General comments

The manuscript by Hajima et al., describes the update in MIROC-ES2L (CMIP6) ESM relative to previous MIROC (CMIP5) ESMs, with emphasis on the land and ocean biogeochemistry components. In addition to the standard piControl and transient Historical simulations, the authors also performed additional sensitivity simulations with the updated model to analyze the impact of new parameterizations and new model improvements on the simulated biogeochemical metrics such as the land carbon content, ocean primary production, nutrient limitations, etc. They also use a set of C4MIP 1%CO2 simulations to quantify the simulated carbon fluxes sensitivity to climate change and atmospheric CO2 increase, and compare these quantities with that from previous MIROC ESM and other CMIP5 ESMs.

The paper is well structured and nicely written, providing a good overview of the new model with key information that would be of interest for GMD readers who need reference to the carbon cycle components in MIROC-ES2L model. Below I have listed several suggestions and comments that the author can consider to further improve the manuscript prior to publication in GMD. I also disclose that I am not an expert in land carbon cycle and therefore can only offer limited reviews on the land biogeochemical part of the paper. It would therefore be good to have additional referee(s) who can offer better perspectives on the land discussions of the paper.

Specific comments

The title gives away an impression that the paper fully describes the MIROC-ES2L model. While it does briefly, most of the content is focused on the land and carbon cycle description and assessments of the performed (biogeochemical-focused sensitivity) simulations. I therefore suggest modifying the title to something like “Description and sensitivity analysis of the carbon cycle components of MIROC-ES2L Earth system model.”

The performed sensitivity experiments are very useful to characterize the model’s response to (some are newly implemented) external forcing and future climate change. Many of the simulated changes in the key metrics, such as the ocean primary production and export production were quantified and discussed (e.g., Page 35). In addition to this useful discussion, I recommend the authors to expand (add additional columns) Tables 3 and 4 to include the same metrics but from the relevant sensitivity experiments

Some information on how stable (or any drift, if exist) the ocean biogeochemistry budgets in the preindustrial control run would be useful. Also, one of the uniqueness of the MIROC-ES2L among other CMIP6 models is the fact that it has a very long model spin up (i.e., >3500 years; albeit some are offline), which is thought to be crucial to reduce model bias (e.g., in the interior ocean biogeochemical tracers such as oxygen; Seferian et al., 2015; GMD). I think it would be of interest to many readers to know whether in fact interior biases is improved through this computationally costly process.

On Page 7, the authors wrote “a static biome distribution”. Please briefly clarify what you meant by this? Do the prescribed PFTs change annually or or static in the sense that it is constant from preindustrial states?

The Figures’ resolutions and qualities need some improvements. Namely, Fig. 2,3,4,10,12: x-labels and y-labels are difficult read in my printouts. Legends in Fig. 10 are also too small to read.

On Page 17, there is a statement that the model might overestimate net carbon uptake by land and/or ocean, or underestimate LUC emissions. This was implied through comparing the diagnosed atmospheric CO2 concentration with the observed values. I found this strange. Since the simulation was performed with prescribed CO2 concentrations, the authors should instead compute the diagnosed anthropogenic emissions, and compared that with the observations (as in Jones et al. 2013), e.g., 403PgC. Based on this diagnosed emissions, the authors can then make a statement whether or not the land and/or ocean sinks are under/overestimated. Furthermore, if the authors have completed the prognostic CO2 simulations, it would be interesting to compare the diagnosed atmospheric CO2 concentrations and determine if the analysis on land/ocean carbon uptake strengths is consistent with the above approach.

Ocean carbon uptakes. While there's discussion on the cumulative carbon sinks over the historical period, there was no discussion on the simulated contemporary CO2 sinks, i.e., annual mean and spatial distributions (only stated in in Table 4). Given the importance of carbon sinks and its application in ESM simulations, I recommend adding some air-sea CO2 flux comparison between the model with e.g., spatial patterns from Landschutzer et al., 2014-https://doi.org/10.1002/2014GB004853; or Takahashi et al., 2009- https://doi.org/10.1016/j.dsr2.2008.12.009), and this discussion can be linked to the currently still limited surface DIC/alkalinity discussions, as shown in Fig. 11.

In addition to the contemporary spatial air-sea CO2 flux patterns, several studies have shown the importance of simulating proper regional seasonal cycle for constraining long-term spread in ESM projections (Kessler and Tjiputra, 2016-ESD; Goris et al., 2018-J. Climate). If the authors agree, it would be great to see how the regional CO2 fluxes in key world ocean regions (North Atlantic, Southern Ocean, etc.) are simulated in MIROC-ES2L and compare that with results from the previous model version.

Similarly for land, if seasonal cycle is an important criteria for constraining future projections (I am not fully aware of such studies on the terrestrial carbon side), then such presentation could be considered as well, e.g., similar to that shown in Tjiputra et al. (2013 - GMD) for GPP across different latitudinal bands.

In Section 3.1.4, MLD bias (too deep) is described as the partial reason for the SST cold bias. One could argue the other way around. On the other hand, in the Southern Ocean, there is a general warm bias, yet the MLD in the model also appears to be too deep. This should be elaborated. In this regard, i would also suggest adding Southern Hemisphere sea ice extent to Fig. 8.

Figure 15 shows that in many regions, there are two (instead of one) nutrients limit the primary production. It is not clear to me how the authors arrive with two limiting nutrients. Shouldn't it be only one, which is the minimum of the three? Some clarification on how it is derived would be useful.

Figure 17 and Table 5 essentially show the same information and appear redundant. I suggest removing Fig. 17 and keep the table.

Technical corrections

In addition to the above comments, please find below a list of minor comments that needs to be clarified as well as edits to improve the manuscript readability.

L25: Comparison … the model could reproduce well the transient global climate change and carbon cycle as well as reproduce the observed large-scale spatial patterns of land carbon cycle and upper ocean biogeochemistry.
L29: … revealed that the simulated ocean biogeochemistry could be altered regionally (and substantially) by …
L35: Suggest removing “model performance in”
L36-7: The MIROC-ES2L could further improve our understanding of climate …
L42: … on simulations using atmosphere- …
L43: had evolved
L48: future climate due to processes …
L52: semicolon between Watanabe and Collins references.
L55-6: As ESMs explicitly simulate …, they can simulate the temporal …
L60: Furthermore, their simulations can be …
L66: … on climate manifests through … OR … on climate is manifested through …
L69: … ocean, and eventually …
L77: transport instead of transportation
L100: you can also cite Kessler and Tjiputra (2016; https://doi.org/10.5194/esd-7-295-2016)
L111: remove ‘nutrient’
L117: as well as the physical response
L162: with existing studies.
L170: used in the CMIP5
L182: … ocean physical model … land physical model …
L183: river routing (?)
L201: what is the ocean horizontal resolution?
L236: prescribed in the forcing data
L236-7: The fluxes of nitrogen out of the land ecosystem are simulated through N2 and N2O production during nitrification …
L277: is transported by rivers
L281: erosion and dissolution of organic carbon in the ocean, these processes are not activated to close the global mass conservation of carbon and nitrogen.
L283: replace ‘simply diagnosed’ with ‘only for diagnostic purpose’
L290: … are simulated with 13 biogeochemical tracers.
L294: .. the Redfield ratio of C.N:P:O= …
L295: remove ‘types of’
L298: The nitrogen cycle in OECO2 is similar ..
L301: denitrification is also …
L330: assumed to be
L334: remove ‘oxygen’ (it is not remineralized, it is consumed)
L347: this is only true when the model simulations is configured to be fully interactive, right? i.e., prognostic (not prescribed) atmospheric CO2.
L376: carbon cycle variation relative to the preindustrial control (CTL). …. climate-carbon feedback (Arora et al. 2013; Schwinger et al., 2014-https://doi.org/10.1175/JCLI-D-13-00452.1)
L386: these experiments
L387: I don’t understand ‘which would be noise in the analysis’. Maybe clarify or remove it.
Table 1 caption: Summary of experimental details.

Table 1, row NO-NRD: replace combination with combining; replace ‘ doesn’t get impact from’ with ‘ is not impacted by’
Table 1, row NO-FD: replace ‘ doesn’t get impact from de’ with ‘ is not impacted by Fe’
L414: replace status with states
L415: … is calculated independently of ocean biogeochemical states.
L427: The first term on the right
L432: replace ‘which is because’ with ‘ since’
L437: second term
L438: This sentence is a bit confusing. TCRE quantifies only the global temperature change in response to emissions, and not the ‘entire climate-carbon cycle response’, right?
L441: … response to atmospheric CO2 increase …
Eqs (8) and (9) are wrong.
L456: … (TOA), and anomalies of new-surface …
L473-4: … internal climate variability (Kosaka et al., 2016).
L492: … positive), and anomalies of (b) …
L508: 166Pg, it looks closer to 200PgC in the figure.
L526: replace alleviated with weakened
L527: … within the independent estimates range of …
L528: … where the estimation of uncertainties take into account both …
L529: remove ‘ by Le Quere et al. (2018) are considered’
L531: … the model shows an increase in carbon …
L542: … conservation in the ocean biogeochemical component.
L548: … would not induce significant global-scale impact …
L570: remove global
L574: In the lower …
L575: replace ‘against’ with ‘driven only by’
L597: .. reveals the annual inputs of …
L598: … increase to 46 …
L606: Shouldn’t 4.5 be 4.2? (Table 3: 13.7-9.5)
L615: … shows a net …
L616: “… stimulates ecosystem nitrogen demand.” Is this implied from the increasing atmospheric N2 fixation in the model?
L623: suggests that the historical …
Fig. 4 caption: Rate of change of global nitrogen budget in (a) land and in (b) ocean …
L660: … attributed to the model bias in simulating cloud cover ….
Fig 6: Why not show the difference between HIST and GPCP as Figs. 5 and 7?
L684-5: suggest revising it to something like:
When projecting future climate change, it is important for an ESM to reproduce the observed climatological patterns of key physical and biogeochemical tracers (Ohgaito and Abe-Ouchi, 2009).
L701: cover. Sea ice …
L701: … summertime minimum concentration is slightly …
L718: Briefly describe how the MLD is computed in the model.
L7738: replace higher with high
L740: is generally still underestimated
L749: Suggest replacing the reference Anav et al. with Jung et al., 2011.
L756: vegetation carbon content including …
L761: define NCSCDv2
L780: replace panels with rows
L784: add units ti the bottom row figures, presumably (g C m^-2)

L797: replace nitrate with nutrient
L799: consistent with the observed
L800: because of the implementation
L801: Ocean circulation also …
L803: … over estimation of nutrients entrainment to the surface and thus …
L804: The simulated global mean vertical profile of nitrate concentration …
L804-6: I am not convince this is the main or only reason of the well fitted model nitrate with observations. Model can simulate correct sources and sinks but still compare poorly with observations if the circulation fields is correct. There are several ways to get analyze whether the circulation fields is reasonable or not, e.g., using apparent oxygen utilization tracer. Nevertheless, you can get correct distribution for the wrong reason (Duteil et al., 2012 - Biogeosciences)
L825: … the equatorial Pacific Ocean, and the Southern Ocean …
L826: … the equatorial Pacific … than the observed …
L828: … much of the low-latitude surface ocean …
L846: … (Fig. 9), which also transports …
L848: replace parameters with tracers
L847-50: You mention alkalinity bias leads to DIC bias. Can you elaborate what cause the alkalinity bias in the first place?
L882: replace Northwest with 'path of north'
L950: remove 'one of'
L991: .. bias, which results in overestimation …
L995: remove 'produce a resultant'
L1002: .. the model is capable …
Fig. 16b shows that for the Neva, Yukon, and Churchill rivers, the DIN flux is larger in HIST-NOLUC than in HIST. This is in contrast to other rivers. Can the authors clarify the mechanism behind these patterns?
L1064: … is simulated to be 0.52 …
L1067: … is within this spread …
L1081: … is the main cause for the lower AF, making …
L1082: … feedbacks more positive and less negative, respectively, …
L1111-2: … plant growth and therefore the change in the land carbon fluxes.
L1115: … thus, the marine productivity is now also affected by the riverine nitrogen input.
L1118: … in response to changes in external iron inputs.
L1120: … We confirmed that …
L1122-3: … the model, the MIROC-ES2L's good performance in simulating … is inherited from its original …
L1114: … as found in some climate models.
L1126: Actually, there is no comparison of global biogeochemical (nitrate, phosphate, oxygen, etc.) budget trends, except for carbon. Suggest revising it to: … capture the observations-based estimates of contemporary air-sea and
air-land carbon fluxes.
L1127: … assessed through comparison …
L1228: the model produced reasonable …
L1133: …. compared with those of the CMIP5 ESMs.
L1138: … This is reduced from the value seen in the model …
L1141-4: Suggest rephrasing the sentence to: A multimodal comparison on feedback strengths using CMIP6 ESMs is necessary to potentially determine whether the climate and carbon cycle sensitivities in MIROC-ES2L are realistic, and furthermore to establish constraints on each feedback process based on observations (e.g., Wenzel et al., 2016; Goris et al., 2018 - J. Climate).

L1148: model alters the land carbon cycle. … carbon content during 1850-2014 is 44
L1154: nitrogen cycle alters the carbon cycle … did not quantify to what extent the soil nutrient ….
L1159: remove 'inputs'. 'external sources' already implies 'inputs'
L1164: replace GPP with primary production. You have used NPP in section 3.1.6 and 3.2.2. Generally, in ESMs, we refer it to simply 'primary production'
L1173: … should also be …
L1177: here you said strongly impact, but on L1160, you states minor contributions. Please clarify.
L1179-80: … radiative balance in the atmosphere. Nitrous …. gases with a long ….
L1191: … a similar set of sensitivity simulations should be …
L1192: remove 'with which'
L1193: replace explorations with quantifications
L1196-7: … ESMs can reproduce some of the dominant long-term environmental changes on Earth …
L1198: ocean acidification …
L1201: replace evolve with improve
L1447-9: Shouldn't GCaCO3 be PrCaCO3 instead (based on Table B2)?
L1467: see eq. (14) of
L1498: (CE^D) should be (CE^P)
L1596: … Cycles, 26, GB2009, doi …
L2022-4: this citation needs to be updated.

Supp Fig 3: If I understood correctly, the 'N2 fixation' box should be labeled 'Diazotroph'.
There should also be arrows from PHY and ZOO to CaCO3?
Arrows from PHY to nutrients should only be labeled (Fast) remineralization, and the arrows from PHY to "DNO3, DFe, (and DPO4)" should then be labelled 'mortality'
What is the purpose of the green label 'Definition of Alkalinity'? I suggest removing it.
Should the 'Nitrification' arrow be reversed?

Finally, if any of my comments are unclear, please feel free to contact me.