used for the iron model to link iron availability to light affinity. It is kept for historical reasons (it was introduced by Nickelsen et al., 2015). Flexible ratios are not introduced to other aspects of the model because the focus of this manuscript is silica cycling. Other researchers (Markus Pahlow, Chia-Te Chien) are currently working on this aspect of the KMBM (e.g., Chien et al. (2020), GMD, doi.org/10.5194/gmd-13-4691-2020).

We have added a new sentence (L 138):

The approximation of \$\theta\$ is simplistic and neglects other factors, e.g. <u>irradiance</u>, which can affect the ratio.

And another at L 132:

Nickelsen et al., (2015) introduced this parameterisation to the model and discuss it at length; interested readers are recommended to read their Section 2.3.2. In the above equation...

Equations 26, 27, and several others later: The authors have the habit of indicating sources of tracers (like in equation 26 the scavenging loss of iron to particles) by an expression \$[Fe]\_{orgads}\$, i.e. with square brackets around the element symbol. The convention with square brackets in the chemical literature, however, is that these denote 'concentration', so \$[Fe]\$ is the concentration of Fe. \$[Fe]\_{orgads}\$ would then be one part of the dissolved Fe pool, not a rate. The same confusing convention is used thoughout the manuscript, e.g. equations 35, 36, 40 and 44. Moreover, it is completely confusing to write the scavenging rate \$k\mathrm{Fe}\_{org}\$ as one could confuse this with a product of a rate \$k\$ with some concentration \$\mathrm{Fe}\_{org}\$. I fell for this and searched for the definition of it.

We agree that this concentration-style formatting might lead to confusion, but it was explicitly requested by a reviewer in the earlier stages of the revision. We have changed the formatting for scavenging.

Equation 39: I think the scavenging by CaCO3 is missing here.

#### It was, thank you. The equation has been corrected.

Lines 301 - 304, equation 44: It is nowhere mentioned where the river input of Si, which is used to balance the Si budget, is applied. Does the model use a prescribed runoff distribution also for the Si input, or is the river input distributed homogeneously over the ocean surface, as e.g. in HAMOCC?

River inputs are scaled against river flow using a mask provided as a standard input with the model. This is now added (L 314):

River inputs of silicate, alkalinity, and DIC are scaled against seasonally-variable river flow using the standard UVic ESCM version 2.9 O\_rivflux.F forcing file.

Line 312: Everywhere else, Si fluxes are given in Pg Si /yr, why here now in Tmol / yr?

We have changed all Si units to Tmol/y in the text.

Lines 379 ff: It is unclear to me what one learns from a Taylor plot of CMIP6 model diatoms relative to the distribution in the model shown here. I would suggest to do the comparison otherwise, not all referenced to one specific model.

We had produced the Taylor plot in this way because the CMIP models are just so different from both each other and the diatom observations that a Taylor plot of biomass referenced to the very sparse dataset was meaningless. By using our model as a reference we tried to show how similar the CMIP models diatom biomass are to each other and our model. However, we have now changed this plot to a comparison between CMIP6 model mean diatom biomass (black dot) and the KMBM3 (red dot). We have changed the text accordingly (from L 390):

We compare KMBM3 model output relative to annual mean CMIP6 model output due to the very low normalised correlation (0.04-0.15) and normalised standard deviation (0.10-0.22) of all models against the sparse Leblanc et al. (2012) diatom biomass dataset. At the time of writing, available CMIP6 simulated annual average diatom biomass shows diverse quantities (maximum concentrations from 0.0035 to 0.03 mol C m-3), and spatial distributions ranging from global maximum concentrations at the Equator (CanESM5-CanOE, Swart et al. 2019), to shallow seas and coastlines (IPSL-CM6A-LR, Boucher et al. 2018), to the high latitudes (CMCC-ESM2, EC-Earth3-CC, GFDL-ESM4, CESM2; Lovato et al. 2021; EC-Earth Consortium 2021; Krasting et al. 2018; Danabasoglu 2019, respectively). Fig. 5 summarises the KMBM3 diatom biomass estimate relative to the mean CMIP6 diatom biomass, simulated at year 2014. The KMBM3 is most closely correlated with GFDL-ESM4 (not shown). As stated above, phytoplankton biomass is difficult to tune for (particularly when multiple functional types are represented) due to the under-constrained parameter space, the theoretical nature of phytoplankton functional categories, and the sparsity of gridded, annually-averaged biomass datasets. Therefore, a wide range in model biomass estimates is expected across the CMIP6 ensemble.

Line 399: What are the units of the given root mean square errors?

#### Units are added to the text. (now L 405).

Line 664: The scavenging of iron on calcite is not new, as far as I know it is used in Moore's BEC model as well.

This is very interesting to know, thank you. We have removed "novel".

Figures 2, 3, 4, 6 and several more: It might make sense to re-do the plots with a proper map projection

While we appreciate that map projections are look sleeker, very few UVic model description papers, or UVic science papers for that matter, have used them. Therefore, we prefer to keep the figures as they are, as it makes it easier to compare them with previously published figures.

Figure 5: What sense does it make to show a Taylor diagram with the model presented hare as the reference data set? Also, I would not call this 'normalized' to KMBM3, but 'referenced'.

Please see our answer given above. We have changed the wording.

Fig. 14: The model is clearly missing the HNLC region wth elevated Si in the North Pacific

The low bias in Si in our model appears to affect representation of all HNLC regions, despite our model performing well relative to the CMIP6 suite. Hopefully this will be resolvable with our future parameter sensitivity studies.