Review comments in black.

Responses in green.

The paper describes the Earth-system model CanESM5, which has performed simulations in support of CMIP6. The paper itself is very well written with very few grammatical or typographical errors. I was particularly impressed with the detailed discussion of how the model is managed with regard to the structure of the various git repositories. The improvements to model throughput are also impressive, going from 4.5 to 16 simulated years per day from the various changes described. The model itself represents some quantities well and others not so well, as might be expected in any ESM. Given the model has a relatively coarse resolution and does not represent all processes the behaviour is not unreasonable.

We thank the reviewer for their constructive comments.

Major Comments

In the Introduction the authors briefly discuss CanESM5 within the modelling system of CCCma, which includes CanRCM, CMAM, and CanSIPS. It might be helpful to describe how these are connected - are they all entirely separate, or do they share components/models?

To clarify the connection we have updated the sentence to read “CanESM, forms the basis of the CCCma modelling system, and shares components with the Canadian Regional Climate Model (CanRCM) for finer scale modelling of the atmosphere (Scinocca et al., 2016), the Canadian Middle Atmosphere Model (CMAM) with atmospheric chemistry (Scinocca et al., 2008), and the Canadian Seasonal to Interseasonal Prediction System which is used for seasonal prediction and decadal forecasts (CanSIPS, Merryfield et al., 2013).”

Figure 1 - I find this figure a little frustrating. While it is clear in naming the different component models that make up CanESM5, it does not describe how they are connected or each model’s complexity. A reader must consult Appendix A, which is somewhat technical and is still not complete for all linkages as this only discusses what goes through the coupler. As it is this figure doesn’t really show the “evolution of components” as it does not describe how complex each component model is in either CanESM2 or CanESM5, although it does show the changes to model version numbers or the change of model used.

We agree this figure was not adding much value, and have removed it in the revision. We have expanded the description of the model in Sections 2.1 and 2.6, providing the reader with more clarity on the complexity of each model component.

What I felt was missing in section 2 was a description of just how complex the atmosphere model CanAM5 is. The description is very brief and readers are directed to Cole et al. (2019), which although will also be in the same special issue as this paper, is currently still in preparation. Also, this paper describing CanESM5 should stand enough on its own such that important details regarding component models should be presented here. I suggest that the authors expand section 2.1 enough to briefly cover the complexity of the atmosphere model and what processes (or not) are represented. There is also no mention of atmospheric chemistry or the complexity of the aerosol scheme.

We have expanded section 2.1 to provide a more complete overview of the complexity of CanAM, including the treatment of aerosols and atmospheric chemistry. We have also added a new section (2.6) to describe the treatment of greenhouse gasses across the ESM components.

Added to Section 2.1.:

“Version 5 of the Canadian Atmospheric Model (CanAM5) employs a spectral dynamical core with a hybrid sigma-pressure coordinate in the vertical. The package of physical parameterizations used by CanAM5 are based on an updated version of its predecessor, CanAM4 (von Salzen et al., 2013). The physics package includes a prognostic cloud microphysics scheme governing water vapour, cloud liquid water, and cloud ice; a statistical layer-cloud scheme; and independent cloud-base mass-flux schemes for both deep and shallow convection. Aerosols are parameterized using a prognostic scheme for bulk concentrations of natural and anthropogenic aerosols, including sulfate, black and organic carbon, sea salt, and mineral dust; parameterizations for emissions, transport, gas-phase and aqueous-phase chemistry, and dry and wet deposition account for interactions with simulated meteorology. CanAM5 employs a triangular truncation at total wavenumber 63 (T63) corresponding to an approximate isotropic resolution of 2.8 degrees in both latitude and longitude. In the vertical, 49 levels are employed with layer thicknesses that increase monotonically from approximately 100 m at the surface to 2km at ~1hPa – the domain lid.”

“2.6 Treatment of greenhouse gases

CanESM5 represents radiative forcing from individual greenhouse gases (GHGs). Aside from CO_2, the concentrations of all radiatively active gases are specified and transiently evolve. Of these, CH_4, N_2O, and families of CFCs are assumed to be well-mixed, while, O_3 is specified as varying spatially – typically employing that prescribed for CMIP6 (Checa-Garcia et al. 2018). CanAM5 offers two modes for modelling CO_2 concentrations - as specified time-evolving concentrations; or as a three-dimensional passive tracer driven by land/ocean surface emissions, prognostically derived through interactive coupling with biogeochemical carbon models in the land and ocean. For example, CanESM5 can be run with prognostic CO_2 in concert with specified anthropogenic fossil fuel emissions to simulate atmospheric CO_2 concentration through the historical and future periods. Wetland methane emissions simulated by CLASS-CTEM, in contrast, are purely diagnostic. While these emissions respond to changes in climate and atmospheric CO_2 concentration (through changes in vegetation productivity), they do not modify atmospheric CH_4 concentration, which are specified.”

The authors should ensure that each component model is described in sufficient detail to understand what processes it can simulate and how these models interact together in the ESM. As this is an Earth-system model, I am interested in seeing just how the component models interact. The tables in Appendix A only cover what goes through the coupler, so methane emissions that are discussed on lines 125-130 aren’t mentioned for instance - are there any others? Here what do the methane emissions do, is there a simple methane oxidation scheme, does it produce water vapour etc.?

We have expanded the description of CanAM (Section 2.1), and we now believe that each component is described in sufficient detail to understand its complexity. We have added a new section to the manuscript titled “2.6 Treatment of greenhouse gases” to specifically address the question of treatment of methane (and other GHGs).

At line 422 the authors mention “five realisations of CanESM2”, and then later discuss a 50 member large ensemble of CanESM2. Is it just that only 5 members were submitted to CMIP5? These are first mentioned on line 844. Is there a reference or more detail that can be provided?

Yes this is correct. Five realizations of CanESM2 were submitted to CMIP5. Later, the 50 member larger ensemble was created by branching those five realizations into ten each in the year 1950 (by perturbing a random seed). To clarify this we have added the text below to the start of Section 6.2, in which the CanESM2 large ensemble data are used:

“Here we make use of the CanESM2 50-member large initial condition ensemble (Kirchmeier-Young et al., 2017; Swart et al., 2018). The 50 realizations in this ensemble were branched in the year 1950 from the five CanESM2 realizations submitted to CMIP5, and were forced by CMIP5 historical (1950 to 2005) and Representative Concentration Pathway (RCP) 8.5 (2006 to 2100) forcing. ”

Minor Comments

Out of interest, why does the numbering system go from CanESM2 to CanESM5?

To clarify this we have added the text to Section 1:

“The leap from version 2 to version 5 was a one-off correction made to reconcile our internal model version labelling with the version label released to the public.”

line 33 - I believe that “earth” should be capitalised in this context.

Changed.

line 149 - I’m not sure why “converted into ice” is in quotes here. Is there a reason for this?

The quotes have been deleted. They were used as ice mass is not explicitly tracked, as described, and hence the conversion into ice is implicit.

Section 2.3 - the way numbers and units are represented here are slightly different from other sections. The × symbol is rather large (compare to line 236), and units are given as m2/s rather than m2s−1 as is done elsewhere in the manuscript.

Changed.

line 280 - why are “bilinear” and “conservative” in quotes? I don’t think this is necessary here.

Changed.

Section 5.4 - two different ways of doing a ±symbol are used (lines 621 and 642).Given the differences in Section 2.3 (mention above) I suggest that the authors double-check how numbers and symbols are presented for consistency.

Changed.

Table B1 (versioning) - should this be “CanESM.vX.Y.Z”?

Changed.

Review comments in black.

Responses in green.

Review of Swart et al. 2019: The Canadian Earth System Model version 5 (CanESM5).

General Comments:

This paper gives a clear and detailed description of the new Canadian Earth System model (CanESM5). The paper nicely describes the main developments relative to earlier Canadian models (e.g. CanESM2) and also provides a good overview of the primary model performance, sensitivities and associated biases, touching on (i) pre-industrial coupled climate stability, (ii) performance over the recent historical past and(iii) first-order sensitivity to increasing CO2 concentrations. It will be a useful reference to users of the model and to climate analysts using CanESM5 CMIP6 results. I recommend the paper be published, conditional on one relatively “major” comment and a number of more “minor” comments being addressed. These are detailed below.

I will structure my comments with some more major ones first followed by a small number of minor errors. I believe that with some small changes to the manuscript the paper would be ready for publication.

We thank the reviewer for their constructive comments.

Major Comment:

The paper assesses the performance of CanESM5 over the latter part of the CMIP6historical simulation period, using a number of standard diagnostics where CanESM5 is compared to observations. A number of model biases/shortcomings are identified, but the likely/probable cause of these biases is never discussed. While I recognise it is not straightforward to fully identify the underlying cause of coupled model errors, it would be a useful addition to the paper if, at least for the most important biases(e.g. excessive surface temperature trends post-1980, precipitation and cloud biases, excess heat uptake in the tropical ocean thermocline etc) some discussion of the cause of these biases was provided. This does not have to be exhaustive and can point to a deeper analysis and discussion to come elsewhere, but some pointers to the main causes of key biases would give the paper a bit more “meat”. This comment also applies to the significant change (increase) in TCR and ECS in CanESM5 relative to CanESM2, a somewhat greater discussion on what is considered the leading cause(s) for this increase seems appropriate.

We have added to the existing descriptions of possible reasons for the largest biases as follows:

On equatorial thermocline biases (Sec. 5.4) we have added, such that it now reads:

“The pattern of excessive heat accumulation in the thermocline is very similar to the pattern of bias seen in CMIP5 models on average (Flato et al., 2013 their Fig. 9.13). Also similar to CMIP5 models there is a cold bias in the ocean below the thermocline. This suggests that the processes controlling the redistribution of heat between the thermocline and the deep ocean play a role in establishing the vertical structure of these temperature biases. For example, Saenko et al. (2012) find that heat redistribution in ocean models can be sensitive to the vertical structure of diapycnal mixing. ”

On the MOC strength (Sec. 5.4), we have added:

“The fairly weak AMOC in CanESM5 is likely associated with excessive sea-ice cover in the Labrador Sea, which inhibits convection.”

On precipitation and cloud biases (Sec. 5.2) we have added:

“The overall pattern of precipitation and cloud biases is very similar to that seen across the CMIP5 models (Flato et al., 2013) and are in part related to the well known issues of a double intertropical convergence zone (ITCZ), and an excessive equatorial Pacific cold tongue (Lin, 2007; Li and Xie, 2014).”

On ECS in section 6.1 we have added to the existing description, such that it now reads:

“A detailed explanation of the reasons behind the increased ECS in CanESM5 over CanESM2 is beyond the scope of this paper. However, the effective radiative forcing (Forster et al, 2016) in CanESM5 due to abrupt quadrupling of CO_2 is very similar to that in CanESM2, suggesting that changes in feedbacks rather than forcings are the source of the higher ECS. Indications are that the increase in ECS is associated with cloud and surface albedo feedbacks, with sea-ice likely playing an important role in the latter effect. The cloud albedo feedback is found to be sensitive to parameter settings in the cloud microphysics scheme. A more detailed examination of the changes in ECS due to cloud microphysics will be provided in a companion paper in this special issue (Cole et al, 2019). ”

On historical warming rates, we discuss the possible causes in Sec. 6.2.1 in which we raise four possibilities, we rule out internal variability, and we suggest that the increase in historical warming is most likely driven by the increases in ECS.

Overall we have tried throughout Sections 5 and 6 to provide concise descriptions of our best understanding of the causes of biases. However as the reviewer notes, it is often extremely difficult to identify the cause of coupled model biases. Indeed, many of the biases noted are common across many models, and their causes remain outstanding questions within the broader community. We feel that it is beyond the scope of this particular paper to explore the causes of biases in further detail. The primary goal of this paper is to provide an overview of the model code and a basic description of model performance to set the stage for further usage and analysis. We note that the remainder of the special issue is available for a more detailed analysis of biases (and we intend to use it as such).

Minor comments:

Minor comments are listed with reference to the line numbers in the paper. Where I think some additional explanation would improve the paper I outline this within each point.

1. L50: “CanESM5 represents a major update to CanESM2”: The reader is left wondering what happened to CanESM3 and 4 ?? Is there some reason naming jumped from “2” to “5”?

To clarify this we have added the text to Section 1:

“The leap from version 2 to version 5 was a one-off correction made to reconcile our internal model version labelling with the version label released to the public.”

2. L86-L87: “the emission of mineral dust and DMS was improved...”. It would be useful to know the degree to which dust, DMS and other natural (aerosol) emissions are fully prognostic in the model, allowing potential future feedbacks to be simulated, versus these using observation-based (and therefore time-invariant) prescribed input fields.

Section 2.1 has been expanded to clarify the treatment of aerosols:

“The aerosols are parameterized using a prognostic scheme for bulk concentrations of natural and anthropogenic aerosols, including sulfate, black and organic carbon, sea salt, and mineral dust; parameterizations for emissions, transport, gas-phase and aqueous-phase chemistry, and dry and wet deposition account for interactions with simulated meteorology."

3. L101-L102: "The introduction of dynamic wetlands and their methane emissions is a new biogeochemical process...". Is methane a prognostic atmospheric variable inCanESM5? e.g. how do the wetland methane emissions influence the model radiation? There is no mention of interactive chemistry in CanESM5, I therefore assume methane is not a prognostic variable and that there are no internal (model) feedbacks between wetland methane emissions and climate. How methane (and other non-CO2 radiatively important gases e.g. ozone) are treated in CanESM5 needs to be more clearly explained.

We have added a new section to the manuscript (2.6) subtitled "The treatment of greenhouse gases", to address this question, which explains that methane is not prognostic (it is specified), as is ozone. For completeness, we have also added comparison of wetland extent simulated by CanESM5 with observation-based estimates and the temporal evolution of annual maximum wetland extent and wetland methane emissions over the historical period in a new section in the appendix.

4. L121: Why does CLASS use 4 PFTs but CTEM use 9? And is vegetation type/amount dynamically predicted in CanESM5 or (externally) prescribed? Please state what is done for the latter and explain the rationale for the former.

The following text is added/modified to Section 2.2 to address this question:

"The reason for separation of PFTs for CTEM is the additional distinction that biogeochemical processes require. For example, the distinction between deciduous and evergreen versions of needleleaf trees is needed to simulate leaf phenology prognostically. Once leaf area index has been dynamically determined by CTEM all CLASS needs to know is that this PFT is needleleaf tree since the physics calculations do not require information about underlying deciduous or evergreen nature of leaves. Similarly, the C_3 and C_4 photosynthetic pathways of crops and grasses determine how they photosynthesize thus affecting the calculated canopy resistance. However, once canopy resistance is known CLASS does not need to know the underlying distinction between C_3 and C_4 crops and grasses to use this canopy resistance in its energy and water balance calculations.

While, the modelled structural vegetation attributes respond to changes in climate and atmospheric CO_2 concentration, the fractional coverage of CTEM's nine PFTs is specified. A land cover data set is generated based on a potential vegetation cover for 1850 upon which the 1850 crop cover is superimposed. From 1850 onwards, as the fractional area of C_3 and C_4 crops changes the fractional coverages of the other non-crop PFTs are adjusted linearly in proportion to their existing coverage, as described in Arora and Boer (2010). The increase in crop area over the historical period is based on LUH2 v2h product (http://luh.umd.edu/data.shtml) of the land use harmonization (LUH) effort produced for CMIP6 (Hurtt et al., 2011)."

5. L304-306: It sound as though the primary tuning for the coupled pre-industrial model simulation was global mean surface temperature. Is this correct? It would seem to me that (a zero) global mean net top of atmosphere radiation would be a more appropriate target. The resulting net TOA balance looks to fairly close to zero (0.11Wm-2, figure 5a). Was the TOA net radiation budget also used as a tuning target?

Yes we absolutely agree. We implicitly meant achieving a reasonable global mean surface temperature with a TOA balance of (near) zero - but we have now made this explicit as follows by adding to this sentence:

"while maintaining a net TOA radiative balance as close to 0 W m^{-2} as possible"

6. L318-319: "The final adjustment was to the carbon uptake over land......achieved via the parameter which controls the strength of the CO2 fertilization effect". Does this indicate some structural problems with the parameterization of CO2 fertilization? I ask because in L813-815 there is mention of; "significantly larger uptake of CO2 by the land biosphere in CanESM5 relative to CanESM2 in the 1pctCO2 experiment". Is this change in land uptake influenced by the tuning applied to the historical period CO2 fertilization term? Allied to this does CanESM5 include any representation of nutrient limitation of the land CO2 fertilization term?

The following text is added to section 2.2 to address this comment:

"Both CanESM5, and its predecessor CanESM2, do not include nutrient limitation of photosynthesis on land since terrestrial nitrogen cycle is not represented. However, both models include a representation of terrestrial photosynthesis downregulation based on Arora et al. (2009) who used results from plants grown in ambient and elevated CO_2 environments to emulate the effect of nutrient constraints. The tunable parameter determining the strength of this downregulation, and therefore the strength of the CO_2 fertilization effect, is higher in CanESM5 than in CanESM2 resulting in higher land carbon uptake in CanESM5. The tuning of this downregulation parameter value, used in CanESM5, is explained in Arora and Scinocca (2016) who evaluate several aspects of modelled historical carbon cycle against observation-based estimates. A land nitrogen cycle model for CTEM is currently being developed which will make the photosynthesis downregulation parameterization obsolete in future versions of the model."

We also add in Section 6.1:

“The reduction in TCRE occurs despite the fact that CanESM5 has a larger temperature response (TCR) than CanESM2 due to significantly larger uptake of CO_2 by the land biosphere in CanESM5 relative to CanESM2 in the 1pctCO2 experiment. As mentioned in Section 2.2, this is due to higher strength of the CO_2 fertilization effect in CanESM5 relative to CanESM2. As shown in Arora and Scinocca (2016) this leads to land carbon uptake in the 1pctCO2 simulation that is higher than in all CMIP5 models compared in Arora et al. (2013).”

7. L320: “Critically, no tuning was undertaken on the climate system response to forcing”. This is not strictly true as the authors say that they tuned the model to improve aspects of the historical simulation, which does include perturbed CO2 forcing.

The tuning based on the historical simulation was for land carbon uptake. In specified CO2 simulations (i.e. most of CMIP6, including the abrupt-4xCO2 simulation), land carbon uptake is purely diagnostic and does not influence climate sensitivity, which is what we are referencing here. To make this more accurate/explicit, we have changed this line to:

“no tuning of climate sensitivity was undertaken”.

8. Figure 4: This figure appears on page 15 but is not discussed until much later. In itself that is okay, but then it might be worth adding a statement where the figure is first shown, saying why this figure is sufficiently important to be shown at the start of this section and that further discussion of the figure results appear in sections xx or yy.

We have added to the text :”Fig. 4 shows the global mean surface temperature for several of the key CMIP6 experiments, which can be used to infer important properties of the model, as discussed further in Sections 4 to 6”

9. Figure 5 (caption). Some of the figures are explicitly listed as being “global” or “global mean” while others are not. Those not classified as “global” or “global mean”, if they are not, for what spatial region are they representative? If they are all global quantities then make this clear in the figure caption.

The Figure caption has been clarified to indicate that all quantities are global means.

10. Figure 6 (caption): “Variables are labelled according to the names in the CMIP6 data request“ As not all readers will be CMIP6 experts, it could be useful to list what this shorthand list means in real word descriptions, possibly in a table in an appendix.

We have inserted table F3, which includes the CMIP6 names and (real world descriptive) long names of each variable used, and we reference this new table in the caption to Figure 6.

11. L520-521: Some of the precipitation biases, while large in percentage terms, are actually very small in absolute terms (e.g. the subtropical East Pacific). Would it be better to plot absolute precipitation biases? as done for temperature and SLP.

We have converted this figure to use absolute biases

12. For Figure 7 (and possibly 8). I suggest showing DJF and JJA seasonal means rather than annual means. This will help focus on biases and reduce the risk of error cancellation within the annual mean, e.g. winter vs summer surface temperature biases, better focus on summer vs winter monsoon precipitation biases etc.

We have converted this figure to show DJF and JJA seasonal means (new Figs. 5, 6, 7).

13. Figure 8 and L530: It would be useful to plot TOA solar and longwave radiation fields and biases (e.g. vs CERES EBAF) alongside the cloud plots, given that the number 1 consequence of cloud errors are in their impact on radiation.

Figure 8 has been expanded to show seasonal means, as requested above. We have elected to leave more detailed comparisons (vs CERES) to the upcoming documenting papers that are specifically focused on CanAM.

14. L589-591: The wording in these lines does not make sense.

Thanks for pointing this out typo. We have adjusted the text to read as follows: "The observations-based temperature and precipitation data used in these plots are from CRU-JRA reanalysis data that were used to drive terrestrial ecosystem models in the TRENDY Intercomparison for the 2018 Global Carbon Budget (Le Quéré et al., 2018)."

15. Figure 13: (caption) “The shading presents the corresponding inter-quartile range”: What range are we referring to here? For the model results is the range across ensemble members? What is the range with respect to the observations?

The shading presents the inter-quartile range that results from the combined effect of inter-annual variability and longitudinal variability. Model results show the ensemble member r1i1p1f1 only. We have adjusted the Figure caption as follows: "Figure 13: Zonal mean values of (a) GPP, (b) HFLS, and (c) HFSS for CanESM5 (r1i1p1f1) (black) and reference data (red) from Jung et al. (2009). The shading presents the corresponding inter-quartile range that results from inter-annual variability as well as longitudinal variability for the period 1982 to 2008."

16. Figure 14: What occurrence of surface temperature and precipitation is plotted? Daily mean values, monthly means? Please state what is plotted.

The values present monthly mean values averaged over the period 1982 to 2008. We have adjusted the Figure caption as follows: "Functional response of GPP to (a) near-surface air temperature and (b) surface precipitation for CanESM5 (r1i1p1f1) and reference data from Jung et al. (2009). Values present monthly mean values averaged over the period 1982 to 2008 and the shading shows the corresponding standard deviation.”

17. L615-616: “Negative annual mean SSS biases occur under the region associated with excessive March sea-ice...” This seems counter to what I would expect. Generally with sea-ice formation there is brine rejection into the surface water, hence with excessive sea ice I would expect excessive brine rejection and a positive bias in SSS.

We have deleted “under the region associated with excessive March sea-ice”. Our statement was based on the visual coherence between the patterns of sea-ice bias, with SSS (and SST). Mechanistically, negative SSS biases could be expected if sea-ice formed elsewhere is being advected into this region where it subsequently melts.

18. L676-678: Please provide references to PIOMAS and GIOMAS.

We inserted in-text references to Zhang and Rothrock (2003) and Schweiger et al., 2011. These are also provided in Table F2.

19. Figure 20: The excessive model sea-ice in the Labrador sea, does this have an impact on deep water formation and potentially help explain the weak (~12Sv) AMOC in CanESM5?

Yes, we expect that it does. The AMOC is stronger in ocean only simulations which do not have this ice cover over the Labrador sea, although, there are also other differences between the coupled and ocean only model (no coupled feedbacks), so this is not definitive.

20. Figure 20 needs a clearer figure caption and explanation. It is not clear whether the plot shows ocean to atmosphere or atmosphere to ocean C flux. I am guessing positive values in the North Atlantic indicate ingassing (e.g. atmosphere to ocean flux of C).The implied direction of the flux with respect to positive and negative values needs to be made clearer. Also figure b looks like the difference in flux CanESM5 minus Landschutzer, the figure caption suggests it is an absolute value of Landschutzer.

We assume here that the reviewer is referring to Fig. 22, which shows the atmosphere-ocean CO2 flux (Fig 20 is sea-ice). We have inserted a clarification on the flux direction into the Fig. 22 caption:

"The flux is positive down (into the ocean)"

and we have clarified that panel b) is the bias, not absolute value of Landschutzer:

"(b) the bias relative to Landschutzer (2009)"

21. L773-774: Why use this definition of the Northern Annular Mode? Is this a standard circulation measure? Would it not be better to plot the NAO for the Northern hemisphere?

We use the standard definition of the Northern Annular Mode, the first EOF of sea level pressure north of 20 degrees north, for example as given in Thompson and Wallace (1999) Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability, J. Climate, 13, and also as at

https://climatedataguide.ucar.edu/climate-data/hurrell-wintertime-slp-based-northern-annular-mode-nam-index

22. Figure 24 (caption): "The colour scale is arbitrary" What does this mean? Presumably the scales are the same across the 4 figures?

The scale is the same across figures. We have removed this statement from the caption.

23. L813-816: The TCR and ECS both increase (by~16% and 50%) in CanESM5relative to CanESM, yet the TCRE decreases by~16%. This is unusual. It is

explained as being due to a “significantly larger uptake of CO2 by the land biosphere in CanESM5relative to CanESM2”. Given the stated importance of TCRE for policymakers, are the authors confident that the increased CO2 uptake is due to realistic and understood reasons?

The land carbon uptake in Earth system models remains highly uncertain as has been noted in Friedlingstein et al. (2006, https://journals.ametsoc.org/doi/10.1175/JCLI3800.1) and Arora et al. (2013, (https://journals.ametsoc.org/doi/full/10.1175/JCLI-D-12-00494.1). As mentioned above, in reply to comment # 6, the tunable parameter determining the strength of this downregulation, and therefore the strength of the CO2 fertilization effect, is higher in CanESM5 than in CanESM2 resulting in higher land carbon uptake in CanESM5. The tuning of this downregulation parameter value, used in CanESM5, is explained in Arora and Scinocca (2016, https://www.geosci-model-dev.net/9/2357/2016/) and uses the land carbon uptake and other aspects of the carbon cycle over the historical period to obtain the best parameter value. Yet, of course, realistic model performance over the historical period is no guarantee of realistic model performance for future periods. Note that modelled historical land carbon uptake in CanESM5 is lower than observation-based estimates from Le Quere et al. (2018) yet its land carbon uptake in the 1pctCO2 simulation is highest amongst all CMIP6 models (manuscript currently in preparation by Vivek Arora). This requires a thorough investigation of CanESM5 results to find reasons behind this apparent conundrum. We suspect that forcing differences between the historical and 1pctCO2 experiments (aerosols in particular), influence rainfall changes, and thereby land carbon uptake. So yes, indeed, the reasons for lower TCRE despite higher TCR and ECS are well understood.

The text in section 6.1 which discusses TCRE is modified as follows:

“The reduction in TCRE occurs despite the fact that CanESM5 has a larger temperature response (TCR) than CanESM2 due to significantly larger uptake of CO_2 by the land biosphere in CanESM5 relative to CanESM2 in the 1pctCO2 experiment. As mentioned in Section 2.2, this is due to higher strength of the CO_2 fertilization effect in CanESM5 relative to CanESM2. As shown in Arora and Scinocca (2016) this leads to land carbon uptake in the 1pctCO2 simulation that is higher than in all CMIP5 models compared in Arora et al. (2013).”

and in Section 6.2.3:

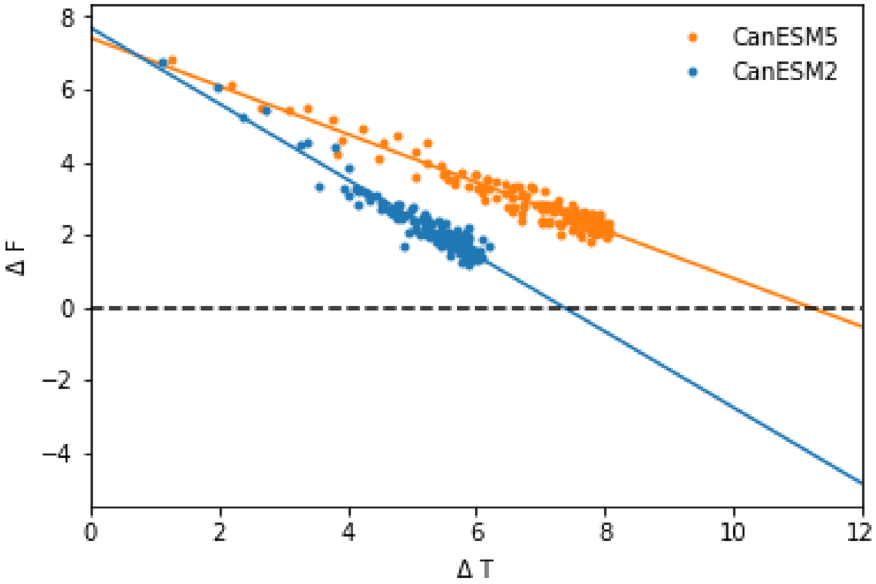
“In contrast, the simulated mean atmosphere-land CO_2 fluxes (Fig. 28c) for the decades of 1960s through to 2000s are lower than their observation-based estimates from Le Quéré et al. (2018). This is despite the fact that the land carbon uptake in the 1pctCO2 simulation for CanESM5 is highest amongst all CMIP5 models reported in Arora et al. (2013). The reason for this conundrum is a topic for future investigation, but might relate to differences in forcing (aerosols) in the historical and 1pctCO2 experiments.”

24. L821-825: “ECS IS 5.7K”: Actually, in figure 4 the global mean surface temperature does not look to have sufficiently stabilized by year 140 to derive an ECS value (e.g.it increase by~0.75C over years 90 to 150). Have the authors tried plotting the so-called Gregory linear regression of nett TOA radiation versus surface temperature to see whether there is a single or multiple linear relationships between these two, or whether a non-linear behaviour is evident?

The ECS value was calculated using the Gregory linear regression approach, not by looking at Figure 4. In the text we state “Given the long equilibration time of the climate system, it is common to estimate ECS from the relationship between surface temperature change and radiative forcing, over the course of the first 140 years of the abrupt-4xCO2 simulation (Gregory et al., 2004).”. To clarify that the number we report was calculated in this way, we have modified the sentence referenced by the reviewer to read:

“Here the ECS is calculated using the Gregory (2004) regression method, after removing linear drift from the piControl following Forster et al. (2013)”.

The plot of TOA radiation vs surface temperature in CanESM2 and CanESM5 over the first 140 years of the abrupt-4xCO2 experiment, together with the regression lines is included below for reference. We note a slight refinement to the CanESM5 ECS value (5.6) in the revision, relative to the original (5.7), owing to the removal of the linear trend from the piControl.



25. Figure 26c: It is difficult to see any differences. This might be better plotted as a model to observation difference or ratio.

We show multiple observations, so a simple difference or ratio is not as intuitive. We agree that differences in low latitudes are harder to see (overlapping lines) as they are small. However where differences are large, such as over the Arctic, they can easily be seen. We choose to leave this figure as it is.

Review comments in black.

Responses in green.

Review of The Canadian Earth System Model version 5 (CanESM5.0.3) by Neil et al. for GMD.

The paper describes the last version of CanESM. The main goal is to provide a reference for people who will analyse CMIP6 model outputs. Like all model description paper, there is a conflict between the need to be as comprehensive as possible, while keeping a reasonable paper length. In this point of view, the paper achieves a good and relevant compromise. It includes a short description of the model components, with all references for the reader who wants to go further in details. It contains a classical choice of model diagnostic to evaluate the model climate. It contains also informations, of the model sensitivity to standardised scenarios of CO2 increase. This content nicely matches the main objective of the paper.

I really appreciated the honesty of the authors, particularly when they described the consequence of the bug in p1 version, corrected in the p2 version. The paper also includes information about model quality control and performance. This is not of interest for CMIP6 data users. But model developers will find useful information on the way other teams work. Information that is hardly presented elsewhere. The general organisation of the paper is good. As non-native English speaker, I won't comment the quality of the syntax. I found very few typos, and thanks the authors for their careful proof reading. The main weakness of the paper is that the model is assessed in comparison with data and with the previous version of the model. There is no comparison with the CMIP5 models. This is not a major concern, as the model outputs will be used during CMIP6, and CamESM will be compared to CMIP5 and CMIP6 model in the next months by other authors. To sum up, this is a very good paper that perfectly fits the reason why GMD has been created for. I think it can be published after a few technical corrections.

We thank the reviewer for their constructive comments.

Major concern

Line 351. I do not understand how "version control, run isolation, strict checking and logging" can insure that the climate is reproducible. I agree that up to now, nobody has observed that lack of bit identical reproducibility in an ESM can drive to a different climate. But we have also theories of deterministic chaos showing that this is possible.

The system described enables us to ensure that we can re-run precisely the same code, in the same way as an original run. On a given machine, the results are bit identical. In some cases, moving across machines (e.g. our current migration from Cray XC40 to Cray XC50) also allows us to maintain bit identity. Hence these runs are precisely numerically reproducible.

Migrations to a different architecture or compiler might result in a bit pattern change. In this case our expectation (and experience) is that the climate will remain the same, but the realization of internal variability will be different. In this sense the run is reproducible in that we are interested in climate not weather. We accept that it is theoretically possible that mulitple-equilibria exist within the model in general. However, within fully coupled modelling this has essentially never been observed for a modern-day like climate, as the reviewer states. The only mulitple-equilibria we know of in complex GCMs involve radical mean state / forcing changes such as under extensive glaciation. There is no evidence we know of that bit-scale changes can lead to a different climate state.

Indeed, we note that different initial condition realizations of the model - with a similar type of small perturbation - result in the same climate. The concept that there are infinitely many possible realizations with the same climate (rather than different equilibrium climates) is also widely employed in coupled modelling and international exercises such as CMIP, which make extensive use of initial condition ensembles. Thousands of such initial condition simulations have been conducted, without the appearance of different equilibrium climates, as far as we know. It is impossible to ever prove that bit-induced multiple equilibria do not exist - but the extensive number of previous simulations are evidence that bit-induced multiple equilibria are exceedingly unlikely. For these reasons we are confident that the CanESM5 climate is robust to bit-pattern changes.

Minor concerns

Line 125 and following. Is there some specific representation of the urbanised areas ?

There is no explicit model for urban areas. Urban areas are represented/parameterized by higher albedo for visible light and lower albedo for near infrared radiation than for natural vegetation. The roughness length over urban areas is higher than that for crops and grasslands but lower than for trees.

Line 151 and following. The melt water of the glaciers goes to the runoff scheme. How did you design a ’river’ routing scheme for the ice sheets ? (from slopes ?)

There is not a dedicated runoff scheme for ice sheets. When the ice melts, the liquid meltwater is treated like runoff and the river routing scheme carries it down stream to the nearest ocean grid cell. This is do-able since river flow directions are specified over glacial cells as well based on their topography. Please see Figure 1 of Arora and Boer (1999).

Line 162. For the NEMO TKE scheme, a better reference than Gaspar et al. 1990 could be Blanke, B. and Delecluse, P.: Variability of the Tropical Atlantic Ocean simulated by a general circulation model with two mixed layer physics, Journal of Physical Oceanography, 23, 1363–1388, 1993.

This reference has been inserted.

Line 218. What is the computing coast of CMOC and CanOE compared to NEMO dynamics and to LIM2 ? It is significant or not ?

Yes it is significant. Turning on CMOC reduces model throughput by a factor of 2, relative to the physical model. Turning on CanOE reduces throughput by a factor of 4 relative to the physical model.

Line 777. Bentson et al. 2013 is not in the bibliography. Probably a typo, for Bentsen.

Corrected.

Line 809. Mathews et al. 2009 is not in the bibliography. Probably a typo, for Matthews.

Corrected.

Line 840. What is "Global Mean Screen Temperature" ? GMST generally stands for Global Mean Surface Temperature.

Changed “Screen” to “Surface”.

Line 1335. Year of the paper is embedded in the URL of the DOI.

Fixed.

The Canadian Earth System Model version 5 (CanESM5.0.3)

Formatted: Left

Neil C. Swart[1,3], Jason N.S. Cole[1], Viatcheslav V. Kharin[1], Mike Lazare[1], John F. Scinocca[1], Nathan P. Gillett[1], James Anstey[1], Vivek Arora[1], James R. Christian[1,2], Sarah Hanna[1], Yanjun Jiao[1], Warren G. Lee[1], Fouad Majaess[1], Oleg A. Saenko[1], Christian Seiler[4], Clint Seinen[1], Andrew Shao[3], Larry Solheim[1], Knut von Salzen[1,3], Duo Yang[1], Barbara Winter[1]

[1]Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, BC, V8W 2P2, Canada

[2]Fisheries and Oceans Canada, Institute of Ocean Sciences, Sidney, BC, Canada.

[3]University of Victoria, 3800 Finnerty Rd, Victoria, BC, V8P 5C2, Canada.

[4]Climate Processes Section, Environment and Climate Change Canada, Victoria, BC, V8P 5C2, Canada.

Correspondence to: Neil C. Swart (neil.swart@canada.ca)

Abstract. The Canadian Earth System Model version 5 (CanESM5) is a global model developed to simulate historical climate change and variability, to make centennial scale projections of future climate, and to produce initialized seasonal and decadal predictions. This paper describes the model components and their coupling, as well as various aspects of model development, including tuning, optimization and a reproducibility strategy. We also document the stability of the model using a long control simulation, quantify the model's ability to reproduce large scale features of the historical climate, and evaluate the response of the model to external forcing. CanESM5 is comprised of three dimensional atmosphere (T63 spectral resolution / 2.8°) and ocean (nominally 1°) general circulation models, a sea ice model, a land surface scheme, and explicit land and ocean carbon cycle models. The model features relatively coarse resolution and high throughput, which facilitates the production of large ensembles. CanESM5 has a notably higher equilibrium climate sensitivity (5.~~7~~6 K) than its predecessor CanESM2 (3.~~8~~7 K), which we briefly discuss, along with simulated changes over the historical period. CanESM5 simulations are contributing to the Coupled Model Intercomparison Project Phase 6 (CMIP6), and will be employed for climate science and service applications in Canada.

1 Introduction

A multitude of evidence shows that human influence is driving accelerating changes in the climate system, which are unprecedented in millennia (IPCC, 2013). As the impacts of climate change are increasingly being felt, so is the urgency to take action based on reliable scientific information (UNFCCC, 2015). To this end, the Canadian Centre for Climate Modelling and Analysis (CCCma) is engaged in an ongoing effort to improve modelling of the global ~~earth~~Earth system, with the aim of enhancing our understanding of climate system function, variability and historical changes, and for making improved quantitative predictions and projections of future climate. The global coupled model, CanESM, forms the basis of the CCCma modelling system, ~~which also includes~~and shares components with the Canadian Regional Climate Model (CanRCM) for finer scale modelling of the atmosphere (Scinocca et al., 2016), the Canadian Middle Atmosphere Model (CMAM) with atmospheric chemistry (Scinocca et al., 2008), and the Canadian Seasonal to Interseasonal Prediction System which is used for seasonal prediction and decadal forecasts (CanSIPS, Merryfield et al., 2013).

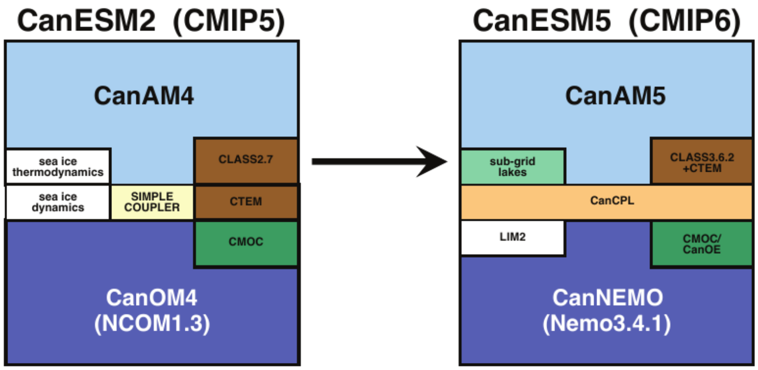
CanESM5 is the current version of CCCma's global model, and has a pedigree extending back 40 years to the introduction of the first atmospheric General Circulation Model (GCM) developed at CCCma's predecessor, the Canadian Climate Centre (Boer and McFarlane, 1979; Boer et al., 1984; McFarlane, et al., 1992). Successive versions of the model introduced a dynamic three dimensional ocean in CGCM1 (Flato et al., 2000; Boer et al. 2000a; Boer et al. 2000b), and later an interactive carbon cycle was included to form CanESM1 (Arora et al, 2009; Christian et al., 2010). The last major iteration of the model, CanESM2 (Arora et al, 2011), was used in the Coupled Model Intercomparison Project phase 5 (CMIP5), and continues to be employed for novel science applications such as generating large initial condition ensembles for detection and attribution (e.g. Kirchmeier-Young et al., 2017; Swart et al., 2018).

As detailed below, CanESM5 represents a major update to CanESM2. The leap from version 2 to version 5 was a one-off correction made to reconcile our internal model version labelling with the version label released to the public. The update includes incremental improvements to the atmosphere, land surface and terrestrial ecosystem models. The major changes relative to CanESM2 are the implementation of completely new models for the ocean, sea-ice, marine ecosystems, and a new coupler. Model developers have a choice in distributing increasing, but finite, computational resources between improvements in model resolution, model complexity and model throughput (i.e. number of years simulated). The resolution of CanESM5 (T63 or ~2.8° in the atmosphere and ~1° in the ocean) remains similar to CanESM2, and is at the lower end of the spectrum of CMIP6 models. The advantage of this coarse resolution is a relatively high model throughput given the complexity of the model, which enables many years of simulation to be achieved with available computational resources. The first major application of CanESM5 is CMIP6 (Eyring et al., 2016), and over 50,000 years of simulation are being conducted for the 20 CMIP6-endorsed MIPs in which CCCma is participating.

The aim of this paper is to provide a comprehensive reference that documents CanESM5. In the sections below, each of the model components is briefly described, and we also explain the approach used to develop, tune and numerically optimize the model. Following that, we document the stability of the model in a long pre-industrial control simulation, and the model's ability to reproduce large-scale features of the climate system. Finally, we investigate the sensitivity of the model to external forcings.

2 Component Models

In CanESM5 the atmosphere is represented by the Canadian Atmosphere Model (CanAM5), which incorporates the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM). The ocean is represented by a CCCma customized version of the Nucleus for European Modelling of the Ocean model (NEMO), with ocean biogeochemistry represented by either the Canadian Model of Ocean Carbon (CMOC) in the standard model version labelled as *CanESM5*, or the Canadian Ocean Ecosystem model (CanOE) in versions labelled *CanESM5-CanOE*. The atmosphere and ocean components are coupled by means of the Canadian Coupler (CanCPL). ~~These~~Each of these components of CanESM5 are~~ summarized schematically in Fig. 1, and~~ described further below.



~~**Figure 1: Model schematic, showing the evolution of components between CanESM2 and CanESM5.**~~

2.1 The Canadian Atmospheric Model version 5 (CanAM5)

Version 5 of the Canadian Atmospheric Model (CanAM5) ~~has several improvements relative to~~employs a spectral dynamical core with a hybrid sigma-pressure coordinate in the vertical. The package of physical parameterizations used by CanAM5 are

Formatted: Left

Formatted: Font color: Black

based on an updated version of its predecessor, CanAM4 (von Salzen et al., 2013~~),~~). The physics package includes a prognostic cloud microphysics scheme governing water vapour, cloud liquid water, and cloud ice; a statistical layer-cloud scheme; and independent cloud-base mass-flux schemes for both deep and shallow convection. Aerosols are parameterized using a prognostic scheme for bulk concentrations of natural and anthropogenic aerosols, including ~~changes to aerosol, clouds, radiation, land surface and lake processes. The model uses a T63~~sulfate, black and organic carbon, sea salt, and mineral dust; parameterizations for emissions, transport, gas-phase and aqueous-phase chemistry, and dry and wet deposition account for interactions with simulated meteorology. CanAM5 employs a triangular ~~spectral~~ truncation ~~in the model dynamical core, with~~ at total wavenumber 63 (T63) corresponding to an approximate ~~horizontal~~isotropic resolution of 2.8 degrees in both latitude~~/~~ and longitude. ~~It uses a hybrid~~In the vertical ~~coordinate system with,~~ 49 levels ~~between the surface and 1 hPa,~~are employed with ~~a vertical resolution of about~~layer thicknesses that increase monotonically from approximately 100 m ~~near~~at the surface~~.~~ ~~Relative~~ to 2km at ~1hPa – the domain lid.

Updates to the ~~35 levels used~~ package of physical parameterizations in ~~CanESM2 most of the additional 14 levels were added~~CanAM5 over those in CanAM4 are as follows. While the ~~upper troposphere and stratosphere. The~~ radiative transfer solution in CanAM5 is similar to that in CanAM4, the representation of ~~radiative processes was~~optical properties were improved through changes to the parameterization of albedos for bare soil, snow and ocean white-caps; cloud optics for ice clouds and polluted liquid clouds; improved aerosol optical properties~~:~~, and absorption by the water vapour continuum at solar wavelengths. For aerosols, the ~~emission~~parameterization for emissions of mineral dust and dimethyl sulfide (DMS) was improved while ~~for clouds~~the bulk stratiform cloud microphysical scheme was modified to include a parameterization of the second indirect effect~~ was activated in the stratiform cloud microphysics.~~

.

Parameterizations of surface processes were improved through an upgrade of the ~~land surface scheme~~ Canadian Land Surface Scheme (CLASS) from ~~CLASS~~version 2.7 to 3.6.2 as well as the inclusion of a parameterization for sub-grid lakes. CanESM5 represents the first coupled model produced by the CCCma in which the atmosphere and ocean do not employ coincident horizontal computational grids. As a consequence, CanAM5 was modified to support a fractional land mask, by generalizing its underlying surface to support grid-box fractional tiles of land and water. This tiling technology was extended to include land surface components of ocean, sea-ice and subgrid scale lakes. In this way, appropriate fluxes ~~are~~can be provided to each ~~model~~ component. A ~~more detailed~~full description ~~of~~ CanAM5 and its relation to CanAM4 will be provided in a companion paper in this special issue (Cole et al., 2019).

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Left

2.2 CLASS-CTEM

The CLASS-CTEM modelling framework consists of the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) which together form the land component of CanESM5. CLASS and CTEM simulate the physical and biogeochemical land surface processes, respectively, and together they calculate fluxes of energy, water, CO_2 and wetland CH_4 emissions at the land-atmosphere boundary. The introduction of dynamic wetlands and their methane emissions is a new biogeochemical process added since the CanESM2~~.~~ (but note these methane emissions are strictly diagnostic, since atmospheric methane concentrations are specified).

CLASS is described in detail in Verseghy (1991), Verseghy et al. (1993) and Verseghy (2000) and version 3.6.2 is used in CanESM5. It prognostically calculates the temperature for its soil layers, their liquid and frozen moisture contents, temperature of a single vegetation canopy layer if it is present as dictated by the specified land cover, and the snow water equivalent and temperature of a single snow layer if it is present. Three permeable soil layers are used with default thicknesses of 0.1, 0.25 and 3.75 m. The depth to bedrock is specified on the basis of the global data set of Zobler (1986) which reduces the thicknesses of the permeable soil layers. CLASS performs energy and water balance calculations and all physical land surface processes for four plant functional types (PFTs) (needleleaf trees, broadleaf trees, crops and grasses), and operates at the same sub-daily time step as the rest of the atmospheric component.

CTEM models photosynthesis, autotrophic respiration from its three living vegetation components (leaves, stem and roots) and heterotrophic respiration fluxes from its two dead carbon components (litter and soil carbon) and is described in detail in Arora (2003), Arora and Boer (2003) and Arora and Boer (2005). ~~Land use change is also modelled on the basis of specified time-varying land cover which incorporates the increase in crop area over the historical period following Arora and Boer (2010).~~ CTEM's photosynthesis module operates within CLASS, at the same time step as rest of the atmospheric component. CTEM provides CLASS with dynamically simulated structural attributes of vegetation including leaf area index (LAI), vegetation height, rooting depth and distribution, and above ground canopy mass~~.~~ which change in response to changes in climate and atmospheric CO_2 concentration. All terrestrial ecosystem processes other than photosynthesis are modelled in CTEM at a daily time step. Terrestrial ecosystem processes in CTEM are modelled for nine PFTs that map directly to the PFTs used by CLASS. Needleleaf trees are divided into their deciduous and evergreen types, broadleaf trees are divided into cold and drought deciduous and evergreen types, and crops and grasses are divided into C_3 and C_4 versions based on their photosynthetic pathways. The reason for separation of PFTs for CTEM is the additional distinction that biogeochemical processes require. For example, the distinction between deciduous and evergreen versions of needleleaf trees is needed to simulate leaf phenology prognostically. Once leaf area index has been dynamically determined by CTEM all CLASS needs to know is that this PFT is needleleaf tree since the physics calculations do not require information about underlying deciduous or evergreen nature of leaves. Similarly, the C_3 and C_4 photosynthetic pathways of crops and grasses determine how they photosynthesize thus affecting the calculated canopy resistance. However, once canopy resistance is known CLASS does not need to know the underlying distinction between C_3 and C_4 crops and grasses to use this canopy resistance in its energy and water balance calculations.

While, the modelled structural vegetation attributes respond to changes in climate and atmospheric CO_2 concentration, the fractional coverage of CTEM's nine PFTs is specified. A land cover data set is generated based on a potential vegetation cover for 1850 upon which the 1850 crop cover is superimposed. From 1850 onwards, as the fractional area of C_3 and C_4 crops changes the fractional coverages of the other non-crop PFTs are adjusted linearly in proportion to their existing coverage, as described in Arora and Boer (2010). The increase in crop area over the historical period is based on LUH2 v2h product (http://luh.umd.edu/data.shtml) of the land use harmonization (LUH) effort produced for CMIP6 (Hurtt et al., 2011).

Formatted: Left

Both CanESM5, and its predecessor CanESM2, do not include nutrient limitation of photosynthesis on land since the terrestrial nitrogen cycle is not represented. However, both models include a representation of terrestrial photosynthesis downregulation based on Arora et al. (2009) who used results from plants grown in ambient and elevated CO_2 environments to emulate the effect of nutrient constraints. The tunable parameter determining the strength of this downregulation, and therefore the strength of the CO_2 fertilization effect, is higher in CanESM5 than in CanESM2 resulting in higher land carbon uptake in CanESM5. The tuning of this downregulation parameter value, used in CanESM5, is explained in Arora and Scinocca (2016) who evaluate several aspects of modelled historical carbon cycle against observation-based estimates. A land nitrogen cycle model for CTEM is currently being developed which will make the photosynthesis downregulation parameterization obsolete in future versions of the model.

Formatted: Left

The calculation of wetland extent and methane emissions from wetlands is described in detail in Arora et al. (2018). In brief, dynamic wetland extent is based on the "flat" fraction in each grid cell with slopes less than 0.2%. As the liquid soil moisture in the top soil layer increases above a specified threshold, the wetland fraction increases linearly up to a maximum value, equal to the flat fraction in a grid cell. The simulated CH_4 emissions from wetlands are calculated by scaling the heterotrophic respiration flux from the model's litter and soil carbon pools to account for the ratio of wetland to upland heterotrophic respiratory flux and the fact that some of the CH_4 flux is oxidized in the soil column before reaching the atmosphere.

~~Specified land cover that includes fractional coverages of CTEM's nine PFTs is generated based on a potential vegetation cover for 1850 upon which the 1850 crop cover is superimposed. From 1850 onwards, as the fractional area of C_3 and C_4 crops changes the fractional coverages of the other non-crop PFTs are adjusted linearly in proportion to their existing coverage, as described in Arora and Boer (2010). The increase in crop area over the historical period is based on LUH2 v2h product (http://luh.umd.edu/data.shtml) of the land use harmonization (LUH) effort (Hurtt et al., 2011).~~

Surface runoff and baseflow simulated by CLASS are routed through river networks. Major river basins are discretized at the resolution of the model and river routing is performed at the model resolution using the variable velocity river routing scheme presented in Arora and Boer (1999). The delay in routing is caused by the time taken by runoff to travel over land in an assumed rectangular river channel and a ground water component to which baseflow contributes. Streamflow (i.e. the routed runoff) contributes fresh water to the ocean grid cell where the land fraction of a CanAM grid cell first drops below 0.5 along the river network as the river approaches the ocean.

In CanESM5, glacier coverage is specified and static. Grid cells are specified as glacier if the fraction of the grid cell covered by ice exceeds 40%, based on the GLC2000 dataset (Bartholomé and Belward, 2005). The combination of this threshold and the model resolution results in glacier covered cells predominantly representing the Antarctic and Greenland ice sheets, with a few glacier cells in the Himalayas, Northern Canada and Alaska. Snow can accumulate on glaciers, and any additional snow above the threshold of 100 kg m^{-2} of snow water equivalent is “converted into ice”~~,~~, and an equivalent mass of freshwater is immediately inserted into runoff – implicitly representing mass balance between accumulation and calving. Snow and ice on glaciers can be melted, with the water exceeding a ponding limit inserted into runoff. There is no explicit accounting for glacier mass balance, or adjustment of glacier coverage. This represents a potentially infinite global source or sink of fresh water in the coupled system, particularly in climates which are far from the state represented by GLC2000. However, in practice the timescales of our centennial-scale simulations are much shorter than the response times of ice sheet coverage, and any imbalances are small (Section 4).

2.3 NEMO modified for CanESM (CanNEMO)

The ocean component is based on NEMO version 3.4.1 (Madec et al. 2012). It is configured on the tripolar ORCA1 C-grid with 45 z-coordinate vertical levels, varying in thickness from ~6 m near the surface to ~250 m in the abyssal ocean. Bathymetry is represented with partial cells. The horizontal resolution is based on a 1° Mercator grid, varying with the cosine of latitude, with a refinement of the meridional grid spacing to 1/3° near the Equator. The adopted model settings include the linear free surface formulation (see Madec et al. 2012 and references therein). Momentum and tracers are mixed vertically using a turbulent kinetic energy scheme based on the model of Gaspar et al. ~~(1990).~~(1990), implemented into NEMO physics by Blanke and Delecluse (1993). The tidally-driven mixing in the abyssal ocean is accounted for following Simmons et al. (2004). Base values of vertical diffusivity and viscosity are 0.5~~×~~$\times 10^{-5}$ and 1.5~~×~~$\times 10^{-4}$ m^2~~/~~s^{-1}, respectively. A parameterization of double diffusive mixing (Merryfield et al., 1999) is also included. Lateral viscosity is parameterized by a horizontal Laplacian operator with eddy viscosity coefficient of 1.0~~×~~$\times 10^4$ m^2~~/~~s^{-1} in the tropics, decreasing with latitude as the grid spacing decreases. Tracers are advected using the total variance dissipation scheme (Zalesak, 1979). Lateral mixing of tracers (Redi 1982) is parameterized by an isoneutral Laplacian operator with eddy diffusivity coefficient of 1~~.×~~.0 $\times 10^3$

m^2~~/~~s^{-1} at the Equator, which decreases poleward with the cosine of latitude. The process of potential energy extraction by baroclinic instability is represented with the Gent and McWilliams (1990) scheme using a spatially-variable formulation for the mesoscale eddy transfer coefficient, as briefly described below.

Two modifications have been introduced to the NEMO's mesoscale and small-scale mixing physics. The first modification is motivated by the observational evidence suggesting that away from the tropics the eddy scale decreases less rapidly than does the Rossby radius (e.g., Chelton et al., 2011). This is taken into consideration in the formulation for the eddy mixing length scale, which is used to compute the mesoscale eddy transfer coefficient for the Gent and McWilliams (1990) scheme (for details, see Saenko et al., 2018). The second modification is motivated by the observationally based estimates suggesting that a fraction of the mesoscale eddy energy could get scattered into high-wavenumber internal waves, the breaking of which results in enhanced diapycnal mixing (e.g., Marshall and Naveira Garabato, 2008; Sheen et al., 2014). A simple way to represent this process in an ocean general circulation model was proposed in Saenko et al. (2012). Here, we employ an updated version of their scheme which accounts better for the eddy-induced diapycnal mixing observed in the deep Southern Ocean (e.g., Sheen et al., 2014).

CanESM5 uses the LIM2 sea ice model (Fichefet and Morales Maqueda, 1997; Bouillon et al., 2009), which is run within the NEMO framework. Some details regarding the calculation of surface temperature over sea-ice are described in the coupling section below.

2.4 Ocean biogeochemistry

Two different ocean biogeochemical models, of differing complexity and expense, were developed in the NEMO framework: CMOC and CanOE. Two coupled models versions will be submitted to CMIP6. The version labelled as *CanESM5* uses CMOC and was used to run all the experiments that CCCma has committed to. The version labelled *CanESM5-CanOE*, described in another paper in this special issue (Christian et al., 2019), is identical to CanESM5, except that CMOC was replaced with CanOE, and this version has been used to run a subset of the CMIP6 experiments, including DECK and historical (see Section 3.4). Both biogeochemical models simulate ocean carbon chemistry and abiotic chemical processes such as oxygen solubility identically, in accordance with the OMIP-BGC protocol (Orr et al., 2017).

2.4.1 Canadian Model of Ocean Carbon (CMOC)

The Canadian Model of Ocean Carbon was developed for earlier versions of CanESM (Zahariev et al., 2008; Christian et al., 2010; Arora et al., 2011), and includes carbon chemistry and biology. The biological component is a simple Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) model, with fixed Redfield stoichiometry, and simple parameterizations of iron limitation, nitrogen fixation, and export flux of calcium carbonate. CMOC was migrated into the NEMO modelling system, and the following important modifications were made: i) oxygen was added as a passive tracer with no feedback on biology; ii) carbon chemistry routines were updated to conform to the OMIP-BGC protocol (Orr et al., 2017); iii) additional passive tracers requested by OMIP were added, including natural and abiotic DIC as well as the inert tracers CFC11, CFC12 and SF6.

2.4.2 Canadian Ocean Ecosystem Model (CanOE)

The Canadian Ocean Ecosystem Model (CanOE) is a new ocean biology model with a greater degree of complexity than CMOC, and represents explicitly some processes that were highly parameterized in CMOC. CanOE has two size classes for each of phytoplankton, zooplankton and detritus, with variable elemental (C/N/Fe) ratios in phytoplankton and fixed ratios for zooplankton and detritus. Each detritus pool has its own distinct sinking rate. In addition, there is an explicit detrital $CaCO_3$ variable, with its own sinking rate. Iron is explicitly modelled, with a dissolved iron state variable, sources from aeolian deposition and reducing sediments, and irreversible scavenging from the dissolved pool. N_2 fixation is parameterized similarly to CMOC with temperature- and irradiance-dependence and inhibition by Dissolved Inorganic Nitrogen, but no explicit N_2-fixer group. In addition, N_2 fixation is iron-limited in CanOE. In CanOE, denitrification is modelled prognostically and occurs only where dissolved oxygen is <6 mmol m^{-3}. Deposition of organic carbon is instantaneously remineralized at the sea floor as in CMOC, and $CaCO_3$ deposited at the sea floor dissolves if the calcite is undersaturated (whereas in CMOC the burial fraction is implicitly 100%). Carbon chemistry and all abiotic chemical processes such as oxygen solubility conform to the OMIP-BGC protocol (Orr et al., 2017) and are identical in CanOE and CMOC, except that in CMOC the carbon chemistry solver is applied only in the surface layer (as there is no feedback from saturation state to other biogeochemical processes in the subsurface layers). CanOE has roughly twice the computational expense of CMOC.

2.5 The Canadian Coupler (CanCPL)

CanCPL is a new coupler developed to facilitate communication between CanAM and CanNEMO. CanCPL depends on Earth System Modeling Framework (ESMF) library routines for regridding, time advancement, and other miscellaneous infrastructure (Theurich et al., 2016; Collins et al., 2005; Hill et al., 2004). It was designed for the Multiple Program Multiple Data (MPMD) execution mode, with communication between the model components and the coupler via the Message Passing Interface (MPI).

The fields passed between the model components are summarized in Tables A1 to A4. In general, CanNEMO passes instantaneous prognostic fields, which are remapped by CanCPL and given to CanAM as lower boundary conditions. These prognostic fields (sea surface temperature, sea-ice concentration and mass of sea-ice and snow) are held constant in CanAM over the course of the coupling cycle. After integrating forward for a coupling cycle, CanAM passes back fluxes, averaged over the coupling interval, which are remapped in CanCPL and passed on to NEMO as surface boundary conditions. An exception is the ocean surface CO_2 flux, which is computed in CanNEMO and passed to CanAM. CanAM and CanNEMO are run in parallel, and the timing of exchanges through the coupler is indicated schematically in Fig. ~~2~~A1.

All regridding in CanCPL is done using the ESMF first order conservative regridding option (ESMF, 2018), ensuring that global integrals remain constant for all quantities passed between component models (but see an exception below). The remapping weights w_{ij}, for a particular source cell i and destination cell j are given by: $w_{ij} = f_{ij} \times A_{si}/(A_{dj} \times D_j)$, where f_{ij} is the fraction of the source cell i contributing to the destination cell j, A_{si} and A_{dj} are the areas of the source and destination cells, and D_j is the fraction of the destination cell that intersects the unmasked source grid (ESMF, 2018).

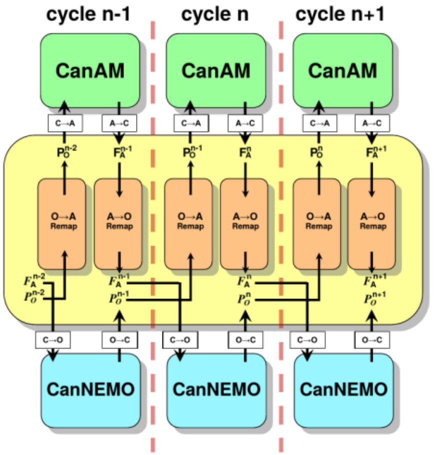


Figure 2: ~~Schematic showing the ordering of exchanges between CanCPL and CanAM and CanNEMO. Prognostic fields (P_O) are passed from NEMO to the coupler, remapped, and passed to CanAM. Fluxes (F_A) are passed from CanAM, remapped in CanCPL, and passed NEMO to complete the next coupling cycle. Superscripts denote the coupling cycle, e.g. Prognostic fields from NEMO are passed to CanCPL at the end of cycle "n", remapped, and used in CanAM during cycle "n+1".~~

Formatted: Left, Indent: Left: 0 cm

Within the NEMO coupling interface the "conservative" coupling option is employed. This option dictates that net fluxes are passed over the combined ocean-ice cell, and the fluxes over only the ice covered fraction of the cell are also supplied, in principle allowing net conservation, even if the distribution of ice has changed given the unavoidable one coupling cycle lag

Formatted: Left

encountered in parallel coupling mode. It was verified that the net heat fluxes passed from CanAM were identical to the net fluxes received by NEMO, to the level of machine precision. Conservation in the coupled model piControl run is discussed further in Section 4.

Sea-ice thermodynamics are computed in the LIM2 ice model, based on the surface fluxes received from CanAM, and the basal heat flux from the NEMO liquid ocean. LIM2 provides the sea-ice concentration, snow and ice thickness to CanAM, via the coupler. The surface flux calculation in CanAM5 requires the ground temperature at the snow/sea-ice interface, GT_{ice}. The GT_{ice} for this purpose can be passed from LIM2 to CanAM once each coupling cycle, or an alternative GT_{ice} can be evaluated in CanAM at every model time step, taking into account evolving surface albedo and atmospheric temperature (e.g. West et al., 2016). As implemented, when computing GT_{ice}, CanAM independently computes the conductive heat flux through sea-ice, and there is no constraint that this flux, or GT_{ice} is the same as that in LIM2. Conservation is maintained because the net heat flux between the atmosphere and sea-ice is computed in CanAM and applied to LIM2, but different ice surface temperatures could result. Both approaches to computing surface fluxes were tested in CanESM5 and no major impacts on sea-ice or the broader climate system relative to the default model were discovered. However, a significantly shorter coupling cycle of one hour was required for convergence when fluxes were computed from the LIM2 GT_{ice} passed through the coupler. The shorter coupling period was required to more physically resolve the response to diurnal variations in radiative and sensible heat fluxes from the atmosphere (see for example West et al., 2016). The evaluation of fluxes from GT_{ice} computed in CanAM, on the other hand, was stable for coupling periods ranging from 1 to 24 hours with no major changes in the mean climate, or variability immediately apparent. A final coupling cycle interval of three hours was implemented for CanESM5 with the computation of fluxes based on the CanAM evaluation of GT_{ice}. These choices represented improved robustness and a compromise between greater efficiency (i.e. longer coupling periods) and maximum "realism", which would be the one hour coupling dictated by the length of the NEMO time step.

After a significant number of CMIP6 production simulations were complete, it was determined that while conservative remapping was desirable for heat and water fluxes, it introduced issues in the wind-stress field passed from CanAM to CanNEMO. Specifically, since CanAM is nominally three times coarser than CanNEMO, conservative remapping resulted in constant wind-stress fields over several NEMO grid-cells, followed by an abrupt change at the edge of the next CanAM cell. This blockiness in the wind-stress results in a non-smooth first derivative, and the resulting peaked wind-stress curl results in unphysical features in, for example, the ocean vertical velocities. Changing regridding of only wind-stresses to the more typical "bilinear" interpolation, instead of "conservative" remapping, largely alleviates this issue. Sensitivity tests indicate no major impact on gross climate change characteristics such as transient climate response or equilibrium climate sensitivity, or on general features of the surface climate. However there is an impact on local ocean dynamics, which led to the decision to submit a "perturbed" physics member to CMIP6. Hence, simulations submitted to CMIP6 labelled as perturbed physics member 1 (“p1”) use conservative remapping for wind stress, while those labelled as “p2” use bilinear regridding (see Section 3.4). A comparison between p1 and p2 runs is provided in Appendix E.

2.6 Treatment of greenhouse gases

CanESM5 represents radiative forcing from individual greenhouse gases (GHGs). Aside from CO_2, the concentrations of all radiatively active gases are specified and transiently evolve. Of these, CH_4, N_2O, and families of CFCs are assumed to be well-mixed, while, O_3 is specified as varying spatially – typically employing that prescribed for CMIP6 (Checa-Garcia et al. 2018). CanAM5 offers two modes for modelling CO_2 concentrations - as specified time-evolving concentrations; or as a three-dimensional passive tracer driven by land/ocean surface emissions, prognostically derived through interactive coupling with biogeochemical carbon models in the land and ocean. For example, CanESM5 can be run with prognostic CO_2 in concert with specified anthropogenic fossil fuel emissions to simulate atmospheric CO_2 concentration through the historical and future periods. Wetland methane emissions simulated by CLASS-CTEM, in contrast, are purely diagnostic. While these emissions respond to changes in climate and atmospheric CO_2 concentration (through changes in vegetation productivity), they do not modify atmospheric CH_4 concentration, which are specified.

3 Model development and deployment

Formatted: Left

3.1 Model tuning and spin up

Each of the CanESM5 component models, CanAM5, CLASS-CTEM and CanNEMO, were initially developed independently under driving by observations in stand-alone configurations - CanAM5 in present-day (2003-2008) AMIP mode and CanNEMO in preindustrial (PI) OMIP-like mode using CORE bulk formulae. In these configurations, free parameters were initially adjusted to reduce climatological biases assessed via a range of diagnostics. Further details of the CanAM5 tuning may be found in Cole et al. (2019). The component models were then brought together in a preindustrial configuration (i.e. the piControl experiment), which was evaluated based on an array of diagnostics. Several thousand years of coupled simulation was run during the finalization of model, and an approach was taken whereby AMIP simulations would be used to derive parameter adjustments in CanAM, which would then be applied to the coupled model.

Initial present-day configurations of CanAM5 that were tuned to give roughly the observed top of the atmosphere net radiative forcing (TOA forcing ~0.7-1.0W m^{-2}) in an AMIP simulation produced coupled piControl simulations that were too cold (global mean near-surface temperatures below 12°C), with extensive sea-ice and a collapsing meridional overturning circulation. One contributor to the tendency of the new coupled model to cool was the inclusion of the thermodynamic consequences of snow melt in the open ocean, which induces an average global cooling of ~0.5 W m^{-2} in the piControl, and was not included in the previous version, CanESM2.

This initial coupled-model cold bias was rectified by adjusting free parameters in CanAM, CLASS and LIM2, in order to achieve a piControl simulation with a global mean screen temperature of around 13.7°C (roughly the absolute value provided for 1850-1900 by the NASA-GISS, Berkeley Earth and HadCRUT4 datasets), and a sea-ice volume within the spread of CMIP5 models~~.~~, while maintaining a net TOA radiative balance as close to 0 W m^{-2} as possible. The specific parameters adjusted were: emissivity of snow (from 1 to 0.97), snow grain size on sea-ice, the drainage parameter controlling soil moisture, the LIM2 parameter controlling the lead closure rate (from 2.0 to 3.0), and most significantly the accretion rate in cloud microphysics. The accretion rate exerted the largest control, and sensitivity to this parameter is described more fully in a companion paper (Cole et al., 2019).

The consequence of the adjustments in CanAM5 was an increase in the present day TOA forcing in AMIP mode from ~1 W/m^2 to ~2.5 W m^{-2}. Nonetheless, historical simulations of the coupled CanESM5 initialized from its equilibrated piControl show an increase in TOA forcing roughly matching the observed values of ~0.7-1.0 W m^{-2} over the 2003-2008 period for which CanAM5 was tuned in AMIP mode. The difference in patterns of SST and sea-ice concentrations between the coupled model and observations are thought to be the cause of these differences in TOA balance between coupled and AMIP mode.

The final adjustment was to the carbon uptake over land so as to better match the observed value over the historical period, and achieved via the parameter which controls the strength of the CO$_2$ fertilization effect (Arora and Scinocca, 2016). No more extensive tuning of CanESM5 was undertaken. Critically, no tuning of climate sensitivity was undertaken~~ on the climate system response to forcing~~ - the transient and equilibrium climate sensitivity of CanESM5 are purely emergent properties. Once the tuned final configuration of CanESM5 was available, ocean potential temperature and salinity fields were initialized from World Ocean Atlas 2009, while CanAM, CLASS-CTEM and CMOC were initialized from the restarts from earlier development runs. The model was spun up for over 1500 years prior to the launch of the official CMIP6 piControl simulation, which extends for a further 2000 years.

3.2 Code management, version control and reproducibility

CanESM5 is the first version of the model to be publicly released, and this code sharing has been facilitated by the adoption of a new version control based strategy for code management. Additional goals of this new system are to adopt industry standard software development practises, to improve development efficiency, and to make all CanESM5 CMIP6 simulations fully repeatable.

To maintain modularity, the code is organized such that each model component has a dedicated *git* repository for the version control of its source code (Table 1). A dedicated *super repository* tracks each of the components as *git submodules*. In this way, the *super repo*. keeps track of which specific versions of each component combine together to form a functional version of CanESM. A commit of the CanESM super repo., which is representable by an 8 character truncated SHA1 checksum, hence uniquely defines a version of the full CanESM source code. The model development process follows an industry standard workflow (Table B1). New model features are merged onto the *develop_canesm* branch, which reflects the ongoing development of the model. Specific model versions, such as that used for CMIP6, are given tags and issued DOIs for ease of reference. We use an internal deployment of *gitlab* to host the model code and associated issue trackers, and we mirror the code to the public, online code hosting platform at gitlab.com/cccma/canesm.

A dedicated ecosystem of software is used to configure, compile, run, and analyze CanESM simulations on ECCC's HPC (Table B2). Several measures are taken to ensure modularity and repeatability. The source code for each run is recursively cloned from gitlab and is fully self contained. A strict checking routine ensures that any code changes are committed to the version control system, and any run-specific configuration changes are captured in a dedicated configuration repository. A database records the SHA1 checksums of the particular model version and configuration used for every run, and these are included in CMIP6 NetCDF output for traceability. Input files for model initialization and forcing are also tracked for reproducibility (Table B1).

Our strategy of version control, run isolation, strict checking and logging ensures that simulations can be repeated in the future, and the same climate will be obtained (bit identical reproducibility is a further step and is dependant on machine architecture and compilers). The implementation of a clear branching workflow, and the uptake of modern tools such as issue trackers, and the gitlab online code-hosting application has improved both collaboration and management of the code. This new system also led to large, unexpected improvements in model performance for two major reasons. The first was democratization of the code – via the promotion of group ownership of the code. The second was the freedom to experiment across the full code base ensured by our isolated run setup (Table B2), which was not possible under the previous system of using a single installed library of code shared across many runs. The performance gains achieved are described in the following section.

Table 1: Code structure and repositories.

Repository	**Purpose**
CanESM	The top-level super-repository, which tracks specific versions of the component *submodules*

Formatted Table

	listed below, to form a function version of the model. Also contains a CONFIG directory with configuration files for the model.
CanAM	The source code for the spectral dynamics and physics of CanAM.
CanDIAG	Diagnostic source code for analyzing CanAM output, this repository also contains various scripting used to run the model.
CanNEMO	The CCCma modified NEMO source code, along with additional utility scripting.
CanCPL	The coupler source code.
CCCma_tools	A collection of software tools for compiling, running and diagnosing CanESM on ECCC's high performance computer.

Formatted: Left

3.3 Model optimization and benchmarking

The ECCC high performance computer system consists of the following components: a "backend" Cray XC40, with two 18 core Broadwell CPUs per node (for 36 cores per node), and roughly 800 nodes in total, connected to a multi-PB lustre file system used as scratch space. This machine is networked to a "frontend" Cray CS5000, with several PB of attached HPFS spinning disk. This whole compute arrangement is replicated in a separate hall for redundancy, effectively doubling the available resources. Finally, a large tape-storage system (HPNLS) is available for archiving model results.

The initial implementation of a CanESM5 precursor on this new HPC occurred around Nov 1, 2017. The original workflow roughly followed that used for CanESM2 CMIP5 simulations. All CanESM5 components (atmosphere CanAM, coupler CanCPL and ocean CanNEMO) were originally running at 64-bit precision. The atmospheric component CanAM was running on two 36-core compute nodes, the coupler was running on a separate node, and the ocean component was running on 3 nodes, resulting in 6 nodes in total. The initial throughout on the system, without queue time, was around 4.6 years of simulation per wall-clock day (ypd), or alternatively 0.02 simulation years per core-day, when normalizing by the number of cores used.

In parallel to the physical model development, significant effort was made to improve the model throughput and eliminate a number of inefficiencies in the older CMIP5 workflow (Fig. ~~3~~1). The largest effort was devoted to improving the efficiency of CanAM5, since this was identified as the major bottleneck. A brief summary of the improvements is given in Table C1 and Fig. ~~3~~1. The most substantial and rewarding change was in converting the 64-bit CanAM component to 32-bit numerics. Since the remaining two components, CanCPL and CanNEMO are still running at the 64-bit precision, the communication between CanAM and CanCPL required the promotion of a number of variables from 32-bit precision to 64-bit and back. The 32-bit CanAM implementation required a number of modifications to maintain the numerical stability of the code.

Calculations in some subroutines, most notably in the radiation code, were promoted to the 64-bit accuracy. Conservation of some tracers, in particular CO$_2$, was compromised at the 32-bit precision, and some additional code changes to conserve CO$_2$ and maintain carbon budgets were implemented. Significant effort was also invested in optimizing compiler options used for NEMO to maximize efficiency, while the scalability of the NEMO code allowed sensibly increasing the node count to keep pace with the accelerated 32-bit version of CanAM.

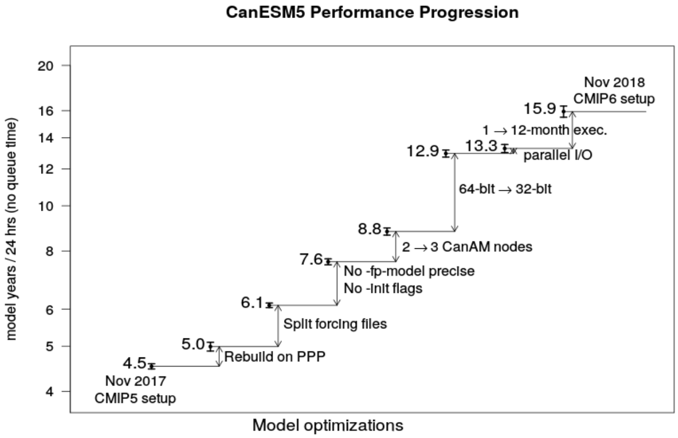


Figure ~~3~~1: Schematic of CanESM5 optimization. See Section 3.3 and Appendix C for details.

In the final setup, the CanAM/CanCPL components are running on three shared compute nodes, and the ocean component CanNEMO is running on 5 nodes, resulting in 8 nodes overall. The combined effect of the improvements listed in Table C1 resulted in more than tripling the original throughput to about 16 ypd (Fig. ~~3~~1). Despite the increase in the total node count from 6 to 8, the efficiency of the model also improved roughly three fold, from 0.02 simulation years per core day of compute to about 0.06 years per core day. This final model configuration can complete a realization of the 165 year CMIP6 historical experiment in just over 10 days, compared to about 36 days had no optimization been undertaken. At the time of writing, over 50,000 years of CMIP6 related simulation had been conducted with CanESM5, consuming about one million core-days of compute time, resulting in about 8 PB of data archived to tape, and over 100 TB of data publicly served on the Earth System Grid Federation (ESGF).

Formatted: Left

3.4 Model experiments and scientific application

This section describes the major experiments and model variants of CanESM5 that are being conducted for the Coupled Model Intercomparison Phase 6 (CMIP6), the first major science application of the model. Fig. ~~4~~2 shows the global mean surface temperature for several of the key CMIP6 experiments, which can be used to infer important properties of the model, as discussed further in Sections 4 to 6. Table 2 lists the variants of CanESM5 which are being submitted to CMIP6. These include the “p1” and “p2” perturbed physics members of CanESM5 (see Section 2.5), and a version of the model with a different ocean biogeochemistry model, CanESM5-CanOE.

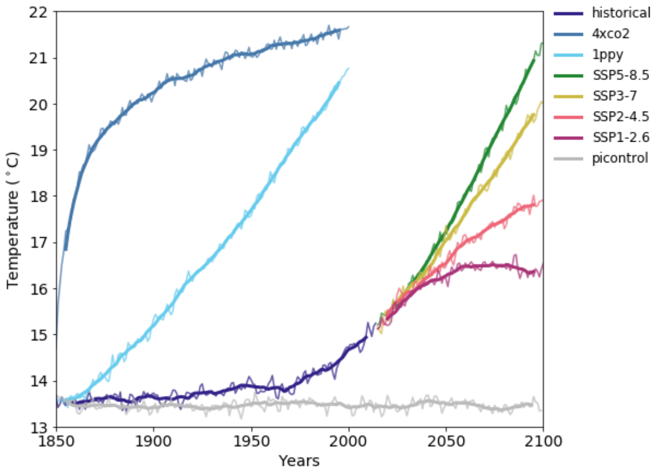


Figure ~~4~~2: Global average screen temperature in CanESM5 for the CMIP6 DECK experiments, as well as the historical and tier 1 SSP experiments (SSP5-85, SSP3-70, SSP2-45 and SSP1-26). Thick lines are the 11 year running means, thin lines are annual means.

Formatted: Left, Indent: Left: 0 cm

Formatted: Left

Table D1 lists the 20 CMIP6 endorsed MIPs in which CanESM5 is participating, and which model variants are being run for each MIP. The volume of simulation continues to grow, and will likely exceed 60,000 years. This is significantly more than the ~40,000 years of CMIP6 simulation estimated by Eyring et al. (2016). The major reason for this is that significantly larger ensembles have been produced than formally requested. For example, CanESM5 will submit at least 25 realizations for the historical and tier 1 SSP experiments, for each the “p1” and “p2” model variant, for a total of 50 realizations, significantly more than the single requested realization. The scientific value of such large initial condition ensembles has become evident (e.g. Kay et al., 2015; Kirchmeier-Young et al., 2017; Swart et al., 2018) and motivates this approach.

Individual historical realizations (ensemble members) of CanESM5 were generated by launching historical runs at 50 year intervals off the piControl simulation. This is the same as the approach used to generate the five realizations of CanESM2, which were submitted to CMIP5. The fifty year separation was chosen to allow for differences in multi-decadal ocean variability between realizations. Below we discuss the properties of the model, including illustrations of the internal variability generated spread across the historical ensemble. All results below are based on the CanESM5 p1 model variant.

Table 2: Model variants

Model variant	**Description**
CanESM5 “p1”	CanESM5 realizations labelled as perturbed physics member 1 (“p1” in the variant label) have conservative remapping of wind-stress fields. The ocean biogeochemistry model is CMOC.
CanESM5 “p2”	CanESM5 realizations labelled as perturbed physics member 2 (“p2” in the variant label), use bilinear remapping of the wind-stress fields. A minor land-fraction change also occurs over Antarctica. The ocean biogeochemistry model is CMOC.
CanESM5-CanOE “p2”	CanESM5-CanOE is exactly the same physical model as CanESM5, but it uses the CanOE ocean biogeochemical model. All CanESM5-CanOE realizations use bilinear remapping of the wind-stress, and hence are labelled as perturbed physics member 2 (“p2” in the variant label). No “p1” variant is submitted. For physical climate purposes CanESM5 and CanESM5-CanOE ~~may be treated as different realizations of the same model.~~are the same model, and simulations with specified CO_2 are bit identical in all realms besides ocean biogeochemistry. In runs with prognostic CO_2 (such as the esm-hist experiment), the physical climate of CanESM5 and CanESM5-CanOE will differ due to the effect of interactive CO_2 in these simulations.

Formatted Table

Formatted: Left

4 Stability of the pre-industrial control climate

The characteristics and stability of the CanESM5 pre-industrial control climate are evaluated using 1000 years of simulation from the CMIP6 piControl experiment, conducted under constant specified greenhouse gas concentrations and forcings for the year 1850 (Eyring et al., 2016). Ideally, a climate model and all its subcomponents would exhibit perfect conservation of tracer mass (e.g. water, carbon), energy and momentum, and would be run for long enough to achieve equilibrium. In this case we would expect to see, on long term average, zero net fluxes of heat, freshwater and carbon at the interface between the atmosphere, ocean and land surface, zero top of atmosphere net radiation, and constant long-term average temperatures or tracer mass within each component. In reality however models are not perfectly conservative due to the limitations of numerical representation (i.e. machine precision) as well as possible design flaws or bugs in the code, and models are generally not run to perfect equilibrium due to computational constraints. Despite imperfect conservation or spin up, models can still usefully be applied, as long as the drifts in the control run are small relative to the signal of interest, in our case historical anthropogenic climate. Below we consider conservation and drift of heat, water and carbon in CanESM5 (Fig. ~~5~~3).

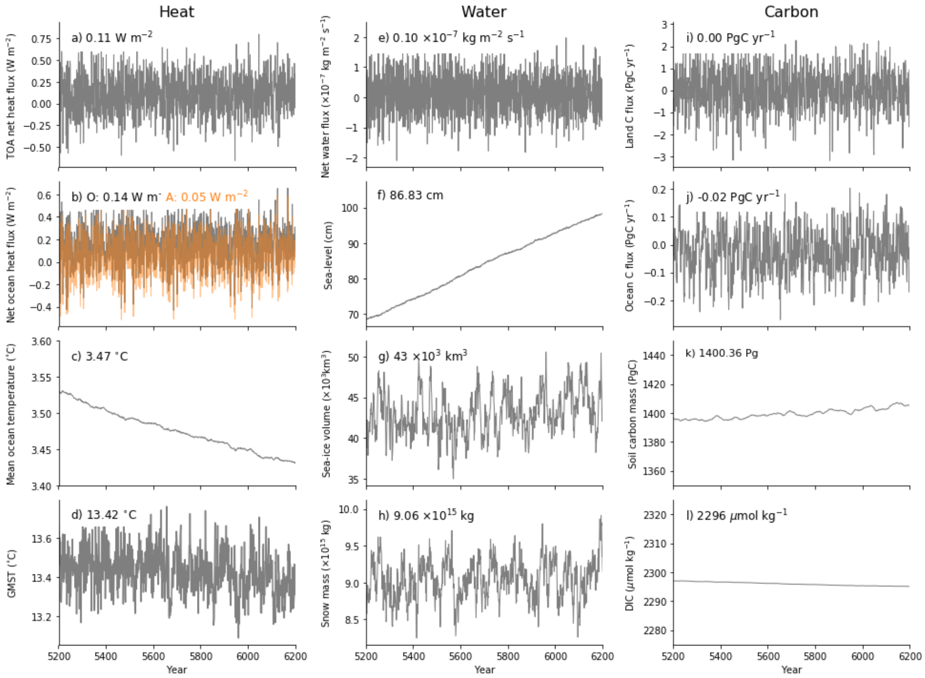


Figure ~~5~~3: Stability of the CanESM5 piControl run, showing global mean (a) top of atmosphere net heat flux, (b) net heat flux at the surface of ocean; (c) volume averaged ocean temperature; (d) ~~global mean~~ screen temperature; (e) net freshwater input at the liquid ocean surface; (f) dynamic sea level; (g) ~~global~~ sea-ice volume; (h) ~~global~~ snow mass; (i) land-atmosphere carbon flux; j) ocean-atmosphere carbon flux; k) terrestrial soil carbon mass and l) ocean dissolved inorganic carbon concentration. Inset numbers are the time-average over the 1000 years shown. Heat fluxes in (a) and (b) are reported per metre squared of global area. The orange line in (b) is the heat flux computed at the bottom of the atmosphere, while the grey line is the heat flux computed at the surface of the liquid ocean (below sea-ice).

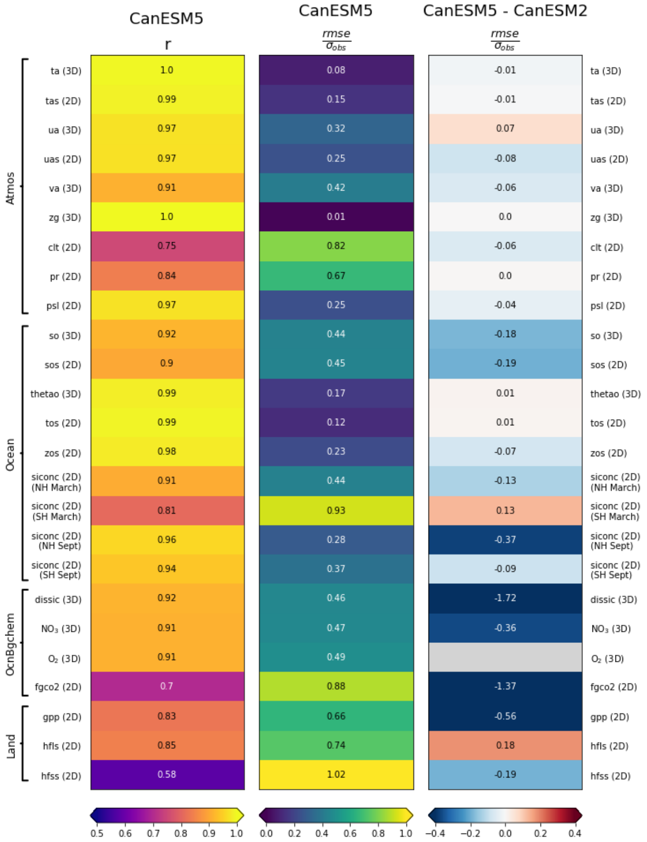
The CanESM5 pre-industrial control shows a stable Top of Atmosphere (TOA) net heat flux of 0.1 Wm^{-2} (fluxes positive down in m^2 of global area, Fig. ~~5a~~3a). The model is close to radiative equilibrium and this control net TOA heat flux is over an order of magnitude smaller than the signal expected from historical anthropogenic forcing (>1 Wm^{-2}). The global mean screen temperature is stable at around 13.4°C (Fig. ~~5d~~3d), indicating thermal equilibrium, and approximately in line with estimates of the temperature in 1850. Half of the net TOA flux is passed from the atmosphere to the ocean (0.05 Wm^{-2}, Fig. ~~5b~~3b). With the conservative remapping in the coupler, the fluxes exchanged between components are identical to machine precision. However, the net heat flux received at the surface of the liquid ocean is 0.14 Wm^{-2}, almost three times higher than the heat flux passed from CanAM to NEMO (Fig. ~~5b~~3b). This discrepancy reflects a non-conservation of heat within the LIM2 ice model. Tests with an ice-free ocean do not suffer this problem. Nonetheless, the discrepancy is relatively small, and ice volume is stable. A further non-conservation occurs within the NEMO liquid ocean. Although the ocean receives a net heat flux of 0.14 Wm^{-2}, the volume averaged ocean is cooling at a rate equivalent to a flux of 0.05 Wm^{-2} (Fig. ~~5c~~3c) implying a total non-conservation of heat in the liquid ocean of about 0.2 Wm^{-2}. Conservation errors of this order are well known in NEMO v3.4.1, likely arise from the use of the linear free surface (Madec et al., 2012), and have been seen in previous coupled models using NEMO (Hewitt et al., 2011). Despite this, the volume averaged ocean temperature drift in CanESM5 is about half the size of the drift in CanESM2. Furthermore the lack of ocean heat conservation in CanESM5 is roughly constant in time, and appears to be independent of the climate (not shown).

At the liquid ocean surface, a small net freshwater flux results in a freshening trend, and a sea-level rise of about 24 cm over 1,000 years (Fig. ~~5e~~3e, f). This rate of drift is more than 20 times smaller than the signal of anthropogenic sea-level rise. The LIM2 ice model appears to be the source of non-conservation: the net freshwater flux provided from CanAM is very close to zero, about six times smaller than that noted above (24 cm / 1000 years). Snow and ice volume are stable, not exhibiting any long term drift, yet they are subject to considerable decadal and centennial scale variability (Fig. ~~5g~~3g, h).

Atmosphere-land carbon fluxes average to zero, and carbon pools within CTEM are stable (Fig. ~~5i~~3i, k). The net ocean carbon flux is fairly close to zero, but remains slightly negative on average at -0.02 Pg yr^{-1} despite a multi-millennial spin up (Fig. ~~5j~~3j). The total mass of dissolved inorganic carbon in the ocean decreases very slightly as a result (Fig. ~~5l~~3l). The rate of ocean carbon drift is approximately an order of magnitude smaller than the modern day anthropogenic signal of ocean carbon uptake (>2 Pg yr^{-1}). The drifts identified above are all far smaller than would be expected from anthropogenically forced trends, confirming that the model is suitably stable to evaluate centennial scale climate change. In the following section, we consider the ability of the model to reproduce large scale features of the observed historical climate.

5 Evaluation of historical mean climate

In this section we use the CMIP6 historical simulations (Eyring et al., 2016) of CanESM5 “p1”, focusing on climatologies computed over 1981 to 2010, unless otherwise noted.



		CanESM5 r	CanESM5 $\frac{rmse}{\sigma_{obs}}$	CanESM5 - CanESM2 $\frac{rmse}{\sigma_{obs}}$
Atmos	ta (3D)	1.0	0.08	-0.01
	tas (2D)	0.99	0.15	-0.01
	ua (3D)	0.97	0.32	0.07
	uas (2D)	0.97	0.25	-0.08
	va (3D)	0.91	0.42	-0.06
	zg (3D)	1.0	0.01	0.0
	clt (2D)	0.75	0.82	-0.06
	pr (2D)	0.84	0.67	0.0
	psl (2D)	0.97	0.25	-0.04
Ocean	so (3D)	0.92	0.44	-0.18
	sos (2D)	0.9	0.45	-0.19
	thetao (3D)	0.99	0.17	0.01
	tos (2D)	0.99	0.12	0.01
	zos (2D)	0.98	0.23	-0.07
	siconc (2D) (NH March)	0.91	0.44	-0.13
	siconc (2D) (SH March)	0.81	0.93	0.13
	siconc (2D) (NH Sept)	0.96	0.28	-0.37
	siconc (2D) (SH Sept)	0.94	0.37	-0.09
OcnBgchem	dissic (3D)	0.92	0.46	-1.72
	NO_3 (3D)	0.91	0.47	-0.36
	O_2 (3D)	0.91	0.49	
	fgco2 (2D)	0.7	0.88	-1.37
Land	gpp (2D)	0.83	0.66	-0.56
	hfls (2D)	0.85	0.74	0.18
	hfss (2D)	0.58	1.02	-0.19

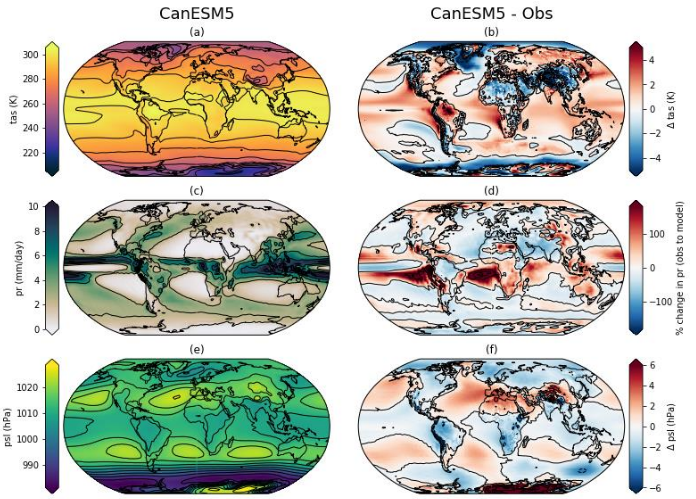
Figure ~~6~~4: Summary statistics quantifying the ability of CanESM to reproduce large scale climate features. Shown are the correlation coefficient (r) between the simulated and observed spatial patterns, the Root Mean Square Error (RMSE) normalized by the (observed spatial) standard deviation (σ), and the difference in normalized RMSE between CanESM5 and CanESM2. The spatial quantities represent temporal means over 1981 to 2010, except as

noted in appendix F. Variables are labelled according to the names in the CMIP6 data request, and are defined in Table F3.

5.1 Overall skill measures

The ability of CanESM5 to reproduce observed large scale spatial patterns in the climate system is quantified using global summary statistics computed over the 1981 to 2010 mean climate (Fig. ~~6~~4). Shown are the correlation coefficient between CanESM5 and observations (r), the Root Mean Square Error (RMSE) normalized by the observed (spatial) standard deviation (σ), and the change in normalized RMSE between CanESM2 and CanESM5. The statistics are weighted by grid cell area for 2D fields, volume for 3D ocean fields, and by area and pressure for 3D atmospheric variables. In general CanESM5 successfully reproduces many observed spatial patterns of the surface climate, interior ocean, and the atmosphere, with correlation coefficients between the model and observations generally above 0.8. Some exceptions are the total cloud fraction (clt, r=0.75), atmosphere-ocean CO_2 flux (fgco2, r=0.7) and the surface sensible heat flux (hfss, r=0.58).

For most variables, normalized RMSE has decreased in CanESM5 relative to CanESM2, indicating an improvement in the ability of the new model to reproduce observed climate patterns over its predecessor. The largest improvements were seen for ocean biogeochemistry variables, while small increases in error were seen for 3D distribution of zonal winds (ua), sea surface temperatures (tos), the March distribution of sea-ice in the Southern Hemisphere (siconc), and surface latent heat flux (hfls). In the following sections individual realms are examined, with a closer look at regional details and biases.



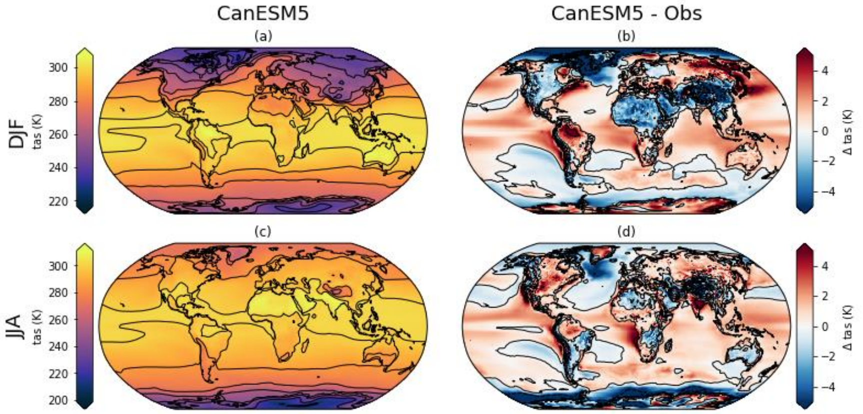


Figure 5: Climatologies over 1981 to 2010 of surface air temperature in CanESM5 (a, c), and their bias from ERA5 over the same period (b, d). Shown for seasons of DJF (a, b) and JJA (c, d).

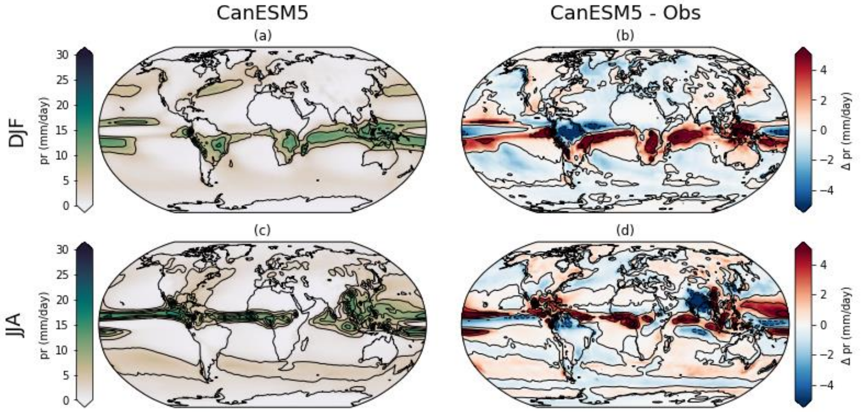


Figure 6: Climatologies over 1981 to 2010 of precipitation (a, c) in CanESM5, and their bias from GPCP (b, d) over the same period. Shown for seasons of DJF (a, b) and JJA (c, d).

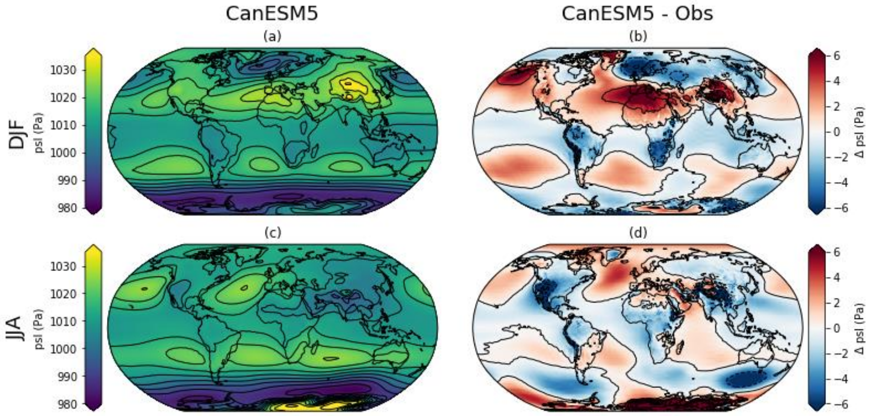
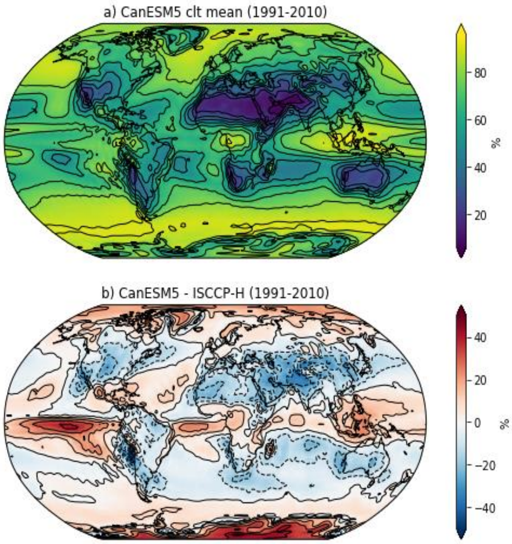


Figure 7: Climatologies over 1981 to 2010 of ~~(a) surface air temperature, (c) precipitation and (e) sea-~~sea level pressure in CanESM5~~,~~ (a, c), and their bias from ~~(b) ERA5, (d) GPCP and (f)~~ ERA5 over the same period~~.~~ (b, d). Shown for seasons of DJF (a, b) and JJA (c, d).

Formatted: Left

5.2 Atmosphere

CanESM5 reproduces the large scale climatological features of surface air temperatures~~,~~ (Fig. 5), precipitation (Fig. 6) and sea-level pressure~~,~~ (Fig. 7), though significant regional biases exist~~ (Fig. 7).~~. CanESM5 is significantly colder than observed over sea-ice covered regions (Fig. ~~7a, b~~5), noticeable in the Southern Ocean, and~~ most obviously~~ in the region surrounding the Labrador sea, which has extensive seasonal sea-ice cover in CanESM5 (see below). The Tibetan plateau, the Sahara and the broader North Atlantic Ocean are also cooler than observed. Warm biases exist over the eastern boundary current systems (Benguela, Humboldt, and California); over the Amazon, eastern North America, much of Siberia, and over broad regions of the tropical and subtropical oceans.



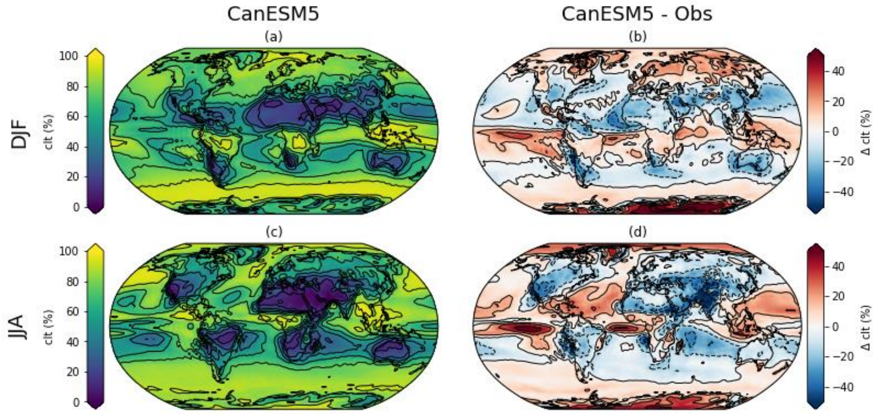


Figure 8: Cloud fraction in ~~(a)~~ CanESM5 ~~and (b) the~~(a, c) their bias with respect to ISCPP-H satellite based observations~~.~~ (b, d). Shown for seasons of DJF (a, b) and JJA (c, d).

Formatted: Left, Indent: Left: 0 cm

Formatted: Font: 10 pt

Formatted: Left

Precipitation biases vary in sign by region (Fig. ~~7d~~6). The largest ~~relative (to mean)~~ biases are ~~excessive simulated precipitation~~ over the ~~eastern~~tropical Pacific and Atlantic oceans, between the equator and extending into the southern subtropics. The overall pattern of precipitation biases is very similar to that seen across the CMIP5 (Flato et al., 2013) and CMIP3 (Lin, 2007) models. The largest land biases are excessive precipitation over much of sub-Saharan Africa, Southeast Asia, Canada, and Peru-Chile. In contrast western Asia, Europe, the North Atlantic and the subtropical to high-latitude Southern Oceans have too little simulated precipitation. The large scale pattern of sea-level pressure is captured by CanESM5 (Fig. ~~7e~~7). Biases relative to ERA5 are largest over the high elevations of Antarctica (Fig. ~~7f~~7), possibly reflecting differences in the extrapolation of surface pressure to sea-level.

Relative to ISCCP-H (Young et al, 2018), version 1.00 (Rossow et al, 2016) the total cloud fraction in CanESM5 is overestimated along the equator, particularly in the eastern tropical Pacific and Atlantic (Fig 8). Too large cloud fraction is also found over Antarctica and the Arctic. Underestimates of total cloud fraction occur over most other land areas, with the largest underestimates over Asia and the Himalayas.

Zonal mean sections of air temperature for the DJF and JJA seasonal means are shown in Fig. 9. In both seasons, CanESM5 is biased warm relative to ERA5 near the tropopause, across the tropics and subtropics. Warm biases also occur in the stratosphere, notably near 60°S above 50 hPa in JJA. Cold biases exist from the subtropics to the high latitudes, where they reach from the surface to the stratosphere, and are strongest in the winter season.

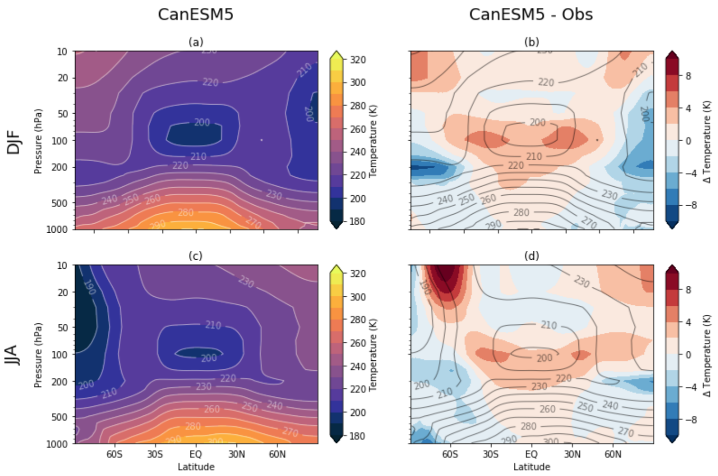


Figure 9: Zonal mean temperature in CanESM5 (a, c) and bias relative to ERA5 (b, d) over 1981-2010, for the DJF (a, b) and JJA (c, d) seasons.

Formatted: Left, Indent: Left: 0 cm

Formatted: Left

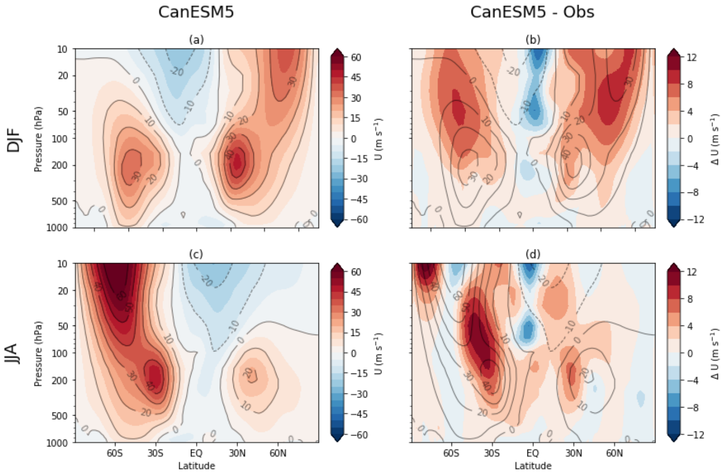


Figure 10: Zonal mean zonal winds (a, c) and bias relative to ERA5 (b, d) over 1981-2010, for the DJF (a, b) and JJA (c, d) seasons.

Formatted: Left, Indent: Left: 0 cm

Zonal mean zonal winds are compared to ERA5 in Fig. 10 for DJF and JJA. The westerly jets in CanESM5 are biased strong, particularly aloft and in the winter hemisphere. Surface zonal winds in CanESM5 are only slightly stronger than observed, and are significantly improved over those in CanESM2 (Fig. 11), which were too strong, particularly over the Southern Hemisphere westerly jet.

Formatted: Left

5.3 Land physics and biogeochemistry

Figures 12 and 13 compare the geographical distribution and zonal averages of gross primary productivity (GPP), and latent and sensible heat fluxes over land with observation-based estimates from Jung et al. (2009). The zonal averages of GPP, and latent and sensible heat fluxes compare reasonably well with observation-based estimates although the latent heat fluxes are somewhat higher especially in the southern hemisphere as discussed below (Fig. 13). Figure 12 shows the biases in the simulated geographical distribution of these quantities. In the tropics biases in GPP, and latent and sensible heat fluxes, broadly correspond to biases in simulated precipitation compared to observation-based estimates (shown in Fig. ~~7~~6).

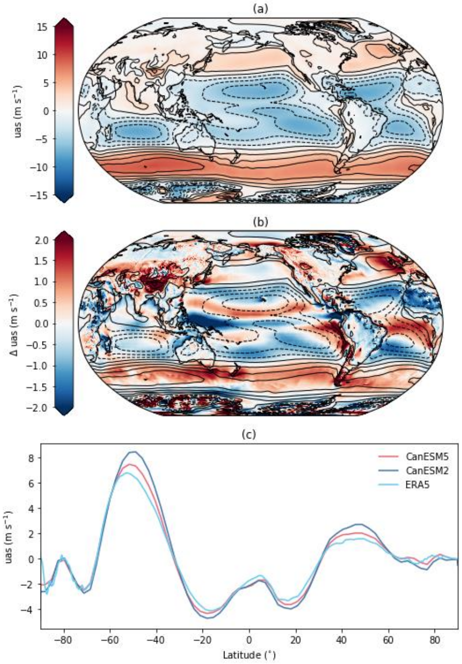


Figure 11: Zonal surface winds in (a) CanESM5, (b) the bias relative to ERA5 and (c) zonal-mean zonal surface winds in CanESM2, CanESM5 and ERA5.

Generally over tropics, as would be intuitively expected, the sign of GPP and latent heat flux anomalies are the same since they are both affected by precipitation in the same way. Sensible heat flux is expected to behave in the opposite direction compared to GPP and latent heat flux in response to precipitation biases. For example, simulated GPP and latent heat fluxes are lower, and sensible heat fluxes higher in the north eastern Amazonian region because simulated precipitation is biased low (Fig. ~~7~~6). The opposite is true for almost the entire African region south of the Sahara desert and most of Australia. Here simulated precipitation that is biased high, compared to observations, results in simulated GPP and latent heat flux that are

Formatted: Left, Indent: Left: 0 cm

Formatted: Left

higher and sensible heat flux that is lower than observation-based estimates. At higher latitudes, where GPP and latent heat flux are limited by temperature and available energy, the biases in precipitation do not translate directly into biases in GPP and latent heat flux as they do in the tropics.

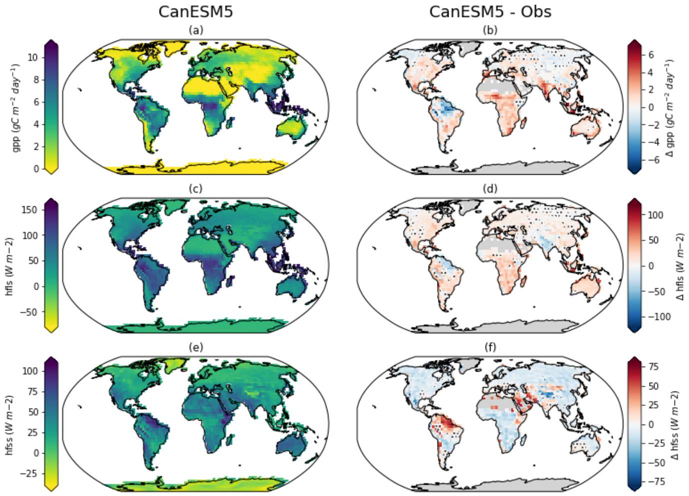


Figure 12: Time-mean values of (a) gross primary productivity (GPP), (c) latent heat flux (HFLS), and (e) sensible heat flux (HFSS) from CanESM5 (r1i1p1f1) (left-hand column) and the corresponding biases with respect to observation-based reference data presented in Jung et al. (2009) (GBAF) (right- hand column). Black dots mark grid cells where biases are not statistically significant at the 5% level using the two-sample Wilcoxon test.

The biases in simulated climate imply that simulated land surface quantities will also be biased which make it difficult to assess if the underlying model behaviour is realistic. This limitation can be alleviated to some extent by looking at the functional relationships between a quantity and its primary climate drivers. This technique works best when a land component is driven offline with meteorological data. In a coupled model, as is the case here, land-atmosphere feedbacks can potentially worsen a model's performance by exaggerating an initial bias. For example, low model precipitation can be further reduced due to feedbacks from reduced evapotranspiration some of which is recycled back into precipitation. Figure 14 shows the functional relationships between GPP and temperature, and GPP and precipitation, for both model and observation-based estimates. The observations-based temperature and precipitation data used in these plots are from CRU-JRA reanalysis data that were used to drive ~~participating~~ terrestrial ecosystem models in the TRENDY Intercomparison for the 2018 Global Carbon Budget (Le Quéré et al., 2018). Figure 14 shows that GPP increases both with increases in precipitation (as would be normally expected) and temperature except at mean annual values above 25 °C when soil moisture

Formatted: English (Canada)

Formatted: English (Canada)

Formatted: English (Canada)

Formatted: English (Canada)

limits any further increases. This threshold emerges both in the model and the observation-based functional relationships. With the caveat mentioned above, the functional relationships of GPP with temperature and precipitation based on simulated data compare reasonably well with those based on observation-based data, although the simulated GPP relationship with precipitation compares much better to its observation-based relationship than that for temperature.

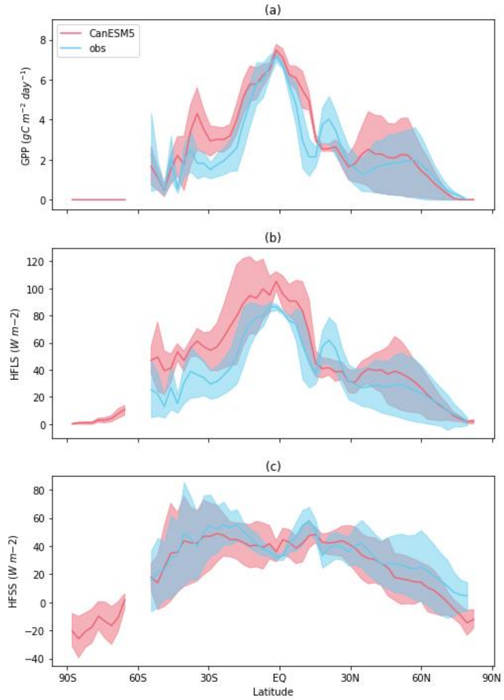


Figure 13: Zonal mean values of (a) GPP, (b) HFLS, and (c) HFSS for CanESM5 (r1i1p1f1) ~~(black)~~ and reference data ~~(red)~~ from Jung et al. (2009). The shading presents the corresponding inter-quartile range ~~(IQR).~~that results from inter-annual variability as well as longitudinal variability for the period 1982 to 2008.

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font: 10 pt

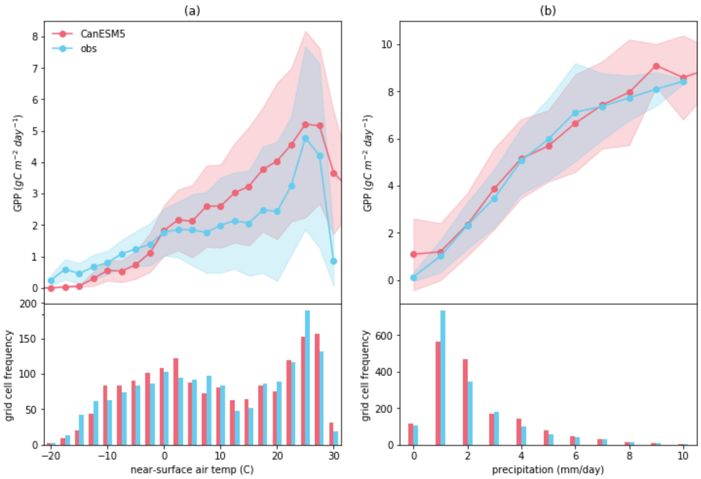


Figure 14: Functional response of GPP to (a) near-surface air temperature and (b) surface precipitation for CanESM5 (r1i1p1f1) ~~(black)~~ and reference data ~~(red)~~ from Jung et al. (2009~~) (GBAF). The~~). Values present monthly mean values averaged over the period 1982 to 2008 and the shading ~~presents~~shows the corresponding standard deviation~~. The grid cell frequencies are shown in the lower part of each plot~~.

As mentioned earlier, dynamically-simulated wetland extent and wetland methane emissions in CanESM5 are purely diagnostic. Figure G1 in appendix G compares zonal distribution of simulated annual maximum wetland extent with observation-based estimates and shows the temporal evolution of annual maximum wetland extent and wetland methane emissions over the historical period.

5.4 Physical ocean

CanESM5 reproduces the observed large scale features of sea surface temperature (SST), salinity (SSS) and height (SSH) (Fig. 15). The largest SST biases are the cold anomalies south east of Greenland and in the Labrador Sea (Fig. 15b). These negative SST biases are associated with excessive sea-ice cover, described further below, and with the surface air temperature biases mentioned above. Positive SST biases are largest in the Eastern Boundary Current upwelling systems, as for surface air temperatures.

Sea surface salinity biases are largest, and positive, around the Arctic coastline, potentially indicating insufficient runoff in this region (Fig. 15d). Negative annual mean SSS biases occur ~~under the region associated with excessive March sea-ice~~ in

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto, Pattern: Clear

Formatted: Font color: Auto, Pattern: Clear

Formatted: Font color: Auto

Formatted: Font: 10 pt

Formatted: Left

the Labrador Sea, and are also found in seas of the maritime continent and eastern tropical Atlantic. Sea-surface height (SSH) is shown as an anomaly from the (arbitrary) global mean (Fig. 15e). Significant SSH biases are associated with the positions of western boundary currents, noticeably for the Gulf Stream and Kuroshio current (Fig. 15f). CanESM5 has too low SSH around Antarctica, and too high SSH in the southern subtropics, with an excessive SSH gradient across the Southern Ocean. This SSH gradient is associated with the geostrophic flow of the Antarctic Circumpolar Current (ACC). The ACC in CanESM5 is vigorous with 190 Sv of transport through Drake Passage. This is larger than observational estimates which range up to 173.3 ~~+/-~~ ± 10.7 Sv (Donohue et al, 2016). In CanESM5 the ACC also exhibits a pronounced, centennial scale variability of about 20 Sv, which is also evident in the piControl simulation (not shown).

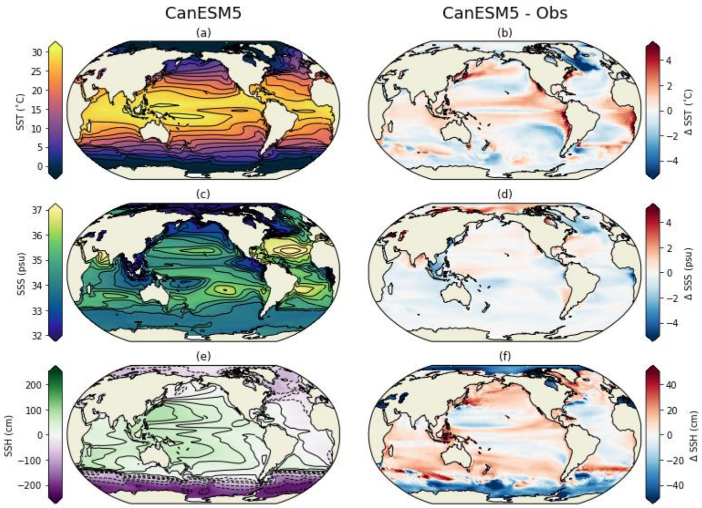


Figure 15. Sea surface (a) temperature, (c) salinity and (e) height averaged over 1981 to 2010, and their biases relative to World Ocean Atlas 2009 (b, d), and the AVISO mean dynamic topography (f).

The CanESM5 interior distributions of potential temperature and salinity are well correlated with observations (Fig. ~~6~~4). In the zonal mean, potential temperature biases are largest within the thermocline, which is warmer than observed, particularly near 50°N (Fig. 16a, b). The deep ocean, the Southern Ocean south of 50°S and the Arctic Ocean are cooler than observed. The pattern of excessive heat accumulation in the thermocline is very similar to the pattern of bias seen in CMIP5 models on average (Flato et al., 2013 their Fig. 9.13). Also similar to CMIP5 models there is a cold bias in the ocean below the thermocline. This suggests that the processes controlling the redistribution of heat between the thermocline and the deep ocean play a role in establishing the vertical structure of these temperature biases. For example, Saenko et al. (2012) find that heat redistribution in ocean models can be sensitive to the vertical structure of diapycnal mixing. The major salinity bias is of excessive fresh waters in the Arctic near 250 m, also typical of the CMIP5 models (Fig. 16d). Sea-surface salinities showed the Arctic to be too salty, but this bias is confined to near the surface, and at all depths below the immediate surface layer the Arctic Ocean is too fresh. The zonal mean salinity also shows a positive salinity bias near 40°N, associated with the Mediterranean outflow.

The Meridional Overturning Circulation in the global ocean, and the Indo-Pacific, as well and Arctic-Atlantic basins is shown in Fig. 17. The global overturning streamfunction shows the expected major features: an upper cell with clockwise rotation, connecting North Atlantic Deepwater formation to low latitude and Southern Ocean upwelling; a vigorous Deacon cell in the Southern Ocean (as a result of plotting in z-coordinates); a lower counter clockwise cell of Antarctic Bottom Water, and vigorous near-surface cells in the subtropics. The upper cell overturning rate at 26°N in the Atlantic is estimated to be 17±4.4 Sv from the RAPID observational array (McCarthy et al. 2015). CanESM5 produces an Atlantic overturning rate of 12.8 Sv at 26°N, below the mean but within the range measured by RAPID. The fairly weak AMOC in CanESM5 is likely associated with excessive sea-ice cover in the Labrador Sea, which inhibits convection. Though we also note that NEMO models have previously been found to underestimate the AMOC (Danabasoglu et al., 2014).

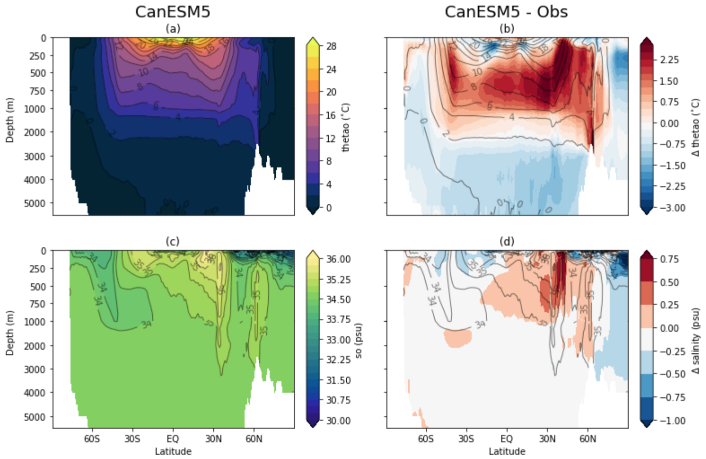


Figure 16: CanESM5 zonal mean ocean (a) potential temperature, (c) salinity averaged over 1981 to 2010, and their biases from World Ocean Atlas 2009 (b, d). Note the depth-scale on the y-axis is non-uniform.

Closely connected to the MOC is the rate of northward heat transport by the ocean (Fig. 18). CanESM5 produces the expected latitudinal distribution of heat transport, but consistent with a weak MOC, slightly underestimates the transport at 24°N, relative to the inverse estimate of Ganachaud and Wunsch (2003). To the north and south, CanESM5 ocean heat transport falls within the observational uncertainties. The MOC and heat transport in CanESM5 are similar to those in CanESM2, as reported in Yang and Saenko (2012)

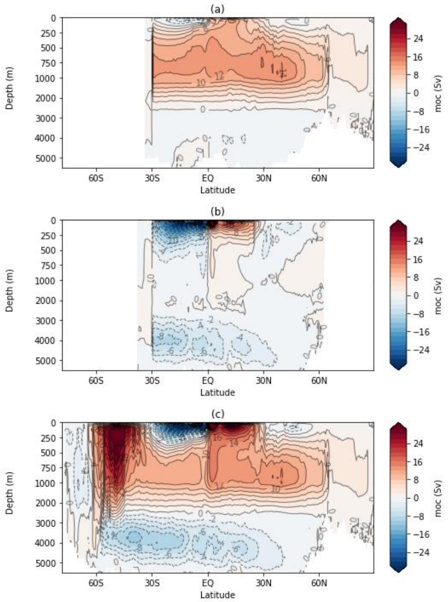


Figure 17: CanESM5 residual meridional overturning circulation in the Atlantic (a), Indo-Pacific (b) and global (c) oceans, averaged over 1981 to 2010 including all resolved and parameterized advective processes. Note the depth-scale on the y-axis is non-uniform.

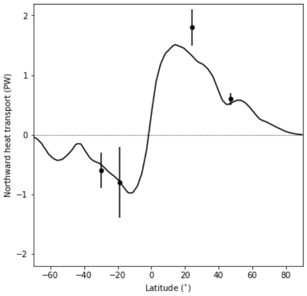


Figure 18: Northward heat transport in the global ocean in CanESM5 (in Petawatts), with error bars showing the inverse estimate of Ganachaud and Wunsch (2003).

Formatted: Left, Indent: Left: 0 cm

Formatted: Left

5.5 Sea-ice

The seasonal cycle of sea-ice extent and volume are shown in Fig. 19. A major change from CanESM2 is seen in the sea-ice volume (Fig. 19b, d). CanESM2 simulated very thin ice, and had about 40% less Northern Hemisphere (NH) ice volume than in the PIOMAS reanalysis~~.~~ (Zhang and Rothrock., 2003; Schweiger et al., 2011). By contrast, CanESM5 has a larger NH ice volume than in CanESM2 ~~and in PIOMAS~~ (Fig. 19b). The amplitude and phase of the annual cycle in NH sea-ice volume in CanESM5 is similar to PIOMAS (Fig. 19b). In the Southern Hemisphere, CanESM5 also has a larger sea-ice volume and seasonal cycle far more consistent with the GIOMAS reanalysis product than CanESM2 (Fig. 19d).

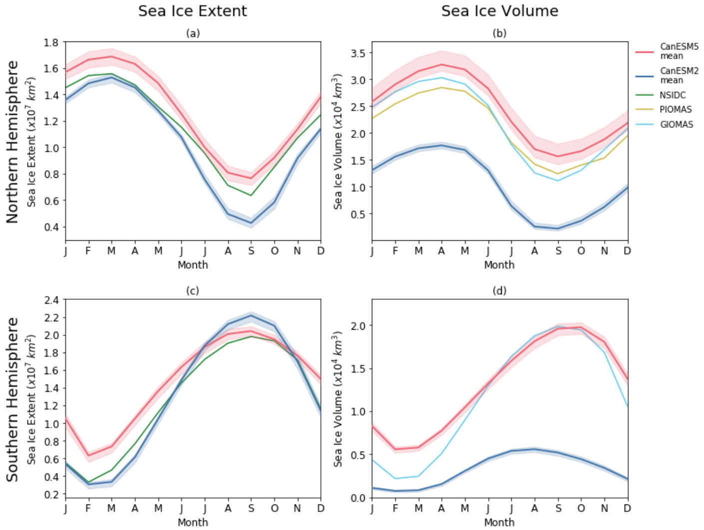


Figure 19: Seasonal cycles of sea-ice extent (a, c) and volume (b, d) in the Northern (a, b) and Southern (c, d) hemispheres averaged over 1981 to 2010. Results are shown for CanESM2, CanESM5, the NSIDC satellite based observations, and the PIOMAS and GIOMAS reanalyses.

While CanESM2 significantly underestimated NH sea-ice extent relative to satellite based observations, CanESM5 generally overestimates the extent (Fig. 19a). The NH sea-ice extent biases are largest in the winter and spring. During the March maximum, excessive sea-ice is present in the Labrador Sea and east of Greenland (Fig. 20a). In the summer and fall, the net NH extent bias is far smaller (Fig. 20c), and results from a cancellation between lower than observed concentrations over the Arctic basin and larger than observed concentrations around northeastern Greenland. Southern Hemisphere sea-ice extent biases are largest during the early months of the year, and in March the positive concentration biases are focused in the northeastern Weddell and Ross Seas (Fig. 20b). In September SH concentration biases between CanESM5 and the satellite observations are focused around the northern ice-edge, and are of varying sign (Fig. 20d).

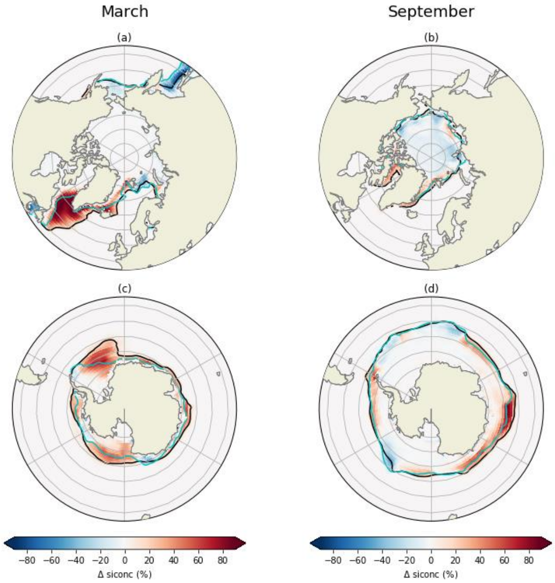


Figure 20: Sea-ice concentration biases between CanESM5 and NSIDC climatologies for the months of March (a, c) and September (b, d), in the Northern (a, b) and Southern (c, d) hemispheres. The solid black contour marks the ice-edge (15% threshold) in CanESM5, and the teal line marks the ice-edge in the observations. Biases are based on the 1981 to 2010 climatology.

5.6 Ocean biogeochemistry

The standard configuration of CanESM5 has a significantly improved representation of the distribution of ocean biogeochemical tracers relative to CanESM2, despite using the same biogeochemical model (CMOC). For the three-dimensional distributions of Dissolved Inorganic Carbon (DIC) and NO3, and the surface CO_2 flux, the Root Mean Square Error (RMSE), relative to observed distributions was reduced by over a factor of two (Fig.~~6~~ 4). Ocean only simulations, whereby NEMO was driven by CanESM2 surface forcing via bulk formulae, show similar skill to the CanESM5 coupled model. From this we infer that changes in interior ocean circulation, rather than boundary forcing, are responsible for the improved representation of biogeochemical tracer distributions.

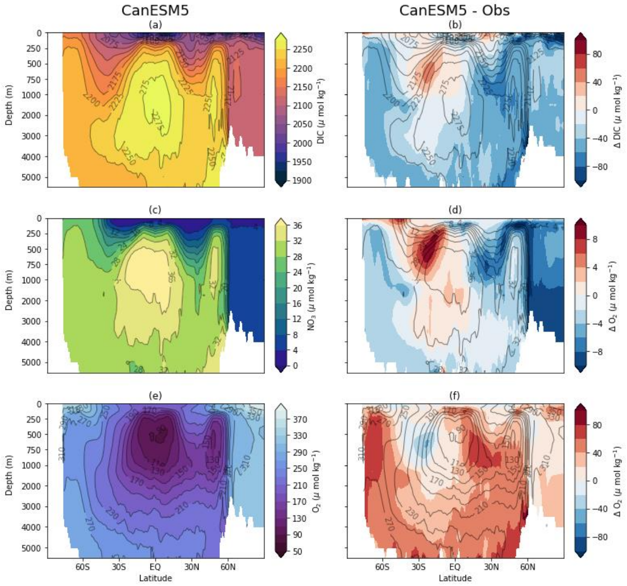


Figure 21: Zonal mean sections of (a) Dissolved Inorganic Carbon, (c) NO_3, and (e) O_2 in CanESM5 averaged over 1981 to 2010, and their biases relative to GLODAP v2 (b, d, f). Note the depth-scale on the y-axis is non-uniform.

In CanESM5 the zonal mean DIC concentration simulated by CMOC is generally lower than observed, by amounts reaching up to about 5% (Fig. 21a, b). One exception to this is in the SH subtropical thermocline, on the northern flank of the Southern Ocean, which shows positive DIC biases between 250 and 1000 m. This area is also one of positive nitrate biases, whose magnitude is close to 30% (Fig. 21d). Elsewhere zonal-mean NO_3 concentrations are generally too low, particularly in the NH thermocline and the Arctic. CanESM5 has higher than observed concentrations of zonal mean O_2 (Fig. 21f). As expected from saturation, biases are largest in the Southern and abyssal ocean, where CanESM5 is colder than observed. However, positive O_2 biases also occur at the base of the thermocline in the NH, where CanESM5 is too warm, suggestive of a biological origin.

The zonal mean NO_3 biases identified at the thermocline level above are the result of partially cancelling biases between the Pacific and Atlantic basins (not shown). The Atlantic has negative NO_3 biases, largest near 1000 m. Meanwhile, there is an excessive accumulation of NO_3 centered at the base of the eastern Pacific thermocline. This buildup occurs due to the simplified parameterization of denitrification in CMOC. Within each vertical column, the amount of denitrification is set to balance the rate of nitrogen fixation, and is distributed vertically proportional to the detrital remineralization rate. In reality nitrogen fixation and denitrification are not constrained to balance within the water column at any one location, but rather denitrification proceeds within anoxic areas. A prognostic implementation of denitrification implemented into CanOE resolves this bias, and will be discussed further in an upcoming article within this special issue.

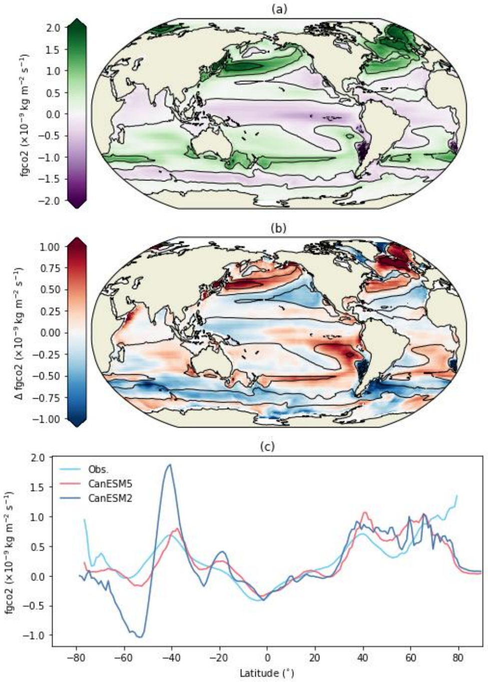


Figure 22: Ocean atmosphere flux of CO_2 in (a) CanESM5 averaged over 1981 to 2010, (b) ~~from~~the bias relative to Landschutzer (2009), and (c) zonal mean CO_2 flux in CanESM2, CanESM5 and Landschutzer (2009) data. The flux is positive down (into the ocean).

The atmosphere-ocean CO_2 flux pattern in CanESM5 correlates significantly better with estimates of the observed flux than CanESM2 (Fig. ~~6~~4). The largest departures from the observations are positive biases in the southeastern Pacific, northwest Pacific and northwest Atlantic (Fig. 22b). These are compensated by negative biases in the Southern Ocean and mid-latitude northeast Pacific. In the zonal mean, CanESM2 had a large flux dipole in the Southern Ocean, which is significantly reduced in CanESM5, and attributable to improved circulation in the new NEMO ocean model and a reduction in Southern Ocean wind speed biases in CanAM5 (Fig. 22c).

5.7 Modes of climate variability

5.7.1 El-Niño Southern Oscillation

The El-Niño Southern Oscillation (ENSO) is a key component of climate variability on seasonal and interannual timescales. To evaluate CanESM5's representation of ENSO, the NINO3.4 index (average monthly SST anomaly in the region bounded by 5S, 5N, 170W, 120W) from the first 10 historical ensemble members is compared against HadISST. The skill of CanESM5 at representing the local and remote effects of ENSO is evaluated by correlating SST anomalies with the resulting NINO3.4 index (Fig. 23a, b). Within the equatorial Pacific, a positive ENSO event in CanESM5 leads to an increase in SSTs across the entire basin whereas observations show negative SST anomalies in the western basin and positive anomalies in the central and eastern Pacific. ENSO in CanESM5 also has weaker teleconnections. The SST within the subtropical North and South Pacific gyres are more weakly anticorrelated to ENSO than observed. HadISST shows a negative North Atlantic Oscillation like pattern associated with ENSO, which is not present in CanESM5. The SST teleconnection in the tropical Indian and Atlantic Oceans is well represented by the model.

The spectral peak in the historical ensemble members (Fig. 23c) occurs at around 3-5 years in general agreement with observations. Variability on decadal time-scales has a large spread between ensemble members likely due to differences in the strength of warming trends over the historical period. Higher frequency variability at monthly to seasonal timescales is significantly lower than observed. The lower monthly variability can also be seen by examining month-by-month interannual variability of NINO3.4 (Fig. 23d). While January remains the month of peak variability, overall the annual cycle of NINO3.4 variability is weaker in CanESM5. In observations, ENSO variability is at its minimum between April and June but in CanESM5 the minimum variability (depending on the ensemble member) tends to be between July and September.

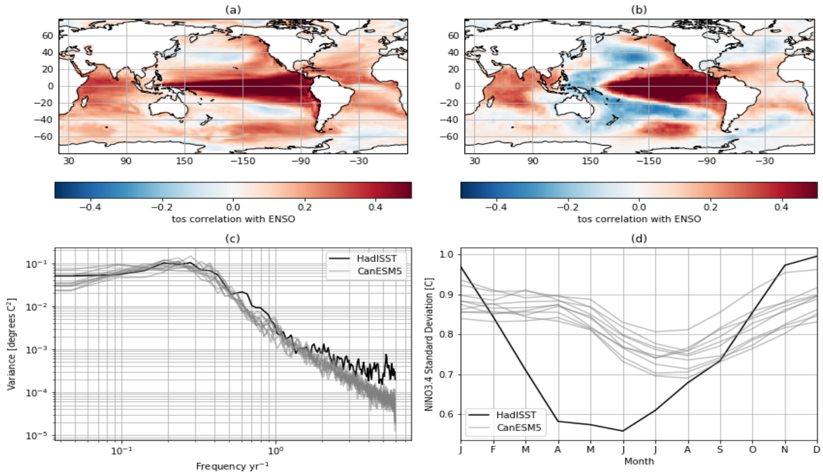


Figure 23: Characteristics of the El Niño Southern Oscillation (ENSO) from and the HadISST observational product. Spatial maps in (a) and (b) are the regression of the SST monthly anomalies from 1850-2014 against the NINO3.4 index from (a) CanESM5 (historical ensemble member r1i1p1f1) and (b) from HadISST. Temporal variability is summarized as power spectra (c) of the NINO3.4 index from HadISST and ten historical ensemble members and the interannual variability of the NINO3.4 index by month (panel d) for CanESM5 and HadISST.

5.7.2 Annular Modes

The Northern Annular Mode is computed as the first EOF of extended winter (DJFM) sea level pressure north of 20°N for CanESM5 and ERA5 (Fig. 24 a, b). The correlation between the CanESM5 and ERA5 patterns is 0.95. Despite the high degree of coherence, some differences between the model pattern and reanalysis are evident (Fig. 24). For example CanESM5 has a positive centre in the north Pacific, not seen in ERA5, and the positive pattern across the North Atlantic is less continuous in CanESM5. This is a typical model bias (e.g. ~~Bentson~~Bentsen et al., 2013). The first EOF in CanESM5 also explains about 8% more variance than in the reanalysis.

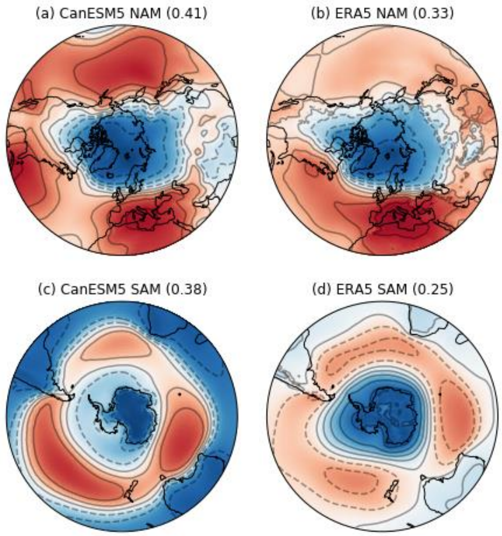


Figure 24: First Empircal Orthogonal Functions (EOFs) of sea-level pressure north of ~~20N~~20°N (a, b), and south of ~~20S~~20°S (c, d), representing the Northern Annual Mode and Southern Annular Mode respectively. The NAM is based on the extended winter DJFM season, and the SAM is based on monthly sea-level pressure. Results are shown for CanESM5 (a, c) and ERA5 (b, d), and the amount of variance explain by each EOF is given in brackets. ~~The color scale is arbitrary.~~

The Southern Annular Mode is the dominant mode of climate variability in the Southern Hemisphere, with significant influences on atmospheric circulation, precipitation, and the Southern Ocean. We compute the SAM pattern as the first EOF of sea level pressure south of 20°S. The CanESM5 and ERA5 pattern correlation is 0.7. In CanESM5, the first EOF accounts for 13% more variance than in the reanalysis. Despite such biases, these results confirm that CanESM5 captures the principal modes of tropical and mid-latitude climate variability.

6 Climate response to forcing

6.1 Response to CO2 forcing

The global mean screen temperature change under the idealized CMIP6 DECK experiments "abrupt-4xCO2" and "1pctCO2" are shown in Fig. ~~4~~2. From these simulations, three major benchmarks of the model's response to CO2 forcing can be quantified (Table 3).

Table 3: Key sensitivity metrics: Transient Climate Response (TCR), Transient Climate Response to Cumulative Emissions (TCRE), and Equilibrium Climate Sensitivity (ECS).

Model	**TCR (K)**	**TCRE (K/EgC)**	**ECS (K)**
CanESM2	2.4	2.3	3.~~8~~7
CanESM5	2.8	1.9	5.~~7~~6

Formatted: Left

Formatted Table

Formatted: Left

Formatted: Left

Formatted: Left

The Transient Climate Response (TCR) of the model is given by the temperature change in the 1pctCO2 experiment, averaged over the 20 years centered on the year of CO$_2$ doubling (year 70), relative to the piControl. For CanESM5 the TCR is 2.8 K, an increase of 0.4 K over that seen in CanESM2. The CanESM5 TCR is larger than seen in any CMIP5 models, and significantly higher than the CMIP5 mean value of 1.8 K (Flato et al., 2013). The likely range ($\rho > 0.66$) of TCR was given by the IPCC AR5 as 1.0-2.5 K (Collins et al., 2013), while more recent observational based estimates quote a 90% range of 1.2 to 2.4 K (Schurer and Hegerl, 2018), again subject to significant observational and methodological uncertainty.

The Transient Climate Response to Cumulative Emissions (TCRE), incorporates the transient climate sensitivity together with the carbon sensitivity of the system (~~Mathews~~Matthews et al., 2009). It is defined as the ratio of global mean surface warming to cumulative carbon emissions, over the 20 years centered on CO$_2$ doubling in the 1pctCO2 experiment, with units K EgC^{-1}. The metric is of major policy relevance, and is widely used to estimate the allowable emissions to reach given temperature targets. The TCRE of CanESM5 is 1.9 K EgC^{-1}, slightly lower than the CanESM2 value of 2.3 K EgC^{-1}. The reduction in TCRE occurs despite the fact that CanESM5 has a larger temperature response (TCR) than CanESM2~~. The reduction occurs owing~~ due to significantly larger uptake of CO$_2$ by the land biosphere in CanESM5 relative to CanESM2 in the 1pctCO2 experiment. As mentioned in Section 2.2, this is due to higher strength of the CO$_2$ fertilization effect in CanESM5 relative to CanESM2. As shown in Arora and Scinocca (2016) this leads to land carbon uptake in the 1pctCO2 simulation that is higher than in all CMIP5 models compared in Arora et al. (2013). Gillett et al. (2013) estimated the TCRE in 15 CMIP5 models to range from 0.8 to 2.4 K EgC^{-1}, and the IPCC AR5 likely range was assessed as 0.8 to 2.5 K EgC^{-1}.

The Equilibrium Climate Sensitivity (ECS) is defined as the amount of global mean surface warming resulting from a doubling of atmospheric CO$_2$, and a key measure of the sensitivity to external forcing. Given the long equilibration time of the climate system, it is common to estimate ECS from the relationship between surface temperature change and radiative forcing, over the course of the first 140 years of the abrupt-4xCO2 simulation (Gregory et al., 2004). Here the ECS is calculated using the Gregory (2004) regression method, after removing linear drift from the piControl following Forster et al. (2013). For CanESM5~~,~~ the ECS is 5.~~7~~6 K, a significant increase over the value of 3.~~8~~7 K in CanESM2. Like TCR, the CanESM5 ECS value is larger than seen in any CMIP5 models, and significantly higher than the CMIP5 mean value of 3.2

K (Flato et al., 2013). The likely range for ECS was given by the IPCC AR5 as 1.5 to 4.5 K (Collins et al., 2013). CanESM5 falls outside this range, although it is worth noting that there are significant uncertainties in observational constraints of ECS. We also note, as above, that ECS is an emergent property in CanESM5 - no model tuning was done on the response to forcing.

A detailed explanation of the reasons behind the increased ECS in CanESM2 over CanESM5 is beyond the scope of this paper. However, the effective radiative forcing (Forster et al, 2016) in CanESM5 due to abrupt quadrupling of CO_2 is very similar to that in CanESM2, suggesting that changes in feedbacks rather than forcings are the source of the higher ECS. Indications are that the increase in ECS is associated with cloud and surface albedo feedbacks, with sea-ice likely playing an important role in the latter effect. The cloud albedo feedback is found to be sensitive to parameter settings in the cloud microphysics scheme. A more detailed examination of the changes in ECS due to cloud microphysics will be provided in a companion paper in this special issue (Cole et al, 2019). The examination of climate change over the historical period in the following section also reveals some further insights.

6.2 Climate change over the historical period

In this section we briefly discuss CanESM5 simulated changes in surface air temperature, sea-ice, and carbon cycle fluxes over the historical period. We choose these as major emblematic variables of climate change. Here we make use of the CanESM2 50-member large initial condition ensemble (Kirchmeier-Young et al., 2017; Swart et al., 2018). The 50 realizations in this ensemble were branched in the year 1950 from the five CanESM2 realizations submitted to CMIP5, and were forced by CMIP5 historical (1950 to 2005) and Representative Concentration Pathway (RCP) 8.5 (2006 to 2100) forcing.

6.2.1 Surface temperature changes

Global Mean ~~Screen~~Surface Temperature (GMST) changes in CanESM2 and CanESM5 are generally consistent with the observations over the period from 1850 to around the end of the 20th century (Fig. 25a). However, from 2000 to 2014, the increase in GMST is larger in the models than observed. Possible reasons for the divergence are i) forcing errors in the CMIP5 and/or CMIP6 forcing datasets, ii) natural internal variability, iii) incorrect partitioning of heat across components of the climate system or iv) a higher climate sensitivity in the model than in the real world. The 25 realizations of CanESM5 (and 50 realization of CanESM2) provide a good estimate the contribution of internal variability in the model. The observations fall outside the range of this variability, and hence this cannot account entirely for the divergence between the model and observations (assuming the model correctly captures the scale of internal variability). Trends computed from 1981 to 2014 show that the models are warming at roughly twice the observed rate over this period (Fig. 25b). The spread across the 25 realizations from CanESM5 and 50 realizations from CanESM2 do not encompass the observations, reinforcing the point above. CanESM5 warms more rapidly than CanESM2 on average, as would be expected from its higher

ECS and TCR. There is however significant overlap across the distribution of warming rates across the CanESM5 and CanESM2 ensembles. Interestingly, the lower tail of the trend probability distribution functions aligns for the two models, but CanESM5 has a broader distribution, and a larger tail of high warming realizations.

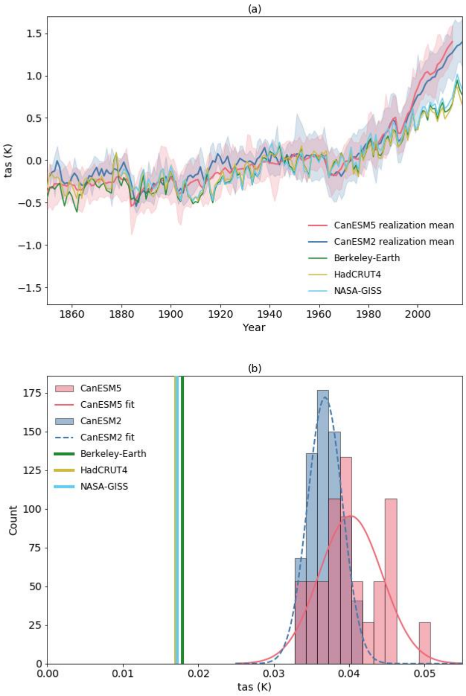


Figure 25 (a) Global mean screen temperature in CanESM5, CanESM2 and various observational products and (b) histogram of historical trends over 1981 to 2014. In (a) the shaded envelopes represent the range over the CanESM2 50 member large ensemble and the CanESM5 25 member "p1" ensemble. In (b) fits of the normal distribution to the CanESM2 and CanESM5 distributions are also shown.

The pattern of surface warming in CanESM5 over the historical period is shown in Fig 26a. The canonical features of global warming are consistent between the model and observations: greater warming over land than ocean, and Arctic amplified warming. The zonal-mean warming trends (Fig 26c) show that both CanESM2 and CanESM5 warmed more than the observations over most latitudes. Divergence between simulated and observed warming rates is largest in the high latitudes, notably over the Southern Ocean and north of 40°N. The larger warming in the CanESM5 ensemble mean, relative to the

CanESM2 ensemble mean, largely occurs over the Arctic. However, there is a very large variability in Arctic warming trends in CanESM5, which most likely are responsible for the spread in GMST trends noted above. Some realizations have lower trends, which overlap with observed warming, while others exhibit considerably higher rates of Arctic warming. Observed warming rates over the Arctic are also some of the most uncertain, due to data sparsity (HadCRUT is masked where observations are not available).

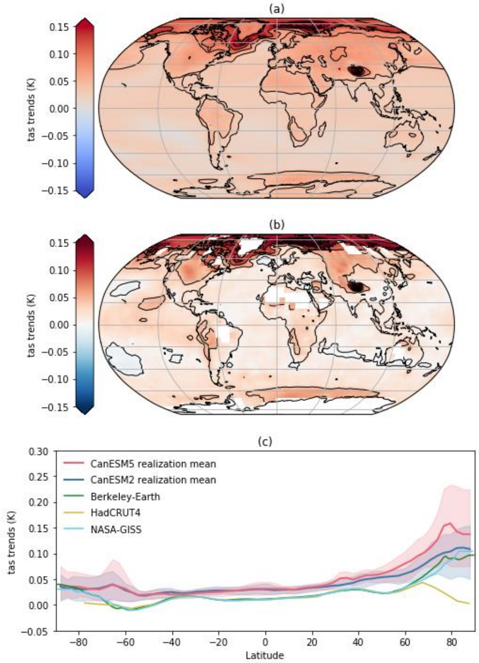


Figure 26: Surface temperature trends in CanESM5 (a), the difference in trend between CanESM5 and HadCRUT4 (b), and zonal mean of trends in CanESM2, CanESM5, and HADCRUT4 over 1981 to 2014 (c). The shaded envelopes in (c) represent the range over the CanESM2 50 member large ensemble and the CanESM5 25 member "p1" ensembles.

6.2.2 Sea-ice changes

CanESM5 closely reproduces the observed reduction in Arctic September sea-ice extent (Fig. 27a). The trends from both the 50 CanESM2 ensemble members, and the 25 CanESM5 ensemble members, show a broad spread due to internal variability

(Fig. 27c). The observed trends lie close to the centre of the model distribution of trends. Given than CanESM5 warms more rapidly than observed, the sea-ice sensitivity (rate of sea-ice decline normalized by the rate of warming) is likely too low (Rosenblum and Eisenman, 2017; Winton, 2011).

In the Southern Hemisphere, observed annual mean Antarctic sea-ice extent showed a tendency to increase, before dramatic declines in the past few years (Fig. 27b). Both CanESM5 and CanESM2 show consistent declines over the historical period, with CanESM2 matching the climatological extent more closely. ~~The spread of trends from the CanESM2 and CanESM5 ensembles suggest that the observed small positive trends in historical Antarctic sea-ice extent could plausibly have been due to internal climate variability (Fig. 27d).~~

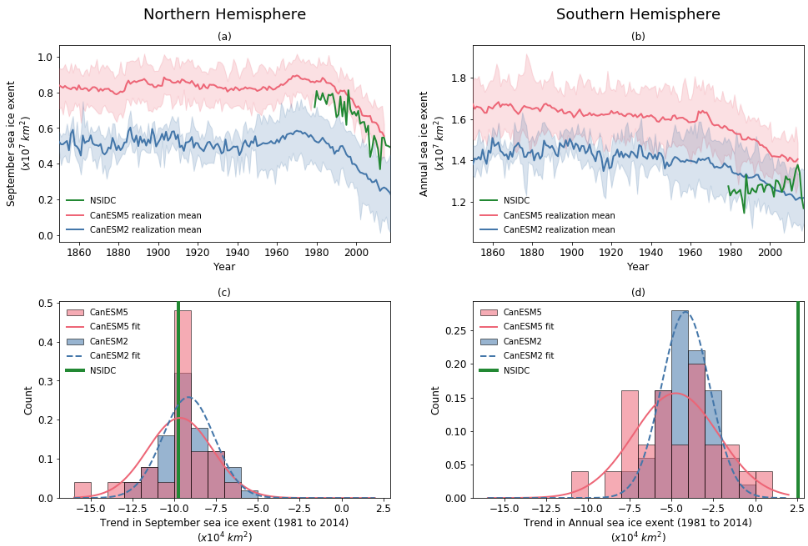


Figure 27: Time series of sea-ice extent during (a) September in the Northern Hemisphere and (b) the annual mean in the Southern Hemisphere in CanESM5, CanESM2, and NSIDC satellite based observations. The histogram of trends over 1981 to 2014 in the lower panels. The shaded envelopes represent the range over the CanESM2 50 member large ensemble and the CanESM5 25 member "p1" ensemble. Fits of the normal distribution to the CanESM2 and CanESM5 histograms are also shown.

6.2.3 Historical carbon cycle changes

The simulated global atmosphere-ocean (F_O) and atmosphere-land (F_L) CO_2 fluxes are shown if Fig. 28 for the historical period, along with their cumulative values over time. Also shown are the diagnosed anthropogenic fossil fuel emissions (E) that are consistent with the specified CO_2 pathway over the historical period, corrected for any drift in model's pre-industrial control simulation (see Appendix F). The simulated values of F_L, F_O, and E are compared against estimates from the Global Carbon Project (Le Quéré et al., 2018).

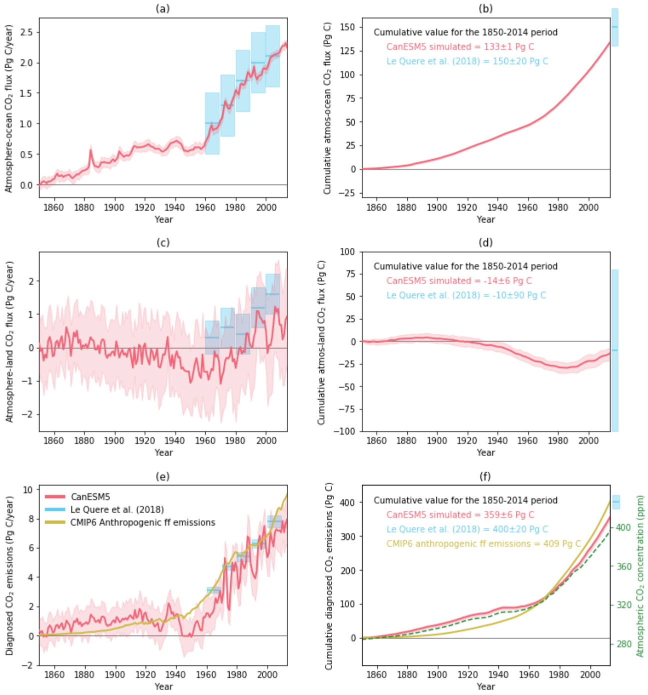


Figure 28: Annual (left column) and cumulative (right column) global values of simulated atmosphere-ocean and atmosphere-land CO2 fluxes, and diagnosed anthropogenic fossil fuel emissions, shown in blue colour. The model values are shown as mean ± 1 standard deviation range and calculated based on the 25 ensemble members of the historical simulation. Model values are compared against estimates from Le Quere et al. (2018).

In Fig. 28a the simulated global atmosphere-ocean CO_2 fluxes compares reasonably well with observation-based estimates from Le Quéré et al. (2018) for the decades of 1960s through to 2000s, although the simulated cumulative value of 133±1 Pg C for the 1850-2014 period is on the lower end of the observation-based estimate of 150±20 Pg C (Fig. 28b). In contrast, the simulated mean atmosphere-land CO_2 fluxes (Fig. 28c) for the decades of 1960s through to 2000s are lower than their observation-based estimates ~~for the decades of 1960s through to 2000s~~ from Le Quéré et al. (2018). This is despite the fact that the land carbon uptake in the 1pctCO2 simulation for CanESM5 is highest amongst all CMIP5 models reported in Arora et al. (2013). The reason for this conundrum is a topic for future investigation, but ~~their~~might relate to differences in forcing (aerosols) in the historical and 1pctCO2 experiments. The cumulative ~~value~~atmosphere-land CO_2 flux of −14±6 Pg C over the 1850-2014 period, however, compares well with the observation-based estimate of −10±90 Pg C (Fig. 28d). The caveat here, of course, is the large uncertainty range in the observation-based estimate of net cumulative atmosphere-land CO_2 flux (Appendix F). The reason the model's simulated cumulative uptake of −14±6 Pg C over the 1850-2014 period compares well with the observation-based estimate of −10±90 Pg C, despite its weaker carbon sink since the 1960s (Fig. 28, panel c) is likely because the carbon source from land use change emissions is also lower.

Panel e and f in Fig. 28 show the allowable diagnosed fossil fuel emissions and their cumulative values for the 1850-2014 period. The cumulative diagnosed fossil fuel emissions of 359±6 Pg C from the model for the period 1850-2014 are somewhat lower than the CMIP6 and Le Quéré et al. (2018) estimates of 409 and 400±20 Pg C, respectively.

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

7 Conclusions

CanESM5 is the latest coupled model from the Canadian Centre for Climate Modelling and Analysis. Relative to its predecessor, CanESM2, the model has new ocean, sea-ice and coupling components, and includes updates to the atmospheric and land surface. The model produces a stable pre-industrial control climate, and notwithstanding some significant biases, CanESM5 is able to reproduce many features of the historical climate. Objective global skill metrics show that CanESM5 improves the simulation of observed large scale climate patterns, relative to CanESM2, for most variables surveyed. A notable feature of CanESM5 is its high equilibrium climate sensitivity of 5.~~7~~6 K, an emergent property of the updated physics described above. This higher climate sensitivity appears to be driven by increased cloud and sea-ice albedo feedbacks in CanESM5. The first major science application of CanESM5 is for CMIP6, with over 50, 000 years of CanESM5 simulation and more than 100 PB of data submitted to the publicly available CMIP6 archive. The model source code is also openly published for the first time. Going forward CanESM5 will continue to be used for climate science applications in Canada.

8 Code availability

The full CanESM5 source code is publicly available at https://gitlab.com/cccma/canesm. The version of the code which can be used to produce all the simulations submitted to CMIP6, and described in this manuscript, is tagged as *v5.0.3*, and has the associated DOI: 10.5281/zenodo.3251113.

9 Data availability

All CanESM5 simulations conducted for CMIP6, including those described in this manuscript, are publicly available via the Earth System Grid Federation (ESGF). All observational data used is publicly available. Data sources and citations are provided in Appendix F.

Appendices

Appendix A: Exchanges through the coupler

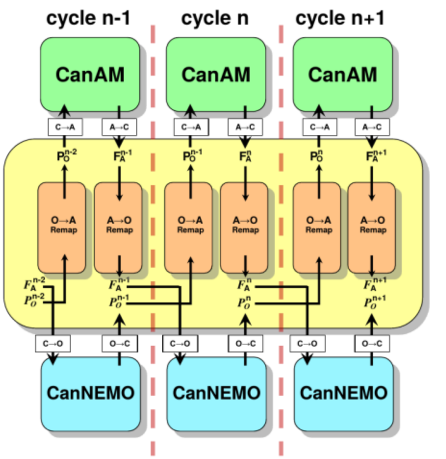


Figure A1: Schematic showing the ordering of exchanges between CanCPL and CanAM and CanNEMO. Prognostic fields (P_O) are passed from NEMO to the coupler, remapped, and passed to CanAM. Fluxes (F_A) are passed from CanAM, remapped in CanCPL, and passed NEMO to complete the next coupling cycle. Superscripts denote the coupling cycle, e.g. Prognostic fields from NEMO are passed to CanCPL at the end of cycle “n”, remapped, and used in CanAM during cycle “n+1”.

Formatted: Left, Indent: Left: 0 cm

Table A1: Fields received by CanAM from CanCPL. The representative area may be the full AGCM grid cell (land, ocean, and ice), “C”, open ocean, “O”, sea-ice, “I”, or the combination. Fields may be instantaneous, “Inst”, or averaged over the coupling cycle, “avg”.

Formatted: Left

Field Received	Field Description	Area	Avg
SICN_atm	sea ice fraction	OI	Inst
SIC_atm	ice water equivalent of sea ice	OI	Inst
SNO_atm	snow water equivalent over sea ice	I	Inst
GT_atm	sea surface temperature	O	Inst
CO2flx_atm	CO2 flux	OI	Inst

Formatted Table

Formatted: Left

Table A2 Fields sent from CanNEMO to CanCPL. Descriptions as in Table A1.

Field Sent	Field Description	Area	Avg
OIceFrc	sea ice fraction	OI	Inst
OIceTck	ice water equivalent of sea ice	OI	Inst
OSnwTck	snow water equivalent over sea ice	OI	Inst
O_SSTSST	sea surface temperature	O	Inst
O_TepIce	sea ice surface temperature	I	Inst
O_CO2FLX	CO2 flux	OI	Inst

Formatted Table

Formatted: Left

Table A3 Fields received by CanNEMO from CanCPL. Descriptions as in Table 1.

Field Received	Field Description	Area	Avg
O_OTaux1	Atm-ocn wind stress (x)	O	avg
O_OTauy1	Atm-ocn wind stress (y)	O	avg
O_ITaux1	Atm-ice wind stress (x)	I	avg
O_ITauy1	Atm-ice wind stress (y)	I	avg
O_QsrMix	solar heat flux mixed over ocean-ice	OI	avg
O_QsrIce	solar heat flux over sea ice	I	avg
O_QnsMix	non-solar heat flux mixed over ocean-ice	OI	avg
O_QnsIce	non-solar heat flux over sea ice	I	avg
OTotEvap	Total evaporation (evap + sublimation)	OI	avg
OIceEvap	sublimation over sea ice	I	avg
OTotSnow	Snow	C	avg
OTotRain	Rain	C	avg
O_dQnsdT	non-Solar sensitivity to temperature	I	avg
O_Runoff	runoff	OI	avg

Formatted Table

O_Wind10	10 meter wind	C	avg
O_TauMod	ocean wind stress modulus	O	avg
O_MSLP	Mean sea level pressure	C	avg
O_AtmCO2	atm CO2 concentration	C	avg

Formatted: Left

Table A4 Fields sent from CanAM to CanCPL. Descriptions as in Table 1.

Formatted Table

Field Sent	Field Description	Area	Avg
UFSO_atm	Atm-ocn wind stress (x)	O	avg
VFSO_atm	Atm-ocn wind stress (y)	O	avg
UFSI_atm	Atm-ice wind stress (x)	I	avg
VFSI_atm	Atm-ice wind stress (y)	I	avg
FSGO_atm	Solar heat flux over ocean	O	avg
FSGI_atm	Solar heat flux over ice	I	avg
BEGO_atm	Total heat flux over ocean	O	avg
BEGI_atm	Total heat flux over sea ice	I	avg
RAIN_atm	Total liquid precipitation	C	avg
SNOW_atm	Total solid precipitation	C	avg
BWGO_atm	ocean freshwater budget (P-E)	O	avg
BWGI_atm	sea ice fresh water budget	I	avg
SLIM_atm	non-Solar sensitivity to temperature	I	avg
RIVO_atm	River discharge	OI	avg
SWMX_atm	Mixed 10 meter wind	C	avg
PMSL_atm	Mean sea level pressure	C	avg
CO2_atm	Atm CO2 concentration	C	avg

Formatted: Left

Appendix B: Code management and model infrastructure

Table B1: Code management

Formatted Table

Item	Description

Source control	Each model component and supporting tools are version controlled in a dedicated git repository. Specific component versions are tracked as submodules by the CanESM super-repo, to define a version of CanESM.
Branching structure / workflow	Development of CanESM5 code follows a *gitflow* like workflow, commonly found in industry. Each logical unit of work is first described by an *issue*. Code changes are implemented on a dedicated feature branch. For simplicity, the feature branch is created in all submodules. Upon completion and acceptance, the feature branch is merged back onto the *develop_canesm* branch, which represents the latest state of the coupled model. Periodic tags on the *develop_canesm* branch mark stable versions of the model, which are then used for production purposes. The model version used for CMIP6 production is tagged as “CanESM.v5.0.0”, and can be used to reproduce all existing CMIP6 simulations. A series of modified git commands is used to aid in working with submodules.
Versioning	Release versions of CanESM are tagged on the *develop_canesm* branch. Tags appear as CanESM.~~vXYZ~~vX.Y.Z, where X is the major version, Y is a minor number, and Z is a bugfix level number. For example, CanESM.v5.0.2. Over the course of CMIP6 development, only bit-pattern preserving changes have been accepted.
Forcing & initialization files	Forcing and initialization files are important for reproducibility, but not directly amenable to version control. An additional repository named *CanForce* contains the source code for scripts which produced the original input files. Input files are also checksummed, and a list of these checksums is tracked in the CanESM super-repository.
External dependencies	Specific versions of third party libraries, such as NetCDF, are loaded via an initialization procedure. Third party library source code is not directly tracked.

Table B2: Process for running CanESM

Item	**Description**

Formatted: Left

Formatted: Left

Formatted Table

Run setup	Runs are setup on the ECCC HPC using a single entry point script (*setup-canesm*), which recursively clones the CanESM super-repository, and extracts some specific run configuration files. Hence, each run has a self-contained, full copy of the CanESM source code. This isolates runs from "external" changes, and also allows experimentation without affecting runs. When generating ensembles, code sharing between members is possible. *setup-canesm* also undertakes logging, recording which specific commit of CanESM was used in the run.
Run time environment	CanESM5 is run under Linux on ECCC's HPC. The user environment begins as only containing the path to *setup-canesm*. A machine-specific environment setup files is extracted from CCCma_tools by this utility script, and is sourced to define the runtime environment. The runtime environment essentially re-defines the PATH variable to point to the locally extracted scripting, as well as defining a host of machine-specific environment variables required at runtime.
Compilation	*setup-canesm* extracts utility compilation scripts. Ultimately, compilation scripts call the make utility to compile the code. The compilation of CanNEMO depends on the makenemo utility included in the source. Compilation of CanAM and CanCPL is done with makefiles, which are generated by the build-exe script, which determines required dependencies.
Configuration	CanESM runs are configured via the *canesm.cfg* file, which is extracted from the CanESM super-repo by *setup-canesm*. The configuration file allows selection of type of experiment (forcing files), start and end dates, diagnostics to be undertaken, and various options like dumping files to tape and deleting files. This configuration file is automatically captured in a dedicated configuration repository for posterity.
Sequencing	A legacy set of sequencing scripting is used to run CanESM simulations. In essence, a script called *cccjob* uses the information in *canesm.cfg* to create a sequential *string* of bash scripts, which run the model, compute diagnostics, and so forth. Such *jobstrings* are submitted to the HPC scheduler, and iterated over in sequence by a series of scripts contained in the *CCCma_tools* repository.
Strict checking	"Strict checking" is implemented during compilation, configuration, and during each increment over which the model is run when in production mode for official activities like CMIP6. Strict checking ensures that any source code changes have been committed, and that any configuration changes are captured in a dedicated repository.

Appendix C: Code optimization

Table C1: Description of optimization improvements to CanESM5. See Fig. ~~3~~1 for a graphical representation.

Description of change	**Throughput improvement (ypd)**
Several I/O heavy operations, such as splitting and repacking files that were running in serial with the model execution were switched to run in parallel on the post-processing machine. In addition, the job submission scripting was simplified.	0.4
Splitting multi-year forcing files into yearly chunks resulted in speed improvement due to the non-sequential access of the CCCma file format.	1.1
Compiler flag optimization. Specifically the "-fp-model precise" was replaced with "-mp1" flag in the final 32-bit version, and the "-init=arrays -init=zero" flags were eliminated. The optimization level was increased from "-O1" to "-O2".	1.5
Adding a node to the CanAM component to speed up spectral transforms, and implement sharing of one node with the coupler (no increase in the overall node count).	1.2
Converting CanAM from 64 to 32 bit numerics	4.1
Writing model output from different cores/tasks into separate files (labeled in Fig. ~~3~~1 as "parallel I/O"), and rebuilding them in parallel on the post-processing machine.	0.4
Changing model execution from occurring in monthly chunks (with re-initialization from restarts at the beginning of every month), to occurring in annual chunks.	2.6

Formatted: Left

Formatted Table

Formatted: Left

Formatted: Justified, Indent: Left: 0 cm

Formatted: Justified

Formatted: Left

Formatted: Justified, Indent: Left: 0 cm

Formatted: Justified

Formatted: Left

Formatted: Justified, Indent: Left: 0 cm

Formatted: Justified

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Appendix D: CMIP6 MIP participation and model variants

Table D1: List of MIPs and model variants of CanESM5 planned for submission to CMIP6.

MIP	Model variant
DECK-historical	CanESM5-p1, CanESM5-p2, CanESM5-CanOE-p2
C4MIP	CanESM5-p1, CanESM5-p2
CDRMIP	CanESM5-p1, CanESM5-p2
CFMIP	CanESM5-p2
DAMIP	CanESM5-p1
DCPP	CanESM5-p2
FAFMIP	CanESM5-p2
GeoMIP	CanESM5-p2
GMMIP	CanESM5-p2
ISMIP6	CanESM5-p1, CanESM5-p2
LS3MIP	CanESM5-p1, CanESM5-p2
LUMIP	CanESM5-p1, CanESM5-p2
OMIP	CanESM5, CanESM5-CanOE (uncoupled).
PAMIP	CanESM5-p2
RFMIP	CanESM5-p2
ScenarioMIP	CanESM5-p1, CanESM5-p2, CanESM5-CanOE-p2
VolMIP	CanESM5-p2
CORDEX	N/A (CanRCM)
DynVar	CanESM5-p2
SIMIP	CanESM5-p1, CanESM5-p2

Formatted: Left

Formatted Table

Formatted: Left

Appendix E: Comparison between p1 and p2

Sections 2.5 and 3.4 described the technical differences between perturbed physics members p1 and p2, submitted to the CMIP6 archive. Here we provide a preliminary analysis of the differences between the two model variants.

Fig. E1 shows surface air temperature and precipitation averaged over 200 years of piControl experiment, for p1, p2 and the difference between them. Notable in the differences are the “cold” spots in Antarctica, which arise from a mis-specified land fraction in p1, and were resolved in p2. Otherwise there are no significant differences.

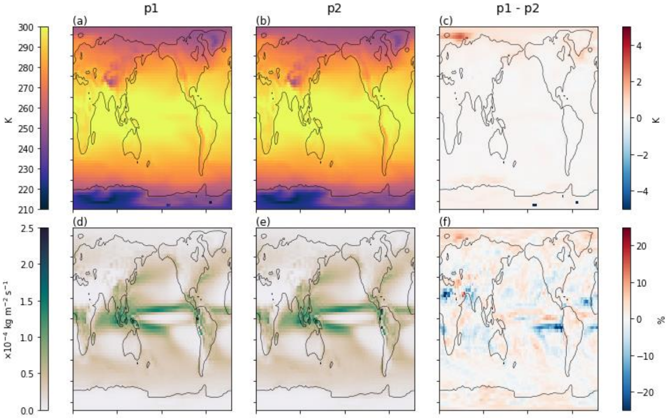


Fig. E1: Climatologies of surface air temperature (a, b) and precipitation (d, e,) computed over 200 years of piControl simulation of the p1 (a, d) and p2 (b, e) model variants, and the differences between p1 and p2 (c, f).

Fig E2 shows the ocean surface wind-stress. The blockiness of the field in p1 is evident, as a result of conservative remapping from CanAM. In p2, bilinear remapping was used and the field is smooth on the NEMO grid. The non-smooth nature of wind-stress in p1 resulted in, for example, banding in vertical ocean velocities at 100 m depth, as also shown in Fig. E2d. This does not occur in p2.

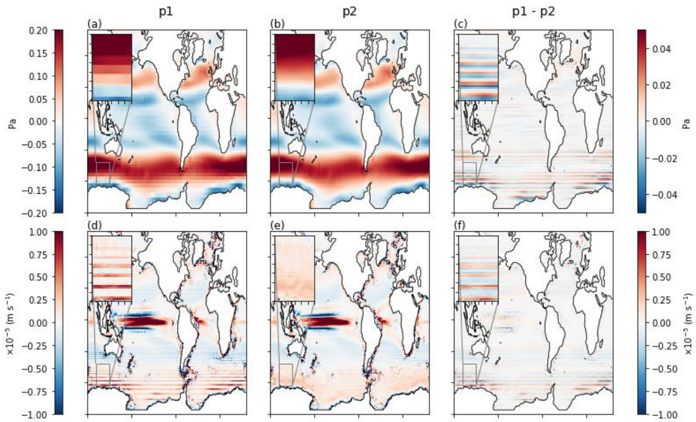


Fig. E2: Climatologies of surface ocean zonal wind stress (a, b) and vertical velocity near 100 m depth (d, e) computed over 200 years of piControl simulation of the p1(a, d) and p2 (b, e) model variants, and the differences between p1 and p2 (c, f). Results are shown on the native NEMO grid. The insets show an enlargement of the Southern Ocean south of Australia.

The response to CO2 forcing in the 1pctCO2 experiments in p1 and p2 is shown in Fig. E3. The global mean top of atmosphere radiation (Fig E3a) and surface air temperature (Fig E3b) responses are indistinguishable, and hence the TCR of these model variants is the same. The ocean is cooler, on average, in p2, but the perturbative response in p1 and p2 are similar (Fig E3c). Ocean surface CO2 flux is also statistically indistinguishable between the variants (Fig E3d).

Maps of the perturbative response, computed as the mean over the 20 years centered on CO_2 doubling in the 1pctCO2 experiments, minus the piControl, are shown in Figs. E4 and E5. There are no fundamental differences in the surface climate response between the two model variants.

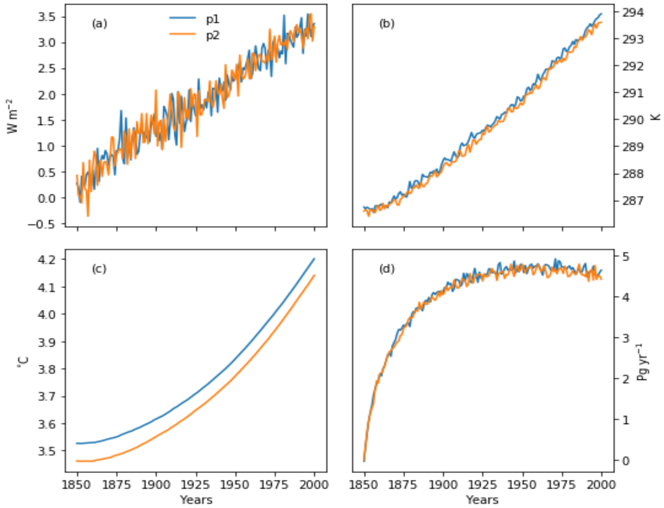


Fig. E3: Global averages of (a) top of atmosphere net radiative flux, (b) surface air temperature, (c) volume averaged ocean temperature and (d) surface ocean CO2 flux in the 1pctCO2 simulations from the p1 and p2 model variants.

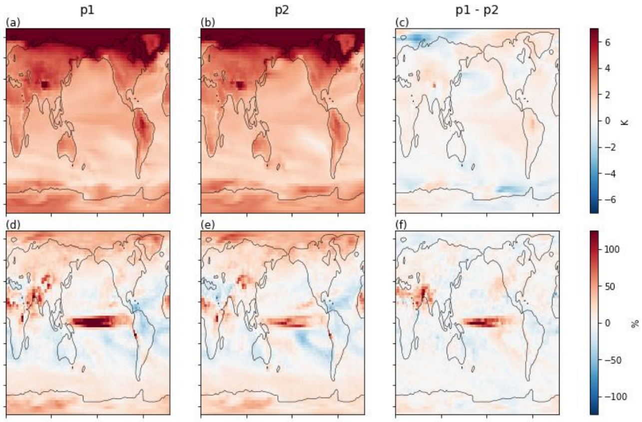


Fig. E4: Perturbation of surface air temperature (a, b) and precipitation (d, e) computed as the mean over the 20 years centered on CO_2 doubling in the 1pctCO2 experiment, minus the mean from 200 years of piControl simulation of the p1 (a, d) and p2 (b, e) model variants, and the differences between p1 and p2 (c, f).

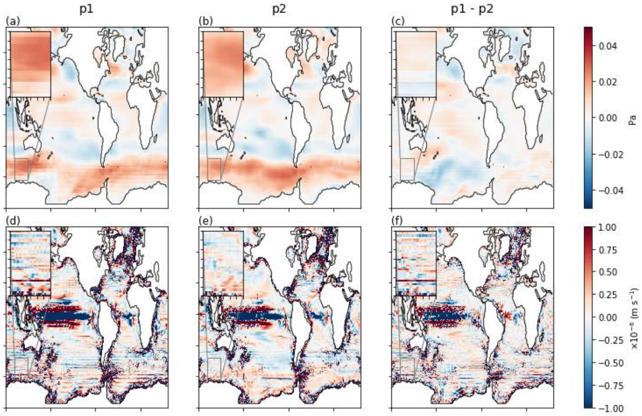


Fig. E5: Perturbations of surface ocean zonal wind stress (a, b) and vertical velocity near 100 m depth (d, e) computed as the mean over the 20 years centered on CO_2 doubling in the 1pctCO2 experiment, minus the mean over 200 years of piControl simulation of the p1 (a, d) and p2 (b, e) model variants, and the differences between p1 and p2 (c, f). Results are shown on the native NEMO grid. The insets show an enlargement of the Southern Ocean south of Australia.

Appendix F: Data sources, variables, and derived quantities

Table F1: List of figures, CanESM5 CMIP6 variables, and observations used, and the time-periods of analysis in the main text. See Table F3 for a definition of variable names.

Fig No	CMIP6 variables	CMIP6 experiment and (variant label)	Observations	Time period of analysis
~~4~~2	tas	piControl, historical, abrupt-4xCO2, 1pctCO2, SSP5-85, SSP (r1i1p1f1)	n/a	1850-2100
~~5~~3	rtmt, hfds, thetao, tas, wfo, zos, sivol, snw, nep, fgco2, cLand, dissic (piControl)	piControl (r1i1p1f1)	n/a	5200-6200
~~6~~4	As labelled.	historical (r1i1p1f1)	ERA5, GPCP, GBAF, WOA09, AVISO, GLODAPv2.2016, ISCCP-H, AVISO MDT, NSIDC, Landschützer et al. (2015)	
5	tas	historical (r1i1p1f1)	ERA5	1981-2010
6	pr	historical (r1i1p1f1)	GPCP	1981-2010
7	~~tas, pr,~~ psl	historical (r1i1p1f1)	ERA5~~, GPCP~~	1981-2010
8	clt	historical (r1i1p1f1)	ISCCP-H	1991-2010 (ISCCP data period)
9	ta	historical (r1i1p1f1)	ERA5	1981-2010
10	ua	historical (r1i1p1f1)	ERA5	1981-2010
11	uas	historical (r1i1p1f1)	ERA5	1981-2010
12	gpp, hfls, and hfss	historical (r1i1p1f1)	GBAF	1982-2009 (GBAF data period)

13	gpp, hfls, and hfss	historical (r1i1p1f1)	GBAF	1982-2009 (GBAF data period)
14	gpp, tas, pr	historical (r1i1p1f1)	GBAF	1982-2009 (GBAF data period)
15	tos, sos, zos	historical (r1i1p1f1)	WOA09, AVISO MDT	1981-2010
16	thetao, so	historical (r1i1p1f1)	WOA09	1981-2010
17	msftmz	historical (r1i1p1f1)	-	1981-2010
18	hfbasin	historical (r1i1p1f1)	Ganachaud and Wunsch (2003)	1981-2010
19	siconc, sithick	historical (r1i1p1f1 to r25i1p1f1)	NSIDC	1981-2010
20	siconc	historical (r1i1p1f1)	NSIDC	1981-2010
21	dissic, no3, o2	historical (r1i1p1f1)	GLODAPv2.2016	1981-2010
22	fgco2	historical (r1i1p1f1)	Landschützer et al. (2015)	1982-2010 (Landschützer data period)
23	tos	historical (r1i1p1f1 to r10i1p1f1)	HadISST	1850-2014
24	psl	historical (r1i1p1f1)	ERA5	1981-2010
25	tas	historical (r1i1p1f1 to r25i1p1f1)	Berkeley-Earth, HadCRUT4, NASA-GISS	Time series: 1850-2014 Trends: 1981-2014
26	tas	historical (r1i1p1f1 to r25i1p1f1)	Berkeley-Earth, HadCRUT4, NASA-GISS	1981-2014
27	siconc	historical (r1i1p1f1 to r25i1p1f1)	NSIDC	Time series: 1850-2014 Trends: 1981-2014
28	fgco2, nep, co2	historical (r1i1p1f1 to r25i1p1f1)	Le Quéré et al. (2018)	1850-2014

Formatted: Left

Table F2: List of observational products used.

Data source	Citation
AVISO MDT	https://www.aviso.altimetry.fr/en/data/products/auxiliary-products/mdt.html
ERA5	Copernicus Climate Change Service (2017)
GPCP	Adler et al. (2017)
ISCCP-H	Young et al. (2018); Rossow et al. (2016)
GBAF	Jung et al. (2009)
World Ocean Atlas 2009 (WOA09)	Locarnini et al. (2009); Antonov et al. (2010)
NSIDC sea-ice concentration	Peng et al. (2013); Meier et al. (2017)
PIOMAS	Zhang and Rothrock (2003); Schweiger et al. (~~2003~~2011)
GIOMAS	Zhang ~~et al.~~and Rothrock (2003~~)~~)); http://psc.apl.washington.edu/zhang/Global_seaice/data.html
GLODAPv2	Lauvset et al. (2016)
Landschützer	Landschützer et al. (2015)
HadISST	Rayner et al. (2003)
Berkley Earth	http://berkeleyearth.org/land-and-ocean-data/
HadCRUT4	Morice et al. (2012)
NASA-GISS	GISSTEMP Team (2019); Lenssen et al. (2019)
Global Carbon Budget 2018	Le Quéré et al. (2018)

Table F3: List of CMIP6 variable names used, and their long names.

cLand	Total Carbon in All Terrestrial Carbon Pools
clt	Total Cloud Cover Percentage
co2	Mole Fraction of CO2
dissic	Dissolved Inorganic Carbon Concentration
fgco2	Surface Downward Flux of Total CO2
gpp	Carbon Mass Flux out of Atmosphere Due to Gross Primary Production on Land
hfbasin	Northward Ocean Heat Transport
hfds	Downward Heat Flux at Sea Water Surface
hfls	Surface Upward Latent Heat Flux
hfss	Surface Upward Sensible Heat Flux
msftmz	Ocean Meridional Overturning Mass Streamfunction
nep	Net Carbon Mass Flux out of Atmosphere due to Net Ecosystem Productivity on Land
no3	Dissolved Nitrate Concentration
o2	Dissolved Oxygen Concentration

Formatted Table

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: English (Canada)

Formatted: English (Canada)

Formatted: English (Canada)

Formatted: Left

Formatted: English (Canada)

Formatted: English (Canada)

Formatted: English (Canada)

Formatted: Left

Formatted Table

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

Formatted: Left

pr	Precipitation
psl	Sea Level Pressure
rtmt	Net Downward Radiative Flux at Top of Model
siconc	Sea-ice Area Percentage (Ocean Grid)
sithick	Sea Ice Thickness
sivol	Sea-Ice Volume per Area
snw	Surface Snow Amount
sos	Sea Surface Salinity
ta	Air Temperature
tas	Near-Surface Air Temperature
thetao	Sea Water Potential Temperature
tos	Sea Surface Temperature
ua	Eastward Wind
uas	Eastward Near-Surface Wind
va	Northward Wind
wfo	Water Flux into Sea Water
zg	Geopotential Height
zos	Sea Surface Height Above Geoid

In Figure 28 the diagnosed allowable anthropogenic fossil fuel emissions are calculated via Equation (F1):

$$\frac{d\,[CO2]}{dt} = E - F_L - F_O = E - (F'_L - E_{LUC}) - F_O \qquad \text{(F1).}$$

In these historical simulations, the concentration of atmospheric CO_2 is specified (that is the term $d[CO_2]/dt$ is known) and the model's land and ocean carbon cycle components simulate atmosphere-land (F_L) and atmosphere-ocean (F_O) CO_2 fluxes, respectively. The F_L=F'_L-E_{LUC} term includes natural atmosphere-land CO_2 flux (F'_L) and the emissions associated with land use change (E_{LUC}) which are calculated interactively in the model in response to the historical increase in cropland area. As a result, the term E can be calculated and represents the allowable anthropogenic fossil fuel emissions.

Le Quéré et al. (2018) do not provide a direct value of net cumulative atmosphere-land CO2 flux (F_L). Instead, they separately provide estimates of cumulative values of F'_L (185±50 Pg C) and E_{LUC} (195±75) in their Table 8. Here, we calculate observation-based value of F_L=F'_L-E_{LUC} =185−195 = −10 Pg C and its uncertainty as 90 Pg C (the uncertainty is calculated as, $\sqrt{(50^2+75^2)} = 90.13$ PgC. The large uncertainty range for the observation-based estimate of cumulative F_L is therefore due to large uncertainties in both land use change emissions and the natural atmosphere-land CO_2 flux.

Formatted: English (Canada)

Formatted: Left

Appendix G: Wetland extent and wetland methane emissions

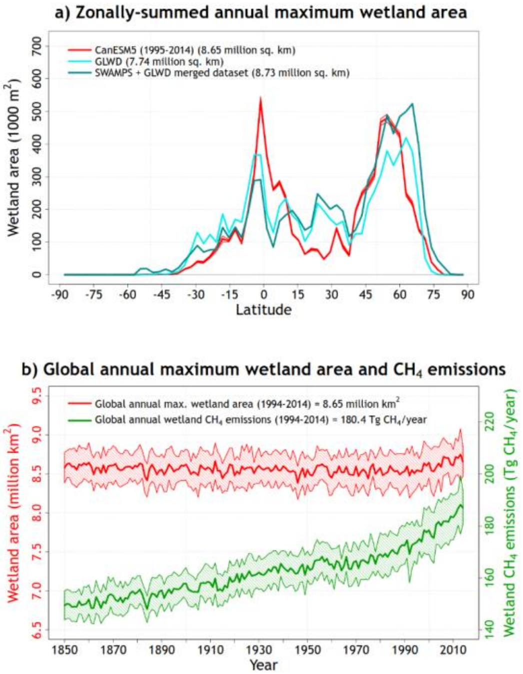


Figure G1: Comparison of simulated zonally-summed annual maximum wetland area with observation-based estimates based on the Global Lakes and Wetland (GLWD; Lehner and Döll, 2004) and a new product that is formed by merging remote sensing based observations of daily surface inundation from the Surface Water Microwave Product Series (SWAMPS; Schroeder et al., 2015) with the static inventory of wetland area from the GLWD as explained in Poulter et al. (2017) (panel a). Panel (b) shows the temporal evolution of simulated global annual maximum wetland extent and wetland methane emissions over the 1850-2014 period of the historical simulation.

Figure G1, panel a, compares the zonally-summed annual maximum wetland area with observation-based estimates based on the Global Lakes and Wetland (GLWD; Lehner and Döll, 2004) and the new product formed by merging remote sensing based observations of daily surface inundation from the Surface Water Microwave Product Series (SWAMPS; Schroeder et al., 2015) with the static inventory of wetland area from the GLWD from Poulter et al. (2017). Maximum wetland fraction from the model (1995-2014) and SWAMPS+GLWD (2000-2012) product is calculated as the maximum of 12 mean monthly values for the time period noted and the multiplied by area of the grid cells to calculate zonally-summed area. The model overall captures the broad latitudinal distribution of wetlands with higher wetland area in the tropics and at northern high-latitudes. The model yields higher wetland area in the tropics than both observation-based estimates due to higher wetland area simulated in the Amazonian region. The Amazonian region is densely forested and the SWAMPS product is unable to map wetlands beneath closed forest canopies. Biases also likely exist in the GLWD data set since parts of the Amazonian region are fairly remote. The global annual maximum wetland extent of 8.65 million km^2 also compares well with observation-based estimates of 7.74 (GLWD) and 8.73 (SWAMPS+GLWD) million km^2.

Panel b of Figure G1 shows the time evolution of simulated annual maximum wetland extent and wetland methane emissions over the 1850-2014 period from the historical simulation. The shaded range represents the standard deviation over the 25 ensemble members of the historical simulation. While the simulated wetland extent does not change significantly over time, the methane emissions increase from about 150 Tg CH_4/yr during the early pre-industrial period to about 180 Tg CH_4/yr for the present day (1995-2014). This increase in wetland methane emissions is caused by increased vegetation productivity driven by increased atmospheric CO_2 concentration over the historical period. The simulated present day wetland methane emissions of 180 Tg CH_4/yr are comparable to central estimate of 166 Tg CH_4/yr from Saunois et al. (2016) and their range of 125-204 Tg CH_4/yr.

Author contributions

Formatted: Left

NCS co-led CanESM5 development, contributed to CanNEMO and CMOC development and the data request, performed simulations, led the creation of the figures and wrote most of the manuscript; JNSC contributed to development of CanAM5, CanCPL and tuning of CanESM5, wrote the CanAM5 section, performed simulations and contributed to the data request; VK contributed to the development of CanAM, notably optimization, contributed to the data request, and performed production simulations; ML contributed to the development of CanAM, CanCPL and the data request; JS co-led CanESM5 development; NG contributed to CanESM5 development and tuning; JA contributed to CanCPL development, the data request, and led publication of data on the ESGF; VA contributed to the development of CLASS and CTEM; JC developed CanOE and contributed to CMOC development; SH produced many of the figures; YJ contributed to the data request and conversion; WL contributed to CanNEMO development and ran simulations; FM contributed to the CanESM5 software infrastructure; OS led ocean physics testing; provided a specific analysis that motivated the p2 variant; ChS contributed analysis of the land component; ClS contributed to CanESM5 software infrastructure; AS created Fig. 23, contributed to CanESM5 development, and performed simulations; LS developed CanCPL; KVS led the development and tuning of atmospheric model parameterizations; DY contributed to ocean model development, ocean and sea ice diagnostics for CMIP6, and performed production simulations; BW processed forcing datasets for CanAM; All authors contributed to editing the manuscript.

Competing interests

No competing interests.

Disclaimer

CanESM has been customized to run on the ECCC high performance computer, and a significant fraction of the software infrastructure used to run the model is specific to the individual machines and architecture. While we publicly provide the code, we cannot provide any support for migrating the model to different machines or architectures.

Special issue statement (will be included by Copernicus)

To be included in the CanESM5 special issue.

Acknowledgements

We acknowledge Dr. Michael Sigmond, Dr. Greg Flato and Dr. William Merryfield for helpful comments on a draft of the paper. CanESM5 was the cumulative result of work by many individuals, who we thank for their contributions. CanESM5 simulations were performed on ECCC's HPC, and CanESM5 data is served via the Earth System Grid Federation.

References

Adler RF, Sapiano M, Huffman GJ, et al.: The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. Atmosphere (Basel), 2018;9(4):138. doi:10.3390/atmos9040138, 2018.

Antonov, J. I., Seidov, D., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Garcia, H. E., Baranova, O. K., Zweng, M. M., and Johnson, D. R.: World Ocean Atlas 2009, Volume 2: Salinity. S. Levitus, Ed. NOAA Atlas NESDIS 69, U.S. Government Printing Office, Washington, D.C., 184 pp, 2010.

Arora, V.K.: Simulating energy and carbon fluxes over winter wheat using coupled land surface and terrestrial ecosystem models~~.~~, Agricultural and Forest Meteorology~~.~~, 118(1-2), 21-47, 2003.

Arora, V.K., and Boer, G.J.: A variable velocity flow routing algorithm for GCMs, J. Geophys. Res., 104, D24, 30965-30979, 1999.

Arora, V.K. and Boer, G.J.: A representation of variable root distribution in dynamic vegetation models. Earth Interactions, Vol. 7, Paper 6, 19 pp, 2003.

Arora, V.K. and Boer, G.J.: A parameterization of leaf phenology for the terrestrial ecosystem component of climate models, Global Change Biology. 11(1), 39-59, 2005.

Arora, V. K., Boer, G. J., Christian, J. R., Curry, C. L., Denman, K. L., Zahariev, K., Flato, G. M., Scinocca, J. F., Merryfield, W. J. and Lee, W. G.: The effect of terrestrial photosynthesis down-regulation on the 20th century carbon budget simulated with the CCCma Earth System Model, J. Clim. 22, 6066-6088, 2009.

Arora, V.K. and Boer, G.J.: Uncertainties in the 20th century carbon budget associated with land use change, Global Change Biology, 16(12), 3327-3348, 2010.

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V., Lee, W. G., and Merryfield, W. J.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophys. Res. Lett., 38, L05805, doi:10.1029/2010GL046270, 2011.

Arora, V. K. and Scinocca, J. F.: Constraining the strength of the terrestrial CO_2 fertilization effect in the Canadian Earth system model version 4.2 (CanESM4.2), Geosci. Model Dev., 9, 2357-2376, https://doi.org/10.5194/gmd-9-2357-2016, 2016.

Arora, V. K., Melton, J. R., and Plummer, D.: An assessment of natural methane fluxes simulated by the CLASS-CTEM model, Biogeosciences, 15, 4683-4709, https://doi.org/10.5194/bg-15-4683-2018, 2018.

Bartholomé, E. and Belward, A. S.: GLC2000: a new approach to global land cover mapping from Earth observation data, International Journal of Remote Sensing, 26:9, 1959-1977, DOI: 10.1080/01431160412331291297, 2005.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjánsson, J. E.: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate, Geosci. Model Dev., 6, 687-720, https://doi.org/10.5194/gmd-6-687-2013, 2013.

Blanke, B. and P. Delecluse: Variability of the Tropical Atlantic Ocean Simulated by a General Circulation Model with Two Different Mixed-Layer Physics, J. Phys. Oceanogr., 23, 1363–1388, https://doi.org/10.1175/1520-0485(1993)023<1363:VOTTAO>2.0.CO;2, 1993.

Boer, G. J., and McFarlane, N. A.: The AES atmospheric general circulation model. Report of the JOC Study Conference on Climate Models: Performance, Intercomparison and Sensitivity Studies, Vol. I, GARP Publ. Ser. No. 22, pp. 409–460, 1979.

Boer, G. J., McFarlane, N. A., Laprise, R., Henderson, J. D., and Blanchet, J.–P.: The Canadian Climate Centre Spectral Atmospheric General Circulation Model, Atmos.–Ocean, 22(4), 397–429, 1984.

Boer, G.J., Flato, G.M., Reader, M.C., and Ramsden, D.: A transient climate change simulation with historical and projected greenhouse gas and aerosol forcing: experimental design and comparison with the instrumental record for the 20th century~~.~~, Climate Dynamics, 16, 405-425, 2000a.

Boer, G.J., Flato, G.M, and Ramsden, D.: A transient climate change simulation with historical and projected greenhouse gas and aerosol forcing: projected climate for the 21st century~~.~~, Climate Dynamics, 16, 427-450, 2000b.

Bouillon, S., Morales Maqueda, M.A., Legat, V., and Fichefet, T.: An elastic-viscous-plastic sea ice model formulated on Arakawa B and C grids~~.~~, Ocean Modelling, 27, 174-184, doi : 10.1016/j.ocemod.2009.01.004, 2009.

Checa-Garcia, R., Hegglin, M. I., Kinnison, D., Plummer, D. A., and Shine, K. P.: Historical tropospheric and stratospheric ozone radiative forcing using the CMIP6 database, Geophysical Research Letters, 45, 3264–3273. https://doi.org/10.1002/2017GL076770, 2018.

Cole et al.: The Canadian Atmospheric Model version 5. GMD special issue on CanESM5, in preparation, 2019.

Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), 2019/05/15. https://cds.climate.copernicus.eu/cdsapp#!/home

Chelton, D. B., Schlax, M. G., and Samelson, R. M.: Global observations of nonlinear mesoscale eddies. Prog. Oceanogr., 91, 167–216, https://doi.org/10.1016/j.pocean.2011.01.002, 2011.

Christian, J. R., et al.: The global carbon cycle in the Canadian Earth system model (CanESM1): Preindustrial control simulation, J. Geophys. Res., 115, G03014, doi:10.1029/2008JG000920, 2010.

Christian, J.R. et al: The Canadian Ocean Ecosystem, GMD special issue on CanESM5, in preparation, 2019.

Collins, N., Theurich, G., DeLuca, C., Suarez, M., Trayanov, A., Balaji, V., Li, P., Yang, W., Hill, C., and da Silva, A.: Design and Implementation of Components in the Earth System Modeling Framework~~.~~, International Journal of High Performance Computing Applications, Vol. 19, No. 3, pp. 341-350, 2005.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W.J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A.J. and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds~~.)].~~.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Danabasoglu, G., Yeager, S.G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Böning, C., Bozec, A., Canuto, V.M., Cassou, C., Chassignet, E., Coward, A.C., Danilov, S., Diansky, N., Drange, H., Farneti, R., Fernandez, E., Fogli, P.G., Forget, G., Fujii, Y., Griffies, S.M., Gusev, A., Heimbach, P., Howard, A., Jung, T., Kelley, M., Large, W.G., Leboissetier, A., Lu, J., Madec, G., Marsland, S.J., Masina, S., Navarra, A., Nurser, A.J.G., Pirani, A., y Mélia, D.S., Samuels, B.L, Scheinert, M., Sidorenko, D., Treguier, A.M., Tsujino, H., Uotila, P., Valcke, S., Voldoire, A., Wang, Q.: North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II).Part I: Mean states, Ocean Modelling, 73, 76-107, 2014.

Donohue K.A, Tracey K.L, Watts D.R, Chidichimo M.P, Chereskin TK: Mean Antarctic Circumpolar Current transport measured in Drake Passage, Geophysical Research Letters, 43, 11,760-11,767, 2016.

ESMF Joint Specification Team: ESMF Reference Manual for Fortran. http://www.earthsystemmodeling.org/esmf_releases/public/last/ESMF_refdoc/ESMF_refdoc.html, 2018.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fichefet, T., and Morales Maqueda, M.A.: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics~~.~~, Journal Geophs. Res., 102, 12609-12646, https://doi.org/10.1029/97JC00480, 1997.

Flato, G.M., Boer, G.J., Lee, W.G., McFarlane, N.A., Ramsden, D., Reader, M.C., and Weaver, A.J.: The Canadian Centre for Climate Modelling and Analysis Global Coupled Model and its Climate. Climate Dynamics, 16, 451-467, 2000.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C. and Rummukainen, M.: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., and Zelinka, M.: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models, J. Geophys. Res. Atmos., 118, 1139– 1150, doi:10.1002/jgrd.50174, 2013.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., Andrews, T., Pincus, R., and Schulz, M.: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, J. Geophys. Res. Atmos., 121, 12,460– 12,475, doi:10.1002/2016JD025320, 2016.

Ganachaud, A. and Wunsch, C.: Large-Scale Ocean Heat and Freshwater Transports during the World Ocean Circulation Experiment~~.~~, J. Climate, 16, 696–705, https://doi.org/10.1175/1520-0442(2003)016<0696:LSOHAF>2.0.CO;2, 2003.

Gaspar, P., Grégoris, Y., and Lefevre, J.-M.: A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at station Papa and long-term upper ocean study site~~.~~, J. Geophys. Res., 95, 16179–16193, https://doi.org/10.1029/JC095iC09p16179, 1990.

Gent, P. R., and McWilliams, J. C.: Isopycnal mixing in ocean general circulation models~~.~~, J. Phys. Oceanogr., 20, 150–155, https://doi.org/10.1175/1520-0485(1990)020,0150:IMIOCM.2.0.CO;2, 1990.

Gillett, N.P., Arora, V.K., Matthews, D., and Allen, M.R.: Constraining the Ratio of Global Warming to Cumulative CO_2 Emissions Using CMIP5 Simulations~~.~~, J. Climate, 26, 6844–6858, https://doi.org/10.1175/JCLI-D-12-00476.1, 2013.

Formatted: author, Font color: Auto

Formatted: Font color: Auto

GISTEMP Team: GISS Surface Temperature Analysis (GISTEMP), version 4~~.~~, NASA Goddard Institute for Space Studies~~.~~, Dataset accessed 2019-04-07 at https://data.giss.nasa.gov/gistemp/, 2019.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophys. Res. Lett., 31, L03205, doi:10.1029/2003GL018747, 2004.

Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren, A. J., and Hunke, E. C.: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, Geosci. Model Dev., 4, 223-253, https://doi.org/10.5194/gmd-4-223-2011, 2011.

Hill, C., DeLuca, C., Balaji, V., Suarez, M., and da Silva, A.: Architecture of the Earth System Modeling Framework~~.~~, Computing in Science and Engineering, Vol. 6, No. 1, pp. 18-28, 2004.

Hurtt, G.C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Goldewijk, K.K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P. and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands~~.~~, Climatic Change 109:117–161 DOI 10.1007/s10584-011-0153-2, 2011.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (eds~~.)].~~.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, doi:10.1017/CBO9781107415324, 2013.

Jung, M., Reichstein, M., and Bondeau, A.: “Towards Global Empirical Upscaling of FLUXNET Eddy Covariance Observations: Validation of a Model Tree Ensemble Approach Using a Biosphere Model~~.~~”, Biogeosciences 6 (10): 2001–13., 2009

Kay, J.E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J.M., Bates, S.C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability~~.~~, Bull. Amer. Meteor. Soc., 96, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1, 2015.

Kirchmeier-Young, M.C., Zwiers, F.W., and Gillett, N.P.: Attribution of Extreme Events in Arctic Sea Ice Extent. J. Climate, 30, 553–571, https://doi.org/10.1175/JCLI-D-16-0412.1, 2017.

Landschützer, P., Gruber, N. and Bakker, D.C.E.: A 30 year observation-based global monthly gridded sea surface pCO2 product from 1982 through 2011 http://cdiac.ornl.gov/ftp/oceans/spco2_1982_2011_ETH_SOM-FFN. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee. doi: 10.3334/CDIAC/OTG.SPCO2_1982_2011_ETH_SOM-FFN, 2015.

Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the 1° × 1° GLODAP version 2, Earth Syst. Sci. Data, 8, 325-340, https://doi.org/10.5194/essd-8-325-2016, 2016.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296(1–4), 1–22, doi:10.1016/j.jhydrol.2004.03.028, 2004.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth Syst. Sci. Data, 10, 2141-2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

Lenssen, N., Schmidt, G., Hansen, J., Menne, M., Persin, A., Ruedy, R., and Zyss, D.: ~~Improvements in the uncertainty model in the Goddard Institute for Space Studies Surface Temperature (GISTEMP) analysis. J.~~Improvements in the uncertainty model in the Goddard Institute for Space Studies Surface Temperature (GISTEMP) analysis, J. Geophys. Res. Atmos., in press, doi:10.1029/2018JD029522, 2019.

Lin, J.: The Double-ITCZ Problem in IPCC AR4 Coupled GCMs: Ocean–Atmosphere Feedback Analysis, J. Climate, 20, 4497–4525, https://doi.org/10.1175/JCLI4272.1, 2007.

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., and Johnson, D. R.. World Ocean Atlas 2009, Volume 1: Temperature. S. Levitus, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp, 2010.

Matthews, H.D., Gillett, N.P., Stott, P.A. and Zickfeld, K~~..~~.: The proportionality of global warming to cumulative carbon emissions~~.~~, Nature, 459 (7248), 829-832, 2009.

Madec, G., and Coauthors: NEMO ocean engine, version 3.4~~.~~, Institut Pierre-Simon Laplace Note du Pole de Modélisation 27, 367 pp., 2012.

Marshall, D., and Naveira Garabato, A.: A conjecture on the role of bottom-enhanced diapycnal mixing in the parameterization of geostrophic eddies~~.~~, J. Phys. Oceanogr., 38, 1607–1613, 2008.

McCarthy, G.D., Smeed, D.A., Johns, W.E., Frajka-Williams, E., Moat, B.I., Rayner, D., Baringer, M.O., Meinen, C.S., Collins, J. and Bryden, H.L.: Measuring the Atlantic Meridional Overturning Circulation at 26°N, Progress in Oceanography,Vol. 130, 91-111, https://doi.org/10.1016/j.pocean.2014.10.006, 2015.

McFarlane, N.A., Boer, G.J., Blanchet, J.-P., and Lazare, M.: The Canadian Climate Centre Second-Generation General Circulation Model and Its Equilibrium Climate~~.~~, J. of Climate, 5, 1013-1044, 1992.

Meier, W. N., Fetterer, F., Savoie, M., Mallory, S., Duerr, R., and Stroeve, J.: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 3~~.~~, Boulder, Colorado USA~~.~~, NSIDC: National Snow and Ice Data Center. doi: https://doi.org/10.7265/N59P2ZTG~~.~~, Last accessed 2019/04/17, 2017.

Merryfield, W. J., Holloway, G., and Gargett, A.E.: A global ocean model with double-diffusive mixing, J. Phys. Ocean., 29, 1124–1142, 1999.

Merryfield, W.J., Lee, W., Boer, G.J., Kharin, V.V., Scinocca, J.F., Flato, G.M., Ajayamohan, R.S., Fyfe, J.C., Tang, Y., and Polavarapu, S.: The Canadian Seasonal to Interannual Prediction System. Part I: Models and Initialization~~.~~, Mon. Wea. Rev., 141, 2910–2945, https://doi.org/10.1175/MWR-D-12-00216.1, 2013.

Formatted: English (Canada)

Formatted: English (Canada)

Morice, C.P., Kennedy, J.J., Rayner, N.A. and Jones, P.D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 dataset~~.~~, Journal of Geophysical Research, 117, D08101, doi:10.1029/2011JD017187, 2012.

Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., Doney, S. C., Dunne, J. P., Dutay, J.-C., Graven, H., Griffies, S. M., John, J. G., Joos, F., Levin, I., Lindsay, K., Matear, R. J., McKinley, G. A., Mouchet, A., Oschlies, A., Romanou, A., Schlitzer, R., Tagliabue, A., Tanhua, T., and Yool, A.: Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP), Geosci. Model Dev., 10, 2169-2199, https://doi.org/10.5194/gmd-10-2169-2017, 2017.

Peng, G., Meier, W. N., Scott, D., and Savoie, M.: A long-term and reproducible passive microwave sea ice concentration data record for climate studies and monitoring, Earth Syst. Sci. Data. 5. 311-318. https://doi.org/10.5194/essd-5-311-2013, 2013.

Poulter, B., Bousquet, P., Canadell, J. G., Ciais, P., Peregon, A., Marielle Saunois, Arora, V. K., Beerling, D. J., Brovkin, V., Jones, C. D., Joos, F., Nicola Gedney, Ito, A., Kleinen, T., Koven, C. D., McDonald, K., Melton, J. R., Peng, C., Shushi Peng, Prigent, C., Schroeder, R., Riley, W. J., Saito, M., Spahni, R., Tian, H., Lyla Taylor, Viovy, N., Wilton, D., Wiltshire, A., Xu, X., Zhang, B., Zhang, Z. and Zhu, Q.: Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics, Environ. Res. Lett., 12(9), 094013, 2017.

Rayner, N. A.; Parker, D. E.; Horton, E. B.; Folland, C. K.; Alexander, L. V.; Rowell, D. P.; Kent, E. C.; Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res~~.~~., Vol. 108, No. D14, 4407 10.1029/2002JD002670, 2003.

Redi, M. H.: Oceanic isopycnal mixing by coordinate rotation~~.~~, J. Phys. Oceanogr., 12, 1154–1158, https://doi.org/10.1175/1520-0485(1982)012,1154:OIMBCR.2.0.CO;2., 1982.

Rosenblum, E. and Eisenman, I.: Sea Ice Trends in Climate Models Only Accurate in Runs with Biased Global Warming~~.~~, J. Climate, 30, 6265–6278, https://doi.org/10.1175/JCLI-D-16-0455.1, 2017.

Rossow, W.B.; Walker, A.; Golea, V.; Knapp, K. R.; Young, A.; Inamdar A.; Hankins, B.; and NOAA's Climate Data Record Program: International Satellite Cloud Climatology Project Climate Data Record, H-Series v1.00 NOAA National Centers for Environmental Information~~.~~, 10 May 2019. doi:10.7289/V5QZ281S, 2016.

Saenko, O.A., Zhai, X., Merryfield, W., and Lee, W.: The combined effect of tidally and eddy-driven diapycnal mixing on the large-scale ocean circulation~~.~~, J. Phys. Oceanogr., 42, 526–538, ~~https://doi.org/10.1175/JPO-D-11-0122.1, 2012.~~ https://doi.org/10.1175/JPO-D-11-0122.1, 2012.

Saenko, O.A., Yang, D., and Gregory, J.M.: Impact of mesoscale eddy transfer on heat uptake in an eddy-parameterizing ocean model~~.~~, J. Climate., 31, 8589-8606, https://doi.org/10.1175/JCLI-D-18-0186.1, 2018.

~~Swart, N.C., Gille, S.T., Fyfe, J.C. and Gillett, N.P.: Recent Southern Ocean warming and freshening driven by greenhouse gas emissions and ozone depletion. Nature Geoscience, 11, 836-841, 2018.~~

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J., Cole, J. N. S., Plummer, D., Verseghy, D., Reader, M. C., Ma, X., Lazare, M. and Solheim, L.: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of Physical Processes, Atmosphere-Ocean, 51:1, 104-125, DOI: 10.1080/07055900.2012.755610, 2013.

Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J. G., Dlugokencky, E. J., Etiope, G., Bastviken, D., Houweling, S., Janssens-Maenhout, G., Tubiello, F. N., Castaldi, S., Jackson, R. B., Alexe, M., Arora, V. K., Beerling, D. J., Bergamaschi, P., Blake, D. R., Brailsford, G., Brovkin, V., Bruhwiler, L., Crevoisier, C., Crill, P., Covey, K., Curry, C., Frankenberg, C., Gedney, N., Höglund-Isaksson, L., Ishizawa, M., Ito, A., Joos, F., Kim, H.-S., Kleinen, T., Krummel, P., Lamarque, J.-F., Langenfelds, R., Locatelli, R., Machida, T., Maksyutov, S., McDonald, K. C., Marshall, J., Melton, J. R., Morino, I., Naik, V., O'Doherty, S., Parmentier, F.-J. W., Patra, P. K., Peng, C., Peng, S., Peters, G. P., Pison, I., Prigent, C., Prinn, R., Ramonet, M., Riley, W. J., Saito, M., Santini, M., Schroeder, R., Simpson, I. J., Spahni, R., Steele, P., Takizawa, A., Thornton, B. F., Tian, H., Tohjima, Y., Viovy, N., Voulgarakis, A., van Weele, M., van der Werf, G. R., Weiss, R., Wiedinmyer, C., Wilton, D. J., Wiltshire, A., Worthy, D., Wunch, D., Xu, X., Yoshida, Y., Zhang, B., Zhang, Z. and Zhu, Q.: The global methane budget 2000–2012, Earth Syst Sci Data, 8(2), 697–751, doi:10.5194/essd-8-697-2016, 2016.

Schurer, A., Hegerl, G., Ribes, A., Polson, D., Morice, C., and Tett, S.: Estimating the Transient Climate Response from Observed Warming~~.~~, J. Climate, 31, 8645–8663, https://doi.org/10.1175/JCLI-D-17-0717.1, 2018.

Schroeder, R., McDonald, C. K., Chapman, D. B., Jensen, K., Podest, E., Tessler, D. Z., Bohn, J. T. and Zimmermann, R.: Development and Evaluation of a Multi-Year Fractional Surface Water Data Set Derived from Active/Passive Microwave Remote Sensing Data, Remote Sens., 7(12), doi:10.3390/rs71215843, 2015.

Scinocca, J. F., McFarlane, N. A., Lazare, M., Li, J., and Plummer, D.: Technical Note: The CCCma third generation AGCM and its extension into the middle atmosphere, Atmos. Chem. Phys., 8, 7055-7074, https://doi.org/10.5194/acp-8-7055-2008, 2008.

Scinocca, J.F., Kharin, V.V., Jiao, Y., Qian, M.W., Lazare, M., Solheim, L., Flato, G.M., Biner, S., Desgagne, M., and Dugas, B.: Coordinated Global and Regional Climate Modeling~~.~~, J. Climate, 29, 17–35, https://doi.org/10.1175/JCLI-D-15-0161.1, 2016.

Sheen, K., and Coauthors: Eddy-induced variability in Southern Ocean abyssal mixing on climatic timescales~~.~~, Nat. Geosci., 7, 577–582, doi:https://doi.org/10.1038/ngeo2200, 2014.

Simmons, H. L., Jayne, S. R., St. Laurent, L. C., and Weaver, A. J.: Tidally driven mixing in a numerical model of the ocean general circulation. Ocean Modell., 6, 245–263, https://doi.org/10.1016/S1463-5003(03)00011-8, 2004.

Schweiger, A., R. Lindsay, J. Zhang, M. Steele, H. Stern, Uncertainty in modeled arctic sea ice volume, J. Geophys. Res., doi:10.1029/2011JC007084, 2011

Swart, N.C., Gille, S.T., Fyfe, J.C. and Gillett, N.P.: Recent Southern Ocean warming and freshening driven by greenhouse gas emissions and ozone depletion, Nature Geoscience, 11, 836-841, 2018.

Theurich, G., DeLuca, C., Campbell, T., Liu, F., Saint, K., Vertenstein, M., Chen, J., Oehmke, R., Doyle, J., Whitcomb, T., Wallcraft, A., Iredell, M., Black, T., Da Silva, A. M., Clune, T., Ferraro, R., Li, P., Kelley, M., Aleinov, I., Balaji, V., Zadeh, N., Jacob, R., Kirtman, B., Giraldo, F., McCarren, D., Sandgathe, S., Peckham, S., and Dunlap IV, R.: The Earth System Prediction Suite: Toward a Coordinated U.S. Modeling Capability~~.~~, Bulletin of the American Meteorological Society, Vol. 97, No. 7, pp. 1229-1247, 2016.

UNFCCC: Paris Agreement. https://unfccc.int/sites/default/files/english_paris_agreement.pdf, last access 10 May 2019, 2015.

Verseghy, D. L.: Class—A Canadian land surface scheme for GCMS. I. Soil model~~.~~, Int. J. Climatol., 11: 111-133~~.~~, doi:10.1002/joc.3370110202, 1991.

Verseghy, D. L., McFarlane, N. A. and Lazare, M.: Class—A Canadian land surface scheme for GCMS, II. Vegetation model and coupled runs~~.~~, Int. J. Climatol., 13: 347-370. doi:10.1002/joc.3370130402, 1993.

Verseghy, D.L.: The Canadian land surface scheme (CLASS): Its history and future, Atmosphere-Ocean, 38:1, 1-13, DOI: 10.1080/07055900.2000.9649637, 2000.

West, A. E., McLaren, A. J., Hewitt, H. T., and Best, M. J.: The location of the thermodynamic atmosphere–ice interface in fully coupled models – a case study using JULES and CICE, Geosci. Model Dev., 9, 1125-1141, https://doi.org/10.5194/gmd-9-1125-2016, 2016.

Winton, M.: Do Climate Models Underestimate the Sensitivity of Northern Hemisphere Sea Ice Cover~~?.~~?, J. Climate, 24, 3924–3934, https://doi.org/10.1175/2011JCLI4146.1, 2011.

Yang, D. and Saenko, O.A.: Ocean Heat Transport and Its Projected Change in CanESM2~~.~~, J. Climate, 25, 8148–8163, https://doi.org/10.1175/JCLI-D-11-00715.1, 2012.

Young, A. H., Knapp, K. R., Inamdar, A., Hankins, W., and Rossow, W. B.: The International Satellite Cloud Climatology Project H-Series climate data record product, Earth Syst. Sci. Data, 10, 583-593, https://doi.org/10.5194/essd-10-583-2018, 2018.

Zahariev, K., Christian, J. R. and Denman, K. L.: Preindustrial, historical, and fertilization simulations using a global ocean carbon model with new parameterizations of iron limitation, calcification, and N2 fixation, Prog. Oceanogr., 77, 56–82, 2008.

Zalesak, S. T.: Fully multidimensional flux corrected transport algorithms for fluids, J. Comput. Phys., 31, 335–362, 1979.

Zhang, J. and Rothrock, D.A.: Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates, Mon. Wea. Rev. 131(5), 681-697, 2003

Zobler, L.: A World Soil File for Global Climate Modelling. NASA Technical Memorandum 87802. NASA Goddard Institute for Space Studies, New York, New York, U.S.A., 1986.